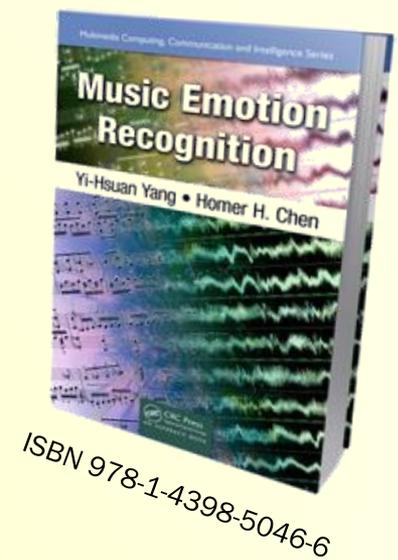


Music Emotion Recognition

Homer H. Chen
National Taiwan University
homer@cc.ee.ntu.edu.tw



Music Listening

Anytime, anywhere, anyone ...

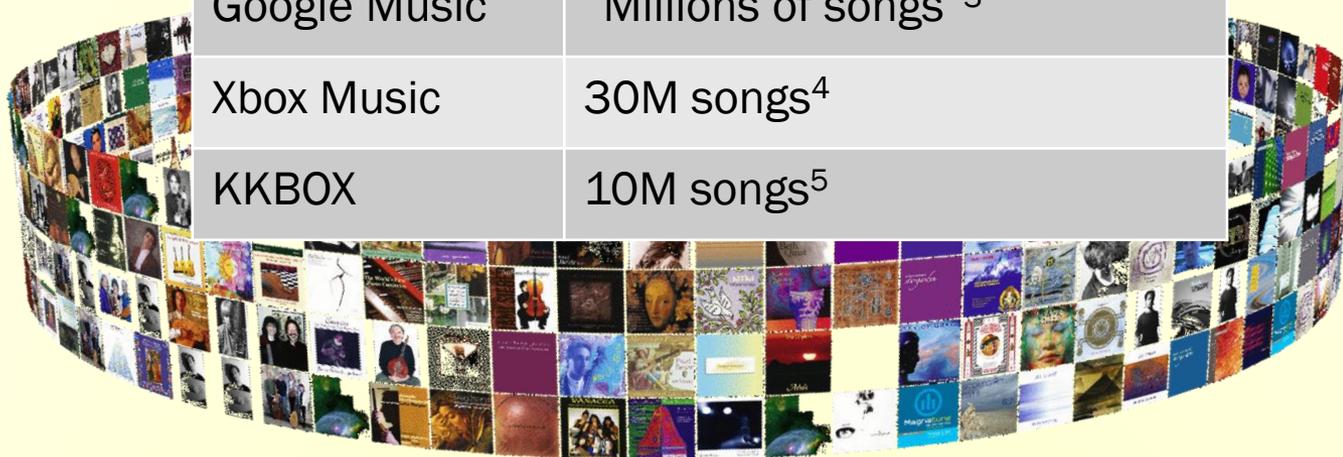


Evolution of Music Playback



Explosive Growing of Digital Music

Music provider	Statistics
iTunes store	26M songs ¹ 25B downloads (Feb. 2013)
Amazon MP3	25M songs ²
Google Music	“Millions of songs” ³
Xbox Music	30M songs ⁴
KKBOX	10M songs ⁵



¹Retrieved Sep. 2012, <http://www.apple.com/pr/library/2012/09/12Apple-Unveils-New-iTunes.html>

²Retrieved Jun. 2013, <http://www.amazon.com/MP3-Music-Download/b?ie=UTF8&node=163856011>

³<https://play.google.com/about/music/>

⁴Retrieved Jun. 2013, <http://www.microsoft.com/en-us/news/Press/2012/Oct12/10-14XboxMusicPR.aspx>

⁵Retrieved Jun. 2013, <http://www.ithome.com.tw/itadm/article.php?c=80653&s=1>

Source :
Government Information Office, Taiwan
<http://tw.kkbox.com>
<http://www.apple.com>

Music & Emotion

- Music expresses emotions

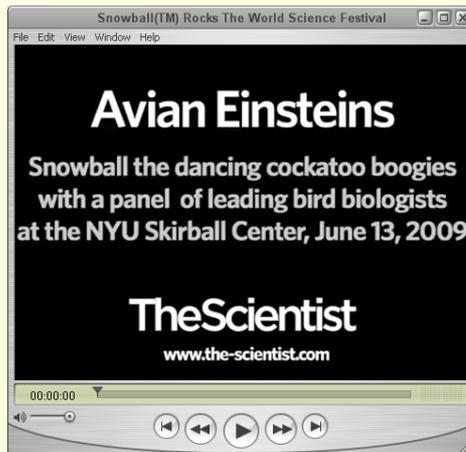
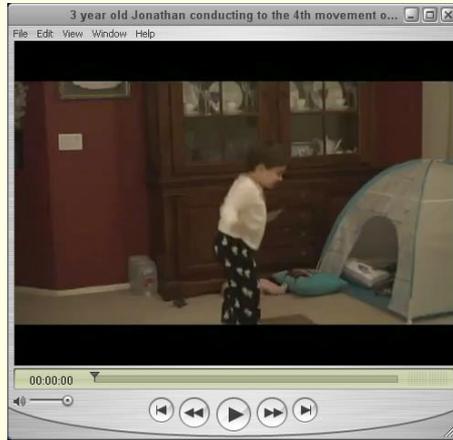


Music & Emotion

- Music induces emotions



Power of Music



Music Emotion Description

- **Categorical Approach:** Divide music into groups and describe each group by an adjective



Exhausted



Confused



Ecstatic



Guilty



Suspicious



Enraged



Ashamed



Cautious



Angry



Hysterical



Frustrated



Sad



Satisfied



Overwhelmed



Cheerful



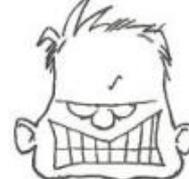
Lonely



Embarrassed



Happy



Mischievous



Disgusted



Frightened



Boring



Surprised



Anxious

Traditional Music Information Retrieval (MIR)

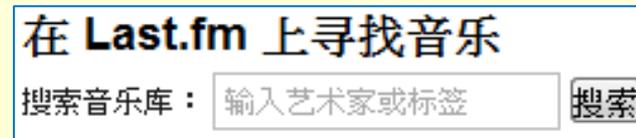
- Text based
 - Artist name
 - Song name
 - Album title
 - Lyrics
 - Genre

- KKBOX



歌曲名稱	歌手/演出者	專輯	音樂類型	播放
詞 突然好想你	五月天 (Mayday)	後青春期的詩	國語歌曲	播放
詞 你不是真正的快樂	五月天 (Mayday)	後青春期的詩	國語歌曲	播放

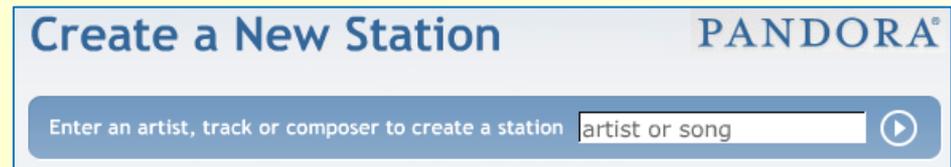
- Last.fm



在 Last.fm 上寻找音乐

搜索音乐库：

- Pandora



Create a New Station PANDORA®

Enter an artist, track or composer to create a station

All Music Guide

Mariah Carey



Photo by Markus Klinko & Indrani

Picture Browser
< Previous Next >

Born
Mar 27, 1970 in Huntington, NY

Years Active
1910 20 30 40 50 60 70 80 90
2000

Genres
R&B

Styles
• Dance-Pop
• Adult Contemporary
• Pop
• Adult Contemporary R&B

Biography by Jason Ankeny
The best-selling female performer of the 1990s, Mariah Carey rose to superstardom on the strength of her stunning five-octave voice. An elastic talent who moved easily from glossy ballads to hip-hop-inspired dance-pop, she earned frequent comparison to rivals Whitney Houston and Celine Dion, but did them both one better by composing all of her own material. Born in Long Island, NY, on March 27, 1970, Carey moved to New York City at the age of 17 -- just one day after graduating high school -- to pursue a music career; there she befriended keyboardist Ben Margulies, with whom she began writing songs. Her big break came as a backing vocalist on a studio session with ... » [Read more](#)

Moods
• Confident
• Party/Celebratory
• Sensual
• Sexy
• Carefree
• Exuberant
• Playful
• Refined/Mannered
• Stylish
• Amiable/Good-Natured
• Bright
• Energetic
• Fun
• Laid-Back/Mellow
• Reflective
• Romantic
• Sentimental
• Slick
• Sparkling
• Sweet

Instruments
• Vocals

Other Entires
• Classical Music Entry
• Movie Entry

AMG Artist ID
P 62404

Watch music videos by this artist!

Gracenote Media Manager

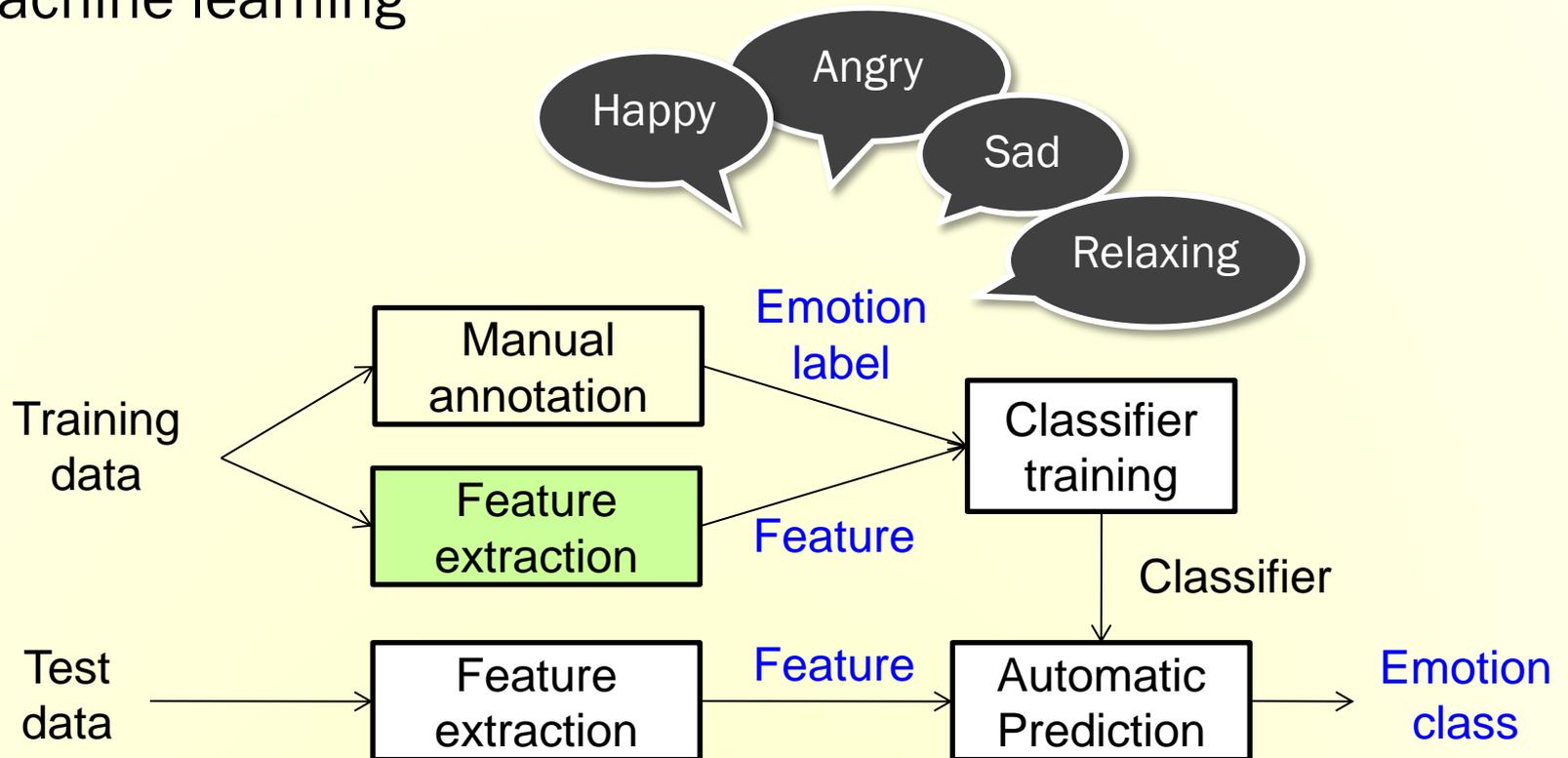
The screenshot shows the Gracenote Media Manager application window. The interface is divided into several sections:

- Top Bar:** Contains the application name 'gracenote.', a search bar, and playback controls like 'Menu', 'Queue', and a play button.
- Navigation:** Includes tabs for 'Library', 'Add Media', 'MusicID', 'More Like This', 'Discover', and 'Enhanced Content'.
- Left Panel:** A tree view showing a hierarchy of media categories such as 'Track Artists', 'Album Artists', 'Albums', 'Classical', 'Years', 'Genre', 'Origin', 'Era', 'Artist Type', and 'Mood'. The 'Mood' category is expanded, showing sub-categories like 'Peaceful', 'Romantic', 'Sentimental', etc.
- Main List:** A table of tracks with columns for 'Track Artist', 'Album', 'Title', 'Mood', and 'Tempo'. The 'Mood' column is highlighted with a red box.
- Bottom Bar:** Shows the current selection: '1 item (9:04) 10.39 MB'.

Track Artist	Album	Title	Mood	Tempo
The Association	The Association's Greatest Hits	Everything That Touches You	Idealistic / Stirring	Medium Fast
The Association	The Association's Greatest Hits	Never My Love	Energetic Groovy / Bitters...	Medium Fast
The Association	The Association's Greatest Hits	Cherish	Tender	Medium
Astor Piazzolla	Sur	Tristeza/Separation	Intimate / Nostalgic / Bitte...	Medium
Astrid Hadad & Los Tar...	Ay!	iAy Qué Dolor Vivir!	Sweet	Medium
Astrud Gilberto	Verve Jazz Masters 9	The Girl From Ipanema [Live]	Sultry / Swank	Medium
Astrud Gilberto	Verve Jazz Masters 9	A Certain Sadness	Tender	Medium Slow
Astrud Gilberto	Verve Jazz Masters 9	The Shadow Of Your Smile	Sophisticated / Lush / Ro...	Medium Slow
The Ataris	Blue Skies, Broken Hearts...N...	San Dimas High School Foot...	Hard Positive Excitement	Fast
Audioslave	Audioslave	Cochise	Hard Positive Excitement	Fast
Audra	The Arbitrary Width Of Shad...	There Are No Snakes In Hea...	Heavy Brooding	Medium Fast
Aurora Feat. Naimee C...	Euro Dance Hits	Ordinary World	Frenetic Euphoric Bliss	Fast
Average White Band	Funkgasm	Pick Up The Pieces	Dark Groovy / Savvy	Medium Fast
Avril Lavigne	The Best Damn Thing -Specia...	Girlfriend	Hard Dark Excitement	Fast
Avril Lavigne	The Best Damn Thing -Specia...	When You're Gone	Loud Strength & Glory	Medium
Avril Lavigne	The Best Damn Thing	Keep Holding On	Loud Strength & Glory	Medium Fast
Aynur Dogan	Crossing The Bridge: The Sou...	Ehmedo	Creepy / Eerie / Ominous	Medium
Aztec Camera	High Land, Hard Rain	Walk Out To Winter	Energetic Groovy / Bitters...	Fast
B.B. King	King Of Blues	When My Heart Beats Like A...	Gritty / Earthy / Soulful	Medium Slow
B.J. Thomas	The Very Best Of Burt Bachar...	Raindrops Keep Fallin' On My...	Cheerful / Playful	Medium Fast
B.T. Express	Disco Classics	Do It ('Til You're Satisfied)	Relaxed Groove	Medium Slow
B5	Radio Disney: Party Jams	Let's Groove	Edgy / Sexy	Medium Fast
The B-52's	The B-52's	52 Girls	Heavy Groove	Medium Fast
Baby Bash Feat. T-Pain	Cyclone	Cyclone	Arrogant / Attitude / Defi...	Medium
Babyshambles	The Blinding E.P.	The Blinding	Energetic Alienation / Mela...	Medium Fast
Bachman-Turner Overd...	BTO's Greatest	You Ain't Seen Nothing Yet	Confident / Celebratory	Medium Fast

Music Emotion Classification (MEC)

Machine learning



Subjects Annotation

- Ask human subjects to annotate music emotion



- Happy
- Sad
- Angry
- Relaxed

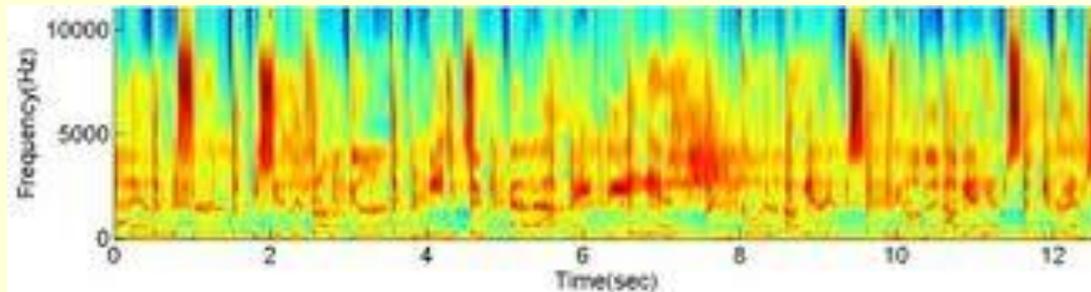


- Happy
- Sad
- Angry
- Relaxed

Music Features

- Spectral
 - Spectral centroid, spectral rolloff, spectral flux, MFCCs
- Temporal
 - Zero-crossing rate, temporal centroid, log-attack time
- Melody/harmony
 - Pitch class profile, key clarity, harmonic change, musical mode
- Rhythmic
 - Beat histogram, average tempo (BPM)

⋮



Spectral Features

- **Spectral centroid**

- average frequency of the signal weighted by magnitude

$$\sum_{n=1}^{R_t} M_t[n] = 0.5 * \sum_{n=1}^N M_t[n].$$

- **Spectral roll-off**

- how much of the frequencies are concentrated below a given threshold

$$\sum_{n=1}^{R_t} M_t[n] = 0.85 * \sum_{n=1}^N M_t[n].$$

- **Spectral flux**

- how much the frequency varies over time

$$F_t = \sum_{n=1}^N (N_t[n] - N_{t-1}[n])^2$$

- **Spectral flatness**

- Whether the spectral power is concentrated

$$\frac{\sqrt[N]{\prod_{n=0}^{N-1} x(n)}}{\left(\frac{\sum_{n=0}^{N-1} x(n)}{N}\right)}$$

Spectral Features (Cont'd)

- MFCCs

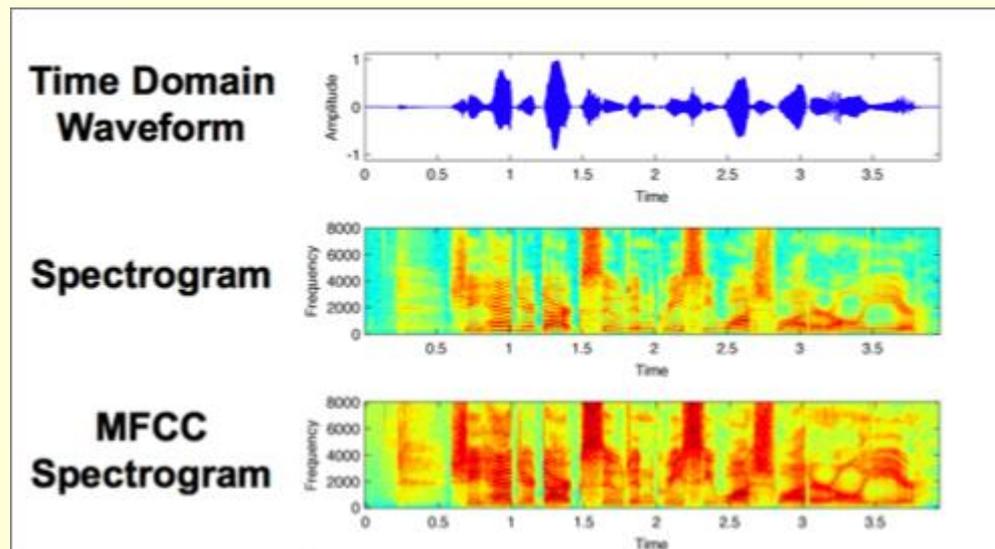
Windowing

FFT

Mel frequency mapping

Log $|\cdot|$

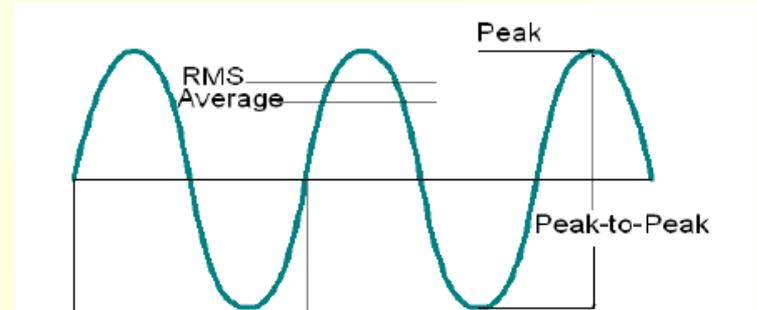
DCT



Loudness

- Root-mean-square energy (RMS)

- $\sqrt{\text{mean}(A^2)}$
- Classifying exciting/relaxing music

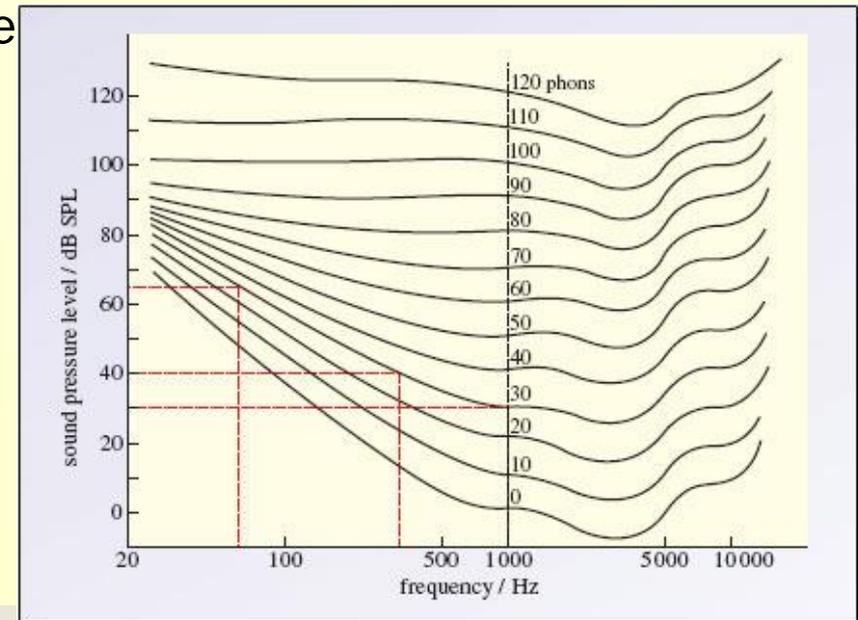


- **Low-energy** feature

- Percentage of “texture window” (1 sec) with RMS value under average
- Example: Vocal music with silence

- Intensity vs. loudness

- Physical / psychological
- Sound pressure level (db SPL) / phone

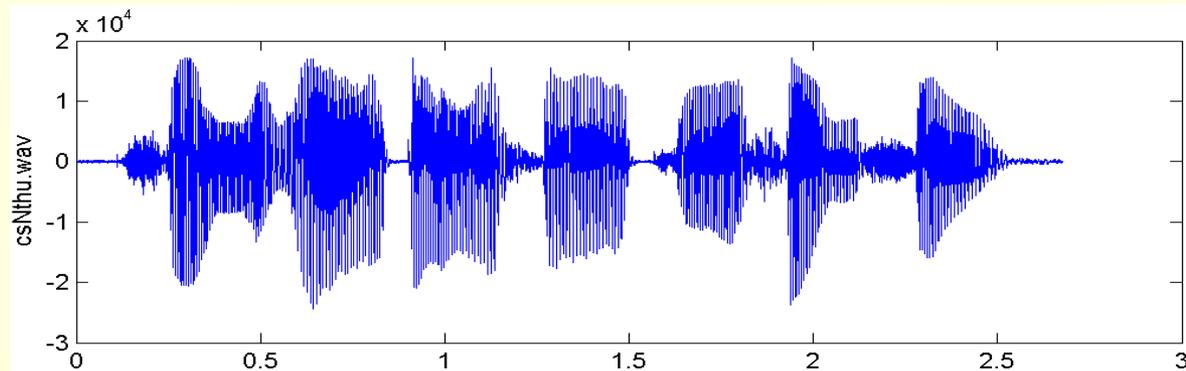


Zero Crossing Rate

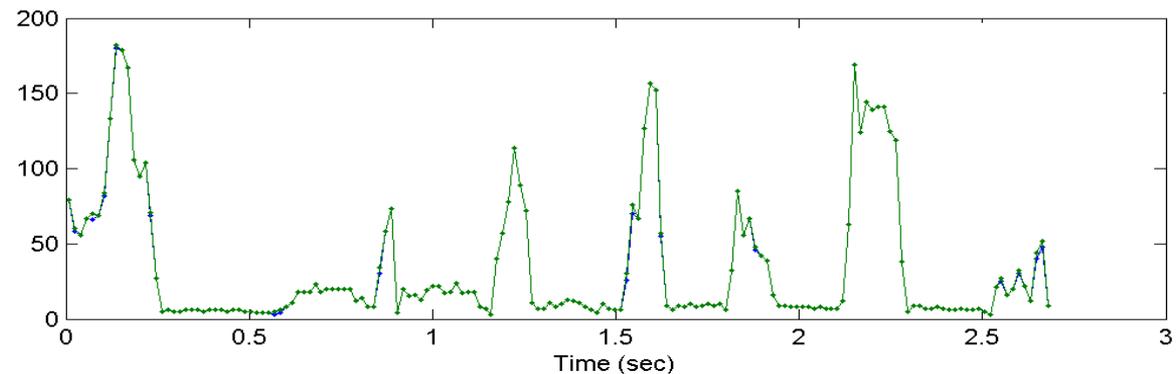
- Number of time domain crossings within a frame

$$Z_t = \frac{1}{2} \sum_{n=1}^N |\text{sign}(x[n]) - \text{sign}(x[n-1])|$$

Signal

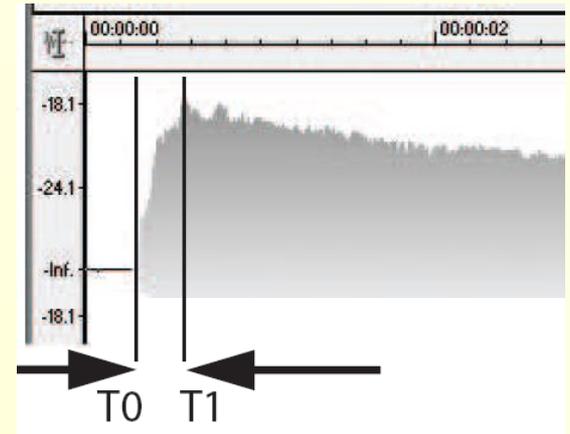
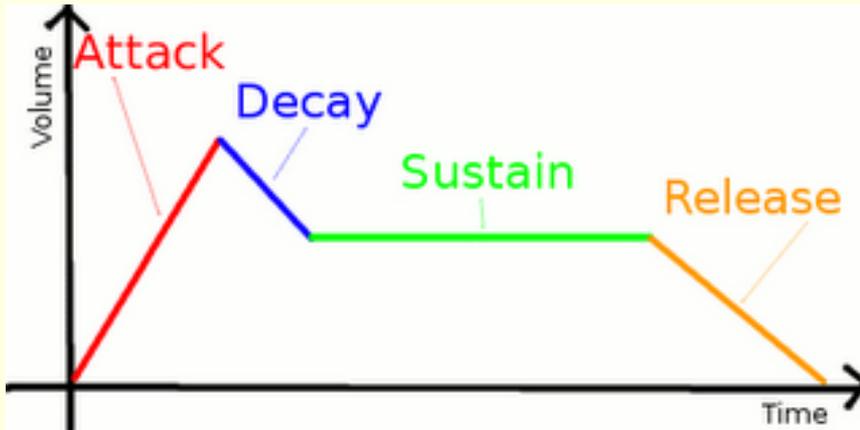


ZCR



ADSR

- Attack, Decay, Sustain, Release



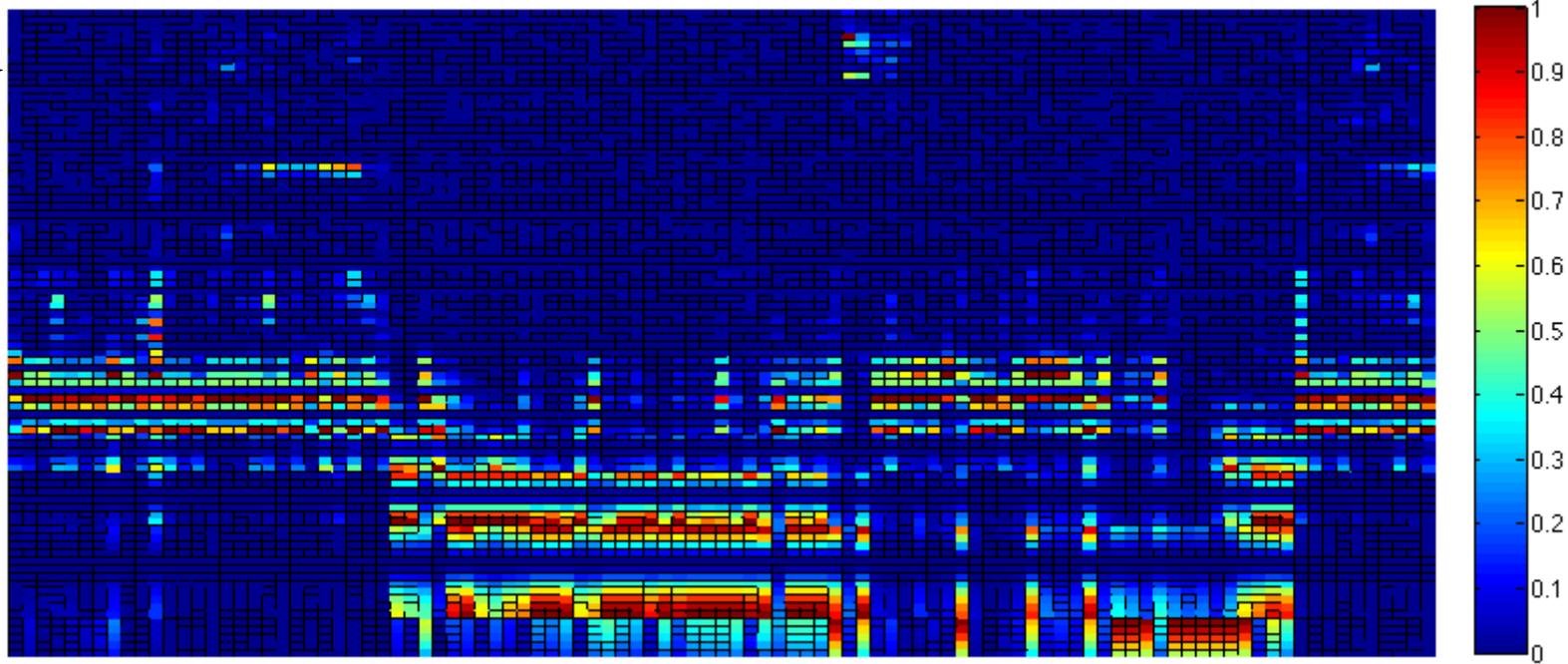
- Log attack time

$$LAT = \log_{10}(T1 - T0)$$

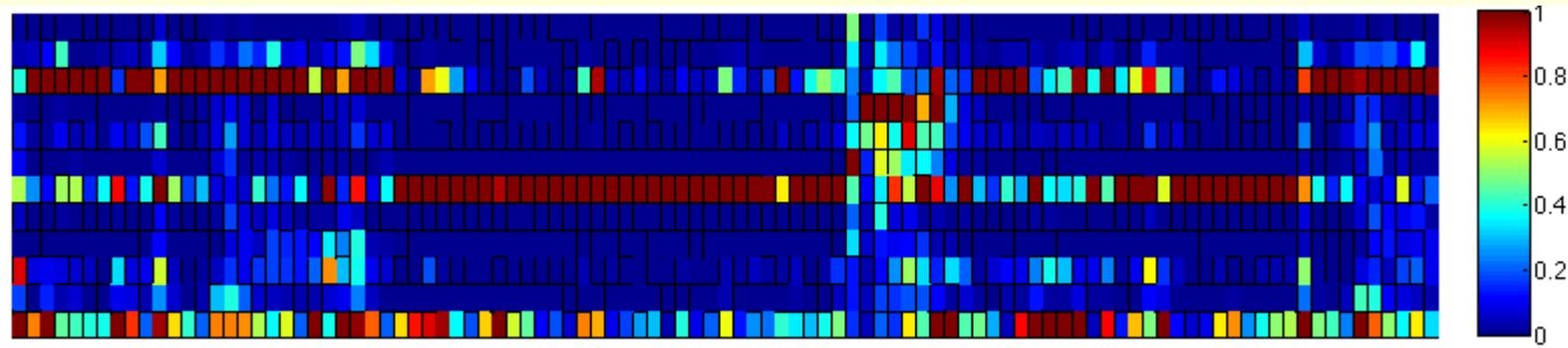
- the time it takes to reach the maximum amplitude of a signal from a minimum threshold time

Pitch Class Profile

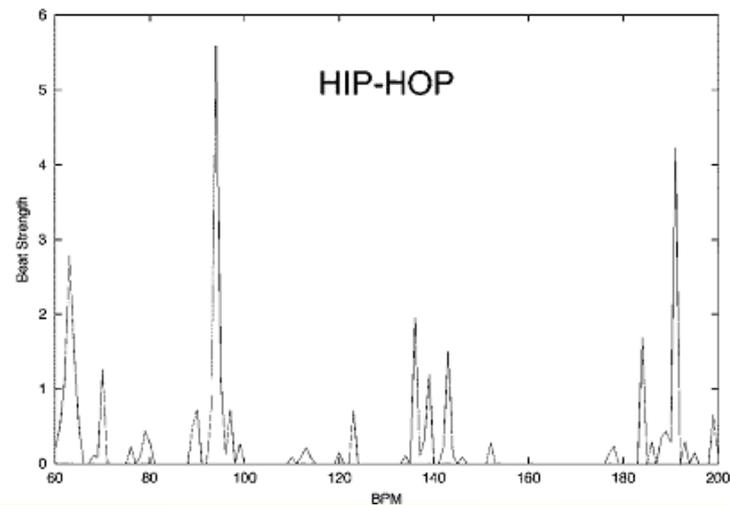
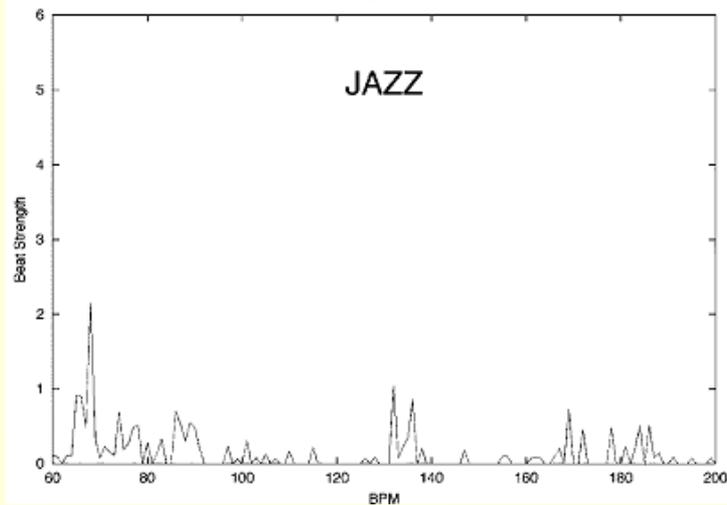
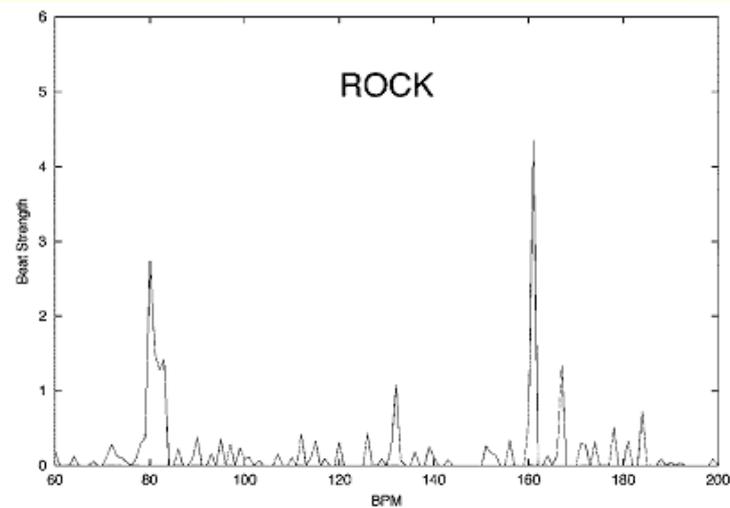
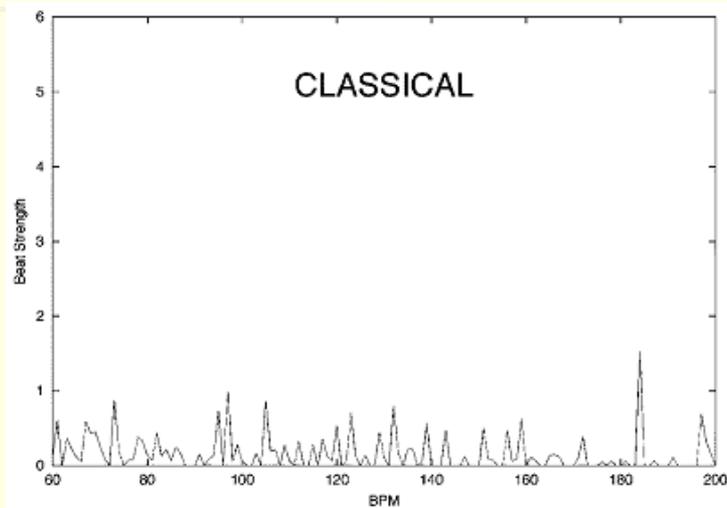
Unfolded



Folded



Beat Histogram



Problems of Categorical Approach

- Ambiguity

- Happy = joyous = cheerful = delighted?
- Sad = sorrowful = depressed = gloomy?

- Granularity

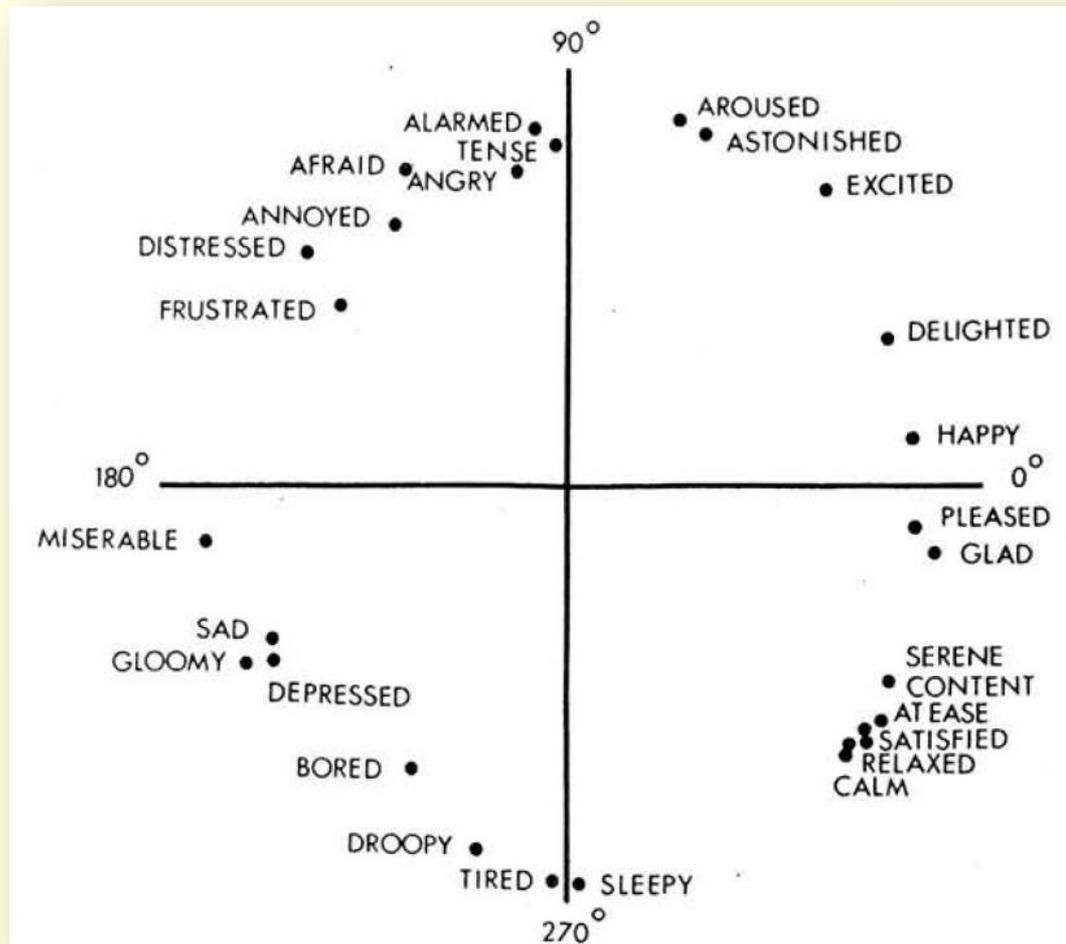
- Small number of emotion classes
 - Insufficient to describe the richness of human perception
- Large number of emotion classes
 - Difficult to obtain ground truth values

abandoned, abashed, abused, aching, admiring, adoring, adrift, affectionate, affronted, afraid, aggravated, aglow, ailing, alarmed, alienated, alienated, alone, ambivalent, anguished, annoyed, annoyed, antagonistic, anxious, apart, apologetic, appalled, appreciative, apprehensive, ardent, ashamed, attached, attentive, awful, awkward...

Dimensional Approach

Arousal

- Activation, activity
- Energy and stimulation level

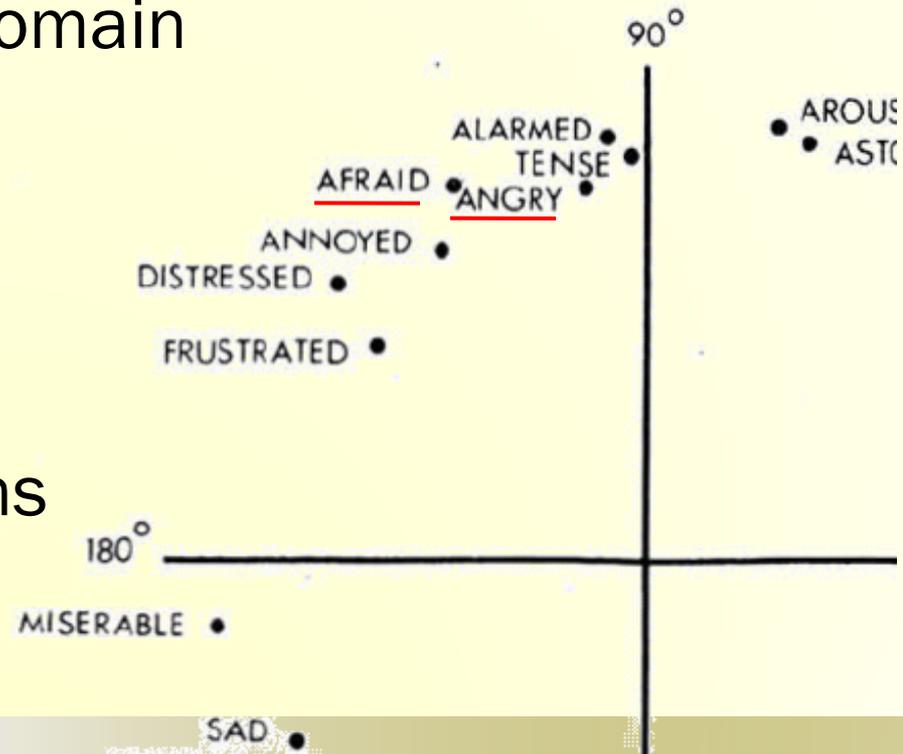


Valence

- Pleasantness
- Positive and negative affective states

Dimensional Approach

- Strength
 - No need to consider which and how many emotions
 - Generalize MER from categorical domain to real-valued domain
 - Provide a simple means for 2D user interface
- Weakness
 - Blurs important psychological distinctions
 - Afraid, angry



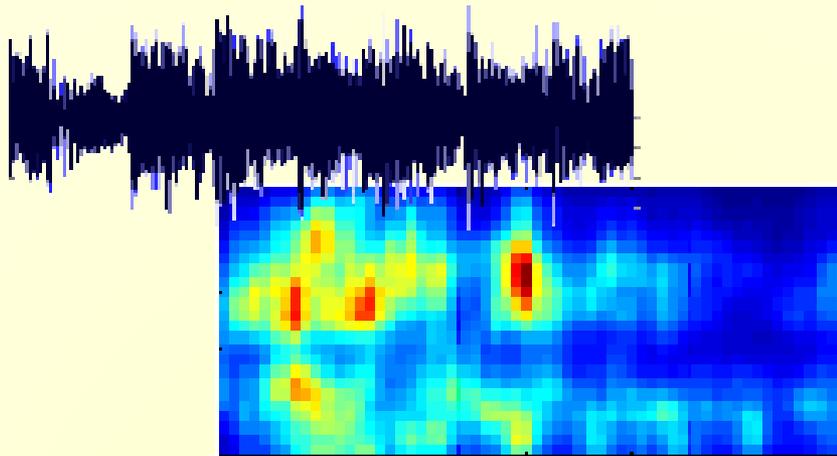
Arousal and Valence Features

- **Arousal**

- **Pitch:** high/low
- **Tempo:** fast/slow
- **Timbre:** bright/soft
- ⋮

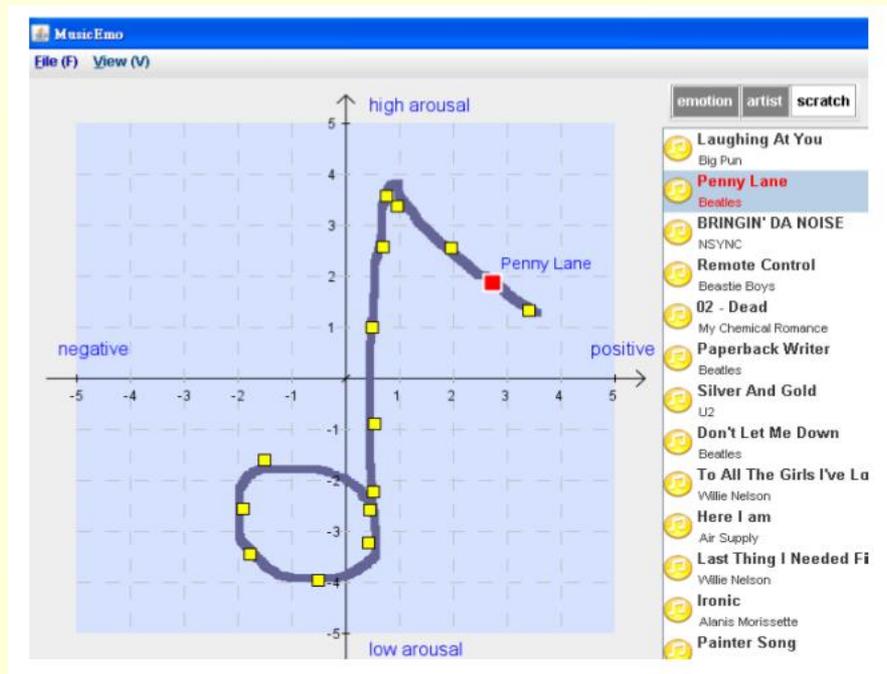
- **Valence**

- **Harmony:** consonant/dissonant
- **Mode:** major/minor
- **Tonality:** tonal/atonal
- ⋮



Mr. Emo

- Developed by our lab at NTU
- Each music piece is a point in the emotion space
- A great app for smart phones



demo.lnk

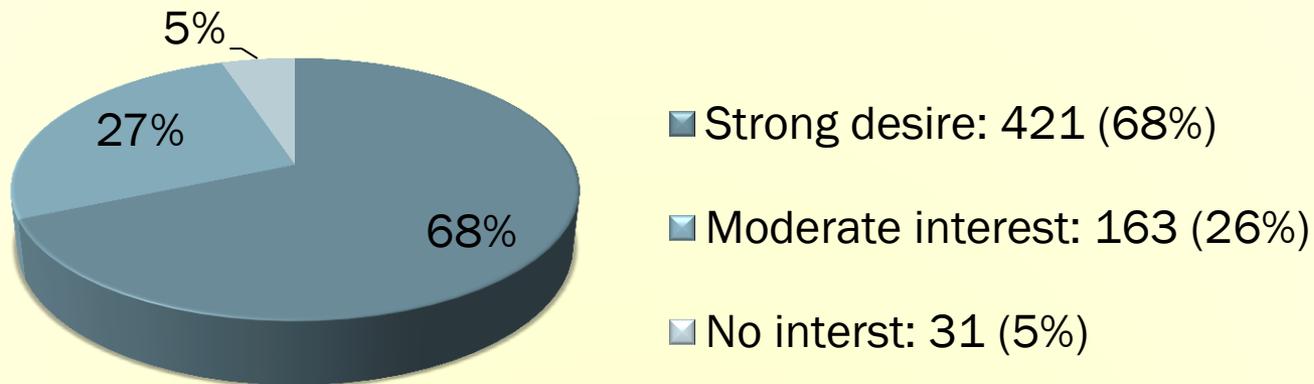
[Live Demo](#)

YouTube

<http://www.youtube.com/watch?v=ra55xO20UHU>

Retrieving Music by Emotion

- Complementary to traditional MIR method
- Survey: Would you like to retrieve music by emotion?
 - 615 subjects (mostly college students):



Musicoverly

The screenshot displays the Musicoverly website interface. At the top left is the Musicoverly logo. A search bar with the placeholder text "Search an artist" is located at the top center. On the top right, there are links for "Log In", "Sign Up", and a dropdown menu for "U.S.A.". Below the search bar, navigation tabs include "Mood", "Dance", "Artist radio", "New", "Widget", and "Lab".

The "Mood" section features a slider with "Energetic" on the left and "Calm" on the right, with "Dark" and "Positive" also indicated. Below the slider is a decade selector with buttons for "50s", "60s", "70s", "80s", "90s", and "00s", with "1958" selected under the "50s" button.

A grid of music recommendations is shown, each with a cover image and artist name:

- Roy Orbison: Fool's Hall of Fame
- Bob Marley: Could you be loved
- Nat King Cole: Blue moon
- Eliane Elias: Desafinado
- Eric Clapton: Layla
- The Verve: Bitter sweet
- O.M.D: Tesla girls
- Ziggy Marley: Legend
- Billie Holiday: Nat King Cole
- Stan Getz & Joao: Live 2004
- Sugar Daddy: Live 1980
- Monty Alexander: Live 1980

At the bottom left, there is a "share my mood with this song" section with Facebook and Twitter icons, and a list of genre checkboxes including Rock, Reggae, Jazz, Pop, Soul, Blues, Soundtrack, Gospel, Vocal Pop, Latino, Disco, World, Rap, Funk, Classical, R & B, Electro, and Metal.

<http://musicoverly.com/>

Moodagent

- Automatically profiles music based on emotion, mood, genre, style, tempo, beat, vocals, instruments and production features
- Create mood-based playlists by setting the mood-sliders or choosing a seed track
- Need to profile a song on PC if online database does not have the entry



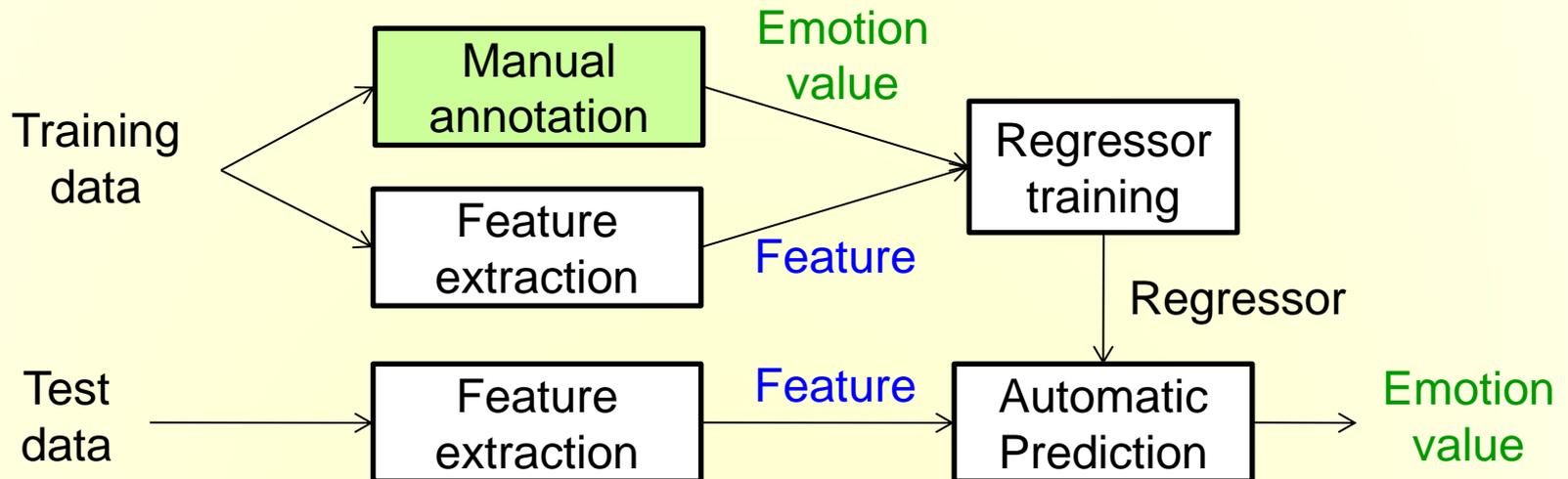
<http://www.moodagent.com/>

Obtain Music Emotion Rating

- Subject annotation
 - Rates the VA values of each song

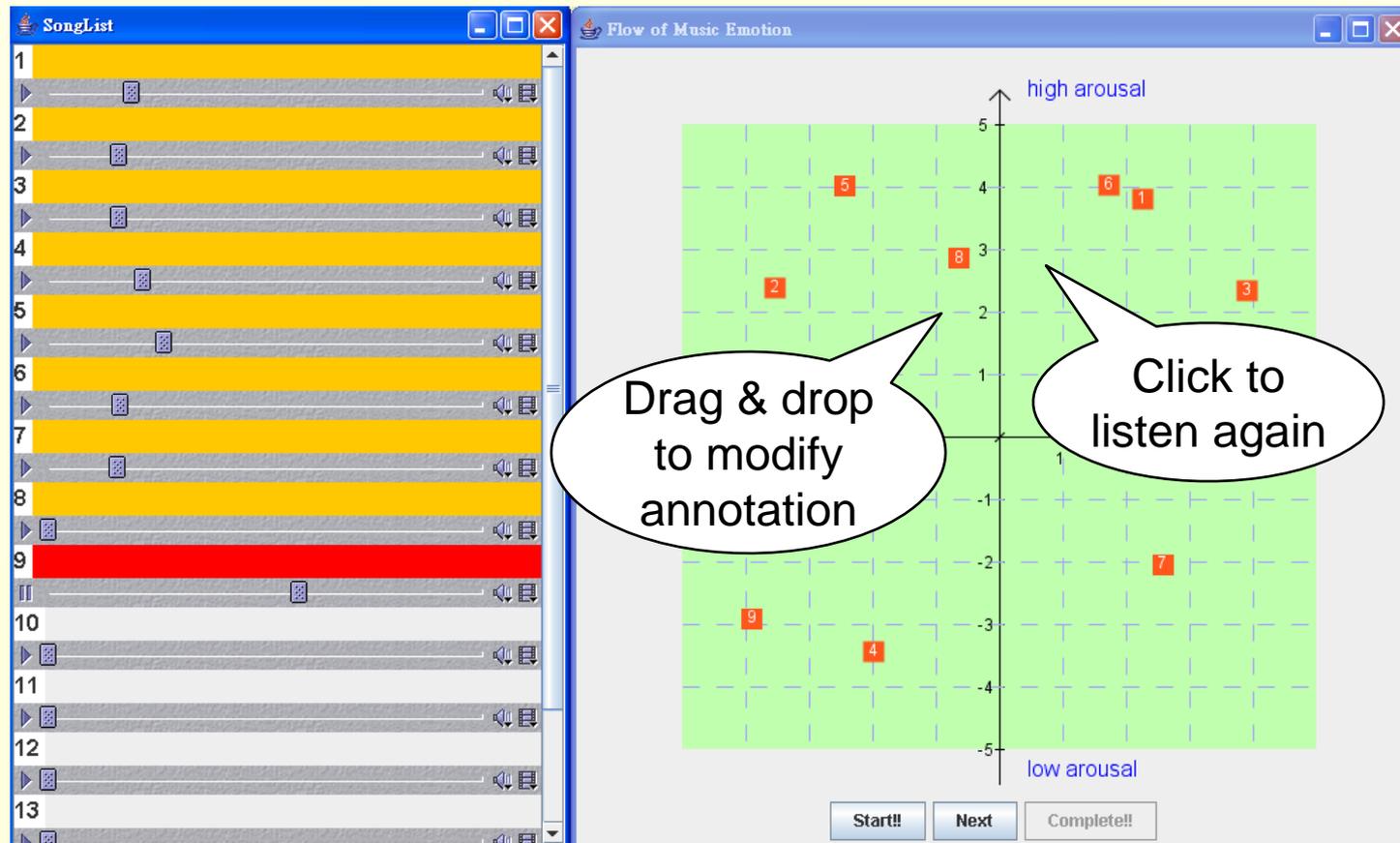
- Ordinal rating scale 

- Scroll bar 



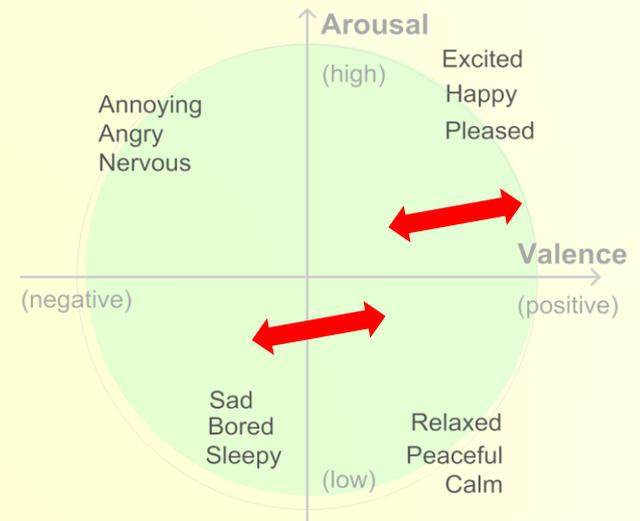
AnnoEmo: GUI for Emotion Rating

- Easy to differentiate



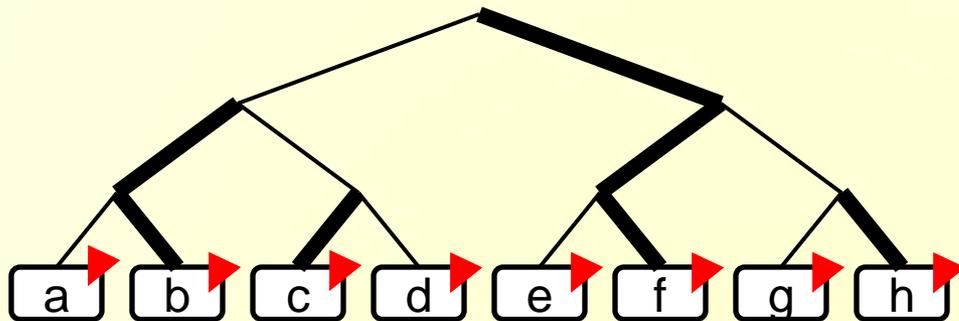
Difficulty of Emotion Annotation

- The cognitive load is high
- Difficult to ensure the rating scale is used consistently
 - Is the distance between 0.6 and 0.8 equals to the distance between -0.1 and 0.1 ?
 - Does 0.7 mean the same for two subjects?



Ranking-Based Emotion Annotation

- Emotion tournament
 - Requires only $n-1$ pairwise comparisons
 - The global ordering can later be approximated by a greedy algorithm



	a	b	c	d	e	f	g	h	
a									0
b	■		■	■					3
c			■	■					1
d				■					0
e									0
f	■	■	■	■	■		■	■	7
g									0
h								■	1

Which songs is more positive?

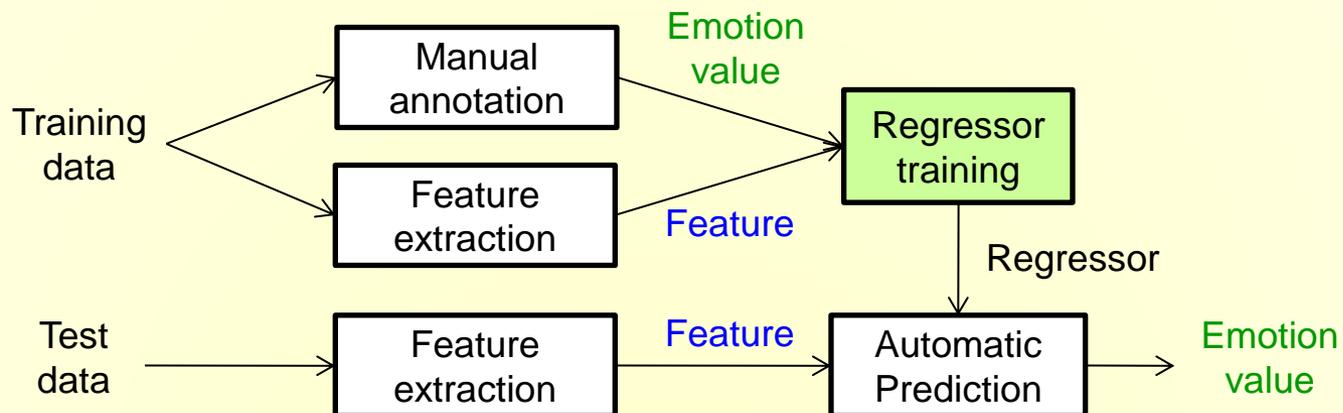
$f > b > c = h > a = d = e = g$

Regression

- Predict the VA values

- Trains a regression model (regressor) $f(\cdot)$ that minimizes the mean squared error (MSE)
- One for valence; one for arousal

- y_i : numerical emotion value
- \mathbf{x}_i : feature (input)
- $f(\mathbf{x}_i)$: prediction result (output)
e.g. linear regression
 $f(\mathbf{x}_i) = \mathbf{w}^T \mathbf{x}_i + b$



Improving Valence Recognition by Lyrics

- Lyrics

- 張惠妹 - 人質

- Without lyrics  Original



- Lyrics - 在我心上用力的開一槍
讓一切歸零 在這聲巨響
如果愛是說什麼都不能放
我不掙扎

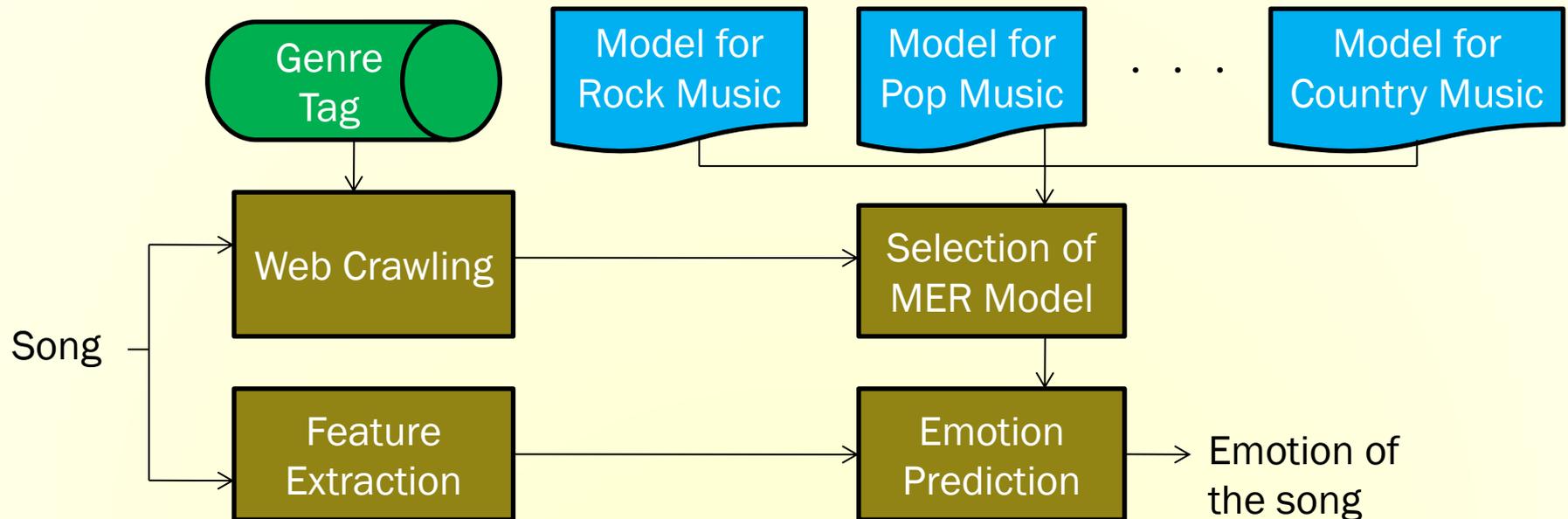
- Without lyrics - neutral

- With lyrics - sad

- Improves accuracy of valence by 19.9%

Improving MER by Genre Tags

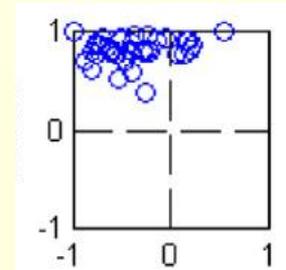
- Genre information makes music retrieval more effective



- MER accuracy increases 13.0%

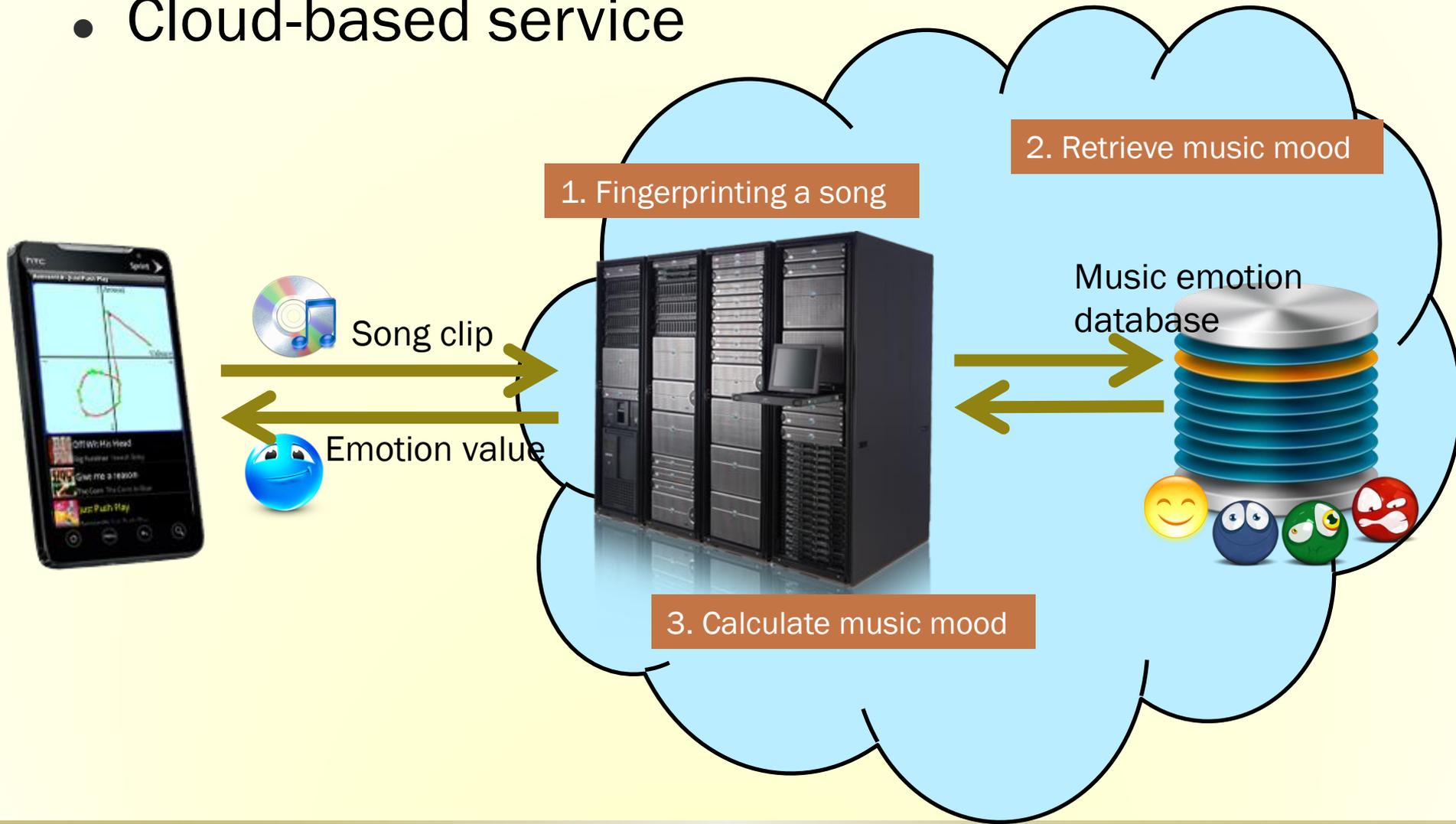
Personalized MER System

- People feel the same song differently
- A general model that fits everyone is almost impossible
 - Smells Like Teen Spirit by Nirvana 
 - Most people annotate it as negative valence
 - Rock music fans annotate it as positive valence
- Our approach
 - Build personalized model by user feedback
 - Choose an MER model according to personal information (gender, age, music preference, etc.)



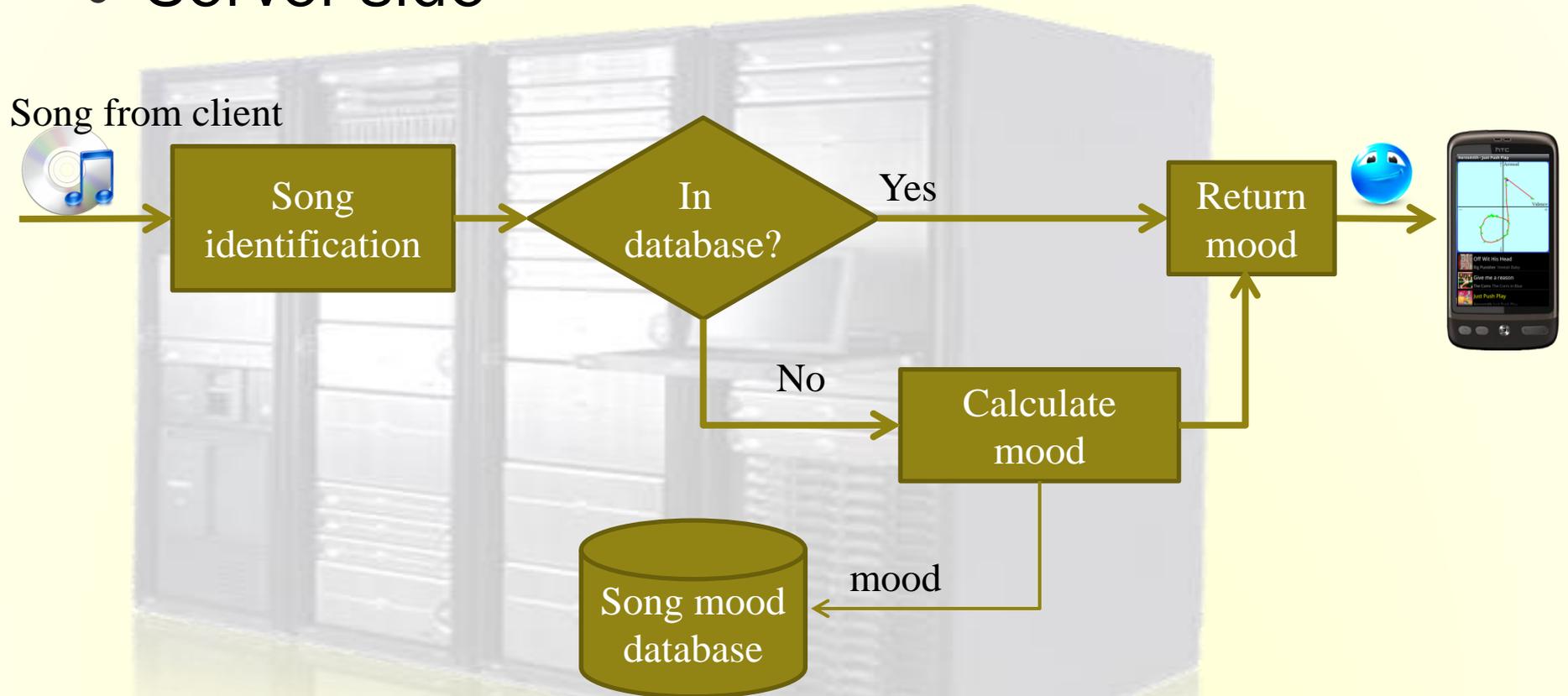
Emotion-Based Music Retrieval System

- Cloud-based service

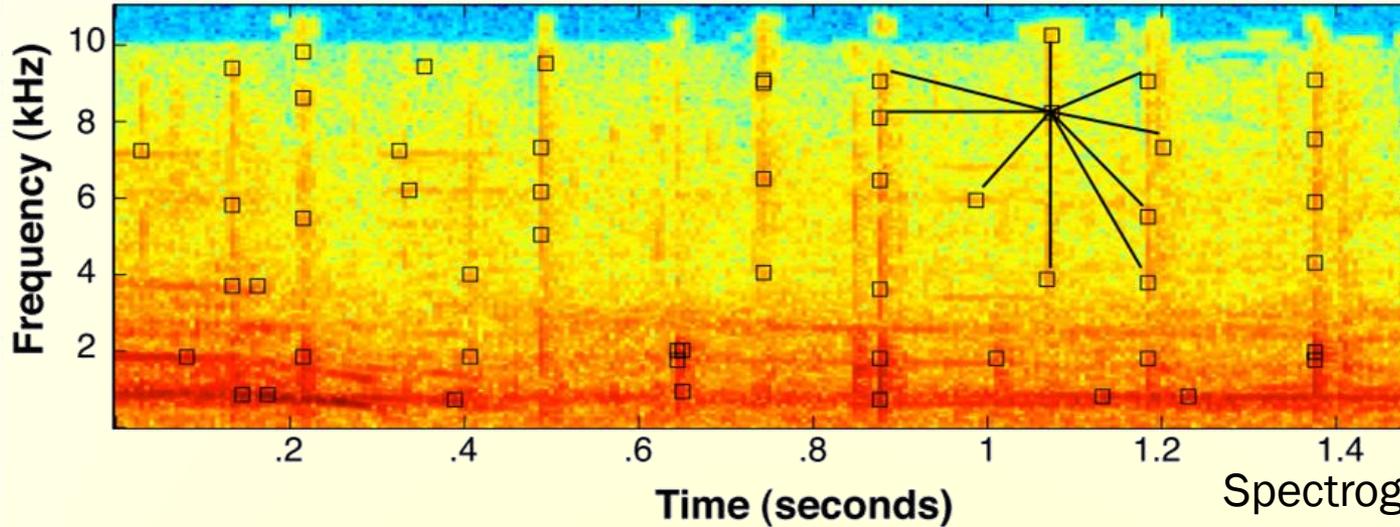


Emotion-Based Music Retrieval System

- Server side



Audio Fingerprinting



Spectrogram of a song clip with peak intensities marked

Frequency (Hz)	Time (seconds)
823.44	1.054
1892.31	1.321
712.84	1.703
...	...
819.71	9.943

Example fingerprint

Emotion-Based Music Retrieval System

- Client side
 - Automatic song profiling
 - Retrieve songs by emotion coordinates
 - Represent playlist by emotion trajectory
 - Show the mood distribution of all songs of an artist
- Ported to Android phone and iPhone



Acknowledgement



Yi-Hsuan Yang



Sighter Liu



Ming-Yen Su



Yu-Ching Lin



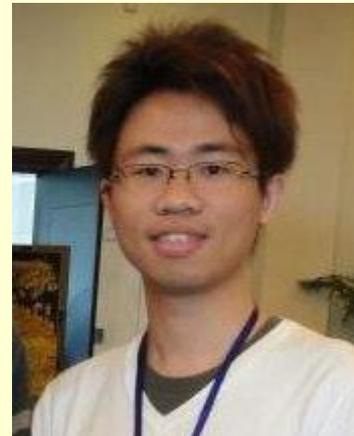
Cheng-Te Lee



Ya-Fan Su



Ann Lee



Keng-Sheng Lin



Cheng-Ya Sha



Pei-Chun Chen

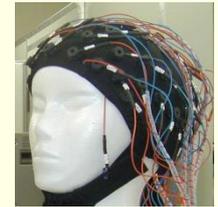


Heng-Tze Cheng

Extensions



Recognize users' emotion by EEG for music recommendation



EEG



Music therapy



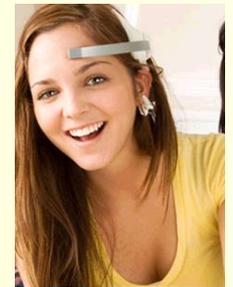
Music accompaniment



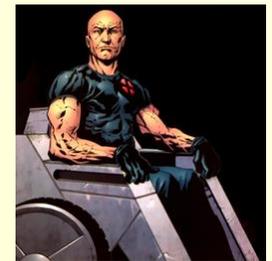
Incidental music selection



Karaoke system



NeuroSky

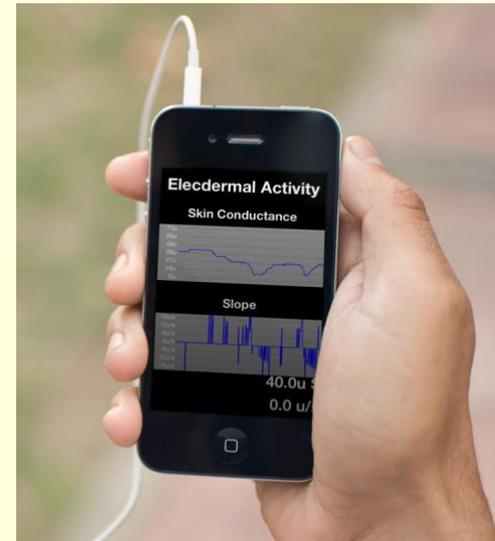


Prof. Charles Xavier

Electrodermal Sensor



Affectiva



NTU

Automatic Video Highlight Extraction

- Traditional system only considers low-level visual features
 - Motion, color
- Use the emotion of incidental music to improve accuracy

Buzzer Beat
34 sec



The flower shop without Roses
22 sec



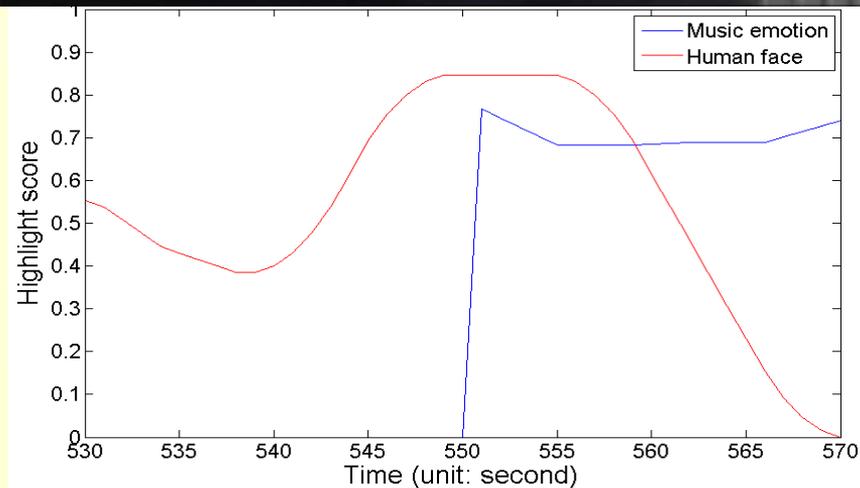
Last Friends
24 sec



K.-S. Lin, A. Lee, Y.H. Yang, and H. H. Chen, "Automatic highlights extraction for drama video using music emotion and human face features," in *Proc. IEEE Workshop on Multimedia Signal Processing*, Nov. 2011.

Romantic Music

- Human face and music



Automatic Transcription

- Classical music
 - No lyrics
 - Arouses emotion by melody
- Automatic Transcription of Piano Music
 - Extracts melody information

Song	Prelude and Fugue No.2 in C Minor	Sonata no. 8 Pathetique in C minor, 3 rd movement	Moments Musicaux No.4	Sonata K.333 in Bb Major, 1 st Movement
Composer	Bach	Beethoven	Schubert	Mozart
Original				
Result				

C.-D. Lee, Y.-H. Yang, and H. H. Chen, "Multipitch estimation of piano music by exemplar-based sparse Representation," *IEEE Trans. Multimedia*, vol. 14, no. 3, pp. 608-618, Jun. 2012.

Singing Voice Timbre Classification

- Using singing voice Timbre to classify music
- Build a new data set for this task
- Empirically validate that
 - Using vocal segment detection and singing voice separation improves the classification accuracy
 - Voice features are remarkably effective
- Applications: singing voice timbre as a high-level feature

C.Y. Sha, Y.-H. Yang, Y.-C. Lin and H. H. Chen, "Singing voice timbre classification of Chinese popular music," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2013.

The KKTIM data set

- KKBOX Timbre
 - 387 Chinese songs (91 singers)
 - 272 Japanese songs (66 singers)
- 6 singing timbre classes
- Multi-label, per-song instead of per-singer

Chinese	#Song	
低沉 (Deep)	74	 黃小琥
沙啞 (Gravelly)	57	 阿杜
渾厚 (Powerful)	70	 那英
甜美 (Sweet)	54	 郭靜
空靈 (Ethereal)	63	 林憶蓮
高亢 (High-pitched)	81	 林志炫

Japanese	#Song	
轉音多、鼻音 (Run riffs/Nasal)	39	 中孝介
沙啞 (Gravelly)	50	 T.M.Revolution
渾厚 (Powerful)	50	 久保田利伸
活力偶像、甜美 (Sweet)	50	 真野惠里菜
乾淨、明亮 (Bright)	40	 奧華子
高亢 (High-pitched)	50	 MISIA

Demo

- Chinese songs
- Probability output of six classifiers

Song ID	低沉 Deep	沙啞 Gravelly	渾厚 Powerful	甜美 Sweet	空靈 Ethereal	高亢 High-pitched	標記 Ground truth
384	0.72	0.49	0.51	0.10	0.56	0.58	1,0,0,0,0,0
117	0.43	0.62	0.67	0.28	0.13	0.21	0,1,0,0,0,0
631	0.56	0.52	0.60	0.30	0.05	0.40	0,0,1,0,0,0
632	0.25	0.50	0.32	0.91	0.20	0.45	0,0,0,1,0,0
443	0.30	0.47	0.45	0.38	0.84	0.45	0,0,0,0,1,0
371	0.67	0.54	0.50	0.31	0.11	0.48	0,0,0,0,0,1
636	0.14	0.54	0.47	0.77	0.83	0.63	0,1,1,0,0,0

Probability of belonging to the class

Emotional Accompaniment Generation

Emotion

Possible Accompaniments

15 progression(s) were found.

Valence Axis

5

Arousal Axis

-3

Find Accompaniment

Choose an accompaniment: 1 (default)

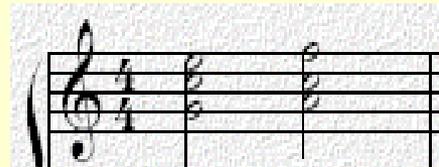
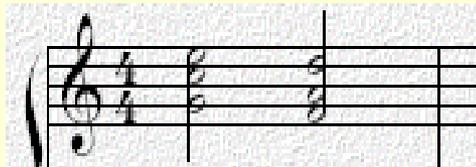
Generate MIDI file

Harmony progression 1: I IV VI II III VI
Harmony progression 2: VI IV V VI III IV
Harmony progression 3: VI II VI II III VI
Harmony progression 4: I VI III VI III IV
Harmony progression 5: I VI II VI III I
Harmony progression 6: VI IV VI II III I

P.-C Chen, K.-S. Lin, and H. H. Chen, "Emotional accompaniment generation system based on harmonic progression," *IEEE Trans. Multimedia*, v. 15, no. 7, pp. 1-11, Nov. 2013

Valence and Harmonic Progression

- Main valence-affecting features
 - ✓ Mode: major key-happy; minor key-sad
 - ✓ Chord: focus is on the consonance of a single chord
- Music is a delicate temporal art with emphasis on the flow of chords =>
Harmonic progression

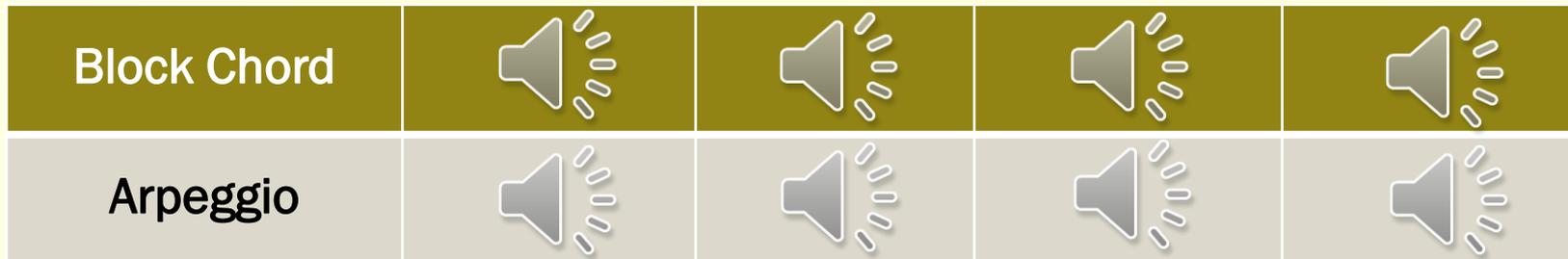
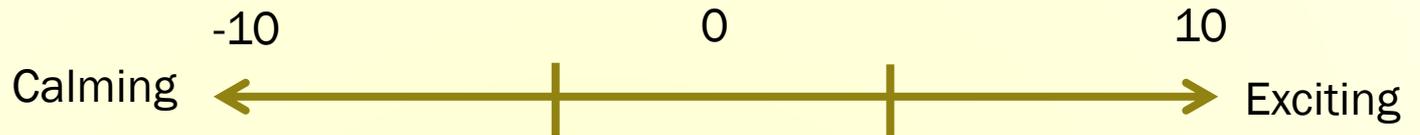


Putting Everything Together

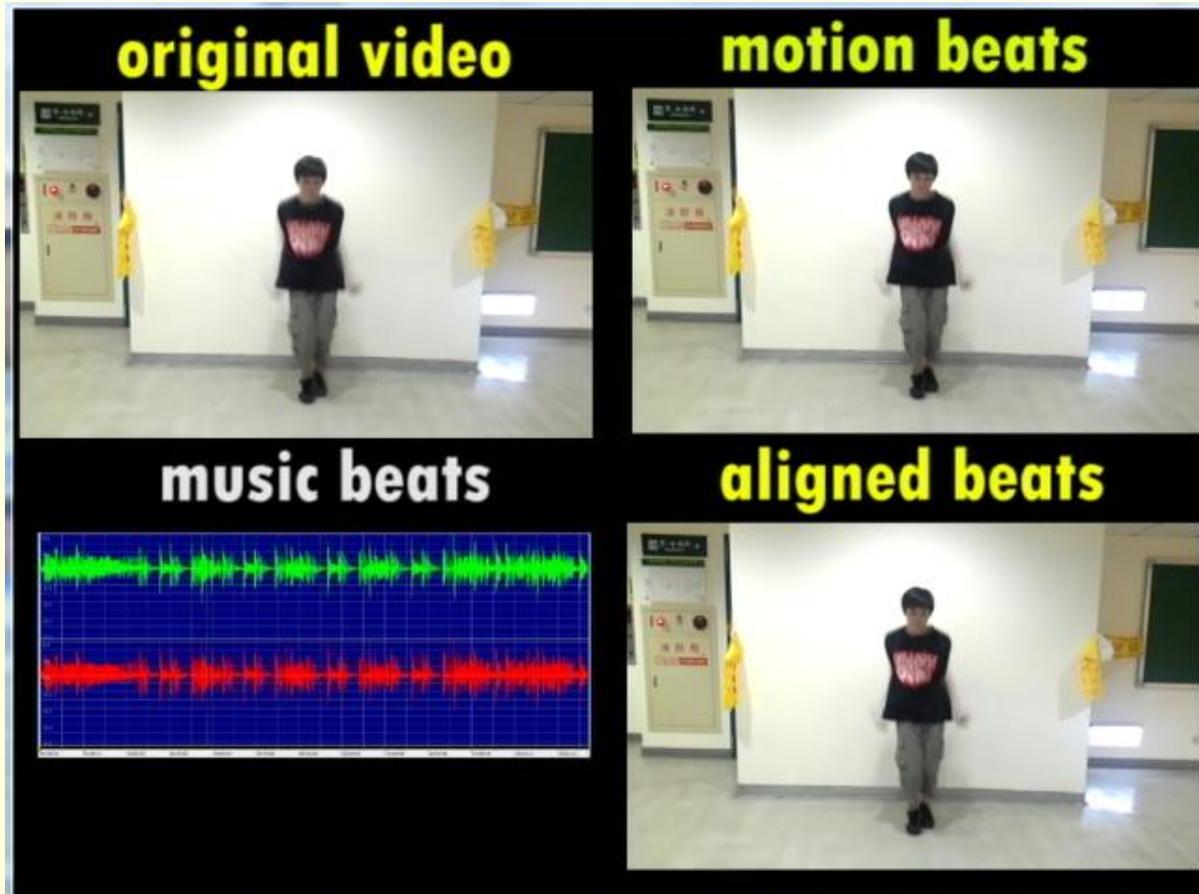
Valence values



Arousal values



Alignment Evaluation of Motion Beats



C. Ho, W.-T. Tsai, K.-S. Lin, and H. H. Chen, "Extraction and alignment and evaluation of motion beats for street dance," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2013.