3D Biometrics

Biometrics have been used to identify humans since the 19th century. Over time, these biometrics became 3D. The main reason for this was the growing need for more features in the images to create more reliable identification models. Three main categories of 3D biometrics were identified. These were face, hand and gait. The corresponding percentages for these categories were 74.07%, 20.37% and 5.56%, respectively. The face is further categorized into facial, ear, iris and skull, while the hand is divided into fingerprint, finger vein and palm. In each category, facial and fingerprint were predominant, and their respective percentages were 80% and 54.55%. The use of the 3D reconstruction algorithms was also determined. These were stereo vision, structure-from-silhouette (SfS), structure-from-motion (SfM), structured light, time-of-flight (ToF), photometric stereo and tomography. Stereo vision and SfS were the most commonly used algorithms with a combined percentage of 51%. The state of the art for each category and the available datasets are also presented. Finally, multimodal biometrics, generalization of 3D reconstruction algorithms and anti-spoofing metrics are the three areas that should attract scientific interest for further research.

1. Introduction

Biometrics are unique body characteristics used to identify people. They were first used at the end of the 19th century with the well-known and globally used fingerprints. According to Jain et al. [1], the most commonly used biometric methods are DNA, ear, facial, hand and finger veins, fingerprint, gait, hand geometry, iris, palmprint, retina, signature and voice. To perform the identification, a device is used to capture the biometric data. Most often, this device captures images, and the quality of the captured images therefore affects the performance of the model. The first identifications were completed manually by experts, but in some cases, the results were controversial due to the human factor. Later, the use of technology for identification has evolved in the form of image processing methods and matching techniques, creating tremendous identification models that take advantage of biometrics. Over the years, these technologies achieved great performance and thus became more popular. Moreover, biometrics for identifying people is nowadays not only used in forensics but also to gain access to certain places or to log in to some smart devices.

As technology has advanced, a biometric security problem has emerged. The technology became very familiar and very vulnerable to malicious acts. This has had a major impact on various security protocols as biometrics have become a part of daily lives. To counter this, a few approaches have been developed in this area. One of them is to increase the robustness of the selected biometric category. Bear in mind that the traditional methods are referred to 2D, and some scientific approaches have led to the development of 3D biometrics. Of course, this was not just about the fancy addition of the third dimension but mainly about increasing the extracted features and creating more efficient systems. These additional features are the key to the desired performance improvement.

Computer vision has always been linked to biometrics as it provides the necessary tools for identification through 3D image analysis. In addition, the technological advancement of computer vision using state-of-the-art Artificial Intelligence methods to achieve the above benefits has led to the need to apply it to identification systems. As the demand for robust models has increased, the transition from 2D to 3D biometric methods has been a one-way street.

The core element for 3D reconstruction is depth information. Various algorithms have been developed to extract the relevant information. The first work published for 3D biometrics in general was from David Zhang and Guangming Lu in 2013 [2]. In their book, they described the image acquisition methods and categorized them into two major categories, the single and multi-view approaches. Another approach is to categorize them into active and passive methods. In active methods, the light source is directed to the desired surface as an essential step for 3D reconstruction and is characterized by low computational cost. Passive methods, on the other hand, are usually very computationally intensive and use ambient light conditions [3]. Furthermore, active methods can be categorized into structured light, time of flight (ToF), photometric stereo and tomography. Passive methods include stereo vision, structure-from-silhouette (SfS), structure-

from-texture (SfT) and structure-from-motion (SfM). The taxonomy shows that the two main approaches, active and passive, have four methods at once. With further approaches, it is possible to create a third category. This category could be semi-active, semi-passive, or even a combination of passive and active methods. Such multimodal approaches should lead to new 3D reconstruction algorithms.

Structured light produces a beam from a light source onto the surface of the object. The wavelength of the light can be in the visible range, the infrared (IR) or even the near infrared (NIR). The calculations from the reflection of the beam provide depth information. Secondly, the ToF method takes into account the reflection time between the surface and a reference point, while the photometric stereo method uses different lights to create different shades of the object and the model is created by combining these lights. Finally, tomography can be either optical coherence tomography or computed tomography (CT). In both cases, the 3D models are created by multiple scans.

In addition, stereo vision creates depth information by comparing image information of the same spot from two different viewpoints. SfS uses images taken from different angles to form the silhouette of the object. In addition, the SfT is applied when the surface has a homogeneous texture and then uses the different orientation of the images to create the depth details. Finally, the SfM uses a video as input, which is usually consisting of frames that capture an object from different angles.

An important factor for successful reconstruction is the sensor used. The RGB-D camera is commonly used in various 3D reconstruction applications. This particular type can provide color and depth information by combining the three primary colors (red–green–blue) and calculating the relative distance to the sensor accordingly. According to an analysis of RGB-D camera technologies for face recognition, Urlich et al. [4] found that stereoscopy (active and passive), followed by structured light, produced the best results. The importance of these cameras was also emphasised by Zollhöfer et al. [5]. In the research, they presented the different approaches and the great performance of these sensors for all aspects of 3D reconstructions, including biometric elements such as the face, etc.

Furthermore, the above eight different methods are used throughout the literature without any correlation between the categories of 3D biometrics, as each category has been studied separately so far. Although some research refer to a group of categories, such as facial and ears, the vast majority refer explicitly to one category. This creates additional barriers to the extraction of cross-biometric information, such as common methods between categories or even similarities in the state of the art. Furthermore, there is no literature research in the field that examines all categories of 3D biometrics at once.

2. Face

The first and most popular category was the face recognition, with the facial and ear being two popular subsections. In some cases, the ear is part of a multimodal approach or a standalone biometric. In 2011, Yuan et al. [6] studied ear recognition. In the research, the researchers had also included the recognition process of 3D images for the first time. The ear has a unique shape and also shows minor deformations over the years. According to the research, the main reconstruction methods are SfS, SfM and Stereo Vision, with the last method being the most effective. Yaun et al. [6] also conclude that the accuracy and robustness of the system are greatly improved when ear features are used in combination with face features.

The first research for facial biometric was by Islam et al. in 2012 [7], which presented a 3D facial pipeline in four stages. These stages were 3D data acquisition, detection, representation and recognition. The researchers provide some details about the data acquisition techniques. For face recognition, there are two main categories: the use of 2D or 3D images. The state of the art for using 2D images was the Support Vector Machine (SVM) with 97.1% on 313 faces [8] and for 3D images was the Point Distribution Model (PDM), which achieved 99.6% on 827 images [9]. For the representation of the face, the balloon image [10] and iso-countours [11] were the most advanced models. When used to reconstruct a face, they achieved 99.6% and 91.4% accuracy in face recognition, respectively. The final step was the recognition. Since the face often changes during emotional expressions, the researchers presented two main categories: rigid and non-rigid, depending on whether the model is considered rigid or not. Although the percentages for the rigid approach were high (the Iterative Closest Point (ICP) algorithm [12] reached 98.31%), some samples were rejected by the algorithm due to different expressions. On the other hand, rigid approaches had similar performance but increased computational cost. For ear reconstruction, the researchers proposed three different approaches. The first was to use landmarks and a 3D mask, but this approach depends on manual intervention. For the second approach, the use of 3D template matching was proposed, with an accuracy of 91.5%. This approach had better performance than the previous one but a higher error rate. The last

and most efficient method was the Ear Shape Model, proposed by Chen and Bhanu [13] in 2005, which achieved an accuracy of 92.5% with an average processing time of 6.5 s.

The use of ToF methods in 3D face recognition was research by Zhang and Lu in 2013 [2], presenting two main approaches for ToF applications. The first is image capture with incoherent light, and the second is based on optical shutter technology. Both use the reflection of NIR light. They also pointed out the disadvantages of these devices, namely the high error rates due to the generation of low-resolution images by the different devices. Of course, they also believe that hardware will be able to support a higher resolution biometric system in the near future. The main advantage of the Tof is that it can provide real-time results, which are very important for biometrics.

In 2014, Subban and Mankame [14] wrote a research focusing on 3D face recognition methods and proposing two different approaches. The first extracts features from the different facial attributes (nose, lips, eyes, etc.), and the second assumes the face as a whole entity. Furthermore, the researchers presented the methods that had the best performance based on recognition rate (RR). In particular, a combination of geometric recognition and local hybrid matching [15] achieved 98.4%, which was followed by the method of local shape descriptor with almost the same performance (98.35%) [16]. The remaining methods 3D morphing [17] and multiple nose region [18] were equally efficient with 97% and 96.6%, respectively. In the same year, Alyuz et al. [19] described the phenomenon of difficulty in identifying 3D faces in the presence of occlusions. These occlusions can be accessories such as hats or sunglasses or even a finger in front of the face. However, Alyuz proposed a method consisting of removing occlusions and then restoring the missing part of the face, achieving a high identification accuracy (93.18%).

In the following year, Balaban et al. [20] researched deep learning approaches for facial recognition. The researchers emphasized that as deep learning models evolve, better datasets are needed. In fact, the state-of-the-art Google FaceNet CNN model [21] had an accuracy of 99.63% in the Labeled Faces in the Wild (LFW) dataset [22]. This very high accuracy somehow shows that the scientific community should create datasets with a lot of additional images. Balaban also believes that such a dataset will be revolutionary and compares it to the transition from Caltech 101 to Imagenet datasets. The next year, in 2016, Liu et al. [23] presented the weaknesses of facial recognition systems when the input images are multimodal. These can be IR, 3D, low-resolution or thermal images, which are also known as heterogeneous data. Liu also suggested that future recognition algorithms should be robust to multimodal scenarios in order to be successfully used in live recognition scenarios. The researchers also highlight the fact that humans can easily perform face recognition with multimodal images. In order to mimic this behaviour, one approach is to make the different models exposed to long-term learning procedures.

Furthermore, Bagga et al. introduced a research of anti-spoofing methods in face recognition, including 3D approaches [24]. In face spoofing, the "attacker" creates fake evidence to trick a biometric system. In 3D, these proofs are fake masks created from real faces. Four techniques are proposed: motion, texture, vital sign and optical flow-based analysis. Motion-based analysis involves analyzing the movement of different parts of the face, such as the chin or forehead, so that any attempt at forgery can be detected. The best approach is a calculation estimate based on illumination invariance by Klaus Kollreider et al. [25]. The second technique extracts the texture and frequency information using Local Binary Pattern (LBP), which is followed by histogram generation. Finally, an SVM classifier classifies whether the face is real or fake. The third method is vital sign recognition analysis. This can be performed either by user interaction such as following some simple commands (e.g., head movements, etc.) or in a passive way, i.e., by detecting mouth and/or eye movements. Lastly, the optical flow-based analysis, proposed by Wei et al. [26], is primarily used to distinguish a fake 2D image from a 3D face.

In 2017, Mahmood et al. [27] presented the state of the art in face recognition, using 3D images as input. According to their research , five recognition algorithms stood out for their high performance. These are ICP [28], Adaptively Selected Model Based (ASMB) [29], Iterative Closest Normal Point (ICNP) [30], Portrait Image Pairs (PIP) [31] and Perceived Facial Images (PFI) [32] with 97%, 95.04%, 99.6%, 98.75% and 98%, respectively. The researchers concluded that despite the high accuracy achieved by applying state-of-the-art algorithms, the models are not yet reliable enough to process data in real-time. Furthermore, they underlined that the performance varies greatly on different datasets, and further research should be conducted in this direction.

Later, in 2019, Albakri and Alghowinem [33] researched various anti-spoofing methods. According to their research, pulse estimation, image quality and texture analysis are the most efficient methods. Meanwhile, they proposed a proof-of-concept research in which the fake input is detected by computing the depth data. Using four different types of attacks (flat image, 3D mask, on-screen image and video) on three different devices, iFace 800, iPhone X and Kinect, they managed to detect the fake attack with very high accuracy (92%, 100% and 100%, respectively). The last two reviews

were also related to anti-spoofing approaches. Wu et al. [34] focused on applications in China and also presented three trends for protecting against facial forgery. The first was a full 3D reconstruction of the face, whose main drawback is low performance. The next was a multimodal fusion approach consisting of a combination of visible and IR light generated by binocular cameras. Finally, the generative model with a proposed new noise model [35] was the state of the art. Finally, Liu et al. [36] reviewed the results of the Chalearn Face Anti-Spoofing Attack Detection Challenge at CVPR2020. Eleven teams participated using the metric ACER standardized on ISO and achieved great performance with the best percentages of 36.62% and 2.71% for single and multimodal approaches, respectively.

3. Fingerprint

The 3D fingerprint was first reviewed in 2013 by Zhang and Lu [2]. First, they present the general flowchart of 3D fingerprint reconstruction from images acquired by touchless devices. The first step is to calibrate the camera, which is a very common procedure for reconstruction applications. The next step is to determine the three different correspondences. These are based on the SIFT feature, the ridge map and the minutiae. After these correspondences, the coordinates are created based on the generated matching points. The final step is to produce an estimation of the shape of the human finger. The researchers concludes that the most difficult process in the whole reconstruction is the establishment of the different correspondences.

In 2014, Labati et al. [37] presented a research on contactless fingerprinting as a biometric. The research described the advantages of capturing the corresponding images without contact between the finger and the capture device. The researchers presented three main approaches to reconstruct the fingerprint: a multiview technique, followed by structured light and stereophotometry. In the multiview technique, several cameras are used to obtain the images. More specifically, two, three and five cameras have been proposed. It should be noted here that although a large number of cameras means higher accuracy, it also increases the computational cost. Therefore, the optimal way to design a system is to find a compromise between accuracy and computational cost. Using a structured light provides an estimate of the ridges and valleys of the fingerprint in addition to texture estimation. The light is projected onto the finger several times in the form of a wave to capture the corresponding image. The last method was photometric stereo imaging. Using a single camera and several LED lights as illuminators, the fingerprint was reconstructed using the photometric stereo technique. The above approaches were promising and opened up new scientific fields. The researchers also emphasized that there is no work yet on the 3D reconstruction of fingerprints that can be used as a biometric.

Later, Jung et al. [38] suggested ultrasonic transducers as a possible way to find depth information. More specifically, the researchers reviewed the various ultrasonic sensors used in microelectromechanical systems (MEMS) and their applications. One of them was to acquire appropriate depth information of a fingerprint and its use as a biometric. With the improvement of MEMS technology, ultrasonic sensors are improving in terms of accuracy and robustness. Moreover, this type of sensor outperforms the more traditional ones, such as optical or capacitive, achieving better performance. For this reason, Jung underlined that the above sensors have great scientific potential for 3D biometrics. The use of ultrasound devices was further highlighted in 2019 by Lula et al. [39]. The research described the use of ultrasound devices as a suitable method for capturing 3D fingerprints that can be used in biometric systems. The main problem with this approach was the long acquisition time of the images. This was overcome by using additional probes, which are usually arranged cylindrically. It should also be noted that the frequency bandwidth varies between 7.5 and 50 MHz and depends on the probe model and circumference chosen. A high ultrasound frequency offers high skin penetration but low resolution, while a lower frequency has the opposite effect.

Finally, Yu et al. [40] presented a research on 3D fingerprints as a biometric using optical coherence tomography (OCT) as an acquisition device. To calculate the depth information, the light beam was directed at the subject and a mirror. The light penetrated the finger and the system correlated the reflection of the finger's and the mirror's beam. Through this process, the system calculated the depth information and then reconstructed the 3D representation. The light penetration also provided the inner fingerprint, which is unaffected by environmental influences. Sweat pores and glands are also visible through this approach. These additional elements provide more features, and the biometric system has become more robust as a result. Despite the fact that OCT has the above advantages, there are also some disadvantages. These include latency, cost, mounting limitations and low resolution

Related work shows that only two 3D biometric categories have been reviewed: the face and the fingerprint. This is probably because these biometrics are very common as 2D biometrics, especially facial. This led to the fact that the other biometric categories such as iris or finger vein were not reviewed. In order to present their state of the art, additional research about them should be conducted.

References

1. Jain, A.K.; Ross, A.; Prabhakar, S. An introduction to biometric recognition. IEEE Trans. Circuits Syst. Video Technol. 2004, 14, 4–20.

2. Zhang, D.; Lu, G. 3D Biometrics; Springer: New York, NY, USA, 2013.

3. Moons, T.; Van Gool, L.; Vergauwen, M. 3D reconstruction from multiple images part 1: Principles. Found. Trends® Comput. Graph. Vis. 2010, 4, 287–404.

4. Ulrich, L.; Vezzetti, E.; Moos, S.; Marcolin, F. Analysis of RGB-D camera technologies for supporting different facial usage scenarios. Multimed. Tools Appl. 2020, 79, 29375–29398.

5. Zollhöfer, M.; Stotko, P.; Görlitz, A.; Theobalt, C.; Nießner, M.; Klein, R.; Kolb, A. State of the art on 3D reconstruction with RGB-D cameras. Comput. Graph. Forum 2018, 37, 625–652.

6. Yuan, L.; Mu, Z.C.; Yang, F. A review of recent advances in ear recognition. In Chinese Conference on Biometric Recognition; Springer: Berlin/Heidelberg, Germany, 2011; pp. 252–259.

7. Islam, S.M.; Bennamoun, M.; Owens, R.A.; Davies, R. A review of recent advances in 3D ear-and expression-invariant face biometrics. ACM Comput. Surv. (CSUR) 2012, 44, 1–34.

8. Osuna, E.; Freund, R.; Girosit, F. Training support vector machines: An application to face detection. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Juan, PR, USA, 17–19 June 1997; pp. 130–136.

9. Nair, P.; Cavallaro, A. 3-D face detection, landmark localization, and registration using a point distribution model. IEEE Trans. Multimed. 2009, 11, 611–623.

10. Pears, N. RBF shape histograms and their application to 3D face processing. In Proceedings of the 2008 8th IEEE International Conference on Automatic Face & Gesture Recognition, Amsterdam, The Netherlands, 17–19 September 2008; pp. 1–8.

11. Mpiperis, I.; Malasiotis, S.; Strintzis, M.G. 3D face recognition by point signatures and iso-contours. In Proceedings of the Fourth IASTED International Conference on Signal Processing, Pattern Recognition, and Applications, Anaheim, CA, USA, 14–16 February 2007.

12. Chen, Y.; Medioni, G. Object modelling by registration of multiple range images. Image Vis. Comput. 1992, 10, 145–155.

13. Chen, H.; Bhanu, B. Shape model-based 3D ear detection from side face range images. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)-Workshops, San Diego, CA, USA, 21–23 September 2005; p. 122.

14. Subban, R.; Mankame, D.P. Human face recognition biometric techniques: Analysis and review. In Recent Advances in Intelligent Informatics; Springer: Cham, Switzerland, 2014; pp. 455–463.

15. Huang, D.; Ardabilian, M.; Wang, Y.; Chen, L. 3-D face recognition using eLBP-based facial description and local feature hybrid matching. IEEE Trans. Inf. Forensics Secur. 2012, 7, 1551–1565.

16. Inan, T.; Halici, U. 3-D face recognition with local shape descriptors. IEEE Trans. Inf. Forensics Secur. 2012, 7, 577–587.

17. Zhang, C.; Cohen, F.S. 3-D face structure extraction and recognition from images using 3-D morphing and distance mapping. IEEE Trans. Image Process. 2002, 11, 1249–1259.

18. Chang, K.; Bowyer, K.W.; Sarkar, S.; Victor, B. Comparison and combination of ear and face images in appearance-based biometrics. IEEE Trans. Pattern Anal. Mach. Intell. 2003, 25, 1160–1165.

19. Alyuz, N.; Gokberk, B.; Akarun, L. Robust 3D Face Identification in the Presence of Occlusions. In Face Recognition in Adverse Conditions; IGI Global: Hershey, PA, USA, 2014; pp. 124–146.

20. Balaban, S. Deep learning and face recognition: The state of the art. In Proceedings of the Biometric and Surveillance Technology for Human and Activity Identification XII, Baltimore, MD, USA, 20–24 April 2015.

21. Schroff, F.; Kalenichenko, D.; Philbin, J. Facenet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 815–823.

22. Huang, G.B.; Mattar, M.; Berg, T.; Learned-Miller, E. Labeled faces in the wild: A database forstudying face recognition in unconstrained environments. In Proceedings of the Workshop on Faces in 'Real-Life' Images: Detection, Alignment, and Recognition, Marseille, France, 12–18 October 2008.

23. Liu, X.; Sun, X.; He, R.; Tan, T. Recent advances on cross-domain face recognition. In Chinese Conference on Biometric Recognition; Springer: Cham, Switzerland, 2016; pp. 147–157.

24. Bagga, M.; Singh, B. Spoofing detection in face recognition: A review. In Proceedings of the 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 16–18 March 2016; pp. 2037–2042.

25. Kollreider, K.; Fronthaler, H.; Faraj, M.I.; Bigun, J. Real-time face detection and motion analysis with application in "liveness" assessment. IEEE Trans. Inf. Forensics Secur. 2007, 2, 548–558.

26. Bao, W.; Li, H.; Li, N.; Jiang, W. A liveness detection method for face recognition based on optical flow field. In Proceedings of the 2009 International Conference on Image Analysis and Signal Processing, Linhai, China, 11–12 April 2009; pp. 233–236.

27. Mahmood, Z.; Muhammad, N.; Bibi, N.; Ali, T. A review on state-of-the-art face recognition approaches. Fractals 2017, 25, 1750025.

28. Jin, Y.; Wang, Y.; Ruan, Q.; Wang, X. A new scheme for 3D face recognition based on 2D Gabor Wavelet Transform plus LBP. In Proceedings of the 2011 6th International Conference on Computer Science & Education (ICCSE), Singapore, 3–5 August 2011; pp. 860–865.

29. Alyuz, N.; Gokberk, B.; Akarun, L. 3-D face recognition under occlusion using masked projection. IEEE Trans. Inf. Forensics Secur. 2013, 8, 789–802.

30. Mohammadzade, H.; Hatzinakos, D. Iterative closest normal point for 3D face recognition. IEEE Trans. Pattern Anal. Mach. Intell. 2012, 35, 381–397.

31. Jahanbin, S.; Choi, H.; Bovik, A.C. Passive multimodal 2-D+ 3-D face recognition using Gabor features and landmark distances. IEEE Trans. Inf. Forensics Secur. 2011, 6, 1287–1304.

32. Huang, D.; Soltana, W.B.; Ardabilian, M.; Wang, Y.; Chen, L. Textured 3D face recognition using biological vision-based facial representation and optimized weighted sum fusion. In Proceedings of the CVPR Workshops, Colorado Springs, CO, USA, 20–25 June 2011; pp. 1–8.

33. Albakri, G.; Alghowinem, S. The effectiveness of depth data in liveness face authentication using 3D sensor cameras. Sensors 2019, 19, 1928.

34. Wu, B.; Pan, M.; Zhang, Y. A review of face anti-spoofing and its applications in china. In International Conference on Harmony Search Algorithm; Springer: Cham, Switzerland, 2019; pp. 35–43.

35. Jourabloo, A.; Liu, Y.; Liu, X. Face de-spoofing: Anti-spoofing via noise modeling. In European Conference on Computer Vision (ECCV); Springer: Cham, Switzerland, 2018; pp. 290–306.

36. Liu, A.; Li, X.; Wan, J.; Liang, Y.; Escalera, S.; Escalante, H.J.; Madadi, M.; Jin, Y.; Wu, Z.; Yu, X.; et al. Cross-ethnicity face anti-spoofing recognition challenge: A review. IET Biom. 2021, 10, 24–43.

37. Labati, R.D.; Genovese, A.; Piuri, V.; Scotti, F. Touchless fingerprint biometrics: A survey on 2D and 3D technologies. J. Internet Technol. 2014, 15, 325–332.

38. Jung, J.; Lee, W.; Kang, W.; Shin, E.; Ryu, J.; Choi, H. Review of piezoelectric micromachined ultrasonic transducers and their applications. J. Micromech. Microeng. 2017, 27, 113001.

39. Iula, A. Ultrasound systems for biometric recognition. Sensors 2019, 19, 2317.

40. Yu, Y.; Wang, H.; Sun, H.; Zhang, Y.; Chen, P.; Liang, R. Optical Coherence Tomography in Fingertip Biometrics. Opt. Lasers Eng. 2022, 151, 106868.