Occlusion Detection and Index-based Ear Recognition

Madeena Sultana, Padma Polash Paul, Marina Gavrilova
Dept. of Computer Science
University of Calgary
2500 University DR NW
T2N 1N4, Calgary, AB, Canada
{msdeena, pppaul, mgavrilo}@ucalgary.ca

ABSTRACT

Person recognition using ear biometric has received significant interest in recent years due to its highly discriminative nature, permanence over time, non-intrusiveness, and easy acquisition process. However, in a real-world scenario, ear image is often partially or fully occluded by hair, earrings, headphones, scarf, and other objects. Moreover, such occlusions may occur during identification process resulting in a dramatic decline of the recognition performance. Therefore, a reliable ear recognition system should be equipped with an automated detection of the presence of occlusions in order to avoid miss-classifications. In this paper, we proposed an efficient ear recognition approach, which is capable of detecting the presence of occlusions and recognizing partial ear samples by adaptively selecting appropriate features indices. The proposed method has been evaluated on a large publicly available database containing wide variations of real occlusions. The experimental results confirm that the prior detection of occlusion and the novel selection procedure for feature indices significantly improve the biometric system recognition accuracy.

Keywords

Biometric images, occlusion detection, ear recognition, partial occlusion, classifier selection, adaptive feature selection.

1. INTRODUCTION

Biometric authentication offers advantage over traditional PIN (Personal Identification Number) or password-based security since it is harder to forge, steal, transfer, or lose biometric data. At present, biometric based person recognition has enormous demand in government services as well as commercial sectors due to availability of biometric data, enhanced recognition accuracy and non-invasive nature of authentication. Over the last few years, ear biometric has received growing attention and proven to be useful for an automated person recognition [Cha03a], [Che07a]. Unlike face biometrics, ear has no sensitivity to facial expression changes [Kum12a] and it remains almost unchanged throughout the lifetime of a person [Yua12a]. Ear biometric is not only a powerful feature to identify individuals, but also to recognize identical twins [Nej12a]. Moreover, ear has high user acceptance because of its nonintrusive nature and a passive acquisition process [Jai99a]. Similar to other passive biometrics, the recognition performance of ear biometrics may deteriorate significantly due to natural constraints such as occlusion, lightning, pose difference etc. [Bus10a]. Among all natural constraints, occlusion happens to be the most common scenario, since ear is often partially or fully occluded by hair, earrings, headphones, scarfs etc. Information loss due to occlusion is irrevocable. Unlike lightning or pose variations, where some image enhancement techniques can be applied to retrieve partially lost information, the occlusion information loss results in a complete disappearance of a portion of ear. Moreover, distortions of important global features of ear biometrics such as shape and appearance occur, which further undermine the overall system recognition performance. For those reasons, occlusion is one of the most detrimental degrading factors of ear recognition. It has been reported that consideration of un-occluded regions during matching increases recognition accuracy [Yua12a], [Yua12b]. In order to determine the un-occluded portion of ear it is necessary to detect occluded regions. However, detection of occlusions in ear biometrics remained understudied at present. Detection of real occlusions is a very challenging problem since occurrence, locations, proportion, and reasons of occlusion are uncertain. For instance, different regions of an ear may be occluded by different objects such as hair or earrings at the same time. Also, during identification stage, an ear sample may be occluded partially or fully, or may not be occluded at all. In addition, determining the proportion of occlusion is important to make a decision whether the sample is sufficient for recognition process or needs to be reacquired. Last but

not the least, not every method performs equally on different proportions and different regions of occlusions. For example, global features such as shape-based descriptor may perform well in cases of partial occlusion by earrings, while local or block-based features may work better on distorted shapes due to occlusions by hair. Thus, prior detection of the location and proportion of occlusion could help in selecting the appropriate features as well as feature extraction methods. For these reasons, it is important to develop an ear recognition method that is capable of prior detection of occlusion and can select appropriate features for classification at identification stage. In a recent review paper, Pflug and Busch [Pfl12a] pointed out the lack of studies on real-world ear occlusions. This paper fills this niche and provides a solution to this problem by investigating how real occlusion factors such as hair, accessories etc. affects the recognition performance. The novel contributions of this paper are three fold:

1. We propose a novel method for ear occlusion detection and estimation of occlusion degree using skin-color model.
2. We analyze the impact of real ear occlusions (hair and accessories) on recognition performance.
3. We propose a novel index-based partial ear recognition method that utilizes occlusion information adaptively to obtain consistent recognition rate.

The rest of the paper is organized as follows. Section 2 summarizes some existing researches on ear recognitions. The proposed methodology for occlusion detection and ear recognition is described in Section 3. Section 4 demonstrates experimental results of the performance and effectiveness of the proposed method. Finally, concluding remarks and future works are presented in Section 5.

2. RELEVANT WORK

Person identification using ear biometric has drawn significant attention of many researchers over the last decade. Ear biometric has the advantage of a non-intrusive acquisition in a less controlled environment. However, there has always been a tradeoff between the non-invasiveness of image acquisition and its impact on its quality. Restricting the acquisition environment of ear biometric compromises its noninvasive nature and wide acceptance of users. Moreover, noninvasive biometrics are mostly acquired by surveillance cameras, where environment cannot be controlled. Therefore, instead of imposing tight controls on the acquisition environment, the recent research is focused on developing robust biometric systems that can obtain high recognition rates under less than ideal conditions. Occlusion has been studied for face biometrics to some extent [Lin07a], [Taj13a]. However, occlusion conditions, type, area, proportion etc. of ear are very different than face. There is a lack of study on real occlusions of ear biometrics during identification stage. In this section, we will discuss some contemporary ear recognition methods.

In 2010, Bustard and Nixon [Bus10a] proposed a robust method for ear recognition using homographies calculated from the Scale Invariant Feature Transform (SIFT) points. Authors also showed that performance of this method degraded with an increasing proportion of occlusions. However, the method did not include an automated occlusion detection as well as proportion calculation. Experimentation was conducted on simulated occluded conditions and the effect of real-world occlusions remained uninvestigated. Efficient feature extraction of an ear biometric has been investigated in many recent works. For instance, Huang et al. [Hua11a] proposed Uncorrelated Local Fisher Discriminant Analysis (ULFDA) method for ear recognition, which obtained better performance than benchmark Principle Component Analysis (PCA) and Linear Discriminant Analysis (LDA) [Mar01a]. In 2012, Kumar and Wu [Kum12a] proposed an ear recognition method based on gray-level shape features which outperformed Gabor and log-Gabor based methods. Sparse representation of local texture has been proposed by Kumar and Chan [Kum13a] in 2013, which obtained high recognition rates on different databases. However, none of the aforementioned three methods was evaluated under occluded conditions.

Occlusion has been considered by Yuan and Mu [Yua12a], where a local information fusion method was proposed to obtain robustness under partial occlusion. In this work, experimentation was conducted in the simulated occluded condition, i.e. a specific amount of occlusion has been applied artificially to a certain location of the ear images. However, results showed that the recognition performance of this method varied according to the location as well as amount of occlusion. In another work, Yuan et al. [Yua12b] proposed a sparse based method to recognize partially occluded ears. Experimentation was conducted by adding synthetic occluded regions to the original unoccluded images. Results showed that this method obtained 70% recognition rate for 30% occluded regions, whereas performance dropped below 15% with the increase of the occluded portion up to 50%. In another recent work, Morales et al. [Mor13a] showed that performance of SIFT and Dense-SIFT based feature extraction methods also degraded significantly due to the presence of real-world occlusions. In their work, recognition error rate of SIFT and Dense-SIFT features were 2.78% and 2.03%, respectively on IITD Database. However, the corresponding error rates increased to 20.52% and 25.76% under real-world occlusions on West Pomeranian University of Technology Ear Database [WPUTED]. The above

discussion demonstrates that existing ear recognition methods lack the following:

1. Detection of occlusion remained uninvestigated, although recognition rate highly depends on the presence of occlusion.
2. Recognition rate varies with location and proportion of occlusion. There is a lack of study on automated localization and proportion calculation of occlusion.
3. Existing methods are mostly experimented on simulated or synthetically occluded ear samples in predefined locations. The robustness of ear recognition methods need to be evaluated under real occlusions since type, location, and proportion of real-world occlusions might be very different than the simulated cases.

The above points indicate that there is a gap between real-world occlusion detection and occluded ear recognition methods. In this paper, our main goal is to bridge the gap between occlusion detection and occluded ear recognition by proposing a novel ear recognition method that can detect real occlusions and utilize occlusion information adaptively during recognition stage.

3. PROPOSED METHOD

In this paper, we presented an automated approach of occlusion detection, estimation, and un-occluded region extraction. We also proposed a novel index-based ear recognition method, which can efficiently utilize the extracted un-occluded portion of ear. In the real scenarios, enrolled or template images are mostly obtained under human supervision. Therefore, if occlusion occurs, human supervisor can direct the person to reacquire the sample. On the other hand, identification stage is mostly unsupervised and the system process occluded image in case of the absence of automated detection mechanism, which may eventually lead to a false match. This is why we were interested in measuring occlusion during identification stage. A basic flow diagram of the proposed system is shown in Fig. 1. During enrollment index-based features are extracted and stored in feature database along with corresponding indices. During test, occluded and un-occluded portion of the ear are detected automatically. Next, index-based features are extracted from un-occluded portion of test ear sample and similarity is measured with the corresponding features of enrolled images. The final decision has been obtained from the maximum similarity matching score of the test and enrolled samples. Detailed explanation of the proposed method can be found in the following subsections.

3.1 Types of Occlusion

Occlusion in ear images may occur anytime during identification stage due to the presence of hair, scarf/hat, earring, headphones, dust, and so on. Both shape and appearance of ear vary in a very different way based on the type, location, and proportion of occlusions.

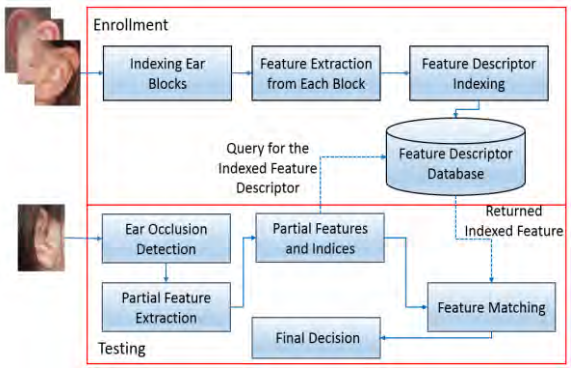

Figure 1. Flow diagram of the proposed ear recognition system.

One reason for the lack of investigation on real occlusions is the unavailability of public database. Researchers [Fre10a] of West Pomeranian University of Technology have created an ear database containing ear samples with different types of real occlusions to facilitate proper validation of ear recognition algorithms. Fig. 2 shows different occluded conditions of ear samples from West Pomeranian University of Technology Ear Database. From Fig. 2, one can see that location, type, and proportion of occlusion are very uncertain and cannot be predefined. Therefore, proper detection of occlusion is indispensable to extract un-occluded features from ear.

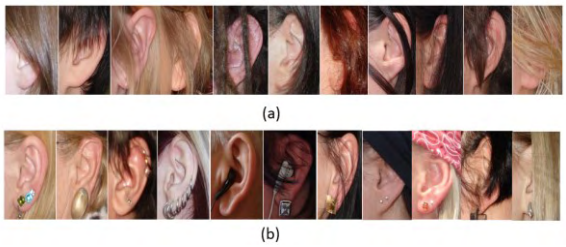

Figure 2. Different types of ear occlusions; a) occlusions by hair; b) occlusions by earrings, scarf, hat, headphones, etc. [WPUTED]

3.2 Ear Enrollment

In this paper, occlusion is considered during test phase to resemble the identification stage. Generally, enrollment is accomplished under human supervision. If occlusion occurs during enrollment human supervisor can reject the biometric sample and reacquire it. Therefore, in this paper, we considered the case that enrolled images are not occluded. Initially, all enrolled images are preprocessed using histogram equalization method and downsampled to 100×80 pixels. Each enrolled image is then partitioned into 10×10 blocks, total 80 blocks. Fig. 3 shows a visual representation of partitioning an ear image into blocks. Next, we applied two-dimensional (2D) Haar

Discrete Wavelet Transform (DWT) to extract local texture features [Sul14a] from each block. Haar wavelet transform decomposes an input block into four sub-bands, one low frequency component (LL) and three detail components (LH, HL, HH). Decomposition to low frequency subband (LL) smoothens image thus reduces noise. Decimated DWT is a popular mathematical tool for image compression since it efficiently reduces image dimensionality at different levels, whereas ensuring seamless reconstructions [DeV92a]. Thus, DWT preserves important information of image while discarding dimensionality. Moreover, DWT is computationally efficient and less sensitive to illumination changes [Sul14a]. The low frequency subband (LL) of DWT contains most of the information of an image. In this work, we applied 1st level Haar DWT to all blocks and considered the low frequency subband of each block as local features. The features of each block are then stored along with its index in feature database.

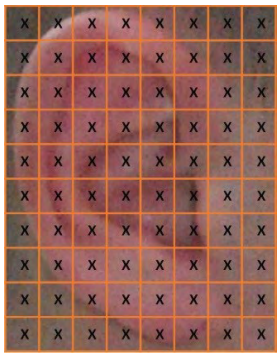

Figure 3. An example of partitioning enrolled ear into indexed-blocks.

3.3 Ear Occlusion Detection

Real-world occlusion detection is a very challenging task because it is uncertain that when and what type of occlusion would arise. There is also no certainty in which portion and what proportion the occlusion would occur. In this section, we propose a novel method of ear occlusion detection and estimation using skin color model. The process in outlined in Algorithm 1.

In our method, the skin color regression model [Pau10a] has been applied for occlusion detection. We utilized skin color model for ear occlusion detection because occlusion obscures skin color information and detection of skin color will allow us to separate occluded and un-occluded regions in ear. The proposed occlusion detection method has four steps: 1) conversion to chromatic color space *r* and *g*, 2) detection of skin regions in *r* and *g* color spaces using skin color likelihood (eq. 5), 3) fusion of *r* and *g* color space images and fill skin regions using morphological operation, and 4) masking un-occluded skin portion from original occluded image. A flow diagram of the steps is depicted in Fig. 5.

Algorithm 1: Occlusion detection and estimation.

Input: Test ear image Y of size $M \times N$.

Output: (B_{Ij}, I_j), un-occluded blocks in Y and corresponding indices.

Step 1: Preprocess Y using histogram equalization method and downsample $M \times N$ to 100×80.

Step 2: Transform image from RGB color space to chromatic color space. Find the value of *r* and *g* as follows [Pau10a]:

$$r = \frac{R}{R+G+B} \quad (1)$$

$$g = \frac{G}{R+G+B} \quad (2)$$

Step 3: Find skin color distribution by 2-D Gaussian model with the following mean vector *A* and covariance matrix *C* [Pau10a]:

$$A = G\{x\}[x = (rg)^T] \quad (3)$$

$$C = \begin{bmatrix} \sigma_{rr} & \sigma_{rg} \\ \sigma_{gr} & \sigma_{gg} \end{bmatrix} \quad (4)$$

Step 4: Estimate likelihood (L) of skin color using the following equation [Pau10a]:

$$L = P(r,g) = \exp[-0.5\,(x-A)^T\, C^{-1}(x-A)\,] \quad (5)$$

where, $x = (r,g)^T$.

Step 5: Find the skin color regions of Y in chromatic color *r* and *g*, denoted as *P(r)* and *P(g)*.

Step 6: Fuse *P(r)* and *P(g)* to obtain resultant binary image, Z = *P(r) AND P(g)*

Step 7: Perform morphological operation using disk shape structuring element of radius 10 to fill the skin regions in Z.

Step 8: Apply Z as a mask on Y to obtain image X containing un-occluded skin portions.

Step 9: Partition X into 10×8 blocks each having 10×10 pixels and construct a block vector $\{B_i | i = 1, 2, \ldots, 80\}$.

Step 10: Construct an index vector $\{I_j | i = 1, 2, \ldots, m\}$, where B_{Ij} contains skin regions (un-occluded) and m is the total number of un-occluded blocks.

Step 11: Estimate total proportion of occlusion, $E = \frac{\sum_{j=1}^{m} B_{Ij}}{\sum_{i=1}^{80} B_i} \times 100$ (6)

Step 12: If E>60%, discard Y and reacquire test image.

Fig. 6 presents some outcomes of occlusion detection of four ear samples from WPUT Ear Database. Fig. 6 (a) shows four original ear samples containing different types of occlusions due to earring,

headphones, and hair. The corresponding ear samples after chromatic color space conversion are shown in Fig. 6 (b). Fig. 6 (c) presents the resultant skin-regions separated from occlusions. After separating occluded and un-occluded regions, the ear image is partitioned into 80 blocks, each containing 10×10 pixels. The estimated occlusion has been calculated as the ratio of the number of un-occluded blocks over total number of blocks (eq. 6). The estimation of occlusion facilitates auto-rejection of unreliable test images, where most of the information is distorted due to occlusion. In the proposed method, if the estimated occlusion is below 60%, the test image will be used for recognition, otherwise it has to be reacquired. In this way, the proposed ear recognition system can reduce false matches by discarding unreliable test samples, automatically.

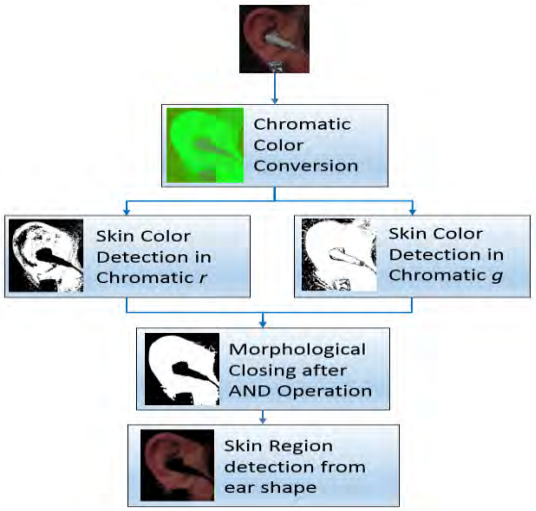

Figure 5. Flow diagram of the four steps of occlusion detection using skin color model.

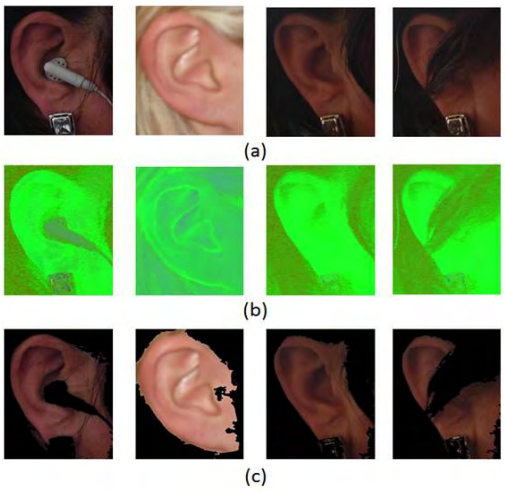

Figure 6. Examples of ear occlusion detection: a) original occluded ears, b) conversion to chromatic color space, c) detected unoccluded skin-regions.

3.4 Partial Feature Extraction and Matching

In the proposed method, partial features are extracted from the detected un-occluded blocks of the test image. 1st level of Haar DWT is applied to all un-occluded blocks (B_{ij}), and four subbands images (LL, LH, HL, HH) are obtained. The low frequency subband (LL) of each block is considered as the local features of corresponding block and converted to a feature vector.

Finally, similarities between the partial features of test ear and corresponding features of enrolled ears are measured for recognition. Fig. 7 shows an example of the corresponding blocks of a test ear and an enrolled ear. The left image in Fig. 7 shows the blocks of skin regions in test image and the right image shows corresponding blocks in enrolled image. Unlike existing methods, we matched the un-occluded blocks of the detected skin regions to the corresponding blocks of enrolled ears. The index vector (I_j) is used to fetch the corresponding blocks of enrolled ears from feature database. A visual representation of the partial feature extraction and similarity matching process is shown in Fig. 8.

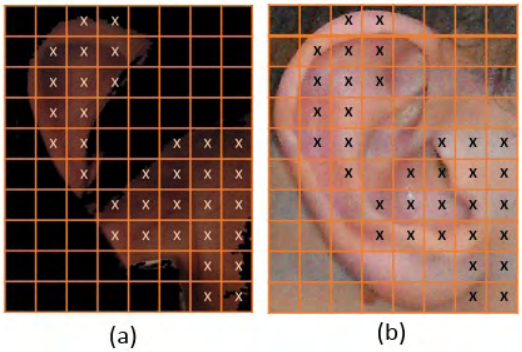

Figure 7. Block indexing, a) unoccluded blocks of test ear, b) corresponding blocks of enrolled ear.

The similarities of the test and enrolled ears are computed using Euclidean distance. Euclidean distance of the indexed features of test ear (V) and enrolled ear (U) can be calculated using the following equation:

$$D = \sqrt{\sum_{j=1}^{m}\sum_{k=1}^{n}\left(u_{jk} - v_{jk}\right)^2} \qquad (7)$$

where u_{jk} and v_{jk} are the kth feature of jth block of U and V, respectively, and m is the total number of un-occluded blocks in V. However, a problem may arise during the index-based matching if the indexed blocks of the test ear do not overlap with the indexed blocks of enrolled ear (in other words, if the test ear is shifted to any direction). There are eight possible directions of shift, which is shown in Fig. 9. We propose to solve this problem by using a matching window in all possible eight directions: B_1, B_2, B_3, B_4, B_5, B_6, B_7, B_8.

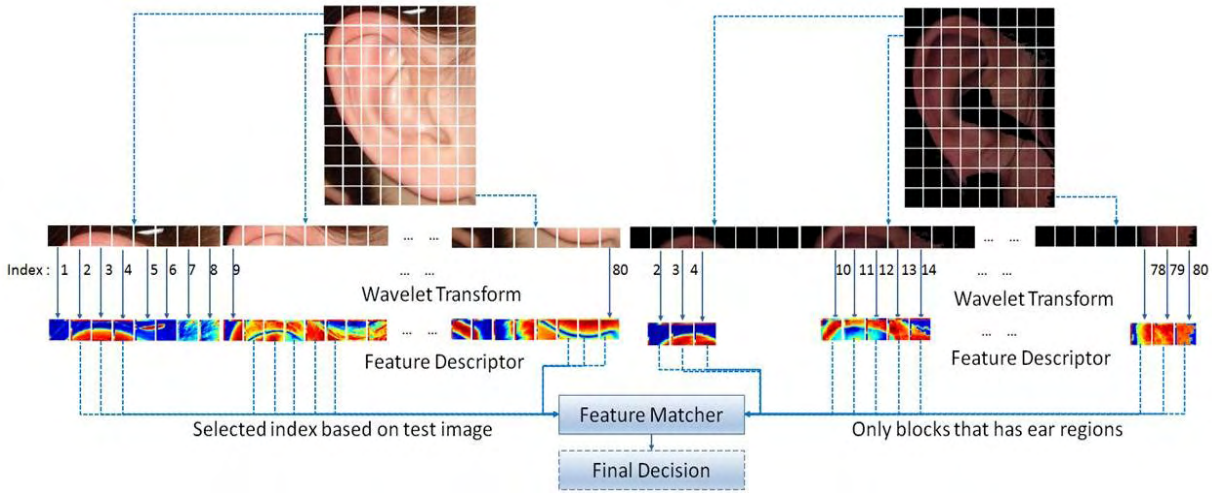

Figure 8. Feature extraction and index-based partial feature matching of test and enrolled ears.

The set of all un-occluded blocks of the test ear is considered as one region. Let us consider $T_{i,j}$ as the un-occluded region of test image and $R_{i,j}$ as the corresponding region in the enrolled sample. The nine similarity score are then calculated using eq. 8 to eq. 16.

$$S_0 = 1 - D(T_{i,j}, R_{i,j}) \quad (8)$$

$$S_1 = 1 - D(T, R_{i,j+1}) \quad (9)$$

$$S_2 = 1 - D(T_{i,j}, R_{i+1,j+1}) \quad (10)$$

$$S_3 = 1 - D(T_{i,j}, R_{i+1,j}) \quad (11)$$

$$S_4 = 1 - D(T_{i,j}, R_{i+1,j-1}) \quad (12)$$

$$S_5 = 1 - D(T_{i,j}, R_{i,j-1}) \quad (13)$$

$$S_6 = 1 - D(T_{i,j}, R_{i-1,j-1}) \quad (14)$$

$$S_7 = 1 - D(T_{i,j}, R_{i-1,j}) \quad (15)$$

$$S_8 = 1 - D(T_{i,j}, R_{i-1,j+1}) \quad (16)$$

The first similarity (S_0) score between the test and enrolled sample is calculated by matching the blocks of test region $T_{i,j}$ and corresponding enrolled region $R_{i,j}$. Next, we calculated the similarity score S_1 along B_1 direction between the test region T_{ij} and training region $R_{i,j+1}$. Then similarity score, S_2 is calculated along B_2 direction between $T_{i,j}$ and $R_{i+1,j+1}$. Similarly, similarity scores S_3 to S_8 are calculated along directions B_3 to B_8 using eq. 11 to eq. 16. The reason for calculating nine similarity scores is that if the test sample is shifted to any of the possible directions, matching score along that direction will be the highest. Thus, calculating similarity scores in all possible directions allow us to find the best matching indices even under shifted condition. Fig. 10 shows pictorial representation of the calculation of nine similarity matching scores from S_0 to S_8. In Fig. 10, S_0 (in middle) represents an example of the corresponding blocks of an enrolled image. The shifted blocks in eights possible directions are represented by S_1 to S_8 in Fig. 10. The shifted blocks were calculated by shifting the whole region (all blocks) towards the eight possible directions as shown in Fig. 9.

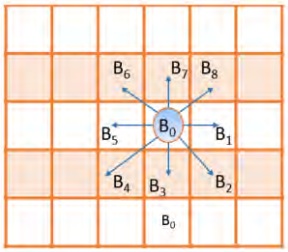

Figure 9. Possible eight directions of image shift.

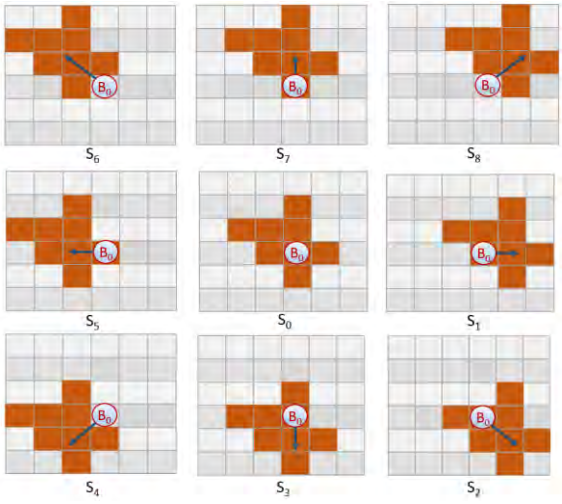

Figure 10. Nine similarity scores (S_0- S_8) calculation by shifting the indexed region of enrolled ear along different directions.

The best matching score is calculated in two ways. First, the highest value among the nine scores is considered as the overall maximum score, S_m (eq. 17). Secondly, we calculated the block-wise maximum score (S_B) among the nine similarity vectors. Calculation of S_B can be shown as eq. 18:

$$S_m = \max_{0 \le i \le 8} S_i \quad (17)$$

$$S_B = \max_{0 \le i \le 8, 1 \le j \le m,} S_{ij} \quad (18)$$

where S_{ij} is the similarity score of jth block of ith similarity vector, m is the maximum number of un-occluded blocks in test image.

4. EXPERIMENTAL RESULTS

Three sets of experiments were conducted to evaluate the performance of the proposed ear occlusion and ear recognition method. All experiments were carried out on Windows 7 operating system, 2.7 GHz Quad-Core Intel Core i7 processor with 16GB RAM. Matlab version R2013a was used for implementation and experimentation of the proposed method. We evaluated our method on WPUT Ear Database [WPUTED] since this is the largest publicly available database containing ear images with wide variations of real occlusions. A brief description of the database is as follows:

WPUT Ear Database [Fre10a]: This database contains 2071 ear images of 254 women and 247 men, total 501 individuals of different ages. There are at least two images per ear of each subject. 15.6% of the images were taken outside and some of them were taken in the dark. 80% of the images are recorded as deformed due to the presence of real occlusions. Ear images of 166 subjects are covered by hair and the presence of earrings are recorded for 147 subjects. The other forms of real-world occlusion in this database are glasses, headdresses, noticeable dirt, dust, birth-marks, ear-pads etc. Many of the samples are simultaneously occluded by different types of occlusion in different proportions.

For our experimentation, the whole database is partition into training and test sets. The training database is created using comparatively un-occluded ear samples. We have single training sample per subject. The occluded images are randomly selected for testing. Each experiment is performed five times and the average recognition accuracy is considered as the recognition performance of the proposed method. Identification rate of the proposed method is analyzed by plotting Cumulative Match Characteristics (CMC) curve. CMC curve is the cumulative probability of obtaining the correct match in the top r positions (ranks). The final matching scores of the test and enrolled images can be obtained in different ways such as block-wise maximum score, overall maximum score, block-wise average score, and overall average score. Therefore, in the first experiment, we compared the performance of the proposed method using different similarity scores to obtain the best performing method of calculating the final similarity score. Fig. 11 shows the CMC curves of the proposed method using block-wise maximum similarity, overall highest similarity, block-wise average score, and overall average similarity scores. From Fig. 11, we can see that the highest performance of the proposed method was obtained by using block-wise maximum similarity score. Consideration of the highest score among the nine scores obtained the 2nd highest performance. Fig. 11 also shows that block-wise average scores performed better than overall average score. However, consideration of the maximum scores are more discriminative than consideration of average scores. The reason for this that not all the similarity scores will find the best match among the test and training blocks and averaging all scores may fade away the best match. Fig. 11 shows that correct matching probabilities of the block-wise maximum similarity, overall maximum similarity, block-wise average similarity, and overall average similarity at rank 1 are 73%, 65%, 57%, and 51%, respectively. Therefore, from the first set of experiments, we found that block-wise maximum score obtained the best results for the proposed method.

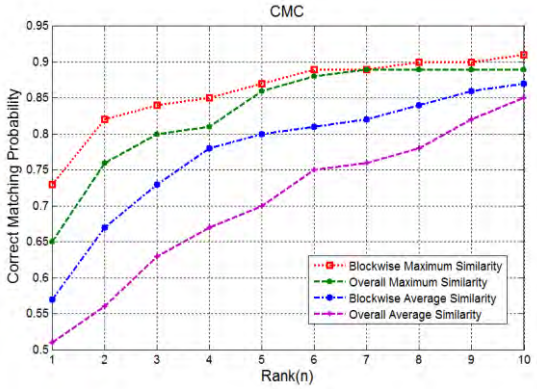

Figure 11. CMC curves of the proposed method using different similarity scores.

In second set of experiments, we compared the performance of the proposed method with a baseline Haar discrete wavelet transform-based method. For the baseline method, ear features were extracted using Haar discrete wavelet transform from test sample without applying any occlusion detection mechanism and the features were matched with the enrolled samples regardless of indices. The CMC curves for the proposed method and the baseline methods are plotted in Fig. 12. From Fig. 12, we can see that the rank 1 recognition rate for the proposed method is 73%, whereas for the baseline method obtained 60% recognition accuracy. Also, 91% recognition rate was obtained by the proposed method within rank 10. The CMC curves in Fig. 12 demonstrate the effectiveness of prior occlusion detection and index-based matching of occluded features.

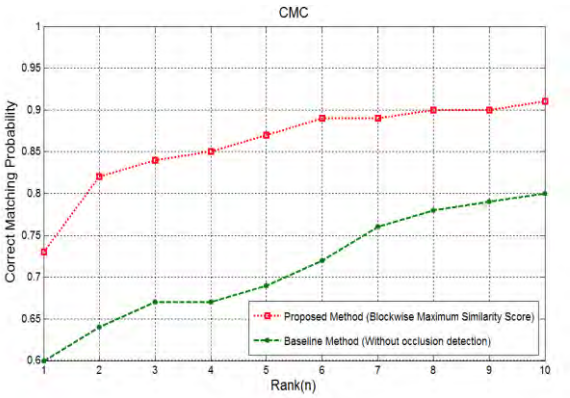

Figure 12. Recognition performance improvement by the proposed method on WPUT database.

In the final set of experiment, we evaluated performance of the proposed method on different amount of occlusions. Amount of occlusion is estimated as the ratio of the occluded blocks over total number of blocks in an ear sample. Fig. 13 shows how the performance of the proposed method varied with different proportion of occlusion. From Fig. 13, we can see that the proposed method can obtain recognition rate as high as 85% with 10% estimated occlusion. Also, the recognition performance remained nearly 80% under 30% occlusion, which is a better result than reported by previous studies. The performance of the proposed method was at 67% even with the 50% of ear image occluded! However, there is simply not enough features for high precision of ear recognition when over 60% of an ear is occluded and in this case ear sample needs to be reacquired.

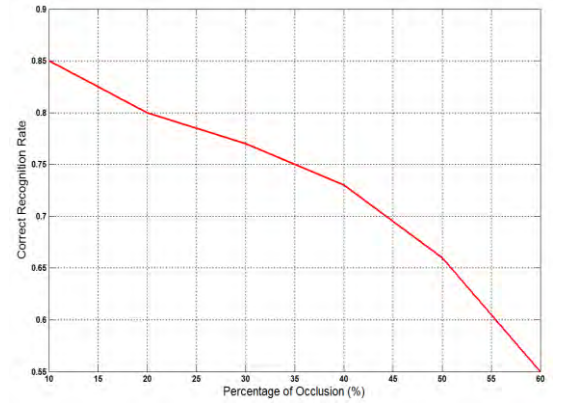

Figure 13. Performance of the proposed method with different degrees of occlusions.

From the above experiments, we can summarize the performance improvement of the proposed method as follows. First of all, automated occlusion detection and estimation allowed us to decide upon whether an ear sample is good enough to be recognized or it is needed to be reacquired. In this way, the proposed method can improve recognition rate by reducing false matches of overly occluded images. Secondly, unlike existing methods where occluded regions were predefined, localization of unoccluded portion in our method is automated. Therefore, the system can adaptively decide upon which portion of the image is unoccluded and good for feature extraction. Thirdly, during recognition, features are extracted from only unoccluded portion of the ear image and matched with corresponding portion of the enrolled samples, which reduces the probability of unreliable matching of the occluded portion. Finally, the problem of shifted indices is solved by using the best block-wise matching scores in eight different scores. For these reasons, the proposed method is capable of obtaining a reliable recognition performance under real occlusions of ears during identification stage.

5. CONCLUSION

A completely automated approach to ear occlusion detection and estimation using skin color model has been proposed in this paper. We also proposed a novel index-based ear recognition method to recognize partially occluded ears effectively. The most important advantage of the proposed method is it can estimate occlusion on ear samples during identification stage and adaptively use this information to select proper indices of the features for recognition purpose. There is a scarcity of occluded ear samples in biometric community and only few publicly available databases contain occluded ear samples. However, the adaptive decision making process of the proposed method doesn't depend on any learning or training of occlusions and thus can be applied to any database. The proposed method of handling ear occlusion was proved to be a very effective in the real world scenarios. Our experiments on real occluded ear images validated the effectiveness of occlusion detection and index-based feature matching for partial ear recognition. Future research will look into incorporating weights into an occlusion estimation process to improve the recognition even further.

6. ACKNOWLEDGMENTS

The authors would like to thank NSERC DISCOVERY program grant RT731064, URGC, NSERC ENGAGE, NSERC Vanier CGS, and Alberta Ingenuity for partial support of this project.

7. REFERENCES

[Bus10a] Bustard, John D., and Nixon, Mark S. Toward unconstrained ear recognition from two-dimensional images. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, 40, No. 3, pp. 486-494, 2010.

[Cha03a] Chang, K., Bowyer, K. W., Sarkar, S., & Victor, B. (2003). Comparison and combination of ear and face images in appearance-based biometrics. IEEE Transactions on Pattern Analysis

and Machine Intelligence, 25, No. 9, pp. 1160-1165, 2003.

[Che07a] Chen, H., & Bhanu, B. (2007). Human ear recognition in 3D. IEEE Transactions on Pattern Analysis and Machine Intelligence, 29, No. 4, pp. 718- 737, 2007.

[DeV92a] De Vore, R. A., Jawerth, B., & Lucier, B. J. (1992). Image compression through wavelet transform coding. IEEE Transactions on Information Theory, 38, No. 2, pp. 719-746.

[Fre10a] Frejlichowski, D., and Tyszkiewicz, N. (2010). The West Pomeranian University of Technology ear database–a tool for testing biometric algorithms. In Image analysis and recognition, pp. 227-234, 2010. Springer Berlin Heidelberg.

[Hua11a] Huang, H., Liu, J., Feng, H., & He, T. (2011). Ear recognition based on uncorrelated local Fisher discriminant analysis. Neurocomputing, 74, No. 17, pp. 3103-3113, 2011.

[Jai99a] Jain, Anil K., Bolle, R., and Pankanti, S. eds. Biometrics: personal identification in networked society. Springer Science & Business Media, 1999.

[Kum12a] Kumar, A., and Wu, C., Automated human identification using ear imaging. Pattern Recognition, 45, No. 3, pp. 956-968, 2012.

[Kum13a] Kumar, A., and Chan, T. S. T., Robust ear identification using sparse representation of local texture descriptors. Pattern Recognition, 46, No. 1, pp. 73-85, 2013.

[Lin07a] Lin, D., & Tang, X. Quality-driven face occlusion detection and recovery. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'07), pp. 1-7, 2007.

[Mar01a] Martínez, A. M., & Kak, A. C. (2001). PCA versus LDA. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23, No. 2, pp. 228-233, 2001.

[Mor13a] Morales, A., Ferrer, M. A., Diaz-Cabrera, M., & Gonzalez, E. (2013, October). Analysis of local descriptors features and its robustness applied to ear recognition. 47th International Carnahan Conference on Security Technology (ICCST), pp. 1-5, 2013. IEEE.

[Nej12a] Nejati, H., Zhang, L., Sim, T., Martinez-Marroquin, E., & Dong, G. (2012, November). Wonder ears: Identification of identical twins from ear images. 21st International Conference on Pattern Recognition (ICPR), pp. 1201-1204, 2012. IEEE.

[Pau10a] Paul, P. P., Monwar, M. M., Gavrilova, M. L., & Wang, P. S. Rotation invariant multiview face detection using skin color regressive model and support vector regression. International Journal of Pattern Recognition and Artificial Intelligence, 24, No. 08, pp. 1261-1280, 2010.

[Pfl12a] Pflug, A., & Busch, C. (2012). Ear biometrics: a survey of detection, feature extraction and recognition methods. Biometrics, IET, 1, No. 2, pp. 114-129.

[Sul14a] Sultana, M., Gavrilova, M., & Yanushkevich, S. Expression, pose, and illumination invariant face recognition using lower order pseudo Zernike moments. 9th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP), pp. 216-221, 2014.

[Yan03a] Yan, C., Sang, N., & Zhang, T. (2003). Local entropy-based transition region extraction and thresholding. Pattern Recognition Letters, 24, No. 16, pp. 2935-2941.

[Yua12a] Yuan, L., and Chun Mu, Z. Ear recognition based on local information fusion. Pattern Recognition Letters, 33, No. 2, pp. 182-190, 2012.

[Yua12b] Yuan, L., Li, C., and Mu, Z. Ear recognition under partial occlusion based on sparse representation. 2012 International Conference on System Science and Engineering (ICSSE), pp. 349-352, 2012.

[WPUTED] http://ksm.wi.zut.edu.pl/wputedb/ last accessed on March 12, 2015.

[Taj13a] Tajima, Y., Ito, K., Aoki, T., Hosoi, T., Nagashima, S., & Kobayashi, K. (2013, June). Performance improvement of face recognition algorithms using occluded-region detection. 2013 International Conference on Biometrics (ICB), pp. 1-8, 2013.