A FRAMEWORK FOR THE VALIDATION OF PRIVACY PROTECTION SOLUTIONS IN VIDEO SURVEILLANCE

Frédéric Dufaux and Touradj Ebrahimi

Multimedia Signal Processing Group, Ecole Polytechnique Fédérale de Lausanne (EPFL)
CH-1015 Lausanne, Switzerland
Email: frederic.dufaux@epfl.ch, touradj.ebrahimi@epfl.ch

ABSTRACT

The issue of privacy protection in video surveillance has drawn a lot of interest lately. However, thorough performance analysis and validation is still lacking, especially regarding the fulfillment of privacy-related requirements. In this paper, we put forward a framework to assess the capacity of privacy protection solutions to hide distinguishing facial information and to conceal identity. We then conduct rigorous experiments to evaluate the performance of face recognition algorithms applied to images altered by privacy protection techniques. Results show the ineffectiveness of naïve privacy protection techniques such as pixelization and blur. Conversely, they demonstrate the effectiveness of more sophisticated scrambling techniques to foil face recognition.

Keywords— Privacy, scrambling, face recognition

1. INTRODUCTION

Privacy protection is quickly becoming a very central issue in video surveillance. While video surveillance can help repress crime and terrorism, hence benefiting the society, the widespread use of security cameras has led to well documented forms of abuse, including: criminal abuse by law enforcement officers, institutional abuse by spying upon and harassing political activists, abuse for personal purpose such as stalking women or estranged girlfriends/spouses, discrimination including racial discrimination, voyeurism where bored male operators spy on women, and release of public camera footage in the public domain. Moreover, its big brother nature is hindering wider acceptance of video surveillance.

The perspective of forthcoming powerful video analytics tools, combined with pervasive networks of dense cameras is further raising the threat of privacy loss.

Fortunately, recent research results have shown that new technologies are emerging with the potential to effectively protect privacy, without hampering video surveillance tasks. These results challenge the common conjecture that increased security should incur a loss in privacy. Recent overviews are given in [1][2].

The system introduced in [3] relies on computer vision to analyze the video content. Depending on users' access-control rights, different versions of the video are then presented where privacy-sensitive information is removed.

Privacy filters are proposed in [4], which operate on sensor data to remove privacy-sensitive information. These filters are specified using a privacy grammar.

Data corresponding to faces is encrypted in [5] in order to conceal identity. The process is reversible for authorized users in possession of the secret encryption key. Similarly, a scheme for secure coding of arbitrarily shaped visual objects is presented in [6].

The methods in [7][8] propose privacy protection solutions for JPEG 2000 video. Conditional access control techniques are proposed in [7] to scramble Regions of Interest (ROI), e.g. people or faces. The scrambling is applied either in wavelet-domain or codestream-domain. In [8], code-blocks corresponding to ROI are trimmed down to the lowest quality layer of the codestream. Subsequently, the quality of the ROI can be decreased by limiting the video bit rate.

Two efficient region-based transform-domain and codestream-domain scrambling techniques are proposed in [9] to hide privacy-sensitive information in MPEG-4 video. In the first approach, the sign of selected transform coefficients is pseudo-randomly inverted during encoding. In the second approach, bits of the codestream are pseudo-randomly flipped after encoding. The region-based transform-domain scrambling is extended to H.264/AVC in [10].

The technique in [11] removes privacy-sensitive information from the video sequence. A perceptually-based compressed-domain watermarking technique is then used to securely embed this data in the video stream. Similarly, a secure reversible data hiding technique is introduced in [12] for privacy data embedding. A framework for privacy data management is also proposed to allow individual users to control access to their private data.

Face recognition techniques pose the threat to automatically identify people in a video surveillance scene. This issue is addressed in [13], where an algorithm is introduced to de-identify faces such that many facial characteristics are preserved but the face cannot be reliably

recognized. It is also shown that simple ad-hoc de-identification methods do not prevent successful face recognition.

However, although the issue of privacy protection has drawn a lot of interest, thorough performance analysis is still lacking. In particular, it is paramount to validate proposed privacy protection solutions against user and system requirements for privacy. Moreover, it is still unclear whether current privacy protection approaches can be efficiently integrated into existing surveillance architecture and deployed in large scale systems.

The objective of this paper is to define a framework to assess the capacity of privacy protection solutions to hide distinguishing facial information and to conceal identity. For this purpose, we use the Colorado State University (CSU) Face Identification Evaluation System [14] to evaluate the performance of face recognition algorithms applied to images altered by privacy protection techniques. By performing extensive and comprehensive experiments on the FERET database [15], we show the ineffectiveness of naïve privacy protection techniques such as pixelization and blur. Conversely, we demonstrate the effectiveness of more sophisticated scrambling techniques to foil face recognition.

This paper is structured as follow. An outline of four privacy protection approaches under consideration is given in Sec 2. Next, a framework for face identification evaluation is presented in Sec. 3. In order to validate privacy protection solutions, performance assessment using this framework is analyzed in Sec. 4. Finally, conclusions are summarized in Sec 5.

2. PRIVACY PROTECTION APPROACHES UNDER CONSIDERATION

In this section, we briefly describe four approaches for privacy protection that we will subsequently evaluate for their capability to hide facial information and to provide anonymity.

As reference, we first consider two naïve methods, applying simple pixelization or Gaussian blur. Note that in both cases, privacy-sensitive information is lost and the process is irreversible.

We also consider two more sophisticated ROI-based transform-domain scrambling methods [10]. Both methods are applied jointly with H.264/AVC encoding [16], which is becoming the prevalent format in video surveillance systems. Two slice groups are defined using Flexible Macroblock Ordering to distinguish between the scrambled ROI and the unscrambled background. The first method pseudo-randomly inverts the sign of transform coefficients of blocks belonging to ROI. The second one applies a pseudo-random permutation of the transform coefficients in blocks corresponding to ROI. These two methods are fully reversible. Namely, authorized users, in possession of a secret encryption key, can reverse the scrambling process and recover the truthful scene. Conversely, other users

obtain a video sequence where ROI have severe noise, concealing privacy-sensitive information.

These four approaches to provide anonymity are detailed in the following subsections.

2.1. Pixelization

We first consider pixelization as a naïve approach for privacy protection. Pixelization consists in noticeably reducing resolution in ROI. In practice, it can be achieved by substituting a square block of pixels with its average.

Pixelization is commonly used in television news and documentaries in order to obscure the faces of suspects, witnesses or bystanders to preserve their anonymity. The same technique is also used to censor nudity or to avoid unintentional product placement on television.

One drawback of this approach is that integrating pixels along trajectories over time may allow to partly recovering the concealed information.

2.2. Gaussian Blur

The second naïve approach for privacy protection removes details in ROI by applying a Gaussian low pass filter. More precisely, the image is convolved with a 2D Gaussian function defined by

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \qquad (1)$$

where σ is the standard deviation. Blurring is sometimes preferred to pixelization in order to obscure privacy-sensitive information.

2.3. Scrambling by Random Sign Inversion

Next, we consider a ROI-based transform-domain scrambling method for H.264/AVC. It scrambles the quantized transform coefficients of each 4x4 block of the ROI by pseudo-randomly flipping their sign [10].

More specifically, defining the vector of quantized transform coefficients $qcoeff[i]$ with $i=0..15$, the scrambling consists in performing the following operation for each i

$$qcoeff[i] = \begin{cases} -qcoeff[i] & \text{if } random\_bit = 1 \\ +qcoeff[i] & \text{otherwise} \end{cases} \qquad (2)$$

Besides being fully reversible, this method offers a number of advantages. The same scrambled stream is transmitted to all users independently from their access rights. The scrambling is confined to ROI, whereas the background remains unaltered. Finally, it has a small impact in terms of coding efficiency, and requires a low computational complexity.

2.4. Scrambling by Random Permutation

Finally, we consider an alternative ROI-based transform-domain scrambling method for H.264/AVC. In this method, a random permutation is applied to rearrange the order of transform coefficients in 4x4 blocks corresponding to ROI [10]. The random permutation is expressed as follow

$$\begin{pmatrix} 0 & 1 & ... & 14 & 15 \\ x_0 & x_1 & ... & x_{14} & x_{15} \end{pmatrix} \quad (3)$$

The Knuth shuffle is used to generate a permutation of *n* items with uniform random distribution. More explicitly, it starts from the identity permutation and scans through each position *i* from 0 to 14, swapping the element currently at position *i* with the element at an arbitrarily chosen positions from *i* through 15.

This method provides the same advantages as the scrambling by random sign inversion.

3. FRAMEWORK FOR FACE IDENTIFICATION EVALUATION

The objective of this paper is to validate the anonymity functionality of privacy protection approaches. For this purpose, we use the CSU Face Identification Evaluation System (FIES), which provides standard face recognition algorithms and standard statistical methods for assessing performances [14]. A brief description of the CSU FIES is given hereafter.

3.1. Face Recognition

We consider two face recognition algorithms, namely, Principal Components Analysis (PCA) [17] and Linear Discriminant Analysis (LDA) [18].

In PCA, also known as eigenfaces, a linear transformation is applied to rotate feature vectors from the initially large and highly correlated subspace to a smaller and uncorrelated subspace. PCA has shown to be effective for face recognition. Firstly, it can be used to reduce the dimensionality of the feature space. Secondly, it eliminates statistical covariance in the transformed feature space. In other words, the covariance matrix for the transformed feature vectors is always diagonal.

LDA aims at finding a linear transformation which stresses differences between classes while lessening differences within classes, where a class corresponds to all images of a given individual. The resulting transformed subspace is linearly separable between classes. In [18], PCA is first performed to reduce the feature space dimensionality. LDA is then applied to further decrease the dimensionality while safeguarding the distinctive characteristics of the classes. The final subspace is obtained by multiplying the PCA and LDA basis vectors.

3.2. Face Identification and Evaluation System

The CSU FIES is composed of four main components: image pre-processing, training, testing and performance analysis [14].

The preprocessing step aims at reducing detrimental variations between images. Faces are firstly aligned using information about the eye coordinates. Then, pixel values are equalized and contrast and brightness are normalized.

The next step is training. Its purpose is to create the subspace into which test images are subsequently projected and matched. Training is performed using a training set of images.

In the testing step, a distance matrix is computed in the transformed subspace for all test images. In our experiments, we use a Euclidian distance for PCA and the soft distance proposed in [18] for LDA. At this stage, two image sets are defined: the gallery set is made of known faces, whereas the probe set corresponds to faces to be recognized.

Finally, face recognition performance is analyzed. More specifically, a cumulative match curve is generated. For this purpose, for each probe image, the recognition rank is computed. Namely, a rank 0 means that the best match is of the same subject, a rank 1 means that the best match is from another person but the second best match is of the same subject, etc. Then, the cumulative match curve is obtained by summing the number of correct matches for each rank.

4. PRIVACY PROTECTION PERFORMANCE ASSESSMENT RESULTS

We now describe experiments carried out in order to assess privacy protection solutions. Results are then reported and analyzed.

4.1. Test data

In this paper, we use the grayscale Facial Recognition Technology (FERET) database [15] to carry out experiments. Indeed, this database is widely used for face recognition research, although it is not representative of typical video surveillance footage. From this database, we consider a subset of 3368 images of frontal faces for which eye coordinates are available. The images have 256 by 384 pixels with eight-bit per pixel. We further consider two series of images denoted by 'fa' and 'fb'. The 'fa' indicates a regular frontal image, and the 'fb' indicates an alternative frontal image, taken within seconds of the corresponding 'fa' image, where a different facial expression was requested from the subject.

In our experiments, we use standard training, gallery and probe sets from the FERET test. More specifically, the training set includes 501 images from the 'fa' series. In turn, the gallery set is composed of 1196 from the 'fa' series, whereas the probe set is made of 1195 images from the 'fb' series.

4.2. Sample Images with Privacy Protection

The results of the four privacy protection approaches described in Sec. 2, namely pixelization, Gaussian blur, scrambling by random sign inversion and scrambling by random permutation are illustrated in Fig. 1 for a sample image of the FERET database.

The four privacy protection techniques could be used in video sequences, and to hide information in ROI. Nevertheless, hereafter we simply apply them to still images and to obscure the whole picture. In particular, both region-based transform-domain scrambling approaches are applied jointly with H.264/AVC encoding. In this paper, as we deal with still images, we straightforwardly use the H.264/AVC Intra coding mode.

Note that the effect of the privacy protection techniques could be slightly different when applied to video sequences.

4.3. Face Recognition Performance Analysis

We now evaluate the capacity of privacy protection solutions to hide distinguishing facial information in order to foil face recognition techniques and hence to conceal the identity of a person.

Similarly to [13], we study different types of attack. In the first round of experiments, we consider a simple attack where training and gallery sets are made of unaltered images. Conversely, probe set corresponds to images with privacy protection. In other words, altered images are merely processed by the face recognition algorithms without taking into account the fact that privacy protection tools have been applied.

Fig. 2 and Fig. 3 show cumulative match curves for PCA and LDA respectively, comparing the recognition rate as a function of the rank for original image data as well as for the four considered privacy protection approaches.

It can be observed that for both PCA and LDA schemes applied on original images, recognition rate is superior to 70% at rank 0 (i.e. the best match is of the same subject as the probe), and superior to 90% at rank 50.

When applying a Gaussian blur, the performance drops radically for LDA. However, recognition rate remains high for PCA with 56% success at rank 0. Pixelization fares worse. The recognition rate is 56% and 13% at rank 0 for PCA and LDA respectively.
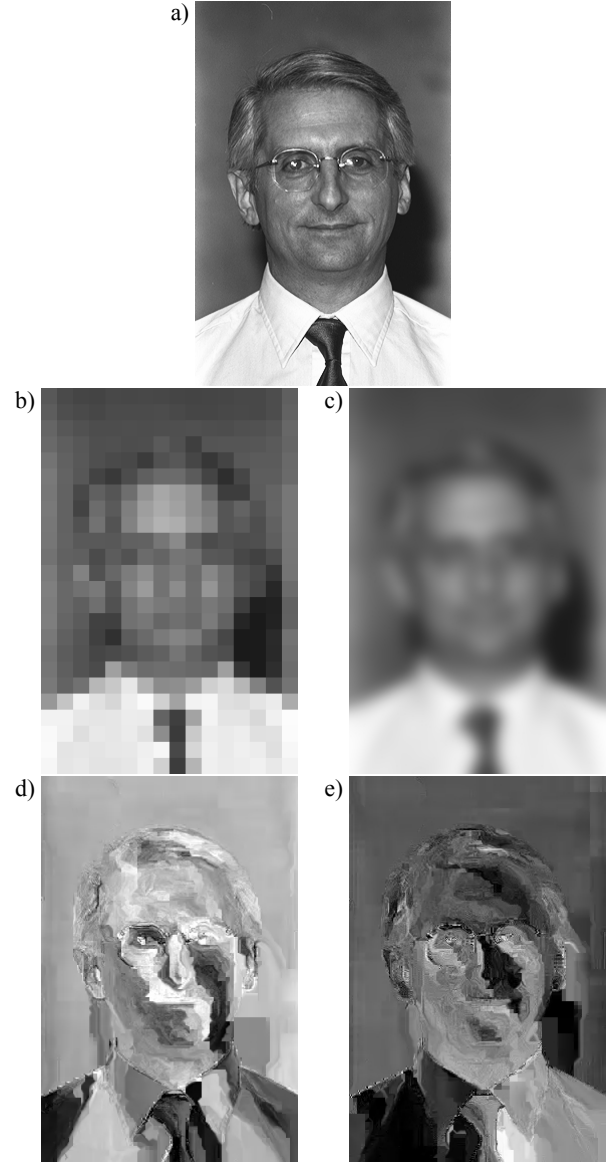

Fig. 1. Examples of privacy protection approaches: a) original image, b) pixelization with a block size of 16, c) Gaussian blur with a standard deviation σ=8, d) scrambling by random sign inversion, e) scrambling by random permutation.

However, results clearly show that both region-based transform-domain scrambling approaches are successful at hiding identity. The recognition rate is nearly 0% at rank 0, and remains below 10% at rank 50, for both PCA and LDA algorithms. In addition, it can be observed that both random sign inversion and random permutation schemes achieve nearly the same performance.

In the second round of experiments, we consider a more sophisticated attack. Namely, privacy protection tools are now applied to all images in the training, gallery and probe sets. This corresponds to an attacker which gets access to protected data. Alternatively, an attacker may attempt replicating the alteration due to privacy protection

techniques on his own training and gallery sets. Fig. 4 and Fig. 5 show corresponding cumulative match curves for PCA and LDA respectively.

With Gaussian blur, the performance remains nearly identical. It even improves slightly for LDA. Pixelization is not much better at hiding facial information. The recognition rate is still 45% and 17% at rank 0 for PCA and LDA respectively.

Finally, both region-based transform-domain scrambling approaches are again successful at hiding identity. The recognition rate is nearly 0% at rank 0 for both PCA and LDA algorithms.
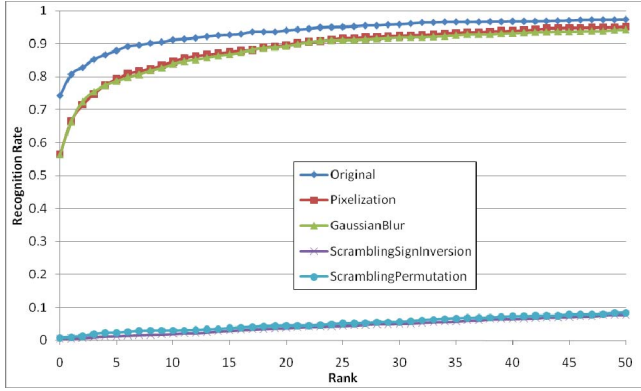

Fig. 2. Cumulative match curve for PCA with Euclidian distance: performance comparison of privacy protection solutions (unaltered images in training and gallery sets, altered images in probe set).
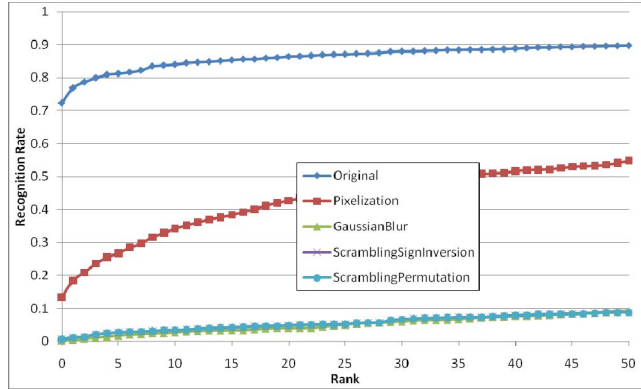

Fig. 3. Cumulative match curve for LDA with soft distance: performance comparison of privacy protection solutions (unaltered images in training and gallery sets, altered images in probe set).
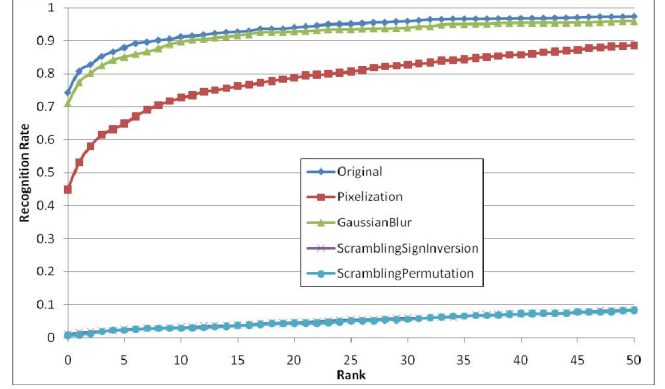

Fig. 4. Cumulative match curve for PCA with Euclidian distance: performance comparison of privacy protection solutions (altered images in training, gallery and probe sets).
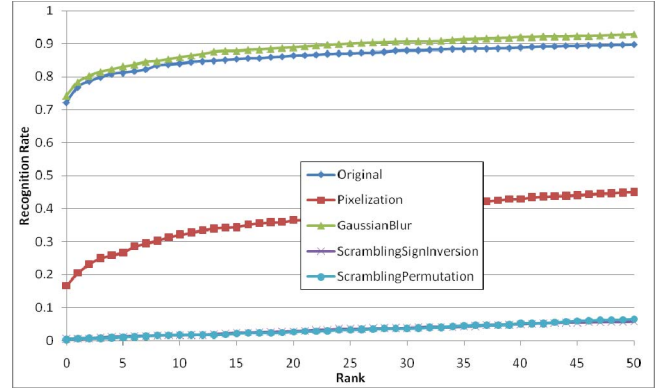

Fig. 5. Cumulative match curve for LDA with soft distance: performance comparison of privacy protection solutions (altered images in training, gallery and probe sets).

5. CONCLUSIONS

This paper is a step forward in the assessment of privacy protection solutions. More specifically, we have described a framework to verify the effectiveness of privacy protection techniques at hiding distinguishing facial information and hence concealing identity.

We have conducted rigorous and comprehensive experiments using PCA and LDA face recognition algorithms on the FERET database. Results have shown that applying Gaussian blur or pixelization is ineffective at providing anonymity. In both cases, the recognition rate remains significant. Finally, results have shown that region-based transform-domain scrambling approaches are successful at hiding identity, with the recognition rate dropping to nearly 0%.

Future work will concentrate in further analyzing the performance of privacy protection solutions, verifying that they can successfully address privacy issues. In addition, it is important to carry out experiments using more realistic video surveillance footage. It is also imperative to better understand user and system requirements regarding privacy protection. Finally, performance analysis should also

include the impact on compression efficiency, complexity, and security against attacks.

6. REFERENCES

[1] A. Cavallaro, "Privacy in Video Surveillance", IEEE Signal Proc. Magazine, vol. 24, no. 2, pp. 168-169, March 2007.

[2] A. Senior, "Protecting Privacy in Video Surveillance" Springer, 2009.

[3] A.W. Senior, S. Pankanti, A. Hampapur, L. Brown, Y.-L. Tian, A. Ekin, J. Connell, C.F. Shu and M. Lu, "Enabling Video Privacy through Computer Vision", IEEE Security and Privacy, vol. 3, no.3, pp. 50-57, May-June 2005.

[4] D. A. Fidaleo, H.-A. Nguyen, M. Trivedi, "The networked sensor tapestry (NeST): a privacy enhanced software architecture for interactive analysis of data in video-sensor networks", Proc. of the ACM 2nd Int. Workshop on Video Surveillance & Sensor Networks, New York, NY, 2004.

[5] T.E. Boult, "PICO: Privacy through Invertible Cryptographic Obscuration", IEEE/NFS Workshop on Computer Vision for Interactive and Intelligent Environments, Nov. 2005.

[6] K. Martin, and K.N. Plataniotis, "Privacy Protected Surveillance Using Secure Visual Object Coding", IEEE Trans. on Circuits and Systems for Video Technology, vol. 18, no. 8, pp. 1152-1162, Aug. 2008.

[7] F. Dufaux, and T. Ebrahimi, "Video Surveillance using JPEG 2000", in SPIE Proc. Applications of Digital Image Processing XXVII, Denver, CO, Aug. 2004.

[8] I. Martinez Ponte, X. Desurmont, J. Meessen, and J.-F. Delaigle, "Robust Human Face Hiding Ensuring Privacy" in Proc. of International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), Montreux, Switzerland, April 2005.

[9] F. Dufaux and T. Ebrahimi, "Scrambling for Privacy Protection in Video Surveillance Systems", IEEE Trans. on Circuits and Systems for Video Technology, vol. 18, no. 8, pp. 1168-1174, Aug. 2008.

[10] F. Dufaux and T. Ebrahimi, "H.264/AVC Video Scrambling for Privacy Protection", in Proc. IEEE International Conference on Image Processing, San Diego, CA, Oct. 2008.

[11] W. Zhang, S.S. Cheung, and M. Chen, "Hiding privacy information in video surveillance system", in Proc. IEEE International Conference on Image Processing, Genoa, Italy, Sept. 2005.

[12] S.S. Cheung, J.K. Paruchuri, T.P. Nguyen, "Managing Privacy Data in Pervasive Camera Networks", in Proc. IEEE International Conference on Image Processing, San Diego, CA, Oct. 2008

[13] E. Newton, L. Sweeney, and B. Malin, "Preserving Privacy by De-identifying Face Images", IEEE Trans. on Knowledge and Data Engineering, vol. 17, no. 2, pp. 232-243, February 2005.

[14] Evaluation of face recognition algorithms web site, http://www.cs.colostate.edu/evalfacerec.

[15] The Facial Recognition Technology (FERET) database, http://www.itl.nist.gov/iad/humanid/feret/feret_master.html.

[16] T. Wiegand, G.J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H.264/AVC Video Coding Standard", IEEE Trans. on Circuits and Systems for Video Technology, vol. 13, no. 7, pp. 560-576, July 2003.

[17] M.A. Turk and A.P. Pentland, "Face Recognition Using Eigenfaces", in Proc. of IEEE Conference on Computer Vision and Pattern Recognition, Maui, HI, June 1991.

[18] W. Zhao, R. Chellappa, and A. Krishnaswamy, "Discriminant analysis of principal components for face recognition", in Wechsler, Philips, Bruce, Fogelman-Soulie, and Huang, editors, "Face Recognition: From Theory to Applications", pp. 73–85, 1998.