

A Fully Automatic Approach for Human Recognition from Profile Images Using 2D and 3D Ear Data

S. M. S. Islam, M. Bennamoun, A. S. Mian and R. Davies

School of Computer Science and Software Engineering, The University of Western Australia
35, Stirling Hwy, Crawley, WA 6009

{shams, bennamou, ajmal, rowan}@csse.uwa.edu.au

Abstract

The use of ear shape as a biometric trait for recognizing people in different applications is one of the most recent trends in the research communities. In this work, a fully automatic and fast technique based on the AdaBoost algorithm is used to detect a subject's ear from his/her 2D and corresponding 3D profile images. A modified version of the Iterative Closest Point (ICP) algorithm is then used for the matching of this extracted probe ear to the previously stored ear data in a gallery database. A coarse-to-fine hierarchical technique is used where the ICP algorithm is first applied on low and then on high resolution meshes of 3D ear data. We obtain a rank one recognition rate of 93% while testing with the University of Notre Dame Biometrics Database. The proposed recognition approach does not require any manual intervention or sharp extraction of ear contour from the detected ear region. No segmentation of the extracted ear is required and more importantly, the system performance does not rely on the presence of a particular feature of the ear.

1. Introduction

Due to instances of fraud with the traditional ID based systems, biometric recognition systems are gaining popularity day by day. In such a system, one or more physiological (e.g. face, fingerprint, palmprint, iris and DNA) or behavioral (e.g. handwriting, gait and voice) traits of a subject are taken into consideration for automatic recognition. Although the ear as a biometric trait is not as accurate as iris or DNA, it is non-intrusive and easy to be collected. The face is also non-intrusive but its appearance is affected by changes in facial expressions, use of cosmetics or eye glasses. The ear is also smaller in size but rich in features and its shape does not change with aging between 8 years and 70 years [7]. It can be used separately or in a multimodal approach with the face for effective hu-

man recognition in many applications including some national IDs, security, surveillance and law enforcement applications. However, in any of these applications, accurate recognition of the ear is an important step.

A biometric recognition system may operate in one or both of two modes: authentication and identification. In authentication, one-to-one matching is performed to compare a user's biometric to the template of the claimed identity. In identification, one-to-many matching is done to associate an identity with the user by matching it against every identity in the database. For both modes, an accurate detection is an essential pre-requisite. However, ear detection from arbitrary profile images is a challenging problem due to the fact that the ear is sometimes occluded by hair and earrings and ear images can vary in appearance under different viewing and illumination conditions. Consequently, most of the existing ear recognition algorithms assume that the ear has been accurately detected [9].

One of the earliest ear detection methods uses Canny edge maps to detect the ear contour [3]. Hurley et al. [6] proposed the "force field transformation" for ear detection. Alvarez et al. [1] used a modified active contour algorithm and Ovoid model for detecting the ear. Yan and Bowyer [14] proposed taking a predefined sector from the nose tip to locate the ear region. The non-ear portion from that sector is cropped out by skin detection and the ear pit was detected using Gaussian smoothing and curvature estimation. Then, they applied an active contour algorithm to extract the ear contour. The system is automatic but fails if the ear pit is not visible. Li Yuan and Mu [15] used a modified *CAMSHIFT* algorithm to roughly track the profile image as the region of interest (ROI). Then, contour fitting is operated on ROI for further accurate localization using the contour information of the ear.

Most recently, Islam et al. [8] proposed an ear detection approach based on the AdaBoost algorithm [12]. The system was trained with rectangular Haar-like features and using a dataset of varied races, sexes, appearances, orientations and illuminations. The data was collected by cropping

and synthesizing from the University of Notre Dame (UND) biometrics database [13, 14], the NIST Mugshot Identification Database (MID), the XM2VTSDB [10], the USTB, the MIT-CBL and the UMIST database. The approach is fully automatic, provides 100% detection while tested with 203 non-occluded images of the UND profile face database and also works well with some occluded and degraded images.

As summarized in the survey of Pun et al. [11] and Islam et al. [9], most of the proposed ear recognition approaches use either PCA (Principal Component Analysis) or the ICP algorithm for matching. Choras [4] proposed a different automated geometrical method. Testing with 240 images (20 different views) of 12 subjects, 100% recognition rate is reported. Genetic local search and the force field transformation based approaches have also been proposed by Yuizono et al. [16] and Hurley et al. [6] respectively. The first ever ear recognition system tested with a larger database of 415 subjects is proposed by Yan and Bowyer [14]. Using a modified version of the ICP, they achieved an accuracy of 95.7% with occlusion and 97.8 % without occlusion (with an Equal-error rate (EER) of 1.2%). The system does not work well if the ear pit is not visible.

In this work, we have adopted the work of Islam et al. [8] for detecting the ear from 2D profile images and then, extended it for cropping the corresponding 3D profile face data. After ear detection, we apply a variant of the Iterative Closest Point (ICP) algorithm for recognition of the ear at different mesh resolutions of the extracted 3D ear data. Using two different resolutions hierarchically, we obtain a rank 1 recognition rate of 93%. The proposed system is fully automatic and does not rely on the presence of a particular feature of the ear (e.g. ear pit). It also does not require an accurate extraction of the ear contour and hence reduces the computational cost. Besides, the ear recognition results can be combined with other biometric modalities such as 2D and 3D faces to obtain a more robust and accurate human recognition system.

The paper is organized as follows. The proposed system for 3D ear detection and that for 3D ear recognition are described in Section 2 and 3 respectively. Results obtained are reported and discussed in Section 4. Section 5 concludes our findings.

2. 3D Ear Detection and Normalization

The proposed 3D ear detection approach utilizes the Adaboost based 2D ear detection technique developed by Islam et al. [8]. Since the 3D profile data are co-registered with corresponding the 2D images, the 2D ear detector is first scanned through the whole profile image to localize the ear. A rectangle is placed covering the ear. The 3D data corresponding to this rectangular region is cropped for use in 3D ear recognition. The complete ear detection process from 2D and 3D face profile image/data is shown in Figure 1. A sample of a profile image and corresponding 2D and 3D ear data detected by our system is also shown in the same figure.

Figure 1. A sample of a profile image and corresponding 2D and 3D ear data detected by our system is also shown in the same figure.

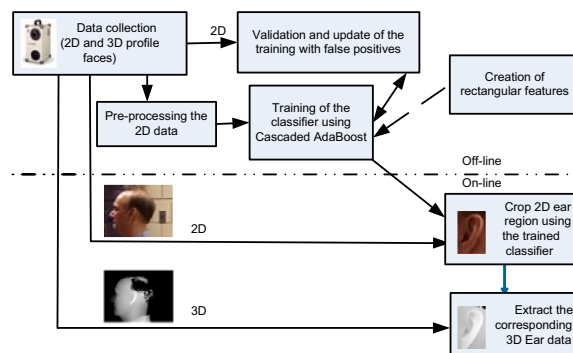


Figure 1. Block diagram of the 3D ear detection approach.

Once the 3D ear is detected, we remove all the spikes and holes by filtering the data. We perform triangulation on the data points, remove edges longer than a threshold of 0.6 and finally, remove disconnected points. The data is then normalized by shifting to its mean and then, uniformly sampled using a grid of size 82 by 56 pixels.

3. 3D Ear Matching and Recognition

The extracted 3D ear data of a subject (also called probe data) is matched with one (for authentication) or all (for identification) 3D ear data (also called gallery data) stored in the gallery database built off-line. Matching can be performed based on the error of registering between the two data sets, more specifically, two clouds of points. The ICP algorithm [2] is considered as one of the most accurate algorithms for this purpose. However, it is computationally expensive and may converge to a local minimum if the two data sets are not nearly registered. To minimize these limitations, we adopted a coarse-to-fine hierarchical technique. ICP is first applied on low and then, on high resolution meshes of the 3D ear data.

The mesh reduction was performed using the surface simplification algorithm of Garland and Heckbert [5] as it preserves features on the surfaces of a mesh (unlike the sampling reduction which removes points without any regard to the features). The degree of mesh reduction can be controlled by fixing the number of triangles to be remained in the reduced mesh or by fixing the surface error approximations (using quadric matrices). We achieved better result using the later option (see Section 4.3).

Initially, reduced meshes (see Figure 2) of the probe (created on-line) and the gallery (created off-line) ear data are used for coarse registration with the ICP. The rotation and translation resulting from this coarse registration are applied to the original data set and then, the ICP algorithm is applied to them to get a finer match. The matching approach which

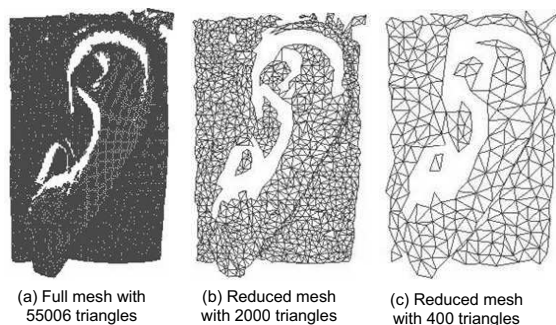


Figure 2. Sample of the full and reduced meshes of the extracted 3D ear data.

we term as the two-step ICP is shown in the flowchart of figure 3. Subscripts 'p' and 'q' in the flowchart are used for the probe and the gallery respectively.

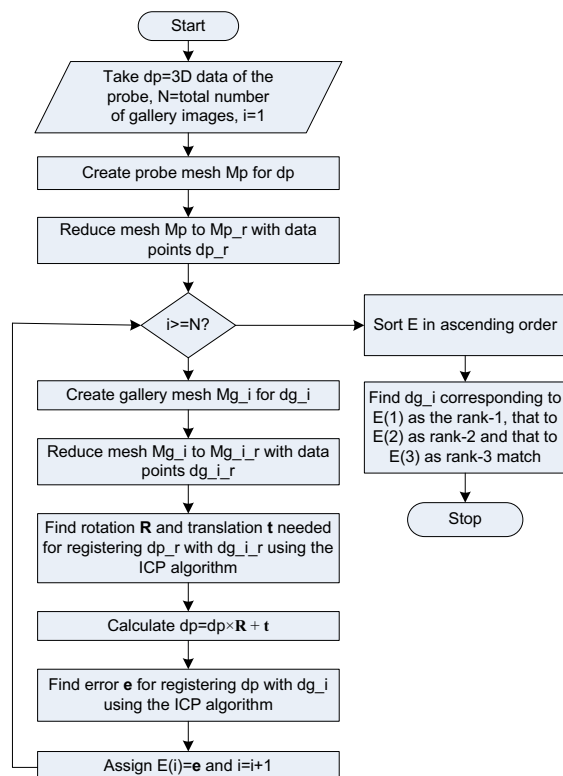


Figure 3. Flowchart of the matching algorithm using coarse-to-fine hierarchical technique with ICP.

4. Results and Discussion

In this section, results of our recognition system using both two and three levels of mesh resolution are reported. The reasons for misclassifications are also discussed.

4.1. Dataset Used

The recognition performance of the proposed system was evaluated against two datasets collected from the UND Biometrics Database [13, 14]. Dataset A consists of arbitrarily selected 200 profile images of 100 different subjects each with the size of 640 by 480 pixels. Among the chosen images, 100 images collected in the year 2003 are used in the gallery database and another 100 images of the same subjects collected in the year 2004 are used for the probe database. No image of this dataset was used in the training of the detection classifiers. In Dataset B, the whole UND database is used excluding the images of two subjects (due to data error). 300 images taken in the year of 2003 are used as gallery. One image out of multiple images of the same subjects is arbitrarily chosen as probe.

4.2. Recognition Rate of the Single-step ICP

Using Dataset A, we obtained recognition rate of 93% and 94% rank one and rank ten respectively with the single-step ICP. Testing with Dataset B provided recognition rate of 93.98%, 94.31%, 95.31% and 96.32% rank one, two, three and ten respectively. Figure 4 shows a sample of the correct recognition. It is noticed that the system works even in the presence of partial occlusions due to hair and ear-rings.

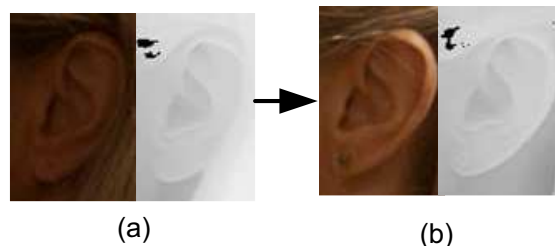


Figure 4. Example of recognition: (a) 2D and range image of the gallery ear (b) 2D and range image of the probe ear with a small ear-ring and hair (This figure is best seen in color).

4.3. Improvement Obtained with the Two-step ICP

The recognition rate with Dataset A improved to 93%, 94% and 95% for rank 1, rank 2 and rank 7 respectively when we applied the two-step ICP as mentioned in Section 3. In our experiment, we choose quadric error factor of 10 for the mesh reduction. The improvement is shown by the plots in Figure 5. It is worth mentioning that the recognition result becomes worse (84%, 88% and 88% respectively) when we fix the number of triangles to be 400 for mesh reduction. This is due to the fact that the size of the detected window is not same for all the ear images.

The two-step ICP provides better results in the cases where the single-step ICP converged to a local minimum and also where the probe and the gallery images differ

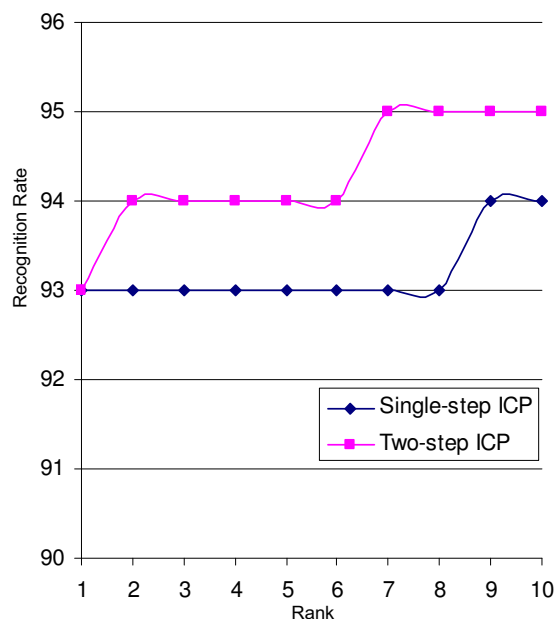


Figure 5. Recognition rates with Single-step and Two-step ICP.

slightly in rotation and translation. An example of this improvement is shown in figure 6.

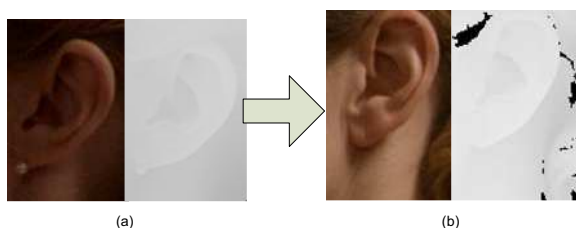


Figure 6. Examples of improvement with the two-step ICP which was not correctly recognized by the single-step ICP [left image is the gallery and right one is the probe].

4.4. Analysis of the Misclassification

Both single-step and two-step ICP failed to recognize some of the probe images. Visual inspection of these probe images and the corresponding gallery images indicates that the ICP algorithm cannot work properly if there is large variation in pose (rotation and translation) as shown in Figure 7-(a, b, c), occlusion by hair or ear-rings as shown in Figure 7-(a, d) or missing of data as in Figure 7-(e) is present. The pose variation might occur during the capture of the probe profile image as mentioned in section 4.1 or by the detection program (see Figure 7-(d)).

5. Conclusion

Our method for ear detection and recognition from profile images is fully automatic and robust to some degrees of occlusion due to hair and ear-rings. One of its major

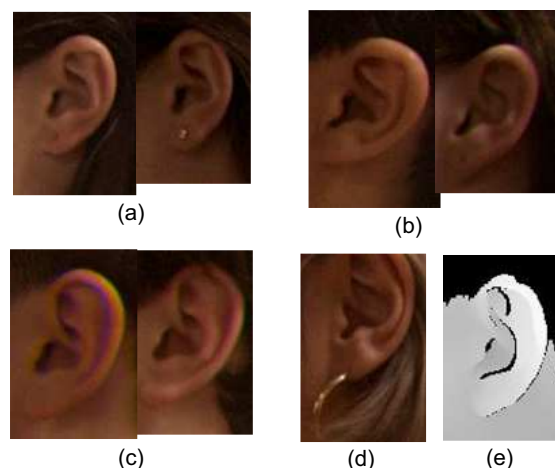


Figure 7. Examples of misclassification:(a, b, c) 2D image of a gallery (left) and the corresponding probe (right) with pose variations, (d) 2D image of a probe with occlusions (e) Range image of a probe having missing data.

strengths is that it does not assume accurate prior detection of the ear. It also does not rely on the presence of particular features like the ear pit. Thus, it is suitable for any non-intrusive biometric application. In this paper, it is also shown that the recognition performance of the ICP algorithm improves with the proposed hierarchical approach. ICP performance may be further improved by providing better initial registration. This can be done by invariant feature matching. The recognition system can be made more robust to occlusion by representing the extracted ear data with some occlusion-invariant representations. These include our future research tasks.

Acknowledgements

Authors acknowledge the use of the University of Notre Dame Biometrics Database for 3D ear detection and recognition. This research is sponsored by ARC grants DP0664228 and DP0881813.

References

- [1] L. Alvarez, E. Gonzalez, and L. Mazorra. Fitting ear contour using an ovoid model. *In Proc. of Int'l Carnahan Conf. on Security Technology, 2005.*, pages 145 – 148, Oct. 2005.
- [2] P. J. Besl and N. D. McKay. A Method for Registration of 3-D Shapes. *IEEE Trans. on PAMI*, 14(2):239–256, 1992.
- [3] M. Burge and W. Burger. Ear biometrics in computer vision. *In Proc. of the ICPR'00*, pages 822 – 826, Sept 2000.
- [4] M. Choras. Ear biometrics based on geometrical feature extraction. *Electronic Letters on Computer Vision and Image Analysis*, Vol. 5:84–95, 2005.
- [5] M. Garland and C. Heckbert. Surface Simplification Using Quadric Error Metrics. *In SIGGRAPH*, 1997.

- [6] D. J. Hurley, M. S. Nixon, and J. N. Carter. Force field feature extraction for ear biometrics. *Computer Vision and Image Understanding*, 98(3):491–512, June 2005.
- [7] A. Iannarelli. *Ear Identification. Forensic Identification Series*. Paramount Publishing Company, Fremont, California, 1989.
- [8] S. Islam, M. Bennamoun, and R. Davies. Fast and Fully Automatic Ear Detection Using Cascaded AdaBoost. *Proc. of IEEE Workshop on Application of Computer Vision WACV 2008*, Jan. 2008.
- [9] S. Islam, M. Bennamoun, R. Owens, and R. Davies. Biometric Approaches of 2D-3D Ear and Face: A Survey. *Proc. of Int'l Conf. on Systems, Computing Sciences and Software Engineering, SCSS 2007*, Dec. 2007.
- [10] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre. Xm2vtsbd: The extended m2vts database. *In Proc. of the 2nd Conf. on Audio and Video-base Biometric Personal Verification, Springer Verlag, New York*, pages 1 – 6, May 1999.
- [11] K. H. Pun and Y. S. Moon. Recent advances in ear biometrics. *In Proc. of the Sixth IEEE Int'l Conf. on Automatic Face and Gesture Recognition*, pages 164 – 169, May 2004.
- [12] R. Schapire and Y. Singer. Improved boosting algorithms using confidence-rated predictions. *Mach. Learn.*, 37(3):297–336, 1999.
- [13] University of Notre Dame Biometrics Database. <http://www.nd.edu/~cvrl/UNDBiometricsDatabase.html>.
- [14] P. Yan and K. W. Bowyer. Biometric recognition using 3d ear shape. *IEEE Trans. on PAMI*, 29(8):1297 – 1308, Aug. 2007.
- [15] L. Yuan and Z.-C. Mu. Ear detection based on skin-color and contour information. *In Proc. of the Int'l Conf. on Machine Learning and Cybernetics*, Vol. 4:2213 – 2217, Aug. 2007.
- [16] T. Yuizono, Y. Wang, K. Satoh, and S. Nakayama. Study on individual recognition for ear images by using genetic local search. *In Proc. of Congress on Evolutionary Computation*, pages 237–242, 2002.