Fast and Robust Biometric Authentication Scheme Using Human Ear

Mozammel Chowdhury[1], Rafiqul Islam[1(✉)], and Junbin Gao[2]

[1] School of Computing and Mathematics, Charles Sturt University, Sydney, Australia
{mochowdhury,mislam}@csu.edu.au
[2] Discipline of Business Analytics, The University of Sydney Business School, Sydney, Australia
junbin.gao@sydney.edu.au

Abstract. Biometric authentication using human ear is a recent trend in security applications including access control, user recognition, surveillance, forensic, and border security systems. This paper aims to propose a fast and robust authentication scheme using ear biometric. In this work, a fast technique based on the AdaBoost algorithm is used to detect the ear of the user from profile images. An efficient stereo matching algorithm is used to match the user's ear data (probe) to the previously enrolled (stored) ear data in a gallery database for verification and recognition. Correspondences are established between extracted features of the probe and gallery image sequences. The performance of the recognition approach is evaluated on different standard ear datasets and compared with other techniques. Experimental results suggest the superiority of the proposed approach over other popular techniques reported in this work.

Keywords: Biometric authentication · Access control · Ear recognition

1 Introduction

Biometric identification and authentication has been gaining popularity for providing safety and security in many applications such as, access control, surveillance system, visa processing, national IDs, border checking, law enforcement applications and so on. Biometric system is a technique that relies on the unique biometric characteristics of individuals to verify or recognize the user for secure access to a system [1]. A biometric system may operate in one or both two modes: authentication and identification. In authentication mode, one-to-one matching is performed to compare a user's biometric data to a specific pattern of the claimed identity enrolled in the system earlier. In identification or recognition process, one-to-many matching is done to identify a user's biometric by comparing it against every identity patterns stored in a large database. The traditional methods for user authentication or identification have deficiencies that restrict their applicability in security systems. The properties used in the traditional authentication methods can be forgotten, disclosed, lost or stolen. Biometric characteristics on the other hand, are unique and not duplicable or transferable. Therefore, biometric trait based security systems have been proven superior to traditional ID based systems [2].

X. Lin et al. (Eds.): SecureComm 2017 Workshops, LNICST 239, pp. 133–144, 2018.
https://doi.org/10.1007/978-3-319-78816-6_11

Most of the biometric systems use traits such as, fingerprint, face, facial components, palm print, hand geometry, iris, retina, gait and voice [3]. In recent years, the use of human ear as a biometric trait is a promising trend in the research community. The ear is quite attractive biometric candidate because, the shape of the ear is unique to individuals and generally unaffected by changing facial expressions, anxiety, use of cosmetics or eye glasses and aging [4]. Moreover, several ear features such as smaller in size, co-location with face, and relatively less change in shape due to aging has made it very popular among biometric communities. An ear also has reduced spatial resolution and uniform distribution of colour. However, due to its complex geometrical shape and often being obscured by hair, ear-ring, head-cover and the similar, developing fast, accurate and robust ear based biometric systems is still very challenging [5, 6].

A typical automatic ear-based biometric system consists of the following steps: detection (or segmentation) of the ear, normalization and enhancement, feature extraction and matching (recognition or verification). Ear detection refers to the localization of the ear shape in a facial profile image. After detection or segmentation, the ear region can be normalized (in orientation or in size) and enhanced to make it simple for further operations such as feature extraction and matching processes. Since the other processing steps like feature extraction, recognition or verification depend on accurate detection of the ear, this stage is crucial in biometric system.

This paper proposes a robust and efficient ear based biometric system using AdaBoost based ear detection, local features extraction and stereo matching based recognition algorithms. Correspondence matching is crucial for meaningful comparisons of two images. The importance of good correspondences is even greater in the case of ear recognition. Standard systems often align the ears or a few other features, using translation, or similarity transformations. However, these can still result in significant misalignments in the ear region. To handle this situation, we use stereo matching. This allows for arbitrary, one-to-one continuous transformations between images, along with possible occlusions, while maintaining an epipolar constraint. In matching correspondences between scan lines in two images, a stereo matching cost is optimized, which reflects how well the two images match. Consequently, we can use the stereo matching cost as a measure of similarity between two ear images (probe and gallery image). Although, stereo matching algorithms have been used in face recognition earlier [20, 21], we are the first to use stereo matching approach for ear recognition. The proposed system does not require training or extraction of the ear contour and hence reduces the computational cost compared to other existing methods. Hence, the low computation time renders its suitability to employ it in real time applications. The obtained ear recognition results can be combined with other biometric modalities such as facial features to develop a more robust and accurate recognition system.

The rest of the paper is organized as follows. In Sect. 2, we have discussed the related works on ear based biometric recognition. The proposed scheme is presented in Sect. 3. Experimental results are reported and discussed in Sect. 4. Finally, Sect. 5 concludes the paper.

2 Related Works

Ear based biometric authentication system is considered as one of the most promising solutions for secure systems. Due to many practical applications, there is currently an increasing demand of biometric technology in the industry. According to the surveys [2–4], most of the proposed ear based recognition approaches use either PCA (Principal Component Analysis) or the ICP algorithm for matching [6–9].

Yaqubi et al. propose a system employing edge features taken over multiple positions and orientations [10]. The extracted features are classified using an SVN and a kNN with recognition accuracy of 96.5%.

Islam et al. [11] find local surface patches (LSP) to select features for their system. PCA is then used to find the most descriptive features in the LSP. The feature extractor repeats selecting LSP until the desired number of features is found. The algorithm is evaluated using UND ear database. They obtain a recognition rate of 93.5%. However, the approach has not been tested with pose variation and different scaling.

Wang et al. [12] employ different feature vectors in their method using seven moment invariants. The feature vectors are used as the input for a back propagation neural network which is trained to classify the moment invariant feature sets.

Gutierrez et al. [13] divide the detected ear regions into three equally sized segments. The upper segment shows the helix, the middle one shows the concha and the lower part shows the lobule. Each of these sub images is decomposed by wavelet transform and then fed into a modular neural network (MNN).

Alaraj et al. [14] use PCA in their work for feature representation. The approach use a multilayer feed forward neural network for classification of the PCA based feature components. They have reported a rank-1 performance of 96%.

3 Proposed Approach

The proposed biometric authentication scheme based on human ear is consisted of the following stages: (i) Acquisition of profile face images, (ii) Ear data extraction and normalization, (iii) Refinement, (iv) Features extraction, (v) Feature matching, and (vi) Recognition/Authentication. The architecture of the proposed ear biometric system is depicted in Fig. 1.

3.1 Ear Detection and Normalization

Ear detection consists of extracting the position of the ear in a facial profile image. Different automatic ear detection methods have been published in recent years. In this work, the ear region is detected on profile face images using the AdaBoost based detector [18]. The motivation behind the selection of this detector is that it possesses high accuracy and speed. After detecting the ear region, the corresponding ear data is then extracted. To ensure the whole ear is extracted, we expand the detected ear regions by an additional 20 pixels around each direction. The extracted ear data varies in dimensions. Therefore, we normalize the extracted ear shape with uniform dimension of 160 by 140.

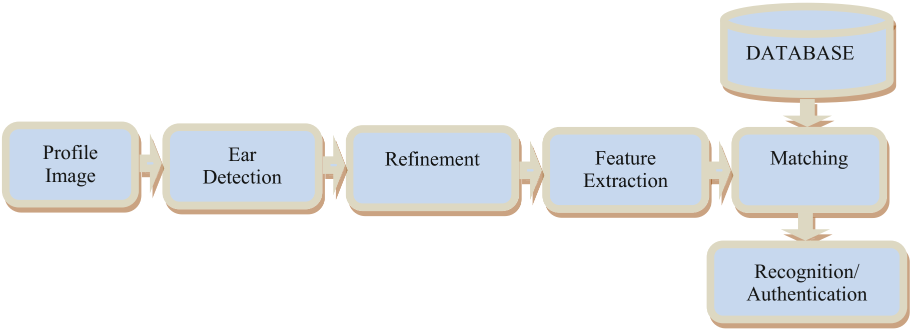


Fig. 1. Architecture of the proposed authentication system using ear biometric.

3.2 Pre-processing of Ear Data

Once the human ear is detected, we employ a fuzzy filter [19] to remove all the spikes and holes from the extracted ear region. We choose this filter because it has the advantage of both median and average filtering and possesses high accuracy and speed. This filter employs fuzzy rules for deciding the gray level of the pixels within a window in the image.

3.3 Features Extraction

One of the crucial tasks in biometric ear recognition is the features extraction. Different types of features commonly used in ear recognition include: intensity and shape features, Fourier descriptors, wavelet-based (i.e. Gabor) features or SIFT points [25]. The extracted ear features are used for matching with the one stored in the gallery database. In this work, we use local edge features since they are invariant to pose variation, occlusion and illumination changes.

Edge or gradient histogram corresponds to the spatial distribution of the edge features in the image. The gradient of an image $f(x, y)$ can be expressed by,

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} G_x \\ G_y \end{bmatrix} \quad (1)$$

where, $G_x = \frac{\partial f}{\partial x}$ is the gradient in x direction, and
$G_y = \frac{\partial f}{\partial y}$ is the gradient in y direction.

The gradient direction can be calculated by the formula:

$$\theta = \tan^{-1}\left[\frac{G_y}{G_x}\right] \quad (2)$$

We use Canny edge detector to extract the edge features from the ear data. The gradient values are normalized to [0, 255].

3.4 Feature Matching and Recognition

The extracted ear features of a user (probe data) are matched with the specific ear data (gallery data) stored in the gallery database built off-line. Matching can be performed based on the error of registering between the two data sets, more specifically, two clouds of points. In this work, we use a stereo algorithm [22] to compute the degree of similarity, which is quite fast and efficient. The stereo algorithm compares two ear data (probe and gallery) and computes the degree of similarity between the probe image and the gallery image and identifies the user's ear that produces the best matching score. Prior to stereo matching, we need to estimate the epipolar geometry. The probe and gallery images are rectified and the similarity score is computed by computing the stereo matching cost of every row of the rectified images.

3.5 Epipolar Geometry and Rectification

The rectification allows the use of epipolar geometry environment where the epipolar lines are horizontal i.e., parallel to the lines of the image sequences [23]. In epipolar geometry, any point lying on an epipolar line in the reference image (i.e., probe image) corresponds to a point lying on the same epipolar line in the target image (i.e., gallery image). After rectification of the two ear images, the matched points have necessarily the same coordinate in the both images. Therefore, in case of searching for corresponding points in two ear images, it is only necessary to search in the same epipolar line, reducing a 2D search space to 1D. In order to achieve rectification, we adopt the algorithm proposed by Fusiello et al. [24].

3.6 Stereo Algorithm and Matching Costs

We employ a robust and fast stereo algorithm for matching correspondences between the probe and gallery images based on fuzzy correlation measure. The aim of matching correspondences is to compute the measure of similarity or matching cost for identification of the user's ear. To determine the correspondences between two images, we match the windows of pixels on the same epipolar lines in the reference (probe) and target (gallery) image. In our method, we assume that the pixels surrounded by a window possess approximately equal disparity. Thus, the matching cost C for a pixel (x, y) in the probe image is estimated by taking a window of pixels centered at (x, y) in the probe image, and placing a similar window of pixels centered at $(x + d, y)$ in the gallery image and computing the difference between these two windows using a fuzzy correlation measure given by the following Eq. (3). Here, d is a searching range over the same epipolar line in the gallery image.

$$C(x,y,d) = \frac{\sum_{x,y\in W} F(x,y)|I_P(x,y)\times I_G(x+d,y)|}{\sqrt{\sum_{x,y\in W} F(x,y)I_P^2(x,y)\times \sum_{x,y\in W} F(x,y)I_G^2(x+d,y)}} \tag{3}$$

where $I_P(x, y)$ and $I_G(x, y)$ are the intensities of the pixels at position (x, y) in the probe and gallery images, respectively; and W is a square window. F(x, y) is the fuzzy measure corresponding to the pixel at position (x, y), has Gaussian distribution which is proportional to fuzzy membership function:

$$F(x,y) = \exp(-\frac{|I_P(x,y) - I_G(x+d,y)|^2}{2\sigma^2}) \tag{4}$$

where, σ is the standard deviation of all pixels within the window.

The matching cost C(x, y) for every pixel (x, y) can be computed by the winner-take all strategy such that,

$$C(x,\ y) \ = \ \arg\max C(x,\ y,\ d) \tag{5}$$

3.7 Final Matching Cost and Recognition

In order to authenticate a user, the matching is performed between the probe ear image and the enrolled gallery pattern. For recognition process, a number of iteration is accomplished for matching the probe image with the stored gallery images. When we match a probe image to a gallery image using our proposed stereo algorithm, we obtain different window costs. We pick the best matching scores and estimate a normalized (average) matching cost for every pair of the probe and the gallery images, by using the following equation:

$$C(I_P, I_G) = \frac{\sum_{i=1}^{n} C(I_{P,i}, I_{G,i})}{\sum_{i=1}^{n} |I_{P,i}| + |I_{G,i}|} \tag{6}$$

where C is the normalized matching cost for the image pair: the probe and a gallery ear image. Thus, we compute normalized costs for all pair of images by comparing the probe with all gallery images. We then identify the gallery ear image that provides best similarity measure given by,

$$\text{Similarity, } S \ = \ \max\{C_n(I_{\mathrm{P}},\ I_{\mathrm{G}})\} \tag{7}$$

where, C_n refers to the normalized cost of n^{th} image pair (the probe and the gallery image), n = 1 … N; and N denotes the total number of images considered in the gallery. The best match is considered for identification when, S > T. Here, T is a predetermined threshold and is set to 0.65 by empirical evaluation.

4 Experimental Evaluation

In this section, we evaluate the performance of our proposed algorithm and compare with other similar techniques reported in this work. To demonstrate the effectiveness of our algorithm, we perform experiment using several real images and standard ear datasets as well. Experiments are carried out on a computer with 2.8 GHz Intel Core i7 processor. The algorithm has been implemented using Visual C++.

4.1 Datasets

In this experiment, we use three different standard datasets of ear images, prepared by the University of Notre Dame (UND) [15], the University of Science and technology in Beijing (USTB) [16], and the Indian Institute of Technology, Delhi (IITD) [17]. The UND dataset includes 942 images of 302 human subjects, the USTB database contains 308 images of 77 subjects, and the IIITD database includes 421 images of 121 subjects. The detailed features of the ear databases are summarized in Table 1. Figure 2 shows some sample images of these ear databases.

Table 1. Features of the ear datasets.

Dataset	Total images	Individuals	Additional features
UND-F	942	302	3D and corresponding 2D profile images from 302 human subjects including some partially occluded images, captured in 2003 and 2004
USTB	308	77	Images were captured from 77 human subjects in 4 different sessions between November 2003 and January 2004
IITD	421	121	3 images were taken per subject in an indoor environment, collected between October 2006 and June 2007

4.2 Results

The authentication process has been successfully evaluated with 100% accuracy using the real image sequences. To evaluate the recognition performance, the algorithm is tested with three standard ear datasets: UND, USTB and IITD. In this work, we use one image for every subject from each dataset as a probe image while the remaining one image is used as the gallery image. Figures 3 and 4 demonstrate ear detection and local feature extraction process, respectively. The performance of our proposed recognition scheme is evaluated using the stereo matching algorithm which is fast and efficient for user identification. The recognition performance of our proposed method is compared with other existing similar methods such as, support vector machine (SVM) [10], AdaBoost [11], neural network (NN) [12], modular neural network (MNN) [13], and NN with principal component analysis (NN + PCA) [14]. The comparisons for different methods are reported in the Fig. 5, which clearly indicates the superiority of our

Fig. 2. Example of profile images with ear of different shapes: left ears (top), and right ears (bottom).

proposed method. We also test our algorithm using different datasets and experimental evaluation indicates that our approach provides better performance for the UND database with a recognition rate of 98.96%, as shown in Fig. 6. The proposed algorithm achieves a very low false positive rate (FPR) which is 0.25%. Figure 7 presents a comparison of FPRs for different methods.

We compare the computation time of our recognition algorithm with other methods. A Visual C++ implementation of our algorithm requires around 0.39 s to extract the local edge features from a probe ear image, and the average time to match a probe-gallery pair in the recognition process is 0.18 s on the UND dataset. Table 2 summarizes the comparison for matching time of the recognition approach of this paper with others on the UND database. Matching times are computed on different machines

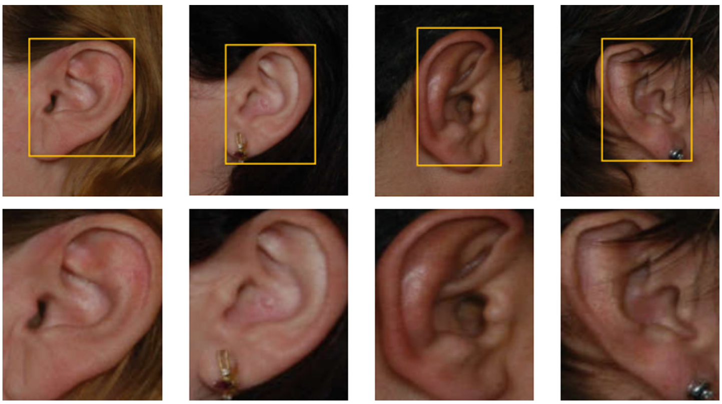
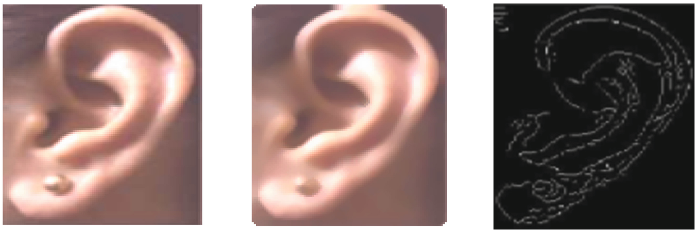


Fig. 3. Detection process: detected ear shape (top) and extracted ear (bottom).



(a) (b) (c)

Fig. 4. Features extraction: (a) initial ear shape, (b) refined ear data, and (c) extracted edge features.

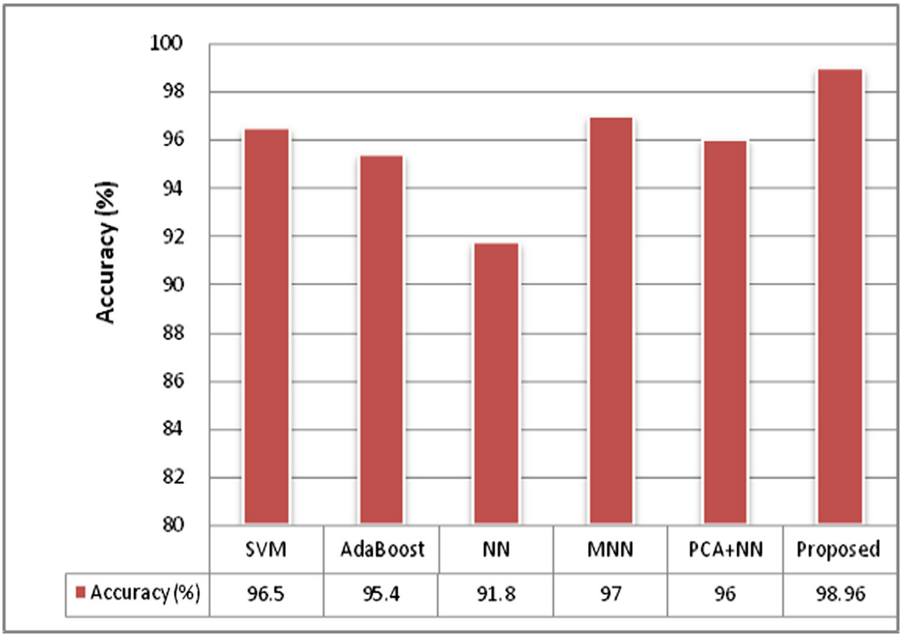


Fig. 5. Comparison for recognition accuracy with different methods.

in different approaches. Results show that our proposed matching algorithm can achieve superior performance with significant reduction of computation time compared to other methods. Empirically we find that a window of size 3 × 3 pixels and a range value of +5 (d) for searching correspondence pixels are good choices for better results.

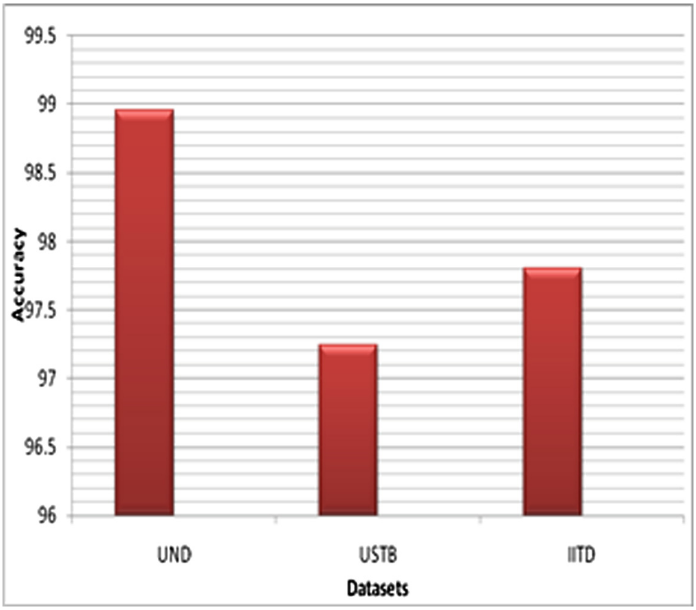


Fig. 6. Recognition accuracy on different datasets.

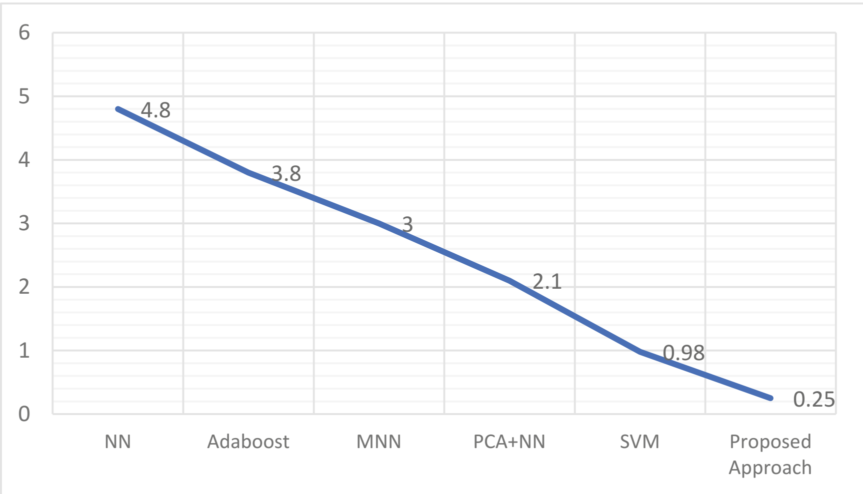


Fig. 7. Comparison of false positive rates for different methods.

Table 2. Computational time for matching a pair of probe-gallery ear image with different methods on UND dataset.

Method	Feature extraction time	Average matching time
Our method (Visual C++)	0.39 s	0.18 s
Islam et al. [11] (MATLAB)	22.2 s	2.28 s
Chen and Bhanu [26] (C++)	N/A	1.1 s

5 Conclusion

This paper proposes a robust and efficient human identification based on ear biometric trait employing a hybrid neural network. Our system could be capable to cope with pose variations, occlusion and illumination changes. The effectiveness of the proposed algorithms has been tested with different standard datasets. Experimental evaluation confirms that our proposed method achieves superior performance comparable to other existing similar methods. Our next target is to extend the algorithm by combining it with other biometric modalities such as facial features to develop a more robust, secure and accurate human recognition system. We believe that this proposed method will be useful for many real-time applications where very fast processing time is important.

References

1. Chowdhury, M., Gao, J., Islam, R.: Biometric authentication using facial recognition. In: Deng, R., Weng, J., Ren, K., Yegneswaran, V. (eds.) SecureComm 2016. LNICST, vol. 198, pp. 287–295. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-59608-2_16
2. Marqués, I., Graña, M.: Image security and biometrics: a review. In: Corchado, E., Snášel, V., Abraham, A., Woźniak, M., Graña, M., Cho, S.-B. (eds.) HAIS 2012. LNCS (LNAI), vol. 7209, pp. 436–447. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-28931-6_42
3. Jain, A., Kumar, A.: Biometric recognition: an overview. In: Mordini, E., Tzovaras, D. (eds.) The International Library of Ethics, Law and Technology, vol. 11, pp. 49–79. Springer, Heidelberg (2012). https://doi.org/10.1007/978-94-007-3892-8_3
4. Islam, S.M.S., Bennamoun, M., Owens, R., Davies, R.: A review of recent advances in 3D ear and expression invariant face biometrics. ACM Comput. Surv. **44**(3), 14:1–14:34 (2012)
5. Islam, S.M.S., Davies, R., Bennamoun, M., Owens, R.A., Mian, A.S.: Multibiometric human recognition using 3D ear and face features. Pattern Recogn. **46**(3), 613–627 (2013)
6. Choras, M.: Ear biometrics based on geometrical feature extraction. Electron. Lett. Comput. Vis. Image Anal. **5**, 84–95 (2005)
7. Yuizono, T., Wang, Y., Satoh, K., Nakayama, S.: Study on individual recognition for ear images by using genetic local search. In: Proceedings of Congress on Evolutionary Computation, pp. 237–242 (2002)
8. Hurley, D.J., Nixon, M.S., Carter, J.N.: Force field feature extraction for ear biometrics. Comput. Vis. Image Underst. **98**(3), 491–512 (2005)
9. Yan, P., Bowyer, K.W.: Biometric recognition using 3D ear shape. IEEE Trans. PAMI **29**(8), 1297–1308 (2007)

10. Yaqubi, M., Faez, K., Motamed, S.: Ear recognition using features inspired by visual cortex and support vector machine technique. In: International Conference on Computer and Communication Engineering (ICCCE), pp. 533–537 (2008)
11. Islam, S., Davies, R., Bennamoun, M., Mian, A.: Efficient detection and recognition of 3D ears. Int. J. Comput. Vis. **95**, 52–73 (2011)
12. Wang, X., Xia, H., Wang, Z.: The research of ear identification based on improved algorithm of moment invariants. In: Third International Conference on Information and Computing (ICIC), p. 58 (2010)
13. Gutierrez, L., Melin, P., Lopez, M.: Modular neural network integrator for human recognition from ear images. In: The 2010 International Joint Conference on Neural Networks (IJCNN) (2010)
14. Alaraj, M., Hou, J., Fukami, T.: A neural network based human identification framework using ear images. In: TENCON (2010)
15. UND (2005) Database. http://www.nd.edu/cvrl/CVRL/DataSets.html
16. USTB (2002) Database. http://www.en.ustb.edu.cn/resb/
17. IIT Delhi ear database. http://www4.comp.polyu.edu.hk/~csajaykr/IITD/Database_Ear.htm
18. Liu, H., Liu, D.: Improving adaboost ear detection with skin-color model and multi-template matching. In: 3rd IEEE ICCSIT, vol. 8, pp. 106–109 (2010)
19. Chowdhury, M., Gao, J., Islam, R.: Fuzzy logic based filtering for image de-noising. In: IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2016), Vancouver, Canada (2016)
20. Castillo, C.D., Jacobs, D.W.: Using stereo matching for 2D face recognition across pose. In: Proceedings IEEE International Conference Computer Vision and Pattern Recognition (2007)
21. Ashraf, A.B., Lucey, S., Chen, T.: Learning patch correspondences for improved viewpoint invariant face recognition. In: Proceedings IEEE International Conference Computer Vision and Pattern Recognition, June 2008
22. Chowdhury, M., Gao, J., Islam, R.: Fast stereo matching with fuzzy correlation. In: IEEE Conference on Industrial Electronics and Applications (ICIEA 2016), Hefei, China (2016)
23. Chowdhury, M., Bhuiyan, M.A.: Fast window based stereo matching for 3D scene reconstruction. Int. Arab J. Inf. Technol. **10**(3), 209–214 (2013)
24. Fusiello, A., Trucco, E., Verri, A.: A compact algorithm for rectification of stereo pairs. Mach. Vis. Appl. **12**, 16–22 (2000)
25. Kumar, R., Selvam, P., Rao, K.N.: Pattern extraction methods for ear biometrics: a survey. In: Proceedings World Congress on Nature & Biologically Inspired Computing (NaBIC 2009), Coimbatore, India, pp. 1657–1660 (2009)
26. Chen, H., Bhanu, B.: Human ear recognition in 3D. IEEE Trans. PAMI **29**(4), 718–737 (2007)