



### A. Preprocessing

Preprocessing mostly is necessary to facilitate further high performance recognition. In this study, we firstly convert the RGB color input image to a 256 grayscale image using Equation 1.

$$A_{GL} = \frac{3A_R + 6A_G + A_B}{10} \quad (1)$$

Where  $A_{GL}$  is the converted gray level image, and  $A_R$ ,  $A_G$ , and  $A_B$  are R, G and B spectrum of the color image respectively. Next, a vertical and horizontal Sobel operator is applied on the image to amplify the corresponding edges. Then an appropriate threshold  $T$  is used to generate the binary images where the edges are highlighted in a black background. (See Equation 2).

$$G(x, y) = \begin{cases} 1 & \text{if } A(x, y) \geq T \\ 0 & \text{other\_wise} \end{cases} \quad (2)$$

The experiments suggest that the threshold  $T$  must be from nearly 0.25 for clear, noise free, good brightness of document photo to almost 0.12 for noisy, darker document photo. The mean value of the main image is used to determine whether it is bright enough or not, while the difference between the energies of the image and its one octave low pass filtered version determines the level of noise. A dilation operator also is used to generate more continues borders and edges. Equation 3 illustrates the dilation operation, where  $A$  and  $B$  are the image and the structuring element respectively.[3]

$$A \oplus B = \{d \in R^2 : d = a + b, \forall a \in A, b \in B\} \quad (3)$$

### B. Passport Code Region Extraction

After the preprocessing stage, a morphological operator is applied to the image for specifying the passport code location. Morphology is a technique in image processing based on shapes [3, 4]. Mathematical morphology is a powerful tool for image analysis based on neighborhood operations. The mask used for the neighborhood operations is called Structuring Element (SE). An SE can have any size and geometrical shape. The shape of SE's widely used is in the form of line (generally horizontal or vertical). The basic morphological operations are erosion and dilation, which are carried out by convolving the SE with the image. Erosion is often used to remove irrelevant details from binary image and dilation is used to fill the gaps or holes. Erosion and dilation operations are usually combined and result in other two operations viz., Opening, i.e., erosion followed by dilation with the same SE and Closing i.e., dilation followed by erosion with the same SE. Opening can erase white holes on dark objects (or can remove small white objects in a dark background). An object will be erased if the SE does not fit within it. Closing removes black holes on white objects. A hole will be removed if the SE does not fit within it. These operations are illustrated in Figure 3. In erosion and dilation the SE is convolved with the image by sliding the SE over the entire image. In the process of convolution the pivotal element in the image (binary) is changed accordingly [13].

Now we can build a morphological operator that is sensitive to a specific shape in the input image. In fact we try to create a structural element with special properties. For instance, a horizontal line of length=5 pixels or a rotated one with  $\Theta=45$  can be some typical structural elements. In this

study, a rectangular box can be employed as a structural element to detect the car plates. (See Figure 2 for details).

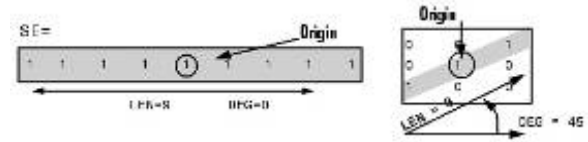


Fig. 2. Structural element for horizontal and  $\Theta=45$  rotated line

Applying the closing operation on edge image in the horizontal direction yields several connected regions which are passport code region candidates. With respect to a passport code region shape, a rectangular structural element (SE) is used. The main motivation to focus on morphology was the rectangle shape of the plates. However, there may be more than one candidate region as passport code region location in the image after closing.

### C. Segmentation

Character segmentation is an important stage in many Passport Recognition Systems. There are many factors that cause the character segmentation task difficult, such as image noise, passport watermarking and rotation and illumination variance [8, 14]. Preprocessing is very important for the good performance of character segmentation. Firstly, image is filtered to enhance the image and removing the noises and unwanted spots.

During the threshold processing on a gray scale Image, many small objects or points accrue in the threshold image due to the problem of different illumination, low quality cameras and motion effect. This kind of noise gives direct effect on segmentation and recognition process. We have used a morphological process which search the entire image for small connected elements and remove it. Then dilation operator is applied to the image for separating the characters from each other if the characters are close.

After this stage, partition scanning is proposed to extract characters from passport code region. This method is conducted by checking the mean of each partition in image (the size of partition of a character or background). In this process, the background is defined as white (1) and characters are defined as black (0). After some experiments, it is concluded that the optimum threshold value is nearly 0.7-0.8.

Partition value larger than 0.7-0.8 is considered as background, otherwise it is considered as character. Then the passport code region is divided into two blocks. The first block contains digits, and the second block contains the letter.

## III. CHARACTER RECOGNITION USING TEMPLATE MATCHING

Before employment of the recognition algorithm, the characters would be normalized. Normalization is to refine the characters into a block containing no extra white spaces(pixels) in all the borders of the characters. Then each character is fit to a 38\*20 block as shown in Figure 5. Fitting approach is also necessary for template matching. For matching the characters with the database, input images must be equalizing to a 22\*10 block with the database characters. In this system the characters are fit to 22\*10. The



in scanned passport photos. The PCA based face recognition method is not very effective under the conditions of varying pose, illumination and watermark trace, since it considers the global information of each face image and represents them with a set of weights. Under these conditions the weight vectors will vary considerably from the weight vectors of the images with normal pose and illumination, hence it is difficult to identify them correctly. On the other hand if the face images were divided into smaller regions and the weight vectors are computed for each of these regions, then the weights will be more representative of the local information of the face. When there is a variation in the pose or illumination, only some of the face regions will vary and rest of the regions will remain the same as the face regions of a normal image. Hence weights of the face regions not affected by varying pose and illumination will closely match with the weights of the same individual's face regions under normal conditions. Therefore, it is expected that improved recognition rates can be obtained by following the modular PCA approach. We expect that if the face images are divided into very small regions the global information of the face may be lost and the accuracy of this method may deteriorate. In this research, each image in the training set is divided into four smaller images. Hence the size of each sub-image will be  $L^2/4$ . These sub-images can be represented mathematically as:

$$I_{ij}(m, n) = I_i \left( \frac{L}{2}(j-1) + m, \frac{L}{2}(j-1) + n \right) \forall i, j \quad (4)$$

where  $i$  varies from 1 to  $M$ ,  $M$  being the number of images in the training set,  $j$  varies from 1 to 4 and  $m$  and  $n$  vary from 1 to  $L/2$ . Figure 4 shows the result of dividing a face image into four smaller images using Eq.(5). The average image of all the training sub-images is computed as:

$$A = \frac{1}{4M} \sum_{i=1}^M \sum_{j=1}^4 I_{ij} \quad (5)$$

The next step is to normalize each training sub-image by subtracting it from the mean as:

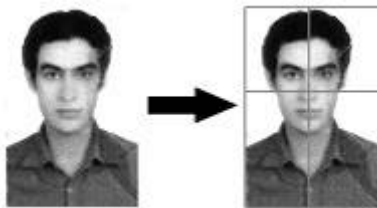


Fig. 6. A face image divided into 4 smaller images

$$Y_{ij} = I_{ij} - A \quad \forall i, j \quad (6)$$

From the normalized sub-images the covariance matrix is computed as:

$$C = \frac{1}{4M} \sum_{i=1}^M \sum_{j=1}^4 Y_{ij} \cdot Y_{ij}^T \quad (7)$$

Next we find the eigenvectors of  $C$  that are associated with the  $M_0$  largest eigenvalues. We represent the eigenvectors as  $E_1, E_2, \dots, E_{M'}$ . The weights are computed from the eigenvectors as shown below:

$$W_{pnjK} = E_K^T \cdot (I_{pnj} - A) \quad \forall p, n, j, K \quad (8)$$

where  $K$  takes the values  $1, 2, \dots, M'$ ,  $n$  varies from 1 to  $\Gamma$ ,  $\Gamma$

being the number of images per individual, and  $p$  varies from 1 to  $P$ ,  $P$  being the number of individuals in the training set. Weights are also computed for the test sub-images using the eigenvectors as shown in the next equation:

$$W_{test jK} = E_K^T \cdot (I_{test j} - A) \quad \forall j, K \quad (9)$$

Mean weight set of each class in the training set is computed from the weight sets of the class as shown below:

$$T_{pjK} = \frac{1}{\Gamma} \sum_{K=1}^{M'} \sum_{n=1}^{\Gamma} W_{pnjK} \quad \forall p, j \quad (10)$$

Next the minimum distance is computed as shown below:

$$D_{pj} = \frac{1}{M'} \sum_{K=1}^{M'} |W_{test jK} - T_{pjK}| \quad (11)$$

$$D_p = \frac{1}{4} \sum_{j=1}^4 D_{pj} \quad (12)$$

$\min(D_p) < \theta_i$  for a particular value of  $p$ , the corresponding face class in the training set is the closest one to the test image. Hence the test image is recognized as belonging to the  $p$ th face class. Figure 7 illustrates this idea graphically.

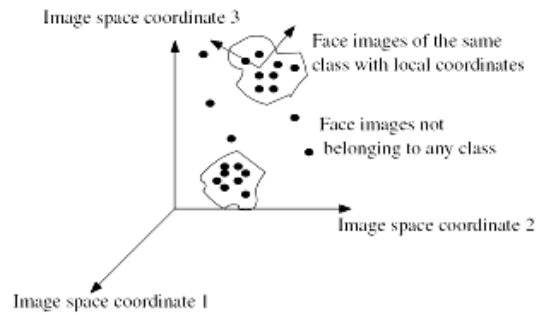


Fig. 7. The image-space and face-space coordinate system.

## VII. CONSTRUCTION OF FACE INFORMATION DATABASE

First, face images are acquired from a lot of passport images by using the extraction method of face area. The face images are learned by using MPCA algorithm, so that the metric of eigenvectors and individual feature vectors corresponding each face image are obtained, which are used in the verification of face information, and the information on two vectors is saved to database with other related user information. The construction process of face information database is shown in Fig.8

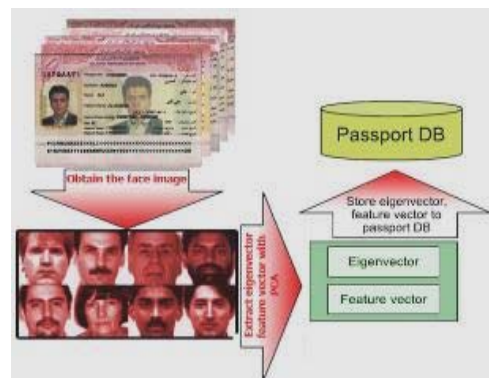


Fig. 8. Construction process of face information database.

### VIII. CONCLUSIONS

The purpose of immigration control is to find forgeries, criminal and illegal immigrants, and people prohibited from departing the country. It is difficult to use only bare eyes to distinguish and control the immigration process because the process takes a long time and the manual approach is inaccurate. To solve these problems, this paper presented a novel method of passports identification and recognition. Firstly, we extracted the passport code location, and then we separated the passport characters by segmentation, and applied a correlation-based template matching scheme for recognition of passport characters. The face is authenticated by calculating the feature vector of facial image in the input passport image using PCA algorithm and measuring the similarity between the calculated feature vector and the feature vector of facial image retrieved from the database.

This system is customized for the Iranian passports. The system is tested over a large number (more than 150) of passport images, where this algorithm performs well on Iranian passports. Finally, it is proved to be %96.8 accurate in the extraction of passport code region, and %94 accurate in the segmentation of the characters and %92 in the recognition of the characters. The accuracy of %98 in the

face detection and %97 correct in face recognition was also achieved.

### REFERENCES

- [1] K.B. Kim, "Intelligent immigration control system by using passport recognition and face verification", *Advances in Neural Networks- ISNN 2005*, New York, 2005, pp. 147-156.
- [2] K.B. Kim, S.S. Kim, S.A. Ha, "Recognition of passports using a hybrid intelligent system" *Image Analysis and Recognition*, New York, 2005, pp. 540-548.
- [3] S.H.Kasaei, S.M.Kasaei, S.A.Monadjemi "A Novel Morphological Method for Detection and Recognition of Vehicle License Plates" *American Journal of Applied Sciences*, pp.2066-2070, 2009.
- [4] Mu-Chun Su and Chao-Hsin Hung "A neural-network-based approach to detecting rectangular objects", *Neurocomputing*, pp.270-283, 2007.
- [5] R. Singh, M. Vatsa, A. Noore, and S. K. Singh. "Age transformation for improving face recognition performance". In *Pattern Recognition and Machine Intelligence (PRMI)*, November 2007.
- [6] Paul A. Viola and Michael J. Jones. Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137-154, 2004.
- [7] S. Kother Mohideen, S. Arumuga Perumal, and M. Mohamed Sathik. "Image de-noising using discrete wavelet transform" *International Journal of Computer Science and Network Security*, January 2008.
- [8] Marios Anthimopoulos, Basilis Gatos and Ioannis Pratikakis, "A two-stage scheme for text detection in video images" *Image and Vision Computing*, Vol28, Issue 9, pp.1413-1426, 2010.