MMAR 10th IEEE International Conference on Methods and Models in Automation and Robotics
2004 30 August - 2 September 2004, Międzyzdroje, Poland

FACE RECOGNITION SYSTEM FOR AN ASSISTANT ROBOT

L.M. BERGASA, J.A. HERNÁNDEZ, R. BAREA, M. E. LÓPEZ, M. ESCUDERO

Departamento de Electrónica. Universidad de Alcalá
28805 Alcalá de Henares, Madrid, SPAIN
email: bergasa@depeca.uah.es

Abstract. In this paper we present a new face recognition system onboard a mobile robot for assisting the elderly. As difference with almost all the existing face recognition systems, in our case, the face images database and the recognition images are obtained from a robot moving. The robot localizes a user in a room and it moves toward him in order to recognize his face. Face images are cropped and resized automatically. We have used 2DPCA (Two Dimensional Principal Component Analysis) as face recognition techniques. We have evaluated the performance of this technique in our system using several real experiments. Then, we have compared results obtained with 2DPCA technique and the traditional PCA method. Experimental results, conclusions and future works have finally presented.

Key Words. Face recognition, PCA, 2DPCA, robotic assistance.

1. INTRODUCTION

In the last years, the number of elderly in need of care is increasing dramatically. In the European Union, about 15% of the total population is over 65 years old. As the baby-boomer generation approaches the retirement age, this number will increase significantly. By 2030, more than 25% of the population will be 65 and over. This increase could collapse the Hospitals and assistance centers. One alternative to this problem could be using assistance at home, but current nursing home costs range between 30,000 € and 60,000 € annually. The dramatic increase of costs poses extreme challenges to society. We need to find new technologies and alternative ways of providing care to this sector of the population, where the need of personal assistance is larger than in any other age group. Aware of this necessity, nowadays there are several projects and research groups working on the development of assistant robots. Among them we can find the "Nursebot project", with robots *Flo* [1] and *Pearl* [2], "I.L.S.A" [3] and "Morpha" [4] projects.

In order to contribute to this research field, the Electronics Department of the University of Alcalá is working on the SIRAPEM project (Spanish acronym of *Robotic System for Elderly Assistance*). The general goal of this project is the development of a robotic assistant (called SIRA) which allows the user to be completely monitored 24 hours a day and tele-diagnosed from the assistance centers. Therefore we guarantee the security of the user and decrease the health-care cost because the care givers only would go to the patient's house if it were strictly necessary [5].

This paper is focused in one of the high level services of SIRA called "automatic monitoring and recognition of the patient" and mainly in the second part of this. The goal of this service is to localize a user in a room, recognize him and monitor his state, sending an alarm when some risk situations are detected (loss of stability or fading).

Respecting the recognition method used, PCA has been widely investigated and has become one of the most successful approaches in face recognition. Recently, a new method called 2DPCA has been developed [6]. This last method has two important advantages over PCA. Firstly, it is easier to evaluate the covariance matrix accurately. Secondly, less time is required to determine the corresponding eigenvectors. In other hands, effectiveness and robustness of 2DPCA is higher to other techniques as *Fisherfaces*, *ICA* (Independent Component Analysis) and their corresponding non-linear techniques, *Kernel PCA and Kernel Fisherfaces*, as it has been shown in [6] using three face images database: ORL, AR and Yale. This is the reason because we have applied the 2DPCA methodology in our recognition system.

In this paper we evaluate performance of 2DPCA method onboard a robot moving and we compare its results with the obtained with the traditional PCA method, in order to show its advantages over the reference method (PCA). In our case, data base generation and recognition stage are developed from the robot moving and images are automatically cropped and resized using a previous algorithm for localizing the user's face. The remainder of this paper is organized as follows: In section 2, the general architecture of SIRA robot is presented. *Automatic monitoring* high level service is briefly described in section 3. In section 4, face recognition system is explained. Experimental results and performance evaluation of our system is presented in section 5. Finally, conclusions and future works are presented in section 6.

2. GENERAL ARCHITECTURE OF SIRA ROBOT

SIRA is based on a commercial platform (the PeopleBot robot of ActivMedia Robotics [7]) with a differential drive mobile base. Its architecture is composed of four main modules: environment perception, navigation, human-machine interface and high-level services. The first module is endowed with encoders, bumpers, two sonar rings (high and low) and a vision system based on a PTZ (pan-tilt-zoom) color camera connected to a frame grabber. The navigation system combines information from the different sensors for global navigation using Partially Observable Markov Decision Processes (POMDPs). This module is controlled by the high-level services. The human-machine interface is composed of speakers, microphone, a tactile screen, the same PTZ camera used in the perception module, and wireless Ethernet link. The system architecture includes two human-machine interaction systems, voice (synthesis and recognition speech) and touch screen for simple command selection (for example, a destination room to which the robot must go to carry out a service task). The high-level services module controls the rest of the systems and includes several tasks of tele-assistance, tele-monitoring, providing reminding and social interaction. SIRA is equipped with two on-board PCs for supporting all its architecture and batteries with a lifetime of 4 hours approximately.

3. AUTOMATIC MONITORING

Though the vast majority of older adults live in the community, many reside with similarly frail relatives, or live alone with little or no outside support. Family members are often widely dispersed and minimally involved in meeting the day to day needs of their elders. In home services from community agencies are generally time-limited and prohibitively expensive for many older adults. These circumstances can pose substantial risks, for example, loss of stability or fading can cause the user falls on the floor. This can have, if undetected by others, severe consequences up to the patient's death. By reducing these risks, we are implementing a system which is able to answer to the call of the user by monitoring his state and sending an alarm if a danger situation for him is detected. Then, the robot achieves an automatic monitoring periodically, localizing the user and detecting his situation.

We have designed a system for localizing, identifying and tracking the patient's face based on computer vision. These abilities are essential for finding a user and for being able to interact with a person while he is moving. Our system follows the scheme shown in figure 1.

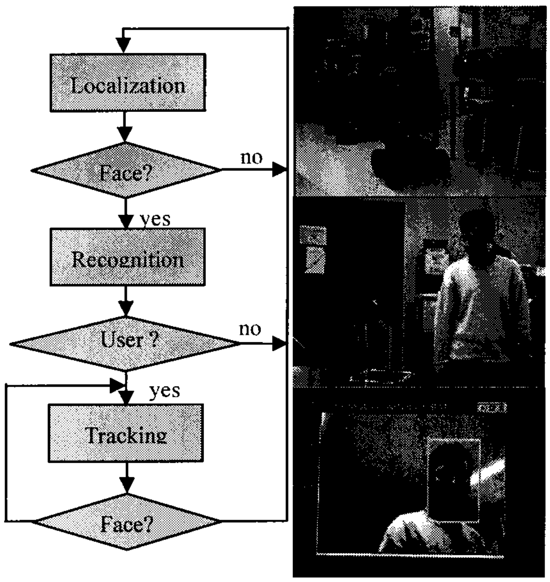

Fig.1. Automatic monitoring system

Firstly, the camera and the robot are moved looking for human faces in a room. The robot navigates towards free zones, using reactive behavior in order to avoid obstacles, until a face is found. Our face localization system is based on a detector of simple features called Haar-like and a classifier based on the learning algorithm AdaBoost [8]. This method generates an image representation called integral image that allows for very fast Haar-like feature evaluation (e.g. dark zones around bright areas in horizontal or vertical). Once the integral image has been computed, any one of these Haar-like features can be computed at any scale or location in constant time. AdaBoost algorithm selects a small number of critical visual features from a larger set and yields extremely efficient classifiers. Once a face has been found, a tracking algorithm based on the explained method applied in a window placed in the last detected frame, tracks the face. While a person's face is tracked, the camera and the robot are continually adjusted to keep the person centered in the camera image, with a predefined size, and the robot aligned with the user. Once the tracking system has centered the user, a face recognition algorithm is run in order to known if the detected person is the wanted person or not. If the answer is true the face tracking is enabled and the robot tracks the user, stopping when he is enough closer to the user for establishing a tele-conference. If the answer is negative the system continues looking for another face in the room.

4. FACE RECOGNITION SYSTEM

4.1 2DPCA Algorithm

Let **X** denote an n-dimensional unitary column vector. The idea is to project image **A**, an m x n random matrix, onto **X** by the following linear transformation:

$$Y = AX \tag{1}$$

Thus, an m-dimensional projected vector **Y** is obtained. This is called the projected feature vector of image **A**. How is determined a good projection vector **X**? Let $\mathbf{A}_j$ (j=1,2,...,M) denote the set of sample images (*m x n*) of a group of users. The total scatter of the projected sample images is a good measure of the discriminatory power of the projected vectors and can be evaluated by the trace of the covariance matrix ($\mathbf{S}_x$) of the feature vectors. The projection vectors are optimal when the total scatter of the projected samples is maximized [6].

$$J(X) = tr(S_x) = X^T G_t X \tag{2}$$

The optimal projection vectors are the set of eigenvectors of $\mathbf{G}_t$, the *image covariance (scatter) matrix* of the sample images. $\mathbf{G}_t$ is an *n x n* nonnegative definite matrix which can be evaluated by:

$$G_t = \frac{1}{M}\sum_{j=1}^{M}\left(A_j - \overline{A}\right)^T\left(A_j - \overline{A}\right) \tag{3}$$

where $\overline{A}$ is the average image of the training samples:

$$\overline{A} = \frac{1}{M}\sum_{j=1}^{M} A_j \tag{4}$$

In general, it is enough to select the first *d* orthonormal eigenvectors corresponding to the first d largest eigenvalues {$\mathbf{X}_1$, $\mathbf{X}_2$,..., $\mathbf{X}_d$}, in order to characterize the whole image.

The optimal projection vectors of 2DPCA, {$\mathbf{X}_1$, $\mathbf{X}_2$,..., $\mathbf{X}_d$}, are use for feature extraction. For a given image sample **A**, let

$$Y_k = A X_k, \quad k = 1,2,\ldots,d \tag{5}$$

Then, a family of projected feature vectors, {$\mathbf{Y}_1$, $\mathbf{Y}_2$,..., $\mathbf{Y}_d$} are obtained, which are called the principal components (vectors) of the sample image **A**. The principal component vectors obtained are used to form an *m x d* matrix **B** =[$\mathbf{Y}_1$, $\mathbf{Y}_2$,..., $\mathbf{Y}_d$], which is called the *feature matrix or feature image* of the image sample **A**.

After a transformation, a feature matrix is obtained for each image. Then, two arbitrary feature images can be compared using a nearest neighbor classifier. The distance between the matrices, $B_i = \left[Y_1^{(i)}, Y_2^{(i)}, \ldots, Y_d^{(i)}\right]$ and $B_j = \left[Y_1^{(j)}, Y_2^{(j)}, \ldots, Y_d^{(j)}\right]$, is defined by:

$$d(B_i, B_j) = \sum_{k=1}^{d}\left\|Y_k^{(i)} - Y_k^{(j)}\right\| \tag{6}$$

where $\left\|Y_k^{(i)} - Y_k^{(j)}\right\|$ represent the Euclidean distance between the two principal component vectors $Y_k^{(i)}$ and $Y_k^{(j)}$.

The feature matrix of the training images are grouped by users, then, a user, $\mathbf{U}^z$, will have associate the features matrix corresponding to his training images, $U^z = \left[B_1^z, B_2^z, \ldots, B_M^z\right]$. Given a test feature matrix **B?**, the recognized user will be that whose feature image B_j^z has the closest distance *d* to **B?**

$$if\ d\left(B?, B_l^w\right) = \min_{\forall j, \forall z} d\left(B?, B_j^z\right) \quad then \quad B? \in U^w \tag{7}$$

4.2 Recognition procedure

The recognition procedure is embedded within the automatic monitoring service explained in section 3. Our system is able to crop and get images of 110 x 80 pixels automatically because it uses a previous localization algorithm. It is able to localize a face in a room using pan, tilt and zoom capabilities of the camera and moving capabilities of the robot. Once the face is found, a tracking algorithm is applied in a window centered at the central position of the detected face in the frame before. While the face tracking process the camera and the robot are adjusted continually in order to keep the face centered in the image, with the optimal size (110 x 80). Once the tracking system has positioned the robot in front of the user and the face size is optimal, the face recognition algorithm is run.

The recognition procedure has two stages: images data base generation and user recognition. Both are run from the robotic platform. For the database generation, the user stands up in front of the robot and a sound alert is emitted when a valid image is taken in order to modify user head position, face expression or body position for the next one. In this way, faces on the images have some tilting, rotation and changes in expression and illumination. Images are obtained using a standalone application. This is useful for taking the desired number of sample images with the optimal size in some seconds. Therefore, the addition of new users to the database can be performed quickly.

In the user's recognition stage, images are obtained directly from the robot in the same way as in the data base generation one. But, in this case, robot and user position will be different. The onboard recognition procedure is applied on 10 consecutive frames and starts when the size of the face is the same as the size of the images used in the training process of the database. The user who has been recognized 6 or more times and whose distance values are under a threshold (that depends on the number of used eigenvectors) is considered to be the recognized user.

5. EXPERIMENTAL RESULTS

The goal of this paper is evaluate performance of the 2DPCA technique in a face recognition system onboard a mobile robotic platform and compare its results with the obtained with the traditional PCA technique. To do that, we have used two different kind of experiments, the first one consists on generating an own images database, taken from the robot, under similar conditions that others well-known face image databases (ORL, AR and Yale), and evaluate the system performance using database images. The second one consists on using the database images for training and get new images from the robot in different time and in different positions inside a room for the recognition stage and evaluate its performance.

5.1 Experiments on the own database

Our database contains frontal color face images from 15 individuals, each providing 30 different images with a tolerance for some tilting and rotation of the face of up to 30 degrees. Moreover, there is also some variation in facial expressions and in light conditions. All acquired images are automatically adjusted to the user's face with a resolution of 100x80 pixels.

Five sample images of one user in the face database are shown in figure 2. In these experiments, the 15 most representative images of each user have been selected among the images in the database for training and the rest have been used for evaluation. This is an important issue because the more images training variety, the better performance is achieved.

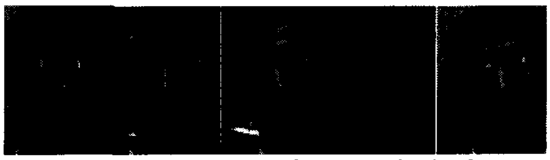

Fig 2.Five sample images of one user in the face database

The results have been obtained varying the number of principal components vectors from 1 to 40 for both methods (PCA, 2DPCA), and the number of training samples from 5 to 15 as can be seen in figure 3. The top recognition accuracy (98,66%) is achieved by 2DPCA with 5 eigenvectors and 15 samples. For the PCA method the maximum recognition accuracy is 98,22% for 20 principal components and 15 samples. As can be observed for the 2DPCA, the best accuracy results are obtained for a low number of components vectors while for the PCA technique is necessary a higher number of principal components in order to obtain similar accuracy values. For the both methods, the more samples training used, the better performance is achieved.

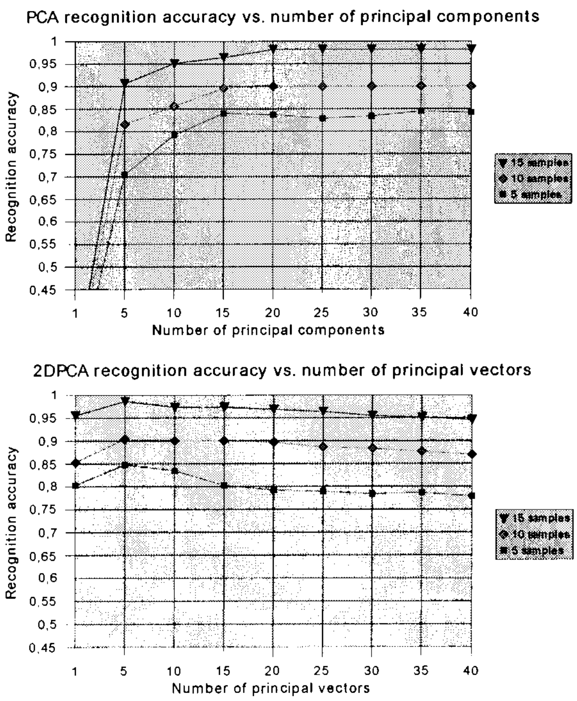

Fig. 3.Recognition accuracy performance for 2DPCA and PCA under the number of principal components and number of samples

The process time (recognition time) and the size of necessary stored data have been too analyzed in order to compare 2DPCA and PCA techniques. For PCA it is necessary to store the eigenvectors matrix, the average matrix and the principal components for every training sample. For 2DPCA, eigenvectors matrix and feature matrix per training sample are needed. Table 1 shows recognition time and size of stored data for PCA and 2DPCA methods and for our database in the case of maximum recognition accuracy. This situation is reached in the case of 2DPCA for 15 training samples and 5 eigenvectors and in the case of PCA for 20 eigenvectors. As can be seen, the training (feature extraction) and recognition time are significantly better for 2DPCA than for PCA. On the other hand, the size of stored data is higher for 2DPCA than for PCA.

In order to obtain a more exhaustive study of the 2DPCA performance we have also compared it with other typical methods, very known in the recognition literature, as *Fisherfaces* (FF) [9], *Kernel PCA (KPCA)*[10] and *Kernel Fisherfaces* (KFF) [11]. The experiment results are listed in Table 1. Respecting the maximum recognition accuracy, 2DPCA was better than other methods excepts for the *Fisherfaces* which results were equal than the obtained for the 2DPCA. Other methods needed a very higher number of eigenvectors in order to obtain a similar recognition accuracy than the 2DPCA one. Then, the size of the stored data and the process time were higher for the other methods respecting the 2DPCA. In conclusion, general performance of the 2DPCA method is better than the obtained for the other typical methods which it was compared.

In a second test seven users repeated the recognition procedure with the same light conditions, facial expressions and backgrounds that in the previous one. The number of eigenvectors was varied from 1 to 40 in steps of 5. In this case the top recognition accuracy was got with 1 eigenvectors. Three users were recognised 100 % of times and one user was not recognised any time. For the other four users, these were not recognised in 10 of 35 times. The total average recognition accuracy was 68.2 %.

6. CONCLUSIONS AND FUTURE WORK

In this paper, a face recognition system onboard a mobile robot for assisting the elderly has been presented. The system is based on 2DPCA technique. We have analyzing performance of this technique, over an own images database generated from our robot, and we have obtained similar results that obtained by others researchers over others well-known face image databases as ORL, AR and Yale. We have compared results obtained with 2DPCA technique with the traditional PCA one and we have concluded that the first one is simpler, computationally more efficient and more accurate, for a low number of components, than the second one. Then, we have compared 2DPCA and other typical methods (*Fisherfaces, Kernel Fisherfaces and kernel PCA*) performance and we have conclude that the first one is better than the second ones. Finally, we have shown that the recognition performance decrease when the recognition stage is run from the robot in different time and in different positions inside a room. We have conclude that the system is sensitive to considerable variation of user's physical aspect (beard, hairstyle) and to light condition variations if these don't have been consider in the training. The system is robust to the tilting and rotation of the face because the previous localization algorithm only gets images for recognition if they have a tilting and rotation tolerance under to 30 degrees.

	Max. Recognition accuracy	Training time (s)	Recognition time (ms)	Size of stored data (KB)	Number of eigenvectors
2DPCA	98.66 % (222/225)	0.74	86.63	489.96	5
PCA	98.22 % (221/225)	2.23	102.11	120.70	20
FF	98.66 % (222/225)	10.92	104,23	1057.60	30
KPCA	98.22% (221/225)	16.13	542.56	2166.70	20
KFF	96.88%(218/225)	5.89	524.32	1995.70	35

Table 1.Comparation between PCA and 2DPCA

5.2 Experiments on the robot in real-time

In a first test one user was selected among users in the database. The user was moving around the lab in 5 different sites with variable light conditions, facial expressions and backgrounds, while the robot was moving towards him. Whole images got from the robot are shown in figure 4, in green we can see the window for the recognition (100x80 pixels) that is adjusted automatically to the user's face.

Eigenvectors number was varied from 1 to 40. In this case the top recognition accuracy was reached with 5 eigenvectors. The average recognition accuracy was 73.6 %. The user was not recognized in 5 of 25 times.

In the early future we have the intention of enhance the user's number of our system and the number of images for training of each user under extreme conditions in indoor environments for improving the performance of the system.

Acknowledgments

This work has been supported by grants DPI2002-02193 (SIRAPEM Project) and TIC2002-10744-E (Special Action of the ADVOCATE II Project) from the Spanish Ministry of Science and Technology (MCyT).

7. REFERENCES

1. Roy N., Baltus G., Fox D., Gemperle F., Goetz J., Hirsch T., Magaritis D., Montemerlo M., Pineau J., Schulte J., and Thrun S. Towards personal service robots for the elderly. In *Proc. of the Workshop on Interactive Robotics and Entertainment (WIRE)*, Pittsburgh, PA. 2000.

2. Montemerlo M., Pineau J., Roy N., Thrun S., and Verma V. Experiences with a Mobile Robotic Guide for the Elderly. In *Proc. of the AAAI National Conference on Artificial Intelligence*. Edmonton, Canada. 2002.
3. Haigh K.Z., Phelps J., and Geib C.W. "An Open Agent Architecture for Assisting Elder Independence". In the *First International Joint Conference on Autonomous Agents and MultiAgent Systems*, pp. 578-586. 2002.
4. Lay, K. Prassler, E. Dillmann, R. Grunwald, G. Hägele, M. Lawitzky, G. Stopp, A. von Seelen,W. "MORPHA: Communication and Interaction with Intelligent, Anthropomorphic Robot Assistants". In *Tagungsband Statustage Leitprojekte Mensch-Technik-Interaktion in der Wissensgesellschaft*. 2001.
5. L. M. Bergasa, R. Barea, E. López, M. S. Escudero, J. A. Hernández. "SIRA, a Robotic System for Elderly Assistance". *ISR 2004. 35th International Symposium on Robotics*. Paris. March 2004.
6. Yang, D. Zhang, A. F. Frangi and Jing-yu Yang. "Two-dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 26, no.1, pp. 131-137, Jan. 2004.
7. ActivMedia Robotics. http://activmedia.com
8. P.Viola, M. Jones. "Rapid Object Detection using a Boosted Cascade of Simple Features", *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition*. 2001.
9. Peter N. Belhumeur, Jo©o P. Hespanha and David J. Kriegman "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no.7, Jul. 2004.
10. B. Schölkopf, A. Smola, K. R. Muller, "Nonlinear Component Analysis as a Kernel Eigenvalue Problem", *Neural Comput. 10 (5) (1998) 1299-1319.*
11. M. H. Yang, "Kernel Eigenfaces vs. Kernel Fisherfaces: Face Recognition using kernel methods", *Proceedings of the fifth IEEE International Conference on Automatic Face and Gesture Recognition (RGR'02), Washington DC, May 2002, pp 215-220.*

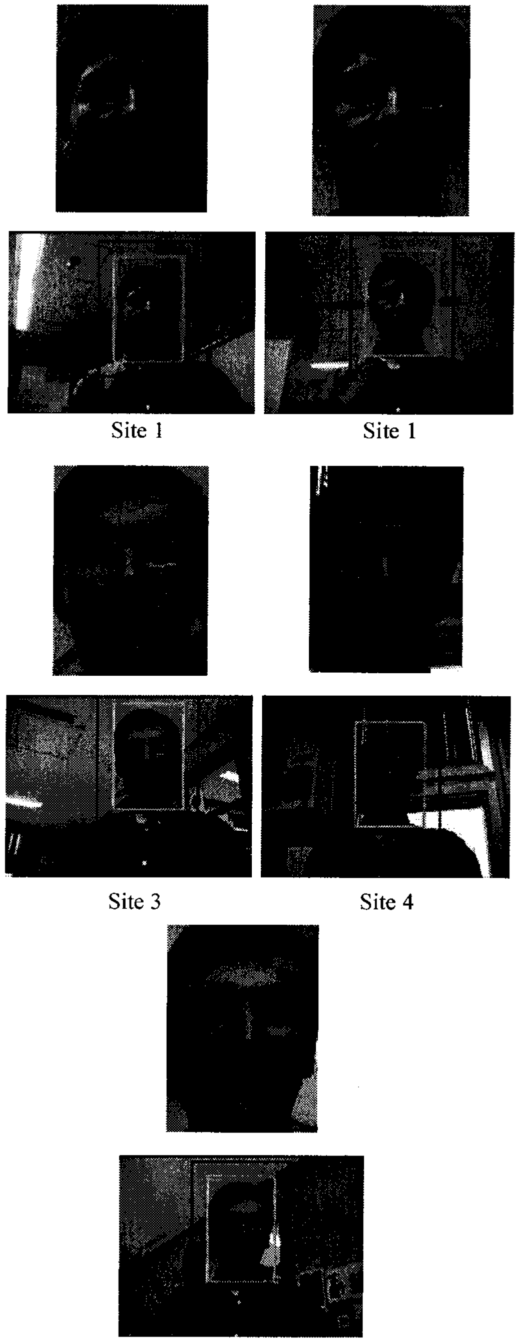

Fig. 4. Images for test 1