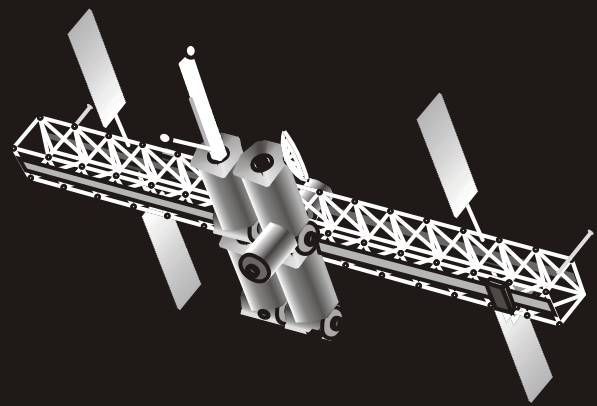


© 2005
ENSC 894
Simon Fraser University
School of Engineering Science

2005 Course Proceedings



Writing the Future:

Interactions between Engineers, Technology, and Language

Vibration Arthrography as a Diagnostic Method for Osteoarthritis of the Knee joint

Pooneh Afkari

Abstract— The technique of vibration arthrography (VAG) permits the non-invasive evaluation of internal derangements of the knee. In this method, the delicate sounds emitted from a joint are recorded by a sensitive detector which is placed on a joint, and the resultant signals are depicted by an oscilloscope. The primary objective in this study was to examine the vibration signal patterns associated with osteoarthritis of the knee. Using the computer-assisted recording and analysis equipment, a major survey was conducted into the signal patterns present in subjects with either symptomatic or normal knee joints. The results of this experiment demonstrated that the VAG technique is a non-invasive and sensitive tool for early detection of osteoarthritis of the knee joint. Additionally, it was shown that by putting an accelerator on the patella of the left knee with osteoarthritis, we will achieve a more characteristic signal of this abnormality; in other words, no significant difference between signals was recorded from medial, lateral, and patella of the right knee ($p > 0.05$). The reason for this finding could be less physiologic arthrokinematic sounds under the patella of the left knee. In order to facilitate this presentation, we start with a brief review of the VAG signal's history and the anatomical structure of the knee joint. At the end of this paper, some new directions for analyzing vibration arthrography signals in the clinical examination also summarized.

Index Terms—Click, knee, osteoarthritis, vibrationarthrography.

I. INTRODUCTION

AMONG all the methods for diagnosing joint disorders, MRI is the most common one; however, because of its high expense, it is not widely used for diagnosis. On the other hand, orthopedic surgery and invasive methods of investigation, which are sometimes inevitable tools for diagnosing pathological conditions, have so many associated complications as to preclude their use in most patients [2]. Another highly developed technique in orthopedic practice is the examination of joints for subtle signs of mechanical malfunctions indicated by sound and crepitus. In this technique, the subtle sounds of a joint are detected and recorded by a sensitive accelerometer which is placed over the joint [4]. Therefore, VAG promises to be a sensitive and non-invasive tool for an objective examination of the human's locomotor system. This technique has been recognized as a reliable method of early diagnosis in neonatal hip displacement [3, 4]. Analyzing acoustic vibrations emitted from the joint was first done by Walter (1929). By using a microphone placed on the knee, he succeeded recording the sounds emitted from this joint; however; there were serious problems with skin friction and ambient noise that diminished the microphone signal-to-noise ratio. Furthermore, microphones integrate all sounds

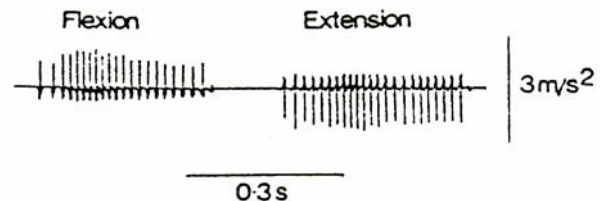


Fig.1 The PPC mirror image waveform, which can be recorded from all normal knees

arising in a region of space with the subtle sounds of the joint, thereby preventing the detection of the vibration's location (Mang et al 1980, Mollan 1981, Mollan et al 1982, 1983).

Further experiments were carried out in which a microphone was replaced with an accelerometer (Mollan, 1981; Mc Carea et al 1985). In 1991, Neely et al could record the physiological patellofemoral crepitus (PPC) by means of a small transducer placed on the patella femoral joint. PPC is the result of microscopic movements of the patella in the trochlear groove. During slow flexion and extension, the patella rocks to and fro in the trochlear groove of the femoral condyle and alternately sticks and slips, producing a PPC vibration (Beverland et al, 1985). The PPC acceleration signal is a periodic complex sinusoidal signal similar to heart signals. It contains two wavelets: 1) during extension of the knee and 2) during

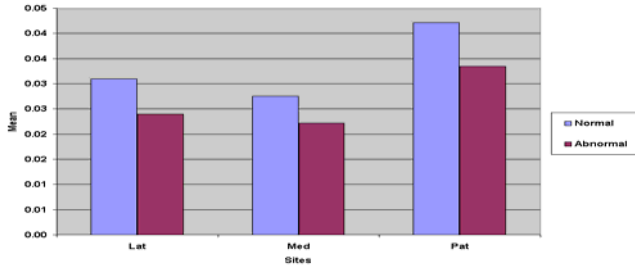


Fig. 2 The average of the Lateral, Medial & Patellar sites of the subjects

flexion of the knee. Since the patella is moving in opposite directions during flexion and extension, these two wavelets mirror image each other (Neely et al, 1991) (Fig 1).

The reason for the sound emissions from the knee joint is suggested to be the rubbing of knee-joint surfaces against one another (Mollan, 1982; Tavathia, 1992). If we put our finger on the knee surface, we may feel subtle vibrations during movement of the knee. These vibrations can produce signals which are called “clicks” and which have the frequency of 40-140 Hz (Fig 2). The source of these sounds is said to be exactly under the patellar surface [7, 8].

In the work reported in the present paper, an attempt is made to examine whether vibrationarthrography can be a diagnostic tool for osteoarthritis of the knee joint.

II. BACKGROUND

A. Anatomy of the human knee joint

The knee joint, which is categorized as a synovial joint, is formed between the femur and the tibia. It has four distinguishing features: a synovial membrane, a joint cavity, an articular cartilage, and a fibrous capsule. This joint consists of three parts: the patellofemoral joint (the joint between the patella and the femur), and the lateral and medial tibiofemoral joints (the lateral and medial sides of the joints between the femoral and tibial condyles). A synovial membrane lines the inner aspect of the articulation cartilages of the tibia and continues superiorly into the suprapatellar bursa of the femur. This bursa, which is located between the femur and the quadriceps tendon, is an extension of the joint cavity (Romans et al, 1989). The synovial membrane also attaches to the periphery of the patella and is separated from the patella ligament by the infrapatellar fat pad. The quadriceps femoris is made up of four muscles, which unite at the quadriceps tendon, (which, in turn, attaches to the patella). The patella is a sesamoid bone that is located on the anterior surface of the knee joint, within the quadriceps femoris tendon. The posterior surface of the patella is divided into areas called facets, which allow the patella to glide smoothly over the trochlea of the femur during flexion. Two types of cartilage exist in the knee joint: hyaline cartilage and menisci. The hyaline

cartilage is located on the articulating surface of bones and acts to transmit loads between them while withstanding repetitive joint motion. This cartilage is vital to joint function because it compresses to protect the underlying bone during movement. The menisci cartilage, which has the wedge shaped fibrocartilaginous structure, is located between the femur and the tibia.

B. The abnormal human knee joint

Osteoarthritis, which is the most common idiopathic condition of the knee, affects both the bone and the articular cartilage of the patella. In this abnormality, the articular cartilage softens, fibrillates, and sheds off the undersurface of the patella. Osteoarthritis has been graded according to the severity of the lesion (Noyes et al, 1989). Grade one involves articular cartilage softening (closed chondromalacia) in which the cartilage may have near or complete loss of resilience. Grade two is characterized by interruption of the cartilage surface, and is often limited to the patellar articular cartilage. In Grade three, the underlying bone is exposed. This includes a normal contour which has been eroded.

The preceding review examines the accuracy of vibrationarthrography, one of the highly developed diagnostic methods for detecting osteoarthritis of the knee joint.

III. METHODS AND MATERIALS

A. Research subjects

This study was carried out as experimental research. The main subjects of this study were patients with osteoarthritis (OA). In order to determine the possibility of using VAG signals for diagnosing OA, we measured the VAG signals of the patients at the Burnaby hospital who were undergoing physiotherapy. The status and the location of OA had been carefully examined by a radiologist. Among 11 patients suffering from OA, 7 were female with an average age of 55 years and 4 were male with an average age of 47 years. 11 people with normal knee were also included to form the control group. The age of the normal subjects ranged from 14 to 34 years. All the abnormal subjects had arthritis in the tibiofemoral joint which had been diagnosed by an orthopedic surgeon.

B. Data Collection

Subjects were requested to sit on a rigid table with both legs suspended. In order to record the VAG signal, an accelerator (Bruel & Kjaer 4375, Denmark) was placed on the knee surface at three positions: the medial, lateral, and patella. The signals were amplified and recorded by an amplifier and an oscilloscope (Bruel & Kjaer, model 2525, Denmark), during swinging movement of the leg from 90 degree flexion to full extension and to 90 degree flexion. The total time period of this movement was approximately 4 seconds. Each person performed this movement three

times. In order to examine the amount of inflation in the swollen knee, the diameter of the knee, 5 cm above, and 5 cm below was measured. The range of motion of the patella and the knee was measured by hand and a set-square respectively. The VAG signals were screened using MATLAB and their root mean square was analyzed using SPSS.

IV. RESULTS

A. Result of the tentative study

In order to evaluate the accuracy of the recorded signals, we performed a tentative study in four normal subjects. The characteristics of these subjects are shown in table 1. The findings of this study showed no significant difference between the results of four repeated swinging movements for each part of the knee ($p > 0.05$). As a result, we concluded that our data were reliable.

B. Results of the main study

The main study was carried out on both knees of 21 subjects (21 normal knees and 21 osteoarthritic knees). Table 2 shows the overall results of comparing the RMS of VAG signals at each part of the normal knee with the same part of the abnormal knee. It can be concluded from this table that by putting the accelerometer on the patellas of subjects with left osteoarthritic knees, we will achieve more characteristic signals for this disease. Additionally, the average of the RMS signals in three parts of the normal knee was higher than those of the abnormal knee (Fig 2). However; this difference was not significant ($p > 0.05$).

V. DISCUSSION

The results of this experiment indicated that the average of the RMS signals of the osteoarthritic knees, including all signals recorded from the lateral, medial, and patella, were less than the average of the RMS signals at the same side of the normal knee. Moreover; no significant difference was

TABLE 1
CHARACTERISTIC OF SELECTED SUBJECTS

Name	Sex	Age	P-value		
			Lat [†]	Med [†]	Pat [†]
NN	M	27	0.3	0.7	0.9
ES	F	20			
HA	M	23			
NA	M	27			

† Lat= Lateral, Med= Medial, Pat= Patella

TABLE 2
RESULTS OF COMPARING THE AVERAGE OF RMS BETWEEN NORMAL AND ABNORMAL KNEE JOINT

Sites	Lateral	Medial	Patella

Knee	L	R	L	R	L	R
Normal Vs. Abnormal	.09	.56	.06	.93	.03	.71

found between VAG signals recorded from three parts of the right osteoarthritic knees. However, it was shown that if the abnormality existed in the left knee, the patella would be the best place for obtaining clear diagnostic signals ($p < 0.05$). We assume that the high frequency arthrokinematic movements in the right knee produces more physiological noises; Consequently, the signals recorded from the right knee are not as clear as those recorded from the left knee, and the left knee can show more reasonable signals. The reason for a higher VAG average signal in normal knees compared to that of symptomatic knees was suggested to be lower background amplitude in osteoarthritic knees [1, 2]. In other words, the amplitude of the physiological patellofemoral movement in the normal knee is more than that of the abnormal knee, which is due to the reduction of arthrokinematic movements in the symptomatic knee (Fig 2).

Beverland et al., (1986) observed this information in various types of meniscal tear. Using vibration sensors attached to bony prominences around the knee of one hundred and seventy-two patients, who had been arthroscopically demonstrated to have a meniscal lesion, provided characteristic meniscal signals in 150 patients [6, 4]. In many cases, the site of meniscal tears could also be determined. For instance, if the posterior region of the meniscus was torn, the signal would have the smallest size. After resection, however, this signal was significantly diminished.

Our findings were also in satisfactory agreement with Rangayyan's studies (1997), which examined the patellar click in one complete cycle of the knee movement. This signal is the transient signal emitted from the patella in the flexion and extension movements. Since the vibration appears when the patella's area of contact moves onto adjacent facets, these signals are produced repeatedly in one complete cycle of movement.

VI. FINAL RESULTS

Presences of osteoarthritis in tibiofemoral joints disturb the rhythm of the VAG signal. This will result in a lower average of the RMS for the abnormal knee, compared with the normal knee, and sharper spikes due to the presence of clicks in the abnormal knee. Further, by placing the accelerometer on the patella of the left osteoarthritic knee, we can receive a more clear diagnostic signal of this abnormality. However, it is too early to judge whether this method can replace the MRI and arthroscopy. The main reason for this is the interventions of noises arise from the

skin and the surrounding environment friction.

REFERENCES

- [1] L.A Neely, W.G Kernohan, D.A.Barrt, C.H.B Mee, and R.A.B Mollant, "Optical measurement of physiological patellofemoral crepitus." Clin. Phys. Physiol.Meas., 1991, Vol. 12, 219-226
- [2] Rangaraj M. Rangayyan, Sridhar Krishnan, F.R. Noyes. & C.L. Stabler, "A system for grading articular cartilage lesions at d arthroscopy," The Journal of Sport Medicine., 1989, vol.17,505-513
- [3] Mollan.R.A.B., Mccullagh G.C., Wilson R.I., "A critical appraisal of auscultation of human joints", Journal of clinical orthopaedics and related research., 1982, Vol. 170, 231-237
- [4] Beverland DE, McCoy GF, Kernohan WG, Mollan RAB, "What is patellofemoral crepitus?" 1986, Vol.68, 288-293
- [5] E. H. Miller, "A note on reflector arrays (Periodical style—Accepted for publication)," *IEEE Trans. Antennas Propagat.*, to be published.
- [6] J. Wang, "Fundamentals of erbium-doped fiber amplifiers arrays (Periodical style—Submitted for publication)," *IEEE J. Quantum Electron.*, submitted for publication.
- [7] C. J. Kaufman, Rocky Mountain Research Lab., Boulder, CO, private communication, May 1995.
- [8] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interfaces (Translation Journals style)," *IEEE Transl. J. Magn.Jpn.*, vol. 2, Aug. 1987, pp. 740-741 [*Dig. 9th Annu. Conf. Magnetics Japan*, 1982, p. 301].

Internet-based Robotic Teleoperation

Renji Chen

Experimental Robotics Laboratory

Engineering Science School, Simon Fraser University, Burnaby, B.C., Canada, V5A 1S6

renjic@sfu.ca

Abstract: This paper presents a critical review of current research approaches on networked robotic teleoperation, especially the Internet-based robotic teleoperation. Basing on this review, we discussed the deficiency of current research approaches and proposed a novel one for modeling and control of teleoperation system. In addition, this paper suggests and summarizes some promising directions about the development of Internet-based robotic teleoperation.

Keywords: robotic teleoperation, Internet-based control, wave variable, predictive control

I. INTRODUCTION

ONLINE robots, i.e., Internet-based robots, go beyond webcams by including some mechanism for action. Although a disagreement exists among experts about how to define robot, a robot is usually defined as a physical device that is capable of moving through its environment, in contrast to a purely software-based “softbot”. An online robot is a robot that is accessible from any computer on the Internet, generally through the HTTP protocol. Since the first online telerobot appeared in August 1994, about forty such systems have been put online by research teams around the world. Online robots take many forms and allow users to perform a variety of operations remotely[1].

Most issues of Internet-based robotic teleoperation originated from the Networked Control System (NCS), so lots of similarities exist between them. And many approaches for the NCS can be easily applied to the Internet-based control. For an excellent review of current control technologies of Networked Control System, refer to [2] and the latest special issue on current Networked Control System[25]. Though Internet-based control, one special kind of NCS in fact, has many similarities with NCS, still it has many key differences with NCS. While almost all of the current NCS are special-purpose

dedicated industrial network such as Fieldbus, CANbus and so on, Internet-based control refers to the control over the biggest global easily accessed Internet. Since the Internet is much more complex than the dedicated industrial network, Internet-based control is also much more challenging than the control of NCS.

This paper firstly gives a general introduction about Internet-based robotic teleoperation. Then a through view of current research approaches in this field. On the base of the research summary, we proposed our research approach. Finally, the conclusion and some promising research direction are presented.

II. PROBLEM STATEMENT

Fig.1 is a typical architecture of teleoperation control. In the Internet-based robotic teleoperation control, the transmission link will be the Internet which makes Internet-based control much more complex and difficult than common Networked Control Systems.

When data are transmitted over the Internet, they are divided into small packets and routed in real time through a possibly large number of intermedia routers. Since the exact route of transmitted data can not be determined beforehand and totally determined by the current network state, Internet traffic exhibits very complex phenomena. And thus the transmitted information undergoes severe effects such as data missing, signal distortion etc. One typical transmitted and corresponding received sine wave together with time delay are shown in Fig.2. It could be seen that the biggest effect the Internet imposes on transmitted information is the time delay, also an obvious distortion effect can be seen. Generally speaking, the effect the Internet imposes on transmitted information can be roughly summarized as following: (1) large or small uncertain time delay, (2) missing data, (3) signal distortion. For an Internet-based discrete control system, vacant sampling is another

important phenomena, which can be classified as effect (2). All of these adverse effects render Internet-based control faced with tremendous challenges.

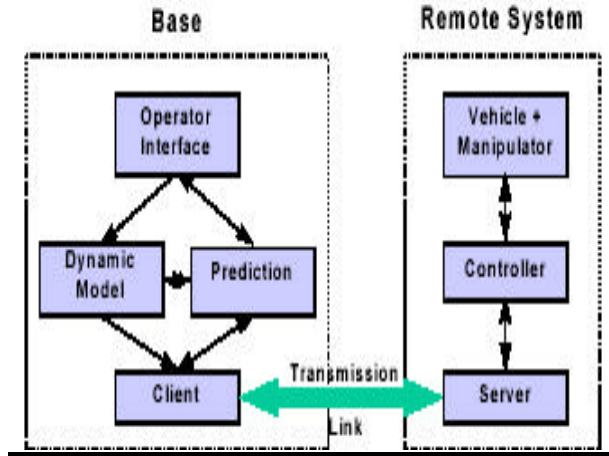


Fig.1 Typical architecture of teleoperation control (the transmission link is Internet in Internet -based control)

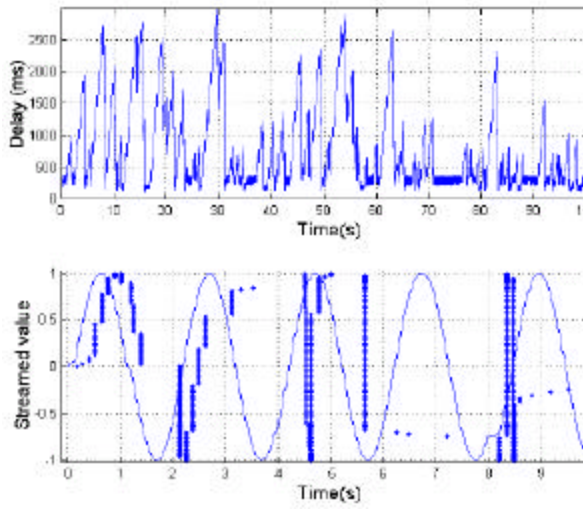


Fig.2 Top: Cross atlantic round trip time delay between Georgia Tech, Atlanta and Metz, France using TCP protocol. Bottom: Transmitted and received sine wave, sampled at 10 milliseconds[20].

In order to realize Internet-based control, the first problem to be solved is the model identification. In particular, the Internet channel identification is the first and most important and difficult problem. From the above we know that Internet traffic exhibits extremely complex phenomena which means that the Internet channel is also extremely complex and challenging to be identified.

III INTERNET -BASED ROBOTIC TELEOPERATION APPROACHES

3.1 Control structure

Currently, roughly two general NCS configurations listed as follows exist, the same is Internet-based teleoperation[2]:

Direct structure

The NCS in the direct structure is composed of a controller and a remote system containing a physical plant, sensors and actuators. The controller and the plant are physically located at different locations and are directly linked by a data network in order to perform remote closed-loop control as illustrated in Fig.3. In this architecture, the local controller and the remote plant interact in a kind of direction way.

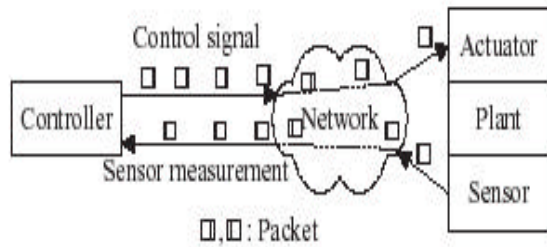


Fig.3 NCS in a direct structure

The control signal is encapsulated in a frame or a packet and sent to the plant via the network. The plant then returns the system output to the controller by putting the sensor measurement into a frame or a packet as well. In a practical implementation, multiple controllers can be implemented in a single hardware unit to manage multiple NCS loops in the direct structure.

Hierarchical structure

The basis hierarchical structure consists of a main controller and a remote closed-loop system as depicted in Fig.4.

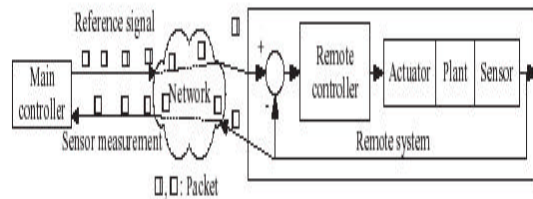


Fig.4 NCS in a hierarchical structure

Periodically, the main controller computes and sends the reference signal in a frame or a packet via a network to the remote system. The remote system then processes the reference signal to perform local closed-loop control and returns to the sensor measurement to the main controller for networked closed-loop control. The networked control

loop usually has a longer sampling period than the local control loop since the remote controller supposes to satisfy the reference signal before processes the newly arrival reference signal. Similar to the direct structure, the main controller can be implemented to handle multiple networked control loop for several remote systems.

3.2 Research approaches for Internet-based robotic teleoperation

Based on the above two control structure, myriad designs exist for constructing Internet-based robotic teleoperation architectures with force feedback or no force feedback. In general, these attempts can be roughly grouped into following several approaches. As a matter of fact, many designs tend to blur these distinctions.

(1) “move and wait” strategy

This is one of the earliest simple approaches which can be applied in the Internet-based robotic teleoperation. In this strategy, the remote user make a small, relatively safe motion, then wait to see that it succeeded before moving again[3]. Obviously, this approach is feasible for some entertainment application, but it can exhibit slow bandwidth and infeasible for most robotic teleoperation case.

(2) Web browser based operation method

Most of the early Internet-based controlled robots used the Common Gateway Interface (CGI)[4] to interface between web browsers and the physical devices being controlled. CGI processes are launched by a web browser in response to certain HTTP requests, and the CGI result is sent as the HTTP response. This approach has undergone major changes since the first online robot in 1994[5], and the introduction of JAVA is probably the most significant as it allows code to be executed on the client side, whereas previously all code had to be in the CGI process.

(3) Supervisory control approach[6]

In supervisory control, lower level tasks are executed autonomously at the remote site, while the operator provides higher level instructions from the local site. This approach provides the operator with the ability to supervise an autonomously executed task without direct intervention into the process. It frees the remote user from the need to command every detail of the operation at the remote site and it could achieve satisfactory performance for Internet-based teleoperation. This relies upon the

telerobot having a degree of autonomous control to pursue individual goals. Conway[7] terms this Teleautonomous control — intelligent action at a distance and defines different operating modes of a telerobot.

These higher levels of control are particularly suited to environments where there are certain common repetitive tasks. In a pure telerobotic system these must be performed repetitively by the operator every time, often leading to tiredness, lack of concentration and mistakes.

(4) Bilateral control

To enhance the controllability of a teleoperation system and to improve its performance, it is important to develop an effective haptic interface[10]. However, the use of haptic interface over the Internet is a much more challenging problem than transferring audio and video data because these devices must remain stable in spite of the performance fluctuations typical of Internet. In particular, Internet data have random time delays and packet losses that depend on the characteristics of the network and the network's load. Current control design tools are inadequate for Internet-based force-reflecting telerobots because they lack identification and analysis procedures for addressing the specific characteristics of Internet time delay[26]. Much research has taken place in controlling telerobots using bilateral control. The intention is to control both contact force and position (velocity) of the telerobot. As a practical matter, doing both simultaneously is not possible. Instead, bilateral control is modeled as a two-port where force is the input and velocity is the output for the teleoperator. Likewise, velocity is the input and force is the output for the telerobot.

(5) predictive control/display

Predictive control is used to refer to a broad range of predictive approaches including traditional predictive control, Kalman filter, and internal model control. Its role in remote teleoperation is to generate a model for the remote system to compensate for the delay. This predicted model may be used graphically to provide a pseudo real-time response to teleoperator commands. The Jet Propulsion Laboratory[12-14] has done much work in using predictive displays to help control space-based robotic systems. One manifestation of their predictive paradigms is coined the “phantom robot”[13].

Predictive display could give user a simulation of remote robot and allow for interaction with it. It allows

the user to see the effect of commands well before they are actually executed by the remote robot. If the remote environment was sufficiently satisfied some modeling conditions, one could even consider simulating the dynamic motion of objects at the remote site. By providing the remote user with a model of the remote robot that reacts instantaneously, one can largely segregate the operator from the effects of the communication delay. In this approach, the operator controls a local simulation of the telerobot with the control signals then sent in parallel to the simulation and the remote telerobot. The simulation is then displayed superimposed over the return video. In this way, the operator can “see” the effects of the control immediately without having to fully wait for the return signal from the telerobot.

(6) Behavior-based robotic teleoperation

In Internet-based teleoperation, the use of supervisory control to address the time delay problem requires a remote manipulator to exhibit some degree of autonomy. This autonomy is necessary to minimize the contribution of communication delay time to task completion time and to relieve the operator of the responsibility for contact control. Matthew R.Stein[22] examines the use of behavior-based or subsumption architecture control to provide the required autonomy for the remote manipulator. Behavior-based controllers demonstrate the desirable features including reliability and robust operation in unstructured environments and, when used in conjunction with operator direction, avoids the need for higher level representations which do not fit well into the architecture.

(7) Event-based approach

Ning Xi and T.J.Tarn[15,16] proposed a new method for action synchronization and control of telerobotic systems. The significance of this method is that it can effectively deal with the random time-delay in the communication channel, such as the Internet. In addition, the result is independent of human operator. With this approach, the ordinary time referenced control scheme becomes a novel non-time referenced action control one. Instead of using time as an action reference variable, a new sensor-based action reference “event” is employed. Thus, the communication delay will have little impact on the operations, in particular the stability of the system.

(8) Wave variables approach

The problem of time delayed force reflecting

teleoperation motivated the original development of wave variables[17]. Wave variable approach evolves from the earlier work of Anderson and Spong[18]. Though based entirely on the framework of passivity, these variables present a modification or extension to the theory of passivity and provide some additional benefits not found in traditional power variables, such as effort and flow or velocity and force. First, they create robustness to arbitrary time delays or phase lags. And second, they combine force and velocity information in a hybrid fashion, so that the resulting systems have an inherent hybrid structure. Thirdly, based only on the concepts of power and energy, they are applicable to difficult nonlinear systems with unknown models and large uncertainties such as Internet-based control system. As such they are well suited for interaction with real physical environments. Also they have proven well suited for implementation, allowing simple cascading of elements without traditional admittance/impedance related causality problems.

Recently wave-based formulation was enhanced using a prediction technique to improve performance (especially the settling time)[20]. In conventional wave-based techniques stability comes at the expense of degrading system performance with increasing time delay. In order to enhance performance, prediction techniques similar to the Smith Predictor and its various modified variations are employed to cut down the lag felt by the user, while simultaneously improving the settling time. Fig.5 shows the wave variable transmission diagram used in the force reflecting teleoperation. The characteristic wave impedance b is a positive constant or a symmetric positive definite matrix and assumes the role of a tuning parameter, which allows matching a controller to a particular environment or task.

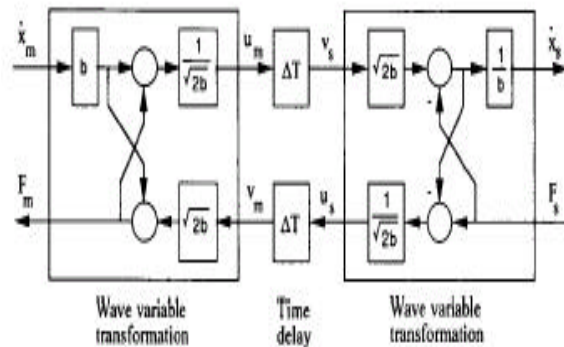


Fig.5 wave-based communication, by transforming velocity-force variables to wave variables before and then back to velocity-force

variables after transmission

(9) Observer-based design approach

Watanabe and Ito[21] proposed an observer for linear feedback control law of multivariable systems with multiple delays in controls and outputs. Their approach overcomes the limitation of modified Smith predictors extended to accommodate multivariable systems. In an alternative scheme the use of a time forward observer derived from conventional state space analysis to ensure stability under a varying delay is made[20]. It is shown that the observer is easy to implement and has performance characteristics superior to passivity-based formalisms. However, the observer is not robust to model mismatches and is unable to reflect remote forces back to the master at steady state conditions while in hard contact. This scheme is primarily geared towards stabilizing a time delayed system, and not force reflection.

(10) Network channel model identification based approach

The uncertainty of Internet-based communications makes modeling real-time systems quite difficult. One key issue of variable time delay compensation in the Internet-based teleoperation is the knowledge of some properties of the delay variation, i.e., the availability of a model of the Internet segment between the two parts of the telerobot. Until now there is little research in modeling such systems. Fiorini and Oboe[23-24] have made some pioneer research on this issue and discussed the various problems and approaches encountered when using the Internet as a communication medium for control systems. They proposed a simple architecture that implements a solution to the three main issues of Internet-based control, namely, the identification of the delay properties of an Internet segment, the design and analysis of a telerobotic controller, and the test of the controller with a planar force reflecting master interacting with virtual objects.

From the above thorough review of current research on Internet-based robotic teleoperation, we can see that research on Internet-based robotic teleoperation are far from mature due to the complexity of the Internet environment, many open problems still exist. Though many experimental system exist since 1994, no essential breakthrough happens in this research field after so many years. While some research approaches such as supervisory control[3], event-based approach[16],

predictive display/control[13] etc which proved to be useful and can realize the control action in some way, they have not involve the essential important issues such as the identification of Internet channel. Roberto Oboe and Paolo Fiorini[23-24] only carried out some early research about the Internet channel modeling issues. It should be noted that bilateral control[8-10] has already been widely accepted as one control mode in current Internet-based robotic teleoperation. The novel concept of wave variable is a promising tool in Internet-based control. Gunter Niemeyer and Saghir Munir[19] and [20] carried out some meaningful research with the use of wave variable concept.

V. PROPOSED RESEARCH APPROACHES

In our proposed approach, we employ the integration approach of control methods combined with the signal processing methods which can be implemented under the Internet exhibiting communication constraints. Several key issues are as follows.

Internet channel identification

The essential difficulty of Internet-based control lies in the Internet environment. Particularly, to model the Internet channel is the utmost key factor in the Internet-based control. Unfortunately, so little research exist about this issue due to lack of mature knowledge about the dynamics of Internet. Generally speaking, Internet channel is highly dynamic, time varying, nonlinear, non-Gaussian etc, and all of these features pose the Internet channel identification as highly challenging research. One approach may be to treat the Internet as one blackbox and do not consider the details of the data transmission on the Internet. Then we can use some statistical learning approach to recursively identify the Internet channel before each control action implemented.

wave variable transformation

In order to guarantee the stability of the complex Internet-based robotic teleoperation, wave variable approach can be employed to stabilize the force feedback teleoperation in which large uncertain time delay exist. Accordingly, the design of the predictive controller and adaptive filter will be carried out in the so called wave domain.

predictive control

Predictive control can be employed as the basic control framework for Internet-based robotic

teleoperation control. The motivation for employing predictive control is because predictive control with such features as above could efficiently deal with the dynamic uncertainty problem existed in Internet-based control system. It should be pointed out that one approach for implementing such scheme can be in the wave domain. For practical application in Internet-based control, it is not feasible to employ predictive control directly due to its limitations. In our system predictive control approach is augmented with sophisticated signal processing methods together with some learning mechanism.

adaptive filtering approach

Since the signals transmitted on the Internet will undergo some small or big changes, mainly the transmission delay, signal missing, signal distortion. In order to ensure the efficient implementation of the designed control algorithm, adaptive filtering approach is employed to deal with such signal restoration issue.

Figure 6 shows a schematic diagram of our proposed integrated control framework. It consists roughly of a local operator, a remote robot, a robot controller, an Internet channel identifier and two wave variable transformation operations. In particular, the local operator include a haptic device and a predictive display interface; The robot controller includes a predictive controller augmented with learning mechanism.

VI. CONCLUSIONS

This paper gives out a thorough review of current research about Internet-based robotic teleoperation control. Despite of so long time research and so many various research methods, many open problems still exist in this field. In order to realize satisfying Internet-based robotic teleoperation performance, various sophisticated research methods need to be integrated to tackle this problem, such as enhance the autonomy capability of local robot with AI techniques, hybrid system approach, and information theory etc. All of these approaches could be some promising research directions .

REFERENCES

[1] Ken Goldberg, Roland Siegwart. Beyond webcams: an introduction to online robots, Massachusetts Institute of Technology, 2002
 [2] Yodyium Tipsuwan, Mo-Yuen Chow. Control methodologies in networked control systems. Control Engineering Practice. 11(2003), pp1099-1111

[3] William R.Ferrell, Thomas B.Sheridan. Supervisory control of remote manipulation. IEEE Spectrum 4(10):81-88, October 1967
 [4] The common gateway interface. <http://www.w3.org/CGI/>
 [5] K.Taylor, J.Trevelyan. Atrilia's telerobot on the web. In 26th International Symposium on Industrial Robotics, pp.39-44, Singapore, October 1995
 [6] Thomas Sheridan. Space teleoperation through time delay review and prognosis. IEEE Transactions on Robotics and Automation, 9(5): 592-606, October 1993
 [7] Lynn Conway, Richard A.Bolz, and Michael W.Walker. Teleautonomous systems: projecting and coordinating intelligent action at a distance. IEEE Transactions on Robotics and Automation, 6(2):146-158, April 1990
 [8] Gary M.H.Leung, Bruce A.Francis. Bilateral controller for teleoperators with time delay via \mathcal{H}_∞ -synthesis. IEEE Transactions on Robotics and Automation, 11(1):105-116, February 1995
 [9] Dale A.Lawrence. Stability and transparency in bilateral teleoperation. IEEE Transactions on Robotics and Automation, 9(5):624-637, October 1993
 [10] Robert J.Anderson. Teleoperation with virtual force feedback. Proceedings of the 1993 SPIE International Symposium on Optical Tools for Manufacturing and Advanced Automation, Boston MA, September 1993
 [11] Park, Jong Hyeon&Cho, Hyun Chul. Sliding mode controller for bilateral teleoperation with varying time delay. International Conference on Advanced Intelligent Mechatronics, Sep.,1999, pp.311-316
 [12] Antal K.Bejczy, Paolo Fiorini, Won S.Kim, and Paul S.Schenker. Toward integrated operator interface for advanced teleoperation under time-delay. Intelligent Robots and Systems, pp.327-348, 1995
 [13] Antal K.Bejczy, Won S.Kim, and Steven C.Venema. The Phantom robot: predictive displays for teleoperation with time delays. IEEE International Conference on Robotics and Automation, pp.546-551, May 1990
 [14] Won S.Kim, Antal K.Bejczy. Demonstration of a high-fidelity predictive/preview display technique for telerobotic servicing in space. IEEE Transactions on Robotics and Automation, 9(5):698-701, October 1993
 [15] Ning Xi, T.J.Tam. Action synchronization and control of Internet-based telerobotic systems. Proceedings of the 1999 IEEE International Conference on Robotics and Automation, Detroit, Michigan, May 1999, pp.219-224
 [16] N.Xi, T.J.Tam and A.K.Bejczy. Intelligent planning and control for multirobot coordination: an event-based approach. IEEE Transactions on Robotics and Automation, 12(3), June 1996
 [17] Niemeyer, Gunter and Slotine, Jean-Jacques E., Stable Adaptive

Teleoperation. IEEE Transactions on Automatic Control, AC-16(1): 152-162, Jan 1991

[18] Gunter Niemeyer. Using wave variables in time delayed force reflecting teleoperation. PhD Thesis, Massachusete Institute of Technology, 1996

[19] Anderson, Robert J.and Spong Mark W. Bilateral control of teleoperators with time delay. IEEE Transactions on Automatic Control, AC-34(5): 494 -501, May 1989

[20] Saghir Munir. Internet-based Teleoperation. PhD thesis, Georgia Institute of Technology, 2001

[21] Watanabe, Keiji and Ito, Masami. An observer for linear feedback control laws of multivariable systems with multiple delays in control and outputs. Systems & Control Letter, 1(1):54-59, July 1981

[22] Matthew R.Stein. Behavior-based control for time delayed teleoperation. PhD thesis, Department of Computer and Information Science, School of Engineering and Applied Science, University of Pennsylvania, 1994

[23] Roberto Oboe, Paolo Fiorini. A design and control environment for Internet-based telerobotics. The International Journal of Robotics Research, 17(4): 433-449, April 1998

[24] Paolo Fiorini, Roberto Oboe. Internet -based telerobotics: problems and approaches. International Conference on Advanced Robotics, Monterey, CA 1997

[25] Special issue on Networked Control Systems, IEEE Transactions on Automatic Control, No.9, 2004

[26] Tissa Mirfakhrai and Shahram Payandeh, A Delay Prediction Approach for Teleoperation over the Internet, Proceedings of IEEE International Conference on Robotics and Automation, 2002

[27] B.Hannaford, J.Hewitt, T. Maneewarn, S.Venema, M.Appleby, and R.Ehresman. Telerobotic macros for remote handling of protein crystals. Proceedings Intl. Conf. On Advanced Robotics, Monterey, CA, July 1997



Renji Chen received his BS degree from Xidian University in 1991, and master degree from Harbin Institute of Technology, China in 1996. Worked in Shenyang Institute of Automation, Chinese Academy of Sciences from 1996 to 1999,

and worked as research assistant in Automation department of Shanghai Jiaotong University, then as a system engineering in Zhongxing Telecommunication Corporation, Shanghai for 8 months before came to University of British Columbia in 2002 to pursue PhD degree. Transferred to SFU since September 2004. His research focus on Internet-based robotic teleoperation, predictive control, signal processing information theory, and statistical learning etc

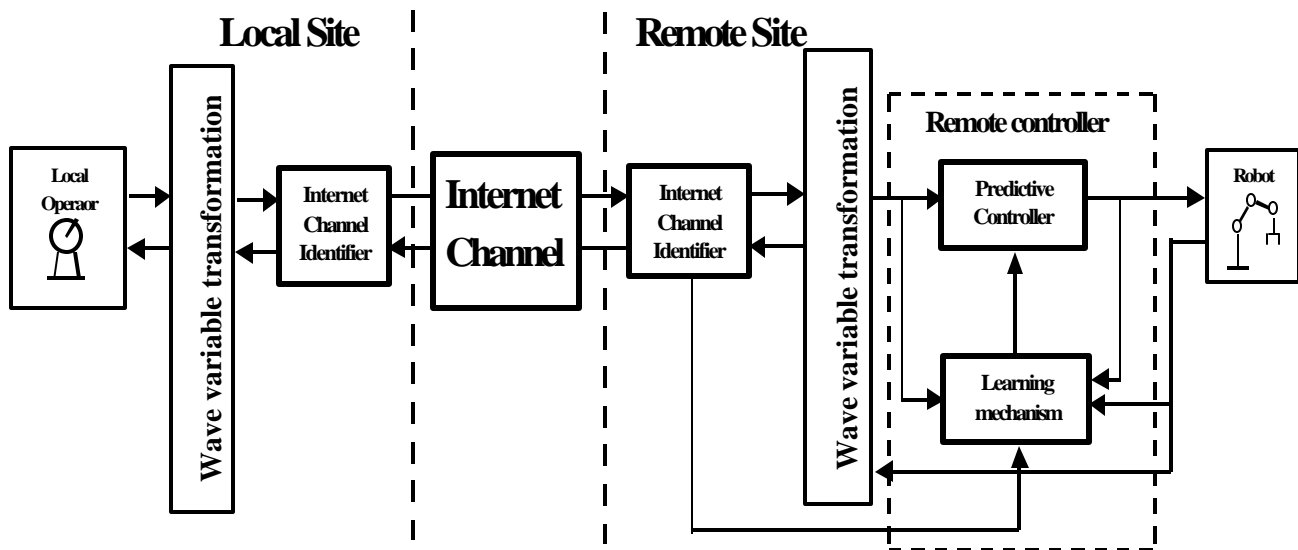


Fig.6 Schematic diagram of integrated Internet-based robotic teleoperation control

Brain Computer Interface (BCI)

Seyed Hossein Faramarzi

Abstract—Although many devices are available to assist disabled individuals, often they have difficulties interacting with the world around them. The assistive devices available in the market currently are accessible to people with a wide range of disabilities from spinal cord injuries to traumatic brain injuries. These devices usually require little and sometimes no movement from the patient. Brain computer interfaces (BCIs) are a new class of devices that are being built with the advancement of new technology and digital processing techniques. To use these devices, the patient simply needs to think to turn on the special device sensors which then activate the device and allow the person to connect with their surroundings. The primary motivation is to help patients of neuromuscular impairments; However, the technology can further develop into other applications including, but not limited to, video games.

I. INTRODUCTION

The Brain-Computer Interface (BCI) is a communication system and enables the user to control special computer applications by using only his or her thoughts. Different research groups have examined and used different methods based on electroencephalography (EEG) recorded from the scalp of the human brain to achieve this result. The EEG is measured and sampled while the user imagines different things (for example, moving their left or their right hand), and depending on the BCI, particular preprocessing and feature extraction methods are applied to the EEG sample [1]. The result is then possible to detect the task-specific EEG signals or patterns from the EEG samples with a certain level of accuracy.

The first BCI research can be dated back to 1960's, but it was in the 1990's when BCI research really got started, and faster computers and better EEG devices offered new possibilities. To date, over 20 BCI research groups have taken different approaches to the subject, some more successful than others. Less than half of the BCI research groups have built an online BCI, which can give feedback to the subject. None of the BCIs have yet become commercial, and only a couple has been tested outside laboratory environments. Despite the technological developments, numerous problems still exist in building efficient BCIs, and the biggest challenges are related to accuracy, speed and usability. The BCIs could provide a new communication tool for people suffering from, called locked-in syndrome, and they are completely paralyzed physically and unable to speak, but cognitively intact and alert. Locked-in syndrome can be caused, for example, by amyotrophic lateral sclerosis (ALS), high-level spinal cord injury or brain stem stroke, and in its severest form, people are not able to move any muscle of their body.

The electroencephalograph (EEG) is used to record the electro-biological signals of the brain which can be analyzed to operate the device(s) that will be used by the patient to interact with his/her environment. The signals are recorded into appropriate tools which convert the signals into data, and the device can read in order to obtain relevant information about what the patient was thinking and intended to do [2]. The machine may display the result faster than human movement depending on the type of internal processor used. Until the advent of computers about sixty years ago, the signals couldn't be processed and converted into communicative signals. Although BCIs are a miracle for many disabled patients, they are not widely adopted because they are still relatively new in the field of biomedical equipment and, consequently, very expensive.

The goal of this project is to create a BCI that controls a computer cursor with horizontal and vertical movements in addition to analyzing the signal frequencies and obtaining their corresponding noise levels. The device is engineered to operate with standard environmental noise and one or more users. This project is developed based on past research and projects undertaken by several other students at Simon Fraser University in the fields of Electrical and Biomedical Engineering [4].

This project has produced a brain computer interface, which is robust, has more room for expansion, and permits further research by future students and researchers. The detected brain activity is analyzed and then processed to move the cursor and activate the click function. Note that some movements are much easier to achieve than others such as moving from the top-right of the screen to the bottom-left rather than moving to particular corners or positions on the screen. However, with further training, patients could quickly learn to move the cursor to their desired location with ease.

The human brain working is one of the most difficult and exciting discoveries and interest in the brain. A critical point of debate between different schools of thought lies in the question that if the explanation of behavior and mental thought contained completely within the physiological make up of the organ itself, or is there something more? The categories of philosophy that answer this question are dualisms (theories where the mind and the brain are two separate phenomena) and monism (theories where the mind and the brain are the same entity). Dualistic theories argue that although the mind and the brain may interact, they can never be reduced so far as to become the other; however, monolists believe that the two entities actually become the same. In this paper, we answer this question from a scientific perspective.

II. ELECTROENCEPHALOGRAPHY

Electroencephalography (EEG) is a form of electro-biological measurement that is based upon electrical signals from the brain [3]. The scalp is an appropriate electrical insulator, and EEG has a poor resolution; however, since all electric currents produce a magnetic field, and the scalp is transparent to magnetic fields, MEG (Magnetoencephalography) can be used to record brain activity with better spatial resolution [3].

The EEG method is the most suitable technique for brain computer interfaces because it is non-invasive and comparatively inexpensive.

The average weight of human brain is around 1400 grams, and the brain can be divided into four structures: cerebral cortex, cerebellum, brain stem, hypothalamus and thalamus. The most relevant of them concerning BCIs is the cerebral cortex, which can be divided into two hemispheres (left and right). The right hemisphere senses information from the left side of the body and controls movements on the left side of the cerebrum is the opposite. The hemispheres, which can be divided into four lobes, are connected with each other via corpus callosum called frontal, parietal, occipital and temporal lobes. Cerebral cortex is responsible for many "higher order" functions such as problem solving, language comprehension and processing of complex visual information and can be divided into several areas, which are responsible for different functions. With respect to EEG measurements, the cerebral cortex is the most important part of the brain compared to other sections of the brain [1].

The primary motor cortex is the region of the brain cortex that is the most essential for producing a brain computer interface. Understanding the properties of scalp recorded EEG signal is critical to understanding and designing brain computer interfaces. Because the electrical signals of the brain must pass through the cerebrospinal fluid, the skull, and the scalp, the most useful signals are those that originate from a population of neurons [14].

Three main categories of brain activity can be detected by EEG (rhythmic brain activity, event-related potentials – ERP, and event-related desynchronization – ERD, and event-related synchronization – ERS). ERPs are changes in EEG that take place because of a particular event or stimulus and usually smaller (by a factor of about 10) in amplitude than rhythmic signals [11]. ERS and ERD are a change to the amplitude of a certain EEG brain rhythm due to a particular event. To precisely measure ERS and ERD the power of a certain frequency band is measured before and after an event, over a certain number of trials [11]. EEG suffers from volumetric smearing of signals, the dependence on vast numbers of neurons acting in unison, and the trial-to-trial variability of brain activity. The layers of cerebrospinal fluid, the skull and the scalp between the brain and electrodes act as an electrical shield, smearing and attenuating signals. Nevertheless, EEG consists of good temporal resolution and is simple to manufacture and used non-invasively, which is the most practical method for creating brain computer interfaces [5].

III. BRAIN COMPUTER INTERFACE

Determining the relationship between brain signals/activities and mental tasks has opened up new fields of research and communication. With the growing discoveries of the human brain, the cost and power of new medical technologies and the increased interest/demand for assistive devices for disabled patients the technology of a brain computer interface (BCI) has created huge interests and caused many researchers to study the human brain more precisely [3].

With a BCI device, no longer human muscle movements is essential to communicate with the outside world; communication and control of the environment can be accomplished via neurophysiological signals of the human body. A BCI involves monitoring brain activity, using a brain imaging technology (EEG), and detecting changes in the activity to communicate with the outside world using digital signal processing algorithms [6].

The reason EEG is the best brain imaging technology for BCI devices is due to its inexpensive price, easy acquisition, high temporal resolution (allowing realtime systems) and direct correlation between brain activity and EEG signals. Therefore, almost all BCI devices developed so far have been exclusively based on EEG, and brain activity patterns controlled by the user is essential to control BCI devices [12]. The main function of the BCI device is to interpret the patterns recorded on the patient's scalp and use the characteristics of the received signals to control other devices that the patient may be involved with such as a wheelchair [12].

IV. BCI METHODOLOGY

At the present time, the ability to discover and analyze human brain activity is severely limited because the current BCI devices cannot read the disabled brain. Patient cannot effortlessly cause the cursor to move to a particular location on the BCI screen or get a wheelchair to turn to a desired direction. From the brain's perspective a handicapped patient's brain signals are rather "vague" and unclear and thus the brain cannot process them as desired [8]. To account for this problem, the BCI takes advantage of brain motions that can be changed slightly in order to produce understandable commands. However, consciously modifying such brain signals requires a huge deal of meditation, effort, preparation, and training by the patient. Examining complex EEG signals to extract the control signals is a difficult digital signal processing challenge, even with a prepared and trained patient who may be completely concentrated on the task. The two major approaches for adjusting a patient's brain activity as input to a BCI device are the pattern recognition (PR) approach and the operant conditioning (OC) approach [9]. The EEG pattern recognition approach is based on different mental tasks and the operant conditioning approach is based on the self-regulation of the EEG response [9].

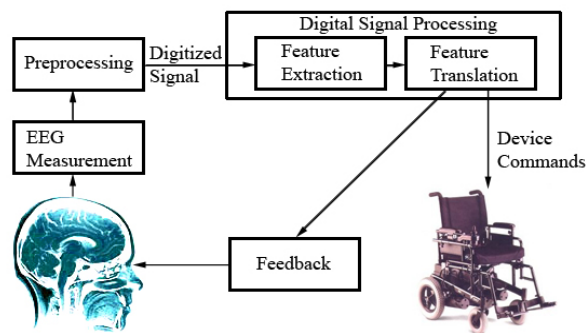


Fig. 1. Brain computer interface structure

V. BCI STRUCTURE

A brain computer interface (BCI) is a device that records EEG signals from the human brain, processes the signals into digitalized data and uses them as inputs to a computer system, as shown in Fig. 1. A BCI device consists of five different stages: measurement, preprocessing, feature extraction, feature translation and feedback [12], [13].

The measurement stage involves the conversion of the electrophysiological signals of the brain into electrical signals. The preprocessing stage is the stage that lies between the electrophysiological signals of the electrodes and the digital signals that will act as input to the computer. The feature extraction stage is one of the most significant parts of the BCI device because in this stage, the digitized EEG signals extract features that can be used to control the BCI. When the features of the EEG signal, to be used for control, are extracted, they are translated into device commands which will cause a desired device to act in a particular way. The feedback stage allows users to observe their performance and use the presented result to improve their current skills. Because of this fact, the feedback stage is sometimes referred to as the integral part of a BCI device.

VI. HARDWARE AND SOFTWARE

The hardware for this project includes all components essential to take electrophysiological signals from the human brain, pre-process these signals and transform them into digital data ready to be used by a computer. The headgear is very rigid and has some amplification circuitry attached to them, so we can ensure that the headgear is very close to the recording electrodes of the BCI.

The software digitized signals are used within the Labview design environment using the data acquisition card and the Labview software. The design environment itself has a visual interface and the building blocks of signal manipulation are inserted and connected as desired.

VII. CONCLUSION

With the field of brain control interfaces increasing everyday, expected that many other products available to aid

and assist disabled patients.

The results of the actual project will exceed expectations with brain activity being clearly detectable from the three regions of concern, the left motor cortex, the right motor cortex, and the ocular area. Patients will improve in every test and with further training they would be able to control the cursor with increased speed and accuracy. Such results, while not necessarily fast enough for a real time device, will be very encouraging for the future of BCIs.

Accuracy, speed, usability and feedback methods should be improved in the current BCI systems. Accuracy is the most important and affects greatly on the performance of the BCI.

Many of the BCI systems operated in a synchronous way, using trials lasting many seconds each, means that time required for making one selection is long. This time should be kept short preferably below one second. Feedback methods could also be improved.

In the future, an exhaustive research about the mental tasks should be done. A study of the left and right hand movements using high-resolution EEG and MEG is planned. Research topics would include the localization of the brain activity during the mental tasks and how the EEG changes in process of time. Other research areas would be feedback methods and online learning.

Many challenges are in the future of the BCI field; however, none of the BCIs are capable of proper cursor control, which could be used to control ordinary computer applications. In the near future, special applications must be developed for BCIs. Today, special writing applications or Internet browser can provide communication tools for severely disabled people. These applications could be improved. In the future, BCIs could be used to control a hand prosthesis. How well that can be achieved with EEG-based BCIs is not yet known. Non-invasive BCIs recording activity directly from the motor cortex may be used for this kind of purpose in the future.

REFERENCES

- [1] Jasper, H.H. 1958. "The ten-twenty electrode system of the International Federation," *Electroencephalography and Clinical Neurophysiology*, Vol. 10.46
- [2] Smith, R.C. 2004. "Electroencephalography based Brain Computer Interfaces". M. Engineering Thesis, University College Dublin, Dublin, Ireland.
- [3] Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G. and Vaughan, T.M. 2002. "Brain-Computer Interfaces for Communication and Control". *Clinical Neurophysiology*. Vol. 113, 767-791.
- [4] Andrew Rawicz, SFU School of Engineering Science, private communication.
- [5] Horowitz, P. 1989. *The Art of Electronics*. New York: Cambridge University Press.
- [6] Restak, R. 1984. *The Brain*. Toronto: Bantam Books.
- [7] Nunez, P. 1981. *Electric Fields of the Brain*. New York: Oxford University Press.
- [8] Carpenter, M. 1991. *Core Text of Neuroanatomy 4th ed*. Baltimore, Maryland: Williams & Wilkins.
- [9] Misulis, K. 1997. *Clinical Neurophysiology, 2nd ed*. Newton, MA: Butterworth-Heinemann.

ENSC 894 COURSE TRANSACTIONS

- [10] The Brain. <http://www.undergrad.ahs.uwaterloo.ca/~vmucci/thebrain.html>.
- [11] Anatomy of the Brain. <http://www.ahaf.org/alzdis/about/AnatomyBrain.htm>.
- [12] Allison, B. 1999. Brain Computer Interface Systems. <http://bci.ucsd.edu/>.
- [13] Ba'ar, E. 1998. Brain Function and Oscillations. Germany: Springer-Verlag.



Seyed Hossein Faramarzi Received the B.A.Sc in Biomedical Engineering in 2000, and he is currently pursuing his Master's degree in the Engineering Science at Simon Fraser University (SFU). Some of the notable achievements during his primary and secondary school years were:

- Ranked fifth in the Tehran Provincial Examination in 1991.
- Ranked seventh in the Tehran Mathematical Examination in 1994.
- Ranked nineteen in the National Mathematical Olympiad of Iran in 1995.

In 1996, he enrolled in Biomedical Engineering at Islamic Azad University of Tehran. He graduated from IAUT in August 2000 with a Bachelor's degree in Biomedical Engineering (Bio-electronics).

He started working at Jam Medical Equipment Company in August 1998. In this company, he was working as an electronic engineer, tasked with designing electronic circuits, and systems. Aside from his main task as an electronic engineer, he also performed investigation of electronic failures, supervision of technicians, development of maintenance and operating standards.

He is currently pursuing graduate studies at the Master's level in Simon Fraser University. He is keenly interested the field of Bio-electronics Engineering. My current area of research is continuing implementation of a two-dimensional cursor with click function using human cortical rhythms.

A Two-Phase Gaussian Filter (TGF) Algorithm for Neighbors Selection in Probabilistic Roadmap

Yifeng Huang, *Student Member, IEEE*, and Kamal Gupta, *Senior Member, IEEE*,

Abstract—When selecting neighbors for nodes inside a sample set in the Probabilistic Roadmap algorithm, we propose a Two-Phase Gaussian Filter (TGF) algorithm, which aims to archive efficient connection in workspace. The first phase, (i.e., Neighborhood Candidates selection), is based on Delaunay Triangulation of a random uniformly distributed points set. In the second phase, TGF can easily combine different kinds of C-space metrics, and then adopts a Gaussian function to sample neighbors from the Neighborhood Candidates set based on the combined C-space metrics. A new neighbors selection criterion, (i.e., Direction Metric), based on direction in C-Space is also introduced. We combine the Direction Metric with the scaled Euclidean Metric in TGF, and show the effectiveness of the algorithm in some preliminary experiments.

I. INTRODUCTION

Probabilistic Roadmap (PRM) planners have been used successfully for many high dimensional motion planning problems [1]. It constructs of a roadmap whose nodes and edges (usually a straight line in Configuration Space) are collision-free. The basic framework of PRM [1] uses uniform random samples in free space as milestones. These milestones are connected by edges, and the connectivity between pairs of milestones is tested by a local deterministic planner. Then a roadmap is created to capture the connectivity of the free C-Space. Researchers have realized PRM’s inability to connect across narrow passages in the configuration space [2]. The main reason for PRM having difficulty dealing with narrow passages is that the probability for a random sample to fall in a small volume of C-Space is small. Furthermore, connecting samples in narrow areas requires more samples being added into space, which may lead to a prohibitively large size of the roadmap.

Many variant techniques have shown encouraging results but the basic two-phase framework of PRM is adapted by most of newer versions, and specific improvement techniques are proposed in each phase of the algorithm. Most researchers [3][4][5][6][2] have focused on the first step of the roadmap construction, sample generation. They share a common feature, in that their performance relies on some global parameters that need to be chosen experimentally and carefully. In [7], the authors create a *visibility set* inside which each milestone works as a guard; thus, wide open areas will have small number of guards.

In this paper, we take a closer look at the second aspect of the Probabilistic Roadmap (PRM) algorithm, (i.e., node connection), to which less attention has been paid. Two crucial points are considered by most researchers with respect to the node connection: (i) distance metrics and (ii) neighborhood selection. Most PRM algorithms use a *Nearest*

Neighborhood Connection (NNC) strategy to determine the neighbors of a given node. In [1], a node will take the k nearest samples (evaluated by the distance metrics selected) as neighbors, with k upper-bounded by a constant parameter *MaxNumNeighbors*.

To save the unnecessary calling of the collision detector in PRM [8], the distance metrics must be chosen carefully such that it fits the requirements for both practical use and efficiency. Intuitively, we prefer those pairs of nodes, whose configurations are “close” to each other in workspace to be connected[1], but good metrics in the workspace usually take more time to compute. As a tradeoff, many researchers accept metrics in C-Space as the first choice for PRM. A commonly accepted distance metric is scaled Euclidean Distance [1], which is calculated in C-space, and the scales are given based on the kinematic characteristics of the robot. Other metrics (Minkowski, Manhattan, etc.) are used occasionally in some special cases. By selecting distance metrics in C-space, we expect that if two nodes are “close” in C-Space, their configurations will be “close” to each other in the workspace as well. However, such a relationship is not always the case for all the pairs of nodes due to the complication of mapping between C-space and workspace[4].

In this paper, we first propose a new distance metric called Direction Metric (DM), which is calculated in C-Space but has a stronger connection to the workspace distance with respect to other C-space metrics. Then, we introduce a framework for a Two Phase Gaussian Filter (TGF) neighborhood selection strategy, where the distance between a pair of nodes is used as the probability for them to be chosen as neighbors. We show how TGF allows us to easily combine the merits of those known C-space metrics along with our new DM. Our combined metric involves more time on distance computation, but it will only be used within a so-called Neighborhood Candidates (NCs) set, which is generated by the first phase of the TGF and upper-bounded by a constant on size. To obtain the NCs set, we adopt the traditional Euclidean Metrics to calculate a bound R_d ; only those nodes inside the bound will be considered as Neighborhood Candidates[9]. The bound R_d is calculated base on Delaunay Triangulation of the uniform random set. Characteristics of this bound are discussed in detail later.

II. PRM REVIEW

The general algorithm, as described in [1] is as follows:

Basic PRM Algorithm:

- 1) $N \leftarrow \emptyset$

- 2) $E \leftarrow \emptyset$
- 3) **loop**
- 4) $c \leftarrow$ a randomly chosen node in C_{free}
- 5) $N_c \leftarrow$ a set of candidate neighbors of c from N
- 6) $N \leftarrow N \cup \{c\}$
- 7) **for** all $n \in N_c$, in order of increasing $D(c, n)$ **do**
- 8) if there exists a local path between c and n **then**
- 9) $E \leftarrow E \cup \{(c, n)\}$

Of most interest here is the *Nearest Neighborhood Connection* (NNC) strategy used in most PRM algorithms. To determine its neighbor nodes (the ones the node is connected to via a collision-free edge), a node will select the neighborhood set N_c based purely on the Euclidean distance function. In step-5 of the **Basic PRM Algorithm**, a threshold $Maxdist$ is used to bound the set of N_c to be the samples located inside the ball that is centered at c and has a radius of $Maxdist$. Another constant, $MaxNumNeighbors$, bounds the size of the set N_c and guarantees that the ratio between the number of the edges and nodes is approximately fixed. In [1] and [10], those numbers are chosen experimentally, $MaxNumNeighbors$ is typically on the order of 30 when dimension is 6.

III. DISTANCE METRICS

The selection of distance metrics is one of the key issues in PRM since we prefer a node to be connected with others “close” to itself to save the unnecessary call to the local planner. Evaluating a distance metric’s quality is a difficult problem. Many different metrics have been introduced in the frame work of PRM based on the kinematics of the robot. Below are some general factors required to be considered when we choose a metric for PRM:

- 1) The metric must be easy to compute. Since the distance calculation is called $O(n^2)$ times (n is the size of the sample set), a complicated distance metric obviously increase the computational burden of the PRM
- 2) PRM works in C-space, but connectivity in workspace should be give more consideration.
- 3) Kinematics of the robot plays an important role for distance metric selection

Because direct evaluation of distance in workspace sometime involves too much computation[4], a widely accepted alternative is to use the distance in C-Space. C-Space metrics are easy to calculate, but some commonly used C-Space distance metrics are sometime do not provide very efficient connections in workspace.

A. A Direction Metric (DM) for Neighbors Selection

We introduce a new Direction Metric (DM) in C-Space. This metric takes advantage of some experimentally found properties of uniform random sample set. Based on those properties, we first select a Neighborhood Candidates (NCs) set for a node. The final neighbors of a node will be selected inside this Neighborhood Candidates set. Our Direction Metric will be used to evaluate the distance between a node and its NCs only.

Let c be the configuration of a node p that we are trying to select neighbors for. and c' any other nodes’ configuration in

C-Space. Let $\bar{c} = c - c'$ and $\bar{c}_n = Norm(\bar{c})$, where $Norm(\cdot)$ is the normalization function. Then the distance under DM of two configuration is given by:

$$dis_u(c, c') = f(c, \bar{c}_n),$$

where function $f(\cdot)$ is selected based on kinematics of the robot.

Different from traditional metrics in C-space, $f(\cdot)$ will lead to the same results for a set of configuration pairs which are on the straight line in C-space. As we mentioned before, we only use this metric to evaluate the distance between a node and its Neighborhood Candidates(NCs) set, which is calculated using the Euclidean Metric. To acquire the NCs set of a node p , we set a bounding ball area which is centered at p and has radius R_d . Only those nodes within the bound (the Euclidean distance between p and other nodes less than R_d) will be considered as Neighborhood Candidates, denoted as N_{rd} of p . Because the nodes are random uniformly distributed, we can select R_d such that the vectors from p to those NCs tend to be approximately along different directions in C-Space (refer section-IV for detail).

The key idea of the distance function $f(\cdot)$ is that given c , we learn precisely the configuration of the robot in workspace. At the same time, unit vector, \bar{c}_n , will tell us how the robot will move in the real world. Thus, we can straightforwardly learn how the robot move when both c and \bar{c}_n are specified.

When using PRM to solve a motion planning problem for a 4-link manipulator, we assume to uniformly distribute nodes in C-Space. Usually, we take the volume swept by the robot as a good heuristics of distance evaluation in the workspace. The volume swept can be coarsely evaluated by the product between the speed that each link sweeps in the workspace and the Euclidean distance between the two node. The sweep speed for each link is fully specified by c and \bar{c}_n . Thus, to locate those NCs “close” to a node p in workspace, we could assign a smaller distance value to those candidate nodes, whose corresponding \bar{c}_n tend to sweep volume “slower”.

For a planar manipulator in our experiment, we simply define $f(\cdot)$ as:

$$f(c, \bar{c}_n) = \sum_{i=0}^{d-1} |x_i + x_{i+1}|, \quad (1)$$

which is independent of c . Here x_i is the i_{th} components of vector $\bar{c}_n = (x_1, x_2, \dots, x_d)$, where d is the dimension of the C-Space. In Equ-1, we will acquire a lower value: when a link moves at a degree of positive value or negative value, and the next link closer to the end effector moves at a degree of negative value or positive value, respectively. In another word, the robot biases to the motion such as “reaching” or “shrink”, and sweeps volume relatively slowly in workspace.

Our idea is based on heuristics similar to applying RRT in humanoid robot [11]: when robot reaches a new configuration, it will first try a set of “smart” movements, which are defined by $f(\cdot)$ in our case. Our work is different from [11], in that we define $f(\cdot)$ in the node connection stage of the PRM. But [11] do not explicitly define $f(\cdot)$, in stead of which a set of predefined configurations are given by the user.

Obviously, specific definition of $f(\cdot)$ depends heavily on the kinematics of the robot. Finding a systematic approach to determine $f(\cdot)$ is a difficult problem. We could imagine a more sophisticated definition of $f(\cdot)$ than Equ-1, by cutting C-space into blocks, and considering the correspond movement in each block separately (thus taking c in Equ-1 into accounts as well). However, we do not need a complicated $f(\cdot)$ in most of practical cases, since we could combine it with other classic C-space metrics to archive more efficient connections for PRM.

B. Combination of Different metrics

Since different C-space metrics have their own pros and cons [4], it may be a good idea to use the combination of different metrics. A simple choice is to use the linear combination of the selected metrics as following:

$$dis_f = \sum_{m=0}^M \alpha_m * dis_m,$$

where $\sum_{m=0}^M \alpha_m = 1$, dis_m stands for the distance given by the m_{th} metrics, and M is the number of the metrics we are using. Here, α_m is chosen experimentally.

The idea mentioned above is easy to implement, but can hardly be used practically since common C-space metrics range much different. In other word, evaluation of the weight α_m is difficult. One alternative is to use the ranking of the nodes when ordering under different metrics.

Let N_{rd} , (i.e., the Neighborhood Candidates set), be the nodes inside the bound R_d with centering at p . We try to decide for node p which nodes inside N_{rd} will be selected as neighbors of p . To evaluate the “distance” from a node, $p_i \in N_{rd}$ to p , we use the ranking instead of the real values given by different metrics.

$$dis_r = \sum_{m=0}^m \alpha_m * P(\mathfrak{R}_m), \quad (2)$$

where \mathfrak{R}_m is the ranking of node p_i when the m_{th} metric is used, the closer p_i is to p , the higher the ranking is. $P(\cdot)$ is a function of the ranking \mathfrak{R}_m that we will discuss below.

C. Gaussian Neighborhood Filter

The combined metric based on ranking can be directly used for neighbors selection by choosing the nearest k neighbors. However, we are more interested in a random neighborhood selection strategy.

When using m_{th} metrics, we assign a probability g_m for each node p_i based on its ranking \mathfrak{R}_m , and acquire a final summation of g_m over all the M metrics (see Equation-2). After we normalize the dis_r for all the nodes in candidate set as dis_{rn} , we select node p_i with a probability corresponding to its dis_{rn} . In our experiments, the value g_m follows the Gaussian distribution:

$$g_m = P(\mathfrak{R}_m) = \frac{1}{\sqrt{2\pi}\mathfrak{R}_m} \exp\left(-\frac{\mathfrak{R}_m^2}{2\sigma^2}\right).$$

We set the parameter σ equal to the size of the sample set $|N_{rd}|$. The idea behind our method is that we give a higher chance for a node to be selected if it is “closer” to p , but we will also give a node, which is “far” away from p , an opportunity to be selected. In technique mentioned above, we expect to compensate the possibility of missing some long distance neighbors (in C-Space), which are preferable in workspace. Such a random policy also has an advantage of increasing the connectivity in C-Space. For instance, it can reduce the possibility of a new node repeatedly connects to a close part of the Roadmap along approximately one direction. When a new node is added very close to an old node, this random method can also avoid connecting the new node to the same neighbors as the old one.

When selecting neighbors for a node p using Gaussian Filter, we need not sort the distance between p and all the other nodes in C-space. We will only rank those nodes inside the Neighborhood Candidates set, whose size is a constant and modest with respect to the sample size. The computation burden increased by this step will not be a significant factor for the whole algorithm, as shown in our experiments.

IV. NEIGHBORHOOD CANDIDATES SELECTION BASED ON DELAUNAY TRIANGULATION

The first step of the two-phase neighbor selection involves a bound calculation. We made this bound as tight as possible but left enough high quality candidates at the same time. We will acquire this bound based on empirically observed properties of Delaunay Triangulation of a random uniformly distributed point set. A detailed explanation of the work can be found in [9]. Here we only give a brief introduction of the key idea. Throughout this paper, we use Euclidean Metrics to calculate the Neighborhood Candidates set.

A. Delaunay Triangulation Based Candidate Selection Strategy

Given a set S of n distinct points in R^d , the Delaunay Triangulation is the geometric dual of the Voronoi Diagram. Delaunay Triangulation and Voronoi Diagram of a points set S have been researched for decades in computational society. A review of related topics can be found in [12].

Our Delaunay Triangulation Candidate Selection (DTCS) Strategy aims to choose the neighbors candidate adjacent to a node, say c , based on Delaunay Triangulation. A node in configuration space will be a Delaunay Neighbor (DN) of node c iff their Voronoi cells are adjacent to each other, or alternatively, they belong to the same Delaunay cell. The properties listed below shed some lights on why we choose DTCS.

When a node $p1$ that is NOT DN of node p , let O_b be a ball centered at midpoint (denote as c) of the line segment $(p, p1)$ with radius that is equal to length of the line segment (c, p) (or $(c, p1)$) then there exists another node p' , which is DN of p , inside the ball O_l [13]. The property above then gives us a reason to consider only DN of a node as Neighbor Candidate. A simple scenario illustrates in Fig-1 (right) that when going from point B to A , which is NOT DN of B ,

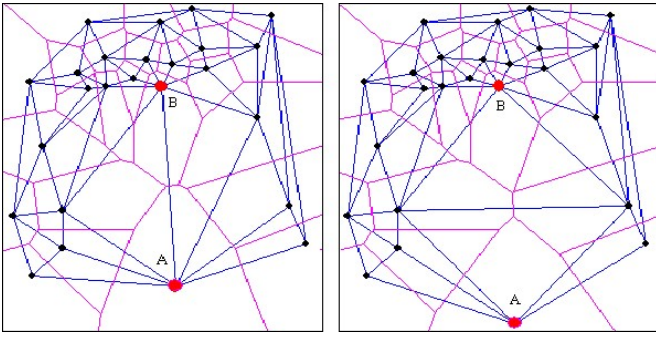


Fig. 1. Delaunay Triangulation build based on both distance and direction information

the robot will “meet” some DNs of B first “approximately” on its way to point A. Then obviously a better policy is to connect to DNs first than to nodes that are NOT DNs. Since the number of the DNs (as shown in [9]) are tend to be a constant that exponential with respect to dimension of the C-space, constructing network based only on DN already lead to a graph that is heavily dense. Thus it will be “sufficient” to only consider the DN as Neighbor Candidates.

Another scenario illustrates that the Delaunay Triangulation network is constructed based not only on “distance” but also on “direction” information. Consider Figure-1. In the left figure, even if point A is far away from B, it will still be connected to B so long as it is the “closest neighbor in that direction”. In the figure on the right, as A is moved farther along the direction, it may get disconnected from node B because more points appear approximately along \overline{BA} . Intuitively if a node c has some neighbors approximately along a particular direction, nodes further away along the same direction will not be direct neighbors of this node. While if a node doesn’t have neighbors along a direction, a node along this direction will be its Delaunay neighbor even though it is sort of far away from c. Hence the Delaunay neighborhood selection for a uniformly distributed point set keeps a *balance* between “distance” and “direction”. We can also see that if we miss some DNs as our Candidates, we lose neighbors in some “wide open” direction, that will obviously “damage”, to some extend, the connectivity of the overall network. This lead our believe the “necessaries” of select DNs as Neighborhood Candidate.

B. An approximation of Delaunay Triangulation Neighborhood Candidates Selection

Explicit calculation of Delaunay Triangulation will output all the Delaunay cells and a data structure that facilitates visiting all the cells, but its complexity is exponential with respect to the dimensionality, and hence intractable when the dimensionality is high [14]. However, we need not build an explicit Delaunay Triangulation. We are only interested in finding those pairs of nodes when they belong to same Delaunay Cell, or alternatively, they are Delaunay Neighbors (DN). The problem of deciding if two nodes are Delaunay Neighbors or not, can be posed as a linear programming (LP) problem. The description of linear programming problem

presented here is roughly based on the algorithm described in [15].

The Voronoi diagram of a set S of n points in R^d is the projection of the following $(d+1)$ polyhedron to R^d space of the first d components:

$$P = \{x \in R^{d+1} \mid \sum_{j=1}^d p_j^2 - \sum_{j=1}^d 2p_j x_j + x_{d+1} \geq 0 \forall p \in S\}.$$

For simplicity, denote it as:

$$P = \{x \in R^{d+1} \mid b - Ax \geq 0\},$$

where A is a given $n \times (d+1)$ matrix and b is an n -vector. Now, we formulate the following LP for any distinct $i, j = 1, 2, \dots, n$

$$\begin{aligned} & \text{minimize} && f(x) := b_j - A_j x \\ & \text{subject to} && b' - Ax \geq 0 \\ & && b_i - A_i x \leq 0, \end{aligned}$$

where A_j stands for the j^{th} row of A and vector b' is equal to b except that the j^{th} component of b' , $b'_j = b_j + 1$.

The *Voronoi cell* of points i and j are adjacent, i.e, i and j will be Delaunay Neighbors, if and only if the objective value $f(x)$ is negative at an optimum solution.

So we can decide whether two nodes are Delaunay Neighbors by solving a LP problem which has $n+1$ constraints and can be solved in polynomial time. Unfortunately, we need to solve $O(n^2)$ LP problems to construct the whole Delaunay network. Because each LP problem includes $n+1$ constraints and every constraint has in turn d variables for d dimensional space, building the Delaunay network exactly is computationally expensive.

We suggest an alternative efficient approach that will build a network that is “similar” to Delaunay network. The motivation behind our approach comes from some empirically observed statistically expected properties of Delaunay Neighbors. A detailed algorithm can be found in our previous work [9]. The key point is that some experimental observations show the relationship between the nearest neighbors and Delaunay Neighbors. Based on this relationship, a set of bounds R_d are calculated for different sample sizes and stored in a look-up table T for each dimension d . To construct network approximately close to Delaunay network, we can simply connect a node with all the other nodes inside a ball (centered at itself) of radius R_d , which is looked up in table T based on size of the sample set.

V. TWO-PHASE RANDOM GAUSSIAN FILTER (TGF) ALGORITHM

The number of Delaunay Neighbors increases exponentially with dimension. Therefore, connecting a node with all its Neighborhood Candidates becomes impractical in high dimension space. Similar to standard PRM, we use a parameter, $MaxNumNeighbors$, as an upper-bound to the number of neighbors that a node will select. In our previous work [9], we introduce a Delaunay Triangulation based Connection (DTC) algorithm, in which we randomly choose

$MaxNumNeighbor$ neighbors from the Neighborhood Candidates set. In this work, we weight each node in Neighborhood Candidate set as shown in Section-III-B, and sample a node as neighbor based on the its associated possibility.

Two Phases Random Gaussian Filter algorithm (TGF):

- 1) For a node i , select the radius R_d from the look-up table T based on current number of uniform nodes.
- 2) Store all the neighbors, whose distance to i is less than R_d , in a list L .
- 3) Connect the node i to its nearest neighbor.
- 4) If length of list L is less than or equal to $MaxNumNeighbors$, connect the node i with all the nodes in L .
- 5) Else, compute the distance dis_r between node i and all the nodes in list L based on equation - 2 and normalize all the dis_r as dis_{rn} . Randomly choose $MaxNumNeighbors - 1$ neighbors from list L as neighbors of i , the probability for a node to be selected is equal to its corresponding normalized dis_{rn} .

VI. EXPERIMENTAL RESULTS

We proceeded tests with TGF algorithm in 4D and 6D configuration spaces, and set the parameter $MaxNumNeighbors$ as 8 and 20, respectively. The planner was implemented using Motion Planning Kernel(MPK) Library [16] developed at Simon Fraser University. All the experiments were run on a 2.0Ghz Pentium4 PC. The collision detector used in all the experimenters is V-Collide library developed at University of North Carolina [17].

In Environment A and B (see Fig-2 left and right respectively), a planar manipulator needs to go through some narrow passages to reach the goal. Environment C (Fig-3) is a 6D planning problem. Configurations in light color represent the start and goal, respectively. Dark colored configurations are some intermediate configurations along the solution path. The sample generation strategies in all our experiments are the same. We combine two basic distance metrics, i.e., scaled Euclidean Metric and Direction Metric, to weight the Neighborhood Candidates using the Gaussian Filter. The weights for the scaled Euclidean Metric are the same as in [1]. The function of Direction Metric is shown in Equation-1. In all the experiments, we assign three different metrics weights for scaled Euclidean Metric: 0.8, 0.2, and 0.

We performed 30 trials for each example. Table I shows the average time and the number of nodes left in the roadmap, when the problems are solved. The combined metric shows improvement compared to single metric and the result tends to favorite a higher weight to our Direction Metric with respect to the traditional scaled Euclidean Metric.

VII. CONCLUSION

To achieve an efficient connection for PRM, we introduce a Two-Phase Gaussian Filter algorithm that can easily combine the merits of different C-space metrics. Our NCs selection

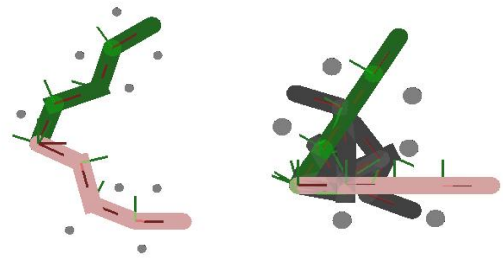


Fig. 2. Experiment A (left) and B (right) – 4-link planner manipulator with fixed base

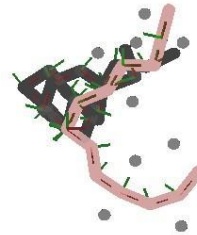


Fig. 3. Experiment C – 6-link planner manipulator with fixed base

in the first phase allows to combined the selected metrics without increasing the computation complexity for PRM. We also define a new C-space Direction Metric, and show its effectiveness when combined with scaled Euclidean Metric. In the second phase of TGF, we randomly sample a node’s neighbors from the Neighborhood Candidates based on the possibility associated with each candidate. Experiment results show clearly the efficiency of the proposed algorithm.

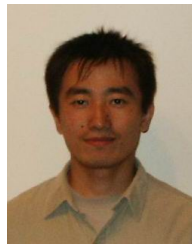
REFERENCES

- [1] L. E. Kavraki, P. Svestka, J. C Latombe, and M. H. Overmars. Probabilistic roadmaps for path planning in high-dimensional configuration spaces. *IEEE Transactions on Robotics and Automation*, 12(4):566–580, Aug. 1996.
- [2] D. Hsu, L. E. Kavraki, J-C. Latombe, R. Motwani, and S. Sorkin. On finding narrow passages with probabilistic roadmap planners. *P.K. Agarwal, L.E. Kavraki, M.T. Mason, editors, Robotics: The algorithmic perspective*, pages 141–154, 1998.
- [3] V. Boor, M. H. Overmars, and A. F. Stappen. The gaussian sampling strategy for probabilistic roadmap planners. In *Proceedings 1999 ICRA. IEEE International Conference on Robotics and Automation*, volume 2, pages 1018–1023. Carnegie Mellon Univ., Pittsburgh, PA, USA, IEEE, May 1999.
- [4] N. M. Amato, O. B. Bayazit, L. K. Dale, C. Jones, and D. Vallejo. OBPRM: An obstacle-based prm for 3d workspaces. In *P.K. Agarwal et al., editors, Robotics: The Algorithmic Perspective: 1998 Workshop on the Algorithmic Foundations of Robotics*, pages 155–168, 1998.
- [5] S. A. Wilmarth, N. M. Amato, and P. F. Stiller. Mapprm: A probabilistic roadmap planner with sampling on the medial axis of the free space. In *Proceedings 1999 ICRA. IEEE International Conference on Robotics and Automation*, volume 2, pages 1024–1031. Carnegie Mellon Univ., Pittsburgh, PA, USA, IEEE, May 1999.
- [6] D. Hsu, T. Jiang, J. Reif, and Z. Sun. The bridge test for sampling narrow passages with probabilistic roadmap planners. In *Proceedings 2003 ICRA. IEEE International Conference on Robotics and Automation*, volume 3, pages 4420–4426. Carnegie Mellon Univ., Pittsburgh, PA, USA, IEEE, May 2003.
- [7] C. Nissoux, T. Simeon, and J-P. Laumond. Visibility based probabilistic roadmap. In *Proceedings. 1999 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 1999)*. Fraunhofer Inst. for Comput. Graphics, Darmstadt, Germany, IEEE, Oct 1999.

Pro.	Connector*	N	Time(sec.)
A	NNC-W	2567	112.3
	DTC	919	76.2
	TGF-W-0.8	1021	78.9
	TGF-W-0.2	884	56.6
	TGF-W-0.0	901	61.4
B	NNC-W	360	16.14
	DTC	273	11.29
	TGF-W-0.8	278	10.82
	TGF-W-0.2	244	9.29
	TGF-W-0.0	280	11.87
C	NNC-W	2492	1500
	DTC	1185	617
	TGF-W-0.8	900	446.9
	TGF-W-0.2	780	400.7
	TGF-W-0.0	1272	569.3

*NNC-W: NNC with Weighted Euclidean Metrics
 DTC: DTC Random Select inside Neighborhood Candidates
 TGF-W-0.8: Two Phase Gaussian Filter inside Neighborhood Candidates (favorite Weighted-Euclidean Metrics)
 TGF-W-0.2: Two Phase Gaussian Filter inside Neighborhood Candidates (favorite Direction Metrics)
 TGF-W-0.0: Two Phase Gaussian Filter inside Neighborhood Candidates (use Direction Metrics only)

TABLE I
 EXPERIMENTAL RESULTS FOR PROBLEM A, B AND C



Yifeng Huang received the M.Sc. degree in the automatic system from the Shenyang Institute of Automation, China Academic of Science, China, in 1999.

From 2000 to 2001, he worked as a system engineer in Shenyang Hua Yan Electric and Power Inc., China. Presently, He is a Ph.D Candidate in Robotic Lab, Engineering Science, Simon Fraser University, CANADA. His main interests are robotic learning, Computational Geometry, and 3D modeling.



Kamal Gupta received the Ph.D. degree in electrical engineering from McGill University, Montreal, P.Q., Canada, in 1987.

Since then he has been a faculty member in the School of Engineering Science, Simon Fraser University, Burnaby, B.C., Canada. His research interests and contributions are in the geometric aspects of robotics and automation, in particular motion planning, manipulation and geometric reasoning, and 3-D vision for robotic tasks. He has also consulted for robotics and automation companies in these areas. He has held visiting scientist positions at INRIA, Rhone-Alpes, France, and at the Robotics Laboratory, Stanford University, Stanford, CA, from 1993 to 1994. He is a co-editor of *Practical Motion Planning in Robotics: Current Approaches and Future Directions* (New York: Wiley, 1999).

Dr. Gupta serves as a co-chair of the IEEE Robotics and Automation Society's Technical Committee on Motion and Path Planning.

[8] N. M. Amato, O. B. Bayazit, L. K. Dale, C. Jones, and D. Vallejo. Choosing good distance metrics and local planners for probabilistic roadmap methods. *IEEE Transactions on Robotics and Automation*, 16:442–447, Aug. 2000.

[9] Y. Huang and K. Gupta. Delaunay triangulation based node connection strategy for probabilistic roadmap planners. In *Proceedings 2004 ICRA. IEEE International Conference on Robotics and Automation*, volume 2, pages 908–914. Carnegie Mellon Univ., Pittsburgh, PA, USA, IEEE, May 2004.

[10] R. Bohlin and L. E. Kavraki. Path planning using lazy prm. In *Proceedings 2000 ICRA. IEEE International Conference on Robotics and Automation*, volume 1, pages 521–528. Carnegie Mellon Univ., Pittsburgh, PA, USA, IEEE, April 2000.

[11] J.J. Kuffner and S.M. Lavalle. Rrt-connect: An efficient approach to single-query path planning. In *Video Proceedings 2000 ICRA. IEEE International Conference on Robotics and Automation*. Carnegie Mellon Univ., Pittsburgh, PA, USA, IEEE, May 2000.

[12] D. F. Watsou. Computing the n-dimensional delaunay tessellation with application to voronoi polytopes. *The Computer Journal*, 24(2):167–171, 1981.

[13] M. D. Berg, M. V. Kreveld, M. Overmars, and O. Overmars. *Computational Geometry*. Springer Verlag, 1997.

[14] K. Mehlhorn and M. Seel. Delaunay triangulations in higher-dimensional space, implementation report. <http://www.mpi-sb.mpg.de/LEDA/www/leps/ddgeokernal.html>, 1998.

[15] K. Fukuda. Frequently asked questions in polyhedral computation. <http://www.ifor.math.ethz.ch/staff/fukuda/polyfaq/polyfaq.html>, 2003.

[16] I. Gipson, K. Gupta, and M. Greenspan. Mpk: An open extensible motion planning kernel. *Journal of Robotic Systems*, 18:433–443, Aug. 2001.

[17] A. Pattekar, J. Cohen, T. Hudson, S. Gottschalk, M. Lin, and D. Manocha. V-COLLIDE User's Manual - Release 1.1. Department of Computer Science, University of North Carolina at Chapel Hill, 1998.

Robot Path Planning Using Harmonic Function-based Probabilistic Roadmaps

Moslem Kazemi, *Student Member, IEEE*, Mehran Mehrandezh, *Member, IEEE*, and Kamal Gupta, *Member, IEEE*

Abstract—A new hybrid robot motion planning approach based on Harmonic Functions (HF) and Probabilistic Roadmaps (PRM) is presented. The proposed approach consists of building a Probabilistic Roadmap using information obtained about the robot's Configuration Space (C-space) topology through a Fluid Dynamic (FD) paradigm based on the HFs. The crux of our approach is to identify narrow passages using a FD paradigm and pass them over to a PRM-like method to build a roadmap that efficiently captures the connectivity of the free C-space, especially in narrow regions. We present two techniques to implement the proposed approach: (1) the original Harmonic Function-based Probabilistic Roadmap (HFPRM) technique where the HFs are computed over the explicit representation of the robot's C-space, and (2) the Incremental HFPRM (IHFP) technique which is an extension of the original HFPRM, and can be applied to virtually any type of robot.

Simulation results obtained using these techniques show that the combination of the HF and the PRM works better than each individual technique in terms of finding a collision-free path in environments where narrow passages exist. Inspired by the promising results obtained through the original HFPRM and incremental update of it, we also present a sensor-based motion planning technique for robotic navigation in environments not known *a priori*. Experimental results obtained using this technique to navigate a mobile robot equipped with ultrasonic range finders in an unknown environment are also presented.

Index Terms—Robot Path Planning, Potential Field, Probabilistic Roadmap, Harmonic Functions

I. INTRODUCTION

Given an environment cluttered with obstacles the basic robot Motion Planning (MP) problem is established as: moving a robot from a start configuration to a goal configuration without colliding with obstacles. This problem can be divided into two main categories: (1) Model-Based MP (MBMP) where the environment is completely known, and the robot's task is to find a collision free path from the start configuration to the goal configuration, and (2) Sensor-Based MP (SBMP) where the environment is unknown and the task for the robot, equipped with on-board sensors, is to simultaneously explore the environment and reach a given goal configuration.

We present a new hybrid robot motion planning technique, combining the potential field approach based on Harmonic Functions (known as harmonic potential fields) and the Probabilistic Roadmaps (PRM). Harmonic potential field is a subclass of the Potential Field (PF) approach, where the potential satisfies Laplace's equation [1]. Harmonic functions constitute a powerful mean to generate local minima free smooth paths. However, possible deviation from one streamline to another when navigating a robot may cause chattering in motion [2]. Moreover searching through the calculated streamlines

for finding the shortest path connecting the start and goal configurations will be computationally cumbersome. The PRM is based on a random sampling scheme to construct an implicit representation of the robot's C-space as a roadmap [3]. The efficiency of the PRM in solving the classical robot motion planning problem is reduced in environments where narrow passages exist.

Simulation results show that the combination of the potential field approach based on the HFs and the PRM can overcome each individual technique's shortcoming in finding a collision-free path between start and goal configurations in environments where narrow passages exist. The proposed MP approach can be utilized in autonomous robot MP with applications in robotic search-and-rescue, and unmanned space exploration and construction. Two techniques are proposed: (1) the Harmonic Function-based Probabilistic Roadmap (HFPRM) that is applicable to MP of robots with low Degrees Of Freedom (DOF), and (2) the Incremental HFPRM (IHFP) technique which is more general and can be applied to virtually any type of robot.

In the HFPRM technique the HFs are computed over the explicit representation of C-space. A PRM with a novel sampling scheme is then constructed based on information obtained about the C-space topology through HFs. The obtained roadmap is searched for the shortest path. Promising results are obtained using this technique for solving shortest-path problem in difficult scenarios.

Since explicit representation of the C-space for robots with many DOFs is not practical, an incremental version of the HFPRM technique has been also developed which is more general. In this technique a roadmap is enhanced and updated in an iterative scheme based on information obtained through HFs. At each iteration HFs are computed over the current representation of the C-space which is incrementally constructed. Inspired by the promising results obtained through the original HFPRM and incremental update of it, the proposed approach was implemented on a mobile robot equipped with ultrasonic range finders moving in an unknown environment.

The rest of the paper is organized as follows. In the next section we report on literature cited on PF approaches, PRM methods, and some hybrid techniques combining these two to address the motion planning problem. We will also briefly mention the literatures which address the SBMP problem for robots in unknown environments. Section III explains the problem formulation pertinent to both model-based and sensor-based motion planning techniques. Section IV presents the proposed solution methodology for model-based motion planning of robots in known environments along with the

simulation results. Section V explains our proposed sensor-based path planning approach for mobile robots moving in environments not known *a priori*. The experimental results obtained using our proposed technique to navigate a mobile robot in an unknown environment are presented in Section VI. Section VII concludes the paper.

II. RELATED WORKS

A. Potential Field Approaches

The basic idea of potential field approaches is to construct a potential field with an attractive global minimum at the goal configuration and a repulsive local maximum at the obstacles. Potential fields have been introduced as a practical and deterministic approach in robot path planning [4]. The harmonic potential field is an important class of the potential field approaches based on HFs, which provides local minima free solutions. Connolly et. al. proposed a robot motion planning method based on harmonic functions. Local minima free paths were obtained in static environments [1]. Latombe and his group proposed a PF approach that consists of incrementally building a graph connecting the local minima of a PF defined in the robot's C-space and concurrently searching this graph until a goal configuration is attained [5].

The ability to incrementally update the harmonic functions allows one to use them in dynamic as well as stationary environments. Masoud et. al. generalized the harmonic potential approach for tracking a maneuvering target that is moving amidst known stationary obstacles [6]. A recursive motion planning scheme using harmonic functions without complete knowledge of the environment *a priori* has been proposed in [7].

Different physical analogies have been employed to adapt the harmonic potential fields for robot path planning. To include the external forces which influence the robot performance in uneven terrains and also to consider clearance between the robot and obstacles, Louste and Liegeois ([8]) presented a path planning method using the incompressible but viscous flow equations. Wang and Chirikjian [9] propose a potential field approach based on a steady-state heat transfer analogy. In their methodology, the optimal path is presented as a heat flow with minimal thermal resistance. An electrostatics terminology has been proposed in [10] and [11]. In our proposed method we adopt a fluid dynamic paradigm based on an incompressible non-viscous fluid flow to bias the distribution of random nodes in a PRM towards narrow regions.

There are some drawbacks in navigating robots based on HFs which limit their usage. For instance, chattering in motion when following a streamline obtained through the harmonic functions may cause deviation from one streamline to another [2]. Besides, finding the shortest streamline needs a global search technique and can be computationally cumbersome.

B. Probabilistic Roadmap Methods

The PRM method comprises two phases; In the first phase, i.e. *pre-processing* phase, a predefined number of random nodes are generated, and a roadmap is constructed by connecting the neighboring nodes based on an Euclidean metric.

In the second phase, i.e. *query* phase, the start and goal configurations are attempted to be connected to the roadmap obtained in the pre-processing phase, and the roadmap is then searched for the shortest path connecting the start and goal configurations.

This technique has been the method of choice for MP of robots with many degrees of freedom. However, constructing a roadmap that can efficiently capture the connectivity of the free C-space is a challenging issue [12], [13]. The PRM technique can simply fail in environments where narrow passages exist [12].

Many variants of the PRM have been proposed to identify narrow regions (e.g. [13], [14], [15]). Amato et. al. evaluate various node generation and connection strategies in an Obstacle-Based Probabilistic Roadmap (OBPRM) [13]. Bayazit et. al. developed a technique to enable a human and an automatic planner to cooperatively solve motion planning problems to discover critical configurations using haptic hints [14]. The narrow passage problem was also addressed in [15]. They developed a medial axis approach based on a sampling heuristic as a guideline for choosing configurations in critical regions. Hsu et. al. utilized a free space dilation concept for widening narrow passages to guarantee free space connectivity in those regions [12]. A bridge test strategy has been proposed in [16] which boosts the sampling density inside narrow passages based on simple tests for local geometry. The bridge test strategy reduces the sampling density in wide-open collision-free regions. Berg and Overmars proposed a method which uses the shape of free workspace to guide the sampling in C-space [17]. This technique can only be used when the C-space resembles the workspace. A variant of PRM method based on a visibility notation has been introduced in [18] and [19] to generate roadmaps with small number of nodes. In their proposed method a random configuration is added to the roadmap if either it connects two connected components of the roadmap or is not "visible" by some so-called guard configurations.

Some other random sampling approaches have been proposed based on the construction of search trees rather than graph-based roadmaps. Rapidly-exploring Random Trees (RRT) and Ariadne's Clew Algorithm (ACA) are two independent examples under this category [20], [21]. Reliable identification of narrow passages using random planners is an open research area.

C. Hybrid Techniques

Due to shortcomings of the PF approach and the PRM method in identifying narrow regions some hybrid techniques have been proposed utilizing the advantages of each individual. A hybrid technique based on Harmonic Potential Functions and random sampling of the non regular grid decomposition of the robot's C-space was presented in [22]. Harmonic functions are utilized to bias the random sampling of the cells, and to explore the C-space. They utilized a hierarchical technique ([23]) to compute the harmonic functions on a non-regular grid decomposition of C-space. The gradient of the potential field is followed to find a sequence of free cells from the

start configuration to the goal configuration. Efficiency of these technique in finding collision-free path in presence of narrow passages has not been addressed. An artificial potential based probabilistic roadmap method has been proposed in [24]. In their technique a partial computation of potential field in workspace is utilized to bias the sampling distribution in a PRM approach in C-space. This yields denser sampling of the C-space close to obstacle boundaries, hence a denser coverage in narrow regions. Nonetheless, a high number of samples can be generated in irrelevant narrow regions (i.e., regions with dead-ends) using this method. Furthermore this technique will not generate a uniform coverage on the wide-open regions.

D. Sensor-Based Motion Planning Techniques

Sensor-based MP for robots has been extensively studied for the past two decades. We mention some of them which are related to our work. *View Planning* problem, i.e. finding the best next configuration where the next scan should be carried out, is one of the main issues in SBMP problem.

Renton et. al. presented a sensor-based robotic system, called Plan-N-Scan, using a wrist mounted laser range finder [25]. In their proposed technique an A^* search is used to select the best view configuration among all the possible free configurations to scan a region (view planning). The A^* search can be costly in both time and memory as the robot's DOF increases. In their system, to explore the unknown regions of the environment, higher priority is given to the region which is closest to the origin of the workspace, i.e. the initial robot's configuration.

Gupta and his group developed a Sensor-Based Incremental Construction of Probabilistic Roadmaps (SBIC-PRM) approach [26]. In their technique a probabilistic roadmap is incrementally constructed as the robot senses new free spaces in the physical space. The roadmap represents the free C-space connectivity within which the robot carries out its motion to further sense the unknown parts of the environment. They successfully implemented their technique on an Eye-in-Hand system consisting of a 6 DOF PUMA 560 manipulator equipped with a laser range finder mounted on robot's end-effector. To address the view planning problem they proposed an information theoretic approach which maximizes the information gained due to a scan for optimal exploration of the unknown environment [27]. They also integrate the information gain criterion with a *greedy* component to bias the search toward a goal configuration.

Cheung and Lumelsky proposed a SBMP technique using a whole-sensitive arm manipulator moving in a three dimensional workspace [28]. In their proposed technique the information obtained in previous scans has little impact on what should be scanned next (i.e., their planner is memoryless). Kutulakos et. al. also proposed a similar algorithm which decomposes the path planning problem in 3 dimensions into two independent subproblems: a planar path planning problem and a three dimensional exploration problem [29]. This technique is applicable for robots with low DOF (i.e., $n \leq 3$) only.

Mehrandezh and Gupta utilized a variant of the probabilistic roadmap, i.e. Lazy PRM, for the SBMP and exploration

of an articulated robot equipped with a large number of proximity sensors mounted on its body [30]. They proposed a novel metric based on the notation of the C-zone map for robot's movement which results in efficient exploration of the environment. This mapping represents the size of the set of configurations associated with each point in the physical space and is calculated off-line prior to planning. A path going through those regions of physical space that correspond to higher C-zone volume are given higher priority for the next scan.

In our proposed SBMP technique information obtained about the C-space topology through the FD paradigm is utilized to find the best next scanning strategy. A PRM-based method is employed to efficiently capture the connectivity of the robot's free C-space using information obtained about the topology of the robot's C-space through the FD paradigm.

III. PROBLEM FORMULATION

A. Model-Based Motion Planning (MBMP)

Pertaining to the basic robot motion planning problem (described in Section I) demanding extra features on the sought path such as its length (i.e., known as the shortest-path problem) will increase the complexity of the problem especially in cases where the shortest path has to pass through narrow passages. The natural space where the motion planning can be carried out would be the robot's C-space where the robot can be represented as an automaton (i.e., a point robot). A hybrid motion planning approach based on HF and PRM is proposed for robotic navigation in both low and high dimensional spaces. In this approach a fluid dynamic paradigm based on harmonic functions is employed to obtain information about the robot's C-space topology. These information are then exploited to bias the distribution of random nodes in a PRM that captures the connectivity of the free C-space. The following sub-problems should be investigated in terms of MBMP of robots.

1) **Potential Field Approach Using FD Paradigm:** A fluid dynamic paradigm is utilized to generate permissible paths (i.e. streamlines) in the robot's C-space connecting the start and the goal configurations [31], [32]. These configurations are considered as the virtual flow source and sink. Obstacles are also considered as impenetrable objects. In particular, the notion of potential flow, an inviscid, incompressible, and irrotational fluid flow, is employed for computing the streamlines. We will describe how this analogy would help us to find the narrow regions (i.e., regions through which are hard to navigate the robot) in an environment cluttered with obstacles, later on.

A potential flow can be mathematically represented by:

$$\begin{cases} \nabla \cdot \vec{v} = 0 & \text{zero compressibility,} \\ \nabla \times \vec{v} = 0 & \text{zero vorticity,} \end{cases}$$

where \vec{v} is the velocity vector of the flow [33]. Zero vorticity implies that the velocity of the potential flow can be stated as the gradient of a scalar potential ϕ :

$$\vec{v} = -\nabla\phi, \quad (1)$$

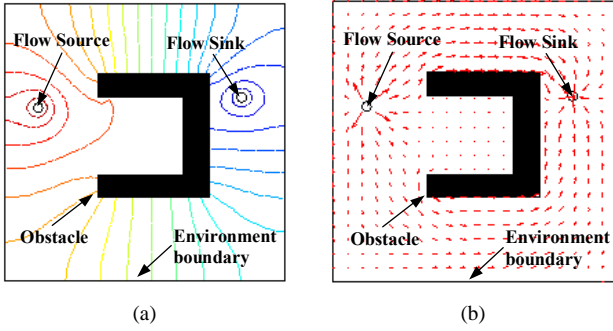


Fig. 1. (a): Potential contours and (b): Velocity vectors pertaining to a potential flow in a sample simulation.

Zero incompressibility of the potential flow combined with 1 yields the Laplace's equation:

$$\nabla^2 \phi = 0, \quad (2)$$

In an n -dimensional space the Laplace's equation can be represented by:

$$\sum_{i=1}^n \frac{\partial^2 \phi}{\partial x_i^2} = 0, \quad (3)$$

Solutions to 3 are in the form of harmonic Functions.

Harmonic Functions (HF) can be used to generate collision-free streamlines without local minima; however, saddle points (aka stagnation points) can occur. Due to Bernoulli's principle for inviscid incompressible fluid flow, higher velocities will appear in narrow regions given that the environment is bounded. Bernoulli's principle is mathematically represented as follow:

$$P + \frac{1}{2} \rho v^2 + \rho gh = \text{constant}, \quad (4)$$

where P denotes the static pressure (pressure induced due to random motions of molecules within the flow), ρ denotes the density, which is constant due to incompressibility of potential flow, v denotes the velocity of the flow, and h denotes the height of fluid particles with respect to a fixed (i.e., zero-potential) surface. Assuming constant potential energy, i.e. ρgh is constant, the Bernoulli's equation then reduces to a simple relation between the velocity and the static pressure of the flow which states that: as the speed of a moving fluid increases the static pressure within the fluid decreases and vice versa. Hence, due to lower static pressure of flow in narrow regions, higher velocities are observed. One should note that no vortex will be generated due to the assumption that the flow is inviscid.

Fig.1(a) and Fig.1(b) show the potential contours (i.e., isopotential lines) and velocity vectors pertaining to a potential flow in a sample simulation, respectively. The magnitudes of vectors shown in Fig.1(b) are proportional to the velocity of the fluid particles. Higher velocity is observed in the narrow passage between the c-shape obstacle and the environment boundary.

Harmonic functions constitute a powerful mean to generate local minima free smooth paths. However, possible deviation from one streamline to another when navigating a robot may cause chattering in motion [2]. Moreover searching through the

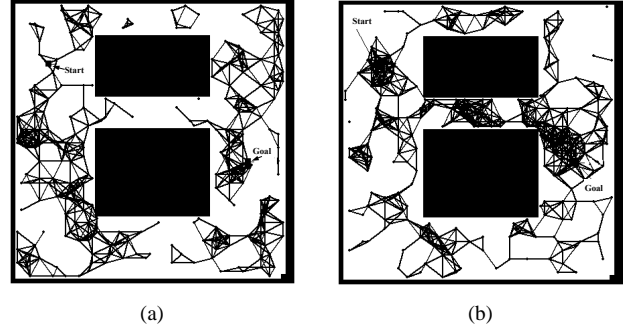


Fig. 2. (a) and (b): Generated roadmaps using original PRM method and HFPRM technique, respectively.

calculated streamlines for finding the shortest path connecting the start and goal configurations will be computationally cumbersome. These drawbacks necessitate to use a rudimentary motion planning technique to overcome these shortcomings. A variation of the PRM method has been chosen for this purpose. This will be explained with more details in the next section.

2) **PRM Construction Biased Toward Narrow Regions:** It will be shown that a rudimentary motion planning technique based on PRM can generate a better coverage of the C-space topology with the help of the information available to it through HFs. This will empower the PRM-based approach in finding the shortest path through narrow passages. It will be also shown that our proposed hybrid approach would perform favorably over the original PRM in finding shortest paths through narrow regions.

Fig.2(a) and Fig.2(b) show the roadmaps obtained for a simple, yet general scenario, using the original PRM method and our proposed hybrid technique, respectively. As can be seen the roadmap shown in Fig.2(b) more efficiently captures the topology of the confined space between two rectangular-shape black obstacles.

In our proposed approach the random sampling scheme utilized for constructing the roadmap is modified based on information obtained on the flow velocity through HFs to bias the distribution of random nodes towards critical regions (i.e., regions inducing high velocity). This guarantees a better coverage of the free C-space (C_{free}) than that using the original PRM.

We describe two different schemes for constructing the PRM:

- 1) **One shot PRM construction:** this technique is suitable for MBMP of robots with low DOF where the explicit representation of the robot's C-space is available or can be easily computed. In this method the PRM is constructed in one shot based on the information obtained using HFs. This scheme is utilized in the HFPRM technique and will be explained in more details in section IV-A.2.
- 2) **Incremental construction of the PRM:** this method is most suitable for MP of robots (model-based and/or sensor-based) with high DOF where the explicit computation of the robot's C-space is not practical. Using this scheme the PRM is incrementally constructed and

interactively updated based on the information obtained through HFs at each iteration. We utilize this scheme in our IHFPRM technique to capture the connectivity of the free C-space. Section IV-B.3 explains the proposed incremental construction of the PRM in more detail.

Extension of this hybrid MP technique to SBMP will be described in the next section.

B. Sensor-Based Motion Planning (SBMP)

The main idea of the proposed sensor-based planning approach is to utilize a fluid dynamic paradigm based on potential flows (as explained in section III-A.1) to identify and prioritize critical regions, i.e. narrow passages and hard-to-navigate regions. The proposed PRM that captures the connectivity of the free space is incrementally expanded as the robot senses the physical environment.

SBMP problem can be decomposed into two main sub-problems: (1) *View Planning* that determines which region should be scanned next and vantage point search that determines from where the scan should be carried out, and (2) *C-space Update/Expansion* that updates/expands the C_{free} as the robot senses the physical environment. In the following sub-sections we explain these sub-problems in more details.

1) View Planning and Vantage Point Search: The view planning and vantage point search problem comprises two steps: (1) identifying the best region that should be scanned next, and (2) finding the best permissible robot's configuration from which the region found in step (1) can be scanned. Pertaining to the first step different criteria have been utilized (e.g., [25], [27], [30]). In our proposed planning framework we address this problem from a FD analogy perspective. As it was explained in section III-A.1, the velocity profile obtained through the FD paradigm based on HFs can be utilized to obtain information about the topology of the environment. These information can be further exploited to prioritize regions of the environment to be scanned next. The main idea of our SBMP approach is to scan the regions associated with higher flow velocity first. Having these regions selected, the best view configuration (i.e., vantage point) among all permissible and reachable free configurations of the robot is then selected in the second step. This procedure is repeated till a collision-free path connecting the start and goal configuration is found. The view planning and vantage point search will be described in more details in Section V-D.

2) C-space Update/Expansion: The C-space update / expansion problem refers to updating and expanding the free C-space (C_{free}) as the robot senses the physical environment. A variation of the PRM approach that outperforms the original PRM is utilized for better coverage of the connected components of the robot's C-space at each iteration. A novel technique based on Fast Fourier Transform (FFT) and convolution theorem has been utilized to avoid the degeneration, i.e. adding new random nodes in the same region over and over when constructing a roadmap. This technique will be described in more details in Section V-C.

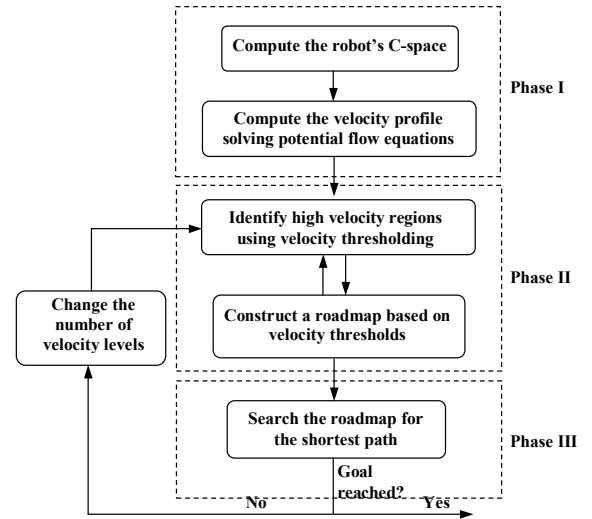


Fig. 3. The HFPRM solution methodology flowchart.

IV. MODEL-BASED MOTION PLANNING: SOLUTION METHODOLOGY

In this section the proposed solution methodology to implement the HFPRM and the incremental version of the HFPRM technique for model-based robot motion planning will be explained.

A. The HFPRM Technique: Solution Methodology

Fig.3 illustrates the overall solution methodology pertaining to the HFPRM technique.

The proposed HFPRM technique comprises three phases: in phase one the Laplace's equation, pertinent to potential flow, is solved in the computed robot's C-space. In phase two a probabilistic roadmap with a novel sampling scheme is constructed based on information obtained about the environment topology through the HF technique developed in phase one. The roadmap is then searched for the shortest path in phase three. More details are given in the following sub-sections.

1) Using FD Paradigm to Identify Narrow Regions: As it was explained earlier in Section III-A.1, a fluid dynamic analogy based on potential flows is used to solve the motion planning problem. Considering the properties of the potential flow, the potential field must satisfy the Laplace's equation as in (2). Gradient of this potential field constitutes the velocity field of the flow. Solution to the Laplace's equation is categorized as harmonic functions.

Numerical solution of the Laplace's equation can be obtained using Finite Difference Method (FDM). FDM imposes a mesh of uniformly distributed grid points in a region. The Laplace's equation at each grid point is then approximated using finite difference equations. Using the *Taylor series* expansion the second order central difference approximation of the velocity potential ϕ , in a 2D Cartesian space, can be represented as follows:

$$\frac{\partial^2 \phi(x, y)}{\partial x^2} = \frac{\phi(x + h, y) - 2\phi(x, y) + \phi(x - h, y)}{h^2} + O(h^2), \quad (5)$$

$$\frac{\partial^2 \phi(x, y)}{\partial y^2} = \frac{\phi(x, y + h) - 2\phi(x, y) + \phi(x, y - h)}{h^2} + O(h^2), \quad (6)$$

where h is the grid step size¹ and $O(h^2)$ is the truncation error which is of order h^2 . Assuming unit step size h as $h = 1$, and neglecting the truncation errors the discrete form of Laplace's equation (known as Laplacian difference equation) in a two dimensional Cartesian space can be rewritten as follows:

$$\phi(x+1, y) + \phi(x-1, y) + \phi(x, y+1) + \phi(x, y-1) - 4\phi(x, y) \simeq 0 \quad (7)$$

This equation can be easily extended to higher dimensions as well.

Expressing the actual physical phenomenon, the Laplace's equation has to be complemented with boundary conditions. These conditions are normally transposed into the conditions on the potential field ϕ . The boundary conditions (BC) have to be satisfied on the obstacle's boundaries, on the boundary of the environment, and at the source and the sink of the flow.

There are two main boundary conditions, namely *Dirichlet* and *Neumann*. The Dirichlet boundary condition can be mathematically stated as $(\phi|_{\partial\Omega} = c)$, where $\partial\Omega$ is the boundary of the domain (or region) of interest, Ω , and c is a constant. In this case the boundary is held at a constant potential and the flow must be along the outward normal of the obstacle surfaces [2]. The Neumann boundary condition can be mathematically described as $(\frac{\partial\phi}{\partial\vec{n}}|_{\partial\Omega} = c)$, where \vec{n} is the vector normal to the boundary of the domain defined as $\partial\Omega$. The Neumann boundary condition constrains the gradient of ϕ to accept a constant value on the boundary. In our case the flow must move tangential to the obstacle's boundaries, thus c must be set to zero. In some cases the Mixed or Robin boundary condition is employed as a linear combination of Dirichlet and Neumann boundary conditions. This can be mathematically represented as $(\alpha\phi|_{\partial\Omega} + \beta\frac{\partial\phi}{\partial\vec{n}}|_{\partial\Omega} = c)$, where α, β and c are constants [2].

In the proposed fluid dynamic paradigm the source and the sink are fixed at high (i.e. $\phi = 1$) and low (i.e. $\phi = 0$) constant potential values, i.e. Dirichlet BC. The boundary of the environment and the obstacles inside the environment are considered impenetrable, i.e. Neumann BC, $\frac{\partial\phi}{\partial\vec{n}}|_{\partial\Omega} = 0$. For all points in the environment other than those on the environment and obstacles boundaries, source and the sink a randomly selected potential value ϕ where $\phi \in [0, 1]$ is assumed as the initial guess. These values are interactively updated till a solution is converged.

The potential field equations (7), the boundary conditions and the initial values of the potential field give a complete description of the potential flow which can be mathematically represented in the form of a system of linear equations. Iterative schemes are normally utilized to solve these systems. The iterative methods employ an initial guess and generate a sequence of approximations to the solution. The iteration continues until the accumulated change in the potential values assigned to grid points is within a threshold from one iteration to the next (a variety of other stopping criteria can be utilized as well). *Jacobi, Gauss-Seidel, Successive Overrelaxation point*

¹For simplicity here it is assumed that the grid size in x and y directions are the same.

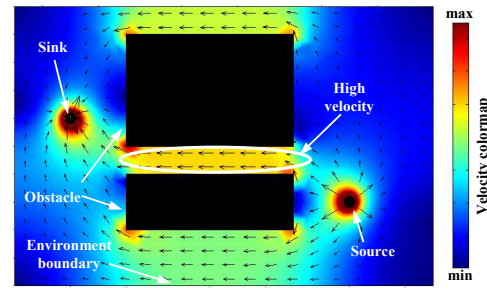


Fig. 4. Velocity profile in narrow region with corresponding velocity colormap

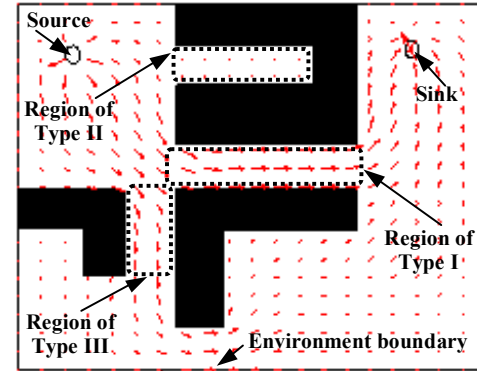


Fig. 5. Narrow regions categories; Type I: high velocity narrow regions, Type II: narrow regions with dead-ends, Type III: low velocity narrow regions

iterations, and *Tri-Diagonal Matrix Algorithm (TDMA)* are most commonly used methods for solving systems of linear equations interactively. The TDMA with a better convergence rate than others was employed in the computer simulations.

The TDMA is a technique for rapidly solving tridiagonal systems of linear equations [34]. TDMA utilizes *forward elimination* and *backward substitution* schemes to calculate the exact solutions of the system with a finite number of algebraic operations. TDMA can be extended to higher dimensions by applying it interactively in a line-by-line fashion in each dimension [34].

Once the potential values have been computed the velocity of flow can be calculated using the gradient of ϕ at each grid point. The velocity profile will be helpful to extract some information regarding the environment topology, i.e. identifying narrow regions. As it was explained in section III-A.1, higher velocities are observed in narrow passages due to Bernoulli law. The concept developed in our proposed method can be better understood through the computer simulation shown in Fig.4. Arrows in Fig.4 represent the velocity of the flow. The magnitude of the arrows is proportional to the value of velocity at each point. Higher velocities can be noticed in the narrow region between two rectangular shape obstacles. It should be noted that the narrow passages in the environment can be divided into different categories. Three different narrow passages may exist in a typical environment (see Fig.5).

In type I region high velocity exists meaning that the possibility of the shortest path going through this region is very high. In the type II region the velocity would be at the

minimal level leading to the fact that these narrow passages are dead-ends. One should note that due to inviscid property of the potential flow no vortex will be generated in these dead end regions therefore low velocities will be observed. In type III regions velocity values are low meaning that flow has a low tendency to pass through them. This information can be used in the next phase of the proposed motion planning technique, where utilizing a PRM-based approach to achieve a more uniform representation of the free C-space (i.e. \mathcal{C}_{free}).

2) **Using PRM to Capture the Connectivity of the Free C-space:** In our proposed PRM-based approach the pre-processing phase of the original PRM is modified considering the velocity regions obtained in phase I by a velocity thresholding technique. In this technique, given the number of velocity levels, each grid point according to its velocity value falls into one of these levels. A point falls into level (i) if and only if its velocity is in the following interval: $(\frac{i-1}{L}V_{max} \dots \frac{i}{L}V_{max}]$, where L is the number of velocity levels and V_{max} is the maximum velocity in the velocity profile obtained using HF. Using this scheme the velocity regions are organized in a predefined number of levels.

In the proposed PRM-based approach more random nodes are generated in high velocity regions in order to achieve a better coverage of the free C-space especially in narrow passages. The number of random nodes selected in each velocity region is determined by a function whose value changes with velocity level and the region area as follows:

$$r(i) = \frac{s(i)}{S}N, \quad (8)$$

where

$$s(i) = (1 - \omega)n_i + \omega i, \quad (9)$$

$$S = \sum_{i=1}^L s(i), \quad (10)$$

and n_i denotes the number of free nodes in level i , ω is a normalized factor, i.e. $\omega \in [0, 1]$, and N is the total number of desired random nodes to be generated. A tradeoff between area and the velocity in each level is established by changing the weight factor, ω . The number of random nodes to be selected for each region with a certain velocity level is computed using (8). A metric, i.e. Euclidean distance between nodes, is then used for connecting neighboring nodes in the roadmap in the pre-processing phase. Start and goal configurations are then added to the roadmap in the query phase. Dijkstra algorithm has been utilized to find the shortest path connecting the start and the goal configurations.

One should note that in the HFPRM technique the PRM is constructed in one shot given that the robot's free C-space can be explicitly computed. A FFT-based technique for fast computation of the robot's C-space has been utilized for this purpose. More details on this method can be found in the next section.

3) **Simulation Results and Discussions:** A Fast Fourier Transform (FFT) based method was employed to compute the robot's C-space. In this method the C-space is represented as the convolution of the robot and the workspace bitmaps [35].

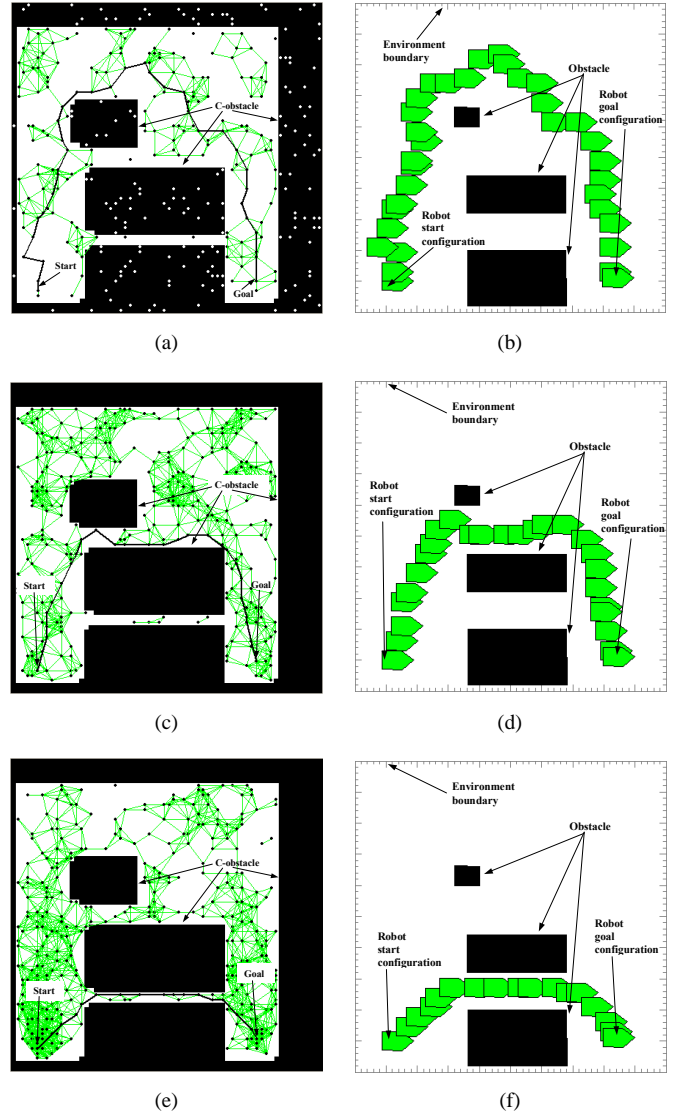


Fig. 6. Simulation results for motion planning of a 2-DOF mobile robot; (a): the roadmap obtained by original PRM, (b): robot's snap shots following the path obtained by original PRM, (c): the roadmap obtained by enhanced PRM, (d): robot's snap shots following the path obtained by enhanced PRM, (e): the roadmap obtained by HFPRM, (f): robot's snap shots following the path obtained by HFPRM.

This convolution can be computed efficiently using FFT. The running time of this method depends only on the resolution of discretization used to create the robot and the workspace bitmaps. For further details one can refer to [35].

The proposed HFPRM motion planning technique has been simulated for planar robots with up to 3 DOF for sake of illustration. 2D and 3D examples are presented in this section.

In the first example the motion planning problem for a 2-DOF mobile robot with only translational mobility has been carried out in an environment where some narrow passages exist. The shortest path must pass through one of them. Fig.6(a)-6(f) show the simulation results obtained using three different motion planning technique, namely the original PRM, the enhanced PRM and the proposed HFPRM. The robot is shown

as a polygon in green color. The roadmaps constructed in C-space based on three motion planning techniques mentioned above are shown in Fig.6(a), Fig.6(c) and Fig.6(e). The number of random nodes generated by these three methods are kept at constant. C-obstacles are presented in black. The snap shots of robot's motion obtained by each technique are also illustrated in Fig.6(b), Fig.6(d), and Fig.6(f). In this example the shortest path connecting the start and goal configurations must pass through the narrow passage between two big rectangular-shape obstacles.

The thick solid line in Fig.6(a) connecting the start and goal configurations shows the generated path obtained through the original PRM. As it is shown narrow passages have not been well covered by the roadmap. Thus, the original PRM fails to find the shortest path. Fig.6(c) and Fig.6(d) show the results obtained by enhanced PRM. In enhanced PRM method the random nodes are generated only within free C-space. As it is shown in Fig.6(c) the constructed roadmap presents a better coverage of the \mathcal{C}_{free} . However, narrow passages have not been well covered by the constructed roadmap. Thus the enhanced PRM fails to find the shortest path either. However one should note that the path found by the enhanced PRM has an overall shorter length than that when using the original PRM. The simulation results of the proposed HFPRM are depicted in Fig.6(e) and Fig.6(f). As it can be seen from Fig.6(e) the roadmap constructed based on HFPRM yields a good coverage² of the free C-space especially in hard-to-navigate regions. This method could find the shortest path going through the narrow regions between two obstacles efficiently.

Fig.7 shows the simulation results for motion planning of a 2-DOF mobile robot with the same mobility as explained in the first example. In this example any collision-free path connecting start and goal configurations has to go through a narrow passage. Fig.7(a) depicts the generated roadmap in the C-space using the proposed HFPRM. As it is shown the roadmap reflects the connectivity of the free C-space very efficiently. The snap shots of robot's motion moving from the start configuration towards the goal configuration are shown in Fig.7(b). The original and the enhanced PRM methods have difficulty to capture the connectivity of the free C-space. Fig.7(c) and Fig.7(d) show the constructed roadmaps based on enhanced PRM and the original PRM, respectively. Both methods fail to find a path as it can be seen from these figures.

Fig.8(a)-8(c) show the simulation results for motion planning of a 3-DOF mobile robot obtained using HFPRM technique. This example has been chosen to show the capability of the proposed HFPRM motion planning technique to find a path through hard-to-navigate regions in 3D. As it is shown in Fig.8(a) the robot has to pass through a confined space to reach the goal. The proposed HFPRM technique provides a good coverage of this narrow region. Thus it succeeds to find a path from the start to the goal configuration passing through this confined space. Part of this collision-free path in robot's C-space is depicted in Fig.8(b). The robot's 3D C-

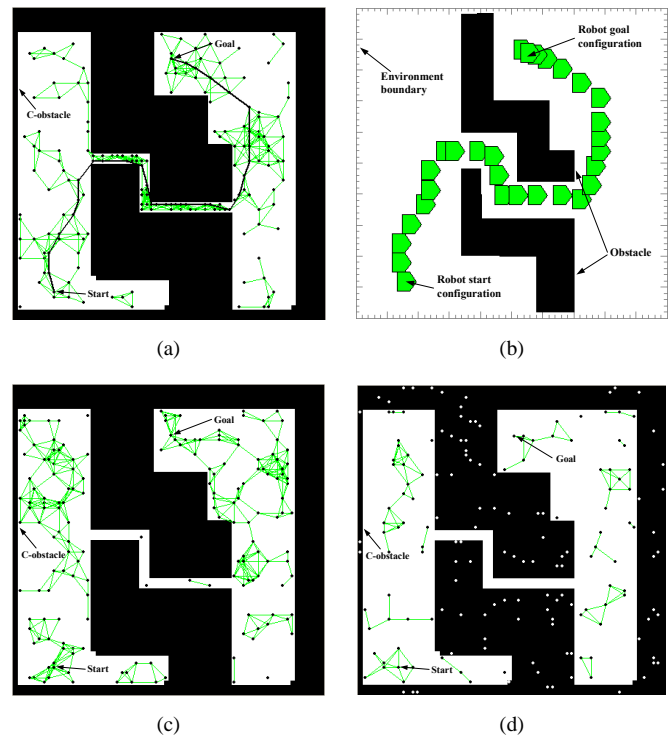


Fig. 7. Simulation results for motion planning of a 2-DOF mobile robot; (a): constructed roadmap based on HFPRM, (b): robot's snap shots following the path obtained by HFPRM, (c): the roadmap obtained by modified PRM, (d): the roadmap obtained by original PRM.

space is also shown in Fig.8(c). One should note that the C-obstacle corresponding to the boundary of the C-space is not illustrated. Otherwise the C-obstacle corresponding to the obstacle within the robot's workspace could not have been shown. The original and the enhanced PRM do not provide a good coverage of the C-space especially in the confined space between two obstacles. Thus they fail in connecting the start node to the roadmap in the query phase.

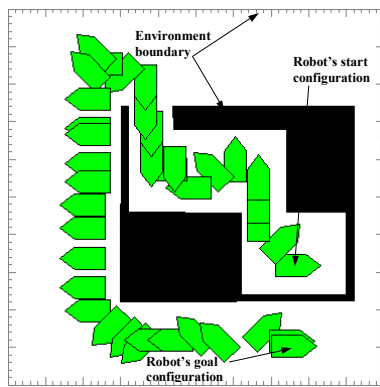
In the next section we present the solution methodology pertaining to the incremental version of the HFPRM technique, i.e. IHFPRM technique, for model-based motion planning of robots with high DOF.

B. Incremental HFPRM Technique: Solution Methodology

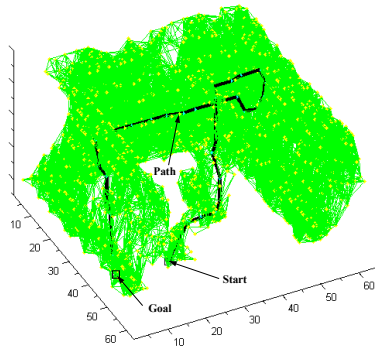
Although the FDM technique used to solve the Laplace's equation can be extended to higher dimensions (i.e., $n > 3$), but the explicit representation of the C-space of robots with many DOF will pose memory storage problems. As an alternative to a grid representation of the C-space one can use triangulated mesh generated by finite element mesh generators. This will drastically decrease memory needed to represent the robot's C-space. Therefore Finite Element Method (FEM) was utilized to solve HFPRM incrementally which will be described below. The overall IHFPRM solution methodology is depicted in Fig.9.

The IHFPRM technique works as follows: first a roadmap is initialized in robot's C-space by generating a pre-defined number of random nodes (the start and the goal configuration included) and connecting them together based on a metric,

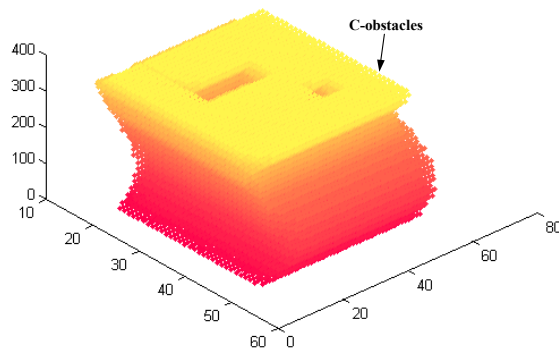
²One should note that good coverage means that a large portion of the narrow passages is covered by the roadmap.



(a)



(b)



(c)

Fig. 8. Simulation results for motion planning of a 3-DOF mobile robot obtained using HFPRM; (a): robot's snap shots in physical space, (b): constructed roadmap in C-space, (c): robot's 3D C-space.

i.e. maximum edge size represented as an Euclidean distance. Then in an iterative scheme the roadmap is enhanced and searched for the shortest path. In each iteration the Laplace's equation, pertinent to potential flow, is solved over the current representation of the C-space using Finite Element Method (FEM) and the velocity of the flow is calculated at each node of the roadmap. The roadmap is then enhanced using a new, yet simple, random sampling scheme based on information obtained about the C-space topology with a *velocity thresholding* technique. This sampling scheme yields a roadmap with better coverage in regions associated with higher velocities (i.e. narrow regions). At the final stage the roadmap is searched by a local planner for the shortest path connecting the start and

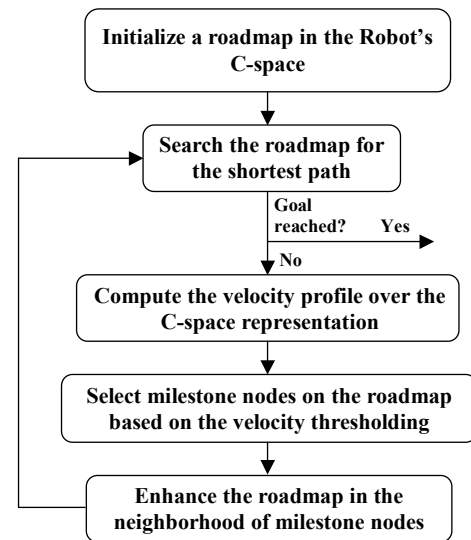


Fig. 9. IHFPRM solution methodology.

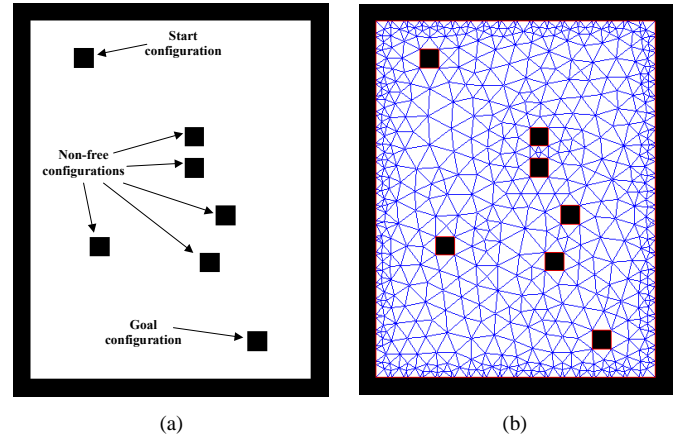


Fig. 10. (a): example of grid-based representation of a 2D C-space, (b): the corresponding mesh generated by a FEA mesh generator.

the goal configurations. The enhancement ends either after a maximum number of iterations has reached or a collision-free path has been found.

First we elaborate on the required C-space representation to compute the HFs using FEM, and then the above mentioned steps will be discussed in more details in the following subsections.

1) C-space Representation: In the proposed technique two different representations of the robot's C-space are maintained: The graph-based representation, i.e. roadmap, and the grid-based representation. The former captures the connectivity of the free C-space in form of a graph structure that is exploited for path planning. The latter captures the topology of the C-space obstacles (C-obstacle) in form of a composite geometry object and is utilized by Finite Element Analyzer (FEA) to compute the harmonic functions over the C-space.

In this representation the start and the goal configurations and each non-free configuration in the C-space are represented by rectangular grids. The corresponding boundary conditions pertinent to the potential flow are assigned to the boundaries of

these grids in the next step. At each iteration of the proposed technique a FEM mesh generator builds a triangulated mesh based on the updated C-space representation³. The generated mesh is then utilized by the FEA to solve the Laplace's equation. An example of this representation and the corresponding mesh for a 2D C-space are depicted in Fig.10. One should note that this representation can be mathematically extended to higher dimensions simply by extending the Dirichlet/Neumann condition to an n -dimensional space.

2) Using FD Paradigm to Identify Narrow Regions:

Similar to the original HFPRM technique a fluid dynamic analogy based on potential flow is used to obtain information about the environment topology (refer to Section IV-A.1).

Numerical solution of Laplace's equation can be obtained using Finite Element Method (FEM) [34]. The FEM discretizes the domain into simple geometry objects, i.e. elements, and the Laplace's equation is approximated by a piecewise linear function over each element.

In the proposed IHFPRM technique a triangulated mesh is generated based on the current grid-based representation of the C-space using a FEM mesh generator. The Laplace's equation, considering the boundary conditions as described earlier in Section IV-A.1, is then approximated on each mesh element by piecewise linear functions. This results in a system of linear equations whose solution yields the potential of the flow at each node of the triangulated mesh.

Having the potential values computed, the velocity of flow can be calculated using the gradient of ϕ at each node of the triangulated mesh. The velocity profile will be helpful to extract some information regarding the environment topology, i.e. identifying narrow regions. Higher velocities are observed in narrow passages due to Bernoulli's law [33] (refer to Section IV-A.1). The information obtained through the velocity profile are used to enhance the PRM at each iteration to achieve a more uniform representation of the free C-space.

3) Incremental Construction and Update of the PRM:

In our proposed IHFPRM technique a coarse roadmap with a small set of randomly generated nodes is initially built in the free C-space. The roadmap is then enhanced in an iterative scheme based on the information obtained about the C-space topology through the FD paradigm. One should note that in the basic PRM the roadmap is constructed in the pre-processing phase without considering the locations of the start and goal configurations. However, in our proposed PRM-like approach the topology of the generated roadmap depends on the locations of the start and goal configurations. This will be shown through computer simulation presented in Section IV-B.4.

Based on the velocity profile computed through the FD paradigm, a velocity thresholding technique is utilized to identify regions with different topologies. At each iteration of the IHFPRM technique, given the number of velocity levels, each node of the roadmap according to the velocity of the potential flow at that node falls into one of these levels⁴A

node falls into level (i) if and only if its velocity would be in the following interval: $(\frac{i-1}{L}V_{max} \dots \frac{i}{L}V_{max}]$, where L is the number of velocity levels and V_{max} is the maximum velocity in the velocity profile. Using this scheme the nodes of the roadmap are categorized in a predefined number of levels according to the corresponding velocity regions. One should note that there may be different number of nodes in each level corresponding to the area associated with different velocities.

The main idea of the proposed PRM approach is to update and enhance the roadmap in regions associated with higher velocities. Update of the PRM is carried out in an iterative scheme by selecting milestone nodes to enhance the existing PRM. The number of milestone nodes selected at each velocity level is determined based on the average velocity and the number of nodes (area) of the corresponding level. This can be mathematically expressed by:

$$r_i = \frac{s_i}{S}R, \quad (11)$$

where

$$s_i = (1 - \omega) \frac{n_i}{N} + \omega \frac{V_{avi}}{V_{max}}, \quad (12)$$

$$S = \sum_{i=1}^L s_i, \quad (13)$$

r_i : number of milestone nodes selected in level i ,

R : total number of milestone nodes,

ω : weight factor,

n_i : total number of nodes in level i ,

N : total number of nodes in the roadmap,

V_{avi} : average velocity in level i ,

V_{max} : maximum velocity in the roadmap.

A tradeoff between area and the velocity at each level is established by changing the weight factor, ω . (12) is the dynamic version of (9). The maximum velocity is calculated online in (12) while in (9) the maximum velocity is considered to be constant. The number of milestone nodes to be randomly selected at each threshold level is computed using (11). At each iteration, having the number of milestone nodes computed, the roadmap is enhanced in the following steps:

Step (1): For $i = 1, \dots, L$, at each level i , r_i nodes with the smallest number of adjacent nodes are selected as the milestone nodes. This results in a uniform distribution of nodes at each level i .

Step (2): A pre-defined number of randomly selected configurations is generated within the neighborhood of each milestone node.

Step (3): The new generated nodes are checked by a local planner for the collision. The free configurations are added to the roadmap based on a metric, i.e. Euclidean distance between nodes, and the non-free configurations are kept in a list of C-obstacle nodes which will be used to update the grid-based representation of the C-space in the next iteration, as explained in IV-B.1.

The updated roadmap is then searched for the shortest path. If no path is found the grid-based representation of the C-space is updated by adding the C-obstacle grids corresponding to the C-obstacle nodes found in the previous iteration and

³However, one can generate a coarse representation of meshes using rectangular elements with the same level of accuracy.

⁴The velocity of the flow at each node of the roadmap, which does not lie on any node of the mesh, is calculated using linear interpolation.

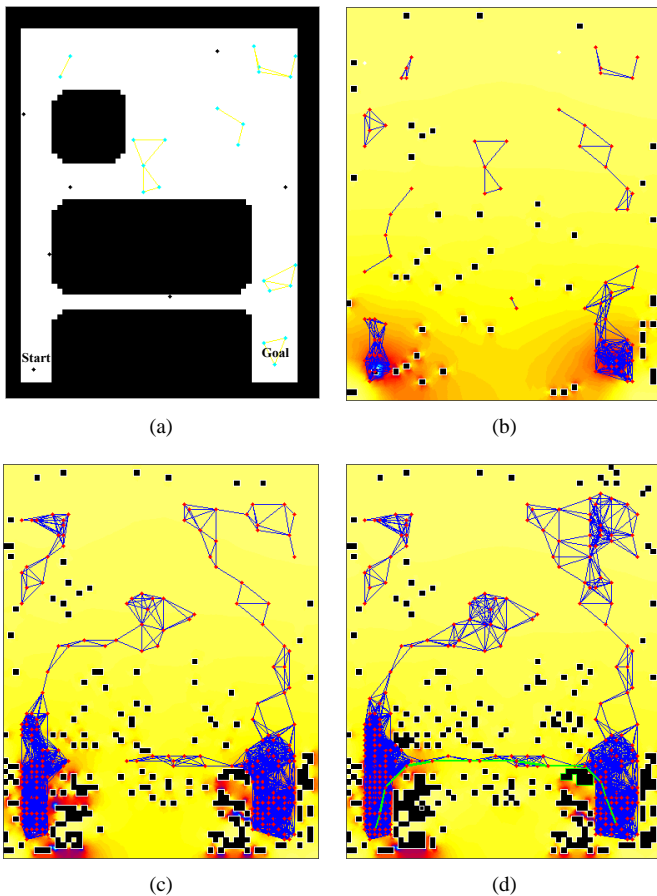


Fig. 11. (a): Explicit representation of the C-space and the initial roadmap, (b) and (c): the enhanced roadmap after the 2nd and the 5th iteration, (d): the shortest path found along with the roadmap after the 7th iteration.

HFs are calculated using FEM based on the updated C-space representation. The roadmap is enhanced based on the new velocity profile found at each iteration, and the procedure continues. The iteration continues until a path is found or a maximum number of enhancements has been reached. In the next section some of the simulation results obtained through the proposed IHFPRM technique are presented.

4) **Simulation Results and Discussions:** The proposed IHFPRM technique has been simulated for 2-DOF mobile robots with only translational mobility for sake of illustration. One should note that since no explicit representation of the C-space is required this technique can be extended to robots with many DOF. In this section we present 2D simulation results for some examples where critical regions, i.e. narrow regions and hard-to-navigate passages, exist. The effect of different parameters, which are set before the planning starts, are also investigated on overall performance of our proposed hybrid motion planning technique.

Fig.11(a)-Fig.11(d) show the simulation results obtained using the IHFPRM for an environment where the robot can pass through more than one passage to reach the goal configuration. However, the shortest path passes through a narrow region which is between two rectangular shape C-obstacles as depicted in Fig.11(a). The explicit representation of the C-space, for sake of illustration, is also shown. The enhanced

IT#	ROADMAP SEARCH (sec.)	FEM SOLVE (sec.)	ROADMAP ENHANCE (sec.)
1	0.00	4.01	0.50
2	0.01	5.81	0.88
3	0.02	9.84	1.61
4	0.04	11.73	3.77
5	0.06	14.45	6.01
6	0.12	20.44	8.09
7	0.37	-	-

TABLE I

RUN TIME STATISTICS PERTAINING TO THE EXAMPLE SHOWN IN FIG. 11.

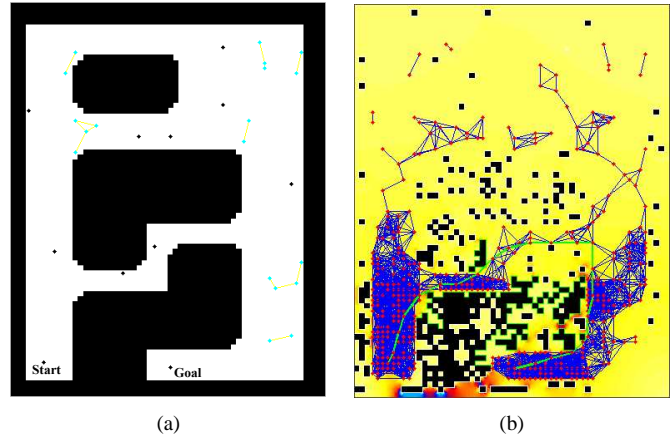


Fig. 12. (a): Explicit representation of the C-space and the initial roadmap, (b): the shortest path along with the roadmap after the 9th iteration.

roadmap after the second and the fifth iteration are shown in Fig.11(b) and Fig.11(c), respectively. The C-obstacle grids are presented in black, and the corresponding velocity profile is shown in color in the background. The shortest path has been found at the 7th iteration as depicted in Fig.11(d). As it is shown the obtained roadmap yields a good coverage in narrow region.

The CPU time associated with main functions of the IHFPRM technique, using a 2.4 GHz Intel Celeron, is summarized in Table I. Almost 75% of the total running time relates to the solving of the Laplace's equation using the FEM.

Fig.12 shows the simulation results for motion planning of a 2-DOF mobile robot with the same mobility as explained in the first example. In this example the shortest path connecting the start and the goal configurations has to go through a hard-to-navigate passage. The shortest path, highlighted in Fig.12(b), has been found through the hard-to-navigate region after 9 iterations.

Fig.13 shows another simulation results obtained in an environment where any collision-free path connecting the start and the goal configurations has to go through a narrow passage. The obtained path along with the enhanced roadmap after 4 iterations is shown in Fig.13(b). As it is depicted the obtained roadmap maintains a good coverage in the narrow passage as well as in the wide-open regions.

As mentioned earlier the performance of the proposed technique will depend on the location of the start and goal configurations. This has been shown through Fig.14. Fig.14(a)

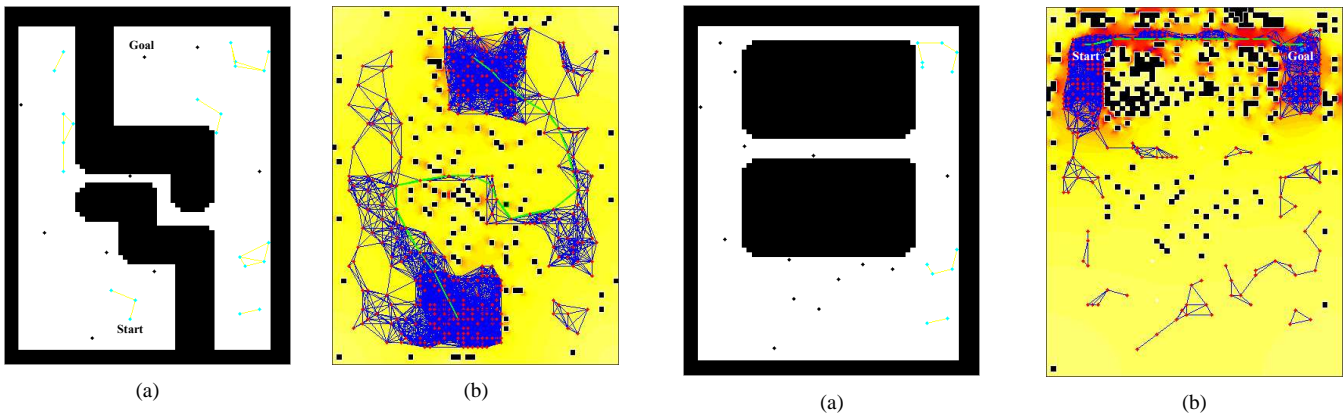


Fig. 13. (a): Explicit representation of the C-space and the initial roadmap,(b): the shortest path along with the roadmap after the 4th enhancement.

shows the C-space of the robot, which is shown only for sake of illustration (no explicit C-space computation has been carried out in the computer simulations). In Fig.14(b) through Fig.14(f) the positions of the start and goal configurations have been changed. As can be seen from these figures the enhancement of the roadmap will be always biased towards regions with higher velocity thresholds. In all cases the planner is capable of finding the shortest path going through narrow passages. In the next section we discuss the effect of some parameters on the performance of the proposed technique through simulation results.

5) Sensitivity Analysis on Motion Planning Parameters:

We examine the effect of three important parameters on the overall performance of the proposed technique in finding the shortest path: (1) L the number of velocity threshold level, (2) ω the velocity weight factor, and (3) the maximum edge size in the roadmap. We discuss the effects of these parameters on the performance of the IHFPRM technique for a 2-DOF mobile robot moving in the environment shown in Fig.14(a) with the position of the start and the goal configuration being set as in Fig.14(d). One should note that in each of the following examples the roadmap is generated from the same seed of random numbers.

1. Effect of the number of velocity threshold levels, L : The number of the velocity threshold levels, L plays an important role in finding paths going through narrow regions. The contribution of this parameter can be better understood through some representative simulation results given in Fig.15(a) and Fig.15(b). The simulation results are based on scenario given in Fig.14(d). The shortest path going through the narrow region between two rectangular-shape obstacles was found after 4 iterations with the total number of random nodes at 472 with $L = 5$ in Fig.15(a). Fig.15(b) shows the results for the case where $L = 15$. The shortest path is found after 4 iterations but with the total number of random nodes at 572.

2. Effect of the velocity weight factor, ω : The velocity weight factor is utilized to make a tradeoff between the average velocity of the flow at a specific level with the number of previously generated nodes in that level. The contribution of this factor on the performance of the proposed technique is

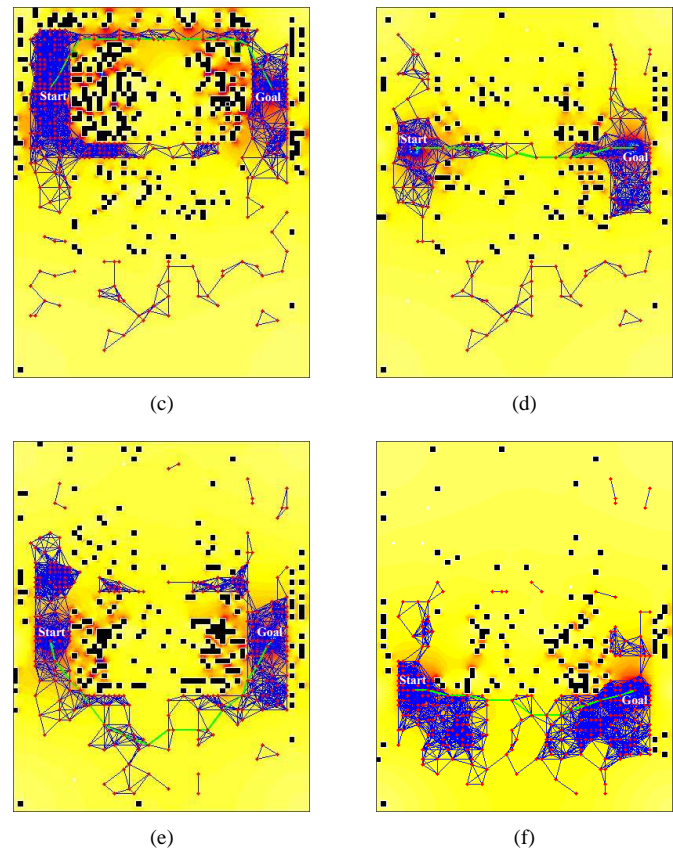


Fig. 14. (a): Explicit representation of the C-space,(b)-(f): the shortest path along with the roadmap pertaining to different positions of the start and goal configurations in the same environment.

illustrated through representative simulations. Again the simulation results are based on the scenario shown in Fig.14(d). Weight factors of $\omega = 0.2$ and $\omega = 0.5$ are chosen in two simulations whose results are presented in Fig.16(a) and Fig.16(b), respectively. As can be seen in Fig.16(a), more random nodes are generated in wide-open regions than that in Fig.16(b). In short, high velocity weight factor will cause generating less random nodes in wide-open regions far from the obstacles. Therefore, finding the shortest path with lower number of nodes can be achieved by choosing higher weight factor. One should also note that a part of the shortest path could go through wide-open regions. A fine coverage of these

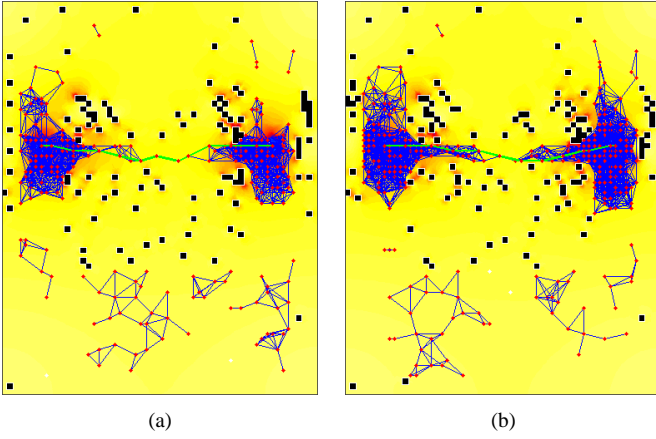


Fig. 15. The roadmap along with the corresponding shortest path with (a): 5 threshold levels, (b): 15 threshold levels.

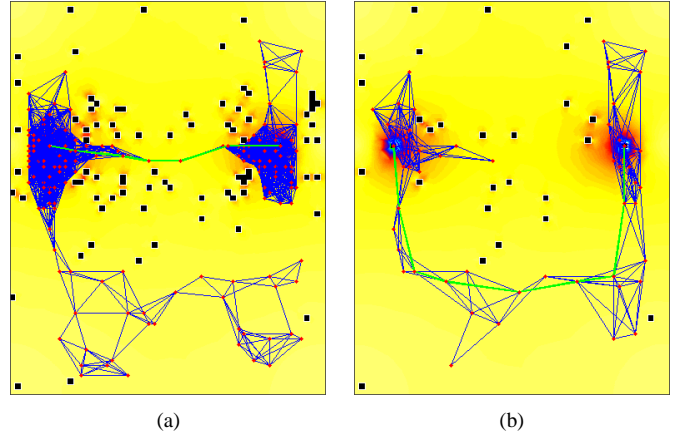


Fig. 17. The roadmap along with the corresponding shortest path with the maximum edge size of (a): 10 and (b): 15.

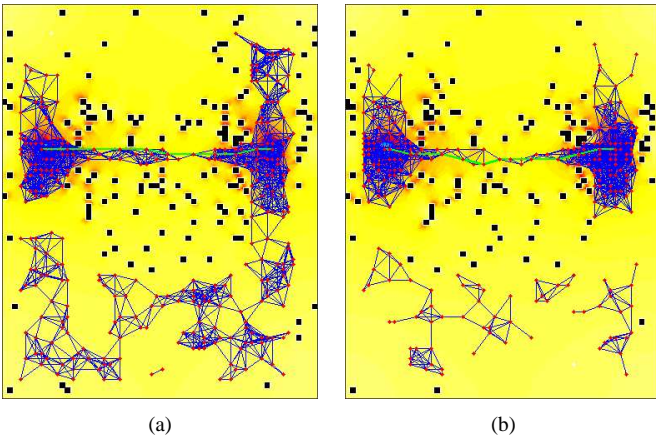


Fig. 16. The roadmap along with the corresponding shortest path with velocity weight factor equal to (a): 0.20 and (b): 0.50.

regions would be needed to make sure the shortest path found in the roadmap closely represents the actual shortest path.

3. Effect of the maximum edge size: Choosing bigger maximum edge size to connect the nodes in the roadmap results in less planning time and coarser roadmap which consequently yields a longer path. Fig.17 illustrates the effect of the edge size on the obtained roadmap. Fig.17(a) and Fig.17(b) show the obtained roadmap after 3 and 2 iterations with maximum edge size of 10 and 15, respectively, for the same scenario as shown in Fig.14(d). As it is shown with bigger maximum edge size the planner needs less enhancement to find the path, however, the obtained shortest path will be most probably very longer than the actual shortest path.

V. SENSOR-BASED MOTION PLANNING: SOLUTION METHODOLOGY

A. Definitions and Notations

Let \mathcal{A} denote a cylindrical-shape robot with 3 DOF (x - y translations and θ rotation) moving in a 2D workspace \mathcal{W} . Considering the circular cross-section of the robot, the robot's C-space \mathcal{C} can be introduced as a 2D space where each point in this space represents the x - y coordinates of the robot's center

of rotation represented in a fixed coordinate frame attached to the robot's workspace.

The robot is equipped with ultrasonic range finders. The \mathcal{W} is incrementally updated as the robot scans the physical workspace and correspondingly the \mathcal{C} representation will be also updated. The free part of the C-space $\Delta\mathcal{C}_{free}$, which is made known free after each scan can be computed comparing the updated \mathcal{C} representations obtained through the last two scans. A probabilistic roadmap \mathcal{R} is constructed within the $\Delta\mathcal{C}_{free}$ after each scan to capture the connectivity of the \mathcal{C}_{free} . The robot is only allowed to move within the \mathcal{C}_{free} represented by the \mathcal{R} .

B. Algorithm

The overall algorithm of the sensor-based motion planning framework is given as the following pseudo-code:

```

SB-IHFPRM
{
     $\mathcal{W} \leftarrow \text{Workspace.Initialize};$ 
     $[\mathcal{C}, \Delta\mathcal{C}_{free}, \mathcal{C}_{unbnd}] \leftarrow \text{Cspace.Initialize}(\mathcal{W});$ 
     $\mathcal{R} \leftarrow \text{Roadmap.Initialize}(\Delta\mathcal{C}_{free});$ 
     $q_{cur} \leftarrow q_s;$ 
     $it \leftarrow 0;$ 
    repeat
         $it \leftarrow it+1;$ 
         $path \leftarrow \text{Roadmap.Search}(q_{cur}, q_g, \mathcal{R});$ 
        if  $path \neq \text{NULL}$ 
             $q_{cur} \leftarrow \text{Robot.Move}(path);$ 
            Exit;
        end;
         $\mathcal{V} \leftarrow \text{Fem.Solve}(q_s, q_g, \mathcal{C});$ 
         $q_v \leftarrow \text{Planner.NextBestView}(\mathcal{V}, \mathcal{C}_{unbnd}, \mathcal{R}, \mathcal{W});$ 
         $path \leftarrow \text{Roadmap.Search}(q_{cur}, q_v, \mathcal{R});$ 
         $q_{cur} \leftarrow \text{Robot.Move}(path);$ 
         $\Delta\mathcal{W} \leftarrow \text{Robot.Scan};$ 
         $\mathcal{W} \leftarrow \text{Workspace.Update}(\Delta\mathcal{W}, \mathcal{W});$ 
         $[\mathcal{C}, \Delta\mathcal{C}_{free}, \mathcal{C}_{unbnd}] \leftarrow \text{Cspace.Update}(\mathcal{C}, \mathcal{W});$ 
         $\mathcal{R} \leftarrow \text{Roadmap.Expand}(\Delta\mathcal{C}_{free}, \mathcal{R});$ 
    until  $it \leq \text{MAXITNUM}$ 
}
    
```

The proposed algorithm works as follow: first the workspace and correspondingly the robot's C-space representations are initialized⁵. A probabilistic roadmap \mathcal{R} is then initialized in the small free region surrounding the robot's start configuration q_s . Then in an iterative "Scan and Go" fashion the roadmap is incrementally enhanced and expanded as the robot senses the workspace after each scan. At each iteration the roadmap is searched for the shortest path between the robot's current configuration q_{cur} and the goal configuration q_g . If such a path exists the robot moves to the goal configuration following the obtained path. Otherwise the robot plans for the next scan. The Laplace's equation, pertinent to the potential flow, is solved over the current representation of the C-space \mathcal{C} using a Finite Element Analyzer (FEA) package and subsequently the flow's velocity profile \mathcal{V} is computed. Planner then finds the next best view configuration q_v where the next scan should be carried out using the information obtained through this velocity profile. This is done through the `Planner.NextBestView()` class which will be explained in more details in Section V-D. The roadmap is then searched for the shortest path from q_{cur} to q_v and the robot moves to the configuration q_v to do the next scan. Based on the new information obtained about the workspace through a scan, i.e. $\Delta\mathcal{W}$, the workspace and correspondingly the C-space representations are updated and the new free part of the C-space $\Delta\mathcal{C}_{free}$ is calculated. This procedure is particularly further explained in Section V-C. The roadmap is enhanced and expanded within the $\Delta\mathcal{C}_{free}$ found in the previous scan. More details on roadmap expansion is given in Section V-C. The updated roadmap is then searched for the shortest path between q_{cur} and q_g . If a path is found the robot moves to the goal configuration and the planner exits, otherwise the robot plans for the next scan and repeats all the above steps. The iterations continue until either a maximum iteration number, i.e. MAXITNUM, is reached or a collision-free path is found. The main parts of the proposed algorithm is explained in more details in the following sub-sections.

C. Workspace and C-space Representations

We utilize a Fast Fourier Transform (FFT)-based method using convolution theorem to compute the $\Delta\mathcal{C}_{free}$ explicitly⁶. Using this technique the robot's C-space can be represented as the convolution of the robot and workspace representations. In this section we describe this technique in details.

Consider a physical workspace $\mathcal{W} = [a, b] \times [c, d] \subset \mathbb{R}^2$. \mathcal{W} can be discretized into an $N \times M$ array W . Each grid w_{ij} of \mathcal{W} is defined as:

$$w_{ij} = \left[a + i \frac{(b-a)}{N}, a + (i+1) \frac{(b-a)}{N} \right] \times \left[c + j \frac{(d-c)}{M}, c + (j+1) \frac{(d-c)}{M} \right], \quad (14)$$

⁵Initially the C-space includes a free region surrounding the robot's start configuration, and the boundary set for the C-space to be considered as C-obstacles. The remaining part of the C-space is assumed to be unknown.

⁶In general there is no need for explicit representation of the $\Delta\mathcal{C}_{free}$ using the PRM. However, explicit representation of $\Delta\mathcal{C}_{free}$ can drastically decrease the CPU-time needed for constructing the roadmap. One should note that $\Delta\mathcal{C}_{free}$ can be computed explicitly only for robots with low DOF (i.e., $n \leq 3$).

where $i = 1, \dots, N$, and $j = 1, \dots, M$. Initially we assume that there is no obstacle in the workspace and all the cells of array W are assigned 0 (i.e., free) other than the boundaries of the workspace which are assigned 1 (i.e., obstacle). Array W is incrementally updated as the robot senses the physical space. If an obstacle is found anywhere in a grid w_{ij} of the workspace we let $W(i, j) = 1$, otherwise 0. Considering that the robot has a circular cross-section, the robot's C-space, \mathcal{C} , can be represented as a set of (x, y) coordinates of the robot's center of rotation. For a planar robot in a 2D workspace the C-space bitmap array C can be explicitly obtained as the convolution of the workspace bitmap array W and the mirror of the robot's bitmap array, A' [35]:

$$C = W * A', \quad (15)$$

where $A'(i, j) = A(-i, -j)$ ⁷. The above convolution can be computed using Fast Fourier Transform (FFT). The running time of this calculation depends only on the resolution of the discretization. For more details on this one can refer to [35].

We also maintain a bitmap array W_{unk} which represents the unknown part of the workspace. Each cell of W_{unk} is set to 1 if the corresponding grid in the workspace is unknown, otherwise 0. Initially we assume that all the cells in the workspace other than the small region around the robot at the start configuration are unknown and the corresponding cells in W_{unk} are set to 1. W_{unk} is incrementally updated as the robot scans the workspace and makes some part of it known. Correspondingly C_{unk} represents the bitmap array pertinent to the unknown C-space \mathcal{C}_{unk} , and is computed using the convolution method explained earlier as follow:

$$C_{unk} = W_{unk} * A'. \quad (16)$$

The set of configurations which constitute the boundary of the unknown C-space, i.e. \mathcal{C}_{unkbnd} the cross-section of the free C-space \mathcal{C}_{free} and the unknown C-space \mathcal{C}_{unk} , can be computed using this technique interactively. In the view planning stage of the proposed technique the next region which should be scanned next is selected based on information obtained through the FD paradigm. This is explained in more details in Section V-D.

The bitmap array C_{free} corresponding to the free C-space at each iteration is computed as follow:

$$C_{free}(i, j) = \begin{cases} 1 & \text{if } C_{unk}(i, j) = 0 \ \& \ C(i, j) = 0 \\ 0 & \text{otherwise,} \end{cases}$$

where $i = 1, \dots, N$, $j = 1, \dots, M$, and each cell of the C_{free} which is assigned 1 shows that the corresponding configuration is free. The C_{free} is utilized to calculate the new free part of the C-space $\Delta\mathcal{C}_{free}$ which is made known after each scan to expand the roadmap. $\Delta\mathcal{C}_{free}$ can be calculated as: $\Delta\mathcal{C}_{free} = C_{free}^i - C_{free}^{i-1}$, where C_{free}^n denotes the C_{free} calculated at the n^{th} iteration. Having this computed, the roadmap is then expanded within this region by generating a pre-defined number of random nodes and connecting them together based on a predefined maximum edge size of the roadmap. The explicit computation of the $\Delta\mathcal{C}_{free}$ prevents the roadmap to be degenerated in the same region.

⁷ A denotes the robot's bitmap array.

D. View Planning Using FD paradigm

The view planning problem is solved in two steps: in the first step a region of the workspace which should be scanned next is selected based on information obtained through the FD paradigm, and in the second step a configuration that can be reached by the robot is selected to scan the region found in the first step from. In the following sub-sections first we explain how FD paradigm can be used to identify critical regions in the environment, and then we will describe our proposed technique to address the steps mentioned above.

1) Using FD Paradigm to Identify Critical Regions:

The same fluid dynamic analogy explained in Section IV-B.2 is used to solve the view planning problem. In the proposed SBMP technique FEM is utilized to solve the Laplace's equation over the grid-based representation of the C-space. A triangulated mesh is generated based on the current grid-based representation of the C-space using a FEM mesh generator. The Laplace's equation is then approximated on each mesh element by piecewise linear functions. This results in a system of linear equations whose solution yields the potential of the flow at each node of the triangulated mesh.

After computing the potential values the velocity of flow can be calculated directly using the gradient of ϕ at each node of the triangulated mesh. This velocity profile is used to extract some information regarding the environment topology, i.e. identifying narrow regions. Higher velocities are observed in narrow passages due to Bernoulli's law [33]. For a complete discussion on this one can refer to our previous paper [36]. The information obtained through the velocity profile is used to prioritize different regions of the environment when it comes to the scan planning stage of our proposed sensor-based motion planning technique, i.e. *Which Region to Scan*. This will be discussed in further details in THE NEXT section.

2) **Which Region to Scan:** The crux of our planning approach is to expand the roadmap into the regions of the C-space associated with higher velocities pertinent to the potential flow. Therefore, in our view planning approach higher priority is given to the high velocity regions (configurations). The boundary of the unknown C-space, C_{unkbnd} is calculated and updated using the convolution method explained in Section V-C. The velocity of the flow is calculated at each configuration of the C_{unkbnd} using a linear interpolation. The configuration q_x with highest velocity among all the configurations in C_{unkbnd} is selected as the central configuration. The region represented by $\mathcal{A}(q_x)$ (i.e., the region occupied by the robot when being at configuration q_x) is scanned next. At this stage it remains to find the best feasible configuration on the roadmap that can be reached by the robot to do the next scan from, hence the name i.e. *From Where to Scan*.

3) **From Where to Scan:** To find the best view configuration to scan the region represented by $\mathcal{A}(q_x)$ obtained in the previous step, all the nodes (configurations) of the connected component of the existing roadmap containing the start configuration q_s are examined and the best view configuration, q_v is selected. The q_v is chosen in a way that if the robot is placed at that configuration the distance between the center of the sensor's field of view \mathcal{V}_s and the center of workspace grid corresponding to the center of $\mathcal{A}(q_x)$ would be minimum

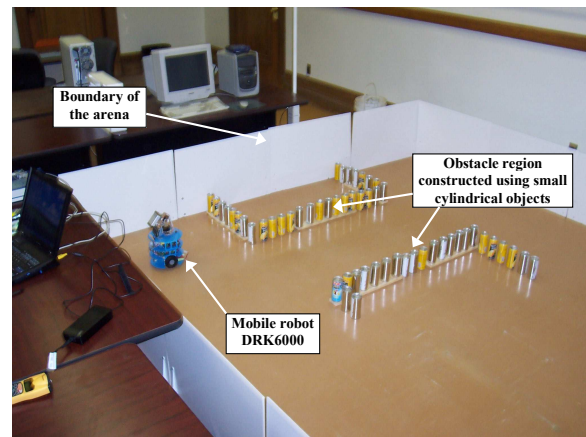


Fig. 18. The robot's workspace (a rectangular-shape arena).

without having any known obstacle obstructing the sensor's view.

VI. IMPLEMENTATION AND EXPERIMENTAL RESULTS

The proposed motion planning technique was implemented on a wireless mobile robot, called DRK6000, manufactured by Dr.Robot Inc. [37]. The robot is equipped with an array of ultrasonic range finders which retrieve range data from the physical space. The sensors' field of view is 30 degrees with a range limited to 100cm. The robot's motion is limited within a 375cm \times 300cm rectangular-shape arena (see Fig.18). Small cylindrical objects are used to construct different shapes of obstacles in this workspace. The best size for these cylindrical shape objects were selected after conducting a series of tests. The Aluminum cans with 5cm of diameter showed the least sonar shadow. Thus they were selected in experiments.

As it was explained earlier in Section V-C a grid is used to represent the robot's workspace. In our experiments we discretize the robot's workspace into a 60 \times 75 grid. This workspace is simulated in MATLAB. Furthermore the simulation updates the grid-based representation of the workspace, the robot's C-space, the roadmap, and the robot's location as it moves. The workspace and correspondingly the robot's C-space representations are updated as the robot scans the environment. Fig.19 shows the simulator initialized for a sample experimental setup. What is represented through this simulator is snapshots of the real robot moving in real environment. More explanation on the simulator will follow.

A large number of experiments have been carried out. However, only two representative experiment runs are presented in this section. Fig.20 shows the workspace and robot's C-space for 14 scans (i.e., iterations). The internal model of the robot's workspace and the path segments that it follows at each iteration are shown in the first column of Fig.20. The second column shows the corresponding robot's C-space along with the roadmap at each iteration. Fig.20(a) shows the initialized workspace along with the robot (represented as a green circle) at its initial configuration. A small region (a 65cm \times 65cm rectangular area) around the robot, shown in white color, is initially assumed to be free to accommodate the first movements of the robot. The grids corresponding

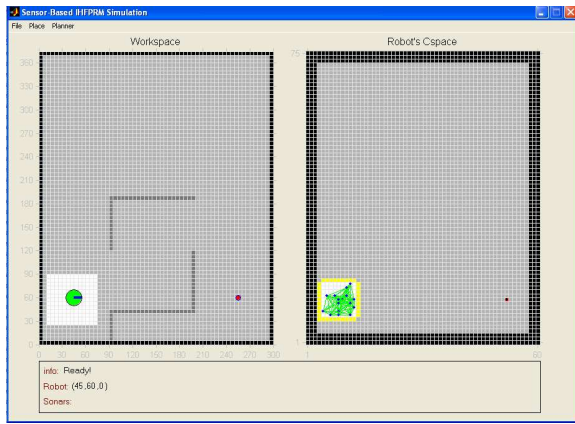


Fig. 19. Sensor-based simulator designed in MATLAB.

to the unknown free regions of the workspace are shown in gray and those corresponding to the unknown obstacles are shown in darker color⁸. The boundary of the workspace is assumed to be known and it is shown in black. The initial robot's C-space is shown in Fig.20(b). The grids corresponding to the unknown configurations are shown in gray. The black region shown in robot's C-space denotes C-obstacles. The roadmap initially generated in the small free part of the C-space corresponding to the small free region around the robot at its initial configuration is shown in green. The grids with yellow colors in the C-space represent the boundary of the unknown C-space, \mathcal{C}_{unkbnd} calculated using FFT.

The robot's task in this experiment is to start from the initial configuration and to scan the workspace interactively till it finds a feasible path (i.e., a collision-free path) to reach the goal configuration. Each time the robot takes a new scan the workspace \mathcal{W} , free C-space, \mathcal{C}_{free} , and the roadmap \mathcal{R} are updated. In the next step the planner finds the next view configuration q_v to scan the area defined by $\mathcal{A}(q_x)$ ⁹. The grids corresponding to the q_v and q_x at iteration i are denoted by v_i and x_i and are displayed in the workspace model, respectively. Fig.20(p) shows the final roadmap obtained at the 14th iteration where goal has been reached. Collision-free path connecting the start and the goal configurations is shown as a thick line on this roadmap as well (see Fig.20(p)). The robot at the goal configuration along with the path traversed by the robot in the workspace are shown in Fig.20(o). As it is shown the planner finds the shortest path passing through the narrow region between the two obstacles successfully.

Fig.21 shows the robot's snapshots obtained using the proposed motion planning technique for a different scenario. The shortest path in this experiment must pass through a hard-to-navigate passage. As it can be seen through Fig.21(a)-21(h) the planner finds the path connecting the start and goal configurations going through the narrow region between two obstacles¹⁰ after 11 iterations.

⁸One should note that the obstacles are plotted for illustration only. The positions of obstacles are not known *a priori*.

⁹ $\mathcal{A}(q)$ denotes the physical space occupied by the robot when the robot is at configuration q .

¹⁰One should note that any other path not going through this narrow region will be longer than the path going through this region in average.

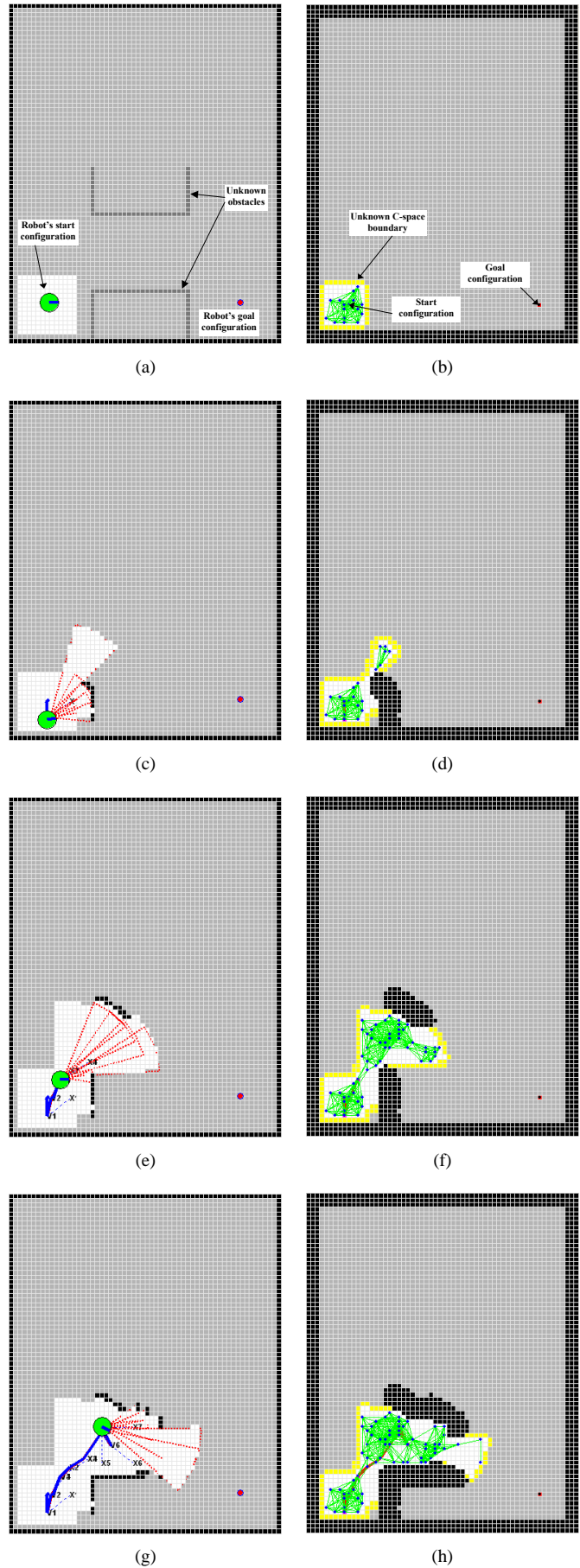


Fig. 20. Robot's snapshots obtained using our sensor-based technique implemented on a mobile robot with 3 DOF. Left column: the workspace models from iteration 0 to 14. Right column: the robot's C-space along with the roadmap obtained from iteration 0 to 14 (continued on next column).

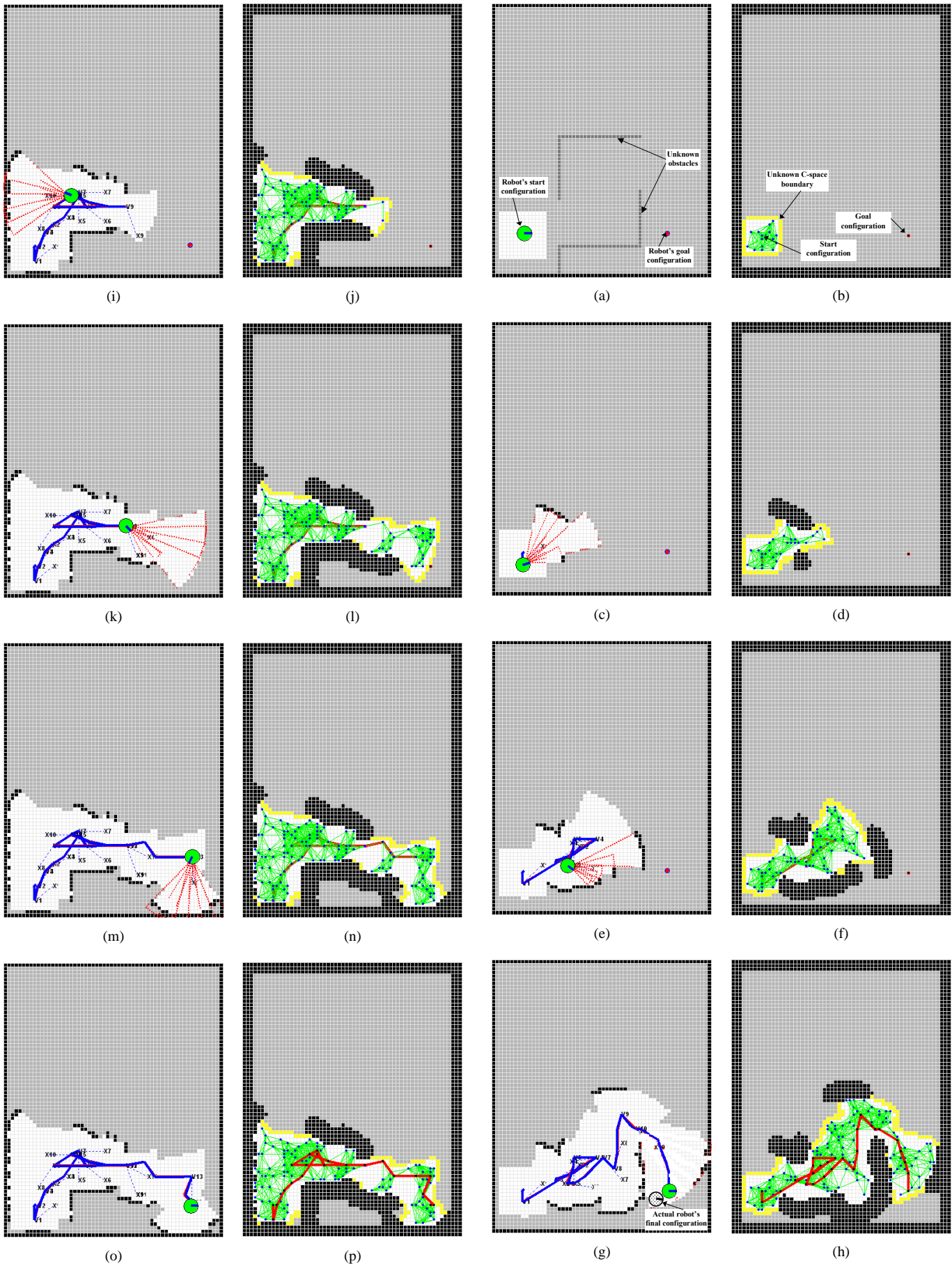


Fig. 20. (continued from last column) Robot's snapshots obtained using our sensor-based technique implemented on a mobile robot with 3 DOF.

Fig. 21. Robot's snapshots obtained using our sensor-based technique implemented on a mobile robot with 3 DOF. (a), (c), (e), and (g): the workspace model at iteration 0, 1, 6, and 11, respectively, (c), (d), (f), and (h): the robot's C-space along with the roadmap obtained at iteration 0, 1, 6, and 11.

It is noteworthy that the robot starts wandering around at a couple of points but it finds its path through the narrow region eventually. One should also note that due to the odometry errors the actual position of the robot in the real workspace will be different from the position calculated based on odometry data. This results in: (1) the map of the environment obtained based on range data through the planning process does not exactly correspond to the real workspace (This can be better understood by comparing the workspace model shown in Fig.21(a) and the actual model of the workspace shown in Fig.21(g)), and (2) the actual position of the robot when the goal is reached would be different from the final position of the robot calculated based on odometry data (This can be better visualized by looking at Fig.21(g) where the position of the robot at its goal configuration is slightly different from the position that the robot is supposed to reach.). In average an error of 10 to 20cm in position and 5 to 10 degrees in orientation are observed throughout the experiments.

These representative experimental results along with a number of other experiments carried out off-line reflect on the credibility of the proposed robot motion planning technique.

VII. CONCLUSIONS AND FUTURE WORKS

A hybrid motion planning technique based on harmonic functions and probabilistic roadmaps was proposed. It was shown through computer simulations and experimental results that this technique can solve MBMP and SBMP problems efficiently. The main idea of the proposed approach is to identify critical regions, i.e. narrow passages and/or hard-to-navigate regions, using information obtained about the environment topology through a fluid dynamic paradigm. Furthermore, these information will be passed to a PRM-based method to build a roadmap which efficiently captures the connectivity of the free C-Space.

To implement the proposed approach, two techniques were presented: (1) the Harmonic Function-based Probabilistic Roadmap (HFPRM) that is applicable to motion planning of robots with low DOF, and (2) the Incremental HFPRM (IHFRM) technique which is applicable to motion planning of robots with high DOF. The proposed HFPRM approach was implemented on a mobile robot with 3 DOF moving in an unknown environment cluttered with unknown obstacles. The robot's task is to gain information about the unknown environment incrementally through iterative scans using its on-board sensors and to find a collision-free path to reach a pre-defined goal configuration. The information obtained through a fluid dynamic paradigm is utilized to prioritize different regions in the workspace at the scan planning stage. A PRM-like method is utilized to capture the free portion of the robot's C-space in form of a graph. This graph is then searched for the shortest path that connects the robot to the goal configuration.

Performance of the proposed motion planning technique was fully investigated through computer simulations and experimental results. It was shown that the proposed technique is capable of solving the target-seeking problem in complex environments not known *a priori*.

This work can be extended in the following areas:

- extension of this work to the sensor-based motion planning of robotic arms with many DOF, and
- sensor-based motion planning in environments with dynamic obstacles.

Furthermore, a more comprehensive study on the error analysis of the system will add to the robustness of the proposed technique.

ACKNOWLEDGMENT

The financial support provided by the Natural Sciences and Engineering Research Council of Canada (NSERC) in form of a discovery grant is greatly acknowledged.

REFERENCES

- [1] C. I. Connolly, J. B. Burns, and R. Weiss, "Path planning using laplace's equation," in *Proc. of the IEEE International Conference on Robotics and Automation*, Cincinnati, OH, Apr. 1990, pp. 2102–2106.
- [2] C. I. Connolly and R. Grupen, "On the applications of harmonic functions to robotics," *Journal of Robotic Systems*, vol. 10, no. 7, pp. 931–946, 1993.
- [3] L. E. Kavraki, P. Svesktka, J. C. Latombe, and M. H. Overmars, "Probabilistic roadmaps for path planning in high-dimensional configuration space," *IEEE Trans. Robot. Automat.*, vol. 12, no. 4, pp. 566–580, June 1996.
- [4] O. Khatib, "Real-time obstacle avoidance using harmonic potential functions," *International Journal of Robotics Research*, vol. 5, no. 1, pp. 90–98, 1986.
- [5] J. Barraquand, B. Langlois, and J. C. Latombe, "Numerical potential field techniques for path planning," *IEEE Trans. Robot. Automat.*, vol. 22, no. 2, pp. 224–241, 1992.
- [6] A. A. Masoud and M. M. Bayoumi, "Intercepting a maneuvering target in a multidimensional stationary environment using a wave equation potential field strategy," in *Proc. of the IEEE International Symposium on Intelligent Control*, Columbus, OH, Aug. 1994, pp. 243–248.
- [7] K. Sato, "Deadlock-free motion planning using the laplace potential field," *Advanced Robotics*, vol. 7, no. 5, pp. 449–461, 1993.
- [8] C. Louste and A. Liegeois, "Near optimal robust path planning for mobile robots: the viscous fluid method with friction," *Journal of Intelligent and Robotic Systems*, vol. 27, no. 1-2, pp. 99–112, 2000.
- [9] Y. Wang and G. Chirikjian, "A new potential field method for robot path planning," in *Proc. of the IEEE International Conference on Robotics and Automation*, San Francisco, CA, Apr. 2000, pp. 977–982.
- [10] J. Guldner, V. Utkin, and H. Hashimoto, "Robot obstacle avoidance in n-dimensional space using planar harmonic artificial potential fields," *Journal of Dynamic Systems, Measurement, and Control*, vol. 119, no. 2, pp. 160–166, 1997.
- [11] V. Utkin, S. Drakunov, and F. H. H. Hashimoto, "Robot path obstacle avoidance control via sliding mode approach," in *Proc. of the IEEE International Workshop on Intelligent Robots and Systems*, Osaka, Japan, Nov. 1991, pp. 1287–1290.
- [12] D. Hsu, L. E. Kavraki, J. C. Latombe, R. Motwani, and S. Sorkin, "On finding narrow passages with probabilistic roadmap planners," in *Proc. of Int. Workshop on Algorithmic Foundation of Robotics (WAFR)*, Houston, TX, Mar. 1998, pp. 141–154.
- [13] N. M. Amato, O. B. Bayazit, L. K. Dale, C. V. Jones, and D. Vallejo, "Obprm: An obstacle-based prm for 3d workspaces," in *Proc. of Int. Workshop on Algorithmic Foundation of Robotics (WAFR)*, Houston, TX, Mar. 1998, pp. 155–168.
- [14] O. B. Bayazit, G. Song, and N. M. Amato, "Enhancing randomized motion planners: Exploring with haptic hints," in *Proc. of the IEEE International Conference on Robotics and Automation*, San Francisco, CA, Apr. 2000, pp. 529–536.
- [15] C. Holleman and L. Kavraki, "A framework for using the workspace medial axis in prm planners," in *Proc. of the IEEE International Conference on Robotics and Automation*, San Francisco, CA, Apr. 2000, pp. 1408–1413.
- [16] D. Hsu, T. Jiang, J. Reif, and Z. Sun, "The bridge test for sampling narrow passages with probabilistic roadmap planners," in *Proc. of the IEEE International Conference on Robotics and Automation*, Taipei, Taiwan, Sept. 2003, pp. 4420–4426.

- [17] J. P. Berg and M. H. Overmars, "Using workspace information as a guide to non-uniform sampling in probabilistic roadmap planners," in *Proc. of the IEEE International Conference on Robotics and Automation*, New Orleans, LA, Apr. 2004, pp. 453–460.
- [18] T. Simeon and J.-P. Laumond, "Visibility based probabilistic roadmap for motion planning," *Advanced Robotics*, vol. 14, no. 6, pp. 477–493, 2000.
- [19] C. Nissoux, T. Simeon, and J.-P. Laumond, "Visibility based probabilistic roadmaps," in *Proc. of the IEEE International Conference on Intelligent Robots and Systems*, Oct. 2003, pp. 382–387.
- [20] J. J. Kuffner and S. M. LaValle, "Rrt-connect: An efficient approach to single-query path planning," in *Proc. of the IEEE International Conference on Robotics and Automation*, San Francisco, CA, Apr. 2000, pp. 995–1001.
- [21] E. Mazer, J. M. Ahuactzin, and P. Bessiere, "The ariadne's clew algorithm," *Journal of Artificial Intelligence Research*, vol. 9, pp. 295–316, 1998.
- [22] P. Iniguez and J. Rosell, "Probabilistic harmonic-function-based method for robot motion planning," in *Proc. of the IEEE International Conference on Intelligent Robots and Systems*, Las Vegas, NV, Oct. 2003, pp. 382–387.
- [23] J. Rosell and P. Iniguez, "A hierarchical and dynamic method to compute harmonic functions for constrained motion planning," in *Proc. of the IEEE International Conference on Intelligent Robots and Systems*, Lausanne, Switzerland, Oct. 2002, pp. 2335–2340.
- [24] D. Aarno, D. Kragic, and H. I. Christensen, "Artificial potential biased probabilistic roadmaps method," in *Proc. of the IEEE International Conference on Robotics and Automation*, New Orleans, LA, Apr. 2004, pp. 461–466.
- [25] P. Renton, M. Greenspan, H. A. Elmaraghy, and H. Zghal, "Plan-scan: A robotic system for collision-free autonomous exploration and workspace mapping," *Journal of Intelligent and Robotic Systems*, vol. 24, pp. 207–234, 1999.
- [26] Y. Yu and K. Gupta, "Sensor-based probabilistic roadmaps: experiments with an eye-in-hand system," *Advanced Robotics*, vol. 14, pp. 515–536, 2000.
- [27] —, "C-space entropy: A measure for view planning and exploration for general robot sensor systems in unknown environments," *International Journal of Robotics Research*, vol. 23, pp. 1197–1223, 2004.
- [28] E. Cheung and V. J. Lumelsky, "Motion planning for a whole-sensitive robot arm manipulator," in *Proc. of the IEEE International Conference on Robotics and Automation*, Cincinnati, Ohio, May 1990, pp. 344–349.
- [29] K. N. Kutulakos, V. J. Lumelsky, and C. R. Dyer, "Vision-guided exploration: A step toward general motion planning in three dimensions," in *Proc. of the IEEE International Conference on Robotics and Automation*, May 1993, pp. 289–296.
- [30] M. Mehrandezh and K. Gupta, "Simultaneous path planning and free space exploration with skin sensor," in *Proc. of the IEEE International Conference on Robotics and Automation*, Washington, DC, May 2002, pp. 3838–3843.
- [31] D. Keymeulen and J. Decuyper, "The fluid dynamic applied to mobile robot motion: the stream field method," in *Proc. of the IEEE International Conference on Robotics and Automation*, Sacramento, CA, Apr. 1994, pp. 790–796.
- [32] —, "Self-organizing system for the motion planning of mobile robot," in *Proc. of the IEEE International Conference on Robotics and Automation*, Minneapolis, Apr. 1996, pp. 3369–3374.
- [33] J. D. Anderson, *Fundamentals of Aerodynamics*. New York: McGraw-Hill, 1991.
- [34] D. A. Anderson, J. C. Tannehill, and R. H. Pletcher, *Computational Fluid Mechanics And Heat Transfer*. New York: McGraw-Hill, 1984.
- [35] L. E. Kavasaki, "Computation of configuration-space obstacles using the fast fourier transform," *IEEE Trans. Robot. Automat.*, vol. 11, no. 3, pp. 408–413, 1995.
- [36] M. Kazemi and M. Mehrandezh, "Robotic navigation using harmonic function-based probabilistic roadmaps," in *Proc. of the IEEE International Conference on Robotics and Automation*, New Orleans, LA, Apr. 2004, pp. 4765–4770.
- [37] Dr robot inc. [Online]. Available: <http://www.drrobot.com/>



Moslem Kazemi received his B.Sc. degree in Computer Engineering (hardware) from Sharif University of Technology, Tehran, Iran, in 2000. He completed his M.Sc. degree on "Robotic Navigation Using Harmonic Function-based Probabilistic Roadmaps" at University of Regina, Regina, Canada, in 2004. He is currently a Ph.D. candidate in the school of Engineering Science (Robotics Lab) at Simon Fraser University, Burnaby, Canada. His research interests include robotics, control, and mechatronics.



Mehran Mehrandezh received his B.Sc. (in mechanical engineering) from Sharif University of Technology in Tehran, Iran, his M.Sc. from Queen's University (Automatic Controls Lab.) in Kingston, Canada, and his Ph.D. from University of Toronto (CIMLab) in Toronto, Canada in 1989, 1995, and 1999, respectively. He also spent two years in the School of Engineering Science at Simon Fraser University in Vancouver, Canada as a Post Doctoral fellow. He is currently an assistant professor in the Faculty of Engineering at the University of Regina, Regina, Canada. His research interests include motion planning of co-operating robotic arms in dynamic environments, autonomous sensor-based motion planning, robotic interception of moving objects, and urban search and rescue robotics.



Kamal Gupta received his Ph.D. degree in Electrical Engineering from McGill University, Canada, in 1987. He is currently a professor in the School of Engineering Science at Simon Fraser University, Burnaby, Canada. His research interests include computational robotics, motion planning, spatial reasoning, interpretation of three dimensional scenes.

Atmospheric Pressure, Low cost, Surface Micromachined Pirani Pressure Sensor

Kourosh Khosraviani

Abstract— An innovative thermal pressure sensor, based on short-distance thermal conduction, and for operation in atmospheric pressure ranges, has been explored. This simple device consists of a heater made of a polysilicon resistor, a heat sink made by the silicon substrate, and a tiny air gap between the heaters and the substrate. With innovative ideas and simple processes, the gap can be made in the order of nanometers in size, inexpensively, which significantly increases the measurement range of the sensor. Our sensor is capable of measuring applied pressures up to 700kPa (7atm), and has been fabricated with a 3-mask, in a 3 μ m standard Complementary Metal Oxide Semiconductor (CMOS) compatible process. The device can be made as small as 50x50 μ m², less than the size of a standard bonding pad.

Index Terms—Micromachining, Pirani, Pressure Sensor

I. INTRODUCTION

THE Pirani gauge, which was invented by -Marcello Pirani- in 1906, utilizes the thermal conductivity of gases to detect pressure changes when the pressure is below a given limit [1]. The sensor simply consists of a hot wire with a temperature-dependent electrical resistance [2]. The temperature of the wire is determined by the constant electrical power that heats the wire and the heat loss to the environment by the thermal conduction of the surrounding gas. Higher pressure brings more gas molecules in contact with the wire and reduces the temperature of the wire. The resistance of the wire changes with the temperature and becomes a measure of the applied pressure.

Significant progress in the miniaturization of Pirani sensors has been made since silicon micromachining has been introduced. The benefits of micromachining are simple design, robust devices, low-cost fabrication, low power consumption, and fast responses. These advantages, as well as the prospect of digital sensors and integrated system-on-chip, have led to the development of many micro machined Pirani sensors over recent years.

Most micro machined Pirani sensors operate in the vacuum pressure range, a pressure range well below 1atm (101kPa), because, a small gap is hard to make without complicating the manufacturing processes and increasing the cost. The common procedure is to shrink the gap and keep the heater membrane straight without touching the heat sink. When the gap reduces to the sub-micron range, this procedure becomes challenging because of the surface tension, and the sticking effect in the structure releasing step.

Herwaarden [3] introduced an integrated pressure sensor based on an integrated heater and a thermopile. The heater and the thermopile were on a cantilever beam, which was suspended over a substrate; the measuring range was 1mPa to 1kPa. Higashi [4] fabricated a sensor with a NiFe suspended resistor over a 25 μ m deep V-groove which was anisotropically etched in the substrate. The operating range was 103.4Pa to 1.03MPa. Robinson and Parameswaran [5] reported a sensor with polysilicon resistors and an oxide platform which was suspended over an anisotropically etched 50 μ m deep trench; that showed 2% sensitivity over the pressure range 1.33Pa to 13.3kPa. Ping [6] fabricated a sensor with platinum resistors on a suspended membrane over a V-groove cavity with the pressure range of 0.01Pa to 800Pa. Oliver and Henry [7] built a CMOS compatible sensor, which was operating over 100Pa to 10⁶Pa, with metal resistors and a 0.65 μ m thick metal sacrificial layer. Alvesteffer [8] introduced a sensor with multilayer metal resistors with a silicon lid on top of the resistors that served as a heat sink. Air flows both on top and below the resistors and the operating range was between 0.013Pa to 101kPa. Bruce [9],[10] used a crucial freeze-drying step to avoid sticking the heater and heat sink and made a device with 300nm gap which had linear output up to 7atm. Puers [11] reported an extremely miniaturized sensor fabricated by Focus Ion Beam (FIB). The heater was a suspended tungsten bridge and a 400nm gap between the bridge and substrate was made by FIB.

The sensor which we made, is based on the same working principle, but measures higher pressure (atmospheric or above) with a simple and low-cost fabrication procedure. Unlike other methods, which hold the heater membrane straight at a great expense, we let the heater membrane collapses. Taking advantage of the surface irregularities of the heater, a tiny air gap can be made inexpensively. The measurable pressure range has been extended to more than 7atm (700kPa), much higher than that of existing sensors. The procedure only needs three masks with a 3 μ m CMOS-compatible technology.

The device consists of four identical polysilicon resistors. Two of these resistors are sensing elements (called heaters), which are suspended over the silicon substrate by a tiny air gap and are used for pressure measurement. The other two polysilicon resistors are dummy resistors that lay directly on the substrate and are used for temperature compensation. All of the resistors have the same structure and dimensions, and are sandwiched by silicon dioxide (SiO₂) films. The sensor is

tested by connecting the four resistors in a Wheatstone bridge configuration, although each heater can be tested individually as well. The gap between the collapsed heater structure and the substrate is estimated to be in the nanometers range in size.

We can still measure pressure by gas thermal conduction, even though the heater collapses to the heat sink, because of the surface irregularities of the contacts between the heater and the heat sink. Surface irregularities consist of macroscopic and microscopic irregularities. Macroscopic irregularities include flatness deviations and waviness while microscopic irregularity is measured by surface roughness [12]. Both irregularities exist in all practical solid surfaces. “Two solid surfaces apparently in contact only touch each other at a few individual spots. Even at relatively high contact pressure (of the order of 10MPa), the actual contact area for most solid surfaces is only about 1 to 2% of the nominal contact area [12]”.

II. MICROSTRUCTURE DESIGN AND FABRICATION

We used n-doped polysilicon to construct the heaters, because of the polysilicon good electrical and thermal properties, compatibility with standard CMOS fabrication processes, and the granular surface to build the gap.

The fabrication process is carrying on a silicon (100) wafer substrate with three different layers on top of it:

- 1) A 0.5 μm thermally grown silicon dioxide, which is used as a sacrificial layer during the release of the heaters.
- 2) A 0.55 μm n-type Low Pressure Chemical Vapor Deposition (LPCVD), which is used to build the heaters.
- 3) A 0.5 μm LPCVD silicon dioxide, which is used as a mask for etching the polysilicon.

In order to increase the performance of the sensor and to compensate for ambient temperature variations, a Wheatstone bridge was made which consists of two released and two unreleased heaters. The two unreleased heaters, unlike the sensing heaters, are not released from the substrate; therefore, their values do not change with pressure and only depend on the substrate temperature.

The basic fabrication process included the following steps:

- 1) Patterning the LPCVD dioxide to make a mask for the polysilicon.
- 2) Patterning the polysilicon by existing dioxide as a mask using TMAH as an etching solution.
- 3) Patterning the LPCVD dioxide in order to open a window for bonding pads and for protecting the unreleased heaters from being undercut during the heater releasing process.
- 4) Depositing aluminum and patterning the aluminum in order to make bonding pads.

- 5) Sacrificially etching the thermal oxide and releasing the heaters, which cause the heaters to collapse to the substrate.

Fig. 1 shows the released and unreleased heaters with aluminum and oxide on top of the unreleased heaters. Heaters have a U-shape cantilever structure about 10 μm wide and 250 μm long. Typical resistance of the heaters at room temperature and atmospheric pressure is about 900 Ω . In order to prevent any physical damage to the released heaters, the heater release step of the fabrication process is done after cutting and separating individual dies. In order to do this, the silicon dioxide etchant should be compatible with aluminum and must not attach it. Standard buffered oxide etchant (BOE) does etch the aluminum, so we used a modified solution for etching silicon dioxide and releasing the heaters. A mixture of 30cm³ Glycerol and 50cm³ BOE, which does not etch aluminum, was used as a silicon dioxide etchant during the sacrificially etching process.

After releasing the heaters by sacrificially etching the thermal oxide, and during the final rinsing process we let the membrane heater collapse to the substrate (the heat sink). A very thin layer of native oxide is made during the final rinsing process of fabricating the sensor, which isolates the released heaters from the substrate. This thin layer of oxide can also be formed easily by a short period dip in 30% H₂O₂ [13]. When the released heaters collapse to the substrate, the surface roughness of the poly heaters makes tiny gaps between the heat sink and the heater. The gap size depends on the polysilicon grain size, which is in the order of a few tens of nanometer. The polysilicon grain size itself depends on the deposition parameter. Therefore, we were able to build a nanometer size gap between the heater and the heat sink,

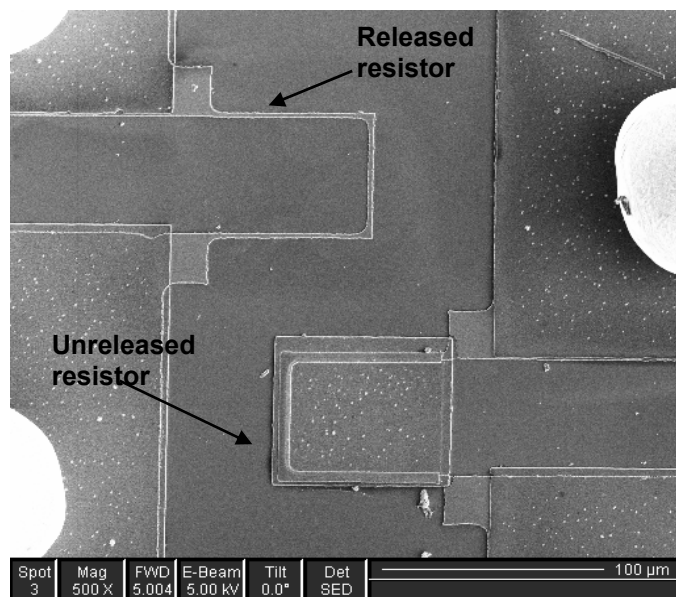


Fig. 1. Released and unreleased heaters with aluminum and oxide on top of the unreleased heaters.

which dramatically increases the operating pressure range of the sensor.

III. EXPERIMENTAL RESULT

The behavior of 5 separate sensors with same heater dimension and structure were investigated. A constant voltage source is connected to the Wheatstone bridge and the passing current is measured in order to calculate the power density of the sensor. Power density is defined as the ratio of applied power to the cross section area of the resistors ($mW/\mu m^2$). When the pressure changes from 100kPa (1atm) to 800kPa (8atm), the sensors output increases from around 2mV to 19mV for the five sensors, respectively. Fig. 2 depicts the variation of the Wheatstone bridge output versus the variation of the sensors applied pressure. The change in output, results from the expected decrease in the heater resistance with rising pressure. Higher pressure introduces more air molecules for heat conduction, and more heat loss leads to a temperature decrease and reduced heater resistances. The change is more significant at the lower end of the applied pressure range, and slows down when pressure is increased and reaches the value of the transition pressure.

The lower limit of the measurement range is determined by solid conduction and thermal radiation. At low pressures, simply not enough gas molecules available for efficient thermal conduction, so the thermal conductivity of gases approaches zero and the thermal conduction through solid contacts dominates.

The upper limit of the measurement range is set by the transition pressure P_{tr} . For example, P_{tr} is about 70psi (483kPa) for a gap size of 50nm between the heater and the heat sink, which agrees with the experimental data from our test result and others [14], [15].

The device output variation versus applied pressure variation is not linear. The two major factors contributing to the non linearity are the pressure-independent solid thermal conduction, and the pressure dependent Knudsen number. The solid thermal conduction is determined by the solid contact area between the polysilicon resistor and the silicon substrate, including the bonding pads and the collapsed part of the resistors. The solid thermal conduction is constant over the entire operating pressure range, and its ratio to the air gap area, affect the nonlinearity.

The Knudsen number (K_N) is a dimensionless parameter that characterizes the gas behavior at low pressure [16]. It is defined as the ratio of the molecular Mean Free Path (MFP), to the characteristic dimension of the container (d), which is the distance traveled by gas molecules [16].

The thermal conductivity of the heater is proportional to the reciprocal of the Knudsen number. Therefore, the nonlinearity introduced to the thermal conductivity of the heaters by the Knudsen number, is more significant at the higher end of the operating pressure range. With rising pressure, the mean free path of the gas molecules becomes smaller as does the Knudsen number; therefore, the thermal conductivity of the heaters is reduced.

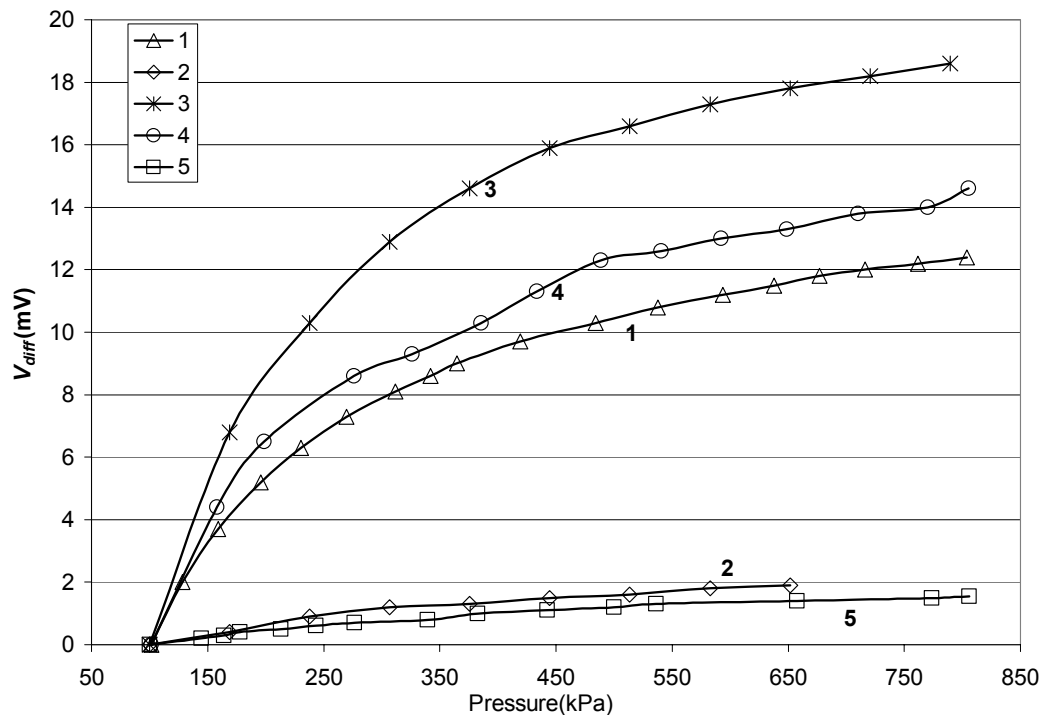


Fig. 2. Output variation versus Pressure under the power density of $0.01 mW/\mu m^2$

Another experiment was conducted to analyze the effect of different applied power densities on the device sensitivity. Device sensitivity (S) is defined as normalized variation in the bridge output voltage (V_{out}) per unit of the applied pressure.

$$S = \frac{V_{out}}{V} \frac{1}{P}, \quad (1)$$

in which V is the bridge excitation voltage and P is the sensor pressure. Increasing the power density showed a significant increase in the device sensitivity. However, the device sensitivity shows more increase in the lower half of the applied pressure.

IV. DISCUSSIONS

Despite the heaters collapsing to the substrate after the final rinse, the experimental results support our expectation of a tiny gap formation between the heaters and the substrate, due to the polysilicon heater's surface roughness.

Heat is conducted away from the heaters in two directions, vertical and horizontal directions. In the horizontal direction, heat flows through the body of the heaters to the bonding pads and finally to the substrate. However, heat conduction in the vertical direction flows through the air gap between the heaters and the substrate, and through the solid contact between them. Base on our heater's analytical model and

theoretical calculation, the horizontal heat conduction is negligible compared to the vertical heat conduction.

Thermal conduction in the vertical direction consists partly through the solid contacts, and partly through the air gap. The magnitude of each term depends on the beam structure. Based on the analysis of the conductions through these two parts, a model has been derived to relate the sensor's sensitivity to the coefficient ξ which denotes the percentage of the heater surface area, making solid contact to the substrate.

In our analysis, we assume that only ξ percent of the heater's surface has a solid contact to the substrate and $(1 - \xi)$ percent of the heater's surface has no contact to the substrate, because of the surface irregularities. Theoretical calculation of the equivalent thermal conductance between the heaters and the substrate, for ξ value between 1% and 5%, are consistent with experimental data from sensors 1, 3, and 4. The equivalent thermal conductance between the heaters and the substrate is calculated by adding the thermal conductance of the air gap, and the thermal conductance of the solid contact between the heaters and the substrate. Fig. 3 depicts the theoretical and the experimental data for the heaters equivalent thermal conductance. Sensor's 2 experimental data for the equivalent thermal conductance corresponds to $\xi = 50\%$ of the numerical calculation, indicate that this heater hasn't been released completely.

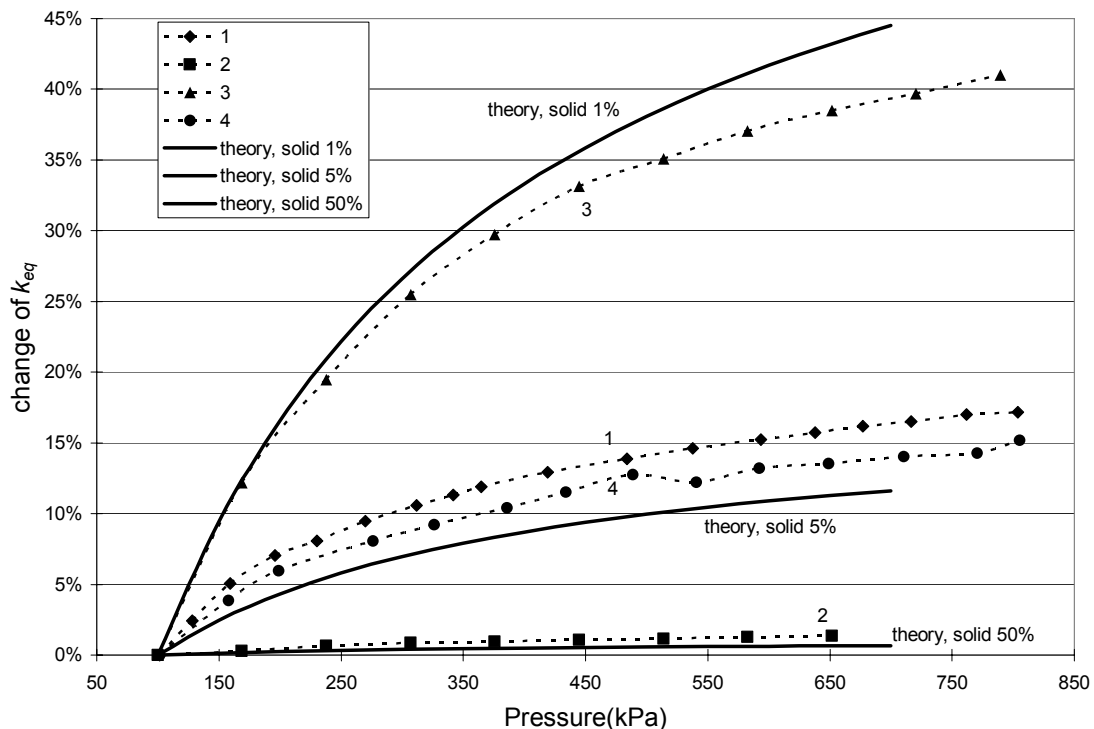


Fig. 3. Variation of equivalent thermal conductance (k_{eq}) versus different values of ξ .

V. CONCLUSION

An innovative short-distance thermal conduction pressure sensor fabricated by surface micromachining technique was presented in this paper. The advantages of the presented device are: the simple and inexpensive fabrication procedure, and the wide measuring pressure range. Our device has been built by using 3 masks in $3\mu\text{m}$ and CMOS compatible fabrication processes. Experimental results show significant sensitivity to the applied pressure. Device output voltage shows variation between 2mv to 19mv by changing the applied pressure between 100kPa to 800kPa with a power density of $0.01\text{mW}/\mu\text{m}^2$. Sensitivity of the sensor increases by increasing the power density which is more significant in the lower half of the pressure range. Although we let the heater's membrane touch the substrate, comparison of analytical and experimental data revealed that only 1 to 5 percent of the heater's surface has solid contact with the substrate which is very important insight that has been facilitated our fabrication process.

[16] J. Baker, M.E. Calvert, D.J. Power, E.T. Chen, M.L. Ramalingam, and T.R. Lamp, "On the role of the Knudsen number with respect to heat transfer in micro-scale flows", Proceedings of the 31st Intersociety Energy Conversion Engineering Conference, Vol. 2, 11-16, pp.1396 – 1401, 1996



Kourosh Khosraviani received his B.Sc in electronic engineering from Shahid Beheshti (Meli) University, Tehran/Iran, in 1995. He is a M.A.Sc student in school of engineering science at Simon Fraser University, British Columbia, Canada. His current research interest is micro electromechanical systems (MEMS). He has several years of work experience in different branches of industry as design engineer and automation supervisor.

REFERENCES

[1] H. Turnbull, R.S. Barton and J.C. Riviere, "An Introduction to Vacuum Technique", Grotge Newnes Limited, 1962
 [2] G.C.M. Meijer and A.W. Herwaarden, "Thermal Sensors", Institute of Physics Publishing, Bristol and Philadelphia, 1996
 [3] Bruce C.S Chou, Chung-Nan Chen and Jin-Shown Shie, "Fabrication and Study of a Shallow-Gap Pirani Vacuum Sensor with a Linearly Measurable Atmospheric Pressure Range", Sensors and Materials, Vol. 11, No.6, pp.383-392, 1999
 [4] S.D. James, R.G. Johnson and R.E. Higashi, "A Broad Range Absolute Pressure Microsensor", Proc. IEEE Solid-State Sensor and Actuator Workshop, pp.107-108, 1988
 [5] A.M. Robinson, P. Haswell, R. P. W. Lawson and M. Parameswaran, "A thermal conductivity microstructural pressure sensor fabricated in standard complementary metal-oxide semiconductor", Review of Scientific Instruments, 63 (3), pp.2026-2029, 1992
 [6] Ping Kuo Weng and Jin-Shown Shie, "Micro-Pirani vacuum gauge", Review of Scientific Instruments, 65 (2), pp.492-499, 1994
 [7] Oliver Paul and Henry Baltes, "Novel fully CMOS-compatible vacuum sensor", Sensors and Actuators A, Vol. 46-47, pp.143-146, 1995
 [8] W. J. Alvesteffer, D. C. Jacobs, and D. H. Baker, "Miniaturized thin film thermal vacuum sensor", The Journal of Vacuum Science and Technology A, 13 (6), pp.2980-2985, 1995
 [9] Bruce C.S. Chou and Jin-Shown Shie, "An Innovative Pirani Pressure Sensor", Transducers '97, 1997 International Conference on Solid-State Sensors and Actuators, pp.1465-1468, 1997
 [10] A.W. Herwaarden and P.M. Sarro, "Integrated vacuum sensor", Sensors and Actuators, Vol. 8, pp.187-196, 1985
 [11] R. Puers, S. Reyntjens and D. De Bruyker, "The NanoPirani--an extremely miniaturized pressure sensor fabricated by focused ion beam rapid prototyping", Sensors and Actuators A, Vol. 97-98, pp.208-214, 2002
 [12] C.V. Madhusudana, "Thermal Contact Conductance", Springer, 1996
 [13] Marc Madou, "Fundamentals of Microfabrication", CRC Press, 1997
 [14] R. Puers, S. Reyntjens and D. De Bruyker, "The NanoPirani--an extremely miniaturized pressure sensor fabricated by focused ion beam rapid prototyping", Sensors and Actuators A, Vol. 97-98, pp.208-214, 2002
 [15] Carlos H. Mastrangelo and R.S. Muller, "Microfabricated thermal absolute-pressure sensor with on-chip digital front-end processor", IEEE J. Solid-State Circuits, Vol. 26, No.12, pp.1998-2006, 1991

Surgical Simulations: It's All in a Game!

Gaming techniques for medical applications.

Vidya Kotamraju*, Shahram Payandeh, and John Dill

Abstract—Computer games have come a long way since A.S. Douglas' Tic-Tac-Toe, evolving into a well-understood set of methods, including recent developments in realism and immersiveness for game scenarios. Surgical Simulation, on the other hand, has barely a decade of technological development. Surgical techniques have evolved from direct hands-on maneuvers to indirect minimal-access procedures. Learning these techniques, though difficult, is possible by the use of virtual training environments. The common component of virtual reality provides an opportunity to apply game programming ideas to such training environments. This paper outlines key techniques in three-dimensional game applications that can be adapted to enhance current surgical simulation technology. A description of common game logic and presentation techniques with an illustration of their potential use in surgical simulations is provided. The choice and implementation of a technique must be based on application requirements.

I. INTRODUCTION

THE first computer game, a version of Tic-Tac-Toe, was created in 1952 by A.S. Douglas. Since then, game programming has evolved into a specialized discipline with well-understood methods forming the core of most game projects. The need for realism and immersiveness in game scenarios has added to the complexity and flexibility of the techniques involved.

Computer games are real-time applications with data acquisition and response being performed under time-critical conditions. These interactive applications consist of a virtual-world simulator that accepts real-time data, control mechanisms for user interaction, and a computation and display module that constantly monitors the simulator to change the parameters of the virtual world. Game programming can be divided into two sections: logic and presentation. Game logic deals with the implementation of actions in response to events in the game while game presentation pertains to the interfaces needed, within the context of the game theme, to give the user a realistic experience.

Surgical Simulation is a relatively new technological revolution taking place due to the advent of minimal access procedures in modern surgery. These procedures, like Laparoscopy, involve the use of video cameras to show the surgeon the instruments and operating site. Learning these techniques is difficult and, thus, represents an important challenge to face. The latest advances in the field of education offer an increasing number of innovative learning tools for such techniques.

Manuscript received June 21, 2005; revised July 21, 2005. Asterisk indicates corresponding author.

The authors are with the School of Engineering Science, Simon Fraser University, Burnaby, BC V5A1S6 Canada (E-mail: vkotamra@sfu.ca).

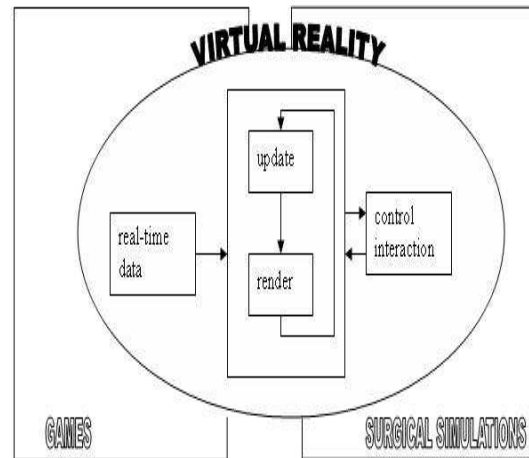


Fig. 1. Virtual reality: common component of game programming and surgical simulations.

Recent progress in computer technology has allowed the development of virtual environments to train the surgeon-to-be. Virtual reality is considered to be a promising area with a high potential of enhancing the training of health-care professionals. Virtual reality can provide an interactive and engaging educational context, thus supporting experiential learning-by-doing. As shown in Fig. 1, the common component of virtual reality provides an opportunity to apply game programming techniques to such training environments.

Reported work has explored the difference in the complexities of game programming and surgical simulations [1]. Much research has also been done in the application of gaming techniques to surgical learning [2] and [3]. This paper outlines key techniques lying at the heart of three-dimensional game applications that have a potential to enhance current surgical simulation technology. Significant research in this field is presented, followed by discussion on key issues concerning current techniques and simulator development directions.

The paper illustrates the main areas of application:

- 1) The use of multiple-degree-of-freedom haptic devices to provide force feedback.
- 2) The application of artificial intelligence for surgical motions and competence analysis.
- 3) Geometrical techniques for simplification and collision detection.

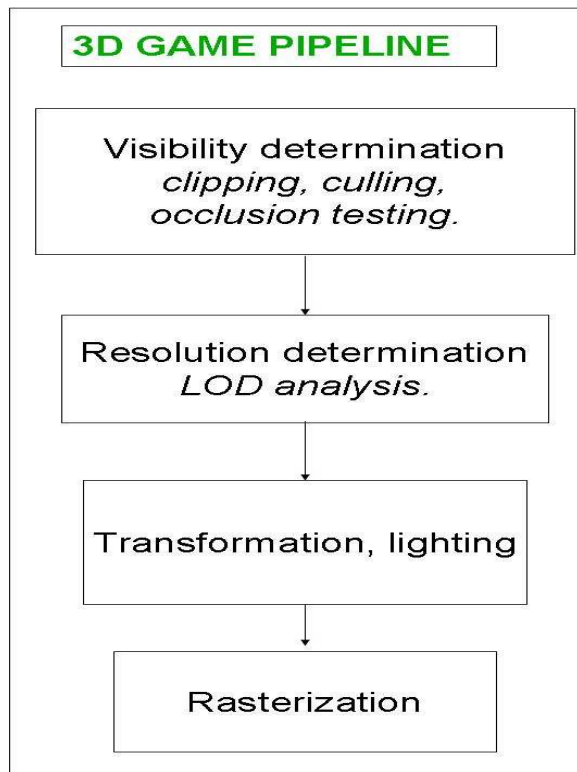


Fig. 2. A programmable graphics pipeline.

- 4) Network programming for remote surgical training.
- 5) Programming patterns, usability patterns and performance tuning for maximizing performance.
- 6) Enhancement of visual details through multiple-resolution scenes and illumination effects.
- 7) Bleeding simulations with dynamic unconnected mass points and data/patient-specific view with image wraps.

II. PIPELINE PARALLELS

A. The Graphics Pipeline

Three dimensional games follow a fixed sequence of steps that is referred to as the graphics pipeline in [4] and is illustrated in Fig. 2.

The graphics pipeline can be described in terms of four essential stages:

- 1) Visibility determination
- 2) Resolution determination
- 3) Transformation and lighting
- 4) Rasterization

Initially, a game application accepts raw geometry data (typically in the form of primitives like triangles, polygons etc.) specified in the world coordinate system. This geometry is then converted into normalized coordinate space, eliminating portions of the scene beyond the user's view port. The elimination is done by means of standard clipping, culling, and occlusion testing algorithms. Next, multi-resolution algorithms are implemented to view different parts of the scene in different resolutions. This technique is more commonly termed level-of-detail and results in rapid rendering. Before scene

display, geometric transformations are performed on the input data. The transformations include translation, rotation, and scaling followed by projection of the three-dimensional transformed data to the two-dimensional screen. The lighting stage implements specific lighting models to increase the realism of the scene. Through the lighting models, light sources, surface properties and scene-illumination can be defined. Lastly, the geometry data is converted into pixel data and rendered on the view-screen.

B. Surgical Simulation Framework

The surgical simulation framework can be divided into three sections:

- 1) deformable models.
- 2) collision detection and response.
- 3) visual and haptic feedback.

The user is provided with a three-dimensional view of the virtual organ and the operation tool, with their interactions shown in real-time. These interactions are enabled via a haptic device, through which the position of the tool can be controlled. As the interactions occur, deformations take place resulting in the update of the position and orientation of the organ. The haptic device then provides the user with the feedback that results from such deformations (see Fig. 1).

III. GAME LOGIC

Game logic delves into user input, artificial intelligence, geometry, network programming, design patterns and performance tuning issues. These issues address user interaction, action and strategic artificial intelligence, multi-player programming and macro-level code analysis.

A. User Input

A good user interaction model is an essential element of game programming. The keyboard, mouse and joystick are the most common peripheral devices providing input to the game user. Force feedback is also used to provide direct perception of virtual objects to the user. Such feedback improves realism and immersion in games.

Haptics is an emerging force feedback technology that adds a sense of tactation by means of high-speed vibrations or other stimuli (see Fig. 3). Deformation modeling, an integral part of surgical simulation, demands the use of accurate visual and haptic feedback. The importance of force feedback for assessing the validity of simulation and training has been demonstrated in [5]. Providing such feedback enables the user to feel contact forces and see graphical deformations computed in real time. PHANToM from Sensable Technology is a haptic device that has been incorporated into several current simulation systems [6] and [7]. Generally, a six degree-of-freedom device providing force and torque feedback is used to transfer organ-force response to the user. Inclusion of a haptic device in the surgical simulation framework can be done in two stages: collision detection and collision response. As the user manipulates the probe of the haptic device, the new position and orientation of the probe are acquired, and collisions with



Fig. 3. Haptic device for force feedback.

the virtual objects are detected. Once a collision is detected, contact forces are updated, and the collision response stage provides the user with the tactual and visual representation of the virtual objects.

Unpublished experiments in [8] indicate a fast update rate of 500 Hz - 1kHz for haptic devices. The authors state that lower rates in haptic displays are undesirable as they create instability. In virtual environment simulations with deformable objects, graphic display frames are updated at rates of about 10 Hz. The difference in the simulation rates can cause an oscillatory behavior in the haptic device. These non desired oscillations may indicate that the interface is active which may become highly unstable and can inflict bodily harm on the operator. A solution for this difference in the frequency update rates between the haptic devices and virtual environments has been presented in [9]. The strategy used is to separate the haptic loop and graphical rendering loop such that the haptic interaction is done between the haptic device and the local topology of the model at a rate frequency of about 1KHz. This model is updated by the parameters sent by the physical simulation of the virtual environment at a rate frequency of about 10 Hz.

B. Design patterns

Design patterns are abstract systems of interaction between classes, objects, and communication flow. These patterns are proven solutions to well-established software engineering problems and are choices based on object-oriented principles. Design patterns are provided as lists, with a complete description of the problem and its optimal solution. The solution is given in terms of design and implementation. Model reusability is an advantage provided by these patterns.

Programming patterns and usability patterns are the major design patterns used in game programming. Programming patterns list common coding problems and their standard solutions while usability patterns provide solutions to human-computer-interaction and interface design problems. Spatial index, factory, and singleton programming patterns, state, focus, and progress usability patterns are a few main strategies identified to solve common surgical simulations' design problems.

Singleton is a global object with a single instance. It is represented as a class with a single public method, which can be used to request an instance of the singleton. Singletons are used to avoid multiple storage of global variables in memory. This problem is avoided as every singleton variable created for access by a class in the application points to the same object in memory. The texture manager, the haptic controller object, and even the user class can be singletons. Spatial indices are used to perform queries on large three-dimensional environments. These search mechanisms are efficient and can be implemented by the use of lists, quadrees or octrees. Collision detection queries can be implemented by spatial indices. Factories centralize construction and destruction of objects through abstract classes. This concept provides a global mechanism for handling objects.

States are visual cues to inform the user of the current configuration. Typically, a mouse cursor depicts states by providing options to the user. An essential issue with states is to keep them distinguishable and small in number. The focus pattern simplifies the identification of states. Focus is used to center the user's attention to a specific action. This implementation can be done by menu deactivation, blurring or use of multiple resolutions. Progress patterns display quantitative information about a process. Typical implementation of a progress is done by using a meter bar which has a beginning and an end. Usability patterns are extremely useful in virtual training environments and provide ample scope to the user for skill enhancement.

C. Geometrical Techniques

The interaction between surgical tools, tissues and organs requires contact loci to be determined. In [10], the authors describe this interaction in terms of two processes. The first process detects occurrence of contacts between the virtual tools and the virtual organs. The second one computes proper organ-response such as deformation fields and interaction forces.

Usually, the collision points are determined by tool trajectories and the geometrical models. In order to realistically simulate interactions between deformable objects, all contact points, including those due to self-collisions, must be considered. Several collision detection approaches meet the requirements of simulation environments with dynamically deforming objects. As shown in Fig. 4, these tests check intersections of an object either against a point or a ray or a bounding box. In most of the techniques for the determination of collision points, the objects are partitioned into a group of small convex cells and stored in the database hierarchically [11].

Collision response in surgical simulation usually involves simulation of soft tissues at lower frequency and rendering of forces at higher frequency. It computes realistic or physically based behavior of soft tissues in response to tool interaction. Among various techniques, the finite element (FEM) and Boundary Element (BEM) methods are widely used for modeling deformable objects [12], including soft tissues for real time surgical simulation.

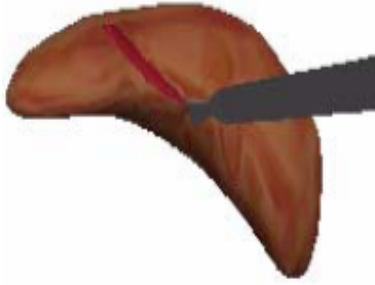


Fig. 4. Line-triangle intersection tests determine collision between virtual tool and deformable organ.



Fig. 6. Particle based simulation of bleeding after cutting into an organ.

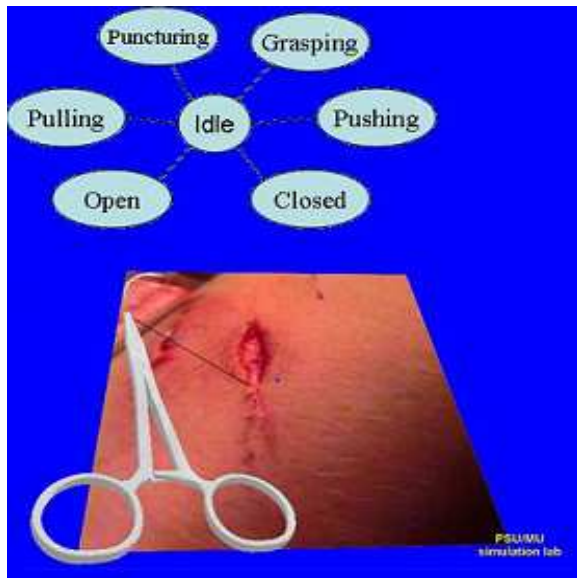


Fig. 5. Screen shot of finite state machine with a subset of states as shown in [11].

D. Artificial Intelligence

Artificial Intelligence is another major component of game programming. Four main AI constructs are most commonly used in games: state machines, rule systems, state-space searches, and genetic algorithms. Finite state machines have been used in [7] to analyze surgical motions. The methods described in [13] were modified to develop a finite state machine model with states and transitions. States included needle grasping, needle pushing and pulling, and skin puncturing. Fig. 5 depicts the finite state machine model developed in [7]. An analysis of the extracted data measures state-to-state transitions and calculates the amount of time spent by the user at each state. Path planning involves computing a collision-free route between two points in virtual space. The focus of path planning in surgical simulations is to assist the user in obtaining the best suitable start point. It can also be used to guide the tool trajectory. Several approaches to solve path planning use the class A^* algorithm [14].

Rule-based expert-like systems use inference procedures with uncertainty since the knowledge provided by the human

expert is imprecise. Fuzzy logic utilizes effective tools to handle and manipulate such data, and to make decisions based on them. There has been extensive research on applying fuzzy set theory in medical fields [15]. Fuzzy relations and inference rules are used to measure surgical competence by providing scores based on surgical decisions.

IV. GAME PRESENTATION

Game presentation explores the use of level-of-detail, shading, texture mapping, and particle systems towards improving realism in games. These issues involve applying texture techniques to the geometry and rendering realistic complex scenarios.

A. Particle systems

A particle system is a mathematical formulation to describe phenomenon representing a group of dynamic, time-dependant, unconnected mass points. Common particle systems used in games include smoke, fog, snowflakes, and birds. In a particle system, particles follow a cycle in which they originate, move, and die. Stochastic modeling can be used to control the behavior of these particles. This modeling is permissible because a cutting process is understood to exhibit some amount of stochastic behavior for tool-tissue pairs [16]. As shown in Fig. 6, surgical simulations can use particle systems to simulate bleeding and clotting after a cutting operation of the tool. The velocity of the blood flow should be based on studied values of blood vessels.

B. Shading

In game programming, shading represent processes involved in providing color to the pixels on the view-screen. Broadly, two processes constitute shading: illumination and textures. Illumination handles light and shadow effects while textures deal with properties of the surface.

1) *Illumination*: Illumination is used to add visual details to the virtual system. Lighting can be of two types: per-vertex or per-pixel. Selection of the lighting equation is based on the density of primitives in the virtual model and involves three components:



Fig. 7. Texture mapping of the eye.

- 1) Ambient: Provides base lighting to the scene by scattering in all directions.
- 2) Specular: Provides surface reflection along the direction of the mirror.
- 3) Diffuse: Provides reflections from all surfaces in the scene.

Normally, the Phong lighting model is used for surgical simulations. This model is developed for curved surfaces and calculates vertex-normal interpolations for each face of the virtual object.

2) *Texture Mapping*: A texture represents the material or surface properties of a virtual object. Texture mapping is based on a predefined bitmap consisting of texture pixels, or texels. This method consists of wrapping the image onto the object to create a realistic representation of the object in three-dimensional space. The mapping is done by establishing a correspondence between vertices in the geometry and texels in the map (see Fig. 7). The advantage is complexity reduction and rendering speedup, because just a single texel read is required for each pixel being displayed on the view screen. Texture mapping is used extensively in games to add realism to the scene.

Texture mapping can be used in surgical simulation to provide data or patient-specific views [17]. A three-dimensional texture array contains the original volume data. Dynamic local two-dimensional arrays contain the visible surface data.

V. CONCLUSION

It is generally accepted that similar to flight simulators for pilot training, surgical simulators have the potential to train medical personnel effectively. As these training procedures involve manipulating virtual organs through instruments, the utilization of gaming techniques seems beneficial. This paper has presented two basic approaches: logic techniques for performance tuning and improved user interaction; and presentation techniques for realistic rendering and adding scene complexity.

It can be seen that, several computer game development techniques can be adapted for surgical simulation. The choice and implementation of a technique must be based on requirements of the application. As described in this paper, the

medical field has seen rapid progress in applying these techniques to training simulations. Although technical challenges remain in creating realistic simulations, recent advances in hardware and computational capabilities have created exciting opportunities.

ACKNOWLEDGMENT

The authors would like to thank Steve Whitmore for coordinating the review of the paper and publishing it.

REFERENCES

- [1] C. Wagner, M.A. Schill, and R. Manner, "Intraocular surgery on a virtual eye," *Communications of the ACM* 45(7), 4549.
- [2] A. Simo, and M. Cavazza, "A virtual human for simulating cardiac emergencies," in *Proc. VSMM*, 2002, pp. 843-852.
- [3] B.D. Mann, B.M. Eidelson, S. G. Fukuchi, S. A. Nissman, S. Robertson, and L. Jardines, "The development of an interactive game-based tool for learning surgical management algorithms via computer," *The American Journal of Surgery* 183, 305-308.
- [4] D. Sanchez-Crespo, "Core Techniques and Algorithms in Game Programming," Indianapolis, Ind.: New Riders Education, 2004.
- [5] C.L. Moody, C. Baber, and T.N. Arvanitis, "The role of haptic feedback in the training and assessment of surgeons using a virtual environment" in *Proceedings of Euro Haptics* (Educational Technology Research Paper Series), ETRP12 (4151), 170173.
- [6] F. Tendick, et al., "A virtual environment test bed for training laparoscopic surgical skills," *Presence* 9(3), 2365.
- [7] R.W. Webster, D.I. Zimmerman, et al., "A prototype haptic suturing simulator," *Medicine meets virtual reality*, 2001.
- [8] B. Chang, and E.J. Colgate, "Real-time impulse-based simulation of rigid body systems for haptic display" in *Proceedings of the ASME International Mechanical Engineering Congress and Exhibition*, 1997, pp. 18.
- [9] C.A. Mendoza, and C. Laugier, "Realistic haptic rendering for highly deformable virtual objects," in *Proc. Virtual Reality Conference*, 2001, pp. 264-269.
- [10] K. Jung, Suvranu De, and A. S. Mandayam, "Computationally efficient techniques for real time surgical simulation with force feedback," in *Proceedings of the 10th Symposium On Haptic Interfaces For Virtual Envir. and Teleoperator Sysys (HAPTICS.02)*, 2002.
- [11] M. Lin, and S. Gottschalk, "Collision detection between geometric models: A survey," in *Proceedings of IMA Conference on Mathematics of Surfaces*, 1998.
- [12] K.J. Bathe, "Finite element procedures," NJ: Englewood cliffs, 1996.
- [13] J. Rosen, M. MacFarlane, C. Richards, B. Hannaford, and M. Sinanan, "Surgeon-tool force, torque signatures - evaluation of surgical skills in minimally invasive surgery," in *Proceedings of Medicine Meets Virtual Reality (MMVR)*, 1999.
- [14] C. Wurl, D. Heinrich, and H. Wm, "Parallel on-line motion planning for industrial robots," *The Third ASCE Specially Conference on Robotics for Challenging Environments, Robotics98*, 1998.
- [15] K.P. Adlassnig, and G. Kolarz, "Computer-assisted medical diagnosis using fuzzy subsets". *Approximate Reasoning in Decision Analysis*, pp. 219, 1982.
- [16] M. Mahvash, and V. Hayward, "HapticRendering of Cutting: A Fracture Mechanics Approach," *Haptics-e*, The Electronic Journal of Haptics Research, [Online], vol.2, 2001. Available: <http://www.haptics-e.org>
- [17] W. Lin, and R.A. Robb, "Dynamic volume texture mapping and model deformation for visually realistic surgical simulation". *Stud Health Technol. Inform.* 62:198-204, 1999.



Vidya Kotamraju received the BE degree in 2001 from Goa University. From 2002 to 2004, she worked at the National Institute of Oceanography, India. Since the fall of 2004, she has been pursuing her MAsC in Simon Fraser University and is currently associated with the Experimental Robotics Laboratory at the School of Engineering Science. Her research interests are distributed computing, human computer interaction, medical imaging, and geographic information systems.



Shahram Payandeh received the B.Sc. degree in 1981 from the University of Akron, and the MASc and PhD degrees in 1986 and 1990, respectively, from the University of Toronto. He is serving as a professor in the School of Engineering Science at Simon Fraser University. He has several publications to his credit. His research interests are in the general area of mechanics, modeling, and control of robotic (physical/virtual) mechanisms.



John Dill received the BA, MS, and PhD degrees in 1962, 1964 and 1970, respectively, from the University of British Columbia, North Carolina State University, and California Institute of Technology. He currently holds the position of a professor in the School of Engineering Science at Simon Fraser University. He has published one book and several papers in journals and conferences. His research interests are broadly in the area of computer graphics, user interfaces, scientific visualization, computer aided engineering and design.

Simulation of Laparoscopic Common Bile Duct Exploration

Qian Liu, John Dill, and Shahram Payandeh

Abstract—Currently, surgical simulation has drawn considerable attention as an alternative surgical training tool. In laparoscopic surgery, virtual environments for training manipulative skills are now well established as research areas. One such skill is laparoscopic cholecystectomy, which involves inserting a catheter into the cystic duct using a pair of laparoscopic graspers. The hardware components of the computer-based training environment include a two handed device and a computer. The two handed device is a training substitute for the laparoscopic instruments. The computer displays not only the virtual models of the instruments and the organs but also the interactions between them. By practicing with the system, surgeons can be trained to grasp and insert a flexible and freely moving catheter into the deformable cystic duct in the virtual environment.

Index Terms—Cholecystectomy, collision detection, laparoscopic surgery, virtual environment.

I. INTRODUCTION

Surgical training is traditionally performed using an apprenticeship model. Before trainees can undertake the real surgeries, they must practice on plastic models, animals, and cadavers. With electromechanical user interface devices, computer-based surgical simulations open new possibilities in surgical training, which offers many benefits compared to traditional training methods. First, it can not only provide different training scenarios easily, but also let the trainee practice on the same scenario as many times as needed. Second, the virtual training environment can objectively quantify surgical performance and simulate the result of an operation. Finally, no animal or patient is at risk of harm. In the laparoscopic surgery, the surgeon only opens small incisions in the patient's abdomen. Instruments, such as graspers, cautery hooks, and staplers, are inserted into the body through those small holes. The operating site is viewed through a laparoscope, which is also inserted through a small incision. Because the patient experiences less trauma and pain than normal surgery, the laparoscopic surgery is also called Minimally Invasive Surgery. The physical setting of a computer-based surgical training environment is quite similar to real laparoscopic surgery, which will help trainees easily and quickly transfer the skills gained in a virtual system to a real operating environment [1].

In this paper, we present a laparoscopic cholecystectomy training environment, which simulates the process of inserting a catheter into the cystic duct using a pair of graspers. We use a novel geometrical algorithm to accomplish simulating the interaction of the instrument and the duct. When being grasped, the tissue of the duct will be deformed by the shafts of the jaws. In previous papers, usually the tips of the jaws are the only part of the instrument to interact with a point on the duct. Our simulator supports three types of collision detection. The interaction of the models and the instrument take place between the grasper and duct, the duct and the catheter, as well as the grasper and the catheter. In addition, we use two methods to represent the common bile duct and the catheter, respectively, according to their different physical behavior. The duct is represented by a mass-spring model which is loaded by a Virtual Reality Modeling Language (VRML) file, and it can be partially deformed or totally moved according to its physical properties. The mesh model has both an inner-surface and an outer-surface to simulate the thickness of the duct. Because the behavior of catheter is less complex than the soft tissue (the duct in our case), we choose a particle-based model to simulate the catheter, which is less sophisticated than the mesh model. These techniques can be integrated into training environments for both open and laparoscopic surgery.

We have also refined a more realistic mesh model to simulate the instrument. This tentative approach to establish a microsurgical training environment will provoke further research in the area of virtual surgery training systems.

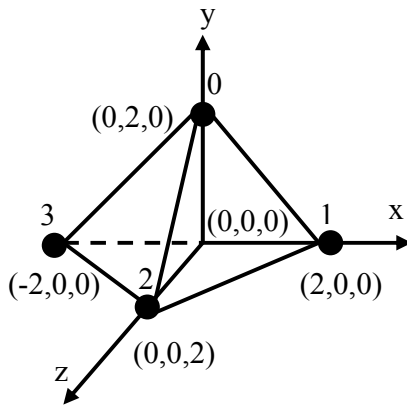
II. GEOMETRIC MODELS OF SURGICAL INSTRUMENTS AND ORGANS

II.1 Modeling and Representing the Duct

A. VRML

In this project, we chose the VRML file format as the means of representing geometric shapes external to our program. VRML is an open standard for virtual reality on the Internet. Our program reads the VRML text file and converts it into an internal data structure to make processing more efficient.

Though VRML models have many other properties, such as material and texture, we mainly use geometry and



```
#VRML V2.0 utf8
geometry IndexedFaceSet{
  coord Coordinate {
    point [
      0 2 0,
      2 0 0,
      0 0 2,
      -2 0 -0,
    ]
  }
  coordIndex [ 0, 20, 1, -1,
    0, 2, 1, -1,
    0, 1, 3, -1,
    0, 3, 2, -1,
    2, 3, 1, -1,
  ]
}
```

Fig. 1. Example of a simple VRML file of a tetrahedron.

topology information (i.e., the vertices' coordinates and how they are connected together). VRML models have an array of points, referred to as vertices in later descriptions. The "point" is defined by x, y and z coordinates. Then the model uses a "coordIndex" to define faces (planar polygons) as a set of indices to points in the vertex array. The last index of a face is "-1", to distinguish it from the next face.

Fig. 1 illustrates an example of a simple VRML model, which defines a tetrahedron. The number above each vertex is its index. For example, the first face of this tetrahedron is composed of points 0—(0,2,0), 2—(0,0,2) and 1—(2,0,0). The "face" in a "coordIndex" object in a VRML model is a polygon, such as a triangle, quadrangle, etc.

To simulate the duct, which is composed of two cylinders, we integrate two cylinder models (outer surface of the duct and inner surface of the duct) in one VRML model file. Here, the inner-surface of the duct is a non-uniform scaled model of the outer-surface. For collision detection of the instruments and the duct, we consider the outer surface of the duct. To detect the collision of the duct and catheter, we consider the inner surface of the duct and the catheter, which is represented by the particle system.

B. Data Structure

Six main classes are defined to represent the geometric object [2]. The basic underlying data structure is the STL (i.e., Standard Template Library) vector; it is easy to insert and delete nodes without changing the whole data array. Because STL has become a standard in C++, the algorithm

has been optimized. The access and operation speed is much faster than the data structure we defined ourselves.

1) CMesh

The whole object is represented by a mesh object. The *CMesh* class has the following main properties:

- vector<*CVertex**> —all the vertices of the object are stored in a vector
- vector<*CPolygon**> —all the polygonal facet of the object are stored in a vector
- vector<*CEdge**> —all the edges of the object are stored in a vector

2) CVertex

The *CVertex* class has some properties of a vertex on the object:

- *VECTOR mLoc*(Float x, y, z) —coordinate of this vertex
- vector<*CPolygon**> —pointers to all the incident polygon of this vertex
- vector<*CEdge**> —pointers to all the edges connected to this vertex
- *VECTOR mVertNormal* —vertex normal, average of its incident polygons' normal

3) CPolygon

The *CPolygon* here is the same as a triangle and face. It has the following properties:

- vector<*CVertex**> —pointers to vertices form this polygon
- vector<*CEdge**> —pointers to all its edges
- vector<*CPolygon**> —pointers to all the polygons incident to this polygon
- *VECTOR mPolyNormal* —polygon's normal

4) CEdge

Each edge is considered to be a line which connects the two end vertices. The properties are as follows:

- *CVertex* mV1, mV2* —pointers to the two end vertices defined this edge
- vector<*CPolygon**> —pointers to two incident polygons of this edge. If the edge is on the border, it just has one incident polygon

5) VECTOR

The *VECTOR* class is used to save the properties of the movement vector and calculate some geometrical equations: dot product, cross product, etc.:

- float x,y,z —normal of the vector
- float length —length of the vector

6) CInterV

The class of intersection vertex is used to save the properties of the intersection point:

- *VECTOR mLoc* —coordinate of this point

Due to the simplicity and low computational requirement, mass-spring models are commonly used to simulate deformable tissue. In mass-spring models, each vertex is a mass point (i.e., a node), each edge is a spring, and the edges form a 3-D mesh. Vertex displacements are calculated by solving differential equations that model the mass-spring

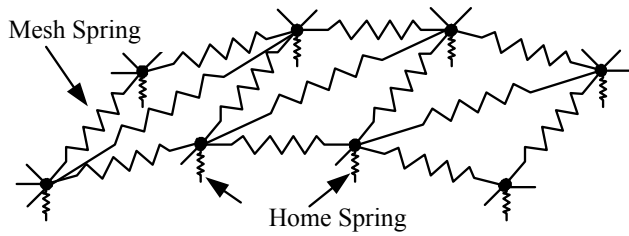


Fig. 2. Part of a mass-spring model.

system. Because we are only interested in the outer surface for display, we choose a surface mass-spring model to simulate the deformable virtual objects in this paper.

In our model, the simulated surface is divided into small triangles, where each vertex is a mass point. A linear spring is defined along each triangle edge. These springs are called “mesh springs” because they model the surface of the object. When a soft elastic object deforms, the interior of the object also contributes to the shape of deformation. To reflect this fact in our model, each node is also connected by a spring to its initial or “un-deformed home” position. We name these springs “home springs” define the internal structure for the object. The set of home positions comprises the rigid kernel of the deformable model, which preserves the object’s shape. [3] [4]. We also add damping to the mesh by adding each mass node a force proportional to the node’s velocity, but in the opposite direction.

Fig. 2 demonstrates part of the modeled surface. In this figure, home springs are drawn with nonzero length to depict their existence. Each edge can have a different spring constant and the nodes can have different masses, but the spring constants are not based on real material properties. For a homogeneous model, the spring constant for each spring should be inversely proportional to spring length. However, the virtual mesh will not have the same elasticity as real material because no well-accepted algorithm is available in literature to set spring constants. This mass-spring model seems better suited for surgical training, which requires more emphasis on visual realism than exact, patient-specific deformation, but requires performing simulations in real-time.

Model deformation results in vertex displacement. Due to spring forces, the displacement propagates to the neighboring area. To simulate the tension of soft tissue, every mesh spring has been initially stretched (i.e., its rest length is less than its initial length).

II.2 Modeling and Representing the Catheter

The catheter is modeled as a set of particles that were uniformly distributed along its centerline [5]. Each particle moves under the influence of internal and external forces. Each particle is connected to its neighbors through a network of springs and dampers.

In our particle-based model, we calculate the total force applied to each particle by using (1).

$$f_i = \begin{cases} f_{spring} = k_s \sum (l - l_0) \\ f_{bending} = k_b \sum (\theta - \theta_0) \\ f_{damping} = -k_d v \\ f_{external} \end{cases} \quad (1)$$

The parameters in (1) are:

- l distance between the particle and its neighbor;
- v velocity of the particle;
- θ relative angle between neighboring segments;
- k_s, k_b stiffness and bending coefficients respectively.
- k_d damping coefficient.

The acceleration (a_i), velocity (v_i), and position of each particle (x_i) are updated using the Euler numerical integration method.



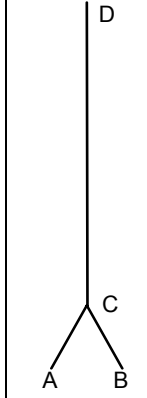
$$\begin{cases} \dot{x}_i = v_i \\ \dot{v}_i = a_i = f_i / m_i \\ x_i(t + \Delta t) = x_i(t) + v_i(t)\Delta t \\ v_i(t + \Delta t) = v_i(t) + a_i(t)\Delta t \end{cases} \quad (2)$$

The parameter m_i in (2) represents the mass of the particle.

III. COLLISION DETECTION

An important part of surgical instrument and soft tissue modeling relies upon the fast collision detection algorithms. During the real-time simulation, we check three collisions: 1) between the instruments and the duct 2) between the instruments and the catheter 3) between the catheter and the duct.

TABLE I
SURGICAL INSTRUMENT: GRASPING FORCEPS

Instrument Name	Real Instrument Picture	Virtual Instrument	Collision Detection Representation
Laparoscopic grasping forceps			

In order to detect the collisions between the catheter and the cystic duct, we modeled the catheter using a group of particles connected to each other with virtual springs and dampers along the catheter centerline. We defined a radius of influence for each particle to detect the collisions between the outer surface of the catheter and the inner surface of the duct in real-time.

For the purposes of graphical display, instruments are represented by three cylinders in order to give the instrument a realistic appearance. The third column in Table 1 is a screenshot of our virtual instrument in the simulator. For collision detection purposes, the instrument representation is simplified to a small set of line segments, as shown in the forth column of Table 1. Collision detection then reduces to detecting intersections between line segments and triangles. We will calculate the exact intersection point of the line segment and triangle.

A. Line-Segment and Triangle Intersection

The simple case of computing the intersection point between a line segment and a triangle is to determine the intersection between the line, on which the given line segment lies, and the plane that contains the triangle. To locate the intersection point, we first check whether the line and plane are parallel. If they are not parallel, there must be an intersection point. Then we need to check whether this intersection point lies along the line segment and lies within the triangle (Fig. 3).

$$p = p_1 + t(p_2 - p_1) \quad (3)$$

$$Ax + By + Cz + D = 0 \quad (4)$$

We represent the line by (1), where P_1 and P_2 are two known points on the line. The plane is represented by (2), where the values of A , B and C are the components of the plane normal, which can be found by taking the cross product of any two normalized edge vectors. However, for graphic rendering, we have already calculated the normal for every triangle. Therefore, we can easily acquire the values of A , B , and C for every triangle. D can be found by substituting one vertex into the equation for the plane as the following:

$$D = -(AP_{ax} + BP_{ay} + CP_{az}). \quad (5)$$

$$t = (D + AP_{1x} + BP_{1y} + CP_{1z}) / (A(P_{2x} - P_{1x}) + B(P_{2y} - P_{1y}) + C(P_{2z} - P_{1z})) \quad (6)$$

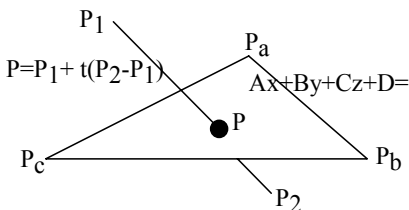


Fig. 3. Line segment and triangle intersection

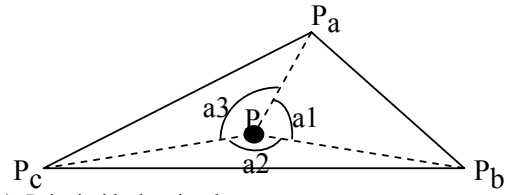


Fig. 4. Point inside the triangle.

To determine the intersection point p , we just need (4) to compute the value t . If the denominator above $A(P_{1x} - P_{2x}) + B(P_{1y} - P_{2y}) + C(P_{1z} - P_{2z})$ is zero, then the line is parallel to the plane and they do not intersect. To satisfy the condition that the intersection point to lie on the line segment, t must be between 0 and 1.

Then we need to check whether this intersection point is inside the triangle. If a point is inside a triangle, the sum of the internal angles is 2π , while a point outside the triangle has lower angle sums.

As illustrated in fig. 4, we want to test whether point P is inside triangle $P_aP_bP_c$. We can calculate the angles as the following:

$$\begin{cases} a1 = \text{acos}(P_{a1} \cdot P_{a2}) \\ a2 = \text{acos}(P_{a2} \cdot P_{a3}) \\ a3 = \text{acos}(P_{a3} \cdot P_{a1}) \end{cases} \quad (7)$$

Where

$$\begin{cases} P_{a1} = (P_a - P) / |(P_a - P)| \\ P_{a2} = (P_b - P) / |(P_b - P)| \\ P_{a3} = (P_c - P) / |(P_c - P)| \end{cases} \quad (8)$$

If the sum of $a1$, $a2$ and $a3$ is equal to 2π , point P is inside the triangle. When using the computer, we need consider the floating error. Therefore, we approximate this algorithm as the following:

```
sum = a1+a2+a3;
if(abs(sum-2*3.14) < eps)
    return true;
where eps is floating error we can accept.
```

IV. SOFT TISSUE DEFORMATION

A. Grasper Touches the Mesh

We have proposed a simple method to simulate the deformation when the grasper is touching the mesh model (i.e., the grasper jaw intersects the non-deformed mesh object at one point) [2]. We achieve this deformation by simply pushing the intersected triangle down. As shown in Fig. 5, when the instrument intersects surface triangle ABC , we simply move all the vertices of triangle ABC by the

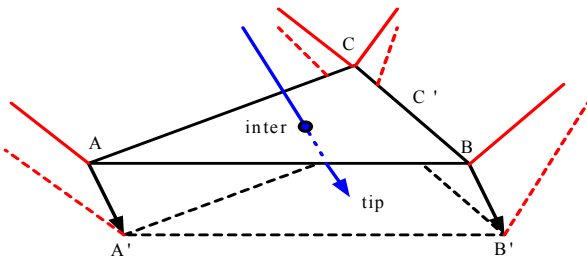


Fig. 5. Mesh surface deformation model, close look.

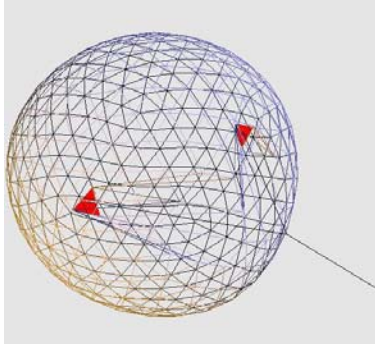


Fig. 6. Global view of the deformed sphere. The two red triangles are moved to the tips of the grasper.

vector \vec{I} from the intersection point P_{inter} to the current grasper tip P_{tip} , as if the instrument displaces the whole triangle down as a plane.

$$\vec{I} = P_{tip} - P_{inter} \quad (9)$$

B. Grasper Pinches the Mesh

In reality, when user employs the grasper to pinch the mesh model, the tissue will be grasped and deformed by the whole jaw, as well as the intersected points on the mesh [6]. In our project, when one grasper jaw intersects with two points on the mesh surface, the triangles outside the grasper jaw and between the two intersected points need to be moved along the grasper jaw (Fig. 7).

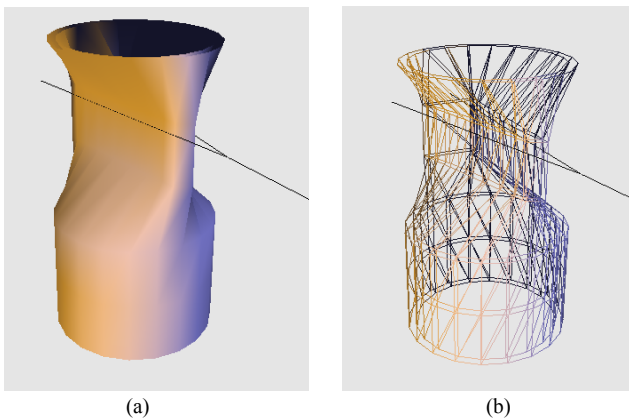


Fig. 7. Simulation of grasper pinching the mesh. (a)Surface mesh (b) Wire frame mesh. From the wire frame we can recognize that the triangles along the grasper jaws are all moved.

Here we propose a simplified approach to find all the triangles which need to be moved. The steps are as follows:

1. Get the parameters of the grasper jaw plane.
2. Search each triangle of the mesh object to mark all the triangles which should be deformed by the grasper jaws.

We judge that the triangle should be moved along grasper jaw AC (refer to Table 1) by two default assumptions: 1) all possible intersected triangles lie in the 3-D space $Z > 0$. 2) The angle between the triangle normal and the grasper jaw normal is less than 90 degree (i.e. the dot product of the triangle normal and grasper jaw normal > 0). Rather, the triangle needs to be moved to the grasper jaw BC if it satisfies two conditions: 1) the Z value of the triangle normal < 0 . 2) The angle between the triangle normal and the grasper jaw BC 's normal is less than 90 degree (i.e., the dot product of the triangle normal and the grasper normal < 0).

3. Project the triangles to the grasper jaw plane.

If the projected triangle intersects with the line segment of the grasper jaw, this triangle will be moved along the grasper jaw. The new vertices of the projected triangles are the new positions of the original triangle.

We determine that the line segment intersects with the projected triangle if the line segment intersects with two edges of the triangle. As shown in Fig. 8, line segment AC intersects with both edges P_aP_b and P_aP_c , so we say that line AC intersects with triangle $P_aP_bP_c$.

For example, if we want to move all the triangles which need to be deformed by grasper jaw AC , the pseudo codes are:

```

Find the grasper jaw plane
While not the end of the triangle on the mesh do
{
  If (the Z value of the triangle normal > 0)
  {
    If (Dot Product of the triangle normal and the
    grasper jaw normal > 0)
    {
      Project the triangle to the grasper jaw plane
      If (line segment AC intersects the projected
      triangle)
        Displace the vertices of the triangle to the
        vertices of the projected triangle
    }
  }
  Search next triangle
}
    
```

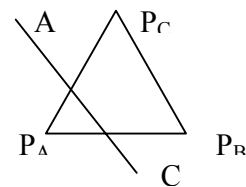


Fig. 8. Line segment AC and triangle $P_aP_bP_c$ intersect.

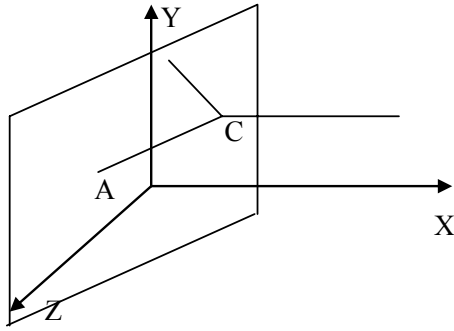


Fig. 9. Grasper jaw plane AC.

3.1 Find the planes where the grasper jaw AC and BC are respectively.

Fig. 9 defines the plane AC which is parallel to the Y axis and has line AC in it. This plane AC can be written as

$$Lx + Nz + D = 0 \quad (10)$$

The coordinates of line segment AC are defined as $A(Xa, Ya, Za)$, $C(Xc, Yc, Zc)$.

Set the uniformed normal of the plane to be L, M, and N.

$$\begin{cases} \sqrt{L^2 + N^2} = 1 \\ LXc + NZc + D = 0 \\ LXa + NZc + D = 0 \end{cases} \quad (11)$$

Parameter L, N and D can be calculated by using (9).

3.2 Project the triangle $V_0V_1V_2$ to the jaw plane (Fig. 10).

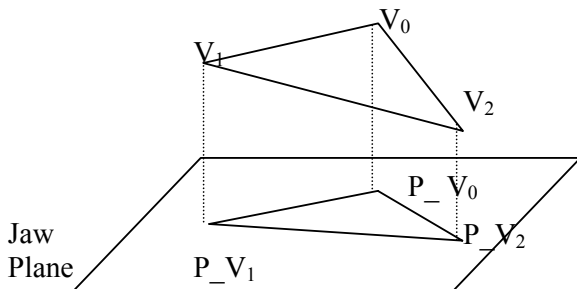
The coordinate of point V_0 is set to (X_0, Y_0, Z_0) , and it satisfies both (9) and (10). Equation (10) is how we formulate the line segment form V_0 to $P_{-}V_0$.

$$\frac{x - X_0}{L} = \frac{y - Y_0}{0} = \frac{z - Z_0}{N} = t \quad (12)$$

We can update the line parameter t from (11).

$$t = \frac{-L \times X_0 - N \times Z_0 - D}{L^2 + N^2} \quad (13)$$

Finally, the coordinate of $P_{-}V_0(x, y, z)$ can be computed by using (12).


 Fig. 10. Triangle $V_0V_1V_2$ projects to jaw plane.

$$\begin{cases} x = L \times t + X_0 \\ y = Y_0 \\ z = N \times t + Z_0 \end{cases} \quad (14)$$

3.3 Calculate the intersection point of two line segments.

The algorithm which we use to judge if the two line segments intersect or not is described as follows:

Let the end points of the line segments are P_1, P_2, Q_1 , and Q_2 respectively, the two line segments can be expressed parametrically as

$$P(t) = A + Bt \quad (15)$$

and

$$Q(s) = C + Ds \quad (16)$$

where

$$\begin{cases} A = P_1 \\ B = P_2 - P_1 \\ C = Q_1 \\ D = Q_2 - Q_1 \end{cases} \quad (17)$$

The intersection of line segment P_1P_2 and line segment Q_1Q_2 occurs when

1. $A + Bt = C + Ds$

2. Both t and s are between 0 and 1.

We can calculate the value of t, and s using the following equations:

$$t = -\frac{(C \times D) \bullet A}{(C \times D) \bullet B} \quad (18)$$

and

$$s = -\frac{(A \times B) \bullet C}{(A \times B) \bullet D} \quad (19)$$

V. CONCLUSION

In laparoscopic surgery and other surgical fields, virtual reality simulators have great potential to train surgeons for their manipulative skills [7]. In our paper, we introduce our approach to simulate the common bile duct exploration: grasping and inserting a flexible and freely moving catheter into the deformable cystic duct in the virtual training environment [5]. We represent the duct by a two layer mesh surface model and the catheter by the particle system. To simulate the mesh deformation, we detect the collisions between the grasper and the duct as well as the catheter; and between the duct and the catheter. This final project will be integrated into our Laparoscopic Training Environment (LTE) in Experimental Robotics Lab (ERL).

ACKNOWLEDGMENT

The authors would like to thank Dr. Peter Doris in Surrey Memorial Hospital to let us establish a training environment, and watch a real laparoscopic surgery. Author Q. Liu thanks her senior supervisors Dr. Shahram Payandeh and Dr. John Dill for their advice and support.

REFERENCES

- [1] A. Faraz, S. Payandeh, *Engineering approaches to mechanical and robotic design for minimally invasive surgery (MIS)*, Boston : Kluwer Academic, c2000.
- [2] H. Zhang, "Simulation Tissue Dissection for Surgical Training," M.A.Sc. Thesis, Dept. Eng. Sc., Simon Fraser University, Vancouver, B.C., 2004.
- [3] Meseure P., Chaillou C., "Deformable Body simulation with Adaptive Subdivision and Cuttings," Proceedings of the WSCG'97, Plzen, 10-14 February 1997, pp 361-370.
- [4] J. Zhang, S. Payandeh and J. Dill, "Haptic subdivision: an Approach to Defining Levels of Details in Haptic Rendering," Proceedings of the 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, Orlando, FL, pp201-208, Mar. 2002. C.
- [5] B. asdogan, C. Ho, M.A. Srinivasan, "Virtual Environments for Medical Training: Graphical and Haptic Simulation of Laparoscopic Common bile Duct Exploration," IEEE Transection, vol. 6, NO. 3, pp. 269-285, Sept. 2001.
- [6] C. Forest, H Delingette, N. Ayache, *Surface Contact and Reaction Force Models for Laparoscopic Simulation*, Epidaurre Research Project.
- [7] E. EreI, B. Aiyenibe, P.E.M. Butler, Microsurgery simulators in virtual reality: review. Wiley InterScience [Online]. Available: <http://www.interscience.wiley.com>



Qian Liu received the Bachelor's degree in electronic engineering from Southeast University, China, in 1998. After receiving the Bachelor's degree, she worked for Sample Technology CO. She participated in developing Electronic Police System, Criminal

Files Retrieving and Displaying System, WAN Digital Monitor and Alarm System, and Vehicle Surveillance and Record System. Currently, she is a Master student of Engineering Science at S.F.U. Her research focuses on computer graphics, real-time simulation, and virtual environments. Her thesis is to develop the computer-based surgical training system.

Increasing the power of functional maps at a cortical activation reorganization experiment

Behrang Nosrat Makouei

Abstract— Large deformation diffeomorphic metric mapping (LDDMM) of Beg (Beg, M.F. et al. 2004) is used to normalize the functional magnetic resonance imagery (fMRI) of cortical effects in learning a second language. We have used LDDMM to map structural scans of multiple subjects into extrinsic atlas coordinates; these same diffeomorphic mappings were used to transfer the corresponding functional data activations to the same extrinsic coordinates. By normalizing the functional images using LDDMM, anatomical variation across the subjects are removed and the statistical power in the averaged LDDMM mapped signals is significantly increased.

Index Terms— fMRI, image normalization, large deformation mapping, nonlinear alignment

I. INTRODUCTION

THE recent discovery that magnetic resonance imaging can be used to map changes in brain hemodynamics that correspond to mental operations extends traditional anatomical imaging to include maps of human brain function. The ability to observe both the structures and also which structures participate in specific functions is due to a new technique called functional magnetic resonance imaging (fMRI). fMRI provides high resolution, noninvasive reports of neural activity detected by a blood oxygen level dependent (BOLD) signal [10, 11]. BOLD signal is used to monitor the increase in oxygen consumption following local brain activation.

Recent advances in structural imagery, providing us with sub 1 mm scale images, has opened new horizons in intrinsic study of brain function, i.e. the comparison of functional responses based on within subject anatomical coordinate systems. Intrinsic study of brain functions is based on the fact that rigid-body transformations are able to establish the link between coarse resolution functional responses and the high resolution structural images [1-4]. Intrinsic approach is in sharp contrast with extrinsic atlas studies requiring the pooling of functional responses across anatomical coordinate systems.

Figure will be replaced later

Fig. 1. will be replaced soon

The low resolution and high baseline noise level of current functional images leaves no choice but averaging multiple images from multiple subjects to get a reasonable signal to noise ratio (SNR). The best approach for averaging cross-subject images is transforming all the functional images to the same extrinsic atlas coordinates. As depicted in figure 1, high resolution structural magnetic resonance (sMR) images play a crucial link in this transformation; transformations from sMR images to the atlas are computed and applied to the corresponding functional images to bring them to the same extrinsic coordinates (this process is usually called normalization).

The success of normalization in overlapping homologous areas is important for intersubject inferences; if functional areas are not well aligned between individuals, then common activated homologous regions might not be visible at results. Most of the conventional normalization algorithms use a mathematical measure of overall image mismatch and a minimization algorithm with iterative changes in transformations to find the best set of transformations for matching the image to the template. Usually, they begin with optimizing linear parameters: translations, rotations, zooms and often shears. Then, they proceed to find the best set of nonlinear (warping) parameters to further match the details of two brains [8, 12]. Our method of normalization, LDDMM, provides a geodesic flow in the space of diffeomorphisms between

anatomical configurations such that connected sets remain connected, disjoint sets remain disjoint, and submanifold structures are preserved. Preserving the connectivity relation between regions is particularly important for averaging functional data where the bijective property of the maps ensures that artifacts, due to superposition of functional data from neighboring regions, are avoided.

All individuals do not have the same brain microstructure; meaning that during minimizing the distance (error) between transformed image and the template, the algorithm might fall into local minima and the result of mapping become obscured and unreliable. To eliminate such errors, we first segment a region of interest (ROI) from all the included images and apply the normalization only to ROI's. Here, the ROI is deduced from our previous information about regions of brain activated during the underlying fMRI experiment [5]. To make the segmentations more accurate, all the images are first aligned to a common template through a rigid-body transformation.

In this paper, we demonstrate the superiority of LDDMM over conventional normalization methods (for the instance the one implemented in SPM2 software) on localizing the brain activation during investigating the cortical effects of learning a second language. The effectiveness is observed by quantifying the increased statistical power in the averaged LDDMM mapped signal. Although LDDMM is used to map a specific ROI across brains, the same method can be used to normalize other parts of brain as well.

II. METHODS

Both high resolution structural MRI and functional sequences were acquired from subjects. Local LDDMM and SPM2 methods for registration (normalization) were applied to optimize the alignment to increase the statistical power and precision of localization within the ROI. The generated transformations were then applied to the functional response map of each subject. For each subject, the mappings minimize a squared error registration distance between the observed imagery and the atlas.

A. Functional Method

Functional magnetic resonance imaging (fMRI) was employed to investigate the cortical effects of learning Mandarin as a second language by native speakers of American English. Specifically, the main goal of this study was to find out if the neural changes associated with learning Mandarin lexical tones occurred within known language-related areas and/or within new regions recruited for these functions. Relation between learning-related neural changes and centroid-shifts/variations in the volumes of activated cortical regions was also studied. Six college-age native speakers of American English who were beginning learners of Mandarin participated in the study.

The study involved two identical fMRI scans: an initial scan prior to training and a follow-up scan 2 weeks after training. Each scan consisted of baseline and stimulation periods. During the stimulation period, subjects performed a tone identification task in which they were required to identify 40 auditory presentations of Mandarin lexical items. Each word bore a tone that belonged to one of four possible tonal categories. In the same session, subjects also performed related control tasks with visual, auditory, and motor components. Results were separated from the lexical tone identification images and served to increase the specificity of the tone-training results. All brain images for all the subjects were computationally normalized to allow direct comparisons between pre- and post-training scans. Anatomical labels, Brodmann's areas, and atlas sectors were assigned to the active regions based on judgments of brain atlas correspondence [13].

Activation of each voxel was determined by a multistage statistical analysis that compared mean amplitude of signals acquired during stimulation and baseline periods. This technique also required statistically significant signal changes on two identical runs [14] resulting in an empirically determined false-positive rate of $p < .0001$ for each condition. To eliminate the nonspecific visual, motor, and auditory information in the tone identification runs, a "logical exclusion", AND NOT operation, was employed to segregate activity associated with the finger-tapping (motor), checkerboard (vision), and the tone sweep (auditory) functions from the lexical tone identification tasks, presumably yielding the activity unique to Mandarin lexical tone identification. Comparison of pre- and post-training images was based on aligned data sets. Two logical operations were performed on the thresholded images: [pre-training AND post-training] presumably reflecting regions activated for all conditions [15] and [post-training NOT pre-training], presumably reflecting activity unique to the post-learning state.

B. LDDMM

The LDDMM algorithm computes a transformation $\varphi : \Omega \rightarrow \Omega$ where $\Omega \subseteq \square^3$ is the 3D cube on which the data (structural and functional) are defined. The transformation computed by LDDMM is the end point $\varphi = \phi_1$ of a flow of vector fields $v_t \in V, t \in [0, 1]$ given by the ordinary differential equation $\dot{\phi}_t = v_t(\phi_t)$, where ϕ_0 is identity $\phi_0(x) = x, x \in \Omega$. Enforcing smoothness on the vector fields $v \in V$ ensures that the solution to the differential equation $\dot{\phi}_t = v_t(\phi_t), t \in [0, 1]$ is in the space of diffeomorphisms [6]. Smoothness is enforced by constraining the L^2 norm of several derivatives of the vector field to be in a Sobolev space with norm-square

$\|f\|_V^2$ [7]. The optimal φ is the minimum of the endpoint of the inexact matching problem

$$\inf_{\varphi=\phi_1; \phi_1(x)=\int_0^1 v_t(\phi_t) dt+x, v \in V} \int_0^1 \|v_t\|_V^2 + \lambda D_\varphi, \quad (1)$$

where D_φ is designed on the ROI or the MRI imagery, and $\lambda > 0$ is the relative weight assignment. We term the mapping based on the ROI segmentations the ROI-LDDMM algorithm.

Here ROI-LDDMM was applied to both ROI hand segmentations of MRI functional and structural images. In order to help the algorithm on initial estimation of parameters, several landmarks were put at ROI segmentations. Note that, as the complexity of the 3D structure increases, the benefit of the high dimensional transformation is further realized.

C. SPM2 based normalization

The normalization algorithm implemented in SPM2 [8] takes a two-step iterative parameter search approach. The first step of registering images from different subjects is determining the optimum 12 parameter affine transformation. Prior knowledge of the variability of head sizes is included within a Bayesian framework in order to increase the robustness and accuracy of the method. The second step is a nonlinear registration for correcting gross differences in head shapes that cannot be accounted for by the affine normalization alone. The nonlinear warps are modeled by linear combinations of smooth discrete cosine transformation basis functions (approximately 1000 lowest frequency ones). The coefficients of this basis function set are found by a Maximum A posteriori (MAP) parameter estimation.

The model for defining nonlinear warps uses deformations consisting of linear combinations of low-frequency periodic basis functions. The spatial transformation from coordinates x_i to coordinates y_i is:

$$\begin{aligned} y_{1i} &= x_{1i} + u_{1i} = x_{1i} + \sum_j q_{j1} d_j(x_i) \\ y_{2i} &= x_{2i} + u_{2i} = x_{2i} + \sum_j q_{j2} d_j(x_i), \quad (2) \\ y_{3i} &= x_{3i} + u_{3i} = x_{3i} + \sum_j q_{j3} d_j(x_i) \end{aligned}$$

where q_{jk} is the j^{th} coefficient for dimension k , and $d_j(x)$ is the j^{th} basis function at position \vec{x} . MAP parameter estimation tries to find the best set of q_{jk} 's based on the least-square Euclidian distance measure. As for affine registration, the optimization minimizes the sum of squared differences between a source (f) and the

template image (g). The images may be scaled differently, so an additional parameter (w) is needed to accommodate this difference. The minimized function is then:

$$\sum_i (f(y_i) - wg(x_i))^2. \quad (3)$$

The MAP parameter estimation evaluates the optimum coefficients for a set of basis functions: it minimizes the sum of squared differences between the template and source image, while simultaneously minimizing the deviation of the transformation from its expected value. In order to adopt the MAP approach, we need estimations of the likelihood of obtaining the fit given the data, which requires prior knowledge of spatial variability, and also knowledge of the variance associated with each observation (voxels of the image). True Bayesian approaches assume that the variance associated with each voxel is already known, whereas the approach used in SPM2 is a type of Empirical Bayesian method, which attempts to estimate this variance from the residual errors. Because the registration is based on smooth images, correlations between neighboring voxels are considered when estimating the variance. The same approach is used for the spatial normalization of both high quality MR images, and low resolution noisy fMRI images.

III. RESULTS

We select the structural MRI image of one of the patients as template. In order to make the induction on activated regions more straightforward, the template was first aligned to the Talairach & Tournoux brain atlas. All the functional images were upsampled to the size of the structural template image; this upsampling is itself a method of smoothing which is crucial for fMRI data analysis [9].

Figure will be replaced later

Fig. 2. Errors between the template and the transformed images using LDDMM and SPM2 normalization methods

Based on prior knowledge of the brain regions activated through underlying experiment [5], ROI's were selected at all of the functional and structural images. For each subject, the fluid match transformation (LDDMM) mapping the ROI of the corresponding structural image to the ROI of the template was calculated. The same transformations (mappings) were used on ROI's of the

Figure will be replaced later

Fig. 3. Calculated mean images of functional sequences after normalizing using LDDMM and SPM2 methods

relating functional images.

Figure 2 shows the comparison between the accuracy of LDDMM and SPM2 normalizations. Figure 3 shows an alternative visualization of the improved alignment accuracy obtained by the LDDMM method. Depicted is the comparison between results of averaging the ROI's of 3 participants' 171 functional images. LDDMM yields a clearer and more detailed mean image where SPM2's approach results in a relatively blurred image. The results indicate that the higher local dimensions of the LDDMM provide increased power in terms of normalization accuracy.

Based on the Box-Car approach, data was fit to a model and activated regions were estimated based on a t-test ($p < 0.001$) [5]. Figure 4 depicts the estimated activated regions and the

corresponding number of involved voxels.

IV. CONCLUSION

The current low SNR of functional images makes it imperative to use averaging of functional data to obtain clearer signal that represent activation of the individual structures. The most important problem on functional data averaging is the large variability of human brain anatomical structures. Methods to remove anatomical variability via representation of average information in extrinsic atlas

Figure will be replaced later

Fig. 4. will be replaced soon

coordinates require the link of high resolution structural imagery with the intrinsic within-coordinate fMRI scans. Common low dimensional transformation methods are limited in their ability to account for local variability in size, shape and location of brain microstructure. Increasing the dimension of the mappings to extrinsic atlas coordinates via LDDMM increases the accuracy in localization of brain responses and therefore increases the

statistical power.

In this paper we compared LDDMM with a widely used normalization method implemented in the SPM2 software. To increase the precision of both transformations, a ROI was selected and transformations were applied to that region only. Same t-test (and p-value) was used to calculate the superimposed activation on the sMR images. Comparing the activation maps resulted from ROI-LDDMM and ROI-SPM2 normalizations, LDDMM based approach demonstrates its superiority in three aspects:

- 1) *A greater number of significant regions*
- 2) *Higher magnitude (higher peak) of activity within regions*
- 3) *Higher SNR of activity (smaller standard error of the mean) within regions*

Improved cross-participant alignment resulted in a less variable and more robust response in the group analysis. This change in the variability was statistically reliable in an analysis of cross-participant t-tests, not isolated to the best voxel, but general in nature.

The LDDMM employs the most detailed transformations of human anatomy possible (with highest DOF), essentially calculating the transformation between two dense continuums attempting to transform the geometry of individual structures into the common extrinsic coordinates. We notice, however, that at least two sources of variability or inaccuracy are not addressed in the current approach. First, the functional neuroanatomy across participants may vary and the precise localization of activities may differ between any two brains. Typical solutions to this problem (for the instance, blurring of the functional data or using independent "localizer" or "reference" scans) will only be quantitatively addressable via methods that incorporate the improved structural alignment presented here. Second, distortions of the mapping between the functional coordinate system and the structural coordinate system are not addressed directly. Thus, deformations of the brain in functional datasets arising from localized magnetic inhomogeneity, or mislocalizations arising from large draining vein signals are not corrected by the current approach. However, advances in neuroimaging techniques that will ameliorate these difficulties will still face the anatomical variability addressed here.

REFERENCES

- [1] Maurer, C. R and Fitzpatrick, J. M. "A review of medical image registration" ed. R.J. Maciunas, R. J. pp. 17-44, 1993.
- [2] J. B. A. Maintz, H. W. M and Viergever, M. A. *SPIE Medical Imaging, Image Processing*, p. 144154, 1998.
- [3] Wells, W, Grimson, E, Kikinis, R, and Jolesz, F. *IEEE Transactions on Medical Imaging*, 15, 429-442, 1996.
- [4] Maes, F, Collignon, A, Vandermeulen, D, Marchal, G, and Suetens, P. *IEEE Trans. Medical Imaging*, p. 187198, 1997.
- [5] Wang, Y, Sereno, J. A., Jongman, A, and Hirsch, J., "fMRI Evidence for Cortical Modification during Learning of Mandarin Lexical Tone", *Journal of Cognitive Neuroscience* 15:7, pp. 1019-1027.

- [6] Dupuis, P, Grenander, U, and Miller, M. I. *Quart. App. Math.* 56, 587-600, 1998.
- [7] Beg, M. F, Miller, M. I, Troune, A, and Younes, L. *Int. Journal of Computer. Vision* 61, 139-157, 2004.
- [8] Ashburner, J and Friston , K.J, “Spatial normalization using basis functions”, In R.S.J. Frackowiak, K.J. Friston, C. Frith, *et. al.*, editors, *Human Brain Function*. Academic Press, 2nd edition, 2003.
- [9] Poline, J.-B. , Friston, K. J., Worsley, K. J, and Frackowiak, R. S. J., “Estimating smoothness in statistical parametric maps: Confidence intervals on p-values”, *Journal of Computer Assisted Tomography*, 19(5):788-796, 1995.
- [10] Ogawa, S., Lee, T. M., Kay, A. R., Tank, D. W. “Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields”, *Magn Reson Med* ,14:68—78, 1998.
- [11] Belliveau, J. *et al.* “Functional mapping of the human visual cortex by magnetic resonance imaging”. *Science* **254**, 716–719, 1991.
- [12] Woods, R.P., Mega, M.S., Thompson, P.M. “Use of Automated Polynomial Warping to Create an MRI Atlas Specific for the Study of Alzheimer’s Disease”, *Annals of Neurology* 44(3):449, 1998.
- [13] J. Talairach and P. Tournoux. “Co-Planar Stereotaxic Atlas of the Human Brain: 3-Dimensional Proportional System -- An Approach to Cerebral Imaging”, *Thieme Medical Publishers*, New York, NY, 1st edition, January 1988.
- [14] Kim, K. H. S., Relkin, N. R., Lee, K.-M., and Hirsch, J. “Distinct Cortical Areas Associated with Native and Second Languages” *Nature*, 388, 171-174, 10 July 1997.
- [15] Hirsch, J., R-Moreno, D., and Kim, K. H. S. “Interconnected Large-Scale Systems for Three Fundamental Cognitive Tasks Revealed by Functional MRI”, *Journal of Cognitive Neuroscience*, 13 (3), 389-405, March 2001.



Behrang Nosrat Makouei was born in Tehran, Iran on April 12, 1982. He received at B.A.Sc degree in Electrical Engineering from Sharif University of Technology, Tehran, Iran in 2004. He was a research assistant at Image Processing Lab at Sharif University of Technology from 2003 to 2004. He is

currently pursuing his M.A.Sc degree at Simon Fraser University, BC, Canada as a research assistant at Medical Image Analysis Lab. His research interests include medical image and communications signal processing.

TCP Performance in Wireless Networks: A Survey of Existing Solutions

Modupe O. Omueti, *Student Member, IEEE*

Abstract—TCP is a connection-oriented protocol that provides reliable transport layer service for Internet applications, and it is designed primarily for wired networks. The Internet has witnessed remarkable growth of wireless IP communications, for which TCP provides transport service. Because wireless networks exhibit characteristics which violate the key assumption in the design of TCP, its performance degrades in wireless networks. Therefore, improving the performance of TCP in wireless networks is important. Several solutions have been proposed based on different criteria. We present different classes of solution to TCP implementation in wireless and heterogeneous networks, along with their advantages and disadvantages. In addition, we identify characteristics of an ideal solution for implementing TCP in wireless and heterogeneous networks. Moreover, we compare the proposed solutions with an ideal solution, and conclude that none of the proposed solutions has all the characteristics of an ideal solution.

Index Terms—TCP, packet loss, congestion control, wireless networks, heterogeneous networks.

I. INTRODUCTION

The Internet provides information exchange among a wide range of clients and servers, and across various networks from wired to wireless. The Transmission Control Protocol (TCP) and the Internet Protocol (IP) are two separately designed protocols that specify the rules of communication between end hosts on the Internet. TCP is an Open Systems Interconnection (OSI) transport layer protocol, while IP is a network layer protocol. TCP and IP are the dominant protocols used on the Internet today, and TCP has become the de facto standard used for most Internet applications such as Telnet, HTTP, and FTP [1]. The Internet has witnessed a dramatic growth in wireless applications, and an increase in wireless IP communications for which TCP provides transport layer service. Therefore, TCP must be able to support these Internet applications in both wired and wireless networks.

Wired and wireless networks have dissimilar characteristics. Originally designed primarily for wired networks, TCP's performance degrades in wireless and heterogeneous networks. Wired networks are characterized by a negligible random bit

error rate (BER) — a key assumption in the design of TCP. Hence, packet loss is attributed mainly to congestion in wired networks. Congestion avoidance and control algorithms were developed in [2] and their use with TCP was standardized in [3] to reduce packet loss due to congestion, enabling TCP to perform well in wired networks. However, packet losses occur more frequently in wireless networks due to their peculiar characteristics and not due to congestion. If packet losses are non-congestion related, invoking congestion avoidance and control algorithms decreases TCP throughput; hence, TCP performance degrades.

Standard versions of TCP include TCP Tahoe [4], TCP Reno [5], and TCP NewReno [6], which differ in their implementation of congestion control algorithms. Tahoe employs slow start, congestion avoidance, and fast retransmit congestion control algorithms. TCP Reno adds the fast recovery mechanism to TCP Tahoe, while TCP NewReno enhances TCP Reno by modifying the fast recovery mechanism and addressing the issue of partial acknowledgements (ACKs).

In section two of this article, we present the TCP Reno implementation of the congestion control algorithms and round-trip time (RTT) estimation. Section 3 describes wireless networks and various features they exhibit which degrade the performance of standard TCP. In section four, we describe various classes of proposed solutions along with their pros and cons. In section five, we propose characteristics of an ideal solution for implementing TCP in wireless networks, and compare the proposed solutions with an ideal solution. Our conclusions are presented in section six.

II. TRANSMISSION CONTROL PROTOCOL

TCP is a reliable connection-oriented end-to-end protocol that provides basic bi-directional transfer of a continuous stream of octets between hosts. It is implemented with mechanisms for connection management, in-order delivery, flow control, congestion avoidance and congestion control. We describe the congestion control algorithms and round-trip time (RTT) estimation, which are key factors in the performance of the Transmission Control Protocol.

A. TCP Congestion Control Algorithms

TCP Reno is one of the most widely adopted TCP versions with the four congestion control algorithms: slow start, congestion avoidance, fast retransmit, and fast recovery. These algorithms are based on sliding window and additive increase

Manuscript received June 16, 2005.

Modupe O. Omueti is with the School of Engineering Science, Simon Fraser University, Burnaby, BC, V5A 1S6 Canada (e-mail: momueti@cs.sfu.ca).

multiplicative decrease (AIMD) [7] algorithms. The congestion window (wnd) and the receiver window ($rwnd$) are two variables added to the TCP per-connection state to implement these algorithms. The wnd is the maximum number of bytes the sender may send before receiving an ACK, while the $rwnd$ is the maximum number of bytes the receiver may receive. The minimum of wnd and $rwnd$ determines the amount of data sent into the network. Another state variable, the slow start threshold ($ssthresh$), is utilized in determining if the slow start or congestion avoidance algorithm is used to control how data is transmitted in the network. The initial value of $ssthresh$ may be arbitrarily high [5], but it can be adjusted in response to congestion.

1) *Slow Start*: After the three-way handshake is completed, the slow start algorithm is used by TCP for probing a network gradually at the start of transmission to determine its capacity. The initial value of the sender's wnd at this stage (i.e., initial window), must be less than or equal to twice the sender maximum segment size ($SMSS$) [5]. During slow start, TCP increments the wnd by at most $SMSS$ bytes for each ACK received that acknowledges new data. The slow start algorithm is used as long as $wnd < ssthresh$, but when $wnd > ssthresh$ the congestion avoidance algorithm is used. When wnd and $ssthresh$ are equal, either slow start or congestion avoidance can be used. In addition, slow start is used after packet loss is detected by the retransmission timeout (RTO) timer mechanism — an indication of congestion. Hence, the TCP sender reduces the size of its wnd to half and linearly increases it as in congestion avoidance.

2) *Congestion Avoidance*: Congestion avoidance is used to probe the network more slowly than in slow start. TCP Reno employs a flow control mechanism based on sliding window, allowing the wnd to be increased linearly by one full-size segment per RTT in the congestion avoidance phase. Congestion avoidance ends when congestion is detected (a packet loss occurs). When a packet loss occurs, either the sender receives duplicate ACKs (DUPACKs) or the sender's RTO timer expires, activating fast retransmit and fast recovery.

3) *Fast Retransmit*: Packet loss is attributed to network congestion in wired networks. The TCP sender uses fast retransmit to detect and repair data packet loss based on incoming DUPACKs. TCP assigns a unique sequence number ($seqno$) to each transmitted data packet. When a packet loss occurs, the TCP receiver issues a DUPACK for any out-of-order data packet received. The fast retransmit algorithm uses the arrival of three DUPACKs as an indication of packet loss and causes the TCP sender to retransmit the lost data packet without waiting for the (RTO) timer to expire. TCP uses the sender's RTO timer when the network cannot trigger a threshold of three DUPACKs: when the $wnd < 4$ or when the network is temporarily disconnected. In the absence of the three DUPACKs, if the RTO timer expires, TCP assumes data packet loss and retransmits the lost data packet. The value of $ssthresh$ is also set to:

$$ssthresh = \max(FlightSize / 2, 2 \times SMSS), \quad (1)$$

where $FlightSize$ is the total outstanding data in the network.

4) *Fast Recovery*: In fast recovery, after the lost data packet is retransmitted, the wnd is set to $ssthresh + 3 \times SMSS$. Thus, the wnd is inflated by the number of segments that have left the network (three DUPACKs), which the receiver has buffered. For each additional DUPACK received, the wnd is incremented by $SMSS$ to reflect the additional segment that has left the network. The new value of wnd and the $rwnd$ may allow the transmission of a new segment. When the next ACK that acknowledges a newly transmitted segment arrives, the sender sets the wnd to $ssthresh$ to deflate the wnd after which TCP resumes the congestion avoidance phase. The TCP Reno congestion algorithms are shown in Fig. 1.

B. Estimation of RTT and RTO

The TCP sender transmits a data segment with a particular $seqno$ and expects an ACK which covers that $seqno$. If the RTO expires before the expected ACK is received, the TCP sender retransmits the data segment and reinitializes the RTO timer to resynchronize transmission. If the calculated value of RTO is too large, timeouts are not detected quickly and bandwidth is wasted. In addition, if the calculated RTO is too small, it causes unnecessary retransmissions, which can lead to congestion and, again, wasted bandwidth.

To compute the retransmission timeout (RTO), the TCP sender maintains two variables: smoothed RTT ($srtt$), which is the moving average of RTT and RTT variation ($rttvar$). Karn's algorithm [8] is used to estimate the value of RTT by taking RTT samples ($sampleRTT$) from data segments that are not retransmitted. Karn's algorithm eliminates retransmission ambiguity. This ambiguity arises when retransmissions have been sent out and, when the ACK is received, it is not clear which packet (original or retransmission) is being acknowledged. The values of $srtt$ and $rttvar$ are computed as:

$$rttvar = (1 - \beta) \times rttvar + \beta \times |sampleRTT - srtt| \quad (2)$$

$$srtt = (1 - \alpha) \times srtt + \alpha \times sampleRTT, \quad (3)$$

with recommended parameter values $\alpha = 0.125$ and $\beta = 0.25$ [3]. The value of $srtt$ used in (2) is its value before it is updated in (3); hence, the values are calculated in the above order [9]. Then RTO is computed as:

$$RTO = srtt + \max(G, 4 \times rttvar), \quad (4)$$

where G is clock granularity in seconds [9].

III. WIRELESS NETWORKS

Wireless networks provide connectivity for users to access information anytime and anywhere through wireless links. Users have access to traditional computing systems and applications such as the Internet and the World Wide Web (WWW) respectively, by using wireless enabled laptops and small screen devices. Therefore, growth in the deployment of wireless networks has been remarkable. From an application point of view, different kinds of wireless networks exist, and each has its own peculiar characteristics.

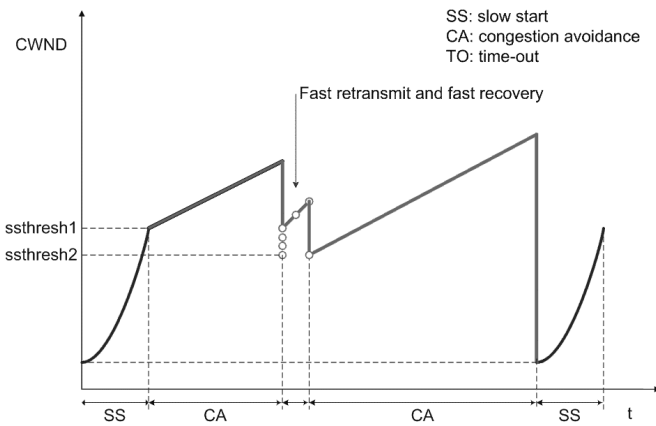


Fig. 1. TCP congestion control algorithms.

A. Types of Wireless Networks

We can broadly classify wireless networks into cellular networks, satellite networks and ad hoc networks. We describe these classes of wireless networks.

1) *Cellular Networks*: In cellular networks, a fixed host (FH) is connected to a mobile host (MH) by a base station (BS). The MH has lower power consumption and processing capacity than the base station, which tends to introduce inefficiency to the network. The MH changes location from time to time necessitating handoff between base stations. Cellular networks have evolved from 1G to 2G to 2.5G to 3G. Examples of 2G networks are Global System for Mobile communication (GSM) and Packet Data Cellular (PDC) networks. General Packet Radio Service (GPRS) is a 2.5G network while wideband code division multiple access (CDMA) and cdma2000 are 3G networks. With this evolution has come an increase in uplink and downlink data rates and voice capacity for users. Cellular networks are an example of heterogeneous networks.

2) *Satellite Networks*: Satellite networks have a satellite link between the sender and the receiver, and the satellites are at a great altitude above the earth's surface. The great altitude causes the inherently long delays experienced in the delivery of messages over satellite links. Messages travel over long distances causing their signal strength to diminish, and leading to a low SNR. Consequently, satellite links have a high occurrence of errors in transmission.

3) *Ad hoc Networks*: Ad hoc networks also known as peer-to-peer networks are set up without wireless infrastructure. Wireless hosts on ad hoc networks communicate directly with each other without making use of an access point (AP) or a wired network connection. The hosts are usually within radio distance from each other, change routes frequently and partition the network. Hence, packet losses occur, which are not due to congestion.

B. Characteristics of Wireless Links

Wireless links use radio or infrared waves as a means of transmission. Because of the peculiarities of these media, wireless links exhibit intrinsic characteristics that affect the performance of transport layer protocols such as TCP in

wireless networks. We discuss these characteristics.

1) *High Bit Error Rate (BER)*: Wireless links use radio as the medium of transmission; hence, they are subject to weather conditions, and multi-path interference. Due to limited radio coverage and mobility of the wireless end devices, wireless links experience temporary disconnections and reconnections. As a result, wireless links exhibit significantly higher BER than wired networks. The high BER results in a higher number of packet losses in wireless and heterogeneous networks than in wired networks.

2) *Large and Varying Latency*: Wireless links have a large and varying latency or propagation delay. A typical round trip time (RTT) of 2.5G/3G links varies between a few hundred milliseconds to one second [10]. Satellite links exhibit variable RTT due to long and varying propagation delays. A sudden increase in the latency of the propagation path is known as a delay spike [10] and 2.5G/3G links experience delay spikes exceeding the typical RTT. Delay spikes are caused by link recovery from temporal loss of radio coverage, handoff, and traffic re-routing by wireless networks. Delay spikes can cause spurious timeouts and unnecessary retransmissions leading to a multiplicative decrease in the *cwnd* size.

3) *Bandwidth Oscillation*: The RF spectrum is limited; hence, the amount of bandwidth available to all types of wireless networks is restricted. Point-to-point satellite networks have 6GHz (uplink) and 4GHz (downlink) [11], and can accommodate a limited number of mobile channels. Cellular networks have high data rate needs (10-20kps uplink and 10-40kbps downlink for 2.5G, 64kps uplink and 384kps downlink for 3G), which require that the limited spectrum be shared dynamically based on the number of concurrent users. As a result, RF resources are repeatedly allocated and de-allocated, which leads to varying bandwidth and data rates available to users. The periodic allocation and de-allocation of RF resources is referred to as bandwidth oscillation [10]. Spurious retransmission, which leads to TCP throughput degradation, is an effect of bandwidth oscillation.

4) *Path Asymmetry*: Wireless networks employ asymmetric uplink and downlink data rates. In cellular networks, this path asymmetry is due to connections between a fast FH and a slow MH. Furthermore, in satellite networks path asymmetry is due to the expense of the equipment used to send data to satellites. Delayed acknowledgements (ACKs) in the slow direction can limit throughput in the fast direction.

IV. SOLUTIONS TO TCP IMPLEMENTATION IN WIRELESS NETWORKS

The fundamental problem of TCP implementation in wireless networks is the misinterpretation of packet losses as an indication of network congestion. Hence, a standard TCP like TCP Reno reacts to packet loss by reducing the size of its *cwnd*. Packet losses are more frequent in wireless networks due to their peculiar characteristics. Analytical results [1] show that a standard TCP like Reno can recover from infrequent

packet losses in one RTT with its fast retransmit and fast recovery mechanisms. However, multiple packet losses occur in short bursts due to the characteristics of wireless links so the probability of their occurrence in one RTT increases. For multiple packets losses in one RTT, the TCP sender halts its transmission due to the retransmission timer's exponential back off. If the multiple packet losses are not due to congestion, the halving of the $cwnd$ and halting of transmission lead to a decrease in TCP throughput.

TCP is implemented in wireless or heterogeneous networks based on a number of considerations. Proposed solutions tackle the issue of misinterpretation of the cause of packet losses, which is common to all wireless or heterogeneous networks. Based on design, TCP is classified as split or end-to-end mode, in which the heterogeneous network is divided into wired and wireless sections. The end-to-end approach can be reactive or proactive. Other proposed solutions target particular wireless networks and address problems specific to such wireless networks. A classification of proposed solutions based on loss responsive flows and implemented at different layers of the OSI reference model is in [12, 13]. We discuss different classes of proposed solutions.

A. Design-based Solutions

Solutions to TCP implementation in wireless networks based on design employ the basis that wired and wireless networks differ significantly in characteristics. The assumption is that wired networks links are more reliable than wireless links in terms of error rate and capacity.

1) *Split Mode Solutions*: The split mode solutions separate the network into wired and wireless sections by terminating the TCP connection prior to the wireless link at a base station or access point. Hence, the split mode solutions shield the wireless links from the wired or fixed network, thus the wireless link events has the least impact on the fixed network. The intermediate node (BS or AP) is a terminal for both wired and wireless sections of the network. End hosts of the wired section communicate separately with the intermediate node with no knowledge of the wireless end hosts and vice versa. Figure 2 illustrates a split mode scheme.

Indirect TCP (I-TCP) [14] is an example of a split mode solution. The connection between the mobile host (MH) and fixed host (FH) is split into two — one between the MH and its mobile support router (MSR) over the wireless link and the other between the MSR and the FH over the fixed or wired network. The MSR communicates with the FH on behalf of the MH and vice versa, and I-TCP maintains the congestion window ($cwnd$) separately for both wired and wireless connections. The center point of the connection moves to the new mobile support router (MSR) when the MH switches cells during the span of an I-TCP connection. The FH is completely unaware and unaffected by the cell switch because the switch is done seamlessly. I-TCP integrates specialized support necessary for the low speed and unreliable wireless link and MH, while the fixed network is left unchanged.

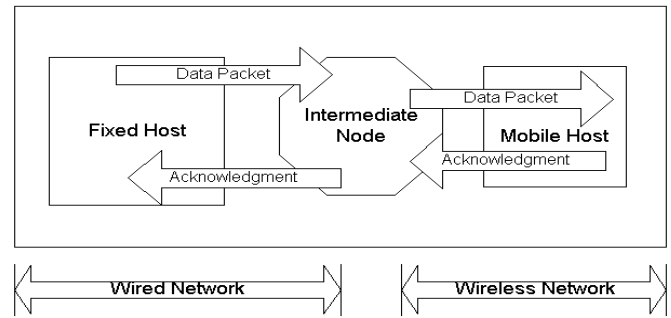


Fig. 2 TCP split mode connection

However, I-TCP requires additional overhead at the MSR for per packet processing from both ends of the network for a single packet: data or ACK. Thus, the intermediate node processes a TCP packet before it reaches its destination and violates the end-to-end semantics of the original TCP. I-TCP also performs poorly when there are lengthy disconnections, and fails with encryption.

M-TCP [15] is a protocol that employs the split connection approach but maintains end-to-end semantics of TCP. It is designed to deal with spurious timeout caused by frequent or lengthy disconnections observed in wireless links. M-TCP operates a three layer hierarchical structure consisting of MH, mobile support station (MSS), and supervisor host (SH). In M-TCP, every TCP connection is split in two at the SH, which is connected to the fixed network. The TCP connection from the SH to the fixed host (FH) uses standard TCP, while the connection between the SH and mobile host (MH) uses a modified version of TCP. The SH serves a gateway function, and has a bandwidth management module that assigns a fixed amount of bandwidth to each connection. Since the available bandwidth varies in heterogeneous networks, the SH recalculates the fixed amount of bandwidth periodically. The SH receives a segment from the FH and forwards it to the MH, and receives the ACK from the MH and forwards it to the FH.

Unlike other split connection schemes, M-TCP saves the ACK of the last byte in the SH in order to prevent loss of outstanding packets. If the SH stops getting ACKs from the MH, it assumes the MH is temporarily disconnected and sends the saved ACK with an advertised window of the MH as zero to the FH. When the supervisor host (SH) receives non-zero window size indication from the MH (i.e., MH is reconnected), the SH replies to the persist packet from the FH with an appropriate window size. Thus, the FH can resume transmission at full rate. However, a major disadvantage of M-TCP is the complexity of the SH.

2) *End-to-end Mode Solutions*: In the end-to-end solutions, only the end hosts participate in flow control and congestion control. The receiver notifies the sender of prevailing network conditions and thus the sender decides whether, or not, to initiate congestion control. The end-to-end approach maintains the end-to-end semantics of TCP and introduces no additional overhead at an intermediate node for processing. However, it requires the modification of transport layer of the TCP sender stack. Figure 3 shows an end-to-end TCP scheme.

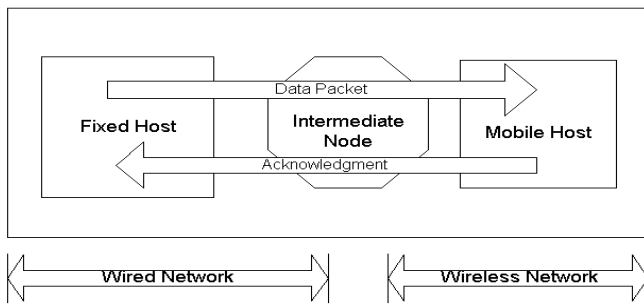


Fig. 3 TCP end-to-end connection

The ability to probe the network for the available bandwidth especially with bandwidth oscillation is key to the performance of end-to-end approach. The available bandwidth of flow is the minimum unused link capacity of the flow's fair share along the path [1]. The end-to-end solutions differentiate congestion losses from non-congestion losses, and implement congestion control by employing reactive or proactive mechanisms.

a) *Reactive congestion control*: In reactive congestion control, the TCP sender adjusts the congestion window ($cwnd$) when the network situation exceeds a threshold or becomes marginal. The standard Reno scheme is an example of reactive congestion control. The scheme responds to packet loss indicated by timeout or DUPACKs. Thus, congestion is unavoidable and TCP will only react to congestion by halving the $cwnd$ and reducing the transmission rate, which may not be necessary for wireless random losses.

Wireless random losses can be contiguous in nature and the fast recovery mechanism of TCP Reno can recover from just one loss in one window. The TCP NewReno modifies the fast recovery mechanism of TCP Reno to cope with multiple losses in one window. However, the TCP NewReno mechanism does not terminate until all multiple losses are recovered, as indicated by the recovery of partial ACKs of one window.

An example of a reactive solution for wireless networks is TCP selective acknowledgement (TCP SACK) [16]. TCP SACK is an extension for TCP that uses SACK instead of the cumulative ACK used in TCP Reno and TCP NewReno. The cumulative ACK scheme forces the sender either to wait for a round trip time (RTT) to ascertain each lost packet, or to retransmit unnecessarily, correctly received packets. When multiple packet losses occur, TCP SACK reacts by sending ACKs for successfully received and queued up data at the receiver. The sender knows the exact packets that are lost and retransmits only the lost packets. Moreover, like TCP Reno, TCP SACK responds to packet loss reactively thus limiting its ability for congestion avoidance.

b) *Proactive congestion control*: In proactive congestion control, the TCP sender adjusts the congestion window ($cwnd$) according to feedback information, which indicates network conditions. Hence, the sender reacts intelligently to the network condition, or the cause of packet losses, and thus prevents the network from going into an undesirable state.

TCP Veno [17] is an example of a proactive scheme for dealing with random losses by using the network congestion

level to indicate the cause of packet losses. The scheme uses TCP Vegas [18] mechanism of estimating network condition to determine the likely cause of a packet loss. If the estimate indicates that the network is not congested, TCP Veno safely declares a loss as random. In addition, TCP Veno modifies TCP Reno by reducing the sending rate less aggressively thus preventing unnecessary throughput degradation. If the estimate indicates network congestion, TCP Veno considers the loss congestive and activates the standard TCP Reno scheme. However, TCP Veno provides little support for mobility.

TCP Westwood [19] is a sender-side proactive approach that uses the end-to-end bandwidth estimate to differentiate the cause of a packet loss — congestion or wireless link characteristic. The TCP sender continuously measures the rate of the connection by monitoring the rate of returning ACKs. Then the TCP sender uses this estimate to compute the $cwnd$ and $ssthresh$ after it detects congestion: three duplicate acknowledgements (DUPACKs) or timeout. TCP Westwood does not require extra overhead for packet inspection at an intermediate node. Though TCP Westwood performs well in heterogeneous networks, the performance degrades when random packet loss rate exceeds a few percent.

TCP-Real [20] is a receiver-oriented proactive scheme that uses a wave-based communication pattern. A wave is a number of fixed-sized data packets sent back-to-back, matching the inherent characteristic of TCP to send packets back-to-back. The wave gives a well-known pattern of exchange, which permits the receiver to decide the wave level based on the forward path. Based on the wave level, the receiver estimates the level of congestion, and the estimate determines the sender's window. TCP-Real eliminates unnecessary idle periods by adjusting to the next wave level after a back off. The forward path-based estimation improves performance in asymmetric paths; the receiver decides rate adjustment regardless of the reverse-path characteristics. However, TCP-Real provides little support for mobility.

TCP-Jersey [21] is a sender-oriented proactive scheme that adjusts the sending rate according to the network condition. This scheme implements the available bandwidth estimation (ABE) algorithm and the congestion warning (CW) router configuration. The TCP sender continuously estimates the bandwidth available for a connection using the ABE algorithm. The router configured with CW marks all packets when congestion is incipient, thus differentiating packet loss caused by congestion from packet loss caused by wireless link errors. However, the router needs to support explicit congestion notification (ECN) in order to implement CW, and modification is required at the TCP sender.

Proactive solutions are generally more robust than reactive solutions in handling random losses because proactive solutions distinguish congestion losses from random losses more quickly.

B. Application-specific Solutions

The application-specific solutions are designed to meet the needs of specific wireless networks. Although all wireless

networks have a common problem of high BER, each has its own peculiar characteristics that affect TCP performance.

1) *Cellular Application*: Handoff by mobile clients is a typical characteristic of cellular network, which a scheme designed specifically for cellular networks should address. Freeze-TCP [22] is an end-to-end solution designed specifically to address TCP throughput degradation caused by frequent disconnection and reconnection due to handoff or temporary blockage of radio signals by obstacles. It assumes the mobile client has knowledge of the radio signal strength, hence can predict an impending disconnection. In Freeze-TCP, when an impending disconnection is detected, the MH sends a zero window advertisement (ZWA), which forces the sender (FH) into persist mode. In persist mode, the FH stops sending more packets but sends zero window probes (ZWP) while keeping its window size unchanged. When the MH reconnects, it responds to the ZWP with triplicate reconnection ACKs, and the FH continues transmission using a window size consistent with the advertised value. However, the TCP layer protocol must be exposed to details of roaming and handoff algorithms implemented in MH. Freeze-TCP is also not useful in case of a disconnection when the FH and MH are not exchanging data.

2) *Satellite Application*: A satellite link experiences an inherently long propagation delay due to the finite speed of light and communication satellite's altitude. During slow start, this long propagation delay increases the time required by the TCP sender to reach a high sending rate. TCP-Peach [23] is a proposed solution to combat TCP performance degradation over satellite links caused by the long propagation delay and high BER. It employs two new algorithms: sudden start and rapid recovery, in combination with the traditional congestion avoidance and fast retransmit algorithms of TCP.

Sudden start introduces dummy packets that are copies of the last sent data packet, which probe the availability of network resources. The TCP sender sends multiple dummy packets between consecutive data packets, and the reception of acknowledgements (ACKs) for the dummy packets indicates the availability of unused bandwidth. The dummy packets are low-priority packets; hence, a router will drop them first if congestion occurs. The rapid recovery algorithm substitutes the traditional fast recovery algorithm to address the high BER. The sender uses dummy packets to interleave data packets and inflates the sending window size upon reception of ACKs for the dummy packets. However, the dummy packets are an additional overhead to this scheme, and modifications are needed at the router and end hosts.

3) *Ad Hoc Application*: Lack of wireless infrastructure, frequent route changes and network partitions are the major causes of packet loss in ad hoc networks. Ad hoc TCP (ATCP) [24] is an end-to-end solution designed for such an environment. It is a thin layer inserted between the IP and TCP layers that listens to network state information provided by ECN and Internet Control Message Protocol (ICMP) messages. Based on the network state information, ATCP puts the TCP sender into the appropriate state: persist state, congestion control state, or retransmit state. When packet loss is due to high BER, Ad hoc TCP retransmits the lost packet without shrinking the congestion window (*cwnd*). If the ICMP

message indicates a route change or transient partition, ATCP puts the TCP sender into the persist state waiting to resume when a new route is found. While ATCP does not alter the TCP code, it requires the modification of the TCP stack.

TCP-Feedback [25] is an algorithm also designed specifically for ad hoc networks. It uses route failure notification (RFN) and route reestablishment notification (RRN) to address route disconnections and reconnections.

4) *Wireless Wide Area Networks (WWANs)*: These are networks characterized by low variable bandwidth, very high and variable delays, path asymmetry, and non-congestion related losses. Wireless TCP (WTCP) [26] is a rate-based end-to-end mechanism that distinguishes random losses from congestion losses using inter-packet delays. The receiver computes the desired sending rate with its rate control mechanism and notifies the sender of this rate in the ACK packet. On reception of the ACK, the sender adjusts its rate accordingly. When the sender waits for ACKs for one or more packet beyond threshold duration, it goes into blackout mode. The sender sends out a probe packet periodically to elicit ACKs from the receiver to recover from the blackout mode.

C. Layer-based Solutions

In the Layer-based approaches, proposed solutions are based on distinguishing between types of losses encountered in a wireless network.

1) *Pure Link Layer*: Pure link layer solutions are based on hiding the unwanted wireless link characteristics from the higher layer protocols. Link layer protocols are implemented, which employ error control techniques: forward error correction (FEC), automatic repeat request (ARQ), and go-back-N to retransmit lost packets. AIRMAIL is an example of a link layer protocol, and it employs a combination of FEC and ARQ for loss recovery. The pure link layer solutions require no modification of the TCP protocol stack. 2.5G and 3G cellular systems have low rate of packet loss because of link-layer protocol [10]. However, recovery at the link layer appears as delay jitter to higher layer protocols increasing the computed RTO value. In addition, error control techniques are applied regardless of the QoS requirements of the individual flows sharing the link. Flow-adaptive link layer approach [12] is proposed to address the QoS issue.

2) *Soft-state Cross Layer*: The soft-state layer solutions make the TCP sender aware of the wireless link through the link or network layer informing the transport layer of particular events. Then the transport layer responds accordingly. Different solutions proposed use control messages: ICMP, ECN, explicit loss notification (ELN) [27] and multiple ACKs [28] to inform the transport layer of a particular event. The soft-state layer solutions distinctly separate congestion losses from error losses. However, the TCP sender cannot be informed when extreme network congestion occurs. Furthermore, these solutions require changes to the transport layer of protocol stack at the sender and some or all intermediate nodes.

3) *Soft-state Transport Layer*: The proposed solutions employ a "TCP-aware local loss recovery, and use caching to save the TCP sender from unnecessary invocation of congestion control. The solutions maintain a "soft" state, (i.e.,

it is not crucial for end-to-end connection, and thus preserves the end-to-end semantics of TCP). However, they require modification of the intermediate node (base station) and optionally, the mobile host (MH).

TCP Snoop [29] is classified as a soft-state solution that provides reliable link layer that is TCP-aware. It uses DUPACKs to identify packet loss, and a negative ACK (NACK) scheme for data transfer from the MH to the fixed host (FH). When TCP Snoop detects a packet loss, it retransmits the lost packet locally. Concurrently, TCP Snoop shields the TCP sender from the wireless link errors especially the transient ones. If congestion occurs when sending to the MH, a packet can drop at a bottleneck link between the MH and wireless link. If a packet drops, the dropped packet is mistaken for damage loss by the TCP-aware cache. Hence, congestion signal, (three DUPACKs) will not be propagated to the sender, which constrains TCP Snoop to edges in the path that are free of congestion. In addition, if any link between the MH and wireless link becomes a bottleneck when sending from the MH, the NACK scheme cannot differentiate between damage losses and congestion losses. Hence, TCP Snoop does not perform well in the presence of frequent or lengthy disconnections, and it fails with encryption.

4) *Hard-state Transport Layer*: The approaches in this group employ splitting the TCP connection into wired and wireless portions. The hard-state transport layer approach was initially proposed for implementation in I-TCP, but has been integrated for use in transit satellite links. The major advantage is shielding the wireless link errors from the end-to-end TCP flow; hence, the flow can fully utilize its share of bandwidth over the entire path. However, hard-state transport layer approaches sacrifice the end-to-end semantics of TCP.

5) *Pure Transport Layer*: In the pure transport solutions, only the end hosts carry out the flow control and congestion control and address losses solely on an end-to-end basis. TCP SACK is an example of a pure transport layer solution. The benefit of transport layer approaches is that they maintain the state that is crucial to end-to-end connections.

They also require no changes or extra processing capabilities at intermediate nodes. Moreover, the transport layer of the TCP stack needs to be modified at the sender.

V. COMPARISON OF SOLUTION WITH IDEAL SOLUTION

An ideal solution for TCP implementation in wireless or heterogeneous networks will possess the following characteristics:

- 1) Robustness against high BER: High BER is common to all wireless networks, and should be addressed effectively.
- 2) Robustness against disconnection: The performance of an ideal solution should not degrade with any type of disconnection: frequent or long.
- 3) TCP end-to-end semantics: An ideal solution should maintain the true end-to-end semantics of TCP.
- 4) Modification requirement: It should be operable with existing infrastructure without requiring extensive modifications at end stations and intermediate nodes
- 5) Support for mobility: Mobility of clients on the wireless or heterogeneous network should be supported.
- 6) Intermediary requirement: The role of intermediaries, if any, should be 100% efficient such that it does not affect the end-to-end TCP performance efficiency.
- 7) Support for encryption: An ideal solution should support encryption, which is crucial to network security.

These characteristics of an ideal solution address problems encountered in all types of wireless or heterogeneous networks. The comparison of the proposed solutions based on characteristics of an ideal solution is shown in Table 1. From the table, no solution has all the characteristics of an ideal solution. The proposed solutions sacrifice at least one characteristic because of specific wireless applications they are designed for, or a specific problem they are designed to address in the wireless or heterogeneous networks.

Table 1. Comparison of proposed solutions with an ideal solution

TCP Solution	Robustness against high BER	Robustness against disconnection	TCP end-to-end semantics	Intermediary requirement	Support for mobility	Support for encryption	Modification Requirement
I-TCP	Yes	No	No	Yes	High	No	Base station
M-TCP	Yes	No (frequent)	Yes	Yes	High	No	Base station
TCP-SACK	Yes	Yes	Yes	No	low	Yes	Sender
TCP Veno	Yes	Yes	Yes	No	Low	Yes	Sender
TCP Westwood	Yes	Yes	Yes	No	Low	Yes	Sender
TCP Real	Yes	No (long)	Yes	No	Low	Yes	End stations
TCP Jersey	Yes	Yes	Yes	No	Low	Yes	Sender and router
Freeze-TCP	No	Yes	Yes	No	High	Yes	Base station
TCP-Peach	Yes	Yes	Yes	Yes	High	Yes	Router and end stations
ATCP	Yes	Yes	Yes	No	High	Yes	TCP stack
WTCP	Yes	No (long)	Yes	No	High	Yes	End stations
TCP Snoop	Yes	No (frequent)	Yes	Yes	Low	No	Base station

VI. CONCLUSION

In this paper, we presented the congestion control algorithms of standard TCP and their performance efficiency in wired networks because of the negligible random bit error rate of wired networks. We also presented various types of wireless networks and their characteristics, which pose problems to the performance of standard TCP. Then we discussed various solutions for TCP implementation in wireless networks, and their pros and cons. We also presented the characteristics an ideal solution should possess.

Although none of the solutions has all the characteristics of an ideal solution, they tackle particular problems in wireless networks, or specific types of wireless networks. While analyzing the performance improvements of each solution over standard TCP is important, making a general conclusion of one solution against another is difficult. This difficulty arises because application scenarios differ from network to network, and different types of wireless networks do not have the same characteristics.

Some proposed solutions target specific wireless network problems and thus cannot solve every wireless network problem. For some solutions, the need for intermediary requirements outweighs every other design consideration. If modification is minimal, the intermediary requirement is an additional overhead. However, exclusion of an intermediary requirement is at the expense of extensive modifications in the in the end stations or TCP stack. Furthermore, split mode solutions sacrifice support for encryption by not maintaining the end-to-end semantics of TCP.

Moreover, proactive solutions that maintain end-to-end semantics if combined with link level approaches could address most problems and manage network bandwidth more efficiently. However, interoperability with existing infrastructure and modification requirements remain a key factor in implementing such solutions.

REFERENCES

- [1] Y. Tian, K. Xu, and N. Ansari, "TCP in wireless environments: problems and solutions," *IEEE Commun. Mag.*, vol. 43, no. 3, pp. S27-S32, Mar. 2005.
- [2] V. Jacobson, "Congestion avoidance and control," in *Proc. ACM SIGCOMM*, Stanford, CA, Aug. 1988, pp. 314-329.
- [3] R. Braden, "Requirements for Internet hosts – communication layers," *IETF RFC 112*, Oct. 1989.
- [4] Information Sciences Institute, "Transmission control protocol," *IETF RFC 793*, Sept. 1981.
- [5] M. Allman, V. Paxson, and W. Stevens, "TCP congestion control," *IETF RFC 2581*, Apr. 1999.
- [6] S. Floyd and T. Henderson, "The NewReno modification to TCP's fast recovery algorithm," *IETF RFC 2582*, Apr. 1999.
- [7] D. Chiu and R. Jain, "Analysis of the increase/decrease algorithms for congestion avoidance in computer networks," *Journal of Computer Networks and ISDN systems*, vol. 17, no. 1, pp. 1-14, June 1989.
- [8] P. Karn and C. Partridge, "Round trip time estimation," in *Proc ACM SIGCOMM*, Stowe, VT, Aug. 1987.
- [9] V. Paxson and M. Allman, "Computing TCP's retransmission timer," *IETF RFC 2988*, Nov. 2000.
- [10] H. Inamura, G. Montenegro, R. Ludwig, A. Gurtov, and F. Khafizov, "TCP over second (2.5G) and third (3G) generation wireless networks," *IETF RFC 3481*, Feb 2003.

- [11] M. Allman, D. Glover, and L. Sanchez, "Enhancing TCP over satellite channels using standard mechanisms," *IETF RFC 2488*, Jan. 1999.
- [12] R. Ludwig, "A case for flow-adaptive wireless links," University of California, Berkeley, CA, Tech. Rep. UCB//CSD-99-1053, May 1999.
- [13] M. Patel, N. Tanna, P. Patel, and R. Banerjee, (2001, Nov.) TCP over wireless networks: issues, challenges and survey of solutions. [Online]. Available: http://bbcr.uwaterloo.ca/~pinhan/ECE710/TCP_wireless.pdf
- [14] A. Bakre and B. R. Badrinath, "I-TCP: indirect TCP for mobile hosts," Department of Computer Science, Rutgers University, Piscataway, NJ, Tech. Rep. DCS-TR-314, Oct. 1994.
- [15] K. Brown and S. Singh, "M-TCP: TCP for mobile cellular networks," *ACM SIGCOMM Computer Communication Review*, vol. 27, no. 5, pp. 19-43, Oct. 1997.
- [16] M. Mathis, J. Mahdavi, S. Floyd, and A. Romanow, "TCP Selective Acknowledgment Options," *IETF RFC 2018*, Oct. 1996.
- [17] C. P. Fu and S. C. Liew, "TCP Veno: TCP enhancement for transmission over wireless access networks," *IEEE J. Select. Areas Commun.* vol. 21, no. 2, pp. 216-228, Feb. 2004.
- [18] L. Brakmo and L. Peterson, "TCP Vegas: end-to-end congestion avoidance on a global internet," *IEEE J. Select. Areas Commun.* vol. 13, no. 8, pp. 1465-1480, Oct. 1995.
- [19] C. Casetti, M. Gerla, S. Mascolo, M.Y. Sanadidi, and R. Wang "TCP Westwood: Bandwidth Estimation for Enhanced Transport over Wireless Links," in *Proc. ACM MobiCom*, Rome, Italy, July 2001, pp. 287-297.
- [20] V. Tsaoussidis and C. Zhang, "TCP-Real: receiver-oriented congestion control," *Computer Networks: The International Journal of Computer and Telecommunications Networking* vol. 40, no. 4, pp. 477-497, Nov. 2002.
- [21] K. Xu, Y. Tian, and N. Ansari, "TCP-Jersey: TCP-Jersey for wireless IP communications," *IEEE J. Select. Areas Commun.* vol. 22, no. 4, pp. 747-756, May. 2004.
- [22] T. Goff, J. Moronski, and V. Gupta, "Freeze-TCP: a true end-to-end TCP enhancement mechanism for mobile environments," in *Proc. IEEE INFOCOM*, Tel-Aviv, Israel, Mar. 2000, pp. 1537-1545.
- [23] I. F. Akyildiz, G. Morabito, and S. Palazzo, "TCP-Peach: A New Congestion Control Scheme for Satellite IP Networks," *IEEE/ACM Trans. Networking*, vol. 9, no. 3, pp. 307-321, June 2001.
- [24] J. Liu and S. Singh, "ATCP: TCP for Mobile Ad Hoc Networks," *IEEE J. Select. Areas Commun.* vol. 19, no. 7, pp. 1300-1315, July 2001.
- [25] K. Chandran, S. Raghunathan, S. Venkatesan, and R. Prakash, "A Feedback Based Scheme for Improving TCP Performance in Ad-Hoc Wireless Networks," in *Proc. ICDCS*, Amsterdam, The Netherlands, May. 1998, pp. 472-479.
- [26] P. Sinha, T. Nandagopal, N. Venkitaraman, R. Sivakumar, and V. Bharghavan, "WTCP: A Reliable Transport Protocol for Wireless Wide-Area Networks," *Wireless Networks*, vol. 8, no.2, pp. 301-316, 2002.
- [27] H. Balakrishnan, V. N Padmanabhan, S. Seshan, and R. H. Katz, "A comparison of mechanisms for improving TCP performance over wireless links," *IEEE/ACM Trans. Networking*, vol. 5, no. 6, pp. 756-769, Dec. 1997.
- [28] S. Biaz, M. Mehta, S. West, and N. Vaidya, "TCP over Wireless Networks Using Multiple Acknowledgements," Texas A & M University, College Station, TX, Tech. Rep. TR97-001, Jan. 1997.
- [29] H. Balakrishnan, S. Seshan, E. Amir, and R. H. Katz, "Improving TCP/IP performance over wireless networks," in *Proc. ACM MobiCom*, Berkeley, CA, Nov. 1995, pp. 2-11.



Modupe O. Omueti received a BSc. in electronic and electrical engineering from the Obafemi Awolowo University, Ile-Ife, Nigeria in 2001. She worked in Procter & Gamble as Electrical and Instrumentation Manager, and Mobile Telephony Network as Switch Engineer before returning to the academia. Currently, she is a MASc. candidate in engineering science at the Simon Fraser University. Her research interests include communication networks, wireless communications and performance analysis of network protocols.

Miss Omueti is a member of the Institute of Electrical Engineers (IEE), and a graduate member of the Nigerian Society of Engineers (NSE).

Epilepsy Monitoring System for Diagnosis and Treatment

Paris S. Lisette, Simon Fraser University, Burnaby, BC

Abstract—Research about epileptic patients requires monitoring neurological information along with other physiological information such as electroencephalogram (EEG), respiration, oxygen level, and recorded videos over a period of time. The physiological parameters are stored and are available to the doctors and clinicians who are conducting the diagnosis and researching the condition. To obtain the required information, an epilepsy monitoring system (EMS) is necessary because it provides all the specialized devices needed. An obsolete analog system encounters problems such as synchronization accuracy, high maintenance costs, and space for storing the recorded information. Replacing it with a new digital system, we benefit from automatic synchronization, maintenance costs decrease, and the patients' recorded information is stored in an organized manner that is easily accessible to the researchers. Specialized storage devices, called Storage Area Networks (SANs) are ideal because patient information is kept in a safe and confidential manner, with a robust backup for effective disaster recovery. Once the information is in the SAN, neurologists and epilepsy researchers can access the information quickly and easily for clinical review and research purposes.

Index Terms—Biomedical equipment. Biomedical monitoring. Biomedical signal analysis. Medical information systems.

I. INTRODUCTION

EPILEPSY is a neurological disease that produces sudden, chronic, and brief changes in brain activity.

No specific cause has been reported; however, any electrical alteration in the neurons of the brain produces an epileptic seizure. A seizure is an abnormal and disproportional electrical impulse of the neurons that produces functional alterations in the brain. Seizures can be categorized into different types, determined by their origin and effects [1]. If the origin and kind of seizure is detected, it can be diagnosed and treated properly; an epilepsy monitoring system (EMS) can help neurologists and clinical staffs make that diagnosis.

The EMS incorporates the following components: 1) one or more epilepsy-monitoring units (EMU), 2) a data display at the nurse's station, 3) a central control unit, and 4) a storage device.

An EMU is a specially equipped room in a clinical facility, where the physical parameters of a patient are obtained, such as electroencephalogram (EEG), respiration, oxygen levels, and video images. Therefore, each EMU is equipped with a bed and washroom, a 128 electrode EEG, respiration and oxygen level sensors, and a video camera.

The monitored parameters are displayed at the nurse's station as well as at the central control unit, which is a room to gather, record, review, and synchronize the gathered data.

To review the data, an authorized person can fast forward, rewind, pause, delete, and play them at any time. Fig. 1. illustrates the layout of the EMS and the connections among its components.

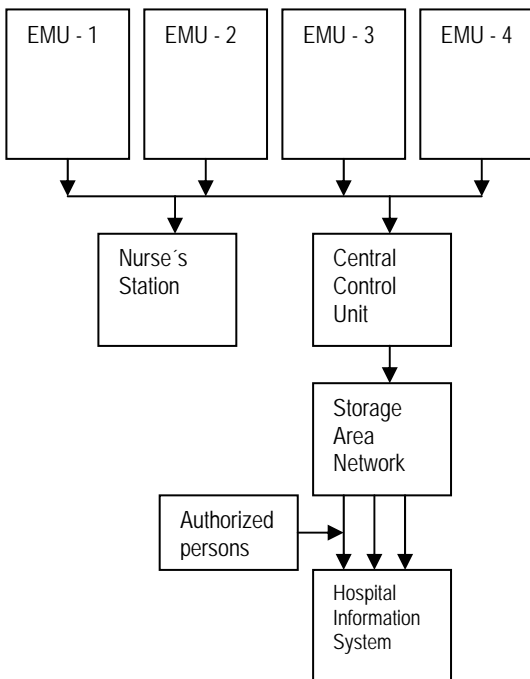


Fig. 1. Epilepsy Monitoring System layout.

A full study to diagnose epilepsy consists of being monitored with an EMU for 24h to 36h; the average amount of information gathered during a session is approximately 20 Gb per patient. Only the data before, during, and after a seizure is clinically valuable because during that period the patient displays abnormal brain and

physical behavior. Then, the neurologist and the clinical team analyze the information and reduce it to approximately 1 GB. All this information must be easily accessible in the patients' clinical history for epilepsy research.

Traditionally, information was stored on writable CDs and each patient would require two or three CDs; this method presents some problems. First, when clinicians have to review patient data from previous monitoring sessions, the search for the appropriate CD where the study was stored is time consuming.

Secondly, CDs are not suitable for saving the patient study in a confidential way, at the same time; information is not included in the patients' clinical history. Finally, there is no backup.

II. MOTIVATION

In this article, we are going to address a specific hospital's EMS that was bought approximately 10 years ago. This system obtains an analog EEG signal and an analog video signal, along with the other parameters required. Because it has been in constant use since it was acquired, its components have deteriorated and are considered an obsolete system compared to the ones now available.

The analog system is composed of four EMUs, each with a 128 electrode EEG, and a video camera. The central control unit records an EEG on a 1.44MB floppy-drive in the CPU and records the video on a conventional VCR using VHS tapes; it is also equipped with a monitor to view the data. Clocks and timers are required to synchronize the two signals. After the proper staff edit the information, it is saved on CDs, encountering the already mentioned storage problems.

Besides being an obsolete system, the maintenance cost of this analog system has become a problem for the hospital. The VHS tapes can store 6h to 8h of data, depending on the recording speed, so in a 24h study, three cassettes are required for one patient, which can become costly over the years. Between numerous videotapes, VCR repairs and cable maintenance, the yearly/overall cost of the system is high compared to the cost of a new digital EMS. Another problem that would be solved with a new digital system is the storage complication. A study of the digital EMS in the market was made.

The digital EMS has an EEG with a 128 electrodes that obtain the brain signals with specific amplifiers and filters to eliminate noise. It is also equipped with several DC inputs, which are able to accept different signals such as oxygen level, respiration, etc., which are also connected to the amplifiers. The use of a digital camera is considered necessary to record the patients' activity as well as their faces while they are being monitored.

The storage disadvantages can be overcome by using a Storage Area Network (SAN), which is a high-speed and high-bandwidth storage network [2]. This network can store up to 1Tb, and it has the capacity to be connected to the central control unit, providing access to stored information more quickly and easily. It can also be set for

periodic automatic backup keeping the information safe in case of a disaster. The SAN has the ability to share information with other systems, being an important property for a device in a hospital.

Hospitals are regulated by a specific database that holds the patients' information; it includes the patients' personal information and their clinical history along with the results of medical studies. Such database is called the hospital information system (HIS) [3] and is useful for doctors, nurses, and other staff of the hospital as a source of updated data. Information contained in the database is the source for decision-making about patients' treatment. By having a digital EMS, the data from the epilepsy study stored in the SAN can be shared with the HIS, bringing the results of the study together with the rests of the patients' clinical history and personal information. Not all staff is allowed to review the epilepsy studies, therefore, a mechanism involving security passwords has to be implemented for safety.

After evaluating the benefits and drawbacks for the acquisition of the new digital EMS, we decided to purchase the new system. In the next section, we will describe the process of installation and testing the new system, followed by the connection of the SAN and the corroboration of real-time acquisition. At the end, we will discuss the results and conclusions.

III. METHOD

A. Installation

Our first step was to visit the hospital in order to familiarize ourselves with the area and the devices used for monitoring in order to coordinate the installation and implementation of a four-bed EMS. We measured the distance from each of the EMUs to the nurse's station and to the central control unit, and designed a layout of the way the cables would travel from one room to another.

Our second step was to install the four EMUs and connect them with the central control unit and the nurse's station. All the cables were installed above the ceiling to maintain space in the halls and in the rooms, providing security to the people around the system and for the safe transference of data.

Connections were completed one unit at a time, making sure each cable arriving at the central control unit was labeled, to avoid confusion while connecting them and for future reference.

B. Test the Epilepsy Monitoring System (EMS)

The EMS was tested in several stages; each test was run at the four EMU. We tested the EEG system with its own verification test, to reassure the system had no problems or failures. As a second test, we employed a signal generator as an input to the electrodes, selecting a sinusoidal wave with a specific frequency and amplitude. Then we analyzed the output by comparing it with the original input, reviewing that the frequency and the amplitude of the sinusoidal were equal.

The digital camera was set on record, testing it's functions capabilities under different conditions of light and different external situations. It was confirmed that the

ENSC 894 COURSE TRANSACTION

correct signal was obtained regardless of the amount of light received from the environment.

Once the EMUs were working properly, we performed some tests on the central control unit to be confident that the data was being recorded, and the software for reviewing it worked correctly. The following list describes the steps of the test:

1. Set an EMU working with all the physiological signals.
2. Open the study from the central control unit and verify the acquisition of data in real time.
3. Switch off the EMU in the room.
4. Verify that the data is still available in the central control unit after its acquisition.
5. Save it under a specific name.
6. Edit the information by using all the review functions.
7. Save the edited study under a second name.
8. Close all opened files.
9. Finally, open the two saved files, and verify that they were saved and retrieved correctly as well as verifying that changes were saved in the edited file.

C. Configuration of the SAN

In order to configure the SAN we developed a storage strategy, which accommodates the way the information is stored [4]. The Department of Informatics at the hospital helped with the planning and the implementation of the SAN.

Our approach to coordinate the connection between the EMS and the SAN is as follows:

1. Connect the central control unit and the SAN with an optical cable.
2. Store the information in the SAN.
3. Open the saved file on a hospital's computer, with a pre-defined password. This gives us certainty that the information is being stored in the SAN and it is available.
4. Measure time of storing and retrieving a full study, 1GB approximately, in the SAN.

D. Corroborate real-time acquisition and proper storing information

To corroborate the acquisition in real-time, the digital camera was focused to a clock at the epilepsy monitoring unit, the clock was synchronized with the one on the display monitor of central control unit. At a specific time, we started a full study just watching the time change on the clock. We saved the full study in the SAN [5] and then we compared the times on the central control unit and edited the information. We reviewed time in which the information had arrived, helping to verified the low latency of the network.

IV. RESULTS

The information coming from each of the four EMUs arrived in real time to the central control unit, and to the nursing station. Confirming that the information coming from the EEG, respirator, oxygen levels, and video camera

of one patient, in a certain instant, is coordinated and reproduced. These successful results grant confidence to the digital EMS study and assure that the information in it is reliable for decision-making.

At the central control unit, the data is easily accessible in real time and after the study is concluded. The review tools also work correctly.

The SAN was connected and configured properly, because the information coming from the EMS was stored and can be searched by authorized staff afterwards easily. This procedure is executed in less time than it used to take for finding a specific study on the CD's, demonstrating the success parameter of overcoming traditional storage methods.

V. RELATED WORKS

The problems we have addressed in this project have been approached in different ways. The storage problem was a conference topic at the Annual Fall Meeting of the Biomedical Engineering Society in 2002. The Lab for Computational Neuroscience proposed and implemented an object based video compression for the data of a full EMU study, this was achieved by separating the video frames into different components and then compressing them individually [6]. Another approach to overcome the accessibility and security of the information obtained in a study was developed by the Dept. of Neurological Surgeon in Pittsburgh [7], with a multimedia interface MPEG-7, that provides a safe platform for the information to be stored and gives the bonus of remote accessibility via the internet.

VI. CONCLUSION

Epilepsy is a disease of the brain caused by abnormal functioning of the neurons, an epilepsy monitoring system (EMS) can be used to diagnose epilepsy and prescribe proper treatment. The EMS study lasts for 24h to 36h and the data of a full study is not needed, because only the information from the seizures and pre and post seizures are important for a clinical diagnose. Therefore, the information is reviewed and edited by experts and the useful data is sorted. In order to store the studies and have the information easily accessible, the application of a SAN is needed. The SAN stores the information in real time, and authorized persons can review it anytime or anywhere from a computer. Proper storage of the obtained data provides neurologists an easy and convenient method to study each patient separately for specific treatment.

REFERENCES

- [1] M. Clomenares, MD., (2005, Nov 13), *Epilepsia*, Abcmedicus, [Online]. Available: <<http://www.abcmedicus.com/articulo/pacientes/id/57/pagina/4/epilepsia.html>>.
- [2] SUN Microsystems, (2005), [Online]. Available: <<http://www.sun.com/aboutsun/media/presskits/netstorage.html>>.

ENSC 894 COURSE TRANSACTION

[3] K. Ohe, (1995, Aug 7), Internet Society, Public Health and Medicine. [Online]. Available: <<http://www.isoc.org/HMP/PAPER/238/>>.

[4] R. Romo Reyes, "Propuesta de una red de computadoras para empresas", México, 1997.

[5] M. Camargo Fernández, "Red local de comunicación para dispositivos de aplicación específica", México, 1990.

[6] Minqui Sun; Quiang Liu; Scheuer, M.L.; Sciabassi, R.J., (2002), **Assessment of object-based video compression for epilepsy monitoring, IEEE**, [Online]. Available: <<http://ieeexplore.ieee.org.proxy.lib.sfu.ca/search/searchresult.jsp?history=yes&queryText=%28%28epilepsy+monitoring%29%3Cin%3Emetadata%29>>

[7] Lin-Sen Pon; Sciabassi, R.J.; Qiang Liu; Scheuer, M.; Mingui Sun, (2004), **A medical EEG/video multimedia content description system, IEEE**, [Online]. Available: <<http://ieeexplore.ieee.org.proxy.lib.sfu.ca/search/searchre>

[sult.jsp?history=yes&queryText=%28%28epilepsy+monitoring%29%3Cin%3Emetadata%29](http://ieeexplore.ieee.org.proxy.lib.sfu.ca/search/searchresult.jsp?history=yes&queryText=%28%28epilepsy+monitoring%29%3Cin%3Emetadata%29)>



Paris Shaadi Lisette. Born in Mexico City, Mexico, August of 1979. She received Bachelors in Applied Science Engineering, biomedical. She studied at the Universidad Iberoamericana in Mexico City, Mexico, 1998.

She interned at the Hospital for Sick Children, Toronto, ON. She worked for a Mexican Company, specializing in medical equipment, as a Hospital Manager, assisting and managing devices for minimal invasive and cardiovascular surgery.

Laser Patterned Sn/In Film and Its Applications

Jun Peng

School of Engineering Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada

Abstract --a new kind of UV and visual transparent and conductive films were created by use laser exposure bilayer of metal of Sn and In in air atmosphere. Sn over In films (15-120nm thick) with a 1:10 thickness ratio were deposited by DC sputtering. An argon laser beam exposed the film into patterns. XRD, SEM, TEM, and Auger have been used to investigate the film's microstructure and composition suggesting ITO like characteristics. XRD indicated a preferred In_2O_3 (222) orientation which is similar to ITO films deposited by other methods. High temperature post-annealing (500 °C) was performed and it was found that transparency and conductivity of the film can be further improved. The film can be developed by a wet solution removes unexposed Sn/In leaving patterned ITO like films on substrate. The large changes in optical density means Sn/In films can be used as a material of the direct write photomasks. The film formed on pSi wafer can create a heterojunction structure, which showed a photovoltaic characteristic when illuminated by ordinary light.

Index Terms: transparent and conductive film, laser exposure, ITO film, photoresist, anisotropic etch, direct write mask, etch resist.

I. INTRODUCTION

The transparent and conductive films like indium tin oxide (ITO) and tin oxide films have been widely used in liquid crystal and solar cells and transparent electrodes. In most of these applications the films were firstly deposited or formed on the substrates and then followed by traditional photolithography, which involves photoresist coating, exposure, developing, wet chemical etching, stripping and baking, to get designed patterns. However, those methods have many drop backs such as chemical environmental impact, the etch was sensitive to the ITO stoichiometry and deposition method and etched ITO films result in poorly-defined edges with undercutting and incomplete etching [1]. In our research, a direct laser patterning process was developed to create transparent and conductive films, which is a single-step, dry writing

process and requires no mask, thereby eliminating all the photolithographic process steps. And, since there is no chemical process and no heating of the substrate, it can be used in some special applications in which chemical process and heating substrate are not allowed.

In our earlier research Bi/In bilayer films demonstrated high transparency and conductivity after laser exposure in air. Under laser exposure Bi and In form an eutectic alloy which appears to create an oxidized material due to the laser heating the film. A development process was also created which removed the unexposed Bi/In film leaving only the exposed material [4,5]. This has lead us to consider using laser exposed Sn/In bilayer films as an alternative for creating layers with similar structure and properties to ITO. The research suggested that a new kind transparent and conductive film could be direct patterned by use laser beam exposing bilayer metal film of Sn/In. Thus in this paper we report this new kind film making, its structure analysis and its applications.

II. EXPERIMENTS OF BILAYER FILMS

2.1 The film deposition

The Sn/In film consists of two sequential DC-sputtered Sn and In thin layers (thickness around 15 –120 nm). The thickness ratio of Sn to In will result in material with different characteristics. The DC sputtering was carried out by use a Corona DC sputter machine, which can deposit multilayer films without an air break. This is needed for our sample because our sample has a bi-layer structure. The target materials used were two inches in diameter 99.99% Sn and 99.99% In. Before the deposition, the substrates of Si (100) wafer or glass slides were RCA1 and RCA2 cleaned and then baked for 20 minutes at 120 °C to remove moisture [5]. The system was first pumped down to a base pressure of 6×10^{-7} Torr before sputtering. Argon was used as the sputtering gas introduced through a needle valve. During deposition the substrate was kept at the room temperature, the chamber pressure was at 4 m Torr with an Ar gas flow rate of 10 sccm. The as-deposited films showed a shiny surface and their thickness could be controlled by the time and power of deposition .

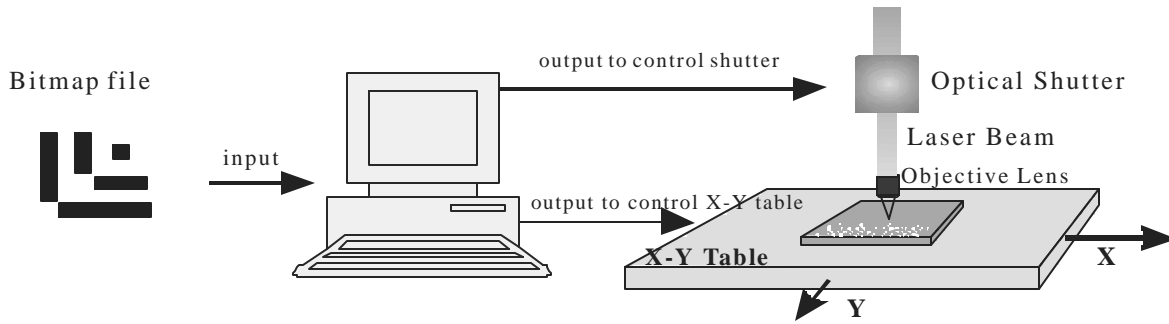


Figure 1. Sn/In bilayer film exposure setup. The script or bitmap files input into X-Y table to control computer. The laser beam was focused by a 50 x objective lens and the sample was placed on an X-Y table.

The deposition rate for tin and indium was: Sn 6.4 Å/Min·W and In 4 Å/Min·W. The film thickness was measured by a stylus profilometer (Model: TENCOR Alpha-Step 500). Similarly, pure indium films with same thickness have been deposited in order to compare the results.

2.2 The laser exposure set up

To write patterns on the Sn/In film, a computer controlled X-Y-Z positioning table and focused 5 W argon ion laser (514/488 nm) was used. The 2.5 mm diameter laser beam was focused by different objective lenses to create variable spot size. Depending on the lenses used, for example 50x objective lens or 50mm lens, the waist of the focused beam ranged from 2 μm to 10 μm. The laser beam was fixed during the scanning and the sample was placed on the X-Y-Z table. The table was moved in a raster scan motion along X direction at a constant speed while there is a constant incremental step along the Y direction. Different designed patterns can be made by writing computer control scripts or by using a specialized mask making program which takes 8bit gray-scale bitmap created by common drawing software as input files [6]. Figure 1 illustrates the picture of the laser exposure system.

III. OPTICAL PROPERTY OF SN/IN FILMS

The most interesting phenomenon is that the exposed Sn/In film becomes almost transparent compared with the highly absorbing unexposed film. Due to an oxidation process described in section 6, the film changes from a shiny, highly reflective, metallic film as deposited into a nearly transparent one. Optical picture of characteristics that was made by laser exposure is shown in Figure 2. Under back lit condition, the exposed area (characters) became highly transparent compared with unexposed area which are highly absorbing (black).

The optical density stabilizes around 0.24 OD at the I-line wavelength for Sn/In (120 nm thickness) on glass slide. With this low absorption levels the Sn/In film is near the needs for a direct write photomask where the converted areas can be used as the mask “openings” without development.

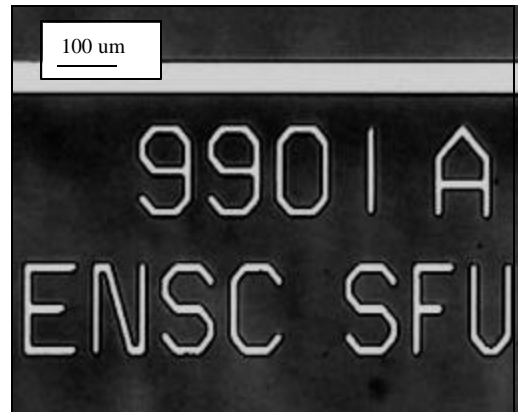


Figure 2. The characters were made by laser exposure, which become highly transparent compared with unexposed dark area.

IV. ELECTRONIC PROPERTIES OF FILMS

The Sn/In oxide films exhibit metallic like conductivity both before and after laser exposure. As the literature on deposited ITO films suggest, this characteristic may be partially attributed to the content of Sn in the film since tin is an intentional dopant [7]. This higher conductivity may come from the oxygen vacancy and excess metal In ions, similar to ITO film [9]. In our experiments, in order to clarify the exposed Sn/In film’s electrical properties, a series of sheet resistance and Hall tests have been performed for films of thickness 120 nm with different Sn ratio and laser scanning power. The sheet resistances (Rs) were measured

by using four-point probe, Model MP0705A made by Wentworth Labs, which is connected to an HP 3478A multi-meter. The 4 probes are arranged in a line with 1 mm space between one and another. It eliminates any contact resistance in the measurement.

In the experiments a series samples with laser exposure power from 0 W to 0.4 W were tested and the average sheet resistance (R_s) data were obtained (with an accuracy of $\pm 5\%$). By assuming that the thickness of the films is uniform, the film resistivity (ρ) was determined using the simple relation $\rho = R_s \cdot d$, where d is the film thickness. At same time, the optical density of exposed film has been tested.

Hall mobility and carrier density measurements were obtained by using the Van der Pauw structure [12] at room temperature. The sample film was 10% Sn with thickness of 120 nm and exposed by laser of 0.4 W. The sample current was 4 mA with the magnetic field strength of 5kG. The results gave the Hall mobility around $5.8 \text{ cm}^2 / \text{V}\cdot\text{S}$ and the carrier concentration was in the range of $7 \times 10^{18} \text{ cm}^{-3}$ to $1 \times 10^{20} \text{ cm}^{-3}$ depending on the laser power and Sn ratio.

V. THE STRUCTURE ANALYSIS

The focused laser light is absorbed in the film creating a temperature rise in that area. The heat energy, in turn, making the shiny and reflecting film into an almost transparent one. In order to understand the mechanisms behind the laser exposure processes, Profilometry, XRD, EDX, Auger, SEM and TEM were used to investigate those structure changes in the processes.

To confirm that Sn/In exposure process is not a laser ablation process in which laser has removed most material in the area it reaches, a profile test has been performed by using a Tencor profilometer. The results show that after laser exposure the area has become slightly thicker than the unexposed area instead of becoming thinner. A slight thickness increase of about 25 nm was observed on the exposed side, which means that a amount of volume increase happened after laser exposure. This increase in the thickness is due to the oxidation.

5.1 XRD structure analysis

To determine the structure of Sn/In laser converted films, X Ray Diffraction (XRD) were performed. The 120 nm films deposited on glass slide substrates. The XRD instrument was Philips PW1730 x-ray generator with CuK α ($\lambda = 1.5418 \text{ \AA}$) and Norelco Diffractometer (Model No. 3-202). In order to compare the results, the XRD was performed for Sn/In films both before and after laser exposure. Figure 3 shows the diffraction peaks of Sn/In (Sn

10%) film with different power scanned. In the figure most peaks have been identified by comparing them with the standard powder diffraction database (JCPDS) [14].

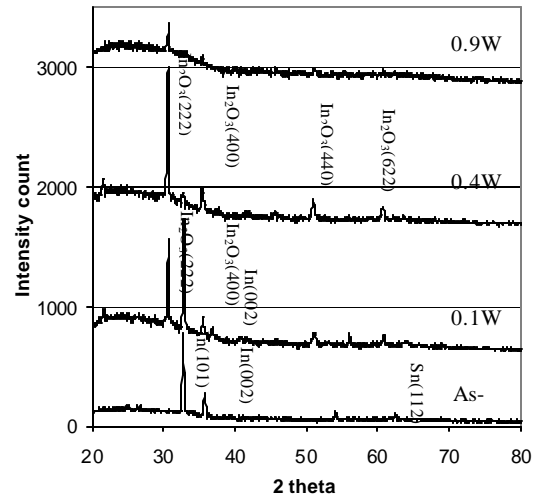


Figure 3. XRD spectrum of Sn/In (Sn 10%), 120 nm thick film exposed with different laser power.

There are four curves in the XRD plot of Figure 3. The bottom curve gives the diffraction spectrum of the film as-deposited. It only consists of strong pure In, Sn peaks and some their compound material weak peaks, showing that large portion of Sn and In were not combined before laser exposure. This result meets with our expectations since the substrate was at room temperature which is much lower than the alloy temperature. The hump from 15° to 35° in the XRD curve is caused by the glass substrate. The curves above the as-deposited are those laser exposed by 0.1 W, 0.4 W, and 0.9 W respectively. With the increase of the laser power, the indium oxide peaks (222) and (400) get stronger, indicating that the laser exposure is an oxidation process in which indium is oxidized from oxygen in air. It is noticed that when the laser power reaches 0.4 W no pure Sn or In peaks remained. This suggests that most pure metals in the film have been converted into oxide materials. However, with further increase of the laser power, only a weak In_2O_3 (222) peak is seen, indicating strong orientation along $\langle 111 \rangle$ direction happened or the film becomes amorphous structure. Comparing the curve of 0.4 W in Figure 3 to the typical XRD pattern of ITO [15] and those reported XRD spectrums of ITO films [15,16], it is found that they are almost same and the In_2O_3 (222) peak at 32 degree is a typical ITO orientation peak, indicating that laser converted Sn/In film has the same structure as ITO.

From above results a conclusion is obtained that process of Sn/In film turning transparent under the laser exposure is a thermal oxidation process and the final

material is a kind of indium tin oxide or indium oxide film, which exhibits transparent and conductive properties like ITO film.

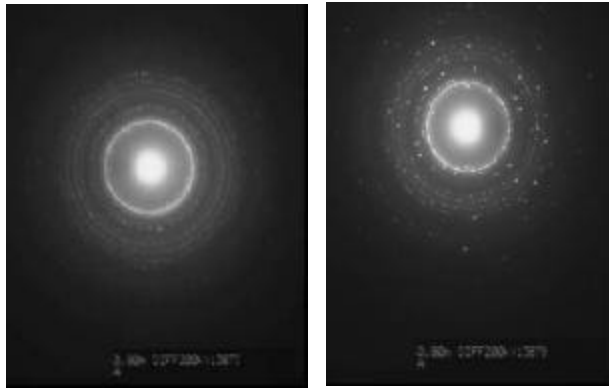


Figure 4. Electronic diffraction patterns of unexposed and exposed Sn/In film on Lacay Carbon coated formvar grids.

5.2. TEM measurement

TEM measurement was performed to identify the structure changes after laser exposure. Since laser exposed film is so thin that the sample cannot be created by use transitional methods. The Lacay Carbon coated formvar TEM grids (300 mesh copper) and SiO-coated copper grids were used to make TEM sample directly. First, Bi and In with thickness of 12/12 nm were DC sputtered on the grids. Then the NdYAG laser working in single pulse condition (4ns) at wavelength of 533 nm was used to expose the film. Since the power on laser beam spot follows the Gauss distribution law, in the exposed area the central section were totally burned because high laser power. The around section had been exposed successfully and the film became transparent. Figure 4 shows the electronic diffraction from unexposed area and exposed area. The difference of those diffraction patterns shows the structure has changed after laser exposure.

Those electronic diffraction rings have been indexed by compare the TEM diffraction data with JCPDS data. The results show that for as-deposited area the possible structures are: In_3Sn , InSn_4 , Sn and In. For the exposed area, the most possible material structure is In_2O_3 . Other possible materials are Sn_2O_3 , SnO_2 , InSn_4 , and In_3Sn . This result is consistent with XRD analysis results, in which same materials were indexed.

VI. ANNEALING IN AMMONIA

We had annealed the 60 nm Sn/In (Sn 10%) films in ammonia ambient at 450 to 500 °C for 30 min. After

annealing it was found that both exposed and unexposed area became more transparent. Figure 5 shows the OD changes after high temperature annealing. The exposed area drops from 0.24 into 0.21 at 365 nm wavelength (I line). This is the lowest OD we have got for 60 nm thick exposed Sn/In films so far. The reason for this dropping may be due to the hydrogen reduction effect during the annealing process since both ammonia and steam ambient consist of hydrogen element [28]. The test showed that the film's conductivity was also improved a little after annealing in those ambient.

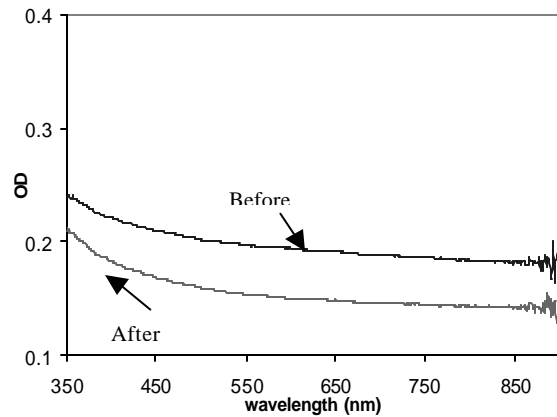


Figure 5. The film OD change after annealing.

VII. APPLICATIONS OF PATTERNED FILMS

The applications of Bi/In laser converted films were reported in our earlier papers [17,18]. In that research Bi/In oxide film was shown to be a new kind of inorganic thermal resists with low energy exposure requirements due to its low eutectic melting point. Further more Bi/In also shows large changes in transparency with laser exposure going from 3 OD to 0.26 OD. It has been used to create good direct write photomask [17,18]. Here we briefly explore Sn/In applications in the same areas.

7.1 Direct write photomask

Photomasks with smaller features, fewer defects, larger areas and lower cost are required by the microfabrication and micromachining industries. Conventional photomasks were made by Cr (chrome) film on transparent glass substrates. The typical Cr photomask manufacturing process involves many steps such as: the blank mask preparation: Cr and Cr oxide deposition, photoresist coating, resist baking, laser or e-beam direct-writing, resist

development, metal layer dry or wet etching, resist stripping and cleaning [20].

In our previous research, based on the OD changes with laser power, the Bi/In bimetallic film has been successfully used as direct write photomask material in our experiments and it reduced the number of manufacture steps [17]. Compared with Bi/In film which OD changes from 3 to 0.26, Sn/In film (Sn 5%) produce a higher transparency (OD changes from 3 to 0.22 at I line). This OD change magnitude is much better than other existing direct write materials such as HEBS, a glass that darkens with e-beam exposure, whose OD changes from 1.9 to 0.4 at the 365 nm [19].

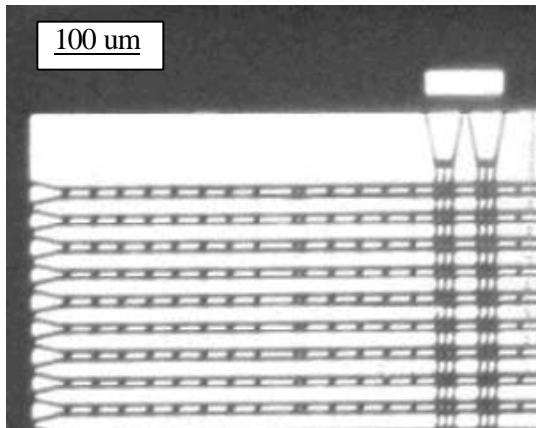


Figure 6. Sn/In direct write mask: thickness = 120 nm, Sn 5%, exposed laser power = 0.45W, the finest feature line is 5 μm wide.

Figure 6 shows the binary photomask with complex patterns made under following conditions: 1). The substrate is RCA1 and RCA2 cleaned glass slides. 2). The laser system script software was used to control the X-Y-Z table. 3). The thickness of Sn/In film is 120 nm with Sn ratio 5%. 4). Argon laser raster-scan power = 450 mW; 5). 50mm objective lens used to focus argon laser beam; 6) X-Y-Z table moving speed (along X direction) = 1 cm/sec; 7). Raster-scan increment (along Y direction) is 4 μm. The finest feature line width is 5 μm. In these examples, which is limited by the laser exposure system used, structures smaller than 1 μm have been produced with other optics on our system.

In order to test the direct write photomask, a “Quintel 4” mask aligner with Hg source was used to expose a Shipley SPR2FX-1.3 photoresist coated on a Si wafer. With a 40 seconds exposure time and a 10 mW/cm² light intensity, a clear pattern was made in the photoresist. These exposure values are the same as requires for the standard chrome mask in this system.

7.2 Sn/In used as an anisotropic etch resist

Silicon anisotropic etching is widely used in the microelectronic and micromachining area. But standard organic photoresist can not work as etch masks for alkaline etching processes such as KOH and TMAH. Our early research showed exposed Bi/In film was highly resistant to alkaline-based anisotropic silicon etching solutions. Hence it has been successfully used as etch mask in alkaline based anisotropic wet etch process [17,18]. Since Sn/In is closely related to ITO, a material with known compatible in Si technology, thus it may be preferably to use Sn in some applications

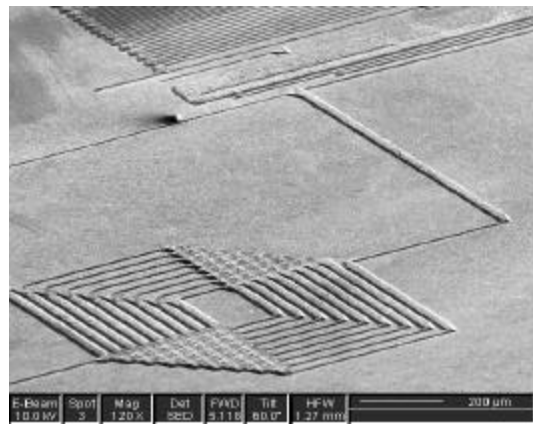


Figure 7. SEM picture of KOH anisotropic etch patterns on Si (100).

Similar to exposed Bi/In film, the exposed Sn/In film can also be used as an etch resist mask in the anisotropic silicon etch process. Detailed research shows that the etch rate of exposed Sn/In in KOH at 85 °C is about 1.4 nm/min and in TMAH is less than 0.5 nm/min. It was also found the etch ratio of Si to exposed Sn/In film can reach 700: 1 in KOH solution. Figure 7 shows the SEM picture of anisotropic etch pattern picture of etched trenches on Si (100) wafer surface. These results suggest that exposed Sn/In film is not only a new direct write photomask material but an anisotropic etch resist mask material.

7.3 Photovoltaic property of Sn/In on p- Si

According to the report [13] that transparent and conductive film like ITO can form a heterojunction with p-Si. In this structure ITO was applied as n type material since most of deposited ITO film exhibits electron rich characteristics. In that structure the photovoltaic effect was obtained so it was used in making solar cell.

Since laser exposed Sn/In film has similar properties to traditional ITO we have tried to use it in those structures. In our experiment the laser exposed Sn/In film was used to replace the traditional ITO in the structure. First, the bilayer of Sn/In film (Sn layer on the top) with thickness of 12/108 nm was sputtered on the chemical cleaned surface of p-Si wafer (2 ohmic-cm).

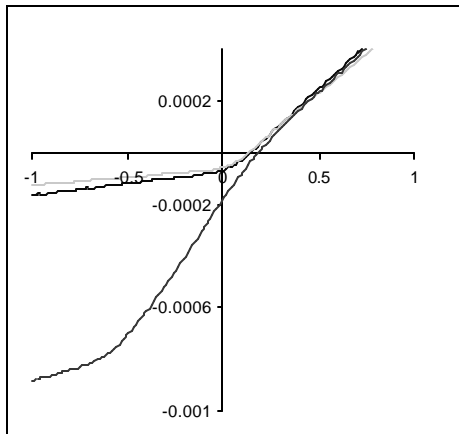


Figure.8 the dark and illuminated I-V characteristic

Aluminum layer with thickness 600 nm was deposited on backside of the wafer for the ohmic contact reason. After this the wafer was loaded into a furnace with the temperature of 300 C for 20 min. the nitrogen gas was filled into the furnace during the alloy process. Then, the 0.4 W argon laser was used to expose the Sn/In film and make a 5mm² window on the surface. Since the exposed Sn/In area become transparent, the illumination light can easily penetrate the top layer and into the interface area. A heterojunction was built by the laser exposed Sn/In film and p-type silicon wafer, in which Sn/In film act as an n-type material. The further study shows that this heterojunction has a photovoltaic characteristic.

Figure 8 shows the dark and illuminated I-V characteristic of the junction of exposed Sn/In with p-Si. From the curve it was noticed that the p-n junction exhibits a typical diode IV characteristic. The open voltage is 0.18V and short current is 0.19 mA. The light-generated current is about 1 mA. In our experiments the illumination source is ordinary light (4 mW/cm²). Since laser exposed film has a high transmission at VU range, it suggests that the junction will be sensitive with the UV.

VIII. CONCLUSION

Laser exposed Sn/In bilayer films can form a transparent and conductive Sn/In oxide film which have the similar properties to ITO. The power used in the process is much lower than that in the laser ablation processes for ITO. The optical density (OD) of SnIn film changed from 3 to 0.22 at 365 nm wavelength after laser exposure, showing it is a good material for direct write photomask application. Profilometry, XRD, SEM and EDX have been used to investigate the film's microstructure and composition. The results confirm that the process of laser exposure is an oxidation process in which Sn, In and oxygen combined at a low threshold temperature. XRD patterns show the laser converted Sn/In film has the similar structures to standard ITO. The electric properties such as sheet resistance and Hall mobility have also been studied. The results show that the film with Sn 10% has the lowest resistivity which is consistent with the composition used in current ITO films. A development etch solution was found that would remove the unexposed Sn/In leaving a patterned ITO layer.

The larger OD change of Sn/In films allows us to make direct write photomasks. Using the photomasks made from Sn/In films, clear patterns with small features have been successfully transferred on to the photoresist. The exposed Sn/In oxide films can also withstand etch attack of alkaline based anisotropic etch solutions, indicating that the film can be used as etch resist mask in microfabrication and micromachining applications. By using exposed Sn/In film as etch resists, different etch trenches have been successfully created on Si (100) wafers.

REFERENCE

- [1] David A. Willis, "thermal mechanisms of laser micromachining of indium tin oxide". Proceedings of SPIE, Volume 5339, 26-29 January 2004.
- [2] Oguz Yavas and Mikio Takai. "Laser processing of materials and structures for flat panel displays"
- [3] James G. Lunney, Richard R. O'Neil, and Karl Schulmeister, "Excimer laser etching of transparent conducting oxides". Appl. Phys. Lett. 59 (6), 5 August 1991.
- [4] Glenn Chapman, Yuqiang Tu, Marinko V. Sarunic, " Bi/In Bimetallic Thermal Resists for Microfabrication, Photomasks and Micromachining Applications" Proc. SPIE Vol. 4690, pp 465-476, San Jose, CA, Mar., 2002.
- [5] Y. Tu, G. Chapman, M. Sarunic, "Bimetallic Thermal Activated Films for Microfabrication, Photomasks and

- DataStorage", Proceedings of SPIE Vol. 4637, pp. 330-340, 2002.
- [6] Glenn Chapman, Yuqiang Tu, "Single Step Direct-Write Photomask Made from Bimetallic Bi/In Thermal Resist", SPIE, V.2 4977A-38 (p.1 of 12) 2003.
- [7] H.Kim, J. S. Horwitz, A. Pique, C. M. Gilmore, D.B. Chrisey. "Electrical and optical properties of indium tin oxide thin films grown by pulsed laser deposition". Appl. Phys. A69 [suppl.] S447 – S 450 (1999) Applied Physics A Materials Science & Processing.
- [8] Sundaram. KB and Bhagavat GK. Phy. Status solidi a, 63, K15.1981.
- [9] John C. C. Fan and John B. Goodenough, " X – ray photoemission spectroscopy studies of Sn -doped indium-oxide films" , J.Appl. Phys. 48, 3524 --3531. 1977.
- [10] K. Enjouji, K. Murata and S. Nishikawa, "The analysis and automatic control of a reactive d.c. magnetron sputtering process" , Thin Solid Films, 108, 1, pp 1--7. 1983.
- [11] Radhouane Bel Hadj Tahar, Takayuki Ban, Yutaka Ohya, and Yasutaka Yakahashi , " Tin doped indium oxide thin films: Electrical properties", Journal of Applied Physics, Volume 83. Number 5. 1998.
- [12] D.K. Schroder: Semiconductor Material and Device Characterization (John Wiley & Sons, New York 1990).
- [13] K B Sundaram and G K Bhagavat, "Chemical vapour deposition of tin oxide films and their electrical properties", J. Phys. D: Appl. Phys., 14, 333-8. 1981.
- [14] JCPDS diffraction database. 1967.
- [15] K. Sreenivas, T. Sudersena Rao, and Abhai Mansingh. "Preparation and characterization of rf sputtered indium tin oxide films", J. Appl. Phys. 57 (2), 15 January. pp. 384 – 392. 1985.
- [16] Kazuhiko Tonooka, Hiroshi Bando, Yoshihiro Aiura, "Photovoltaic effect observed in transparent p-n heterojunctions based on oxide semiconductors", Thin Solid Films 445 (2003) 327 –331.
- [17] Glenn Chapman, Yuqiang Tu, "Single Step Direct-Write Photomask Made from Bimetallic Bi/In Thermal Resist", SPIE, V.2 4977A-38 (p.1 of 12) 2003.
- [18] Glenn Chapman, Yuqiang Tu and Jun Peng, "Bi/In Thermal Resist For Both Si Anisotropic Wet Etching And Si / SiO₂ Plasma Etching", SPIE Micromachining and Microfabrication, Jan., 2004.
- [19] Chuck Wu, US Patent Nos. 4567104, 4670366, 4894303, 5078771 and 5285517.
- [20] Yamazaki Teruhiko, et al, "Manufacture of photomask", Japan Patent No. 55-143560, Aug.11,1980
- [21] A Subrahmanyam and N Balasubramanian, "Studies of the photovoltaic behavior of indium tin oxide (ITO)/silicon junctions prepared by the reactive thermal evaporation technique" Semicond. Sci. Technol. 7 (1992) 324-327.
- [22] K. Moore "Preparation and characterization of rf sputtered indium tin oxide films" J. Phys. D: Appl. Phys., 14, 333-8. 1981.



Jun Peng received the M. Sc degree in Physic from Lanzhou University, CHINA, in 1992. From 1998 to 2001, he worked in Lanzhou microelectronics Ltd.

Haptics: Interaction between Humans and Computers

Haibo Sun

Abstract— In this paper, we introduce a new area of Virtual Reality: Haptics, which is a science about interaction between computers and humans via tactile sensation. Former and current research activities in it are shown in the paper, and different haptic interfaces and their requirements are also discussed. As well, applications of haptic interfaces and future research fields are presented.

Index Terms—Haptics, haptic interfaces, parallel mechanisms, spherical mechanisms.

I. INTRODUCTION

WE live in an era of increased computer use in all fields of life, ranging from personal finance, to healthcare, to manufacturing, and to entertainment. Nowadays computers are not only faster and less expensive, but also communicate with us in more sophisticated ways. The CD-based multimedia has allowed human-computer interaction through text, compressed graphical animation, stereo sound, and live video images.

Virtual Reality (VR) is superior to other forms of human-computer interaction because it provides a real-time environment, integrating several new communication modalities. These include stereo graphics, 3-D sound, tactile feedback, and even taste and smell. However, the realism of current VR simulators is hampered by the lack of good physical simulation, which means the touch and force feedback. These two factors involve the science of Haptics: which is the science of applying tactile sensation to human interaction with computers. In the real world, our primary interface with the physical is our haptic sense (sense of touch).

In order to complete the imitation of real space, one would expect that a haptic interface would be necessary. A haptic interface is a device which allows a user to interact with a computer through receiving tactile feedback. It achieves this feedback by applying a degree of opposing force to the user along the x, y, and z axes. It involves physical contact between the computer and the user, usually through an I/O device, such as a joystick or data gloves, that senses the body's movements. By using haptic devices, users can not only feed information to

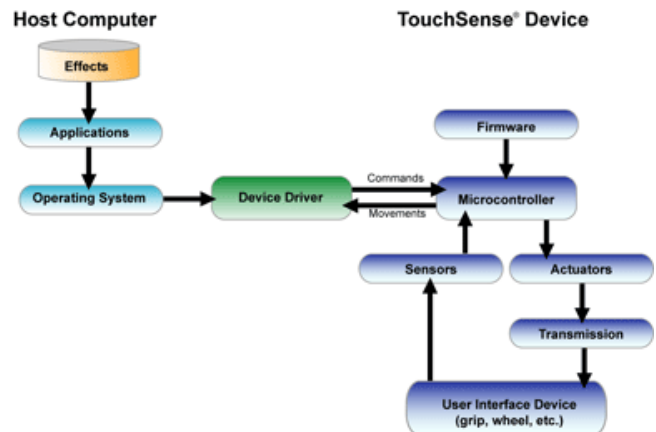


Fig. 1. Illustration of a haptic device system

the computer, but also receive information from the computer in the form of a felt sensation on some part of the body. Fig. 1 gives an illustration of a typical haptic device system.

Haptic interfaces have been in use for over 40 years in a number of applications, ranging from force-feedback teleoperation systems to scientific data visualization to virtual training systems. Due to the variety of applications, haptic interface designs span a wide range, from interfaces with non-dextrous, power grips to many recent designs with more sensitive and dextrous finger grips.

II. RELATED WORKS

Earlier research on integrating force feedback to VR simulations was undertaken by the robotics community, which adapted existing teleoperation servo-arms [1]. Steps towards less complex systems were taken by integrating force feedback to joysticks [2]. For instance, Millman et Al. [3] designed a four degrees of freedom haptic device using a parallel structure. The parallel design enables the direct drive actuators to remain fixed to the ground, thus minimizing inertia.

Iwata et Al. [4] designed a device with six degrees of freedom based on a parallel mechanism to apply reaction forces to the operator's hand. One of the key design aspects of this device is the manipulator handle that is composed by three sets of pantographs instead of linear actuators, to increase workspace volume.

A well known haptic interface with an increased workspace, is the PHANToM, developed by Massie and Salisbury [5]. It is a serial three degrees of freedom interface, that provides translational workspace in three dimensions. The cable drive of the device provides a smooth transmission with relatively low inertia.

Recent specific VR systems and haptic interfaces have been developed for medical applications. One of these simulated systems was developed by Picinbono et al. [6] to explore and sense human tissues. Burdea et al. [7] developed a haptic interface for ankle rehabilitation. Other dedicated haptic devices have been designed, to allow operators to interactively manipulate the modeled objects and execute surgical tasks [8], [9], [10].

Iwata [11] built a 9-DOF Compact Master, which combines a 6-DOF parallel positioning device in series with three 1-DOF finger actuators. In [12], the 6-DOF hybrid SMARTee interface uses a parallel combination of three 2-DOF serial linkages. In [13], a 6-DOF joystick is described, which contains three pantograph linkages and Yokoi *et al.* [14] presents a 6-DOF serial hand controller with three prismatic actuators for translation and three rotary actuators for orientation. A 5-DOF (3 translation, 2 rotation) haptic stylus is described in [15] which uses five linear actuators in parallel. A 4-DOF (3 translations, 1 rotation) hybrid device using rotary actuators is presented in [16] while a 4-DOF Manipulandum that combines a 3-DOF planar parallel robot in series with a linear vertical motion stage is described in [3]. In [5], the hybrid PHANToM has 3 active translational degrees of freedom and 3 passive rotational degrees of freedom. A 2-DOF planar pantograph is optimized in [17] and the 2-DOF linear actuated Magic Mouse is presented in [18]. Finally, Vertut [19] presents a historical survey of earlier hand controllers, articulated arms, and exoskeletons.

In terms of quick motion, the magnetic levitation haptic interface of Berkelman et al. [20] is very appealing, but its workspace is small. Quick motion can be also achieved through a wired system, (e.g., SPIDER [21]). Such devices occupy, however, a large space. Building a compact haptic device with the help of a 6-DOF parallel mechanism is possible, as shown by Long [13]. But in general, fully 6-DOF parallel mechanisms (e.g., Stewart Platform [22], Pantograph Linkage [23], HEXA [24]) are characterized with a restricted workspace for orientation. To increase the workspace, parallel mechanisms with kinematic redundancy can be employed [25], [26].

III. DIFFERENT MECHANISMS USED FOR HAPTIC INTERFACES

A. parallel mechanisms

From the work done by the researchers listed in the last section, we find different mechanisms which could be used for haptic interfaces. Since the master-slave system proposed by Goertz in the 1950's [27], many researchers have developed various types of haptic display, such as the exoskeleton type

master arm by M. Bergamasco [28], PHANToM [5] by Massie and Salisbury, and the MagLev Wrist [20] by Hollis in CMU.

Although many haptic interfaces have been developed as above, serial structures cannot achieve a hard contact feeling. Other devices having a parallel structure are heavy, have relatively small workspaces, and require considerable power because of floating actuators.

To overcome these difficulties, a new mechanism was proposed by Gosselin [29] and Lee, et al. [30] which is lighter by employing non-floating actuators and has a relatively large workspace as compared to other parallel mechanisms.

Parallel mechanisms have many merits as haptic devices due to high structural stiffness, high force bandwidth, high force dynamic range, etc.; however, it is known that parallel reflection performance in haptic applications. At singular points, a haptic system can't completely generate a reflecting force, and moreover, actuator saturation happens. So singularity-free algorithms for maintaining system performance and generating reflecting force, need to be investigated.

Parallel mechanisms are excellent candidates for advanced robotic applications. However, in applications such as space robotics or telescope mechanisms, dynamic balancing is required or at least highly desirable. A mechanism is said to be dynamically balanced if, for any motion of the mechanism, the reaction forces and torques at its base are, at all times, identically equal to zero. In space applications, this condition is required in order to preserve the momentum of the moving base (satellite, space station).

Over the past decades, many researchers have studied the dynamic balancing of the four-bar linkage. At the end of the 60s, the static balancing of four- and six-bar mechanisms was treated by Berkof and Lowen [31], providing the foundations necessary for dynamic balancing. Thereafter, the same authors treated the optimization of the shaking moment of a statically balanced four-bar mechanism [32]. Later, dynamic balancing was obtained using idler loops [33] or with counter-rotations [34]. Another approach described by Abu-Abed and Papadopoulos [35] and Kochev [36] constrains the mechanism to follow a predefined trajectory, chosen in such a way that no forces or moments are generated at the base. Yet another approach is the active balancing proposed in [37] using redundant actuation. Finally, in [38] the dynamic balancing of the four-bar mechanism is carried out without counter-rotations. In the same article, the dynamic balancing of a planar 3-DOF parallel mechanism is also shown by using a cascade of dynamically balanced four-bar mechanisms to build its legs. By doing so, no counter-rotations are necessary. On the other hand, the balancing constraints governing the four-bar linkages are very restrictive, thus complicating the design of such a mechanism. In order to simplify the design, it is proposed here to use five-bar linkages to build the legs of the 3-DOF mechanism. This approach will, however, have the disadvantage of requiring additional mechanical elements to the system (counter-rotations). Nevertheless, in some applications, this compromise might be desirable.

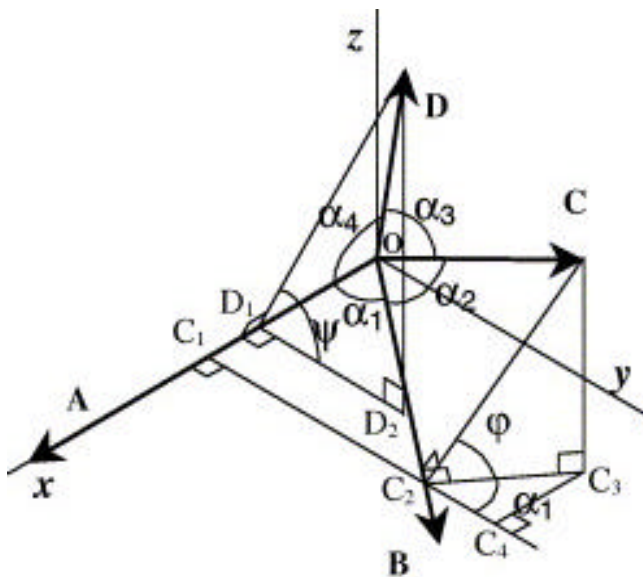


Fig. 2. Position vectors of a spherical four-bar mechanism

B. Spherical mechanisms

Spherical mechanisms are linkages which generate the spherical motion, in which the displacement of any point on the body is constrained to the surface of a sphere. In contrast, planar mechanisms generate two-dimensional motion. For this reason, their design is compatible with using conventional drafting tools while the synthesis of spherical mechanisms is three-dimensional and is not well suited for these drafting techniques. The spherical mechanism designer must be able to visualize the entire problem in three-dimensions, and computer graphics can be an effective tool for providing this necessary visualization. Efforts have been made to create computer graphics based software packages for spherical four-bar mechanism designs.

Both planar and spherical linkage mechanisms are constrained mechanisms in three-dimensional Euclidean space \mathbb{R}^3 . Their links are constrained to move on planar and spherical surfaces, respectively. Two independent coordinates are necessary and sufficient to describe a point on both planar and spherical surfaces. From this point of view, planar and spherical mechanisms are similar. On the other hand, moving links of both spherical and spatial mechanisms generate three-dimensional movements. That is to say, spherical mechanisms are an intermediate stage between planar and spatial mechanisms [39]. Actually, because all of their joint axes intersect at infinity, planar mechanisms are a special case of spherical mechanisms. Similarly, because all of their joint axes intersect at the same point, spherical mechanisms are a special case of spatial mechanisms. Applications of spherical mechanisms to robotic mechanical systems and to rehabilitation treatment of the anatomic joints, have been proposed by Chablat and Angeles [40].

In Fig. 1 a general spherical four-bar mechanism is shown. Vectors of revolute pairs of spherical four-bar mechanism were

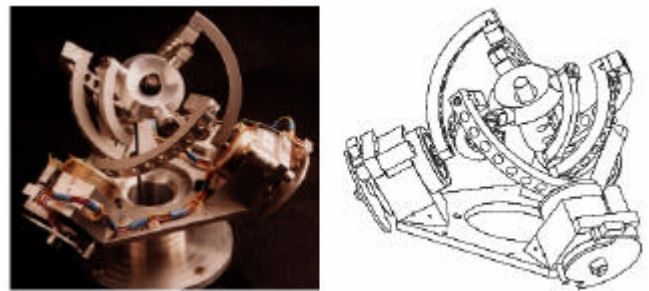


Fig. 3. The Agile Eye

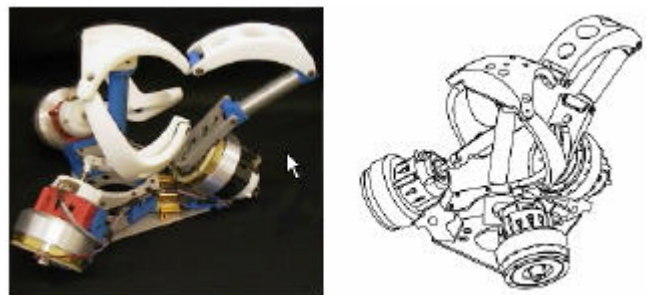


Fig. 4. The 3 DOF haptic device “SHaDe”

shown as position vectors and they are named as **A**, **B**, **C** and **D**, respectively. Vector **A** has been aligned to the *x*-axis, vector **B** has been put on *x*-*y* plane and all vector lengths are assumed as 1 to simplify the problem. Because the locus of points in space having a given fixed distance from a given point *O*, is viewed as a sphere, all heads of vectors **A**, **B**, **C** and **D**, are on the surface of the sphere. To derive one simple equation using the terms a_1, a_2, a_3, a_4, Φ and ψ , some perpendicular lines to coordinate axes were dropped and formed right-angled triangles.

C.M. Gosselin’s group at University of Laval has done much work on haptic devices design using spherical mechanisms. They developed a 3-dof haptic device: “SHaDe” [41]. As shown in Fig. 4, it allows a human operator to control motions while being subjected to force feedback. The mechanism presents the particularity of having only three degrees of freedom, leading to a simpler design and a more ergonomic utilization. Because of the spherical geometry, such a haptic device has several advantages, namely, a pure rotation around a point located inside the user’s hand (no translations at this point), a large workspace, a comfortable usage, and precise manipulation with arm resting. In Fig. 3, Gosselin’s group also developed a 3-DOF 3-RRR spherical parallel manipulator “Agile Eye” [42], used for the rapid orientation of a camera. Its mechanical architecture allows the achievement of high operational speeds. One important feature of it, is that the singularity loci of the agile eye are independent from the chosen branch set.



Fig. 5. The haptic pen proposed by Stocco et Al.

IV. REQUIREMENTS FOR HAPTIC INTERFACES

Haptic devices present a difficult mechanical design problem, which includes a lot of requirements: as P. Fischer et Al. pointed out [43] that, to achieve good static and dynamic performance, such as workspace, force transmissibility, isotropy of the force and motion, backdrivability, high force bandwidth, and high force dynamic range.

A brief explanation of some of these requirements is given next. “Workspace” means the range of motion of a haptic device, like the largest displacement of the end-effector from the center of the mechanism; “Force Transmissibility” is the ratio of the transmitted force (output force) to the input force; “Backdrivability” is the ability of the device to drive the input from the output; and “Force Bandwidth” means the frequency range of input and output forces. “Isotropy” here, means that forces and motions of the device are identical in all directions of the workspace.

A haptic device generates feedback by presenting an impedance to the human hand. Generally, “feedback” means that, when you try to move the device (like moving the finger tip), you will feel force or torque preventing you from doing this, if not following the designated route. And “Impedance” here means that displacement of mechanisms result in the generation of forces or torques. Stocco et Al. [44], just as Fig. 5 above, show that at present 2 types of such devices can be used to implement an intelligent variable impedance: an impedance device and an admittance one. The former one is easily backdrivable in its

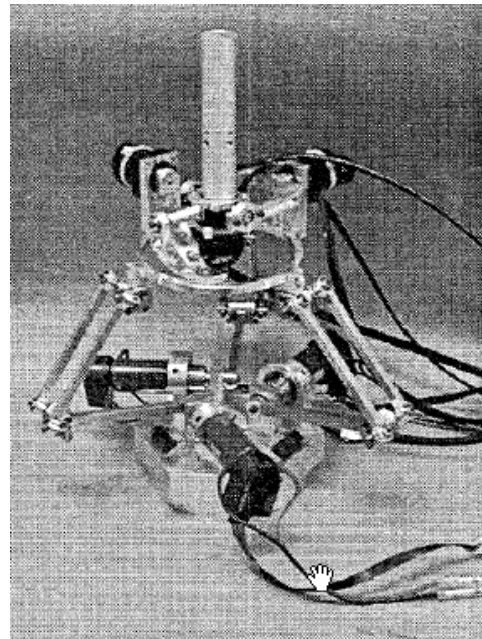


Fig. 6. Overview of a 6-DOF haptic device

passive state, and can use direct drive or moderately geared actuators to adjust its effective impedance by applying a force, based on position and velocity.

V. APPLICATIONS OF HAPTIC INTERFACES

The previous sections described advantages and disadvantages of different mechanisms for haptic devices and the requirements of them, and all of these have prompted developers to look for applications in various fields.

One of the domains, where haptic devices are already attracting significant interest, is medical training, especially surgical training. There has been an extraordinary increase of medical knowledge in the second half of last century. This prompted medical schools to modernize their teaching methodology for topics ranging from human anatomy to physiology and pathology. Interactive CD-ROM anatomy trainers have replaced earlier static atlases, and are widely used today. More recently, the first VR-based teaching systems have been developed. Surgical training systems, especially for minimally invasive surgery (MIS) have also become commercially available. These systems use head-mounted displays (HMD) to immerse the student or resident inside a virtual patient, providing a very intuitive and powerful instruction environment. However, the present VR MIS training systems lack haptic feedback, which limits their usefulness. This deficiency makes advanced surgical training impractical or the teaching of related areas such as diagnosis, anesthesiology, or rehabilitation procedures. Some of the aspects of the application in medical training include tissue palpation, anesthesia, minimally invasive surgery and microsurgery.

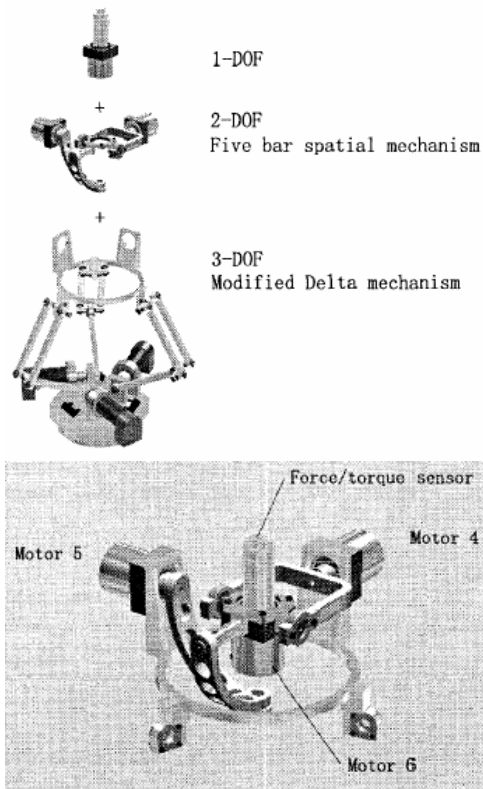


Fig. 7. Distribution of degrees of freedom of the above device and its orientation sub-system

Another application domain for haptic devices is VR entertainment. This is a multibillion dollar industry, which has traditionally used visual and sound effects, in response to the user's position input. At the present time, force and tactile feedback are being added to various video games to enhance their interactivity and realism. Fig. 6 shows a typical parallel 6-DOF haptic device and Fig. 7 gives the distribution of its 6 DOFs.

In previous sections we have shown many robots used as haptic feedback devices, so another important application area is telerobotics, where haptic feedback can enhance the operator's telepresence, overcome stability problems (due to arm dynamics or time delay), and improve overall task performance. Virtual haptic feedback can help program the robot arm locally or at a distance as well.

The above application list is by no means exhaustive. Other potential domains are education, business, architectural design, and mechanical design. The spectrum of VR simulations using haptic feedback will increase as the underlying technology matures and its price becomes more affordable.

VI. FUTURE RESEARCH

In the foreseeable future, major improvements in feedback actuators, sensors, and computing hardware will lead to miniaturization, less cumbersome haptic interface devices, and an increase in the user's safety and freedom of motion. These improvements should in turn lead to more natural, realistic, and

useful simulations.

Now there is a clear need for a small, powerful, responsive, clean, silent, and safe actuator. Researchers are currently investigating a number of novel technologies, that may result in such an actuator. The miniaturization of haptic feedback actuators will make possible the construction of light and comfortable full-body suits, and these will replace present upper-body prototypes, which are very complex, heavy, and expensive. Basic technological advances in actuator miniaturization, light and long-life power sources, and miniature and wearable computers will make tethering of haptic feedback suits unnecessary. This will open the door for unencumbered and natural simulations, over a large physical space and with multiple participants.

The above discussion on the future of haptic feedback devices concentrated on technology advances and left out the sociological impact these may have. The whole area of force and haptic feedback is under active research worldwide, and it is difficult to predict exactly where the current advances will lead. Nevertheless, it is safe to say that, if used correctly and fully, force and haptic feedback devices should be very beneficial.

REFERENCES

- [1] G. Burdea and P. Coiffet, "La réalité virtuelle", Hermès, 1993.
- [2] G. Berdea, "Force and touch feedback for virtual reality", John Wiley and sons, 1996.
- [3] P. Millman and J. E. Colgate, "design of four degree of freedom force reflection manipulandum with specified force/torque workspace", IEEE Intl. conference on robotics and automation, Sacramento, CA April 1991, pp 1488-1493.
- [4] H. Iwata, "Pen based haptic virtual environment", IEEE Annual virtual reality international symposium, Piscataway, NJ, 1993 pp 287-292.
- [5] T. Massie and k. Salisbury, "PHANTOM haptic interface a device for probing virtual objects", Dynamic systems and control ASME, New York, NY, 1994, pp 295-299.
- [6] G. Picinbono and J.C. Lombardo, "Extrapolation: a solution for force feedback", Colloque scientifique international réalité virtuelle et prototype, 3-4 Juin 1999, Laval France.
- [7] <http://www.caip.rutgers.edu/vrlab/ankle.html>
- [8] <http://robotics.eecs.berkeley.edu/mcenk/medical/#endo>
- [9] <http://rcs.ee.washington.edu/BRL/devices/surgical/>
- [10] <http://iregtl.iai.fzk.de/DYNAMIK/html/hifec.html>
- [11] H. Iwata, "Artificial reality with force-feedback: Development of desktop virtual space with compact master manipulator", Proc. SIGGRAPH, vol.24, no. 4, pp. 165-170, Aug. 1990.
- [12] K. Cleary and T. Brooks, "Kinematic analysis of a novel 6-DOF parallel manipulator", in Proc. IEEE Intl. Conf. Robotics and Automation, May 2-6, 1993, pp. 708-713.
- [13] G. L. Long and C.L. Collins, "A pantograph linkage parallel platform master hand controller for force-reflection", in Proc. IEEE Intl. Conf. Robotics and Automation, vol. 4, May 9-12, 1994, pp. 3211-3216.
- [14] H. Yokoi, J. Yamashita, Y. Fukui, and M. Shimojo, "Development of 3D-Input device for virtual surface manipulation", in Proc. 3rd IEEE Int. Workshop Robot and Human Communication, July 18-20, 1994, pp. 134-139.
- [15] C. D. Lee, D. A. Lawrence, and L.Y. Pao, "Guaranteed convergence rates for five degree of freedom in parallel haptic interface

- kinematics”, in Proc. IEEE Int. Conf. Robotics and Automation, May 10-15, 1999, pp. 3267-3274.
- [16] T. Kotoku, K. Komoriya, and K. Tanie, “ A force display system for virtual environments and its evaluation” , in Proc. IEEE Int. Workshop Robot and Human Communication, Sept. 1-3, 1992, pp. 246-251.
- [17] V. Hayward, J. Choksi, G. Lanvin, and C. Ramstein, “ Design and multi-objective optimization of a linkage for a haptic interface” , in Proc. ARK '94, Int. Workshop Advances Robot Kinematics, June 1994, pp. 352-359.
- [18] A. J. Kelley and S. E. Salcudean, “ The development of a force-feedback mouse and its integration into a graphical user interface” , in Proc. Int. Mechanical Engineers Congr. Expo., vol. DSC-55-1, Nov. 6-11, 1994, pp. 287-294.
- [19] J. Vertut, “ Control of master slave manipulators and force feedback” , in Proc. 1977 joint Automation Control Conf., San Francisco, CA.
- [20] P. J. Berkelman, R. L. Hollis and S. E. Salcudean, “ Interacting with Virtual Environment Using a Magnetic Levitation Haptic Interface” , Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, Pittsburgh, Pennsylvania, 1995, pp. 117-122.
- [21] Y. Hirata and M. Sato, “ 3-Dimensional Interface Device for Virtual Workspace” , Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, Raleigh, North Carolina, 1992, pp. 889-896.
- [22] D. Stewart, “ A Platform with Six degrees of Freedom” , Proc. of the Institution of Mechanical Engineers 1965-1966, vol. 180, Part 1, no. 15, 1965, pp. 371-386.
- [23] H. Inoue, Y. Tsusaka and T. Fukuizumi, “ Parallel Manipulator” , Robotics Research, The Third International Symposium, O. Faugeras and G. Giralt (Eds), The MIT Press, 1986, pp. 321-327.
- [24] F. Pierrot, M. Uchiyama, P. Dauchez and A. Fournier, “ A New Design of a 6-DOF parallel Robot” , J. of Robotics and Mechatronics, vol. 2, no. 4, 1991, pp. 308-315.
- [25] T. Asano, H. Yano and H. Iwata, “ Basic Technology of Simulation System for Laparoscopic Surgery in Virtual Environment with Force Display” , Medicine meets Virtual Reality, Global Healthcare Grid, Studies in Health Technology and Informatics, IOS Press, vol. 39, 1997, pp. 207-215.
- [26] V. Hayward, “ Toward a Seven Axis Haptic Device” , Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, Pittsburgh, Pennsylvania, 1995, pp. 133-139.
- [27] R. C. Goertz, “ Fundamentals of General-Purpose Remote manipulation” , Journal of Nucleonics, no. 11, 1952, pp. 36-42.
- [28] M. Bergamasco, et al., “ An Arm Exoskeleton System for Teleoperation and Virtual Environments Applications” , Proc. of Intl. Conf. on Robotics and Automation, San Diego, CA, 1994, pp. 1449-1454.
- [29] Gosselin, C.M., Allan, J.F., and Lalibert, T., “ A New Architecture for the JFToMM Tenth World Congress on the Theory of Machine and Mechanisms, Oulu, Finland, 1999, pp. 1140-1145.
- [30] Lee, J.H., Eom, K.S., Yi, B.J. and Suh I.H., “ Design of A New 6-DOF Parallel Haptic Device” , Proc. of Intl. Conf. on Robotics and Automation, 2001, pp. 886-891.
- [31] Berkof, R.S., and Lowen, G.G., “ A new method for completely force balancing simple linkage” , ASME J. Eng. Ind., 1969, pp. 21-26.
- [32] Berkof, R.S., and Lowen, G.G., “ Theory of sheaking moment optimization of force-balanced four-bar linkages” , ASME, J. Eng. Ind., 1971, pp. 53-60.
- [33] Bagci, C., “ Complete balancing of linkage using complete dynamical equivalents of floating links: CDEL Method” , in DE Flexible Mechanisms, Dynamics and Analysis, 22nd Biennial Mechanisms Conf. vol. 47, 1992, pp. 477-488.
- [34] Tsay, D.M., Ho, H.C., and Wang, K.C., “ Design of torque balancing cams for globoidal cam indexing mechanisms” , ASME J. Mech. Des., 2002, 124, pp. 441-447.
- [35] Abu-Abed, A., and Papadopoulos, E., “ Design and motion planning for a Zero-Reaction manipulator” , IEEE Intl. Conf. on Robotics and Automation, 1994, San Diego, CA, pp. 1554-1559.
- [36] Kochev, I.S., “ Full shaking momentbalancing of planar linkages by a prescribed input speed fluctuation” , Mech. Mach. Theory, 1990, 25, pp. 459-466.
- [37] Yu, Y.Q., and Lin, J., “ Active balancing of a flexible linkage with redundant drives” , ASME J. Mech. Des., 2003, 125, pp. 119-123.
- [38] Ricard, R., and Gosselin, C.M., “ On the development of reactionless parallel manipulators” , Proc. of the ASME 26th Biennial Mechanisms and Robotics Conf., Baltimore, MA, 2000, No. MECH-14098.
- [39] A. Castellet and F. Thomas, Characterization of the self-motion set of the orthogonal spherical mechanism, *Mech. Mach. Theory* **34** (1999), pp. 59–88.
- [40] D. Chablat and J. Angeles, The computation of all 4R serial spherical wrists with an isotropic architecture, *ASME J. Mech. Design* **125** (2003), pp. 275–280.
- [41] [Lionel Birglen](#), [Clément M. Gosselin](#), [Nicolas Pouliot](#), [Bruno Monsarrat](#) and [Thierry Laliberté](#), SHaDe, a new 3-dof haptic device, *IEEE Transactions on Robotics and Automation*, 2002, Vol. 18, No. 2, pp. 166-175.
- [42] C.M. Gosselin, E.S. Pierre and M. Gagné, On the Development of the Agile Eye, *IEEE Robotics & Automation Magazine*, Dec. 1996, pp.29-37.
- [43] P. Fischer, R. Daniel, and K.V. Siva, “Specification and design of input devices for teleoperation,” in *Proc. IEEE Int. Conf. Robotics and Automation*, 1990, pp. 540-545.
- [44] Leo J. Stocco, Septimiu E. Salcudean and Farrokh Sassani, “Optimal Kinematic Design of a Haptic Pen,” *IEEE/ASME Mechatronics* 2001, Vol. 6, No. 3, pp. 210-220.
- [45] Burdea G. C., “Force and Touch Feedback for Virtual Reality”, John Wiley & Sons, INC., 1996.
- [46] Craig J. J., “ Introduction to Rototics”, Addison-Wesley Publishing Company, Inc., 1986.
- [47] Faraz A. and Payandeh S., “ Engineering Approaches To Mechanical and Robotic Design for Minimally Invasive Surgery” , Kluwer Academic Publisher, 2000.



Haibo Sun received the B.Eng. and M.Eng. degrees in Mechanical Engineering from Kunming University of Science and Technology, Kunming, Yunnan, China in 2002 and 2004. Now he is a M.A.Sc student at the School of Engineering Science, Simon Fraser University, Burnaby, B.C., Canada. His research interests include haptic devices design, robotics and mechanism kinematic analysis.

Basic Issues in Fault Detection and Diagnosis for Dynamic Systems

Esmail Tafazzoli

Abstract—Fault detection and diagnosis is the process of determining the existence of faults in a system as well as their type, size, and location. Symptom generation and symptom evaluation are two key steps in every fault detection and diagnosis system. This paper introduces these steps and presents the main developed methods in systems fault detection and diagnosis. For the first step, we focus on Parameter estimation, observer state estimation, and parity equations as the basic methods of model-based fault detection. For the second step, we introduce fault diagnosis methods, such as fuzzy rule-based systems and neural networks. Finally, we present the concept of robustness, which is one of the most important issues in fault detection.

Index Terms—Fault detection, fault diagnosis, symptom generation.

I. INTRODUCTION

Fault detection and diagnosis (FDD) is the most important issue in technical process supervision. Detecting system abnormalities in the early stages prevents severe damage and further extra costs. By the rapid development of complex processes and control systems, the need for accurate fault detection systems with high reliability has received considerable attention.

The theoretical development of FDD began in the early 1970s. In 1976, Willsky [1] presented the first major survey paper in FDD, which was followed by a more recent survey by Isermann [2], who summarized the development of fault detection methods based on modeling, parameter estimation, and state estimation. Later, Gertler [3] reviewed techniques for detecting and isolating faults in complex plants. Further efforts were contributed by Patton and Chen [4] in the field of FDD, analyzing parity equation methods.

During the last twenty years, many different methods have been proposed and developed for FDD systems. Based on publications in the last five years, it can be seen that parameter estimation and observer-based methods are the most frequently applied techniques for fault detection. In addition, the use of fuzzy methods as well as Neural Networks (NN) has increased in the fields of systems modeling, classification, and rule-based reasoning [5].

II. FAULT DETECTION AND DIAGNOSIS

A. Definitions

The terminology in the field of FDD is not consistent. Therefore, in order to avoid further misunderstandings, some definitions frequently used in the literature are given below [5]:

1) States:

–Fault: An unpermitted deviation of at least one characteristic property or parameter of the system from the acceptable, usual, or standard condition.

–Symptom: A change of an observable quantity from normal behavior.

–Disturbance: An unknown and uncontrolled input, acting on the system.

–Residual: A fault indicator, based on the deviation between measurements and model-based computation.

2) Functions:

–Fault detection: Determining the presence of faults in a system and the time of detection.

–Fault diagnosis: determining type, size, and location of faults in the system.

The task of the fault detection and diagnosis system is divided into two major phases: 1) analytic (or heuristic) symptom generation and 2) fault diagnosis. In the first phase, the detection component of the FDD system determines the existence of any fault in the process through residual generation and change detection methods, which is followed by symptom generation. Once the symptoms are obtained, the diagnostic part of the FDD system determines the type, size, and location of the fault by using different methods (e.g., classification, clustering, etc.). Fig. 1 shows the overall scheme of the FDD procedure.

B. Analytic symptom generation

The analytic knowledge about the process is used to generate symptoms for fault diagnosis. This task can be performed by different methods, depending on the available process information. Residuals can be produced by:

1) limit value checking of measurable signals.

2) signal analysis of directly measurable variables, using signal models (e.g., correlation functions, frequency spectra,

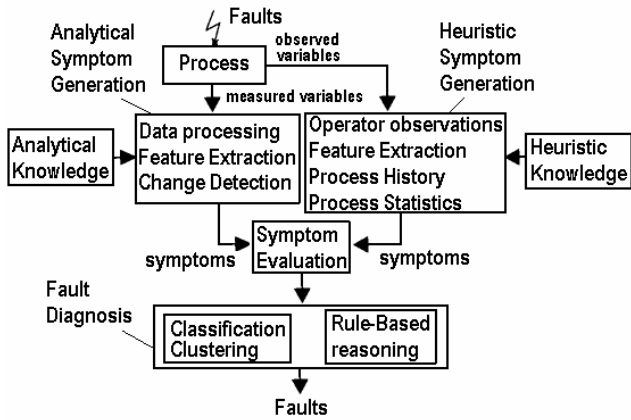


Fig. 1. Fault detection and diagnosis scheme. Any possible information is processed in order to achieve the best result [6].

autoregressive moving average), or obtaining characteristic values (e.g., variances, amplitudes, frequencies, or model parameters)

3) process analysis, using mathematical process models together with parameter estimation, state estimation, and parity equation methods. Here, the characteristic values are parameters, state variables, or residuals [6].

When residuals are obtained, some features are extracted and compared with the normal features of the fault-free process. Subsequently, any change in the measured signal, signal model, or process model is considered as an analytic symptom.

C. Heuristic symptom generation

In addition to analytic symptom generation, which uses quantifiable information, heuristic symptoms can be generated by using qualitative knowledge from human information [8]. The qualitative knowledge includes operator observations and inspections in the form of color, acoustic noise, wear, tear, or vibration. Another source of heuristic knowledge is the process history in the form of maintenance performance, repairs, and lifetimes. Moreover, statistical data obtained from experience can be considered as qualitative knowledge. Based on these information sources, heuristic symptoms are produced in the form of linguistic variables (e.g., small, medium, large).

III. MODEL-BASED FAULT DETECTION

Based on recent publications, model-based fault detection systems are the most frequently applied methods [7]. Therefore, in this paper, we present three basic methods of model-based fault detection in detail, followed by a brief introduction to fuzzy and NN methods. The basic approaches in model-based fault detection are known as:

- 1) *Parameter estimation approach*
- 2) *Observers approach*
- 3) *Parity equation approach*

The task of the model-based fault detection system via these approaches is to identify any changes in the process and generate symptoms by using a mathematical model of the

process. Generating effective residuals from the process information is the main phase in the procedure of fault detection. Fig. 2 describes the basic structure of the model-based fault detection system.

A. Residual generation via parameter estimation

In this case, the idea is based on estimating actual process parameters and comparing them with the reference model parameters. The process parameters are estimated by measuring the input and output signals and applying a parameter estimation method. The occurrence of any fault will affect the process parameters and will appear in the form of parameter changes. Therefore, any discrepancy would indicate the presence of a fault in the process.

The process can be presented in the form of a transfer function as:

$$G = \frac{B(s)}{A(s)}$$

and the model as:

$$G = \frac{\hat{B}(s)}{\hat{A}(s)}$$

So the error would be:

$$e = \hat{B}(s)u(s) - \hat{A}(s)y(s)$$

where u and y are the input and output of the system. The system parameters are estimated by minimizing the error (e.g., using least square method). Fig. 3 shows two different approaches for generating residuals by using model-based parameter estimation. In this figure, Θ is the parameter vector. If any fault affects the parameters, then the estimator would indicate the parameter changes.

B. Residual generation via observers

The main idea behind the observer techniques is to estimate the outputs and states of the process from available

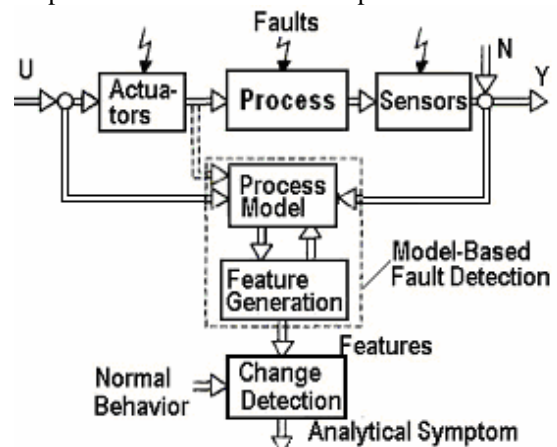


Fig. 2. General Model-Based fault detection procedure[6].

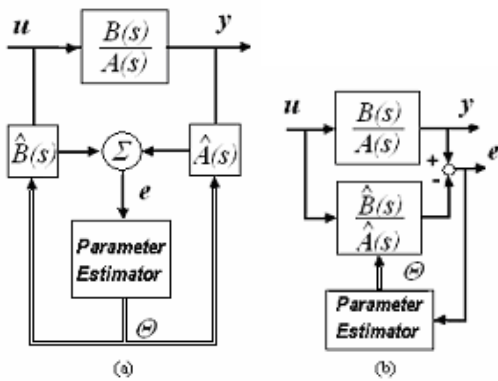


Fig. 3. Residual generation using parameter estimation [8]:
a) equation error, b) output error.

measurements. Therefore, the weighted values of the estimated errors are used as residuals. Consider the process representation in the state space form as below:

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (1)$$

$$y(t) = Cx(t) \quad (2)$$

where A , B and C are process parameter matrices. The state observer builds unmeasurable state variables from input and output. The following equations defined the observer as:

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + He(t) \quad (3)$$

$$e(t) = y(t) - C\hat{x}(t) \quad (4)$$

where H is the observer gain matrix and $e(t)$ is the output error. Fig. 4 illustrates the observer structure and the error construction scheme. If the observer is stable, in a fault-free condition, the error asymptotically reaches zero. Whenever a fault appears in the process, the equations change as (5) and (6), which will result in a nonzero output error.

$$\dot{x}(t) = Ax(t) + Bu(t) + L_i f_i(t) + Mv(t) \quad (5)$$

$$y(t) = Cx(t) + L_o f_o(t) + Nw(t). \quad (6)$$

Fault signals at the input and output of the process are denoted by f_i and f_o , acting through matrices L_i and L_o . In addition, nonmeasurable disturbances are represented by v and w in (5) and (6). Letting $v(t)=w(t)=0$ and $\tilde{x} = \dot{x} - \dot{\hat{x}}$, the state estimation error can be computed from (4), (5), and (6) as:

$$\dot{\tilde{x}}(t) = [A - HC]\tilde{x}(t) + L_i f_i(t) - HL_o f_o(t) \quad (7)$$

and the output error would be:

$$e(t) = C\tilde{x}(t) + L_o f_o(t). \quad (8)$$

Both the state variables, $x(t)$, and output error, $e(t)$ are influenced by faults; therefore, they can be chosen as residuals (see Fig. 4).

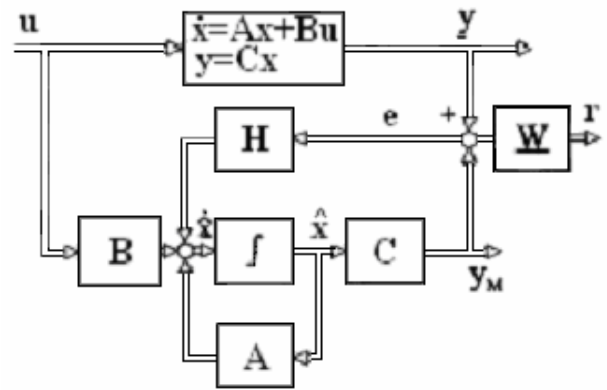


Fig. 4 State Observer schematic [6].

The faults illustrated above that affect the system by being added to the output or input, are called additive faults. In another way, if they appear in the process parameters A , B or C , then the residual equations become:

$$\dot{\tilde{x}}(t) = [A - HC]\tilde{x}(t) + [\Delta A - H\Delta C]x(t) + \Delta Bu(t) \quad (9)$$

$$e(t) = C\tilde{x}(t) + \Delta Cx(t) \quad (10)$$

where ΔA , ΔB and ΔC are the faults appearing in the form of a parameter change. Since these changes are multiplied to the process states, they are called multiplicative faults.

C. Residual generation via output observers

An alternative way of generating residuals involves the use of output observers to estimate the output signal. In this case, the estimation of state variables is not of primary interest. Compare Fig. 4 with Fig. 5 for the difference between the output observer and state observer.

D. Residual generation via parity equations

In this class of residual generators, the idea is to run a process model in parallel with the process itself and to obtain the difference between the model and process [8].

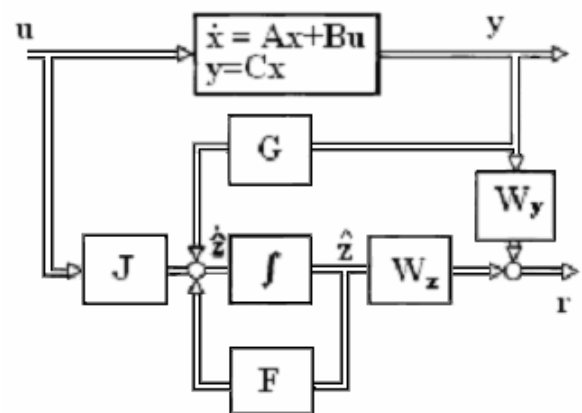


Fig. 5. Output observer[6].

A linear process is described by the transfer function

$$G_p(s) = \frac{y(s)}{u(s)} = \frac{B(s)}{A(s)} \quad (11)$$

and the process model is assumed to be

$$G_m(s) = \frac{y_m(s)}{u_m(s)} = \frac{B_m(s)}{A_m(s)}. \quad (12)$$

where m stands for model and p denotes process. Based on equations (11) and (12), if the input and output of the process are affected by fault signals, $f_u(t)$ and $f_y(t)$, then the error can be obtained as:

$$y(s) = G_p(s)u(s) + G_p(s)f_u(s) + f_y(s) \quad (13)$$

$$e(s) = y(s) - y_m(s) = G_p(s)f_u(s) + f_y(s) \quad (14)$$

$$e(s) = A_m(s)y(s) - B_m(s)u(s) \quad (16)$$

$$e(s) = B_p(s)f_u(s) - A_p(s)f_y(s). \quad (17)$$

The error deviates from a zero state when any fault affects the process. The results of equations 14 and 17 are called parity equations [9]. A schematic of the parity equation method is shown in Fig. 6.

E. Fuzzy and Neural Network Modeling

One approach toward residual generation is the use of fuzzy modeling techniques. Fuzzy models can be seen as logical models with IF-THEN rules as the core elements of the model structures. This property of the fuzzy models allows engineers to build models based on natural language statements, which are more transparent, compared to mathematical models. In addition, fuzzy methods are capable of handling uncertain information in modeling nonlinear systems. Moreover, fuzzy models can be regarded as flexible mathematical structures that approximate a large class of nonlinear systems to a desired degree of accuracy[5], [8], [11].

The combination of fuzzy methods with neural networks is another alternative in fault detection approaches.

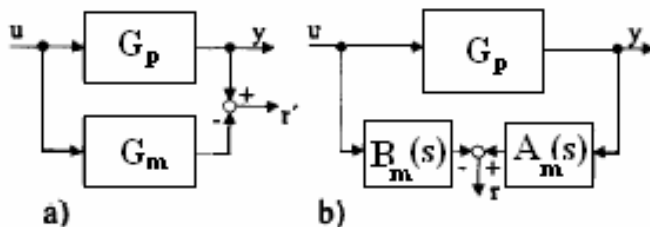


Fig. 6 Parity equations method [8]. a) Generating output error, b) Generating equation error.

NNs are known for approximating any nonlinear dynamic function. Further, online training is a significant property of NNs that allows changing the fault detection system when changes are applied to the process operating conditions. In addition, NNs can be used when no mathematical model of the system is available.

A disadvantage of using NNs is the lack of transparency in modeling systems. In fact, a NN system operates as a black box with no quantitative/qualitative information available. In contrast, fuzzy systems offer a linguistic model of the system dynamics that can be easily interpreted and analyzed. Therefore, combining fuzzy systems and NN (neuro-fuzzy systems) would improve the properties of the system [10], [11].

IV. FAULT DIAGNOSIS METHODS

When the fault detection phase is completed, the second phase of the FDD system (Fault diagnosis) continues the procedure. At this phase, analyzing the generated symptoms determines the type, size, and location of the fault. If several symptoms change differently for certain faults, then the first way of determining a fault is to use classification methods to indicate changes of symptom vectors. Some classification methods are:

- Geometrical distance and probabilistic methods
- Artificial neural networks methods
- Fuzzy clustering methods

If more information about symptom-fault relations is available, then methods of reasoning can be applied. Diagnostic models exist in the form of symptom-fault causalities. These causalities are expressed as IF-THEN rules; therefore, both analytic and heuristic symptoms can be processed concurrently [7], [8].

When both heuristic and analytic symptoms are available, using a fuzzy system is an appropriate choice, because these systems can process both kinds of symptoms. In fuzzy FDD, the generated symptoms are fuzzified and then evaluated by an inference mechanism which consists of IF-THEN rules. Finally, the results are defuzzified in order to be used for further actions. Fig. 7 illustrates the general steps of symptom evaluation and fault diagnosis with a fuzzy system.

V. ROBUSTNESS IN FAULT DETECTION SYSTEMS

Model-based Fault detection systems use mathematical models of the process. Therefore, the accuracy of the FDD system depends on the accuracy of the model. The better the system is modeled, the better its reliability and performance would be. However, modeling errors and disturbances can not be completely eliminated. Hence, robust method seems to be necessary. A robust fault detection system is only sensitive to faults, even in the presence of disturbances or modeling errors.

Robustness can be achieved by producing residuals that are insensitive to uncertainties. Both uncertainties and faults affect the residuals, so the task of an FDD system is to discriminate between these two by de-coupling disturbances

REFERENCES

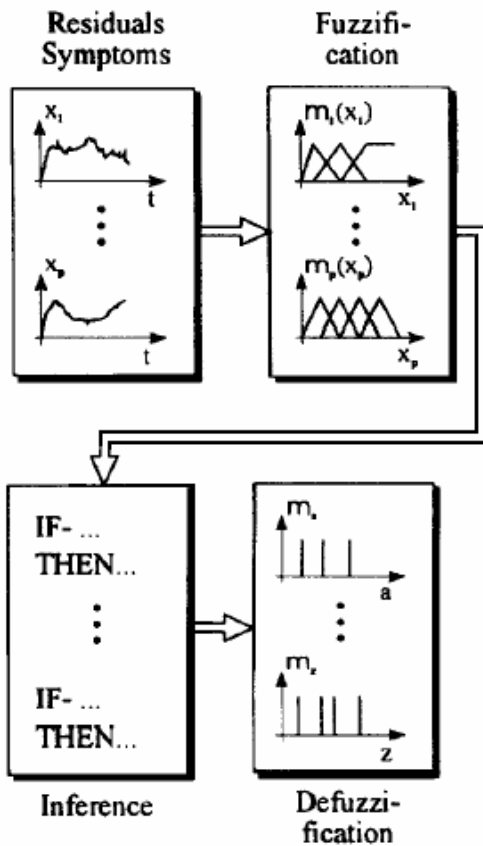


Fig. 7. General steps of symptom evaluation with fuzzy fault diagnosis system [12].

from the residuals. A complete de-coupling would be the best case. However, if a perfect de-coupling is not achievable, one can obtain an approximate de-coupling by minimizing a performance index.

In the case of modeling errors, the problem is more difficult because modeling errors appear in the same way that multiplicative faults do. One solution to this problem is the use of adaptive thresholds at the decision making stage. Another solution is to find an approximate structure for uncertainties and then consider the approximate structure as disturbances and apply residual de-coupling methods [13].

VI. CONCLUSION

This paper has presented basic fault detection and diagnosis concepts and methods. Three major methods for generating residuals have been introduced in model-based fault detection disciplines. These methods include parameter estimation, state variable estimation, and parity equations that are used for symptom generation. A fuzzy modeling approach has been addressed as an alternative modeling method, especially, for generating heuristic symptoms. Fault diagnosis definition and methods of symptom evaluation have been introduced. In this category, A fuzzy symptom evaluation structure has been addressed. Finally, a brief definition of robustness in Fault detection systems has been presented and discussed.

- [1] A. S. Willsky, "A survey of design methods for failure detection systems," *Automatica*, vol. 12, pp. 601–611, 1976.
- [2] R. Isermann, "Process fault detection based on modeling and estimation methods – a survey," *Automatica*, vol. 20, no. 4, pp.387–404, 1984.
- [3] J. Gertler, "Survey of model-based failure detection and isolation in complex plants," *IEEE Control system magazine*, page 3–11, 1988.
- [4] R. Patton, J. Chen, "A review of parity space approaches to fault diagnosis for aerospace systems," *AIAA Journal of Guidance, Control and Dynamics*, vol. 17, no.2, 278–285, 1994.
- [5] S. Simani, C. Fantuzzi, R. J. Patton, *Model-based fault diagnosis in dynamic systems using identification techniques*. London: Springer, 2003, ch. 2.
- [6] R. Isermann, "Supervision, fault detection and fault diagnosis methods: an introduction," *Control Engineering Practice*, vol. 5, no. 5, 639–652, 1997.
- [7] R. Isermann, P. Balle, "Trends in the application of model-based fault detection and diagnosis of technical processes," *Control Engineering Practice*, vol. 5, no. 5, 709–719, 1997.
- [8] R. J. Patton, P. M. Frank, R. N. Clark, *Issues of fault diagnosis for dynamic systems*. London: Springer, 2000, ch. 1.
- [9] J. Gertler, "Generating directional residuals with dynamic parity equations," *Proc. IFAC/IMACS Symp. SAFEPROCESS*, Baden Baden Germany, 1991.
- [10] R. J. Patton, C. J. Lopez, F. J. Uppal, "Artificial intelligence approaches to fault diagnosis," *Journal of applied mathematics and computer science*, vol. 9, no. 3, 471–518, 1995.
- [11] L. Wang, *A course in fuzzy systems and control*. NJ: Prentice Hall International Inc., 1997, ch. 14.
- [12] D. M. Frank, B. K. Seliger, "Fuzzy logic and neural networks application to fault diagnosis," *Int. J. of approximate reasoning*, 16, 67–88, 1997.
- [13] R. J. Patton, J. Chen, "Observer based fault detection and isolation: robustness and application," *Control Engineering Practice*, vol. 5, no. 5, 671–682, 1997.



Esmail Tafazzoli was born in Tehran, Iran, in 1975. He received the B.A.Sc. degree in mechanical engineering from Khajeh Nasirodin Toosi University of Technology in 1999, Tehran, and the M.Sc. degree from Ferdowsi University of Mashhad in 2002, Mashhad, Iran.

From 2002 to 2004, he was a research assistant at Ferdowsi University of Mashhad. His research interests include fault detection and diagnosis in dynamic systems, tool condition monitoring in machining processes and fuzzy logic.

Mr. Tafazzoli is the 1999 recipient of the top student university award from KNT university of technology.

Silhouette based 3D Shape Reconstruction from multiple 2D Images

Lila Torabi, ENSC 894, summer 2005, SFU

Abstract— Numerous computer vision researchers have dealt with the construction of 3D solid models, when multiple 2D images are available. Among other image cues, silhouettes are attractive because they could be easily extracted, and they are good sources for shape information; also silhouette-based algorithms are fast enough for real time processing. The traditional framework for silhouette-based methods was volume intersection, but recently, image-based modeling has been proposed as a practical alternative. Here we present an improved technique for obtaining the 3D shape of an existing object, utilizing a single video camera and a turntable; The algorithm is an efficient image-based approach to computing visual hulls from image silhouette data. This method does not suffer from the computation complexity, limited resolution, or quantization artifacts of previous volumetric approaches.

Index Terms— Epipolar Geometry, Image-based modeling, Silhouette, Visual Hull

I. INTRODUCTION

To understand the 3D content of a scene is a central problem in computer vision [1] [2] [3], since it allows computer-driven equipment to perform tasks such as navigation, manipulation, inspection, and remodeling of existing objects autonomously.

Many computer vision research have dealt with the construction of 3D models when 2D images are available. These approaches are usually categorized as shape-from-X, according to the information is used, such as stereo, texture, motion, or shadow. Along with other image cues, silhouettes have been used to build approximate geometric models of static objects and scenes. The importance of these techniques is not only because silhouettes of a 3D object are good sources of shape information, but also extracting and working with silhouettes is easy, and silhouette-based algorithms are fast enough for real time processing.

A silhouette is essentially a binary segmentation of an object image, wherein pixels are labeled either “foreground” if they belong to the object or “background” if they don’t.

In silhouette-based approaches, shape reconstruction can be performed using volume intersection, to recover the volumetric description of the object. Silhouette images observed from a coarse set of sampling points determine a region of space as the shape of the object. Each image along with its projection center, defines a cone of rays, formed by back projecting the silhouette to the 3D space. The shape of the object is described as the volume, lying

within all the cones.

Octree was the first important method for volume intersection [4]. The idea originated in 1980, but in 1986, Chen proposed a practical solution for this technique [5]. The octree is a method of describing a volume by recursive subdivision of a single cube, surrounding the object. The universe cube is split into eight smaller cubes of equal volume called octants. Any octant lying partially within the object is split again, and this process is repeated until either a desired level of accuracy is achieved or no octant lies partially within the object [6]. The central problem with this approach is, its non object-oriented description; also the consequence of storing the model in a tree structure, requires large amount of storage and is time consuming.

Since then a number of silhouette-based algorithms have been proposed in order to enhance this method in different ways. Potmesil describes a method for computing a voxel representation of objects from sequences of silhouettes. He uses an Octree data structure to represent a binary volume of scene. Szeliski has implemented a similar idea, using a turntable to rotate objects in front of a real camera [7]. He computes an Octree-based voxel representation of the object by projecting octree nodes on the silhouette images. This technique, due to its 2D computations, not only decreased the processing time, but also the required storage memory.

Other volumetric reconstructions related to silhouette techniques, have also been suggested, which are not well suited to real-time implementation [8][9][10].

Although it seems intuitive that the shape of an object can be recovered from many silhouettes, clearly not all shapes can be constructed from silhouettes alone. If every possible viewpoint were used, the volume that lay within all of the cones would be the object’s *visual hull*. In the silhouette-based 3D modeling, visual hull is the best approximation of the object’s shape. Laurentini introduced the concept of the Visual Hull along with formalism for analyzing object reconstruction from silhouettes. His framework is useful for understanding the limitation of silhouette methods [11] [12].

Since volumetric approaches suffer from resolution problems and memory intensity, is difficult to retain the full precision of the original silhouette images using the standard volumetric representation.

Image-based modeling has been proposed as a practical alternative to the traditional modeling framework. In

image-based modeling, one starts with images and directly produces new images from this data. This technique avoids traditional polygonal modeling process, and often leads to the modeling algorithms whose running time is independent of the scene's geometric and photometric complexity. In 1997 McMillan introduced warping equations to define the basis of image-based modeling [13]. Chen's QuickTime VR is one of the first commercial acquired systems [14].

The image-based representation has a number of advantages in terms of computational efficiency, storage requirements, and ease of rendering.

In this paper, a new method for extracting the visual hull has been proposed. This method is based on the image-based modeling. Section 2, will describe the precise definition and mathematic equations for image-based modeling, epipolar geometry and visual hull. Section 3 includes the proposed technique, and the next section contains the experimental and implementation results. Finally, in section 5, we will have suggestion and conclusion.

II. IMAGE-BASED MODELLING

A. Visual Hull

By definition, for each point 'p' in the visual hull of an object, the line between p and the projection center, hits the object at least at one point. This definition illustrates that visual hull of an object consists of all points in the space, which their projection appears in all the silhouettes. As you can see in fig. 1, every projection center along with its corresponding silhouette, defines a cone in the space. The common space among all these cones, defines the visual hull.

The image-based visual hull is constructed with respect to a desired viewpoint in the viewing region of the visual hull. In fact, an image-based visual hull is the projection of the actual visual hull at a desired point of view, rather than an absolute approximation for the 3D shape of an object. Thus, the final result is stored and presented in 2D [15].

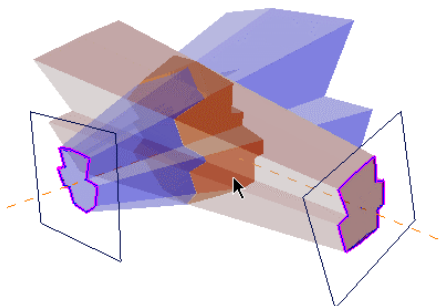


Fig. 1. Visual Hull is the smallest convex hull of an object.

The image-based visual hull of an object is a 2D silhouette image together with its depth information. In image-based approaches, the visual hull constructing process is carried out in 2D. The details of the technique for computing the image-based visual hull, will be explained further

B. Epipolar Geometry

According to projection geometry, each image, taken from a particular point of view, is specified by two parameters: the projection center, and the projection transformation matrix. Based on these two parameters, for each pixel in the image, a corresponding ray in the space is determined. And the Epipolar geometry explains how this ray appears as a line, in another point of view. This line is called an epipolar line [16]. The relationship between two projection centers, a point 'x' in space and the corresponding epipolar line, is shown in fig. 2.

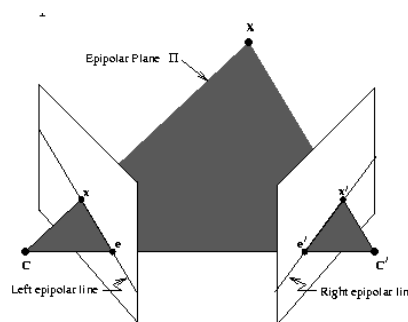


Fig. 2. Epipolar geometry explains the mapping between points in images from different points of view.

We show the mapping function for a pixel in image 'a', (\bar{x}_a) and the corresponding epipolar line (l_b) , in image 'b' from a different point of view in (1).

$$l_b = F_{ba} \bar{x}_a . \tag{1}$$

As is shown in (2), the transformation matrix F_{ba} is a combination of the projection matrixes for each point of view.

$$F = E_b P_b P_a^{-1} . \tag{2}$$

According to the above equations, each point in the image 'a' is back-projected to space as a ray, and this ray is projected to image 'b' as a line. The mapping function between a point in one image and the corresponding epipolar line in another image, constructs the basis of our image-based algorithm.

Consider the image 'a' as the image from the desired point of view, and b_i 's ($i=1...n$) as the reference images. For each pixel (\bar{x}_a) in the image 'a', the cross-section

between the corresponding epipolar line (l_{b_i}) in image ' b_i ', and the silhouette of the image ' b_i ', will be saved as the occupancy interval for (\bar{x}_a). Consequently for each pixel in this silhouette image, a list of occupancy intervals is stored. If a pixel does not belong to the silhouette, then the list is empty; otherwise the list contains intervals of space that are occupied by the visual hull of the object. These intervals, extruded over the solid angle subtended by the pixel, represent the region of the visual hull that projects to that pixel. The union of all such slices gives the visual hull as sampled from the desired viewpoint [17].

The occupancy intervals can be stored as pairs of real numbers (where the numbers represent the minimum and maximum depth of the interval). Thus, while the volume is discretized in two dimensions, the third dimension is continuous, allowing for higher resolution volumes than a voxel approach. Note also that this representation can be used for an arbitrary volume; it is not specialized for a visual hull. Computing a visual hull using the image-based techniques is much simpler than previous approaches, because the three-dimensional cone intersections and the volumetric carving operations are replaced with the 2D interval intersections [18]. These interval intersections are fast and robust, allowing a real-time calculation of the visual hull. If the image-based visual hull is derived from m images, each with (n,n) resolution, the running time will be $O(mn^2)$. The image-based representation of the visual hull alleviates some of the problems with the standard voxel approaches. By definition, such a representation preserves the full resolution of the input images and contains no more or no less information than that provided by the silhouettes [19].

III. IMPROVEMENT IN THE IMAGE-BASED MODELING METHOD

As mentioned in the previous section, the image-based approaches extract the occupancy intervals for each pixel of the desired image, from m reference images.

Precisely reviewing the algorithm demonstrates, several pixels in the desired image, correspond to equivalent epipolar line in each reference image. Our proposed algorithm is based on avoiding this redundancy in order to decrease the process time. Considering an arbitrary point of view as the desired projection center; for every reference silhouette image, the epipolar point corresponding to the desired projection center C' , will be achieved as in (3):

$$e_i = P_i C' . \tag{3}$$

Here, e_i is the epipolar point for image ' i ', defined as

the projection of the C' on this image .

Providing a list of all points on the contour of the silhouette, of a reference image ' i ', an arbitrary point x_1 is selected from the list. The line between this point and the epipolar point is drawn, and as shown in fig. 3, this line may cross the silhouette contour at several other points.

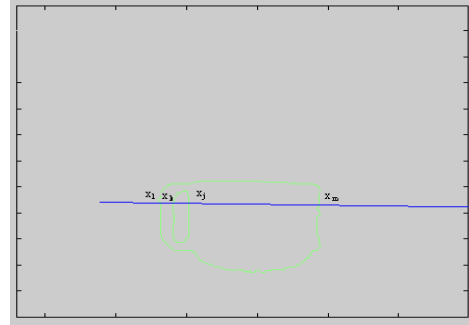


Fig. 3. The epipolar line intersects the silhouette in multiple points

As you have seen in fig. 1, the projection centers and their equivalent epipolar points are on the same plane. Thus, according to triangle rule, the occupancy intervals of x_1x_2 or x_1x_2 , can be derived. The result intervals are back-projected into the space, and saved as the occupancy intervals for the points on the desired image, which correspond to this epipolar line.

The same process is executed for every point on the silhouette contour, on the each reference image. At this stage the image-based visual hull is accomplished.

IV. IMPLEMENTATION AND RESULTS

Our 3D modeling system, utilizes a single video camera to achieve a sequence of images of an object, rotating about a single axis on a turntable. We used a turntable that could rotate in fixed 15 degrees. The algorithm is an efficient image-based approach for computing visual hulls from silhouette images, which takes advantage of *Epipolar geometry*.

In order to convert a series of images into a 3D model, first the camera should be calibrated. Here the camera is considered calibrated, if the mapping between the images coordination and directions, relative to center of the camera are known. However, the position of the camera in space (i.e., its translation and rotation with respect to world coordinates) is not necessary. For an ideal pinhole camera, delivering a true perspective image, this mapping can be characterized completely by *intrinsic parameters* of the camera [20]. In contrast, a camera's *extrinsic parameters* represent its location and rotation in space [21].

Since our calibration method is based on a photometric technique, we use a calibration object [22]. The calibration object (a checkerboard) is placed on the turntable, and at least 5 pictures should be taken from different viewpoints. After the camera calibration, we put the main object on the turntable (without changing the camera position), and take series of images from the object.

The sequence of the images taken from a calibrated camera, are the appropriate source for the silhouette extraction phase. In this project, we used one of the most robust techniques for silhouette extraction, called *Blue Screen Matting* method [23]. In this technique, a constant color should be used as the background, so we covered the background and the turntable with blue.

As the algorithm was declared in last section, the visual hull for the desired viewpoint is constructed based on the silhouette images and the camera calibration information. This algorithm has been implemented on a cup. Reference images as shown in fig. 4(a), are nine images which are taken from the cup in 30 degree rotation steps. By processing these images, three images from three different viewpoints, accompanying depth information for each image, are accomplished. The resolutions of the output images are fairly near to the source images.

V. CONCLUSIONS AND SUGGESTIONS

3D modeling and rendering is one of the most interesting topics in computer vision. Various methods and techniques have been proposed in this area, attempting to achieve a real-time, robust, and high resolution algorithm, with the minimum memory storage requirements. Image-based approaches provide a real time solution for the problem, with minimum memory storage and processing complexity.

This research has provided good results according to the running time and the storage requirement. However, some practical problems affected the accuracy of the result. In future, more effort should be undertaken to determine the source of these problems and solve them.

The proposed method is only implemented on a simple object, more experiences with complex objects will precisely determine the advantages and disadvantages of this algorithm.

As we have mentioned, in this research different images were taken from the object using a turntable, other experimental systems, with larger variety in selecting projection centers will lead us to more realistic problems.

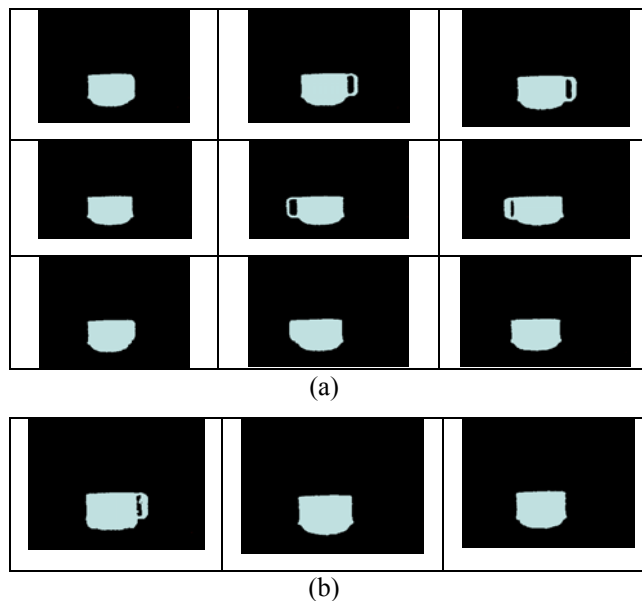


Fig. 4. a) The reference images for constructing the image-based visual hull for a cup, b) Three images from three desired point of views.

REFERENCES

- [1] Debevec, p. C., Talor, and J. Malik, "Modeling and rendering architecture from photographs," *SIGGRAPH*, pp. 29-38, 1995.
- [2] Debevec, P. E., Y. Yu, and G. D. Borshukov, "Efficient view dependent image-based rendering with projective texture mapping," *Proc. OfEGRW*, 1998.
- [3] Faugeras, O., *Three-dimensional computer vision: A geometric*, MIT Press, 1993.
- [4] Potmesil, M., "Generating octree of 3D objects from their silhouettes in a sequence of images," *CVGIP* 40, 1-29, 1987.
- [5] C. H. Chien, and J. K. Aggarwal, "Identification of 3D objects from multiple silhouettes using quadtree/octree," *Comp., Vision, Graphics and Image Processing* 36, pp. 256-273, 1986
- [6] Narendra AHUJA, "Generating Octree from Object Silhouettes in Orthographic Views," *IEEE PAMI*, vol. 11, pp. 137-149., 1989.
- [7] Szeliski, R., "Rapid octree construction from image sequences," *CVGIP Image Understanding* 58, 23-32, 1993.
- [8] M.Jones and J.P.Oakley, "A New Shape Representation for fast Reconstruction from multitude 2D perspective Images," *IEE Proceeding Vision, Image and Signal Processing*, May. 1995.
- [9] Jiang Yu Zheng, "Acquiring 3D models from sequences of contours," *IEEE PAMI*, vol. 16, Feb. 1994.
- [10] M.Jones and J.P.Oakley, "An Efficient Representation Object Shape for Silhouette Intersection," *IEE Proceeding Vision, Image and Signal Processing*, March. 1994.
- [11] Laurentini, A., "The visual hull concept for silhouette based image understanding," *IEEE PAMI* 16, 2, pp. 150-162, 1994.
- [12] Laurentini, A., "How far 3D shapes Can be understood from 2D silhouettes," *IEEE PAMI* 12,5, 1995
- [13] McMallin, L., *An image-based approach to three-dimensional computer graphics*, Ph.D. Thesis, university of North Carolina at Chapel Hill, Dept. of Computer Science, 1997.
- [14] Chen, S. E., "Quick time VR-an image-based approach to virtual environment navigation," *SIGGRAPH*, pp. 29-38, 1995
- [15] Chris Buehler, Wojciech Matusik, Leonard McMillan, and Steven J. Gortler, "Creating and rendering image-based Visual Hull," <http://graphics.lcs.mit.edu/~wojciech/uh/780-tr>, 1999.

- [16] *W. Niem*, "Robust and fast modeling of 3D natural objects from multiple views," *SPIE Proc.*, vol. 2182, pp. 388-397, 1994.
- [17] *Debevec, P.*, Modeling and rendering architecture from photographs, *Ph.D. Thesis, University of California at Berkeley Computer Science Division, Berkeley, CA, 1996*
- [18] *McMillan, L., and G. Bishop*, "Plenoptic modeling: An image-based rendering system," *SIGGRAPH 95*, 39-46.
- [19] *Xianfeng Gu, Steven J. Gortler, Hugues Hoppe, Leonard McMillan, Benedict J. Brown, and Abraham D. Stone*, "Silhouette mapping," <http://cs.cmu.edu/project/v-groups.htm>, 2000.
- [20] *Tsai, R.Y.*, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," *IEEE Journal of Robotics and Automation*, vol. RA-3, No. 4, Aug. 1987.
- [21] *Faugeras, O. D., Q. T. Luong, S. J. Maybank*, "Camera self-calibration: theory and experiments," *Proc. of the 2nd European Conference on Computer Vision*, pp. 563-578, 1992.
- [22] *Zhengyov, Zhang*, "A flexible new technique for camera calibration," WWW.research.microsoft.com/~zhang/publications.htm.
- [23] *Smith, A. R., J. F. Blinn*, "Blue screen matting," *ISIGGRAPH*, 1996.



Lila Torabi received her B.S. in communication engineering from Sharif University, Tehran/Iran, in 1996. After 2 years of engineering experience, she continued her education in AI and Robotics, and got her M.S. in 2000 from Amir Kabir University, Tehran/Iran. Now she is a PHD student in school of engineering science at Simon Fraser University, British Columbia,

Canada. Her current research interest is Path Planning in Robotics. She has several years of work experience in different branches of industry as design engineer and automation supervisor.

Investigation of Resolution Limit for Bi/In Bimetallic Films Using Two-beam Pulsed-laser Interference Patterning

Glenn H. Chapman, Jun Wang, Chinheng Choo and Yuqiang Tu

School of Engineering Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada

Abstract— This paper introduces a method to explore the resolution of Bi/In bimetallic films. Bi/In films are promising cost-effective inorganic materials for laser direct-write grayscale photomasks with advantages over traditional organic microlithography materials. However, in our pervious research, we found that the smallest feature created in the films is not less than 2 μ m by single-beam argon laser writing system. In this project, a two-beam interference technology was used to investigate the resolution at nanometer-scale in Bi/In films. Using pulsed Nd:YAG laser(532nm/266nm), we created parallel lane pattern on 30nm thick Bi/In films. The smallest pitch size is less than 500nm and line width within 70-100m. The experimental results show the strong ability of Bi/In films for microfabrication and micromachining photolithography materials.

Index Term—Bi/In film, diffraction, laser, interference, photomask, resolution.

I. INTRODUCTION

BI/IN film is a bilayer structure of Bismuth film over Indium film created by the DC-sputter deposition process, and it can be exposed using lasers with a wide range of wavelength and can be developed by diluted RCA2 solutions[1]. After an argon laser (514/488nm) full exposure, the optical density (OD) of films changes from original 3OD to 0.2OD [2]. The results of X-ray diffraction analysis (XRD) and profilometry measurement demonstrate oxidation process takes place in the film area under the

laser exposure. Using a calibration process, we can fabricate a grayscale pattern in a film, which can further be an analogue photomask for microfabrication process [3]. ANSYS simulation shows that the temperature at laser spots on Bi/In films is over their melting points. Bi/In X-ray diffraction analysis (XRD) indicates that the oxidation take place in the Bi/In bimetallic films during laser scanning.

As shown in the figure 1, Bismuth layer and Indium layer (30nm to 120nm for each layer) are DC-sputter deposited on a substrate (figure 1(a)). The step is a dry step without baking or edge bead removing process. Then the slide with film is exposed using lasers. The part under higher power density scanning becomes more transparent than the one under lower density laser-scanning (figure 1 (b)). After RCA2 development, the fully-exposed part will remain in the substrate while the bismuth and indium metal react with RCA2 and dissolve.

Our previous experiments have demonstrated that Bi/In bimetallic films can be a cost-effective solution to traditional organic microlithography materials[4].Bi/In films have two outstanding advantages over normal inorganic photomask materials. First, the process is much simpler than the conventional photolithography procedure using organic mask materials [5]. The pattern in a Bi/In film can be transfer to a photoresist by UV exposure, or convert the pattern to a silicon wafer with etch. After they are used as photomasks, Bi/In films can still be remained while the traditional photomask must be removing. So the photolithography process with Bi/In photomasks has a high potential for low-cost production without tedious alignment and complicated etching processes. Secondly, Bi/In films are independent to the exposure laser wavelength from DUV to IR while the normal organic photomask materials must undergo a big alternation on their compositions to fit to DUV and EUV range light exposures [6].

In order to explore the potential applications of Bi/In films in the microfabrication for micron even nanometer magnitude devices, we must know the minimum feature size that can be fabricated on Bi/In films. The resolution of a pattern created by laser-scanning in a film is dependent on

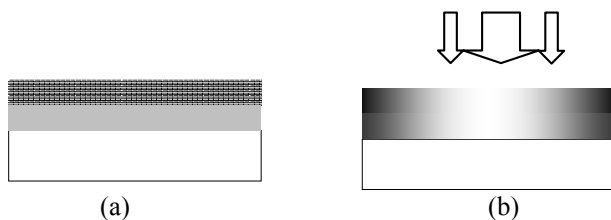


Fig. 1. The process of Bi/In bimetallic film
(a) DC-sputter films;(b)laser-exposure.

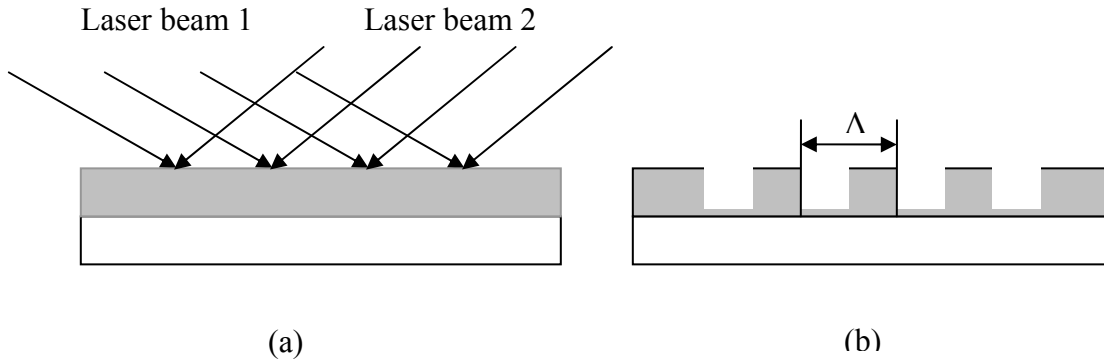


Fig. 2. Schematic illustration for two laser beam setup (a) and interference pattern with period Λ after laser exposure and RCA2 development (b).

the related optical systems, including lens, laser writing-systems, mirrors, and shutters. Currently our focused (50X lens) laser spot is around $2\mu\text{m}$ using a CW Argon laser (514nm/488nm). When the one-beam argon laser writes on the film, the smallest feature is not less than the resolution ($2\mu\text{m}$) of the argon laser-writing system. The value means that finding the resolution capacity of films below $2\mu\text{m}$ by the normal optical systems is difficult.

A two-beam interference setup using a Nd:YAG laser was built in our lab with the goal of making finer structures in Bi/In films. The pitch size (500 nanometer and below) can be achieved, so the two-beam interference technique can help us investigate the minimum dimension of patterns (can be less than $2\mu\text{m}$) in Bi/In films and further understand the resolution limitation of Bi/In films.

II. THE TWO-BEAM INTERFERENCE TECHNIQUE

Interference is the overlapping superposition of two or more waves resulting in a new wave pattern [7]. The two-beam interference technique is used to create high aspect and nanometer-scale gratings in positive photoresist polymer films applied in polymer-based optical communication systems [8]. And interferometric lithography (IL) also provides very simple, inexpensive techniqueto print extreme sub-micron structures [9]. In our

two-beam interference experiment, two coherent lights, with same intensity but different phases, are from a same laser source, pulsed Nd:YAG laser which produces high power light pulses at the IR (1064 nm), Green (533 nm), UV (266 nm) and Vacuum UV (213 nm) wavelengths.

As shown in Bragg's law (formula 1), the angle of two beams and the wavelength of lasers decide the interference-pattern size in films.

$$\Lambda = \frac{\lambda}{2 \sin \theta} \quad (1)$$

Here Λ is the grating period, and λ is the wavelength of the Nd: YAG laser (266nm/532nm), and θ is the angle of two incident beams on Bi/In film

The interference lithography is a better way to show resolution ability of films than using single-beam laser scanning because the latter can not create any features below $2\mu\text{m}$. Ideally, the smallest pitch by the two-beam interference technology can be half of wavelength of the Nd: YAG laser, meaning the range of around 532-133nm.

III. EXPERIMENTS FOR RESOLUTION STIGATION

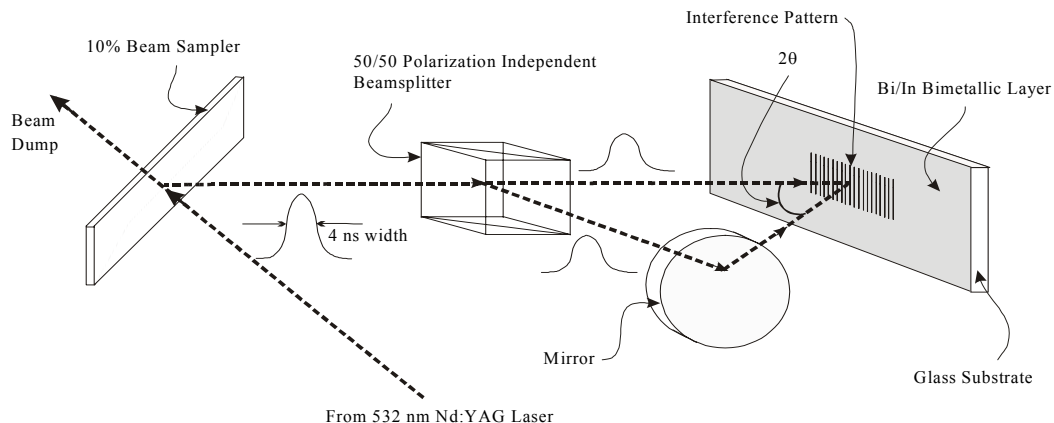


Fig. 3. Illustration of two-beam interference pattern system for Bi/In film

TABLE I EXPERIMENTAL VALUES

Substrate	Glass	Silicon
Laser (532nm)	25	270
scanning power(mJ)		
RCA2 concentration	1:1:48	3:3:44
RAC2 developing time(sec)	25~40	300~600

In the two-beam interference lithography experiment, we used an Nd: YAG pulsed-laser at wavelengths of 532nm and 266nm (shown in fig 3) to explore the resolution limitation of our Bi/In films. A 10% beamsplitter allows the laser to operate at higher powers ensuring better beam quality than the one at lower laser powers. A 50/50 polarization independent beamsplitter is applied for one coherent light, and the mirror reflects the laser from the beamsplitter as another beam of coherent light to make parallel-line-pattern on Bi/in samples.

A. Experiments

The Nd:YAG laser can produce high power light pulses at the IR (1064 nm), green (533 nm), UV (266 nm) and Vacuum UV (213 nm) wavelengths. In this experiment, 533nm and 266nm were used. In addition, the influence of substrates (glass slides and silicon wafers) of Bi/In films was also included. Because of the difference of thermal conductivity of glass and silicon, the laser power for Bi/In films with different substrates is different. The experimental results will demonstrate the influences of substrates, laser wavelengths and laser powers to the features on structures created.

After the parallel line patterns are created on the bilayer films, the unexposed parts need to be removed by a developing step and leave the parts which converted to bismuth oxides and indium oxides still on the substrates. A diluted RCA2 solution with HCL: H₂O₂:H₂O at the room temperature was employed as the developer to remove the bismuth and indium areas in films. Concentrations and developing time (shown in table 1) were experimented with using films on different substrates.

The Scanning Electron Microscope (SEM) pictures allow us to observe the interference lines in details. Because SEM requires films of good conductivity, our exposed Bi/In films should be coating by carbon or gold in vacuum evaporation. The thickness of the carbon coating is around 5-10nm.

IV. RESULTS

A. Substrate influence

The interference pattern varies on Bi/In films with different substrates and different laser dose as well as different wavelength laser.

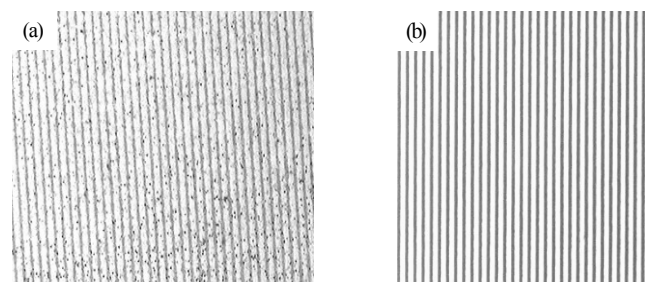


Fig. 4.1200X optical microscope pictures of Bi/In film immediately after exposure. (a) Glass substrate, 25 mJ pulse with 1.72 μm pitch, 1200X magnification. (b) Silicon

In the figure 4, two 1200X optical microscope pictures show the parallel line interference patterns with different substrates (silicon and glass). Two points should have been noticed. First, the reason we used various laser power (25mJ and 270mJ) on different substrates to create the interference pattern is that the heat transfer properties of two materials are not same. The thermal conductivity of silicon (150w/mK.) is over 100 times of the one of glass slide (around 1w/mK) at room temperature. Secondly, the bridging occurs between the lines on the film with glass substrate while the film with silicon substrate display much cleaner features. Therefore, the more precise structure can be created on silicon wafers than that on glass slides.

B. Wavelength influence

As indicated in equation 1, laser wavelength has an influence on the line pitch. We found that the finer periods can be created using 266nm laser than the ones using 532nm laser, which is consistent with the principle from the formula.

On figure 5, we can see that after single-pulse 532nm argon laser exposure, the minimum feature fabricated on the film with glass substrate is 720nm line with 1650nm pitch while on the film with silicon substrate is 190-220nm line with 555nm pitch. After the exposure of 266nm Nd:YAG laser, the minimum line has a width of $85 \pm 15\text{nm}$ with the pitch at 430nm using 266nm laser on a 30nm Bi/In (50/50) film.

The minimum feature created by the 532nm laser with two-beam interference method is 1/10 of that single-pulse Nd: YAG laser at 532nm exposure while 266nm Nd:YAG laser can create less than 1/20-1/30 of size which fabricated by one-beam argon laser exposure. Bi/In bimetallic film with silicon substrate and this size is only 9 percent of the smallest feature created by the laser at 532nm. The result clearly demonstrates the two-beam inference can much better help us to explore the capability of resolution Bi/In films. And the value of resolution we created in the Bi/In bimetallic film also show that the promising ability to apply in the microfabrication for micro-optics devices and nanometer devices.

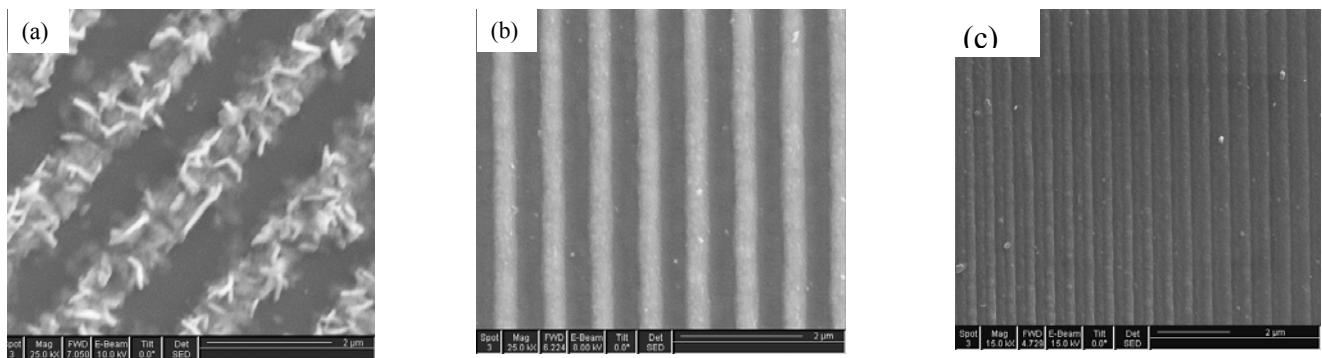


Fig. 5. SEM picture of carbon coated Bi/In. (a) Glass substrate, 35 second development time, line widths are $0.72 \mu\text{m}$ and the pitch is $1.65 \mu\text{m}$. (b) Silicon substrate, 6 minute development time, incident angle of 19.4 degrees, line widths are between 280 and 390 nm and the pitch is 802 nm . (c) Carbon coated Bi/In on Silicon. The incident angle was 28.8 degrees, the development time was 7 minutes, the measured pitch was 555 nm and the line widths were in the range 190 to 220 nm .

V. FUTURE WORK

In order to explore the thermal process in detail within Bi/In films during laser scanning, we plan to do the further work in four aspects shown below:

- 1) Simulating the temperature distribution within Bi/In films with different substrates (glass and silicon) and different laser power (25mJ and 270mJ) respectively.
- 2) Cooling glass substrates while the stronger laser exposures
- 3) Making bigger angle of beams for small fringes should benefit the further investigative work on resolution of Bi/In films.
- 4) Investigating the resolution limit of other inorganic films such as Sn/In films, Al films and Zn films.

VI. CONCLUSION

A two-beam interference technology was successfully used in investigation for the resolution limit of 30nm Bi/In films. Compare to other ways to create fine structure by laser in films, the two-beam interference experiment is straight, cheap and controllable. Previous research the small features created by single-beam cw argon laser ($514/488\text{nm}$) in Bi/In films is $2\mu\text{m}$, which is 30times of the minimum size fabricated by two-beam pulsed Nd: YAG laser ($532/266\text{nm}$). This paper also discusses the influence of substrates to the pattern quality in films after laser exposure. The experimental results show that the finer and cleaner structure can be created on silicon wafers than that on glass slides. And the smaller pattern can be generated by 266nm Nd: YAG laser than that by 532nm laser, which is identical with the Bragg's law. The smallest pitch and line so far we have created in a 30nm Bi/In ($50/50$) film is 430nm and 70nm respectively. The values promise the capacity of Bi/In films to show nanometer-scale structures and the application in microfabrication and micromachining industry.

ACKNOWLEDGMENT

Jun Wang thanks Dr. Yang Li for SEM scanning and part of the measurements.

REFERENCE

- [1] G.H. Chapman and Y. Tu, "Bi/In Thermal Resist for Both Si Anisotropic Wet Etching and Si / SiO₂ Plasma Etching", *Proc. SPIE Micro04, Photonics West, Micromachining and Microfabrication Process Technology IX*, San Jose, Jan. 2004
- [2] G.H. Chapman, J. Dykes, D. Poon, C. Choo, J. Wang, Y. Tu, and J. Peng, "Creating Precise 3D Microstructures Using Laser Direct-write Bimetallic Thermal Resist Grayscale Photomasks", *Proc. SPIE Photonics West: Photon Processing in Microelectronics and Photonics IV*, v 5713, pg. 247-258, San Jose, Jan. 22, 2005.
- [3] Y. Tu, G.H. Chapman J. Peng, J. Dykes, and D. Poon, "Calibrating Gray-scale Direct Write Bimetallic Photomasks to Create 3D Photoresist Structures", *Proc. SPIE BACUS Symposium on Photomask Technology 24th annual*, v 5567, pg 245-256, Monterey, CA, Sept. 2004
- [4] Y. Tu, G.H.Chapman, "Bi/In as patterning and masking layers for alkaline-base Si anisotropic etching", *Proc. SPIE Micro03, Photonics West 2003, Micromachining and Microfabrication Process Technology VII*, Vol. 4979, Pg 87-98, January 2003
- [5] G.H. Chapman, Y. Tu, "Single Step Direct-Write Photomask Made From Bimetallic Bi/In Thermal Resist", *Proc. SPIE LASE 2003 Photonics West, Photon Processing in Microelectronics and Photonics II*, v4977, pg 257-268, San Jose, Jan 2003.
- [6] G.H. Chapman, Y. Tu, J. Peng, "Wavelength Invariant Bi/In Thermal Resist As A Si Anisotropic Etch Masking Layer And Direct Write Photomask Material", *Proc. SPIE Microlithography Conference, Advances in Resist Technology and Processing XX*, v5039, pg 472-483, Santa Clara, CA, Feb. 2003.
- [7] Dong Jun Kang, Jin-Ki Kim, and Byeong-Soo Bae, "Simple fabrication of diffraction gratings by two-beam interference method in highly photosensitive hybrid sol-gel films", *Optics Express*, Vol. 12, No. 17 August 23, 2004, pp. 3947 – 3953
- [8] Shinya Shibata, Yanlong Che, Okihiro Sugihara, Naomichi Okamoto and toshikuni Kaino, "Fabrication of high-resolution periodical structure in photoresist polymers using laser interference technique", *Japanese Journal of Applied Physics*, Vol. 43, No.4B, 2004, pp.2370-2371
- [9] Xiaolan Chen, A. Frauenglass and S.R.J.Brueck, "Interferometric lithography pattern delimited by a mask image", *SPIE*. Vol 3331, pp496-502

ENSC 894 COURSE TRANSACTIONS



Glenn H. Chapman (S'72–M'80) was born on August 28, 1948. He received the B.Sc. degree in engineering physics and the M.Sc. degree from Queen's University, Kingston, ON, Canada, in 1972 and 1975, respectively, and the Ph.D. degree from McMaster University, Hamilton, ON, Canada, in 1982. From 1980 to 1990, he was a Research Staff member with the Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, where he worked on developing laser redundancy techniques for the Wafer Scale Integration project. Since 1990, he has been a Professor with the School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada, specializing in the areas of large-area laser restructurable silicon systems, microfabrication technology, and micromachined sensors involving lasers. He is the author of 27 journal papers, 68 conference papers, two book chapters, and 20 patents.

Prof. Chapman is a Senior Fellow of the British Columbia Advanced System Institute.



Jun Wang received the B.Sc degree, Master degree and Ph.D in physical electronics and opto-electronics from University of Electronic Technology and Science of China (UESTC), in 1990, 1993 and 2002, respectively. From 1993 to 2002, she works as a research staff at Information Display Device Lab of UESTC and published 10 papers at journal and international conference for the project “ the high-luminance and high-resolution YAG projection tube”. Currently, Jun Wang is a graduate of Prof. Glenn Chapman at the Simon Fraser University in Burnaby, Vancouver, BC, Canada.

A Primal Dual Algorithm for View Planning with Traveling Cost Based on the Unified Problem Formulation

Pengpeng Wang, *Student Member, IEEE*

Abstract—The problem of view planning with traveling cost (*Traveling VPP*) refers to finding a sequence of sensing actions with minimum total costs by the robot sensor system to completely inspect an object surface with known geometry. The costs to minimize include sensing cost, which is proportional to the number of views planned, and the traveling cost by the robot to realize them. The common practices in the field of robotics is to divide the problem into two subproblems, namely to solve first the minimum views problem, the *VPP* problem, and to plan the shortest itinerary for the robot, the *Shortest Itinerary (SI)* problem, to realize the planned views. However, as we show in a counterexample, this approach could result in arbitrarily poor approximation ratios. In this paper, to improve the performance, we take a unified approach to the problem (i.e., to consider the two subproblem as a whole and design algorithms accordingly).

In this paper, we formulate the *Traveling VPP* problem as an integer program. This formulation provides the foundation of solving the problem optimally. Inspired by its close relationship to the connected facility location problem (*Connected FL*) [3], we propose a fast polynomial time algorithm to solve the *Traveling VPP* problem using the primal dual schema. A nice feature of this algorithm is that it addresses nicely the tradeoff between the minimum number of views and the traveling cost to realize each view. By solving the counterexample to the two-level approach, we show that this algorithm works significantly better than the two-level approach.

Index Terms—View planning; traveling VPP, shortest itinerary; facility location; primal dual scheme.

I. INTRODUCTION

For object inspection tasks, the robot-sensor system is required to take minimum views to completely inspect the surface of an object of interest. At the same time, since the sensor poses are achieved by the movements of the robot mechanism on which

the sensor is mounted, from a motion planning point of view, we would also like to minimize the total robot movements. The consideration of both the minimum views and the minimum movements motivates the problem of view planning with traveling cost, denoted by *Traveling VPP* (i.e., to plan a sequence of sensing actions with minimum total costs by the robot sensor system to completely inspect an object surface with known geometry). These costs include both the total sensing cost, proportional to the number of views planned, and the traveling cost, the itinerary distances to connect these views.

Motivated by the two components of the problem: to plan a set of views to cover the whole object surface and to connect these views by a minimum length itinerary, in [1], the authors propose to decompose the problem accordingly into two levels. At the first level, the optimization problem of minimizing the number of views, the view planning problem (*VPP*), is solved without considering the robot traveling costs. At the second level, these planned views are fed into a shortest itinerary planning optimization problem (*SI*), and the shortest path movements (tour) to connect all these views is solved. As shown in Fig. 1, even under the assumption that at each level the respective optimization problem is solved optimally (as opposed to approximately), this two-level approach can perform arbitrarily poorly. This issue occurs because the planned views at the first level are so far apart for the robot to realize efficiently since no traveling cost is considered at the first stage. By making the leftmost view even further apart from the rightmost vies, we can show the two-level approach performs arbitrarily poorly.

This observation motivates the approach we are taking in this paper: to formulate both components as a whole and design fast algorithms.

Manuscript received June 21, 2005; revised July 21, 2005. This work was supported in part by the NSERC under PGS B scholarship.

The author is with School of Engineering Science, Simon Fraser University, Burnaby, B.C., Canada (e-mail: pwangf@cs.sfu.ca).

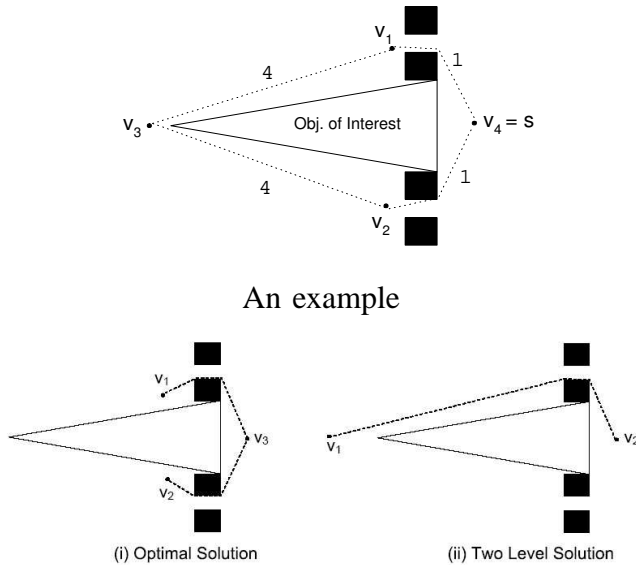


Fig. 1. An example showing the performance of the two-level approach. In the figure, s_i 's are the surface patches of the object to inspect. v_j 's are the view points planned. The dashed lines shown are the planned paths connecting with the view points. The two-level approach, (i), although being able to minimize the number of views, needs a much longer traveling path than the optimal solution, (ii), since these planned views are far apart.

In the integer and combinatorial optimization research area, the Traveling VPP problem relates closely to the classical facility location (FL) problem. FL problem refers to planning the minimum number of “facilities” to serve all the clients considered. By reduction to a Set Cover problem¹, VPP is shown to be NP-hard [2]. It is widely believed that it takes exponential time to solve it optimally. This theoretical result motivates enormous efforts to solve the VPP problem approximately using a polynomial time algorithm.

In [3], the authors use a primal dual schema based algorithm to the connected FL problem, a variant of the FL problem that has the connection (between the planned facilities) requirement on top of the constraint that each location is served by at least one facility. The idea of primal dual schema is to investigate the duality between the linear program (LP) relaxation of the original problem and its corresponding dual program (DP). Using the fact that the solutions of both programs are bounded by each other, the primal dual schema helps show

¹The Set Covering problem, to find a subset of set \mathcal{A} with minimum cardinality, with which a collection of subsets of \mathcal{A} all have non-empty intersection, is another classic combinatorial optimization problem.

the approximation ratios of the algorithmic solution. Using the primal dual schema, the authors prove that the approximation ratio, the ratio between the algorithmic and the optimal solution in the worse case, is constant [3].

Because of the relevance of the algorithm [3] to Traveling VPP, here we give an intuitive explanation of the result here. The dual variables associated (with the clients) are interpreted as the prices paid for both being served and constructing the facility connection. The algorithm picks (opens) only those facilities that can serve a large number of clients and then constructs the Steiner Tree to connect them. By the primal dual schema, the algorithm guarantees that the total costs, the primal solution, is bounded at a constant ratio by the dual solution, thus guaranteeing the constant performance ratio.

The rest of this paper is organized as follows. First, we formulate the overall problem as an integer program (IP), which is the emphasis of this paper and serves as the foundation of designing fast and good algorithms. Then we show how to use primal dual schema based algorithm to address the tradeoff between traveling and sensing costs. Finally, by solving the counterexample in Fig. 1, we show our new algorithm performs better than the two-level approach [1].

II. TRAVELING VPP PROGRAM FORMULATION

To introduce the formulation of Traveling VPP, in the following, we first formulate its two components, VPP and SI, and then show how to modify them to formulate Traveling VPP.

A. View Planning Problem (VPP)

To help readers understand the formulation, we illustrate it using a two dimensional setting. However, we emphasize that this formulation and the algorithms designed accordingly apply to the general case. In this simple setting, an object of interest (of polygonal shape) resides in a two dimensional space populated with polygonal obstacles. A point robot equipped with a omnidirectional range sensor with line of sight is required to inspect the complete (polygonal) surface of the object. The omnidirectional and line of sight nature of the sensors reduces the view or viewpoint configuration space into the same 2D space² where the objects reside. We can then

²For general cases, the view point space may be of higher dimensionality than the physical space.

decompose the space into view cells according to the visibility of the surface. This idea is illustrated in Fig. 2.

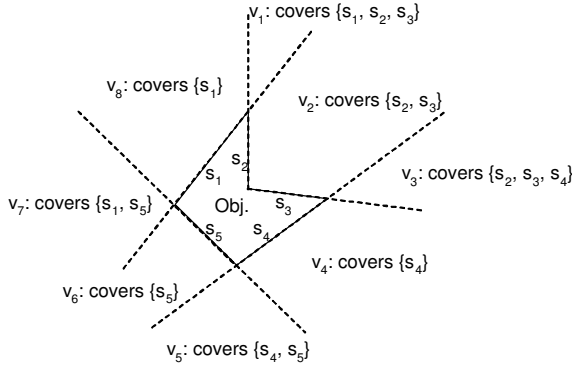


Fig. 2. An example showing how to decompose the view point space into view cells according to the geometry of the object of interest. The object of interest, labeled by *Obj.*, consists of five surface patches, s_1, \dots, s_5 . The view point space is decomposed into eight cells, v_1, \dots, v_8 . The surface patches that can be covered at these cells are also labeled inside each cell.

We denote the set of all cells by \mathcal{V} and further label the cells by i . We denote the set of surface patches by \mathcal{S} and label them by j . We use the notation of $i \in \mathcal{V}$ and $j \in \mathcal{S}$ to imply the “ i th view cell” and “ j th surface patch”, respectively.

The VPP problem refers to planing the minimum number of views subject to the constraint that each surface patch is covered by at least one view. We can easily see the close connection to the well-known facility location problem. Following is The VPP problem in a integer program (IP) setting. Using binary variable $y_i \in \{0, 1\}$, for $i \in \mathcal{V}$, to indicate whether to take the view in the i th view cell ($y_i = 1$) or not ($y_i = 0$), and c_i as the cost of taking that view³, we have:

VPP Problem (IP)

$$\begin{aligned}
 & \text{Minimize} && \sum_{i \in \mathcal{V}} c_i y_i \\
 & \text{Subject to:} && \\
 & \forall j \in \mathcal{S} : && \sum_{i \in \mathcal{V} \text{ "sees" } j} y_i \geq 1 \\
 & && y_i \in \{0, 1\}
 \end{aligned} \tag{1}$$

³Different values of c_i usually imply the different sensing efforts associated with different views. For general cases, these costs are the same for all the views. But we want to make it as general as possible.

B. Shortest Itinerary (SI) Problem

The second component of Traveling VPP is to plan the shortest path itinerary (SI problem) to connect the planned views and the starting position s of the robot.

We construct the graph $G = \{\mathcal{V}, E\}$ that encodes the robot movements to achieve different views. The node set \mathcal{V} is the set of all view cells and s , the starting position of the robot. The edge e between two views v_{i_1} and v_{i_2} represents the shortest path from v_{i_1} to v_{i_2} . Thus, the edge cost, c_e , is this shortest path length.⁴

In the following, instead of giving an IP formulation for the shortest itinerary problem, we instead give a Steiner tree formulation [4]. The reason is twofold. First, the Steiner tree formulation can be generalized to the case where the views to take are not explicitly specified, which is exactly the case for the Traveling VPP problem where the views are implicitly specified. Second, although this formulation is not exactly the same as the itinerary problem, the solution set of the Steiner tree problem nevertheless includes the shortest itinerary solution, and the optimal solution to the Steiner tree problem can be easily modified to become an itinerary, and this approximation solution cost is bounded by a constant factor, 2 or 1.5.

We use z_e to indicate whether to include edge e in the solution ($z_e = 1$), or not ($z_e = 0$). Without loss of generality, identifying the planned views as $\mathcal{V}_p = \{v_1, v_2, \dots, v_k\}$, the IP formulation of the Steiner tree problem is as follows,

Steiner Tree Problem (IP):

$$\begin{aligned}
 & \text{Minimize} && \sum_{e \in E} c_e z_e \\
 & \text{Subject to:} && \\
 & \forall i \in \mathcal{V}_p, \forall T : && \sum_{e \in \delta(T)} z_e \geq 1 \\
 & && z_e \in \{0, 1\}
 \end{aligned} \tag{2}$$

In the above, we use $\delta(T)$ to denote the set of edges that cross a cut/partition, T and $\mathcal{V} \setminus T$, of the graph. The inequality constraints thus can be interpreted as follows: for each planned view, i , and any subset of \mathcal{V} containing i , at least one edge

⁴For implementation, we could first construct a graph/roadmap that encodes the robot local movements between neighboring views, which is similar to the local planner in the PRM setting [5].

crossing the cut is picked. This guarantees there always exist connections from the planned views to the known starting node, s .

C. Traveling VPP Problem

For the Traveling VPP problem, a weighted linear combination of the path distance and number of view points/cells is needed. The weights are set to adjust for the tradeoff between the movement costs and the sensing costs. In the following, we give an IP formulation using this combined cost as the objective function. By modifying the Steiner tree formulation shown in the previous section, we can get the IP formulation for Traveling VPP. Note that unlike the previous case where the view cells to visit are all known, here we do not know them in advance; in stead, we know the requirement that the surface patches, which can be viewed by those visited, should cover the whole surface.

In the following formulation, the graph structure that connects the view cells is exactly the same as in the shortest itinerary formulation above. The graph $G = (\mathcal{V}, E)$ represents the connectivity between the view cells constructed according to the object geometries and robot movement (local) to travel from one cell to another neighboring cell. The cost function $c_e, e = \langle i_1, i_2 \rangle \in E$ represents the traveling cost of e from cell i_1 to i_2 . (We could make some metric assumptions such as the triangle inequalities although it is not necessarily the situation, especially for general cases where the robot has complex kinematics.)

In the following, we define the binary variable, y_i , as the indicator whether to take a view at view cell i ; we define the binary variable, z_e , as the indicator whether to include the edge, $e = \langle i_1, i_2 \rangle$, in the constructed Steiner tree. Thus, the IP formulation for the traveling VPP problem is given as: (assuming that we know the start position s of the robot)

Traveling VPP (IP):

$$\begin{aligned} & \text{Minimize} && w_v \sum_{i \in \mathcal{V}} y_i + w_p \sum_{e \in E} c_e z_e \\ & \text{Subject to:} && \\ & \forall j \in \mathcal{S} : && \sum_{i \in \mathcal{V}: i \text{ 'sees' } j} y_i \geq 1 \\ & \forall i \in \mathcal{V}, \forall T : && \sum_{e \in \delta(T)} z_e \geq y_i \\ & && y_i, z_e \in \{0, 1\} \end{aligned}$$

The first set of constraints above is the same as the VPP problem that requires each surface patch of the object, $j \in \mathcal{S}$, is covered by at least one view, ≥ 1 . The second set of constraints, modified from the SI problem, indicates that for view planned ($y_j = 1$), there must be some connection to the start position of the robot s .

In the following, as is the common practice in the integer programming research area, we relax the integral variables to be positive reals⁵ and derive the linear program (LP) formulation:

$$\begin{aligned} & \underline{\text{LP:}} \\ & \text{Minimize} && w_v \sum_{i \in \mathcal{V}} y_i + w_p \sum_{e \in E} c_e z_e \\ & \text{Subject to:} && \\ & \forall j \in \mathcal{S} : && \sum_{i \in \mathcal{V}: i \text{ 'sees' } j} y_i \geq 1 \quad (3) \\ & \forall i \in \mathcal{V}, \forall T : && \sum_{e \in \delta(T)} z_e \geq y_i \quad (4) \\ & && y_i, z_e \geq 0, i \in \mathcal{V}, e \in E \end{aligned}$$

Now that we associate the dual variables, α_j , to the constraints (3), and associate the dual variables, β_{iT} , to constraints (4). The corresponding dual program (DP) is given as:

$$\begin{aligned} & \underline{\text{DP:}} \\ & \text{Maximize} && \sum_{j \in \mathcal{S}} \alpha_j \\ & \text{Subject to:} && \\ & \forall i \in \mathcal{V} : && \sum_{j \in \mathcal{S}: i \text{ 'sees' } j} \alpha_j - \sum_{T: i \in T} \beta_{iT} \\ & && \leq w_v \quad (5) \\ & \forall e \in E : && \sum_{i \in \mathcal{V}} \sum_{T: i \in T \wedge e \in \delta(T)} \beta_{iT} \\ & && \leq w_p c_e \quad (6) \\ & && \alpha_j, \beta_{iT} \geq 0, j \in \mathcal{S}, T \subseteq \mathcal{V} : s \notin T \end{aligned}$$

Following [3], we interpret the dual variables, α_j , as the price that surface patch j would like to pay to take the view, at the cost of w_v , and to construct a Steiner tree connection. The part for constructing the Steiner tree connection is given by β_{iT} .

⁵In the integer programming research domain, a common practice is to relax integral variables to positive reals, to solve the corresponding problem accordingly, and to recover integer solutions.

The unified formulation given above provides us with a way of solving optimally the solution, as opposed to previous work based on a two-level approach. However, the larger number of constraints above makes it rather difficult and time consuming for practical implementations. In the following, we introduce a fast primal dual algorithm and analyze its performance.

III. PRIMAL DUAL ALGORITHM

A nice feature of the unified formulation is that the two sets of constraints above provide a way of dividing the problem into two stages, namely, the view planning stage and the Steiner tree construction stage, while maintaining the tradeoff between traveling and sensing costs. To maintain the tradeoffs between the two types of constraints is a major advantage over to treat them independently in the previous two-level approach. This motivates our primal dual algorithm to modify the constraints to address nicely the tradeoff described below.

The algorithm we propose works only in DP, the dual part of the problem, and finds feasible (integral) solution to the primal. Instead of the original problem, we solve a modified version, the view planning without traveling cost problem. However, we modify the costs of the views according to the distance from this view cell to the robot start position. That is, we modify the constraints of DP (5) to be:

$$\forall i \in \mathcal{V} : \sum_{j \in \mathcal{S}: i \text{ 'sees' } j} \alpha_j \leq w_v + w_p c_{is} \quad (7)$$

In the above, we use c_{is} to denote the path distance between view cell, i , and the robot starting position, s . A nice feature of this modification is that we weights differently the views according to the distance for the robot to travel. Intuitively, this modification means that if we decide to take a view at a far distance, it should be able to cover a large number of surface patches. In essence, this modification works similarly to the primal dual algorithm for the connected FL problem [3], where a facility is chosen to (tentatively) open if it can serve a lot of clients.

The algorithm works in two phases. In Phase I, we increase all the dual variables α_j synchronously from value 0 until one of the inequalities

(7), say for $i' \in \mathcal{V}$, becomes equality. (In case more than one equalities become tight, we can arbitrarily choose one of them.) We then choose i' as the view planned and freeze all the dual variables associated with it. We keep increasing all the other unfrozen dual variables until all the surface patches are covered/frozen. In Phase II, we construct the Steiner tree connections between the views planned in Phase I. The method we use is the one in [4]. We refer to [4] for details.

In the following, we illustrate this algorithm using the simple example in Fig. 3. For a triangle object, we want to choose a subset of $\{v_1, v_2, v_3, v_4\}$, connected to the starting position $s = v_4$ by a path, to completely cover the three surface patches $\{s_1, s_2, s_3\}$. (The dual variables associated are α_1, α_2 , and α_3 respectively.) The weights of sensing and movement for this example are equal.

The constraints (7) are changed to be:

$$\text{For } v_1 : \alpha_1 \leq 2$$

$$\text{For } v_2 : \alpha_2 \leq 2$$

$$\text{For } v_3 : \alpha_1 + \alpha_2 \leq 7$$

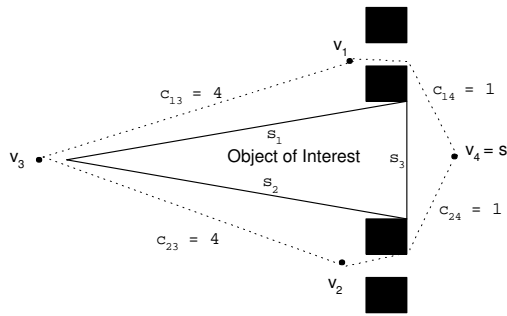
$$\text{For } v_4 : \alpha_3 \leq 1$$

At Phase I, we initialize α_1, α_2 , and α_3 to 0. We then increase all of them simultaneously until one of the above inequalities becomes tight. Obviously, when $\alpha_3 = 1, \alpha_1 = \alpha_2 = 2$, the constraints associated with view cells v_1, v_2 , and v_4 are tight. We output these three as the results of Phase I (i.e., we plan to take views v_1, v_2 , and v_4 , Fig. 3 (b)). At Phase II, we plan the itinerary between these three views, Fig. 3 (c).

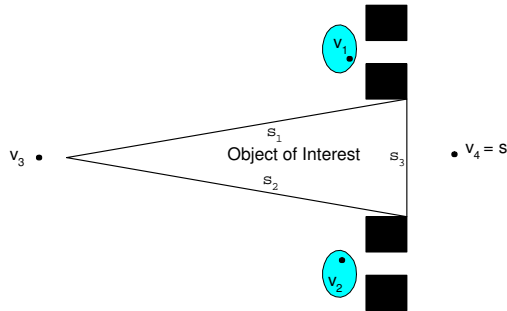
As shown in Fig. 1, our primal dual algorithm is able to find the optimal solution while the two-level approach performs poorly. The algorithmic performance verifies the unified formulation of Traveling VPP and the power of our primal dual algorithm.

IV. CONCLUSION AND FUTURE WORK

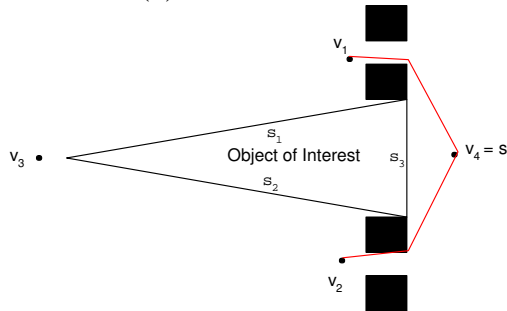
In this paper, we addressed the Traveling VPP problem of planning the set of connected views to completely inspect an object with minimum costs. The costs considered include the sensing cost, which is proportional to the number of views planned, and traveling cost, which is bounded by a constant ratio of the length of the Steiner connection between these views. As opposed to the common approach



(a) The problem setting



(b) Phase I result



(c) Phase II result

Fig. 3. Illustration of the primal dual algorithm. (a) A simple example: three surface patches, $\{s_1, s_2, s_3\}$, are to be covered by a subset of views, $\{v_1, v_2, v_3, v_4\}$. (b) The Phase I result by the primal dual algorithm: v_1 and v_2 are chosen (circled in the figure) as the views to take. (c) The Phase II result: the path (line segments in the figure) to connect the robot start position s and the two planned views v_1 and v_2 .

of decomposing the problem into two VP and SI problems and solving independently, we proposed a unified approach. By attacking the problem using a unified formulation, first as IP then as LP relaxation, and deriving its DP, we showed how to address the tradeoff between sensing and traveling by the DP constraint modifications. We designed a primal dual algorithm to consolidate this idea, which solves optimally a counterexample to the two-level approach.

It remains to verify theoretically that our algo-

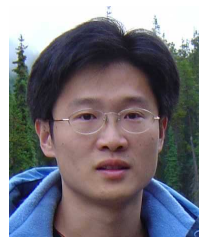
rithm indeed gives a constant approximation ratio for worse case scenarios. We are currently investigating the primal dual schema and (relaxed) complement slackness conditions for a proof. We would also like to improve our solution by devising a more sophisticated algorithm. For example, one possible improvement is to iteratively adjust the dual Steiner tree constraints to approach the optimal solution.

ACKNOWLEDGMENT

The author would like to thank the editor and the anonymous reviewer for valuable comments and suggestions.

REFERENCES

- [1] T. Danner and L. Kavraki. "Randomized Planning for Short Inspection Paths". In *Proc. of IEEE International Conference on Robotics and Automation*, 2002.
- [2] M. Garey and D. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. San Francisco: W. H. Freeman, 1979.
- [3] C. Swamy and A. Kumar. *Primal-Dual Algorithms for Connected Facility Location Problems*. *Algorithmica* 40(4): 245-269 (2004).
- [4] M. Goemans and D. Williamson. *A General Approximation Technique for Constrained Forest Problems*. *SIAM Journal on Computing*. 24(2), pp. 296-317.
- [5] L. Kavraki, P. Svestka, J.C. Latombe, and M. Overmars. *Probabilistic Roadmaps for Path Planning in High-Dimensional Configuration Spaces*. *IEEE Transactions on Robotics and Automation*, 12(4):566-580, 1996.



Pengpeng Wang (S'00) received the Bachelor of Science degree from University of Science and Technology of China, Hefei, P.R. China in 1998, and the Master of Applied Science in 2003 from Simon Fraser University, Burnaby, Canada, where he is currently working towards the Ph.D. degree in electrical engineering.

His research interests include sensor and model based robot motion/manipulation planning, mathematical optimization, computer vision, pattern recognition, and robotics.

Mr. Wang is a recipient of the Postgraduate Scholarships from the National Sciences and Engineering Research Council of Canada.

Robust Fault Diagnosis for Satellite Attitude Systems Using Neural State Space Models

Qing Wu, *Student Member, IEEE*

Abstract—In this paper, a robust fault detection and diagnosis (FDD) scheme using neural state space (NSS) models has been developed for a class of nonlinear dynamic systems. The neural state space models are adopted to estimate the modeling uncertainties in the states and outputs of the system. Subsequently, a residual signal is generated to identify the characteristics of the fault. Moreover, the robustness, sensitivity and stability properties of the proposed fault detection and diagnosis scheme are rigorously investigated in a unified framework. Finally, the NSS model-based FDD scheme is applied to a satellite attitude control system, and the simulation results demonstrate its good performance.

Index Terms—Fault diagnosis, neural networks, aerospace applications.

I. INTRODUCTION

MODERN automation technologies are not only required to ensure the stability and desired performance of a system in normal operating conditions, but also need to guarantee the safety and reliability of the system in the presence of faults. A fault is characterized by an unpermitted and unpredictable deviation in the system dynamics. Prompt fault detection and diagnosis can gain more time to accommodate the fault by using hardware redundancy or analytical redundancy techniques. Therefore, it is crucial to develop a fault diagnosis strategy capable of promptly detecting the occurrence of the fault (fault detection), correctly determining the location of the fault (fault isolation), and precisely estimating the characteristics of the fault (fault identification).

Due to these specifications, the research activities dealing with fault detection and diagnosis (FDD) have received significant attention over the last two decades. Books [1], [2], [3] and survey papers [4], [5], [6] provide detailed overviews of the contributions of fault detection and diagnosis using analytical redundancy techniques.

With respect to the FDD methodologies, model-based methods and process history based methods are two main categories. In the model-based FDD methods, observer-based and system identification-based approaches [7] are most frequently used (e.g. adaptive observer [8], neural adaptive observer [9], [10], [11] and neural networks [12], [13], [14], [15], [16], etc). Additionally, in practice, nonlinearities and modeling uncertainties universally exist, which motivates the development of robust fault detection and diagnosis for nonlinear systems [3], [17].

Moreover, due to the increasing complexity and autonomy of spacecraft, more sophisticated control systems are utilized

in the flight launches, orbit control, and attitude control of the spacecraft. This trend inevitably raises the probability of the occurrence of faults. Therefore, considering the safety and economic issues, fault diagnosis technologies for the spacecraft become increasingly important.

In this paper, a nonlinear observer, which is based on neural state space (NSS) models, is developed for the purpose of robust fault detection and diagnosis. The advantage of this FDD scheme is that it is able not only to detect and identify various faults with high accuracy, but also to analyze the analytical properties of the FDD scheme conveniently. In order to test the NSS model-based FDD scheme, a multi-input multi-output (MIMO) satellite attitude control system with abrupt and incipient faults is studied. Simulation results show the effectiveness of the proposed FDD scheme.

II. PROBLEM STATEMENT

The class of dynamic systems under study is represented by the following discrete-time state space model:

$$\begin{aligned} x(k+1) &= Ax(k) + B(Cx(k), u(k)) \\ &\quad + \eta(x(k), u(k), k) + f_x(y(k), u(k), k) \\ y(k) &= Cx(k) + \mu(x(k), u(k), k) + f_y(y(k), u(k), k) \end{aligned}$$

where $x \in \mathbb{R}^n$ is the state vector; $u \in \mathbb{R}^m$ is the input vector; $y \in \mathbb{R}^p$ is the measurable output vector of the system, respectively. The smooth vector fields contain $\eta : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^+ \rightarrow \mathbb{R}^n$, $\mu : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^+ \rightarrow \mathbb{R}^p$, $f_x : \mathbb{R}^p \times \mathbb{R}^m \times \mathbb{R}^+ \rightarrow \mathbb{R}^n$, and $f_y : \mathbb{R}^p \times \mathbb{R}^m \times \mathbb{R}^+ \rightarrow \mathbb{R}^p$, and $A \in \mathbb{R}^{n \times n}$, $C \in \mathbb{R}^{p \times n}$ are known state and output matrices. The vector $\eta = [\eta_1, \dots, \eta_n]^\top$ includes plant modeling uncertainties and disturbances, $\mu = [\mu_1, \dots, \mu_p]^\top$ represents sensor modeling uncertainties and noise, f_x denotes a state fault function caused in the actuator and/or in the system dynamics, and f_y specifies an output fault function caused in the sensor. In general, it is reasonable to make no further assumption about the characteristics of faults and uncertainties but rather to consider them as nonlinear functions of the system input, output, and time. For simplicity of presentation, in the rest of this paper, f_x , f_y , η and μ are only written as functions of the discrete time, k .

Throughout this paper, for analyzing the properties of the FDD scheme conveniently, we introduce the following assumptions.

Assumption 1: The observability matrix related to the pair (A, C) is full rank.

Assumption 2: Each element of the state and output uncertainty η and μ are unstructured and bounded (i.e., there are

two known vectors $\bar{\eta} = [\bar{\eta}_1, \dots, \bar{\eta}_n]^\top$ and $\bar{\mu} = [\bar{\mu}_1, \dots, \bar{\mu}_p]^\top$ such that $|\eta_i(k)| \leq \bar{\eta}_i$, and $|\mu_i(k)| \leq \bar{\mu}_i$, where $k \in \mathcal{K}$. $\mathcal{K} \subset \mathbb{R}^+$ is the time interval prior to the occurrence of any fault (i.e., $\mathcal{K} = [0, \min(K_x, K_y)]$, where K_x and K_y denote the beginning time of a state fault and an output fault, respectively). Hence, $\|\eta\| \leq (\sum_{i=1}^n \bar{\eta}_i^2)^{1/2} = \eta_0$ and $\|\mu\| \leq (\sum_{i=1}^p \bar{\mu}_i^2)^{1/2} = \mu_0$.

III. FAULT DIAGNOSIS STRATEGY

In this section, a robust fault diagnosis scheme is developed for detecting and identifying faults of the class of systems described by (1). The proposed fault diagnosis scheme is constructed such that it is capable of not only detecting, but also of identifying a fault through observing the residual. The development consists of two steps: First, a NSS model-based observer is designed to estimate the uncertainties. Then, a residual is generated to detect and diagnose faults.

A. NSS Model-Based Observer

Based on the plant representation (1), a diagnostic observer is proposed as follows

$$\begin{aligned} \hat{x}(k+1) &= A\hat{x}(k) + B(Cx(k), u(k)) + K_0(y(k) - \hat{y}(k)) \\ &\quad + \hat{\theta}_x(k), \quad \hat{x}(0) = x_0 \\ \hat{y}(k) &= C\hat{x}(k) + \hat{\theta}_y(k) \end{aligned} \quad (2)$$

where $\hat{x} \in \mathbb{R}^n$ and $\hat{y} \in \mathbb{R}^p$ denote the state and output vectors of the observer, respectively. The i th element of the saturation function vectors $\hat{\theta}_x = [\hat{\theta}_x^{(1)}, \dots, \hat{\theta}_x^{(n)}]^\top$ and $\hat{\theta}_y = [\hat{\theta}_y^{(1)}, \dots, \hat{\theta}_y^{(p)}]^\top$ are only used to approximate the η_i and μ_i respectively. They are expressed as

$$\hat{\theta}_x^{(i)}(k) = \begin{cases} \theta_x^{(i)}(k) & \text{if } |\theta_x^{(i)}(k)| < \bar{\eta}_i \\ \bar{\eta}_i & \text{if } \theta_x^{(i)}(k) \geq \bar{\eta}_i \\ -\bar{\eta}_i & \text{if } \theta_x^{(i)}(k) \leq -\bar{\eta}_i \end{cases} \quad (3)$$

and

$$\hat{\theta}_y^{(i)}(k) = \begin{cases} \theta_y^{(i)}(k) & \text{if } |\theta_y^{(i)}(k)| < \bar{\mu}_i \\ \bar{\mu}_i & \text{if } \theta_y^{(i)}(k) \geq \bar{\mu}_i \\ -\bar{\mu}_i & \text{if } \theta_y^{(i)}(k) \leq -\bar{\mu}_i \end{cases} \quad (4)$$

where $\theta_x^{(i)}$ and $\theta_y^{(i)}$ are two neural state space models [18]

$$\begin{aligned} \theta_x^{(i)}(k+1) &= W_x^{(i,1)}(k)\theta_x^{(i)}(k) + W_x^{(i,2)}(k) \\ &\quad \cdot \sigma(W_x^{(i,3)}(k)\theta_x^{(i)}(k) + W_x^{(i,4)}(k)e_x^{(i)}(k)) \end{aligned} \quad (5)$$

and

$$\begin{aligned} \theta_y^{(i)}(k+1) &= W_y^{(i,1)}(k)\theta_y^{(i)}(k) + W_y^{(i,2)}(k) \\ &\quad \cdot \sigma(W_y^{(i,3)}(k)\theta_y^{(i)}(k) + W_y^{(i,4)}(k)e_y^{(i)}(k)). \end{aligned} \quad (6)$$

In (5) and (6), $W_x^{(i,j)}$, $W_y^{(i,j)}$, for $j = 1, \dots, 4$ are corresponding weight vectors. $\sigma(x) = (1 - e^{-x})/(1 + e^{-x})$ is the hyperbolic tangent activation function. $e_x^{(i)}$ is the i th element of the state estimation error $e_x(k) = x(k) - \hat{x}(k)$ and $e_y^{(i)}$ is the i th element of the output estimation error $e_y(k) = y(k) - \hat{y}(k)$.

Subtracting the dynamics of the observer (2) from the system dynamics (1) obtains the dynamics of the state estimation error e_x and the output estimation error e_y

$$\begin{aligned} e_x(k+1) &= Ae_x(k) - K_0e_y(k) + \tilde{\theta}_x(k) + f_x(k) \\ e_y(k) &= Ce_x(k) + \tilde{\theta}_y(k) + f_y(k) \end{aligned} \quad (7)$$

where K_0 is chosen such that $A_0 = A - K_0C$ has all its eigenvalues in the unit circle. $\tilde{\theta}_x(k) = \eta(k) - \hat{\theta}_x(k)$ represents the state uncertainty estimation error and $\tilde{\theta}_y(k) = \mu(k) - \hat{\theta}_y(k)$ denotes the output uncertainty estimation error.

The residual used to diagnose faults is selected to be

$$\begin{aligned} r(k+1) &= e_x(k+1) - A_0e_x(k) \\ &= \tilde{\theta}_x(k) + f_x(k) - K_0\tilde{\theta}_y(k) - K_0f_y(k). \end{aligned} \quad (8)$$

Hence, if the accurate estimations of η and μ are achieved, the effect of the fault on the residual becomes evident. Moreover, if either a state fault or an output fault occurs individually, the residual is able to characterize the dynamics of the fault. However, if both of them occur simultaneously, the isolation of the state fault and output fault entails further work.

According to the properties of the neural state space models in [18], [19], there are two NSS models θ_x^* and θ_y^* with optimal weights W_x^* and W_y^* which can completely characterize the state uncertainty η and sensor uncertainty μ . Thus, the state uncertainty estimation error is $\tilde{\theta}_x = \theta_x^* - \hat{\theta}_x$, and the output uncertainty estimation error is $\tilde{\theta}_y = \theta_y^* - \hat{\theta}_y$.

In order to investigate the dynamics of the estimation error, the i th elements of the uncertainty estimation error are also approximated by NSS models as follows

$$\begin{aligned} \tilde{\theta}_x^{(i)}(k+1) &= \tilde{W}_x^{(i,1)}(k)\tilde{\theta}_x^{(i)}(k) + \tilde{W}_x^{(i,2)}(k) \\ &\quad \cdot \sigma(\tilde{W}_x^{(i,3)}(k)\tilde{\theta}_x^{(i)}(k) + \tilde{W}_x^{(i,4)}(k)h^{(i)}(e_x(k))) \end{aligned} \quad (9)$$

$$\begin{aligned} \tilde{\theta}_y^{(i)}(k+1) &= \tilde{W}_y^{(i,1)}(k)\tilde{\theta}_y^{(i)}(k) + \tilde{W}_y^{(i,2)}(k) \\ &\quad \cdot \sigma(\tilde{W}_y^{(i,3)}(k)\tilde{\theta}_y^{(i)}(k) + \tilde{W}_y^{(i,4)}(k)g^{(i)}(e_y(k))) \end{aligned} \quad (10)$$

where the i th element of the state uncertainty estimation error and the output uncertainty estimation error are written as a nonlinear function of the state estimation error $h^{(i)}(e_x(k))$ and a nonlinear function of output estimation error $g^{(i)}(e_y(k))$, respectively.

Using the properties of the NSS models, (9) and (10) can be expressed as

$$\tilde{\theta}_x(k+1) = \tilde{M}_x(k)\tilde{\theta}_x(k) + \tilde{N}_x(k)h(e_x(k)) \quad (11)$$

$$\tilde{\theta}_y(k+1) = \tilde{M}_y(k)\tilde{\theta}_y(k) + \tilde{N}_y(k)g(e_y(k)) \quad (12)$$

where

$$\tilde{M}_x(k) = \tilde{W}_x^{(1)}(k) + \tilde{W}_x^{(2)}(k)(I - \Delta_x)\tilde{W}_x^{(3)}(k)$$

$$\tilde{N}_x(k) = \tilde{W}_x^{(2)}(k)(I - \Delta_x)\tilde{W}_x^{(4)}(k)$$

$$\tilde{M}_y(k) = \tilde{W}_y^{(1)}(k) + \tilde{W}_y^{(2)}(k)(I - \Delta_y)\tilde{W}_y^{(3)}(k)$$

$$\tilde{N}_y(k) = \tilde{W}_y^{(2)}(k)(I - \Delta_y)\tilde{W}_y^{(4)}(k)$$

in which, $\Delta_x = \text{diag}\{\delta_x^{(i)}\}$, and $\Delta_y = \text{diag}\{\delta_y^{(i)}\}$ ($i = 1, \dots, n$) with

$$\delta_x^{(i)} = 1 - \frac{\sigma(\tilde{W}_x^{(i,3)}(k)\tilde{\theta}_x^{(i)}(k) + \tilde{W}_x^{(i,4)}(k)h^{(i)}(e_x(k)))}{\tilde{W}_x^{(i,3)}(k)\tilde{\theta}_x^{(i)}(k) + \tilde{W}_x^{(i,4)}(k)h^{(i)}(e_x(k))}$$

$$\delta_y^{(i)} = 1 - \frac{\sigma(\tilde{W}_y^{(i,3)}(k)\tilde{\theta}_y^{(i)}(k) + \tilde{W}_y^{(i,4)}(k)g^{(i)}(e_y(k)))}{\tilde{W}_y^{(i,3)}(k)\tilde{\theta}_y^{(i)}(k) + \tilde{W}_y^{(i,4)}(k)g^{(i)}(e_y(k))}.$$

By taking (7) and (11)-(12) into account, the dynamics of the estimation error in the absence of faults is

$$\begin{aligned} e_x(k+1) &= Ae_x(k) - K_0e_y(k) + \tilde{\theta}_x(k) \\ e_y(k) &= Ce_x(k) + \tilde{\theta}_y(k) \\ \tilde{\theta}_x(k+1) &= \tilde{M}_x(k)\tilde{\theta}_x(k) + \tilde{N}_x(k)h(e_x(k)) \\ \tilde{\theta}_y(k+1) &= \tilde{M}_y(k)\tilde{\theta}_y(k) + \tilde{N}_y(k)g(e_y(k)). \end{aligned} \quad (13)$$

B. Update Law and Its Convergence

The parameters of the NSS models in this work are updated using the extended Kalman filter (EKF) algorithm [18] as follows

$$\begin{aligned} K_i(k) &= P_i(k)H_i(k)[H_i(k)^\top P_i(k)H_i(k) + R_i(k)]^{-1} \\ P_i(k+1) &= P_i(k) - K_i(k)H_i(k)^\top P_i(k) \\ W_i(k+1) &= W_i(k) + K_i(k)E_i(k) \end{aligned} \quad (14)$$

where W_i is the associated weight of a NSS model. $E_i(k)$ is selected according to the NSS model in use. For example, as for the $\hat{\theta}_x(k)$, $E_i(k) = e_x^{(i)}(k)$, and as for the $\hat{\theta}_y(k)$, $E_i(k) = h^{(i)}(e_x(k)) - \tilde{\theta}_x^{(i)}(k)$. The superscript i represents the i th component of the vector. $K_i(k)$ is known as the Kalman gain, $P_i(k)$ is the covariance matrix of the state estimation error, and $R_i(k)$ is the estimated covariance matrix of noise. For SISO systems, $R_i(k)$ is calculated according to the following equation in [20]

$$R_i(k) = R_i(k-1) + [E_i(k)^2 - R_i(k-1)]/k. \quad (15)$$

$H_i(k)$ is the derivative of the output of a NSS model with respect to its weight W_i .

Convergence of the parameter update law is analyzed in the following theorem.

Theorem 1: The parameter update process is convergent, provided the following condition is guaranteed.

$$H_i(k)^\top P_i(k)H_i(k) > 0 \quad (16)$$

Proof: $R_i(k)$ keeps nonnegative in all iterations since it is the estimated variance of noise for SISO systems. Therefore, (16) yields

$$0 < H_i(k)^\top P_i(k)H_i(k) < 2[H_i(k)^\top P_i(k)H_i(k) + R_i(k)]$$

which is equivalent to

$$0 < H_i(k)^\top P_i(k)H_i(k)[H_i(k)^\top P_i(k)H_i(k) + R_i(k)]^{-1} < 2.$$

Based on the update law (14), (17) becomes

$$0 < H_i(k)^\top K_i(k) < 2. \quad (17)$$

A positive Lyapunov function candidate is considered:

$$V(k) = \frac{1}{2}E_i(k)^2 \quad (18)$$

The first order difference of (18) is

$$\begin{aligned} \Delta V(k) &= V(k+1) - V(k) \\ &= \frac{1}{2}[E_i(k+1)^2 - E_i(k)^2] \\ &= \Delta E_i(k) \left[E_i(k) + \frac{1}{2}\Delta E_i(k) \right]. \end{aligned} \quad (19)$$

The difference of $E_i(k)$ can be approximated by

$$\begin{aligned} \Delta E_i(k) &= \left[\frac{\partial E_i(k)}{\partial W_i} \right]^\top \Delta W_i \\ &= -H_i(k)^\top K_i(k)E_i(k). \end{aligned} \quad (20)$$

Thus, based on (17) and (20), (19) is further derived to be

$$\begin{aligned} \Delta V(k) &= \Delta E_i(k) \left[E_i(k) + \frac{1}{2}\Delta E_i(k) \right] \\ &= -H_i(k)^\top K_i(k) \left[1 - \frac{1}{2}H_i(k)^\top K_i(k) \right] E_i(k)^2 \\ &< 0. \end{aligned}$$

Therefore, if (16) is guaranteed, the parameter update process using the EKF algorithm is convergent.

Remark 1: In numerical computation, (16) is guaranteed via carefully choosing the initial value of $P_i(k)$. Usually, $P_i(0)$ is set to a large identity matrix.

IV. ANALYTICAL PROPERTIES

A. Robustness

The robustness property of a fault diagnosis method refers to its ability to prevent false alarms. In [17], the robustness is achieved by using a dead-zone operator in the learning algorithm, which is aimed to be insensitive to the system uncertainties. However, in this work, saturation functions are used to confine the estimators only to approximate the system uncertainties so that the residual is able to characterize the dynamics of the fault.

The estimation error dynamics prior to any fault (i.e., $k < \mathcal{K}$) is described by

$$\begin{aligned} e_x(k+1) &= Ae_x(k) - K_0e_y(k) + \tilde{\theta}_x(k), \quad e_x(0) = 0 \\ e_y(k) &= Ce_x(k) + \tilde{\theta}_y(k). \end{aligned} \quad (21)$$

Based on Assumption 1 and (3)-(4), we have

$$\|\tilde{\theta}_x(k)\| \leq 2\eta_0 \quad (22)$$

$$\|\tilde{\theta}_y(k)\| \leq 2\mu_0. \quad (23)$$

Suppose there exists a finite k_e ($0 < k_e < \mathcal{K}$) such that $\|e_y(k)\| < \epsilon$ for $0 \leq k \leq k_e$ and

$$\|e_y(k_e+1)\| \geq \epsilon \quad (24)$$

where $\epsilon = 2\kappa(\|K_0\|\mu_0 + \eta_0) + 2\mu_0$, and $\sum_{j=0}^{\infty} \|CA_0^j\| < \kappa$. The constant κ exists because the eigenvalues of A_0 are all in a unit circle. ϵ is used to indicate the occurrence of a fault.

Based on the estimation error dynamics (21), in the interval $0 \leq k \leq k_e + 1$, the output estimation error satisfies

$$e_y(k+1) = CA_0 e_x(k) - CK_0 \tilde{\theta}_y(k) + C\tilde{\theta}_x(k) + \tilde{\theta}_y(k+1). \quad (25)$$

Solving the discrete time equation and taking the norm of the output estimation error, we obtain

$$\begin{aligned} \|e_y(k_e+1)\| &\leq \left\| C \sum_{j=0}^{k_e} A_0^{k_e-j} K_0 \tilde{\theta}_y(j) \right\| + \|\tilde{\theta}_y(k_e+1)\| \\ &\quad + \left\| C \sum_{j=0}^{k_e} A_0^{k_e-j} \tilde{\theta}_x(j) \right\| \\ &\leq C \sum_{j=0}^{k_e} A_0^{k_e-j} \|K_0\| \|\tilde{\theta}_y(j)\| + \|\tilde{\theta}_y(k_e+1)\| \\ &\quad + C \sum_{j=0}^{k_e} A_0^{k_e-j} \|\tilde{\theta}_x(j)\| \\ &< 2\kappa(\|K_0\|\mu_0 + \eta_0) + 2\mu_0 = \epsilon. \end{aligned}$$

This inequality clearly contradicts (24), which implies that for $k < K$, the output estimation error remains within a bound.

B. Sensitivity

The sensitivity of a FDD scheme involves the capability to correctly determine the existence of the fault. The following theorem specifies the set of state and output fault that can be detected by the proposed FDD scheme.

Theorem 2: Consider the fault diagnosis scheme described by (2).

- 1) If there exists a discrete time $k_x > 0$ such that the state fault $f_x(k)$ satisfies condition

$$\left\| C \sum_{j=K_x}^{K_x+k_x} A_0^{K_x+k_x-j} f_x(j) \right\| \geq 2\epsilon \quad (26)$$

then the state fault will be detected, that is,

$$\|e_y(K_x + k_x + 1)\| \geq \epsilon.$$

- 2) If there exists a discrete time $k_y > 0$ such that the sensor fault $f_y(k)$ satisfies condition

$$\left\| f_y(K_y + k_y + 1) - C \sum_{j=K_y}^{K_y+k_y} A_0^{K_y+k_y-j} K_0 f_y(j) \right\| \geq 2\epsilon \quad (27)$$

then the sensor fault will be detected, that is,

$$\|e_y(K_y + k_y + 1)\| \geq \epsilon.$$

Proof: When only a state fault occurs, after the occurrence of the fault ($k > K_x$), the output estimation error is

$$\begin{aligned} e_y(K_x + k + 1) &= -C \sum_{j=0}^{K_x+k} A_0^{K_x+k-j} K_0 \tilde{\theta}_y(j) + \tilde{\theta}_y(K_x + k + 1) \\ &\quad + C \sum_{j=0}^{K_x+k} A_0^{K_x+k-j} \tilde{\theta}_x(j) + C \sum_{j=K_x}^{K_x+k} A_0^{K_x+k-j} f_x(j). \end{aligned}$$

Using the triangle inequality, we obtain

$$\begin{aligned} \|e_y(K_x + k_x + 1)\| &\geq - \left\| C \sum_{j=0}^{K_x+k_x} A_0^{K_x+k_x-j} K_0 \tilde{\theta}_y(j) \right\| \\ &\quad - \|\tilde{\theta}_y(K_x + k_x + 1)\| \\ &\quad - \left\| C \sum_{j=0}^{K_x+k_x} A_0^{K_x+k_x-j} \tilde{\theta}_x(j) \right\| \\ &\quad + \left\| C \sum_{j=K_x}^{K_x+k_x} A_0^{K_x+k_x-j} f_x(j) \right\| \\ &\geq -\epsilon + \left\| C \sum_{j=K_x}^{K_x+k_x} A_0^{K_x+k_x-j} f_x(j) \right\|. \end{aligned}$$

Therefore, if the condition (26) is satisfied, then $\|e_y(K_x + k_x + 1)\| \geq \epsilon$, which implies that the state fault is detected.

Similarly, after the occurrence of a sensor fault ($k > K_y$), the output estimation error is

$$\begin{aligned} e_y(K_y + k + 1) &= -C \sum_{j=K_y}^{K_y+k} A_0^{K_y+k-j} K_0 f_y(j) + f_y(K_y + k + 1) \\ &\quad - C \sum_{j=0}^{K_y+k} A_0^{K_y+k-j} K_0 \tilde{\theta}_y(j) + \tilde{\theta}_y(K_y + k + 1) \\ &\quad + C \sum_{j=0}^{K_y+k} A_0^{K_y+k-j} \tilde{\theta}_x(j). \end{aligned}$$

Hence, based on the triangle inequality, we obtain

$$\begin{aligned} \|e_y(K_y + k_y + 1)\| &\geq - \left\| C \sum_{j=0}^{K_y+k_y} A_0^{K_y+k_y-j} \tilde{\theta}_x(j) \right\| \\ &\quad - \left\| C \sum_{j=0}^{K_y+k_y} A_0^{K_y+k_y-j} K_0 \tilde{\theta}_y(j) \right\| - \|\tilde{\theta}_y(K_y + k_y + 1)\| \\ &\quad + \left\| f_y(K_y + k_y + 1) - C \sum_{j=K_y}^{K_y+k_y} A_0^{K_y+k_y-j} K_0 f_y(j) \right\| \\ &\geq -\epsilon + \left\| f_y(K_y + k_y + 1) - C \sum_{j=K_y}^{K_y+k_y} A_0^{K_y+k_y-j} K_0 f_y(j) \right\|. \end{aligned}$$

Therefore, if the sensor fault satisfies condition (27) then $\|e_y(K_y + k_y + 1)\| \geq \epsilon$ and the output fault is detected.

C. Stability

In order to analyze the stability of the fault diagnosis scheme, another two assumptions are needed.

Assumption 3: In the parameter update process, the weights of the NSS models are bounded and guarantee

$$\begin{aligned} \alpha_1 I &\leq I - 2\tilde{M}_x(k)^\top \tilde{M}_x(k) \leq \alpha_2 I \\ \beta_1 I &\leq (1 - 2\zeta^2)I - 2\tilde{M}_y(k)^\top \tilde{M}_y(k) \leq \beta_2 I \end{aligned}$$

where $\alpha_1, \alpha_2, \beta_1$ and β_2 are all positive constants and ζ is defined below.

Assumption 4: If the parameter update process is convergent, following two inequalities are satisfied.

$$\begin{aligned} h(e_x(k))^\top \tilde{N}_x(k)^\top \tilde{N}_x(k) h(e_x(k)) &\leq \xi^2 \|e_x(k)\|^2 \\ g(e_y(k))^\top \tilde{N}_y(k)^\top \tilde{N}_y(k) g(e_y(k)) &\leq \zeta^2 \|e_y(k)\|^2 \end{aligned}$$

where ξ and ζ are both positive constants.

Considering the dynamics of the estimation error (13), we have the following theorem regarding the convergence of the estimation error.

Theorem 3: The equilibrium point of the estimation error dynamics (13), defined as $z_{eq} = [e_x^\top \tilde{\theta}_x^\top \tilde{\theta}_y^\top]^\top = 0$, is globally uniformly stable.

Proof: The Lyapunov function candidate is constructed by synthesizing three positive Lyapunov function candidates

$$\tilde{V}(k) = \tilde{V}_1(k) + \tilde{V}_2(k) + \tilde{V}_3(k).$$

The Lyapunov function candidate $\tilde{V}_1(k)$ is designed as

$$\tilde{V}_1(k) = e_x(k)^\top P_e e_x(k) \quad (28)$$

where P_e is a solution to the Lyapunov equation

$$P_e - A_0^\top P_e A_0 = Q_e \quad (29)$$

for a given symmetric and positive matrix Q_e . The solution of (29) exists since A_0 has all its eigenvalues inside the unit circle.

By taking (13) into account, the first order difference of $\tilde{V}_1(k)$ is given by

$$\begin{aligned} \Delta \tilde{V}_1(k) &= e_x(k+1)^\top P_e e_x(k+1) - e_x(k)^\top P_e e_x(k) \\ &= [A_0 e_x(k) + \tilde{\theta}_x(k) - K_0 \tilde{\theta}_y(k)]^\top P_e \\ &\quad \cdot [A_0 e_x(k) + \tilde{\theta}_x(k) - K_0 \tilde{\theta}_y(k)] - e_x(k)^\top P_e e_x(k) \\ &= e_x(k)^\top (A_0^\top P_e A_0 - P_e) e_x(k) + \tilde{\theta}_x(k)^\top P_e \tilde{\theta}_x(k) \\ &\quad + \tilde{\theta}_y(k)^\top K_0^\top P_e K_0 \tilde{\theta}_y(k) + 2e_x(k)^\top A_0^\top P_e \tilde{\theta}_x(k) \\ &\quad - 2e_x(k)^\top A_0^\top P_e K_0 \tilde{\theta}_y(k) - 2\tilde{\theta}_x(k)^\top P_e K_0 \tilde{\theta}_y(k) \\ &\leq -e_x(k)^\top Q_e e_x(k) + \tilde{\theta}_x(k)^\top P_e \tilde{\theta}_x(k) \\ &\quad + \tilde{\theta}_y(k)^\top K_0^\top P_e K_0 \tilde{\theta}_y(k) + 2\|e_x(k)\| \|A_0^\top P_e\| \|\tilde{\theta}_x(k)\| \\ &\quad + 2\|e_x(k)\| \|A_0 P_e K_0\| \|\tilde{\theta}_y(k)\| + 2\|\tilde{\theta}_x(k)\| \|P_e K_0\| \|\tilde{\theta}_y(k)\| \\ &\leq e_x(k)^\top [-Q_e + (\|A_0^\top P_e\| + \|A_0^\top P_e K_0\|) I] e_x(k) \\ &\quad + \tilde{\theta}_x(k)^\top [P_e + (\|A_0^\top P_e\| + \|A_0^\top P_e K_0\| + \|P_e K_0\|) I] \tilde{\theta}_x(k) \\ &\quad + \tilde{\theta}_y(k)^\top [K_0^\top P_e K_0 + (\|A_0^\top P_e K_0\| + \|P_e K_0\|) I] \tilde{\theta}_y(k) \\ &\leq -\lambda_{\min}(Q_1) \|e_x(k)\|^2 + \lambda_{\max}(Q_2) \|\tilde{\theta}_x(k)\|^2 \\ &\quad + \lambda_{\max}(Q_3) \|\tilde{\theta}_y(k)\|^2 \end{aligned} \quad (30)$$

where

$$\begin{aligned} Q_1 &= Q_e - (\|A_0^\top P_e\| + \|A_0^\top P_e K_0\|) I \\ Q_2 &= P_e + (\|A_0^\top P_e\| + \|A_0 P_e K_0\| + \|P_e K_0\|) I \\ Q_3 &= K_0^\top P_e K_0 + (\|A_0^\top P_e K_0\| + \|P_e K_0\|) I \end{aligned}$$

are three positive definite matrices.

Design a positive function $V_2(k) = \tilde{\theta}_x(k)^\top \tilde{\theta}_x(k)$, by using the assumption 3 and 4, its first order difference is

$$\begin{aligned} \Delta V_2(k) &= \tilde{\theta}_x(k+1)^\top \tilde{\theta}_x(k+1) - \tilde{\theta}_x(k)^\top \tilde{\theta}_x(k) \\ &= \tilde{\theta}_x(k)^\top (\tilde{M}_x(k)^\top \tilde{M}_x(k) - I) \tilde{\theta}_x(k) + \xi^2 \|e_x(k)\|^2 \\ &\quad + 2(\tilde{\theta}_x(k) \tilde{M}_x(k))^\top \tilde{N}_x(k) h(e_x(k)) \\ &\leq \tilde{\theta}_x(k)^\top (2\tilde{M}_x(k) \tilde{M}_x(k) - I) \tilde{\theta}_x(k) + 2\xi^2 \|e_x(k)\|^2 \\ &\leq -\alpha_1 \|\tilde{\theta}_x(k)\|^2 + 2\xi^2 \|e_x(k)\|^2. \end{aligned} \quad (31)$$

Accordingly, the $\tilde{V}_2(k)$ is designed to be

$$\tilde{V}_2(k) = \frac{\tilde{\theta}_x(k)^\top \tilde{\theta}_x(k)}{\alpha_1} \lambda_{\max}(Q_2). \quad (32)$$

Similarly, another positive function is constructed to be $V_3(k) = \tilde{\theta}_y(k)^\top \tilde{\theta}_y(k)$. Based on Assumption 3 and 4, the first order difference of $V_3(k)$ satisfies

$$\begin{aligned} \Delta V_3(k) &= \tilde{\theta}_y(k+1)^\top \tilde{\theta}_y(k+1) - \tilde{\theta}_y(k)^\top \tilde{\theta}_y(k) \\ &= \tilde{\theta}_y(k)^\top (\tilde{M}_y(k)^\top \tilde{M}_y(k) - I) \tilde{\theta}_y(k) + \zeta^2 \|e_y(k)\|^2 \\ &\quad + 2(\tilde{\theta}_y(k) \tilde{M}_y(k))^\top \tilde{N}_y(k) g(e_y(k)) \\ &\leq \tilde{\theta}_y(k)^\top (2\tilde{M}_y(k) \tilde{M}_y(k) - I) \tilde{\theta}_y(k) + 2\zeta^2 \|e_y(k)\|^2. \end{aligned}$$

Based on (21), there is

$$\zeta^2 \|e_y(k)\|^2 \leq 2\zeta^2 (\|C\|^2 \|e_x(k)\|^2 + \|\tilde{\theta}_y(k)\|^2).$$

Therefore, $\Delta V_3(k)$ is further derived to be

$$\begin{aligned} \Delta V_3(k) &\leq \tilde{\theta}_y(k)^\top [2\tilde{M}_y(k)^\top \tilde{M}_y(k) + (2\zeta^2 - 1)I] \tilde{\theta}_y(k) \\ &\quad + 2\zeta^2 \|C\|^2 \|e_x(k)\|^2 \\ &\leq -\beta_1 \|\tilde{\theta}_y(k)\|^2 + 2\zeta^2 \|C\|^2 \|e_x(k)\|^2. \end{aligned} \quad (33)$$

So, the $\tilde{V}_3(k)$ can be designed to be

$$\tilde{V}_3(k) = \frac{\tilde{\theta}_y(k)^\top \tilde{\theta}_y(k)}{\beta_1} \lambda_{\max}(Q_3). \quad (34)$$

Summarizing above results, the first order difference of the Lyapunov function $\tilde{V}(k)$ is as follows

$$\begin{aligned} \Delta \tilde{V}(k) &= \Delta \tilde{V}_1(k) + \Delta \tilde{V}_2(k) + \Delta \tilde{V}_3(k) \\ &\leq \left[-\lambda_{\min}(Q_1) + \frac{2\xi^2 \lambda_{\max}(Q_2)}{\alpha_1} \right. \\ &\quad \left. + \frac{2\zeta^2 \|C\|^2 \lambda_{\max}(Q_3)}{\beta_1} \right] \cdot \|e_x(k)\|^2. \end{aligned} \quad (35)$$

Hence, provided

$$\lambda_{\min}(Q_1) \geq \frac{2\xi^2 \lambda_{\max}(Q_2)}{\alpha_1} + \frac{2\zeta^2 \|C\|^2 \lambda_{\max}(Q_3)}{\beta_1} \quad (36)$$

is satisfied, $\Delta \tilde{V}(k) \leq 0$, that is, the equilibrium points of the estimation error dynamics $z_{eq} = [e_x^\top \tilde{\theta}_x^\top \tilde{\theta}_y^\top]^\top = 0$ is globally uniformly stable.

V. CASE STUDY

In this section, the proposed fault diagnosis scheme is applied to the dynamics of a satellite model [21], [22]. The moments of inertia about principal axes of this satellite are $J_1 = 17.581\text{kg m}^2$, $J_2 = 16.533\text{kg m}^2$, and $J_3 = 3.529\text{kg m}^2$, respectively. The orbital rate ω_0 is $7.2722 \times 10^{-5}\text{rad/s}$ and the altitude of the orbit is 37786km from the Earth's surface. When the satellite is a small one without flexible appendages, its Euler's equations of motion are given by

$$\begin{aligned} J_1\dot{\omega}_1 - (J_2 - J_3)\omega_2\omega_3 &= T_{u1} + T_{d1} \\ J_2\dot{\omega}_2 - (J_3 - J_1)\omega_3\omega_1 &= T_{u2} + T_{d2} \\ J_3\dot{\omega}_3 - (J_1 - J_2)\omega_1\omega_2 &= T_{u3} + T_{d3} \end{aligned} \quad (37)$$

where $\omega_1 \sim \omega_3$ are the body-axis components of angular velocity, $T_{u1} \sim T_{u3}$ are the control torques, and $T_{d1} \sim T_{d3}$ are the environmental disturbance torques.

For small attitude deviation from local vertical local horizontal orientation, the angle velocity can be expressed by yaw angle ψ , pitch angle θ , roll angle ϕ and orbital rate ω_0 . Thus, through transformation, the nominal model of a satellite attitude dynamics is written as

$$\begin{aligned} J_1\ddot{\phi} - \omega_0(J_1 - J_2 + J_3)\dot{\psi} + 4\omega_0^2(J_2 - J_3)\phi &= T_{u1} + T_{d1} \\ J_2\ddot{\theta} + 3\omega_0^2(J_1 - J_3)\theta &= T_{u2} + T_{d2} \\ J_3\ddot{\psi} + \omega_0(J_1 - J_2 + J_3)\dot{\phi} + \omega_0^2(J_2 - J_1)\psi &= T_{u3} + T_{d3}. \end{aligned} \quad (38)$$

When the state vector is chosen as $x = [\phi \ \dot{\phi} \ \theta \ \dot{\theta} \ \psi \ \dot{\psi}]^\top$, the (37) in the presence of uncertainties can thus be represented by (1) with matrices (A, C) :

$$A = \begin{bmatrix} 1 & \tau & 0 & 0 & 0 & 0 \\ a_{21} & 1 & 0 & 0 & 0 & a_{26} \\ 0 & 0 & 1 & \tau & 0 & 0 \\ 0 & 0 & a_{43} & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \tau \\ 0 & a_{62} & 0 & 0 & a_{65} & 1 \end{bmatrix}, \quad C^\top = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

where $a_{21} = \frac{4\omega_0^2(J_3 - J_2)\tau}{J_1}$, $a_{26} = \frac{\omega_0(J_1 - J_2 + J_3)\tau}{J_1}$, $a_{43} = \frac{3\omega_0^2(J_3 - J_1)\tau}{J_2}$, $a_{62} = \frac{\omega_0(J_2 - J_1 - J_3)\tau}{J_3}$, and $a_{65} = \frac{\omega_0^2(J_1 - J_2)\tau}{J_3}$. τ denotes the sampling time interval. While the state uncertainty is represented by

$$\eta = \begin{bmatrix} 0 & \frac{T_{d1}\tau}{J_1} & 0 & \frac{T_{d2}\tau}{J_2} & 0 & \frac{T_{d3}\tau}{J_3} \end{bmatrix}^\top$$

and the sensor uncertainty takes the form:

$$\mu = [\varsigma\phi \quad 0 \quad \varsigma\theta \quad 0 \quad \varsigma\psi \quad 0]^\top$$

where $\varsigma = 0.02\%$ is the coefficient of sensor uncertainties.

The performance of the proposed fault diagnosis scheme is demonstrated by the cases when an actuator fault or a sensor fault individually occurs at the different state and output of the nominal system. Figure 1 shows the diagnosis results when the abrupt fault f_{a2} occurs at state ϕ . Figure 2 displays the diagnosis results when the incipient fault f_{a5} occurs at output ψ .

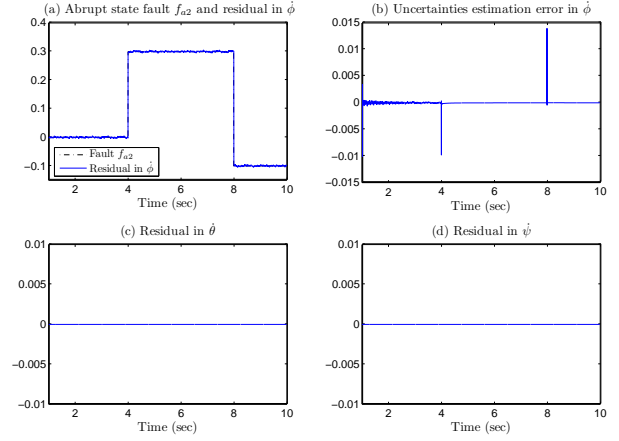


Fig. 1. (a) Time-behaviors of the fault f_{a2} (dash-dotted line) and the corresponding residual in the faulty state ϕ . (b) Time-behaviors of the uncertainty estimation error in the faulty state ϕ . (c) Time-behaviors of the residual in the healthy state θ . (d) Time-behaviors of the residual in the healthy state ψ .

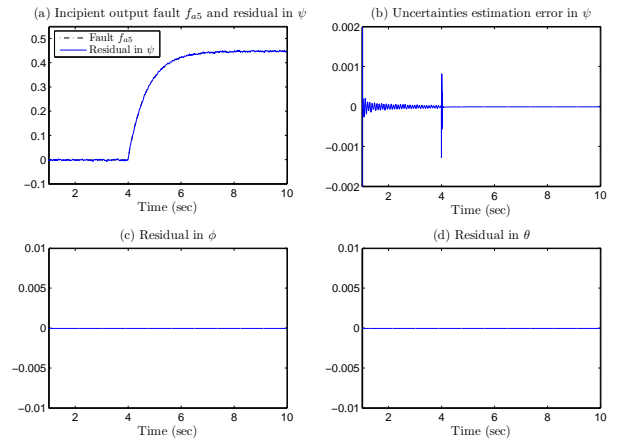


Fig. 2. (a) Time-behaviors of the fault f_{a5} (dash-dotted line) and the corresponding residual in the faulty output ψ . (b) Time-behaviors of the uncertainty estimation error in the faulty output ψ . (c) Time-behaviors of the residual in the healthy output ϕ . (d) Time-behaviors of the residual in the healthy output θ .

In these two figures, (1a) and (2a) illustrate that the residual is able to specify the characteristics of the fault with satisfactory accuracy in the presence of modeling uncertainties. (1b) and (2b) portray the uncertainty estimation errors in state ϕ and output ψ , respectively. The uncertainty estimation error can be used to indicate the occurrence of a fault because the uncertainty estimation error in the faulty state or faulty output significantly changes. Moreover, in (1c), (2c), (1d) and (2d), the residuals in the channels where no fault occurs remain zero or close to zero. These phenomena exhibit that the proposed FDD scheme is a reliable method to locate and identify the fault.

VI. CONCLUSIONS

This work investigates a robust fault diagnosis scheme using neural state space models for a class of nonlinear systems. The proposed FDD scheme uses a nonlinear observer to estimate

the system uncertainties in the states and outputs, and a residual is generated to locate and identify various faults. Theoretically, the robustness and sensitivity properties have been analyzed rigorously. Moreover, the stability of the FDD system is demonstrated by using the properties of the NSS models. Finally, the application of this FDD scheme to a satellite attitude system demonstrates its effectiveness.

REFERENCES

- [1] R.J. Patton, P.M. Frank and R.N. Clark, *Fault Diagnosis in Dynamic Systems: Theory and Applications*, Englewood Cliffs, NJ: Prentice-Hall, 1989.
- [2] R.J. Patton, R. Clark, R.N. Clark, *Issues of Fault Diagnosis for Dynamic Systems*, London: Springer-Verlag, 2000.
- [3] J. Chen and R.J. Patton, *Robust Model-based Fault Diagnosis for Dynamic Systems*, Boston: Kluwer Academic Publishers, 1999.
- [4] V. Venkatasubramanian, R. Rengaswamy, K. Yin and S.N. Kavuri, A Review of Process Fault Detection and Diagnosis: Part I: Quantitative Model-based Methods, *Computers and Chemical Engineering*, vol. 27, pp. 293-311, 2003.
- [5] V. Venkatasubramanian, R. Rengaswamy, K. Yin and S.N. Kavuri, A Review of Process Fault Detection and Diagnosis: Part II: Qualitative Models and Search Strategies, *Computers and Chemical Engineering*, vol. 27, pp. 313-326, 2003.
- [6] V. Venkatasubramanian, R. Rengaswamy, K. Yin and S.N. Kavuri, A Review of Process Fault Detection and Diagnosis: Part III: Process History Based Methods, *Computers and Chemical Engineering*, vol. 27, pp. 327-346, 2003.
- [7] S. Simani, C. Fantuzzi, and R.J. Patton, *Model-based Fault Diagnosis in Dynamic Systems Using Identification Techniques*, Berlin: Springer, 2003.
- [8] F. Caccavale and L. Villani, "An Adaptive Observer for Fault Diagnosis in Nonlinear Discrete-time Systems," in *Proc. of the 2004 American Control Conference*, Boston, Massachusetts, pp. 2463-2468, 2004.
- [9] T. Marcu, L. Mirea and P.M. Frank, "Neural Observer Schemes for Robust Detection and Isolation of Process Faults," *UKACC International Conference on CONTROL '98*, pp. 958-963, 1998.
- [10] A. Alessandri and T. Parisini, "Neural State Estimators for Direct Model-based Fault Diagnosis," in *Proc. of the American Control Conference*, Philadelphia, Pennsylvania, pp. 2874-2878, 1998.
- [11] A. Alessandri, "Fault Diagnosis for Nonlinear Systems Using A Bank of Neural Estimators," *Computers in Industry*, vol. 52, pp. 271-289, 2003.
- [12] S. Naidu, E. Zafirou and T.J. Mcavoy, Use of Neural-Networks for Failure Detection in A Control System, *IEEE Control Syst. Mag.*, vol. 10, pp. 49-55, 1990.
- [13] T. Marcu, and L. Mirea, Robust Detection and Isolation of Process Faults Using Neural Networks, *IEEE Control Syst. Mag.*, vol. 17, pp. 72-79, 1997.
- [14] Y.M. Chen and M.L. Lee, "Neural Networks Based Scheme for System Failure Detection and Diagnosis," *Mathematics and Computers in Simulation*, vol. 58, pp. 101-109, 2002.
- [15] M.M. Polycarpou, "Stable Learning Scheme for Failure Detection and Accommodation," in *Proc. of IEEE International Symposium on Intelligent Control*, Columbus, Ohio, pp. 315-320, 1994.
- [16] M.M. Polycarpou and A.J. Helmicki, "Automated Fault Detection and Accommodation: A Learning Systems Approach," *IEEE Trans. Syst., Man, Cybern.*, vol. 25, pp. 1447-1458, 1995.
- [17] A.B. Trunov and M.M. Polycarpou, "Automated Fault Diagnosis in Nonlinear Multivariable Systems Using A Learning Methodology," *IEEE Trans. Neural Networks*, vol. 11, pp. 91-101, 2000.
- [18] J.A.K. Suykens, J. Vandewalle and B.D. Moor, *Artificial Neural Networks for Modelling and Control of Nonlinear Systems*, Boston: Kluwer Academic Publishers, 1996.
- [19] J.A.K. Suykens, B.D. Moor, and J. Vandewalle, "Robust Local Stability of Multilayer Recurrent Neural Networks," *IEEE Trans. Neural Networks*, vol. 11, pp. 222-229, 2000.
- [20] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification*, Massachusetts: MIT Press, 1983.
- [21] H. Bang and Hyung Don Choi, "Attitude Control of A Bias Momentum Satellite Using Moment of Inertia," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 38, pp. 243-250, 2002.
- [22] C.D. Yang and Y.P. Sun, "Mixed H_2/H_∞ State-feedback Design for Microsatellite Attitude Control," *Control Engineering Practice*, vol. 10, pp. 951-970, 2002.



Qing Wu received the Bachelor's degree in automatic control engineering from the Huazhong University of Science and Technology (HUST), Wuhan, Hubei, P.R. China, in 1999 and the Master's degree in control theory and control engineering from HUST, Wuhan, Hubei, P.R. China, in 2002.

He is currently a Ph.D. candidate with the School of Engineering Science, Simon Fraser University, Vancouver, BC, Canada. His current research interests include fault detection and diagnosis, fault tolerant control, neural networks, learning control.

Self-Motion Graph in Path Planning with End-effector Path Specified

Zhenwang Yao

School of Engineering Science
 Simon Fraser University
 Burnaby, B.C, Canada, V5A 1S6
 Email: zyao@cs.sfu.ca

Abstract—Redundant robots have good dexterity in avoiding obstacles when performing tasks. In this paper, an enhanced probabilistic method is proposed to solve the problem of path planning along a specified end-effector path. Introducing a data structure named *Self-Motion Graph*, we allow robot self-motion along the end-effect path. Computer simulations show that this enhancement improves performance in some simulated cases.

Index Terms—Path planning, end-effector path, Self-motion.

I. INTRODUCTION

REDUNDANT manipulators have good dexterity and flexibility in avoiding obstacles when performing tasks, and an extensive literature discusses redundant robots in various motion planning problems. In this paper, we focus on the path planning problem with the end-effector path specified, also known as the *tracking problem* [1], which is formulated as:

Given an end-effector path in workspace, $x_g(t), t \in [0, T]$, and a start robot configuration q_0 , determine an entire joint space path $q(t)$, such that $F(q(t)) = x_g(t), t \in [0, T]$, where $F(q)$ is the robot forward kinematics function.

The existing approaches used to solve this problem can be classified into two types. One is on-line control, which is used for real-time control, and the other is off-line planning, which computes a feasible joint path before execution. Most Jacobian-based approaches fall in the former category. A probabilistic approach is also proposed, which fall into the latter category. This paper proposes an enhanced probabilistic method by taking advantage of robot self-motion.

Jacobian-based approaches work on instantaneous velocity level. The end-effector velocity, \dot{x} , can be represented in term of joint velocities, \dot{q} : $\dot{x} = J\dot{q}$, where J is the Jacobian matrix. Consider J is not a square matrix for redundant robots,

$$\dot{q} = J^\dagger \dot{x} + (I - J^\dagger J)z \quad (1)$$

where I is the identity matrix and $J^\dagger = J^T(JJ^T)^{-1}$. The first term is the minimum norm solution for $\dot{x} = J\dot{q}$, and the second term is the homogeneous solution. z is an arbitrary vector, and $(I - J^\dagger J)z$ is its projection in the null-space of J , also known as *self-motion*, meaning the movement caused by $(I - J^\dagger J)z$ does not affect the end-effector pose (i.e., end-effector position and orientation). Self-motion can be used to satisfy additional constraints, including obstacle avoidance [2], joint limit avoidance [3], and singularity avoidance [4]. Our

proposed method also uses self-motion to avoid obstacles, but instead of using an explicit form for self-motion (as in above-mentioned methods), we explore self-motion in a probabilistic fashion.

The previous approaches are local approaches. In some environments especially where obstacles clutter, the robot may get struck at some local minimum. In [5], a global approach for this problem was proposed. Therein, the problem is transferred to a finite-time non-linear control problem which can be solved by the Newton type of iterative substitution algorithms. However, every sampling point along the path needs to be considered in every iteration, which means the dimension of the transformed problem is very high, and consequently, the computation is expensive. The other drawback is that it requires all constraints to be expressed in explicit forms, which is difficult in most complicated environments.

As more probabilistic methods are used in path planning problems, especially for those high dimensional problems, Oriolo et al. [6] extends probabilistic mechanism to solve the tracking problem. The extended probabilistic method establishes connections among generated configurations for the sequential sampling poses, and to randomly generate configurations for a given pose, the techniques for close-chain robots [7], [8] are adopted. Several planning strategies (planners) are proposed therein, including *Greedy*, *RRT-Like* which uses the concept of RRT (Rapidly-exploring Random Tree) [9], and *RRT-Connect-Like* which uses the concept of RRT-Connect [10]. All the planners proposed perform well in practice, however not allowing self-motion along the given path, the planners only search paths in a relatively small area around the start configuration, which is sometimes inefficient. More details are discussed in Section II.

Based on [6], we propose an enhancement on current probabilistic planners for the problem by allowing robot self-motion, and demonstrate that for some applications, this enhancement help us solve the problems more efficiently.

The paper is organized as follows. In Section II, we first review the current probabilistic method, and propose our enhancement. In Section III, implementation and some consideration of the enhancement is discussed. To demonstrate the performance of the enhancement, computer simulations are presented in Section IV, followed by our conclusions.

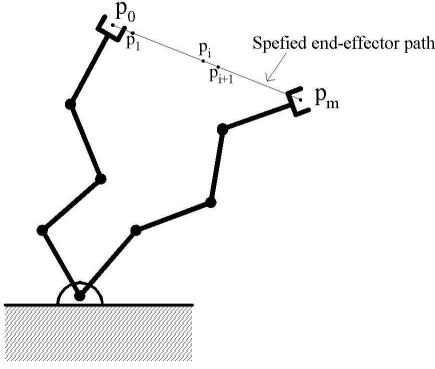


Fig. 1. Sampling along the specified end-effector path.

II. CURRENT PROBABILISTIC APPROACH AND PROPOSED IMPROVEMENT

As mentioned above, current probabilistic approach establishes connections among generated configurations for the sampling poses. The given end-effector path is discretized into a sequence of poses for the end-effector, $\{p_0, \dots, p_m\}$, as shown in Fig. 1. Then, the goal of the planners is to find a collision-free joint path, which is a sequence of configurations, $\{q_0, \dots, q_m\}$, such that $p_i = F(q_i)$, for all $0 \leq i \leq m$.

A. Configuration Generation

To generate configurations for a given end-effector pose, the configuration generation methods for close-chain robots [7], [8] are employed. Clearly, when the end-effector of an open-chain robot is fixed to a certain pose, generating a configuration for the robot is equivalent to generating a configuration for a closed-chain robot. The active-passive link decomposition [8] and the random loop generator [7] are efficient methods to generate configurations for closed-chain robots. The idea of the methods is to break the kinematics chain into the active sub-chain, whose joint variables (active configuration) are generated randomly, and the passive sub-chain, whose joint variables (passive configuration) are computed via closed-form inverse kinematics.

B. Current Probabilistic Planners

In [6], different probabilistic planners are proposed, in which *Greedy* and *RRT-Like* are the basic ones. All the others are combinations of these two using certain sets of parameters.

1) *Greedy Planner*: The *Greedy* planner uses a depth-first search strategy. Starting from the start configuration q_0 , it consecutively generates a configuration for each pose in the vicinity of its predecessor. For example, q_{i+1} for p_{i+1} is generated in the vicinity of q_i corresponding to pose p_i ; if it succeeds, then q_{i+2} for p_{i+2} is generated in the same way; if it fails to extend further after retries, then it restarts from the very beginning, pose p_0 . Thus, only one configuration is generated for every pose, and the data structure storing the explored path is a singly-linked list, as shown in Fig. 2.

The *Greedy* planner performs well when the environment is simple and the path is short (i.e., m is small). When the

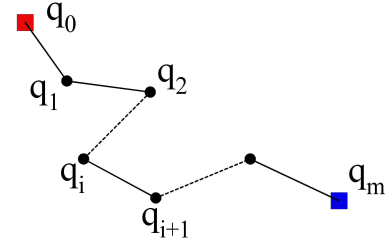


Fig. 2. Data structure for the *Greedy* planner: list.

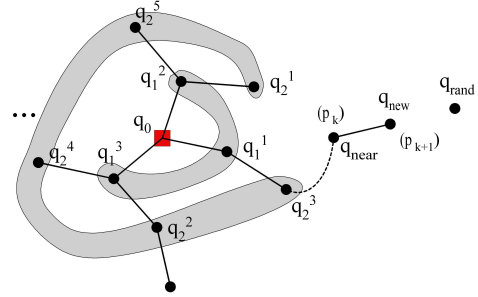
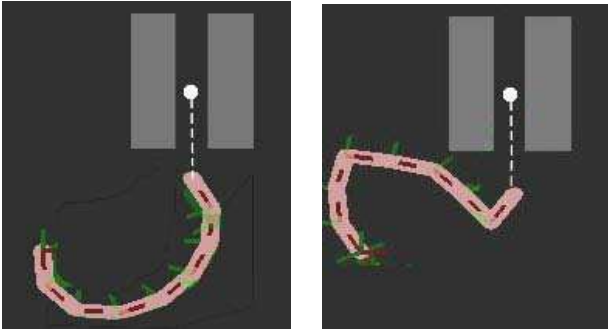


Fig. 3. Data structure for the *RRT-Like* planner: random tree.

number of poses increases and the environment becomes more complicated, the planner is easy to fail due to the depth-first search characteristic. For example, since the exploration for q_{i+1} (corresponding to p_{i+1}) is based on q_i (for p_i), if q_i is a “bad” configuration, it may be difficult to generate a feasible configuration for pose p_{i+1} , in the vicinity of q_i .

2) *RRT-Like Planner*: To avoid the depth-first search limitation in the *Greedy* planner, the *RRT-Like* planner generates more than one configuration for every pose. It applies the *RRT* [9] strategy upon active configurations, and passive configurations are determined according to the corresponding active configurations and the tree hierarchy. As shown in Fig. 3, each level of the tree corresponds to a particular pose in the end-effector path. For instance, the root of the tree is a configuration q_0 for the first pose, p_0 , the configurations in the second level of the tree, $\{q_1^1, q_1^2, \dots\}$, are configurations for pose p_1 , and $\{q_i^1, q_i^2, \dots\}$ are configurations for pose p_i . With a randomly chosen configuration, q_{rand} , the tree is extended from the closest node, q_{near} to a new configuration, q_{new} . Its active configuration q_{new}^a and passive configuration q_{new}^p are computed separately. First, as a typical RRT extension, q_{new}^a is computed by a linear displacement from q_{near}^a toward q_{rand}^a , where q_{near}^a and q_{rand}^a are active configuration of q_{near} and q_{rand} . Then, q_{new}^p is computed with q_{new}^a and its pose is determined from the tree hierarchy; for instance, if $F(q_{near}) = p_k$, then $F(q_{new}) = p_{k+1}$. If the height of the tree reaches the number of sample poses, m , planning succeeds.

The random tree grows in many directions, therefore prevents the robot from getting stuck at a bad configuration. On the other hand, *RRT-Like* has a drawback: as it computes the configurations too randomly, it takes a relatively long time to get to the goal, especially when the number of samples is large.



(a)Start with a “good” conf. (b)Start with a “bad” conf.

Fig. 4. Using a bad start configuration may lead to failure to find a path.

TABLE I

A COMPARISON OF USING DIFFERENT START CONFIGURATION

	Given start Conf. in (a.) (Success/Failure)	Given start conf. in (b.) (Success/Failure)
Greedy	8/12	0/20
RRT-C	15/5	0/20
RRT-G-C	19/1	0/20

3) *Combined Planners*: The limitations of *RRT-Like* and *Greedy* are overcome by combining these two planners. The proposed combinations from [6] include *RRT-Connect-Like*, *RRT-Greedy*, and *RRT-Greedy+Connect*. These planners interleave *Greedy* and *RRT-Like* with specified parameters. For example, the *RRT-Connect-Like* planner combines them in the following way:

- a. First call the *RRT-Like* planner to make an extension in a random direction.
- b. Second, if a feasible configuration is generated for pose, p_k , call the *Greedy* planner to explore the sub-path from p_k to the goal, p_m .
- c. Terminate if the goal is achieved, and planning succeeds; otherwise, repeat from step (a) until the maximum number of retries is exceeded.

C. *Observations*

With current planners, we observed that the start configuration significantly affects the result of planning. The following experiment demonstrates this observation. As shown in Fig. 4, under the same parameter, including sampling discretization, iteration/retry, we try to find a path for both scene (a) and (b). The only difference between the scenes is that they start with different configurations for the first pose. The experimental result (for 20 runs) in Table I demonstrates that a “bad” configuration may result in failure to find a path.

D. *Proposed Improvement*

The above observation can be explained as follows. Since all the planners proposed in [6] assume that self-motion is not allowed, when the given end-effector path is sampled into m poses, the final joint path consists of exactly m configurations. At the same time, to let the movement between successive configurations satisfy the end-effector path constraint, we must

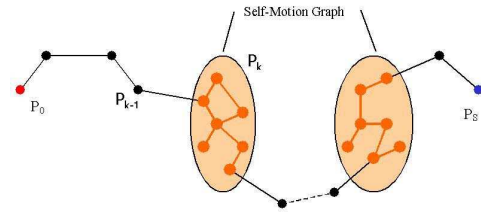


Fig. 5. Expansion of tree nodes for the *Greedy* planner to include self-motion graph.

ensure the difference between successive configurations is small enough, such that moving linearly in the configuration space approximately results in a linear movement in the task space. With these requirements, it may be difficult sometimes for a robot to move out of a bad configuration by limited number of small movements.

This limitation originated our enhancement on current planners: to allow self-motion along the end-effector path. In the enhanced planners, a set of configurations may be explored for a sampled pose, and with these configurations, the robot can make self-motion to avoid obstacles and move out of a bad configuration without changing the end-effector pose. In the next section, details of adaption on current planners and related data structure for enabling self-motion are presented.

III. IMPLEMENTATION WITH SELF-MOTION

A. *Data Structure*

The current planners uses a tree-type data structure to store the path, and each level of the tree corresponds to a pose along the end-effector path. Self-motion is incorporated into the planner by further expanding a node in the tree into an equivalent group of nodes corresponding to the same pose. As shown in Fig. 5, this group of configurations for a pose is represented by a connected graph, called the *Self-Motion Graph (SMG)* in short hereafter. All configurations in $SMG(p_k)$ are connectable to some configuration for the last pose p_{k-1} , and once a configuration in $SMG(p_k)$ is connected to a configuration for pose p_{k+1} , the path is successfully extended to the next pose p_{k+1} .

B. *Self-Motion Graph Exploration*

As we know how to generate configurations for an end-effector pose, the exploration of *SMGs* is straightforward. An *SMG* is explored by repeatedly calling *Explore_SMG()*. As shown in Fig. 6, *Explore_SMG(k)* returns a connectable collision-free configuration q_{smg} for p_k . It applies an *RRT* like strategy upon active configurations. *Random_Config()* generates a random active configuration, q_r^a ; *Nearest_Node()* finds the closest configuration q_c in the current *SMG* (w.r.t active configurations); *Extend_Config()* displaces a new active configuration from q_c^a toward q_r^a with a fixed distance; *ClosedForm_InvKin()* computes the passive configuration. To ensure the end-effector does not move between two configurations, q_{smg} should be close to q_c in every dimension, which is guaranteed by *Extend_Config()* for active configurations and by *ClosedForm_InvKin()* by passive configurations.

```

Function: Explore_SMG( $k$ )
IF ( $SMG(p_k).conf\_num > MAX\_SMG\_NODE$ ) THEN
    RETURN failure;
 $retry \leftarrow 0$ ;
DO
     $q_r^a \leftarrow Random\_Config()$ ;
     $q_c^a \leftarrow SMG(p_k).Nearest\_Node(q_r^a)$ ;
     $q_{smg}^a \leftarrow Extend\_Config(q_c^a, q_r^a)$ ;
     $q_{smg}^p \leftarrow Closedform\_InvKin(p_k, q_{smg}^a, q_c^p)$ ;
     $q_{smg} \leftarrow (q_{smg}^a, q_{smg}^p)$ 
    IF  $Is\_Collision\_Free(q_{smg}, q_c)$  THEN
         $SMG(p_k).Add\_Edge(q_c, p_{smg})$ ;
        RETURN  $q_{smg}$ ;
    ELSE
         $retry \leftarrow retry + 1$ ;
WHILE( $retry \leq MAX\_SMG\_RETRY$ );
RETURN failure;
    
```

Fig. 6. Pseudo-code for SMG exploration.

$Explore_SMG()$ may return *failure* if (1) it can not find a connectable collision-free configuration after a fixed number of retries; or (2) the number of configurations in the SMG reaches a predefined limit. In our implementation, the data structure for an SMG is a tree. The root node of $SMG(p_k)$ is the configuration connecting a configuration for p_{k-1} .

Do we have to explore the SMG for every pose? To reduce the computation to build and search the explored tree, we only create the SMG when required. For example, the *Greedy* planner terminates the exploration from q_k (for p_k) to q_{k+1} (for p_{k+1}) after a certain number of retries. To enable self-motion, we do not give up at this point, and instead, an SMG is propagated for pose p_k . For the poses where the planner can extend to the next pose within a fixed number of retries, simple nodes are retained.

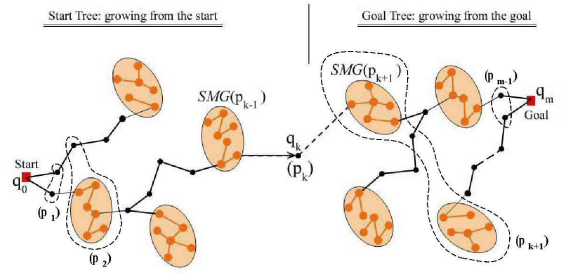
C. SMG in Greedy Planner

Knowing how to explore the SMG , we can now integrate the self-motion enhancement into current planners. Integrating $SMGs$ into the *Greedy* planner is straightforward, and the idea is as follows.

- 1) An SMG is created for pose p_k if the planner fails to extend the path from q_i (for p_i) to q_{i+1} (for p_{i+1}) after a fixed number of retries.
- 2) With the SMG , every time a new configuration q_{smg} is explored in $SMG(p_i)$, the planner tries to extend the path to the next pose, p_{i+1} . Once a valid configuration q_{i+1} is obtained, the planner stops exploring the SMG .
- 3) Otherwise, it repeatedly explores the SMG for p_i , until the number of retries or the number of configurations in $SMG(p_i)$ exceeds the predefined limit.

D. SMG in RRT-Connect-Like Planner

For many basic motion planning problems, the *RRT-Connect* planner [10] performs better than the *RRT* planner [9]. *RRT-Connect* has several differences compared to *RRT*: (a) it has greedy heuristic to the goal, and (b) it has two random trees, one grows from the start configuration, and the other grows from the goal configuration. For the end-effector path tracking


 Fig. 7. Random trees with $SMGs$. q_k for p_k is extended from the start tree, then the nodes (including SMG) for pose p_{k+1} is checked for connectivity.

problem, the *RRT-Connect-Like* [6] planner only adopts the greedy heuristic in (a). With $SMGs$, we implement a more sophisticated planner. The idea is as follows.

1. Two random trees are grown from the start configuration and the goal configuration, respectively. In each tree, configurations and $SMGs$ on the same level correspond to the same pose. The goal configuration for the goal pose is chosen randomly or is specified by the users.
2. After every RRT extension, it obtains a new configuration which corresponds to a certain pose. It greedily extends the path from the newly-extended pose to the goal or start pose, depending on which tree is being extended. For a pose where the planner fails to extend further after a maximum number of retries, an SMG is created for the pose, and the planner tries to explore this SMG , and make connection to next pose. As shown in Fig. 7, nodes in new random-tree can be an SMG , and basically the leaves are $SMGs$, because that is where the planner fails to extend further.
3. As shown in Fig. 7, assume from the start tree, a new configuration for p_k is explored in the previous greedy extension, it checks whether there exists a connectable $SMG(p_{k+1})$ in the random tree growing from the other direction; if so, then make the connection, and succeed if connected.
4. If it can not reach the last pose after exploring the $SMGs$, then manipulate the other three and repeat from step 2.

IV. COMPUTER SIMULATIONS

To show performance of the enhancement, we created various scenes to compare the original planners and the enhanced ones. The simulations are carried on a Pentium-4 2.0G PC, with our in-house developed software library MPK, the Motion Planning Kernel [11].

A. Planners and Objectives

In these simulations, six planners are compared: three planners are proposed in [6], the other three are enhanced with SMG .

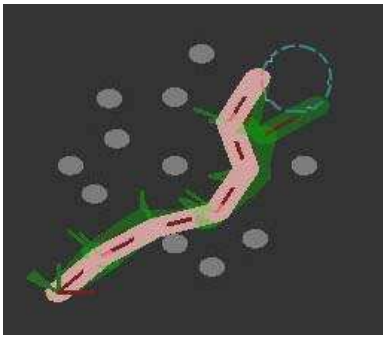


Fig. 8. Case 1.

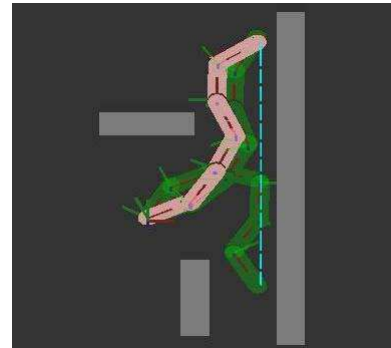


Fig. 9. Case 2.

 TABLE II
 PERFORMANCE OF DIFFERENT PLANNERS IN CASE I

Planner	Time(s)	C-D-P	C-D-L
Greedy	19.4	24457	2121
RRT-C	3.8	6038	451
RRT-G-C	5.7	8659	731
SMG-Greedy	2.5	3528	373
SMG-RRT-C	3.7	5462	614
SMG-RRT-C2	3.0	3852	347

Greedy	Greedy planner in [6]
RRT-C	RRT-Connect-Like Planner in [6]
RRT-G-C	RRT-Greedy+Connect Planner in [6]
SMG-Greedy	Greedy planner with <i>SMG</i>
SMG-RRT-C	RRT-Connect-Like Planner with <i>SMG</i> , but only has one tree growing from the start configuration.
SMG-RRT-C2	RRT-Connect-Like Planner with <i>SMG</i> , and with two random trees from the start and the goal respectively.

In each scene, the result is based on the average performance over 20 runs, and the following performance was measured for each planner.

<i>Time(s)</i>	Planning time, second as unit.
<i>C-D-P</i>	Collision Detection for Point. Number of collision checks for a single configuration.
<i>C-D-L</i>	Collision Detection for Linear Connection. Number of collision checks for a connection between two configurations.

Note that, to do the collision check between two configurations, the linear connection between two configurations is sampled and every intermediate configuration is checked for collision. In other words, one collision check for connection (C-D-L) results in multiple checks for point (C-D-P). However, other than the checks introduced by C-D-L, C-D-P also includes collision checks for randomly generated configurations.

B. Cases and Results

Three cases are reported to demonstrate the performance of different planners.

1) *Case 1*: As shown in Fig. 8, the robot moves the end-effector along a circle in an environment with obstacles. The result is shown in Table II.

 TABLE III
 PERFORMANCE OF DIFFERENT PLANNERS IN CASE 2

Planner	Time(s)	C-D-P	C-D-L
Greedy	43.5	99366	4080
RRT-C	43.6	80573	4129
RRT-G-C	11.6	21380	1086
SMG-Greedy	8.7	16007	1288
SMG-RRT-C	9.4	15330	1509
SMG-RRT-C2	11.1	11210	990

2) *Case 2*: As shown in Fig. 9, the robot moves the end-effector along the surface of one of obstacles. The result is shown in Table III.

3) *Case 3*: As shown in Fig. 10, the robot moves the end-effector into a narrow passage. The result is shown in Table IV.

C. Remarks

With *SMGs*, the enhanced planners have better performances in these cases, they normally have shorter planning time and involve less collision checks.

In these three cases, *SMG-Greedy* always achieves the best running time. Similarly to *RRT-Connect-Like*, *SMG-Greedy* holds the breadth-first search characteristic in the sense that they generate more than one configuration for a pose; however, unlike *RRT-Connect-Like*, by using *Self-Motion Graph*, *SMG-Greedy* only utilizes the breadth-first search when it has to.

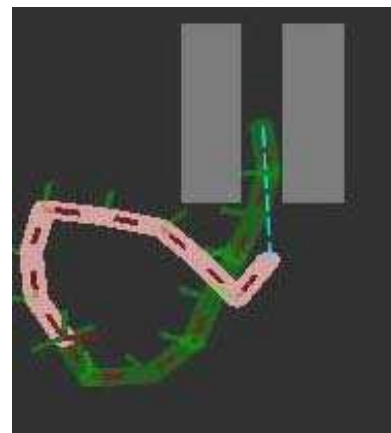


Fig. 10. Case 3.

TABLE IV
PERFORMANCE OF DIFFERENT PLANNERS IN CASE 3

Planner	Time(s)	C-D-P	C-D-L
Greedy	29.9	69552	3032
RRT-C	18.1	35818	1918
RRT-G-C	7.6	14242	786
SMG-Greedy	6.2	11061	1126
SMG-RRT-C	6.8	10812	1355
SMG-RRT-C2	13.0	10818	1237

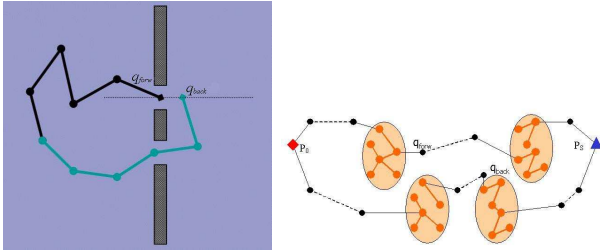


Fig. 11. A special case: not connectable q_{forw} and $SMG(q_{back})$.

In the second and third case, although *SMG-RRT-C2* involves less collision checks, the running time is longer than *SMG-RRT-C*. The reasons are (1) *SMG-RRT-C2* needs to check the connectivity between two trees, and (2) *SMG-RRT-C2* needs to generate a configuration for the last pose to grow the tree from the goal, which is hard in these two cases (this generation can be avoided if the goal configuration is specified by users.)

When using the *SMG-RRT-C2* planner, we always check the connectivity between two random-tree growing from different directions. This connection checking takes relatively large portion of computation, and in some cases, these two might not be connectable at all. For instance, as shown in Fig. 11, although pose p_{back} and p_{forw} are consecutive poses, the node q_{forw} is not connectable to $SMG(p_{back})$.

SMGs help in applications where time requirement along the end-effector path is loose, like inspection robots. On the other hand, some applications may not benefit from this enhancement. For example, in spray painting applications, a constant end-effector velocity is required along the end-effector path, and self-motion generates an inconstant end-effector velocity (thereby uneven paint deposition may result).

V. CONCLUSIONS

In this paper, by introducing the *Self-Motion Graph*, an enhancement is proposed based on current probabilistic planners for path planning problem with end-effector path specified. To explore robot self-motion, we adopt configuration generation techniques for closed-chain robots. Computer simulations prove that this enhancement indeed improves the performance, in terms of finding a collision-free path.

REFERENCES

[1] J. Ahuactzin and K. Gupta. The kinematic roadmap: A motion planning based global approach for inverse kinematics of redundant robots. *IEEE Transactions on Robotics and Automation*, 15:653–669, 1999.

[2] A. Maciejewski and C. Klein. Obstacle avoidance for kinematically redundant manipulators in dynamically varying environments. *International Journal of Robotics Research*, 4:109–117, 1985.

[3] Z. Guo and T. Hsia. Joint trajectory generation for redundant robots in an environment with obstacles. In *IEEE International Conference on Robotics and Automation*, pages 157–162, 1990.

[4] T. Yoshikawa. Manipulability of robotic mechanisms. *International Journal of Robotics Research*, 4:3–9, 1985.

[5] S. Seereeram and J.T. Wen. A global approach to path planning for redundant manipulators. *IEEE Transactions on Robotics and Automation*, 11:152–160, 1995.

[6] G. Oriolo, M. Ottavi, and M. Vendittelli. Probabilistic motion planning for redundant robots along given end-effector paths. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1657–1662, 2002.

[7] J. Cortes, T. Simeon, and J.P. Laumond. A random loop generator for planning the motions of closed kinematic chains with PRM methods. In *IEEE International Conference on Robotics and Automation*, pages 2141–2146, 2002.

[8] L. Han and N. Amato. A kinematics-based probabilistic roadmap method for closed kinematic chains. In B. Donald, K. Lynch, and D. Rus, editors, *Workshop on Algorithmic Foundations of Robotics*, pages 233–246, March 2000.

[9] S. M. LaValle. Rapidly-exploring random trees: A new tool for path planning. Technical Report TR 98-11, Computer Science Dept. Iowa State University, Oct. 1998.

[10] J. Kuffner and S. LaValle. RRT-connect: An efficient approach to single-query path planning. In *IEEE International Conference on Robotics and Automation*, pages 995–1001, 2000.

[11] I. Gipson, K. Gupta, and M. Greenspan. MPK: An open extensible motion planning kernel. *Journal of Robotic Systems*, 18(8):433–443, Aug. 2001.



Zhenwang Yao received the B.S. degree from in Electrical Engineering from University of Science and Technology of China in 1998, and the M.A.Sc. degree in Electrical Engineering from Simon Fraser University, Canada, in 2005. He is currently working toward the Ph.D. degree. His research interests include path planning, distributed robotics, and mobile sensor networks.