Subdivided Ear Recognition

DAVID ROMERO, PETER PEER, ŽIGA EMERŠIČ

Abstract: The present paper addresses the performance of different ear feature extractors on partial ear images. The main goal is to find out which parts of the ear have major influence on successful recognition for each extractor. In this sense, a whole ear recognition pipeline has been simulated using Annotated Web Ears (AWE) Toolbox and dataset [1]. Ears have been divided in up, down, internal and external parts and results have been compared. It has been demonstrated the existence of performance gaps between different ear parts and extractors. Trying to exploit that, a score level distance fusion approach has been tested combining separately obtained distances by means of weighted averaging.

Keywords: • 1. Ear Recognition • 2. Biometrics • 3. Computer Vision

---

ADDRESS OF THE AUTHORS: University of Ljubljana, Faculty of Computer and Information Science, Večna pot 113, 1000 Ljubljana, Slovenia, e-mail: dr7289@student.uni-lj.si, {peter.peer, ziga.emersic}@fri.uni-lj.si

1. INTRODUCTION

Image based biometric subject recognition has always been an active research topic in computer vision field. During the years, several physical metrics (face, iris…) and feature extractor algorithms have been explored. The present work is focused on ear recognition, a promising topic that still has several research issues unsolved.

According to [1], feature extraction, used for ear recognition, can be divided into four different categories: geometric, holistic, local and hybrid. In this paper local methods have been addressed. They are based on the principle of evaluating the whole ear texture without looking for specific points or shapes. Because of that, it is interesting to address if different ear parts have different influence on recognition performance of these extractors.

This information is especially useful when facing one of the most typical issues: occlusion. In some cases, clear images of ears cannot be acquired and recognition must be performed only on ears occluded by hair, glasses or earrings. Furthermore, depending on the occluded ear part recognition systems could suffer different performance degradation rates. Thus, it is important to investigate if these differences actually exist and if so, how big they are for each feature extractor.

Despite helping to better overcome occlusion it is also interesting to address if this performance asymmetry between different ear parts could be somehow exploited to improve recognition rates when whole ear images are available

2. SUBDIVIDED EAR RECOGNITION

We perform recognition using two different subdivision approaches: up-down and internal-externa. For internal-external subdivision the whole dataset has been first divided into left-right images. The idea is to simulate separately up-down and external-internal parts of left and right ears and look for asymmetries in the performance.

In order to get an overall and consistent picture of the topic, simulations have been carried out using several feature extractors. The following ones have been tested: Local Binary Patterns (LBP) [2], Binarized Statistical Image Features (BSIF) [3],[4] Local Phase Quantization (LPQ) [5], Rotation Invariant LPQ (RLPQ) [6], Patterns of Oriented Edge Magnitudes (POEM) [7], Histogram of Oriented Gradients (HOG) [8],[9] and Dense Scale Invariant Feature Transform (DSIFT) [10]. All of them are already implemented in AWE toolbox and belong to local approach category.

For score level fusion distances from different parts of the ear have been combined based on a simple weighted averaging (1) where α defines the averaging weight and goes from 0 to 1.

$$d_{comb} = \alpha\, d_1 + (1-\alpha) d_2 \tag{1}$$

3. RESULTS

To obtain partial ear images for the experiments, we used *AWE* dataset. It contains 1000 images of 100 subjects that have been subdivided following the criteria shown in Figure 1. Identification results based on the AWE dataset and ear division is described in the continuation of this section.

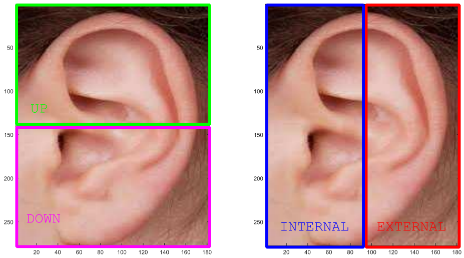

Figure 1: Ear subdivision approaches.

3.1 Up-Down comparison

For up-down comparison Protocol 2 of AWE toolbox has been used. A previously defined distribution of a train set of 600 images has been used to test the performance of each algorithm based on 5 K-Fold and chi square distance.

Obtained Rank-1 absolute and degradation values are displayed in Table 1. The Gap column shows the difference between upper and down parts. At the same time, Cumulative Match Characteristic (CMC) curves are plotted in Figure 2.

Table 1: Rank-1 values for Up-Down comparison

DATASET	*AWE*	*Up*	*Deg.*	*Down*	*Deg.*	*Gap*
LBP	43.6 ± 7.2	36.3 ± 5.0	7.3	34.7 ± 7.5	8.9	+1.6
BSIF	48.5 ± 6.9	42.9 ± 3.5	5.6	36.8 ± 9.4	11.7	+6.1
LPQ	43.3 ± 7.7	39.2 ± 5.1	4.1	32.4 ± 9.0	10.9	+6.8
RILPQ	43.5 ± 9.3	36.2 ± 5.1	7.3	31.2 ± 7.4	12.3	+5.0
POEM	49.0 ± 6.9	40.1 ± 5.7	8.9	36.4 ± 7.9	12.6	+3.7
HOG	43.5 ± 8.0	30.9 ± 3.3	12.6	34.0 ± 8.0	9.5	-3.1
DSIFT	43.6 ± 8.5	34.1 ± 6.1	9.5	33.9 ± 13.5	9.7	+0.2

It is clear that, for most of the algorithms, upper part of the ear has a higher influence in successful recognition. HOG and DSIFT are the only descriptors where upper part is equal or worse than down part. Especially for HOG, the performance degradation for both up and down parts is highly noticeable.

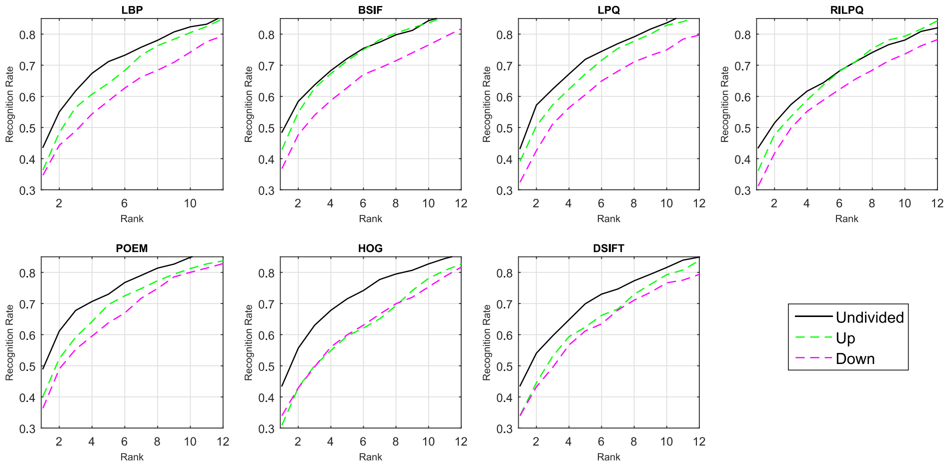

Figure 2: CMC curves for Up-Down comparison

On the other hand, for BSIF, LPQ and RILPQ obtained upper part curves are surprisingly close to ones with the whole ear.

3.1 External-Internal comparison

In this case, internal and external parts of left and right ears have been separately tested following AWE Protocol 1. Thus, the entire datasets have been tested with a random distribution, 5 K-Fold evaluation and chi square distance. Obtained Rank-1 results for both left and right ears are shown in Tables 2 and 3 respectively.

Table 2: Rank-1 values for Left Internal-External comparison

DATASET	*Left*	*Left Int.*	*Deg.*	*Left Ext.*	*Deg.*	*Gap*
LBP	57.1 ± 6.3	55.0 ± 7.1	2.1	42.9 ± 4.0	14.2	+12.1
BSIF	59.4 ± 6.9	62.7 ± 6.8	-3.3	44.4 ± 4.6	15.0	+18.3
LPQ	58.2 ± 6.0	54.0 ± 6.8	4.2	38.6 ± 4.9	19.6	+15.4
RILPQ	56.7 ± 6.5	47.2 ± 5.8	9.5	40.6 ± 5.3	16.1	+6.6
POEM	62.7 ± 4.5	47.7 ± 6.4	15.0	43.5 ± 6.3	19.2	+4.2
HOG	57.9 ± 5.3	41.7 ± 5.0	16.2	41.9 ± 4.8	16.0	-0.2
DSIFT	56.9 ± 5.2	51.7 ± 6.9	5.2	44.4 ± 2.8	12.5	+7.3

According to the results, internal part appears to be more relevant than external part. All extractors have achieved better or at least equal recognition rates for this part of the ear and this difference is consistent for both left and right ears.

The performance gap is not constant for all the extractors. HOG shows again practically no gap while BSIF and LPQ show the biggest one. Again, BSIF achieves almost the same performance for internal part and whole ear images. LPQ and DSIFT also get quite close to the whole ear curve.

Table 3: Rank-1 values for Right Internal-External comparison

DATASET	*Right*	*Right Int.*	*Deg.*	*Right Ext.*	*Deg.*	*Gap*
LBP	52.4 ± 9.9	48.3 ± 2.1	4.1	38.5 ± 1.9	13.9	+9.8
BSIF	57.7 ± 6.3	55.2 ± 3.7	2.5	38.7 ± 2.7	19.0	+16.5
LPQ	52.9 ± 5.2	48.8 ± 4.3	4.1	34.2 ± 2.2	18.7	+14.6
RILPQ	49.4 ± 6.5	42.9 ± 4.0	6.5	34.4 ± 3.6	15.0	+8.5
POEM	58.9 ± 7.1	44.2 ± 3.6	14.7	39.4 ± 3.7	19.5	+4.8
HOG	54.8 ± 8.1	40.0 ± 3.2	14.8	36.5 ± 4.7	18.3	+3.5
DSIFT	51.2 ± 7.5	47.3 ± 3.0	3.9	30.2 ± 3.9	21.0	+17.1

3.1 Distance fusion

Separately obtained distances for up and down parts have been combined according to equation (1) for three different values of α: 0.5, 0.65 and 0.8. Obtained CMC curves for up-down are depicted in Figure 3. As it is shown, proposed distance combination always works equal or better than partial ear images separately. However, in general these improvements are not big enough to outperform recognition rates obtained with whole ear. Interestingly, BSIF is the only feature extractor that achieves a global improvement.
Same distance fusion approach has been applied to internal-external ear parts with similar results.

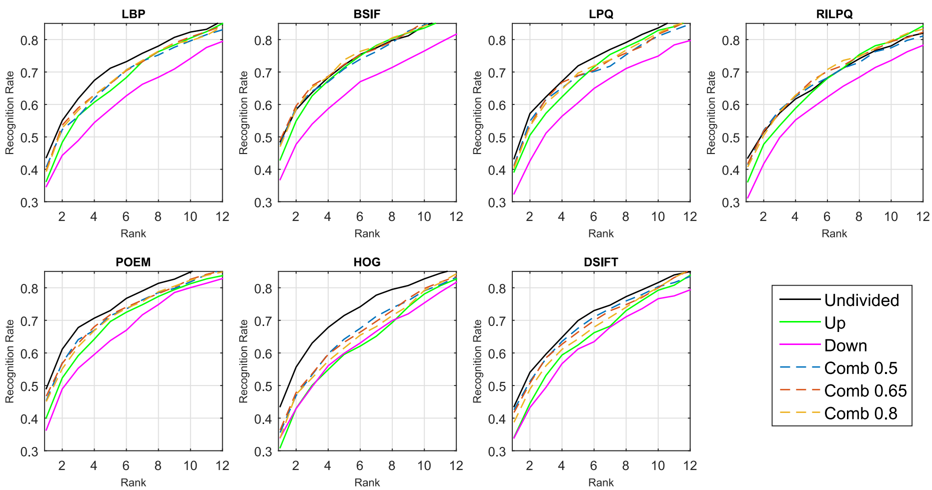

Figure 3: CMC curves for Up-Down distance combining

4. DISCUSSION AND FUTURE WORK

Results show that, according to Rank-1, there is a performance gap of up to +7 between upper and down parts of the ear. At the same time, the gap between internal and external parts is up to +18.3. However, these differences are variable between different feature extractors. The different nature of each extractor could be an explanation for these differences.

Although simulated weighted averaging distance fusion method showed small improvements, the existence of a performance gap suggests that it could be possible to exploit it to improve the overall performance. In this sense, some future research steps can be identified. First, smaller and more complex ear parts could be tested. Secondly, new fusion and feature extractor methods could be developed to exploit performance asymmetry between ear parts. Finally, feature extractors that have shown smaller degradation should be strongly considered in future occluded ear recognition systems.

5. CONCLUSIONS

Throughout this paper, the influence of different ear parts over several feature extractor algorithms has been addressed. Based on AWE toolbox and dataset, simulations have been carried out with images of only one part of the ear (up, down, external or internal).

According to obtained results, we can conclude that, for most of the feature extractors, upper and internal parts have stronger influence than down and external parts. This performance gap between internal-external parts is bigger than the one for up-down parts. Both performance gap and degradation differ significantly depending on used feature extractor. We propose to use here presented degradation values as a metric of the sensitivity to occlusion of a feature extractor.

At the same time, separately calculated distances have been combined based on a basic weighted averaging. A small improvement has been achieved mostly for BSIF extractor but results show that, for most of the cases there is no significant improvement. However, this should be considered as a first step to exploit existing performance gap since a lot of research work remains to be done: different subdivision strategies or development of new feature extractors or fusion models based on obtained results.

LITERATURE

1. Ž. Emeršič, V. Štruc, and P. Peer, “Ear recognition: More than a survey,” Neurocomputing, 2017.

2. Y. Guo and Z. Xu, “Ear recognition using a new local matching approach,” in Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on. IEEE, 2008, pp. 289–292.

3. A. Benzaoui, A. Hadid, and A. Boukrouche, “Ear biometric recognition using local texture descriptors,” Journal of Electronic Imaging, vol. 23, no. 5, p. 053008, 2014.

4. A. Benzaoui, N. Hezil, and A. Boukrouche, “Identity recognition based on the external shape of the human ear,” in 2015 International Conference on Applied Research in Computer Science and Engineering (ICAR). IEEE, 2015, pp. 1–5.

5. A. Pflug, C. Busch, and A. Ross, “2d ear classification based on unsupervised clustering,” in Biometrics (IJCB), 2014 IEEE International Joint Conference on. IEEE, 2014, pp. 1–8.

6. V. Ojansivu, E. Rahtu, and J. Heikkila, “Rotation invariant local phase quantization for blur insensitive texture analysis,” in Pattern Recognition, 2008. ICPR 2008. 19th International Conference on. IEEE, 2008, pp. 1–4.

7. N.-S. Vu and A. Caplier, “Face recognition with patterns of oriented edge magnitudes,” in European conference on computer vision. Springer, 2010, pp. 313–326.

8. A. Pflug, P. N. Paul, and C. Busch, “A comparative study on texture and surface descriptors for ear biometrics,” in Security Technology (ICCST), 2014 International Carnahan Conference on. IEEE, 2014, pp. 1–6.

9. N. Damer and B. Führer, “Ear recognition using multi-scale histogram of oriented gradients,” in Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2012 Eighth International Conference on. IEEE, 2012, pp. 21–24.

10. Križaj, V. Štruc, and N. Pavešić, “Adaptation of sift features for robust face recognition,” in International Conference Image Analysis and Recognition. Springer, 2010, pp. 394–404.