Biometric Authentication by Ear and Face Using Image Fusion and SIFT Through Alpha Blending

Subhranil Som
Department of Computer Application
JIS College of Engineering

Jhelum Das
Department of Computer Application
JIS College of Engineering

Akash Pal
Department of Computer Science & Engineering
JIS College of Engineering

Aritra Dey
Department of Electrical Engineering
JIS College of Engineering

Abstract—**Biometric surveillance is a technology that measures and analyzes human physical and/or behavioral characteristics for authentication, identification, or screening purposes. In this paper an approach has been proposed combining the ear and face biometrics for the purpose of identification of a person. At first, two separate images of a human ear and side face have been taken and Discrete Wavelet Transform has been applied to de-noise the images. The two separate synthesized images have been fused using Alpha blending and on fused images SIFT has been applied to extract corners that can be saved as feature points in the database for matching, tracking and recognizing the subject.**

Keywords: SIFT, DWT, Alpha blending, Surveillance, Discrete, Wavelet

I. INTRODUCTION

In the field of identification Biometric identification techniques such as Fingerprint recognition, Facial recognition, Voice pattern, Handwritten Signature, Retina recognition, Iris recognition are well established. “The Imposter” is a 2012 British-American documentary film about the 1997 case of the French confidence trickster Frederic Bourdin, who impersonated Nicholas Barclay, a Texas boy who disappeared at the age of 13 in 1994. The impersonation was eventually unearthed as a result of the suspicions of a private investigator, Charles (Charlie) Parker, and an FBI agent, Nancy Fisher. In this film it has been shown that the imposter Frederic Bourdin was busted by Charles (Charlie) Parker. Charlie compared Nicholas Barclay ear image from a photograph with Frederic Bourdin’s ear image. He found the shapes of the two ears different. Eventually this fact went on to become important evidence [27]. This is one of the few cases where ear biometrics have been utilized successfully. The morphology of the outer ear is simple compared to t he rich texture of the iris or the random distribution of minutiae in fingerprints. The structure of the ear is fairly stable and robust to changes in facial expressions [7]. Human ear is a new class of relatively stable biometrics that has drawn researchers’ attention recently. A single reference 3D ear shape model and locates the ear helix and the antihelixes parts in registered 2D color and 3D range images. The experimental results on the UCR data set of 155 subjects with 902 images under pose variations and the University of Notre Dame data set of 302 subjects with time-lapse gallery-probe pairs are presented to compare and demonstrate the effectiveness of the proposed algorithms and the system by Hui Chen; Bhanu, B. [1]. Ping Yan; Bowyer, K.W. represent a complete system for ear biometrics, including automated segmentation of the ear in a profile view image and 3D shape matching for recognition to evaluate the system with the largest experimental study to date in ear biometrics, achieving a rank-one recognition rate of 97.8 percent for an identification scenario and an equal error rate of 1.2 percent for a verification scenario on a database of 415 subjects and 1,386 total probes [2]. In many real-world applications, unimodal biometric systems often face significant limitations due to sensitivity to noise, intra-class variability, data quality, non-universality, and other factors. Attempting to improve the performance of individual matchers in such situations may not prove to be highly effective. Multi-biometric systems seek to alleviate some of these problems by providing multiple pieces of evidence of the same identity [3]. Research on biometrics has noticeably increased. However, no single bodily or behavioral feature is able to satisfy acceptability, speed, and reliability constraints of authentication in real applications. The present trend is therefore toward multimodal systems [4]. The use of otoacoustic emissions (OAE) is identity verification biometric. OAE could be important as a biometric identifier in applications where users wear headsets since it is discrete and difficult to spoof. OAE are very low level [~17 dB sound pressure level (SPL)] sounds emitted from the human ear as part of the normal hearing process. They can occur spontaneously or be invoked by a suitable stimulus, these being known as transient evoked otoacoustic emissions (TEOAE) and distortion product otoacoustic emission (DPOAE). The level of background noise is the most significant practical factor that affects biometric performance [5].

In section II, Discrete Wavelet Transform [14] has been discussed which has been applied to de-noise the images of ear and side face. Section III describes Alpha Blending, a technique to fuse the synthesized de-noised images to extract more number of features than available from individual images of ear and side face [13, 9], using the SIFT descriptor [15,16, 17, 18, 19, 20, 25], discussed in section IV. The result of image fusion will be a new image which is more suitable for human and machine perception or further tasks of image processing such as image segmentation, feature extraction and object recognition [8]. Section V gives a brief overview of the advantages of SIFT over Harris Corner Detection [11, 12, 24]. The proposed work has been sequenced in section VI and an allied flow chart is given in section VII. A brief conclusion has been discussed in section VIII followed by the future scope and references in section IX and X respectively.

II. DISCRETE WAVELET TRANSFORM

The Discrete Wavelet Transform (DWT) is a linear transformation that operates on a data vector whose length is an integer power of two, transforming it into a numerically different vector of the same length. It is a tool that has been separated data, say x, into different frequency components, and has been studied each component with resolution matches to its scale. DWT [14] is computed with a cascade of filtering followed by a factor 2-sub sampling (Figure: 1).

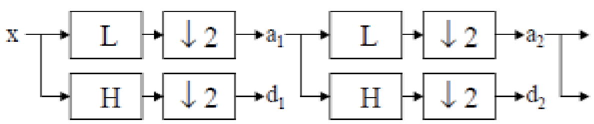


Figure 1 DWT Tree

H and L denote high and low-pass filters respectively, ↓ 2 denotes sub sampling. Outputs of this filter are given by equations (1) and (2)

$$a_{j+1}[p] = \sum_{n=-\infty}^{+\infty} l[n-2p]a_j[n] \qquad \text{.............} (1)$$

$$d_{j+1}[p] = \sum_{n=-\infty}^{+\infty} h[n-2p]a_j[n] \qquad \text{..............} (2)$$

Elements aj are used for next step (scale) of the transform and elements dj, called wavelet coefficients, determine output of the transform. l[n] and h[n] are coefficients of low and high-pas filters respectively. It can be assumed that on scale j+1 there is only half from number of a and d elements on scale j. Results that DWT can be done until only two aj elements remain in the analyzed signal. These elements are called scaling function coefficients.

DWT algorithm for two-dimensional pictures is similar. The DWT is performed firstly for all image rows and then for all columns (Figure: 2).

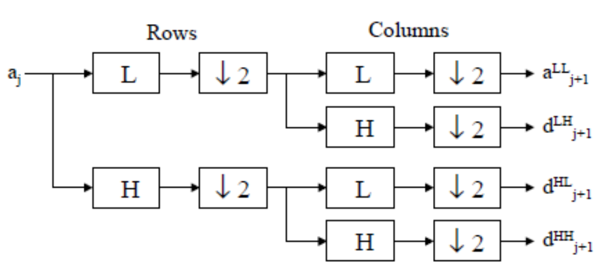


Figure 2: Wavelet decomposition for two-dimensional pictures

The main feature of DWT is multi-scale representation of function. By using the wavelets, given function can be analyzed at various levels of resolution. The DWT is also invertible and can be orthogonal.

II. ALPHA BLENDING

Alpha blending is the process of combining a translucent foreground color with a background color, thereby producing a new blended color. The degree of the foreground color's translucency may range from completely transparent to completely opaque. If the foreground color is completely transparent, the blended color will be the background color. Conversely, if it is completely opaque, the blended color will be the foreground color. The translucency can range between these extremes, in which case the blended color is computed as a weighted average of the foreground and background colors. Fast graph provides alpha blending functions that work on RGB color values, on direct color bitmaps, and on direct color virtual buffers. The alpha blending functions do not work when a 256-color virtual buffer is active.

IV. SCALE-INVARIANT FEATURE TRANSFORM (SIFT) DESCRIPTOR

The Scale-Invariant Feature Transform (SIFT) [15, 16, 17, 18, 19, 20, 25] descriptor aims to achieve robustness to lighting variations and small positional shifts by encoding the image information in a localized set of gradient orientation histograms. The descriptor computation starts from a scale and rotation normalized region extracted with one of the detectors like Hessian detector or the Harris detector [11, 12]. As a first step, the image gradient magnitude and orientation are sampled around the key point location using the region scale to select the level of Gaussian blurs (i.e. the level of the Gaussian pyramid at which, computation is performed). Sampling is performed in a regular grid of 16×16 locations covering the interest region. For each sampled location, the gradient orientation is entered into a coarser 4×4 grid of gradient orientation histograms with 8 orientation bins each, weighted by the corresponding pixel's gradient magnitude and by a circular Gaussian weighting function with a σ of half the region size. The purpose of this Gaussian window is to give higher weights to pixels closer to the middle of the region, which is less affected by positional shifts. The motivation for this choice of representation is that the coarse spatial binning allows for small shifts due to registration errors without overly affecting the descriptor. At the same time, the high-dimensional representation provides enough discriminative power to reliably distinguish a large number of key points. When computing the descriptor, it is important to avoid all boundary effects, both with respect to spatial shifts and to small orientation changes. Thus, when entering a sampled pixel's gradient information into the 3-dimensional spatial/orientation histogram, its contribution should be smoothly distributed among the adjoining histogram bins using trilinear interpolation [21,22]. Once all orientation histogram entries have been completed, those entries are concatenated to form a single 4 × 4 × 8 = 128 dimensional feature vector. Final illumination normalization completes the extraction procedure. For this, the vector is first normalized to unit length, thus adjusting for changing image contrast. Then all feature dimensions are threshold to a maximum value of 0.2 and the vector is again normalized to unit length. This last step compensates for non-linear illumination changes due to camera saturation or similar effects [26].

V. ADVANTAGES OF SIFT OVER HARRIS CORNER DETECTION

While shown to be remarkably robust to image plane rotations, illumination changes, and noise, the locations returned by the Harris Corner detectors are only repeatable up to relatively small-scale changes. The reason for this is that both detectors rely on Gaussian derivatives computed at a certain fixed base scale σ. If the image scale differs too much between the test images, then the extracted structures will also be different. For scale invariant feature extraction, it is thus necessary to detect structures that can be reliably extracted under scale changes [15]. An important aspect of SIFT is that it generates large numbers of features that densely cover the image over the full range of scales and locations. SIFT gives feature stability to noise; it can match features after random change in image scale & orientation, with differing levels of image noise. Also it gives feature stability to affine change [24]; it can match features after image affine distortion also.

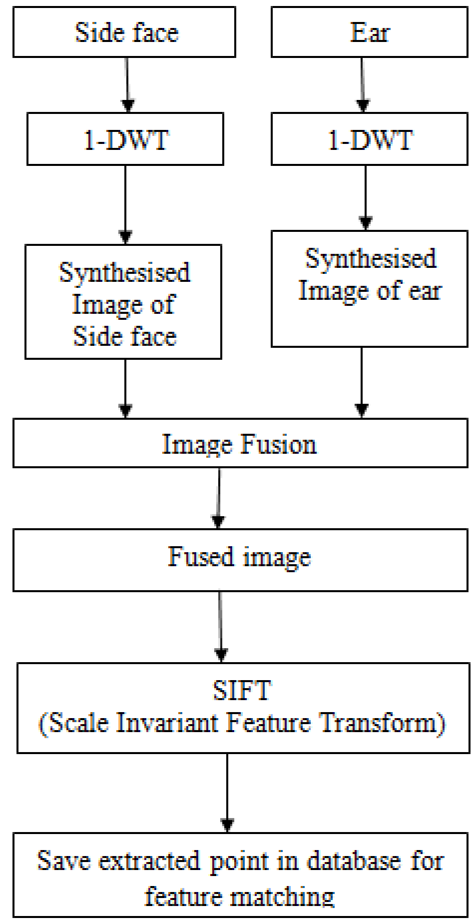
VI. PROPOSED METHOD

Step 1: Ear and side face of the same person has been taken for experiment.
Step 2: Two images (i.e. side face and ear) of same size were analyzed and 1-level Discrete
Wavelet Transform (DWT) was to decompose the pictures.
Step 3: Fuse synthesized images using Alpha-Blending technique.
Step 4: SIFT on the fused images has been performed.
Step 5: Corners has been saved as a feature point for tracking and recognizing objects in the database for matching.

.

VII. FLOWCHART AND EXPERIMENT WITH THE PICTURES OF THE PROPOSED METHOD

The block diagram of the flowchart of the proposed method is described below.



The picture of Ear and Side Face of a person has been taken for experiments in the Figure: 4 and Figure: 5 respectively. Decomposition at level one is given in the Figure: 6 and Figure: 7. Synthesized image of ear and side face is represented in the Figure: 8 and Figure: 9 respectively. Fused image of ear and side face is given in Figure: 10 and SIFT key points o f fused image is given in Figure: 11.

Figure 3: Face

Figure 4: Ear

Figure 5: Side-face

Figure 8: Synthesised image of ear

Figure 9: Synthesised image of side face

Figure 10: Fused image

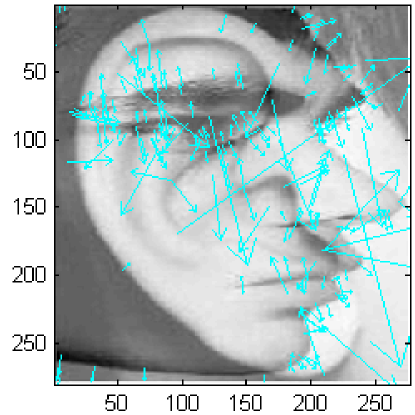


Figure 11: SIFT extracted features

VIII. STEPS OF MATCHING

1. Matching function reads two images, finds their SIFT features, and displays lines connecting the matched Keypoints. A match is accepted only if its distance is less than distance Ratio times the distance to the second closest match. It returns the number of matches displayed.

2. Matching function finds SIFT (Scale Invariant Fourier Transform) Keypoints for each image.

3. Assume some distance ratio for example suppose distance ratio=.4 it means that it only keep matches in which the ratio is less than distance Ratio. 5. Now for each descriptor in the first image, it selects its match to second image.

4. Then it creates a new image showing the two images side by side with lines joining the accepted matches.

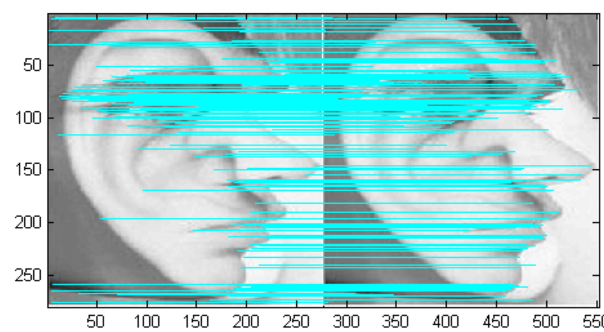


Figure 12: Normal matching

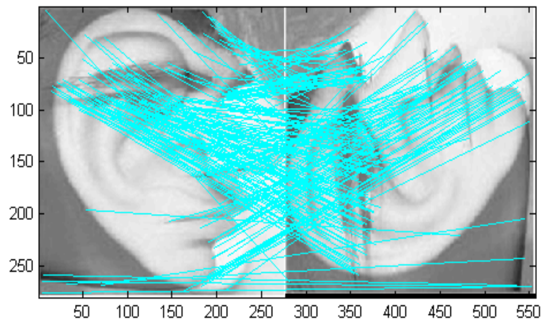


Figure 13: Image rotated by 90 degree

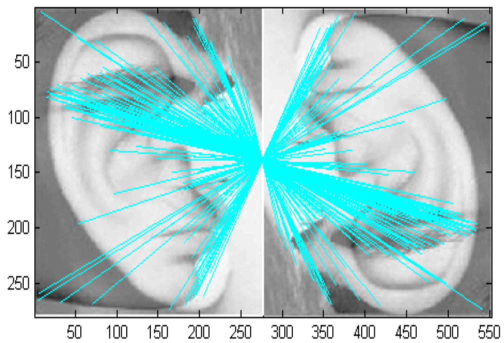


Figure 14: Image rotated by 180 degree

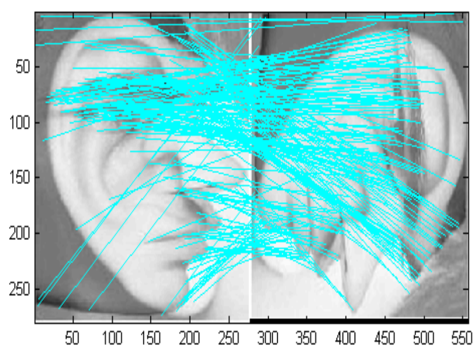


Figure 15 Image rotated by 270 degree

IX. RESULT

Number of key points found in the fused image is 206.

Image Transformation	**Match found**
0 degree	206
90 degree	189
180 degree	187
270 degree	190

Table 1: Result of matching along with the degree of rotation

IX. CONCLUSIVE DISCUSSION

In the proposed approach the advantages and comparisons have been made among the existing techniques and following the most suitable techniques were adopted for the proposed approach. The total number of key points extracted can be saved in the database for identification and for future image processing operations like tracking or recognition of objects.

X. FUTURE SCOPE

Wavelets seem to be effective for analysis of textures recorded with different resolution. It is very important problem in Nuclear Magnetic Resonance (MNR) imaging, because high-resolution images require long time of acquisition. This causes an increase of artefacts caused by patient movements, which should be avoided. There is an expectation that the proposed approach will provide a tool for fast, low resolution NMR medical diagnostic in near future.

REFERENCES

[1]. Hui Chen; Bhanu, B., “Human Ear Recognition in 3D”, Pattern Analysis and Machine Intelligence, IEEE Transactions on Volume:29, Issue: 4, ISSN : 0162-8828, INSPEC Accession Number: 9370992, pp. 718 - 737 , April 2007

[2]. Ping Yan; Bowyer, K.W., “Biometric Recognition Using 3D Ear Shape”, Pattern Analysis and Machine Intelligence, IEEE Transactions on Volume:29 , Issue: 8, ISSN : 0162-8828, INSPEC Accession Number: 9621358, Digital Object Identifier : 10.1109/TPAMI.2007.1067, pp. 1297 –1308, Aug. 2007.

[3]. Monwar, M.; Gavrilova, M.L., “Multimodal Biometric System Using Rank-Level Fusion Approach”, Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on Volume:39 , Issue: 4, ISSN :1083-4419 INSPEC Accession Number:10665721, Digital Object Identifier: 10.1109/TSMCB.2008.2009071, pp. 867 - 878 Aug. 2009.

[4]. De Marsico, M.; Nappi, M.; Riccio, D.; Tortora, G., “NABS: Novel Approaches for Biometric Systems”, Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on Volume:41, Issue: 4, Biometrics Compendium, IEEE, ISSN :1094-6977 pp. 481 - 493 July 2011.

[5]. Grabham, N.J.; Swabey, M.A.; Chambers, P.; Lutman, M.E.; White, N.M.; Chad, J.E.; Beeby, S.P., “An Evaluation of Otoacoustic Emissions as a Biometric “, Information Forensics and Security, IEEE Transactions on Volume: 8, Issue: 1, Biometrics Compendium, IEEE, ISSN: 1556-6013, pp. 174 – 183, November 2012, Issue Date: Jan. 2013

[6]. Cadavid, S.; Abdel-Mottaleb, M., “3-D Ear Modeling and Recognition From Video Sequences Using Shape From Shading”, Information Forensics and Security, IEEE Transactions on Volume: 3, ISSN: 1556-6013, pp. 709 - 718

Dec.

[7]. "The ear as a Biomertic", Hurley and Nixon, EUSIPCO 2007.

[8]. Shekhar Karanwal , Davendra Kumar , Rohit Maurya ,"Fusion of Fingerprint and Face by using DWT and SIFT", International Journal of Computer Applications (0975 – 8887) Volume 2 – No.5, June 2010

[9]. A.K.Jain and A. Ross, "Multibiometric systems", Communications of the ACM, vol. 47, no.1, pp.34 - 40, 2004.

[10]. Akhil Pratap Shing,Agya Mishra,"Wavelet Based Watermarking on Digital Image",Indian Journal of Computer Science and Engineering,Vol 1 No 2, 86-91

[11]. Harris, C., Stephens, M., 1988, A Combined Corner and Edge Detector, Proceedings of 4th AlveyVision Conference.

[12]. KonstantinosG. Derpanis, 2004, The Harris Corner Detector.

[13]. Nilanjan Dey, Subhendu Das, Pranati Rakshit, "A Novel Approach of Obtaining Features Using Wavelet Based Image Fusion and Harris Corner Detection", International Journal of Modern Engineering Research , Vol.1, Issue.2, pp-396-399.

[14]. S. Mallat: A wavelet Tour of Signal Processing, Academic Press, San Diego 1998.

[15]. D. G. Lowe, "Distinctive image features from scale invariant Keypoints", International Journal of Computer Vision, vol. 60, no. 2, 2004.

[16]. Brown, M. and Lowe, D.G. 2002. Invariant features from interest point groups. In British Machine Vision Conference, Cardiff, Wales, pp. 656-665.

[17]. Crowley, J. L. and Parker, A.C. 1984. A representation for shape based on peaks and ridges in the difference of low-pass transform. IEEE Trans. on Pattern Analysis and Machine Intelligence, 6(2): 156-170.

[18]. Fergus, R., Perona, P., and Zisserman, A. 2003. Object class recognition by unsupervised scaleinvariant learning. In IEEE Conference on Computer Vision and Pattern Recognition, Madison, Wisconsin, pp. 264-271.

[19]. Funt, B.V. and Finlayson, G.D. 1995. Color constant color indexing. IEEE Trans. on Pattern Analysis and Machine Intelligence, 17(5): 522-529.

[20]. Hough, P.V.C. 1962. Method and means for recognizing complex patterns. U.S. Patent 3069654.

[21]. Lowe, D.G. 1991. Fitting parameterized 3- dimensional models to images. IEEE Trans. on Pattern Analysis and Machine Intelligence, 13(5): 441-450.

[22]. Lowe, D.G. 2001. Local feature view clustering for 3D object recognition. IEEE Conference on Computer Vision and Pattern Recognition, Kauai, Hawaii, pp. 682-688.

[23]. Lowe, D.G. 1999. Object recognition from local scale-invariant features. In International Conference on Computer Vision, Corfu, Greece, pp. 1150-1157.

[24]. Mikolajczyk, K. 2002. Detection of local features invariant to affine transformations, Ph.D. thesis,Institut National Polytechnique de Grenoble, France.

[25]. Torr, P. 1995. Motion Segmentation and Outlier Detection, Ph.D. Thesis, Dept. of Engineering Science, University of Oxford, UK.

[26]. Schmid, C., andMohr, R. 1997. Local grayvalue invariants for image retrieval. IEEE Trans. on Pattern Analysis and Machine Intelligence, 19(5): 530-534.

[27]. The Imposter, a 2012 British-American documentary film by Bart Layton.

Authors Profile

Dr. Subhranil Som received his Master degree in Computer Application in 2003. His PhD in Computer Science and Engineering Technology from University of Kalyani, West Bengal, India in the year of 2012. He is an empanelled PhD supervisor in the area of technology in the West Bengal University of Technology. He is working as a Principal Investigator of an UGC funded project. He holds a distinction in Physics and Mathematics in Graduation. His fields of interest include Cryptography and Network Security, Robotics, Core Java, C++, C. He is currently Asst. Professor in the Department of Computer Application, JIS College of Engineering, and West Bengal, India. He was attached with a WHO's International Research Project on "e-Health for Health Care Delivery", University of New South Wales, Sydney, Australia. He has finished several courses related to computer Application, object oriented analysis and design, Software Engineering and Project Management. He has more than 8 years teaching and research experience

Ms. Jhelum Das pursuing Master degree in Computer Application 2012-2015 in JIS College of Engineering under the West Bengal University of Technology. She has participated in several technical workshop, seminar

Mr. Akash Pal pursuing Bachelor in Technology degree in Computer Science & Engineering 2011-2015 in JIS College of Engineering under the West Bengal University of Technology. He has participated in several technical workshop, seminar etc. He also has some conference papers and journals.

Mr. Aritra Dey pursuing Bachelor in Technology degree in Electrical Engineering 2012-2016 in JIS College of Engineering under the West Bengal University of Technology. He has participated in several technical workshop, seminar etc. He also has some conference papers and journals.