Ear Biometric Database in the Wild

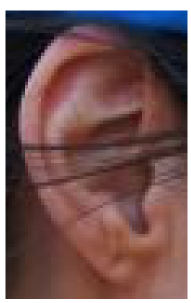
Žiga Emeršič, Peter Peer
Computer Vision Laboratory, Faculty of Computer and Information Science, University of Ljubljana
Večna pot 113, SI-1000 Ljubljana, Slovenia
Email: {ziga.emersic, peter.peer}@fri.uni-lj.si

***Abstract*—Ear biometrics is gaining on popularity in recent years. One of the major problems in the domain is that there are no widely used, ear databases in the wild available. This makes comparison of existing ear recognition methods demanding and progress in the domain slower. Images that were taken under supervised conditions and are then used to train classifiers in ear recognition methods can in effect cause these classifiers classifiers to fail under application in the wild. In this paper we propose a new database which consists of ear images in the wild of known persons taken from the Internet. This ensures different indoor and outdoor lightning conditions, different viewing angles, occlusions, and a variety of image sizes and quality. In experiments we demonstrate that our database is more challenging than others.**

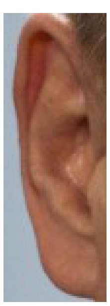
I. INTRODUCTION

Ear as a physiological characteristic presents unique [1] and stable data source for verification and identification methods. It is a non-invasive source of biometric features unlike palms, fingerprints or eyes, which require close user cooperation. Compared to widely used non-invasive data source – facial biometric features, it is insensitive to human facial expression and largely also to aging [1]. This represents a strong case for wider use of ear biometric data for person recognition.

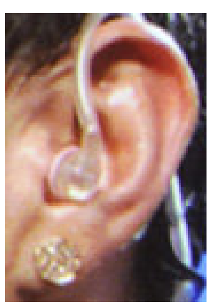
(a) No occlusion



(b) Occlusion by hair



(c) No occlusion but significant angle



(d) Occlusion by earphones and earrings

Fig. 1: Problems that occur when ear biometric data is captured in the wild.

When building any recognition system it is important to be able to impartially verify the performance of a system. To do that two aspects regarding data must be fulfilled. Data used must be as close to data in the wild as possible. This means that the built system needs to use as similar data as possible during evaluation and during final operation. The second aspect is that when comparing different recognition systems same impartial data source is used. Currently this is not the case as different authors use different data sources, which makes a comparison of results difficult. The reason for this is that there is no widely used, freely available, and annotated ear image database in the wild available. This slows down the progress in this domain because it is not always clear, which methods improve the overall performance and how much or if the data used for evaluation is too laboratory-like with ears in perfect alignment, photographs taken under the same or similar illumination using the same photographic equipment. Figure 1 shows possible difficulties that arise when taking ear images from cases in the wild. Only Figure 1a represents an ear without any disturbances, other three (1b, 1c, 1d) include hair, earrings, significant angle and earphones.

When dealing with ear biometric data there are three main steps to address [1]: ear detection, ear features detection and person recognition. For the person recognition system in the wild it is important to perform well in all of three steps – each step facilitates the next one. However, we put emphasis on the last two and dealt with the databases that already contain only ear images. In that case, ear detection does not represent a necessity (even though it might increase overall performance [1]).

CVL (Computer Vision Laboratory, Faculty of Computer and Information Science, University of Ljubljana) ear database is set to change this and simplify the evaluation process. In this paper we present our own (CVL) annotated ear database and use SIFT [2], [1], [3] and HOG [4] [1], [5] for feature detection and ear description. Further on, with the help of SVM [6] we execute the process of verification on our own CVL ear database and two existing databases WPUTEDB [7] and IIT Delhi ear database [8].

In Section II we discuss existing databases and their main properties. Section III presents CVL ear database, its main features and main steps made during database preparation. In Section IV the experiments that were undertaken to compare the new CVL ear database with two existing ones are described. The paper concludes in Section V with plans and suggestions for future work.

978-1-4673-7846-8/15/$31.00 ©2015 IEEE

II. Existing databases

In order to evaluate CVL ear database properly, existing publicly available databases need to be overviewed. Some databases require license, which can be obtained freely for research purposes. We tested WPUTEDB [7] and IIT Delhi ear database [8] and compared them with CVL ear database. These two databases represent good examples due to their properties: while WPUTEDB contains color images with occlusions, varying angles and both left and right ears, the IIT Delhi ear database contains only grayscale images of left ears with no major occlusions present.

WPUTEDB
: The database of The West Pommeranian University of Technology contains 3348 images of 421 subjects [7]. There are around 4 to 10 images per subject. Occlusions are present including earrings and hair. Images were taken under different indoor lightning conditions and angles ranging from approximately 70° to 120°. Images contain left and right ears (annotated in file names – L and R, respectively), subject are of different age groups.

IIT Delhi ear database
: The IIT Delhi ear database contains 493 grayscale images of 125 subjects [8]. Number of images per subject ranges from 3 to 6. No major occlusions are present. Images were taken at different indoor lightning conditions and contain only subjects' right ears. All images were taken from the same profile angle. The database also contains normalized images (equal image dimensions, ears centered, tightly cropped and aligned with axes), but in our tests we used the default (non-normalized) images because this presents a bigger challenge and is closer to the images taken in the wild.

University of Notre Dame databases
: The ear databases of The University of Notre Dame consist of multiple separate databases. They contain 3480 3D and corresponding 2D profile ear images of 952 subjects, averaging 3 to 4 images per subject, as well as 2D-only dataset containing 464 images from 114 subjects [1], [9] with the average of 4 images per subject. Images were taken under different lightning conditions and angles, and contain subjects' left ears.

UBEAR dataset
: The UBEAR dataset contains 4429 images from 126 subjects taken from both left and right side [1] [10] with the average of 35 images per subject. Images were taken under varying lightning condition, varying angles and contain occlusions. This database is interesting because images were taken while subjects were moving.

While authors of some of the databases tried to simulate conditions in the wild, images were still taken under similar conditions using similar or the same tools. Images in existing databases are also of high quality with rather large resolutions (over 200×200 pixels in most cases). This is something that cannot be guaranteed in the wild conditions, where ear represents only a small part of a captured image and where pose and illumination vary significantly. Another property that existing databases share is that subjects were never photographed at different life periods – time differences span within days, weeks or mostly a year – not decades. In CVL ear database presented in this paper and described in Section III, these aspects were addressed to enable a thorough evaluation of current and future ear recognition methods.

III. CVL ear database

The Computer Vision Laboratory ear database consists of 804 images of 16 subjects. Images per subject range from 19 to 94. Subjects are well known persons with a lot of freely available images, where ears are clearly visible and sufficient image sizes are available. Images were taken under different lightning conditions, indoor and outdoor. Images in the database vary in size and quality. Majority of the images are in color (11 images in grayscale) and under 200×200 pixels in size, with the smallest image having dimensions of 18×27 pixels. All images are stored in a Portable Network Graphics format. Subjects were taken at different angles ranging from 0° to 90° and beyond, to approximately 150°; where 0° represents frontal image, 90° profile image and 150° image from behind at an approximate angle of 30°. Occlusions are present in the images, including earrings, hair, earphones etc. Another important aspect is that images of subject were taken not only at different times but also at subjects' different life periods – something that we were not able to find amongst existing databases. Although we do not have the exact data regarding subjects' ages at the time the images were taken, we estimate that the differences of some of the images are up to 30 years.

The process of building CVL ear database consists of two main steps, as shown in Figure 2: web crawl that includes image acquisition (described in Subsection III-A) and image preparation that includes filtering of images and annotation (described in Subsection III-B).

Annotation data is stored in a JSON format, which provides smaller overhead than XML, but if needed, can still be transformed into an arbitrary format. Images were manually cropped and annotated. Each annotation data consist of a corresponding tragus center point, image dimensions, ear direction and a file name. The center point is the location of the outer area of tragus, as shown in the Figures 3a and 3b.

Image direction attribute describes the direction in which the center of the ear is, and not necessarily whether the ear is from the left or from the right side of the head. The reason for this is that images uploaded on the Internet are often mirrored. During image acquisition we noticed that sometimes background behind subjects strongly indicated that the image is mirrored – various labels and captions were mirrored. Nevertheless, we still consider that ear direction could prove to be useful parameter in the future, as in majority of cases still indicates, which ear is in the image. Left and right ear differ and this could be an important factor for a successful person recognition [1] [10].

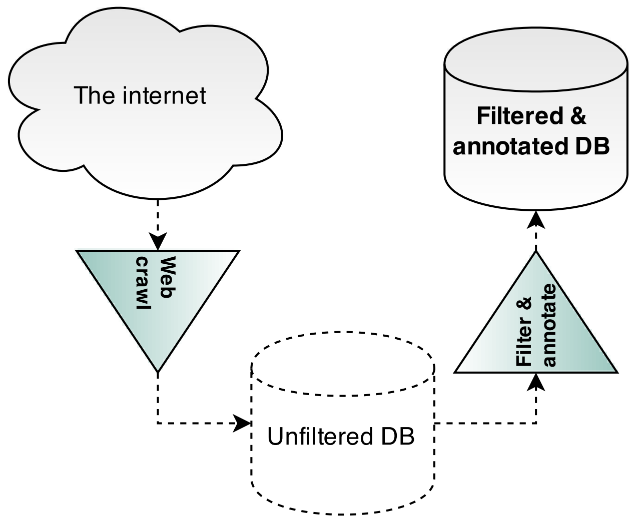


Fig. 2: Diagram showing CVL ear database creation process

(a) Right direction (b) Left direction

Fig. 3: Two annotated images from the database with different directions and annotated with center point. The point lies in the outer area of tragus horizontal-wise and in the center vertical-wise.

Tragus point could be proven useful when using concentric circles as holistic descriptors [1] or when distance between tragus and outline of the ear is used [11]. This also enables researchers to put emphasis on ear recognition performance and not ear (or ear-center) detection.

Of course, if a recognition system is built with the purpose of operating, let us say, a real-time identification system, annotation data would not be available. It would be on authors to calculate the tragus center point if needed and whether the ear is on the left or on the right side of the head. Another important aspect is ear detection itself. CVL ear database already contains cropped images with ears relatively centered, but in a described environment the authors would again be forced to calculate that if needed.

But as aforementioned our goal is to provide a background for a faster and better development of ear recognition techniques as such and not also ear detection.

CVL ear database is freely available per request.

A. *Image acquisition*

Images were acquired with the help of Google and Bing search engine APIs. We used two methods for image acquisition: plain image search and web search with additional web crawling. The second option was proposed in [12], promising better results and a possibility of acquiring more images since both Google and Bing search APIs provide a finite set of image results. Both methods were implemented using our scripts written in PHP scripting language.

During the plain image search method, Bing and Google APIs for image search are called. The returning result set contains a list of URLs of the images. Using our scripts, we then visit those URLs and store the images.

During the web search, Google and Bing APIs return a list of URLs of web sites where the data regarding the search term is located. The whole process consists of fetching URLs and then checking and storing content. When URLs are fetched we begin with the recursive procedure:

1) *Meta-information analysis and storing of images:* The meta-information in the header of the website is fetched and checked. If the meta-information suggests that the content is text based and contains HTML document we continue to the next step of web crawling, if the meta-information suggests the content is an image we store it and if it is anything else we stop the process.

2) *Web site analysis and URL fetching:* From Document Object Model of the website all URLs are parsed and filtered whether the initial part of the search query is present in the URL (or the corresponding description). When URLs are obtained, for each one the step one is called and the procedure recursively repeats until the wanted image number is reached.

The procedure taken by the authors in [12] includes fetching images during web crawling and then applying classification method such as SVM [6] for final filtering of the results. By using the web crawler we were able to acquire more images but also at a smaller initial accuracy. This recursive procedure resulted in a high amount of unusable images, where out of 5000, approximately 100 (2%) were usable, even though we did some prefiltering during the web crawl: only URLs that meet the criteria described in the previous paragraph were used.

Our goal was to acquire 500 initial images per person and then manually reject the obtained results and crop and annotate the rest of them. We found that for both APIs using this type of search did not present any problems. The aforementioned procedure was useful if the search query was ”ear”: the size of results after rejection was around 5.000 images. But the procedure did not bring any overall gain when the search query was ”x y side” (where x represents first and y last name of a person), therefore we did not fine tune SVM [6] to use it to filter final results of web crawling, but instead used image search results for the majority of the images.

For the purpose of searching for a specific person, image search was sufficiently successful. Also, executing a local filtering (with the help of SVM or other learning model) is harder when a category represents a person A or a person B etc. instead of a penguin or a tiger as was the case in [12].

B. Image preparation

Methods described in Section III-A provided us with approximately 8.000 images. Although images at this stage are already grouped by the persons or search terms (which in our case means "x y side", where x denotes a person's first name and y a person's last name), they still needed to be annotated and appropriately filtered. For this purpose a simple tool was developed that enabled quick cropping, annotation and rejection of unusable images. Note that both Google and Bing image search APIs yielded good results that contained only a few images with unrelated content – meaning they contained objects or other persons. The main reasons for image rejection were that ears in the images were too small, too occluded or not visible enough due to bad lightning conditions or inappropriate viewing angles.


```
"284": {
  "file": "284.png",
  "x": 20,
  "y": 42,
  "d": "l",
  "w": 61,
  "h": 74
}
```

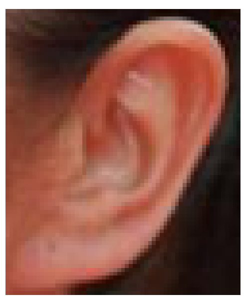



Fig. 4: An example image with the corresponding annotation data from the CVL ear database, where x and y represent coordinates of the tragus, d direction (l = left, r = right), and w and h image dimensions.

When building a database it is important to manually check and filter the content – even if we would use the best state-of-the-art classifiers and would optimally tune them, it would still be impossible to guarantee 100% accuracy without manually checking the results. This also applies to annotating ear areas, or in our case cropping images.

Next step is the image format conversion. The images that we acquired in previous steps are in different formats. For the final database this is not desired because it is hard to work with a database where each image is in an arbitrary format. We chose Portable Network Graphics format because it enables easier manipulation for the final users due to its lossless properties.

As the last step, tragus centers and ear directions were added to the annotation data.

In the resulting annotated database some images are deliberately repeated with only minor differences, i.e. original images were the same (in some cases with small image quality difference) but the ear was cut out under slightly different coordinates. It should also be noted that images are not normalized in any way (color, size etc.) – it is the job of a researchers to do so, if needed. An example image with the corresponding annotation data from the CVL ear database is shown in Figure 4.

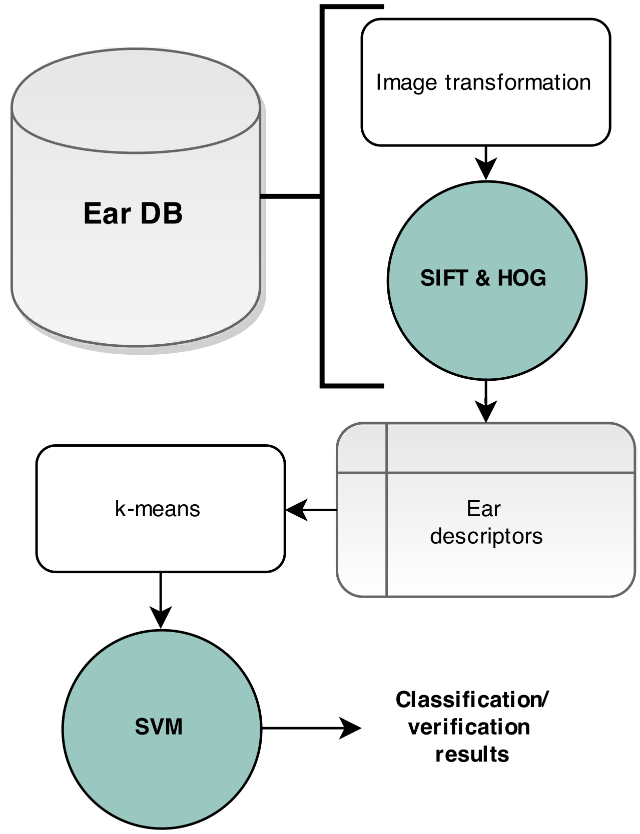


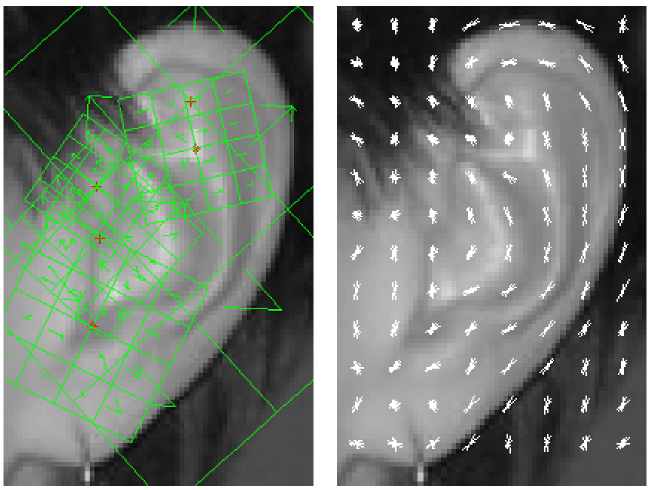
Fig. 5: Diagram showing the process of evaluation.

IV. Experiments

To compare and evaluate CVL ear database with the existing ones, we decided to extract ear features from the images, feed them to a classifier, perform verification and compare the results as shown in Figure 5. We did not focus on evaluating SIFT or HOG itself but rather on the difference in classification between different databases.

The performance when using CVL ear database is expected to be worse than existing databases, because we use images in the wild that are generally of lower overall quality and taken under non-laboratory like conditions. The deviation nevertheless should not be too big (even when using basic methods) because that would mean that the database is either too difficult to use or that images are of insufficient quality for use in ear biometrics. The experiments therefore verify the applicability of the database.

As the initial step, ear features need to be extracted. We decided to use two widely used methods for feature detection and description: Scale-Invariant Feature Transform (SIFT) [2] and Histograms of Oriented Gradients (HOG) [4]. Inputs to both methods were grayscale images. Examples of ear descriptors of both methods are displayed in Figure 6a and 6b. We also tried applying gradient to both x and y coordinates, with the goal of enhancing performance on blurred images or images where ear features are not as distinct. We did not achieve better performance, so these experiments are not shown in the final results.



(a) Scale-Invariant Feature Transform (b) Histogram of Oriented Gradients

Fig. 6: An image from the database with displayed descriptors after grayscale conversion was made.

A. Acquiring ear descriptors using SIFT

SIFT is a local method that is invariant to image scaling, rotation and partially to change in illumination and 3D transformation [3], [13]. The mentioned properties make SIFT a good method for handling ear biometric data and it has been used in ear detection and recognition [13], [14], [15]. It consists of four major steps: scale-space extrema detection, key point localization, orientation assignment, key point descriptor [3], and we added the fifth: k-means clustering for dimension reduction [16], [17]. The results of steps two and three can be seen in Figure 6a, where orientation histograms on the 4×4 regions are shown in green. The last step of k-means clustering is needed because we need to transform data so that it is of an appropriate dimensions for later classification using Support Vector Machine (SVM) [6] as described in Subsection IV-C.

In our experiments we used a basic implementation of SIFT and many things could be optimized. Process of classification could be enhanced using color SIFT as proposed in [13] or the input data could be normalized.

B. Acquiring ear descriptors using HOG

The HOG method is useful when illumination variations or shadowing are present [18]. It has been used in object and ear detection and recognition [19], [20], [21], [22].

The first step in the HOG calculation process was image size transformation to fixed 100×100 pixels. Since HOG returns fixed number of features for a given image size this enables us to feed the results directly into SVM [6] without applying k-means clustering (or other method) first. The experiments have shown that this did not significantly influence the performance.

Even though the results vectors were of the same size, we could still enhance performance by additionally reducing number of attributes using PCA or k-means clustering [16], [23], [17], [20]. Another viable improvement would be to use different size classes or the so-called Pyramid Histogram of Oriented Gradients as proposed in [21], [24], [25], [26].

C. Classification

As the final learning model Support Vector Machine [6] was used for the verification procedure. Because input vectors need to be of the same size, certain transformations were needed while calculating descriptors as mentioned in Subsections IV-A and IV-B. The training set that was used to learn SVM consisted of an ear descriptor for each ear and a true/false value whether the corresponding descriptor belongs to the class or not. After the learning process, SVM then predicted the most probable value for each new ear descriptor in the test set. The procedure was repeated for all the persons in the database.

During learning and testing processes we did not differentiate between left or right ears, because according to [9] 90% of people's right and left ears are symmetric. Performance was evaluated using repeated random sub-sampling validation: dataset was randomly divided into test and train groups with ratios of 3 to 7 and repeated three times with the final results being an average over all runs.

D. Results

Results are presented in Table I for HOG descriptors and Table II for SIFT descriptors. Performance is the overall ratio of correctly classified subject vs. all subjects. Specificity is defined as $\frac{N_T}{N}$, where N_T represents true negatives (correctly classified as negatives) and N all negatives. Sensitivity is defined as $\frac{P_T}{P}$, where P_T represents true positives and P all positives. It is important to notice that in verification process we want specificity to be as high as possible, while sensitivity is not that much of an issue. However, emphasis on the sensitivity should be put when the false rejection rate is undesired as well. The experiments show that classification on CVL ear database performs slightly worse that other two (97.53% vs. 99.92% and 99.85% using HOG descriptors and 87.29% vs. 99.85% and 99.41% using SIFT descriptors), which is as expected because we used in the wild data. At the same time the performance is not much worse – if it would be, it would mean that the database is not applicable when feature extraction methods like SIFT or HOG are used. Thus, CVL ear database provides important and challenging source of data for ear biometric experiments.

V. Conclusion

In this paper we presented first publicly available, ear database in the wild, with the in the wild lightning, occlusions, poses and distortions. The experiments have shown that although images were acquired from the Internet, meaning the photographs were taken under unsupervised conditions,

TABLE I: Results using HOG descriptors and SVM classifier

	CVL ear database	WPUTEDB	IIT Delhi ear database
Performance [%]	97.53	99.92	99.85
Specificity [%]	99.59	100	100
Sensitivity [%]	58.77	59.98	51.06

TABLE II: Results using SIFT descriptors and SVM classifier

	CVL ear database	WPUTEDB	IIT Delhi ear database
Performance [%]	87.29	99.85	99.41
Specificity [%]	93.59	99.94	99.93
Sensitivity [%]	19.55	57.39	37.75

the database is comparable with the existing ones, while still offering additional challenge for ear recognition methods. We hope that this database will enable easier comparison of different ear biometric verification and identification methods in the research community.

Our plan is to upgrade current database. A version 2 of the database with much more subjects (hundreds), where each of them is represented with around 10 images is under construction to present even bigger challenge to the community. We are also developing a publicly available toolkit for ear biometric recognition methods with integrated classification and comparison methods. This should even further improve the evaluation and comparison of ear biometric recognition methods.

References

[1] A. Pflug and C. Busch, "Ear biometrics: a survey of detection, feature extraction and recognition methods," *Biometrics, IET*, vol. 1, no. 2, pp. 114–129, June 2012.

[2] D. Lowe, "Object recognition from local scale-invariant features," in *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, vol. 2, 1999, pp. 1150–1157 vol.2.

[3] ——, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004. [Online]. Available: http://dx.doi.org/10.1023/B%3AVISI.0000029664.99615.94

[4] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1, June 2005, pp. 886–893 vol. 1.

[5] ——, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1, June 2005, pp. 886–893 vol. 1.

[6] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995. [Online]. Available: http://dx.doi.org/10.1023/A:1022627411411

[7] D. Frejlichowski and N. Tyszkiewicz, "The west pomeranian university of technology ear database – a tool for testing biometric algorithms," in *Image Analysis and Recognition*, ser. Lecture Notes in Computer Science, A. Campilho and M. Kamel, Eds. Springer Berlin Heidelberg, 2010, vol. 6112, pp. 227–234. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-13775-4_23

[8] A. Kumar and C. Wu, "Automated human identification using ear imaging," *Pattern Recogn.*, vol. 45, no. 3, pp. 956–968, Mar. 2012. [Online]. Available: http://dx.doi.org/10.1016/j.patcog.2011.06.005

[9] P. Yan and K. Bowyer, "Empirical evaluation of advanced ear biometrics," in *Computer Vision and Pattern Recognition - Workshops, 2005. CVPR Workshops. IEEE Computer Society Conference on*, June 2005, pp. 41–41.

[10] R. Raposo, E. Hoyle, A. Peixinho, and H. Proenca, "Ubear: A dataset of ear images captured on-the-move in uncontrolled conditions," in *Computational Intelligence in Biometrics and Identity Management (CIBIM), 2011 IEEE Workshop on*, April 2011, pp. 84–90.

[11] L. Lu, Z. Xiaoxun, Z. Youdong, and J. Yunde, "Ear recognition based on statistical shape model," in *Innovative Computing, Information and Control, 2006. ICICIC '06. First International Conference on*, vol. 3, Aug 2006, pp. 353–356.

[12] F. Schroff, A. Criminisi, and A. Zisserman, "Harvesting image databases from the web," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 4, pp. 754–766, April 2011.

[13] J. Zhou, S. Cadavid, and M. Abdel-Mottaleb, "Exploiting color sift features for 2d ear recognition," in *Image Processing (ICIP), 2011 18th IEEE International Conference on*, Sept 2011, pp. 553–556.

[14] D. Kisku, H. Mehrotra, P. Gupta, and J. Sing, "Sift-based ear recognition by fusion of detected keypoints from color similarity slice regions," in *Advances in Computational Tools for Engineering Applications, 2009. ACTEA '09. International Conference on*, July 2009, pp. 380–385.

[15] J. Bustard and M. Nixon, "Robust 2d ear registration and recognition based on sift point matching," in *Biometrics: Theory, Applications and Systems, 2008. BTAS 2008. 2nd IEEE International Conference on*, Sept 2008, pp. 1–6.

[16] A. Likas, N. Vlassis, and J. J. Verbeek, "The global k-means clustering algorithm," *Pattern Recognition*, vol. 36, no. 2, pp. 451 – 461, 2003, biometrics. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0031320302000602

[17] T. Santhanam and M. Padmavathi, "Comparison of k-means clustering and statistical outliers in reducing medical datasets," in *Science Engineering and Management Research (ICSEMR), 2014 International Conference on*, Nov 2014, pp. 1–6.

[18] N. Damer and B. Fuhrer, "Ear recognition using multi-scale histogram of oriented gradients," in *Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2012 Eighth International Conference on*, July 2012, pp. 21–24.

[19] H. Ameur, A. Helali, M. Nasri, H. Maaref, and A. Youssef, "Improved feature extraction method based on histogram of oriented gradients for pedestrian detection," in *Computer Information Technology (GSCIT), 2014 Global Summit on*, June 2014, pp. 1–5.

[20] N. Bardeh and M. Palhang, "New approach for human detection in images using histograms of oriented gradients," in *Electrical Engineering (ICEE), 2013 21st Iranian Conference on*, May 2013, pp. 1–4.

[21] T. Anakavej, A. Kawewong, and K. Patanukhom, "Internet-vision based vehicle model query system using eigenfaces and pyramid of histogram of oriented gradients," in *Signal-Image Technology Internet-Based Systems (SITIS), 2013 International Conference on*, Dec 2013, pp. 179–186.

[22] H.-C. Tsai, W.-K. Fan, B.-W. Chen, J.-F. Wang, and P.-C. Lin, "A real-time awareness system for happiness expression based on the multilayer histogram of oriented gradients," in *Awareness Science and Technology (iCAST), 2012 4th International Conference on*, Aug 2012, pp. 289–293.

[23] C. Ding and X. He, "K-means clustering via principal component analysis," in *Proceedings of the Twenty-first International Conference on Machine Learning*, ser. ICML '04. New York, NY, USA: ACM, 2004, pp. 29–. [Online]. Available: http://doi.acm.org/10.1145/1015330.1015408

[24] Z. Li, J. Imai, and M. Kaneko, "Facial-component-based bag of words and phog descriptor for facial expression recognition," in *Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on*, Oct 2009, pp. 1353–1358.

[25] F. Lu, X. Yang, R. Zhang, and S. Yu, "Image classification based on pyramid histogram of topics," in *Multimedia and Expo, 2009. ICME 2009. IEEE International Conference on*, June 2009, pp. 398–401.

[26] D. Abdullah, I. Murtza, and A. Khan, "Feature extraction and reduction strategy based on pyramid hog and hierarchal exploitation of cortex-like mechanisms," in *Multi Topic Conference (INMIC), 2013 16th International*, Dec 2013, pp. 160–165.