An annotation-free method for evaluating privacy protection techniques in videos

An annotation-free method for evaluating privacy protection techniques in videos

Tahir Nawaz and James Ferryman
Computational Vision Group, School of Systems Engineering, University of Reading, UK
{t.h.nawaz,j.m.ferryman}@reading.ac.uk

Abstract

While several privacy protection techniques are presented in the literature, they are not complemented with an established objective evaluation method for their assessment and comparison. This paper proposes an annotation-free evaluation method that assesses the two key aspects of privacy protection that are privacy and utility. Unlike some existing methods, the proposed method does not rely on the use of subjective judgements and does not assume a specific target type in the image data. The privacy aspect is quantified as an appearance similarity and the utility aspect is measured as a structural similarity between the original raw image data and the privacy-protected image data. We performed an extensive experimentation using six challenging datasets (including two new ones) to demonstrate the effectiveness of the evaluation method by providing a performance comparison of four state-of-the-art privacy protection techniques.

1. Introduction

With the increasing use of the surveillance applications in public places [11], the need of protecting the privacy of individuals is also growing [8, 25]. Privacy protection may involve hiding or masking out image regions that would otherwise reveal object identity. Several privacy protection techniques have been presented in the literature [5, 10, 12, 16, 20]. These techniques essentially apply different image filtering operations to provide a different level of identity protection [25].

While evaluation criteria exist for assessing methods in other areas of computer vision including optical flow estimation [6], stereo correspondence estimation [22] and video tracking [15, 17], there is an absence of an established method for the performance evaluation of different aspects of privacy protection methods. Some works exist that used subjective methods for evaluating privacy protection techniques [7, 13, 21]. Boyle *et al.* [7] presented a methodology that involved applying global (full-frame) privacy protection on a set of video sequences and showing them to subjects, and in turn assessing the privacy protection techniques based on the collected subjects' responses using questionnaires. Saini *et al.* [21] and Korshunov *et al.* [13] also used a similar subjective methodology except that they applied privacy protection locally (only on sensitive image regions) in video sequences. The performance evaluation based on the above methods rely on subjective judgements and hence could lack objectivity. An evaluation framework was proposed that did not rely on subjective judgement and used the face detection and face recognition accuracies on the privacy-protected data as measures of the privacy protection [14]. While an interesting contribution, this framework is target-dependent (i.e. aimed at image data with face targets) and depends also on the performance of detection and recognition algorithms used.

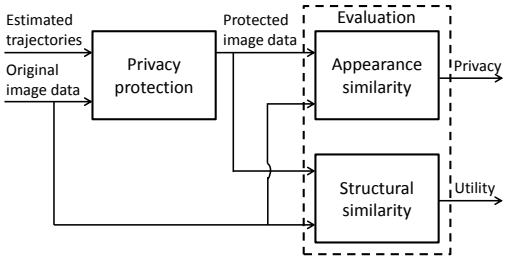


Figure 1. Proposed method for objectively evaluating a privacy protection technique in the context of video tracking by quantifying the *privacy* and *utility* aspects.

To objectively[1] evaluate a privacy protection method the two key aspects to consider are as follows [10, 21]: first, a quantification of the extent of identity information hidden by it that is the *privacy*; second, a quantification of the preservation of the behavioral and structural information that is the *utility*. An ideal privacy protection technique may aim to maximize the privacy as well as the utility.

This paper presents an evaluation method (Fig. 1) for ob-

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 312784.

[1]Here the term 'objective' means the non-reliance of an evaluation method on subjective judgements.

jectively assessing the privacy and utility aspects of the privacy protection techniques. The evaluation method is target independent and annotation free. Privacy is measured by quantifying the appearance similarity between the original and privacy-protected image data. Utility is measured by quantifying the structural similarity between the original and privacy-protected image data. We demonstrate the effectiveness of the proposed evaluation criteria with an extensive experimentation by assessing and comparing four privacy protection techniques in the context of video tracking on six challenging datasets (including two new ones).

2. Problem definition

Consider a video sequence V consisting of K frames: $V = (f_k)_{k=1}^{K}$, where f_k denotes the frame k. Let $\mathscr{X}$ be a set of trajectories (or tracks) estimated by a tracker in V: $\mathscr{X} = \{\mathfrak{X}_j\}_{j=1}^{J}$, where J is the total number of estimated trajectories. $\mathfrak{X}_j$ is the estimated trajectory for target j: $\mathfrak{X}_j = (X_{k,j})_{k=k^j_{start}}^{k^j_{end}}$, where k^j_{start} and k^j_{end} are the first and final frame numbers of $\mathfrak{X}_j$, respectively. $X_{k,j} = (x_{k,j}, y_{k,j}, A_{k,j}, l_j)$, where $(x_{k,j}, y_{k,j})$ and $A_{k,j}$ denote the position and the occupied area information of target j on the image plane and l_j defines its ID. Without the loss of generality $A_{k,j}$ is considered in the form of a bounding box in which case $X_{k,j}$ can be re-written as: $X_{k,j} = (x_{k,j}, y_{k,j}, w_{k,j}, h_{k,j}, l_j)$, where $w_{k,j}$ and $h_{k,j}$ denote the width and height of the bounding box for target j at f_k. The number of estimated targets at f_k is denoted as n_k, which are defined as $\{X_{k,1}, \ldots, X_{k,j}, \ldots, X_{k,n_k}\}$. Let $B_{k,j}$ denotes the image data within the bounding box $X_{k,j}$. B_k is the set containing the image data within all the bounding boxes in f_k: $\mathrm{B}_k = \{B_{k,1}, \ldots, B_{k,j}, \ldots, B_{k,n_k}\}$. Let $B'_{k,j}$ denotes the privacy-protected image data obtained by applying a privacy protection method on $B_{k,j}$. Therefore, B'_k is the set containing the privacy-protected image data within all the bounding boxes in f_k: $\mathrm{B}'_k = \{B'_{k,1}, \ldots, B'_{k,j}, \ldots, B'_{k,n_k}\}$. The evaluation procedure compares B'_k with respect to B_k, the original unprotected data, to assess the privacy protection method under consideration in the form of a score, S_k, at f_k.

3. Evaluation method

The proposed evaluation method is aimed to assess the two key aspects of privacy protection that are privacy and utility. Unlike [7, 13, 21] the method does not rely on subjective judgement. Additionally, unike [14] the method does not require the application of privacy protection methods on an image data with a particular target type. Moreover, the method is annotation free.

Privacy is assessed in terms of the appearance similarity between B_k and B'_k. A smaller appearance similarity between B_k and B'_k alludes to a greater impact of the applied privacy protection. For quantifying the appearance similarity we use the widely-used Bhattacharyya distance that is a metric (unlike the Kullback-Leibler divergence that is non-symmetric and hence not a metric) and does not assume the same variance for $B_{k,j}$ and $B'_{k,j}$ (which Mahalanobis distance does). At frame k we compute the amount of achieved privacy, P_k, in B'_k as follows:

$$P_k = \frac{1}{n_k} \sum_{j=1}^{n_k} D_{k,j}(q^{B_{k,j}}, q^{B'_{k,j}}), \tag{1}$$

where $D_{k,j}(q^{B_{k,j}}, q^{B'_{k,j}})$ is the Bhattacharyya distance at frame k between the probability distributions (normalized histograms) of $B_{k,j}$, $q^{B_{k,j}}$, and $B'_{k,j}$, $q^{B'_{k,j}}$. $D_{k,j}(q^{B_{k,j}}, q^{B'_{k,j}}) = \sqrt{1 - BC_{k,j}(q^{B_{k,j}}, q^{B'_{k,j}})}$: where the Bhattacharyya coefficient $BC_{k,j}(q^{B_{k,j}}, q^{B'_{k,j}}) = \sum_{z=0}^{Z} q^{B_{k,j}}(z) q^{B'_{k,j}}(z)$. $Z = 255$ as we use 255 bins (which is equal to the number of intensity levels) in computing the normalized histograms for $B_{k,j}$ and $B'_{k,j}$. In the case of RGB image, $D_{k,j}(\cdot) = \sqrt{D^2_{k,j,red}(\cdot) + D^2_{k,j,green}(\cdot) + D^2_{k,j,blue}(\cdot)}$: $D^2_{k,j,red}(\cdot)$, $D^2_{k,j,green}(\cdot)$ and $D^2_{k,j,blue}(\cdot)$ are the Bhattacharyya distances between the corresponding probability distributions of the red, green and blue channels of $B_{k,j}$ and $B'_{k,j}$. $P_k \geq 0$: the higher P_k the greater the amount of achieved privacy. While the computation of P_k enables analyzing the achieved privacy at each frame, to facilitate the performance comparison between different privacy protection methods we provide the overall achieved privacy, P, in the form of a single score as follows:

$$\mathrm{P} = \frac{1}{K} \sum_{k=1}^{K} P_k. \tag{2}$$

Utility is quantified in terms of the structural similarity between B_k and B'_k. A smaller structural similarity refers to a lower preservation of structural information. For computing the structural similarity between $B_{k,j}$ and $B'_{k,j}$ we use the well-known Structural Similarity Index (SSIM) [23] that was also employed in [9, 10]. At frame k the utility, U_k, is computed as follows:

$$U_k = \frac{1}{n_k} \sum_{j=1}^{n_k} \mathrm{MSSIM}_{k,j}(B_{k,j}, B'_{k,j}), \tag{3}$$

where $\mathrm{MSSIM}_{k,j}(B_{k,j}, B'_{k,j})$ is the mean SSIM value between $B_{k,j}$ and $B'_{k,j}$ for a variation of local windows [23]. $\mathrm{MSSIM}_{k,j}(B_{k,j}, B'_{k,j}) = \frac{1}{M} \sum_{m=1}^{M} \mathrm{SSIM}^m_{k,j}(B_{k,j}, B'_{k,j})$, where $\mathrm{SSIM}^m_{k,j}(B_{k,j}, B'_{k,j})$ is the SSIM value for mth window and is given as follows [23]: $\mathrm{SSIM}^m_{k,j}(B_{k,j}, B'_{k,j}) = \frac{(2\,{}^m\mu_{B_{k,j}}\,{}^m\mu_{B'_{k,j}} + C_1)(2\,{}^m\sigma_{B_{k,j}B'_{k,j}} + C_2)}{({}^m\mu^2_{B_{k,j}} + {}^m\mu^2_{B'_{k,j}} + C_1)({}^m\sigma^2_{B_{k,j}} + {}^m\sigma^2_{B'_{k,j}} + C_1)}$.

Table 1. Summary of the datasets. Key. K: number of frames; Occ: occlusion; SC: scale changes; IC: illumination changes; Cr: crowdedness; PC: pose changes.

Dataset	K	Frame size	Target type	Challenges
ETH Bahnhof	999	480×640	Person	Occ, SC, IC, Cr
ETH Sunnyday	354	480×640	Person	Occ, SC, IC, Cr
iLids Easy	5220	576×720	Person	Occ, SC, IC
PETS 2000	160	576×768	Vehicle	SC, PC
OKG	243	960×1280	Vehicle	SC, PC
CAST	150	960×1280	Vehicle	SC, PC

SSIM is computed on grey-scale data [23] such that ${}^{m}\mu_{B_{k,j}}$ and ${}^{m}\mu_{B'_{k,j}}$ are the mean intensity values, and ${}^{m}\sigma_{B_{k,j}}$ and ${}^{m}\sigma_{B'_{k,j}}$ are the standard deviations in $B_{k,j}$ and $B'_{k,j}$, respectively, for the local window m; ${}^{m}\sigma_{B_{k,j}B'_{k,j}}$ is the correlation coefficient; and C_1 and C_2 are the constants. $U_k \in [0,1]$: the higher U_k the larger the utility retained. For the reason described in the computation of P (Eq. 2) we compute the overall retained utility, U, in the form of a single score as follows:

$$\mathrm{U} = \frac{1}{K}\sum_{k=1}^{K} U_k. \tag{4}$$

4. Experimental validation and analysis

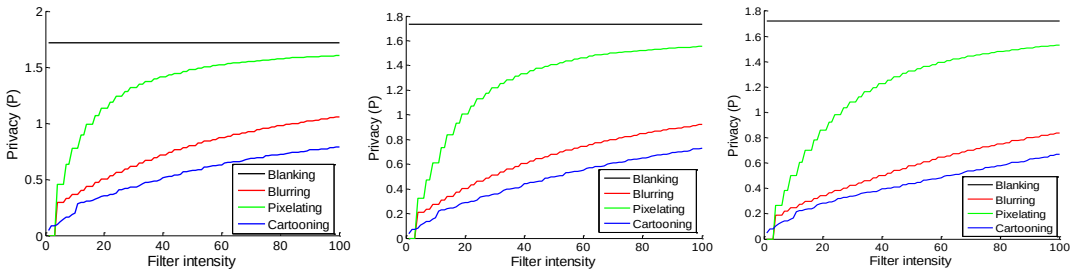
We use six challenging datasets in experiments (Table 1). Among the datasets, four are well known and publicly available including ETH Bahnhof [1], ETH Sunnyday [1], iLids Easy [3] and PETS 2000 [4]. The other two are new and recorded outside the OKG nuclear power plant site in Sweden and in a Centre for Applied Science and Technology (CAST) site in UK under the EU project: P5 [2]. The datasets contain the *full-person body* and the *vehicle* as target types. ETH Bahnhof, ETH Sunnyday and iLids Easy contain multiple targets and we use tracking results from a multi-target tracker [19] in these scenes. PETS 2000, OKG and CAST contain a single target and we generated tracking results using a single-target tracker [18] in these scenes.

We demonstrate the effectiveness of the proposed evaluation criteria by evaluating and comparing four well-known privacy protection methods including cartooning, blurring, pixelating and blanking. Cartooning involves applying an initial blurring on the input image data followed by mean-shift filtering and edge recovery using the already generated gradient mask with sobel edge detector [10]. The kernel size at the initial blurring stage (A) and the spatial radius (sp) and color radius (sr) at the mean-shift filtering stage are given as follows [10]: $A_i = [i \cdot A_{orig}/50]$; $sp_i = [i \cdot sp_{orig}/50]$; $sr_i = [i \cdot sr_{orig}/50]$; where i is the filter intensity: $i \in [1,100]$ and the parameters $A_{orig} = 7$, $sp_{orig} = 20$ and $sr_{orig} = 40$ [10]. Additionally, as done in [10], for establishing some correspondence and a fair comparison among different techniques the kernel size used in the case of blurring and pixelating for a particular filter intensity, i, is equal to sp_i as defined above for cartooning. Blanking completely masks out the privacy-sensitive information in the image data. We therefore apply each privacy protection technique on the tracking results in all datasets for a full variation of filter intensity, i; of course blanking remains unaffected over a variation of i. Fig. 2 and Fig. 3 plot the privacy (P) and utility (U) scores, respectively, of the four privacy protection techniques for a variation of i on all datasets. Fig. 4 shows some sample qualitative results for the privacy protection techniques on all datasets with an increasing i.

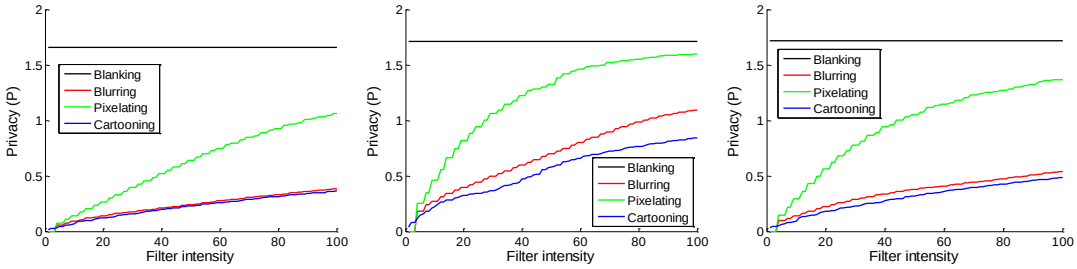
The trends of the privacy scores (P) obtained by the four techniques over a variation of i are similar across all datasets (Fig. 2). Expectedly, on all datasets blanking provides the highest privacy (highest P) as it masks out the entire information in the image data (Fig. 4). Among the remaining methods, pixelating consistently obtain the highest P followed by blurring and cartooning on all datasets for the entire variation of i as shown in Fig. 2 (note that for $i \in [1,3]$, $\mathrm{P} = 0$ for blurring and pixelating because according to the equation of sp_i their kernel size is 1×1 thus leaving the image data unaltered by these two methods). Interestingly, the ranking trend of pixelating, blurring and cartooning techniques in terms of their privacy scores over a variation of i is similar to that reported in [10] (the authors of [10] used however a different dataset to show the results).

As in the case of P scores the trends of the utility scores (U) obtained by the four methods over a variation of i are also alike across all datasets (Fig. 3). Blanking, which consistently achieves the highest privacy, provides the least utility (smallest U) due to a total loss of structural information. Cartooning, on the other hand, preserves the structural information better than the rest of the methods (Fig. 4). Indeed, on all datasets cartooning shows the highest U for the entire variation of i followed by blurring and pixelating, which corresponds to the conclusions made in [10].

Indeed, the aim for a privacy protection technique would be to provide an appropriate trade off between U and P. To this end in Fig. 5 we also plot U (as computed in Fig. 3) *vs.* P (as computed in Fig. 2) on all datasets. U *vs.* P plot would be desirable for choosing among different privacy protection techniques for a specific application. For example, for a specific application on ETH Bahnhof, ETH Sunnyday and iLids Easy (Fig. 5(a-c)), for a desired $\mathrm{P} = 0.60$ pixelating would be the best choice as it provides the highest U. Likewise, on OKG, CAST and PETS 2000 (Fig. 5(d-f)), for a desired $\mathrm{P} = 0.26$ cartooning would be the best choice due to the highest U. In general pixelating is found to provide a better trade off between U and P on datasets with *person* target (Fig. 5(a-c)), and cartooning provides a better trade off on datasets with *vehicle* target (Fig. 5(d-f)). We also checked the statistical significance of the P and U scores obtained by the four privacy protection methods using the



(a) ETH Bahnhof (b) ETH Sunnyday (c) iLids Easy



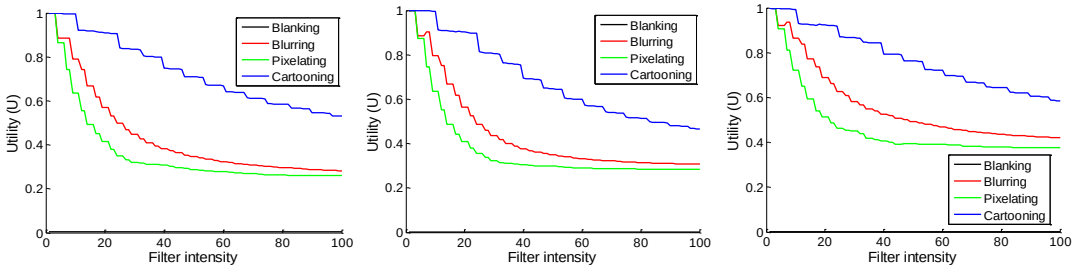
(d) OKG (e) CAST (f) PETS 2000

Figure 2. Privacy score (P) obtained by different privacy protection techniques for a variation of filter intensity on all datasets.

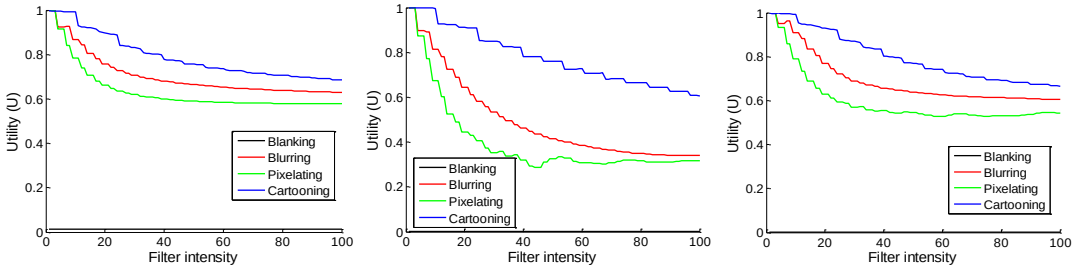
Welch ANOVA test [24]. Statistical significance is achieved at the standard 5% significance level both for the case of P and U scores on each dataset.

5. Conclusions

We presented an annotation-free and target-independent evaluation method for objectively assessing privacy protection techniques. The evaluation method assesses *privacy* by measuring the Bhattacharyya distance-based appearance similarity and *utility* by quantifying the SSIM-based structural similarity between the original image data and the privacy-protected image data. Through an extensive experimentation on six datasets (including two new ones) we showed the usefulness of the proposed evaluation method by providing a statistically-significant comparison of four privacy protection techniques: blanking, blurring, pixelat-



(a) ETH Bahnhof (b) ETH Sunnyday (c) iLids Easy



(d) OKG (e) CAST (f) PETS 2000

Figure 3. Utility score (U) obtained by different privacy protection techniques for a variation of filter intensity on all datasets.

(a) ETH Bahnhof dataset [Filter intensity: $i = 20$]

(b) ETH Sunnyday dataset [Filter intensity: $i = 30$]

(c) iLids Easy dataset [Filter intensity: $i = 40$]

(d) OKG dataset [Filter intensity: $i = 50$]

(e) CAST dataset [Filter intensity: $i = 60$]

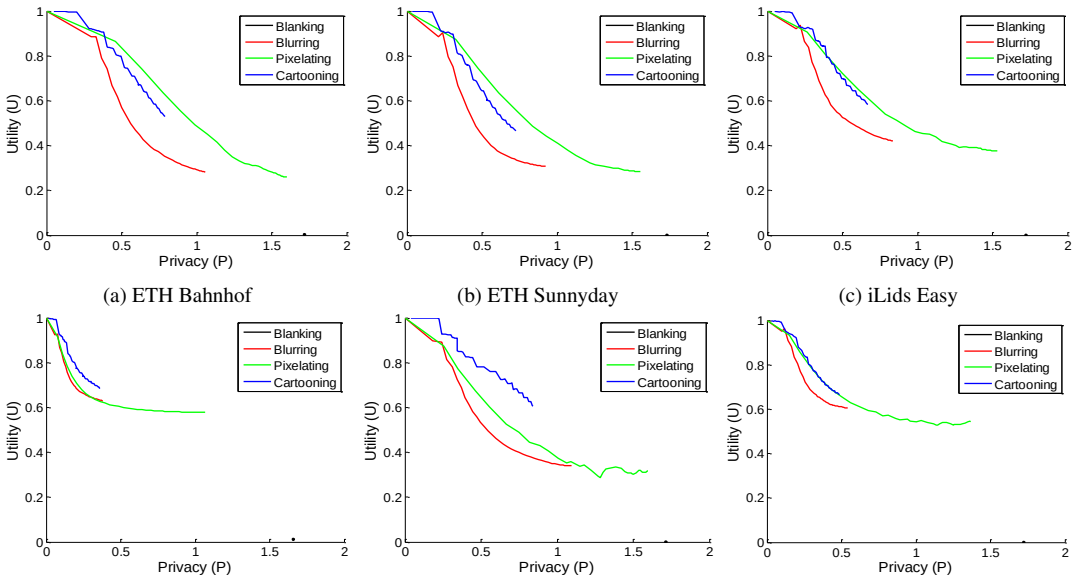
(f) PETS 2000 dataset [Filter intensity: $i = 70$]

Figure 4. Sample qualitative results for different privacy protection techniques on all datasets with an increasing filter intensity ($i = 20, 30, 40, 50, 60, 70$). Column 1: original frame; column 2: blanking; column 3: blurring; column 4: pixelating; column 5: cartooning.

ing and cartooning. Blanking is not desirable in general because it results in a complete loss of visual (and hence the structural) information thus providing a very low utility. Among the remaining techniques, pixelating achieves a higher privacy and cartooning provides a higher utility over a variation of filter intensity. Moreover, pixelating and cartooning are generally found to provide a better trade off between utility and privacy on datasets with person and vehicle targets, respectively. Finally, in this study we evaluated the privacy protection techniques by applying them locally only on the estimated target bounding boxes. The proposed method is indeed generic and could also be used for evaluating a globally-applied privacy protection.

References

[1] ETH Bahnhof and Sunnyday Datasets. http://www.vision.ee.ethz.ch/~aess/iccv2007/. Accessed March 2015.



(a) ETH Bahnhof (b) ETH Sunnyday (c) iLids Easy

(d) OKG (e) CAST (f) PETS 2000

Figure 5. Utility score (U) plotted *vs.* privacy score (P) obtained by different privacy protection techniques for a variation of filter intensity on all datasets.

[2] EU project P5. http://www.foi.se/p5. Accessed June 2015.

[3] http://www.eecs.qmul.ac.uk/~andrea/avss2007_d.html. Accessed March 2015.

[4] PETS 2000 dataset. ftp://ftp.cs.rdg.ac.uk/pub/PETS2000/. Accessed March 2015.

[5] A. J. Aved and K. A. Hua. A general framework for managing and processing live video data with privacy protection. *Multimedia systems*, 18(2):123–143, 2012.

[6] S. Baker, D. Scharstein, J. Lewis, S. Roth, M. J. Black, and R. Szeliski. A database and evaluation methodology for optical flow. *IJCV*, 92(1):1–31, 2011.

[7] M. Boyle, C. Edwards, and S. Greenberg. The effects of filtered video on awareness and privacy. In *Proc. of CSCW*, 2000.

[8] A. Cavallaro. Privacy in video surveillance. *IEEE SPM*, 24(2):165–166, 2007.

[9] F. Defaux. Video scrambling for privacy protection in video surveillance: recent results and validation framework. In *Proc. of SPIE*, 2011.

[10] A. Erdelyi, T. Barat, P. Valet, T. Winkler, and B. Rinner. Adaptive cartooning for privacy protection in camera networks. In *Proc. of IEEE AVSS*, 2014.

[11] S. Fleck and W. Strasser. Smart camera based monitoring system and its application to assisted living. *Proceedings of IEEE*, 96(10):1698–1714, 2008.

[12] B.-J. Han, H. Jeong, and Y.-J. Won. The privacy protection framework for biometric information in network based cctv environment. In *Proc. of ICOS*, 2011.

[13] P. Korshunov, C. Araimo, F. D. Simone, C. Velardo, J.-L. Dugelay, and T. Ebrahimi. Subjective study of privacy filters in video surveillance. In *Proc. of IEEE Work. MMSP*, 2012.

[14] P. Korshunov, A. Melle, J.-L. Dugelay, and T. Ebrahimi. Framework for objective evaluation of privacy filters. In *Proc. of SPIE*, 2013.

[15] M. Kristan, R. Pflugfelder, A. Leonardis, J. Matas, F. Porikli, L. Cehovin, G. Nebehay, G. Fernandez, and T. Vojir. The vot2013 challenge: overview and additional results. In *Proc. of CVWW*, 2014.

[16] A. Martinez-Balleste, H. A. Rashwan, D. Puig, and A. P. Fullana. Towards a trustworthy privacy in pervasive video surveillance systems. In *Proc. of IEEE PerCom*, 2012.

[17] T. Nawaz, F. Poiesi, and A. Cavallaro. Measures of effective video tracking. *IEEE TIP*, 23(1):376–388, 2014.

[18] J. Ning, L. Zhang, D. Zhang, and C. Wu. Robust mean-shift tracking with corrected background-weighted histogram. *IET Comp. Vis.*, 6(1):62–69, 2012.

[19] H. Pirsiavash, D. Ramanan, and C. C. Fowlkes. Globally-optimal greedy algorithms for tracking a variable number of objects. In *Proc. of IEEE CVPR*, 2011.

[20] F. Z. Qureshi. Object-video streams for preserving privacy in video surveillance. In *IEEE AVSS*, 2009.

[21] M. Saini, P. Atrey, S. Mehrotra, and M. Kankanhalli. Anonymous surveillance. In *Proc. of IEEE ICME*, 2011.

[22] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *IJCV*, 47(1/2/3):7–42, 2002.

[23] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE TIP*, 13(4):600–612, 2004.

[24] B. L. Welch. On the comparison of several mean values: An alternative approach. *Biomet.*, 38(3-4):330–336, 1951.

[25] T. Winkler and B. Rinner. Security and privacy protection in visual sensor networks: A survey. *ACM Comp. Surv.*, 47(1), 2014.