Subdivided Ear Recognition

ROSUS, Maribor

2019

David Romero, Matej Vitek, Blaž Meden, Peter Peer, Žiga Emeršič
{matej.vitek, blaz.meden, peter.peer, ziga.emersic}@fri.uni-lj.si

University of Ljubljana
Faculty of Computer and Information Science
Computer Vision Laboratory

Introduction

In ear recognition we use ear images to recognize subjects.

However, there are some issues:

- occlusion,
- ear accessories,
- missing portions of ears,
- background noise.

To tackle this problem we need to analyze which parts of ears contribute to (affect) recognition the most. In this work we propose a possible solution using local feature extraction techniques.

Subdivided Ear Recognition

- We perform recognition using two different subdivision approaches: up-down and internal-external.
- Score level fusion distances from different parts of the image combined using plain weighted averaging:

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

where α goes from 0 to 1.

Feature Extractors Used

We have used the following feature extractors:

- Local Binary Patterns (LBP),
- Binarized Statistical Image Features (BSIF),
- Local Phase Quantization (LPQ),
- Patterns of Oriented Edge Magnitudes (POEM),
- Histogram of Gradients (HOG),
- Dense (Grid-wise, without keypoint detection) Scale Invariant Feature Transform (DSIFT).

Experiments – Intro

We used AWE dataset with 1000 2D images of 100 subjects where images were taken in unconstrained environments.

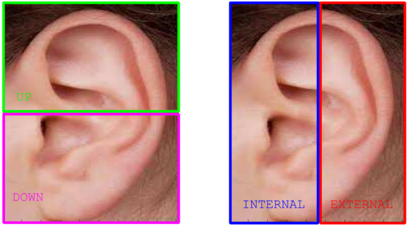

Figure: Ear subdivision.

For initial algorithm evaluation 600 images were used and for the final analysis the remainder of 400 images.

Up-Down Comparison

	AWE	*Up Deg.*	*Down Deg.*	*Gap*
LBP	43.6 ±7.2	7.3	8.9	+1.6
BSIF	48.5±6.9	5.6	11.7	+6.1
LPQ	43.3±7.7	4.1	10.9	+6.8
RILPQ	43.5±9.3	7.3	12.3	+5.0
POEM	49.0±6.9	8.9	12.6	+3.7
HOG	43.5±8.0	12.6	9.5	-3.1
DSIFT	43.6±8.5	9.5	9.7	+0.2

Figure: Comparison of the upper (Up) vs Lower (Down) comparison and performance degradation (Deg.).

External-Internal Comparison

	Left	*Deg. Int.*	*Deg.Ext.*	*Gap*
LBP	57.1±6.3	2.1	14.2	+12.1
BSIF	59.4±6.9	-3.3	15.0	+18.3
LPQ	58.2±6.0	4.2	19.6	+15.4
RILPQ	56.7±6.5	9.5	16.1	+6.6
POEM	62.7±4.5	15.0	19.2	+4.2
HOG	57.9±5.3	16.2	16.0	-0.2
DSIFT	56.9±5.2	5.2	12.5	+7.3

	Right	*Deg. Int.*	*Deg.Ext.*	*Gap*
LBP	52.4±9.9	4.1	13.9	+9.8
BSIF	57.7±6.3	2.5	19.0	+16.5
LPQ	52.9±5.2	4.1	18.7	+14.6
RILPQ	49.4±6.5	6.5	15.0	+8.5
POEM	58.9±7.1	14.7	19.5	+4.8
HOG	54.8±8.1	14.8	18.3	+3.5
DSIFT	51.2±7.5	3.9	21.0	+17.1

Figure: Comparison of internal (Int.) vs external (Ext.) comparison and performance degradation (Deg.).

Discussion & Conclusion

- The performance differences between the separate extractors are visible.
- **Inner-outer:** Internal parts seem to be more relevant than the outer parts – consistently for left and right ears.
  - HOG achieves almost the same performance for the internal part as the outer part, while BSIF and LPQ experience the highest performance drop when comparing internal to external parts of ear images.
- **Upper-lower:** The results are not conclusive, however, the upper part of the ear generally seems to have a higher influence than lower part on recognition.
  - HOG and DSIFT are the only descriptors where upper part is similar or worse than down part.

Discussion & Conclusion

- Fusing separate parts together in the score level does not improve overall performance.
- However, we have shown the differences that arise when using different parts of ears.
- For the future, analysis using the recently very popular CNN-based approaches will be interesting to see.