H-034

PERFOMANCE IMPROVEMENT OF PDLDA BASED FACE RECOGNITION USING FUSION FACE DESCRIPTOR

I Gede Pasek Suta Wijaya†,‡ Keiichi Uchimura† Gou Koutaki†

1. Introduction

Suppose we have large variability face images due to pose variations, as shown in Fig, 1. Our recent research, which did the face recognition using predictive linear discriminant analysis (PDLDA) and holistic features, shows good enough performance in both off-line and real-time test for the face pose variations labeled P_4, P_5, and P_6 of Fig 1. In off line test, the proposed method gave 99.52% recognition rate and in real-time test, it provided the recognition rate, false rejection rate, and false acceptance rate by about: >98%, 2%, and 4%, respectively with short processing time[11]. Those achievements were achieved with considering the chrominance components of face image, which were extracted using YCBCr color space transformation.

In addition, the PDLDA based face recognition which is an alternative approach of LDA, can solve the main problem of the LDA in term of retraining problem which always comes when new data enter to the system continuously (incrementally). However, it remains have several problems in term of the large variability of face due to the lighting condition and large variability due to the face pose variations. In other words, it does not work at all for face pose variations labeled P_1-P_3, and P_7-P_9 of data in Fig. 1.

Regarding to large variability of face images due to the lighting condition problem, several solution have been proposed such as Linear binary pattern (LBP), modified LBP, SQI, low frequency removal based illumination compensation. Regarding to large variability due to the face pose variations, several solutions have been proposed, such as using large training samples and holistic features based on frequency analysis. However, to getting large sample variations of face in single person is hard to be done and requires large memory space. The proposed method based on holistic features just works for non-large face pose variations (in the range of -20^0 and $+20^0$ face movement from the camera point). Therefore, we have to find other solution, which can solve this problem especially for multi-pose variations, as shown in Fig. 1.

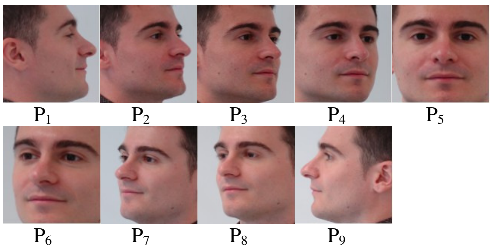

Figure 1. Large variability of face due to pose variations.

† Computer Science and Electrical Engineering GSST Kumamoto University, Kumamoto-Shi, JAPAN.
‡ Electrical Engineering Dept., Engineering Faculty, Mataram University, Mataram-NTB, INDONESIA.

In this paper, we presents an improvement of our previous research [4,5] to solve the large face variability due to pose variations. To achieve the improvements, we apply multistage classifiers, i.e. PDLDA as the first stage classifier and the fusion face descriptors as the second stage classifier. The function of the first stage classifiers is to determine several class candidates that are closely to the query image and the function of the second stage classifier is to determine the best similarity of query image to the class candidates. The PDLDA is chosen because it is simple and fast method that gave better performance than that of CPCA and CLDA. The fusion face descriptors which is an abstracts representation of 3D face images extracted using scale invariant features transforms (SIFT) is chosen because it can provide any face pose variations estimation using one descriptors.

2. PDLDA Based Face Recognition

The PDLDA based recognition was proposed to overcome main problems of the conventional LDA[2,3] in terms of large time processing for retraining and recalculating the global covariance or the between class scatter, S_b. The PDLDA works the same as the original LDA except on the defining constant global mean (μ_a) as a constant value for avoiding the S_b recalculation, as follows

$$S_b^p = \sum_{k=1}^{L} P(x_k)(\mu_k-\mu_p)(\mu_k-\mu_p)^T + P(x_{new})(\mu_{new}-\mu_p)(\mu_{new}-\mu_p)^T = S_b^{old}+S_b^{new} \qquad (1)$$

Then, within class scatter, S_w, also can be simplified as:

$$S_w = \frac{1}{N}\left\{\sum_{k=1}^{L-1} S_w^k + S_w^L\right\} = \frac{1}{N_{new}}\left\{S_w^{old}+S_w^{new}\right\} \qquad (2)$$

By substituting the S_b with the S_b^P of LDA eigen analysis then we get the optimum projection matrix called as PDLDA projection matrix (W_{PDLDA}) which has to satisfy the following criterion

$$J(W_{PDLDA}) = \arg\max_{W_{PDLDA}} \frac{\left|W_{PDLDA}^T S_b W_{PDLDA}\right|}{\left|W_{PDLDA}^T S_w W_{PDLDA}\right|} \qquad (3)$$

By using the W_{PDLDA}, the projected features of the both training and querying data set can be performed as done by the LDA.

For matching process, the Euclidean distance based on nearest neighbor rule is implemented for face classification.

3. The Proposed Methods

3.1 Fusion Face Descriptors

The fusion face descriptors is a fusion of selected pose variations features which represents the 3D image information. In this case, the features are extracted from 2D face images. Therefore, to realize this idea, we require a set of 2D face images which represent sub-sampling of 3D face images, as shown in Fig. 1.

From these images, we extract the dominant face information of each image called as sub-3D face features and then fuse them into descriptors. This descriptor is called as fusion face descriptor which is an abstracts representation of 3D face images. In this case, the SIFT from Ref. [6] is implemented for getting the sub-3D face features. The block diagram our proposed fusion face descriptor is presented in Fig. 2.

The proposed algorithm works as follows:

1. Suppose, we select P_2, P_5, and P_8 of the face images set of Fig. 1 for face input set.
2. ***Step 1***: from the selected images, sub-3D face features are extracted using SIFT algorithm, which is denoted by F_1, F_b, and F_2, respectively.
3. ***Step 2***: Removing the redundant features of the training faces are done step-by-step using intersection (∩) and subtraction operation as the following equation:
$$F_i^{'} = F_i - (F_i \cap F_b) \quad (4)$$
Where, $F_i^{'}$ is non-redundant features of i-th training image, F_i is the original features of i-th face image, and F_b is the features bases extracted from the frontal face image. The intersection operation is implemented to detect the location of the redundant features correspond to features bases. Then, the redundant features are removed by subtracting the original features with the detected redundant features.
4. ***Step 3***: fusing all of the non-redundant features ($F^{'}_i$, i=1, 2, 3, …, n-1, where n is number of training faces) with the features bases into a descriptor using the union (U) operation:
$$D_k = F_b \cup F_1^{'} \cup F_2^{'} \cup \cdots \cup F_{n-1}^{'} \quad (5)$$
Where, D_k is fusion face descriptor or a descriptor of multipose face images of k-th class.

In this case, the fusion face descriptor is a two dimensional data which is represented using a matrix. The benefit of this representation is simple and requires less memory space compare to the real 3D data. Furthermore, the more 2D-face images are included for building the descriptor the more rich face descriptors will be gotten. Consequently, if the more rich face descriptors are used for recognition, the higher recognition rate will be achieved.

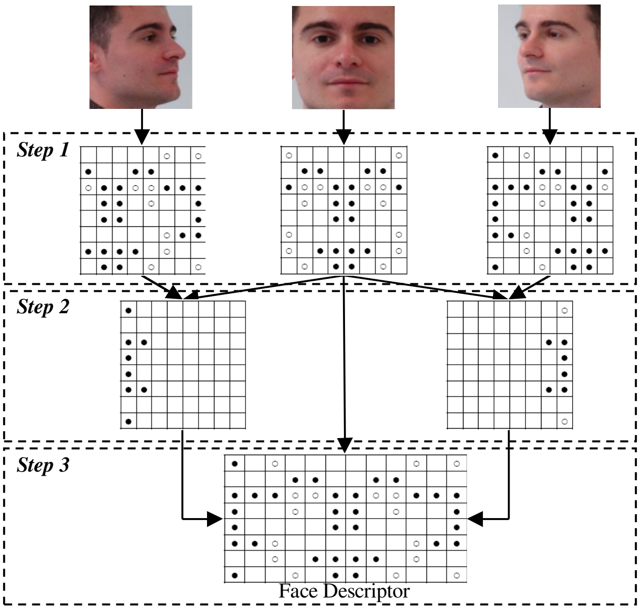

Figure 2. 3D face description block diagram.

3.2 Multi Stage Recognition

The proposed face recognition algorithm can be illustrated briefly in Fig. 3. There are two main processes: PDLDA processing as the first stage of face features classifier, and the fusion face descriptor (FFD) processing as the second stage of classifier which classify just first five the PDLDA-determined class candidates.

The PDLDA process performs the first classification in order to define class candidates for the next process. In this case, the first five smallest score of training set ID keep as class candidates. Next, from the first stage classifier, we got the five ID candidates, which is much similar to the query face images. The second stage classifiers try to find out the best ID correspond to the face the query.

The second stage verification is done using the following step:

1. Loading all of the face training set which correspond to the candidates ID.
2. Extracting the FFD of each face image from training set using the algorithm as presented in sub-section 3.1 and representing them as [D_1, …D_5] .
3. Extracting the FFD of query face image and representing it as representing D_q

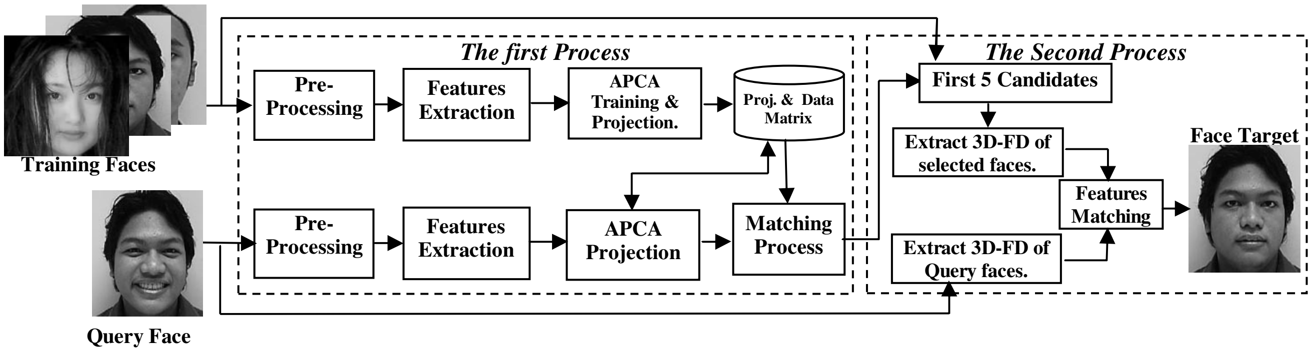

Figure 3. Face recognition diagram block.

1. Matching the D_q with $[D_1, \ldots D_5]$ to find out the number of matching key points. To find out the number matching key point of two descriptors (D_q and D_1 can be done using the following procedure.

```
func noOfMatch (D1 ∈ ℜ^{rxc1}, D2 ∈ ℜ^{rxc2}, thresh)
    nOfMatch=0;
    for i = 1 to c1 do
        Good=10000;
        Best=10000;
        idx=-1;
        for j = 1 to c2 do
            f=D1(:,i)-D2(:,j);
            d=f'*f;
            if (d<Good )
                Best=Good;
                Good=d;
                idx=j;
            elseif(d<Best)
                Best=d;
            endif
        endfor
        if ((thresh*Good <= Best)&& idx>0)
            nOfMatch=nOfMatch+1;
        endif
    endfor
    return (nOfMatch)
endfunc
```

2. Finally, verification criterion is defined based on number of matching key points. The face descriptor which has the largest number of matching points is concludes as the best likeness.

4. Experimental and Result

4.1 Experiment Setup

In order to know the performance of the proposed method, several experiments using data from some challenges databases: the ITS-Lab.[11], ORL[12], CVL[13], and GTV databases[14] were done in PC with specification: Core-Duo Processor 1.7 GHz and 2 GB RAM. The face pose variations examples from the mentioned databases are shown in Fig. 4. In addition, the detail explanation of each database can be found in the Refs. [11]-[14].

In addition, all experiments were done using the following condition:

1. The first there face images (P_1-P_3) were selected for training set and the remaining face images were selected for testing.
2. Both training and querying face images size was setup on 128 x 128 pixels.
3. The image stretching was employed before holistic features extraction and the holistic features size were setup 53 elements per images for the first stage classifier.

4.2 Results and Discussion

The first experiment was done to know the effect of the second stage classifier on the recognition rate improvement of the base-line method. The experimental result shows that the proposed method provides sufficiently improvement for ITS-Lab and ORL databases and significant improvement for CVL and GTAV databases than that of base-line methods (see Fig. 5). The significant improvement can be achieved because the CVL and GTAV contain larger face pose variations due to face pose variations than that of the ITS-Lab. and ORL., as shown in Fig. 4(c) and 4(d). It means the FFD-based face recognition can overcome the multi-pose difficulty because the FFD contains any estimation of poses variations face features.

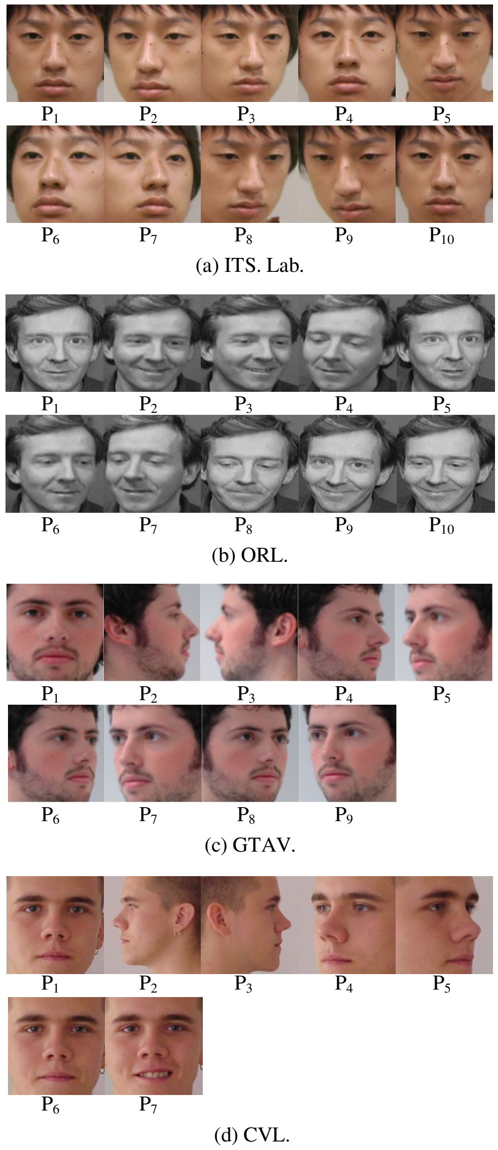

Figure 4. 3D face description block diagram.

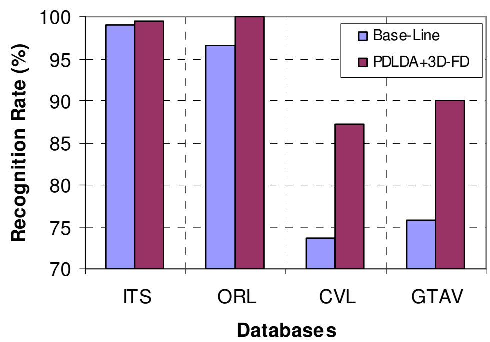

Figure 5. The effect of the second stage classifier on the recognition rate improvement.

The next experiment was done to know the performance of the proposed method on recognizing the 3D data. In this test, we compare our proposed method with the recent 3D methods which work based on combination multi-features (MF) and multi-feature fusion (MFF) of 3D face image with PCA and PCALDA (see Refs. [4,5]. In this test, we compare three best variants of those methods called as MF+PCA, MF+PCALDA and MFF+PCALDA. The experiment was done in the ITS-Lab face database version 1 which is 3D face database containing 40 classes which each classes consist of 10 face pose variations. The face images were acquired by Konica Minolta 3D-camera series VIVID 900. The testing parameters were setup the same as done in Refs. [4,5]: five images of each classes was chosen for training set and the remaining images. In addition, we also compared the performance of our proposed method with base-line method on the experiment using 3 face images as training set.

The experimental results show that our proposed method can achieve almost the same recognition rate as MFF+PCALDA methods (99.70% and 99.98% respectively) for testing using 3 face images for training. When the experiment using 5 face images for training, our proposed method recognize all of the testing data (see Table. 1). It means our proposed method is an alternative solution for building multi-pose face recognition with having reasonable achievements compared to 3D-based face recognition. Even though the recognition of FFD+LDA is not much different with the MFF+PCALDA but our proposed method does not require multi-features consisting of three features vectors and does not need 3D camera sensors for making the features at all. In other words, our proposed method is cheapest 2D-based face recognition approach which can be implemented for real time multi-pose face recognition with 2D web-camera as the image capturing.

In addition, the more images are considered (trained) for building the face descriptor, the higher recognition rate is achieved as shown in Table 1. The proposed method can recognize all testing images, which mean each training class, contain the richer face descriptor when the more face images are trained.

5. Conclusion and Future Works

The multi-stage face recognition approach based on fusion face descriptor can be used to improve the PDLDA-based recognition especially to improve the large face variability due to pose variations. The proposed method gives sufficient improvement in term of recognition rate against face pose variations, and provide better recognition rate over recent 3D-face recognition methods. It means, our proposed method is potentially to be implemented for 2D-multi pose face recognition approach which can be implemented for real time face recognition with 2D web-camera as face image sensor.

This research will be continued to be implemented for realtime face recognition and develops for security system. In addition, we will try to visualize the 3D-based face descriptor into 3D face descriptor and to modify the FSIF-based face extraction for decreasing time complexity requirements.

Table 1. The recognition rate of the proposed methods vs. that of Multi-features and PCALDA based 3D face recognition[4,5].

No	Methods	Recognition Rate as function Number of Training	
		3	5
1	MF+PCA	NA	94.08
2	MF+PCALDA	NA	99.34
3	MFF+PCALDA	NA	99.98
4	PDLDA	97.69	98.50
5	PDLDA+FFD	**99.70**	**100.00**

Acknowledgements

I would like to send my appreciation and my great thank to Innovation Project of Kumamoto University, which give many supports to this research, and my great thank to owner of ORL, CVL, GTAV databases.

References

[1]. W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, Face Recognition: A Literature Survey, ACM Computing Surveys, vol.35, no.4, pp.399-458, December 2003.
[2]. K. I. Chang, K. W. Bowyre, and P. J. Flyn, An Evaluation of Multi-modal 2D+3D Face Biometrics, IEEE Transactions on PAMI, vol.27, pp.619-624, April 2005.
[3]. K. I. Chang, K. W. Bowyre, and P. J. Flyn, Face recognition using 2D and 3D facial data, Multimodal User Authentication Workshop, pp. 25-32, 2003.
[4]. Z. Cuicui, IGPS. Wijaya, G. Koutaki, and K. Uchimura, 3D Face Recognition Based on Multi-Features of Face Image, Proceedings of IEE-Japan conference, Kumamoto Japan, October 2010, (CDROM).
[5]. Z. Cuicui, K. Uchimura, C. Zhang, G. Koutaki, 3D Face Recognition Using Multi-level Multi-feature Fusion, Proceedings of the 4th Pacific-Rim Symposium on Image and Video Technology (PSIVT 2010), Singapore, pp.21-26, November 2010.
[6]. D. G. Lowe, Distinctive Image Features from Scale-Invariant Keypoints, International Journal of Computer Vision, vol.60, no.2, pp. 91-110, 2004.
[7]. D. G. Lowe, Object Recognition from Local Scale-Invariant Features, Proceedings of the International Conference on Computer Vision, Corfu, September 1999.
[8]. M. Turk and A. Pentland, Eigenfaces for Recognition, Journal of Cognitive Neuroscience, vol.3, no.1, pp.71-86, 1991.
[9]. Yu_DLDA} H. Yu and J. Yang, A Direct LDA algorithm for High-DimensionalData-with Application to Face Recognition, Pattern
[10]. W. Chen, J-E. Meng, and S. Wu, PCA and LDA in DCT Domain, Pattern Recognition Letter, no.26, pp.2474-2482, 2005.
[11]. IGPS. Wijaya, K. Uchimura, and Z. Hu, Improving the PDLDA Based Face Recognition Using Lighting Compensation, Proceedings of Workshop of Image Electronics and Visual Computing 2010, Nice France, March 2010, (CDROM).
[12]. F. Samaria and A. Harter, Parameterization of a stochastic model for human face identification, in The 2nd IEEE Workshop on Applications of Computer Vision, Sarasota Florida, pp. 138-142 (1994).
[13]. http://lrv.fri.uni-lj.si/facedb.html
[14]. http://gps-tsc.upc.es/GTAV/ResearchAreas/ UPCFaceData-base/ GTAVFaceDatabase.htm