International Journal of Engineering and Technical Research (IJETR)
ISSN: 2321-0869, Volume-3, Issue-4, April 2015

Music Player Based on Facial Expressions

Rambeer Singh, Sumit Kumar Sheorain, Ajit Singh,Ved Prakash Pandey, Prof. N.K.Bansode

***Abstract*— The most expressive way humans display emotions is through facial expressions.Facial Expression Recognition is one of the important task in the machine learning. This emerging field has been research interest for scientists for several different scholastic tracks, i.e. computer science, engineering, psychology and neuroscience. The development of an automated system that accomplishes this task is rather difficult. Various techniques are being developed to perform this task but the biggest challenge is to show accuracy in detecting the expressions of the face. So in this article, a similar system is proposed to tackle this issue and using system to play music based on the facial expression. For this system Gabor filters is applied to the available datasets. More specifically, the proposed framework plays the music based on facial expression captured in the WebCam.**

***Index Terms*— Designed Framework, Method-Analysis, Performance.**

I. INTRODUCTION

Facial expression is one of the most natural and powerful way that human use to communicate their emotions. Automated facial expression analyser is very useful for various vision systems, speech processing, airport security and access control, intelligent human machine interaction etc.

We are applying the result of the facial expression for playing music according to the mood of the person capturing the input from the WebCam. This system basically be applied to 7 expressions of the human face (i.e. happy, sadness, disgust, fear, surprize and neutral). To built this system investigation are carried out on various facial expression recognition engines, feature selection techniques and machine learning methods. Compression on various platforms is also carried out such as MATLAB and OpenCV.

For this system selecting subset of Gabor filter[1][3] using AdaBoost[1][2] and training SVM[1] (Support Vector Machines) on the output of the filter is found particularly promising in OpenCV. The speed and accuracy of the system is increased using the combination of AdaBoost and SVM.

The system is fully automated and works at a high level of accuracy (about 93% to new subjects on 7 forced choices).

Manuscript received March 30, 2015.

Rambeer Singh, Department of Computer Engineering, Army Institute of Technology, Savitribai hule Pune University, Pune – Maharashtra – 411015

Sumit Kumar Sheorain, Department of Computer Engineering, Army Institute of Technology, Savitribai hule Pune University, Pune – Maharashtra – 411015

Ajit Singh, Department of Computer Engineering, Army Institute of Technology, Savitribai hule Pune University, Pune – Maharashtra – 411015

Ved Prakash Pandey, Department of Computer Engineering, Army Institute of Technology, Savitribai hule Pune University, Pune – Maharashtra – 411015

Prof. N.K.Bansode, Department of Computer Engineering, Army Institute of Technology, Savitribai hule Pune University, Pune – Maharashtra – 411015

This paper is structured as follow. Section II presents related research work. Section III provides a description of the components of the proposed system. Section IV describes the system design that is to be built. Section V describes implementation of system in a programming environment. Details about system dataset are described in Section VI and expected results are described in in Section VII. Section VIII summarizes the conclusions and Section IX summarizes the future work.

II. LITERATURE SURVEY

For developing a new application, first we need to study existing methods which are performing same task. Some of the similar Tools that we have referred for our work are as follow:

Bayesian Network

In this method facial expression is generated by activation of facial muscles. The visual results of muscle activation are changing contours of the mouth, eye and eyebrow. We can also observe the change in texture and position of wrinkles on face.

For studying facial movements define special Region Of Interests (ROIs) in that one muscle activity and corresponding AU is limited to one ROI but more than

one muscle can be active in one ROI. For movement in facial expression movement of pixels in consecutive frames is considered.

This method has various shortcomings such as selection of parameters as input of our Bayesian network[4] is a complex process, choice of ROI is arbitrary. The average classification rate for this method is between 80-90%.

Hidden Marcov Model

This system utilizes facial animation parameters (FAPs), supported by the MPEG-4 standard, as feature for facial expression classification. HMM can be applied in the single stream and multi stream. The FAPs describe the movement of the outer lip counters and eyebrows.

The stream weights are determined based on the facial expression recognition results obtained when this is applied individually.

Based on FAPs the overall expression recognition performance for outer lips is 87%and that of eyebrows is 58%.

Census Transformation

It is a geometrical feature based approach. The face geometry is extracted using modified active shape model. Each part of the face geometry is effectively represented by the Census Transformation (CT) based feature histogram. The facial expression is classified by the SVM classifier with exponential chi-square weighted merging kernel. In this method face consists of 68 landmark point and additional points on the forehead region of the face.

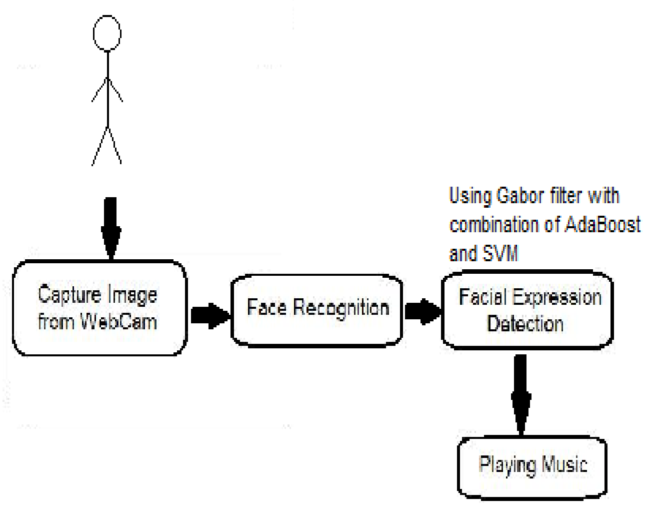
This method also uses the database of the facial expression such as JAFFE database.
This method achieves the accuracy of 83% using the database.

Gabor Filter using AdaBoost and SVM
This method is shown from the next section of the article.

III. PROPOSED MODEL

Model proposed in this article is a music player based on the mood of the user. This model is implemented in 4 stages:

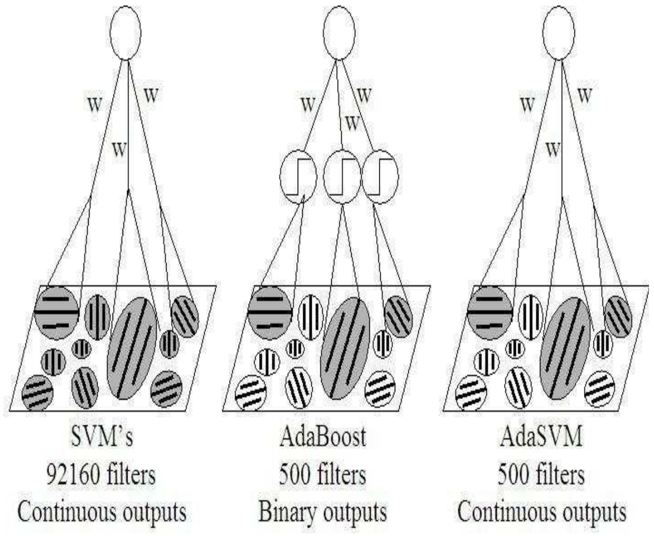
1. **Capture Image**
   In this stage image of the user is captured using the WebCam and that image is used for detecting the mood of the user. There are various challenges faced in the image such as brightness, different shades in the image, clarity, etc.
2. **Face Recognition**
   In this stage captured image is processed by the system to get the face from the complete image.
3. **Facial Expression Detection**
   In this stage expression of the detected face are recognised into one of the 7 expressions (i.e. happy, sad, fear, surprize, angry, disgust and neutral) using the Gabor filter method using the combination of the AdaBoost and the SVM.
4. **Playing Music**
   In this last stage the recognised expression is used for playing music in the player.



IV. SYSTEM DESIGN

The system use various techniques for facial expression recognition such as Gabor Filter, AdaBoost, SVM, Facial Action Coding Sequence (FACS). We basically uses combining feature selection of the AdaBoost and the classification by the SVM. AdaBoost is not only fast classifier but is also a feature selection technique. Features that are selected contingent on features that are already selected is the biggest advantage of the AdaBoost.

In feature selection by Adaboost, each Gabor filter is a treated as a weak classifier. Adaboost picks the best of those classifiers, and then boosts the weights on the examples to weight the errors more. The next filter is selected as the one that gives the best performance on the errors of the previous filter. At each step, the chosen filter can be shown to be uncorrelated with the output of the previous filters.



In the above figure SVM's learn weights for the continuous outputs of all 92160 Gabor filters. AdaBoost selects a subset of features and learns weights for the threshold outputs of those filters. AdaSVM's learn weights for the continuous outputs of the selected filters.

We explored training SVM classifiers on the features selected by Adaboost. When the SVM's were trained on the threshold outputs of the selected Gabor features, they performed no etter than AdaBoost. However, we trained SVM's on the continuous outputs of the selected filters. We informally call these combined classifiers AdaSVM. AdaSVM's outperformed straight AdaBoost by 3.8 percent points, a difference that was statistically significant ($z=1.99$, $p=0.02$). AdaSVM's outperformed SVM's by an average of 2.7 percent points, an improvement that was marginally significant ($z = 1.55$, $p = 0.06$).

The Gabor features selected by AdaBoost provide one indication of the spatial frequencies that are important for this task. Examination of frequency distribution suggested that a wider range of spatial frequencies, particularly in the high spatial frequencies, could potentially improve performance. Indeed, by increasing from 5 to 9 spatial frequencies (2:32 pixels per cycle at 0.5 octave steps), performance of the AdaSVM improved to 93.3% correct. At this spatial frequency range, the performance advantage of AdaSVM's was greater. AdaSVM's outperformed both AdaBoost ($z=2.1$, $p=.02$) and SVM's ($z=2.6$, $p<.01$). Moreover, as the input size increases, the speed advantage of AdaSVM's becomes even more apparent. The full Gabor representation was 7 times larger than before, whereas the number of Gabor selected by Adaboost only increased by a factor of 1.7. The result of 93% accuracy for a user-independent 7-alternative forced choice.

V. SYSTEM IMPLEMENTATION

Environmental setup required for implementing proposed system includes Ubuntu as platform, OpenCV as programming language.

Dependencies on various languages are as shown below:

- CMake >=2.8
- Python >=2.7, <3.0
- OpenCV >=2.4.5

The procedure to compile on Linux platform is:

1. mkdir build
2. cd build
3. cmake.. ; make ; make install - now the asset folder should be populated

For compilation of system in windows following steps are followed:

- Using CMake or CMakeGUI, select emotime as source folder and configure.
- If it complains about setting the variable OpenCV_DIR set it to the appropriate path so that:
  - C:/path/to/opencv/dir/ contains the libraries (*.lib)
  - C:/path/to/opencv/dir/include contains the include directories (OpenCV)
  - If the include directory is missing the project will likely not be able to compile due to missing reference to OpenCV or similar.
- Then generate the project and compile it.
- This was tested with Visual Studio 12 64 bit.

Usage of WebCam to capture image using CAM gui is:

- ./emotimegui_cli FACEDETECTORXML (EYEDETECTORXML|none) WIDTH HEIGHT NWIDTHS NLAMBDAS NTHETAS (svm|ada) (TRAINEDCLASSIFIERSXML)+

For successful compilation, system is to be trained using following steps:
After mkdir build; cd build; cmake ..; make; make Install go to the assets folder and:

- Initialize a dataset using
- Then fill it with your images or use the Cohn-Kanade importing script
- Now you are ready to train models

VI. SYSTEM DATASET

The Cohn-Kanade database is one of the most used faces database. Its extended version (CK+) contains also FACS code labels (i.e. Action Units) and emotion labels (neutral, anger, disgust, fear, happy, sadness, surprise). This dataset consists of 100 university students ranging in age from 18 to 30 years. 65% were female, 15% were African-American, and 3% were Asian or Latino. Videos were recorded in analog S-video using a camera located directly in front of the subject. Subjects were instructed by an experimenter to perform a series of 23 facial expressions. Subjects began and ended each display with a neutral face. Before performing each display, an experimenter described and modeled the desired display. Image sequences from neutral to target display were digitized into 640 by 480 pixel arrays with 8-bit precision for gray scale values.

For our study, we selected the 313 sequences from the dataset that were labeled as one of the 7 basic emotions. The sequences came from 90 subjects, with 1 to 7 emotions per subject. The first and last frames (neutral and peak) were used as training images and for testing generalization to new subjects, for a total of 625 examples. The trained classifiers were later applied to the entire sequence.

VII. EXPECTED RESULT

Proposed model is expected to detect the facial expression and play music as per the mood shown by the expression. The results of applying the Cohn-Kanade database to the system are as follows:

Resultant expression / Expected expression	Sadness	Neutral	Disgust	Angry	Surprise	Fear	Happy
Sadness	67%	-	-	17%	17%	-	-
Neutral	-	90%	-	3%	1%	2%	-
Disgust	-	-	100%	-	-	-	-
Angry	-	36%	9%	45%	-	-	-
Surprise	-	6%	-	-	94%	-	-
Fear	-	-	-	-	17%	67%	17%
Happy	-	-	-	-	-	-	100%

.

VI. CONCLUSION

In this article, we proposed a system for playing music based on the facial expression. We presented a systematic comparison of machine learning methods applied to the problem of fully automatic recognition of facial expressions, including AdaBoost and support vector machines. Best results were obtained by selecting a subset of Gabor filters using AdaBoost and then training Support Vector Machines on the outputs of the filters selected by AdaBoost. The combination of Adaboost and SVM's enhanced both speed and accuracy of the system. The generalization performance to new subjects for a 7-way forced choice was 93.3% and 97% correct on two publicly available datasets, the best performance reported so far on these datasets.

VII. FUTURE WORK

Initially system proposed is Linux based System and is accurately applicable for the limited dataset. Further it can be extended as platform independent application so that can be used in heterogeneous environment. Further this system can be enhanced to recognise minute change in the facial expressions and this system can also be extended for real time system. The dataset for this system is to be expanded and its accuracy is to be improved. This system is aligned to face in 2D plane and it can be extended to work for the faces in 3D plane. We are presently exploring applications of this system.

REFERENCES

[1] Marian Stewart Bartlett, Gwen Littlewort, Mark Frank, Claudia Lainscsek, Ian Fasel, Javier Movellan. Article on Fully Automatic Facial Action Recognition in Spontaneous Behavior.

[2] Ra ul Rojas. Article on AdaBoost and the Super Bowl of Classiers A Tutorial Introduction to Adaptive Boosting.

[3] Javier R. Movellan. Tutorial on Gabor Filters

Kwang-Eun Ko, Kwee-Bo Sim. 2010 International Conference article on Development of a Facial Emotion Recognition Method based on combining AAM with DBN