Proceedings of the International MultiConference of Engineers and Computer Scientists 2018 Vol I
IMECS 2018, March 14-16, 2018, Hong Kong

Face Classification using Adjusted Histogram in Grayscale

Weenakorn Ieosanurak, and Watcharin Klongdee

***Abstract*—This paper focuses histogram in grayscale image for a face recognition. We propose an easy and effective classification method for a face grayscale image based on histogram in grayscale of face images. We find histogram of image pixels as an input, then we modify histogram and classify by Euclidean distance. The proposed method tests on the Grimace, ORL and Jaffe databases. The recognition rate is compared with Kullback-Leibler divergence (KLD). The proposed algorithms show high performance and they have recognition rate over 98%, 95% and 100% for ORL, Jaffe and Grimace database, respectively.**

***Index Terms*—face classification, face recognition, histogram, Euclidean distance**

I. Introduction

THERE are several methods to extract feature extraction for face human such as Principle Component Analysis (PCA) [1], Linear Discriminant Analysis (LDA) [2], Independent Component Analysis (ICA) [3]. Although there are several methods for face recognition, there is some publications on application of histogram or the probability distribution functions based methods in face recognition. For example: Yoo and Oh [4] presented a method which used chromatic histograms of faces. Rodriguez and Marcel [5] proposed a method that divided a face into several blocks and extracted the local binary pattern (LBP) feature histograms from each block and concatenated into a single global feature histogram to represent the face image; the face was recognized by a simple distance-based grey-level histogram matching. Moreover, Demirel and Anbarjafari [6] proposed face recognition based on the probability distribution functions (pdf) of pixels in different color channels by using the Kullback-Leibler distance (KLD) between the probability distribution functions of the input face and the PDFs of the faces in the training set for face recognition. There is one researcher presents a classification method which histogram is done by the feature extraction and is classified by KLD.

In this paper, we proposed a classification method which is classified by Euclidean distance. We used histogram as input before face classification. Moreover, the histogram which used for an input for this paper, was applied to modify the histogram of value which is mean among of $j-1, j$ and $j+1$ value.

Manuscript received December 1, 2017; revised December 26, 2017. This work was supported in part by the develop and promotion of science and technology talents (DPST), Thailand.

Weenakorn Ieosanurak is with Khon Kaen University Thailand (e-mail: weenakorn.i@ kkumail.com)
Watcharin Klongdee is with Khon Kaen University Thailand. (corresponding author to provide phone: +66-83356-9335; e-mail: kwatch@ kku.ac.th).

II. MATERIALS AND METHOD

A. Kullback-Leibler divergence

Kullback-Leibler divergence (KLD), which is a measuring the difference between two probability distributions over the same variable x, is closely related to relative entropy, information for divergence and information for discrimination. Let $P(x)$ and $Q(x)$ be two probability distribution of discrete random variable x. That is both $P(x)$ and $Q(x)$ sum up to 1, and $P(x) > 0$ and for any x in X.

$D_{KLD}(P(x)||Q(x))$, which is a measure of the information lost when $Q(x)$ is used to approximate $P(x)$, is defined by

$$D_{KLD}(P(x)||Q(x)) = -\sum_{x\in X} P(x)\ln\frac{Q(x)}{P(x)}, \qquad (1)$$

and (1) is equivalent to

$$D_{KLD}(P(x)||Q(x)) = -\sum_{x\in X} P(x)\ln\frac{P(x)}{Q(x)}. \qquad (2)$$

The test image is classified into class C,

$$C = \arg\min\{D_{KLD}(P(x)||Q(x))\} \qquad (3)$$

where $P(x)$ is the training image and $Q(x)$ is the tested image.

B. Euclidean distance

Let X, Y be vectors of length n such that $X = [x_1, x_2, x_3, \dots, x_n], Y = [y_1, y_2, y_3, \dots, y_n]$. The Euclidean distance is defined

$$d_2(X,Y) = \left(\sum_{i=1}^{n} |x_i - y_i|^2\right)^{\frac{1}{2}}. \qquad (4)$$

III. Proposed algorithm

The training database contains M images and N class. Each image matrix is normalized by converting to the equivalent image vector (column matrix) x_i . The training matrix X contains the image vectors as $X = [x_1, x_2, \dots, x_M]$. The extracted features using histogram are used as input then we classify by Euclidean distance. Before start, we convert the entire color face image to the grayscale image.

Step 1. Input the training image and the extracted features using histogram are used, that is the input, sample k class i , is defined by

$$Y_k^i = [q_0, q_1, q_2, \dots, q_{255}],$$

where $q_j = \frac{h(j)}{\sum_{j=0}^{255} h(j)}$ and $h(j)$ is represented the number of pixels having grayscale j. We adjust histogram by defining a new histogram as follow

$$X_k^i = [p_0, p_1, p_2, \dots, p_{255}],$$

where $p_0 = q_0, p_{255} = q_{255}$ and

$$p_j = \frac{q_{j-1} + q_j + q_{j+1}}{3},$$

$j = 1, 2, 3, \ldots, 254$, and let X be the tested image, which is defined by

$$X = [p_0, p_1, p_2, \ldots, p_{255}].$$

Step 2. Classify the tested image by Euclidean distance. We calculate Euclidean distance for each image as follow:

$$g_i(X) = ||X - X_k^i||,$$

where $k = 1, 2, 3, \ldots, n_i$ and n_i is the total number of samples in the i^{th} class.

Step 3. The tested image X is in class $C(X)$:

$$C(X) = \arg_i \min\{g_i(X), i = 1, 2, 3, \ldots, N\}.$$

IV. Experiments and Results

We evaluated the proposed algorithms based on three datasets from Grimace [8], Jaffe [9] and Olivetti Research Laboratory (ORL) database [10] which are shown in Table I.

TABLE I
FACE RECOGNITION DATABASES

DATABASE	IMAGE	# CLASS	# SAMPLES PER CLASS
GRIMACE	180 x 200	18	20
JAFFE	256 x 256	10	20
ORL	112 x 92	40	10

The criterions of the experiments are as follows: Euclidean distance and Kullback - Leibler divergence (KLD). The databases are used by Grimace, Jaffe and ORL. Before we verify class of face image by proposed algorithm, we transform the training image and test image as column vector and find a new histogram then classify by using Euclidean distance. We divide the experiment into four parts. Firstly, the training set is 60% of all images in the class and the remaining 40% of all images was the test set. Secondly, the training set is 70% of all images in the class and the remaining 30% of all images was the test set. Thirdly, the training set is 80% of all images in the class and the remaining 20% of all images was the test set. Finally, the training set is 90% of all images in the class per class and the remaining 10% of all images was the test set; we chose the test image from the test set. The recognition rate is shown in Table II-IV. Moreover, we find a new histogram for Jaffe database. That is $p_0 = q_0,\ p_1 = q_1,\ p_{254} = q_{254},\ p_{255} = q_{255}$ and

$$p_j = \frac{q_{j-2} + q_{j-1} + q_j + q_{j+1} + q_{j+1}}{5},$$

where $j = 2, 3, \ldots, 253$.

TABLE II
THE RECOGNITION RATE OF GRIMACE DATABASE

Proportion (#train-#test)	Proposed algorithm	KLD
60-40	100	100
70-30	100	100
80-20	100	100
90-10	100	100

TABLE III
THE RECOGNITION RATE OF JAFFE DATABASE

Proportion (#train-#test)	Proposed algorithm	KLD
60-40	95	95
70-30	95	93.33
80-20	95	95
90-10	95	95

TABLE IV
THE RECOGNITION RATE OF ORL DATABASE

Proportion (#train-#test)	Proposed algorithm	KLD
60-40	98.12	97.5
70-30	98.33	98.33
80-20	98.75	98.75
90-10	100	100

V. Conclusion

In this letter, we introduced an algorithm for face recognition system based on adjusted histogram and Euclidean distance. Euclidean distance is used to perform face classification. Minimum Euclidean distance between histogram of a tested face and adjusted histogram of the faces in the database was used to perform histogram matching. The performance of proposed algorithm was compared using the Kullback-Leibler divergence (KLD). Meanwhile, the experiment results on Grimace, Jaffe and ORL face database.

Mostly, it is evident that the recognition rate of proposed algorithm is as same as KLD; besides, the recognition rate of proposed algorithm is better than KLD in case of Jaffe database (70-30 proportion) and ORL database (60-40 proportion). Moreover, if proposed algorithm is applied for color image, the recognition is more effective classification.

Acknowledgment

Authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

References

[1] M. Turk, A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.
[2] P. Relhumeur, J. Hespanha, and D. Kreigman, "Eigenfaces vs. fisher- faces: Recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 19, no. 7, pp. 7 11–720, Jul. 1997.
[3] K. Karande and S. Talbar, "Face recognition under variation of pose and illumination using independent component analysis," *ICGST - GVIP 2008*,, vol. 8, no.4, pp. 1–6.
[4] T.-W. Yoo and I.-S. Oh, "A fast algorithm for tracking human faces based on chromatic PDFs," *Proc. 9th European Conference on Computer Vision(ECCV)*, vol. 20, pp. 967–978, 1999.
[5] Y. Rodriguez and S. Marcel, "Face authentication using adapted local binary pattern pdfs," in *Proc. 9th European Conference on Computer Vision(ECCV)*, Graz, Austria, May 7–13, 2006, pp. 321?332.
[6] H. Demirel and G. Anbarjafari, "Pose invariant face recognition using probability distribution functions in different color channels," *IEEE Signal Process. Letter* 2008, vol. 15, pp. 537-540.
[7] H. Demirel and G. Anbarjafari, "Improved pdf based face recognition using data fusion," *International journal of electronics; mechanical and mechatronics engineering*, vol. 2, no. 2, pp. 195-200.
[8] L. Spacek *Description of the collection of facial images.* Available: http://cswww.essex.ac.uk/mv/allfaces.

[9] M.J. Lyons, S. Akemastu, M. Kamachi and J. Gyoba, "Coding Facial Expressions with Gabor Wavelets," *3rd IEEE International Conference on Automatic Face and Gesture Recognition*, Nara, Japan, 1998, pp. 200?205,1998.

[10] F. Samaria and A. Harter, "Parameterisation of a stochastic model for human face identification," *2nd IEEE Workshop on Applications of Computer Vision*. Dec. 1994, Sarasota (Florida).

ISBN: 978-988-14047-8-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)