

Video-based face recognition in color space by graph-based discriminant analysis

S. Shafeipour Yourdeshahi^{1*}, H. Seyedarabi¹ and A. Aghagolzadeh²

1. Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran.

2. Faculty of Electrical and Computer Engineering, Babol University of Technology, Babol, Iran.

Received 05 March 2015; Accepted 18 May 2016

*Corresponding author: shafeipour@tabrizu.ac.ir (S. Shafeipour).

Abstract

Video-based face recognition has attracted significant attention in many applications such as media technology, network security, human-machine interfaces, and automatic access control system in the past decade. The usual way for face recognition is based upon the grayscale image produced by combining the three color component images. In this work, we consider grayscale image as well as color space in the recognition process. For key frame extractions from a video sequence, the input video is converted to a number of clusters, each of which acts as a linear subspace. The center of each cluster is considered as the cluster representative. Also in this work, for comparing the key frames, the three popular color spaces RGB, YCbCr, and HSV are used for mathematical representation, and the graph-based discriminant analysis is applied for the recognition process. It is also shown that by introducing the intra-class and inter-class similarity graphs to the color space, the problem is changed to determining the color component combination vector and mapping matrix. We introduce an iterative algorithm to simultaneously determine the optimum above vector and matrix. Finally, the results of the three color spaces and grayscale image are compared with those obtained from other available methods. Our experimental results demonstrate the effectiveness of the proposed approach.

Keywords: *Face Recognition; Key Frame; Intra-class; Inter-class; Color Component.*

1. Introduction

Face recognition has been one of the most popular research areas in computer vision during the recent decades due to its many applications such as media technology, surveillance systems, access control, and human-machine interfaces. However, by far, no comprehensive technique has been proposed to provide a robust solution to all problems. Most of the algorithms proposed are still image-based, and they have been proven successful in solving some of the problems such as pose, lighting and expression variations, occlusion, and low image resolutions. Recently, due to the much development in media technology and the increased demands on security, video-to-video face recognition has received a significant attention.

One of the most important properties of a video sequence is its temporal continuity. Using this

property, we can increase the recognition rate. In this regard, in the past decade, researchers have begun to use the spatiotemporal data [1]. A wide variety of popular approaches in this field try to represent a video sequence as a linear subspace or a nonlinear manifold [2]. Then the similarity between two sequences is obtained by measuring the distance between two subspaces. Most of these methods compute principal angles between two spaces. In these methods, the recognition process relies mainly on the representation of the video sequence as a linear subspace [3]. In [4], manifold has been partitioned into a number of linear subspaces by Maximal Linear Patch, and manifold to manifold distance has been converted to measure the similarity between pairwise linear models. In [5], Wang et al. have introduced a discriminative learning method called Manifold

Discriminant Analysis (MDA) to solve the recognition problem. Therefore, MDA tries to determine the mapping matrix in the grayscale images, where local models from the manifolds with different class labels can be better separated, and meanwhile, the local data compactness within each manifold can be better enhanced.

Some methods are based upon the still image techniques, which try to select key frames from a video sequence and then apply still image-based face recognition algorithms. Therefore, these techniques focus on extracting all key-frames from the video sequence as the representative frames.

On the other hand, color plays an effective role in the biometrics works. The RGB color space is an additive color model, and all the other color spaces that may be more appropriate for many types of visual task can be derived from this color space [6-10]. For example, the HSV color space has been chosen in image retrieval [11], and it has been demonstrated that the YCbCr color space is more useful than the RGB color model for the face detection tasks [12, 13]. A conventional way of face recognition is to convert the RGB color space into a grayscale image by combining the three color component images. In [14], Wang et al. have shown that the human visual system uses color cues for recognition, and using color space can improve the recognition rate.

In this paper, we compare the three color spaces RGB, YCbCr, and HSV in video-based face recognition using the graph-based discriminant analysis. Then we introduce our proposed method, which includes face detection in video sequence, clustering in order to obtain the key frames, and graph-based discriminant analysis in color space. Extensive experimental evaluation of the proposed model and its comparison for the three color spaces and other available methods are discussed in section 3, followed by discussion of the results.

2. Proposed algorithm

The proposed algorithm is based upon converting video sequences into some clusters, each of which acting as a linear subspace. The center of each cluster is considered as the cluster representative. The recognition process is completed by comparing the key frames. In this paper, two intelligent ways are proposed for clustering. For comparing the key frames, we consider the grayscale image as well as the color space. The three popular color spaces RGB, YCbCr, and HSV are used for the mathematical representation

of each frame and the recognition process used for the graph-based discriminant analysis.

By introducing the intra-class and inter-class similarity graphs to the color space, the problem changes to determining the color component combination vector and mapping matrix. In this paper, two equations are achieved using a number of mathematical operations, and an optimal solution is achieved by designing an iterative algorithm. Figure 1 shows a block diagram of the proposed algorithm. In the following parts, the details of our proposed method are explained.

2.1. Face detection in video sequence

Face detection is an essential early step in the face recognition systems. Real-time face detection is very significant in a video-based face recognition. The Cascade AdaBoost face detector, proposed by Viola and Jones [15], and face detection based on color information [16] are two main methods for a real time face detection. A better face region is normalized via finding the distance between two eyes. There are many methods for eye detection. The most important ones consist of template matching, eigen space, and Hough transform. For face detection in other frames, this point is used that the neighbouring regions within successive video frames tend to be highly correlated. Therefore, a search window is determined around the face region in the first frame, and by calculating the correlation between the current face region and the next frame into the search window, in which the highest correlation to conclude is the face region in the next frame. This procedure is repeated for the entire frames. Determining the size of the search window has a major role in the calculation.

2.2. Clustering and key frames

Several approaches have been presented for clustering and constructing local linear models from nonlinear manifold [17-19]. Most of them use iterative-based clustering methods such as k-means to classify a given data through a certain number of clusters. The most important problem with this method is the number of clusters required to be specified by the user. Also the linearity of the local model has a low accuracy, and therefore, the clustering result is not optimal. Two different algorithms have been proposed for clustering videos. In the first algorithm, data containing face detection in successive frames in the video sequence is used. In other words, the motion vectors obtained by the successive frames can predict the movements of the head and perform clustering.

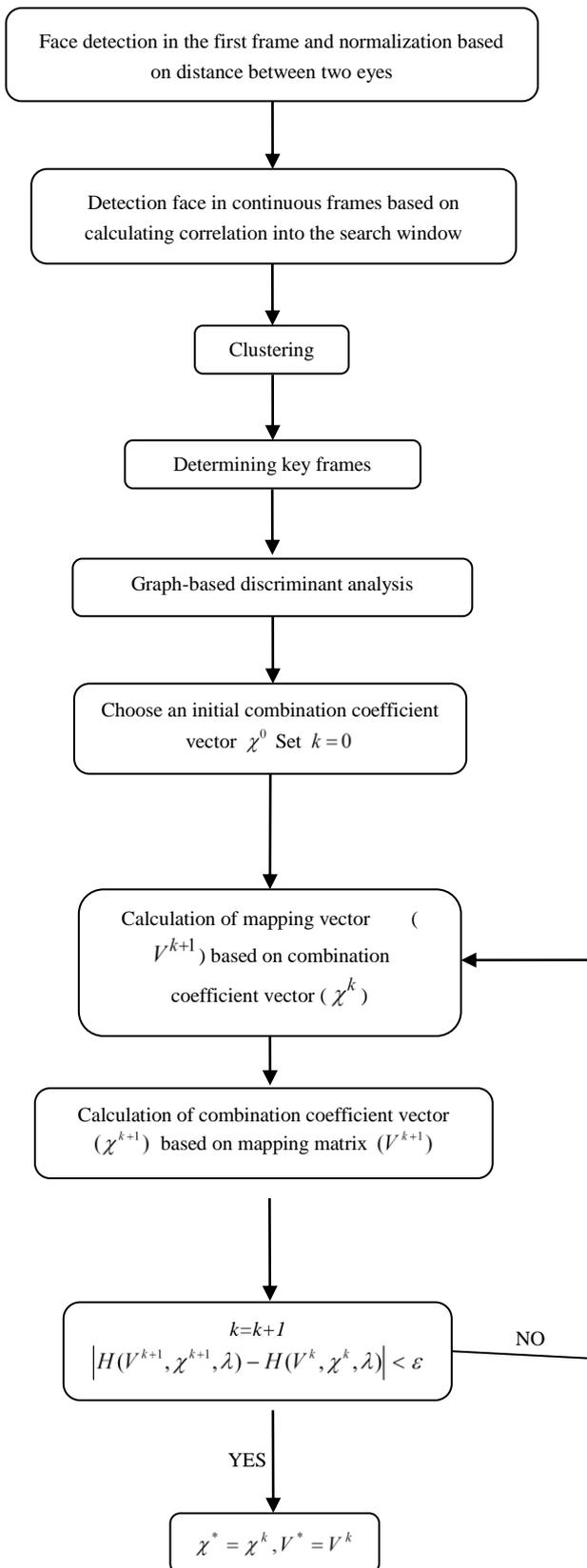


Figure 1. The block diagram of the proposed algorithm.

This method has the following advantages:

- Determining the number of clusters and linear model is quite intelligent, and with the implementation of this method, the number of clusters depends on moving the head in each video sequence.
- Linearity of the model has a high accuracy.
- This method is very simple because of the existence of the motion vector information.
- Due to changes in the direction of motion vector, we can predict the head direction.

In this clustering method, the key frames are selected based on the motion vectors with the most changes in the horizontal and vertical directions.

In the second clustering method, the first frame of each video is chosen as the center of the first cluster. Then by calculating the correlation between the first key frame and the next one and its comparison with the suitable threshold, it can be determined whether the second frame belongs to the first cluster. To obtain the second cluster, the first frame that does not belong to the first cluster is determined as the center of the second cluster. Similarly, this process is continued to successive frames until all frames are assigned to the clusters. It should be noticed that after obtaining all clusters, their key frames are compared with a specified threshold for merging. An appropriate choice of the threshold is important in determining the number of clusters.

2.3. Graph-based discriminant analysis in color space

The usual way for face recognition is to use grayscale images. In the present study, we used color information, and compared the three popular color spaces RGB, YcbCr, and HSV. The RGB color space is an additive color model, and all the other color spaces can be derived from it. In this color space, each pixel is represented by the three components red, green, and blue. If we can represent an image by separating the intensity of the color information, some processing steps would be faster.

MPEG compression, which is used in digital component video standard, is coded in the YCbCr color space. In this color space, the luminance and color information are separated. The Y component describes brightness, and Cb and Cr describe a color difference rather than a color. The HSV (hue, saturation, value) color space was designed with an emphasis on the visual perception. In the following part, face recognition is achieved using the graph-based discriminant analysis.

Let A_i be a color image with three columns, and each color component be a column vector. We suppose that the resolution of each color component is $m \times n$. Therefore, the color image A_i is expressed as an $N \times 3$ matrix: $A_i = [\alpha_{1i}, \alpha_{2i}, \alpha_{3i}]$, where $\alpha_{1i}, \alpha_{2i}, \alpha_{3i}$ are three color components related to a color space (RGB, HSV or YCbCr) and $N = m \times n$.

In this work, we tried to find the optimal coefficients to combine the three color components of the above color spaces. Suppose that Z_i is a combined image, given by (1):

$$Z_i = x_1\alpha_{1i} + x_2\alpha_{2i} + x_3\alpha_{3i} = [\alpha_{1i}, \alpha_{2i}, \alpha_{3i}]\chi = A_i\chi \quad (1)$$

where, x_1, x_2, x_3 are their coefficients. Therefore, our attempt was to find the coefficients so that Z_i was the optimal description of A_i for face recognition.

A graph in our model refers to a collection of vertices or nodes and a collection of edges that connect pairs of vertices. Given n points $\{Z_1, \dots, Z_n\}$ from the underlying manifold M , and each Z_i belonging to class c_i , the similarity score between two points with three color components is defined by:

$$K_{i,j} = \min \|\alpha_{li} - \alpha_{lj}\|^2 \quad l = 1, 2, 3 \quad (2)$$

where, $\|\bullet\|$ denotes the Euclidean norm. The two-adjacency graphs for inter-class and intra-class are defined by:

$$S_b(i, j) = \begin{cases} 1 & \text{if } K_{i,j} \geq \varepsilon \quad c_i \neq c_j \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

$$S_w(i, j) = \begin{cases} 1 & \text{if } K_{i,j} \leq \varepsilon \quad c_i = c_j \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

where, c_i is the class label for Z_i , and c_j is the class label for Z_j , and S_w and S_b are the intra-class similarity graph and the inter-class similarity graph, respectively.

Our aim was to minimize the distance between points of the connected points of S_w , and to maximize the distance between points of the connected points of S_b in the feature space. Therefore, the points on M manifold were mapped to a new manifold M' , i.e., $\phi: Z_i \rightarrow Y_i$. The matching score between two points in the new manifold can be computed using the Euclidean distance measures. Such a mapping can

be expressed by minimizing and maximizing the following two objective functions, respectively [20]:

$$\min \frac{1}{2} \sum_{i,j} (Y_i - Y_j)^2 S_w(i, j) \quad (5)$$

$$\max \frac{1}{2} \sum_{i,j} (Y_i - Y_j)^2 S_b(i, j) \quad (6)$$

where, $Y_i = V^T Z_i = V^T A_i \chi$. The minimization problem from (5) can be written as follows:

$$\begin{aligned} & \min \frac{1}{2} \sum_{i,j} ((V^T A_i \chi)^2 + (V^T A_j \chi)^2 - 2(V^T A_i \chi V^T A_j \chi)) S_w(i, j) \\ & = \sum_{i,j} V^T A_i \chi S_w(i, j) \chi^T A_i^T V - \sum_{i,j} V^T A_i \chi S_w(i, j) \chi^T A_j^T V \\ & = V^T A_\chi (D_w - W_w) A_\chi^T V \\ & = V^T A_\chi L_w A_\chi^T V \end{aligned} \quad (7)$$

and the maximization problem from (6) can be written as:

$$\begin{aligned} & \max \frac{1}{2} \sum_{i,j} ((V^T A_i \chi)^2 + (V^T A_j \chi)^2 - 2(V^T A_i \chi V^T A_j \chi)) \\ & \quad \cdot S_b(i, j) = V^T A_\chi L_b A_\chi^T V \end{aligned} \quad (8)$$

where, $L_w = D_w - W_w$ and $L_b = D_b - W_b$ are the Laplacian matrix, W_w and W_b are the weight matrices, D_w and D_b are the diagonal matrices, and $A_\chi = [A_1 \chi | A_2 \chi \dots | A_N \chi]$. It should be noticed that this calculation shows that L_w is a positive semi-definite. Equation 7 can be written as:

$$\begin{aligned} & \arg \min_{V\chi} V^T A_\chi L_w A_\chi^T V \\ & = \arg \min_{V\chi} (V^T A_\chi D_w A_\chi^T V - V^T A_\chi W_w A_\chi^T V) \\ & = \arg \min_{V\chi} (1 - V^T A_\chi W_w A_\chi^T V) \\ & \text{subject to } V^T A_\chi D_w A_\chi^T V = 1 \end{aligned} \quad (9)$$

The constraint $V^T A_\chi D_w A_\chi^T V = 1$ eliminates the arbitrary scaling factor. Therefore, Equation 9 can be re-written as follows:

$$\begin{aligned} & = \arg \max_{V\chi} (V^T A_\chi W_w A_\chi^T V) \\ & \text{subject to } V^T A_\chi D_w A_\chi^T V = 1 \end{aligned} \quad (10)$$

Similarly, Equation 8 can be re-written as a maximization problem, as follows:

$$\arg \max_{V\chi} (V^T A_\chi L_b A_\chi^T V) \quad (11)$$

Integrating (10) and (11), we have:

$$\arg \max_{V, \chi} (V^T A_\chi \beta L_b A_\chi^T V + V^T A_\chi (1 - \beta) W_w A_\chi^T V) \quad (12)$$

$$= \arg \max_{V, \chi} (V^T A_\chi (\beta L_b + (1 - \beta) W_w) A_\chi^T V) \\ \text{subject to } V^T A_\chi D_w A_\chi^T V = 1 \quad (13)$$

$$H(V, \chi, \lambda) = V^T A_\chi (\beta L_b + (1 - \beta) W_w) A_\chi^T V - \\ - \lambda (V^T A_\chi D_w A_\chi^T V - 1) \quad (14)$$

where, β is a suitable constant and $0 \leq \beta \leq 1$. To find V and χ , we first took the derivative of $H(V, \chi, \lambda)$ in (14) with respect to V :

$$\frac{\partial H(V, \chi, \lambda)}{\partial V} = 2A_\chi (\beta L_b + (1 - \beta) W_w) A_\chi^T V - 2\lambda A_\chi D_w A_\chi^T V \quad (15)$$

Equating the derivative to zero, $\frac{\partial H(V, X, \lambda)}{\partial V} = 0$, and thus we can write:

$$A_\chi (\beta L_b + (1 - \beta) W_w) A_\chi^T V = \lambda A_\chi D_w A_\chi^T V \quad (16)$$

Secondly, Equations 5 and 6 can be re-written as follow ($0 \leq \delta \leq 1$):

$$H(V, \chi, \lambda) = \sum_i V^T A_i \chi \chi^T A_i^T V \beta D_b(i, i) + \\ + \sum_{i, j} V^T A_i \chi \chi^T A_j^T V [(1 - \beta) S_w(i, j) - \beta S_b(i, j)] - \\ - \lambda (\sum_i V^T A_i \chi \chi^T A_i^T V D_w(i, i) - 1) \quad (17)$$

Therefore, we can write:

$$\frac{\partial H(V, \chi, \lambda)}{\partial \chi} = 2[\sum_i A_i^T V \beta D_b(i, i) V^T A_i + \sum_{i, j} A_j^T V \\ ((1 - \beta) S_w(i, j) - \beta S_b(i, j)) V^T A_i] \chi - \\ - 2\lambda \sum_i A_i^T V D_w(i, i) V^T A_i \chi \quad (18)$$

Equating the derivative to zero, $\frac{\partial H(V, \chi, \lambda)}{\partial \chi} = 0$,

therefore we have the following equation:

$$[\sum_i A_i^T V \beta D_b(i, i) V^T A_i + \sum_{i, j} A_j^T V ((1 - \beta) S_w(i, j) - \\ - \beta S_b(i, j)) V^T A_i] \chi = \lambda \sum_i A_i^T V D_w(i, i) V^T A_i \chi \quad (19)$$

Therefore, we have the following two equations:

Equation I:

$$A_\chi (\beta L_b + (1 - \beta) W_w) A_\chi^T V = \lambda A_\chi D_w A_\chi^T V \\ \text{subject to } V^T A_\chi D_w A_\chi^T V = 1 \quad (20)$$

Equation II:

$$[\sum_i A_i^T V \beta D_b(i, i) V^T A_i + \sum_{i, j} A_j^T V ((1 - \beta) S_w(i, j) - \\ - \beta S_b(i, j)) V^T A_i] \chi = \lambda \sum_i A_i^T V D_w(i, i) V^T A_i \chi \\ \text{subject to } \sum_i V^T A_i \chi \chi^T A_i^T V D_w(i, i) = 1 \quad (21)$$

For solving the above two equations, we used the generalized Eigen-equation $P\eta = \kappa Q\eta$ [21]. From the above theorem, the solutions to the above equations can be chosen as the eigenvectors of the generalized equation (for equation II, associating to the largest eigenvalue). Therefore, we can easily calculate the optimum points χ and V by an iterative algorithm. Suppose that $\chi = \chi^k$ is the initial value for χ in the k-th iteration. First we calculated $V = V^{k+1}$ from (20). In the second step, we calculated $\chi = \chi^{k+1}$ based on $V = V^{k+1}$. In the next iteration, $\chi = \chi^{k+1}$ was used as the initial value. This algorithm performs the above two steps successively until it converges. It should be pointed out that convergence may be determined after n+1 times of iterations, if $|H(V^{n+1}, \chi^{n+1}, \lambda) - H(V^n, \chi^n, \lambda)| < \epsilon$. Then we chose χ^{n+1} and V^{n+1} .

Given a test vide sequence, we first partitioned it into some clusters, and then using χ^{n+1} and V^{n+1} the center of each cluster was mapped to the new space. Therefore the similarity between manifolds could be obtained by calculating the distance between the cluster representatives in the new space.

3. Experimental results

The described algorithm was evaluated using the Honda/UCSD database [22] and the CMU MoBo database [23]. The first video database, Honda/UCSD, has been collected by K. C. Lee et al., and is used for video-based face recognition. The spatial and temporal resolution of each video sequence is 640×480 and 15 frames per second, respectively. In this database, every person turns his/her head in different speeds and rotations. Some of these sequences include a partial occlusion. The second database, MoBo (Motion of Body), was originally collected for the purpose of human identification from distance. The considered subset contained 96 sequences of 24 different subjects walking on a treadmill. Each person had 4 videos.

For each video sequence, after the detection of face and eye, the face area was marked by a

square with the dimensions $2.8d \times 2.8d$, where d is the distance between the two eyes. Figure 2 shows this face area. Figure 3 illustrates the face regions detected in some frames of the Honda/UCSD database. As described earlier, in the first method of clustering, the face regions obtained at the position of the frames within the search window are used.

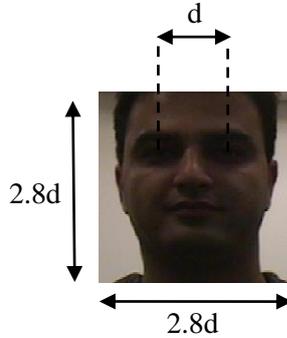


Figure 2. Face area marked by distance between two eyes.



Figure 3. Faces detected in one of the video sequences of Honda/UCSD database.

The search window dimension is 10×10 pixels. Changes in the horizontal and vertical directions, compared to the previous frame, are shown, respectively, with dx and dy . In other words, dx and dy are the motion vectors.

Figure 4 illustrates the variation of dx and dy in consecutive frames of a video sequence in the Honda/UCSD database. As an example shown in the figure, with the shift to the right human head, the motion vectors are moved to the right. Accordance with the motion vectors, clustering can be done. As described in the first clustering method, the key frames are selected based on the highest $|dx| + |dy|$ within the cluster. These frames in addition to the first frame of a video sequence are considered as the key frames.

The second approach to clustering was performed in successive frames by comparing the correlations. The number of clusters was determined according to the threshold value. In this simulation, this value was set as 0.7. Figure 5 illustrates seven clusters obtained by this method.

In the second clustering method, centers of each cluster were selected as the key frames.

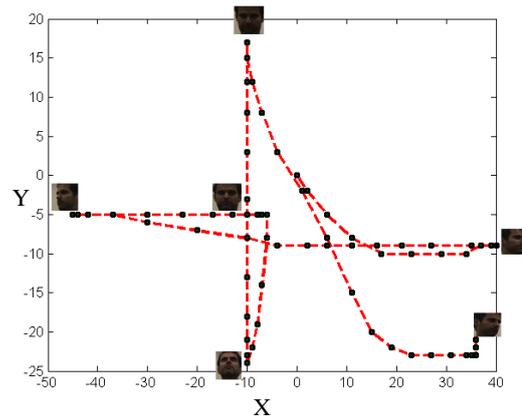


Figure 4. Variation of dx and dy in one video sequence of the Honda/UCSD database.

Figure 6 shows the key frames of two video sequences in the Honda/UCSD database. Experiments were performed for 5 randomly selected training/test combinations to report the identification rates. In order to evaluate the proposed approach, all faces were normalized, followed by histogram equalization to eliminate the lighting effects, and then face recognition experiments were conducted on the three color spaces (RGB, YcbCr, and HSV) and the grayscale image.

The initial value for χ was set as $\chi^0 = [1/3, 1/3, 1/3]$ (averaging the three color component images). After describing the iterative process for training, the algorithm generated one optimal color component combination coefficient vector $\chi^l = [x_{1l}, x_{2l}, x_{3l}]^T$ and the optimal discriminant mapping matrix. The vector χ^l determined one discriminating color component $Z^l = x_{1l}\alpha_1 + x_{2l}\alpha_2 + x_{3l}\alpha_3$.

We designed the algorithm on the grayscale images and the three popular color spaces (RGB, YCbCr and HSV), and then compared the performance of our algorithm with two still image techniques (Eigenface and Fisherface) [24], NN matching in LLE + k-means clustering [18], Mutual Subspace Method (MSM) [25], Manifold to Manifold Distance (MMD) method [26], Regularized Nearest points (RNP) [27], Collaboratively Regularized Nearest points (CRNP) [28] and Collaborative Mean Attraction (CMA) [29].

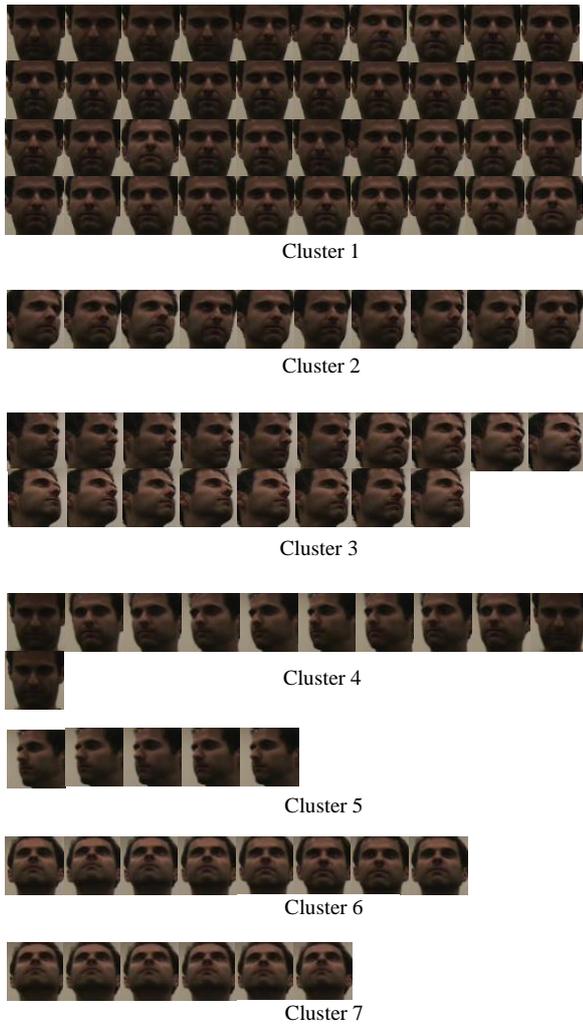


Figure 5. Clusters obtained by comparing correlation.

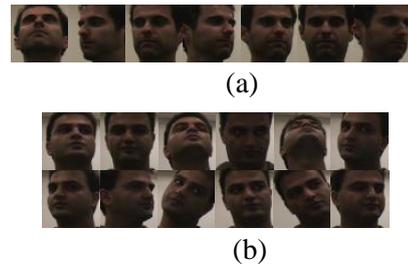


Figure 6. Center of clusters for two video sequences of Honda/UCSD database.

Table 1 shows the average recognition rate by different color space databases and the seven methods in the Honda/UCSD and CMU MoBo databases. It also shows that the proposed algorithm, in color space, achieves a better performance than the classical methods using grayscale images. Therefore, color plays an effective role in face recognition. In the best performance, the algorithm achieves an average recognition rate of 97.6% for the YCbCr color space (in Honda/UCSD database), which is a nearly 7% increase compared with the proposed method using the grayscale images.

4. Conclusions and future work

In this work, we proposed a novel method for a video-based face recognition in color space. For extracting the key frames from the video, the input sequences were converted into a number of clusters, each one acting as a linear subspace. Also to compare the center of each cluster, three popular color spaces (RGB, YCbCr and HSV) were used for the mathematical representation of each frame, and graph-based discriminant analysis was applied for the recognition process.

Table 1. Average recognition rate by different color spaces in Honda /UCSD database and CMU MoBo database.

Different methods	Average rate of recognition Honda/UCSD	Average rate of recognition CMU MoBo	Ref.
Eigenface	74.2	81.0	[24]
FisherFace	79.2	88.3	[24]
LLE+K-means	91.8	89.8	[18]
MSM	88.2	85.1	[25]
MMD	96.9	93.6	[26]
RNP	87.1	94.7	[27]
CRNP	97.4	94.4	[28]
CMA	94.9	94.6	[29]
Proposed Grayscale	90.9	88.7	Proposed
Proposed RGB	93.8	91.1	Proposed
Proposed HSV	96.4	93.8	Proposed
Proposed YCbCr	97.6	95.1	Proposed

We introduced an iterative algorithm to simultaneously determine the optimal color component combination vector and mapping matrix. Our experimental results also indicate that the recognition performance with color images is significantly better than grayscale images.

In the future works, we intend to explore new methods for computing the manifold to manifold distance in color space, which mainly emphasize the similarity of data variation modes between two manifolds, and compare them with the result of the proposed method.

References

- [1] Aggarwal, G., Chowdhury, A. R. & Chellappa R. A. (2004). system identification approach for video-based face recognition. In 17th International Conference on Pattern Recognition, vol. 4, pp. 175-178.
- [2] Lee, K. C., Ho, J., Yang, M. H. & Kriegman, D. (2003). Video-based Face Recognition Using Probabilistic Appearance Manifolds. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 313-320.
- [3] Wang, R., Shan, S., Chen, X., Dai, Q. & Gao, W. (2012). Manifold-Manifold Distance and its Application to Face Recognition With Image Sets. IEEE Trans. on Image Processing, vol. 21, no. 10, pp. 4466-4479.
- [4] Wang, R., Shan, S., Chen, X., Chen, J. & Gao, W. (2011). Maximal Linear Embedding for Dimensionality Reduction. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 33, no. 9, pp. 1776-1792.
- [5] Wang, R. & Chen, X. (2009). Manifold discriminant analysis. In CVPR, pp. 429-436.
- [6] Torres, L., Reutter, J. Y. & Lorente, L. (1999). The importance of the color information in face recognition. In Proc, IEEE Int. Conf. Image Processing, vol. 3, pp. 627-631.
- [7] Rajapakse, M., Tan J. & Rajapakse J. (2004). Color channel encoding with NMF for face recognition. In Proc. IEEE Int. Conf. Image Processing, vol. 3, pp. 2007-2010.
- [8] Yang, J., Zhang, D., Xu, Y. & Yang, J. Y. (2006). Recognize color face images using complex eigenfaces. In D. Zhang, A. K. Jain, (ed.): Advances in Biometrics, Int. Conf. Lecture Notes in Computer Science, Springer-Verlag Berlin Heidelberg, vol. 3832, pp. 64-68.
- [9] Phung, S. L., Bouzerdoum, A. & Chai, D. (2002). A novel skin color model in YCbCr color space and its application to human face detection. In IEEE International Conference on Image Processing (ICIP'2002), vol. 1, pp. 289-292.
- [10] Yang, J. & Liu, C. (2008). Color Image Discriminant Models and Algorithms for Face Recognition. IEEE Transactions on Neural Networks, vol. 19, no. 12, pp. 2088-2098.
- [11] Vadivel, A., Sural, S. & Majumdar, A. (2005). Human color perception in the HSV space and its application in histogram generation for image retrieval. In SPIE Proceedings settings, San Jose CA, United states of America, pp. 598-609.
- [12] Hsu, R. L., Abdel-Mottaleb, M. & Jain, A. K. (2002). Face detection in color images. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 24, no. 5, pp. 696-706.
- [13] Dios, J. J. de. & Garcia, N. (2003). Face detection based on a new color space YCgCr. In Proc. IEEE Int. Conf. Image Processing, vol. 3, pp. 909-912.
- [14] Wang, S. J., Yang, J., Zhang, N. & Zhou, C. G. (2011). Tensor Discriminant Color Space for Face Recognition. IEEE Transactions on Image Processing, vol. 20, no. 9, pp. 2490-2501.
- [15] Viola, P. & Jones, M. (2001). Rapid Object Detection Using a Boosted Cascade of Simple Features. Proc. Conf. Computer Vision and Pattern Recognition, vol. 1, pp. 511-518.
- [16] Wang, Y. & Yuan, B. (2001). A novel approach for human face detection from color images under complex background. Pattern Recognition, vol. 34, no. 10, pp. 1983-1992.
- [17] Fan, W. & Yeung, D. Y. (2006). Locally Linear Models on Face Appearance Manifolds with Application to Dual-Subspace Based Classification. CVPR, vol. 2, pp. 1384-1390.
- [18] Hadid, A. & Pietikäinen, M. (2004). From Still Image to Video-Based Face Recognition: An Experimental Analysis. Proceedings of the Sixth IEEE international conference on Automatic face and gesture recognition, pp. 813-818.
- [19] Kim, T. K., Arandjelović, O. & Cipolla, R. (2005). Learning over Sets Using Boosted Manifold Principal Angles (BoMPA). Proc. British Machine Vision Conf., pp. 779-788.
- [20] Yan, S., Xu, D., Zhang, B., Zhang, H. J., Yang, Q. & Lin, S. (2007). Graph embedding and extensions: A general framework for dimensionality reduction. IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 1, pp. 40-51.
- [21] Lancaster, P. & Tismenetsky M. (1985). The Theory of Matrices (Second Edition). Academic Press, INC. Orlando, Florida.
- [22] Lee, K., Ho, J., Yang, M. & Kriegman, D. (2005). Visual tracking and recognition using probabilistic appearance manifolds. Computer Vision and Image Understanding, vol. 99, no. 3, pp. 303-331.
- [23] Gross, R. & Shi, J. (2001). The cmu motion of body (mobo) database. Technical Report CMU-RI-TR-

01-18, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA.

[24] Belhumeur, P. N., Hespanha, J. P. & Kriegman, D. J. (1997). Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection. PAMI, vol. 19, no. 7, pp. 711–720.

[25] Yamaguchi, O., Fukui, K. & Maeda, K. (1998). Face Recognition Using Temporal Image Sequence. FG, pp. 318–323.

[26] Wang, R., Shan, S., Chen, X., Dai, Q. & Gao, W. (2008). Manifold-Manifold Distance with Application to Face Recognition based on Image Sets. IEEE Conference computer vision and pattern recognition. pp. 1-8.

[27] Yang, M., Zhu, P., Gool, L. V. & Zhang, L. (2013). Face recognition based on regularized nearest points between image sets. In The 10th IEEE International Conference on Automatic Face and Gesture Recognition. FG, pp. 1-7.

[28] Wu, Y., Minoh, M. & Mukunoki, M. (2013). Collaboratively Regularized Nearest Points for Set Based Recognition. Proceedings of the British Machine Vision Conference, BMVA Press, pp. 1-10.

[29] Wu, Y., Mukunoki, M. & Minoh, M. (2014). Collaborative Mean Attraction for Set Based Recognition. The 17th Meeting on Image Recognition and Understanding (MIRU).

شناسایی چهره در رشته‌های ویدیویی با استفاده از تجزیه و تحلیل جداسازی بر اساس گراف در فضای رنگی

سجاد شفیعی پور یوردشاهی^۱، میرهادی سیدعربی^۱ و علی آقاگل زاده^{۲*}

^۱ دانشکده مهندسی برق و کامپیوتر، دانشگاه تبریز، تبریز، ایران

^۲ دانشکده مهندسی برق و کامپیوتر، دانشگاه صنعتی نوشیروانی بابل، بابل، ایران

ارسال ۲۰۱۵/۰۳/۰۵؛ پذیرش ۲۰۱۶/۰۵/۱۸

چکیده:

در طول دهه گذشته شناسایی چهره در رشته‌های ویدیویی به علت کاربردهایی در موضوعاتی همچون فن‌آوری رسانه، امنیت شبکه، واسط‌های انسان با ماشین و سیستم کنترل دسترسی خودکار مورد توجه زیادی قرار گرفته است. روش‌های معمول برای شناسایی چهره، بر اساس تصاویر خاکستری بوده که با ترکیب سه رنگ اصلی تولید می‌شوند. در این مقاله، در فرآیند شناسایی، علاوه بر تصاویر خاکستری از تصاویر رنگی نیز استفاده شده است. برای استخراج فریم‌های کلیدی از رشته ویدیویی، هر رشته ورودی به تعدادی خوشه تبدیل می‌شود که هر کدام به عنوان یک زیرفضای خطی عمل می‌کند. مرکز هر خوشه به عنوان نماینده آن خوشه در نظر گرفته می‌شود. همچنین در این مقاله، برای مقایسه فریم‌های کلیدی از سه فضای رنگی متداول (RGB، YCbCr و HSV) جهت نمایش ریاضی، استفاده شده و برای انجام فرآیند شناسایی، تجزیه و تحلیل جداسازی بر اساس گراف اعمال می‌شود. با معرفی شباهت درون کلاسی و شباهت برون کلاسی در فضای رنگی نشان داده می‌شود که مشکل، به تعیین بردار ترکیب اجزای رنگی و ماتریس نگاشت تبدیل می‌شود. ما با معرفی یک الگوریتم تکراری به‌طور همزمان بردار و ماتریس بهینه مذکور را تعیین می‌کنیم. در نهایت نتایج به‌دست آمده از تصاویر سه فضای رنگی مذکور و تصاویر خاکستری با روش‌های دیگر مقایسه می‌شوند. نتایج آزمایشات کارآیی الگوریتم ارائه شده را اثبات می‌کند.

کلمات کلیدی: شناسایی چهره، فریم کلیدی، درون کلاسی، برون کلاسی، اجزای رنگ.