Study of DRMF and ASM facial landmark point for micro expression recognition using KLT tracking point feature

R A Asmara, P Choirina*, C Rahmad, A Setiawan, F Rahutomo, R D R Yusron and U D Rosiani

1 Department of Information Technology, State Polytechnic of Malang, Indonesia
2 Electrical Engineering Department, State Polytechnic of Malang, Indonesia

*priska choirina@polinema.ac.id

Abstract. Micro-expression recognition is one of the popular researches in analysing expressions on the face. Micro-expression is a facial movement that occurs in a short time and is difficult to identify manually by human eyes. In general research, facial landmarks are used to form a large size ROI for each facial feature for the feature extraction process. In this study, we track the subtle motions of micro expressions by using point features. This approach aims to get feature extraction from tracking results and then analyse micro-expression. We compared the Active Shape Model and Response Map Fitting methods to produce accurate points and fast time on facial features. To measure the subtle motion tracking of facial features in each frame tracking is done using the Kanade-Lucas-Tomasi method. To estimate the rationality of our method, we conducted an experiment on CASME II and SAMM dataset for micro-expressions. The results show that the points on DRMF are more accurate with point-to-point error is 7.9 and the time taken is faster which requires time is 0.02 second. We evaluated the method proposed for evaluation showed that using CASME II - Naive Bayes (79.3%) and SAMM - Naive Bayes (74.6%).

1. Introduction

Psychological studies have shown that a person may deliberately hide the original emotion, but it cannot falsify his expression it is called micro-expressions. Micro-expression occurs when the muscle movements in the face and short to deliberately hide his real emotions [1]. The term micro itself usually has two meanings: a very brief and very small expression change. According to Ekman and Friesen in his research, micro-expressions have a very quick description that goes around 1/25 to 1/3 s [2], and usually appears at a time when someone wants to hide or suppress the real emotions. The micro-expression occurs spontaneously and cannot be controlled by the mind of a person, then it is considered an important clue to reveal the actual expression on someone and be able to detect deception [3].

Micro-expression recognition currently has attracted much attention of the researchers in the last few years, and currently is one of the focus points in the field of computer vision. Many computer vision techniques that have been developed to track or identify facial activity into 3 levels [4]. The initial level is the tracking of facial features to detect movement of the face area, facial recognition further actions in FACS system to recognize a face and the last activity analysis of facial expression for recognizing facial expressions in humans. In some studies, tracking point facial features have been studied for facial expression recognition, especially macro-expressions and give good results. The advantage of using tracking points feature is that less time is needed because at some time it can also be used for expression recognition.
recognition [2]. Micro-expression studies generally use the entire face [5,6] but in fact, the movement of micro expression only occurs in a few parts [7]. When micro-expression occurs, movement in the eye area dominates more than the other areas [3].

This research tries to use the tracking points on the face area for micro-expression recognition. Shuoqing have been doing research on the micro-expressions recognition by tracking points feature on the utilization of the area of the mouth [2]. In this study, utilizing the methods of KLT for tracking points on the eyebrows, eyes, and mouth [7]. To get the point features on face area that is fast and accurate, we will do a comparison of DRMF [8,9] and ASM methods for detect facial points features [10,11]. Tracking at any point of the face will be used to acquire the extraction feature for the micro expression recognition process.

2. Proposed method
The block diagram of the micro-expression recognition with point features is illustrated in Figure 1. In this section, we explain this step in Micro-Expression Recognition that we propose.

![Figure 1. The block diagram of the micro-expression recognition with point features tracking.](image)

2.1. Facial landmark detection
To recognize the micro-expressions, locating facial features such as the location points of the eyes, eyebrows, nose, and mouth plays a significant role because if facial point features are located accurately, the extraction and classification stages of the following features will be stronger and more efficient. The first step is to recognize the facial area by utilizing facial landmark method. By comparing ASM and DRMF methods [8], to find position and feature area of the face at the first frame in the image sequences.

![Figure 2. Landmark points in a facial region: (a) 49 point features using the DRMF, (b) 79 point features using the ASM.](image)
2.1.1. **Active Shape Model (ASM).** ASM are statistical models, contains global models and a lot of local feature models [9]. Objects that are detected in various features must be consistent from one image to another is one of the main requirements of ASM. In this study, there were 79 points (see Figure 2, (b)) defined for details of the face shape which included the eyebrows, eyes, mouth, nose, and jaw. There are \( n \) point features and each is located on a clear facial contour. Position \( n \) of this point is arranged in vector form, it is \( X = [x_1, y_1; x_2, y_2; \ldots; x_n, y_n]^T \). Where \( x_n \) and \( y_n \) are horizontal coordinate and vertical coordinate of the \( k^{th} \) feature point respectively. PCA operation is used eigenvectors of the covariance matrix corresponding to main shape variations can be generated. The ASM searching algorithm uses an iteration process to find the best landmarks [9].

2.1.2. **Discriminative Response Map Fitting (DRMF).** DRMF is a novel discriminative regression-based approach within the CLM framework [10]. This method shows impressive performance in general face fitting scenarios and is very suitable for handling dynamic backgrounds, large amounts of occlusion, and hanging illumination conditions [11]. With lower computational time and real-time capabilities [12]. First, the Viola-Jones face detector is used to find the face area in each frame. Second, a set of initial point features is calculated by extracting patch responses followed by low dimensional projections. Third, the DRMF repeats this initial feature point by correlating with the target frame. The resulting features that control the variation in shape and appearance are learned from a training set. Statistical models of appearance patterns and patterns of variability, which are presented in, are applied in the DRMF method in a strong and accurate way to find. In our application, we use 49 point features (see Figure 2 (a)) obtained from the DRMF method.

2.2. **Tracking point features**

From the previous process, we get the coordinates of the points on the face of the ASM and DRMF methods. Furthermore, from these points the process of determining the facial features will be used. In this study, we used 20 facial points in the right eyebrow, left eyebrow, right eye, left eye and mouth area. The following Table 1 determines the points of the face area used:

| Left Eyebrow | Right Eyebrow | Left Eye | Right Eye | Mouth |
|--------------|--------------|----------|-----------|-------|
| ASM          |              |          |           |       |
| 22-25        | 16-19        | 28-30, 69| 33-35, 74 | 49, 55, 52, 58 |
| DRMF         |              | 2-5      | 7-10      | 20-23 | 26-29 | 32, 35, 38, 41 |

Table 1 shows the determination of the index points of the results of facial landmark detection in each face area (see Figure 1). In the left eyebrow area with the DRMF method, we determine the 2nd landmark coordinate to the 5th landmark coordinate. For the ASM method, the left eyebrow area uses the 22nd to 25th landmark coordinates. Examples of points for determining landmark points can be seen in Figure 3.

![Figure 3. A landmark point for the DRMF (a) and ASM (b) tracking point feature.](image-url)
After determining face area points, subsequent tracking point features based on the Kanade-Lucas-Tomasi (KLT) were adopted for tracking points feature in our approach [13]. This method calculates the number of differences in the square of the intensity between the features in the current and previous frames. Feature movement is defined as a displacement that minimizes the amount of difference and is carried out continuously in the image sequences until all features can be traced. First, initiate the points feature on the onset frame and save it for the next process. In the iteration process, the point feature will be tracked until the offset frame and in this process, it is possible to lose the feature point. To avoid this, we process facial landmark detection to restore the appropriate point features.

2.3. Feature extraction

After tracking the points feature, there are two types of extraction features based on the coordinates of the point features in each frame transferring the coordinates can be formulated as follows:

\[ X = (x_{p,f} - x_{p,f-1}), \quad Y = (y_{p,f} - y_{p,f-1}) \]  

Where \((x_{p,f}, y_{p,f})\) for current frame and \((x_{p,f-1}, y_{p,f-1})\) for the previous frame. From Equation 1 generate two features is \(X\) and \(Y\) for each displacement distance. There are 20 point features in every frame. So, the dimension of the vector feature is also 20.

The point displacement at each frame produces a vector that has a magnitude and orientation (theta) component. Magnitude is the distance between the initial point \(P\) and the end point \(Q\). In symbols magnitude of \(\overrightarrow{PQ}\) is written as \(|\overrightarrow{PQ}|\). If the coordinates of the initial point and end point of the vector are known, the distance calculation can be used to find the magnitude. The magnitude can be formulated as follows:

\[ |\overrightarrow{PQ}| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \]  

The direction of a vector is the measure of the angle it makes with a horizontal line. One of the following formulas can be used to find the direction of a vector:

\[ O(x,y) = \tan^{-1}\left(\frac{y}{x}\right) \]

3. Experimental results

3.1. Dataset

To evaluate the performance of the proposed algorithm, the experiments were carried out on two recent spontaneous micro-expression databases, namely CASME II [14] and SAMM [15]. In this experiment, we only use frame onset to offset to increase the accuracy of the micro-expressions recognition and balance the data for the classification process. From the micro expression label, we use the label disgust, happiness, and surprise. In the CASME II disgust label there are 512 frames, happiness there are 514 frames and 516 frames of 6 videos on each label that has been extracted frame onset until the offset. For SAMM, the disgust label amounts to 508 frames, happiness 505 frames and surprise 506 frames. Meanwhile, the total videos we use are 18 videos in each dataset.

3.2. Execution time

Micro-expressions recognition needs to be performed in real-time, computational time is an important aspect of this method [16]. The code has been applied in MATLAB and executed on Intel Core i7-6700T CPU that runs at 2.8GHz, uses a 64-bit Windows OS and 4GB RAM. The most time-consuming part is landmark detection, ASM method takes 3.79s in both datasets. DRMF takes much faster than ASM, which is 0.02s. Since they need a lot of time, landmarks are only detected in the first frame (onset frame) and when feature point is lost during the tracking process. The tracking process for each frame takes 0.04 s in both databases.
3.3. Accuracy of DRMF vs ASM facial landmark points

In this experiment, we conducted an experiment to determine accuracy of point features from ASM and DRMF. Because in micro-expression recognition, the accuracy of area features is important. We used 3 videos in each dataset and use 20 point features on the onset frame. For measurement of feature point accuracy, we use the Euclidean distance formula.

Figure 4 displays the result of point-to-point errors (the distance between model points and the corresponding points marked in frame). The length of point-to-point error is from 2.54 to 89.95. The average of point-to-point error for 20 landmarks using both algorithms, the result shows that difference between ASM and DRMF in CASME II is equal 14.26. In SAMM, the difference between ASM and DRMF equal to 11.76. So, the DRMF algorithm locates the points more accurately than ASM algorithm.

![Figure 4. Shows point-to-point errors for the DRMF and ASM models.](image)

3.4. Micro-expression recognition classification

In this research, the results of the feature extraction would be classified by machine learning using Naïve Bayes and Multi-Layer Perceptron (MLP) with popular training algorithm backpropagation. For backpropagation, the activation function used is logistic sigmoid. This experiment uses learning rate is 0.001 (values between 0-1). The value of the epochs passed for training process with value is 200.

| Dataset | MLP-Backpropagation | Naïve Bayes |
|---------|---------------------|-------------|
|         | ASM                 | DRMF        | ASM         | DRMF        |
| CASME II| 53.5%               | 61.9%       | 66.0%       | 79.3%       |
| SAMM    | 96.1%               | 64.1%       | 82.0%       | 74.6%       |

Table 2. The recognition accuracies with MLP and Naïve Bayes.

Table 3 shows that with ASM in SAMM dataset it gives good results with backpropagation classifier is 96.1% and naïve bayes is 82.0%. Whereas in CASME II, DRMF gives good results is 61.9% in backpropagation classifier and 79.3% in naïve bayes. However, experimental accuracy from the previous landmark points showed that the ASM method produced was not accurate enough to detect the face area (see Figure 2(b)). Conversely, with the DRMF the exact point of the landmark gives good results with a minimal minimum error. Shuoqing Yao on the micro-expressions recognition resulted in an accuracy of 84.0% on the label happiness and 74.5% on the disgust label by using the CASME dataset [2].

4. Conclusion

In this paper, we present an algorithm for micro-expression recognition by utilizing facial points feature. With facial landmark methods from ASM and DRMF, we compare the two methods to get the exact point area of the face and give a quick time in processing. Because the movement on the micro-
expression is very subtle and fast, it requires the accuracy of the face area and speed of processing time. We use points on the eyebrow, eye and mouth area as point’s features. The process of tracking point features uses the KLT method to get movement in the face area. Using datasets from SAMM and CASME II, we get the results that processing time on detection facial landmark using DRMF method (~0.02s) faster than ASM (~3.79s). In experiments, the classification of micro-expressions, ASM gives high results in both datasets, but the accuracy of the points given to this method is quite poor. With DRMF method the accuracy is given at each point achieve good result, with the classification result with CASME II - Naive Bayes (79.3%) and SAMM - Naive Bayes (74.6%).

Acknowledgments
The authors would like to thank State Polytechnic of Malang for supporting this research.

References
[1] Iwasaki M and Noguchi Y 2016 Hiding true emotions: Micro-expressions in eyes retrospectively concealed by mouth movements Scientific Reports 6 22049
[2] Yao S, He N, Zhang H and Yoshie O 2014 Micro-expression recognition by point features tracking IEEE International Conference on Communications 1–4
[3] Duan X, Dai Q, Wang Q, Wang Y and Hua Z 2016 Recognizing spontaneous micro-expression from eye region Neurocomputing 217 27–36
[4] Li Y, Wang S, Zhao Y and Ji Q 2013 Simultaneous Facial Feature Tracking and Facial Expression Recognition IEEE Transactions on Image Processing 22 2559–2573
[5] Liu Y, Li B and Lai Y 2018 Sparse MDMO: Learning a Discriminative Feature for Spontaneous Micro-Expression Recognition IEEE Transactions on Affective Computing, pp. 1–1, 2018.
[6] Wang S J, Wu S, Qian X, Li J and Fu X 2017 A main directional maximal difference analysis for spotting facial movements from long-term videos Neurocomputing 230 382–389
[7] Wang S J, Yan W J, Zhao G, Fu X and Zhou C G 2015 Micro-Expression Recognition Using Robust Principal Component Analysis and Local Spatiotemporal Directional Features Computer Vision - ECCV 2014 Workshops 325–338
[8] Oh Y H, See J, Le Ngo A C, Phan R C W and Baskaran V M 2018 A Survey of Automatic Facial Micro-Expression Analysis: Databases, Methods, and Challenges Frontiers in Psychology 9 1128
[9] Iqtait M, Mohamad F S and Mamat M 2018 Feature extraction for face recognition via Active Shape Model (ASM) and Active Appearance Model (AAM) IOP Conf. Ser.: Mater. Sci. Eng. 332 012032
[10] Asthana A, Zafeiriou S, Cheng S and Pantic M 2013 Robust Discriminative Response Map Fitting with Constrained Local Models IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA 3444–3451
[11] Rosiani U D, Ririd A R T H, Choirina P, Sooai A G, Sumpeno S and Purnomo M H 2018 Micro Expression: Comparison of Speed and Marking Accuracy in Facial Component Detection International Conference on Computer Engineering, Network and Intelligent Multimedia (CENIM) 221–226
[12] Liong S T, See J, Wong K and Phan R C W 2016 Automatic Micro-expression Recognition from Long Video Using a Single Spotted Apex Computer Vision – ACCV 2016 Workshops 345–360
[13] Tomasi C and Kanade T 1991 Detection and Tracking of Point Features International Journal of Computer Vision
[14] Yan W J 2014 CASME II: An Improved Spontaneous Micro-Expression Database and the Baseline Evaluation PLOS ONE 9 1 e86041
[15] Davison A K, Lansley C, Costen N, Tan K and Yap M H 2018 SAMM: A Spontaneous Micro-Facial Movement Dataset IEEE Trans. Affective Comput. 9 1 116–129
[16] Huang X, Zhao G, Hong X, Zheng W and Pietikäinen M 2015 Spontaneous Facial Micro-
Expression Analysis using Spatiotemporal Completed Local Quantized Patterns

*Neurocomputing* 175