HDR IMAGING FOR FEATURE DETECTION
ON DETAILED ARCHITECTURAL SCENES

G. Kontogianni, E. K. Stathopoulou*, A. Georgopoulos, A. Doulamis

Laboratory of Photogrammetry, School of Rural and Surveying Engineering, National Technical University of Athens –
15780 Zografou Athens, Greece
georgia.kontog@gmail.com - elliestath@central.ntua.gr - drag@central.ntua.gr - adoulam@cs.ntua.gr

Commission V, WG V/4

KEY WORDS: Feature detection, High dynamic range, Tone mapping, Architectural Scenes

ABSTRACT:

3D reconstruction relies on accurate detection, extraction, description and matching of image features. This is even truer for complex architectural scenes that pose needs for 3D models of high quality, without any loss of detail in geometry or color. Illumination conditions influence the radiometric quality of images, as standard sensors cannot depict properly a wide range of intensities in the same scene. Indeed, overexposed or underexposed pixels cause irreplaceable information loss and degrade digital representation. Images taken under extreme lighting environments may be thus prohibitive for feature detection/extraction and consequently for matching and 3D reconstruction. High Dynamic Range (HDR) images could be helpful for these operators because they broaden the limits of illumination range that Standard or Low Dynamic Range (SDR/LDR) images can capture and increase in this way the amount of details contained in the image. Experimental results of this study prove this assumption as they examine state of the art feature detectors applied both on standard dynamic range and HDR images.

1. INTRODUCTION

3D reconstruction of architectural and cultural heritage objects is commonly used for documentation, visualization, navigation and dissemination purposes. Available Structure from Motion (SfM) algorithms are able to reconstruct large 3D scenes relatively fast, with the use of consequent or randomly taken images of the object. Standard SfM techniques rely on the accurate detection, extraction, description and matching of image features (e.g. keypoints). A huge advantage of such methods is that no prior information about the object scene or the camera path is needed. High quality images are though essential for the performance of such algorithms. Keypoint detection may be complete and more accurate if High Dynamic Range (HDR) images of the scene are used, due to the presence of additional information with respect to standard images (Chermak and Aouf, 2012; Chermak et al. 2014; Jagadish and Sinzinger, 2008). Higher number of detected keypoints will potentially increase the amount of inliers during image matching, resulting consequently to improved SfM results.

Architectural assets are often complex in terms of geometry as well as texture and due to their location are usually exposed to extreme illumination conditions i.e. containing dark shadows or bright sunlight. Under these conditions, important details and colour information of the image may be lost, degrading the result of feature detection and extraction. HDR images could be useful in such cases as they broaden the limits of luminance range that standard images can capture and increase in this way the amount of characteristic features contained in the image. Detail maintenance is crucial towards 3D reconstruction especially in such objects, making them excellent study cases for research in the field of HDR imaging.

This study focuses on the use of HDR images towards optimizing feature detection in images of high detailed architectural scenes such as church altars with complex frescoes, highly decorated arches, columns and other buildings of special architectural style. Tests are performed to investigate the behaviour of common state of the art feature detectors while using HDR images. The rest of the paper is organised as follows: Section 2 is a review of related previous work concerning HDR imaging and feature detection algorithms. Section 3 analyses our approach, followed by experimental results with self-captured images while conclusions are presented in Section 4.

2. RELATED WORK

2.1 Feature Detection

Feature detection is a fundamental research topic in many applications in computer vision and photogrammetry. Camera geometry can be calculated when a sufficient number of correct feature matches between the images is known. Thus, path estimation and 3D reconstruction are strongly influenced by the quality of the correspondences between image pairs and based consequently on reliable key point extraction. A variety of detector operators of distinct points (corners) have been presented in the past (e.g. Moravec, Forstner, Harris). More recent research investigates feature detection and description. We briefly present some important detectors and descriptors below.

SIFT (Scale Invariant Feature Transformation) is an operator introduced by Lowe (Lowe, 1999; Lowe, 2004) which can detect, describe and extract features with the use of Gaussian

---

* Corresponding author. This is useful to know for communication with the appropriate person in cases with more than one author.

scale space pyramids. Thus, SIFT features are invariant to image scale and rotation and partially invariant to illumination changes.

Bay et al. (Bay et al 2006) presented an algorithm named SURF (Speed Up Robust Features) that outperformed the existing state of the art detectors and descriptors in terms of repeatability, distinctiveness and robustness as well as speed. The detection part of the algorithm uses integral images through Hessian matrix approximation. The description part, on the other hand, describes the intensity of the neighbourhood around the pixel using Haar wavelets.

FAST (Features from Accelerated Segment Test) comes from the machine learning field (Rosten and Drummond, 2006). It uses a high speed test to distinguish key points from other points, by examining the pixels that lie along a circle around each point. If the pixels are found a lot darker or brighter that the candidate point, it is considered to be a key point. For the sake of speed, instead of checking all pixels, just four of them are tested.

BRIEF (Binary Robust Independent Elementary Features) was introduced by Calonder et al. (Calonder et al., 2010) and is a short binary descriptor using the Hamming distance. The obtained descriptor is not invariant to scale and rotation changes (Alahi et al., 2012).

ORB (Oriented FAST and Rotated BRIEF), introduced by Rublee (Rublee et al., 2011) is a combination of FAST (to detect stable points) and BRIEF (to describe them). Both these operators produce satisfactory results relatively fast, but BRIEF provides also invariance to rotation and robustness to noise (Alahi et al., 2012).

Other efficient detectors and descriptors have also been proposed in the literature (e.g. DAISY, BRISK, FREAK (Tola et al., 2010; Leutenegger et al., 2011; Alahi et al., 2012)). However, few studies regarding feature detection and tracking in HDR images have been published in the past (Chermak and Aouf, 2012; Chermak et al. 2014; Jagadish and Sinzinger, 2008; Cui et al., 2011). Evaluation studies about the behaviour feature detectors also exist. Schmid et al. compared five different detectors of the then state of the art using repeatability rate and information content as evaluation criteria (Schmid et al., 2000). Mikolajczyk et al. introduced in 2005 a reference test set of images and evaluated ten detectors and descriptors (Mikolajczyk et al., 2005). Pribyl et al. compared Harris, Shi-Tomasi, FAST and Fast Hessian detectors on HDR images (Pribyl et al., 2013). Jagadish and Sinzinger evaluated their proposed method for image matching and SIFT on HDR images (Jagadish and Sinzinger, 2008).

2.2 HDR Imaging

High dynamic range imaging (HDR) in image processing and photography is the method of generating images that contain wider dynamic range than one standard digital image can record (Reinhard et al., 2010). This is possible either by using special sensors with extended dynamic range (HDR sensor) or by merging multiple images of the same scene taken with varying exposure settings. Relative literature with techniques for the creation of HDR images exists (Debevec and Malik, 2008; Robertson et al., 1999; Mitsunaga and Nayar, 1999; Rovid et al., 2007).

An HDR image is actually a radiance map whose pixels reach a maximum of 32-bit floating point image representation (Chermak and Aouf, 2012). Common displays can visualise images in a 16- or 8-bit format but are unable to render properly the ones in 32-bit representation. To this end, a procedure called tone mapping was introduced for the compression of HDR image to fit the dynamic range of the display device while preserving the detail. Tone mapping operators can be global or local, empirical or perceptually based, static or dynamic (Cui et al., 2011). Several such operators have been developed in the recent years (Mantiuk et al., 2006; Drago et al., 2003; Durand and Dorsey, 2003; Reinhard et al., 2002; Ashikhmin, 2002; Mertens et al., 2007).

Mantiuk et al. created an algorithm that transforms the luminance values of an image to contrast using gradients on all levels of Gaussian pyramid and consequently transforms contrast values to HVS and scales the response. Finally the image is reconstructed back to luminance values (Mantiuk et al., 2006). On the other hand, Drago et al. proposed a global tone mapping operator that scales the image in the logarithmic domain, using a bias parameter and further radiometrical corrections (Drago et al., 2003). Durand and Dorsey's algorithm decomposes the SDR image into two layers: base layer and detail layer using an edge-preserving bilateral filter. The contrast of the base layer is then compressed, preserving all the details (Durand and Dorsey, 2002). A global, relatively simple tonemapping operator based on photographic practices was introduced by Reinhard et al. in (Reinhard et al., 2002). Another technique of creating HDR images is the so-called exposure fusion according to which the multiple exposures are merged into a high-quality low dynamic range image, ready for the display. Best pixel values from the sequence are selected based on a quality measure and combined into the final result. In this technique the tone mapping procedure is omitted (Mertens et al., 2007).

Although a broadly used technique in the fields of photography and image processing, HDR imaging in architectural heritage is not extended enough according to the existing literature. Ntregka et al. investigate the usage of HDR images in photogrammetric applications in the field of cultural heritage such as calibration and orthoimages (Ntregka et al., 2013). Guidi et al. study the optical pre-processing with HDR imaging that may improve the automated 3D modeling pipeline based on SfM and image matching with special emphasis on optically non-cooperative surfaced of shiny and dark materials (Guidi et al., 2014). Another recent study uses HDR images in the 3D documentation of cultural heritage and investigates how HDR images affect the 3D models, geometrically through point clouds and radiometrically through textured models (Kontogianni and Georgopoulos, 2014).

3. EXPERIMENTAL RESULTS

3.1 Test Datasets

The images used in these tests were taken with a Canon DSL camera EOS 1Ds Mark III. The datasets were produced using exposure bracketing (Figures 1, 2). During the photo shooting the camera was mounted on a tripod. Scene 1 depicts a church altar with highly decorated frescoes and other sculpture details. Scene 2 depicts the columns of an ancient altar, captured from inside. Both scenes are cases where illumination conditions of

high contrast occur, i.e. darkness and/or bright sunshine. Sequential images of scene 1 were captured in 5 different exposure layers (1-stop increment, -2EV to +2EV), while for each image of scene 2 we took 7 different ones (1-stop increment, -3EV to +3EV).
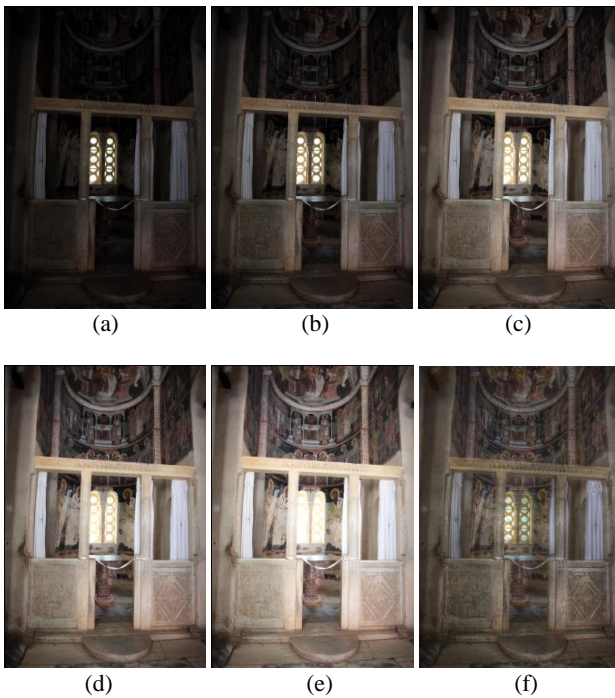

(a) (b) (c)

(d) (e) (f)

Figure 1: Scene1, (a) – (e) 5 frames with exposure values from -2EV to +2EV (f) tone mapped image according to Mantiuk's algorithm (Mantiuk et al., 2006).

(a) (b)

(c) (d)

(e) (f)

(g) (h)

Figure 2: Scene 2, (a) – (g) 7 frames with exposure values from -3EV to +3EV (h) tone mapped image according to Mantiuk's algorithm (Mantiuk et al., 2006).

3.2 HDR Image Creation

HDR image fusion was implemented according to Debevec and Malik's algorithm (Debecev and Malik, 2008). This approach uses the constraint of sensor reciprocity to linearly recover the camera response function and fuse the multiple images into a high dynamic radiance map. For the tone mapping, Mantiuk's algorithm was used (Mantiuk et al., 2006). They use a gradient domain approach, based on a low pass (Gaussian) pyramid of high contrast values. The main reason of selecting this operator was that it preserves the details needed for feature detection while enhancing the colors of the images.

3.3 Testing Detectors

In our experiments, feature detectors' behaviour is tested on SDR and HDR (their LDR equivalent) images. For this reason, state of the art feature detectors and descriptors are investigated, such as Scale Invariant Feature Transform (SIFT) (Lowe, 2004), Speeded-Up Robust Features (SURF) (Bay et al., 2006), Features from Accelerated Segment Test (FAST) (Rosten and Drummond, 2006) and the relatively new Oriented FAST and Rotated BRIEF (ORB) (Rublee et al., 2011).

Figure 3: Features detected with SIFT (detail) (a) on the SDR image and (b) on the HDR image.

Figure 4: Features detected with FAST (a) on the SDR image (14073 points) and (b) on the HDR image (68001 points).
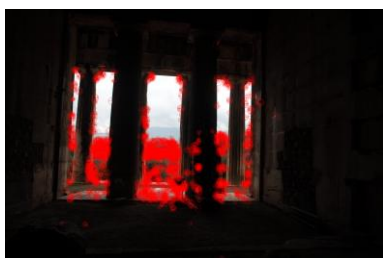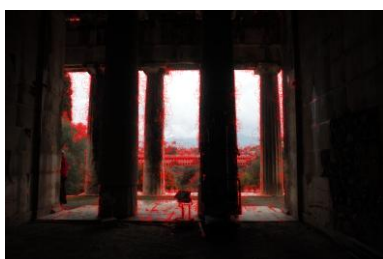

(a)

(b)

Figure 5: Features detected with ORB (a) on the SDR image (8144 points) and (b) on the HDR image (20000 points).

(a)

(b)

Figure 6: Features detected with SURF (a) on the SDR image (5047 points) and (b) on the HDR image (16843 points).

Figures 3-6 present some example results of the tested feature detectors on SDR and tone mapped HDR images for both scenes. In terms of objectiveness, the SDR images used in these tests were the ones with 0 EV. As it is shown in the figures, the use of HDR imaging for feature detection certainly increases the number of detected features. In particular, SURF increases its performance by an average of 100% while applied on HDR images. Processing time remains almost stable as it increases slightly by 2%. FAST performs even better, as it detects 350% more features, while in terms of time it is kept on the same level (around half a second). HDR imaging increases the number of detected points by almost 70% using ORB detector while needs around 14% more time, but still less than two seconds. SIFT detects approximately 200% more features on HDR images than on the standard ones, while it needs the same time. Tables 1 and 2 include the average detected feature points for all captured images and the time needed for their detection respectively. As observed in the test images, the distribution of the detected feature points remains the same for both SDR and HDR.

Table 1: Average feature points detected on SDR and HDR images for both scenes.

		SIFT	SURF	FAST	ORB
Scene 1	SDR	7062	6292	2730	19106
	HDR	13639	7663	6752	32627
Scene 2	SDR	5722	6492	2385	16902
	HDR	19013	15391	9315	23893

Table 2: Average time needed (in sec.) to detect feature points on SDR and HDR images for both scenes.

		SIFT	SURF	FAST	ORB
Scene 1	SDR	84.59	27.20	0.31	1.56
	HDR	82.74	27.79	0.32	1.76
Scene 2	SDR	75.72	29.34	0.29	1.53
	HDR	79.84	30.09	0.33	1.76

4. CONCLUSIONS

This paper has presented the usage of HDR images in feature detection by testing some state of the art operators. Comparisons are made between SDR and tone mapped HDR images in terms of performance and speed. The results show a marked increase in the number of detected feature points in the same time frame while using HRD imaging. It is believed that a larger number of correctly detected key points will potentially increase the matching performance and consequently the 3D reconstruction of such scenes. Future work will include further investigation towards this direction and evaluate matching results and 3D models while using HDR imaging on such scenes.

ACKNOWLEDGEMENTS

REFERENCES

Alahi, A., Ortiz, R., & Vandergheynst, P. (2012, June). Freak: Fast retina keypoint. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on* (pp. 510-517). IEEE.

Ashikhmin, M. (2002, July). A tone mapping algorithm for high contrast images. In *Proceedings of the 13th Eurographics workshop on Rendering* (pp. 145-156). Eurographics Association.

Bay, H., Tuytelaars, T., & Van Gool, L. (2006). Surf: Speeded up robust features. In *Computer Vision–ECCV 2006* (pp. 404-417). Springer Berlin Heidelberg.

Calonder, M., Lepetit, V., Strecha, C., & Fua, P. (2010). Brief: Binary robust independent elementary features. In *Computer Vision–ECCV 2010* (pp. 778-792). Springer Berlin Heidelberg.

Chermak, L., & Aouf, N. (2012, August). Enhanced feature detection and matching under extreme illumination conditions with a HDR imaging sensor. In*Cybernetic Intelligent Systems (CIS), 2012 IEEE 11th International Conference on* (pp. 64-69). IEEE.

Chermak, L., Aouf, N., & Richardson, M. (2014). HDR imaging for feature tracking in challenging visibility scenes. *Kybernetes*, *43*(8), 1129-1149.

Cui, Y., Pagani, A., & Stricker, D. (2011). Robust point matching in HDRI through estimation of illumination distribution. In *Pattern Recognition* (pp. 226-235). Springer Berlin Heidelberg.

Debevec, P. E., & Malik, J. (2008, August). Recovering high dynamic range radiance maps from photographs. In *ACM SIGGRAPH 2008 classes* (p. 31). ACM.

Drago, F., Myszkowski, K., Annen, T., & Chiba, N. (2003, September). Adaptive logarithmic mapping for displaying high contrast scenes. In *Computer Graphics Forum* (Vol. 22, No. 3, pp. 419-426). Blackwell Publishing, Inc.

Durand, F., & Dorsey, J. (2002, July). Fast bilateral filtering for the display of high-dynamic-range images. In *ACM Transactions on Graphics (TOG)* (Vol. 21, No. 3, pp. 257-266). ACM.

El Gamal, A., High Dynamic Range Image Sensors, 2002.

Fattal, R., Agrawala, M., & Rusinkiewicz, S. (2007). Multiscale shape and detail enhancement from multi-light image collections. *ACM Trans. Graph.*,*26*(3), 51.

Guidi G., Gonizzi S., Micoli L.L., 2014. Image pre-processing for optimizing automated photogrammetry performances. ISPRS Technical Commission V Symposium, 23-25 June, Riva del Garda, Italy.

Jagadish, K., and Sinzinger, E.D, 2008. ISVC 1, volume 5358 of Lecture Notes in Computer Science, Springer, page 359-369.

Kontogianni, G. and Georgopoulos, A., 2014. Investigating the effect of HDR images for 3D documentation of Cultural Heritage. In: *Proceedings of EuroMed 2014 – International Conference on Cultural Heritage, 3-8 November, Limassol, Cyprus.*

Leutenegger, S., Chli, M., & Siegwart, R. Y. (2011, November). BRISK: Binary robust invariant scalable keypoints. In *Computer Vision (ICCV), 2011 IEEE International Conference on* (pp. 2548-2555). IEEE.

Lowe, D. G. (1999). Object recognition from local scale-invariant features. In*Computer vision, 1999. The proceedings of the seventh IEEE international conference on* (Vol. 2, pp. 1150-1157). IEEE.

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

Mantiuk, R., Myszkowski, K., & Seidel, H. P. (2006). A perceptual framework for contrast processing of high dynamic range images. *ACM Transactions on Applied Perception (TAP)*, *3*(3), 286-308.

Mertens, T., Kautz, J., & Van Reeth, F. (2007, October). Exposure fusion. In*Computer Graphics and Applications, 2007. PG'07. 15th Pacific Conference on*(pp. 382-390). IEEE.

Mikolajczyk, K., & Schmid, C. (2005). A performance evaluation of local descriptors. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*,*27*(10), 1615-1630.

Mitsunaga, T., & Nayar, S. K. (1999). Radiometric self calibration. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*(Vol. 1). IEEE.

Ntregka, A., Georgopoulosa, A., & Quinterob, M. S. (2013). Photogrammetric Exploitation of HDR Images for Cultural Heritage Documentation. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *5*, W1.

Přibyl, B., Chalmers, A., & Zemčík, P. (2013, March). Feature point detection under extreme lighting conditions. In *Proceedings of the 28th Spring Conference on Computer Graphics* (pp. 143-150). ACM.

Reinhard, E., Heidrich, W., Debevec, P., Pattanaik, S., Ward, G., & Myszkowski, K. (2010). *High dynamic range imaging: acquisition, display, and image-based lighting*. Morgan Kaufmann.

Reinhard, E., Stark, M., Shirley, P., & Ferwerda, J. (2002, July). Photographic tone reproduction for digital images. In *ACM Transactions on Graphics (TOG)* (Vol. 21, No. 3, pp. 267-276). ACM.

Robertson, M. A., Borman, S., & Stevenson, R. L. (1999). Dynamic range improvement through multiple exposures. In *Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on* (Vol. 3, pp. 159-163). IEEE.

Rosten, E., & Drummond, T. (2006). Machine learning for high-speed corner detection. In *Computer Vision–ECCV 2006* (pp. 430-443). Springer Berlin Heidelberg.

Rovid, A. S., Varkonyi-Koczy, A. R., Hashimoto, T., Balogh, S., & Shimodaira, Y. (2007, May). Gradient based synthesized multiple exposure time HDR image. In *Instrumentation and*

Measurement Technology Conference Proceedings, 2007. IMTC 2007. IEEE (pp. 1-6). IEEE.

Rublee, E., Rabaud, V., Konolige, K., & Bradski, G. (2011, November). ORB: an efficient alternative to SIFT or SURF. In *Computer Vision (ICCV), 2011 IEEE International Conference on* (pp. 2564-2571). IEEE.

Schmid, C., Mohr, R., & Bauckhage, C. (2000). Evaluation of interest point detectors. *International Journal of computer vision*, *37*(2), 151-172.

Tola, E., Lepetit, V., & Fua, P. (2010). Daisy: An efficient dense descriptor applied to wide-baseline stereo. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *32*(5), 815-830.