Real-time 3D face recognition using low-cost acquisition devices

Maurício Pamplona Segundo[1,*]

[1]Department of Computer Science – Universidade Federal da Bahia
Av. Adhemar de Barros, S/N – Campus de Ondina – 40170-110 – Salvador/BA, Brazil

mauricio@dcc.ufba.br

Abstract. *Biometric recognition seeks to distinguish humans by physical or behavioral characteristics, but there is not a biometric feature that solves all challenges in this problem. Face biometrics show up as a practical option, since they are well accepted and the cost is fairly low. Among different facial properties that could be used for recognition, the geometry stands out for its invariance to pose and illumination. This doctoral work proposes a complete framework for the problem of recognizing people using 3D images. What differs this work from others is that real-time performance and low-cost acquisition were added as requirements. To accomplish this, we designed a novel face detection method, which was thoroughly evaluated and compared to the state-of-the-art. In addition, we optimized the normalization, description and matching stages of the recognition process. We show the operation of our recognition framework in a recent biometric application, which consists in continuously authenticating the identity of a user to offer a more secure session for high security environments. By doing so, we developed the first continuous authentication system based on facial geometry, which is robust to a wide range of facial variations. Finally, the compatibility between 3D recognition and current forms of identification (e.g. ID photos) was addressed. In order to match 3D images to ID photos, we designed a 3D reconstruction method that uses a single or multiple 2D images of a face to retrieve its geometry. The method can effectively create realistic 3D face models, which are suitable for recognition purposes.*

1. Introduction

Since the beginning of human history people have intuitively used physical and/or behavioral characteristics as a mean of recognition, which is called biometric recognition. In the last century there was an increasing interest in developing biometric-based systems for person identification due to their many applications in security, accessibility and law enforcement, among others. Although there is not a perfect biometric feature, face recognition stands out for its trade-off between accuracy (*i.e.* distinctiveness, permanence) and

[*]This doctoral work was developed by Maurício Pamplona Segundo at Universidade Federal do Paraná (UFPR) with a 1-year sandwich period at the University of South Florida (USF), and was approved by UFPR's Department of Informatics in December 13, 2013. His advisor was Dr. Luciano Silva and his co-advisors were Dr. Olga R. P. Bellon and Dr. Sudeep Sarkar (international advisor). Dr. Silva and Dr. Bellon are with the IMAGO Research Group, Department of Informatics, UFPR, Brazil. Dr. Sarkar is with the Department of Computer Science and Engineering, USF, USA.

practicality (*i.e.* acceptability, collectability, universality) [Jain et al. 2004]. For this reason, according to Hietmeyer [Hietmeyer 2000], the face is the most likely biometric to be used in a global traveler identification system.

Face recognition based on 2D images was the main focus of researches on face biometrics for many years. However, recognizing people in 2D images is a challenging problem due to variations in pose, illumination, facial expressions and age. So, there was an increasing interest in using 3D images to recognize faces in the past decade due to its invariance to changes in pose and lighting [Mian et al. 2007, Pamplona Segundo et al. 2010, Queirolo et al. 2010]. The main challenge for a viable 3D face recognition system use to be the acquisition (*i.e.* high cost, low-speed, limited capture area). Fortunately, this situation is changing with the advances in 3D sensing. A new generation of acquisition devices (*e.g.* Microsoft Kinect, ASUS Xtion PRO, Primesense Carmine) are able to capture up to 100 frames per second (fps), reach objects up to five meters away, and cost about US$ 200 on average. Their accuracy is not as high as in expensive 3D sensors, but the gap is narrowing fast with new devices being released every year.

These devices opened up a vast array of real-time applications that were not possible with earlier sensors. Some examples are user interaction, action recognition and live object reconstruction. This is also true for 3D face analysis, which can now be performed in real-time for video applications, such as expression spotting [Shreve et al. 2014] and continuous authentication [Pamplona Segundo et al. 2013a]. Thus, there is a high demand for real-time 3D analysis techniques to take full advantage of 3D videos.

In this context, the main goal of this doctoral work is the design of a 3D face recognition framework that is fully automatic, runs in real-time and uses low-cost acquisition devices. By doing so, we allow the development of 3D face recognition systems that are economically viable in industry. There is also a secondary goal, which is to provide compatibility between a 3D face recognition system and the current forms of identification (*e.g.* ID cards, passports, driver licenses). This is done by reconstructing the geometry of a face using 2D images, which allows matching 3D images to ID photos (*e.g.* match travelers to their passport photos using facial geometry).

1.1. Contributions

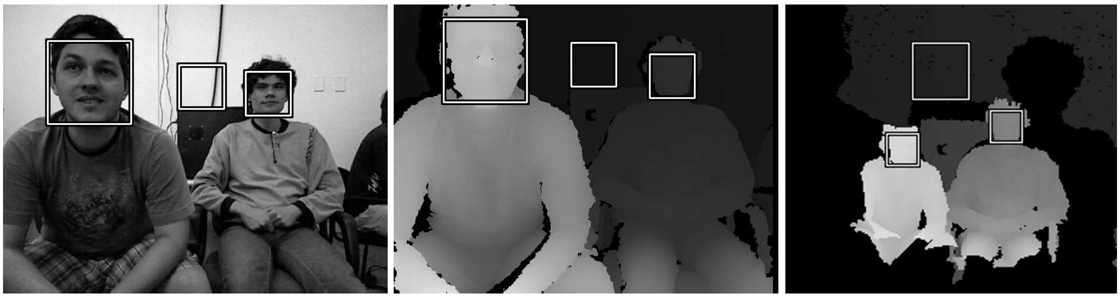
To achieve our goals, a number of contributions have been made:

- we have designed, to the best of our knowledge, the first real-time 3D face recognition framework that handles multiple fps and uses a low-cost device for image acquisition [Pamplona Segundo et al. 2013a];
- as far as we know, we have created the most accurate 3D face detector in the literature, which works successfully for different acquisition scenarios, including substantial variations in resolution, noise, pose, and facial expressions [Pamplona Segundo et al. 2011, Pamplona Segundo et al. 2013b];
- we have designed a scale-invariant image representation, named orthogonal projection images, that allows using the size of an object to optimize the detection process [Pamplona Segundo et al. 2011, Pamplona Segundo et al. 2013b];
- we have successfully applied our 3D face recognition framework to the continuous authentication problem, making it the first continuous authentication system based on 3D face images [Pamplona Segundo et al. 2013a];

- we have shown a more intuitive way of evaluating continuous authentication systems using well-known biometric terms [Pamplona Segundo et al. 2013a];
- we have developed a new 3D face reconstruction method that uses only a single 2D face image with arbitrary pose as input [Pamplona Segundo et al. 2012, Choi et al. 2010];
- we have corroborated neuropsychology works [Hole and Bourne 2010], showing quantitatively that half-frontal face images have more information about the geometry of the face than frontal and profile ones [Pamplona Segundo et al. 2012].

2. Achieved results

We presented a novel 3D face detector that uses orthogonal projection images to eliminate the need for multiple scans in a same image, as illustrated in Figure 1. Figures 1(a) and 1(b) show faces with different sizes in 2D and 3D images due to the perspective distortion, while faces in Figure 1(c) have similar size. We also show how to use our detector to find faces across large pose variation with a frontal face detector by using multiple orthogonal projection images from different viewpoints of the same scene.



(a) (b) (c)

Figure 1. Illustration of the detection process using different input images: (a) color images, (b) depth images and (c) orthogonal projection images.

The proposed detector presents the following advantages over other works in the literature:

- images are no longer scanned in multiple scales, so the computation time is considerably reduced;
- 3D images are more invariant than 2D images and only regions with a pre-specified size are tested, what reduces the number of non-face candidates and makes the detector much more reliable;
- robustness against pose variation can be achieved by increasing the number of orthogonal projection images, and our detector returns not only the face location but the face pose as well.

Our detector was tested in more than 13,000 face images from six different databases, and most of these images presented at least one artifact (*e.g.* facial expressions, noise, pose, occlusion). Nevertheless, and it was able to detect 99% of the faces, with less than 1% of false alarms. Moreover, our detector outperformed the most used detectors in the literature [Colombo et al. 2006, Viola and Jones 2004]. These results represent the state-of-the-art in 3D face detection. They were published in the IEEE International Conference on Systems, Man, and Cybernetics [Pamplona Segundo et al. 2011] and in Pattern

Recognition Letters [Pamplona Segundo et al. 2013b]), which are a Qualis B2 conference a Qualis A1 journal, respectively.

Our 3D face recognition framework was successfully applied to the continuous authentication problem, which is illustrated in Figure 2. The acquisition was performed by a Kinect sensor, and the proposed approach automatically detects, normalizes, describes and matches 3D images in real-time. Our experiments on continuous authentication have used more than 10 hours of genuine and impostor accesses, and it obtained a 0.8% Equal Error Rate (EER), which is the lowest EER obtained so far by a continuous authentication system in the literature [Janakiraman et al. 2005, Sim et al. 2007]. Also, the system was able to detect most of the impostors within a one-second window. At the best of our knowledge, this is the first continuous authentication system that uses 3D face images to check and make sure that the accessing user is the allowed one. The results were published in the IEEE Computer Society Workshop on Biometrics [Pamplona Segundo et al. 2013a], held in conjunction with the IEEE Computer Society Conference Conference on Computer Vision and Pattern Recognition, which is a Qualis A1 conference.

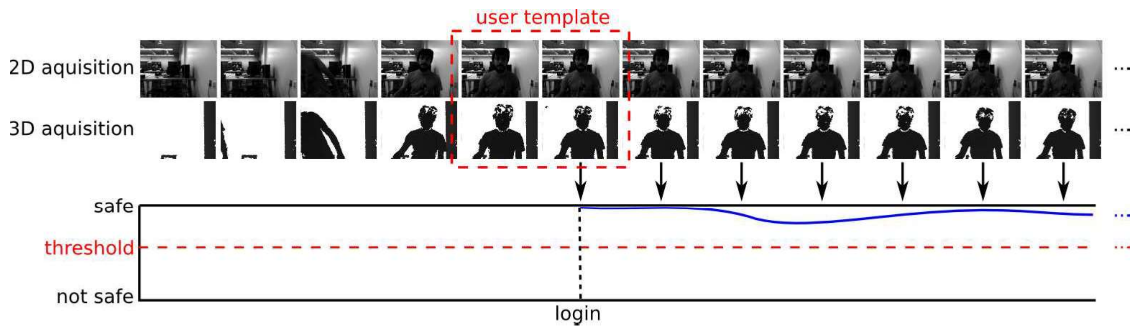


Figure 2. Illustration of the operation of a continuous face authentication system based on face images.

Finally, we presented a new 3D face reconstruction method that uses a single or multiple 2D face images with arbitrary pose as input. It uses a sparse deformable model of 3D faces to guide the simultaneous estimation of pose and deformation parameters. The average reconstruction error ranged from 3.3 to 7mm, depending on the pose and the noise in landmarks location. Our results corroborated previous neuropsychology works [Hole and Bourne 2010] that stated that humans better perceive facial geometry in half-frontal face images, showing quantitatively that half-frontal faces have more 3D information than frontal and profile ones. These results were published in the IAPR International Conference on Pattern Recognition [Choi et al. 2010] and in the IEEE International Conference on Image Processing [Pamplona Segundo et al. 2012], which are both Qualis A1 conferences.

3. Final remarks

The goals of this work were both the design of a fully automatic, real-time 3D face recognition framework using low-cost acquisition devices and to provide compatibility between a 3D face recognition system and the current forms of identification. The Microsoft Kinect is used for 3D acquisition, but other depth sensors with equal or better accuracy than the Kinect can also be used with small or no changes. To the best of our knowledge, this is the first complete framework to use low-cost acquisition devices and handle multiple fps. In addition, the presented framework was successfully applied to the continuous authentication problem, which consists in monitoring the identity of users during the entire access and not only at login.

3.1. Impact

Brazil will hold the World Cup in 2014 and the Olympic Games in 2016, but the security in Brazilian airports, stadiums and other public places is far from ideal. Some private enterprises are interested in security systems based on face recognition to prevent unauthorized access, assault and vandalism. However, only 2D face recognition systems are commercially available, and such systems are subject to many problems (*e.g* pose, illumination, facial expressions). With the rapid advances in 3D imaging, we believe that in the near future 3D sensors will be more accurate and will cover larger areas, and our framework could be applied to video surveillance and access control of highly crowded areas with better performance than the current systems.

3.2. Future directions

We intend to investigate if 2D and/or infrared images could be combined to 3D images to get more accurate recognition results in a multimodal scenario. Moreover, parallelism could be used to handle several faces simultaneously when highly crowded areas are considered.

We would like to apply our 3D detector to other objects whose size presents a small intraclass variation, and try different detection methods on orthogonal projection images since they would also become scale-invariant as well. We could also use orthogonal projection images for other purposes, such as object segmentation and recognition.

Finally, we would like to extend our 3D face reconstruction method to include a dense deformable face model. This way, the results would be more realistic and more accurate for recognition purposes.

Acknowledgements

The authors would like to thank CAPES, CNPq and UFPR for the financial support.

References

Choi, J., Medioni, G., Lin, Y., Silva, L., Bellon, O., **Pamplona Segundo, M.**, and Faltemier, T. C. (2010). 3D face reconstruction using a single or multiple views. In *Proceedings of the 20th International Conference on Pattern Recognition*, pages 3959–3962.

Colombo, A., Cusano, C., and Schettini, R. (2006). 3d face detection using curvature analysis. *Pattern Recognition*, 39(3):444–455.

Hietmeyer, R. (2000). Biometric identification promises fast and secure processings of airline passengers. *The Internationl Civil Aviation Organization Journal*, 55(9):10–11.

Hole, G. and Bourne, V. (2010). *Face processing: Psychological, neuropsychological, and applied perspectives*. Oxford University Press.

Jain, A. K., Ross, A., and Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):4–20.

Janakiraman, R., Kumar, S., Zhang, S., and Sim, T. (2005). Using continuous face verification to improve desktop security. In *Proceedings of the 7th IEEE Workshops on Application of Computer Vision*, volume 1, pages 501–507.

Mian, A., Bennamoun, M., and Owens, R. (2007). An efficient multimodal 2d-3d hybrid approach to automatic face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(11):1927–1943.

Pamplona Segundo, M., Sarkar, S., Goldgof, D., Silva, L., and Bellon, O. (2013a). Continuous 3d face authentication using rgb-d cameras. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 64–69.

Pamplona Segundo, M., Silva, L., and Bellon, O. R. P. (2011). Real-time scale-invariant face detection on range images. In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, pages 914–919.

Pamplona Segundo, M., Silva, L., and Bellon, O. R. P. (2012). Improving 3d face reconstruction from a single image using half-frontal face poses. In *Proceedings of the 19th IEEE International Conference on Image Processing*, pages 1797–1800.

Pamplona Segundo, M., Silva, L., Bellon, O. R. P., and Queirolo, C. C. (2010). Automatic face segmentation and facial landmark detection in range images. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 40(5):1319–1330.

Pamplona Segundo, M., Silva, L., Bellon, O. R. P., and Sarkar, S. (2013b). Orthogonal projection images for 3d face detection. *Pattern Recognition Letters*.

Queirolo, C. C., Silva, L., Bellon, O. R. P., and **Pamplona Segundo, M.** (2010). 3d face recognition using simulated annealing and the surface interpenetration measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(2):206–219.

Shreve, M., **Pamplona Segundo, M.**, Luguev, T., Goldgof, D., and Sarkar, S. (2014). High-resolution 3d surface strain magnitude using 2d camera and low-resolution depth sensor. *Pattern Recognition Letters*.

Sim, T., Zhang, S., Janakiraman, R., and Kumar, S. (2007). Continuous verification using multimodal biometrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(4):687–700.

Viola, P. and Jones, M. J. (2004). Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154.