

# Iris Super-Resolution Using Iterative Neighbor Embedding

Fernando Alonso-Fernandez  
IS-Lab/CAISR  
Halmstad University (Sweden)  
feralo@hh.se

Reuben A. Farrugia  
Department of CCE  
University of Malta (Malta)  
reuben.farrugia@um.edu.mt

Josef Bigun  
IS-Lab/CAISR  
Halmstad University (Sweden)  
josef.bigun@hh.se

## Abstract

*Iris recognition research is heading towards enabling more relaxed acquisition conditions. This has effects on the quality and resolution of acquired images, severely affecting the accuracy of recognition systems if not tackled appropriately. In this paper, we evaluate a super-resolution algorithm used to reconstruct iris images based on iterative neighbor embedding of local image patches which tries to represent input low-resolution patches while preserving the geometry of the original high-resolution space. To this end, the geometry of the low- and high-resolution manifolds are jointly considered during the reconstruction process. We validate the system with a database of 1,872 near-infrared iris images, while fusion of two iris comparators has been adopted to improve recognition performance. The presented approach is substantially superior to bilinear/bicubic interpolations at very low resolutions, and it also outperforms a previous PCA-based iris reconstruction approach which only considers the geometry of the low-resolution manifold during the reconstruction process.*

## 1. Introduction

Low image resolution can reduce the effectiveness of iris biometric systems to recognize individuals. Unfortunately, this problem arises in a number of real-world biometric applications that are becoming ubiquitous, such as those making use of surveillance or smart-phone cameras [13]. In this context, super-resolution (SR) techniques can be used to enhance the quality of low resolution iris images to improve the recognition performance of existing systems.

Two main categories of SR methods are distinguished in the literature: reconstruction- and learning-based [21]. Reconstruction-based methods fuse several low resolution (LR) images to obtain a high resolution (HR) image, with the disadvantage that multiple LR images are needed as input. On the other hand, learning-based methods use coupled dictionaries to learn the mapping relations between LR and HR image pairs in order to hallucinate a HR image from

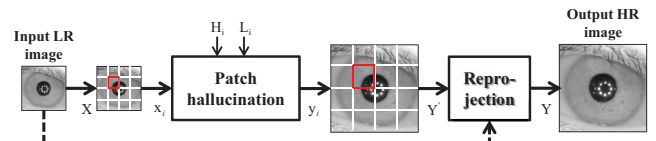


Figure 1. Block diagram of patch-based hallucination.

the observed LR one. Learning-based methods have the advantage of only needing one image as input, and generally allows to achieve higher magnification factors than reconstruction-based methods [21].

Super-resolution in biometrics is relatively recent, with a lot of research in face reconstruction [24]. Despite the vast literature on image SR, one reason of such limited research might be that most SR approaches are general-scene, designed to produce overall visual enhancement, but the aim of biometrics is a better recognition accuracy [19]. Therefore, adaptation of super-resolution techniques to the particularities of images from a specific biometric modality is needed to achieve a more efficient up-sampling [5]. Reconstruction-based methods to improve iris images from videos include for example the work [11], where authors compute the pixel-wise average of a number of aligned iris images, or the work [15], where authors apply PCA to unwrapped iris images in order to highlight the variance information among the pixel intensity vectors, and then compute the pixel-wise average of the resulting images. Both methods select as input images the frames with best quality from a given iris video stream. Learning-based methods include for example [23], which uses Multi-Layer Perceptrons, or [9], which employs frequency analysis. A major limitation of these two learning-based works is that they try to develop a prototype iris using combination of complete images. Patch-based methods, which models a local patch using collocated patches from the training dictionary, instead of using the whole image, have been also proposed. The work [12] for example employs Markov networks for this purpose, while the work [2] employs PCA. In these methods, each patch is hallucinated separately, having its own optimal reconstruction coefficients, which provides better

quality reconstructed prototypes with better local detail and lower distortions. Local methods are also generally superior in recovering texture than global methods, which is essential due to the prevalence of texture-based methods in ocular biometrics [20].

Existing iris super-resolution methods assume that low- and high-resolution manifolds of iris images have similar local geometrical structure, hence learn local models on the low-resolution manifold, which are then used to combine the high-resolution patches. For example, the previous work [2] uses PCA to project the low-resolution patches onto a low-dimensional sub-space. The weights suitable to reconstruct the low-resolution patch are then used to synthesize the high-resolution patch. However, the geometrical structure of the low-resolution manifold is distorted by the one-to-many relationship between low- and high-resolution patches [16]. Therefore, the reconstruction weights estimated on the low-resolution manifold do not correlate well with the actual weights needed to reconstruct the unknown high-resolution patch. In this work, we iteratively seek for the optimal neighbourhood on the high-resolution manifold, which is not affected by the degradation process. This method first approximates the high resolution patch using weights approximated on the low-resolution manifold, and then exploits the geometrical structure of the high-resolution manifold to improve the reconstruction quality [14]. Incorporating the HR space to the regularization process has the advantage that this space is unaffected by the image degradation process.

In our experiments, we employ the CASIA-IrisV3-Interval database [7] of NIR iris images. 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. We conduct verification experiments with two iris comparators based on Log-Gabor wavelets [18] and SIFT key-points [17]. Log-Gabor exploit texture information globally (across the entire iris image), while SIFT exploit local features (in discrete key points), therefore our motivation is to employ features that are diverse in nature, and reveal if they behave differently. Despite the patch-based SR approach used is not new [14], we contribute with its implementation to iris images, and particularly with the application (and fusion) of these two iris comparators to the reconstructed images. Reported results demonstrate the superiority of the presented method at very low resolutions w.r.t. bicubic/bilinear interpolations and the PCA method of [2]. The best individual matcher achieves an EER of  $\sim 4\%$  for an iris image size of only  $15 \times 15$ , with the fusion of the two systems pushing down this value to  $\sim 3.6\%$ . In addition, the improvement of the presented method w.r.t. the other methods is even more pronounced at high security regions (low FAR), where the iris modality is usually employed [6].

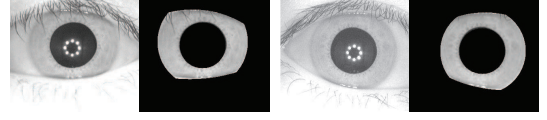


Figure 2. Example of images of the CASIA Interval v3 database with the annotated circles modeling iris boundaries and eyelids.

## 2. Low Resolution Iris Reconstruction

The method employed is based on the reconstruction algorithm for face images of [14], which is described next. Given an input low resolution (LR) image  $\mathbf{X}$ , the goal is to reconstruct its high resolution (HR) counterpart  $\mathbf{Y}$ . The LR image can be modeled as the HR image manipulated by blurring ( $B$ ), warping ( $W$ ) and down-sampling ( $D$ ) as  $\bar{\mathbf{X}} = DBW\mathbf{Y} + \bar{n}$  (where  $\bar{n}$  represents additive noise). For simplicity, the warp matrix and noise are usually omitted, leading to  $\bar{\mathbf{X}} = DB\mathbf{Y}$ .

In local patch-based methods, input LR images are first separated into  $N = N_v \times N_h$  overlapping patches  $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$  according to the predefined patch size and overlap pixels. Parameters  $N_v$  and  $N_h$  are the vertical and horizontal number of patches respectively. Since we will consider square images in our experiments, we can assume that  $N_v = N_h$ . Each individual LR patch  $\mathbf{x}_i$  is then hallucinated separately, producing the corresponding HR patch  $\mathbf{y}_i$ . Patch-based approaches provide better quality reconstructed prototypes than global approaches, with better local detail and lower distortions. The structure of the hallucination method employed is shown in Figure 1.

### 2.1. Training Dictionary

Two super sets of basis patches  $\mathbf{H}_i$  and  $\mathbf{L}_i$  are computed for each patch  $\mathbf{x}_i$  from collocated patches of a training database of  $M$  high resolution images  $\{\mathbf{H}\}$ . Super set  $\mathbf{H}_i = \{\mathbf{h}_i^1, \mathbf{h}_i^2, \dots, \mathbf{h}_i^M\}$  is obtained from collocated patches of  $\{\mathbf{H}\}$ . By degradation (low-pass filtering and down-sampling), a low-resolution database  $\{\mathbf{L}\}$  is obtained from  $\{\mathbf{H}\}$ , and the other super set  $\mathbf{L}_i = \{\mathbf{l}_i^1, \mathbf{l}_i^2, \dots, \mathbf{l}_i^M\}$  is obtained similarly, but from  $\{\mathbf{L}\}$ .

### 2.2. Multilayer Locality-Constrained Iterative Neighbour Embedding (M-LINE)

This work assumes that input images reside on a low-dimensional manifold where we try to find the optimal combination weights suitable to synthesize a high-resolution representation. The authors in [2] proposed a global approach where all entries in the dictionary are used to estimate the optimal weights. This, however, results in over-smooth reconstructed images which lacks important texture detail, which is essential for iris recognition, especially at very low-resolutions. Moreover, the authors in [16] empirically show that the structure of the low- and high-resolution

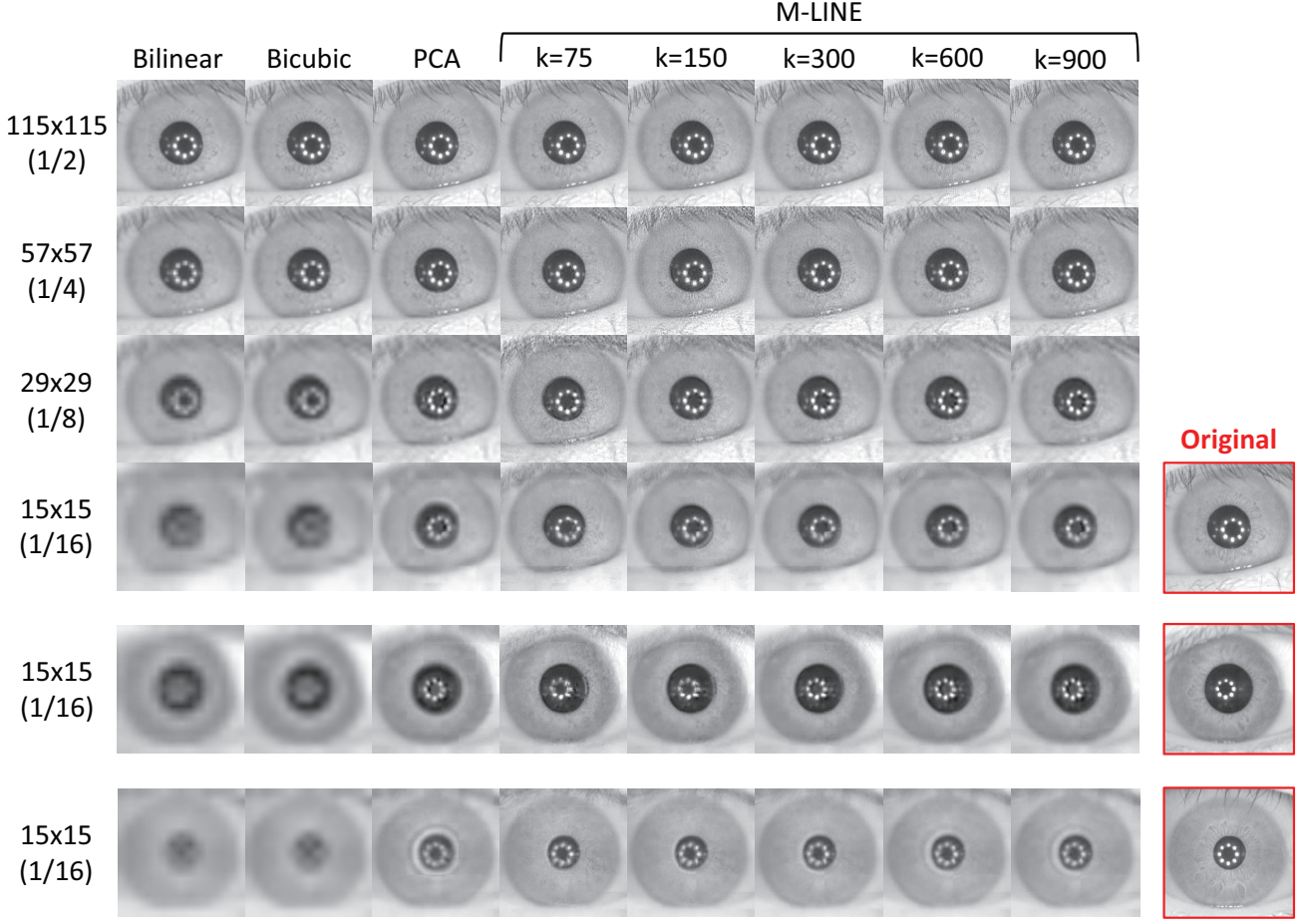


Figure 3. Resulting HR hallucinated images for different down-sampling factors. The original HR image is also shown (bottom right).

manifold is not well preserved, especially at very low resolutions.

The M-LINE method was originally proposed for face super-resolution [14] to *i*) to preserve the neighbourhood selected for neighbour embedding, and *ii*) mitigate the inconsistencies between the low- and high-resolution manifolds. The first problem is addressed by iteratively update of the neighbours used for neighbour embedding by exploiting the geometrical structure of the high resolution manifold, which is not affected by noise due to resolution degradation. The second step involves updating the low-resolution dictionary to reduce the modality gap between low- and high-resolution patches contained within the dictionary, although this step is not yet implemented for our experiments here.

The first estimate of the high-resolution patch  $\mathbf{v}_{i,0}$  is initialized by up-scaling the low-resolution patch  $\mathbf{x}_i$  using bicubic interpolation. This iterative method has a loop indexed by  $j \in [0, J - 1]$ . For every iteration  $j$ , the support  $\mathbf{s}$  of  $\mathbf{H}_i$  that minimizes the distance

$$\mathbf{d} = \|\mathbf{v}_{i,j} - \mathbf{H}_i(\mathbf{s})\|_2^2 \quad (1)$$

is computed using  $K$ -nearest neighbours ( $K \leq M$ ). The combination weights are then derived using

$$\mathbf{w}_{i,j}^* = \arg \min_{\mathbf{w}_{i,j}^*} (\|\mathbf{x}_i - \mathbf{L}_i(\mathbf{s}) \mathbf{w}_{i,j}^*\|_2^2 + \tau \|\mathbf{d}(\mathbf{s}) \odot \mathbf{w}_{i,j}^*\|_2^2) \quad (2)$$

where  $\tau$  is a regularization parameter. Operator  $\odot$  denotes the element-wise multiplication, and it is used to penalize the reconstruction weights with the distances between  $\mathbf{v}_{i,j}$  and its closest neighbors in the training dictionary  $\mathbf{H}_i$ . Note that Equation 2 jointly considers the LR manifold (via  $\mathbf{x}_i$ ,  $\mathbf{L}_i(\mathbf{s})$ ) and the HR counterpart (via  $\mathbf{d}(\mathbf{s})$ ). This optimization problem can be solved by an analytic solution [14]. The estimated high-resolution patch is then updated using

$$\mathbf{v}_{i,j+1} = \mathbf{H}_i(\mathbf{s}) \mathbf{w}_{i,j}^* \quad (3)$$

and the loop is then repeated. The final estimate of the high-resolution patch is then derived using

$$\mathbf{y}_i = \mathbf{v}_{i,J} \quad (4)$$

We employ  $\tau=1e^{-5}$  and  $J=4$ , as recommended in [14]. Finally, once the overlapping reconstructed patches  $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N\}$  are obtained, they are stitched together by averaging, resulting in the reconstructed HR image  $\mathbf{Y}'$ .

### 2.3. Image Re-projection

Inspired by [2], we incorporate a re-projection step to  $\mathbf{Y}'$  in order to reduce artifacts and make the output image  $\mathbf{Y}$  more similar to the input image  $\mathbf{X}$ . The image  $\mathbf{Y}'$  is re-projected to  $\mathbf{X}$  via  $\mathbf{Y}^{t+1} = \mathbf{Y}^t - vU(B(DB\mathbf{Y}^t - \mathbf{X}))$  where  $U$  is the up-sampling matrix. The process stops when  $|\mathbf{Y}^{t+1} - \mathbf{Y}^t|$  is smaller than a threshold. We use  $v=0.02$  and  $10^{-5}$  as the difference threshold, as proposed in [2].

## 3. Baseline Reconstruction Method and Iris Comparators

The reconstruction method described in Section 2 is evaluated against bilinear and bicubic interpolation, as well as the eigen-patch hallucination method of [2]. In the latter method, input LR patches are projected to collocated LR training patches via PCA eigen-transformation. Once the optimal reconstruction weights of a given patch are obtained in the LR manifold, the corresponding HR patch is super-resolved using the same weights, but with collocated HR training patches. Contrarily to the method of Section 2, the reconstruction weights are obtained only from LR training patches, and they are simply transferred to the HR manifold. Thus, the HR manifold is not used to find the optimal weights.

We also conduct iris recognition experiments using two different systems based on 1D Log-Gabor filters (LG) [18] and the SIFT operator [17]. In LG, the iris region is first unwrapped to a normalized rectangle of  $20 \times 240$  pixels using the Daugman’s rubber sheet model [8] and next, a 1D Log-Gabor wavelet is applied plus phase binary quantization to 4 levels. Comparison between binary vectors is done using the normalized Hamming distance [8], which incorporates noise mask, so only significant bits are used. In the SIFT method, 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 key-points in the two images under comparison. The LG implementation is from Libor Masek code [18], using its default parameters (optimized for CASIA images, which we employ here as well). The SIFT method uses a free toolkit for feature extraction and matching<sup>1</sup>, with the adaptations described in [4] (particularly, it includes a post-processing step to remove spurious matching points using geometric constraints). The iris region and corresponding noise mask for feature extraction and matching is obtained by manual annotation of the database used,

as shown in Figure 2 (more information is provided in the experimental setup).

## 4. Experimental Framework

### 4.1. Database and Protocol

We use the CASIA Interval v3 iris database [7]. It has 2,655 NIR images of  $280 \times 320$  pixels from 249 contributors captured in 2 sessions with a close-up iris camera, totalling 396 different eyes (the number of images per contributor / per session is not constant). Manual annotation of the database is available [1, 10], which is used as input for our experiments. All images are resized via bicubic interpolation to have the same sclera radius (we choose as target radius the average sclera radius  $R=105$  of the whole database, given by the groundtruth). Then, images are aligned by extracting a region of  $231 \times 231$  around the pupil center (corresponding to about  $1.1 \times R$ ). If extraction is not possible (for example if the eye is close to an image side), the image is discarded. After this procedure, 1,872 images remain, which will be used for our experiments.

The dataset of aligned images has been divided into two sets, a training set comprised of images from the first 116 users ( $M=925$  images) used to train the hallucination methods, and a test set from the remaining 133 users (947 images) which is used for validation. We perform verification experiments with the iris comparators in the test set. We consider each eye as a different user. Genuine comparison trials are obtained by comparing each image of a user to the remaining images of the same user, avoiding symmetric comparisons. Impostor trials are obtained by comparing the 1<sup>st</sup> image of a user to the 2<sup>nd</sup> image of the remaining users. With this procedure, we obtain 2,607 genuine and 19,537 impostor scores.

### 4.2. Results

The 947 test images are used as our high resolution (HR) reference images. We then down-sample these images via bicubic interpolation by a factor of  $2n$  (i.e. the image is resized to  $1/(2n)$  of the original HR size), and the down-sampled images are used as input LR images, from which hallucinated HR images are extracted. This simulated down-sampling is the approach followed in most of the previous super-resolution studies [24], mainly due to the lack of databases with low-resolution and corresponding high-resolution reference images. We test until a down-sampling factor of 16 (corresponding to a LR image size of  $15 \times 15$ ). We also extract the normalized iris region (size  $20 \times 240$ ) from both the hallucinated HR and the reference HR images, as well as LG and SIFT features, according to the algorithms of Section 3. The performance of the hallucination algorithm is measured by computing the Peak Signal-to-Noise Ratio (PSNR, in dBs) and the Structural

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



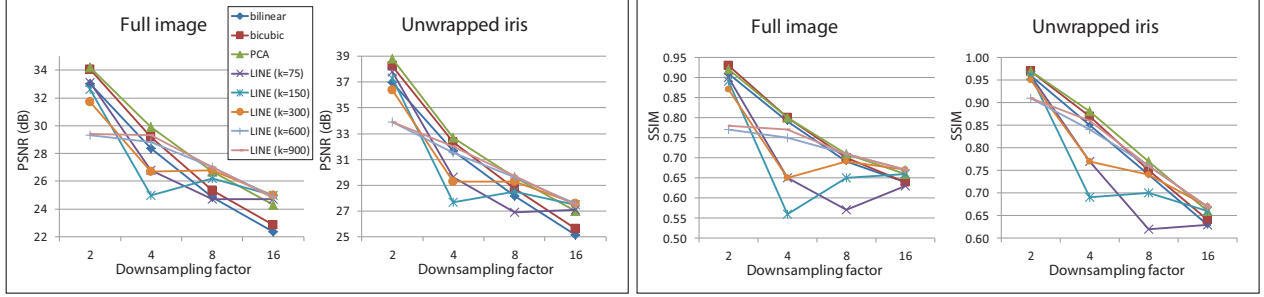


Figure 4. Hallucination results with different down-sampling factors (left: PSNR, right: SSIM). Best seen in color.

		Full image				Unwrapped iris region			
method		115×115 (1/2)	57×57 (1/4)	29×29 (1/8)	15×15 (1/16)	115×115 (1/2)	57×57 (1/4)	29×29 (1/8)	15×15 (1/16)
bilinear	psnr	33	28.36	24.86	22.39	36.94	31.64	28.18	25.16
	ssim	0.91	0.79	0.69	0.64	0.96	0.85	0.74	0.63
bicubic	psnr	34.04	29.18	25.33	22.86	38.22	32.35	28.74	25.64
	ssim	<b>0.93</b>	<b>0.8</b>	0.7	0.64	<b>0.97</b>	0.87	0.75	0.64
PCA	psnr	<b>34.2</b>	<b>29.9</b>	26.7	24.3	<b>38.8</b>	<b>32.7</b>	29.6	27
	ssim	0.92	<b>0.8</b>	<b>0.71</b>	0.66	<b>0.97</b>	0.88	0.77	0.66
M-LINE k=75	psnr	33.1	26.8	24.7	24.7	37.8	29.6	26.9	27.1
	ssim	0.9	0.65	0.57	0.63	0.96	0.77	0.62	0.63
M-LINE k=150	psnr	32.6	25	26.2	<b>25</b>	37.3	27.7	28.5	27.5
	ssim	0.89	0.56	0.65	0.66	0.96	0.69	0.7	0.66
M-LINE k=300	psnr	31.7	26.7	26.8	<b>25</b>	36.4	29.3	29.3	<b>27.6</b>
	ssim	0.87	0.65	0.69	<b>0.67</b>	0.95	0.77	0.74	<b>0.67</b>
M-LINE k=600	psnr	29.3	28.8	<b>27</b>	24.9	33.9	31.5	29.6	<b>27.6</b>
	ssim	0.77	0.75	<b>0.71</b>	<b>0.67</b>	0.91	0.84	<b>0.76</b>	<b>0.67</b>
M-LINE k=900	psnr	29.4	29.3	<b>27</b>	24.8	33.9	32	<b>29.7</b>	27.5
	ssim	0.78	0.77	<b>0.71</b>	<b>0.67</b>	0.91	0.86	<b>0.76</b>	<b>0.67</b>

Table 1. Hallucination results with different down-sampling factors for the different reconstruction methods (average values of test dataset).

Similarity index (SSIM) between the hallucinated HR image and the corresponding HR reference image. Results are shown in Table 1 and Figure 4. We also compare our method with bicubic and bilinear interpolation, as well as the eigen-patch (PCA) method of [2]. Figure 3 shows the hallucinated images. In PCA and M-LINE, we employ a patch size of 1/4 of the LR image size. This value is motivated by [2], where better results were obtained in general with bigger patch sizes. The patch size is defined in proportion to the dimensions of the LR image to ensure that they cover the same relative size across different scaling factors. Overlapping between patches is 1/3 of the patch size. The M-LINE method is tested using different values of  $K$ , from  $K=75$  (small neighbors set) to  $K=900$  (which corresponds to nearly the whole training set of images).

From the results of Table 1 and Figure 4, it can be seen

that the advantages of employing M-LINE are evident at very low resolutions (image size of  $15 \times 15$ ). Here, PSNR of M-LINE is 0.7 dB better than of PCA (on the whole image), and 2.14 dB better than bicubic interpolation. At a LR image size of  $29 \times 29$ , M-LINE is still better than PCA in 0.3 dB, and 1.67 dB better than bicubic. If down-sampling is not so severe (i.e. image size of  $115 \times 115$  or  $57 \times 57$ ), PCA becomes the best performing method. Regarding the neighbor number  $K$  of M-LINE, it can be observed that a bigger set is preferred, with higher PSNR and SSIM values obtained for  $K \geq 300$ . For  $K = 75$  or  $K = 150$ , it can be seen from Figure 4 that the performance of M-LINE is much poorer than the other methods, specially at intermediate down-sampling factors. If we look Figure 3, we observe that smaller values of  $K$  produces sharper reconstructed images, while a bigger  $K$  produces blurrier images. This is ex-

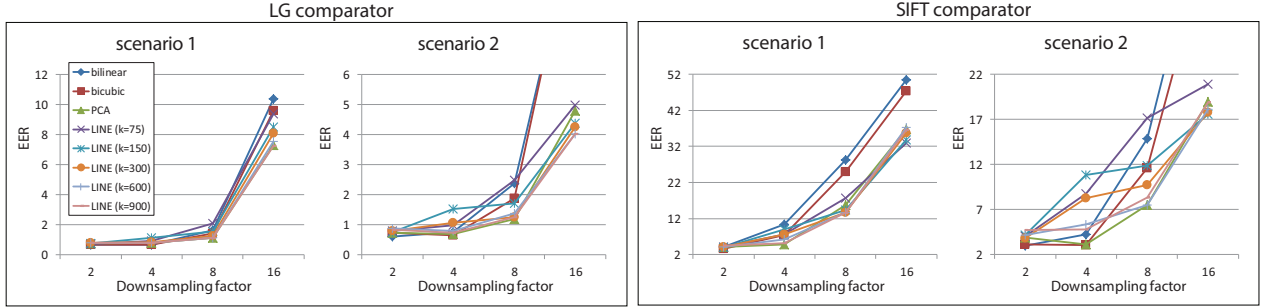


Figure 5. Verification results (EER) of the two scenarios employed. Best seen in color.

	Scenario 1				Scenario 2			
	115×115 (1/2)	57×57 (1/4)	29×29 (1/8)	15×15 (1/16)	115×115 (1/2)	57×57 (1/4)	29×29 (1/8)	15×15 (1/16)
<b>method</b>								
<b>LG COMPARATOR</b>								
bilinear	<b>0.69%</b>	0.69%	1.61%	10.39%	<b>0.61%</b>	0.76%	2.38%	11.03%
bicubic	<b>0.69%</b>	<b>0.68%</b>	1.42%	9.59%	0.73%	<b>0.65%</b>	1.88%	11.25%
PCA	0.76%	0.8%	<b>1.11%</b>	<b>7.29%</b>	0.73%	0.69%	<b>1.18%</b>	4.79%
M-LINE k=75	0.72%	0.93%	2.07%	9.38%	0.79%	0.99%	2.49%	4.97%
M-LINE k=150	0.76%	1.15%	1.53%	8.5%	0.76%	1.53%	1.71%	4.37%
M-LINE k=300	0.8%	0.85%	1.26%	8.08%	0.8%	1.07%	1.26%	4.26%
M-LINE k=600	0.8%	0.81%	1.15%	7.56%	0.87%	0.8%	1.38%	<b>4.02%</b>
M-LINE k=900	0.8%	0.79%	<b>1.11%</b>	<b>7.25%</b>	0.84%	0.73%	1.26%	<b>4.02%</b>
<b>method</b>								
<b>SIFT COMPARATOR</b>								
bilinear	4.05%	10.42%	28.23%	50.52%	<b>3.01%</b>	4.26%	14.82%	41.66%
bicubic	<b>3.51%</b>	7.41%	24.99%	47.33%	3.13%	<b>3.08%</b>	11.6%	36.37%
PCA	4.17%	<b>4.74%</b>	15.65%	36.67%	3.9%	3.11%	<b>7.46%</b>	18.98%
M-LINE k=75	4.13%	7.55%	17.62%	<b>32.99%</b>	4.04%	8.71%	17.16%	20.87%
M-LINE k=150	3.93%	8.95%	14.43%	33.7%	4.14%	10.85%	11.9%	<b>17.54%</b>
M-LINE k=300	4.12%	7.51%	13.73%	35.69%	3.81%	8.27%	9.73%	17.76%
M-LINE k=600	4.18%	6.11%	<b>13.61%</b>	37.17%	4.16%	5.32%	<b>7.47%</b>	18.13%
M-LINE k=900	4.51%	5.18%	13.72%	36.56%	4.7%	4.77%	8.34%	18.85%

Table 2. Verification results (EER) with different down-sampling factors for the different reconstruction methods.

pected, since a bigger value of  $K$  implies that more patches are being averaged (under their respective weights), so the output image patch will be smoother.

Next, we report verification experiments using hallucinated HR images. We consider two 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 Table 2 and Figure 5.

The following observations can be made from these results:

- While PCA and M-LINE have similar or worse performance than bilinear/bicubic interpolations at small down-sampling factors, their performance is much more better at very low resolutions, highlighting the benefits of trained reconstruction methods. In addition, M-LINE goes one step further, outperforming PCA when the image resolution is very small. Regarding performance of the individual comparators, SIFT is much more sensitive to resolution reduction. While EER of the LG comparator is kept at  $\sim 4\%$  in scenario 2, SIFT goes above 17%.

Scenario 1				Scenario 2			
115×115 (1/2)	57×57 (1/4)	29×29 (1/8)	15×15 (1/16)	115×115 (1/2)	57×57 (1/4)	29×29 (1/8)	15×15 (1/16)

method	FUSION LG+SIFT COMPARATORS							
bilinear	<b>0.61%</b>	0.76%	1.65%	10.28%	<b>0.61%</b>	0.76%	2.3%	10.98%
bicubic	0.69%	<b>0.7%</b>	1.36%	9.59%	<b>0.61%</b>	<b>0.7%</b>	1.73%	10.97%
PCA	0.65%	<b>0.7%</b>	1.12%	7.25%	0.65%	<b>0.7%</b>	<b>1.03%</b>	4.49%
M-LINE k=75 (LG, SIFT)	0.69%	0.82%	1.9%	9.28%	0.66%	1%	2.3%	4.88%
M-LINE k=150 (LG, SIFT)	0.69%	1.07%	1.51%	8.09%	0.65%	1.27%	1.65%	4.11%
M-LINE k=300 (LG, SIFT)	0.73%	0.84%	1.23%	7.9%	0.65%	1%	1.23%	3.96%
M-LINE k=600 (LG, SIFT)	0.7%	0.73%	<b>1.07%</b>	7.52%	0.76%	0.8%	1.09%	3.69%
M-LINE k=900 (LG, SIFT)	0.73%	0.76%	<b>1.07%</b>	7.21%	0.73%	0.76%	1.07%	3.83%
M-LINE k=900 (LG), k=75 (SIFT)	0.73%	0.73%	1%	<b>7.18%</b>	0.76%	0.73%	1.19%	3.83%
M-LINE k=900 (LG), k=150 (SIFT)	0.77%	0.76%	1.11%	7.25%	0.76%	0.73%	1.23%	3.68%
M-LINE k=900 (LG), k=300 (SIFT)	0.73%	0.73%	1.11%	7.21%	0.73%	0.76%	1.15%	<b>3.58%</b>

Table 3. Fusion results (EER) of the LG and SIFT comparators.

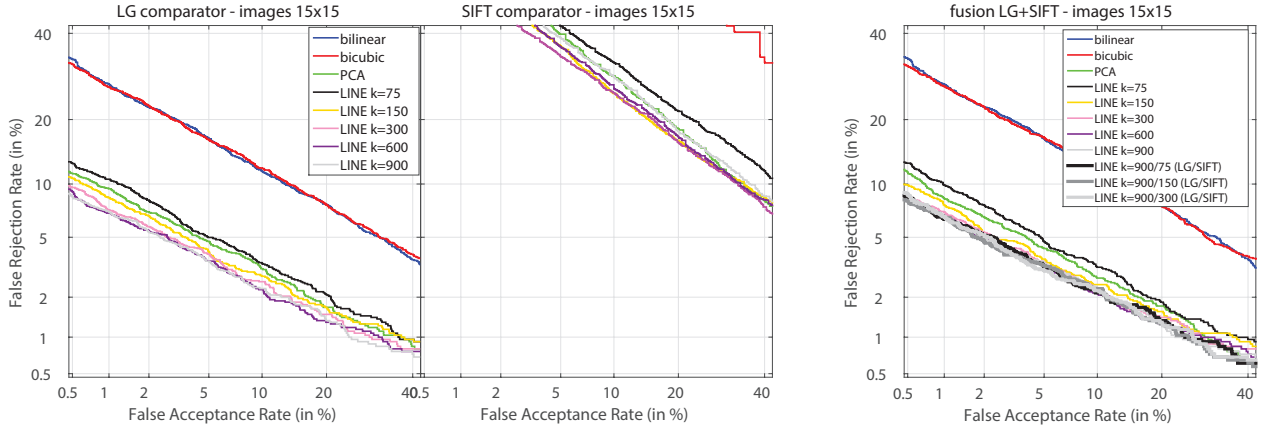


Figure 7. Fusion results (DET curves). Results are given for scenario 2 and LR image size of 15×15 only. Best seen in color.

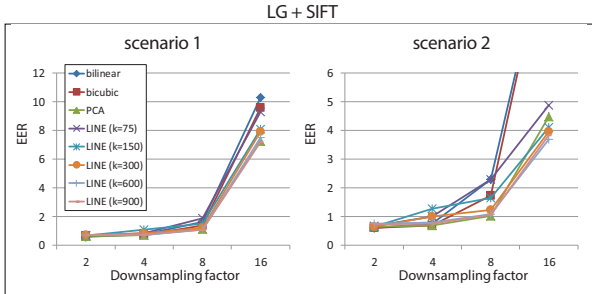


Figure 6. Fusion results (EER) of the two scenarios employed. Best seen in color.

- Interestingly, the preferred neighbor number  $K$  of M-LINE is different for each matcher. While LG prefers a bigger set ( $K > 300$ ), SIFT shows better results with a smaller set. This contradicts the re-

sults with PSNR/SSIM indicators, where a bigger set was preferred. These general-scene quality indicators are widely employed in the super-resolution literature, but they are more tailored to measure visual enhancement; however, the aim of applying super-resolution to biometrics is enhancing recognition performance [19]. For this reason, looking at the recognition performance is necessary, since the performance of an individual comparator may not follow.

- Both scenarios has a relatively similar performance up to a certain down-sampling factor. But at low resolutions, both comparators show a much more better performance in scenario 2. In the latter scenario, both gallery and probe images undergo the same down-sampling and reconstruction procedure. In scenario 1, the gallery image employed is the original HR image,

which seems to have fairly different feature properties than corresponding reconstructed LR images (at least with the feature extraction methods used here). Similarly, performance is not degraded until a down-sampling factor of 1/8, suggesting that the size of stored query images can be kept low without jeopardizing performance, as well as the size of the test image if it has to be sent through a communication channel.

We then carry out fusion experiments using linear logistic regression. Given  $N$  comparators ( $N=2$  in our case) 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 [3]. We use this trained fusion approach because it has shown better performance than simple fusion rules (like the mean or the sum rule) in previous works, as in the one reported above. Results are given in Table 3 and Figure 6. Since the preferred neighbor number  $K$  of M-LINE is different for each matcher, we have tested both the case where  $K$  is the same for both comparators, and also the case where the LG comparator uses  $K=900$  and SIFT varies from  $K=75$  and 150. As it can be observed, M-LINE also outperforms PCA here when the image has very low resolution. While the best EER with LG only is  $\sim 4\%$ , its fusion with SIFT pushes the EER until  $\sim 3.6\%$ , even if the individual performance of SIFT is  $\sim 17.5\%$ . We further analyze the benefit of M-LINE w.r.t. the other reconstruction methods in Figure 7, where we plot the DET curves (scenario 2) of the individual comparators, and of the fusion of the two systems for an input LR image size of  $15 \times 15$ . In this extreme case of very low resolution, it can be observed that the performance of PCA (green curves) is systematically improved by M-LINE for nearly any FRR/FAR value. The improvement of M-LINE w.r.t. the other methods is even more pronounced at high security regions (low FAR), where the iris modality is usually employed [6].

## 5. Conclusion

More relaxed acquisition environments are pushing the iris modality towards the use of low resolution imagery. Here, we apply an iris super-resolution reconstruction method based on Multilayer Locality-Constrained Iterative Neighbour Embedding of local image patches (M-LINE) [14] to increase the resolution of near-infrared (NIR) iris images. Previous works on patch-based iris reconstruction only consider the LR manifold to reconstruct input LR patches, and the reconstruction weights are simply transferred to the HR manifold [2]. In the method described in this paper, on the contrary, reconstruction starts in the HR manifold (which is unaffected by the image degradation process), and distances to the HR training dictionary are used to penalize the reconstruction weights of the input

LR patch in the LR space. This approach is compared to bilinear/bicubic interpolation, and to the mentioned method of [2], which is based on projecting input LR patches to the LR training set using a PCA eigen-transformation. We also carry out iris verification experiments on the reconstructed images using two iris comparators based on Log-Gabor (LG) wavelets and SIFT key-points. Experimental results show that the M-LINE method has a superiority over the other reconstruction methods when the resolution of the image is very small ( $15 \times 15$  in this paper), with an EER of  $\sim 4\%$  using the LG comparator, and an EER of  $\sim 3.6\%$  after the fusion of the two comparators. Another observation is that performance of the comparators is not degraded significantly until input images are down-sampled by 1/8 (image size of  $29 \times 29$ ), meaning that stored query images or transmitted test images can be kept low without sacrificing performance.

The PCA and M-LINE methods evaluated here assume that the low- and high-resolution manifolds have similar local geometrical structure, hence the same reconstruction weights are used in both manifolds. While this simplifies the problem, it is not usually the case, since the degradation process of the LR image results in one-to-many relationship between low- and high-resolution patches. To cope with this effect, we are exploring the update strategy of the LR dictionary proposed in the M-LINE method [14], which is not implemented in this paper. Another directions include evaluating the resiliency of other iris recognition algorithms [22] as well as employing low-resolution imagery captured in visible range (e.g. with smart-phones).

## Acknowledgements

Author F. A.-F. thanks the Swedish Research Council for funding his research. Authors acknowledge the CAISR program and the SIDUS-AIR project of the Swedish Knowledge Foundation.

## References

- [1] F. Alonso-Fernandez and J. Bigun. Near-infrared and visible-light periocular recognition with gabor features using frequency-adaptive automatic eye detection. *IET Biometrics*, 4(2):74–89, 2015.
- [2] F. Alonso-Fernandez, R. A. Farrugia, and J. Bigun. Eigen-patch iris super-resolution for iris recognition improvement. *Proc European Signal Processing Conference, EUSIPCO*, Sep 2015.
- [3] F. Alonso-Fernandez, J. Fierrez, D. Ramos, and J. Gonzalez-Rodriguez. Quality-based conditional processing in multi-biometrics: Application to sensor interoperability. *IEEE Trans. on Systems, Man and Cybernetics-Part A: Systems and Humans*, 40(6):1168–1179, 2010.
- [4] F. Alonso-Fernandez, P. Tome-Gonzalez, V. Ruiz-Albacete, and J. Ortega-Garcia. Iris recognition based on sift features.



*Proc IEEE Intl Conf Biometrics, Identity and Security, BIDS*, 2009.

- [5] S. Baker and T. Kanade. Limits on super-resolution and how to break them. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(9):1167–1183, Sep 2002.
- [6] K. Bowyer, K. Hollingsworth, and P. Flynn. Image understanding for iris biometrics: a survey. *Computer Vision & Image Understanding*, 110:281–307, 2007.
- [7] CASIA databases. <http://biometrics.idealtest.org/>.
- [8] J. Daugman. How iris recognition works. *IEEE Trans. on Circuits and Systems for Video Technology*, 14:21–30, 2004.
- [9] A. Deshpande, P. P. Patavardhan, and D. H. Rao. Super-resolution for iris feature extraction. *Proc Intl Conf on Computational Intelligence and Computing Research, IC-CIC*, pages 1–4, 2014.
- [10] H. Hofbauer, F. Alonso-Fernandez, P. Wild, J. Bigun, and A. Uhl. A ground truth for iris segmentation. *Proc Intl Conf Pattern Recognition, ICPR*, 2014.
- [11] K. Hollingsworth, T. Peters, K. Bowyer, and P. Flynn. Iris recognition using signal-level fusion of frames from video. *IEEE Transactions on Information Forensics and Security*, 4(4):837–848, 2009.
- [12] J. Huang, L. Ma, T. Tan, and Y. Wang. Learning based resolution enhancement of iris images. *Proc. BMVC*, 2003.
- [13] A. K. Jain and A. Kumar. *Second Generation Biometrics*, chapter Biometrics of Next Generation: An Overview. Springer, 2011.
- [14] J. Jiang, R. Hu, Z. Wang, and Z. Han. Face super-resolution via multilayer locality-constrained iterative neighbor embedding and intermediate dictionary learning. *IEEE Transactions on Image Processing*, 23(10):4220–4231, Oct 2014.
- [15] R. Jillela, A. Ross, and P. Flynn. Information fusion in low-resolution iris videos using principal components transform. *Proc. IEEE Workshop on Applications of Computer Vision, WACV*, pages 262–269, Jan 2011.
- [16] B. Li, H. Chang, S. Shan, and X. Chen. Locality preserving constraints for super-resolution with neighbor embedding. *Proc Intl Conf Image Processing, ICIP*, pages 1189–1192, 2009.
- [17] D. Lowe. Distinctive image features from scale-invariant key points. *Intl Journal of Computer Vision*, 60(2):91–110, 2004.
- [18] L. Masek. Recognition of human iris patterns for biometric identification. Master’s thesis, School of Computer Science and Software Engineering, University of Western Australia, 2003.
- [19] K. Nguyen, S. Sridharan, S. Denman, and C. Fookes. Feature-domain super-resolution framework for gabor-based face and iris recognition. In *Proc IEEE Conf on Computer Vision and Pattern Recognition, CVPR*, pages 2642–2649, Jun 2012.
- [20] I. Nigam, M. Vatsa, and R. Singh. Ocular biometrics: A survey of modalities and fusion approaches. *Information Fusion*, 26(0):1 – 35, 2015.
- [21] S. C. Park, M. K. Park, and M. G. Kang. Super-resolution image reconstruction: a technical overview. *Signal Processing Magazine, IEEE*, 20(3):21–36, May 2003.
- [22] C. Rathgeb, A. Uhl, and P. Wild. *Iris Biometrics - From Segmentation to Template Security*, volume 59 of *Advances in Information Security*. Springer, 2013.
- [23] 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 *Intl Conf Ubiquitous Information Technologies Applications, ICUT*, pages 1–5, Dec 2009.
- [24] N. Wang, D. Tao, X. Gao, X. Li, and J. Li. A comprehensive survey to face hallucination. *Intl Journal of Computer Vision*, 106(1):9–30, 2014.