Crowdsourcing-based Evaluation of Privacy in HDR Images

Pavel Korshunov[a], Hiromi Nemoto[a], Athanassios Skodras[b], and Touradj Ebrahimi[a]

[a]Multimedia Signal Processing Group, EPFL, Switzerland;
[b]DECE, University of Patras, Greece;

ABSTRACT

The ability of High Dynamic Range imaging (HDRi) to capture details in high-contrast environments, making both dark and bright regions clearly visible, has a strong implication on privacy. However, the extent to which HDRi affects privacy when it is used instead of typical Standard Dynamic Range imaging (SDRi) is not yet clear. In this paper, we investigate the effect of HDRi on privacy via crowdsourcing evaluation using the Microworkers platform. Due to the lack of HDRi standard privacy evaluation dataset, we have created such dataset containing people of varying gender, race, and age, shot indoor and outdoor and under large range of lighting conditions. We evaluate the tone-mapped versions of these images, obtained by several representative tone-mapping algorithms, using subjective privacy evaluation methodology. Evaluation was performed using crowdsourcing-based framework, because it is a popular and effective alternative to traditional lab-based assessment. The results of the experiments demonstrate a significant loss of privacy when even tone-mapped versions of HDR images are used compared to typical SDR images shot with a standard exposure.

Keywords: HDR imaging, privacy protection, crowdsourcing, evaluation

1. INTRODUCTION

The legal term "invasion of privacy" typically refers to a person's right to keep his or her life private and free from the intrusion of others. It is widely agreed that photography, as a device for invading privacy, can be extraordinarily intrusive. Digital photography, in particular, has led to a massive creation of new privacy issues for many people. It is estimated that the average Londoner is photographed 300 times a day by CCTV cameras*. Modern invasion of privacy laws essentially protect people in four different ways: intrusion of solitude, public disclosure of private facts, false light, and appropriation. The widespread use of image and video recording devices makes protection of privacy a challenging problem.

High Dynamic Range imaging (HDRi or HDR) represents a set of techniques used in imaging and photography to reproduce a greater dynamic range of luminosity than possible using standard digital imaging or photographic techniques. HDR images or video can represent more accurately the range of intensity levels found in real scenes, from direct sunlight to faint starlight. The most commonly used technique to capture HDR images is called 'image fusion' and involves taking several shots at different exposures, to capture the maximum dynamic range, then combining them into one image. This technique allows to create HDR with existing sensors leading to its increasing popularity, as it is already built into the smart phones software.[1]

HDRi allows higher detail and contrast that are missing from typical Standard Definition Range (SDR) images. This feature opens new perspectives in digital photography but at the same time brings along serious concerns in privacy intrusion issues. Private life activities or facts can be easier exposed now to common view due to equally high quality details in dark and bright image regions offered by the new technology.

Therefore, in this paper, we study the extent of privacy intrusion incurred by the HDR imaging as opposed to SDR imaging. We subjectively evaluate the effect of HDR on privacy using crowdsourcing approach. Crowdsourcing is a low cost and practical alternative to lab-based subjective evaluations, especially, for such subjective and environment independent issues as privacy.[2] We employ Microworkers † crowdsourcing platform and, since todays devices are not able to display HDRi directly, we compare several tone-mapped image versions with its

emails: pavel.korshunov@epfl.ch, hiromi.nemoto@epfl.ch, skodras@upatras.gr, touradj.ebrahimi@epfl.ch

*http://news.bbc.co.uk/2/hi/uk_news/2071496.stm

†http://microworkers.com/

typical SDR counterpart. To this end, we created a privacy-aware image dataset, which consists of 20 HDR images obtained by fusing 5 bracketed images with different exposures, depicting groups of people under highly variable lighting conditions (see Figure 1 for the samples), including deep shades, sunny outdoor and dark indoor, and lens sun flare.

We show to the crowdsourcing workers an SDR variant of each image (a zero level exposure in the camera used, Nikon D70) in the dataset and four tone-mapped versions obtained using the following representative tone-mapping operators: exponential (a simple global operator), by Drago *et al.*[3] (a complex global operator), by Reinhard and Devlin[4] (a complex local-based operator), and by Ward and Simmons[5] (a global operator adopted in JPEG-HDR format). To evaluate privacy intrusiveness of HDR images in our dataset, we ask 5 privacy-related questions about race, gender, details of clothing, and age. The answers to these questions can help us determine how much privacy-related details are visible and rank different variations of the image according to privacy intrusiveness.

Overall, the following are the main contributions of this paper:

- First to our knowledge, HDR privacy-aware image dataset;
- Subjective methodology and protocol for evaluation of privacy in HDR images;
- Crowdsourcing approach to evaluate privacy intrusiveness in images.

2. BACKGROUND AND RELATED WORK

2.1 Privacy

A lot of research effort in the past was focused on approaches to incorporate privacy protection into existing security systems and frameworks, typically via implementing access rights management and policies.[6–9] Another large body of work is on development of algorithms and methods to protect visual privacy, such as using watermarking to hide visual personal information,[10] scrambling techniques to reversibly distort privacy sensitive regions,[11] removal of unauthorized personnel from the video feed,[12] encoder independent geometrical-based reversible distortions.[13,14] However, little was done towards understanding of privacy issues in practical multimedia applications, analysis of the effect of privacy protection on the main system utility, and development of effective evaluation methodologies that take into account both the context and content.

Some work was done for the evaluation of privacy filters (tools that protect visual privacy) objectively. The objective evaluation of several primitive privacy filters was first performed by Newton *et al.*,[15] where the authors demonstrated that such filters cannot adequately protect from the successful face recognition, because recognition algorithms are robust. The robustness of face recognition and detection algorithms to primitive distortions is also reported in Ref. 16. In the work by Dufaux *et al.*,[17] a framework is defined to evaluate the performance of face recognition algorithms applied to images altered by various obfuscation methods. Another study[18] considered an automated video surveillance system and focused on finding the privacy-intelligibility tradeoff using objective metrics such as face detection and face recognition.

However, since privacy is a subjective notion, Ref. 19 argued that the evaluation should be done subjectively. The authors considered the problem of finding the balance between the ability of human guards to perform a surveillance task and adequate protection of privacy. A subjective methodology and protocol were defined for evaluation of privacy protection tools, focusing on two important aspects: (i) how much of the privacy is protected by such a tool and (ii) how much it impacts the efficiency of the underlying surveillance task (intelligibility). The pixelization filter showed the best performance in terms of balancing between privacy protection and allowing sufficient intelligibility. Masking filter, instead, demonstrated the highest privacy protection with low incorrectness and high uncertainty, which can be suitable for the higher security surveillance applications.

2.2 HDR imaging

High Dynamic Range (HDR) imaging is expected, together with ultra-high definition (UHD) and high frame rate (HFR) video, to become a technology that may change photo, TV, and film industries. Many different subjective evaluations have been previously performed to compare different tone-mapping operators for HDR images and video. Main focus of these studies was either on determining a more superior approach to tone-mapping or establishing an evaluation methodology for subjective evaluation of HDR content. One of the first subjective evaluations of HDR images was performed by Ledda *et al.*[20] The authors used paired comparison to evaluate the perceptual quality of six different tone-mapping algorithms. An HDR display was used as reference display for 48 subjects. The focus of this work was on the evaluation methodology for the subjective comparison of HDR images in a controlled environment. The evaluations provided the performance ranking of different tone-mapping algorithms leading to different perceptual qualities in color and gray images. Similar studies were conducted to determine the appeal of HDRi,[21] usefulness of HDR for astronomical images,[22] how accurately tone-mapping algorithms represent reality,[23] on objective metrics of HDR,[24] and on using HDR for 3D content.[25]

However, as opposed to the above work, the goal of our study is not to find the best tone-mapping algorithm but to demonstrate the invasive nature of HDR imaging, which threatens privacy of people. We therefore need an effective evaluation methodology and tool that would allow us to test an image on privacy intrusiveness subjectively in a practical environment.

2.3 Crowdsourcing

Crowdsourcing is a specific instance of the outsourcing approach, when a larger task is split into a set of small independent tasks that can be performed online in a short time without long-term commitment. Typically, such tasks are repetitive and are usually grouped in larger units, referred to as campaigns. Social networks can be a possible source of workers, which are usually unpaid in this case, for performing the crowdsourcing tasks. This alternative was explored for evaluation of privacy tradeoff in video surveillance[2] and proved being a viable option for such assessments. An online application VideoRate has been built which uses Facebook ID for login and a more reliable user authentication. Compared to laboratory based evaluations, this approach provides some flexibility to users, as one could stop participation in the experiments at any time, as well as it better simulates the real-time scenario, since the users evaluate videos in different lighting conditions and using monitors with different resolutions and color settings. The call to participate was disseminated in the subjective test using the VideoRate application via such social networks as Facebook, Twitter, and LinkedIn, as well as various research mailing lists. With an estimated outreach to more than 1500, some 120 among them used the application and submitted subjective scores. These subjective results were then compared with the results from a similar evaluation conducted by a conventional approach in a designated research test laboratory at EPFL. The results demonstrate a high correlation with only minor differences favoring the crowdsourcing method, which means that it can be considered as a reliable and effective approach for subjective evaluation of visual privacy filters.

A more conventional crowdsourcing approach is to use a specialized mediator, an online crowdsourcing platform, which allows setting up and distributing tasks, manages the workers for the campaigns, handles payments, etc. Notable examples of such platforms are Amazon's Mechanical Turk (MTurk)[‡] and Microworkers. MTurk is the largest crowdsourcing platform and is often used in research, as well as in commercial third-party applications; however, it allows only US residents or companies to submit tasks to the platform. Microworkers, on the other hand, allows international employers and it also provides employers the ability to choose the country of origin of workers. In this paper, we employ Microworkers as the commercial crowdsourcing platform, as oppose to the social networks,[2] and evaluate different tone-mapped versions of HDR images testing their privacy intrusiveness compared to typical SDR images. To display images to different workers provided by Microworkers and to collect evaluation results, we used the *QualityCrowd* framework.[26] It is an open-source platform designed for QoE evaluation with crowdsourcing. We choose this framework, because it is easy to modify for our privacy evaluation task using the provided simple scripting language for creating campaigns, training sessions, and control questions.

[‡]https://www.mturk.com/mturk/

3. EVALUATION FRAMEWORK

To evaluate privacy intrusiveness in HDR images, one needs a dataset of such containing various people with a visible privacy-related information. There are many datasets for evaluation of video analytics (detection, recognition, and tracking algorithms), such as FERET dataset§ and Labeled Faces in the Wild (LWF)¶, PETS 2007‖, etc. There are also some of the HDR dataset of images and video commonly used for evaluation of tone-mapping algorithms or appeal of HDRi compared to SDRi. But since these datasets were not designed with privacy issues in mind, they are not suitable for evaluation of privacy related aspects. To the authors knowledge, only one dataset for evaluation of privacy exists, called PEViD,[27] but it is not an HDR dataset and therefore cannot be used for this evaluation.

3.1 Dataset

The privacy dataset of HDR images should emphasize two main aspects:

- It should demonstrate the variety of privacy-related issues and notions, such as gender, age, race, and personal identifiable items.
- It should bring out the specifics of HDR imaging, i.e., images dynamic range should be high enough for HDR to manifest itself.

Following the above principles, we have created am HDR image dataset for privacy evaluation (see Figure 1 for some sample images), which contains 20 HDR images of 6000×4000 resolution. The images were shot with Nikon D70 camera using 5 bracketed images with different exposures ($-2, -1, 0, 1, 2$ exposure settings of the camera). The images depict several (from 3 to 15) people standing in normal poses and mostly facing the camera in different contrasting lighting conditions. The following conditions and scenarios were used:

- Camera in sunny outdoor is shooting inside dark indoor
- Camera inside dark indoor is shooting outside sunny outdoor
- Camera is in sunny outdoor shoots towards a dark shade
- Camera is facing a sunset

When creating the dataset, the aim was to have deep shades, sunny outdoor and dark indoor, lens sun flare, etc.

For crowdsourcing evaluations, to accommodate the typical display sizes of workers, all images were resized to 800×600 resolutions.

3.2 Methodology

Since people on crowdsourcing platform do not own HDR monitor, their displays are only able to reproduce SDR images. Therefore, beside an SDR variant of each image, corresponding to a zero level exposure in the camera Nikon D70, four tone-mapped versions were also shown to the crowdsourcing workers. The following four representative tone-mapping operators were used: exponential (a simple global operator), by Drago *et al.*[3] (a complex global operator), by Reinhard and Devlin[4] (a complex local-based operator), and by Ward and Simmons[5] (a global operator adopted in JPEG-HDR format).

To evaluate privacy intrusiveness of HDR images in our dataset, for each image, we ask workers the following five questions:

- How many people do you see in the picture?

§http://www.itl.nist.gov/iad/humanid/feret/feret_master.html
¶http://vis-www.cs.umass.edu/lfw/
‖http://pets2007.net/

(a) SDR version

(b) Exponential TMO

(c) SDR version

(d) Reinhard TMO

Figure 1: Examples of images from the dataset. The top images are more 'true' HDR image showing the higher dynamic range compared to the bottom images.

- What is the GENDER of the person inside red box?
- What is the RACE of the person inside red box?
- What is the AGE of the person inside red box?
- What is the COLOR of T-SHIRT/TOP/DRESS of the person inside red box?

The answers to these questions can help us determine how much privacy-related details are visible and rank different variations of the image according to privacy intrusiveness.

Special care was taken to make sure that a particular worker does not see twice the same person appearing in different images of the dataset. For this purpose, the dataset was split in two subsets with 10 images each shown in such order that a person shown in the red box (about whom the questions are asked) does not appear in the previous images. Therefore, a worker always answers questions about the person in the red box that did not appear before. This insures that workers answers are not influenced by the previous knowledge and makes the subjective results reliable and fair. For the same reason, each campaign has one version of content (SDR and four tone-mapped) appearing only once. And a single worker was not allowed to evaluate more than one campaign. Having two subsets of images and five versions of each content resulted in 10 different campaigns of 10 images.

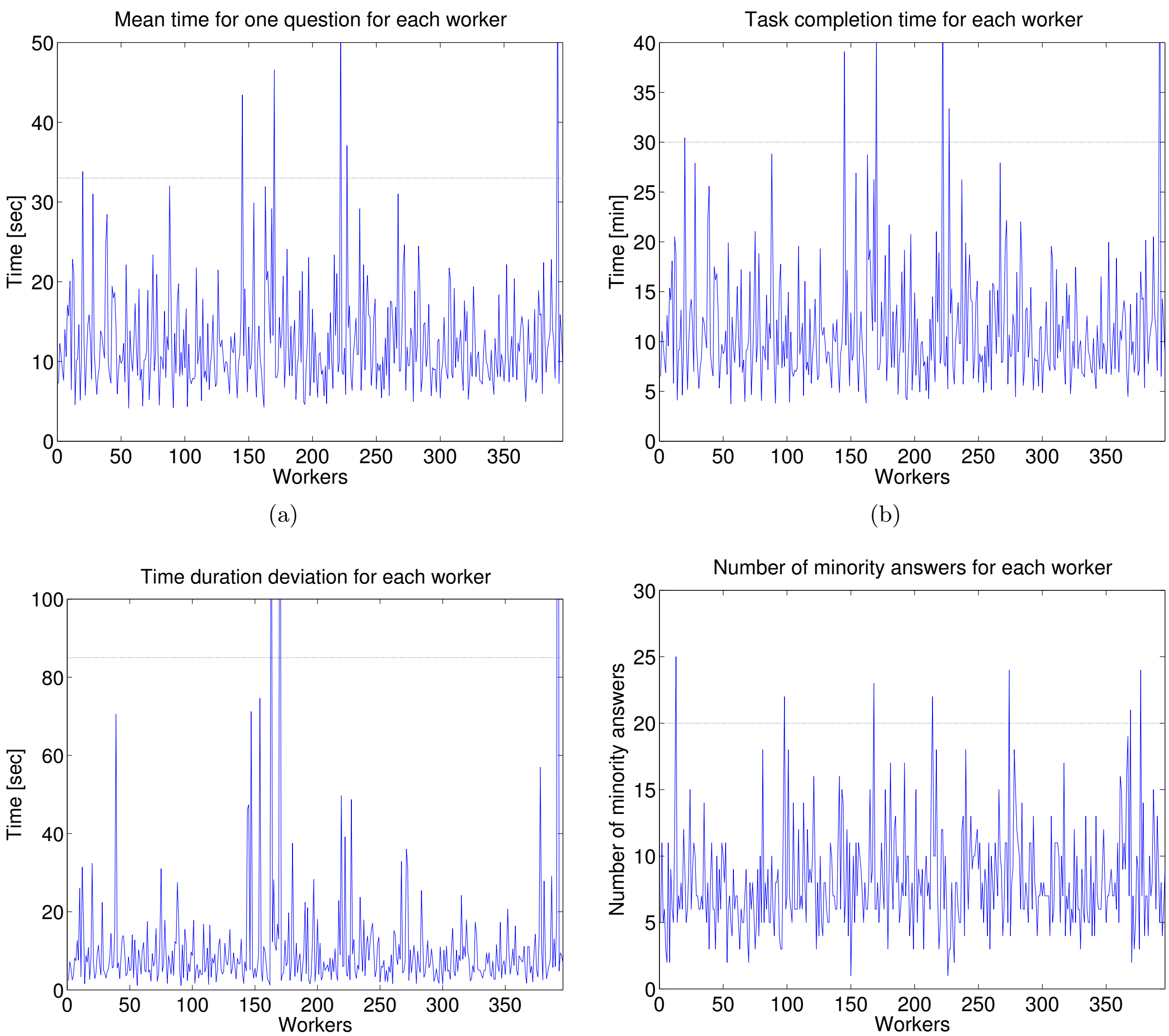
The evaluation required a relatively lager amount of workers, because a worker could not evaluated more than one campaign. Since the largest number of workers on Microworkers come from Asia, this region was selected for the evaluation. Only countries where English is a dominant language were chosen, with either more than 50% of population or more than 10 mil of people speaking English, according to the Wikipedia.

Each campaign started with a message of what is required from the worker and a training session explaining the evaluation process. Also, a brightness test was performed for each worker using method similar to the one described in Ref.28.

4. EVALUATION RESULTS

In total, 396 workers participated in the evaluation experiment. The answers were first checked for the reliability and only those that passed the filter were used in the result analysis.

4.1 Reliable workers detection



(a) (b) (c) (d)

Figure 2: Different methods of detecting unreliable workers.

Unlike lab-based subjective experiment where all subjects can be observed by experiment operators and its test environment also can be controlled, the major shortcoming of the crowdsourcing-based subjective evaluation is the inability to supervise participants behavior and to restrict their test conditions. When using crowdsourcing for evaluation, there is a risk of including untrusted data into analysis due to the wrong test conditions or unreliable behavior of some workers who try to submit low quality work in order to reduce their effort while maximizing their received payment.[28] For this reason, unreliable workers detection is an inevitable process in crowdsourcing-based subjective evaluation. To identify a worker as 'trustworthy', the following four factors were used in our experiment:

- Task completion time;
- Mean observation time per question;
- Observation duration deviation;
- Number of minority answers.

The objective of the first three factors is to filter out the workers who have strange behaviors in the middle of their task, because they are either not serious or have poor concentration. The observation time per question is measured as the time from when the question is displayed until the time the answer is given by the worker. The task completion time, mean observation time and observation duration deviation can be calculated using this data. If the task completion time or mean observation time per question is too long compared to their averages of all workers, it can be deduced that they did not take the test seriously or were distracted during their tasks. The task completion time and mean response time per question for each worker are given in Figures 2a and 2b respectively. As indicated by the figures, some workers demonstrate large such values compared to the corresponding mean. In order to filter out unreliable workers, the participants with three times larger than the standard deviation of task completion time or mean observation time per question were excluded. These values were chosen to account for about 99% ordinary workers with the assumption that the distribution of these numbers follows the normal distribution, since the task was simple.

The observation duration deviation for each worker is shown in Figure 2c. This number shows how consistent a worker is in answering the questions, hence, we excluded the workers with large deviation number, since it means they did not concentrate on performing the task consistently. The value which is three times larger than the standard deviation of all workers was used to remove the untrustworthy participants.

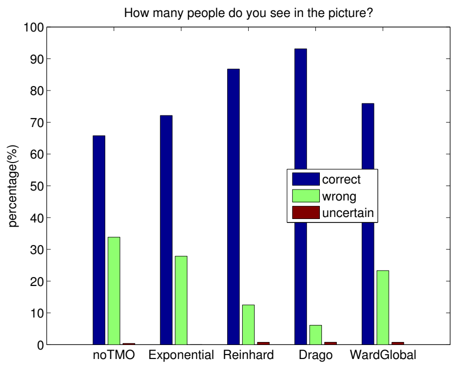
Finally, unreliable workers were also identified using an approach similar to a typical outlier detection method, commonly used in most subjective quality evaluations. However, typical subjective tests use scoring methods like five-grade evaluation, and outlier detection is performed on mean opinion score.[29] Our experiments do not have opinion scores because of the specific privacy-oriented questions we used. Therefore the number of minority answers has been used for the outlier detection instead (see Figure 2d). The assumption is that a participant who has a lot of different answers compared to the majority of workers is unreliable. The threshold was determined in the same way as for the other three factors.

The above unreliable worker detection methods filtered out 14 workers out of total 396, resulting in 382 scores used in the following analysis.

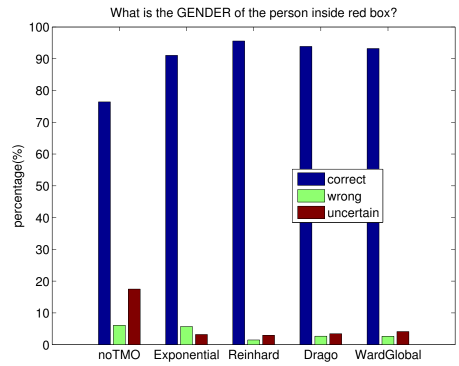
4.2 Evaluation results

Since not all HDR images actually have high dynamic range of luminance, we divided all images of the dataset into two clusters: (i) images that while being HDR have range of luminance similar to SDR, termed *SDR-like HDR images* (see Figures 1c,1d) and (ii) HDR images, termed *true HDR images* (see Figures 1a,1b), that have luminance range significantly larger than SDR. This division was done to demonstrate the difference in the level of privacy intrusiveness when using true HDR images and HDR images exhibiting dynamic range similar to a 'normal' SDR images.

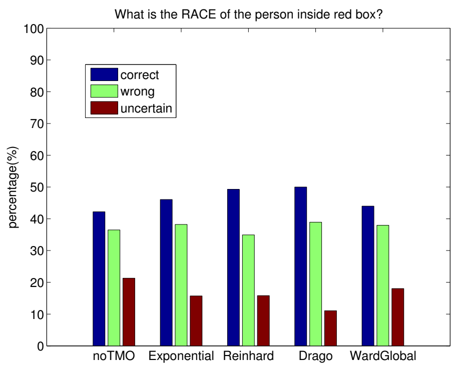
Figure 3 and Figure 4 show the averaged answers to the questions given in Section 3.2 for the two clusters of images. From the figures, it is clear that the higher number of workers answered questions correctly for tone-mapped versions of the true HDR images compared to SDR version with no tone-mapping, which is not the case for SDR-like HDR images, where all versions yield the same scores. Also, true HDR images were more



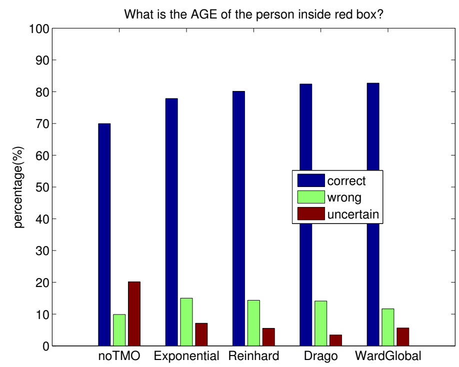
(a)



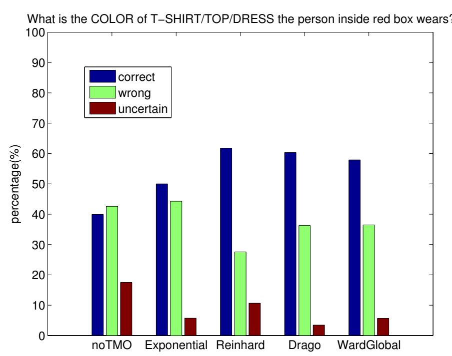
(b)



(c)



(d)



(e)

Figure 3: Subjective scores for true HDR images. SDR (noTMO) and HDR images tone-mapped with different algorithms were evaluated.

confusing for workers overall with correct scores for all such images being lower and uncertainty higher than for SDR-like HDR images. Such confusion was due to many dark or over-saturated bright areas with low visibility of the true HDR images.

When considering only SDR-like HDR images, we can still notice that questions about Race (Figure 4c), Age (Figure 4d), and Color (Figure 4e) result in a number of wrong answers. The main reason for these is the ambiguity of the choices we gave to workers as the answers to these questions. The answers to these kind of questions are subjective and may depend of such factors like culture of the worker, origin, social background, etc. For instance, the answers to the question about race included 'White', 'Asian', 'Other', and 'I don't know'. And several people in the images were from South America with darker skin, which the workers, being from

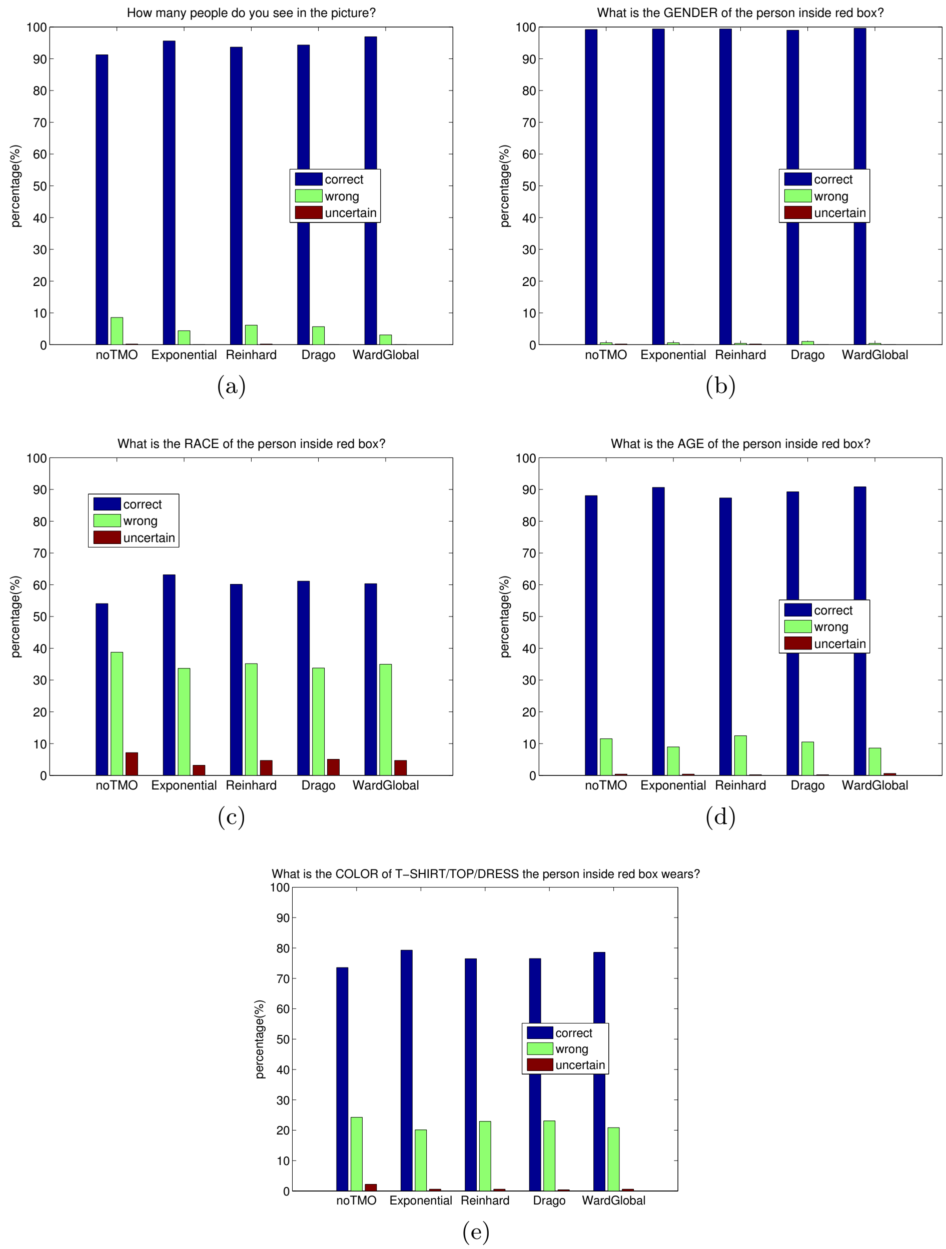


Figure 4: Subjective scores for the set of SDR-like HDR images. SDR (noTMO) and HDR images tone-mapped with different algorithms were evaluated.

the Asian region, assumed to be also Asian origin, leading to a few wrong answers. However, these differences in interpretation of the answers are not important. What is important is the different in how people answers the same questions for different versions of the image contents when comparing true HDR images and SDR-like HDR images.

Therefore, the main conclusion we can draw from Figure 3 and Figure 4 is that the tone-mapped versions of the true HDR images show more privacy details than SDR versions, as the number of correct answers for tone-mapped versions is higher and the uncertainty is lower. Some tone-mapping algorithms, such as Drago, also show more visible details, and hence they are more privacy intrusive, than others.

5. CONCLUSION

This paper evaluated the privacy intrusiveness of HDR imaging via crowdsourcing evaluation of differently tone-mapped images vs. a typical SDR. For that purpose, a privacy-aware dataset of 20 HDR images of different groups of people were created using image fusion technique. The dataset contains both 'true' HDR images with high dynamic range and SDR-like images with luminance range similar to a 'typical' SDR image. Using the proposed evaluation methodology of privacy intrusiveness, we conducted crowdsourcing experiments employing 396 workers from Microworkers crowdsourcing platform. The results of workers, checked for reliability, demonstrated that images tone-mapped from HDR are more intrusive than SDR versions.

Since crowdsourcing is still not considered as reliable as the controlled experiments done in a laboratory environment that is designed for the subjective evaluation, the findings of this paper need to be confirmed by the lab-based assessments. Also, displaying tone-mapped version on a typical monitor is a poor substitution to the actual HDR image displayed on an HDR monitor. Therefore, privacy evaluation experiments need to be repeated by comparing actual HDR content with SDR version on a real HDR monitor.

ACKNOWLEDGMENTS

This work was conducted in the framework of COST Action IC1003 – Qualinet with partial support from the EC funded Network of Excellence VideoSense. Special thanks to Lukáš Krasula and Dragana Đorđević for their valuable discussions and help in the dataset creation.

REFERENCES

[1] Mantiuk, R., Cichowicz, M., and Smyk, M., "Implementation of HDR photographic pipeline in mobile devices," *Image Analysis and Recognition* **7324**, 367–374 (2012).

[2] Korshunov, P., Cai, S., and Ebrahimi, T., "Crowdsourcing approach for evaluation of privacy filters in video surveillance," in [*Proceedings of the ACM Multimedia 2012 Workshop on Crowdsourcing for Multimedia*], *CrowdMM'12*, 35–40 (Oct. 2012).

[3] Drago, F., Myszkowski, K., Annen, T., and Chiba, N., "Adaptive logarithmic mapping for displaying high contrast scenes," in [*Proc. of EUROGRAPHICS 2003*], Brunet, P. and Fellner, D. W., eds., *Computer Graphics Forum* **22**(3), 419–426, Blackwell, Granada, Spain (2003).

[4] Reinhard, E. and Devlin, K., "Dynamic range reduction inspired by photoreceptor physiology," *IEEE Transactions on Visualization and Computer Graphics* **11**, 13–24 (Jan. 2005).

[5] Ward, G. and Simmons, M., "JPEG-HDR: a backwards-compatible, high dynamic range extension to JPEG," in [*ACM SIGGRAPH 2005 Courses*], *SIGGRAPH'05*, ACM, New York, NY, USA (Aug. 2005).

[6] Ebrahimi, T., Abdeljaoued, Y., Ventura, R. F. I., and Escoda, O. D., "MPEG-7 camera," in [*Proceedings of International Conference on Image Processing (ICIP 2001)*], **3**, 600 –603 (2001).

[7] Senior, A., Pankanti, S., Hampapur, A., Brown, L., Tian, Y.-L., Ekin, A., Connell, J., Shu, C., and Lu, M., "Enabling video privacy through computer vision," *IEEE Security and Privacy* **3**, 50–57 (May 2005).

[8] Park, S.-W., Han, J. W., and Shin, S.-U., "Secure service mechanism of video surveillance system based on h.264/svc," in [*International Conference on Information Technology and Multimedia (ICIM 2011)*], 1 –4 (Nov 2011).

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

[10] Zhang, W., Cheung, S., and Chen, M., "Hiding privacy information in video surveillance system," in [*Proc. IEEE International Conference on Image Processing*], (Sep 2005).

[11] Dufaux, F. and Ebrahimi, T., "Scrambling for privacy protection in video surveillance systems," *IEEE Trans. on Circuits and Systems for Video Technology* **18**, 1168–1174 (Aug 2008).

[12] Wickramasuriya, J., Datt, M., Mehrotra, S., and Venkatasubramanian, N., "Privacy protecting data collection in media spaces," in [*Proceedings of the 12th annual ACM international conference on Multimedia*], 48–55 (Oct. 2004).

[13] Korshunov, P. and Ebrahimi, T., "Using warping for privacy protection in video surveillance," in [*18th International Conference on Digital Signal Processing (DSP)*], 1–6 (2013).

[14] Korshunov, P. and Ebrahimi, T., "Using face morphing to protect privacy," in [*IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS)*], 208 – 213 (Aug. 2013).

[15] Newton, E., Sweeney, L., and Malin, B., "Preserving privacy by de-identifying face images," *IEEE Trans. on Knowledge and Data Engineering* **17**, 232–243 (Feb 2005).

[16] Korshunov, P. and Ooi, W. T., "Video quality for face detection, recognition, and tracking," *ACM Trans. Multimedia Comput. Commun. Appl.* **7**, 14:1–14:21 (Sept. 2011).

[17] Dufaux, F. and Ebrahimi, T., "A framework for the validation of privacy protection solutions in video surveillance," in [*Proceedings of IEEE International Conference on Multimedia & Expo (ICME 2010)*], (July 2010).

[18] Korshunov, P., Melle, A., Dugelay, J.-L., and Ebrahimi, T., "A framework for objective evaluation of privacy filters in video surveillance," in [*SPIE Applications of Digital Image Processing XXXVI*], **8856** (Aug. 2013).

[19] Korshunov, P., Araimo, C., De Simone, F., Velardo, C., Dugelay, J., and Ebrahimi, T., "Subjective study of privacy filters in video surveillance," in [*2012 IEEE 14th International Workshop on Multimedia Signal Processing (MMSP)*], 378–382 (Sept. 2012).

[20] Ledda, P., Chalmers, A., Troscianko, T., and Seetzen, H., "Evaluation of tone mapping operators using a high dynamic range display," in [*ACM SIGGRAPH 2005 Papers*], *SIGGRAPH'05*, 640–648, ACM, New York, NY, USA (Aug. 2005).

[21] Yoshida, A., Blanz, V., Myszkowski, K., and Seidel, H.-P., "Perceptual evaluation of tone mapping operators with real-world scenes," **5666**, 192–203, Human Vision and Electronic Imaging X, IS&T/SPIE's 17th Annual Symposium on Electronic Imaging (2005).

[22] Park, S. H. and Montag, E. D., "Evaluating tone mapping algorithms for rendering non-pictorial (scientific) high-dynamic-range images," *J. Vis. Comun. Image Represent.* **18**(5), 415–428 (2007).

[23] Kuang, J., Yamaguchi, H., Liu, C., Johnson, G. M., and Fairchild, M. D., "Evaluating HDR rendering algorithms," *ACM Trans. Appl. Percept.* **4**, 9:1–9:27 (July 2007).

[24] Annighöfer, B., Tajbakhsh, T., and Grigat, R.-R., "Prediction of results from subjective evaluation of real-time-capable tone-mapping operators applied to limited high-dynamic-range images," *Journal of Electronic Imaging* **19**, 011015–011015–12 (Jan. 2010).

[25] Mai, Z., Doutre, C., Nasiopoulos, P., and Ward, R., "Subjective evaluation of tone-mapping methods on 3D images," in [*17th International Conference on Digital Signal Processing (DSP)*], 1–6 (July 2011).

[26] Keimel, C., Habigt, J., Horch, C., and Diepold, K., "Qualitycrowd - a framework for crowd-based quality evaluation," in [*Picture Coding Symposium (PCS), 2012*], 245–248 (May 2012).

[27] Korshunov, P. and Ebrahimi, T., "PEViD: privacy evaluation video dataset," in [*SPIE Applications of Digital Image Processing XXXVI*], **8856** (Aug. 2013).

[28] Hossfeld, T., Keimel, C., Hirth, M., Gardlo, B., Habigt, J., Diepold, K., and Tran-Gia, P., "Best practices for QoE crowdtesting: QoE assessment with crowdsourcing," *IEEE Transactions on Multimedia* **PP**(99), 1–1 (2013).

[29] ITU-R BT.500-13, R., [*Methodology for the subjective assessment of the quality of television pictures*], International Telecommunication Union, Geneva, Switzerland (2012).