PRELIMINARY EVALUATION OF HDR TONE MAPPING OPERATORS FOR CULTURAL HERITAGE

Rossella Suma[a, *], Georgia Stavropoulou[b], Elisavet K. Stathopoulou[c], Luc van Gool[b], Andreas Georgopoulos[c], Alan Chalmers[a]

[a] WMG, University of Warwick, CV4 7AL, Coventry, UK. r.suma@warwick.ac.uk; alan.chalmers@warwick.ac.uk

[b] ESAT - PSI, VISICS, KU Leuven, Kasteelpark Arenberg 10, 3001 Heverlee, Belgium. georgina.stavropoulou@esat.kuleuven.be; luc.vangool@kuleuven.be

[c] Laboratory of Photogrammetry, National Technical University of Athens, Heroon Polytechneiou 9, 15780, Athens, Greece. elliestath@central.ntua.gr; drag@central.ntua.gr

Abstract:

The ability of High Dynamic Range (HDR) imaging to capture the full range of lighting in a scene has led to an increasing interest in its use for Cultural Heritage (CH) applications. Photogrammetric techniques allow the semi-automatic production of 3D models from a sequence of images. Current photogrammetric methods are not always effective in reconstructing objects under harsh lighting conditions, as significant geometric details may not have been captured accurately in under- and over-exposed regions of the images. HDR imaging offers the possibility to overcome this limitation. In this paper we evaluate four different HDR tone-mapping operators (TMOs) that have been used to convert raw HDR images into a format suitable for state-of-the-art photogrammetric algorithms, and in particular keypoint detection techniques. The evaluation criteria used are the number of keypoints and the number of valid matches achieved. The comparison considers two local and two global TMOs.

Key words: high dynamic range imaging, HDR tone mapping, features detection, 3D reconstruction

1. Introduction

High Dynamic Range (HDR) is an imaging technique that enables the acquisition, storage and display of a wider range of luminance values than normal cameras allow (Banterle *et al.* 2011). The ability of HDR to capture all the detail in a scene, even in cases of harsh lighting conditions, makes it a very useful method for providing robust data for photogrammetric reconstruction and photorealistic texturing. This is especially true with Cultural Heritage (CH) environments, which are often characterised by highly reflective or shiny materials and intense presence of shadows and bright areas. Using HDR imaging for generating the input of photogrammetric processes such as Structure from Motion (SfM) could significantly improve the final results. Focusing on the keypoint detection, in cases of Low Dynamic Range (LDR) images with strongly shadowed or very bright areas, the detected keypoints tend to cluster only in the well-exposed areas of the image. This non-uniform spatial distribution of points, can increase, as a result, the image registration error and potentially compromise the stability of the reconstructed geometry. HDR imaging is capable of overcoming these limitations and therefore represents a more reliable input for the SfM pipeline.

One problem with providing HDR input for photogrammetric purposes is that most 3D reconstruction algorithms have been designed to work with traditional 8-bit LDR images. Therefore HDR images need first to be converted to LDR equivalents, a process commonly achieved by applying Tone mapping operators (TMOs). Various TMOs have been developed in the latest years, generally divided into global and local, depending on their working principle.

Although researchers have suggested the use of HDR images as a possible enhancement in the photogrammetric pipeline, especially for archaeological applications (Wheatley 2011); (Ntregka *et al.* 2014); (Guidi *et al.* 2014), HDR has not been frequently used in CH and photogrammetry up until now. Kontogianni *et al.* (2015) conducted a comparative study of feature detectors on tone mapped images with respect to the number of detected points but only one tone mapper was applied. Přibyl *et al.* (2016) evaluated the suitability of original Low Dynamic Range (LDR), native HDR and tone mapped HDR images with feature point detectors. An important aspect of tone mapping research field is that not all tone-mapping techniques have been developed for the same purpose. With the aim of documentation, reuse and processing using photogrammetry and computer vision techniques, it is

---

[*] Corresponding Author: Rossella Suma, r.suma@warwick.ac.uk

important that the TMO preserves reality-such as radiometry, while also being efficient for the applications.

In this paper we address the problems related to the initial steps of SfM with an in depth analysis on how feature detection is affected when applied on images that have been tone-mapped with four different TMOs: the Gradient Domain HDR Compression (Fattal *et al.* 2002), the local and global version of Reinhard *et al.* (2002); and a video Tone Mapping Operator called Display Adaptive Tone Mapping (Mantiuk *et al.* 2008). Regarding the detection, three of the most popular feature detectors were used; Difference of Gaussians (DoG), a keypoint detection method used by SIFT (Lowe 1999), Fast Hessian approximation, a process integrated in the SURF package (Bay *et al.* 2006) and, finally, FAST, that is based on accelerated segment test (AST) (Rosten and Drummond 2006).

2. Methodology

The images were captured with the exposure bracketing technique in various cultural heritage sites with the use of DSLR cameras and tripods. Depending on the dynamic range of the scene, either five or seven different exposures were taken with the full frame camera. The RAW data was processed with the Matlab HDR Toolbox (Banterle *et al.* 2011), which merges the different exposures into an HDR image using Debevec and Malik's method (1997). These HDR images were subsequently tone mapped with each of the four TMOs under study using pfstools library (Mantiuk *et al.* 2007*).* The tone mappers operators applied are: Fattal *et al.* (2002), Reinhard *et al.* (2002) in its two variants and Mantiuk *et al.* (2008) from now on referred to as Fattal, ReinhardGlobal ReinhardLocal and Mantiuk. In particular Mantiuk video TMO was selected for our tests because photogrammetry requires a sequence of images, and as we are tone mapping this entire sequence, we wanted to investigate its performance in terms of brightness, colour appearance, and coherence preservation for the entire sequence.

The criteria under which the tested TMOs and the respective LDR images were evaluated in this paper are the number of detected keypoints and the number of matches achieved between stereo pairs. The number of detected keypoints in each tone-mapped image alone may be a misleading criterion when it comes to photogrammetry applications, as a large number of keypoints does not necessarily offer a noise-free and homogenous coverage of the image or good and sufficient matches between images. Therefore, visual inspection was also performed in order to evaluate the distribution of the keypoints as well as their density. In the next step, keypoint description and matching were performed in order to evaluate the TMOs based on the number of matches. For this test, only the SURF descriptor was applied so that the results would be comparable only in terms of detection. For the matching process, the Flann-based (Muja *et al.* 2009) matcher was used and the resulting matches were filtered in order to keep the best ones, based on Lowe's ratio test criterion (Lowe 2004).

3. Results

Figure 1 illustrates the average number of keypoints and how three different TMOs and the original LDR image perform under the three selected feature detection methods. As can be seen, all detectors perform better with Mantiuk TMO than with ReihardLocal and ReinhardGlobal and the middle exposed original LDR image. ReinhardLocal is marginally better than ReinhardGlobal. Fattal was not included in Figure 1 as it detects significantly more keypoints compared to any of the other 3 TMOs (700000 for FAST, 176451 for SIFT and 108183 for SURF). The excessive number of keypoints possibly stems from Fattal's ability to intensify the contrast in dark regions, enhancing previously not visible details and texture**.** However**,** a larger number of keypoints does not guarantee a better 3D reconstruction, and, specifically in the case of Fattal TMO, image noise may also have been enhanced, leading to erroneous detection of keypoints. The combination of FAST and Fattal supports this argument and can produce a number of keypoints that is higher than 5% of the image pixels (around 1 out of 20 pixels is detected as an interest point). The results were also inspected visually by plotting all the detected points on the images.
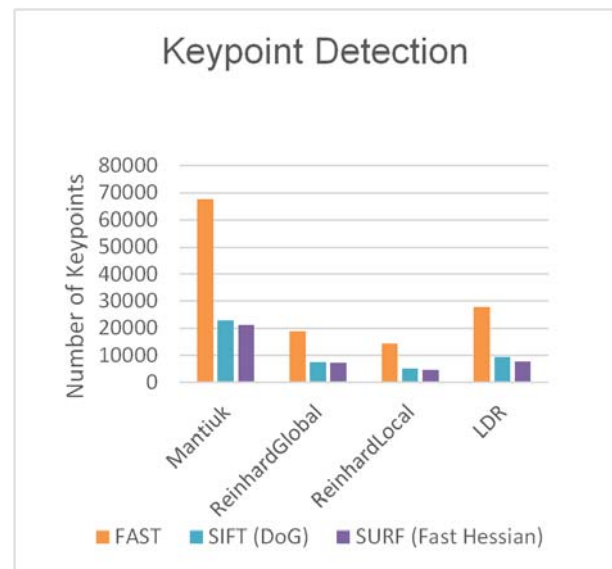

Figure 1: Comparison of the average number of keypoints detected with FAST, SIFT and SURF for Mantiuk, ReinhardGlobal, ReinhardLocal and LDR.

In Figure 2 only a detail of a fresco belonging to the Asinou Church in Cyprus is used to illustrate the density of the points and their distribution in the image space across different combinations of TMOs and detectors. In terms of density, the combination of FAST and Fattal produces an excessive amount of points that almost completely covers the image. Although less intense results are produced with the coupling of Fattal and the Fast Hessian (SURF) and DoG (SIFT) detectors, the density of points is still high even in regions where the fresco lacks in relevant details. Both Reinhard's methods produce images with low contrast and this is also reflected in the density of keypoints. Regarding the points' spatial distribution, apart from Fattal, the rest of the TMOs produce more points in highly textured areas.
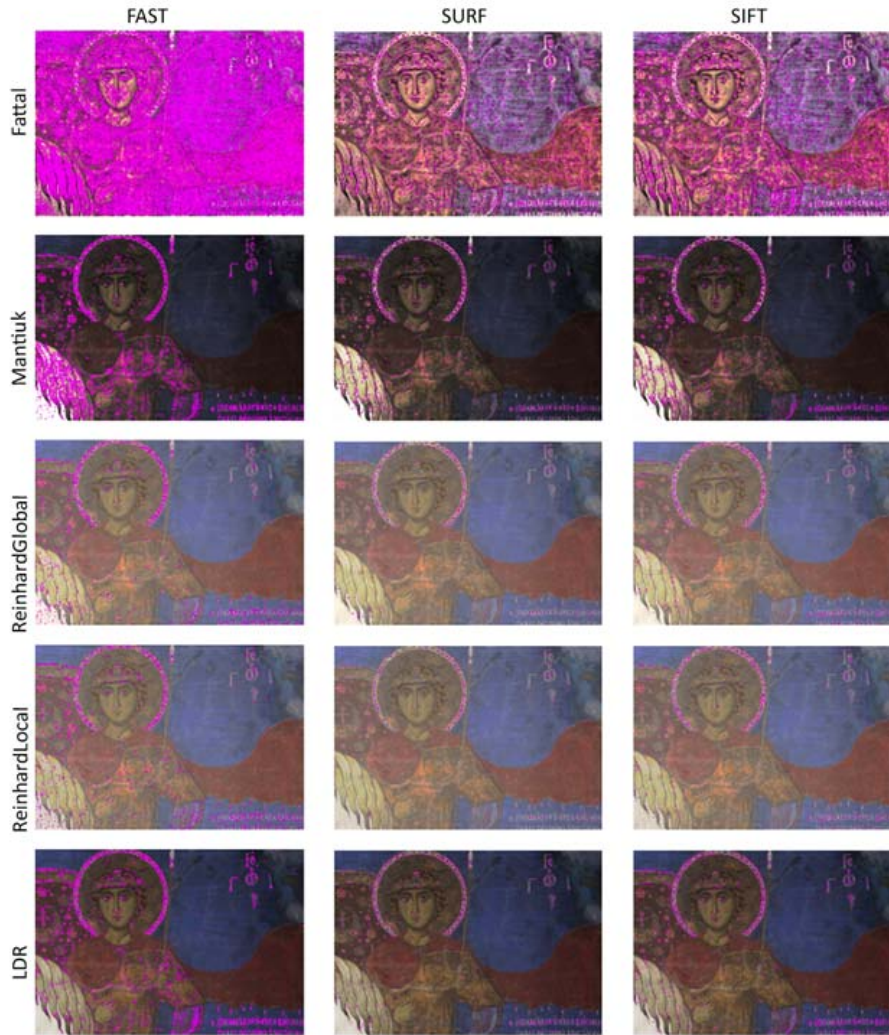
Figure 2: Particular of Asinou Church fresco with keypoints for all the TMOs and LDR tested with all three detectors.

Figure 3 shows the number of matches that have succesfully passed Lowe's ratio test (good matches). Again, Mantiuk TMO outperforms the other three, while the number of succefull matches is much bigger when the SURF (Fast Hessian) detector is being used. ReinhardGlobal and standard LDR images follow, with ReinhardGlobal performing slightly better, whereas ReinhardLocal has the lowest scores. The number of good matches can be considered as a more reliable criterion than the number of detected keypoints, as it reensures keypoint repeatability between sequential frames and subsequently produces more stable results for 3D reconstruction.

4. Conclusions

In this paper four different tone-mapping methods were compared in terms of their suitability for keypoint detection and matching. The results suggest that Mantiuk's method is a more suitable TMO for photogrammetric procedures since both the number of keypoints and the number of matches are high enough but not excessive. In the case of Fattal, in fact, the quantity of detected keypoints is higher but this does not imply better quality in the reconstruction, but rather computationally-intensive, time-consuming and prone-to-

error during image matching and registration. Regarding the Reinhard's methods, surprisingly, more points are detected in the original LDR images than in the tone-mapped images.
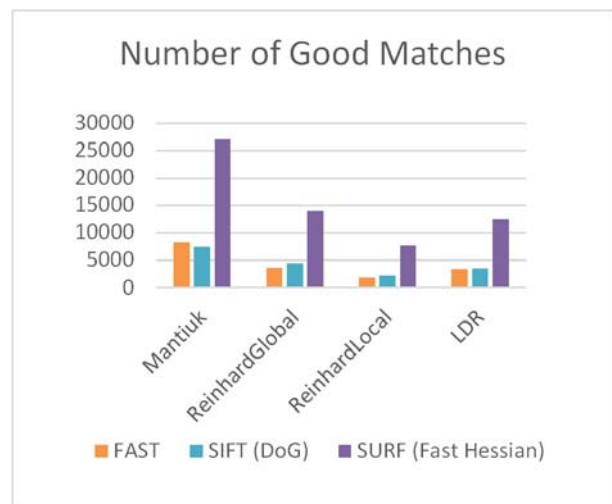

Figure 3: Comparison of number of good matches achieved on strereopairs tone-mapped with different TMOs.

This paper has shown that, despite the availability of a plethora of HDR TMOs in literature, not all of them are suitable for the specific task of cultural heritage site documentation and reconstruction. Moreover, most of the tone mapping algorithms have previously been evaluated based only on the visual appeal of their results. Therefore, our future work will keep exploring the potential of HDR TMOs specifically tailored for photogrammetric applications, and a more in depth and quantitative evaluation of the TMOs will be intended by including further criteria and more case studies, which demonstrate the lighting difficulties in cultural heritage environments. The future work will involve the use of a common criterion, such as the Repeatability Rate by Schmid *et al.* (2000) to evaluate the robustness of the different detectors over TMOs.

Acknowledgements

References

BANTERLE, F., ARTUSI, A., DEBATTISTA, K. and CHALMERS, A., 2011. Advanced high dynamic range imaging: theory and practice. CRC Press.

BAY, H., TUYTELAARS, T. and VAN GOOL, L., 2008. Surf: Speeded up robust features. Computer Vision and Image Understanding, 110(3), pp. 346–359.

DEBEVEC, P.E. and MALIK, J., 1997. Recovering high dynamic range radiance maps from photographs. In proceeding of the SPIE: Image Sensors (Vol. 3965, pp. 392-401).

FATTAL, R., LISCHINSKI, D. and WERMAN, M., 2002. Gradient domain high dynamic range compression. In ACM Transactions on Graphics (TOG), 21(3), ACM, pp. 249-256.

KONTOGIANNI, G., STATHOPOULOU, E.K., GEORGOPOULOS, A. and DOULAMIS, A., 2015. HDR imaging for feature detection on detailed architectural scenes. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 40(5), p. 325.

GUIDI, G., GONIZZI, S. and MICOLI, L.L., 2014. Image pre-processing for optimizing automated photogrammetry performances. ISPRS Annals of The Photogrammetry, Remote Sensing and Spatial Information Sciences, 2(5), p. 145.

LEDDA, P., CHALMERS, A., TROSCIANKO, T. and SEETZEN, H., 2005. Evaluation of tone mapping operators using a high dynamic range display. ACM Transactions on Graphics (TOG), 24(3), pp. 640-648.

LOWE, D.G., 2004. Distinctive image features from scale-invariant keypoints. International journal of computer vision, 60(2), pp. 91-110.

MANTIUK, R., DALY, S. and KEROFSKY, L., 2008. Display adaptive tone mapping. ACM Transactions on Graphics (TOG), 27(3), p. 68.

MANTIUK, R., KRAWCZYK, G., MANTIUK, R., and SEIDEL, H., 2007. High dynamic range imaging pipeline: Perception-motivated representation of visual content. Human Vision and Electronic Imaging XII. doi:10.1117/12.713526

MUJA, M. and LOWE, D.G., 2009. Fast approximate nearest neighbors with automatic algorithm configuration. VISAPP (1), 2, pp. 331-340.

NTREGKA, A., GEORGOPOULOS, A. and QUINTERO, M. S., 2014. Investigation on the use of HDR images for cultural heritage documentation. International Journal of Heritage in the Digital Era, 3(1), pp. 1-18.

PŘIBYL, B., CHALMERS, A., ZEMČÍK, P., HOOBERMAN, L. and ČADÍK, M., 2016. Evaluation of feature point detection in high dynamic range imagery. Journal of Visual Communication and Image Representation.

REINHARD, E., STARK, M., SHIRLEY, P. and FERWERDA, J., 2002. Photographic tone reproduction for digital images. In ACM Transactions on Graphics (TOG), 21(3), ACM, pp. 267-276.

ROSTEN, E. and DRUMMOND, T., 2006. Machine learning for high-speed corner detection. In Computer Vision–ECCV 2006, Springer Berlin Heidelberg, pp. 430-443.

SCHMID, C., MOHR, R. and BAUCKHAGE, C., 2000. Evaluation of interest point detectors. International Journal of computer vision, 37(2), pp. 151-172.

WHEATLEY, D., 2011. High dynamic range imaging for archaeological recording. Journal of Archaeological Method and Theory, 18(3), pp. 256-271.