

Feature Detection using KAZE and Harris Detectors for Ear Biometrics

Mrunmayi Sunil Sawant

Department of Electronics and Telecommunication
V.E.S.I.T

Chandan Singh Rawat

Associate Professor, Department of Electronics and
Telecommunication V.E.S.I.T

Abstract—The importance of human ear as a suitable biometric feature for human identification was contrived a long time ago. There are significant proofs that demonstrate that human ear alone can be utilized as a biometric trait because it overcomes the constraints of other biometric traits. Human ear can also be used in integrated manner with some other biometric features like fingerprint, iris for a better performing biometric system known as multi-modal biometric technique. Accuracy of a featured oriented image classification system is mainly depended on feature detector – extractor used. Therefore, choosing an efficient feature detector is of utmost importance. This paper proposes feature detection using KAZE (blobs) and Harris (corners) algorithms.

Keywords—Biometrics, identification, KAZE, Harris, feature detection, matching

I. INTRODUCTION

The appearance of the outer ear in human beings (or pinna) is formed by the outerhelix, the lobe, the tragus, the antihelix, the antitragus, and the concha. Figure 1 illustrates the anatomy of human ear. The countless ridges and valleys on the outer ear's surface serve as acoustic resonators. At low frequencies the pinna reflects the acoustic signal towards the ear canal. At higher frequencies the pinna reflects the sound waves thus causing neighbouring frequencies to drop. In addition, the outer ear allows us humans to recognize the origin of a sound. The shape of the outer ear develops during the embryonic stage from six growth nodules. The structure of the ear therefore, is not completely random, but still subjected to cell segmentation. Though there are similarities between the right and left ear they are not symmetrical [1].

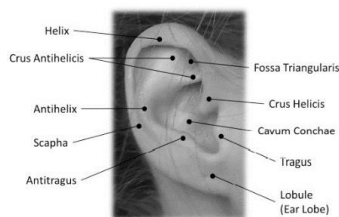


Figure 1: Human ear anatomy [1]

Human ear has remarkable characteristics which makes it a good trait for biometric systems. Medical studies have proven that vital changes in the shape of the ear take place before age of 8 years and after 70 years of age. Hence, ear is considered as a stable feature. Ear is also a great example of passive biometric trait where much support is not needed from the subject. Ear can be captured even without the knowledge of the subject from farther distance. Size of the ear is also one of the most important aspect. Size of the ear is smaller than

that of the face but larger than the iris and the fingerprints and hence image of the ear can be collected without much hassle. Ear is not changed under the influence of cosmetics or with the use of eye glasses. Moreover, as there is absolutely no need to touch any equipment, there are no issues related to hygiene. Ear images are also much more secured than face mainly because it is very difficult to associate ear image with a particular person. This makes the ear database extremely secure with low risk of attacks.

Iannarelli (1989) [2] states that the human ear is an unique feature of each individual. Dividing the ear into eight parts, twelve measurements are taken around the ear and distances are measured by placing a compass over an enlarged image. *Hurley and Carter* [3] described multiple methods for identification combining results from different neural classifiers. *Nixon and Carter (2000a, 2000b)* have used a different technique called force field transformation where the image is treated as gaussian attractors. The method used by *Choras (2005)* [4] was based on the new co-ordinate system in the centroid, making any rotation of the image irrelevant for the purpose of identification. Later, *Choras (2007)* [5] added additional experiments to further expand his earlier study by recommending multi-modal (hybrid) biometric system. *Middendorff et al. (2007)* [6] emphasized that the type of data that is used whether 2D or 3D, the type of classification algorithm performed on each type of data, the output of the performed algorithm, the fusion type performed to combine them can improve the performance of a biometric system. *Kisku et al. (2009)* [7] proposed a multi-modal recognition system using ear and fingerprints based on Scale Invariant Feature Transform (SIFT). Another research in ear biometric by *Zhou et al. (2001)* [8] includes a robust technique 2D ear recognition using colour SIFT features.

The IIT Delhi ear database is used which is provided by the Hong Kong Polytechnic university. (https://www4.comp.polyu.edu.hk/~csajaykr/myhome/databas e_request/ear/). Figure 2 shows some sample images from database.

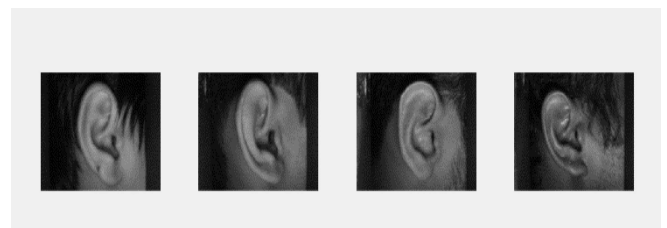


Figure 2: Sample from ear database

II. METHODOLOGY

Figure 3 shows a generalised block diagram of the biometric identification system based on the ear images.

Pre-processing:

Normally, the input images are affected by unwanted signal such as, noise in the capturing device, lighting variation, changes in background because of which the required details cannot be extracted satisfactorily. Hence, the input image needs to be pre-processed to remove unwanted signals in order to preserve the required details.

Pre-processing is the first step implemented in the system. It is mainly done to segment the ear portion from the rest of the entire image. First the image is cropped into suitable dimensions as shown in figure 4. Ideally the next step would be RGB to gray scale conversion but since the IIT Delhi database already consists of gray scale images this step has been skipped. The next step is to perform histogram equalization which is a technique for adjusting image intensities to enhance contrast as shown in figure 5. Both grayscale and color images contain a lot of noise or random variation in brightness among pixels. Gaussian filtering is done to reduce this noise that is added during the acquisition phase as shown in figure 8.

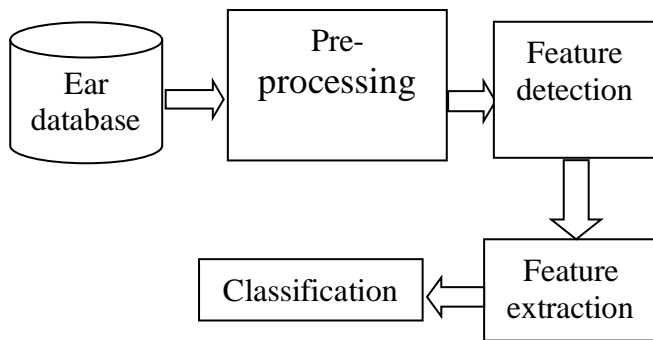


Figure 3: General block diagram of biometric system.



Figure 4: Cropped image



Figure 5: Histogram equalized image

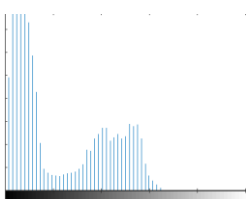


Figure 6: Original histogram image

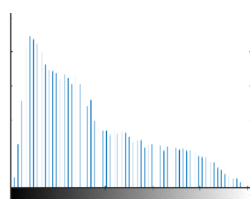


Figure 4: Equalized histogram of cropped image



Figure 5: Gaussian filtering



Figure 6: Sharpened image

Feature detection:

A feature detection algorithm detects feature points also called as interest points or key points in a picture. Features are generally detected as corners, blobs, lines, edges etc. The key to feature detection is to find features that remain locally invariant so that they can be detected even in the presence of rotation or scale change. The features which are detected are described in different ways on the basis of distinctive patterns owned by their neighboring pixels. KAZE and Harris features are detected

KAZE feature detector:

KAZE features are novel multiscale 2D feature detectors. These features were proposed by *P.F Alcantarilla* in 2012 [9] that make use of non-linear scale space through non-linear diffusion filtering [9]. This makes blurring in images locally adaptive to interest points thereby decreasing the noise and also holding on to the boundaries of region in images. This detector is based on scale normalized determinant of Hessian matrix which is calculated at multiple scale levels. [10]

$$L_{Hessian} = \sigma^2(L_{xx}L_{yy} - L_{xy}^2) \quad n \quad 1$$

Where (L_{xx}, L_{yy}) are second order horizontal and vertical derivatives respectively and L_{xy} is the second order cross derivative. The maxima of the detector response are picked up as feature-points with the help of a moving window. Feature description establishes the characteristic of rotation invariance by searching dominant orientation in a circular neighborhood around each recognized feature. KAZE features are rotation and scale invariant, have limited affine and have more distinctiveness at varying scale. Equation (2) shows the standard non-linear diffusion formula.

$$\frac{\partial L}{\partial t} = \text{div}(c(x, y, t) \cdot \nabla L) \quad 2$$

Where “c” is conductivity function. “div” is divergence. “∇” is gradient operator and “L” is image luminance.

Harris corner detector:

In the year 1988, *C.Harris* and *M.J.Stephens* proposed Harris algorithm which is based on the Moravec algorithm [11]. The corner points are the prime key feature information in the image. The Harris algorithm is based on the gradient of two pixel points, which determines whether the pixel point is a corner point or not. The Taylor series expansion method is used to calculate the gray change trend of the window, which moves in any direction [11]. And then detect the corner point by the mathematical analysis equation.

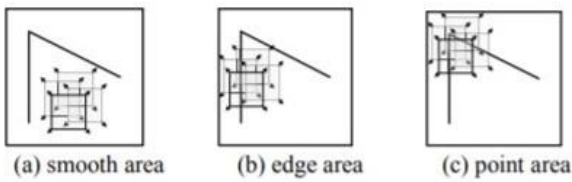


Figure 7: Harris window for corner detection

As shown in Figure 10, three cases are considered

1. If the window image patch is smooth then all shifts will result in small changes.
2. If the window mounts on edge then a shift results in small change but a shift perpendicular to edge will result in large changes.
3. If the window patch is a corner then all shifts will be large. Hence a corner will be detected.

Harris algorithm pseudocode [12]:

- Calculate the gradient of image $I(x,y)$ in x and y direction.

$$I_x = [-1,0,1] \otimes I(x,y);$$

$$I_y = [-1,0,1]^T \otimes I(x,y) \quad 3$$

- Calculate the product of the two directions I_x^2, I_y^2 and $I_x I_y$
- Calculate the correlation matrix M for every pixel points.

$$M = \begin{bmatrix} A & B \\ B & C \end{bmatrix},$$

$$A = I_x^2 \otimes w_{u,v},$$

$$B = I_{xy} \otimes w_{u,v},$$

$$C = I_y^2 \otimes w_{u,v} \quad 4$$

- Calculate the corner response (R)

$$R = Det - k(Tr)^2 \quad 5$$

where

$$Tr(M) = A + C \quad 6$$

$$Det(M) = AC - B^2 \quad 7$$

(k – empirical constant, $k = 0.04-0.06$)

- Get non-maximal suppression in the d of 3×3 or 5×5 . If the neighbourhood response value is bigger than the threshold value, the maximum value point in the neighbourhood is the corner point of the image.

Feature extraction using Histogram of Oriented Gradient (HOG) :

Feature extraction involves computing a descriptor which is usually done on regions centered around detected features. Descriptors depend on image processing to remodel an area pixel neighborhood into a compact vector representation. Histogram of Oriented Gradients (HOG) are successfully and coherently used to solve the issues related to object detection and recognition, precisely where illumination variations are considered. [12]

Histogram of Oriented Gradients calculates the occurrences of different gradient orientations within an area neighborhood of a picture. Each pixel within a cell or local neighborhood contributes with a weighted vote for an orientation histogram channel supported the values found within the gradient computation. A bunch of spatially connected cells are called a block where the gradient strengths are locally normalized within each block to account for changes in illumination and contrast.

HOG descriptor involves five steps, which are the gradient computation, orientation binning, histogram computation, histogram normalization and concatenation of local histograms. The algorithm starts by computing the local gradient by convolving a 3×3 region (HOG cells) with two one-dimensional filters (-101) and (101) . The local orientation at the middle of every HOG cell is that the weighted sum of filter responses of every pixel.

Classification :

Support vector machine (SVM) is one amongst the foremost powerful and successful statistical classification technique. The support vector machines are supervised learning models with its associated learning algorithms that analyze data and recognize the patterns.

- The computational complexity of SVMs doesn't rely upon the dimensionality of the input space.
- SVM training always finds a worldwide minimum and their simple geometric interpretation provides fertile ground for further investigation.
- The SVM approach doesn't try and control model complexity by keeping the quantity of features small.
- In comparison with traditional multilayer perception neural networks that suffer from the existence of multiple local minima solutions, convexity is a very important and interesting property of nonlinear SVM classifiers [14].

III. SIMULATION RESULTS

Matlab-2018 has been used for performing the simulation in this paper. Specifications of computer system used are: Intel (R) Core (TM) i5-4200U CPU @ 1.60 Hz, 4GB RAM. The database comprises of 493 grayscale images of 125 subjects. The number of images used per subject varies between 3 to 6. For the simulation purpose images of 35 subjects are used.

The database is further divided into training set and testing set to calculate the accuracy. The number of images in training set varies from 3 to 4 per person and the number of images in testing set varies from 1 to 2 per person.

After pre-processing steps , the KAZE and Harris features are detected. The number of detected features are 620 and 542 for KAZE and Harris detector respectively as shown in table 1 and figures 11 & 12 respectively.

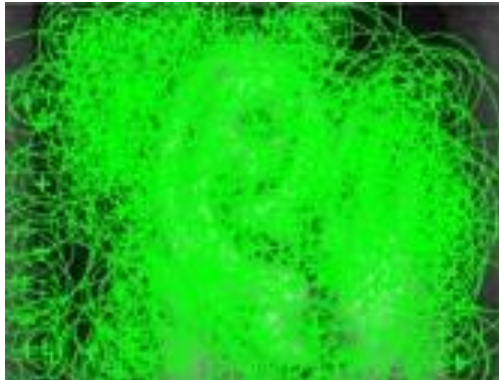


Figure 11: Detected KAZE features



Figure 12: Detected Harris corners

The detected KAZE AND Harris features are then fed into feature extraction technique using the Histogram of Oriented Gradients (HOG). HOG is the feature descriptor used to detect objects in computer vision and image processing. this method counts occurrences of gradient orientation in localized portions of a picture – region of interest. This feature is important to coach and test the algorithm which is employed to acknowledge the ear. For simulations 8×8 , HOG cells and 9 bin histograms are used as shown figure 13. Figure 14 & 15 shows the extracted HOG features on KAZE and Harris detectors respectively..

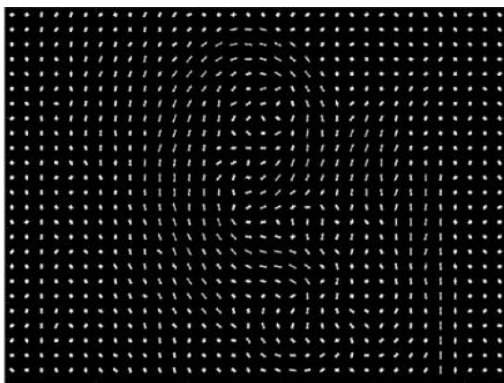


Figure 13:8x8 cell size of ear image

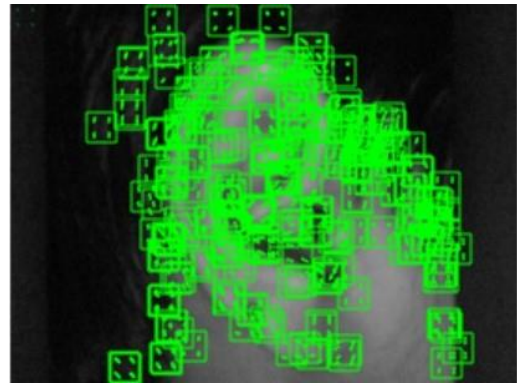


Figure 14: Extracted HOG features on KAZE detectors



Figure 15: Extracted HOG features on Harris detectors

In the last step (classification) support vector machine is employed for verification. SVM classifier is employed to calculate the accuracy of the system. From the confusion matrix of the tested ears average class accuracy is obtained for SVM classifiers. The Y-axis of the confusion matrix corresponds to the anticipated class (Output Class) and also the X-axis corresponds to truth class (Target Class). The diagonal cells correspond to observations that are correctly classified. The off-diagonal cells correspond to incorrectly classified observations. Figure 16 & 17 shows the plot of the confusion matrix for KAZE and Harris detector using HOG extractor and SVM classifier. The obtained accuracy is 91.86% for KAZE detector and 83.72% for Harris detector.

Table 1: Number of detected KAZE and Harris features

Sr.no	Detector used	No. of features detected	Detected images/Total images	Recognized rate
1.	KAZE	620	79/86	91.86%
2.	Harris	542	72/86	83.72%

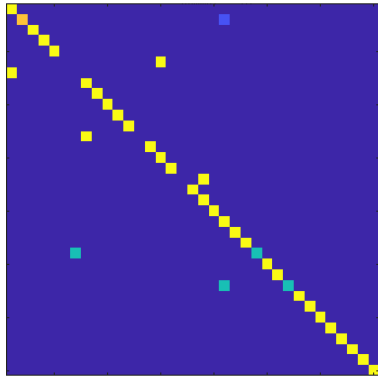


Figure 16: Plot of confusion matrix using KAZE-HOG and SVM classifier

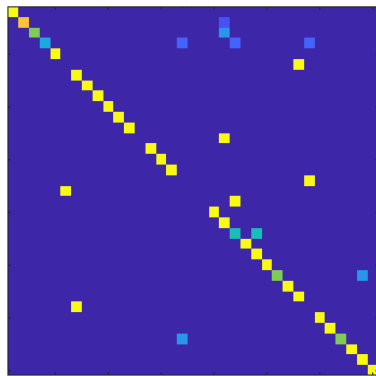


Figure 17: Plot of confusion matrix using Harris-HOG and SVM classifier

IV. CONCLUSION AND FUTUREWORK

In this paper KAZE and Harris features are detected and are extracted using Histogram of Oriented Gradients (HOG) extractor. Detectors that depend on gradient-based and intensity variation approaches detect good local features. These features include edges, blobs, and regions. Good local features exhibit the following properties: **Repeatable detections, Distinctive, Localizable.** KAZE and Harris detectors exhibit all the above properties. For simulation two completely opposite detectors are taken. KAZE detector is a blob detector whereas Harris is a corner detector. KAZE detector is scale and rotation invariant while Harris corner detector is not scale and rotation invariant hence, KAZE detected the highest feature points. Also, KAZE features have limited affinity and have more distinctiveness at varying scales. On the opposite hand, Harris corner detector is simply rotation invariant. They are not invariant to scale changes. Also, the standard of the detected Harris features degrade when the scale changes and they are partially invariant to affine intensity change. KAZE-HOG detector extractor pair is found to be the most accurate with recognition rate of 91.86%.

Ear biometric incorporates a scope within the research field under which there is a scope to develop many real time .The future work is to increase the number of ear images and use more databases to perform detection and also to evaluate the performance of the system in real time domain. Also, to pander challenges like occlusions like earrings, hairs etc.

V REFERENCES

- [1] Abaza, Ayman, Arun Ross, Christina Hebert, Mary Ann F. Harrison, and Mark S. Nixon. "A survey on ear biometrics." *ACM computing surveys (CSUR)* 45, no. 2 (2013): 1-35.
- [2] Iannarelli, A. "Ear Identification, Forensic Identification Series. 1989." *Fremont, Calif: Paramount Publishing.*
- [3] Hurley, David J., Mark S. Nixon, and John N. Carter. "Automatic ear recognition by force field transformations." (2000): 7-7.
- [4] Choras, Michal. "Ear biometrics based on geometrical feature extraction." *ELCVIA Electronic Letters on Computer Vision and Image Analysis* 5, no. 3 (2005): 84-95.
- [5] Chorás, Michal. "Emerging methods of biometrics human identification." In *Second International Conference on Innovative Computing, Informatio and Control (ICICIC 2007)*, pp. 365-365. IEEE, 2007.
- [6] Middendorff, Christopher, Kevin W. Bowyer, and Ping Yan. "Multi-modal biometrics involving the human ear." In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-2. IEEE, 2007.
- [7] Kisku, Dakshina Ranjan, Phalguni Gupta, and Jamuna Kanta Sing. "Feature level fusion of biometrics cues: Human identification with Doddington's Caricature." In *International Conference on Security Technology*, pp. 157-164. Springer, Berlin, Heidelberg, 2009.
- [8] H. T. Zhou,Z. "Multimodal Surveillance: Sensors, Algorithm and Systems," in *Artech House*, Lomdom, 2001.
- [9] Alcantarilla, Pablo Fernández, Adrien Bartoli, and Andrew J. Davison. "KAZE features." In *European Conference on Computer Vision*, pp. 214-227. Springer, Berlin, Heidelberg, 2012
- [10] Tareen, Shaharyar Ahmed Khan, and Zahra Saleem. "A comparative analysis of sift, surf, kaze, akaze, orb, and brisk." In *2018 International conference on computing, mathematics and engineering technologies (iCoMET)*, pp. 1-10. IEEE, 2018
- [11] Harris, Christopher G., and Mike Stephens. "A combined corner and edge detector." In *Alvey vision conference*, vol. 15, no. 50, pp. 10-5244. 1988.
- [12] Guiming, Shi, and Suo Jidong. "Multi-Scale Harris Corner Detection Algorithm Based on Canny Edge-Detection." In *2018 IEEE International Conference on Computer and Communication Engineering Technology (CCET)*, pp. 305-309. IEEE, 2018.
- [13] Pflug, Anika, Pascal Nicklas Paul, and Christoph Busch. "A comparative study on texture and surface descriptors for ear biometrics." In *2014 International Carnahan Conference on Security Technology (ICCST)*, pp. 1-6. IEEE, 2014
- [14] Classification of Human Ear by Extracting Hog Features and Support Vector Machine Naveena M.1 , G Hemantha Kumar 2 1 Scientist, HCPC, UPE, 2 Vice Chancellor 12 University of Mysore, Mysore-570006