



The IV2 Multimodal Biometric Database (Including Iris, 2D, 3D, Stereoscopic and Talking Face Data) and the IV2-2007 Evaluation Campaign

Dijana Petrovska-Delacrétaz, Sylvie Lelandais, Joseph Colineau, Liming Chen, Bernadette Dorizzi, Emine Krichen, Mohamed Anouar Mellakh, Anis Chaari, Souhila Guerfi, Moshen Ardabilian, et al.

► To cite this version:

Dijana Petrovska-Delacrétaz, Sylvie Lelandais, Joseph Colineau, Liming Chen, Bernadette Dorizzi, et al.. The IV2 Multimodal Biometric Database (Including Iris, 2D, 3D, Stereoscopic and Talking Face Data) and the IV2-2007 Evaluation Campaign. 2nd IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS 2008), Sep 2008, Crystal City, Washington DC, United States. (elec. proc), 10.1109/BTAS.2008.4699323 . hal-00765334

HAL Id: hal-00765334

<https://hal.science/hal-00765334>

Submitted on 22 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The IV² Multimodal Biometric Database (Including Iris, 2D, 3D, Stereoscopic, and Talking Face Data), and the IV²-2007 Evaluation Campaign

D. Petrovska-Delacrétaz, S. Lelandais, J. Colineau, L. Chen, B. Dorizzi, M. Ardabilian,
E. Krichen, M.-A. Mellakh, A. Chaari, S. Guerfi, J. D'Hose and B. Ben Amor

Abstract—Face recognition finds its place in a large number of applications. They occur in different contexts related to security, entertainment or Internet applications. Reliable face recognition is still a great challenge to computer vision and pattern recognition researchers, and new algorithms need to be evaluated on relevant databases. The publicly available IV² database allows monomodal and multimodal experiments using face data. Known variabilities, that are critical for the performance of the biometric systems (such as pose, expression, illumination and quality) are present. The face and subface data that are acquired in this database are: 2D audio-video talking-face sequences, 2D stereoscopic data acquired with two pairs of synchronized cameras, 3D facial data acquired with a laser scanner, and iris images acquired with a portable infrared camera.

The IV² database is designed for monomodal and multimodal experiments. The quality of the acquired data is of great importance. Therefore as a first step, and in order to better evaluate the quality of the data, a first internal evaluation was conducted. Only a small amount of the total acquired data was used for this evaluation: 2D still images, 3D scans and iris images. First results show the interest of this database. In parallel to the research algorithms, open-source reference systems were also run for baseline comparisons.

I. INTRODUCTION

Three major considerations were taken into account for the design of the IV² database. First, our focus was to acquire data related to the face modality. Face could be seen under different aspects: 2D, 3D, stereoscopic vision, and it encompasses also information about speech and iris. Therefore, face and subface modalities were chosen to be recorded. Our second concern was to acquire data that reflect also the current challenges of that modality (such as illumination or pose variability). Our last concern was related to multimodality, so that this database could serve for multimodal experiments related to different kinds of facial and subfacial data.

D. Petrovska, B. Dorizzi, E. Krichen, and M.-A. Mellakh are with Institut TELECOM, T&M SudParis, 9 Rue Charles Fourier, Evry, FRANCE; dijana.petrovska, bernadette.dorizzi, emine.krichen, mohamed.anouar_mellakh@it-sudparis.eu

S. Lelandais-Bonade, A. Chaari, and S. Guerfi are with Laboratoire IBISC-CNRS FRE 2873, Université d'Evry Val d'Essonne, Evry, FRANCE; s.lelandais, chaari, guerfi@iut.unif-evry.fr

J. Colineau and J. D'Hose are with THALES Research and Technology, RD 128 91767 Palaiseau, Cedex, FRANCE; joseph.colineau, johan.dhose@thalesgroup.com

L. Chen and M. Ardabilian are with Ecole Centrale de Lyon, Lyon, FRANCE; liming.chen, mohsen.ardabilian@liris-edu.eu

B. Ben Amor is with Telecom-Lille1, Lille, FRANCE; boubaba.benamor@telecom-lille1.fr

The IV² database was designed to permit evaluations of biometrics algorithms in regard to different known variabilities that influence the results, such as quality of the images, pose, expression and illumination. The existing databases offer one or sometimes more variabilities but not all at the same time. Or there are not always a big number of subjects present, which make the results extremely dependent on the subset of subjects present in the database. For practical purposes, we adopted the strategy of having “one variability” at a time. For example, we consider mostly only one session (same day acquisition) data, and record data that have either pose, illumination, or camera resolution variability.

The face and subface data that are present in the IV² database are: 2D audio-video talking-face sequences, 2D-stereoscopic data acquired with two pairs of synchronized cameras (frontal and left profile), 3D-facial data acquired with a laser scanner, and iris images acquired with a portable infrared camera. The motivations for choosing these data are the following.

Although a great deal of effort has been devoted to 2D face recognition, however, it still remains a challenging problem and successful 2D face recognition systems have been deployed only under constrained situations. As a conclusion of research reports, such as [18], one can consider that the most challenging factors are those that cause intra-class variations such as: 3D head pose, illumination (including indoor/outdoor), occlusion due to other objects and accessories (sunglasses, etc.) and facial expressions. Regarding 2D face recognition, the majority of system are using still images. Results are good when controlled still images are used. If variabilities such as pose, illumination or image quality are present, the recognition results are much worse.

Improvements for face recognition in unconstrained condition could be achieved in different ways. One possibility is to use more still images, or even a sequence of images. Combining 2D and 3D information in different ways is another solution. 3D information gives morphological information, which is more stable during a short time than 2D information. Another advantage in using 3D information is to make imposture more difficult. Or one could use different face and subface data (such as iris) in order to improve the results.

With the IV² database, it is now possible to test those hypothesis within a single database. For example this database provides a significant number of facial expressions, and full

3D face models for each subject. Audio-video sequences including high- and low- quality cameras and with pose variabilities are also present. 3D models obtained with a laser scanner or with two pairs of stereoscopic high resolution cameras, could also be compared. Fusion with iris is also possible, in order to measure the advantage of multimodality.

The sensors used to acquire the data are the following:

- A digital camera and a webcam are used to acquire 2D face sequences and synchronized face and speech data (denoted as the talking-face data);
- Two pairs of stereoscopic video cameras are used for the frontal- and left-profile stereoscopic data;
- The 3D faces are acquired with a laser scanner;
- Iris images are acquired with a portable infrared camera.

This multimodal database, metadata and experimental protocols defined for the 2007 evaluation, are publicly available (at a production cost, related to the database storage requirements of 1 TeraB)[12].

This paper is organized as follows. The acquisition cabin is presented in Sect. II. The recorded audio-video streams are described in Sect. III. Data for 3D face reconstruction, including 3D laser scans and stereoscopic data are described in Sect. IV. Sect. V presents the iris data. The summary of the acquired data and its partitioning are presented in Sect. VI. The first set of evaluations conducted within an internal evaluation campaign are summarized in Sect. VII. Conclusions and perspectives are presented in Sect. VIII.

II. THE ACQUISITION CABIN

The sensors necessary for the acquisition of the IV² multimodal biometric database were placed in a cabin, shown on Fig. 1, that resembles a so called “photomaton” device designed to collect citizens’ biometric data. In order to have the required lighting conditions, four halogen lamps are positioned on the top structure and directed towards the ground. Two other spots are installed in front of the person in order to limit the shades formed by the nose. This functionality is used under the controlled illumination conditions. The halogen lamps are controlled automatically, in order to create variable illumination conditions at precise moments of the acquisition protocol.

The majority of the IV² database contains data acquired within a single session. A small part of the subjects, came for a second session (between some days to a month later). A disjoint development data of 52 subjects is also part of the IV² database. During the acquisition of the development data the acquisition protocols, described below, were not strictly followed. In the following sections more details regarding the biometric data and acquisition protocols of the evaluation part of the database for each modality are given. For all the modalities present, around 300 subjects participated to the data collection. 77 of those subjects, participated to a second session, with a short time interval (some days to one month). For each modality, there are small variations regarding the number of subjects, due to acquisition problems.



Fig. 1. The IV² acquisition cabin.

III. THE IV² AUDIO-VISUAL STREAMS WITH POSE, QUALITY, ILLUMINATION AND EXPRESSION VARIABILITIES

Regarding existing face and audio-visual databases, without being exhaustive, we can mention the audio-visual BANCA database [1], [17], XM2VTS [16], PIE, FERET, and FRGC (Face Recognition Grand Challenge) [20] databases. The experiments and evaluations conducted with these and other databases, showed the importance of illumination, pose and image quality on the performance on face recognition algorithms. In the IV² database, the acquisition protocol is designed in order to allow such variabilities. The audio-visual streams (denoted also as talking-face data) are acquired with high and low quality cameras.

A. Audio-visual Streams with High and Low Quality Cameras

Two audio-video streams are recorded in parallel with a high-quality digital camera (denoted as DVCAM) and a low-quality Webcam (denoted as WEBCAM). The resolution of the DVCAM images is 780×576 pixels, with a mean distance of 90 pixels between the centers of the eyes. For the WEBCAM the resolution is 640×480 pixels, with a mean distance of 65 pixels between the centers of the eyes. During the audio-video acquisition the subject is reading 15 French sentences (with a duration of about 100 s). The same text is read by all the subjects. Then the person is invited to look at left (respectively right) in order to capture his left (respectively right) profile, and finally to look up and down. At the end of this acquisition step, two audio-video sequences are available (taken simultaneously): one from the camcorder and the other from the webcam. At this step the low resolution acquisition is stopped. This part of the acquisition is denoted as the talking-face sequence. Examples of still images extracted from the audio-video sequences with the high-quality digital camera and the low-quality Webcam are shown in Fig. 2.

The acquisition continues with only the high resolution camera (DVCAM). The protocol is concentrated on the illumination, expression and pose variabilities, and can be

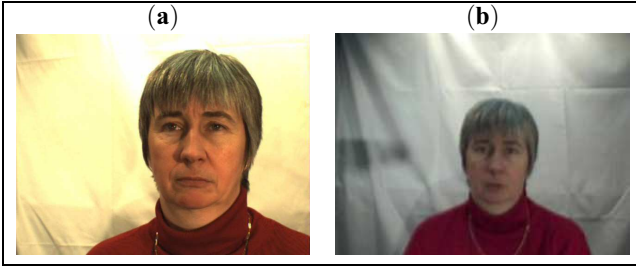


Fig. 2. Examples of still images taken from the talking-face sequences, with: (a) the high quality digital camera (DVCAM), and (b) the low quality camera (WEBCAM).

broken down into the following steps: the subject is first invited to look at the camera for approximately five seconds while the illumination intensity is set to its maximum. He is then invited to follow a visual path, and imitate a smiling, disgust and fear expressions. The duration of this part is 10 s. Next, the subject looks at the camera, and the illumination decreases by four steps during the next 10 s. Finally, the illumination changes from the right to the left side during another 10 s. From those sequences face images could be extracted that have the following variabilities for the high quality camera: pose, expression and illumination. Examples of such images are shown in Fig. 3.

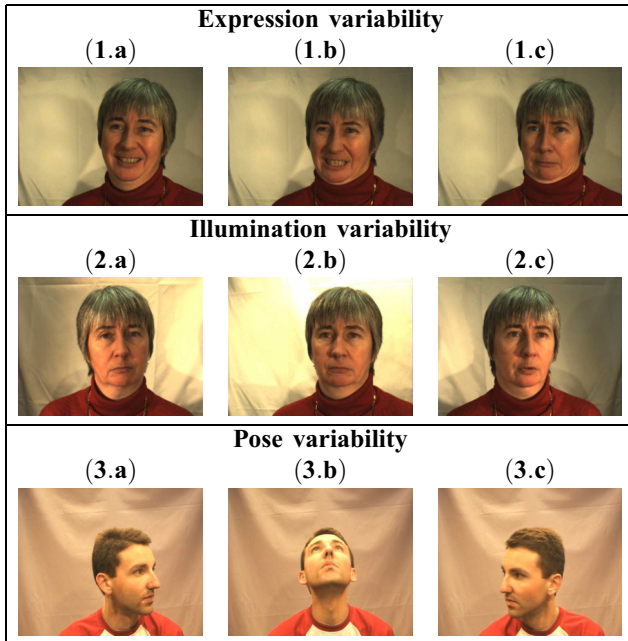


Fig. 3. Examples of variabilities present in the DVCAM video stream, related to: (1. a-c) expression, (2. a-c) illumination, and (3. a-c) pose.

IV. IV² DATA FOR 3D FACE RECONSTRUCTION

An emerging solution for face recognition is to use 3D face data. As the technology to acquire 3D facial surfaces becomes simpler and cheaper, the use of 3D facial scans could be considered. The research community has made very significant progress and there are today several commercial

systems based on different approaches using 3D information which are available on the marketplace. Nevertheless, a lot of acquisition systems present imperfections, such as missing data for the eye or hair regions and artifacts related to the disturbance of the structured light by particular ambient lighting conditions.

The second challenging problem for the 3D face modality is the presence of expression variation. The face muscles' deformations are known as a source of major difficulties in 3D-face recognition. In fact, 3D deformations due to facial expression can change considerably the 3D surface and thus a simple use of rigid registration algorithm is insufficient to build an accurate face recognition system. In addition, pose variation, time lapse between enrollment and test data (also known as gallery and probe data), presence of eyeglasses, and other factors are also important present challenging difficulties for 3D face biometrics.

Trying to use 3D information has become an emerging research direction in hope to make face recognition more accurate and robust. However, the number of publicly available 3D face databases to the research community is up to now very small. An example of a publicly available biometric database including 3D laser scans is the Face Recognition Grand Challenge (FRGC) [20] database. The version two of this database contains 4007 sessions of 466 subjects acquired under controlled lighting conditions and present also some facial expressions. Only frontal views are acquired in this database.

The main challenge of the 3D-face data present in the IV² database is to provide more significant facial expressions and some other variations such as pose and illumination in order to make tests under these variations, and the availability of 3 partial 3D scans (left profile, frontal and right profile), that permit a full 3D reconstruction of the face. Such 3D reconstructed data are present for each of the subjects.

A. 3D Laser Scans

The 3D face acquisition is carried out using the Konica-Minolta Vivid 7000 sensor. This device is a structured-light sensor that takes a 400×400 range sampling and a registered texture image; both channels are exported together in a VRML2.0 file format. Complete 3D data scans can be obtained in a post-processing step, starting from three 2.5 images taken from different view points: left profile, frontal, and right profile data.

During one acquisition session, taken in indoor arbitrary illumination conditions, a set of eight scans of a subject is collected. Each session contains also variations of lighting conditions, facial expressions and orientation of the head. In fact, it consists of five facial expressions (neutral, closed eyes, disgust, happiness, and surprise), two different lighting conditions and three poses (frontal, left profile and right profile).

Fig. 4 shows the partial 3D scans (left, frontal and right), that are used for the reconstruction of a full 3D face model, shown on Fig. 5, including the shape and texture channels.

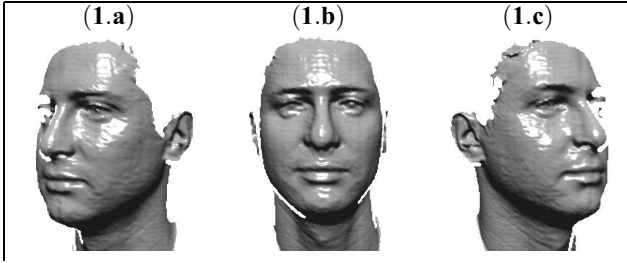


Fig. 4. Partial 3D data, (1. a) left profile, (1. b) frontal, and (1. c) right profile.

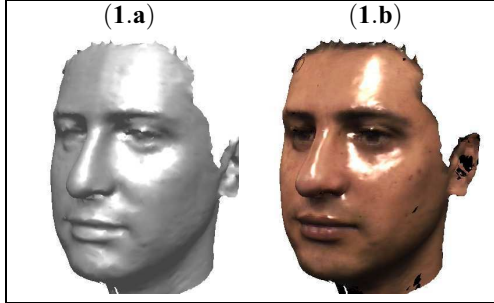


Fig. 5. Example of a reconstructed 3D full face model: (1. a) shape part, and (1. b) texture part.

Fig. 6 shows the variations present in the partial scans provided in the IV² database. Five different facial expressions are present in each session. While, some of them (disgust, happy and surprise) show a significant deformation, the neutral and closed eyes data show only small deformations of the face.

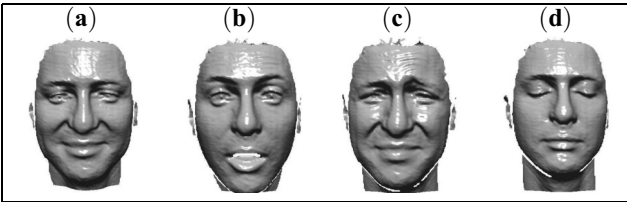


Fig. 6. 3D scans with expression variability: (a) smile, (b) surprise, (c) disgust, and (d) closed eyes.

B. 2D Stereoscopic Data

Another way to obtain 3D information is to use stereoscopic sensors. Such a sensor is composed of two cameras. The 2D stereoscopic data are acquired using two stereoscopic benches. The first one, placed in front of the subject, takes frontal face sequences. The second, placed on the left, takes images of the left profile of the subject. Figs. 7 and 8 show examples of images acquired from these two stereoscopic pairs. They are denoted as STEREO-Frontal and STEREO-Left cameras.

The color video cameras used for the stereoscopic acquisition are very good quality cameras with a resolution of 640×480 pixels, with a low distortion objectives. Each pair of these cameras is synchronized. This synchronization is essential to obtain the 3D face information later. A calibration

process of the stereoscopic benches is also developed, and the metadata needed for these 3D reconstructions are also present.

During 40 s, we acquire one sequence from each of the four cameras. During the acquisition, light, pose and expression variabilities are present, the same as for the 2D sequences acquired with the high-resolution DVCAM.



Fig. 7. : Example of synchronized frontal images from the stereoscopic STEREO-Frontal pair of cameras.



Fig. 8. : Example of synchronized left profile images from the stereoscopic STEREO-Left pair of cameras.

V. IV² IRIS DATA

The IV² iris data is acquired with a portable infrared camera. The number of subjects (around three hundred) that participated to this data collection, is bigger than in the other publicly available databases. Multimodal experiments could also be made combining iris with different facial data, such as 2D and 3D. The camera has an integrated quality control module. Subjects wearing glasses were asked to take them away for the first session acquisition. For the subject that had come more than once, they were asked to keep their glasses on. Therefore, for the people having more sessions, the glasses variability is present in our database, besides the quality variability.

VI. SUMMARY OF ACQUIRED DATA AND ITS PARTITIONING

In this section, a summary of the acquired data is presented first. Next, the partitioning of the database is given. This partitioning is necessary in order to define development and evaluation sets.

A. Summary of acquired data

For each session, the subjects participated to the following acquisitions:

- Simultaneous acquisition of talking faces (100 s) including 5 s of pose variability with high and low quality cameras (DVCAM and WEBCAM respectively).

- Simultaneous acquisition of expression, illumination (different illumination levels), illumination (from left to right), and pose variability (total duration 40 s) with the DVCAM and STEREO-Frontal and STEREO-Left cameras.
- Acquisition of 3D data with the laser scanner, comprising pose, illumination and expression variabilities.
- Acquisition of iris images with a portable (low quality) iris infrared sensor.

In addition to the acquired data, some “metadata” are present. They are related to the age class, color of skin, color of eyes, color of hair, and presence of beard, mustache and glasses, and the calibration information of the stereoscopic heads.

The total mean size of the recorded data for one complete session is about 2.2 GigaB for each subject, composed of:

- 2D stereoscopic sequences: 2000 MB;
- 3D faces (Minolta Scanner): 11 MB;
- Talking 2D faces with the DVCAM: 200 MB;
- Talking 2D faces with the WEBCAM: 15 MB;
- Iris images: 15 MB.

B. Partitioning of the IV^2 Database

The IV^2 multimodal database (1 TeraB), is split into a development and an evaluation set, as follows:

- IV^2 -development data, contains data from 52 subjects. These acquisitions were done at the beginning of the project, when the acquisition protocol was not yet finalized. This data could be used as development data, for example to construct a PCA face space, or to tune some parameters.
- IV^2 -evaluation data: In this set, there are about 300 subjects with one session. Among those subjects 77 come for a second session and 19 for a third session.

VII. THE IV^2 -2007 EVALUATION CAMPAIGN

From the IV^2 -evaluation data, a first set of data (of approximate size of 100 GB) was extracted for the IV^2 -2007-evaluation. The data were semi-automatically extracted in order to allow the first set of monomodal and multimodal experiments, described in the following paragraphs. The evaluation campaign for IV^2 was conducted in September 2007, on the same way as the NIST speaker evaluation: each team conducted a list of tests with no indication on their type (intraclass or interclass). Once their results submitted to the evaluator (THALES), the ground truth for the conducted test was made available to the participants. The goal of this first evaluation campaign (mostly focused on monomodal evaluations) was to measure the difficulty of the acquired database, and to find out the most appropriate and challenging datasets that could be used for further monomodal and multimodal evaluations. Baseline algorithms (already benchmarked on other databases, developed within the European BioSecure project[6], described and compared in [19], and available at [7]) are also used in order to calibrate the difficulty of this database.

A. 2D Face Evaluation

For the 2D face evaluation part (described in more detail in [15]), only one still image per subject was used for enrollment and one still image for the test. All the enrollment images are frontal, with high resolution, taken under maximum illumination and with neutral expressions. Table I gives a description of the test images according to the four experiments defined for the 2007 evaluation. The characteristics of these experiments can be summarized as follows. All the enrollment (gallery) images are good quality still face images with no variability. IV^2 -2D-Exp1 could be considered as an experiment where there is not a lot of variability present, and corresponds to “controlled scenarios”, including some small expression variability for the test data. For the IV^2 -2D-Exp2 the test data is chosen in order to include illumination variability. For the IV^2 -2D-Exp3, the test data comes from the WEBCAM, while for the IV^2 -2D-Exp4, the data were chosen in order to measure the influence of time variability (of a small set of 77 subjects).

TABLE I
2D EXPERIMENTS. PART ONE: DESCRIPTION OF THE IMAGES USED FOR THE 2D EXPERIMENTS (V MEANS VARIATION AND N STANDS FOR NUMBER). PART TWO: EER RESULT OF THE BASELINE SYSTEM AND THE LOWER AND UPPER BOUNDS OF THE SUBMITTED SYSTEMS

Experiment	2D-Exp1	2D-Exp2	2D-Exp3	2D-Exp4
Sessions	Mono	Mono	Mono	Multi
Quality	High	High	Low	High
Expression V.	Small	No	No	No
Illumination V.	No	Yes	No	No
N. Intraclass	2595	2502	1654	1796
N. Interclass	2454	2362	1598	1796
BioSecure (PCA)	7.3	21.6	16.3	13.6
Participants EER%	4–7	12–22	11–22	8–20

Five appearance based methods were evaluated on the IV^2 database. The results are reported with the Equal Error Rate percentage. The first system is the baseline Principal Component Analysis (PCA) system, developed within the European BioSecure project[6], and available at [7]. For comparison purposes, the same baseline system was run on the FRGCv2 and on the BANCA database, allowing for comparisons of databases and different algorithms (see also [19] for a detailed description of the benchmarking PCA system and comparisons with other research prototypes). The results of the baseline PCA system (with the same parameters, such as data for train the PCA face space) on the FRGCv2 database give for the FRGC-Exp1 (controlled vs controlled data) an EER of 5.6% and for the FRGC-Exp4 (controlled vs uncontrolled data) an EER of 24.0%. Results of the baseline PCA system confirm that the difficulty of the IV^2 database is comparable to the difficulty of the FRGCv2 database. The results of the baseline PCA are comparable for IV^2 -2D-Exp1 and FRGCv2-Exp1 (EER of 7.3% and 5.6%, respectively), and for IV^2 -2D-Exp2 and FRGCv2-Exp4 (EER of 24.0% and 21.6%, respectively).

The results from the IV^2 -2007 2D-still face evaluation

confirm the state of art: the Linear Discriminant Analysis (LDA) performs well in controlled scenarios, the choice of the training set for PCA analysis is important for the generalization properties of the PCA, and performance quickly decreases with the wide range of variabilities to be handled. The LDA applied to a Gabor transformation of the face, is more robust than the classical graylevel approach. More results and comparisons of this approach on other databases can be found in [19].

In the 2007 evaluation (using only still images), only a small subset of the recorded data in the IV² database is exploited. Future work will concentrate on preparing more data for new experiments, including video sequences, and multimodal data. These 2D face evaluations have shown the contribution of this database for new evaluations in real life adverse conditions.

B. 3D Face Evaluation

For the IV²-2007-3D evaluation, the influence of facial expressions on the performance of facial surface matching based 3D face recognition algorithms was of interest.

From the available data, three experiments are designed for the 2007 evaluation, denoted as 3D-Exp1, 3D-Exp2 and 3D-Exp3. Data from the first session (S1) were split into enrollment data (the neutral faces) and the rest of the 3D scans, with time, illumination and expressions variabilities are used as test samples. The 3D experiments using scans from the first (S1) and second (S2) session can be summarized as follows:

- IV²-3D-Exp1: neutral faces from S1 against neutral faces from S2; 96 subjects with 760 intraclass among 1532 total comparisons.
- IV²-3D-Exp2: neutral faces from S1 against neutral faces from S1; 300 subjects with 1604 intraclass among 4001 total comparisons.
- IV²-3D-Exp3: neutral faces from S1 against expressive faces from S1; 300 subjects with 2380 intraclass among 4724 comparisons.
- IV²-3D-Exp3-a: neutral faces from S1 against expressive faces from S1; 300 subjects with 2244 intraclass comparisons (including bad quality enrollment data).

IV²-3D-Exp1 allows comparisons between neutral faces of two different sessions. Neutral faces include three representations: neutral expression, neutral expression in low ambient light and closed eyes.

IV²-3D-Exp2 allows comparisons between neutral faces from the same session.

Because 3D scans are very sensitive to defects (such as holes in the nose or eyebrow region), IV²-3D-Exp3 is designed to measure the influence of bad quality enrollment data, when illumination and expressions variabilities are present in the test data. In IV²-3D-Exp3 (the 7% of the bad quality enrollment scans were discarded). Exp-3D-3 is a comparison between neutral faces in the gallery (enrollment data) and expressive faces (probe) in the same session. Expressions include: smile, surprise, and disgust.

For IV²-3D-Exp3-a the bad quality enrollment data were kept, leading to 2444 intraclass comparisons (the rest of the data being the same as in IV²-3D-Exp3).

More details about these 3D evaluations can be found in [10]. Three facial surface matching based algorithms, namely Iterative Closest Point (ICP) [2], [3], L-ICP which is a variant of ICP applied to localized facial surface centred at the nose comparable to the work described in [9], and R-ICP which is a region-based facial surface algorithm [4], [5] are compared. The experimental results show that facial expressions account for a major intraclass variations and that the region-based facial surface matching algorithm (R-ICP), while giving more weight to stable facial regions in the similarity score computation process, gives promising results.

C. Iris Evaluation

The iris data acquired within the IV² database is compared to the publicly available ICE iris database [11] (that contains 2953 images from 132 persons). Some degradations are present in the ICE database, like blur, occlusion and off angles images. The IV² iris database is acquired following the same acquisition protocol as the Casia-v2 database, using the OKI device [8]. Both right and left iris are acquired. Subjects can wear eyeglasses in the second session which is a noteworthy difference from the ICE database. For the 2007 evaluation, we created a subset of IV² iris data, with more than 4,000 images from 120 persons, leading to 16,000 intraclass and 16,000 interclass comparisons. Fig. 9 shows the distribution of a probabilistic global quality measure defined in [13] for images from the IV² and ICE databases. It shows that, using this quality measure, the quality of the iris images from IV² is lower than ICE. Indeed, 5% of the IV² images have lower quality measure than 0.5, while only 2% of the ICE database have this quality value. 20% of the IV² iris images have a lower quality metric than 0.6 and 50% lower than 0.7 while in ICE only 8% (respectively 23%) of the images have 0.6 (respectively) 0.7 quality measure. Therefore IV² database is a good choice to evaluate performance of systems on degraded iris images.

For the IV²-2007 evaluation, three different algorithms were compared: two open-source reference systems, OSIRIS (available at [7], and described in [19]) and Masek [14], and a research system based on correlation and quality measures called INT [13]. The INT system outperforms OSIRIS and Masek which have almost similar behavior (the two systems are based on the Daugmans algorithm). For OSIRIS, no prior eyelids detection is used. The gap between the three systems is higher than the one obtained using the ICE database on which better performance for the same three systems was obtained.

D. Multimodal Evaluation

Some preliminary multimodal experiments were conducted. Score level fusion of 2D, 3D and iris modalities showed that as expected, improvements that can be obtained when score fusion is done.

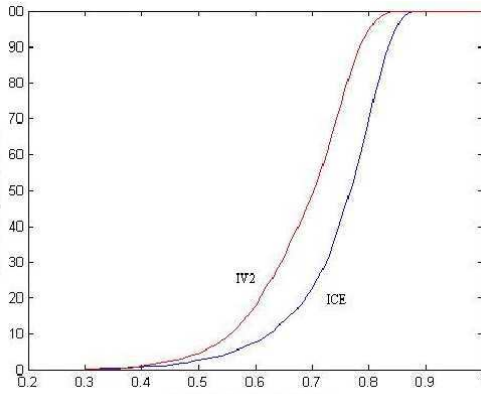


Fig. 9. Percentage of data present in the ICE and IV2 databases, as a function of the quality measure (spanning from 0–1, from low-quality to high-quality measures).

VIII. CONCLUSIONS AND PERSPECTIVES

The IV² multimodal biometric database has many different situations such as variations of pose, expressions, illumination or quality. A first internal evaluation campaign was conducted in 2007, showing that the multiple test conditions this database can offer, related to degraded acquisition conditions, represent challenging datasets for the actual biometrics systems. Within this campaign only a small part of the database was used and annotated (100 GigaB from 1 TeraB of recorded data). With this database, it is possible to further define new experiments and evaluation campaigns for the scientific community. The new experimental protocols will focus on exploiting different aspects (datasets) from the database, such as video sequences, talking faces, stereoscopic data, full 3D scans and multimodal fusion.

The database is publicly available. It is protected by the French Data Protection Authority. If a research laboratory wants to acquire the IV² database, including the 2007 evaluation package, all the necessary information is available on the website of the project [12].

IX. ACKNOWLEDGMENTS

This work was partially supported by the French Research Ministry and the Ministry of Defense, realized during the Techno Vision program. We would like to thank all the partners of the IV² project, who took part in the conception of the acquisition cabin and the experimental protocols: S. De Sousa, G. Chollet, H. Bredin, S. Salicetti-Garcia, J.-L. Dugelay, C. Mallauran, A. Mikhaleukov, C. Bernard, J. Dujin, F.-X. Espiau, C. Couteaud, and N. Boujema.

The authors gratefully acknowledge the reviewers' comments.

REFERENCES

- [1] E. Bailly-Baillière, S. Bengio, F. Bimbot, M. Hamouz, J. Kittler, J. Mariéthoz, J. Matas, K. Messer, V. Popovici, F. Porée, B. Ruiz, J.-P. Thiran, *The BANCA Database and Evaluation Protocol*, 4th International Conference on Audio- and Video-Based Biometric Person Authentication, AVBPA 2003, Guildford, UK
- [2] P.J. Besl and Neil D. McKay, *A Method for Registration of 3D Shapes* Pattern Analysis and Machine Intelligence, Volume 14 Number 2, p. 239-255, February 1992
- [3] B. Ben Amor, M. Ardabilian, L. Chen, *New Experiments on ICP-based 3D Face Recognition and Authentication*, 18th International Conference on Pattern Recognition, (ICPR 2006), pp. 1195-1199
- [4] B. Ben Amor, M. Ardabilian, L. Chen, *Enhancing 3D Face Recognition by Mimics Segmentation*, Proceedings of the Sixth International Conference on Intelligent Systems Design and Applications (ISDA 2006): 150-155, October 16-18, 2006, Jinan, China. IEEE Computer Society 2006, ISBN 0-7695-2528-8
- [5] B. Ben Amor, M. Ardabilian, L.L. Chen, *Toward a region-based 3D face recognition approach*, 2008 IEEE International Conference on Multimedia & Expo (ICME 2008), June 23-26 2008, Hannover, Germany
- [6] BioSecure NoE, <http://biosecure.info>
- [7] BioSecure Benchmarking Framework, <http://share.int-evry.fr/svnview-eph/>
- [8] Casia database, <http://www.sinobiometrics.com>
- [9] K. I. Chang, K. W. Bowyer, and P. J. Flynn. Multiple Nose Region Matching for 3D Face Recognition under Varying Facial Expression, IEEE Transactions on Pattern Analysis and Machine Intelligence, 28(10): 1695-1700, October 2006
- [10] J. Colineau, J.D.Hose, B. Ben Amor, M. Ardabilian, L. Chen, B. Dorizzi, *3D Face Recognition Evaluation on Expressive Faces Using the IV2 Database*, submitted to ACIVS 2008.
- [11] ICE database: <http://iris.nist.gov/ice/>
- [12] IV² database *Identification par l'Iris et le Visage via la Video*, <http://iv2.ibisc.fr/PageWeb-IV2.html>
- [13] E. Krichen, S. Garcia-Salicetti, and B. Dorizzi, *A New Probabilistic Iris Quality Measure for Comprehensive Noise Detection*, IEE Conference on Biometrics: Theory, Applications and Systems; Sept. 27th to 29th, 2007, Washington DC
- [14] L. Masek, *Recognition of Human Iris Patterns for Biometric Identification*, University of Western Australia, 2003
- [15] A.M. Mellakh, A. Chaari, S. Guerfi, J. D'Hose, J. Colineau, S. Lelandais, D. Petrovska-Delacrétaz, B. Dorizzi; *2D Face Recognition in the IV²-2007 Evaluation Campaign*, Internal Report.
- [16] K. Messer, J. Matas, J. Kittler, J. Luettin and G. Maitre, *XM2VTSDB: The extended M2VTS database*, Proc. Second International Conference on Audio- and Video-based Biometric Person Authentication (AVBPA), 1999
- [17] K. Messer and al., *Face Authentication Test on the BANCA Database*, ICPR 2004: Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04) volume 4, 2004, Washington, DC, USA
- [18] A. J. O'Toole, P. J. Flynn, K.W. Bowyer, C. L. Schott, and M. Sharpe. *FRVT 2006 and ICE 2006 Large-Scale Results* National Institute of Standards and Technology, NISTIR 7408, <http://face.nist.gov>, 2007
- [19] D. Petrovska-Delacrétaz, G. Chollet and B. Dorizzi; *Biometric Reference Systems and Performance Evaluation*, Springer Verlag 2008
- [20] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek; *Overview of the Face Recognition Grand Challenge*. Computer Vision and Pattern Recognition (CVPR 2005), San Diego, June 2005, 1:947-954