Development of Information Systems to Verify Lecture Presence in the Lecture Room using a Stereo Vision Camera

J Jumi[1], Achmad Zaenuddin[2]
Departement of Bussiness Administration
Politeknik Negeri Semarang
Semarang, Indonesia

Tedjo Mulyono[3]
Department of Civil Engineering
Politeknik Negeri Semarang
Semarang, Indonesia

Abstract—**The face is often used as a feature of human biometrics which can be categorized as unique. This makes the face image as one of the unique fields that can be used as a key field in the detection of image-based personal information. In the process of verifying the attendance of lectures manually through signature detection, it is still possible for counterfeiting and requires a long time. This condition can be anticipated through verification of college attendance using face detection. The conditions during lectures that affect the accuracy of face detection include students changing seat positions, variations in face poses, classroom lighting and the distance of the camera to the object. This requires a method to improve detection accuracy so that college attendance verification has a high level of validation. The method used in this study is face recognition through matching similarity of shapes, colors and textures. Shape feature extraction using invariant moments, color feature extraction using color moments and texture feature extraction using statistical textures. In this study a lecture attendance verification model was developed through the recognition of many faces in the classroom or lecture hall using a stereo vision camera. The test produces a system that has a detection accuracy level that supports verification of the presence of valid lectures. By using a variety of seating positions, variations in poses, variations in illumination brightness (illumination) and variations in the distance of the camera with the object. Matching testing results achieved an accuracy of more than 75%.**

Keywords—*Stereo vision camera; face recognition; illumination brightness; verification; presence*

I. INTRODUCTION

Technology can replace humans in carrying out routine activities with low error rates and is difficult to manipulate because the machine will carry out their duties in accordance with the instructions given. This will support the success of work that requires a high level of honesty such as verification of college attendance. The use of manual presence in college attendance is not safe against manipulation of attendance levels in attending lectures and requires special time to sign.

Biometric characteristics in humans have advantages and disadvantages in its application. The use of face image as the verification key was chosen because the face is a human biometric characteristic that is classified as unique or not the same even if it is born twins. Face detection system is one of the fields of Image Processing and Computer Vision that is currently being developed. This science is the development of a system that approaches the visual system in the human eye. In the face detection system, facial image recognition can be carried out on a still image or a moving recording called a video [2]. Research on facial recognition based on facial symmetry patterns using two-dimensional face images as training images [3]. In this research, a facial image reconstruction with a frontal position through the process of centering and division of the image into two equal parts, namely half the left face image and half the right face image. The reversing process is carried out on one of the facial images and subsequently averaged over the merging of the left image and the right image. The test uses a benchmark database from the Yale database and AR database, but it also compares it with a 3D database from MU-2 stereo imaging from technologies Ltd. (Atlanta, GA). The research shows that facial recognition techniques used are able to work well with each method used. The recognition accuracy reaches a high percentage at the frontal face position or 0^0 angle. This study did not do a test that discusses the variation of illumination, variations in the distance of the camera with the object and variations in the position of the object.

Some research on face image recognition and verification that has been done still uses a single face image input in the sense of not using a lot of face images in a video recording. In addition, also has not used a stereo vision camera as an input image capture media. The study also has not measured accuracy in variations of illumination, variations in the distance of the camera with the object and variations in the position of the object. Accuracy measurement functions to test and determine a reliable model in the verification process. The lecture attendance verification system through the recognition of many facial images in the classroom built in this study is a verification system that has an accuracy of more than 95% recognition and is reliable for variations in illumination, pose variations, variations in camera distance with objects and variations in object positions. Changes or variations caused by light, pose, and position and distance of student seating will change at each lecture meeting.

II. RELATED WORK

A. Image Acquisition

In the first year of research, image capture was carried out using a camera with 12 Megapixel specifications. The resulting face images have different expressions. Examples of data used in this test can be seen in Table 3.1.

TABLE I. TABLE TEST IMAGERY
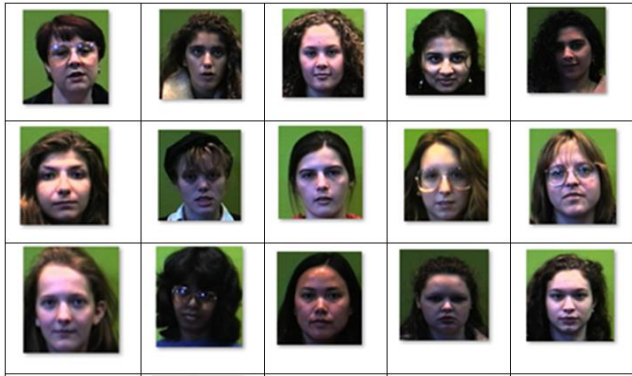
B. Content Based Image Retrieval (CBIR)

Some similar studies relating to image-based image retrieval based on CBIR (Content Based Image Retrieval) using the image feature weighting method include image search using a combination of color features and weighted texture feature vectors [1]. In this research, measurement of image retrieval accuracy has been carried out using Manhattan as a measurement of similarity distance. Weighting 0.1 and 0.2 in color and texture features produces the highest retrieval accuracy.

Search by weighting the texture features using DCT (Discrete Cosine Transform) has also been carried out for CBIR-based image search [8]. The value of the feature will be the weight coefficient of the image that will be used as a basis for searching. Image indexing based on the co-occurrence matrix weighted color feature, results in relatively high search accuracy values [9]. In this research, the weighting of color and indexing feature values is based on the similarity of diagonal and non-diagonal elements.

Preprocessing is the stage of improving image quality before feature extraction with the aim of increasing the accuracy of the image feature extraction results. There are differences in preprocessing in the form feature extraction with the color and texture feature extraction. The difference in preprocessing is to get a quality image before feature extraction. This system has a different preprocessing stage, there are:

C. Preprocessing

1) Preprocessing Shape Features.

a) Resizing : The size of the image will affect the length of feature extraction time, the smaller the size, the faster the computation of feature extraction.

b) Grayscale: Especially in calculating the features of color elements are not taken into account, so that the image is changed to grayscale so that the computation process becomes faster.
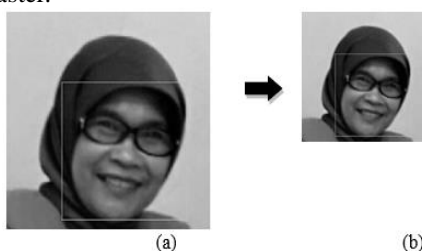

Fig. 1. Preprocessing Stages of Shape Features, (a) Grayscale Image, (b) Resize

2) Preprocessing color features and texture features.

a) Resize: The larger the image size, the longer the extraction time, so that the resize stage is needed to speed up the computation process. At this stage the image is resized to 200 x 200 pixels.

b) Grayscale: Grayscale preprocessing is done for texture feature extraction so that the texture is more visible, while grayscale color feature extraction is not done because the color value that will be taken into account.

c) Histogram Equalization: At this stage histogram leveling is done so that the image quality becomes more contrasted to get the value of quality color and texture features. Histogram leveling results can be seen in Fig. 2 and Fig. 3.

Fig. 2. Comparison of images before and after the histogram process, (a) Original Image, (b) Image of Histogram Equalization (HE)

Fig. 3. Comparison of histograms, (a) Original Image Histogram, (b) Image Histogram of HE Results

D. Feature Extraction of Shape, Color and Texture

Extraction is carried out on all three features with the following stages:

a) Shape Feature Extraction: Shape feature extraction is done using the invariant moment method. This method is used because it is not susceptible to image changes caused by Rotation, Scale and Translation (RST) [8]. the diversity of face images through shape features are shown in Fig. 4. The result of shape feature extraction in Fig. 5 are seven invariant moment values as in Table II. The seven values represent the shape features of each image in the form of numbers. Testing is done to prove that invariant moment is not susceptible to RST (Rotate-Scale-Translation).
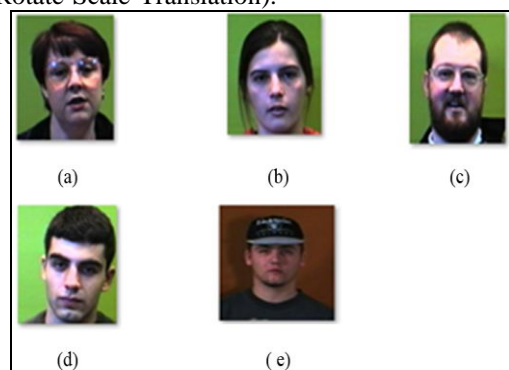

Fig. 4. Diversity of face images with various shapes and types (a) Boxed female image, (b) Female oval image, (c) Male oval image (d) Male square image, (e) Triangle Male image
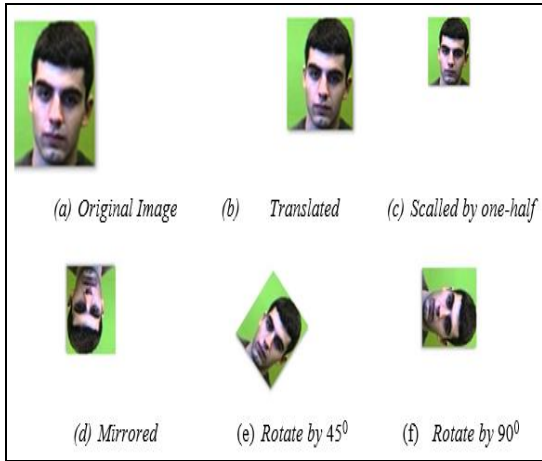
Fig. 5. RST face images, (a) Original Image, (b) Translated, (c) Scalled by one-half, (d) Mirrored, (e) Rotated by 450, (f) Rotated by 900

TABLE II RESULT OF FEATURE EXTRACTION USING INVARIANT MOMENT

Moment	Original	Rotate 90^0	Scalled by one-half	Translated
M1	3.13E-04	2.24E-04	3.15E-04	2.42E-04
M2	4.74E-06	6.73E-08	4.74E-07	2.94E-07
M3	8.43E-11	4.47E-12	8.43E-10	2.78E-14
M4	7.23E-11	5.69E-12	7.11E-11	7.37E-14
M5	4.83E-22	2.34E-24	4.69E-22	-2.05E-28
M6	2.43E-13	2.54E-11	2.55E-15	-2.10E-18
M7	3.65E-22	1.79E-24	3.85E-24	-4.86E-32

From the seven feature values above, it can be seen that the moment value 4, 5, 6 and 7 is the moment value which is relatively consistent with the difference in the value of the moment which is relatively small or less than 2 with the existence of rotate, scale and translation. Examples of the results of feature feature extraction in the asset image are shown in Table III.

TABLE III VALUE FEATURES OF FACE IMAGE SHAPES.

M1	M2	M3	M4	M5	M6	M7
1.45000195	0.188975334	0.0762499	0.02363947	0.000947725	0.00376677	0.0003303
2.48348357	0.823986086	3.0933703	2.15046114	5.390506501	1.61300548	1.3058714
3.48416923	1.039665923	1.8335181	1.02325626	1.000085438	0.92708062	-0.9819713
2.21618706	0.537978445	1.6756548	0.79498076	0.913934495	0.53380865	-0.0813140
2.47919286	0.380511046	0.7509339	2.30678952	2.991816666	1.37385046	-0.5165248
2.60737587	1.23599042	3.8528044	2.80684864	9.000996332	2.94092083	2.0447390
1.86594301	0.727555306	0.2769835	0.19100703	0.042842856	0.15677254	-0.0097308
3.52586089	3.570411335	21.328541	14.8063178	249.6122036	26.8028648	-83.216753
1.45000195	0.188975334	0.0762499	0.02363947	0.000947725	0.00376677	0.0003303
.....	.....	.....	.....	.....	.....	.....
.....	.....	.....	.....	.....	.....	.....
2.60737587	1.23599042	3.8528044	2.80684864	9.000996332	2.94092083	2.0447390
1.45000195	0.188975334	0.0762499	0.02363947	0.000947725	0.00376677	0.0003303
2.48348357	0.823986086	3.0933703	2.15046114	5.390506501	1.61300548	1.3058714

Table III shows that the value of the shape feature components has a long range of values between one moment value and another moment value in one image. This shows the dominance between the components of the moment value. This dominance can result in no other moment playing a role in determining the accuracy of similarity or it can be said that a component of little value has a small role in determining the accuracy of an image's similarity. To eliminate the

dominance between the value of the feature, normalization of the feature value is done in the next step.

b) Color Feature Extraction: Preprocessing with histogram equalization in Figure 3 produces a more contrasting image. The image is an input image for color feature extraction. The extraction method used is the color moment. This method is able to distinguish images based on color features [11].

c) Texture Feature Extraction: Texture is the regularity of certain patterns formed from the arrangement of pixels in an image. Texture value can be used as one of the variables for image similarity measurement.

d) Normalization Feature: Normalization is important to group the ranges or intervals of different feature values in the same scale with smaller ranges. Normalization is carried out to give the same weight to different feature values from the extraction results. Normalization in this study uses equation (1).

$$D'(i) = \frac{D(i) - \min(D)}{\max(D) - \min(D)} \tag{1}$$

where D '(i) is the normalized i value feature, D (i) is the i value feature before normalized, min (D) is the minimum value of each feature and max (D) is the maximum value each feature.

e) Feature weighting: The weighting of the different feature values in this study was used for retrieval accuracy analysis. Weighting varying feature values will result in different levels of accuracy of the retrieval.

E. K-Mean Clustering

The use of clustering in this study aims to speed up the retrieval process. K-Means is a clustering technique using centroid (cluster center point) to represent a cluster [5]. In addition, K-Means is able to group large data quickly [4].

F. Measurement of Similarity

Measurement of query image similarity with database image uses the euclidean distance concept and is done only in one cluster.Measurement of similarity using equation (2) using eleven dimensions or features are three shape features including the invariant moment values of 3, 5 and 7, then 3 color features using Hue values and 5 texture features namely smoothness, standard deviation, skewness, uniformity and entropy.

$$D(Q,M) = \sqrt{\sum_{n=1}^{k} (Q_n - M_n)^2} \tag{2}$$

Where Q_n and M_n are features of the query image and database image in the nth dimension. The results of the similarity calculation are sorted so that the similarity value near zero has the highest level of similarity.

G. Retrieval Akuration

Retrieval accuracy measurements on images before and after weighting and clustering are done using equation (3).

$$AK = \frac{AR}{AS} \times 100\% \tag{3}$$

AR (Actual Relevant): the amount of data deemed relevant by the user. AS (Actual Search): the amount of data retrieved by the system.

III. METHOD

In this research, a lecture attendance verification model was developed using a stereo vision camera. The use of the camera is to anticipate differences in seating position, face poses, lighting variations and the distance of the camera to the object at each lecture meeting. The use of face image database for college participants is matched with the recording of lectures in the lecture hall attended by students as lecturers. Students in attending lectures do not always sit in the same position or seat, they tend to change seats in each meeting. So are the pose, lighting and distance of the object with the camera at each lecture meeting also changes. It is necessary to measure the level of accuracy of attendance verification so that the verification model has an optimal level of accuracy.

The tools used in this study are equipment in the form of hardware and software. The hardware used is a stereo vision camera, a computer with 8GB RAM specifications, 1 TB of memory and a printer equipped with a scanner. The software used for this research is the operating system Windows XP, C #, OpenCV, Mathlab 2012 and Visual Studio 2010.

In a number of studies that have been carried out above, no one has used a combination of feature weighting and clustering methods based on shape, color and texture features for the CBIR-based image search process. In this research, image retrieval has been carried out with the research stages shown in Fig. 6.
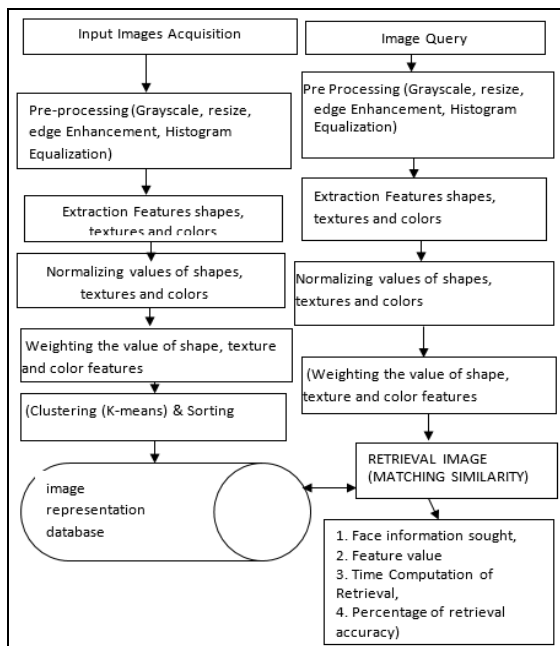

Fig. 6. Overview of Image Search Stages

IV. RESULT AND DISCUSSION

A. Retrival Results

Retrieval testing is done with variations in the number of clusters and variations in the percentage of feature weights.The results of the test with 50% weighting for shape feature weight, 30% color feature weight and 20% texture feature weight with 10 cluster variations are shown in Fig. 7.
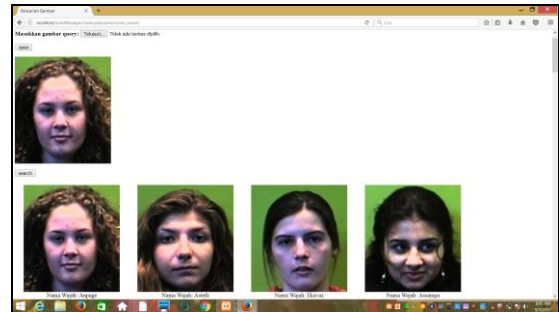

Fig. 7. Retrieval Using the Weight of Shape, Color and Texture (0.5,0.3 and 0.2)

Retrieval testing is done with variations in the number of clusters and variations in the percentage of feature weights.The results of the test with 50% weighting for shape feature weight, 30% color feature weight and 20% texture feature weight with 10 cluster variations are shown in Fig. 7. Fig. 7 shows the results of retrieval ranking of similarities 1 to 5 from the face image database. Tests with weighting variations and variations in the number of clusters will produce different retrieval output on the same data.

B. Before Clustering

Retrieval accuracy analysis at this stage uses a face image database before clustering with weighting variations in shape, color and texture features. The accuracy value of the retrieval results with the same feature weights and varying feature weights can be seen in Table IV.

TABLE IV. PERCENTAGE OF RETRIEVAL ACCURACY WITH SAME FEATURE WEIGHT AND VARIABLE FEATURE WEIGHT IN FACE IMAGE DATABASE

Image of Name	Equal Weighted	W1	W2	W3	W4
I1	75,55%	78,59%	86,75%	84,38%	61,33%
I2	74,67%	79%	87%	83,87%	60,67%
I3	75,05%	79,45%	87,10%	83,33%	60%
I4	74,69%	79,15%	87,50%	86,67%	59,33%
I5	75,65%	78,75%	86,67%	84,38%	60,67%
I6	75%	78,50%	86,36%	84%	60%
I7	75,76%	79,45%	86,67%	83,87%	59,33%
I8	74,89%	78,65%	87,33%	83,33%	60,67%
I9	74,87%	79,55%	86%	84%	61,33%
I10	75,35%	79,25%	87,33%	84,38%	60,67%

Table IV shows that the retrieval accuracy with the variation of W2 produces the highest average accuracy and the variation of the weight of the W4 produces the lowest accuracy. Retrieval accuracy graphs with the same feature

weights and feature weights vary in the image database before clustering are shown in Fig. 8.
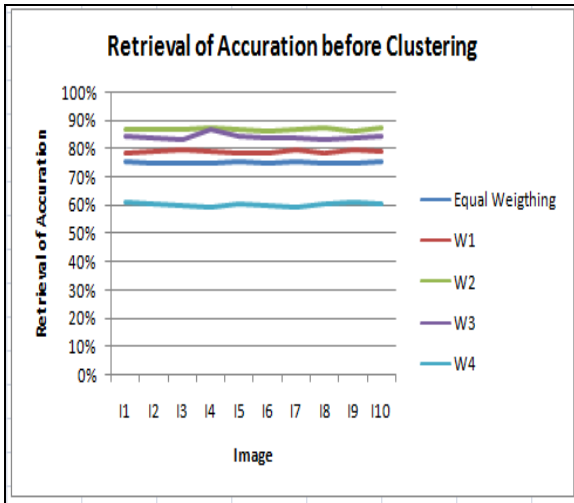

Fig. 8. Retrieval Accuracy Graph with Same and Variable Feature Weighting Before Clustering in the Face Image Database

From Fig. 8, it is shown that the variation in W2 weights yields the highest retrieval accuracy compared to other weight variations.

C. After Clustering

Grouping based on similarity will further narrow the search space for image data information so that it will shorten retrieval time and improve retrieval accuracy. Retrieval accuracy with variations in weights and number of clusters can be seen in Table V.

TABLE V. PERCENTAGE OF RETRIEVAL ACCURACY WITH VARIATION IN FEATURE WEIGHT AND VARIATION IN NUMBER OF CLUSTERS

Number of Clusters	Same Weight	W1	W2	W3	W4
3	72,55%	75,59%	92,67%	82,38%	63,33%
4	72,67%	75%	92,33%	82,33%	63,67%
5	73,05%	76,45%	93,33%	83,36%	64,00%
6	73,69%	76,15%	93,67%	83,67%	64,33%
7	74,65%	77,75%	94,33%	84,38%	65,67%
8	74%	77,50%	94%	84%	65,00%
9	74,76%	77,45%	94,67%	83,87%	64,33%
10	75,89%	79,65%	95,67%	86,67%	66,93%
11	74,47%	79,15%	93,55%	84%	66,23%
12	74,15%	77,25%	93%	85,33%	65,67%
13	74,20%	76,15%	92,23%	82,63%	64,63%
14	73,25%	75,59%	92,13%	82,44%	64,55%
15	73,13%	75,29%	92,09%	82,23%	63,43%

From Table V shows that the weighting variation of W2 with a weight of 50% of the shape features, 30% of the color features and 20% of the texture features and variations in the number of clusters 10 results in the highest retrieval accuracy of more than 95%. While the weighting variation of W4 with 30% of the weight of the shape features, 50% of the weight of the color features and 20% of the weight of the texture features with variations in the number of clusters 3 results in the lowest accuracy which is less than 64%. Retrieval accuracy graphs in the face image database using the same and varied feature weights namely W1, W2, W3 and W4 with variations in the number of clusters ranging from 3 to 15 can be seen in Fig. 9.
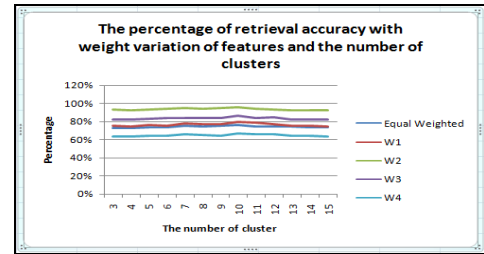

Fig. 9. Retrieval Accuracy Graph with Feature Weight Variation and Cluster Number Variation

Fig. 9. shows that retrieval accuracy tends to increase to variations in the number of clusters of 10 and tends to decrease after the number of clusters is greater than 10.

D. Attendance Database Design

The use of face detection systems for attendance attendance lectures are built using attendance databases that have 10 tables in them, there are: login tables, face training, photos, students, attendance, minutes, classes, courses, lecturers, and schedules. When the lecturer starts to open the attendance system, the system asks the user to fill in a category from the user, the username that is filled in with nip or the username that has been created and the password. In the future, when taking photos in lectures for attendance systems do face testing by matching face training data for which the photo data has been stored in the face training table. Photos of the faces of each student will be detected by a camera that leads to their faces and displays the attendance link that is connected to the attendance table. The attendance table is linked to the official news table allowing the user to check the attendance of students who attend the lecture meeting at that time by checking their details. The news table seen in Fig. 10 linked to the schedule allows the user to know the classes and courses that are taking place, the capable lecturer.
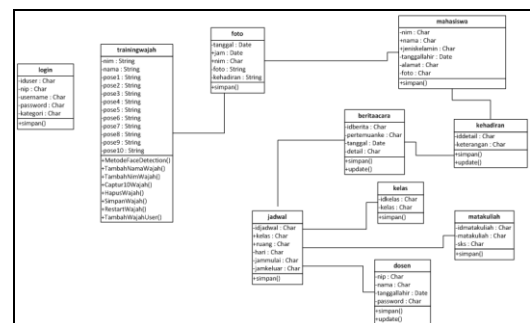

Fig. 10. Database Design

V. CONCLUSSIONS

In testing the retrieval accuracy of feature values before and after the normalization of feature values shows that there is a change in the percentage of retrieval accuracy in normalized image data. The difference in the number of clusters in the image database has a significant effect on the time to search for CBIR-based information in this case using office face data as training and testing data.

REFERENCES

[1] P. N. Belhumeur,., Joau, R., Hespanha and david, J.K, Egienfaces vs Fisherfaces, 1997.

[2] G.A.N. Martinez and R.V. Ayala, Real Time face Detection Using Neural Network In Artificial Intelligent, 10th Mexican International Conference (pp. 144-149). IEE, 2011.

[3] J. Harguess and J. K. Aggarwal, he average half face 2D and 3D face recognition, Pattern Recognition and Machine, 135-148, 2009.

[4] Jumi, A. Ashari, Analisis Ektraksi Ciri Menggunakan Zernike Moment dan Invariant Moment Pada Citra Aset, Prosiding Seminar Nasional FMIPA UNNES, Semarang,, 2012.

[5] Jumi, S. N. Azhari, Akurasi Kombinasi Ekstraksi Ciri Warna dan Bentuk Untuk Penelusuran Citra Aset, Proceeding International Conference on Information Technology and Electrical Engineering di Fakultas teknik Elektro Universitas Gadjahmada, Yogyakarta, 2012.

[6] Jumi, A. Harjoko, Pengukuran Kemiripan Citra dengan Metode Invariant Moment dan Euclidean Distance, Prosiding Seminar Nasional di Universitas Surakarta pada tanggal 10 Maret 2012.

[7] Jumi and A. Harjoko, Image Similarity Analysis Based on Shape, Color and Texture Feature of Asset Image, International Conference on Computer Science Electronics and Instrumentation, Yogyakarta, Indonesia, 2012.

[8] Y. Siagian, K. Anwar, J. Hutahean, Pengenalan Wajah pada Sistem Absensi Secara Realtime dengan Metode Wavelet, Jurnal Royal, Vol 1 No.1, 2014.

[9] M. Turk and A. Petland, Egigenfaces for Recognition, Journal cognitive neuroscience vol 3.no.1, pp. 71-86, 1991.

[10] P. Viola and M. Jones, Rapid object detection using a boosted cascade of simple features, In proceeding IEEE conference on computer vision and pattern recognition, 2001 (1-9) , 2001.

[11] S. A. Wibowo, H. Lee, E. K. Kim, T. Kwon And S. Kim, Tracking detection and correcction face tracking by detection approach based on fuzzy coding histogram and point representation, International conference on fuzzy theory and It's apllications, IEE, 34-39, 2015.