Facial Expressions Recognition Using Markov Stationary Feature - Vector Quantization and Support Vector Machine Method

I Maliki, F S Jarockohir
Department of Informatics Engineering, Faculty of Engineering and Computer Science, Universitas Komputer Indonesia, Jln. Dipatiukur No 112-116 Bandung 40132 Indonesia

Email: irfan.maliki@email.unikom.ac.id

Abstract. Facial expression is a form of nonverbal communication that can convey the emotional state of someone to the person who observes it. Research on the recognition of facial expressions is one of the interesting fields in computer science. This research aims to improve the accuracy of recognition performance. The process carried out in this research is to perform feature extraction and image classification. Markov Stationary Feature - Vector Quantization (MSF-VQ) method is used for feature extraction and Support Vector Machine (SVM) for image classification. Data set in used is 1440 data with six classifications of facial expressions. The results of the testing showed 97.41% which stated that this method could be recommended to be applied in the facial expressions recognition.

1. Introduction
Facial expressions recognition is one of many research fields that is still interesting to date. Facial expressions are one or more movements or positions of muscles under the facial skin. Facial expressions are very important part of communication. In general, facial expressions have 6 types of types, such as Happy, Sad, Surprised, Fear, Anger and Disgust [1] [2].

Several researchers have carried out several studies related to facial expressions recognition. Firstly, Saputra who used the Local Binary Pattern (LBP) method to get an accuracy of 65.1% [3]. Secondly, Pengututan [4] who used the Discrete Cosine Transform (DCT) and Principal Component Analysis (PCA) got an accuracy of 50%. Lastly, Saumil [5] who used Linear Discriminant Analysis (LDA) got an accuracy of 60%.

From these studies obtained several weaknesses, namely the results of accuracy of several methods used have not achieved maximum results. Based on this, this study aims to apply methods that are able to improve accuracy of the recognition of facial expressions. In facial expressions recognition, two steps must be done, first feature extraction from an input image and then classification of that image into a particular class of expression.

This study used the Markov Stationary Feature - Vector Quantization (MSF-VQ) method on feature extraction based on research [6] for face recognition with an accuracy of 99%. At the image classification stage, using the SVM method to determine the expression class of the image.
2. Method
The method of this research is shown in Figure 1.

![Flow methodology](image)

**Figure 1. Flow methodology**

In general, the stages carried out in this study consist of three stages, namely:

1. **Preprocessing**
   Preprocessing process is a step to determine the face area to be selected and reduce noise in order to get better accuracy of feature extraction. The face detection process is using the Viola & Jones method and the lowpass filter to reduce noise. In the process of face detection there are several processes carried out [7]: Grayscale Haar-Like Feature - Integral image - Adaboost (Adaptive Boosting) - Cascade Classifier.

2. **Feature Extraction**
   Feature extraction is the process of getting values that are characteristic, and as input from machine learning, in this study feature extraction has 2 gradual process:
   a. **Vector Quantization (VQ)**
      VQ is the process of forming compressed images, how VQ works is to map the codebook [9]. In figure 2, illustration of VQ. The process of VQ is as follows (see Figure 2):

![Vector Quantization illustration](image)

**Figure 2. Vector Quantization illustration**
Step 1: Random codebook initials

\[ c = \{ c_1, c_2, c_3, \ldots, c_{Nc} \} \]

Step 2: Shape the Image into blocks

\[ x = \{ x_1, x_2, x_3, \ldots, x_{NB} \} \]

Step 3: Look for the index value, to find the index value, look for the Euclidean distance value between the initial codebook and the image block to search for distance.

\[
d(x_b, c_j) = \sqrt{\sum_{t=0}^{n-1} (x_t - c_t)^2}
\]

To find the index value, determine the following equation:

\[
\text{indeks} \in X: d(x_b, c_{j-1}) < d(x_i, c_j)
\]

Step 4: Update the codebook

From the occurrence of the index on each block, the values of \( x_t \) and \( N_j \) have been accommodated, then calculate the following equation:

\[
c_t = \frac{x_t}{N_j}
\]

This value is the value of each of these values is the content of the new codebook. After that, look for distortion values by calculating the distance between the new codebook and the block:

\[
d(X, C) = \sum_{t=0}^{n-1} (x_t - c_t)^2
\]

After the distortion value is generated, the distortion value is divided by the number of \( Nb \times n \).

\[
D_m = \frac{d(X, C)}{Nb \times n}
\]

After getting the distortion value, update the codebook until it meets and stops the codebook update process until it meets the conditions:

\[
\frac{D_{m-1} - D_m}{D_m} \leq \varepsilon = 0.001
\]

After the codebook update process is complete, map the index values that are accommodated. By taking the value that is on the updated codebook.

b. Markov Stationary Feature (MSF)

The process of MSF is as follows [5]:

1. Generating Spatial Co-occurrence Matrix
   
   This stage is the stage of generating co-operation. The matrix \( p_l \) is the stage of calculating the number of matrix values for the occurrence of coordinates, co-occurrence matrix denoted as \( C = (c_{ij})_{KxK} \)

2. After obtaining the Co-assurance Matrix, the corresponding Matrix Transition \( p = (p_{ij})_{KxK} \) is derived from Spatial co-operation matrix \( C \)

\[
p_{ij} = \frac{c_{ij}}{\sum_{j=1}^{K} c_{ij}}
\]

3. Distribution status is required after step n is \( \pi(n) \) and the distribution initials are invariant measures of the Markov chain, which can be accumulated
\[ \pi \approx \frac{1}{K} \sum_{i=1}^{K} \widetilde{\alpha}_{ij} \]

where \( A_n = [\alpha_1, ..., \alpha_n]^T \)

Here \( \pi \) is a stationary distribution that satisfies \( \pi = \pi P \).

\[ A_n = \frac{1}{n+1} (I + P + ... + P^n) \]

4. Full features that include a combination of initial distributions and a stationary distribution

\[ \pi(0) = \frac{c_{i1}}{\sum_{i=1}^{K} c_{i1}} \]

\[ MSF = [\pi(0), \pi]^T \]

3. Classification

The next stage will be learning and testing on training data with histogram test data, using the M-SVM method on the Multiclass Support Vector Machine (M-SVM) method. The M-SVM method can only do binary classifications (two classes). The approach used is one-against-all (OAA) [10].

3. Results and Discussion

Performance testing is done by using a confusion matrix to calculate how much accuracy is obtained. Confusion Matrix is a table that maps the performance results of an algorithm. Each column in the matrix represents the predicted class, while each row represents the actual class. The dataset used for training was 1170 data, and 270 testing data from 15 subjects. In this test using three types of filters consist of 3x3, 5x5 and 7x7 (see Tables 1-3).

| Expression | Happy | Sad | Surprised | Fear | Anger | Disgust | Accuracy |
|------------|-------|-----|-----------|------|-------|---------|----------|
| Happy      | 39    | 0   | 1         | 3    | 2     | 0       | 86.67%   |
| Sad        | 7     | 29  | 0         | 5    | 0     | 4       | 64.44%   |
| Surprise   | 1     | 2   | 41        | 1    | 0     | 0       | 91.11%   |
| Fear       | 6     | 0   | 1         | 32   | 6     | 0       | 71.11%   |
| Anger      | 1     | 4   | 1         | 5    | 30    | 4       | 66.67%   |
| Disgust    | 2     | 4   | 0         | 2    | 4     | 33      | 73.33%   |

| Mean accuracy | 75.56% |

From the confusion matrix test using 270 times the detection, the highest accuracy in the 5x5 and 7x7 filters was 97.41%. The causes of errors in classifying using the Markov Stationary Feature - Vector Quantization method because it is very dependent on the
characteristics of the image and the amount of data training. From the test results show that the level of accuracy of recognition increased compared to previous studies [3][4][5].

| Expression | Result | Accuracy |
|------------|--------|----------|
| Happy      | 45     | 0        | 0        | 0        | 0        | 0        | 100.00%  |
| Sad        | 0      | 45       | 0        | 0        | 0        | 0        | 100.00%  |
| Surprise   | 1      | 0        | 44       | 0        | 0        | 0        | 97.78%   |
| Fear       | 0      | 0        | 0        | 45       | 0        | 0        | 100.00%  |
| Anger      | 0      | 1        | 0        | 0        | 43       | 1        | 95.56%   |
| Disgust    | 0      | 1        | 0        | 0        | 3        | 41       | 91.11%   |

Mean accuracy 97.41%

| Expression | Result | Accuracy |
|------------|--------|----------|
| Happy      | 45     | 0        | 0        | 0        | 0        | 0        | 100.00%  |
| Sad        | 0      | 45       | 0        | 0        | 0        | 0        | 100.00%  |
| Surprise   | 1      | 0        | 44       | 0        | 0        | 0        | 97.78%   |
| Fear       | 1      | 0        | 0        | 44       | 0        | 0        | 97.78%   |
| Anger      | 0      | 0        | 0        | 44       | 1        | 97.78%   |
| Disgust    | 0      | 1        | 0        | 0        | 3        | 41       | 91.11%   |

Mean accuracy 97.41%

4. Conclusion
Based on the results of tests that have been conducted, the level of accuracy achieved using the MSF-VQ method reached 97.14%. This shows that the MSF-VQ method can be recommended as a good method for facial expressions recognition. For further research that can be developed to improve the accuracy and speed of the process by segmenting the eyes, nose and mouth.

Acknowledgment
Authors acknowledged UNIKOM for supporting this research.

References
[1] Calvo, M. G., & Nummenmaa, L. (2016). Perceptual and affective mechanisms in facial expression recognition: An integrative review. Cognition and Emotion, 30(6), 1081-1106.
[2] Ding, C., & Tao, D. (2016). A comprehensive survey on pose-invariant face recognition. ACM Transactions on intelligent systems and technology (TIST), 7(3), 37.
[3] Mujib, K., Hidayatno, A., & Prakoso, T. (2018). Pengenalan Wajah Menggunakan Local Binary Pattern (Lbp) Dan Support Vector Machine (Svm). Transient, 7(1), 123-130.
[4] Hasdiansyah, G., Novianty, A., & Anbarsanti, N. (2016). Perancangan Dan Implementasi Sistem Pengenalan Bahasa Isyarat Indonesia Dengan Mengkombinasikan Rgb Dan Skeleton Kinect Menggunakan Hidden Markov Model. eProceedings of Engineering, 3(2).

[5] Uddin, M. Z., Hassan, M. M., Almogren, A., Alamri, A., Alrubaian, M., & Fortino, G. (2017). Facial expression recognition utilizing local direction-based robust features and deep belief network. IEEE Access, 5, 4525-4536.

[6] Yan, Y. (2016, November). Improved face recognition algorithm using extended vector quantization histogram features. In 2016 IEEE 13th International Conference on Signal Processing (ICSP) (pp. 1046-1050). IEEE.

[7] Siagian, E. F. (2018). Sistem Pendeteksi Kantuk pada Pengendara Mobil Menggunakan Haar Cascade Classifier dan Sobel Edge Filtering.

[8] Suryowinoto, A., & Hamid, A. (2017). Penggunaan Pengolahan Citra Digital Dengan Algoritma Edge Detection Dalam Mengidentifikasi Kerusakan Kontur Jalan. Institut Teknologi Adhi Tama Surabaya.

[9] Tiwari, A., Sharma, M., & Tamrakar, R. K. (2017). Watermarking based image authentication and tamper detection algorithm using vector quantization approach. AEU-International Journal of Electronics and Communications, 78, 114-123.

[10] Jain, V. (2017). Perspective analysis of telecommunication fraud detection using data stream analytics and neural network classification based data mining. International Journal of Information Technology, 9(3), 303-310.