LIVENESS DETECTION BASED ON 3D FACE SHAPE ANALYSIS

Andrea Lagorio, Massimo Tistarelli, Marinella Cadoni (University of Sassari, Italy)
Clinton Fookes, Sridha Sridharan (Queensland University of Technology, Brisbane, Australia)

ABSTRACT

In recent years face recognition systems have been applied in various useful applications, such as surveillance, access control, criminal investigations, law enforcement, and others. However face biometric systems can be highly vulnerable to spoofing attacks where an impostor tries to bypass the face recognition system using a photo or video sequence.

In this paper a novel liveness detection method, based on the 3D structure of the face, is proposed. Processing the 3D curvature of the acquired data, the proposed approach allows a biometric system to distinguish a real face from a photo, increasing the overall performance of the system and reducing its vulnerability.

In order to test the real capability of the methodology a 3D face database has been collected simulating spoofing attacks, therefore using photographs instead of real faces. The experimental results show the effectiveness of the proposed approach.

Index Terms — 3D Face recognition, liveness detection, Biometrics, Pattern recognition, Shape analysis;

1. INTRODUCTION

In practical biometric applications, the data acquisition phase can be crucial to identify spoofing attacks. For example, in a face recognition system, a synthetic biometric feature attack can be simply performed by showing a face portrait to the camera (Figure 1). This can be achieved either with a photograph or a mask portraying a genuine user.

Most face recognition systems operate on 2D images using visible or infrared light. Their sensitiveness to spoofing attacks can vary significantly [1]. However, without the intervention of a robust spoofing detection front-end, any recognition system could be quite easily fooled.

To avoid this potential danger a biometric system should be able to detect the movement of the face. This can be achieved holistically or through the extraction and tracking of facial features in order to detect genuine expressions or natural involuntary movements (smile or eye blinking) [2]. All these countermeasures can be by-passed by showing a video clip of a genuine user to the camera.

In order to increase the robustness of the system to similar attacks, more elaborate anti-spoofing techniques can verify the three-dimensionality of the face captured by the device, for example by laser scanning. Another approach to the problem involves the use of specialized extra hardware (e.g. thermal sensors) allowing the detection of user liveness. Obviously, the adoption of additional tools increases the cost and the complexity of the recognition system.

In the following, a methodology based on optoelectronic 3D scanning is presented. This approach has the advantage of using full 3D face information, however, in contrast to laser scanning, the 3D model is acquired in a few milliseconds minimising the user cooperation.

Holding a perfect frontal pose for seconds, as is often required by different imaging techniques, can be avoided in favour of more natural behaviour.

Our proposed method could be implemented in different scenarios: either as an anti-spoofing tool, coupled with 2D face recognition systems known to be vulnerable to such attacks [7]; or to be integrated within an automatic 3D face recognition systems to perform an early detection of spoofing attacks. In this latter case, by avoiding to attempt the recognition phase, the computation time is reduced and accidental recognition errors are avoided.

Experimental results show that the proposed spoofing detection approach is able to detect spoofing attacks with high accuracy.

Figure 1: Example of a sensor spoofing attack.

978-1-4673-4989-5/13/$31.00 ©2013 IEEE

2. RELATED WORKS

A pioneering work on liveness detection based on 3D structure was presented by Choudhury et al. in 1999 [3]. Kollreider et al. [4] proposed a method based on optical flow to detect non-rigid local deformations of facial components. Both these systems were very sensitive to the image quality of the video stream.

Image quality plays an important role also in the method proposed by Li et al. [5]. They based their work on the observation that the Fourier spectrum generated by a photograph presents significant differences with respect to the one generated by a real face. Määttä et. al [7] proposed a method based on the analysis of the facial texture using the Local Binary Pattern descriptor (LBP) to distinguish between images acquired from a real person and images acquired from a photo.

Other approaches in the literature are based on multimodal analysis or additional interaction with the subject. For instance, Frischholz and Dieckmann [6] fuse face and voice. These methods, however, inevitably increase the complexity and the computational time of the system.

De Marsico et al. [8], proposed a method based on the computation of geometric invariants to determine if the captured image is the projection of a real three-dimensional face. The most robust face spoofing detection techniques proposed to date all rely on two main activities: (i) the verification of the real 3D face shape, and (ii) an active interaction with the user. Both processes may require additional hardware and the design of complex algorithms to enable active cooperation with the subject. While the former can incur an additional cost for the system, the latter cannot be achieved in a covert application or whenever the subject does not want to be identified.

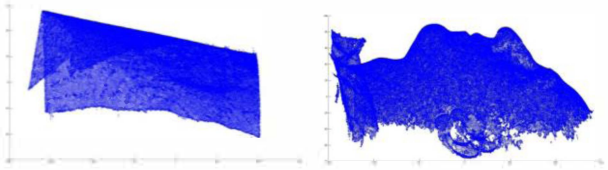

Figure 2: Example 3D acquisition of a picture of a human face (left), and a real 3D face (right).

3. THE PROPOSED APROACH

The aim of the proposed technique is to determine if an impostor is employing a 2D image of a genuine user to fool a face recognition system. The proposed method computes the 3D features of the captured face data to determine if a human face has been presented to the acquisition camera.

Figure 2 (left) shows an example of a 3D acquisition from a bended 2D photographic source, i.e. a projection of the point cloud on the XY plane. The lack of surface variation in the scan (i.e. very low surface curvature) is clear evidence the acquisition comes from a 2D source.

Figure 2 (right) illustrates the acquisition of a real 3D face. To compare the two 3D scans, a simple and fast method can be implemented, based on the computation of the mean curvature of the surface. Given a 3D scan F, an approximation of the actual curvature value at each point $\mathbf{p}$ is computed from the principal components of the Cartesian coordinates within a given neighborhood. The singular value decomposition, or PCA, is computed from the covariance matrix of the Cartesian coordinates of all points lying within a spherical neighborhood Ω_r of radius r centered at point $\mathbf{p}$. The value of the curvature at $\mathbf{p}$ is computed as:

$$C = \frac{(\mathbf{p} - \mathbf{b}) * \mathbf{v}}{d^2}$$

where $\mathbf{v}$ is the eigenvector corresponding to the smallest eigenvalue of the decomposition, $\mathbf{b}$ is the baricenter of the Cartesian coordinates of the points within Ω_r, and d is the mean distance of all points within Ω_r. As the radius r controls the degree of smoothness of the curvature along the 3D surface, the localization accuracy can be varied by changing the value of r [10, 11]. The mean curvature of the 3D points lying on the face surface is computed as the arithmetic average of the curvature values of all the points. In Figure 3, a comparison between the 3D data acquired from a photograph and from a real 3D face is shown. The color codes the curvature values: blue represents low curvature values and red represents high curvature values. The mean curvature computed from the photograph in Figure 3, with a value of r equal to 5, is equal to 0.004634. The mean curvature computed from the real face is equal to 0.036957. The large difference (an order of magnitude) between the two curvature values clearly indicates the discriminative power of the proposed method.

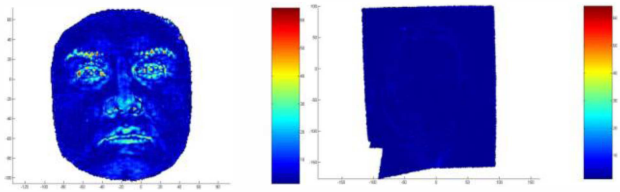

Figure 3: Curvature values computed from the 3D data captured from a real human face (left) and a printed picture of the same (right).

The proposed approach has several advantages over other liveness detection techniques:

- It does not require active interaction with the subject, nor smiling, speaking or responding to any external prompt;
- It does not require additional hardware to be used (such as a microphone). The 3D acquisition system used for the recognition phase can be simply adapted for this pre-processing step. If 2D faces are used for

recognition, a set of acquired face images can be used to reconstruct the 3D shape;

- No constraints are required on the head pose and orientation.

4. THE 3D SPOOFING TEST DATABASE

To test the effectiveness of the proposed liveness detection technique a set of 3D face scans was acquired using an optoelectronic stereo system (VECTRA 3D CRT, see [9]). It is a structured light system, consisting of two high resolution cameras for texture acquisition, two projectors, and four frontal calibrated cameras. An area of about 40x40 cm can be captured with a single 3D scan. The acquisition speed is about 2 milliseconds. Flashing lights are incorporated into the system so that ambient light does not need to be controlled. The output is a 3D virtual model of the face, consisting of a mesh of approximately 80,000 points and the associated texture. A picture of the VECTRA system is shown in Figure 4.

Figure 4: Vectra 3D CRT system.

A database composed of mixed 3D scans from real human faces and 2D pictures has been collected using the Vectra system. The spoofing attacks were simulated by 70 3D scans acquired from different photographs with different head pose and orientation[1], which will be called the "fake" set (FS).

For the set of "genuine users", 22 scans from 11 subjects were acquired, which will be called the "genuine" Vectra set (GVS).

To test the efficacy of the method on a larger database, 70 frontal 3D scans of 70 subjects from the Bosphorus database were also considered. The Bosphorus Database [12, 13] is a 3D face dataset including a rich set of expressions, systematic pose variations and different types of occlusions. The 3D facial data was acquired using a structured-light based 3D system. Acquisitions are single view, and subjects were made to sit at a distance of about 1.5 meters away from the 3D digitizer. The sensor resolution along the X, Y & Z (depth) axes are 0.3mm, 0.3mm and 0.4mm respectively, and the colour texture images have a resolution of 1600x1200 pixels. A 1000W halogen lamp was used in a dark room to obtain homogeneous lighting for good quality texture images.

5. EXPERIMENTAL RESULTS

Two experiments were designed. In the first one, the sets FS and GVS were used, with the size r of the spherical neighborhood set to 6. The distribution of the mean curvature values for the two sets were well separated, and the value of the False Rejection Rate (FRR), was equal to zero.

In the second experiment we used the FS and the Bosforus sets. In order to determine the sensitivity of the algorithm to variations in the size r, several experiments with different values of r, ranging from 4 to 20, were performed. The distribution of the mean curvature values of the 3D scans, both from pictures (impostors) from the "fake" set and real subjects (clients), for different values of the radius r, is shown in Figure 5. The value of the False Rejection Rate (FRR) at rank 1, for different values of r, is always equal to zero. As it can be noticed, the separation of the genuine client and impostor distributions increases with the radius of the neighbourhood. This demonstrates the robustness of the algorithm to parameterization.

To test the sensitivity of the algorithm to artificial surface variations, the same tests were performed with the same pictures but creased, i.e. with a surface which is not perfectly planar and uniform. In this case we expected the system to fail or to report a lower discrimination capability. As from the performed experiments, the separation between the client and impostor distribution does decrease, but still increases with the radius r of the considered spherical neighbourhood. Indeed, as long as the "wrinkles" on the paper (or other material used to impress the picture) are not too severe to impair recognition from the 2D geometry of the depicted face, the system is still capable of easily distinguishing between a real face and a (creased) picture. As soon as the separation between the two distributions is not sufficient to robustly discriminate a true face, the picture on the paper can not be recognized as well. In case of severely "wrinkled" picture paper, an EER very close to 10% is obtained. Therefore, the liveness detection is guaranteed as long as recognition can be still performed from the 2D picture. Due to space limitations the full results for the creased picture could not be included in this version of the manuscript. Nonetheless, the full results and related graphs will be included in the final version of the paper.

[1] It is worth noting that even though different pictures of different users have been used, in this case the subject on the photo is not relevant as only the 3D structure of the acquisition is analyzed, not the iconic information.

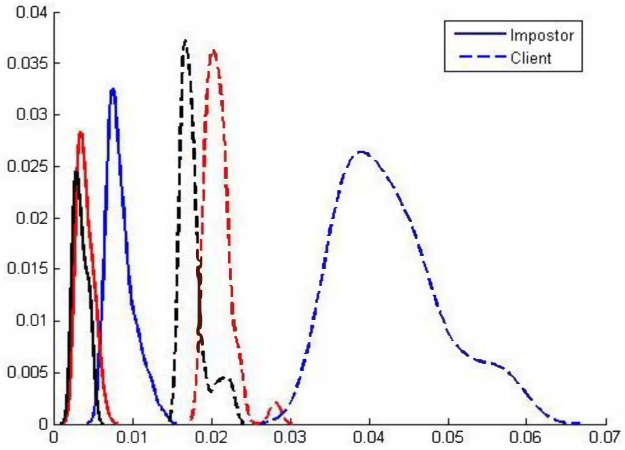

Figure 5: Distribution of the second test results for varying values of the neighborhood radius r equal to 4 (blue line), 12 (red line) and 20 (black line). Both the genuine client (dotted line) and the impostor (solid line) distributions are reported. The value of the mean curvature is reported on the horizontal axis.

6. CONCLUSIONS

This paper has addressed the problem of sensor spoofing attacks for face recognition. The considered scenario may well involve a 3D or a 2D recognition system. In the former case the 3D scan can be directly used by the liveness detection algorithm before any further processing is performed. In the latter case 3D data could be computed from two or more 2D face images captured by the camera. In either case, the proposed method is based on the assumption that a real face has a characteristic 3D structure, with notable variations in the local surface curvature. Even though a printed picture or a computer screen has its own 3D structure, and can be even bend to bear a non-flat 3D shape, it can never exhibit the surface curvature variations which are typical of a human face. Based on this assumption, an effective and automatic method has been proposed to determine if a genuine user is presented to the acquisition sensor. The technique, based on the estimation of the first order statistics of the surface curvature, can be very effective and does not require the active cooperation of a user. As such, it can be effectively used in covert applications or whenever the subject does not cooperate or does not want to be recognized. The experiments, performed on real data and also in adverse conditions, clearly demonstrate the effectiveness and robustness of the proposed methodology.

7. ACKNOWLEDGMENTS

This work has been partly supported by the Italian National PRIN Project 2008XCSNEW "Advanced face analysis methods for security and protection of biometric systems" and by the Australian ARC Dicovery project DP110100827.

8. REFERENCES

[1] K. A. Nixon, V. Aimale, R. K. Rowe, *Spoof Detection Schemes*, in A.K. Jain, P. Flynn and A.A. Ross (eds.), Handbook of Biometrics, Springer, 2007

[2] G. Pan, Z. Wu and L. Sun, *Liveness Detection for Face Recognition*, Recent Advances in Face Recognition, pp. 236, December 2008, I-Tech, Vienna, Austria.

[3]. T. Choudhury, B. Clarkson, T. Jebara, A. Pentland, Multimodal person recognition using unconstrained audio and video, In Proc. of Int. Conf. on Audio- and Video-Based Biometric Person Authentication, pp.176–181, 1999.

[4] K. Kollreider, H. Fronthaler, J. Bigun, *Evaluating liveness by face images and the structure tensor*, IEEE Work. on Automatic Identification Advanced Technologies, pp.75–80, 2005.

[5] J. Li, Y. Wang, T. Tan, A. Jain, *Live Face Detection Based on the Analysis of Fourier Spectra*, Biometric Technology for Human Identification, Proc. SPIE, vol. 5404, pp. 296–303, 2004.

[6] R. W. Frischholz, U. Dieckmann, *BioID: A Multimodal Biometric Identification System*, IEEE Computer, vol. 33, no. 2, pp.64–68, February 2000.

[7] J. Määttä, A. Hadid, M. Pietikäinen, *Face Spoofing Detection From Single Images Using Micro-Texture Analysis*, Proc. International Joint Conference on Biometrics (IJCB 2011), Washington, D.C., USA

[8] M. De Marsico, M. Nappi, D. Riccio and J.L. Dugelay, *Moving Face Spoofing Detection via 3D Projective Invariants,* Proc. of the 5th IAPR International Conference on Biometrics, March 29 - April 1, 2012 New Delhi, India

[9] Vectra Web site: http://www.canfieldsci.com/imaging_systems/research_systems/VECTRA-CR_3D/VECTRA_Specs.html

[10] M. Cadoni, A. Lagorio, E. Grosso and M. Tistarelli, *Exploiting 3d faces in biometric forensic recognition,* Proc. of European Signal Processing Conference, 23-27 August 2010, Aalborg, Denmark

[11] M. Cadoni, M. Bicego and E. Grosso, *3D Face Recognition Using Joint Differential Invariants*, Proc. of The 3rd IAPR/IEEE International Conference on Biometrics, ICB 2009, 2-5 June 2009, Alghero, Italy

[12] A. Savran, N. Alyüz, H. Dibeklioğlu, O. Çeliktutan, B. Gökberk, B. Sankur, and L. Akarun, *Bosphorus database for 3D face analysis*. Biometrics and Identity Management, pages 47–56, 2008.

[13] A. Savran, N. Alyüz, H. Dibeklioğlu, O. Çeliktutan, B. Gökberk, B. Sankur, L. Akarun, Bosphorus Database for 3D Face Analysis, The First COST 2101 Workshop on Biometrics and Identity Management (BIOID 2008), Roskilde University, Denmark, May 2008.