Biometric Sensor Interoperability: A Case Study in 3D Face Recognition

Javier Galbally and Riccardo Satta
European Commission - Joint Research Centre, IPSC, Via Enrico Fermi 2749, 21027 Ispra, Italy

Keywords: 3D Face Recognition, Interoperability, 3D Face Database.

Abstract: Biometric systems typically suffer a significant loss of performance when the acquisition sensor is changed between enrolment and authentication. Such a problem, commonly known as sensor interoperability, poses a serious challenge to the accuracy of matching algorithms. The present work addresses for the first time the sensor interoperability issue in 3D face recognition systems, analysing the performance of two popular and well known techniques for 3D facial authentication. For this purpose, a new gender-balanced database comprising 3D data of 26 subjects has been acquired using two devices belonging to the new generation of low-cost 3D sensors. The results show the high sensor-dependency of the tested systems and the need to develop matching algorithms robust to the variation in the sensor resolution.

1 INTRODUCTION

In recent decades, we have witnessed the evolution of the biometric technology from the first pioneering works in signature and voice recognition to the current state of development where a wide spectrum of highly accurate systems may be found, ranging from largely deployed modalities like face, fingerprint or iris, to more marginal ones like the ear or the keystroke. This path of technological evolution has naturally led to the analysis of biometric-related issues different from the mere improvement of the systems accuracy. Among these relatively novel problems, biometric sensor interoperability stands out as one which has concentrated significant attention from the biometric community.

Ideally, the biometric feature set extracted from the raw data is expected to be an invariant representation of a persons trait. However, in reality, the feature set is sensitive to several factors including the change in the sensor used for acquiring the raw biometric samples. In this context, sensor interoperability refers to the ability of a biometric system to adapt to the data obtained from a variety of sensors. Most biometric systems are designed to compare data originated from the same sensor, but fail to give a reliable performance when the acquisition device is changed between the enrolment and the authentication phase.

Note that the problem of sensor interoperability as defined above is a challenging one, which cannot be solved by simply adopting a common biometric data exchange format (ISO/IEC, 2011; ANSI-INCITS, 2004), which aids in the exchange of images or feature sets between systems but does not provide a method to compare feature sets obtained from different sensors. Over the last years, researchers have analysed the impact of sensor interoperability in biometric performance trying to estimate the loss coming from it. Such studies include traits like fingerprints (Ross and Jain, 2004; Alonso-Fernandez et al., 2006), face (Khiyari et al., 2012), signature (Alonso-Fernandez et al., 2005), voice (NIST, 2014) or multimodal approaches (Alonso-Fernandez et al., 2008).

Similarly to what was done some years ago in other more mature modalities such as fingerprints (Ross and Jain, 2004), the present study represents an initial step to explore sensor interoperability in a relatively new biometric field like 3D face recognition. Opposed to its 2D counterpart, face authentication based only on the 3D morphology is claimed to be more robust to illumination and pose changes, however, its resilience to sensor changes has only been considered before in a very preliminary work (Faltemier and Bowyer, 2006). In this contribution, we take advantage of the new generation of affordable 3D acquisition sensors, to study the impact of using devices with different resolution, on two largely used 3D face matchers. With this objective, we have generated the first database with 3D facial data of the same individuals acquired with two different sensors.

The rest of the work is organized as follows. The experimental protocol with its main three components: data, systems and experiments, is described in Sect. 2. Experimental results are reported in Sect. 3.

Galbally, J. and Satta, R.
Biometric Sensor Interoperability: A Case Study in 3D Face Recognition.
DOI: 10.5220/0005682501990204
In *Proceedings of the 5th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2016)*, pages 199-204
ISBN: 978-989-758-173-1

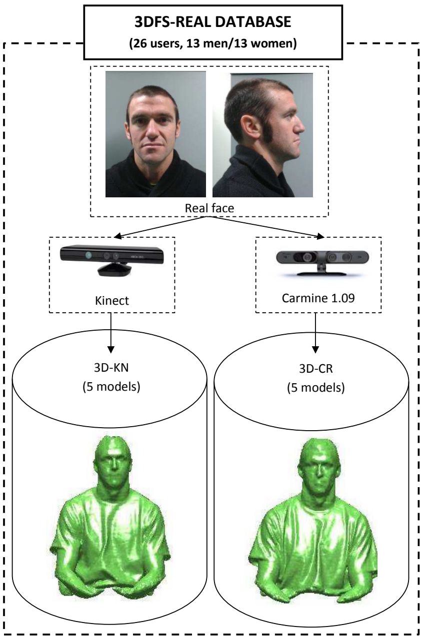

Figure 1: General diagram of the structure and generation process of the new 3DFS-REAL DB.

Conclusions are finally drawn in Sect. 4.

2 EXPERIMENTAL PROTOCOL

The experimental protocol has been designed to fulfill the main objective set in the present work, that is, determine the potential interoperability of low-cost 3D sensors based on Light Coding technology. In order to report as unbiased and meaningful results as possible the protocol includes:

- *Data*. The new 3D-Face Spoofing Database (3DFS-DB), which contains 3D, 2.5D and 2D real and spoofing data that allow to perform a very wide range of different tests, including interoperability performance evaluations. The DB is composed of two datasets: real and fake. In the present work only the 3DFS-REAL dataset will be used.
- *Systems*. Two proprietary implementations of state-of-the-art 3D recognition systems.
- *Experiments*. In order to fully characterize the interoperability of the two sensors considered in the experiments, two different scenarios are considered: *i*) performance evaluation under the standard operation scenario; and *ii*) performance evaluation under the sensor interoperability scenario.

All these three elements, database, systems and experiments, are described in the next subsections. Then, results are presented in Sect. 3.

2.1 The 3DFS-REAL Database

The 3D Face Spoofing Database (3DFS-DB) contains real and fake facial data of 26 subjects, 13 men and 13 women, all Caucasian between 25 and 55 years of age. It is composed of two datasets of real (3DFS-REAL) and fake (3DFS-FAKE) data. The present work only makes use of the real 3D data and, therefore, will focus on the description exclusively of this data subcorpus.

The 3DFS-REAL dataset contains 3D models in .stl format acquired using two low-cost standard 3D scanners (the price is around 200$): the Microsoft Kinect[1] and the PrimeSense Carmine 1.09[2].

Although several other 3D-face databases are currently available for research purposes including different pose, illumination and expression (Phillips et al., 2005; Zafeiriou et al., 2011; Min et al., 2014), to the best of our knowledge, this is the first 3D face database that contains samples of the same subjects acquired with two different sensors, allowing this way to perform interoperability experiments.

Both sensors contain a standard RGB camera that captures 2D 640×480 pixel color data and an infrared projection system which detects the depth in the picture (i.e., 2.5D data). Both sensors incorporate the Light Coding technology developed by the Israeli based company PrimeSense (recently acquired by Apple), however, the Carmine 1.09 scanner has a shorter range of operation (between 0.3-1.5 meters with respect to 0.8-4 meters of the Kinect) which enables it to achieve a maximum depth resolution of around 0.5mm compared to the 1mm resolution of Kinect.

Before the acquisition of the dataset all users where informed of the nature of the experiments and the processing of their data and were invited to sign a consent form in compliance with the applicable EU data protection legislative framework[3]. The dataset was acquired in an office like scenario with no specific illumination control and no constraints on the background except that no other object was allowed within the acquisition range. Data were captured as follows: The user sat in front of the sensor on a revolving chair fixed to the ground and rotated 180°from left to right at a regular speed with a neutral face expression. The 3D models were acquired using the 90$ license ap-

[1]http://en.wikipedia.org/wiki/Kinect

[2]http://en.wikipedia.org/wiki/PrimeSense

[3]Regulation (EC) No 45/2001 of the European Parliament and of the Council of 18 December2000 on the protection of individuals with regard to the processing of personal data by the Community institutions and bodies and on the free movement of such data.

plication Skanect[4] and saved in .stl format. For each user a total 10 models were acquired: five with Kinect and five with Carmine 1.09. The general structure and generation process of the database is depicted in Fig. 1.

In compliance with the EU personal data protection regulation only indirect access to the data is possible for research purposes upon request to the authors[5]. Such an indirect access implies that interested researchers can run their algorithms on the database remotely but they are not allowed to download the data. It is envisaged that such access will be automatized in the future through the use of the new open source BEAT platform (BEAT, 2012).

2.2 3D Face Recognition Systems

Bear in mind that, as mentioned before, the objective of the work is not to develop new and more precise face recognition systems, but to evaluate the interoperability of new low-cost 3D sensors providing reliable baseline results using reasonably accurate implementations of well known solutions for face authentication. For this purpose, two different popular systems were considered in the experiments:

- **3D Proprietary Implementation 1: HD-based**. The system carries out the next preprocessing steps before computing the similarity scores between two 3D models: 1) head detection; 2) head segmentation from the rest of the body; 3) head rotation so that the eyes are aligned with the x axis; 4) face segmentation from the rest of the head; 5) face normalization forcing the nose tip to be at point $(0,0,0)$. The similarity score between two normalized 3D face models is computed as the Hausdorff distance (HD) (Barnsley, 1993; Henrikson, 1999), which measures how far two subsets (not necessarily composed of the same number of points) are from each other within a given metric space (in our case a three-dimensional space). In brief, two sets are close according to the Hausdorff distance if every point of either set is close to some point of the other set. The Hausdorff metric had already been successfully used in previous works to compare 2D images (Huttenlocher and Rucklidge, 1992), 3D meshes (Cignoni et al., 1998), and in 3D face recognition (Achermann and Bunke, 2000; Wang and Chua, 2006), showing a remarkable performance in the Face Recognition Grand Challenge (FRGC) (Phillips et al., 2005).

- **3D Proprietary Implementation 2: ICP-based**. The same preprocessing steps followed by system 1 are performed prior to the computation of the similarity score. Then, the score is generated according to the Iterative Closest Point (ICP) algorithm, which is a well-established technique used for rigid registration of 3D surfaces (Besl and McKay, 1992). In order to minimize the distance between two cloud points (which is the sum of distances calculated for all points in one of the surfaces, finding the closest point on the other), ICP computes and revises the translation and rotation iteratively. This registration is used to establish point-to-point correspondences between two face models. The final minimized distance, ICP error, is used by the system as the similarity score between the two compared faces (Amor et al., 2006; Lu et al., 2004).

  Two limitations of the ICP-based approach are that it needs a good initialization for an accurate result and that it does not consider nonrigid transformations which is required in the presence of surface deformations, such as occlusions or facial expressions. In the particular case of the present study, such two challenges are addressed at the acquisition of the database, allowing only frontal samples and neutral expression.

2.3 Experiments

Defining a clear methodology and its associated metrics to assess the interoperability of biometric sensors is not a straight forward problem, as there are different variables and evaluations involved when the interoperability dimension is introduced. The evaluation protocol usually followed for the assessment of biometric sensor interoperability defines two possible working scenarios as shown in Fig. 2:

- **Standard Scenario**, where both enrollment and test samples are acquired using the *same* sensor. This scenario serves as the baseline with which to compare the interoperability results. It considers genuine access attempts (i.e., regular access attempt in which a user logs in as himself) and zero-effort impostor access attempts (i.e., access attempts in which the attacker uses his own real biometric trait but claims to be a different user). In this scenario performance is typically reported in terms of the FRR (False Rejection Rate, number of genuine access attempts wrongly rejected) and the FAR (False Acceptance Rate, number of zero-effort impostor access attempts wrongly accepted). The working point where both the FRR and the FAR take the same value is the Equal

[4] www.skanect.com

[5] For further details on the distribution of the DB please contact: javier.galbally@jrc.ec.europa.eu

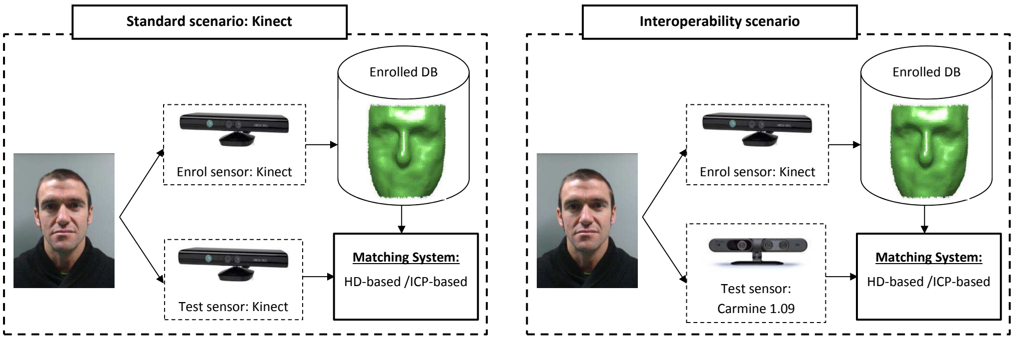

Figure 2: Diagram showing the different enrolment/test sensor configurations for the standard scenario (with the Kinect sensor) and the interoperability scenario considered in the work.

Error Rate (EER) and is generally accepted as a good estimation of the overall performance of the system.

- **Interoperability Scenario**, genuine and impostor attempts are defined as before, however, in this case, the sensors used for enrollment and test are different, leading in general to poorer results. Although the metrics used to evaluate the systems in this scenario are the same as in the standard one, for clarity we will refer to them as FAR-I, FRR-I and EER-I, where the "I" stands for Interoperability.

All these four metrics (i.e., FRR, FAR and FRR-I,FAR-I) should be strictly assessed to determine the real performance variation experimented by a given system between the two scenarios.

For each of the systems considered in the experiments and described in Sect. 2.2, the sets of scores (i.e., genuine scores and zero-effort impostor scores) were computed as follows:

- *Standard Scenario*. The same protocol was used for the two systems and for the models produced with the Kinect and the Carmine 1.09 sensors. *Genuine scores* were computed using successively all five processed 3D face models for enrollment (i.e., one each time), and testing with the remaining four models of the same sensor avoiding repetitions, leading this way to $26 \times 10 = 260$ genuine scores. *Zero-effort impostor scores* were computed matching the first model from the 25 remaining users to the first model of a given subject (acquired with the same sensor), that is $26 \times 25 = 650$ zero-effort impostor scores. Therefore, in this scenario, for each system, two sets of FRR/FAR curves are available, one for the Kinect and one for the Carmine 1.09.
- *Interoperability Scenario*. In this case *genuine scores* were computed matching, for each user, all 5 models acquired with the Kinect sensor to all 5 models acquired with the Carmine 1.09, leading this way to $26 \times 5 \times 5 = 650$ genuine scores. *Zero-effort impostor scores* were computed matching all five Carmine 1.09 models of each user to the fist Kinect model of the remaining 25 users, that is $26 \times 5 \times 25 = 3,250$ zero-effort impostor scores. Therefore, in this scenario, for each system, there is one set of curves FRR-I/FAR-I.

3 RESULTS

The experimental protocol described in Sect. 2 allows to objectively compare the performance of 3D face recognition systems in the standard and interoperability scenarios and, therefore, to fully characterize the performance variation experimented by the two considered recognition systems.

The genuine and zero-effort impostor sets of scores described in 2.3 are used to compute the metrics FRR/FAR in the licit scenario and FRR-I/FAR-I in the interoperability case. Each of these two metric tuples are plotted in the form of Detection Error Trade-off (DET) curves in Fig. 3, so that the performance of the systems may be visually compared in the two considered working scenarios. For each of the charts, the x axis represents either the FAR or the FAR-I depending on the scenario selected (licit or interoperability). A quantitative comparison between the two scenarios may be obtained from the EER shown in the charts legend. Two different curves are presented for the standard scenario, one for each sensor used in the acquisition: Kinect (KN) and Carmine 1.09 (CR).

Several interesting conclusions may be extracted from the results shown in Fig. 3:

- Regarding the standard scenario results, it may be observed that the performance of the 3D proprietary systems considered in the work, based only on the face geometry/shape, is still a step behind

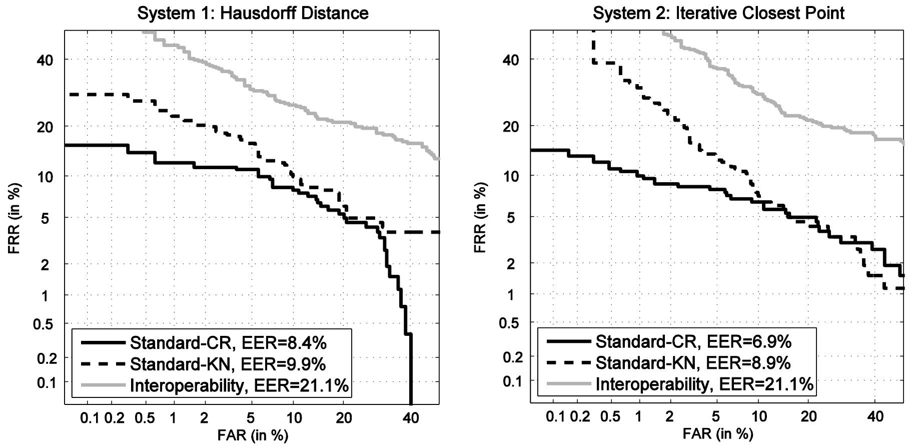

Figure 3: DET curves for the two systems considered in the work and for the two different scenarios: standard (with the Kinect and Carmine 1.09 sensors) and interoperability. CR stands for Carmine 1.09 while KN stands for Kinect.

that of top-ranked 2D face recognition systems under good acquisition conditions (i.e., controlled illumination, pose and background). This corroborates the results obtained in past independent competitions (Phillips et al., 2005; Bowyer et al., 2006), and shows that, in spite of the obvious advances in terms of size and price, off-the-shelf 3D sensing technology still needs to improve its accuracy to reach really competitive recognition results in the field of face authentication.

- Also worth noting that, as expected, in the standard scenario the higher resolution of the Carmine 1.09 sensor with respect to Kinect translates into better performance, decreasing the EER from 9.9% to 8.4% in the case of the Hausdorff distance system and from 8.9% to 6.9% in the ICP-based case.
- Under the data acquisition conditions (i.e., uncontrolled office-like illumination, frontal samples and neutral pose), the ICP-based matcher seems to consistently achieve a better performance than the Hausdorff distance system, independently of the scenario or the sensor considered.
- Both systems are equally affected by the change in the acquisition sensor (i.e., resolution) between enrolment and test, with a relative increase of their EER of over 100%.

Overall, the results depicted in Fig. 3 show the need to take into account the interoperability effect in the design of 3D face recognition systems. The variation in the sensor resolution clearly poses a big challenge to standard state-of-the-art 3D face recognition systems which experiment a significant decrease in their accuracy when two different devices are used for enrolment and test.

4 CONCLUSIONS

In the present work we have presented the first study on 3D face recognition interoperability using the new generation of low-cost 3D acquisition sensors. For this purpose, we have acquired a unique gender-balanced database which contains 3D face models of the same 26 subjects, captured with two sensors based on the same technology but with different resolution and acquisition ranges.

The results have shown the lack of robustness of two popular recognition systems to the change in the acquisition device between enrolment and test. The experiments have also confirmed previous evaluations were it was seen that, even though it is supposed to be more robust to illumination and pose changes (Bowyer et al., 2006), pure 3D face recognition technology (including acquisition and matching based only on the face geometry) is still not as mature and developed as 2D facial authentication.

Although the statistical significance of the study is limited due to the relatively small amount of data considered (i.e., 26 subjects), we believe that, from a qualitative point of view, the results show the high sensor-dependency of the assessed systems. Future work includes enlarging the database with further subjects and other low-cost 3D sensors as well as testing more advanced commercial algorithms for 3D face recognition. However, at its present stage, the work may still be seen as a reliable proof of concept of the studied interoperability problem.

In summary, the current study may be understood as a consistent and rigorous practical example which shows that, although many advances have been reached in the field of 3D face recognition, there are still open issues, such as the interoperability problem, which have been extensively explored in other more mature biometric modalities, but that still need to be properly addressed in this relatively young technology. In addition to the sensor issue, among the challenges that lie ahead the biometric community regarding 3D face interoperability, is the development of a data interchange standard similar to those already defined for other modalities (ISO/IEC, 2011), which would certainly help to maintain and homogenize performance across applications.

The same that, more than a decade ago, previous pioneering works initiated the discussion in the fingerprint trait (Ross and Jain, 2004), we believe that the present research can stimulate the community to look into the interoperability topic in 3D face recognition, in order to find ways to mitigate the problem and to develop algorithms intrinsically robust to the exchange of the acquisition sensors.

REFERENCES

Achermann, B. and Bunke, H. (2000). Classifying range images of human faces with Hausdorff distance. In *Proc. of ICPR*, pages 809–813.

Alonso-Fernandez, F., Fierrez, J., et al. (2008). Dealing with sensor interoperability in multi-biometrics: The UPM experience at the biosecure multimodal evaluation 2007. In *Proc. SPIE BTHI*.

Alonso-Fernandez, F., Fierrez-Aguilar, J., and Ortega-Garcia, J. (2005). Sensor interoperability and fusion in signature verification: A case study using tablet pc. In *Proc. ABPA*, Springer LNCS-3781, pages 180–187.

Alonso-Fernandez, F., Veldhuis, R. N. J., et al. (2006). Sensor interoperability and fusion in signature verification: A case study using tablet pc. In *Proc. ICARCV*, pages 422–427.

Amor, B. B., Ardabilian, M., and Chen, L. (2006). New experiments on ICP-based 3D face recognition and authentication. In *Proc. of ICPR*, page 11951199.

ANSI-INCITS (2004). ANSI INCITS 385-2004 face recognition format for data interchange.

Barnsley, M. F. (1993). *Fractals everywhere*, chapter Metric spaces; Equivalent spaces; Classification of subsets; and the space of fratals, pages 5–41. Morgan Kaufmann.

BEAT (2012). BEAT: Biometrics Evaluation and Testing. http://www.beat-eu.org/.

Besl, P. J. and McKay, N. D. (1992). Method for registration of 3-D shapes. *IEEE Trans. on PAMI*, 14:239–256.

Bowyer, K. W., Chang, K., and Flynn, P. (2006). A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition. *Computer Vision and Image Understanding*, 101:1–15.

Cignoni, P., Rocchini, C., and Scopigno, R. (1998). Metro: Measuring error on simplified surfaces. *Computer Graphics Forum*, 17:167–174.

Faltemier, T. and Bowyer, K. (2006). Cross sensor 3D face recognition performance. In *Proc. SPIE BTHI*.

Henrikson, J. (1999). Completeness and total boundedness of the hausdorff metric. *MIT Undergraduate Journal of Mathematics*, pages 69–80.

Huttenlocher, D. and Rucklidge, W. (1992). A multiresolution technique for comparing images using the hausdorff distance. Technical Report Technical Report 1321, Cornell University.

ISO/IEC (2011). ISO/IEC 19794-5:2011 information technology biometric data interchange formats.

Khiyari, H. E., Abate, A. F., et al. (2012). Biometric interoperability across training, enrollment, and testing for face authentication. In *Proc. IEEE BIOMS*.

Lu, X., Colbry, D., and Jain, A. (2004). Matching 2.5D scans for face recognition. In *Proc. ICPR*, pages 362–366.

Min, R., Kose, N., and Dugelay, J.-L. (2014). Kinectfacedb: A kinect database for face recognition. *IEEE Trans. on SMCS*.

NIST (2014). NIST speaker recognition evaluation (SRE) series. http://www.itl.nist.gov/iad/mig/tests/spk/.

Phillips, P. J., Flynn, P. J., Scruggs, T., Bowyer, K. W., Chang, J., Hoffman, K., Marques, J., Min, J., and Worek, W. (2005). Overview of the face recognition grand challenge. In *Proc. IEEE ICVPR*, pages 947–954.

Ross, A. and Jain, A. (2004). Biometric sensor interoperability: A case study in fingerprints. In *Proc. BioAW*, pages 134–145.

Wang, Y. and Chua, C.-S. (2006). Robust face recognition from 2D and 3D images using structural hausdorff distance. *Image and Vision Computing*, 24:176–185.

Zafeiriou, S., Hansen, M., et al. (2011). The photoface database. In *Proc. IEEE ICVPRW*, pages 132–139.