Anonymization and De-identification of Personal Surveillance Visual Information: A Review

Cesar Pantoja, Virginia Fernandez Arguedas, Ebroul Izquierdo

Multimedia and Vision Research Group
School of Electronic Engineering and Computer Science
Queen Mary, University of London,
Mile End Road, E1 4NS, London, United Kingdom
{cesar.pantoja;virginia.fernandez;ebroul.izquierdo}@eecs.qmul.ac.uk

Keywords: De-identification, anonymization, privacy, surveillance, video filters.

Abstract

The recent widespread adoption of video surveillance systems implies an invasive proactive approach to ensure citizen's security. The ever-increasing amount of recorded information, implies a direct threat to citizen's privacy and their right to preserve their personal information. Thus, a general social concern has raised for the citizen's lost of privacy, demanding new approaches to preserve and protect their privacy, ensuring their anonymity and freedom of action whilst maintaining the surveillance performance. Several approaches have been proposed to preserve sensitive information. In this paper, a review of the existing anonymization and de-identification techniques is presented, categorising them by the domain in which the anonymization is applied and evaluating them with a common framework which takes into account the features and characteristics of each method.

1 Introduction

Recent technological development and the expansion of digital acquisition systems, such as digital cameras or smart phones, have created a digital revolution, where any multimedia content can be accessed, uploaded or downloaded from the internet. The free access to content and the social networking share of content through digital platforms provides enough resources to enhance knowledge acquisition and personal development. However, the lack of privacy-legislation has raised new concerns regarding the privacy of individuals. Revealing a breach between the in-growing digital applications and privacy prevention and protection legislation.

The ubiquitous presence of video cameras in public and private environments enhances the need to address human privacy protection in an effective manner, preserving humans privacy but enabling protection forces to access to the content if needed for an investigation. Thus, different privacy levels and access criteria should be envisioned.

In addressing privacy protection, different techniques have been developed. However, a general tendency proposes to de-identify or anonymize individuals from multimedia content. De-identification or anonymization is a process which aims to remove all identification information of the person from an image or video, while maintaining as much information on the action and its context [1]. Despite hiding the identity is a rather easy task, by replacing the region of interest (ROI), typically the face, with black pixels, such substitution not only hides the identity but also the contextual information of the image, which does not harm the human privacy. Thus, the goal is to protect the privacy of the individuals whilst preserving as much information as possible of the image. Additionally, de-identification techniques must tackle the anonymization of the image providing a natural feeling, so preserving the intelligibility. Finally, privacy protection should be immune to recognition from humans as well as robust enough against computer vision techniques.

In this paper, a survey of the existing anonymization techniques is presented, categorizing them in *Transform-domain* and *Pixel-level*, according to the level where the anonymization is applied. The remainder of the paper is organized as follows. Section 2 is the core of the paper, where the existing anonymization techniques are classified and analyzed. Whilst Section 3 draws some conclusions and states future research lines.

2 Literature Review

De-identification techniques fall into two broad categories, depending on which stage the process is applied. On the one hand, there are techniques that work in the codec, namely *transform-domain* anonymization techniques. These techniques intend to protect regions of interest (ROI) by using the scalability provisions of the used codec. As a result, errors or random information are introduced in the coding process of the ROI, resulting in artifacts and other alterations in the decoding process. On the other hand, some techniques work at a pixel level, namely *pixel-level* anonymization techniques, which deal with the pixels' intensities and chrominances to hide the identity of the detected person.

When assessing the performance of an anonymization method, two factors have to be taken into account [8], (i) the privacy preserving capabilities and (ii) the intelligibility. The former reflects the ability to conceal particular features of the video, allowing the identification of certain personal characteristics, such as gender, race or age. Whilst the latter demonstrate the ability to distinguish the actual events or features of interest. Thus, anonymization presents a trade-off challenge, where higher privacy means less ability to detect events or features of interest. Moreover, Korshunov and Ebrahimi enlisted a set of fundamental characteristics for a practical privacy protection method [10]. During this paper, we shortlisted a set to be used in the methods' evaluation based on their impact on the results, including:

- Reversibility or the ability of the method to recover the original contents of the ROI. Such property is required in viable surveillance systems where certain suspects or events need to be de-anonymized to further inspections/investigations.
- Security enables to reverse the anonymization only when certain conditions are met, i.e. the presence of a secret shared key.
- Variable strength granularity allows modifying certain parameters of the method to yield different degrees of strength and privacy-preservation.

In the following paragraphs, an exhaustive review of the existing anonymization techniques categorized into *transform-domain* and *pixel-level* is presented.

2.1 Transform-domain anonymization techniques

Transform-domain anonymization techniques take advantage of the scalability provisions of the used codec to introduce random information in the coding of the ROIs. As a result, the encoded video is scrambled. Such process can be reversed by transmitting the keys over a secure channel. However, Transform-domain approaches can only be applied to scenarios where a specific codec is used, and new codecs would require the integration of the anonymization technique into the codec.

Dufaux and Ebrahimi presented a method for scrambling regions of interest in transform-coding based codecs (such as the Discrete Cosine Transform - DCT or Discrete Wavelet Transform - DWT) [3]. The proposed approach works by pseudo-randomly flipping the coefficients (inverting the signs) of the ROI during encoding, while the rest of the scene remains constant. The scrambled parts remain understandable enough to recognize the class of the subject while most traces to identify the subject are lost, achieving a balance between privacy and intelligibility. The de-encryption key, which allows the process to be reversed, is securely handed over to the law-enforcement authorities or a trusted third party with legal capabilities enabling the unscrambled video feed. Two implementations of the approach were presented, one for MPEG-4 coded videos (which uses DCT) and the other for Motion JPEG 2000 (based on DWT). In the former, the AC coefficients are flipped pseudo-randomly. While in the latter the quantized wavelet coefficients belonging to the AC sub-bands are flipped with the same process. Variable anonymization strength can be achieved by restricting the scrambling to less AC coefficients in the DCT codecs or limiting to less resolution levels in the DWT codecs. The proposed approach presents a secure technique due to the large flipping possibilities, making impractical any attack to correctly decode the ROI. Dufaux and Ebrahimi's work[3] was extended in [5], where a codestream scrambling was introduced. The same principle applied in the transform-domain was used to the codestream domain. Hence, the bits are flipped pseudo-randomly on the AC or DC coefficients of the DCT based codec (such as MPEG-4). Since the scrambling is performed in the coding of the stream, no extra computational cost is incurred, and bitstreams are standard compliant, showing the scrambled ROI to anybody without the encryption key. In [4], authors further extended their scrambling method in the transform domain by using a pseudo-random permutation of the coefficients, instead of a flip. In all the cases, the impact on coding performance is small, and the computational complexity increase is negligible. Figure 1 shows the scrambling technique for DCT-based codecs (left) and applied to DWT-based codecs (right). This scramble technique affects uniquely to the ROI and offers several degrees of scrambling.

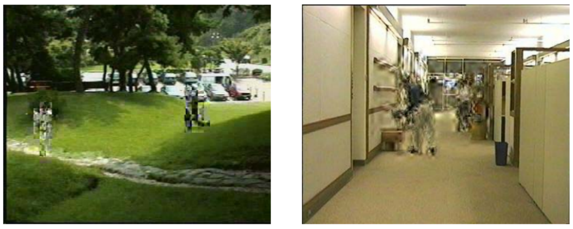


Figure 1. Scrambling of an MPEG-4 ROI (left) and an M-JPEG2000 ROI (right)[5]

In [13], authors took advantage of JPEG XR scalability provisions to produce a stream where ROIs are securely concealed, but format compliant. Only with the de-encryption key the stream can be correctly decoded, achieving security and reversibility. Similarly to the approach in [5], only the luminance channels are scrambled in order to keep the impact on the coding efficiency and computational complexity low. Moreover, authors proposed three different ROI scrambling techniques: pseudo-randomly shifting the level of the DC coefficients, pseudo-randomly permuting the ordering of LP coefficients, and pseudo-randomly inverting the signs of the HP transform coefficients. All of these techniques can be used alone or combined to achieve different levels of concealment and codec efficiency. Thus, ensuring the variable strength granularity and codec efficiency (due to the decrease of the entropy in the contents, the coding efficiency decreases with the bitrate). Figure 2 presents their result of scrambling on an ATM scene. As it can be observed the intelligibility of the scene remains while preserving the privacy of the subjects.

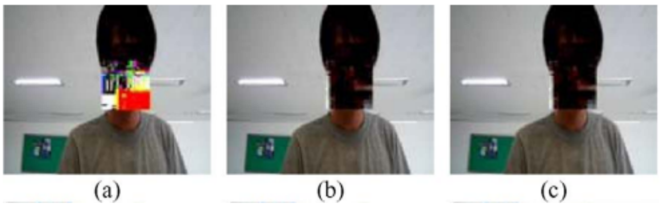


Figure 2. Visual effect of ROI-based, subband-adaptive scrambling[13]. (a) all subbands. (b) DC+LP+HP. (c) DC+LP.

2.2 Pixel-level anonymization techniques

Pixel-level anonymization techniques modify the intensity of the pixels inside the ROI to protect the privacy of the detected person. Two categories can be detected according to the isolated modification of the pixels or the use of the environmental pixels. If the surrounding pixels are taken into account, a maximum intelligibility is obtained. Whilst if the pixels are modified in an independent-basis, the privacy is maximized. Pixel-level anonymization techniques normally have variable strength granularity. However, most of them are irreversible unless the original content of the ROI is stored. In that case, security is compromised and associated to an access-system which deals with different access privilege levels.

An approach to anonymize visual surveillance information by modifying the intensities of the pixels was proposed in [11]. Authors present a novel method to de-identify face images, namely *k*-same, so that face-identification software is ineffective. *K*-same determines the similarity between faces based on a distance metric and creates new faces by averaging image components, which may be the original image pixels (*k*-Same-Pixel) or eigenvectors (*k*-Same-Eigen). As a result, effectively a face is still present, however, the face is no longer the same seen as in the original. Hence, both intelligibility and privacy are ideally high. Variable strength granularity can be achieved by modifying the number of faces included in the averaging, represented by the *k* parameter. This method is based on the *k*-anonymity concept stating that for every piece of anonymized data, there must be *k* pieces of data in the original dataset to which the piece of data could be representative of. For the *k*-Same model, this means that a face image presented to a face recognition system would find *k* matches in the gallery. Although their tests show that face recognition algorithms are thwarted by this approach, the approach does not have the reversibility property, so once the process is applied to a face, it cannot be reverted without storing the original (refer to Figure 3).

The *k*-same concept is further explored in [7] and [6] by Gross *et al.*. In particular, their concept of *utility* is roughly related to what we defined as *intelligibility* or the ability to distinguish particular features of the scene/object that would assist in the actual event identification. So in [7], authors enhance the method by taking into account facial features, such as facial expression or gender, referred as *k*-Same-Select. The insertion of these features integrate their novelty respect the previous method. Thus in [11], a smiling face would be *k*-anonymized with neutral faces, causing also artifacts and ghosting to the resulting face. In [7],the desired features are manually annotated. In [6], authors expanded the method to automatically classify these features by creating Active Appearance Models (AAMs) of the face. The improved method, namely *k*-Same-M, applies *k*-Same to the AAMs instead of the pixel data. Generally, the methods based on *k*-same are very robust in terms of privacy-preserving capabilities and intelligibility, present variable strength granularity (by adjusting the *k* paramter), and produce better quality images. However, the source material is fully modified and lost, preventing its reversion, unless the source image is also stored.

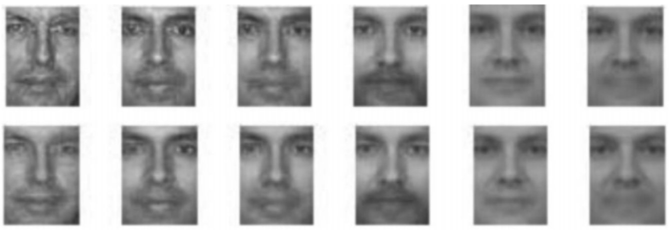


Figure 3. Face images using *k*-Same-Pixel (top) and *k*-Same-Eigen (bottom) with varying degrees of strength by enlarging the *k*-parameter (k = 2, 3, 5, 10, 50, and 100 from left to right) [11]

In [14], authors carrying out a survey of what people perceived as privacy from the viewpoint of the relationship between a viewer and a subject. This survey showed that: (1) A subject's disclosable privacy is positively correlated with subject's closeness to a viewer, (2) A viewer's responsibility expected by a subject is positively correlated with subject's closeness to the viewer, (3) A subject's disclosable privacy is positively correlated with a viewer's responsibility expected by the subject, and (4) A subject's disclosable privacy is individual. Based on these conclusions, authors developed a system and presented two additional anonymization methods. The first proposed method was based on a set of abstraction operations to gradually control visual information. Figure 4 shows the twelve proposed operators. The second proposed method consists of building an AAM of the face and layering a mask on top of it, as a result, the facial expression is preserved (refer to Figure 5). Unfortunately, specific privacy and intelligibility performance evaluation for each abstraction operator and mask faces is not available. Qualitatively, it can be observed that some masked faces have non or very low privacy preserving capabilities (refer to Figure 4 b, c), whilst others present non or very low intelligibility (refer to Figure 4 g, h, i, j, k, l and Figure 5 g, h, i). However, some examples have a different balance of privacy vs. intelligibility (refer to Figure 5 b, c, d). Others are *ad hoc* techniques that have been evaluated before (refer to Figure 4 d, e and Figure 5 e, f). Variable strength granularity was not evaluated. Finally, as most pixel-level methods, the source material is destroyed and additional methods for securely reversing the anonymization should be envisioned and employed.

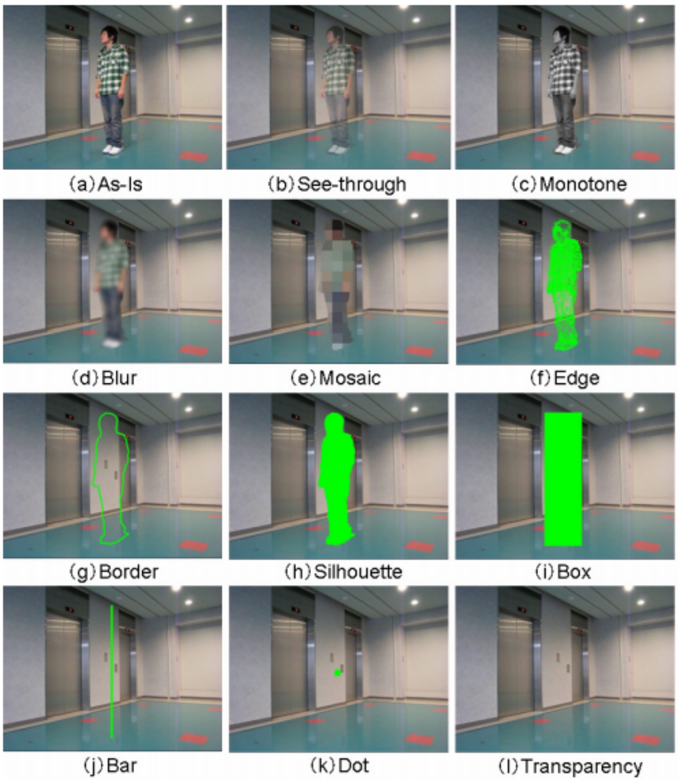


Figure 4. Abstraction operators proposed by [14]

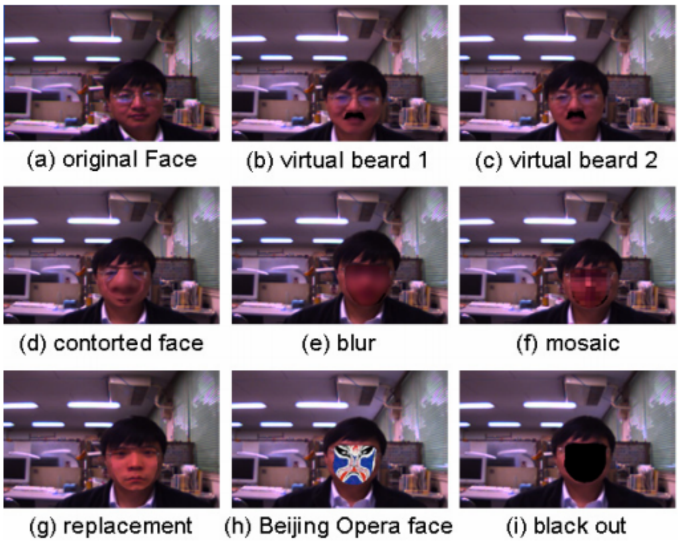


Figure 5. Original face and resulting masked faces using [14]

A reversible anonymization method was presented in [2] by using pixel relocation. The approach relocates the pixels inside the ROI in a specific, known way, such that the process can be reverted by carrying out the inverse relocation. As a result, the ROI is completely unintelligible and no control on the strength of the anonymization is presented. Thus, privacy is completely preserved at the cost of no intelligibility. As stated before, the method is reversible, but its security level is not high due to its dependence to a known and unique relocation operation (the same relocation operation is applied to all the images). Thus, it only takes for an attacker to know the defined relocation operations carried out to reverse the anonymization of any image. Additionally, due to the nature of the approach, only rectangular ROIs can be anonymized. Moreover, since the reversibility depends on the exact location of the pixels, when an image is coded in a lossy format, the pixel values can change, causing a loss of information in the reversion process. A similar reversible approach was presented in [12]. This method encrypts the contents of the ROI with a chaos cryptography approach, generating similar visual results as in [2]. These two methods preserve the privacy, but share similar drawbacks: no intelligibility, no variable strength granularity and possible loss of information in the reversal process if an image is encoded in a lossy format. However, [12] presents the advantage of being secure, due to its encryption technique, preventing attackers to decode the image.

Another reversible pixel based anonymization method was presented in [10], where ROIs are "warped" using a set of key points. Warping algorithm consists on the following steps (with unwarping being the same algorithm inversely applied to the warped image):

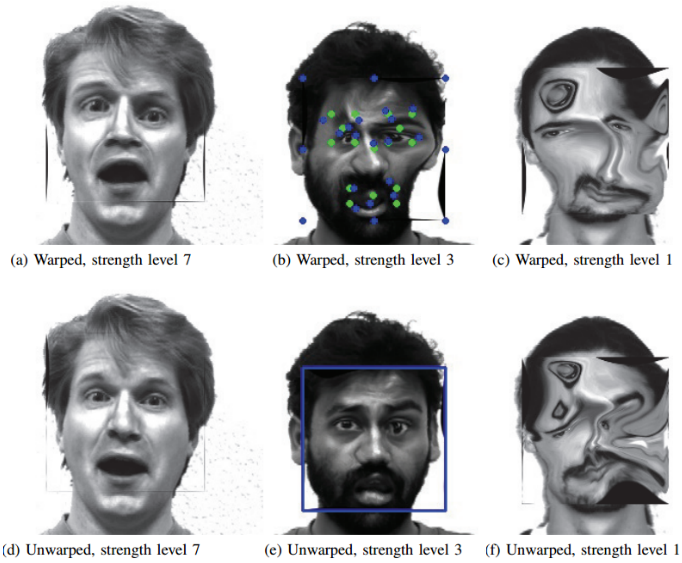
1. Select a set of key points in the image
2. Randomly shift these points (i.e., change their coordinates) by adding or subtracting random value with weight depending on the warping strength.
3. The resulted coordinates constitute the desired destinations for the selected point in the target warped image.
4. Based on the original and destination coordinates of the key point, compute transformation matrix.
5. Apply the transformation to each pixel in the image, using "cubic" interpolation.

Face detection algorithms accuracy for the unwarped/warped images with low warp applied is the same as for the original images. However, when warping is to high, the face detection performs poorly compared to the original images. As stated before, this method is reversible and also secure, because the key points are randomly generated and unwarping can only be applied if these points are known. The method presents also variable strength granularity, as the amount of warping can be modified to yield different levels of anonymization. The balance between privacy protection and intelligibility is not explored, but qualitatively it can be observed that some of the features can still be recognized in the warped images. However, due to the loss of some original values of the pixels, replaced by interpolations of the new pixels, the unwarping performance is affected, resulting in the reconstructed image to be an approximation of the original (The higher the warping level, the higher the error introduced in the unwarping process) (refer to Figure 6).

Ad hoc methods for anonymization of surveillance visual information, like blurring, pixelation, and masking, were evaluated by Korshunov *et al.* in [8] and [9] using subjective and

	Method	Ref.	Reversibility	Security	Variable strength granularity	Privacy	Intelligibility
Transform-domain	AC pseudo-random coefficients flipping	[3]	Y	High	Yes	High	High
	AC sub-bands pseudo-random coefficient flipping	[3]	Y	High	Yes	High	High
	Codestream coefficient flipping	[5]	Y	High	Yes	High	High
	Coefficient permutation	[4]	Y	High	Yes	High	High
	Subband-adaptive acrambling	[13]	Y	High	Yes	High	High
Pixel-level	Blurring	[8][9]	N	N/A	Yes	Low	High
	Pixelization	[8][9]	N	N/A	Yes	High	High
	Masking	[8][9]	N	N/A	No	High	Low
	k-same	[11]	N	N/A	Yes	High	High
	k-same-select	[7]	N	N/A	Yes	High	High
	k-same-M	[6]	N	N/A	Yes	High	High
	Abstraction operators	[14]	N	N/A	N/A	N/A	N/A
	Mask faces	[14]	N	N/A	N/A	N/A	N/A
	Pixel relocation	[2]	Y	Low	No	High	Low
	Chaos cryptography	[12]	Y	High	No	High	Low
	Warping	[10]	Y	High	Yes	N/A	N/A

Table 1. Methods comparison


(a) Warped, strength level 7 (b) Warped, strength level 3 (c) Warped, strength level 1

(d) Unwarped, strength level 7 (e) Unwarped, strength level 3 (f) Unwarped, strength level 1

Figure 6. Examples of warping and unwarping for different faces and different levels of warping strength. Green dots in Figure (b) are original locations of key pixel grid, and blue dots are these pixels randomly shifted[10]

crowdsourcing methods, respectively. In particular, their objective was to evaluate the effectiveness of several methods regarding their privacy preserving capabilities and intelligibility. To evaluate the methods, 9 video sequences were used and different indoor video surveillance scenarios were considered. The scenarios were classified in *normal* (i.e. such as a person walking towards and away from the camera) or *suspicious*, (i.e. blinking into the camera or wearing a scarf around the mouth to hide personal identity). Subjects or participants (in the crowdsourcing evaluation) were asked to watch a video sequence and then answer several questions related to the achieved privacy or intelligibility. The privacy-related questions included “What is the gender of the person?”, “What is the race of the person?” and “Does the person wear glasses?”. While the intelligibility-related questions included “Does the person wear sunglasses?”, “Does the person wear a scarf?” and “Does the person blink into the camera?”. Their results show that (i) blurring filter yields the highest intelligibility while providing the lowest privacy protection, (ii) masking filter shows the highest privacy protection, having the lowest intelligibility and (iii) pixelization filter demonstrates high privacy protection while still yielding high degree of the activities recognition (refer to Figure 7). Thus, pixelization filter presents the best balance of privacy and intelligibility. Additionally, the strength of these anonymization techniques could be adjusted to achieve granularity. However, the main drawback is that Pixel-level methods destroy the original content to ensure privacy. So separate methods to securely reverse the anonymization must be devised. Figure 7 shows examples of these techniques.

3 Conclusions

In this paper, an exhaustive review of the existing anonymization and de-identification techniques was presented. The review categorized such techniques according to the level where the privacy is applied, into *Transform-domain* and *Pixel-level*. Qualitative evaluation demonstrated that the inclusion of contextual information offers the best balance between privacy and intelligibility (i.e. *k*-same based approaches). Whilst

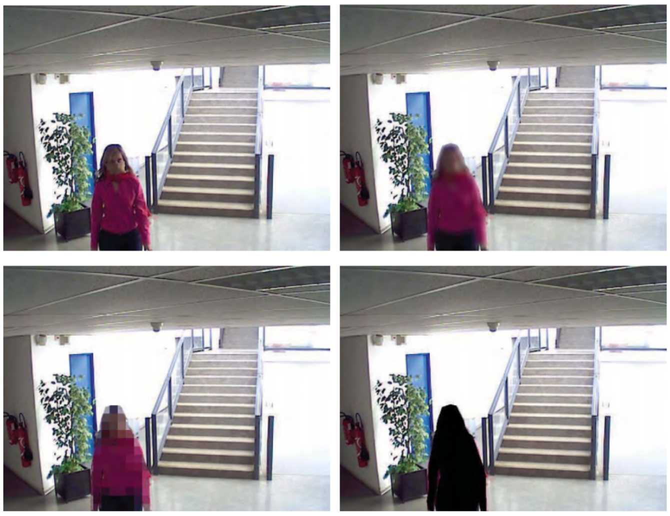


Figure 7. Examples of the different pixel-level anonymization techniques evaluated in [8]. From left to right, top to bottom: original frame, blurring, pixelization, masking

Transform-domain approaches offer an overall better balance between the features and privacy/intelligibility. A summary of the comparison is presented in table 1. To sum up, an ideal solution would offer a trade-off between the importance of intelligibility and privacy. As future work, we intend to expand the study and to propose a methodological evaluation framework enabling quantitative performance evaluation. Additionally, a pixel-level anonymization technique addressing the enhancement of intelligibility in privacy preserving applications is under development.

Acknowledgements

The research presented in this paper was supported by the European Commission under contract FP7-SEC 261743VideoSense and FP7-SEC 285024 Advise.

References

[1] P. Agrawal and P. J. Narayanan. Person de-identification in videos. *IEEE Transactions on Circuits and Systems for Video Technology*, 21(3):299–310, 2011.

[2] J. Cichowski and A. Czyzewski. Reversible video stream anonymization for video surveillance systems based on pixels relocation and watermarking. In *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, pages 1971–1977, 2011.

[3] F. Dufaux and T. Ebrahimi. Scrambling for video surveillance with privacy. In *Conference on Computer Vision and Pattern Recognition Workshop, 2006. CVPRW '06.*, pages 160–160, 2006.

[4] F. Dufaux and T. Ebrahimi. H.264/avc video scrambling for privacy protection. In *15th IEEE International Conference on Image Processing, 2008. ICIP 2008.*, pages 1688–1691, 2008.

[5] F. Dufaux and T. Ebrahimi. Scrambling for privacy protection in video surveillance systems. *IEEE Transactions on Circuits and Systems for Video Technology*, 18(8):1168–1174, 2008.

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

[7] Ralph Gross, Edoardo Airoldi, Bradley Malin, and Latanya Sweeney. Integrating utility into face de-identification. In *5th international conference on Privacy Enhancing Technologies*, PET'05, pages 227–242, Berlin, Heidelberg, 2006. Springer-Verlag.

[8] 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)*, pages 378–382, 2012.

[9] Pavel Korshunov, Shuting Cai, and Touradj Ebrahimi. Crowdsourcing approach for evaluation of privacy filters in video surveillance. In *ACM multimedia 2012 workshop on Crowdsourcing for multimedia*, CrowdMM '12, pages 35–40, New York, NY, USA, 2012. ACM.

[10] Pavel Korshunov and Touradj Ebrahimi. Using warping for privacy protection in video surveillance. In *18th International Conference on Digital Signal Processing. DSP2013.*, 2013.

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

[12] Sk.Md.Mizanur Rahman, M.Anwar Hossain, Hussein Mouftah, Abdulmotaleb El Saddik, and Eiji Okamoto. Chaos-cryptography based privacy preservation technique for video surveillance. *Multimedia Systems*, 18(2):145–155, 2012.

[13] Hosik Sohn, W. De Neve, and Yong-Man Ro. Privacy protection in video surveillance systems: Analysis of subband-adaptive scrambling in jpeg xr. *IEEE Transactions on Circuits and Systems for Video Technology*, 21(2):170–177, 2011.

[14] Xiaoyi Yu, K. Chinomi, T. Koshimizu, N. Nitta, Y. Ito, and N. Babaguchi. Privacy protecting visual processing for secure video surveillance. In *15th IEEE International Conference on Image Processing, 2008. ICIP 2008.*, pages 1672–1675, 2008.