TOWARDS OPTIMAL DISTORTION-BASED VISUAL PRIVACY FILTERS

Pavel Korshunov and Touradj Ebrahimi

Multimedia Signal Processing Group, EPFL, Lausanne, Switzerland

ABSTRACT

The widespread usage of digital video surveillance systems has increased the concerns for privacy violation. Since video surveillance systems are invasive, it is a challenge to find an acceptable balance between privacy of the public under surveillance and security related features of the systems. Many privacy protection tools have been proposed for preserving privacy, ranging from such simple methods like blurring or pixelization to more advanced like scrambling and geometrical transform based filters. However, for a given filter implemented in a practical video surveillance system, it is necessary to know the strength with which the filter should be applied to protect privacy reliably. Assuming an automated surveillance system, this paper objectively investigates several privacy protection filters with varying strength degrees and determines their optimal strength values to achieve privacy protection. To this end, five privacy filters were applied to images from FERET dataset and the performance of three recognition algorithms was evaluated. The results show that different privacy protection filters influence the accuracy of different versions of face recognition differently and this influence depends both on the robustness of the recognition and the type of distortion filter.

Index Terms— Privacy protection, video surveillance, security optimization, privacy strength

1. INTRODUCTION

More and more video surveillance systems are adopting privacy protection mechanisms in response to the public concern over unwelcome intrusion into personal lives. These protection techniques or privacy filters vary from such simple approaches like blurring, pixelization, or masking to more advanced methods, including encryption [1], scrambling [2], anonymization [3], and geometrical-based [4] approaches.

Most of the privacy filters distort a specified visual region in image or video with the goal to obfuscate sensitive personal information. Typical filter has a certain level of strength, for instance, the value of standard deviation of a Gaussian blur. The strength of such filter is often chosen in an ad hoc manner, by following common sense considerations such that the resulted image would be pleasant, less distracting, or would retain some general visual form. Another approach [5] is to choose a privacy filter's strength in such a way that intelligibility of video surveillance is preserved. However, in scenarios when the privacy protection has higher priority, for instance in people counting surveillance applications, highway monitoring, or crowd analysis, a filter's strength should be constrained to insure certain minimal level of privacy protection.

In this paper, we assume an automated video surveillance system that relies on video analytics for its operation. This assumption is reasonable, given how widespread and largely deployed the surveillance systems are today, which forces them to rely on video analytics in order to reduce the cost of the surveillance and increase their scalability. In such systems, privacy can be considered to be well protected if the performance of a privacy intrusive video analytic, such as face recognition, drops below an acceptable level. Hence, in a practical surveillance system, given a privacy filter, one needs to determine the constraint value of the filter's strength, which would insure the required level of privacy protection. And once this privacy level is achieved, other considerations, such as intelligibility or pleasantness, can be accommodated for.

Therefore, we consider five distortion-based privacy filters with varying strength parameters, including Gaussian blurring with varying standard deviation, pixelization with varying size of the averaging block, masking with varying opacity, warping [6] with varying number of tiles, and morphing [4] with varying intensity level. We investigate the influence of the filters' strength parameters on performance of three state-of-art face recognition algorithms implemented in OpenCV: based on Principal Component Analysis (PCA) [7], referred to as 'Eigen' in the paper, based on Linear Discriminant Analysis (LDA) [8], referred to as 'Fisher', and based on local features (LBP) [9], referred to as 'LBPH'. We use publicly available FERET dataset [10], which is a common dataset to test the performance of face recognition algorithms.

The paper is organized as follows. Section 2 presents the related work. Section 3 describes the experimental methodology of finding strength for each pair of privacy filter and recognition algorithm. Section 4 provides and discusses the evaluation results. Section 5 concludes the paper.

This work has been conducted in the framework of EC funded Network of Excellence VideoSense and COST IC1003 European Network on Quality of Experience in Multimedia Systems and Services QUALINET.

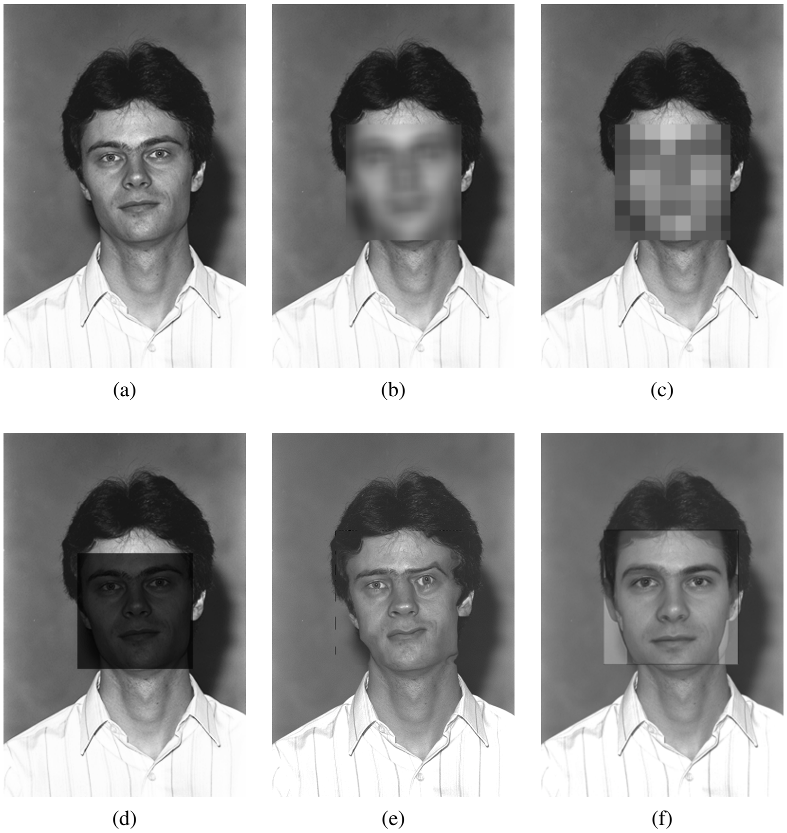

Fig. 1: An example image (a) from FERET dataset with the following privacy protection filters applied: (b) blurring with 17 kernel size, (c) pixelization with 16 averaging block size, (d) masking with 0.7 opacity, (e) warping with strength 10, and (f) morphing with intensity 0.4.

2. RELATED WORK

Newton *et al.* [11] argued that several primitive privacy filters cannot adequately protect from the successful face recognition, because recognition algorithms are robust. The robustness of face recognition and detection algorithms to primitive distortions is also reported in [12]. In the work by Dufaux *et al.* [13], a framework is defined to evaluate the performance of face recognition algorithms applied to images altered by various obfuscation methods, based on the Face Identification Evaluation System (FIES). Experiments using the FERET database showed the ineffectiveness of naïve face obfuscation techniques such as pixelization and blurring in hindering recognition performance. The authors argue that more sophisticated scrambling techniques are more effective in impeding face recognition. The above studies assume a certain strength level of the privacy protection tools such as blurring or pixelization. However, in a practical surveillance system, one would use the largest possible strength of, for instance, pixelization filter to make sure a given recognition algorithm still fails. It is also important to understand that each recognition algorithm would have different tolerance levels for different privacy protection filters.

Several works [5, 14] also considered the problem of finding the balance between the ability of human guards to perform a surveillance task and adequate protection of privacy. The authors argued that since privacy is a subjective notion, the evaluation should be done subjectively. Then, the authors define a subjective methodology for evaluation of privacy protection tools and propose a subjective evaluation protocol, focusing on two important aspects: (i) how much of the privacy is protected by such a tool and (ii) how much it impacts the efficiency of the underlying surveillance task (intelligibility). The pixelization filter shows the best performance in terms of balancing between privacy protection and allowing sufficient intelligibility. Masking filter, instead, demonstrates the highest privacy protection with low incorrectness and high uncertainty, which can be suitable for the higher security surveillance applications. This paper differs in the definition of privacy protection, because the goal now is to find a strong enough privacy filter that decreases the accuracy of an intrusive video analytic. This assumption results in privacy being an objective notion, which is applicable in automated surveillance systems.

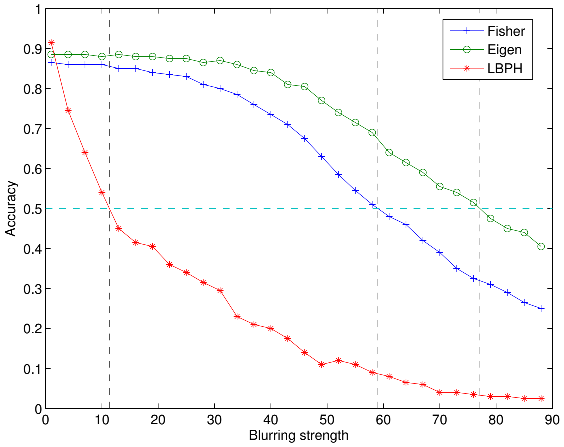

Fig. 2: Recognition results for blurring filter

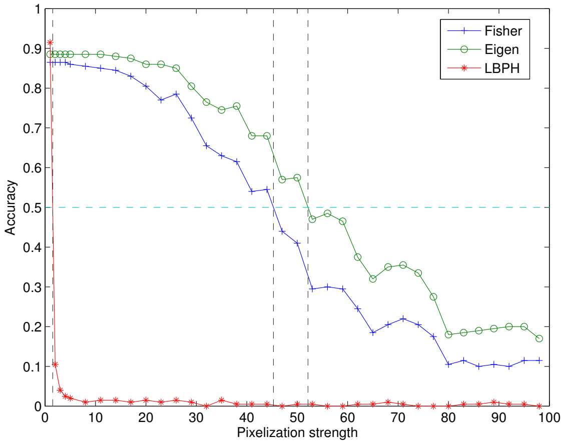

Fig. 3: Recognition results for pixelization filter

Another study [15] considered video surveillance system using video analytics but the work was focused on finding the privacy-intelligibility tradeoff using objective metrics without the consideration to insure the reliability of privacy protection algorithms. This paper considers privacy protection as a practical goal and demonstrates the importance of carefully selecting the strength of privacy protection filters that are adequate to the video analytics employed by the surveillance system.

3. EVALUATION FRAMEWORK

We investigate the influence of strength parameters of several privacy filters, such as Gaussian blurring, pixelization, masking, warping [6], and morphing [4], on the performance of three recognition algorithms: PCA based [7] referred to as ‘Eigen’, LDA based [8] referred to as ‘Fisher’, and LBP based [9] referred to as ‘LBPH’, to determine the minimal threshold strength values that would ensure adequate privacy protection. Figure 1 demonstrates the result of privacy filters applied to an example image.

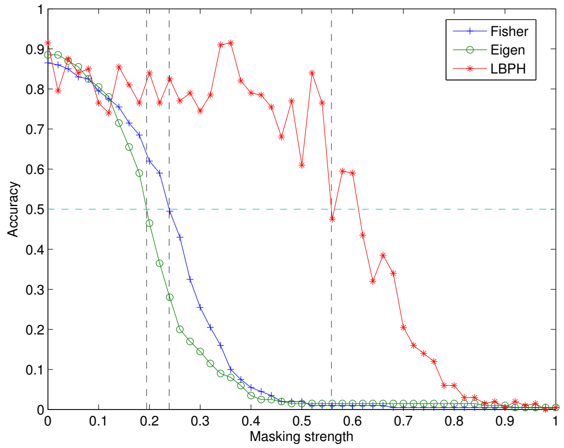

Fig. 4: Recognition results for masking filter

In the experiments, we use a subset of 200 images from FERET dataset [10], which consists of $14'051$ gray scale images of persons with clearly visible faces and different facial expressions under various environmental conditions, including illumination, orientation, appearance and age variations. We use the recommended subset ‘fa’ as the gallery set and subset ‘fb’ as the probe set for face recognition algorithms in all evaluation tests. The recognition is performed on the facial regions detected by Viola-Jones face detection [16].

The experiments are conducted as follows. For a given privacy filter, we change the strength of the filter with a small step and for each strength value, the faces from the ‘fb’ subset of the dataset are distorted by this filter accordingly. A given recognition algorithm is then applied to the distorted faces and rank 1 value of the Cumulative Matching Characteristic (CMS) is computed, which determines algorithm’s accuracy for the current filter’s strength. Testing different strength values, we vary the value of the standard deviation of Gaussian for blurring filter from 1 to 90 with step 3, the size of the block for pixelization filter from 1 to 100 with step 3, the opacity of masking filter from 0 to 1.0 with step 0.02, the number of tiles in warping filter from 25 to 1 with step 1 (the lower number of tiles results in a stronger filter), and the intensity parameter of morphing filter from 0 to 1.0 with step 0.04, using fixed interpolation value of 0.8, as recommended in [4].

4. RESULTS

Figures 2-6 show recognition results for all filters and recognition algorithms. In each figure, vertical axis corresponds to the accuracy of recognition, and horizontal axis corresponds to different strength values of a privacy filter. For each filter, there exist a strength value, termed *critical strength*, that leads to a significant decrease in accuracy of a given recognition algorithm and, consequently, to an increase in privacy protection. We selected 0.5 as a significantly low accuracy value, since only half of the faces are correctly recognized, which is indicated by the dashed horizontal line in the figures. Hence, the critical strength values for each filter-recognition algorithm pair are indicated by vertical dashed lines.

From the figures, we can note that all privacy filters can significantly decrease accuracy of the recognition, however, different filters affect different recognition algorithms differently. For instance, for simple filters (blurring, pixelization, and masking), critical strength values vary drastically across recognition algorithms, and this behavior needs to be taken into account in a practical surveillance system. LBPH demonstrates significantly faster decrease compared to Eigen and Fisher algorithms for blurring and pixelization (see Figures 2 and 3), while for masking filter the situation is reversed. On the other hand, more advanced warping and morphing filters are more stable in the way they affect the accuracy of recognition, as the similarity between curves in each of Figures 5 and 6 and the less varying critical strength values illustrate. Overall, masking seems to be the most suitable among simple filters for privacy protection, because it can reduce recognition accuracy to near zero when using opacity value larger than 0.8. Morphing filter (see Figure 6) demonstrates similarly strong effect on the recognition for intensity strength values larger than 0.8. However, because morphing is reversible (the original face can be recovered securely) and it decreases accuracy of all recognition algorithms in a similar and nearly linear fashion, it is the best choice for privacy protection among all evaluated privacy filters. Warping (see Figure 5) seems to be the least suitable filter for automated video surveillance, because it affects recognition only at very high distortion levels (low number of tiles), whereas it is significantly more complex than simple blurring, pixelization, and masking filters.

Considering performance of recognition algorithm, LBPH stands out from others as it 'reacts' in an unpredictable way to different privacy filters. For instance, from Figure 2 it is clear that LBPH is very sensitive to blurring and quite sensitive to pixelization, which, as a side result, means that LBPH may not perform well on images compressed with JPEG (pixelization artifacts) and JPEG 2000 (blurring artifacts) encoders. This is somewhat surprising since LBPH algorithm is popular among researchers for its high recognition accuracy. LBPH is also unstable to the changing strengths values of masking (see Figure 4) and morphing (see Figure 6) filters. A possible explanation of such instability is that both filters uniformly change the intensity values of all pixels in an image, which, in turn, non-uniformly affects histograms of local pixel patterns (the core of LBPH recognition), resulting in the high variations in recognition accuracy even for small changes of a strength value.

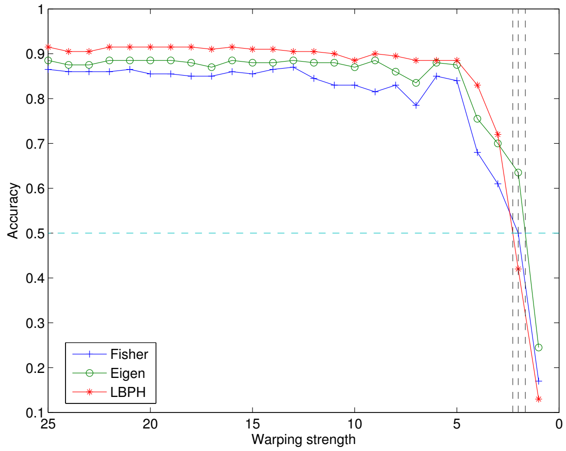

Fig. 5: Recognition results for warping filter

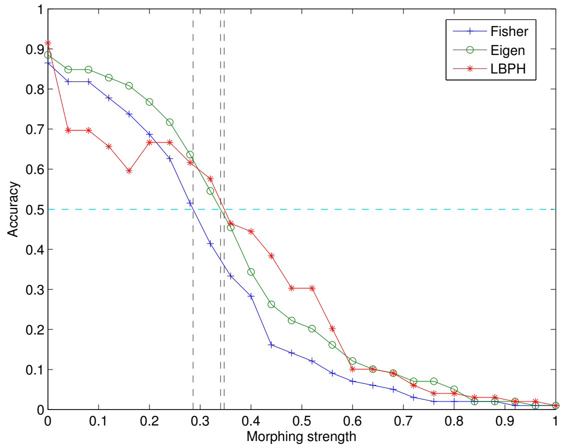

Fig. 6: Recognition results for morphing filter

5. CONCLUSION AND FUTURE WORK

In this paper, by using five different privacy filters and three recognition algorithms, we have demonstrated the importance of careful selection of strength values for each filter-recognition algorithm pair. For the practical systems this findings means that one has to constraint a selected privacy protection filter with a determined strength value to ensure privacy protection. It also means that the value highly depends on a given video analytic, which in case of our evaluations is a recognition algorithm. We also demonstrated that filters affect recognition algorithms differently and morphing filter seems to be the best choice among the evaluated privacy filters.

As a future extension, the subjective evaluations should also be conducted to understand the difference between 'perception' of privacy by video analytics and human visual system. The investigation can also include other surveillance scenarios such as object tracking or event detection.

6. REFERENCES

[1] T. Winkler and B. Rinner, "Trustcam: Security and privacy-protection for an embedded smart camera based on trusted computing," in *Proceedings of IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS)*, Sept. 2010, pp. 593–600.

[2] 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.

[3] J. Wickramasuriya, M. Datt, S. Mehrotra, and N. Venkatasubramanian, "Privacy protecting data collection in media spaces," in *Proceedings of the ACM International Conference on Multimedia (MM)*, New York, NY, USA, Oct. 2004, pp. 48–55.

[4] P. Korshunov and T. Ebrahimi, "Using face morphing to protect privacy," in *Proceedings of IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS)*, Krakow, Poland, Aug. 2013.

[5] P. Korshunov, C. Araimo, F. De Simone, C. Velardo, J. Dugelay, and T. Ebrahimi, "Subjective study of privacy filters in video surveillance," in *2012 IEEE 14th International Workshop on Multimedia Signal Processing (MMSP)*, Sept. 2012, pp. 378–382.

[6] P. Korshunov and T. Ebrahimi, "Using warping for privacy protection in video surveillance," in *18th International Conference on Digital Signal Processing (DSP)*, Santorini, Greece, June 2013.

[7] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1991, pp. 586–591.

[8] P. N. Belhumeur, J. P. Hespanha, and D. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 19, no. 7, pp. 711–720, 1997.

[9] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 12, pp. 2037–2041, 2006.

[10] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 10, pp. 1090–1104, Oct. 2000.

[11] 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, Feb. 2005.

[12] P. Korshunov and W. T. Ooi, "Video quality for face detection, recognition, and tracking," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 7, no. 3, pp. 14:1–14:21, Sept. 2011.

[13] F. Dufaux and T. Ebrahimi, "A framework for the validation of privacy protection solutions in video surveillance," in *Proceedings of IEEE International Conference on Multimedia & Expo (ICME)*, Singapore, July 2010.

[14] P. Korshunov, S. Cai, and T. Ebrahimi, "Crowdsourcing approach for evaluation of privacy filters in video surveillance," in *Proceedings of the ACM Multimedia Workshop on Crowdsourcing for Multimedia (CrowdMM)*, Nara, Japan, Oct. 2012, pp. 35–40.

[15] P. Korshunov, A. Melle, J.-L. Dugelay, and T. Ebrahimi, "A framework for objective evaluation of privacy filters in video surveillance," in *SPIE Applications of Digital Image Processing XXXVI*, San Diego, California, USA, Aug. 2013, vol. 8856.

[16] P. Viola and M. Jones, "Robust real-time face detection," in *Proceedings of the Workshop on Statistical and Computation Theories of Vision (ICCV)*, Vancouver, Canada, July 2001, vol. 2, p. 747.