A New Approach in Solving Illumination and Facial Expression Problems for Face Recognition

Yee Wan Wong, Kah Phooi Seng, Li-Minn Ang

[a] *The University of Nottingham Malaysia Campus*
Jalan Broga
43500 Semenyih, Selangor
Tel : 03-89248358, Fax : 03-89248017
E-mail : yeewan.wong@nottingham.edu.my

ABSTRACT

In this paper, a novel dual optimal multiband features (DOMF) method is presented to increase the robustness of face recognition system to illumination and facial expression variations. The wavelet packet transform first decomposes image into low-, mid- and high-frequency subbands and the multiband feature fusion technique is incorporated to select the subbands that are invariant to illumination and expression variation separately. These subbands form the optimal feature sets. Parallel radial basis function neural networks are employed to classify these feature sets. The scores generated by the neural networks are combined by an adaptive fusion mechanism where the level of illumination variations of the testing image is estimated and the weights are assigned to the scores accordingly. The experimental results show that DOMF outperforms other algorithms and also achieves promising performance on illumination and facial expression variation conditions

Keywords

Face recognition, multiband features, wavelet packet transform, illumination variation, adaptive fusion, neural network

1.0 INTRODUCTION

Face recognition system is an automatic system that identifies a person identity using human facial characteristics. It has received significant attention because of its wide range of applications (Chellappa, Wilson, & Sirohey, 1995). However, uncontrolled factors such as illumination variation, facial expression, facial occlusion, pose variation and aging impose challenges to face recognition system. Among these factors, illumination variation is one of the significant factors affecting the performance of the face recognition system [].

Methods have been proposed to solve the illumination problem. However, most of these existing methods consider only illumination problem (Ekenel & Sankur, 2005; Shin, Lee & Kim, 2008, Zhang et al. 2009) but not with facial expression problem. This is because illumination variation affects the low-frequency component or global appearance of a face image (Adini, Moses & Ullman, 1997), whereas facial expression variation affects the high-frequency components (Naster & Ayach, 1996). Hence, the compensation for one kind of variation causes adverse effect on the other. For example, the experimental results in (Ekenel & Sankur, 2005) showed that the wavelet low-frequency subband achieved high recognition rate against changes in expression but the performance degraded with the presence of illumination variations. Within the eigen-subspace domain, the three significant principle components are discarded to reduce the illumination variation (Belhumeur et al., 1997). However, in order to maintain system performance for image without illumination variation and improve performance for image with illumination variation, the first three principle components are assumed to capture illumination variation only.

Most recently, Jadhav and Holambe (2008) presented a face recognition system based on combination of Radon and wavelet transform, which is invariant to illumination and facial expression variations. The DC component of the low-frequency subband was removed when testing the algorithm performance in illumination variation. They showed that their proposed system achieved high recognition accuracy in the variation of facial expression and illumination separately. However, the system performance may degrade when it is tested against the combination of variations due to the removal of the DC component.

In this paper, a novel dual optimal multiband feature (DOMF) method for face recognition is presented. The wavelet packet transform (WPT) (Primer, 1998) decomposes the image into frequency subbands to represent the facial features of the face image. The multiband feature fusion technique which was proposed in one of our research papers (Wong et al., 2009) is incorporated to search for subbands that are invariant to illumination and facial expression variation separately. Parallel radial basis

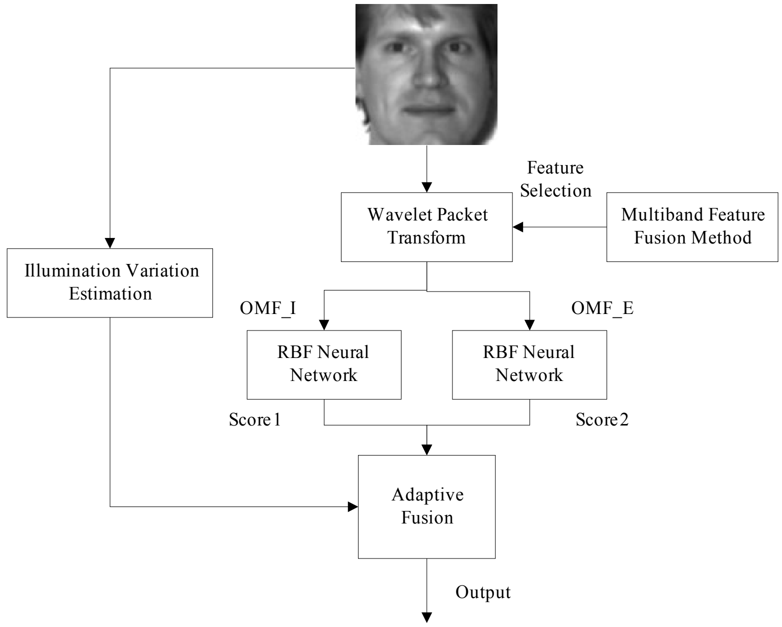


Figure 1: Block diagram of the proposed DOMF method for face recognition system

function (RBF) neural network (Ranganath & Arun, 1997) is used to classify the two sets of feature. An adaptive fusion mechanism is proposed to combine the two feature sets where the weights assigned to both feature set are based on the illumination variation level of the input image. Therefore, our system, DOMF method, can reduce the effect of expression and illumination, and achieve a good recognition performance under different variation without causing adverse effect. Experimental result based on different databases show that DOMF method outperforms radon and wavelet and principle component analysis (PCA) (Turk & Pentland, 1991) under various image conditions.

This paper is organized as follows. The proposed DOMF method is presented in Section 2. Experimental results are given in Section 3, which compare the performance of DOMF to other face recognition algorithms based on the Extended Yale database B (Lee et al., 2005), the YaleB database (Georghiades, Belhumeur & Jacobs, 2001), the AR database (Martinez & Benavente, 1998) and the ORL database (The ORL in Cambridge, UK, http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html). Finally, concluding remarks are given in Section 4.

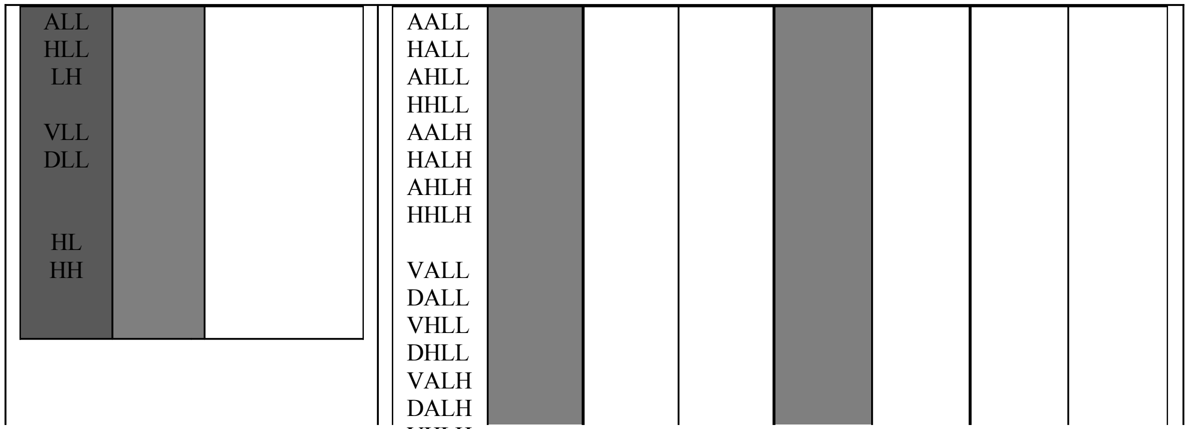
2.0 Face Recognition using DOMF method

This section describes the proposed DOMF method. Section 2.1 provides an overview of the proposed DOMF method. Section 2.2 describes the multiband feature fusion method which is used to search for features that are invariant to illumination and facial expression variations. The classification and fusion method which incorporate the illumination variation estimation and adaptive fusion are discussed in Section 2.3.

2.1 Overview

Figure 1 shows the block diagram of the proposed DOMF method for face recognition system. The aims of the proposed DOMF are to extracts the optimal sets of subband that are invariant to facial expression and illumination and to avoid the adverse effect by introducing an adaptive fusion method to combine the optimal feature sets. The multiband feature technique is incorporated in the system to search for expression and illumination invariant facial features. WPT fully decomposes the face image into low-, mid- and high-frequency subbands. Statistical method (Feng, Yuen & Dai, 2000) and recognition accuracy are used to test the discriminative performance and recognition performance of the subbands. The subbands that obtain high discriminative and recognition performance will be selected. The best performing subbands are combined to further increase the recognition performance of the system. The optimal set of subband is named as Optimal Multiband Feature (OMF). The OMF for expression and illumination are named as OMF_E and OMF_I respectively. Parallel RBF neural networks are used to classify the OMF_E and OMF_I. The RBF neural network generates two sets of score for both the OMFs. The scores are linearly combined through a set of fusion weight.

In the proposed method, the weights are determined by the illumination variation estimator where the illumination variation factor will be assigned based on the illumination variation level of the input image. Then, the final decision is generated with respect to the weights of the system which



are adaptively assigned based on the illumination variation factors.
The details of the DOMF are described in the following subsections.

2.2 Multiband Feature Fusion Technique

We first locate the facial features that are invariant to facial expression and illumination variation. Naster et al. (Naster & Ayach, 1996; Naster, Moghaddam & Pentland, 1997) found that changes in illumination affect the low-frequency spectrum. This statement indicates that high-frequency components are invariant to illumination. On the other hand, the facial expression variation only affects the high-frequency spectrum. This means that compensation for one variation may have an adverse effect on another. To avoid this scenario, we therefore proposed a multiband feature technique to search for the frequency subbands that are invariant to illumination variation and facial expression variation separately. It is important to note that other than LL being decomposed further, LH, HL and HH are also being decomposed further in the multiband feature fusion technique.

Recognition accuracy and statistical analysis based on class separation are used to evaluate the recognition performance of the subbands. Recognition accuracy shows how well the system can match images from the same people and class separation shows how well the system can distinguish images from different people (Feng, Yuen & Dai, 2000). To test the class separation, *N* face images from a database, one image per person are used. The face images are chosen randomly from the training and testing sets. The similarity matrix with the size *N*x*N* records the similarity between image *i* and image *j*. For a good representation, $\rho(i,j)$ should be closed to one if $i = j$ and should be close to zero if $i \neq j$. The Average Unmatched Similarity Value, AUMSV (Feng, Yuen & Dai, 2000) that is defined as below,

$$AUMSV = \frac{1}{(N^2 - N)} \sum_{i=1}^{N} \sum_{j=1}^{N} \rho(i,j) \qquad i \neq j \quad (1)$$

is used to give a single numerical value to the similarity performance of the subband. This term shows how well the subband representation distinguishes the images from different people, and it ranges from 0 to 1, which means the higher the discriminatory power, the smaller the AUMSV value.

Below are the steps proposed to select the optimal subbands:

Step 1: Compute the AUMSV and recognition accuracy in each subbands from level 1 and 2.

Step 2: The subbands that obtain AUMSV that is lower than 0.5 and recognition accuracy that is higher than half of the recognition rate of the original image will be selected for further decomposition to level 3. The threshold values determine the computational complexity of the system. This step reduces the computational complexity by avoiding decomposition of all subbands from level 2 to level 3.

Step 3: Further decompose subbands that fulfill the subband selection criteria to level 3 decomposition.

Step 4: Two best performing subbands in terms of AUMSV in level-3 decomposition will be concatenated.

The optimal feature set for facial expression and illumination are named as OMF_E and OMF_I respectively. The location of the frequency subbands OMF_E and OMF_I are shown in Fig. 2 (a) and (b) respectively. The figure shows that the combination of LL and ALL forms the OMF_E and the combination of HALL and AALH forms the OMF_I.

2.3 Classification and Fusion

In the previous section, the multiband feature fusion technique which is used in obtaining the OMF_E and

OMF_I was discussed. In this section, the illumination variation estimation is presented. This estimator is implemented prior to the combination of the OMF_I and OMF_E to obtain the illumination variation factor which influences the weight assignment of the system. In order to be robust to face image under different illumination variation, we first take the logarithm transform of the image. Hence, we have $I' \approx \log(I)$ where I represents the original image. One reason for transforming I into logarithm domain is that the logarithm of the image can reduce the effect of luminance. Another reason is to reduce the pixel value of the original image.

The illumination variation estimation based on morphological opening is then applied on the I'. By using morphological opening, the facial features of the image can be removed and therefore the illumination variation of the image can be estimated. The level of illumination variation can be described as the illumination variation factor k. Assuming that the face image that contains illumination variation has one side of the image brighter than the other, the k can be determined as the difference between the mean pixel value at the left and the right sides of the image.

After obtaining the illumination variation factor , the weight of the system can be determined adaptively based on the value of the testing image. As discussed ealier, parallel RBF neural networks will generate two sets of score. The sum rule is incorparated in this system to combine the scores. Sum rule computes the final score from (Alexandre, Campilho & Kamel, 2001)

$$s = \sum_{i=1}^{J} w_i s_i$$

(2)

where is the number of modalities (which is two in this case), w_i are a set of fusion weights and are the scores obtained from the modalities. The fusion weights are adaptive in the sense that the weights assigned to the modalities are based on the illumination variation factor of the testing image. The weight for each image is determined with the following definition

$$w_i = \begin{cases} w & k \geq T \\ 1 - w & k < T \end{cases}$$

(3)

T denotes the threshold of the illumination variation factor where it is determined as the maximum value of the illumination variation factors of the training images. The variation w is fixed and will be obtained experimentally and the results will be shown in the next section.

3.0 EXPERIMENTAL RESULTS

In this section, we first evaluate the recognition performance of the OMF_E and OMF_I in facial expression and illumination variations respectively. Then the recognition performance of DOMF will be evaluated. The databases used included the EYaleB database, the YaleB database, the AR database and the ORL database. The EYaleB and YaleB database are often used to investigate the effect of illumination variation in face recognition. It contains ligting from different angles. In the AR database, other than ligting from left and right, it also includes lighting from both sides of the face image. Besides, it contains facial expression variations. The ORL database includes facial expression and perspective variations.

3.1 OMF_E and OMF_I performance evaluation

The first part of this experiment is to select the OMF_E using multiband feature fusion technique. The nearest neighbor classifier based on Euclidean distance was used for classification in this experiment only. In Wong, Seng & Ang (2009), we found that the combination of HALL (high-frequency subband) and AALH (mid-frequency subband) forms the OMF_I. Hence, in the following experiment, OMF_I refers to this form of subband combination. The recognition performance of the OMF_I in EYaleB and AR database are shown in Table 1.

The ORL database and AR database were used in locating OMF_E. The ORL and AR database were divided into sub-classes that only contain facial expression variation. In ORL database, 9 subjects were chosen. There were 18 samples images with normal facial expression were used for training and 63 samples images with different facial expressions were used for testing. In AR database, 100 subjects were used. There were 200 samples images with normal facial expression were used for training and 200 samples images with facial expression variation were used for testing. All images were scaled to 32×32 pixels resolution. To test AUMSV, N=9 for ORL and N=100 for AR database.

The two best performing subbands in ORL and AR databases were LL and ALL. The LL achieved the lowest AUMSV of 0.299 and 0.379 and the highest recognition rate of 96.8% and 79.5% in ORL and AR databases respectively. The second best performing subband is ALL, which is the low-frequency subband from the second level of decomposition. The ALL achieved AUMSV of 0.356 and 0.383 and recognition rate of 95.2% and 76.5% in ORL and AR databases respectively. Combining LL and ALL forms the OMF_E. Table 2 shows that OMF_E achieved recognition rate of 99% and 81.5% in ORL and AR database respectively.

Table 1: *Recognition rate (%) of the OMF_I in EYaleB and AR database*

Subband	EYaleB	AR
OMF_I	81.6	84.5

Table 2: *Recognition rate (%) of the OMF_E in ORL and AR database*

Subband	ORL	AR
OMF_E	99.0	81.5

3.2 DOMF performance evaluation

After obtaining the OMF_E and OMF_I, they will be combined by the DOMF method. In this experiment, the performance of DOMF with different weight values were tested based on the different databases. The databases used were not divided into sub-classes. The number of distinct subjects, the number of training images and the number of testing images in the respective databases are tabulated in Table 3. The face images in different databases are captured under different conditions, such as facial expression varations, illumination variation, etc. In each database, the training images were chosen to be images with face image of frontal view, and under even illumination and neutral facial expression, others formed the testing set. To future reduce the effect of uneven illumination, the logarithm transform was applied to the testing images in YaleB database.

The weight value w was chosen based on the experimental results. The weight values were from 0 to 1 with the interval of 0.1. Simulation result showed that the weight value of 0.6 achieved the highest recognition rate for all three databases. Hence, the 0.6 was applied as the weight value in the proposed system.

Table 3: *The databases used in this experiment*

	AR	ORL	YaleB
Number of subjects	100	40	10
Number of training images	300	80	50
Number of testing images	300	320	600

After obtaining the weight, the performance of the PCA, radon and wavelet and DOMF methods were tested on the different databases. For PCA, all the eigenfaces available for each testing of database were used. For the radon with wavelet technique, 180 radon projections and three level of Daubechies wavelet transform were used as the setting. RBF neural networks were used for the classification in this section. The relative performances of the different methods were shown in Table 4. The results show that the proposed DOMF achieved high recognition rate in different databases. This shows that the proposed DOMF method is robust to illumination and facial expression variations. The adverse effect of compensating one kind of the variations was avoided.

The DOMF achieved recognition rate of 98%, 88.4% and 91% in AR, ORL and YaleB database respectively. From the Table, we see that DOMF outperformed all the other algorithms tested in terms of recognition rate on the different databases. The result also shows that the DOMF achieved the high recognition rate in ORL database which contains faces rotated out of the image plane.

Table 4: *Face Recognition rates under different facial expression on different databases*

Recognition rate (%)	PCA	Radon and wavelet	DOMF
AR	86.0	50.0	98.0
ORL	16.3	74.4	88.4
YaleB	85.5	37.7	91.0

4.0 CONCLUSION

In this paper, a novel dual optimal multiband feature (DOMF) method for face recognition was presented. This method aimed to increase the robustness of face recognition system to illumination and facial expression variations. In our approach, the multiband feature fusion technique was incorporated to search for subbands that were invariant to illumination and facial expression variation separately. The optimal multiband features that were found to be invariant to illumination variation and facial expression variation, namely OMF_I and OMF_E respectively were combined by the adaptive fusion method. Parallel RBF neural networks were employed for the classification of the two set of features. The adverse effect of compensating one kind of the variations was avoided by estimating the level of illumination variations of the input image and the weights were assigned to the modalities accordingly. Experimental results showed that OMF_I and OMF_E achieved high recognition rate under different illumination and facial expression variation respectively. The paper also compared the recognition performance of DOMF method with other face recognition algorithms on different databases. The experimental results showed that DOMF outperformed other algorithms tested and also achieved consistent and promising performance on different conditions.

REFERENCES

Adini, Y., Moses, Y., and Ullman, S. (1997). Face recognition: the problem of compensating for changes in illumination direction, *IEEE Trans. Pattern Anal. Mach. Intell*., 18(11), 1067-1079.

Alexandre, L. A., Campilho, A. C. and Kamel, M. (2001). On combining classifiers using sum and products rules. *Pattern Recognition Letters,* 22, 1283-1289.

Belhumeur, P. N., Hespanha, J. P., and Kriegman D. J. (1997). Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Trans. Pattern Anal. Mach. Intell*., 19, 711-720.

Chellappa, R., Wilson, C. L., & Sirohey, S. (1995). Human and machine recognition of faces: A survey. *Proceedings of the IEEE*, 83, 705-740.

Chen, L. F., Liao, H.-Y. M., Ko, M.-T., Lin, J.-C. and Yu G.-J. (2000). A new-LDA-based face recognition system which can solve the small sample size problem. *Pattern Recog*. 33, 1713-1726.

Ekenel, H. K. and Sankur, B. (2005). Multiresolution face recognition. *Image and Vision Computing*, 23, 469-477.

Fang, T., Yuan, B., Y., Tang, Y. Y., Shang, Z., Lui, D., Lang, F. (2008). Multiscale facial structure representation for face recognition under varying illumination. *Pattern Recognition*, 42, 1-15.

Feng, G. C., Yuen, P.C. & Dai, D. Q. (2000). Human face recognition using PCA on wavelet subband. *Journal of Electronic Imaging*, 9, 226-233.

Georghiades, A. S., Belhumeur, P. N., and Jacobs, D. W. (2001). From few to many: illumination cone models for face recognition under variable lighting and pose. *IEEE Trans. Pattern Anal. Mach. Intel*., 23(6), 630-660.

Gonzalez, R. C. and Wood, R. E. (1993). *Digital Image Processing*. Addison-Wesley, Reading, MA.

Jadhav, D. V. and Holambe, R. S. (2008). Feature extraction using Radon and wavelet transforms with application to face recognition. Neurocomputing, 1-9.

Lee, K. C., Ho. J., and Kriegman D. (2005). Acquiring Linear Subspaces for Face Recognition under Variable Lighting. *IEEE Trans. on PAMI*, Vol. 27, 5, 684-698.

Martinez, A.M. and Benavente, R. (1998). The AR face database. *CVC Tech. Report,* 24.

Naster, C., and Ayache N. (1996). Frequency-based non-rigid motion analysis. *IEEE Trans. Pattern Anal. Mach. Intell*., 18, 1067-1079.

Naster, C., Moghaddam, B., Pentland, A. (1997). Flexible images: matching and recognition using learned deformations. *Comput. Vision Image Understanding*, 65(2), 179-191.

Primer, A. (1998). *Introduction to wavelet and wavelet transform*. Prentice Hall.

Ranganath, S., Arun, K. (1997). Face Recognition Using Transform Features and Neural Networks. *Pattern Recog*, 10, 1615-1622.

Shin, D., Lee, H.-S., and Kim, D. (2008). Illumination-robust face recognition using ridge regressive bilinear models. *Pattern Recognition Letters*, 29, 49-58.

Turk, M., and Pentland, A. (1991). Eigenfaces for recognition. *J. Cogn. Neurosci*, 3 (1), 71–86.

Wong, Y. W., Seng, K. P. and Ang, L.-M. (2009). The Audio-visual Authentication System Over Internet Protocol. *Proceedings of the International MultiConference of Engineers and Computer Scientists 2009 (IMECS 2009)*, *The 2009 IAENG International Conference on Imaging Engineering (ICIE'09),* Vol. I, 938-943.

The ORL in Cambridge, UK, <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>.