International Journal of Latest Engineering Research and Applications (IJLERA) ISSN: 2455-7137

Volume – 02, Issue – 10, October – 2017, PP – 72-77

Human Emotion Based Song Player

P Rajitha Nair[1], HeyShanthini PandiyaKumari.S[2], Vinod Unnikrishnan[3]

[1]*(Department of Computer Science &Engg, New Horizon College of Engg/ Bangalore, India)*

[2]*(Department of Computer Science &Engg, New Horizon College of Engg/ Bangalo re, India)*

[3]*(IT, Dell International Services/ India)*

Abstract: Song Players with automatic song shuffling capability for mobile/personal computer/handheld computer are widely available and most of them accepts user feedback to identify user's mood and play songs as per their mood. A key improvement area in this approach is with the requirement for manual user input to the application to determine the current emotional state of the user. The onus is thus on the user to mark his present emotional state and hence doesn't cater for any dynamism in the emotions of the user.

This paper introduces an approach to add automated human emotion recognition mechanism with an active updated music provider which provides for the user to get an automated and seamless Song Shuffler. Facial Action Coding System devised by Carl-Herman Hjortsjöis the basis of the human emotion recognition aspect of this system. Music content used will be reviewed both by the user and also be based on the user's emotional change as a feedback to the music.

"Face is the mirror of one's soul"

Keywords: Song Shuffler, FACS, Human Emotion Recognition.

I. INTRODUCTION

1.1 FACIAL ACTION CODING SYSTEM (FACS)

Facial Action Coding System (FACS) has classified face movements of humans by their visibility. This is conceptualized on a system developed by a Swedish anatomist Carl-Herman Hjortsjö.[1] This was thenenhanced by Wallace V. Friesen & Paul Ekman.[2] Friesen , Ekman, and Joseph C. Hager provided a majorenhancement to FACS in the year 2002.[3] Individual facial muscles' movements are encoded by FACS with minor various instant modifications in facial appearance as foundation.[4] This is a usual standard to categorize expression of emotions, and has been useful to animators and psychologists. FACS leads amongst methodsin detecting human faces in images and videos, extracting features of faces, and creating profile of the facial movements.[4]

1.2 APPLICATION OF FACS

FACS can be used to recognize any expression of the face that is anatomically possible. The facial expression is deconstructed to Action Units (AU) and the expression producer's temporal segments. AUs are not dependent on any interpretation, that is, an Action Unit itself doesn't mean an emotion. Rather it is the collection of Aus in a given context that helps to decode the underlying human emotion. FACS manual already provides detailed interpretation of the Aus meaning. Aus as per FACS is the contraction and relaxation of one or more facial muscle. FACS can also distinguish subtle differences in the same resultant expression.

The below given human face anatomy will provide a quick reference to the facial muscles which is referred to in this article.

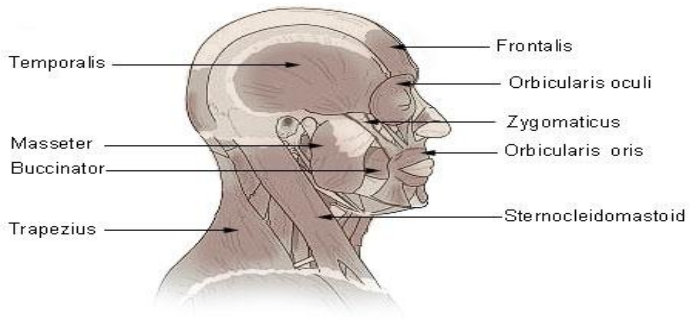

Fig 1: Reference Picture depicting Major Facial Muscles – Cross Section
Source: http://training.seer.cancer.gov/module_anatomy/images/illu_head_neck_muscle.jpg

1.3 ACTION UNITS AND ACTION DESCRIPTORS

Action Units (AUs) are the fundamental actions of groups of muscles or individual muscles.

Action Descriptors (ADs) are movements that may be due to the actions of several muscle groups (e.g.: uplifted eyebrows along with open jaw).

Intensity Scoring

FACS' Intensities are marked by prefixing letters A–E (for minimal-maximal intensity) to the Action Unit number (e.g. AU 3A is the weakest of AU 3 and AU 3E is the maximum intensity that is possible).

A - Trace
B - Slight
C - Marked or Pronounced
D - Severe or Extreme
E - Maximum

Examples of Emotions that FACS distinguishes accurately,

1. Pan-am Smile – Insincere and Voluntary identified by contraction of "Zygomatic Major" alone
2. Duchenne Smile – Involuntary and Sincere – identified by "Zygomatic Major" and inferior part of "Orbicularis Oculi"

The below given table gives a quick reference of the major emotions and the corresponding Action Units that FACS Manual describes.

Emotion	**Action Units**
Happiness	6+12
Sadness	1+4+15
Surprise	1+2+5B+26
Fear	1+2+4+5+7+20+26
Anger	4+5+7+23
Disgust	9+15+16
Contempt	R12A+R14A

Table 1: Reference Table of major emotions and corresponding Action Units as described in FACS Manual

1.4 SAMPLES OF AUS AND ADS

Main Codes

U	**FACS Name**	**Muscular Basis**
	Neutral face	
	Inner Brow Raiser	frontalis (pars medialis)
	Outer Brow Raiser	frontalis (pars lateralis)
	Brow Lowerer	corrugator supercilii, depressor supercilii, depressor glabellae,
	Upper Lid Raiser	superior tarsal muscle, levatorpalpebraesuperioris
	Cheek Raiser	orbicularis oculi (pars orbitalis)

Table 2: Main Codes

Head Movement Codes

AU	FACS Name	Action
51	Head Turn Left	
52	Head Turn Right	
53	Head Up	
54	Head Down	
M55	Head Tilt Left	Symmetrical 14 is preceded or accompanied by a tilt of head to left.

Table 3: Head Movement Codes

Eye Movement Codes

AU	FACS Name	Action
61	Eyes Turn Left	
M61	Eyes Left	The onset of the symmetrical 14 is immediately preceded or accompanied by eye movement to the left.
62	Eyes Turn Right	
M62	Eyes Right	The onset of the symmetrical 14 is immediately preceded or accompanied by eye movement to the right.
63	Eyes Up	

Table 4: Eye Movement Codes

Visibility Codes

AU	FACS Name
70	Brows and forehead not visible
71	Eyes not visible
72	Lower face not visible
73	Entire face not visible
74	unsociable

Table 5: Visibility Codes

Gross Behavior Codes

These codes are reserved for recording information about gross behaviors that may be relevant to the facial actions that are scored.

AU	FACS Name
40	Sniff
50	Speech
80	Swallow

Table 6: Gross Behavior Codes

II. USAGE OF FACS FOR EMOTION RECOGNITION – SONG SHUFFLER

The method used to recognize Human Emotions would be FACS in the system being developed. To analyze the user's emotions and keep track of the emotions the user interaction device's (mobile phone, handheld computing device, personal computers) camera module would be used real time by the Human Emotion Recognition engine component of the Song Shuffler system.

Fig 2: Quick View of the Key Components of Song Shuffler

The above image - fig 2 gives a quick view of the key components of the Song Shuffler app. The mobile device is just shown for representation and the Song Shuffler is expected to work with any mobile/computing device/personal computer/handheld computer that has a front facing camera in built.

The key parts of the Song Shuffler are,

1. Human Emotion Recognition Engine
2. Feedback Capture Engine
3. Music Content database – cloud based or locally stored
4. Music Player – to play the music files
5. Front facing Camera inbuilt in the computing device

2.1 HUMAN EMOTION RECOGNITION ENGINE

The Human Emotion Recognition Engine is the key module of the Song Shuffler and will be used to capture, record and provide input to other modules about the user's emotional state. It will directly interact with the computing device's front facing camera to capture the human facial expression.

The camera would continuously capture sequence of images as the user starts to interact with the Song Shuffler. Continuous engagement of the user with the user interaction device is not required for the proposed system to work as the proposed system would automatically capture the images using the device's front facing camera. These captured images will be initially converted to gray scale images and the focal point of the facial muscles (refer picture below for a sample) will be used to identify the AUs and ADs involved.

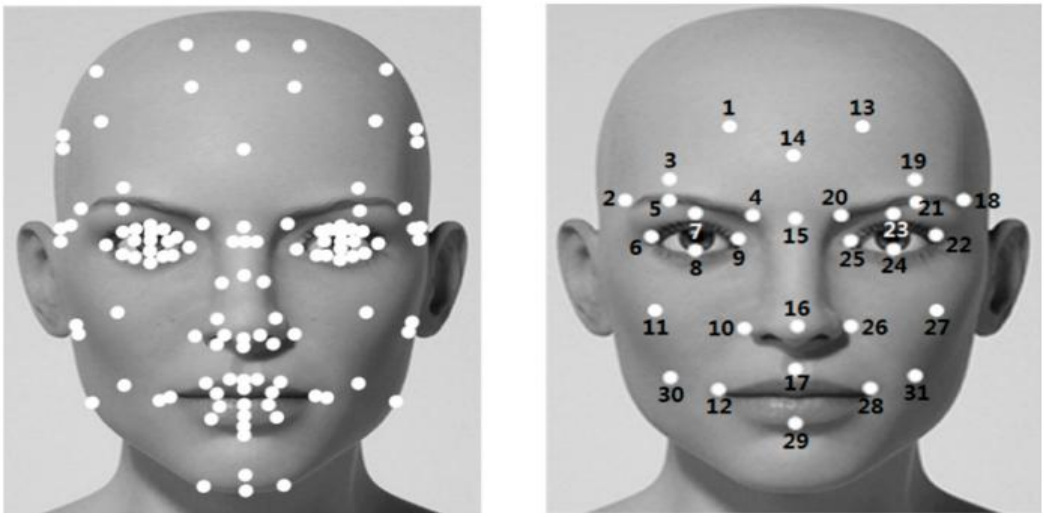

Fig 3: Sample gray scale image with muscle focal points marked

The images will be analyzed real time and the FACS analysis result of the image made available to the Song Shuffler system so that the next step of action, that is shuffling of songs or continuation of the songs can be performed.

2.2 FEEDBACK CAPTURE ENGINE

The Feedback Capture Engine would obtain the input from the Human Emotion Recognition Engine periodically when the Song Shuffler is playing songs.

The captured feedback will then be used for two main purposes,

2.2.1 Shuffle Song

When human hear songs, the mood of the listener is also impacted either positively, negatively or neutrally. That is the user's mood is either improved – that is the user becomes happier or the mood deteriorates – that is the user becomes sadder, or the mood doesn't change.

These mood swings result in change of emotion and change of emotion results in change of facial expression of the user. These changes will be recognized by the Emotion Recognition engine and the input will be made available to the Feedback Capture Engine. The Feedback Capture Engine in turn would compare the recorded emotion with the song that is being played and confirm whether the song befits the mood of the listener. If the song doesn't fit in, then it will be shuffled and a better fit song will be played next. To reduce drastic change in the music which could be unpleasant the songs would also be marked on a scale of 10 how fit it is to its mood. Say Music Track 1 fits for happy but only at a 5 point. Music Track 2 fits for sad but at a 10 point and so on.

2.2.2 Record Feedback of the Played Music

The other function of the Feedback capture engine would be to record the user's emotion as feedback for the song. The Human Emotion Recognition Engine's provided input would also be recorded as a feedback for the song and the scale rating would be also based on the emotion shown by the user whilst listening to the music track.

In addition to the real time feedback captured from the user the engine would also use the feedback already available for the music tracks from reviewers and from other users who use the cloud service to store their songs.

Both the available feedbacks would be used in cohesion to arrive at the correct song to be played for the user.

2.3 MUSIC PLAYER

The function of the Music Player would be to read through the commands sent by the Feedback Capture Engine via the Music Content database and keep playing the next available song.

The data-source for the music player could be locally held music tracks in the user's computing device or cloud service where the music as purchased or free content available for the user are available.

The detailed UI features of the Music Player are out of context for this article as the key focus is on the usage of Human Emotion Recognition.

III. CONCLUSION

The discussed method for recognition of human emotion using FACS is not completely error proof. However, since the application of the method is for an entertainment purpose the errors would only result in reduced acceptance of the system and not have any other negative impact.

Even these limitations can be improved upon using various methods – example: storage of reference human emotions images, continuous analysis of the mood swings of the user which can better help predict mood swings and help predict the actual current emotional state in a more accurate manner.

Similarly, a training module for the Song Shuffler's emotion recognition engine could also be planned that could be used by the user to train the engine with the user's major emotions – that is allow the system to capture reference images of the user's happy, sad, ecstatic, agitated, angry faces. These reference images could in turn be used an input for the real time emotion recognition engine to accurately identify the user's current emotion.

Further research of the methods involved is required before the discussed system can be developed. Similarly the feedback mechanism that uses the human emotions for dynamic, implicit feedback recording could also be impacted by the limitations of the FACS emotion analysis method. This in turn can be overcome temporarily and efficiently by integrating with multiple music review providers and also by providing an option to the user to manually record the feedback for a given music track.

To conclude, human-computer emotional interaction is one of the most important step ahead in the computing world and such practical implementation of the human-computer emotional interaction methods would only aid the further development of this field.

To quote da Vinci metaphorically, "For once you have tasted flight you will walk the earth with your eyes turned skywards, for there you have been and there you will long to return"

IV. REFERENCES

[1]. *Hjorztsjö, CH (1969). Man's face and mimic language*. free download: Carl-Herman Hjortsjö, Man's face and mimic language"

[2]. P. Ekman and W. Friesen. Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, Palo Alto, 1978.

[3]. Paul Ekman, Wallace V. Friesen, and Joseph C. Hager. Facial Action Coding System: The Manual on CD ROM. A Human Face, Salt Lake City, 2002.

[4]. *Hamm, J.; Kohler, C. G.; Gur, R. C.; Verma, R. (2011). "Automated Facial Action Coding System for dynamic analysis of facial expressions in neuropsychiatric disorders". Journal of Neuroscience Methods.* ***200*** *(2): 237–256. doi:10.1016/j.jneumeth.2011.06.023. PMC 3402717 . PMID 21741407*

