

# Reconstruction of Smartphone Images for Low Resolution Iris Recognition

Fernando Alonso-Fernandez\*, Reuben A. Farrugia†, Josef Bigun\*

\* Halmstad University. Box 823. SE 301-18 Halmstad, Sweden

† Department of CCE, University of Malta, Msida, Malta

Emails: feralo@hh.se, reuben.farrugia@um.edu.mt, josef.bigun@hh.se

**Abstract**—As iris systems evolve towards a more relaxed acquisition, low image resolution will be a predominant issue. In this paper we evaluate a super-resolution method to reconstruct iris images based on Eigen-transformation of local image patches. Each patch is reconstructed separately, allowing better quality of enhanced images by preserving local information. We employ a database of 560 images captured in visible spectrum with two smartphones. The presented approach is superior to bilinear or bicubic interpolation, specially at lower resolutions. We also carry out recognition experiments with six iris matchers, showing that better performance can be obtained at low-resolutions with the proposed eigen-patch reconstruction, with fusion of only two systems pushing the EER to below 5-8% for down-sampling factors up to a size of only  $13 \times 13$ .

## I. INTRODUCTION

Among all biometric techniques, iris recognition is regarded as one of the most reliable and accurate identification system available [1]. In recent years, there has been an increasing pressure towards allowing acquisition ‘at a distance’ and ‘on the move’ by means of relaxed acquisition conditions [2]. It poses additional problems to the quality of acquired iris images, with the lack of pixel resolution being the most evident. Here, we address the problem of up-sampling, or increasing the size of a low-resolution image, due to for example a long acquisition distance to the object of interest. Low resolution is also an issue when bandwidth or storage limitations exist, such as in remote surveillance systems, personal devices or smartcards. In these cases, images are usually compressed e.g. via JPEG2000 [3], but image dimensions are kept constant. Nevertheless, image quality losses due to compression are not within the scope of this paper.

Super-resolution (SR) techniques aim to reconstruct the missing high resolution (HR) image  $\bar{Y}$  given a low resolution (LR) image  $\bar{X}$ . The LR image is modeled as the corresponding HR image manipulated by blurring ( $B$ ), warping ( $W$ ) and down-sampling ( $D$ ) as  $\bar{X} = DBW\bar{Y} + \bar{n}$  ( $\bar{n}$  represents additive noise). For simplicity, some works omit the warp matrix and noise, leading to  $\bar{X} = DB\bar{Y}$ . Two main SR approaches exist: reconstruction- and learning-based [4]. In reconstruction-based, sub-pixel shifts among multiple LR images are fused to obtain a HR image, therefore several LR images are needed as input. Alternatively, in learning-based approaches, the relationship between LR and HR images of a training database is modeled, and reconstruction can be done with only one LR image as input. Learning approaches generally outperform reconstruction methods and achieve higher magnification factors. Recently, SR techniques have been applied

to biometric systems. Since Baker and Kanade first suggested SR algorithms specifically for faces [5], there has been a lot of research in face super-resolution (also called hallucination) [6]. Learning approaches have been also used with other biometrics including iris [7], [8], [9], [10]. However, one major limitation is that they try to develop a prototype iris using combination of complete images. Eigen-patches is a strategy which models a local patch using collocated patches from a dictionary, instead of using the whole image. Each patch is hallucinated separately, providing better quality reconstructed prototypes with better local detail and lower distortions. Local methods are also generally superior in recovering texture than global methods.

Here, we apply an iris super-resolution technique based on PCA Eigen-transformation of local image patches inspired by the system of [11] for face images. A PCA Eigen-transformation is conducted in each patch of the input LR image. The HR patch is then reconstructed as a linear combination of collocated HR patches of the training database. This way, every patch has its own optimal reconstruction coefficients, allowing to preserve local image information. Prior to the hallucination process, iris images are aligned with respect to the pupil center, since alignment is critical for the performance of SR systems. In our experiments, we use the Visible Spectrum Smartphone Iris (VSSIRIS) database [12], captured with two smartphones. We also conduct verification experiments with six different iris matchers. The presented hallucination method is superior to bilinear and bicubic interpolation, showing more resiliency to down-sampling. Regarding the recognition rates, better performance can be obtained at low-resolutions with the proposed eigen-patch reconstruction method w.r.t. bilinear or bicubic interpolation. It is also observed that recognition performance is not significantly degraded with any matcher until a down-sampling factor of 1/8 (image size of  $39 \times 39$ ). In addition, by fusion of just two matchers, performance at low resolutions can be improved as well, with EER pushed to below 5-8% for any given down-sampling factor (which in our experiments includes iris images of up a size of only  $13 \times 13$ ). This means for example that storage or data transmission requirements can be reduced, two important features for biometric technologies to achieve massive adoption [2].

## II. EIGEN-PATCH IRIS HALLUCINATION

The structure of the hallucination method is shown in Figure 1. It is based on the eigen-patch hallucination method for face images of [11], adapted for iris images at [13].

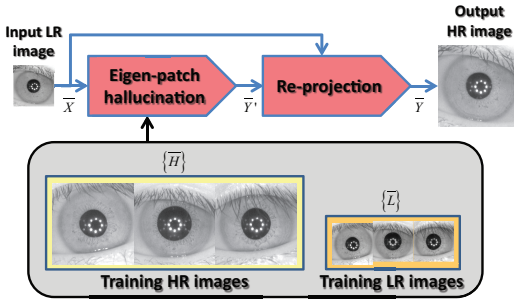


Fig. 1. Structure of the eigen-patch iris hallucination system.

Given an input low resolution (LR) iris image  $\bar{X}$ , the system first separates the image into  $N$  overlapping patches  $\{\bar{x}\} = \{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_N\}$ . Two super sets of basis patches are computed for each LR patch  $\bar{x}_i$ , using collocated patches of a training database of high resolution images  $\{\bar{H}\}$ . One of the super sets,  $\{\bar{h}_i^1, \bar{h}_i^2, \dots, \bar{h}_i^M\}$ , is obtained from collocated high-resolution (HR) patches of the training database. By degradation (low-pass filtering and down-sampling) a low resolution database  $\{\bar{L}\}$  is obtained from  $\{\bar{H}\}$ , and the other super set,  $\{\bar{l}_i^1, \bar{l}_i^2, \dots, \bar{l}_i^M\}$ , is obtained similarly, but for  $\{\bar{L}\}$ .  $M$  is the size (number of images) of the training set. A PCA Eigen-transformation is then conducted in each input LR patch  $\bar{x}_i$  using the collocated patches  $\{\bar{l}_i^1, \bar{l}_i^2, \dots, \bar{l}_i^M\}$  of the LR facial training images to obtain the optimal reconstruction weights  $\bar{c}_i = \{c_i^1, c_i^2, \dots, c_i^M\}$  of each patch (see Figure 2). By allowing each LR patch of the input image to have its own optimal reconstruction weights, the HR patch will be closer to the input LR patch, therefore more local information can be preserved and less reconstruction artifacts appear. Once the reconstruction weights  $\bar{c}_i$  of each patch are obtained, the HR patches are rendered using the collocated patches of the HR images of the training set  $\{\bar{H}\}$ . The reconstruction coefficients of the input image  $\bar{X}$  using the LR patches is carried on to weight the HR basis set, which yields the preliminary reconstructed HR iris image  $\bar{Y}'$ , after averaging the overlapping regions. Additional details of this Eigen-transformation procedure can be obtained in [11].

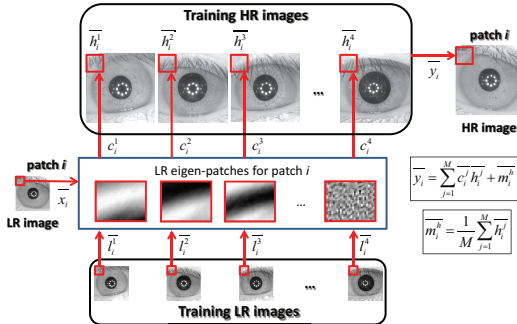


Fig. 2. Eigen-patch hallucination step.

A re-projection step is further applied to  $\bar{Y}'$  to reduce artifacts and make the output image  $\bar{Y}$  more similar to the input image  $\bar{X}$ . The image  $\bar{Y}'$  is re-projected to  $\bar{X}$  via  $\bar{Y}^{t+1} =$

$\bar{Y}^t - \tau U \left( B \left( DB\bar{Y}^t - \bar{X} \right) \right)$  where  $U$  is the upsampling matrix. The process stops when  $|\bar{Y}^{t+1} - \bar{Y}^t|$  is smaller than a threshold ( $10^{-5}$  in our experiments). We also use  $\tau=0.02$ .

### III. IRIS RECOGNITION SYSTEMS

We conduct iris matching experiments using six different systems based on 1D log-Gabor filters (LG) [14], SIFT operator (SIFT) [15], local intensity variations in iris textures (CR) [16], Discrete-Cosine Transform (DCT) [17], cumulative-sum-based grey change analysis (KO) [18], and Gabor spatial filters (QSW) [19]. In LG, CR, DCT, KO and QSW, the iris region is first unwrapped to a normalized rectangle using the Daugman's rubber sheet model [20]. Normalization produces a 2D array of  $20 \times 240$ , height  $\times$  width, (LG) and  $64 \times 512$  (others), with horizontal dimensions of angular resolution and vertical dimensions of radial resolution. Feature encoding is implemented according to the different feature extraction methods, leading to fixed-length templates which are matched using distance measures. Rotation is accounted for by shifting the grid of the 2D array of the query image in counter- and clock-wise direction and selecting the lowest distance, which corresponds to the best match between the two templates. In the SIFT matcher, SIFT key points are directly extracted from the iris region (without unwrapping), and the recognition metric is the number of matched key points, normalized by the average number of detected keypoints in the two images under comparison. The LG implementation is from Libor Masek code [14]; the SIFT method uses a free toolkit for feature extraction and matching<sup>1</sup>, with the adaptations described in [21] (particularly, it includes a post-processing step to remove spurious matching points using geometric constraints); and the remaining algorithms are from the University of Salzburg Iris Toolkit software package (USIT) [22].



Fig. 3. Sample images from VSSIRIS database (taken from [12]).

### IV. DATABASE AND EXPERIMENTAL PROTOCOL

We use the Visible Spectrum Smartphone Iris (VSSIRIS) database [12], which consists of images from 28 subjects (56 eyes) captured using the rear camera of two different smartphones (Apple iPhone 5S and Nokia Lumia 1020). Images from the iPhone 5S have  $3264 \times 2448$  pixels, while images from the Lumia 1020 have  $3072 \times 1728$  pixels. Images have been obtained in unconstrained conditions under mixed illumination consisting of natural sunlight and artificial room

<sup>1</sup><http://vision.ucla.edu/vedaldi/code/sift/assets/sift/index.html>

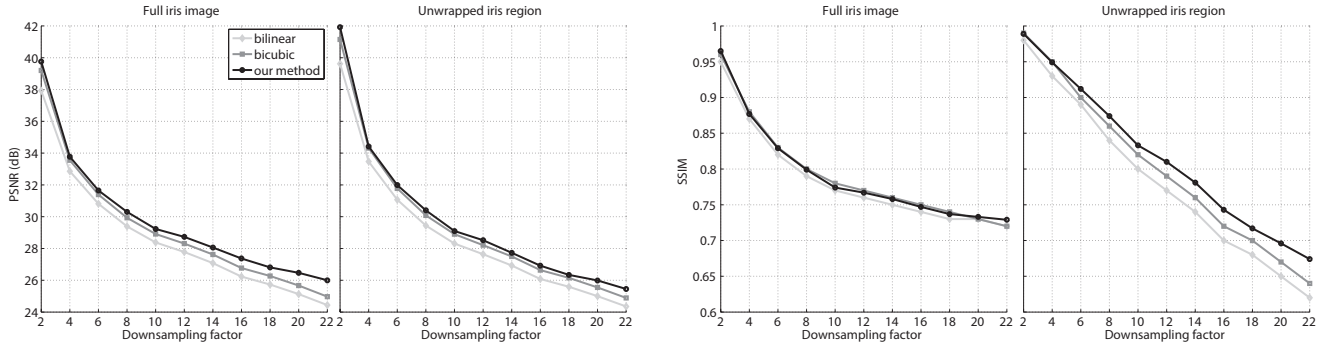


Fig. 4. Hallucination results with different down-sampling factors (left: PSNR values, right: SSIM values).

light. Each eye has 5 samples per smartphone, thus totalling  $5 \times 56 = 280$  images per device (560 in total). Acquisition is done without flash, in a single session and with semi-cooperative subjects. Figure 3 shows some example images. Iris segmentation data is also available, which has been used as input for our experiments. All images of the database have been resized via bicubic interpolation to have the same sclera radius (we choose as target radius the average sclera radius  $R=145$  of the whole database, given by the groundtruth). Then, images are aligned by extracting a square region of  $319 \times 319$  around the sclera center (corresponding to about  $1.1 \times R$ ). Two sample images can be seen in Figure 5, bottom.

Aligned and normalized HR iris images are then down-sampled via bicubic interpolation by a factor of  $2n$  (i.e. the image is resized to  $1/(2n)$  of the original HR size), and down-sampled images are used as input LR images, from which hallucinated HR images are computed. Given an input LR image, we use all available images from the remaining eyes (of both smartphones) to train the eigen-patch hallucination method (leave-one-out strategy). Training images are mirrored in the horizontal direction to duplicate the size of the training dataset, thus having  $55 \text{ eyes} \times 10 \text{ samples} \times 2 = 1100$  images for training. We also extract the normalized (unwrapped) iris region (size  $20 \times 240$ ) from both the hallucinated HR and the reference HR images, according to the algorithm of Section III. In addition, we perform verification experiments with the mentioned iris matchers. Experiments are done separately for each smartphone. We consider each eye as a different user. Genuine matches are obtained by comparing each image of an eye to the remaining images of the same eye, avoiding symmetric matches. Impostor matches are obtained by comparing the  $1^{st}$  image of an eye to the  $2^{nd}$  image of the remaining eyes. With this procedure, we obtain  $56 \times 10 = 560$  genuine and  $56 \times 55 = 3,018$  impostor scores per smartphone.

## V. RESULTS

Performance of the hallucination algorithm is measured by computing the PSNR (in dBs) and SSIM values between the hallucinated HR image and the corresponding HR reference image, with results given in Figure 4. We also compare our method with bicubic and bilinear interpolation. Figure 5 shows the hallucinated images (only for a selection of down-sampling factors for the sake of space).

As can be observed from Figure 4, the eigen-patch hallucination method outperforms bilinear and bicubic interpolations,

with a higher difference for bigger down-sampling factors. The advantage of eigen-patch hallucination is more evident when the resolution of the LR image becomes very low. This can be also appreciated in the examples of Figure 5, where the blurring of bilinear and bicubic interpolations at very low resolutions is clearly evident. The only exception is the SSIM measure on the full iris image, where the three interpolation methods show similar performance. It is also worth noting that the biggest drop in PSNR and SSIM values is observed for small down-sampling factors (1/4 to 1/8), but further down-sampling results in smaller drops (again, there is one exception with the SSIM measure on unwrapped iris images, but PSNR on the same images does show the mentioned behavior).

We now report verification experiments using hallucinated HR images. We consider two different scenarios: 1) enrolment samples taken from original HR input images, and query samples from hallucinated HR images; and 2) both enrolment and query samples taken from hallucinated HR images. The first case simulates a controlled enrolment scenario with good quality images, while the second case simulates a totally uncontrolled scenario (albeit for simplicity, enrolment and query samples have similar resolution in our experiments). Results are given in Figures 6 and 7. Several experimental findings can be put forward as result:

- Verification performance using the proposed hallucination method is similar to bilinear and bicubic interpolations for small down-sampling factors, but with some matchers (SIFT, CR, DCT) better performance is obtained with our hallucination method at very low resolutions. With the LG matcher, the difference in performance is already evident with just a down-sampling of 1/2. Interestingly, with the QSW matcher, bilinear and bicubic interpolations are better at low resolution, indicating that the artifacts introduced by the eigen-patch method due to the division of the image in patches may be jeopardizing performance (observe the ‘tessellation’ effect at very low resolutions in Figure 5). On the other hand, with the KO matcher, no particular interpolation method stands out.
- The performance of scenarios 1 and 2 is pretty similar up to a certain down-sampling factor but for some matchers, performance of scenario 2 is much better than scenario 1 at very low resolutions (specially SIFT and DCT matchers, but also CR, LG and QSW). When the size of input LR images is very low, the

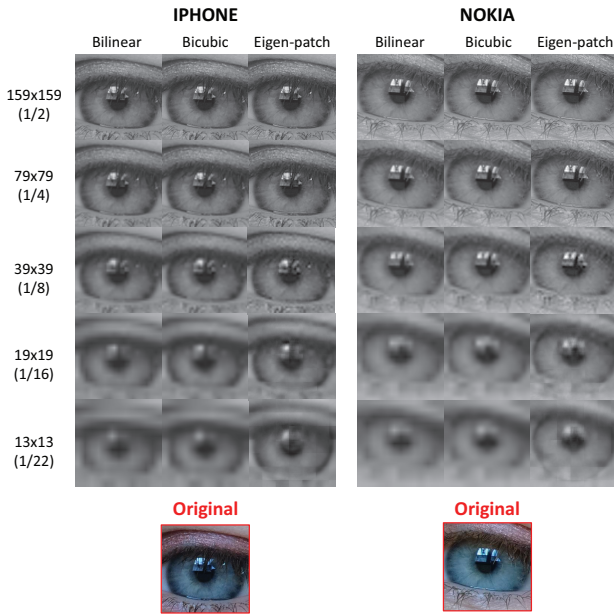


Fig. 5. Resulting HR hallucinated images for different down-sampling factors. The original HR image is also shown (bottom).

quality of hallucinated HR images is fairly different than the quality of original HR images (see Figure 5). As a result, it is better to match hallucinated HR images among themselves (scenario 2) than matching an hallucinated HR image against an original HR image (scenario 1). On the other hand, the KO matcher is shown to be insensitive to this phenomenon.

- Performance with any given matcher is fairly constant up to a factor of 1/8, which corresponds to an image size of  $39 \times 39$ . This suggests that the size of both query and test images can be kept very low without sacrificing performance, which is relevant for instance with devices having low storage capabilities or with low speed communication channels (recall that original HR images have a size of  $319 \times 319$ ). The best performing matcher (SIFT) does not show a significant degradation until a down-sampling factor of 1/14 (image size  $23 \times 23$ ).
- The best absolute performance is obtained by the SIFT and LG matchers with eigen-patch interpolation. The SIFT matcher has an EER smaller than 1% until a down-sampling factor of 1/14 (scenario 2), then the EER increases up to 10-11%. The LG matcher, on the other hand, has an EER of 7-8% for any given down-sampling factor. Performance of the other matchers is always above 10% EER, although their performance is pretty constant for any down-sampling given factor, at least in scenario 2 (the DCT and CR matchers shows some tendency to a worse EER in lower resolutions).

We then carry out fusion experiments using linear logistic regression. Given  $N$  matchers which output the scores  $(s_{1j}, s_{2j}, \dots, s_{Nj})$  for an input trial  $j$ , a linear fusion is:  $f_j = a_0 + a_1 \cdot s_{1j} + a_2 \cdot s_{2j} + \dots + a_N \cdot s_{Nj}$ . The weights  $a_0, a_1, \dots, a_N$  are trained via logistic regression as described in [23]. We use this trained fusion approach because it has shown

better performance than simple fusion rules (like the mean or the sum rule) in our previous works. Results are given in Figures 8 and 9. For the sake of time and page space, we have only tested fusion combinations of two matchers, reporting here the combinations where there is a clear improvement in performance due to the fusion. The following observations can be done from these results:

- The matcher with the best absolute EER (SIFT) can be further supported by others to improve EER at low resolutions (see LG+SIFT, SIFT+KO, SIFT+QSW). In particular, the fusion of the two best systems (LG and SIFT) pushes the EER to below 5-8% (scenario 2) for any down-sampling factor. This is not to say that the best combination always involves the best individual systems. In this case, the two matchers are based on different features, so it is expected a high complementarity between them. In particular, LG works with the unwrapped iris image by applying a 1D Log-Gabor filter to the whole image, while SIFT uses the original iris image to compute orientation histograms around selected key points only.
- The DCT matcher also appears to be well complemented by other matchers (CR, KO and QSW). As with SIFT, DCT is a matcher that has tendency to a worse EER as resolution decreases. When DCT is fused with the mentioned matchers, such tendency is reduced (scenario 1) or even disappeared (scenario 2).

## VI. CONCLUSION

Variability in resolution will be a common situation as iris recognition evolves towards more relaxed acquisition conditions or is applied to forensic scenarios [24]. Here, we apply a iris super-resolution technique based on PCA Eigen-transformation of local image patches [11] to increase the resolution of iris images. Experiments are conducted on a database of images in visible spectrum captured with two smartphones. Experimental results show that the eigen-patch approach is superior to traditional bilinear or bicubic interpolations, being more resilient as image resolution decreases. We also carry out iris matching experiments on the reconstructed images with six different matchers according to state-of-the art techniques. Two operational scenarios are considered, one where original high-resolution images are matched against hallucinated high-resolution images (controlled enrolment), and another scenario where only hallucinated images are used (uncontrolled scenario). Experiments show that at low resolutions, better performance can be obtained with the proposed eigen-patch reconstruction method w.r.t. bilinear or bicubic interpolation. In addition, recognition performance with any matcher is not degraded significantly until image is down-sampled by 1/8 or higher factors, while one of the matchers (SIFT) does not show significant degradation until a down-sampling of 1/14. This allows to use images of reduced size, which is of importance under low storage or data transmission capabilities. We also carry out fusion experiments, with results showing that by appropriate combination of matchers, performance at low resolutions can be improved as well, with EER of the best combination of two matchers pushed to below 5-8% EER for any given down-sampling factor (which in our experiments



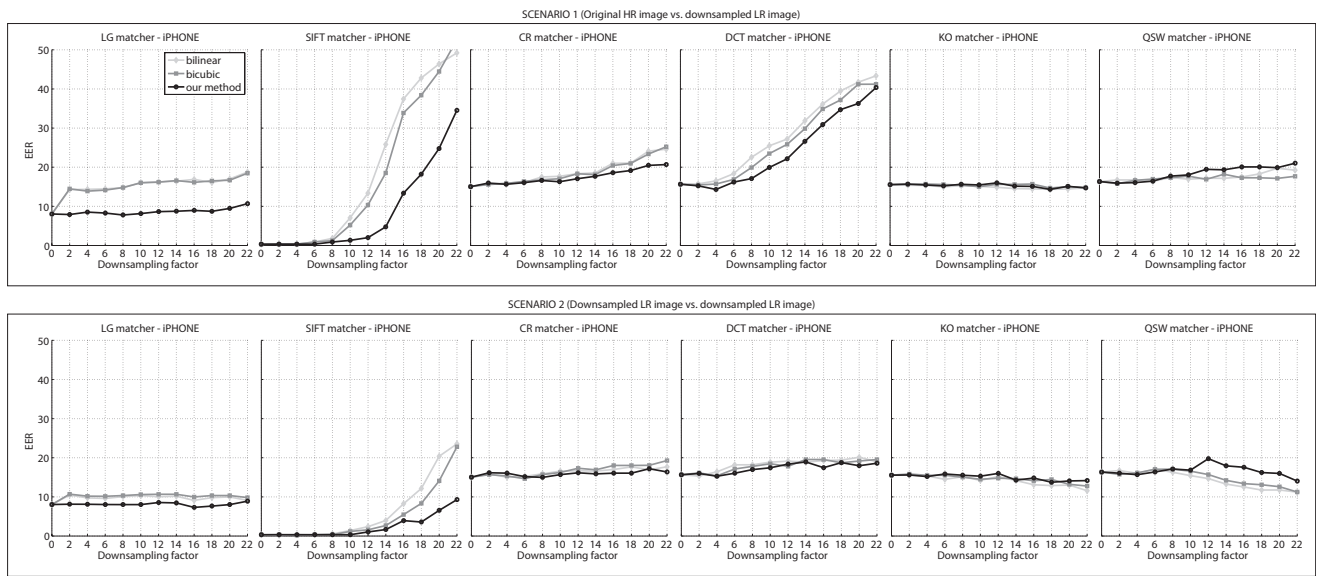


Fig. 6. **iPhone 5S**: Verification results (EER) of the two scenarios considered with different down-sampling factors.

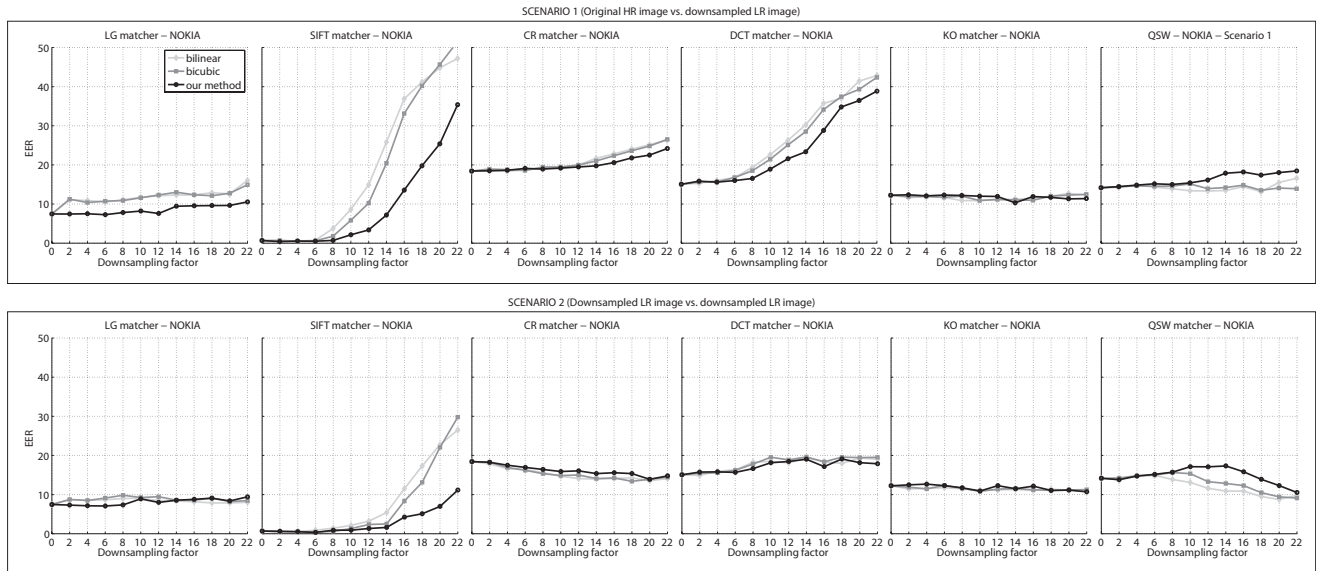


Fig. 7. **Nokia Lumia**: Verification results (EER) of the two scenarios considered with different down-sampling factors.

includes iris images of up to a down-sampling factor of 1/22, or an image size of only  $13 \times 13$ .

Future work includes improving the hallucination method by Manifold Learning approaches [25], which model non-linear relations between low and high resolution images, through which we expect to cope with artifacts appearing at low resolutions due to division of the image in patches. Another direction is increasing the number of systems involved in the fusion, as well as the use of training sets smaller than the one used here. Lastly, we plan to analyze sensitivity of the proposed method to misalignment of the eye, since alignment is very critical to proper hallucination output.

#### ACKNOWLEDGMENT

Authors thank K. B. Raja and R. Ramachandra of Gjøvik University College (GUC) in Norway for granting access to the VSSIRIS database while

F. A.-F. was a visiting researcher at GUC funded by EU COST Action IC1106. Author F. A.-F. also thanks the Swedish Research Council for funding his research, and the CAISR program of the Swedish Knowledge Foundation.

#### REFERENCES

- [1] A. Jain, P. Flynn, A. Ross, *Handbook of Biometrics*. Springer, 2008.
- [2] A. K. Jain, A. Kumar, *Second Generation Biometrics*. Springer, 2010, ch. Biometrics of Next Generation: An Overview.
- [3] G. W. Quinn, P. Grother, M. Ngan, "IREX IV: Part 2 - compression profiles for iris image compression," *NISTIR 7978*, 2014.
- [4] S. C. Park, M. K. Park, M. G. Kang, "Super-resolution image reconstruction: a technical overview," *IEEE SPM*, vol. 20, no. 3, 2003.
- [5] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," *IEEE TPAMI*, vol. 24, no. 9, pp. 1167–1183, Sep 2002.
- [6] N. Wang, D. Tao, X. Gao, X. Li, and J. Li, "A comprehensive survey to face hallucination," *IJCV*, vol. 106, no. 1, pp. 9–30, 2014.

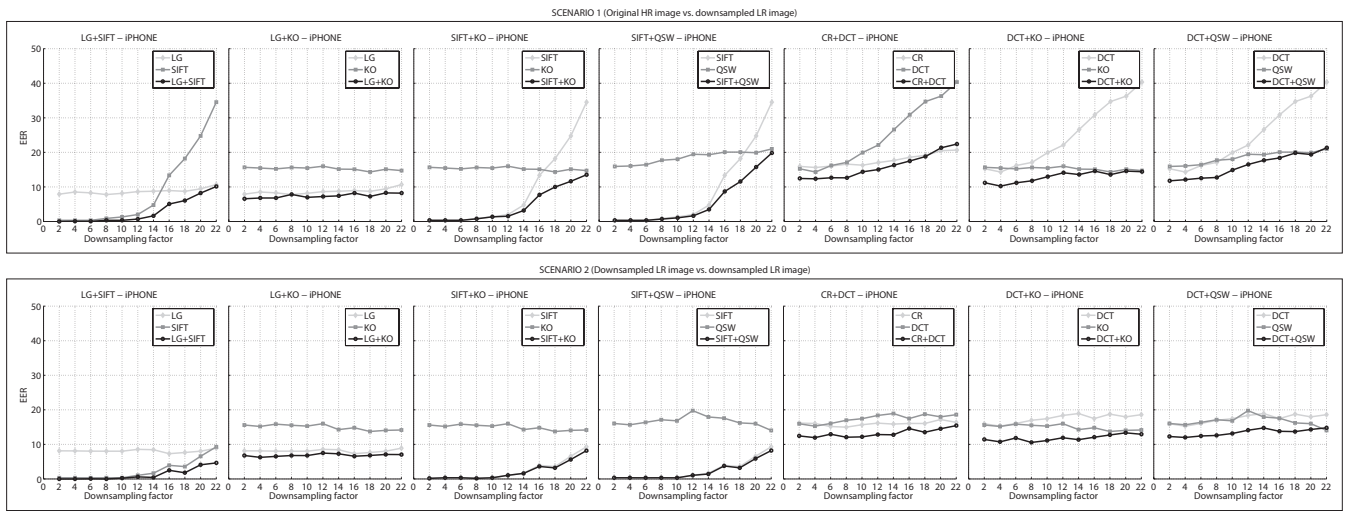


Fig. 8. iPhone 5S: Fusion results (EER) of the two scenarios considered with different down-sampling factors (eigen-patch reconstruction only).

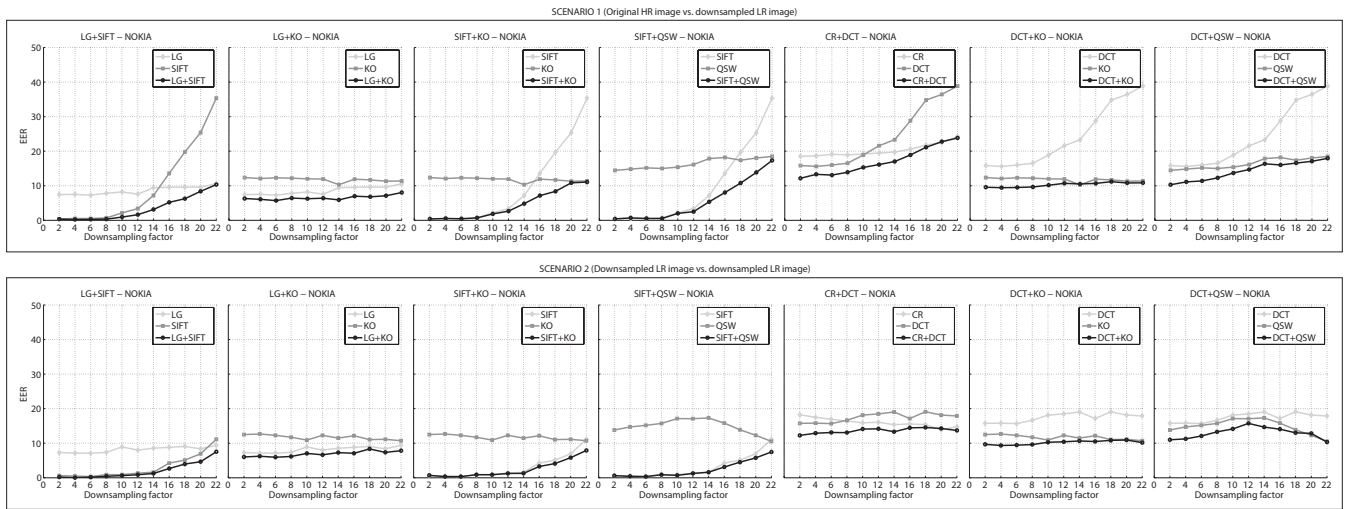


Fig. 9. Nokia Lumia: Fusion results (EER) of the two scenarios considered with different down-sampling factors (eigen-patch reconstruction only).

- [7] K. Y. Shin, K. R. Park, B. J. Kang, and S. J. Park, "Super-resolution method based on multiple multi-layer perceptrons for iris recognition," in *Proc ICUT*, Dec 2009, pp. 1–5.
- [8] K. Nguyen, C. Fookes, S. Sridharan, and S. Denman, "Quality-driven super-resolution for less constrained iris recognition at a distance and on the move," *IEEE TIFS*, vol. 6, no. 4, pp. 1248–1258, Dec 2011.
- [9] K. Nguyen, S. Sridharan, S. Denman, and C. Fookes, "Feature-domain super-resolution framework for gabor-based face and iris recognition," in *Proc CVPR*, June 2012, pp. 2642–2649.
- [10] R. Jillela, A. Ross and P.J. Flynn, "Information fusion in low-resolution iris videos using Principal Components Transform," *Proc WACV*, 2011.
- [11] H.-Y. Chen, S.-Y. Chien, "Eigen-patch: Position-patch based face hallucination using eigen transformation," in *Proc ICME*, July 2014.
- [12] K. B. Raja, R. Raghavendra, V. K. Vemuri, C. Busch, "Smartphone based visible iris recognition using deep sparse filtering," *Patt Recogn Lett*, vol. 57, pp. 33–42, 2015.
- [13] F. Alonso-Fernandez, R. A. Farrugia, J. Bigun, "Eigen-patch iris super-resolution for iris recognition improvement," *Proc EUSIPCO*, 2015.
- [14] L. Masek, "Recognition of human iris patterns for biometric identification," Master's thesis, University of Western Australia, 2003.
- [15] D. Lowe, "Distinctive image features from scale-invariant key points," *Intl Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [16] C. Rathgeb and A. Uhl, "Secure iris recognition based on local intensity variations," in *Proc ICIA*, vol. 6112, pp. 266–275, 2010.
- [17] D. Monro, S. Rakshit, and D. Zhang, "DCT-Based iris recognition," *IEEE TPAMI*, vol. 29, no. 4, pp. 586–595, April 2007.
- [18] J.-G. Ko, Y.-H. Gil, J.-H. Yoo, K.-I. Chung, "A novel and efficient feature extraction method for iris recognition," *ETRI Journal*, vol. 29, no. 3, pp. 399–401, 2007.
- [19] L. Ma, T. Tan, Y. Wang, D. Zhang, "Personal identification based on iris texture analysis," *IEEE TPAMI*, vol. 25, no. 2, pp. 1519–1533, 2003.
- [20] J. Daugman, "How iris recognition works," *IEEE TCSVT*, vol. 14, pp. 21–30, 2004.
- [21] F. Alonso-Fernandez, P. Tome-Gonzalez, V. Ruiz-Albacete, J. Ortega-Garcia, "Iris recognition based on sift features," *Proc BIDS*, 2009.
- [22] C. Rathgeb, A. Uhl, and P. Wild, *Iris Biometrics - From Segmentation to Template Security*, Springer, 2013.
- [23] F. Alonso-Fernandez, J. Fierrez, D. Ramos, J. Ortega-Garcia, "Dealing With Sensor Interoperability in Multi-biometrics: The UPM Experience at the Biosecure Multimodal Evaluation 2007," *Proc BTHI*, 2008.
- [24] K. Bowyer, K. Hollingsworth, and P. Flynn, "Image understanding for iris biometrics: a survey," *CVIU*, vol. 110, pp. 281–307, 2007.
- [25] J. Jiang, R. Hu, Z. Wang, Z. Han, "Face super-resolution via multilayer locality-constrained iterative neighbor embedding and intermediate dictionary learning," *IEEE TIP*, vol. 23, no. 10, pp. 4220–4231, Oct 2014.