Automatic human ear detection approach using modified adaptive search window technique

Raad Ahmed Hadi¹, Loay Edwar George², Zainab Jawad Ahmed³

¹Department of Computer Science, College of Education, Al-Iraqia University, Baghdad, Iraq ²Department of Remote Sensing and Geographic Information Systems, University of Baghdad, Baghdad, Iraq ³Department of Biology Science, College of Science, University of Baghdad, Baghdad, Iraq

Article Info

Article history:

ABSTRACT

Received Jul 31, 2020 Revised Oct 1, 2020 Accepted Oct 21, 2020

Keywords:

Adaptive search window Automatic ear detection Biometric Computational time Ear recognition

The human ear biometric recognition plays an important role in the forensics specialty and has significant impact for biometrician scientists and researchers. Actually, many ear recognition researches showed promised results, but some issues such as manual detection process, efficiency and robustness aren't attained a certain level of maturity. Therefore, the enhancement developing approaches still continuous to achieve limited successes. We propose an efficient, reliable and simple automatic human ear detection approach. This approach implement two stages: preprocessing and ear landmarks detection. We utilized the image contrast, Laplace filter and Gaussian blurring techniques to made enhancement on all images (increasing the contrast, reduce the noisy and smoothing processes). After that, we highlighted the ear edges by using the Sobel edge detector and determining the only white pixels of ear edges by applying the image substation method. The improvement focused on using the modified adaptive search window (ASW) to detect the ear region. Furthermore, our approach is tested on Indian Institute of Technology (IIT) Delhi standard ear biometric public dataset. Experimental results presented a well average detection rate 96% for 493 image samples from 125 persons and computational time almost ≈ 0.485 seconds which is evaluated with other previous works.

This is an open access article under the <u>CC BY-SA</u> license.



Corresponding Author:

Raad Ahmed Hadi Department of Computer Science, College of Education Al-Iraqia University Adamiyah, Baghdad, Iraq Email: raad.hd@gmail.com

1. INTRODUCTION

Ear biometrics authentication is a form of security which has an important role in the forensics field to define the unique physical characteristics for distinguishing the people's identity, for identification or verification task [1-4]. Nowadays, a majority of the well-known ear biometric approaches focused on recognizing of cropped ear manually without using an automatic ear detection and segmentation techniques. So, the task automated ear localization represents a big challenge for many researchers whose main concern is in human ear biostatistics studies. In addition, issues such as efficiency and robustness are still need for improvements by biometrician scientists and researchers and need to make it in harmony with each other [5-7]. In fact, few introduced techniques that can be used to distinguish the ear landmark automatically. Murukesh *et al.* [2] suggested a new ear recognition approach that utilized contourlet transform and appearance shape model (ASM) techniques for feature extraction, followed by classification and ear matching processes

507

using fisher linear discriminant analysis (FLDA). The proposed method tested on Indian Institute of Technology (IIT) Delhi ear database by using of 50 ear images from 10 persons and own ear databases. This approach indicated well overall performance with accuracy as high as 97%. In fact, the time evaluation not mentioned in this approach.

Ghoualmi *et al.* presented a new ear biometrics system based on using an artificial bee colony (ABC) and scale invariant feature transform (SIFT) techniques. De facto, histogram equalization (HE) and contrast limited adaptive histogram equalization (CLAHE) were tested in this approach on three standard ear image databases (IIT Delhi, USTB 1 and USTB 2). Moreover, it showed well results with an average accuracy of 97.15% [4]. Also, a new, automatic and efficient ear recognition approach recommended by Hadi and George. This approach is based on using the color skin, sobel edge detector, image subtraction and region-growing techniques. This approach is tested on the IIT Delhi ear image dataset, and presented a well performance with average detection rate almost 91.8% and computational time around ≈ 1.33 seconds in terms of efficiency [5].

Interestingly, using ensemble of convolutional neural network (CNN) techniques for an ear localizing system is proposed by Ganapathi *et al.* [7]. In this system, three models of CNN trained the given dataset. In fact, these models proved better performance in case used together. The proposed ear localizing system is tested on two databases, IIT indore-collection A (IIT-Col A) database and annotated web ear (AWE) database with the existence of pose variations, occlusion and illumination conditions matters which takes on average 2.1 seconds. In addition, an innovative ear recognition algorithm based on using extraction of geometrical features such as (shape, mean, centroid and Euclidean distance between pixels) suggested by Anwar *et al.* [8]. Although the experimental results showed that the proposed approach gives well outcomes and achieved average accuracy around 98%, it is computationally complex. Furthermore, it requests for manual initialization for successful execution of detection process.

Additionally, an efficient ear recognition technique based on neural networks (NN) is demonstrated by Zhang and Mu [9]. The authors utilized multiple scale faster region-based convolutional neural networks (faster R-CNN) as a tool to detect the 2D ear region from the profile image automatically. This proposed technique is tested on a set of 200 web images under variant photographic conditions, and it is achieved 98% detection rate. Likewise, an automated human ear identification system proposed by Tariq and Akram. This system encompassed from three stages: preprocessing, features extraction and identification processes respectively. The experimental results illustrate an average accuracy of 97.2% and 95.2% that are evaluated on the USTB and IIT Delhi ear image databases respectively [10].

Benzaoui *et al.* proposed an ear description and recognition that used a robust elliptical local binary pattern (ELBP) and discrete wavelet transform (DWT) techniques to depict the adequate details of the two-dimensional ear images [11]. However, the evaluation results showed a success recognition rate around 94% when tested on 500 images from 100 persons from the IIT Delhi database. Besides, an efficient online ear-based personal identification system presented by Meraoumia *et al.* [12]. In this paper, each ear had specific features set which is extracted by using Gabor filter. In order to realize an ideal multi-representation system, the fusion phase is applied by trying of several combinations of using these features (phase, module and real (imaginary) parts mixtures). This system tested on IIT Delhi database of 221 users and yields a well performance of ear identification process.

Another ear detection model suggested by [13]. The authors used two techniques as follows: snake-based background removal (SBR) and snake-based ear localization (SEL). Conversely, this model shows well result but it suffers from the high computational time; it is around 3.86 s per image. In [14], the authors proposed approach for ear localization based on color (YCbCr color space) detection and edge mapping techniques. Indeed, the ear detection average time is around 7.95s which is consumed more resources from computational complexity aspect. Alternative work was suggested by Hourali and Gharravi [15]. They used a modified form of discrete cosine transform (transformed DCT). It is tested on two datasets USTB subset II and IIT Delhi subset II and evaluated with good efficiency. Nevertheless, this work assumed that the ear region is cropped manually by specified ear detector advancely. So, it isn't an automated ear localization system.

The main idea of our work is to present an efficient, reliable and simple automatic human ear detection approach which is based on modified ASW. The proposed approach which is recapitulated by two phases preprocessing and ear landmarks detection. Additionally, experiminal tests show promised results of the proposed approach. This paper is structured as follows: The proposed approach of ear biometric detection is presented in section 2. During the section 3, the results and dissection is illustrated. Finally, conclusion and future work suggestions are summarized in section 4.

2. THE PROPOSED APPROACH OF EAR BIOMETRIC DETECTION

Here, a proposed approach of ear biometric detection is illustrated that is extracted efficiently and simplify in implementation the ear's landmarks information as shown in Figure 1. More specifically, this

approach implements two stages as follow: firstly, we made a preprocessing for image enhancement; three operations are used for increasing the contrast (stretching), reduce or blur the noisy (Gaussian blur) and smoothing (laplace filter) of all ear images. Secondly, we applied a Sobel edge detector technique to highlight the ear landmark edges and yield a binary mask image. After that, we used the image subtraction process to determine the only white pixels of ear edges. In fact, we subtracted the resultant image of Sobel edge detection technique which is inspired from [16, 17] applied to detect the ear landmarks region. This new technique used four detectors to scan, collect white pixels of ear edges and recognized it. During the testing process, we applied our approach on 493 sample belong to IIT Delhi standard ear biometric public dataset. Besides, two performance measures are evaluated the proposed approach as indicator of its efficiency as follow the accuracy analysis and computational time respectively. The average accuracy of the present work shows promised result but little bit less from some previous works, but in contrary the efficiency of real-time depicts faster result in ear detection through computing each process per image.



Figure 1. Layout of proposed ear detection approach

2.1. Preprocessing

During the preprocessing, the whole images read one by one and convert to grayscale images; the image contrast stretching is applied to adjust the intensities contrast of images for attaining more sharpening of the ear edges. After that, gaussian blur and laplace filter are applied to enhance the quality of ear images (increase the disparity of ear landmarks and remove the noise or blur of image). Here, one of the common image illumination enhancement techniques is contrast stretching which is used. It works by spreading the gray-levels (brightness values) of the handled image into dynamic range. Then, the resultant image will be giving more information for analysis process [18]. Actually, calculating the pixel illumination values are illustrated by computing of the power interval that is selected the minimum and maximum values and stretched main power interval of the histogram to the full range (0-255) as shown in (2) [19]:

$$Gs(x,y) = \begin{cases} 255 \quad if G(x,y) \ge Max \\ 0 \quad if G(x,y) \le Min \\ 255 * \left(\frac{G(x,y) - Min}{Max - Min}\right) \quad Otherwise \end{cases}$$
(1)

where G(x, y) is the image coordination, max and min calculated as follows:

$$\begin{aligned} Min &= \mu - \alpha\sigma \\ Max &= \mu + \alpha\sigma \end{aligned} \tag{2}$$

where μ and σ symbols indicated the mean and standard deviation values of the image, respectively. The parameter α is applied to control the strength of implemented linear extent.

After that, the Gaussian blurring operation means simply to transform one color value to the other very smooth. Also, it considers one of common essential operations before many tasks such as edge detection in image processing specialty. The Gaussian smoothing averaging operator or filter represents a 2D conventional operator that takes the shape of a Gaussian (bell-shaped) hunch, it will remove the image's noise with the high spatial frequencies and yields a smoothing outcome. In two dimensions, an isotropic (i.e., circularly symmetric) Gaussian blur filter function has the form [20, 21]:

$$G(x,y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x^2 + y^2)/2\sigma^2}$$
(3)

where σ symbol is the standard deviation of the 2D distribution and it controls the degree of image smoothing.

Also, the Laplace filter will be used in this proposed approach. It is used as a measure of the second spatial derivative of an image which is a 2D identical (isotropic) degree. However, it is a conventional operator that highlights the changing of rapid intensity areas. So, it is often used after smoothing approximating process of an image by Gaussian blur filter to reduce sensitivity to noise effect [22, 23]. The following equation depicts of an image which is denoted by Laplacian L(x,y) with pixel intensity values I(x,y) is assumed by:

$$L(x,y) = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2}$$
(4)

Additionally, we compare every pixel of the resultant image (Lap_image(x)) from Laplace filtration process with a range from (70>=Lap_image(x)<=200) to get more white pixels, and highlighting the ear edges (binary mask image) for next step (Sobel edge detection process).

2.2. Ear landmarks detection

The main idea and contribution of the proposed approach is to utilize new technique for recognition the ear landmarks which is adapted by [16, 17]. After the preprocessing phase, we have got an enhanced image with sharped edge and noiseless of ear landmarks. Firstly, a Sobel edge detector applied on the enhanced image to discriminate the human ear edges. An edge is the boundary between regions of two images, which has the distinct properties according to some features such as (gradient, color, texture or gray level). It implies a pair of orthogonal gradient operator, it works to find the edge strength and direction at location (x, y) of an image f, it is called the gradient, denoted by ∇f , and is defined as a vector [5].

$$\nabla f(x,y) = grad(f) = [G_x G_y] = \left(\frac{\partial f}{\partial x} \frac{\partial f}{\partial y}\right)$$
(5)

This vector of continuous function f(x, y) has important geometrical properties of location (x, y). The value of changing rate in the direction of gradient vector is called the magnitude (length) of the vector ∇f , and is denoted as M (x, y), where:

$$M(x,y) = mag(\nabla f) = \sqrt{G_x^2 + G_y^2}$$
(6)

In addition, the direction angle of gradient vector is declared as:

$$\theta(x,y) = \tan^{-1} \left[\frac{G_y}{G_x} \right] \tag{7}$$

Sobel edge detector is worked by convolution concept; it uses a set of 3x3 convolution kernels. One of the (Gx) kernels set is used to distinguish the brightness strength of edges in the horizontal direction, and the other one (Gy) is used to distinguish the brightness strength of edges in the vertical direction [5, 23-26]. Secondly, an arithmetic operation is a process performed between pixel-to-pixel is called image subtraction. It is often used to detect the differences between two images (e.g. removing of relevant contents from the image or detect the relevant object motion between two frames of a video sequence). Therefore, the resultant image from previous process will be subtracted from the enhanced image to get the final binary mask image with only the same edges (white pixels) between the two-image mentioned before. Given a 2D array of (X) image and another 2D array of (Y) image, the resulting is a new array, scalar (Z), is obtained by calculating [5, 24]:

$$X - Y = Z \tag{8}$$

Finally, a modified technique that is kind of similar to [16, 17], it will be used for efficiently and simplicity distinguishing the ear landmark from the face side region. However, it is a variable-size search window that is used to detect the ear landmarks region by adjusting the size scales (height and width). Actually, it is a flexible search window by using four directions (top, bottom, left and right detectors) for locating and detection the ear landmarks height and wight dimensions. Practically, this technique is applied on the resultant low-level binary image (pixel colour values with white and black) from image subtraction process, it's scanned the image from four directions to find the white pixels by using flexible detectors as shown in Figure 2. These detectors initialized by using thresholds as training values to scan the image pixels and reduce the search window process with accumulators to count white pixels for ear landmarks localization from the face side region. Finally, these accumulators will be selected based on the higher one of white pixels counting values, and it will determine the corrdinates of the detected ear region.



Figure 2. Adaptive search window technique for ear detection

3. RESULTS AND DISCUSSION: EVALUATION

The performance of proposed approach had been tested on the Indian Institute of Technology (IIT) Delhi ear image dataset and evaluated in terms of processing time and detection accuracy. The IIT dataset consists of 493 RGB ear image, which was taken from 125 different (subjects) persons that are of ages between 18 and 58 years. In addition, these images have a resolution of 272x204 pixels for each. The proposed approach is carried out using Microsoft visual C# 2017 software on a pentium IV Core i5 (1.60 GHz) laptop. Figure 3 present some samples from the dataset images.



Figure 3. IIT Delhi ear image samples

3.1. Detection accuracy

The performance of ROI (Region of Interest) or ear detection can be measured [3] as follows:

$$DetectionAccuracy = \frac{Number of Correct Detections \times 100}{Total Test Samples}\%$$
(9)

However, the ear detection accuracy of the proposed approach showed a well result almost (96.5%) compared to other previous works as shown in Table 1. Additionally, we can notice in the table mentioned before that the most of researchers tested their works on own datasets, which are mostly not standard dataset. Moreover, small numbers of samples of their own datasets and IIT Delhi datasets are used for testing and evaluation their ear recognition systems.

Publication	Approach	Name of Dataset	Ear image samples	Accuracy
Murukesh.C et al. [2]	Contoulet and PCA	IIT Delhi	50	96%
Hadi and George [5]	Color skin, edge detection and image subtraction	IIT Delhi	493	91%
Anwar, A. S. et al. [8]	Geometrical features	IIT Delhi	450	98%
Tariq and Akram [10]	Haar wavelets and normalized cross correlation (NCC)	IIT Delhi	125 subject	95%
Jitendra, B. [27]	Geometrical feature	Their Own	30	00%
		Dataset	50	9070
Alaraj <i>et al</i> . [28]	Principal components analysis (PCA) and	Their Own	85	96%
	MLFFNNs	Dataset	05	2070
The proposed	Gaussian, Laplace, edge detection, image			
approach	subtraction and modified adaptive search window	IIT Delhi	493	96%
	(ASW)			

T 1 1 T 1 1	1 1 1 1	11	•
Table I Har and oth	er nhugiologica	I human traite ee	mnaricor
raule r. Lai and ou		i numan itans o	JIIIDai 1501

3.2. Processing time

The processing time of each procedure for the proposed approach is measured in second which is utilized to evaluate the computational time cost of the proposed approach. The average time for every procedure is computed via six procedures/operations such as stretching, Gaussian, Laplace, Sobel edge detection, image subtraction and ear detection. In this regard, our proposed approach achieved an efficient computational time (≈ 0.485 seconds). Figure 4 shows the calculation of the procedures of the proposed approach that are used together to detect ear landmarks. Table 2 summarizes the previous woks results compared with the efficient acquired computational time result from the proposed approach.

In spite of our results give little bit less accuracy of detection relation to some other researchers, but it is more efficient in terms of computational time as shown in Table 2. On other hand, the approach had some misdetection results which is shown in Figure 5 for more clarification. Actually, we can notice that the problem of misdetection is the disjoint pixels of ear edges. In fact, the density of white pixels is affected on the work of ear detection process using ASW technique to segment the ROI object correctly, because the detectors depended on counting these white pixels of ear edge and if there are many gaps (black pixels) in shape of ear's edge; the result of detection accuracy rate will be decreased. In this regard, we will work on make some enhancements on the approach to yield more accurate detection accuracy rate in harmony with efficient computational time in the future.



Figure 4. Ear detection approach's computational time result



Figure 5. Disjoint pixels of ear region (some samples)

Table 2. Comparison between the proposed approach and other previous works in terms of computational time

in terms of computational time					
Publication	Approach	Computational time (seconds)			
Hadi and George [5]	Color skin, edge detection and image subtraction	1.33			
Tariq and Akram [10]	Haar wavelets and Normalized Cross Correlation (NCC)	0.60			
The proposed approach	Gaussian, Laplace, edge detection, image subtraction and				
	modified Adaptive search window (ASW)	0.48			

4. CONCLUSION

An efficient, reliable and simple approach has successfully proposed for automatic human ear detection; this approach is based on using modified ASW. It is consisted of two phases: preprocessing and ear landmarks detection. Firstly, three operations (image contrast stretching, Gaussian blur, and Laplace filter) of image enhancement are utilized to make improvement on all images (increasing the contrast, reduce the noisy and smoothing processes). Secondly, two operations (Sobel edge detector and image subtraction) are used highlighted and determined the only white pixels of the ear edges respectively. In addition, a modified method which is adopted from ASW is used to detect the ear landmarks region. Likewise, our approach is tested on IIT Delhi standard ear biometric public dataset. Experimental results presented a well average detection rate 96% for 493 image samples from 125 persons and computational time almost ≈ 0.485 seconds which is evaluated with other previous works. In the future, we will expand our test in many ear databases with different scenarios such as pose variation, occlusion, scale and illumination changes matters.

ACKNOWLEDGEMENTS

The authors are grateful and appreciative to the Ministry of Higher Education and Scientific Research (MOHESE), Iraq, for supporting this research grant.

REFERENCES

- [1] Vélez J. F., S. Á., Moreno B. and Sural S., "Robust Ear Detection for Biometric Verification," *IADIS International Journal on Computer Science and Information Systems*, vol. 8, no. 1, pp. 31-46, 2013.
- [2] Murukesh C., A. Parivazhagan, and K. Thanushkodi, "A Novel Ear Recognition Process Using Appearance Shape Model, Fisher Linear Discriminant Analysis and Contourlet Transform," *Proceedia Engineering*, vol. 38, pp. 771-778, 2012.
- [3] Prakash S., and Gupta P., "Ear Biometrics in 2D and 3D," *Augmented Vision and Reality*, ed. L.B.W. Riad I. Hammoud, vol. 10, 2015.
- [4] Ghoualmi L., A. Draa, and S. Chikhi, "An ear biometric system based on artificial bees and the scale invariant feature transform," *Expert Systems with Applications*, vol. 57, pp. 49-61, 2016.
- [5] Hadi, R.A. and L.E. George, "An Automated Ear Detection Method based on Sobel Edge Detection and Image Subtraction Techniques," 2018 Thirteenth International Conference on Digital Information Management (ICDIM 2018), 2018.
- [6] Cintas C., *et al.*, "Automatic Ear Detection and Segmentation over Partially Occluded Profile Face Images," *Journal of Computer Science and Technology*, vol. 19, no. 01, pp.81-90, 2019.

- [7] Ganapathi I. I., et al., "Unconstrained ear detection using ensemble-based convolutional neural network model," Concurrency and Computation: Practice and Experience, vol. 32, no. 1, 2020.
- [8] Anwar A. S., K. K. A. Ghany, and H. Elmahdy, "Human Ear Recognition Using Geometrical Features Extraction," *Procedia Computer Science*, vol. 65, pp. 529-537, 2015.
- [9] Zhang Y., and Z. Mu, "Ear Detection under Uncontrolled Conditions with Multiple Scale Faster Region-Based Convolutional Neural Networks," *Symmetry*, vol. 9, no. 4, 2017.
- [10] Tariq, A. and Akram M.U., "Personal identification using computerized human ear recognition system," *TELKOMNIKA Telecommunication Computing Electronics and Control*, vol. 10, no. 2, pp. 321-326, 2012.
- [11] Benzaoui A., A. Kheider, and A. Boukrouche, "Ear description and recognition using ELBP and wavelets," 2015 International Conference on Applied Research in Computer Science and Engineering (ICAR), 2015.
- [12] Meraoumia A., S. Chitroub, and A. Bouridane, "An automated ear identification system using Gabor filter responses," 2015 IEEE 13th International New Circuits and Systems Conference (NEWCAS), 2015.
- [13] Deepak R., A.V. Nayak, and K. Manikantan, "Ear detection using active contour model," 2016 International Conference on Emerging Trends in Engineering, Technology and Science (ICETETS), 2016.
- [14] Prakash S., and P. Gupta, "An efficient ear localization technique," *Image and Vision Computing*, vol. 30, no. 1, pp. 38-50, 2012.
- [15] Hourali F. and Gharravi S., "An Ear Recognition Method Based on Rotation Invariant Transformed DCT," International Journal of Electrical and Computer Engineering (IJECE), vol. 7, no. 5, pp. 2895-2901, 2017.
- [16] Hadi R. A., "Moving Vehicles Detection and Tracking Approach Via Modified Background Subtraction and Innovative Adaptive Serach Window," PhD thesis, Faculty of Computing Universiti Teknologi Malaysia: Malaysia, pp. 215, 2015.
- [17] Hadi R. A., L. E. George, and M. J. Mohammed, "A Computationally Economic Novel Approach for Real-Time Moving Multi-Vehicle Detection and Tracking toward Efficient Traffic Surveillance," *Arabian Journal for Science and Engineering*, vol. 42, no. 2, pp. 817-831, 2017.
- [18] Acharya T. and A. K. Ray, "Image Processing: Principles and Applications," Hoboken, NJ: John Wiley and Sons, 2005.
- [19] Abdul-Nasir A. S., M. Y. Mashor, and Z. Mohamed, "Modified global and modified linear contrast stretching algorithms: new colour contrast enhancement techniques for microscopic analysis of malaria slide images," *Hindawi Publishing Corporation Computational and Mathematical Methods in Medicine*, vol. 2012, pp. 1-16, 2012.
- [20] Nixon M. and A. S. Aguado, "Feature Extraction & Image Processing for Computer Vision," Third Edition, *Academic Press, Inc,* 2012.
- [21] Shih F. Y., "Image Processing and Pattern Recognition: Fundamentals and Techniques," Wiley, 2010.
- [22] Haralick R. M. and L. G. Shapiro, "Computer and Robot Vision," *Addison-Wesley Longman Publishing Co., Inc*, 2002.
- [23] Gonzalez R. C., and R. E. Woods, "Digital Image Processing," Upper Saddle River, NJ: Prentice Hall, 2018.
- [24] Marques O., "Practical Image and Video Processing using MATLAB," *Hoboken, New Jersey: John Wiley & Sons, Inc*, 2011.
- [25] Gao W., et al., "An Improved Sobel Edge Detection," 2010 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT), Chengdu, China pp. 67-71, 2010.
- [26] Pratt W. K., "Digital Image Processing: PIKS Inside," Fourth ed. John Wiley & Sons, Inc, 2007.
- [27] Jitendra B. Jawale and S. M. T. Anjali S. Bhalchandra, "The Human Identification System Using Multiple Geometrical Feature Extraction of Ear An Innovative Approach," *International Journal of Emerging Technology and Advanced Engineering*, vol. 2, no. 3, pp. 662-666, 2012.
- [28] Alaraj M., Hou J., and Fukami T., "A Neural Network Based Human Identification Framework Using Ear," Proceedings of TENCON 2010 IEEE Region 10 Conference, Fukuoka, Japan, 2010.