2019 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)

Multiple Face Image Feature Extraction Using Geometric Moment Invariants Method

1st Muhammad Fachrurrozi
Informatics Engineering
Universitas Sriwijaya
Palembang, Indonesia
mfachrz@unsri.ac.id

2nd Saparudin
Informatics Engineering
Universitas Sriwijaya
Palembang, Indonesia
saparudin@unsri.ac.id

3rd Ayu Lestari
Informatics Engineering
Universitas Sriwijaya
Palembang, Indonesia
lestariayu669@gmail.com

4th Osvari Arsalan
Informatics Engineering
Universitas Sriwijaya
Palembang, Indonesia
osvari2017@gmail.com

5th Samsuryadi
Informatics Engineering
Universitas Sriwijaya
Palembang, Indonesia
syamsuryadi@unsri.ac.id

6th Ermatita
Information System
Universitas Sriwijaya
Palembang, Indonesia
ermatita@unsri.ac.id

***Abstract*— Research on human facial expression recognition has become a growing field. One important step in the recognition of facial expressions is feature extraction. This research uses Geometric Moment Invariants (GMI) as a feature extraction method. Research on facial expression recognition using either the GMI method or another method use single face image as the dataset. Therefore, in this study uses GMI feature extraction to classify facial expressions on multiple face images. Face detection process uses Viola-Jones method on OpenCV and classification process uses Multi Class SVM method. The results are features for each expression and a small average accuracy of 5 times. Therefore, the classification is also done with the k-fold cross validation technique with another classification method. The average accuracy results are still small. It caused by the training image also using outer area of face in the image, so the background included as the image features. It is tested from k value 2 to10, and produce Multi Class SVM 10.2%, Decision Tree Classifier 14.73%, Random Forest Classifier 14.78%, Gaussian Naive Bayes 14.73%, Nearest Centroid 14.66%, MLP Classifier 11.09%, and Stochastic Gradient Descent Classifier 14.19%. The highest accuracy result is Random Forest Classifier method 14.78%. In Random Forest method, the best k value obtained is 4 with an average accuracy 16.18%.**

Keywords*— *Geometric Moment Invariants, Feature Extraction, Facial Expressions, Multiple Face

I. Introduction

Research in facial expression recognition is now growing. The facial expression of recognition consists of several components, namely face detection, facial feature extraction, and facial expressions classification [1]. One of component that being focus of this research is feature extraction. Feature extraction method used is Geometric Moment Invariants (GMI) method. The GMI method has invariant values that are not affected by deformation changes and get high object recognition results. Several studies on facial expression recognition using both the GMI method and other methods, they use single face image dataset. Basu et al [2] and Geethu & Kamatchi [3] use the Multi Class SVM classification method, as well as Tsai & Chang who use the SVM classification method. It begs the question, what if facial expression recognition uses multiple face image dataset. Therefore, this research aims to apply the GMI method in feature extraction on multiple face images for classification of facial expressions. The expected results are collection of features for each expression and the results of the classification of facial expressions. Research on facial expression recognition uses a lot of single-face images as data for the training and testing process. This raises the question, what is the result of facial expression feature extraction if the image used has multiple faces. Therefore, in this research, Geometric Moment Invariants feature extraction will be applied in the classification of facial expressions on multiple face images.

II. Related Work

Research on human facial expression recognition is growing. One important step in the introduction of facial expressions is feature extraction [1]. The feature extraction method used in this study was Geometric Moment Invariants (GMI) method. These following are the research that uses GMI method. Research by Samad, Haq, & Khan [4] which resolves deformation problems in object recognition. The results of his research get 96% average accuracy. The research by Zhang et al [5] carried out feature extraction of brain images in the process of detection of alcoholic patients. The results of his research get 91% accuracy. In addition, research by Lukic, Tuba, and Tuba [6] about the leaf recognition, feature of leaf derived from combination of GMI method with Local Binary Patterns method, that research obtained 94.15% accuracy. The research on human facial expressions recognition on single-face images was carried out by Basu et al [2], who obtained 87.5% accuracy. Based on the description above, it can be concluded that using GMI method got high accuracy in object recognition. Therefore, this research will focus on extracting GMI features in classification of facial expressions. But research on facial expression recognition, either using feature extraction of GMI or not, the research used a single-face image. The research used single-face images as dataset in facial expressions recognition conducted by Basu et al [2], Geethu & Kamatchi [3] and Tsai & Chang [1]. This bring out question, what if the human facial expression recognition uses multiple face images.

III. Facial Expession Classification

A. Face Detection

The face detection method in this research uses Viola-Jones on OpenCV. The Viola-Jones method is a face detection method that provides high result of face detection [7]. Fig.1 shows a flow of the Viola-Jones method in face detection.

Image pre-processing used for image changes according to feature extraction requirements. In this research required binary image.

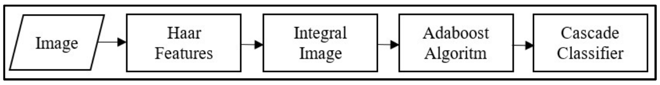


FIGURE 1. PROCESS FLOW OF VIOLA-JONES METHOD

B. *Facial Feature Extraction with the Geometric Moment Invariants (GMI)*

Feature extraction is process of taking important features of an object that distinguish an object class with other classes. In 1962, Hu [8] first introduced an invariant moment. The invariant value, a value that does not change from image that go through certain deformations [9]. The following are step to calculate the GMI method.

Feature extraction is process of taking important features of an object that distinguish an object class with other classes. In 1962, Hu first introduced an invariant moment. The invariant value, a value that does not change from image that go through certain deformations [9]. The following are step to calculate the GMI method.

- Calculate the Value Moments
- The GMI method begins by calculating the two-dimensional moment value with an order $(p + q)$ from a digital image with an MxN size defined as:

$$m_{pq} = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} x^p y^q \ f(x,y) \tag{1}$$

where m_{pq} is a two-dimensional moment of the function *f(x,y)*, *x* is the image coordinate on *x* axis, *y* is the image coordinate on *y* axis, *p* and *q* is the order moment = 0, 1, 2 and 3, *f(x,y)* is the intensity of the binary value at pixels between 0 and 1, *dx* is the large shift in the direction of *x*, and *dy* is the large shift in the direction *y*.

- The next moment is a central moment that is invariant to translation [5]. At the central moment *f(x, y)* is defined as:

$$\mu_{pq} = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} (x-\bar{x})^{\,p} \, (y-\bar{y})^{\,q} \, f(x,y) \tag{2}$$

$$\bar{x} = \frac{m_{10}}{m_{00}} \tag{3}$$

$$\bar{y} = \frac{m_{01}}{m_{00}} \tag{4}$$

where $\bar{x}$ and $\bar{y}$ are the mean of central point, as well as m_{10}, m_{01}, and m_{00} : order of moments.

- The next moment is a normalized central moment. This moment is invariant for scaling [5]. The normalized central moment is defined as:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}} \tag{5}$$

$$\gamma = \frac{p+q}{2} + 1 \tag{6}$$

Where, η_{pq} is the normalized central moment at orders p and q, μ_{pq} is the central moment in orders p and q, p and q are order of moments = 0, 1, 2 and 3, and γ is the value of scaling power.

- Calculate the Seven Features of GMI

The last moment added rotation invariant [5]. The seven moments of Geometric Moment Invariants are defined as follows [10]:

$$\phi_1 = \eta_{20} + \eta_{02} \tag{7}$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4{\eta_{20}}^2 \tag{8}$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \tag{9}$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \tag{10}$$

$$\begin{aligned}\phi_5 = (\eta_{30} - \ 3\eta_{12})(\eta_{30} \\ + \eta_{12})\,[(\eta_{30} + \eta_{12})^2 \\ - 3(\eta_{21} + \eta_{03})^2] \\ + (3\eta_{21} - \eta_{03})(\eta_{21} \\ + \eta_{03})[3(\eta_{30} + \eta_{12})^2 \\ - (\eta_{21} + \eta_{03})^2]\end{aligned} \tag{11}$$

$$\begin{aligned}\phi_6 = (\eta_{20} - \eta_{02})\,[(\eta_{30} + \eta_{12})^2 \\ - (\eta_{21} + \eta_{03})^2] \\ + 4\eta_{11}[(\eta_{30} + \eta_{12}) \\ - (\eta_{21} + \eta_{03})]\end{aligned} \tag{12}$$

$$\begin{aligned}\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} \\ + \eta_{12})\,[(\eta_{30} + \eta_{12})^2 \\ - 3(\eta_{21} + \eta_{03})^2] \\ + (3\eta_{12} - \eta_{30})(\eta_{21} \\ + \eta_{03})[3(\eta_{30} \\ + \eta_{12})^2 \\ - (\eta_{21} + \eta_{03})^2]\end{aligned} \tag{13}$$

where ϕ_1 , ϕ_2 , ϕ_3 , ϕ_4 , ϕ_5 , ϕ_6 , ϕ_7 is moment invariant, η_{pq} is normalized central moment at orders p and q, and p and q are order of moments = 0, 1, 2 and 3.

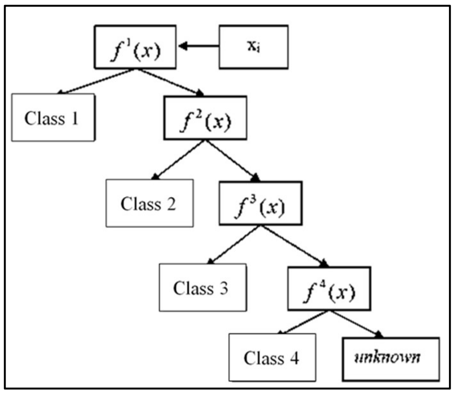


FIGURE 2. EXAMPLE OF CLASSIFICATION WITH A 1-AGAINST-ALL SCHEME

C. *Classification of Facial Expressions*

Basically SVM is a binary classifier, which classifies data in two classes [11]. If there are more than two classes, then the SVM Multi-class method is used. The SVM Multi-class method that will be used in this research has a 1-against-all scheme. The 1-against-all scheme is a scheme where there is one binary SVM for each class to separate the class members from other class members. Fig. 2 shows the 1-against-all scheme [11].

From the discussion above, the design of this research process is shown in Fig. 3.

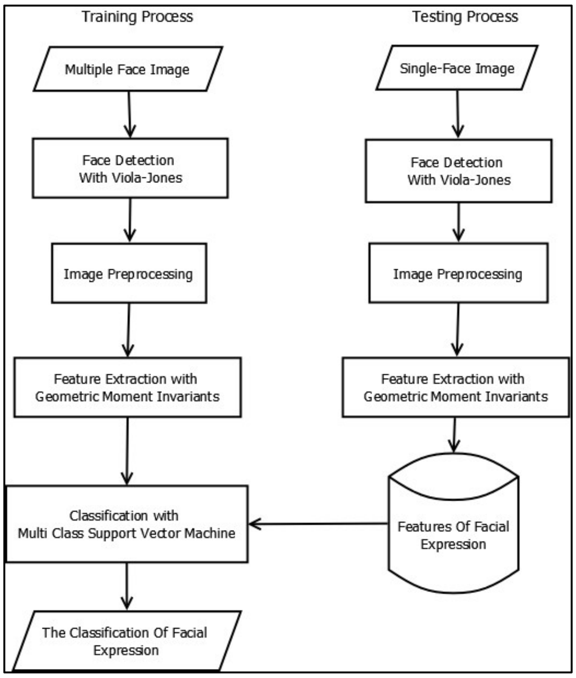


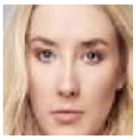
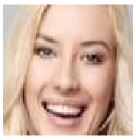
FIGURE 3. CURRENT PROPOSED DESIGN PROCESS

This research consists of two processes, namely the training process and the testing process. In the training process, a single face image is done by face detection. Then, the face is pre-processed into a binary image and then extracted features. Then, a collection of features is used as class model of facial expression, which will be used in the process of classifying facial expressions. In the testing process, face detection on multiple face images. Then, the face is pre-processed into a binary image. Then, the face image becomes input for feature extraction process. The features obtained from previous process, become a parameter in classification of facial expressions. After the classification process, the results of expressions of the faces detected will be known.

IV. RESULTS

Data are single face and multiple face images. Single face images from the Extended Cohn-Kanade / CK + dataset [12]. The data used as many as 68 subjects for each expression of happiness, sadness, anger, disgust, shock, fear, and natural with size of 640 x 490 pixels. Multiple face image data taken from the *https://www.shutterstock.com/* site as many as 100 data. In the first test, 100 multiple face images were tested 5 times. The test results can be seen in Tab 1. Some of the results of the first test in the first experiment can be seen in Tab 2 and the results of the extraction of the characteristics of the second test data can be seen in Tab 3. Visualization of distribution of training data points can be seen in Fig. 4.

TABLE 1. RESULTS OF FEATURE EXTRACTION OF GMI METHOD IN FIRST TEST DATA

Images	Features	Results
	1	0.0007752348506367586
	2	2.133888262004688e-09
	3	8.00390223998433e-13
	4	1.982532313942676e-12
	5	-2.4936863415611943e-24
	6	-9.074632869324559e-17
	7	1.3542401859574542e-25
	1	0.000753409234923105
	2	6.91084334018544e-10
	3	4.90897852386785e-13
	4	1.40268262616154e-12
	5	-4.07626685871931e-25
	6	-3.14892386568329e-17
	7	1.09023959934028e-24

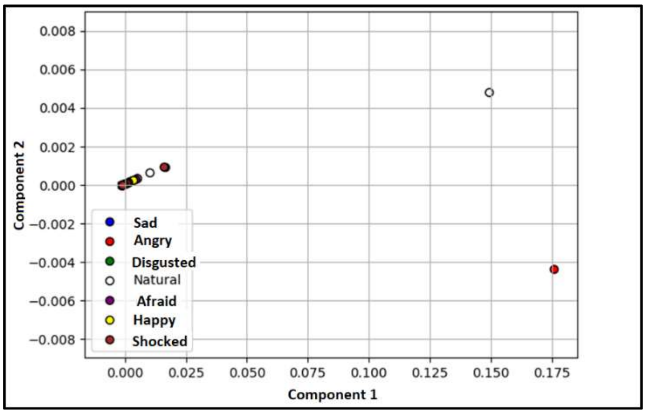


FIGURE 4. DISTRIBUTION OF TRAINING DATA POINTS ON 2D CHARTS

WS is the face accuracy detected, B is the accuracy of happy expressions, S is the accuracy of sad expressions, M is the accuracy of angry expressions, J is the accuracy of disgusted expressions, K is the accuracy of shocked expressions, T is the accuracy of fear expressions, and N is the accuracy of expression natural.

In Tab. 2, it can be concluded that the results of each experiment in the first test get an average face detection accuracy of 96.02% and each expression has a small average accuracy value except sad expressions. Sad expressions get the accuracy of 77.26%. The expressions that are not properly classified are expressions of disgust, shock, fear, and natural. In addition, the average accuracy of happy expressions was 18.74%, and angry expressions was 1.09%. The results of the classification of facial expressions on multiple face images obtained small accuracy. This can occur due to the following factors.

TABLE 2. AVERAGE ACCURACY OF FIRST TEST RESULTS ON EACH EXPERIMENT

Average Accuracy of Each Expression					Total Average Accuracy
Experiment 1	Experiment 2	Experiment 3	Experiment 4	Experiment 5	
WS=96,02% B=22,17% S=77,89% M=1,09% J=0,0% K=0,0% T=0,0% N=0,0%	WS=96,02% B=17,46% S=79,82% M=1,09% J=0,0% K=0,0% T=0,0% N=0,0%	WS=96,02% B=20,25% S=71,49% M=1,09% J=0,0% K=0,0% T=0,0% N=0,0%	WS=96,02% B=14,71% S=76,67% M=1,09% J=0,0% K=0,0% T=0,0% N=0,0%	WS=96,02% B=19,1% S=80,42% M=1,09% J=0,0% K=0,0% T=0,0% N=0,0%	WS=96,02% B=18,74% S=77,26% M=1,09% J=0,00% K=0,00% T=0,00% N=0,00%

- In the process of face detection using Viola-Jones method on OpenCV. The face that is detected is a little incorrect because there are parts that detected but not needed such as hair and image background. This can affect feature extraction which should be extracted to distinguish each class. An example can be seen in Tab 1.
- Using the Geometric Moment Invariants feature extraction method in this case of facial expressions is not strong enough to take the features that represent each expression. That is because in the facial expressions recognition, there are micro expressions such as eyes, nose, mouth, eyebrows, and others that are not involved in this research.
- Seen from Fig. 4, the data distribution features of Geometric Moment Invariants in each class have similarities.

The first test results get the classification accuracy of small facial expressions. So, the second test was also carried out. In the second test, the training data will be tested using the k-fold cross validation technique. Training data totalled 68 single face images for each facial expression. So that the number of training data is 476 single face images. The test results with the k-fold cross validation technique with several classification methods can be seen in Tab 3. Visualization of the comparison between accuracy and the method used can be seen in Fig. 5.

TABLE 3. SECOND TEST RESULTS

Classification Method	Average Results of Accuracy of Each K-Fold Cross Validation (%)									Average (%)
	K=2	K=3	K=4	K=5	K=6	K=7	K=8	K=9	K=10	
MCS	12,4	10,92	11,34	10,93	10,29	9,03	9,87	7,99	9,04	10,2
DT	14,92	13,23	15,76	14,91	14,7	14,92	15,34	14,28	14,5	14,73
RF	15,13	13,03	16,18	15,13	14,08	14,71	15,76	15,12	13,88	14,78
GNB	14,71	15,34	14,08	13,87	15,14	14,92	14,71	14,71	15,13	14,73
NC	15,34	15,12	14,29	13,45	15,77	14,5	14,29	14,71	14,49	14,66
MLP	12,6	10,92	10,92	10,93	10,5	11,34	12,18	9,87	10,52	11,09
SGD	14,29	14,28	14,29	13,87	14,27	14,29	15,75	12,81	13,84	14,19

GNB is Gaussian Naive Bayes, NC is Nearest Centroid, MLP is Multi-Layer Perceptron Classifier, and SGD is Stochastic Gradient Descent Classifier.

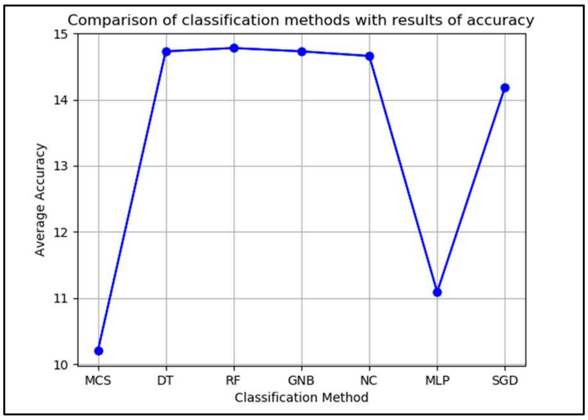


FIGURE 5. COMPARISON OF CLASSIFICATION METHODS WITH RESULTS OF ACCURACY

The classification methods used in the facial expressions classification also get small average accuracy. The average results of the accuracy obtained were the results of the average accuracy of k = 2 to k = 10 obtained, namely MC 10.2%, DC 14.73%, RF 14.78%, Gaussian Naive Bayes 14.73%, Nearest Centroid 14.66%, MLP Classifier 11.09%, and Stochastic Gradient Descent Classifier 14.19%. The highest accuracy results are in the Random Forest Classifier method of 14.78%. In the Random Forest method, the best k value obtained is 4 with an average accuracy of 16.18%.

V. CONCLUSION

Facial image feature extraction was carried out using the Geometric Moment Invariants method to classify facial expressions on multiple face images. The results of feature extraction using the GMI method for facial expressions have similar values. Therefore, the average accuracy of classification of facial expressions obtained is small. It caused by the training image also using outer area of face in the image, so the background included as the image features. In the test results obtained in the form of characteristics and average accuracy of each expression. Expressions that got the greatest accuracy were sad expressions of 77.26%. This study also conducted a second test with the k-fold cross validation technique from the values k = 2 to k = 10 which was implemented with several classification methods. The second test results obtained the highest average accuracy on the Random Forest Classifier method which is 14.78%.

REFERENCES

[1] H.-H. Tsai and Y.-C. Chang, “Facial expression recognition using a combination of multiple facial features and support vector machine,” *Soft Comput.*, vol. 22, no. 13, pp. 4389–4405, 2018.

[2] A. Basu, A. Routray, S. Shit, and A. K. Deb, “Human emotion recognition from facial thermal image based on fused statistical feature and multi-class SVM,” in *2015 Annual IEEE India Conference (INDICON)*, 2015, pp. 1–5.

[3] Geethu G S and K. T, “Recognition of Facial Expressions in Image Sequence using Multi-Class SVM,” *Int. J. Innov. Res. Comput. Commun. Eng.*, vol. 4, no. 8, pp. 14630–14638, 2016.

[4] S. Samad, A. Haq, and A. Shoab, “Orientation Invariant Object Recognitions Using Geometric Moments Invariants and Color Histograms,” *Int. J. Comput. Electr. Eng. Orientat.*, vol. 7, no. 2, pp. 101–108, 2015.

[5] Y.-D. Zhang *et al.*, “Alcoholism detection by medical robots based on Hu moment invariants and predator–prey adaptive-inertia chaotic particle swarm optimization,” *Comput. Electr. Eng.*, vol. 63, pp. 126–138, 2017.

[6] M. Lukic, E. Tuba, and M. Tuba, “Leaf recognition algorithm using support vector machine with Hu moments and local binary patterns,” in *2017 IEEE 15th International Symposium on Applied Machine Intelligence and Informatics (SAMI)*, 2017, pp. 485–490.

[7] P. Viola and M. J. Jones, “Robust real-time face detection,” *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, 2004.

[8] Ming-Kuei Hu, “Visual pattern recognition by moment invariants,” *IRE Trans. Inf. Theory*, vol. 8, no. 2, pp. 179–187, Feb. 1962.

[9] J. Flusser, B. Zitova, and T. Suk, *Moments and Moment Invariants in Pattern Recognition*. Wiley Publishing, 2009.

[10] R. C. Gonzalez and R. E. Woods, *Digital image processing*. Upper Saddle River, N.J.: Prentice Hall, 2008.

[11] D. Verma and D. Mukhopadhyay, “Age driven automatic speech emotion recognition system,” in *2016 International Conference on Computing, Communication and Automation (ICCCA)*, 2016, pp. 1005–1010.

[12] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, “The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression,” in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, 2010, pp. 94–101.

978-1-7281-2930-3/19/$31.00 ©2019 IEEE

978-1-7281-2930-3/19/$31.00 ©2019 IEEE