Overview of MediaEval 2012 Visual Privacy Task

Atta Badii[1], Tomas Piatrik[2], Mathieu Einig[1], Chattun Lallah[1]

[1] University of Reading
Intelligent Systems Research Laboratory,
School of Systems Engineering
United Kingdom
{atta.badii; m.l.einig; c.lallah}@reading.ac.uk

[2] Queen Mary,
University of London
Multimedia & Vision Research Group
United Kingdom
tomas.piatrik@elec.qmul.ac.uk

ABSTRACT

In this paper, we describe the Visual Privacy Task, including its aim, its related dataset and the evaluation methods.

Keywords

Privacy, Privacy Protection, Video Analytics, Image Processing, Tracking, Evaluation

1. INTRODUCTION

The advances in artificial intelligence and video surveillance are leading to more and more complex surveillance systems of growing scale and capabilities. This ubiquity and intelligence poses great threats to privacy, and new mitigation technologies must be found in order to ensure an appropriate level of privacy protection.

The Visual Privacy Task (VPT) is a brand new task in MediaEval 2012 exploring how image processing, computer vision and scrambling techniques can deliver technological solutions to some visual privacy problems. The goal of privacy protection is to prevent potential access to information, the divulgement of which can amount to a (perceived) intrusion of an individual's privacy. The extent of such a (perceived) loss of privacy depends on the individual as well as the context and as such can only be determined by reference to the user ("data-subject") in each case. Accordingly this task focuses on human face privacy protection techniques for videos without its context-dependent variations, the consideration of which are excluded for the purpose of this particular task. The evaluation is performed using both video analytics algorithm and user study in order to evaluate both objectively and subjectively.

2. THE VPT DATASET

The data set consists of videos collected from a range of standard and high resolution cameras and contains clips of different scenarios showing one or several persons walking in front of the cameras. People may also wear specific wardrobe elements which could potentially reveal their identity. For this year, the list of elements considered in the task is reduced to 3 items: scarf, hat, and glasses. People may be at a distance from the camera or near the camera, making their faces vary considerably in pixel size and quality. The ambient lighting conditions of the videos vary across widely and half of the clips were recorded at night. The dataset contains 200 video clips and associated annotations in xml form. The video clips are divided into subgroups: morning session, evening session and clips containing people wearing accessories. The clips are in mpeg format with a frame rate of either 25 or 30 fps, and an image size of either 640x480 or 704x576 pixels. More details about the dataset can be found in [1].

MediaEval 2012 Workshop, October 4-5, 2012, Pisa, Italy

3. VISUAL PRIVACY TASK

This task explores how image processing, computer vision and scrambling techniques can deliver technological solutions to some visual privacy problems [2] [3] [4]. The goal of privacy protection is to prevent potential access to information, the divulgement of which can amount to a (perceived) intrusion of an individual's privacy. The extent of such a (perceived) loss of privacy depends on the individual as well as the context and as such can only be determined by reference to the user ("data-subject") in each case. Accordingly this task focuses on human face privacy protection techniques for videos without its context-dependent variations, the consideration of which are excluded for the purpose of this particular task.

The goal of the task is to propose ways, in which human faces occurring in digital imagery can be obscured so as to render them unrecognisable. This will ensure that a person appearing in a video frame cannot be visually identified, unless certain over-riding criteria prevent such masking of the individual's face. Such criteria can be based on a priori detected anomalies, critical events, alerts etc. or on prior official permission granted by higher authorities to suspend the protection of the identity of an individual in specific cases. Since the resulting partly obscured videos are still to be available for viewing, an optimal balance should be struck so that despite the extent of such masking of the facial identity as may be necessary, the categorical identity of any masked data subject therein, as a human being, can still be recognisable to the viewer. Thus the facial identity obscuring techniques should not result in artefacts that are 'socially inappropriate/offensive' and unacceptable to the human users.

Therefore, the goal of the Visual Identity Protection task is to automatically detect and make unidentifiable, the faces appearing in digital imagery. The participants should also demonstrate that their choice of obscuring technique is such that the resulting obscured (e.g. pixelated) faces do not become the focal point of a viewers' attention (standing out, distracting) thus adversely impacting the acceptability-usability of the resulting obscured/scrambled images, or the whole frame, from the perspective of both the data-subject as well as that of the person watching the video. Participants are provided with videos containing faces from different camera angles. The goal of the task is to identify which technique is more acceptable for the task of Face Privacy Protection in practical situations, for example, in live video surveillance scenarios.

4. EVALUATION METHODOLOGIES

The ground truth consists of annotations of faces and defined personal accessories which are specifically confined to three accessory types: scarf, hat and glasses. The evaluation of the results is based on the following techniques and approaches.

4.1 Objective metrics

4.1.1 Face Detection (Privacy)

Face detection algorithm will be run on the obscured videos submitted for the evaluation using Viola-Jones face detection from OpenCV library. Ideally, no faces should be found, since they all should be obscured. The faces found by the face detection algorithm are matched against the ground truth to avoid taking into account false positives of the detection algorithm.

Possible false positives of the detection algorithm are verified manually. Furthermore, the xml files created by the participants, listing the positions of the obscured areas, are compared against the ground truth. The area of overlap between the detected bounding boxes and the ground truth will be used as the main score measure. A small error threshold will be set for the comparison of the bounding boxes' sizes in order to make sure that participants will not be penalised unfairly for having obscured areas that are either slightly too large or too small. However, as the main objective is the protection of privacy, having smaller than required bounding boxes will incur more penalties than having larger ones.

4.1.2 Object Tracking (Intelligibility)

The intelligibility is measured by applying the Histogram of Oriented Gradient as a human detector to the scene. Successful detections mean that even though the sensitive areas have been obscured, the resulting video can still be used for Video Analytics. These detections are compared against the detections from the raw video.

4.1.3 Metric for Visual Appropriateness

Obscuring of faces and personal item as may be worn by a data subject (more than one such item can be present in each frame) will be evaluated using SSIM and PSNR metrics evaluating image quality based on human eye perception of salience in the image. A successful system should have a minimal impact on the global quality of the image with modifications occurring only on the sensitive areas which should be anonymised.

4.2 User Study for Assessment of Appropriateness of the Visual Privacy

As a complement to the above-mentioned objective official metrics, a random subset of videos from the submitted runs will also be evaluated through a user study aimed at developing a deeper understanding of user perceptions of appropriateness in privacy protection. This subjective evaluation will take into account three main aspects of any obscured (element of) image, namely intelligibility, privacy, and appropriateness. Given the context in the surveillance scenario under consideration, the questions related to whether a person wears personal items that can be used for identification (i.e. glasses, hat, and scarf) will be considered as relevant to privacy and intelligibility. Thus such accessories as glasses, hats and scarf will also have to be subjected to obscuring techniques just as is the face of the wearer of such accessories.

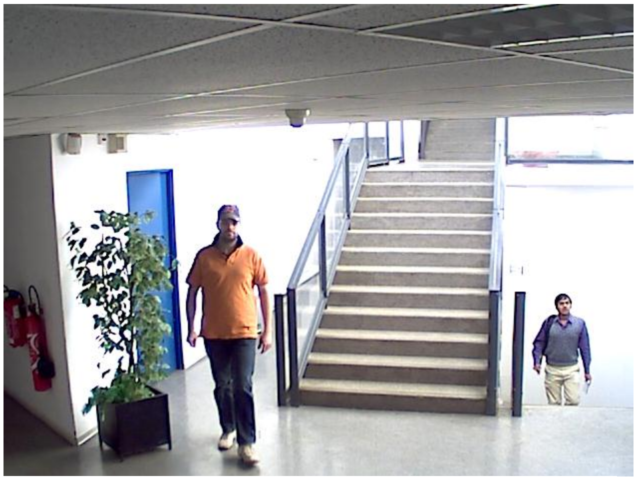

Figure 1. Sample frame from the VPT Data Set [1]

The visual appropriateness of the obscured images will be evaluated taking into account the various aspects such as pleasantness, distraction, and acceptance for video surveillance, etc. This visual appropriateness criterion will essentially map, on the one hand, the category "recognisability" of an obscured image as a member of a particular species, and, on the other hand the obscuring Effects, and the obscuring Side Effects on the perception of the image by a viewer and the extent of any resulting negative or positive emotion or distraction in the mind of the viewer of an image that has been subjected to such obscuring. Insight from this user study will serve as a baseline for refining the metrics and thereby adjust the specification of the follow-on task to be set next year.

5. ACKNOWLEDGMENTS

This Visual Privacy MediaEval task was supported by the European Commission under contracts FP7-261743 VideoSense.

6. REFERENCES

[1] Carmelo, V., Claudia, A., Dugelay, J.-L., "Synthetic and privacy-preserving visualization of video sensor network outputs," ICDSC'11, 5th ACM/IEEE International Conference on Distributed Smart Cameras, August 22-25, 2011, Ghent, Belgium

[2] Dufaux, F. & Ebrahimi, T., "Scrambling for Privacy Protection in Video Surveillance Systems," IEEE Transaction on Circuits and Systems for Video Technology, Vol. 18, Nr. 8 (2008), p. 1168-1174

[3] Dufaux, F. & Ebrahimi, T., "A framework for the validation of privacy protection solutions in video surveillance," 2010 IEEE International Conference on Multimedia and Expo (ICME), pp.66-71, 19-23 July 2010.

[4] Senior, A., "Privacy Protection in a Video Surveillance System," Privacy Protection in Video Surveillance, Springer, 2009