International Journal of Education and Science Research Review
ISSN 2348-6457
Volume-2, Issue-1 February- 2015
www.ijesrr.org Email- editor@ijesrr.org

A Review: Detection and Analysis of Facial Micro-Expression

Mr. Ravi Kumar Bhall
M.Tech.
Department of Computer Science Engineering
DIT University Dehradun
Uttarakhand

Mr. Vishal Sharma
Associate Professor
Department of Computer Science Engineering
DIT University Dehradun
Uttarakhand

ABSTRACT

In psychology facial expressions are used to analyze the behavioural aspects of a subject to know the suppressed feelings such as anger, sadness, happiness, contempt, surprise, disgust, and fear. These expressions are proven to be more successful in identifying the mood and the real intentions of the subject but the main problem with facial expressions is that they can be faked which Leeds to misjudging the subject.

Recent studies show that there are some leaked facial micro-expressions which occur for very small duration i.e. 1/25 to 1/3 second and can't be controlled thus can't be faked. These micro-expressions can only be detected by a high speed camera having high frame rate such as 100 fps or higher.

Facial micro-expressions were proven to be an important source of information for hostile intent and danger detection. The increase in violence and extreme terrorist activities around the globe urges for the better and faster technological solutions that can help to detect and prevent these actions. The fusion of computer technology and psychology research has a potential for developing such solution to the next level. This paper is a review of various techniques used in detecting facial micro-expressions.

Keywords – deception detection, facial micro-expression, micro- expression detection techniques

1. INTRODUCTION

After thirty years of research by Ekman, Frank and O'Sulliva [1] in addition to in depended group of Portet [2] micro-expressions were found to be an essential behavioural source for detecting deception and can be used for danger demeanour detection as well [1]. Facial micro-expression is a brief, involuntary expression shown on the face of humans when they are trying to conceal or repress an emotion. Micro-expressions usually occur in high-stakes situations, where people have something to lose or gain [3].

From the technical point of view the detection of facial micro-expressions is not an easy task using the traditional approaches. Their duration is 1/25 to 1/3 seconds, and they appear with low muscle intensity. The need for analyzing such momentary expressions requires a use of a high-speed camera.

2. METHODS USED FOR FACIAL MICRO-EXPRESSION DETECTION

The general approach towards automatic facial expressions detection system (micro-expressions as well) consists of three steps: (1) face acquisition, (2) facial data extraction and representation, and (3) facial expression recognition. Face acquisition (1) includes an automatic detection and tracking of the face in the input video. Extraction of the face direction could be added to this step.

In facial data extraction and representation for expression analysis, two main approaches exist: geometric feature-based methods and appearance-based methods.

The geometric facial features are presented by the shape and location of facial components (such as mouth, eyes, eyebrows, and nose). The facial components and facial feature points are extracted by some computer vision techniques that form a feature vector that represents the face geometry [5].

Superior research results were reported on **Active Appearance Model (AAM) by Kanade** [6] group. However, there are two disadvantages of AAM. First, this approach requires extensive dataset with large

amount of manually tagged points of the face. Second, the accuracy of facial feature tracking significantly decreases in the faces that were not included in the training set.
Another approach is based on direct tracking of 20 facial feature points (e.g. eye and mouth corner, eyebrow edges) by particle filter [7]. This approach delivers good results for some facial motions, but fails in detecting subtle motions, that can be detected only by observing skin surface. The performance of this and similar approaches strongly rely on the accuracy of the facial feature points tracking. In practice, facial feature points tracking algorithm cannot deliver the necessary accuracy for micro-expression recognition task.

In Appearance-Based Methods, image filter, such as **GABOR WAVELETS** are applied to either the entire face or specific regions in the face, to extract a feature vector. This method was applied for spontaneous facial motion analysis and considered to be the most popular [8]. However, this method is based on analyzing the video frame by frame, without considering correlation between frames. In addition, applying this approach for facial surface analysis requires large datasets for training an enormous number of filters.

Using 3D-Gradient Descriptor: In this method [4] 3D gradient oriented histogram is selected for describing the facial motions, due to its ability to capture the correlation between the frames. Maraszalek [9] summaries several works that showed a good results in classifying motions in video signal using 3D gradient descriptors. In Dollar work [10] local descriptors such as normalized pixel values, brightness gradients, and windowed optical flow were evaluated and compared for action recognition. Experiments on three datasets: facial expressions, mouse behaviour, and human activity, show best results for gradient descriptors [10]. Those descriptors, however, were computed by concatenating all gradient vectors in a region.

Using Spatio-Temporal Strain: In this method [11] for the automatic spotting (temporal segmentation) of facial expressions in long videos comprising of macro- and micro-expressions. The method utilizes the strain impacted on the facial skin due to the non- rigid motion caused during expressions. The strain magnitude is calculated using the central difference method over the robust and dense optical flow field observed in several regions (chin, mouth, cheek, and forehead) on each subject’s face. This approach is able to successfully detect and distinguish between large expressions (macro) and rapid and localized expressions (micro).
Using Gabor Filters & GentleSVM (Support Vector Machine): In this approach [12] for automatic micro-expression recognition The system is fully automatic and operates in frame by frame manner. It automatically locates the face and extracts the features by using Gabor filters. GentleSVM is then employed to identify micro-expressions. As for spotting, the system obtained 95.83% accuracy. As for recognition, the system showed 85.42% accuracy which was higher than the performance of trained human subjects.
It has been shown that **SVMs** can provide state-of-the-art accuracy for the facial expression recognition problem [12] [13]. To further improve the performance of our proposed algorithm, we chose to use the Gentleboost algorithm as a feature selector preceding the SVM classifiers. This combination is called GentleSVM. In feature selection by Gentleboost, each Gabor filter is treated as a weak classifier. Gentleboost picks the best of those classifiers, and then adjusts the weight according to the classification results. The next filter is selected as the one that gives the best performance on the errors of the previous filter. All the selected features are then united to form a new representation. Then the 1-norm soft margin linear SVMs are trained on the selected Gabor features. For multiclass classification, the one-against-all method is implemented. The expression category decision is implemented by choosing the classifier with the maximum geometric margin for the test example.

Using Temporal Interpolation Model: This model [14] shows how to use graph embedding to interpolate images at arbitrary positions within a micro-expression. This allows inputting a sufficient number of frames to the feature descriptor even for very short expressions with very small number of frames. It also enables to achieve more statistically stable feature extraction results by increasing the number of frames used for extraction. Zhou et al. [15] previously proposed a similar method for synthesising a talking mouth.

Using Discriminant Tensor Subspace Analysis (DTSA) Plus Extreme Learning Machine (ELM): DTSA [16] treats a gray facial image as a second order tensor and adopts two-sided transformations to reduce dimensionality. One of the many advantages of DTSA is its ability to preserve the spatial structure information of the images. In order to deal with micro-expression video clips, DTSA is extending to a high-order tensor. Discriminative features are generated using DTSA to further enhance the classification performance of ELM classifier. Another notable contribution of the proposed method includes significant improvements in face and micro-expression recognition accuracy.

Extreme learning machine (ELM) as a new learning algorithm for single layer feed forward neural networks (SLFNs) as shown in Fig. 2, was first introduced by Huang el al. [17][18]. ELM seeks to overcome the challenging issues faced with the traditional SLFNs learning algorithms such as slow learning speed, trivial parameter tuning and poor generalization capability. ELM has demonstrated great potential in handling classification and regression tasks with excellent generalization performance. The learning speed of ELM is much faster than conventional gradient based iterative learning algorithms of SLFNs like back propagation algorithm while obtaining better generalization performance. ELM has several significant features [18] which distinguish itself from the traditional learning algorithms of SLFNs:

i. ELM is easily and effectively used by avoiding tedious and time-consuming parameter tuning.
ii. ELM has extremely fast learning speed.
iii. LM has much better generalization performance than the gradient based iterative learning algorithms in most cases.
iv. ELM is much simpler and without being involved in local minima and over-fitting.
v. ELM can be used to train SLFNs with many non-differentiable activation functions.

3. CONCLUSION

S.NO	METHOD	ADVANTAGE	DISADVANTAGE
1.	Active Appearance Model (AAM)	Good results on using large training sets.	Requires extensive dataset with large amount of manually tagged points of the face.
2.	Gabor Wavelets	Good in spontaneous facial motion analysis.	Requires large datasets for training an enormous number of filters.
3.	Using 3D-Gradient Descriptor	Ability to capture the correlation between the frames & good results in classifying motions in video signal.	Manual point tagging.
4.	Using Spatio-Temporal Strain	Can differentiate between macro and micro facial expression	Controlled environment needed.
5.	Using Gabor Filters & GentleSVM (Support Vector Machine	Operates in frame by frame manner. Automatically locates the face and extracts the features.	Inaccurate image alignments may impair the performance.
6.	Using Temporal Interpolation Model	Detection is possible in small number of frames.	Less real-time recognition.
7.	Using Discriminant Tensor Subspace Analysis (DTSA) Plus Extreme Learning Machine (ELM):	High detection rate on 100fps.	Highly accurate data of The facial movement is required.

REFERENCES

1. P. Ekman, "Telling Lies" 2nd Edition, Norton, (2009).
2. S. Porter and L. Brinke, "Reading between the lies: Identifying concealed and falsified emotions in universal facial expressions", Psychological Science, vol. 19, pp. 508-514, (2008).
3. P. Ekman, "Facial Expressions of Emotion: an Old Controversy and New Findings", Philosophical Transactions of the Royal Society, vol. B335, pp. 63-69, (1992).
4. S. Polikovsky, Y. Kameda, and Y. Ohta. Facial micro-expressions recognition using high speed camera and 3D- gradient descriptor. In ICDP, 2009.
5. Z. Li, Handbook of Face Recogntion. Springer, (2005).
6. S. Lucey, "Investigating Spontaneous Facial Action Recognition through AAM Representations of the Face", Face Recognition Book. Pro Literatur Verlag, (2007).
7. M. Pantic, "Detecting Facial Actions and their Temporal Segments in Nearly Frontal-View Face Image Sequences", Systems, Man and Cybernetics, vol. 4, pp. 3358-3362, (2005).
8. M. Bartlett, "Automatic Recognition of Facial Actions in Spontaneous Expressions", Journal of Multimenia, vol. 1, no. 6, pp. 22-35, (2006).
9. M. Marszalek, "A spatio-temporal descriptor based on 3D-gradients", British Machine Vision Conference, (2008).
10. P. Dollar, "Behavior recognition via sparse spatiotemporal features", Visual Surveillance and Performance Evaluation of Tracking and Surveillance, pp. 65-72, (2005).
11. M. Shreve, S. Godavarthy, D. Goldgof, and S. Sarkar. Macro-and micro-expression spotting in long videos using spatio-temporal strain. In FG, 2011.
12. Qi Wu, Xunbing Shen, and Xiaolan Fu. The machine knows what you are hiding: an automatic micro-expression recognition system. In Proceedings of the 4th international conference on Affective computing and intelligent interaction - Volume Part II, pages 152–162, 2011.
13. Whitehill, J., Littlewort, G., Fasel, I., Bartlett, M., Movellan, J.: Toward Practical Smile Detection. IEEE Trans. Pattern. Anal. Mach. Intell. 31, 2106–2111 (2009).
14. T. Pfister, X. Li, G. Zhao, and M. Pietikinen. Recognising spontaneous facial micro-expressions. IEEE International Conference on Computer Vision 2011.
15. Z. Zhou, G. Zhao, and M. Pietik ̈ainen. Towards a Lip reading System. In CVPR, 2011.
16. Wang, S., Chen, H.L., Yan, W.J., Chen, Y.H., Fu, X.: Face recognition and micro-expression recognition based on discriminant tensor subspace analysis plus extreme learning machine. Neural Processing Letters Neural Processing Letters February 2014, Volume 39, Issue 1, pp 25-43-27 Feb 2013.
17. Huang, G., Zhu, Q., Siew, C.: Extreme learning machine: a new learning scheme of feed forward neural networks. In: 2004 IEEE International Joint Conference on Neural Networks, vol. 2, pp. 985–990. IEEE (2004)
18. Huang, G., Zhu, Q., Siew, C.: Extreme learning machine: theory and applications. Neuro computing **70**(1-3), 489–501 (2006)