Student Attendance System in Classroom Using Face Recognition Technique

Samuel Lukas[1], Aditya Rama Mitra[2], Ririn Ikana Desanti[3], Dion Krisnadi[4]
[1,4]Informatics Department, [2]Computer System Department, [3]Information System Department
Universitas Pelita Harapan
Karawaci, Indonesia
{samuel.lukas, aditya.mitra, ririn.desanti, dion.krisnadi}@uph.edu

Abstract—**Authentication is one of the significant issues in the era of information system. Among other things, human face recognition (HFR) is one of known techniques which can be used for user authentication. As an important branch of biometric verification, HFR has been widely used in many applications, such as video monitoring/surveillance system, human-computer interaction, door access control system and network security. This paper proposes a method for student attendance system in classroom using face recognition technique by combining Discrete Wavelet Transforms (DWT) and Discrete Cosine Transform (DCT) to extract the features of student's face which is followed by applying Radial Basis Function (RBF) for classifying the facial objects. From the experiments which is conducted by involving 16 students situated in classroom setting, it results in 121 out of 148 successful faces recognition.**

Keywords-Discrete Wavelet Transform,Discrete Cosine Transform, Radial Basis Function Neural Network

I. INTRODUCTION

In traditional face-to-face (F2F) class setting, student attendance record is one of the important issues dealt with any school, college and university from time to time. To keep the student attendance record valid and correct, the faculty staff should have a proper mechanism for verifying and maintaining or managing that attendance record on regular basis. In general, there are two types of student attendance system, i.e. manual attendance system (MAS) and automated attendance system (AAS). By practicing manual recording, faculty staff may experience difficulty in both verifying and maintaining each student's record in classroom environment on regular basis, especially in classes attended by a large number of students. In practice, the manual system also requires more time for recording and calculating the average attendance of each enrolled student. On the other hand, the automated attendance system may offer some benefits to to the faculty, at least it may lessen the administrative burden of its staff. Particularly, for attendance system which adopts human face recognition (HFR) technique, such a system commonly involves the process of extracting key features from any facial image of student captured at the time he/she is entering the classroom, or when everybody already occupies his/her seat in the classroom. Upon its succesful recognition, it proceeds to marking that recognized student's attendance automatically. Following that general idea, the discussion of this paper is based on the known face recognition techniques in its endeavor to develop a specific computer application which can be used for recognizing any enrolled student automatically from the digital images captured in classroom.

In general, there are two known approaches to HFR, i.e. feature-based and brightness-based approach. The feature-based approach uses key point features of the face, such as edges, eyes, nose, mouth, or other special characteristics. Therefore, the calculation process only covers some parts of the given image that have been extracted previously. On the other hand, the brightness-based approach calculates all parts of the given image. It is also known as holistic-based or image-based approach.

Since all parts of the image have to be considered, the brightness-based approach takes longer time to process the image and is also more complicated. To make the process short and simple, the image has to be transformed into a certain model. Among many models which are available, there is one that was introduced by Turk and Pentland in 1991. They developed their recognition system based on the Principle Component Analysis (PCA) method [1][2][3]. Other proposed models which are popular are Discrete Wavelet Transform (DWT) [4] and Discrete Cosine Transform [5][6][7][8].

This paper covers the combination use of DWT and DCT in HFR and also refers to the key results obtained from the preliminary research of the same project as published in [13].

II. FEATURE EXTRACTION AND RECOGNITION SYSTEM

A. Discrete Wavelet Transform

The Discrete Wavelet Transfor (DWT) has been introduced to be a highly efficient and flexible method for dealing with subband decomposition of a signal. Nowadays, 2D-DWT is established as a key operation in image processing. It is known also as a multi-resolution analysis which decomposes any image into its wavelet coefficients and scaling function [10]. It is easily seen that wavelets also have rough edges. Nevertheless, they are able to render pictures better by eliminating the blockiness. This elimination process is done with the help of filters with different cut-off frequencies at different scales. Furthermore, it is easy to implement which in turn may reduce the computation time and resources required [11]. 2-D DWT operates in a straight forward manner by inserting array transposition between the two 1-D DWTs. The rows of the array are processed first with only one level of decomposition. This essentially divides the array into two vertical halves. The first half stores the average coefficients, while the second half stores the detail coefficients. This process then is repeated in the orientation of its columns which results in four subbands for each decomposition level within the array which is defined by the filter output (Fig. 1). Figure 1 shows a two-level 2-D DWT decomposition of a given image [12].

Figure 1.Two level decomposition for 2D-DWT

B. *Discrete Cosine Transforms*

Discrete Cosine Transforms (DCT) is actually derived from the Discrete Fourier Transform (DFT). Since objects are in the form of images then consequently 2D-DCT can be implemented [7][8]. In this context, spatial domain of any image $I(x,y)$ will be transformed to frequency domain as $C(u,v)$ as expressed in Eq. 1 & Eq. 2; while its inverse is expressed in Eq. 3.

$$C(u,v) = \alpha(u)\alpha(v)\sum_{r=0}^{row}\sum_{c=0}^{col} I(r,c)\cos\left[\frac{(2r+1)u\pi}{2N}\right]\cos\left[\frac{(2c+1)v\pi}{2N}\right] \quad (1)$$

$$\alpha(u),\alpha(v) = \begin{cases} \sqrt{\frac{1}{N}} & for\ u,v = 0, \\ \sqrt{\frac{2}{N}} & for\ u,v = 1,2,\dots,N-1 \end{cases} \quad (2)$$

$$I(r,c) = \sum_{u=0}^{row}\sum_{v=0}^{col} \alpha(u)\alpha(v)C(u,v)\cos\left[\frac{(2r+1)u\pi}{2N}\right]\cos\left[\frac{(2c+1)v\pi}{2N}\right] \quad (3)$$

Properties of DCT are de-correlation, energy compaction, domain scaling, separability, and symmetry. De-correlation means that there is no correlation in calculating among all the DCT coefficients. Therefore, all DCT coefficients can be calculated independently. DCT exhibits excellent energy compaction for highly correlated images. Efficacy of a transformation scheme can be directly gauged by its ability to pack input data into as few coefficients as possible without introducing visual distortion in the reconstructed image significantly. However, DCT is not scaling invariant. This implies that all images used, either for training or identification, have to be uniform in size. By separability it means that the DCT coefficients can be computed in two steps through successive 1-D operations which are applied on rows and columns of an image. This computation is expressed in Eq. 4. Another look at the row and column operations in that equation reveals that these operations are functionally identical. Such a transformation is called a *symmetric transformation* [9].

$$C(u,v) = \alpha(u)\alpha(v)\sum_{r=0}^{row}\cos\left[\frac{(2r+1)u\pi}{2N}\right]\sum_{c=0}^{col} I(r,c)\cos\left[\frac{(2c+1)v\pi}{2N}\right] \quad (4)$$

C. *Radial Basis Function Network*

Radial Basis Function (Neural) Network (RBFN) is essentialy derived from Multi-Layer Perceptron Network (MLPN), but RBFN takes a slightly different approach. The hidden nodes in RBFN implement a set of radial basis functions (e.g. Gaussian functions), while the output nodes implement linear summation functions as similarly found in an MLP. The network training is divided into two stages: first, the weights from the input to hidden layer are determined, and then the weights from the hidden to output layer are calculated. The training or learning process of RBFN is known to be very fast. The general configuration of RBFN for *P* input nodes with *Q* hidden nodes and *R* output nodes can be consulted in [10]. The example scheme/structure of such a RBFN is illustrated in Figure 2.

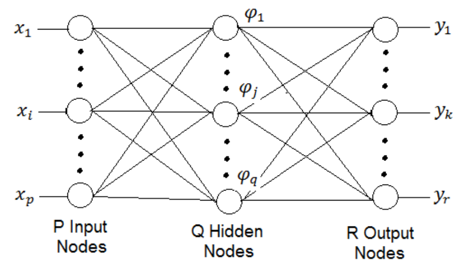

Figure 2. Structure of RBFN

The primary goal of RBFN is to find such a function $f: x^p \to y^r$ that it can interpolate a set of *N* data points in a *p*-dimensional input space, $\boldsymbol{x} = (x_i x_2 \dots x_p)$, which are to be mapped onto the *r*-dimensional output space, $\boldsymbol{y} = (y_i y_2 \dots y_r)$. For every hidden node, it has a center vector, $\boldsymbol{x_c} = (x_{c_1} x_{c_2} \dots x_{c_p})$ and a variance, σ_c^2. The output of every hidden node is expressed in Eq. 5. By composing a linear combination with the weights, W_{kj}, from hidden nodes to the output nodes, the output of the RBF can be expressed as in Eq. 6.

$$\varphi(\|\boldsymbol{x} - \boldsymbol{x_c}\|),\ \varphi(a) = exp\left(-\frac{a^2}{2\sigma^2}\right) \quad (5)$$

$$\boldsymbol{y_k} = f(\boldsymbol{x}) = \sum_{j=1}^{q} W_{kj}\varphi_j\left(\left\|\boldsymbol{x} - \boldsymbol{x_{c_j}}\right\|\right) \quad (6)$$

III. SYSTEM DESIGN

The block diagram of the proposed system is presented in Figure 3. In the shown diagram, facial training image is a set of student facial image as training data. In this research, 186 student facial images are used which are created from 16 students. The number of student facial images of each student varies among others. In Figure 4 some of those training data is exhibited.

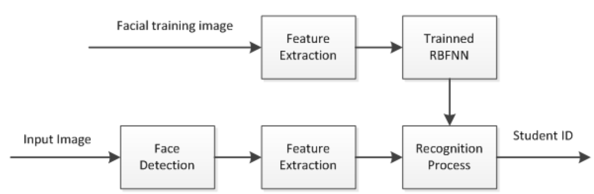

Figure 3. Block diagram of the system

Figure 4. Five example facial training images for each student

As its name reveals, feature extraction is involved with a purpose to extract features from any student's facial image that is required to come in a uniform size, in this case, 64x64 pixels. The process is completed actually by performing grayscale normalization, histogram equalization, Discrete

Wavelet Transform (DWT), and Discrete Cosine Transform (DCT). This process is illustrated in Figure 5.

Figure 5. Three student facial training image

Suppose p-DCT coefficients are used as the features of an image, and there are n images in training set with k classes which are labeled as $c_i, i = 1,2,\dots,k$, then the input data set can be represented as a $n \times k$ matrix denoted as matrix $X = \{x_{ij} \mid i = 1,2,\dots,n \;;\; j = 1,2,\dots,k\}$.

The outputs of the trained RBFN are the average and the variance of each hidden layer nodes, μ_i, σ_i^2 and also the weights of RBFN from hidden layer nodes to the output layer nodes. μ_i and σ_i^2 are parameters of i^{th} node in the hidden layer. They are called as the vector centres (Eq. 7) and variances (Eq. 8) of the training data of i^{th} class from k classes respectively.

$$\mu_{ij} = \frac{1}{|C_i|}\sum_{i \in C_i} x_{ij} \qquad i = \{1,\dots,k\} \;;\; j = 1,2,\dots,p \tag{7}$$

$$\sigma_i^2 = \frac{1}{p|C_i|}\sum_{i \in C_i}\sum_{j=1}^{p}(x_{i,j} - \mu_{ij})^2 \quad i = 1,2,\dots,k \tag{8}$$

The weights of RBFN is denoted as a square matrix W of size $k \times k$. For each data in the training set, the output of q^{th} node in the hidden layer is expressed in Eq. 9, and the target output of node q is formulated in Eq. 10. It is clear that $\{h_q\}$ can be formed as a matrix H of size $n \times k$, and $\{t_q\}$ in the form of T of size $n \times k$. RBFN weights from hidden layer to output layer are $W = \{w_{ij}\}$ as expressed in Eq. 11.

$$h_q = \varphi_q(x_{ij}) = e^{\left(-\frac{\sum_{j=1, i \in C_i}^{p}(x_{ij} - \mu_{qj})^2}{2\sigma_q^2}\right)} \qquad q = 1,2,\dots,k \tag{9}$$

$$t_q(x_{ij}) = \begin{cases} 1 & if\ q = c_i \\ 0 & others \end{cases} \qquad q = 1,2,\dots,k \tag{10}$$

$$W = (H^T H)^{-1} H^T T \tag{11}$$

Face detection is performed to get student's facial image from the input image. By input image, it is an image of students captured while they are in the classroom listening to the lecture (Fig. 6).

Figure 6. Example of an input image

In the recognition process, any single facial image is extracted by p DCT coefficients to produce a vector $\boldsymbol{x} = (x_1 x_2 \dots x_p)$. While the output of hidden layer is presented in the form of $\boldsymbol{h} = (h_1 h_2 \dots h_k)$ (Eq. 12), and the RBFN output is $\boldsymbol{o} = (o_1 o_2 \dots o_k)$ (Eq. 13).

$$h_i = \varphi_i(\boldsymbol{x}) = e^{\left(\frac{\|x - \mu_i\|^2}{2\sigma_i^2}\right)} \qquad i = \{1,\dots,k\} \tag{12}$$

$$\boldsymbol{o} = \boldsymbol{h}W \tag{13}$$

It is clear that the index j with the highest value for component o_j indicates that it is the index of the expected class of the given facial image. Each index is associated to a unique student ID.

IV. Experiment Result and Discussion

In the experiment design, the number of features of any facial student image is set to be constant, i.e. 16 DCT coefficients. With this setting the experiment is run to provide the answer to the question of which level of DWT that will achieve the best recognition level from the given data. From the results presented in Table 1, it can be seen that out of 148 student facial images with level of DCT of 2, 121 facial images can be recognized succesfully giving a total level of recognition of 82%. This figure is perceived as the best recognition rate which can be obtained from the data.

TABLE I. Level of DWT and Level of Recognition

Student Name	#	Level of DWT			Level of Recognition		
		1	2	3	1	2	3
Alvin	11	10	10	9	0.91	0.91	0.82
Evan	6	2	4	4	0.33	0.67	0.67
Felix	10	7	9	9	0.70	0.90	0.90
Gerry	9	8	8	8	0.89	0.89	0.89
Harry	7	6	5	5	0.86	0.71	0.71
Ivan	11	9	9	7	0.82	0.82	0.64
Jason	7	5	5	5	0.71	0.71	0.71
Johan	11	8	8	6	0.73	0.73	0.55
Keren	12	9	10	10	0.75	0.83	0.83
Kim	12	10	11	9	0.83	0.92	0.75
Leo	5	4	4	4	0.80	0.80	0.80
Leon	9	7	7	8	0.78	0.78	0.89
Livia	9	8	7	6	0.89	0.78	0.67
Marcel	11	9	10	10	0.82	0.91	0.91
Veri	13	8	10	10	0.62	0.77	0.77
Wong	5	4	4	4	0.80	0.80	0.80
Total	148	114	121	114	0.77	0.82	0.77

Further analysis of the failure in recognizing the rest facial images indicates that a student may be possibly recognized as other student(s). Two facial images of Jason are accidentally recognized as Leo and Kim; while three of Veri's images are recognized as Kim, Ivan and Johan (Fig. 7). However, there is also possibility in which case two facial images of Evan are recognized falsely as Leo, Harry as Alvin, and Leon as Ivan (Fig. 8).

Figure 7. Multiple recognitions of one student

Figure 8. One-to-one recognition of one student

By considering the total level of recognition as resulted from the experiment which does not meet high expectation, further research should be dealing with the problem exhibited from the choice of the techniques used in this research. More precisely, the special attention should be paid to the improvement of the feature extraction or recognition process technique to be used.

V. CONCLUSION

It can be concluded that automated student attendance system in classroom using human face recognition technique works quite well. Certainly, it can be improved for yielding a better result particularly by paying attention in feature extraction or recognition process. This improvement may help the recognition process become more robust. The success rate of the proposed system in recognizing facial images of the students who are seated in classroom is about 82%.

ACKNOWLEDGMENT

This research is funded by Indonesian Ministry of Research, Technology and Higher Education, Directorate General of Higher Education, Kopertis Wilayah III, No. 788/K3/KM/SPK.LT/2016 and 168/LPPM-UPH/VII/2016

REFERENCES

[1] M. Turk and A. Pentland, “Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp.71-86, 1991

[2] K. I. Diamantaras and S. Y. Kung, *Principal Component Neural Networks: Theory and Applications*, John Wiley & Sons, Inc., 1996.

[3] A. Pentland, B. Moghaddam, and T. Starner, “View-based and modular eigenspaces for face recognition,” *IEEE Conf. on Computer Vision and Pattern Recognition*, MIT Media Laboratory Tech. Report No. 245 1994

[4] M. Alwakeel and Z. Shaaban, “Face recognition based on haar wavelet transform and principal component analysis via Levenberg-Marquardt back propagation neural network,” *European Journal of Scientific Research*, vol.42, No.1, pp.25-31, 2010.

[5] V. V. Kohir andU. B. Desai, *Face Recognition Using a DCT-HMM Approach*, Indian Institute of Technology, Mumbai, India, 1998.

[6] S. E. Handy, S. Lukas, and H. Margaretha, “Further tests for face recognition using discrete cosine transform and hidden Markov model,” in *Proc. International Conference on Electrical Engineering and Informatics (MICEEI)*, Makasar, 2012.

[7] S. A. Khayam, *The Discrete Cosine Transform (DCT): Theory and Application*, Michigan State University, 2003

[8] M. H. Muchri, S. Lukas, and D. H. Hareva, “Implementation discrete cosine transform and radial basis function neural network in facial image recognition,” in *Proc. International Conference on Soft Computing Intelligent Systems, and Information Technology*, 2015

[9] B. J. Oh, “Face recognition using radial basis function network based on LDA,” *International Journal of Computer, Information Science and Engineering*, vol. 1, pp. 401-405, 2007.

[10] A.A. Blessie, J. Nalini, and S. C. Ramesh, “Image compression using wavelet transform based on the lifting scheme and its implementation,” *International Journal of Computer Science Issues*, vol. 8, no. 1, 2011.

[11] Y. S. Bute and R.W. Jasutkar, “Implementation of discrete wavelet transform processor for image compression,” *International Journal of Computer Science and Network* (IJCSN), vol. 1, issue 3, 2012

[12] D. Gupta and S. Choubey, “Discrete wavelet transform for image processing,” *International Journal of Emerging Technology and Advanced Engineering*, vol. 4, issue 3, March, 2015

[13] L. Samuel., A. R. Mitra, R. I. Desanti., D. Krisnadi., “Implementing Discrete Wavelet and Discrete Cosine Transform with Radial Basis Function Neural Network in Facial Image Recognition”, Journal of Image and Graphics, Vol 4, No 1, June 2016