An ear recognition system based on local wavelet subband energy distribution

Ruaa Isam Fadhil[1], Loay E. George[2]
[1]Ministry of education, Baghdad, Iraq
[2]University of Information Technology and Communication, Assistant President for Scientific Affairs, Baghdad, Iraq

Article Info

Article history:

Received May 15, 2021
Revised Sep 2, 2021
Accepted Sep 13, 2021

Keywords:

Central moment coefficients nearest classifier
Ear recognition
Haar wavelet transform

ABSTRACT

The outer ear features have been used for many years in forensic science of recognition. Human ear is a valuable information provenance of data for individual identification/authentication. Ear meets biometric characteristic (universality, distinctiveness, permanence and collectability). Biometric system depending on ear image facing two major challenges; the first one is the localization of human ear area in given profile face image, and the second one is the selection of proper features to distinguish between individuals. In this work, we propose an alogorithm for ear recognition based on the local spatial energy distribution of wavelet sub-bands, because of wavelet transform has the ability to analyze the local feature of 2-D image by determining where the low frequency and high frequency areas are and it provides full description of the spatial distribution of the ear image. Nearest classifier are used to make a recognition decision in matching stage. The system was tested over a public database consist of 493 images. The attained recognition rate was (95.28%) and the achieved minimum equal error rate (EER) is 0.02%.

Corresponding Author:

Ruaa Isam Fadhil
Ministry of education
Baghdad, Iraq
Email: Roaaess87@gmail.com

1. INTRODUCTION

As there is an increasing need to automatically recognize individuals, many methods are used for personal identification, and it has been an efficient field of research over the last decade [1], [2]. Passwords, and ID cards represented traditional methods for personal identification, but they can be purloin, forge, or forgotten, while biometric method has many characteristics, such as: universal, unique, perpetual, and could be measured [3]-[5]. The shape of the outer ear known for many years as a valuable resource for personal identification by criminal investigators, Alphonse Bertillon the French criminologist was the earliest to recognize the potential deployment of ear shape as distinctive characteristic for identifying humans, more than a century ago [6]. The ear can be captured easily from a distance, and don't require a person to entirely be cooperating [7]. This produces ear recognition as an interesting technique for smart monitoring functions and for forensic image analysis. It is worth taking into consideration that ear images is a more reliable uni-modal biometric recognition technique than face biometric recognition techniques, basically since the association of ear image with a given individual is very difficult in fact, most of individuals are not capable of recognizing their own image, subsequently, the ear databases do not require being secured as the face databases, since the possibility of attacks is much lowering [8], and it is require less computation time than other biometric techniques, since the size of ear images are relatively small [9]. Furthermore, ear shape didn't affected by expression, mode, or health. But, recognition systems based on ear images, still suffer from many issues such as illumination, pose, and obstruction [10]-[15]. All these challenges should be taken in consideration when design ear recognition system.

Biometric systems based on ear images facing many challenges can be grouped mainly into two main parts, first is allocating ear region and eliminating unwanted skin and hair area. Furthermore, the images are captured in different illuminance circumstances which produce images with many problems such as: noise, bluer, and low illuminance, which makes allocating ear area very complicated. The second one is finding proper features to represent ear image for distinguishing individuals. In this paper we handle these challenges by pre-processing of ear images to improve the ear image data (features) throw suppressing unwanted data (surrounding skin and hair region) and enhancement of some important ear image features so that ear recognition system can benefit from this improved data in feature extraction and decision making stages. representing in: image enhancing, after that image size normalization required to unified features in feature extraction step. In feature extraction the local spatial energy distribution of wavelet sub-bands of ear image is applied, to decompose ear image into different resolutions. For reducing the number of wavelet coefficients, and preserveing image information the produced image is divided into blockes with overlap, central moment are calculated for each block to represent the ear image features. The suggested system could used as a tool for extracting ear region and features of different ear images in colors, shapes, and size. The paper rest of paper is organized as follows, Section 2 discusses the research methodology, Section 3 describes the experimental resultsand descussion, and finally conclusion are provided in Section 4.

2. RESEARCH METHOD

During the last few years, researches paid a lot of attention to the ear biometric system due to its characteristics. Recent studies have introduced different methods for biometric recognition. Geometrical measures based on ear edge images are used, because of its invariant to parallel move, scale and rotation, the feature vector composed of multiple geometrical feature, such as (shape, Euclidean distances of side of a triangle, and angles of a triangle), but the images may suffer from a problem with the outer shape of the ear, which may cause the failure of the whole system [12], [13]. Other studies combine multible technique to improve recognition results such as, appling of a back propagation (BP) artificial neutral network with geometrical features [14]. Then researchs attended to use methods for change the space and data representation, to decrease the dimensionalities, or to choose only the valuable information for feature extraction. A combination of elliptical local binary pattern operator and haar wavelets transform as a method for characterizing the specific details of the two dimensional ear images in [15] were proposed, this approach is based on pixel information, the pixels of the ear image are arranged, and processed in one vector, while the size of the vector represented by the total number of the pixels, principal component analysis (PCA) [16], color spaces fusion [17], 2D Gabor filter [18] are similar common techniques.

This paper produces an automated ear recognition system of both spatial and geometrical features. There are three stages of the proposed method. First, preprocessing is applied to allocate ear region and unify ear image size in order to improve the feature vector. Next, extracting the features by apply 2-D Haar wavelet transform, then, image portioning into blockes with overlap for local feature extraction inorder to generate statistical norm to build the feature vector. Finally, compares the extracted feature set (vector) with the feature sets that are already extracted from training samples and saved as template vectors in a database to define the identity or authenticity of a person whose ear is being tested as shown in Figure 1.

2.1. Preprocessing

Preprocessing is an important stage that afact the outcoming data, and it is consider to be a challenging one, so it requires many steps to overcome the artifact in the acquired image. The involved steps of this stage are the followings.

2.1.1. Allocation of ear region

This is an important step and the most challenging step; it aims to define the ear area from all surrounding regions. The accuracy of the ear region allocation process greatly affects the whole process of identification or verification task. The allocation process implies the following image processing steps; it is consist of the steps given.

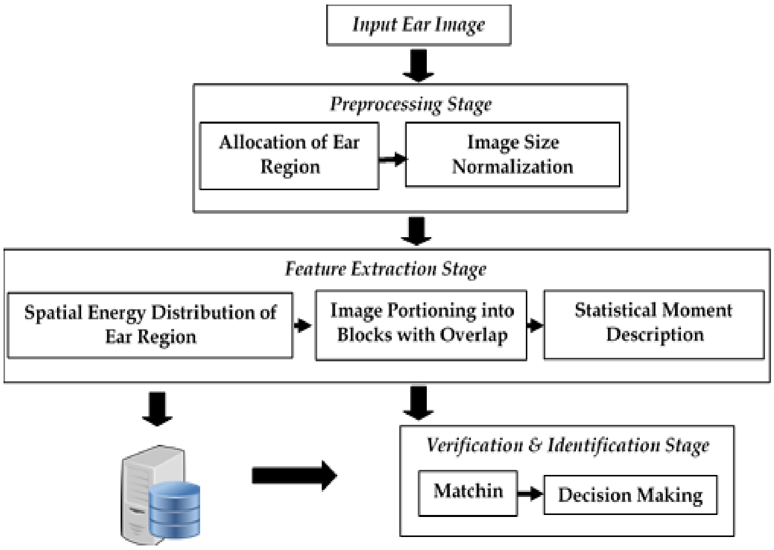
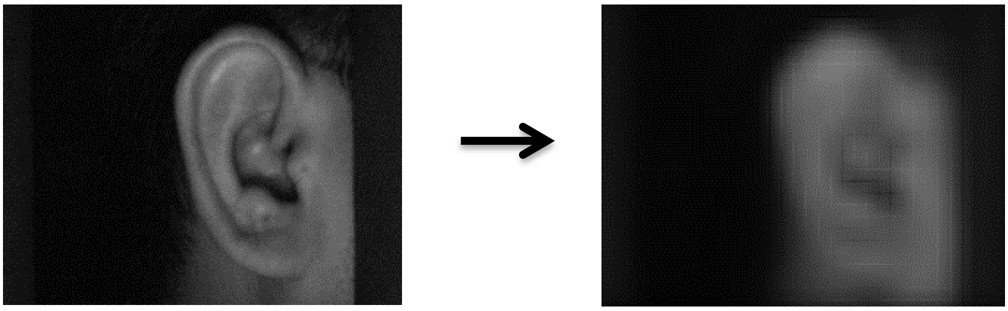


Figure 1. The proposed system layout

a) Cubic spline

The original ear image is captured as part of the side part of the face; it holds unwanted area which increases the required computation complexity and scale down the accuracy of matching. In order to enhance the ear region allocation task a cubic spline interpolation was applied. Cubic spline interpolation is a piecewise continuous curve, with continuous derivatives of first and second order [19]. Cubic spline produces a smooth ear image in such a way that we can correctly define skin region from unwanted region (e.g. hair region). Figures 2(a) and (b) shows the smoothness of the output image after applying the cubic spline.



(a) (b)

Figure 2. Image smoothing; (a) original image, (b) cubic spline image

b) Ear image enhancement using histogram equalization

The original ear image is enhanced using histogram equalization method. This method leads to redistribute the original image histogram in order to obtain more contrasted image whose histogram is wider [20], [21]. The expansion of luminosity distribution is procured by gathering the adjacent grey values to specific value. So, the grey levels number of the enhanced image is less than the grey levels number that belongs to the original image. This step has efficient effect to discriminate the ear regions from the surrounding skin area as shown in the Figure 3(a).

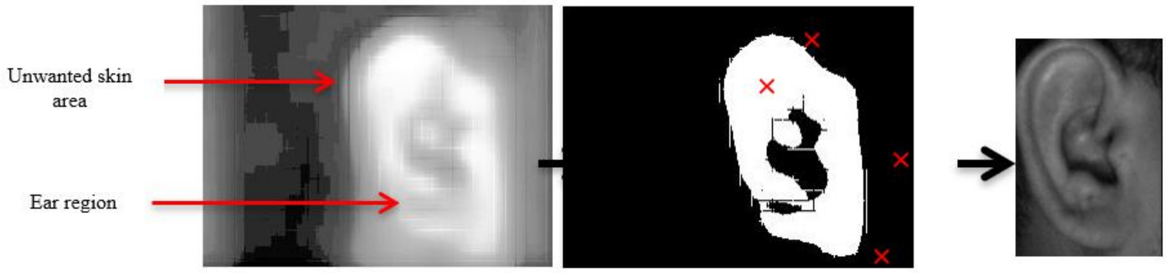
c) Binarization

Binarization is a process of transforming a gray image to a binary image which contains only two classes, black (pixel value =0), and white (pixel value =1). Since the image brightness was enhanced, the binarization process become simple; whereas the enhanced image consists of separated intensity levels that facilitated the binarization process. In the proposed system the binarization process depends on the value of global threshold; it was selected depending on the highest intensity found in the closest region to the capture device as shown in Figure 3(a). Where max represents the maximum intensity value and thr is the threshold value.

$$g'(i,j) = \begin{cases} 1 \; if \; g(i,j) \geq max - thr \\ 0 \; otherwise \end{cases} \quad (1)$$

d) Ear region localization

The purpose of this step is to extract ear region, without unnecessary parts like skin, and hair parts. Firstly, make a scanning for the coordinates of a skin area in order to find the minimum and maximum points in x and y - coordinates (first and last hit to white pixel), as shown in Figure 3(b). Second clipping ear region according to the rectangle represented by the allocated points as shown in Figure 3(c).



(a) (b) (c)

Figure 3. Ear region segmentation; (a) histogram equalization image, (b) binarized image, (c) cropped ear region

2.1.2. Image size normalization

Ear image size, and shape suffer from many changes during image capturing, which cause serious problems in designing ear recognition algorithms. In the image database when we observing the ear image samples we can notice many size variations in the collected ear image patterns. Furthermore, ear image cropping causes more variation in ear images. So size normalization is a necessary step for excluding size invariance on ear images before feature extraction. It is mapped into a standard window size as shown in Figure 4(a). To apply this mapping the affine transformation technique is used on the ear images with the Bilinear Interpolation algorithm using four nearest neighbors for interpolation [21].

2.2. Feature extraction

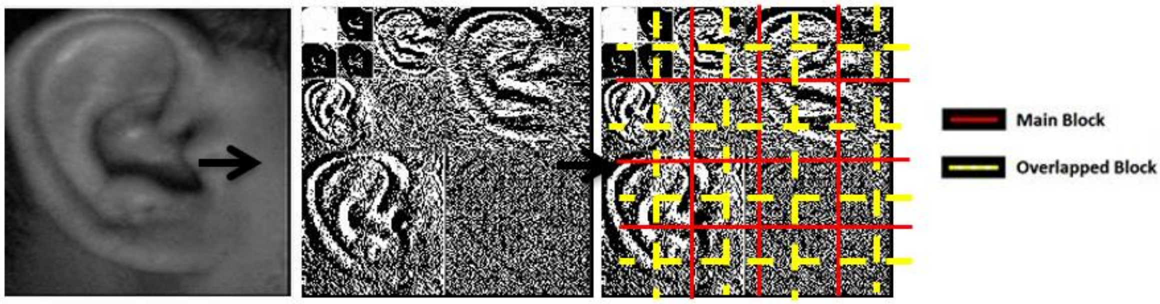
The major challenge for biometric systems that established on computer vision is to extract such features that will characterize individual ears in a distinctive technique. Discrete wavelet transform (DWT) is considered to be one of the common used image processing techniques in computer vision for object detection, analysis and classification [22]. The Implementation of DWT as an image processing method used for producing the transformation values (wavelet coefficient). In this stage the critical point is how to interpret the wavelet coefficient to symbolize individual for classification or detection. In this study, wavelet coefficients will be used in the processing and analyses of ear images since DWT decompose ear image into variant stages of resolution. By appling DWT, we can produce a new feature set depending on wavelet coefficient analyses of. The technique helps in reducing the required coefficients for feature vectors. The involved steps for determining the spatial distribution of sub-band wavelet energy are the followings as shown in Figure 4:

- Step 1: apply 2-D Haar wavelet transform in order decompose ear signal into four sub-images, where LL represents the low frequency (approximation) sub-band, HL refers to high frequency component along the horizontal direction, LH refers to high frequency component along the vertical direction and HH represents the diagonal high frequency component. The band LH, HL and HH are called detail (or wavelet) sub-bands. After first wavelet decomposition, the approximation (LL) sub-band is fed for next wavelet decomposition. Then, the second LL sub-band is submitted again for next decomposition is shown in Figure 4(b).
- Step 2: the wavelet image is divided into blocks with overlap as described in Figure 4(c).
- Step 3: Generate statistical norm to build the feature vector of the image. Image moments are utilitarian for describing objects after segmentation [23]. The adopted moments are the central moments instead of

the ordinary moments. They are computed in terms of deviations from the mean instead from the origin. The function of such moments is, mostly, selected to have some attractive property or feature.

$$\|Norm\|_p = \frac{1}{N\times N}\sum_{x=0}^{N-1}\sum_{y=0}^{M-1}\,(I(x,y)-mean)^p\,sign(I(x,y)-mean) \tag{2}$$

where, 0 < p < 1.



(a) (b) (c)

Figure 4. Feature extraction; (a) size normalized image, (b) 3-Passes DWT, (c) partitioning sub-band to overlapping blocks

2.3. Matching and decision making

In this stage, an input ear image is feeded to the system to calculate the degree of matching. The input ear image is processed to obtained features list that will be straightly matched with the previously saved templates using KNN classifier (k-nearest neighbor), it is considered as widely known algorithms for supervised learning in pattern recognition and, classification.

KNN classifier has many features: efficiency, simplicity, intuitiveness and competitive classification functionality in many area [24]. The KNN classifier is uses basically the euclidean distance for comparing samples, a test sample (input one) and the set of training samples (stored templates) with K value equal to one. In 1-nearest neighbor algorithm, the portend class of test sample *x* is adjust equal to the actual class ω of its nearest neighbor, where *mi* is a closest neighbor to *x* if the distance:

$$d(m_i,x) = max_j\{d(m_j,x)\} \tag{3}$$

3. RESULTS AND DISCUSSION

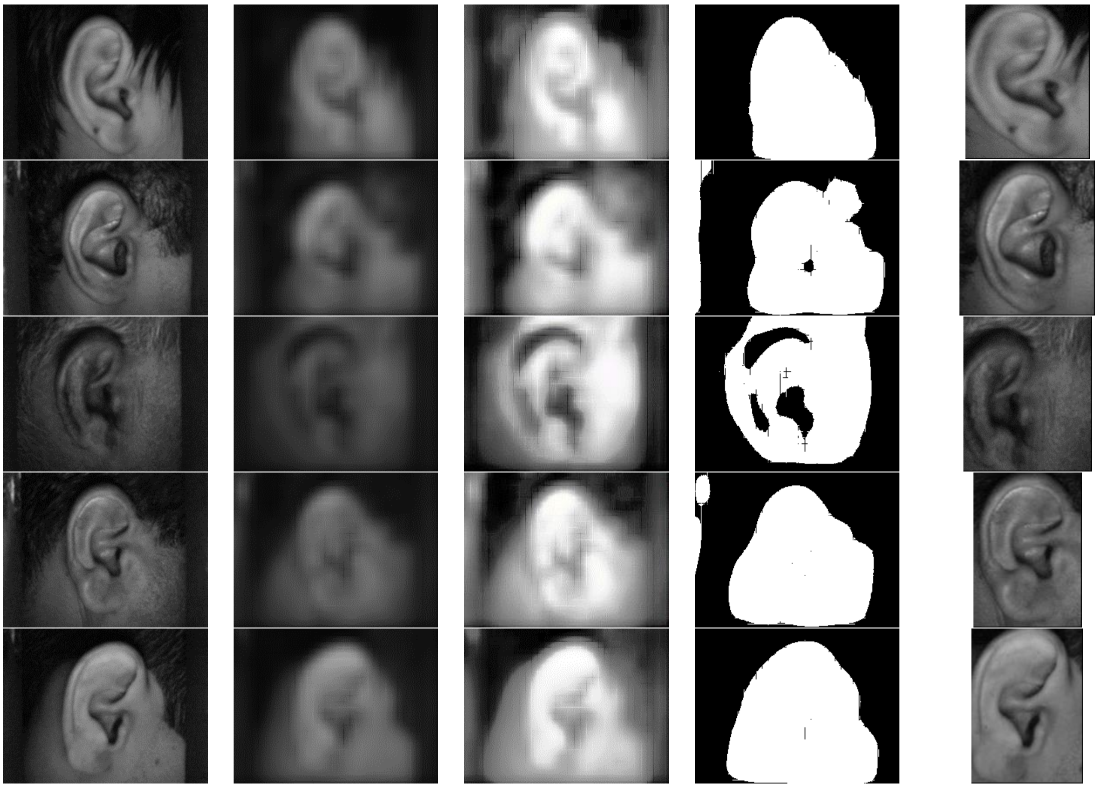
The dataset used for testing in this research is taken from Delhi ear image database that is publicly available. The ear images are obtained from a distance (touchless). The database is obtained from 125 individuals, and each one has at least three ear images. The resolution of the obtained images is 272 × 204 pixels, and are available in bmp format. In the following experiments, the data set had been divided into two sets, one for training, and the other for testing. A first set consist of 280 samples had been used for training to build the nearest neighbor (NN) classifier, and the second set consist of 213 samples had been used for testing the proposed system. All the images are preprocessed and the ear region are allocated from the acquired image as shown in Figure 5, where Figure 5(a) represent the aquired image, Figure 5(b) represent the image after applying cubic spline for image smooth, Figure 5(c) histogram equalization image to adjust the brightness, Figure 5(d) converting to binry image to allocate the ear region, Figure 5(e) cropped image.

3.1. Identification (recognition) results

The performance of identification system is evaluated by applying the correct recognition rate (CRR); which represented the ratio between the number of corrects recognition decisions (n_c) and the totals number of tried tests (n_T):

$$CRR = \frac{n_c}{n_T} \tag{4}$$

Tables 1, 2 and 3 illustrates the attained recognition results when applying statistical norm3/4 on the ear image. The attended results shows that the recognition rate is increased with respect to increasing the wavelet passes, increasing block number causes delay in the system and dispersion in the block information, while increasing overlap ratio lead to increasing block size.



(a) (b) (c) (d) (e)

Figure 5. The results of ear region allocation; (a) original image, (b) cubic spline image, (c) histogram equalization image, (d) converting to binry image, (e) cropped image

Table 1. The recognition rate for different wavelet passes with block size (9×9) and overlap ratio (0.2)

Wavelet Passes	Recognition Rate (%)
1-pass	86.32
2-pass	87.26
3-pass	95.28

Table 2. The recognition rate for different block size

Block Size	Recognition Rate (%)
5×5	82.07
7×7	84.9
9×9	95.28
11×11	90.56
13×13	92.45

Table 3. The recognition rate for different overlap ratio

Overlap Ratio	Recognition Rate
0.0	85.37
0.1	91.03
0.2	95.28
0.3	92.04
0.5	90.09

3.2. Verification (authentication) results

The receiver operating characteristic (ROC) curve is used to evaluate the performance of verification system, it performs the false rejection rate (FRR) against the false acceptance rate (FAR) at various thresholds on the matching score. The system threshold value is obtained according to the equal error rate (EER) criteria, where FAR = FRR.

$$FAR = \frac{A}{B}\ , FRR = \frac{C}{D} \quad (5)$$

where, A is the number of successful authentications by impostors, B is the number of attempts at authentication by unauthorized users, C is the number of failed attempts at authentication by authorized users, and D is the number of attempts at authentication by genuine users. Furthermore accuracy parameter can be used to evaluate the performance of biometric systems (i.e., the proportion of correct predictions) and it does not need to take into consideration what is positive (P) and what is negative (N) [25], [26].

$$ACC = \frac{TP+TN}{P+N} \tag{6}$$

where true positive (TP) is the number of genuine users that identified correctly, true negative (TN) is the number of impostor users attempts that rejected by the system. Table 4 shows FAR, FRR and accuracy values for different threshold values, the verification results is performed using the best parameters setup that lead to best recognition rate. The ROC curve between the FAR and FRR with various thresholds is plotted in Figure 6. The equal error rate is 0.02% for the threshold value is equal to 0.015.

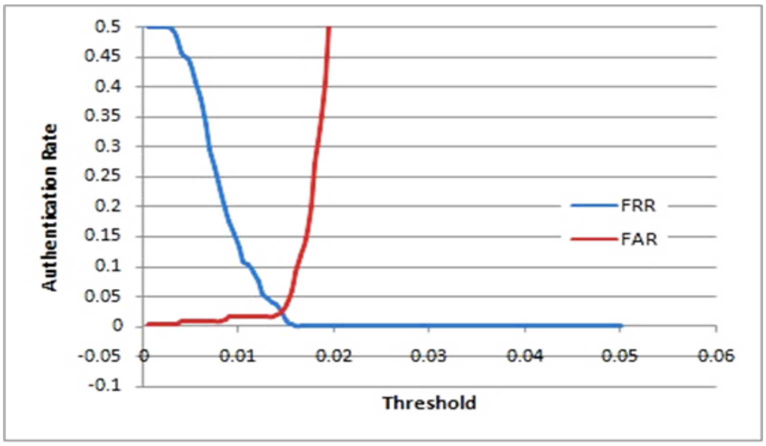


Figure 6. The ROC curve

Table 4. FAR, FRR and accuracy versus threshold values

Threshold	FRR(%)	FAR(%)	Accuracy(%)
0.014	0.02	0.036	0.972
0.0145	0.024	0.024	0.976
0.015	0.036	0.008	0.978
0.0155	0.06	0.004	0.968
0.016	0.092	0.001	0.954

4. CONCLUSION

In this paper, a personal verification and identification system based on the ear image is introduced. Allocating ear region is achallenging step, because of all following steps are depending on it. So it required many tasks to do it starting with image preprocessing with cubic spline and histogram equalization, then image binarazation is applied, after that skin region allocation is performed. This step may applied for such systems that require skin region localization. A new feature set is proposed in this work; it depends on the "local spatial energy distribution of wavelet sub-bands". The conducted results indicated that the proposed system achieved high recognition rate CCR 95.28%, and EER equal to 0.02%, which indicates high performance in verification.

REFERENCES

[1] G. M. Shivakumar and C. M. Patil, "An Investigation towards Effectiveness of Present State of Biometric-Based Authentication System", *International Journal of Research and Scientific Innovation (IJRSI)*, vol. 4, no. 7, pp. 44-55, July 2017.
[2] Z. Wang, J. Yang, and Y. Zhu, “Review of Ear Biometrics,” *Arch Computat Methods Eng.*, vol. 28, pp. 149-180, 2021, doi: 10.1007/s11831-019-09376-2.

[3] S. B. Kutemate and R. U. Shekokar, "Secure and Reliable Human Identification Based on Finger-Vein Patterns," *International Journal of Engineering Research & Technology (IJERT)*, vol. 4, no. 3, pp. 978-980, 2015, doi:10.17577/IJERTV4IS030938.

[4] T. F. Kiely, "*Forensic Evidence: Science and the Criminal Law*," Second Edition, chapter Ear Impressions, pp. 368-370, CRC Press, 2005.

[5] S. Nikose and H. K. Meena, "Ear-biometrics for human identification," *IEEE*, 06 October 2020, doi: 10.1109/ACCTHPA49271.2020.9213190.

[6] A. Bertillon, "*La Photographie Judiciaire: Avec Un Appendice Sur La Classification Et L'Identification Anthropometriques*," Gauthier-Villars, Paris; 1890.

[7] L. Yuan and Z. Mu, "Ear Recognition Based on Gabor Features and KFDA," *The Scientific World Journal*, vol. 2014, Article ID 702076, p. 12, 2014, doi: 10.1155/2014/702076.

[8] M. Khosrow-Pour, "*Encyclopedia of Information Science and Technology*," Third addition, IGI Global, 2014, doi: 10.4018/978-1-5225-2255-3.

[9] B. S. El-Desouky, M. El-Kady, M. Z. Rashad, and M. M. Eid, "Ear Recognition And Occlusion," *International Journal of Computer Science & Information Technology (IJCSIT)*, vol. 4, no. 6, pp. 97-104, 2012, doi: 10.5121/ijcsit.2012.4608.

[10] C. Sforza, G. Grandi, M. Binelli, D. G. Tommasi, R. Rosati, and V. F. Ferrario, "Age- And Sex-Related Changes In The Normal Human Ear," *Forensic Science International*, vol. 187, no. 1-3, pp. 110.e1-110.e7, 2009, doi: 10.1016/j.forsciint.2009.02.019.

[11] L. Meijerman, C. V. D. Lugt and G. J. R. Maat, "Cross-Sectional Anthropometric Study of the External Ear," *Journal of Forensic Sciences*, vol. 52, no. 2, pp. 286-293, 2007, doi: 10.1111/j.1556-4029.2006.00376.x.

[12] Z. Mu, L. Yuan, Z. Xu, D. Xi, and S. Qi, "Shape and Structural Feature Based Ear Recognition," *Lecture Notes in Computer Science (LNCS)*, vol. 3338, pp. 663-670, 2004, doi: 10.1007/978-3-540-30548-4_76.

[13] B. Jawale and 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.

[14] X. Q. Wang, H. Y. Xia, and Z. I. Wang, "The Research of Ear Identification Based on Improved Algorithm of Moment Invariants," In *Proceedings of Third Int. Conference on Information and Computing (ICIC)*, Wuxi Jiang Su, pp. 58-60, 2010, doi: 10.1109/ICIC.2010.21.

[15] Benzaoui, A. Kheider, and A. Boukrouche, "Ear description and recognition using ELBP and wavelets," in: *Proceedings of the International Conference on Applied Research in Computer Science and Engineering*, pp. 1-6, 2015, doi: 10.1109/ARCSE.2015.7338146.

[16] K. Chang, K. W. Bowyer, S. Sarkar, and B. Victor, "Comparison and Combination of Ear and Face Images in Appearance-Based Biometrics," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1160-1165, 2003, doi: 10.1109/TPAMI.2003.1227990.

[17] L. Nanni and A. Lumini, "Fusion of color spaces for ear authentication", *Pattern Recognition (Elsevier)*, vol. 42, pp. 1906-1913, 2009, doi: 10.1016/j.patcog.2008.10.016.

[18] C. Wu Kumar, "Automated human identification using ear imaging," *Pattern Recognition (Elsevier)*, vol. 45, pp. 956-968, 2012, doi: 10.1016/j.patcog.2011.06.005.

[19] J. S. Jimmy Li and S. Randhawa, "colour filter array demosaicking using cubic spline interpolation," *Conference Paper in Acoustics, Speech, and Signal Processing*, 1988. ICASSP-88, 1988 International Conference, no. 1, pp. I-865-I-868, 2007, doi: 10.1109/ICASSP.2007.366045.

[20] S. S. Bagade and V. K. Shandilya, "Use of Histogram Equalization in Image Processing for Image Enhancement," *International Journal of Software Engineering Research and Practices*, vol. 1, no. 2, pp. 6-10, 2011.

[21] C. L. He, P. Zhang, J. Dong, C. Y. Suen, and T. D. Bui, "The Role of Size Normalization on the Recognition Rate of Handwritten Numerals," *Centre for Pattern Recognition and Machine Intelligence*, Concordia University Montreal, Quebec, Canada H3G 1M8.

[22] Y. Ernest, C. M. Lam Yuan, and Y. Tang, "Feature Extraction Using Wavelet And Fractal", *Pattern Recognition Letters*, vol. 22, no. 3-4, pp. 271-287, 2001, doi: 10.1016/S0167-8655(01)00003-4.

[23] M. Nixon and A. Aguado, "*Feature Extraction & Image Processing for Computer Vision*," Third Edition, Academic Press, 2012.

[24] J. Gou a, L. Du, Y. Zhang, and T. Xiong, "A New Distance-weighted k-nearest Neighbor Classier," *Journal of Information & Computational Science*, vol. 9, no. 6, pp. 1429-1436, 2012.

[25] R. M. Polli, A. V. Maran, A. T. Z. Jouglas, E. Silva, P. S. Brandi, and D. I. Hass, "A Proposal for the Hand Palm Identification, Using Local Binary Pattern," *International Journal of Advanced Engineering Sciences and Technology (IJAEST)*, vol. 9, no. 2, pp. 302-309, 2011.

[26] R. E. Fadhil and L. E. George, "The Use of Spatial Distribution of the Local Histogram Based Features for Finger's Veins Biometrics," *British Journal of Mathematics and Computer Science*, vol. 15, no. 4, no. BJMCS.24729, pp. 1-11, 2016, doi: 10.9734/BJMCS/2016/24729.