Heterogeneous iris image hallucination using sparse representation on a learned heterogeneous patch dictionary

Yung-Hui Li *[a], Bo-Ren Zheng [a], Dai-Yan Ji[b], Chung-Hao Tien[c], Po-Tsun Liu[c]
[a] Dep. of Computer Science and Information Engineering, National Central Univ./300 Jhongda Rd., Taoyuan Taiwan 320, R.O.C.
[b] Advanced Analog Technology Inc./17 Industry E. 2nd Rd., Hsinchu Taiwan 300, R.O.C.
[c] Dep. Of Photonics, National Chiao Tung Univ./1001 University Rd., Hsinchu Taiwan 300, R.O.C.

ABSTRACT

Cross sensor iris matching may seriously degrade the recognition performance because of the sensor mis-match problem of iris images between the enrollment and test stage. In this paper, we propose two novel patch-based heterogeneous dictionary learning method to attack this problem. The first method applies the latest sparse representation theory while the second method tries to learn the correspondence relationship through PCA in heterogeneous patch space. Both methods learn the basic atoms in iris textures across different image sensors and build connections between them. After such connections are built, at test stage, it is possible to hallucinate (synthesize) iris images across different sensors. By matching training images with hallucinated images, the recognition rate can be successfully enhanced. The experimental results showed the satisfied results both visually and in terms of recognition rate. Experimenting with an iris database consisting of 3015 images, we show that the EER is decreased 39.4% relatively by the proposed method.

Keywords: sensor mis-match, patch-based heterogeneous dictionary, sparse representation

1. INTRODUCTION

Biometric identification is to identify a person through his/her physiological or behavioral characteristics [1]. Among all possible biometric modalities, iris recognition [2] achieves the highest recognition rate and hence is highly valued in both research community and industry. Recently, due to the booming of the interest of widely using iris recognition in real life scenario, more and more companies start making new image sensors for iris acquisition. When a system designer decides to replace an old iris camera with a newer model, it is nearly impossible to call back all existing users to collect their enrollment data with the new image sensors. Therefore, there will be situations where we have to match iris images captured by new sensors to the one that is captured by the old sensor. Such problem can be called "heterogeneous iris recognition" or "cross-sensor iris matching problem".

Iris recognition under cross-sensor condition is an important issue in practical situation because it may decrease the recognition performance. Therefore, the typical impression of the "extremely high recognition rate" of iris recognition may not hold under cross-sensor condition.

In this paper, we propose two approaches to solve the problem of heterogeneous iris recognition. We implemented both approaches and tested the performance on large-scale heterogeneous iris database. Our contribution includes:

(1) To the best knowledge of the authors, it is the first work that applies the latest sparse representation theory into the problem of heterogeneous iris recognition.
(2) Both of the proposed ideas are intuitive and easy to interpret, compared to the existing work.
(3) Two learning based approaches are proposed for the problem.
(4) Performance evaluation on large-scale heterogeneous iris database.

The rest of the paper is organized as following. The previous work is reviewed in section 2. The two proposed methods are described in section 3. The experimental procedure and results are presented in section 4, which is followed by conclusion in section 5.

Applications of Digital Image Processing XXXVII, edited by Andrew G. Tescher, Proc. of SPIE Vol. 9217, 92171V
 · doi: 10.1117/12.2060838

2. PREVIOUS WORK

There are not too many existing publications that address the issue of heterogeneous iris recognition. Bowyer et al. [3, 4] investigated the interoperability of iris sensors from different manufacturers using multiple available matching algorithms. Pillai et al. [5] used a kernel learning method [6] for learning transformations from iris images captured by one sensor to another and applied such framework for sensor adaptation.

For the research work about heterogeneous face recognition, Li et al. [7, 8, 9] proposed a face-sketch heterogeneous space eigenface method that is able to synthesize face images based on its sketch counterpart. In recognition stage, an advanced correlation filter is built in order to perform illumination tolerant face recognition.

3. PROPOSED METHOD

In this work, we propose two patch-based dictionary-learning methods for the purpose of heterogeneous iris image hallucination. They are described in the following sub-sections, respectively.

3.1 Heterogeneous dictionary learning by sparse representation

We propose a novel method to attack this problem, which is a patch-based heterogeneous dictionary learning method using sparse representation. Given a heterogeneous iris database that consists of two iris image sets, captured by two iris image sensors A and B, we denote these two datasets I^A and I^B. Specifically,

$$I^A = \{I_1^A, I_2^A, \dots, I_M^A\} \tag{1}$$
$$I^B = \{I_1^B, I_2^B, \dots, I_M^B\} \tag{2}$$

where I_k^A and I_k^B denotes the k^{th} iris images in image set I^A and I^B, respectively. Note that these two iris images subsets are preprocessed so that

(a) The corresponding iris images I_k^A and I_k^B are coming from the same subject
(b) I_k^A and I_k^B are globally aligned.

Here, the "globally aligned" means that iris feature extraction and matching algorithm has been applied to two iris images I_k^A and I_k^B, and the best circular shift amount between them has been computed. Next, one of the two images has been circularly shifted so that the iris texture patterns between I_k^A and I_k^B are aligned globally.

Next, the iris images are all broken down into overlapped patches. The patch-based heterogeneous iris database is represented as P_A and P_B.

$$P^A = \{P_1^A, P_2^A, \dots, P_N^A\} \tag{3}$$

$$P^B = \{P_1^B, P_2^B, \dots, P_N^B\} \tag{4}$$

where P_k^A and P_k^B denotes the k^{th} iris images patch in image set P^A and P^B, respectively. Note that $N>>M$.

The next step is to form a heterogeneous dictionary for iris patches. In this stage, we create a new heterogeneous patch set Θ from P^A and P^B. Specifically,

$$\Theta = \{HP_i | HP_i = \begin{bmatrix} P_i^A \\ P_i^B \end{bmatrix}, \forall\ 1 \le i \le N\} \tag{5}$$

The set Θ can be viewed as iris image patch set in a heterogeneous space, which is composed by combining image patches from different optical sensors. Therefore, in this work, we call Θ as "heterogeneous iris dictionary".

During the test stage, given a test iris image I_{test}^B captured by image sensor B, our goal is to hallucinate its corresponding image I_{test}^A so that it looks as if it is captured by sensor A and has the same image quality as all images in set I^A. Here the basic assumption is that the image quality of set I^A is much higher than that of I^B, therefore, in order to achieve higher recognition rate, it is highly desired to hallucinate I_{test}^A based on the given image I_{test}^B .

First, the given test image I_{test}^B is broken into overlapped patches. For each patch p_i^{test}, we perform sparse decomposition using $\{P_i^B\}$ as dictionary D to compute the sparsest reconstruction coefficient α_i:

$$\alpha_i = arg\min_{\beta_i}(\|p_i^{test} - D\beta_i\|_2^2 + \mu\|\beta_i\|_0) \tag{6}$$

Equation (6) can be solved by Orthogonal Matching Pursuit (OMP) [10, 11]. Thus, α_i contains information indicating which atoms in D should be used to reconstruct p_i^{test}, under the constraint that the number of the reconstruction atoms is minimized. Therefore, the index of the non-zero element in α_i gives us a hint about which element in D has the highest resemblance to p_i^{test}. Suppose the index of the element with the largest value in α_i is j, then we are confident to declare that the atom P_j^B has the highest resemblance to p_i^{test}. Using P_j^A which is the counterpart of P_j^B in the upper part of the heterogeneous dictionary Θ to represent p_i^{test} in reconstructed space, and continuing applying such method $\forall 1 \leq i \leq N$, we are able to hallucinate I_{test}^A.

Figure 1 and 2 shows the proposed idea in training and test stage, respectively.

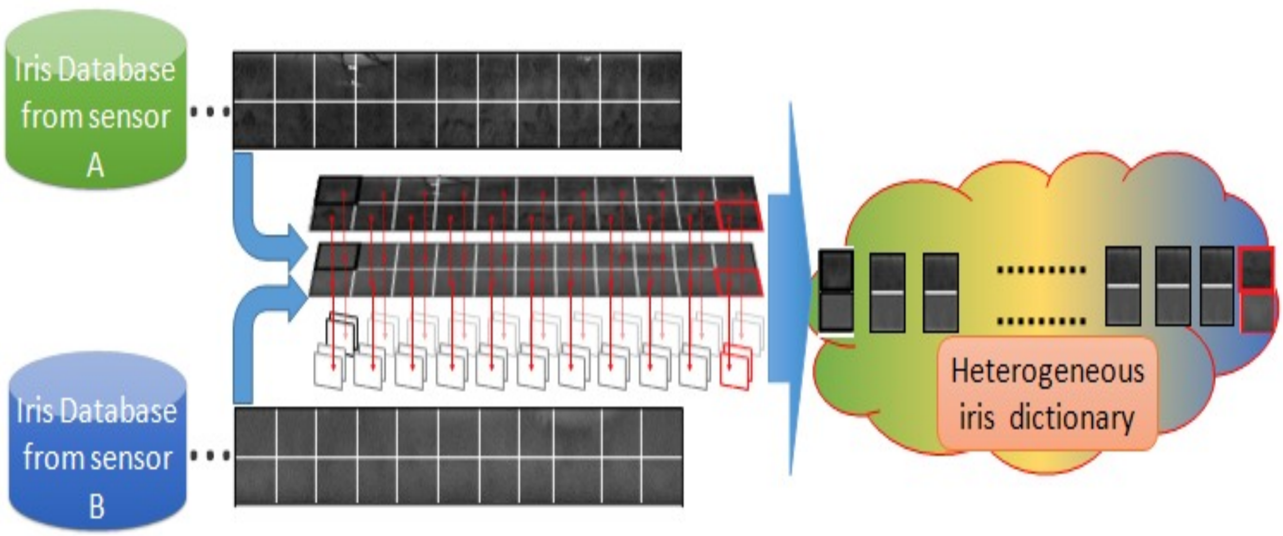

Figure 1: Illustration of the proposed method 1 (Heterogeneous dictionary learning by sparse representation) during training stage.

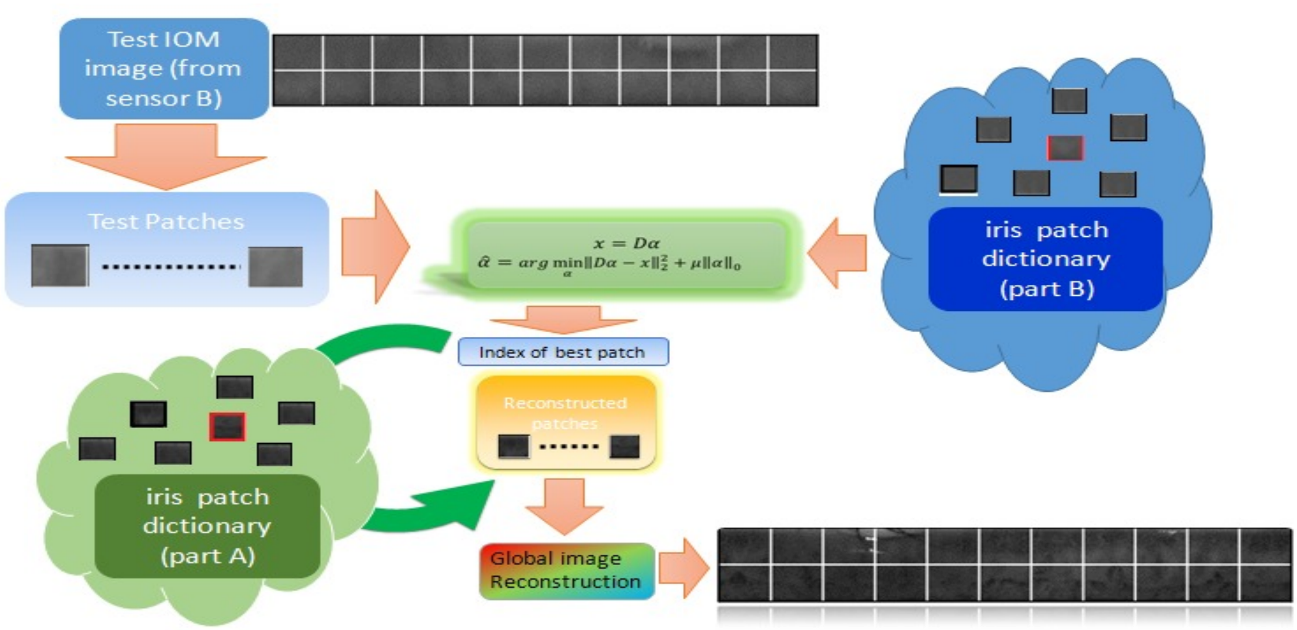

Figure 2: Illustration of proposed method 1 (Heterogeneous dictionary learning by sparse representation) during test stage.

3.2 Heterogeneous space eigeniris approach

The second method we propose is inspired by the work in [7, 8, 9]. Therefore, we call it "heterogeneous space eigeniris" approach.

Given iris image pair database I^A and I^B, again, we build a heterogeneous dictionary Θ for local patch. Inspired by [7, 8, 9], we would like to train the heterogeneous space eigeniris by using this heterogeneous patch set. Applying PCA on Θ, we get a set of heterogeneous space eigeniris images **Ψ.**

$$\mathbf{\Psi} = \{\psi_i | \psi_i = \begin{bmatrix} \phi_i^A \\ \phi_i^B \end{bmatrix}, \forall\, 1 \le i \le N\} \tag{7}$$

where ψ_i is the i^{th} eigenvector computed by solving the eigenvalue/eigenvector problem of the covariance matrix derived from Θ. Note that each ψ_i can be viewed as a combination of two eigen-patch images ϕ_i^A andϕ_i^B, belonging to the pseudo eigen-patch set of patch set P^A and P^B, respectively. The word "pseudo" here means that the eigen-patch set $\{\phi_i^A\}$ and $\{\phi_i^B\}$ does not really span the subspace P^A and P^B, because the property of orthonormality does not hold for either of them. Only after they are combined together (i.e.,**Ψ**) then does it have the orthonormality.

During test stage, given a test iris image I_{test}^B, again, we first broke the image into a set of local patches. Second, since we already have $\{\phi_i^B\}$ which can be viewed as pseudo eigen-patch set, we can project every patch in I_{test}^B to the subspace spanned by $\{\phi_i^B\}$ and compute their coordinate in this subspace. However, as described in the previous paragraph, $\{\phi_i^B\}$ does not have the orthonormality property. Therefore, we need to use pseudo-inverse to compute the projection coefficients. Specifically, for each patch p_i^{test}sampled from I_{test}^B, assuming ϕ^B is a matrix with ϕ_i^B being its i^{th} column, the projection coefficients pc_i^{test} can be computed as:

$$pc_i^{test} = ((\phi^B)^T \phi^B)^{-1} (\phi^B)^T p_i^{test} \tag{8}$$

Once pc_i^{test} is computed, it can be used to hallucinate the corresponding patch image p_i^{hal} using linear combination:

$$p_i^{hal} = \phi^A pc_i^{test} \tag{9}$$

where ϕ^A is a collection of $\{\phi_i^A\}$ whose i^{th} column is ϕ_i^A

After all patch images $\{p_i^{hal}\}$ are hallucinated, the corresponding global iris image I_{test}^A can be generated by overlapping the patch images $\{p_i^{hal}\}$ in their corresponding location.

4. EXPERIMENT

4.1 Database

In order to measure the iris recognition performance based on the proposed patch-based heterogeneous dictionary learning algorithm, experiments have to be performed on databases which contain both high quality and low quality iris images for the same iris class. The database we used in our experiment collected at Carnegie Mellon University during March and April in 2009. The iris images are captured by two kinds of iris acquisition devices: 1) IOM [12], whose image quality is low; 2) SecuriMetrics PIER 2.3 [13], whose image quality is much better than IOM. The details of the IOM and PIER database are given in Table 1. The total number of iris images is 3015.

Table 1: Statistics about IOM and PIER

Database Properties	IOM	PIER
Number of iris classes	111	
Size of the picture	640x480	
Maximal number of images per session	54	3
Minimal number of images per session	10	3
Average number of images per session	24	3
Total number of images	2682	333

4.2 Procedures

In practical situations, suppose we have collected a few iris images as training data, it is reasonable for us to further tuning the training dictionary with optimized data. Therefore, in our experiment, we choose training data with the following two schemes and measure their performance separately:

(1) Random selection: randomly select IOM/PIER image pairs to be the training data, the rest left as test testing data.
(2) Optimal selection: perform intra-class iris recognition, then select the iris image which has the lowest HD to all of the other images for the same subject. In this way, the training IOM/PIER image pairs would be the optimal one to use.

Therefore, we have a set of PIER iris images $I^A = \{I_1^A, I_2^A, \dots, I_M^A\}$, and a set of corresponding IOM iris images $I^B = \{I_1^B, I_2^B, \dots, I_M^B\}$, where$I_k^A$ and I_k^B is column vector. For test data, we choose all IOM iris images except the chosen image (training data) for each iris class. All both training data and test data will be pre-segmented and normalized to the size of 30x180.

All training data are divided into patches and stored in the corresponding heterogeneous iris dictionary. Figure 3 shows the heterogeneous iris dictionary including both high quality (upper parts captured by PIER) and low quality (lower parts captured by IOM) iris images.

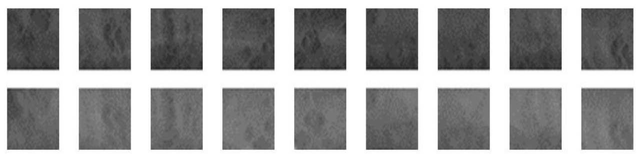

Figure 3: Illustration of heterogeneous iris dictionary during training stage

4.3 Patch size optimization

Because all training and test images are divided into patches, the size of the patch may affect the performance. In order to analyze the impact of the patch size to the recognition performance, we perform an experiment of patch size optimization.

For all classes, we perform heterogeneous iris image hallucination (hallucinating PIER images given IOM images) using proposed method, with various patch sizes. After the PIER images are hallucinated, we perform intra-class iris matching and record the average HD. The patch size which results in the lowest HD is the optimal size to use because under such patch size, the hallucinated iris image has the highest similarity to the training data.

The patch size we tried ranges from 3x3 to 30x30. The value of Hamming Distance (HD) for the training and testing matching varies in different sizes as shown in Figure 4. We can see that the best patch size is 17x17 and 29x29 for proposed method 1 (heterogeneous dictionary learning method using sparse representation) and 2 (heterogeneous space eigeniris method), respectively.

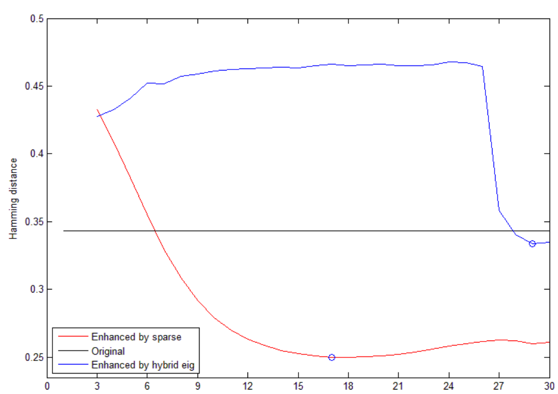

Figure 4: Experimental result of patch size optimization.

4.4 Large-scale heterogeneous iris recognition results

In Figure 5 we show the ROC curves that are based on three different methods, as well as the baseline. The baseline curve represents the iris recognition performance when directly matching training and testing images without using any algorithm to improve iris image quality. The red and black curves represent the iris recognition performance after using the patch-based heterogeneous dictionary learning method in the random and optimal selection situations, and the pink curve represents the iris recognition performance after using the heterogeneous space eigeniris method in the optimal selection situation to enhance the test image quality.

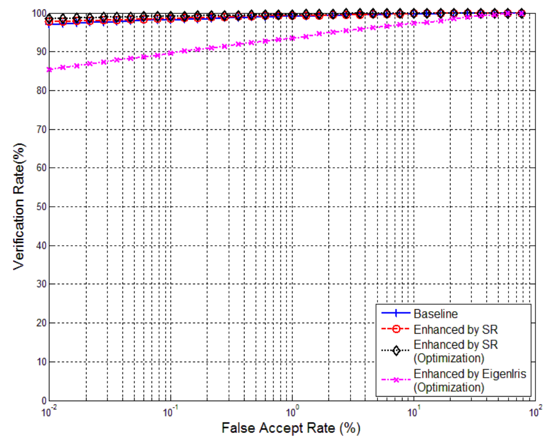

Figure 5: ROC curves comparison of the baseline, heterogeneous space eigeniris method, and the heterogeneous dictionary learning method using sparse representation.

We can see that when FAR = 10^{-2} %, the heterogeneous dictionary learning method using sparse representation is much better than the method using heterogeneous space eigeniris (which only achieves 85.46%) in the verification rate. The result reveals that the method using heterogeneous dictionary learning by sparse representation is more suitable for approaching the sensor mis-match problem.

In Figure 6 we show the zoom-in version of Figure 5, where we only plot the baseline and the result of using sparse representation (two schemes of training data selection). We can see that when FAR = 10^{-2} %, the verification rate of the heterogeneous dictionary learning method using sparse representation achieves 98.5% and 97.7% in the optimal and random selection schemes respectively, both of which are superior than that of baseline (which is about 96.9%).

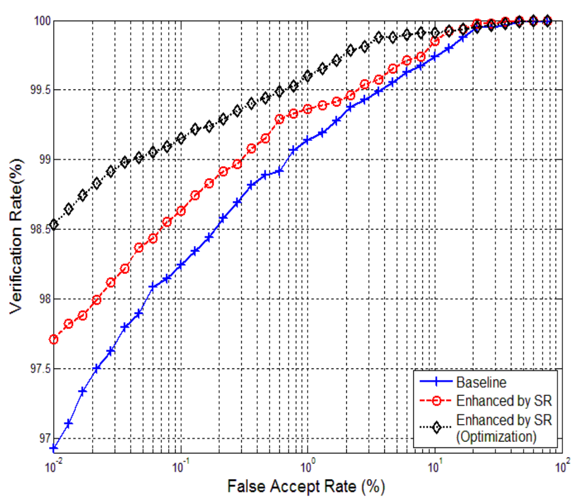

Figure 6: ROC curves comparison of the baseline and the heterogeneous dictionary learning method using sparse representation in the random and optimal selection schemes.

Figure 7 shows the histogram of HD distribution for the authentic and impostor comparison, before (which is baseline) and after applying the proposed method. We can see that the authentic score distribution obviously being moved toward left side, while the imposter score distribution remains almost the same. Moreover, the EER of the proposed method achieves 0.5344%, compared to EER=0.8824% in the baseline experiment. It shows our proposed method is able to make EER decrease 39.4% relatively, demonstrating the effectiveness of the proposed method.

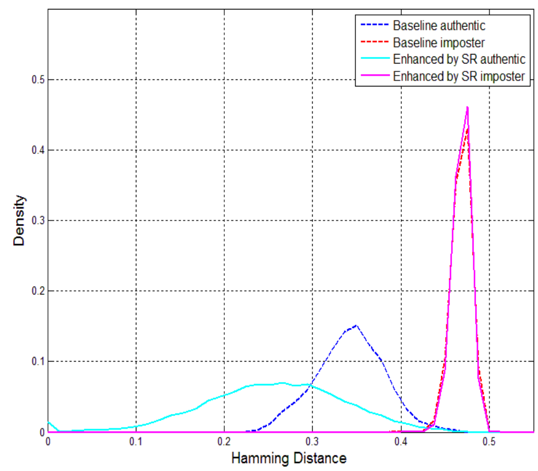

Figure 7: The density of Hamming distance of baseline and the heterogeneous dictionary learning method using sparse representation in the optimal selection scheme.

Figure 8 shows the example iris images hallucinated by the proposed methods SR (heterogeneous dictionary learning method using sparse representation) and eigeniris (heterogeneous space eigeniris method), respectively. From these three examples, given test IOM image whose quality is low, we can see that heterogeneous dictionary learning method using sparse representation can synthesize high quality image that look as if it is captured by PIER device.

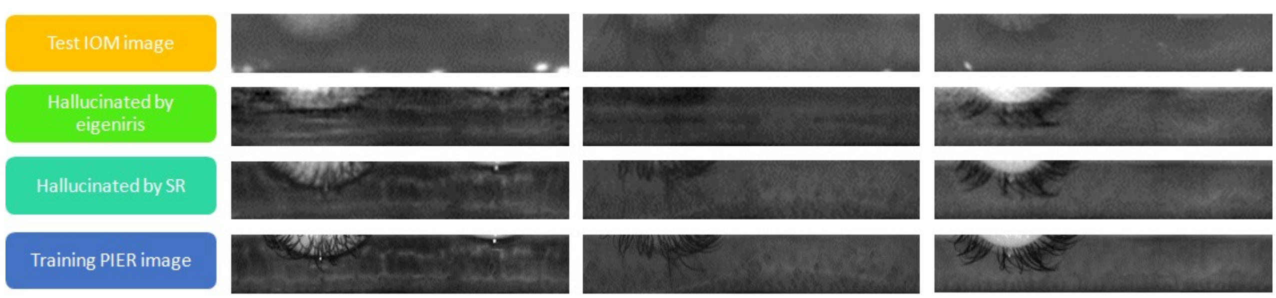

Figure 8: Comparison of the iris images that are synthesized by both proposed methods.

5. CONCLUSION

In this paper, the problem we would like to solve is the iris sensor mis-match problem. The problem is particularly serious when the following conditions are met: (1) the iris images for training and testing are acquired by different iris image sensors and (2) the training set images have higher quality while the test images have lower quality. We proposed two patch-based dictionary-learning methods. Among them, the proposed method using heterogeneous dictionary learning method by sparse representation achieves better recognition performance. It is able to reduce the EER up to 39.4% relatively. Future work includes using more delicate algorithm (for example, k-SVD [14, 15]) for dictionary atom update and collecting more heterogeneous iris images for even larger-scale experiment.

ACKNOWLEDGEMENT

This work was financially supported by the National Science Council of Taiwan under contract no. NSC 102-2221-E-008 -115.

REFERENCES

[1] Jain A.K., Ross A., and Prabhakar S., “An introduction to biometric recognition,” IEEE Transactions on Circuits and Systems for Video Technology 14(1), 4-20 (2004)

[2] Bowyer K. W., Hollingsworth K., and Flynn P. J., “Image understanding for iris biometrics: A survey,” Computer Vision and Image Understanding 110(2), 281-307 (2008)

[3] Bowyer K., Baker S., Hentz A., Hollingsworth K., Peters T., and Flynn P., “Factors that degrade the match distribution in iris biometrics,” Identity in the information Society 2(3), 327-343 (2009)

[4] Connaughton R., Sgroi A., Bowyer K. W., and Flynn P. J., “A cross-sensor evaluation of three commercial iris cameras for iris biometrics,” in IEEE Computer Society Workshop on Biometrics, 90-97 (2011)

[5] Pillai J. K., Puertas M., and Chellappa R., "Cross-sensor Iris Recognition through Kernel Learning,", IEEE Transactions on Pattern Analysis and Machine Intelligence 36(1), 73-85 (2013)

[6] Weinberger K. Q., Sha F., and Saul L. K., “ Learning a kernel Matrix for nonlinear dimensionality reduction,” in International Conference on Machine learning, 839-846 (2004)

[7] Li Y., Savvides M. and Bhagavatula V., "Illumination Tolerant Face Recognition Using a Novel Face From Sketch Synthesis Approach and Advanced Correlation Filters," Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, (2006)

[8] Li Y. and Savvides M., "Faces from sketches: a subspace synthesis approach," SPIE Defense and Security Symposium. International Society for Optics and Photonics, (2006)

[9] Savvides M., Bhagavatula R., Li Y and Abiantun R., “Frequency Domain Face Recognition,” Face Recognition, (2007)

[10] Davis G., Mallat S., and Avellaneda M., “Adaptive Greedy Approximations,” Constructive Approximation 13(1), 57-98 (1997)

[11] Pati Y., Rezaiifar R., and Krihnaprassad P., “Orthogonal Matching Pursuit: Recursive Function Approximation with Applications to Wavelev Decomposition,” Conference Record of The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers, 40-44 (1993)

[12] Matey J., Naroditsky O., Hanna K., Kolczynski R., LoIacono D., Mangru S., Tinker M., Zappia T., and Zhao W., “Iris on the move: Acquisition of images for iris recognition in less constrained environments,” Proceedings of the IEEE 94(11), 1936-1947 (2006)

[13] “Securimetrics pier device,” securiMetrics Inc., http://www.securimetrics.com/solutions/pier.html.

[14] Aharon M., Elad M., and Bruckstein A. M., “K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation,” IEEE Transactions on Signal Processing 54(11), 4311-4322 (2006)

[15] Donoho D. L., Elad M., and Temlyakov V.N., “Stable recovery of sparse overcomplete representations in the presence of noise,” IEEE Transactions on Information Theory 52(1), 6-18 (2006)