ISSN 2229-5518

Recognition of Multi-Pose Face In Colour Images Using Gabor Filters Based On SVM Concept

V Santhosh Kumar, A. Nageshwar Rao, S Anand Kumar, T Santosh Kumar

Abstract—Human face recognition plays an important role in applications such as video surveillance, human computer interface, and face image database management. This paper presents an improved face recognition method for multi-pose face recognition in color images, which addresses the problems of illumination and poses variation. At first, color multi-pose faces image features were extracted based on Gabor wavelets with different orientations and scales filters, then the mean and standard deviation of the filtering image output are computed as features for face recognition. In addition these features were fed up into support vector machine (SVM) for face recognition over a wide range of facial variations in the color, position, scale, orientation, 3D, pose, and expression in images from stereo-pair database.

Index Terms—File systems, overlay networks, denial-of-service attacks, performance and scalability, location hiding, gabor filters, svm concept.

I. Introduction

Human face recognition plays an important role in applications such as video surveillance, human computer interface, and face image database management. In recent years face recognition has received substional attention, but still remained very challenging in real applications. Despite the variety of approaches and tools studied, face recognition is not accurate or robust enough to be used in uncontrolled environments.

There are two approaches for face recognition, holistic and geometric. Geometric approaches dominated in the 1980's where simple measurements such as the distance between the eyes and shapes of lines connecting facial features. We used to recognize faces, while holistic methods became very popular in 1990'svwith the well known approach of eigenface. Even though holistic methods such as neural networks are more complex to implement then their geometric counterparts, their application is much more straight forward, whereby an entire image segment can be reduced to a few key values for comparison with other store key values and no exact measures or knowledge such as eye locations needs to be known. The problem with this "grab all" approaches was that noise, occlusions such as glasses and any other non-face image attribute could be learned by the holistic algorithm and become part of the recognition result even though such factors are not unique to faces.

- *V Santhosh Kumar is currently working as a associate professor at Anurag College of Engineering, Hyderabad E-mail: vsanthosh46@gmail.com*
- *A Nageshwar Rao is working as a assistant professor at Visveswaraya College of Engineeringand Technology. E-mail: nageswar.arla@gmail.com*
- *S Anand Kumar is currently doing his master's in technology at JPNE, Hyderabad*
- *T Santosh Kumar is currently working as a associate professor at Anurag College of Engineering, Hyderabad E-mail: santosh.tula@gmail.com*

Color representation is a problem of significant importance in the fields of computer vision and image processing. Color-based image classification is invariant to transformation in the image that is due to rescaling, translation, distortion, and rotation. In this paper, we first convert the original RGB image to the different color spaces image. Multi-pose face representation and based on the output of Gabor filters in different color space is computed. Finally they obtained Gabor feature vectors are fed up into support vector machine for face recognition and the results of face recognition in different color space have been compared.

Face representation (FR) plays a typically important role in face recognition and methods such as principal component analysis (PCA) and linear discriminate analysis (LDA) have been received wide attention recently. However, despite of the achieved successes, these FR methods will inevitably lead to poor classification performance in case of great facial variations such as expression, lighting, occlusion and so on, due to the fact that the image gray value matrices on which they manipulate are very sensitive to these facial variations.

II. Face Recognization

A real-time face recognition system can be separated into three modules:

1. Face Location
2. Face Normalization
3. Face Recognition.

Standardized Face Database:

The input into the system and consequently, the input into the first face location module will be either a still image or a frame from a video sequence. Analysis of the complex scene will then be performed to output an estimated location of the face.

The face normalization module will then aim to transform the image into a standardized format, where differing scaling factors and rotational angles of faces, differing lighting conditions and background environment, as well as differing facial expressions of faces will be considered. The normalized faces can then be entered into the face recognition module, either to add a new face into the database or to recognize a face from the existing database.

Research and focus has been placed upon the face recognition stage, and most developmental work has been performed using photo databases and still images that were created under a constant predefined environment. Since a controlled environment removes the need for extensive normalization adjustments, reliable techniques have been developed to recognize faces with reasonable accuracy provided the database contains perfectly aligned and normalized face images.

Thus the challenge for face recognition systems lie not in the recognition of the faces, but in the normalizing all input face images to a standardized format that is compliant with the strict requirements of the face recognition module. Although there are models developed to describe faces, such as texture mapping with the Candide model there are currently no developed definitions or mathematical models defining what the important and necessary components are in describing a face.

Due to the complexity of the face recognition problem, a modular approach was taken whereby the system was separated into smaller individual stages. Focus was then placed upon making each stage a reliable section before integrating the modules into a complete system. The face recognition system includes three major tasks – face detection, face normalization and face recognition – and each of these tasks are further broken down into separate stages. This chapter details the design of each of these stages, providing a detailed description of how the system was constructed.

A. ***Face Recognition:***

When designing a complex system, it is important to begin with strong foundations and reliable modules before optimizing the design to account for variations. Provided a perfectly aligned standardized database is available, the face recognition module is the most reliable stage in the system. The biggest challenge in face recognition still lies in the normalization and preprocessing of the face images so that they are suitable as input into the recognition module. Hence, the face recognition module was designed and implemented first.

The task of implementing a face recognition module is therefore accomplished first by considering a set of predefined face inputs rather than using variable images. It is important to begin with as little variable parameters as possible, and a pre-processed face database omits any possible uncertainty from the detection and normalization modules.

Given a perfect set of faces such that the scale, rotation, background and luminance is controlled, the recognition module can be designed to work with the optimal ideal inputs, since it is crucial that the performance of this foundation module be as optimized as possible. Its ability to recognize an ideal database will determine the best possible performance attainable by the overall complete system. Any subsequent development and implementation of the face detection and normalization module will therefore be aimed at providing this ideal set of database.

B. ***Face Detection:***

The first step in trying to build an ideal database for the recognition modules is to locate the exact position of the face in the image. Whether an input image is to be added into the system or be tested for identity, before any processing can be applied at the recognition stage, the face needs to be detected, extracted and normalized. The process of face detection and face normalization is intertwined and supported by one another. The two modules do not work completely as a stand-alone system and require the intermediate outputs of each module to continue with any further processing. In the case of face detection, two separate stages have been designed – the initial coarse detection phase, and the refined face search stage.

The coarse detection phase involves a quick scan over the complete image analyzing the color content of the input. The purpose of this stage is to reduce the search space by identifying the skinned regions of the image, so that the second face detection stage, which is performed after normalization of the luminance, can apply refined search techniques in order to locate the exact position of the face in the image.

The second stage has been designed to use two different algorithms depending on the status of the database. When the database is empty and no faces have been processed by eigenface decomposition, normalized cross correlation is used to determine the centre location of the face.

C. ***Skin Detection:***

All images or complex scenes from a video sequence enter the face recognition system through the face detection module, confronting the first stage – skin detection. The purpose of this first stage is to perform a fast coarse search of the scene in order to locate

Probable skinned regions in the image, so that non-skinned backgrounds can be eliminated with the knowledge that the face will not be located in those regions of the image.

A smaller image can then be extracted from the scene, such that subsequent searches for the exact location of the face can be performed on a reduced search space rather than on the entire image. This not only helps with increasing the speed of processing, but also the accuracy of the location of the face by eliminating the probability of error and reducing the possibility of erroneous false matches.

It has been demonstrated that human skin tones form a special category of colors, distinctive from the colors of most other natural objects. Although skin colors vary between different people and different races, it was found in YUV color representation; human skin colors remain distributed over a very small region on the chrominance plane.

D. ***Color Segmentation:***

To perform the skin region analysis, each pixel is firstly classified as being either skin or non-skin. In order to increase the speed of this module and realizing that it only acts as a fast coarse search of the scene, a down sampled version of the image is used. For an input image of size 240 by 320, a down sampled rate of four is adequate, such that the skin detection module only needs to operate on a 60 by 80 image. Down sampling by a factor higher than four has been observed to degrade the accuracy of the search dramatically.

Recall that the purpose of the color segmentation is to reduce the search space of the subsequent modules, thus, it is important to determine as tight a box as possible without cutting off the face. It is common during the color segmentation to return values that are closely skin but non-skin, or other skin-like colored regions that is not part of the face or the body. These noisy erroneous values are generally isolated pixels or group of pixels that are dramatically smaller than the total face regions, which would be represented by a big connected region in the binary image. Inclusion of these noisy pixels would result in a box that is much larger than intended and defeat the purpose of the segmentation.

E. ***Refined Face Search:***

The input into the second refined face detection stage is a luminance normalized skinned region of the image. This stage is the final stage of the face detection module, involving a refined search for the location of the face. The precision of the output from the refined search modules will determine how perfectly aligned the images are. Subsequently, as the recognition module requires ideal inputs that satisfy a list of restrictive criteria, the accuracy of these stages will be a measure of how successful the recognition and the overall system will be. It is therefore extremely important and a high accuracy rate is of utmost concern. Depending on the status of the face database, two methods have been designed for this important stage of the detection module – Normalized Cross Correlation (NCC) and Face Space Projection.

NCC involves finding the best match between a template and a sequence of windows, while face space search involves projecting the sequence of windows into face space and measuring how "face-like" each window is. Therefore, one major difference between the two searches is the input data required and the choice of technique used is dependent on what information is available.

The NCC search requires a template; therefore, a typical face or average face needs to be available in order for NCC to work. Face Space Projection requires a set of eigenface so that each window can be projected into face space. Hence, unless the set of dominant basis vectors has been calculated and available for use, the projection technique cannot be used.

F. ***Face Normalization:***

The process of face detection and face normalization is intertwined and the success of the system is dependent on how well the two modules work together. Face detection will only be successful if the input faces are normalized so that the orientation and lighting of the images are similar to the stored templates of the system. In other words, the input images have to be adjusted in such a way that it would seem as if all input images are captured under the exact same conditions and environment. This remains the biggest challenge yet to be solved by the many research groups in this area, and is the major reason why a complete solution to the Problem of face recognition has not been presented. The face normalization module can be divided into three main categories – scaling normalization, rotation normalization and lighting normalization.

1) ***Lighting Normalization:***

Lighting normalization involves adjusting the luminance of the image so that all images can be regarded as taken under the same lighting conditions. Since the refined face search modules rely on the intensity values on each pixel, careful attention needs to be paid in conserving the same amount of luminance on all images such that no particular bias is placed upon lighting differences. The recognition should be based upon the relationship and the variances between faces and not on the background lighting conditions or the time of day that the image was captured.

The lighting normalization module is applied in two places: once before the refined detection module and once prior to entering the face images into the recognition module. The first normalization is applied to enhance the accuracy of locating the face, by ensuring that all selected regions contain the same total energy as the average face. This is to avoid extreme biasing towards a particular lighting condition in the image when the distance functions are calculated. There was a particular problem with NCC search when light was incident from one end of the image to the other, and biasing was placed upon the brighter side since it would naturally record a higher correlation value regardless of the template.

ISSN 2229-5518

The addition of a lighting normalization stage relieved the problem.

The second lighting normalization stage was applied so that all images inputting into the recognition module will have the same total energy, so that no bias is placed upon a particular face during recognition. Since all images at this stage will be the same size, a conservation of energy would act as if all images of the heads were taken under the same lighting conditions. The total energy of an image is the sum of the squares of the intensity values.

$$Energy = \sum^{image} (intensity)^2 \qquad (1)$$

This makes the ratio between intensities a square rooted relationship between two images. That is, if we compute the ratio of energy between image A and image B to be *e*, we would multiply the total energy of image B by *e* so that image A and B will have the same total energy. We would however, only multiply each intensity value by the square root of *e*, $\sqrt{e}$, so that the sum of the squares of ($B*\sqrt{e}$) is the same as the norm of A. In our recognition system, all images are stored as intensity values. Therefore, inside the luminance normalization module, the energy of the selected window will first be calculated and then compared to the energy of the average window. The ratio of the energy is then computed, and the intensity of each pixel in the window will be multiplied by the square root of that ratio. The resultant image will be such that the total energy of the image is normalized.

2) *Scaling Normalization:*

A standardized scale is essential for the correct operation of the two refined search stages and the recognition module. In order for accurate face detection, it is a requirement that the face in the input image matches the size of the faces that is in the stored database; and for recognition, comparison between faces is not possible unless the faces are matched in scale. Scaling problems are a result of faces being photographed at different distances from the camera. Hence, whether the head appears bigger due to the face being extremely close to the camera, or smaller because the face is far away, all the faces need to be normalized to a standard size that conforms to the database.

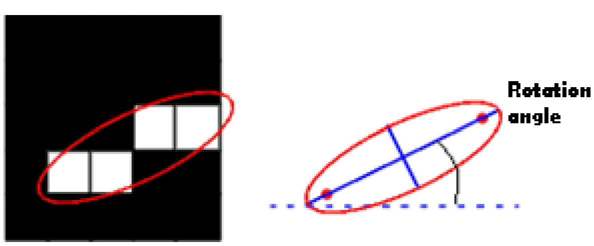


Figure1: Illustration of the determination of the rotational angle

3) *Rotation Normalization:*

In a video sequence, the faces in the scenes will typically be subjected to many different rotational orientations. In order to match the recognition ability of human beings, which is currently under vigorous research, a face recognition system would need to be able to identify faces without placing restrictions upon the allowable rotational orientations of the faces. Due to the three dimensional rotary motion of heads, forward tilting, planar rotations (sideway tilting of heads) and turning rotations (twisting of heads) are all major problems encountered by face normalization modules. However, due to the complexity, only planar rotation – sideways tilting of the head – has been investigated in this system.

III. FACE FEATURES EXTRACTED USING GABOR FILTERS

One of the crucial aspects of texture classification is extraction of proper and representative textural features. Here we consider the features generated by Gabor filters. Basically Gabor filters are a group of wavelets with each wavelet capturing energy at a specific frequency and a specific direction. Expanding a signal using this basis provides a localized frequency description, therefore capturing local features energy of the signal. Multipose faces images features van be extracted from this group of energy distributions.

A. *Wavelets:*

The approximation coefficients give a broad picture of the signal highlighting the major features. In applying wavelet analysis to sampled signals, a down sampling operation is performed after each level of decomposition. This simply means the number of data points in the components at level j approximation or detail will be reduced by a factor of two compared to the corresponding number of data points at level (j-1). Thus the advantage of using a wavelet transform is it reduces the size of the inputs to a neural network while at the same time providing "good features" by using the approximation coefficients.

Decompose a signal into component parts

B. *Fourier analysis:*

Signal can be represented as a (possibly infinite) sum of sine and cosine functions

Signal becomes a set of wavelet coefficients

Coefficients represent features of signal

Given the basis functions ----- we can represent f(t) as

$$= \sum_{k\varepsilon z} \alpha_{jo,k} \phi_{jo,k(t)} + \sum_{j=jo}^{\alpha} \sum_{kez} d \qquad (2)$$

Where α's are the approximation coefficients and d's are the detail coefficients and defined to be

ISSN 2229-5518

$$\alpha_{jv,k} = \int_R f(t)\varphi_{jv,k}(t)dt \quad (3)$$

$$d_{j,k} = \int_R f(t)\varphi_{j,k}(t)dt \quad (4)$$

In the discrete signal case we compute the Discrete Wavelet Transform by successive low pass and high pass filtering of the discrete time-domain signal. This is called the Mallat algorithm or Mallat-tree algorithm

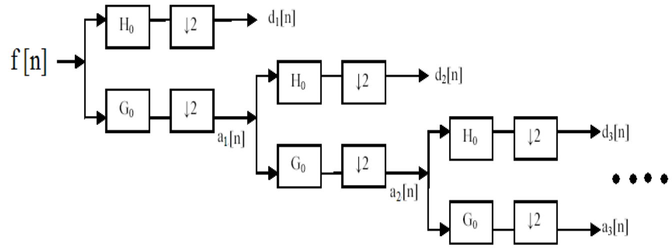


Figure 2: Mallat-tree

C. ***Algorithm:***

Capture features of images e.g. edges in coefficients

Use Haar wavelets

Square basis functions

Calculate coefficients, truncate, quantize

Easy to implement and compute

A Particular Gabor elementary function can be used as the mother wavelet to generate a whole family of Gabor wavelets. A two dimensional Gabor function g(x, y) and its Fourier transform G (u, v) van be written as:

$$g(x,y)=\left[\frac{1}{2\Pi\sigma_x\sigma_y}\right]\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma^2 x}+\frac{y^2}{\sigma^2 y}\right)\right]+2\Pi j\omega x \quad (5)$$

$$G(u,v)=\exp\{-\frac{1}{2}\left(\frac{(u-v)^2}{\sigma^2}+\frac{v^2}{\sigma^2 v}\right) \quad (6)$$

Where $j=\sqrt{-1}, \sigma_u = \frac{1}{2\Pi\sigma_v}, and \sigma_u = \frac{1}{2\Pi\sigma_v},$

Gabor functions form a complete but non orthogonal basis set. Expanding a signal using this Basis provides a localized frequency description. Let g(x, y) be the mother Gabor wavelet, then this self similar filter dictionary can be obtained by appropriate dilations and rotations of g(x,y) through the generation function:

$$g_{mn}(x,y) = a^{-m} g(x,y), \quad (7)$$

Where a>1, m, n=integer

$$y' = a^{-m}(\cos\theta + \sin\theta)x' \quad (8)$$

$$x' = a^{-m}(\cos\theta + \sin\theta) \quad (9)$$

$$\theta = \frac{n\Pi}{k}$$

Where n and k indicate the orientation and scale of the filter respectively K is the total number of orientations and n is the total number of scales in the filter bank the scale factor is meant to ensure that the energy is independent of m.

For a given image I(x, y), the decomposed image at scale k and orientation n is:

$$= \int I(x,y) g^*_{mn}(x - x_1, y - y^1)\mathrm{dx1dy1} \quad (10)$$

Where * indicate the complex conjugate.

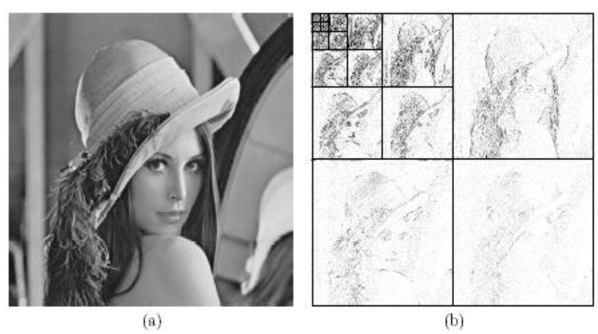
It is assumed that the local texture regions are spatially homogeneous and the mean $\mu_{\mu mn}$ and the standard deviation α_{mn} of the magnitude of the transform coefficients are used to represent the region for classification purposes:

$$\mu_{\mu mn} = \iint |w_{wn}(x,y)| dxdy \quad (11)$$

And

$$\alpha_{mn} = \iint (|W_{mn}(x,y) - \mu_{mn}|)^2 dxdy \quad (12)$$

A feature vector is now constructed using and as feature



components

Figure 3: wavelets for vision

D. ***Three mother wavelets***

$$\psi^d_{j,m,n}(x,y) := \psi_{j,m}(x)\psi_{j,n}(y),$$

$$\psi^v_{j,m,n}(x,y) := \psi_{j,m}(x)\phi_{j,n}(y), \quad (13)$$

$$\psi^h_{j,m,n}(x,y) := \phi_{j,m}(x)\psi_{j,n}(y),$$

The texture description for k scales and n orientations is given by:

Features = $\{\mu_{00}, \sigma_{oo}, \mu_{11}, \sigma_{11}\}$

E. Gabor Features:

A Gabor wavelet is a complex planar wave restricted by 2D Gaussian envelope. Features are Gabor wavelets at different scales and orientations convolved with the face image.After the Gabor features of Multi pose faces image have been extracted which had described, the next task is that is that recognition of color face image using support vector machine (SVM)

IV. SUPPORT VECTOR MACHINES

A. History:

SVM is inspired from statistical learning theory. SVM was first introduced in 1992. SVM becomes popular because of its success in handwritten digit recognition -1.1% test error rate for SVM. This is the same as the error rates of a carefully constructed neural network, LeNet 4.SVM is now regarded as an important example of "kernel methods", arguably the hottest area in machine learning

B. Introduction:

What is benefits SV learning?

Based on simple idea

High performance in practical applications

Characteristics of SV method

Can dealing with complex nonlinear problems

(Pattern recognition, regression, feature extraction) But working with a simple linear algorithm (by the use of kernels)

C. Applications

This maximum-margin separator is determined by a subset of the data points. Data points in this subset are called "support vectors". It will be useful computationally if only a small fraction of the data points are support vectors, because we use the support vectors to decide which side of the separator a test case is on. To find the maximum margin separator, we have to solve the following optimization problem. This is tricky but it's a convex problem. There is only one optimum and we can find it without fiddling with learning rates or weight decay or early stopping. Don't worry about the optimization problem. It has been solved. It's called quadratic programming. It takes time proportional to N^2 which is really bad for very big datasets so for big datasets we end up doing approximate optimization!

$$\mathbf{w}.\mathbf{x}^c + b > +1 \quad \textit{for positive cases}$$

$$\mathbf{w}.\mathbf{x}^c + b < -1 \quad \textit{for negative cases}$$

$$\textit{and} \quad \|\mathbf{w}\|^2 \quad \textit{is as small as possible}$$

D. Support Vector Machine Algorithm:

1. Choose a kernel function

2. Choose a value for C

3. Solve the quadratic programming problem (many software packages available)

4. Construct

V. CONCLUSION

This paper proposes a multi-Pose faces image recognition algorithm, in which Gabor feature representation in different color space is extracted an SVM classifier are employed. After extracting Multi-Pose faces image features, Gabor filter the image has been filtered with four orientations and six scales filters, and then the mean and standard deviation of the image output are computed. Finally, the obtained Gabor feature vectors are fed up into support vector machine (SVM) for classification. This paper also demonstrates that when a bank of Gabor filters is applied to an image, there are strong relationships between the outputs of the different filters. These relationships are used to devise a new feature which is capable of describing Multi-Pose faces image information in a concise manner. The performance of color space features is found to be better than that of the features which just extracted from gray image. The experimental results show this approach can be used to automatically identify the plant categories. The future study will focus on how to extract more efficient features from bark images to improve the classification accuracy further.

VI. REFERENCES

[1] V. Bruce, P.J.B. Hancock and A.M. Burton (1998) "Comparisons between human and computer recognition of faces", Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition, 14(16),pp.408-413

[2] G. Yongsheng and M.K.H. Leung. (2002) "Face recognition using line edge map", IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(6), 764-779.

[3] M. Turk and A. Pentland. (1991) "Eigenfaces for recognition". Journal of Cognitive Neuroscience, 3, pp.71- 86.

[4] Zhi-Kai Huang, Hui-Ming Huang, Yun-Min,Xie, Ling- Ying Hou. (2008) "Face Recognition Using Model of the Scene Shape Unsupervised Image Segmentation Using EM Algorithm by Histogram", Journal of ComputationalInformation Systems: Vol. 4,pp.919-928.

[5] J.E. Meng, W. Chen and W. Shiqian. (2005) "High speed face recognition based on discrete cosine transform and RBF neural networks", IEEE Transactions on Neural Networks, 16(3),pp.679 – 691.

[6] R. Fransens, C. Strecha, L. Van Gool.(2005) "Parametric Stereo for Multi-Pose Face Recognition and 3D-Face Modeling", ICCV 2005 workshop "Analysis and Modeling of Faces and Gestures".

[7] Zhi-Kai Huang and Zhi-Feng Wang. (2007) "Bark Classification Using RBPNN in Different Color Space", Neural Information Processing – Letters and Reviews Vol. 11, No. 1, pp.7-13.

[8] M. Kirby, L. Sirovich, Application of the Karhunen–Loeve procedure for the characterization of human faces, IEEE Trans. Pattern Anal. Mach. Intell. 12 (1990) 103–108.

[9] M. Turk, A. Pentland, Eigenfaces for recognition, J. Cognitive Neurosci. 3 (1) (1991) 71–96.

[10] J. Yang, D. Zhang, et al., Two-dimensional pca: a new approach to appearance-based face representation and recognition, IEEE Trans. Pattern Anal. Mach. Intell. 26 (1) (2004) 131–137.

[11] J. Ye, Generalized low rank approximation of matrices, in: International Conference on Machine Learning, pp. 2004, pp. 887–894.

[12] J. Liu, S. Chen, Non-iterative generalized low rank approximation of matrices, Pattern Recognition Lett. 27 (9) (2006) 1002–1008.

[13] P. Belhumeur, J. Hespanha, D.J. Kriegman, Eigenfaces vs. Fisherfaces:recognition using class specific linear projection, IEEE Trans. Pattern Anal. Mach. Intell. 19 (7) (1997) 711–720.

[14] D. Swets, J. Weng, Using discriminant eigenfeatures for image retrieval, IEEE Trans. Pattern Anal. Mach. Intell. 18 (1996) 831–836.

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

[16] J. Liu, S. Chen, Discriminant common vecotors versus neighbourhood components analysis and laplacianfaces: a comparative study in small sample size problem, Image Vision Comput. 24 (3) (2006) 249–262.

[17] J. Liu, S. Chen, Resampling lda/qr and pca+lda for face recognition, in: The 18th Australian Joint Conference on Artificial Intelligence, 2005, pp. 1221–1224.

[18] J. Ye, Q. Li, A two-stage linear discriminant analysis via qrdecomposition, IEEE Trans. Pattern Anal. Mach. Intell. 27 (6) (2005) 929–941.

[19] Z. Hong, Algebraic feature extraction of image for recognition, Pattern Recognition 24 (1991) 211–219.