TELKOMNIKA, Vol.15, No.4, December 2017, pp. 1794~1807
ISSN: 1693-6930, accredited **A** by DIKTI, Decree No: 58/DIKTI/Kep/2013
DOI:10.12928/TELKOMNIKA.v15i4.6576

Contradictory of the Laplacian Smoothing Transform and Linear Discriminant Analysis Modeling to Extract the Face Image Features

Arif Muntasa*, Indah Agustien Siradjuddin
Informatics Engineering Department, Enginering Faculty, University of Trunojoyo, Madura
Ry Telang Po. Box 2 Kamal, Bangkalan, Indonesia
*Corresponding author, e-mail: arifmuntasa@trunojoyo.ac.id

Abstract

Laplacian smoothing transform uses the negative diagonal element to generate the new space. The negative diagonal elements will deliver the negative new spaces. The negative new spaces will cause decreasing of the dominant characteristics. Laplacian smoothing transform usually singular matrix, such that the matrix cannot be solved to obtain the ordered-eigenvalues and corresponding eigenvectors. In this research, we propose a modeling to generate the positive diagonal elements to obtain the positive new spaces. The secondly, we propose approach to overcome singularity matrix to found eigenvalues and eigenvectors. Firstly, the method is started to calculate contradictory of the laplacian smoothing matrix. Secondly, we calculate the new space modeling on the contradictory of the laplacian smoothing. Moreover, we calculate eigenvectors of the discriminant analysis. Fourth, we calculate the new space modeling on the discriminant analysis, select and merge features. The proposed method has been tested by using four databases, i.e. ORL, YALE, UoB, and local database (CAI-UTM). Overall, the results indicate that the proposed method can overcome two problems and deliver higher accuracy than similar methods.

Keywords: *Feature extraction, face image, contradictory, and laplacian transform*

1. Introduction

Appearance-based method to extract the image features is still conducted by many researchers such as Principal Component Analysis (PCA). PCA is the oldest method to extract the image features using statistical methods [1]–[5]. The method also has been developed into many methods, i.e. multi linear PCA [6], Non-Linear Principal Component Analysis, Improved principal component regression [7], and kernel subspace [8].

The problem of the image computational is high dimensionality, for both video and still images. Many researchers have tried to obtain the best solution related to dimensionality reduction. However, many methods have been improved, but they just leave another problem. The oldest appearance method to extract the features is Principal Component Analysis (PCA). PCA projects new space into eigenvector based on decreasing ordered-eigenvalues, where the results can reduce meaningfully the image samples as the training sets, where the dimensionality of the image as the training has same size or smaller compared to the number of training samples. In this case, PCA only performs efficiently when the number of training sets is fewer than image dimensionality. Unfortunately, PCA will fail to project the training sets when the number of training sets is more than the image dimensionality. In addition, PCA also represents global manifolds of the object as the training samples, whereas local manifolds cannot be well represented by PCA.

LDA can overcome the limitation of PCA, where LDA is able to extract features up to a number of classes. Another result of a researcher is Linear Discriminant Analysis (LDA) [8]–[11]. LDA maximized between class scatter and minimized within class scatter. LDA projects eigenvectors based on the ordered-eigenvalues of between and within class scatter. LDA is better than PCA, because LDA can reduce more significant than PCA, i.e. maximum features for LDA is number of classes, while maximum features for PCA is number of training samples, where training samples are greater than number of classes [11], [12]. In addition, LDA has also maximized the distance between classes to each another, and minimized the distance of

Received Mar 5, 2017; Revised September 18, 2017; Accepted September 30, 2017

member on the same class, therefore the member of class will be easly separated. However LDA also leaves a new problem, the inability to reduce the dimensions of the image, when the number of classes exceeds the dimensions of the image.

Another approach is also conducted by using laplacian smoothing, where the Kronecker matrix operation is employed and selected to earn the optimum values. As we know that laplacian smoothing uses Neuman discrete and Kronecker matrix to obtain the new space, where the value used is -1 the first or the last diagonal and -2 for other diagonal elements. The negative values will generate negative new spaces, and the negative new spaces can remove the dominant characteristics. The Kronecker matrix result usually cannot be solved to obtain the eigenvalues and eigenvectors because of singular matrix, therefore it is necessary to appear the features by replacement the negative into positive emenets.

In this research, we improve two limitations of the laplacian smoothing. The first, we propose contradictory of the laplacian smoothing transform to generate the positive diagonal matrix. The second we propose to overcome singularity matrix on the laplacian smoothing transform. We also propose to combine between the contradictory of the laplacian smoothing and the linear discriminant to earn the best features before classification applied.

The rest of the paper is composed as follows, proposed method, experimental material and scheme, results and discussion, compare to similar methods, and conclusions. On the proposed method, we explain the framework of the proposed approaches, i.e. calculate contradictory of the laplacian smoothing matrix, calculate the new space modeling on the contradictory of the laplacian smoothing, calculate eigenvectors of the discriminant analysis, calculate the new space modeling on the discriminant analysis, select and merge features.

2. Research Method

The Laplacian transform modeling has been implemented to extract the image features. However, this method can be also improved by Neuman Discrete, where the original of the method uses the negative values, which are -1 and -2. The results showed that the resulting features can eliminate the dominant characteristics. Therefore, we proposed to reverse the values (as the first contradictory) to appear the dominant characteristics. Another weakness of the laplacian smoothing is singularit, where it usually appears when eigenvalues and eigenvectors are calculated. We proposed the model to overcome the singularity (as the second contradictory). The proposed approach can be displayed in Figure 1.

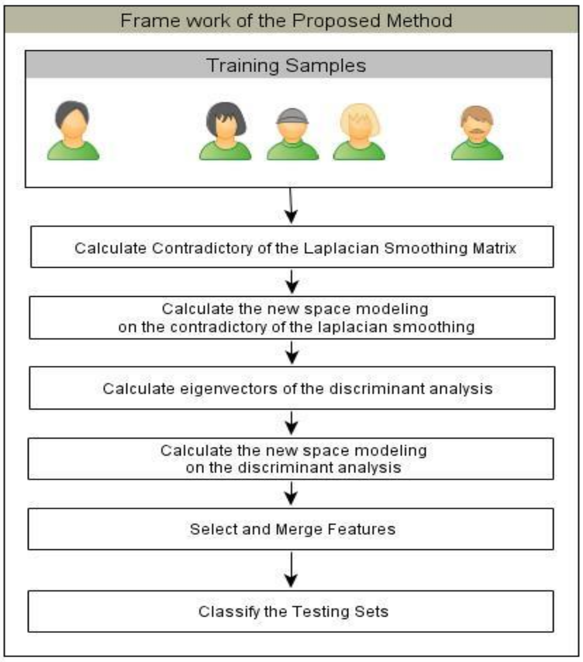

Figure 1. Framework of the proposed Method

Suppose, training image sample is symbolized by $F_{w,\,h}$, while testing image sample is represented by $G_{w,h}$. The image samples have size with w and h as width and height. To process an image, the size of image must be resized into one dimensional model, an original image size will be changed into row matrix model, which is matrix with width $w \times h$ and height 1. The row matrix model of an image can be shown as follows

$$F_{1,n} = (f_{1,1} \quad f_{1,2} \quad \cdots \quad f_{1,h\times w-1} \quad f_{1,h\times w}) \tag{1}$$

If number of training samples is m, then the matrix training sample can be composed into matrix with width $w \times h$ and height m as follows

$$F_{m,h\times w} = \begin{pmatrix} f_{1,1} & f_{1,2} & \cdots & f_{1,h\times w-1} & f_{1,h\times w} \\ f_{2,1} & f_{2,2} & \cdots & f_{2,h\times w-1} & f_{2,h\times w} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ f_{m-1,1} & f_{m-1,2} & \cdots & f_{m-1,h\times w-1} & f_{m-1,h\times w} \\ f_{m,1} & f_{m,2} & \cdots & f_{m,h\times w-1} & f_{m,h\times w} \end{pmatrix} \tag{2}$$

The testing sample must be also represented by one dimensional matrix, which is row matrix model in Equation (3). Equation (2) and (3) will be applied to earn the features of the training and the testing samples.

$$G_{1,w\times h} = (g_{1,1} \quad g_{1,2} \quad \cdots \quad g_{1,w\times h-1} \quad g_{1,h\times w}) \tag{3}$$

2.1. Calculate Contradictory of the Laplacian Smoothing Matrix

Contradictory of the Laplacian Smoothing Matrix is a new approach to produce the new projection based on the opposite of the Laplacian Smoothing Transform. We have developed the new model based on the Laplacian smoothing matrix. At first, we compose the matrix with zero values except on last position, where an index of j applies to $\forall j, j \in 1, 2, , \cdots, h \times w$ and index of k instructs to $\forall k, k \in 1, 2, , \cdots, h \times w$ as shown as follows

$$L_1(j,k) = \begin{cases} 1, & j = h \times w \text{ and } k = h \times w \\ 0, & Otherwise \end{cases} \tag{4}$$

Furthermore, the value 1 will be inserted on the above matrix using Equation (5) and (6), where the index of j is $\forall j, j \in 1, 2, 3, \cdots, w-1$

$$L_1(h \times j + 1, h \times j) \leftarrow 1 \tag{5}$$

$$L_1(h \times j, h \times j + 1) \leftarrow 1 \tag{6}$$

The Equation (4), (5) and (6) produces the sparse matrix, where only the several positions have element with value 1. The results of Equation (4), (5), and (6) are added for each column as shown in the following equation

$$L_2(1,k) = \sum_{j=1}^{h\times w} L_1(j,k) \tag{7}$$

Equation (7) produces row matrix model, where it has $h \times w$ columns and a row. The result of Equation (7) is changed into diagonal matrix. The matrix diagonalization result is subtracted by Equation (5) as shown as follows

$$L_3(j,k) = \mathbb{D}\big(L_2(1,k)\big) - L_1(j,k) \tag{8}$$

In this case $\mathbb{D}$ denotes the formation of the diagonal matrix, where all matrix elements are 0 values except the diagonal elements is filled by $L_2(1,k)$. The results of Equation (8) can be also represented as follows

$$L_3(j,k) = L_2(j,k) - L_1(j,k) \tag{9}$$

Where the index of j and k are applicable for $\forall j, j \in 1,2,,\cdots,h \times w$ and $\forall k, k \in 1,2,,\cdots,h \times w$. Moreover, we create the first contradictory, all matrix elements are filled by 2 except for $(1,1)$ and also $(\lceil h \times w/2 \rceil, \lceil h \times w/2 \rceil)$ as show follows

$$L_4(j,j) = \begin{cases} 1, \; if \; j = 1 \; or \; j = \left\lceil \frac{h \times w}{2} \right\rceil \\ 2, \; otherwise \end{cases} \tag{10}$$

In this case, $\forall j, j \in 1,2,,\cdots, \left\lceil \frac{h \times w}{2} \right\rceil$, the results of Equation (9) is applied to compute the Newman Discrete matrix as follows

$$D_{0.5 \times h, \; 0.5 \times w} = \rho \times h \times w \times \left(L_4(j,k)\right) \tag{11}$$

Output of Equation (11) is employed to calculate $\Delta_{h,w}$ through Kronecker operation as shown in the following equation

$$\Delta_{h,w} = D_{0.5 \times h, \; 0.5 \times w} \otimes I_{2,2} + I_{2,2} \otimes D_{0.5 \times h, \; 0.5 \times w} \tag{12}$$

The second contradictory of the Laplacian smoothing matrix can be used to avoid singularity matrix through as the follows

$$L_5 = \left(L_4 + \rho \times \Delta_{h, \; w} + (L_4)^T\right) \tag{13}$$

The result of Equation (13) is applied to compute the Eigenvalues and eigenvectors by using characteristic equation

$$A \times L_5 = \lambda \times L_5 \tag{14}$$

The Eigenvalues generated by Equation (14) must be ordered decreasingly and followed by a change of column index of the eigenvectors as follows

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_m \tag{15}$$

$$\Lambda_{h \times w,1}, \Lambda_{h \times w,2}, \Lambda_{h \times w,3}, \cdots, \Lambda_{h \times w,m} \tag{16}$$

2.2. Calculate the New Space Modeling on the Contradictory of the Laplacian Smoothing

The results of equation (16) are required to earn the new space as object features, for both on the training and testing sets. The new space of the training sets can be calculate by multiplication of the training sets $F_{m,h \times w}$ and eigenvectors of the contradictory of the laplacian smoothing $\Lambda_{h \times w,m}$using equation

$$\mathcal{X}_{m,m} = F_{m,h \times w} \times \Lambda_{h \times w,m} \tag{17}$$

While the new space of the testing set can be easily obtained using multiplication between the testing sets $G_{1,h \times w}$ and eigenvectors of the contradictory of the laplacian smoothing $\Lambda_{h \times w,m}$using equation

$$\mathcal{Y}_{1,m} = G_{1,h \times w} \times \Lambda_{h \times w,m} \tag{18}$$

2.3. Calculate Eigenvectors of the Discriminant Analysis

The discriminant analysis is produced by rationalization of the fisher's discriminant analysis [13]–[15]. The result of the discriminant analysis can be used as linear classifier or well known as dimensionality reduction. It can be earned to maximize between scatter and to minimize within scatter. Between and within scatter can be defined as follows

$$S_B = \sum_{j=1}^{c} N_j \times \left(\mu_j - \mu\right) \times \left(\mu_j - \mu\right)^T \tag{19}$$

$$S_W = \sum_{j=1}^{c} \frac{1}{N_j - 1} \times \sum_{F \in \omega_j} (F - \mu_j) \times (F - \mu_j)^T \quad (20)$$

In this case N_j describes the number of training samples on the *j* class, μ_j is the average of training sample on the *j* class, where $\forall j, j \in 1, 2, 3, \cdots, m$ and *m* states the number of class. The average for all training samples is depicted by μ, while $F \in \omega_j$ is training samples on the *j* class. The Eigenvalues and eigenvectors of the discriminant analysis can be easily obtain by using equation

$$\left(\frac{(\vec{w}^T \times \sum_B \times \vec{w})}{(\vec{w}^T \times \sum_W \times \vec{w})}\right) \times A = \vec{\lambda} \times A \quad (21)$$

The calculation results of the Equation (21) deliver non-ordered eigenvalues and eigenvectors. In order to the best characteristics, the ordered-eigen values must ob ordered decreasingly and supported by column index of the eigenvector [1], [6], [16]. The results are follows

$$\vec{\lambda}_1 \geq \vec{\lambda}_2 \geq \vec{\lambda}_3 \geq \cdots \geq \vec{\lambda}_m \quad (22)$$

$$\vec{\Lambda}_{h\times w,1}, \vec{\Lambda}_{h\times w,2}, \vec{\Lambda}_{h\times w,3}, \cdots, \vec{\Lambda}_{h\times w,m} \quad (23)$$

2.4. Calculate the New Space Modeling on the Discriminant Analysis

The new space as features can be earned by multiplication between the training sample $F_{m,h\times w}$ and the eigenvector $\vec{\Lambda}_{h\times w,m}$ (Equation (23)). In this case, the matrix results have a size of *m*, where *m* is the number of training sets

$$\vec{\mathcal{X}}_{m,m} = F_{m,h\times w} \times \vec{\Lambda}_{h\times w,m} \quad (24)$$

While the new space of the testing set can be easily obtained by multiplication of the testing sets $G_{1,h\times w}$ and eigenvectors $\vec{\Lambda}_{h\times w,m}$ of the discriminant analysis as follows

$$\vec{\mathcal{Y}}_{1,m} = G_{1,h\times w} \times \vec{\Lambda}_{h\times w,m} \quad (25)$$

2.5. Select and Merge Features

Before classification process, the features delivered by contradictory of the laplacian transform and discriminant analysis must be selected first. Based on the Equation (17), we can re-write it into the matrix form as follows

$$\mathcal{X}_{m,m} = \begin{pmatrix} \mathcal{X}_{1,1} & \mathcal{X}_{1,2} & \cdots & \mathcal{X}_{1,m-1} & \mathcal{X}_{1,m} \\ \mathcal{X}_{2,1} & \mathcal{X}_{2,2} & \cdots & \mathcal{X}_{2,m-1} & \mathcal{X}_{2,m} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathcal{X}_{m-1,1} & \mathcal{X}_{m-1,2} & \cdots & \mathcal{X}_{m-1,m-1} & \mathcal{X}_{m-1,m} \\ \mathcal{X}_{m,1} & \mathcal{X}_{m,2} & \cdots & \mathcal{X}_{m,m-1} & \mathcal{X}_{m,m} \end{pmatrix} \quad (26)$$

At first, the features of the contradictory of the laplacian transform are selected *u* features, where *u*=*m*-*q*, and $\exists q, q \in 1, 2, 3, \cdots, m-1$. It means that, we remove *q* features. The feature selection results can be written as follows

$$\mathcal{X}_{m,m} = \begin{pmatrix} \mathcal{X}_{1,1} & \cdots & \mathcal{X}_{1,u} \\ \mathcal{X}_{2,1} & \cdots & \mathcal{X}_{2,u} \\ \vdots & \ddots & \vdots \\ \mathcal{X}_{m-1,1} & \cdots & \mathcal{X}_{m-1,u} \\ \mathcal{X}_{m,1} & \cdots & \mathcal{X}_{m,u} \end{pmatrix} \quad (27)$$

The feature extraction of the discriminant analysis as shown in Equation (24) can be also expressed on the matrix form with *m* columns and *n* rows as shown follows:

$$\vec{x}_{m,m} = \begin{pmatrix} \vec{x}_{1,1} & \vec{x}_{1,2} & \cdots & \vec{x}_{1,m-1} & \vec{x}_{1,m} \\ \vec{x}_{2,1} & \vec{x}_{2,2} & \cdots & \vec{x}_{2,m-1} & \vec{x}_{2,m} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vec{x}_{m-1,1} & \vec{x}_{m-1,2} & \cdots & \vec{x}_{m-1,m-1} & \vec{x}_{m-1,m} \\ \vec{x}_{m,1} & \vec{x}_{m,2} & \cdots & \vec{x}_{m,m-1} & \vec{x}_{m,m} \end{pmatrix} \quad (28)$$

Moreover, the features generated by discriminant analysis are also selected *v* features, and *v*=*m*-*s*, where $\exists s, s \in 1,2,3,\cdots,m-1$. In this case, *s* features are removed as shown follows

$$\vec{x}_{m,m} = \begin{pmatrix} \vec{x}_{1,1} & \cdots & \vec{x}_{1,v} \\ \vec{x}_{2,1} & \cdots & \vec{x}_{2,v} \\ \vdots & \ddots & \vdots \\ \vec{x}_{m-1,1} & \cdots & \vec{x}_{m-1,v} \\ \vec{x}_{m,1} & \cdots & \vec{x}_{m,v} \end{pmatrix} \quad (29)$$

The results of the feature selection are merged into a matrix, where the features of the discriminant analysis are put on the right of the contradictory of the laplacian transform features as shown follows:

$$\mathbb{X} = \left(\begin{array}{ccc|ccc} x_{1,1} & \cdots & x_{1,u} & \vec{x}_{1,1} & \cdots & \vec{x}_{1,v} \\ x_{2,1} & \cdots & x_{2,u} & \vec{x}_{2,1} & \cdots & \vec{x}_{2,v} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ x_{m-1,1} & \cdots & x_{m-1,u} & \vec{x}_{m-1,1} & \cdots & \vec{x}_{m-1,v} \\ x_{m,1} & \cdots & x_{m,u} & \vec{x}_{m,1} & \cdots & \vec{x}_{m,v} \end{array}\right) \quad (30)$$

After selection and merging of the features on the training sets, the similar process is also employed on the testing sets. The Equation (18) as the features of the testing sets of the contradictory of the laplacian transform and discriminant analysis are formulated follows:

$$y_{1,m} = (y_{1,1} \quad y_{1,1} \quad \cdots \quad y_{1,m-1} \quad y_{1,m}) \quad (31)$$

$$\vec{y}_{1,m} = (\vec{y}_{1,1} \quad \vec{y}_{1,1} \quad \cdots \quad \vec{y}_{1,m-1} \quad \vec{y}_{1,m}) \quad (32)$$

Furthermore, the number of feature selection on the testing sets must take the same size with training sets. In this case, we have selected *u* features for the contradictory of the laplacian transform and *v* features for the discriminant analysis. The selection and merging results are:

$$\mathbb{Y} = (y_{1,1} \quad \cdots \quad y_{1,u} | \vec{y}_{1,1} \quad \cdots \quad \vec{y}_{1,v}) \quad (33)$$

2.6. Classify the Testing Sets

The feature selection and merging results are compared between the Equation (33) and Equation (30). In this research, we employ the Euclidian distance to obtain the nearest neighbor as follows

$$D = \|\mathbb{X} - \mathbb{Y}\| \quad (34)$$

3. Experimental Material and Scheme

In this research, four different databases have been utilized to test the robustness of our proposed method, which are ORL [13], YALE [14], UoB [15], and CAI-UTM. The ORL has four hundred images, where ten people were taken with ten different accessories, poses, and expressions. In this case, an image of the ORL has 112 and 92 pixels for height and width [13].

The YALE only involves fifteen persons, and for each person was captured from eleven different models, they are lighting, view angle, expressions, and also pose. The size of the image on the YALE has been changed into 136 pixels for height and 104 pixels for width [14]. The third database is UoB face image, where thirty persons have been involved to be created the image samples with ten poses, where the size of image is 140 pixels for height and 120 pixels for width. This size is the result of resizing from the original size [15]. The last database is CAI-UTM, University of Trunojoyo Madura. We have taken a thousand images from a hundred persons, where for each person has support ten different models, which are expression, accessories, angle of view, and pose. Four databases are shown in Figure 2. The ORL, YALE, and UoB samples are shown on the first, the second, and the third row respectively, whereas on the fourth row is sample of the CAI-UTM database.

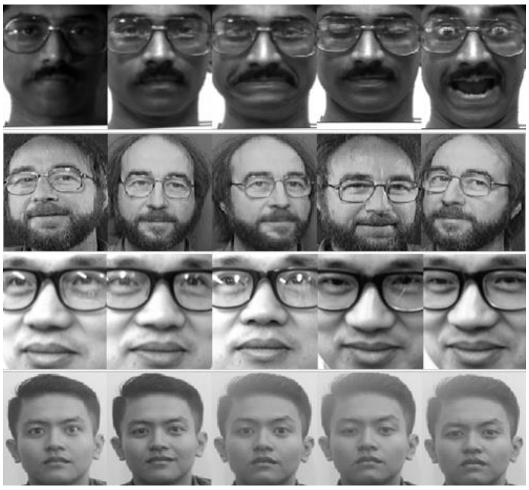

Figure 2. Samples of Face Image, ORL [13], YALE [14], UoB [15], and CAI-UTM

To proove the robustness of the proposed method, we employed four schemes for each database as shown in Table 1. We use *P*(*c*, *t*, *e*) as scheme models for the testing, where *c* represents *c*-fold cross validation. Variable of *t* depicts number of training sets used, while *e* symbolizes number of testing sets employed. For example *P*(4, 3, 7) means that cross validation is conducted four times with four different indexes, number of training sets employed are three images, and number of testing sets performed are seven images.

Table 1. Testing Scheme Used in This Research

Scheme	Training Set Used Database			
	ORL	YALE	UoB	CAI-UTM
1^{st}	P(4, 2, 8)	P(4, 2, 9)	P(4, 2, 8)	P(4, 2, 8)
2^{nd}	P(4, 3, 7)	P(4, 3, 8)	P(4, 3, 7)	P(4, 3, 7)
3^{rd}	P(4, 4, 6)	P(4, 4, 7)	P(4, 4, 6)	P(4, 4, 6)
4^{th}	P(4, 5, 5)	P(4, 5, 6)	P(4, 5, 5)	P(4, 5, 5)

4. Results and Discussions

We have randomized the training set index for each scheme based on Table 1. For all schemes, cross validation utilized is 4-folds, whereas number of features selected and number of experiments conducted are depend on number of images on the database image applied and k-fold cross validation implemented as shown in Table 2.

Table 2. Experiments Conducted for each Scheme

Database Image	Features Used	Number of Experiments for 1-fold	Number of Experiments for 4-fold	Number of Experiments for All Schemes
ORL	1-20, 1-21, . . ., 1-39	20	20 × 4=80	80× 4=320
YALE	1-6, 1-7, . . . , 1-14	9	9 × 4 = 36	36× 4=144
UoB	1-21, 1-22, . . ., 1-29	9	9 × 4=36	36× 4=144
CAI-UTM	1-70, 1-71, . . ., 1-99	30	30 × 4=120	120× 4=480

The results will be discussed sequentially, which are the results of the ORL, the YALE, the UoB, and the CAI-UTM face image database.

4.1. The Results of the ORL Database

We have carried using four schemes on the ORL database, which are experimental using two, three, four, and five training samples. For each scheme, we conducted different eighty times. The results are demonstrated in Figure 3. It indicates that the different results are delivered from all schemes, where the worst accuracy occurred on the fourth scheme (using two images as training sets), while the best accuracy is obtained by using four images as training sample.

From the experimental results it can be seen that, the difference in outcomes between the first, second, third, and fourth schemes. In the first scheme twenty experiments have been performed and the results of each experiment are not much different, therefore the accuracy deviation on the first scheme is also very small, that is 0.009.

Experimental results in the second scheme show that all experiments are better than the first scheme. The addition of an image as the training sample has provided an increase in accuracy across all experiments. It can be proved by the best, the average, and the worst accuracies also increase for all experimental, which are 87.86%, 85.71%, and 86.70% for the best, the average, and the worst accuracy respectively. The standard deviation obtained is also better than previous scheme that is 0.006. This shows that the result difference of the experiments in the first scheme is smaller than the first scheme. In the third and fourth schemes also occurred the same phenomena as in the first and second scheme. The results of all schemes show that the best results have been achieved by the last scheme, i.e. 97.50% for maximum accuracy, 94.50% for minimum accuracy and 96.18% for average accuracy, while the best standard deviation appeared on the second scheme as demonstrated in Table 3. Errors in facial image recognition occur because the training data selected does not represent all models of the tested image, this is evidenced by the addition of training images with different models can improve the accuracy of face image recognition as shown in Figure 3

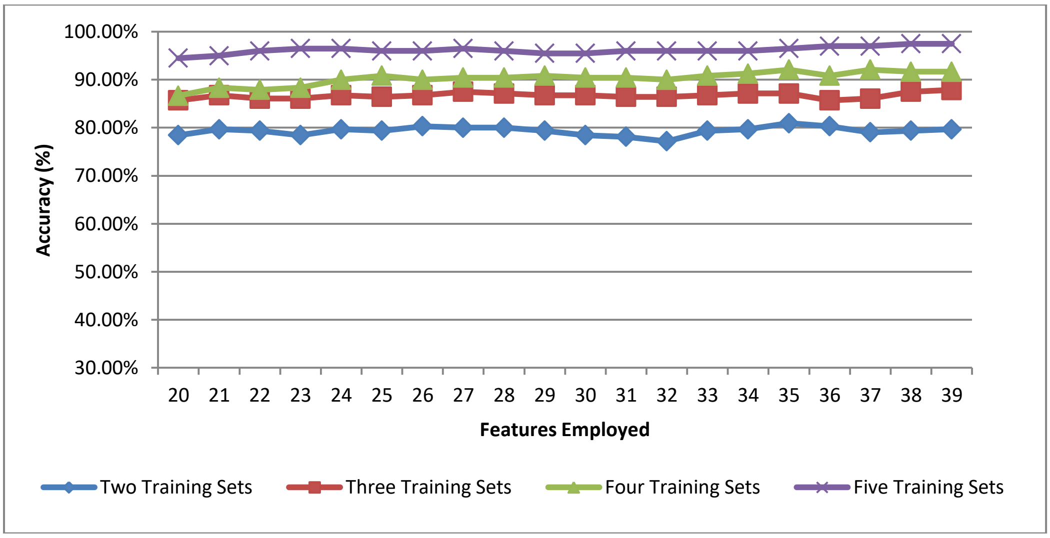

Figure 3. Accuracy on the ORL

Table 3. Experiments Conducted for each Scheme on the ORL

	Training Sets Employed				
	Two	Three	Four	Five	The best
Max	80.94%	87.86%	92.08%	97.50%	97.50%
Min	77.19%	85.71%	86.67%	94.50%	94.50%
Average	79.34%	86.70%	90.25%	96.18%	96.18%
Stdev	0.009	0.006	0.014	0.007	0.007

4.2. The Results of the YALE Database

The images on the YALE database are influenced by lighting, therefore the face images are harder recognized than face images on the ORL database, though the YALE database has fewer images than the images on the ORL database. Because the YALE database has fewer images than the ORL database, therefore we only conducted 144 experimental to evaluate the robustness our proposed method. We take six until fourteen features to characterize an object. In this case, the YALE also is evaluated by using four schemes, and for each scheme is conducted thirty six experiments, i.e. nine experiments is repeated by four-folds cross validation. The performance results of our proposed method on the YALE database can be depicted in Figure 4. The anomaly results are delivered on the second scheme as shown in Figure 4, the results decreased accuracy when number of features is more. It appeared when features implemented are ten until fourteen features, whereas for other schemes have produce higher accuracy for more features. The best accuracy is delivered by using five images as the training samples, i.e. 97.78%, 85.56%, and 92.84% for the maximum, minimum, and average accuracies respectively as shown in Table 4. Similar phenomenon also occurs in the YALE database, where lower accuracy has been delivered by the experimental results with smaller samples for the training process. This can be proven in Figure 4, where the average accuracy for 5 training data is higher than the others. In general, the used features also affected to the recognition results. Accuracy has a tendency to rise along with the number of features used, although there are some anomalies in some points.

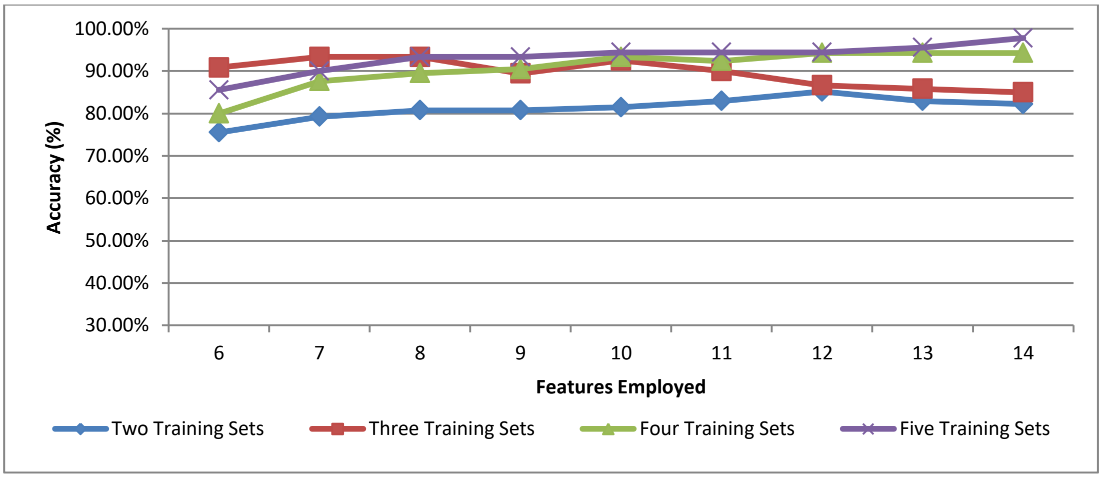

Figure 4. Accuracy on the YALE

Table 4. Experiments Conducted for each Scheme on the YALE

	Training Sets Employed				The Best
	Two	Three	Four	Five	
Max	85.19%	93.33%	94.29%	97.78%	97.78%
Min	79.26%	85.00%	80.00%	85.56%	85.56%
Average	81.98%	89.66%	90.16%	92.84%	92.84%
Stand Dev	0.027	0.032	0.047	0.035	0.035

4.3. The Results of the UoB Database

The ORL database is rare utilized to analyze the image, but in this research we apply this database to evaluate the robustness our proposed method. Experiments conducted are same with the YALE database, i.e. 144 experiments. Our proposed method delivered clear results, where the difference of the 1st, 2nd, 3rd, and 4th schemes are very clear, i.e. the use of two samples as training is not better than three, four and five samples as training sets. Similar to other schemes, three samples as training has delivered worse results than four and five samples as training, where the results of four and five samples as training are similar, but it is better to use five samples instead of four samples as training as demonstrated in Figure 5. For each scheme, the difference of experimental results is not significant. It indicates that the use of features as parameter in similarity measurements does not bring effect with accuracy results. It can be also shown from standard deviation delivered is very small, especially on the fourth scheme, i.e. 0.005. The best accuracy of all schemes is found on the last scheme, i.e. 94% for the maximum, 93.00% for the minimum, and 92.25% for the average accuracy. It indicates that the use of samples as the training has played the essential role to produce the accuracy of the recognition. Table 5 clarifies clearly the results in the maximum, the minimum, and the average.

In this database, it is clear that the difference of facial image recognition results in each of the planned schemes. The addition of images to the training process has contributed to the accuracy of the recognition process, although the addition of features does not impact the increase in recognition accuracy as shown in Figure 5. This proves that the features used have the same level, where the associated eigenvalue has almost the same value

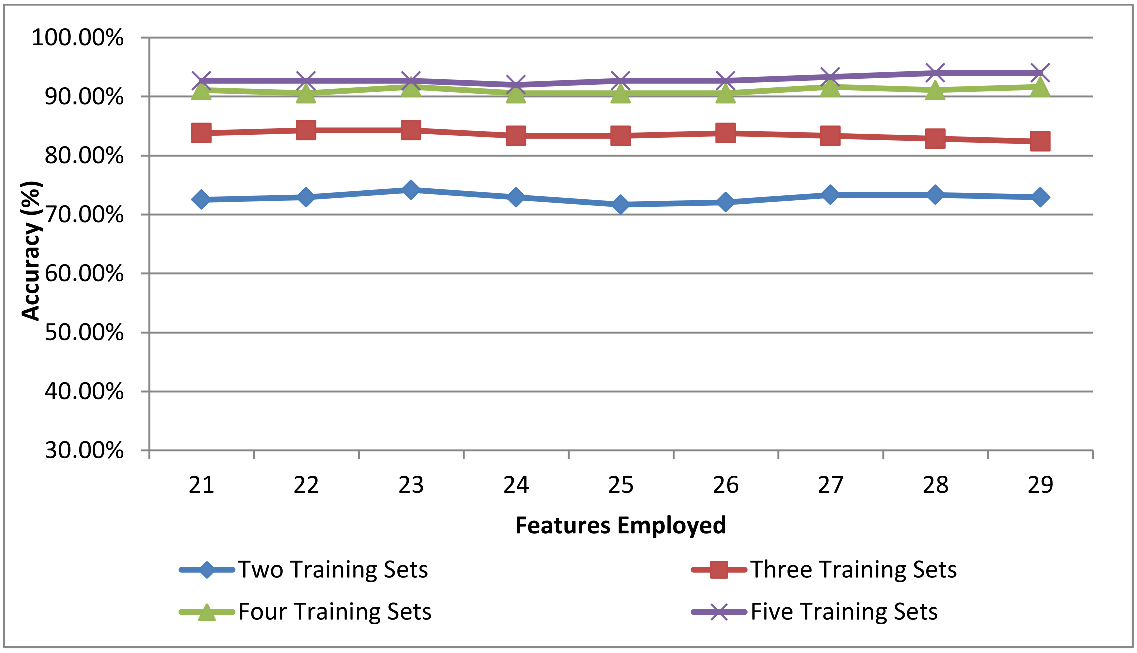

Figure 5. Accuracy on the UoB

Table 5. Experiments Conducted for each Scheme on the UoB

	Training Sets Employed				The Best
	Two	Three	Four	Five	
Max	74.17%	84.29%	91.67%	94.00%	94.00%
Min	71.67%	82.38%	90.56%	92.00%	92.00%
Average	72.87%	83.49%	91.05%	92.96%	92.96%
Stdev	0.007	0.006	0.005	0.007	0.007

4.4. The Results of the CAI-UTM Database

The last schema of experiments is to evaluate our proposed method using the CAI-UTM database. It is the local database taken from our university, i.e. University of Trunojoyo Madura. Figure 6 presented in detail the results, where the first scheme has delivered anomaly results, i.e. the less features measured, the higher resulting accuracy, though the difference not significant. The same phenomenon is also found in the second scheme, the result also decreases when the measured feature is added, although the decreasing of accuracy is extremely small. The second scheme, the increasing of the features measured does not show significant increasing in accuracy, and even the accuracy tends to decrease even though the decline is extremely small. Unlike experimental results on the third and last scheme, the similarity measurement results tend to be stable, in which the result difference of each experiment is very small. Table 6 is summary of the experiments on the CAI-UTM database, where the best performance is delivered by the last scheme, i.e. the maximum accuracy is 91.40%, the minimum accuracy is 90.60%, while the average accuracy is 90.97%.

In this database, Experimental results look different when compared to the previous three databases. In the first scheme, it appears that the addition of features actually decreases the accuracy, although the decreasing of accuracy is not significant, this is because the training image used has the same expression. While in the second until the last scheme shows an almost stable accuracy despite the number of features added. Overall, the results demonstrated that the little difference between the maximum, minimum, and average accuracies produced very small standard deviation, i.e. 0.002. This indicates that all experiments produce the similar accuracies as shown in Figure 6.

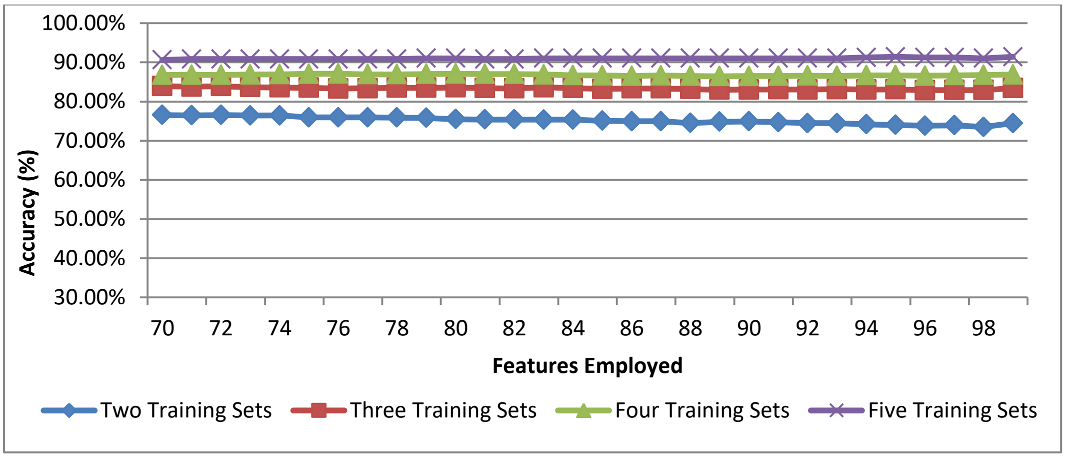

Figure 6. Accuracy on the CAI-UTM

Table 6. Experiments Conducted for each Scheme on the CAI-UTM

	Training Sets				The Best
	Two	Three	Four	Five	
Max	76.53%	83.91%	87.10%	91.40%	91.40%
Min	73.50%	82.89%	86.43%	90.60%	90.60%
Average	75.18%	83.33%	86.76%	90.97%	90.97%
Stdev	0.009	0.003	0.002	0.002	0.002

5. Compare to Similar Methods

The results on the ORL, the YALE, and the UoB of our proposed method are also compared to similar methods, i.e. PCA, LDA, combining LST+LDA, S-LDA, LPP, Simplification of LPP, S-MFS, and S-NPE. The first, the results of our proposed method on the ORL database are compared to PCA, LDA, combining LST+LDA, S-LDA, LPP, Simplification of LPP, S-MFS, and S-NPE. The comparison results displayed that the method of S-LPP, LST+LDA, S-MFS, S-NPE, and S-LDA outperformed to our proposed method for two, three, and four training samples, but our proposed method is better than the PCA, LDA, combining LST+LDA, S-LDA, LPP, Simplification of LPP, S-MFS, and S-NPE when five images are applied as training samples as shown in Figure 7. It shows that our proposed method is suitable for more images than fewer images as the training samples

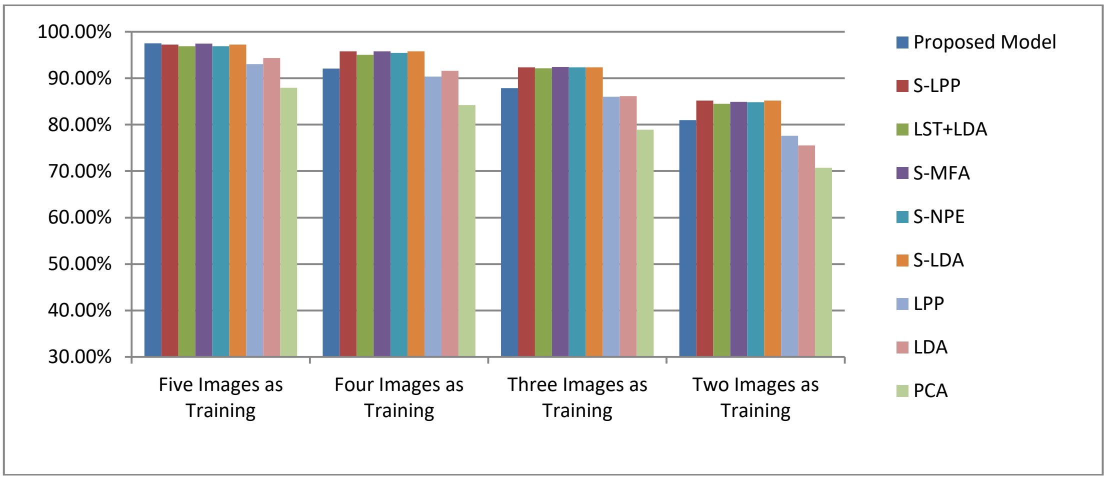

Figure 7. Comparison Results on the ORL Database

The second comparison is conducted on the YALE database. Our proposed is also compared to similar method, such as PCA, LDA, combining LST+LDA, S-LDA, LPP, Simplification of LPP, S-MFS, and S-NPE. The results show that experiments for all schemes are better than similar method, i.e. PCA, LDA, combining LST+LDA, S-LDA, LPP, Simplification of LPP, S-MFS, and S-NPE as shown in Figure 8. Only LST+LDA delivers accuracy close to our proposed method results, but our proposed method is still better than LST+LDA method. Our proposed method is suitable for face images with occlusion effects.

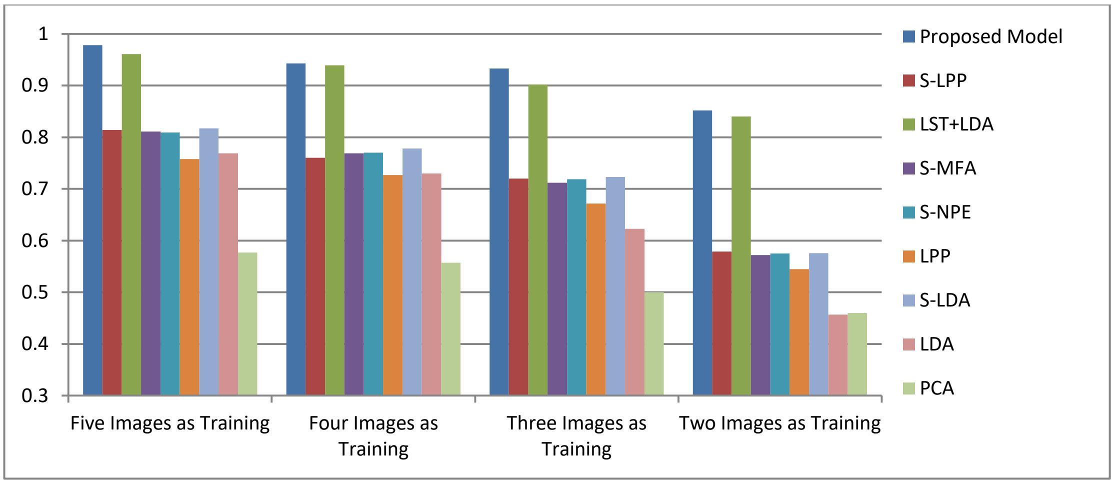

Figure 8. Comparison Results on the YALE Database

The last comparison is performed to UoB database as shown in Figure 9. The results indicates that our proposed method also delivered higher accuracy than others, i.e. Principal

Component Analysis (PCA), Linear Discriminant Analysis (LDA), Locality Preserving Projection (LPP), and Orthogonal Locally Preserving Projection (O-LPP) for all schemes, and even our proposed method produces much higher accuracy than PCA, LDA, LPP, and O-LPP. The comparison results indicate that our proposed method can be accepted as feature extraction method on the face image.

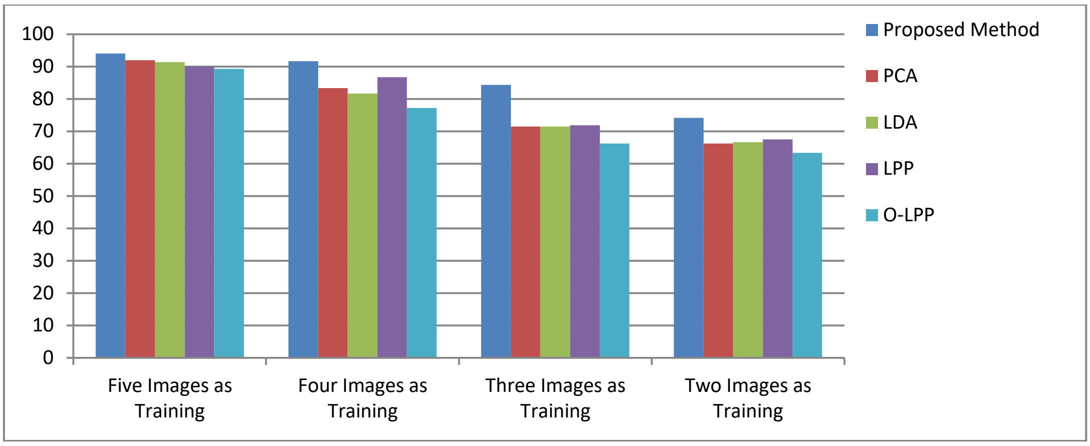

Figure 9. Comparison Results on the UoB Database

6. Conclusion

Proposed method has been evaluated by using four face image databases, and it can be concluded that the contradictory of the laplacian smoothing transform has extract the object features through positive diagonal elements. The experimental results also prove that singularity cases never appear when the eigenvectors and eigenvalues are computed. The performance results displayed that the maximum accuracies for the ORL, the YALE, the UOB, and CAI-UTM are 97.50%, 97.78%, 94.00%, and 91.40% respectively, whereas the best standard deviation are 0.006 for the ORL, 0.027 for the YALE, 0.005 for the UOB, and 0.002 for the CAI-UTM database.

The proposed method also efficienly worked on occlusion condition. It can be proved when the proposed method was evaluated by the YALE database. It has delivered higher accuracy than the others. The evaluation results showed that our proposed method is better than others, i.e. as PCA, LDA, combining LST+LDA, S-LDA, LPP, Simplification of LPP, S-MFS, and S-NPE.

Acknowledgement

This research was fully subsidized by the Directorate General of Higher Education through competence grand with contract number 1225/UN46.31/PN/2017

References

[1] H Zou, T Hastie, R Tibshirani. “Sparse Principal Component Analysis”. *J. Comput. Graph. Stat.* 2006; 15(2): 265–286.
[2] Q Ding, ED Kolaczyk. “A Compressed PCA Subspace Method for Anomaly Detection in High-Dimensional Data”. *IEEE Trans. Inf. Theory*. 2013; 59(11): 7419–7433.
[3] IGPS Wijaya, K Uchimura, G Koutaki. “Face Recognition Using Holistic Features and Linear Discriminant Analysis Simplification”. *Telkomnika*. 2012; 10(4): 775–787.
[4] E Setiawan, A Muttaqin. “Implementation of K-Nearest Neighbors Face Recognition on Low-power Processor”. *Telkomnika*. 2015; 13(3): 949–954.
[5] A Muntasa, IA Siradjuddin, RT Wahyuningrum. “Multi-Criteria in Discriminant Analysis to Find the Dominant Features”. *Telkomnika*. 2016; 14(3): 1113–1122.
[6] H Lu, ANVKN Plataniotis. “MPCA: multilinear principal component analysis of tensor objects”. *IEEE Trans. Neural Networks*. 2008; 19(1).

[7] Huang, JFSM Yang. "Improved principal component regression for face recognition under illumination variations". *IEEE Trans. Signal Process. Lett.* 2012; 19(4): 179–182.

[8] Jian Huang, PC Yuen, Wen-Sheng Chen, Jian Huang Lai. "Choosing Parameters of Kernel Subspace LDA for Recognition of Face Images Under Pose and Illumination Variations". *IEEE Trans. Syst. Man Cybern. Part B*. 2007; 37(4): 847–862.

[9] A Muntasa. "A new approach: The local feature extraction based on the new regulation of the locally preserving projection". *Appl. Math. Sci.* 2015; 9(101–104): 5065–5078.

[10] J Chen, Z Luo, T Takiguchi, Y Ariki. "Multithreading cascade of SURF for facial expression recognition". *EURASIP J. Image Video Process.* 2016; 2016: 1–13.

[11] J Ye, Q Li. "A two-stage linear discriminant analysis via QR-decomposition". *IEEE Trans. Pattern Anal. Mach. Intell.* 2005; 27(6): 929–941.

[12] A Muntasa, IA Sirajudin, MH Purnomo. "Appearance global and local structure fusion for face image recognition". *Telkomnika*. 2011; 9(1): 125–132.

[13] H Hu, P Zhang, F la Torre. "Face recognition using enhanced linear discriminant analysis". *IET Comput. Vis.* 2010; 4(3): 195.

[14] J Ye, Q. Li. "A two-stage linear discriminant analysis via QR-decomposition". *IEEE Trans. Pattern Anal. Mach. Intell.* 2005;. 27(6): 929–941.

[15] J Liu, S Chen. "Discriminant common vectors versus neighbourhood components analysis and Laplacianfaces: A comparative study in small sample size problem". *Image Vis. Comput.* 2006; 24(3): 249–262.

[16] H Cevikalp, M Neamtu, M Wilkes, A Barkana. "Discriminative common vectors for face recognition". *IEEE Trans. Pattern Anal. Mach. Intell.* 2005; 27(1): 4–13.