Multi-Object Face Recognition Using Content Based Image Retrieval (CBIR)

Muhammad Fachrurrozi
Informatics Engineering Department
Faculty of Computer Science, Universitas Sriwijaya
Palembang, Indonesia
mfachrz@unsri.ac.id

Erwin
Computer Engineering Department
Faculty of Computer Science, Universitas Sriwijaya
Palembang, Indonesia
erwin@unsri.ac.id

Saparudin
Informatics Engineering Department
Faculty of Computer Science, Universitas Sriwijaya
Palembang, Indonesia
saparudin@unsri.ac.id

Mardiana
Law Department
Faculty of Law, Universitas Sriwijaya
Palembang, Indonesia
mardiana_rachman@yahoo.com

***Abstract*— Real-time face recognition system process divided into three steps, feature extraction, clustering, detection, and recognition. Each step uses a different method that is Local Binary Pattern (LBP), Agglomerative Hierarchical Clustering (AHC) and Euclidean Distance. Content Based Image Retrieval (CBIR), an image searching techniques based on image feature, is implemented as the searching method. Based experiments and the testing result, recall and precision values are 65.32% and 64.93% respectively.**

***Keywords—Face Recognition; LBP; AHC; Real-time; Multi-object; CBIR; Euclidian Distance*;**

I. INTRODUCTION

In the multi-object image, human face recognition is one of the things that can be identified or recognized. In addition to fingerprints, eyes, and sounds, biological information possessed by humans is the face. In real-time image recognition process, firstly the system will perform face detection process by using Open-cv Library. The detected face then extracted using the Local Binary Pattern (LBP). The vector features resulted from the extracted training image are grouped using Agglomerative Hierarchical Clustering (AHC), and then compared with the vector feature of test image using Euclidean distance. This process based on CBIR because to search the vector features in the database is done by looking the visual content of the image itself [1].

In order to get the good result, CBIR has several processes, one of them is to extract the features to get the values of human face characteristic. Local Binary Pattern (LBP) algorithm is able to characterize and distinguish surface texture. This algorithm can achieve a high accuracy if implemented on a texture with low variance [2]. In face recognition system, feature extraction uses LBP based on the texture of the face image.

Hierarchical Clustering algorithm implemented as a clustering method because it can improve the speed and accuracy of image matching in CBIR [3] [4]. In addition, Agglomerative Hierarchical Clustering methods used in this face recognition system are Single Linkage, Complete Linkage, and Average Linkage.

Euclidean distance accuracy to search the minimum distance between test image and training image from eigenface extraction result is 95% [5]. In this research, Euclidean Distance used to search distance between testing image vector feature and training image feature vector.

II. FACE RECOGNITION

A. Local Binary Pattern (LBP)

LBP represents a pixel which formed by a 3x3 matrix as a comparison between the center pixel and its surrounding pixel which then converted into binary numbers. The comparison assumes that if the surrounded pixel value is greater than the central pixel value than it will be 1 otherwise 0. After we get 8 binary numbers in each pixel then it will be replaced with the decimal form to get the result.

The LBP algorithm equation can be expressed as the following formula:

$$LBP(x_c, y_c) = \sum_{p=0}^{7} f\,(g_p - g_c)2^p \quad (1)$$

Information:

g_p : central pixel value

g_c : the pixel value around the center

p : number of pixels around the center

And the function f (x) is defined as follows:

$$f(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases}$$

B. *Hierarchical Clustering*

Hierarchical clustering is used to group similar imagery into clusters to increase the speed of searching image [2]. In this algorithm, there is a hierarchal tree which provides a view of several levels of data abstraction known as dendogram[3]. In Hierarchical Clustering, there are two ways to cluster the data, namely *agglomerative* and *divisive.* Agglomerative clustering process based on the amount of data grouped into some hierarchy, it will be clustered into a hierarchical unit. Divisive performs the reverse process of agglomerative, that is cluster one hierarchy into several hierarchies. The description of hierarchical clustering can be seen in Fig. 1.

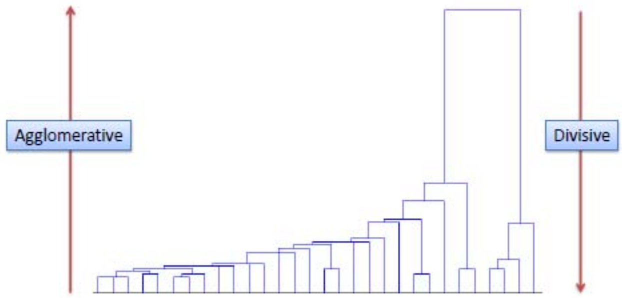

Fig. 1. *Hierarchical Clustering*

Agglomerative Hierarchical clustering is a clustering algorithm based on the proximity distance between two images into a hierarchy. This process repeats itself until it gets some hierarchy. The hierarchy with the closest distance is combined into one hierarchy. The proximity to the new hierarchy then recalculated and the closest hierarchy is merged again. The process is repeated until all the data (object) clustered into one hierarchy.

Calculating the spacing between two images using the *Manhattan Distance* formulated in equation (2):

$$d = \sum_{i=1}^{n} |X_i - y_i| \qquad (2)$$

where:

d : the distance between the image of x and y.

n : number of variables.

x_i : the value of x on i variable.

y_i : the value of y on i variable

The distance between images is written into a matrix called distance matrix. In order to determine the distance between the two clusters, Agglomerative Hierarchical clustering has 3 methods of grouping data, namely:

1. *Single Linkage*

Single Linkage classifies data based on the closest distance (Min) between the hierarchy. Single Linkage can be formulated in equation (3):

$$d(uv)w = Min[d(uw), d(vw)] \qquad (3)$$

where :

$d(uv)w$: the distance between the hierarchy uv and w.

$d(uw)$: the distance between the hierarchy u and w.

$d(vw)$: the distance between the hierarchy v and w.

2. *Complete Linkage*

Complete Linkage categorizes data by the furthest distance (Max) or the maximum distance between hierarchies. Complete Linkage can be formulated in equation (4):

$$d(uv)w = Max[d(uw), d(vw)] \qquad (4)$$

where:

$d(uv)w$: the distance between the hierarchy uv and w.

$d(uw)$: the distance between the hierarchy u and w.

$d(vw)$: the distance between the hierarchy v and w.

3. *Average Linkage*

Average Linkage classifies data based on the average distance between the hierarchy. Average Linkage can be formulated in equation (5):

$$d(uv)w = \frac{d(uw)+d(vw)}{2} \qquad (5)$$

where:

$d(uv)w$: the distance between the hierarchy uv and w.

$d(uw)$: the distance between the hierarchy u and w.

$d(vw)$: the distance between the hierarchy v and w.

C. *Euclidian Distance*

Euclidean distance is an algorithm to calculate the difference or minimum distance between the test image and training image. An Object that will be used as a recognition object is the one which has the minimum distance. To calculate the minimum distance between the vector value of the test image and the training image present in the database, the Euclidean distance algorithm equation can be expressed by the formula:

$$J(A,B) = \sqrt{\sum_{e=1}^{p}(a_e - b_e)^2} = \|A - B\| \quad (6)$$

which :

J(A, B): the distance between imagery A and B

A: the value of the image vector present in the database

B: value of the test image vector

P: number of variables

a_e : value a on the e-variable

b_e : value b on the e-variable

III. METHODOLOGY

Figure 2 is a block diagram of a multi-object face recognition system based on image retrieval, where the path of the system is divided into 2 stages namely the phase of data capture phase of face recognition.

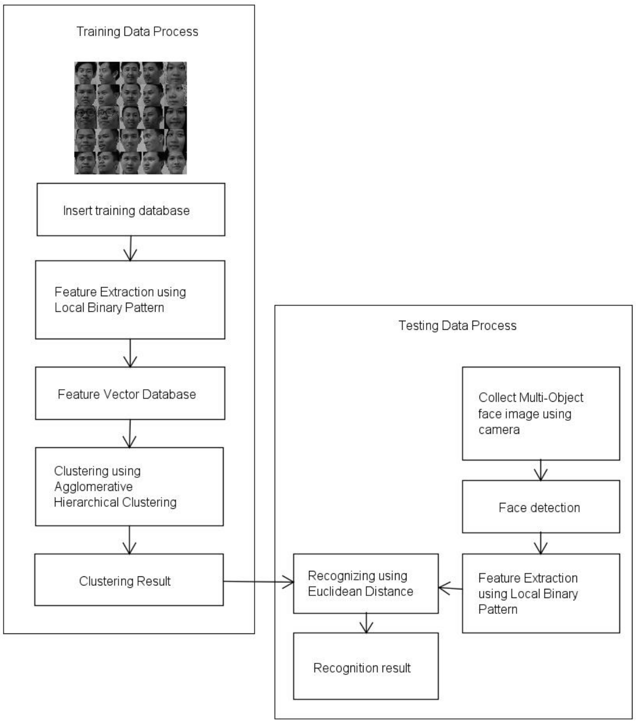

Fig. 2. General System Diagram

In general, the steps of data retrieval in this research is as follows:

1. Collect face image.
2. The feature extraction process using LBP method to get the characteristic of face image then transformed into vector feature form which will be stored in the database.
3. Then do the clustering process using the AHC method on the vector of the face image in the database. Clustering results that have been obtained later used as a comparator value of calculating the distance for face recognition.

In the face recognition phase the process of the system is as follows:

1. Open webcam to detect faces. The process of face detection and recognition is done in real-time.
2. The detected face then captured.
3. The feature extraction process is done using Local Binary Pattern method to get the feature of the face image then transformed into the vector feature form.
4. Then, do the process of recognition by calculating the distance between the new face image features and features of the existing on the database by using Euclidian distance which then matched with the clustering results.

IV. IMPLEMENTATION AND RESULT

Data used as many as 170 images from 17 people, each person took 10 images with different sides and same background. The image is used as a training data with dimensions 150x150 pixels. Examples of face image data can be seen in Figure 3.

Fig. 3. Example of Image Face Data

In Figure 4 is the main form of the system, where the system has two menus that are browse and start the camera. Browse menu is a menu to do the recognition in offline and start menu camera is a menu to do recognition in real-time.

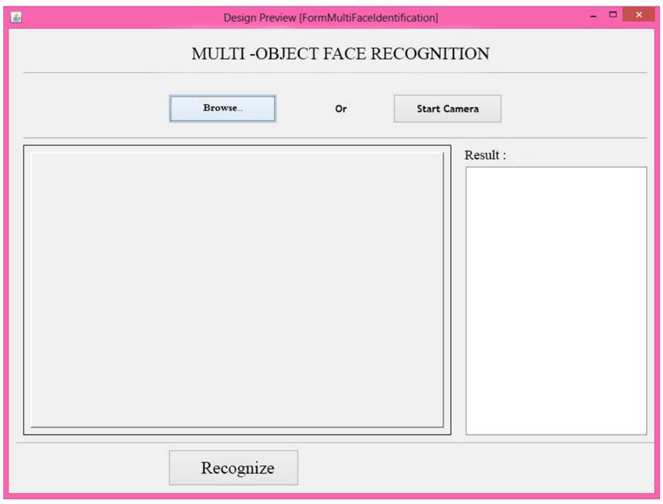

Fig. 4. Main Form

Face image recognition done by comparing the distance from the test vector image with training image feature value using Euclidean. The smallest distance of vector feature between the test image and the training image will determine the outcome of the recognition. Figure 5 is a multi-object face image recognition form.

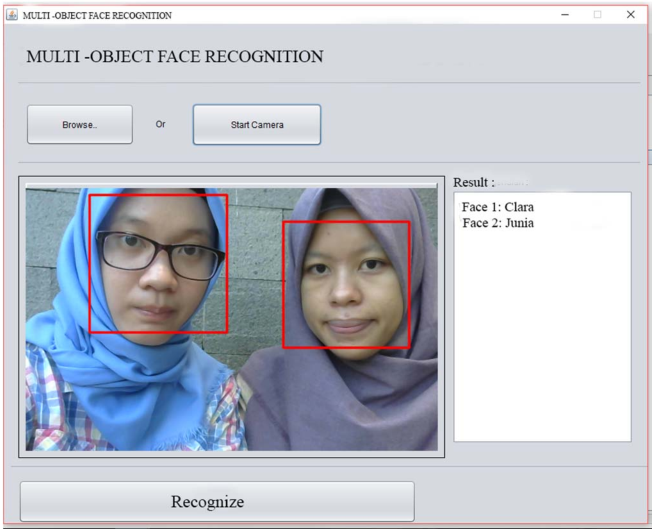

Fig. 5. Face Image Recognition Form

Image detection and recognition testing were are done by calculating the value of recall and precision by using 4 attributes ie True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

TABLE I. TESTING OF OFFLINE FACE DETECTION

Object category	Mean (%)	
	Recall	Precision
Object 1	90%	90%
Object 2	88%	85%
Object 3	86,7%	85,5%
Object 4	86,5%	84,7%
Object 5	85,4%	83%
Object 6	78,7%	76,3%
Average total	85,88%	84,08%

TABLE II. TESTING OF REAL – TIME FACE DETECTION

Object category	Mean (%)	
	Recall	Precision
Object 1	83,4%	83,4%
Object 2	74,2%	72,3%
Object 3	71,8%	70,5%
Object 4	71,8%	70,3%
Object 5	74,5%	72,6%
Object 6	70,6%	70,2%
Average total	73,38%	72,9%

Based on testing table 1 for offline face detection, it is known that the recall and precision values are 85.88% and 84.08%. While for testing table 2 face detection in real-time, recall and precision value is 73,38% and 72,9%.

TABLE III. TESTING OF OFFLINE FACE DETECTION

Object category	Mean (%)	
	Recall	Precision
Object 1	93%	93%
Object 2	82,3%	81,8%
Object 3	82,2%	82%
Object 4	81,4%	80,7%
Object 5	80,9%	80,2%
Object 6	79,3%	78,7%
Average total	83,2%	82,73%

TABLE IV. TESTING OF REAL-TIME FACE DETECTION

Object category	Mean (%)	
	Recall	Precision
Object 1	70%	70%
Object 2	68,9%	68,3%
Object 3	65%	64,6%
Object 4	64,5%	63,9%
Object 5	62,8%	62,3%
Object 6	60,7%	60,5%
Average total	65,32%	64,93%

Based on testing table 3 for offline face recognition, it is known that the value of recall and precision is 83.2% and 82.73%. While for testing 4 face recognition table in real-time, recall and precision value is 65,32% and 64,93%.

V. CONCLUSION

From the results of tests conducted, it can be concluded that:

1. The system can recognize a single or multi object with value of recall and precision equal to 65,32% and 64,93%.
2. The process of feature extraction of the input image plays an important role in determining the success rate of face image recognition.
3. In the testing process, the system could recognize successfully if the testing data has the same lighting, distance, and other effects during training.

REFERENCE

[1] D. H. Z. and P. D. D. F. Dr. Fuhui Long, "Fundamentals of Content-Based Image Retrieval," *Multimed. Inf. Retr. Manag. Technol. Fundam.*, pp. 1–26, 2003.

[2] E. Prakasa, "Ekstraksi Ciri Tekstur dengan Menggunakan Local Binary Pattern Texture Feature Extraction by Using Local Binary Pattern," vol. 9, no. 2, pp. 45–48, 2016.

[3] S. Pandey, P. Khanna, and H. Yokota, “Clustering of hierarchical image database to reduce inter-and intra-semantic gaps in visual space for finding specific image semantics,” *J. Vis. Commun. Image Represent.*, vol. 38, pp. 704–720, 2016.

[4] V. S. V. S. Murthy, E. Vamsidhar, J. N. V. R. S. Kumar, and P. Sankara Rao, “Content Based Image Retrieval using Hierarchical and K-Means Clustering Techniques,” *Int. J. Eng. Sci. Technol.*, vol. 2, no. 3, pp. 209–212, 2010.

[5] M. üg. Çarıkçı and F. Özen, “A Face Recognition System Based on Eigenfaces Method,” *Procedia Technol.*, vol. 1, pp. 118–123, 2012.

[6] A. Katare, S. K. Mitra, A. Banerjee, D. Microsystems, and P. Limited, “Content Based Image Retrieval System for Multi Object Images using Combined Features,” 2007.

[7] N. Alajlan, M. S. Kamel, and G. Freeman, “Multi-object image retrieval based on shape and topology,” *Signal Process. Image Commun.*, vol. 21, no. 10, pp. 904–918, 2006.

[8] T. Ahonen, S. Member, A. Hadid, S. Member, and M. Pietika, “Face Description with Local Binary Patterns : Application to Face Recognition,” vol. 28, no. 12, pp. 2037–2041, 2006.