Serious Fun: Cartooning for Privacy Protection

Ádám Erdélyi,
Institute of Networked and Embedded Systems
Alpen-Adria Universität Klagenfurt
Lakeside Park B02b,
9020 Klagenfurt, Austria
adam.erdelyi@aau.at

Thomas Winkler
Institute of Networked and Embedded Systems
Alpen-Adria Universität Klagenfurt
Lakeside Park B02b,
9020 Klagenfurt, Austria
thomas.winkler@aau.at

Bernhard Rinner
Institute of Networked and Embedded Systems
Alpen-Adria Universität Klagenfurt
Lakeside Park B02b,
9020 Klagenfurt, Austria
bernhard.rinner@aau.at

ABSTRACT

This paper presents our solution to the 2013 MediaEval Visual Privacy Task [3]. We apply cartoon-like effects to captured video such that identities of persons are protected while behavioural information and hence system intelligibility are maintained. We present our processing pipeline which includes additional protection steps such as re-colouring or additional blurring and discuss early evaluation results.

1. MOTIVATION AND GOALS

Privacy in visual surveillance is one of the outstanding open issues currently investigated by several research groups. Privacy protection goes along with a loss of intelligibility (i.e., utility of the video's content) since filters such as blanking or pixelization reduce the amount of visible information. An ideal approach should achieve a tradeoff which preserves behavioural information while protecting identities.

The 2013 MediaEval Visual Privacy Task [3] provides a set of video sequences, the PEViD data-set [5], together with annotations of sensitive regions (e.g., faces, persons, carried items) in separate data files. The privacy protection algorithms developed by participants should obscure people and their personal belongings while maintaining sufficient information for observers to monitor behaviour and actions.

With the limited robustness of state-of-the art object detection techniques in mind, we designed our approach to be extendable to cope with misdetections and implicit privacy leakage channels [6]. Our solution focuses on the whole body of monitored people including their carried items instead of obscuring their faces only.

2. PRIVACY PROTECTION APPROACH

Our approach consists of four sequential steps (cp. Figure 1) including (1) pre-processing, (2) item-recolouring, (3) cartooning and (4) additional face blurring. The individual steps are discussed throughout Sections 2.1 to 2.4. Our prototype is implemented in C++ using OpenCV [1]. For parsing the provided annotation files we use pugixml [2].

2.1 Preprocessing

Before applying the privacy filters, some preprocessing is required. First, the motion regions are detected by using background extraction. A preliminary blurring with a 7×7 size kernel is done before the motion region detection in order to reduce noise. Secondly, the edges are detected with a Sobel edge detector. Both the motion regions and the edges are used as masks in the privacy filter.

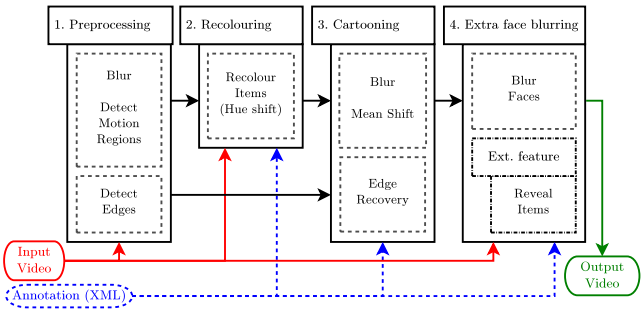

Figure 1: Pipeline of the proposed privacy filter.

2.2 Recolouring Items

Before applying the cartooning effect, all personal items are recoloured in order to distort their original characteristics as much as possible. Instead of simply using the rectangular bounding box from the annotation, we merge it with the motion region mask by using a bitwise-AND operation. This way the colour change is done only to the item and the viewer's attention is not distracted by the sharp edges of a recoloured regular rectangle. The colour change itself is performed by shifting the hue level by 180°. The RGB (Red-Green-Blue colour model) frame is converted to HLS (Hue-Lightness-Saturation colour model) beforehand and back to RGB after the hue shift.

2.3 Cartooning

The cartooning effect is applied to the ROI (Region of Interest) in the image frame, namely to the bounding boxes of people and personal items. A preliminary blur is also applied in order to reduce noise and achieve visually better end results. Furthermore, the ROIs are again merged with the contour mask as described earlier in order to avoid viewer distraction. The cartooning is done in two steps:

1. Applying a Mean Shift Filter [4] with a spatial window radius of 20 and a colour window radius of 40. This makes the image smooth and reduces the number of colours as if it was drawn like a cartoon.

MediaEval 2013 Workshop, October 18-19, 2013, Barcelona, Spain

2. To enhance the cartoon look, the original image is restored along object contours. A bitwise weighted copy from the original frame to the processed frame is performed using the gradient mask from the Sobel edge detector. This makes the image less blurry after the Mean Shift Filter and more similar to a cartoon where contours of the drawings are usually emphasized. Optionally, a single colour (e.g., black) could be used for contours instead of copying the original image. However, this was found to be too distractive due to the added noise.

2.4 Additional Face Blurring

In cases where faces are relatively big (i.e., close to the camera), cartoonization does not provide sufficient privacy protection. Therefore, we introduced an additional, adaptive face blurring effect. A Gaussian blur is applied to the faces in an elliptical ROI with a kernel size that is proportional to the size of the face. The elliptical ROIs have gradient edges in order to avoid viewer distraction.

3. EXTENDED FUNCTIONALITY

A central aspect of video surveillance is the detection of unusual events such as abandoned luggage. Even though it is not required in MediaEval's Visual Privacy Task, we integrated an extra feature into our processing pipeline which prevents protection from supposedly critical image regions. Critical events leading to uncovering of such regions are:

- A bag or backpack is dropped (i.e., not moving).
- An umbrella or a bottle is used as a weapon in a fight.
- A wallet is being stolen.

This function further enhances intelligibility and helps the operator of the surveillance system in determining when to react.

4. RESULTS

Figure 2 presents two pairs of frames (original and modified) which show the results of the applied cartooning effect. The re-colouring of personal items is nicely visible at the scarf and the bag in Figure 2b.

(a) Original. (b) Modified. (c) Original. (d) Modified.

Figure 2: Comparison of original frame and cartoonized frame with re-colouring (bag and scarf) and face blur.

Figure 3 visualizes the evaluation results provided by the Visual Privacy Task organizers. In the objective (i.e., automated) evaluations our approach consistently achieves higher scores than the average over all 9 participating teams in all three evaluation categories (intelligibility, privacy and appropriateness). For the subjective (i.e., human) evaluation our approach is well above average for intelligibility and appropriateness but below average for privacy protection. This indicates that a stronger cartooning effect is required to provide adequate privacy in case of human observers.

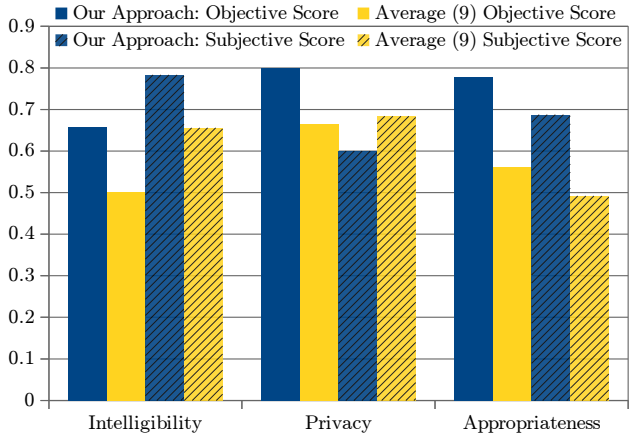

Figure 3: Evaluation scores for intelligibility, privacy and appropriateness. Blue bars show the scores for our proposed approach; Yellow bars show the average scores over the 9 Privacy Task participants. The plain bars show the result for objective (i.e., automated) evaluation while the hatched bars show the results for subjective (i.e., human) evaluation.

5. CONCLUSION AND FUTURE WORK

The performance bottleneck of our approach, that currently prevents real-time applications, is OpenCV's Mean Shift Filter which is relatively slow. There are alternative and possibly faster algorithms that can produce a similar cartooning effect which will be explored in the future work.

Assuming that such a cartooning algorithm sufficiently hides facial identity, it would be possible to apply it globally to the whole image. This way we do not depend on feature detectors that typically are not accurate and reliable enough.

Based on the evaluation criteria of the Visual Privacy Task we created a tool that compares original and filtered videos in terms of privacy and intelligibility. In future work we will use it for a comprehensive evaluation of global approaches.

6. REFERENCES

[1] OpenCV – Open Source Computer Vision. `http://opencv.org` (last visited: Sept. 2013).

[2] pugixml – Light-weight, simple and fast XML parser for C++ with XPath support. `http://pugixml.org` (last visited: Sept. 2013).

[3] A. Badii, M. Einig, and T. Piatrik. Overview of the MediaEval 2013 Visual Privacy Task. In *Proceedings of the MediaEval Workshop*, Barcelona, Spain, 2013.

[4] Yizong Cheng. Mean shift, Mode Seeking, and Clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8):790–799, 1995.

[5] P. Korshunov and T. Ebrahimi. PEViD: Privacy Evaluation Video Dataset at Applications of Digital Image Processing XXXVI. In *Proceedings of SPIE*, volume 8856, San Diego, CA, 25-29 August 2013.

[6] M. Saini, P. K. Atrey, S. Mehrotra, and M. S. Kankanhalli. Considering Implicit Channels in Privacy Analysis of Video Data. *IEEE Communications Society E-Letters*, 6(11):27–30, 2011.