International Journal of
Emerging Research in Management &Technology
ISSN: 2278-9359 (Volume-4, Issue-7)

Research Article

July 2015

Facial Emotion Detection by Minima and Re-Generation

Ayan Chatterjee[1], Bachcha Lal Pal[2]
[1]Master of Technology Research Scholar, [2]Assistant Professor and Head
[1, 2] Department of Computer Science, Mewar University,
Rajasthan, India

Abstract—

Automatic facial recognition systems have come a long way since 1970. Today, even mobile application developers have hands on APIs on platforms like Visual Studio for face detection to be implemented and distributed commercially in the hands on millions of end users. Emotion detection is achieved in 3 phases, namely face detection, feature extraction and expression classification. This paper discusses emotion detection using local minima positions starting from an approximate result obtained from the mouth and directing towards a more precise result with left eye, right eye, nose and mouth.

Keywords— Emotion Detection, Emotion Recognition, Expression Classification, Facial Expression Recognition

I. INTRODUCTION

There are several approaches to solve facial emotion detection taking inputs from 2 dimensional images from geometric methods to neural network to 3 dimensional recognition using Microsoft Kinect. Recently face recognition has made its way to enterprise authentication, security systems, credit card verification, criminal identification, device authentication in Windows Hello[9] on Windows 10.

Recent research include feature based and model based technique[2], geometric alignment and local binary patterns[3], principle component analysis[4], affect recognition methods[5], and self-organizing map[6] to name a few.

Filters applied on image information often tend to modify the original data resulting in loss of image information. Our goal was to keep it intact and extract as much information as possible using minima position of the pixel information.

II. MOUTH'S LOCAL MINIMA

Emotion detection by mouth's local minima[1] starts by exercising the position of minimum pixel values (information) of each column marked down. The positions together form a polynomial trend-line as shown in figure 1.

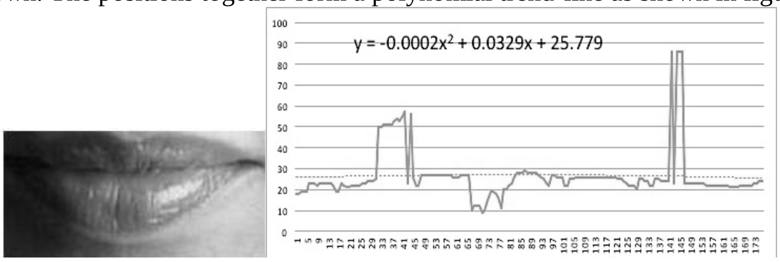

Fig. 1. Emotion detection by mouth's local minima

III. RE-GENERATION

Re-generation or mirroring is filtering incorrect detection of facial feature regions and duplicating/mirroring with the acceptable ones through facial axis. Here, we try to understand what is on our plate and what can be deduced on application of

1. Facial feature proportion approximation, and
2. Facial symmetry

Implementation of the same is demonstrated in subsection F in section IV.

IV. FACIAL FEATURES LOCAL MINIMA

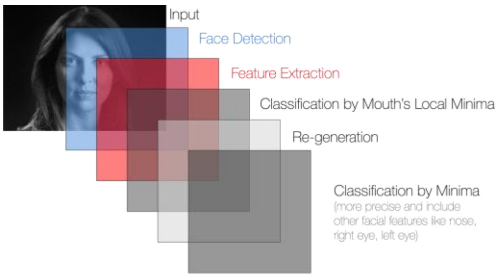

Fig. 2. From input to expression classification

We shall be looking closely at the effect of local minima on the 4 facial features – left eye, right eye, nose and mouth. For explanation purposes, figure 3 has been taken whose detected facial features are acceptable by use of Viola Jones algorithm[8]. To summarize and classify the above statements, we employ facial emotion detection as explained in figure 2.

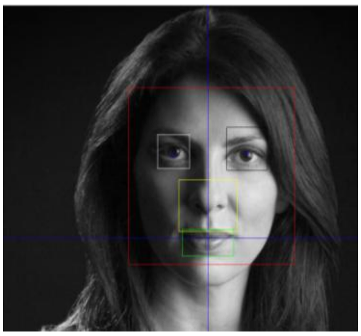

Fig. 3. Acceptable detection of facial features by use of Viola Jones algorithm

A. Facial Axis

Facial axis is the centre of the human face that divides it into two equal halves, taking advantage of the symmetrical nature of the human face. The goal of the further subsection B, C, D and mouth ends is to determine the facial axis, re-generate or in common terms mirror and conclude a more precise classification from what was determined by mouth's local minima[1] only.

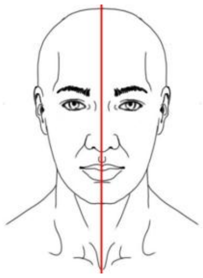

Fig. 4. Illustration showing facial axis

B. Nose

This subsection discusses the facial feature, nose. In this, the areas of interest are the two openings called nostrils or nares. Nostrils form a cluster of minimum pixel information as compared to the rest.

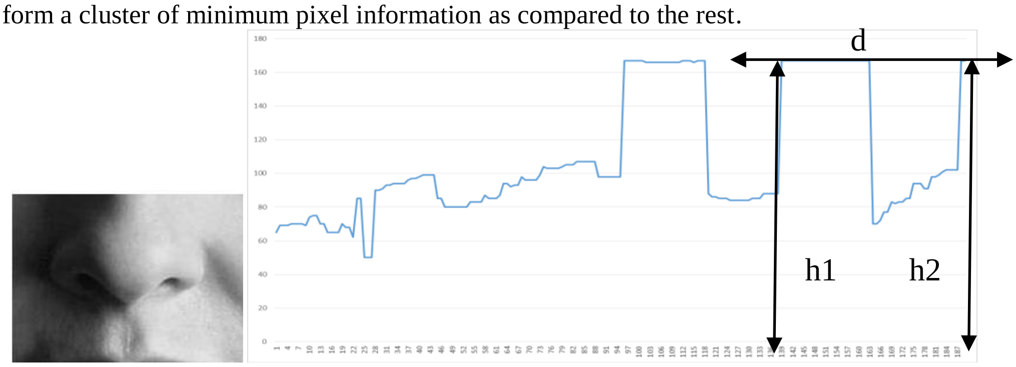

Fig. 5. Extracted nose

It can be seen in figure 5, that each nostril forms a distinctive plateau with average height h1 and h2 respectively. The difference between the 2 heights h1 and h2 shows the angle of tilt i.e. if h1 ≅h2, face is horizontal as captured from a camera. Similarly, if h1>h2, face is tilted towards the right and if h1<h2, face is tilted towards the left.

In figure 5, the average h1 and h2 are 166.522 and 167, showing h2 is greater by 0.478 hence the face is almost straight with a minute tilt towards the left.

On continuous captures and video capture, 'd' can be monitored for flared nostrils showing distinct displeasure emotions like anger, agitated, flight response, fear, and anxiety to name a few. d/2 is also one of the points where facial axis should pass through.

This is a good data set to work with. In general, when polynomial trend-line of degree 2 is constructed i.e. $\alpha x^2 + \beta x + \gamma$, the sign of α will be negative.

C. Left Eye

The eyes follow a somewhat different approach. We take the whole real estate of the detected left eye region and divide it into 9 quadrants as per rule of thirds as shown in figure 6. The 4 intersecting centre points are called pivots. Quadrant 7, 8 and 9 will consist of an eyebrow, quadrant 5 will cover eye ball where the concentration of local minima will be maximum. If eye ball is present in quadrant 4 or 6, it signifies the person is looking towards the left or right. We have taken 5 consecutive minima for the same.

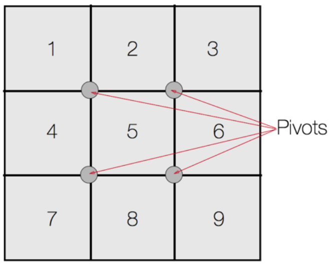

Fig. 6. Rule of thirds

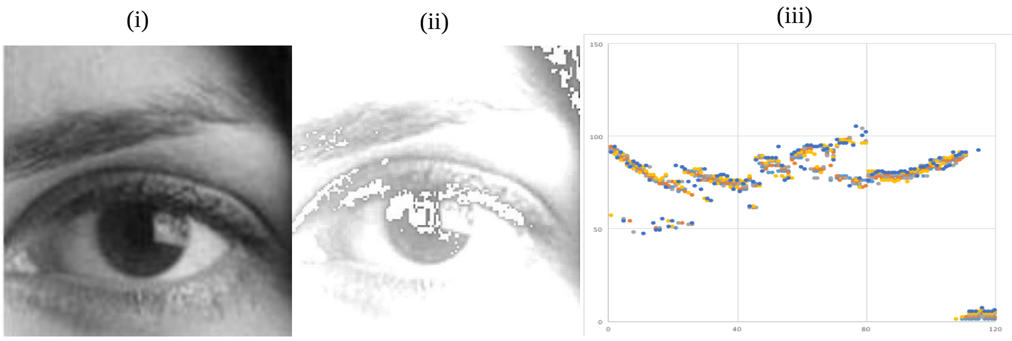

Fig. 7. From left to right, (i) extracted left eye, (ii) local minima applied (5 iterations), and (iii) the positions of those minima

If the concentration of the local minima is found on quadrant 1, 2, 3, 7, 8, or 9, then the detected eye region is not completely acceptable and it needs to be adjusted.

D. Right Eye

Similar to the left eye, the right eye of human face follows similar pattern. Its analysis extracted from figure 3 is shown in figure 8.

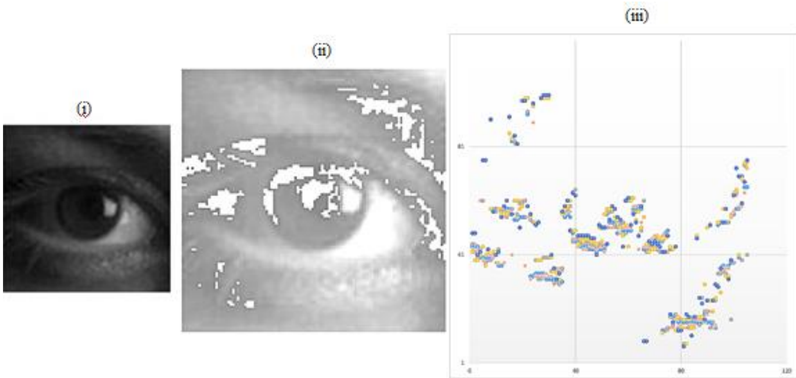

Fig. 8. From left to right, (i) extracted right eye, (ii) local minima applied (5 iterations), and (iii) the positions of those minima

E. Verifiable Information

Before moving towards a more precise and adding computational weights, here, we take a moment to see the feasibility of re-generation. The feasibility can be detected as shown in figure 9.

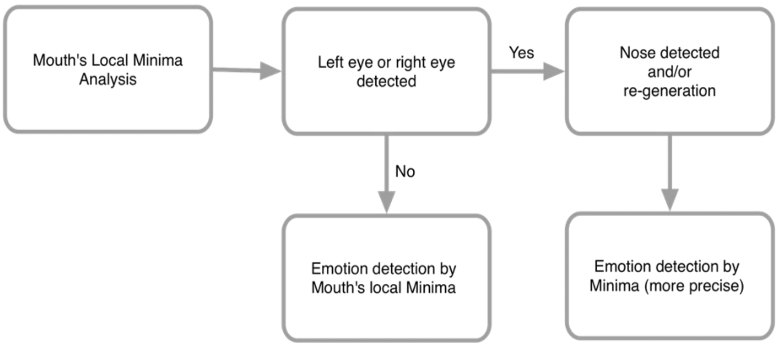

Fig. 9. Flow for verifiable information to expression classification

F. Re-Generation

In this phase, we take the information we have and try to correct inaccurately determined facial feature region as shown in figure 10. We do that by looking at the trends of the minima positions of the facial features derived, and find out any sudden drops or misalignments in the trend, giving us the end of features like mouth end. We, the duplicate it according to pairs such as

1. Left eye and right eye
2. Mouth ends

After this is done, we move on to check that the feature regions fall in order such as

1. Right eye falls on the right of facial axis
2. Left eye falls on the left of facial axis
3. Nose is below the eyes and above mouth
4. Mouth falls below detected nose

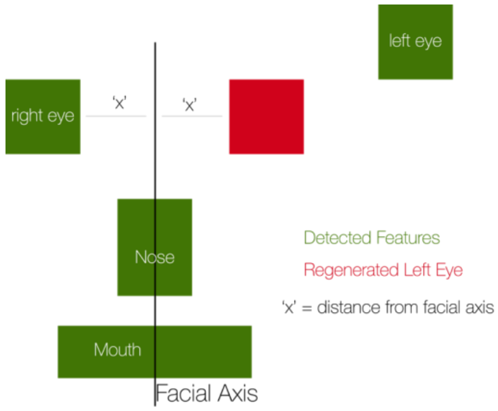

Fig.10. example of re-generation

G. Classification

TABLE I CLASSIFICATION CRITERIA OF THIS APPROACH

Acceptable	Re-generation	Classifiable	Precision
Both Eyes, Nose, Mouth	-	Yes	High
Left Eye, Nose, Mouth	Right Eye	Yes	High
Right Eye, Nose, Mouth	Left Eye	Yes	High
Left Eye, Mouth	Right Eye, Nose	Yes	High
Right Eye, Mouth	Left Eye, Nose	Yes	High
Both Eyes, Mouth	Nose	Yes	High
Nose, Mouth	-	Yes	Low
Mouth	-	Yes	Low
Both Eyes, Nose	-	No	-
One Eye, Nose	-	No	-
Nose	-	No	-

Mouth, being the most important feature contributing in classification, table 1 shows the criteria that determine whether a face is emotionally classifiable. Table 2 shows up to 14 classifiable emotions that can be derived of this approach using machine language approach by monitoring mouth, eyes and nose. The emotion derived in figure 3 is that of a happy face, as derived from mouth's trend-line.

TABLE II CLASSIFIABLE EMOTIONS

Single Capture (Image)		Continuous Capture (Video)
Without Machine Learning	With Machine Learning	
Happy, Sad, Straight face, Surprise, Wink, and degree of happiness or sadness	Happy, Sad, Straight Face, Wink, Dizzy, Focus, Confusion, Surprise, Anger, Fear, Silly, Laughter, Grin	Happy, Sad, Straight Face, Wink, Dizzy, Focus, Confusion, Surprise, Anger, Fear, Silly, Laughter, Grin, Sleepy

V. EXPERIMENTAL RESULTS

This approach has been tested on 50 data sets taken at random from FACS[7] database and Viola Jones algorithm[8] was taken for face detection. Among those, 3 were reported unclassifiable by re-generation due to improper detection of the mouth, 1 false positive occurred when chin was detected as mouth and emotion state happy was concluded, and the rest were acceptable conclusions. One example of each are demonstrated in the figures below.

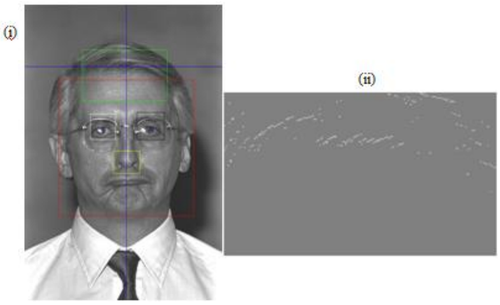

Fig. 11. Improper mouth region detection resulting as unclassifiable. Left to right, (i) face and feature detection, and (ii) extracted mouth with local minima positions

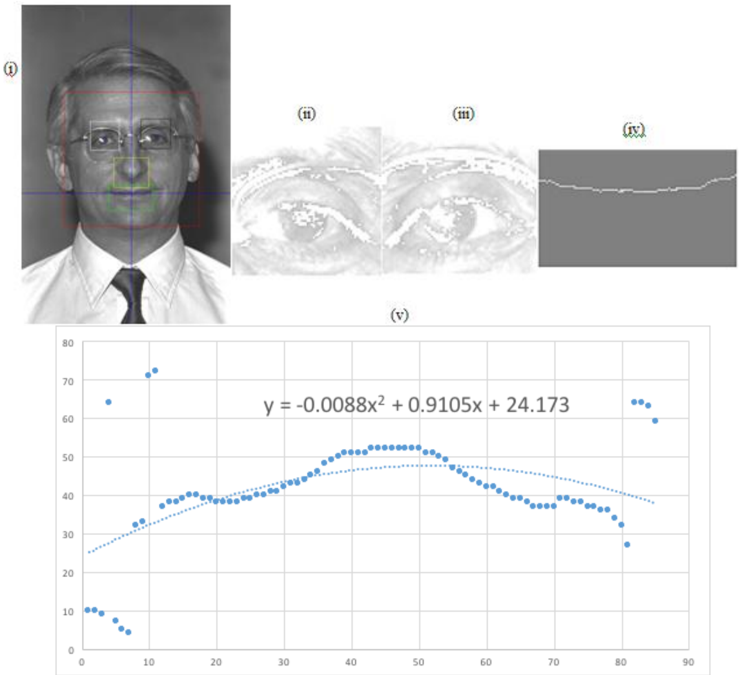

Fig. 12. Classified as happy. All features are acceptable. Left to right, (i) face and feature detection, (ii) extracted left eye with local minima positions, (iii) extracted right eye with local minima positions, (iv) extracted mouth with local minima positions, and (v) minima positions and trend-line on nose.

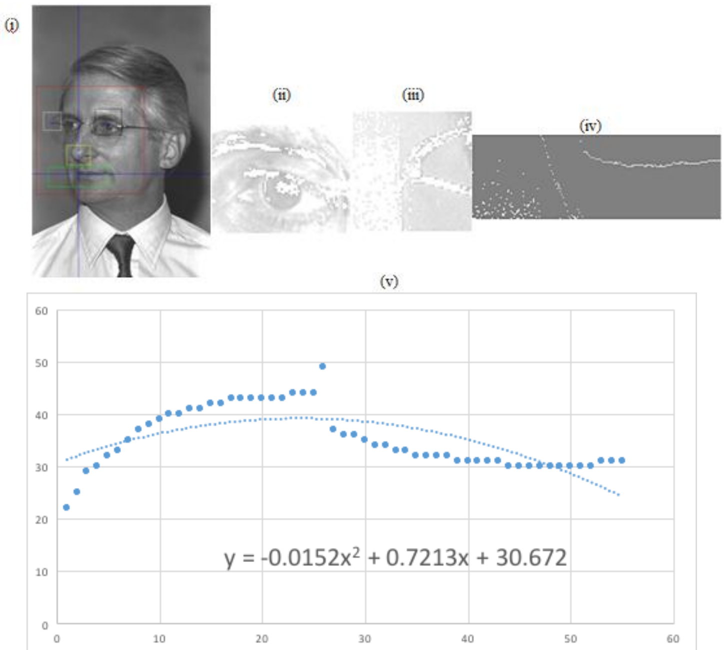

Fig. 13. Classified as happy. Re-generation to occur. From left to right, (i) face and feature detection, (ii) accepted left eye with local minima positions, (iii) unacceptable right eye position declaring re-regeneration on the region, (iv) acceptable mouth position declaring re-generation on the mouth end, and (v) minima positions and trend-line on nose

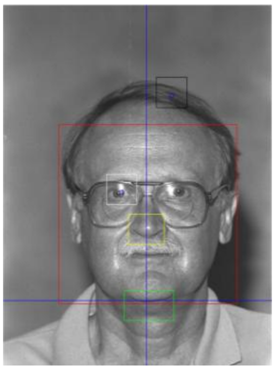

Fig. 14. Classified as happy. False positive result where chin is detected as mouth region.

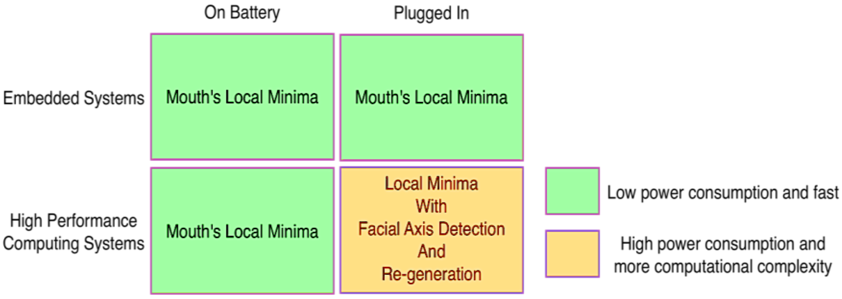

Fig. 15. Recommendation on the use of this approach

Recommendation of the use of this approach on figure 15 is based on power consumption and limited computational power of embedded systems.

VI. CONCLUSIONS

Facial emotion detection by minima along with re-generation acts as a great way to refine and enhance the accuracy of detection. This approach breaks down the complete problem into smaller sub-problems which can be handled easily. At first we arrive at a ballpark with mouth's local minima and then slowly proceed to increase accuracy of the result. A machine learning approach could be further introduced for user personalization and globalization factors.

ACKNOWLEDGMENT

The authors would like to thank the teaching and non-teaching staff at Mewar University for their continuous support in this research.

REFERENCES

[1] Chatterjee et al.,"Effect of Mouth's Local Minima on Emotion Detection", International Journal of Advanced Research in Computer Science and Software Engineering 5(4), April- 2015, pp. 1384-1388

[2] Mishra, B., Fernandes, S.L., Abhishek, K., Alva, A., Shetty, C., Ajila, C.V., Shetty, D., Rao, H., Shetty, P, "Facial expression recognition using feature based techniques and model based techniques: A survey", Electronics and Communication Systems (ICECS), 2015.

[3] Xun Wang,Xingang Liu, Lingyun Lu, Zhixin Shen, "A New Facial Expression Recognition Method Based on Geometric Alignment and LBP Features", Computational Science and Engineering (CSE), 2014 IEEE 17th International Conference, 19-21 Dec. 2014.

[4] Agrawal, S., Khatri, P., "Facial Expression Detection Techniques: Based on Viola and Jones Algorithm and Principal Component Analysis", Advanced Computing & Communication Technologies (ACCT), 2015 Fifth International Conference, 21-22 Feb. 2015.

[5] Zeng, Z., Pantic, M., Roisman, G.I., Huang, T.S., "A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions",Pattern Analysis and Machine Intelligence, IEEE Transactions (Volume:31, Issue: 1), 07 March 2008.

[6] Anima Majumder, Laxmidhar Behera, Venkatesh K. Subramanian, "Emotion recognition from geometric facial features using self-organizing map", Volume 47, Issue 3, March 2014, Pages 1282–1293.

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

[8] P. Viola & M. Jones (2001), "Rapid object detection using a boosted cascade of simple features", Computer Vision and Pattern Recognition(CVPR) (Volume:1).

[9] Joe Belfiore, Windows Blog, "http://blogs.windows.com/bloggingwindows/2015/03/17/making-windows-10-more-personal-and-more-secure-with-windows-hello"