Wonder Ears: Identification of Identical Twins from Ear Images

Hossein Nejati†, Li Zhang†, Terence Sim

National University of Singapore

nejati@nus.edu.sg, {lizhang,tsim}@comp.nus.edu.sg

Elisa Martinez-Marroquin[1]‡, Guo Dong[2]

[1] *University of Canberra, Australia*

[2] *Facebook Inc., Palo Alto, CA*

Abstract

While identical twins identification is a well known challenge in face recognition, it seems that no work has explored automatic ear recognition for identical twin identification. Ear image recognition has been studied for years, but Iannarelli (1989) appears to be the only work mentioning the twin identification (performed manually). We here explore the possibility of automatic twin identification from their ear images based on a psychological model for face recognition in humans, known as Exception Report Model (ERM). We test our approach on 39 pairs of identical twins (78 subjects), with several levels of resolution, occlusion, noise, left vs. right ear, and feature optimization which verifies the robustness of the introduced features.

1 Introduction

The incidence of twins has progressively increased in the past decades. Twins birth rate has risen to 32.2 per 1000 birth with an average 3% growth per year since 1990 [8]. Since the increase of twins, identical twins are becoming more common, which in turn, is urging biometric identification systems to accurately distinguish between twin siblings. The significant similarity between identical twins is known to be a great challenge for face recognition systems and the performances of current 2D face recognition systems on twins have been recently questioned [12]. Several researchers have shown encouraging results in automatic recognition systems that using other features such as fingerprint, palmprint, iris, speech, and combinations of some of the above biometrics [12].

In this work we focus on identification of identical twins based on ear images. Ear images have several ad-

vantages over many other features. The ear shape does not change significantly after adulthood, its surface has a relatively uniform color distribution, it is invariant to expression, and ear images are more robust to illumination and head pose changes than features like faces [2]. The early studies mostly addressed the question of uniqueness of ears. The most well-known pioneer seems to be Iannarelli (1989), in which he performed manual identification over 10,000 ears and found no indistinguishable ears. This work along with previous works such as Hirschi (1970), Rother (1976), Hunger and Hammer (1987), and more recent works such as Van der Lugt (1998) mostly indicate that the variability between ears is large enough to assume ears as unique identification features.

On the basis of the above studies, automatic ear recognition techniques have been introduced, mostly employing methods used in other biometric fields. Eigen-ears [11] could provide high accuracy in recognition in closely controlled conditions, otherwise, having dramatic performance reduction. In order to handle rotation and illumination changes in ear images, Abate et al. [1] introduced a method based on Generic Fourier Descriptors. Yan also presented a complete system [14] including automated segmentation of the ear in a profile view image and 3D shape matching for recognition under constrained conditions with specialized cameras. Bustard and Nixon [3] recently proposed an ear registration method that utilized SIFT features followed by a homography transformation, to cope with the occlusion and pose changes. The transformed images are then masked and matched using Euclidean distance. Despite high performance, their semi-automatic ear masking procedure occasionally fails to match correctly to the ear area. Although several aspects of ear recognition have been explored, there seems to be no work on automatic twins identification using ear images.

We propose our approach based on a psychological model, originally suggested for face perception in humans, known as Exception Report Model (ERM) [13]. The ERM consists of two main suggestions: (1) the pos-
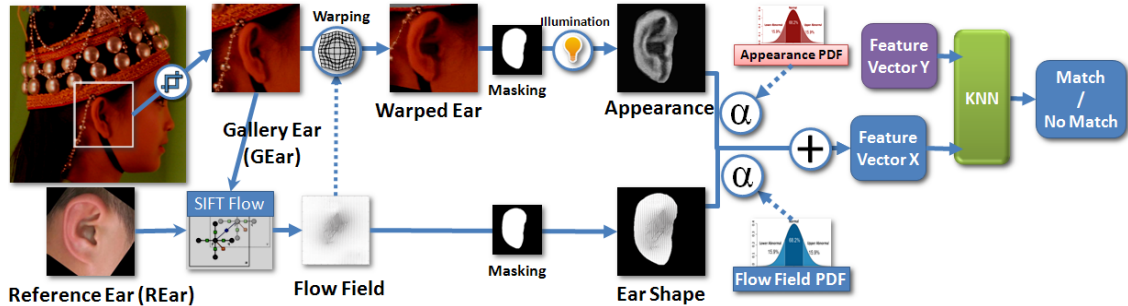

Figure 1. The block diagram of our algorithm.

sibility of accurate face recognition by focusing more on abnormal features and less on normal features (ERM possibility), and (2) the optimality of the use of only about 10% of the features (only abnormal features) for rapid and accurate recognition (ERM optimality). The ERM has been used in some automatic face recognition methods such as [5, 9], but not for ear recognition. Our proposed system consists of two parts, namely, ear image normalization, and feature weighting and verification. Our continuation in the first part is to normalize and use *both shape and appearance* of the ear for recognition. Our contribution in the second part is to weight points in the ear shape and appearance *based on their level of abnormality*. We finally train a K-Nearest Neighbor classifier to verify whether two given ear images belong to the same subject.

We evaluate the performance of our algorithm on a dataset of 39 pairs of identical twins (ERM possibility) against 5 resolution levels, 4 occlusions levels, and 4 noise levels, as well as left ear vs. right ear training-testing sets. We also test the ERM optimality for left and right ears. Performances of our algorithm in these experiments suggest the applicability of ERM to a wider range of automated visual tasks than only faces.

2   Ear Image Normalization

In first part of our algorithm, ear normalization (see Fig. 1), given a gallery ear image (GEar), we crop the ear out of the profile view, and then normalize the rotation, scale, and illumination, based on a reference ear image (REar). Normalization in previous works has been performed based on both manually *sparse* point registration, both manually [6, 4] and automatically, using methods such as SIFT feature matching [3] and graph matching [2]. However, when all ears images are transformed into a single reference image coordinates, the 3D structure of the ear (i.e. ear shape) is lost and merely the intensity values (i.e. ear appearance) is remained. In contrast, we calculate and store the dense correspondence between each GEar and the REar using SIFTFlow. We not only use this dense flow for the scale and rotation normalization, but also treat the flow itself as the relative shape information of each GEar.

We normalize the GEar images in five steps. In the first step we loosely crop a window of 300×300 pixels out of the profile view (originally 1728×1152 pixels), around the ear-hole, detected using a simple image correlation (Fig. 1, crop symbol). The cropping is only to reduce the search window of the SIFTFlow algorithm in the next step. In the second step we apply the SIFTFlow to calculate accurate dense flow field between the cropped window and the REar (Fig. 1, flow field). Then in the third step we warp the GEar image to the REar image coordinates based on the flow field, thus normalizing its scale and rotation (Fig. 1, warping). As we are only interested in the pixels corresponding to the ear, in the fourth step we mask out the non-ear pixels in both the flow field and the warped ear using a single pre-defined binary image, manually defined for the REar image (Fig. 1, masking). Finally, in the fifth step we normalize the illumination of warped image using Contrast-limited adaptive histogram equalization (CLAHE) [15] (Fig. 1, illumination). At the end of this part of our algorithm, we have normalized ear shape (i.e. the masked flow field) and ear appearance (i.e. the masked, illumination normalized warped image), shown in Fig. 1.

3   Feature Weighting and Verification

In the second part of our algorithm, feature weighting and verification, we apply the ERM concept to weight each point. The ERM suggests that the importance of a feature has a direct relationship with the abnormality of that feature and the further a point is from its related mean value, the more abnormal it is. Thus, we first estimate a normal probability density function (PDF) for each shape and appearance point, based on the distribution of their values in the entire dataset. Then we define the abnormality strength (weight) of point k, as the distance of that point from the mean,
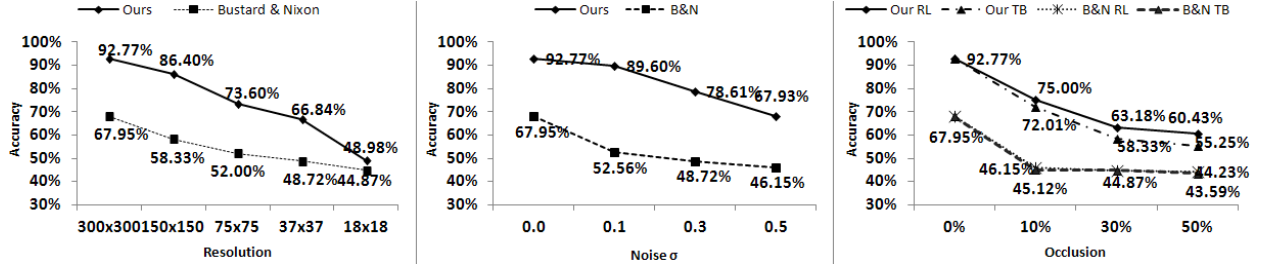
Figure 2. Results of sibling verification across different resolution, noise, and occlusion levels

normalized by the sigma, in the corresponding PDFs:

$$w_{shape,k} = \sqrt{(\frac{x_k - \mu_{x,k}}{\sigma_{x,k}})^2 + (\frac{y_k - \mu_{y,k}}{\sigma_{y,k}})^2}$$

$$w_{int,k} = (\frac{int_k - \mu_{i,k}}{\sigma_{i,k}})$$

where $w_{shape,k}$ is the shape weight; $w_{int,k}$ is the intensity weight; x_k and y_k are the X and Y coordinates; int_k is intensity value of point k; and $(\mu_{x,k}, \sigma_{x,k})$, $(\mu_{y,k}, \sigma_{y,k})$, and $(\mu_{i,k}, \sigma_{i,k})$ are mean and sigma values of the corresponding X coordinate, Y coordinate, and intensity PDFs. Finally, we form the feature vector by concatenating weighted shape and appearance points, to represent a GEar image (see Fig. 1):

$$
\begin{aligned}
FeatureVector_i &= W_{shape}^T \times S || W_{int}^T \times I \\
W_{shape} &= [w_{shape,1}, w_{shape,2}, \ldots, w_{shape,n}] \\
W_{int} &= [w_{int,1}, w_{int,2}, \ldots, w_{int,n}]
\end{aligned}
$$

where $FeatureVector_i$ is the concatenated vectors representing GEar image i, S is the ear shape values, and I is the ear intensity values.

Given a pair of weighted feature vectors, we now train a KNN classifier, using Mahalanobis distance, to verify whether the vectors representing the two ears, belong to the same subject. Our choice of a simple classifier such as KNN is to observe discriminative power of the proposed abnormal features in ear recognition.

4 Experiments

We evaluate the performance of our algorithm in verification of ear images from 39 pairs of twins (78 subjects). Our Twin dataset is the largest publicly available image dataset of twins, obtained in the Sixth Mojiang International Twins Festival, China, 2010, containing Chinese, Canadian and Russian subjects, each having 2 to 4 real and 20 synthesized images. Real images are captured from profile view, containing the head and shoulder, with some translation and 3D rotation. Synthesized images are obtained from real images by adding random noise, translation, 3D rotation, and realistic motion blur (Xu & Jia 2010).

In our tests, the task is given a pair of ear images, we verify whether both ears belong to the same subject (siblings are treated as different subjects). We tested five experiments on totally 5000 pairs (1792 positive pairs 36%, and 3208 negative pairs 64%). We tested five resolutions (300×300, 150×150, 75×75, 37×37, and 18×18 pixels), four noise levels (Gaussian noise with $\mu = 0$ and $\sigma = 0$, 0.1, 0.3, and 0.5), four occlusion levels (simulated 0%, 10%, 30%, and 50%, in right-to-left and top-to-bottom directions). We also tested different ear side training-testing sets (two cases of training and testing on the same side (right or left), and two cases of training and testing on different sides). Finally, motivated by the optimality claim of the ERM, we test accuracy of our algorithm by applying feature optimization on the point abnormality strength (when only points with at least a minimum abnormality strength, $dist$, is used), pruning the points as follows:

$$w_{int,k}(dist) = \begin{cases} w_{int,k} \ if & w_{int,k} > dist \\ 0 & o.w. \end{cases}$$

5 The Results and Discussion

We compare accuracy results of our algorithm compared with ear recognition in [3] (B&N) in Fig. 2 and 3, and Table 1. Our algorithm performs up to 92% on the Twins dataset, constantly better than B&N. Results also show robustness of our algorithm regarding resolution, noise, and occlusion. Fig. 2 indicates that even with noise $\sigma = 0.5$, our accuracy is almost the same as the B&N without noise. This may be because of the SIFT-Flow dense point registration, but it also indicates that the abnormal features are robust against noise.

Comparing occlusion variation accuracy results in Fig. 2 with other tests, it seems that our algorithm is robust towards resolution and noise than the occlusion. This can be because of the loss of strong features (trained in the non-occluded images) in the occlusion variations, while these features, although weakened, are still present in the resolution and noise. In addition, as the accuracy drops more rapidly in the top-to-bottom occlusion curve, it seems that strong features are lo-

Training-Testing	L-L	R-R	L-R	R-L
Accuracy %	92.77	92.76	54.78	53.40

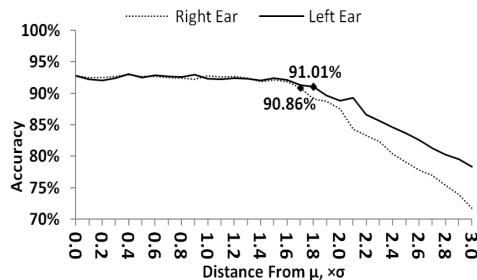**Table 1. Results of training and testing with left and right ears.**

Figure 3. Results of feature optimization.

cated more at the top of the ears in our dataset, rather than right of the ears.

Results of training and testing on same or different side ears, presented in Table 1, show that the left and right ears in our subjects do not share much of their abnormal features. This means that one cannot train only on one side ears and hope to accurately recognize ears from both sides.

Finally, the feature optimization results are presented in Fig. 3 that agrees with the Exception Report Model (ERM) optimality claim that using only about 10% of the features, the brain can accurately and rapidly recognize faces. Similarly, Fig. 3 shows that even using only features with distance more than 1.7σ from the μ (almost only the top 5%), we can still achieve more than 90% accuracy.

6 Conclusions

We are the first to address the problem of automatic twin ear verification by using both shape and appearance of ears, and motivated by Exception Report Model (ERM), a psychological framework for the perception of faces by the brain [13], which has shown good results in face recognition before (see e.g. [7, 9]). We showed that, similar to face recognition, by focusing on abnormal (exceptional) features of the ear shape and appearance, we can accurately identify twins (up to 92%). In our experiments on 39 pairs of twins with different age, gender, and ethnicity, we showed the robustness of our algorithm against variations of resolution, noise, and occlusion. These experiments also showed the abnormality features are not the same in the right and left ears. Finally feature optimization showed that with only

the top 5% features we can achieve a fast and accurate recognition.

In conclusion, our results suggest that ears are not only a powerful identification feature for regular subjects, but also for identical twins, where many other approaches e.g. face recognition have failed [10]. In addition to addressing the twin identification from ear images, our work in this paper suggests that the ERM, although originally suggested for face recognition, may be applicable to a wider range object recognition problems, which may also help simulating new frameworks in human visual system studies.

References

[1] A. F. Abate, M. Nappi, D. Riccio, and S. Ricciardi. Ear recognition by means of a rotation invariant descriptor. In *ICPR, Vol 04*, pages 437–440, 2006.

[2] M. Burge and W. Burger. Ear biometrics in computer vision. In *ICPR*, volume 2, pages 822 –826 vol.2, 2000.

[3] J. D. Bustard and M. S. Nixon. Toward unconstrained ear recognition from two-dimensional images. *Trans. Sys. Man Cyber. Part A*, 40:486–494, May 2010.

[4] K. Faez, S. Motamed, and M. Yaqubi. Personal verification using ear and palm-print biometrics. In *SMC*, pages 3727–3731, 2008.

[5] M. F. Hansen and G. A. Atkinson. Biologically inspired 3d face recognition from surface normals. *ICEBT*, 2:26–34, 2010.

[6] A. Iannarelli. Ear identification. 1989.

[7] L. N. S. M. F. Hansen, G. A. Atkinson and M. L. Smith. 3d face reconstructions from photometric stereo using near infrared and visible light. *ICEBT*, 114:942–951, 2010.

[8] J. Martin, H. Kung, T. Mathews, D. Hoyert, D. Strobino, B. Guyer, and S. Sutton. Annual summary of vital statistics: 2006. *Pediatrics*, 2008.

[9] H. Nejati, T. Sim, and E. Martinez-Marroquin. Do you see what i see?: A more realistic eyewitness sketch recognition. *IJCB*, 2011.

[10] P. Phillips, P. Flynn, K. Bowyer, R. Bruegge, P. Grother, G. Quinn, and M. Pruitt. Distinguishing identical twins by face recognition. In *FG 2011*, pages 185 –192, 2011.

[11] M. I. Saleh, W. C. L., and P. Schaumont. Eigen-face, eigen-ears. using ears for human identification. Master's thesis, Virginia Polytechnic Institute and State University, 2007.

[12] Z. Sun, A. Paulino, J. Feng, Z. Chai, T. Tan, and A. Jain. A study of multibiometric traits of identical twins. *SPIE*, 2010.

[13] M. Unnikrishnan. How is the individuality of a face recognized? *J. Theor. Biology*, 261(3):469 – 474, 2009.

[14] P. Yan and K. Bowyer. Biometric recognition using 3d ear shape. *TPAMI*, pages 1297–1308, 2007.

[15] K. Zuiderveld. Contrast limited adaptive histogram equalization. pages 474–485, 1994.