De-Identifying Facial Images Using Singular Value Decomposition

P. Chriskos, O. Zoidi, A. Tefas and I. Pitas
Department of Informatics
Aristotle University of Thessloniki
Thessaloniki 54124 Greece
chriskos@csd.auth.gr, {ozoidi,tefas,pitas}@aiia.csd.auth.gr

***Abstract*—In this paper, a method is proposed that manipulates images in a manner that hinders face recognition by automatic recognition algorithms. The purpose of this method, is to partly degrade image quality, so that humans can identify the person or persons in a scene, while common classification algorithms fail to do so. The approach used to achieve this involves the use of singular value decomposition (SVD). From experiments it can be concluded that, the method reduces the percentage of correct classification rate by over 90% . In addition, the final image is not degraded beyond recognition by humans.**

I. Introduction

With the increasing amount of visual media that is shared, viewed and stored on-line, it is incontestable that privacy is a main concern for all users. The free access that is granted to all this visual information may carry many dangers concerning the privacy of the creators and of the subjects in these media. Face recognition algorithms are able to identify faces in videos and images without much effort, thus violating the privacy of the subjects. Malicious users can use video sharing sites and social media to collect information about specific individuals and groups fast and effortlessly. Moreover, the wide use of video surveillance in public places, in conjunction with face identification software, is a major threat of privacy, since, all persons can be identified regardless of suspicion level. Other examples of contributors to the problem include Google Street View and EverySpace among others, whose attempt to provide services which include visual data inevitably invade our everyday privacy, although not intentionally. As such, the necessity arises to develop methods that protect the subject's privacy, while maintaining a level of quality. This quality is not only limited to the visual quality of the final product, but the viewer must also be able to recognize the number of individuals in a scene, possibly even the individuals themselves and what actions are taking place in the image or video frame.

With this in mind, suppose a malicious user has trained a classifier in order to recognize images of targeted individuals or groups in a set of images available online. New images that are modified by a certain method, will not be recognized by the trained classifier, tackling the attempt of a malicious user searching new images of his targets and rendering further activities of the targets safe.

In order to achieve this, the common approach is to develop algorithms designed to hinder face recognition by both human viewers and face recognition methods. These algorithms aim to destroy the majority if not all visual data. Some of these methods achieve this by completely or partially blackening the face area [8]. Other methods in this category include the use of 3D morphable models [3] to cover the initial face in the image. Some methods have also been proposed that exploit the weaknesses of various face recognition algorithms, in order to manipulate images in a way that reduces the effectiveness of these algorithms [4]. Generally, most available methods for face de-identification fall in one of two groups, those that involve ad-hoc distortion [1] [8] [10] and the k-Same family of algorithms [6], implementing the k-anonymity protection model. The ad-hoc methods use simple methods, such as, smoothing the image using a Gaussian filter and subsampling an image leading to pixelation. It can be shown that these algorithms are easy to defeat and, also, degrade the utility of the images. The k-Same family of algorithms implements k-anonymity. From a set of initial images in which each subject is represented by only one image (known as person-specific), this method computes the de-identified image set so that all of the de-identified images indiscriminately relate to at least k elements of the initial image set. This method works either in the image space or in the Principal Component Analysis coefficient space. It can be shown that the best possible success rate of this method is $\frac{1}{k}$. Another method, proposed by Phillips [9], protects privacy by reducing the number of eigen vectors used in reconstructing images from basis vectors.

In this paper a novel method is proposed that aims to reduce the percentage of positive face identification of common recognition algorithms, while retaining enough visual information to characterize the end product as visually acceptable. The proposed method utilizes the singular value decomposition method (SVD), manipulating the values of the coefficients, in order to alter the initial image. The purpose is to enable human viewers to identify the individual pictured, while hindering common identification methods from achieving a high identification rate.

The rest of the paper is organized as follows. Section II, provides an overview of the singular value decomposition method. Section III, describes the proposed method. Experimental results are presented in Section IV. Finally, the conclusions are drawn in Section V.

II. Singular Value Decomposition

Singular value decomposition (SVD) [4] [7] [2] [5]is a matrix factorization method that approximates a matrix $\mathbf{A} \in \Re^{n \times p}$ with the product of three matrices $\mathbf{U} \in \Re^{n \times n}, \mathbf{S} \in \Re^{n \times p}$ and $\mathbf{V} \in \Re^{p \times p}$. The SVD theorem, states that any real matrix $\mathbf{A} \in \Re^{n \times p}$ can be decomposed uniquely as

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T \tag{1}$$

Matrices $\mathbf{U}$ and $\mathbf{V}$ must abide to the following conditions:

$$\mathbf{U}^T\mathbf{U} = \mathbf{I}_{n \times n} \tag{2}$$

and

$$\mathbf{V}^T\mathbf{V} = \mathbf{I}_{p \times p}, \tag{3}$$

where $\mathbf{I}_{n \times n}$ and $\mathbf{I}_{p \times p}$ are the identity matrices of size n and p, respectively, meaning that $\mathbf{U}$ and $\mathbf{V}$ are orthogonal. Matrix $\mathbf{S}$ is a diagonal matrix with the same dimensions as the input matrix $\mathbf{A}$, containing the singular values of $\mathbf{A}$.

In order to perform singular value decomposition it is necessary to find the eigenvalues and eigenvectors of the symmetric matrices $\mathbf{A}\mathbf{A}^T$ and $\mathbf{A}^T\mathbf{A}$. The eigenvectors of $\mathbf{A}^T\mathbf{A}$ consist the columns of matrix $\mathbf{V}$. The matrix $\mathbf{A}^T\mathbf{A}$ can be written as:

$$\mathbf{A}^T\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T\mathbf{V}\mathbf{S}\mathbf{U}^T = \mathbf{U}\mathbf{S}^2\mathbf{U}^T \tag{4}$$

The eigenvectors of $\mathbf{A}\mathbf{A}^T$ make up the columns of matrix $\mathbf{U}$. The matrix $\mathbf{A}\mathbf{A}^T$ can be written as:

$$\mathbf{A}\mathbf{A}^T = \mathbf{U}\mathbf{S}\mathbf{V}^T\mathbf{U}\mathbf{S}\mathbf{V}^T = \mathbf{V}\mathbf{S}^2\mathbf{V}^T \tag{5}$$

Finally, the singular values in $\mathbf{S}$ are the square roots of the matrix $\mathbf{A}\mathbf{A}^T$ or $\mathbf{A}^T\mathbf{A}$ eigenvalues. The singular values are arranged in descending order in the primary diagonal of matrix $\mathbf{S}$. These singular values are real numbers. More explicitly, if $\mathbf{A}$ is a matrix with real values, then the values in $\mathbf{U}$ and $\mathbf{V}$ are also real.

Having calculated the matrices of the singular value decomposition, the initial rectangular matrix can be derived, by using the formula (1) mentioned in the SVD theorem. Matrix $\mathbf{A}$ can also be alternatively computed as

$$\mathbf{A} = \sum_{i=1}^{rank} \sigma_i \mathbf{u}_i \mathbf{v}_i^T, \tag{6}$$

where $\mathbf{u}_i \in \Re^n$, $\mathbf{v}_i \Re^p$ denote the columns of matrices $\mathbf{U}$ and $\mathbf{V}$, respectively. $rank$ denotes the rank of matrix $\mathbf{A}$ and it is equal to the number of its positive singular values. As a result, the rank of matrix $\mathbf{A}$ is equal to the rank of matrix $\mathbf{S}$, since the rank of a diagonal matrix is equal to the number of its nonzero diagonal entries. Finally, σ_i denotes the i^{th} singular value of $\mathbf{A}$.

III. Person de-identification based on SVD

The workhorse of the proposed method is SVD applied on facial images. As described in Section II, SVD factorizes the input matrix (in our case a facial image) $\mathbf{A}$ as a product of three matrices: the singular values matrix $\mathbf{S}$ and the eigenvectors matrices $\mathbf{U}$ and $\mathbf{V}$. The proposed person de-identification method utilizes the SVD to manipulate facial images in order to reduce facial identification by software agents. This method alters the values in the matrices produced by the decomposition.

In order to reduce the correct identification rate, the following steps are followed. First, the coefficients (singular values) of matrix $\mathbf{S}$ with the largest values are reduced to zero. Next, the matrices $\mathbf{U}$ and $\mathbf{V}$ are blurred using an averaging filter. Finally, the same matrices are sharpened using a modified Sobel filter. The logic behind this course of action, is described below.

A. SVD Coefficient Zeroing (SVD-CZ)

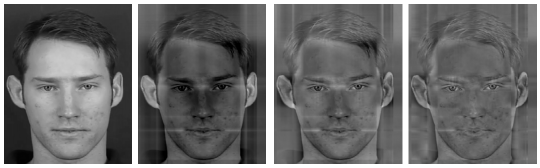
The most discriminative visual information in an image lies in the coefficients (singular values) with the largest values. Therefore, in the first step, the idea is to remove this information contained in the first coefficients, in the form of pixel luminosity. Since we are removing the first N coefficients, we are actually removing those coefficients that contain the majority of information that a face recognition algorithm would use to successfully identify a subject. This is achieved by setting the first N singular values in $\mathbf{S}$ to zero. Equivalently, we remove the first N primary coefficients used in recomposing the final image. This process produces a new $\mathbf{S}$ matrix referred to as $\mathbf{S}_{CZ}$. Dropping SVD coefficients other than the first was also attempted, but the error rates for the mentioned classifiers was low.

By setting the N largest singular values to zero, the final image tends to darken with respect to the input image. In order to preserve adequate visual data for easy face identification by human viewers, we increase the luminosity of all pixels in the end of the process, by adding a fixed value to the pixels of the output image. This darkening effect is due to the fact that the largest coefficients in matrix $\mathbf{S}$ are reduced to zero. These values are subsequently used in the calculation of the output image through matrix multiplication. Since matrix multiplication involves summing of coefficients some of which are set to zero instead of having their initial positive values, the result is smaller in numerical value. As a result, the output image is darker.

The effect of SVD coefficients zeroing can be viewed in Figure 1, where the darkening effect was reduced by adding luminosity 100 in each pixel of the final images.

B. SVD Coefficient Averaging (SVD-CA)

As we have previously discussed, the method goal is to allow human viewers to recognize with relative ease the subject in an image and, at the same time, fool automatic classifiers trying to identify specific individuals. This difficulty will arise from the fact that these classifiers where trained with



(a) (b) (c) (d)

Fig. 1. (a) Original Frame, (b) SVD-CZ with $N = 1$, (c) SVD-CZ with $N = 2$, (d) SVD-CZ with $N = 4$

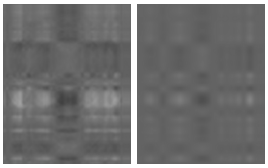


Fig. 2. Left: Result for SVD-CA with r = 4 Right: Result for SVD-CA with r=10 (composed with $\mathbf{U}_{averaged}$ and $\mathbf{V}_{averaged}$)

Fig. 3. Left: Result for SVD-CA with $r = 10$, Right: Result for SVD-CA with $r = 20$

clean versions of the images and ,subsequently, will falsely identify the manipulated images. To achieve this, the entries of the eigenvectors in matrices $\mathbf{U}$ and $\mathbf{V}$ are mixed by a blurring filter. The averaging filter employed is the $m \times m$ circular averaging filter, with $m = 2r + 1$, where r is the radius of the circular filter. By applying the averaging filter to matrices $\mathbf{U}$ and $\mathbf{V}$, we obtain the matrices $\mathbf{U}_{averaged}$ and $\mathbf{V}_{averaged}$. Recomposing the image solely from the averaged matrices, leads to poor visual quality, as portrayed in Figure 2. From Figure 2 we notice that the output images are degraded beyond recognition. In order to counterbalance this effect, only a percentage of the values from the new matrices is used. The final matrices $\mathbf{U}_{CA}$ and $\mathbf{V}_{CA}$ utilized to calculate the output image are given by the following equations:

$$\mathbf{U}_{CA} = \frac{\alpha * \mathbf{U}_{averaged} + \mathbf{U}}{1 + \alpha} \tag{7}$$

and

$$\mathbf{V}_{CA} = \frac{\alpha * \mathbf{V}_{averaged} + \mathbf{V}}{1 + \alpha}, \tag{8}$$

where the parameter α adjusts the equilibrium between visual quality and privacy protection. Similarly to the previous method, this step also introduces a darkening effect in the resulting image. This effect is adjusted as in the first step. The visual result of equations (7), (8) is displayed in Figure 3, individually for SVD-CA with added luminosity 100.

C. SVD Modified Sobel Filtering (SVD-MSF)

The final step utilizes a modified Sobel filter in order to manipulate matrices $\mathbf{U}_{CA}$ and $\mathbf{V}_{CA}$. Sobel filtering is generally used for edge detection in images. Edge detection is used to remove part of the previous blurring, while mixing the coefficient values even further.This modified filter, contains values different from the classic Sobel filter. More specifically, the filter $\mathbf{G}$ used is a 3×3 matrix of the form:

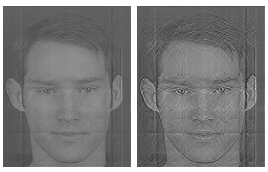


Fig. 4. Left: Result for SVD-MSF $d = 0.2$, Right: Result for SVD-MSF $d = 1.0$

$$\mathbf{G} = \begin{bmatrix} d & 2d & d \\ 0 & 0 & 0 \\ -d & -2d & -d \end{bmatrix} \tag{9}$$

where the parameter d was empirically determined to be in the range $[0.2, 0.8]$. Edge detection when applied to matrices $\mathbf{U}_{CA}$ and $\mathbf{V}_{CA}$ results in matrices $\mathbf{U}_{final}$ and $\mathbf{V}_{final}$. Similar to the SVD-CA step, only a percentage of the resulting matrix is used in computing the output image according to (7), (8). The output of this individual step is shown in Figure 4. After applying the above steps, the output image $\mathbf{P}$ is calculated, through the matrices $\mathbf{U}_{final}$, $\mathbf{S}_{CZ}$ and $\mathbf{V}_{final}$ using the formula:

$$\mathbf{P} = \mathbf{U}_{final}\mathbf{S}_{CZ}\mathbf{V}_{final}^{T} \tag{10}$$

In the following, this series of steps will be referred to as the SVD-DID method.

IV. Experimental Results

A. Database Description

The experiments where run on the XM2VTS and the Yale B facial image databases. From the XM2VTS database, 16 individuals where selected and used in the experimental process from the first recording. These individuals are facing the camera with a neutral background. The frontal images where isolated, and cropped to the face area. Finally, the images where converted to grayscale. This process resulted in a dataset with 388 train and 265 test samples from the 16 videos. Each sample has 128721 dimensions (401×321), with both test and train samples converted in vertexes with dimensions 1×128721. The Extended Yale B database includes images from 38 individuals under different lighting conditions. Train and test cases where defined having 1209 and 1205 samples, respectively, by randomly selecting half the images of each individuals. Each image has 1200 dimensions (30×40). These train and test samples where used to train classifiers in order to measure the efficiency of the proposed method. The two classifiers used are the K-Nearest Neighbor (KNN) Classifier with 3 neighbors and the Naive Bayes Classifier (NBC). These classifiers base recognition on the image texture. Other classifiers exist that employ information for the facial image structure based on the detection of a set of fiducial points. These classifiers do not perform as well as texture-based

classifiers. However, we still expect the proposed method to affect these classifiers since the accurate detection of the fiducial points will be hindered from texture manipulation.

B. Significance of each step in de-identification accuracy

In this section, we present and analyze the results from training and testing the efficiency of each of the steps described in Section III. The results are presented for each step with error percentages and the mean Mean Square Error (mMSE) for the test set of images, compared to the initial set.

To calculate the mMSE we assume that the images are in vector form with dimensions $np \times 1$, where np is the number of pixels in each image. With this in mind, this metric, is calculated using the following formula:

$$mMSE = \frac{1}{N_{im}} \sum_{i=1}^{N_{im}} \left[\frac{1}{np} \sum_{j=1}^{np} \left(M_i(j) - \hat{M}_i(j) \right)^2 \right], \quad (11)$$

where N_{im} is the total number of images, np is the number of image pixels, M_i is the i^{th} original image and finally $\hat{M}_i$ is the i^{th} output image of the applied method. All calculations for the mMSE are done with the images having values in the range $[0, 1]$, after they where divided by 255.

As mentioned above, the necessity to increase the luminosity of all pixels in the final image arises in order to counterbalance the darkening effect introduced by the algorithm steps. In the experiments, the values 0, 100 and 150 were used for reducing the darkening effect.

1) Results for SVD-CZ: Experimental results of setting the N largest singular values to zero are depicted in Tables I and II. It can be observed that, the increase of the number of zeroed singular values tends to increase the mMSE while, at the same time, the error rate is increased for both classifiers. Altering the number of nearest neighbors in the KNN classifier such as 1 and 5, yields the same results. These results are displayed for different number of zeroed coefficients and for different amounts of brightness added to the final image. Visual results can be seen in Figure 1. It can be easily seen from these figures that this method alone does not provide an acceptable output image, since too many visual artifacts are introduced that decrease the overall image quality, even by zeroing only a couple of the first singular values.

2) Results for SVD-CA: For the circular averaging filter, the error rates are displayed in Tables III and IV. The error rates where calculated in relation with the radius r of the circular filter and the amount of brightness that is applied for both databases. The mMSE in this case does not increase by increasing radius r. However, is shows a relevance to the added luminosity as well. On the other hand, error rates increase by increasing the radius value. Resulting images can be seen in Figure 3.

For this step, it was mentioned that only a percentage of the newly calculated matrices is used. By varying parameter α, we obtain the results in Table V. We conclude that parameter α affects the error rate of both classifiers. The parameters where $r = 10$, $\alpha = 0.8$. From this table it can be observed that by increasing parameter α the mMSE increases along with the error rate of the classifiers.

TABLE V
ERROR RATES FOR SVD-CA $r = 10$

Param. α	KNN (K=3)	Naive Bayes	mMSE
α = 0.5	52.08 %	68.30 %	0.0549
α = 0.8	53.21 %	83.02 %	0.0477
α = 1.0	59.25 %	86.79 %	0.0468

TABLE VIII
ERROR RATES FOR SVD-MSF $d = 0.5$

Param. α	KNN (K=3)	Naive Bayes	mMSE
α = 0.5	52.08 %	68.30 %	0.0512
α = 0.8	50.56 %	86.79 %	0.0454
α = 1.0	55.47 %	90.57 %	0.0453

3) Results for SVD-MSF: Applying the modified Sobel filter to the matrices, we obtain the error rates displayed in Tables VI and VII. The results are related with parameter d and the added luminosity. By increasing the value of parameter d we obtain higher mMSE but, generally, the error rates remain unchanged. As before, parameter α was set to 0.8. Image results of the method are displayed in Figure 4 for parameters $d = 0.5$, $\alpha = 0.8$.

In this method, altering parameter α, leads to the error rates in Table IX. The error rates are for the parameter d value $d = 0.5$ and added luminosity 100. In this case, altering α leads to a decrease of the mMSE and varying error rates.

Summarizing the results for each phase independently, we observe that some of these phases either degrade image quality to a great extend, or provide insufficient privacy protection. By merging all these phases in one method we obtain the results shown in the following section.

C. Results for SVD-DID

The SVD-DID method as a whole includes the three steps described in the previous sections (III-A,III-B,III-C). By applying these in the following order, i.e. SVD-CZ, SVD-CA and SVD-MSF, we derive this method that encompasses the advantages of all phases which are image quality and privacy protection. The defined parameters of this method can be altered to adjust the equilibrium between image quality and privacy protection, depending on the purpose of applying this method. The results for the full application of this method are displayed in Tables IX and X and Figures 5 and 6. The results in the tables are displayed in relation with parameter α, added luminosity and number of zeroed coefficients. Other visual results are displayed in Figure 7 for higher luminosity added to the image at 150. Figure 8 shows the result of applying a circular filter and a modified Sobel filter with inappropriate parameters.

In the SVD-DID method, numerous parameters are involved in composing the final image. These parameters give this method the flexibility required for modern privacy needs. A parameter that plays major role in the error rates of this method is the added luminosity. The value of 100 was generally chosen, since it displayed good visual output and low mMSE. Increasing luminosity over about 150, leads pixel values over

TABLE I
ERROR RATES FOR NUMBER OF ZEROED COEFFICIENTS (XM2VTS)

Zeroed Coefficients	Luminosity +0			Luminosity +100			Luminosity +150		
	KNN	**NBC**	**mMSE**	**KNN**	**NBC**	**mMSE**	**KNN**	**NBC**	**mMSE**
1	69.34 %	69.34 %	0.1903	55.47 %	55.47 %	0.0484	52.45 %	52.45 %	0.0927
2	90.57 %	90.57 %	0.1959	72.45 %	72.45 %	0.0523	72.45 %	72.45 %	0.0959
4	90.57 %	90.57 %	0.2001	83.02 %	83.02 %	0.0552	78.49 %	78.49 %	0.0981
8	93.21 %	93.21 %	0.2023	93.21 %	93.21 %	0.0569	79.25 %	79.25 %	0.0996

TABLE II
ERROR RATES FOR NUMBER OF ZEROED COEFFICIENTS (YALEB)

Zeroed Coefficients	Luminosity +0			Luminosity +100			Luminosity +150		
	KNN	**NBC**	**mMSE**	**KNN**	**NBC**	**mMSE**	**KNN**	**NBC**	**mMSE**
1	92.53 %	96.43 %	11.373 e-4	79.25 %	92.53 %	5.5735 e-4	90.12 %	97.51 %	13.426 e-4
2	93.61 %	97.34 %	11.910 e-4	97.26 %	96.51 %	6.0820 e-4	93.94 %	97.51 %	13.921 e-4
4	95.60 %	97.34 %	12.187 e-4	97.34 %	96.51 %	6.3574 e-4	95.93 %	97.51 %	14.195 e-4
8	96.93 %	97.34 %	12.312 e-4	97.34 %	97.51 %	6.4748 e-4	97.26 %	97.51 %	14.309 e-4

TABLE III
ERROR RATES FOR CIRCULAR AVERAGING FILTER (XM2VTS)

Filter Radius	Luminosity +0			Luminosity +100			Luminosity +150		
	KNN	**NBC**	**mMSE**	**KNN**	**NBC**	**mMSE**	**KNN**	**NBC**	**mMSE**
5	85.66 %	67.55 %	0.0815	49.43 %	83.02 %	0.0518	80.38 %	72.08 %	0.1524
10	86.04 %	69.06 %	0.0895	53.21 %	83.02 %	0.0477	80.38 %	72.08 %	0.1422
20	90.57 %	71.70 %	0.0935	50.06 %	86.79 %	0.0459	80.38 %	72.08 %	0.1375

TABLE IV
ERROR RATES FOR CIRCULAR AVERAGING FILTER (YALEB)

Filter Radius	Luminosity +0			Luminosity +100			Luminosity +150		
	KNN	**NBC**	**mMSE**	**KNN**	**NBC**	**mMSE**	**KNN**	**NBC**	**mMSE**
5	94.85 %	96.18 %	1.1050 e-4	90.54 %	97.51 %	1.4683 e-4	97.01 %	97.51 %	4.0146 e-4
10	94.61 %	96.43 %	1.2097 e-4	89.88 %	97.51 %	1.4204 e-4	96.93 %	97.51 %	3.8905 e-4
20	94.77 %	96.43 %	1.2632 e-4	89.88 %	97.51 %	1.3986 e-4	96.85 %	97.51 %	3.8310 e-4

the 255 limit of 8 bit images increasing the mMSE and the error rates of the classifiers. Hence, the attacker cannot identify the transformed images by increasing the image luminosity. Luminosity is dependent on image capture settings and the value of parameter N, which can be selected depending on the size of the image as well as the level of target privacy. Generally, reducing the first four coefficients is sufficient to achieve high error rates. In the case of parameter α, the value of 0.8 is preferred. Further increasing α, a larger portion of the distorted matrix is used to compose the final image. The value for parameter d should be selected in conjunction with parameter r. These parameters can be selected in order to introduce more visual artifacts in the output image (low r, high d). This can be used to hinder recognition from the mentioned classifiers as well as others based on fiducial points giving false negatives on automatic fiducial point recognition algorithms. The proposed values are 0.5 for d and 10 for parameter r.

From the above results and discussion, we observe that with the correct selection of parameter values, we can attain high levels of privacy, while maintaining acceptable image quality. Error rates for both classifiers are high for both databases with maximum error rates at 97.36 % (NBC) for the XM2VTS database and 97.51 % (NBC) for the YaleB database.

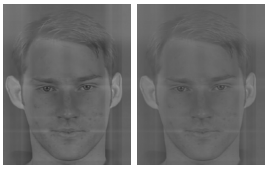


Fig. 5. Left: Result for SVD-DID for N=1, luminosity +100, α=0.5, Right: Result for SVD-DID for N=1, luminosity +100, α=0.8 (r = 10 and d = 0.5)

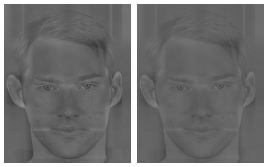


Fig. 6. Left: Result for SVD-DID for N=2, luminosity +100, α=0.5, Right: Result for SVD-DID for N=2, luminosity +100, α=0.8 (r = 10 and d = 0.5)

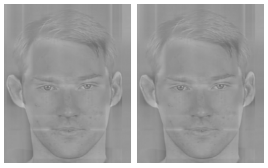


Fig. 7. Left: Result for SVD-DID for N=2, luminosity +150, α=0.5, Right: Result for SVD-DID for N=2, luminosity +150, α=0.8 (r = 10 and d = 0.5)

TABLE VI
ERROR RATES FOR MODIFIED SOBEL FILTERING (XM2VTS)

Value of d	Luminosity +0			Luminosity +100			Luminosity +150		
	KNN	**NBC**	**mMSE**	**KNN**	**NBC**	**mMSE**	**KNN**	**NBC**	**mMSE**
0.2	90.57 %	67.17 %	0.0978	50.57 %	86.79 %	0.0447	84.53 %	72.08 %	0.1335
0.5	90.57 %	67.17 %	0.0988	50.57 %	86.79 %	0.0447	85.66 %	72.08 %	0.1342
1.0	69.43 %	67.17 %	0.1041	49.81 %	86.79 %	0.0509	85.66 %	72.08 %	0.1397

TABLE VII
ERROR RATES FOR MODIFIED SOBEL FILTERING (YALEB)

Value of d	Luminosity +0			Luminosity +100			Luminosity +150		
	KNN	**NBC**	**mMSE**	**KNN**	**NBC**	**mMSE**	**KNN**	**NBC**	**mMSE**
0.2	94.11 %	96.35 %	1.3880 e-4	89.38 %	97.51 %	1.4197 e-4	96.85 %	97.51 %	3.8252 e-4
0.5	95.19 %	96.10 %	1.6637 e-4	90.04 %	97.34 %	1.7403 e-4	96.93 %	97.51 %	4.1433 e-4
1.0	95.52 %	95.44 %	4.4246 e-4	90.04 %	97.01 %	4.4916 e-4	96.93 %	97.51 %	6.8898 e-4

TABLE IX
ERROR RATES FOR SVD-DID (XM2VTS)

Zeroed Coefficients	Luminosity +0						Luminosity +100					
	$\alpha = 0.5$			$\alpha = 0.8$			$\alpha = 0.5$			$\alpha = 0.8$		
	KNN	**NBC**	**mMSE**	**KNN**	**NBC**	**mMSE**	**KNN**	**NBC**	**mMSE**	**KNN**	**NBC**	**mMSE**
0	65.66 %	59.25 %	0.0631	90.57 %	87.92 %	0.0977	52.83 %	68.30 %	0.0507	50.57 %	86.79 %	0.0447
1	90.57 %	97.36 %	0.1947	90.57 %	93.74 %	0.1971	76.60 %	93.21 %	0.0508	90.57 %	93.21 %	0.0527
2	90.57 %	97.36 %	0.1985	90.57 %	97.36 %	0.2000	90.57 %	93.21 %	0.0539	90.57 %	93.21 %	0.0551
4	93.21 %	97.36 %	0.2014	93.71 %	97.36 %	0.2022	93.21 %	93.21 %	0.0562	93.21 %	93.21 %	0.0569

TABLE X
ERROR RATES FOR SVD-DID (YALEB)

Zeroed Coefficients	Luminosity +0						Luminosity +100					
	$\alpha = 0.5$			$\alpha = 0.8$			$\alpha = 0.5$			$\alpha = 0.8$		
	KNN	**NBC**	**mMSE**	**KNN**	**NBC**	**mMSE**	**KNN**	**NBC**	**mMSE**	**KNN**	**NBC**	**mMSE**
0	93.03 %	93.94 %	8.3750 e-4	92.78 %	93.86 %	8.3746 e-4	89.21 %	97.51 %	1.5260 e-4	89.13 %	97.51 %	1.5266 e-4
1	93.53 %	97.34 %	2.5675 e-4	94.85 %	97.34 %	2.6036 e-4	97.01 %	97.18 %	1.2868 e-4	97.34 %	97.51 %	1.3220 e-4
2	95.85 %	97.34 %	2.6490 e-4	95.93 %	97.34 %	2.6653 e-4	97.34 %	97.51 %	1.3655 e-4	97.34 %	97.51 %	1.3814 e-4
4	96.68 %	97.34 %	2.6912 e-4	96.76 %	97.34 %	2.6972 e-4	97.34 %	97.51 %	1.4075 e-4	97.34 %	97.51 %	1.4131 e-4

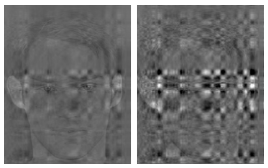


Fig. 8. Left: Result for SVD-DID for N=1, luminosity +100, α=0.8, $r = 10$ and d = 5 Right: Result for SVD-DID for N=1, luminosity +100, α=0.8, $r = 10$ and d = 10

V. CONCLUSIONS

We have proposed method that aims to limit the effectiveness of face identification methods, while retaining part of the initial visual quality. From the results above, it can be deducted that using the appropriate parameter values in each step, a high level of privacy can be attained. In the YaleB database, the highest error rate achieved was 97.51% and the highest error rate for the XM2VTS database was 97.36%. Despite the high error rate, the end product of these methods can be characterized as acceptable for everyday use.

This method, when applied to the initial images, tend to have a smoothing effect on the image, while introducing visual artifacts. Also, by applying the various methods and filters there exists the tendency to darken the image, which is counterbalanced, by adding a constant value to the output image, in order to preserve adequate visual information so that the faces can be identified by human viewers. The combination of these effects reduces the identification accuracy of automatic face identity classifiers.

From the error rates and visual results we can conclude that the proposed SVD-DID method serves the purpose of protecting privacy and providing a visually acceptable output. A drawback of the proposed SVD-DID method is that it is irreversible, i.e., once the image is filtered it cannot return to its original form. Future work is directed towards the implementation of reversible de-identification methods.

ACKNOWLEDGMENT

The research leading to these results has been partially supported from COST, Action IC1206.

REFERENCES

[1] P. Agrawal and P. J. Narayanan. Person de-identification in videos.
[2] G. Bebis. Singular value decomposition (svd). http://www.cse.unr.edu/~bebis/MathMethods/SVD/lecture.pdf.
[3] V. Blanz, S. Romdhani, and T. Vetter. Face identification across different poses and illuminations with a 3d morphable model.
[4] B. Driessen and M. Drmuth. Achieving anonymity against major face recognition algorithms.
[5] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, MD, 2013.
[6] R. Gross, L. Sweeney, J. Cohn, F. de la Torre, and S. Baker. Face de-identification.
[7] Perspectives in Biological Engineering Course MIT BE.400 / 7.548. Singular value decomposition (svd) tutorial. http://web.mit.edu/be.400/www/SVD/Singular_Value_Decomposition.htm.
[8] E. Newton, L. Sweeney, and B. Malin. Preserving privacy by de-identifying facial images.
[9] P. J. Phillips. Privacy operating characteristic for privacy protection in surveillance applications.
[10] S. Tansuriyavong and Shin ichi Hanaki. Privacy protection by concealing persons in circumstantial video image.