# Multi-model Ear Database for Biometric Applications

**Atanas Nikolov, Virginio Cantoni, Dimo Dimov, Andrea Abate and Stefano Ricciardi**

**Abstract** We present the 3DEarDB, a multi-model ear database, characterized by different types of ear representation, either 2D or 3D, depending on the acquisition device used. The main objective is to provide the biometrics community with a unified tool for testing and comparing of classification algorithms not only on 2D intensity and/or depth images, or videos, but also on detailed 3D mesh models of human ears. The 3DEarDB features accurate 3D mesh models of right ear captured from more than 100 subjects, with a resolution of 1 mm and an accuracy of 0.05 mm, collected via the VIUscan 3D laser scanner, available at the Smart Lab of IICT-BAS, in the AComIn project frames. Two more ear acquisition modalities are also included: 3D Kinect ear depth maps and 2D high-definition video clips, associated to the basic mesh models. To extend 3DEarDB compatibilities with known methods for 2D/3D ear detection and/or recognition, we provide two more ear model types. Namely, a set of 2D ear intensity projections (of different orientations and/or lightening directions), and a set of 2D depth map projections can be generated by demand from the basic 3D ear models. Finally, we report about preliminary experiments conducted by means of Extended Gaussian Image approach that confirm the consistency of the proposed 3D-Ear-Data-Base.

A. Nikolov · D. Dimov (✉)
Institute of Information and Communication Technologies (IICT),
Bulgarian Academy of Sciences (BAS), Sofia, Bulgaria
e-mail: dtdim@iinf.bas.bg

A. Nikolov
e-mail: a.nikolov@iinf.bas.bg

V. Cantoni
Department of Industrial and Information Engineering, Pavia University, Pavia, Italy
e-mail: virginio.cantoni@unipv.it

A. Abate · S. Ricciardi
Department of Computer Science, Salerno University, Salerno, Italy
e-mail: abate@unisa.it

S. Ricciardi
e-mail: sricciardi@unisa.it

## 1 Introduction

The usage of biometric identifiers as a reliable and convenient way to verifying a
person's identity has become common worldwide in the last decade, with particular
regard to the most established ones like fingerprint, face and, more recently, iris.
A key factor in diffusion of a biometric entity is its acceptability, since this char-
acteristic directly affects the range of applications and the extent of the provided
advantages in the context of both validation and identification [1]. In addition,
aspects like stability over time and reduced intra class variations have been proved
relevant in determining the success of biometrics-based id-check solutions. To these
regards, ear seems to be a convenient biometric feature since it combines good
distinctiveness, as indirectly proved by the high recognition accuracy achieved [2–
4], with high acceptability (since is captured without the need for a physical con-
tact) and permanence. The human ear was first hypothesized as a salient identifier in
the end of XIX century by the French criminologist A. Bertillon [5], but only in
1949 A. Iannarelli proposed, with a more scientific approach, a set of twelve
measurements characterizing the ear geometry [6]. The clear advantages in using
ear biometrics are related to its tridimensional (3D) structure protruding from the
overall head surface/profile (when observed frontally) that allows for simple and
contactless capture by means of 2D and 3D techniques. Ear is characterized by
easily recognizable ridges and valleys, whose configuration is relatively immune to
variation due to ageing [7]. The almost complete absence of shape changes rep-
resents another advantage of this biometrics whose main intra-class variations
derive by occlusions caused by hair, hats, earrings, etc., [8].

Though the number of contributions delivered by the research community on the
topic of ear recognition are not comparable to the effort produced so far for face,
fingerprint or even iris, many different methods and algorithms have been proposed
with both 2D and 3D approaches over the last 15 years. 2D methods have exploited
a variety of descriptors, including Principal Component Analysis (PCA) [9, 10],
Independent Component Analysis (ICA) [11], Active Shape Model (ASM) [12],
sparse representations [13], force fields [2, 14, 15], ear geometries [16, 17], Generic
Fourier Descriptor (GFD) [18], wavelet transforms [3, 19, 20], Local Binary
Patterns (LBP) [21], Gabor filters [22] and Scale-Invariant Feature Transform
(SIFT) [23, 24].

The first 3D method [25] was proposed in 2004 and exploited the Local Surface
Patch (LSP) representation and the Iterative Closest Point (ICP) algorithm, that was
also used [4, 26, 27] for matching ears models obtained as range images or 3D
mesh. A 2.5D approach was explored using surveillance videos and pseudo 3D
information extracted by means of Shape-from-Shading (SFS) scheme [28]. It is
worth to mention also two recent approaches to 3D ear recognition, based on the

EGI representation of 3D ear models [29], and on the 2D appearance 3D multi-view approach [30], in which additional related works are surveyed. A detailed and recent survey on Ear processing and recognition can be found in [31], as well as in [32, 33].

A crucial aspect of the research around ear biometrics is represented by the availability of public ear databases to be used as a reference to test and stress proposed methods on a common set of images captured in known conditions, and to highlight the strengths and the weaknesses of each method and/or approach in terms of recognition accuracy and robustness. To this regard, a number of ear datasets have been publicly released through the last 10 years, along with the research works that led to their creation. They typically provide 2D pictures of the ear(s) isolated or as a part of face profiles (mostly captured in laboratory), and in a limited number of cases also 3D scans of the face region near to the ear. We provide details on the existing ear datasets in Sect. 2 of this paper. Since, currently there is still a lack of a multi-model ear database, providing a full spectrum of capturing modalities for each of the enrolled subjects, in this paper we present such a kind of ear dataset that features high resolution 3D scans for each subject (both, row data and a segmented, cleaned polygonal mesh), also high resolution color pictures, high resolution video capture from variable angles, color pictures captured by last-generation mobile devices and other indirect modalities derived by the 3D data (2D intensity, and depth images).

The rest of the paper is organized as follows. Section 2 presents a description of the existing, publicly available, ear datasets. Section 3 provides a detailed description of a new dataset developed with regard to all the provided models and their capture. Section 4 presents the results of the first batch of experiments conducted on the proposed dataset and, finally, Sect. 5 draws some conclusions.

## 2 Publicly Available Ear-Specific Datasets—A Brief Review

As recalled in the previous section, there is a small number of publicly-available ear-specific datasets released so far, at least if we do not consider well known face database like, the FERET database [34], the CAS-PEAL database [35], the UMIST database [36], the NIST Mugshot Identification Database (MID) [37] or the XM2VTS database [38] which, though not originally aimed at ear biometrics, have been used and cited in literature mostly for testing ear detection algorithms. The ear-specific datasets are the AMI Ear Database [39], the UBEAR dataset [40], the University of Notre Dame (UND) databases [41], the University of Science and Technology Beijing (USTB) Databases [42], as well as the most recent OpenHear database [43], and the SYMARE database [44]. They are briefly described in the following lines.

**AMI Ear Database** [39] consists of ear images collected from students, teachers and staff of the Computer Science department at Universidad de Las Palmas de Gran Canaria (ULPGC), Las Palmas, Spain. The 700 images provided have been captured solely in an indoor environment from 100 different subjects in the age range of 19–65 years. For each individual, seven images (six right ear images and one left ear image) are taken under the same lighting conditions, at a capture resolution of $492 \times 702$ pixels, with the subject seated at a distance of about 2 m from the camera. Five of the captured images are right side profile (right ear) with the individual facing forward, looking up and down, and looking left and right (Fig. 1).

**UBEAR Dataset** [40] represents the result of a research study focused on capturing ear images on the move in uncontrolled conditions, including ample variations of posing, lighting and presence of occlusions, to the aim of providing a real-world set of samples that should result very challenging for detection and recognition algorithms. The dataset is built by means of four high-resolution ($1280 \times 960$ pixels at 15 fps) video captures, two for each ear across two different sessions, requiring each subject to undergo the same enrollment protocol. From each video 17 frames (5 frames for stepping ahead and backwards + 12 frames for head movements in four directions, namely, 3 upwards, 3 downwards, 3 outwards, and 3 towards) are selected for each of the 126 subjects, acquired of whom 44.62 % are males and 55.38 % are females. The result database contains 4430 uncompressed gray-scale images, a few is shown in Fig. 2.

**UND Databases** [41] of the University of Notre Dame include a variety of biometric data in various modalities, organized in collections. The following four collections are relevant for ear biometrics:



BACK          FRONT          UP          DOWN          LEFT          RIGHT          ZOOM

**Fig. 1** Seven samples of two subjects captured from different directions (from AMI dataset)



**Fig. 2** Samples of different posing in the UBEAR dataset

- Collection E: 464 visible-light face side profile (ear) images from 114 human subjects captured in 2002.
- Collection F: 942 3D (+corresponding 2D) profile (ear) images from 302 human subjects captured in 2003 and 2004.
- Collection G: 738 3D (+corresponding 2D) profile (ear) images from 235 human subjects captured between 2003 and 2005.
- Collection J2: 1800 3D (+corresponding 2D) profile (ear) images from 415 human subjects captured between 2003 and 2005.

**USTB Databases** [42] of the University of Science and Technology Beijing represent four databases dedicated to ear biometrics:

- Image Database I (dated: July–Aug 2002) contains 180 grayscale images of right ear from 60 subjects, each one photographed three times including one frontal image, another one with slight angle and one more with different lighting condition.
- Image Database II (dated: Nov 2003–Jan 2004) contains 308, 300 × 400 pixels, 24bit color images of right ear from 77 subjects, each one photographed four times with one profile image, two different form angles and one with different lighting conditions.
- Image Database III (dated: 20 Nov–30 Dec 2004) contains two ear datasets, a dataset with regular ear images and another one with occluded ear images. The first dataset includes right side profiles captured at 768 × 576 pixels, 24 bit colors from 79 subjects captured from variable rotations: 22 rotation steps to the right and 18 to the left. The second dataset contains 144 images of partially occluded ears from 24 subjects. They obey three conditions: partial occlusions (disturbance from some hair), trivial occlusions (little hair), and regular (natural) occlusions.
- Image Database IV (dated: Jun 2007–Dec 2008) contains both grayscale and color ear images, 500 × 400 pixels each, from 500 subjects acquired from multiple angles by 17 CCD cameras distributed around the volunteer at a 15° step from each other.

**OpenHear**, the Open head and ear database [43], is an open database of 3D surface scans of human heads and ears. Its purpose is to be used for acoustical simulation in aid design. The dataset contains head and ear 3D models of 20 subjects (10 men, 7 women, 1 baby boy, and 2 girls), see part of them in Fig. 3. The scans (available in VTK format) are acquired using a 3dMD cranial scanner, placed



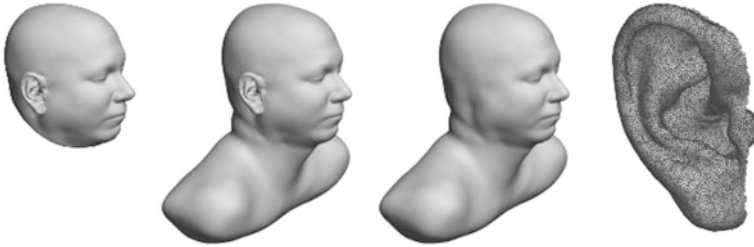**Fig. 3** Samples from the current version of OpenHear dataset

**Fig. 4** Samples from SYMARE: the four types of surface meshes provided per subject

at the 3D Craniofacial Image Research Laboratory at the University of Copenhagen. The initial 3D point clouds are created via 3dMD stereo-algorithms, while surface reconstructions are obtained using the authors algorithm to create complete head and ear models from initial captured data.

**SYMARE** [44], the Sydney York Morphological and Acoustic Recordings of Ears database, supports acoustics research exploring the relationship between the morphology of human outer ears and their acoustic filtering properties for purpose of improving the individualization of 3D audio for personal audio devices in the future. The database includes multiple mesh models (upper torso, head and ears) at varying resolutions for 61 listeners (48 male and 13 female) in order to accommodate acoustic stimulations at different frequencies. The 3D data are collected using a Philips 3T Achieva MRI scanner. For each of the 61 subjects in the database, high-resolution (sub-millimeter) surface meshes are provided for: (i) the head and ears, (ii) the head, upper torso and ears, (iii) the head and upper torso (no ears), (iv) the separated left and right ears, see Fig. 4. The number of surface elements involved in an average head and torso mesh is about 130 K elements.

## 3   Overview of Our 3D Ear Database

The announced 3D Ear Database, called here 3DEarDB, was collected mainly during the middle of 2015 at the Institute of Information and Communication Technologies at Bulgarian Academy of Sciences (IICT-BAS) in the frames of AComIn[1] project. We have gathered more than 100 precise 3D mesh models of right ears of persons, who differ in gender as well as in age (25–65). A scan resolution of 1 mm between neighboring 3D points and accuracy of 0.05 mm for each 3D point was chosen for simplicity of the data gathering, considering it to be enough for near future experiments. The first version of 3DEarDB (dated May, 2014) contained 3D ear models of the same precision but for 11 persons only, and

---

[1]http://www.iict.bas.bg/acomin/.

was designed for initial experiments with both our approaches to 3D ear classification and/or recognition, [29, 30].

The recent objective of 3DEarDB is to provide, in a consistent way, many different output formats for the given human (subject, person) ear represented. These includes: (i) a raw 3D ear mesh model, (ii) a processed 3D ear mesh, (iii) Kinect 3D ear depth (range) images, (iv) accompanying 2D ear video clips, (v) generated structures of 2D ear intensity projections, and (vi) generated structures of 2D ear depth images. This consistent variety of ear capturing formats could be very useful for ear biometrics community to test and compare algorithms accuracy on possibly different input scenarios—from the ideal case of precise (and static) 3D mesh to more realistic (and dynamic) case of 2D video data and/or still images.

By our best knowledge, cf. also Sect. 2, among the existent Ear Datasets, the only DB, which provides corresponding 2D and 3D data for the same subject's ear is that of UND Collections F, G, and J2, [41]. The UND 3D ear data do not represent real polygonal 3D meshes, but only 3D range images containing depth information. Moreover, the ear video data, which could be used for performing 3D ear reconstruction as an alternative to 2D range images, are missing there. The recent 3D databases, OpenHear [43] and SYMARE [44], really concern 3D ear data, but they are not designed especially for visual ear biometrics. Besides, neither OpenHear (only 20 face models), nor SYMARE even with its 61 listeners recorded and scanned, could be considered statistically enough representative at present.

An essential requirement of the large biometrics community is that such a DB has to top 100, or more, persons represented. We also consider ear biometrics based on video data as the most realistic case according to the contemporary technology development, especially if it is intended to be build-in the portable electronics of personal use. For this reason, it is useful to provide accurate 3D ear mesh representation as reference for evaluation of 3D video reconstruction errors, and for comparing between ideal and real recognition performances of investigated descriptors and classifiers. Because of we consider colors a non-informative ear feature for classification, we do not scan it at present. Colors are kept in the accompanying 2D ear video clips.

Next section contains a more detailed description of our multi-model Ear DB, considering two main types of ear data—hardware acquired and software generated. Hardware acquired ear representations are composed by raw and post-processed 3D ear meshes (from 3D laser scanners), 3D depth maps (from Kinect cameras), and 2D Video clips (from photo cameras). The software generated ear representations from each 3D mesh model are also two types at present, namely: (i) structures of images, i.e. 2D intensity projections with different lightening and/or orientation (using MeshLab[2]); and (ii) corresponding structures of 2D depth map projections with different orientation (using Wolfram Mathematica[3]).
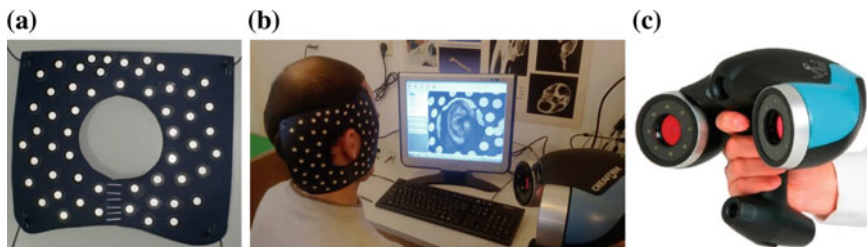
---

[2]http://meshlab.sourceforge.net/.

[3]https://www.wolfram.com/mathematica/.

**Fig. 5 a** The cartoon "helmet". **b** A person under scanning. **c** VIUscan 3D scanner

## 3.1 Data Acquisition

The three types of devices we use to collect ear data are described below. Only right ears data are gathered, and only one 3D ear model per subject is represented in 3DEarDB, because of limited people resource, for the time being. For more detail on this matter see also discussions in Sects. 4.2 and 5.

**VIUscan 3D Laser Scanner**. This hand-scanner of Creaform (Fig. 5c) was bought by the AComIn project for the Smart Lab of IICT-BAS in the end of 2013. Well computer assisted, it can reproduce a 3D mesh model of the scanned solid as well as respective textures and/or colors. Although, we have not used the maximal resolution (0.1 mm) and any color data, they could be very useful in other applications, where 3D objects have variable texture with fine surface details, [45].

This type of scanners require specific markers (retro-reflective targets) regularly situated on or around the object of scanning. The scanner needs to "see" at least four targets, which should not move in respect to the object of scan. VIUscan uses these targets to position itself in the space. To facilitate our work, we created a special "helmet" of cartoon with enough markers on it. The helmet is to be placed on the subject's head around the ear before scanning (Fig. 5a, b).

Omitting of color data makes the procedure of scanning faster, up to 10 min per ear, as well as more comfortable, because of no need of special lightening—possible shadows do not disturb scanning.

**Kinect Xbox One Sensor**.[4] This motion sensor of Microsoft is an upgraded version of its predecessor Xbox 360. Available as a standalone version since October 2014, it has an infrared array and a $512 \times 424$ pixels time-of-flight camera that resolves scene depth and allows for motion tracking and gesture recognition. This new Kinect also includes a Full HD ($1920 \times 1080$) video camera with increased field of view.

We plan to use Kinect for obtaining real depth maps of ears and to apply its accompanying software for 3D reconstruction (using video and/or depth maps).

---

[4]https://en.wikipedia.org/wiki/Kinect_for_Xbox_One#Specifications.
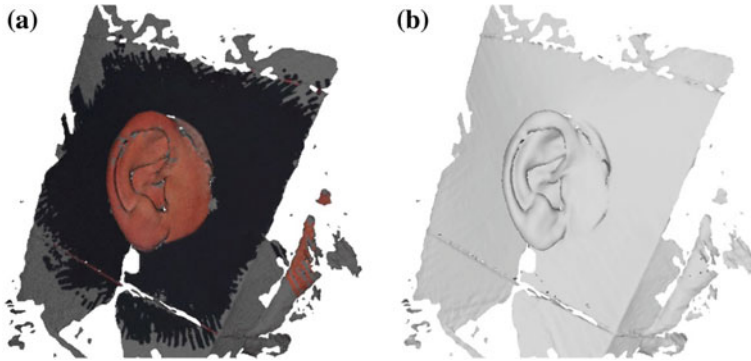
**Fig. 6 a** Raw scanned ear with color data. **b** Only the surface of the raw ear data

**Olympus Photo Camera**.[5] The Olympus SH-21 photo camera with its 16 MP CMOS sensor of 1/2.3″ format has been used for producing Full HD (1920 × 1080) video clips for each subject's ear, generally in a MP4 format file.

## 3.2 Raw (Unprocessed) Ear Data

A raw scanned ear, as shown on Fig. 6b, appears from VXelements software usually accompanying VIUscan scanners, [45]. The primary output file format is CSF, which size, in our case is about 64 MB per ear. VXelements help to convert each CSF to an OBJ format (an ASCII text) file for the ear geometry, and to an accompanying BMP file for the ear colors. In Fig. 6a we illustrate a colored ear scan, only for giving an idea of how it looks like, although not using it for now, as already mentioned. We use OBJ files at next (half-tone) post-processing, see Fig. 6b. Of course, color data could be successfully used for an automatic 3D ear segmentation, what is outside this work.

## 3.3 Raw Ear Data Post-processing

To create a complete and appropriately smooth 3D mesh model for each ear, we describe a post-processing of six steps using either VXelements [45] or MeshLab [46].

---

[5]http://www.olympus-global.com/en/news/2011b/nr111110sh21e.jsp.

**Step 1: Coarse Segmentation (by VXelements)**

- Apply the filter called *Remove Isolated Patches* on the input CSF data.
- Perform coarse manual segmentation of the ear surface from the surrounding background using the *Brush Selection*, *Reverse Selection*, and *Delete Facets* tools.

**Step 2: Holes Filling (by VXelements)**

- Run the *Optimize Surface* reconstruction algorithm each time when choosing a different size of ear holes to be filled-in. This procedure is the most time consuming, because of better results could not be predicted but experimented.
- After filling the appropriate holes, save the result CSF file (its size here is about 49 MB per ear). To continue with MeshLab processing, convert CSF to OBJ file that results in about 600 KB (per ear).

**Step 3: Fine Editing of Mesh-Facets (by MeshLab)**. It includes finer background segmentation, as well as removing unpleasant sharp peaks (Fig. 7a) in the current 3D mesh model resulting from the Optimize Surface tool of the previous step. Of course, the peak facets removal leads to new holes to fill-in (Fig. 7c), but of much smaller size (Fig. 7b), that is usually no problem for MeshLab.

**Step 4: Mesh Extra Smoothing (by MeshLab)**. After holes filling (Fig. 7c), the final step is smoothing the complete 3D object (Fig. 7d). The MeshLab function we prefer to this aim, is the *HC Laplacian Smooth*, based on the paper of Vollmer et al. [47]. At this final stage of manipulation, each ear mesh consists of about 6–8 thousands of (triangular) facets, determined by about 3–4 thousands of vertexes (3D points). Omitting the normal vectors data, considered here derivative and redundant ones for simplicity, the size of the respective OBJ file is reduced up to about 240 KB (per ear).

**Step 5: Mesh Decimation and Subdivision (by MeshLab)**. This step is necessary for creation of test data for our EGI classification approach [29], which we use to prove experimentally the 3DEarDB functionality. The MeshLab function for increasing the facets number (Fig. 8c) is called *Subdivision Surfaces: LS3 Loop*, based on [48], and the function reducing this number (Fig. 8a) is *Quadratic Edge Collapse Decimation*.
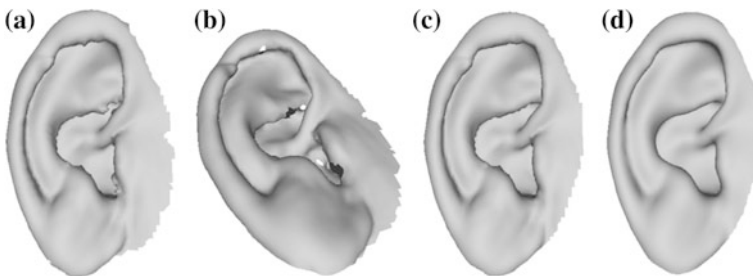


**Fig. 7** **a** Sharp peaks. **b** New holes created. **c** All holes filled. **d** Final smooth
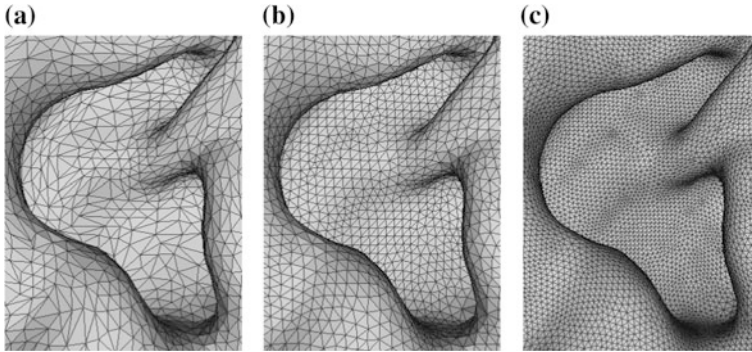
**Fig. 8 a** Decimated facets. **b** Original scan resolution. **c** Subdivided (refined) facets

**Step 6: Geometric Normalization (in MATLAB)**. It includes translation, orientation and scale of each ear model separately:

- Translate the Cartesian origin into the model barycenter, i.e. the averaged (x, y, z) coordinates of all 3D points (vertexes) of the mesh. After subtracting it from all vertexes, the new barycenter becomes (0, 0, 0).
- Rotate Principal axes, i.e. the eigenvectors of the covariance matrix over the whole mesh (all the vertexes). To normalize by rotation, the vertexes are rotated back to the already centralized Cartesian coordinate system, see also Fig. 9.
- Scale: The three eigenvalues (associated to principal axes, they should be already rotated) are used to normalize the mesh model by scale, so that the bounding box of the model (or its equivalent ellipsoid) to reach predefined sizes, e.g. 1-s (units). The three scale coefficients (reciprocal to eigenvalues) for each model have to be saved, if the real ear size will be further essential.

**Fig. 9** A normalized ear model
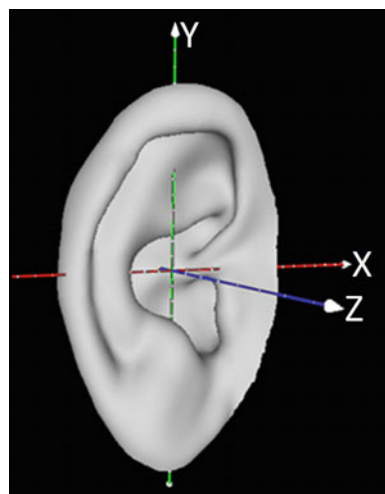
### 3.4  Kinect 3D Depth (Range) Images

At present, we do not give 3D ear data gathered by a Kinect camera. Instead, we have generated 2D depth-map images from 3DEarDB, as described in Sect. 3.7.

### 3.5  Full HD Ear Video Clips

A 1920 × 1080 video is made over each ear, uniformly filming it by azimuth from −80° to +80°, for 3 different altitude rows (upper, central, and lower ones) towards the center of the ear frontal view (Fig. 10), in the same laboratory, immediately after the 3D ear scan. Each clip is about 20 s long, at 30 fps that costs about 45 MB per clip, written in MP4 file format.

### 3.6  2D Intensity Projections

The 2D ear projections are produced in MeshLab, by loading a number of layers, one for each 3D rotation of an ear. Then, 2D snapshots of all these layers are made and recorded in JPEG format. The artificial lightening chosen is frontal and coherent.

The 2D intensity projections are taken according to a rotations scheme of 100 frontal view directions, uniformly distributed towards the ear barycenter, i.e. on 10 declinations and 10 azimuths uniformly chosen in the interval (−45°, +45°), cf. also Fig. 11. Of course, the angle step could be smaller or larger, in this way to manipulate the density of the resultant set of 2D projections, i.e. the size of output JPG files.
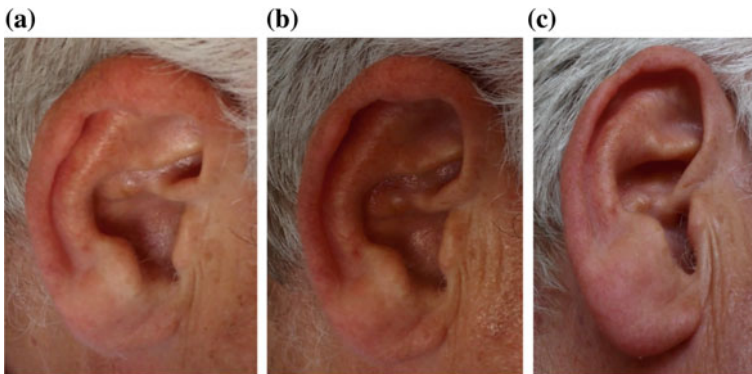


**Fig. 10**  Representative frames for the three horizontal rows of an ear video clip. **a** View from above. **b** A central view; and **c** view from bellow
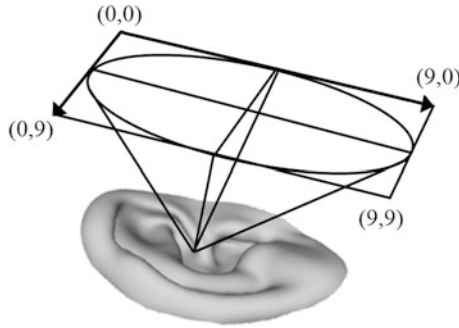
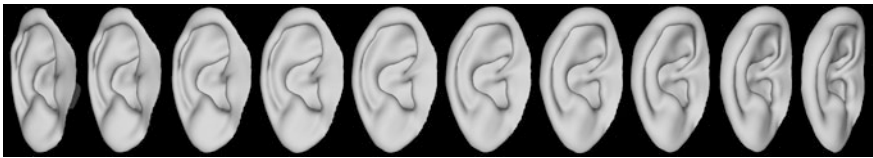**Fig. 11** A scheme of multi-view 3D modeling of a given ear



**Fig. 12** 2D ear images from a row of the ear model rotation scheme, cf. also Fig. 11

This type of 3D ear representation, we call it *Multi-view 3D modeling*, has been developed for our experiments in [30]. We needed there a random access to the Multi-view datasets, but the same datasets could be arbitrary ordered, e.g. top-down and left-right, like the video clips of Sect. 3.5.

An illustration of ten 2D ear images generated from a 3D ear model (for a given central row, cf. Fig. 11), is shown in Fig. 12.

## 3.7 2D Depth Map Images

The build-in functions of Wolfram Mathematica software was used to render 2D depth images from a 3D mesh, where instead of intensity values, the $z$-coordinates of the 3D points are recorded into the 2D image grid (Fig. 13). For consistence with



**Fig. 13** Ear depth maps under orthographic projections of a given 3D ear model

previous section, the depth maps correspond to rotation scheme illustrated on Fig. 11.

## 3.8  Web Access to 3DEarDB

The current version of 3DEarDB will be placed at a free of charge disposal of academic and non-profit research people interested in it. An extended description of the 3DEarDB structure, build-in functions, other potentialities, and license agreements will appear on the web site of IICT-BAS very soon.

## 4  3DEarDB Consistency Experiments

To test the current 3DEarDB functionality, we have experimented using our EGI based approach to ear classification and/or recognition [29]. The EGI representation squeezes appropriately the 3D mesh model data into a sphere, so that it can be visualized and/or used like a 2D (histogram) image, and even like an 1D histogram, by an appropriate re-indexing of facets, e.g. by a spiral, see also [29].

The EGI (Extended Gaussian Image) was initially proposed by B.K.P. Horn, in 1984, [49], see also [50]. Formally, the EGI of a 3D surface represents a histogram of all orientations of the modeled surface on a unit (Gaussian) sphere. Because of surface usual representation by a discrete mesh, every facet from the modeling 3D mesh will be accumulated into the respective point on the Gaussian sphere, according to the unit normal vector and the area of each facet. I.e. the total weight of each EGI point equals the cumulative area of all the mesh facets with the same normal vector direction. In practice, the Gaussian sphere is also discretized by a triangular tessellation, most often based on icosahedron (20 triangular facets). Depending on the level $n$ of the sphere discretization, the number $m$ of 3-angle-facets equals: $m = 4^n 20, n = 0, 1, \ldots$

In our experiments, we have chosen the following three levels: $n = 1, 2, 3$ corresponding to $m = 80, 320$, and 1280, see Table 1.

The opportunity of using the simpler EGI representation of 3D ear mesh models (in deviance of their convex/concave ambiguity) was experimentally demonstrated on a small ear DB, containing only 11 ears models, see [29]. The current version of our 3DEarDB consists of more than 100 ear models that by our best knowledge is enough statistically representative. A hundred of these models, obtained at scan resolution of 1 mm, in similar laboratory conditions, and well post-processed as described here, has been experimented (see Table 1), similarly to [29], to believe

**Table 1** EGI accuracy results: true recognition rate (TRR)

| TRR (%) | | 0.5 mm (a higher resolution recalculated) | | | 1.0 mm (the original 3D scanning resolution) | | | 1.4 mm (a lower resolution recalculated) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Noise (mm)** | | **0.05** | **0.10** | **0.15** | **0.10** | **0.20** | **0.30** | **0.20** | **0.30** | **0.40** |
| ∼ % on width | | 0.16 | 0.31 | 0.47 | 0.31 | 0.62 | 0.93 | 0.62 | 0.93 | 1.24 |
| ∼ % on height | | 0.10 | 0.20 | 0.30 | 0.20 | 0.40 | 0.60 | 0.40 | 0.60 | 0.80 |
| ∼ % on depth | | 0.38 | 0.76 | 1.14 | 0.76 | 1.52 | 2.27 | 1.52 | 2.27 | 3.03 |
| **80** facets | $E_2$ | 100 | 93 | 17 | 100 | 92 | 16 | 100 | 88 | 40 |
| | $E_{BC}$ | 100 | 100 | 67 | 100 | 100 | 70 | 100 | 100 | 80 |
| **320** facets | $E_2$ | 100 | 84 | 25 | 100 | 83 | 27 | 100 | 80 | 44 |
| | $E_{BC}$ | 100 | 100 | 80 | 100 | 100 | 66 | 100 | 100 | 80 |
| **1280** facets | $E_2$ | 100 | 63 | 13 | 100 | 52 | 16 | 100 | 56 | 30 |
| | $E_{BC}$ | 100 | 100 | 62 | 100 | 95 | 48 | 100 | 96 | 78 |

one more again in the proposed 3DEarDB plausibility. For evaluation of similarity between EGI histograms, we have considered again the two geometrical scores:

- the Euclidean distance: $E_2 = \sqrt[2]{\sum_{i=1}^{m} (M_i - S_i)^2}$, and

- the Bray Curtis figure of merit [51]: $E_{BC} = \frac{\sum_{i=1}^{m} |M_i - S_i|}{\sum_{i=1}^{m} (M_i - S_i)}, 0 \le E_{BC} \le 1$;

  where $M_i$ and $S_i$ are both the histogram bins under comparison (of the model and the input objects), $i = 1, 2 \ldots m$; $m = 80$ or 320, or 1280, see Table 1.

## 4.1 Additional Notes to Table 1

- Nearest-neighbor method has been performed for tests, where each processed 3D ear model is considered a center of a class, i.e. the number of classes now is 100.
- Each 3D ear model in the 3DEarDB has been additively noised before using it for test recognition (retrieving the most similar one from 3DEarDB). Three versions of 3DEarDB, i.e. for 3 scan resolutions have been tested: 1.0 mm that is the original one, and two more, 0.5 and 1.4 mm that are recalculated from the original (see Step 5 in Sect. 3.3).
- The noise is artificially generated randomly in the used intervals of 3D scan, i.e. on average: width = 32.3 mm (on O$x$), height = 50.3 mm (on O$y$), and depth = 13.2 mm (on O$z$). These 3 intervals have been simply averaged using respective eigenvalues at the normalization processing (Step 6 in Sect. 3.3).

- To be comparable with other (or further) experiments, the noise intervals are expressed in percents, respectively towards the averaged width, height and depth.

## *4.2 Experiment Analysis*

The following generalization can be done analyzing the conducted experiments:

- Experiments conducted on the current 3DEarDB (100 ear models) confirm the possibility of using the EGI representation for the unambiguous identification of ears nevertheless of their surface mixture of concavities and convexities. This is confirmed by the evaluated noise limits for each of the three experimented resolutions (0.05, 0.10, 0.20 mm, see leftmost columns of Table 1, where TRR = 100 %) that well overcome 0.05 mm, the declared accuracy of used 3D scanner VIUscan.
- As expected, the Bray-Curtis distance ($E_{BC}$) is more robust to the corresponding level of noise, than the Euclidean distance ($E_2$), giving higher TRR.
- A "phenomenon" can be observed for the rest of results of the type TRR < 100 % (at higher level of noise, see middle and rightmost columns), where improvements of either EGI representation (80 → 320 → 1280) or 3D scanning resolution (0.5 ← 1.0 ← 1.4) give an unexpected decrease of TRR at similar levels of noising.
- This "phenomenon" of TRR behavior is considered outside the main positive result for 3DEarDB functionality. Besides of concavities-convexities-mixture of ear surfaces, it can be explained also with combinations of other nonlinearities, like: (i) triangulation irregularities of 3D models, (ii) EGI representation irregularities, (iii) smoothing effect of software manipulation of resolution, etc.
- Because of the opportunities of reducing either the geometric resolution of 3D scanning or the complexity of EGI representation, are always approaching to real time processing, we will keep attention on this phenomenon in our future work.

## 5   Discussion and Conclusion

The current paper describes and proposes to the ear biometric research community a novel multi-model Ear Database, called 3DEarDB. It is composed from different corresponding sets of ear representations from about 100 subjects of Caucasian race acquired by various capturing devices: 3D Laser Scanner, Kinect Xbox One sensor, and a Digital Photo Camera.

The 3DEarDB distinguishes from the currently known similar DBs for its completeness in ear representations of different formats—3D meshes, 3D depth (range) images, 2D video clips, 2D intensity projections. For this reason, it could be useful for comparative analyses among a large variety of known 2D/3D ear recognition approaches and new ones as well, based on the 3D mesh information itself.

A few extra notes about the 3DEarDB near future:

- The current 3DEarDB consists of more than 100 3D ear models. It will be systematically extended in accordance with the feedback from potential users from biometric community in the country and abroad.
- At present, the 3DEarDB consists of only one 3D ear model per subject. The optimal number of (repeated) models per subject will be evaluated soon on the base of a few model versions for a small number of subjects represented (by their right ear). The same is also intended for the left human ear.
- In order to speed up the model acquisition, besides of Kinect camera, we are planning to experiment also with a 3D scanner of structured light type, perhaps on the price of some precision reduction.

# References

1. Day, D.: Biometric applications, overview. In: Li, S.Z., Jain, A.K. (eds.) Encyclopedia of Biometrics, pp. 169–174. Springer, Heidelberg (2015)
2. Hurley, D., Nixon, M., Carter, J.: Ear biometrics by force field convergence. In: Proceedings of the 5th International Conference on Audio- Video- Biometric Person Authentication, pp. 386–394 (2005)
3. Wang, Y., Mu, Z., Zeng, H.: Block-based and multi-resolution methods for ear recognition using wavelet transform and uniform local binary patterns. In: Proceedings of the 19th IEEE International Conference on Pattern Recognition (ICPR), pp. 1–4 (2008)
4. Yan, P., Bowyer, K.: Empirical evaluation of advanced ear biometrics. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 41–42. San Diego, CA, USA, ISBN 0-7695-2372-2 (2005)
5. Bertillon, A.: Signaletic Instructions Including: The Theory and Practice of Anthropometrical Identification (1896)
6. Iannarelli, A.: Ear Identification, Forensic Identification Series. Paramount Publ. Company, Fremont, CA (1989)
7. Cummings, A.H., Nixon, M.S., Carter, J.N.: A novel ray analogy for enrolment of ear biometrics share. In: Proceedings of IEEE Fourth Conference on Biometrics: Theory, Applications and Systems, Washington DC, USA, pp. 1–6 (2010)
8. De Marsico, M., Nappi, M., Riccio, D.: HERO: human ear recognition against occlusions. In: IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 178–183. June 13–18 2010

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

10. Victor, B., Bowyer, K.W., Sarkar, S.: An evaluation of face and ear biometrics. In: Proceedings of 16th IEEE International Conference on Pattern Recognition (ICPR), pp. 429–432 (2002)

11. Zhang, H., Mu, Z., Qu, W., L Iu, L., Zhang, C.: A novel approach for ear recognition based on ICA and RBF network. In: Proceedings of the 4th IEEE International Conference on Machine Learning and Cybernetics, pp. 4511–4515 (2005)

12. Yuan, L., Mu, Z.: Ear recognition based on 2D images. In: First IEEE International Conference on Biometrics: Theory, Applications, and Systems, pp. 1–5 (2007)

13. Naseem, I., Togneri, R., Bennamoun, M.: Sparse representation for ear biometrics. In: Proceedings of the 4th International Symposium on Advances in Visual Computing (ISVC), Part II, pp. 336–345 (2008)

14. Hurley, D., Nixon, M., Carter, J.: Automatic ear recognition by force field transformations. In: Proceedings of the IEEE Colloquium on Visual Biometrics, pp. 7/1–7/5 (2000)

15. Hurley, D., Nixon, M., Carter, J.: Force field feature extraction for ear biometrics. Comput. Vis. Image Underst. **98**(3), 491–512 (2005)

16. Choras, M., Choras, R.: Geometrical algorithms of ear contour shape representation and feature extraction. In: Proceedings of the 6th IEEE International Conference on Intelligent Systems Design and Applications, pp. 451–456 (2006)

17. Choras, M.: Ear biometrics based on geometrical feature extraction. Electron. Lett. Comput. Vis. Image Anal. **5**(3), 84–95 (2005)

18. Abate, A., Nappi, M., Riccio, D., Ricciardi, S.: Ear recognition by means of a rotation invariant descriptor. In: Proceedings of the 18th IEEE International Conference on Pattern Recognition (ICPR), pp. 437–440 (2006)

19. Hailong, Z., Mu, Z.: Combining wavelet transform and orthogonal centroid algorithm for ear recognition. In: Proceedings of the 2nd IEEE International Conference on Computer Science and Information Technology, pp. 228–231 (2009)

20. Sana, A., Gupta, P.: Ear biometrics: a new approach. In: Proceedings of the 6th International Conference on Advances in Pattern Recognition, 06 Sep. 2006, pp. 1–5 (2007)

21. Nanni, L., Lumini, A.: A multi-matcher for ear authentication. Pattern Recogn. Lett. **28**(16), 2219–2226 (2007)

22. Watabe, D., Sai, H., Sakai, K., Andnakamura, O.: Ear biometrics using jet space similarity. In: Proceedings of the IEEE Canadian Conference on Electrical and Computer Engineering, pp. 1259–1264. Niagara Falls, ON. e-ISBN 978-1-4244-1643-1, May 4–7 2008

23. Dewi, K., Yahagi, T.: Ear photo recognition using scale invariant keypoints. In: Proceedings of the International Computational Intelligence Conference, pp. 253–258 (2006)

24. Kisku, D.R., Mehrotra, H., Gupta, P., Sing, J.K.: SIFT-based ear recognition by fusion of detected key-points from color similarity slice regions. In: Proceedings of the IEEE International Conference on Advances in Computational Tools for Engineering Applications (ACTEA), pp. 380–385 (2009)

25. Chen, H., Bhanu, B.: Human ear detection from side face range images. In: Proceedings of the IEEE International Conference on Pattern Recognition (ICPR), pp. 574–577 (2004)

26. Islam, S., Bennamoun, M., Mian, A., Davies, R.: A fully automatic approach for human recognition from profile images using 2D and 3D ear data. In: Proceedings of the 4th International Symposium on 3D Data Processing, Visualization and Transmission, pp. 131–135. Atlanta, Georgia, USA (2008)

27. Yan, P., Bowyer, K.: Biometric recognition using 3D ear shape. IEEE Trans. Pattern Anal. Mach. Intell. **29**(8), 1297–1308 (2007)

28. Cadavid, S., Abdelmottaleb, M.: 3D ear modeling and recognition from video sequences using shape from shading. IEEE Trans. Inf. Forens. Secur. **3**(4), 709–718 (2008)

29. Cantoni, V., Dimov, D.T., Nikolov, A.: 3D ear analysis by an EGI representation. In: Cantoni, V., Dimov, D.T., Tistarelli, M. (eds.) Proceedings of the 1st International Workshop on Biometrics, BIOMET June 23–24, 2014, Sofia, Bulgaria. Biometric Authentication, LNCS, vol. 8897, pp. 136–150. Springer, Heidelberg (2014)
30. Dimov, D.T., Cantoni, V.: Appearance-based 3D object approach to human ears recognition. In: Cantoni, V., Dimov, D.T., Tistarelli, M. (eds.) Proceedings of the 1st International Workshop on Biometrics, BIOMET June 23–24, 2014, Sofia, Bulgaria. Biometric Authentication, LNCS, vol. 8897, pp. 121–135. Springer, Heidelberg (2014)
31. Barra, S., De Marsico, M., Nappi, M., Riccio, D.: Unconstrained Ear processing: what is possible and what must be done. In: Scharcanski, J., Proença, H., Du, E. (eds.) Signal and Image Proceeding for Biometrics, LNEE, vol. 292, pp. 129–190. Springer, Berlin (2014)
32. Pflug, A.: Ear recognition: biometric identification using 2- and 3-dimensional images of human ears. ISBN: 978-82-8340-007-6, Ph.D. thesis, 205p., Gjøvik Univ. College, 2-2015
33. Prakash, S., Gupta, P.: Ear biometrics in 2D and 3D—localization and recognition. In: Hammoud, R.I., Wolff, L.B. (eds.) Augm. Vision & Reality, vol. 10. Springer, Singapore (2015)
34. Phillips, P.J., Wechsler, H., Huang, J., Rauss, P.J.: The FERET database and evaluation procedure for face recognition algorithms. Image Vis. Comput. **16**(5), 295–306 (1998)
35. Gao, W., Cao, B., Shan, S., Zhou, D., Zhang, X., Zhao, D.: CAS-PEAL database (2004). http://www.jdl.ac.cn/peal/
36. UMIST database (1998). http://www.shef.ac.uk/eee/research/iel/research/face.html
37. MID. NIST mugshot identification database (1994). http://www.nist.gov/srd/nistsd18.cfm
38. XM2VTSDB database (1999). http://www.ee.surrey.ac.uk/CVSSP/xm2vtsdb/
39. AMI Ear Database. http://www.ctim.es/research_works/ami_ear_database/
40. Raposo, R., Hoyle, E., Peixinho, A., Proença, H.: UBEAR: a dataset of ear images captured on-the-move in uncontrolled conditions. In: IEEE Workshop on Computational Intelligence in Biometrics and Identity Management (CIBIM), pp. 84–89. Paris, France (2011)
41. UND Databases. http://www.cse.nd.edu/∼cvrl/CVRL/Data_Sets.html
42. USTB Databases. http://www1.ustb.edu.cn/resb/en/index.htm
43. OpenHear Database. http://www2.imm.dtu.dk/projects/OpenHear/
44. SYMARE Database. http://www.ee.usyd.edu.au/carlab/symare.htm
45. HANDY SCAN 3D: The portable 3D scanners for industrial application. http://www.creaform3d.com/sites/default/files/assets/brochures/files/handyscan/Handyscan3D_Brochure_EN_HQ_22052012.pdf
46. Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., Ranzuglia, G.: MeshLab: an open-source mesh processing tool. In: Proceedings of Eurographics Italian Chapter Conference, pp. 129–136 (2008)
47. Vollmer, J., Mencl, R., Müller, H.: Improved laplacian smoothing of noisy surface meshes. Int. Conf. Eurographics **18**(3), 131–138 (1999)
48. Boyé, S., Guennebaud, G., Schlick, C.: Least squares subdivision surfaces. Comput. Graph. Forum. **29**(7), 2021–2028 (2010)
49. Horn, B.K.P.: Extended Gaussian images. Proc. IEEE. **72**, 1671–1686 (1984)
50. Kang, S.B., Horn, B.K.P.: Extended gaussian image (EGI). In: Ikeuchi, K. (ed.) Computer Vision—A Reference Guide, pp. 275–278. Springer, New York (2014)
51. Bray, J.R., Curtis, J.T.: An ordination of upland forest communities of southern Wisconsin. Ecol. Monogr. **27**, 325–349 (1957)