Biometric Analysis of Human Ear Matching Using Scale and Rotation Invariant Feature Detectors

Soumyajit Sarkar and Guanghui Wang

Department of EECS, University of Kansas, Lawrence, KS 66045
ghwang@ku.edu

Abstract. Biometric ear authentication has received enormous popularity in recent years due to its uniqueness for each and every individual, even for identical twins. In this paper, two scale and rotation invariant feature detectors, SIFT and SURF, are adopted for recognition and authentication of ear images; an extensive analysis has been made on how these two descriptors work under certain real-life conditions; and a performance measure has been given. The proposed technique is evaluated and compared with other approaches on two data sets. Extensive experimental study demonstrates the effectiveness of the proposed strategy.

Keywords: Biometrics, image matching, SIFT, SURF, ear recognition.

1 Introduction

Biometric authentication of people based on various anatomical characteristics, like eye, ear, face, iris, and fingerprint have attracted lots of attention during the past few decades, and some of these techniques have already been successfully applied for recognition and authentication. However, many systems are not very robust and may fail to work under certain conditions. Biometric ear recognition is a relatively new technique that may surpass the existing systems due to several significant reasons. For example, the acquisition of ear images is relatively easy and, unlike iris, can be captured without the co-operation of individuals [1].

Human ear contains rich and stable features which are more reliable than face features, as the structure of the ear is not subject to change with age. It has also been found out that no two ears are exactly the same even for identical twins [3]. The detailed structure of ear is not only very unique but also permanent, since the shape of a human ear never shows drastic changes over the course of life. The research on ear identification was first conducted by Bertillon, a French criminologist, in 1890. The process was refined by American police officer, Iannarelli [20], who divided the ear based on various distinctive features of seven parts: i.e. helix, concha, antihelix, crux of helix, intertragic notch, tragus, and antitragus [3].

In this Paper, we propose to use two scale and rotation invariant feature detectors, i.e. SIFT (scale invariant feature transform) and SURF (speed up robust features), for ear recognition. Both SIFT and SURF extract specific interest points from an image and generate descriptors for the feature points to a form a reliable matching results.

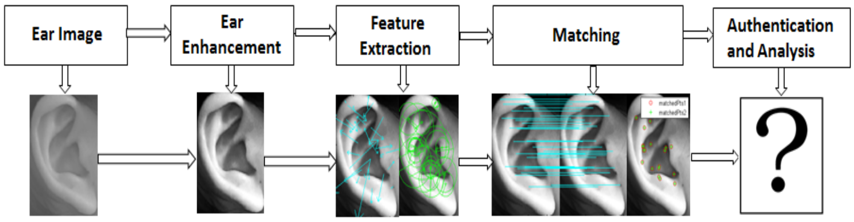


Fig.1 The pipeline of the propose ear recognition system

Extensive experiments have been carried out on two different sets of databases to evaluate their performance with respect to various rotations and scales. One of the most important feature of ear images is its easiness in acquisition, however, the acquired images may be in different scales, rotations, and illumination. The scale and rotation invariant property of the SIFT and SURF algorithms makes them perfect for ear authentication under various circumstances.

The rest of the paper is organized as follows. Some background and related research are discussed in Section 2; the proposed method is presented in details in Section 3; some experimental results and analysis are given in Section 4; and the paper is concluded in Section 5.

2 Background and Related Research

Human ears start to develop between fifth and seventh weeks of pregnancy. At this stage, the embryo face takes on more definition as mouth perforation, nostrils and ear indentations become visible. Forensic science literature reports that ear growth after the first four months of birth is highly linear [20]. The rate of stretching is five times greater than normal during the period from 4 months to the age of 8, after which, it is constant until the age of seventy when it again increases. Thus it can be said that ear remains almost unchanged during a substantial period of 62 years and, thus, it is one of the strong points of considering ear for biometric authentication.

An ear biometric system can be viewed as a typical pattern recognition problem, where the input image is reduced to a set of features that are subsequently used to compare against the feature sets of the other images in the database in order to find a best match to determine identity [2]. Based on this finding, a good amount of research has been done on ear recognition [5][8][10].

Haar-based methods have given fairly better results for face detection as it is robust and fast. The different types of ear recognition systems include those of intensity-based, force-field based, 2D curves geometry, wavelet transformation, Gabor filters, SIFT, and 3D features. The force-field transforms gained popularity due to its uniqueness and efficiency [22]. Similar methods have also been implemented on other kinds of ear recognition systems [8][10].

3 The Proposed Method

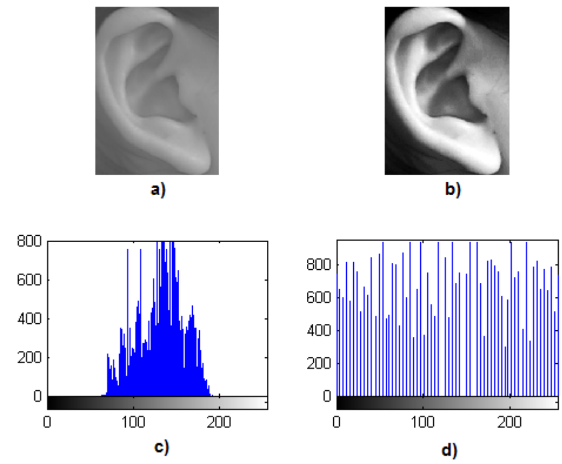
Real-life ear images can be acquired in various formats with different scaling and rotation conditions. In this paper, we propose to use scale and rotation invariant feature detectors to describe interested features and match them with other images in the databases. The proposed ear recognition technique is shown in Figure 1. Below is a brief description of each function block.

3.1 Ear Image Enhancement

The ear enhancement process starts with contrast enhancement, where we apply histogram equalization to improve the contrast in an image in order to stretch out its intensity range, from which, we get an enhanced version of the original image by maximizing the contrast level of an image, as shown in Figure 2.

It has been experimentally found that after contrast stretching, both the SIFT and SURF detectors are able to find more feature points. Thus, image enhancement is an essential step of the entire process. As shown in Figure 3, 10 SIFT keypoints were detected in the original image, while 32 features were detected in the enhanced image.

Fig.2 Image enhancement result.
a) Original image;
b) enhanced image;
c) histogram distribution of the original image;
d) histogram distribution of the enhanced image



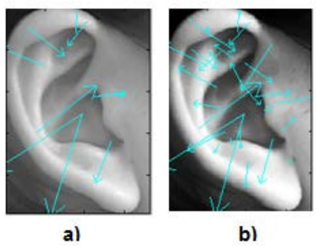


Fig.3 a) 10 SIFT features detected in the original image; b) 32 SIFT features detected in the enhanced image.

3.2 Feature Extraction & Matching

Feature Extraction is the process of extracting salient features from the image, and each feature is described by a vector which summarizes the required information for that point [2]. Features are extracted exclusively in order for the image to be matched with the features of the input image to authenticate the ear so that a decision can be made. In this paper, two rotation and scale invariant features are studied.

Speed up robust features (SURF): SURF is a high performance, fast scale and rotation invariant point detector and descriptor. It outperforms previously proposed schemes with respect to repeatability, distinctiveness and robustness [9]. The detector is based on the Hessian matrix and uses a very basic Laplacian-based detector, called difference of Gaussian (DoG). The implementation of SURF can be divided into three main steps. First, interest points are selected at distinctive locations in the image, such as corners, blobs, and T-junctions. Then, the neighborhood of every interest point is represented by a feature vector. This descriptor has to be distinctive and robust to noise, detection errors, and geometric and photometric deformations. Finally, the descriptor vectors are matched between different images. When working with local features, the issue that needs to be settled is the required level of invariance. Here the rotation and scale invariant descriptors seem to offer a good compromise between feature complexity and robustness to commonly occurring deformations, skew, anisotropic scaling, and perspective effects [9].

Given a point in an Image, the Hessian matrix is defined as follows:

$$H(x,\sigma) = \begin{bmatrix} L_{XX}(x,\sigma) & L_{Xy}(x,\sigma) \\ L_{Xy}(x,\sigma) & L_{yy}(x,\sigma) \end{bmatrix} \tag{1}$$

where $L_{XX}(x,\sigma)$ is the convolution of the Gaussian second order derivative $\frac{\partial^2}{\partial x^2}g(\sigma)$ at the point. This method leads to a novel detection, description and subsequent matching steps. Using relative strengths and orientations of gradient reduces the effect of photometric changes. Figure 4 shows the detection results with respect to rotation and scale change. As shown in Section 4, it has been found that though SURF is rotation invariant, its performance in matching, i.e. matching score, decreases sharply when the images are rotated or scaled. The SURF features are not stable over various rotation angles and scale changes.

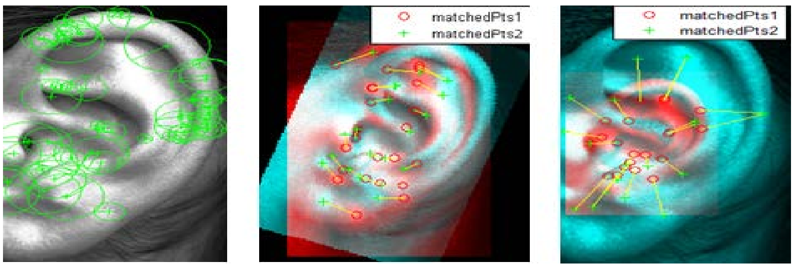


Fig.4 The detected SURF features (left) and matching result under rotation (middle) and scale change (right).

Scale Invariant Feature Transform (SIFT): The SIFT features are invariant to image scaling and rotation and shown to provide robust matching across a substantial range of affine distortion, change in 3D viewpoint, addition of noise, and change in illumination. The computation stages of SIFT are as follows.

Step 1. Scale space extrema detection: The first step is to construct a Gaussian scale over all the locations. It is implemented efficiently by using a difference of Gaussian (DoG) to identify potential interest points. The 2D Gaussian operator $G(x,y,\sigma)$ is convolved with the input image $I(x,y)$:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{2}$$

where the DoG images are obtained by subtracting the subsequent scales in each octave.

$$G(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \tag{3}$$

Step 2. Accurate keypoint localization: Once a keypoint has been detected, a detailed model is fitted to determine its location and scale. The keypoints are selected based on measures of their stability. Further details can be found in [16].

Step 3. Orientation assignment: One or more orientations are assigned to each keypoint location based on local image gradient directions. All future operations are performed on image data that has been transformed relative to the assigned orientation, scale, and location for each feature.

Step 4. Keypoint descriptor: The local image gradients are measured at selected scale in the region around each keypoint. They are transformed into a certain representation that allows for significant levels of local shape distortion and shape illumination.

Figure 5 shows an evaluation of the SIFT detector. It is evident the SIFT keypoints are very stable when the images are rotated and scaled. The scaling results are much better compared to the rotation results in our experiments.

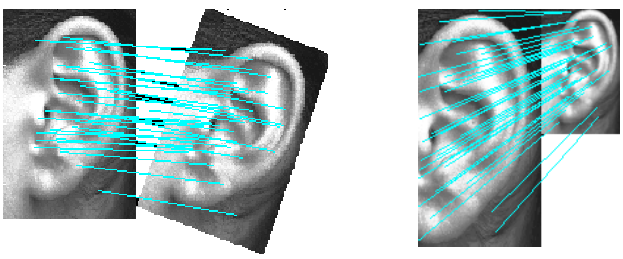


Fig.5 The matching results of SIFT detectors under rotation (left) and scale change (right).

Matching of Ear Images: Image matching is the process by which the features extracted from the SIFT and the SURF descriptors of the input image are being matched with the features already computed and stored in the database. Figure 6 shows the matching results using the SURF and SIFT detectors, where the nearest neighbor is defined as the keypoint with the minimum Euclidean distance for the invariant descriptor vector.

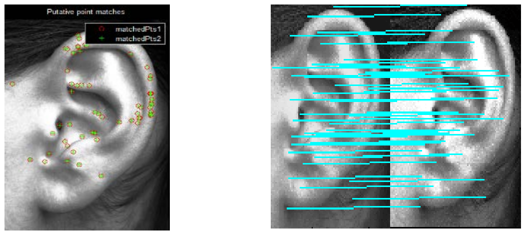


Fig.6 The matching results of SURF detector (left) and SIFT detector (right).

4 Experimental Results and Analysis

The proposed approach has been evaluated on two data sets. One is the AMI database [7], which consists of 175 ear images; and the other is the IIT Delhi database [5], which consists of 494 images of 125 distinct persons. The images were all converted to gray-scale images for ease of work. It has also been found out that contrast enhancement is an important factor for feature detection and matching, because it makes the feature detectors find better set of keypoints and increase the effectiveness of matching.

According to the experiments performed, it has been found that upper helix, antihelix, and tragus are the most important regions for feature selection compared to others. These regions contribute to about 64% of the feature points.

Figure 7 shows some sample images from the two databases we used for our experiments. The graphs in Figure 8 indicates the average number of keypoints found and matched by SIFT and SURF detectors when the images are rotated from a range of 0 to 180 degrees. The results suggest that the SIFT detector is fairly stable over a variation of angles from 20 to 160 degrees, whereas the SURF detector, though faster and rotation invariant, is not very stable.

Table 1 shows the keypoints detected and matched by the SIFT and SURF detectors, where the performance ratio is the ratio of the number of matched points to that of detected features. It is obvious that the SIFT algorithm performs better when the sizes of images are decreased, while the SURF algorithm performs better when the image sizes are increased. However, the amount of detected keypoints by the SIFT detector is always higher than that by the SURF detector.

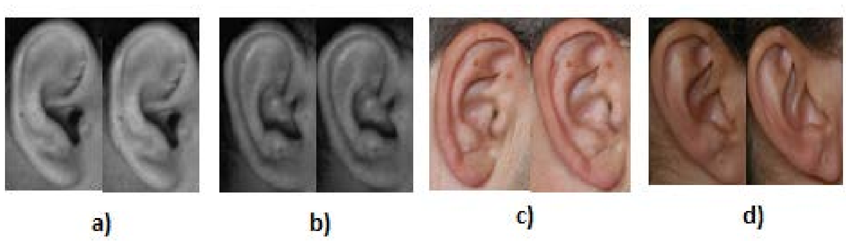


Fig.7 Sample images from the IIT Delhi database a) & b) and the AMI database c) & d)

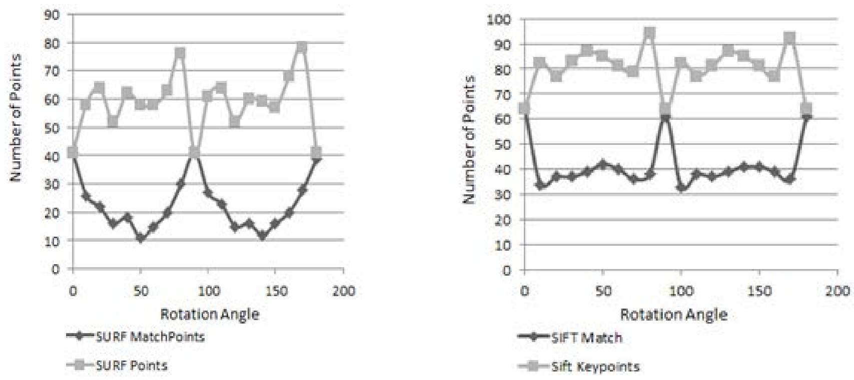


Fig.8 A comparison result of the detected and matched keypoints by SURF and SIFT

Table 2 shows an overview of how the two detectors work in real-life conditions where some images are not matched due to illumination changes as those images were mostly taken at night and at different angles. Thus, the descriptors fail to find enough feature keypoints for matching. The overall recognition rates of the SIFT and SURF algorithms on the IIT Delhi database are 96.8% and 94.4%, respectively. As a comparison, we also implemented other methods for ear recognition. The template matching technique yields a recognition rate of 93% for [24], and 92.6% for [23], whereas the recognition rate by the contour extraction technique [25] is 85%. It is evident that the proposed technique yields a higher recognition rate.

Table 1. SIFT and SURF detection and matching results at different scales

Scaling		0.25	0.5	0.75	1.0	2.0	3.0	4.0
Number of Features	SIFT	28	53	58	64	170	247	233
	SURF	3	12	30	41	39	41	44
Number of Matches	SIFT	24	45	52	64	53	47	51
	SURF	2	9	23	41	20	21	16
Performance Ratio	SIFT	0.85	0.85	0.89	1	0.32	0.20	0.22
	SURF	0.67	0.75	0.75	1	0.51	0.51	0.30

Table 2. Experimental results on the IIT Delhi database

Method	Number of images	Matched images	Unmatched images	Time for matching (s)	Recognition rate (%)
SIFT	125	121	4	0.21	96.8
SURF	125	118	7	0.183	94.4

5 Conclusion

In this paper, we have studied two scale and rotation invariant feature detecters and their application to ear recognition. Although both the SIFT and the SURF are invariant under scale and rotation changes, their performance decreases under certain conditions. The SIFT detector is more stable than the SURF detector under rotation changes. It is also found that the SIFT algorithm performs better for image decreasing, in contrast, the SURF algorithm performs better for image increasing. Experimental evaluations have demonstrated the effectiveness of the proposed techniques in ear recognition. In future study, we will further investigate how to increase the performance and reliability of the proposed approach.

References

1. A. Pflug, C. Busch ,Ear Biometrics: A survey of Detection, Feature Extraction and recognition Methods, IET Biometrics, Volume 1, Issue 2, June 2012, p. 114 – 129.
2. A. Abaza, A. Ross, C. Hebert, M. Ann, F. Harrison, M.S.Nixon, A Survey on Ear Biometrics, ACM Computing Surveys(CSUR), Volume 45, Issue 2, February 2013, Article No. 22.
3. A. Tariq, M Usman Akram, Personal Identification using Ear Recognition, TELKOMNIKA, Vol.10, No.2, June 2012, pp. 321-326.
4. C. Harris and M.J. Stephens, A combined corner and edge detector, 4th Alvey Vision Conference, Manchester, UK, 1988, pp. 147-151.
5. A. Kumar and C. Wu, Automated Ear Identification using Ear Imaging, Pattern Recognition, vol. 41, no. 5, March 2012.
6. M. Burge, W. Burger, Ear Biometrics in Computer Vision. 15th International Conference of Pattern Recognition. 2000; 822-826.
7. E. Gonzalez, L. Alvarez, L. Morazza, AMI Ear Database, Centro de I+D de Tecnologias de la Imagen.
8. M. Zichun, Y . Li, X. Zhengguang, Shape and Structural Feature Based Ear Recognition,Advances in Biometric Person Authentication,Guangzhou,China,2004; 663-670.
9. H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. "SURF: Speeded Up Robust Features." Computer Vision and Image Understanding (CVIU).Vol. 110, No. 3, pp. 346–359, 2008.
10. S. A. Daramola, O. D Oluwaninyo, Automatic Ear Recognition System using Back Propagation Neural Network, International Journal of Video & Image Processing and Network Security, IJVIPNS-IJENS. 2001;11(1):28-32.
11. D. G. Lowe, Object recognition from local scale-invariant features, International Conference on Computer Vision, Corfu, Greece(September 1999),pp. 1150-1157.
12. P. Scovanner, S. Ali, M. Shah, A 3-dimanesional sift descriptor and its applications to action recognition, In proceedings of the 15th international conference on Multimedia'07. ACM, New York, NY, USA, 357-360.
13. A. Gilbert, J. Illingworth, R. Bowden, Fast Realistic multi action recognition using mined dense spatio-temporal features, Computer Vision, 2009 IEEE 12th International Conference, pgs 925-931.

14. G. Lema, L. Di Martino, S. Berchesi, A .Fernandez, F. Lecumberry, J Preciozzi, Evaluation of a face recognition system performance's variationon a citizen passport database, IEEE Computing Conference(CLEI), pg 1-6.
15. F. Alonso-Fernandez, P. Tome-Gonzalez, V. Ruiz-Albacete, J. Ortega-Garcia, Iris recognition based on SIFT Features, pg 1-8,IEEE 2009 International Conference on Biometrics.
16. D. G. Lowe, Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision, 60,2(2004),pp. 91-110.
17. A. Jain, A. Ross, S. Prabhakar, An introduction to biometric recognition, IEEE Transactions on Circuits and Systems for Video Technology,14(1),4-20.
18. R. C. Gonzalez, R. C. Woods, Digital Image Processing, 2nd Edition, Prentice Hall, 2002.
19. T. Lindeberg, Feature detection with automatic scale selection, IJCV,30(2):79-116,1998.
20. A. Iannarelli, Ear Identification, Forensic Identification Series, Paramount Publishing Company, 1989, Fremont, CA.
21. S. U. Park, S. Pankanti, A. K. Jain, Fingerprint verification using SIFT Features, Defense and Security Symposium, Biometric Technologies for Human Identification, BTHI, Proc. SPIE,2008.
22. D. Hurley, M. Nixon, J. Carter, Automatic ear recognition by force field transformations. In Proceedings of the IEEE Colloquium on Visual Biometrics. 7/1–7/5.
23. H. Chen and B. Bhanu, "Human ear detection from side face range images," in Proceedings of International Conference on Pattern Recognition, ICPR' 04, vol. 3, pp. 574–577, IEEE Computer Society, 2004.
24. S. Ansari and P. Gupta 2007. Localization of ear using outer helix curve of the ear. In Proceedings of the IEEE International Conference on Computing: Theory and Applications. 688–692.
25. P. Yan and K. Bowyer, Biometric recognition using 3D ear shape. IEEE Trans. Pattern Anal. Mach. Intell. 29, 8, 1297–1308.
26. L. Yuan and F. Zhang, Ear detection based on improved adaboost algorithm. In Proceedings of the 8th IEEE International Conference on Machine Learning and Cybernetics (ICMLC).