Learning to Invert Local Binary Patterns

Felix Juefei-Xu
felixu@cmu.edu
Marios Savvides
msavvid@ri.cmu.edu

Carnegie Mellon University
Pittsburgh, Pennsylvania 15213
USA

Abstract

In this work, we have proposed to invert the local binary patterns (LBP) descriptor. The success of the inversion gives rise to two applications: face de-appearance and re-appearance. The de-appearance, based on image-LBP forward mapping, is thorough in the sense that not only the identity information but also the soft-biometric information of the subject is removed. The re-appearance yields face reconstruction with high fidelity and also enables secure application with a unique encryption key. The re-appearance is achieved by learning the inverse mapping of the LBP descriptors through an ℓ_0-constrained coupled dictionary learning scheme that jointly learns two overcomplete dictionaries in both the pixel and the LBP domains such that inverse mapping from the LBP domain to the pixel domain is made possible without knowing the mapping function explicitly. The procedure also comes naturally with high selectivity when reconstructing the faces with various LBP encryption keys. We have shown the effectiveness of our proposed approach on the FRGC ver 2.0 database which involves large-scale fidelity test and face verification experiments using the state-of-the-art commercial and academic face matchers.

1 Introduction

Local feature descriptors such as SIFT [39, 40], HOG [6], BRIEF [5], FREAK [1], and LBP [45] are extremely popular in computer vision community nowadays. They map a local image patch to a highly non-linear representation that is usually invariant or robust to certain variations, giving rise to higher discriminability or encoding more robust local information that is suitable for certain computer vision tasks such as detection and recognition.

Since 2011, researchers have been trying to invert such a mapping, going from the descriptor to the image. One of the earliest work is pioneered by Weinzaepfel *et al*. [58] on reconstructing an image from its SIFT descriptors. In 2012, E. d'Angelo *et al*. [7] reconstruct images from BRIEF and FREAK descriptors, with their later work in [8] showing improved methods. In 2013, Vondrick *et al*. [56, 57] visualize HOG descriptor. A more recent work in 2015 attempts to invert the CNN feature [41] for better understanding the deep image representations.

There have not been studies on inverting the LBP [45] descriptor. In this work, we will invoke a coupled dictionary learning paradigm with an ℓ_0-constrained optimization to learn the inverse mapping from the LBP glyph to the original image. Since LBP is invented and gains popularity in face recognition community [20, 35, 36, 51], we decide to utilize face

Pages 29.1-29.14
DOI: https://dx.doi.org/10.5244/C.30.29

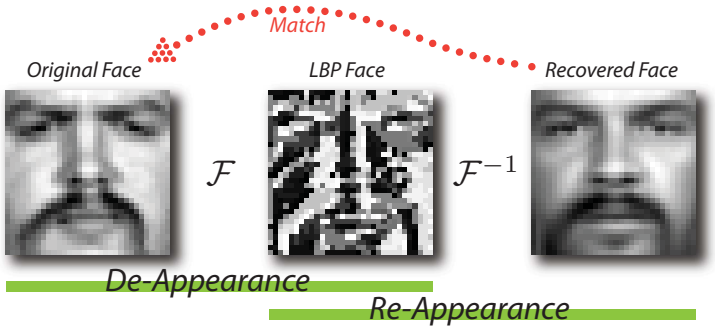


Figure 1: Flowchart of the method. The de-appearance step uses LBP for the forward mapping $\mathcal{F}$ and obtains the LBP face (glyph) in the middle. The re-appearance step tries to learn the inverse mapping $\mathcal{F}^{-1}$ from the LBP domain to the pixel domain. The recovered face is of high fidelity as compared to the original face.

images for showing the effectiveness of our proposed method. It turns out, different LBP encoding schemes lead to very much different LBP glyphs, making the inverse mapping that we learn highly selective. This implies that, if the inverse mapping is learned on one particular LBP encoding scheme, it will only well recover the original image for a query LBP glyph under the *same* encoding scheme, but not a different one. Also, LBP glyph is naturally an ideal way of hiding some of the image details, while preserving rough structure of the underlying content so that people are still able to tell what is been encoded, but not the fine details such as the identity and attributes of the subject. Therefore, we will apply the entire pipeline (forward mapping from image to LBP glyph and inverse mapping from LBP glyph to image) to face de-appearance and re-appearance tasks.

The study of privacy has gained prominence in the past decade along with the advancement of both the camera hardware and surveillance technology as well as more robust facial recognition systems (FRS) which can deal with more unconstrained scenarios. Latest FRS can easily identify subjects even from a very low resolution footage with off-angle faces and other non-ideal variations such as harsh lighting and facial occlusions. Therefore, applications that require identities of people in the scene be hidden shall resort to de-appearance techniques, *e.g.* Google Streetview and sensitive medical face databases.

Current methods of face de-appearance can be categorized as follows. The naive approach is to blur or blackout the entire face or the identifiable portion (periocular region [16, 17, 25, 27, 33, 37]) of the face. Such method either loses the entire information (blackout), or can be easily attacked by super-resolution or de-blurring techniques to recover the true identity. A family of more sophisticated approach, called *k-same*, de-appearances the query face image by anonymizing it among at least k candidates through low rank approximation in some domains such as the pixel domain, eigenvector domain, or active appearance model coefficients domain [43], [11], [10], [12], [13], [50]. Such methods yield much better visual appearance for the de-appearanced face image. However, one of the biggest caveats is that we can almost never recover the true identity of the subject from the de-appearanced face. Moreover, many of these approaches choose not to perform a thorough de-appearance. For example, soft-biometric traits such as skin color, ethnicity, gender, age, hair style, *etc.* are still showing in the de-appearanced image, which may be a disadvantage when it comes to sensitive information related applications. A more recent work [9] utilizes various soft-biometric classifiers to preserve high utility after de-appearance. However, the recovery based on the de-appearanced face image is, again, next to impossible.

Scope: Our face de-appearance paradigm is fundamentally different from many other de-

identification methods [9, 15] which preserve the face utility or attribute, in the following two major areas: (1) using others' method, once the face is de-appearanced, though the utility is preserved, the original face or subject information can never be recovered, while ours can still recover the identity information with very high fidelity. (2) We aim for 'thorough' face de-appearance such that the de-appearanced face contains no soft-biometrics traits of the subject, while theirs choose to preserve the face utility. The de-appearance problem we study in this work is quite different from previously studied ones, although we still use quite similar terminology. It is hard to put our method and others' under the same rooftop for comparison, because we aim at different goals and each is suited for different applications.

The re-appearance studied in this work also differs from other re-identification work. In literature, especially in surveillance community, re-identification usually refers to the following: "Given an image / video of a person taken from one camera, re-identification is the process of identifying the person from images / videos taken from a different camera [3]." Here, our 'face re-appearance' relates more to 'face recovery' which aims at reconstructing the pixel domain faces from their LBP glyphs, which were first de-appearanced using LBP forward mapping.

Specifically, the two applications to be discussed in this work are (1) a thorough face de-appearance method, and (2) a secure face re-appearance method. Generally speaking, the more thorough the de-appearance step is, the harder it is for re-appearance to recover the identity of the face. One extreme example would be blackout. There exists a trade-off between the thoroughness of the de-appearance step and the recovery capability of the re-appearance step. Our method provides a very high level of thoroughness in de-appearance step (both the identity and all soft-biometric information is removed), as well as high fidelity among re-appearanced faces. Also, we harness the high selectivity of the learned inverse mapping for different LBP encoding to make the re-appearance step secure, where high-confidence re-appearance only comes if the correct re-appearance key is applied. All of these are achieved by the forward mapping from image to LBP glyph and the learned inverse mapping through coupled dictionary learning.

To the best of our knowledge, this is the very first work that attempts to invert LBP, as well as its applications on face de-appearance and re-appearance with an emphasis on the thoroughness of the de-appearance and the security of the re-appearance. Figure 1 shows the flowchart of our proposed method where $\mathcal{F}$ is the forward mapping from image to LBP glyph, and $\mathcal{F}^{-1}$ is the learned inversion mapping. As can be seen visually, the recovered face shows high fidelity as compared to the original face.

Previous Work: In the work by Weinzaepfel *et al*. [58] on reconstructing an image from its SIFT descriptors, the authors first obtain the SIFT representation of patches from a large-scale image database, and when the query appearance descriptor comes in, nearest neighbor search and image-descriptor correspondence can provide the recovered image patch. The procedure is completed after seamless stitching and smooth interpolation for empty zones.

E. d'Angelo *et al*. make an attempt on reconstructing two local binary descriptors: BRIEF and FREAK in [8], which is an extension of their earlier work [7]. In their approach, two algorithms are proposed. The first algorithm can work on real-valued difference descriptors by tackling the reconstruction problem as a regularized deconvolution problem. The second algorithm harnesses some recent findings on 1-bit compressive sensing to reconstruct image parts from binarized difference descriptors. The authors have shown that the knowledge of the particular measurement layout of a local binary descriptor is sufficient to infer the original image patch without any external information or databases. It is also worth mentioning that these work [7, 8] apply only to BRIEF and FREAK, but not LBP descriptor. The au-

thors have provided an explanation in Section 2.2 of [8]. This is due to the fact that LBP, as conventionally used [45], requires a histogram step after the binarization, and such a step undermines the spatial awareness of the LBP glyph, which makes any algorithm very hard, if not at all impossible, to recover the original image. For this same reason, the LBP descriptor discussed in this paper does not have the final histogram step. We are interested in inverting the LBP glyph (or the LBP face) as shown in Figure 1.

2 Proposed Method and its Applications

Why is inverting a descriptor hard? It is not an easy task because descriptors usually compacts visual information to achieve invariance or robustness. Invariance and robustness means that the forward mapping from an image patch to descriptor is many-to-one, which makes the inversion hard and ill-posed. Additional constraints and / or appropriate priors are required to arrive at a solution for such a problem. In this work, we invoke a coupled dictionary learning paradigm with an ℓ_0-norm constraint to learn the non-linear mapping between pixel and LBP representations, and vice versa.

Thorough Face De-Appearance:

The intuition behind using LBP for de-appearance is straightforward because LBP is a local difference operator and most of the COTS FRS cannot deal with LBP faces / glyphs. They either can not locate the LBP faces from the scene, or the match scores are horribly low. Therefore, we resort to this simple yet very widely used LBP descriptor for face de-appearance purposes.

The traditional LBP operator was first introduced by Ojala *et al.* [45, 49]. All neighbors that have values higher than the value of the center pixel (pivot point) are given value 1 and 0 otherwise. The binary numbers associated with the neighbors are then read sequentially to form an 8-bit (or 24-bit for 5×5 case) string. The binary number is then converted to a decimal number as the feature assigned to the center pixel. The LBP feature for the center point (x_c, y_c) can be represented as: $LBP(x_c, y_c) = \sum_{n=0}^{L-1} s(i_n - i_c)2^n$ where i_n denotes the intensity of the n^{th} surrounding pixel, i_c denotes the intensity of the center pixel, L is the length of the sequence, and $s = 1$ if $i_n \geq i_c$, otherwise, $s = 0$.

When converting the bit string to a decimal number, there are many ways to do so depending on the ordering of the neighboring pixels. For instance, the most significant bit (MSB) is assigned to the top left pixel in the 3×3 patch and the following bits are obtained sequentially in a counter-clockwise fashion. Apparently, there are $8! = 40,320$ ways to order the 8 surrounding pixels, each one corresponding to a unique key for secure re-appearance to be discussed.

Secure Face Re-Appearance:

The de-appearance step using LBP is easy to implement, however, the re-appearance step that maps the LBP glyphs (we call it LBP domain faces) back to the face images (we call it pixel domain faces) requires more effort. Here, we propose to learn the inverse mapping of LBP via a coupled dictionary learning scheme.

The LBP operation is a nonlinear mapping[1] $\mathcal{F} : \mathbb{R}^d \longmapsto \mathbb{R}^d$ because of the thresholding of the neighboring pixels when compared to the center pixel within each local patch. This forward mapping is many-to-one because different images, as long as the local partial ordering of the pixels stays the same, will lead to the same LBP representation. Thus, using the

[1] d is the vectorized image dimension of the original image crop, which is the same as its LBP glyph.

inverse mapping $\mathcal{F}^{-1}$ to obtain one single re-appearanced image requires more constraints or prior information since this is an ill-posed problem.

The ideas of our method are as follows. We can jointly learn an overcomplete dictionary for the pixel domain faces and another for the LBP domain faces such that the sparse approximation coefficients for a query LBP face can be shared by the pixel domain faces which can lead to the reconstruction of the re-appearanced face. By doing this, a mapping between pixel domain image and LBP domain image can be established without explicitly knowing the mapping function. The enforcement of the sparsity level as well as the sharing of sparse coefficient between the two domains are the added constraints that can uniquely determine the inverse mapping $\mathcal{F}^{-1}$. The first objective is to learn a dictionary for the pixel domain faces: $\underset{\mathbf{D},\mathbf{X}}{\text{minimize}} \|\mathbf{Y}-\mathbf{D}\mathbf{X}\|_F^2$ subject to $\forall i, \|\mathbf{x}_i\|_0 < K$. Similarly, the second objective is to learn a dictionary for the LBP domain faces: $\underset{\mathbf{D}_{\text{LBP}},\mathbf{X}}{\text{minimize}} \|\mathbf{Y}_{\text{LBP}}-\mathbf{D}_{\text{LBP}}\mathbf{X}\|_F^2$ subject to $\forall i, \|\mathbf{x}_i\|_0 < K$. Combining the two objectives and solving them jointly allows us to enforce a common K-sparse representation. Our primary problem is therefore:

$$\underset{\mathbf{D},\mathbf{D}_{\text{LBP}},\mathbf{X}}{\arg\min} \|\mathbf{Y}-\mathbf{D}\mathbf{X}\|_F^2 + \|\mathbf{Y}_{\text{LBP}}-\mathbf{D}_{\text{LBP}}\mathbf{X}\|_F^2 \text{ subject to } \forall i, \|\mathbf{x}_i\|_0 < K \tag{1}$$

Here $\mathbf{Y} \in \mathbb{R}^{d\times N}$ and $\mathbf{Y}_{\text{LBP}} \in \mathbb{R}^{d\times N}$ are the pixel domain and the LBP domain training images respectively in matrix form where each column is a vectorized image. Here, d is the dimensionality of the data and N is the number of training images. $\mathbf{D} \in \mathbb{R}^{d\times M}$ and $\mathbf{D}_{\text{LBP}} \in \mathbb{R}^{d\times M}$ are the two overcomplete dictionaries for the pixel domain and the LBP domain faces, where $M \gg d$ is the number of dictionary atoms. $\mathbf{X} \in \mathbb{R}^{M\times N}$ is the sparse coefficient matrix shared between the two domains.

Obtaining a consistent sparse encoding between the two domains allows for a more meaningful reconstruction. Given a novel de-appearanced image $\mathbf{y}_{\text{LBP}}$ in the LBP domain, we first obtain the sparse representation $\mathbf{x}$ in $\mathbf{D}_{\text{LBP}}$. We then obtain the reconstruction using $\mathbf{D}\mathbf{x}$. By forcing consistent sparse representations $\mathbf{x}$ during training, we optimize for a low reconstruction error for both domains jointly and simultaneously. A simple rearrangement can lead to solving the formulation using the standard K-SVD dictionary learning approach as previously observed [14]:

$$\underset{\mathbf{D},\mathbf{D}_{\text{LBP}},\mathbf{X}}{\arg\min} \left\| \begin{pmatrix} \mathbf{Y} \\ \mathbf{Y}_{\text{LBP}} \end{pmatrix} - \begin{pmatrix} \mathbf{D} \\ \mathbf{D}_{\text{LBP}} \end{pmatrix} \mathbf{X} \right\|_F^2 \text{ subject to } \forall i, \|\mathbf{x}_i\|_0 \leq K \tag{2}$$

which translates to the standard K-SVD problem where we $\underset{\mathbf{D}',\mathbf{X}'}{\text{minimize}} \|\mathbf{Y}'-\mathbf{D}'\mathbf{X}\|_2$ under $\|\mathbf{x}_i\|_0 \leq K$, with $\mathbf{Y}' = (\mathbf{Y}^\top, \mathbf{Y}_{\text{LBP}}^\top)^\top$ and $\mathbf{D}' = (\mathbf{D}^\top, \mathbf{D}_{\text{LBP}}^\top)^\top$. This method is open set, enabling reconstruction of any face that is not present in the training set. To learn the optimal reconstruction sparsity level for the task, we conduct a pilot experiment in which we measure the average peak signal-to-noise ratio (PSNR) [21, 22, 24, 25, 29, 31, 38] between the re-appearanced face and the original face (prior to de-appearance) while increasing sparsity. The optimal choice of sparsity for reconstruction is $K_r = 8$ via cross validation.

3 Experiments

We use the **Target Set** of the large-scale NIST's FRGC ver 2.0 database [48], containing 466 different subjects, with a total of 16,028 images.

Face Re-Appearance Fidelity Experiments:

Figure 2: Re-appearance results on the FRGC target set.

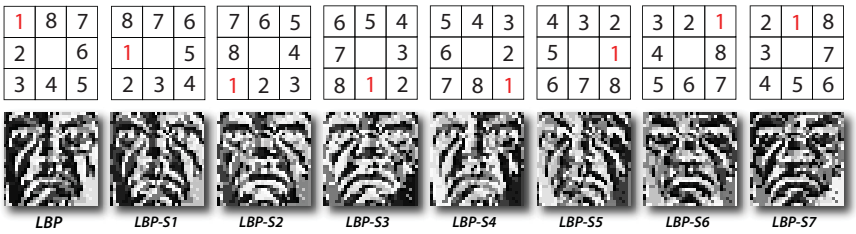


Figure 3: 8 different orderings and their corresponding LBP images for the same input face. The MSB is shown in red.

Following the procedure discussed in Section 2, we first de-appearance all the 16,028 face crops of size 32×32 from the FRGC target set using LBP. The two dictionaries were pre-trained using 200,000 images from a separate mugshot-type database. Not a single FRGC face image is observed during the training of the dictionaries. For each de-appearanced face, we use OMP [47] to get the sparse approximation coefficient from the LBP domain dictionary, and then use the same coefficient to reconstruct the image in the pixel domain, thus re-appearance. Figure 2 shows the re-appearance results for 20 random subjects from the FRGC target set. The average PSNR for the entire target set (16,028 face images) is 16.0015 dB. As can be seen, the proposed dictionary-based re-appearance method yields high fidelity in reconstruction.

We have previously mentioned that different orderings in formulating the LBP descriptor can play a role in secure face re-appearance. Figure 3 shows 8 different sequential counter-clockwise orderings and their corresponding LBP images for the same input face.

We call these 7 LBP variants LBP-S1 to LBP-S7 in addition to the original LBP discussed in Section 3. The 8 LBP glyphs look entirely different, which gives rise to the secure re-appearance capability. There are $8! = 40,320$ possible orderings for 3×3-patch LBP encoding, and each ordering is essentially an encryption key. The basic idea is that, if the joint dictionary is trained with one particular LBP variant, such dictionary pair can only be used for re-appearancing, with high fidelity, the de-appearanced face using the *same* LBP variant. Figure 4 shows the idea of secure re-appearance. In this case, the dictionary pair is trained on standard LBP, and when de-appearanced query faces using various LBP variants come, only the queries using the *same* LBP variant can yield high PSNR, and others would yield much lower PSNR. The average PSNR values are shown along with various re-appearance results in Figure 4 as well as tabulated in Table 1. One can imagine that for encoding LBP using

Type of Ordering	***LBP***	***LBP-S1***	***LBP-S2***	***LBP-S3***
Average PSNR	**16.0015**	14.2344	10.7258	8.9185
Type of Ordering	***LBP-S4***	***LBP-S5***	***LBP-S6***	***LBP-S7***
Average PSNR	8.2662	8.4797	11.9588	14.2227

Table 1: Average PSNR over 16,028 FRGC Target Set face images.

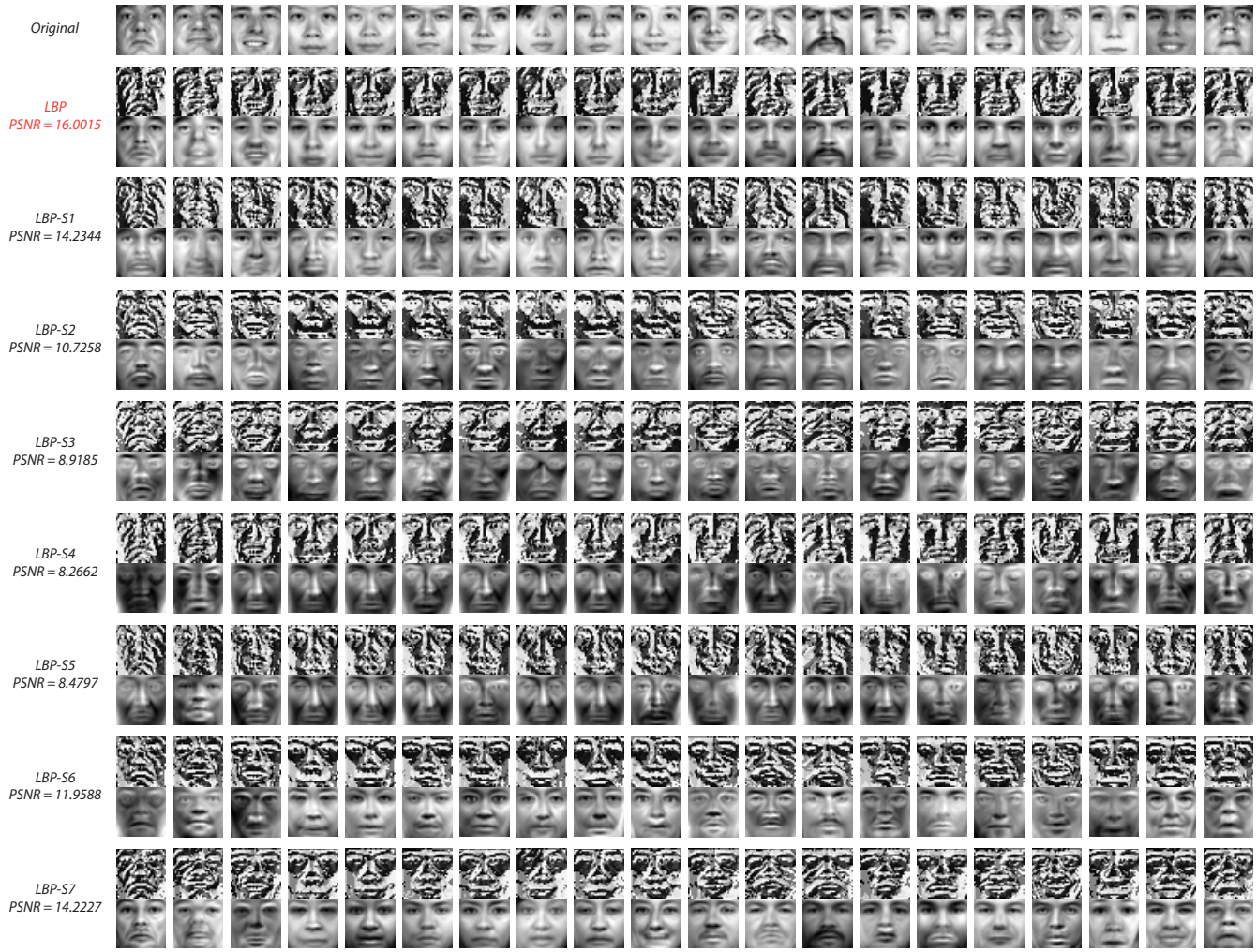


Figure 4: Re-appearance results of various de-appearanced images using different LBP variants. High PSNR only occurs when the dictionary and query use the same LBP variant.

larger patch size, *e.g.* 5×5, there are $24! = 6.2045 \times 10^{23}$ possible orderings, or unique encryption keys. To make it even more secure, we can make the key diverse according to people, instead of a fixed key for everyone in the database. Also, further security measures are possible by making connections to cancelable biometrics such as [52].

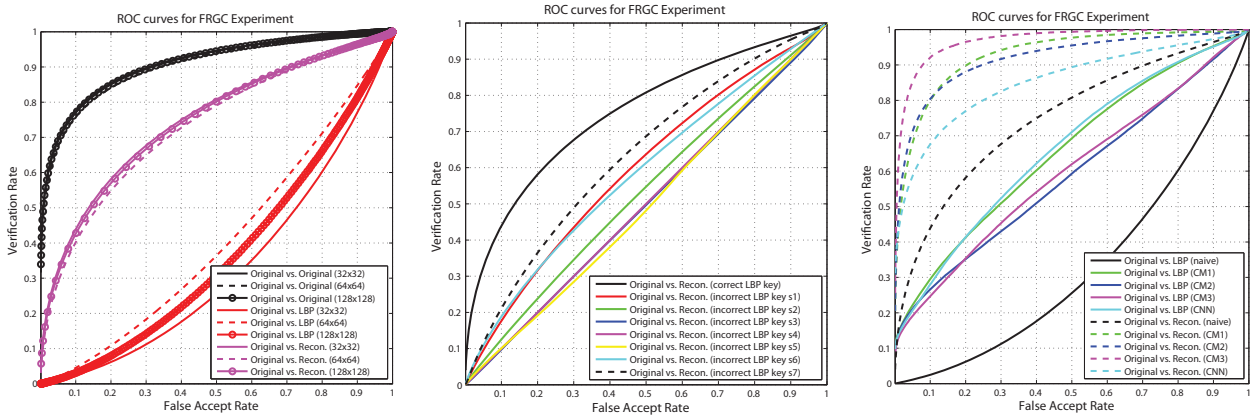
Not surprisingly, we can observe that similar orderings lead to similar re-appearance performance. The converted decimal number from the bit string is dominated by the most significant bit (MSB), and therefore, when rotating the bit string (from LBP-S1 to LBP-S7), the difference (in decimal number) between two bit strings is relatively small if the location of the MSB is close to each other. That is exactly why, LBP-S1 and LBP-S7 have relatively higher PSNR compared to the rest of the LBP variants because their MSB are next to that of the original LBP.

Here, we choose not do the histogram step in LBP because it kills the spatial resolution and make it impossible to recover. Actually, the way we conduct the procedure is a special case of histogramming with each bin at every pixel location, which is the finest possible way of getting the histogram. Also, for applications that the original faces are not accessible by the end user, the proposed method should be favored. It also makes sure that only the ones with the right keys can recover the faces. As can be seen in Figure 1, 2, and 4, the recovered faces greatly resemble the original faces, which is again confirmed by high PSNR. We resort to this approach due to its simplicity and universality.

Face Verification Experiments:

	VR at 0.1% FAR	*VR at 1% FAR*	*VR at 10% FAR*	*EER*	*AUC*
Orig. vs. Orig. 32×32	0.349	0.524	0.777	0.170	0.9109
Orig. vs. Orig. 64×64	0.324	0.567	0.757	0.175	0.9030
Orig. vs. Orig. 128×128	0.339	0.513	0.769	0.172	0.9085
Orig. vs. LBP 32×32	0.000	0.002	0.025	0.621	0.3297
Orig. vs. LBP 64×64	0.000	0.003	0.045	0.568	0.4076
Orig. vs. LBP 128×128	0.000	0.002	0.032	0.595	0.3678
Orig. vs. Recon. 32×32	0.046	0.174	0.440	0.320	0.7501
Orig. vs. Recon. 64×64	0.048	0.152	0.402	0.332	0.7323
Orig. vs. Recon. 128×128	0.057	0.161	0.427	0.325	0.7440

Table 2: Performance on FRGC Experiment 1 Evaluation. Experiments using a naive classifier with three image resolutions are evaluated.



(a) (b) (c)

Figure 5: 5(a): ROC curves for matching FRGC target set to themselves, their LBP glyphs, and the reconstructed counterparts using correct encryption key. 5(b): ROC curves for matching FRGC target set to the reconstructed counterparts, using various encryption keys. 5(c): ROC curves for matching FRGC target set to their LBP glyphs, and to the reconstructed counterparts, using various face matchers.

A. Face verification using a naive classifier: We have conducted face verification experiments to see how well can the re-appearanced images be matched to the original ones. The performance is analyzed using verification rate (VR) at 0.1%, 1% and 10% false accept rate (FAR), equal error rate (EER), the receiver operating characteristic (ROC) curves, and the area under the ROC curve (AUC) consolidated in Table 2 and Figure 5(a). We strictly follow NIST's **FRGC Experiment 1** protocol which involves 1-to-1 matching of $16,028$ target images to themselves (~ 256 million pair-wise face match comparisons). The similarity matrix is of size $16,028 \times 16,028$. We use the normalized cosine distance (NCD) [4, 18, 19, 23, 26, 28, 30, 32, 34, 46, 54, 55, 60] as the metric.

Our baseline is original image vs. original image which should yield the highest possible performance. To show the effectiveness of using LBP for de-appearance, we match original images to LBP images. To showcase the effectiveness of re-appearance, we match original images to the reconstructed images. Three image resolutions[2] are used: 32×32, 64×64, and 128×128.

As can be seen, de-appearance yields abysmal verification performance as expected, when matched against original images, which shows the thoroughness of the proposed approach. Also, we have demonstrated the effectiveness of the proposed dictionary-based re-

[2]For 64×64 crop, 500k mugshots are used for dictionary learning, and for 128×128 crop, 1 million are used.

	VR at 0.1% FAR	*VR at 1% FAR*	*VR at 10% FAR*	*EER*	*Rank-1 ID Rate*	*AUC*
Orig. vs. Recon. (correct LBP key)	0.046	0.174	0.430	0.320	0.3456	0.7501
Orig. vs. Recon. (incorrect LBP key s1)	0.002	0.029	0.163	0.426	0.0153	0.5945
Orig. vs. Recon. (incorrect LBP key s2)	0.002	0.023	0.128	0.478	0.0088	0.5325
Orig. vs. Recon. (incorrect LBP key s3)	0.001	0.009	0.102	0.499	0.0018	0.4961
Orig. vs. Recon. (incorrect LBP key s4)	0.001	0.012	0.107	0.513	0.0011	0.4900
Orig. vs. Recon. (incorrect LBP key s5)	0.001	0.012	0.101	0.509	0.0021	0.4931
Orig. vs. Recon. (incorrect LBP key s6)	0.002	0.047	0.178	0.441	0.0076	0.5843
Orig. vs. Recon. (incorrect LBP key s7)	0.002	0.030	0.218	0.402	0.0107	0.6324

Table 3: Performance on FRGC Experiment 1 Evaluation. Matching original images with reconstructions using various LBP keys.

appearance method by showing good performance when matching the re-appearanced faces to the original ones, while previous methods do not have the capability at all. In addition, we see that image resolution does not seem to affect the verification performance too much.

For the ROC curves, as we all know that a random chance would give a diagonal line, and any method that leads to a curve below the line can be improved by simply flipping the decision rule. However, under the circumstance that the system is *agnostic* about whether the input image is raw pixel image or LBP glyph, it will produce consistent decision regardless and it is possible to have an ROC curve below the diagonal line. The access system does not have the prior information whether the incoming query image is a LBP glyph or a pixel domain image. It also does not have the ensemble knowledge and characteristics of the entire testing set, like what the ROC curve tries to capture. All it sees is a single matching score, and there is no reason for it to 'flip' the decision strategy based upon just a single query image at run-time.

In addition, we carry out experiments that match original images with reconstructed images using different encryption keys, following the ones shown in Figure 3. The results are reported in Table 3 and the ROC curves are shown in Figure 5(b). We can observe that, only reconstructing using the correct encryption key can yield high face verification performance, which is in line with the reconstruction fidelity observed in Figure 4 and Table 1.

B. Face verification using the state-of-the-art commercial and academic face matchers: In the following set of experiments, we utilize three state-of-the-art commercial matchers (CM) and one academia matcher for face verification. The three commercial matchers are paraphrased as CM1, CM2, and CM3. The academic face matcher we use is OpenFace [2], which is an open-source Python and Torch implementation of Google's FaceNet convolutional neural network (CNN) architecture [53]. With a slight modification on the network structure, we train our model by combining the 3 largest publicly-available face recognition datasets: FaceScrub [44], CASIA-WebFace [59], and MegaFace [42]. Once the model is trained, we pass the testing image through the network and extract a 4096-dimensional feature vector from the top-most fully connected layer. It is also worth noting that CM2 cannot work with tightly cropped faces as those shown in Figure 2 and 4. Therefore, we plaster the LBP (de-appearanced) face and the reconstructed (re-appearanced) face back to the original loosely cropped image, as depicted in Figure 6. As can be seen, CM2 may or may not harness the strong background information outside the square tight crop, among the original image, the plastered-back LBP glyph, and the plastered-back recovered face. This is something we don't have control over and we report the results from the CM2 as is.

We first match the original faces from the FRGC to the LBP counterparts, and then we match the original faces to the recovered one. The ROC curves are shown in Figure 5(c)

	Naive	*CM1*	*CM2*	*CM3*	*CNN*
Original vs. LBP	0.3297	0.6567	0.5893	0.5984	0.6624
Original vs. Recon.	0.7501	0.9289	0.9181	0.9691	0.8544

Table 4: Performance (AUC) on FRGC. Matching original images with LBP glyphs and reconstructions using a naive classifier, 3 commercial and 1 academic face matchers.

Figure 6: For CM2, loosely cropped faces are provided. We plaster back the square LBP glyph and the recovered face to the original face image.

and the AUC for each method is consolidated in Table 4. From the results we can have the following observations: (1) even the state-of-the-art commercial and academic face matchers cannot successfully match the original faces to the LBP faces with high accuracy, which shows the efficacy of our thorough face de-appearance step; (2) when matching the original faces to the recovered ones, using sophisticated matchers does dramatically improve on the naive classifier and we have reached AUC over 90% using CM1, CM2, and CM3, which demonstrate the high fidelity of our face re-appearance step.

4 Conclusions

We have proposed a novel method for inverting the LBP descriptor. The success of the inversion gives rise to two face-related applications: face de-appearance and re-appearance. The de-appearance based on image-LBP forward mapping is thorough in the sense that not only the identity information, but also the soft-biometric information of the subject is removed. The re-appearance yields face reconstruction with high fidelity and also enables secure application with unique encryption key. The re-appearance involves leaning the inverse mapping of the LBP descriptors through an ℓ_0-constrained coupled dictionary learning paradigm that jointly learns two overcomplete dictionaries in both the pixel and the LBP domain such that inverse mapping $\mathcal{F}^{-1}$ from the LBP to the pixel domain is made possible. We have showcased the effectiveness of our proposed approach on the FRGC ver 2.0 database which involves large-scale fidelity test and face verification experiments using the best commercial and academic face matchers. Future work may include developing an approach that can blindly invert the LBP descriptor without knowing the encoding schemes at all.

References

[1] Alexandre Alahi, Raphael Ortiz, and Pierre Vandergheynst. Freak: Fast retina keypoint. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 510–517, 2012.

[2] Brandon Amos, Jan Harkes, Padmanabhan Pillai, Khalid Elgazzar, and Mahadev Satyanarayanan. Openface 0.1.1: Face recognition with google's facenet deep neural network, October 2015.

[3] Apurva Bedagkar-Gala and Shishir K. Shah. A survey of approaches and trends in person re-identification. *Image and Vision Computing*, 32(4):270 – 286, 2014. ISSN 0262-8856. doi: http://dx.doi.org/10.1016/j.imavis.2014.02.001. URL http://www.sciencedirect.com/science/article/pii/S0262885614000262.

[4] P. Buchana, I. Cazan, M. Diaz-Granados, F. Juefei-Xu, and M.Savvides. Simultaneous Forgery Identification and Localization in Paintings Using Advanced Correlation Filters. In *IEEE International Conference on Image Processing (ICIP)*, pages 1–5, Sept 2016.

[5] Michael Calonder, Vincent Lepetit, Christoph Strecha, and Pascal Fua. Brief: Binary robust independent elementary features. In *Computer Vision–ECCV 2010*, pages 778–792. Springer, 2010.

[6] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886–893, 2005.

[7] Emmanuel d'Angelo, Alexandre Alahi, and Pierre Vandergheynst. Beyond bits: Reconstructing images from local binary descriptors. In *Pattern Recognition (ICPR), 2012 21st International Conference on*, pages 935–938, 2012.

[8] Emmanuel d'Angelo, Laurent Jacques, Alexandre Alahi, and Pierre Vandergheynst. From bits to images: Inversion of local binary descriptors. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 36(5):874–887, 2014.

[9] Liang Du, Meng Yi, Erik Blasch, and Haibin Ling. Garp-face: Balancing privacy protection and utility preservation in face de-identification. In *Biometrics (IJCB), 2014 IEEE International Joint Conference on*, pages 1–8, Sept 2014.

[10] R. Gross and L. Sweeney. Towards real-world face de-identification. In *Biometrics: Theory, Applications, and Systems, 2007. BTAS 2007. First IEEE International Conference on*, pages 1–8, Sept 2007.

[11] R. Gross, L. Sweeney, F. de la Torre, and S. Baker. Model-based face de-identification. In *Computer Vision and Pattern Recognition Workshop, 2006. CVPRW '06. Conference on*, pages 161–161, June 2006.

[12] R. Gross, L. Sweeney, F. De la Torre, and S. Baker. Semi-supervised learning of multi-factor models for face de-identification. In *IEEE CVPR*, pages 1–8, June 2008.

[13] Ralph Gross, Latanya Sweeney, Jeffery F. Cohn, Fernando De la Torre, and Simon Baker. *Protecting Privacy in Video Surveillance*, chapter Face De-identification, pages 129–146. Springer Publishing Company, Incorporated, 2009. ISBN 978-1-84882-300-6.

[14] Zhuolin Jiang, Zhe Lin, and L.S. Davis. Label consistent K-SVD: Learning a discriminative dictionary for recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35 (11):2651–2664, Nov 2013.

[15] A. Jourabloo, Xi Yin, and Xiaoming Liu. Attribute preserved face de-identification. In *Biometrics (ICB), 2015 International Conference on*, pages 278–285, May 2015.

[16] F. Juefei-Xu and M. Savvides. Can Your Eyebrows Tell Me Who You Are? In *Signal Processing and Communication Systems (ICSPCS), 2011 5th International Conference on*, pages 1–8, Dec 2011.

[17] F. Juefei-Xu and M. Savvides. Unconstrained Periocular Biometric Acquisition and Recognition Using COTS PTZ Camera for Uncooperative and Non-cooperative Subjects. In *Applications of Computer Vision (WACV), 2012 IEEE Workshop on*, pages 201–208, Jan 2012.

[18] F. Juefei-Xu and M. Savvides. An Image Statistics Approach towards Efficient and Robust Refinement for Landmarks on Facial Boundary. In *Biometrics: Theory, Applications and Systems (BTAS), 2013 IEEE Sixth International Conference on*, pages 1–8, Sept 2013.

[19] F. Juefei-Xu and M. Savvides. An Augmented Linear Discriminant Analysis Approach for Identifying Identical Twins with the Aid of Facial Asymmetry Features. In *Computer Vision and Pattern Recognition (CVPR) Workshops, 2013 IEEE Conference on*, pages 56–63, June 2013.

[20] F. Juefei-Xu and M. Savvides. Subspace Based Discrete Transform Encoded Local Binary Patterns Representations for Robust Periocular Matching on NIST's Face Recognition Grand Challenge. *IEEE Trans. on Image Processing*, 23(8):3490–3505, aug 2014.

[21] F. Juefei-Xu and M. Savvides. Pokerface: Partial Order Keeping and Energy Repressing Method for Extreme Face Illumination Normalization. In *Biometrics: Theory, Applications and Systems (BTAS), 2015 IEEE Seventh International Conference on*, pages 1–8, Sept 2015.

[22] F. Juefei-Xu and M. Savvides. Encoding and Decoding Local Binary Patterns for Harsh Face Illumination Normalization. In *IEEE International Conference on Image Processing (ICIP)*, pages 3220–3224, Sept 2015.

[23] F. Juefei-Xu and M. Savvides. Pareto-optimal Discriminant Analysis. In *IEEE International Conference on Image Processing (ICIP)*, pages 611–615, Sept 2015.

[24] F. Juefei-Xu and M. Savvides. Single Face Image Super-Resolution via Solo Dictionary Learning. In *IEEE International Conference on Image Processing (ICIP)*, pages 2239–2243, Sept 2015.

[25] F. Juefei-Xu and M. Savvides. Fastfood Dictionary Learning for Periocular-Based Full Face Hallucination. In *Biometrics: Theory, Applications and Systems (BTAS), 2016 IEEE Seventh International Conference on*, pages 1–8, Sept 2016.

[26] F. Juefei-Xu and M. Savvides. Multi-class Fukunaga Koontz Discriminant Analysis for Enhanced Face Recognition. *Pattern Recognition*, 52:186–205, apr 2016.

[27] F. Juefei-Xu, Khoa Luu, M. Savvides, T.D. Bui, and C.Y. Suen. Investigating Age Invariant Face Recognition Based on Periocular Biometrics. In *Biometrics (IJCB), 2011 International Joint Conference on*, pages 1–7, Oct 2011.

[28] F. Juefei-Xu, C. Bhagavatula, A. Jaech, U. Prasad, and M. Savvides. Gait-ID on the Move: Pace Independent Human Identification Using Cell Phone Accelerometer Dynamics. In *Biometrics: Theory, Applications and Systems (BTAS), 2012 IEEE Fifth International Conference on*, pages 8–15, Sept 2012.

[29] F. Juefei-Xu, Dipan K. Pal, and M. Savvides. Hallucinating the Full Face from the Periocular Region via Dimensionally Weighted K-SVD. In *Computer Vision and Pattern Recognition (CVPR) Workshops, 2014 IEEE Conference on*, pages 1–8, June 2014.

[30] F. Juefei-Xu, K. Luu, and M. Savvides. Spartans: Single-sample Periocular-based Alignment-robust Recognition Technique Applied to Non-frontal Scenarios. *IEEE Trans. on Image Processing*, 24(12):4780–4795, Dec 2015.

[31] F. Juefei-Xu, D. K. Pal, and M. Savvides. NIR-VIS Heterogeneous Face Recognition via Cross-Spectral Joint Dictionary Learning and Reconstruction. In *Computer Vision and Pattern Recognition (CVPR) Workshops, 2015 IEEE Conference on*, pages 141–150, June 2015.

[32] F. Juefei-Xu, D. K. Pal, K. Singh, and M. Savvides. A Preliminary Investigation on the Sensitivity of COTS Face Recognition Systems to Forensic Analyst-style Face Processing for Occlusions. In *Computer Vision and Pattern Recognition (CVPR) Workshops, 2015 IEEE Conference on*, pages 25–33, June 2015.

[33] F. Juefei-Xu, E. Verma, P. Goel, A. Cherodian, and M. Savvides. DeepGender: Occlusion and Low Resolution Robust Facial Gender Classification via Progressively Trained Convolutional Neural Network with Attention. In *Computer Vision and Pattern Recognition (CVPR) Workshops, 2016 IEEE Conference on*, June 2016.

[34] Felix Juefei-Xu and Marios Savvides. Facial Ethnic Appearance Synthesis. In *Computer Vision - ECCV 2014 Workshops*, volume 8926 of *Lecture Notes in Computer Science*, pages 825–840. Springer International Publishing, 2015.

[35] Felix Juefei-Xu and Marios Savvides. Weight-Optimal Local Binary Patterns. In *Computer Vision - ECCV 2014 Workshops*, volume 8926 of *Lecture Notes in Computer Science*, pages 148–159. Springer International Publishing, 2015.

[36] Felix Juefei-Xu, M. Cha, J. L. Heyman, S. Venugopalan, R. Abiantun, and M. Savvides. Robust Local Binary Pattern Feature Sets for Periocular Biometric Identification. In *Biometrics: Theory Applications and Systems (BTAS), 4th IEEE Int'l Conf. on*, pages 1–8, sep 2010.

[37] Felix Juefei-Xu, Miriam Cha, Marios Savvides, Saad Bedros, and Jana Trojanova. Robust Periocular Biometric Recognition Using Multi-level Fusion of Various Local Feature Extraction Techniques. In *IEEE 17th International Conference on Digital Signal Processing (DSP)*, pages 1–7, 2011.

[38] Felix Juefei-Xu, Dipan K. Pal, and Marios Savvides. Methods and Software for Hallucinating Facial Features by Prioritizing Reconstruction Errors, 2014. U.S. Provisional Patent Application Serial No. 61/998,043, June 17, 2014.

[39] David G Lowe. Object recognition from local scale-invariant features. In *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, volume 2, pages 1150–1157, 1999.

[40] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[41] A. Mahendran and A. Vedaldi. Understanding deep image representations by inverting them. In *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*, pages 5188–5196, June 2015.

[42] Daniel Miller, Evan Brossard, Steven M Seitz, and Ira Kemelmacher-Shlizerman. Megaface: A million faces for recognition at scale. *arXiv preprint arXiv:1505.02108*, 2015.

[43] E.M. Newton, L. Sweeney, and B. Malin. Preserving privacy by de-identifying face images. *Knowledge and Data Engineering, IEEE Transactions on*, 17(2):232–243, Feb 2005.

[44] Hong-Wei Ng and S. Winkler. A data-driven approach to cleaning large face datasets. In *Image Processing (ICIP), 2014 IEEE International Conference on*, pages 343–347, Oct 2014.

[45] Timo Ojala, Matti Pietikainen, and Topi Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(7):971–987, 2002.

[46] D. K. Pal, F. Juefei-Xu, and M. Savvides. Discriminative Invariant Kernel Features: A Bells-and-Whistles-Free Approach to Unsupervised Face Recognition and Pose Estimation. In *Computer Vision and Pattern Recognition (CVPR), 2016 IEEE Conference on*, June 2016.

[47] Y. Pati, R. Rezaiifar, and P. Krishnaprasad. Orthogonal matching pursuit: Recursive function approximation with application to wavelet decomposition. In *Asilomar Conf. on Signals, Systems and Comput.*, 1993.

[48] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, Jin Chang, K. Hoffman, J. Marques, Jaesik Min, and W. Worek. Overview of the face recognition grand challenge. In *CVPR*, volume 1, pages 947–954, jun 2005.

[49] M. Pietikainen, G. Zhao, A. Hadid, and T. Ahonen. *Computer Vision Using Local Binary Patterns*. Springer, 2011.

[50] B. Samarzija and S. Ribaric. An approach to the de-identification of faces in different poses. In *Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2014 37th International Convention on*, pages 1246–1251, May 2014.

[51] Marios Savvides and Felix Juefei-Xu. Image Matching Using Subspace-Based Discrete Transform Encoded Local Binary Patterns, September 2013. US Patent US 2014/0212044 A1.

[52] Marios Savvides, BVKV Kumar, and Pradeep K Khosla. Cancelable biometric filters for face recognition. In *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, volume 3, pages 922–925. IEEE, 2004.

[53] F. Schroff, D. Kalenichenko, and J. Philbin. Facenet: A unified embedding for face recognition and clustering. In *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*, pages 815–823, June 2015.

[54] K. Seshadri, F. Juefei-Xu, D. K. Pal, and M. Savvides. Driver Cell Phone Usage Detection on Strategic Highway Research Program (SHRP2) Face View Videos. In *Computer Vision and Pattern Recognition (CVPR) Workshops, 2015 IEEE Conference on*, pages 35–43, June 2015.

[55] S. Venugopalan, F. Juefei-Xu, B. Cowley, and M. Savvides. Electromyograph and Keystroke Dynamics for Spoof-Resistant Biometric Authentication. In *Computer Vision and Pattern Recognition (CVPR) Workshops, 2015 IEEE Conference on*, pages 109–118, June 2015.

[56] Carl Vondrick. Visualizing object detection features. Master's thesis, Massachusetts Institute of Technology, 2013.

[57] Carl Vondrick, Aditya Khosla, Tomasz Malisiewicz, and Antonio Torralba. Hoggles: Visualizing object detection features. In *Computer Vision (ICCV), 2013 IEEE International Conference on*, pages 1–8, 2013.

[58] Philippe Weinzaepfel, Hervé Jégou, and Patrick Pérez. Reconstructing an image from its local descriptors. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 337–344, 2011.

[59] Dong Yi, Zhen Lei, Shengcai Liao, and Stan Z. Li. Learning face representation from scratch. *CoRR*, abs/1411.7923, 2014. URL http://arxiv.org/abs/1411.7923.

[60] N. Zehngut, F. Juefei-Xu, R. Bardia, D. K. Pal, C. Bhagavatula, and M. Savvides. Investigating the Feasibility of Image-Based Nose Biometrics. In *IEEE International Conference on Image Processing (ICIP)*, pages 522–526, Sept 2015.