Development of a fast panoramic face mosaicking and recognition system

Fan Yang
Michel Paindavoine, MEMBER SPIE
Université de Bourgogne
Laboratoire Le2i
Aile de l'Ingénieur
Mirande, BP 400
21011 Dijon cedex, France
E-mail: fanyang@u-bourgogne.fr

Hervé Abdi
The University of Texas at Dallas
Program in Cognition and Neurosciences
MS: Gr.4.1
Richardson, Texas 75083-0688

Anthony Monopoli
Université de Bourgogne
Laboratoire Le2i
Aile de l'Ingénieur
Mirande, BP 400
21011 Dijon cedex, France

Abstract. We present some development results on a system that performs mosaicking of panoramic faces. Our objective is to study the feasibility of panoramic face construction in real time. To do so, we built a simple acquisition system composed of five standard cameras, which together can take simultaneously five views of a face at different angles. Then, we chose an easily hardware-achievable algorithm, consisting of successive linear transformations, in order to compose a panoramic face from these five views. The method has been tested on a large number of faces. In order to validate our system, we also conducted a preliminary study on panoramic face recognition, based on the principal-component method. Experimental results show the feasibility and viability of our system. This allows us to envisage later a hardware implementation. We also are considering using our system for other applications, such as human expression categorization and fast 3-D face reconstruction.
 [DOI: 10.1117/1.2009707]

Subject terms: artificial vision; image mosaicking; face recognition; principal-component analysis; FFT.

Paper 040490RR received Jul. 21, 2004; revised manuscript received Feb. 4, 2005; accepted for publication Feb. 4, 2005; published online Aug. 18, 2005.

1 Introduction

Biometry is currently a very active area of research, which comprises several subdisciplines such as image processing, pattern recognition, and computer vision (see, e.g., Ref. 1). The main goal of biometry is to build systems that can identify people from some observable characteristics such as their faces, fingerprints, airises. Faces seem to have a particularly strong appeal for human users, in part because we routinely use facial information to recognize each other. Different techniques have been used to process faces such as neural network approaches,[2,3] eigenfaces,[4,5] and Markov chains.[6] As the recent DARPA-sponsored vendor test showed, most systems use frontal facial images as their input patterns.[7] As a consequence, most of these methods are sensitive to pose and lighting conditions. One way to override these limitations is to combine modalities (color, depth, 3-D facial surface, etc.).[8–12]

Most 3-D acquisition systems use professional devices such as a traveling camera or a 3-D scanner.[10,11,13] Typically, these systems require that the subject remain immobile during several seconds in order to obtain a 3-D scan, and therefore these systems may not be appropriate for some applications, such as human-expression categorization using movement estimation, or real-time applications. Also, their cost can easily make these systems prohibitive for routine applications. In order to avoid using expensive and time-intensive 3-D acquisition devices, some face recognition systems generate 3-D information from stereo vision.[14] Complex calculations, however, are needed in order to perform the required self-calibration and 2-D projective transformation.[15] Another possible approach is to derive some 3-D information from a set of face images, but without trying to reconstitute the complete 3-D structure of the face.[8,16]

For the last ten years, our laboratory has worked on face processing and obtained results for 2-D face tracking and recognition.[17–19] The goal of the present paper is to describe a system that is simple and efficient and that also can potentially process 3-D faces in real time. Our method creates panoramic face mosaics, which give some 3-D surface information. The acquisition system is composed of five cameras, which together can obtain simultaneously five different views of a given face. One of its main advantages is easy setup and very low cost.

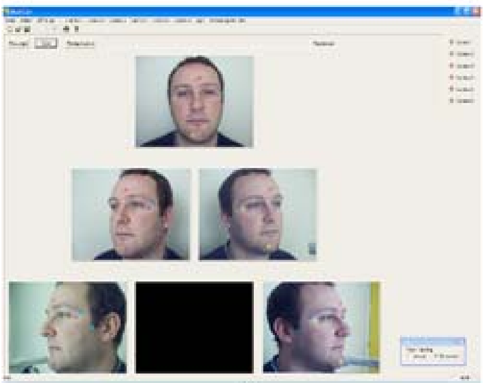
This paper is organized as follows. First, we describe our acquisition system. Then, we describe the method for creating panoramic face mosaics using successive linear transformations. Next, we present experimental results on panoramic face recognition. Finally, we conclude and explore possible follow-ups and improvements.

2 Acquisition System

Our acquisition system is composed of five Logitech 4000 USB cameras with a maximal resolution of 640×480 pixels. The parameters of each camera can be adjusted independently. Each camera is fixed on a height-adjustable sliding support in order to adapt the camera position to each individual (see Fig. 1, top). The acquisition program grabs images from the five cameras simulta-

0091-3286/2005/$22.00 © 2005 SPIE

(a)



(b)

Fig. 1 The acquisition system: the left panel shows the five cameras and their support; the right panel shows the five images collected from a subject. Each image size is 240×320 pixels.

Fig. 2 Distribution of ten markers in five views: image 1, image 2, image 3, image 4, and image 5.

neously (see Fig. 1, bottom). These five images are stored in the PC with a frame data rate of $20 \times 5 = 100$ images per second.

The human subject sits in front of the acquisition system, directly facing the central camera (camera 3). Different color markers are placed on the subject's face. These markers are used later on to define common points between different face views. The positions of these color markers correspond roughly to the face fiduciary points. Figure 2 shows the position of the chosen points. There are ten markers on each face, with at least three markers in common between each pair of face views.

With the cameras used, each pixel corresponds to an area of 1 mm^2. The area of each marker used is approximately 20 mm^2, which is sufficient for our application because it does not require high precision (a precision of 1 mm in the marker position measurement is sufficient for our purposes).

3 Panoramic Face Construction

Several panoramic image construction algorithms have been already introduced.[16,20–23] For example, Jain and Ross[23] have developed an image-mosaicking technique that constructs a more complete fingerprint template using two impressions of the same finger. In their algorithm, they initially aligned the two impressions using the corresponding minutiae points. Then, this alignment was used by a modified version of the iterative closest point (ICP) algorithm in order to compute a transformation matrix that defines the spatial relationship between the two impressions. A resulting composite image is generated using the transformation matrix, which has six independent parameters: three rotation angles (α, β, γ) about the x, y, and z axes, respectively, and three translation components (t_x, t_y, t_z) along the three axes.

For faces, Liu and Chen[16] have proposed using facial geometry in order to improve the face mosaicking result. They used a spherical projection (instead of a cylindrical projection), because it works better with the head motion in both horizontal and vertical directions. They developed a geometric matching algorithm in order to describe the correspondences between the 2-D image plane space QUV and the spherical surface space $O\alpha\beta$. In order to accomplish this, the two matching parameters $[\Delta\alpha, \Delta\beta]^T$ are found using the Levenberg-Marquardt algorithm. In order to improve the computational efficient, Liu and Chen approximated the mapping function with a triangular mesh, representing a face as a set of triangles.

In general, the methods using nonlinear transformations and iterative algorithms obtain very correct results in terms of geometric precision. However, these methods require a large number of computations and therefore cannot be easily implemented in real time. Because ultimately we want to be able to build a real-time system, we decided to use simple (and therefore fast) linear methods. Our panoramic face construction algorithm is performed in three stages (see Fig. 3):

1. marker detection and marker coordinate calculation

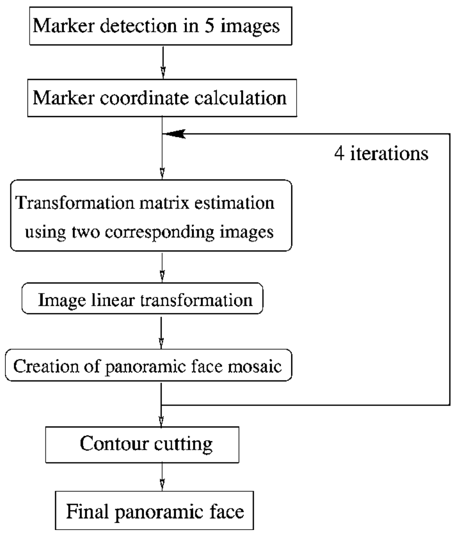


Fig. 3 Block diagram of proposed panoramic face construction algorithm.

2. transformation matrix estimation and image linear transformation
3. creation of panoramic face mosaics.

3.1 *Marker Detection and Marker Coordinate Calculation*

The first step of the algorithm corresponds to the detection of the markers put on the subject's face. The markers were made of adhesive paper (so that they would stick to the subject's face). We used three colors to create ten markers (four blue, three yellow, and three violet ones; see Fig. 2 for an illustration). These three colors were chosen after several tests because they are easily identifiable against to the face hue. These markers were used as reference points for pasting the different views of the face.

In order to detect the markers, we used color segmentation based on the hue and saturation components of each image. This procedure allows strong color selectivity and small sensitivity to luminosity variation. Figure 4 illustrates the steps used for the yellow-marker detection process. First, color segmentation gives, from the original image (Fig. 4, top left), a binary image that contains the detected markers (Fig. 4, top right). Then, in order to find the marker coordinates, we used a logical AND operation, which was performed between the binary image and a grid including white pixels separated by a fixed distance (Fig. 4, bottom left). This distance was chosen in relation to the marker area. A distance of 3 pixels allows us to capture all white zones (detected markers). Finally, we computed the centers of the detected zones (Fig. 4, bottom right). These centers gives the coordinates of the markers in the image.

3.2 *Transformation-Matrix Estimation and Image Linear Transformation*

We decided to represent each face as a mosaic. A mosaic face is a face made by concatenation of the different views pasted together as if they were on a flat surface. So, in order to create a panoramic face we combine the five different views. We start with the central view and paste the lateral views one at a time (see Fig. 3). Our method consists of transforming the image to be pasted in order to link common points between it and the target image. We obtain this transformed image by multiplying it by a linear transformation matrix. This matrix is calculated as a function of the coordinates of three common markers between the two images. $\mathbf{C}_1$ and $\mathbf{C}_2$ represent, respectively, the coordinates of the first and second images:

$$\mathbf{C}_1 = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix}, \tag{1}$$

$$\mathbf{C}_2 = \begin{bmatrix} x'_1 & x'_2 & x'_3 \\ y'_1 & y'_2 & y'_3 \end{bmatrix}. \tag{2}$$

We obtain the transformation matrix as follows:

$$\mathbf{T} = \mathbf{C}_1 \times (\mathbf{C}_2^\star)^{-1} \tag{3}$$

with

$$\mathbf{C}_2^\star = \begin{bmatrix} x'_1 & x'_2 & x'_3 \\ y'_1 & y'_2 & y'_3 \\ 1 & 1 & 1 \end{bmatrix} \tag{4}$$

and

$$\mathbf{T} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{bmatrix}. \tag{5}$$

Then, we generalize this transformation to the whole image:

$$x = a_1 x' + b_1 y' + c_1,$$

$$y = a_2 x' + b_2 y' + c_2. \tag{6}$$

This linear transformation corresponds to a combination of image rotation, image translation, and image dilation (see Fig. 5). Figure 6 displays the superposition of image 3 (not transformed) and image 4 (transformed using the coordinates of the yellow markers as common points).

3.3 *Creation of Panoramic Face Mosaics*

We begin the panoramic face construction with the central view (image 3; see Fig. 2). From the superposition of the original image 3 and transformed image 4 (see Fig. 6), we remove redundant pixels in order to obtain a temporary panoramic image 3-4 (see Fig. 7, left). In order to eliminate redundant pixels, we create a cutting line that goes through two yellow markers. This panoramic image 3-4 temporarily becomes our target image.

We repeat this operation for each view. First, image 2 is pasted on the temporary panoramic image 3-4 in order to obtain a new temporary panoramic image 2-3-4 (see Fig. 7, right). The corresponding transformation matrix is generated using three common violet markers. Then, we compute the transformation matrix that constructs image 2-3-4-5

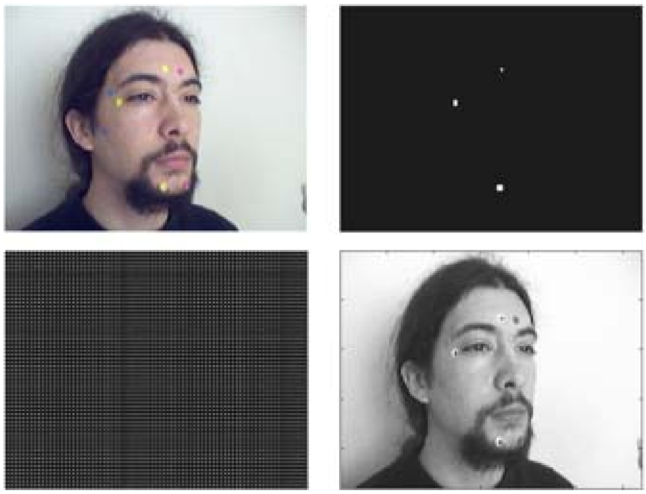


Fig. 4 Yellow-marker detection (reading left to right): original image; binary image after color filtering applied to hue and saturation components; a grid of white pixels; and marker coordinate localization.

Fig. 5 Image 4 before and after the linear transformation. The transformation matrix is computed using image 4 and image 3 (central view; see Fig. 2 and 6).

Fig. 6 Superposition of images 3 and 4: original image 3 (left), and superposition of transformed image 4 and original image 3 (right).

Fig. 7 Mosaicking results: image 3-4 (left), and image 2-3-4 (right).

Fig. 8 Mosaicking results: image 2-3-4-5 (left), and image 1-2-3-4-5 (right).

(see Fig. 8, left) using two blue markers and one yellow marker (topmost). Finally, image 1 is pasted to the temporary panoramic image 2-3-4-5 with the help of two blue markers and one violet marker (topmost) (see Fig. 8 right).

Figure 9 (left) displays the final panoramic face composition from five views. This composition preserves some of the face shape. For example, the chin of a human face possesses more curvature than other parts; therefore the bottom part of the panoramic face is composed of five views: 1, 2, 3, 4, and 5. On the other hand, three views (1, 3, and 5) suffice to compose the top part.

Figure 9 right shows the final mosaic face obtained after automatic contour cutting. For this, we first surround the panoramic face with a circle that passes by the extreme points of the ears in order to eliminate the background. Then, we replace segments of this circle by polynomial curves using extreme-point coordinates located with the help of the marker positions.

Note that these ten markers allow us to link common points between five views. The coordinates of the markers are computed in the marker detection process (see Sec. 3.1), and arranged in a table. Then, all ten markers are erased from all five views, using a simple image-processing technique (local smoothing). This table of marker coordinates is regenerated for each temporary panoramic image construction. The goal of marker elimination is to use panoramic faces for face recognition or 3-D face reconstruction.

As compared to the method proposed by Liu and Chen,[16] panoramic faces obtained using our model are less precise in geometry. For example, Liu and Chen used a triangle mesh in order to represent a face. Each triangle possesses its own transformation parameters. In our system, a single transformation matrix is generated for a complete image. Liu and Chen have also established a statistical

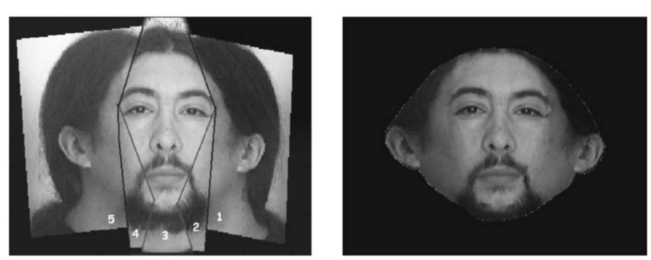


Fig. 9 Mosaicking results: panoramic face composition from five views (left), and final mosaic (right).

Fig. 10 A sampling of panoramic faces from the first session's database.

modeling containing the mean image and a number of "eigenimages" in order to represent the face mosaic.

Our objective is to study an efficient and simple algorithm for later hardware implantations. Methods necessitating a large calculation volume and a large memory space are not adapted to embedded systems. In order to test and validate our panoramic face mosaicking algorithm, we propose, in the next sections, a study of face recognition based on the *eigenface* model proposed by Turk and Pentland.[4]

We created a panoramic face database composed of 12 persons $\times$ 4 expressions $\times$ 2 sessions$=$96 panoramic faces. The two acquisition sessions were performed over an interval of one month. The four expressions were: neutral, smile, deepened eyebrows, and eyes closed (see Fig. 10). We implemented a face recognition procedure using this database.

4 Face Recognition Description: PCA

Over the past 25 years, several face recognition techniques have been proposed, motivated by the increasing number of real-world applications and also by the interest in modeling human cognition. One of the most versatile approaches is derived from the statistical technique called principal-component analysis (PCA) adapted to face images.[24] Such a approach has been used, for example, by Abdi[25] and Turk and Pentland[4] for face detection and identification. PCA is based on the idea that face recognition can be accomplished with a small set of features that best approximates the set of known facial images. Application of PCA for face recognition proceeds by first performing a PCA on a well-defined set of images of known human faces (every person is represented by a number of different images, in which various expressions are captured). From this analysis, a set of K principal components is obtained, and the projection of the new faces on these components is used to compute distances between new faces and old faces. These distances, in turn, are used to make predictions about the new faces (i.e., are these new or old faces?).

Technically, PCA on face images proceeds as follows. The K face images to be learned are represented by K vectors $\mathbf{a}_k$, where k is the image number.[26] Each vector $\mathbf{a}_k$ is obtained by concatenating the rows of the matrix storing the pixel values (here, gray levels) of the k'th face image. This operation is performed using the vec operation, which transforms a matrix into a vector (see Ref. 27 for more details).

The complete set of patterns is represented by a $I \times K$ matrix noted $\mathbf{A}$, where I represents the number of pixels of the face images and K the total number of images under consideration. Specifically, the learned matrix $\mathbf{A}$ can be expressed as

$$\mathbf{A} = \mathbf{P}\boldsymbol{\Delta}\mathbf{Q}^T, \tag{7}$$

where $\mathbf{P}$ is the matrix of eigenvectors of $\mathbf{AA}^T$, $\mathbf{Q}$ is the matrix of eigenvectors of $\mathbf{A}^T\mathbf{A}$, and $\boldsymbol{\Delta}$ is the diagonal matrix of singular values of $\mathbf{A}$, that is, $\boldsymbol{\Delta} = \boldsymbol{\Lambda}^{1/2}$, with $\boldsymbol{\Lambda}$, the matrix of eigenvalues of $\mathbf{AA}^T$ and $\mathbf{A}^T\mathbf{A}$. The left singular eigenvectors $\mathbf{P}$ can be rearranged in order to be displayed as images. In general, these images are somewhat facelike,[25] and they are often called *eigenfaces*.[4]

Given the singular vectors $\mathbf{P}$, every face in the database can be represented as a weight vector in the principal-component space. The weights are obtained by projecting the face image onto the left singular vectors, and this is achieved by a simple inner product operation:

$$\mathrm{PROJ}_x = \mathbf{x}^T\mathbf{P}\boldsymbol{\Delta}^{-1}, \tag{8}$$

where $\mathbf{x}$ is a facial vector, corresponding to an example face in the training process or a test face in the recognition process. Therefore, when a new test image whose identification is required is given, its vector of weights also represents the new image.

Identification of the test image is done by locating the image in the known face database whose weights have the smallest Euclidean distance from the weight of the test image. This algorithm, employed by Turk and Pentland,[4] is called the *nearest neighbor classification rule*.

Figure 11 displays the projections of the learned faces stored in the matrix $\mathbf{A}$ on principal components 2 and 3. For this illustration, 48 panoramic faces were analyzed. Each panoramic face is labelled by its identity (12 persons, numbered 1 to 12), its expression (four expressions, coded N for neutral, S for smile, R for deepened, and F for eyes closed), and its acquisition session (two sessions, a and b). Here only the first session (session a) is displayed.

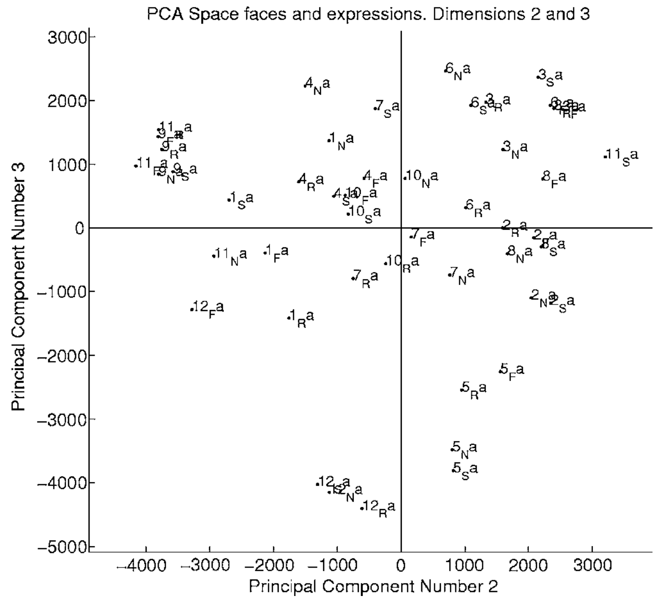


Fig. 11 2-D projection representation of learned matrix **A** on principal components 2 and 3; each panoramic face is differentiated by its identity (number 1 to 12), its expression (N,S,R,F), and its acquisition session (a, b).

5 Experimental Results on Panoramic Face Recognition

5.1 *Spatial Representation*

For these first tests, panoramic faces were analyzed using the original 240×320-pixel image (spatial representation) without preprocessing. The database consisted of 12 persons $\times$ 4 expressions $\times$ 2 sessions $= 96$ panoramic faces, and was divided into two subsets. One subset served as the training set, and the other subset as the testing set. As illustrated in Fig. 10, all these panoramic faces possess a uniform background, and the ambient lighting varied according to the daylight.

From the panoramic face database, one, two, three, or four images were randomly chosen for each individual in order to create the training set (number of patterns for learning per individual, $p = 1$, 2, 3, 4). The rest of the pan-

Table 1 Results of panoramic face recognition with spatial representation. The number of eigenvectors used corresponds to the mean values obtained with the discriminant analysis algorithm during several executions.

No. of training examples per individual, p	Total no. of training examples	No. of eigenvectors used	No. of tests for recognition	Recognition rate (%)
1	12	9	84	70
2	24	13	72	85.08
3	36	18	60	90.1
4	48	25	48	93.21

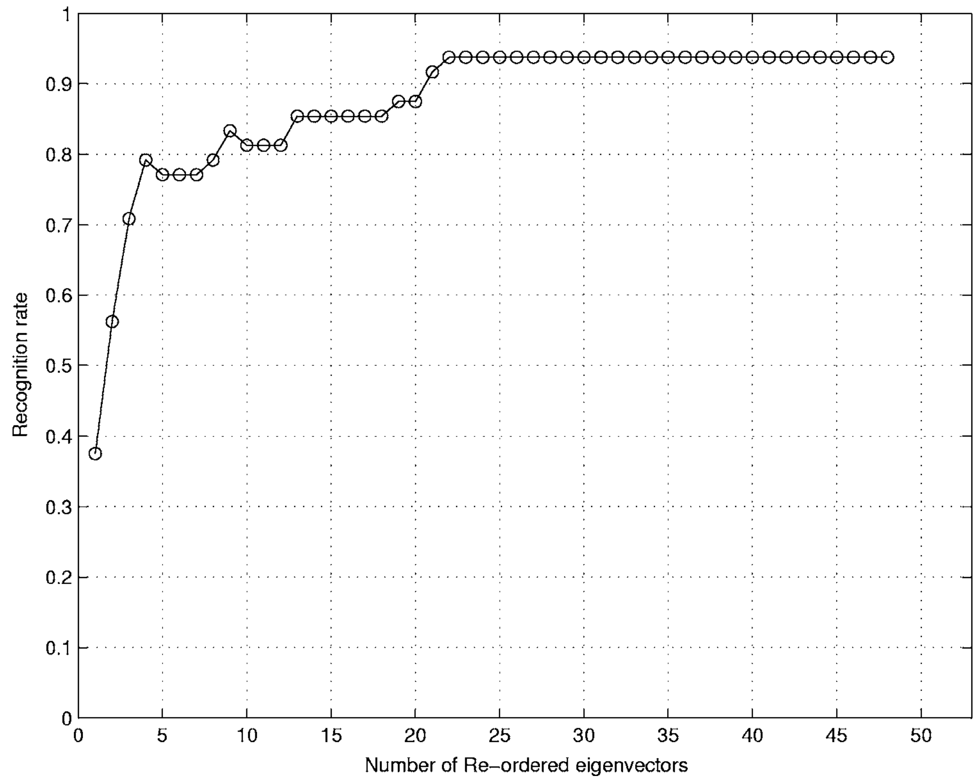


Fig. 12 Recognition rate as a function of the number of reorganized eigenvectors.

oramic faces were used in order to test the face recognition method. For example, when $p=1$, the total number of training examples is equal to 1×12 persons$=12$, and the number of test samples for recognition is equal to $96-12=84$. Therefore, for each individual, only one panoramic face is learned in order to recognize seven other images of this person. Several executions of our MATLAB program were run for each value of p, using randomly chosen training and testing sets. Then we computed the mean performance. The results are presented in Table 1.

Using the nearest neighbor classification rule, the panoramic face identity test is done by locating the closest image in the known face database. Therefore, the system can make only confusion errors (i.e., associating the face of one person with a test face of another). In Table 1, the recognition rate corresponds to correct panoramic face recognition rate.

We added a discriminant analysis stage in the face recognition process so as to determine the number of necessary eigenvectors. This analysis, called the *jackknife*,[28] reorders eigenvectors, not according to their eigenvalues, but according to their importance for identification. Specifically, we computed the ratio of the between-group inertia to the within-group inertia for each eigenvector. This ratio expresses the quality of the separation of the identity of the subject performed by this eigenvector (a similar approach was taken by O'Toole et al.[29]). The eigenvector with the largest ratio performs the best identity separation, the ei-

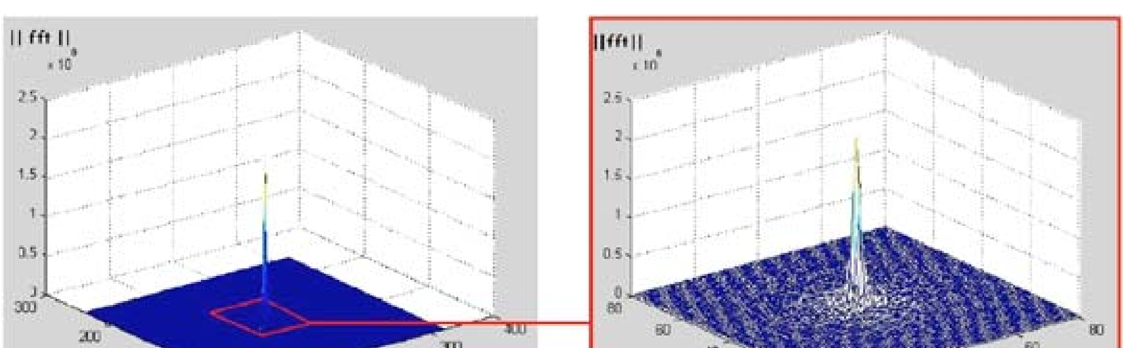


Fig. 13 3-D representation of original FFT amplitude (left), and lowpass-filtered FFT amplitude of only $80\times 80=6400$ values, which are presented to the recognition system (right).

Table 2 Results of panoramic face recognition with frequential representation.

No. of training examples per individual, *p*	Total no. of training examples	No. of eigenvectors used	No. of tests for recognition	Recognition rate (%)
1	12	8	84	76.83
2	24	13	72	91.26
3	36	18	60	93.25
4	48	24	48	97.46

genvector with the second largest ratio performs second best, etc.

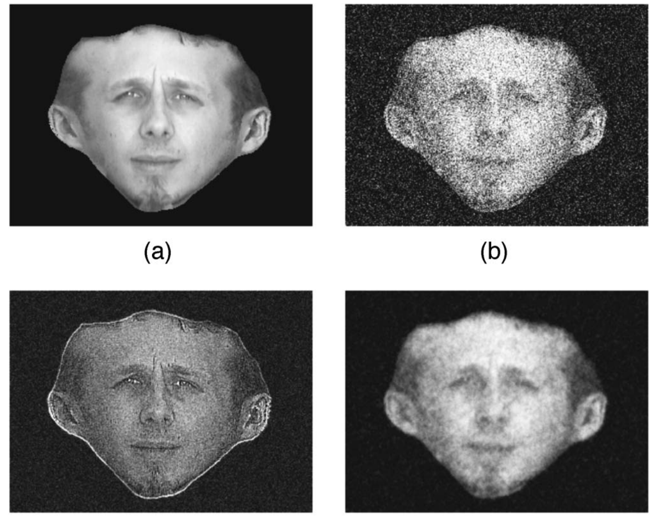
For the example shown in Fig. 12, we observe that it suffices to use only 23 eigenvectors to reach the maximum recognition rate (93.75%). Additional eigenvectors do not add to the quality of the identification.

5.2 *Frequential Representation*

We also tested the frequential behavior of our recognition system. Figure 13 (left) displays the FFT amplitude of a panoramic face. We can observe that the spectra are well centered at low frequencies. This allows us to apply a low-pass filter in order to reduce the size of the data set to process (see Fig. 13, right). Only 80×80 FFT amplitude values of low frequencies were used for the recognition system.

We applied the same training and testing process as used in spatial representation. Test results are given in Table 2. We obtain a better recognition rate with the frequential representation (97.46%) than with the spatial representation (93.21%). This advantage of the frequential representation is due to the fact that for face images, the spectrum amplitude is less sensitive to noise (or variations) than the spectrum phase. We confirmed this interpretation by using a panoramic face image to which noise was added. Figure 14(a) shows a original panoramic face. Figure 14(b) displays the same panoramic face image with added noise. We first obtained the FFTs of these two images and then their inverse FFTs in the two following manners:

1. Using the spectrum amplitude of the noised image and the spectrum phase of the original image [see Fig. 14(c)]



(a) (b) (c) (d)

Fig. 14 Amplitude is less sensitive to noise than phase is. (a) An original panoramic face image. (b) Original image with added Gaussian noise (mean 0 and variance 0.05). (c) IFFT image using the spectrum amplitude of (b) and the spectrum phase of (a). (d) IFFT image using the spectrum amplitude of (a) and the spectrum phase of (b). The image (c) is more similar to the image (a) than the image (d) is.

2. Using the spectrum phase of the noised image and the spectrum amplitude of the original image [see Fig. 14(d)].

These results show that the face obtained with the first configuration is closer to the original face than the face obtained with the second configuration. This confirms that the spectrum amplitude is less sensitive to noise than the spectrum phase.

Tsalakanidou et al.[8] described a study designed to evaluate three different approaches (color, depth, and combination of color and depth) for face recognition and quantity the contribution of depth. The color images were stored in portable pixmap format (ppm) with a resolution of 720 $\times 576$ pixels. In addition, for each person the *structured light* approach was used for capturing the 3-D facial surface and thus creating the depth map. Their experimental results show significant gains of 5% with the use of depth information. They have obtained a recognition rate of 97.5% using the depth map and color frontal view. The face recognition technique used is based on the implementation of the PCA algorithm (similar to ours). By using a simple and easily built system, we obtain some 3-D face information.

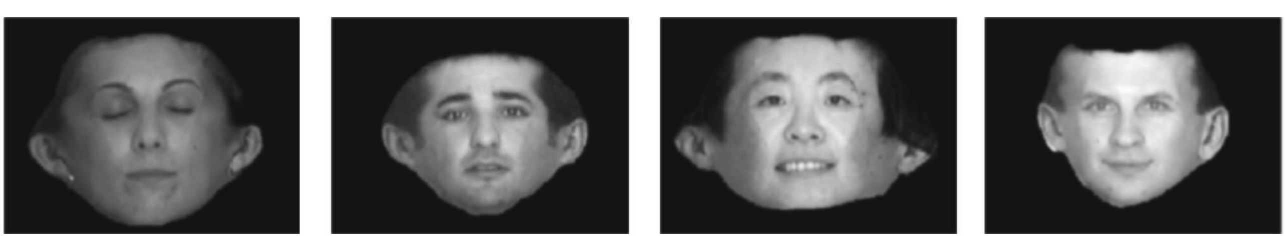


Fig. 15 Panoramic faces of four unknown persons.

Table 3 Results of panoramic face recognition with negative samples. These performance results were obtained using the frequential representation. Performance declined in comparison with tests without negative samples.

No. of training examples per individual, *p*	Total no. of training examples	No. of eigenvectors used	No. of tests for recognition	Non-recognition rate (%)	Confusion rate (%)	Recognition rate (%)
1	12	8	116	25.4	5.85	68.75
2	24	13	104	12.74	4	83.26
3	36	18	92	7.58	3.5	88.92
4	48	24	80	4.82	2.8	92.38

The obtained performance of panoramic face recognition is very close to that of Tsalakanidou et al.

5.3 *Panoramic Face Recognition with Negative Samples*

In order to evaluate the behavior of our system for unknown people, we added four people to the test database (see Fig. 15). These panoramic faces were obtained as described in Sec. 3.

Table 3 displays the performance of different tests. We added $4\ \text{persons} \times 4\ \text{expressions} \times 2\ \text{sessions} = 32$ panoramic faces in each test set. In order to reject these unknown faces, we established a threshold of Euclidean distance. Because we are working on applications of typical access control, where confusion is more harmful than nonrecognition, we decided to use a severe acceptance threshold in order to reject intruders. Note that the acceptance threshold is constant for all tests. Efficiency is defined as follows:

- *Recognition:* Correct recognition of a panoramic face.
- *Nonrecognition:* A panoramic face has not been recognized.
- *Confusion:* A panoramic face is confused with an intruder.

6 Conclusions and Perspectives

In this paper, we have proposed a fast and simple method for panoramic face mosaicking. The acquisition system consists of several cameras followed by a series of fast linear transformations of the images. The simplicity of the computations makes it possible to envisage real-time applications.

In order to test the recognition performance of our system, we used the panoramic faces as input to a recognition system based on PCA. We tested two panoramic face representations: spatial and frequential. We found that a frequential representation gives the better performance, with a correct recognition rate of 97.46%, versus 93.21% for spatial representation. An additional advantage of the frequential representation is that it reduces the data volume to be processed and this further accelerates the calculation speed. We used negative samples for the panoramic face recognition system, and the correct recognition rate was 92.38%. Experimental results show that our fast mosaicking system provides relevant 3-D facial surface information for recognition application. The obtained performance is very close or superior to published levels.[2,3,6,8]

In the future, we plan to simplify our acquisition system by replacing the markers with a structured light. We also hope to use our system without markers. For this, we will detect control points on faces (corners, points of maximum curvature, etc.). Another line of development is to improve the geometry quality of our panoramic face mosaic construction.[16,30] For this, we will use realistic human face models. We are also exploring processing panoramic face recognition using other classifiers with more variable conditions. Then, 3-D face applications are envisaged, such as real-time human expression categorization using movement estimation and fast 3-D facial modeling for compression and synthesis, as in videoconferencing.

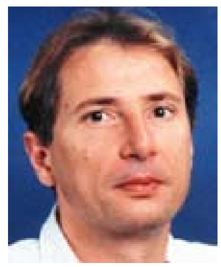
References

1. S. Y. Kung, M. W. Mak, and S. H. Lin, *Biometric Authentification: A Machine Learning Approach*, Prentice-Hall, Upper Saddle River, NJ (2005).
2. A. J. Howell and H. Buxton, “Learning identity with radial basis function networks,” *Neurocomputing* **20**, 15–34 (1998).
3. T. Sim, R. Sukthankar, M. Mullen, and S. Baluja “Memory-based face recognition for visitor identification,” in *Proc. 4th IEEE Int. Conf. on Automatic Face and Gesture Recognition* (2000).
4. M. Turk and A. Pentland, “Eigenfaces for recognition,” *J. Cogn Neurosci.* **3**, 71–86 (1991).
5. H. Abdi, D. Valentin, and A. O’Toole, “A generalized auto-associator model for face semantic process,” in *Optimization and Neural Networks*, D. Levine, Ed., Erlbaum, Hillsdale NJ (1997).
6. M. Slimane, T. Brouard, G. Venturini, and J. P. Asselin de Beauville- “Unsupervised learning of pictures by genetic hybridization of hidden Markov chain,” *Signal Process.* **16** (6), 461–475 (1999).
7. P. J. Phillips, P. Grother, R. J. Micheals, D. M. Blackburn, E. Tabassi, and J. M. Bone “Face recognition vendor test 2002,” presented at *IEEE International Workshop on Analysis and Modeling of Faces and Gestures (AMFG)* (2003).
8. F. Tsalakanidou, D. Tzovaras, and M. G. Strintzis, “Use of depth and colour eigenfaces for face recognition,” *Pattern Recogn. Lett.* **24**, 1427–1435 (2003).
9. C. Beumier and M. Acheroy, “Face verification from 3D and grey level clues,” *Pattern Recogn. Lett.* **22**, 1321–1329 (2001).
10. C. Hehser, A. Srivastava, and G. Erlebacher, “A novel technique for face recognition using range imaging,” in *Proc. 7th Int. Symp. on Signal Processing and Its Applications (ISSPA)*, IEEE (2003).
11. X. Lu, D. Colbry, and A. K. Jain, “Three-dimensional model based face recognition,” in *Proc. Int. Conf. on Pattern Recognition*, pp. 362–366, IEEE (2004).
12. K. W. Bowyer, K. Chang, and P. Flynn, “A survey of 3D and multi-

modal 3D+2D face recognition," in *Proc. Int. Conf. on Pattern Recognition (ICPR)*, IEEE (2004).
13. V. Blanz and T. Vetter, "Face recognition based on fitting a 3D morphable model," *IEEE Trans. Pattern Anal. Mach. Intell.* **25**, 1063–1074 (Sep. 2003).
14. J. G. Wang, R. Venkateswarlu, and E. T. Lim, "Face tracking and recognition from stereo sequence," *Comput. Sci.* **2688**, 145–153 (2003).
15. R. Hartly and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed., Cambridge Univ. Press (2003).
16. X. Liu and T. Chen, "Geometry-assisted statistical modeling for face mosaicing," in *IEEE Int. Conf. on Image Processing (ICIP)*, Vol. **2**, pp.883–886 (2003).
17. F. Yang, M. Paindavoine, and H. Abdi, "A new filtering technique combining a wavelet transform with a linear neural network: application to face recognition," *Opt. Eng.* (11) **39**, 2894–2899 (2000).
18. J. Mitéran, J. P. Zimmer, F. Yang, and M. Paindavoine, "Access control: adaptation and real time implantation of a face recognition method," *Opt. Eng.* (4) **40**, 586–593 (2001).
19. F. Yang and M. Paindavoine, "Implementation of a RBF neural network on embedded systems: real time face tracking and identity verification," *IEEE Trans. Neural Netw.* **14**(5), 1162–1175 (2003).
20. Y. Kanazawa and K. Kanatani, "Image mosaicing by stratified matching," *Image Vis. Comput.* **22**, 93–103 (2004).
21. Y. Zhou, H. Xue, and M. Wan, "Inverse image alignment method for image mosaicing and video stabilization in fundus indocyanine green angiography under confocal scanning laser ophthalmoscope," *Comput. Med. Imaging Graph.* **27**, 513–523 (2003).
22. P. F. McLauchlan and A. Jaenicke, "Image mosaicking using sequential bundle adjustment," *Image Vis. Comput.* **20**, 751–759 (2002).
23. A. K. Jain and A. Ross, "Fingerprint mosaicking," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)* (2002).
24. D. Valentin, H. Abdi, A. J. O'Toole, and G. W. Cottrell, "Connectionist models of face processing: a survey," *Pattern Recogn.* **27**, 1208–1230 (1994).
25. H. Abdi, "A generalized approach for connectionist auto-associative memories: interpretation, implications and illustration for face processing, in *Artificial Intelligence and Cognitive Sciences*, J. Demongeot (Ed.), Manchester Univ. Press, Manchester (1988).
26. H. Abdi, D. Valentin, and B. Edelman, *Neural Networks*, Sage, Thousand Oaks, CA (1999).
27. H. Abdi, D. Valentin, B. E. Edelman, and A. J. O'Toole, "More about the difference between men and women: evidence from linear neural networks and the principal component approach," *Perception* **24**, 539–562 (1995).
28. M. C. K. Yang and D. H. Robinson, *Understanding and Learning Statistics by Computer*, World Scientific, Singapore (1986).
29. A. J. O'Toole, F. Jiang, H. Abdi, and J. V. Haxby, "Partially distributed representations of objects and faces in ventral temporal cortex," *J. Cogn Neurosci.* **17** (in press).
30. W. Puech, A. G. Bors, I. Pitas, and J.-M. Chassery, "Projection distortion analysis for flattened image mosaicking from straight uniform generalized cylinders," *Pattern Recogn.* **34**, 1657–1670 (2001).

Fan Yang received her BS degree in electrical engineering from the University of Lanzhou (China) in 1982. She was a scientific assistant in the Department of Electronics at the University of Lanzhou. She received her MS (D.E.A.) degree in computer science and her PhD degree in image processing from the University of Burgundy (France) in 1994 and in 1998, respectively. She is currently a full *maître de conférences* and member of LE2I CNRS-UMR (Laboratory of Electronic, Computing and Imaging Sciences). Her research interests are in the areas of pattern recognition, neural network, motion estimation based on spatiotemporal Gabor filters, parallelism, and real-time implementation, and more specifically, automatic face image processing (algorithms and architectures).

Michel Paindavoine received his PhD in electronics and signal processing from Montpellier University, France, in 1982. He was with Fairchild CCD Company for two years as an engineer specializing in CCD sensors. He joined Burgundy University in 1985 as *maître de conférences* and is currently full professor and director of LE2I UMR-CNRS, Laboratory of Electronic, Computing and Imaging Sciences, Burgundy University, France. His main research topics are image acquisition and real-time image processing. He is also a member of ISIS (a research group in signal and image processing of the French National Scientific Research Committee).



Hervé Abdi received his MS in psychology from the University of Franche-conté (France) in 1975, his MS (D.E.A.) in economics from the University of Clermond-Ferrand (France) in 1976, his MS (D.E.A.) in neurology from the University Louis Pasteur in Strasbourg (France) in 1977, and his PhD in mathematical psychology from the University of Aix-en-Provence (France) in 1980. He was an assistant professor at the University of Franche-Comté (France) in 1979, an associate professor at the University of Bourgogne at Dijon (France) in 1983, and a full professor at the University of Bourgogne at Dijon (France) in 1988. He is currently a full professor in the School of Human Development at the University of Texas at Dallas. He was a visiting associate professor of cognitive and linguistic sciences at Brown University in 1986 and 1987, and a Fulbright scholar. His interests include neural network and cognitive modeling, experimental design, and statistics.

Anthony Monopoli: Biography and photograph not available.