

Improving Very Low-Resolution Iris Identification Via Super-Resolution Reconstruction of Local Patches

Fernando Alonso-Fernandez¹, Reuben A. Farrugia², Josef Bigun³

Abstract: Relaxed acquisition conditions in iris recognition systems have significant effects on the quality and resolution of acquired images, which can severely affect performance if not addressed properly. Here, we evaluate two trained super-resolution algorithms in the context of iris identification. They are based on reconstruction of local image patches, where each patch is reconstructed separately using its own optimal reconstruction function. We employ a database of 1,872 near-infrared iris images (with 163 different identities for identification experiments) and three iris comparators. The trained approaches are substantially superior to bilinear or bicubic interpolations, with one of the comparators providing a Rank-1 performance of $\sim 88\%$ with images of only 15×15 pixels, and an identification rate of 95% with a hit list size of only 8 identities.

Keywords: Iris, biometrics, super-resolution, low resolution.

1 Introduction

While the literature on image super-resolution is ample, its application to biometrics is relatively recent, with most research concentrated on face reconstruction [Wa14]. However, a number of applications which are becoming ubiquitous, such as surveillance or smartphone biometrics, have the lack of pixel resolution as one of their most evident problems when acquisition is done distantly. One reason of such limited research might be that most super-resolution approaches are general-scene, aimed at producing overall visual enhancement, which does not necessarily correlate with better recognition performance [Ng12]. Thus, adaptation of super-resolution techniques to the particularities of images from a specific biometric modality is needed to achieve a more efficient up-sampling [BK02].

This paper investigates two trained super-resolution approaches based on PCA Eigen transformation (eigen-patches) [AFB15] and Locality-Constrained Iterative Neighbor Embedding (LINE) of local image patches [Ji14] in the context of iris identification. The methods employed make use of coupled dictionaries to learn the mapping relation between low- and high-resolution image pair in order to hallucinate a high-resolution image from the observed low-resolution one. This *learning-based* strategy has the advantage of only needing one low-resolution image as input, and usually allow higher magnification factors than *reconstruction-based* methods, which fuse several low-resolution images into a high-resolution one [PPK03]. Another particularity of the evaluated methods is that they

¹ School of ITE, Halmstad University, Sweden, feralo@hh.se

² Department of CCE, University of Malta, Malta, reuben.farrugia@um.edu.mt

³ School of ITE, Halmstad University, Sweden, josef.bigun@hh.se

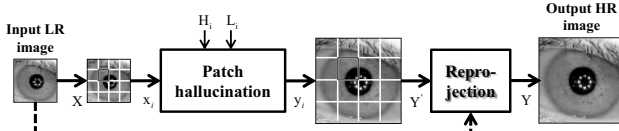


Fig. 1: Block diagram of patch-based hallucination.

use a patch-based approach, where overlapped local image patches are reconstructed separately, and then stitched together. This better represents local details and preserves texture than if reconstruction of the complete image was done at a time, since each patch has its own optimal reconstruction function. In our experiments, we employ the CASIA-IrisV3-Interval database [CA] of NIR iris images, with low-resolution images having a size of only 15×15 pixels. Identification experiments are conducted with three iris comparators based on 1D Log-Gabor filters (LG) [Ma03], SIFT key-points [Lo04], 5 and local intensity variations of iris textures (CR) [RU10]. LG and CR exploit texture information globally (across the entire image), while SIFT exploits local features in discrete key-points. Thus, one motivation is to employ features that are diverse in nature. Despite the patch-based approaches used are not new [AFB15, Ji14], we contribute with its evaluation in the context of iris identification, and particularly with the application of these three iris comparators to the reconstructed images. Reported results show the superiority of the two trained reconstruction approaches w.r.t. bicubic or bilinear interpolations, with an impressive Rank-1 performance of $\sim 88\%$ with the LG comparator under such very low resolution.

2 Reconstruction of Low Resolution Iris Images

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 $\bar{\mathbf{X}} = DBW\bar{\mathbf{Y}} + \bar{\mathbf{n}}$ (where $\bar{\mathbf{n}}$ represents additive noise). For simplicity, W and $\bar{\mathbf{n}}$ are usually omitted, leading to $\bar{\mathbf{X}} = DB\bar{\mathbf{Y}}$. In local patch-based methods (Figure 1), 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 a predefined patch size and overlap pixels (N_v and N_h are the vertical and horizontal number of patches). Since we will consider square images, we assume that $N_v = N_h$. 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 from $\{\mathbf{L}\}$. Each individual LR patch \mathbf{x}_i is then hallucinated using the dictionaries \mathbf{H}_i and \mathbf{L}_i , producing the corresponding HR patch \mathbf{y}_i .

Eigen-Patch Reconstruction Method (PCA). This method is described in [AFB15], which is based on the algorithm for face images of [CC14]. Here, a PCA eigen-transformation is conducted in the set of LR basis patches \mathbf{L}_i . Given an input LR patch \mathbf{x}_i , it is then projected onto the eigenpatches of \mathbf{L}_i , obtaining the optimal reconstruction weights $\mathbf{c}_i = \{c_i^1, c_i^2, \dots, c_i^M\}$ of \mathbf{x}_i w.r.t. \mathbf{L}_i . The reconstruction weights are then carried on to

weight the HR basis set, and the HR patch is super-resolved as $\mathbf{y}_i = \mathbf{H}_i \mathbf{c}_i^T$. 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 preliminary reconstructed HR image \mathbf{Y}' .

Locality-Constrained Iterative Neighbour Embedding Method (LINE). This is based on the algorithm for face images of [Ji14]. Instead of using all entries of the training dictionary to estimate the reconstruction weights, a set of $K < M$ entries is used. Using all entries can result in over-smooth reconstructed images which lacks important texture information, which is essential for iris. Given a LR patch \mathbf{x}_i , a first estimate of the HR patch $\mathbf{v}_{i,0}$ is initialized by bicubic up-scaling. Then, an iterative loop indexed by $j \in [0, J-1]$ is started. For every iteration, the support \mathbf{s} of \mathbf{H}_i that minimizes the distance $\mathbf{d} = \|\mathbf{v}_{i,j} - \mathbf{H}_i(\mathbf{s})\|_2^2$ is computed using K -nearest neighbours. The combination weights are then derived using

$$\mathbf{w}_{i,j}^* = \arg \min_{\mathbf{w}_{i,j}^*} (\|\mathbf{x}_i - \mathbf{L}_i(\mathbf{s}) \mathbf{w}_{i,j}^*\|_2^2 + \tau \|\mathbf{d}(\mathbf{s}) \odot \mathbf{w}_{i,j}^*\|_2^2) \quad (1)$$

where τ is a regularization parameter. Operator \odot (element-wise multiplication) 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 . This optimization problem can be solved by an analytic solution [Ji14]. The estimated HR patch is then updated using $\mathbf{v}_{i,j+1} = \mathbf{H}_i(\mathbf{s}) \mathbf{w}_{i,j}^*$ and the loop is repeated. The final estimate of the HR patch is then derived using $\mathbf{y}_i = \mathbf{v}_{i,J}$. We employ $\tau=1e^{-5}$ and $J=4$ [Ji14]. Contrarily to the PCA method, where reconstruction weights are obtained in the LR manifold and then simply transferred to the HR manifold, note that Equation 1 jointly considers the LR manifold (via $\mathbf{x}_i, \mathbf{L}_i(\mathbf{s})$) and the HR counterpart (via $\mathbf{d}(\mathbf{s})$) during the reconstruction. In addition, reconstruction starts in the HR manifold, which is not affected by the degradation process, and computation of the K nearest neighbors employed for reconstruction is done in this manifold as well.

Image Reprojection. Inspired by [AFB15], we incorporate a re-projection step to \mathbf{Y}' 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 $\mathbf{Y}^{t+1} = \mathbf{Y}' - vU(B(DB\mathbf{Y}' - \mathbf{X}))$ where U is the up-sampling matrix. The process stops when $|\mathbf{Y}^{t+1} - \mathbf{Y}'| \leq \varepsilon$. We use $v=0.02$ and $\varepsilon = 10^{-5}$ [AFB15].

3 Experimental Framework

We use CASIA Interval v3 iris database [CA]. It has 2,655 NIR images of 280×320 pixels from 249 contributors captured with a close-up camera. Manual annotation is available, including iris circles and noise mask (Figure 2) [AB15, Ho14], which is used as input for our experiments. All images are resized by bicubic interpolation to have the same sclera radius ($R=105$, average of the database given by the ground-truth). Then, images are aligned by extracting a region of 231×231 around the pupil center (corresponding to $\sim 1.1 \times R$). If extraction is not possible (for example if the eye is close to a boundary), the image is discarded. After this procedure, 1,872 images remain, which are then divided into two sets, a training set with 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) for

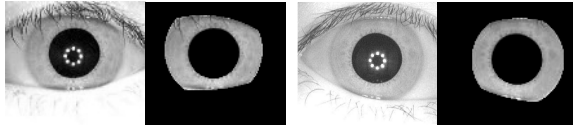


Fig. 2: Example of images of the CASIA Interval v3 database with the annotated circles modeling iris boundaries and eyelids.

validation. We carry out identification experiments with three iris comparators in the test set. From the 133 users, we select those eyes having at least two samples, resulting in 163 different eyes (i.e. identities) and 927 images. The first sample of each eye is considered as enrolment sample, and the remaining 764 samples are used as input for identification. This results in $764 \times 163 = 124,532$ comparisons. Given an input sample, identification is done by outputting the N closest identities of the enrolment set. An identification is considered successful if the correct identity is among the N outputted ones.

The iris comparators used are based on 1D Log-Gabor filters (LG) [Ma03], SIFT operator [Lo04], and local intensity variations in iris textures (CR) [RU10]. In LG, the iris region is first unwrapped to a normalized rectangle of 20×240 pixels [Da04] 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 [Da04]. 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 CR method starts by unwrapping the iris to a rectangle of 64×512 pixels, and then it traces intensity variations across horizontal stripes of distinct height, encoding the paths where the minimum and maximum grey values of each column occur. The LG implementation is from Libor Masek [Ma03], using its default parameters. The SIFT method uses a free toolkit³, with adaptations described in [Al09] to remove spurious matchings. The CR algorithm is from the University of Salzburg Iris Toolkit (USIT) [RUW13].

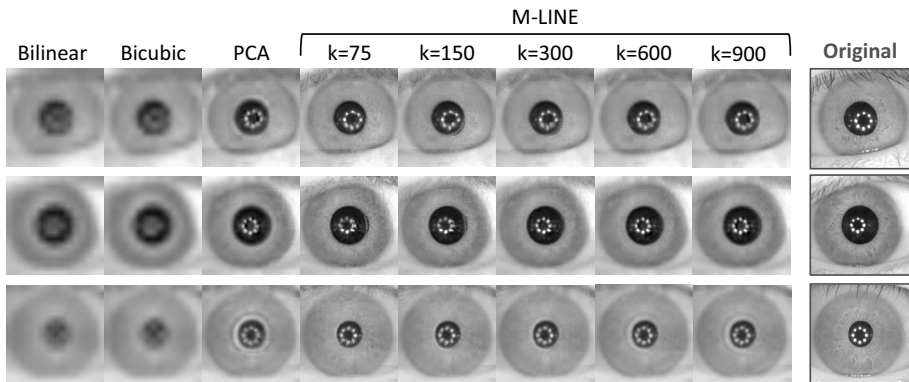


Fig. 3: Resulting hallucinated HR images. The original HR image is also shown.

³ <http://vision.ucla.edu/vedaldi/code/sift/assets/sift/index.html>

4 Results

The two reconstruction methods are evaluated together with bilinear and bicubic interpolations. The 947 validation images are used as HR reference images. They are down-sampled via bicubic interpolation to a size of 15×15 , corresponding to a down-sampling factor of 16, and then used as input LR images of the reconstruction methods, from which hallucinated HR images are computed. This simulated down-sampling is the approach followed in most previous studies [Wa14], mainly due to the lack of databases with LR and corresponding HR reference images. In PCA and LINE, we employ a patch size of $1/4$ of the LR image size. This is motivated by [AFB15], where better results were obtained with bigger patch sizes. Overlapping between patches is $1/3$ of the patch size. We also extract the LG, SIFT and CR features from both the hallucinated HR and the reference HR images. Figure 3 shows some examples of reconstructed images with the different methods tested here. It can be observed that smaller values of K results in sharper reconstructed images, while a bigger K produces blurrier images. This is expected, since a bigger K implies that more patches are being averaged, so the output image patch will be smoother.

The performance of the reconstruction methods is measured by reporting identification experiments using hallucinated images. We do not report other measures traditionally used in super-resolution literature (e.g. PSNR) since the aim of applying these algorithms in biometrics is enhancing recognition performance [Ng12]. Two scenarios are considered: 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, while the second case simulates a totally uncontrolled scenario (albeit for simplicity, both samples have similar resolution). We first test the LINE method using different values of K , from $K=75$ (small neighbors set) to $K=900$ (nearly the whole training set). Identification results are given in Figure 4. It can be seen that the preferred neighbor size K is different for each comparator. While LG and CR prefer a bigger set ($K > 300$), SIFT shows better results with a smaller set ($K = 150$). This highlights the need of looking into the performance of individual comparators, rather than into general scene indicators such as PSNR, since the image properties recovered by a particular algorithm may not be relevant for a comparator, even if visual appearance of the reconstructed image can be referred as ‘good’.

We then select the best LINE configurations for each comparator, and report identification results together with the other reconstruction methods (Figure 5). Our first observation is the superior performance of PCA and LINE w.r.t. bilinear or bicubic interpolation, highlighting the benefits of trained reconstruction. Also, LINE is superior to PCA in some cases, while in others, both methods show similar performance. In this sense, PCA can be pre-trained in advance using the set \mathbf{L}_i of basis patches, since eigenpatches are the same for any input patch \mathbf{x}_i , so higher computational speeds can be expected. LINE on the other hand needs to compute the set of nearest neighbors specific of a particular input patch.

Regarding performance of individual comparators, LG is clearly superior to the others. Rank-1 performance of LG is above 70% (scenario 1) and 84% (scenario 2). Also, an identification rate of 95% with this comparator is obtained for a hit list size of just $N=8$ (sce-

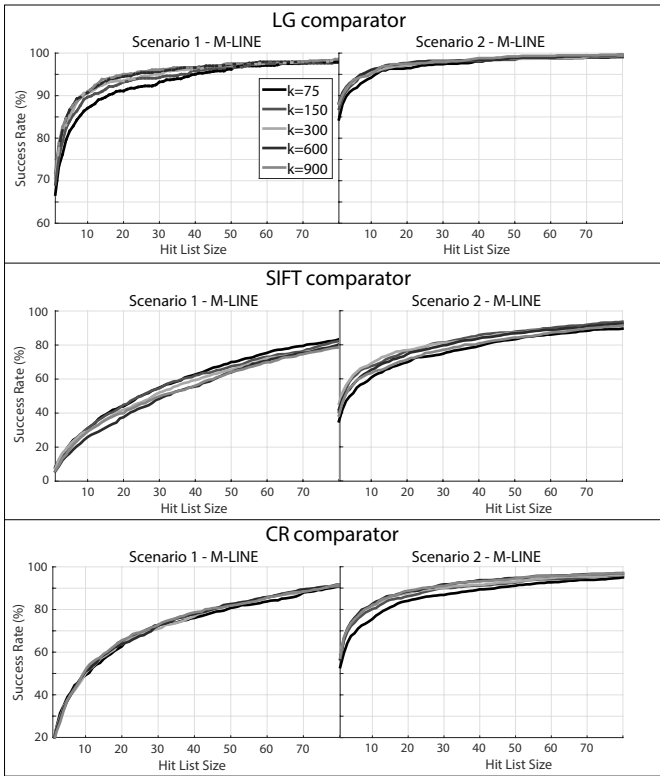


Fig. 4: Identification results (LINE method). Best seen in colour.

nario 2) using LINE. Rank-1 of SIFT is very poor (less than 10% in scenario 1 and $\sim 40\%$ in scenario 2), while an identification rate of 95% cannot be achieved even if $N > 80$. The CR comparator only does a little bit better than SIFT. It should be noted however that the size of the LR images is very small (15×15). With respect to the two scenarios evaluated, scenario 2 has much better performance. In scenario 2, both enrolment and query images undergo the same down-sampling and reconstruction. It seems that when the two images do not suffer the same degradation process (i.e. scenario 1), they have fairly different feature properties, at least with the features employed here. This result has been observed in previous verification studies [AFB15] as well.

5 Conclusions

While more relaxed acquisition environments are pushing image-based biometrics (e.g. face or iris) towards the use of low resolution imagery, it can pose significant problems in terms of reduced performance if not addressed properly. Here, we apply two trained super-resolution approaches based on PCA transformation [AFB15] and Locality-Constrained Iterative Neighbor Embedding (LINE) of local patches [Ji14] to improve the resolution

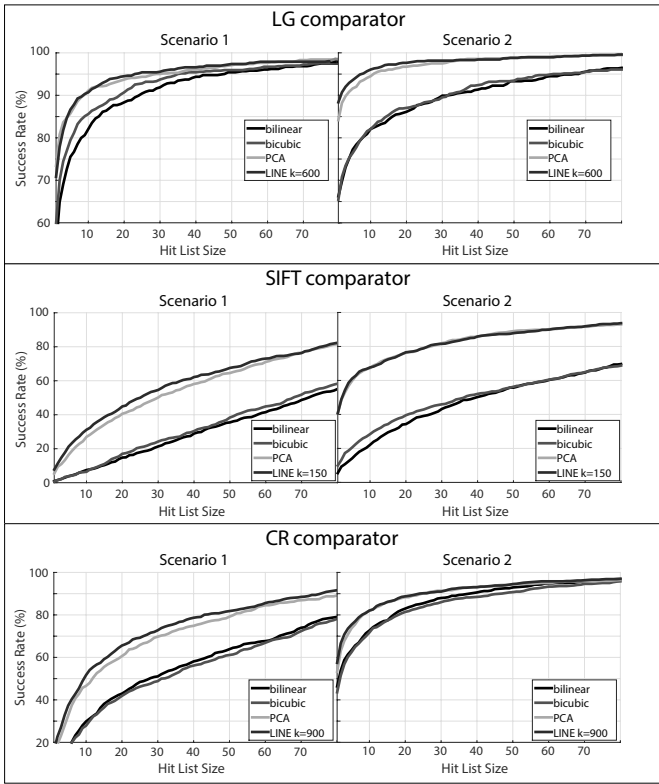


Fig. 5: Identification results of the different image reconstruction methods employed (LINE method: best case according Figure 4 is shown). Best seen in colour.

of iris images under infra-red lightning. We carry out identification experiments on the reconstructed images with three iris comparators based on Log-Gabor wavelets (LG), SIFT keypoints, and local intensity variations of iris textures (CR). Low resolution images are simulated by down-sampling high-resolution irises to a size of just 15×15 . Experimental results show a clear superiority of trained approaches under such challenging conditions w.r.t. bilinear or bicubic methods. Even under such low resolution, a Rank-1 performance of $\sim 88\%$ is obtained with one of the comparators (LG), and an identification rate of 95% is obtained with a hit list size of just 8. Another observation is that the LINE method is superior to PCA in some cases, but their performance is in general very similar. This allows computational savings by using PCA, since PCA models are the same for any input image, so they can be trained in advance.

An avenue of improvement is removing the assumption that reconstruction weights are the same in the low- and high-resolution manifolds. While this simplifies the problem, the LR manifold is usually distorted by the one-to-many relationship between LR and HR patches [Wa14]. Another simplification is the assumption of linearity in the combination of

patches from the training dictionary. We will also consider including additional recognition methods [RUW13] and employing imagery in visible range (e.g. 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

- [AB15] Alonso-Fernandez, F.; Bigun, J.: NIR & visible-light periocular recognit with Gabor features using frequency-adaptive automatic eye detection. *IET Biometrics*, 4(2), 2015.
- [AFB15] Alonso-Fernandez, F.; Farrugia, R. A.; Bigun, J.: Eigen-Patch Iris Super-Resolution for Iris Recognition Improvement. *Proc EUSIPCO*, Sep 2015.
- [Al09] Alonso-Fernandez, F.; Tome-Gonzalez, P.; Ruiz-Albacete, V.; Ortega-Garcia, J.: Iris Recognition Based on SIFT Features. *Proc IEEE BIDS*, 2009.
- [BK02] Baker, S.; Kanade, T.: Limits on super-resolution and how to break them. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(9):1167–1183, Sep 2002.
- [CA] CASIA databases: <http://biometrics.idealtest.org/>.
- [CC14] Chen, Hong-Yuh; Chien, Shao-Yi: Eigen-patch: Position-patch based face hallucination using eigen transformation. *Proc IEEE*, 2014.
- [Da04] Daugman, J.: How iris recognition works. *IEEE Trans. on CSVT*, 14:21–30, 2004.
- [Ho14] Hofbauer, H.; Alonso-Fernandez, F.; Wild, P.; Bigun, J.; Uhl, A.: A Ground Truth for Iris Segmentation. *Proc Intl Conf Pattern Recognition, ICPR*, 2014.
- [Ji14] Jiang, Junjun; Hu, Ruimin; Wang, Zhongyuan; Han, Zhen: 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.
- [Lo04] Lowe, D.: Distinctive image features from scale-invariant key points. *Intl Journal of Computer Vision*, 60(2):91–110, 2004.
- [Ma03] Masek, Libor: Recognition of human iris patterns for biometric identification. MSc thesis, School of Computer Science / Software Engineering, Univ Western Australia, 2003.
- [Ng12] Nguyen, K.; Sridharan, S.; Denman, S.; Fookes, C.: Feature-domain super-resolution framework for Gabor-based face and iris recognition. *Proc CVPR*, 2012.
- [PPK03] Park, Sung Cheol; Park, Min Kyu; Kang, Moon Gi: Super-resolution image reconstruction: a technical overview. *Signal Processing Magazine, IEEE*, 20(3):21–36, May 2003.
- [RU10] Rathgeb, Christian; Uhl, Andreas: Secure Iris Recognition Based on Local Intensity Variations. *Proc ICIAR*, 6112:266–275, 2010.
- [RUW13] Rathgeb, Christian; Uhl, Andreas; Wild, Peter: Iris Biometrics - From Segmentation to Template Security, volume 59 of *Advances in Information Security*. Springer, 2013.
- [Wa14] Wang, N.; Tao, D.; Gao, X.; Li, X.; Li, J.: A Comprehensive Survey to Face Hallucination. *Intl Journal of Computer Vision*, 106(1):9–30, 2014.