Design Space Exploration for Adaptive Privacy Protection in Airborne Images

Omair Sarwar[1,2], Bernhard Rinner[1], Andrea Cavallaro[2]
[1]Institute of Networked and Embedded Systems, Alpen-Adria-Universität Klagenfurt, Austria
[2]School of Electrical Engineering, Queen Mary University of London, UK
omair.sarwar@aau.at, bernhard.rinner@aau.at, a.cavallaro@qmul.ac.uk

Abstract

Airborne cameras on low-flying unmanned vehicles introduce new privacy challenges due to their mobility and viewing angles. In this paper, we focus on face recognition from airborne cameras and explore the design space to determine when a face in an airborne image is inherently protected, that is when an individual is not recognizable. Moreover, when individuals are recognizable by facial recognition algorithms, we propose an adaptive filtering mechanism to lower the face resolution in order to preserve privacy while ensuring a minimum reduction of the fidelity of the image. In particular, we estimate the resolution of faces captured at different altitudes and tilt angles using the data from navigation sensors and ascertain when the captured face is inherently protected. When the face is unprotected, we define a mechanism that automatically configures the strength of a privacy protection filter to improve the trade-off between privacy protection and fidelity of an aerial image or video.

1. Introduction

Micro Aerial Vehicles (MAVs) equipped with high-resolution cameras are becoming common in public places for recreational and business video capturing [2, 14]. To protect the privacy of individuals who happen to be in the vicinity and are captured in the videos, it is desirable to identify and redact portions of the images, such as faces, which would reveal their identity.

Unlike several applications of surveillance imagery such as people counting, perimeter protection and behaviour analysis whose utility is not compromised by a full redaction of the identity-related data (sensitive regions) [7], recreational or business videos require a minimal distortion of the image content in order to be usable. For these videos, utility can be defined as the fidelity of the protected images with respect to the originally captured images.

The redaction or distortion of sensitive regions can be obtained through fixed or adaptable privacy filters. *Fixed privacy filters* generally remove sensitive regions in images or replace them with a de-identified representation. Examples of fixed privacy filters include masking (blanking) [6, 23] and replacing regions representing people with silhouettes or avatars [24, 8, 9, 21]. *Adaptable privacy filters* can be configured depending on privacy and fidelity of the targets. Examples of adaptable privacy filters include pixelation [9], blurring [28], cartooning [12], scrambling [11] and warping [18]. These are adaptable privacy filters whose distortion strength can be adjusted automatically during run time, for example by configured pixelation and blurring filters based on the size of the detected objects [22]. However, new challenges *e.g.* due to mobility and the tilt angle of the airborne cameras arise for these privacy filters as most have been developed for Closed Circuit Television (CCTV) or hand-held cameras.

Recent frameworks that support privacy-preservation in airborne cameras are based on geo-fencing, generic data encryption and on-board visual data processing. For *geo-fencing*, the MAV may either contain the coordinates of restricted geographical areas in their navigation software (*e.g.* the community-generated database NoFlyZone [20]) before a mission or during a mission through beacons over WiFi to determine whether flying in certain areas is allowed or re-routing is necessary [26]. However, simple geo-fencing is insufficient to protect the privacy with oblique images in the presence of powerful optics. Alternatively, MAVs can send *encrypted data* to a privacy server that blanks, blurs or mosaics sensitive regions [17]. This approach could cause significant latency due to data transfer between the airborne camera, the privacy server and the end-user. A desirable approach is to exploit on-board visual processing to detect and then blur faces *prior* to sending the images to the end-user. UAS-VPG [3] is a specific approach that extends traditional privacy-preserving filters. However this approach does not consider (i) the tilt angle of the camera and, most importantly, (ii) how to configure a blurring filter in a moving camera. In fact, faces can be captured from various angles and distances, thus resulting in a high variation of face orientations and resolutions.

978-1-5090-3811-4/16/$31.00

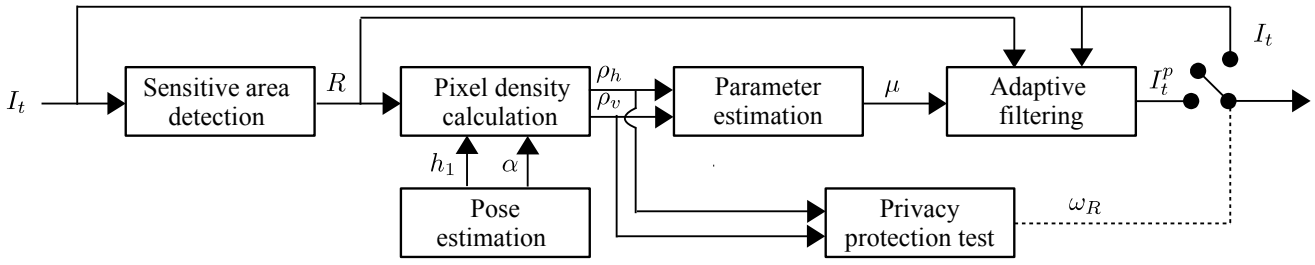


Figure 1: Block diagram of the proposed adaptive privacy filtering framework for airborne cameras.

In this paper, we focus on models for the on-board adaptive protection of faces captured from airborne cameras. We define a mechanism for *privacy design space exploration* that allows us to automatically configure an adaptive privacy filter. The mechanism uses the resolution of the detected sensitive region to determine when it is inherently protected. We use the auxiliary data from the on-board navigation sensors (Global Positioning System (GPS) and Inertial Measurement Unit (IMU)) to determine when a face is not inherently protected and then apply an adaptive privacy filter that distorts a sensitive region depending upon its captured resolution. The block diagram of the proposed approach is shown in Figure 1.

The rest of the paper is organised as follows: Section 2 formulates the privacy design space and discusses its analytical results, while Section 3 describes the design of an adaptive privacy filter. Section 4 presents experimental results and Section 5 concludes this paper.

2. Privacy Design Space

Let an MAV fly at an altitude of h_1 meters. Let the principal axis $\boldsymbol{P}$ of its on-board camera be tilted by α from the nadir direction $\boldsymbol{N}$ (see Figure 2). A value of $\alpha \neq 0$ generates an oblique image. Each frame I_t at time t could contain N_t faces. However, for simplicity but without loss of generality we will consider in this paper only the case of $N_t = 1$. We represent the face region in the image as R, which is viewed at an angle β. Finally, let h_2 be the height of the face above ground.

Let the *pixel density* be defined as the number of pixels (px) per unit distance (cm) in a certain portion of the image I_t at time t. Let ρ_h and ρ_v represent the pixel density (px/cm) around the centre C_R of R in the horizontal and vertical direction, respectively (see Figure 3).

Our goal is to determine whether ρ_h and ρ_v are sufficiently large to facilitate the recognition of the identity of the person whose face is captured in R and then to apply an adaptive privacy filter in case recognition is considered possible.

Using reliable commercial-off-the-shelf differential GPS and IMU units, an MAV can estimate its position with an error that is smaller than $10 - 20$ cm [25]. In this work, we assume that on-board navigation sensor's data is available for each I_t with high accuracy and therefore use these auxiliary data to estimate h_1 and α.

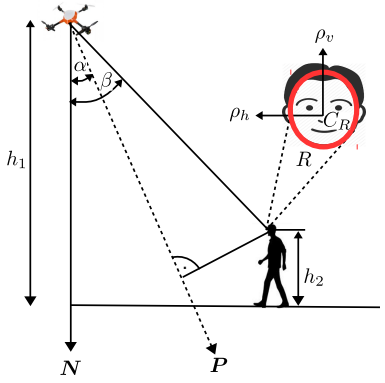


Figure 2: Capturing an image with an airborne camera at height h_1. The principal axis $\boldsymbol{P}$ of the camera is tilted by α from the nadir direction $\boldsymbol{N}$. The sensitive region R, a face at height h_2 above the ground, is viewed at an angle β. The variables ρ_h and ρ_v represent the horizontal and vertical pixel density of R at its centre C_R in the captured image.

Given a face detection result R in I_t, we determine β using C_R. For $\alpha \neq 0$, a pixel of the image sensor covers an area of the 3D world of different size depending on its distance and orientation [15], [19]. Let p_h and p_v represent the physical dimensions of a pixel in the horizontal and vertical direction, respectively. If f is the focal length of the camera, we can determine the horizontal density ρ_h for a pixel around C_R as

$$\rho_h = \frac{f cos(\beta)}{p_h(h_1 - h_2)}, \tag{1}$$

and the vertical density ρ_v as

$$\rho_v \approx \frac{f cos(\beta) sin(\beta)}{p_v(h_1 - h_2)}. \tag{2}$$

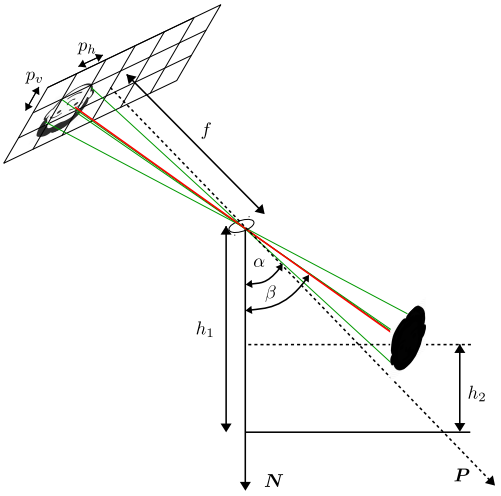


Figure 3: Mapping of a face at height h_2 to an image plane with oblique imagery. The height h_2 is considered between the ground and the center of the face, C_R, which is seen from the back in the illustration.

Let the binary status variable $\omega_R \in \{0, 1\}$ define whether R is naturally protected ($\omega_R = 0$) because of a low horizontal and vertical density, or not ($\omega_R = 1$):

$$\omega_R = \begin{cases} 1 & if \quad \rho_h > \rho_h^0 \quad \& \quad \rho_v > \rho_v^0 \\ 0 & otherwise \end{cases} \tag{3}$$

where ρ_h^0 and ρ_v^0 are experimentally defined thresholds. If $\omega_R = 0$, then the original frame I_t can be transmitted without any modifications.

Figures 4a and 4b show an example of ρ_h and ρ_v if we consider as camera a Canon EOS 5D Mark II in HD mode (1920×1080 pixels), whose sensor dimensions are 36 x 24 mm^2, thus $p_h = 18.75$ μm and $p_v = 22.22$ μm. We chose $h_2 = 1.7$ m and h_1 from 3 m to 150 m. Typical lenses have a focal length from 10 mm to 200 mm. Finally, β is assumed to vary from $0°$ to $90°$. If we accept to compromise on fidelity, we can globally filter I_t using ρ_h and ρ_v determined at $\beta = 45°$. This would lead in certain image-capturing conditions to an unnecessarily high level of blurring that reduces the fidelity of I_t.

Figure 4c depicts the boundary between the privacy sensitive space and the inherently protected space. The privacy sensitive space is determined by intersecting the individual segmentations based on ρ_h and ρ_v using the thresholds $\rho_h^0 = 1$ px/cm and $\rho_v^0 = 1$ px/cm.

3. Adaptive Privacy Filter

When R is unprotected (i.e. the face is recognizable), we want to apply an adaptive privacy filter $\mathcal{G}$ so that the fidelity of the images can be increased, while protecting the identity, with respect to the use of a fixed privacy filter. Let I_t^p be a frame with a protected sensitive region generated as

$$I_t^p = \mathcal{G}(I_t, R, \mu), \tag{4}$$

where μ represents the parameter of an adaptive privacy filter. The value of μ is such that the corresponding filtering with $\mathcal{G}$ turns ρ_h and ρ_v to be smaller than ρ_h^0 and ρ_v^0, respectively. For example, in pixelation μ is determined by the averaging kernel size [9], in blurring by the standard deviation of a Gaussian [28], in cartooning by the kernel size of a mean shift filter [12], in scrambling by the number of transform coefficients [11] and in warping by the relocation distance of pixels with respect to the calculated values of ρ_h and ρ_v [18]. Due to the difference in their protection operation, these techniques will have different thresholds ρ_h^0 and ρ_v^0 for the same sensitive region R.

In this work, we use the Gaussian blur that is widely used for outdoor imagery [13]. $\mathcal{G}$ therefore becomes a convolution operation on R. We use an approximated anisotropic Gaussian kernel defined as

$$g(h, v) = \frac{1}{2\pi\sigma_h\sigma_v} e^{-\left(\frac{h^2}{2\sigma_h^2} + \frac{v^2}{2\sigma_v^2}\right)}, \tag{5}$$

where σ_h and σ_v are the standard deviations of the Gaussian in the horizontal and vertical direction, respectively. The estimation of σ_h and σ_v depends on their spatial cut-off frequencies f_h^c and f_v^c, respectively, which in turn depend on the selected point on the filter response (i.e. 3 dB, "half width half maximum" or "close to zero"). In this work, we chose "close to zero" (i.e. three times of a Gaussian function's standard deviation in the frequency domain) to suppress frequency components higher and equal than f_h^c and f_v^c sufficiently. After transforming f_i^c with $i \in \{h, v\}$ from frequency to space domain (i.e. threshold ρ_i^0), the estimated values for σ_h and σ_v are given by

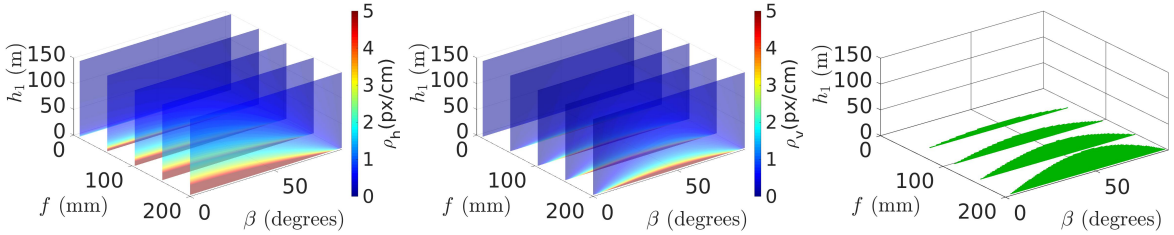
$$\sigma_i = \frac{3\rho_i}{\pi\rho_i^0}. \tag{6}$$

Finally, μ is determined by sampling the Gaussian function upto three times of σ_i as

$$\mu_i = 2\lceil 3\sigma_i \rceil + 1. \tag{7}$$

After convolving with a kernel of size μ_i, the useful information is reduced to ρ_i^0 (px/cm). Figure 5 shows sample results of adaptive Gaussian blurring, estimated standard deviations σ_h and σ_v, and kernel sizes μ_h and μ_v. From the figure, it is apparent how μ_h and μ_v decrease with decreasing ρ_h and ρ_v, respectively. This adaptive behaviour of μ_h and μ_v aims at maintaining the fidelity of the images. The required values of ρ_h^0 and ρ_v^0 depend on the expected ability of a face recognizer to discriminate identities.

In the next section we quantify the benefits of the proposed blurring approach.



(a) (b) (c)

Figure 4: Horizontal and vertical pixel densities for EOS 5D Mark II in HD video mode at different heights, with different focal lengths and tilt angles. (a) Variation of ρ_h for different values of h_1, f and β. The largest values of ρ_h occur at $\beta = 0°$. (b) Variation of ρ_v for different values of h_1, f and β. The largest values of ρ_v occur at $\beta = 45°$. (c) Separation of the privacy sensitive space from the inherently protected space using $\rho_h^0 = \rho_v^0 = 1$ px/cm as segmentation threshold. The privacy sensitive space (indicated by the green slices) corresponds to the intersection of the segmentation performed in (a) and (b).

Figure 5: Comparison between various degrees of Gaussian blur to protect a facial image from an LDA face classifier. The numbers following $a*$ represent the value of ρ_h and ρ_v, respectively. The numbers following $b*$, $c*$, $d*$ represent the value of μ_h and μ_v, respectively. (a1-a7) Original images from [16] with decreasing pixel densities (from left to right). (b1-b7) Gaussian blur on (a1-a7) with $\rho_h^0 = \rho_v^0 = 0.5$ that results in slightly under-blurred images (i.e. the recognition rate is higher than that of a random classifier). (c1-c7) Adaptive Gaussian blur on (a1-a7) based on our proposed framework ($\rho_h^0 = \rho_v^0 = 0.4$) resulting in an adequate blurring of the faces for privacy preservation. With an adaptive Gaussian blur the kernel is selected depending on the pixel density for R. (d1-d7) Gaussian blur on (a1-a7) with $\rho_h^0 = \rho_v^0 = 0.3$ that results in slightly over-blurred images that, although they make the recognition rate equivalent to that of a random classifier, they unnecessarily decrease the fidelity of the facial images. (e1-e7) Fixed Gaussian blur of (a1-a7) using a *safe kernel* designed considering the highest possible pixel density in order to make the recognition accuracy of a classifier equivalent to that of a random classifier, irrespective of the pixel density for R, the face. This fixed Gaussian blur ($\mu_h = 121$, $\mu_v = 105$) significantly deteriorates the fidelity of lower resolution faces.

4. Experimental Results

In this section we first study different target resolutions to determine the threshold values for separating inherently protected and unprotected spaces. We then analyse the trade-off between privacy and fidelity for adaptive and fixed privacy filters using standard face recognition algorithms by measuring their recognition performance with images from airborne cameras. We use the LDA [5] and LBPH [1] algorithms for face recognition. The LDA face recognizer reduces the class dimension by maximising the inter-class to intra-class scatter ratio. In contrast, the LBPH face recognizer encodes a local structure instead of a full image to reduce the class dimension.

To measure fidelity, we apply the Structural Similarity Index (SSIM) and the Peak Signal to Noise Ratio (PSNR). The SSIM measures the image quality by comparing the degradation in the structure of a protected image with the original [27], while the PSNR represents the power ratio of the original image with respect to the error introduced by protection.

4.1. Experimental Setup

We use an outdoor dataset emulating an MAV dataset with the availability of auxiliary data, created with a camera placed at different heights and distances from faces [16]. In this dataset, the principal axis of the camera is parallel to the ground, i.e. $\alpha = 90°$. In order to compute the pixel densities for this particular setup, we modify Equations 1 and 2 as $\rho_h = \frac{f cos(\gamma)}{p_h d}$ and $\rho_v \approx \frac{f cos^2(\gamma)}{p_v d}$, where $\gamma = 90° - \beta$ and d represents the horizontal distance between the face and the camera.

We consider an image as privacy protected when the face recognition algorithms achieve an accuracy similar to that of a random classifier and therefore look for threshold values ρ_h^0 and ρ_v^0 resulting in a random classifier accuracy. The accuracy of a face recognizer corresponds to the rank 1 value of the cumulative match curve.

The data set contains 11 subjects and thus the accuracy of a random classifier for this dataset is 0.091 $(1/11)$. We chose data from 63 different positions for each individual resulting in a total of 693 test images (63×11).

To train the LDA and LBPH face recognizers, we used separate training images, which consisted of 11 images for each subject. For the detection of the face region, we annotated the images with the ground truth of the eye locations. As described in [4], we pre-processed all training and test images by (i) applying an affine transformation to compensate for scale and face rotation, (ii) using a bilateral filter to reduce noise and to compensate for light variations and finally (iii) masking to remove non-facial portions.

For the test images, we determined the values of ρ_h and ρ_v both by using auxiliary data (see Equations 1 and 2) and by manually counting the number of pixels on a face and normalising it by the standard face dimensions, i.e. the bitragion breadth of 15.9 cm and menton-crinion length of 21.9 cm [10]. As shown in Figure 6a, there is a small difference between the calculated pixel densities using these two methods with a Root Mean Square Error (RMSE) of 0.39 px/cm for ρ_h and 0.74 px/cm for ρ_v, respectively. The main reason for this error lies in our assumption of the identical height (h_2) and the identical face dimension for all 11 subjects.

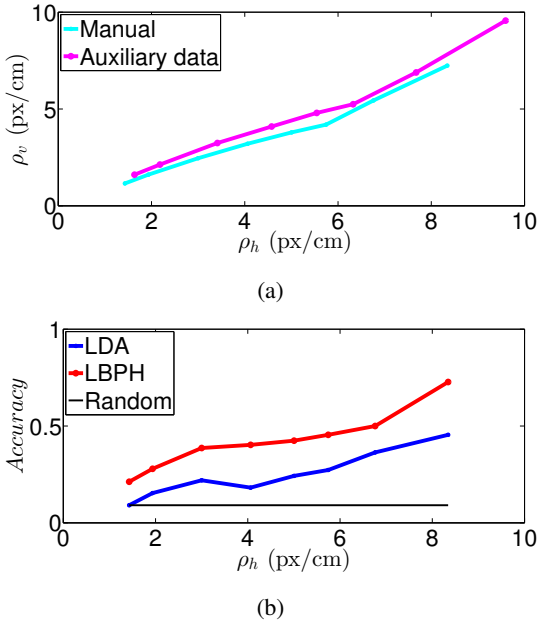


Figure 6: (a) The pixel density variation (ρ_h and ρ_v) of the 693 images in the data set based on auxiliary data and manually counting the pixels. (b) The achieved accuracy of the LDA and LBPH face recognizers over the 693 raw images.

4.2. Privacy Design Space

Figure 6b shows the recognition accuracy of LDA and LBPH over the original 693 test images as well as the response of a random classifier for reference. The LBPH face recognizer clearly outperforms the LDA face recognizer. However, the clear identification of the threshold values for the separation between the inherently protected and unprotected space is not possible with this data. For LBPH, the separating boundary lies *below* $\rho_h = 1.43$ px/cm (and corresponding $\rho_v = 1.15$ px/cm). Although LDA touches the random classifier level at $\rho_h = 1.43$ px/cm, this resolution cannot be considered as threshold with high confidence, because the accuracy of LDA also dropped at $\rho_h = 4.1$ px/cm and then again increased at $\rho_h = 2.99$ px/cm. Such behaviour may be due to the limited population size of the dataset. Thus, the separating boundary lies *at or below*

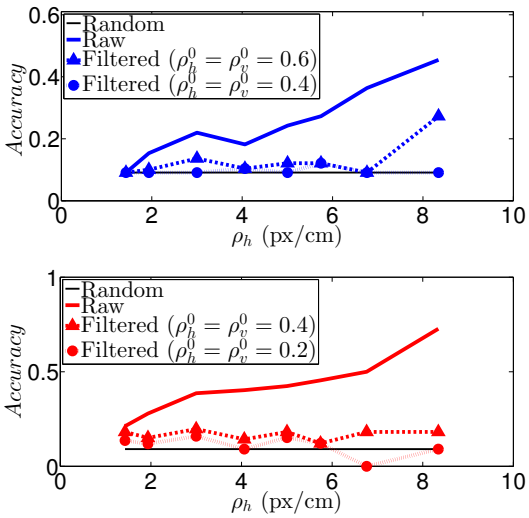


Figure 7: Face recognition accuracy of adaptively filtered data with different threshold values ρ_h^0 and ρ_v^0. (Top) LDA: The accuracy is similar to a random classifier at $\rho_h^0 = 0.4$ px/cm, $\rho_v^0 = 0.4$ px/cm. (Bottom) LBPH: The average response is close to a random classifier at $\rho_h^0 = 0.2$ px/cm, $\rho_v^0 = 0.2$ px/cm.

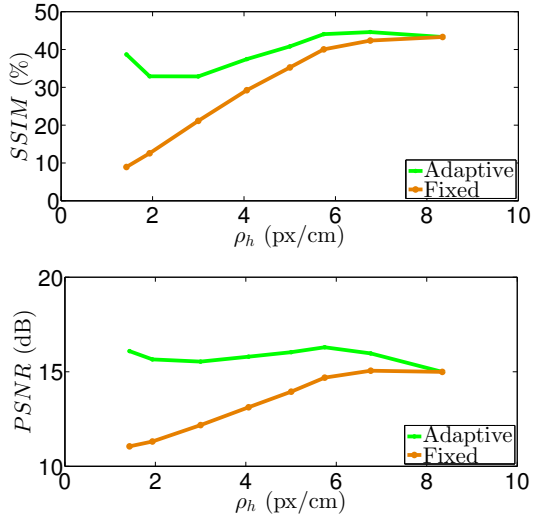


Figure 8: Fidelity achieved with adaptive and fixed privacy filtering over 693 images using the (top) Structural Similarity Index (SSIM) and the (bottom) Peak Signal to Noise Ratio (PSNR).

$\rho_h = 1.43$ px/cm (and corresponding $\rho_v = 1.15$ px/cm) for LDA.

4.3. Privacy Adaptive Filtering

To explore the boundary further, we filter all test images with adaptive Gaussian blur to degrade the pixel resolution to the specified levels of $0.6, 0.4$ and 0.2 px/cm for both ρ_h^0 and ρ_v^0. The kernel size of the Gaussian blur filter ($\mu_h \times \mu_v$) is computed as described in Section 3. We then determine the recognition accuracy over the degraded images. Figure 7 shows that the accuracy of LDA remains mostly around the random classifier accuracy of 0.091 at $\rho_h^0 = 0.4$ px/cm and $\rho_v^0 = 0.4$ px/cm. Thus, we define $\rho_h^0 = 0.4$ px/cm and $\rho_v^0 = 0.4$ as the boundary between the inherently protected and unprotected spaces for LDA. Although the accuracy of LBPH fluctuates, its average response is close to a random classifier at $\rho_h^0 = 0.2$ px/cm and $\rho_v^0 = 0.2$ px/cm. We therefore use $\rho_h^0 = 0.2$ px/cm and $\rho_v^0 = 0.2$ px/cm as the separating boundary for the LBPH.

Finally, we compare the fidelity of the adaptively filtered images with images filtered with a fixed safe kernel. Therefore, we apply adaptive Gaussian blurring with $\rho_h^0 = 0.4$ px/cm and $\rho_v^0 = 0.4$ px/cm and fixed Gaussian blurring with kernel size $\mu_h = 121 \times \mu_v = 105$ to all images and measure the SSIM and PSNR values. Figure 8 shows that adaptive filtering provides a higher fidelity in terms of SSIM and PSNR as compared to fixed safe filtering.

5. Conclusions

We presented a privacy protection framework for airborne cameras on recreational unmanned aerial vehicles. The horizontal and vertical resolutions of sensitive regions can be determined for tilted cameras from different heights using data from navigation sensors. This knowledge of the resolution can be used to determine whether a sensitive region is inherently protected or unprotected. For an unprotected region, we exploit this prior knowledge of the resolution to configure an adaptive privacy filter that keeps the fidelity of aerial images high while preserving privacy. In the specific implementation discussed in this paper, we showed the effectiveness of the proposed approach on faces by measuring their fidelity. However, the proposed approach can be applied to other sensitive regions, such as car number plates and can also be extended to other than fidelity.

As future work we will expand the experimental evaluation with larger population size and we will extend the study to incorporate recent face recognition algorithms. Moreover, we will investigate potential privacy leakage caused by super resolution filters.

6. Acknowledgements

We thank Hwai-Jung Hsu and Kuan-Ta Chen for sharing their data set [16] for our experimental analysis.

O. Sarwar has been supported in part by Erasmus Mundus Joint Doctorate in Interactive and Cognitive Environment, which is funded by the Education, Audio-visual & Culture Executive Agency under the FPA no 2010-0015.

References

[1] T. Ahonen, A. Hadid, and M. Pietikainen. Face Description with Local Binary Patterns: Application to Face Recognition. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 28(12):2037–2041, December 2006.

[2] AirDog. https://www.airdog.com/. [Last accessed: 2016-04-14].

[3] R. Babiceanu, P. Bojda, R. Seker, and M. Alghumgham. An onboard UAS visual privacy guard system. In *Proc. Integrated Communication, Navigation, and Surveillance Conf. (ICNS)*, pages J1:1–J1:8, Herdon, VA, USA, April 2015.

[4] D. L. Baggio, S. Emami, D. M. Escriva, K. Ievgen, N. Mahmood, J. Saragih, and R. Shilkrot. *Mastering OpenCV with Practical Computer Vision Projects*. Packt Publishing, Limited, December 2012.

[5] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19(7):711–720, July 1997.

[6] A. Cavallaro. Adding privacy constraints to video-based applications. In *Proc. European Workshop on the Integration of Knowledge, Semantics and Digital Media Technology*, page 8, London, UK, November 2004.

[7] A. Cavallaro. Privacy in Video Surveillance. *IEEE Signal Processing Magazine*, 24(2):168–169, July 2007.

[8] D. Chen, Y. Chang, R. Yan, and J. Yang. Tools for Protecting the Privacy of Specific Individuals in Video. *EURASIP Journal on Applied Signal Processing*, 2007(1):107–107, January 2007.

[9] K. Chinomi, N. Nitta, Y. Ito, and N. Babaguchi. PriSurv: Privacy Protected Video Surveillance System Using Adaptive Visual Abstraction. In *Proc. Int. Conf. on Advances in Multimedia Modeling*, pages 144–154, Kyoto, Japan, January 2008.

[10] Human Engineering Design Data Digest, Department of Defense Human Factors Engineering Technical Advisory Group (DOD HFE TAG). http://www.acq.osd.mil/rd/hptb/hfetag/products/documents/HE_Design_Data_Digest.pdf, April 2000.

[11] F. Dufaux and T. Ebrahimi. Scrambling for Video Surveillance with Privacy. In *Proc. Computer Vision and Pattern Recognition Workshops*, pages 160–160, New York, USA, June 2006.

[12] A. Erdelyi, T. Barat, P. Valet, T. Winkler, and B. Rinner. Adaptive cartooning for privacy protection in camera networks. In *Proc. Int. Conf. on Advanced Video and Signal Based Surveillance (AVSS)*, pages 44–49, Seoul, Korea, August 2014.

[13] A. Frome, G. Cheung, A. Abdulkader, M. Zennaro, B. Wu, A. Bissacco, H. Adam, H. Neven, and L. Vincent. Large-scale Privacy Protection in Google Street View. In *Proc. IEEE Int. Conf. on Computer Vision*, pages 2373–2380, Kyoto, Japan, October 2009.

[14] Hexo+. https://hexoplus.com/. [Last accessed: 2016-04-14].

[15] J. Höhle. Oblique aerial images and their use in cultural heritage documentation. In *Proc. Int. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XL-5/W2, pages 349–354, Strasbourg, France, September 2013.

[16] H.-J. Hsu and K.-T. Chen. Face Recognition on Drones: Issues and Limitations. In *Proc. First Workshop on Micro Aerial Vehicle Networks, Systems, and Applications for Civilian Use*, DroNet '15, pages 39–44, Florence, Italy, May 2015.

[17] Y. Kim, J. Jo, and S. Shrestha. A server-based real-time privacy protection scheme against video surveillance by unmanned aerial systems. In *Proc. Int. Conf. on Unmanned Aircraft Systems (ICUAS)*, pages 684–691, Orlando, FL, USA, May 2014.

[18] P. Korshunov and T. Ebrahimi. Using warping for privacy protection in video surveillance. In *Proc. Int. Conf. on Digital Signal Processing (DSP)*, pages 1–6, Fira, Santorini, Greece, July 2013.

[19] P. Meixner and F. Leberl. Interpreting building facades from vertical aerial images using the third dimension. In *Proc. Joint Symposium of ISPRS Technical Commission IV & AutoCarto*, pages 15–19, Orlando, FL, USA, November 2010.

[20] NoFlyZone. https://www.noflyzone.org/. [Last accessed: 2016-04-14].

[21] F. Qureshi. Object-Video Streams for Preserving Privacy in Video Surveillance. In *Proc. Int. Conf. on Advanced Video and Signal Based Surveillance (AVSS)*, pages 442–447, Genova, Italy, September 2009.

[22] M. Saini, P. K. Atrey, S. Mehrotra, and M. Kankanhalli. Adaptive Transformation for Robust Privacy Protection in Video Surveillance. *Advances in Multimedia*, 2012:1–14, Feburary 2012.

[23] A. 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*, 3(3):50–57, June 2005.

[24] S. Tansuriyavong and S. Hanaki. Privacy Protection by Concealing Persons in Circumstantial Video Image. In *Proc. Workshop on Perceptive User Interfaces*, pages 1–4, Orlando, FL, USA, November 2001.

[25] D. Turner, A. Lucieer, and L. Wallace. Direct Georeferencing of Ultrahigh-Resolution UAV Imagery. *IEEE Trans. on Geoscience and Remote Sensing*, 52(5):2738–2745, May 2014.

[26] T. Vaidya and M. Sherr. Mind your (R, Φ)s: Location-Based Privacy Controls for Consumer Drones. In *Proc. Intern. Workshop on Security Protocols*, pages 91–104, Cambridge, UK, March 2015.

[27] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. on Image Processing*, 13(4):600–612, April 2004.

[28] J. Wickramasuriya, M. Datt, S. Mehrotra, and N. Venkatasubramanian. Privacy Protecting Data Collection in Media Spaces. In *Proc. Int. Conf. on Multimedia*, pages 48–55, New York, NY, USA, October 2004.