

Ear Detection and Recognition Techniques: A Comparative Review



Pallavi Srivastava, Diwakar Agrawal and Atul Bansal

Abstract Among several types of biometric systems, ear recognition is a bustling research area. Due to the minimal cooperation of the user, this biometric trait proves to be a good application in security and surveillance. Over the period of last two decades, various contributions have been reported with robust techniques and approaches in ear biometrics. This paper provides an overview of various ear recognition and detection techniques using 2D ear images, among which some are automated and some are not. Also, a comparative review of the available databases for research purposes is provided. A comparative vision of ear detection and recognition is presented in this paper in chronological order.

Keywords Detection · Recognition · Feature extraction · Biometrics · Databases

1 Introduction

With the increasing invasion of technology in every regard to living nowadays, the world is becoming more and more digitized. This makes it difficult to protect confidential information. Conventional keys and passwords are not any more secure to corroborate that the data is out of reach of unauthorized users. This has brought biometric authentication in focus, as it is a productive way to authenticate an individual's identity. Biometric authentication is the procedure of validating an individual's identity based on some unique and measurable traits of that individual. These traits are innate and distinctive to each person and can be classified into physical and behavioral characteristics like, face, fingerprints, gait, palm print, ear, voice, keystroke

P. Srivastava (✉) · D. Agrawal · A. Bansal
GLA University, Mathura 281406, India
e-mail: pallavi.srivastava_mtec17@gla.ac.in

D. Agrawal
e-mail: diwakar.agarwal@gla.ac.in

A. Bansal
e-mail: atul.bansal@gla.ac.in

© Springer Nature Singapore Pte Ltd. 2020
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,
Lecture Notes in Networks and Systems 94,
https://doi.org/10.1007/978-981-15-0694-9_50

dynamics, and signature dynamics; among all the physical characteristics, ear recognition has emerged as an active research area. For ear detection and recognition, many advanced techniques and approaches have evolved. Various databases that are required in robust training and testing purposes of ear recognition and detection are available publicly. Images required for ear authentication in an automated system can be extracted from video sequences or profile headshots. One of the reasons because of which ear biometrics has gained immense interest is that, to seize the ear image user's cooperation is not needed. Burge and Burger [1] proposed a passive identification machine vision system. This system localizes and segments the ear of a subject applying deformable counters on a Gaussian pyramid representation of an image gradient. After this, a graph model is constructed using the edges and curves within the ear and then for classification, a graph matching algorithm is used. Moreno et al. [2] in 1999 were the first to build a fully automated ear recognition system. The features used were ear shapes, wrinkles, and outer ear point. Mu et al. [3] expanded this approach. They combined the inner ear structure and outer ear shape as a feature vector. For classification, neural network was used. Yuizono et al. [4] exploited genetic local search to reduce the error between the training and testing image. They noted the registrant recognition rate approximately 100%. This paper gives an overview of existing ear detection and recognition techniques and surveys databases available publicly.

2 Background of Ear Detection and Recognition

2.1 *The Ear Framework*

The evolution of the external ear is a complex process. The embryonic period of ear starts from the fifth week of pregnancy and continues to the postnatal period. The structural support for evolution of the external ear is contributed by pharyngeal arch apparatus. The ear is developed into different segments namely inner ear, outer ear, and middle ear. The outer ear consists of pinna which is the unification of six auricle hillocks and ectoderm, which is the external auditory meatus. The middle ear and eustachian tube are formed by pharyngeal pouch endoderm. The basic terminology of the external ear is shown in Fig. 1. The visible and prominent part of the outer ear is called pinna. It is formed by helix which unites into the lobe. Antihelix is parallel to the helix, then there is concha which is a conch-formed space and merges into incisura which has two side ridges called tragus and antitragus.

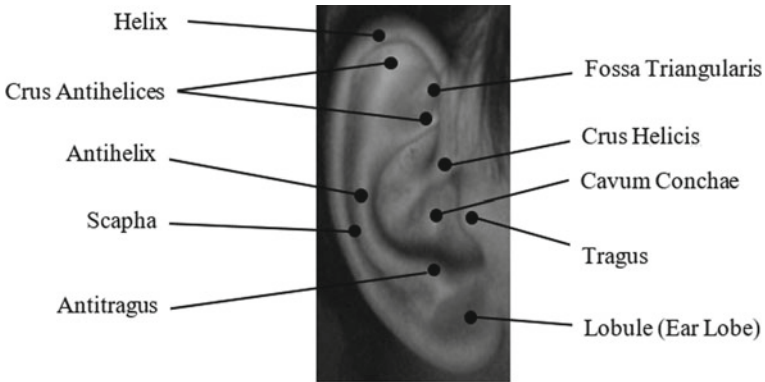


Fig. 1 Structure of the ear [22]

2.2 Operation of a Traditional Ear Biometric System

An ear biometric system has two phases of operation, namely enrolment and recognition as shown in Fig. 2. In the enrolment phase, biometric sensor examines the ear image of the user in order to obtain the digital pattern. Then this pattern undergoes feature extraction techniques to produce better demonstrative presentation called feature set or feature vector, which is stored in the database and called template. In the recognition phase, a new sample of the user to be authenticated is scanned and the pattern of the sample is generated using feature extraction techniques. Then this sample pattern is compared to the template stored in the database. This comparison is performed by a classification technique which results in distance or score which

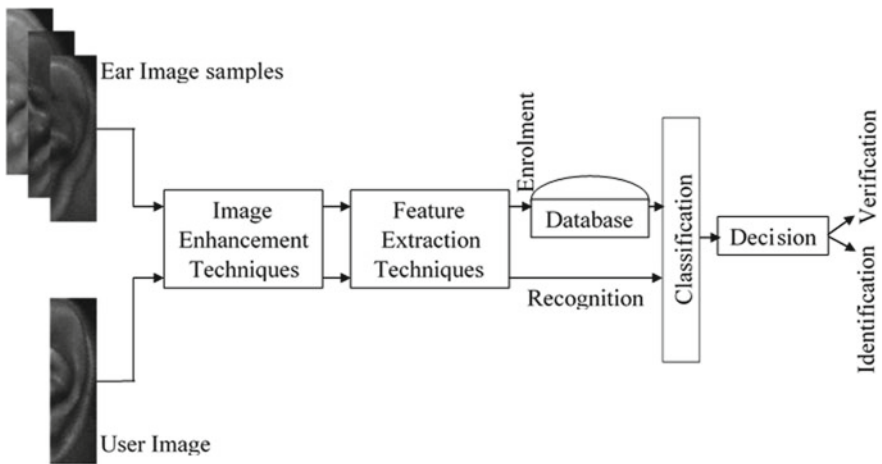


Fig. 2 A traditional ear biometric system

determines the similarity/dissimilarity between the sample image pattern and template. The sample image pattern is assigned to the template which has the minimum score or distance unit, and the identity of the query image is divided.

3 State-of-the-Art in Ear Recognition and Detection Systems

3.1 Ear Recognition Approaches

Burge and Burger [5] proposed a method with the ability to work successfully are passive identification system. This paper shows the biometric system abstractly including labels like uniqueness and estimation. The algorithm uses a graph matching method for classification for user identification. Hurley et al. [6] proposed a method which converts the ear image into the force field. The structure of the force field has peaks that are called potential energy well. Every energy well has a potential channel linked with it. This model including both the steps, that is, force field transform and potential well and channel extraction forms the basis of ear description. Victor et al. [7] exploited principal component analysis (PCA) for ear identification. This approach is applied to ear and face images in the dataset.

Zavar et al. [8] built a system to automatically enroll ear images and recognize ear. It uses part-wise ear description incorporating SIFT. Further, the model is extended using Log-Gabor filter for wavelet analysis. Prakash and Gupta [9] proposed an ear recognition technique to minimize the problem like the pose, dire (poor), contrast, and illumination variation. They executed the algorithm using three image upgrading techniques to nullify the poor low contrast, light, and noise effect. For feature extraction, SURF was carried out. Kumar and Chan [10] introduced an approach based on the sparse representation of Radon transform; the adjacency relation between the gray scale of the image are converted as the superior gray-scale characteristic orientations in the local region. Basit and Shoaib [11] presented an ear recognition technique constructed using curvelet transform. The feature extraction step is performed by applying fast discrete curvelet transform and for classification, KMM is exploited.

Nigam and Gupta [12] proposed an ear recognition system that takes a major problem. This system preprocesses the image and performs a Canny edge detection method. For image enhancement, Contrast Limited histogram equalization (CLAHE) is exploited. Image transformation is conducted using Gradient Ordinal Relation Pattern (GORP) and STARGORP (SGORP). Pflug et al. [13] applied certain texture and surface descriptors and presented their performance. Different texture and surface descriptors utilized are LBP, LPQ, HOG, and BSIF. They also proposed a histogram-based descriptor that can be utilized when the fusion of two different information

passages are required. Anwar et al. [14] proposed an algorithm for geometrical features based on ear recognition. The preprocessed ear images are used by the snake model for detecting the ear.

Youbi et al. [15] presented a human ear recognition algorithm which utilizes MLBP-based feature extraction and for capturing the similarity and dissimilarity, KL distance is used. This system gave 95% of rank 1 identification rate. Features are extracted by first dividing the image into blocks. Ghoualmi et al. [16] proposed a system that performs better image enhancement using an artificial bee colony (ABC) algorithm. Features are extracted using scale-invariant feature transform (SIFT) and for matching/classification, Euclidean distance is used. Emersic et al. [17] proposed an ear detection technique taking into account problems like low illumination and occlusions. For this purpose, they used convolution encoder–decoder networks (CEDs) which are based on SegNet architecture.

Chowdhury et al. [18] presented an ear recognition methodology which utilizes the invariable edge local features and for classification neural network was used. This approach was applied to different databases and performance was compared with some state-of-the-art methods. Sarangi et al. [19] proposed a new ear recognition scheme which utilizes PHOG and LDA. Local features are extracted using pyramid histogram of oriented gradients (PHOG) and for dimension reduction of the PHOG descriptor, linear discriminant analysis (LDA) was used. The training and testing images are classified using nearest neighbor.

Alqaralleh and Toygar [20] presented a 2D ear recognition method. Features are extracted from tragus and non-occluded part of the ear by local binary pattern (LBP) texture descriptor. Then the score between training and testing samples of the tragus and ear image are calculated separately. After this, the match scores of both tragus and ear image is fused and finally classified using KNN. Alshazly et al. [21] proposed an ear recognition approach based on gradient features, namely Histogram of Oriented Gradients (HOG), Local Optimal Oriented Patterns (LOOP), Local Directional Patterns (LDP), and Weber Local Descriptor (WLD). For classification, chi-square similarity was incorporated. Emersic et al. [22] presented a wide overview of automatic ear recognition techniques which are mainly descriptor based. They also presented several datasets for research work in the same area.

3.2 Comparison of Existing Databases for Ear Detection and Recognition

In order to evaluate the performance parameters after training and testing of the detection or recognition techniques of an Ear biometric system, image databases of sufficient size are required. In this section, a comparison of several databases that have been used in the literature for estimating the performance of Ear detection and recognition systems is given. Some of the datasets have both raw images and preprocessed images which are present in the normalized form. Table 1 represents a

Table 1 A comparative summary of the available databases and their features. The column “Database” provides the name of the database. “No. of Subjects” and “No. of Images/Videos” presents the number of subjects and a total number of images in that particular database. The column “Gender” indicates whether both genders are present or not. And the last column “Occlusion” indicates whether the images are occluded or not

Database	No. of subjects	No. of Images/Videos	Gender (Male\Female)	Occlusion
WVU [23]	402	460 Videos	Both	Yes
USTB [24]				
I	60	80 Images	Both	No
II	77	308 Images		
III	79	1738 Images		
IV	500	8500 Images		
UCR [25]	55	902 Images	Both	Yes
UND [26]				
Collection F	114	464 Images	Both	No
Collection E	302	942 Images		
Collection G	235	738 Images		
Collection J2	415	1800 Images		
UMIST [27]	20	564 Images	Both	No
XM2VTS [28]	295	4 Videos	Both	No
FERET [29]	1199	14,126 Images	Both	No
CAS-PEAL [30]	1040	99,594 Images	Both	Yes
IIT Delhi [31]	221	793 Images	Both	No
IIT Kanpur [32]				
Subset I	190	801 Images	Both	No
Subset II	89	801 Images		
AWE [33]	100	1000 Images	Both	Yes
UBEAR [34]	126	4420 Images	Both	Yes
NCKU [35]	90	3330 Images	Both	Yes
YSU [36]	259	2590 Images	Both	No

comparative summary of the available databases and their features. Databases may have videos or images with a number of images with different subjects. Most of the databases are freely available or can be provided by applying for a license.

3.3 Comparison of Various Ear Detection and Recognition Methods

After discussing the several existing techniques and algorithms in the domain of ear detection and recognition, also about the available databases for applying the training and testing approach and measuring the performance. In this section, Table 2 presents a comparative summary of the approaches and techniques that have been surveyed in this paper. Techniques are sequenced in chronological order with a brief description of them and the reported results. It indicates the databases used by authors; label “Own” indicates that authors have used the database created or collected on their own. Tag “NA” indicates that the information is not provided by the author. Table 2 provides information about the number of subjects and the total number of images contained by that database. It indicates the degree of occlusion in the images provided by the database. The column “Result” indicates the performance evaluation of the techniques incorporated in the references. Performance of the techniques is indicated in terms of rank 1 recognition rate (I), equal error rate (EER), identification rate (i), and detection rate (D).

Table 2 A comparative overview of several Ear Detection and Recognition techniques. Column “Dataset” provides the name of the database used in that reference. “No. of Subjects” and “No. of Images” presents the number of subjects and a total number of images in that particular database. Column “Result” indicates the noted performance of the system. Then the column “Description” gives a short description of the techniques used by the authors

References	Dataset	No. of subjects	No. of images	Result	Description
Burge and Burger [5]	Own	NA	NA	NA	Ear images represented as Voronoi diagram
Hurley et al. [6]	Own	NA	NA	NA	Ear images are dealt as a Gaussian attractors
Victor et al. [7]	Own	294	808	40(I)	Eigen vector techniques are applied to ear images
Chang et al. [37]	UNDE	114	NA	71.6(I)	Training images are stored as “Ear pace”
Yuan et al. [38]	USTB II	77	NA	91(I)	Nonnegative matrix factorization ear recognition
Choras [39]	Own	188	NA	86.2(I)	Contour of ear images extracted as features
Zavar and Nixon [8]	XM2VTS	63	NA	99.5(I)	Utilizes SIFT and Log-Gabor filter for feature extraction
Prakash and Gupta [9]	IITK	190	NA	2.8(EER)	Dealt with pose, low contrast, and illumination variation

(continued)

Table 2 (continued)

References	Dataset	No. of subjects	No. of images	Result	Description
Islam et al. [40]	UND-F	NA	429	95.4(I)	Utilizes 2D AdaBoost detector along with 3D local feature extraction
Kumar and Chan [10]	IITD	125	NA	97.56(I)	Sparse representation feature extraction
	IITD	221		96.9(I)	
	UND	110		92.6(I)	
Basit and Shoaib [11]	IITD II	221	442	96.2(I)	Features extracted using FDCT via wrapping technique
Nigam and Gupta [12]	IITD	125	493	99.2(I)	Reference point method and GORP for image normalization and transformation
	UND-E	114	443		
Pflug et al. [13]	UND J2	158	NA	98.7(I)	Feature extraction performed by LPQ, BSIF, LBP, and HOG
	AMI	100		100(I)	
	IITK	72		99.2(I)	
Jiajia Lei et al. [41]	UND- F	302	942	100(I)	3D ear landmark localization, detection, and pose classification
	UND-G	113	512		
	UND-J2	415	1800		
Anwar et al. [14]	IITD I	50	150	98(I)	Snake Model for ear detection and Canny edge for image enhancement
Zineb Youbi et al. [15]	IITD	121	471	95(i)	Feature extraction by MLBP and classification using KL divergence
Ghoulmi et al. [16]	IITD	125	421	94.8(I)	ABC algorithm[] and SIFT perform the feature extraction
	USTB1	60	180		
	USTB2	77	308		
Emersic et al. [17]	AWE	100	1000	99.21(D)	Gives a PED-CED ear detection technique
Chowdhury et al. [18]	UND	302	942	98.2(I)	Detection using AdaBoost-based detector
Sarangi et al. [19]	UND E	114	464	96.6(I)	PHOG and LDA for feature extraction and dimensionality reduction
				3.395(EER)	
Alqaralleh and Toygar [20]	USTB-3	NA	NA	92.3(I)	Fusion of features from tragus and another part of the ear is done using LBP descriptor
Alshazly et al. [21]	IITD	221	793	97(I)	Gradient-based feature extraction methods are used

4 Conclusion

This paper features the use of computer vision and image processing technology in the field of Ear detection and recognition. It discusses various existing Ear identification, segmentation, detection, and recognition techniques, and approaches. It also provides a comparison of the performance of several approaches. Table 2 gives a comparative review of several existing techniques and their features. It shows the databases used by the author and the corresponding performances noted by them in terms of recognition rate, equal error rate, identification rate, and detection rate. Table 1 provides a comparative overview of the databases that are available for the researchers who wish to test or propose new Ear recognition and detection techniques. It includes features like number of subjects, number of images corresponding to the subjects; it also indicates whether the images are occluded or not.

References

1. Burge, M., & Burger, W. (1997, May). Ear biometrics for machine vision. In *21st workshop of the Austrian association for pattern recognition* (pp. 275–282).
2. Moreno, B., Sanchez, A., & Vélez, J. F. (1999). On the use of outer ear images for personal identification in security applications. In *Proceedings IEEE 33rd Annual 1999 International Carnahan Conference on Security Technology* (pp. 469–476).
3. Mu, Z., Yuan, L., Xu, Z., Xi, D., & Qi, S. (2004). Shape and structural feature based ear recognition. *Advances in biometric person authentication* (pp. 663–670). Berlin: Springer.
4. Yuizono, T., Wang, Y., Satoh, K., & Nakayama, S. (2002, May). Study on individual recognition for ear images by using genetic local search. In *Proceedings of the 2002 Congress on Evolutionary Computation* (Vol. 1, pp. 237–242).
5. Burge, M., & Burger, W. (2000). Ear biometrics in computer vision. In *Proceedings 15th International Conference on Pattern Recognition* (Vol. 2, pp. 822–826).
6. Hurley, D. J., Nixon, M. S., & Carter, J. N. (2002). Force field energy functionals for image feature extraction. *Image and Vision Computing*, 20, 311–317.
7. Victor, B., Bowyer, K., & Sarkar, S. (2002). An evaluation of face and ear biometrics. In *Object recognition supported by user interaction for service robots* (Vol. 1, pp. 429–432).
8. Arbab-Zavar, B., & Nixon, M. S. (2011). On guided model-based analysis for ear biometrics. *Computer Vision and Image Understanding*, 115, 487–502.
9. Prakash, S., & Gupta, P. (2013). An efficient ear recognition technique invariant to illumination and pose. *Telecommunication Systems*, 52, 1435–1448.
10. Kumar, A., & Chan, T. S. T. (2013). Robust ear identification using sparse representation of local texture descriptors. *Pattern Recognition*, 46, 73–85.
11. Basit, A., & Shoaib, M. (2014). A human ear recognition method using nonlinear curvelet feature subspace. *International Journal of Computer Mathematics*, 91, 616–624.
12. Nigam, A., & Gupta, P. (2014, November). Robust ear recognition using gradient ordinal relationship pattern. In *Asian conference on computer vision* (pp. 617–632). Cham: Springer.
13. Pflug, A., Paul, P. N., & Busch, C. (2014, October). A comparative study on texture and surface descriptors for ear biometrics. In *2014 international carnahan conference on security technology* (pp. 1–6).
14. Anwar, A. S., Ghany, K. K. A., & Elmahdy, H. (2015). Human ear recognition using geometrical features extraction. *Procedia Computer Science*, 65, 529–537.

15. Youbi, Z., Boubchir, L., Bounneche, M. D., Ali-Chérif, A., & Boukrouche, A. (2016, June). Human ear recognition based on multi-scale local binary pattern descriptor and KL divergence. In *2016 39th international conference on telecommunications and signal processing* (pp. 685–688).
16. Ghoulmi, L., Draa, A., & Chikhi, S. (2016). An ear biometric system based on artificial bees and the scale invariant feature transform. *Expert Systems with Applications*, *57*, 49–61.
17. Emeršič, Ž., Gabriel, L. L., Štruc, V., & Peer, P. (2017). Pixel-wise ear detection with convolutional encoder-decoder networks. [arXiv:1702.00307](https://arxiv.org/abs/1702.00307).
18. Chowdhury, M., Islam, R., & Gao, J. (2017, June). Robust ear biometric recognition using neural network. In *2017 12th IEEE conference on industrial electronics and applications* (pp. 1855–1859).
19. Sarangi, P. P., Mishra, B. S. P., & Dehuri, S. (2017, February). Ear recognition using pyramid histogram of orientation gradients. In *2017 4th international conference on signal processing and integrated networks* (pp. 590–595).
20. Alqaralleh, E., & Toygar, Ö. (2018). Ear recognition based on fusion of ear and tragus under different challenges. *International Journal of Pattern Recognition and Artificial Intelligence*, *32*, 1856009.
21. Alshazly, H. A., Hassaballah, M., Ahmed, M., Ali, A. A. (2018, September). Ear biometric recognition using gradient-based feature descriptors. In *International conference on advanced intelligent systems and informatics* (pp. 435–445). Cham: Springer.
22. Emeršič, Ž., Štruc, V., & Peer, P. (2017). Ear recognition: More than a survey. *Neurocomputing*, *255*, 26–39.
23. West Virginia University Libraries. Retrieved January 15, 2019 from <https://lib.wvu.edu/databases/>.
24. Ear Recognition Laboratory Homepage at University of Science & Technology Beijing (USTB). Retrieved January 20, 2019 from <http://www1.ustb.edu.cn/resb/en/index.htm>.
25. UCR Library. Retrieved January 20, 2019 from <https://library.ucr.edu/research-services/databases>.
26. UND Chester Fritz Library. Retrieved January 15, 2018 from <https://library.und.edu/databases>.
27. UMIST Database. Retrieved January 20, 2019 from <https://www.sheffield.ac.uk/eee/research/iel/research/face>.
28. The XM2VTS Database. Retrieved January 17, 2019 from <http://www.ee.surrey.ac.uk/CVSSP/xm2vtsdb/>.
29. Face recognition technology (FERET). Retrieved January 15, 2019 from <https://www.nist.gov/programs-projects/face-recognition-technology-feret>.
30. The CAS-PEAL face database. Retrieved January 17, 2019 from <http://www.jdl.ac.cn/peal/top.htm>.
31. IIT Delhi Ear Database. Retrieved January 20, 2019 from http://www4.comp.polyu.edu.hk/~csajaykr/IITD/Database_Ear.htm.
32. IIT Kanpur Ear Database. Retrieved January 15, 2019 from <http://home.iitk.ac.in/~ganil/>.
33. AWE Database. Retrieved January 15, 2019 from <https://www.awedatabase.com/>.
34. UBEAR. Retrieved January 17, 2019 from <http://ubear.di.ubi.pt/>.
35. NCKU Database. Retrieved January 15, 2019 from http://robotics.csie.ncku.edu.tw/Databases/FaceDetect_PoseEstimate.htm.
36. YSU Database. Retrieved January 20, 2019 from <http://lib.ysu.am/libraries.html>.
37. Chang, K., Bowyer, K. W., Sarkar, S., & Victor, B. (2003). Comparison and combination of ear and face images in appearance-based biometrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*, 1160–1165.
38. Yuan, L., Mu, Z. C., Zhang, Y., & Liu, K. (2006, August). Ear recognition using improved non-negative matrix factorization. In *18th international conference on pattern recognition* (Vol. 4, pp. 501–504).
39. Choraś, M. (2008). Perspective methods of human identification: Ear biometrics. *Opto-Electronics Review*, *16*, 85–96.

40. Islam, S. M., Davies, R., Bennamoun, M., & Mian, A. S. (2011). Efficient detection and recognition of 3D ears. *International Journal of Computer Vision*, *95*, 52–73.
41. Lei, J., You, X., & Abdel-Mottaleb, M. (2016). Automatic ear landmark localization, segmentation, and pose classification in range images. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *46*, 165–176.