Face Recognition and Retrieval across Age Using GLCM Technique

[1]Lakshmi Bhavana P,[2]Myna A N
[1]PG Student,[2]Assistant Professor
Department of Information Science and Engineering,
M.S.Ramaiah Institute of Technology, Bengaluru, India

***Abstract*—This paper proposes an algorithm for face recognition and retrieval across the variation of age factor in human faces by making use of the Gray Level Co-occurrence Matrix (GLCM) technique. A cross age facial image dataset will be used to implement this system. A large scale celebrity image dataset called cross age celebrity dataset (CACD) is available on the Internet merely for educational and research purposes with an age invariant reference space. Feature extraction is a method by which the visual content of the images are captured for recognition and retrieval in huge databases. This paper presents how the gray level co-occurrence matrix can be used to extract the second order statistical features of an image. The texture feature of every image is extracted to perform the implementation. This method requires only a linear projection for feature encoding and is highly scalable, which is observed in the retrieval phase.**

***Index Terms*— Cross Age Face Recognition, GLCM, Texture Feature.**

I. INTRODUCTION

Face recognition has been an important aspect in the field of image processing since a very long time. It is a very challenging task in analysing images and in pattern recognition. In a world where tremendous advancements have been made in computer vision and pattern recognition, face recognition across age via facial images has become a subject of interest.

Detecting the presence of a human face in the sample image is called face detection whereas identifying a particular individual in the same image is referred to as face recognition. One has to master the differences between face detection and face recognition to bring about new methods, ideas technologies and algorithms that help in accuracy of face recognition. Factors that affect the accuracy of a facial image fall into categories like illumination, the pose of the person in the image, facial expression and factor of aging [1]. Previous researches have made efforts in solving the face recognition problem considering one or more or a combination of these factors. Based on the face verification standards in the unconstrained environments a good performance was achieved by using the dataset called Labelled Faces in the Wild (LFW) [2]. Nevertheless the dataset contains little changes in the age factor.

In the face recognition exemplar the pre-specified data sets comprise images of a group of people whose identity is known. The challenge lies in deciding to which person in the whole dataset the query image refers to especially with factors like age variance. Most of the studies are related to age estimation and simulation while facial images across age can vary and thus face recognition or retrieval becomes a more challenging task.

II. RELATED WORK

Recognizing a person using digital face images is the task of face recognition. Many face recognition systems have been designed and developed to measure the similarity between two or more images. Automated face recognition systems find the facial landmarks to develop an accurate and scalable matching scheme. A major problem in the field of face recognition is facial aging, scars and marks of the query image and matching these low quality images to a photograph database. Anil K Jain [3] proposed a solution to this problem by developing preprocessing methods with the ability to improve pose, illumination and image resolution.

Face Datasets

It is a known fact that the amount of data available is directly proportional to the amount of research that can be done. In face recognition a number of datasets are available to the present day researchers as a benchmark, facilitating a good platform for further inventions and discoveries.

The Labeled Faces in the Wild (LFW) by Hu han [4] is one such database that was designed to study the problem of face recognition under unconstrained environments where the facial images show diversity in pose, expression, background, age, color, hairstyles. The Face Recognition Grand Challenge (FRGC) and FG-NET-Aging Dataset [5] database include high resolution facial image data along with 3D scans and sequentially arranged images.

This database particularly contains two images of a single person- one with a smiling face and the other with a neutral expression. Another database proposed by the researchers called the BioID Face Database contains different views of the same scene. The human face undergoes a lot of changes that are intrinsic [6] like pimples, saggy skin, fine lines, warts etc and extrinsic changes that may occur due to a person's smoking habits, pollution or due to the environment that he lives in. Such reasons make it difficult for the researcher to introduce novelty in the field of face recognition.

Face Recognition Models

Measuring the similarity between two images under different settings is a tedious task. Due to the intrapersonal changes face recognition process becomes unbelievably hectic and stringent.

An associate predict model was proposed by Qi Yin, Xiaou Tang [7] which associated an input face image with identities that were alike from a dataset that was generic. Once the association was completed, prediction of whether two images look alike or not was done by the model itself creatively. Keeping in mind the wide range of features that could be drawn out of a facial image many researches were conducted. Features like color and shape are concentrated more upon but one feature that catches the eye of the researchers is the texture feature. Timo Ahonen, Abdenour Hadid[8] proposed the Local Binary Pattern model for texture description which is tolerant to monotonic gray-scale changes.

Age estimation methods can be grouped into three categories such as anthropometric, age pattern subspace and age regression methods. Albert Montillo, Haibin Ling [9] exploits the machine learning method called the random forest where a challenging task is to define the predictor features. Forensic experts often use facial marks to verify a suspect against a candidate face image. Park and Jains face-image matching and retrieval system provides tools, such as manual and automatic mark labeling; image retrieval using facial marks like scars, moles and freckles[10]. Dihong Gong proposed a hidden factor analysis [11] approach to address the challenging problem of age invariant face recognition. The basic idea of the Hidden Factor Analysis model is to separate the aging variations from the person's specific features for pursuing the robust age-invariant face features.

III. GRAY LEVEL CO-OCCURRENCE MATRIX (GLCM)

A property that represents the surface and structure of an image is called Texture [12]. It can also be defined as regular repetition of a pattern on a surface. Textures of an image are detailed visual patterns that comprise of regions with sub-patterns having the characteristics of brightness, color, shape, size. Statistical texture analysis deals with texture features computed on the basis of statistical distribution of pixel intensity at a given position relative to others in a matrix of pixel representing image. Keeping in mind the number of pixels or dots in each combination, statistics is classified into first-order statistics, second-order statistics or higher-order statistics. Higher order statistics consider the relationship between three or more pixels.

Feature extraction based on gray-level co-occurrence matrix (GLCM) falls under the second-order statistics that can be use to analyze images as a texture. GLCM is a table that shows of how often a combination of pixel brightness values in an image occurs. A simple 1D histogram might not be entirely useful in characterizing texture features as it is a spatial property. Hence, GLCM matrix is used in texture analysis.

The GLCM is a square matrix that reveals specific properties about the spatial distribution of the pixels in a texture image as defined by Haralick in 1973[13]. It shows how often a reference pixel which has the intensity 'i' occurs in a unique relationship to another pixel called neighbor pixel with the intensity 'j'. Each element of the matrix denotes the number of occurrences of pixel element pair (i,j) with value 'i' and a pixel with value 'j' separated by a distance 'd' relative to each other. The spatial relationship between two neighboring pixels can be represented in several ways with different offsets and angles, the default one being between a pixel and its immediate neighbor to its right. Due to their large dimensionality, the GLCM's are very sensitive to the texture samples on which they are estimated. Hence the number of gray levels is often reduced.

Figure 1(a) represents the formation of the GLCM of the grey-level (4 levels) image at the distance d = 1 and the direction of 0°.

0	0	1	1	1
0	0	1	1	1
0	2	2	2	2
2	2	2	3	3
2	2	3	3	3

	0	1	2	3
0	2	2	1	0
1	0	4	0	0
2	0	0	6	2
3	0	0	0	3

Figure 1(a) and (b) Formation of the GLCM matrix

It is an example matrix of pixels intensity representing image with four levels of grey. Note the intensity level intensity level 0 and 1 are marked with a thin box. The thin box represents pixel-intensity 0 with pixel intensity 1 as its neighbor (in the horizontal direction or the direction of 0 º). There are two occurrences of such pixels. Therefore, the GLCM matrix formed in Fig 1(b) shows a value 2 in row-0, column-1. Likewise, GLCM matrix row-0 column-0 is also given a value of 2, because there are two occurrences in which pixels with value 0 has pixels 0 as its neighbor (horizontal direction). In this manner the pixels matrix representing in the first figure can be transformed into GLCM as in second figure. In addition to the horizontal direction (0 º), GLCM can also be formed for the direction of 45 º, 90 º and 135 º as shown in Fig. 2.

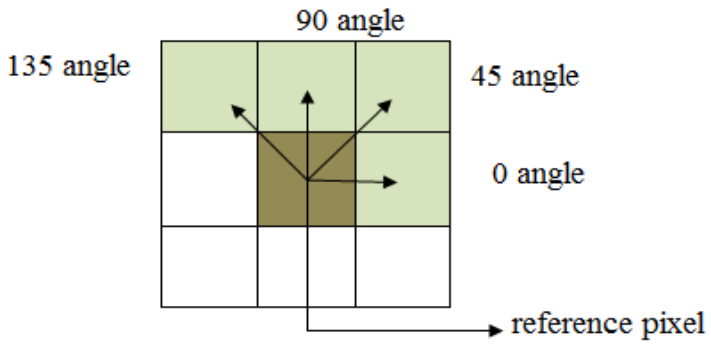

Figure 2 Angles at which the reference pixel is chosen

As told by Haralick [13]14 features from the co-occurrence matrix can be extracted for an image that is: Entropy, Sum Entropy, Difference Entropy, Inverse Difference Moment, Contrast, Homogeneity, Inertia, Correlation, Maximum Correlation Coefficient, Angular Second Moment, Sum Average, Variance, Sum Variance, Difference Variance. These features can be used in a combination as a vector. Although co-occurrence matrices capture the texture properties, it never uses a tool for analysis, such as comparing the two textures. The matrix of data must be extracted again to get the numbers that can be used to classify the texture.

GLCM Steps

STEP 1	:	Get width and height of image.
STEP 2	:	Get the packed pixel content of whole image as an RGB array
STEP 3	:	Transform the RGB array into a gray scale array
STEP 4	:	Initialize the Co-occurrence Matrix with gray level 256
STEP 5	:	For each pixel compare the gray level of another pixel
STEP 6	:	Store the value in Co-Occurrence Vector and return the vector

IV. USING GLCM FOR FEATURE EXTRACTION

As suggested by Haralick [13] a number of texture features can be extracted from the GLCM. In general the following notations are used:

G is defined as the number of gray levels used.
μ is defined as the mean value of matrix element P.
μ_x, μ_y, σ_x and σ_y are defined as the means and standard deviations of P_x and P_y.

The 14 textural features proposed by Haralick provide the required information about spatial distributions and texture characteristics of an image such as: homogeneity of the image, linear dependencies of gray tone, pixel contrast levels, number of boundaries present and the complexity of the image. Haralick et al introduced the use of co-occurrence probability matrix for extracting various texture features from an image. GLCM of an image is computed using a displacement vector d, defined by the radius delta (δ) and pixel orientation theta (θ).

Features

1. *Contrast*

Contrast is the measure of the spatial frequency of an image and is also known as difference moment of GLCM. The difference between the highest and the lowest values of a contiguous set of pixels is given by this metric. Local variations present in the image and the amount of it can be determined by measuring the contrast of the image.

$$CONTRAST = \sum_{n=0}^{G-1} n\text{^}2\{\sum_{i=1}^{G}\sum_{j=0}^{G} P(i,j)\}, \quad |i-j| = n \tag{1}$$

2. *Inverse Difference Moment (IDM)*

Local homogeneity of the image influences the IDM. Due to the weighting factor in the denominator IDM gets small contributions from non-homogeneous areas. The result for non-homogeneous images is a low IDM value whereas homogeneous images get a relatively high value.

$$IDM = \sum_{i=0}^{G-1}\sum_{i=0}^{G-1} \frac{1}{1+(i-j)^2} P(i,j) \tag{2}$$

3. *Entropy*

The complexity of an image is measured by the statistic called entropy. If the image is not uniform in its texture, entropy of the image gains a high value and the GLCM values are low. Entropy is strongly proportional to the complexity of an image but inversely correlated to energy. Homogeneous scenes usually have high entropy values.

$$ENTROPY = -\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} P(i,j) * \log\left(P(i,j)\right) \tag{3}$$

4. *Correlation*

Linear dependencies of gray tone in any image are calculated by a feature called correlation. It measures the linear dependence between two or more pixels at a given position that are relative to one another.

$$CORRELATION = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1} \frac{\{i*j\}*P(i,j) - \{\mu_x * \mu_y\}}{\sigma_x * \sigma_y} \tag{4}$$

5. *Variance*

Variance of an image is also sometimes known as sum of squares. It measures the heterogeneity and is closely related to the standard deviation which is one of the first order statistics. When a fluctuation of gray values occurs from their mean value the variance of an image also increases. All the elements that differ from the average value of P(i,j) are weighted by the variance value.

$$VARIANCE = \frac{1}{G^2}\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} (P(i,j) - \mu)^2 \tag{5}$$

6. *Angular Second Moment (ASM)*

ASM is a measure of homogeneity of an image and is also known as Uniformity. In a homogeneous scene only a few gray levels are seen thus giving a GLCM with only a few values of P(i, j). Thus, the sum of squares taken will be high.

$$ASM = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1} \{P(i,j)\}^2 \tag{6}$$

The textural uniformity of an image i.e pixel pair repetition is calculated by this metric. The maximum energy value that can be reached is 1. Constant or a periodic form of the gray levels results in high values of ASM. The other features are the ones that are derived from the above mentioned 6 features.

The input image is recognised by using the open source tool called the OpenCV. The algorithm works in three steps namely face recognition, feature extraction and classification. The recognition process is aided by an open source tool. The feature extraction is implemented by a high level language like Java. Once the features of the image are extracted, it is compared with the images available in the database for the classification purpose. Classification is further implemented by comparing the input image against the images in the database by measuring the Euclidean distance between the images.

Euclidean Distance

The Euclidean distance is the distance between two points in Euclidean space. The distance between two points P and Q is defined as the square root of the sum of the squares of the differences between the corresponding coordinates of the points. The two-dimensional Euclidean geometry, the Euclidean distance between two points is given by:

$$EUCLIDEAN\ DISTANCE = \sqrt{\sum_{i=1}^{n} (Pi - Qi)^2} \quad (7)$$

A certain threshold is set, below which the result images are displayed in the ascending order of their distances to the input image.

V. PROPOSED APPROACH

Figure 3 shows the proposed approach used for implementing GLCM in feature extraction process. The query image Fi is given as an input and the OpenCV detects or recognises the presence of a human face. Further feature extraction where the features of the images present in the database F (i,j) are retrieved in the form of a co-occurrence feature vector. The number of categories of images is given by variable N and the number of images in each category is given by variable M. Setting of threshold value is made completely automated for the computer. Methods to assign threshold are categorized into the six groups: Histogram shape-based methods, Clustering-based methods, Entropy-based methods, Object Attribute-based methods, spatial methods, Local methods.

We make use of the histogram shape based method where peaks, valleys and curvatures of the smoothed histogram are analyzed. There are two types of histograms namely local and global. To set a global threshold or to adapt a local threshold to an area, we usually look at the histogram to see if we can find two or more distinct modes—one for the foreground and one for the background. A histogram is a probability distribution: p(g) = ng/n where ng are the non gray levels and n is the total number of pixels. In a Local Color Histogram (LCH) information regarding the color distribution of regions is described [14]. When comparing two images, we calculate the distance, using their histograms, between a region in one image and a region in same location in the other image. The distance between the two images will be determined by the Euclidean distance. The category of images that are less than or equal to the threshold value are retrieved and displayed as output to the user.

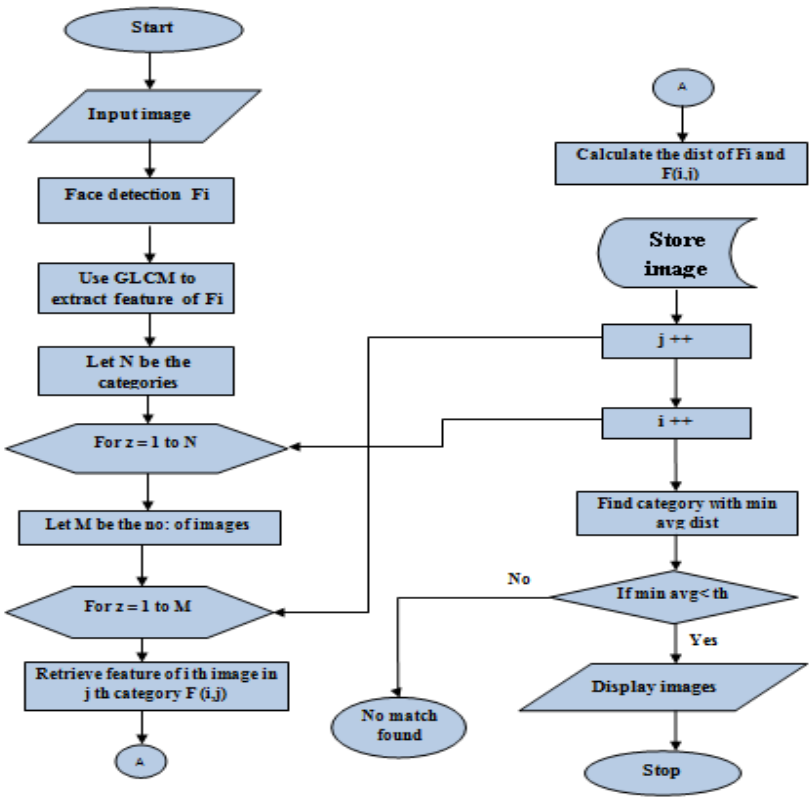

Figure 3 Proposed approach to retrieve the cross-age images from the database

VI. Results

The experimental results of our approach to face recognition and retrieval across age are shown in Fig.4. An average of about 200 images was used to conduct the experiment. We make use of the celebrity database that is available freely on the internet called cross-age celebrity dataset (CACD) as it contains images of various celebrities at all ages.

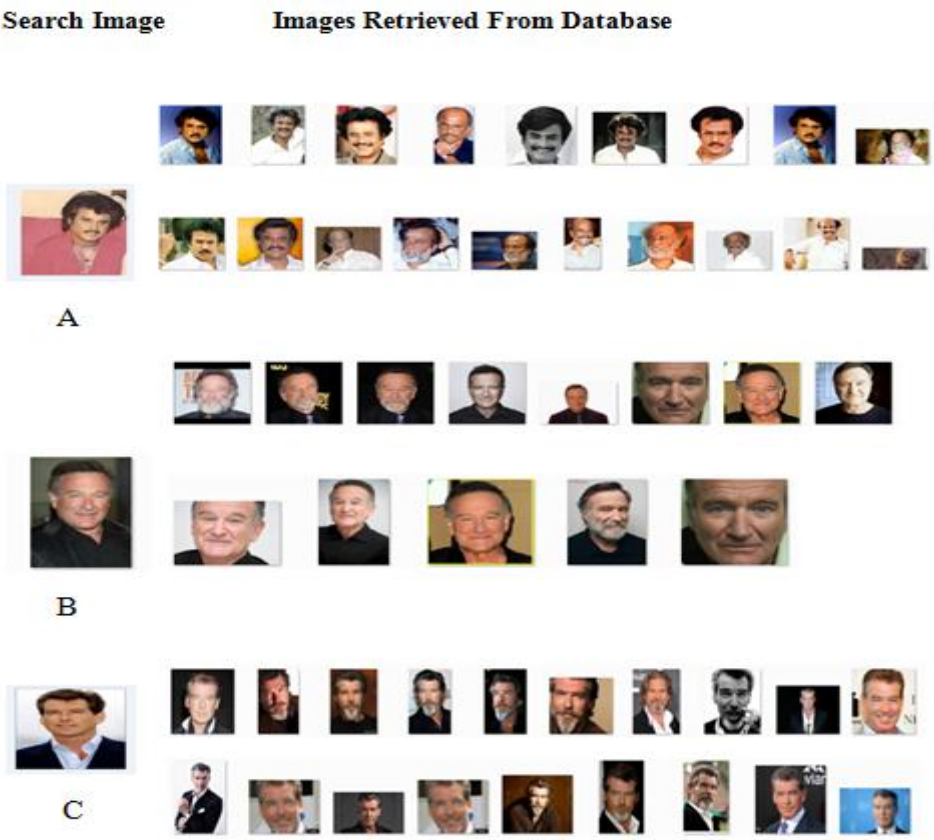

Figure 4 Sample Images of three celebrities retrieved from the database

The images were resized to a dimension of 250 x 250, i.e to a width of 250 pixels and height of 250 pixels before experimentation. The image on the left hand side represents the query image while the ones on the right represent the retrieval results. The results of features extracted for a sample input images are shown in the Table 1 and the Euclidean distance measured between the query image and images in dataset is shown in Table 2. The features extracted for the images in the dataset are shown in Table 3.

Table 1 Texture features obtained for input celebrity

Image	Contrast	Angular second moment	First order IDM	Second order IDM	Entropy
A	0.0018	23.1293	0.1277	8.1754	8.4229
B	0.0019	12.5272	0.2405	8.2277	8.7504
C	0.0325	17.6856	0.2049	7.5199	7.8677

Table 2 Calculation of Euclidean distance values for images

Image	Euclidean Distance
A	23.91674
B	07.21204
C	10.53051

Table 3 Features extracted for images in the dataset

Image A	Contrast	Angular second moment	First order IDM	Second order IDM	Entropy
1	0.0033	15.3414	0.2043	8.3724	8.8763
2	0.0160	12.9270	0.2415	8.1679	8.7620
3	0.0012	13.1050	0.2054	8.4470	8.8632
4	0.0013	12.8547	0.2063	8.5154	8.8910
5	0.0013	12.8547	0.2063	8.5154	8.8912
6	0.0029	12.7077	0.2371	8.1448	8.6086
7	9.0000	13.7103	0.1834	8.5831	8.8518
8	0.0015	18.0517	0.1788	8.5388	8.7352
9	0.0015	17.5905	0.1710	8.4939	8.6728
10	0.0012	20.2043	0.1442	8.5959	8.6104
11	0.0014	11.6499	0.1999	8.3873	8.8151
12	0.0033	15.3414	0.2047	8.3724	8.8763
13	0.0109	19.1180	0.2202	8.3553	8.8490
14	0.0108	19.1180	0.2202	8.3553	8.8491
15	8.0101	26.1757	0.1346	8.8188	8.8398
16	0.0017	15.2426	0.1905	8.4577	8.6097
17	0.0028	12.7477	0.2011	8.1698	8.4425
18	0.0034	11.7770	0.2279	8.1605	8.5007
19	0.0022	13.9880	0.1950	8.3555	8.5846

The time taken for the whole process was found to be 9018 ms for the sample query image. We look forward for further improvements and experimentation to this technique which can be of great help in the field of forensics, social networking and research programmes.

VII. Conclusion

Motivated by Haralick work, we have presented how GLCM can be used for face recognition and retrieval across the aging factor in human faces. GLCM can be a useful technique in extracting features of an image. This method has two important steps apart from the face recognition performed by the OpenCV tools i.e. feature extraction and classification of the image to which data it belongs to in the database. We have seen that the goal to extract similar images as the query image across age factor has been successfully achieved by the proposed algorithm. To our best beliefs the method has been tested for 147 images with a very few false positives. As more effective and efficient prior knowledge can be built and more improvements can be made, we hope that eventually this work can be extended in fields like forensics, social networking and many more.

Acknowledgment

We would like to express our deepest appreciation to M.S. Ramaiah Institute of Technology for providing all the facilities to carry out this work.

References

[1] U.Park “Face matching and retrieval in forensics applications,” IEEE Multimedia Mag.,vol. 19, no. 1, p. 20, Jan.2012.

[2] G. B. Huang, M. Ramesh, T. Berg, and E.Learned-Miller, “Labeled faces in the wild: A database for studying face recognition in unconstrained environments,” Univ. Massachusetts, Amherst, MA, USA, Tech. Rep. 07-49, Oct. 2007, pp. 1–11.

[3] S. Z. Li and A. K. Jain, Handbook of Face Recognition, 2nd ed. New York, NY, USA: Springer, 2011.

[4] Age, Gender and Race Estimation from Unconstrained Face Images Hu Han, Member, IEEE and Anil K. Jain, Fellow, IEEE.

[5] The fg-net aging database. In http://www.fgnet.rsunit.com/.

[6] C. Hill, C. Solomon, and S. Gibson, “Aging the Human Face—A Statistically Rigorous Approach,” Proc. IEE Int’l Symp.Imaging for Crime Detection and Prevention, pp. 89-94, 2005.

[7] Q. Yin, X. Tang, and J. Sun, “An associate-predict model for face recognition,” in Proc. IEEE Conf. Comput. Vis. Pattern Recog., Jun.2011, pp. 497–504.

[8] Face Description with Local Binary Patterns:Application to Face Recognition Timo Ahonen, Student Member, IEEE, Abdenour Hadid, and Matti Pietika¨ inen, Senior Member, IEEE.

[9] A. Montillo and H. Ling, "Age regression from faces using random forests," in Proc.IEEE Int. Conf. Image Process., Nov. 2009, pp.2465–2468.

[10] U. Park and A.K. Jain, ''Face Matching and Retrieval Using Soft Biometrics,'' IEEE Trans. Information Forensics and Security, vol. 5, no. 3, 2010, pp.406-415.

[11] D. Gong, Z. Li, D. Lin, J. Liu, and X. Tang,"Hidden factor analysis for age invariant face recognition," in Proc.IEEE Int. Conf. Comput. Vis., Dec. 2013, pp. 2872– 2879.

[12] C. Rosenberger and C. Cariou, "Contribution to Texture Analysis", In Proc. International Conference on Quality Control by Artificial Vision,vol. 1, pp. 122-126, Le Creuzot, 2001.

[13] R.E. Haralick, K.Shanmugam, I.Dinstein, Textural Features for Image Classification, IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-3, No.6, Nov 1973.

[14] Content Based Image Retrieval Using Local Color Histogram Metty Mustikasari, Sarifuddin Madenda, EriPrasetyo, Djati Kerami, Suryadi Harmanto Gunadarma University STMIK Jakarta, University of Indonesia metty@staff.gunadarma.ac.id, sarif@jak-stik.ac.id, eri@staff.gunadarma.ac.id