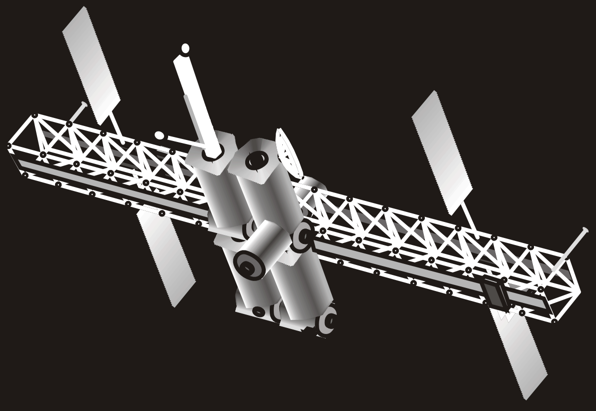
© 2006
ENSC 894
Simon Fraser University
School of Engineering Science

2006 Course Proceedings



Graduate Communication:

Applying Strategies of Writing in Technical Journals

Exploring depth perception in virtual reality environments

Caitlin Akai
School of Interactive Arts & Technology
Simon Fraser University
Vancouver, British Columbia
Email: cakai@sfu.ca

***Abstract*— Virtual reality environments (VREs) are becoming increasingly common in a wide range of applications including training and simulation, physical rehabilitation, and industrial design. The use of virtual reality displays in the industrial design of automobiles and airplanes is gaining popularity due to the potential to reduce laborious and time intensive design processes requiring high levels of detail and realism. Unfortunately, some users perceive distortions in size and distance while using virtual environments for design tasks, and the cause of these perception errors is an active area of research. This paper discusses some of the current research into depth perception, examines the methods being used to study depth perception in virtual environments, and speculates on potential future work.**

I. Introduction

Over the last fifteen years, virtual reality environments have begun to move out of the realm of science fiction and into real world applications and research. Current applications of virtual reality include: health-related issues such as rehabilitation and phobia reduction , military training and simulation, and perception research [1]–[3]. Virtual reality (VR) is also commonly used in industrial design processes to replace laborious and time-intensive physical models [4]. Virtual reality can allow for faster and more frequent design iterations, reducing the total time required for the design process and allowing for quicker ẗime to marketf̈or new vehicles. While many companies have invested in VR technology, virtual reality displays have yet to reach their full potential because some users perceive distortions in depth and distance [5]. Distance estimates in virtual environments tend to be contracted with closer distances being overestimated while farther distances are underestimated. Design tasks require very high levels of detail and realism, and even small distortions can affect a designer's ability to evaluate a design. Therefore, finding solutions to the perceptual distortions caused by virtual reality displays is critical to industries that have yet to receive much return on their investments.

Many researchers have examined depth perception in virtual environments, but there is no simple answer to the cause of the perception problems. This paper describes the types of virtual displays used in typical industrial processes, discusses the perceptual problems triggered by virtual displays, examines the methods used to study depth perception in virtual environments, and speculates on directions for future work.

II. What Is Virtual Reality?

Virtual reality environments are displays that allow viewers to perceive a three-dimensional (3D) image from two-dimensional (2D) images. The displays take advantage of the slight separation of the human eyes (interpupillary distance), which provides the cue of binocular disparity. Binocular disparity is the different view of the world seen by each eye. In virtual environments, binocular disparity is recreated by projecting two stereo images taken from slightly different perspectives. The images are set laterally apart on the screen and must be viewed with special glasses. Once these images are projected onto the viewer's retina, they are recombined in the visual cortex of the brain, and a three-dimensional image is perceived despite the projected image being only two-dimensional. The process used by the visual cortex to combine the images (known as the correspondence problem) is still poorly understood, though pattern matching likely plays a role [6].

Industrial design applications commonly use two types of virtual reality displays: displays viewed with stereo glasses and head-mounted displays. Displays viewed with glasses include small screen displays like FishTank VR and large screen projection displays composed of one or more screens [7]. Multi-screen displays are often arranged in a u-shape to form a Cave Automatic Virtual Environment (CAVE) [8]. Using a CAVE configuration provides a more immersive experience for the viewer, because the screens are large enough to fill the viewer's entire field of view. Screen displays can be either passive or active. Passive displays use analyglyph 3D, in which the stereo images are projected in different colours and require glasses with different coloured lenses (e.g., most 3D movies use red and blue lenses), or are based on circular polarization, which allows certain wavelengths of light to enter each eye when polarizing glasses are worn. To achieve a sense of depth, both systems ensure that each eye views a slightly different image. Active stereo displays use shutterglasses that automatically flicker on and off in sync with the projected image. Most displays viewed with glasses also track the viewer's head position so they receive the correct image for their viewpoint. Thus, only a single viewer is provided with the correct viewpoint, while others will see slight distortions of the scene.

The second type of virtual display commonly used is the head-mounted display (HMD), which provides an extremely immersive 3D experience. Head mounted displays are worn as helmets with glasses, with the stereo images projected directly onto the lenses of the glasses. HMD's are similar to the active stereo displays in that they project synchronous stereo images. Each HMD is worn by a single viewer and usually occludes all vision of the outside world, providing for a highly immersive experience. Since HMD's do not allow for multiple simultaneous viewers, they are less likely to be used as a collaborative tool. Thus, large-screen active stereo displays are more commonly used for industrial design applications.

III. Perceptual Problems in Virtual Reality

While virtual reality can provide a compelling sense of three-dimensionality from a two-dimensional image, some users are not able to perceive depth in the displays. Approximately 5-10% of people are stereoblind and unable to see 3D in VR because they cannot use binocular disparity as a depth cue [9]. In the real world, there are many monocular cues which stereoblind people can rely on to provide a sense of depth. The virtual world also uses monocular cues but places a stronger emphasis on binocular disparity as a cue to depth. However, for people who are not stereoblind virtual reality can also beproblematic causing some to see distortions in depth and size in virtual environments [5], [10].

Another significant problem affecting users of virtual environments is the high incidence of eyestrain and cybersickness (i.e., nausea and dizziness caused by exposure to VR) [10], [11]. These side-effects severely limit the amount of time that users can comfortably spend in virtual environments. Research on perceptual adaptation has found that subjects can adapt to virtual environments over time, reducing the severity of cybersickness symptoms with increased exposure [12].

Compounding the problems in depth perception in VR are individual differences. Innate differences in individual perception seem to cause perceptual problems in virtual environments but are relatively insignificant when perceiving depth in the real world. Individual differences manifest themselves in varying perceptions of size, distance and depth of objects in 3D virtual environments.

IV. Research on Depth Cues

Researchers have used virtual reality as a tool to learn about depth perception in general, and to investigate how depth perception in virtual environments differs from the real world. In particular, virtual environments are used to understand distance and size perception under various conditions. To provide a sense of depth, virtual reality environments rely on many of the same depth cues available in the real world. Depth cues can be either binocular (i.e., cues that can only be viewed with two eyes) or monocular (i.e., cues that can be viewed with one eye). Current research on depth perception in virtual environments investigates both the influence of individual depth cues and the combination of multiple depth cues. Other factors affecting depth perception in virtual reality displays include proprioceptive feedback from the ocular muscles (extraretinal inflow), and hardware properties such as blur.

A. Binocular Depth Cues

The two major binocular cues to depth are binocular disparity and the vergence position of the eyes. Binocular disparity (or stereopsis) provides a perception of relative depth from the disparities in the images seen from each eye. Disparity varies with the distance between objects being viewed. Objects that are closer together will have a smaller disparity than those that are farther away. Humans are extremely sensitive to differences in binocular disparity, and have an average disparity threshold of 5 arcsec, a difference of 0.1mm at arm's length [6]. The advantages of binocular disparity include: improved visual detection, resolution and discrimination [13]. Binocular vision has also been shown to provide a more accurate perception of distance than monocular vision [14].

The second binocular cue to depth is vergence, the inward and outward movement of the two eyes. To focus on an object, our eyes move together so the image is projected onto the fovea, the most sensitive area of the eye's retina. To bring an image into focus we accommodate (focus) on it by adjusting the crystalline lens using the eye's ciliary muscles. Vergence provides a strong cue to depth because different distances require different amounts of vergence. Vergence is a reliable cue from 10 cm to 6 m but is unreliable at large fixation distances and can lead to a contraction bias in reduced cue conditions [15]. The weight of vergence as a depth cue decreases with distance and when retinal cues decrease [16].

Vergence and accommodation provide information in the form of proprioceptive feedback from the ocular eye muscles. Signals sent from the ocular muscles to the brain are known as extraretinal inflow. Sources of extraretinal inflow include muscular feedback and internal monitoring of the muscle position [17]. Extraretinal inflow from vergence is an important depth cue for distance perception, but can be perturbed by extending the eye muscles using an eccentric (angled) gaze [18]. Holding an eccentric gaze for 30 seconds causes errors in perceived visual direction as well as pointing and throwing [19].

B. Monocular Depth Cues

Virtual reality environments provide several monocular depth cues observable with a single eye. Monocular cues include motion, blur, occlusion, aerial and linear perspective, familiar size, relative height, texture, accommodation and shadow. The majority of VR research has focused on particular cues like motion, size, accommodation, and blur.

1) Motion: Motion parallax, the relative motion of different points on an object at different distances, is caused by movement of the object (kinetic depth) or movement of the viewer (motion perspective) [20], [21]. In virtual reality, motion perspective is available to users being head-tracked in stereo displays or those wearing HMD's. Kinetic depth cues are only available when the virtual scene is animated, but are important cues to three-dimensional shape and interact with stereo disparity during early depth processing [22]. Rogers & Graham found that motion parallax produced by observer or object movement provides a reliable and unambiguous perception of relative depth [23]. Though work by Beall et al., found that observer-produced absolute motion parallax was a weak determinant of distance and size perception of nearby objects in both the real and virtual environments [24].

2) Size: Size is an important factor in depth perception in virtual environments because according to Emmert's law perceived size is a function of perceived distance. This law also accounts for size constancy, i.e., perception of size remains constant although the size of the image projected on the retina varies as the object moves in distance. The ability to interpret moving objects as unchanging in size seems obvious for real world tasks, but in virtual worlds, where rules are less clear, size constancy may break down. In reduced-cue conditions, a size-distance paradox can be observed, causing viewers to perceive smaller closer objects as farther away than more distant targets [25]. Using a task that required both a verbal estimate of distance and pointing to a target, Mon-Williams & Tresilian found that verbal reports led to a response consistent with the size-distance paradox but pointing did not [26]. They concluded that the paradox was therefore a cognitive phenomenon.

One of the most well known size constancy studies is that of Holway & Boring, who examined size constancy in monocular viewing conditions and binocular viewing [27]. The task required subjects to adjust the size of a circle projected on a screen to match that of a circle set at a different distance. They found that the binocular condition resulted in a slight overestimation of target size, while monocular cues consistently resulted in underestimation.

Familiar size and relative size are two important size depth cues used in virtual reality. The size of familiar objects can provide an estimate of distance in uncertain situations, while relative size allows for size comparison between different objects and is reliable over a range of distances [21].

3) Accommodation: Accommodation is the eye's ability to focus by adjusting the crystalline lens with the ciliary muscles. By itself, accommodation is only effective for 2 metres or less and declines with age, though it can interact with other sources of information to provide a stronger perception of depth [25]. In virtual displays, accommodation is problematic because all images are projected onto the screen (i.e., on a single focal plane), but perceived at different depths requiring varying amounts of vergence. In the real world, our eyes converge on the object we accommodate on, creating a synkinnetic link to vergence and accommodation. In virtual reality, this link is broken, though its influence on depth perception is still unclear [28]. Akeley et. al have attempted to address this problem by creating a display with multiple focal distances so that the correct vergence and accommodation cues are available at several pre-determined distances [29]. While the initial work is exploratory, their approach may one day be applicable to head-mounted virtual displays, thereby reducing some perceptual problems.

4) Blur: Blur is a general cue to depth related to accommodation, because objects on the plane of fixation are in focus while those at other distances are blurred. Blur is a relatively unreliable depth cue, since its magnitude varies with pupil diameter and refractive state, as well as with depth [30]. Blur discrimination has a relatively large 'Just Noticeable Difference', making it a course measure of depth. However, blur can provide important ordinal depth information at borders of objects at extreme blur values [31]. In virtual displays when blur is combined with the depth cue of binocular disparity, disparity is the dominant cue [30]. The potential usefulness of blur cues are interesting to stereoscopic display researchers because they are absent in both HMD's and projection-based displays and are related to the accommodation/vergence conflict discussed previously [29].

C. Cue Combination

Examining how cues work together to produce our perception of depth is an active area of research. In both the real and virtual worlds, many simultaneous cues to depth are available, but understanding how different cues interact is a complicated problem. Researchers recognize that no single depth cue is dominant and that depth perception is more accurate when more cues are available [20], [24], [32]. Cues can be combined through summation (averaging), multiplication (interactions between cues), and selection (a single cue is used) [32]. Various combinations of cues are studied to determine how they interact. For example, research on vergence has found that if vergence conflicts with other cues or there is less vergence demand, less weight will be given to it perceptually [33]. A study by Bradshaw et al. found that differential perspective and vergence angle are additive when combined as cues for scaling depth from horizontal disparities [34]. Hillis et al. concluded that single cue information could be lost when cues from the same modality are combined, but not when different modalities are combined (e.g., haptics and vision) [35].

V. Research Methods In Depth Perception

In order to quantify depth perception, research typically focuses on tasks related to distance and size estimation. Experimental design for research on depth perception requires careful consideration on the cue to measure, the environment to conduct the study in, and the method of measurement (metric) used.

The most common approach to depth perception research is a psychophysical approach. Psychophysics measures absolute thresholds (i.e., the minimum stimulus intensities detectable) and difference thresholds or Just Noticeable difference (i.e., the smallest change of stimulus intensity required to be perceived as a change) [36]. To determine absolute and difference thresholds, a large number of trials are averaged because of individual fluctuations in sensitivity to stimuli.

Psychophysics requires isolating cues of interest and measuring subjects' responses to cues during specific tasks. However, the act of isolating a cue can change its effect, making it difficult to generalize results to more ecologically valid full-cue environments [6]. Accommodation is a weak cue in full cue conditions (since it gives course ordinal information) but is weighted more heavily in reduced-cue conditions [26].

Philbeck et al. found that distance perception in reduced cue conditions consistently showed systematic error, while perception was essentially accurate with full cues [37]. Loomis et al. found that in reduced cue conditions subjects overestimated target distances of less than 2 metres, but underestimated targets over 3 metres away [14].

Virtual reality displays are reduced cue environments since they almost always require a darkened environment. Some cues, like accommodation and vergence, are perceived differently under darker conditions than in full light. In a study that used a reaching task in a dark environment, Bingham et al. found that subjects under-reached to targets, while Johnston found that cylindrical objects viewed in dark were perceived as being expanded or compressed depending on viewing distance [38], [39].

A. *Metrics*

Deciding on the appropriate metric for investigating depth perception is a critical issue in research. The metric used to measure the subjects' response can add bias to the results, making it difficult to know if an effect was caused by the cue being measured or the measurement itself. When choosing a metric, the distance of interest is a key consideration. A common way of expressing distance in depth perception research is in terms of exocentric and egocentric space. Exocentric space is the distance between objects being seen by the viewer. Egocentric space is measured in relation to the observer. Cutting has divided egocentric space into 3 further regions – personal space (0-1.5 or 2 metres), action space (2-30 metres), and vista space (30 metres or more) [40]. Most tasks in depth research focus on egocentric distance, since exocentric distance estimates are more error prone [14]. Absolute distance information, as opposed to relative distance estimations, is required for egocentric distance [26].

Metrics for judging distance in personal space often use pointing tasks or related motor tasks [15], [41], [42]. The majority of VR depth research has explored the mid-range of action space using visually-directed action metrics, by allowing visual input before the task, but removing it once the task is underway [43]. Walking metrics are the most common form of visually-directed action metrics. A variety of walking metrics are used in depth research including: visually directed walking, triangulated walking and pointing, and triangulated walking, blindfolded walking, and walking on treadmills [14], [43], [44]. Other less common metrics include throwing and imagined walking [45], [46].

Metrics common to psychophysics that do not use walking include: two-alternative forced choice, method of adjustment, and verbal/written report. Tasks using two-alternative forced choice present subjects with two stimuli and requires them to chose the one most representative of the cue being studied (e.g., subjects might say which of the two are bigger or closer). The method of adjustment requires the subject to adjust the intensity of the stimulus and can be done with one or two stimuli. With two stimuli, one stimuli is a standard that the subject must match the other stimuli to by manipulating the variable of interest. If only a single stimulus is used, the subject must adjust the intensity of the stimulus until it is detectable. Tasks using verbal and written reports require subjects to make verbal or written estimates of the stimulus intensity.

VI. Future Directions

Although a tremendous amount of information on depth perception in virtual environments has been uncovered through research, the cause of perceptual distortions in virtual reality and possible solutions remain unknown. Future depth perception research in virtual reality environments will expand its current focus to include more research on complex conditions including multiple cues. While research conducted in complex environments can obscure the exact cause of effects, reduced cue conditions may not be applicable to the typical environment encountered by virtual reality users. Direct comparisons between real and virtual stimuli could be very informative about how virtual reality modifies the cues available for depth perception.

Future research will also focus on the interaction of visual and non-visual cues in virtual environments. Though in the real world we receive many cues to depth (with only a fraction of them being visual), most virtual reality systems rely solely on visual cues (and some proprioceptive cues). Research has already begun to look more closely at haptic and auditory cues as feedback. The use of tactile augmentation (using real objects for haptic feedback with virtual visual feedback) has been shown to increased the perception of weight and realism of objects viewed in VR [1]. Applications for haptic feedback in VR include distractions for burn victims during wound cleaning, and treatment of phobias like arachnophobia.

The role of adaptation in virtual environments to increase user comfort and reduce the incidence of cybersickness and eyestrain will also require further work [47]. Accommodation and vergence show signs of adaptation after time spent in virtual environments, which may lead to solutions for the accommodation-vergence conflict [48].

VII. Conclusion

The promise of virtual reality suggests tremendous possibilities for refining and improving the efficiency of industrial design processes. Yet, the perceptual distortions that plague these environments are a severe limitation to the widespread use of virtual reality. Finding solutions to these perceptual problems could revolutionize current industrial design processes and open up new applications for virtual reality. Research has considerably increased our understanding of depth perception both in and out of virtual environments, but more work is needed to find ways to address the effects of individual differences. Research has also found that cues are weighted differently individually than when combined, which suggests there is no simple answer to the depth perception problems plaguing virtual environments. Further work in the areas of cue combination, multimodal cues, and adaptation may bring us closer to solving the perceptual errors common to virtual displays.

References

[1] H. G. Hoffman, "Physically touching virtual objects using tactile augmentation enhances the realism of virtual environments," in *IEEE Virtual Reality Annual International Symposium*, 1998, pp. 59–63.

[2] R. Earnshaw, J. Vince, and H. Jones, Eds., *Virtual reality applications*. London: Academic Press, 1995.

[3] K. Ukai and Y. Kato, "The use of video refraction to measure the dynamic properties of the near triad in observers of a 3-d display," *Ophthalmic Physiological Optics*, vol. 22, pp. 385–388, 2002.

[4] F. Dai, Ed., *Virtual reality for industrial applications*. Berlin, Germany: Springer, 1998.

[5] L. Baitch and R. Smith, "Physiological correlates of spatial perceptual discordance in a virtual environment," in *Fourth International Immersive Projection Technology Workshop*, Iowa, jun 2000.

[6] J. M. Harris, "Binocular vision: Moving closer to reality," *Philosophical Transactions of the Royal Society of London A*, vol. 362, pp. 2721–2739, 2004.

[7] C. Ware, K. Arthur, and K. S. Booth, "Fishtank virtual reality," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Amsterdam, 1993, pp. 37–42.

[8] C. Cruz-Neira, D. Sandin, T. DeFanti, R. Kenyon, and J. Hart, "The cave: Audio visual experience automatic virtual environment," *Communications of the ACM*, vol. 35, no. 6, pp. 733–754, 1992.

[9] R. Blake and R. Sekular, *Perception*, 5th ed. McGraw-Hill, 2006.

[10] J. P. Wann and M. Mon-Williams, "Health issues with virtual reality displays: What we do know and what we don't," *Computer Graphics*, pp. 53–57, May 1997.

[11] K. Stanney, "Realizing the full potential of virtual reality: Human factors issues that could stand in the way," in *IEEE Proceedings of Virtual Reality Annual International Symposium*, 1995, pp. 28–34.

[12] E. Regan, "Some evidence of adaptation to immersion in virtual reality," *Displays*, vol. 16, no. 3, pp. 135–139, 1995.

[13] I. P. Howard, *Seeing in Depth Volume 1: Basic Mechanisms*. Thornhill, Ontario: I. Porteous, 2002.

[14] J. M. Loomis, J. A. D. Silva, J. Philbeck, and S. S. Fukusima, "Visual perception of location and distance," *Current Directions in Psychological Science*, vol. 5, pp. 72–77, 1996.

[15] M. Mon-Williams and H. C. Dijkerman, "The use of vergence information in the programming of prehension," *Experimental Brain Research*, vol. 128, pp. 578–582, 1999.

[16] J. Tresillian, M. Mon-Williams, and B. Kelly, "Increasing confidence in vergence as a distance cue," *Proceedings of the Royal Society of London B*, vol. 266, pp. 39–44, 1999.

[17] W. Shebilske, "Extraretinal information in corrective saccades and inflow vs outflow theories of visual direction constancy," *Vision Research*, vol. 16, pp. 621–628, 1976.

[18] M. Mon-Williams and J. Tresillian, "A framework for considering the role of afference and efference in the control and perception of ocular position," *Biological Cybernetics*, vol. 79, pp. 175–189, 1998.

[19] W. Shebilske, "Ecologicala efference mediation theory and motion perception during self-motion," *Behavioral and Brain Sciences*, vol. 17, no. 2, pp. 330–331, 1994.

[20] I. P. Howard, *Seeing in Depth Volume 2: Depth Perception*. Thornhill, Ontario: I. Porteous, 2002.

[21] J. Cutting, "How the eye measures reality and virtual reality," *Behaviour research, methods, instruments and computers*, vol. 29, no. 1, pp. 27–36, 1997.

[22] L. L. Kontsevich, "Defaults in stereoscopic and kinetic depth perception," *Proceedings of the Royal Society of London B*, vol. 265, no. 1615-1621, 1998.

[23] B. Rogers and M. Graham, "Motion parallax as an independent cue for depth perception," *Perception*, vol. 8, no. 2, pp. 125–134, 1979.

[24] A. C. Beall, J. M. Loomis, J. W. Philbeck, and T. G. Fikes, "Absolute motion parallax weakly determines visual scale in real and virtual environments," in *SPIE Proceedings on Human Vision, Visual Processing and Digital Display*, vol. 2411, 1995, p. 10.

[25] S. Fisher and K. Ciuffreda, "Accommodation and apparent distance," *Perception*, vol. 17, pp. 609–621, 1988.

[26] M. Mon-Williams and J. Tresillian, "Some recent studies on the extraretinal contribution to distance perception," *Perception*, vol. 28, pp. 167–181, 1999.

[27] A. Holway and E. Boring, "Determinants of apparent visual size with distance variant," *American Journal of Psychology*, vol. 54, pp. 21–37, 1941.

[28] A. Eadie, L. Gray, P. Carlin, and M. Mon-Williams, "Modelling adaptation effects in vergence and accommodation after exposure to simulated virtual reality stimulus," *Ophthalmic Physiological Optics*, vol. 20, no. 3, pp. 242–251, 2000.

[29] K. Akeley, S. J. Watt, A. R. Girshick, and M. S. Banks, "A stereo display prototype with multiple focal distances," *Proceedings of SIGGRAPH 2004*, vol. 23, no. 3, pp. 804–813, 2004.

[30] G. Mather and D. R. R. Smith, "Depth cue integration: stereopsis and image blur," *Vision Research*, vol. 40, pp. 3501–3506, 2000.

[31] G. Mather and D. Smith, "Blur discrimination and its relation to blur-mediated depth perception," *Perception*, vol. 31, pp. 1211–1219, 2002.

[32] N. Bruno and J. Cutting, "Minimodularity and the perception of layout," *Journal of Experimental Psychology: General*, vol. 117, no. 2, pp. 161–170, 1988.

[33] J. Tresillian and M. Mon-Williams, "Getting the measure of vergence weight in nearness perception," *Experimental Brain Research*, vol. 132, pp. 362–368, 2000.

[34] M. F. Bradshaw, A. Glennerster, and B. F. Rogers, "The effect of displays size on disparity scaling from differential perspective and vergence cues," *Vision Research*, vol. 36, no. 9, pp. 1255–1264, 1996.

[35] J. Hillis, M. Ernst, M. Banks, and M. Landy, "Combining sensory information: Mandatory fusion within, but not between senses," *Science*, vol. 298, pp. 1627–1630, 2002.

[36] D. Rose, *Research methods in psychology*, 2nd ed. London: Sage Publications, 2000, ch. Psychophysical Methods, pp. 194–210.

[37] J. W. Philbeck and J. M. Loomis, "Comparison of two indicators of perceived egocentric distance under full-cue and reduced cue conditions," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 23, no. 1, pp. 72–85, 1997.

[38] G. Bingham and C. Pagano, "The necessity of a perception-action approach to definite distance perception: monocular distance perception to guide reaching," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 24, no. 1, pp. 145–168, 1998.

[39] E. Johnston, "Systematic distortions of shape from stereopsis," *Vision Research*, vol. 31, no. 5, pp. 813–826, 1991.

[40] J. Cutting and P. Vishton, "Perceiving layout: The integration, relative dominance, and contextual use of different information about depth," in *Handbook of Perception and Cognition: Vol. 5: Perception of Space and Motion*, W. Epstein and S. Rogers, Eds. NY: Academic Press, 1995.

[41] G. Bingham, A. Bradley, M. Bailey, and R. Vinner, "Accommodation, occlusion and disparity matching are used to guide reaching: A comparison of actual and real environments," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 27, no. 6, pp. 1314–1334, 2001.

[42] D. Knill, "Reaching for visual cues to depth: The brain combines depth cues differently for motor control and perception," *JOV*, vol. 5, pp. 103–115, 2005.

[43] J. M. Loomis, N. Fujita, J. Da Silva, and S. Fukusima, "Visual space perception and visually directed action," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 18, no. 4, pp. 906–921, 1992.

[44] D. Proffitt, J. Stefanucci, T. Banton, and W. Epstein, "The role of effort in perceiving distance," *Psychological Science*, vol. 14, no. 2, pp. 106–112, 2003.

[45] C. S. Sahm, S. H. Creem-Regehr, W. B. Thompson, and P. Willemsen, "Throwing versus walking as indicators of distance perception in similar real and virtual environments," *ACM Transactions on Applied Perception*, vol. 2, no. 1, pp. 35–45, jan 2005.

[46] J. M. Plumert, J. K. Kearney, J. F. Cremer, and K. Recker, "Distance perception in real and virtual environments," *ACM Transactions on Applied Perception*, vol. 2, no. 3, pp. 216–233, 2005.

[47] R. B. Welch, *Handbook of virtual environments: Design, implementation, and applications*. Lawrence Erlbaum Assoc., 2002, ch. Adapting to virtual environments, pp. 619–636.

[48] P. Jansen-Osmann and B. Berendt, "Investigating distance knowledge using virtual environments," *Environment and Behavior*, vol. 34, no. 2, pp. 178–193, 2002.

Applying photovoice methods to community based media

Lorna Boschman
Simon Fraser University
School of Interactive Arts and Technology
Surrey BC Canada
lboschma@sfu.ca

ABSTRACT

When social networks of individuals are detached from modern digital communications tools, their ability to represent themselves and their community is affected. If technological access and education is limited, the ability to create media artifacts is lessened, resulting in a marginalized community becoming further isolated. Two strategies to assist individuals to find their voice are examined. *Photovoice* is a public health research practice that encourages participants to photograph what is important to them; community based video projects empower individual members, and allow novice directors to find expression audio-visually.

Photovoice methods can inform community media projects in establishing these practices as standard: Through a community partner, individuals are chosen to be trained in media production and critical analysis; group members discuss their media projects with the group and with researchers; and the finished work is contextualized and exhibited so that policy-makers and community members can view it.

Author Keywords

Developmental disabilities, cognitive disabilities, community-based participatory media, photovoice, technology adaptation.

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

The adoption of digital technology has become more commonplace for those with educational or financial resources; at the same time, it has tended to marginalize communities who lack access to higher education or disposable income. As a result of this exclusion from interaction with emergent information and new media technologies, the opportunity for expression by individuals within these communities is limited. In response to this communications deficit, two distinct tendencies have emerged that typify the movement to supplement polyphonic articulation. The first approach, *photovoice*, is a well-established community based public health method while the second is based on observations of a community based media arts initiative. *This Ability Media Club* is a partnership between a national media agency, a community partner, and an artist-technician.

Photovoice is a participatory action research (PAR) method, first developed by public health researchers Wang and Burris for a project involving village women in Yunnan Province, China [16]. PAR projects have common approaches: The participants are partners with the researcher, the entire group decides which images represent the concerns of their constituency, the images reflect social as well as individual concerns and the members are empowered by creating media objects based on critical reflection [1]. The *photovoice* method allows individuals who are lacking in social power to influence public policy through photography, utilizing the representational power of the visual to give voice to common concerns [7]. *Photovoice* encourages participants to record the strengths and weaknesses of their local community, to discuss issues of interest or concern with other participants, and to influence public policy by involving community leaders and policymakers. Caroline C. Wang, co-creator of the method said, “What experts think is important may not match what people at the grassroots think is important [7].”

The National Film Board of Canada (NFB), a co-sponsor of *This Ability Media Club*, has a history of working to define national identity through media production. Earlier in the 20th century, NFB projectionists toured rural areas of the country, showing the latest documentary productions. In the late 1960’s, the NFB took an important step in participatory media production and encouraged adult literacy through their *Challenge for Change* program [9]. Groups representing financially disadvantaged local communities, whose members were excluded from traditional media production, were given video equipment and support technicians in order to record their concerns. The *Challenge for Change* method exists in contrast to a more traditional documentary approach: A well-meaning director visits a community, makes observations, and records a sympathetic portrait.

Often, this type of production is made without consulting the individuals involved. Thus, the director may be insensitive to local concerns. In contrast, *Challenge for Change* asks local participants to create their own portraits of the community, resulting in works that more closely represent actual local issues. Participants are supported in developing their ability to voice authentic concerns in contemporary media language.

Photovoice asset mapping/analysis and community based media projects allow participants to express themselves in ways that are personally empowering and socially constructive. The areas of social concern that have been addressed by *photovoice* research projects range from women's health issues in a low income community near San Francisco, California to the social participation of youth in Flint, Michigan [7, 10, 12-16]. In this paper, *photovoice* methods and best practices will be discussed and compared to initial observations of *This Ability Media Club*, a Burnaby, BC community based media workshop for people with developmental disabilities. The author directs this project and serves as a technician-in-residence for the productions. Figure 1 shows most of the core participants, gathered around a camera while setting it up.

In the *photovoice* method, researchers work in partnership with community organizations to involve participants. Each individual is given an inexpensive camera and is asked to record images in order to raise issues within the group, and with policymakers, and to affect social change within the area. The expertise of people from the neighborhood is acknowledged; their viewpoints supplement those of professionals who work within the community. As well as sharing the photographs with researchers and policymakers, public exhibition of the images allows the community to come together and celebrate a common vision.

Recently, the new English Program Director General of the NFB, Tom Perlmutter, asked the regional administrators of the organization to initiate projects that followed in the spirit of the *Challenge for Change* era. The executive director of the Pacific Region, Rina Fraticelli, identified several groups - people with disabilities, elders, low-income residents in a Vancouver suburb, and northern indigenous people, as among those who could benefit from such a program. In each case, a local partner was chosen to host the workshops. Fraticelli commissioned a 2004 report that investigated how the organization could provide technical support and guidance to people with cognitive disabilities through a partnership with the Burnaby Association for Community Inclusion (BACI), a local organization ranked among the highest in North America for service and community development [9].

Figure 1: Setting up the camera: program director with This Ability Media Club members

At the same time as the NFB was looking for a way to revitalize the spirit of *Challenge for Change*, BACI was going through an organizational transformation, based on the philosophy of exploring citizenship in the context of disability [9]. As a result, when *This Ability Media Club* formed, the first subject question that participants in the group grappled with was, "What does citizenship mean to you?" It was hoped that exploring this question would result in a deeper understanding of citizenship for people with developmental disabilities.

The initial goals of *This Ability Media Club* were to empower the participants and to allow them to have an opportunity to speak about their lives using media tools and the powerful NFB distribution network. Most of the participants had a little media experience, including working as extras in the local film industry, and acting in community theatre productions.

Those with cognitive disabilities have been defined by the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) as "significantly limited in at least two of the following areas: self-care, communication, home living social/interpersonal skills, self-direction, use of community resources, functional academic skills, work, leisure, health and safety" [4]. BACI works with individuals who have been diagnosed with Downs syndrome, cerebral palsy, autistic spectrum disorder, and other cognitive or developmental disabilities.

Rather than referring to individuals by their medical condition, representatives from within these communities refer to themselves as "self-advocates," a term that will be used to describe these people for the remainder of the paper.

THEORETICAL BASIS

Photovoice was created in response to theoretical literature and participatory education or photography projects. The *photovoice* process developed through application of the practice in the Yunnan Women's Reproductive Health and

Development Program. Wang cites feminist theory as another basis for the *photovoice* method, especially for those projects involving the health concerns of women and children [14, 16]. On this theoretical basis, ordinary women become the experts on their own subjective experience and that expertise is valued. Part of the *photovoice* researcher's role is to assist in communicating this knowledge to a wider audience. Through taking photographs and discussing the themes that emerge, the concerns of women and others with less social power will be brought to the attention of those individuals who shape public policy [10, 12-16].

While professionals may be well intentioned, their concerns are not always those of the grassroots community. For example, public health officials in Conta Costa, a low-income area in the San Francisco Bay area, had identified specific concerns regarding health issues within the community. The concerns typically included measurable quantities such as low birth weight and maternal mortality. However, through a *photovoice* project, residents expressed the need for safe recreational outlets for their children and for community development within the neighborhood [14]. By combining professional with community perspectives, an opportunity is provided for self-expression, and the resulting strategies for health improvement are more likely to meet local needs.

Photovoice also draws on the tradition of education for critical consciousness, as formulated by the Brazilian educator and theorist Paulo Freire. According to this philosophy, it is important to encourage people to speak about the conditions in their own environment and how this relates to the lives of those around them [14]. In developing a critical consciousness, one begins to understand that recognizing words is only a part of literacy. If the learner is still unable to discern misleading statements about one's own condition, a state of illiteracy continues.

Rather than basing their policies directly on academic discourse, the NFB looks to their own history of encouraging people without social power to find their "authentic voice" and to learn to speak out. In a half-hour film about *This Ability*, Rina Fraticelli states that women have been encouraged to speak for themselves, indigenous people have found their voice, and now it is time for people with disabilities to do the same [11]. In the past, this tendency in documentary filmmaking resulted in the formation of NFB studios that specifically encouraged women or indigenous people to direct their own productions. This perspective is based on the idea that a member of any community is in a better position to express the concerns and the inspirations of that community than an outsider.

Both the *photovoice* projects and the NFB community-based initiatives are based on empowering individuals within a community by creating an opportunity for expression through media. Mainstream media outlets reflect the ideology of those with social power, while those without resources are silenced through their lack of education and inadequate access to communications technology. Both types of programs aim to enrich the public discourse through encouraging those whose perspectives have not been heard to speak out through visual media.

In comparing the two methods more closely, the differences in approach are most clearly shown in expected outcomes and in the final ownership of the cultural artifact. The methods will also be examined to determine if providing opportunities for expression among the participants increases their ease of use and comfort with media equipment.

Methods used in *photovoice* projects led by Caroline C. Wang will be compared to an ongoing participatory media workshop, *This Ability Media Club*, led by Lorna Boschman and jointly sponsored by the NFB and BACI. *Photovoice* is conducted as part of a research project, in order to facilitate the transfer of knowledge from community members to those who have social power and the ability to transform the community through their decisions. As a by-product of media creation, participants in *This Ability Media Club* create knowledge about the target community. The goal in participatory media projects is to nurture novice directors who act as representatives of both individual and community concerns and attributes. The resulting media works are able to give outsiders a clearer understanding of the lives of those subjects and contribute to a collective understanding of the group. In both projects, the exhibition of these resulting media projects is a source of pride for the community.

PROCESS AND METHODS

Recruitment and initial training

Before a *photovoice* project begins, researchers form a partnership with a local grassroots organization. This facilitates the successful recruitment of participants. In order to undertake a *photovoice* program, known issues within the target community are conceptualized and the broad goals are identified. Local leaders from the *Neighborhood Violence Prevention Collaborative*, a coalition representing 265 local clubs, provided impetus for a *photovoice* project in Flint, Michigan, USA. Like Detroit, Flint was a city of automobile manufacturing, but economic changes forced the community to re-examine its economy, culture and race relations. Although *photovoice* practice generally recruits policymakers to view the finished work, in this case the community requested that those who set policy be involved in filming as well. In this way, their perspectives on the community's assets and liabilities could be compared to those of the other groups, which consisted of youth leaders, youth who were deemed the most at-risk by the professionals who worked with them, and adult community representatives.

Photovoice participants were recruited through the host Flint organizations. The training workshops began with a discussion of the ethical considerations involved in portrait photography. Participants were warned not to take pictures without permission. Although taking anonymous photos of strangers can produce images that do not look rehearsed, fears for the safety of participants were addressed. In addition to being ethically questionable, an irate subject might attempt to retaliate physically. Participants were instructed to ask subjects to sign a release before their image was recorded. They also gave copies of the portrait to the subject and promised that another release would have to be signed before the image was shown publicly. Part of the *photovoice* philosophy involves showing appreciation to the community by returning the images.

In order for a community media project to be established, a strong alliance was built between the NFB and BACI. In addition to hosting *This Ability Media Club* at their Still Creek Centre, BACI provided a part-time liaison worker to the project. Begun by parents of children with developmental disabilities, BACI still provides social and advocacy services to self-advocates and their families in Burnaby, BC, Canada. A committee comprised of representatives from BACI, the NFB and the self-advocates of Burnaby hired the director of the program. An Advisory Committee met regularly during the first year of the program and included parents of some of the participants, as well as the key stakeholders.

After consulting with BACI staff, a time for the weekly workshop was chosen that had few conflicting events. *Media Club* was also scheduled at a time when self-advocates could arrange transportation. Although most of *This Ability* participants use public transportation independently, the work, training or recreational schedules of others at BACI are coordinated around the bus or van schedules of those who are less mobile. The transportation schedules and daily routines of potential participants should be taken into consideration when initiating a community project. Once participants added *This Ability Media Club* to their weekly schedule, almost all of them returned. The core group consists of 7 self-advocates and 2 paid researchers, smaller than the *photovoice* projects cited but scaled to the available resources.

Before formal workshops began at BACI, a professional digital video camera was introduced informally at the Still Creek Centre. Articles in BACI's newsletter, announcements on BACI's website, open screenings of the raw footage at the Still Creek Centre, and word of mouth helped to publicize *This Ability*. When participants were recorded, they were asked to sign a model release that allowed the NFB to use the footage worldwide. BACI also held copyright of the footage and by negotiating, became the solo copyright holder of the short films directed by self-advocates from *This Ability Media Club*. Some of the participants, such as the director shown in Figure 2, chose to be filmed at work, a source of pride for the participants.

Figure 2: This Ability director at work, in a still from her short film "I Love My Job!"

Selection and discussion of the imagery

In the Flint *photovoice* project, the groups met monthly to return exposed rolls of film and met five times to discuss the images. Professional photographers were assigned to teach participants technical skills. Ten local *photovoice* facilitators also worked with the participants in order to direct the discussion toward critical dialogue, exploring the roots of problems that were noted. Participants were asked to choose one or two photos from each roll and to write informally about it. In the *photovoice* method, questions are centered on the mnemonic "SHOWeD": "What do you See here? What is really Happening? How does this relate to Our lives? Why does this problem or strength exist? What can we Do about it? [15]" If more than 4 photos and stories by participants were related to the same subject, it was defined as a "theme" [13].

Part of the *photovoice* method is to share selected photos with the group and with the researchers. The group discusses what has been shot and what these images say about the community. If the images incriminate the subjects or portray the subjects in a negative light, they are not exhibited through the *photovoice* project. One photo illustrates a baby about to stick a pin into an electrical socket while the mother chats on the phone. The photographer was minding the child, and yelled out to the baby while taking the shot. The baby dropped the pin. The image was captured as part of public health research centered, in part, on young mothers who weren't ready for the responsibilities of parenthood. While the mother in question may have agreed that her image could be used, she may not have realized that she could be presented as an example of irresponsible parenthood [15].

Both *photovoice* and the community media project encourage participants to focus on interpretations of their own position within the community. In the case of This Ability, the focus was not just on the positive or negative but also something in between, related to a shared human experience and to the social contributions of people with different abilities [9]. BACI and This Ability Advisory Committee suggested a theme: "What does citizenship mean to you?" When self-advocates asked for clarification of the language, the question was rephrased as, “What does community mean to you?”

The term “community” is historically significant to self-advocates, as it refers to a time when they were able to leave their previous lives, housed in large institutions, and go out to live in the wider community. In British Columbia, the last large institution that housed those with cognitive disabilities, Woodlands, was finally closed in 1996. The *Cognitive Levers* research group at the University of Colorado notes that closure of large institutions has been undertaken in the US over the past 30 years, in part due to legislative and social impetus [3]. Cognitively disabled individuals who had lived in institutions in the past were now in community settings with varying degrees of independence.

During the first year, *This Ability Media Club* members learned basic documentary production techniques. In the beginning, the Advisory Committee instructed the *Media Club* to engage individuals in the process of creating media as a way of building self-esteem, and media literacy skills and to find a way to include individuals in the BACI community who were not verbally articulate. As participants began to gain confidence, they filmed other self-advocates in volunteer positions with the Red Cross and at a local recycling centre. Individuals who were filmed were asked to sign a release, which allowed the NFB to later use the footage.

Much of the footage that was shot was shared first with the program participants. During the media workshop, participants helped set up the television monitor and took notes while they viewed the footage. Their comments helped participants to see how others responded to their work. A selection of footage was presented at each Advisory Committee meeting, along with a brief verbal report by the two paid staff. Over the course of the first year, stakeholders attending the meetings included *This Ability* participants, parents of the individuals and other BACI self-advocates, BACI Executive Directors, the NFB producer and executive producer, marketing staff, web-exhibition liaison and interns and *Philia – a dialogue on caring citizenship* strategic planners. The directors were in artistic control of their productions, except where they violated copyright laws.

Public exhibition

In the *photovoice* method, the later stages of the project involve showing the photographs to a wider audience, in an attempt to influence public policy. Researchers work with participants to plan public exhibitions of the photos and writing. Community leaders and policy makers are invited to these events, in order to allow direct access by *photovoice* participants. A powerful photo by a 17-year-old participant showed a bullet hole in the window of his school bus. He wrote “I can tell that the bus I ride in is different because the bullet holes are always in different windows. [13]” The image deeply moved policymakers and health officials, and was reproduced in the both local and national media.

During the first year of *This Ability Media Club*, six short video productions were completed. Directed by the self-advocate participants, the films were also under their artistic control. Technical assistance during filming and editing was provided by the program director. Professional audio and video post-production technicians provided the finishing touches. Although the ownership of the films was not in the hands of the directors, this policy is standard when a project is produced by the NFB, a Canadian film agency. Even by their standards, the films used a relatively low budget, which was underwritten by the NFB. As a result of negotiations with their collaborating agency, copyright for the six short films now rests with BACI.

The *This Ability Media Club* films were launched at a large community event, hosted by BACI, with over a hundred people in the audience at the Still Creek Centre. Directors were recognized at the event and were given awards that resemble the Oscars. In Figure 3, the directors are shown with their trophies. Awards were also given to participants who were not able to direct their own production. The six finished films premiered to an international audience on the CitizenShift website, a forum sponsored by the NFB for “free-range media”. During the initial run, when *This Ability* was linked from the main page of the site, over 2,000 distinct visitors watched the films daily. The online dossier is a permanent record of the group, including the films, photos from the *Media Club*, directors’ bios, short interview clips with the directors, planning documents related to the project, and films by other artists with disabilities [11].

To date, the film that has received the most critical acclaim is Michelle McDonald’s “Be Kind to Spiders”. In it, the director talks about her love of spiders and asks the audience not to kill them. She reminds us that over-consumption led to the demise of the buffalo, delivering a powerful ecological message about living in harmony. Although some of the films can stand on their own and have the potential to reach a wide audience, the NFB decided that a “wrap-around” film would showcase the perspectives of the group more effectively. A half hour film, “This Ability”, was written and directed by Lorna Boschman, the program director, and includes the six shorter works, interviews with the short-film directors and other stakeholders and group interaction during workshops [11]. Copyright for the longer film is retained by the NFB, which intends to promote the work to the festival and educational markets internationally.

DISCUSSION

Choice of technology

In comparing Wang’s *photovoice* projects to *This Ability Media Club*, differences begin to emerge due to the technological medium. The *photovoice* projects use an inexpensive camera with black and white film. Each week, participants are asked to shoot a roll and hand it in for processing, which is paid for by the project. Two copies are made, one for the participant and the other for the project analysts. Copyright is retained by the photographer. Public exhibition of the work requires an additional release be signed by the subject and by the photographer. Point and shoot analogue photography is simpler than digital, since a computer is not required for refinishing or printing.

In the first four months of *This Ability Media Club*, participants came to the workshop weekly and practiced their video production skills. Unlike the simple camera used by *photovoice*, group members were expected to learn and be able to operate an entire package of equipment. It consisted of a professional Sony PD170 digital video camera, mounted on a Manfrotto tripod, a shotgun microphone and a boom pole. One great advantage of this camera, from a usability perspective, is that the controls are buttons on the outside housing, rather than within a complicated and more conceptual digital menu. The ease of control that a allows a cinematographer to quickly make adjustments while recording are the same ones that allow self-advocates to see a direct relationship between a physical button and the corresponding image or sound. Several group members have poor motor control. Using a tripod to mount the camera, they are able to keep a steady image while looking in the LCD screen to monitor what is being recorded.

In the first four months of the program, participants in the group became subjects, as they learned to interview each other during *Media Club*. Because of concerns over equipment safety and the risk of loss, the equipment was always conveyed to the location by the paid staff of the program. When not in use, the items were locked in secure storage in BACI's vault. Because the equipment was on loan from the NFB, a federal agency, it was not possible to insure it. Participants were not free to use the equipment except under the supervision of the program director. As a result, the shooting times and locations were all pre-planned and not likely to be the result of spontaneous gestures or experiments by the novice directors.

Teaching media skills to self-advocates

The BC Ministry of Health, Special Education, and school psychologists suggest that instructions for cognitively disabled learners be less abstract and more concrete [5, 8]. Following this principle, training in *This Ability Media Club* followed step-by-step instructions on how to assemble the equipment and to begin recording. Learners were trained to use the equipment in automatic settings to avoid unnecessary complications. The stated goal of the first six months of training was that the self-advocate participants would be able to assemble the equipment unassisted.

Figure 3: This Ability members receive their awards at the BACI launch

If the optional manual features were mentioned as points of interest, they risked being seen by the participants as too abstract. In a strategy to make these options more concrete, they were raised as part of the solution to a technical problem that confronted a group member. When the group reviewed footage that had been shot during previous outings, someone might comment that something looked “wrong” to them. In response, the program director would explain, technically, what the problem was, and would suggest a possible solution. For example, in a shooting situation where the camera faced a bright light, the concept of manual exposure was introduced. In a situation where the group noted that the sound quality was poor, the addition of a second microphone to the audio recording was suggested. The process of setting up the equipment was broken down into ten easy steps and each participant received these instructions. However, remembering the movements through repetition seemed to be a more effective learning strategy for this group than reading the more abstract written instructions.

After a six-month program review, the group changed direction, resulting in new educational objectives. Each individual in the group was asked to direct their own production, rather than participating in a group project. It was at this point that members of the group began to speak for themselves, rather than trying to interpret the needs of their community. In order to re-enforce their new role as directors, personalized letterhead was created for everyone in the group. On a white page, a small photo of the participant and their name followed the title “Director”. These pages were used to write comments, draw pictures or create other reminders for the critique that followed viewing video footage. The round of discussion encouraged everyone to comment on the footage, or as time progressed, on the rough edits of each person's video. Reviewing and commenting on the work of others is a vital part of active learning in the workshop and could be compared to the process of developing a critical consciousness in *photovoice*.

FUTURE WORK

The critical examination of *photovoice* methodologies and the experience of meeting weekly with self-advocate directors in *This Ability Media Club* have raised two important issues. First, how is the technological form of expression chosen and whose needs does it meet? Second, how can the program become sustainable? Although *This Ability* began without a defined output format and the participants were encouraged to engage in the process, the funding agency eventually asked for a product that was in the familiar form of a documentary short. The group members were trained with professional video gear and learned to edit using Final Cut Pro on a Mac computer. However, once the funding for the program has ended, the video equipment, Mac computer, and their personal trainer will be gone. Although the participating self-advocates will have learned valuable skills, they will not be able to practice and will lose a portion of what they have learned as a result.

In order to prepare for the end of the current phase of the project, the group will be interviewed and asked to complete a short survey to determine if they have access to computer technology in their homes or through BACI. Rather than continuing to learn to make films, one future direction may be to encourage group members to explore participatory media that is readily available to the public at little or no cost. If the survey reveals that everyone wants to have their own webpage, for example, they could learn how to take digital stills, write for the net and upload their creations. If the group wants to venture into storytelling, they could learn to construct short narratives from scanned stills or digital photos with a voice-over. By learning to adopt new technology to their needs, participants will continue to develop their voices as active citizens, sharing their perspective with the public in order to further general understanding of the lives of self-advocates.

RECCOMENDATIONS

1. In order for a project to be more accessible to participants, their work and transportation schedules should be determined through close interaction with local partners in the community. Planning around the life of the community will increase their ease of participation.

2. The choice of digital media tools should be ones that are readily available to workshop participants after the project has ended. Although individuals will feel empowered by the work they do in the group, the effects should be sustainable if the long-term goal of the project is to increase use of digital or new media tools within that community.

CONCLUSION

The three phases of the *photovoice* method can be applied successfully to community media arts projects. Establishing standards provides guidelines to media practitioners working with communities and allows them to build on the work of others in similar projects. Education that is customized for the group should accompany access to the technology, and technical mentorship should be available. In order to refine the images and text, other participants should critically examine the creative works. When the media works are exhibited, emphasis should be placed on inviting those with political power to attend. At the same time, the individuals are a source of pride to their community, in part because they speak for themselves and others, slowly breaking down their marginalization.

ACKNOWLEDGMENTS

Thanks to the National Film Board of Canada who funded and facilitated This Ability Media Club and to the Burnaby Association for Community Inclusion, who hosted and co-funded the workshop. Thanks also to SSHRC, CGS Master's Scholarship program, which has supported Lorna Boschman's research in the second year of the project.

REFERENCES

1. Berg, Bruce L. Qualitative Research Methods for the Social Sciences. *Chapter 7: Action Research* (1992), 195-208.
2. CitizenShift – Reel Community – This Ability. Home page of This Ability Media Club dossier with links to finished films, director's biographies and related links. http://citizen.nfb.ca/onf/info?did=1581.
3. Dawe, M., Fischer, G., Gorman, A., Kintsch, A., Konemi, S., Sullivan, J., Taylor, J., and Wellems, G. Smart care: the importance and challenges of creating life histories for people with cognitive disabilities. *Proceedings of the HCI International Conference (HCII) Las Vegas (2005).*
4. Dawe, Melissa. Desperately Seeking Simplicity: How Young Adults with Cognitive Disabilities and Their Families Adopt Assistive Technologies. *CHI 2006 Proceedings* (2006), 1143-1152.
5. Government of British Columbia, Ministry of Education. Special Education Services: A Manual of Policies, Procedures and Guidelines. Students with Intellectual Disabilities, including Students with Moderate to Profound Intellectual Disabilities. Retrieved June 26, 2006 from http://www.bced.gov.bc.ca/specialed/ppandg/planning_2.htm.
6. Harrison, Barbara. Seeing health and illness worlds – using visual methodologies in a sociology of health and illness: a methodological review. *Sociology of Health & Illness 24, 6* (2002), 856-872.
7. *Photovoice*: Social Change Through Photography http://www.*photovoice*.com/.
8. Shaw, Steven R. Academic Intervention for Slow Learners. National Association of School Psychologists *NASP Communiqué* 28, 5. Retrieved July 12, 2006 from

http://www.nasponline.org/publications/cq285slowlearn.html.

9. Smith, Brian. This Ability: Revealing the Contributions of People with Disabilities. Linked from "A Preliminary Report on the NFB Collaboration with the Burnaby Association for Community Inclusion and Philia." Retrieved from *CitizenShift* on July 12, 2006 from http://citizen.nfb.ca/onf/info?aid=6363&atid=24.

10. Strack, R. W., Magill, C., and McDonagh, K. Engaging Youth Through *Photovoice*. *Health Promotion Practice 5, 1*(2004), 49-58.

11. This Ability, half-hour documentary about This Ability Media Club. Available in September 2006 from The National Film Board of Canada.

12. Wang, C. C., Cash, J.L., and Powers, L. S. Who Knows the Streets as Well as the Homeless? Promoting Personal and Community Action Through *Photovoice*. *Health Promotion Practice 1, 1* (2000), 81-89.

13. Wang, C.C., Morrel-Samuels, S., Hutchison, P. M., Bell, L., and Pestronk, R.M. Flint *Photovoice*: Community Building Among Youths, Adults, and Policymakers. *American Journal of Public Health 94, 6* (2004), 911-913.

14. Wang, Caroline C., and Pies, Cheri A. Family, Maternal, and Child Health Through *Photovoice*. *Maternal and Child Health Journal 8, 2* (2004), 95-102.

15. Wang, C., and Redwood-Jones, Y.A. *Photovoice* Ethics. *Health Education and Behavior 28, 5* (2001), 560-572.

16. Wang, Caroline C. *Photovoice*: A Participatory Action Research Strategy Applied to Women’s Health. *Journal of Women’s Health 8, 2* (1999), 185-192.

X-ray quantum counting pixel architecture for digital tomosynthesis

Amir H. Goldan and Karim S. Karim, *Member, IEEE*

Abstract—**Quantum counting is emerging as an alternative detection technique to conventional quantum integration. In quantum counting systems, the value of each image pixel is equal to the number of photons that interact with the detector. The proposed pixel architecture provides a method for energy windowing and false-count elimination. Each pixel is comprised of a radiation detector and integrated analog and digital circuitry. A low noise operational amplifier (opamp) was designed in 0.18 μm CMOS technology with high gain over the entire input common-mode voltage range. A prototype was also developed on a printed circuit board (PCB) using discrete electronic components. Measurements show that the quantum counting operation of the proposed pixel architecture is successful for high x-ray energies.**

Index Terms—**CdZnTe photoconductor, complementary folded cascode, current control circuitry, energy windowing, false-count elimination, low noise, photon counting, quantum counting, radiation detector, silicon PIN photodiode.**

I. Introduction

CURRENTLY, most digital mammography tomosynthesis detectors are based on integrating the x-ray quanta (photons) emitted from the x-ray tube for each frame. This technique is vulnerable to noise due to variations in the magnitude of the electric charge generated per x-ray photon. Higher energy photons deposit more charge in the detector than lower energy photons so that in a quantum integrating detector, the higher energy photons receive greater weight [1]. In mammography, the higher part of the 30 kVp energy spectrum provides lower differential attenuation between tissues, and hence, these energies yield images of low contrast.

X-ray quantum counting detectors solve the noise problem associated with photon weighting by providing better weighting of information from x-ray quanta with different energies. In an x-ray quantum counting system, all photons detected with energies above a certain predetermined threshold are assigned the same weight (i.e., unity). Adding the energy windowing capability to the system (i.e., counting photons within a specified energy range) theoretically eliminates the noise associated with photon weighting and decreases the required x-ray dosage by up to 40% compared to quantum integrating systems [2].

Background noise is also present due to false-counts, even when the detector is not irradiated with photons [3], [4], and this noise source decreases the contrast resolution of the digital image. False-counts are caused by four phenomena: 1) integration of noise-electrons, 2) integration of the radiation detector dark current, 3) integration of electric charge generated as a result of continuous low-energy photon-detector interactions, and 4) cosmic radiation. The third phenomenon relates to the case where continuous photons with energies lower than the specified energy range generate an electric charge with the same weight as a photon with an energy that falls within the range and cause a false-count. The severity of the first three phenomena depends on the system noise level, the radiation detector dark current, and the energy spectrum of the incident photons.

In this research, we introduce the design and implementation of an x-ray quantum counting pixel with energy windowing and false-count elimination for digital tomosynthesis. Each pixel is comprised of: a radiation detector, a low-noise charge amplifier (CA), a set of comparators, a decision-making unit (DMU), a mode selector, and a pseudo-random counter. The integrated mixed-signal components were designed and simulated based on the Taiwan Semiconductor Manufacturing Company (TSMC) 0.18 μm N-WELL process. We also present the design of a complementary folded-cascode (CFC) opamp, with current control circuitry, to achieve high gain over the entire input common-mode voltage range. The improvement in performance of the modified architecture compared to the conventional single-stage CFC opamp was analyzed, and theoretical calculations and simulation results of the opamp's gain and noise are presented.

A quantum counting pixel was developed on a printed circuit board. Two radiation detectors were used for direct conversion of x-rays to electric charge: a plastic silicon PIN photodiode from Fairchild Semiconductor (QSE773), and a single-crystalline cadmium-zinc-telluride (CdZnTe) photoconductor developed by Redlen technologies. Individual photon detection was successful using both detectors.

II. Pixel Architecture

The quantum counting pixel has two distinct modes of operation: 1) detection mode, and 2) readout mode. When a photon is incident on the radiation detector in detection mode,

Manuscript received June 30, 2006. This work was supported in part by the Breast Cancer Society of Canada and the Natural Sciences and Engineering Research Council of Canada.

A. H. Goldan is with the Engineering Science Department, Simon Fraser University, Burnaby, BC V5A 1S6 CANADA (phone: 604-922-3356; fax: 604-925-3357; e-mail: agoldan@sfu.ca).

K. S. Karim is with the Engineering Science Department, Simon Fraser University, Burnaby, BC V5A 1S6 CANADA (phone: 604-268-6859; fax: 604-291-4951; e-mail: kkarim@sfu.ca).

the photon-generated charge in the pixel is integrated. After charge integration, if the output of the on-pixel integrator falls within the specified energy window, the counter is incremented. When all the photons are counted, the pixel is switched to readout mode where counter bits are serially shifted out.

As illustrated in Fig. 1, each pixel is comprised of a radiation detector, a charge amplifier (CA), a set of comparators, a decision-making unit (DMU), a mode selector, and a pseudo-random counter.

For quantum counting systems, the radiation detector is required to have carrier multiplication gain, and the number of carriers generated per photon-detector interaction must be greater than that of the input noise. Amorphous selenium (a-Se) is a good example of a highly developed photoconductor in the x-ray imaging field. The overall conversion gain of a-Se depends on three processes: 1) the initial process of electron-hole pair (ehp) generation, followed by 2) the recombination process, and 3) the avalanche multiplication process due to impact ionization. For example, for the conversion gain of 20 ehp/keV at an electric field of E = 10 V/μm, approximately 1000 electrons will contribute to the input signal after a 50 keV photon is fully absorbed in a-Se. To increase detector sensitivity, a low-noise charge amplifier is required at each pixel. Also, increasing the electric field across the a-Se photoconductor increases the conversion gain, and hence, it increases detector sensitivity [5].

Cadmium-Zinc-Telluride (CdZnTe) is another compound semiconductor material that is promising for quantum counting systems. Because of CdZnTe's high density and high atomic number, thin CdZnTe photoconductors yield high x-ray absorption efficiencies. For example, 150-µm-thick CdZnTe yields an absorption efficiency of 85.1% for 20 keV x-ray photons. Also, CdZnTe is very sensitive to x-rays and has a conversion gain of 200 ehp/keV, which is ten times that of a-Se. However, CdZnTe photoconductors have low charge collection efficiency due to their relatively poor mobility-lifetime products [6], [7].

In detection mode, the photon-generated charge in the pixel is integrated by the CA, while the output of the CA is being compared to three thresholds: the reference threshold voltage, V_{ref}, the lower window threshold voltage, V_{lo}, and the upper window threshold voltage, V_{hi}. The relation amongst the threshold voltages is given as

$$V_{\mathrm{hi}} > V_{\mathrm{lo}} > V_{\mathrm{ref}} > V_{\mathrm{no}} , \tag{1}$$

where V_{no} is the peak output noise voltage of the charge amplifier. The two thresholds, V_{lo} and V_{hi}, are used to form the energy window, and V_{ref} is used by the DMU for automatically discharging the unwanted charge on the feedback capacitor, C_{f}, which could otherwise cause a false-count.

After the comparisons are performed, outputs of the three comparators are passed to the DMU. The DMU is responsible for correctly deciding when to increment the counter, and also when to discharge C_{f}. The DMU contains two delay elements, where the delay time, t_{delay}, depends on the DMU clock period, $T_{\mathrm{DMU\text{-}clk}}$. Fig. 2 depicts the high-level operation of the DMU by means of a flowchart.

The mode selector in Fig. 1 switches the pixel operating mode using two multiplexers based on the logic level of the *shutter* signal. In detection mode, shutter is set to logic "1", the output of the DMU serves as the clock for the counter, and the input to the first shift register, R_1, is the output of the XOR logic gate. However, in readout mode, shutter is set to logic "0", an external clock is generated for the pseudo-random counter, and the input to R_1 is the output of the last shift regis-

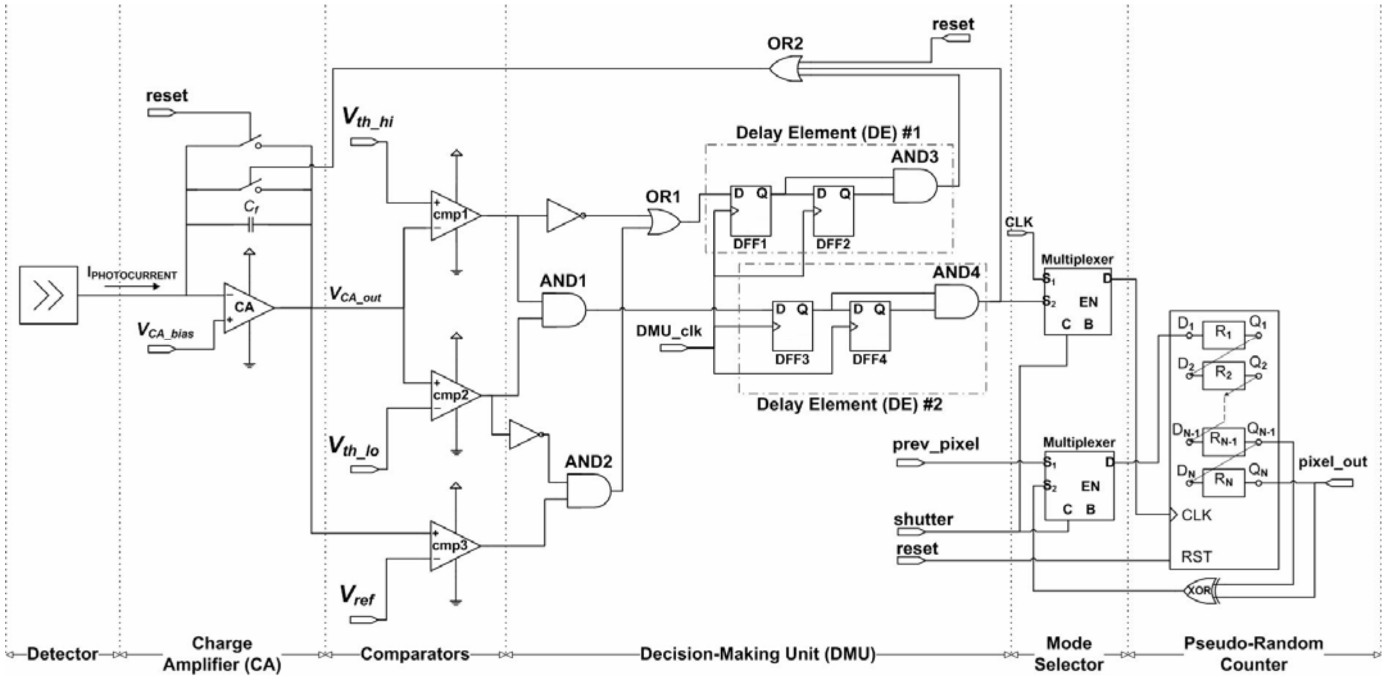


Fig. 1. The quantum counting pixel architecture with energy windowing and false-count elimination.

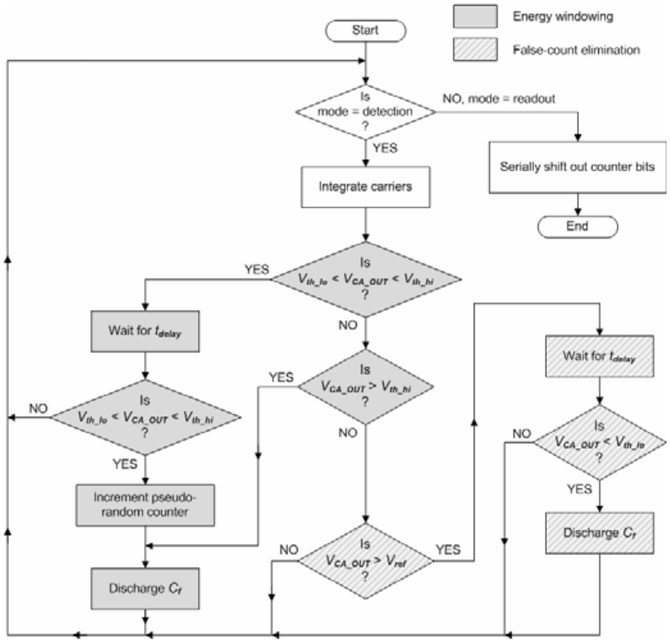

Energy windowing
False-count elimination
Start
Is mode = detection ?
NO, mode = readout
YES
Serially shift out counter bits
End
Integrate carriers
YES
Is $V_{th\_lo} < V_{CA\_OUT} < V_{th\_hi}$?
NO
Wait for t_{delay}
YES
Is $V_{CA\_OUT} > V_{th\_hi}$?
NO
Is $V_{th\_lo} < V_{CA\_OUT} < V_{th\_hi}$?
YES
Increment pseudo-random counter
Discharge C_f
NO
Is $V_{CA\_OUT} > V_{ref}$?
YES
Wait for t_{delay}
NO
Is $V_{CA\_OUT} < V_{th\_lo}$?
YES
Discharge C_f

Fig. 2. Illustration of the high-level operation of the DMU by means of a flowchart.

ter, R_N, of the previous pixel.

The pixel counter must satisfy three criteria: 1) operate at high speed for fast readout rate, 2) occupy a small area to minimize pixel size, and 3) have the capability of serial read-out. The pseudo-random counter satisfies all of the above criteria. The design is comprised of N shift registers, and on every clock pulse, counter bits are logically shifted to the right, and the least two significant bits are XORed in order to generate the most significant bit of the counter value. The counter behaves in a pseudo random fashion as it counts, and theoretically, it will return to its initial state after 2^N-1 counts. Note that the initial state of the counter must be known. In our design, the most significant bit of the counter is initialized to logic "1" and the rest are initialized to logic "0" after the counter is reset. With the current chip architecture, this reset state is hard coded into the chip.

III. LOW NOISE CHARGE AMPLIFIER

The input-stage charge amplifier (CA) increases detector sensitivity by its low noise operation, and by amplifying the electric charge generated as a result of a photon-detector interaction [5]. To have a constant charge gain that is independent of the detector junction capacitance, C_d, and the opamp's parasitic input capacitance, C_s, the open-loop gain of the opamp must be high enough to minimize charge loss to less than 10%.

The major components of noise in the quantum counting pixel are flicker and thermal noise of the CA's opamp, and shot noise caused by the radiation detector dark current, I_D. In most cases, the opamp is the dominant noise source, and the dark current shot noise is negligible due to the radiation detector's low dark current. Thus, we must minimize the input voltage-noise density of the input-stage opamp.

A. Design Overview

Two 1.8V CMOS rail-to-rail complementary folded cascode (CFC) operational transconductance amplifiers (OTA) have been designed and simulated. The simulations are based on the Taiwan Semiconductor Manufacturing Company (TSMC) 0.18 μm N-WELL process. Fig. 3 shows the transistor-level schematic of a conventional rail-to-rail CFC OTA. The rail-to-rail input operation is achieved using both N-input (M1-M2) and P-input (M3-M4) differential pairs at the input-stage. The gain-stage (or the output-stage) consists of a wide-swing current mirror (M5-M8), two cascode transistors (M9-M10), and two current sources (M11-M12). The wide-swing current mirror is used to achieve high output resistance and wide output voltage swing.

The opamp shown in Fig. 3 suffers from varying amplifier transconductance, G_M, for varying input common-mode voltages, V_{ICM}, because both the differential input pairs operate simultaneously only over the intermediate input common-mode voltage range (i.e., approximately for $V_{DD}/3 < V_{ICM} < 2V_{DD}/3$). Over the remainder of the input common-mode voltage range, only one of the two differential pairs will be operational. Thus, G_M drops to half, assuming that the two differential input pairs have the same transconductances.

The problem of varying amplifier transconductance can be solved by incorporating a transconductance-control circuitry at the input-stage, as explained in more detail in [8]. The circuit also suffers from dramatic decrease in gain over the upper input common-mode voltage range (i.e., approximately for $2V_{DD}/3 < V_{ICM} < V_{DD}$), as the P-input pair transistors switch the operating mode from saturation, to triode, to cutoff. The decrease in gain is due to the increase in the value of the channel-length modulation parameter, λ, of transistors M5 and M6, which decreases the output resistance of the opamp. The novel current control circuitry, shown in Fig. 4, provides a mechanism for solving this problem.

B. Design Analysis

The opamp of Fig. 3 experiences high reduction in gain over the upper input common-mode voltage range. This reduction is due to the increase in the biasing currents of the gain-stage transistors as the P-input pair transistors change their operating mode from saturation, to triode, to cutoff. The added current control circuitry in Fig. 4, however, ensures a constant biasing current through the gain-stage transistors over the intermediate and upper input common-mode voltage range.

The current control circuitry uses M18 to duplicate the operation of the P-input pair transistors. The drain current of M18, which is identical to the drain current of M3 and M4, is mirrored twice using the two current mirrors, M19-M20 and M19-M22. The mirrored currents are then subtracted from the constant currents sourced by M21 and M23. The remainder of the currents after this subtraction is sourced to M11 and M12. Thus, the total current sourced to M11 and M12 by both the P-input pair transistors and the current control circuitry, remains

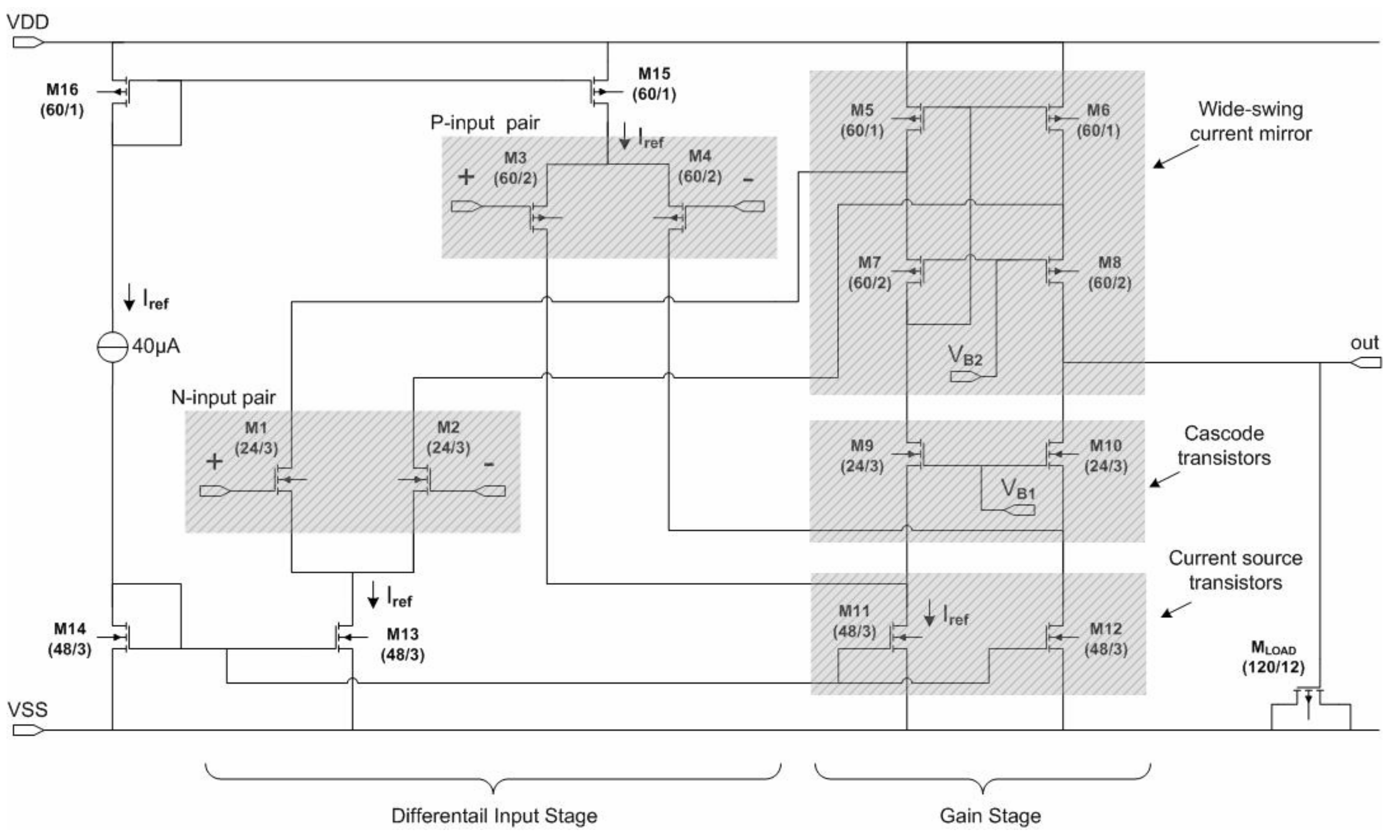


Fig. 3. The circuit schematic of a conventional single-stage CFC opamp.

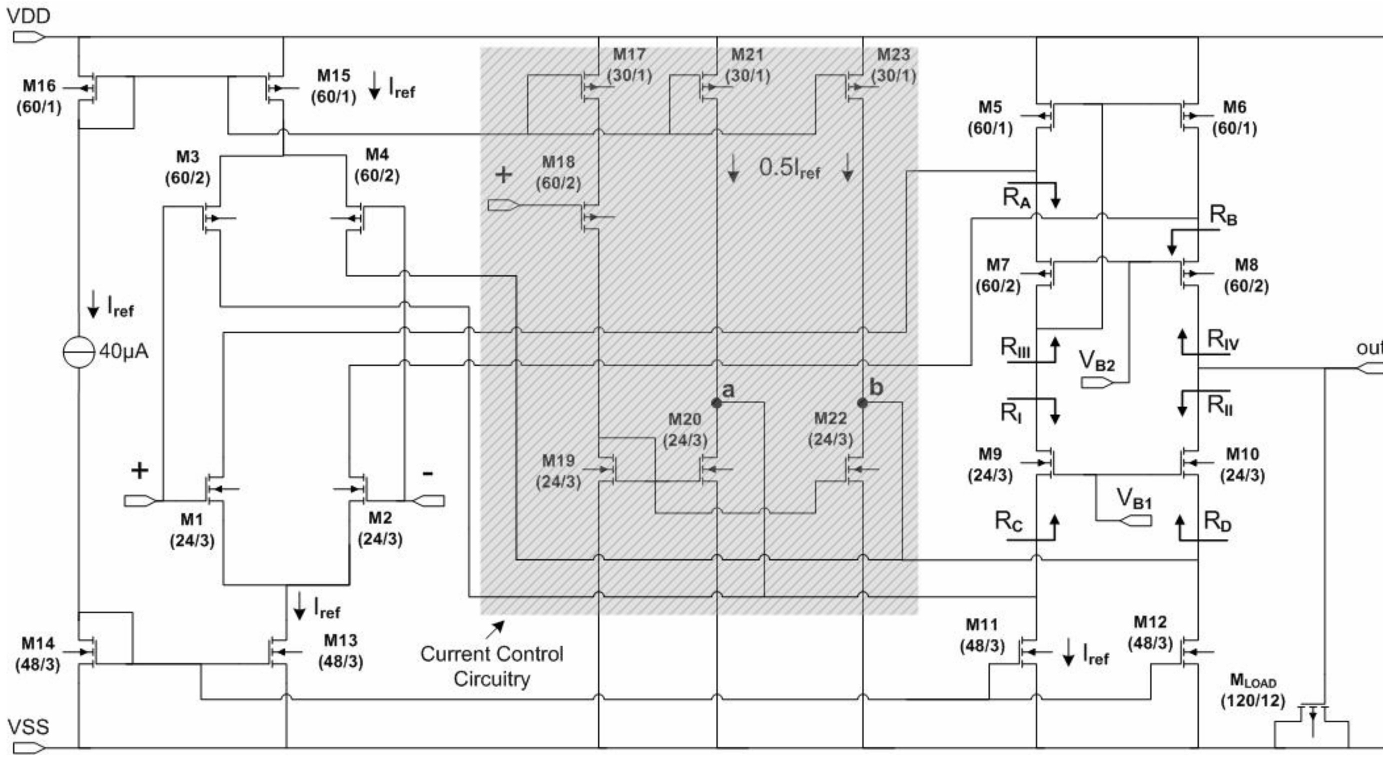


Fig. 4. The circuit schematic of a CFC opamp with the current control circuitry.

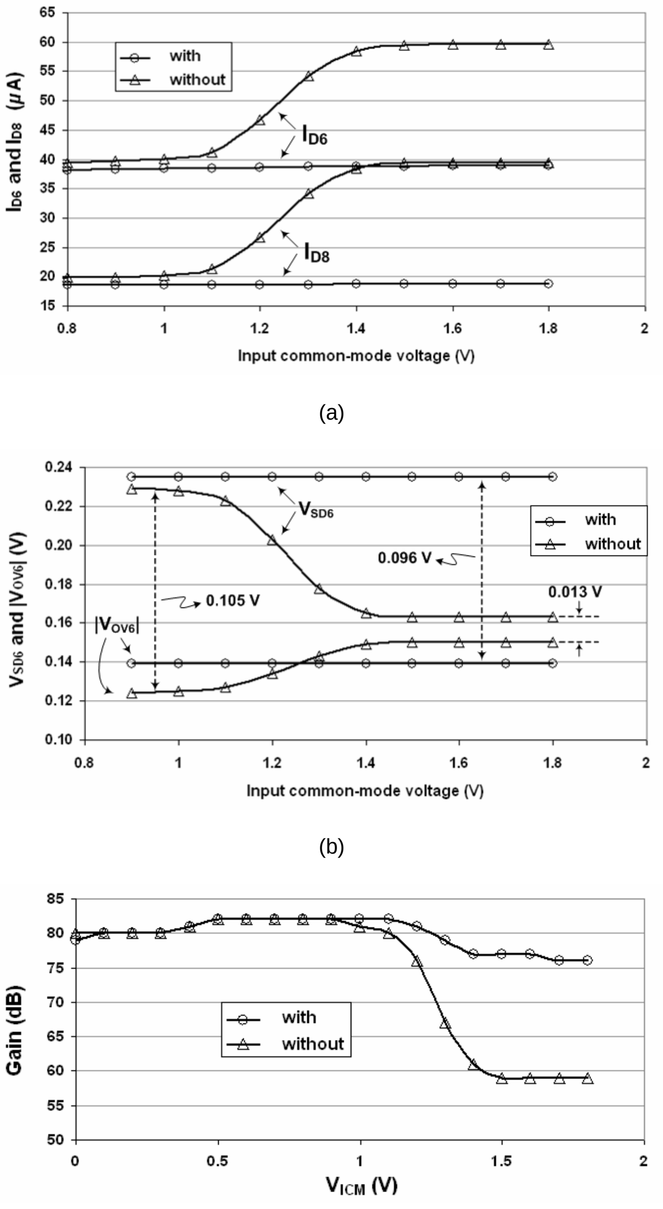
constant over the intermediate and upper input common-mode voltage range, and it is equal to $0.5I_{\text{REF}}$.

For the opamp of Fig. 4, the biasing currents through M6 and M8 is relatively constant in Fig. 5(a), the difference between V_{SD6} and $|V_{\text{OV6}}|$ is remained unchanged in Fig. 5(b), and the high reduction in gain is eliminated in Fig. 5(c). However, the varying opamp transconductance, G_M, for varying input common-mode voltages, V_{ICM}, is still reducing the gain when the N-input/P-input pair transistors switch off.

C. Noise Analysis

The low-frequency small-signal model for the CFC opamp is shown in Fig. 6 [9]. The resistances designated as R_A, R_B, R_C, and R_D are the ones looking into the source of M7, M8, M9, and M10, respectively

$$R_A = \frac{g_{m9} r_{ds9} r_{ds11}}{g_{m7} r_{ds7}}, R_B = \frac{g_{m10} r_{ds10} r_{ds12}}{g_{m8} r_{ds8}}, \quad R_C = \frac{1}{g_{m9}}, R_D = \frac{g_{m8} r_{ds8} r_{ds6}}{g_{m10} r_{ds10}}. \tag{2}$$



(a)

(b)

(c)

Fig. 5. (a) The biasing currents of M6 and M8 versus the input common-mode voltage with and without the current control circuitry, (b) The source-to-drain voltage and the overdrive voltage of M6 versus the input common-mode voltage with and without the current control circuitry, and (c) The differential-input voltage gain of the CFC opamp versus the input common-mode voltage with the without the current control circuitry.

Also, the resistances designated as R_I, R_{II}, R_{III}, and R_{IV} are the ones looking into the drain of M9, M10, M7, and M8, respectively

$$R_I = g_{m9} r_{ds9} r_{ds11}, R_{II} = g_{m10} r_{ds10} r_{ds12}, \quad R_{III} = \frac{1}{g_{m5}}, R_{IV} = g_{m8} r_{ds8} r_{ds6}. \tag{3}$$

The small-signal voltage-transfer function of the CFC opamp of Fig. 4 can be found as follows. The small-signal currents, i_5 and i_8, in Fig. 6(a) are written as

$$i_5 = \frac{g_{m1} v_{in}}{2} \left[\frac{r_{ds1} \| R_A}{\left(r_{ds1} \| R_A \right) + g_{m5}^{-1}} \right] \approx \frac{g_{m1} v_{in}}{2}, \tag{4}$$

and

$$i_8 = \frac{g_{m1} v_{in}}{2} \left[\frac{r_{ds2} \| r_{ds6}}{\left(r_{ds2} \| r_{ds6} \right) + R_B} \right] = \frac{g_{m1} v_{in}}{2(1 + k_1)}, \tag{5}$$

where k_1 is a low-frequency unbalance factor

$$k_1 = \frac{R_B}{r_{ds2} \| r_{ds6}}. \tag{6}$$

Also, the small-signal currents, i_9 and i_{10}, in Fig. 6(b) are given as

$$i_9 = \frac{g_{m3} v_{in}}{2} \left[\frac{r_{ds3} \| r_{ds11}}{\left(r_{ds3} \| r_{ds11} \right) + R_C} \right] \approx \frac{g_{m3} v_{in}}{2}, \tag{7}$$

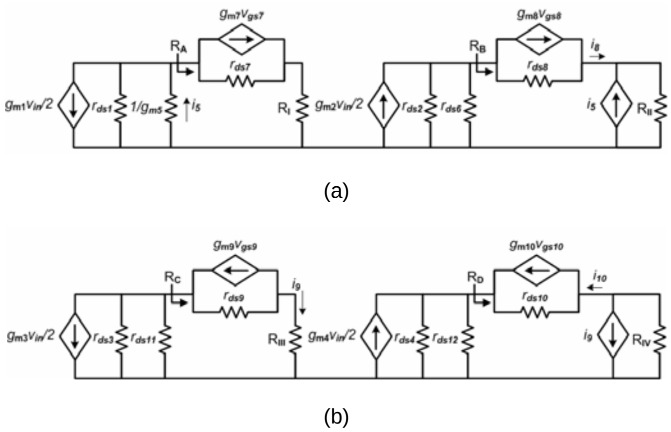
and

$$i_{10} = \frac{g_{m3} v_{in}}{2(1 + k_2)}, \tag{8}$$

where k_2 is defined as

$$k_2 = \frac{R_D}{r_{ds4} \| r_{ds12}}. \tag{9}$$

The small-signal output voltage, v_{out}, at the intermediate input common-mode voltage is equal to the sum of the currents i_5, i_8, i_9, and i_{10}, flowing through the opamp's small-



(a)

(b)

Fig. 6. The small-signal model of the CFC opamp when only (a) the N-input pair is operating, and (b) the P-input pair is operating.

signal output resistance, r_{out}. Thus, the voltage-transfer function is

$$\left(\frac{v_{out}}{v_{in}}\right)_{@V_{ICM}=V_{DD}/2} = \left[\left(\frac{2+k_1}{2+2k_1}\right)g_{m1}+\left(\frac{2+k_2}{2+2k_2}\right)g_{m3}\right]r_{out}\,, \tag{10}$$

where r_{out} is written as

$$r_{out} \approx \left[g_{m10}r_{ds10}\left(r_{ds12}\|r_{ds4}\right)\right]\|\left[g_{m8}r_{ds8}\left(r_{ds6}\|r_{ds2}\right)\right]. \tag{11}$$

The output current-noise density, $i_{no}^2(f)$, for the opamp of Fig. 4, is estimated as [10]

$$i_{no}^2(f)=\sum_{k=1}^{23}g_{mk}^2e_{nk}^2(f). \tag{12}$$

Note that $e_{nk}^2(f)$ is the mean-square voltage-noise of a MOS transistor

$$e_{nk}^2(f)=\left[e_{nk}^2(f)\right]_{thermal}+\left[e_{nk}^2(f)\right]_{flic\,ker} = 4kT\frac{2}{3}\frac{1}{g_{mk}}+\frac{KF}{2fC_{ox}\left(WLK'\right)_k}\,, \tag{13}$$

where k is the Boltzmann constant, T indicates temperature, C_{ox} is the capacitance per unit area of the gate oxide, KF is the flicker noise coefficient, and K' is known as the process transconductance for an nfet/pfet of a given CMOS technology. The output current-noise density can also be expressed in terms of the equivalent input-referred voltage-noise density, $e_{ni}^2(f)$

$$i_{no}^2(f)=\left(\frac{2+k_1}{2+2k_1}+\frac{2+k_2}{2+2k_2}\right)^2 g_{m1}^2e_{ni}^2(f). \tag{14}$$

Here we are assuming that the transconductance of the differential input pairs are the same and equal to g_{m1}. Thus, from eqn. (12) and eqn. (14), $e_{ni}^2(f)$ becomes

$$e_{ni}^2(f)=\left(\frac{2+k_1}{2+2k_1}+\frac{2+k_2}{2+2k_2}\right)^{-2} g_{m1}^{-2}\sum_{k=1}^{23}g_{mk}^2e_{nk}^2(f). \tag{15}$$

Figure 7 depicts the theoretical and extraction-based simulation results of the input-referred voltage-noise density for the opamp of Fig. 4.

IV. QUANTUM COUNTING PIXEL PROTOTYPE

The circuit schematic of the prototype is shown in Fig. 8. The PIN photodiode is reverse biased to create a wide depletion region. When an x-ray photon is incident on the radiation detector, electric charge is generated in proportion to the photon's energy. The charge is then integrated using the charge amplifier, and the output signal is amplified and filtered. Two comparators are utilized to form the energy window. The microcontroller initially operates in input-capture mode, and detects any transition at the output of the two comparators. In detection mode, any transition at the output of the first comparator (cmp1) increments the counter value, and any transition at the output of the second comparator (cmp2) decrements the counter value. In readout mode, the microcontroller uses the built-in serial communication interface (SCI) to serially shift counter values out.

The circuit of Fig. 8 was tested using both QSE773 silicon PIN photodiode and CdZnTe photoconductor. A polyenergetic 120 kVp x-ray source was continuously operated at 2.5 mA, and the results of individual photon detection are shown in Fig. 9. The silicon PIN photodiode was also irradiated with 600 keV gamma-ray photons. Figure 10 shows the output of

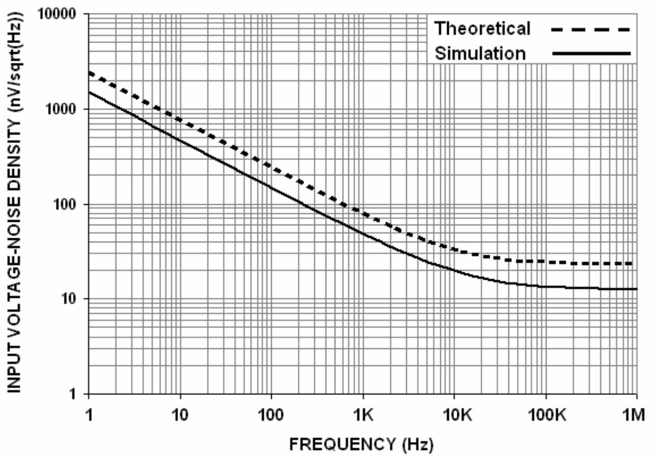


Fig. 7. The input-referred voltage noise spectral density of the CFC opamp with the current control circuitry.

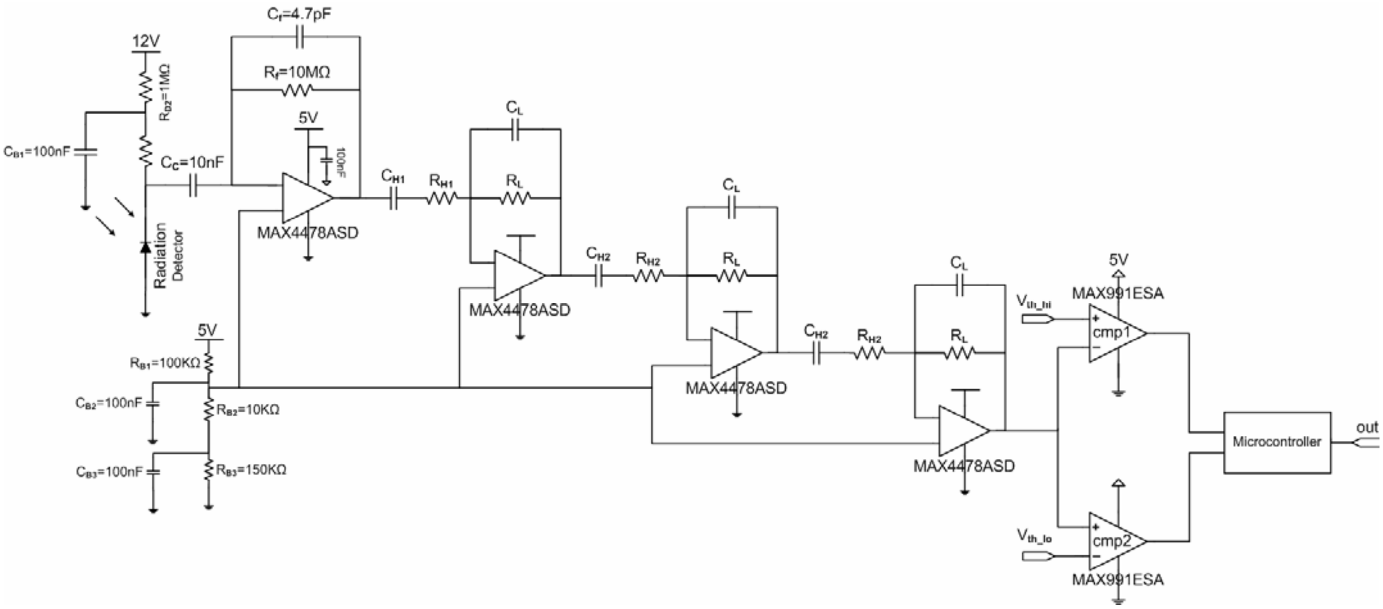
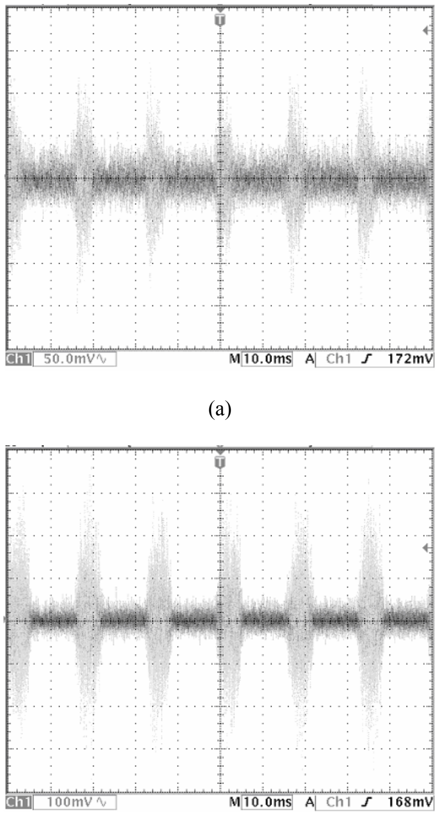


Fig. 8. The circuit schematic of the quantum counting pixel prototype.



(a)

Fig. 9. The output of the filter during x-ray exposure using (a) CdZnTe photoconductor, and (b) QSE773 silicon PIN photodiode.

the filter when a 600 keV photon was incident on the photodiode.

V. Conclusion

We have presented a pixel architecture for a novel quantum counting system with an intelligent decision-making unit to achieve energy windowing and false-count elimination. A low voltage, low noise, and high gain CFC opamp was designed for the pixel's charge amplifier. For the gain-stage transistors operating close to the edge of saturation, their channel length modulation parameters are extremely sensitive to biasing current variations. This sensitivity reduced the gain of the opamp in Fig. 3, when the biasing currents through M6 and M8 increased by $0.5I_{\mathrm{REF}}$. However, the added current control circuitry in Fig. 4 provided a constant biasing current through the gain-stage transistors, and ensured a high gain over the entire input common-mode voltage range.

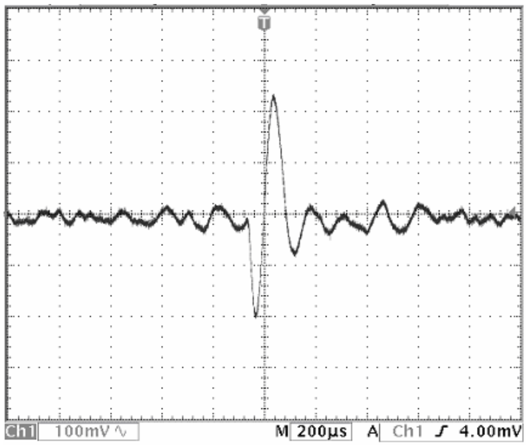


Fig. 10. The output of the filter when a 600 keV photon is incident on the QSE773 silicon PIN photodiode.

We also presented the circuit schematic for the quantum counting pixel prototype. The experimental results showed successful quantum counting operation for photons with energies higher than 50 keV. The work presented in this research has the potential to expedite the development of quantum counting imagers for digital tomosynthesis.

ACKNOWLEDGMENT

The authors thank Li Ng for assistance with the development of the quantum counting pixel prototype.

REFERENCES

[1] Mats Lundqvist, Björn Cederström, Valery Chmill, Mats Danielsson and Bruce Hasegawa, "Evaluation of a photon-counting X-ray imaging system," in *IEEE Trans. Nucl. Sci*., vol 48, no. 4, pp. 1530-1536 (Aug. 2001).

[2] Mats Danielsson, Hans Bornefalk, Björn Cedeström, Valery Chmill, Bruce Hasegawa, Mats Lundqvist, David Nygren, and Tamas Tabár, "Dose-Efficient System for Digital Mammography," in *Proc. SPIE*, vol. 3977, pp. 239-249 (2000).

[3] B. Mikulec, M. Campbell, E. Hiejne, X. Lopart, L. Tlustos, "X-ray Imaging Using Single Photon Processing with Semiconductor Pixel Detectors," in *Nuclear Instruments & Methods in Physics Research A*, Vol. 511, 282-286 (2003).

[4] Lukas Tlustos, Michael Campbell, Erik H.M. Heijne, Xavier Llopart, and Medipix2 Collaboration, "Imaging by photon counting with 256x256 pixel matrix," in *High-Energy Detectors in Astronomy 2004, Proc. SPIE*, Vol. 5501, 78-88 (2004).

[5] Dylan C. Hunt, Sean S. Kirby, and J. A. Rowlands, "X-ray imaging with amorphous selenium: X-ray to charge conversion gain and avalanche multiplication gain," in *Medical Physics*, vol. 29, issue 11, pp. 2464-2471 (Nov. 2002).

[6] James G. Mainprize, Nancy L. Ford, "A CdZnTe slot-scanned detector for digital mammography," in *American Association of Physics in Medicine*, 2767-2781 (November 2002).

[7] James G. Mainprize, Nancy L. Ford, Shi Yin, Tümay Tümer, Eli Gordon, William J. Hamilton, Martin J. Yaffe, "Semiconductor Materials for Digital Mammography," in *Proc. SPIE*, vol. 3977 (2002).

[8] Ron Hogervorst, Johan H. Huijsing, "Design of low-voltage, low-power operational amplifier cells," Kluwer Academic Publishers, 1996, ch. 3.

[9] Phillip E. Allen, Douglas R. Holberg, "CMOS analog circuit design," New York: Oxford University Press, 2002, ch. 6, pp. 302-305.

[10] Amir H. Goldan, Karim S. Karim, John A. Rowlands, "Selective photon counter for digital x-ray mammography tomosynthesis," in *Proc. SPIE*, vol. 6142, 61425B, Mar. 2006.

Amir H. Goldan is pursuing the M.A.Sc. degree at the University of Simon Fraser, Burnaby, BC, Canada.

He has received his bachelor of applied science in computer engineering with a minor in computing science at Simon Fraser University. His research interests include analog and digital circuits, and the development of photon counting imagers for digital tomosynthesis.

Mr. Goldan has received the Natural Sciences and Engineering Research Council (NSERC) of Canada Postgraduate Scholarship and was the recipient of the Honorable Mention Poster Award at the SPIE Medical Imaging Symposium in 2006.

Karim S. Karim received the Ph.D. degree in electrical engineering at the University of Waterloo, Waterloo, ON, Canada.

His research interests include circuit, device, and process development for biomedical imaging applications. Currently, his research focus is on amplified pixel architectures for large-area diagnostic medical X-ray imaging.

Dr. Karim has received two national Natural Sciences and Engineering Research Council (NSERC) of Canada Postgraduate Scholarships and was the recipient of the Michael B. Merickel Best Student Paper award at the SPIE Medical Imaging Symposium in 2001.

A Review of Different Distorted View Algorithms

Mohammad Yasser Khan, *School of Interactive Arts & Technology, Simon Fraser University, Canada*

Abstract—**This paper presents a review of approaches towards visualization of large information spaces, which are broadly categorized as distorted view methods. Distorted view methods visualize data by distorting the view, either by magnifying the areas of interest or by de-magnifying areas which are not of interest. Thus, distorted view methods provide focus and context at the same time. This technique works on the principle of providing more space for important points and reducing that of the unimportant ones.**

Index Terms— **Bifocal display, continuous zoom, fisheye view, visualization.**

I. Introduction

VISUALIZING complex and large information spaces has been a challenge to researchers, especially when these spaces are to be viewed on a small screen. The main concern is to represent this information in a meaningful way such that a user can to identify her-self/him-self within the context of this information.

"People tend to perceive the world using both local detail and global context…. Yet we rely on global context for orientation and to understand local detail" [1]. The authors [1] further state that advancement in the visual representation of large information spaces is governed by the importance of achieving detail-in-context [1]. According to Ware [2] "finding detail in a larger context" also termed as "focus-context problem" has "already been solved by human visual system" (p. 339). "The brain continuously integrates detailed information from successive fixations of the fovea with the less detailed information that is available at the periphery." [2] (p. 339). By exploiting human perceptual capabilities we can develop effective visualization approaches for providing detail-in-context.

Distorted view methods have been proposed for effective and meaningful representation of large information spaces: Bifocal Display [3], Continuous Zoom [1], Fisheye View [4], Perspective Wall [5] and Stretchable Rubber Sheet [6].

The *bifocal display* was originally conceived as an approach for visualization of one dimensional data space that could not be covered on a single screen and was one of the earliest distorted view methods [3]. In this approach, the detailed view is displayed at the center, with distorted views on the two sides.

Dill et al. presented a new distorted view method called *continuous zoom* [1]. This method was developed for the representation of hierarchical, two-dimensional data and shows both focus and context at the same time. This method not only provides a solution for the problems encountered with traditional visualization approaches, but also has advantages over the other distorted view methods.

The *fisheye view* was designed for the representation of hierarchical data [4]. In this method, the information that is more relevant is always displayed in much greater detail as compared to less relevant information. The relative relevance of information is determined by calculating the degree of interest (DOI) and the distance from the current area of focus. Many variants of the Fisheye View have been developed, and these are commonly used in the visualization of large information spaces.

The *perspective wall* was conceptualized as a variant of the Bifocal Display [5]. It was designed for one-dimensional data, to display a view that was a uniform integration of both detail and context. In this case, the distorted view on the sides directly depends on their distance from the viewer.

II. Bifocal Display

The *bifocal display* [3] was one of the earliest distorted view methods. The *bifocal display* was originally conceived as an approach for visualization of one dimensional data space that could not be covered in a single screen. In this approach, the detailed view is displayed at the center, while the distorted views are displayed on the two sides.

The Bifocal display is based on the principal of spatial information management. Here, all the focus is placed on the area of interest. The algorithm [3] divides the screen into different sections. The area of interest is placed at the center of the screen and is displayed in full detail. The sections of the screen outside the center are either compressed or display images of lower quality.

An important feature of this algorithm is that the area of interest can be changed. When any other section is selected or gets focus, the selected section is shifted to the center of the screen and is shown in a detailed view.

The authors [3] point out that in this algorithm the screen is usually divided into three sections, with the middle section having the focus and the detailed view.

Bifocal display is achieved with the help of two mathematical functions: a transformation function and a magnification or a demagnification function. The transformation function determines the manner in which an undistorted image would be represented in the distorted view. On the other hand the role of magnification or a demagnification function is limited to deciding how this magnification and demagnification will take place. This function is applied over the entire area of the image to achieve the desired result.

Despite all its advantages *bifocal display* has a major drawback. This approach fails to maintain the continuity of magnification at the periphery of the magnified and the de-magnified views.

III. The Continuous Zoom

The *continuous zoom* (CZ) [1] presents an approach for the visualization of a 2-dimensional hierarchical data. Dill et al. [1] state, "The continuous zoom displays a network in a rectangular 2-D display space by recursively breaking it up into smaller rectangular areas, creating a hierarchy of nested rectangles." This method provides great deal of flexibility to the users [1]. Figure 1 shows an application CZWeb [4], based on CZ algorithm.

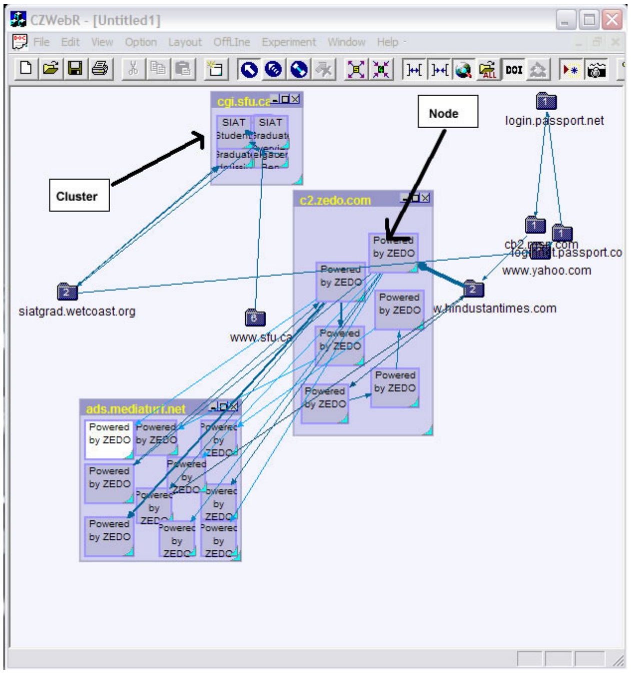


Figure 1. CZWeb application

The *continuous zoom* algorithm allows users to expand and collapse the clusters of information, thereby allowing the users to decide whether or not they want to view the detailed hierarchical view of the information at any given time (see figure 1 above). A user can decide the measure of detail that would be displayed on the screen by simply opening, closing, resizing or reorganizing any of the clusters or the nodes [1]. Additionally, a user can perform automated resizing of both, the clusters and the nodes.

"When a cluster changes size, its contents are resized accordingly. Whenever a node shrinks, it gives up display space to siblings so that they may grow" [1]. By controlling various aspects of the clusters, such as opening them to get a detailed view and closing them to save space, a user can always maintain detail-in-context. Furthermore, in this algorithm the space is allocated to the open clusters on the basis of the details contained in them. Dill et al. [1] add that. "As the nodes shift around on the display, the link vertices are transformed with them, preserving adjacency". Due to the above mentioned reasons, a user can administer the entire display and can easily decide on the details to be shown.

Dill et al. [1] state that, "The algorithm uses a 'budgeting' process to distribute space among nodes of a network. It calculates the amount of space requested from each node and then distributes the fixed overall space budget according to the size of their requests". Figure 2 represents the working of CZ algorithm.

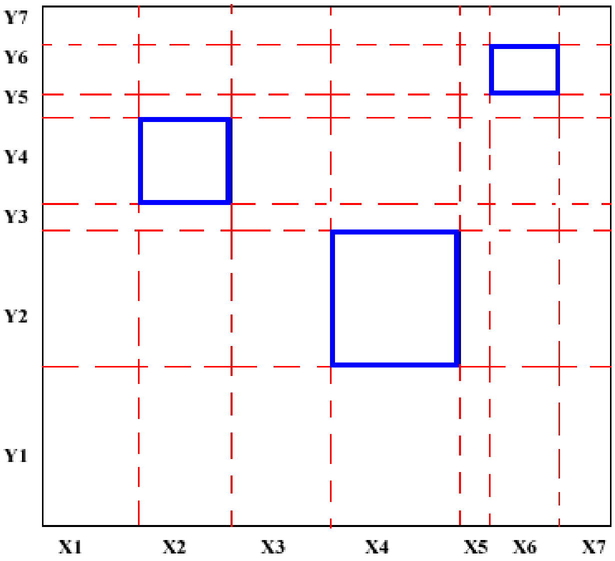


Figure 2. Screen is divided by X and Y intercepts in CZ algorithm.

Intercepts on the X and Y-axes are determined by mapping the node terminals on both of the axes (see figure 2. above). Moreover, the screen or window is divided into intervals, based on successive intercepts on both the X and Y-axes. This concept of intervals is the core of the continuous zoom algorithm. As Dill et al. [1] point out "Since by definition, intervals never overlap, the North-South and East-West geometric relationships between sibling nodes are always maintained".

Using CZ we can zoom in/out the nodes and the clusters by multiplying the intervals with their respective scale factors, and then, by repositioning the nodes within their corresponding containing intervals. This is achieved by placing the nodes within their corresponding intervals by their center points and then rescaling them to fit the interval size. "A (multiplicative) scale factor is defined for each node to specify its size change (since the algorithm works independently in X and Y axis, a separate scale for each axis is needed)" [1].

Another important feature of this algorithm is that we don't have to explicitly adjust the size of the containing cluster once the nodes within it are modified. The authors [1] point out, that in this algorithm uniform change in the size of the child nodes result in corresponding change in the size of the containing clusters. In this algorithm the increase in size of a node can result in a uniform increase in its parent's size to accommodate the change, and the changes can make their way to the root. By specifying a scale factor for the nodes, both in x and y-axes, we can continuously increase or decrease their size. Dill et al. [1] recommend that gap intervals between the nodes should be multiplied by a scale factor equal to the "maximum scale of its neighboring intervals" [1], so that the gaps grow along with the nodes.

The authors [1] have also defined a hybrid variant of the continuous zoom. This variant of continuous zoom algorithm contains favorable aspects of both local and global zoom. The authors [1] state this continuous zoom is superior to the other methods owing to the fact that "continuous zoom network viewing method provides multiple focus points in context and flexible control over node size" [1].

IV. Fisheye Views

An ever-increasing number of applications and programs now generate massive structures of information, and this information is usually displayed on small displays. As a consequence, it has become relatively difficult for a common user to retain her/his sense of comprehension while analyzing this data. The problem occurs as local details require being presented within their global contexts, the lack of which results in the feeling of being lost [4].

Furnas [4] state that, "humans often represent their own "neighborhood" in great detail, yet only major landmarks further away". By presenting the local details along with the small but important global context, much of this problem can be solved. There are several techniques that can be used to achieve this goal. Furnas [4] present a new method called *Fisheye's view*, which makes it possible to show localized detail in its proper context or with respect to the surrounding world. This is achieved by displaying the remote regions in lesser details and giving an in-depth view of the nearby regions and, also by maintaining a balance between the detail and the context.

The author [4] believed that *fisheye view* is inherently used by humans and thus, can be used to build better interfaces. Learning how complex structured information is represented in human brain can facilitate development of better techniques for representation of such information. Furnas [4] conducted an experiment to test the effectiveness of the *fisheye view* and to investigate the additional features that can be added to the *fisheye view* based interfaces.

Furnas [4] provided the subjects with a problem and asked them to list 10 most closely related answers to the problem. Author [4] hoped that the results would validate the fisheye approach, as subjects would either point towards information that is either very important or is more close to them. Author [4] examined people belonging to the different groups such as academicians, employees of a corporation etc. and found that the results verified the Fisheye View, as the subjects had an in-depth understanding of local details, and an understanding of only significant features, when it comes to a global context.

The results verified that Fisheye View was a common phenomenon, often used by people in their daily understanding of things around them. Furnas [4] argue that due to this fact the *fisheye view* technique would prove really effective in interface design.

Furnas [4] provided a formal definition to the Fisheye View method. In this technique the interest of a user is defined by assigning a DOI (degree of interest) to the concerned information. And based on this DOI, information can be organized. By displaying the most interesting information, the problem with the small displays can be considerably reduced, as the users will get a view containing information of high relevance. Furnas [4] further points out that the success of a given display would depend on the choice of DOI.

Furnas [4] state, "Generalized fisheye views arise by decomposing the DOI into two components: *a priori* importance and distance". According to the author [4], "the interest increases with *a priori* importance and decreases with distance". Author [4] states that the success of this algorithm lies in the tradeoff between *a priori* interest and the distance. Moreover, Furnas [4] argues that this method is highly suitable for lists, trees, acyclic graphs etc., owing to the fact that the definition of the Fisheye View "allows interfaces to be defined and explored in any structure where distance and some display-relevant notion of *a priori* importance can be defined". The output of the *fisheye view* is not restricted to a graphical view and can also exist as a natural language text.

Furnas [4] carried out another test to verify the effectiveness of the Fisheye View method. The aim of this experiment was to determine whether the Fisheye View can be used to examine the unexplored parts of large files or documents. Furnas [4] employed 20 participants for this study. The participants were asked to navigate through an unfamiliar hierarchical structure. The participants were also asked to identify relative positions for two different parts of the structure. The purpose of this experiment was to investigate the support for cognitive tasks, as the user proceeded towards the intended target in the structure. The experiment proved that the *fisheye view* was more accurate than other views, as it generated the required structure without losing its context.

V. Conclusion

In this paper I have presented a literature review of the various distorted view methods for visualization of large information spaces. I have provided a description of the Bifocal Display [1], the Continuous Zoom [3], and the Fisheye View

[4]. These distorted view methods play a very important role in meaningful visualization of the complex information spaces.

Many approaches have been proposed for effective and meaningful representation of large information spaces, but the distorted view methods are considered better then the rest owing to their relative advantages over the others techniques. They not only provide the ability to pan/zoom into the data but also provide the context. Thus, a user is able to get a detailed local view along with the overall global context. Non-distortion oriented methods usually only provide simple pan and zoom operations and thus when users zoom into a section of data they loose the surrounding context. Some non-distortion oriented techniques also provide ability to have multiple windows to provide some context. But switching between the windows to obtain context and the detailed view limit a user's ability to relate the both the views properly. Another approach called Map view provides a rectangular area that can be moved on the screen. This rectangular area displays the detailed view, but at the same time limits a user's ability to comprehend the information as it requires mentally figuring out the proper context. The distorted view methods provide a solution for such limitations.

REFERENCES

[1] Dill, J. Bartram, L., Ho, A., and Henigman, F. A continuously variable zoom for navigating large hierarchical networks. Proceedings of the 1994 Conference on Systems, Man and Cybernetics, pages 386-390, 1994.

[2] Ware C. (2004), *Information Visualization: Perception for Design*, Morgan Kaufmann Publishing, pages 317-350.

[3] Apperley M. D., Tzavaras I. and Spence R. (1982), "A Bifocal Display Technique for Data Presentation", Eurographics'82 Proceedings, 27-43.

[4] Collaud, Gé rald, John Dill, Christopher Jones and Paul Tan The Continuously Zoomed Web -a Graphical Navigation Aid for WWW. Proc. IEEE Visualization '96, Late Breaking Hot Topics, pp.1-3, 1996.

[5] Furnas, G.W. "Generalized Fisheye Views." *Proc. ACM SIGCHI '86 Conference on Human Factors in Computing Systems*, pp. 16-12, Apr. 1986.

[6] Mackinlay J. D., Robertson G.G. and Card S. K. (1991), "The perspective wall: Detail and context smoothly integrated", Proceedings of ACM CHI'91 Conference, 173-179.

[7] Sarkar M. and Brown M.H. (1994), "Graphical Fisheye Views" Communications of the ACM 37:12, 73-84.

[8] Ware C. (2004), *Information Visualization: Perception for Design*, Morgan Kaufmann Publishing

Mohammad Yasser Khan received Bachelor of Engineering degree in Computer Science & Engineering in 2005 from Rajiv Gandhi Technical University, Bhopal, MP, INDIA.

Currently, he is pursuing M.Sc. in Interactive Arts & Technology form School of Interactive Arts & Technology, Simon Fraser University, BC, Canada. He is working in field of Information Visualization and Visual Analytics.

Analysis of user behavior in a hybrid satellite-terrestrial network

Savio Lau and Ljiljana Trajković

***Abstract*—Satellite data networks have received much attention, due to their capabilities in providing broadband access for areas not served by traditional broadband technologies. In this paper, we describe measurements of traffic data from a satellite Internet service provider, including a set of billing records and a set of *tcpdump* traces. From the billing records, we investigate the user behavior with respect to uploaded and downloaded traffic volume. The daily and weekly cycles exhibited in the data are examined, as well as holiday's effects on traffic patterns. Analysis of the *tcpdump* traces found the majority of the data traffic to use the TCP transport protocol and we compared the TCP options recorded in the trace with the recommended practices for the satellite environment. Lastly, we present anomalies in the captured traffic, which includes invalid TCP flag combinations and port scans.**

***Index Terms*—Satellite-terrestrial networks, TCP options, user behavior, traffic measurements.**

I. Introduction

DEMAND for broadband Internet access has continued to grow during the past decade and results in the growth in traffic volume, development of new protocols, and development of new access technologies. For a given network, network traffic measurements are used to characterize workloads and evaluate network performance. These measurements allow the detection of changing data traffic dynamics, as well as the proposal of new models that accurately describe the types of data traffic present in the networks. Thus, measurement and analysis of genuine network traffic traces are an important and ongoing task.

In the past decade, researchers have collected and characterized terrestrial Internet traffic [1], [2]. In addition, some of the collected Internet traffic traces have been made available for public analysis [3]. However, the majority of these data are collected from university campuses or research institutions. Conversely, few traces are collected from commercial networks, especially from wireless and satellite environments. In this paper, we describe our measurements from a hybrid satellite-terrestrial network. This network is operated by ChinaSat, a commercial satellite Internet service provider in China that provides broadband access to through the DirecPC system. We analyze the patterns and statistical properties of the collected traffic data in order to understand the network users' behaviors.

This paper is organized as follows: in Section II and III, we describe the DirecPC system and the techniques used in satellite environments to improve performance. The collection of data is described in Section IV. The analysis of billing records and the analysis of the *tcpdump* traces are presented in Section V and VI, respectively. We conclude with Section VII.

II. DirecPC system

Satellite systems broadcast information over a large geographical area, and solve the last mile access problem for less assessable areas. One satellite system, DirecPC, is an asymmetric satellite system deployed by Hughes Network System, through a constellation of geosynchronous satellites. This collection of satellites provides both television and data services, including: DirecTV, a satellite television service; DirecPC, a unidirectional satellite data service; and DirecWay, a bidirectional satellite data service that is replacing DirecPC. The Internet access component of DirecPC is called Turbo Internet. Turbo Internet provides broadband access through a satellite downlink and a return path through a terrestrial dialup modem. This service has an advertised rate of 400kbps for the downlink path.

Through the DirecPC system, ChinaSat provides Internet access to over 200 Internet cafés across Chinese provinces. In addition, ChinaSat also provides Internet access to individual users and businesses.

Since DirecPC works through geosynchronous satellites, which are 35,800 kilometers above the Earth, the long propagation delay (~250 ms) and the high bit error rate (BER) in the satellite link need to be addressed. For the DirecPC system, IP spoofing and TCP splitting are used to improve network performance. These two techniques will be described in detail in Section III.

When a user wishes to browse a website, the request is not sent directly. Instead, the DirecPC software installed on the user's computer adds a "tunneling header" to the request IP packet, and the packet is sent to the satellite Network Operations Center (NOC), using terrestrial dialup modem from a local Internet service provider. At the NOC, the tunneling header is removed and the request is forwarded to the website using a high-speed link. The NOC receives the reply from the website, and the reply is forwarded to the user's satellite receiver through the DirecPC satellite. The data paths of the

Manuscript received July 27, 2006. The authors are with the School of Engineering Science, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada (e-mail: {saviol, ljilja}@cs.sfu.ca).

DirecPC system are illustrated in Fig. 1. IP headers at the user and at the website, including the “tunneling header”, are shown in Fig. 2.

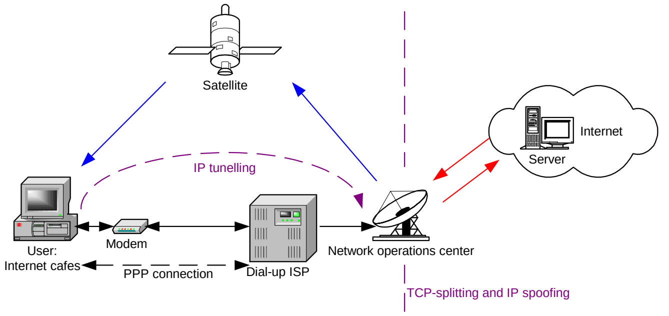


Fig. 1. Data path of the DirecPC system.

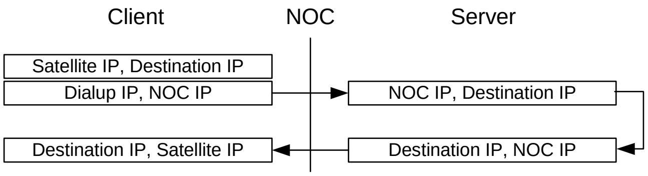


Fig. 2. IP headers used in the DirecPC system. Each box indicates a pair of source and destination IPs. Note the “tunneling header” sent by the client.

III. TCP EXTENSIONS FOR SATELLITE ENVIRONMENTS

TCP [4], the most common transport layer protocol on the Internet, was originally designed for terrestrial networks. However, satellite environments have two characteristics that are undesirable with respect to TCP: long propagation delays (~250 ms for each geosynchronous satellite link) and high bit error rates.

Long propagation delays along with large bandwidth results in large bandwidth-delay product. This bandwidth-delay product measures the amount of data required to be in transit (unacknowledged) in order to maximize the transfer rate between two connection endpoints. This maximum amount of unacknowledged data in TCP is determined by the TCP sliding window. Standard TCP uses a 16-bit field for the window size, which limits the receive windows size to 65,535 bytes, or 64 Kbytes. Given a roundtrip time value RTT, the maximum theoretical throughput of a link is given by the equation:

$$throughput = \frac{window\ size}{RTT}. \qquad (1)$$

For a network with roundtrip time of 400 ms, the maximum throughput would be 164 Kbytes per second or 1.31 Mbps.

High bit error rates have the effect of slowing the recovery of packets losses. Standard TCP, through the use of fast retransmit and fast recovery algorithms, can only correct one missing segment per round trip time. Additional missing segments will cause TCP to enter the slow start phase and throughput would suffer as a result [5], [6].

Proposals have been made to improve TCP in the satellite environment [7]. In addition, researchers have analyzed on TCP performance over satellite links using these proposals [8]. Some of the extensions specific to satellite environment are described in [5], which states the current best practice for TCP over satellites. These extensions include: increasing the TCP’s initial congestion window size [9], employing TCP’s sliding window scale option [6], using selective acknowledgements (SACK) [10], and sending path MTU discovery [11]. Although these extensions have been described as best current practice, not all are widely deployed. Further to the TCP extensions described in [5], performance enhancing proxies (PEPs) [12] have also been successfully deployed to improve TCP performance in satellite systems such as the DirecPC.

A. Increasing TCP’s initial congestion window size

Standard TCP avoids transmitting a large burst of traffic inappropriately through the slow start algorithm. This algorithm begins by setting the size of congestion window (cwnd) to one segment. During the TCP slow start phase, each received acknowledgement (ACK) will increase cwnd by one. For a long-delay network, the slow start algorithm will require a significant amount of time before the sending rate approaches the maximum throughput possible. Thus, the authors of [9] recommend setting the initial window to the size of roughly 4 Kbytes to increase the speed of congestion window growth. The improvement from increasing TCP’s initial window for a satellite environment is described in [13].

B. TCP’s sliding window scale option

Instead of standard TCP’s 16-bit sliding window, TCP’s sliding window scale option expands the field size to 32-bits. This is accomplished by specifying a scale factor in the handshaking SYN segments through an 8-bit field. The field value can be different for each direction. A TCP stack using the scale factor will right-shift the sliding window by the number of bits contained in the scale factor. The sliding window scale option improves the data throughput in satellite environments by eliminating the throughput limit imposed by the 16-bit sliding window [6].

C. Selective acknowledgments (SACK)

With the implementation of fast retransmit and fast recovery algorithms, standard TCP can recover at most one lost segment per RTT. However, in high BER and long delay environments, additional lost segments will result in transmission timeouts and the subsequent retransmission will use the slow start algorithm. As in the case of a small initial congestion window, the use of the slow start algorithm can be time consuming over satellite channels. One of the mitigation methods is to use selective acknowledgements (SACK) [10]. SACK allows a TCP receiver to explicitly specify the segments that have been received. When a sender receives notification of lost segment(s) through SACK, it can retransmit the lost segment(s) earlier and avoid the performance penalty associated with transmission timeouts and the slow start algorithm. In [13], Henderson et al. show the use of SACK improves TCP throughput when RTT is large.

D. *Path MTU discovery*

Path MTU discovery [11] is a technique used to determine the maximum packet size that can be supported in the links between two endpoints without the packet being subjected to IP fragmentation. This technique allows TCP to maximize the ratio of data bytes to overhead bytes. In addition, since the TCP congestion window is increased based on segments, larger segments enable TCP senders to increase congestion window more rapidly.

E. *Performance enhancing proxies*

One of the newer extensions to TCP is the performance enhancing proxies (PEPs). PEPs are a collection of techniques "employed to improve degraded TCP performance caused by characteristics of specific link environments" [12]. As Borders et al. [12] state, PEPs are not intended for general use, as its use results in an undesirable property of breaking the TCP end-to-end principle. With respect to the DirecPC system, two techniques are used for its PEP: TCP splitting and IP spoofing [14]-[18]. An example of TCP-splitting and IP spoofing is shown in Fig 3.

The NOC acts as the intermediary between a satellite user (client) and a website (server). It is at the NOC where the TCP connection is split. The 3-way TCP handshake (SYN, SYN/ACK, and ACK) works in a way identical to standard TCP connections. However, subsequent TCP segments from both endpoints are acknowledged by the NOC on behalf of the other endpoint, using a technique known as IP spoofing. These ACKs, shown in blue, long dash lines, are returned to the two endpoints more quickly in compared to the standard end-to-end connection. This allows TCP congestion window to grow faster, and results in improved performance. In addition, the normal ACKs transmitted by the two endpoints (short, red dashes) are not forwarded by the NOC.

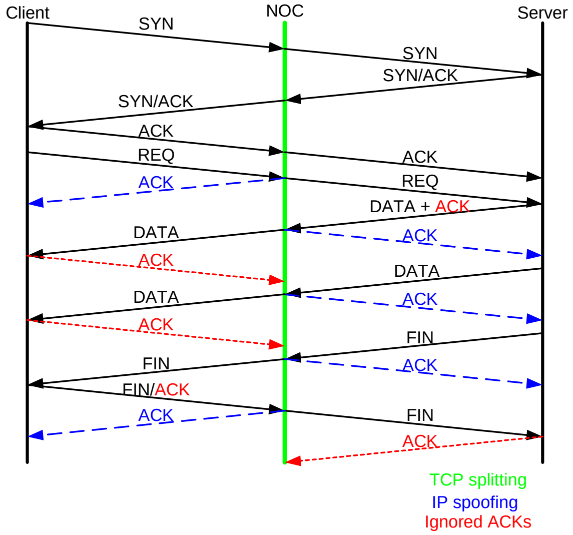


Fig. 3. TCP-splitting and IP spoofing. The TCP split occurs at the NOC (center vertical line). IP spoofed ACKs from the NOC are shown in blue, long dash lines. NOC ignores ACKs shown with red, short dash lines.

Although PEP with TCP splitting and IP spoofing improves performance, the technique imposes considerable memory requirements at the NOC. All segments prematurely acknowledged by the NOC must be kept in local buffers until segments are acknowledged by the endpoints (short, red dashes). As a result of IP spoofing, the NOC is also responsible for retransmitting all lost segments.

IV. Data Collection Method

The NOC is the ideal location to collect traffic traces because all satellite users' traffic is re-routed through the NOC with the use of PEP. Traffic traces were collected with the open-sourced network monitor program *tcpdump* using a Linux PC equipped with a 100 Base-T Ethernet adaptor and a high-resolution (100 µs) timer. *tcpdump* was configured to capture the first 68 bytes from each packet to ensure user privacy and to minimize storage requirements while preserving the IP and TCP headers for analysis.

The network access point for the trace collection was a port on the primary Cisco router at the NOC located in the Northwest rural area of Beijing, China. The router provided access to the inbound and outbound packets sent among the hosts using the 100 Mbps NOC's local area network (LAN). There is a 10 Mbps connection between the NOC and the Internet backbone. In addition to the *tcpdump* traffic traces, we have also obtained two months of billing records from the DirecPC system.

V. Analysis of Billing Records

We obtained billing records for a continuous period between Oct. 31, 2002 and Jan. 10, 2003. The records contain a collection of files generated very hour detailing the connection time, number of downloaded and uploaded packets, volume of downloaded and uploaded bytes and a hexadecimal ID for each active user during the recorded period. Hence, the billing records capture the network dynamics at an hourly level. In total, 1691 hours and 69 full days of record has been obtained.

A. *Hourly and daily traffic volume*

Figs. 4-11 show the aggregated downloaded and uploaded traffic volume in terms of bytes and packets for each hour and for each day, respectively. The downloaded traffic volume in bytes is larger than the uploaded traffic volume in bytes by an order of magnitude. More uploaded packets were captured compared to downloaded packets. This difference may be attributed to the contribution of User Datagram Protocol (UDP) packets.

Trends observed from Figs. 4-7 exhibit a regular pattern that repeats every 24 hours with an exception around 1300th hour. Around the 1300th hour, the daily minimum traffic volume is much higher compared to all other days. This day corresponds to Christmas Eve, Dec. 24, 2002. In Figs. 8 to 11, the maximum number of downloaded bytes is recorded on day 54, which again is Dec. 24, 2002. This behavior indicates that holidays change the dynamics of traffic.

In Figs. 4-7, around hour 1520, the traffic volume dropped to almost zero, followed by a large increase in traffic volume. This observation is noticeable in Figs. 6 and 7, where the packet traffic volume has the highest recorded value near hour 1530. The change in the traffic volume pattern during these hours was caused by a network outage followed by recovery, when the queued emails during the outage were processed.

Furthermore, we observed a drastic reduction in traffic volume between day 59 and day 70 in Figs. 8-11. These days corresponds to a period between Jan. 1, 2003 and Jan. 10, 2003. Although Jan. 1 is a public holiday in China, Jan. 2 to 10 are not public holidays. We believe the reduced volume is related to users who extend their holiday vacation into the first week of 2003.

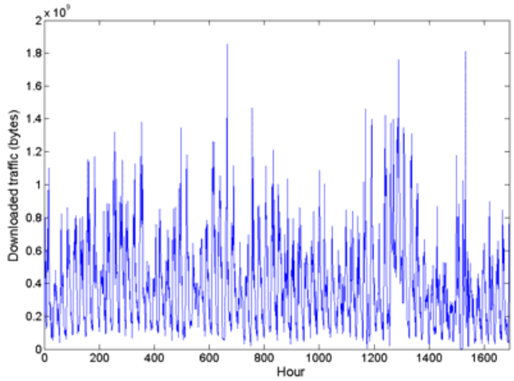


Fig. 4. Aggregated traffic volume per hour (downloaded bytes).

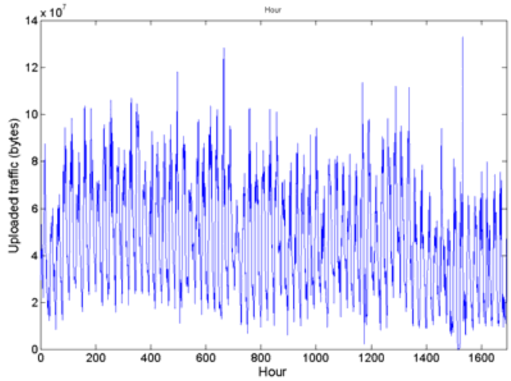


Fig. 5. Aggregated traffic volume per hour (uploaded bytes).

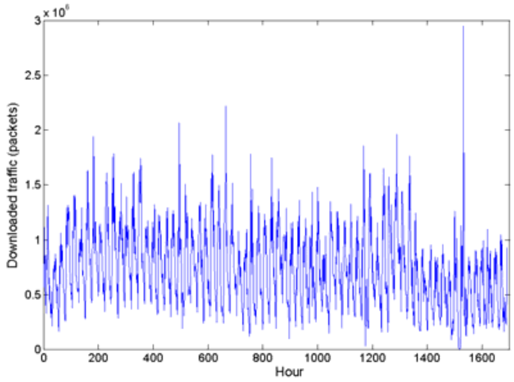


Fig. 6. Aggregated traffic volume per hour (downloaded packets)

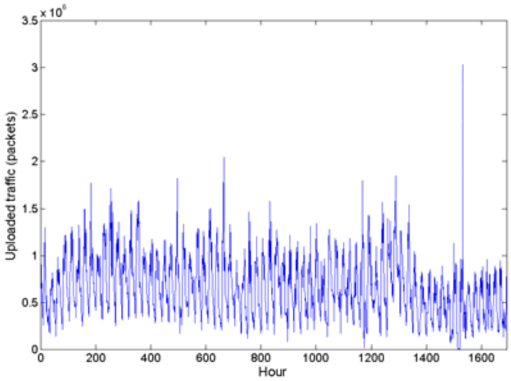


Fig. 7. Aggregated traffic volume per hour (uploaded packets)

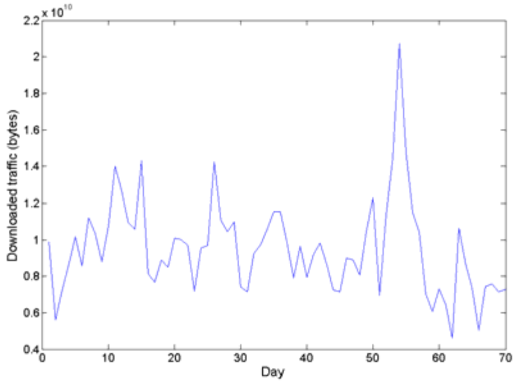


Fig. 8. Aggregated traffic volume per day (downloaded bytes)

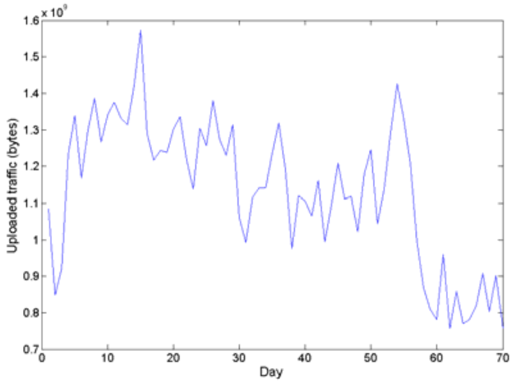


Fig. 9. Aggregated traffic volume per day (uploaded bytes)

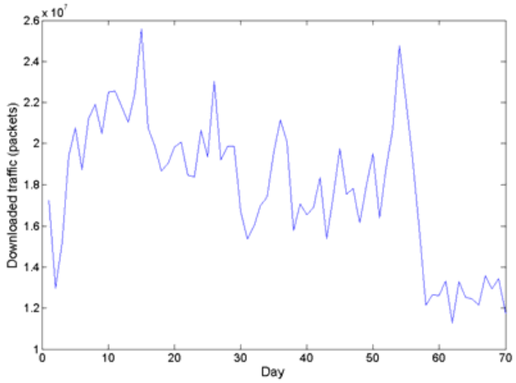


Fig. 10. Aggregated traffic volume per day (downloaded packets).

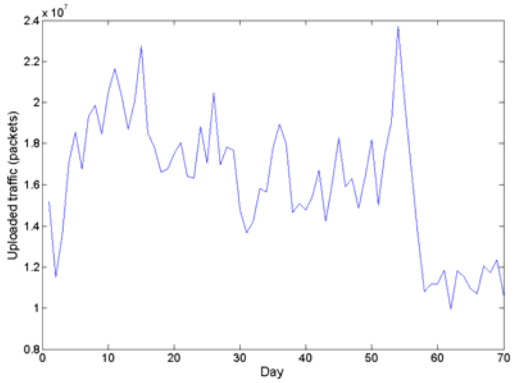


Fig. 11. Aggregated traffic volume per day (uploaded packets).

B. Daily (Diurnal) and weekly cycles

We also gained further insights by compressing the billing records into a day or a week, through averaging the data traffic volumes for the same hour over all days or over the same days of the week. Figs. 12-14 show the daily cycle for downloaded bytes, uploaded bytes, and downloaded and uploaded packets, respectively.

In all three figures, there is a daily minimum at 7 AM. From the daily minimum, the data traffic volume rises rapidly until the three daily maximums at 11 AM, 3 PM and 7 PM are reached. From 7 PM, the data traffic volume drops monotonically until 7 AM. In [19], similar traffic patterns has been recorded, with the difference of the third daily maximum occurring later into the evening (9-10 PM), rather than 7 PM, as recorded in the billing data.

Figs. 15-17 illustrate the traffic volume averaged over a week, with each day of the week exhibiting the same traffic pattern as Figs. 12-14. As expected, traffic volumes on weekends are lower than on weekdays when people work. In Fig. 15, the 3 daily maximums for Wednesdays are not as distinct as other days. This observation may be caused by the fact that both Dec. 24 and Dec. 31, 2002 fall on a Wednesday. As described in Section IIA, traffic volume on holidays have different dynamics and may cause the observed behavior.

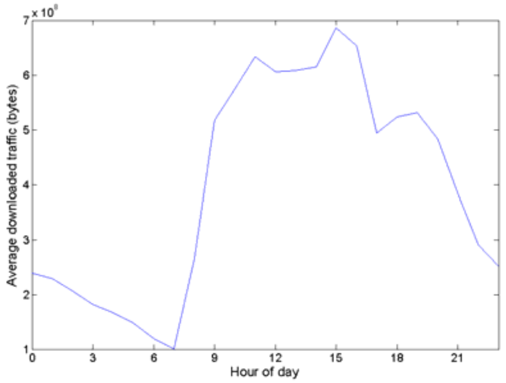


Fig. 12. Downloaded traffic volume (bytes) over a day by averaging all recorded values for the same hour.

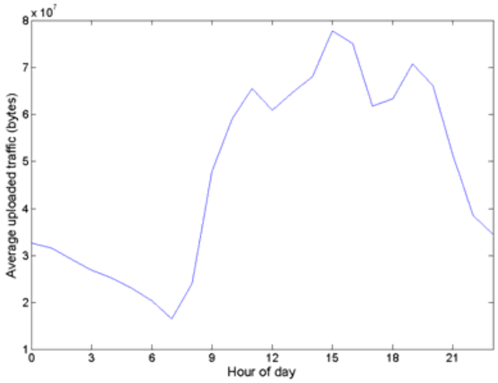


Fig. 13. Uploaded traffic volume (bytes) over a day by averaging all recorded values for the same hour.

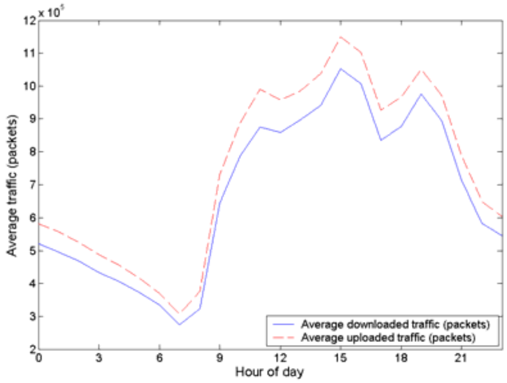


Fig. 14. Downloaded and uploaded traffic volume (packets) over a day by averaging all recorded values for the same hour.

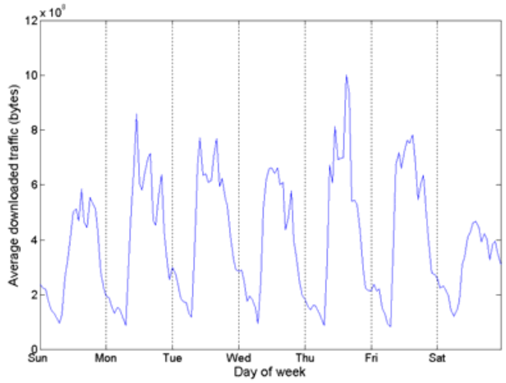


Fig. 15. Downloaded traffic volume (bytes) over a week by averaging all recorded values for the same hour.

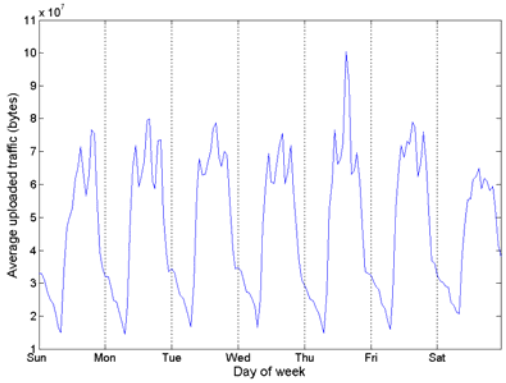


Fig. 16. Uploaded traffic volume (bytes) over a week by averaging all recorded values for the same hour.

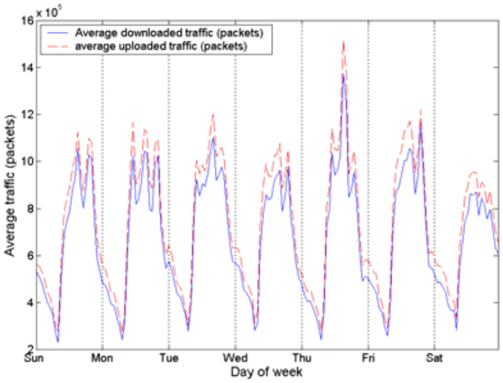


Fig. 17. Downloaded and uploaded traffic volume (packets) over a week by averaging all recorded values for the same hour.

VI. TCPDUMP TRAFFIC TRACE ANALYSIS

The *tcpdump* traffic traces was a continuous set of traces collected between December 14, 2002 and January 10, 2003. The data were stored in 127 collected files, containing ~63 Gbytes of data. Compared to the collected billing records, *tcpdump* trace has a much finer time granularity, in the order of several milliseconds. The *snaplen* option of the *tcpdump* is set to be 68 bytes. This setting allows the full TCP and IP headers to be recorded for analysis while preserving user privacy.

A. *Protocols and Applications*

Since Internet Protocol (IP) [20] is the most widely used network layer protocol on today's network, it is not surprising that the collected traffic traces contain only IP packets.

Our analysis of the *tcpdump* traces indicates that TCP packets accounts for 84.3% of the overall number of packets and accounts for 94.5% of the overall number of bytes sent. UDP accounts for 14.2% of packets and 5.06% of bytes. Lastly, ICMP accounts for 1.45% of packets and 0.45% of bytes.

Since TCP accounts for majority of the packet trace, we analyzed the activity by TCP port numbers. Table I summarizes the traffic in terms of connections and bytes. HTTP/WWW traffic (port 80) is the most widely used TCP application on the ChinaSat network, followed by FTP in terms of number of bytes. In addition, approximately 10% of all connections use unknown ports.

TABLE I. CHARACTERISTICS OF A TRAFFIC TRACE BY TCP APPLICATION.

Applications	Connections (%)	Bytes (%)
WWW (80)	90.0	76.8
FTP-data (20)	0.2	10.7
IRC (194)	0.8	0.008
SMTP (25)	0.1	0.01
POP3 (110)	0.03	0.02
Telnet (23)	0.02	0.002
Others	8.9	12.5
Total	100	100

In contrast, there are few known applications that use a standard UDP port. UDP, an unreliable protocol, is mainly used for real-time applications such as video streaming and Internet telephony. As many of these applications use random ports, we could not apply the same classification technique as used for TCP. The only application that we are certain to be in use over UDP is the Routing Information Protocol (RIP). It's used to communicate between the different hosts on a local network for routing. Although we have captured a large number of packets that uses RIP, they're unrelated to the DirecPC users on the ChinaSat network. Thus, we did not analyze these packets further.

B. *TCP options*

In Section III, we have discussed a number of extensions for TCP that improves TCP performance. Some of these extensions, such as SACK and sliding window scale option, are requested during the TCP 3-way handshake. Thus, the usage of these options for each connection can be determined by examining the first segments of the TCP 3-way handshake, which contains the TCP SYN flag.

We found that SACK is widely used on the ChinaSat network. Over 80% of the connections support SACK. On the other hand, less than 5% of connections use the sliding window scale option.

Further research shows that the most commonly deployed operating system, Microsoft Windows, supports and enables SACK by default since Windows 98 [21]. On the other hand, although sliding window scale option is included in all versions of Microsoft Windows since Windows 98, it is disabled by default [21]. Thus, Microsoft Windows TCP implementation explains the prevalent usage of SACK and the lack of usage of sliding window scale option in the recorded *tcpdump* trace.

While we are able to find the use of SACK and sliding window scale option, we are not able to determine if the other TCP extensions such as increasing the initial congest window size, Path MTU discovery are in use.

C. *Data traffic anomalies*

From our analysis of the *tcpdump* trace, we have observed three types of data traffic anomalies: packets with invalid TCP flag combinations, large number of connections that are closed through TCP reset, and port scans.

1) *Packets with invalid TCP flag combinations*

In the TCP protocol specification [4], three flags are used for opening and closing connections (SYN, RST, FIN). TCP SYN, TCP FIN, and TCP RST are used to open connections, to close connections regularly, and to close connections when an error occurs, respectively. These flags cannot be used in combination with each other. During the early days of the Internet, invalid flag combinations can cause TCP/IP stacks to misbehave or fail, and are used to test the robustness of TCP/IP stacks [22]. Thus, it is unusual to find packets with combinations of the TCP open/close flags from a deployed network. These packets may be caused by the use of malicious software or virus and worms. Table II summarizes the number of discovered packets with invalid TCP flag combinations. In total, 0.3% packets with TCP open/close flags have invalid combinations.

TABLE II. PACKETS WITH INVALID TCP FLAG COMBINATIONS.

TCP flag	Packet count	% of Total
SYN only	19,050,849	48.5%
RST only	7,440,418	18.9%
FIN only	12,679,619	32.3%
SYN+FIN	408	0.001%
RST+FIN (no PSH)	85,571	0.2%
RST+PSH (no FIN)	18,111	0.05%
RST+FIN+PSH	8,329	0.02%
Total number of packets with invalid TCP flag combinations	112,419	0.3%
Total packet count	39,283,305	

2) *Large number of TCP Resets*

In the previous section, we stated that TCP's normal procedure is to open connections with the SYN flag and to close connections with the FIN flag. However, data from Table II indicates that 37% of connections are closed by TCP reset (RST).

Further investigation shows that this is a "feature" of Microsoft Internet Explorer, the most common web browser during the period when the trace was captured. In [23], the authors discovered that Microsoft Internet Explorer uses TCP RST instead of TCP FIN to close connections. This intended "feature", contrary to the TCP protocol specification, was implemented to improve web browsing performance for Microsoft browsers.

3) *Port scans*

We discovered a significant amount of UDP port 137 activities both originating from the ChinaSat network users and directed to the ChinaSat network. The application used on UDP port 137 is the Microsoft NETBEUI protocol. This protocol enables file and printer sharing in a local network for Microsoft Windows PCs. In addition, NETBEUI's normal behavior is to connect to and from UDP port 137. Thus, repeated traffic from UDP port 137 to other UDP ports on a network or traffic from other UDP ports to UDP port 137 would indicate abnormal behavior. In Fig. 18, we show an example of host on the ChinaSat network (IP address 192.168.2.30) that transmits packets to a number of hosts on the Internet from UDP port 137. Note that for certain destination IPs (202.y.y.226), the host computer transmitted multiple ports (1025, 1027, 1028, and 1029). This behavior is known as port scans and usually indicates malicious intent.

Fig. 19 shows an example a host external to the ChinaSat network (IP address 210.x.x.23) transmitting packets to a ChinaSat hosts at the destination UDP port 137 from UDP port 1035. This type of behavior could also indicate malicious intent.

At the time the *tcpdump* traces was captured, two computer worms are prevalent: Bugbear and Opasoft. Both of these worms use the NETBEUI protocol to propagate themselves to other hosts. Without the TCP payload due to truncation, we are unable to determine if these port scans are indeed generated by these two worms.

```
192.168.2.30:137 - 195.x.x.98:1025
192.168.2.30:137 - 202.x.x.153:1027
192.168.2.30:137 - 210.x.x.23:1035
192.168.2.30:137 - 195.x.x.42:1026
192.168.2.30:137 - 202.y.y.226:1026
192.168.2.30:137 - 218.x.x.238:1025
192.168.2.30:137 - 202.y.y.226:1025
192.168.2.30:137 - 202.y.y.226:1027
192.168.2.30:137 - 202.y.y.226:1028
192.168.2.30:137 - 202.y.y.226:1029
192.168.2.30:137 - 202.y.y.242:1026
192.168.2.30:137 - 61.x.x.5:1028
192.168.2.30:137 - 219.x.x.226:1025
192.168.2.30:137 - 213.x.x.189:1028
192.168.2.30:137 - 61.x.x.193:1025
```

Fig. 18. Port scan originating from the ChinaSat network.

```
210.x.x.23:1035 - 192.168.1.121:137
210.x.x.23:1035 - 192.168.1.63:137
210.x.x.23:1035 - 192.168.2.11:137
210.x.x.23:1035 - 192.168.1.250:137
210.x.x.23:1035 - 192.168.1.25:137
210.x.x.23:1035 - 192.168.2.79:137
210.x.x.23:1035 - 192.168.1.52:137
210.x.x.23:1035 - 192.168.6.191:137
210.x.x.23:1035 - 192.168.1.241:137
210.x.x.23:1035 - 192.168.2.91:137
210.x.x.23:1035 - 192.168.1.5:137
```

Fig. 19. Port scan directed to the ChinaSat network.

VII. CONCLUSIONS

In this paper, we describe traffic collection from a commercial hybrid satellite-terrestrial network and analyzed collected traffic traces and billing records. From the billing records, we have shown that the downloaded and uploaded traffic volumes are highly regular, exhibiting both daily and weekly cycles. We have found that a daily minimum occurs at 7 AM and three daily maximums occur at 11 AM, 3 PM and 7 PM.

Analysis of *tcpdump* traces shows the trace to be dominated by TCP traffic where HTTP/WWW packets on port 80 form the majority of the captured data. Furthermore, by examining the TCP SYN packets, we are able to determine the SACK TCP extension is widely used to improve the TCP performance over satellites. Lastly, we were able to describe anomalies detected in the trace, including invalid TCP flag combinations, a large number of TCP resets, port scans, and have provided possible explanation for the origin of these behaviors.

While the collected traffic data captured only a snapshot of the satellite network, its analysis contributes to better understanding of deployed networks and may be of benefit to commercial network traffic management agencies.

ACKNOWLEDGMENT

The authors would like to thank the managers of the DirecPC system at ChinaSat, for their help with data collection.

REFERENCES

[1] S. McCreary, "Trends in wide area IP traffic patterns," *Proc. 13th ITC Specialist Seminar on Measurement and Modeling of IP Traffic*, Monterey, California, Sept. 2000, pp. 1–11.

[2] K. Thompson, G. J. Miller, and R. Wilder, "Wide-area Internet traffic patterns and characteristics," *IEEE Network Magazine*, vol. 11, no. 6, pp. 10–23, Nov. 1997.

[3] The Internet Traffic Archive. (2004, May). [Online]. Available: http://ita.ee.lbl.gov/.

[4] J. Postel, "Transmission control protocol, " RFC 791, Sep. 1981.

[5] M. Allman, D. Glover, and L. Sanchez, "Enhancing TCP over satellite channels using standard mechanisms," RFC 2488, Jan. 1999.

[6] V. Jacobson, R. Braden and D. Borman "TCP extensions for high-performance," RFC: 1323, May. 1992.

[7] M. Allman, S. Dawkins, D. Glover, J. Griner, D. Tran, T. Henderson, J. Heidemann, J. Touch, H. Kruse, S. Ostermann, K. Scott, and J. Semke, "Ongoing TCP research related to satellites," RFC 2760, Feb. 2000.

[8] S. Oueslati-Boulahia, A. Serhrouchni, S. Tohmé, S. Baier, and M. Berrada, "TCP over satellite links: problems and solutions," *Telecommunication Systems,* vol. 13, no.2-4, 2000, pp. 199-212.

[9] M. Allman, S. Floyd, and C. Partridge, "Increasing TCP's initial window," RFC 2414, Sep. 1998.

[10] M. Mathis, J. Mahdavi, S. Floyd, and A. Romanow, "TCP selective acknowledgement options," RFC 2018, Oct. 1996.

[11] J. Mogul and S. Deering, "Path MTU discovery," RFC 1191, Nov. 1990.

[12] J. Border, M. Kojo, J. Griner, G. Montenegro, and Z. Shelby, "Performance enhancing proxies intended to mitigate link-related degradations," RFC 3135, June, 2001.

[13] T. R. Henderson, R. H. Katz, "Transport protocol for Internet-compatible satellite networks," *IEEE J. Select. Areas Commun.*, vol. 17, no. 2, Feb. 1999, pp. 326–344.

[14] A. D. Falk, V. Arora, N. Suphasindhu, D. Dillon, and J. S. Baras, "Hybrid Internet access," *Conference on NASA Centers for Commercial Development of Space*, *AIP Conference Proceedings,* vol. 325, Jan. 1995, pp. 69–74.

[15] V. Arora, N. Suphasindhu, D. Dillon, and J.S. Baras. "Effective extensions of internet in hybrid satellite-terrestrial networks," Proc. 1st Conference of Commercial Development of Space, Albuquerque, NM, Jan. 1996, pp. 339–344.

[16] V. Arora, N. Suphasindhu, J. S. Baras, and D. Dillon, "Asymmetric Internet access over satellite-terrestrial networks," *Proc. AIAA: 16th International Communications Satellite Systems Conference and Exhibit,* , Washington, DC, Feb. 1996, pp. 476–482.

[17] J. S. Baras, S. Corson, S. Papademetriou, I. Secka, and N. Suphasindhu, "Fast asymmetric Internet over wireless satellite-terrestrial networks," *Proc. MILCOM '97*, Monterey, CA, Nov. 1997, pp. 372–377.

[18] Hughes Network performance enhancing technology technical specifications. (2006, June). [Online]. Available: http://www.hns.com/.

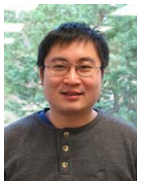
[19] V. Paxson, "Empirically-derived analytic models of wide-area TCP connections," *IEEE/ACM Transactions on Networking*, vol. 2, no. 4, pp. 316–336, Aug. 1994.

[20] J. Postel, "Internet protocol, " RFC 791, Sep. 1981.

[21] Microsoft Windows 2000 TCP/IP implementation details. (2006, June). [Online]. Available: http://www.microsoft.com/technet/itsolutions/network/deploy/depovg/tcpip2k.mspx.

[22] J. Postel, "TCP and IP bake off," RFC 1025, Sept. 1987.

[23] M. Arlitt and C. Williamson, "An analysis of TCP reset behavior on the Internet," *Computer Communications Review*, vol. 35, no. 1, pp.37–44, Jan. 2005.



Savio Lau received the B.A.Sc. degree in computer engineering from Simon Fraser University, Burnaby, BC, Canada, in 2003. He is currently working towards the M.A.Sc. degree in the Communications Networks Laboratory, School of Engineering Science, at the same university. His research interests include network traffic analysis and simulation.

Motion Planning of Multiple Agents in Virtual Environments on Parallel Architectures

Yi Li, *Student Member, IEEE,* and Kamal Gupta, *Member, IEEE*

***Abstract*— We proposed in a previous paper a two-layered approach to motion planning of multiple agents in static virtual environments, consisting of open spaces connected by multiple narrow passages. The Discrete Generalized Voronoi Diagram (GVD) of the environment is used to identify all narrow passages, and plan the global path of each agent automatically. As each agent moves along its global path, it is locally modified using the hybrid technique of combining steering behaviors with Coordination Graphs (CG). It is computationally expensive to construct coordination graphs and then compute optimal joint actions for deadlock avoidances in narrow passages. The planner was single-threaded, and it was able to plan motions of thirty agents moving around in a simple virtual environment with four narrow passages. If more agents are moving in a more complex virtual environment (i.e., with more narrow passages), optimal joint actions can no longer be computed in real-time. In this paper, we parallelize the original sequential code in a supervisor-worker paradigm. We show that significant, scalable speedups are obtained by processing coordination graphs in parallel on a Symmetric Multiprocessing (SMP) machine.**

***Index Terms*— Motion Planning, Virtual Environments, Parallel Programming, SMP.**

I. Introduction

We proposed in [1] a two-layered approach to motion planning of multiple agents in static virtual environments, consisting of open spaces connected by multiple narrow passages. The discrete general Voronoi diagram of the static environment is used to identify all narrow passages automatically. The global path of each agent is also planned using the GVD. To avoid collisions and deadlocks in the narrow passages, each agent's global path is then locally modified in real-time using a hybrid technique combining steering behaviors [2], [3] with coordination graphs [4].

Given a virtual environment with n narrow passages, n tasks are performed at each time step, where each task constitutes constructing a coordination graph for a particular narrow passage and then performing a variable elimination algorithm [4] to compute optimal joint actions (for deadlock avoidance). These tasks were performed sequentially on a single thread in [1]. However, if there are many narrow passages, a single processor may not be able to process all coordination graphs and compute optimal joint actions sequentially in real-time. Since tasks associated with different narrow passages are independent of one another, they can be performed in parallel on Symmetric Multiprocessing (SMP) machines in order to speed up the simulation. We do not consider distributed computers in this paper, because our research is related to computer games and all next-generation game consoles (i.e., Xbox 360, Nintendo Wii, and Sony PlayStation 3) are SMP machines containing multi-core processers.

Motion planning on parallel and distributed architectures have been studied extensively, especially parallelization of probabilistic path planners, such as *Probabilistic Roadmap* (PRM) [5] and *Rapidly-exploring Random Trees* (RRT) [6]. All parallel versions [7]–[11] of these two planners aim to solve high-dimensional problems and/or yield speedups by distributing work to multiple processors. All these planners are *offline* planners. Instead, we coordinate motions of the agents *online* by performing multiple tasks at each time step. These tasks are relatively small, and they are repeated as quickly as possible to achieve maximum frame rate. So, the parallel overhead must be minimized, otherwise the parallelized version may run slower than the sequential version.

There are two main parallel programming models: OpenMP and Message Passing Interface (MPI). With OpenMP, the original sequential code can be converted in a stepwise fashion by adding directives incrementally. However, OpenMP supports only loop-level parallelism, not task parallelism. MPI supports task parallelism on either SMP computers or distributed computers, but it requires more programming changes to go from sequential to parallel. Instead, we implement the task parallelism in a *supervisor-worker* paradigm with System V *Interprocess Communication* (IPC) mechanism.

This paper is organized as follows. We formally define the problem to be solved in section II. The two-layered approach is presented in section III. We show in section IV the importance of minimizing the parallel overhead. A parallel procedure for construction of multiple coordination graphs and computation of multiple optimal joint actions is presented in section V. A simple scheduler is presented in section VI to improve the parallel performance on multiple processors. In section VII, we describe the experiment, and the results obtained. We conclude in section VIII.

II. Problem Definition

Given are k agents $A_1, \ldots, A_k$, and a static 2D polygonal virtual environment in which the agents can move. Each agent is modeled by a disc of radius r with two degrees of freedom x and y given in the world coordinate space. The agents' start positions $\mathbf{P}_s = \{P_{s_1}, \ldots, P_{s_k}\}$, and goal positions $\mathbf{P}_g = \{P_{g_1}, \ldots, P_{g_k}\}$ are also defined. The virtual environment consists of open spaces connected by n narrow passages and for now we assume that no narrow passages intersect with each other.

The task is to move each agent from its start position to its goal position without colliding with other agents and static polygonal obstacles. At each time step, a coordination graph is generated for each narrow passage, and an optimal joint action is then computed for each coordination graph. All n optimal joint actions have to be computed in real-time.

III. The Two-Layered Approach

First, we compute the discrete generalized Voronoi diagram from the frame buffer of the graphics hardware [12]. The distance from any point on the GVD to its nearest obstacle (called *clearance*) can be obtained in the Z-buffer. With the GVD and the distance information, we not only identify all narrow passages in the environments, but also find all agents' global paths.

A. Narrow Passage Identification

We first use the GVD to identify those portions of GVD whose clearance is less than a user-defined threshold, then all narrow passages and their openings in the workspace can be easily identified. In addition, there are expanded openings at both ends of the narrow passage. They are used to control the number of agents inside their respective adjacent openings [13]. In order to be consistent with the terminology used in [13], we designate one opening as the upper one and the other opening as the lower one. All narrow passages, openings, and expanded openings shown in Fig. 1 have been identified automatically. Observe that all narrow passages' openings and expanded openings are drawn in different colors. Openings are drawn in darker gray (two different green) compared to the expanded openings (yellow and orange) surrounding them. One pair of opening and expanded opening is the upper one, and the other pair is the lower one.

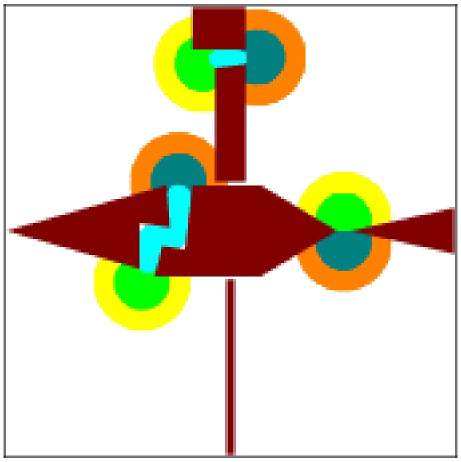


Fig. 1. Virtual environment with three narrow passages.

B. Motion Coordination with Coordination Graph

Since the agents are moving in a 2D virtual environment, the medial axis of the free space is a one dimensional entity; hence the GVD can be used as a roadmap. Using the GVD, the global paths between $\mathbf{P}_s$ and $\mathbf{P}_g$ are computed sequentially and independently of each other.

Having generated a global path for each agent, we want it to follow its path while avoiding obstacles and other agents. This can be accomplished by steering behaviors [2]. All steering behaviors are based on local information, therefore they work fine as long as the agents are located in open spaces, but the agents get easily stuck if they are located in cluttered environments, such as narrow passages. To avoid deadlocks in narrow passages, we presented in [13] a hybrid technique, combining local steering behavior and an efficient AI technique for decision making and planning cooperative multi-agent dynamic systems called a *Coordination Graph* (CG) [4]. We will not discuss details here, just present some broad essentials.

Suppose that m agents $\mathbf{A} = \{A_1, \ldots, A_m\}$ are located inside a narrow passage and its two openings at time t. These agents must coordinate their actions in order to pass through while avoiding deadlocks. All agents are acting in a space described by a set of discrete state variables, $\mathbf{X} = \{X_1, \ldots, X_n\}$. A state $\mathbf{x} = \{x_1, \ldots, x_n\}$ is an assignment to each state variable X_i. Each agent A_j chooses an action a_j from a finite set of possible actions $Dom(A_j)$. Before an optimal joint action $\mathbf{a} = \{a_1, \ldots, a_m\} \in Dom(\mathbf{A})$ is computed, each agent is assigned a local value function $Q_j(\mathbf{x}, \mathbf{a})$. A local value function is basically a set of value rules, where each value rule describes a context — an assignment to state variables and actions — and a value increment. We will not discuss how the rules were derived here. An optimal joint action is simply one joint action that maximizes the total utility function $Q = \sum_j Q_j(\mathbf{x}, \mathbf{a})$, where $Q_j(\mathbf{x}, \mathbf{a})$ is agent A_j's local value function.

Suppose that agent A_1 is located in a narrow passage, and agents A_2 and A_3 are located inside the narrow passage's lower opening and moving toward the passage. Follow the procedure presented in [13], the agents' local value functions are

$$Q_1(\mathbf{x}, \mathbf{a}) = \begin{cases} \langle a_1 \wedge x_1 : 100 \rangle \\ \langle \bar{a}_1 \wedge \bar{x}_1 : 100 \rangle \end{cases} \tag{1}$$

$$Q_2(\mathbf{x}, \mathbf{a}) = \begin{cases} \langle a_1 \wedge a_2 \wedge x_2 : 50 \rangle \\ \langle \bar{a}_1 \wedge a_2 \wedge x_2 : -50 \rangle \\ \langle a_1 \wedge a_2 \wedge \bar{x}_2 : -50 \rangle \\ \langle \bar{a}_1 \wedge a_2 \wedge \bar{x}_2 : 50 \rangle \end{cases} \tag{2}$$

$$Q_3(\mathbf{x}, \mathbf{a}) = \begin{cases} \langle a_2 \wedge a_3 \wedge x_1 \wedge x_3 : -200 \rangle \\ \langle a_2 \wedge a_3 \wedge \bar{x}_1 \wedge \bar{x}_3 : -200 \rangle \end{cases} \tag{3}$$

where

$$\begin{aligned} a_1 &= up \\ a_2 = a_3 &= enter \\ x_1 &= moves\_up\_in\_passage(A_1) \\ x_2 &= inside\_lower\_opening(A_2) \\ x_3 &= inside\_lower\_opening(A_3) \end{aligned}$$

All action and state variables are binary variables in [13]. $\bar{a}$ and $\bar{x}$ are negates of a and x, respectively.

$$\begin{aligned} \bar{a}_1 &= down \\ \bar{a}_2 = \bar{a}_3 &= not\_enter \\ \bar{x}_1 &= moves\_down\_in\_passage(A_1) \\ \bar{x}_2 &= inside\_upper\_opening(A_2) \\ \bar{x}_3 &= inside\_upper\_opening(A_3) \end{aligned}$$

The value rules in the agents' local value functions give awards (i.e., positive values) to desired behaviors, and penalties (i.e., negative values) to non-desired behaviors. For example, the value rules in (1) encourage A_1 to keep going in its current direction, whereas the value rules in (3) prevent A_2 and A_3 from entering a narrow passage at the same time, if they are located in the same opening. The value rules in (2) only allow A_2 to enter the narrow passage, where A_1 is currently located in, if no deadlock will occur.

Next, a coordination graph is automatically constructed with the information in the local value functions. Each node of the coordination graph represents one agent, and two nodes are connected only if the corresponding agents have to coordinate their actions; hence the coordination graph can capture local coordination requirements between agents and the action space is reduced. The coordination graph shown in Fig. 2 is constructed using the local value functions given in (1, 2, 3). Agent A_2's local value function Q_2 as seen in (2) shows clearly that decisions made by agent A_1 affect agent A_2; hence node 1 is connected with node 2 by a directed edge.

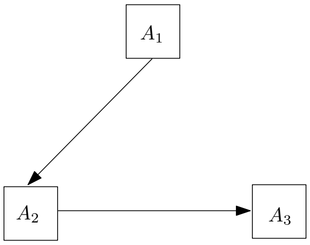


Fig. 2. Coordination graph for a coordination problem with 3 agents.

Once all agents's states have been observed, all value rules which are not consistent with the current states are deleted. The variable elimination algorithm is then used to find optimal joint actions of the agents in combination with a message passing scheme [4]. In order to compute an optimal joint action in real-time, we reduce the number of nodes in a coordination graph by considering all agents inside the narrow passage as one virtual agent and representing them with one node [13]. Agents that just exited from the narrow passage, but are still located inside one of its openings, are not represented by nodes in the coordination graph. The biggest coordination graph in [13] has 5 nodes.

For the coordination problem shown in Fig. 1 and 2, optimal joint action is simply $\{\bar{a}_1, \bar{a}_2, \bar{a}_3\}$, if the current state of agent A_1 is $\bar{x}_1$ (i.e., agent A_1 keeps moving down, A_2 and A_3 are not allowed to enter the narrow passage). Clearly, deadlock is avoided by this optimal joint action.

C. Sequential Code

The single-threaded planner in [1] is an OpenGL program. A *draw* function is invoked as often as possible, because we want to update the window at the highest frame rate possible. Main steps of the sequential code inside the draw function are:

1) Observed each agent's state (i.e., its location and status).
2) Construct n coordination graphs sequentially, where n is the number of narrow passages in the virtual environment.
3) For each coordination graph, compute an optimal joint action.
4) Update the window.

All optimal joint actions have to be computed in real-time regardless of the number of coordination graphs. The single-threaded planner in [1] can handle a virtual environment with three narrow passages, when there are 30 agents moving around in the virtual environment. Once there are more than 30 agents in the same environment, the window can no longer be updated at a sufficient rate. In order to handle more agents, and more complicated virtual environments, we have to speed up the computation of optimal joint actions, by processing coordination graphs in parallel on a SMP machine.

IV. OpenMP

Parallelization of the sequential code in [1] can be easily done using directives-based parallel programming language such as *OpenMP* [14]. The sequential code is parallelized by adding a few directives, which appear as comments in the sequential code. OpenMP uses the fork-join model of parallel execution: the parallelized program still begins execution on a single thread, called the master thread of execution. A team of parallel threads are created by the master thread, when a parallel construct inside the draw function is encountered and the statements (i.e. construction of coordination graphs and computation of optimal joint actions) enclosed by the parallel region construct are executed in parallel among the team of threads. Upon completion of the parallel construct, the threads synchronize and all threads except the master thread are terminated. However, it is expensive to create and destroy threads [14], especially when this is repeated each time the draw function is invoked. Also, the number of narrow passages is limited (less than a couple of hundred) and each task of construction of a coordination graph and computation of an optimal joint action is small. All these lead to excessive parallel overhead. In fact, the parallelized version is slower than the sequential version.

V. Supervisor-Worker Paradigm

Instead of using OpenMP, we implemented the task parallelism in a *supervisor-worker* paradigm (Fig. 3): a single *supervisor*, also called *master*; asks multiple *workers*, also called *slaves*, to perform the tasks.

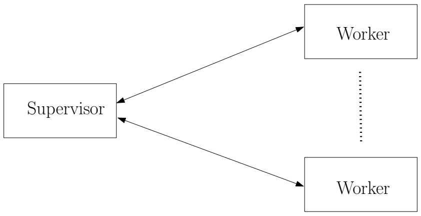


Fig. 3. The Supervisor-work paradigm.

The supervisor/worker paradigm can be implemented with either threads or processes. Threads should be used when the same complex data structures have to be processed concurrently, and processes should be used instead for less tightly couple applications [15]. Therefore, we choose processes for implementation of the supervisor/worker paradigm, because all workers perform their tasks independently of one other; consequently, no data is passed between workers. Another advantage of processes over threads is that it is easier to develop and test processes separately. However, each process has its own set of memory pages; hence, two processes need to use *Interprocess Communication* (IPC) to communicate to one another. We use System V IPC for interprocess communication. POSIX IPC is much newer than System V IPC, but it is not as widely available as System V IPC [15]. There are three kinds of System V IPC objects: *message queues*, *semaphore sets*, and *shared memory segments*.

We divide the single-threaded program in [1] into two programs: a supervisor program and a worker program. Once the supervisor program is launched, we make a system call *fork* to create a second process. The original process is called the *parent*, and the new one is called *child*. The cost of a *fork* is potentially enormous, but it is called only once. In order to create multiple worker processes, we launch the worker program multiple times. The number of the worker processes depends on the number of available processors. The supervisor processes communicate with the worker processes through two message queues (Fig. 4): the supervisor sends messages to the workers through message queue A, and the workers send back the results through message queue B.

At each time step, the parent process iterates through all narrow passages. For each narrow passage, the parent process finds all agents inside the narrow passage and all agents inside the neighboring two openings who are moving toward the narrow passage. The agents' states are then stored inside a message, before the message is appended to the message queue A. As soon as a worker process is free, it pops the first message in the message queue A. Using the data stored inside the message, the worker process performs the task. The optimal joint action is sent back to the supervisor through message queue B, and received by the child process. The process is repeated until the message queue A is empty and the length of the message queue B is equal to the number of narrow passages n in the virtual environment. The parent process and the child process communicate with each other through shared memory (Fig. 4). Once the child process has received all n messages from the workers, it writes the data (i.e., n optimal joint actions) in the message queue B into the same memory it shares with the parent process.

The main advantage of our approach is that all processes are created only once and destroyed only once; hence, the parallel overhead is reduced to a minimum. We use a message queue for communication between the supervisor and the workers, because multiple worker processes can fetch messages from the same message queue (i.e., the message queue A), and small messages (100 bytes or so) can be passed between two processes quickly [15]. However, there are limits on the size of a message and the number of messages in the queue [15]. For example, we are able to send 40 message to the message queue A on an SGI UltimateVision. If there are more than 40 narrow passages in the virtual environment, shared memory can be used instead of message queue for communication between the supervisor and its workers. Shared memory should also be used for big messages [15].

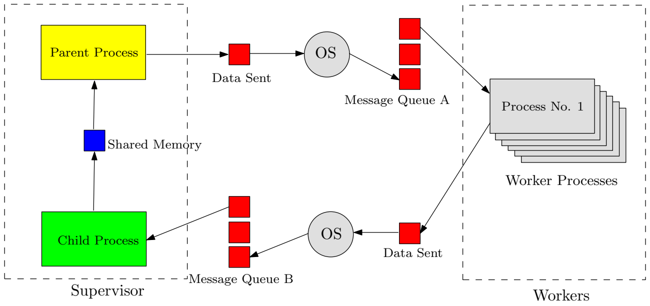


Fig. 4. The Supervisor and its workers communicate through System V message queues and shared memory.

VI. Job Scheduling

At each time step, we have to perform n tasks (i.e., construct n coordination graphs, and then compute an optimal joint action for each one of them). We want to minimize the length of time required to complete all tasks: *makespan* M. Assume that there are m identical processors available for processing in a SMP machine. Assume also that a task can only be processed by one processor, and a task cannot be interrupted. This is a classical *scheduling problem of parallel identical processors and independent jobs* (tasks). There is no algorithm that constructs an optimal schedule [16]. Even if such algorithm exists, it would be much too slow for us; however, a fast and effective heuristic procedure for minimizing M exists: *Longest Processing Time* (LPT), and the schedules it produces are close-to optimal [16].

In general, processing time grows with the number of nodes in a coordination graph. Therefore, an LPT ordering of the tasks can be obtained by sorting the coordination graphs according to the numbers of the nodes. We use *insertion sort*, because it is very simple to implement and efficient on a small data set. The average runtime of insertion sort is $O(n^2/4)$. In the worst case, the sorting takes $O(n^2)$ time. Since n is a relatively small number, runtime of the insertion sort is very short and can be neglected.

VII. Experiments and Results

A. Hardware and Software Setup

The code was written in ANSI C/C++ and compiled using MIPSpro Compilers Version 7.41. The experiments were performed on an SGI UltimateVision with 24 64-bits MIPS R16000 processors at 700 megahertz and running IRIX 6.5.28. Each MIPS processor has 32 kilobyte L1 instruction cache, 32 kilobyte L1 data cache, and 4 megabyte L2 cache.

The SGI UltimateVision is based on SGI's NUMAflex architecture. All memory inside the machine are interconnected by the SGI NUMAlink interconnect technology to a single, system-wide, shared-memory space called the global shared memory. The SGI UltimateVision has 14 gigabyte of global

TABLE I

RUNTIME OF ONE TASK.

Task Size	**1**	**2**	**3**	**4**	**5**	**6**
Time in msec	2	10	20	64	107	160

shared memory and it is visible to all 24 processors. The NUMAlink enables each processer to access its local and remote memory with low-latency.

B. Performance Analysis

An important performance metric is *speedup* S [17], which is defined as

$$S(N, P) = \frac{T_{seq}(N)}{T(N, P)}, \tag{4}$$

where N is the size of the problem, P is the number of processors, T_{seq} is the runtime of the sequential program, and $T(P)$ is the runtime of the parallel program.

To measure the parallel efficiency of our implementation of the supervisor/worker paradigm, we created a simplified supervisor that enables us to perform some experiments in a controlled environment. The simplified supervisor can generate jobs of different sizes. A job consists of one or multiple tasks. The size of the job varies not only with the number of the tasks, but also with the size of each task, which is equal to the number of nodes of the coordination graph inside the task, and it ranges from 1 to 6 in our experiments. Depending on the experiment, the simplified supervisor generates a job with a certain number of tasks, where the sizes of the tasks maybe equal to one another or different. To measure runtime, a timer is started before the parent process in the simplified supervisor sends messages to the workers through the message queue A; it is stopped once the parent process receives all optimal joint actions from the child process.

First, we measured how the runtime of one task varies with the size of the task. For each job, the simplified supervisor sent 40 tasks of same size to the workers for processing. The average runtime for 1 task is shown in Table I.

Next, we investigated how the size of a task (i.e., the number of nodes in a coordination graph) affects the speedup. Three different numbers of nodes were tested: 2, 4, and 6. For each test, the simplified supervisor sent 40 tasks of equal size to the workers. Runtimes (averaged over 3 runs) for 1, 2, 5, 10, and 20 workers are shown in Table II and Fig. 5. The speedups for coordination graphs with 4 and 6 nodes are almost linear. However, for coordination graphs with 2 nodes, the speedup is just over 10 when using 20 workers, due to the *communication overhead*. The effect of the communication overhead is especially evident when the jobs are small, and they are distributed to many processors (e.g., when 40 coordination graphs with 2 nodes are processed by 20 workers). Table I shows that it only takes around 10 msec to process a task with size 2, compared to 64 msec and 160 msec for tasks with sizes 4 and 6, respectively.

TABLE II

RUNTIMES OF 40 EQUAL-SIZE TASKS.

No. of Processors (P)	**1**	**2**	**5**	**10**	**20**
Time in msec (2 nodes)	412	210	86	51	33
Time in msec (4 nodes)	2568	1279	521	268	149
Time in msec (6 nodes)	6326	3148	1272	647	338

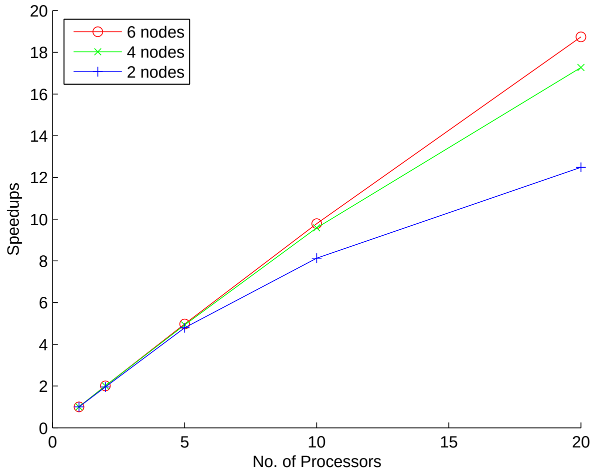


Fig. 5. Speedups for 40 equal-size tasks.

Speedup S is not only limited by the communication overhead, but also by the speed of the slowest processor; hence, S is also reduced from its ideal value of P, when loads over the processors are not balanced. Thus, we have to ensure that each processor performs the same amount of work (if possible). Given multiple tasks with mixed sizes, we use the LPT scheduler to distribute these tasks evenly among the available processors so that each processor performs about the same amount of work.

We tested effect of the LPT scheduler with 20, 30, and 40 tasks. The size of each task was determined randomly by the simplified supervisor; and it ranged from 1 to 6. The tests were performed first without scheduler, and then with the LPT scheduler. Runtimes (averaged over 10 runs) are shown in Table III and Table IV, respectively. Speedups are shown in Fig. 6 and Fig. 7, respectively.

It is clear from the data in Table III, Table IV, Fig. 6, and Fig. 7 that the parallel performance improves significantly when the LPT scheduler is used. Fig. 6 and Fig. 7 show that

TABLE III

RUNTIMES OF NOT EQUAL-SIZE TASKS (WITHOUT SCHEDULING).

No. of Processors (P)	**1**	**2**	**5**	**10**	**20**
Time in msec (20 CGs)	1166	581	323	188	163
Time in msec (30 CGs)	1729	850	439	278	187
Time in msec (40 CGs)	2473	1177	521	347	204

TABLE IV

RUNTIMES OF NOT EQUAL-SIZE TASKS (WITH SCHEDULING).

No. of Processors (P)	**1**	**2**	**5**	**10**	**20**
Time in msec (20 CGs)	1276	652	273	165	160
Time in msec (30 CGs)	1913	844	343	204	167
Time in msec (40 CGs)	2276	1223	485	274	169

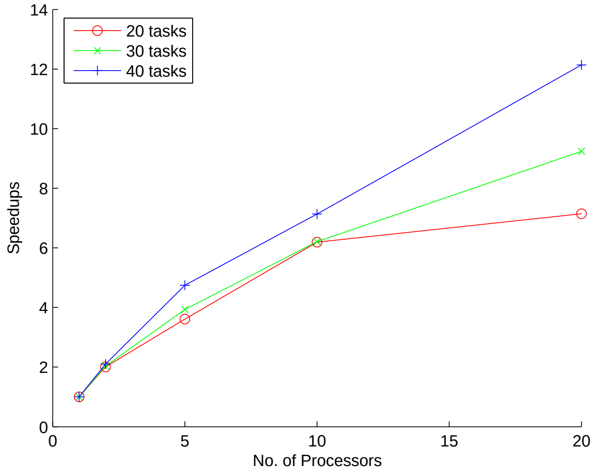


Fig. 6. Speedups for not equal-size tasks (without scheduling).

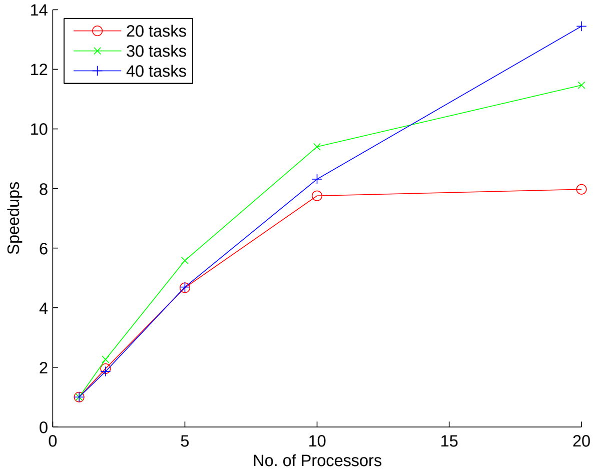


Fig. 7. Speedup for not equal-size tasks (with scheduling).

speedups of 20 tasks are clearly sub-linear. The reason is that multiple tasks cannot be processed in less than 160 msec, if the size of at least one task is 6, as a task with size 6 (Table I) takes about 160 msec to process. The last column of Table IV clearly shows this limit, where all runtimes are about 160 msec.

VIII. Conclusions

We implemented the task parallelism in the supervisor/worker paradigm and let the supervisor processes communicate the worker processes through System V IPC. Since all processes are created and destroyed only once, the parallel overhead is minimized. The experiments show that significant, scalable speedups are obtained. By processing multiple coordination graphs in parallel on a SMP machine, the two-layered approach we proposed in [1] can now handle more agents and more complicated virtual environments.

Acknowledgment

This research has been enabled by the use of WestGrid computing resources, which are funded in part by the Canada Foundation for Innovation, Alberta Innovation and Science, BC Advanced Education, and the participating research institutions. WestGrid equipment is provided by IBM, Hewlett Packard and SGI. The authors would like to thank Brian Corrie at the *Interdisciplinary Research in the Mathematical and Computational Sciences* (IRMACS) Centre, Simon Fraser University for his help and support, and for fruitful discussions.

References

[1] K. Anka, “A hybrid two-layered approach to real-time motion planning of multiple agents in virtual environments,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006.

[2] C. W. Reynolds, “Flocks, herds and schools: a distributed behavioural model,” *Computer Graphics*, vol. 21, no. 4, pp. 25–34, 1987.

[3] ——, “Steering behaviors for autonomous characters,” in *Game Developers Conference*, 1999, pp. 763–782.

[4] C. E. Guestrin, “Planning under uncertainty in complex structured environments,” 2003. [Online]. Available: http://ai.stanford.edu/~guestrin/Publications/Thesis/thesis.pdf

[5] 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*, vol. 12, no. 4, pp. 566–580, 1996.

[6] S. M. LaValle and J. J. Kuffner, “Randomized kinodynamic planning,” in *IEEE International Conference on Robotics and Automation*, vol. 1, 1999, pp. 473–479.

[7] M. Akinc, K. E. Bekris, B. Y. Chen, A. M. Ladd, E. Plaku, and L. E. Kavraki, *Probabilistic Roadmaps of Trees for Parallel Computation of Multiple Query Roadmaps*, ser. Robotic Research: The Eleventh International Symposium. Springer, STAR 15, 2005, pp. 80–89.

[8] E. Plaku, K. E. Bekris, B. Y. Chen, A. M. Ladd, and L. E. Kavraki, “Sampling-based roadmap of trees for parallel motion planning,” *IEEE Transactions on Robotics*, vol. 21, no. 4, pp. 597–608, 2005.

[9] S. Carpin and E. Pagello, “On parallel RRTs for multi-robot systems,” in *the 8th Conference of the Italian Association for Artificial Intelligence*, 2002, pp. 834–841.

[10] N. M. Amato and L. K. Dale, “Probabilistic roadmap methods are embarrassingly parallel,” in *IEEE International Conference on Robotics and Automation*, vol. 1, 1999, pp. 688–694.

[11] P. Isto, “A parallel motion planner for systems with many degrees of freedom,” in *International Conference on Advanced Robotics*, 2001, pp. 339–344.

[12] K. E. Hoff, T. Culver, J. Keyser, M. Lin, and D. Manocha, “Interactive motion planning using hardware-accelerated computation of generalized voronoi diagrams,” in *IEEE International Conference on Robotics and Automation*, vol. 3, 2000, pp. 2931–2937.

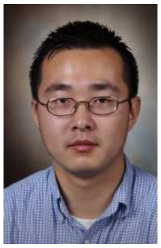
[13] K. Anka, “Motion planning of multiple agents in virtual environments using coordination graphs,” in *IEEE International Conference on Robotics and Automation*, 2005, pp. 380–385.

[14] R. Chandra, L. Dagnum, D. Kohr, D. Maydan, J. McDonald, and R. Menon, *Parallel programming in OpenMP*. San Francisco, CA: Morgan Kaufmann Publishers, 2001.

[15] M. J. Rochkind, *Advanced UNIX Programming*. Addison-Wesley Professional, 2004.

[16] K. R. Baker, *Introduction to Sequencing and Scheduling*. John Wiley and Sons Inc, 1974.

[17] G. Fox, M. Johnson, G. Lyzenga, S. Otto, J. Salmon, and D. Walker, *Solving Problems on Concurrent Processors: General Techniques and Regular Problems*. Englewood Cliffs, N.J.: Prentice Hall, 1988, vol. I.



Yi Li received the M.S. degree in electrical engineering from Royal Institute of Technology (KTH), Stockholm, Sweden, in 2001. In 1999/2000, Mr. Li. was an exchange student at University of Houston, TX. From 2001 to 2002, he worked at Ericsson AB in Kista, Sweden as a software engineer and developed software for Ericsson's WCDMA Radio Base Stations (RBS). Mr. Li is currently on leave from Ericsson and working towards the Ph.D. degree in robotics at Simon Fraser University, Burnaby, BC, Canada.

Kamal Gupta Kamal K. 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, Rhône-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 Societys Technical Committee on Motion and Path Planning.

The Continuously Adaptive Mean-SHIFT (CAMSHIFT) Algorithm for Color-based Visual Tracking

Yan Lu

***Abstract*— Visual tracking is the task of following moving targets through image sequences. It has such wide practical applications that techniques for visual tracking grow increasingly versatile and sophisticated. Three categories of tracking techniques are reviewed in this paper. In consideration of computational cost and speed, a color-based tracking algorithm – Continuously Adaptive Mean-SHIFT (CAMSHIFT) is chosen to build up a single, active camera tracking system. The algorithm converts the raw image of each video frame to a color probability distribution image via a histogram model of the color being tracked. By operating on the probability image, the location and the size of the target are found and determined . This paper explains in detail the principle of the CAMSHIFT algorithm and its implementation. Moreover, the paper points out a problem of legacy induced by mechanical constraints inherent with the pan-tilt-zoom (PTZ) camera. This problem is solved by modifying the CAMSHIFT algorithm. Through analysis, the color-based visual tracking system has proven to be robust, quick, and versatile with a modest computational cost. It can be utilized for such visual tracking tasks as surveillance, human-machine interface, and smart environment.**

***Index Terms*— CAMSHIFT, color-based tracking, computer vision.**

I. Introduction

As a useful robotic sensor, computer vision mimics the human sense of vision and allows for non-contact measurement of the environment. Visual tracking is the task of following moving targets of interest through image sequences. By extracting useful data from the sequences, visual tracking systems provide rich information about positions and orientations of a target, and enable estimation of its future poses. A problem well motivated by practicality, visual tracking has been widely applied in the area of surveillance, monitoring, and medical imaging. It is also an essential portion for such complex systems as visual servoing and Human-Computer Interaction (HCI). In a visual servoing system, results obtained from the tracker are input to the controller and the actuator to perform target manipulation tasks. HCI systems utilize tracking results containing characteristics of human hands or faces to understand gestures and facial emotions.

Tens or even hundreds of different techniques are currently available to meet increasingly complex and diverse tracking requirements. Most of them share common fundamental principles, and can be categorized into three classes: feature-based tracking, motion-based tracking, and model-based tracking. Feature-based tracking locates the position of a target by extracting from image sequences such characteristics of the target as edges, corners, or colors. Laplacian operators, Canny edge detector, and Hough transform are popular methods applied on raw images to obtain geometric primitives [1] [2] [3]. The CAMSHIFT algorithm [4] and Meanshift algorithm [5] are methods of tracking based on color. Instead of dependent on feature extraction, motion-based tracking detects motions of an object through frame comparison or optical-flow [6] [7]. Model-based tracking explicitly builds a model for the target, and it is able to estimate the pose of the target in the next time step by iterating current and past target status using predefined models [8]. Different tracking methods are applied to accomplish various tasks related with computer vision.

The selection of vision sensors or cameras is another key factor determining the performance of a tracking system. Along with the advancement of tracking techniques is the improvement of camera performance. Compared with fixed ones, active cameras with the appealing ability of pan, tilt, and zoom not only are more flexible for manipulation, but also perceive a wider and deeper view of the scene. Therefore, choosing an active camera, in some cases, avoids building a multiple, fixed camera system, which would otherwise involve extra efforts on the calibration and camera-to-camera geometric correlations.

[9] [10] present motion-based tracking methods using active cameras; however, their performance is largely limited by implementing real time filters whose computational cost is not trivial. In order to maintain the agility brought by using an active camera, it is desirable to choose a tracking technique with low computational burden, and thus a quick response. This paper implements the CAMSHIFT algorithm [4] (CAMSHIFT) in a single, active camera tracking system. As a color-based method, CAMSHIFT is fast and computationally efficient for object representation and tracking. It is especially appealing for tracking tasks where the spatial structure of the tracked objects exhibits such a dramatic variability that methods based on a space-dependent appearance reference would quickly break down.

II. The CAMSHIFT Algorithm

CAMSHIFT converts the raw image of each video frame to a color probability distribution image via a histogram model of the color being tracked. The center and the size of the color object are found by operating on the color probability image. The current location and size are used to set the location and the size of the search window in the next video image. The process is then repeated for continuous tracking. In fact, CAMSHIFT is derived from the mean shift algorithm. The latter operates on a static color probability distribution that

Manuscript received June 29, 2006; revised July 27, 2006.

Yan Lu is with School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada (luyanl@sfu.ca).

is represented by color histograms. The mean shift algorithm must be modified to adapt dynamically to the probability it is tracking, because color distributions obtained from video image sequences change over time. Fig. 1 summarizes the algorithm described above, and shows how the mean shift algorithm (indicated by the gray box) is developed into CAMSHIFT.

A. *Color Model*

To say that CAMSHIFT is model-free is not entirely accurate, because it incorporates a probabilistic model based on color. For every possible color in the scene captured by the camera, the algorithm assigns a probability that the color will appear in the object to be tracked. Explained alternatively, CAMSHIFT uses the model to compute the probability that a pixel in a frame is part of the object in the scene.

Several aspects should be considered before choosing a proper color model for tracking. Most digital images are handled using the Red, Green, and Blue (RGB) color space, but the individual R, G, and B components will vary widely under changing illuminations. The Hue, Saturation, and Value (HSV) color space, however, represents color in a different way. Saturation is the grayness of a color, and Value means the brightness or intensity of the color. Hue is determined by the dominant wavelength of a color, and it contains the most important color information. Moreover, Hue is the component in HSV color space that is free from illumination or lighting impact. Therefore, CAMSHIFT creates color models by taking 1-D histograms from the hue channel.

The hue histogram model of an object is computed by counting the number of object pixels that have corresponding hue values. Divided by total pixel numbers occupied by the object, the histograms are normalized, representing the possibility of each hue value the object has. The histogram model is much like a lookup table. CAMSHIFT first extracts hue values from the incoming video pixels, then converts them to corresponding probabilities in the table, and finally outputs the result known as the backproject image. The location and the size of the object are found based on the backproject image. This process is discussed in subsection B.

With different targets in the scenes, a green card and human skin respectively, Fig. 2 and Fig. 3 illustrate the procedure to obtain backproject images. Hue histogram models of the targets are built by extracting and sorting hue values of pixels within the red rectangle areas, as shown in (a) of each figure. Subfigures (b) show the model, with hue values as the independent variable along the x axis, and the probability of each hue value as the dependent variable along the y axis. For better illustration purpose, histogram bars are colored according to the objects in RGB space in order to build visual relationships between the targets and the histograms. Subfigures (c) are backproject images where areas with high probabilities of having the objects are highlighted.

B. *Algorithm Operation*

After the backproject image of a frame from input video sequences is computed, the main work of CAMSHIFT takes place. As implied by the meaning of its acronym (Continuously Adaptive Mean Shift), CAMSHIFT is a variation of the mean shift algorithm, a robust metric used to compute the mean value of a variable, given the probabilities associated with each possible value. The mean shift algorithm uses the zeroth and first moments of the backproject image to compute the center of an area with high probability. If $I(x, y)$ denotes the pixel probability value at position (x, y) in the backproject image, the zeroth moment, M_{00}, is

$$M_{00} = \sum_x \sum_y I(x, y), \tag{1}$$

and the first moments, M_{10} and M_{01}, for x and y are

$$M_{10} = \sum_x \sum_y xI(x, y), \tag{2}$$

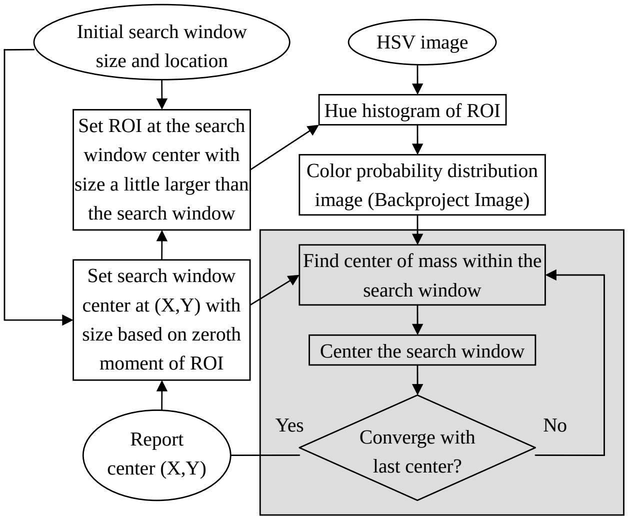

Initial search window size and location
HSV image
Set ROI at the search window center with size a little larger than the search window
Hue histogram of ROI
Color probability distribution image (Backproject Image)
Set search window center at (X,Y) with size based on zeroth moment of ROI
Find center of mass within the search window
Center the search window
Yes
No
Converge with last center?
Report center (X,Y)

Fig. 1. Block Diagram of the CamShift Algorithm

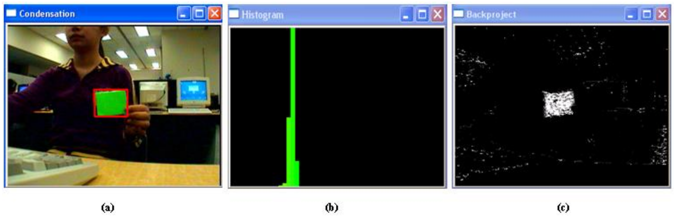


Fig. 2. Histogram and Backproject Images of a Frame with the Green Card as the Target

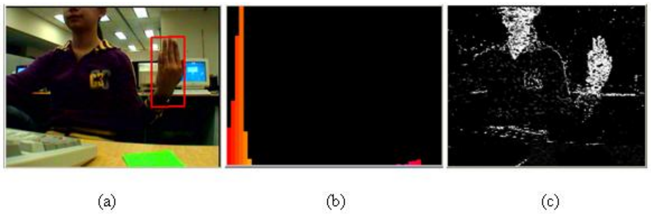


Fig. 3. Histogram and Backproject Images of a Frame with Human Skin as the Target

$$M_{01} = \sum_x \sum_y yI(x, y). \tag{3}$$

Therefore, the coordinates of the mean search window location are

$$x_c = \frac{M_{10}}{M_{00}}, \tag{4}$$

$$y_c = \frac{M_{01}}{M_{00}}. \tag{5}$$

The search window is then centered at (x_c, y_c), and its size is adjusted as a function of the zeroth moment. The algorithm repeats the computation of the center, and locates and resizes the search window until the result converges, which means the window moves for a distance less than the preset threshold. The steps of the mean shift algorithm listed below further clarify its idea. Step 1 and step 2 select the target to be tracked in the scene, and the rest of the steps are looped to realize automatic positioning.

THE MEAN SHIFT ALGORITHM

1. Choose an initial search window size to fit the target.
2. Choose an initial location of the search window based on the location of the target.
3. Compute the mean location (x_c, y_c).
4. Center the search window at the mean location computed in Step 3.
5. Repeat steps 3 and step 4 until the search window center converges.

CAMSHIFT uses the mean shift algorithm as the core, but expands it to deal with dynamically changing distributions. When the target in video sequences moves, the size and the location of the probability distribution change in time accordingly. CAMSHIFT adjusts the search window size during the tracking operation. Initial window size can be set at any reasonable value. During the course of tracking, CAMSHIFT relies on the zeroth moment M_{00} to determine the search window height and width. M_{00} can be treated as the distribution area found under the search window. Therefore, the larger M_{00} is, the bigger the search window size will be. Incorporating the mean shift algorithm, the steps of CAMSHIFT are listed below.

THE CAMSHIFT ALGORITHM

1. Choose an initial location of the search window.
2. Apply the mean shift algorithm as above, and store the zeroth moment.
3. Set the search window size equal to a function of the zeroth moment.
4. Repeat step 2 and step 3 until the search window center converges.

To conclude, CAMSHIFT is based on the following principle: the current frame is searched for a region, a variable-shape and variable-size window, whose color content best matches a reference color model. Starting from the final location in the previous frame, the algorithm proceeds iteratively at each frame so as to minimize the distance between the search window and the target.

III. Implementation

A. Hardware and Software Setup

Our tracking system built to implement CAMSHIFT consists of one active camera with pan, tilt, and zoom (PTZ) capability, a frame grabber, and a personal computer. The PTZ camera captures and follows the scene with moving targets, the frame grabber is responsible for real-time transfer of full resolution color images or sequences of images to computer memory, and the computer supports image data processing and the display of results.

The Sony EVI-D100 PTZ camera (shown in Fig. 4) features 380,000 effective picture elements and $1/4$ type super CCD (Charge-Coupled Device). The camera is capable of pan, tilt, and zoom that can be controlled by the computer via an IN-VISCA cable. One port of the cable connects to a RS-232C jack (a popular serial port) on the computer and the other port connects to the VISCA IN jack on the camera. The VISCA OUT jack is used for a multi-camera system, which can be connected with the VISCA IN port of another camera in serial, so the computer can send commands to, and get acknowledgement from, several cameras for pan, tilt, and zoom control.

The frame grabber used in the system is Euresys Picolo PCI Frame Grabber (shown in Fig. 5). The grabber card is fit into the PCI (Peripheral Component Interconnect) slot and provides the plug where the S-Video cable connects to the PC. The other port of the cable is connected to the S Video jack on the camera. The card supports the acquisition and real-time transfer of full resolution color images or sequences of images to the PC memory. The square-pixel resolution (640 x 480 or 768 x 576) is achieved at the full frame rate (30 or 25 frame per second) for raw image refreshment. Once connected, the camera can output to the computer the video stream through the S Video cable in NTSC format (National Television System(s) Committee).

As to software setup, our system utilizes OpenCV [11] (Open Computer Vision) as a tool for assisting image data processing and analysis. It provides many high level functions to realize image data manipulation, video input and output, various dynamic data structures, and motion analysis.

B. Realization and Results

The tracking system takes as the input an initial window containing the target, and the window size and position are

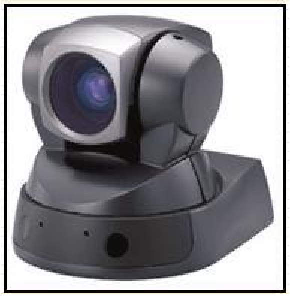


Fig. 4. Sony EVI-D100 PTZ Color Video Camera

Fig. 5. Euresys Picolo PCI Frame Grabber

defined by the user. Processed by CAMSHIFT, the search window is updated to incorporate and fit the target in a rectangle. The algorithm is iterated until the user chooses to track a new target in the scene through mouse input. Then the whole process is repeated, tracking a new target. Apart from CAMSHIFT, two additional issues should be considered before a robust tracking system with an active camera is built up. One issue is how to control the pan and tilt of the camera when the arbitrarily moving target tends to go beyond the current camera view. The other issue is how to deal with the situation when, for some reason, the object no longer exists in the scene.

In order to solve the first issue, the system should be attentive of the target position and its velocity. When any border of the search window is within a preset distance from the borders of the view, the camera is mechanically activated. The amount and the direction of panning or tilting the camera in each time step are determined by the current target velocity in the x and y directions. If $c_n(x_n, y_n)$ and $c_{n-1}(x_{n-1}, y_{n-1})$ denote the search window centers of the current and last time step, then the current target velocity, v_n, is the vector

$$v_n = (x_n - x_{n-1}, y_n - y_{n-1}). \tag{6}$$

The component of v_n along the x direction is $x_n - x_{n-1}$, whose sign decides the pan direction and whose magnitude decides the pan amount. The tilting direction and amount are obtained from $y_n - y_{n-1}$ using the same method.

The pan and tilt command format that Sony EVI-D100 recognizes is shown in Fig.6. In the first bit, x determines which camera will receive the command. VV is the pan speed in the range from 01 to 18, and WW is the tilt speed in the range from 01 and 14. YYYY is the pan position between FA60 to 05A0 (-1440 to 1440 decimal), and ZZZZ is the tilt position between FE98 to 0168 (-360 to 360 decimal). The command is sent from the computer to the camera through the RS-232C port. A pan-tilt command is a 15-byte command, and a writing function is called fifteen times to send the command to the camera.

The target will disappear from the scene when it moves too fast for the camera to adjust the view, which is the second issue mentioned in the first paragraph of this subsection. If so, ideally, pixel values in the backproject image will be all zero, because the probability of having the target in the scene is zero. In a real implementation, however, the overall black backproject image is covered by small white dots due to the inevitable noise inherent with the image acquisition of the camera. However, the noise is a small portion of the whole image, and the backproject image can still be treated as black with constant pixel value $I(x, y) = c$ over all (x, y) positions. According to the CAMSHIFT algorithm in (1), (2), and (3), x_c and y_c, in this case should be

$$\begin{aligned} x_c = \frac{M_{10}}{M_{00}} &= \frac{\sum_x \sum_y xI(x,y)}{\sum_x \sum_y I(x,y)} = \frac{c\sum_x Nx}{cMN} \\ &= \frac{cMN\frac{(M+1)}{2}}{cMN} = \frac{M+1}{2}, \end{aligned} \tag{7}$$

$$y_c = \frac{N+1}{2}, \tag{8}$$

where M and N are the height and width of the image. Therefore, the search window center is the same as the image center, and the window size will expand to the image size. However, the effect is not desirable because, reasonably, the search window size should be zero when the target is not in the scene. For this reason, CAMSHIFT is revised to show only a point in the image center when the target is out of the scene, in which case, the average intensity of the backproject image should be smaller than a threshold value.

8x	01	06	02	VV	WW	0Y	0Y	0Y	0Y	0Z	0Z	0Z	0Z	FF

Fig. 6. Pan and Tilt Command Data Format

Fig.7 shows the flowchart of the program run in the tracking system, with both the consideration of the camera view shifting and absence of the target in the scene. Fig.8 shows a series of snapshots when targets are being tracked. Images (a) - (c) are tracking segments of a green object with proper view shifting to fit the object in the scene, (d) - (f) are the cases of hand tracking, and (g), having a red spot in the center, shows the situation when the scene is absent from the object.

IV. Discussion and Future Work

Color, as a distinct feature, is superior to edges or corners for providing a clear representation of the target. Edges or corners are such general features that it is not easy to sperate them from the background. CAMSHIFT, a color-based tracking technique, is of less computational cost than other elaborate feature extraction methods, such as tracking contours with snakes ([12] [13] [14]), using Eigenspace matching techniques [15], maintaining large sets of statistical hypotheses [16], or convolving images with masks to detect features [17]. Edge or corner detection usually involves filtering, an image processing methods of high calculation burden. Moreover, with its simple algorithm, CAMSHIFT indicates not only where the target is, but also it indicates to which direction the target is moving. The movement direction is obtained from the difference between the center of the current and the last search window. The calculation is general and straightforward, and it avoids using filtering masks, pixel by pixel, around the image to calculate the local gradient. For an attentive tracking system, tracking speed is a key factor impacting the overall system performance. In consideration of both accuracy and

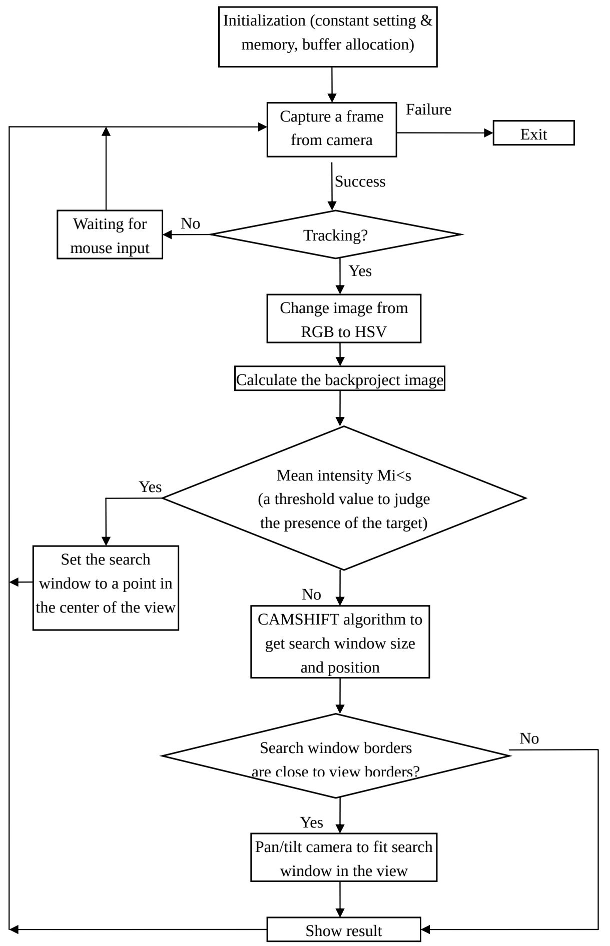

Initialization (constant setting & memory, buffer allocation)
Capture a frame from camera
Failure
Exit
Success
Tracking?
No
Waiting for mouse input
Yes
Change image from RGB to HSV
Calculate the backproject image
Mean intensity Mi<s (a threshold value to judge the presence of the target)
Yes
Set the search window to a point in the center of the view
No
CAMSHIFT algorithm to get search window size and position
Search window borders are close to view borders?
No
Yes
Pan/tilt camera to fit search window in the view
Show result

Fig. 7. Program Flow Chart of Color-based Tracking Using the CAMSHIFT Algorithm

computational cost, CAMSHIFT is a desirable candidate for an attentive tracking system.

Apart from the tracking algorithm, the reaction speed of the active camera also affects the performance of the attentive tracking system. Due to the mechanical constraints and legacy, the execution of camera panning and tilting are always slower than the change of electrical signals. Therefore, if the target is moving faster than the reaction speed of panning and tilting, the system will lose track of the target. In such condition, CAMSHIFT is modified by reducing the search window to a point located at the center of the scene. This modification will prevent the camera from panning and tilting irregularly and unceasingly to search for the object.

The single-camera, visual tracking system is a foundation for building a more complex system with multiple active cameras. One possible architecture can be a two-camera system, in which one camera is fixed to get a general view of the scene, and the other is active with pan, tilt, and zoom capabilities. When the target appears in the view of the fixed camera, the second camera will track it by panning, tilting, or zooming accordingly, and obtain detailed information of the target that is not provided by the fixed camera. More interesting work on camera calibration and noise reduction will be dealt with to set up the complex tracking system.

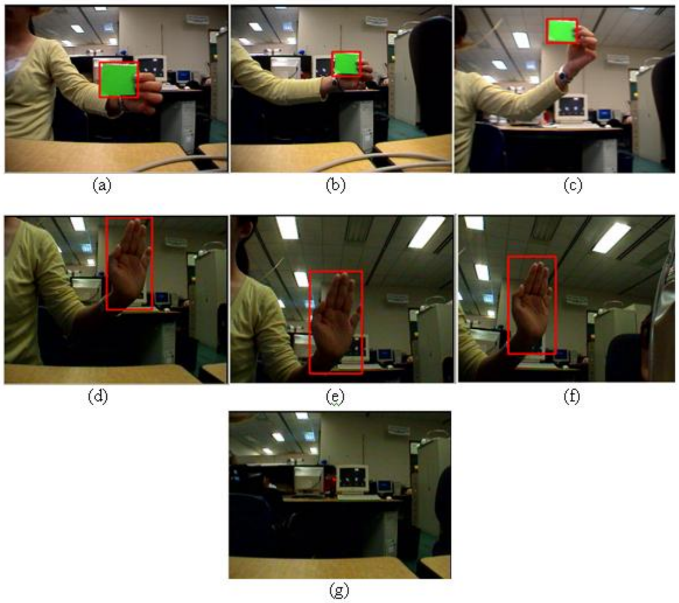


Fig. 8. Snapshots of Tracking Series

V. Conclusion

CAMSHIFT is a color-based tracking algorithm with a high performance-cost ratio. It has been implemented in our single, active camera tracking system with satisfactory tracking results and speed. Modifications have been made to CAMSHIFT in order to deal with the situation when the target is out of the scene. Visual tracking is still expanding and enriching its territory despite the wide and versatile applications. Our system can be treated as a stepping stone for the exploration into more complex systems equipped with advanced and compound tracking techniques.

Acknowledgment

The authors would like to thank Eli Gibson and Jack Cha for sharing their knowledge about building a tracking system using active cameras.

References

[1] A.Prez, M.L.Crdoba, A.Garca, R.Mndez, M.L.Muoz, J.L.Pedraza, and F.Snchez. *A Precise Eye-Gaze Detection and Tracking System*, In WSCG POSTERS proceedings, Plzen, Czech Republic, February 3-7, 2003.

[2] M. Yokoyama and T. Poggio. *A Contour-Based Moving Object Detection and Tracking*, In Proceedings 2nd Joint IEEE International Workshop on VS-PETS, Beijing, October 15-16, 2005.

[3] M. Greenspan, L. Shang, and P. Jasiobedzki. *Efficient Tracking with the Bounded Hough Transform*, CVPR04: IEEE Computer Society International Conference on Computer Vision and Pattern Recognition, vol. 1, Washington D.C., 27th June - 2nd July 2004, pp 520-527.

[4] G. R. Bradski. *Computer Vision Face Tracking For Use in a Perceptual User Interface*, Intel Technology Journal Q2'98, USA, 1998.

[5] D. Comaniciu, V. Ramesh, and P. Meer. *Real-time tracking of non-rigid objects using mean shift*, Proc. of 1st Conf. Comp. Vision Pattern Recogn., 2000.

[6] A. M. McIvor. *Background Subtraction Techniques*, Reveal Ltd, Remuera, Auckland, New Zealand, 2001.

[7] B. Lucas and T. Kanade. *An iterative image registration technique with an application to stereo vision*, Proc. 7th Internat. Joint Conf. on Artificial Intelligence, (IJCAI), 1981.

[8] J.M. Rehg, T. Kanade. *Model-based tracking of self-occluding articulated objects*, iccv, p. 612, Fifth International Conference on Computer Vision (ICCV'95), 1995.

[9] D. Murray, A. Basu. *Motion Tracking with an Active Camera*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 16, no. 5, pp. 449-459, May, 1994.

[10] H. Wu, Q. Chen, H. Oike, C. Hua, T. Wada, and T. Kato. *High Performance Object Tracking System Using Active Cameras*, Faculty of Systems Engineering, Wakayama University, Wakayama City, Wakayama, 640-8510, Japan, 2005.

[11] *OpenCV Reference Manual*, Intel Corporation, Issued in U.S.A., 1999-2001.

[12] K. Sobottka and I. Pitas. *Segmentation and tracking of faces in color images*, Proc. Of the Second Intl. Conf. On Auto. Face and Gesture Recognition, pp. 236-241, 1996.

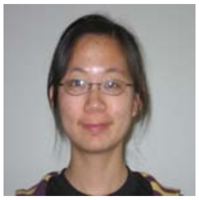
[13] M. Kass, A. Witkin D.Terzopoulos, *Snakes: Active contour Models*, Int. J. o f Computer Vision (1) #4, pp. 321-331, 1988.

[14] C. Vieren, F. Cabestaing, J. Postaire. *Catching moving objects with snakes for motion tracking*, Pattern Recognition Letters (16) #7, pp. 679-685, 1995.

[15] A. Pentland, B. Moghaddam, T. Starner. *View-based and Modular Eigenspaces for face recognition*, CVPR94, pp. 84-91, 1994.

[16] M. Isard, A. Blake. *Contour tracking by stochastic propagation of conditional density*, Proc. 4th European Conf. On Computer Vision, Cambridge, UK, April 1996.

[17] T. Maurer, and C. von der Malsburg. *Tracking and learning graphs and pose on image sequence of faces*, Proc. Of the Second Intl. Conf. On Auto. Face and Gesture Recognition, pp. 176-181, 1996.



Yan Lu received the B.Eng. degree in Electrical Engineering from Shanghai University, Shanghai, P.R.China, in 2005. She worked in Intel Products (Shanghai) as a product engineer. Since January 2006, she has been pursuing her M.A.Sc degree at School of Engineering Science in Simon Fraser University, Burnaby, B.C., Canada. Her research interests include visual tracking, camera calibration, and image analysis.

Implementation of Radio Link Control/Medium Access Control in a GPRS model

Renju Narayanan and Ljiljana Trajković

Abstract— **In this paper, we describe a General Packet Radio Service (GPRS) OPNET simulation model and the implementation of the Radio Link Control/Medium Access Control (RLC/MAC). The RLC/MAC protocol is added to an existing GPRS OPNET model. We have enhanced the existing model by implementing the unacknowledged mode of RLC and two-phase access mechanisms. We have verified the effect of the new implementation on the end-to-end delay and cell update mechanism by performing OPNET simulations.**

***Index Terms*— GPRS, RLC/MAC, cell update, network simulation.**

I. Introduction

GENERAL Packet Radio Service (GPRS) is a packet-switched service based on the Global System for Mobile Communications (GSM), an extensively deployed voice technology . GSM provides data transmission rates of 9.6 kbps, which are inefficient for variable bit rate data services such as web browsing and email [1], [2]. In order to provide higher data rates and efficient data transmission, the European Telecommunication Standards Institute (ETSI) introduced the GPRS over the existing GSM infrastructure. GPRS, considered as a 2.5 G cellular network, may offer data rates up to 171.2 kbps. In GPRS, resources (radio channels) are allocated to users on demand and hence, billing is based on the amount of data transferred rather than on the connection time.

In this paper, we describe the implementation of the Radio Link Control/Medium Access Control (RLC/MAC) protocol in an existing OPNET GPRS simulation model. The existing model contains the implementation of the following GPRS communication-specific protocols: Subnetwork Dependent Convergence Protocol (SNDCP) [3], GPRS Tunneling Protocol (GTP) [3], Mobile Application Part (MAP) [4], and Logical Link Control (LLC) [5]. The cell update procedure [5] and the Base Station Subsystem GPRS Protocol (BSSGP) are also implemented in the existing model. The implementation of the RLC/MAC protocol enables the efficient allocation of radio resources to the users.

This paper is organized as follows. In Section II, we provide an overview of GPRS. We describe the OPNET implementation of RLC/MAC protocol in Section III, and the simulation scenarios and the results validating the implementations in Section IV. Finally, we conclude with Section V.

II. GPRS Overview

A. *System Architecture*

The system architecture of GPRS is shown in Fig. 1. In order to enable GPRS services in the existing GSM infrastructure, several modifications are made to the existing nodes, and two new nodes are introduced [1], [2], [6]: Serving GPRS Support Node (SGSN) and Gateway GPRS Support Node (GGSN). The various nodes shown in the architecture are explained below.

The Mobile Station (MS) consists of a Mobile Equipment (ME) and a Subscriber Identity Module (SIM), and in addition to voice data, these MSs support packet data. MSs that support GPRS may be classified as follows: Class A, Class B, and Class C. Class A MSs simultaneously support the GSM and GPRS services, whereas, Class B and Class C MSs support either GSM or GPRS services at a given time. For Class B MSs, the ongoing GPRS services may be suspended to initiate or receive GSM services. However, Class C MSs must explicitly disconnect from the ongoing GPRS services to enable GSM services.

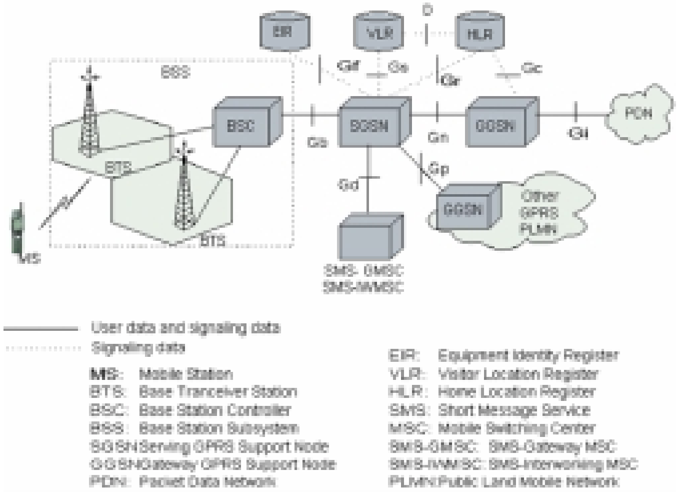

EIR
VLR
D
HLR
BSS
Gf
Gs
Gr
Gc
BSC
Gb
SGSN
Gn
GGSN
Gi
PDN
BTS
Gd
Gp
GGSN
Other GPRS PLMN
BTS
MS
SMS- GMSC
SMS-IWMSC
User data and signaling data
Signaling data
MS: Mobile Station
BTS: Base Tranceiver Station
BSC: Base Station Controller
BSS: Base Station Subsystem
SGSN Serving GPRS Support Node
GGSN Gateway GPRS Support Node
PDN: Packet Data Network
EIR: Equipment Identity Register
VLR: Visitor Location Register
HLR: Home Location Register
SMS: Short Message Service
MSC: Mobile Switching Center
SMS-GMSC: SMS-Gateway MSC
SMS-IWMSC: SMS-Interworking MSC
PLMN Public Land Mobile Network

Fig. 1. GPRS system architecture. Shown are data and signaling paths and GPRS interfaces between various network nodes.

Manuscript received June 29, 2006.

Renju Narayanan and Ljiljana Trajković are with Simon Fraser University, Burnaby, BC, V5A 1S6, Canada. (E-mail: {rsn, ljilja} @cs.sfu.ca).

The Base Station Subsystem (BSS) consists of a Base Station Controller (BSC) and one or more Base Transceiver Stations (BTSs). A logical entity, known as Packet Control Unit (PCU), to manage RLC/MAC functions is also introduced in the system, and this entity may be located at the BTS, BSC, or SGSN. The SGSN exchanges messages between MSs within its service area and GGSN. Its functions include authentication, ciphering, session management, mobility management, logical management, and billing. The GGSN acts as a gateway between the GPRS system and the external Packet Data Networks (PDNs) (IP or X.25 networks).

The GPRS system employs various registers to store information regarding subscribers and the ME. The Home Location Register (HLR) stores the subscriber information, the current SGSN address, and the Packet Data Protocol (PDP) addresses for each user in the Public Land Mobile Network (PLMN). Visitor Location Register (VLR) stores the current location and related information of a visiting subscriber. Equipment Identity Register (EIR) stores information regarding the ME.

B. GPRS Protocol Stack

The GPRS protocol stack for user data transmission is shown in Fig. 2. Um (air interface), Gb, and Gn are the interfaces located between MS and BSS, BSS and SGSN, and SGSN and GGSN respectively. The SNDCP protocol encapsulates the IP packets in GPRS specific packet formats, and the LLC layer provides a reliable logical link, independent of the underlying radio interface protocols, to these data units. The GTP tunnels the user data between the two GSNs in the GPRS backbone network [7]. The Base Station Subsystem GPRS Protocol (BSSGP) layer conveys routing and QoS-related information between the BSS and the SGSN. The RLC layer provides a reliable radio link for data transfer between the MS and the BSS. The MAC layer controls the multiplexing of signaling and data messages from various GPRS users. The GSM RF (Radio Frequency) layer controls the physical channel management, modulation, demodulation, transmission, power control, and channel coding/decoding.

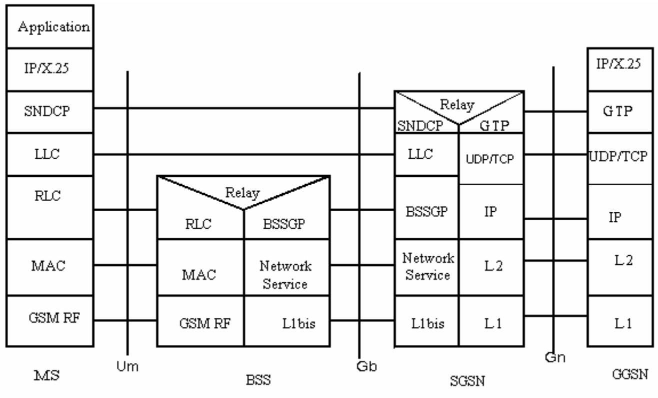


Fig. 2: GPRS transmission plane protocol stack.

C. Air Interface

The air interface provides radio channel connection between an MS and the BTS [8], [9], [10]. GPRS employs distinct frequencies in the uplink (radio link from MS to BTS) and the downlink (radio link from BTS to MS) directions, and a combination of frequency division and time division multiple access (FDMA and TDMA) schemes to allocate radio resources (physical channels). GPRS employs a 52-frame multiframe structure: each multiframe consists of 52 TDMA frames, and four TDMA frames constitute a radio block. Each TDMA frame consists of eight time slots. The Protocol Data Units (PDUs) exchanged between the RLC/MAC entities in the MS and the BTS are called RLC/MAC blocks, and each PDU is transmitted in the same time slot over four continuous TDMA frames (in one radio block). In order to provide higher throughputs, an MS supporting GPRS may transmit or receive in several time slots of a TDMA frame. This capability is indicated by the multislot class of the MS [7].

In GPRS, the physical channels used for packet logical channels are called Packet Data Channels (PDCHs). GPRS employs two types of PDCHs as shown in Fig. 3: traffic and control. The PDCHs used to transfer data during uplink or downlink transmission are called the Packet Data Traffic Channels (PDTCHs). The control channels may be further classified as follows: broadcast, common, and dedicated. The Packet Broadcast Control Channel (PBCCH) broadcasts information related to the serving BTS and the neighboring BTSs. The Packet Common Control Channel (PCCCH) consists of a Packet Random Access Channel (PRACH) used for random access, a Packet Access Grant Channel (PAGCH) used for notifying the MS about access grant, and a Packet Paging Channel (PPCH) used for paging [7]. The Packet Associated Control Channel (PACCH) is used to carry signaling messages during uplink or downlink data transfer. The Packet Timing Control Channel (PTCCH) is used to send timing advance information. The PDTCHs employ four coding schemes: CS-1, CS-2, CS-3, and CS-4, providing data rates of 9.05 kbps, 13.4 kbps, 15.6 kbps, and 21.04 kbps, respectively. Coding schemes CS-1 to CS-4 are mandatory for MSs supporting GPRS, whereas, coding scheme CS-1 is mandatory for the GPRS network.

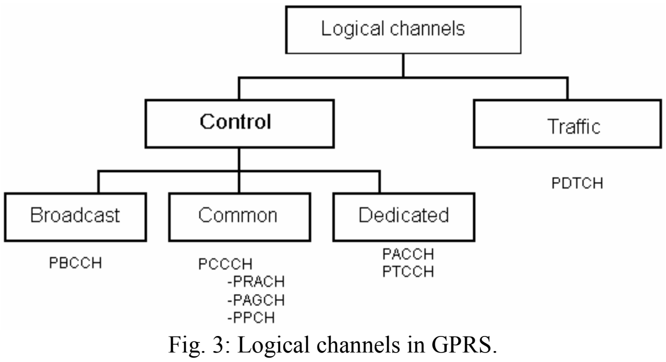


Fig. 3: Logical channels in GPRS.

D. RLC/MAC Procedures

RLC layer segments the LLC PDUs into RLC/MAC blocks and reassembles them [11]. The RLC protocol provides acknowledged and unacknowledged modes of operation. In acknowledged mode, it performs the Backward Error Correction

(BEC) procedures to enable selective retransmission mechanism. The MAC protocol enables multiple MSs to share a common transmission medium, and provides contention resolution for the data transfers originated by the MSs. In order for an MS (or BSS) to transfer data in the uplink (or downlink) direction, a physical connection called Temporary Block Flow (TBF) is established between the two RLC/MAC entities. The TBFs are unidirectional and are established only for the period of data transfer, after which they are released. The BSS assigns a unique Temporary Flow Identity (TFI) to each TBF [11].

GPRS supports three medium allocation modes [7], [11]:

- Fixed allocation: The BSS assigns a fixed allocation of radio blocks and PDCHs to the MS using bitmaps.
- Dynamic allocation: The BSS assigns radio blocks to MSs on a block-by-block basis. An Uplink State Flag (USF) is assigned to the MS for each allocated block.
- Extended dynamic allocation: The BSS assigns USF for a PDCH, but the MS is allowed to transmit not only in that PDCH, but also in all the higher numbered PDCHs.

The GPRS network may support either fixed allocation mode or dynamic allocation mode.

1) Uplink TBF Establishment

GPRS employs two mechanisms for establishing uplink TBF: one-phase access and two-phase access procedures. Even though an MS may request either one of the procedures, the BSS decides the procedure for TBF establishment.

a) One-phase Access Procedure

In the one-phase access procedure, the MS sends a "packet channel request" message to the network indicating the radio priority and the number of resources required. Then, the MS waits for an "uplink assignment" message from the BTS, which contains the time slot and the physical channel allocated to it. When the MS receives the uplink assignment message, it starts sending data including its Temporary Logic Link Identity (TLLI) in the first few blocks, until it receives an uplink ACK/NACK from the network. If the uplink ACK/NACK contains the TLLI of the MS, then the contention is resolved, and it continues sending data. Otherwise, the MS stops sending data, and repeats the packet access procedure. The one-phase access procedure is shown in Fig. 4.

b) Two-phase Access Procedure

In the two-phase access procedure, similar to one-phase access procedure, the MS first sends a "packet channel request" and waits for the "packet uplink assignment" message. However, the MS does not indicate the required number of resources in the channel request message. The network then sends the uplink assignment indicating that only one radio block is allocated. On receiving the uplink assignment, the MS sends a "packet resource request" message indicating its TLLI and the number of resources required. Then, the network assigns the resources according to the request and the available resources, and sends an uplink assignment message including the TLLI to the MS. When the MS receives the uplink assignment with its TLLI, the contention is resolved. The two-phase access procedure is shown in Fig. 5.

2) Downlink TBF Establishment

When the BSS receives an LLC PDU from the SGSN, it initiates the establishment of a downlink TBF by sending a "packet downlink assignment" message to the MS. When the MS responds with an acknowledgement message, the BSS commences sending the PDUs.

3) TBF Release

In the uplink data transfer, the MS performs a count down procedure to indicate the end of a TBF. When it starts sending the last sixteen blocks, it starts decrementing the Countdown Value (CV) at each transmission. The last block of the TBF is sent with CV equals zero. When the BSS receives the last block, it sends a "packet uplink ACK/NACK" message to confirm the release of the TBF, and the MS responds with a "packet control acknowledgement" message releasing the TBF. However, in the case of downlink data transfer, the BSS can easily anticipate the end of a TBF. When the BSS sends the final data block, it sets a flag indicating the end of the TBF. The MS replies with a "packet downlink ACK/NACK" message, and releases the TBF.

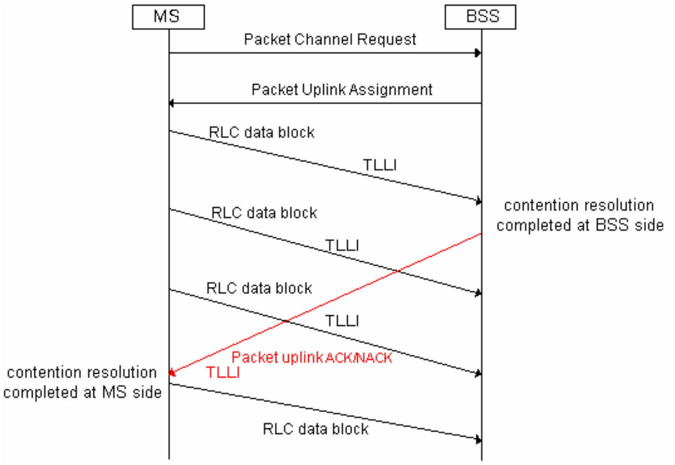


Fig. 4: One-phase access procedure and contention resolution.

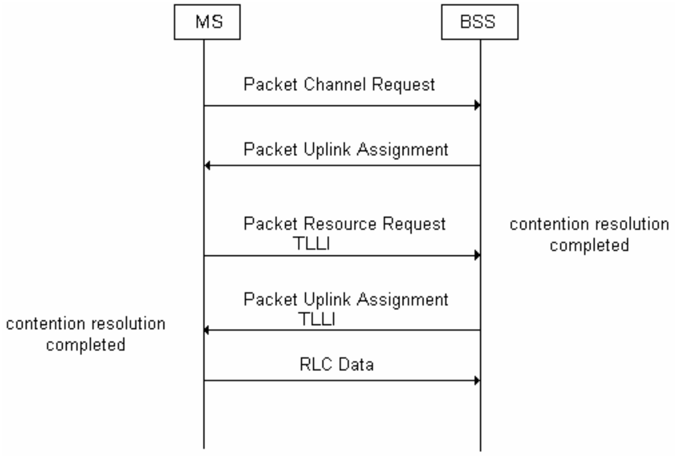


Fig. 5: Two-phase access procedure and contention resolution.

E. Cell Update

When an MS that is attached to an SGSN, moves between coverage areas of BTSs, it performs the cell update. The cell update is performed based on the received signal level (RXLEV) measurements performed by the MS in the network. The MS periodically measures the RXLEV from the BTS in the serving cell and in the neighboring cells. Three cell update modes have been defined [7]:

- NC0: The MS performs autonomous cell reselection and does not send RXLEV measurement reports to the network.
- NC1: The MS performs autonomous cell reselection and periodically sends RXLEV measurement reports to the network.
- NC2: The network controls the cell reselection, and the MS sends the RXLEV measurement reports to the network.

III. OPNET IMPLEMENTATION

We developed a simulation model for GPRS using the OPNET [12] network simulator. Unlike the two described upgrades of the GPRS OPNET contributed model [13], [14], the model described in this paper contains explicit implementation of GPRS-specific protocol layers. The basic GPRS model shown in Fig. 6 includes models for MS, BTS, BSC, SGSN, GGSN, HLR, and a sink. The sink represents the external PDN, and hence, the data flow in this model is unidirectional. However, the signal flow is bidirectional. Only class C MSs in GPRS mode have been modeled, and the MSs in the developed model support single slot operation and raw traffic generation.

Even though an MS measures RXLEV from the BTS of its serving cell and from the neighboring cells, it only stores the information for the six most powerful BTSs [7]. Hence, the developed model supports only six BTSs. Each cell has only one BTS, and each BTS has a coverage area in the range of 15–20 km. The GPRS model supports cell update in the NC0 mode (autonomous). The model supports GPRS Mobility Management (GMM) signaling procedures such as Attach, Activate, Detach, and Deactivate [3].

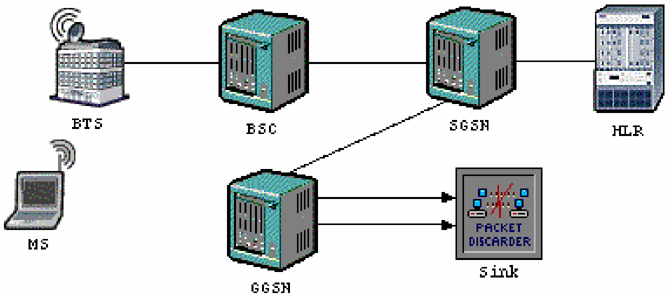


Fig. 6. An example of an OPNET GPRS model connected to an external PDN (sink).

We have implemented the RLC/MAC layer in the MS and BTS models of the existing GPRS model. The following features are implemented in the RLC/MAC layer: unacknowledged mode of RLC, fixed allocation medium access mode, two-phase access procedure, and CS-1 coding scheme. The node model for the MS showing the RLC/MAC node and the Power_Monitor nodes is shown in Fig. 7. In the node model, the first six channels in the receiver are dedicated to receive the PBCCH information from the BTS. The uplink frequency corresponding to the PBCCH frequency is considered as the PRACH frequency, and the MS has a dedicated channel for sending packet channel requests. The *Power_Monitor* node receives the PBCCH information from the BTSs and measures the power of the received messages, and selects the BTS with the highest power level as the serving BTS.

The RLC/MAC process model for the MS is shown in Fig. 8. The variables and buffers are initialized in the *init* state, and the process remains in the *idle* state until it receives a packet from either the BTS or the LLC layer. When the RLC layer receives a higher layer packet, it segments the packets into RLC/MAC blocks and buffers them (*pkt_encap)*. The MS then initiates the packet access procedure by sending a “packet channel request” message and waits for an uplink assignment (*pkt_access*). In *Resource_req* state, when a “packet uplink assignment” message is received, the MS sends a “packet resource request” message. When the MS receives an uplink assignment, it verifies the TLLI included in the message for contention resolution. If contention is resolved, the process waits until its assigned time in *TBF_wait* state and then commences sending data (*send*). When the data block with CV equals zero has been sent, the process enters a forced state (*T3182*) and wait for an “uplink Ack” message. When the ack is received, the process enters *TBF_release* and releases the resources.

The node model for BTS is shown in Fig. 9. The PBCCH_source sends the PBCCH information to the MSs. The RLC/MAC is implemented as a dynamic process. The parent process invokes appropriate child process upon receipt of packets from the MS or the BSC. The parent and child processes are shown in Fig. 10 and Fig. 11, respectively. We have implemented single slot operation of MSs only. The BTS employs a first-in-first-out (FIFO) mechanism to allocate resources to the MSs.

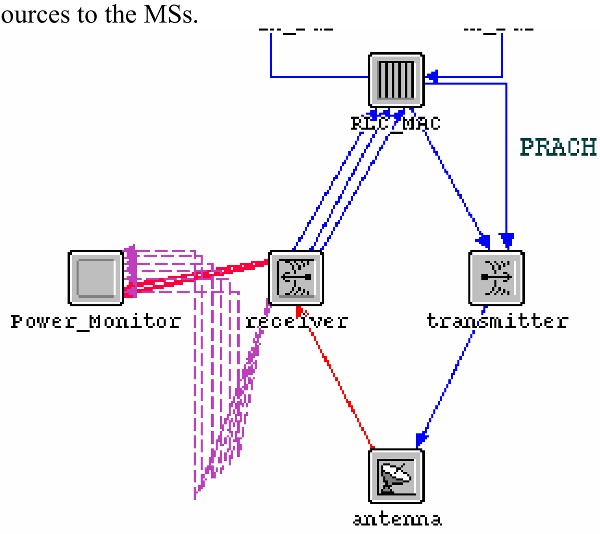


Fig. 7: Node model for MS.

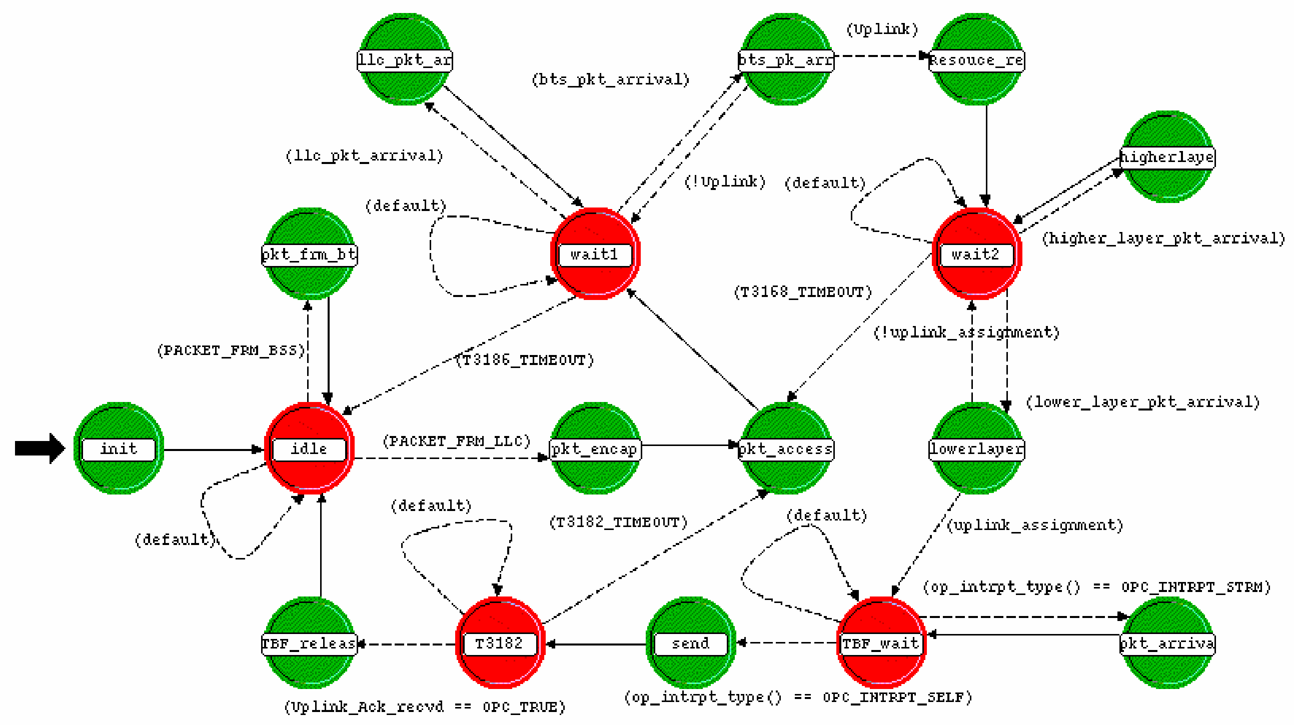


Fig. 8: RLC/MAC process model for MS.

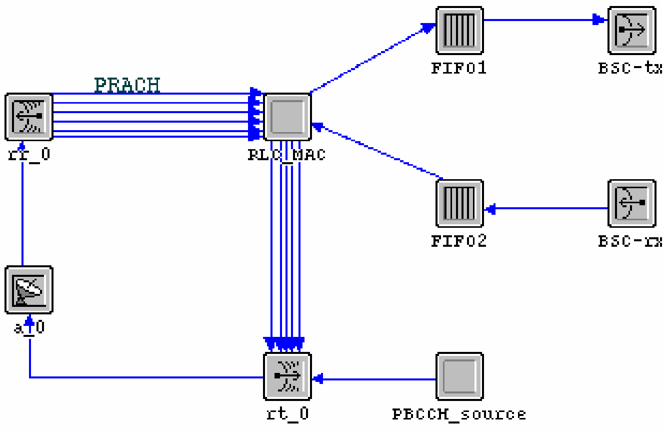


Fig. 9: Node model for BTS.

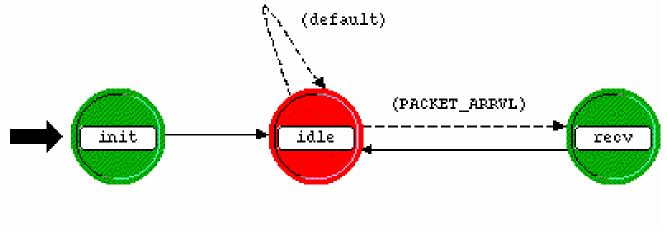


Fig. 10: RLC/MAC process for BTS (parent).

IV. SIMULATION SCENARIOS AND RESULTS

We simulated two scenarios to verify the implementation of RLC/MAC. In the first simulation scenario, we compared the end-to-end delay experienced by a packet with and without the RLC/MAC protocol. This scenario consists of two MSs and a BTS, and the MSs transmit data throughout the 10 minutes of simulation time. The end-to-end delay experienced by packets originating from MSs, shown in Fig. 14, is higher (approximately 10%) in the case of GPRS model with RLC/MAC because of the buffering of data and the higher number of signaling messages.

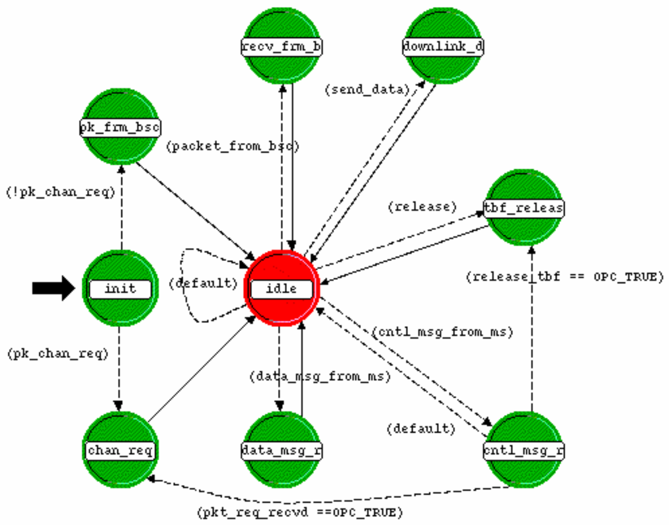


Fig. 11: RLC/MAC process for BTS (child).

In order to verify the cell update mechanism, we simulated a scenario, shown in Fig. 13, where an MS performs the cell update. At the beginning of the simulation, the MS, *mobile_node_1*, is in the coverage area of *Base_Station_0.* As the simulation progresses, *mobile_node_1* moves into the coverage area of *Base_Station_1* and performs the cell update. The throughput (number of packets correctly received or transmitted at the transceiver) statistics shown in Fig. 14 verifies that, at the beginning of the simulation, *mobile_node_1* was transmitting to *Base_Station_0* and later changed transmission to *Base_Station_1*.

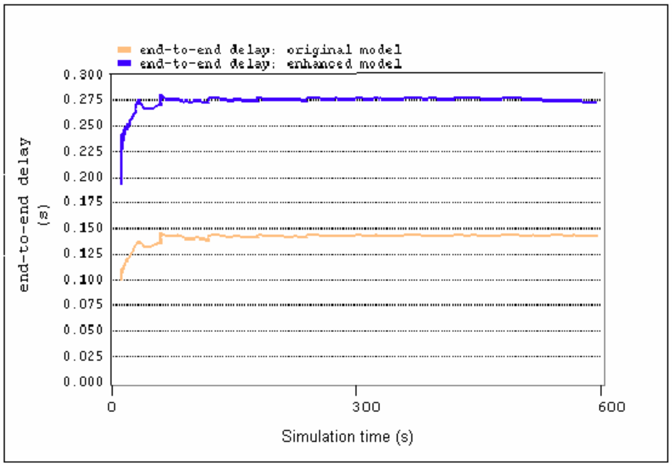


Fig. 12: Comparison of end-to-end delays.

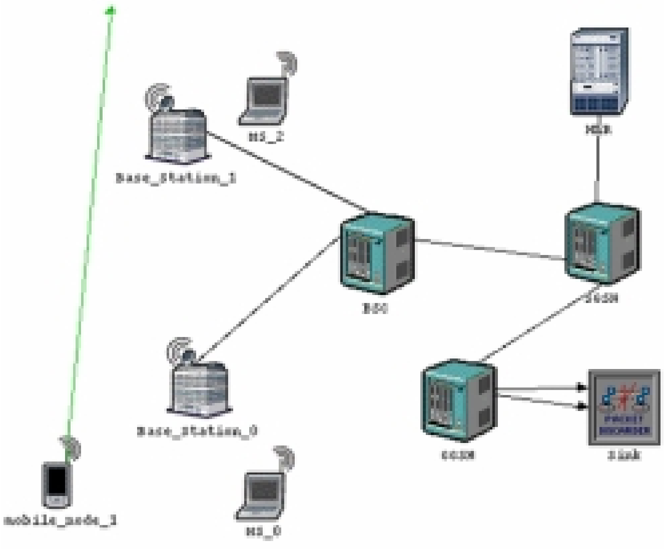


Fig. 13: Simulation scenario for cell update.

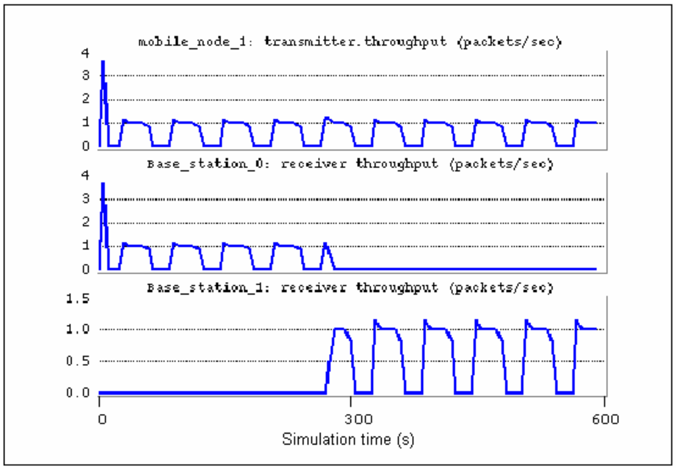


Fig. 14: Throughput at the transmitters and receivers of BTSs and MS.

V. CONCLUSION

In this paper, we described an OPNET model for GPRS, which contains the implementation of various GPRS-specific protocols. In addition, the model includes the implementation of GMM signaling procedures and the cell update procedure. We described the implementation of RLC/MAC layer and presented various simulation scenarios and results that validate the implementation. The developed model could further be enhanced by implementing various resource allocation algorithms deployed in the real world. The performance of the developed model could be evaluated by simulating real-life scenarios and using genuine traffic traces.

REFERENCES

[1] G. Sanders, L. Thorens, M. Reisky, O. Rulik, and S. Deylitz, *GPRS Networks*. Hoboken, NJ: Wiley, 2003.

[2] S. Hoff, M. Meyer, and A. Schieder, “A performance evaluation of Internet access via the general packet radio service of GSM,” *in Proc. 48th IEEE Vehicular Technol. Conf,* Ottawa, ON, May 1998, vol. 3, pp. 1760–1764.

[3] R. Ng and Lj. Trajković, “Simulation of general packet radio service network,” *OPNETWORK,* Washington, DC, Aug. 2002.

[4] V. Vukadinovic and Lj. Trajković, “OPNET implementation of the Mobile Application Part protocol,” *OPNETWORK,* Washington, DC, Aug. 2003.

[5] R. Narayanan, P. Chan, M. Johansson, F. Zimmermann, and Lj. Trajković, “Enhanced general packet radio service OPNET model,” *OPNETWORK,* Washington, DC, Aug. 2004.

[6] A. Mishra, “Performance and architecture of SGSN and GGSN of general packet radio service (GPRS),” in *Proc. IEEE GLOBECOM*, San Antonio, TX, Nov. 2001, vol. 6, pp. 3494–3498.

[7] E. Seurre, P. Savelli, and P. Pietri, *GPRS for Mobile Internet*. Norwood, MA: Artech House, 2003.

[8] C. Bettstetter, H. J. Vögel, and J. Eberspächer, “GSM phase 2+ general packet radio service GPRS: architecture, protocols, and air interface,” *IEEE Commun. Surv.*, vol. 2, no. 3, pp. 2–14, Aug. 1999.

[9] G. Brasche and B. Walke, “Concepts, services, and protocols of the new GSM phase 2+ general packet radio service,” *IEEE Commun. Magazine*, vol. 35, no. 8, pp. 94–104, Aug. 1997.

[10] J. Rendon, F. Casadevall, L. Garcia, and R. Jimenez, “Simulation model for performance evaluation of Internet applications using GPRS radio interface,” *IEEE Electron. Lett.,* vol. 37, no. 12, pp. 786–787, June 2001.

[11] 3rd Generation Partnership Project, TS 04.60 version 8.25.0 Radio Link Control/Medium Access Control.

[12] OPNET Modeler software [Online]. Available: http://www.opnet.com/products/modeler/home.html.

[13] G. Jain and P. Shekhar, “GPRS model enhancements,” *OPNETWORK,* Washington, DC, Aug. 2003.

[14] Y. Sawant , K. Sastry, R. Krishnamoorthy, and S. Taparia, “GPRS model enhancements,” *OPNETWORK,* Washington, DC, Aug. 2004.

R. Narayanan received the Bachelor of Technology degree in Electronics and Communication Engineering from the University of Kerala, India, in 2000. She is currently working towards her M.A.Sc. degree in the School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada.

From 2000 to 2002, she worked as a Software Engineer for Network Systems and Technologies, India. Her current research interest is wireless networks.

Detecting Small Slow-moving Sonar Targets Using Bottom Reverberation Coherence

Jinyun Ren, *Member, IEEE,* and John S. Bird, *Member, IEEE*

***Abstract*— The detection of small targets that appear suddenly or are moving slowly in strong bottom reverberation is a challenging problem for sonar surveillance in shallow water. Based on a new reverberation model, this paper proposes a target detection scheme that provides target sub-clutter visibility in the presence of reverberation. Experimental evidence shows that the bottom reverberation as seen by a stationary sonar is coherent, or at least partially coherent from ping to ping. Therefore, the bottom reverberation from a particular range cell is modeled as a complex signal composed of a stationary or slowly varying coherent component, plus a rapidly varying diffuse component. The coherent component is easily estimated using a recursive mean estimator and then removed by a simple subtraction so that the target need only compete with the diffuse component. Experimental results show a detection gain, as measured by the coherent-to-diffuse ratio, as high as 30dB.**

***Index Terms*— target detection, bottom reverberation coherence, small slow-moving target, sub-clutter visibility, sonar surveillance.**

I. Introduction

A challenging problem for sonar surveillance in shallow water is the detection of small, slow-moving targets embedded in strong bottom reverberation. Shallow-water environments in many regions are filled with rocks and coral, and therefore, compared with the returns of small targets, the bottom reverberation from these regions is much stronger. Conventional thresholding detection techniques are ineffective because they require that the target return be large enough to compete with the bottom reverberation.

Instead of directly detecting targets in strong bottom reverberation, most existing techniques suppress bottom reverberation before the data enter the detection process. These techniques are broadly classed as Moving Target Indication (MTI) [1], [2]. The most common configuration of MTI rejects undesired reverberation by exploiting the difference of Doppler frequency shift between the moving target and the bottom [3], [4].

However, if the target is moving very slowly (i.e., the Doppler frequency shifts of the target and the bottom are similar), and if the target is small, (i.e., the bottom reverberation level is much stronger compared with the target return), Doppler-based MTI techniques are ineffective because the bottom reverberation easily masks the target return in the frequency domain. Moreover, most of these Doppler-based MTI techniques model bottom reverberation as nonstationary, colored noise, which leads to computationally intensive detection algorithms due to the complexity of reverberation.

One non-Doppler-based MTI technique models bottom reverberation as a sum of echoes issued from the transmit signal [5]. Instead of the Doppler shift difference, the power difference between the bottom reverberation and the target return is employed to estimate and to suppress the reverberation. This algorithm distinguishes the different power levels by principle component analysis, and works when the target and the bottom have a similar Doppler shift. However, this algorithm requires a prior knowledge of target power, which limits its applications in practical systems.

We propose a detection scheme that takes advantage of ping-to-ping bottom reverberation coherence to detect small, slow-moving targets in strong bottom reverberation.[1] Ping-to-ping bottom reverberation coherence enables us to track the trajectory of the signals received from a range cell and then to suppress the reverberation by simply subtracting the estimate of the reverberation from the current return. A moving target need only compete with the residues after the subtraction for detection, and therefore sub-clutter visibility is achieved.

A similar scheme was proposed in radar to detect slow-moving targets against heavy ground clutter [6]. Experimental results showed that a significant improvement in detectability was achieved [7]. This idea was proposed again for a VHF radar in [8], where it was argued that the detectability of slow-moving targets embedded in strong clutter was improved by subtracting the coherent clutter from the received signal. However, to the best of our knowledge, this scheme has never been proposed for sonars. Historically, most platforms in sonar applications are moving, and therefore the data from two pings do not generally come from the same part of the bottom. A second reason might be the presence of a surface bounce multipath, which, if the surface is moving, would decorrelate the reverberation from one ping to the other and destroy the ping-to-ping bottom reverberation coherence.

The rest of this paper is organized as follows. In Section II, we show the trajectories of real bottom reverberation in the complex plane and provide evidence that the bottom reverberation is coherent, or at least partially coherent from ping to ping as seen by a stationary sonar. We then show how the ping-to-ping bottom reverberation coherence is employed to enhance target detection. Based on a reverberation model determined from our experimental data, we present a detection scheme to obtain sub-clutter visibility in Section III

[1] Ping-to-ping bottom reverberation coherence in our research implies that if the same waveform is transmitted many times from a given position, and if the propagation medium has not changed during these transmissions, except for noise, the signals received from a particular range cell will be similar from ping to ping, regardless of the distribution of the scatterers in that range cell. The similarity amongst the received signals includes both the amplitude and phase if the received signal is represented by its complex envelope at the base band.

that includes the following three key components: a coherent component estimator, a bottom reverberation suppressor, and a threshold estimator. In Section IV, we define a detection gain as measured by the coherent-to-diffuse ratio (CDR) and present the CDR achieved for a typical range cell. Finally, in Section V we present the results obtained from experimental data with a real target present and draw conclusions in Section VI.

II. Bottom Reverberation Coherence and Target Detection

A. Coherence Evidence

To establish ping-to-ping bottom reverberation coherence, we conducted a field experiment in Sasamat Lake in BC on February 7, 2005. A 210kHz, monostatic sonar was mounted on the bottom of the lake. Both the transmit and receive transducer have fan-shaped beams, and the beams were tilted to ensonify the bottom with the wide side. No multipath was present under this configuration. One sample of the complex envelope was collected for each range cell when a ping was transmitted. Several sets of data were recorded when the water surface was calm.

We analyzed the behavior of the bottom reverberation from a data set with 100 pings. The time duration for this data set was about 130 seconds. Fig.1 illustrates the trajectory of the 100 consecutive returns from a typical range cell in the complex plane, where the I and Q axes respectively represent the inphase and quadrature components of the received signal. The values on the I and Q axes are the digital numbers corresponding to signal voltages from a 16-bit A/D converter. We see that the bottom reverberation from this range cell follows a slowly moving trajectory. The small drift is possibly due to the slight movement of the sonar transducer, and/or the presence of small underwater currents during the collection of the data. Nevertheless, the bottom reverberation from one range cell is similar from ping to ping over a short period of time; in other words, the bottom reverberation from one range cell as seen by a stationary sonar is coherent from ping to ping to certain degree.

B. Target Detection

Given the experimental evidence, we assume the following models for the reverberation and target signals. We have shown that the bottom reverberation from a particular range cell is coherent from ping to ping to certain degree, so we model the bottom reverberation from one range cell as a complex signal composed of a stationary or slowly varying coherent component plus a rapidly varying diffuse component. The coherent component is represented in the phase diagram by an offset from the origin of the complex plane, while the diffuse component is represented as a circular contour, as illustrated in Fig.2(a). The target is assumed to be circularly symmetric Gaussian, and the presence of a target increases only the contour size of the diffuse component, as shown in Fig.2(b).

Conventional thresholding detection techniques estimate the total reverberation power and establish a threshold based on that estimate. The threshold must be large enough to encompass the entire coherent component no matter how small the diffuse component is, as illustrated in Fig.2(c). To declare a target present, the received signal must be outside the threshold contour. However, if we know the coherent component for the range cell, we can set up a threshold contour that is larger than only the diffuse component contour. The threshold is circularly symmetric around the origin of the complex plane if we can remove the coherent component, as shown in Fig.2(d). In other words, if we suppress the bottom reverberation by taking advantage of its ping-to-ping coherence, a moving target need only compete with the diffuse component of the reverberation for detection; hence sub-clutter visibility is achieved. A detailed design of a detection scheme to obtain sub-clutter visibility is presented in next section.

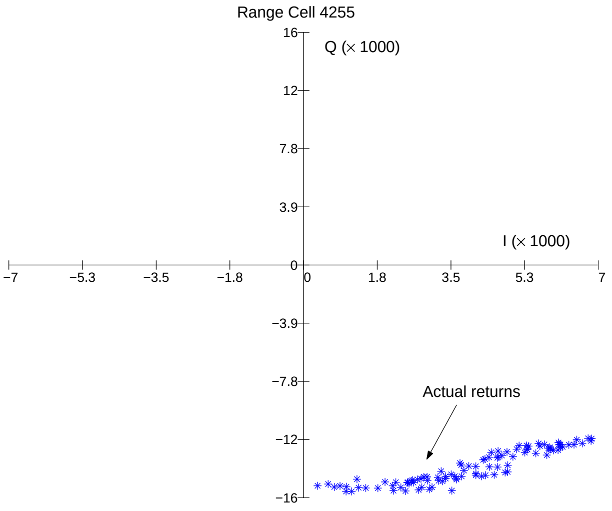


Fig. 1. Returns from one typical range cell

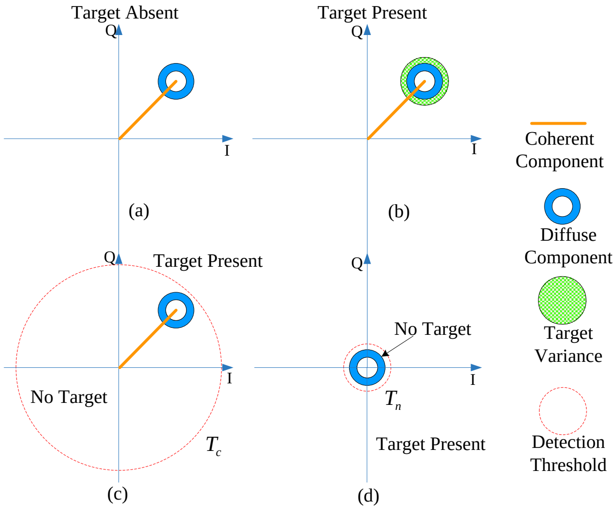


Fig. 2. Phase diagram for target detection

III. Design of a Detection Scheme for Sub-clutter Visibility

In order to design the detection scheme for sub-clutter visibility, we need to estimate both the inphase and quadrature

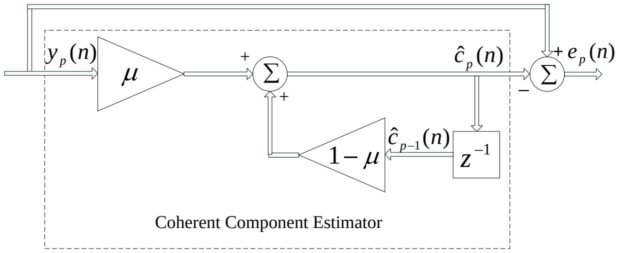


Fig. 3. Coherent component estimator and suppressor

components of the coherent component, the magnitude of the diffuse component, and the threshold. In this section, we first present the design of three key parts of the scheme: a coherent component estimator, a bottom reverberation suppressor, and a threshold estimator. We then present an implementation of the proposed detection scheme and analyze its performance.

A. Coherent Component Estimator

For a range cell, n, the bottom reverberation is modeled as

$$y_p(n) = c_p(n) + d_p(n), \tag{1}$$

where p is the ping index, n is the range cell index, y is the received signal, c is the coherent component, or the predictable part of received signals, and d is the diffuse component, or the unpredictable disturbance or noise. Note that the coherent component is a complex number that can be either constant or slowly varying but predictable, while the diffuse component is a complex random variable that is circularly symmetric Gaussian with variance σ_d^2.

A simple recursive mean estimator is adopted to estimate the coherent component for each range cell and is given by

$$\hat{c}_p(n) = (1-\mu)\hat{c}_{p-1}(n) + \mu y_p(n), \tag{2}$$

where $\hat{c}_p$ is the estimate of c when the p^{th} sample is received, and μ is the estimator parameter over the range $(0, 1)$ and is used to tune the estimator to suit the drift rate of the coherent component. The higher the drift rate, the smaller μ should be. For our data a value for μ of 0.5 produced good results. The implementation of the coherent component estimator is depicted in Fig.3, where z^{-1} represents one sample time delay.

B. Bottom Reverberation Suppressor

The strong bottom reverberation is suppressed by subtracting the estimate, $\hat{c}_p$, from the received signal at p^{th} ping, as shown in Fig.3. The residue is given by

$$e_p(n) = y_p(n) - \hat{c}_p(n), \tag{3}$$

where $y_p(n) - \hat{c}_p(n)$ is known as the bottom reverberation suppressor. When the estimator parameter, μ, is adjusted to suit the drift rate of the coherent component, the residue is reduced to be nearly the diffuse component, $d_p(n)$, alone. Fig.4 illustrates the result after the bottom reverberation suppressor is applied to the data set shown in Fig.1, where $\mu = 0.5$. Notice that the residues are around the origin in the complex plane, which means most of the bottom reverberation power is removed.

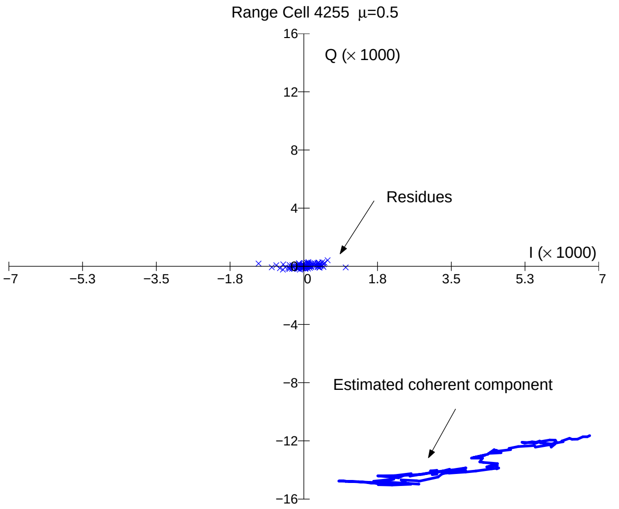


Fig. 4. Returns from a typical range cell after reverberation suppression

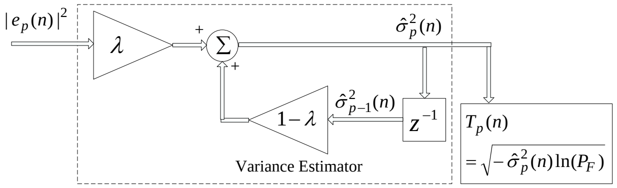


Fig. 5. CFAR Threshold Estimator

C. Threshold Estimator

For a constant false alarm rate (CFAR) detection scheme, when the signals entering the detection process are assumed to be circularly symmetric Gaussian, the threshold is given by

$$T = \sqrt{-\sigma_n^2 \ln(P_F)}, \tag{4}$$

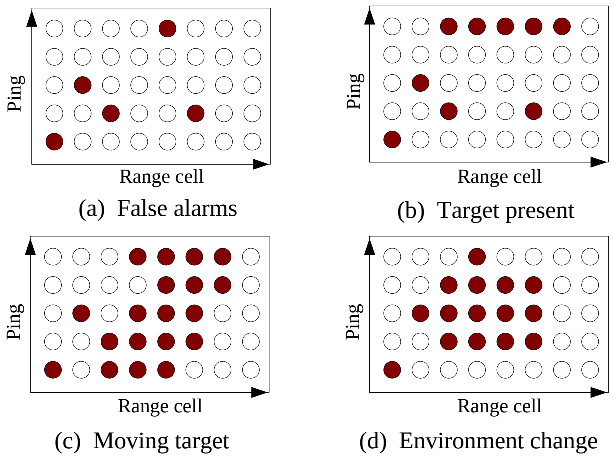
where σ_n^2 is the signal variance, and P_F is the false alarm probability required by the system. Given P_F, the threshold estimator is equivalent to a variance estimator.

The signal variance can be estimated either by taking the average of several consecutive samples or through adaptively updating the output with the latest received sample. Using the latter approach, we employ a recursive method to estimate the variance estimator given by

$$\hat{\sigma}_p^2(n) = (1-\lambda)\hat{\sigma}_{p-1}^2(n) + \lambda|e_p(n)|^2, \tag{5}$$

where $\hat{\sigma}_p^2$ is the estimated variance when p^{th} sample is received, $|e_p(n)|$ is the magnitude of the residue, $e_p(n)$, and λ is the estimator parameter ranging from 0 to 1, exclusively. Fig.5 illustrates the implementation of the variance estimator. Given each estimated variance, the threshold is obtained by using (4).

The threshold estimator updates its output with each received sample. The threshold for one range cell varies slightly when the target is absent. However, the threshold gradually adapts to a larger value when the target is present because the presence of the target increases the variance of the returned signals; on the other hand, the threshold gradually decreases when the target moves out of this range cell or becomes stationary and is part of coherent component.



(a) False alarms

(b) Target present

(c) Moving target

(d) Environment change

Fig. 6. Possible detection patterns

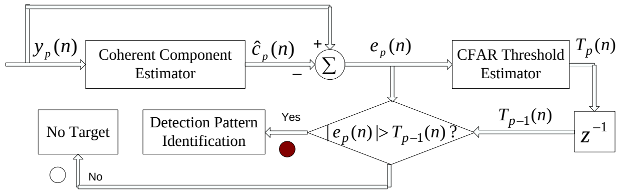


Fig. 7. Detection scheme for sub-clutter visibility

D. Detection Scheme for Sub-clutter Visibility

The detection scheme first suppresses the coherent component of the bottom reverberation and then compares the residue magnitude with latest threshold estimate. The detection result for a range cell is declared to be positive if the magnitude of its residue is greater than the threshold; otherwise, the detection result is declared to be negative. However, the detection decision from one range cell is not enough to determine whether a positive decision is due to: a false alarm; the presence of a target; or a change in the environment.

The detection scheme requires the decisions from consecutive range cells or consecutive pings to come to a conclusion. Some possible detection patterns are illustrated in Fig.6, where each circle represents the detection result for a range cell given one ping. A empty circle and a filled circle represent a negative decision and a positive decision, respectively. If positive decisions are scattered, we declare that they are only false alarms, as illustrated in Fig.6(a). If positive decisions appear for consecutive range cells, we declare that a target might be present, as in Fig.6(b). Meanwhile, if the the consecutive positive decision pattern keeps occurring for several pings and for different sets of range cells, we declare that the target is moving, as illustrated in Fig.6(c); however, if the consecutive positive decision pattern appears for several pings and then disappears, as illustrated in Fig.6(d), we might declare that the environment has changed (e.g., a rock is dropped into the water and settles on the bottom).

The schematic diagram for this detection scheme is depicted in Fig.7. Compared with the conventional thresholding detection technique, our detection scheme adds a coherent component estimator and a bottom reverberation suppressor for each range cell at the front end of the detection process. The two added parts are the key for our detection scheme to achieve sub-clutter visibility.

E. Performance of the Proposed Detection Scheme

The performance of the proposed detection scheme depends on how well the coherent component estimator tracks the bottom reverberation drift. Let us now analyze how the estimator given by (2) changes the power of its input, $y_p(n)$. Note here p represents the ping index, and n represents the range cell index. n is suppressed for convenience in this subsection since we deal with the returns from only one range cell.

We know from (1) that the received signal, y_p, consists of the coherent component, c_p, and the diffuse component, d_p. Provided that only c_p is present, the coherent component estimator output, denoted by c^o, is

$$c_p^o = (1-\mu)c_{p-1}^o + \mu c_p. \tag{6}$$

Note that the more closely μ suits the drift rate of c_p, the better c_p^o represents c_p. Similarly, provided that only d_p is present, the coherent component estimator output, denoted by d^o, is

$$d_p^o = (1-\mu)d_{p-1}^o + \mu d_p. \tag{7}$$

Note that $\hat{c}_p$ is the output of the coherent component estimator when its input contains both c_p and d_p. Therefore, the output of the bottom reverberation suppressor is represented by

$$e_p = y_p - \hat{c}_p = c_p - c_p^o + d_p - d_p^o. \tag{8}$$

Evaluating the power of e_p, we have

$$\begin{aligned} \sigma_{e_p}^2 &= E\left\{e_p e_p^*\right\} \\ &= |c_p^e|^2 + \sigma_d^2 + \sigma_{d^o}^2, \end{aligned} \tag{9}$$

where $E\{\bullet\}$ represents the expectation operation, $*$ represents the conjugate operation, $|c_p^e|^2$ is the power of the difference between the true and the estimated coherent component (i.e.,$c_p - c_p^o$), σ_d^2 is the variance of the diffuse component (i.e., d_p), and $\sigma_{d^o}^2$ is the variance of d^o, which is defined by

$$\begin{aligned} \sigma_{d^o}^2 &= E\left\{d_p^o d_p^{o*}\right\} \\ &= (1-\mu)^2\sigma_{d^o}^2 + \mu^2\sigma_d^2, \end{aligned} \tag{10}$$

or equivalently,

$$\sigma_{d^o}^2 = \frac{\mu}{2-\mu}\sigma_d^2. \tag{11}$$

Notice from (9) that the output noise power of the bottom reverberation suppressor consists of three parts: the signal error $|c^e|^2$ due to the estimation inaccuracy for the coherent component, the diffuse component power σ_d^2, and the noise power $\sigma_{d^o}^2$ from the diffuse component due to the introduction of the coherent component estimator. Amongst these three noise sources, σ_d^2 cannot be changed, but $|c^e|^2$ and $\sigma_{d^o}^2$ can be varied by adjusting the estimator parameter, μ, between 0 and 1.

According to (11), the smaller μ is, the smaller the noise introduced by the coherent component estimator. However, to minimize $|c^e|^2$, we should adjust μ to suit the bottom reverberation drift rate. Therefore, a rule of thumb for selecting

μ is to make it as small as possible provided that it can track the bottom reverberation trajectory. The output noise power is around $1.33\sigma_d^2$ plus resulting $|c^e|^2$ when $\mu = 0.5$ is applied to our experimental data. If the reduction in power due to the suppression scheme is much larger than the output noise power, the detection scheme achieves a high detection gain.

IV. Detection Gain

The performance of a detection algorithm in sonar is generally evaluated by the required signal-to-noise ratio (SNR) to achieve a given probability of detection, P_D, for an assumed probability of false alarm, P_F. The detection gain in our research is calculated as the SNR penalty of the conventional thresholding detection technique over our proposed detection scheme. [9] proved that for large SNR, the SNR penalty is

$$SNR_G = \frac{T_c^2}{T_n^2}, \tag{12}$$

where T_c and T_n are the thresholds for the conventional thresholding detection technique and our proposed detection scheme, respectively. The thresholds are obtained by replacing σ_n^2 in (4) with $|c|^2 + \sigma_d^2$ for the conventional thresholding detection technique, and with σ_d^2 for the proposed detection scheme. Here $|c|^2$ and σ_d^2 are the power of the coherent component and the variance of the diffuse component of the received signals, respectively.

Therefore, the detection gain achieved by our proposed detection scheme is approximately equal to

$$SNR_G = \frac{|c|^2 + \sigma_d^2}{\sigma_d^2} = cdr + 1, \tag{13}$$

with cdr defined by

$$cdr = \frac{|c|^2}{\sigma_d^2}. \tag{14}$$

Notice that cdr is an equivalent gain in signal-to-noise ratio. The larger the cdr, the smaller the target's SNR needs to be to achieve a given detection probability. We use cdr, or $CDR = 10\log(cdr)$, as the detection gain measure for our proposed detection scheme.

Fig.8 illustrates the estimated CDR using 100 returns from the range cell shown in Fig.1. The number of consecutive returns in each group for calculating CDR is 5. Clearly, even though the CDR, or the detection gain, for one range cell changes with time, it is around 30dB for this typical range cell and sometimes achieves a maximum value of 38dB. Therefore, we conclude that the knowledge of the coherent component of bottom reverberation can be used to enhance the detection of targets buried in strong bottom reverberation.

V. Experimental Results

An experiment was conducted under the configuration described in Section II-A, except that a small Remotely Operated Vehicle (ROV) was moved in and out of the sonar beam. Fig.9 illustrates the returns from a typical range cell with the ROV present over the period of the last 10 pings. The returns with the target present are represented by diamonds while the returns without the target present are represented by stars. T_c (dashed circle) and T_n (solid-line circle) are the thresholds established based on the conventional thresholding detection technique and our proposed detection scheme, respectively. Both T_c and T_n are obtained using Fig.5 with given P_F of 10^{-4} and λ of 0.1. For comparison, we display T_n with a displacement determined by the averaged coherent component of these returns without the target present.

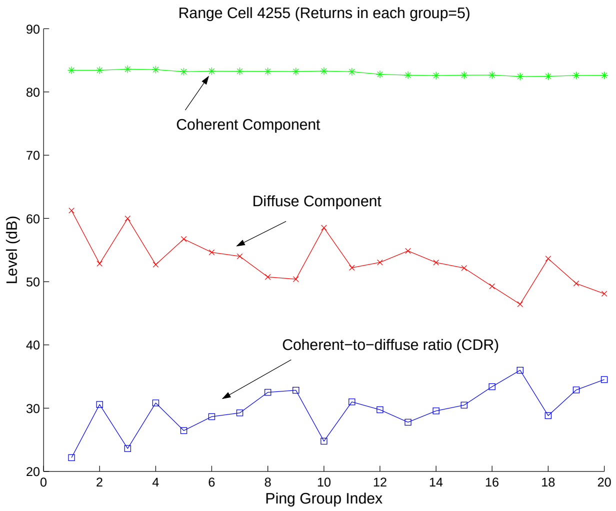


Fig. 8. CDR (detection gain) for a typical range cell

The signal, when the target is present, is a complex sum of the bottom reverberation and the target return. Therefore, these returns with the target present still contain a significant coherent component even though they appear clustered around the origin in the complex plane. Comparing Fig.9 with Fig.2, we conclude that the experimental data match our models for the bottom reverberation and for the target. Note that the P_F determined by T_c is actually less than 10^{-4} because the returns are not circularly symmetric Gaussian as a result of the coherent component, and therefore (4) does not apply. We also observe that the target returns are outside T_n but inside T_c, which means our detection scheme is effective in detecting the small ROV in strong bottom reverberation while the conventional thresholding detection technique is not.

VI. Conclusions and Future Work

This paper shows the trajectories of real bottom reverberation in the complex plane and presents a detection scheme that takes advantage of bottom reverberation coherence to achieve sub-clutter visibility. This paper also demonstrates experimental detection results with a small target present and discusses the detection gain achieved.

However, the results shown in this paper are based on the data obtained by a stationary sonar in a calm environment. Next, we need to study how the ping-to-ping bottom reverberation coherence is degraded by possible transducer motion, environment variations, and multipath interference; and how to achieve a degree of coherence that is close to the one achieved in the calm environment when these disturbances are present.

In conclusion, This paper shows that the proposed detection scheme for sub-clutter visibility is effective in detecting small,

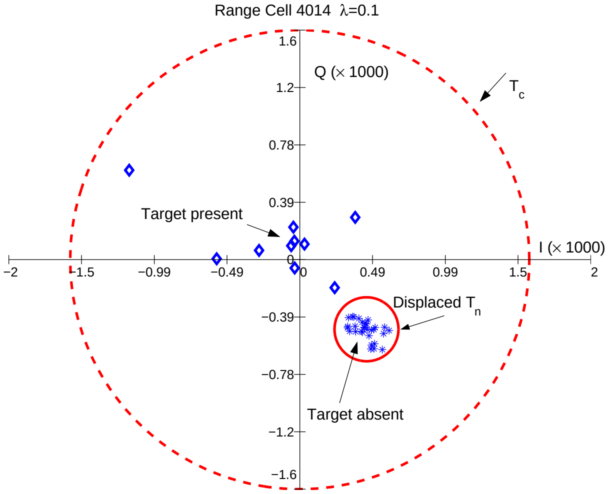


Fig. 9. Returns from a range cell with ROV present

slow-moving sonar targets. This scheme can be integrated into many applications, one of which is the detection of potential threats such as divers. The detection gain is high compared with the conventional thresholding detection technique if the bottom reverberation contains a significant coherent component.

References

[1] J. A. Scheer and J. L. Kurtz, *Coherent Radar Performance Estimation*. Artech House, 1993.

[2] D. C. Schleher, *MTI and pulsed doppler radar*. Artech House Boston, 1991.

[3] S. Kay and J. Salisbury, “Improved active sonar detection using autoregressive prewhiteners,” *IEEE J. Acoust. Soc. Am.*, vol. 87, no. 4, pp. 1603–11, Apr. 1990.

[4] K. M. Kim, C. Lee, and D. H. Youn, “Adaptive processing technique for enhanced CFAR detecting performance in active sonar systems,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 2, pp. 693–700, Apr. 2000.

[5] G. Ginolhac and G. Jourdain, “” principal component inverse” algorithm for detection in the presence of reverberation,” *IEEE J. Ocean. Eng.*, vol. 27, no. 2, pp. 310–321, Apr. 2002.

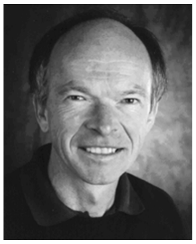
[6] J. S. Bird, “Ground clutter suppression using a coherent clutter map,” in *Radar-82 International Conference*, London, UK, Oct. 1982, pp. 491–495.

[7] ——, “Subclutter visibility for low-doppler targets,” in *Proc. of the 1984 International Symposium on Noise and Clutter Rejection in Radars and Imaging Sensors*, Tokyo, Japan, Oct. 1984, pp. 47–52.

[8] H. Kuschel and W. FGAN-FHR, “Measurement, analysis and processing of VHF ground clutter,” in *The Record of the IEEE 2000 International Radar Conference*, May 2000, pp. 352–358.

[9] J. S. Bird, “The application of circular symmetric signal therory to the detection of singals in noise and clutter,” Ph.D. dissertation, Carleton University, 1980.

Jinyun Ren received the B.S. degree and M.Sc. degree in electrical engineering from Tianjin University, Tianjin, China in 1996 and 1999, respectively. She worked as an electrical engineer for Dongyu Industrial Technology Institute, Shenyang, China, from 1999 to 2001, where she was engaged in hardware design and assembly language development for instrumentation. She has been working towards the Ph.D. degree in School of Engineering Science, at Simon Fraser University, Vancouver, BC, Canada, since 2002. Her current research interests include statistical signal processing and sonar surveillance applications.



John S. Bird received the B.S. degree in electrical engineering in 1973 from the University of British Columbia,Vancouver, BC, Canada, and the Ph.D. degree in electrical engineering from Carleton University, Ottawa, ON, Canada, in 1980. He worked for the Defence Research Establishment Pacific, Victoria, BC, from 1973 to 1981, where he was engaged in signal processing and detection studies related to passive sonar. From 1981 to 1987, he continued these studies but in the context of land- and space-based surveillance radars while on loan to the Communications Research Centre from the Defence Research Establishment Ottawa. During this time, he also worked in the area of satellite spread spectrum communications. He is currently with the School of Engineering Science at Simon Fraser University in Burnaby, BC, where his research interests include signal processing, underwater acoustics, sonar, bottom mapping, autonomous underwater vehicles, limnology, and lake exploration.

Real Time Knot-Tying Simulation

Fuhan Shi
E-mail: fuhans@sfu.ca
School of Engineering Science
Simon Fraser University, Burnaby, BC V5A 1S6

Abstract— **Suturing simulations, of which real time knot-tying is the most challenge part, are essential to today's surgical training systems. In this project, we present a new approach to simulate deformable linear objects (DLOs), with visual and force feedback, by introducing six force components: environmental forces, linear spring force, linear damper, torsional spring, torsional damper, and swivel damper. In our physics-based rope model, which covers the mechanical properties of a real thread such as stretching, compressing, bending, and twisting, we simulate not only external forces, but also internal forces. We developed a simulator to allow users to grasp and smoothly manipulate a virtual rope, and to tie an arbitrary knot. With the assistance of haptic devices, our simulator provide visual and force feedback, and makes our surgical training system more realistic.**

Index Terms— **deformable linear objects (DLOs), knot-tying, collision detection, surgical training system, haptic device, force feedback, bounding-volume hierarchy(BHV).**

I. Introduction

LAPAROSCOPICsurgery (minimally invasive surgery) makes surgery less traumatic to the patient than is to open surgery. However, the requirements for surgical skills are greatly increased. While performing an operation, surgeons cannot rely on traditional eye-hand coordination because they see a 2D image on screens rather than directly seeing the real operating site. In addition, camera views of the operating site can be unusual and unnatural compared to open surgery, which makes the operation even more demanding and difficult to master.

Surgical training is traditionally performed in a master-and-apprentice style, which requires skilled surgeons to spend much time training novice surgeon. The novice surgeon in training watches an expert surgeon performing an operation on real patients, and after sufficient experience, the trainee then perform operations under expert guidance. The models used in traditional training are usually plastic models, animals, and cadavers etc, which either can only demonstrate a limited range of anatomy, cannot reflect the mechanical properties of living tissue, or may not reflect human anatomy. Computer-based surgical simulations, using computers and electromechanical user interface devices, open new possibilities in surgical training, offering many benefits compared to traditional training methods.

Knot-tying simulation, which is a key component of suturing in surgical training systems raises unique and difficult issues for computer-based simulations because of the suture's deformability, difficulty of collision detection and management, and the demanding requirements of force feedback output. In this project, we developed a simulator in our surgical training environment to allow users to tie any kind of knot.

II. Previous Work

A few researchers have made some contributions to the simulation of deformable linear objects (DLOs). Most of these previous models can be put into two categories: physics-based models as in [1], [2] and geometry-based models [3], [4], [5]. Physics-based models are often more accurate, but they may shift continuously until the forces in the system reach an equilibrium, which makes them difficult to manipulate; and the computation is also more complicated generally consuming too much CPU time. Geometric models are slightly less accurate because they do not model internal forces, only displacements, but creating a knot planner is easier for them. In a surgical training system, because our purpose is to enable users to feel force feedback to make it more realistic, we need to calculate both external and internal forces to determine the force output; thus geometric models are obviously inappropriate.

In [2], [6] and [7], researchers have medelled the strands based on the Cosserat theory of elastic rods, but none provide details about how to integrate these models into surgical training systems, how to undertake knot tying, and how to integrate these with haptic devices.

A physics-based model of a rope is represented in [1] by overlapping spheres representing mass-points, which are connected by simple springs. Each mass-point can collide with other mass points as in the instantaneous elastic collision model, but the author only considers the linear spring force and does not allow users to tie a knot. In [8], [9], [10], more force components are taken into consideration, but the authors did not provide any details for knot-tying, and the all the models they used demonstrate instability.

Building our suture model in a similar manner to [10], we provide a user-interface to enable users tie an arbitrary knot. Our collision detection and management methods can help the knot remain stable.

III. Model Description

Based on the finite element method (FEM), our model is a mass-spring system which consist of a spline of linear springs with mass points at the end points of the springs (see Fig. 1).

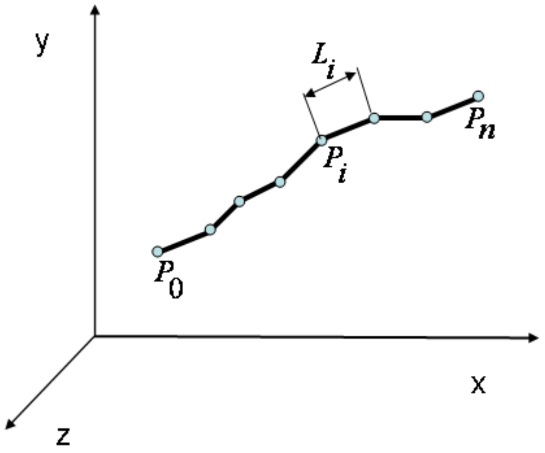


Fig. 1. Suture Model

We define our surgical suture model as a fixed set, S, of P_i $(i = 0, 1, 2, ...n)$ in $3D$ space, connected by cylinder segments. From the simulation function's point of view, these cylinders do not exist and the whole computation is based on the position of these mass point, P_i, and their distance to the next and previous point.

IV. Force Computation

To calculate the shape of our rope, we compute the total force acting at each point, P_i, and update its position based on the computed force. Once the total force at each of the nodes has been calculated, the following equations are use to find the next position of each node. Let us define:

dt = seconds between time i and $i + 1$
F_i = total force on the mass point at time i
M = Mass of the mass point
V_i = mass point's velocity at time i
P_i = the position of mass point at time i

Then,

$$V_{i+1} = V_i + \frac{F_i}{M} dt, \tag{1}$$

$$P_{i+1} = P_i + V_{i+1} dt, \tag{2}$$

The springs and dampers each contribute some force to F_i. The various different springs and dampers all behave differently and we calculate their force contributions using their own particular equations. We computed six force components (internal forces and external forces) to identify the positions of our suture model. Each spring and damper is described as follows:

A. Environmental forces

Environmental forces include contact forces with the grippers or obstacles as well as the gravitational force:

$$F_g = G * m, \tag{3}$$

where $G = 9.8N/kg$, and m is the mass of the mass point.

B. Linear spring force

The linear spring force is computed by comparing the current distance, L_i, between point, P_i and P_{i+1}, with the rest length between point, P_i and P_{i+1}, and by projecting the resulting difference on the direction from point P_i to P_{i+1}.

$$L_i = ||P_{i+1} - P_i||, \tag{4}$$

$$\Delta L = \frac{L_i - L_r}{L_r}, \tag{5}$$

where, L_r, is the rest length between point ,P_i and P_{i+1},(see Fig. 2).

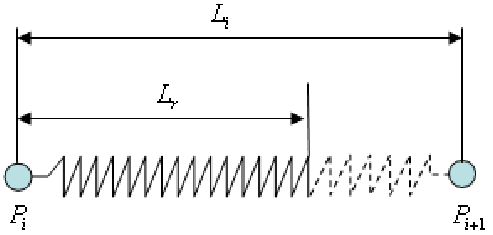


Fig. 2. Linear Spring

Let E_i be the unit vector whose direction is from point, P_i to P_{i+1}, then,

$$E_i = \frac{P_{i+1} - P_i}{||P_{i+1} - P_i||}, \tag{6}$$

Then,

$$F = K * \Delta L * E_i, \tag{7}$$

where K is the linear spring constant.

C. Linear damper (linear friction)

Other forces try to stop the spring as it moves: the friction of air resists the movement, and the internal friction between the mass points of the spring and an external damper also resist the movement. All of these can be combined into a constant called the damping factor, B. This force also opposes the direction of movement and is proportional to the velocity of the moving mass. Notice that when the system is at rest (V=0), no damping force is involved.

$$F = B * (v_{i+1} - v_i) * E_i, \tag{8}$$

where B is the linear damper constant. E_i is the unit vector where the direction is from point, P_i to P_{i+1}. E_i can be calculated as follows:

$$E_i = \frac{P_{i+1} - P_i}{||P_{i+1} - P_i||}, \tag{9}$$

v_{i+1} and v_i in 2.8 are the components of the velocity of point P_{i+1} and P_i on the direction E_i. They can be calculate as follows:

$$v_{i+1} = V_{i+1} \cdot E_i, \tag{10}$$

$$v_i = V_i \cdot E_i. \tag{11}$$

D. Torsional spring

The torsional spring is derived from the angle, α, between two connected segments of the rope. The basic idea is to model each two connected segments as a triangle with a spring as the hypothesis pushing the end points to the full expanded position. The length of the two connected segments remain unchanged. Only the force component orthogonal to the segments is used for the end points (See Fig. 3).

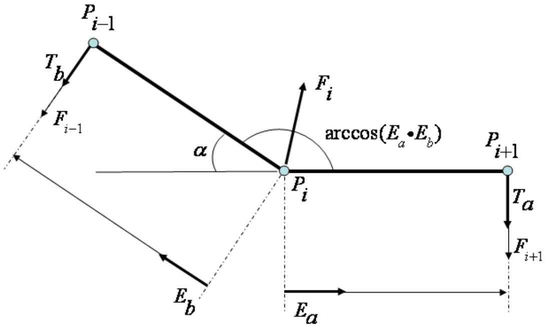


Fig. 3. Torsional Spring

The force can be computed as follows:
Let E_b and E_a be the unit vectors with directions from point, P_i to P_{i-1}, and from, P_i to P_{i+1}, respectively. Then,

$$E_a = \frac{P_{i+1} - P_i}{||P_{i+1} - P_i||}, \tag{12}$$

$$E_b = \frac{P_{i-1} - P_i}{||P_{i-1} - P_i||}, \tag{13}$$

Let T_b and T_a be the unit vectors with directions the same as the force applied at the two endpoints and therefore, orthogonal to E_b and E_a respectively. Then,

$$T_a = E_a \times (E_a \times E_b), \tag{14}$$

$$T_b = E_b \times (E_b \times E_a), \tag{15}$$

We can derive α from:

$$\alpha = \pi - arccos(E_a \cdot E_b), \tag{16}$$

Then,

$$F_{i-1} = K_t * \frac{\alpha}{\pi * ||P_{i-1} - P_i||} * T_b, \tag{17}$$

$$F_{i+1} = K_t * \frac{\alpha}{\pi * ||P_{i+1} - P_i||} * T_a, \tag{18}$$

$$F_i = -(F_{i-1} + F_{i+1}). \tag{19}$$

E. Torsional damper

The torsional damper works against the torsional spring to prevent any harmonic motion from accumulating. Similar to the linear damper, it also models the internal friction that resists bending in regular objects.

Let, v_{i-1}, v_{ib}, be the velocity components of, V_{i-1}, and, V_i, on the direction of, T_b, and let, v_{i+1}, v_{ia}, be the velocity components of, V_{i+1}, and, V_i, on the direction of, T_a,

$$v_{i-1} = V_{i-1} \cdot T_b, \tag{20}$$

$$v_{ib} = V_i \cdot T_b, \tag{21}$$

$$v_{i+1} = V_{i+1} \cdot T_a, \tag{22}$$

$$v_{ia} = V_i \cdot T_a, \tag{23}$$

Then, the torsional damper on the point, P_{i-1}, can be computed by:

$$F_{i-1} = B_t * (\frac{(v_{i-1} - v_i)}{||P_{i-1} - P_i||} + \frac{(v_{i+1} - v_i)}{||P_{i+1} - P_i||}) * \frac{T_b}{||P_{i-1} - P_i||}, \tag{24}$$

$$F_{i+1} = B_t * (\frac{(v_{i-1} - v_i)}{||P_{i-1} - P_i||} + \frac{(v_{i+1} - v_i)}{||P_{i+1} - P_i||}) * \frac{T_a}{||P_{i+1} - P_i||}, \tag{25}$$

$$F_i = -(F_{i-1} + F_{i+1}). \tag{26}$$

F. Swivel damper

Node, P_{i-1}, has a velocity relative to the center node, P_i. So far, two components of that relative velocity have been dampened; however, there still remains a component perpendicular to those two. It isn't shown on the picture below because this component comes straight in/out of the page. Without the dampening node, P_{i-1} could indefinitely orbit the line formed by extending the edge connecting node, P_{i+1} and node, P_i (See Fig. 4).

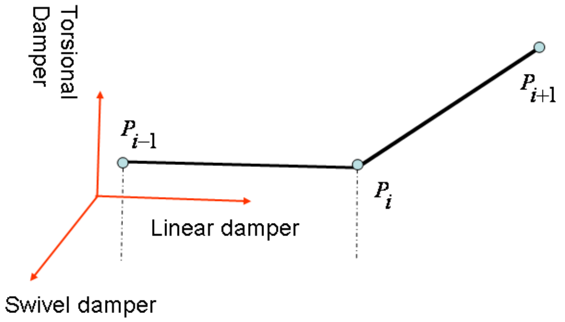


Fig. 4. Swivel Damper

Let S_a and S_b be the unit vectors whose directions are the swivel damper of point, P_{i-1} and P_{i+1},

$$S_a = E_a \times T_a, \tag{27}$$

$$S_b = E_b \times T_b, \tag{28}$$

$$F_{i-1} = B_s * \frac{(V_{i-1} - V_i) \cdot S_b}{||P_{i-1} - P_i||} * S_b, \tag{29}$$

$$F_{i+1} = B_s * \frac{(V_{i+1} - V_i) \cdot S_a}{||P_{i+1} - P_i||} * S_a, \tag{30}$$

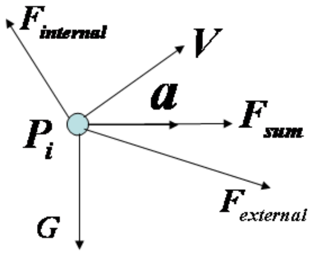


Fig. 5. Haptic Force Output

If the mass of the mass point is m, then

$$F_{sum} = m * a, \tag{32}$$

where a is the acceleration. As we know:

$$F_{sum} = F_{internal} + G + F_{external}, \tag{33}$$

where, G, is gradational force, then we can derive the user input force:

$$F_{external} = F_{sum} - (F_{internal} + G). \tag{34}$$

Reverse the direction of $F_{external}$, output it through Phantom Omni, we can enable users feel right force feedback.

VI. Collision Detection and Management

A. Collision Detection

First, we build a bounding-volume hierarchy (BVH) from the bottom up representing the shape of the rope at successive levels of detail (see Fig. 6).

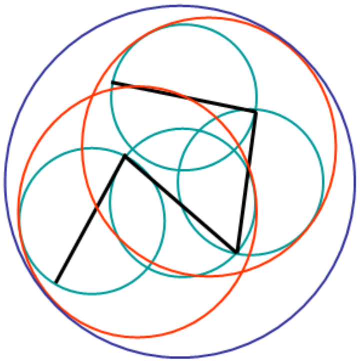


Fig. 6. Bounding-Volume Hierarchy

We bound each rope node by its minimal enclosing sphere. The spheres thus obtained are installed as the leaves of the BVH, in the same order as the nodes along the rope. Then, we bound pairs of successive spheres by new spheres to form each next level of the hierarchy. Hence, the resulting BVH is a balanced binary tree. Each intermediate sphere bounds tightly its two children, So it also encloses all the leaf spheres below it. The root sphere encloses the entire rope and all the other spheres.

To find the self-collisions of the rope, we explore two copies of the BVH from the top down. Whenever two BVHs (one from each copy) are found to not overlap, we know that they cannot contain colliding segments, and hence, we do not explore their contents. When two leaf spheres overlap, the distance between the two centers of the nodes is computed. If it is less than the node diameter, 2R, then the two segments are reported to collide. However, no node is ever considered to be in collision with itself or its immediate neighbors along the rope chain.

To find the collisions between the rope and grippers, we consider the gripper as line segments with a given radius, and check if the BHV of the rope has any overlap with the segments.

The topology of the BVH is computed once before any simulation and then remains fixed. During simulation, only the positions of the bounding spheres and the positions and radii of the higher level spheres are re-computed at each cycle, and this computation is done from the bottom up, so that no sphere is updated more than once.

B. Collision Detection Management

Because we set a threshold for the spring, the maximum length of each segment can not be more than $L_r + K_s$, where $k_s = 1/2 * L_r$; therefore, the rope will never pass through itself.

When two control points collide, they affect each other's trajectories, disrupting the momentum of the system and changing the boundary conditions. The points must not pass through each other. In order to simulate the momentum loss during collision, a constant factor, c, was introduced, such that:

$$||V_{after}|| = c * ||V_{before}||. \tag{35}$$

Our algorithm also ensures that the control points' new velocities repel them from each other (See Fig. 7)

When two nodes are detected to be at a distance, $d < 2R$, from each other, then an equal (but opposite) displacement vector is applied to each node. This displacement is just long enough to romove the segments from collision, with a slight "safety margin" to allow for a sliding motion discussed below. Hence, each node is shifted away by, $R - d/2 + \varepsilon/2$. Since the colliding nodes have been pushed slightly out of contact, they will not be in collision during the following cycle, allowing the rope to slide along itself. After this new motion step, a

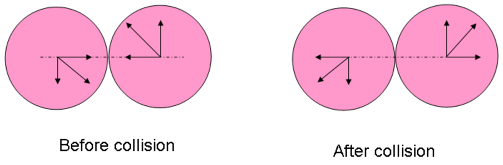


Fig. 7. Collision Detection

new collision between the two nodes may, or may not, happen again (See Fig 8).

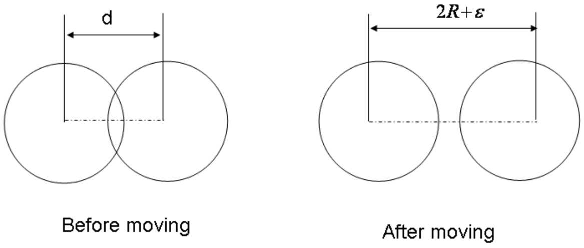


Fig. 8. Overlapping Management

VII. Experiments and Results

We build five different models with different component forces which we have discussed in section II. Comparing the results from these different models, we can obtain the most realistic model.

A. Model 1

This model contains only a linear spring and a linear damper. It is the fastest model to calculate, but the least realistic. The edges between the nodes can bend effortlessly to any angle, whereas on a real thread the internal forces would work against the thread making sharp bends (see Fig. 9).

B. Model 2

This model is almost the same as the model one, but also contains a torsional spring which adds a lot more realism, but is also a lot more computationally intensive. It uses an ’acos’ function and creates some harmonic wave motion in the thread, making it look more like an elastic band (see Fig. 10).

C. Model 3

Compared to model 2, model 3 has been added a torsional damper. This damper stops the harmonic motion present in model 2, but there still remained one obvious problem: if you move the top of the rope in a small circle, the rope will start to ’orbit’ around, and will never stop (see Fig. 11).

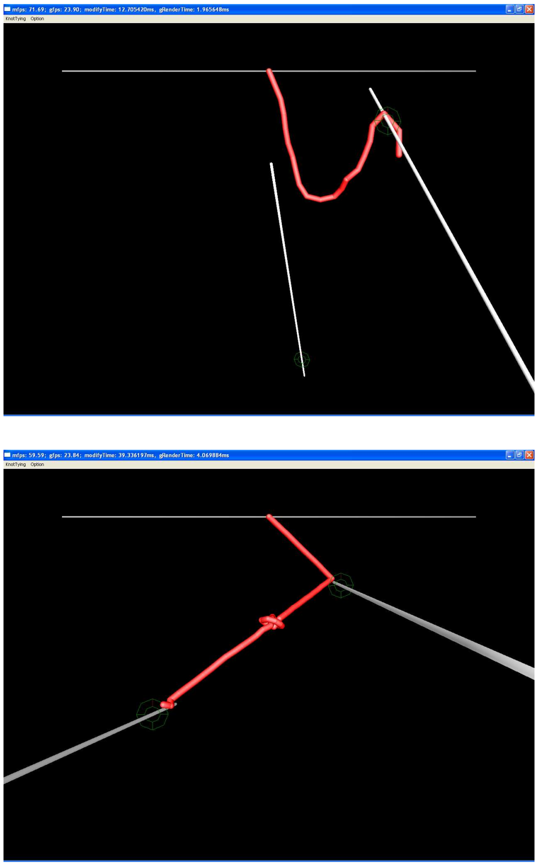

mfps: 71.69; gfps: 23.90; modifyTime: 12.705420ms, gRenderTime: 1.965648ms
KnotTying
Option
mfps: 59.59; gfps: 23.84; modifyTime: 39.336197ms, gRenderTime: 4.069884ms
KnotTying
Option

Fig. 9. Model 1

D. Model 4

Model 4 includes a swivel damper to fix the problem of perpetual orbiting. The result is a thread that looks more like a real thread and less like an elastic band, or a chain of masses tied together with springs (see Fig. 12).

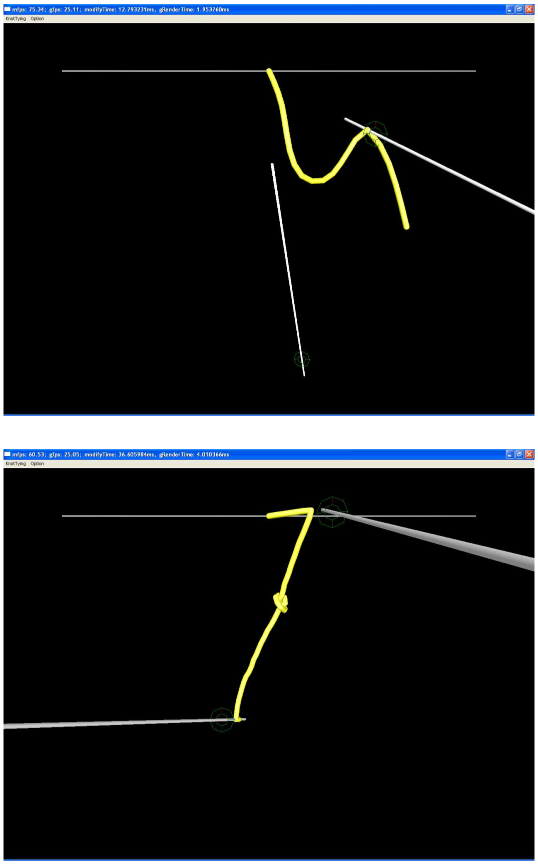
E. Model 5

This model has all the components of model 4. The only difference is that the linear spring’s force is based quadratically on the difference between its current length and rest length, instead on linearly, as it normally is. This made the thread appear a lot less stretchy, which is more realistic since real threads stretch very little (see Fig. 13).

VIII. Conclusion and future work

We presented a fast and simple approach to compute 3D DLO simulations. The key component of our model involves the computations of all six force components. While our simulation cannot produce physically exact shapes and forces, the simulated forces are realistic enough to drive a haptic feedback device like ”Phantom-Omni” from Sensable Inc.

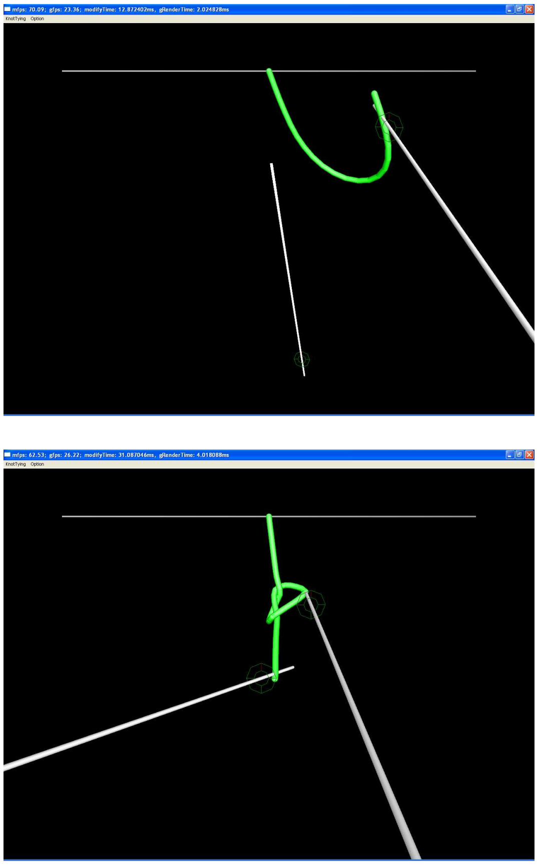
One remaining problem is that, to make sutures more realistic, we must add more segments and more mass points to the model, which may cause the program run more slowly (the more mass points the model has, the more time we need to complete dynamics computation and collision detection). Therefore, we cannot guarantee the haptic



```
mfps: 75.34; gfps: 25.11; modifyTime: 12.793231ms, gRenderTime: 1.953760ms
```

```
mfps: 60.53; gfps: 25.05; modifyTime: 36.605984ms, gRenderTime: 4.010366ms
```

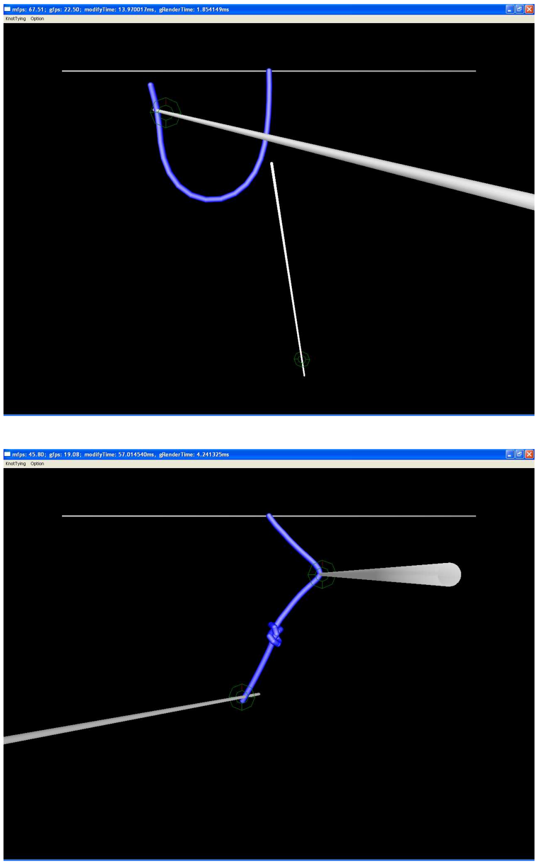
Fig. 10. Model 2



```
mfps: 70.09; gfps: 23.36; modifyTime: 12.872402ms, gRenderTime: 2.024828ms
```

```
mfps: 62.53; gfps: 26.22; modifyTime: 31.087046ms, gRenderTime: 4.018088ms
```

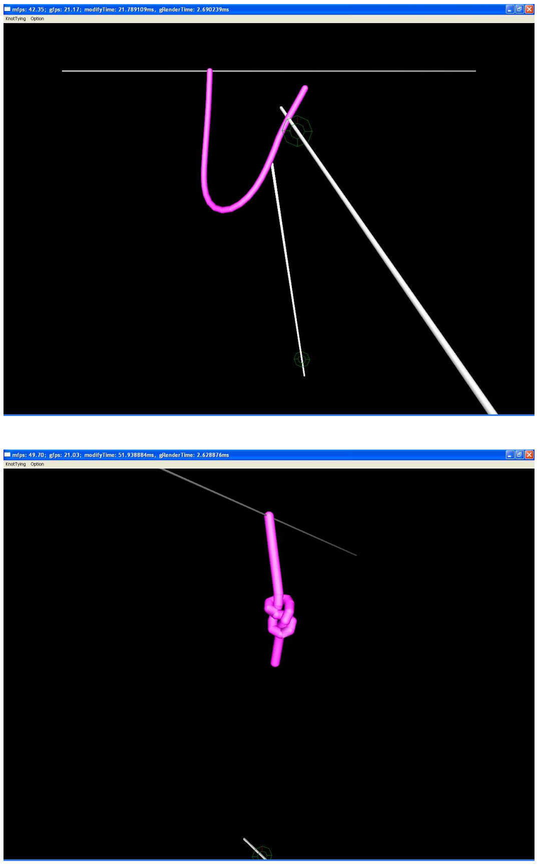
Fig. 11. Model 3



```
mfps: 67.51; gfps: 22.50; modifyTime: 13.970017ms, gRenderTime: 1.854149ms
```

```
mfps: 45.80; gfps: 19.08; modifyTime: 57.014540ms, gRenderTime: 4.241325ms
```

Fig. 12. Model 4



```
mfps: 42.35; gfps: 21.17; modifyTime: 21.789109ms, gRenderTime: 2.690239ms
```

```
mfps: 49.70; gfps: 21.03; modifyTime: 51.938884ms, gRenderTime: 2.628876ms
```

Fig. 13. Model 5

rendering rate to be at least 1000Hz. Users may feel the force output less smoothly, as if the haptic devices might be kicking.

To solve the problem mentioned above, we may introduce level set methods to the modeling and undertake some optimization of dynamic computation and collision detection methods.

REFERENCES

[1] A. Ladd, *Simulated knot tying*, Proceedings of the 2002 IEEE International Conference on Robotics and Automation (ICRA), Washington D.C.

[2] D. Pai, *Strands: Interactive simulation of thin solids using cosserat models*, Computer Graphics Forum, 21(3):347–352, 2002. Proceedings of Eurographics'02.

[3] J. Brown, J. Latombe, K. Montgomery, *Real-time knot tying simulation*, The Visual Computer: International Journal of Computer Graphics, 20(2): 165-179.

[4] J.Takamatsu, T.Morita, K.Ogawara, H.Kimura, K.Ikeuchi,*Representation for Knot-Tying Tasks*,Robotics, IEEE Transactions,Volume: 22, Issue: 1,On page(s): 65- 78,Feb. 2006

[5] Hidefumi Wakamatsu, Eiji Arai,*Knotting/Unknotting Manipulation of Deformable Linear Objects*, International Journal of Robotics Research, Volume 25 , Issue 4 (April 2006), Pages: 371 - 395, ISSN:0278-3649

[6] WANG, C., RICHARDSON, A.M.D., LIU, D., ROSING, R., TUCKER, R. and DE MASI, *Construction of Nonlinear Dynamic MEMS Component Models Using Cosserat Theory*, Proc. SPIE Design, Test, Integration & Packaging of MEMS Symposium, 2003

[7] D. Q. Cao, Dongsheng Liu, Charles H.-T. Wang,*Three Dimensional Nonlinear Dynamics of Slender Structures: Cosserat Rod Element Approach*, eprint arXiv:math/0410286,10/2004

[8] Bjrn Kahl, and Dominik Henrich,*Manipulation of deformable linear objects: Force-based simulation approach for haptic feedback*, 12th International Conference on Advanced Robotics (ICAR 2005), July 18th-20th, 2005

[9] F. Wang, E. Burdet, A. Dhanik, T. Poston, , C. L. Teo , *Dynamic Thread for Real-Time Knot-Tying*, Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems - Volume 00 2005

[10] M. LeDuc, S. Payandeh, J. Dill, *Toward modeling of a suturing task*, Graphics Interface (GI), pp 273-279, Halifax, Nova Scotia, 2003.

[11] J. Lenoir, P. Meseure, L. Grisoni, C. Chaillou, *A suture model for surgical simulation*, 2nd International Symposium on Medical Simulation (ISMS'04) pp. 105-113, Cambridge, Massachusetts (USA).

[12] J. Lenoir, P. Meseure, L. Grisoni, C. Chaillou, *Surgical thread simulation, Modelling & Simulation for Computer-aided Medicine and Surgery*, (MS4CMS), November 2002

2D Filter Banks and Directional Vanishing Moments Filter Banks

Guoqian Sun, *Student Member, IEEE,*

.

ABSTRACT

The fundamentals of multidimensional filter(MD) bank theory is introduced and the perfect reconstruction condition of MD filter banks is discussed. Some methods which simplify the design of 2D filter banks, such as by using the cayley transform [1], are discussed in this parper. Two channel 2D filter banks with directional vanishing moments are discussed in more detail. The simulation result for the method is illustrated. The result shows the compression efficiency can be improve by applying the proper designed filter bank.

***Index Terms*— multiple dimensional, filter banks, directional vanishing moments**

I. Introduction

THE filter bank theory is widely applied in the image coding and compression. According to the dimension of the filters which are used to construct the filter bank, the filter bank can be divided into two different groups: one-dimensional (1D) filter banks, and multi-dimensional filter banks. The 1D filter banks are the most widely used filter banks, because the 1D filters are simple and flexible, and they can be easily extended to 2D images by applying them to rows and columns separately. Therefore, the complexity of encoding can be dramatically reduced by using 1D filters. However, with improvements to the hardware, image compression systems can afford more complexity. The efficiency of the coding system is receiving increasing attention. Differing from the 1D filter, which can only capture the horizontal or vertical geometric characters in the image, 2D filter banks can capture the geometric characters of the image in any direction. Therefore, by using the 2D filter banks, the coding efficiency of the system is higher than that of 1D filter banks system.

Nonseparable filter banks can capture geometric structures in MD data and offer more freedom and better frequency selectivity than traditional separable filter banks constructed from 1D filter banks. The MD filter banks gain more attention. Because traditional design methods for one-dimensional filter banks can not be extended directly to higher dimension and a multidimensional factorization theorem is lacking, designing a nonseparable MD filter banks is difficult. some methods which simplify the design of 2D filter banks, such as by using the cayley transform [2] are in this paper. Among those methods, two channel 2D filter banks with directional vanishing moments[3][4] are discussed in more detail. Such filter banks can capture the edge along the desired direction in the image. One design example is shown, and the performance of this filter is also shown.

In Section II, we introduce the fundamentals of Multidimensional Multirate systems. In Section III, we discuss the methods used to design multidimensional filter banks. In Section IV, directional vanishing moments filter banks are discussed in detail. in Section V, A design example and simulation result are shown. The paper is summarized in Section VI.

II. Fundamentals of Multidimensional Multirate Systems[5]

A. Z Transform in MD

In the D-dimensional case, the signal is denoted as $x(\mathbf{n})$, where the time index, $\mathbf{n}$, is a column vector

$$\mathbf{n} = [n_1, n_2, \dots, n_{D-1}]^T. \tag{1}$$

Thus, $x(\mathbf{n})$ is a function of D integer variables $n_k, 0 \leq k \leq D-1$. The Z transform is defined analogous to the 1D case. The Z transform is defined by

$$X(\mathbf{z}) = \sum_{\mathbf{n}\in\mathcal{N}} x(n) z_0^{-n_0} z_1^{-n_1} \dots z_{D-1}^{-n_{D-1}}. \tag{2}$$

Here, $\mathbf{z}$ denotes the vector

$$\mathbf{z} = [z_1, z_2, \dots, z_{D-1}]^T. \tag{3}$$

Let

$$z_k = e^{j\omega_k}, 0 \leq k \leq .D-1. \tag{4}$$

The z-transform reduces to the Fourier transform $X(\omega)$. Introducing the notation

$$Z(\mathbf{n}) = z_0^{-n_0} z_1^{-n_1} \dots z_{D-1}^{-n_{D-1}}. \tag{5}$$

The z-transform becomes

$$X(\mathbf{z}) = \sum_{\mathbf{n}\in\mathcal{N}} x(\mathbf{n}) Z(-\mathbf{n}). \tag{6}$$

Now, we can write the multidimensional z-transform in a manner which resembles the 1D z-transform.

The key properties of Fourier and z-transform in 1D, still satisfy the multidimensional case.

Linearity

$$a_1 x_1(\mathbf{n}) + a_2 x_2(\mathbf{n}) \Longleftrightarrow a_1 X_1(\mathbf{z}) + a_2 X_2(\mathbf{z}). \tag{7}$$

Shifting

$$y(\mathbf{n}) = x(\mathbf{n} - \mathbf{k}) \Longleftrightarrow Z(-\mathbf{k}) X(\mathbf{z}). \tag{8}$$

Convolution theorem

$$y(\mathbf{n}) = \sum_{\mathbf{m}\in\mathcal{N}} h(\mathbf{m})x(\mathbf{n}-\mathbf{m}) \Longleftrightarrow Y(\mathbf{z}) = H(\mathbf{z})X(\mathbf{z}). \quad (9)$$

From equation (9), we know the multidimensional filter has the form of

$$H(\mathbf{z}) = \sum_{\mathbf{n}\in\mathcal{N}} h(\mathbf{n})Z(-\mathbf{n}). \quad (10)$$

We can divide the multidimensional filters into two groups: separable filters and non-separable filters. A multidimensional filter is said to be separable if

$$H(\omega) = H_0(e^{j\omega_0})H_1(e^{j\omega_1})\ldots H_{D-1}(e^{j\omega_{D-1}}). \quad (11)$$

In this case, the transform function is the product of D one-dimensional transform function. We can apply the multidimensional filter by applying D times the i-th 1D filter along i-th dimension. For instance, in the two dimension case, we apply the row transform and then apply column transform to the image. The separable filters can be easily designed from 1D filters. If the filter can't be written as the product of 1D transform filters, this filter is said to be a non-separable. The design of the non-separable filters has more freedom than the design of separable filter. The support region of the frequency response of the non-separable filters can be an arbitrary shape, which is cubic in the separable filters. The non-separable filters have more desirable properties. Most of our filter bank design discussed later in this paper are all non-separable filters.

Multidimensional filters also have the linear phase property. A multidimensional filter is said to be linear phase if

$$H(\omega) = ce^{-j\mathbf{k}^T\omega}H_R(\omega), \quad (12)$$

where $\mathbf{k}$ is a real constant vector, $H_R(\omega)$ is a real function, and c is a constant.

B. Downsampling and Upsampling in Multidimensional Signal

1) Downsampling: The sampling parameter for multidimensional signals is a matrix called sampling matrix, similar to the 1D sample rate. For sampling of one dimensional signals, the coordinates of the sampled points are the integer times of the sampling rate. But for the sampling of the multidimensional signal, the sampled points are located on the lattices of the sampling matrix. The coordinates of the points on the lattice are the integer linear combinations of the column vector of the sampling matrix. We can write the sampling signal as

$$y(\mathbf{n}) = x(\mathbf{Mn}). \quad (13)$$

The set of the vectors, $\mathbf{Mn}$, is the lattice generated by $\mathbf{M}$, denoted as $\mathbf{LAT}(\mathbf{M})$. Fig.1 shows the lattice generated by the sampling matrix

$$V = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}.$$

But in multidimensional signal sample, the basic lattice matrix is not unique. The $\mathbf{V}$ is a sampling matrix which generates the lattice $\mathbf{LAT}(\mathbf{V})$, and $\mathbf{U} = \mathbf{VE}$. If $\mathbf{E}$ is a unimodular integer matrix, U can also generate the lattice LAT(V). The integer matrix E is said to be unimodular if $[det\mathbf{E}] = \pm 1$. Therefor the inverse $\mathbf{E}^{-1}$ is also a integer matrix. Two sampled 2D images are shown in Fig.2. The two sampling matrix is equivalent to each other by the unimodular matrix, $\mathbf{E}$. As we can see from Fig.2, the two sampled image are the same, but only the coordinates of the points in the images are different. Therefore, any image sampled by equivalent sampling matrices generates the same image, but those sampled images are reordered.

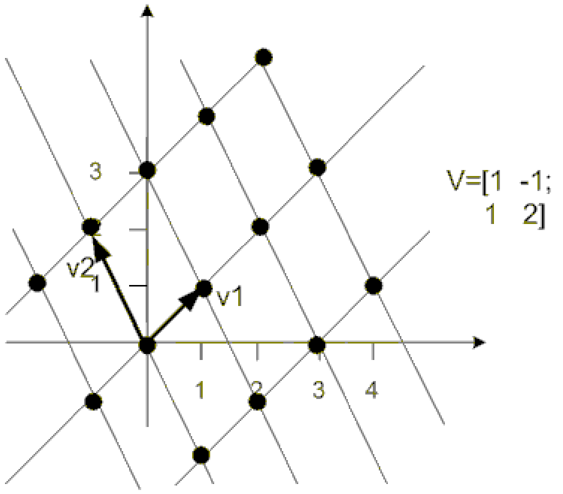


Fig. 1. The set of sample points ($\mathbf{LAT}(\mathbf{V})$) generated by sampling matrix $\mathbf{V}$

For a sampling matrix $\mathbf{V}$, we sketch all the basis vector v_k (columns of $\mathbf{V}$) in a D-dimensional coordinate space, as demonstrated in Fig.3. We can complete a parallelepiped with these vecotors as edges as shown. It is called fundamental parallelepiped, denoted as $\mathbf{FDP}(\mathbf{V})$ of sampling matrix $\mathbf{V}$. We can also define the fundamental parallelepiped as

$$\mathbf{FPD}(\mathbf{V}) = \text{ set of all points } \mathbf{Vx} \text{ with } x \in [0,1)^D. \quad (14)$$

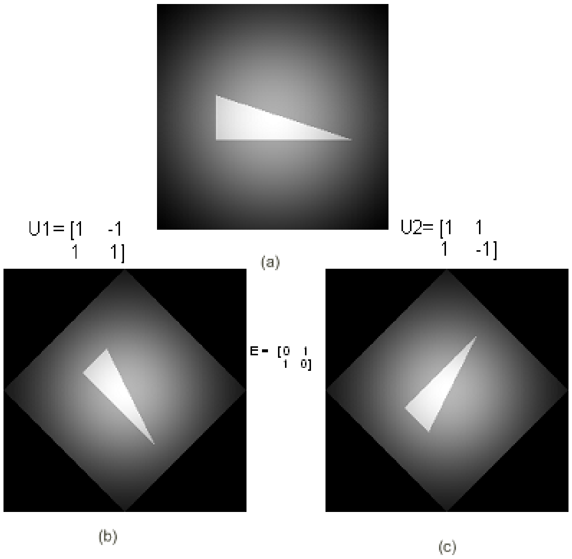


Fig. 2. Sampling in 2D and unimodular matrix

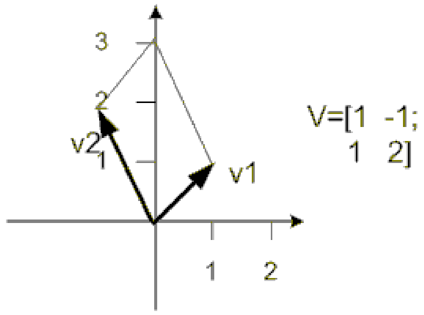


Fig. 3. The fundamental parallelepiped $\mathbf{FPD}(\mathbf{V})$

The $\mathbf{FPD}(\mathbf{V})$ contains a finite number of integer vectors. We denote the set of integer vectors as

$$\mathcal{N}(\mathbf{V}) = \text{ set of integer vectors in } \mathbf{FPD}(\mathbf{V}). \tag{15}$$

The number of vectors in the set $J(\mathbf{V})$ is given by

$$J(\mathbf{V}) = |det\mathbf{V}|. \tag{16}$$

The sampling retains only one sample point in the $\mathbf{FPD}(\mathbf{V})$. After sampling, the number of points in sampled signal reduces to $1/\mathbf{J}(\mathbf{V})$ number of points in the original signal. The $\mathbf{J}(\mathbf{V})$ is said to be the density of the sampling matrix $\mathbf{V}$ or the sampling ratio.

For one dimensional sampling, the frequency response is extended by the sampling ratio M times, and then shifted by $2k\pi/M, k = 0, ..M-1$. The frequency response of the sampled signal is the the summation of all the shifted frequency response. The frequency response of the 2D sampled signal is analogous to the 1D case. The frequency response can be written as

$$Y(\omega) = \frac{1}{J(\mathbf{V})} \sum_{\mathbf{k}\in\mathcal{N}(\mathbf{V}^T)} X(\mathbf{V}^{-T}(\omega - e\pi\mathbf{k})). \tag{17}$$

In equation(17), the number of terms in the summation is equal to $J(\mathbf{V})$. The term $X(\mathbf{V}^{-T}\omega)$ is the stretched version. The other terms in which $\mathbf{k} \neq 0$ are the shifted versions which can generate the alias effect. But the stretching in the multidimensional signal is not as even as the 1D case, because the stretched extent in a direction is depend on sampling matrix. Further, the shifts are vector shifts and not along the dimension axis.

2) Upsampling: Upsampling in multidimensional signal is simply the inverse process of down sampling. In stead of taking out the points on the lattice of sampling matrix in down sampling case, we put the signal onto the lattice of the sampling matrix, leave the rest of points to be zero. This process can be written as

$$y(\mathbf{n}) = \begin{cases} x(\mathbf{V}^{-1}\mathbf{n}), & \text{if } \mathbf{n} \in \mathbf{LAT}(\mathbf{V}) \\ 0, & \text{otherwise} \end{cases} \tag{18}$$

Fig.4 shows the upsampling example of a 2D image.

The frequency response of the upsampled signal is

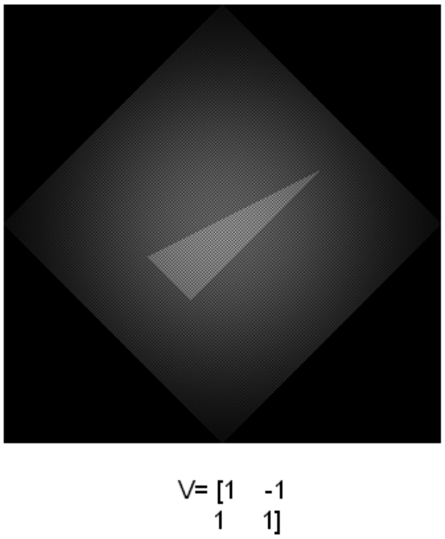


Fig. 4. Upsampling example in 2D image

$$Y(\omega) = X(\mathbf{V}^T\omega), \tag{19}$$

or in the z-transform

$$Y(\mathbf{z}) = X(\mathbf{z}^{\mathbf{V}}), \tag{20}$$

where the $\mathbf{z}^{\mathbf{V}}$ is

$$\mathbf{z}^{\mathbf{V}} = [\mathbf{z}_{\mathbf{0}}^{\mathbf{v}}, \mathbf{z}_{\mathbf{1}}^{\mathbf{v}}, \ldots \mathbf{z}_{\mathbf{D-1}}^{\mathbf{v}}]^T. \tag{21}$$

The column vectors of $\mathbf{V}$ are denoted by v_i, and the spectrum is compacted by $J(V)$ times.

3) polyphase decomposition: Like the 1D signal, the multidimensional signal can also be written in polyphase decomposition form. Given the sampling matrix, $\mathbf{M}$, we can write a multidimensional signal as

$$X(\mathbf{z}) = \sum_{\mathbf{k}\in\mathcal{N}(\mathbf{M})} Z(-\mathbf{k})X_{\mathbf{k}}(\mathbf{z}^{\mathbf{M}}). \tag{22}$$

Here, $Z(-\mathbf{k})$ is analogous to the 1D delay. $X_{\mathbf{k}}(\mathbf{z})$ is the z-transform of the $\mathbf{k}$-th polyphase component $x_{\mathbf{k}}(\mathbf{n})$ of $x(\mathbf{n})$. The polyphase component $x_{\mathbf{k}}(\mathbf{n})$ is

$$x_{\mathbf{k}}(\mathbf{n}) = x(\mathbf{Mn} + \mathbf{k}), \mathbf{k} \in \mathcal{N}(\mathbf{M}). \tag{23}$$

4) Alias-free for sampling: Due to the shift versions in the spectrum of the sampled signal, different shift versions may overlap, and the overlap generates alias. In order to avoid the alias effect, we should constrain the spectrum of the original multidimensional signal. The symmetric parallelepiped, $\mathbf{SPD}(\mathbf{V})$, is defined as the set of vector of the form

$$\mathbf{Vx}, \mathbf{x} \in [-1, 1)^D. \tag{24}$$

Fig.5 shows the example of the symmetric parallelepiped. We can determine the SPD in this figure is combined by the FPD and 3 shift versions. We can verify that $SPD(\mathbf{V})$ can be obtained by appropriately shifting and scaling $FPD(\mathbf{V})$. More specifically, we have

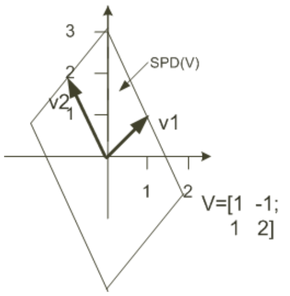


Fig. 5. The symmetric parallelepiped

$$SPD(\mathbf{V}) = FPD(2\mathbf{V}) - \mathbf{V}\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}. \tag{25}$$

If we want the sampled spectrum with no alias, the spectrum of the original signal should have the support region restricted in the $SPD(\pi\mathbf{V}^T)$; therefore, the stretched shifted version doesn't have overlap with each other.

C. Nobel Identities

The nobel identities in the 1D case can be extended to the multidimensional case. Fig.6 shows those identities.

D. Multirate Filter Design

The design of analysis and synthesis filters for a general sampling matrix, M, is a non-trivial problem. The most common frequency response of the filter is of passband region $SPD(\pi\mathbf{V}^T)$. This type of filter can avoid the alias effect after downsampling.

1) Design of Diamond- and Fan-shaped Filters: The diamond shaped filter is useful in some cases, because this filter is more accurate than the rectangle shaped filter which is a separable filter. In [5], a method of designing the diamond shaped filter from one dimensional filter is introduced. The design steps are as follows

1) Design the 1D low pass filter $G(z)$,
2) Define the 2D transfer function $G(z_0 z_1)$(replace z with $z_0 z_1$),
3) Let $K(z_0, z_1) = G(z_0 z_1)G(z_0 z_1^{-1})$,
4) Let $H(z_0, z_1) = \frac{1}{2}[K(z_0, z_1) + K(-z_0, z_1)]|_{z_i \to z_i^{1/2}}$
5) $H(z_0, z_1)$ is the diamond shaped filter,
6) $H(-z_0, z_1)$and $H(z_0, -z_1)$ are fan-shaped filters, which can be obtain by shifting π the Diamond-shaped filter at different direction.

Fig.7 shows the frequency response of the designed filter when n=10.

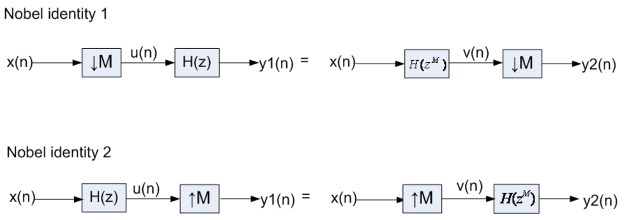


Fig. 6. The multidimensional multirate nobel identities

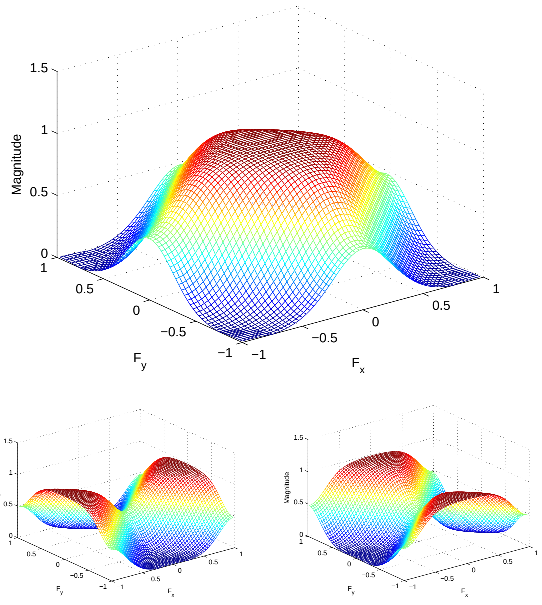


Fig. 7. The nonseparable 2D diamond shaped filter and fan shaped filter

2) Multidimensional filters from the 1D filters: For a more general case, we want a filter which satisfies the alias free condition mentioned in the last section. We can design such a filter from the 1D filter. The procedure for designing the decimation filter $H(\omega)$ for sampling matrix M is as follows

1) Design a 1D low pass filter $P(\omega)$, p(n) denotes its impulse response
2) Define the 2D separable filter as
$$h^{(s)}(\mathbf{n}) = p_0(n_0)p(n_1)\ldots p(n_{D-1}), \tag{26}$$
3) Finally, obtain the impulse response, $h(\mathbf{n})$, of $H(\omega)$ by downsampling $h^{(s)}(\mathbf{n})$ with matrix, $\hat{\mathbf{M}}$, and scaling, such as
$$h(n) = ch^{(s)}(\hat{M}\mathbf{n}), \tag{27}$$
where$\hat{\mathbf{M}} = J(\mathbf{M})\mathbf{M}^{-1} = \pm[Adj\mathbf{M}]$, and $c = [J(\mathbf{M})]^{D-1}$

We can verify the support region of the filter is $\mathbf{SPD}(\pi\mathbf{M^{-T}})$. Fig.8 shows the spectrum of the designed filter. Here,the sampling matrix is $M = \begin{bmatrix}1 & 1\\2 & -1\end{bmatrix}$, and the order of filter, n, is 16.

E. Multidimensional Filter Banks Perfect Reconstruction Condition

Similar to the 1D filter banks, the perfect reconstruction condition of 2D filter banks has two parts, perfect reconstruction and alias free conditions. As we mentioned before, if the support region of the decimate filter is restricted in the $SPD(\mathbf{V})$ of the sampling matrix, there is no alias after downsampling. But this condition only covers a small part of filter banks, so we need more general filters to obtain a perfect reconstruction.

The perfect reconstruction condition for multidimensional filter banks are as follows, for simplification, we only consider

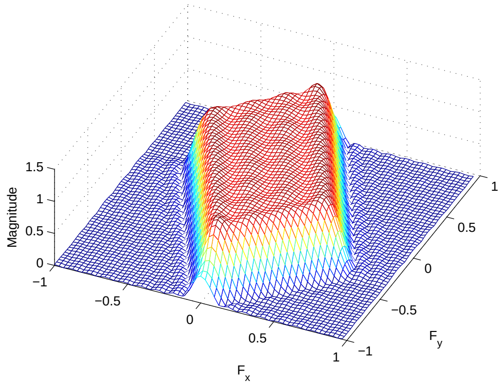


Fig. 8. The multidimensional filters from 1D filters

the 2D and 2 channel filter banks
Perfect reconstruct

$$\sum_{\mathbf{k}\in\mathcal{N}(\mathbf{M}^T)} F_{\mathbf{k}}(\omega)H_{\mathbf{k}}(\omega) = ce^{-j\omega^T\mathbf{n}_o}, \qquad (28)$$

where the $\mathbf{n}_0$ is an integer, c is a constant.

Alias free

$$\sum_{\mathbf{k}\in\mathcal{N}(\mathbf{M}^T)} F_{\mathbf{k}}(\omega)H_{\mathbf{k}}(\omega - 2\pi\mathbf{M}^{-T}\mathbf{m}) = 0, \mathbf{m}\neq 0, \mathbf{m}\in\mathcal{N}(\mathbf{M}^T). \qquad (29)$$

III. Design of Multidimensional Filter Banks[1][2][6]

For the MD filter banks, although we have the perfect reconstruction condition of the filter bank, we still can not extend the traditional design methods for 1D filter banks to the MD filter banks design, because of the lack of multidimensional factorization theorem.

In the polyphase domain, the polyphase synthesis matrix of an orthogonal filter bank is a paraunitary matrix, $U(\mathbf{z})$, if

$$U^T(\mathbf{z^{-1}})U(\mathbf{z}) = \mathbf{I}. \qquad (30)$$

For the 1D filter, we can characterize the paraunitary matrix via a lattice[5] factorization. However, in the multiple dimension case, the lattice structure is not a complete characterization. Therefore, to solve such a problem, we need to solve the nonlinear equations. Zhou introduces a novel method[1] to transform the problem into the simpler linear equations by Cayley transform. The Cayley transform of matrix $\mathbf{U}(\mathbf{Z})$ is defined as

$$\mathbf{H}(\mathbf{z}) = (I+\mathbf{U}(\mathbf{Z}))^{-1}(I-\mathbf{U}(\mathbf{Z})), \qquad (31)$$

and the inverse of CT is

$$\mathbf{U}(\mathbf{z}) = (I+\mathbf{H}(\mathbf{Z}))^{-1}(I-\mathbf{H}(\mathbf{Z})). \qquad (32)$$

The Cayley transform maps a paraunitray matrix to a para-skew-Hermitian(PSH) matrix $\mathbf{H}(\mathbf{Z})$ that satisfies:

$$\mathbf{H}(\mathbf{Z^{-1}}) = -\mathbf{H}(\mathbf{Z}),\ \text{for real coefficients.} \qquad (33)$$

From the observation of PSH condition, we find the PSH condition amounts to linear constraints on the matrix entries. It leads to an easier problem. Therefore, we can first design a PSH matrix and then map this matrix back to a paraunitary matrix by the CT. But 3 problems still need to be solved. First, how to guarantee that the matrix inverse exists. Second, CT destroys the FIR property because of the inverse. The CT of an FIR matrix is in general no longer FIR. Third, how to impose certain filter bank condition in the Cayley domain. In his paper[1], Zhou address these issues.

Proposition 1: Suppose $\mathbf{U}(\mathbf{z})$ is an $N\times N$ matrix and Λ is an $N\times N$ diagonal matrix whose diagonal entries are either 1 or -1. Then at least one Λ exists such that $\mathbf{I}+\Lambda\mathbf{U}(\mathbf{z})$ is nonsingular.

By Proposition 1, we can always find an equivalent paraunitary polyphase matrix $\mathbf{U}(\mathbf{z})$ for any orthogonal filter bank such that $\mathbf{I}+\mathbf{U}(\mathbf{z})$ is invertible and thus its CT, $\mathbf{H}(\mathbf{z})$ exists.

Proposition 2: suppose that $\mathbf{H}(\mathbf{z})$ is the Cayley transform of $\mathbf{U}(\mathbf{z})$. Then

$$\begin{aligned}\mathbf{H}(\mathbf{z}) &= 2(\mathbf{I}+\mathbf{U}(\mathbf{z}))^{-1}-\mathbf{I},\\ \mathbf{U}(\mathbf{z}) &= 2(\mathbf{I}+\mathbf{H}(\mathbf{z}))^{-1}-\mathbf{I}. \qquad (34)\end{aligned}$$

From proposition 1 and proposition 2, we can associate any orthogonal filter bank with a paraunitary matrix $\mathbf{U}(\mathbf{z})$ and PSH matrix $\mathbf{H}(\mathbf{z})$ such that both $\mathbf{I}+\mathbf{U}(\mathbf{z})$ and $\mathbf{I}+\mathbf{H}(\mathbf{z})$ are invertible. Now the first problem is solved.

lemma 1 Suppose $\mathbf{U}(\mathbf{z})$. is a paraunitary FIR matrix of McMillan degree $\mathbf{k}$. Then its Cayley transform $\mathbf{H}(\mathbf{z})$ can be written as $D(z)^{-1}\mathbf{H}'(z)$, where $D(z)$ is an FIR filter and $\mathbf{H}'(z)$ is an FIR matrix, and they satisfy the following conditions:

$$\begin{aligned}&D(z)^{-1} = cz^{\mathbf{k}}D(z),\\ &\mathbf{H}'^T(z^{-1}) = -cz^{\mathbf{k}}\mathbf{H}'(z),\\ &2D(z)^{N-1} = \det(D(z)\mathbf{I}+\mathbf{H}'(z)).\end{aligned}$$

Now we formulate the complete characterization of paraunitary FIR matrices in the Cayley domain.

Theorem 1 The CT of a matrix $\mathbf{H}(\mathbf{z})$ is a paraunitary FIR matrix if and only if it can be written as$\mathbf{H}(\mathbf{z}) = D(z)^{-1}\mathbf{H}'(z)$, where $D(z)$ is an FIR filter and $\mathbf{H}'(z)$ is an FIR matrix, and they satisfy the following four conditions:

1) $D(z^{-1}) = cz^{\mathbf{k}}D(z)$
2) $\mathbf{H}'^T(z^{-1}) = -cz^{\mathbf{k}}\mathbf{H}'(z)$,
3) $2D(z)^{N-1} = \det(D(z)\mathbf{I}+\mathbf{H}'(z)$.
4) $D(z)^{N-2}$ is a common factor of all minors of $D(z)\mathbf{I}+\mathbf{H}'(z)$.

Moreover, the CT of $\mathbf{H}(z)$ can be written as

$$\mathbf{U}(z) = \frac{\text{adj}(D(z)\mathbf{I}+\mathbf{H}'(z))}{D(z)^{N-2}} - \mathbf{I}.$$

Therefore, our problem of designing a paraunitary FIR matrix $\mathbf{U}(z)$ is converted to a problem of designing a PSH matrix $\mathbf{H}(z) = D(z)^{-1}\mathbf{H}'(z)$, where $D(z)$ and $\mathbf{H}'(z)$ satisfy the condition given in the theorem 1. The second problem is solved.

For the third problem, no general theorem can be used to solve it, because the solution depends on the real problem.

IV. Directional Vanishing Moments Filter Bank

In the one dimensional case, a filter H(z) is said to be a vanishing moments filter, if H(z) can be factorized as

$$H(z) = (1 - z)^d R_1(z). \tag{35}$$

. Therefore, this filter has d zeros at z=1 or $\omega = 0$ on the unit circle; d here is the order of vanishing moments. If the discrete polynomial signal of degree less than d, such as $x(n) = \sum_{j=0}^{i} \alpha_j n^j$, filtering x(n) with H(z) produces a zero output. The filter completely annihilates the discrete polynomials of degree less than d. This vanishing moments filter can also be found in the 2 dimensional filter banks. But 2D vanishing moments have direction. We can define the directional vanishing moments filter as follows:

define 1 Let $C(\mathbf{z})$ be a discrete filter and $\mathbf{u} = (u_1, u_2)^T$ be a 2-D vector of coprime integers. We say $C(\mathbf{z})$ has a DVM of order d along the direction $\mathbf{u}$ if it factors as

$$C(z_1, z_2) = (1 - z_1^{u_1} z_2^{u_2})^d R(z_1, z_2). \tag{36}$$

A particular signal is also annihilated by the 2D DVM filter, as in 1D case. To find the character of such a signal, we introduce the directional polyphase representation.

Lemma 1: Suppose that $\mathbf{u} \in \mathbb{Z}^2$ and $u_2 \neq 0$. Then for every $\mathbf{n} \in \mathbb{Z}^2$, there exists a unique pair $(k, \mathbf{r})$ where $k \in \mathbb{Z}$, $\mathbf{r} \in \mathcal{R} : \mathbb{Z} \times \{0, 1, \ldots, |u_2| - 1\}$ such that

$$\mathbf{n} = k\mathbf{u} + \mathbf{r}. \tag{37}$$

This lemma allows us to partition any 2-D signal $x(\mathbf{n})$ into a set of disjoint 1-D signals, $\{x_{\mathbf{r}}(\mathbf{k})\}_{\mathbf{k}\in\mathcal{R}}$, with

$$x_{\mathbf{r}}(\mathbf{k}) = x(k\mathbf{u} + \mathbf{r}). \tag{38}$$

Proposition 1: Let $C(z_1, z_2)$ be a 2-D filter with a factor $(1 - z_1^{u_1} z_2^{u2})^d$. Then a signal $x(\mathbf{n})$ is annihilated by $C(\mathbf{z})$. if each 1-D signal $x_{\mathbf{r}}(\mathbf{k})$ defined in (38) is a discrete polynomial of degree less than d.

From proposition 1, we know that the discrete signal sampled from a continuous-time signal, which is smooth along a given direction, is also annihilated by $C(\mathbf{z})$. To illustrate, we filter a piece wise smooth image with a 2-D filter with a third order DVM along the direction, $u = (1, 2)^T$. This image has two types of edges, one is along the direction, u, another is along the direction perpendicular to u. The image is well approximated by a piecewise polynomial image of sufficiently large degree d. As shown in Fig.9, the edges along the direction, $\mathbf{u}$, are annihilated, but the other ones are preserved. The frequency response of the DVM filter is shown in Fig.10. You can clearly see the vanishing points along the direction, $\mathbf{u}$. Now, we discuss the general two channel 2D filter bank with a valid sampling matrix, S, whose downsampling ratio is 2. In this setting, given a set of analysis/synthesis filters $\{H0(\mathbf{z}), H1(\mathbf{z}), G0(\mathbf{z}), G1(\mathbf{z})\}$ the perfect reconstruction condition is

$$H_0(\mathbf{z})G_0(\mathbf{z}) + H_1(\mathbf{z})G_1(\mathbf{z}) = 2, \tag{39}$$

$$H_0(\mathbf{W}^{\mathbf{k_1}}_{\mathbf{S^{-T}}} \circ \mathbf{z})G_0(\mathbf{z}) + H_1(\mathbf{W}^{\mathbf{k_1}}_{\mathbf{S^{-T}}} \circ \mathbf{z})G_1(\mathbf{z}) = 0, \tag{40}$$

where k1 is the nonzero integer vector in the set $\mathcal{N}$, and $\circ$ denotes the direct product between vector entries. The modulation term $\mathbf{W}^{\mathbf{k_1}}_{\mathbf{S^{-T}}}$ is a function of the sampling lattice generated by $\mathbf{S}$ and has the form $\mathbf{W}^{\mathbf{k_1}}_{\mathbf{S^{-T}}} = (e^{j\pi n_1}, e^{j\pi n_2})^T$, where $(n_1, n_2)^T = \mathbf{S^{-T}k_1}$. Note that all matrix generators of a given lattice are equivalent, up to right multiplication by a unimodular integer matrix. All the sampling matrix can be equivalent to the matrix

$$S_0 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad S_1 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \tag{41}$$

or the matrix generated by exchanging the two diagonal entries

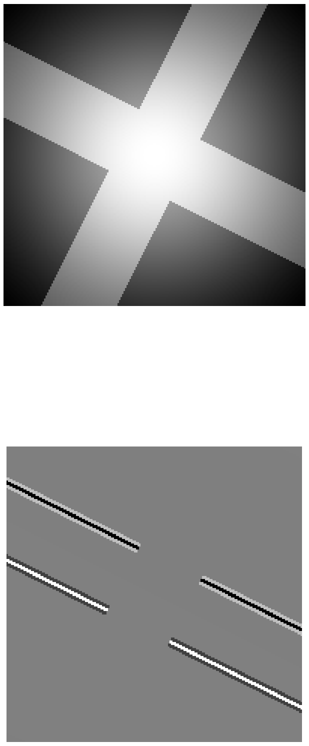


Fig. 9. The illustration of DVM as an edge annihilator.

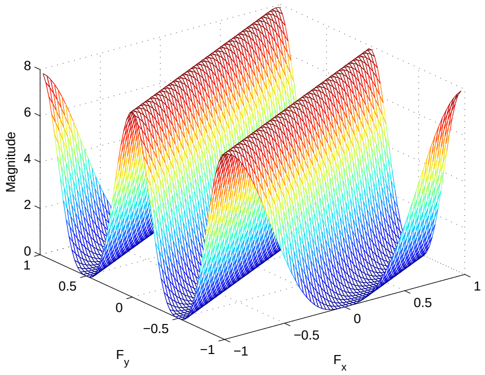


Fig. 10. The frequency response of the DVM filter

of S_1. We can check that $\mathbf{k}_1 = (1,0)^T$ for both S_0 and S_1 so that $\mathbf{W}^{\mathbf{k_1}}_{\mathbf{S^{-T}}} = (-1,-1)^T$ and $\mathbf{W}^{\mathbf{k_1}}_{\mathbf{S^{-T}}} = (-1,1)^T$ for S_0 and S_1 respectively.

If we assume the filter is FIR, we can show the synthesis filters are completely determined from the pair $(H_0(\mathbf{z}), G_0(\mathbf{z}))$ through the relation

$$(H_1(\mathbf{z}), G_1(\mathbf{z})) = (\mathbf{z}^{\mathbf{k_1}} G_0(\mathbf{W}^{\mathbf{k_1}}_{\mathbf{S^{-T}}} \circ \mathbf{z}), \mathbf{z}^{-\mathbf{k_1}} H_0(\mathbf{W}^{\mathbf{k_1}}_{\mathbf{S^{-T}}} \circ \mathbf{z}). \tag{42}$$

The perfect reconstruction condition reduces to

$$H_0(\mathbf{z})G_0(\mathbf{z}) + H_0(\mathbf{W}^{\mathbf{k_1}}_{\mathbf{S^{-T}}} \circ \mathbf{z})G_0(\mathbf{W}^{\mathbf{k_1}}_{\mathbf{S^{-T}}} \circ \mathbf{z}) = 2. \tag{43}$$

In general, given the desired direction of the zero moment, the analysis filter $H_0(\mathbf{z})$ takes the form

$$H_0(\mathbf{z}) = (1 - z_1^{u_1} z_2^{u_2})^L R_{H_0}(\mathbf{z}). \tag{44}$$

where L denotes the order of the DVM. Substituting this equation in (42), we obtain the design equation

$$(1-\mathbf{z}^{\mathbf{u}})^L R(\mathbf{z}) + (1-(\mathbf{W}^{\mathbf{k_1}}_{\mathbf{S^{-T}}} \circ \mathbf{z})^{\mathbf{u}})^L R(\mathbf{W}^{\mathbf{k_1}}_{\mathbf{S^{-T}}} \circ \mathbf{z}) = 2. \tag{45}$$

where $R(\mathbf{z} = R_{H_0}(\mathbf{z})G_0(\mathbf{z})$. If the filter is FIR, $\mathbf{u}^T 2\mathbf{S}^{-T}\mathbf{k}_1$ should be odd. The perfect reconstruction condition reduce to

$$(1 - \mathbf{z}^{\mathbf{u}})^L R(\mathbf{z}) + (1 - \mathbf{z}^{\mathbf{u}})^L R(\mathbf{W}^{\mathbf{k_1}}_{\mathbf{S^{-T}}} \circ \mathbf{z}) = 2. \tag{46}$$

Proposition 2: Consider the filter equation (42) where $\mathbf{u}$ has coprime entries and $\mathbf{u}^T 2\mathbf{S}^{-T}\mathbf{k}_1$ is odd. Then there exists a unimodular integer matrix $\mathbf{U}$ such that if $R(\mathbf{z})$ solves (42) then $\tilde{R}(\mathbf{z}) = R(\mathbf{z}^{\mathbf{U}})$ solves

$$(1 - z_1)^L \tilde{R}(\mathbf{z}) + (1 + z_1)^L \tilde{R}(\mathbf{W}^{\mathbf{k_1}}_{\mathbf{\tilde{S}^{-T}}} \circ \mathbf{z}) = 2. \tag{47}$$

where $\tilde{\mathbf{S}} = \mathbf{US}$. Conversely, if $\tilde{R}(\mathbf{z})$ is a solution to (47) with $\tilde{S}$ and $\mathbf{u}$ as above, then there is a matrix $\tilde{U}$ such that $R(\mathbf{z}) = \tilde{R}(\mathbf{z}^{\tilde{\mathbf{U}}})$ is a solution to (42) with $\mathbf{S} = \tilde{\mathbf{U}}\tilde{\mathbf{S}}$. The integer unimodular matrix $\mathbf{U}$ is a resampling matrix. Hence, the mapping of U is just resampling operation of the filter $R(\mathbf{z})$. From the Proposition 2, we can simplify the design of an arbitrary directional filter to the problem of designing a horizontal direction DVM. Just following the step

1) Design the DVM filter $\tilde{H}(\mathbf{z})$ for the horizontal direction.
2) Find the unimodular matrix U which can transform the direction $\mathbf{u}$ into horizontal direction.
3) Substitute $\mathbf{z}$ with $\mathbf{z}^{\mathbf{U}^{-1}}$ and $H(\mathbf{z}) = \tilde{H}(\mathbf{z}^{\mathbf{U}^{-1}})$
4) The filter H(z) is the desired DVM filter

Proposition 3: Consider a two-channel 2-D filter bank with FIR filters $\{H_0(\mathbf{z}), H_1(\mathbf{z}), G_0(\mathbf{z}), G_1(\mathbf{z})\}$ and the downsampling matrix S. Let $\mathbf{u} = (u_1, u_2)^T$ and $\mathbf{v} = (v_1, v_2)^T$ be two distinct admissible directions. Then the filter bank cannot be perfect reconstruction if one of the follows is true:

1) The filter $H_0(\mathbf{z})$ factors $(1 - \mathbf{z}^{\mathbf{u}})$ and $H_1(\mathbf{z})$ factors $(1 - \mathbf{z}^{\mathbf{v}})$ leaving FIR remainders.
2) One of the filters, say $H_0(\mathbf{z})$ factors $(1-\mathbf{z}^{\mathbf{u}})$ and $(1-\mathbf{z}^{\mathbf{v}})$ simultaneously leaving an FIR remainder.

This proposition shows that DVMs can only be in one branch of filter bank.

Proposition 3: Let $s = \pm 1$, An FIR filter$\tilde{R}(z_1, z_2)$ is solution to the equation

$$(1 - z_1)^L \tilde{R}(z_1, z_2) + (1 + z_1)^L \tilde{R}(-z_1, sz_2) = 2, \tag{48}$$

if and only if it has the form

$$\tilde{R}(z_1, z_2) = R_L(z_1) + (1 + z_1)^L R_o(z_1, z_2) \tag{49}$$

with $R_L(z_1)$ being a univariate solution given explicitly by

$$R_L(z_1) = \sum_{i=0}^{L-1} \left(\begin{array}{c} L+i-1 \\ L-1 \end{array} \right) 2^{-(L+i-1)} (1 + z_1)^i, \tag{50}$$

and $R_o(\mathbf{z})$ satisfying

$$R_o(z_1, z_2) + R_o(-z_1, sz_2) = 0. \tag{51}$$

Due to the lack of factorization theorem for 2-D polynomials, the design of the non-separable 2D filter banks is harder than the 1-D counterpart. But we can design such filters by mapping from the one dimension filter banks. The procedure used for designing such DVM filter from the 1D filter is as follows:

Design 2-D filters $H_0(\mathbf{z})$ and $G_0(\mathbf{z})$ satisfying the perfect reconstruction condition (42) and such that $H_0(\mathbf{z})$ factors $(1 - z_1)^{N_a}$ and $G_0(\mathbf{z})$ factors $(1 - z_1)^{N_s}$

1) Design 1-D filter $H_0^{(1D)}(z)$ and $G_0^{(1D)}(z)$ with N_a/L andN_s zeros at some point $c_0 \in \mathbb{C}$ respectively, and such that $P^{(1D)}(z) = H_0^{(1D)}(z)G_0^{(1D)}(z)$ satisfies
$$P^{(1D)}(z) + P^{(1D)}(-z) = 2. \tag{52}$$
2) Let $M(\mathbf{z}) = (1 - z_1)^L \tilde{R}(\mathbf{z}) + c_0$ with
$$\tilde{R}(\mathbf{z} = \tilde{R}_L(z_1) + (1 + z_1)^L \tilde{R}_o(z_1, z_2), \tag{53}$$
and $R_o(z_1, z_2) + R_o(-z_1, sz_2) = 0$
3) Set $H_0(z) = H_0^{(1D)}(M(\mathbf{z}))$ and $G_0(z) = G_0^{(1D)}(M(\mathbf{z}))$ to obtain the desired 2-D filters.

V. DVM Design Example

In this example, we chose the one dimension phototype filter as

$$\begin{array}{l} H_0^{(1D)}(z) = \frac{1}{2}(1 + z)(2 - (2 - \sqrt{2})z) \\ G_0^{(1D)}(z) = \frac{1}{2}(1+_z)(2 - (6 - 4\sqrt{2})z + (4 - 3\sqrt{2}z^2) \end{array} . \tag{54}$$

And chose the sampling matrix S_0 and

$$M(z_1, z_2) = (1 - z_1)^2 R(z_1, z_2) - 1. \tag{55}$$

We also chose

$$R(z_1, z_2) = -\frac{z_1^{-1}}{2} + (1 + z_1^{-1})^2 R_o(z_1, z_2). \tag{56}$$

We can also chose $R_o(z_1, z_2)$ has following format

$$R_o(z_1, z_2) = r_0(z_1+z_1^{-1}) + r_1(z_2+z_2^{-1}) + r_2(z_1+z_1^{-1})(z_2+z_2^{-1})^2, \tag{57}$$

and optimize the coefficients r_0, r_1, r_2 so that the filter approximate the idea fan response. The optimized parameter is $[r_0, r_1, r_2] = [0.1172, -0.1612, -0.0908]$

Fig.11 shows the frequency response of analysis filter and synthesis filter in the filter bank.

Next, we apply those filter to the images to see the effect of the filter bank. The new filter bank has two stages, each stage has DVMs along the different direction, as shown in Fig.12. we chose the direction for the second stage to be $\mathbf{u} = [-3, 1]$. To obtain the second directional vanishing moments filter, we chose unimodular matrix

$$U = \begin{bmatrix} -1 & -2 \\ -1 & -3 \end{bmatrix}; \qquad (58)$$

First, we apply the filter bank to the image which is shown in Fig.2. The result is shown in Fig.13 and Fig.14.

We also apply the same filter bank to the image barb. Fig.15 and Fig.16 show the filter result.

From the result, we can see each subband at least passes the edge information along one direction except for the low-pass filter at the upper-left corner. Therefore those filter banks can help us capture the edge information of an image. This property is useful in many application.

VI. Conclusion

We have discussed the fundamentals of multidimensional filter banks. Methods which can reduce the complexity of designing a MD filter bank are discussed. The directional vanishing moments filter bank, is introduced. The design example and simulation result show that the DVM filter banks can easily capture the geometric structure of an image, and give a better compression performance.

References

[1] J. Zhou, M. N,Do, and J. kovaeevic "Multidimesional orthogonal filter bank characterization and design using the cayley transform" *IEEE Trans. Image Processing.*,Vol. 14, No.6, June 2005

[2] J. Zhou and M. N. Do, "Multidimensional multichannel FIR deconvolution using Grobner bases", *IEEE Transactions on Image Processing*, to appear.

[3] A.L.Cunha M.N.Do ,"On two-channel filter banks with directional vanishing moment",*IEEE Tran. Img. Processing*, 2005

[4] A.L.Cunha and M.N.Do,"Bi-orthogonal filter banks with directional vanshing moments" ICASSP, Philadelphia, PA, 2005

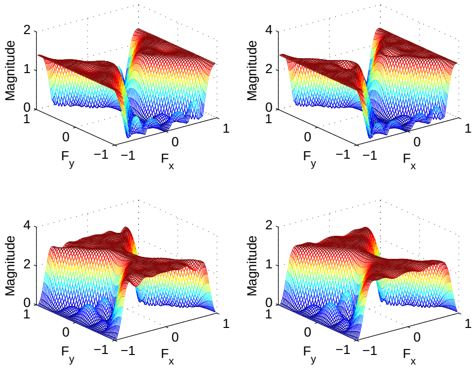


Fig. 11. The frequency response of four filters of the DVM 2D filter bank

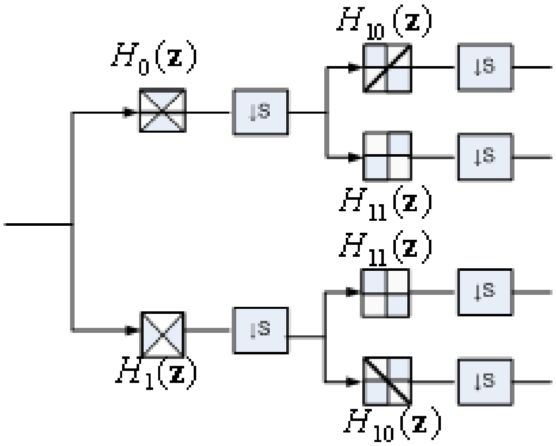


Fig. 12. The DVM directional filter bank with two stages.

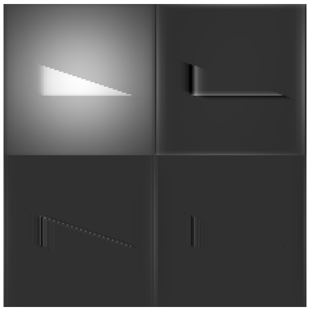


Fig. 13. The result of passing the triangle image through the filter bank once.

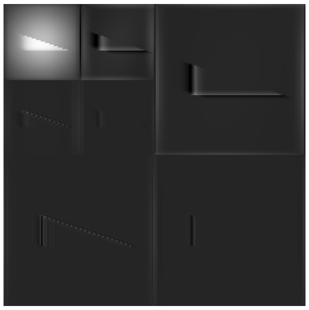


Fig. 14. The result of passing the triangle image through the filter bank twice.

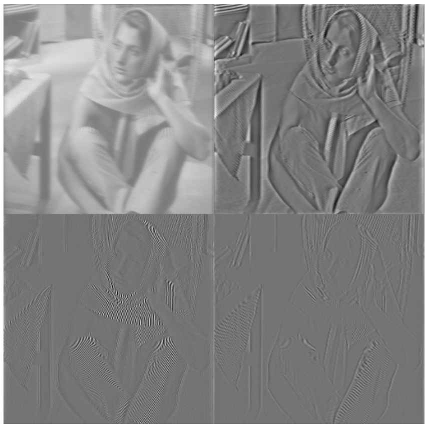


Fig. 15. The result of passing the Barb image through the filter bank once.

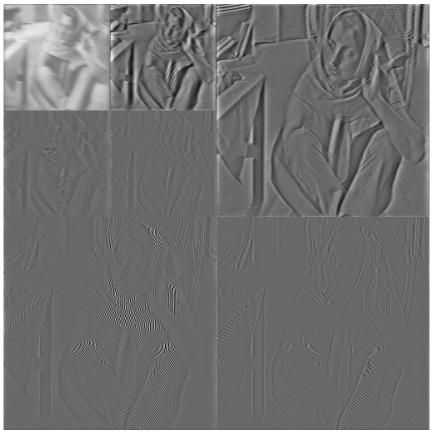


Fig. 16. The result of passing the Barb image through the filter bank twice.

[5] P.P. Vaidyanathan, *Multirate systems and filter banks*, Prentice-Hall, 1993

[6] J. Zhou, M. N. Do, and J. Kovacevic, “Special paraunitary matrices, Cayley transform, and multidimensional orthogonal filter banks,” *IEEE Transactions on Image Processing*, vol. 15, no. 2, pp. 511-519, Feb. 2006.

[7] Ansari, R., and Lau C.L.: “Two-dimonsional IIR filters for exact reconstruction in tree-strucutred subband decomposition” *Electonics Letters.*, vol. 23, June 1987

Guoqian Sun received the B.E. and M.E. degrees from Southeast University, China, in 2000 and 2003, Since March 2005, he has been a Master student at the School of Engineering Science, Simon Fraser University, Canada. He was software engineer with research center in Nanjing of ZTE. Cooperation from 2003 to 2005.

His current research interests include multirate signal processing, image/video coding and processing.

A Novel Magnetic Valve Device for Glaucoma Drainage Devices

Kouhyar Tavakolian, *Student Member IEEE*, kouhyar@ieee.org
Computational and Integrative BioEngineering Research (CIBER)
Faculty of Applied Sciences, Simon Fraser University
8888 University Drive, Vancouver, BC, Canada, V5A 1S6

Abstract— **In this paper glaucoma drainage devices are reviewed and compared, and a novel valve device is proposed for treatment of glaucoma. By opening or closing, the proposed valve can be used to reduce high pressures in the eye (hypertony), or prevent low pressures in the eye (hypotony) respectively. The device works based on the buckling of fluid channel connected to the anterior chamber of the eye and includes a magnetically driven valve.**

Index Terms— **Glaucoma drainage devices (GDD), Intraocular pressure (IOP), MEMS, PDMS, and Microfabriation.**

I. INTRODUCTION

Glaucoma is a common disease of the eye and is the second leading cause of blindness in the world, according to the World Health Organization [1]. There are presently three main therapies for glaucoma: medical treatment with drugs, glaucoma filtration surgery (GFS), and implantation of glaucoma drainage device (GDD). The choice of therapy depends on the patient and is influenced by several factors, including the risks associated with the treatment, the likelihood of its success, and the clinical presentation of the patient [2].

Devices that are surgically inserted into the eye in order to cure glaucoma by increasing the outflow of the aqueous humor fluid are called by a variety of names. Perhaps the most common name is (GDD). Other names include glaucoma filtration implant (GFI) and combinations including the words implant, shunt or drain. Common issues with GDDs include: not opening when they should, not closing when the eye pressure drops, and having varying opening set points. These problems necessitate more research work on development of GDDs [3]. Improving the design of GDDs will lead to increased success rates in glaucoma treatment, consequently GDDs will be more often selected as a replacement for medical treatment or surgery and this is the main motive behind our proposed GDD mechanism [2].

This paper discusses the main problems of previous GDD designs and finally proposes a novel magnetic device as a solution. The proposed device takes advantage of the state-of-the-art micro-electro-mechanical sensor (MEMS) technology and advancements in micro fabrication technology. In section two, there is an introduction to glaucoma and GDDs. Section three of the paper explains the magnetic mechanism which is going to be used as the valve. Section four explains the experiments, which is followed by discussion and conclusion sections.

II. GLAUCOMA AND GDD

In this section the glaucoma disease is briefly introduced and glaucoma drainage devices are explained as the main context of this research.

A. Glaucoma and related eye anatomy

The small volume in front of the eye's lens, as can be seen in Fig.1, is of most interest in this research. This volume is divided into two parts by the iris, which controls the aperture of the lens. The volume between the iris and the cornea is the anterior chamber (250 µl) and the volume between the iris and lens is the posterior chamber (60 µl). Both chambers are filled with a clear liquid known as aqueous humour [2]. With reference to Fig.1, the aqueous humor passes through the pupil into the anterior chamber, and thereafter drains out of the eye into the canal of Schlem. Normal intraocular pressure

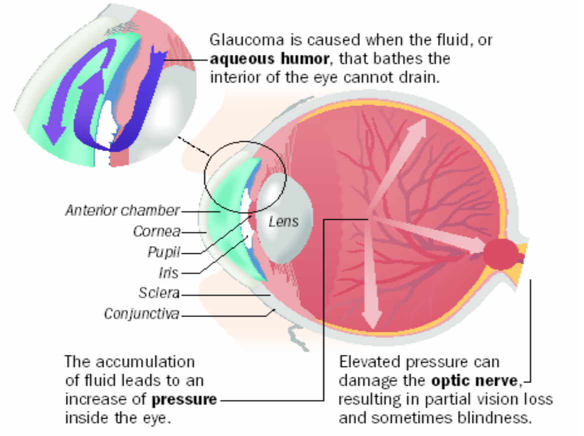

Glaucoma is caused when the fluid, or **aqueous humor**, that bathes the interior of the eye cannot drain.
Anterior chamber
Cornea
Pupil
Iris
Sclera
Conjunctiva
Lens
The accumulation of fluid leads to an increase of **pressure** inside the eye.
Elevated pressure can damage the **optic nerve**, resulting in partial vision loss and sometimes blindness.

Fig.1. Anatomy of an eye with glaucoma [8]

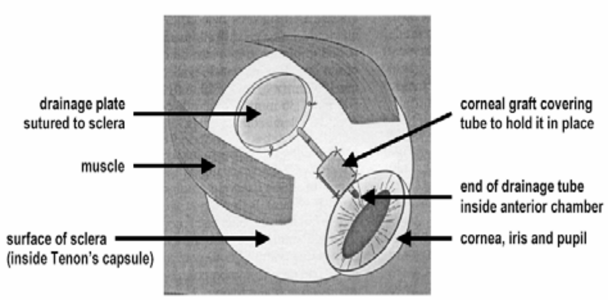


Fig.2. Schematics representation of a GDD [3]

is typically about 15±4 mm Hg, but may rise to 21 mm Hg. Pressures within the eye significantly above this range are considered abnormal. This high pressure is one of the main causes and symptoms of Glaucoma.

The balance between the rate of formation and drainage of aqueous adjusts the pressure in the eye. The rate of aqueous formation is relatively constant at 2-2.5 µl/min though the outflow varies. Many forms of glaucoma are caused by a reduction in the outflow of aqueous, leading to an increase in intraocular pressure (IOP). Thus, increasing the rate of outflow is desirable as it can reduce IOP [4].

B. Glaucoma Drainage Devices

GDDs are surgically inserted into the eye to increase the outflow of the aqueous. GDDs create an alternate aqueous pathway by channeling it from the anterior chamber through a long tube to a plate inserted under the conjunctiva [3]. Most tubes used so far are made of silicone but differences between the devices lie in the nature of the arrangements for drainage and pressure control [2]. Fig.2 shows the general arrangement of a typical drainage device and its positioning in the eye.

After the GDD is implanted, the body tissues react to it as a foreign object and encapsulate it with fibrous tissue that forms a bleb. The bleb is particularly important to the operation of the device, because it is the principal source of resistance to the flow of aqueous through the device. Aqueous humor flows freely out of the anterior chamber, through the tube and onto the drainage plate, where it spreads across the surface of the sclera with minimum resistance. The bleb must grow to a suitable size to regulate the pressure. This is a major feature which determines the success or failure in an individual case [3].

The first GDD was proposed by Molteno and its design inspired other researchers to develop similar devices [5].Other common GDDs are Ahmed glaucoma implant [6], and Baerveldt [7]. The basic designs of these devices are similar in that a silicone tube, channels aqueous humor from the anterior chamber to a fibrous capsule surrounding the plate. The capsule serves as a reservoir for aqueous drainage. Studies have demonstrated that these devices are comparable and are effective in treating patients with Glaucoma [3].

Each GDD design has a group of supporters in the surgical community. The Baerveldt is the most popular among glaucoma specialists. This is because of its bigger surface that reduces the chance of post inflammation. The Ahmed GDD is the easiest to implant, so is more popular among surgeons who have less experience dealing with the common problems following glaucoma surgery. The Molteno has the longest track record, and no GDD has ever been found to give better results. However the Molteno requires surgery in both superior quadrants of the eye, which may explain the popularity of Ahmed and Baerveldt [3].

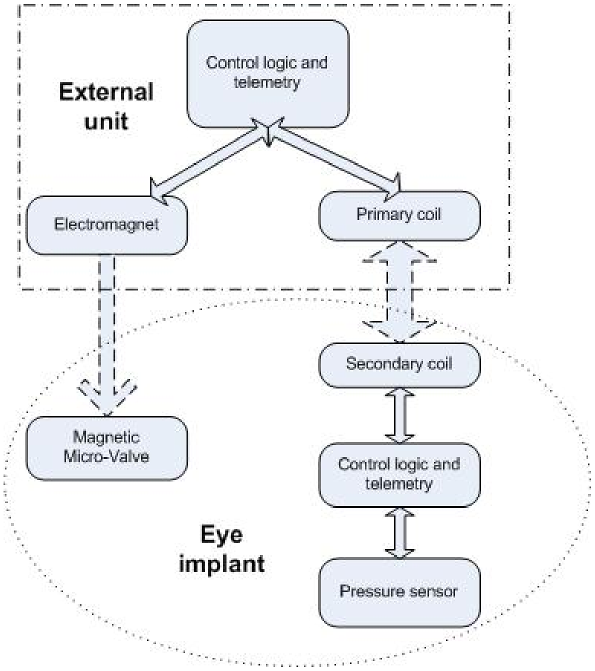


Fig.3. The block diagram of the proposed system for the magnetic actuated GDD. The IOP is sensed and sent out to the external unit in which the decision is made to open or close the magnetic micro-valve according to the measured IOP.

Unfortunately no GDD is ideal and there are several issues still need to be resolved. The most important problems affecting the design of GDDs are immediate hypotony (loss of IOP), overgrowth of the bleb, tube fit and blockage, and valves not closing [2][3][4].

Prior to bleb formation another source of resistance in the flow path is needed. Techniques for creating resistance include implanting the tube several weeks after the drainage plate and temporarily obstructing the tube either by clamping it shut or filling it with material. These solutions have themselves a number of complications: variable delay in lowering IOP and need for second operation [4]. A GDD capable of preventing hypotony would be a valuable device. This research work addresses this problem by proposing a device which will be explained in the next section.

III. DEVICE MECHANISM

In this section the magnetic device is introduced. The proposed magnetic valve can be seen in the block diagram of Fig.3. The pressure sensor can be embedded in the same system or an already designed pressure sensor can be used for sensing IOP [13], [14] [15]. Whenever the IOP falls lower than the 5mmHg, the electromagnet instructs the magnetic valve to close, thereby avoiding hypotony in the eye

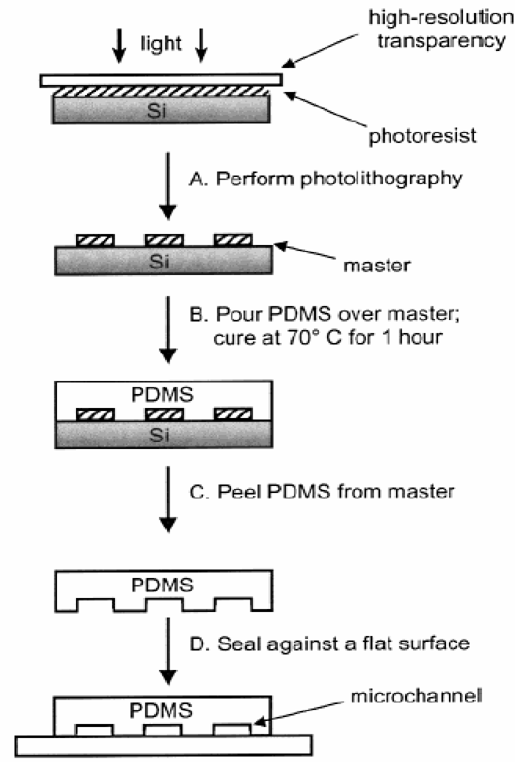

light
high-resolution transparency
Si
photoresist
A. Perform photolithography
Si
master
B. Pour PDMS over master; cure at 70° C for 1 hour
PDMS
Si
C. Peel PDMS from master
PDMS
D. Seal against a flat surface
PDMS
microchannell

Fig.4. The fabrication process followed in this project to fabricate the PDMS channels which is based on soft lithography idea [11]

The proposed valve mechanism is created through fabrication of elastomeric components in microfluidic systems that is based on deformation of elastic materials that have been impregnated with magnetic components [10]. Controllable miniature electromagnets can be used to activate switching valves according to the eye pressure.

Whitesides and Xia [11] proposed an alternative, non-photolithographic set of microfabrication methods named soft lithography. This process can be seen in Fig.4.

In [12] a technique called “multilayer soft lithography” was introduced that combines soft lithography with the capability to bond multiple patterned layers of elastomer. Multilayer structures are constructed by bonding layers of elastomer, each of which is separately cast from a micromachined mold.

The proposed device in this paper is an extension of methods for rapid prototyping of fluidic microchannels based on replica molding of elastomers, such as PDMS [12]. Silicone polymers such as PDMS have high compliance, high elongation, and good sealing properties, making them useful valving materials. Elastomer-based, valving strategies have been reported by [11], who described pneumatic methods for deforming elastomeric microchannels to form microvalves and micropumps. By combining elastomers with magnetic materials and using electromagnets in the substrate, active valves can be fabricated. This approach is simpler and allows for miniaturization, by removing the need for macroscopic, externally switched pneumatic supplies [10].

For ferromagnetic force, the magnetic pressure *F/A (*force per unit area) between an electromagnet and a ferromagnetic material is approximately $F/A=B^2/2\mu_0$, where B is the flux density, and μ_0 is the permeability of free space. The flux density B can be estimated with magnetic circuit approximations or calculated using finite element analysis software [10].

IV. EXPERIMENTS

For the fabrication of the valve, PDMS was used as explained previously because of its mechanical properties, its ability to be shaped in soft lithography process and the ease to be impregnated by magnetic materials in the next steps of the experiments.

In this stage of the project the micro channels were fabricated according to the Fig.5 process. GE RTV 615 was used as the PDMS pre-polymers and SU8-10 [18] was used as the master mold. Different channel widths of 75, 100, 150 and 200 µm were fabricated as network of parallel channels. The channel height was designed to be 10 µm by spinning the SU8 photoresist in 3000 rpm for 30 seconds. The mixture of the prepolymers A and B of RTV 615 was poured on this SU8 master mold and was cured in 70°C for one hour. Another network of PDMS was sealed on the top of this network which had a different mixture ratio of prepolymers. This different ratios of primary prepolymers made the adhesion of the two layers possible as can be seen in Fig.6.

This has been so far very similar to [12] which resembles a pneumatic valve, the next step is to impregnate the prepolymers with Iron powder to be able to close the valve magnetically [10]. This way by sensing the IOP in the eye, while it drops under 5mmHg, the valve is closed magnetically to avoid hypotony as seen in Fig.3.

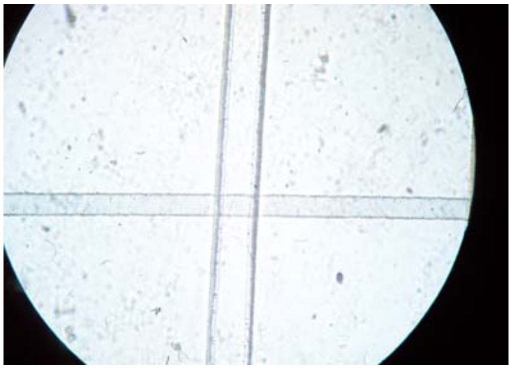


Fig.5. The PDMS channels the horizontal channel is 75 and the vertical is 200 micron widths. The channels heights are 10 micron

V. DISCUSSION

The proposed valved device is based on the buckling of fluid channels connected to the anterior chamber of the eye as seen in Fig.2. The fluid channels were fabricated for the magnetic device using a soft lithography technique. These channels are the main body of the proposed magnetic valve.

The next step is the impregnation of the PDMS with iron. This will give magnetic properties to the polymer. The channels were implemented as a network of parallel channels, so if one of them is buckled by aqueous humor particles the others will still continue functioning.

After the completion of device fabrication there is a need for an in-vitro setup to test the functionality of the device before using it in the eye. A proper test setup can be found in

[17]. The test setup consists of a syringe pump connected to a glass standpipe via a 3-way stopcock as can be seen in Fig.6. The third outlet of the stopcock is connected by a 27 gauge needle to a GDD, which is water sealed inside a glass test tube to avoid any surface tension effect from gas-liquid interface. A 21-gauge needle from the tube is then connected to an Anopore™ filter.

The Anopore™ membrane has a filtration area of 80 mm2 (flow resistance 64 mmHg.min/mL) when placed in its holder. In order to get an area of about 1.5 mm2 (flow resistance 3400 mmHg.min/mL), corresponding to the physiological pressure drop across the in vivo fibrous capsule, an ultraviolet curable epoxy is used to seal the extra area in the Anopore™ membrane. Filtered saline (the same viscosity as aqueous humor) was used as the working fluid throughout these experiments, all performed at room temperature [17].

As the proposed magnetic valve is supposed to be used with a microelectronic system, its operation requires electrical power. This power will be supplied wirelessly as shown in Fig.5. On the other hand, the ability to continuously monitor the IOP can give other possibilities to the eye specialist to monitor the eye condition and gives more accurate control over the pressure.

VI. CONCLUSION

In this paper a new device was proposed as a valve for glaucoma treatment. In this research work the main focus is the fabrication of the magnetic microvalve for GDD. The development of the general system of Fig.3 will be the focus of future research. The buckling mechanism used in this research is not necessarily limited to the magnetic mechanisms and other methods such as electrowetting are also among feasible options [18].

ACKNOWLEDGEMENTS

I would like to thank Dr. Bonnie Gray and Fartash Vasefi for their important helps in SFU microfabrication laboratory.

REFERENCES

[1] Andre Mermoud “Bulletin of the World Health Organization” November 2004, 82 (11).

[2] D J Howorth, “Feasibility Study for a Micromachined Glaucoma Drainage Device” M.Sc thesis, Cranfield University, 2002.

[3] Lim, K.S., Allan, B.D.S., Lloyd, A.W., Muir, A. and Khaw P. T., “Glaucoma drainage devices; past, present, and future” British Journal Ophthalmology, 82, 1083-1089, 1998.

[4] Cristina Rodica Neagu “A Medical Microactuator based on an Electrochemical Principle” M.Sc. thesis, Twente University, 1998.

[5] Molteno GDD: http://www.molteno.com/

[6] Ahmed valve: http://www.ahmedvalve.com/

[7] Baerveldt GDD: http://www.baerveldt.com/

[8] http://www.optonol.com

[9] T. Pan, M. S. Stay, J. D. Brown, B. Ziaie, David Brown, “Modeling and Characterization of a Valved Glaucoma Drainage Device With Implications for Enhanced Therapeutic Efficacy” IEEE Transactions on Biomedical Engineering, Vol. 5, May 2005.

[10] William C. Jackson, Michael J. O’Brien, Emmanuil Rabinovich, and Gabriel P. Lopez “Rapid prototyping of active microfluidic components based on magnetically modified elastomeric materials” J. Vac. Sci. Technol. Apr 2001.

[11] Younan Xia, G.M. Whitesides, “Soft Lithography” Annual Review of Material Science”, pp.153-184, 1998.

[12] Unger, M. A.; Chou, H.-P.; Thorsen, T.; Scherer, A. “Monolithic Microfabricated Valves and Pumps by Multilayer Soft Lithography” Science2000, 288, 113-116.

[13] Rosengren, L., Backlund, Y., Sjostrom, T., Hok, B. and Svedbergh, “A system for wireless intra-ocular pressure measurements using a silicon micromachined sensor” J. Micromech. Microeng., 2, 202–204, 1992.

[14] Schnakenberg, U., Walter, P., Bögel, v.G., Krüger, C., Lüdtke-Handjery, H.C., Richter, H.A., Specht, W., Ruokonen, P. and Mokwa W. “Initial investigations on a system for measuring intraocular pressure” Sensors and Actuators A, 85, 287-291, 2000.

[15] Puers, R., Vandevoorde, G. De Bruyker, “Electrodeposited copper inductors for intraocular pressure elemetry”. J Micromech & Microeng , 10(2), 124-129, 2000.

[16] http://www.microchem.com/

[17] Tingruri Pan, J. David Brown, Babak Ziaei, “A Microfold Test-Bed with Nanopore Membranes for In-Vitro Simulations of Flow Characteristics of Glaucoma Drainage Devices” Proceeding of the second joint EMBS/BMES Confrence, Houston, USA, Oct 2002.

[18] Frieder Mugele, Jean-Christophe Baret, “Electrowetting: from basics to applications” Institute of Physics, Journal of Physics, 2005.

Kouhyar Tavakolian was born on July 12th,1979 in Kermanshah, Iran. He received his B.Sc. (Biomedical Engineerin), from Tehran Polytechnics, in 2000 and his MSc. (Bio-Electrical Engineering, from University of Tehran, 2003 focusing on the designing of brain computer interface systems. He got another M.Sc. degree in computer science from University of Northern British Columbia in 2005. He is currently a Ph.D. candidate at Simon Fraser University in the area of Biomedical engineering. He has been IEEE student member since 2000 and has 18 reviewed publications and has been paper reviewer for different IEEE conferences.

Molding PDMS Channels and an Embedded Detector Chamber

Takaya Ueda and Bonnie Gray

***Abstract*— In this paper, the preliminary work for a cell retention system with an integrated photo-detector is shown. The completed system will be able to retain a single cell in its channel and detect the fluorescence reaction of the cell to light stimulus. The scope of the paper consists of fabricating SU-8 structures that serve as mold masters for the cell channels and photo-detector mount, and using the SU-8 masters to mold the system parts with poly-dimethysiloxane (PDMS). The 53 µm high channel structures were fabricated using PDMS. Most structures were fabricated successfully, but some structures were cracked or washed away. The SU-8 structures for the detector mount were fabricated by spinning two thick layers of SU-8 100. The structures were fabricated successfully in terms of adhesion to the wafer, but the second layer was misaligned relative to the first layer. The PDMS was poured onto the SU-8 masters and cured for 2 hours at 100 °C. The molding process resulted in an imprint of the SU-8 structures, but the bubbles trapped in PDMS cause optical interference for the photo-detector and excitation light. In conclusion, both the SU-8 and the PDMS processes need to be improved to yield useable parts for the final cell retention system.**

***Index Terms*—fluorescence, micro-channels, molding, PDMS, SU-8**

I. Introduction

For biomedical purposes, micro-channels can be used to conduct experiments on single cells, as extensively shown in the work of Dr. Paul Li at Simon Fraser University [1] [2]. The frequently used material for micro-fluidic channels is poly-dimethylsiloxane (PDMS) [3]-[5], which is a flexible silicone elastomer that can be used for molding. Once cured, PDMS is stable and can be used in wide range of temperature environments (-45 to 200°C) [6] [7]. Unlike a process that creates micro-channels in silicon substrate [8], PDMS can be used to create multiple channels using just one mold master [9] [10]. This feature makes PDMS ideal for biomedical applications because creating disposable channels to avoid contamination becomes viable. The main goal of this project is to fabricate a PDMS structure to a hold macro-scale photo-detector. In previous work done on PDMS channels, cell reactions or fluorescence are detected by a photomultiplier tube (PMT) [2], which is expensive and very large compared to the size of the channel. If a small photo-detector can be mounted directly underneath the cell channel to detect cell fluorescence, the cost and size of the cell detection method can be greatly reduced.

This paper describes the design and fabrication of micro-channels to be used for single cell retention and fluorescence detection. The channels were first patterned onto SU-8 photo polymer, and the actual channels were molded using PDMS. One important feature for the design is the photo-detector mount used to hold a macro-scale photo-detector underneath the micro-channel. In order to fabricate the mount, the SU-8 structures must be as thick as possible to conform to the relatively large photo-detector. Another important objective of this paper involves the molding of the structures with PDMS to create the actual micro-channel and the photo-detector mount. Testing of the channels for cell flow and detection of cell fluorescence remains to be explored.

II. Design Overview

Fig. 1 shows the overview for the cell fluorescence detector. The detector consists of a channel layer, a detector mount, and the photo-detector. The channel layer and the detector mount are made from PDMS by molding them onto the SU-8 masters. The channel layer contains channels and a retention chamber for cells to be held at a constant position. Fig. 2 shows the design and dimensions of the channel structure.

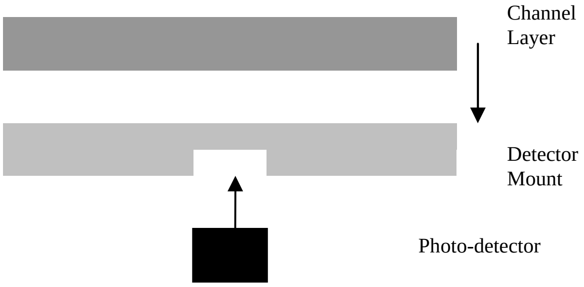


Fig. 1: Cell Fluorescence Detector Overview

Manuscript received June 29, 2006.

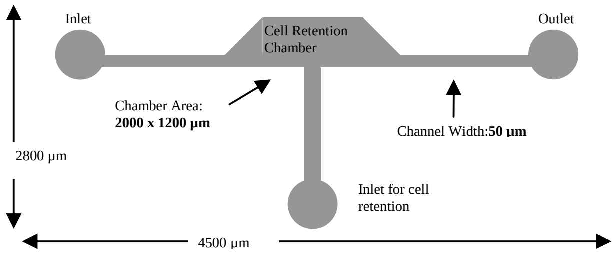


Fig. 2: Cell Retention Channel

As shown in Fig. 2, the channel structure consists of 3 channels. Two are used for the entrance and exit of cells (left and right) and another for holding the cells in the chamber groove by increasing the water pressure (center).

.

III. FABRICATION

The following sections summarize the process steps to create the SU-8 structures for the channels and detector chamber, and the PDMS molding. Before starting the fabrication, <111> silicon wafers were cleaned with acetone and isopropyl-alcohol (IPA). After the wafers were cleaned, they were placed into a 100°C convection oven for 10 minutes to dry off all the moisture.

A. Su-8 channel mold master fabrication

To fabricate 50 µm thick SU-8 channels, SU-8 50 (Microchem) was spun at 2000 rpm for 30 seconds [11]. Then the SU-8 coated wafer was set on the hot plate (Torrey Pines Scientific Inc., HP30) to remove excess SU-8 solvent. After baking, the wafer was left at room temperature for 45 minutes to set the SU-8. Next, the wafer and the transparency mask with channel patterns were set onto the Quintell Aligner for exposure. Then the wafer was returned to the hot plate again for a post-exposure bake to prepare the wafer for development. Finally, the SU-8 coated wafer was developed using a SU-8 Developer (Microchem). With constant agitation, the wafer was developed and then rinsed with IPA. Following the development, the wafer was dried with a nitrogen blower, thus completing the SU-8 channel fabrication (Fig. 3).

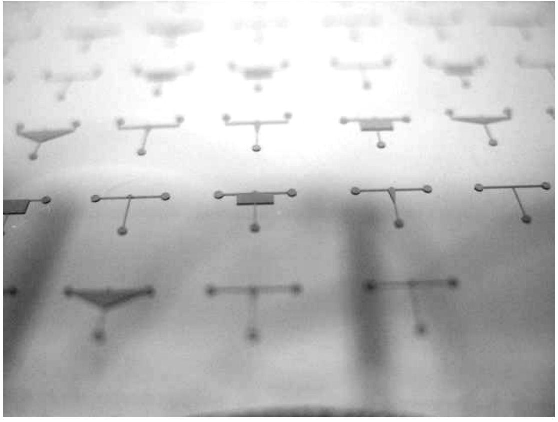


Fig. 3: SU-8 Channels

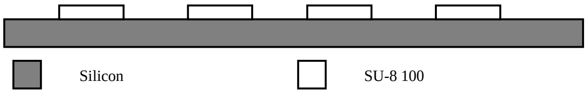


Fig. 4: 1st Layer of SU-8 for the Detector Mount

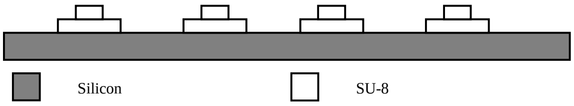


Fig. 5: 2nd Layer of SU-8 for the Detector Mount

B. Fabrication of the Thick SU-8 Mold Master

The process steps for fabricating the thick SU-8 detector mount mold master are similar to that of the channels. However, to create thicker structures, SU-8 100 was spun at 1000 rpm for 45 seconds. After exposure and development, the first layer of the detector was created as shown in Fig. 4. Processing the second layer of SU-8 follows similar process steps as the first layer. The only difference is that the SU-8 100 was spun onto the wafer twice at 1000 rpm for 30 seconds (Fig. 5).

C. PDMS Molding

For molding with PDMS, the *Sylgard 184 Silicone Elastomer Kit*'s (Dow Corning) base and curing agent were thoroughly mixed in 10 to 1 ratio by weight. After a thorough mixing, the mixture was set in a vacuum system to extract the air bubbles. Next, the PDMS mixture was poured onto the SU-8 master. After pouring, the mixture was again set at rest for 15 minutes to release the bubbles which formed during the pour. As shown in Fig. 6, the PDMS coated wafer was covered with a transparency film and placed on the hot plate with an aluminum plate on top to secure the film. The curing time and temperature were 2 hours and 100 °C. Finally, the PDMS film was peeled from the substrate when the curing was complete.

IV. RESULTS AND DISCUSSION

Both the SU-8 structures and the molding with PDMS worked successfully. Parameters such as bake and exposure times affected the result. Sections A and B discuss the results and parameters affecting the SU-8 structures. Then Section C discusses the PDMS molding results.

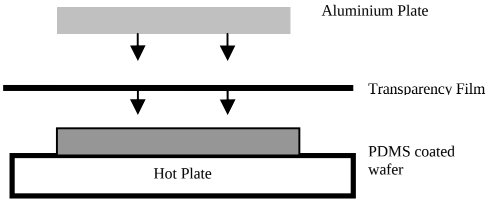


Fig. 6: PDMS Curing Setup [3]

A. SU-8 Structures: Channels

As seen in Fig. 7, the 50 µm thick SU-8 channels are present with most of the structures intact as designed. The height was measured using the profilometer, showing it to be about 53 µm. However, structures like the one circled in Fig. 7 did not turn out successfully. The structure has a streak in middle of the chamber, which means no SU-8 adhered on that area. This problem could have been caused by adhesion problem due to an unclean wafer or by SU-8 beads causing streaks during the spin.

B. SU-8 Structures: Detector Mount

The detector mount was much more difficult to fabricate than the thin channels. The one major problem encountered during fabrication was the thickness of the structures. The SU-8 structures were so thick (~450 µm), it was not possible to align the second layer with the aligner, because the SU-8 coated wafers stick to the mask. Thus, the alignment was done by eye without using the micro-scope built into the aligner. The resulting structures are shown in Fig. 8. The dimensions of the detector mount are 4.5 x 2 mm for the long first layer and 3 x 2 mm for the short second layer. Fig. 8 shows the most aligned structures which were off by 33 µm in the horizontal direction.

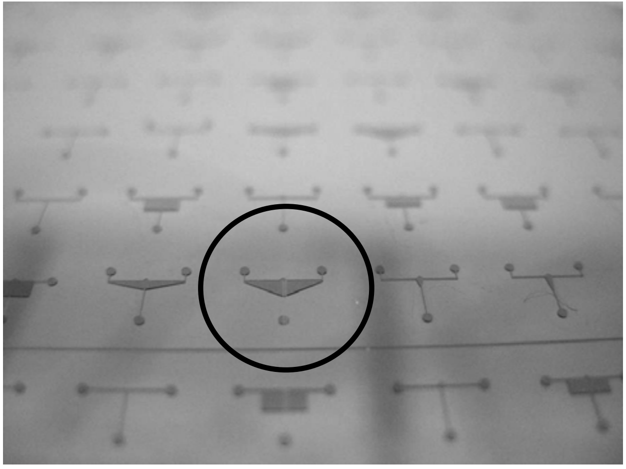


Fig. 7: SU-8 Channel Structures

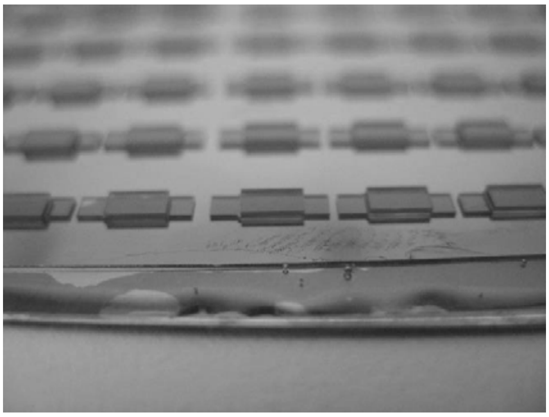


Fig. 8: Aligned SU-8 Structues

C. PDMS Molding

Obtaining clean results using PDMS molding was difficult, so more work needs to be done to characterize the PDMS process. Difficulties encountered were air bubbles trapped in the PDMS and determination of proper curing time. Fig. 9 shows a close-up view of the PDMS mold of the detector mount. Fig. 9 shows the detector mount is clearly imprinted on the PDMS. However, as seen from Fig. 9, air bubbles are present in the PDMS mold which is caused by not placing the PDMS mixture in a vacuum condition for long enough time. Additionally, some of the detector mount structures resulted with holes in the PDMS film, because they were not entirely covered with PDMS. This flaw was likely caused by the uneven thickness of the SU-8 structures, indicating the hot plate must have not been properly leveled during the pre and post exposure bakes.

Fig. 10 and Fig. 11 are photographs of the channel structures in PDMS. Fig. 10 is the center structure featuring the cell retention chamber, and Fig. 11 is the inlet/outlet channel.

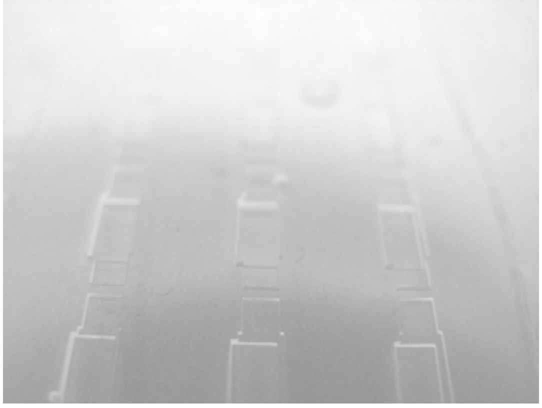


Fig. 9: PDMS of Detector Mount (close-up)

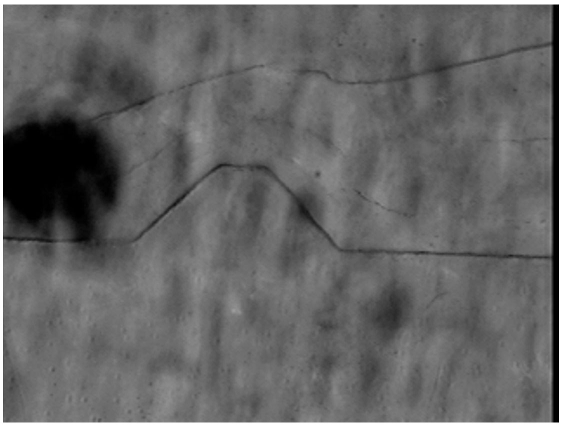


Fig. 10: Channels (Center)

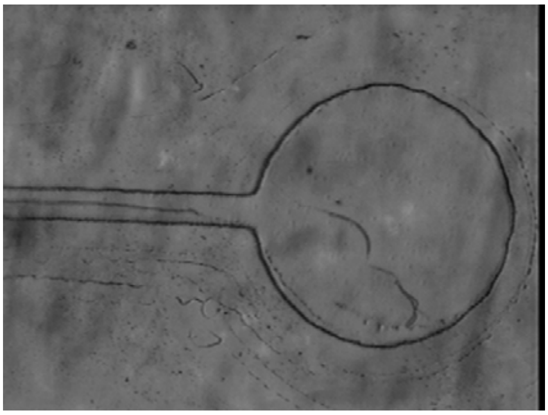


Fig. 11: Channel (Side)

The black spots in both Fig. 10 and Fig. 11 are air bubbles caused by similar problems as the detector mount. The possibility of bubbles affecting the performance of the channels remains to be explored, but the above images clearly indicate that the bubbles will impede the view of the channels. Unfortunately, due to the soft nature of PDMS, the actual depth of the structures could not be measured because the profilometer needle will be trapped in PDMS. Assuming the PDMS did conform to the SU-8 master, the channel depth should be about 53 µm.

V. Conclusion

In this paper, the process of creating SU-8 structures for PDMS mold was explored. Both channel and detector mount SU-8 structures had acceptable results. SU-8 channels were fabricated with a 53 µm thickness. However, during fabrication, many channel structures either cracked or washed away during development. Determining the correct bake and exposure times must be explored further. Unlike the channels, SU-8 detector mount structures were fabricated with almost all of the structures in tact. The problems with the thick SU-8 involved uneven thickness of the structures and misalignment of the second layer. The uneven thickness can be improved by leveling the hot plate during the pre and post exposure bakes. Alignment can be improved if chrome masks are used so that the mask does not bend when in contact with the SU-8.

PDMS molding resulted in a successful imprint of the channels and the detector mount structures. However, further process improvement is required. In particular, improving the process for eliminating the bubbles in the PDMS structures is crucial because they interfere with the optical characteristics of PDMS.

References

[1] Peng, X. and Li, P., *A three-dimensional flow control concept for single-cell experiments on a microchip. 1. Cell selection, cell retention, cell culture, cell balancing, and cell scannning.* Anal. Chem, 2004. **76**: p. 5273-5281.

[2] Li, X. and Li, P., *Microfluidic selection and retention of a single cardiac myocyte, on-chip dye loading, cell contraction by chemical stimulation, and quantitative fluorescent analysis of intracellular calcium.* Anal. Chem, 2005. **77**: p. 4315-4322.

[3] Jo, B. and et al., *Three-dimensional micro-channel fabrication in polydimethylsiloxane (pdms) elastomer.* J. MEMS, 2000. **9**(1): p. 76-81.

[4] Anderson, J. and et al., *Fabrication of topologically complex three-dimensional microfluidic systems in pdms by rapid prototyping.* Anal. Chem, 2000. **72**: p. 3158-3164.

[5] Geschke, O. and et al., *Microsystem engineering of lab-on-chip devices*. 2004: Wiley-VCH.

[6] DowCorning, *Information about dow corning brand silicone encapsulants*. 2005.

[7] Fujii, T., Sando, Y., Higashino, K., and Fujii, Y., *A plug and play microfluidic device.* Lab on Chip, 2003. **3**: p. 193-197.

[8] Boer, M. and et al., *Micromachining of buried micro channels in silicon.* J. MEMS, 2000. **9**(1): p. 94-103.

[9] Duffy, D. and et al., *Rapid prototyping of microfluidic systems in poly(dimethylsiloxane).* Anal. Chem, 1998. **70**: p. 4974-4984.

[10] Tan, A., Rodgers, K., Murrihy, J., Mathuna, C., and Glennon, J., *Rapid fabrication of microfluidic devices in poly(dimethylsiloxane) by photocopying.* Lab on Chip, 2001. **1**: p. 7-9.

[11] Microchem, *Nano su-8 negative tone photoressit fromulations 50-100*. 2002.

Takaya Ueda is an MASc stuudent at Simon Fraser University (SFU) in Burnaby, British Columbia Canada. He obtained BASc in Electronics Engineering in 2004 at SFU. His research interests include fabrication of micro-fluidic structures and mechanical devices. Currently, he is researching methods of transferring metal onto PDMS to fabricate conductive polymers.

An Optical Imaging Technique Using Deep Illumination in the Angular Domain

Fartash Vasefi

Abstract— Deep illumination in the angular domain imaging (ADI) employs micromachined collimators detecting photons emitted from Deep illumination provided by backscattered source light subsurface of the turbid medium within small acceptance angle. These micro-scale angular filters are composed of arrays of collimator filters which are fabricated by silicon bulk micromachining. Two dimensional phantom test objects were observed in high scattering media up to 3 mm deep in the medium at effective SR (scattered to ballistic ratios) from 1:1 to greater than 3E12:1. Results from carbon coated collimators which is fabricated by a sputtering system and roughed-surface collimator to decrease internal reflectivity have been shown.

Index Terms— Optical tomography, angular domain imaging, back-scattering, silicon micromachined collimating array (SMCA), carbon coating

I. Introduction

Optical reflection, absorption, scattering, and fluorescence in living tissues and blood can reveal significant information about the structure and constitution of the living bodies. Within the visible and infrared spectra, tissues and bioliquids are low absorption but highly scattering media [1]. Scattering describes the spectral and angular characteristics of light interacting with living material, as well as its penetration depth; thus, the optical properties of tissues and blood are heavily influenced by changes in scattering properties. Optical imaging is the product of these interactions on light, which is input from various sources, to generate a composite effect that can be recorded and viewed as the "image."

Short wavelength energy emitted from sources carries greater energy than long wavelength emissions. If the wavelength is very short, the energy greatly increases, which can damage the tissue because electrons are dislodged (ionization) from atoms and molecules [1]. Far ultraviolet, X-rays, and gamma rays are all ionizing radiation (short high-energy wavelengths), while visible light, microwaves, and radio waves (longer wavelengths) are non-ionizing radiation.

The first X-ray photograph was created by Roentgen in 1895, and subsequent development provided a breakthrough in medical imaging beneath skin level (subcutaneous).

This technique yielded a wide range of possibilities in terms of applications and consequently gave rise to X-ray tomography: A method which relies on a series of photographs created with the imaging assistance of computers through the tissue, namely, Computerized Axial Tomography (CAT). In spite of the wide range of applications, this method has some limitations, the complications of which disrupt and destroy the chemical structure of living tissue causing tissue damage and increasing cancer risks proportionate to the cumulative dose of radiation applied. Thus, they are classified as ionizing image techniques (X-ray and fluoroscopy). The attractive alternative of non-ionizing radiation imaging has consequently been of high interest, resulting in such techniques as ultrasound (acoustic wavelengths), magnetic resonance imaging (MRI - radio wavelengths), and optical tomography (visible and infrared light).

Another important distinction between imaging techniques relates to still images (photographs) versus real time images (cinema). Magnetic Resonance Imaging (MRI) is a non-ionizing method that requires intensive computing to create off-line tomographic 'still' slices. Real time MRI is not yet possible. Continuous non-ionizing imaging is possible with the ultrasound technique but the obtained resolution (at acoustic wavelengths) is rather poor and the image depends on gross tissue differences in acoustic impedance rather than color and structure closer to molecular dimensions that can be measured at light wavelengths. The problem heretofore for light imaging is that scattering of light through tissue makes optical imaging of structures "fuzzy" and difficult to separate from the dominate scattered light.

There are many situations in which the detection of objects in a highly scattering turbid medium using backscattered light is highly desirable. For example, backscattered light may be utilized to detect a tumor embedded within tissue, such as breast tissue. Various Types of microscopes use backscattered light to display the surface image of a medium with high resolution. A confocal arrangement can extend the image to less than 200 μm below the surface. The conventional Optical Coherent Tomography (OCT) technique, which uses backscattered light, can only image the internal structure of an eye and tissue down to about 600 μm below the skin surface. No clear image of the medium structure in a deeper depth, however, can be formed using the direct backscattered light signals. This is due to multiple-light scattering within a medium, which contributes to noise, loss of coherence, and reduces the intensity of light directly

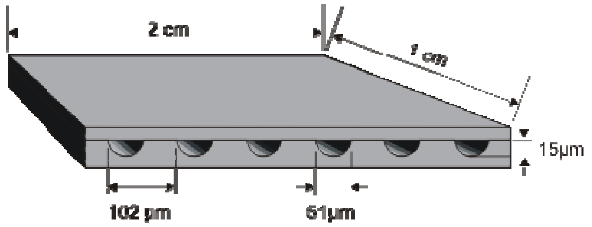


Figure 1: Silicon Micromachined Collimator Arrays (SMCA)

backscattered from the hidden object. The conventional diffusion optical tomography method has several disadvantages. For example, the diffusion method only uses diffusive photons which have suffered many scattering events. Therefore, the signals received by detectors are less sensitive to changes in the structure of the turbid medium, which makes it difficult to obtain high-resolution image reconstruction.

To perfect an optical image of living tissue or blood beneath the skin, the limiting effect of optical scattering must be removed. This paper offers a new technique for doing so and provides high-resolution images of the micro scale phantom in highly scattering medium.

II. Materials and Methods

A. Fabrication of the Silicon Micromachined Collimating Array

To obtain fine object resolution and detection, a collimating array must have relatively small holes with small spacing between them. In the design, as shown in Figure 1, we used 51 µm diameter holes with 102 µm spacing to produce a parallel array of collimators with a predicted object resolution in the 100-200 µm range.

To observe an image, this collimator was aligned to a CMOS imaging array detector in such a way that the hole spacing of 102 µm matched the spacing requirement to be integer numbers of the CMOS detector. Each block of grooves covered 20 × 10 mm squares, fabricated on a 100 mm silicon wafer. When combined with an encasement, these grooves became the Silicon Micromachined Collimator Array (SMCA), which had a very high aspect ratio (200:1) resulting in a small angle of acceptance (0.29^o). The large length of the array (10 mm) combined with the small hole size suggested that silicon surface micromachining could best generate the structure. For these initial experiments, only a linear collimating array was created. The basic steps of the collimator microfabrication are shown in detail in [20]. These collimators started with a furnace <100> silicon wafer oxidation (0.5 µm thick), which is then photolithographicly patterned and etched with HF to create the masking layer of the collimator grooves (see Figure 2). The silicon was etched in HF, HNO_3, CH_3COOH [18], using the oxide openings to produce a groove width of 51 µm after isotropic etching, with a 15 µm undercut on each channel side (see Figure 2). The oxide was then stripped, leaving the grooved structure.

The microchannel array was coated with a thin layer of carbon (~200nm and 600nm) in the sputtering deposition system in order to get better attenuation of the light reflected from the channel walls. The carbon deposited on the bottom and the side wall are approximately 75% and 50% of the thickness of carbon at the top edge, respectively which gives us good deposition uniformity. The wafer was cleaved into 20 x 10 mm sections along the cutting grooves between each section. For the first setup, a carbon coated silicon wafer was bonded to the etched wafer top, creating tubes in the collimator with half-circular cross-sections (as can be seen in the SEM picture in Figure 1).

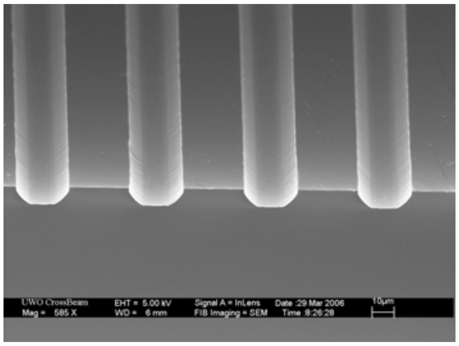


Figure 2: the etched side of the Silicon Micromachined Collimating

B. Backscattering angular domain imaging (ADI) operation

Generally, light entering the tissue undergoes both absorption and scattering. In its simplest form, the laser beam intensity follows an exponentially decaying Beer-Lambert law along its path through the media:

$$I_{out} = I_{in} \exp[-(\mu_a + \mu_s)d] = I_{in} \exp[-\mu_{eff}\, d]\,, \qquad (1)$$

where, for typical mammography values, the absorption coefficient is $\mu_a = 0.7$ cm^{-1}, the scattering coefficient $\mu_s = 130$ cm^{-1} and the depth is d = 5 cm [25]. Unscattered light becomes “Ballistic Photons.” For this example, the ratio of scattered to ballistic photons is 6.7×10^{283}:1. Fortunately, most of the light is not scattered uniformly in all directions, but instead tends to divert mostly toward the laser beam’s direction of motion. The measurements are characterized using the Scattering Ratio, SR (ratio of scattered to unscattered light):

$$SR = \frac{I_0}{I_{bq}} - 1\,, \qquad (2)$$

where I_0 is the initial light intensity, and I_{bq} is the combined ballistic and quasi-ballistic light intensity.

This forward scattering creates an effective absorption anisotropic coefficient [25], $\mu_{eff} = 4.2$ cm^{-1}, for the so called “quasiballistic or snake photons” (ones that are mostly scattered forward). Since these quasi-ballistic photons also contain desired optical information, their scattering ratio of about $10^{11}:1$ represents a significant target for detection in this research.

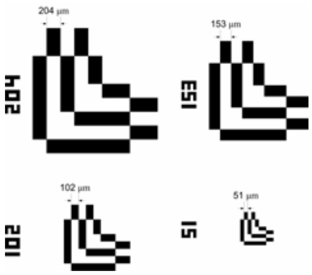


Figure 3: Two-Dimensional phantom test structures

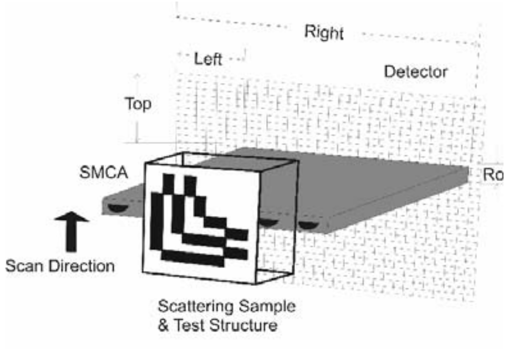


Figure 4: Scan of sample medium across SMCA

If a photon emerges from a scattering medium with an angle greater than the acceptance angle, then it either fails to enter the collimator or becomes absorbed within it. Hence, only the ballistic and quasiballistic photons, undergoing no scattering or small scattering, will be detected. When very small acceptance angles are used, our experiments confirm [25] greatly enhanced detection, which is competitive with existing methods but with *a much simpler detection system*. The method is called Angular Domain Imaging, as introduced in the previous section and [25].

To measure the resolution of the system, aluminum thin film on glass substrate phantoms, shown in Figure 3, were fabricated using photolithography consisting of parallel lines and spaces of 51 µm, 102 µm, 153 µm, and 204 µm line widths. Test phantoms used for this research are placed in a container holding a liquid scattering medium. The test objects are located at various positions within the container.

Previous Angular Domain Imaging experiments [25] have focused on using a laser source aligned to the collimating array for the maximum depth penetration with transillumination in the turbid medium. However, the same technique will work if the light source is not aligned to the collimator, though at much lower depth, if the light source is not aligned to the collimator. This approach allows detection of objects at moderately shallow depths by illumination from the front (i.e., the same side as the collimator [25]). The experiments were undertaken with angle illumination of the scattering medium, as shown in Figure 5. The resulting scattered light behind the samples then acts as the source of light for the collimator microchannel. To create an easily controllable scattering medium, a solution was created by mixing specific amounts of 2% partially skimmed milk with de-ionized (DI) water to achieve a desired scattering level. Milk was chosen because it has properties similar to tissue and exhibits good scattering characteristics as well as low absorption coefficient.

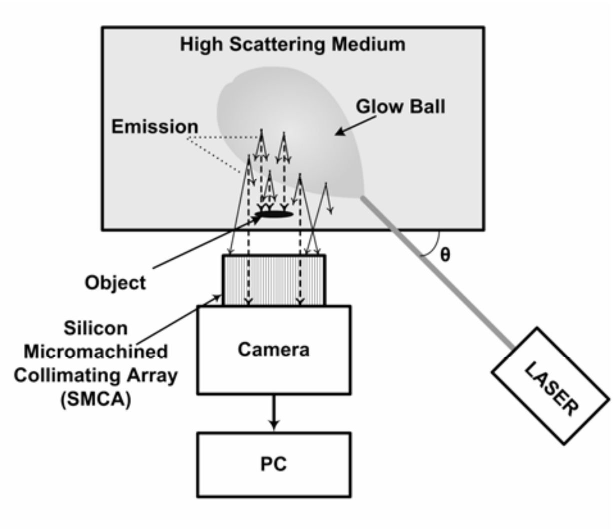


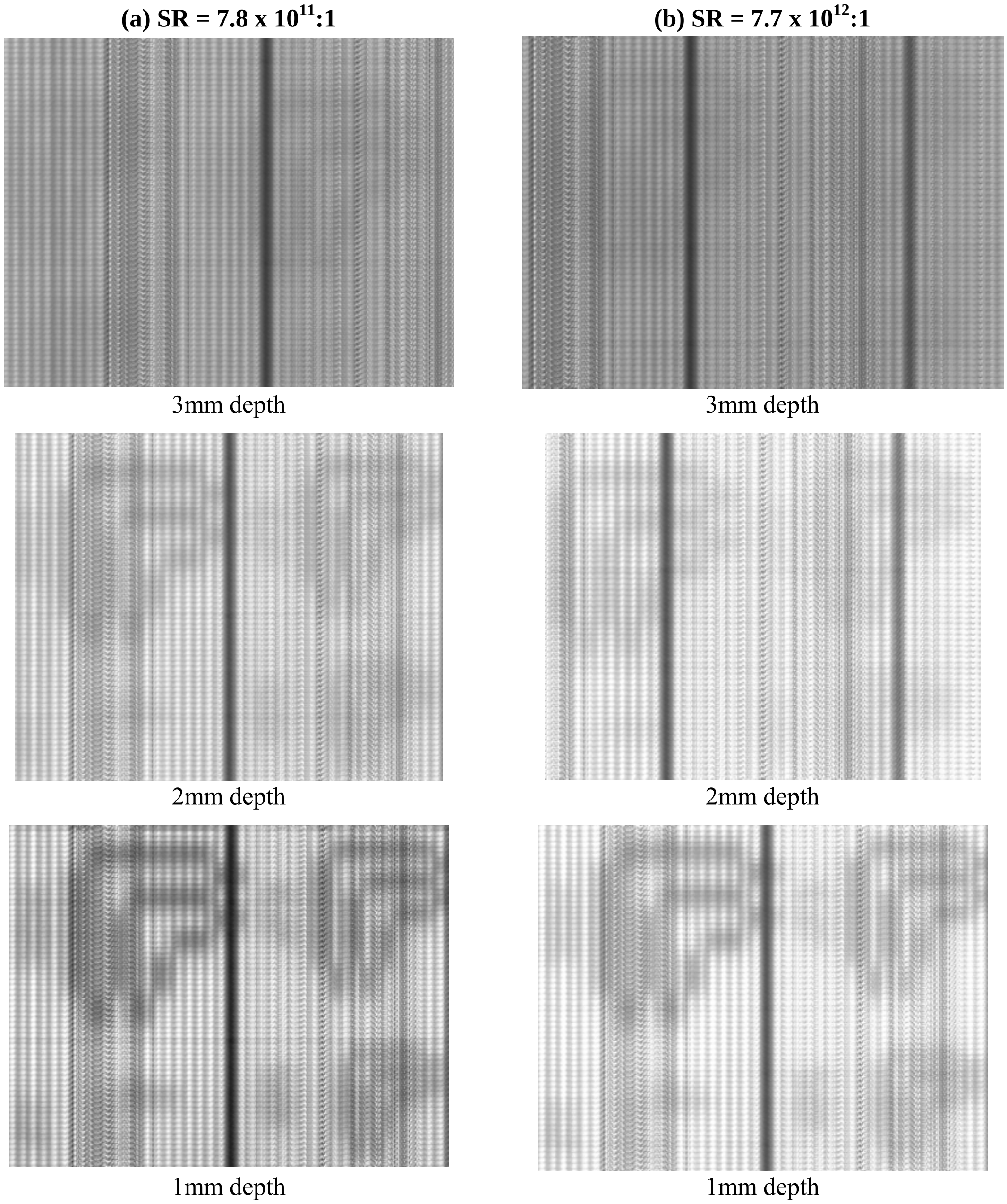
Figure 5: Angled illumination of a test setup

The principle here is that the scattered light, which starts out aligned to the collimator, will have the highest probability of being detected, provided it is not scattered by the medium or blocked by the test objects. Thus, the test structures can be seen as high absorption areas over this uniform illumination. The depth of the test structures should be small enough in order for the scattering to be intangible after the test structure.

The SMCA is placed in front of a CMOS imaging array with a pixel size (5.2x5.2 microns) smaller than the collimating hole diameter of 51 microns. The photons reaching the pixels are filtered by the SMCA so that only those with a trajectory that lies within the collimating hole's acceptance angle reach the CMOS imager. Scanning the collimator along the sample is used to create a two-dimensional image from the linear array (see Figure 4).

III. Results

The test structure slide is placed at distances of 3 mm, 2mm and 1mm back from the collimator side of the scattering medium. As can be seen in Figure 6, the setup required to illuminate the scattering medium from the same side of the image detector is provided. The laser beam has a beam angle of 54° from side of the container. We want to explore how much background illumination comes from light scattering along the front of the tissue. For this reason, the light passes through the scattering medium from the same surface as it is being viewed. Also, this setup will affect the way the light forms the sphere of illumination behind the sample.



(a) SR = 7.8 x 10^{11}:1

(b) SR = 7.7 x 10^{12}:1

3mm depth

3mm depth

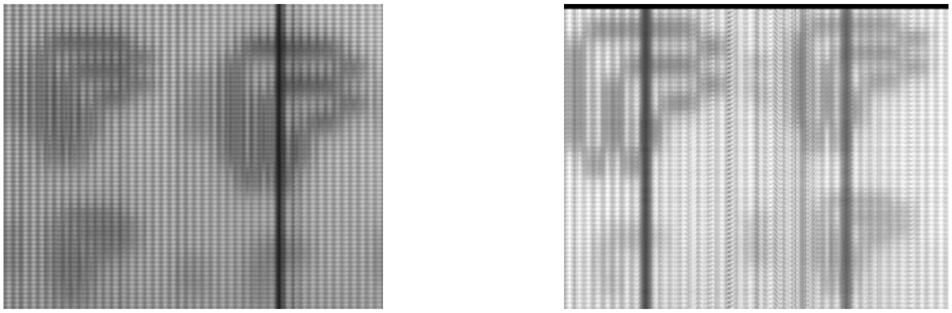
2mm depth

2mm depth

1mm depth

1mm depth

Figure 6: Comparison of the images captured by silicon *SMCA* in the various depth and two different scattering Ratio(SR).



a) Carbon coated SMCA, SR: 3.2×10^{10}:1

b) Silicon SMCA, SR: 1.6×10^{10}:1

Figure 7: Comparison of the images captured by carbon coated *SMCA*

The scattering medium is illuminated from the angle by a 2.5mm diameter laser beam from our argon ion laser (wavelength of λ=488nm). The L shape sample has been placed approximately 3mm deep in the turbid medium. The static images were obtained using different depths and SR values. The depths were one, two and three millimeters, as in the columns of Figure 6. The two scattering ratios shown are in the range of SR=10^{10}:1 to SR=10^{12}:1. For the non-coated silicon tunnels, we can faintly resolve the 204µm-wide test structure lines at the 3mm depth. At the 1mm depth, contrast and definition improve, and we can detect the presence of the small 51µm test structure.

Measurements show that the surrounding light is 68% of the background level through the collimator hole. Hence, removing the reflected light would significantly reduce background scattering, and thus, significantly enhance the detectability of imaging. Thus, adding an absorption coating to the SMCA tunnel sidewalls was expected to reduce the background level and was demonstrated to be true in our experiments. The carbon layer on silicon obtained by sputtering deposition has been compared to a non-coated silicon surface using a simple set of reflection experiments. The first setup with a small angle -- 15^o-- the reflection has been decreased by the 200nm carbon layers from 40% to 20%. The reflection increases more with the 600nm carbon film (up to 13%.) Although in a shallow beam, the reflection increases for all cases considerably. Figure 7 shows the scan results from both the silicon and carbon coated tunnels and indicate no significant difference in contrast and feature definition between the two.

IV. Conclusion

In this paper, the idea of backscattering angular domain imaging was proposed, and the related experiments were presented. This novel method of deep illumination can enable new application areas for optical imaging, such as subcutaneous tissue mapping. To implement this method, the samples were illuminated with a 54^o laser beam. Two dimensional phantom test structures have been successfully detected up to three millimeters deep in the highly scattering medium with a scattering ratio of over 10^{12}:1. In future work, two improvements will be explored to improve the image quality. First, the half circle collimators will be replaced by full circle channels using buried channel technology. This is a method for the fabrication of micro channels under the bulk silicon [28]. Second, multi-wavelength laser and image processing techniques will be used and explored.

The backscatter ADI technique was introduced in this project, and experimental results show that it is a viable technique for imaging objects at depths of 1mm to 3mm in scattering media equivalent to up to SR = 10^{15}:1 over 5cm. Scattering at these levels is similar to some biological tissues; thus, these results are promising for applying the backscatter ADI technique to image tissues at shallow depths. Surface enhancements to the SMCA collimator tunnels were explored, including depositing a carbon film and roughening the silicon surface with an NH_4OH solution at 80ºC. Reflectivity tests on these surfaces and others indicate that the carbon film was not of a good quality for attenuation, and may be suffering in purity (e.g. hydrocarbon film). Tests indicate that roughened topology surfaces, such as the NH_4OH roughened silicon, show good attenuation of reflected light over a wide range of angles. Backscatter ADI results comparing both the silicon and carbon coated SMCA devices confirm that the latter does not provide significantly better rejection of noise during imaging over the normal silicon SMCA. Analysis of how SR values increase for shallow depths (i.e. 1mm to 4mm) as the effective 5cm SR of the entire scattering medium is increased indicate that SR values remain at fairly low levels when dealing with such shallow thicknesses.

References

[1] F. Vasefi, B. Kaminska, G. H. Chapman, P. K.Y. Chan, “ Angular Domain Imaging for Tissue Mapping”, 2nd ASM - IEEE EMBS Conference on Bio, Micro and Nano- systems, January 15-18, 2006, San Francisco, USA.

[2] A.H. Hielscher, R.E. Alcouffe, "OSA Trends in Optics and Photonics on Advances in Optical Imaging and Photon Migration. Vol.2. From the Topical Meeting," in OSA Trends in Optics and Photonics on Advances in Optical Imaging and Photon Migration. Vol.2. From the Topical Meeting, 18-20 March 1996, 1996, pp. xii+406.

[3] Britton Chance, Univ. of Pennsylvania School of Medicine and Robert R. Alfano, CUNY/City College, "Optical Tomography and Spectroscopy of Tissue: Theory, Instrumentation, Model, and Human Studies II," Proceedings of the SPIE - the International Society for Optical Engineering, vol. 2979, /. 1997.

[4] J.R. Mourant, et. all., "Mechanisms of light scattering from biological cells relevant to noninvasive optical-tissue diagnostics," Appl.Opt., vol. 37, pp. 3586-93, 06/01. 1998.

[5] B. Beauvoit, T. Kitai and B. Chance, "Contribution of the mitochondrial compartment to the optical properties of the rat liver: a theoretical and practical approach," Biophys.J., vol. 67, pp. 2501-10, 12/. 1994.

[6] T. Vodinh, "In-vivo Cancer-Diagnosis of the Esophagus Using Differential Normalized Fluorescence (DNF) Indexes," Lasers Surg.Med., vol. 16, pp. 41-47, 1995

[7] Mourant JR, Bigio IJ, Boyer J, Conn RL, Johnson T and Shimada T, "Spectroscopic diagnosis of bladder cancer with elastic light scattering," Lasers in Surgery and Medicine., vol. 17, pp. 350, 01/01. 1995.

[8] J.C. Hebden and D.T. Delpy, "Enhanced time-resolved imaging with a diffusion model of photon transport," Opt.Lett., vol. 19, pp. 311-13, 03/01. 1994.

[9] B.B. Das, K.M. Yoo and R.R. Alfano, "Ultrafast time-gated imaging in thick tissues: a step toward optical mammography," Opt.Lett., vol. 18, pp. 1092-4, 07/01. 1993.

[10] A.F. Fercher, W. Drexler, C.K. Hitzenberger and T. Lasser, "Optical coherence tomography-principles and applications," Reports on Progress in Physics, vol. 66, pp. 239-303, 02/. 2003.

[11] Gang Yao and L.V. Wang, "Monte Carlo simulation of an optical coherence tomography signal in homogeneous turbid media," Phys.Med.Biol., vol. 44, pp. 2307-20, 1999.

[12] Gang Yao and L.V. Wang, "Two-dimensional depth-resolved Mueller matrix characterization of biological tissue by optical coherence tomography," Opt.Lett., vol. 24, pp. 537-9, 04/15. 1999.

[13] M.A. O'Leary, D.A. Boas, B. Chance and A.G. Yodh, "Experimental images of heterogeneous turbid media by frequency-domain diffusing-photon tomography," Opt.Lett., vol. 20, pp. 426-8, 03/01. 1995.

[14] S.B. Colak, D.G. Papaioannou, G.W. 't Hooft, M.B. van der Mark, H. Schomberg, J.C.J. Paasschens, J.B.M. Melissen and N.A.A.J. van Asten, "Tomographic image reconstruction from optical projections in light-diffusing media," Appl.Opt., vol. 36, pp. 180-213, 01/01. 1997.

[15] Prasad, Paras N, “Introduction to Biophotonics”, Hoboken, NJ: Wiley-Interscience, 2003.

[16] G.H. Chapman, M.S. Tank, G. Chou and M. Trinh "Optical imaging of objects within highly scattering mediums using Silicon Micromachined Collimating Arrays", Proceedings SPIE Photonics West, BIOS Optical Fibers and Sensors in Biomedical Applications II (BO10), v4616, pp. 187-198, San Jose, CA, Jan. 2002.

[17] G. H. Chapman, M. Trinh, D. Lee, N. Pfeiffer, and G. Chu, "Angular Domain Optical Imaging of Structures Within Highly Scattering Material Using Silicon Micromachined Collimating Arrays", Optical Tomography and Spectroscopy of Tissue V, B. Chance et al (eds), Proceedings of SPIE Vol. 4955, pp. 462-473, San Josa, CA, Jan. 2003.

[18] G.H. Chapman, M. Trinh, N. Pfeiffer, G. Chu, and D. Lee, "Angular Domain Imaging of Objects Within Highly Scattering Media Using Silicon Micromachined Collimating Arrays", IEEE J. Special Topics on Quantum Electronics, V9, No. 2, pp. 257-266, 2003.

[19] N. Pfeiffer, B. Wai, and G.H. Chapman," Angular Domain Imaging of phantom objects within highly scattering mediums", Proc. SPIE Photonics West: Laser Interaction with Tissue and Cells XV, v 5319, San Jose, CA pg 135-145, Jan. 2004

[20] M. S. Tank, "Development of a Silicon Micromachined Collimator Array to Detect Objects within Highly Scattering Mediums", MSc Thesis, School of Engineering Science, Simon Fraser University, Burnaby, BC Canada, 2001

[21] Teresa Wong, "An alternative approach to multi-chip module interconnections: Laser-welding micro cantilevers", 27, B.A.Sc Thesis, School of Engineering Sciences, Simon Fraser University, Burnaby, BC Canada, 1995

[22] M.S. Tank and G.H. Chapman, "Micromachined Silicon Collimating Detector Array to View Objects in Highly Scattering Medium", Can Jour Elec. & Comp. Eng, .25, no. 1, 13-18, Jan. 2000,

[23] M. Trinh, "Pushing the limits of Optical Tomography using a Silicon Micromachined Collimator Array" B.A.Sc Thesis, School of Engineering Sciences, Simon Fraser University, Burnaby, BC Canada, 2001

[24] G.H. Chapman, P.K.Y. Chan, J. Dudas, J. Rao and N. Pfeiffer, "Angular Domain Image Detectability with Changing Turbid Medium Scattering Coefficients", Proc. SPIE Photonics West: Laser Interaction with Tissue and Cells XVI, v, 5695, pg 160-171, San Jose, CA, Jan. 2005

[25] G. H. Chapman, Josna Rao, Ted Liu, Paulman K.Y. Chan, Fartash Vasefi, Bozena Kaminska, and Nick Pfeiffer, "Enhanced Angular Domain Imaging in Turbid Medium using Gaussian Line Illumination", Proceedings of SPIE Volume: 6084, Optical Interactions with Tissue and Cells XVII, BIOS06, Jan 2006

[26] J Beuthan et al., "IR-Diaphanoscopy in Medicine", Medical Optical Tomography: Functional Imaging and Monitoring, G. Muler et al. (eds), SPIE IS11, pp. 263-282, 1993.

[27] S. Jacques, "Introduction to Biomedical Optics", Oregon Graduate Institute, http://omlc.ogi.edu/classroom/ece532/ (Feb 2001)

[28] Meint J. de Boer, et al., R. "Micromachining of Buried Micro Channels in Silicon" Journal of MEMS, Vol. 9. No. 1, March 2000.

Fabrication of Cell Platforms in SU-8

Stephanie Westwood
Microinstrumentation Laboratory
Simon Fraser University, 8888 University Dr., Burnaby, BC, Canada, V5CS 1A6
swestwoo@sfu.ca

Abstract—**Cell platforms are an integral part of shrinking the traditional biology laboratory. The premise behind cell platforms is the ability to compartmentalize cell analysis into smaller pieces that can be connected and disconnected from the larger system at will. In this paper, we present cell platforms composed of microchannels for examining the membrane elasticity of red blood cells (RBCs) and endothelial cells (ECs). We fabricated the first channels on silicon wafers using SU-8 thick photoresist. Channel widths were chosen to range from two to ten microns and from 50 to 210 microns for RBCs and ECs, respectively. Lengths ranged from four to 100 microns and from 100 to 2100 microns; these dimensions reflect two to ten times their corresponding channel widths. Future work for this project includes assessing the system by growing cells in the channels, enclosing the channels using acrylic or glass, testing the channel pressure using a syringe pump, and integrating the channels into an LOC system using interlocking structures.**

Index Terms— **Cell platform, Endothelial Cells, Microfluidics, SU-8**

I. Introduction

Lab-on-a-chip (LOC) systems represent a sizeable advance in the ability for clinicians and researchers to quickly and cost effectively investigate issues like disease detection, DNA testing, and blood analysis. The shrinking of the conventional lab into an LOC system enables smaller required fluid samples, better process control, and substantial parallelization, which allows for high throughput analysis. We define the integration of several laboratory processes as an LOC system, while the closely related Micro Total Analysis System (µTAS) is defined as the miniaturization of the overall analytic process.

Microfluidics factors into µTAS since it is the study of fluidics at the micro- and nano- scale, but microfluidics is moving beyond the simple ability to make networks of interconnecting channels and into higher level MEMS systems. We can presently incorporate more features into our microfluidics devices including electrical contacts and electromagnetic forces. For example, where mechanical pumps and valves were once used to manipulate fluids, we may now use electroosmotic flow because it is much easier to make an electrical contact than a robust mechanical pump [1].

Currently, most cellular study is performed *in vitro*, with conditions dissimilar to those *in vivo*. Cell platforms would help to mimic *in vivo* conditions in order to better understand cellular functions in the body. Microfluidics could benefit the study of cells in several ways: greater interest in the biochemical analysis of living cells, straightforward integration into a larger system, various methods for large numbers of cell manipulations on the system, sizing of cells fits easily within microfluidic devices (10-100µm), manipulation of single cells workable with MEMS devices, and heat and mass transfer very fast in fluidic systems [2].

Membrane elasticity in cells is an important indicator of disease. For example, research has demonstrated that RBC deformability is an important factor for determining the severity of malaria cases. The less deformable the membrane is, the more likely the malaria will be fatal [3]. In another example, it has been shown that arteries with cuboidal (or rounded) endothelial cells develop atherosclerotic lesions more quickly than arteries containing elongated ECs [4]. Artherosclerosis is clogging, narrowing, and hardening of the body's large arteries and medium-sized blood vessels.

II. Design

A. Silicon vs. SU-8 Microchannels

Traditionally, the simplest way to form a microchannel is to etch a trench in silicon, either by wet or dry etching, and then to adhesively bond a silicon wafer on top to close the channel. Evidently, this technique has its limitations because silicon is opaque and observations cannot be made optically about the fluid in the channels. Improvements to these channels were made by anodically bonding a glass wafer on to the silicon substrate [5]. Some problems with bonding wafers are that they must be very clean and smooth or microvoids could be created.

Instead of bonding wafers together to close the channels, researchers have devised a method to create buried microchannels. The technique is complex consisting of deep reactive-ion etching (DRIE), oxidation, a reactive-ion etch (RIE) to reveal the bottom of the trench, an isotropic etch with KOH to form the channel, and finally closing the channel using low pressure chemical vapour deposition (LPCVD) [6]. The limitations from this technique include its intricate fabrication and once again, that it does not provide the possibility of optical observations.

SU-8 is an epoxy-based negative photoresist. It has so far proven to be biocompatible [7], making it ideal for fabricating low-cost medical devices. SU-8 is also a transparent material, and it has the advantage of being capable of achieving very high aspect ratios (<18) [8]; however, there is no simple process of making a channel cavity. With positive resist, we can develop away the unexposed resist cavity, whereas with negative resist, part of the channel would become cross-linked and we would be unable to develop it away. In this paper, we present one technique for fabricating microchannels from SU-8.

B. *Channel Shape*

The aspects that we wish to examine with the cell platforms are membrane deformability and elasticity of red blood cells and endothelial cells, specifically, the maximum amount of cell membrane deformation and the relaxation time. Thus, we modelled the channels as an inlet structure, a transition, and a constricted channel, followed by another transition, and outlet structure. Figure 1 outlines this channel design.

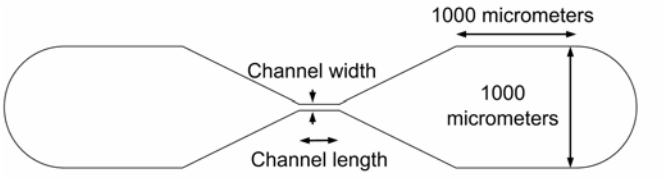


Fig. 1. Microchannel shape. Channel is composed of inlet, transition, constricted channel, transition, and outlet.

C. *Channel Size*

The diameter of a red blood cell (RBC) is between seven and eight micrometers and its shape is that of a biconcave disk [9]. It has been found that when stress is applied isotropically the membrane is highly resistant but would easily stretch in one direction; making is simple for the cell to change its shape in order to squeeze through capillary vessels [10]. Endothelial cells have a diameter of about 20 microns before they plate and between 25 to 65 microns when they plate [11] but can be lengthier when elongated.

Based on the cell diameter, we fabricated a range of channels with varying widths and lengths. The height of the channel must be set because this depends on the thickness of SU-8 we spin. We have chosen a thickness of 10 micrometers for the RBCs and 35 micrometers for the ECs. The widths and length were varied to maximize channel variants in order to achieve greater design success. Table 1 is a summary of the channel dimensions chosen.

TABLE I
CHANNEL DIMENSIONS

RBC Channel		EC Channel	
Width (μm)	Range of Length (μm)	Width (μm)	Range of Length (μm)
2	4 – 20	50	100 - 500
3	6 – 30	70	140 – 700
4	8 – 40	90	180 – 900
5	10 – 50	110	220 – 1100
6	12 – 60	130	260 – 1300
7	14 -70	150	300 – 1500
8	16 – 80	170	340 - 1700
9	18 – 90	190	380 – 1900
10	20 – 100	210	420 – 2100

III. FABRICATION AND RESULTS

A. *Process Flow*

After cleaning the wafer, we dehydrated it by baking it in a convection oven at 100°C for 10 minutes. We spun on a 35 micrometer layer of SU-8 for the endothelial cell channels, and then we patterned the layer using our mask and developed. Next, we spun on a 10 micrometer layer of SU-8, pattern, and develop. Figure 2 outlines the entire process flow.

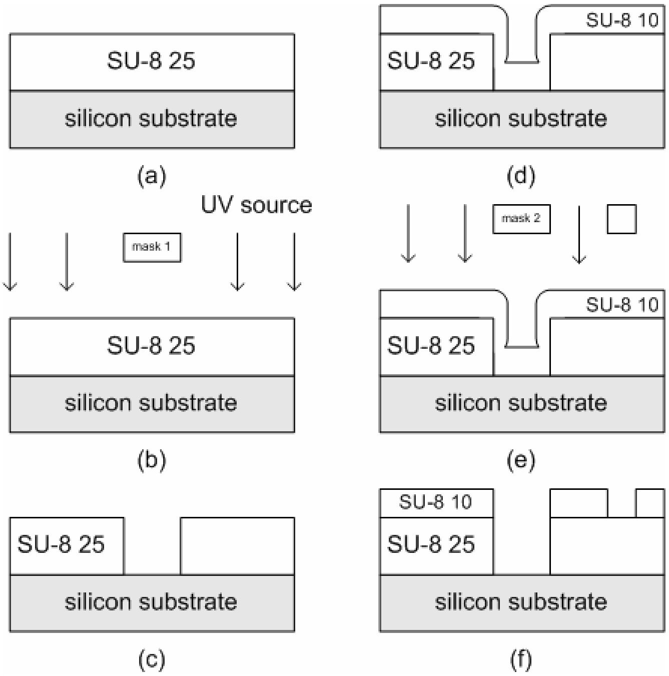


Fig. 2. Process flow for EC and RBC channels. (a) Spin on SU-8 25. (b) Expose SU-8 using first mask. (c) Develop un-crosslinked SU-8. (d) Spin on SU-8 10. (e) Expose using second mask. (f) Develop to reveal final structure.

B. *Results*

We successfully fabricated microchannels in SU-8 on silicon; however, some problems were encountered in the initial stages. To begin with, we were having difficulty developing all the SU-8 away from the microchannels. Through experimentation, we deduced that smaller exposure times were necessary to prevent cross-linking in the channels. Cracks in the SU-8, particularly around the channels, were observed. Using three stages of ramping during pre- and post-bake enabled us to remove most of these excess stress marks. Figure 3 illustrates the unreleased SU-8, and stress cracks, Figure 4 demonstrates a completely released microchannel, and Figure 5 illustrates a wafer-level view of the endothelial cell channels.

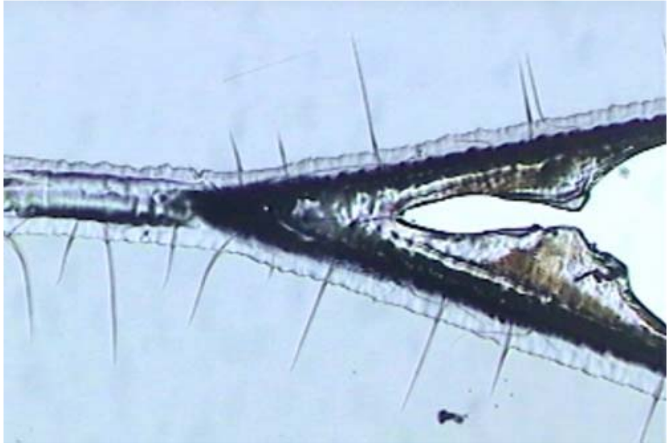


Fig.3. Partially released SU-8 in microchannel. Observe the stress cracks around the exterior of the microchannel.

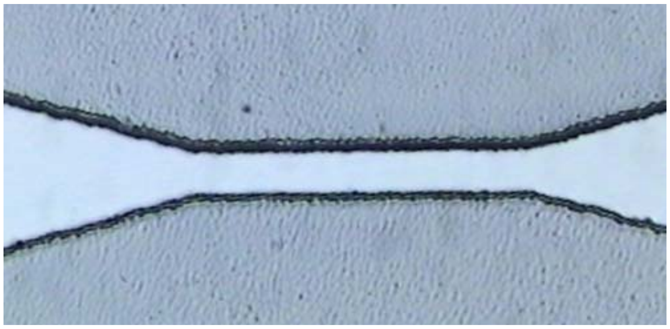


Fig. 4. Released SU-8 microchannel.

IV. CURRENT WORK

The proof-of-concept channels that we fabricated provided a building block for our cell platform system. Currently, we are fabricating endothelial channels on glass slides and growing cells in these channels. We chose to fabricate channels on glass slides instead of glass wafers for cost effectiveness, and for ease of cellular experimentation.

Because of difficulty spinning a uniform layer of SU-8 on a glass slide due to its rectangular shape, we fixed the slides to a handle wafer for the fabrication process. We glued the slides on using a thick photoresist, SC1827 that dissolves in SU-8 developer (at which point our slides and channels release). A successfully fabricated glass microchannel is illustrated in Figure 5.

Currently, our collaborators at the University of California, Davis, are working to grow endothelial cells in the channels. The cell data will enable us to determine if our system is viable and if we should continue in this stream of research.

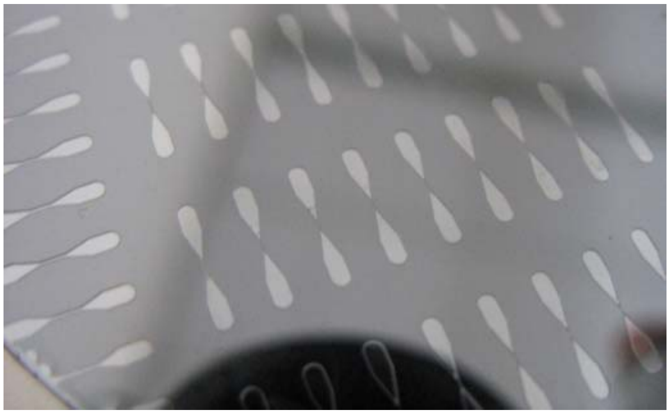


Fig. 5. Fully released endothelial cell microchannels.

V. FUTURE WORK

The work examined in this paper provides a background to a larger cell platform system that includes a network of channels for moving fluids within the chip, and interlocking structures for transferring components on and off chip. To move towards this cell platform, we need to enclose our channels by affixing a sheet of acrylic or glass using an epoxy similar to SU-8 [12] or PMMA (poly-methylmethacrylate) [13]. Following this procedure, we should examine the channel pressure using a syringe pump to ensure the channel walls can withstand the fluid flow of the cells. SU-8 has an ultimate strength of approximately 120 MPa for an 80 micrometer thick sample, which is higher than typical plastics that have ultimate strength less than 70MPa[14]. Finally, interlocking structures should be designed, fabricated and tested to enable system modularity. Design ideas include cylindrical posts, lock and key components, and hexagonal ports [15].

REFERENCES

[1] E. Verapoorte and N. De Rooij, “Microfluidics meets MEMS,” *Proceedings of the IEEE*, vol. 91, no. 6, pp. 930-953, June 2003.

[2] H. Andersson and A. van den Berg, "Microfluidics devices for cellomics: a review," *Sensors And Actuators B-Chemical*, vol. 92, no. 3, pp. 315-325, July 15. 2003.

[3] A. M. Dondorp, P. A. Kager, J. Vreeken, and N. J. White, “Abnormal Blood Flow and Red Blood Cell Deformability in Severe Malaria,” *Parasitology Today*, vol. 16, no. 6, pp. 288-232, June 2000.

[4] R.M. Nerem, “Vascular fluid mechanics, the arterial wall, and atherosclerosis,” *Journal Of Biomedical Engineering*, vol. 114, no. 3, pp. 274–282, August 1992.

[5] R. W. Tjerkstra, P. Ekkels, G. Krijnen, S. Egger, E. Berenschot, K. C. Ma, and J. Bugger, “Etching technology for microchannels,” in *TRANSDUCERS, Solid-State Sensors, Actuators and Microsystems, 12th International Conference on*, 2003, pp 147-152.

[6] M. J. de Boer, R. W. Tjerkstra, J. W. Berenschot, H. V. Jansen, G. J. Burger, J. G. E. Gardeniers, M. Elwenspoek, and A. van den Berg, “Micromachining of buried micro channels in silicon,” *Journal of Microelectromechanical Systems*, vol. 9, no. 1, pp.94-103, March 2000.

[7] G. Voskerician, M. S. Shive, R. S. Shawgo, H. Recum, J. M. Anderson, M. J. Cima, and R. Langer, “Biocompatibility and biofouling of MEMS drug delivery devices,” *Biomaterials*, vol. 24, no. 11, pp 1959-1967, May 1993.

[8] J. Zhang, M.B. Chan-Park, J. Miao, and T.T. Sun, “Reduction of diffraction effect for fabrication of very high aspect ratio microchannels in SU-8 over large area by soft cushion technology,” *Microsystem Technologies*, vol. 11, no. 7, pp 519-25, July 2005.

[9] C. Starr and R. Taggart, "Cell structure and function," in *Biology: The Unity and Diversity of Life* ,10th ed.M. Julet, Ed. California: Brooks/Cole, 2004, pp. 54-79.

[10] R. M. Hochmuth, "Measuring the mechanical properties of individual human blood cells," *Journal of Biomechanical Engineering,* vol. 115, no. 4B, pp 515-519, 1993.

[11] B. Gray, Simon Fraser University, Burnaby, BC, private communication, April 2006.

[12] N. Kuan, D. Liepmann, and A. P. Pisano, "Fabrication and packaging of microfluidic devices with su-8 epoxy," in *Micro-Electro-Mechanical Systems (MEMS), 2000 ASME International Mechanical Engineering Congress and Exposition,* 2000, pp. 591-594.

[13] B. Bilenberg, T. Nielsen, B. Clausen, and A. Kristensen, "PMMA to SU-8 bonding for polymer based lab-on-a-chip systems with integrated optics," *Journal of Micromechanics and Microengineering,* vol. 14, no. 6, pp. 814-818, June 2004.

[14] A. McAleavey, G. Coles, R. L. Edwards, and W. N. Sharpe, Jr, "Mechanical properties of SU-8," in *Materials Science of Microelectromechanical Systems (MEMS) Devices Symposium,* 1999, pp. 213-218.

[15] B. L. Gray, D. K. Lieu, S. D. Collins, R. L. Smith, and A. I. Barakat, "Microchannel platform for the study of endothelial cell shape and function," *Biomed Microdevices,* vol. 4, no. 1, pp. 9-16, March. 2002.

Stephanie Westwood Born in Vancouver, Canada in 1983, Stephanie earned her Bachelors of Applied Sciences degree at the University of Ottawa in Ontario, Canada in 2001. She graduated Cum Laude in Electrical Engineering with a concentration in Communications. Currently, she is working towards a Masters degree in Engineering Science at Simon Fraser University, Burnaby, BC, in the field of microinstrumentation.