

Robust Log-Gabor Filter for Ear Biometrics

Banafshe Arbab-Zavar, Mark S. Nixon
University of Southampton, Southampton SO17 1BJ, UK
baz05r, msn@ecs.soton.ac.uk

Abstract

Ears are a new biometric with major advantage in that they appear to maintain their structure with increasing age. Expanding on our previous parts-based model, we propose a new wavelet approach. In this, the log-Gabor filter exploits the frequency content of the ear boundary curves. Extending our model description, a specific aim of the new approach is to capture information in the ear's outer structures. Ear biometrics is also concerned with the effects of partial occlusion, mostly by hair and earrings. By localization, intuitively a wavelet can offer performance advantage when handling occluded data. We also add a more robust matching strategy to restrict the influence of erroneous wavelet coefficients. Significant improvement is observed when we combine the model and the log-Gabor filter, and we will show that this improvement is maintained as the ears get occluded.

1. Introduction

Ears have long been considered as a potential means of personal identification, yet it is only in the last 10 years that machine vision researchers started to tackle using ears as a biometric. Ears have appealing properties for personal identification: they have a rich structure that appears to be consistent with age from a few months after birth. Clearly, ears are not affected by facial expressions. Images of ears can be acquired without the subject's participation and ears are big enough to be captured from a distance. However there exists a big obstacle — the potential occlusion by hair and earrings, which is almost certain to happen in uncontrolled environments.

One of the first ear biometric works was introduced by Burge and Burger [5]. They modeled each ear with an adjacency graph. Hurley et al. [9] used force field feature extraction to map the ear to an energy field. Chen et al. [6] and Yan et al. [12] exploited the 3D

structure of the ear. Though using 3D can improve performance, using 2D images is consistent with deployment in surveillance or other planar image scenarios. An up-to-date survey of ear biometrics has recently been provided by Hurley et al. [8].

Recognizing the likelihood that an ear can be occluded, in our earlier work [3], we have developed the first model-based ear biometric method. A model is explicit in its approach to identification and it is known for being robust in noise and occlusion. Our data-driven model is built by detecting the clusters of SIFT points in a training set of ear images. In the following section we identify a potential improvement for the model, and describe how a wavelet technique can be used to exploit it. A wavelet approach is chosen specifically to extract the frequency content of the fluctuating surface which gives the shapes of the Helix and the Anti-helix in an ear. Furthermore the wavelet process is localized, meaning that it can provide us with a maximum amount of uncorrupted information when occlusion occurs. Prior to applying the wavelet transform, we automatically rearrange and align the image data into templates. We have chosen the log-Gabor filter as our wavelet. We shall describe this wavelet and the details of the input templates in section 3, and in keeping with the objective of handling occlusion, we also add a more robust matching strategy. The results of recognizing occluded and unoccluded ear images are discussed in section 4, followed by conclusions.

2. Improving the parts-based model

We have previously built a model for ears [3] using a stochastic clustering on the SIFT [10] interest points which were detected on the training images. The clusters of SIFT points were described by their descriptors, locations, scales and orientations. However, the inherent disadvantage of using SIFT points is that the model is unable to convey the information which is not accurately captured in a SIFT descriptor. Such is the case with boundaries and stretched curves which are not

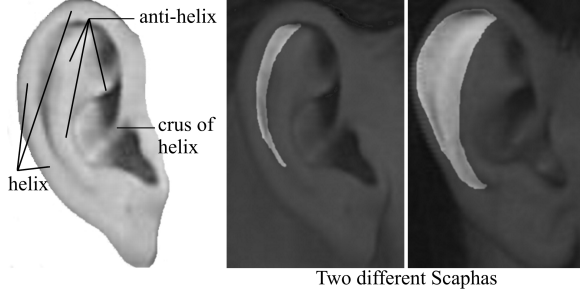


Figure 1. Ear anatomical parts

well captured in a circular-based area descriptor such as SIFT, and thereby these are under-presented in the model. Two such curves are the Helix and the Anti-helix — the loss of which might be significant. The Helix is the outer ear rim and it is attached to the Anti-helix along one side (see figure 1). However, there is also the Scapha, which is a concave surface of free portion lying between the Anti-helix and the Helix. The Scapha allows the Helix to have rather an independent shape from the small curves and ridges within it. Thus the hope of finding new and discriminant information arises from the possibility of independence in Helix. Figure 1 shows an example of two different Scaphas.

Given an estimate of the Helix location, a wavelet approach can be used to describe this curve. The model parts are used to vote for the position of the Helix. A new template is then formed by sampling the image intensities along lines which are mostly normal to the Helix curve.

3. Approach

Wavelets can obtain the frequency content of a signal while being localized in spatial domain. This decimation in spatial domain as well as in frequency is the main advantage of wavelets compared to the Fourier transform.

3.1. Preparing the templates

Prior to applying the wavelets, we prepare aligned templates of image data. These templates are built in such a way to facilitate the detection of helix-related features. The template is the sampled image intensities in a semi-circular region which includes the Helix. The chosen centre of this semi-circle is where the Crus of helix curves inwards, which is almost the midpoint of the ear height and is situated on the outermost part of the ear, opposite to the Helix. This is detected using some of the more reliably detected model parts which

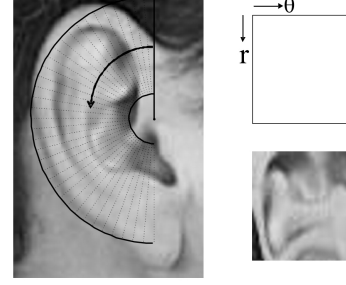


Figure 2. Input template

vote for its position. Once the centre is determined, the image is sampled along radial lines which sweep the ear (see figure 2).

In this template the ridge of the Anti-helix and the Scapha reside in the rows, and thus the columns exhibit the variation between the ridge of the Anti-helix and the Helix in each specific angle. Small discrepancies in rotation and scale of the templates can be accounted for in matching, with little loss of information, via shifting the template horizontally or/and vertically.

3.2. Log-Gabor filter

The log-Gabor filter [7] is a derivative of the standard Gabor filter, which is widely used in image processing. The log-Gabor frequency response is Gaussian in logarithmic frequency scale as opposed to the standard Gabor which has Gaussian frequency response in linear scale. The log-Gabor frequency response is described by:

$$G(\omega) = e^{-\frac{(\log(\omega/\omega_0))^2}{2(\log(k/\omega_0))^2}} \quad (1)$$

where ω_0 is the centre frequency of the filter and the bandwidth is determined by the k/ω_0 term. The main advantage of the log-Gabor filter is that its DC component is always zero, thus the filter bandwidth is not limited to 1 octave, as it is for Gabor filters, and thereby a lesser number of filters can be used to cover a desired spectra. We use a one dimensional log-Gabor filter on the columns of the templates, $T(r, \theta)$. Thus the projected image $P(r, \theta)$ is obtained by:

$$P_n(r, \theta) = F^{-1}[F(T(r, \theta)) \times G_n(\omega)], \theta = \frac{\pi}{2} \cdot \frac{3\pi}{2} \quad (2)$$

where F and F^{-1} denote the Fourier and the inverse Fourier transforms, and $G_n(\omega)$ is a log-Gabor filter at a scale n . This corresponds to convolving a template with the wavelet in the spatial domain. Both even and odd symmetric responses are used in matching.

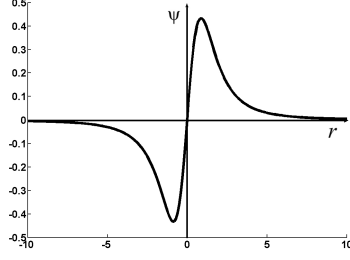


Figure 3. The robust ψ function

3.3. Robust matching

The log-Gabor filter provides us with localized information. Thus the local frequency is not contaminated by the clutter or occlusion of surrounding regions. However the local information of occluded or otherwise corrupted regions is still present in the result. In classification, we use a simple nearest neighbour approach, which includes a distance measure. We could use the Euclidean distance. However this arrangement corresponds to the least squares estimation, which is known to be intolerant of outliers. The least squares method minimizes the error function E :

$$E(\mathbf{r}) = \sum_i \rho(r_i) \quad , \quad \rho(r_i) = r_i^2 \quad (3)$$

where \mathbf{r} is the residual error. To see the influence of the outliers on the error function more clearly, an *influence function* ψ is defined as:

$$\psi(r_i) = \frac{d\rho(r_i)}{dr_i} \quad . \quad (4)$$

In the square error case with $\psi = 2r_i$, the data points retain a weight proportional to their residual value, no matter how far from the mean they are. Thereby an outlier imposes a big influence on the evaluation.

Black and Jepson [4] identified a similar problem in PCA, and used a more robust ρ function with good effect. They showed that the standard mapping of images into the eigenspace corresponds to the least squares estimation, and thus the mapping of noisy images produces poor results. Their solution is to use a more robust mapping. The ρ function of their choice is:

$$\rho(r_i, \sigma) = \frac{r_i^2}{\sigma^2 + r_i^2} \quad ; \quad \psi(r_i, \sigma) = \frac{2r_i\sigma^2}{(\sigma^2 + r_i^2)^2} \quad , \quad (5)$$

where ρ is a function of σ as well as the residuals. In this the value for σ determines the point at which the influence of potential outliers begins to decrease. It can be seen in figure 3 that the new ψ function eliminates

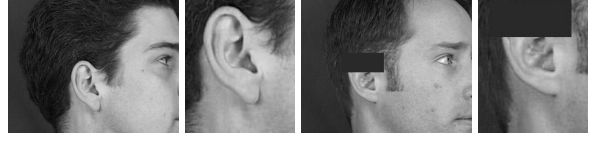


Figure 4. Automatic enrolment

the influence of big residuals. Inspired by this, we use the same ρ function (5) in our distance measure, D :

$$d_{r,\theta,n}(i, j) = [P_n(r, \theta)]_i - [P_n(r, \theta)]_j$$

$$D(T_i, T_j) = \sum_{r,\theta,n} \rho[d(i, j), \sigma] \quad . \quad (6)$$

4. Results

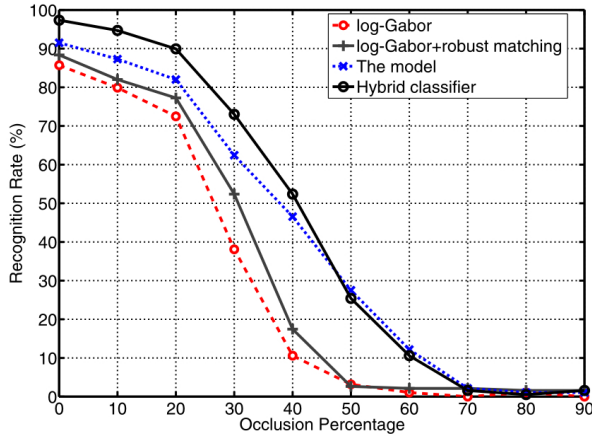
To evaluate our techniques, we have used 252 images from 63 individuals, selected from the head-profiles of the XM2VTS [11] database, wherein these 63 individuals are those whose ear is not obscured by hair. There are 4 images per individual in our database, which are taken in 4 different sessions over a period of five months, and thereby ensuring natural variation between the images of the same subject. One image from each subject is used for training of the model, whilst the remainder are used for performance evaluation.

We have developed an automatic ear enrolment method [2] which finds the position of the ear; an image including the ear is then cropped (see figure 4). In this, the ear is detected using a reduced Hough Transform for ellipses [1] accruing tolerance to noise and occlusion. Further cues from the orientation and size of the ear as well as the characteristics of its edge points help eliminating the erroneous votes in the Hough accumulator space and thereby pinpoint the correct ear position. Our automatic enrolment finds the ear in all 252 images. The details of this work can be found in [2].

Recognition using the nearest-neighbour algorithm on the model parts, obtains 91.5% correct recognition. Exploiting the new log-Gabor coefficients alone, we achieve an 85.7% recognition rate. In this we have used nearest-neighbour algorithm on the Euclidean distances of the log-Gabor coefficients. Exchanging the Euclidean distance for the more robust distance metric(6) improves the performance to 88.4%. We then combine the two, log-Gabor and the model, using a simple decision fusion technique of sum of the normalized scores. The hybrid classifier exhibits a significant improvement with a 97.4% recognition rate, thus suggesting that although the log-Gabor does not perform as well as the model, it contains information which is new and independent of that which is presented in the model.

Table 1. Recognition results

Model	Log-Gabor	Robust log-Gabor	Hybrid
91.5%	85.7%	88.4%	97.4%

**Figure 5. Recognizing occluded ears**

4.1 Occlusion test

Handling occlusion is one of our main objectives in dealing with ear biometrics and at each stage we have chosen our techniques accordingly. We have shown that our automatic enrolment performs well in the presence of occlusion [2], and in recognition we have compared the performance of the model against PCA, and shown that the model is superior at recognizing occluded ears [3]. Figure 5 shows the recognition results against the increasing occlusion. It can be seen that the robust matcher improves the log-Gabor outcome consistently, and thus it is a worthwhile addition. For this test, synthetically occluded probes are automatically enrolled and compared against a gallery of unoccluded images. We occlude the ears from the top, similar to the occlusion which is caused by hair. Figure 4 includes an example in which a 40% occluded ear has been enrolled.

By the new information added from the log-Gabor process, we have seen that the hybrid classifier is superior to both, the model and log-Gabor, separately. This improvement is maintained as the ears are increasingly occluded. However the log-Gabor measures seem to be more affected by the occlusion and thus the hybrid performance gradually approaches that of the model and it eventually drops below it at around 50% of occlusion.

5. Conclusions

We have shown that a log-Gabor filter can be used to further improve our previous model-based approach for

ear biometrics. The log-Gabor filter exploits the variations between the boundary curves of the ear, which are under-presented in the model. The hybrid classifier exhibits a significant improvement with a 97.4% recognition rate. The destructive effects of occlusion were moderated first via the localized approach of the log-Gabor filter and then by applying a more robust matching strategy which restricts the influence of erroneous coefficients. Thereby the hybrid classifier advantage over the model-based approach is consistent even for the occluded samples.

References

- [1] A. S. Aguado, E. Montiel, and M. S. Nixon. On using directional information for parameter space decomposition in ellipse detection. *Pattern Recognition*, 29(3):369–381, 1996.
- [2] B. Arbab-Zavar and M. S. Nixon. On shape-mediated enrolment in ear biometrics. In *International Symposium on Visual Computing (ISVC'07)*, Nevada/California, November 2007.
- [3] B. Arbab-Zavar, M. S. Nixon, and D. J. Hurley. On model-based analysis of ear biometrics. In *IEEE Conference on Biometrics: Theory, Applications and Systems (BTAS'07)*, Washington DC, September 2007.
- [4] M. J. Black and A. D. Jepson. Eigentracking: Robust matching and tracking of articulated objects using a view-based representation. *International Journal of Computer Vision*, 26(1):63–84, 1998.
- [5] M. Burge and W. Burger. Ear biometrics. In A. Jain, R. Bolle, and S. Pankanti, editors, *BIOMETRICS: Personal Identification in a Networked Society*, pages 273–286. Kluwer Academic, 1998.
- [6] H. Chen and B. Bhanu. Human ear recognition in 3d. *IEEE Trans. PAMI*, 29(4):718–737, 2007.
- [7] D. J. Field. Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America*, 4(12):2379–2394, 1987.
- [8] D. J. Hurley, B. Arbab-Zavar, and M. S. Nixon. The ear as a biometric. In A. Jain, P. Flynn, and A. Ross, editors, *Handbook of Biometrics*. 2008.
- [9] D. J. Hurley, M. S. Nixon, and J. N. Carter. Force field feature extraction for ear biometrics. *Computer Vision and Image Understanding*, 98:491–512, 2005.
- [10] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.
- [11] K. Messer, J. Matas, J. Kittler, J. Luetttin, and G. Maitre. Xm2vtsdb: The extended m2vts database. In *Proc. AVBPA*, Washington D.C., 1999.
- [12] P. Yan and K. W. Bowyer. Biometric recognition using 3d ear shape. *IEEE Trans. PAMI*, 29(8):1297–1308, 2007.