ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 12, December 2015

Literature Review of Online Facial Expression Recognition Systems

Kiran P. Suryawanshi[1], Sonali Rangdale[2]

ME Student, Dept. of I.T., Siddhant College of Engineering, Savitribai Phule Pune University,Pune, India[1]

Professor, Dept. of I.T., Siddhant College of Engineering, Savitribai Phule Pune University,Pune, India[2]

ABSTRACT: The paper is intended to present an overview of literature in particularly focusing over a machine learning aspect of facial expression recognition. The facial expression identification finds its use in all the scenarios where human-computer interface plays the central role. Moreover, the market is interested in knowing buyers response from its facial expressions. To add further, online interactive machine chat would gain an edge if the machine or computer system can detect the emotions of user person so that it can serve it better. There had been quite a considerable amount of research happening in this field from late nineties and continuing till date. The paper overlooks the major efforts, algorithms, techniques established so far for the purpose. Going in detail, the paper points some of well perceived efforts of researchers published literature in early 21st century. Researchers to include a test dataset of sever distinct facial expressions and observed a quite successful recognition rate of their component based expression classifier algorithm. This paper presents detailed ideas behind these articles, their algorithms, challenges. Additionally, an effort has been made to summarize the general advantages and drawbacks of the reported methodologies so far.

KEYWORDS: Facial Expression Identifier, Seven Universal Expression, Support Vector Machine, Local Binary Pattern, Component Based Approach

I. INTRODUCTION

Humans are blessed by the nature with the art of spoken language communication which rates them higher among all the animals in the entire bio system. However, the communication would not have meaning if it is not supported by human capacity of facial expression recognition. Facial expressions are a form of nonverbal communication. The facial expressions convey the emotions of the persons participating in a communication exercise. In the modern world which is getting more and more tech savvy, computers have made their place in human lives. The one of aim of embedded software technologies is to make human computer interaction smoother than and perhaps as easy as human-human interface.

Automatic identifying the human facial expressions is the key of computer human interface success. Numerous benefits of such an effective interface would include to betterment of computers and robots to serve better to humans. Not only that but at a day to day level, where humans are involved in auto chats with computing systems, would find it very useful . Today's age of online market, where sellers contact to human customers through computer interface over the internet, it would be of huge gain for sellers to know human customer emotions by identifying the facial expressions.

A facial expression is one or more motions or positions of the muscles beneath the skin of the face. These movements convey the emotional state of an individual to observers. Facial expression recognition is often an emotional experience for the brain and the amygdala is highly involved in the recognition process. The eyes are often viewed as important features of facial expressions. Aspects such as blinking rate can be used to indicate whether or not a person is nervous or whether or not he or she is lying. Also, eye contact is considered an important aspect of interpersonal communication. Similar to eyes, the brows, lips, and chic muscles show key movements during an expression. As computers technology is getting mature day over day, its percolation in human life is advancing as well. Machine learning is one of important section where, researchers striving their best to improve on human-computer interface. The budding ability of computers to detect human facial expression and there behind emotion is one of the primary importance. And so is its improvisation attracts researchers and programmers.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 12, December 2015

Paul Ekman [1], an American psychologist was the pioneer behind identifying and classifying various human facial expressions. His contribution to creating the humongous database of human expressions proved to be an asset for many modern machine learning researchers and scientists.

Fig.1 The Seven Basic Emotions and their Universal Expressions [2]

As per Ekman's historical reports, human facial expressions have been classified In 7 key expressions. Ekman termed these set of seven facial expressions as "Seven Universal Facial Expressions of Emotion". Thereby, any automatic computer software system for facial expression classifier is expected to clearly distinguish between these seven expressions for a given data over a neutral face. These seven expressions are listed as displayed in referred image in Fig. 1.

An expression of contempt is very tricky as it is as narrated, by Ekman, to have similarities between neutral, sad and anger. Basically, a contempt expression involves an asymmetric movement of upper lips. Most of the research in this area is based on measuring the facial muscle movements. A lot early era researchers started by Ekman's idea, to mark the pivotal points on the neutral face and then measuring the distances of such pivotal points over the face as an expression comes up.

The present paper layout is as follows. Post Introduction, section II focuses on summarizing the key contributions in literature for facial expression identification. Also forming the core part of this paper, section II tries to summarize some established algorithms from literature. Section III provides additional discussion from author's point of view on applications of facial expression recognition, and advantages and drawbacks of reviews systems.

II. RELATED WORK

This section would bring forth the key literature milestones in the field facial expressions and their recognition. Here onwards in this section, all of the surveyed literature is arranged in sub-topic of interest fashion, to be more helpful for reader.

A. *Emotions and Their Relation to Facial Expressions: From a Psychologist's View*

Dr. Ekman has been getting a lot of attention lately, due to the fact that he is the scientific consultant for the new show on Fox television named "Lie to Me". The show is even based off of his science. In his book "Emotions Revealed" [3], Ekman discusses how a person's face can be "read" to determine what kind of emotions s/he is feeling. The author then proceeds to focus on emotions such as contempt, disgust, sadness, happiness, and anger. In each chapter he has a model that shows different expressions. He explains in great detail how to read the facial expressions as well as what they seem to mean. He also has an exercise that people can do to use facial expressions to invoke feelings. Overall, it's a fascinating read, which shows how much the face is integral to feeling emotion as well as expressing it. At times, the book is dry and can be a bit of a slog to read through, but Ekman does a fairly comprehensive job of explaining the subject. I'm already eager to see how I can apply the concepts in my everyday communication.

In Chapter 16 of the famous "Handbook of Cognition and Emotion" [4], Ekman had presented a detailed study of challenges and evidences which dwell around the idea to find a relation between expression and emotion. The study

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 12, December 2015

continues to search if same expression can denote multiple emotions of same emotion can be expressed with multiple expressions. Also if the fabricated expression differs from a spontaneous expression, moreover, do similar expressions mean different emotions in different cultures?

The Facial Action Coding System (FACS) is a method for measuring facial expressions in terms of activity in the underlying facial muscles [5]. FACS is a system to taxonomies human facial movements by their appearance on the face, based on a system as adopted by Ekman et al [5]. It is an index of facial expressions, but does not actually provide any bio-mechanical information about the degree of muscle activation. One has to understand that muscle activation is not part of FACS. Movements of nostrils, eyes, cheeks, lips, lip corners, head, neck, tongue, dimples, winks etc., in positive and negative directions are encoded to specific numbers. Also, sometimes, their intensity of movements is accounted. Thus each of the expression would obtain a specific code or a summation of codes, as illustrated in this article.

Last, but not the least from Ekman's works [6], in identifying the expression of contempt among many diverse cultures with two parallel experiments. They showed how the expression of contempt stands out as a distinctive expression than the other known expressions by the time such as anger, sad, happy, disgust, surprise and fear. The experiments were conducted so as to include specimen readers and database from American, Japanese, Indonesian and Sumatra cultures.

B. *Generic Facial Expression Identifiers*

Thereafter, many other automatic facial expression recognition systems stood up among literature. A noteworthy of them was Lanitis et al [7], Lyons and Akamatsu [8] and Wang and Yin [9].

Lanitis et al [7] presented a system which can be used for locating facial features, coding, and reconstruction, recovering 3D pose, recognizing gender and expression, and identifying the individual in an image. Their results for most of the applications were promising, even though considerable variation in 3D pose, lighting, and expression were allowed, demonstrating the potential for use in real life applications. The distinctive feature of their system is that it can cope successfully with almost all aspects of face image processing, within a unified framework. The face interpretation procedures described are fully automatic; errors for the classification experiments may be caused either by failure in locating landmarks accurately, or by failure of the classification algorithm. They do not distinguish between the two cases, since we believe that locating facial characteristics automatically is an important aspect of an integrated system.

Lyons and Akamatsu [8] presented a method for extracting information about facial expressions from images. Facial expression images were coded using a multi-orientation, multi-resolution set of Gabor filters which are topographically ordered and aligned approximately with the face. Gabor wavelets are wavelets using complex functions constructed to serve as a basis for Fourier transforms in information theory applications. They are also closely related to Gabor filters. The important property of the wavelet is that it minimizes the product of its standard deviations in the time and frequency domain. Put another way, the uncertainty in information carried by this wavelet is minimized. The similarity space derived from this representation is compared with one derived from semantic ratings of the images by human observers. The results show that it is possible to construct a facial expression classifier with Gabor coding of the facial images as the input stage. The Gabor representation showed a significant degree of psychological plausibility, a design feature which may be important for human-computer interfaces.

Wang and Yin [9] presented a topographic modelling approach to recognize and analyse facial expression from single static images. The so-called topographic modelling was developed based on a novel facial expression descriptor, Topographic Context (TC), for representing and recognizing facial expressions. This proposed approach applies topographic analysis that treats the image as a 3D surface and labels each pixel by its terrain features. Topographic context captures the distribution of terrain labels in the expressive regions of a face. It characterizes the distinct facial expression while conserving abundant expression information and disregarding most individual characteristics. Experiments on person-dependent and person-independent facial expression recognition using two public databases show that TC is a good feature representation for recognizing basic prototypic expressions. The experimental results show that system achieved the best correct rate at 82.61% for the person-independent facial expression recognition.

C. *Facial Expression Identifiers using Support Vector Machines (SVM) and Local Binary Patterns (LBP)*

Support vector machine (SVM) are the commonly used tool for effective classification of facial expressions. In machine learning, SVMs, are supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis. Given a set of training examples, each marked for

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 12, December 2015

belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other, making it a non-probabilistic binary linear classifier. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on. Most of the researchers who have used it include Chuang and Shih [10], Shan et al [11], Zhao and Pietikainen [12].

In their paper [10], Chuang and Shih focused on recognizing facial action units (AUs), which represent the subtle change of facial expressions. They adopted ICA (independent component analysis) as the feature extraction and representation method and SVM (support vector machine) as the pattern classifier. By comparing with three existing systems, such as Tian, Donato, and Bazzo, their proposed system could achieve the highest recognition rates. Furthermore, the proposed system was fast since it took only 1.8 ms for classifying a test image, as reported by them.

Shan et al.[11] empirically evaluated facial representation based on statistical local features, Local Binary Patterns, for person-independent facial expression recognition. Also, different machine learning methods were systematically examined on several databases. Here it is to be noted that they used extensive database by Kanade et al. [13]. Extensive experiments conducted by them illustrated that LBP features are effective and efficient for facial expression recognition. They further formulated Boosted-LBP to extract the most discriminant LBP features, and the best recognition performance was obtained by using Support Vector Machine classifiers with Boosted-LBP features. Moreover, they investigated LBP features for low-resolution facial expression recognition, which is a critical problem but seldom addressed in the existing work. They observed in their experiments that LBP features perform stably and robustly over a useful range of low resolutions of face images, and yield promising performance in compressed low-resolution video sequences captured in real-world environments.

A spatiotemporal local binary pattern operator from three orthogonal planes (LBP-TOP) was found to be proposed in literature for describing and recognizing dynamic textures and applied to facial expression recognition. Zhao and Pietikainen [12], in their work extended the LBP-TOP features to multi-resolution spatiotemporal space and used them for describing facial expressions. Moreover, AdaBoost was utilized to learn the principal appearance and motion, for selecting the most important expression-related features for all the classes, or between every pair of expressions. Finally, they also employed a support vector machine (SVM) classifier to the selected features for final recognition.

Very recently in 2013, Hong et al. [14], tried to include a test dataset of seven distinct facial expressions and observed a quite successful recognition rate of their component based expression classifier algorithm. The datasets used by Hong et al. are generated using the JACFEE and JACNEUF dataset, which is certified by the Paul Ekman Group LLC [15]. They specifically chose this dataset because it has discrete expression classes (the 7 universal facial expressions of emotion as well as neutral faces). The dataset consists of 280 colored face images, including 140 neutral faces and 20 images of 7 different facial expressions. They have employed four different image pre-processing techniques are used listed as –

1. Grayscale Transformation
2. Local Binary Patterns (LBP)
3. Edges using Canny and Sobel
4. Feature points

The facial expressions were classified using a standard multiclass Support Vector Machine (SVM) and a pairwise adaptive multiclass SVM (pa-SVM) – which used pairwise adaptive model parameters. The significance of their study was to provide a thorough understanding of the choices that impact a classifier's performance and the performance dynamics between expression pairs when the contemptuous expression was then considered. In particular, they extended the issues and justified the use of the contemptuous expression by noting its universal presence in human expression and its asymmetric nature, which was not shared with other expressions. Using their face model, which used feature points around the eyes, brows and mouth, they obtained a best correct classification rate of 98.57% with the contemptuous expression included.

The original LBP operator was introduced by Ojala et al. [16], and was proved a powerful means of texture description. Local Binary Pattern (LBP) is a simple yet very efficient texture operator which labels the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number. Due to its discriminative power and computational simplicity, LBP texture operator has become a popular approach in various

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 12, December 2015

applications. It can be seen as a unifying approach to the traditionally divergent statistical and structural models of texture analysis. Perhaps the most important property of the LBP operator in real-world applications is its robustness to monotonic gray-scale changes caused, for example, by illumination variations. Another important property is its computational simplicity, which makes it possible to analyze images in challenging real-time settings.

The basic idea for developing the LBP operator was that two-dimensional surface textures can be described by two complementary measures: local spatial patterns and grey scale contrast. The original LBP operator (Ojala et al. [16]) forms labels for the image pixels by thresholding the 3 x 3 neighborhood of each pixel with the center value and considering the result as a binary number. The histogram of these 28 = 256 different labels can then be used as a texture descriptor. This operator used jointly with a simple local contrast measure provided very good performance in unsupervised texture segmentation (Ojala and Pietikäinen 1999 [17]). After this, many related approaches have been developed for texture and color texture segmentation.

Another extension to the original operator is the definition of so-called uniform patterns, which can be used to reduce the length of the feature vector and implement a simple rotation-invariant descriptor. This extension was inspired by the fact that some binary patterns occur more commonly in texture images than others. A local binary pattern is called uniform if the binary pattern contains at most two bitwise transitions from 0 to 1 or vice versa when the bit pattern is traversed circularly. For example, the patterns 00000000 (0 transitions), 01110000 (2 transitions) and 11001111 (2 transitions) are uniform whereas the patterns 11001001 (4 transitions) and 01010010 (6 transitions) are not. In the computation of the LBP labels, uniform patterns are used so that there is a separate label for each uniform pattern and all the non-uniform patterns are labeled with a single label.

Ojala et al. [18] noticed in their experiments with texture images that uniform patterns account for a little less than 90% of all patterns when using the (8,1) neighborhood and for around 70% in the (16,2) neighborhood. Each bin (LBP code) can be regarded as a micro-texton. Local primitives which are codified by these bins include different types of curved edges, spots, flat areas etc.

III. Discussion

As described in introduction earlier, the facial expressions do convey a lot of non-verbal information of a person's emotion and moods. Thereby an automatic facial expression recognition system would find its applications in a lot places here human computer interface would exist. A few of them are listed here –

• For robots, to know what the emotions of human prescribing a service command are would mean, to serve him better.

•An automatic video feedback recording system, to know the facial expression and thereby identifying true and false feedbacks. Also to recognize the intensity of feedbacks.

•A.I. chat browser.

•Online marketing, to know and understand customer requirements and pain points.

•Facial identifiers replacing password security system

Face recognition is not perfect and struggles to perform under certain conditions. Researchers described one obstacle related to the viewing angle of the face: "face recognition has been getting pretty good at full frontal faces and 20 degrees off, but as soon as towards profile there have been problems".

Other conditions where face recognition will include poor lighting, sunglasses, long hairs, or other objects partially covering the subjects face and low resolution of images. Another serious disadvantage is that many systems are less effective if facial expressions vary. Even a big smile can render the system less effective. For instance Canada now allows only neutral facial expression in passport photos. At several occasions firms like Google Flikr and Nikon had been criticized for their software lack of ability to recognize faces with other skin colors than light.

As The limitations of identified with some of pre-established facial expression recognizing systems are as follows –

1. Image quality: Image quality affects how well facial-recognition algorithms work. The image quality of scanning video is quite low compared with that of a digital camera. Even high-definition video is, at best, 1080p (progressive scan); usually, it is 720p.

2.Image size: When a face-detection algorithm finds a face in an image or in a still from a video capture, the relative size of that face compared with the enrolled image size affects how well the face will be recognized. An already small image size, coupled with a target distant from the camera, means that the detected face is only 100 to 200 pixels on a side. Further, having to scan an image for varying face sizes is a processor-intensive activity. Most algorithms allow specification of a face-size range to help eliminate false positives on detection and speed up image processing.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 12, December 2015

3.Face angle: The relative angle of the target's face influences the recognition score profoundly. When a face is enrolled in the recognition software, usually multiple angles are used (profile, frontal and 45-degree are common). Anything less than a frontal view affects the algorithm's capability to generate a template for the face. The more direct the image (both enrolled and probe image) and the higher its resolution, the higher the score of any resulting matches.
4.Processing and storage: Even though high-definition video is quite low in resolution when compared with digital camera images, it still occupies significant amounts of disk space. Processing every frame of video is an enormous undertaking, so usually only a fraction (10 percent to 25 percent) is actually run through a recognition system. To minimize total processing time, agencies can use clusters of computers. However, adding computers involves considerable data transfer over a network, which can be bound by input-output restrictions, further limiting processing speed.

As However, among the different bio-metric techniques, facial recognition may not be the most reliable and efficient but, one key advantage is that it does not require the co-operation of the test subject to work. Properly designed systems installed in airports multiplexes and other public places can identify individuals among the crowd, without passer-by even being aware of the system. Other bio-metrics like finger prints, eye-lids scan and speech recognition cannot perform this kind of mass identification.

IV. CONCLUSION

The paper has presented a detailed literature survey of the facial expressions and facial expression recognition systems. With growing human-computer interactions, a system for expression recognition would only add to the superiority of the relation. Facial expression identifiers with facial muscle tagging to use of linear binary patterns and to state-of-the-art systems employing advanced machine learning algorithms such as Support vector machines had been presented in decent details. Moreover, the paper also tried to recollect the applicability of the system to various scenarios, describing its advantages and limitations with core thoughts.

REFERENCES

1. http://www.paulekman.com/paul-ekman/
2. David Matsumoto and Hyi Sung Hwang, "Reading facial expressions of emotion," Science Brief, American Psycological Association. May 2011. Available Online: http://www.apa.org/science/about/psa/2011/05/facial-expressions.aspx
3. Paul Ekman. Emotions Revealed. Times Books, Henry Holt and Company 115 West 18th Street New York, New York 10011, 2003.
4. Paul Ekman. Handbook of Cognition and Emotion, chapter 16, pages 301–320. New York: John Wiley and Sons Ltd, 1999.
5. P Ekman and W.V Friesen. Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, 1978.
6. P. Ekman and K. G. Heider. The universality of a contempt expression: A replication. Motivation and Emotion, 12(3):303–308, 1988.
7. Andreas Lanitis, Chris J. Taylor, and Timothy F. Cootes. Automatic interpretation and coding of face images using flexible models. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7):743–756, 1997.
8. Michael J. Lyons, Shigeru Akamatsu, Miyuki Kamachi, and Jiro Gyoba. Coding facial expressions with gabor wavelets. In Proceedings of Third IEEE International Conference on Automatic Face and Gesture Recognition, pages 200–205, 1998.
9. Jun Wang and Lijun Yin. Static topographic modeling for facial expression recognition and analysis. Computer Vision and Image Understanding, 108(1-2):19–34, 2007.
10. Chih-Chung Chang and Chih-Jen Lin. LIBSVM: a Library for Support Vector Machines, 2001.
11. Caifeng Shan, Shaogang Gong, and Peter W. McOwan. Facial expression recognition based on local binary patterns: A comprehensive study. Image and Vision Computing, 27(6):803–816, 2009.
12. Guoying Zhao and Matti Pietikainen. Boosted multi-resolution spatiotemporal descriptors for facial expression recognition. Pattern Recognition Letters, 30(12):1117–1127, 2009.
13. T. Kanade, J.F. Cohn, and Yingli Tian. Comprehensive database for facial expression analysis. In Proceedings of Fourth IEEE International Conference on Automatic Face and Gesture Recognition, pages 46–53, 2000.
14. K. Hong, S. Chalup and R. King "A Component Based Approach for Classifying the Seven Universal Facial Expressions of Emotion" of Emotion". IEEE Symposium on Computational Intelligence for Creativity and Affective Computing (1)4, Apr 2013.
15. Paul Ekman Group LLC. JACFEE and JACNEUF dataset. http://www.humintell.com/for-use-in-research/.
16. Ojala, T., Pietikäinen, M. and Harwood, D., A Comparative Study of Texture Measures with Classification Based on Feature Distributions. Pattern Recognition 19(3):51-59. 1996.
17. Ojala, T. and Pietikäinen, M., Unsupervised Texture Segmentation Using Feature Distributions. Pattern Recognition 32:477-486. 1999.
18. Ojala, T., Pietikäinen, M. and Mäenpää, T., Multiresolution Gray-scale and Rotation Invariant Texture Classification with Local Binary Patterns. IEEE Trans. Pattern Analysis and Machine Intelligence 24(7): 971-987 2002.