3D Ear Analysis by an EGI Representation

Virginio Cantoni[1], Dimo T. Dimov[2], and Atanas Nikolov[2(✉)]

[1] Department of Industrial and Information Engineering,
Pavia University, Pavia, Italy
virginio.cantoni@unipv.it
[2] Institute of Information and Communication Technologies,
Bulgarian Academy of Sciences, Sofia, Bulgaria
{dtdim,a.nikolov}@iinf.bas.bg

Abstract. In this paper, a new method to represent human ear for biometrics purposes is introduced. Even if ear has a uniform distribution of color, human external ear characteristics are considered unique to each individual and permanent during the lifetime of an adult. For these reasons ear biometrics approaches are relying on morphological ear properties. Even if ear biometrics is a young topic a variety of approaches have been proposed to characterize the ear geometry and topology. Moreover, note that the ear morphology is the biggest human head concavity, and that its convex hull complement is mainly convex. In this connection, the matching potential for ear discrimination can be effectively exploited through an Extended Gaussian Image (EGI) representation. The original EGI representation and its correspondent concrete data-structure are here applied to ear description and discussed for human authentication and identification purposes.

Keywords: Ear authentication · Ear biometrics · Ear identification · Ear recognition · Ear verification · Extended Gaussian Image

1 Introduction

In the computer society personal identification is emerging as a crucial problem: financial institutions, general computer networks, cellular phones, personal workstations, etc. have an ever-growing need to authenticate individuals. Traditionally the identity is established by means of passports, identity cards, badges, keys, or by userid, electronic passwords and personal identification numbers (PIN). These methods are based on possession or knowledge, but possessions can be lost, or stolen, and knowledge can easily be forgotten, or observed. For these reasons the biometrics science is rapidly evolving under the pressure of a large range of application in the civilian computer society. It can be applied in transactions conducted via telephone (exploiting voice recognition technology) and in e-business (exploiting cryptography and public/private password). Biometric methodologies offer today a much higher accuracy than the more traditional ones, and are now normally applied in a variety of applications ranging from personal laptop access to international border control.

V. Cantoni et al. (Eds.): BIOMET 2014, LNCS 8897, pp. 136–150, 2014.
DOI: 10.1007/978-3-319-13386-7_11

In this last decade, more than twenty databases of ears have been collected. These candidate benchmarks differ consistently by number of images, subjects and features. Some of them consider only right ear, or are characterized by variable lighting conditions, rotations, postures, occlusions, yaw and pitch poses, headdresses and earrings, side face, multi-scale, multi-race, ground truth ear's position, or exploitation of depth images (3D), successive sessions, indoor versus outdoor scene, multi-camera, cropped from video streams, etc. Each DB usually considers only a few of these features. For this reason the performances evaluated with the above variability are often not comparable, and the weighting with DB cardinality is certainly not sufficient. Only a few approaches and new proposals have been evaluated and compared on the same benchmark.

Following a very popular taxonomy most of the proposed solutions up-to now successfully proposed are classified as active modalities in which the tested candidate is conscious of the identification or authentication action. These require personal cooperation and will not work if one denies participation. The alternative modality does not require user's active participation, it is relating instead to a passive analysis which usually exploits approaches such as ear recognition or behavioral ones as gait analysis. These modalities can be successful also without that people even know that they are analyzed.

The human outer ear (or pinna) is usually segmented in six basic components: i) the outer helix; ii) the antihelix; iii) the lobe; iv) the tragus; v) the antitragus and finally vi) the concha. This shape in fact evolves during the embryonic state from six growth nodules; its structure therefore is not completely random. Moreover, the detailed structure of the human ear is considered universal and unique (however, it is still to be demonstrated that ears of all people are unique). Furthermore it is considered averagely permanent (the ear appearance does not change consistently when a person ages [1] and is normally collectable. Nevertheless, ear biometrics is not commonly used.

Face recognition has advantages; it can routinely be used in a covert manner, since a person's face is easily captured by video technology and individuals are identified by analyzing certain facial features, such as the medial and lateral corners of the eye or sides of the mouth, nose etc. But it has as well drawback: the face is the most changing part of the body due to facial expressions, during speech and when expressing emotions, and its appearance is often altered by make-up, spectacles, and beards and moustaches and hair styling, moreover there is the effect of age that brings changes in the facial morphology. The ear does not move and only has to support earrings, glasses frames, hearing aids, and it is often occluded by hair. As such, the ear is much less susceptible to interference than many other biometrics, with particular invariance to age.

From the visual complexity viewpoint, face and ear are roughly the same; it is accepted by the researchers that with the decreasing cost of the required 3D scanner and the increasing performance of the ear recognition techniques, ear biometric will be very useful in most practical applications in the near future.

2 Ear Biometrics Approaches

Even if ear biometrics is a young topic a variety of approaches have been proposed, from simple appearance-based methods such as principal component analysis to a whole new perspective based on scale-invariant feature transforms, local binary patterns, force fields, etc. Another proposed taxonomy is related to the general strategies pursued: based on shapes versus contours analysis; on the domain, spatial or transformed by Gabor, Fourier, Hough, Ray, Haar, or by wavelet transformations; considering the pursued ear alignment technique e.g. rigid motion evaluated with the concha area [2, 3] versus the external triangle [4, 5, 6].

Following another taxonomy proposed by Pflug and Busch [7] the proposed solutions are classified on the bases of the input ear image dimensionality (2D or 3D) and then on a quad hierarchy: holistic, local, statistical or hybrid approaches.

The holistic approach is characterized by description of the components as they are mutually interconnected and integrated to compose the ear. A method developed by Burge and Burger [8] is based on the ear representation through a graph model built by the Voronoi diagram on the edge and curve segments extracted from the intensity image, and applies graph matching as discriminant technique. A different model to describe the ear, [9] is built by treating every pixel as an attractor following the Newton's law of gravitation (pixels have a mutual attraction proportional to their intensities and inversely to the square of the distance between them), and the ear is represented by a force field. The discrimination is founded on force field comparison. In [10] the ear image is subdivided into a number of equally large tiles and the self-similarity is evaluated by affine transform of image sub regions. If one tile is occluded, the other tiles may contain a sufficiently distinctive set of features and this make the approach robust to occlusion. In [11] the authors compose six different feature vectors by using seven moment invariants. The moment invariants are robust against changes in scale and rotation. The feature vectors are applied to a back propagation neural network which is trained to classify the feature sets.

Among the holistic methods a large number of proposals exploit classical computer vision transforms. In [12] the generalized Hough transform is used to detect the edges distribution. The cumulative approach make the ear detection Hough transform-based robust to extraction edge misplacement and to pose variation. Extra edges can be due to earing and glasses or hair (mainly strait lines). In [13] a method exploiting the ray transform, which is robust to detect ear in different poses and extra straight edges, is developed. The ray transform is based on the light ray analogy; the simulated ray is reflected by the curved structures like the outer helix in bright regions, hence highlighting these regions in the transformed image. In [14] a Fourier descriptor into frequency space for rotation and scale invariant feature representation is adopted. The ear images have to be aligned and (as in other approaches) the concha region is used to fix a reference point for the alignment step. In [15] a multi-resolution Trace transform and the Fourier transform are exploited to build a feature vector invariant to rotations and scale. In [16] the feature vector consists of some selected wavelet coefficients from Haar-wavelet compression. Applying iteratively a four-level wavelet transform on the ear image, at each iteration new derived coefficients are produced

and stored. In [17] the distinctiveness of different feature extraction methods is compared. In particular, the performance of Fourier descriptors, Gabor transform, moment invariants and statistical features are compared, and the conclusion is that the highest recognition rate is achieved by using moment invariants and Gabor transform.

It is worth to point out the original proposal [18] which develops an ear biometric system based on the acoustic properties of the ear. The method is founded on the estimation of the acoustic transfer function of the ear by stimulating the ear though a sound wave and evaluating the reflected signal.

Basic primitives of the local approach are landmark assessment and local binary pattern. Scale invariant feature transform (SIFT) is known to be a robust way for landmark extraction and can also be used for estimating the rotation and translation between two normalized ear images. A proposal exploiting this approach is [12] in which a reference landmark model, containing a small number of non-redundant landmarks, constitutes the training set. This landmark model is used for filtering the SIFT landmarks, which were initially detected in the reference ear; it is then possible to assign the landmarks with its counterpart. This assignment becomes critical with pose variations and in highly structured regions.

In [19] is proposed 2D ear detection based on edge segmentation through concavity and convexity and then represented in a connectivity graph. A convex hull is applied to the edge in order to detect the ear region. This approach has been extended (with some updating) to the 3D ear analysis.

The keywords of the Statistical approach are principal component analysis (PCA), independent component analysis (ICA), and locally linear embedding (LLE). PCA is by far the most widely adopted method used in ear biometrics; the goal is to reduce the feature vectors dimension. In [20] the performances of PCA when applied on face and ear recognition are compared. In their experiments, the performance of face based recognition overcomes the one of ear based recognition. However, in [21] it is reached (in similar experiments) a different conclusion: no appreciable difference was found between face and ear in terms of recognition performance.

In [22] the performance in ear identification by neural network classifiers is investigated. The ear image is represented by outer ear points, information collected from ear shape and folds, and macro features assessed by compression network. The conclusion was that compression network support the best performance.

In [4] the outer contour of the ear is located by searching for the top, bottom and left points of the detected ear boundary; these points form a triangle and its barycenter is selected as reference point for image alignment and consequently for the matching process.

An example of the hybrid approach is given in [23, 24] that use the active shape model for extracting the outline of the ear. In [23] manually cropped ear images are used. A feature extractor stores selected points on the outline of the ear together with their distance to the tragus which is selected as reference. Before applying a linear classifier, the dimensionality of the feature vectors is reduced by PCA.

3 Extended Gaussian Image

A 3D mesh model, approximating a 3D object, is represented by a set of triangles (see Fig. 1a):

$$\boldsymbol{T} = \{T_1, \ldots, T_N\},\ T_i \subset R^3, \tag{1}$$

where N is the number of triangles of the object mesh.

Each triangle T_i consists of a set of three vertices:

$$T_i = \{\boldsymbol{P}_{A_i}, \boldsymbol{P}_{B_i}, \boldsymbol{P}_{C_i}\} \tag{2}$$

Being $\boldsymbol{w_i}$ and $\overline{\boldsymbol{n}}_i$ respectively the center and the normal of a triangle T_i (see Fig. 1b), the surface area A_i of triangle T_i, and the area A of the mesh are given by:

$$w_i = (P_{A_i} + P_{B_i} + P_{C_i})/3 \tag{3}$$

$$\overline{\boldsymbol{n}}_i = (P_{C_i} - P_{A_i}) \times (P_{B_i} - P_{C_i}) \tag{4}$$

$$A_i = \frac{1}{2}\left|(P_{C_i} - P_{A_i}) \times (P_{B_i} - P_{C_i})\right| \tag{5}$$

$$A = \sum_{i=1}^{N} A_i \tag{6}$$

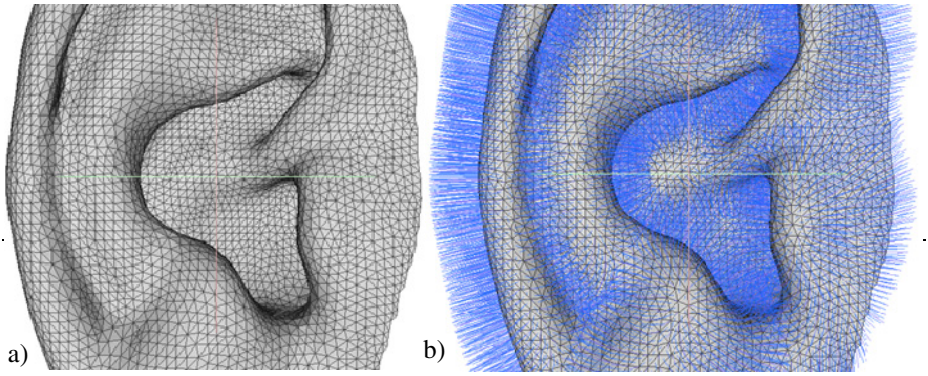

Fig. 1. a) A 3D triangular mesh model; b) normal vectors to mesh triangles; Both images are scaled and respectively cropped for better visualization of the triangular mesh

The Extended Gaussian images (EGI) of a 3D object or shape is the histogram of orientations that represents the distribution of surface area with respect to surface orientation (see Fig. 2) [25].

Each surface patch is mapped to a point on the unit Gaussian sphere according to its surface normal. The weight for each surface normal (represented by a point on the Gaussian sphere) is the total sum of area of all the surface patches having that surface normal.

The EGI can be easily built from needle or depth maps generated by range or stereo devices. In fact, for an effective digital representation the Gaussian sphere is discretized by a triangular tessellation (sometimes called geodesic dome). Starting with a

regular polyhedron (e.g. the icosahedron herewith adopted), recursively in a more detailed description level each triangle is split into four smaller triangles (see Fig. 3). Being k the number of iterative subdivision steps, the number of triangles is $m = 2^{2k}K_0$, where K_0 is the number of faces of the starting polyhedron (20 for the icosahedron) and the area (solid angle) of the single cell is $A_\Delta = \pi/(2^{2(k-1)}K_0)$ respectively [26].

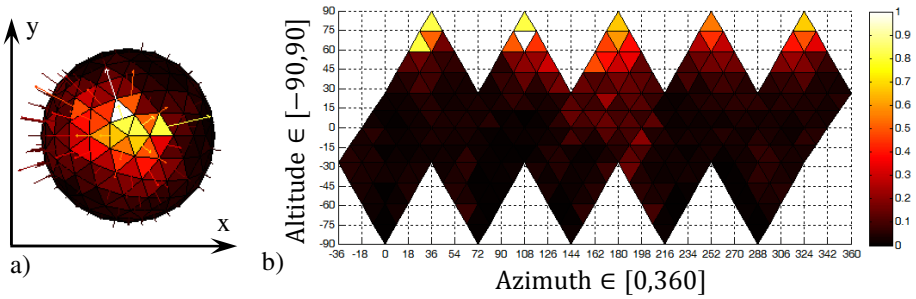

Fig. 2. a) & b) 3D/2D EGI histograms: in each EGI bin (oriented triangle from discretized polyhedron) the triangle areas with the same orientation were accumulated.

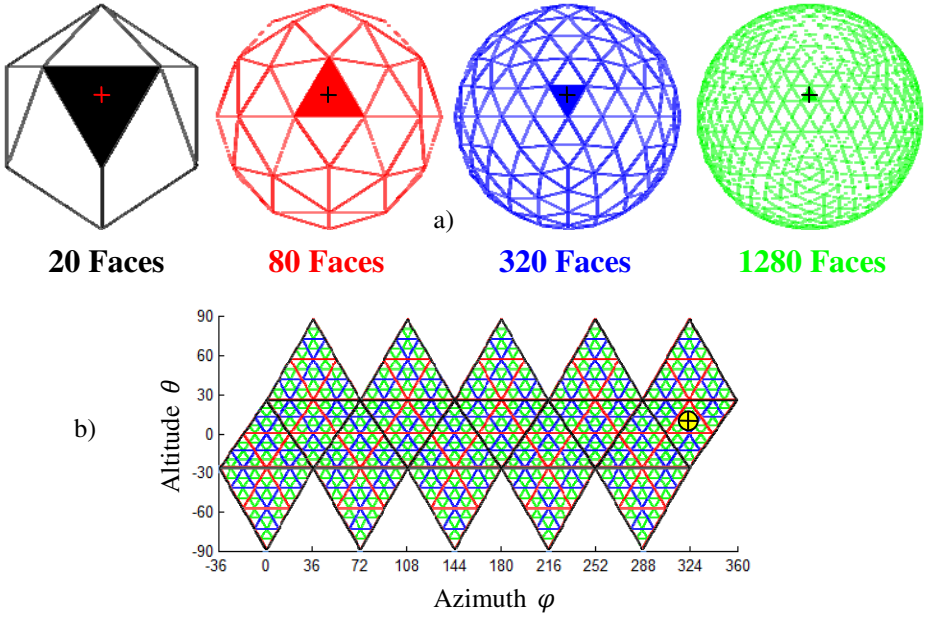

Fig. 3. a) From left to right: hierarchical refinement of successive searching for maximum dot product between the normal vector of an input patch and the icosahedron ones or of the three polyhedrons with 80, 320, and 1280 faces; b) 2D representation of the face positions of the four polyhedrons (expressed by azimuth and altitude of triangle vertices). The representation of a given input orientation ⊕ is also shown.

Main properties of the EGI for convex polyhedrons and for general convex objects are:

- Rotation of the polyhedron corresponds to an equal rotation of the EGI, and vice versa, since the surface normal vectors rotate with the object.
- Being the total mass of the EGI obviously just equal to the total surface area of the polyhedron, and being the same the projected area when viewed from any pair of opposite directions, the center of mass of an EGI has to lie at the origin.
- Herman Minkowski in 1897 demonstrated that a convex object is fully described by the area and orientation of its faces, that is, two different convex polyhedrons have different EGIs. Vice versa two different EGIs represent two different polyhedrons. Moreover, this property is maintained for a general convex object: in case of convex objects, there is an injective correspondence with their EGI.
- The EGI is invariant to translation being a distribution with respect to surface orientation. In registering two 3D objects, the translation can be ignored and the rotation can be evaluated minimizing $e(\boldsymbol{R})$ i.e. just comparing the EGI of the model $M_{\hat{n}}$ and the EGI $S_{\hat{n},\boldsymbol{R}}$ of the shape rotated by $\boldsymbol{R}$:

$$e(\boldsymbol{R}) = \sum_{\hat{n}\in m} \left(M_{\hat{n}} - S_{\hat{n},\boldsymbol{R}}\right)^2 \tag{7}$$

To this regard, many approaches are known to solve or to lighten the above minimization problem, e.g. by the Principal Component Analysis (PCA) method, both the 3D objects can be preliminary normalized by position and size.

3.1 Describing the Process of Creation an Icosahedron and Its 2D Mapping

In brief, the process of building a 3D/2D icosahedron consists of the following steps (see also Fig. 4):

- determine the azimuth φ_i for each vertex;
- estimate the radius ρ from a known relationship between edge length a of the icosahedron and the radius r of the circumscribed sphere;
- express each (x, y) coordinate of the vertices of the icosahedron by ρ and φ_i ;
- compute the z coordinate of the vertices by the radiuses ρ and r, $(r = 1)$;
- find both altitude levels $+\theta_1$ and -θ_1, which are necessary for the 2D representation.

A more detailed description of the creation of an icosahedron and its 2D mapping is given here (see Fig. 4):

1. The azimuth to the vertices with altitude θ_1 is equal to: $\varphi_i^+ = i\,\Delta\varphi = i\,2\pi/5$, and respectively for the vertices with altitude -θ_1, the azimuth is: $\varphi_i^- = i\,\Delta\varphi - \Delta\varphi/2 = \pi/5(2i-1)$, where $i = 0 \div 4$. The azimuth of the top and bottom vertex is zero.

2. It is known that if the edge length of an icosahedron is a, the radius r of the circumscribed sphere around the icosahedron is: $r = a\ \sin\left(\frac{2\pi}{5}\right)$.
   In our case $(r = 1) \Rightarrow a = \frac{1}{\sin\left(\frac{2\pi}{5}\right)} \approx 1.0515$ and the radius ρ:
   $\rho = a\ \left(2\sin\left(\frac{\Delta\varphi}{2}\right)\right)^{-1} = \left(2\sin\left(\frac{\pi}{5}\right)\sin\left(\frac{2\pi}{5}\right)\right)^{-1} \approx 0.8944.$
3. The (x, y) coordinates of vertices on θ_1 altitude level are: $x_i^+ = \rho\cos(\varphi_i^+)$, $y_i^+ = \rho\sin(\varphi_i^+)$. Respectively for vertices on -θ_1 altitude level: $x_i^- = \rho\cos(\varphi_i^-)$, $y_i^- = \rho\sin(\varphi_i^-)$, $i = 0 \div 4$. The (x, y) coordinates of the top (**N**) and bottom (**S**) vertex are $(0, 0)$.
4. The distance $\overline{OO_+} = \sqrt{r^2 - \rho^2} = \sqrt{1 - \rho^2} \approx 0.4472$, which actually is $\rho/2$. So the z coordinates of vertices lying on θ_1 and -θ_1 levels are $\pm\rho/2$. The top and bottom vertex has $z = \pm 1$.

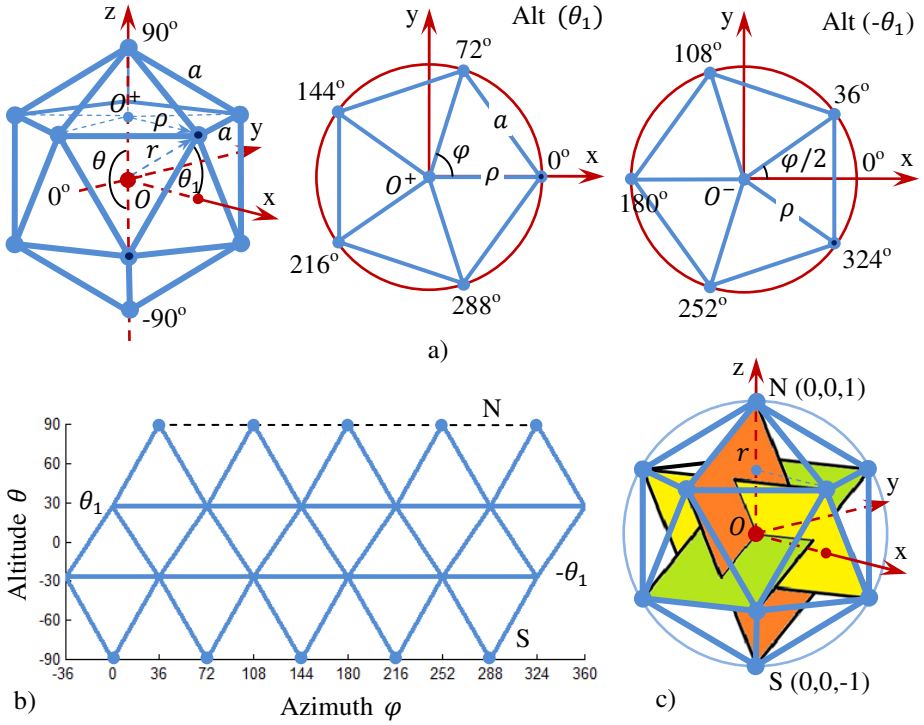

Fig. 4. a) Parameters description of an icosahedron: radius $(r = 1)$ of the circumscribed unit sphere; icosahedron edge length a; azimuth $\varphi \in [0, 360]$; altitude $\theta \in [-90, 90]$; radius ρ of the O_{xy}^+ and O_{xy}^- 2D sections; altitude levels θ_1 and -θ_1 of the vertices belonging to the both 2D sections of the icosahedron; b) the same icosahedron mapped into 2D plane, represented by azimuth φ and altitude θ of its triangular faces; c) the icosahedron, inscribed in a unit sphere, and formed by three mutually orthogonal 'golden rectangles', whose sides ratio is equal to the golden ratio $(1 + \sqrt{5})/2 \approx 1.6180$

5. The altitude levels $\pm\theta_1$, can be estimated as: $\pm\theta_1 = \arccos\left(\frac{\sqrt{x^2+y^2}}{r}\right) \approx 26.57^o$, for any (x, y) point on these levels, where $r = 1$.

The above data is enough for building a 3D icosahedrons and its 2D mapping. It is worth mentioning that higher discretization levels from an icosahedron can be achieved by computing the vertices coordinates, dispatching each triangle into four new ones.

3.2 Evaluation of Similarity Between EGI Histograms

Among the matching indexes adopted to determine a geometrical score of similarity with EGI applications, we have considered the following [27]:

- the Minkowski distance:

$$E_M = \sqrt[p]{\sum_{\hat{n}=1}^{m} |M_{\hat{n}} - S_{\hat{n}}|^p} \tag{8}$$

the Manhattan ($p = 1$) and the Euclidean ($p = 2$) distances are obtained respectively; $M_{\hat{n}}$ and $S_{\hat{n}}$ are histograms under comparing (of model and input object).

- the Bray Curtis figure of merit:

$$E_{BC} = \frac{\sum_{\hat{n}=1}^{m} |M_{\hat{n}} - S_{\hat{n}}|}{\sum_{\hat{n}=1}^{m} |M_{\hat{n}} + S_{\hat{n}}|} \tag{9}$$

obviously $0 \leq E_{BC} \leq 1$;

- a new figure of merit is defined in [29] as the ratio $\mathrm{E_Z} = \mathrm{g/m}$, where g is the number of histograms bins, for which this condition is fulfilled:

$$\left\{(M_{\hat{n}} = S_{\hat{n}} = 0) \cup \left(\frac{|M_{\hat{n}} - S_{\hat{n}}|}{[M_{\hat{n}}, S_{\hat{n}}]_{\max}} \leq \theta\right)\right\}_{\forall \hat{n}, 1 \leq \hat{n} \leq m} \tag{10}$$

being θ a suitable threshold, $0 \leq \theta \leq 1$.

In our experimentations the distances of type (8) has been normalized by the number of polyhedron faces.

The EGI has been initially proposed for applications of photometry by B.K.P. Horn [25] in the '80 and has been extended by K. Ikeuci (the Complex-EGI) [29] in the '90 to overcome the ambiguity that are introduced by the concave parts. Later other improvements have been proposed always with the purpose to reduce the quoted ambiguities, among the others in chronological sequence: the More Extended Gaussian Image (MEGI) in 1994 [30], the Multi-Shell Extended Gaussian Image (MSEGI) and the Adaptive Volumetric Extended Gaussian Image (AVEGI) in 2007 [31], and the Enriched Complex Extended Gaussian Image (EC-EGI) in 2010 [32].

In this preliminary work our experiments are limited to the EGI, because it constitutes a compact and effective representation of a 3D object. Besides, the feeling is that being an ear basically a cavity (certainly with convexities) the EGI can be both effective and efficient [33]. After suitable experiments, other more precise solutions, but not only limiting to derivations of the EGI, can be adopted too.

4 Experimental Analysis

The Institute of Information and Communication Technologies (IICT) of the Bulgarian Academy of Sciences (BAS) has collected an ear dataset with the goal of providing more high definition data than comparable collections. The dataset now represents 11 subjects of various ages and consist of 66 3D ear models in total. For each subject, the dataset contains 6 3D ears (where 5 are intentionally noised). All 3D ear models are taken under optimal lighting conditions through a VIUscan 3D scanner. This type of scanners is composed of a laser cross beam and of two HD cameras surrounded by a set of LEDs, thus allowing the laser triangulation and 3D data acquisition. The scanner can reach a geometry resolution of 0.1mm, an accuracy of 50μm, and 24 bits of texture colors. In Fig. 5, some examples of 2D frontal projections from the 3D dataset are shown.

The preprocessing of the ears consists of cropping the ear from the background and holes filling by the VIUscan 3D scanner's software (VXelements). Then in a post-processing phase by an open source system for the processing and editing of 3D triangular meshes - MeshLab, the final result of Fig. 5 is obtained. Note that this phase is applied just for the model construction that is an off-line procedure.

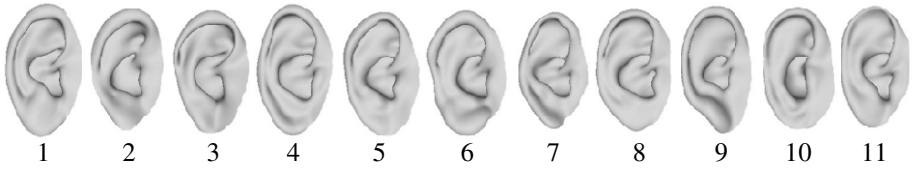

Fig. 5. The current test dataset of 11 original 3D ear models

One of the aims of our experiments is to determine the most appropriate distances for similarity evaluation between EGI histograms (see section 3), as well as their robustness to noise for object recognition. This noise is introduced to represents different accuracy of 3D scanning systems. For this purpose, a uniform noise in a given range was generated, and added to each 3D vertex coordinate of the scanned objects (Fig. 6a). It could be easily seen that: the higher noise, the more uniform the orientation histogram (see Fig. 6b, c), and the more challenging the recognition process.

For each ear in the dataset the corresponding EGI represented by 3D/2D histograms are built. There, in each EGI bin (see Fig. 2) all the areas of object's triangle having the same orientation are accumulated. In order to do this, the histogram bin is selected by the maximum dot product between the input patch orientation and the coarse-to-fine set of triangles' normals of the polyhedron (see Fig. 3). Thus, the total area of the EGI histogram is equal to the object's area.

Before forming the 3D/2D EGI histograms, the PCA method was used to equalize position and scale of the 3D objects in a global coordinate system. Thus, all ears models become invariant to scaling and only their morphology is taking into account. Obviously the eigenvectors and eigenvalues can be applied (and are more and more applied) for recognition purposes, but in our experimentation we use them just for alignment and scaling using as discriminant characteristics only the EGI.

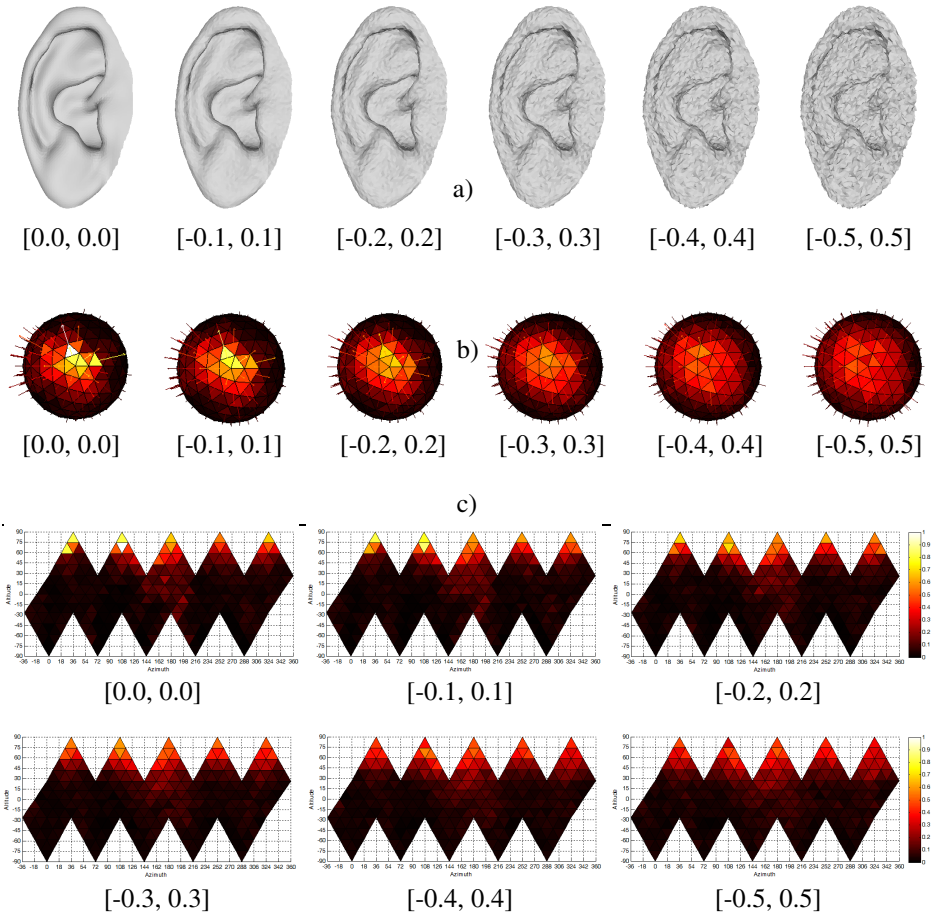

Fig. 6. a) An ear model corrupted with uniform noise in a given range, added to each of its (x, y, z) coordinates, thus simulating the varying accuracy of a 3D scanning device; b) & c) the corresponding 3D/2D EGI histograms of the noisy ears

So far, our primary recognition strategy is based on the classical approach of the nearest neighbor, i.e. the shortest distance is pursued between an EGI histogram of given noisy input ear, compared to all ear model histograms, using (8), (9) and (10). As an example, in Table 1, on rows, the distances between input noisy ears with uniformly distributed noise in range [-0.1, 0.1] and all original ears (models) are given, using matching index E_{BC}. To have true recognition of given noisy ear, the minimal distance in every row must lie on the main diagonal (green cells). If this is not the case, a false detection is attained – as in the case in Table 2 (the red cells). Furthermore, in addition to have true recognition, it is important to evaluate its reliability. Therefore, an evaluation of the recognition reliability is given in the last column of each table. It represents the ratio between the true matches and the first false one. The

smaller its value, the better the reliability. If the reliability value $\eta \geq 1$, it is a false recognition detection.

A set of analogous 20 tables for each of the 4 mentioned distances ($E_{M(p=1|2)}$, E_{BC} and E_Z) and for 5 ranges of the noise have been computed in order to view the impact of noise on these distances reliability. All the results of true and false detections are summarized in Table 3, where the True Recognition Rate (TRR) and average reliability $\tilde{\eta}$ for each of 20 tables (cases) is presented.

Table 1. The matching index E_{BC} is used between input noisy [± 0.1] ears and model ears. The last column represents the ratio η between the true matches and the first false one (in gray)

Noisy \ Model	1	2	3	4	5	6	7	8	9	10	11	η
1	0.099	0.321	0.276	0.211	0.224	0.259	0.255	0.235	0.255	0.264	0.218	0.468
2	0.318	0.099	0.327	0.307	0.308	0.251	0.294	0.280	0.306	0.285	0.330	0.394
3	0.285	0.321	0.089	0.297	0.273	0.305	0.286	0.298	0.309	0.307	0.290	0.325
4	0.222	0.305	0.288	0.079	0.242	0.262	0.234	0.218	0.232	0.275	0.213	0.370
5	0.230	0.316	0.275	0.234	0.087	0.223	0.262	0.237	0.255	0.240	0.243	0.390
6	0.250	0.246	0.299	0.257	0.231	0.093	0.252	0.251	0.265	0.251	0.243	0.404
7	0.263	0.290	0.298	0.234	0.254	0.266	0.089	0.257	0.222	0.274	0.255	0.401
8	0.230	0.285	0.295	0.210	0.227	0.253	0.253	0.096	0.243	0.261	0.249	0.458
9	0.235	0.293	0.292	0.206	0.235	0.246	0.214	0.234	0.104	0.236	0.218	0.507
10	0.267	0.280	0.308	0.264	0.224	0.239	0.268	0.269	0.249	0.101	0.246	0.452
11	0.237	0.321	0.284	0.212	0.238	0.243	0.241	0.250	0.242	0.244	0.090	0.426
											AVG	**0.418**

Table 2. The matching index E_{BC} is used between input noisy [± 0.4] ears and model ears. The last column represents the ratio η between the true matches and the first false one (gray | red)

Noisy \ Model	1	2	3	4	5	6	7	8	9	10	11	η
1	0.215	0.298	0.235	0.226	0.228	0.249	0.228	0.237	0.258	0.283	0.232	0.954
2	0.282	0.208	0.296	0.279	0.283	0.233	0.265	0.261	0.291	0.283	0.293	0.892
3	0.263	0.308	0.213	0.272	0.268	0.286	0.260	0.276	0.289	0.307	0.293	0.821
4	0.249	0.310	0.262	0.203	0.237	0.259	0.224	0.247	0.245	0.284	0.237	0.905
5	0.230	0.311	0.247	0.230	0.208	0.236	0.220	0.240	0.258	0.279	0.247	0.946
6	0.256	0.265	0.264	0.252	0.245	0.206	0.232	0.252	0.271	0.286	0.262	0.887
7	0.283	0.315	0.281	0.267	0.265	0.271	0.204	0.276	0.263	0.311	0.268	0.775
8	0.230	0.302	0.259	0.218	0.227	0.232	0.223	0.216	0.245	0.281	0.241	0.988
9	0.235	0.309	0.265	0.222	0.240	0.249	0.211	0.255	0.223	0.271	0.228	1.055
10	0.233	0.285	0.269	0.212	0.225	0.236	0.244	0.238	0.246	0.211	0.233	0.994
11	0.248	0.315	0.265	0.229	0.246	0.259	0.221	0.256	0.259	0.282	0.219	0.992
											AVG	**0.928**

Table 3. True Recognition Rate (TRR) and corresponding average reliabilities $\tilde{\eta}$ (the smaller, the better), based on different combinations of the investigated distances and ranges of noise

Noise / Distance	± 0.1		± 0.2		± 0.3		± 0.4		± 0.5	
$E_M(p=1)$	100%	0.415	100%	0.655	100%	0.828	72.7%	0.939	36.4%	1.002
$E_M(p=2)$	100%	0.427	100%	0.705	63.6%	0.917	45.5%	1.031	36.4%	1.101
E_{BC}	100%	0.418	100%	0.657	100%	0.830	90.9%	0.928	54.5%	0.987
E_Z	100%	0.776	100%	0.847	90.9%	0.927	72.7%	0.966	36.4%	1.019

According to these experiments, it seems that the figure of merit E_{BC} is the most robust to this kind of uniform noise with the highest TRR and the best average reliability $\tilde{\eta}$ of recognition. A little bit worse for high noise data are the results for the $E_M(p=1)$ metric.

The E_Z index shows good performance, but it ranks on third place by goodness of TRR. This result is subject to the experimentally chosen threshold which is here set to 0.5.

The last $E_M(p=2)$ distance gives the lowest TRR.

The experiments show that the results are very promising, i.e. even the simplest EGI representation of ear models could distinguish them very well each other, no matter that their surface is not entirely convex. Also it would be interesting to use 2D mapping of EGI representation not only for overall better observing the resulting histograms, but also for applying the well-known or new adopted 2D recognition approaches on it.

5 Conclusion

In this paper a new approach suitable for ear authentication and identification has been proposed. The first results look promising for considering these new strategies among the candidates for a practical exploitation. Certainly it is necessary to extend the investigation to more general conditions for the acquisition and experiment also in the large variety of cases and people. Moreover, the computer demanding preprocessing phase is easily supported because it is offline, certainly it must be investigated the amount of preprocessing necessary for the online authentication and identification to be applied to the ear under test. Nevertheless, these preliminary results show robustness to image degradation that looks very encouraging.

The near future activity is related to a few tuning aspects of the current implementation. In particular we will consider: i) relationship between input 3D ear image quality and EGI resolution; ii) speed-up analysis considering simpler data representations, e.g. 2D EGI descriptions as in Fig. 2 b); iii) analysis if there are cases for which more complex EGI representations are required (e.g. Complex EGI, Enriched C-EGI, etc.).

Acknowledgements. This research is partly supported by the project AComIn "Advanced Computing for Innovation", grant 316087, funded by the FP7 Capacity Programme (Research Potential of Convergence Regions).

References

1. Sforza, C., Grandi, G., Binelli, M., Tommasi, D.G., Rosati, R., Ferrario, V.F.: Age- and sex-related changes in the normal human ear. Forensic Sci. Int. **187** (1–3), 110e1–110e7 (2009)
2. Yan, P., Bowyer, K.W.: Biometric recognition using 3D ear shape. Pattern Anal. Mach. Intell. **29**, 1297–1308 (2007)
3. Choras, M.: Perspective methods of human identification: ear biometrics. Opto-Electron. Rev. **16**, 85–96 (2008)
4. 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. (eds.) ACIVS 2008. LNCS, vol. 5259, pp. 1030–1037. Springer, Heidelberg (2008)
5. Mu, Z., Yuan, L., Xu, Z., Xi, D., Qi, S.: Shape and Structural Feature Based Ear Recognition. In: Li, S.Z., Lai, J.-H., Tan, T., Feng, G.-C., Wang, Y. (eds.) SINOBIOMETRICS 2004. LNCS, vol. 3338, pp. 663–670. Springer, Heidelberg (2004)
6. Rahman, M., Islam, R., Bhuiyan, N.I., Ahmed, B., Islam, A.: Person identification using ear biometrics. Int. J. Comput. Internet Manage. **15**, 1–8 (2007)
7. Pflug, A., Busch, C.: Ear biometrics: a survey of detection, feature extraction and recognition methods. IET Biometrics. **1**(2), 114–129 (2012)
8. Burge, M., Burger, W.: Ear biometrics in computer vision. Proc. Int. Conf. on Pattern Recognition **2**, 822–826 (2000)
9. Hurley, D.J., Nixon, M.S., Carter, J.N.: Force field energy functionals for image feature extraction. Image Vision Comp. J. **20**, 311–317 (2002)
10. De Marsico, M., Michele, N., Riccio, D.: HERO: human ear recognition against occlusions. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition Workshops, p. 178 (2010)
11. Wang, X.q., Xia, H.y., Wang, Z.l.: The research of ear identification based on improved algorithm of moment invariants. Third Int. Conf. on Inf. and Computing, p. 58 (2010)
12. Arbab-Zavar, B., Nixon, M.S.: On Shape-Mediated Enrolment in Ear Biometrics. In: Bebis, G., Boyle, R., Parvin, B., Koracin, D., Paragios, N., Tanveer, S.-M., Ju, T., Liu, Z., Coquillart, S., Cruz-Neira, C., Müller, T., Malzbender, T. (eds.) ISVC 2007, Part II. LNCS, vol. 4842, pp. 549–558. Springer, Heidelberg (2007)
13. Alistair, H., Cummings, A.H., Nixon, M.S., Carter, J.N.: A novel ray analogy for enrolment of ear biometrics. Theory Applications and Systems, Fourth IEEE Int. Conf. on Biometrics (2010)
14. Abate, A.F., Nappi, M., Riccio, D., Ricciardi, S.: Ear recognition by means of a rotation invariant descriptor. 18th Int. Conf. on Pattern Recognition **4**, 437–440 (2006)
15. Fooprateepsiri, R., Kurutach, W.: Ear based personal identification approach forensic science tasks. Chiang Mai J. Sci. **38**(2), 166–175 (2011)
16. Sana, P.P.R., Gupta, A.: Ear biometrics: a new approach. Advances in Pattern Recognition, pp. 1–5 (2007)
17. Wang, X., Yuan, W.: Human ear recognition based on block segmentation. Int. Conf. on Cyber-Enabled Distributed Computing and Knowledge Discovery, pp. 262–266 (2009)
18. Akkermans, A.H.M., Kevenaar, T.A.M., Schobben, D.W.E.: Acoustic Ear Recognition for Person Identification. Fourth IEEE Workshop on Automatic Identification Advanced Technologies, pp. 219–223 (2005)
19. Gupta, P., Prakash, S.: An efficient technique for ear detection in 3D: invariant to rotation and scale. Fifth IAPR Int. Conf. on Biometrics, p. 97–102 (2012)

20. Victor, B., Bowyer, K.W., Sarkar, S.: An evaluation of face and ear biometrics. Proceedings of International Conference on Pattern Recognition, pp. 429–432 (2002)
21. Chang, K., Bowyer, K.W., Sarkar, S., Victor, B.: Comparison and combination of ear and face images in appearance-based biometrics. IEEE Transaction on Pattern Analysis of Machine Intelligence. **25**, 1160–1165 (2003)
22. Moreno, B., Sanchez, A., Velez, J.F.: On the use of outer ear images for personal identification in security applications. In: Proceedings of IEEE Conference on Security Technology, pp. 469–476 (1999)
23. Lu, L., Xiaoxun, Z., Youdong, Z., Yunde, J.: Ear recognition based on statistical shape model. First Int. Conf. on Innovative Computing, Inf. and Control, pp. 353–356 (2006)
24. Yuan, L., Mu, Z.: Ear recognition based on 2D images. First IEEE Int. Conf. on Biometrics: Theory, Applications, and Systems, pp. 1–5 (2007)
25. Horn, B.K.P.: Extended Gaussian images. Proc. of the IEEE. **72**, 1671–1686 (1984)
26. Cantoni, V., Gaggia, A., Lombardi, L.: Essay: Extended Gaussian Image for pocket-ligand matching. Definitions: EGI; CEGI e ECEGI. In: W. Dubitzky, O. Wolkenhauer, K. Cho & H. Yokota (eds.) Encyclopedia of Systems Biology. LLC. Springer (2012)
27. Gaggia, A.: System for protein-ligand interaction analysis, PhD Thesis, Pavia University (2013)
28. Zhang, J.: Content-based 3D Model Retrieval Based on Volumetric Extended Gaussian Image Shape Representation. PhD Thesis of the City University of Hong Kong (2007)
29. Kang, S.B., Ikeuchi, K.: Determining 3-D object pose using the complex extended Gaussian image. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 580–585 (1991)
30. Matsuo, H., Iwata, A.: 3-D Object Recognition Using MEGI Model from Range Data. In: Proc. 12th Int. Conf. Pattern Recognition, pp. 843–846 (1994)
31. Wang, D., Zhang, J., Wong, H.-s., Li, Y.: 3D Model Retrieval Based on Multi-Shell Extended Gaussian Image. In: Qiu, G., Leung, C., Xue, X.-Y., Laurini, R. (eds.) VISUAL 2007. LNCS, vol. 4781, pp. 426–437. Springer, Heidelberg (2007)
32. Hu, Z., Chung, R., Fung, K.S.M.: EC-EGI: enriched complex EGI for 3D shape registration. Machine Vision and Applications. **21**, 177–188 (2010)
33. Cantoni, V., Gaggia, A., Lombardi, L.: A data structure for protein-ligand morphological matching. In: New Tools and Methods for Pattern Recognition in Complex Biological Systems, Nuovo Cimento C, vol. 35, no. 5, suppl. 1, pp. 89–97 (2012)
34. Iannarelli, A.: Ear Identification. Paramont Publishing Company, Forensic Identification Series (1989)