A Survey of 3D Ear Recognition Techniques

IYYAKUTTI IYAPPAN GANAPATHI, C2PS, Khalifa University, UAE
SYED SADAF ALI, NGOC-SON VU, École Nationale Supérieure de l'Électronique et de ses Applications (ENSEA), France
SURYA PRAKASH, Indian Institute of Technology Indore, India
NAOUFEL WERGHI, C2PS & KUCARS, Khalifa University, UAE

Human recognition with biometrics is a rapidly emerging area of computer vision. Compared to other well-known biometric features such as the face, fingerprint, iris, and palmprint, the ear has recently received considerable research attention. The ear recognition system accepts 2D or 3D images as input. Since pose, illumination, and scale all affect 2D ear images, it is evident that they all impact recognition performance; therefore, 3D ear images are employed to address these issues. The geometric shapes of 3D ears are utilized as rich features to improve recognition accuracy. We present recent advances in several areas relevant to 3D ear recognition and provide directions for future research. To the best of our knowledge, no comprehensive review has been conducted on using 3D ear images in human recognition. This review focuses on three primary categories of 3D ear recognition techniques: (1) registration-based recognition, (2) local and global feature-based recognition, and (3) a combination of (1) and (2). Based on the above categorization and publicly available 3D ear datasets, this article reviews existing 3D ear recognition techniques.

CCS Concepts: • **Computing methodologies → Biometrics**; **Shape representations**; **Matching**; **Interest point and salient region detections**; **Object recognition**.

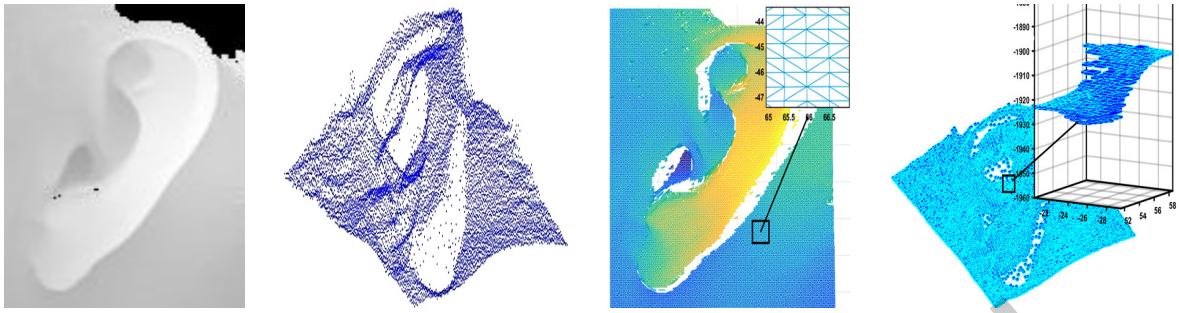
Additional Key Words and Phrases: Biometrics, 2D/3D Ear, Verification/Identification, Local/Global Features, ICP, Age invariant, Inheritance, Data quality

1 Introduction

Recognition is an essential task in the computerized world, and the traditional recognition systems which depend on tokens and passwords are unreliable and extremely harmful. The token may be lost, stolen, or exchanged, and lengthy passwords are hard to remember; hence, recognition systems based on individual biometrics can overcome the limitations of traditional systems [6, 7, 13, 14, 24, 45, 46, 70, 85, 87, 88, 99, 112, 145, 169, 189]. "There's real power in using the appearance of an ear for computer recognition, compared to facial recognition. It's roughly equivalent, if not better", said Professor Kevin Bowyer of the University of Notre Dame [165]. Among the many biometrics that has emerged in recent years, the ear has emerged as one of the most reliable [3, 15, 20, 22, 47, 60, 67, 71, 75, 86, 131, 138]. Unlike the face, the ear does not change its shape with different expressions and is not affected by cosmetics and wrinkles due to aging. Further, the ear size is greater than the fingerprint and iris and smaller than the face; hence, it can be easily acquired and processed. Ear biometrics is

Authors' addresses: Iyyakutti Iyappan Ganapathi, C2PS, Khalifa University, P.O.Box: 127788, Abu Dhabi, UAE, iyyakutti.ganapathi@ku.ac.ae; Syed Sadaf Ali, Ngoc-Son Vu, École Nationale Supérieure de l'Électronique et de ses Applications (ENSEA), Cergy, France, 95014, sadaf.ali, son.vu@ensea.fr; Surya Prakash, Indian Institute of Technology Indore, Indore, MP, India, 453552, surya@iiti.ac.in; Naoufel Werghi, C2PS & KUCARS, Khalifa University, P.O.Box: 127788, Abu Dhabi, UAE, Naoufel.Werghi@ku.ac.ae.

0360-0300/2022/9-ART $15.00
https://doi.org/10.1145/3560884



(a) (b) (c) (d)

Fig. 1. Different representations of 3D ear images. (a) Depth image, (b) point cloud, (c) polygon mesh, and (d) voxel.

non-intrusive, and the data acquisition requires less user cooperation. Recently, researchers developed many new biometric traits for human identification, including body odor, skull, back posterior posture, fingernails, toe print, and nose pores, with the ear being one of the most promising[1]. Additionally, independent embedded systems have been successfully prototyped to show the ear's reliability in real-time applications [90]. Besides the above advantages, Ianneralli's experiment [76] demonstrates that twins and triplets have distinct ears. The empirical evidence presented in [120] validates Ianneralli's assertion and demonstrates the potential role of the ear's performance in privacy preservation [33], kinship verification [44], identifying twins [120], infant recognition [115], and gender classification [100, 121, 168] where other biometrics such as face recognition fail [129]. Besides, we can also use ear biometrics as a supplement/multimodal with other biometrics to enhance security and recognition performance [42, 114]. For instance, ear can be combined with other biometric traits such as palm [54, 73, 108], palm and lips [31], fingerprint [66], face [77, 109, 110, 141, 142, 167, 168, 177], gait [122], finger knuckle [135], ECG and iris [137], fingerprint and iris [115], gait and face [94], palm vein [160], 3D face and finger [166], and 3D face [21, 80, 84, 156, 157]. A 3D ear image is an abstract representation of an ear and can be represented as depth, point cloud, polygon mesh, or voxel. Figure 1 shows the example of each of these ear representations. In literature, existing methods have used these representations to extract features and perform 3D ear recognition. However, each representation has its limits. In surface modeling methods, like mesh, the topological information (connectivity between the points) can be obtained, while in the point cloud, the data is unstructured, and the topological information is absent. The voxel image is a volumetric representation of each point where the change in volume affects the resolution of the 3D image.

In 3D ear recognition, significant works have been carried out using point cloud and depth images compared to the other two representations. Point clouds are widely used in ear matching since most ear recognition techniques rely on registration to obtain matching scores [26, 59, 63, 83, 130, 194]. Further, the depth image is a 2D image that provides the distance information of the object from a viewpoint. It is also widely used in 3D ear recognition, as 2D operations such as convolution can effectively extract features such as curvatures and edges to match ear pairs [185].

The ear recognition framework outlined in Figure 2 involves the following steps.

- **Enrollment**: This step enrolls a user in the recognition system. A camera or 3D scanner is used to capture the user's ear, and a template is constructed using the features derived from the ear image. The obtained user template is enrolled (stored) in the database to facilitate the recognition of the user in the future.

[1]https://www.bbc.com/future/article/20170109-the-seven-ways-you-are-totally-unique

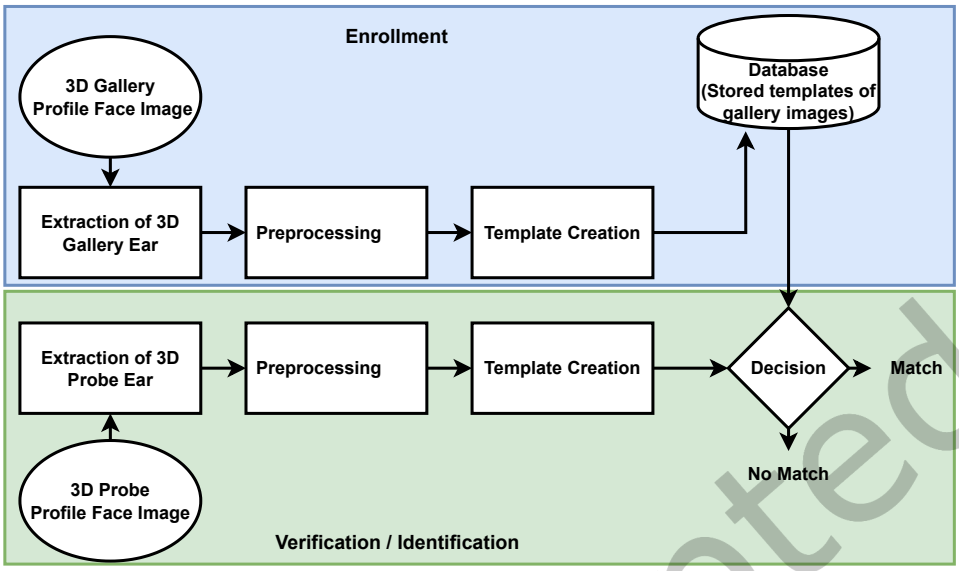
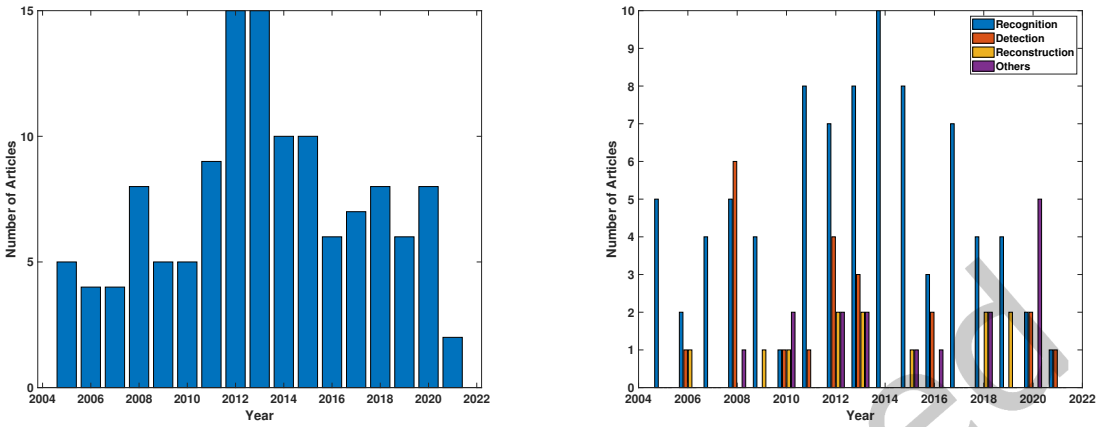


Fig. 2. Outline of enrollment and verification in 3D ear recognition.

- **Verification**: To create a user template, a token and biometric data are combined. Further, the token helps locate a unique template in the database compared to the actual user template to determine if the user is genuine. This procedure uses one-to-one matching.
- **Identification**: A user template is created with only biometric data and compared to all other templates in the database using one-to-many comparisons to determine the best match.

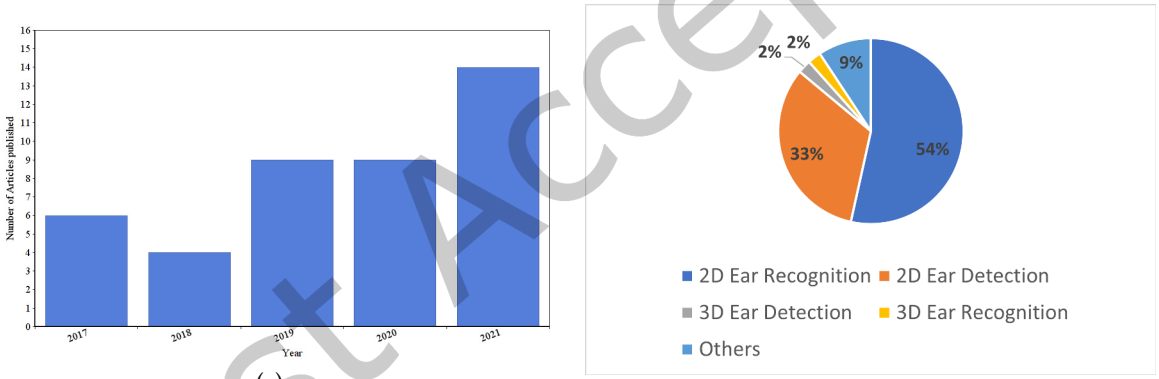
Overall, ear recognition is an intriguing area of study, and Figure 3 illustrates the impact of the ear on human recognition since 2005. We utilized dblp: computer science bibliography[2], Google Scholar[3], Scopus[4], and ACM digital library[5] to compile the list of published articles in the domain of 3D ear recognition. It is noted that there has been a slight decline in the number of publications in recent years, and one major reason for this is the scarcity of publicly available large databases of 3D ear images. Currently, only two major 3D ear databases (UND-J2 and IIT Indore) with a limited number of subjects are available for evaluation. However, since the field of 3D ear biometrics is promising and offers many good features (such as invariance to aging, expression, etc.) compared to the face, we anticipate a rise in the number of publications with large databases in the near future. Also, a competition is held worldwide to inspire researchers in the field of ear biometrics[6] [53], which will also lead to fruitful publications in the future. Recently deep learning-based approaches dominate the classical approaches; therefore, we included, as shown in Figure 4, separate publication statistics to show the impact of deep learning in ear recognition.

[2]https://dblp.org/
[3]https://scholar.google.com/
[4]https://www.scopus.com/
[5]https://dl.acm.org/
[6]http://awe.fri.uni-lj.si/competitions.html



(a) (b)

Fig. 3. Published articles since 2005 on 3D Ear. (a) The number of articles published per year, (b) categorizing published articles based on applications.



(a) (b)

Fig. 4. Since 2017, the impact of deep learning on ear biometrics (2D and 3D). (a) The number of articles published per year, and (b) the role of deep learning in different fields.

1.1 Differences from other related surveys

To the best of our knowledge, there has not been a systematic literature review on 3D ear recognition. A few existing ear recognition surveys are listed in Table 1. It contains long and short surveys, with the majority focusing on 2D ear recognition techniques and only a few on 3D ear recognition techniques. These surveys have been divided into three categories to highlight how they differed from ours.

(1) The first group [2, 51, 92, 128, 162] discusses ear detection, segmentation, and recognition techniques based on local and global features computed using classical and learning-based approaches. In [51], ear recognition in constraint and unconstrained environments, toolboxes for 2D ear recognition, as well as a thorough

Table 1. A summary of related survey papers in ear biometrics.

Year	Authors	Survey on	Domain	Category	Remarks
2012	Pflug *et al.* [128]	Ear	2D	Group - I	Long survey on 2D ear
2013	Abaza *et al.* [2]	Ear	2D/3D		Long survey focuses more on 2D ear
2017	Žiga *et al.* [51]	Ear	2D		Long survey on 2D ear
2021	Wang *et al.* [162]	Ear	2D		Long survey on 2D ear
2021	Kamboj *et al.* [92]	Ear	2D		Long survey on 2D ear and has a detailed review of deep learning techniques.
2008	Islam *et al.* [81]	Ear/Face	2D/3D	Group - II	Short survey on face & ear
2012	Islam *et al.* [82]	Ear/Face	2D/3D		Long survey on 3D face & ear
2018	Rajalakshmi *et al.* [135]	Ear/Finger knuckle	2D		Short survey on 2D ear & finger knuckle
2020	Yichao *et al.* [109]	Ear/Face	2D		Long survey on 2D face & ear
2007	Choras *et al.* [32]	Ear	2D	Group - III	Short survey on 2D ear
2009	Selvam and Rao [144]	Ear	2D		Short survey on 2D ear
2011	Li *et al.* [178]	Ear	2D		Short survey on 2D ear
2014	Singh *et al.* [147]	Ear	2D/3D		Short survey focuses more on 2D ear
2020	Pallavi *et al.* [150]	Ear	2D		Short survey on 2D ear

description of evaluation protocols are given, and in [92], a dedicated survey on deep learning-based techniques is discussed. However, each survey includes only a brief discussion on 3D ear recognition.

(2) The second group discusses multimodal recognition using ear and other biometric traits. A review of 3D face and ear recognition techniques is presented in [81, 82]; however, it does not include the most recent 3D ear recognition techniques. Similarly, multimodal 2D recognition using ear, face, and finger knuckle and the impact of noise and occlusion is discussed in [109, 135]. Overall, this group lacks a focus on 3D ear recognition techniques, and their review scope is slightly broader than ours.

(3) The third group, short surveys, also focused more on 2D than the 3D ear [32, 144, 147, 150, 178].

1.2 Motivation

Ear recognition is a rapidly growing field in biometrics, where the 3D ear has significant performance advantages over 2D. We are especially interested in the 3D ear for the following reasons.

(1) The geometric shape of 3D ear models overcomes the challenges, viewpoint, and illumination mentioned above.

(2) The richness of geometrical properties in 3D has high discriminating features and is superior to its 2D counterpart.

(3) 3D ear can be a better choice to analyze the open issues such as the age-invariant nature of the ear, symmetry property of the ear, and the inheritance nature of the ear.

1.3 Contributions

This survey is the first comprehensive and systematic review of 3D ear recognition techniques for the biometric research community. We hope this survey will help the research community understand major challenges and drawbacks in the state-of-the-art methods and future research opportunities. In summary, the following are the main contributions:

(1) The available datasets and data preprocessing are discussed in detail.

(2) A unified way of categorizing recognition techniques consolidates 3D ear recognition techniques as registration-based and feature-based. The benefits and drawbacks of each method are discussed in depth.

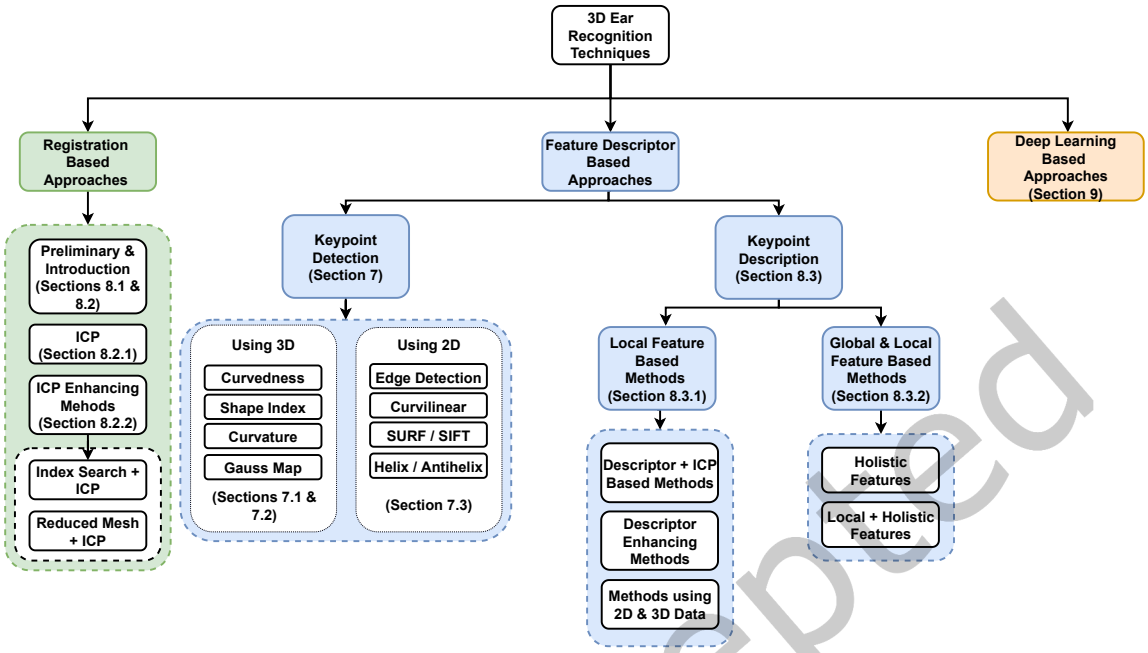


Fig. 5. Categories of approaches in 3D ear-based human recognition.

(3) Visual results demonstrating qualitative analysis of a few 3D ear keypoint detection and recognition methods are presented.
(4) For completeness, a few techniques on the 2D ear have been presented in the context of co-registered 2D and 3D ears.
(5) The current issues in 3D ear recognition approaches and the open problems as future research directions are discussed.

1.4 Structure of this survey

The following is the structure of this survey. Section 2 gives a hierarchical overview of recognition techniques, and Section 3 discusses data acquisition techniques which detail the methods used in 3D ear data collection. Section 4 discusses data preprocessing, which includes ear cropping from 3D profile images, Section 5 provides a full summary of 3D ear datasets, and Section 6 details the metrics used to evaluate recognition techniques. Section 7 discusses 3D keypoint detection techniques, and Section 8 provides a detailed survey of 3D ear recognition techniques. A separate Section 9 is included on learning-based techniques in 3D ear recognition to underline the importance of deep learning. Finally, the article addresses the potential challenges and opportunities in 3D ear recognition in Section 10, followed by conclusions in Section 11.

2 Taxonomy

The taxonomy of different 3D ear recognition techniques is presented in Figure 5 hierarchically. It encompasses recognition strategies based on registration, feature descriptor, and deep learning. We first discuss the registration-based strategy in detail and explain how a registration algorithm determines the match between an ear pair. Further, we present existing techniques for boosting the convergence of registration algorithms. Following

(a) (b)

Fig. 6. 3D scanners. (a) Minolta vivid 910[7], and (b) Artec[8].

that, we describe feature-based techniques and present how the similarity of two ear images is computed using the extracted features. Additionally, we discuss how feature-based techniques are more efficient than registration-based techniques. Finally, we discuss deep learning-based techniques for 3D ear recognition.

3 Data acquisition

To capture the shape of the ear in 3D, commonly used are non-contact 3D laser scanners. The datasets [174], [26] in the literature used Minolta Vivid[7] 910 scanners to collect 3D ears, while a few others used Artec scanners[8] [62]. Figure 6 depicts both scanners. Vivid 910 scanners are expensive; however, they are more precise than many non-contact digitizers. On the other hand, Artec is less costly than Minolta and slightly less accurate. However, the Artec scanner is convenient and could be a good choice for capturing small objects such as ears in portability. In the Minolta scanner, the lower window emits laser light on the subject, while the upper window receives the returning light from the subject. A minor drawback of laser-based scanners is that users may be hesitant to participate in data collection due to exposure to laser light. 3D commercial scanners are expensive, so it is challenging to utilize in real-time. Few techniques have generated 3D ear data from 2D ear data using a 3D reconstruction approach to address the issue. However, these techniques generate volumetric images, and the image resolution is determined by the size of each voxel, which impacts recognition accuracy. It also necessitates a series of 2D ear images and calibration to obtain accurate 3D ear data [12]. As a result, researchers in [183] built a custom scanner that collects data by laser triangulation. The data acquisition process is similar to that of Minolta, generating 2D and 3D data. Figure 7 illustrates the custom scanner in which laser light is emitted onto a 2D ear, caught by a CCD camera and then sent to a computer to build a 3D ear point cloud. However, the method for capturing ears in 3D varies depending on the dataset. In most situations, data collection is accomplished by placing a cap over the subject's head, as the sensor cannot sense the hair surfaces. A few 3D acquisition methods for obtaining various biometric traits are outlined in [182], and more focus on ear data acquisition is detailed in [40, 105, 106, 119].

[7]https://www.konicaminolta.com/
[8]https://www.artec3d.com/

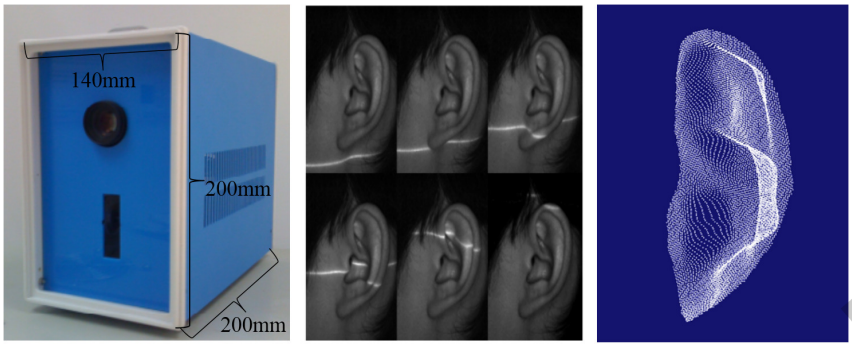


Fig. 7. 3D custom scanner and the generated 3D ear data. (a) Custom scanner, (b) laser light projected on the ear, and (c) generated 3D ear point cloud [183].

4 Data preprocessing

Data preprocessing includes noise removal and hole filling that occurs during data collection, as well as cropping the ear from the face profile image.

4.1 Noise and holes removal

Spikes and holes are the two common errors in the 3D ear data due to subjects' oily skin and sensor, where holes are more frequent than spikes [26]. At the time of data acquisition, the inner part of the ear, which is not visible to the sensor, contributes to the holes. Vertex-based anisotropic diffusion [186] is used to remove spikes [59, 64] as it preserves geometric information while smoothing. Mean, or median filters are commonly used to fill the holes, and a few other works use interpolation to fill holes [26, 171].

4.2 Normalization

Another step in preprocessing is to pay attention to the size of the ear sample. Though the size is independent of the subject's distance from the camera, at the time of data acquisition, a shift in the distance between the subject and the camera causes the irregular distribution of points in the x and y planes. Thus, the distance between x and y must be normalized to obtain a standard ear sample.

4.3 3D ear detection

Recognition performance in an ear biometrics system is highly dependent on the performance of the ear detection module. The trade-off between system complexity and false-positive detection is a crucial factor, as system complexity is proportional to the detection of zero false positives.

4.3.1 Classical approaches Most of the 3D ear dataset available is in the form of profile face images, and it has to be preprocessed to crop the ear for further study. However, cropping 3D ears from 3D profile face images is challenging. Many datasets employ manual cropping for better data quality, as accurate detection and extraction are critical for subsequent recognition. Traditionally, 3D ear detection systems relied heavily on 2D ear images to detect ears from 3D profile face images. We discuss a simple method to crop the ear from 3D profile images, despite mentioning detection and segmentation as open issues in Section 10. Since object detection algorithms in 2D domains are highly efficient at detecting objects in a scene, we can train any detector to detect an ear from a 2D profile face image. Given that each 3D model has its co-registered 2D image, a trained detector can be

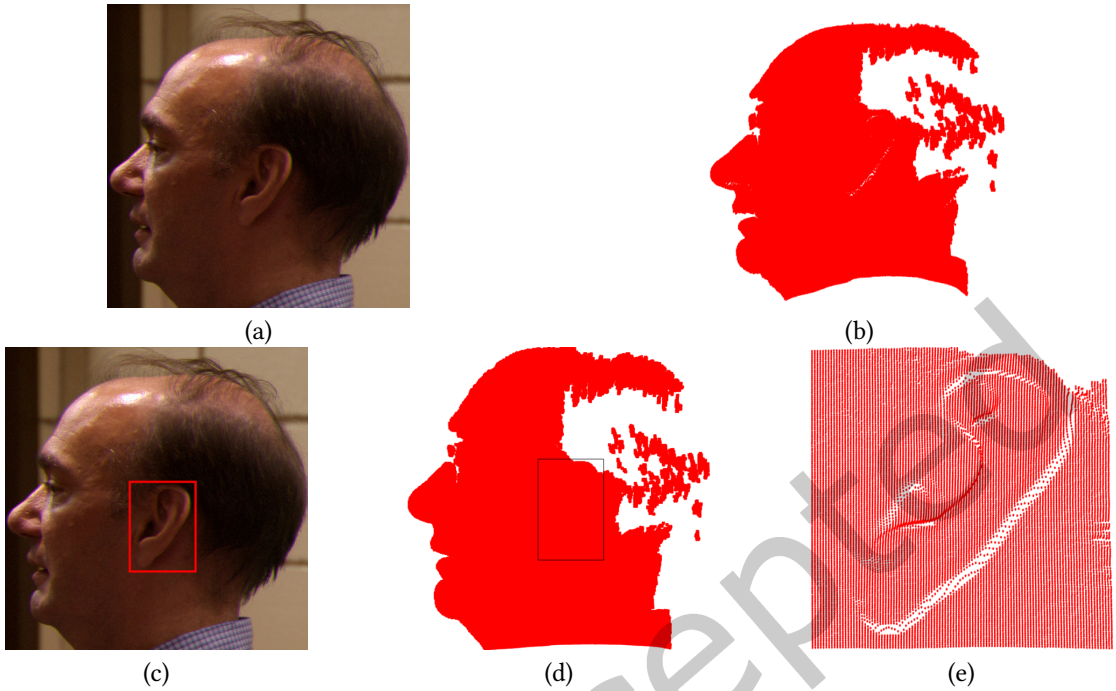


Fig. 8. Cropping of ear from a 3D profile face using a 2D co-registered image. (a) 2D profile face image, (b) co-registered 3D profile face image, (c) 2D ear detection (highlighted in red bounding box), (d) mapping of detected bounding box coordinates from (c) onto 3D ear, and (e) cropped 3D ear [174].

applied to the co-registered 2D images to detect the ear. Consequently, we can map the 2D coordinates of the detected ear to 3D to crop the ear from 3D profile face images [26, 79]. This method is valid for datasets with co-registered images. It is worth noting that points in 3D images that are not mapped to 2D images should be removed carefully before proceeding. The other method is orientation-dependent, where the distance between the nose and ear pit is utilized to crop the ear. We can empirically determine the nose and ear pit distance using a few training profile face images; however, the orientation of all profile images is not assured. An outline of ear detection in 3D using 2D images is demonstrated in Figure 8, where Figures 8(a & b) show the co-registered 2D and 3D profile face image, and Figure 8(c) shows the bounding box ear detection in 2D. The detected coordinates are mapped to the 3D profile image as shown in Figure 8(d) to crop 3D ear as demonstrated in Figure 8(e). Though the majority of techniques use co-registered images in ear detection, a technique is available using a classical approach that can be applied directly to the 3D profile image to detect the 3D ear [102].

4.3.2 Deep learning based approaches These are more promising than classical ones in detecting ears in a complex unconstrained environment [97]; therefore, ear detection in 2D profile images can be performed using any deep learning-based techniques [36, 52, 93]. Similarly, a few techniques in the literature can be applied directly on 3D profile images to localize ears [116, 117, 194]. EpNet [116] is one of the recent techniques in 3D ear detection that uses PointNet [132] as the backbone. Its architecture is shown in Figure 9, where the network accepts point cloud data as input, computes local features at each point and then computes a global feature for the entire 3D ear using a max-pooling operation. Finally, a multi-layer perceptron (MLP) network combines the local and global features to generate output vectors. Further, this technique has employed a statistical model [126] to generate synthetic

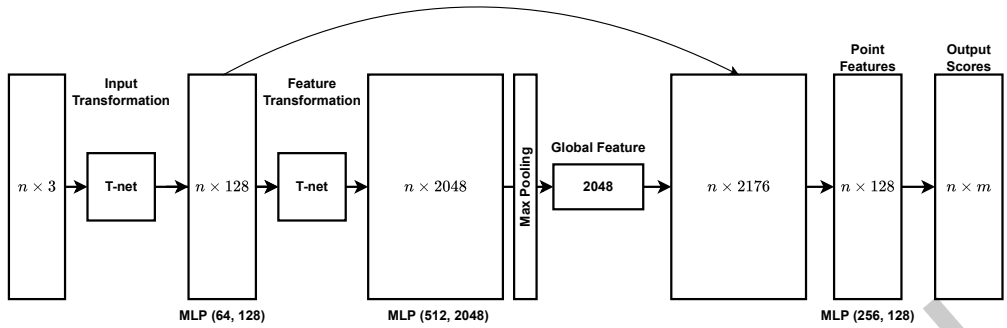


Fig. 9. The architecture of EpNet for 3D ear detection in point cloud data. The T-net and MLP in the architecture stand for transformation network and multi-layer perceptron, respectively [116].

Fig. 10. Visual results of ear detection for EpNet on the UND-J2 dataset where the points highlighted in black represent the ear region [116].

training data as there was a lack of data for handling the data annotation. This technique makes it possible to use the model parameters to determine the ear component automatically. The data is normalized before training so that all coordinate values fall between 0 and 1. After training the technique on synthetic data, it has been evaluated on the UND-J2 dataset to check its performance on a real dataset, and Figure 10 illustrates a few ear detection results for the technique. The other technique presented in [117] is similar to [116], where PointNet++ [133] is used as the backbone architecture. Similarly in [194], PointNet++ [133] is used to segment the ear.

Table 2. Summary of 3D ear datasets.

No.	Dataset Name	Data Format	Acquisition Device	# Subjects	# Images
1	UND-J2 [26]	Range Images (480 x 640), Co-registered 2D	Minolta Vivid 910	415	1800
2	IIT Indore [62]	Point cloud, Mesh	Artec-Eva	Phase-1 - 188 Phase-2 - 176 Phase-3 - 188	1509 1380 1478
3	UCR [27]	Range Images (200 x 200), Co-registered 2D	Minolta Vivid 300	155	902
4	Ear Parotic [107]	Range Images (140 x 200)	Custom Device	250	2000

5 Datasets

This section discusses the various 3D ear datasets that are freely available or require a license agreement. The available 3D ear datasets are summarized in Table 2. A few 3D ear datasets include 2D co-registered images; however, this is not true for all datasets. Moreover, most datasets include profile face images, and 3D ears must be cropped from them. In a few other datasets, West Virginia University (WVU) and University of Miami (UM) generated 3D ear images from video using shape from shading and structure from motion approaches with 48 and 13 subjects. We have not included separate sections for these datasets due to the lack of available information [23].

5.1 UND dataset

The University of Notre Dame, Collection-J2 (UND-J2) dataset[9] is the largest publicly accessible ear dataset. It also provides separate UND-E, UND-F, and UND-G ear datasets, the subsets of the UND-J2 dataset. UND-E has 114 subjects with 464 samples, UND-F has 302 subjects with 942 samples, UND-G has 235 subjects with 738 samples, and UND-J2 has 415 subjects with 1800 samples. All subsets have co-registered 2D ear images except UND-E, and a few samples of the 3D ear from the UND-J2 dataset are shown in Figure 11. The images in the dataset were collected by Minolta Vivid 910 3D scanner in two sessions, maintaining a time gap of 17 weeks between the two sessions during data collection. The scanner captures data in 0.3 seconds in fine resolution with a precision of $\pm 0.008mm$ and accuracy of $\pm 0.10mm$. It is observed that the images in the dataset are affected by pose variations, scaling, and occlusion due to earrings and hair. The dataset is available as profile face images, which would require additional cropping to obtain ear. However, the availability of co-registered 2D and 3D images in the dataset enables us to simplify the cropping process to a certain degree.

[9] https://cvrl.nd.edu/projects/data/

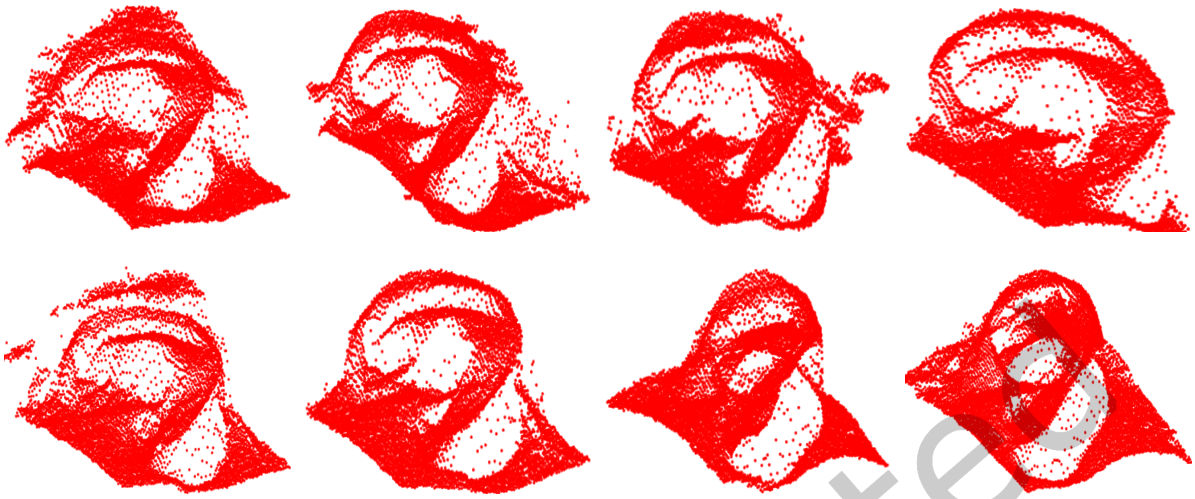


Fig. 11. A few 3D ear samples from the UND-J2 dataset [174].

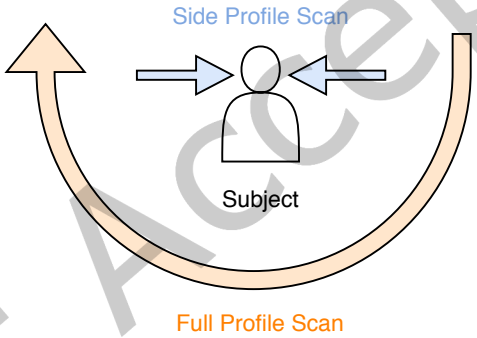


Fig. 12. Setup of full and side profile scanning [40].

5.2 IIT Indore dataset

Indian Institute of Technology Indore[10](IIT Indore) 3D ear dataset is our in-house dataset collected using Artec-Eva® scanner [62]. The scanner provides high resolution and easy handling without using any additional equipment. It captures data with a resolution of $\pm 0.5mm$ and accuracy of $\pm 0.10mm$. The scans are acquired using a wig cap to avoid occlusion and unwanted reflection from the subject's hair. We also ensure that there is no other reflective surface in the vicinity of the subject while capturing the data. The scanner encounters difficulties owing to the optical properties of the eyes and teeth and captures eyes and teeth (if visible) geometrically inward, resulting in holes. Figure 12 shows the side profile and full profile scan setup. The side profile scan contains right and left ear images in 3D, and the complete profile scan includes face, right, and left ear images. Samples of a few processed, scanned images are shown in Figure 13.

[10]http://iiti.ac.in/people/surya/research/IIT_Indore_Ear_Database/index.html

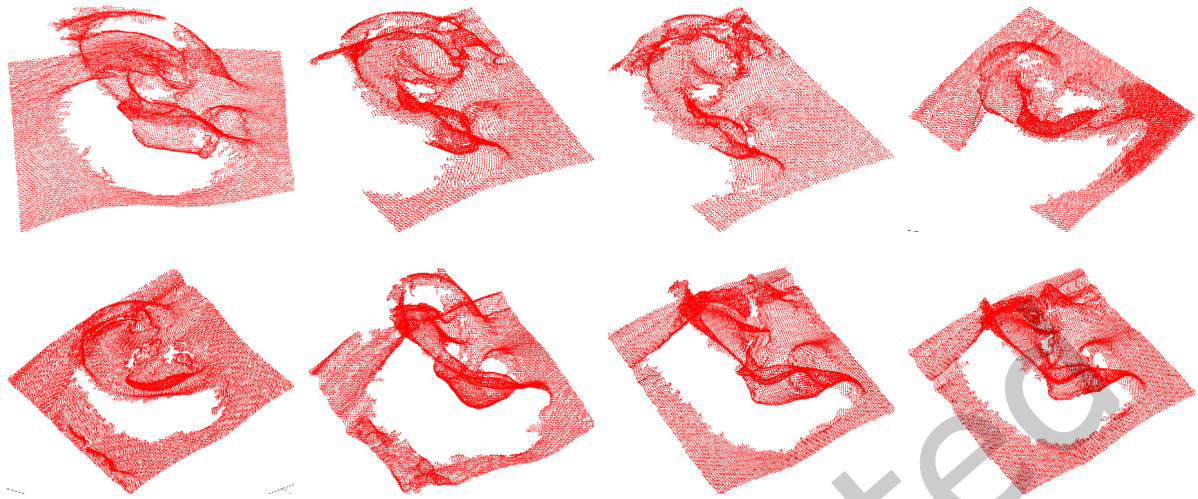


Fig. 13. A few 3D ear samples from the IIT Indore dataset [62].

The ear images have been scanned in normal mode (geometric + texture), and the dataset currently consists of scans in two formats, .ply and .asc. Scans can be visualized, processed, and transformed using the software included with the Artec Eva 3D scanner and open-source software such as MeshLab[11] [35] and CloudCompare[12] [65]. The ear images have been sequentially numbered (from right to left ear) for each subject with an integer identification number, *e.g.,* 1000xx, where 1000 represents the subject number, and xx represents the sample number. The IIT Indore dataset is collected in three phases, approximately one year between two subsequent phases. The data collected in Phase-1, phase-2, and phase-3 has 188 subjects, 176 subjects, and 188 subjects, respectively, with at least three 2D & three 3D ear samples; however, 2D ear images are not co-registered. Among these subjects, 110 subjects are common in all three phases, and the dataset is available in both raw and cropped formats.

5.3 UCR dataset

The University of California Riverside (UCR) dataset[13] has been acquired using the Minolta Vivid 300 3D scanner. The scanner provides range images and co-registered 2D color images. The dataset consists of 155 subjects with 902 samples. The samples from each subject are collected on the same day in three different poses, *viz.* left, right, and front. The dataset contains images with pose variations and occlusions due to hair and earrings.

5.4 Ear parotic face angle

This dataset contains 2000 ear samples from 250 subjects ranging in age from 20 to 60 [107]. The data was collected using a custom laser scanner with a $\pm 0.5mm$ accuracy, and the designed scanner was much cheaper than available scanners in the market. The ear data was collected in two sessions at a time gap of one month. Before data collection, subjects were asked to remove ornaments, to acquire clean data free of occlusions.

[11] https://www.meshlab.net/
[12] https://www.danielgm.net/cc
[13] https://vislab.ucr.edu/RESEARCH/sample_research/Biometrics/ear.php

6 Evaluation Metrics

The following is a collection of assessment methods to evaluate the performance of an ear recognition technique.

- **False Rejection Rate (FRR):** The percentage of times a technique incorrectly classifies a genuine user as an imposter.
- **False Acceptance Rate (FAR):** The percentage of times an imposter is identified as a genuine user.
- **Equal Error Rate (EER):** It is the value when FRR and FAR both have the same value.
- **Genuine Acceptance Rate (GAR):** It represents the proportion of times a technique correctly recognizes an authentic user.
- **Receiver Operating Characteristics (ROC) Curve:** A probability curve shows how much a biometric system is capable of distinguishing the subjects. It plots FAR vs. GAR values where GAR is defined as $(100 - FRR)\%$.
- **Decidability Index (DI):** It measures the separation of the mean of intra-class (similarity scores) and inter-class (dissimilarity scores) probability distributions.
- **Rank-*k*:** It is used to evaluate a biometric system's performance and indicates the frequency with which a correct sample occurs within the top k matches.

Verification accuracy of a recognition technique is used to measure its verification performance and is defined as $Accuracy = (100 - \frac{FAR+FRR}{2})\%$. The optimal combination of FAR and FRR for any recognition technique provides the highest verification accuracy.

7 Feature keypoint extraction

Due to the large number of points on a 3D ear, it is difficult to consider all points during recognition. For example, using a complete point cloud of a 3D ear in registration-based techniques may increase matching time due to slow convergence. To avoid this, in many techniques, keypoints are used to efficiently represent the 3D ear [43, 179]. In many features-based and registration-based ear recognition techniques [27, 59, 130, 190], detecting these keypoints is the initial step in matching. This section presents prominent techniques for detecting the keypoints in a 3D ear, viz., local surface variation (LSV) based, shape index (SI) based, and curvedness-based techniques. This section reproduces and summarizes the visual outcomes of a few techniques to provide readers with a qualitative overview. Table 3 summarizes the techniques used to detect keypoints in a 3D ear.

7.1 Techniques based on curvature

The shape index (SI) is a curvature-based measure for detecting keypoints [28]. SI at a point P is defined as

$$SI(P) = \frac{1}{2} - \frac{1}{\pi} tan^{-1} \frac{k_1(P) + k_2(P)}{k_1(P) - k_2(P)} \tag{1}$$

where k_1 and k_2 are the maximum and minimum principal curvatures, and $k_1 > k_2 \; \forall P$. Formally, k_1 and k_2 are given as follows.

$$\begin{aligned} k_1 &= H(P) + \sqrt{H^2(P) - K(P)} \\ k_2 &= H(P) - \sqrt{H^2(P) - K(P)} \end{aligned} \tag{2}$$

where $H(P)$ and $K(P)$ are the mean and Gaussian curvatures at the point P. For a given ear image, SI is computed for a local region within a window of size $m \times m$. If the window's center point's SI is an extremum compared to all its neighbours, it is classified as a keypoint. The value of SI is also helpful in categorizing the nature of the shapes. A range of SI values and corresponding shape categories are listed in Table 4. The SI values computed at each vertex of a 3D ear model using Equation 1 are illustrated in Figure 14. The color scheme indicates the

Table 3. A summary of 3D ear feature keypoint detection techniques.

No.	Reference	Approach based on	Data Type
1	Chen *et al.* [28]	Curvature	Range Image
2	Chen and Bhanu [26]	Curvature, Edge detection	Range Image
3	Chen and Bhanu [27]	Curvature	Range Image
4	Zeng *et al.* [179]	Curvature	point cloud
5	Zhou *et al.* [190]	LSV	Range Image
6	Islam *et al.* [83]	LSV	Range Image
7	Islam *et al.* [84]	LSV	Range Image
8	Sun *et al.* [152]	Curvature	Range Image
9	Prakash and Gupta [130]	SURF	Range Image
10	Maity *et al.* [111]	Curvature	Range Image
11	Zhang *et al.* [188]	LSV	Range Image
12	Ganapathi *et al.* [59]	LSV	Point cloud
13	Ganapathi *et al.* [64]	Curvilinear	Point cloud
14	Ganapathi *et al.* [61]	Gauss map	Point cloud
15	Zhu *et al.* [193]	Curvature	Range Image

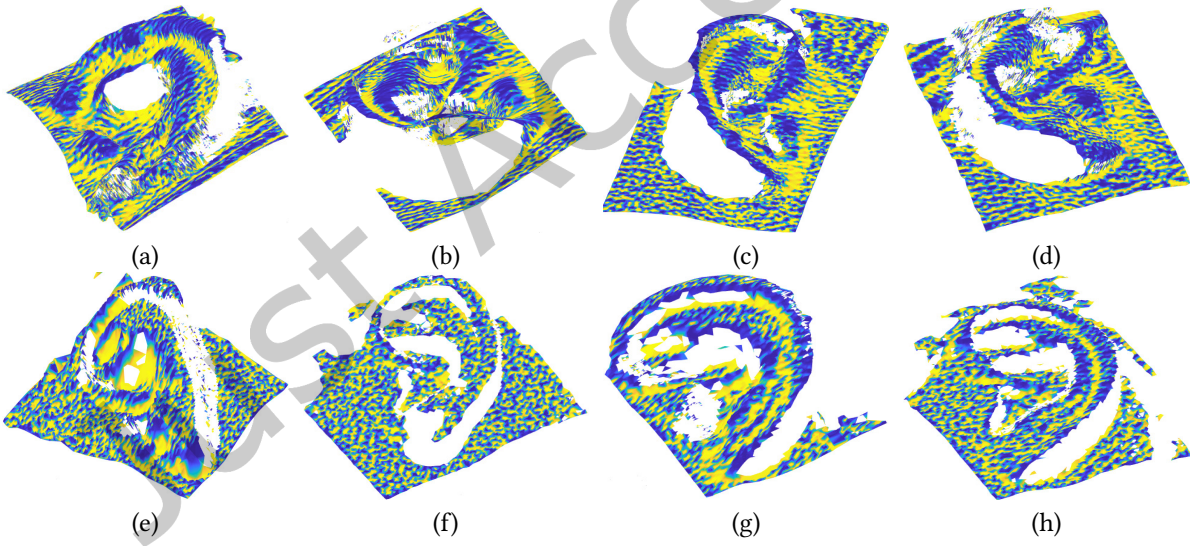


Fig. 14. SI demonstration, with the largest and smallest values represented in yellow and blue, respectively. (a)-(d) Ear samples from the IIT Indore dataset, and (e)-(h) ear samples from the UND-J2 dataset.

increasing order of the SI values from blue to yellow. Filtering the obtained SI values enables the detection of high-quality keypoints using a predefined threshold value where an SI value less than the threshold is eliminated. For instance, points with an SI value greater than a threshold of 0.75, considered as keypoints are highlighted in

Table 4. Shape classification based on SI values [26].

No.	SI	Shape category
1	[0, 1/16)	spherical cup
2	[1/16,3/16)	trough
3	[3/16,5/16)	rut
4	[5/16, 7/16)	saddle rut
5	[7/16, 9/16)	saddle
6	[9/16, 11/16)	saddle ridge
7	[11/16, 13/16)	ridge
8	[13/16, 15/16)	dome
9	[15/16, 1]	spherical cap

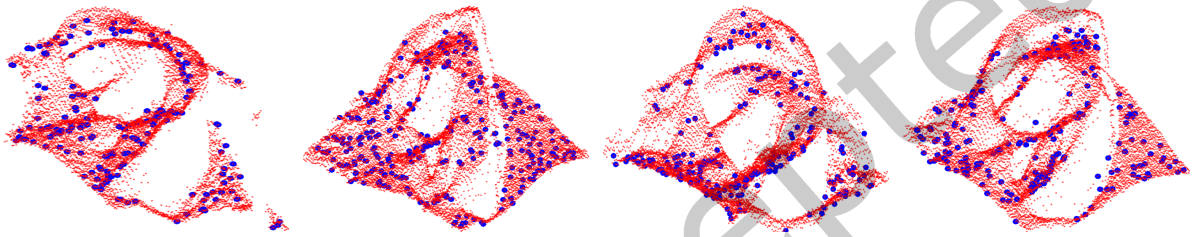


Fig. 15. A few examples of keypoints detected on UND-J2 dataset ear samples using the SI are shown in blue.

blue, as shown in Figure 15. Keypoints have also been detected using another measure called curvedness, $C_{cur}(P)$, which is also based on k_1 and k_2. Formally, $C_{cur}(P) = \sqrt{\frac{k_1^2+k_2^2}{2}}$, where the value of $C_{cur}(P)$ describes the nature of the surface at point P and is found to be invariant to rotation and scaling. If a selected keypoint has neighbours belonging to the boundary points, the keypoints are usually discarded.

7.2 Techniques based on surface variation

This section discusses keypoint detection using surface variations in a neighbourhood at each point P for a fixed radius. Further, these neighbours are used to construct a covariance matrix C defined as $C = \sum_{i=1}^{N}(p_i - P)(p_i - P)^T$, where p_i is one of the N neighbours of point P that lie in a chosen radius. Further, eigenvalues $\lambda_1, \lambda_2, \lambda_3$ where $\lambda_1 > \lambda_2 > \lambda_3$ and eigenvectors V_1, V_2, V_3 are obtained from C [59, 190]. A keypoint is defined as any vertex point that satisfies $\frac{\lambda_3}{\lambda_1+\lambda_2+\lambda_3} > t_1$ and $\frac{\lambda_1}{\lambda_1+\lambda_2+\lambda_3} < t_2$. For predefined thresholds, the detected keypoints are shown in Figure 16, where most of the keypoints (highlighted in black) are found in curved regions, and very few are found in flat areas. Though the surface variations-based approach performs better at detecting keypoints, it produces false positive keypoints in the presence of noise. Figure 17 depicts instances in which keypoints are detected on flat regions of the ears (left upper corner of the image) due to noise. A technique for finding keypoints using Gaussian map clustering is proposed in [61]. This technique is inspired by [163], where a cluster is generated on a unit sphere using the normals at point P and its neighbour. Based on the number of clusters generated, a point P is classified as a keypoint. Figure 18 shows an example of Gauss map clustering for two points present on flat and edge regions produces a different number of clusters, two for flat and four for edge regions. Sun *et al.* [152] proposed a method for locating key points based on a saliency value calculated at each point using a Gaussian

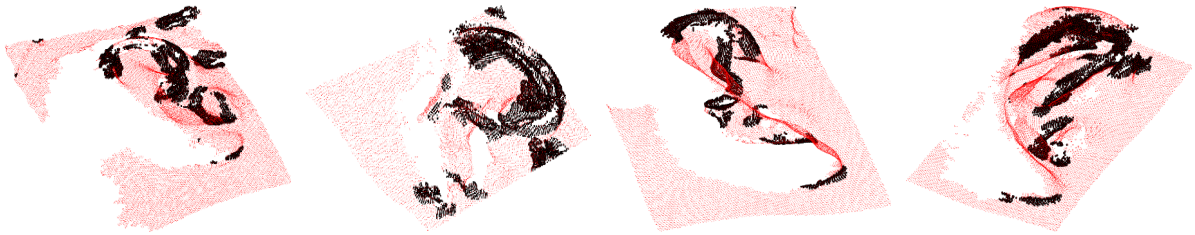


Fig. 16. Keypoint detection using surface variation approach on samples from the IIT Indore dataset. The detected keypoints are highlighted in black.

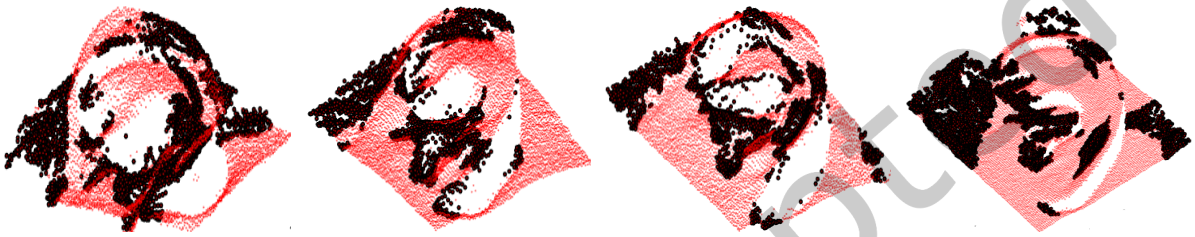


Fig. 17. The effect of noise on keypoint detection using surface variation approach on samples from the UND-J2 dataset. The detected keypoints are highlighted in black.

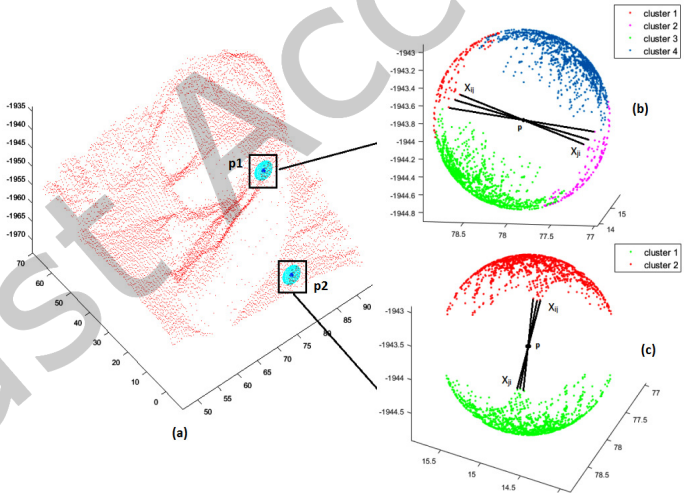
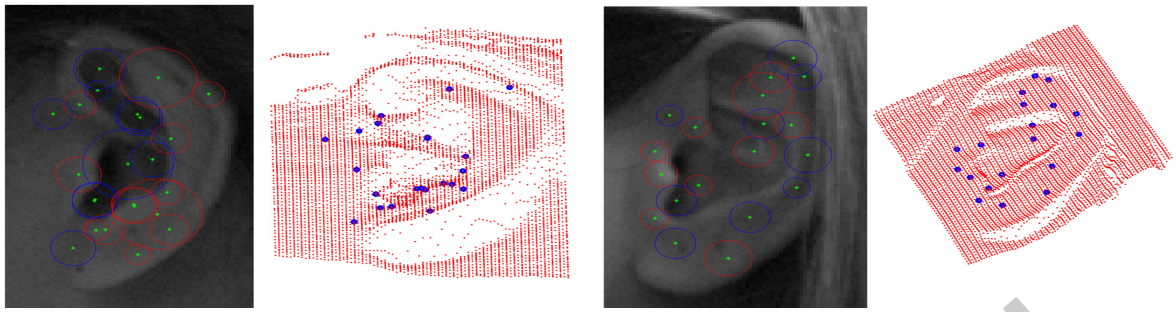


Fig. 18. Keypoint detection using Gauss map clustering. (a) 3D ear data with two selected points p_1 located on the edge and p_2 on the flat region, (b) Gauss map shows two pairs of opposite clusters for a point on the edge region, and (c) Gauss map shows a pair of opposite clusters for a point on the flat region [61].

weighted average of mean curvature values. The saliency values are further filtered via point disk sampling to verify the robustness of the keypoints.



(a) (b)

Fig. 19. Keypoint detection in a 3D ear using a co-registered 2D ear image. (a) & (b) Examples of keypoints detected on ear samples from the UND-J2 dataset where the gray image is the co-registered 2D ear of the 3D ear image. The green points indicate keypoints detected in the 2D domain, whereas the blue points are keypoints generated by mapping the keypoints of the 2D ear onto a co-registered 3D ear [130].

(a) (b) (c)

Fig. 20. Mapping 2D curvilinear structure (feature) points onto a co-registered 3D ear image. (a) Detected curvilinear features in 2D, (b) thinned curvilinear features (red) superimposed on the original 2D ear, and (c) mapped curvilinear features (red) onto 3D [64].

7.3 Techniques using 2D co-registered ear image

Keypoint detection is performed using 2D co-registered images of the 3D ear. Since each 2D ear point corresponds to a 3D ear point, 2D images are utilized to locate keypoints in 3D images. Thus, in literature, popular 2D keypoint detection techniques such as speeded up robust features (SURF) [16] are used to locate keypoints in 2D, which are then mapped to 3D. In [130], the keypoints in the 2D ear images are detected using SURF [16]. Further, the detected keypoints are mapped onto the co-registered 3D ear images. A demonstration is shown in Figure 19, where the figure on the left depicts detected keypoints on a 2D ear image and the figure on the right illustrates mapped keypoints on a co-registered 3D ear image. Similarly, in [64], the technique detects the curvilinear structure in a 2D ear image using a 1D polynomial filter with varying width and rotations. The obtained binary image and the detected curvilinear structure are shown in Figures 20(a & b). These points serve as keypoints for the 2D ear image, and they are subsequently mapped onto a co-registered 3D ear image to locate keypoints in the 3D ear image, as shown in Figure 20(c). In [63], a detailed 3D keypoint detection is demonstrated using

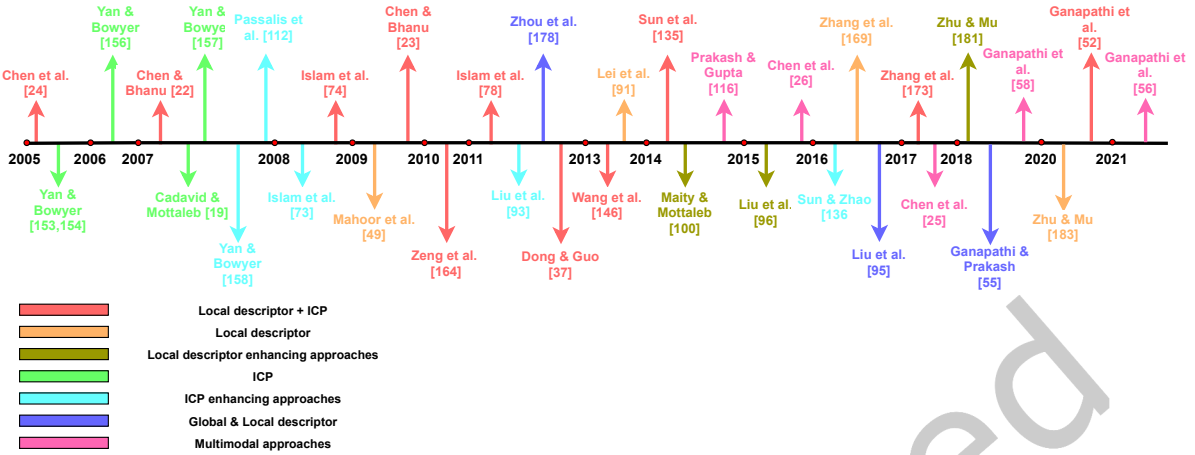


Fig. 21. Chronological overview of the most relevant 3D ear recognition techniques.

different 2D keypoint detectors. However, the techniques that employ co-registered 2D ear images to detect keypoints in 3D ear images face a few challenges. For instance, 2D ear images that are less illuminated or have non-frontal ears may produce erroneous keypoints. Moreover, co-registered 2D images for 3D ear images may not be available in some scenarios.

7.4 Summary

(1) Surface variations, curvatures, and curvedness features can be used to locate keypoints in a 3D ear. The majority of techniques reviewed so far have relied on one of these properties to detect keypoints. Although these properties effectively detect keypoints in the 3D ear, they are susceptible to noise and may detect false keypoints.
(2) A second approach uses a co-registered 2D image to locate the keypoint. This approach does have a few disadvantages. As previously mentioned, a co-registered image of the 3D ear is not always possible. Additionally, variations in lighting and posture in 2D can affect the detection of keypoints in 2D images.

8 Ear recognition

This section discusses recognition techniques that use feature descriptor vectors and registration. First, we discuss registration-based techniques, their drawbacks in terms of convergence, and available methods for improving their performance. We then discuss feature descriptors and their discriminative power in recognition, including the local and global descriptors. A chronological overview of the most relevant 3D ear recognition techniques is given in Figure 21.

8.1 Preliminary: Iterative closest point

Let probe P and gallery G be finite sets with m and n points where $P \in \mathcal{R}^{m\times 3}$ and $G \in \mathcal{R}^{n\times 3}$. For every point in set P, the nearest neighbour is calculated in set G. These points are known as correspondence points, and a translation and a rotational matrix are determined using these points to minimize the registration error E.

$$E(\phi, t) = \sum_{i=1}^{m} \sum_{j=1}^{n} ||p_i - (\phi(g_i) + t)||^2 \tag{3}$$

where $p_i \in P$ and $g_i \in G$. The matrices ϕ and t are computed using a covariance matrix $H = \bar{P}\bar{G}^T$, where $\bar{P}$ and $\bar{G}$ are the mean subtracted matrices. Further, by decomposing the matrix $H = U\Lambda V^T$, a rotation matrix $\phi = VU^T$ and a translation matrix $t = P - \phi(G)$ are computed. These matrices are used in iteration to compute registration errors until the error obtained exceeds a predefined threshold.

8.2 Recognition using registration

This section discusses techniques for 3D ear recognition that focus on rigid where two ear point clouds are fitted together using a sequence of rotation and translation operations. However, in a non-rigid transformation, scaling of the point cloud is used in addition to the above operations. In existing ear recognition techniques, such as [170, 171, 173–176], rigid transformation is used and is mainly achieved by the iterative closest point algorithm (ICP) [17]. The ICP algorithm computes a sequence of transformations that register two ear point clouds (G, P) by finding the closest point in one point cloud to the other point cloud, called correspondence points. The registration error between ear point clouds is used to calculate its matching score, with a lower error indicating a better match. ICP algorithm is accurate in registering ear point clouds; however, it does necessitate a coarse alignment of two ear point clouds to ensure convergence. Also, the computational cost of registration increases with the size of the point cloud. However, if the ICP algorithm uses the *k-d* tree to find the closest point in the gallery, the search time reduces to $O(logN)$, where N is the number of points in the gallery image.

8.2.1 Independent ICP This section discusses techniques that use only the ICP algorithm to match 3D ear. A large-scale ear experiment is conducted in [171] for recognition using depth images and point clouds. An edge-based approach and an ICP algorithm are utilized to perform ear matching. Their findings indicate that an ICP-based approach is more stable and scalable to large datasets. They also presented a technique in [170], an extension of [171] with data refinements by removing false correspondence matches. Similarly, Yan and Bowyer proposed an automatic ear recognition based on ICP in [173, 174], whose primary contribution is automatic ear extraction from profile faces and matching 3D ears using the ICP algorithm. They also introduced another technique based on the ICP for 3D ear recognition that analyses the effect of posture and sub-sampling on registration and the use of point-to-point correspondences and point-to-triangle correspondences in finding the closest points [176]. However, as mentioned previously, all of the above approaches require an initial alignment between ear pairs to perform effectively. In [23], an ICP-based ear recognition technique using video frame sequences is proposed. This technique utilized structure from motion (which utilizes multiple images) and shape from shading (which utilizes a single image) to reconstruct 3D images. This method necessitates several images for reconstruction, and the process is time-consuming. A few examples of ear registration using ICP are shown in Figure 22. The probe ear (blue) is registered with ear samples from the gallery (red) to find the best match, and the error obtained is used as the matching score. The registration error obtained using the same subject can be stored as intra-class matching scores and otherwise stored as inter-class matching scores to compute ROC and EER.

8.2.2 ICP-enhancing techniques To minimize computations and speed up ICP, [175] employs an indexed search approach where a fixed volume is generated and subdivided into voxel elements. Further, each vertex (x, y, z) in a given 3D ear is mapped to the nearest voxel (X, Y, Z) in the generated fixed volume. The index of the mapped vertex point is stored using a data structure. To match a probe image, first, it is indexed using the same data structure, and further, the closest point in the gallery for each point in the probe is computed. This approach has enhanced the registration process by reducing the convergence time. However, it is challenging to fix the voxel size because it affects the resolution and is directly proportional to the data loss. Similarly, to speed up the ICP matching process, a coarse alignment between the ears is performed using reduced mesh structures, and the obtained transformation is further used to fine-tune the mesh structure as a whole [79]. This

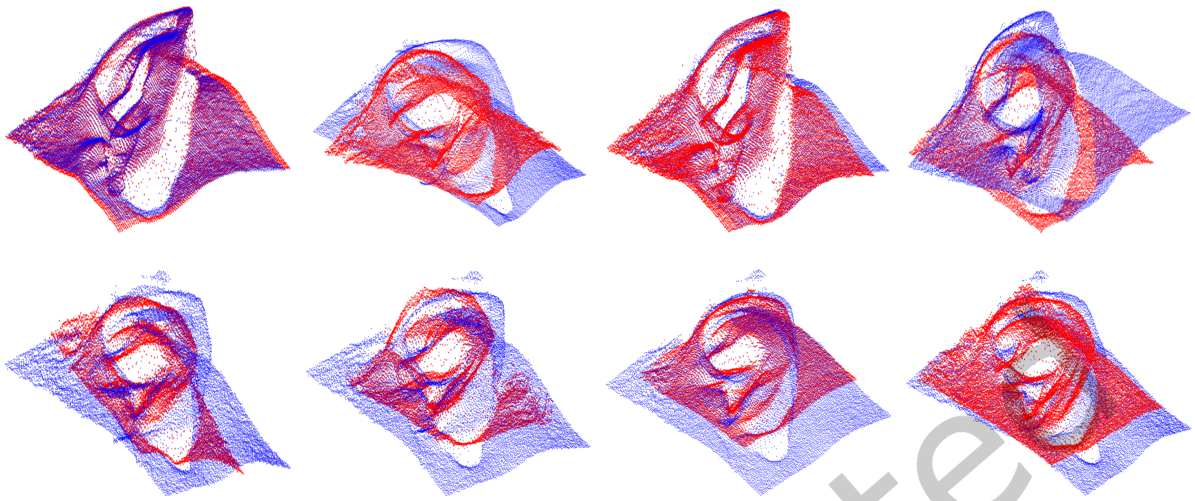


Fig. 22. A few examples of registration of probe (blue) and gallery (red) images using the ICP algorithm [17].

Table 5. Summary of registration-based methods in 3D ear recognition.

No.	Technique	dataset	Data	Gallery, Probe	Recognition Rate (%)	Remarks
1	Islam *et al.* [79] [2008]	UND[a]	2D/3D	300, 300	93.98	It uses reduced mesh with modified ICP
2	Cadavid *et al.* [23] [2007]	WVU[b] + UM[c]	Video	61, 25	84.00	Highly suitable for real-time applications
3	Yan & Bowyer [175] [2007]	UND	3D	369, 369	98.70	It uses pre-computed voxel nearest neighbour
4	Yan & Bowyer [174] [2007]	UND	2D/3D	415, 1386	97.80	Robust to partial occlusion
5	Yan & Bowyer [173] [2006]	UND	2D/3D	415, 415	97.60	Automated ear recognition uses 2D for detection and 3D for recognition
6	Yan *et al.* [176] [2005]	UND	2D/3D	302, 302	98.80	Investigates variants of ICP
7	Yan & Bowyer [171] [2005]	UND	2D/3D	404, 404	97.50	Robust to increase in gallery size
8	Yan & Bowyer [170] [2005]	UND	2D/3D	302, 302	84.10	Robust to increase in gallery size

[a] university of Notre Dame, [b] West Virginia University, [c] University of Miami

technique's performance is evaluated using single- and two-step ICP algorithms. In both cases, the rate of rank-1 recognition remains constant. On the other hand, a coarse mesh structure is equivalent to choosing random points as correspondence points, and due to this, the rough alignment may not always be the same. As a result, the registration error and time can differ. In [153], a cascaded ICP is used to avoid convergence into local minima, leading to faster algorithm convergence than the standard ICP. However, the process is time-consuming in a cascaded setting, where the rigid transformation computed for one level serves as the starting point for the next level. Table 5 summarizes the performance of available registration-based 3D ear recognition techniques with a clear indication of the number of gallery and probe samples.

8.3 Recognition using feature descriptor

Global and local features are used in feature-based 3D ear recognition systems. Global feature-based methods generate feature descriptors for the entire 3D ear. On the other hand, a local feature descriptor is constructed by encoding the geometrical information of neighbourhood of a point. As a result, local feature descriptors are

immune to occlusion and are frequently preferred over global descriptor-based approaches in 3D recognition. Local features are also advantageous in ear recognition since they help detect correspondence between probe and gallery ear images. Most ear recognition approaches that use local features involve a two-step registration process. The first step involves coarse registration using local features, followed by fine registration with the help of ICP. This section examines ear recognition techniques that utilize global and local features.

8.3.1 Techniques using local descriptors In [101], a shape-based keypoint descriptor is introduced where the surface around a keypoint is divided equally in the radial and longitudinal dimensions to compute the descriptor. Further, a feature vector is constructed by concatenating the mean curvedness and SI of the vertices in each partition. Since this technique relies on curvedness and SI, both are prone to noise, which may be affected in the presence of noise. A technique described in [56] is attempted to recognize a 3D ear using a generalized 3D descriptor, the SHOT [140]. However, many additional 3D descriptors have been introduced recently [59, 72, 187], and therefore there is a strong possibility that the performance may improve by utilizing one among them. In [185], supervised dictionary learning is presented for ear matching, which divides the input ear image into blocks and encodes each block as a histogram using surface type information. Finally, the feature vector is constructed by concatenating the histograms of all blocks. However, there is no practical tool for determining the image block size. Further, knowing the block numbers for ears of different scales and occluded images is also challenging. Chen and Bhanu proposed a technique in [25] using helix and antihelix structures with two-step ICP. The helix and antihelix structure is used to find a coarse alignment, followed by ICP to find a fine alignment to match an ear pair. Further, in [28], Chen *et al.* proposed a local surface patch (LSP) descriptor. It is computed using geometrical surface information, including normal and curvature. It comprises surface type, centroid, and a 2D histogram were used to compute the correspondence points for coarse alignment of an ear pair. Then, a fine alignment is achieved using an ICP algorithm. Using LSP, they also propose another recognition technique [26] that employs two approaches; the first uses the helix and antihelix structure of the ear to find correspondence points, and the second use LSP to determine correspondence points between the 3D ears being matched. Further, the probe and gallery are aligned coarsely using point correspondences and then a modified ICP algorithm. Both techniques use computationally expensive features of higher dimensional [26, 28]. To address this, they proposed another technique in [27], which involves mapping higher-dimensional features to a lower-dimensional space; these reduced feature vectors are used to compute a coarse alignment, followed by an ICP alignment.

Similarly, in [180], LSP is used as a local descriptor to locate keypoints and is then fed into a modified ICP to match an ear pair. Similar to helix/antihelix, another technique uses the auricle structure of the ear for coarse alignment, and ICP for fine alignment [161]. However, as discussed, detecting auricle structure may be challenging in the presence of noise and occlusions. Islam *et al.* [83, 84] introduced a local feature descriptor where the keypoint's neighbourhood is cropped and used as a feature descriptor. They rely on PCA-derived local coordinates and change the local axes of the cropped surface if the viewpoint varies significantly. Likewise, in [152, 154], a feature vector is generated by fitting a quadratic surface using [195] to the neighbours of each keypoint. Further, the fitted surface is used to construct a feature descriptor vector. Smoothing is involved in this technique and may lead to loss of geometric information. In [179] a 3D Center-Symmetric Local Binary Pattern (3D CS-LBP) is introduced. At each keypoint, a range image is generated by projecting the neighbours with the help of local axes. Further, 3D CS-LBP is applied to the range image to derive local features, which is used in matching an ear pair. In [59], a descriptor based on geometric statistics is introduced where the descriptor is constructed in three phases at each key point, based on the distribution of neighbours and the normal angles of neighbours in annular regions of concentric spheres, and the distance of neighbours from a plane perpendicular to the keypoint's normal. This technique finds correspondence points using the computed features and then finds the match using ICP. Dong *et al.* [43] generated depth images by rotating a 3D point cloud around local axes and used a 2D descriptor to compute feature vectors for the obtained depth images to match the 3D ear. As the original image's dimension

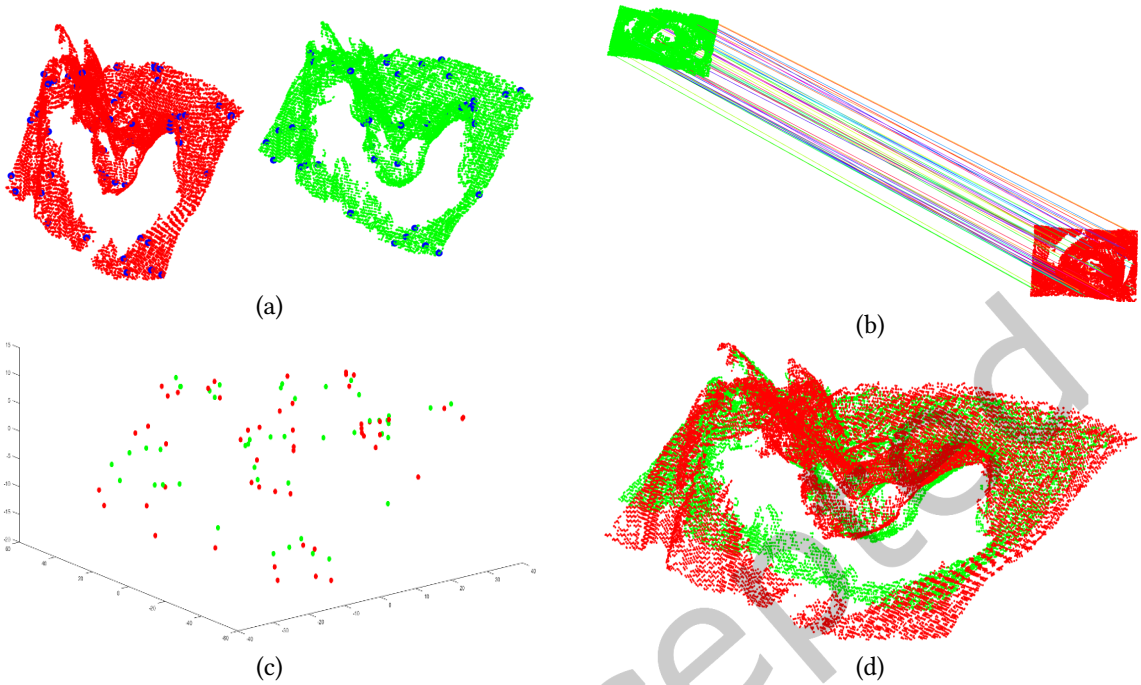


Fig. 23. Outline of 3D ear recognition using local descriptors and ICP [59]. (a) Keypoints detection (blue highlights) on probe and gallery images, (b) local feature computed at each keypoint in probe and gallery image is matched to find the correspondence points, (c) coarse alignment using correspondence points, and (d) ICP algorithm on the entire ear pairs for fine registration.

is reduced, information may be lost while transforming to depth images. Figure 23 depicts a visual example of the contribution of local feature descriptor and ICP in ear matching. First, keypoints (highlighted in blue) are detected in both images, followed by obtaining correspondences between the probe and gallery images using the local features computed at each keypoint to get an initial transformation. Finally, the ICP algorithm is utilized to obtain a fine registration.

8.3.1.1 Performance enhancing approaches using local descriptors A fast index-based ear recognition method to reduce the gallery search space by finding the closest match among all gallery images for a given probe ear image in a constant search time is proposed in [107, 111]. The technique presented in [111] extracts feature vectors of each keypoint using the method described in [190], and the primary objective is to reduce the time required for matching, which is accomplished by using a balanced and unbalanced data split. The technique claims a 50% reduction in search space, which has decreased matching time. Similarly, [107] proposed an ear parotic face angle to reduce the search space. The ear parotic face angle is calculated using the angle between the normals of two planes, the ear and face planes. Further, the obtained angle is used to index each ear image in the gallery. To match a probe ear, the ear parotic face angle of the probe ear is computed and searched for the closest angle match in the gallery images. It retains the matching time constant and scalable to large datasets. In [192], a global index search-based technique matches ears. They used ear shapes to create the KD tree for the gallery set. Consequently, an indexed-based search lists the gallery images that match the probe image. Since the listed

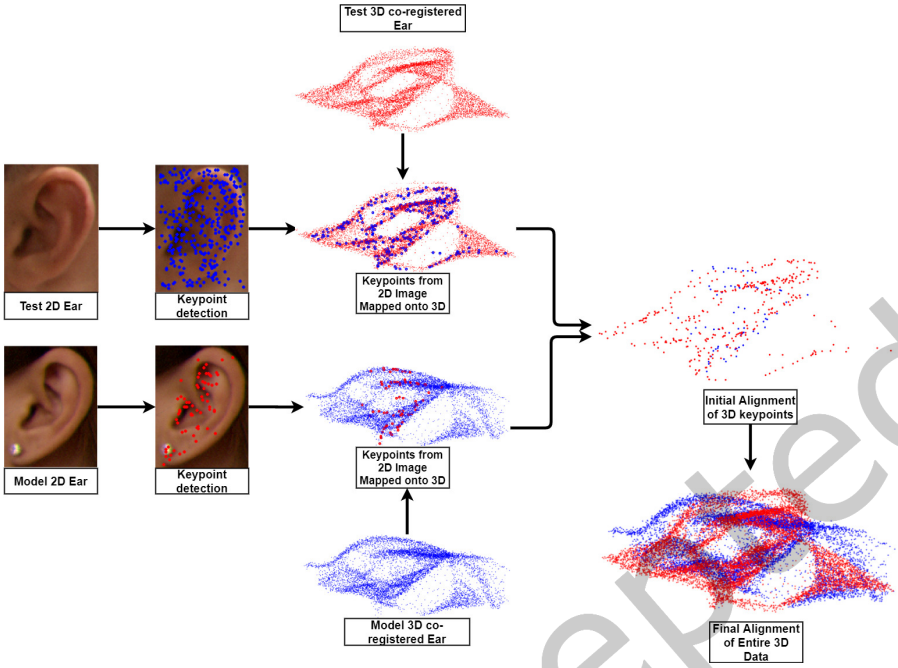


Fig. 24. Ear recognition using 3D and co-registered 2D ear images [63].

images are small in relation to the total number of images in the gallery, the search process is accelerated, and performance has improved. In [188], Zhang *et al.* presented a technique that employs the ICP-LSV algorithm to avoid coarse to fine registration to achieve fast and accurate convergence. Local surface variations are used to remove the flat non-ear data to generate normalized and refined ear data. The convergence during the registration process has improved because the undesired portion has been deleted from the refined ear data. However, the method is noise-prone, and it may be difficult to remove non-ear data in the presence of noise. Flat structures like the ear lobe, which are important for gender identity [113], may be removed as non-ear data cause data loss and degrades the recognition performance.

8.3.1.2 Techniques using 2D and 3D ears This section reviews the techniques that use 3D ear images and co-registered 2D images for ear recognition. In literature, 2D and 3D ears have been associated differently, such as in detection, recognition, and segmentation. An outline of the role of 2D and 3D ears in recognition is illustrated in Figure 24.

(1) Several works employ a multimodal biometrics approach, where 2D and 3D ears are used as different modalities [29, 172]. These studies demonstrate that fusing the matching scores obtained from 2D ear images and 3D ear models can improve recognition performance.
(2) A few other works have used 2D ear images to extract feature keypoints from 3D ear images. The reason is that 2D keypoint detection algorithms are more mature than 3D keypoint detection algorithms, which are still in their early stage. For instance, in [63, 64, 130], keypoints detected on co-registered 2D ear images were used to compute feature keypoints in corresponding 3D ear images.

(3) There are a few other works, such as [26, 79, 173, 174], where co-registered 2D ear images are used to detect and segment ears in 3D profile images.

Yan and Bowyer examined the ear's recognition performance in [172] using multimodal, multi-algorithm, and multi-instance approaches. This technique is similar to [170] and [171], except it includes a detailed 2D and 3D data analysis and algorithms. They conclude that the multimodal approach performs better than the other two: PCA on 2D ear images and ICP on 3D ear images. Prakash and Gupta [130] detected feature keypoints in a 2D image and then mapped them onto a co-registered 3D image. The mapped keypoints are used to align two 3D ears coarsely, and then a fine alignment is obtained using a modified ICP algorithm. Ganapathi *et al.* [64] presented a technique to detect feature keypoints in 2D ear images using curvilinear structures. These structures are used as keypoints in 2D and are later mapped onto the co-registered 3D ear image to get the feature keypoints. Further, each keypoint in the 3D ear is described using a descriptor [68], and the obtained feature descriptor vectors are then used to match an ear pair. Texture and depth scale-invariant feature transform (TDSIFT), a modified SIFT descriptor proposed in [29], where texture information from a 2D ear with depth information from a 3D ear is combined to create a descriptor. Ear pairs are matched using these descriptor vectors. Similarly, [30] proposed a sparse dictionary-based descriptor that uses 2D texture information and 3D shape information to construct a descriptor. Overall, ear data from 2D and 3D domains are required to implement the discussed approaches. However, there is no guarantee that 3D ear data and its co-registered 2D ear data will always be available in real-time. Table 6 summarizes the performance of local feature-based 3D ear recognition techniques with a clear indication of the number of gallery and probe samples.

8.3.2 Techniques using global and local descriptors Global features are computed using the entire 3D ear model. Although ICP-based approaches are discussed separately in Section 8.2; we can also classify them as global since they use the entire image of the probe and gallery ears for registration. Passalis *et al.* created a unique global signature for an ear using an annotated ear model (AEM). The data is fitted with this model, and an image is extracted. Further, signatures are created from the extracted image and utilized to match an ear pair [125, 157]. In [104], a matching algorithm based on a global sliced curve is computed using principal components. These curves are compared using the nearest neighbour classifier and show a five-fold reduction in matching time compared to ICP matching. However, in contrast to ICP, recognition efficiency has deteriorated slightly. Global features are generally less effective in matching ears than local features. On the other hand, global features are more robust to noise than local features. Consequently, global features are often combined with local descriptors to enhance performance when noise and occlusion are present [62, 190].

Zhou *et al.* [190, 191] introduced a match score fusion of local and holistic features for 3D ear recognition. A local surface is cropped and divided into four sub-regions for a predefined radius to compute the local descriptor at each keypoint. Using the SPHIS, each sub-region is encoded as a feature vector and concatenated to construct a final feature descriptor vector. Further, 3D ears are converted into voxels and arranged as a single-column vector to construct a global descriptor. The global features are matched using cosine similarities to get a matching score. Since the feature descriptor uses voxelization as a global feature, changes in voxel resolution may impact the recognition performance. Likewise, 3D ear recognition is performed in [106] using a combination of global features, such as empty centers and angles, and local features, using points, lines, and areas. Normalized coordinates are used to calculate these features, however, the coordinates are not stable in the presence of occlusions and noise, and it may affect the performance of the descriptor. Similarly, [62] proposed a technique based on weighted combinations of four 3D local descriptors [68, 98, 139, 140] and a global descriptor. This technique has enhanced recognition performance by combining the discriminative strength of each local descriptor. The global descriptor, local sphere geometry pattern (LSGP), is based on histograms, similar to the local binary pattern [123] is used along with the local descriptors to match ears. Further, methods for 3D ear recognition based on local joint structural (LJS), a modified version of surface patch histogram of indexed shape (SPHIS) [190], is introduced in

Table 6. Summary of local feature-based methods in 3D ear recognition.

No.	Technique	dataset	Data	Gallery, Probe	Recognition Rate (%)	Remarks
1	Ganapathi *et al.* [59] [2020]	UND-J2	3D	404, 1376	98.60	Robust to noise and occlusion
2	Ganapathi *et al.* [58] [2019]	UND-J2	3D	404, 1376	98.00	Recognition using a hybrid descriptor
3	Zhu & Mu [192] [2018]	UND-J2	3D	415, 415	98.50	Reduces the search space
4	Ganapathi *et al.* [64] [2018]	UND-J2	2D/3D	404, 1376	98.69	3D ear recognition using 2D ear
5	Zhang *et al.* [188] [2017]	UND-J2	2D/3D	415, 415	98.55	ICP convergence has improved
6	Chen *et al.* [29] [2017]	UND-J2	2D/3D	415, 415	95.90	Modified SIFT is used
7	Chen *et al.* [30] [2015]	UND-J2	3D	415, 415	96.40	Dictionary based hybrid descriptor
8	Sun *et al.* [154] [2015]	UND-J2	3D	415, 415	89.61	Graph-based recognition technique
9	Maity *et al.* [111] [2015]	UND-J2	3D	415, 415	98.50	Improved matching time
10	Zhang *et al.* [185] [2015]	UND-J2	3D	468, 168	100.00	Suitable for large-scale data
11	Zeng *et al.* [180] [2014]	UND-J2	3D	100, 100	94.00	Reduces the iteration time of ICP
12	Prakash & Gupta [130] [2014]	UND-J2	2D/3D	404, 1376	98.30	Enhanced performance using modified ICP
13	Sun *et al.* [152] [2014]	UND-J2	3D	415, 415	95.10	Reduces the matching time
14	Wang & Mu [161] [2013]	UND-J2	3D	415, 415	97.59	Robust to pose variation
15	Lei *et al.* [101] [2013]	UND-J2	3D	415, 415	97.40	Captures macro-shape patterns
16	Islam *et al.* [83] [2011]	UND-F UND-J2	2D/3D	302, 302 415, 415	95.40 93.50	3D ear detection using 2D ear images and recognition using 3D ear images
17	Dong & Guo *et al.* [43] [2011]	UND-J2	3D	415, 415	98.87	SIFT based matching
18	Zeng *et al.* [179] [2010]	UND-J2	3D	415, 415	96.39	3D local binary pattern
19	Chen & Banu [27] [2009]	UND-F UCR	3D	302, 302 155, 155	96.70 94.90	Low dimension embedding-based ear matching
20	Chen & Bhanu [26] [2007]	UCR UND-G	3D	155, 155 302, 302	96.77 96.36	Robust to rigid transformation
21	Chen *et al.* [28] [2005]	UCR	3D	52, 52	90.40	Robust to noise
22	Chen & Bhanu [25] [2005]	UCR	3D	30, 30	93.30	Introduced a two-step ICP for ear matching

[193, 194]. SPHIS is a shape-indexed dependent descriptor that is not robust to noise, so the effectiveness of this technique in the presence of noise is unknown. Further, voxelization of 3D ear is used as the global feature and uses game-theoretic matching to compute the similarity of ear pairs. Table 7 summarizes the performance of global and local feature-based 3D ear recognition techniques with a clear indication of the number of gallery and probe samples.

8.4 Summary

(1) ICP-based registration is a widely used technique for matching ears. Variants of ICP are available in the literature; however, the majority of the techniques have used the base version of ICP.
(2) The ICP algorithm is highly accurate and performs well in recognition; however, the time required to compare a pair of 3D ears increases as the number of vertices in the 3D ear data increases. To address this, a few approaches are introduced to speed up the ICP algorithm, and it has been observed that these methods significantly reduce the amount of time necessary for matching.
(3) Feature-based techniques can also be used to accelerate the matching process. Combining a feature-based technique with ICP results in a more optimal matching if it can accurately locate correspondence points.
(4) 2D and 3D co-registered ear data are employed, with 2D data primarily used to detect keypoints in 3D.

Table 7. Summary of global and/or local feature-based 3D ear recognition techniques.

No.	Technique	dataset	Data	Gallery, Probe	Recognition Rate (%)	Remarks
1	Zhu and Mu [194] [2020]	UND-J2	3D	415, -	98.82	SPHIS based descriptor, ICP is not used in matching phases (global and local)
2	Zhu and Mu [193] [2018]	UND-J2	3D	-, -	98.80	Robust to occlusion (global and local)
3	Ganapathi *et al.* [62] [2018]	UND-J2 IIT-Indore[a]	3D	404, 1376 70, 350	98.69 98.90	Fusion of global and local scores in matching (global and local)
4	Zhang *et al.* [106] [2016]	In-house	3D	500, 1500	-	A fast approach to match 3D ear (global and local)
5	Zhou *et al.* [191] [2012]	UND-J2	3D	415, 1386	98.00	Robust to pose variations (global and local)
6	Zhou *et al.* [190] [2011]	UND-J2	3D	415, 415	98.60	Robust to pose variations and fast in matching ear pairs (global and local)
7	Theoharis *et al.* [157] [2008]	UND - F & G	3D	324, 324	95.00	It uses the annotated deformable model, and a signature is constructed for ear matching (global).
8	Passalis *et al.* [125] [2007]	UND + Inhouse	3D	525, 506	94.40	It uses annotated ear model (global)

[a] Indian Institute of Technology Indore

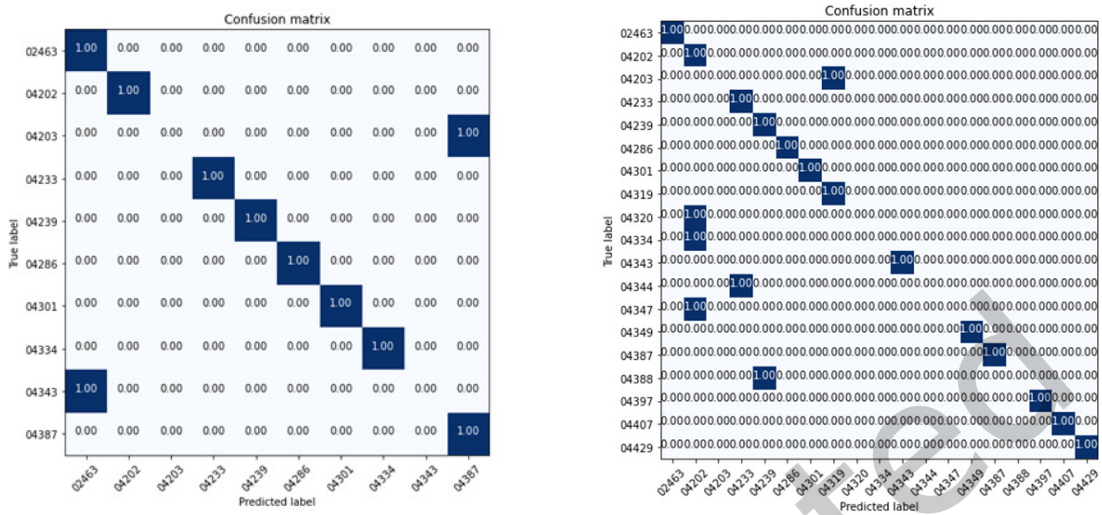
9 Recognition using deep learning

Compared to classical features, learning-based features demonstrate better performance and lead to the publication of several works in ear detection and recognition utilizing deep learning [48, 57, 92]. However, the lack of datasets and the structure of the 3D data is challenging and has become an obstacle to a deep neural network to attain reasonable accuracy compared to traditional approaches. Nevertheless, there have been very few deep learning-based approaches in 3D ear recognition, and given the importance of deep learning, we have dedicated this section to discussing a few available techniques. In [58], an auto-encoder is used to extract features from a 3D ear for recognition. In this technique, 3D ears are transformed into depth images, and each image is fed to a deep auto-encoder to compute a reduced feature. Depth images are obtained at each keypoint by projecting the neighbouring points onto a set of computed local axes. Despite using a learning-based approach to extract 3D ear descriptors, the features obtained are from projected 2D images. Though the direct application of deep learning to the 3D ear is lacking, we evaluated performance using the seminal recognition 3D network, PointNet [132], on two small sets of 3D ear subjects, choosing two groups of ten and nineteen subjects at random for the experiment. Further, an ear sample from a subject is used as a probe, while other samples are used as a gallery. For 100 epochs, we achieved an accuracy of 80% for the first set and 68.40% for the second set. Figure 25 depicts the obtained results as a confusion matrix. The results show that the performance has been dramatically lowered with a slight scaling of the data volume. Although preprocessing and augmentation can increase deep learning performance, an extensive dataset is necessary to achieve the performance of state-of-the-art classical techniques.

10 Future research directions

Despite the voluminous work developed in ear biometrics, there are still vital issues, open problems, and future research directions, which we shed some light on in this section.

- **Ear detection and segmentation** from profile face images remain an open issue in 3D ear biometrics. There are techniques in the literature for segmenting the ear in 3D profile face images using co-registered 2D profile face images [79]. For instance, in [174], the nose pit is used as a reference point to locate the ear, and in [25, 26], helix and antihelix ear structures are used to locate the ear. These techniques rely on 2D co-registered images to detect the ear in 3D. These techniques face challenges where pose changes may affect the nose pit of the reference subject, and as a result, it cannot always be used as a valid reference point for 3D ear segmentation [174]. Similarly, the detection of helix/antihelix structures is based on edge detection methods that may become unstable in the presence of noise [26]. As a result, novel 3D



(a) (b)

Fig. 25. Performance of PointNet on two sets of data constructed from UND-J2. (a) Confusion matrix for the first set, and (b) confusion matrix for the second set.

ear detection approaches are required. Additionally, all existing detection algorithms are valid on images collected in constrained environments; thus, there is a need to investigate methods for 3D ear detection with little or no controlled environments. Apart from these detection approaches, only a few approaches are described in the literature that can detect the ear in 3D without using co-registered 2D images [116, 118]. Therefore, it is essential to examine ear localization in the 3D domain to construct reliable 3D ear detectors.

- **Age invariant nature of the ear** is still a subject of debate, with limited scientific studies. A few scientific articles discuss the aging effect and its consequences [55]. However, these works confine their examination of the age-invariant to a short period. For instance, the study published in [78] utilized a dataset collected over eleven months to examine the recognition efficiency in month-to-month changes in the ear; however, such a limited time frame is insufficient to demonstrate the age-invariant nature of the human ear. The dataset must be collected over a more extended period in multiple phases with a minimum of one year between consecutive steps to determine the ear's age-invariant nature. Additionally, the focus is to track the ear changes due to age alone; therefore, the data used in the analysis should be free from the pose, lighting, and scale variations. Hence, 3D images are suitable since they are relatively free of these influences. As a result, analyzing the age invariance of the human ear would be more approachable using 3D data.
- **Inheritance** is an exciting area of study in ear biometrics, and according to Iannarelli's study [76], it may inherit a few ear features. This argument, however, should be thoroughly investigated because this could pose a serious threat to ear-based biometric systems by jeopardizing the ear's uniqueness. It is essential to create an extensive ear dataset comprised of multiple individuals of various families to conduct a study on this topic.
- **Morphable models (MM)** is an effective mathematical technique for describing 3D shapes. It provides a useful encoding and prior statistical distribution for the shape and texture of the 3D data. New data samples can be generated by exploiting the space spanned by the shape and texture of the training data. In [39], [38],

a dataset[14] comprising 3D morphable models (3DMM) of the head, face, and ear is introduced, and in [21] morphable models of face profile and ear using noisy, incomplete, and occluded data are created. However, A large dataset with annotation is required to build a 3DMM. Therefore, 3DMM should be explored in greater depth to develop different sets of datasets that can be used for gender identification and symmetry analysis and to study the effect of inheritance and age invariance on ear biometrics.

- **Deep learning** paradigm gained popularity and wide deployment in several biometric modalities [4, 18, 89, 155, 159]; however, its usage in 3D ear biometrics is still in its infancy due to several reasons. One of the most significant issues is the scarcity of samples per subject in the datasets used to train the network. Current 3D ear datasets contain a small number of samples, usually two or three per subject. This issue can be addressed using data augmentation and synthetic data. However, augmenting point cloud data of world coordinates is challenging. Leveraging the power of a deep network, a recent work using a scene representation network [149] addresses this problem where any test data could be rendered for unseen views [127]. Further, synthetic data may be another solution, but its resemblance to actual data remains unclear. Due to this, despite having several ear recognition techniques in 2D using deep learning [5, 8–11, 19, 49, 50, 53, 91, 95, 96, 124, 134, 136, 143, 146, 148, 151, 184], only limited techniques available for recognition in 3D. Considering the large arsenal of 3D recognition techniques [69], we believe there is enormous potential in leveraging 3D recognition deep learning models through domain adaptation and meta-transfer learning [74]. This paradigm can also help address the scalability problem. Additionally, deep learning-based point cloud registration could be used more efficiently to register ear images than traditional registration techniques. As demonstrated in [34], registration based on deep learning has a lower computing cost than classical approaches. Similarly, 3D point cloud completion [164], and upsampling [103] from deep learning may be used in preprocessing 3D ears to fill the missing portion of the ear and create upsampled versions of the ear samples.
- **Data quality assessment** is a well-studied subject in the 2D domain. However, very little research has been conducted to assess the quality of 3D biometrics data. Additionally, the quality assessment in 3D is application-dependent. For example, we cannot use the same quality standards on 3D biometrics data as on 3D building data. The author of [41] addresses the effect of data quality on performance; however, further research in establishing proper quality measures for assessing 3D ear data for the biometric application is still necessary. The interoperability of sensors should also be considered when setting such quality standards.
- **Symmetry** is another fascinating field of study in ear biometrics, which may be advantageous for recognition, occluded ear construction, and multimodal recognition. Symmetry refers to the similarity between an individual's left and right ears. For example, if a person enrolls in a biometric system using their left ear, can they be verified using their right ear? This question has been discussed in a few research publications in the 2D domain [1] and [113]. Symmetry studies typically match a subject's ear to its mirror image. However, [1] reported that a significant portion of a person's left and right ear regions are asymmetrical. To corroborate this result, an experiment with left and right profile faces, and left and right ears was performed in [158], which discovered that the left and right profile faces resemble the left and right ears. Recently, a comprehensive investigation was conducted in [113] to determine which portion of the ear is responsible for gender identity and symmetry. In contrast to [1], this study concluded a strong likelihood of symmetry between the left and right ears; therefore, it can be used for recognition. However, it also emphasizes using the middle portion of the ear in recognition. Most research exploring symmetry in ear biometrics has been performed on 2D images. The rich geometric features of the 3D ear could be an excellent choice for

[14]https://www-users.cs.york.ac.uk/ nep/research/YEM/

investigating symmetry [37]; however, only a single work has been reported so far in this context [181]. We believe there is still wide room to explore and leverage ear symmetry property in 3D ear recognition.

11 Conclusions

This paper has presented a contemporary survey of the state-of-the-art methods for 3D ear recognition. First, the approaches for keypoint detection in 3D ears are discussed, and to give a qualitative view, results for a few approaches are reproduced. Further, 3D ear recognition techniques based on registration and feature-descriptor-based techniques are discussed. A comprehensive taxonomy and chronological overview of these methods have been presented. Merits and demerits of various methods are also covered, with potential research directions being listed.

12 Acknowledgement

This research is supported by the Visvesvaraya Ph.D. Scheme for Electronics & IT of Digital India Corporation (formerly Media Lab), Ministry of Electronics & Information Technology, Government of India, and a research fund from Khalifa University, UAE (Ref: CIRA-2019-047). We thank Mr. Akhilesh Mohan Srivastava and Mr. Uttam Sharma for their unconditional support in bringing the IIT Indore dataset into shape. We also thank Mr. Ishan R Dave and Mr. Vivek Baghel for proofreading the manuscript and the statistical analysis of the IIT Indore dataset.

References

[1] Ayman Abaza and Arun Ross. 2010. Towards understanding the symmetry of human ears: A biometric perspective. In *Proc. of Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*. 1–7.

[2] Ayman Abaza, Arun Ross, Christina Hebert, Mary Ann F Harrison, and Mark S Nixon. 2013. A survey on ear biometrics. *ACM Computing Surveys (CSUR)* 45, 2 (2013), 1–35.

[3] Mohamed Abdel-Mottaleb and Jindan Zhou. 2006. Human ear recognition from face profile images. In *Proc. of International Conference on Biometrics (ICB)*. 786–792.

[4] Essam Abdellatef, Eman M Omran, Randa F Soliman, Nabil A Ismail, Salah Eldin SE Abd Elrahman, Khalid N Ismail, Mohamed Rihan, Fathi E Abd El-Samie, and Ayman A Eisa. 2020. Fusion of deep-learned and hand-crafted features for cancelable recognition systems. *Soft Computing* 24, 20 (2020), 15189–15208.

[5] Oussama Aiadi, Belal Khaldi, and Cheraa Saadeddine. 2022. MDFNet: an unsupervised lightweight network for ear print recognition. *Journal of Ambient Intelligence and Humanized Computing* (2022), 1–14.

[6] Syed Sadaf Ali, Vivek Singh Baghel, Iyyakutti Iyappan Ganapathi, Surya Prakash, Ngoc-Son Vu, and Naoufel Werghi. 2021. Toe Prints: An Application Study for Biometric Verification in Adults. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 1418–1424.

[7] Syed Sadaf Ali, Iyyakutti Iyappan Ganapathi, Surya Prakash, Pooja Consul, and Sajid Mahyo. 2020. Securing biometric user template using modified minutiae attributes. *Pattern Recognition Letters* 129 (2020), 263–270.

[8] Hammam Alshazly, Christoph Linse, Erhardt Barth, Sahar Ahmed Idris, and Thomas Martinetz. 2021. Towards explainable ear recognition systems using deep residual networks. *IEEE Access* 9 (2021), 122254–122273.

[9] Hammam Alshazly, Christoph Linse, Erhardt Barth, and Thomas Martinetz. [n.d.]. Handcrafted versus CNN Features for Ear Recognition. *Symmetry* 11, 12 ([n.d.]).

[10] Hammam Alshazly, Christoph Linse, Erhardt Barth, and Thomas Martinetz. 2019. Ensembles of Deep Learning Models and Transfer Learning for Ear Recognition. *Sensors* 19, 19 (2019).

[11] Hammam Alshazly, Christoph Linse, Erhardt Barth, and Thomas Martinetz. 2020. Deep convolutional neural networks for unconstrained ear recognition. *IEEE Access* 8 (2020), 170295–170310.

[12] Teresa CS Azevedo, João Manuel RS Tavares, and Mário AP Vaz. 2010. Three-dimensional reconstruction and characterization of human external shapes from two-dimensional images using volumetric methods. *Computer methods in biomechanics and biomedical engineering* 13, 3 (2010), 359–369.

[13] Lucas Ballard, Seny Kamara, Fabian Monrose, and Michael K Reiter. 2008. Towards practical biometric key generation with randomized biometric templates. In *Proc. of 15th ACM Conference on Computer and Communications Security (CCS)*. 235–244.

[14] Salil P Banerjee and Damon L Woodard. 2012. Biometric authentication and identification using keystroke dynamics: A survey. *Journal of Pattern Recognition Research* 7, 1 (2012), 116–139.

[15] Silvio Barra, Maria De Marsico, Michele Nappi, and Daniel Riccio. 2014. Unconstrained ear processing: What is possible and what must be done. In *Signal and Image Processing for Biometrics*. Springer, 129–190.
[16] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. 2006. Surf: Speeded up robust features. In *Proc. of European conference on computer vision (ECCV)*. 404–417.
[17] Paul J Besl and Neil D McKay. 1992. Method for registration of 3-D shapes. In *Sensor Fusion IV: Control Paradigms and Data Structures*, Vol. 1611. 586–606.
[18] Bir Bhanu and Ajay Kumar. 2017. *Deep learning for biometrics*. Springer.
[19] Aimee Booysens and Serestina Viriri. 2020. Exploration of Ear Biometrics with Deep Learning. In *Proc. of International Conference on Computer Vision and Graphics (ICCVG)*. 25–35.
[20] Mark Burge and Wilhelm Burger. 1996. Ear biometrics. In *Biometrics*. Springer, 273–285.
[21] John D Bustard and Mark S Nixon. 2010. 3D morphable model construction for robust ear and face recognition. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2582–2589.
[22] John D Bustard and Mark S Nixon. 2010. Toward unconstrained ear recognition from two-dimensional images. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* 40, 3 (2010), 486–494.
[23] Steven Cadavid and Mohamed Abdel-Mottaleb. 2007. Human identification based on 3D ear models. In *Proc. of IEEE International Conference on Biometrics: Theory, Applications, and Systems (BTAS)*. 1–6.
[24] Akhilesh Chandra and Thomas Calderon. 2005. Challenges and constraints to the diffusion of biometrics in information systems. *Commun. ACM* 48, 12 (2005), 101–106.
[25] Hui Chen and Bir Bhanu. 2005. Contour matching for 3D ear recognition. In *Proc. of Seventh IEEE Workshops on Applications of Computer Vision (WACV/MOTION)*, Vol. 1. 123–128.
[26] Hui Chen and Bir Bhanu. 2007. Human ear recognition in 3D. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29, 4 (2007), 718–737.
[27] Hui Chen and Bir Bhanu. 2009. Efficient recognition of highly similar 3D objects in range images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31, 1 (2009), 172–179.
[28] Hui Chen, Bir Bhanu, and Rong Wang. 2005. Performance evaluation and prediction for 3D ear recognition. In *Proc. of International Conference on Audio-and Video-Based Biometric Person Authentication (AVBPA)*. 748–757.
[29] Long Chen, Zhichun Mu, Bingfei Nan, Yi Zhang, and Ruyin Yang. 2017. TDSIFT: a new descriptor for 2D and 3D ear recognition. In *Proc. of Eighth International Conference on Graphic and Image Processing (ICGIP)*, Vol. 10225. 102250C–102250C–5.
[30] Long Chen, Zhichun Mu, Baoqing Zhang, and Yi Zhang. 2015. Ear recognition from one sample per person. *PloS One* 10, 5 (2015).
[31] Michal Choras. 2007. Emerging methods of biometrics human identification. In *Proc. of International Conference on Innovative Computing, Information and Control (ICICIC)*. 365–365.
[32] Michal Choras. 2007. Image feature extraction methods for ear biometrics–a survey. In *6th International Conference on Computer Information Systems and Industrial Management Applications (CISIM'07)*. 261–265.
[33] Debbrota Paul Chowdhury, Sambit Bakshi, Chiara Pero, Gustavo Olague, and Pankaj Kumar Sa. 2022. Privacy Preserving Ear Recognition System Using Transfer Learning in Industry 4.0. *IEEE Transactions on Industrial Informatics* (2022).
[34] Christopher Choy, Wei Dong, and Vladlen Koltun. 2020. Deep Global Registration. In *Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2511–2520.
[35] Paolo Cignoni, Marco Callieri, Massimiliano Corsini, Matteo Dellepiane, Fabio Ganovelli, and Guido Ranzuglia. 2008. Meshlab: an open-source mesh processing tool. In *Proc. of Eurographics Italian Chapter Conference*, Vol. 2008. 129–136.
[36] Celia Cintas, Mirsha Quinto-Sánchez, Victor Acuña, Carolina Paschetta, Soledad De Azevedo, Caio Cesar Silva de Cerqueira, Virginia Ramallo, Carla Gallo, Giovanni Poletti, and Maria Catira Bortolini. 2016. Automatic ear detection and feature extraction using geometric morphometrics and convolutional neural networks. *IET Biometrics* 6, 3 (2016), 211–223.
[37] P. Claes, J. Reijniers, M. D. Shriver, J. Snyders, P. Suetens, J. Nielandt, and D. De Tré, G.and Vandermeulen. 2015. An investigation of matching symmetry in the human pinnae with possible implications for 3D ear recognition and sound localization. *Journal of anatomy* 226, 1 (2015), 60–72.
[38] Hang Dai, Nick Pears, Patrik Huber, and William AP Smith. 2020. 3D Morphable Models: The Face, Ear and Head. In *3D Imaging, Analysis and Applications*. Springer, 463–512.
[39] Hang Dai, Nick Pears, and William Smith. 2018. A Data-augmented 3D Morphable Model of the Ear. In *Proc. of IEEE International Conference on Automatic Face and Gesture Recognition (FG)*. 404–408.
[40] Ishan R Dave, Iyyakutti Iyappan Ganapathi, Surya Prakash, Syed Sadaf Ali, and Akhilesh Mohan Srivastava. 2018. 3D Ear Biometrics: Acquisition and Recognition. In *Proc. of 15th IEEE India Council International Conference (INDICON)*. 1–6.
[41] Guy De Tré, Robin De Mol, Dirk Vandermeulen, Peter Claes, Jeroen Hermans, and Joachim Nielandt. 2016. Human centric recognition of 3D ear models. *International Journal of Computational Intelligence Systems* 9, 2 (2016), 296–310.
[42] Lavinia Mihaela Dinca and Gerhard Petrus Hancke. 2017. The fall of one, the rise of many: a survey on multi-biometric fusion methods. *IEEE Access* 5 (2017), 6247–6289.

[43] Xin Dong and Yin Guo. 2011. 3D ear recognition using SIFT keypoint matching. *Energy Procedia* 11 (2011), 1103–1109.
[44] Grega Dvoršak, Ankita Dwivedi, Vitomir Štruc, Peter Peer, and Žiga Emeršič. 2022. Kinship Verification from Ear Images: An Explorative Study with Deep Learning Models. In *2022 International Workshop on Biometrics and Forensics (IWBF)*. 1–6.
[45] Simon Eberz, Kasper B Rasmussen, Vincent Lenders, and Ivan Martinovic. 2016. Looks like eve: Exposing insider threats using eye movement biometrics. *ACM Transactions on Privacy and Security (TOPS)* 19, 1 (2016), 1–31.
[46] Simon Eberz, Kasper B Rasmussen, Vincent Lenders, and Ivan Martinovic. 2017. Evaluating behavioral biometrics for continuous authentication: Challenges and metrics. In *Proc. of ACM on Asia Conference on Computer and Communications Security (ACM ASIACCS)*. 386–399.
[47] Susan El-Naggar, Ayman Abaza, and Thirimachos Bourlai. 2016. On a taxonomy of ear features. In *Proc. of IEEE Symposium on Technologies for Homeland Security (HST)*. 1–6.
[48] Meden B. Emeršič, Ž. and P. Peer. 2020. Evaluation and analysis of ear recognition models: performance, complexity and resource requirements. *Neural computing and applications* (2020), 15785–15800.
[49] Žiga Emeršič, Janez Križaj, Vitomir Štruc, and Peter Peer. 2019. Deep ear recognition pipeline. In *Recent Advances in Computer Vision*. Springer, 333–362.
[50] Žiga Emeršič, Dejan Štepec, Vitomir Štruc, and Peter Peer. 2017. Training convolutional neural networks with limited training data for ear recognition in the wild. In *Proc. of IEEE International Conference on Automatic Face Gesture Recognition (FG)*. 987–994.
[51] Žiga Emeršič, Vitomir Štruc, and Peter Peer. 2017. Ear recognition: More than a survey. *Neurocomputing* 255 (2017), 26–39.
[52] Žiga Emeršič, Diego Sušanj, Blaž Meden, Peter Peer, and Vitomir Štruc. 2021. ContexedNet: Context–Aware Ear Detection in Unconstrained Settings. *IEEE Access* 9 (2021), 145175–145190.
[53] Žiga Emeršič, A Kumar SV, BS Harish, W Gutfeter, JN Khiarak, A Pacut, E Hansley, M Pamplona Segundo, S Sarkar, and HJ Park. 2019. The unconstrained ear recognition challenge 2019. In *Proc. of International Conference on Biometrics (ICB)*. 1–15.
[54] Karim Faez, Sara Motamed, and Mahboubeh Yaqubi. 2008. Personal verification using ear and palm-print biometrics. In *Proc. of IEEE International Conference on Systems, Man and Cybernetics (SMC)*. 3727–3731.
[55] Michael Fairhurst. 2013. *Age factors in biometric processing*. Institution of Engineering and Technology.
[56] Keisirou Fukuda, Takanari Minamidani, Hideyasu Sai, and Daishi Watabe. 2017. An attempt of 3D ear recognition using SHOT. *IEICE Technical Report* 117, 42 (2017), 29–34.
[57] Pedro Luis Galdámez, William Raveane, and Angélica González Arrieta. 2017. A brief review of the ear recognition process using deep neural networks. *Journal of Applied Logic* 24 (2017), 62–70.
[58] Iyyakutti Iyappan Ganapathi, Syed Sadaf Ali, and Surya Prakash. 2019. Multi-resolution Local Descriptor for 3D Ear Recognition. In *Proc. of 18th International Conference of the Biometrics Special Interest Group (BIOSIG)*. 221–228.
[59] Iyyakutti Iyappan Ganapathi, Syed Sadaf Ali, and Surya Prakash. 2020. Geometric statistics-based descriptor for 3D ear recognition. *The Visual Computer* 36, 1 (2020), 161–173.
[60] Iyyakutti Iyappan Ganapathi and Surya Prakash. 2016. False mapped feature removal in spin images based 3D ear recognition. In *Proc. of International Conference on Signal Processing and Integrated Networks (SPIN)*. 620–623.
[61] Iyyakutti Iyappan Ganapathi and Surya Prakash. 2017. 3D Ear Based Human Recognition Using Gauss Map Clustering. In *Proc. of ACM India Compute Conference (COMPUTE)*. 83–89.
[62] Iyyakutti Iyappan Ganapathi and Surya Prakash. 2018. 3D ear recognition using global and local features. *IET Biometrics* 7, 3 (2018), 232–241.
[63] Iyyakutti Iyappan Ganapathi, Surya Prakash, and Syed Sadaf Ali. 2021. Secure Multimodal Access with 2D and 3D Ears. In *Machine Learning for Intelligent Multimedia Analytics*. Springer, 1–20.
[64] Iyyakutti Iyappan Ganapathi, Surya Prakash, Ishan Rajendra Dave, Piyush Joshi, Syed Sadaf Ali, and Akhilesh Mohan Shrivastava. 2018. Ear recognition in 3D using 2D curvilinear features. *IET Biometrics* 7, 6 (2018), 519–529.
[65] Daniel Girardeau-Montaut. 2016. CloudCompare. *Retrieved from CloudCompare: https://www.danielgm.net/cc* (2016).
[66] P Gnanasivam and S Muttan. 2011. Ear and fingerprint biometrics for personal identification. In *Proc. of International Conference on Signal Processing, Communication, Computing and Networking Technologies (ICSCCN)*. 347–352.
[67] Dayanand Bharat Gore. 2019. Comparative Study on Feature Extractions for Ear Recognition. *International Journal of Applied Evolutionary Computation* 10, 2 (2019), 8–18.
[68] Yulan Guo, Ferdous Sohel, Mohammed Bennamoun, Min Lu, and Jianwei Wan. 2013. Rotational projection statistics for 3D local surface description and object recognition. *International Journal of Computer Vision* 105, 1 (2013), 63–86.
[69] Yulan Guo, Hanyun Wang, Qingyong Hu, Hao Liu, Li Liu, and Mohammed Bennamoun. 2020. Deep learning for 3D point clouds: A survey. *IEEE transactions on pattern analysis and machine intelligence* (2020).
[70] Omar Hamdy and Issa Traoré. 2011. Homogeneous physio-behavioral visual and mouse-based biometric. *ACM Transactions on Computer-Human Interaction (TOCHI)* 18, 3 (2011), 1–30.
[71] M Hassaballah, Hammam A Alshazly, and Abdelmgeid A Ali. 2019. Ear recognition using local binary patterns: A comparative experimental study. *Expert Systems with Applications* 118 (2019), 182–200.

[72] Yu He, Shengyong Chen, Hongchuan Yu, and Thomas Yang. 2021. A cylindrical shape descriptor for registration of unstructured point clouds from real-time 3D sensors. *Journal of Real-Time Image Processing* 18, 2 (2021), 261–269.
[73] Nabil Hezil and Abdelhani Boukrouche. 2017. Multimodal biometric recognition using human ear and palmprint. *IET Biometrics* 6, 5 (2017), 351–359.
[74] Timothy M Hospedales, Antreas Antoniou, Paul Micaelli, and Amos J Storkey. 2021. Meta-learning in neural networks: A survey. *IEEE transactions on pattern analysis and machine intelligence* (2021).
[75] David J Hurley, Banafshe Arbab-Zavar, and Mark S Nixon. 2008. The ear as a biometric. In *Handbook of Biometrics*. Springer, 131–150.
[76] A. Iannarelli. 1989. *Ear Identification.* Paramount Publishing Company, Freemont, CA.
[77] H Ibrahim, N Hamdy, and M El-Habrouk. 2009. Personal Identification Using Combined Biometrics Techniques. In *Proc. of International Conference on Systems, Signals and Image Processing (IWSSIP)*. 1–4.
[78] Mina Ibrahim, Mark Nixon, and Sasan Mahmoodi. 2011. The effect of time on ear biometrics. In *Proc. of International Joint Conference on Biometrics (IJCB)*. 1–6.
[79] S Islam, M Bennamoun, A Mian, and R Davies. 2008. A fully automatic approach for human recognition from profile images using 2D and 3D ear data. In *Proc. of International Symposium on 3D Data Processing Visualization and Transmission (3DPVT)*. 131–141.
[80] Syed MS Islam, Mohammed Bennamoun, Ajmal S Mian, and Rowan Davies. 2009. Score level fusion of ear and face local 3D features for fast and expression-invariant human recognition. In *Proc. of International Conference Image Analysis and Recognition (ICIAR)*. 387–396.
[81] Syed MS Islam, Mohammed Bennamoun, Robyn Owens, and Rowan Davies. 2008. Biometric Approaches of 2D-3D Ear and Face: A Survey. In *Advances in Computer and Information Sciences and Engineering*. Springer, 509–514.
[82] Syed MS Islam, Mohammed Bennamoun, Robyn A Owens, and Rowan Davies. 2012. A review of recent advances in 3D ear-and expression-invariant face biometrics. *ACM Computing Surveys (CSUR)* 44, 3 (2012), 1–34.
[83] Syed MS Islam, Rowan Davies, Mohammed Bennamoun, and Ajmal S Mian. 2011. Efficient detection and recognition of 3D ears. *International Journal of Computer Vision* 95, 1 (2011), 52–73.
[84] Syed MS Islam, Rowan Davies, Mohammed Bennamoun, Robyn A Owens, and Ajmal S Mian. 2013. Multibiometric human recognition using 3D ear and face features. *Pattern Recognition* 46, 3 (2013), 613–627.
[85] Anil Jain, Lin Hong, and Sharath Pankanti. 2000. Biometric identification. *Commun. ACM* 43, 2 (2000), 90–98.
[86] Anil K Jain, Patrick Flynn, and Arun A Ross. 2007. *Handbook of biometrics.* Springer Science & Business Media.
[87] Anil K Jain, Karthik Nandakumar, and Arun Ross. 2016. 50 years of biometric research: Accomplishments, challenges, and opportunities. *Pattern Recognition Letters* 79 (2016), 80–105.
[88] Gaurav Jaswal, Amit Kaul, and Ravinder Nath. 2016. Knuckle Print Biometrics and Fusion Schemes–Overview, Challenges, and Solutions. *ACM Computing Surveys (CSUR)* 49, 2 (2016), 1–46.
[89] Andrew Teoh Beng Jin and Lu Leng. 2020. Special Issue on Advanced Biometrics with Deep Learning. *Applied Sciences* 10, 13 (2020).
[90] Umit Kacar, Murvet Kirci, Murat Kus, and Ece Olcay Gunes. 2013. An embedded biometric system. In *Proc. of International Conference on Information Fusion (FUSION)*. 736–742.
[91] Aman Kamboj, Rajneesh Rani, and Aditya Nigam. 2021. CG-ERNet: a lightweight Curvature Gabor filtering based ear recognition network for data scarce scenario. *Multimedia Tools and Applications* 80, 17 (2021), 26571–26613.
[92] Aman Kamboj, Rajneesh Rani, and Aditya Nigam. 2021. A comprehensive survey and deep learning-based approach for human recognition using ear biometric. *The Visual Computer* (2021), 1–34.
[93] Aman Kamboj, Rajneesh Rani, Aditya Nigam, and Ranjeet Ranjan Jha. 2021. CED-Net: context-aware ear detection network for unconstrained images. *Pattern Analysis and Applications* 24, 2 (2021), 779–800.
[94] Rohit Katiyar and Vinay Kumar Pathak. 2011. Recognition based on fusion of gait, ear and face features using KPCA method. In *Proc. of International Conference on Intelligent Computing (ICIC)*. 412–419.
[95] Yacine Khaldi and Amir Benzaoui. 2021. A new framework for grayscale ear images recognition using generative adversarial networks under unconstrained conditions. *Evolving Systems* 12, 4 (2021), 923–934.
[96] Yacine Khaldi, Amir Benzaoui, Abdeldjalil Ouahabi, Sebastien Jacques, and Abdelmalik Taleb-Ahmed. 2021. Ear recognition based on deep unsupervised active learning. *IEEE Sensors Journal* 21, 18 (2021), 20704–20713.
[97] Aicha Korichi, Sihem Slatnia, and Oussama Aiadi. 2022. TR-ICANet: A Fast Unsupervised Deep-Learning-Based Scheme for Unconstrained Ear Recognition. *Arabian Journal for Science and Engineering* (2022), 1–12.
[98] Marcel Körtgen, Gil-Joo Park, Marcin Novotni, and Reinhard Klein. 2003. 3D shape matching with 3D shape contexts. In *Proc. of Central European Seminar on Computer Graphics (CESCG)*, Vol. 3. 5–17.
[99] Ruggero Donida Labati, Angelo Genovese, Enrique Muñoz, Vincenzo Piuri, Fabio Scotti, and Gianluca Sforza. 2016. Biometric recognition in automated border control: a survey. *ACM Computing Surveys (CSUR)* 49, 2 (2016), 1–39.
[100] Jiajia Lei, Jindan Zhou, and Mohamed Abdel-Mottaleb. 2013. Gender classification using automatically detected and aligned 3D ear range data. In *Proc. of International Conference on Biometrics (ICB)*. 1–7.
[101] Jiajia Lei, Jindan Zhou, and Mohamed Abdel-Mottaleb. 2013. A novel shape-based interest point descriptor (SIP) for 3D ear recognition. In *Proc. of IEEE International Conference on Image Processing (ICIP)*. 4176–4180.

[102] Jiajia Lei, Jindan Zhou, Mohamed Abdel-Mottaleb, and Xinge You. 2013. Detection, localization and pose classification of ear in 3D face profile images. In *Proc. of IEEE International Conference on Image Processing (ICIP)*. 4200–4204.
[103] Ruihui Li, Xianzhi Li, Chi-Wing Fu, Daniel Cohen-Or, and Pheng-Ann Heng. 2019. PU-GAN: A Point Cloud Upsampling Adversarial Network. In *Proc. of IEEE/CVF International Conference on Computer Vision (ICCV)*. 7202–7211.
[104] Heng Liu and David Zhang. 2011. Fast 3D point cloud ear identification by slice curve matching. In *Proc. of International Conference on Computer Research and Development (ICCRD)*, Vol. 4. 224–228.
[105] Yahui Liu, Guangming Lu, and David Zhang. 2015. An effective 3D ear acquisition system. *PloS One* 10, 6 (2015).
[106] Yahui Liu, Bob Zhang, Guangming Lu, and David Zhang. 2016. Online 3D ear recognition by combining global and local features. *Plos one* 11, 12 (2016).
[107] Yahui Liu, Bob Zhang, and David Zhang. 2015. Ear-parotic face angle: A unique feature for 3D ear recognition. *Pattern Recognition Letters* 53 (2015), 9–15.
[108] MS Lohith and MN Eshwarappa. 2021. Multimodal biometrics for person identification using ear and palm print features. *IETE Journal of Research* 67, 1 (2021), 90–97.
[109] Yichao Ma, Zengxi Huang, Xiaoming Wang, and Kai Huang. 2020. An Overview of Multimodal Biometrics Using the Face and Ear. *Mathematical Problems in Engineering* (2020).
[110] Mohammad H Mahoor, Steven Cadavid, and Mohamed Abdel-Mottaleb. 2009. Multi-modal ear and face modeling and recognition. In *Proc. of IEEE International Conference on Image Processing (ICIP)*. 4137–4140.
[111] Sayan Maity and Mohamed Abdel-Mottaleb. 2014. 3D ear segmentation and classification through indexing. *IEEE Transactions on Information Forensics and Security* 10, 2 (2014), 423–435.
[112] Judith A Markowitz. 2000. Voice biometrics. *Commun. ACM* 43, 9 (2000), 66–73.
[113] Di Meng, Mark S Nixon, and Sasan Mahmoodi. 2021. On Distinctiveness and Symmetry in Ear Biometrics. *IEEE Transactions on Biometrics, Behavior, and Identity Science* 3, 2 (2021), 155–165.
[114] Christopher Middendorff, Kevin W Bowyer, and Ping Yan. 2007. Multi-modal biometrics involving the human ear. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 1–2.
[115] Yaseen Moolla, Anton De Kock, Gugulethu Mabuza-Hocquet, Cynthia Sthembile Ntshangase, Norman Nelufule, and Portia Khanyile. 2021. Biometric Recognition of Infants using Fingerprint, Iris, and Ear Biometrics. *IEEE Access* 9 (2021), 38269–38286.
[116] Md Mursalin and Syed Mohammed Shamsul Islam. 2020. EpNet: A Deep Neural Network for Ear Detection in 3D Point Clouds. In *Proc. of International Conference on Advanced Concepts for Intelligent Vision Systems (ACIVS)*. 15–26.
[117] Md Mursalin and Syed Mohammed Shamsul Islam. 2021. Deep Learning for 3D Ear Detection: A Complete Pipeline From Data Generation to Segmentation. *IEEE Access* 9 (2021), 164976–164985.
[118] Elisa Mussi, Michaela Servi, Flavio Facchini, Rocco Furferi, Lapo Governi, and Yary Volpe. 2021. A novel ear elements segmentation algorithm on depth map images. *Computers in Biology and Medicine* 129 (2021).
[119] Michele Nappi, Stefano Ricciardi, and Massimo Tistarelli. 2017. Real Time 3D Face-Ear Recognition on Mobile Devices: New Scenarios for 3D Biometrics "in-the-Wild". In *Human Recognition in Unconstrained Environments*. Elsevier, 55–75.
[120] Hossein Nejati, Li Zhang, Terence Sim, Elisa Martinez-Marroquin, and Guo Dong. 2012. Wonder ears: Identification of identical twins from ear images. In *Proc. of International Conference on Pattern Recognition (ICPR)*. 1201–1204.
[121] Huy Nguyen-Quoc and Vinh Truong Hoang. 2020. Gender recognition based on ear images: a comparative experimental study. In *Proc. of the International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*. 451–456.
[122] Mark S Nixon, Imed Bouchrika, Banafshe Arbab-Zavar, and John N Carter. 2010. On use of biometrics in forensics: gait and ear. In *Proc. of European Signal Processing Conference (EUSIPCO)*. 1655–1659.
[123] Timo Ojala, Matti Pietikainen, and David Harwood. 1994. Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. In *Proc. of 12th International Conference on Pattern Recognition (ICPR)*, Vol. 1. 582–585.
[124] Ibrahim Omara, Ahmed Hagag, Guangzhi Ma, Abd El-Samie, E Fathi, and Enmin Song. 2021. A novel approach for ear recognition: learning Mahalanobis distance features from deep CNNs. *Machine Vision and Applications* 32, 1 (2021), 1–14.
[125] Georgios Passalis, Ioannis A Kakadiaris, Theoharis Theoharis, George Toderici, and Theodoros Papaioannou. 2007. Towards fast 3D ear recognition for real-life biometric applications. In *Proc. of IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS)*. 39–44.
[126] Pascal Paysan, Reinhard Knothe, Brian Amberg, Sami Romdhani, and Thomas Vetter. 2009. A 3D face model for pose and illumination invariant face recognition. In *Proc. of sixth IEEE international conference on advanced video and signal based surveillance*. 296–301.
[127] Zhi-Yi Chin Chieh-Ming Jiang Pei-Tse Chiang, Meng-Hsun Tsai. 2022. 3D Point Cloud Data Augmentation via Scene Representation Network. https://github.com/joycenerd/3D_Augmentation
[128] Anika Pflug and Christoph Busch. 2012. Ear biometrics: a survey of detection, feature extraction and recognition methods. *IET Biometrics* 1, 2 (2012), 114–129.
[129] P Jonathon Phillips, Patrick J Flynn, Kevin W Bowyer, Richard W Vorder Bruegge, Patrick J Grother, George W Quinn, and Matthew Pruitt. 2011. Distinguishing identical twins by face recognition. In *Proc. of IEEE International Conference on Automatic Face and Gesture*

Recognition (FG). 185–192.
[130] Surya Prakash and Phalguni Gupta. 2014. Human recognition using 3D ear images. *Neurocomputing* 140 (2014), 317–325.
[131] Surya Prakash and Phalguni Gupta. 2015. *Ear biometrics in 2D and 3D: localization and recognition*. Vol. 10. Springer.
[132] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. 2017. Pointnet: Deep learning on point sets for 3D classification and segmentation. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 652–660.
[133] Charles Ruizhongtai Qi, Li Yi, Hao Su, and Leonidas J Guibas. 2017. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. In *Proc. of 31st International Conference on Neural Information Processing Systems (NIPS)*, Vol. 30. 5105–5114.
[134] K Radhika, K Devika, T Aswathi, P Sreevidya, V Sowmya, and KP Soman. 2020. Performance Analysis of NASNet on Unconstrained Ear Recognition. In *Nature Inspired Computing for Data Science*. Springer, 57–82.
[135] B Rajalakshmi and S Sumathi. 2018. Survey of multimodal biometric using ear and finger knuckle image. In *Proc. of International Conference on Communication, Computing and Internet of Things (IC3IoT)*. 48–53.
[136] Solange Ramos-Cooper, Peru Arequipa, and Guillermo Camara-Chavez. 2022. Domain Adaptation for Unconstrained Ear Recognition with Convolutional Neural Networks. *CLEI electronic journal* 25, 2 (2022).
[137] Meryem Regouid, Mohamed Touahria, Mohamed Benouis, and Nicholas Costen. 2019. Multimodal biometric system for ECG, ear and iris recognition based on local descriptors. *Multimedia Tools and Applications* 78, 16 (2019), 22509–22535.
[138] Arun Ross and Ayman Abaza. 2011. Human ear recognition. *Computer* 44, 11 (2011), 79–81.
[139] Radu Bogdan Rusu, Nico Blodow, and Michael Beetz. 2009. Fast point feature histograms (FPFH) for 3D registration. In *Proc. of IEEE International Conference on Robotics and Automation (ICRA)*. 3212–3217.
[140] Samuele Salti, Federico Tombari, and Luigi Di Stefano. 2014. SHOT: Unique signatures of histograms for surface and texture description. *Computer Vision and Image Understanding* 125 (2014), 251–264.
[141] Partha Pratim Sarangi, BS P Mishra, and Sachidanada Dehuri. 2018. Multimodal biometric recognition using human ear and profile face. In *Proc. of International Conference on Recent Advances in Information Technology (RAIT)*. 1–6.
[142] Partha Pratim Sarangi, Deepak Ranjan Nayak, Madhumita Panda, and Banshidhar Majhi. 2021. A feature-level fusion based improved multimodal biometric recognition system using ear and profile face. *Journal of Ambient Intelligence and Humanized Computing* (2021), 1–32.
[143] Sudeep Sarkar, Mauricio Pamplona Segundo, and Earnest Eugene Hansley. 2019. Unconstrained ear recognition using a combination of deep learning and handcrafted features. US Patent 10,423,823.
[144] Ramesh Kumar Panneer Selvam and K. N. Rao. 2009. Pattern extraction methods for ear biometrics - A survey. *World Congress on Nature & Biologically Inspired Computing (NaBIC)* (2009), 1657–1660.
[145] Abdul Serwadda and Vir V Phoha. 2013. Examining a large keystroke biometrics dataset for statistical-attack openings. *ACM Transactions on Information and System Security (TISSEC)* 16, 2 (2013), 1–30.
[146] Maha Sharkas. 2022. Ear recognition with ensemble classifiers; A deep learning approach. *Multimedia Tools and Applications* (2022), 1–27.
[147] Durgesh Singh and Sanjay K Singh. 2014. A survey on human ear recognition system based on 2D and 3D ear images. *Open Journal of Information Security and Applications* 1, 2 (2014), 21–30.
[148] Harsh Sinha, Raunak Manekar, Yash Sinha, and Pawan K Ajmera. 2019. Convolutional neural network-based human identification using outer ear images. In *Soft Computing for Problem Solving*. Springer, 707–719.
[149] Vincent Sitzmann, Michael Zollhöfer, and Gordon Wetzstein. 2019. Scene representation networks: Continuous 3D-structure-aware neural scene representations. *Advances in Neural Information Processing Systems* 32 (2019).
[150] Pallavi Srivastava, Diwakar Agrawal, and Atul Bansal. 2020. Ear detection and recognition techniques: a comparative review. *Advances in Data and Information Sciences* (2020), 533–543.
[151] Dejan Štepec, Žiga Emeršič, Peter Peer, and Vitomir Štruc. 2020. Constellation-based deep ear recognition. In *Deep Biometrics*. Springer, 161–190.
[152] Xiaopeng Sun, Guan Wang, Lu Wang, Hongyan Sun, and Xiaopeng Wei. 2014. 3D ear recognition using local salience and principal manifold. *Graphical Models* 76, 5 (2014), 402–412.
[153] X. Sun and W. Zhao. 2016. Parallel Registration of 3D Ear Point Clouds. In *Proc. of International Conference on Information Technology in Medicine and Education (ITME)*. 650–653.
[154] Xiao-Peng Sun, Si-Hui Li, Feng Han, and Xiao-Peng Wei. 2015. 3D ear shape matching using joint α-entropy. *Journal of Computer Science and Technology* 30, 3 (2015), 565–577.
[155] Kalaivani Sundararajan and Damon L Woodard. 2018. Deep learning for biometrics: A survey. *ACM Computing Surveys (CSUR)* 51, 3 (2018), 1–34.
[156] Sumegh Tharewal, Hanumant Gite, and KV Kale. 2017. 3D face & 3D ear recognition: Process and techniques. In *Proc. of International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC)*. 1044–1049.
[157] Theoharis Theoharis, Georgios Passalis, George Toderici, and Ioannis A Kakadiaris. 2008. Unified 3D face and ear recognition using wavelets on geometry images. *Pattern Recognition* 41, 3 (2008), 796–804.

[158] Önsen Toygar, Esraa Alqaralleh, and Ayman Afaneh. 2018. Symmetric ear and profile face fusion for identical twins and non-twins recognition. *Signal, Image and Video Processing* 12, 6 (2018), 1157–1164.
[159] Mayank Vatsa, Richa Singh, and Angshul Majumdar. 2018. *Deep Learning in Biometrics.* CRC Press.
[160] M Vijay and G Indumathi. 2018. Multimodal Biometric System Using Ear and Palm Vein Recognition Based on GwPeSOA: Multi-SVNN for Security Applications. In *Proc. of International Conference On Computational Vision and Bio Inspired Computing (ICCVBIC).* 215–231.
[161] Kai Wang and Zhi-chun Mu. 2013. A 3D ear recognition method based on auricle structural feature. *International Journal of Computer Science Issues (IJCSI)* 10, 5 (2013).
[162] Zhaobin Wang, Jing Yang, and Ying Zhu. 2021. Review of ear biometrics. *Archives of Computational Methods in Engineering* 28, 1 (2021), 149–180.
[163] Christopher Weber, Stefanie Hahmann, and Hans Hagen. 2010. Sharp feature detection in point clouds. In *Proc. of Shape Modeling International Conference (SMI).* 175–186.
[164] Xin Wen, Peng Xiang, Zhizhong Han, Yan-Pei Cao, Pengfei Wan, Wen Zheng, and Yu-Shen Liu. 2021. Pmp-net: Point cloud completion by learning multi-step point moving paths. In *Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).* 7443–7452.
[165] Wired. 2010. Ears could make better unique ids than fingerprint. https://www.wired.com/2010/11/ears-biometric-identification/.
[166] Damon L Woodard, Timothy C Faltemier, Ping Yan, Patrick J Flynn, and Kevin W Bowyer. 2006. A comparison of 3D biometric modalities. In *Proc. of Conference on Computer Vision and Pattern Recognition Workshop (CVPRW).* 57–61.
[167] Xiaona Xu and Zhichun Mu. 2007. Feature fusion method based on KCCA for ear and profile face based multimodal recognition. In *Proc. of IEEE International Conference on Automation and Logistics (ICAL).* 620–623.
[168] Dogucan Yaman, Fevziye Irem Eyiokur, and Hazım Kemal Ekenel. 2021. Multimodal soft biometrics: combining ear and face biometrics for age and gender classification. *Multimedia Tools and Applications* (2021), 1–19.
[169] Roman V Yampolskiy and Venu Govindaraju. 2008. Behavioural biometrics: a survey and classification. *International Journal of Biometrics* 1, 1 (2008), 81–113.
[170] Ping Yan and KevinW Bowyer. 2005. Empirical evaluation of advanced ear biometrics. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW).* 41–48.
[171] Ping Yan and Kevin W Bowyer. 2005. Ear biometrics using 2D and 3D images. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW).*
[172] Ping Yan and Kevin W Bowyer. 2005. Multi-biometrics 2D and 3D ear recognition. In *Proc. of International Conference on Audio-and Video-Based Biometric Person Authentication (AVBPA).* 503–512.
[173] Ping Yan and Kevin W Bowyer. 2006. An automatic 3D ear recognition system. In *Proc. of International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT).* 326–333.
[174] Ping Yan and Kevin W Bowyer. 2007. Biometric recognition using 3D ear shape. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29, 8 (2007), 1297–1308.
[175] Ping Yan and Kevin W Bowyer. 2007. A fast algorithm for ICP-based 3D shape biometrics. *Computer Vision and Image Understanding* 107, 3 (2007), 195–202.
[176] Ping Yan, Kevin W Bowyer, and Kyong Jin Chang. 2005. ICP-based approaches for 3D ear recognition. In *Biometric Technology for Human Identification II*, Vol. 5779. 282–291.
[177] Li Yuan, Zhichun Mu, and Ying Liu. 2006. Multimodal recognition using face profile and ear. In *Proc. of International Symposium on Systems and Control in Aerospace and Astronautics (ISSCAA).* 887–891.
[178] Li Yuan, Zhi-Chun Mu, and Fan Yang. 2011. A review of recent advances in ear recognition. In *Proc. of Chinese Conference on Biometric Recognition (CCBR).* 252–259.
[179] Hui Zeng, Ji-Yuan Dong, Zhi-Chun Mu, and Yin Guo. 2010. Ear recognition based on 3D keypoint matching. In *Proc. of IEEE International Conference on Signal Processing (ICSP).* 1694–1697.
[180] Hui Zeng, Rui Zhang, Zhichun Mu, and Xiuqing Wang. 2014. Local feature descriptor based rapid 3D ear recognition. In *Proceedings of the 33rd Chinese Control Conference.* 4942–4945.
[181] David Zhang and Guangming Lu. 2013. Two Significant Characteristics in 3D Ear. In *3D Biometrics.* Springer, 51–68.
[182] David Zhang and Guangming Lu. 2015. *3D Biometrics: Systems and Applications.* Springer Publishing Company, Incorporated.
[183] David Zhang, Guangming Lu, and Lei Zhang. 2018. Online 3D Ear Recognition. In *Advanced Biometrics.* springer, 309–328.
[184] Jie Zhang, Wen Yu, Xudong Yang, and Fang Deng. 2019. Few-shot learning for ear recognition. In *Proc. of International Conference on Image, Video and Signal Processing (IVSP).* 50–54.
[185] L. Zhang, L. Li, H. Li, and M. Yang. 2016. 3D Ear Identification Using Block-Wise Statistics-Based Features and LC-KSVD. *IEEE Transactions on Multimedia* 18, 8 (2016), 1531–1541.
[186] Ying Zhang and A Ben Hamza. 2006. Vertex-based anisotropic smoothing of 3D mesh data. In *Proc. of Canadian Conference on Electrical and Computer Engineering (CCECE).* 202–205.

[187] Yuhe Zhang, Chunhui Li, Bao Guo, Chenhao Guo, and Shunli Zhang. 2021. KDD: A kernel density based descriptor for 3D point clouds. *Pattern Recognition* 111 (2021).
[188] Yi Zhang, Zhichun Mu, Li Yuan, Hui Zeng, and Long Chen. 2017. 3D ear normalization and recognition based on local surface variation. *Applied Sciences* 7, 1 (2017), 104/1–21.
[189] Nan Zheng, Aaron Paloski, and Haining Wang. 2016. An efficient user verification system using angle-based mouse movement biometrics. *ACM Transactions on Information and System Security (TISSEC)* 18, 3 (2016), 1–27.
[190] Jindan Zhou, Steven Cadavid, and Mohamed Abdel-Mottaleb. 2011. A computationally efficient approach to 3D ear recognition employing local and holistic features. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition Workshop (CVPRW)*. 98–105.
[191] Jindan Zhou, Steven Cadavid, and Mohamed Abdel-Mottaleb. 2012. An efficient 3-D ear recognition system employing local and holistic features. *IEEE transactions on Information Forensics and Security* 7, 3 (2012), 978–991.
[192] Qinping Zhu and Zhichun Mu. 2018. An Efficient 3D Ear Recognition System Based on Indexing. In *Proc. of Chinese Conference on Biometric Recognition (CCBR)*. 507–516.
[193] Qinping Zhu and Zhichun Mu. 2018. Local and Holistic Feature Fusion for Occlusion-Robust 3D Ear Recognition. *Symmetry* 10, 11 (2018).
[194] Qinping Zhu and Zhichun Mu. 2020. PointNet++ and Three Layers of Features Fusion for Occlusion Three-Dimensional Ear Recognition Based on One Sample per Person. *Symmetry* 12, 1 (2020), 78.
[195] A Yu Zinovyev. 2003. Method and software for fast construction of principal manifolds approximations. In *Tech. Rep., Institute of Advanced Scientific Studies, Bures-sur-Yvette, France.*