Biometrics

3D-face-recognition

Multimodal

Semester Project - Final report

Eurecom Fall 2005

Written by : Stephan Singh, Ayoub Bacherki

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The aim of this project is to provide a 3D recognition system using Eigenfaces. It has to support multimodal techniques taking advantages of the texture and the 3D shape of the faces.
First we will talk about the state of the art in the multimodal face recognition. Then we will apply several recognition method, such as score fusion and distance calculation to have a range of result and compare them. Finally, we will introduce our Matlab identification system.

Chapter 2

Enrollment systems

The majority of companies building 3D cameras provide hardware and / or software solutions for the 3D-facial-recognition. We found many other brand, but most of them are used for medical and not for biometrics.

2.1 A4 Vision

A4 Vision [1], based in Geneva, Switzerland, offers an enrolment machine, it is used to perform enrolment and the building of biometric 3D face templates. The output is a 3D Biometric template and a standard 2D color image (the texture). The main feature is using infrared light to have Light Invariance of the 3D model and exploit all its benefits. This machine can be used in a verification or identification mode (a new model with an LCD screen exists).

Figure 2.1: A4 Vision: 3D Enrollment camera

The web site doesn't give any information about algorithms used for enrolment and recognition. These algorithms are proprietors.

Figure 2.2: A4 Vision: 3D Face reader LCD

Figure 2.3: A4 Vision: 3D Face reader

A4vision gives some performances of the system:

- Verification Time: < 1 second
- identification time: < 1 second
- Enrolment time: 3-5 seconds
- Platform : Windows

2.2 Genex Technologies

Genex Technologies [2] offers a bunch of hardware and software used in 3D face recognition. The 3D Face Cam for Surematch 3D suite is a camera with 3 horizontal lenses. Using these lenses, it constructs a 3D image, the projection of which is 4Mpixels 2D image (High rates).

Figure 2.4: Genex Technologies: 3D Enrollment camera

After capturing 3D images, the system constructs fake images changing some features like lighting, the position and even the face expression (smile, anger...) modifying face muscles. The Genex' series contains for this purpose some software, like 3D sketch match. For the recognition, Genex offers software with its series which is 3D Face Match, a 3D to 3D recognition. Unfortunately, there is no information about algorithms and techniques.

2.3 IMCSecurity

IMCSecurity [3] 3D FaceCam: Coming soon
The main feature is the use of multimodal recognition. The site was under construction, will be available soon.

2.4 Biovisec

Biovisec [4] product: Nemesis 3D. The system gets a 2D image and constructs a virtual 3D image used for enrolment. It is a software solution which enables recognition with glasses, hair and takes into account change in time.
Some performances:

- Tracks 4 people at the same time with 30 images per second, per person.
- Verification done in less than one second

2.5 Geometrix

As used for other research work within the Eurécom Institute, we had to use Geometrix [5] (see fig. 2.5). It is a 2-cameras system, taking a picture for the texture and recombine the both picture to compute a 3D model.

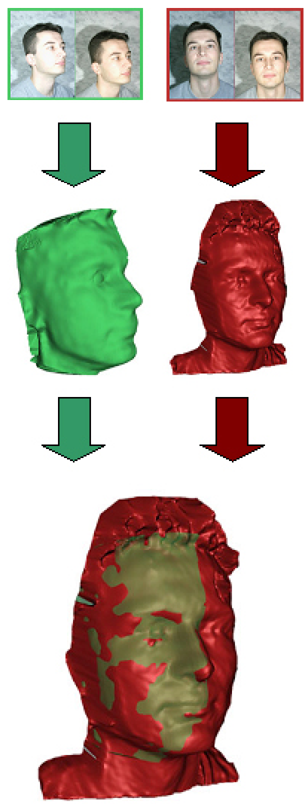

Figure 2.5: Geometrix: 3D Enrollment camera

Chapter 3

Face database

3.1 State of the art

3.1.1 2D face databases

There are many existing 2D face databases. Each has his own characteristics. A lot of them are enumerated on the Face Recognition Homepage [6]. We have done an recapitulating table (see Tables 3.1 and 3.2) contenting the most important used in research.

Different 2D face databases

Base	Nb	Nb pics	pix	Comments
The Color FERET Database, NIST, USA	N/A	N/A	N/A	One of the most used face database
The Yale Face Database	15	165	11	Different facials expressions, glasses. Lighting: front left, right
The Yale Face Database B	10	5760	576	64 illuminations,9 poses (photos)
PIE Database, CMU	68	41368	2236	4 expressions, 43 illuminations,13 poses (photos)
Cohn-Kanade AU Coded Facial Expression Database	100	2000	N/A	23 facial displays, they ranged in age from 18 to 30 years. Sixty-five percent were female, 15 percent were African-American, and three percent were Asian or Latino
Face Recognition Data, University of Essex, UK	395	7900	20	glasses, beards,... male and female subjects, people of various racial origins, first year undergraduate students, so the majority of indivuals are between 18-20 years old but some older individuals are also present.
NIST Mugshot Identification Database		1495 m, 78 f	N/A	front , profile
NLPR Face Database	N/A	450	N/A	27 different pics/person or so unique people under with different lighting/expressions/backgrounds.
M2VTS Multimodal Face Database (Release 1.00)	37	185	5	one week intervals or when drastic face changes occurred in the meantime. During each shot, people have been asked to count from '0' to '9' in their native language (most of the people are French speaking), rotate the head from 0 to -90 degrees, again to 0, then to +90 and back to 0 degrees. Also, they have been asked to rotate the head once again without glasses if they wear any.

Table 3.1: 2D face Databases

Different 2D face databases (Cont.)

Base	Nb	Nb pics	pix	Comments
The AR Face Database, Purdue University, USA	70 m, 56 f	4000		clothes, glasses, make-up, hair style, etc.
Caltech Faces	N/A	450	N/A	27 different pics/person or so unique people under with different lighting/expressions/backgrounds.
The UMIST Face Database	20	564	N/A	Subjects cover a range of race/sex/appearance, range of poses from profile to frontal views
Psychological Image Collection at Stirling (PICS)	N/A	N/A	N/A	N/A
BioID Face DB - HumanScan AG, Switzerland	23	N/A	N/A	N/A
Japanese Female Facial Expression (JAFFE) Database	10 f	213	22	7 facial expressions (6 basic facial expressions + 1 neutral)
3drma	106m 14f	720	6	different orientations, 2 sessions, 3 pics each

Table 3.2: 2D face Databases (Cont.)

3.1.2 3D face databases

When we began our project, we haven't found a lot of 3D face databases. For this reason, we wanted to create a new one. But as one goes along we were researching, we found some new ones and ordered some to apply our system and algorithms.

University of Notre Dame Database [7]

It contains near-frontal range images of 277 individuals. For each individual, there are between three to ten range-images, all taken at different times. The database contains considerable variations in hairstyles of individuals.

MIT database [8]

It contains face images of 10 subjects. It provides two training sets (High resolution pictures, frontal, half-profile and profile view; and Synthetic images (324/subject) rendered from 3D head models of the 10 subjects). The head models were generated by fitting a morphable model to the high-resolution training images. The illumination, pose (up to about 30 degrees of rotation in depth) and the background are varied.

XM2VTS database [9]

It is a large multimodal database. It includes lots of face images, face videos and 3D face data of 295 persons, The 3D data was generated using an special system and was converted to VRML format. It supported by European ACTS projects' Multi Modal Verification for Tele services and Security applications.

The BJUT-3D Large-Scale Chinese Face Database [10]

The BJUT-3D database includes 500 Chinese people. There are 250 females and 250 males in the database. Everyone has the 3D face data. We acquire original high-resolution human 3D face data by the CyberWare 3D scanner in special environment. To build 3D face database, the original 3D data should be preprocessed, and cut the redundant parts. This face database is now made available (named by BJUT-3D-R1) for research purpose only on a case-by-case basis only. The Multimedia and Intelligent Software Technology Beijing Municipal Key Laboratory in Beijing University of Technology is serving as the technical agent for distribution of the database and reserves the copyright of all the data in the database.

GavabDB [11]

It contains 427 three-dimensional facial surface images. 61 persons (45 male and 16 female), and there are 7 different images each person. Each image consists in a three-dimensional mesh representing a face surface. There are systematic variations over the pose and facial expression of each person.

3D RMA [12]

Created in the Signal and Image Center (SIC) of Brussels, The 3D acquisition system was based on structured-light, being constructed using a camera and a projector. It contains 120 individuals.

3D face database of York University [13]

This database has images from 97 individuals. It contains 10 captures per individual including different poses.

Different 3D face databases				
Base	**Nb**	**Nb pics**	**pix**	**Comments**
University of Notre Dame		951		pairs of photographs 2D 3D
MIT-CBCL Face Recognition Database	10	N/A	N/A	N/A
The Extended M2VTS Database, University of Surrey	N/A	N/A	N/A	Each recording contains a speaking head shot and a rotating head shot. Sets of data taken from this database are available including high quality colour images, 32 KHz 16-bit sound files, video sequences and a 3D model.
The BJUT-3D Large-Scale Chinese Face Database	500			No accessories, only a swimming hat to eliminate hair.
GavabDB	61	427		
3drma	106m 14f	720	6	different orientations, 2 sessions, 3 pics each
3D face database of York University	97	1455	15	no effort was made to control the lighting conditions

Table 3.3: 3D face Databases

3.2 Requirements for a new 3D face database

The database will be available on the Internet and have the Eurecom Institute label. To make this successful, we should design a web site with a simple domain name, for ex biometrix.fr, and an ergonomic interface to enable users to download several ”packs” (basic, light, heavy...), like what has been done for the Multimodal IR database Equinox [14].

3.2.1 Database creation

In the chapter 2, we compared in a state of the art several existing databases commonly used in the Research. Unfortunately, most of these databases are 2D, and there are a few 3D, but rarely available. Therefore, our aim is to construct a 3D database containing 3D objects and corresponding 2D image textures.
During the construction, we will apply a mask on each image to delete hair and dark zones.
We will also test its accuracy each time we update.

3.2.2 Choice of People

People are chosen randomly among the Eurecom students (They will just have to sign an authorization).

Subjects :

- 30 people from both sexes.
- between 22 and 25 years old
- from different ethnic origins

3.2.3 Accessories

Subjects will have to get off their hats, glasses because accessories may deform 3D models. The snap shop will be head uncovered.

3.2.4 Shooting

The snap shots will be done in a room with outdoor light; the subject will be in front of the window. Each snap shot will take about four minutes as we have to let the subject sit, to calibrate the camera for the shots, to get the photos and export them. The visage expression has to be Neutral.

For each subject, we will change the following parameters:

Lighting:

Dark A dark snap shot in the multimedia laboratory.

Outdoor- A dark snap shot with only the outdoor light.

Outdoor+: A snap shot using outdoor light and camera's spotlight.

Time :

The snap shots will be done at different times, one in the morning, one at the midday pause, and the last one at the end of the afternoon. A time-slot will be reserved for each type of lighting.

Special :

It will be one snapshot taking into account different features of a person: for example a man who did not shave.

3.2.5 Nomenclature:

At the beginning of the base construction, each subject will have his/her number for the entire construction. We will name the different lightings with 1, 2, 3 for dark, outdoor-only and spotlight and 4 will stand for the special photo. Each session will have its number and finally we will add to the name the gender of the subject.
For example, the third snapshot for the tenth person (man) with an outdoor only, the name will be as follow:
2_10_3_m.jpg

3.2.6 Distribution

The database will be free to use and download via a web page, as we said before. Users will be able so select some parameters to have a personalized version of the database(lights, parts or complete database, . . .) .
For example, a person may choose to download the snapshots with outdoor light only.

3.2.7 Planning

We plan to to snap each two weeks. Due to the camera problems (see Appendix B) , we had to cancel the snapshots.

3.2.8 Conclusion

At the moment, there are few 3D databases. It will be a good opportunity for the Eurecom Institute to make its own database, accurate, available and free to use.

Chapter 4

Fusion algorithms

4.1 State of the art

The term *information fusion* encompasses any area which deals with utilizing a combination of different sources of information, either to generate one *representational format*, or to *reach* a decision.
When looking from the point of decision making, there are several motivations for using information fusion:

- Utilizing complementary information (e.g., audio and video) can reduce error rates.
- Use of multiple sensors (i.e., redundancy) can increase reliability.
- Sensors can be physically separated, allowing the acquisition of information from different points of view.

In our project, the aim of fusion is to combine complementary information, 2D and 3D features.
In this part, we will make a review and a classification of all the fusion techniques at the moment, after that we will present in depth the methods we have implemented in our Matlab code.
We can classify information fusion techniques into two main categories, the pre-mapping fusion and the post-mapping fusion.
The pre-mapping consists of combining information before any use of classification (the classification in our project is to recognize, computing distances from models), while the post-mapping fusion consists of combining information after classifications, into the decision/opinion space (after computing the distances for example).
In each of these categories there are several methods that we will develop and explain, giving some examples.

4.1.1 Pre-mapping fusion

Acquisition fusion:

The aim is to construct a unique model after the acquisition, for example a weighting model of an infrared image and a texture image. After getting this model, we extract the needed features and train and enrol the system, and finally have a recognition based on this scheme.
The main drawback of this fusion method is the sensor. As we want to construct a unique model, just after the acquisition, we have to get the information from the same sensor, to be able to have commensurate data.
In our project, the acquisition is made with Geometrix cameras, two cameras that take two pictures and construct a 3D object and a 2D image, standing for the texture. As this data is not commensurate, we cannot use this technique, but it might be an interesting method if, for example, we had a grey-scale depth image during the acquisition.

Feature fusion:

The feature fusion is a fusion done *before* the training. After the acquisition, we extract the most important features, using different extraction techniques from the data that we have. Then, we combine these features to construct a vector model, either by using a weighting scheme if they are commensurate, or by some concatenation technique if they are not. [17, 18, 19, 20].
After this step, we can train, do the enrolment and the recognition.
For example, in our project, we have to find out what are the important features of the 3D model and the 2D model, extract them (some zones in the images) and construct a vector that will stand for our model.
This method has several drawbacks the most important of which is that we cannot know how each feature contributes in our final decision [20].

4.1.2 Post-mapping fusion

We can also call this category *decision fusion*. We first use different classifiers to classify the models that we have, for example in our project, we do the recognition with the 3D models and the 2D models apart, and combine the scores of these recognitions. The inspiration of the post-mapping fusion is to overcome the bad properties of classifiers.
There are several subcategories in the decision fusion which are: the operators AND, OR, ranked list, majority voting and a weighting scheme.

AND Fusion:

The decision is made only if each classifier reaches the same decision, if they all agree. It means that recognition (identification or verification) is made only if we recognized the same subject with the 3D and the 2D recognitions.

OR Fusion:

In OR fusion [22, 23], a decision is made as soon as one of the classifiers makes a decision. In comparison to AND fusion, this type of fusion is very relaxed, providing multiple possible decisions in multi-class problems. Since in most multi-class problems this is undesirable, OR fusion is mainly useful where one would like to detect the presence of an event/object with a low false rejection bias (in a person verification scenario, where we would like to detect the presence of a true claimant, this translates to a low FRR and high FAR).

Majority voting:

In this method, we have some classifiers that each one makes a decision. The combination consists of taking into account the majority selected decision. The drawback of this decision is that we must have more than two classifiers. We can use this method in 3D recognition combined with 2D texture and infrared recognition.

Ranked lists combination:

In ranked list combination [24, 25, 26], each classifier provides a ranked list of class labels, with the top entry indicating the most preferred class and the bottom entry indicating the least preferred class. The ranked lists can then be combined via various means [25], possibly taking into account the reliability and discrimination ability of each classifier. The decision is then usually reached by selecting the top entry in the combined ranked list.

Weighting fusion:

The idea of the weighting fusion is to combine the scores of each classifier by doing a weighting summation, and after that reaching a decision with the MAX or MIN factor. For example, we combine all the distances of the 2D and the 3D model with some weighting scheme and select with the MIN factor, as we want to get the nearest model from the subject we want to recognize.

The weighting is done to weight a classifier more than another. There are several methods to select the weight factors: having the mean value, the Chang-Bowyer fusion, the alpha-weighting scheme. We can also use several methods of computing the distances and combine them: computing an Euclidian distance, Cosine distance, Lp distance or Mahalanobis distance and weight them.

4.2 Fusion algorithm implementation

4.2.1 EDF-Eurecom database

We use the 3D database EDF-Eurecom in all our fusion tests. We built (trained) the spaces with the Eurecom faces (134 faces) and enroll and test with the EDF faces (250 and 50 faces resp.).

4.2.2 Normalization

In our project, we are using recognition with 2D and 3D, which are different spaces using different models. Each recognition is based on a computing method; the choice of the distance, and the results are given by each classification are not commensurate. That's the purpose of Normalization: to get the score of each recognition and bring it to the same space enabling us to compare them, to make them commensurate and to combine the scores with information' fusion method.
We found a way to normalize the results : The idea is to apply an affine transformation on the distances and have them in [0, 1] interval.
Let's consider a 2D recognition. We have a texture image of a person to recognize. We project this image on the space and compute its distance from all the models enrolled in this space, which are saved in an array Distance. Consider d_{min} and d_{max} the minimal and the maximal distance in this array. Each distance is exactly in the $[d_{min}, d_{max}]$ interval. We will now apply an affine transformation f, with the interval $[d_{min}, d_{max}]$ and the image interval [0, 1].
f is defined as follow:

$$f(d_{min}) = 0$$
$$f(d_{max}) = 1$$

$$f(x) = \frac{1}{d_{max} - d_{min}} x - \frac{d_{min}}{d_{max} - d_{min}}$$

Thanks to this normalization, we were able to use the weighting fusion scheme and code it into MATLAB. Details are shown later. Before doing any fusion, we first have applied the recognition of the 2D and the 3D database separately to have an idea about recognition rates. With the Eurecom/EDF database we had found a recognition rate under 80% for the 2D (see fig. 4.1) and above 80% for the 3D (see fig. 4.2), using all the eigenfaces.

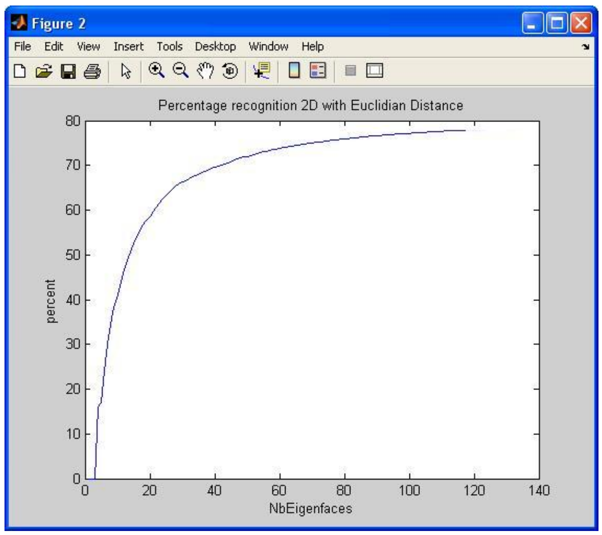

Figure 4.1: Recognition rate for 2D with Eurecom/Edf database

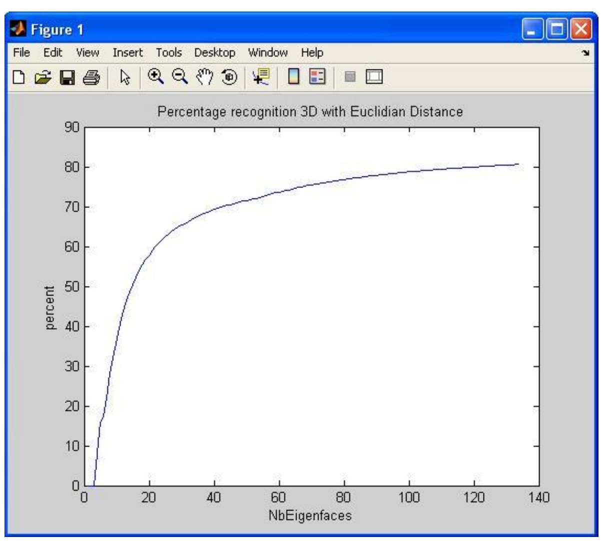

Figure 4.2: Recognition rate for 3D with Eurecom/EDF database

We think that these low rates are linked to the accuracy of the database. When we use the Feret Database, we get better recognition rates, below is the graph of different rates for a certain number of eigenfaces. We see that we reach a recognition rate about 80% with only 10 eigenfaces, and a recognition rate that stays at about 88%. (see fig. 4.3).

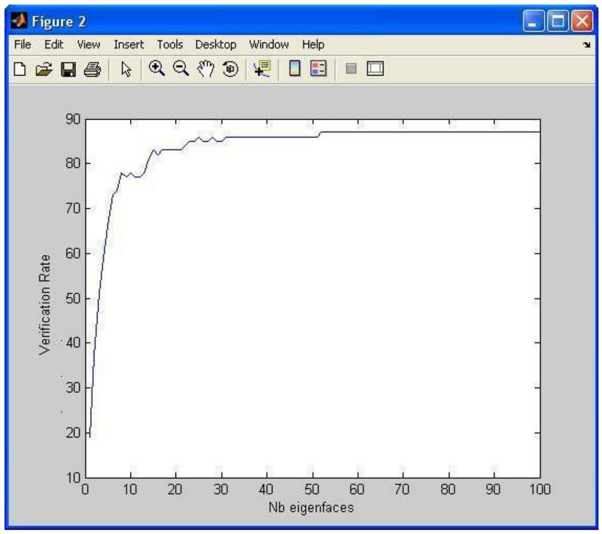

Figure 4.3: Recognition rate with FERET database

For all the following tests, we used a removing eigenfaces technique. See paragraph 5.5

The second part of our strategy is to test, as introduced above, different distances. We computed the mean results of four different distances with the Eurecom/EDF database to select which are the best ones.

4.2.3 Distance calculation

We tried several distances because the choice of a distance may improve the results as described in this paper [28]. We implemented all the following distances. The results are available in the Chapter 4

Euclidian

$$d(X,Y) = \sqrt{\sum_i (X_i - Y_i)^2}$$

We remark that we reach a recognition rate about 92% for the Euclidian Distance (see fig. 4.4).

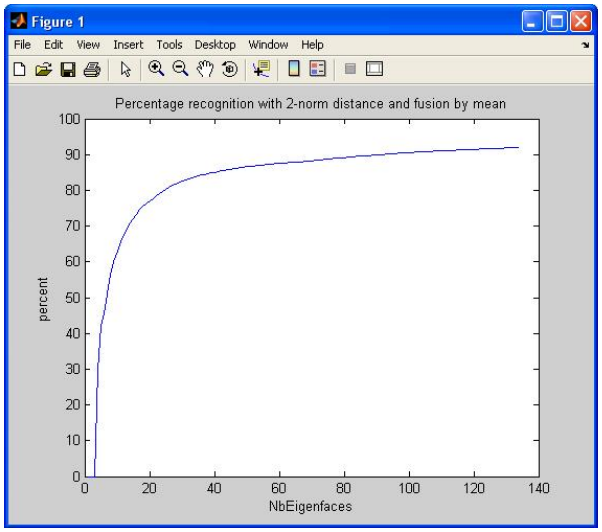

Figure 4.4: Recognition rate with Euclidian distance (mean fusion)

p-norm

$$d(X,Y) = \left(\sum_i |X_i - Y_i|^p \right)^{\frac{1}{p}}$$

The recognition rate for the 1- and 3-norm Distance (see fig. 4.6) is about 85%, which is better than using 3D or 2D alone.

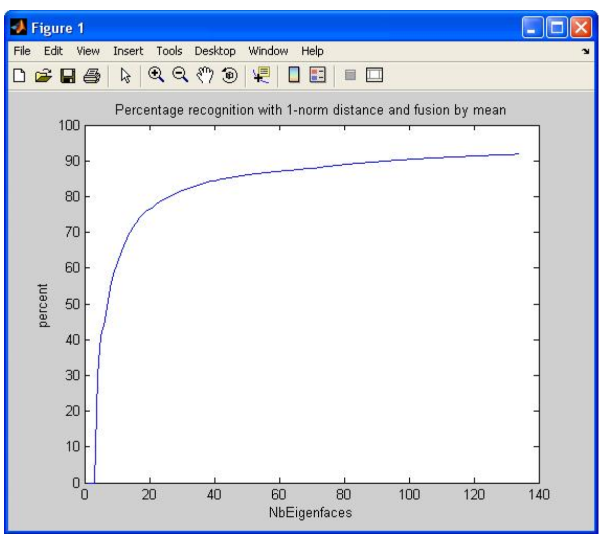

Figure 4.5: Recognition rate with 1-norm distance (mean fusion)

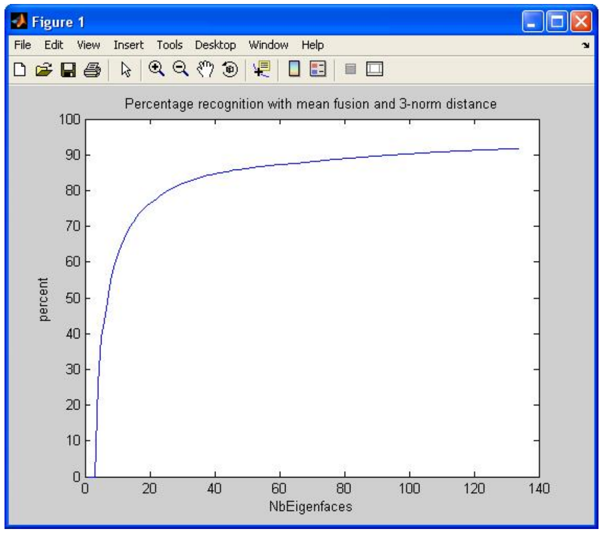

Figure 4.6: Recognition rate with 3-norm distance (mean fusion)

Cosine normalized

$$d(X,Y) = -\frac{\sum_i (X_i Y_i)}{\|X\|\|Y\|}$$

The cosine distance leads us to worse results than the other distances, because it doesn't take into account the norms of the vectors. It computes only the cosine of the angle between them, alignment only.

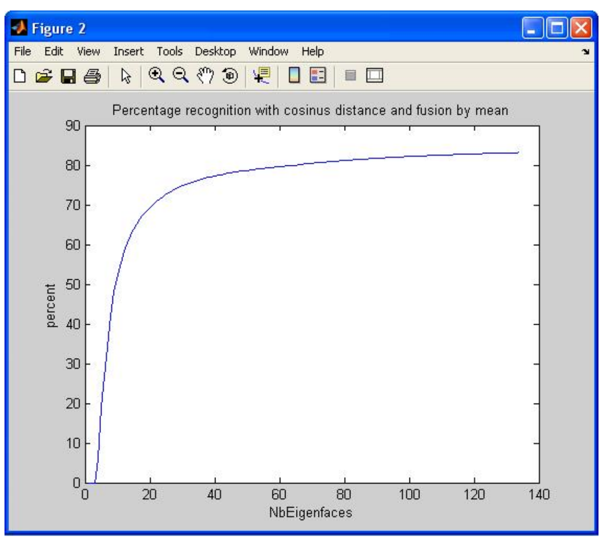

Figure 4.7: Recognition rate with cosines distance (mean fusion)

Mahalanobis Distance

$$d(X,Y) = -\sum_i \frac{1}{\sqrt{\lambda_i}}(X_i Y_i)$$

The recognition rate for Mahalanobis is about (see fig. 4.8) 84%, which is better than using 3D or 2D alone. We introduced this distance because it lead to significant increase of identification rates with FERET database [28] energy (see fig. 4.9).

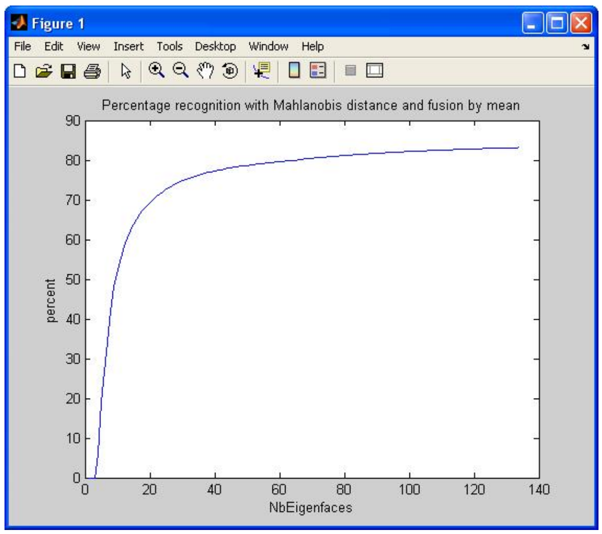

Figure 4.8: Recognition rate with Mahalanobis distance (mean fusion)

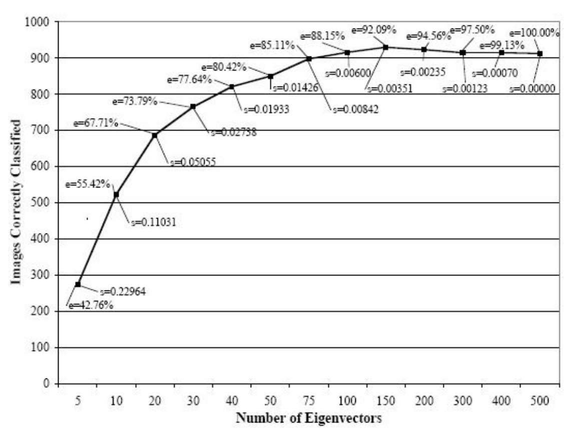

Figure 4.9: Energy and stretching dimension on FERET

We remark that the d-norm distance (d=3) leads to the best results and consequently we applied fusion techniques on this distance. In the following paragraphs, we will present three weighting scheme, and the results using them with the d-norm distance.

4.2.4 Fusion with the mean

The first step is to do the 3D and 2D recognitions separately and get the arrays of distances. We normalize these distances and compute their means and finally keep the minimum normalized mean distance as the best score. This method gives us the following graph. We can notice that we have better results than with 2D and 3D only (see fig. 4.4). We reach a recognition rate about 92% .

4.2.5 Chang-Bowyer

The second method used for the fusion is the one described in the Notre-Dame article [16]. It uses the following weighting:

$$weight = \frac{2^{nd}dist - 1^{st}dist}{3^{rd}dist - 1^{st}dist}$$

If this ratio gets bigger, the first result in the selected classifiers corresponds to the actual recognized subject. We see that the results (see fig. 4.10) are better with this weighting scheme using the Eurecom/EDF database.

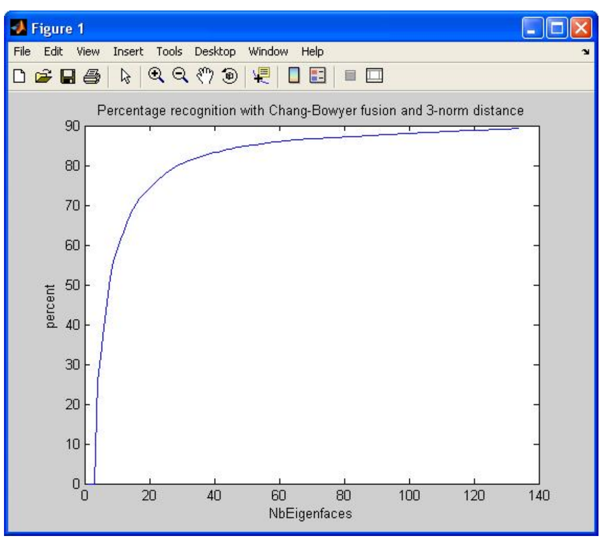

Figure 4.10: Identification rate with Chang-Boywer fusion method

4.2.6 α-weight Method

The idea is to test the recognition using an α weighting to have an optimal result.

We will try different α coefficients during this test to find which one leads to an optimal result.

$$MinDist = \alpha Dist1 + (1 - \alpha)Dist2$$

To get the best α, we test different values between 0 and 1, with a step of 0.01. After that, we test the recognition on known results. We found the optimal weight (see fig. 4.11) for EDF-Eurecom: $\alpha = 0.7$ for texture, which means that the information included in the texture is more important but the contribution of the 3D information is significant, otherwise, we would have found $\alpha = 1$.

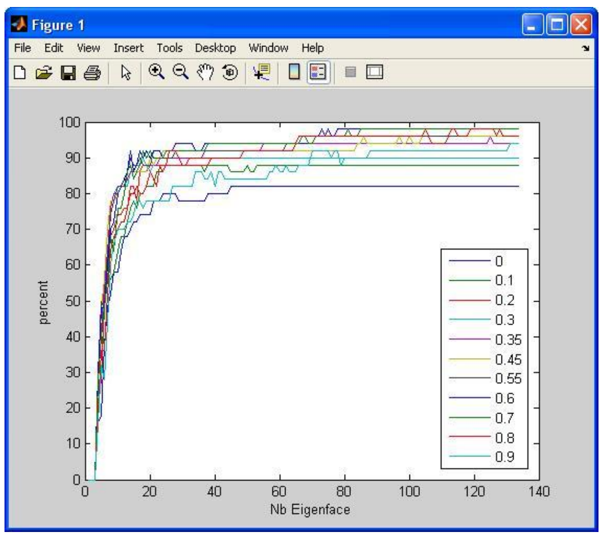

Figure 4.11: Curves of identification rates with different values for alpha

4.2.7 Tables of results

Here we summarize the results we obtained after several tests
We can see that the best results are reached with the heuristical method of weighting using the 3-Norm distance. This fusion technique lead to a significant increase of our recognition rates, the 2D and 3D are below 86%, and their fusion is above 93%.
The combination shows that 2D contributes more than depth image as its weight is 70%.

4.3 Conclusion

Our fusion was based on the decision combination. In all our tests, we have reached better results than using 2D or 3D recognition only. These results showed that the

	Euclidian Distance	1-norm	3-norm	Cosine	Mahlanobis
Eigenface					
2D only	81	86	86	85	91
3D only	78	81	82	84	88
Mean fusion	91	91	91	82	83
Chang-Bowyer fusion	90	88	89	86	88
Alpha Fusion (alpha = 0.7)	92	91	93	87	89

Table 4.1: Best Results in percent

contribution of 2D is more important than 3D's, as its weight (70%) is higher during the decision.
But, even if 3D's weight is less important, its contribution to 2D recognition heightened the identification rate from 86% to 93%.
A lead would be to test our code on other databases to find out if this weighting scheme is the best combination.

Chapter 5

Coding work

5.1 Introduction

This project is based on the Roberto Tron's " Face recognition Matlab system using eigenfaces " [15]. It's a graphical application that enables space construction, people enrolment, identification and verification using eigenfaces.
The starting idea was to modify this software in order to execute the same functions for a multimodal 2D-3D recognition.

5.2 Modifications

Our starting point was an existent work, and thus, the first step is difficult; understanding the global structure and each minor detail. For this reason, we began the work by reading the Roberto Tron's report, followed by testing the application.
To make the software support a multimodal recognition, we had to locate each part of the code needing a doubling. The task was drawn-out and perilous but after many attempts, here are the main modifications:
First, we added a "Color to Black and White" method because the EDF-database used has been developed supporting color images. Then, we had to create a second space, standing for gray scale depth images, and all its attached parameters.
We finally doubled each variable used for the space, such as means, eigenfaces etc, (for example: Space.mat became Space1.mat and Space2.mat).
Roberto Tron has used a system of structure to save some values corresponding to the Model and Space (such as the number of eigenfaces, the scale factor, ...). We add to these parameters two other ones : the eigenvalues, to make them available everywhere, and an identifier (1 or 2).
The enrolment consists of loading the directory containing images to enroll and put them, in an alphabetical order, into a matrix. Each image has its index corresponding to its alphabetical position in the directory.
To recognize people (i.e. to know which person has been recognized) we use this index.

When we project a subject image on the eigenspace to recognize it, we associate it with an index, that corresponds to the index of the same subject in the enrolment matrix.
We compute the distance between this projection and all the models in the space. When the index of the nearest model is the same as the index of the subject, it's a successful recognition, otherwise it's a failure.
We do the same "indexation" to the 2D and 3D. In fact, we had to recombine results for a multimodal recognition. Thus, we can know which subject has the best score for the texture and which one for the depth, and the fusion.
We also add several fusion methods, implement some distances in order to calculate the recognition scores.

5.3 GUI

The software is called by the command: `> startmain`.
The main choice menu (see fig. 5.1) gives access to each sub-menu, almost the same than the Robert Tron's ones:

- Base construction
- System enrolment
- Examples
- Identification
- Verification

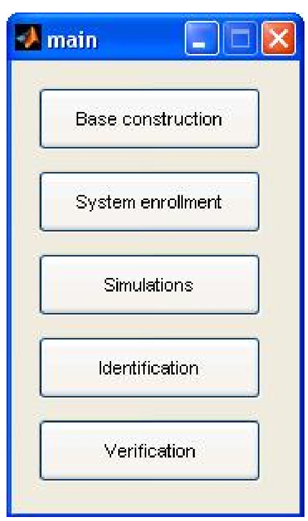

Figure 5.1: Main choice menu

Each button launches the module relative to a step in the recognition or identification process:

Base construction
This module (see fig. 5.2) creates two bases (eigenfaces) for two face spaces from a set of images (for example: texture and depth) and save them in the files space1.mat and space2.mat.

- Texture / 3D-Shape database directory: select a directory containing the images to use for the estimation of the face space
- Specify number of bits for images (input and mean) (unchanged)
- Specify number of bits for space representation (unchanged)
- Type of overflow (unchanged)
- Resize images (unchanged)
- Save Config.: saves all the parameters in config.m.
- Save Config. and Build Space: saves all the parameters in config.m, build the space (i.e. the /Matlab/CommandLine/ spaceFromImages.m script is called) and saves it to the file space.mat.

System enrollment
This module (see fig. 5.3) trains the system by computing users models (Models1 and Models2) using the space files (created with the previous step) and a set of users' images. The models of user are composed by the mean of all his/her images projected into the space spanned by eigenfaces. The results are stored in the files Models1.mat and Models2.mat

- Texture / 3D-Shape database directory: select a directory containing the images to use for the computation of the users models in the face spaces.
- Texture / 3D-Shape Space file select the file containing the face space base (created with the appropriate module).
- Number of images for each user (unchanged)
- Specify number of bits for input images (unchanged)
- Save Config.: saves all the parameters in config.m.
- Save Config. and Build Space: saves all the parameters in c onfig.m, build the space (i.e. the /Matlab/CommandLine/ spaceFromImages.m script is called) and saves it to the file space.mat.

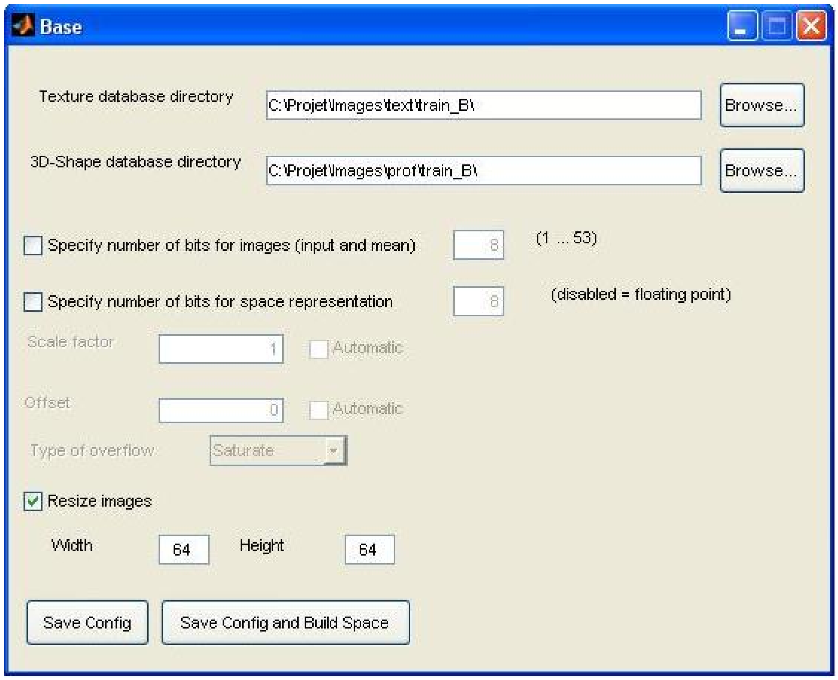

Figure 5.2: Base choice menu

Best identification
This module (see fig. 5.4) is used to find the best user model corresponding to a test image, for both tested images, and display them. After choosing test image, space and models files, the program computes the best matching (nearest) model present in the database and display the image corresponding to it. The input image can be scaled and clipped before the identification as Roberto Tron's explanations.

- Texture / 3D-Shape file: select the test images for the identification
- Texture / 3D-Shape Space file: select the file containing the face space base (created with the appropriate module).
- Texture / 3D-Shape Model file: select the file containing the face model (projection) (created with the enrollment module).
- Test image: shows a preview of the image which will be used for the identification. It reflects each change in the parameters (x and y offsets, scale, number of bits).
- x offset, y offset: (unchanged)
- Scale: (unchanged)
- Number of eigenfaces: (unchanged)

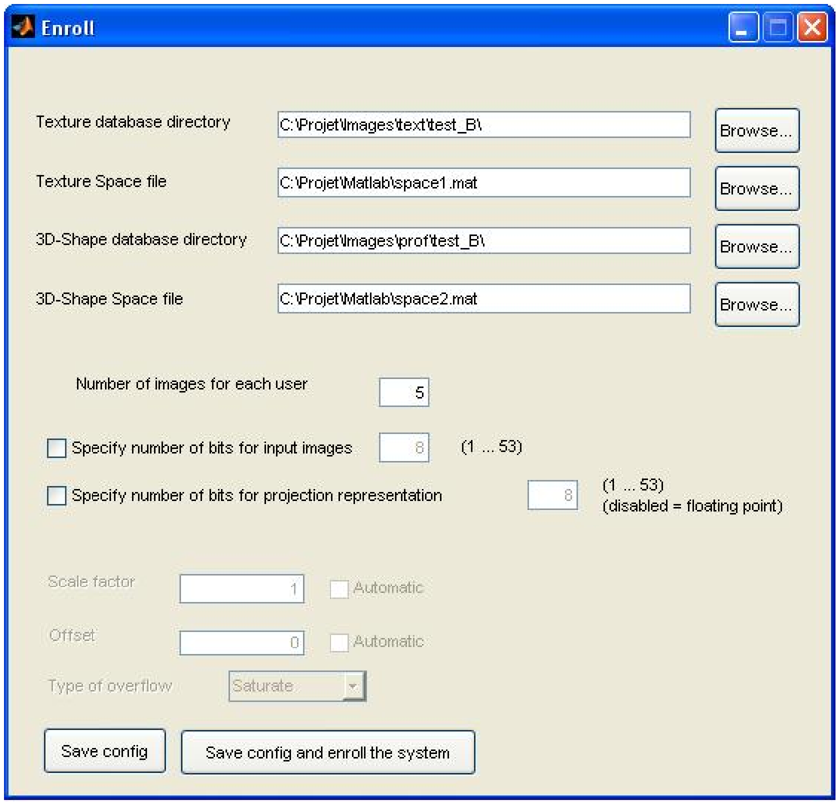

Figure 5.3: Enroll choice menu

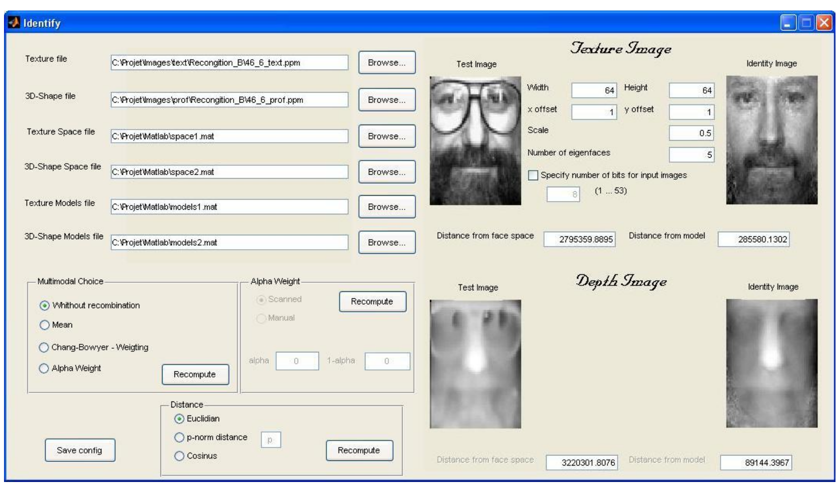

Figure 5.4: Identification choice menu

- Specify number of bits for images (input and mean) (unchanged)
- Distance from space: read only, shows the distance between the projection of the test image the face subspace.

- Identity image: the image reconstructed from the best model by adding the products between projections and eigenfaces.
- Distance from model: read only, shows the distance between the projection of the test image and the best matching (nearest) model.
- Multimodal Choice

Multimodal Choice

The distances from the model are saved in a vector. The first value is the first face, the second too, ... This is done for the texture and the depth map. We can use these values for the recombination. We can choose in this part of the module (see fig. 5.5) between four different fusions:

- Without recombination
- Mean fusion
- Chang-Bowyer
- Alpha-weighting: the alpha-weight could be modified by the user

We can also choose between four distances calculations:

- Euclidian
- p-norm
- Normalized cosines
- Mahalanobis

Verification

This module (see fig. 5.6) is almost the same than the Identification's one, but here we take account of impostors. When an impostor tries to be recognized, the baboon's picture appears.

5.4 Useless Eigenfaces

In several articles [28], we have found that removing first eigenfaces, with the highest eigenvalues, lead to an improvement of the recognition rates, because they represent the illumination differences. We also found that we may remove the last eigenfaces, with the lowest eigenvalues, that correspond to noise and do not increase the recognition rates. There are several methods to remove the last eigenvalues.

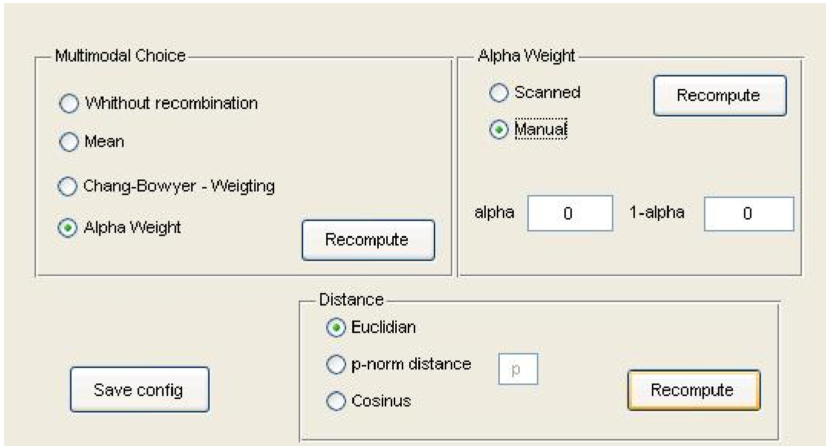

Figure 5.5: Multimodal choice submenu

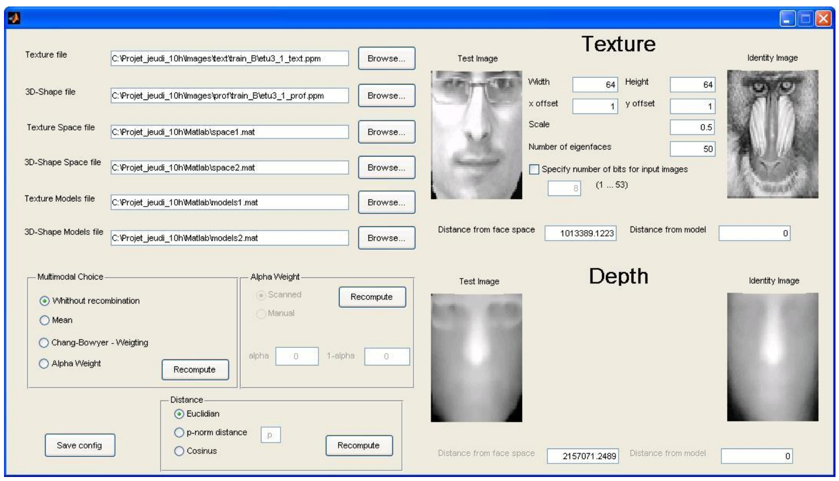

Figure 5.6: Verification choice menu

5.4.1 First Eigenfaces

The following graph show the percentage evolution removing the first Eigenfaces (see fig. 5.7). This is a heuristic method: we found that the threshold is to remove the fourth first coordinates corresponding to the fourth first eigenfaces.

5.4.2 Last Eigenfaces

The last eigenfaces [28] correspond, as we said to noise. They do not contribute in a significant increase of the recognition rates, they only add useless computational time. This is the main reason that lead us to look for a removal method.

We found different ideas to remove the last Eigenvectors. The first of these variations

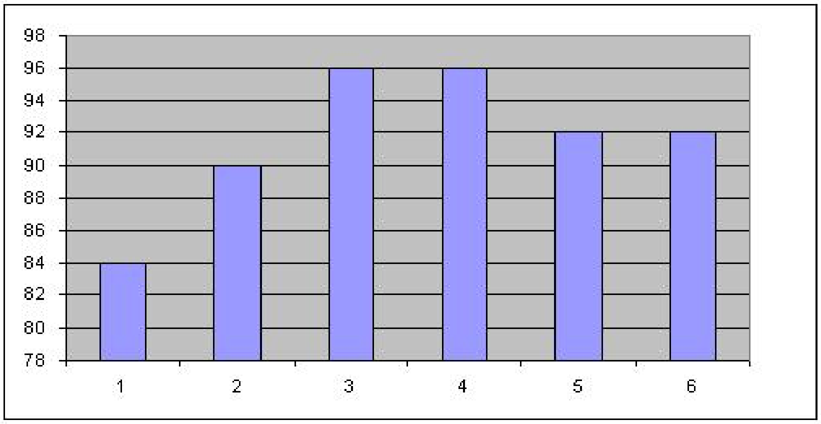

Figure 5.7: Identification rate removing the 1st to the 6th eigenface

removes the last 40 % of the Eigenvectors [27]. This is a heuristic threshold selected by experience, but in our experiences, when we removed only the 40 %.
The second scheme [28] uses the minimum number of Eigenvectors to guarantee that energy e is greater than a threshold. A typical threshold is 0.9. e_i is defined as the energy of the i^{th} Eigenvector:

$$e_i = \frac{\sum_{j=1}^{i} \lambda_j}{\sum_{j=1}^{k} \lambda_j}$$

This formula shows that the last eigenvalues do not have an important weight and thus may be removed. The energies of eigenvalues, which are above the 0.9 threshold, do not have a *significant weight* and consequently the coordinates linked to corresponding eigenfaces may be not token into account.

The second method is to compute $\frac{\lambda_i}{\lambda_1}$, with a threshold of 0.01. This means that the lowest eigenvalues, comparing to first ones, have not an important weight and may then be removed.

5.5 Eigennose

We decided to test how does nose information contribute in the face recognition. For this purpose, we apply eigennose recognition, nose recognition. We extract a zone containing the nose. This was possible because the base is normalized, all the noses are in the middle of texture and depth images.

5.5.1 Nose extraction

We make the extraction using a free software, photofiltre (www.photofiltre.org). This software enables to cut the images in a directory with the same method. The snap shots below shows the extraction of the nose from a texture image (see fig. 5.8).

After this step, we export these images to corresponding directory /Depth/Enrollement for example.
Unfortunately, this software exports the images in different format , but not to .ppm.
We had to find a solution to convert automatically nose .jpg image to .ppm images, supported by our recognition software.
ImageMagick is an open source solution enabling several transformation on photos, especially conversion, using Windows Command or Shell scripts.
Here is the CMD script applied to the directories containing all the nose images in .jpg format.

```
> for %f in (*.jpg) do convert %f %~ nf.ppm
```

5.5.2 Nose recognition

After nose extraction, we test the recognition using PCA algorithm, eigennoses.
We proceed in the same way as we have done with eigenfaces, space construction, enrolment, recognition and useless eigennoses removal.

Eigennoses useless

We remove heuristically the first eigennoses as we did for eigenfaces, by computing the recognition rates and their evolution. We found that we could remove the first five eigennoses (see fig. 5.9).

5.5.3 Eigennoses texture and depth

First we tested the texture and the depth separately to compare with the Eigenfaces(see fig. 5.10 and 5.11).
We can notice that the results are good. We have almost 85 % recognition for 3D shape.
This leads us to apply some fusion techniques, as follow :

5.5.4 Eigennoses with mean fusion

We have a better result, about 85% (see fig. 5.12). It is almost the same than the 3D-shape result.

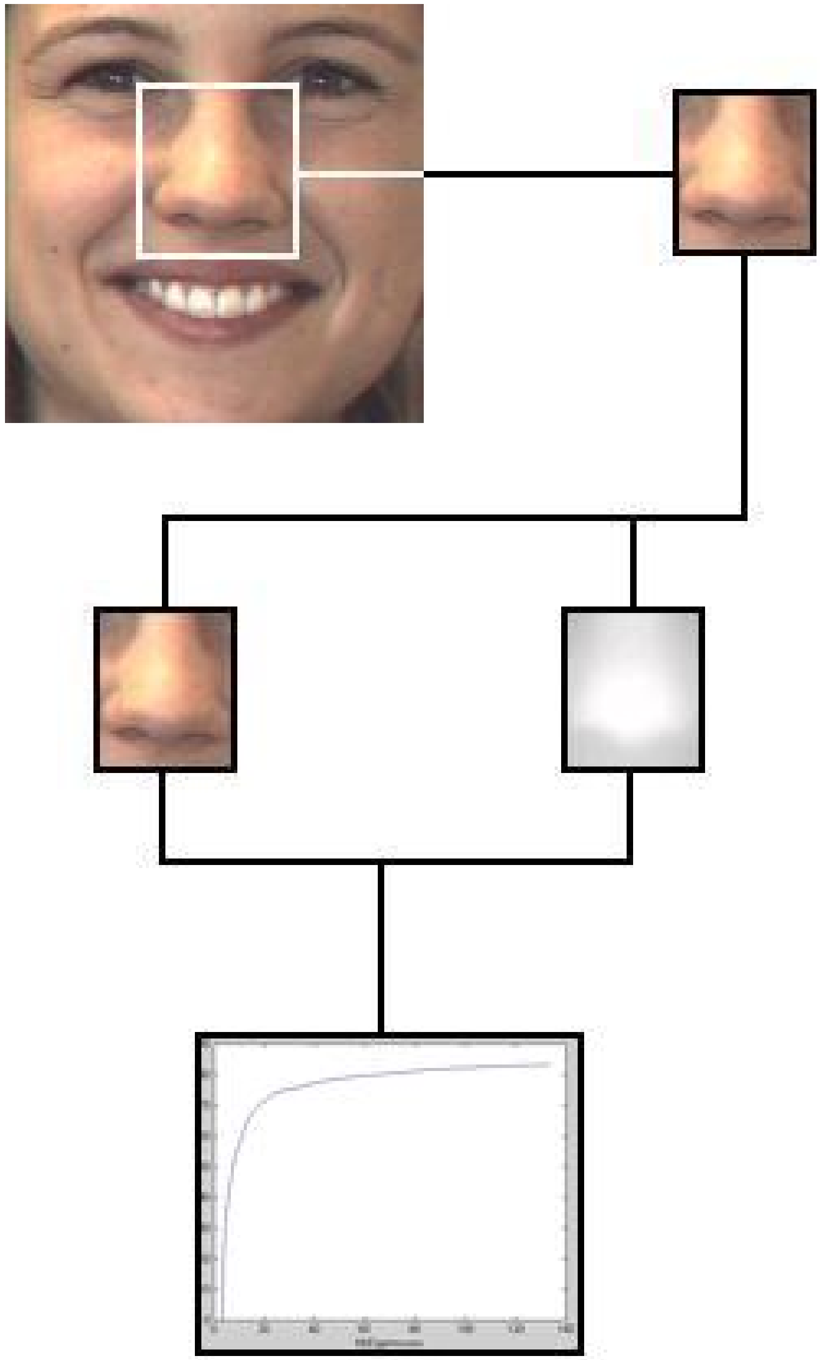

Figure 5.8: How we cut the nose

5.5.5 Eigennoses with Chang-Boywer fusion

We have worse results, about 80% (see fig. 5.13).

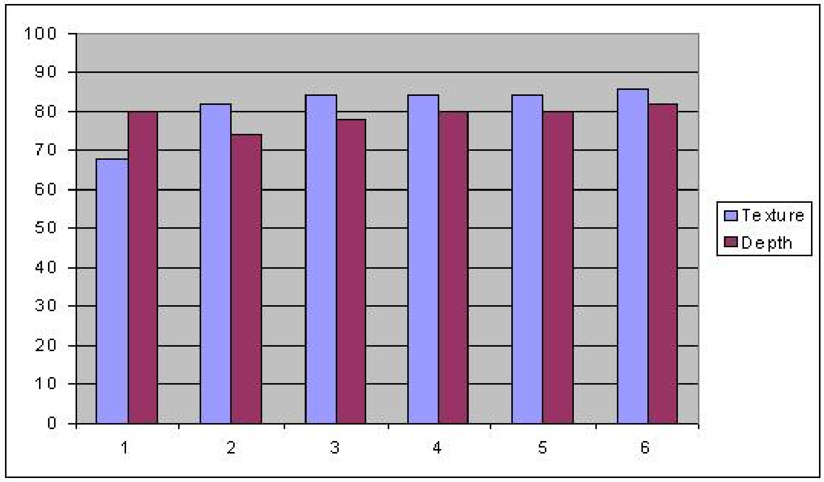

Figure 5.9: We can see we can remove the first five Eigennoses

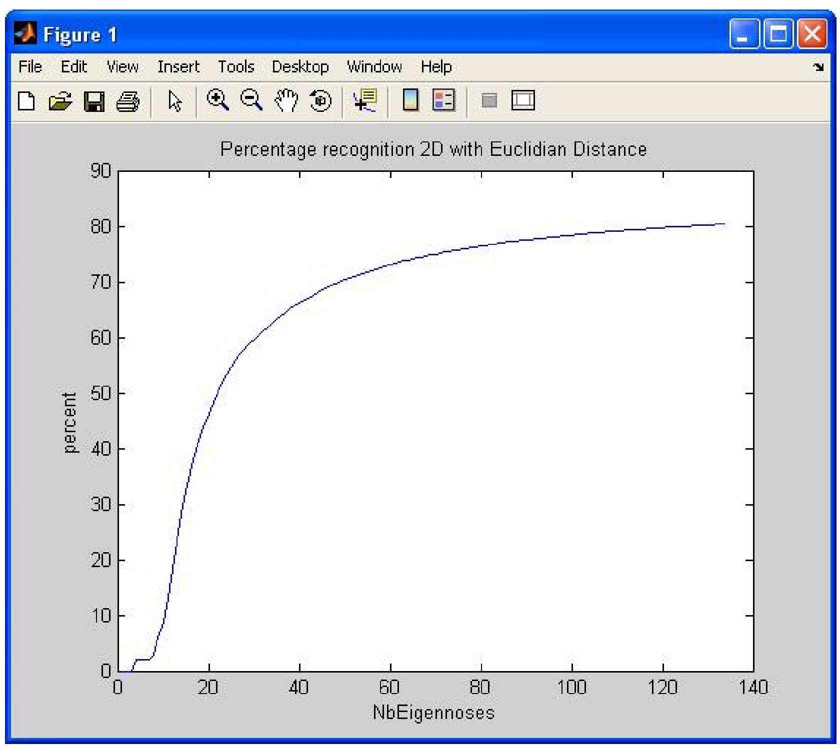

Figure 5.10: Percentage recognition for texture only

5.5.6 Eigennoses conclusion

The eigennose algorithm shows that the nose only lead to quite good results. We think that a way to improve the face recognition is to use extraction of local features, such as noses, eyes, mouth, which contain significant information about subjects.

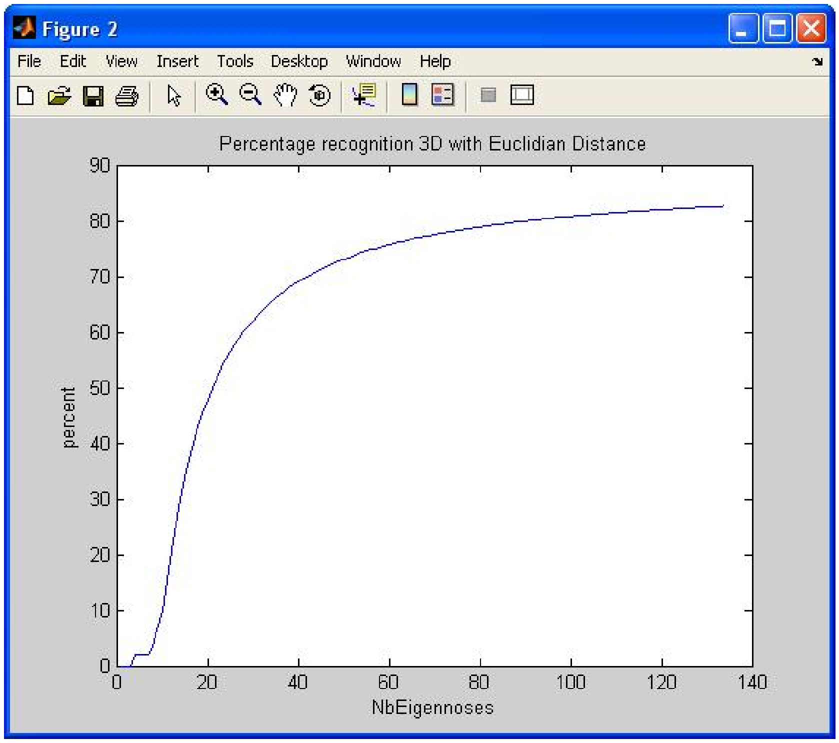

Figure 5.11: Percentage recognition for depth only

We may use a Feature fusion technique (see paragraph 4.1.1) by finding a solution to combine these data, or to use Decision fusion by doing eigeneyes, eigennoses and eigenmouth recognition, and combine the results.

5.6 Problems

First of all, many problems occurred because of a move from Linux to Windows. Secondly, the EDF-database is a poor-quality one, so it is difficult to have a good identification. Finally, it was very difficult to improve the program because has been not though to be amended.

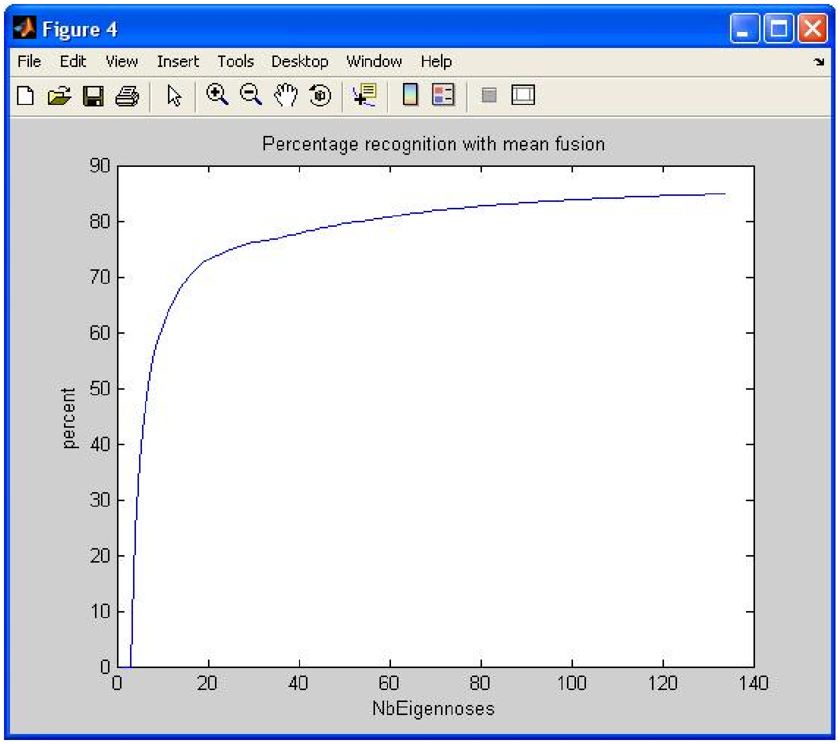

Figure 5.12: Percentage recognition with mean fusion

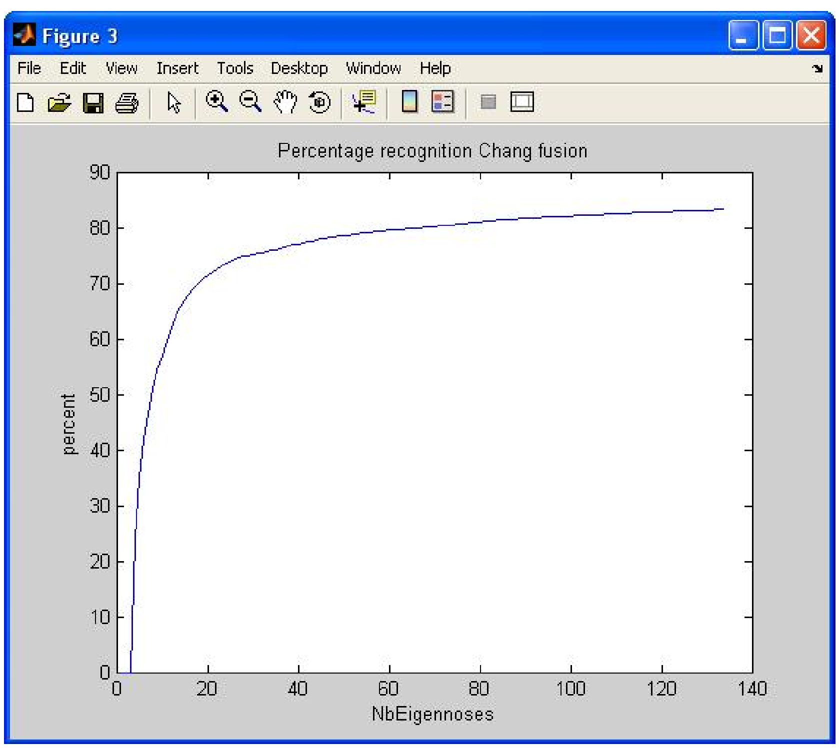

Figure 5.13: Percentage recognition with Chang-Boywer fusion

Chapter 6

Conclusion

In Chapter 2, we have seen that there are more and more 3D databases that available for the research. This leads to think that recognition based on 3D is an interesting field.

In Chapter 3, we have proven that the fusion of 2D and 3D information is much more efficient than the use of a unique mode recognition. In our experiments, we used a decision fusion technique based on a different weighting scheme, and we remarked that the identification rate increases with 7% more than each mode.

A lead would be the use of several distances and possibly combine them. In fact, each distance takes advantage of a particular feature, for example the Cosine with the angle, and the Mahalanobis tries to make all the vectors at the same level.
The combination of these distances may then improve the recognition by using a weighting fusion as showed in [28].
We may also think of applying different distances on each mode, for example Euclidian distance on the 2D and the 3-norm on the 3D. These distances has to be the ones that optimize results on each mode.

Our tests on eigennoses lead us to think that a feature fusion may give better results. Indeed, with eigennoses recognition we obtain about 85% identification rate.

Nevertheless, during our research, we have never found comparisons between fusions. One of our goals was to provide an evolutive system to compare different recognition techniques. It is easy to improve and add new fusions techniques. This system could be considered as an independent tester for new methods, and able to provide comparisons in the same ”referential”.

Bibliography

[1] www.a4vision.com

[2] www.genextech.com

[3] www.imcsecurity.com

[4] www.biovisec.com

[5] www.geometrix.com

[6] www.face-rec.org

[7] www.nd.edu/ cvrl/UNDBiometricsDatabase.html

[8] cbcl.mit.edu/software-datasets/heisele/facerecognition-database.html

[9] www.ee.surrey.ac.uk/Research/VSSP/xm2vtsdb/

[10] www.bjpu.edu.cn/sci/multimedia/mul-lab/3dface/overview.htm

[11] gavab.escet.urjc.es/gavab/recursos_en.html

[12] www.sic.rma.ac.be/˜ beumier/DB/3d_rma.html

[13] www-users.cs.york.ac.uk/˜ tomh/3DFaceDatabase.html

[14] www.equinoxsensors.com/products/HID.html

[15] Roberto Tron, Face recognition system using eigenfaces (with Matlab and C++), September 2005, Eurécom Institute

[16] Kyong I. Chang Kevin W. Bowyer Patrick J. Flynn Xin Chen, Multi-biometrics Using Facial Appearance, Shape and Temperature, May 2004, Notre Dame University

[17] A. Adjoudani, C. Benoit, Audio-visual speech recognition compared across two architectures, in: Proc. 4th European Conf. Speech Communication and Technology, Madrid, Spain, vol. 2, 1995, pp. 1563-1567

[18] D.L. Hall, J. Llinas, Multisensor data fusion, in: D.L. Hall, J. Llinas (Eds.), Handbook of Multisensor Data Fusion, CRC Press, 2001, pp. 1-10.

[19] J. Luettin, Visual speech and speaker recognition, PhD thesis, Department of Computer Science, University of Sheffield, 1997

[20] A. Ross, A. Jain, Information fusion in biometrics, Pattern Recogn. Lett. 24 (2003) 2115-2125.

[21] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, Wiley, 2001.

[22] R.C. Luo, M.G. Kay, Introduction, in: R.C. Luo, M.G. Kay (Eds.), Multisensor Integration and Fusion for Intelligent Machines and Systems, Ablex Publishing Corporation, Norwood, NJ, 1995, pp. 1-26.

[23] P.K. Varshney, Distributed Detection and Data Fusion, Springer-Verlag, New York, 1997.

[24] B. Achermann, H. Bunke, Combination of classifiers on the decision level for face recognition, Technical report IAM-96-002, Institut für Informatik und angewandte Mathematik, Universität Bern, 1996

[25] T.K. Ho, J.J. Hull, S.N. Srihari, Decision combination in multiple classifier systems, IEEE Trans. Pattern Anal. Machine Intell. 16 (1994) 66-75.

[26] V. Radova, J. Psutka, An approach to speaker identification using multiple classifiers, in: Proc. IEEE Conf. Acoustics, Speech and Signal Processing, Munich, vol. 2, 1997, pp. 1135-1138.

[27] H. Moon and J. Phillips. Analysis of pca-based face recognition algorithms. In K. Boyer and J. Phillips, editors, Empirical Evaluation Techniques in Computer Vision. IEEE Computer Society Press.

[28] Wendy S. Yambor Bruce A. Draper J. Ross Beveridge, Analyzing PCA-based Face Recognition Algorithms: Eigenvector Selection and Distance Measures, July 2000, Colorado State University

Appendix A

Database creation for an European project

A.1 Database creation

Most of existing face database are restricted to 2D images. We propose here to build a 3D face database. For the enrollment, we use Geometrix (www.geometrix.com): It is a stereo 3D camera (it takes two pictures of a person and create its 3D model). We can export 3D models in different extensions: .gvf, .mts, .3ds, .obj, .wrl and .stl.

A.2 Requirements

A.2.1 People

We will randomly choose people from Eurecom (having their approval).
Patients:

1. 50 people (male and female).
2. two separate groups A and B arbitrary chosen.
3. 3 classes of age: Students (22-25 years), Staff (26-40 years), Professors (40-60 years).
4. origin: Eurecom is an international institute; many nationalities are represented.

A.2.2 Accessories

To avoid face distortions, people will be uncovered. We will ask to remove all hat, cap, glasses, . . .

A.2.3 Snapping

We will need five minutes per enrollment (installation, snap, exportations). Neutral face expression. For each candidate, we will snap under the following variations:

1. Lighting

    - *Outdoor-lighting+Spotlight:* The photographs will be taken in front of a window (natural light), and with a spotlight.
    - *Outdoor-lighting-only:* Same as previous but with no spotlight.
    - *Darkroom:* Snap in a room with only neon light.

2. Schedule
   We will take photographs at different day time: in the morning, noon, and afternoon.

3. Special
   If somebody forget to shave himself, or have a new scar, or anything special on his face, he will be photographed.

A.2.4 Nomenclature

Each people belonging to A or B will receive a personal number during the first session. Then, we will number (1, 2, 3) each lighting (outdoor, spotlight, darkroom) and the special picture 4, and each session will be numbered. For example, the third session for the 5th group-A-male, with the outdoor-only-lighting will be:
A_5_1_3_m.wrl

A.3 Interface

The database will have an interface admitting a customization of the database: We could choose to use the complete or only some parts of the database. For example, we could manipulate separately sub databases A and B.

A.4 Schedule

- from the 7th to the 11th November
  send mail to find guinea-pigs.

- from the 21th to the 25th November
  Finalize the legal and organizational aspects (sign the form to allow the utilization of the photographs, schedule).

- from the 28th November to the 2dn December
  Noon snap - A Group.
- from the 5th to the 9th December
  Noon snap - B Group.
- from the 12th to the 16th December
  Morning snap - A Group.
- from the 9th to the 13th January
  Morning snap - B Group.
- from the 16th to the 20th January
  Afternoon snap - A Group.
- from the 23th to the 27th January
  Afternoon snap - B Group.
- from the 30th January to the 3rd February
  Darkroom - A and B Group.

Appendix B

Correspondence with Geometrix lab

B.1 Wednesday 30 th November, to Geometrix

```
Hello,

We are students at Eurecom Institute working on a 3D project.

We have a geometrix camera with the software FaceVision 200, but we
have now some problems to make cameras work.

AS an attached file, you'll find an image of the error when we
launch the software.

We are waiting for your answer to solve this problem and are
available if you need any further information.

Best regards,

Ayoub Bacherki - Stephan Singh
```

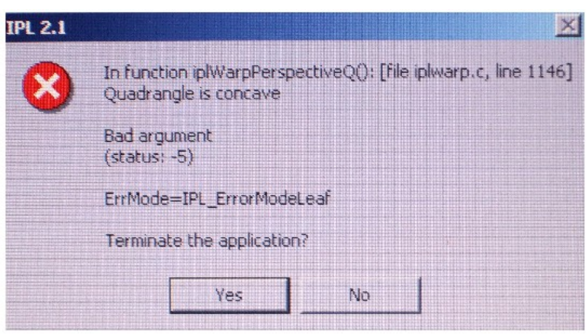

Figure B.1: Error with Geometrix

B.2 Thursday 1 st December, from Geometrix

```
Hi Ayoub,

The message has to do with a bad image for the camera.

Reasons: Not enough power to camera Bad USB (hub) Bad circuit in
camera

We need to do a little troubleshooting.

1.Was the system working fine then stopped or was it moved and set
back up?

2. How do you have the cameras plugged in to the PC? Are you using
the USB hub Geometrix sent with the system? The cameras need to run
through a USB hub that is powered.

The Camera will not work to full power if they are plugged directly
into the PC or in a USB hub that is not powered and/or not of good
quality. It may work you will not be getting good image and the
system will not work to the max

If you are not using a USB hub this will most likely be the problem.
Use the one we sent or purchase a good quality "USB2" that is
powered.

3. Have you tried unplugging the cameras and re plugging them back
in. Make sure to wait at least 10 seconds before plugging in the
second camera.

Please reply to this question and then we go from there.

Regards,

Steve
```

B.3 Wednesday 7 th December, to Geometrix

```
Hello,

We still have the Camera problem.

We tried many configurations on many PCs (and windows OS).

As you asked us, we plugged the camera USB cables directly on the PC
(USB 2) but it did not Work. We tried to use a powered USB hub, but
it still gives the same previous error (attached document).

Please tell us what we can do to use the cameras.

Best regards,

Ayoub
```

B.4 Thursday 8 th December, from Geometrix

```
Bonjour,

Je suis un ingenieur qui travaille avec Steve Macdonald. Il m'a
transmit la description de vos problemes. Afin de les resoudre le
plus vite possible, j'ai besoin que vous repondiez presisement a une
liste de questions:

1. Le systeme fonctionnait correctement sur un PC, puis soudainement
ces messages d'erreurs sont apparus sur ce meme PC? OU 2. Le systeme
a ete transferre sur un autre PC, et ces erreurs ont commence a
apparaitre?

3. Utilisez-vous toujours le meme hub USB2 que nous vous avons
fourni?

4. Dans le cas contraire, pouriez-vous nous transmettre les
specifications du hub que vous utilisez aujourd'hui?

5. Les specifications du hub que nous utilisons peuvent etre
trouvees sur ce site web:
http://catalog.belkin.com/IWCatProductPage.process?Merchant_Id=&Section_Id=200433&pcount=&Product_Id=123917


Cordialement, David Guigonis
```

B.5 Thursday 8 th December, to Geometrix

```
Cher Monsieur,

> 1. Le systeme fonctionnait correctement sur un PC, puis soudainement
> ces messages d'erreurs sont apparus sur ce meme PC?

Exactement. On a donc essaye apres de changer de machine (plusieurs
machines differentes, plusieurs OS differents, et toujours la meme
erreur). Nous avons meme recupere un "ghost" (datant d'une periode
ou l'aquisition fonctionnait) de la machine sur laquelle cela
marchait, sans succes.

> 3. Utilisez-vous toujours le meme hub USB2 que nous vous avons fourni?

Oui, et nous avons egalement essaye de les brancher directement sur
les USB 2.0 de nos machines selon les directives de Steve Macdonald,
sans effet...


Merci d'avance pour votre aide,

Cordialement,

Stephan Singh
```

B.6 Friday 9 th December, from Geometrix

```
1. Quels sont les peripheriques branches sur le hub USB2? Seulement
les cameras du systeme, ou quelque chose d'autre?

2. Si quelque chose d'autre est branche sur le hub, le debrancher et
ne laisser que les cameras. Tester le system.

3. Si le systeme ne fonctionne pas, nous avons un programme qui
permet de capturer des images a partir d'une seule camera. Vous
pouvez le telecharger sur notre site ftp:

ftp://ftp3.geometrix.com login: eurecom password: lucam

L'executable dont vous avez besoin est: LuCam.exe

- Copiez ce fichier dans le repertoire ou FaceVision est installe
(C:\Program Files\Geometrix\FaceVision 200\) - Debranchez toutes les
cameras du systeme, et n'en rebranchez qu'une. - Demarrez LuCam.exe.
- Selectionnez une camera dans la liste (il ne devrait y en avoir
qu'une), cliquez "OK" - Cliquez sur "Start Preview"

Une fenetre devrait s'ouvrir et afficher de la video provenant de la
camera.

4. La video est-elle de bonne qualite? 5. Fermez l'application. 6.
Debranchez la camera, et branchez la 2eme camera 7. Redemarrez
l'application. 8. La video est-elle de bonne qualite? 9. Branchez
toutes les cameras, et verifiez que la video est toujours de bonne
qualite pour les 2 cameras.


Cordialement, David Guigonis
```

B.7 Friday 9 th December, from Geometrix

```
Cher Monsieur,

Il ne me sera pas possible de vous aider jusqu'au 3 Janvier 2006. Si
votre probleme persiste, and vous avez besoin d'aide avant cette
date, veuillez contacter Steve Macdonald (SteveM@geometrix) en
anglais.

Cordialement, David Guigonis
```

B.8 Thursday 19 th December, to Geometrix

```
Dear Sir,

Since the last mail, nothing has changed.

We have bought a brand new Powered USB 2.0 HUB and we still have the
Camera problem...

Is there a known workaround ?

Thanks,

Sebastien Sanchez.
```