Ear Biometrics: A Survey of Detection, Feature Extraction and Recognition Methods

Anika Pflug, Christoph Busch *

July 2, 2012

The possibility of identifying people by the shape of their outer ear was first discovered by the French criminologist Bertillon, and refined by the American police officer Iannarelli, who proposed a first ear recognition system based on only seven features.

The detailed structure of the ear is not only unique, but also permanent, as the appearance of the ear does not change over the course of a human life. Additionally, the acquisition of ear images does not necessarily require a person's cooperation but is nevertheless considered to be non-intrusive by most people.

Because of these qualities, the interest in ear recognition systems has grown significantly in recent years. In this survey, we categorize and summarize approaches to ear detection and recognition in 2D and 3D images. Then, we provide an outlook over possible future research in the field of ear recognition, in the context of smart surveillance and forensic image analysis, which we consider to be the most important application of ear recognition characteristic in the near future.

1 Introduction

As there is an ever-growing need to automatically authenticate individuals, biometrics has been an active field of research over the course of the last decade. Traditional means of automatic recognition, such as passwords or ID cards, can be stolen, faked, or forgotten. Biometric characteristics, on the other hand, are universal, unique, permanent, and measurable.

The characteristic appearance of the human outer ear (or pinna) is formed by the outer helix, the antihelix, the lobe, the tragus, the antitragus, and the concha (see Figure 1).

*Hochschule Darmstadt - CASED, Haardtring 100, 64295 Darmstadt, Germany

The numerous ridges and valleys on the outer ear's surface serve as acoustic resonators. For low frequencies the pinna reflects the acoustic signal towards the ear canal. For high frequencies it reflects the sound waves and causes neighbouring frequencies to be dropped. Furthermore the outer ear enables humans to perceive the origin od a sound.

The shape of the outer ear evolves during the embryonic state from six growth nodules. Its structure, therefore, is not completely random, but still subject to cell segmentation. The influence of random factors on the ear's appearance can best be observed by comparing the left and the right ear of the same person. Even though the left and the right ear show some similarities, they are not symmetric [1].
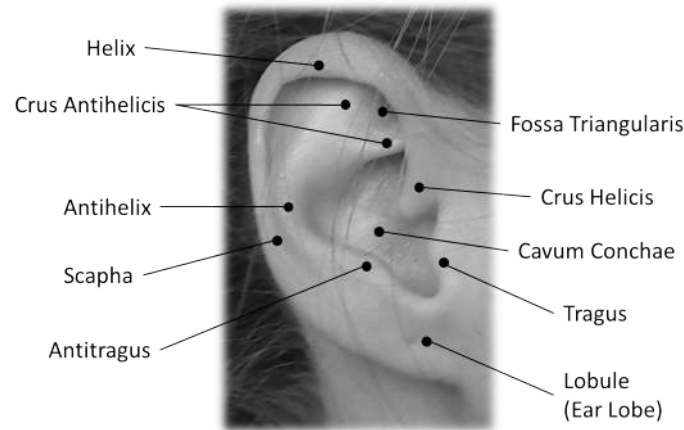

Figure 1: Characteristics of the human ear the German criminal police uses for personal identification of suspects

The shape of the outer ear has long been recognized as a valuable means for personal identification by criminal investigators. The French criminologist Alphonse Bertillon was the first to become aware of the potential use for human identification through ears, more than a century ago [2]. In his studies regarding personal recognition using the outer ear in 1906, Richard Imhofer needed only four different characteristics to distinguish between 500 different ears [3]. Starting in 1949, the American police officer Alfred Iannarelli conducted the first large scale study on the discriminative potential of the outer ear. He collected more than 10 000 ear images and determined 12 characteristics needed to unambiguously identify a person [4]. Iannarelli also conducted studies on twins and triplets, discovering that ears are even unique among genetically identical persons. Even though Iannarelli's work lacks a complex theoretical basis, it is commonly believed that the shape of the outer ear is unique. The studies in [5] and [6] show that all ears of the investigated databases posses individual characteristics, which can be used for distinguishing between them. Because of the lack of a sufficiently large ear database, these studies can only be seen as hints,not evidence, for the outer ear's uniqueness.

Research about the time-related changes in the appearance of the outer ear has shown, that the ear changes slightly in size when a person ages [7] [8]. This is explained by the fact that with ageing the microscopic structure of the ear cartilage changes, which

reduces the skin elasticity. A first study on the effect of short periods of time on ear recognition [9] shows that the recognition rate is not affected by ageing. It must, however, be mentioned that the largest time elapsing difference in this experiment was only 10 months, and it therefore is still subject to further research whether time has a critical effect on biometric ear recognition systems or not.

The ear can easily be captured from a distance, even if the subject is not fully cooperative. This makes ear recognition especially interesting for smart surveillance tasks and for forensic image analysis. Nowadays the observation of characteristics is a standard technique in forensic investigation and has been used as evidence in hundreds of cases. The strength of this evidence has, however, also been called into question by courts in the Netherlands [10]. In order to study the strength of ear prints as evidence, the Forensic Ear identification Project (FearID) was initiated by nine institutes from Italy, the UK, and the Netherlands in 2006. In their test system, they measured an Equal Error Rate (EER) of 4% and came to the conclusion that ear prints can be used as evidence in a semi-automated system [11]. The German criminal police use the physical properties of the ear in connection with other appearance-based properties to collect evidence for the identity of suspects from surveillance camera images. Figure 1 illustrates the most important elements and landmarks of the outer ear, which are used by the German BKA for manual identification of suspects.

In this work we extend existing surveys on ear biometrics, such as [12], [13], [14], [15] or [16]. Abaza *et al.* [17] contributed an excellent survey on ear recognition in March 2010. Their work covers the history of ear biometrics, a selection of available databases and a review of 2D and 3D ear recognition systems. This work amends the survey by Abaza *et al.* with the following:

- A survey of free and publicly available databases.

- More than 30 publications on ear detection and recognition from 2010 to 2012 that were not discussed in one of the previous surveys.

- An outlook over future challenges for ear recognition systems with respect to concrete applications.

In the upcoming Section we give an overview of image databases suitable for studying ear detection and recognition approaches for 2D and 3D images. Thereafter, we discuss existing ear detection approaches on 2D and 3D images. In Section 4 we go on to give an overview of ear recognition approaches for 2D images, and in Section 5 we do the same for 3D images. We will conclude our work by providing an outlook over future challenges and applications for ear recognition systems.

2 Available Databases for Ear Detection and Recognition

In order to test and compare the detection or recognition performance of a computer vision system, in general, and a biometric system in particular, image databases of

3

sufficient size must be publicly available. In this section, we want to give an overview of suitable databases for evaluating the performance of ear detection and recognition systems, which can either be downloaded freely or can be licensed with reasonable effort.

2.1 USTB Databases

The University of Science and technology in Beijing offers four collections (`http://www1.ustb.edu.cn/resb/en/doc/Imagedb\_123\_intro\_en.pdf`) (`http://www1.ustb.edu.cn/resb/en/doc/Imagedb\_4\_intro\_en.pdf`) of 2D ear and face profile images to the research community. All USTB databases are available under license.

- Database I: The dataset contains 180 images in total, which were taken from 60 subjects in 3 sessions between July and August 2002. The database only contains images of the right ear from each subject. During each session, the images were taken under different lighting conditions and with a different rotation. The subjects were students and teachers from USTB.

- Database II: Similarly to database I, this collection contains right ear images from students and teachers from USTB. This time, the number of subjects is 77 and there were 4 different sessions between November 2003 and January 2004. Hence the database contains 308 images in total, which were taken under different lighting conditions.

- Database III: In this dataset 79, students and teachers from USTB were photographed in different poses between November 2004 and December 2004. Some of the ears are occluded by hair. Each subject rotated his or her head from 0 degrees to 60 degrees to the right and from 0 degrees to 45 degrees to the left. This was repeated on two different days for each subject, which resulted in 1600 images in total.

- Database IV: Consisting of 25500 images from 500 subjects taken between June 2007 and December 2008, this is the largest dataset at USTB. The capturing system consists of 17 cameras and, is capable of taking 17 pictures of the subject simultaneously. These cameras are distributed in a circle around the subject, who is placed in the center. The interval between the cameras is 15 degrees. Each volunteer was asked to look upwards, downwards and eyelevel, which means that this database contains images at different yaw and pitch poses. Please note that this database only contains one session for each subject.

2.2 UND Databases

The University of Notre Dame (UND) offers a large variety of different image databases, which can be used for biometric performance evaluation. Among them are five databases containing 2D images and depth images, which are suitable for evaluation ear recognition systems. All databases from UND can be made available under license (`http://cse.nd.edu/~cvrl/CVRL/Data\_Sets.html`).

- Collection E: 464 right profile images from 114 human subjects, captured in 2002. For each user, between 3 and 9 images were taken on different days and under varying pose and lighting conditions.

- Collection F: 942 3D (depth images) and corresponding 2D profile images from 302 human subjects, captured in 2003 and 2004.

- Collection G: 738 3D (depth images) and corresponding 2D profile images from 235 human subjects, captured between 2003 and 2005

- Collection J2: 1800 3D (depth images) and corresponding 2D profile images from 415 human subjects, captured between 2003 and 2005 [18].

- Collection NDOff-2007: 7398 3D and corresponding 2D images of 396 human subject faces. The database contains different yaw and pitch poses, which are encoded in the file names [19].

2.3 WPUT-DB

(a) Good quality   (b) Occlusion by hair   (c) Sparse lighting   (d) Occlusion by jewelery
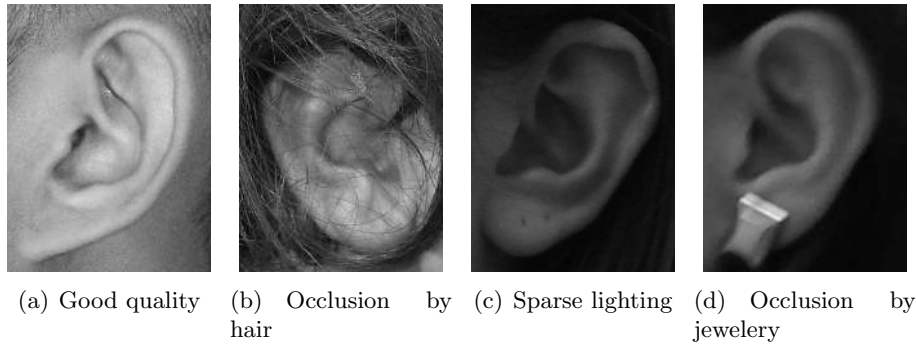
Figure 2: Example images from the WPUT ear database [20]. The database contains ear photographs of varying quality and taken under different lighting conditions. Furthermore the database contains images, where the ear is occluded by hair or by earrings.

The West Pommeranian University of Technology has collected an ear database with the goal of providing more representative data than comparable collections (`http://ksm.wi.zut.edu.pl/wputedb/`) [20]. The database contains 501 subjects of all ages and 2071 images in total. For each subject, the database contains between 4 and 8 images, which were taken on different days and under different lighting conditions. The subjects are also wearing headdresses, earrings and hearing aids, and in addition to this, some ears are occluded by hair. In Figure 2, some example images from the database are shown. The presence of each of these disruptive factors is encoded in the file names of the images. The database can be freely downloaded from the given URL.

2.4 IIT Delhi

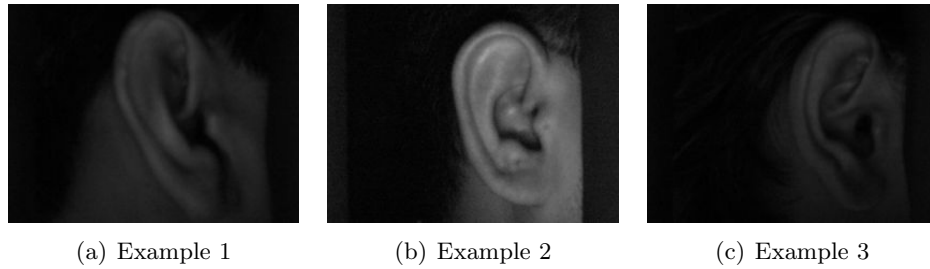(a) Example 1   (b) Example 2   (c) Example 3

Figure 3: Example images fron IIT Delhi ear database [21].

The IIT Delhi Database is provided by the Hong Kong Polytechnic University (`http://www4.comp.polyu.edu.hk/~csajaykr/IITD/Database\_Ear.htm`) [21]. It contains ear images that were collected between October 2006 and June 2007 at the Indian Institute of Technology Delhi in New Delhi (see Figure 3). The database contains 121 subjects, and at least 3 images were taken per subject in an indoor environment, which means that the database consists of 421 images in total.

2.5 IIT Kanpur

The IITK database was contributed by the Indian Institute of Technology in Kanpur (`http://www.cse.iitk.ac.in/users/biometrics/`) [22]. This database consists of two subsets.

- Subset I: This dataset contains 801 side face images collected from 190 subjects. Number of images acquired from an individual varies from 2 to 10.

- Subset II: The images in this subset were taken from 89 individuals. For each subject 9 images were taken with three different poses. Each pose was captured at three different scales. Most likely, all images were taken on the same day. It is not stated whether subset II contains the same subjects as subset I.

2.6 ScFace

The SCface database is provided by the Technical University of Zagreb (`http://www.scface.org/`) [23] and contains 4160 images from 130 subjects. The aim of the database is to provide a database, which is suitable for testing algorithms under surveillance scenarios. Unfortunately, all surveillance camera images were taken at a frontal angle, such that the ears are not visible on these images. However the database also contains a set of high resolution photographs from each subject, which show the subject at different poses. These poses include views of the right and left profile, as shown in Figure 4. Even though the surveillance camera images are likely to be unsuitable for ear recognition studies, the high resolution photographs could be used for examining resistance to pose variations of an algorithm.

(a) Half left profile   (b) Full left profile   (c) Full right profile

Figure 4: SCface example images [23]. These images show examples for the photographed pictures, not for the pictures collected with the surveillance camera system.

2.7 Sheffield Face Database

This database was formerly known as the UMIST (`http://www.sheffield.ac.uk/eee/research/iel/research/face`) database and consists of 564 images of 20 subjects of mixed race and gender. Each subject is photographed in a range of different yaw poses, including a frontal view and profile views.

2.8 YSU

The Youngston State University collected a new kind of biometric database for evaluation forensic identification systems [24]. For each of the 259 subjects, 10 images are provided. The images are grabbed from a video stream and show the subject in poses between zero and 90 degrees. This means that the database contains right profile images and a frontal view image for each subject. It also contains hand drawn sketches from 50 randomly selected subjects from a frontal angle. However this part of the database is not of interest for ear recognition systems.

2.9 NCKU

The National Cheng Kung University in Taiwan has collected an image database, which consists of 37 images for each of the 90 subjects. It can be downloaded from the university's website (`http://robotics.csie.ncku.edu.tw/Databases/FaceDetect_PoseEstimate.htm#Our_Database_`). Each subject is photographed in different angles between -90 degrees (left profile) and 90 degrees (right profile) in 5 degree steps. Figure 5 some examples
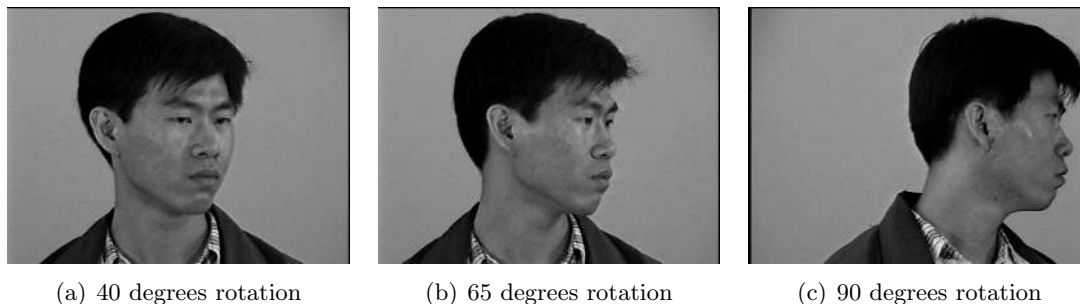
|(a) 40 degrees rotation|(b) 65 degrees rotation|(c) 90 degrees rotation|

Figure 5: Some example images from the NKCU face database, showing the same subject at different angles.

for this are displayed. Such a series of images is collected at two different days for each of the subjects. All images were taken under the same lighting conditions and with the same distance between the subject and the camera.

As this data was originally collected for face recognition, some of the ears are partly or fully occluded by hair, which make this data challenging for ear detection approaches. Consequently, only a subset of this database is suitable for ear recognition.

2.10 UBEAR dataset

The dataset presented in [25] contains images from the left and the right ear of 126 subjects. The images were taken under varying lighting conditions and the subjects were not asked to remove hair, jewelry or headdresses before taking the pictures. The images are cropped from video stream, which shows the subject in different poses, such as looking towards the camera, upwards or downwards.

Additionally, the ground truth for the ear's position is provided together with the database, which makes it particularly convenient for researches to study the accuracy of ear detection and to study the ear recognition performance independently from any ear detection.

3 Ear Detection

This section summarizes the state of the art in automatic ear detection in 2D and 3D images respectively. Basically all ear detection approaches are relying on mutual properties of the ears morphology, like the occurrence of certain characteristic edges or frequency patterns. Table 1 gives a short overview of the ear detection methods outlined below. The upper part of the table contains algorithms for 3D ear localization, whereas the lower part lists algorithms designed for ear detection in 2D images.

Chen and Bhanu propose three different approaches for ear detection. In the approach from [26] Chen and Bhanu train a classifier, which recognizes a specific distribution of shape indices, which are characteristic for the ear's surface. However this approach only works on profile images and is sensitive to any kind of rotation, scale and pose variation.

Table 1: Summary of Automatic ear detection methods for 2D and 3D images

Publication	Detection Method	Database		Perf.
		#Img	Type	
Chen & Bhanu [28]	Shape model and ICP	700	3D	87.71%
Chen & Bhanu [27]	Helix Shape Model	213	3D	92.6%
Zhou et al. [37]	Histograms of Categorized Shapes	942	3D	100%
Prakash & Gupta [32]	connectivity graph	1604	3D	99.38%
Abaza et al. [33]	Cascaded adaboost	940	2D	88.72%
Ansari and Gupta [30]	Edge detection and curvature estimation	700	2D	93.34%
Alastair et al. [39]	Ray transform	252	2D	98.4%
Alvarez et a.l [36]	Ovoid model	NA	2D	NA
Arbab-Zavar & Nixon [38]	Hough Transform	942	2D	91%
Arbab-Zavar & Nixon [45]	Log-Gabor filters and wavelet transform	252	2D	88.4%
Attarchi et al. [29]	Edge detection and line tracing	308	2D	98.05%
Chen & Bhanu [26]	Template Matching with Shape index histograms	60	2D	91.5%
Islam et al. [85]	Adaboost	942	2D	99.89%
Jeges & Mate [62]	Edge orientation pattern	330	2D	100%
Kumar et al. [40]	Edge clustering and active contours	700	2D	94.29%
Liu & Liu [86]	Adaboost and skin colour filtering	50	2D	96%
Prakash & Gupta [31]	Skin Colour and graph matching	1780	2D	96.63%
Shih et al. [87]	Arc-Masking and AdaBoost	376	2D	100%
Yan & Bowyer [18]	Concha Detection and active contours	415	2D	>97.6
Yuan & Mu [35]	CAMSHIFT and aontour fitting	Video	2D	NA

(a) Original image, edge image and Hough transform [38]

(b) Original image and ray transform [39]

(c) Original image, edge enhanced image and corresponding
edge orientation model [62]

Figure 6: Examples for different ear detection techniques

In their later ear detection approaches Chen and Bhanu detected image regions with a large local curvature with a technique they called step edge magnitude [27]. Then a template, which contains the typical shape of the outer helix and the anti-helix, is fitted to clusters of lines. In [28] where Chen and Bhanu narrowed the number of possible ear candidates by detecting the skin region first before the helix template matching is applied on the curvature lines. By fusing color and curvature information, the detection rate could be raised to 99.3% on the UCR dataset and 87.71% on UND collection F and a subset of collection G. The UCR dataset is not publicly available and is hence not covered in Section 2. For a description of this dataset see [17].

Another example for ear detection using contour lines of the ear is described by Attrachi *et al.* [29]. They locate the outer contour of the ear by searching for the longest connected edge in the edge image. By selecting the top, bottom, and left points of the detected boundary, they form a triangle with the selected points. Further the barycenter of the triangle is calculated and selected as reference point for image alignment. Ansari *et al.* also use an edge detector in the first step of their ear localization approach [30]. The edges are separated into two categories, namely convex and concave. Convex edges are chosen as candidates for representing the outer contour. Finally the algorithm connects the curve segments and selects the figure enclosing the largest area for being the

outer ear contour. It should be noted that the IITK database and USTB II already contain cut-out ear images. Hence it can be put into question, whether the detection rates of 93.34% and 98.05% can be reproduced under realistic conditions.

A recent approach on 2D ear detection using edges is described by Prakash and Gupta in [31]. They combine skin segmentation and categorization of edges into convex and concave edges. Afterwards the edges in the skin region are decomposed into edge segments. These segments are composed to form an edge connectivity graph. Based on this graph the convex hull of all edges, which are believed to belong to the ear, is computed. The enclosed region is then labeled as the ear region. In contrast to [29], Prakash and Gupta prove the feasibility of edge-based ear detection on full profile images, where they achieved a detection rate of 96.63% on a subset of the UND-J2 collection. In [32] propose the same edge connectivity for ear recognition on 3D images. Instead of edges, they use discontinuities in the depth map for extracting the initial edge image and then extract the connectivity graph. In their experiments, they use the 3D representations of the same subset as in [31] and report a detection rate of 99.38%. Moreover they show that the detection rate of their graph-based approach is not influence by rotation and scale.

Jedges and Mate propose another edge-based ear detection approach, which is likely to be inspired by fingerprint recognition techniques. They train a classifier with orientation pattern, which were previously computed from ear images. Like other naive classifiers, their method is not robust against rotation and scale. Additionally the classifier is likely to fail under large pose variations, because this will affect the appearance of the orientation pattern.

Abaza *et al.* [33] and Islam *et al.* [34] use weak classifiers based on Haar-wavelets in connection with AdaBoost for ear localization. According to Islam *et al.*, the training of the classifier takes several days, however once the classifier is set up, ear detection is fast and effective. Abaza *et al.* use a modified version of AdaBoost and report a significantly shorter training phase. The effectiveness of their approach is proved in evaluations on five different databases. They also include some examples of successful detections on images from the internet. As long as the subject's pose does not change, weak classifiers are suitable for images which contain more than one subject. Depending on the test set Abaza *et al.* achieved a detection rate between 84% and 98.7% on the Sheffield Face database. On average, their approach successfully detected 95% of all ears.

Yan and Browyer developed an ear detection method which fuses range images and corresponding 2D color images [18]. Their algorithm starts by locating the concha and then uses active contours for determining the ear's outer boundary. The concha serves as the reference point for placing the starting shape of the active contour model. Even though the concha is easy to localize in profile images, it may be occluded if the head pose changes or if a subject is wearing a hearing aid or ear phones. In their experiments Yan and Browyer only use ear images with minor occlusions where the concha is visible; hence it could neither be proved nor disproved whether their approach is capable of reliably detecting ears if the concha is occluded.

Yuan and Mu developed a method for real-time ear tracking in video sequences by applying Continuously Adaptive Mean Shift (CAMSHIFT) to video sequences [35]. The CAMSHIFT algorithm is frequently used in face tracking applications and is based on

region matching and a skin color model. For precise ear segmentation, the contour fitting method based on modified active shape models, which have been proposed by Alvarez *et al.* is applied [36]. Yuan and Mu report a detection rate of 100%, however the test database only consisted of two subjects. Nevertheless their approach appears to be very promising for surveillance applications but needs to be further evaluated in more realistic test scenarios.

Shih *et al.* determine ear candidates by localizing arc-shaped edges in an edge image. Subsequently the arc-shaped ear candidates are verified by using an Adaboost classifier. They report a detection rate 100% on a dataset, which consists of 376 images from 94 subjects.

Zhou *et al.* train a 3D shape model in order to recognize the histogram of shape indexes of the typical ear [37]. Similarly to the approaches of Abaza *et al.* and Islam *et al.*, a sliding window of different sizes is moved over the image. The ear descriptor proposed by Zhou *et al.* is built from concatenated shape index histograms, which are extracted from sub-blocks inside the detection window. For the actual detection, a Support Vector Machine (SVM) classifier is trained to decide whether an image region is the ear region or not. As far as we know, this is the first ear detection approach, which does not require having corresponding texture images in addition to the range image. Zhou *et al.* evaluated their approach on images from the UND collections and report a detection rate of 100%. It should be noted that this approach was not tested under rotation, pose variations and major occlusions, but under the impression of the good performance, we think this is an interesting task for future research.

Ear detection methods based on image transformations have the advantage of being robust against out-of-plane rotations. They are designed to highlight specific properties of the outer ear, which occur in each image where the ear is visible no matter in which pose the ear has been photographed. In [38] the Hough transform is used for enhancing regions with a high density of edges. In head profile images, a high density of edges especially occurs in the ear region (see Figure 6(a)). In [38] it is reported that the Hough transform based ear detection gets trapped when people wear glasses since the frame introduces additional edges to the image. This especially occurs in the eye and nose region. The ear detection approach based on Hough transform was evaluated on the images in the XM2VTS database (see [17] for a detailed database description), where a detection rate of 91% was achieved.

The ray transform approach proposed in [39] is designed to detect the ear in different poses and to ignore straight edges in the image, which can be introduced by glasses or hair. Ray transform uses a light ray analogy to scan the image for tubular and curved structures like the outer helix. The simulated ray is reflected in bright tubular regions and hence these regions are highlighted in the transformed image. Since glass frames have straight edges, they are not highlighted by the ray transform (see Figure 6(b)). Using this method Alastair *et al.* achieved am impressive recognition rate of 98.4% on the XM2VTS database. Hence, the ray transform approach by Alastair *et al.* outperforms Hough transform, most likely because it is more robust against disruptive factors such as glasses or hair.

A recent approach for 2D ear detection is described in [40]. Kumar *et al.* propose to

extract ears from 2D images by using edge images and active contours. They evaluate their approach on a database, which consists of 100 subjects with 7 images per subject. A special imaging device was used for collecting the data. This device makes sure that the distance to the camera is constant and that the lighting conditions are the same for all images. Within this setting a detection rate of 94.29% is reported.

4 2D Ear Recognition

Each ear recognition system consists of a feature extraction and a feature vector comparison step. In this survey we divide ear recognition approaches into four different subclasses namely holistic approaches, local approaches, hybrid approaches and statistical approaches.

In the Tables 2 and 3 all 2D ear recognition approaches mentioned in this paper are summarized in chronological order.

4.1 Holistic Descriptors

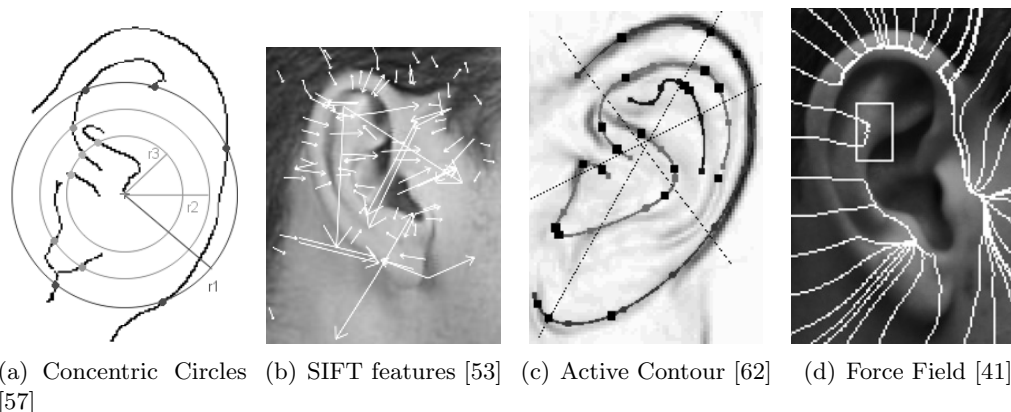(a) Concentric Circles [57] (b) SIFT features [53] (c) Active Contour [62] (d) Force Field [41]

Figure 7: Examples for feature extraction for 2D ear images.

Another approach, which has gained some popularity is the Force Field Transform by Hurley [41]. The Force Field transformation approach assumes that pixels have a mutual attraction proportional to their intensities and inversely to the square of the distance between them rather like Newton's universal law of gravitation. The associated energy field takes the form of a smooth surface with a number of peaks joined by ridges (see Figure 7(d)). Using this method, Hurley *et al.* achieved a rank-1 performance of more than 99% on the XM2VTS database (252 images). Building on these results, Abdel-Mottaleb and Zhou use a 3D representation of the force field for extracting points lying on the peak of the 3D force field [42]. Because the force field converged at the outline of the ear, the peaks in the 3D representation basically represent the ear contour. Nonetheless, the force field method is more robust against noise than other edge detector,

13

Table 2: Summary of approaches for 2D ear recognition approaches, part 1. Unless stated differently, performance always refers to rank-1 performance.

| Publication | Summary | Database | | Perf. |
		#Subj	# Img	
Burge and Burger [88]	Vornoi Distance Graphs	NA	NA	NA
Yuan and Mu [66]	Full Space LDA with Outer Helix Feature Points	79	1501	86.76%
Hurley [41]	Force Field Transform	63	252	99%
Moreno et al. [89]	Geometric features with Compression Network	28	268	93%
Yuizono et al. [73]	Genetic Local Search	110	660	99%
Victor et al. [67]	PCA	294	808	40%
Chang et al. [68]	PCA	114	464	72.7%
Abdel-Mottaleb and Zhou [42]	Modified Force Field Transform	29	58	87.9%
Mu et al. [58]	Geometrical measures on edge images	77	308	85%
Abate et al. [46]	General Fourier Descriptor	70	210	88%
Lu et al. [65]	Active Shape Model and PCA	56	560	93.3%
Yuan et al. [90]	Non-Negative Matrix Factorization	77	308	91%
Arbab-Zavar et al. [53]	SIFT points from ear model	63	252	91.5%
Jedges and Mate [62]	Distorted Ear Model with feature points	28	4060	5.6% EER
Liu et al. [63]	Edge-based features from different views	60	600	97.6%
Nanni and Lumini [72]	Gabor Filters and SFFS	114	464	80%
Rahman et al. [59]	Geometric Features	100	350	87%
Sana et al. [47]	Haar Wavelets and Hamming Distance	600	1800	98.4%
Arbab-Zavar and Nixon [45]	Log-Gabor Filters	63	252	85.7 %
Choras [57]	Geometry of ear outline	188	376	86.2%

Table 3: Summary of approaches for 2D ear recognition approaches, part 2. Unless stated differently, performance always refers to rank-1 performance.

Publication	Summary	Database #Subj	# Img	Perf.
Dong and Mu [43]	Force Field Transform and NKFDA	29	711	75.3%
Guo and Xu [60]	Local Binary Pattern and CNN	77	308	93.3%
Nasseem *et al.* [76]	Sparse representation	32	192	96.88%
Wang *et al.* [91]	Haar Wavelets and Local Binary Patterns	79	395	92.41%
Xie and Mu [71]	Locally Linear Embedding	79	1501	80%
Yaqubi *et al.* [74]	HMAX and SVM	60	180	96.5%
Zhang and Mu [70]	Geometrical Features, ICA and PCA with SVM	77	308	92.21
Badrinath and Gupta [54]	SIFT landmarks from ear model	106	1060	95.32%
Kisku *et al.* [55]	SIFT from different Color Segments	400	800	96.93%
Wang and Yuan [50]	Low-Order Moment Invariants	77	308	100%
Alaraj *et al..* [69]	PCA with MLFFNNs	17	85	96%
Bustard *et al.* [52]	SIFT Point Matches	63	252	96%
De Marisco *et al.* [48]	Partitioned Iterated Function System (PIFS)	114	228	61%
Gutierrez *et al.* [75]	MNN with Sugeno Measures and SCG	77	308	97%
Wang *et al.* [49]	Moment Invariants and BP Neural Network	NA	60	91.8%
Wang and Yuan [44]	Gabor Wavelets and GDA	77	308	99.1%
Fooprateepsiri and Kurutach [92]	Trace and Fourier Transform	68	68	97%
Prakash and Gupta [22]	SURF and NN classifier	300	2066	2.25% EER
Kumar *et al.* [40]	SIFT	100	700	95% GAR, 0.1% FAR
Wang and Yan [61]	Local Binary Pattern and Wavelet Transform	77	308	100%
Kumar and Wu [21]	Phase encoding with Log Gabor filters	221	753	95.93%

15

such as Sobel or Canny. Using this approach, Abdel-Mottaleb and Zhou achieved a rank-1 performance of 87.93% on a dataset with consists of 103 ear images from 29 subjects.

Dong and Mu [43] add pose invariance to the edges, which are extracted by using the force field method. This is achieved with null space kernel fishier discriminant analysis (NKFDA), which has the property of representing non-linear relations between two datasets. Dong and Mu conducted experiments on the USTB IV dataset. Before feature extraction, the ear region was cropped out manually from the images and the pose is normalized. For pose variation of 30 they report a rank-1 recondition rate of 72.2%. For pose variations of 45 the rank-1 performance dropped to 48.1%.

In a recent publication of Kumar and Wu [21] they present an ear recognition approach, which uses the phase information of Log-Gabor filters for encoding the local structure of the ear. The encoded phase information is stored in normalized grey level images. In the experiments, the Log-Gabor approach outperformed force field features and a landmark-based feature extraction approach. Moreover, different combinations of Log-Gabor filters were compared with each other. The rank-e performance for the Log-Gabor approaches ranges between 92.06% and 95.93% on a database which contains 753 images from 221 subjects.

The rich structure of the outer ear results in specific texture information, which can be measured using Gabor filters. Wang and Yuan [44] extract local frequency features by using a battery of Gabor filters and then select the most distinctive features by using general discriminant analysis. In their experiments on the USTB II database, they compared the performance impact of different settings for the Gabor filters. Different combinations of orientation and scales in the filter sets are compared with each other and it was found that neither the number of scales nor the number of orientations has a major impact on the rank-1 performance. The total rank-1 performance of Wang and Yuan's approach is 99.1%. In a similar approach Arbab-Zavar and Nixon [45] measured the performance of Gabor filters in the XM2VTS database where they report a rank-1 performance of 91.5%. A closer look at the Gabor filter responded showed that the feature vectors are corrupted by occlusion or other disruptive factors. In order to overcome this, more robust comparison method is proposed, which resulted in an improved recognition rate of 97.4%.

Abate *et al.* [46] use a generic Fourier descriptor for rotation and scale invariant feature representation. The image is transformed into a polar coordinate system and then transformed into frequency space. In order to make sure, that the centroid of the polar coordinate system is always at the same position, the ear images have to be aligned before they can be transformed into the polar coordinate system. The concha serves as a reference point for the alignment step, such that the center point of the polar coordinate system is always located in the concha region. The approach was tested on a proprietary dataset, which contains 282 ear images in total. The images were taken on two different days and in different roll and yaw poses. The rank-1 performance of the Fourier descriptor varies depending on the pose angle. For 0 pose variation the rank-1 performance is 96%, but if different poses are included in the experiments, it drops to 44% for 15 and 19% for 30.

In the work of Fooprateepsiri and Kurutach exploit the concepts of multi-resolution

Trace transform and Fourier transform. The input images from the CMU PIE database are serialized by using the trace transform and stored in a feature vector. The advantage of the trace transform is that the resulting feature vector is invariant to rotations and scale. Furthermore Fooprateepsiri and Kurutach show that their descriptor is also robust against pose variations. In total they report a rank-1 performance of 97%.

Sana *et al.* use selected wavelet coefficients extracted during Haar-Wavelet compression for feature representation [47]. While applying the four level wavelet transform several times on the ear image, for each iteration they store one of the derived coefficients in a feature vector. The reported accuracy of their algorithm is 96% and was achieved on the basis of the IITK database and on the Saugor database (350 subjects).

A feature extraction system called PIFS is proposed by De Marisco *et al.* [48]. PIFS measures the self-similarity in an image by calculating affine translations between similar sub regions of an image. In order to make their system robust to occlusion, De Marisco *et al.* divided the ear image into equally large tiles. If one tile is occluded, the other tiles still contain a sufficiently distinctive set of features. De Marisco *et al.* could show that their approach is superior to other feature extraction methods under the presence of occlusion. The experiments of De Marisco *et al.* have been conducted in order to assess the system performance in different occlusion scenarios. The basis for these tests was the UND collection E and the first 100 subjects of the FERET database. If occlusion occurs on the reference image, a rank-1 performance of 61% (compared to 40% on average with other feature extraction methods) is reported. Without occlusion, the rank-1 performance is 93%.

Moment invariants are a statistical measure for describing specific properties of a shape. Wang *et al.* [49] compose six different feature vectors by using seven moment invariants. They also show that each of the moment invariants is robust against changes in scale and rotation. The feature vectors are used as the input for a back propagation neural network which is trained to classify the moment invariant feature sets. Based on a proprietary database of 60 ear images, they report a rank-1 performance of 91.8%. In [50] Wang and Yuan compare the distinctiveness of different feature extraction methods on the USTB I database. They compare the rank-1 performance of Fourier descriptors, Gabor-Transform, Moment Invariants and statistical features and come to the conclusion that the highest recognition rate can be achieved by using moment invariants and Gabor transform. For both feature extraction methods Wang and Yuan report a rank-1 performance or 100%.

4.2 Local Descriptors

Scale invariant Feature Transform (SIFT) is known to be a robust way for landmark extraction even in images with small pose variations and varying brightness conditions [51]. SIFT landmarks contain a measure for local orientation; they can also be used for estimating the rotation and translation between two normalized ear images. Bustard *et al.* showed that SIFT can handle pose variations up to 20 degrees [52]. However it is not a trivial to assign a SIFT landmark with its exact counterpart, especially in the presence of pose variations. In highly structured image regions, the density and redundancy of

SIFT landmarks is so high, that exact assignment is not possible. Hence the landmarks have to be filtered before the actual comparison can start. Arbab-Zavar *et al.* [53] as well as Badrinath and Gupta [54] therefore train a reference landmark model, which only contains a small number of non-redundant landmarks. This landmark model is used for filtering the SIFT landmarks, which were initially detected in the probe and reference ear. Having the filtered landmarks it is possible to assign each of the landmarks with its matching counterpart. Figure 7(b) shows an example for SIFT landmarks extracted from an ear images, which were used as training data for the reference landmark model in the work of Arbab-Zavar *et al.*. Because Arbab-Zavar *et al.* also used the XM2VTS database for evaluation, their results can be directly compared to the rank-1 performance reported by Bustard and Nixon. Arbab-Zavar *et al.* achieved a rank-1 performance of 91.5%. With the more recent approach by Bustard and Nixon the performance could be improved to 96%. Using the IIT Delhi database Kumar *et al.* report a GAR of 95% and a FAR of 0.1% when using SIFT feature points.

Kisku *et al.* address the problem of correct landmark assignment by decomposing the ear image into different color segments [55]. SIFT landmarks are extracted from each segment separately, which reduces the chance of assigning SIFT landmarks that are not representing the same features. Using this approach, Kisku *et al.* achieve a rank-1 performance of 96.93%.

A recent approach by Prakash and Gupta [22] fuses Speeded Up Robust Features (SURF) [56] feature points from different images of the same subject. They propose to use several input images for enrolment and to store all SURF feature points in the fused feature vector, which could be found in the input images. These feature sets are then used for training a nearest neighbor classifier for assigning two correlated feature points. If the distance between two SURF feature points is less than a trained threshold, they are considered to be correlated. The evaluation of this approach was carried out on the UND collection E and the two subsets of the IIT Kanpur database. Prakash and Gupta tested the influence of different parameters for SURF features and for the nearest neighbor classifier. Depending on the composition of the parameters the EER varies between 6.72% and 2.25%.

Choras proposes a set of geometric feature extraction methods inspired by the work of Iannarelli [57]. He proposes four different ways of feature location in edge images. The concentric circles method uses the concha as reference points for a number of concentric circles with predefined radii. The intersection points of the circles and the ear contours are used as feature points (see Figure 7(a).). An extension of this is the contour tracing methods, which uses bifurcations, endpoints and intersecting points between the ear contours as additional features. In the angle representation approach, Choras draws concentric circles around each center point of an edge and uses the angles between the center point and the concentric circles intersecting points for feature representation. Finally the triangle ratio method determines the normalized distances between reference points and uses them for ear description. Choras conducted studies on different databases where he reported recognition rates between 86.2% and 100% on a small database off 12 subjects and a false reject rate between 0% and 9.6% on a larger database with 102 ear images.

Similar approaches which are using the aspect ratio between reference points on the ear contours are proposed by Mu *et al.* with a rank-1 performance of 85% on the USTB II database [58] and Rahman *et al.* [59]. Rahman *et al.* evaluated their approach on a database, which consists of 350 images from 100 subjects. They report a rank-1 performance of 90%. For images, which were taken on different days the rank-1 performance dropped to 88%.

Local binary patterns (LBP) are a technique for feature extraction on the pixel level. LBP encode the local neighborhood of a pixel by storing the difference between the examined pixel and its neighbors. Guo *et al.* [60] extract LBP from the raw ear images and create histograms describing the distribution of the local LBP. Then a cellular neural network of trained to distinguish between the LBP of different subjects in the USTB II database.

In the by Wang and Yan [61] the dimensionality of the feature vector is reduced with linear discriminant analysis before a Euclidean distance measure quantifies the similarity of two feature vectors. Wang and Yan evaluated their approach on the USTB II dataset and report a rank-1 performance of 100%.

4.3 Hybrid Approaches

The approach of Jedges and Mate is twofold [62]. In a first feature extraction step they generate an average edge model from a set of training images. These edges represent the outer helix contour as well as the contours of the antihelix, the fossa triangularis and the concha. Subsequently each image is enrolled by deforming the ear model until it fits the actual edges displayed in the probe ear image. The deformation parameters, which were necessary for the transformation, are the rst part of the feature vector. The feature vector is completed by adding additional feature points lying of intersections between a predefined set of axes and the transformed main edges. The axes describe the unique outline of ear. Figure 7(c) shows the edge enhanced images with fitted contours together with the additional axes for reference point extraction. They report an EER of 5.6% using a database with cropped images and without pose variations.

Liu *et al.* combine front and backside view of the ear by extracting features using the triangle ratio method and Tchebichef moment descriptors [63]. Tchebichef moments are a set of orthogonal moment functions based on discrete Tchebichef polynomials and have been introduced as a method for feature representation in 2001 [64]. The backside of the ear is described by a number of lines that are perpendicular to the longest axis in the ear contour. These lines measure the local diameter of the auricle at predefined points. The rank-1 performance of this combined approach is reported to be 97.5%. If only the front view is used, the rank-1 performance is 95% and for the backside images, Liu *et al.* report 86.3% rank-1 performance.

Lu *et al.* [65] as well as Yuan and Mu [66] use the active shape model for extracting the outline of the ear. Lu *et al.* are using manually cropped ear images from 56 subjects in different poses. A feature extractor stores selected points on the outline of the ear together with their distance to the tragus. Before applying a linear classifier, the dimensionality of the feature vectors is reduced by principal component analysis (PCA).

Lu *et al.* compare the rank-1 performance of pipelines where only the left or the right ear was used for identification and also show that using both ears increases the rank-1 performance from 93.3% to 95.1%. In the USTB III database Yuan and Mu report a rank-1 performance of 90% is the head rotation is lower than 15. For rotation angles between 20 and 60 the rank-1 performance drops to 80%.

4.4 Classifiers and Statistical Approaches

Victor *et al.* were the first research group to transfer the idea of using the eigenspace from face recognition to ear recognition [67]. They reported that the performance of the ear as a feature is inferior to the face. This may be due to the fact that in their experiments Victor *et al.* considered the left and the right ear to be symmetric. They used the one ear for training and the other ear for testing, which could have lowered the performance of PCA in this case. The reported rank-1 performance is 40%. With a rank-1 performance of 72.2% in the UND collection E, Chang *et al.* [68] report a significantly better performance than Victor *et al.*. Alaraj *et al.* [69] published another study, where PCA is used for feature representation in ear recognition. In their approach a multilayer feed forward neural network was trained for classification of the PCA based feature components. The observed a rank-1 performance of 96%, and hence improved the previous results by Victor *et al.* and Chang *et al.*. However it should be noticed that this result is only based on a subset of one of the UND collections, which consists of 85 ear images from 17 subjects.

Zhang and Mu conducted studies on the effectiveness of statistical methods in combination with classifiers. In [70] they show that independent component analysis (ICA) is more effective on the USTB I database than PCA. They first used PCA and ICA for reducing the dimensionality of the input images and then trained an SVM for classifying the extracted feature vectors. Furthermore the influence of different training set sizes on the performance was measured. Depending on the site of the training set the rank-1 performance for PCA varies between 85% and 94.12%, whereas the rank-1 performance for ICA varies between 91.67% and 100%.

Xie and Mu [71] propose an improved locally linear embedding (LLE) algorithm for reducing the dimensionality of ear features. LLE is a technique for projecting high-dimensional data points into a lower dimensional coordinate system while preserving the relationship between the single data points. This requires the data points to be labeled in some way, so that their relationship is fixed. The improved version om LLE by Xie and Mu eliminated the problem by using a different distance function. Further Xie and Mu show, that LLE is superior to PCA and Kernel PCA, if the input data contains pose variations. Their studies were conducted on the USTB III database showed that the rank-1 performance of regular LLE (43%) is improved significantly by their method to 60.75%. If the pose variation is only 10, the improved LLE approach achieved a rank-1 performance of 90%.

In their approach Nanni and Lumini [72] propose to use Sequential Forward Floating Selection (SFFS), which is a statistical iterative method for feature selection in pattern recognition tasks. SFFS tries to find the best set of classifiers by creating a set of rules,

which best fits the current feature set. The sets are created by adding one classifier at a time and evaluating its discriminative power with a predefined fitness function. If the new set of rules outperforms the previous version, the new rule is added to the final set of rules. The experiments were carried out on the UND collection E and the single classifiers are fused by using the weighted sum rule. SFFS selects the most discriminative sub-windows which correspond to the fittest set of rules. Nanni and Lumini report a rank-1 recognition rate of 80% and a rank-5 recognition rate of 93%. The EER varies between 6.07% and 4.05% depending on the number of sub-windows used for recognition.

Yiuzono *et al.* consider the problem of finding corresponding features in ear images as an optimization problem and apply genetic local search for solving it iteratively [73]. They select local sub windows with varying size as the basis for the genetic selection. In [73] Yiuzono *et al.* present elaborated results, which describe the behavior of genetic local search under different parameters, such as different selection methods and different numbers of chromosomes. On a database of 110 subjects they report a recognition rate of 100%.

Yaqubi *et al.* use features obtained by a combination of position and scale-tolerant edge detectors over multiple positions and orientations of the image [74]. This feature extraction method is called HMAX model and is inspired by the visual cortex of primates and combines simple features to more complex semantic entities. The extracted features are classified with an SVN and a kNN. The rank-1 performance on a small dataset of 180 cropped ear images from 6 subjects varies between 62% and 100% depending on the kind of basis features.

Moreno *et al.* implement a feature extractor, which locates seven landmarks on the ear image, which correspond to the salient points from the work of Iannarelli. Additionally they obtain a morphology vector, which describes the ear as a whole. These two features are used as the input for different neural network classifiers. They compare the performance of each of the single feature extraction techniques with different fusion methods. The proprietary test database is composed of manually cropped ears from 168 from 28 subjects. The best result of 93% rank-1 performance was measures using a compression network. Other configurations yielded error rates between 16% and 57%.

Gutierrez *et al.* [75] divide the cropped ear images into three equally sized parts. The upper part shows the helix, the middle part shows the concha and the lower part shows the lobule. Each of these sub images is decomposed by wavelet transform and then fed into a modular neural network. In each module of the network a different integrators and learning functions was used. The results of each of the modules are fused in the last step for obtaining the final decision. Depending on the combination between integrator and learning function, the results vary between 88.4% and 97.47% rank-1 performance on the USTB I database. The highest rank-1 performance is achieved with Sugeno measure and conjugate gradient.

In [76] Nasseem *et al.* propose a general classification algorithm based on the theory of compressive sensing. They assume that most signals are compressible in nature and that any compression function results in a sparse representation of this signal. In their experiments in the UND database and the FEUD database, Nasseem *et al.* show that their sparse representation method is robust against pose variations and varying lighting

Table 4: Summary of approaches for 3D ear recognition. Performance (Perf.) always refers to rank-1 performance.

| Publication | Comparison Method | Database | | Perf. |
		#Subj	#Img	
Cadavid *et al.* [84]	ICP and Shape from shading	462	NA	95%
Chen and Bannu [28]	Local Surface Patch	302	604	96.36%
Chen and Bhanu [27]	ICP Contour Matching	52	213	93.3%
Liu *et al.* [63]	Mesh PCA with neural network	60	600	NA
Liu and Zhang [81]	Slice Curve Matching	50	200	94.5%
Islam *et al.* [82]	ICP with reduced meshes	415	830	93.98%
Islam *et al.* [93]	Local Surface Features with ICP-Matching	415	830	93.5%
Passalis *et al.* [80]	Reference ear model with morphing	525	1031	94.4%
Yan and Bowyer [77]	ICP using voxels	369	738	97.3%
Yan and Bowyer [18]	ICP using Model Points	415	1386	97.8%
Zheng *et al.* [83]	Local Binary Patters	415	830	96.39%
Zhou *et. al.* [79]	Surface Patch Histogram and voxelization	415	830	98.6%, 1.6% EER

conditions. The rank-1 performance varied between 89.13% and 97.83%, depending on the dataset used in the experiment.

5 3D Ear Recognition

In 2D ear recognition pose variation and variation in camera position, so-called out-of-plane-rotations, are still unsolved challenges. A possible solution is using 3D models instead of photos as references, because a 3D representation of the subject can be adapted to any rotation, scale and translation. In addition to that, the depth information contained in 3D models can be used for enhancing the accuracy of an ear recognition system. However, most 3D ear recognition systems tend to be computationally expensive. In Table 4 all 3D ear recognition systems described in this section are summarized.

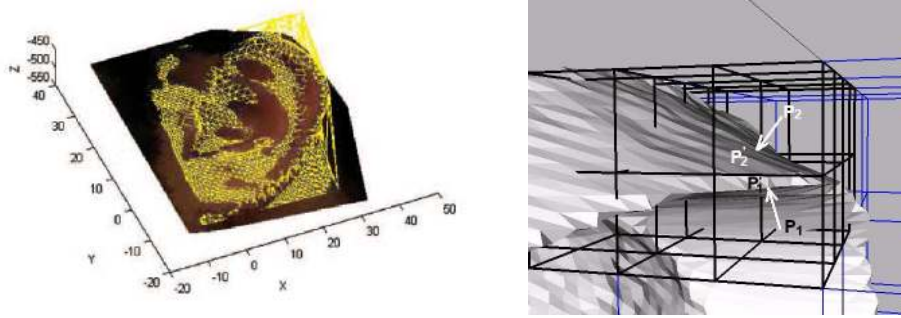Although ICP is originally designed to be an approach for image registration, the

Figure 8: Examples for surface features in 3D ear images. The upper image shows an example for ICP-based comparison as proposed in [27], whereas the lower figure illustrates feature extraction from voxels as described in [77].

registration error can also be used as a measure for the dissimilarity of two 3D images. Because ICP is designed to be a registration algorithm, it is robust against all kinds of translation or rotations. However ICP tends to stop too early, because it gets stuck in local minima. Therefore ICP requires the two models to be coarsely pre-aligned before fine alignment using ICP can be performed. Chen and Bhanu extract point clouds from the contour of the outer helix and the register these points with the reference model by using ICP [27]. In a later approach Chen and Bhanu use local surface patches (LSP) instead of points lying on the outer helix [28]. As the LSP consist or less points than the outer helix, this reduces the processing time while enhancing the rank-1 performance from 93.3% with the outer helix points to 96.63 % with LSP.

Yan and Browyer decompose the ear model into voxels and extract surface features from each of these voxels. For speeding up the alignment process, each voxel is assigned an index in such a way that ICP only needs to align voxel pairs with the same index [77] (see Figure 8). In [18] Yan and Browyer propose the usage of point clouds for 3D ear recognition. In contrast to [27] all points of the segmented ear model are used. The reported performance measures of 97.3% in [77] and 97.8% in [18] is similar but not directly comparable, because different datasets were used for evaluation.

Cadavid *et al.* propose a real-time ear recognition system, which reconstructs 3D models from 2D CCTV images using the shape from shading technique [78]. Thereafter the 3D model in compared to the reference 3D images, which are stored in the gallery. Model alignment as well as the computation of the dissimilarity measure is done by ICP. Cadavid *et al.* report a recognition rate of 95% on a database of 402 subjects. It is stated in [78] that the approach has difficulties with pose variations. In [79] Zhou *et. al.* use a combination of local histogram features voxel-models. Zhou *et. al.* report that their approach is faster and with an EER of 1.6% it is also more accurate than the ICP-based comparison algorithms proposed by Chen and Bhanu and Yan and Browyer.

Similarly to Cadavid *et al.*, Liu *et al.* reconstruct 3D ear models from 2D views [63]. Based on the two images of a stereo vision camera, a 3D representation of the ear is derived. Subsequently the resulting 3D meshes serve as the input for PCA. However Liu

et al. do not provide any results concerning the accuracy of their system but since they did not publish any further results on their PCA mesh approach, it seems that it is no longer pursued.

Passalis *et al.* go a different way for comparing 3D ear models in order to make comparison suitable for a real-time system [80]. They compute a reference ear model which is representative for the average human ear. During enrolment, all reference models are deformed until they fit the reference ear model. All translations and deformations, which were necessary to fit the ear to the reference model are then stored as features. If a probe for authentication is given to the system, the model is also adapted to the annotated ear model in order to get the deformation data. Subsequently the deformation data is used to search for an associated reference model in the gallery. In contrast to the previously described systems, only one deformation has to be computed per authentication attempt. All other deformation models can be computed before the actual identification process is started. This approach is reported to be suitable for real-time recognition systems, because it takes less than 1 msec for comparing two ear templates. The increased computing speed is achieved by lowering the complexity class from $O(n)^2$ for ICP-based approaches to $O(n)$ for their approach. The rank-1 recognition rate is reported to be 94.4%. The evaluation is based on non-public data, which was collected using different sensors.

Heng and Zhang propose a feature extraction algorithm based on slice curve comparison, which is inspired by the principles of computer tomography [81]. In their approach the 3D ear model is decomposed into slices along the orthogonal axis of the longest distance between the lobule and the uppermost part of the helix. The curvature information extracted from each slice is stored in a feature vector together with an index value indicating the slice's former position in the 3D model. For comparison the longest common sequence between two slice curves with similar indexes is determined. Their approach is only evaluated on a non-public dataset, which consists of 200 images from 50 subjects. No information about pose variations or occlusion during the capturing experiment is given. Heng and Zhang report a rank-1 performance is 94.5% for the identification experiment and 4.6%EER for the verification experiment.

Islam *et al.* reconnect point clouds describing 3D ear models to meshes and iteratively reduce the number of faces in the mesh [82]. These simplified meshes are then aligned with each other using ICP and the alignment error is used as the similarity measure for the two simplified meshes. In a later approach Islam *et al.* extract local surface patches as shown in Figure 9 and use them as features [34]. For extracting those LSP, a number of points is selected randomly from the 3D model. Then the data points which are closer to the seed point than a defined radius are selected. PCA is then applied to find the most descriptive features in the LSP. The feature extractor repeats selecting LSP until the desired number of features has been found. Both approaches were evaluated using images from UND. The recognition rate reported for [82] is 93.98% and the recognition rate reported for [34] is 93.5%. However, none of the approaches has been tested with pose variation and different scaling.

Zheng *et al.* extract the shape index at each point in the 3D model and use it for projecting the 3D model to 2D space [83]. The 3D shape index at each pixel is represented

by a grey value at the corresponding position in the 2D image. Then SIFT features are extracted from the shape index map. For each of the SIFT points a local coordinate system is calculated where the z-axis correspondents to the feature point's normal. Hence the z-values of the input image are normalized according to the normal of the SIFT feature point they were assigned to. As soon as the z values have been normalized, they are transformed into a grey level image. As a result, Zheng *et al.* get a local grey level image for each of the selected SIFT features. Next LBP are extracted for feature representation in each of these local grey level images. Comparison is first performed by coarsely comparing the shape indexes of key pints and then using Earth mover's distance for comparing LBP histograms from the corresponding normalized grey images. Zheng *et al.* evaluated their approach on a subset of the UND-J2 Collected and achieved a rank-1 performance of 96.39%.
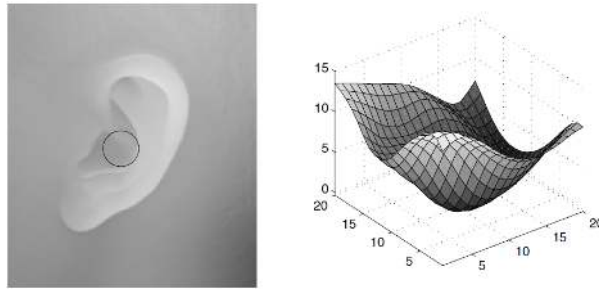

Figure 9: Example for local surface (LSP) patch features as proposed in [34]

6 Open challenges and future applications

As the most recent publications on 2D and 3D ear recognition show, the main application of this technique is personal identification in unconstrained environments. This includes applications for smart surveillance, such as in [84] but also the forensic identification of perpetrators on CCTV images or for border control systems. Traditionally these application fields are part of face recognition systems but as the ear is located next to the face, it can provide valuable additional information to supplement the facial images.

Multi modal ear and face recognition systems can serve as a means of achieving pose invariance and more robustness against occlusion in unconstrained environments. In most public venues surveillance cameras are located overhead in order to capture as many persons as possible and to protect them from vandalism. In addition, most of the persons will not look straight into the camera, so in most cases no frontal images of the persons will be available. This fact poses serious problems to biometric systems, using facial features for identification. If the face is not visible from a frontal angle, the ear can serve as a valuable additional characteristic in these scenarios.

Because of the physical proximity of the face and the ear, there are also many possibilities for the biometric fusion of these two modalities. Face and ear images can be fused on the feature level, on the template level and on the score level. Against the background

of this application, there are some unsolved challenges, which should be addressed by future research in this field.

6.1 Automatic Ear Localization

The fact that many systems presented in literature use pre-segmented ear images shows, that the automatic detection of ears especially in real-life images is still an unsolved problem. If ear recognition systems should be implemented in automatic identification systems, fast and reliable approaches for automatic ear detection are of importance. As a first step towards this goal, some research groups have published data collections, which simulate typical variations in uncontrolled environments such as varying lighting conditions, poses and occlusion. Based on these datasets, existing and future approaches to ear recognition should be tested under realistic conditions in order to improve their reliability.

Moreover, 3D imaging systems become increasingly cheap in the last years. Consequently 3D ear recognition becomes important and with it the need of locating ears in depth images or 3D models. Currently, only one approach for ear detection in depth has been published, which is a first step towards ear detection in 3D images.

6.2 Occlusion and Pose Variations

In contrast to the face, the ear can be partially or fully covered by hair or by other items such as headdresses, hearing aids, jewelry or headphones. Because of the convex surface of the outer ear, parts of it may also be occluded if the subject's pose changes. In some publications, robustness against occlusion is explicitly addressed, but there are no studies on the effect of the effect of certain types of occlusion like hair or earrings on the recognition rate of an ear recognition system. Once more, the availability of public databases which contain occluded ear images is likely to foster the development of solutions for pose invariant and robust algorithms for ear detection and feature extraction.

Moreover to our best knowledge there are no studies about the visibility of the outer ear in different public environments. In order to develop algorithms for ear detection and recognition, further information about commonly occluded party of the ear is needed.

Occlusion due to pose variations is another unmet challenge in ear recognition system. Similarly to face recognition, parts of the ear can become occluded if the pose changes. Recently, some feature extraction methods have been proposed, which are robust against pose variations to some degree. However, this issue is not fully solved yet. Another possibility for compensating pose variations could be the usage of 3D models instead of depth images of photographs.

6.3 Scalability

Currently available databases only consist of less than 10 000 ear images. The only exception is the USTB IV collection, which has not been released for the public yet. In realistic environments the size of the database will be significantly larger, which makes exhaustive search in identification scenarios infeasible. Therefore, not only the accuracy

but also the comparison speed of ear recognition systems will be interesting for future research.

In order to make ear recognition applicable for large scale systems, exhaustive searches should be replaced by appropriate data structures allowing logarithmic time complexity during the search. This could for example be achieved by exploring the possibilities of organizing ear templates in search trees.

6.4 Understanding Symmetry and Ageing

Because ear recognition is one of the newer fields of biometric research, the symmetry of the left and the right ear has not been fully understood yet. The studies of Iannarelli indicate that some characteristics of the outer ear can be inherited and ageing slightly affects the appearance of the outer ear. Both assumptions could be confirmed in more recent studies, but because of a lack of sufficient data, the effect of inheritance and ageing on the outer ear's appearance is not fully understood yet. Furthermore, there are no large scale studies of the symmetry relation between the left and the right ear yet.

Therefore another interesting field for future research could be, to gain a deeper understanding of the effect of inheritance any symmetry on the distinctiveness of biometric template. Moreover, long term studies on the effect of time on ear templates are needed in order to get a better understanding of the permanence of this characteristic.

7 Summary

We have presented a survey on the state of the art in 2D and 3D ar biometrics, covering ear detection and ear recognition systems. We categorized the large number of 2D ear recognition approaches into holistic, local, hybrid and statistical methods, discussed their characteristics and reported their performance.

Ear recognition is still a new field of research. Although there is a number of promising approaches, none of them has been evaluated under realistic scenarios which include disruptive factors like pose variations, occlusion and varying lighting conditions. In recent approaches, these factors are taken under account, but more research on this is required until ear recognition systems can be used in practice. The availability of suitable test databases, which were collected under realistic scenarios, will further contribute to the maturation of the ear as a biometric characteristic.

We have collected a structured survey of available databases, existing ear detection and recognition approaches and unsolved problems for ear recognition in the context of smart surveillance system, which we consider to be the most important application for ear biometrics. We think that this new characteristic is a valuable extension for face recognition systems on the way to pose invariant automatic identification.

References

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

[2] Bertillon A. 'La Photographie Judiciaire: Avec Un Appendice Sur La Classification Et L'Identification Anthropometriques'. Gauthier-Villars, Paris; 1890.

[3] Imhofer R. 'Die Bedeutung der Ohrmuschel für die Feststellung der Identität'. Archiv für die Kriminologie. 1906;26:150–163.

[4] Iannarelli AV. 'Ear identification'. Paramont Publishing Company; 1989.

[5] Meijerman L, Sholl S, Conti FD, Giacon M, van der Lugt C, Drusini A, et al. 'Exploratory study on classification and individualisation of earprints'. Forensic Science International. 2004;140(1):91 – 99.

[6] Singh P, Purkait R. 'Observations of external earAn Indian study'. HOMO - Journal of Comparative Human Biology. 2009;60(5):461 – 472. Available from: http://www.sciencedirect.com/science/article/pii/S0018442X09001164.

[7] Sforza C, Grandi G, Binelli M, Tommasi DG, Rosati R, Ferrario VF. 'Age- and sex-related changes in the normal human ear'. Forensic Science International. 2009;187(1-3):110.e1 – 110.e7. Available from: http://www.sciencedirect.com/science/article/pii/S0379073809000966.

[8] Meijerman L, Van Der Lugt C, Maat GJR. 'Cross-Sectional Anthropometric Study of the External Ear'. Journal of Forensic Sciences. 2007;52(2):286–293.

[9] Ibrahim MIS, Nixon MS, Mahmoodi S. 'The effect of time on ear biometrics'. In: International Joint Conference on Biometrics (IJCB); 2011. p. 1 –6.

[10] Hoogstrate AJ, Heuvel HVD, Huyben E. 'Ear identification based on surveillance camera images'. Science & Justice. 2001;41(3):167 – 172.

[11] Alberink I, Ruifrok A. 'Performance of the FearID earprint identification system'. Forensic Science International. 2007;166(2-3):145 – 154.

[12] Islam SMS, Bennamoun M, Owens R, Davies R. 'Biometric Approaches of 2D-3D Ear and Face: A Survey'. In: Sobh T, editor. Advances in Computer and Information Sciences and Engineering. Springer Netherlands; 2008. p. 509–514.

[13] Choras M. 'Image Feature Extraction Methods for Ear Biometrics–A Survey'. In: Computer Information Systems and Industrial Management Applications, 2007. CISIM '07. 6th International Conference on; 2007. p. 261 –265.

[14] Pun KH, Moon YS. 'Recent Advances in Ear Biometrics'. In: Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on; 2004. p. 164 – 169.

[15] Lammi HK. 'Ear Biometics'. Lappeenranta University of Technology, Department of Information Technology; 2004.

[16] Ramesh KP, Rao KN. 'Pattern extraction methods for ear biometrics - A survey'. In: Nature Biologically Inspired Computing, 2009. NaBIC 2009. World Congress on; 2009. p. 1657 –1660.

[17] Abaza A, Ross A, Herbert C, Harrison MAF, Nixon M. A Survey on Ear Biometrics; 2011. Accepted at ACM Computing Surveys. Available from: `http://eprints.ecs.soton.ac.uk/22951/`.

[18] Yan P, Bowyer KW. 'Biometric Recognition Using 3D Ear Shape'. Pattern Analysis and Machine Intelligence. 2007 August;29:1297 – 1308.

[19] Faltemier TC, Bowyer KW, Flynn PJ. 'Rotated Profile Signatures for robust 3D feature detection'. In: 8th IEEE International Conference on Automatic Face and Gesture Recognition; 2008. .

[20] Frejlichowski D, Tyszkiewicz N. 'The West Pomeranian University of Technology Ear Database A Tool for Testing Biometric Algorithms'. In: Campilho A, Kamel M, editors. Image Analysis and Recognition. vol. 6112 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg; 2010. p. 227–234.

[21] Kumar A, Wu C. 'Automated human identification using ear imaging'. Pattern Recogn. 2012 March;45(3):956–968. Available from: `http://dx.doi.org/10.1016/j.patcog.2011.06.005`.

[22] Prakash S, Gupta P. 'An Efficient Ear Recognition Technique Invariant to Illumination and Pose'. Telecommunication Systems Journal, special issue on Signal Processing Applications in Human Computer Interaction. 2011;30:38–50.

[23] Grgic M, Delac K, Grgic S. 'SCface – surveillance cameras face database'. Multimedia Tools Appl. 2011 February;51(3):863–879. Available from: `http://dx.doi.org/10.1007/s11042-009-0417-2`.

[24] Al Nizami HA, Adkins-Hill JP, Zhang Y, Sullins JR, McCullough C, Canavan S, et al. 'A biometric database with rotating head videos and hand- drawn face sketches'. In: Proceedings of the 3rd IEEE international conference on Biometrics: Theory, applications and systems. BTAS'09. Piscataway, NJ, USA: IEEE Press; 2009. p. 38–43. Available from: `http://dl.acm.org/citation.cfm?id=1736406.1736412`.

[25] Raposo R, Hoyle E, Peixinho A, Proenca H. 'UBEAR: A dataset of ear images captured on-the-move in uncontrolled conditions'. In: Computational Intelligence

in Biometrics and Identity Management (CIBIM), 2011 IEEE Workshop on; 2011. p. 84 –90.

[26] Chen H, Bhanu B. 'Shape Model-Based 3D Ear Detection from Side Face Range Images'. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops (CVPR); 2005. p. 122.

[27] Chen H, Bhanu B. 'Contour Matching for 3D Ear Recognition'. In: Proceedings of the Seventh IEEE Workshop on Applications of Computer Vision (WACV/MOTION); 2005. .

[28] Chen H, Bhanu B. 'Human Ear Recognition in 3D'. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2007 April;29(4):718 –737.

[29] Attarchi S, Faez K, Rafiei A. 'A New Segmentation Approach for Ear Recognition'. In: Blanc-Talon J, Bourennane S, Philips W, Popescu D, Scheunders P, editors. Advanced Concepts for Intelligent Vision Systems. vol. 5259 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg; 2008. p. 1030–1037.

[30] Ansari S, Gupta P. 'Localization of Ear Using Outer Helix Curve of the Ear'. In: International Conference on Computing: Theory and Applications; 2007. p. 688 –692.

[31] Prakash S, Gupta P. An Effient Ear Localization Technique. Image and Vision Computing. 2012;30:38–50.

[32] Prakash S, Gupta P. 'An Efficient Technique for Ear Detection in 3D: Invariant to Rotation and Scale'. In: The 5th IAPR International Conference on Biometrics (ICB); 2012. .

[33] Abaza A, Hebert C, Harrison MAF. 'Fast Learning Ear Detection for Real-time Surveillance'. In: Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS 2010); 2010. p. 1 –6.

[34] Islam S, Davies R, Bennamoun M, Mian A. 'Efficient Detection and Recognition of 3D Ears'. International Journal of Computer Vision. 2011;95:52–73.

[35] Yuan L, Mu ZC. 'Ear Detection Based on Skin-Color and Contour Information'. In: International Conference on Machine Learning and Cybernetics. vol. 4; 2007. p. 2213 –2217.

[36] Alvarez L, Gonzalez E, Mazorra L. 'Fitting ear contour using an ovoid model'. In: 39th Annual 2005 International Carnahan Conference on Security Technology (CCST '05); 2005. .

[37] Zhou J, Cadavid S, Abdel-Mottaleb M. 'Histograms of Categorized Shapes for 3D ear detection'. In: International Conference on Biometrics: Theory Applications and Systems; 2010. .

[38] Arbab-Zavar B, Nixon M. 'On Shape-Mediated Enrolment in Ear Biometrics'. In: Bebis G, Boyle R, Parvin B, Koracin D, Paragios N, Tanveer SM, et al., editors. Advances in Visual Computing. vol. 4842 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg; 2007. p. 549–558.

[39] Alastair H, Cummings AH, Nixon MS, Carter JN. 'A Novel Ray Analogy for Enrolment of Ear Biometrics'. In: Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS); 2010. .

[40] Kumar A, Hanmandlu M, Kuldeep M, Gupta HM. 'Automatic Ear Detection for Online Biometric Applications'. In: Third National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG); 2011. p. 146 –149.

[41] Hurley DJ, Nixon MS, Carter JN. 'Force Field Energy Functionals for Image Feature Extraction'. Image and Vision Computing. 2002;20(5-6):311 – 317.

[42] Abdel-Mottaleb M, Zhou J. 'Human Ear Recognition from Face Profile Images'. In: Zhang D, Jain A, editors. Advances in Biometrics. vol. 3832 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg; 2005. p. 786–792.

[43] Dong J, Mu Z. 'Multi-Pose Ear Recognition Based on Force Field Transformation'. In: Second International Symposium on Intelligent Information Technology Application (IITA). vol. 3; 2008. p. 771 –775.

[44] Wang X, Yuan W. 'Gabor Wavelets and General Discriminant Analysis for Ear Recogniton'. In: 8th World Congress on Intelligent Control and Automation (WCICA); 2010. p. 6305.

[45] Arbab-Zavar B, Nixon M. 'Robust Log-Gabor Filter for Ear Biometrics'. In: International Conference on Pattern Recognition (ICPR); 2008. .

[46] Abate AF, Nappi M, Riccio D, Ricciardi S. 'Ear Recognition by means of a Rotation Invariant Descriptor'. In: 18th International Conference on Pattern Recognition, ICPR 2006.. vol. 4; 2006. p. 437 –440.

[47] Sana PPR Anupam; Gupta. 'Ear Biometrics: A New Approach'. In: Advances in Pattern Recognition; 2007. .

[48] De Marsico M, Michele N, Riccio D. 'HERO: Human Ear Recognition against Occlusions'. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW); 2010. p. 178.

[49] Wang Xq, Xia Hy, Wang Zl. 'The Research of Ear Identification Based On Improved Algorithm of Moment Invariants'. In: Third International Conference on Information and Computing (ICIC); 2010. p. 58.

[50] Wang X, Yuan W. 'Human Ear recognition Based on Block Segmentation'. In: International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC); 2009. p. 262 –266.

[51] Lowe GD. 'Object recognition from local scale-invariant features'. In: IEEE International Conference on Computer Vision (ICCV 1999). vol. 2. IEEE Computer Society; 1999. p. 1150 –1157.

[52] Bustard JD, Nixon MS. 'Toward Unconstrained Ear Recognition From Two-Dimensional Images'. Systems, Man and Cybernetics, Part A: Systems and Humans. 2010 April;40:486.

[53] Arbab-Zavar B, Nixon MS, Hurley DJ. 'On Model-Based Analysis of Ear Biometrics'. In: First IEEE International Conference on Biometrics: Theory, Applications, and Systems, 2007. (BTAS 2007); 2007. p. 1 –5.

[54] Badrinath G, Gupta P. 'Feature Level Fused Ear Biometric System'. In: Seventh International Conference on Advances in Pattern Recognition (ICAPR); 2009. p. 197 –200.

[55] Kisku DR, Mehrotra H, Gupta P, Sing JK. 'SIFT-based ear recognition by fusion of detected keypoints from color similarity slice regions'. In: International Conference on Advances in Computational Tools for Engineering Applications (ACTEA); 2009. p. 380 –385.

[56] Bay H, Tuytelaars T, Gool LV. 'SURF: Speeded Up Robust Features'. In: Proceedings of the 9th European Conference on Computer Vision; 2006. .

[57] Choras M. 'Perspective Methods of Human Identification: Ear Biometrics'. Opto-Electronics Review. 2008;16:85–96.

[58] Mu Z, Yuan L, Xu Z, Xi D, Qi S. 'Shape and Structural Feature Based Ear Recognition'. In: Li S, Lai J, Tan T, Feng G, Wang Y, editors. Advances in Biometric Person Authentication. vol. 3338 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg; 2005. p. 311–364.

[59] Rahman M, Islam R, Bhuiyan NI, Ahmed B, Islam A. 'Person Identification Using Ear Biometrics'. International Journal of The Computer, The Internet and Management. 2007 August;15:1–8.

[60] Guo Y, Xu Z. 'Ear Recognition Using a New Local Matching Approach'. In: 15th IEEE International Conference on Image Processing (ICIP); 2008. p. 289 –292.

[61] Wang Zq, Yan Xd. 'Multi-Scale Feature Extraction Algorithm of Ear Image'. In: International Conference on Electric Information and Control Engineering (ICEICE); 2011. p. 528.

[62] Jeges E, Mt L. 'Model-Based Human Ear Localization and Feature Extraction'. IC-MED. 2007;1(2):101–112.

[63] Liu H, Yan J. 'Multi-view Ear Shape Feature Extraction and Reconstruction'. In: Third International IEEE Conference on Signal-Image Technologies and Internet-Based System (SITIS); 2007. p. 652 –658.

[64] Mukundan R, Ong SH, Lee PA. 'Image analysis by Tchebichef moments'. IEEE Transactions on Image Processing. 2001 September;10(9):1357 –1364.

[65] Lu L, Xiaoxun Z, Youdong Z, Yunde J. 'Ear Recognition Based on Statistical Shape Model'. In: First International Conference on Innovative Computing, Information and Control (ICICIC); 2006. p. 353 – 356.

[66] Yuan L, Mu Z. 'Ear Recognition Based on 2D Images'. In: First IEEE International Conference on Biometrics: Theory, Applications, and Systems (BTAS); 2007. p. 1 –5.

[67] Victor B, Bowyer K, Sarkar S. 'An evaluation of face and ear biometrics'. In: 16th International Conference on Pattern Recognition (ICPR). vol. 1; 2002. p. 429 – 432 vol.1.

[68] Chang K, Bowyer KW, Sarkar S, Victor B. 'Comparison and Combination of Ear and Face Images in Appearance-Based Biometrics'. IEEE Transactions in Pattern Anallysis and Machine Intelligene. 2003 September;25:1160–1165.

[69] Alaraj M, Hou J, Fukami T. 'A neural network based human identification framework using ear images'. In: TENCON 2010 - 2010 IEEE Region 10 Conference; 2010. .

[70] Zhang H, Mu Z. 'Compound Structure Classifier System for Ear Recognition'. In: IEEE International Conference on Automation and Logistics (ICAL); 2008. p. 2306 –2309.

[71] Xie Z, Mu Z. 'Ear Recognition Using LLE and IDLLE Algorithm'. In: 19th International Conference on Pattern Recognition (ICPR); 2008. p. 1 –4.

[72] Nanni L, Lumini A. 'A Multi-Matcher For Ear Authentication'. Pattern Recognition Letters. 2007 December;28:2219–2226.

[73] Yuizono T, Wang Y, Satoh K, Nakayama S. 'Study on Individual Recognition for Ear Images by using Genetic Local Search'. In: Proceedings of the 2002 Congress on Processing Scociety of Japan (IPSJ) Kyushu Chapter Symposium; 2002. p. 237–242.

[74] Yaqubi M, Faez K, Motamed S. 'Ear Recognition Using Features Inspired by Visual Cortex and Support Vector Machine Technique'. In: International Conference on Computer and Communication Engineering (ICCCE); 2008. p. 533 –537.

[75] Gutierrez L, Melin P, Lopez M. 'Modular Neural Network Integrator for Human Recognition From Ear Images'. In: The 2010 International Joint Conference on Neural Networks (IJCNN); 2010. .

[76] Naseem I, Togneri R, Bennamoun M. 'Sparse Representation for Ear Biometrics'. In: Bebis G, Boyle R, Parvin B, Koracin D, Remagnino P, Porikli F, et al., editors. Advances in Visual Computing. vol. 5359 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg; 2008. p. 336–345.

[77] Yan P, Bowyer KW. 'A Fast Algorithm for ICP-Based 3D Shape Biometrics'. In: Fourth IEEE Workshop on Automatic Identification Advanced Technologies; 2005. p. 213 – 218.

[78] Cadavid S, Mahoor MH, Abdel-Mottaleb M. 'Multi-Modal Biometric Modeling and Recognition of the Human Face and Ear'. In: IEEE International Workshop on Safety, Security Rescue Robotics (SSRR); 2009. p. 1 –6.

[79] Zhou J, Cadavid S, Abdel-Mottaleb M. 'A computationally Efficient Approach to 3D Ear Recognition Employing Local and Holistic Features'. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW); 2011. p. 98 –105.

[80] Passalis G, Kakadiaris IA, Theoharis T, Toderici G, Papaioannou T. 'Towards Fast 3D Ear Recognition For Real-Life Biometric Applications'. In: IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS 2007); 2007. p. 39 –44.

[81] Liu H, Zhang D. 'Fast 3D Point Cloud Ear Identification by Slice Curve Matching'. In: 3rd International Conference on Computer Research and Development (ICCRD); 2011. p. 224.

[82] Islam SMS, Bennamoun M, Mian AS, Davies R. 'A Fully Automatic Approach for Human Recognition from Profile Images Using 2D and 3D Ear Data'. In: Proceedings of 3DPVT - the Fourth Internatinoal Symposium on 3D Data Processing, Visualization and Transmission; 2008. .

[83] Zeng H, Dong JY, Mu ZC, Guo Y. 'Ear Recognition Based on 3D Keypoint Matching'. In: IEEE 10th International Conference on Signal Processing (ICSP); 2010. p. 1694.

[84] Cadavid S, Abdel-Mottaleb M. '3D ear modeling and recognition from video sequences using shape from shading'. In: 19th International Conference on Pattern Recognition (ICPR); 2008. p. 1 –4.

[85] Islam SMS, Bennamoun M, Davies R. 'Fast and Fully Automatic Ear Detection Using Cascaded AdaBoost'. In: Applications of Computer Vision, 2008. WACV 2008. IEEE Workshop on; 2008. p. 1 –6.

[86] Liu H, Liu D. 'Improving Adaboost Ear Detection with Skin-color model and Multi-template Matching'. In: 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT). vol. 8; 2010. p. 106 –109.

[87] Shih HC, Ho CC, Chang HT, Wu CS. 'Ear Detection Based on Arc-Masking Extraction and AdaBoost Polling Verification'. In: Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP). vol. ear detection; 2009. p. 669 –672.

[88] Burge M, Burger W. 13. In: Jain AK, Bolle R, Pankanti S, editors. 'Ear Biometrics'. Springer US; 1998. p. 273–285.

[89] Moreno A B andSanchez, Velez JF. 'On The Use of Outer Ear Images for Personal Identification in Security Applications'. In: IEEE 33rd Annual 1999 International Carnahan Conference on Security Technology; 2002. p. 469 – 476.

[90] Yuan L, chun Mu Z, Zhang Y, Liu K. 'Ear Recognition using Improved Non-Negative Matrix Factorization'. In: 18th International Conference on Pattern Recognition (ICPR). vol. 4; 2006. p. 501 –504.

[91] Wang Y, chun Mu Z, Zeng H. 'Block-Based and Multi-Resolition methods for Ear recognition Using Walelste Transform and Uniform Local Binary Patterns'. In: 19th International Conference on Pattern Recognition (ICPR); 2008. p. 1 –4.

[92] Fooprateepsiri R, Kurutach W. 'Ear Based Personal Identification Approach Forensic Science Tasks'. Chiang Mai Journal of Science. 2011 January;38(2):166–175.

[93] Islam SM, Davies R, Mian M A S an Bennamoun. 'A Fast and Fully Automatic Ear Recognition Approach Based on 3D Local Surface Features'. In: Proceedings of the 10th International Conference on Advanced Concepts for Intelligent Vision Systems. ACIVS '08. Berlin, Heidelberg: Springer-Verlag; 2008. p. 1081–1092.