Face Detection & Recognition based on Fusion of Omnidirectional & PTZ Vision Sensors and Heteregenous Database

Redouane Khemmar
Research Institute for Embedded Systems
(IRSEEM), ESIGELEC,
76801 Saint Etienne du Rouvray Cedex
Rouen, FRANCE

Jean Yves Ertaud
Research Institute for Embedded Systems
(IRSEEM), ESIGELEC,
76801 Saint Etienne du Rouvray Cedex
Rouen, FRANCE

Xavier Savatier
Research Institute for Embedded Systems
(IRSEEM), ESIGELEC,
76801 Saint Etienne du Rouvray Cedex
Rouen, FRANCE

ABSTRACT

Large field of view with high resolution has always been sought-after for Mobile Robotic Authentication. So the Vision System proposed here is composed of a catadioptric sensor for full range monitoring and a Pan Tilt Zoom (PTZ) camera together forming an innovative sensor, able to detect and track any moving objects at a higher zoom level. In our application, the catadioptric sensor is calibrated and used to detect and track Regions Of Iinterest (ROIs) within its 360 degree Field Of View (FOV), especially face regions. Using a joint calibration strategy, the PTZ camera parameters are automatically adjusted by the system in order to detect and track the face ROI within a higher resolution and project the same in facespace for recognition via Eigenface algorithm.
Face recognition is one important task in Nomad Biometric Authentication (NOBA[1]) project. However, as many other face databases, it will easily produce the Small Sample Size (SSS) problem in some applications with NOBA data. Thus this journal uses the compressed sensing (CS) algorithm to solve the SSS problem in NOBA face database. Some experiments can prove the feasibility and validity of this solution. The whole development has been partially validated by application to the Face recognition using our own database NOBA.

Keywords:

face recognition, compressed sensing, nomad biometric authentication, eigenface recognition, omnidirectional sensor.

[1]This work is part of the NOmad Biometric Authentication (NOBA) project funded by ERDF under the Interreg IVA program (Ref. No. 4051) in collaboration with the University of Kent.

1. INTRODUCTION

Ample amount of development has been released with multiple-camera sensor systems to meet the rapidly growing demands in monitoring mobile robot applications. One of the recent examples is NOBA system that considers the field of biometrics and mobile robotics which are currently dissociated despite being based on common technological foundations (perception, detection and classification). A unique examples is the use of an omnidirectional camera in combination with PTZ camera, referred to as a dual camera system (as shown in Fig. 1). Omnidirectional cameras are able to exploit a wide Field Of View (FOV) within its 360 degrees for full range monitoring. However, low and non-uniform resolution of these catadioptric sensors make close observations of particular targets, especially in biometric authentication applications difficult. PTZ cameras with high mobility and zoom ability, compensates the deficiencies of omnidirectional cameras. Based on a unified model projection introduced by Geyer [6], the catadioptric sensor is calibrated, in order to generate correct perspective images. By using Viola and Jones algorithm [12] in the resulting images, the program detects the face ROI, and then applies the tracking algorithm based on a correlation approach. Then a joint calibration method is performed to localize the face ROI in order to generate a zoomed in face image with high resolution.
Face recognition can be considered as a kind of computer research or application for automatically identifying one person from a digital image or a video frame from a video source. Popular recognition algorithms can be divided into two main approaches: 1) geometric, which look at distinguishing features; 2) photometric, which is a statistical approach that distills an image into values and comparing the values with templates to eliminate variances. In this research area, there are some famous algorithms have been applied, such as Principal Component Analysis (PCA), Linear Discriminate Analysis (LDA), Local Binary Patterns (LBP), and so on. A comparative

Fig. 1. Prototype of vision system.

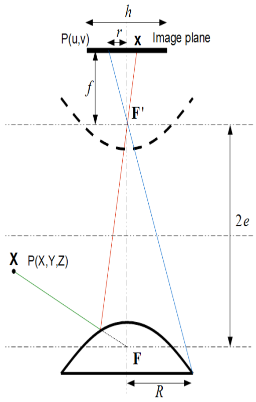

Fig. 2. Image Formation with an hyperbolic mirror.

study of several algorithms in face recognition is presented in this journal. Firstly, the extracted zoomed face is projected into face space for face recognition using Eigenface algorithm. Secondly, a particular attention has been given to a more robust approach for face recognition. Compressed sensing with small sample size problem (combined with other recognition algorithms) will be selected for face recognition and applied to heteregenous database.
This journal is organized as follows. Section 2 describes System architecture and processing, modeling and the calibration of catadioptric sensor. Section 3 focuses on the proposed face detection and tracking algorithms. The fusion procedure of the omnidirectional and PTZ cameras is described in section 4. Section 5 describes about the implementation of Eigenface algorithm on our NOBA local database. In Section 6, we will use the CS algorithm to solve face recognition task in which we introduce our research background and related theory. Section 7 illustrates the experiment/analysis results and Section 8 concludes this journal.

2. SYSTEM ARCHITECTURE AND PROCESSING

In the presented work, we focus on the issues related to face detection, extraction and recognition in a two camera serial architecture. The developed system is characterized by unbalanced sensors functionalities : A data exchange program from the processing unit is designed to detect the domain of interest through the wide-field catadioptric sensor. The main program is performed to manage the processing and the pointing of the PTZ camera into the corresponding region of interest. Commanding and controlling position of the PTZ camera is established by using a network connection through http sockets (virtual channels), where messages and control information are sent and received.

2.1 The catadioptric Sensor

Fixed view point constraint. : The architecture of catadioptric sensor adheres to the Single-View-Point (SVP)theory [1]. The SVP constraint enables to generate correct perspective images. In fact, the optical center of the camera has to coincide with the second focus F of the hyperbola located at distance 2e from the mirror focus as illustrated in Fig.2. The eccentricity "e" is a parameter of the mirror given by the manufacturer. To realize this task, we first calibrate our camera with a standard calibration tool to determine the central point and the focal length. Knowing the parameters of both the mirror and the camera, the image of the mirror on the image plane can be easily predicted if the SVP constraint is taken into consideration, as shown in Fig.2. The expected mirror boundaries are superposed on the image and the mirror has then to be moved manually to fit this estimation as shown in Fig.3.

2.2 Sensor calibration

Sensor calibration is the process to determine the optical and geometrical features which are generally addressed as intrinsic and extrinsic parameters and they allow to estimate the correspondence between the 3D points of the scene and their projection into the image plane (pixel coordinates). The camera calibration used is based on a generic model introduced by [6] and [2], and then, modified by [9], who generalized the projection matrix and took into consideration distortions. Further calibration details can be found in [4] and [10]. Figure 5 represents the projection process.
As described in [2] and [9] as well as mentioned in [7], the projection of a 3D point can be done by projecting the 3D point X[w y z] onto the unit sphere centered on C_m : then Points $X_s = [x_s \; y_s \; z_s]^T$ are, then, projected onto the new frame with the origin $C_p = [0 \; 0 \; \xi]^T$, the obtained point $(X_s)_{C_p}$ onto a normalized plane and finally the last step enables us to find the camera projection matrix K expressed according to γ_u and γ_v, which are respectively, the generalized horizontal and vertical focal length, and (u_0, v_0) the coordinates of the principal point on the image point and the skew α :

$$p = K.m = \begin{pmatrix} \gamma_u & \gamma_u.\alpha & u_0 \\ 0 & \gamma_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} .m \qquad (1)$$

In our model, we consider that the impact of the parameter α, often null, is irrelevant. Parameters to be estimated in that model are : $\xi, \gamma_u, \gamma_v, u_0$ and v_0.

With the tool developed by [3], calibration is achieved by observing a planar pattern at different positions. The pattern can be freely moved (the motion does not need to be known) and the user needs to select the four points corners pattern. This calibration process is similar to that of Mei [9]. It consists of a minimization over all the model parameters of an error function between the estimated

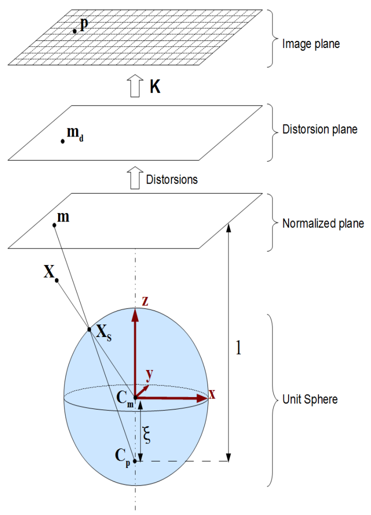

Fig. 3. Unified projection model.

projections of the pattern corners and the measured projection using Levenberg-Marquardt algorithm [8]. This minimization process enables to find the parameters combination that reduces the error of pattern retro projection.

3. FACE DETECTION & TRACKING BASED ON MERGER OF CATADIOPTRIC SENSOR AND PTZ CAMERA

Most of image processing techniques are performed on conventional images, i.e. perspective images. Actually, deformations caused by the catadioptric system do not give us the opportunity to perform the existing face detection algorithms on raw images (Fig 4.). As a consequence, geometrical transformations should be performed to obtain a panoramic image close to perspective images, where the face detection algorithm will be performed.

3.1 Panoramic images unwrapping

The existence of the model of the unit sphere, simplifies the unwrapping problem. Under the fixed view point constraint and by performing a retro projection, we are able to project the pixels of the panoramic pictures onto the unit sphere used in the unified model. Then, these pixels are projected onto the image plane. Thereby, we obtain the mapping between the pixels on the panoramic image and their corresponding on the camera retinal plane. Fig. 5 shows the result of the spherical transformation of the considered catadioptric image. Face detection algorithm is applied to these unwrapped images.

3.2 Face detection and tracking algorithms

Face detection techniques have been researched for years and much progress has been proposed in literature. However, in 2001, Paul Viola and Michael Jones [12] achieved a robust real time method for face detection, which was fifteen times quicker than the methods existing at that time. The technique relies on the use of simple Haar-like features that are evaluated quickly through the use of a new image representation called “integral image” that allows fast

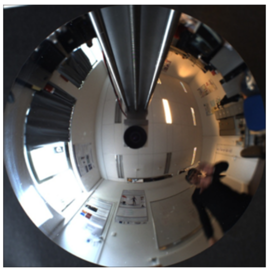

Fig. 4. Original catadioptric image.

Fig. 5. Unwrapped image obtained by spherical transformation.

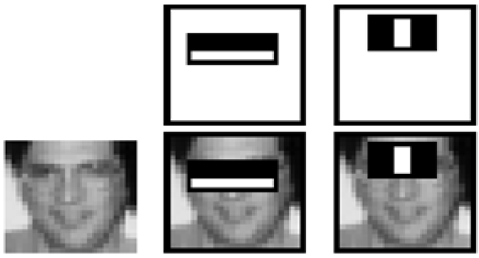

Fig. 6. Haar Like Features.

feature evaluation. Fig. 6 represents two of the 60000 Haar-like features available.
To select the best filtering feature, Adaboost, the machine learning introduced in [5], is used. In fact, given a set of weak classifiers, not much better than random, if we iteratively combine their output, the training error will quickly converge to zero. Fig. 7.(a) & Fig. 7.(b) shows the detected and zoomed face image. As explained in [7], the face ROI so obtained is zoomed in with a high resolution which further can be processed for face recognition.

4. MERGER OF OMNIDIRECTIONAL AND PTZ CAMERAS

4.1 Joint Calibration Strategy

The joint calibration method is based on defining a reference position (x_0, y_0) for the PTZ camera on the catadioptric 360° image. This position has to coincide with the default orientation of the PTZ camera and is chosen as the starting point for pan angle evaluation (Fig. 7). In fact, given a point $p(x, y)$ and the image catadioptric image center $C(x_0, y_0)$, we can compute the pan angle θ_p in the

(a) (b)

Fig. 7. (a). Detected face Image (b). Zoomed Image

omnidirectional referential:

$$\theta_p = arctan\left(\frac{y - y_0}{x - x_0}\right) \quad (2)$$

$$\tan(\alpha_{pan}) = \frac{v_y}{v_x} \quad (3)$$

The detailed explanation and formulae is inspired by paper [7].

4.2 Face detection with PTZ camera

As explained in the previous sections, the catadioptric sensor is able to detect and track face ROIs. By using the ROI data localization from the catadioptric images, the PTZ camera can detect and also make a zoom in the face ROI. Brief descriptions of the different steps for the face ROI localization are as follow: Firstly, perform Viola and Jones face detection algorithm on the unwrapped catadioptric image to identify and localize the face ROI. Secondly, using the pan angle of the ROI center, calculate the pan angle to be directed to the PTZ camera. We use a constant tilt angle and a minimum zoom value. This enables us to point the PTZ camera on a large area where the face is probably located. Thirdly, perform again, Viola and Jones face detection algorithm to detect the face ROI in the obtained PTZ image. Finally, command the PTZ camera in order to center the face ROI detected in the PTZ image (pan and tilt calculation) and then apply the zoom factor computed according to both image and ROI widths and heights, as expressed below:

$$Zoom = Min\left(\frac{Image_{width}}{ROI_{width}}, \frac{Image_{height}}{ROI_{height}}\right) \quad (4)$$

This last step enables us to obtain a zoomed in face image with high resolution, is useful for face recognition processing.

5. FACE RECOGNITION BASED ON FUSION OF CATADIOPTRIC AND PTZ VISION SENSORS

Face recognition is a very active area of research in computer vision and biometric fields since late 1980s. Among the plethora of techniques available Eigenfaces technique is one of the earliest appearance-based face recognition methods, which was developed by M. Turk and A. Pentland in [11]. The study and evaluation of the performances of this method for the Fusion based System may be interesting to implement in Real Time face recognition. In this section, we study the NOBA face recognition based Eigenface algorithm.

Fig. 8. Average Images.

5.1 The Eigenface Algorithm

The eigenfaces technique for face recognition, developed by M. Turk and A. Pentland [11], consists of two main phases: • Learning: This phase utilizes the idea of the Principal Component Analysis (PCA) and decomposes face images into set of characteristic feature images called eigenfaces. • Recognition: This phase is then performed by projecting a new face into a low dimensional linear space defined by the generated eigenfaces in order to analyze and then to recognize it.

5.1.1 Learning phase. The learning phase for face recognition technique involves the following steps:

—The first step is to obtain a set S with M training face images. Let a face image be a two-dimensional N by N array of intensity values. After obtaining our set of face images, we calculate the mean images Ψ.

—The next step consists in calculating the difference faces by subtracting the average face from each input image.

—This set of very large vectors is then subject to the principal component analysis method, which seeks a set of M orthonormal vectors μ_n, which best describes the distribution of the data. The k_{th} vector μ_k is chosen such that : We note that the vectors and scalars are the eigenvectors and eigenvalues, respectively, of the covariance matrix.

$$\lambda_k = \frac{1}{M}\sum_{n-1}^{M}(\mu_k^T\phi_n)^2 \quad (5)$$

The covariance matrix C is defined as follows :

$$C = \frac{1}{M}\sum_{n-1}^{M}\Phi_n\Phi_n^T = AA^T \quad (6)$$

With A=$[\Phi_1\Phi_2\Phi_3.......\Phi_M]$

Following this equation set, we construct the M by M matrix $L = A^TA$, where $L_{mn} = \Phi_m^T\Phi_n$ and find the M eigenvectors $v_n$of L. These vectors determine linear combinations of the M training set face images to form the eigenfaces μ_n :

$$\mu_n = \sum_{k=1}^{M} v_{nk}\Phi_k = Av_n, n = 1.......M \quad (7)$$

With this method, the calculations are greatly reduced from the number of pixels in the images (N^2) to the other number of images in the training set (M).

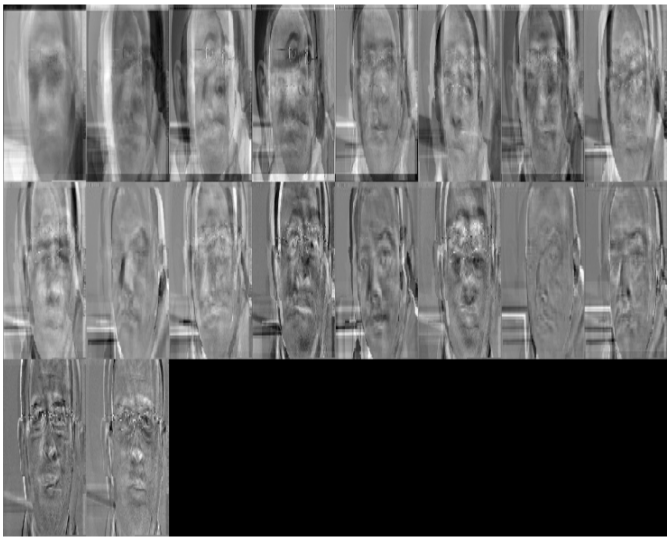

Fig. 9. Output of Eigenfaces.

5.1.2 Recognition Phase. The usefulness of the eigenvectors varies according to the associated eigenvalues. In fact, we choose only the most meaningful eigenvectors corresponding to the highest eigenvalues and we ignore the rest. As a consequence, the number of basis functions is reduced from M to M' (M' less than M) and the computation is reduced. The recognition procedure is summarized as follows:

(1) A new face, to be identified, is transformed to its eigenface components. First, we compare our input image with the mean image and then multiply their difference with each eigenvector of the obtained L matrix. Each value would represent a weight and would be saved on a vector Ω :

$$\omega_k = \mu_k(\Gamma - \Phi) \tag{8}$$

where $\Omega^T = [\omega_1, \omega_2, \omega_3 ....\omega_M]$

(2) We determine which face class provides the best description for the input image. This is done by minimizing the Euclidean distance and also the distance between the image and the face space :

$$\xi_k = \|\Omega - \Omega K\|^2 \ \xi = \|\Phi - \Phi_f\|^2 \tag{9}$$

(3) We choose a threshold that defines the maximum allowable distance from any face class, and a threshold that defines the maximum allowable distance from face space. If the minimum distance and the distance , classify the input face as the individual associated with class vector. If the minimum distance but the image may be classified as unknown.

(4) If the new face is classified as known individual, this image may be added to the original set of familiar set images, and the eigenfaces may be recalculated. This gives the opportunity to modify the face space as the system encounters more instances of known faces.

The entire eigenface program is implemented in C language by using OpenCV library and the main functions developed to learn and recognize faces with OpenCV eigenface methods.

Table 1. Noba Database.

Image Features
Image 1: Emotional Expression of Smiling
Image 2: Laughter
Image 3: Anger
Image 4: Surprise
Image 5: Eyes Closed
Image 6: Fcae with high illumination
Image 7: Tilt UP 30°
Image 8: Tilt Right 30°
Image 9: Tilt Left 90°
Image 10: Tilt Left 60°
Image 11: Tilt Left 30°
Image 12: Tilt Right 60°
Image 13: Tilt Right 90°
Image 14: Tilt Down 30°
Image 15: Frontal View Neutral Face
Image 16: Face with Low illumination

5.2 NOBA Database

The Nomad Biometric Authentication (NOBA) project focuses on the development of biometric technologies for strong authentication and secured data exchange. This project is undertaken as part of the Franco-British cooperation INTERREGIVA and the GRR EEM (major research in Electronics, Energy and Materials). It brings together two research teams, IRSEEM and the Department of Electronics of the University of Kent. The ultimate goal of the NOBA project is the development of a smart system that cooperatively uses solutions based on vision rely on fixed devices embedded on a mobile platform. Its main applications would be in Electronic surveillance, ambient control of restricted access areas or other security systems. NOBA database includes 3 Subsets : Gait database, Face database and Iris database.
In NOBA Face DataBase, we have 10 subjects (or persons). The face data are captured from 2 sensors : Pan Tilt Zoom (PTZ) camera & Camera with infrared light (2D digitizer). During building the Face database, we take into account variations of Poses of head and eyes direction (various poses), Facial expressions (various expressions), Illuminations conditions (various illuminations). The combination of expressions under illumination and poses under expressions, 4 classes are defined :

—C1 : Facial expressions descriptors,

—C2 : Mouth movements descriptors,

—C3 : Pose of head and eye direction descriptors,

—C4 : Various Illuminations.

16 face images for each subject with 4 frontal images, 8 images for different tilts, 2 facial expressions images, 2 images with various illuminations. Image features are mentioned in Table 1. Our idea is to test and verify the results of Eigenface on our own database and see the efficiency and accuracy in terms of implementation.

5.3 Results obtained under the heterogeneous database

To test performances of the developed application, we used a free publicly available face database, the Olivetti Research Laboratories (ORL). This face database provides 10 sample images of each of 40 subjects. For some of the subjects, the images were taken at different times, varying lighting slightly, facial expressions (open/closed eyes, smiling/non-smiling) and facial details (glasses/no-glasses).

Table 2. Results of test subjects s45 and s50 with non-trained faces.

Nearest	Truth	Wrong/Right	Confidence
50	50	correct	0.925854
45	45	correct	0.938389
50	50	correct	0.882469
50	50	correct	0.945845

Table 3. Results of different trained and non-trained faces.

Nearest	Truth	Wrong/orrect	Confidence
50	50	Correct	0.924261
45	45	Correct	1.000000
23	23	Correct	0.971391
26	23	Correct	0.936796

All the images are taken against a dark homogeneous background and the subjects are in up-right, frontal position (with tolerance for some side movement). And we merged our subjects (under the same conditions) with this database to obtain a Heterogeneous database. The experimental results are performed with 12 subjects (10 ORL + 2 NOBA database) for the time being. Better results and testing to be done on extended number of subjects in the near future.

5.3.1 Various scenarios. Results obtained show that the developed eigenface program recognizes correctly the considered people belonging already to the face database (training data). The index of the nearest face image corresponds to the ground truth and the confidence level is at 100%. The confidence level is calculated based on the Euclidean distance, so that similar images should give a confidence between 0.5 to 1.0, and very different images should give a confidence between 0.0 to 0.5. The confidence level has to be higher than 0.6 because face images of this person are included to the database but they are not acquired in the same conditions. Excluded images outside of the training database with a confidence level varying from 0.87 until 0.95. Sometimes, this level confidence can be higher until 0.99 and lower until 0.6.
Although the testing images don't correspond to any person, the algorithm provides a confidence level very high (greater than 0.8) in both cases. With results obtained in this test and several tests performed, we conclude that the confidence level is not a good criterion on which we have to rely, in order to recognize people. The algorithm works only when the index of the nearest training face image corresponds exactly to the one of the testing one. The results for the test images (which are mentioned in Table 2 and 3) of two subjects which are trained but doesnot have the exact match. So confidence level is not 1.
Training Subjects s20 to s30 and two of NOBA database s45 & s50 are trained except s23/6,s23/7,s50/8,s50/9,s45/2,s50/2 & s26/1. The results obtained shows the difference in Confidence level depending on whether that face was present in the training data or not. Better the match, higher the confidence level approaching 1.
Accuracy can be seen to be almost closing 100% as the trained data is able to recognize the test image if the subject is trained previously no matter image do not match exactly.

6. COMPRESSED SENSING FACE RECOGNITION METHOD IN HETEROGENEOUS DATABASE WITH SMALL SAMPLE SIZE PROBLEM

The traditional approach to accomplish one face recognition task is by comparing selected facial features from the image and a facial database. Actually, the fundamental of these methods is based on Euclidean distance. However, the recognition result of this idea may be not satisfactory in some situation. One common problem in many databases is so-called Small Sample Size (SSS) problem. When the sample size is smaller than the sample dimensionality, the within-class scatter matrix is singular, which is also named the SSS problem. This section of paper aims to study the performance of Compressed Sensing (CS) algorithm for face recognition task in NOBA database. There will be SSS problem in some special situation.

6.1 Background and Theory

6.1.1 NOBA Face database. During the construction of NOBA database, the acquisition conditions are verified like variations of poses of head and eyes direction and illuminations conditions. However, as many other face databases (where the images are medium quality even worse), there will be SSS problem in NOBA when the training samples can not be acquired adequately. Thus we need to use some new methods to improve this situation.

6.1.2 What is compressed sensing?. As a novel signal processing method, CS has been a hot point topic in many research fields. The fundamental theory of CS has been respectively proved by Candès, Tao and Remberg [1, 2], and Donoho [3]. On the contrary of traditional signal processing, this method means that the signal can be recovered from much fewer measurements than what is usually considered necessary. From the research of Candès *et al.* [1, 2], one CS framework can be composed of two main parts. The first part is CS sampling process which is that the original signal $x \in R^N$ can be acquired through the following linear random projections as :

$$y = \Phi x = \{< x, \varphi_i >\}\,|_{i=1}^{M} x \in R^M \;\; with \;\, x \in R^N \tag{10}$$

where Φ is an underdetermined sensing matrix, $M \ll N$, and y is the so-called sensing data of x. The second part is CS recovery process which is essentially an improvement of pseudoinverse operation. Although Φ and y can be acquired after CS sampling, Φ is underdetermined which is difficult to find the inverse. In fact, if x is sparse ($K \ll M$) and the RIP (Restricted Isometry Property) of Φ is fulfilled, the inverse problem of Eq.(10) can be transformed as :

$$\min \|x\|_0 \quad s.t. \quad \Phi x = y \tag{11}$$

In theory, the solution of Eq.(11) can be global optimal. But the $l_0 - norm$ induces a complete metric topology for the space of measureable functions. And the $l_0 - norm$ minimization belongs to NP (Non-deterministic Polynomial) problem which is hard to solve.

$$\min \|x\|_1 \quad s.t. \quad \Phi x = y \tag{12}$$

Fortunately, Candès, Tao and Romberg [1, 2] give and prove the critical point of CS that the optimal solution in $l_1 - norm$ minimization of Eq.(12) can be equivalent to in $l_0 - norm$ minimization of Eq.(11) under some conditions. This will guarantee the feasibility and stability of CS recovery problem.

6.2 CS in Face Recognition

In the domain of Face Recognition (FR), some new methods based on CS have been presented. There are two typical FR approach with CS. One is Sparsity Preserving Projections (SPP) [4]. Unlike many existing techniques such as Local Preserving Projection (LPP) and Neighborhood Preserving Embedding (NPE), where local neighborhood information is preserved during the Dimensionality Reduction (DR) procedure, SPP aims to preserve the sparse reconstructive relationship of the data, which is achieved by minimizing a l_1 regularization-related objective function. The obtained projections are invariant to rotations, rescalings and translations of the data, and more importantly, they contain natural discriminating information even if no class labels are provided. Moreover, SPP chooses its neighborhood automatically and hence can be more conveniently used in practice compared to LPP and NPE. Another is the famous Sparse Representation (SP) [5]. Based on a SP computed by l1-minimization, a general classification algorithm is proposed for (image-based) object recognition. With sparsity properly harnessed, the choice of features becomes less important than the number of features used. Moreover, occlusion and corruption can be handled uniformly and robustly within the same classification framework. One can achieve striking recognition performance for severely occluded or corrupted images by a simple algorithm with no special engineering. Recent development in the emerging theory of SP or CS [5] reveals that if the solution x_0 sought is sparse enough, the following problem can be solved in polynomial time by standard linear programming methods :

$$\min \|x\|_1 \quad s.t. \quad \Phi x = y$$

Even more efficient methods are available when the solution is known to be very sparse. For each class i, let $\delta_i : R^N \rightarrow R^N$ be the characteristic function which selects the coefficients associated with the $i-th$ class. For $x \in R^N$, $\delta_i(x) \in R^N$ is a new vector whose only nonzero entries are the entries in x that are associated with class i. Using only the coefficients associated with the $i-th$ class, one can approximate the given test sample y as $\hat{y_i} = A\delta_i(\hat{x_1})$. We then classify y based on these approximations by assigning it to the object class that minimizes the residual between y and $\hat{y_i}$:

$$Identity(y)_{i=1,\ldots,k} = \min \|y - A\delta_i(\hat{x_1})\|_2 \tag{13}$$

6.3 Experiments and Analysis

6.3.1 Experiment context. Fig.10 shows the flowchart of two face recognition methods LBP and CS. In fact, the function of LBP can be considered as a kind of feature extraction method; and CS is not the feature extraction technique but a special classification strategy which is based on CS theory. Therefore, the goal of this experiment is to actualize the contrastive experiment of traditional method (LBP) and CS method in some face image dataset with SSS problem.

6.3.2 CS under heterogeneous face database. In this experiment, some face images will be selected from several different databases as the experimental samples to compose a heterogeneous face database, including NOBA, ORL and CSU as it shown in Fig. 11. There are ten different images of each of 31 distinct subjects. For

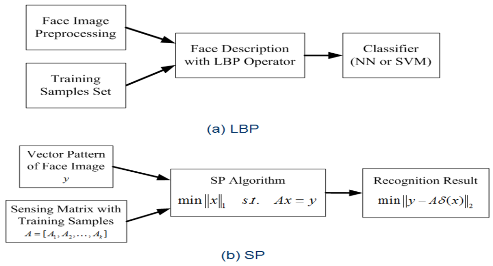

Fig. 10. The flowchart of LBP and CS with face recognition.

Fig. 11. Several examples in the heterogeneous face database.

some subjects, the images were taken at different times, varying the lighting, facial expressions and facial details, etc. The size of each image is 92x112 pixels, with 256 grey levels per pixel. Generally, there are at least two pivotal steps in any pattern recognition process. One is feature extraction and another is classification technique. In this experiment, Local Binary Patterns (LBP) and Sub-Sampling (SS) will be used as two feature extraction methods; and Nearest Neighbor (NN) and CS will be selected as two classification methods. Therefore, this experiment consists of three combinations which are 1) LBP-NN, 2) SS-CS and 3) LBP-CS.

For this heterogeneous database, two strategies about how to select test sample will be actualized. Firstly, is to select one image as test sample randomly form the whole database. In other words, the test sample may be from NOBA, ORL or CSU. Secondly, is that only NOBA image will be chose as test sample. No matter which strategies, three combinations about face recognition will be carried into execution. And recognition rate is the final index to evaluate the performance of different three recognition methods. The code of LBP is cited from the work of Ahonen *et al.* [6]. The code of CS is referred from the research work of Koh et al. [7]. In each group, there are 10 time iteration operation with three recognition strategies (LBP-NN, SS-CS and LBP-CS). And the number of realization is 100 in each iteration operation. 1) Fig.12 (a) shows the result when the test sample is selected from the whole heterogeneous database randomly. Obviously, the recognition rate of SS-CS is much better than LBP-NN and LBP-CS. The difference between SS and LBP is that the former is redundancy but the latter is not. It implies that the redundancy feature extraction method (such as SS) will more suit to SP face recognition. 2) Fig.12 (b) shows the result when only NOBA images can be randomly selected as the test sample. In this group, it shows that both LBP-NN and SS-CS are good enough for face recognition task. However, the result of LBP-CS is not satisfactory which is only about 0.7. This indicates that LBP can be good at the traditional recognition or classification problem but it may not work well in the CS method.

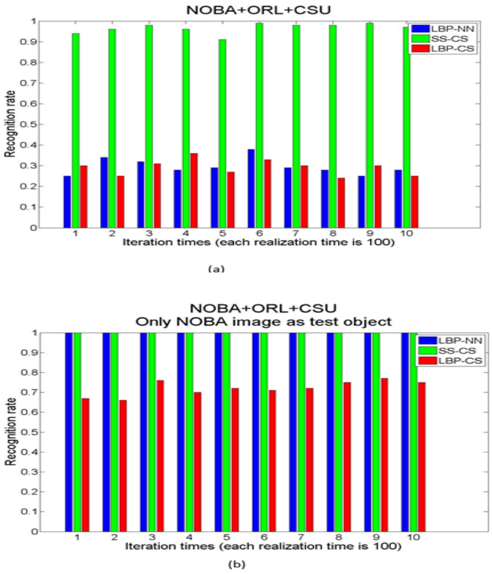

Fig. 12. Experimental results based on the heterogeneous database : (a). NOBA+ORL+CSU. (b). NOBA+ORL+CSU Only NOBA Image as test Object.

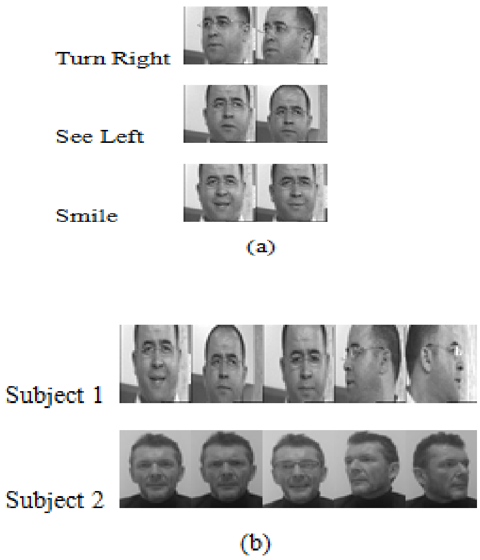

Fig. 13. Two classification standards for NOBA database : (a). Different behaviors. (b). Different subjects.

6.3.3 CS under NOBA face database. This experiment aims to study the performance of CS algorithm for face recognition task with SSS problem in NOBA database. There are two classification standards which one is by behavior and another is by subject. Fig.13 shows the two classification examples. As same as the last experiment, three recognition combinations will be selected, including 1) LBP-NN, 2) SS-CS and 3) LBP-CS. However, only one strategy about the selection of test sample will be actualized. It is that the test sample number will be generated randomly in each system running. As the above mentioned, there are two classification standards for NOBA database. For the one standard with behavior, the number of each class N_1 is only 2. For another standard with subject, the number of each class N_2 is 15. In NOBA database, the size of each image S is 92x112 pixels. Because of $x \in R^N$, it is obviously that there is the so-called SSS problem in NOBA. Actually, the SSS problem is known to have significant influences on the design and performance of a statistical pattern recognition system. In other words, the execution of traditional recognition method (such as LBP-NN) may encounter computational difficulty. Therefore, the purpose of this experiment is not only to compare the recognition rate based on LBP or CS in NOBA but also to research the performance of the two different recognition methods in the SSS problem.

7. RESULTS

7.1 Face Recognition based Eigenface Algorithm Results

The preliminary results of the actual implementation of the initial prototype of the system is discussed in this section. In order to test the proposed system, the omnidirectional camera has been placed at 20 cm under the dome Axis PTZ camera [7]. Viola Jones [12] algorithm can detect only frontal and semi-profile faces.Once the face ROI localization is determined from the unwrapped image, the program sends the corresponding pan angle via http sockets. Then, a separated process is created to detect the face in the current PTZ image (Fig.7a) by applying Viola and Jones algorithm. Fig. 7(a) illustrates the resulting image after performing this face detection algorithm. Finally, the program control automatically the camera in order to center and zoom in the detected face localization. As a consequence, we obtain a high resolution face picture (Fig.7 b). The localized PTZ image is then downsampled and converted to grayscale image of 92x112 (low-dimensional linear subspace defined by eigenfaces) and saved into the test database for recognition and eigenface algorithm is implemented to check for the authenticity or match of the given subject with that of the training or registered database or subject. A new face is compared to known face classes by computing the distance between their projections onto the face subspace.

Test results obtained (as shown in Fig.9) demonstrates that the algorithm provides good results when the testing images correspond to known people belonging to the training face database. Further one system limitation is represented by the poor resolution in the unwrapped images (Fig.8). This can affect performances of the face detection algorithm. Actually, at relatively high distances from the optical axis of the omnidirectional camera (more than 2 meters), Viola and Jones detector is not able to detect face location in these unwrapped images. We can notice that the Viola and Jones algorithm performs more efficiently when the target face is closer to the catadioptric sensor and also when assuming adequate illumination conditions. The current system is also limited by the large number of threads, the program runs. This affects the processing time of the application and the frame rate of the video stream. Therefore, to further optimize these parameters, improvements in the system architecture and improvements related to the decision criteria should be implemented during the last phase of recognition are objectives of the upcoming studies.

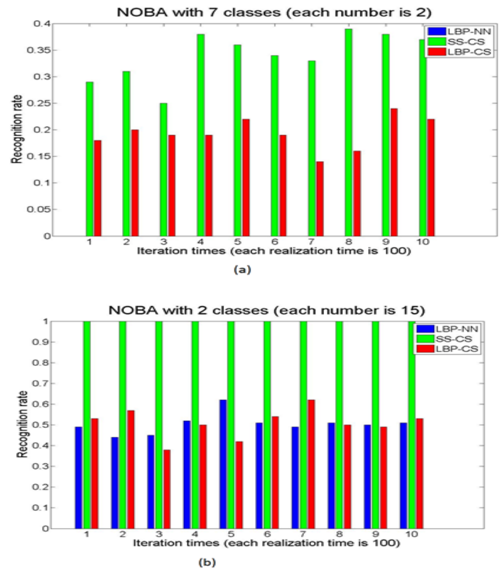

Fig. 14. Experimental results for two classification standards: (a). NOBA with 7 classes (each number is 2). (b). NOBA with 2 classes (each number is 15).

7.2 Face Recognition based Compressed Sensing with SSS Problem Results

LBP-NN, SS-CS and LBP-CS still are used as the recognition strategy for NOBA database. As two main parameters, the system of each strategy will have 10 iterations respectively and each iteration will has 100 realizations.

Fig.14 (a) is the classification result with behavior information in NOBA. It shows that either of SS-CS and LBP-CS work better than LBP-NN. It is obvious that there is serious SSS problem in this classification standard. Although the recognition rate of SS-CS and LBP-CS is not good enough, it implies that the performance of recognition system with SSS problem can be improved effectively by using CS algorithm.

Fig.14 (b) shows the result based on the subject´s difference. It can easily be seen that SS-CS has the best recognition rate in all three. Although the level of SSS problem is subdued, it still has a negative influence to the recognition method with LBP (LBP-NN and LBP-CS). Because SS can reserve redundancy information which is the pivotal factor in CS method. This means that SS-CS may be one useful solution to SSS problem in face recognition.

8. CONCLUSION

Firstly, in the first part of this paper, we proposed a unique vision system, effecient to automatically detect and track ROI at a higher zoom level. Experimental results using robust calibration methods and real-time detection and tracking algorithms demonstrates a significantly improved accuracy in providing a closer look of the target for recognition purposes. A panoramic FOV eliminates the need for more cameras or mechanically turnable camera. The integration of authentication processes like face detction, tracking and recognition makes the system self sustained for biometric authentication.The advantages of omnidirectional sensing are obvious for application like survelliance and immersive telepresence.

Secondly, according to the experimental result, it can be seen that CS face recognition method (SS-CS) usually can be better than the traditional method (LBP-NN) which is based on Euclid distance. And the combination of LBP and CS also is not satisfactory in the face recognition task for the heterogeneous database because redundancy is one important factor in CS method but LBP is not. In short, SS-CS will have a good performance in the face recognition task for heterogeneous database. On the other side, the redundancy is important to the recognition method with CS technique. This idea will be opposite to the traditional recognition method (such as LBP) which will be more concerned about specificity. However, this experiment shows that CS algorithm with redundancy feature extraction can have positive significance for accomplishing the pattern recognition task especially with SSS problem.

Our future work focus on the improvement of the current system architecture in order to make the detection and tracking algorithms more robust and faster. Better and effecient methods for object detection in the unwrapped pictures and recogntion of face with higher and faster recognition rate are the heart of forthcoming studies. Moreover, our upcoming work will be oriented toward merging more biometric analysis like iris and gait to have a robust and mobile multimodal biometric system.

Acknowledgments

The authors would like to thank all people who are given volunteers for the construction of NOBA Database. We thank them for their availability during image data acquisition.

9. REFERENCES

[1] S. Baker and S.K. A theory of single-viewpoint catadioptric image formation. International Journal of Computer Vision., 103(3):175-196, 2006.

[2] J. Barreto. A unifying geometric representation for central projection systems. Computer Vision and Image Understanding., 103(3):208-217, September 2006.

[3] R. Boutteau. Reconstruction tridimensionnelle de l´environnement d´un robot mobile a` partir d´informations de vision omnidirectionnelle pour la preparation. PhD Thesis., University de rouen, 2009.

[4] R. S. D. Scaramuzza, and A. Martinelli. A flexible technique for accurate omnidirectional camera calibration and structure from motion. Proceedings of the International Conference on Computer Vision., 45-52, January 2006.

[5] Y. Freund and E. Schapire. A decision-theoritic generalization of on-line learning and an application to boosting. Journal of Computer and System Science., 1997.

[6] Geyer and K. Daniilidis. A unifying theory for central panoramic systems and practical implications. Proceedings of the European Conference on Computer Vision., 445-461, 2000.

[7] A. Iraqui. H, Y. Dupuis, R. Boutteau, J.-Y. Ertaud, and X. Savatier. Fusion of omnidirectional and ptz cameras for face detection and tracking. In International Conference on Emerging Security Technologies (EST)., 2010.

[8] Levenberg. A method for the solution of certain probless in least squares. Quarterly of Applied Mathematics,. 2:164-168, 1994.

[9] C. Mei and P. Rives. Single view point omnidirectional camera calibration from planar grids. In Proceedings of the International Conference on Robotics and Automation (ICRA)., 3945-3950, 2007.

[10] S. R. Ramalingram, and P. Lodha Sturn. Towards complete generic camera calibration. In Proceedings of the International Conference on Computer Vision and Pattern Recognition., 767-769, June 2005.

[11] M. Turk and A. Pentland. Eigenfaces for recognition. In Vision and Modeling Group The Media Laboratory., 1994.

[12] P. Viola and M. Jones. Robust real-time object detection. International Journal of Computer Vision., 2001.

[13] E. Candès, J. Romberg, and T. Tao. Robust uncertainty principles : Exact signal reconstructionfrom highly incomplete frequency information. IEEE Transactions on Information Theory., 52(2):489-509, 2006.

[14] E. Candès, J. Romberg,and T. Tao. Near optimal signal recovery from random projections : Universal encoding strategies. IEEE Transactions on Information Theory., 52(12):5406-5425, 2006.

[15] D.L. Donoho. Compressed sensing. IEEE Transactions on Information Theory., 52(4):1289-1306, 2006.

[16] L. Qiao, S. Chen, X. Tan. Sparsity preserving projections with applications to face recognition. Pattern Recognition., 43(1):331-341, 2010.

[17] J. Wright, A.Y. Yang, A. Ganesh, S.S. Sastry, and Y. Ma. Robust face recognition via sparse representation. IEEE Transactions on Pattern Analysis and Machine Intelligence., 31(2):210-227, 2009.

[18] T. Ahonen, A. Hadid, and M. Pietikainen. Face Description with Local Binary Patterns: Application to Face Recognition. IEEE PAMI., 28(12):2037-2041, 2006.

[19] K. Koh, S. J. Kim, and S. Boyd. An Interior-Point Method for Large-Scale l1-Regularized Logistic Regression. The Journal of Machine Learning Research., 8:1519-1555, 2007.