International Journal of Information Technology and Management (IJITM), Special Issue on Advances and Trends in Biometrics, Vol. x, No. y, 000-000 1

Identities, forgeries and disguises

Gérard Chollet and Patrick Perrot

CNRS-LTCI, Télécom ParisTech,
46 rue Barrault, 75634 PARIS cedex 13, France
Tel: +33 1 45 81 78 84, Fax: +33 1 45 81 37 94
E-mail: gerard.chollet@telecom-paristech.fr
E-mail: perrot@telecom-paristech.fr

Walid Karam and Chafic Mokbel

Departments of Computer Science and Computer Engineering,
University of Balamand, Al-Kurah, Lebanon
Tel: +961 6 930 250, Fax: +961 6 930 278
E-mail: walid.karam@balamand.edu.lb
E-mail: chafic.mokbel@balamand.edu.lb

Sanjay Kanade and Dijana Petrovska-Delacrétaz

Institut Télécom, Télécom SudParis
9 Rue Charles Fourier, 91011 Évry Cedex, FRANCE
Tel: +33 1 60 76 40 40, Fax: +33 1 60 76 43 37
E-mail: sanjay.kanade@it-sudparis.eu
E-mail: dijana.petrovska@it-sudparis.eu

Abstract: The preservation of your identity could become a major concern. In many situations, you need to claim an identity and this claim needs to be verified somehow. The technology called biometrics may help. But, what if a deliberate impostor claims your identity? Will this forgery be always detected? Biometric identity verification is imperfect. This paper reviews some of the techniques that a deliberate impostor could use to defeat a biometric verification system. It focuses on audio-visual forgeries using voice conversion and face animation. It also describes identity disguise as a means of falsifying and concealing one's identity. The recovery of an identity and cancelable biometrics are also useful techniques to protect from identity theft. Such techniques also find useful applications in multimedia.

Keywords: Identity verification, forgery, audio-visual imposture, voice conversion, face animation, identity disguise, identity recovery, cancelable biometrics.

Reference to this paper should be made as follows: Chollet, G., Perrot, P., Karam, W., Mokbel, C., Kanade, S., and Petrovska-Delacrétaz, D., 'Identities, forgeries and disguises', *International Journal of Information Technology and Management (IJITM), Special Issue on: "Advances and Trends in Biometrics,"* Vol. 9, No. 4, pp.xxx–xxx.

Biographical notes: Gérard Chollet is a CNRS research scientist within Télécom ParisTech (formerly ENST). He has supervised over 30 doctoral thesis in the area of Speech and Signal Processing. Until the doctoral level, his education was centered on Mathematics (DUES-MP), Physics (Maîtrise), Engineering and Computer Sciences (DEA). He studied Linguistics, Electrical Engineering and Computer Science at the University of California, Santa Barbara where he was granted a PhD in Computer Science and Linguistics. He joined CNRS (the french public research agency) in 1978 at the Institut de Phonétique in Aix en Provence. In 1981, he was asked to take charge of the speech research group of Alcatel. In 1983, he joined a newly created CNRS research unit at ENST, and headed the speech group before leaving temporarily for IDIAP. In 1992, he was asked to participate to the development of IDIAP, a new research laboratory of the "Fondation Dalle Molle" in Martigny, Switzerland. He initiated a successful collaboration with the Swiss Telecom-PTT and attracted some funding from the Swiss Confederation (from National and European programs). His main research interests include phonetics, automatic speech processing, speech dialog systems, multimedia, pattern recognition, digital signal processing, speech pathology, and speech training aids.

Patrick Perrot is an Officer of the French Gendarmerie and an engineer specialized in telecommunications. He was employed by Philips and Lucent Technologies before joining the French Gendarmerie in 1996. He has been working in the Signal, Speech and Image Department of the Forensic Research Institutes of the French Gendarmerie since 2001. He holds a PhD in Telecommunications from Télécom ParisTech (Paris, France).

Walid Karam has been teaching computer science and computer engineering courses at the University of Balamand (Al-Kurah, Lebanon) since 1993. His academic experience includes software development, Java technologies, operating systems, and various senior topics in computing. Other experiences include network management and e-learning initiatives. He is currently coordinating a UNESCO/UNDP funded project (ma3bar) to promote and develop Free and Open Source Software in the Arab region. He holds a Bachelor of Electrical Engineering from Georgia Institute of Technology (Atlanta, GA), a Master of Science in Computer Engineering from South Dakota State University (Brookings, SD), and a PhD in Telecommunications from Télécom ParisTech (Paris, France). His research interests include biometrics, face recognition, speech recognition, and digital signal processing.

Chafic Mokbel is an associate professor and a researcher at the University of Balamand since 2001. He is also the acting secretary

general of the University of Balamand Research Council. He graduated as electrical engineer from the Lebanese University (1988) and holds a PhD from Télécom-ParisTech (ENST) (1992) in Signal and Image Processing. He joint France Télécom Research and Development (CNET) in 1992 as an expert in voice technologies. He spent 1999 at IDIAP as the head of the Speech Group. He has supervised and/or co-supervised more than 15 PhD students in his fields of interest that cover, statistical signal processing, speech and image processing, graph theory, biometrics, cryptography, handwriting recognition, language modeling, statistical machine translation, multisensors signal processing, biomedical, brain-computer interface and, communication. He has developed several software tools like Becars, HCM, a VoiceXML gateway. Several of those tools have been evaluated in international evaluation compains. Since 1992, he has participated to several European projects. He has also been for 4 years assitant editor of the "Speech Communication" Elsevier Journal. Chafic Mokbel is a senior member of the IEEE.

Sanjay Kanade received a Masters degree in Instrumentation engineering from SGGS Institute of Engineering and Technology, Nanded, India. He holds a PhD degree from Institut Télécom SudParis, France (2010). His research interests include signal and image processing, biometrics, coding theory, and cryptography. His PhD research was in the field of information security by combining biometrics with cryptography.

Dijana Petrovska-Delacrétaz obtained a degree in Physics and a Doctoral degree from the Swiss Federal Institute of Technology (EPFL) in Lausanne, in 1982 and 1990 respectively. After a break for family reasons, she received a grant for women re-insertion of the Swiss National Science Foundation and started a new research activity in speech processing at the EPFL-LANOS (formerly CIRC) laboratory. After a year (starting in 1999) as a consultant at AT&T Speech Research Laboratories, and another year as a Post-doc at TELECOM-ParisTech (formerly GET-ENST) in Paris, she was working as a senior scientist in DIVA (Document, Image and Voice Analysis) group, Informatics Department, of Fribourg University, in Switzerland. Since September 2004 she holds an Assistant Professor position in the Intermedia group of the Electronics and Physics Department of Institut TELECOM-SudParis (ex GET-INT). Her research activities are mainly oriented towards pattern recognition, signal processing, and data-driven machine learning methods, that are exploited for different applications such as speech, speaker and language recognition, very low-bit speech compression, biometrics (2-D and 3-D face and voice), and crypto-biometrics (including privacy preserving biometrics). As per September 2010, her full list of publications (see also http://webspace.it-sudparis.eu/~petrovs) is composed of three patents, one book, 16 book chapters and journal papers, and 47 publications in conferences proceedings. She participated actively to the collection of two publicly available databases (for speaker recognition and biometrics evaluations), and one collection of open-source software.

1 Introduction

Biometrics comprises techniques to verify the identity of a person. These techniques rely on physiological and/or behavioral traits. This paper focuses on audio-visual traits and in particular, on forgeries and disguises of these traits.
In a first part, a review of automatic techniques to transform the voice and the face of an impostor is examined. Such techniques have not been originally developed for forgeries and have other applications. For instance, voice conversion is a very useful tool for the development of new voices in a Text-To-Speech system. Talking-face applications make use of voice and face transformations. Here, the main interest resides on the use of such techniques in forgeries which introduce specific constraints. Some proposals are made for voice conversion and face morphing. Evaluations of such techniques are discussed.
In a second part, we discuss techniques that could be used by someone who wants to hide his/her identity. Again, voice conversion and face animation is quite efficient to disguise one's identity. We review some of the techniques to detect voice disguise, which could be used in forensic applications.
Finally, we take a look at "cancellable biometrics", a set of techniques that allow users to renew their biometric templates and models in case these have been stolen and used by impostors. This is a possible solution to the potential problem of cross database matching.

2 Audio-visual forgery

The identity of a person is primarily determined visually by his face and audibly by his voice. These two modalities, i.e. face and voice, are used naturally by people to recognize each other. They are also employed by many identity recognition systems to automatically verify or identify humans for commercial, security and legal applications, including forensics. Audio-visual identity verification is introduced in [2]. However, altering the features of the face and/or the voice can be effectively used to trick an audio-visual identity verification system so as to have an impersonator be accepted as a genuine user.

Audio-visual forgery, or imposture, is the process of modifying both the face and the voice of an impostor to make them look and sound like those of an authentic client. It is reasonable to assume that an impostor has knowledge of the audio-visual recognition system techniques used on one hand, and, on the other hand, has enough information about the target client (face image, video sequence, voice recording). It would then be possible to use techniques of voice transformation and face animation as deliberate imposture methods to defeat the audio-visual system. Karam et al. [25, 26] propose voice transformation and face animation techniques to evaluate the effects of deliberate imposture on identity verification systems.

At the audio level, a voice transformation technique may be used. Several approaches have been proposed in the past decades, often for Multimedia or text-to-speech applications. State-of-the-art voice transformation algorithms take into consideration, in addition to the short-term spectra transformation, the pitch and the prosody of speech. Both short-term and long-term information included in the

speech signals are processed. The parameters of the short-term voice conversion function are generally estimated on a large amount of aligned speech from the source and the target speakers. In addition, state-of-the-art speaker recognition systems rarely consider long-term features such as prosodic or pitch information. Moreover, limited amount of client speech, if any, is generally available for training the system. To cope with those constraints, a voice transformation technique (MixTrans) [25] has been employed to change the perceived speaker identity of the speech signal of the impostor to that of the target client. MixTrans requires a limited amount and not necessarily aligned speech from source and target speakers.

At the visual level, an animation of the impostor face is achieved using a thin-plate spline warping [25]. Face animation is equally achieved using commercial animation packages such as CrazyTalk [17]. Abboud et al. [1] proposed appearance-based lip tracking and cloning on speaking faces as a means of face transformation.

Imposture results indicate an increase in equal error rates from 5.1% to 15.39% on the performance of the audio-visual verification system, implying a higher impostor acceptance rate.

The drop in performance observed when applying audio-visual forgery stresses the need to increase the robustness of the system. The imperfections in the transformed audio-visual scene may help to detect forgeries. To overcome the challenges imposed by deliberate imposture on audio-visual identity verification systems, Bredin et al. in [5, 8, 6, 7] provide a study on the level of audio visual synchronization as a means of imposture detection, which helps make talking face authentication robust to deliberate imposture.

Audio-visual identity verification systems and effects of deliberate imposture are reported within the framework of the BioSecure project [11, 8, 14]. A guide to biometric reference systems and performance evaluations is provided in [34].

3 Identity disguise

Voice disguise is considered as a deliberated action of the speaker who wants to falsify or to conceal his identity [33, 32]. A relevant way to mask one's identity is to use a simple but efficient disguise. Lots of possibilities are offered to a speaker to change his voice. In the field of automatic speaker recognition application, one of the most efficient threats is voice disguise. Based on crime analysis in Germany, it is noted in [28] that there was "... an overall occurrence of voice disguise in 52 percent of the cases where the offender used his/her voice and may have expected to have it recorded during the criminal action. This percentage includes cases of blackmailing, where the specific percentage was as high as 69 percent." And in [12] it is noted that "Disguised speech is typically found in situations in which the criminal thinks he is being recorded. This situation is very common in cases of kidnapping, a kind of crime whose incidence has increased considerably in the past years in Brazil."

This section focuses on four specific non-electronic and deliberated disguises according to the classification proposed by [37]: High pitched voice, low pitched voice, covered mouth, and pinched nostrils [32]. An experiment based on a 50 speaker database is proposed. 30 speakers utter the same sentence and those utterances are used to train the classifiers. 10 speakers utter sentences with a wide phonetic coverage and are used for development. The remaining speakers are used

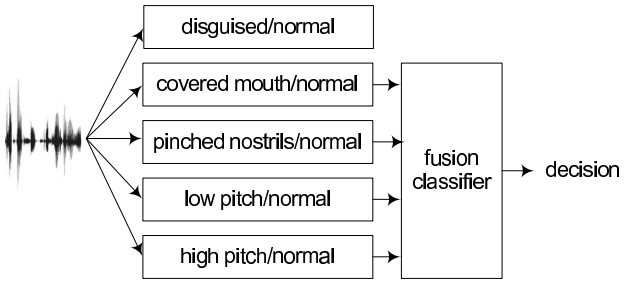
for testing. A set of acoustic features in the speech signal, including the formants F_1 and F_2, meanF_0, minF_0, maxF_0, 12 MFCC and their first derivatives are extracted from each speech segment. Different classification techniques are evaluated to detect disguise: k-nearest neighbours and Support Vector Machine (SVM). Two different architectures of SVM classifiers fusion have been experimented. The first architecture consists in a parallel combination of five SVM classifiers as proposed in Fig. 1(a). The second architecture consists in a hybrid combination of classifiers as proposed in Fig. 1(b). The level of performance of each classifier is based on the analysis of ROC (Receiver Operating Characteristic) curve and the criterion linked to the curve, the Area Under Curve (AUC). The ROC graph is a useful technique for organizing classifiers and visualizing their performance.

Fig. 2 proposes the results of each classifier between normal voice and a combined of four disguises composed by the different disguises previously described. This curve reveals a good level of performance for the parallel fusion architecture and the SVM classifier with an AUC of 0.79 and 0.78 respectively. The hybrid fusion architecture presents an AUC of 0.72 and the 5 nearest neighbours an AUC of 0.67. In the area of forensic speaker recognition it could be interesting to realize a step of disguise detection as pre-processing in order to avoid directing the investigation toward unlikely suspects and away from likely ones.

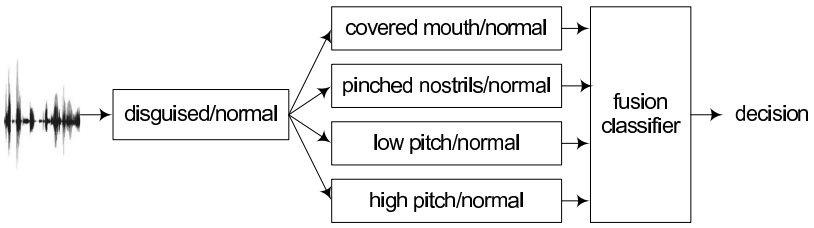
A more sophisticated method to disguise his voice is to use voice conversion considered as an electronic-deliberated technique. In [28] a proposition of original voice conversion technique to trick an automatic speaker recognition system is proposed where a degradation of near than 50% is noticed on the level of performance after the conversion. Compared to the audio-visual forgery methods, these methods make use of sufficient aligned data. In [31] three different methods of voice conversion are compared with the forensic application of imitating a French politician.

4 Identity Recovery

In addition to voice and face recognition, 3-D face model may be used as a biometrics modality. It is believed that the animation and disguise of a 3-D face model is harder than face recognition forgery. In this section we present an algorithm for constructing a 3-D textured model of a person's face from two photos. This subject has been lately studied extensively. The proposed approaches often make use of a determined number of views and/or strict conditions on how the 2-D views are taken. Our goal is to obtain a realistic 3-D model from an undetermined number of photos. From two views, the 3-D model can then be refined.
We present applications of this work in forensic sciences, the aim being to recover the identity of an individual. Video surveillance produces hours of video data, however it often lacks quality and, frontal or profile views may not be available in the data. As mentioned earlier, estimating a 3-D textured model can be a first step in a robust person identification. Indeed, the frontal view of the model can be compared to frontal views of other persons making face recognition part of it. Another lead is the comparison between two 3-D models, for instance by computing 3-D distances between specific points of the face.



(a) Parallel classifier fusion



(b) Hybrid classifier fusion

Figure 1 Classification architectures for voice disguise detection

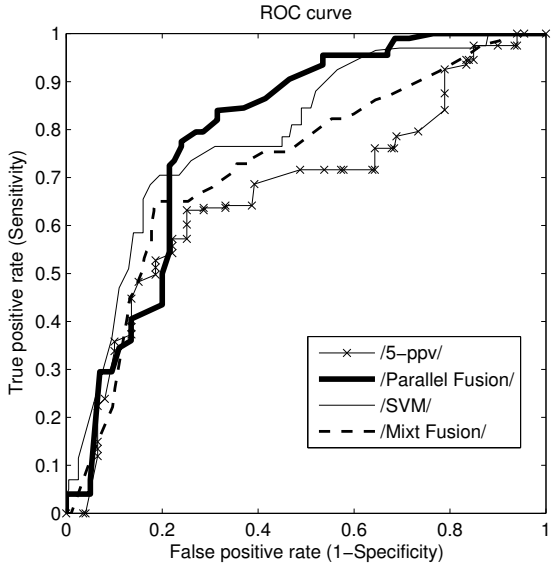


Figure 2 Performance of normal and disguised voices

4.1 3-D face reconstruction methods

In comparison with 2-D face images, 3-D face models have the advantage of being illumination and pose invariant. This is particularly relevant when handling changing environments in practical surveillance. Many different approaches tackle the reconstruction of a 3-D model from one or several 2-D views of a face. Feature detection plays an important role in the process and directly affects the accuracy and robustness of the resulting reconstruction. The features can be detected automatically (which is faster) or manually (which may be more accurate, especially on low-resolution images). When video sequences are available, the temporal coherence of the frames is an added bonus in the estimation of the pose. Particular frames, for instance orthogonal ones, can be extracted from the video when available [10]. The 3-D model of the face is produced by deformation of a generic 3-D model or by a combination of several models from a database. The reconstruction can be the result of an optimization problem [3] or simple deformation of a generic model from the correspondence between feature points on the 3-D generic model and the 2-D images [30]. Such a 3-D model could also be obtained from a single 2-D image. Whereas it may suffer from a lack of information compared to other methods using more than two different views, the use of databases, for instance of eigenhead models [39], or a statistic model [15], show promising results. Finally, 3-D information could also be available from 3-D scans. The surface may then be obtained by minimization of a set of constrained equations [13]. We use a predefined generic 3-D geometric model which will match the specific face from two or more images by matching points. The features will be manually selected for more accuracy, especially true for low-resolution 2-D data. The textures are then retrieved from the 2-D images by ray tracing. The maps of coordinates and color that are obtained will then be used together to produce a global smooth surface.

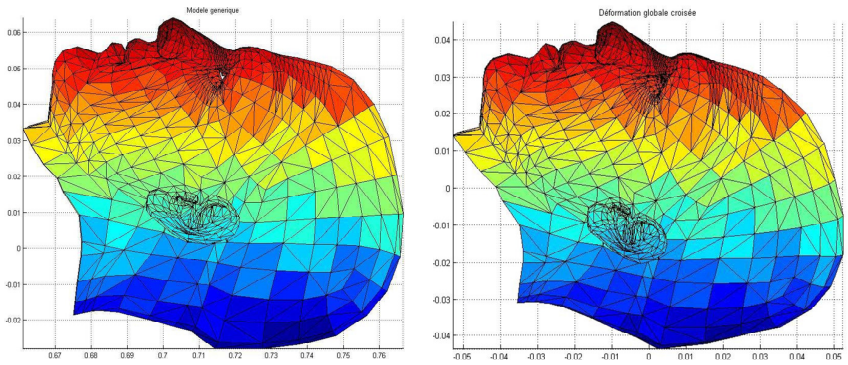
4.2 3-D model from a frontal view and a profile view

After a preliminary step to adjust the characteristics of both photos such as light, contrast, etc. We outline three main steps:

- retrieving peculiar data, for instance the positions of the eyes, the chin, the mouth, the nose, the contour of the face, and the ears;
- estimating a geometrical 3-D model of the face;
- texturing this geometrical model.

4.2.1 Retrieving the data

In [30], algorithms are presented to automatically detect facial feature points. In forensic applications it is not so crucial to automatically detect facial features. Therefore in our work the user is required to select pairs of matching points. For each pair, one of the points is taken on the 2-D view (a photo), the other one on a generic 3-D geometrical model for the same face feature (e.g. the center of one eye)



(a) Generic model (b) Deformed model

Figure 3 Profile views

4.2.2 Deformed 3-D geometrical model

Following the method described in [10], a deformation matrix is computed for each view from the pairs of matching 2-D and 3-D points. A RBF (radial-based functions) model is used, the coefficients or the matrix are the coefficients of the model. A global deformation matrix, taking into account the two sets of matching pairs, is then applied to the whole 3-D geometrical generic model to obtain a geometrically deformed 3-D model of the face, which satisfies both 2-D views.

4.2.3 Textured 3-D model

If geometry already carries essential data about a face, it is far from sufficient to lead to successful identification, especially if the comparison is visual. Colors must be retrieved from the photos and applied to the deformed 3-D model. However, one must be cautious as to how the mapping is done in order to obtain a realistic rendering. Blending will be an essential step and color data must thus be stored in a practical way.

In order to build color maps, we resorted to ray tracing. Just interpolating the 3-D coordinates (on the deformed geometrical 3-D model) as well as the color of the different points (on the photos) manually defined by the user would have been far too less sufficient. We use a rectangular grid with rays orthogonally projected onto the deformed 3-D model. The impacts give us the spatial coordinates, the corresponding position on the grid corresponds to a pixel, or a group of pixels, of the photo, which gives us the matching colors. The spatial coordinates are normalized by a spherical projection, and the maps (one map for each spatial coordinate, one another for the color data) are internally filled by a k-nearest neighbors approach.

The four triplets of maps (one third for a profile view, one another for the frontal view, the last third for the other profile view simply being the symmetry of the first one) are then blended together to form four final maps. A frontier is first computed between the first profile view maps and the frontal view ones, and the frontal view maps and the second profile view ones. Finally, the maps are smoothed around the two frontiers. The four final maps give us the textured 3-D model.

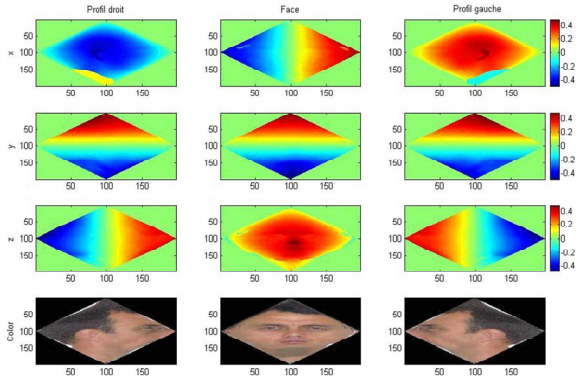


Figure 4 The 12 maps: x,y,z,Color (rows); right profile, frontal, left profile views (columns)

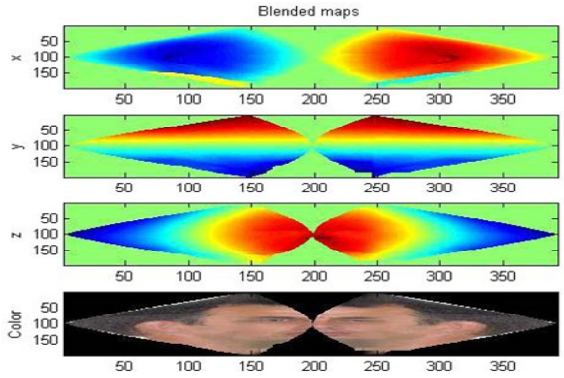


Figure 5 Blended and smoothed final maps

Whereas the symmetry of the human face is a general assumption, it may be partially incorrect. For instance, the hair in the area of the forehead will not often be symmetric. Therefore there will be discontinuities between the color maps, maybe even the coordinates maps, in those areas. One solution is to discard the second profile. The 3-D textured model will be more incomplete but without any discontinuities. If two profiles (one left, one right) are available, it will be possible to obtain a more accurate geometric deformed 3-D model. A ray tracing on the frontal and the two profile views will produce 3 sets of maps. The blending of those 3 sets will generate a correct complete result without discontinuities.

4.3 Conclusions on Identity Recovery from 2-D Images

The proposed method can produce a realistic 3-D textured model of a face from two photos of the face. Applications in forensic sciences and especially concerning identification appear very relevant. However, a lack of accuracy concerning the global dimensions of the face if the profile and frontal views are not available, can lead to a dilated model, which could hinder visual identification as well as automatic 2-D or 3-D comparison. A more robust approach than the selection of extreme points for solving this problem may be the use of specific anthropometrical points. Another drawback of our approach is the selection by the user of points on the 2-D view and 3-D generic model: while it benefits from an increased accuracy, it also slows the whole process. Finally, although the results obtained with uncovered faces are convincing, some difficulties with occluded faces may appear. It may be efficient to detect such difficult zones and discard them (another view bringing the missing data) or pay particular attention to them, using symmetry or using generic data for missing patches.

5 Cancelable biometrics

With more and more applications using biometrics, new privacy and security risks arise. Personal (including biometric) information could be tracked from one application to another by cross-matching between biometric databases, thus compromising privacy. A crucial issue is the potential misuse of collected data. Questions like "What can I do if my biometric data has been stolen or misused?" require urgent attention not only to reassure users with regards to privacy intrusion but also to prevent misuse and improve accuracy. Moreover, since standard biometric templates are permanently associated with an individual, they could not be used anymore in case they are compromised. Since they cannot be replaced, they are also inherently non-revocable. This makes "classical" biometric systems unfit for privacy and security critical applications. Therefore, these major issues of biometric systems that guarantee the rules of privacy protection should be solved urgently.

In France, it is the C̈ommission Nationale de l'Informatique et des Libertés(C̈NIL) organization that guarantees that data privacy laws are applied to the collection, storage, and use of personal data, by issuing authorizations for biometrics applications.

Over the last decade, a new innovative multidisciplinary research field has emerged, that combines biometrics and cryptography, and that has the capability

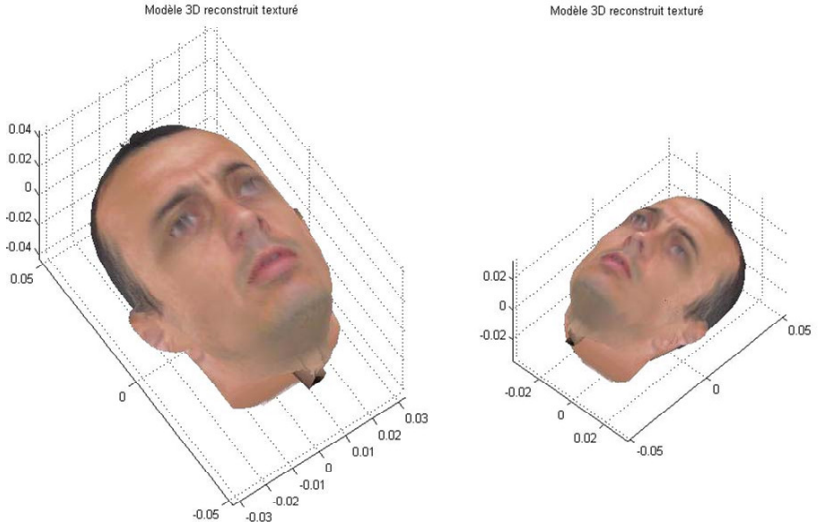


Figure 6 3-D textured model, two different views

to guarantee biometric data privacy in an algorithmic way. The resulting innovative hybrid systems have the following important properties: they confer to biometric characteristics the needed capabilities of revocability, privacy, and diversity, and provide cryptographic systems with a strong link to the user through biometrics.

Many solutions, termed as cancelable biometrics, are proposed in order to overcome the problems of revocability and cross-database matching of biometrics [35, 36, 18, 27, 38, 4]. These solutions basically involve combination of some assigned user specific secret with the biometric characteristics. The use of an assigned secret allows the revocation of the template if the template is compromised.

Kanade et al. [21, 22] proposed a simple shuffling scheme which randomizes the biometric data with the help of a shuffling key. The data to be shuffled is divided into blocks and these blocks are rearranged according to the bit values of the shuffling key. The advantages of this scheme are: (a) possibility of revocability in biometric systems, (b) improvement in the verification performance (nearly 80% decrease in equal error rate) because it increases the impostor Hamming distance without changing the genuine Hamming distance, (c) template diversity, (d) impossibility of cross-matching and therefore protection of privacy.

There are also a number of systems in the literature with which a long and stable bit-string can be derived from biometrics [20, 19, 16, 29]. Such systems can also possess the properties of revocability, template diversity, and privacy protection.

Kanade et al. proposed such key regeneration systems using uni-biometrics (iris [21]) as well as multi-biometrics (two iris system [23] and multi-modal system using a combination of iris and face [24]).

Such systems can help if an impostor has stolen the biometric data. In this case, these systems allow revocation of the compromised template and re-enrollment of the user with the same biometric data which is not possible with classical biometric systems.

6 Conclusions and perspectives

Biometric recognition systems have known important development in the last decades. We believe that forgeries and disguises are threats to those systems. Dummy finger and irises can compromise such systems and Audio-visual forgeries are simple and efficient. We have shown that voice conversion and face animation technologies can challenge seriously the audio-visual biometric systems even when limited amount of biometric data are available from either the source of the target person. A drop that may reach 50% in the performance of such systems is observed. This makes the audio-visual biometrics non robust to such forgery attacks and limit its application fields.

We started exploring two solutions to increase the robustness of the existing biometric systems: 3-D face models and cancelable biometrics. With 3-D face models constructed from one or several 2-D views of a person, the face recognition becomes a particular case that uses a distance of likelihood on the frontal view alone. However, the construction of such models is not trivial and has to overcome a non controlled measurement of the 2-D views.

Cancelable biometrics is a possible solution to counteract the possibilities of cross-matching biometric databases. We have proposed key-regeneration systems based on uni-biometrics and multi-biometrics. Such algorithms allow revocation of the compromised templates using the same biometric data in case such data have been stolen. We do believe that the development of this approach will certainly increase the robustness of biometrics recognition systems to forgeries and disguises.

Acknowledgements

The authors would like to acknowledge the support of the French "Agence Nationale de la Recherche" (ANR) within the project BIOTYFUL (ANR-06-TCOM-018), the Franco-Lebanese project "Programme de Coopération pour l'Évaluation et le Développement de la Recherche (CEDRE), and the Biometrics for Secure Authentication (BioSecure Network of Excellence, an FP6 project).

References

[1] Abboud, B. and Chollet, G. (2005) 'Appearance based lip tracking and cloning on speaking faces', In *Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis, 2005. ISPA 2005.*, pp. 301-305.

[2] Abboud, B., Bredin, H., Aversano, G. and Chollet, G. (2007) 'Audio-visual identity verification: an introductory overview', In *Progress in NonLinear Speech Processing*, pp. 118-134, Springer-Verlag, Berlin, Heidelberg.

[3] Amin, S. and Gillies, D. (2007) 'Analysis of 3-D face reconstruction', In *ICIAP'07: Proceedings of the 14th International Conference on Image Analysis and Processing*, pp. 413-418, Washington, DC.

[4] Boult, T.E., Scheirer, W.J. and Woodworth, R. (2007) 'Revocable fingerprint biotokens: Accuracy and security analysis', In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-8, June 2007.

[5] Bredin, H. and Chollet, G. (2006) 'Measuring audio and visual speech synchrony: Methods and applications', In *International Conference on Visual Information Engineering, VIE 2006*, pp. 255-260.

[6] Bredin, H. and Chollet, G. (2007) 'Audiovisual speech synchrony measure: application to biometrics', In *EURASIP J. Appl. Signal Process.*, pp. 179-182, Vol. 1.

[7] Bredin, H. and Chollet, G. (2008) 'Making talking-face authentication robust to deliberate imposture', In *IEEE Conference on Acoustics, Speech and Signal Processing, ICASSP 2008*, pp. 1693-1696, Vol. 1, April 2008.

[8] Bredin, H., Aversano, G., Mokbel, C. and Chollet, G. (2006) 'The BioSecure talking-face reference system', In *2nd Workshop on Multimodal User Authentication.*

[9] Bredin, H., Miguel, A., Witten, I.H. and Chollet, G. (2006) 'Detecting replay attacks in audiovisual identity verification', In *IEEE Conference on Acoustics, Speech and Signal Processing, ICASSP 2006*, pp. 621-624.

[10] Chen, M. and Hauptmann, E. (2004) 'Toward robust face recognition from multiple views', In *International Conference on Multimedia and Expo (ICME'04)*, pp. 27-30, 2004.

[11] Chollet, G., Aversano, G., Dorizzi, B. and Petrovska-Delacrétaz, D. (2005) 'The first BioSecure residential workshop', In *4th International Symposium on Image and Signal Processing and Analysis*, Zagreb, Croatia.

[12] de Figueiredo, R.M. and de Souza Britto, H. (1996) 'A report on the acoustic effects of one type of disguise', In *Forensic Linguistics*, pp. 168-175, Vol. 3(1).

[13] Elyan, E. and Ugail, H. (2007) 'Reconstruction of 3-D human facial images using partial differential equations', In *JCP*, pp. 1-8, Vol. 2(8).

[14] Fauve, B., Bredin, H., Karam, W., Verdet, F., Mayoue, A., Chollet, G., Hennebert, J., Lewis, R., Mason, J., Mokbel, C. and Petrovska-Delacrétaz, D. (2008) 'Some results from the BioSecure talking face evaluation campaign', In *IEEE Conference on Acoustics, Speech and Signal Processing, ICASSP 2008*, pp. 4137-4140.

[15] Guan, Y. (2007) 'Automatic 3-D face reconstruction based on single 2d image', In *MUE'07: Proceedings of the 2007 International Conference on Multimedia and Ubiquitous Engineering*, pp. 1216-1219, Washington, DC.

[16] Hao, F., Ross, A. and Daugman, J. (2006) 'Combining crypto with biometrics effectively', In *IEEE Transactions on Computers*, pp. 1081-1088, Vol. 55(9).

[17] Reallusion Inc. (http://www.reallusion.com/crazytalk/).

[18] Beng Jin, A.T., Ngo, D., Ling, C. and Goh, A. (2004) 'Biohashing: two factor authentication featuring fingerprint data and tokenised random number', In *Pattern Recognition*, pp. 2245-2255, Vol. 37(11), November 2004.

[19] Juels, A. and Sudan, M. (2002) 'A fuzzy vault scheme', In *Proc. IEEE Int. Symp. Information Theory*, A. Lapidoth and E. Teletar, editors, pp. 408, IEEE Press, 2002.

[20] Juels, A. and Wattenberg, M. (1999) 'A fuzzy commitment scheme', In *Proceedings of the Sixth ACM Conference on Computer and communication Security (CCCS)*, pp. 28-36, 1999.

[21] Kanade, S., Camara, D., Krichen, E., Petrovska-Delacrétaz, D. and Dorizzi, B. (2008) 'Three Factor Scheme for Biometric-Based Cryptographic Key Regeneration Using Iris', In *The 6th Biometrics Symposium (BSYM)*, September 2008.

[22] Kanade, S., Petrovska-Delacrétaz, D. and Dorizzi, B. (2009) 'Cancelable Iris Biometrics and Using Error Correcting Codes to Reduce Variability in Biometric Data', In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 2009.

[23] Kanade, S., Petrovska-Delacrétaz, D. and Dorizzi, B. (2009) 'Multi-Biometrics Based Cryptographic Key Regeneration Scheme', In *IEEE International Conference on Biometrics: Theory, Applications, and Systems (BTAS)*, September 2009.

[24] Kanade, S., Petrovska-Delacrétaz, D. and Dorizzi, B. (2010) 'Obtaining Cryptographic Keys Using Feature Level Fusion of Iris and Face Biometrics for Secure User Authentication', In *IEEE CVPR Workshop on Biometrics*, June 2010.

[25] Karam, W., Bredin, H., Greige, H., Chollet, G. and Mokbel, C. (2009) 'Talking-face identity verification, audiovisual forgery, and robustness issues', In *EURASIP Journal on Advances in Signal Processing*, 2009.

[26] Karam, W., Mokbel, C., Greige, H. and Chollet, G. (2009) 'Audio-visual identity verification and robustness to imposture', In *Advances in Biometrics, Third International Conference, ICB 2009*, Massimo Tistarelli and Mark S. Nixon, editors, Vol. 5558, pp. 796-805, Lecture Notes in Computer Science, June 2009.

[27] Lumini, A. and Nanni, L. (2007) 'An improved biohashing for human authentication', In *Pattern Recognition*, pp. 1057-1065, Vol. 40(3), March 2007.

[28] Masthoff, H. (1996) 'A report on a voice disguise experiment', In *Forensic Linguistics*, pp. 160-167, Vol. 3(1).

[29] Nandakumar, K., Jain, A.K. and Pankanti, S. (2007) 'Fingerprint-based fuzzy vault: Implementation and performance', In *IEEE Transactions of Information Forensics and Security*, Vol. 2(4), pp. 744-757, December 2007.

[30] Park, I., Zhang, H. and Vezhnevets, V. (2005) 'Image-based 3-D face modeling system', In *EURASIP J. Appl. Signal Process.*, pp. 2072-2090, 2005.

[31] Perrot, P., Morel, M., Razik, J. and Chollet, G. (2009) 'Vocal forgery in forensic sciences', In *Forensics in Telecommunications, Information and Multimedia, Second International Conference, e-Forensics 2009*, Vol. 8, pp. 179-185, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Springer, January 2009.

[32] Perrot, P. and Chollet, G. (2008) 'The question of disguised voices', In *Acoustics 08*, July 2008.

[33] Perrot, P., Preteux, C., Vasseur, S. and Chollet, G. (2007) 'Detection and recognition of voice disguise', In *proceedings of International Association for Forensic Phonetics and Acoustics Conference 2007*, pp. 3, Plymouth, UK, July 2007.

[34] Petrovska-Delacrétaz, D., Chollet, G., and Dorizzi, B. (2009) 'Guide to Biometric Reference Systems and Performance Evaluation', Springer Publishing Company, Incorporated, 2009.

[35] Ratha, N.K., Connell, J.H. and Bolle, R.M. (2001) 'Enhancing security and privacy in biometrics-based authentication systems', In *IBM Systems Journal*, Vol. 40(3), pp. 614-634, 2001.

[36] Ratha, N.K., Chikkerur, S., Connell, J.H. and Bolle, R.M. (2007) 'Generating cancelable fingerprint templates', In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29(4), pp. 561-572, April 2007.

[37] Rodman, R.D. (1998) 'Speaker recognition of disguised voices', In *Consortium on Speech Technology Conference on Speaker Recognition by Man and Machine: Directions for Forensic Applications COST250*, 1998.

[38] Savvides, M., Vijaya Kumar, B.V.K. and Khosla, P.K. (2004) 'Cancelable biometric filters for face recognition', In *Proceedings of the 17th International Conference on Pattern Recognition (ICPR04)*, Vol. 3, pp. 922-925, August 2004.

[39] Wang, S.F. and Lai, S.H.. (2008) 'Efficient 3-D Face Reconstruction from a Single 2D Image by Combining Statistical and Geometrical Information', In *Computer Vision ACCV 2006*, Lecture Notes in Computer Science, Vol. 3852, pp. 427-436, Springer, 2006.