TUB @ MediaEval 2013 Visual Privacy Task: Reversible Scrambling with colour-preservative Characteristic

Sebastian Schmiedeke, Pascal Kelm, and Thomas Sikora
Communication Systems Group
Technische Universität Berlin, Germany
{schmiedeke,kelm,sikora}@nue.tu-berlin.de

ABSTRACT

This paper describes our participation in the Visual Privacy Task of MediaEval 2013, which aims to obscure human occurrence in image sequences. As a result the recorded person should be unrecognisable. We use an approach which pseudo-randomly scrambles pixels within specified regions. This technique is reversible and preserves the colour characteristic of each region. So, colour-based approaches will still be able to automatically distinguish between differently dressed individuals. The evaluations of our results show that the privacy aspect got a very high score in both objective and subjective metrics. Our approach has a lack of intelligibility since it was measured by applying the Histogram of Oriented Gradients which might be fail on scrambled areas since edges are not preserved.

Keywords

Privacy, Privacy Protection, Video Analytic, Image Processing, Filtering

1. INTRODUCTION

Video surveillance of public spaces is expanding. Consequently, individuals are increasingly concerned about the "invasiveness" of such ubiquitous surveillance and fear that their privacy is at risk. The demands of stakeholders to prevent criminal activities are often seen to be in conflict with the privacy requirements of individuals. The main challenge is to preserve the anonymity of the surveyed individuals and also to fulfil the stakeholders needs. The problem of privacy protection in video surveillance is concerned in this year's MediaEval Visual Privacy Task [1]. A typical way to protect privacy in images and videos is to apply techniques such as blurring or masking. Since these techniques are irreversible, scrambling is introduced in [2]: A transform-domain scrambling technique, where pixels in the respective regions are pseudo-randomly scrambled based on a secret key. Our approach is quite similar, but applied on the pixel domain to be independent in the encoder used.

2. METHODOLOGY

Our proposed privacy-protection approach consists of a scrambling module that obfuscates regions of detected persons. A human detector was not built since these regions were provided by the organisers.

These regions are then vertically partitioned into slices having a width of one pixel. Our scrambling module obfuscates these regions by shuffling these slices. So, an obfuscated region differs from its original version in a changed sequence of these slices.

For this scheme, we choose a column-wise swapping pixels to recognise background structures and action. As seen in Figure 1, the action of stealing can be coarsely recognised by the extended arm visible through the horizontal black line in the scrambled image.

The shuffle algorithm is based on a modified variant of the Fisher-Yates method [3] which generates "random" permutations. The original sequence c consisting of N slices is changed by swapping each i-th slice with the j-th slice, where j is defined by a pseudo-random number generator. The algorithm is implemented as depicted in Algorithm 1:

Algorithm 1 Modified Fisher-Yates shuffle

procedure SHUFFLE(i, j)
    $i \leftarrow 0$
    for $i < N$ **do**
        $j \leftarrow$ random number $\in i \leq j < N$
        $t \leftarrow c(i)$
        $c(i) \leftarrow c(j)$
        $c(j) \leftarrow t$
    end for
end procedure

So, the permutation of the slices of the detected regions is determined by the order generated by a pseudo-random sequence. The pseudo-random sequence is repeatable due to the characteristics of the pseudo-random number generator. This generator produces a random, but repeatable sequence of integer numbers by specifying a certain, but fixed seed. This seed is generated from the hash value of a chosen password. This value is fixed for all regions in each frame and video sequence. Since the pseudo-random sequence is repeatable thru the given seed, the permutation of slices is reversible. So, the scrambled image regions can be recovered by knowing the password and the coordinates of each region. Overlapping regions were united before scrambling to ensure its reversibility. We choose for scrambling instead of cryptography to be robust against image compression artefacts and transmission errors. If these errors occur parts of the image will be also distorted, but these errors are not propagated.

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

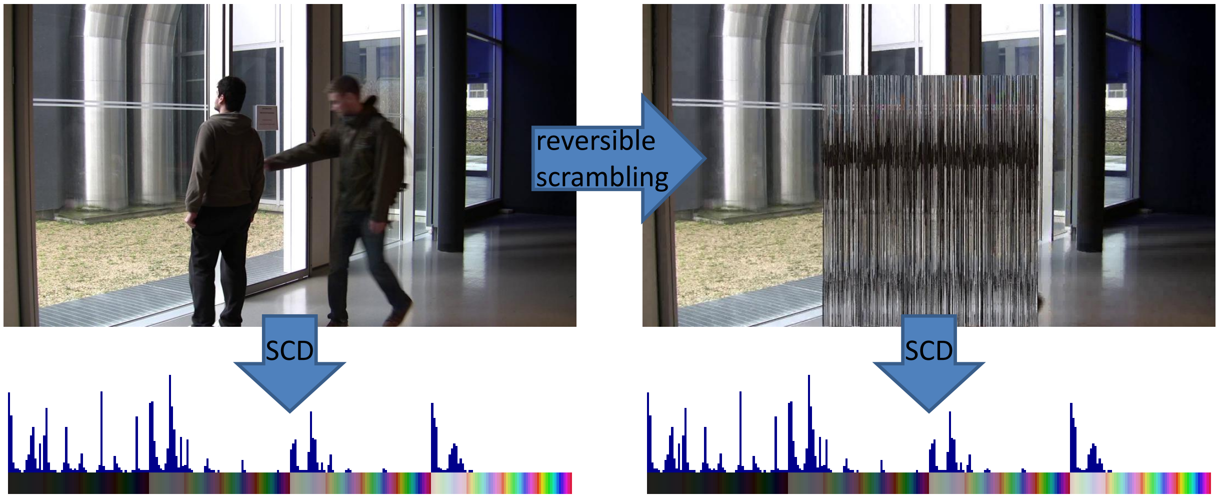

Figure 1: Original frame and its scrambled version together with their colour histogram.

3. EXPERIMENTS

The video sequences of the VPT dataset [1] [4] are obscured by scrambling the areas of detected persons instead of detected faces or parts within. We are sure that individuals can be identified not only by their face but also their clothes or accessories. So, the individuals are anonymised at best and a colour-based cluster algorithm may also be able to group regions depicting the same person.

The evaluation of the obscured videos took place using objective and subjective procedures. The objective metrics like face detection, object tracking and person re-identification were compared for each sequence pair.

Table 1: Objective and subjective evaluation (average values of the task in brackets)

	Objective	Subjective
Intelligibility	0.3514 (0.5124)	0.1767 (0.6557)
Privacy	0.7991 (0.6649)	0.9721 (0.6838)
Appropriateness	0.4073 (0.5605)	0.1533 (0.4921)

Three metrics are used to evaluated the results: Appropriateness, intelligibility, and privacy.

Appropriateness stands for the influence of the obscuring filter on the human perception of the image distortion. The objective score is measured using SSIM and PSNR, here a score of ca. 0.41 is reached which is below the average. A reason might be that quite large regions are scrambled which affect the pixel-based calculation of SSIM and PSNR. The subjective score is based on aspects such as pleasantness, distraction, and user acceptance. Here the score of ca. 0.15 is quite low resulting from distraction of the users.

Intelligibility stands for the ability of tracking obscured persons. The objective score is measured by applying the Histogram of Oriented Gradients as a human detector. Since our approach scrambles the slices of detected regions an edge-based human detector is unlikely able to detect person on these scrambled slices. So, a sub-standard value is expected: 0.35. The subjective score is 0.18.

The privacy metric concerns about the identification of individuals through their faces or personal accessories. The objective score is measured by a re-identification score from a visual model [1] and a face detection using the Viola-Jones algorithm from OpenCV. A score of ca. 0.80 is achieved which is above the average due the scrambled person regions. The subjective score is 0.97 even higher which is also above the average. A high subjective score was excepted, since it is very hard for the human eye to recognise a body shape or even a specific individual on a scrambled image.

Table 1 provides the evaluation scores for these three metrics.

4. CONCLUSION

We propose a reversible approach for scrambling regions within images or videos to obscure its content. This approach ensures a maximum of privacy, but fails in the other aspects, like appropriateness and intelligibility. In our mind, coarse action within these scrambled regions are still recognisable as depicted in Figure 1. The clue is that these regions can be de-scrambled, if the coordinates of the regions and the password which generated the seed for the pseudo-random number generation are known.

5. ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Community's FP7 under grant agreement number FP7-261743 (VideoSense). We would also like to thank Martha Larson for organizing this event.

6. REFERENCES

[1] A. Badii, M. Einig, and T. Piatrik. Overview of the MediaEval 2013 Visual Privacy Task. In *MediaEval 2013 Workshop*, Barcelona, Spain, October 18-19 2013.

[2] F. Dufaux and T. Ebrahimi. Video surveillance using jpeg 2000. In *Optical Science and Technology, the SPIE 49th Annual Meeting*, pages 268–275. International Society for Optics and Photonics, 2004.

[3] R. Durstenfeld. Algorithm 235: Random permutation. *Commun. ACM*, 7(7):420–, July 1964.

[4] P. Korshunov and T. Ebrahimi. PEViD: privacy evaluation video dataset. *Applications of Digital Image Processing XXXVI*, 25-29 August 2013.