



International Journal of Advance Engineering and Research Development

Special Issue on Recent Trends in Data Engineering

Volume 4, Special Issue 5, Dec.-2017

Emotion Based Music Player With The Help Of SIFT And k-NN Algorithms

Prof M. P Nerkar, Neha Chhajed¹,Surykant Wankhade²

¹Computer Department, AISSMS Institute of Information Technology

²Computer Department, AISSMS Institute of Information Technology

Abstract —This project puts forth a framework for real time face and voice recognition and related emotion detection system based on facial features, their actions and intensity of voice. The key elements of Face are considered for prediction of face emotions of the user. The differences in each facial feature which are in-variant to scaling as well as rotation are used to determine the different emotions of face. Machine learning algorithms are used for recognition and classification of different classes of face emotions by providing training of different set of images. In this context, by implementing herein algorithms would contribute in several areas of identification and many real-world problems. The proposed algorithm is implemented using open source computer vision (OpenCV) and Machine learning. We have given a set of images and sounds each image containing multiple faces and is associated with some names in the associated caption. The aim of naming faces is to detect right emotions for each face where as in case of sound we sound meter is used to take input and then classification the input is done with the help of various classification algorithms.

Keywords—Open Source Computer Vision(OpenCV);Machine Learning; object recognition; Sound Meter;

II. INTRODUCTION

When machine analysis was commenced on human cases, it noticed great progress. Facial expression and speech are two main aspects of human emotional expressions, since people mostly rely on facial expression and speech to understand some one's present state. In recent years, recognizing human emotions from facial expression and speech i.e., audio visual emotion recognition has increased, attracting extensive attention towards artificial intelligence. As Audio Visual emotion recognition is motivated for establishing a reliable and less complex relationship between humans and computers, this increases its importance to human computer interaction(HCI). This research mainly focuses to ease human work by increasing interaction between humans and computers which in-turn helps to increase use of computers in day-to-day work. As computers have become an important asset of our speedy life, need for meaningful and easy communication between computers and humans has also increased. In addition to this, speech recognition is also successfully established area of research, but its main limitation is that it cannot respond appropriately to emotions of different people. To overcome this drawback, many computers are developed that are capable to detect, understand and reply to multiple emotional states of various people similar to how human being does. Hence, this audio-emotion recognition (AER) is a latest field of study which is providing a great advancement in the field of Human-Computer Interaction(HCI).

II. SURVEY DETAILS

1. Robust Object Detection and Tracking Using Sift Algorithm [1]

This paper states that, the SIFT algorithm is now-a-days more successful algorithm that recognizes multiple objects as same objects even if the object is complex or hidden in any of the objects or it has altered its location, even if the object is observed from fairly distinct angles. Hence in this paper they transformed an image into a set of local feature vectors. Hence, they came to a conclusion that SIFT algorithm is fairly successful to identify the same emotion for an object which is translated, scaled and rotation but is partially accurate to change in illumination. They have included a counter, with the help of which accuracy percentage of algorithm can be tracked. Hence, with the help of such algorithm there can be great improvement in computer vision.

2. Novel Approach to Face Expression Analysis in Determining Emotional Valence and Intensity with Benefit for Human Space Flight Studies (2015) [10]

This research presents a version of a simple algorithm for face expression analysis which targets to determine emotional valence and intensity with the help of less number of points, i.e eight Facial Characteristic Points and two types of textures in the face which results in better processing time. This is used as an alternative for statistical pattern recognition which is used to characterize the many software. Hence, it is concluded that the prospective implementation of herein algorithm for facial expression analysis would be an important add-on to the existent solutions. This solution can be used to perform evaluation, selection, training and recovery of individuals in complex professions, especially, of astronauts before, during and after prolonged human space flights.

3. A New Signal Classification Technique by Means of Genetic Algorithms and k-NN. [2]
This paper states that the k-NN algorithm is reliable and can be used because it does not need to be trained. The k-NN algorithm is a supervised i.e a set of example patterns and their respective classes is needed and non-parametric i.e it does not require prior training of data, classification algorithm. With the help of k-NN algorithm, pattern can be simply classified by comparing input pattern with the example set one by one serially, until each of example set present in database is tried and compared. Out of the example sets, the items that are chosen at the end are those whose distance to the pattern which is to be classified is small. And at the end the k-patterns which has the shortest distance are selected. Similarly, the class to which the pattern belongs is recognized as the class of the k patterns which is closer to the input pattern. In this paper they have computed the computational cost for k-NN algorithm when each new pattern is classified, as comparison is done with each of the example items present inside database.
4. Overview of Face Recognition System Challenges [12]
In this paper the challenges that are likely to occur while searching for an object match within the database are thoroughly discussed. For developing a computational model that recognizes and classifies objects is bit difficult. The stated that the main reason of difficulty is that different faces are complex and multidimensional and each of their expression has meaningful visual stimuli. There are many algorithms that can solve this problem. Hence, it is concluded that the magic of matching real-time objects with the database faces sometimes does not work even though many face recognition algorithms with high success rates are developed. Software used for Face recognition fails at times even though the suspect images were present in public and private database entries. Hence, to get a successful match of suspect, our database should have some process of training, because most of algorithms which are developed are non-parametric, which can be major drawback of the developed system
5. Emotion recognition from audiovisual information [8]
In this paper, they have tried to investigate the system by integrating both the audio as well as visual information for machine which in turn will help to recognize different expressions which might or might not be the real internal emotion of the suspect. They have assumed that the user has knowingly shown his/her emotions with the help of his/her facial and vocal expression in order to communicate successfully with the computer. Hence this helps to analyze human's ability to recognize some basic emotions by evaluating subjectively. Hence, they concluded that machines can perform better or we can say they work like humans with little limitations. They also discovered feature clustering of computer which were similar to the confusions of human beings. Further, with the help of these confusions they concluded that these two modalities complement each other. This contrasting relation between humans and computers made it possible to gain high accuracy for recognizing emotion than either of modality alone.
6. Comparing Accuracy of K-Nearest-Neighbor and Support-Vector-Machines for Age Estimation [3].
Comparing Accuracy of K-Nearest-Neighbor and Support-Vector-Machines for Age Estimation. In this paper two algorithms are compared in terms of accuracy. Support Vector Machine (SVM) and K Nearest Neighbor (KNN). Hence, they concluded that Support vector machine is classification and regression analysis method which can be used for measurements and software engineering with the help of supervised learning. Hence supervised methods are used for analyzing data and recognizing various patterns. In the same way k-nearest neighbor algorithm is also a classification algorithm but it performs classification data using training data.
 - Advantages and dis-advantages of SVM –
 - 1) SVM performs well on datasets that have huge number of attributes, even when there are just a limited cases that can be accessed by the training process.
 - 2) SVM classification has limitations in terms of speed and size during both training and testing phase of the algorithm.
 - Advantages and dis-advantages of KNN
 - 1) Even though the target class is multi-modal, the algorithm leads to best precision.
 - 2) Major stand-back of the KNN algorithm is that it utilise every feature similarly in computing a part of processing. Accuracy of k-NN is kept high in most of the cases. But as size of dataset increases, accuracy of SVM as well as KNN based system decreases.
 - 3) But in overall accuracy KNN works more effectively.
 - 4) As size of training dataset increases, time consumed for predicting values of KNN system increases dramatically. But in case of SVM time required for predicting values remains constant.

7. Expressions Invariant Face Recognition Using SURF and Gabor Features [6].

In this paper they have presented SURF features in order to improve face recognition through expressions invariance. They considered the facial points as key-points and with the help of detected points. They described detected features using SURF algorithm. They also combined Gabor filtering with the SURF features and developed a system which is orientation invariant. They concluded that system works well with Gabor features and SURF in combination and proposed work can be further extended multiple object detection. Hence with the help of SURF, asymmetric Facial feature can be easily extracted based on the run time information.

8. Facial Expression Based Music Player [13].

In this paper they have presented a system that recognizes facial expression and then plays songs according to the expression that is detected. They have used PCA approach for extracting facial features. In combination to this, Euclidean distance classifier has also been used which classifies these expressions and then plays the song according to the detected emotion. In this work, they have used real images i.e. user dependent images as an input through in-built laptop camera. They have taken their own training datasets from 4 different individuals and then input image was compared with training datasets and then with the help of Euclidean distance classifier the dataset with smallest distance was the output emotion of that person. The final result showed the accuracy level obtained with the help of PCA approach is best

9. SIFT Vs SURF

Parameters	SIFT	SURF
<u>Accuracy</u>	Highly accurate	Moderate
<u>Variances</u>	Key-points remains unchanged even if the input image is scaled or rotated	Slight changes in input image can lead to change of key-points
<u>Computation complexity</u>	Heavy	Expensive
<u>Speed</u>	Slower than SURF	Faster

Though SIFT is slow while detecting emotion but as we focus on the complexity of system, SIFT is less complex because even if the input image is scaled or rotated SIFT does not recalculate its key-points and this tells SIFT is better to use when more samples of same kind of image is present.

10. K-NN Vs SVM

Parameters	k-Nearest Neighbour	Support Vector Machine
<u>Compatibility</u>	Compatible with linearly separate classes.	Not compatible with linearly separate classes.
<u>Robustness</u>	Highly robust with noisy input data.	Can't classify audio in a noisy environment
<u>Accuracy</u>	High when training data is less, but accuracy decreases as the training data increases.	Remains constant even if the training data size increases
<u>Features</u>	Large number of features can be identified.	Less Feature can be identified

As our requirement is to develop a music player which has large training data set, so for identifying large number of features kNN(k-Nearest Neighbour) is accurate as compared to svm(Support Vector Machine).

IV. PROPOSED SYSTEM

The proposed system is tested on datasets which consists of wide range of face images with different expressions, poses, illumination conditions and genders. We used database for training of emotion detection system. Both databases are open source and our algorithm performed well on both datasets For face recognition, we have used a webcam for capturing of faces. The implemented algorithm is capable of recognizing different persons in a single window. If the recognition environment is under proper lighting condition and less background noises then the recognition rate will be high. And also recognized sound and play similar music

➤ Advantage of proposed system

- 1) Required less time when we have to detect emotion and detect sound
- 2) Accuracy is higher compared to other methods
- 3) The accuracy of both face recognition and emotion detection can be increased by increasing the number of images

IV. CONCLUSION AND FUTURE SCOPE

This method is secure, reliable and easy to use. Automated Face and sound Systems based on face recognition techniques thus proved to be time saving and secured. This system can also be used to identify an different emotions. The future work includes the implementation of the system in android improves the availability of the system to more users.

REFERENCES

- [1] Robust Object Detection and Tracking Using Sift Algorithm Shweta Yakkali*, Vishakha Nara, Neelam Tikone, Darshan Ingle SIESGST, Mumbai University Maharashtra, India
- [2]A New Signal Classification Technique by Means of Genetic Algorithms and kNN Daniel Rivero, Enrique Fernandez-Blanco, Julian Dorado, Alejandro Pazos Department of Information and Communications Technologies University of A Corua A Corua, Spain
- [3] Comparing Accuracy of K-Nearest-Neighbor and Support-Vector-Machines for Age Estimation Anchal Tomar *1, Anshika Nagpal *2 *1 M.tech student at M.I.ET College, Dr. A.P.J. Abdul Kalam Technical University (APJAKTU) Uttar Pradesh, India *2 Assistant Professor at M.I.E.T, APJAKT University, India
- [3]F. Dellaert, T. Polzin and A. Waibel, Recognizing Emotion in Speech, in Proc. International Conf. on Spoken Language Processing, Philadelphia, PA, USA, pp. 1970-1973, October 3-6, 1996.
- [4]Alexandre Alahi, Raphael Ortiz, Pierre Vanderghyest. FREAK: Fast Retina Keypoint. In IEEE Conference on Computer Vision and Pattern Recognition, 2012.
- [5]D.G. Lowe, Object Recognition from Local Scale-Invariant Features , Proc. Seventh Intl Conf. Computer Vision, pp. 1150-1157, 1999
- [6]Expressions Invariant Face Recognition Using SURF and Gabor Features Barun Kumar Bairagi ARC India Pvt. Ltd. Kolkata, India .
- [7]Z. Zeng and T.S. Huang are with the Beckman Institute, University of Illinois at Urbana-Champaign, 405 N. Mathews Ave., Urbana, 61801.
- [8]Li Wern Chew, Kah Phooi Seng, Li-Minn Ang, Vish Ramakonar, and Amalan Gnanasegaran, Audio-Emotion Recognition System using Parallel Classifiers and Audio Feature Analyzer, in 2011 3rd Int. Conf. on Computational Intell., Modelling Simulation, 2011, pp. 210-215.
- [9]M.R. Ahsan, M.I. Ibrahimy, and O.O. Khalifa, EMG Signal Classification for Human Computer Interaction: A Review, European Journal for Scientific Research, vol. 33, no. 3, 2009, pp. 480-501.
- [10]Adrian Dinculescu, Cristian Vizitiu, Alexandru Nistorescu, Mihaela Marin, Andreea Vizitiu, Novel Approach to Face Expression Analysis in Determining Emotional Valence and Intensity with Benefit for Human Space Flight Studies, 5th IEEE International Conference on E-Health and Bioengineering - EHB, pp:1 - 4, November 2015.

- [11]Rajesh K M, Naveenkumar M, An Adaptive-Profile Modified Active Shape Model for Automatic Landmark Annotation Using Open CV, International Journal of Engineering Research in Electronic and Communication Engineering (IJERECE), Vol.3, Issue.5, pp:18-21, May 2016.
- [12]Ambika Ramchandra , Ravindra Kumar, Overview Of Face Recognition System Challenges, International journal of scientific technology research , Vol.2,pp: 234-236, August 2013.
- [13]Facial Expression Based Music Player Sushmita G. Kamble and Asso. Prof. A. H. Kulkarni Department Of Computer Science and Engineering KLSGIT Belagavi, Karnataka, India