Evaluation of a Local Descriptor for HDR Images

Artur Santos Nascimento[a], Welerson Augusto Lino de Jesus Melo[b],
Beatriz Trinchão Andrade[c] and Daniel Oliveira Dantas[d]
Departamento de Computação, Universidade Federal de Sergipe, São Cristóvão, SE, Brazil

Keywords: High Dynamic Range Images, Feature Point Detection, Feature Point Description.

Abstract: Feature point (FP) detection and description are processes that detect and extract characteristics from images. Several computer vision applications rely on the usage of FPs. Most FP descriptors are designed to support low dynamic range (LDR) images as input. However, high dynamic range (HDR) images can show details in bright and shadowed areas that LDR images can not. For that reason, the interest in HDR imagery as input in the detection and description processes has been increasing. Previous studies have explored FP detectors in HDR images. However, none have presented FP descriptors designed for HDR images. This study compares the FP matching performance of description vectors generated from LDR and HDR images. The FPs were detected and described using a version of the SIFT algorithm adapted to support HDR images. The FP matching performance of the algorithm was evaluated with the mAP metric. In all cases, using HDR images increased the mAP values when compared to LDR images.

1 INTRODUCTION

In computer vision, characteristics are image regions with properties such as lines, borders, or high contrast. These characteristics are extracted in two steps: detection and description of feature points (FP). Applications such as object recognition, scene reconstruction, and biometric systems rely on FP detection and description (Andrade et al., 2012; Schmid et al., 2000; Se et al., 2002).

Detection algorithms search for image FPs that can be found even when the image undergoes geometric or photometric transformations. Based on the image and its FPs, description algorithms extract discriminative invariant signatures from these FPs that can identify and distinguish a given FP from another (Rana et al., 2019). If a description is good enough, it can identify the same FP in a different capture of the same scene.

Most detection and description algorithms receive low dynamic range (LDR) images as input. An LDR image usually has 8 bits per sample. Thus, the color intensity values are limited to integers in the range

$[0, 255]$. For this reason, scenes with substantial lighting variation can result in underexposed and overexposed areas. In these cases, detectors and descriptors may fail as details are hidden by bright or shadowed areas.

On the other hand, HDR images use more than 8 bits per sample, allowing a higher dynamic range and greater color accuracy in the overexposed and underexposed areas. Knowing that feature extraction from a scene is highly dependent on the scene lighting at capture time, this study hypothesizes that the FP description will detect more FPs, especially in bright or shadowed areas.

Previous studies have evaluated FP detection and shown that FP detection performance improves when using HDR images as input (Melo et al., 2018; Ostiak, 2006). Therefore, detection algorithms adapted for HDR images were created, such as Harris Corner for HDR and DoG for HDR (Melo et al., 2018). However, few studies have evaluated FP descriptors that use HDR images as input.

In this context, this study evaluates and compares the FP description and FP matching performance in HDR and LDR images. Using the mean average precision (mAP) metric, two datasets, and an adapted version of the SIFT algorithm, we observed that using HDR images as input increased the mAP by 81.80% (from 0.22 to 0.40) with the first dataset (3D lighting),

[a] https://orcid.org/0000-0003-0419-6170
[b] https://orcid.org/0000-0003-0644-2427
[c] https://orcid.org/0000-0002-1407-8250
[d] https://orcid.org/0000-0002-0142-891x

and by 61.50% (from 0.39 to 0.63) with the second one (2D lighting).

This paper is organized as follows. Section 2 describes the related works. Section 3 briefly describes the SIFT algorithm. Section 4 describes the dataset used, the evaluation metric, and the proposed methodology. Section 5 shows the experimental results. The conclusions are drawn in Section 6.

2 RELATED WORKS

The number of studies using HDR and tone-mapped images as input for detection and description algorithms is significantly smaller than those using LDR images as input. In this section, we classify studies that use HDR and tone-mapped images into four groups: studies that propose datasets with complex lighting configurations and that offer captures in both LDR and HDR formats; studies focused on FP detection algorithms; studies focused on FP description algorithms; and studies that adapt FP detection or description algorithms for specific applications.

2.1 Datasets

Přibyl (Přibyl et al., 2012) studied the use of tone mapping (TM) in HDR images to improve FP detection. The main idea is to capture more details in regions that conventional LDR images can not correctly represent due to the lack of details in bright and shadowed areas of the scene. There is no need to change the detection algorithms in this approach as the TM algorithm maps the tones from the HDR image into an LDR image.

With that in mind, Přibyl generated a dataset whose images have abrupt lighting changes. Two scenes were captured: a planar scene (2D) containing three posters in A4 sheet format; and a 3D scene containing several non-planar rigid objects. Both scenes were placed in a dark room with controlled lighting. The scenes were captured in three different sequences, consisting of changing viewpoints, distances, and lighting.

- **Viewpoint Changing Sequence:** the camera is moved following a circular trajectory with its center in the scene with a step of $2.5°$. Thereby, 21 captures were made, resulting in a $50°$ total trajectory;

- **Distance Changing Sequence:** the scene was captured seven times and the distance between the camera and the scene increased progressively, yielding the distance sequence of 100 cm, 103 cm, 109 cm, 122 cm, 147 cm, 197 cm, and 297 cm;

- **Lighting Changing Sequence:** the scene was also captured seven times, each time with a different combination of three light sources being on or off, with at least one of them on.

They conclude that images generated from local TM algorithms give better results than images from global TM algorithms when used as input to the feature detection process (Přibyl et al., 2012).

In order to produce a series of studies exploring the use of TM in HDR images to improve the detection and description processes, Rana et al. (Rana et al., 2015) proposed a new dataset. They also evaluated TM algorithms not considered by Přibyl et al. (Přibyl et al., 2012).

The dataset proposed by Rana et al. consists of two capture sequences, named project room sequence (PRS) and light room sequence (LRS). PRS is composed of eight different lighting configurations created by blocking light coming from a projector with the help of different objects. The scene is composed of several bright and dark-colored objects arranged to create sharp shadows and bright areas. LRS comprises seven different natural lighting conditions created by changes in global lighting through the opening and closing of window blinds, room ambient illumination, and diffuse lighting from a tungsten lamp. In this scene configuration, there are also dark and light objects with different types of surfaces.

2.2 FP Detection Algorithms

Přibyl et al. (Přibyl et al., 2016) use the dataset proposed in their previous studies (Přibyl et al., 2012) to expand their evaluation of the impact of TM during FP detection. They include more TM algorithms and the DoG detection algorithm in their experiments. The conclusion is that the TM algorithms proposed by Fattal et al. (Fattal et al., 2002) and Mantiuk et al. (Mantiuk et al., 2006) detect more FPs than other ones.

Two studies by Rana et al. focus on FP detection. In the first study (Rana et al., 2016b), Rana et al. propose a TM operator to optimize FP detection. As a result, using their TM operator improves the correlation coefficient (CC) and repeatability rate (RR) compared to other TM operators. In the second study (Rana et al., 2017a), Rana et al. develop a new adaptive local TM operator that uses support vector regression (SVR) to predict optimal modulation maps to improve FP detection. As a result, the TM operator by Rana et al. showed better RR when compared to other TM operators.

Melo et al. propose a modification for FP detectors in order to improve FP detection on HDR images.

A local mask based on the coefficient of variation (CoV) of sets of pixels is proposed as an additional step in FP detection. Also, the Uniformity metric is introduced as a new criterion to evaluate FP detection in areas under different lighting intensities. The obtained results show better uniformity and repeatability rate in most tested HDR images when compared to standard FP detectors. Moreover, they indicate that HDR images have great potential to be explored in applications that rely on FP detection (Melo et al., 2018).

2.3 FP Description Algorithms

There are several studies from Rana et al. (Rana et al., 2016a; Rana et al., 2017b; Rana et al., 2019) focused on image description based on HDR images. The first one (Rana et al., 2016a) compared image matching using TM in HDR images and LDR images using SIFT (Lowe, 2004), SURF (Bay et al., 2008), FREAK (Alahi et al., 2012), and BRISK (Leutenegger et al., 2011) to describe the FPs. By using several global and local TM algorithms, it is observed that all combinations that used tone-mapped images as input performed better than with LDR images. In their second study (Rana et al., 2017b), Rana et al. proposed an adaptive TM operator that uses SVR to predict optimal modulation maps to improve FPs description, specifically to make image matching invariant to day or night scene illumination. Finally, in the third study (Rana et al., 2019), Rana et al. used the adaptive TM operator to improve FP description in image matching.

Khwildi and Zaid (Khwildi and Zaid, 2018) introduce a descriptor based on LDR image expansion. In this approach, the LDR image is converted to an HDR image using a reverse TM operator (TM-HDR image). Then, the resulting TM-HDR image is tone mapped back to LDR image (TM-LDR) and SIFT descriptor is used to describe FPs. As a result, it is demonstrated that features described from the TM-LDR image are more descriptive than those extracted from LDR and TM-HDR images. In future works, they consider exploring local TMs and improving the efficiency of local descriptors directly in HDR images.

2.4 Applications

Ige et al. (Ige et al., 2016) developed a facial expression recognition algorithm using support vector machines (SVMs) and local binary patterns (LBP). First, a TM algorithm is used to convert HDR into LDR images. Then, the resulting tone-mapped image is used as input to the SURF algorithm. Finally, the tone-mapped images and the LDR images are compared. As a result, the approach using tone-mapped images showed better results than LDR images. Tone-mapped images reach 79.8% accuracy, while the traditional methods that use LDR images range between 31.3% and 70.8% accuracy.

Ostiak et al. (Ostiak, 2006) used HDR images to execute an image stitching process to generate a panorama. In their study, a tone-mapped panorama is generated to make the shadowed and bright areas of the image visible. They mention a modification of the SIFT algorithm to describe FPs in HDR images without giving further details. The discussion is based on a visual analysis of the generated panoramas, and the results are subjective. The algorithm presents a better performance in static scenes than in dynamic scenes, when stitching shadowed and bright areas that are not visible in LDR images.

The related works show that FP detection and description using HDR images as input is a field yet to be explored. Most HDR-related studies use tone-mapped images generated from HDR. Tone-mapped images bring more details in shadowed and bright areas than LDR images and there is no need to adjust detection and description algorithms.

On the other hand, few studies using HDR images as input were found and just Melo et al. (Melo et al., 2018) detailed the modifications made in their FP detection algorithm to support HDR images. Using HDR images brings more information but requires adaptations in detection and description algorithms to support floating-point values used to represent pixels.

All studies showed improvements when using tone-mapped images as input to the detection and description algorithms. Specifically, Melo et al. (Melo et al., 2018) showed that using HDR images as input resulted in better detection performance in terms of repeatability rate and uniformity, especially in shadowed areas of the image.

3 SIFT

The Scale Invariant Feature Transform (SIFT) detects and describes FPs from an image. SIFT is one of the most popular FP extractors, invariant to scale and orientation (Rey Otero and Delbracio, 2014), and shows good performance in detection and extraction, as shown in various studies (Přibyl et al., 2016; Rana et al., 2016a; Rana et al., 2017a; Rana et al., 2017b; Rana et al., 2019).

To detect the FPs, SIFT creates an image representation from the input image called *scale-space*, which

consists of a collection of increasingly blurred images. Each FP consists of a blob-like structure whose coordinates (x, y) and characteristic scale σ are located with subpixel accuracy. SIFT calculates the dominant orientation Θ of each FP detected. In this way, the tuple (x, y, σ, Θ) defines the center, size, and orientation of a normalized patch where the description vector is computed. The SIFT detector is known as Difference of Gaussian (DoG).

The SIFT description vector has 128 dimensions and is built from a histogram of local gradient directions around the FP. The local neighborhood size is normalized to obtain scale-invariance in the description. Next, a dominant orientation in the FP neighborhood is calculated and used to orient the grid over which the histogram is determined to make the description rotation-invariant. The SIFT FP description vector is theoretically invariant to image translation, rotation and scale changes (Lowe, 2004; Rey Otero and Delbracio, 2014).

4 METHODS

To evaluate and compare FP description when using LDR and HDR images, we chose datasets that provide scenes with complex lighting and scene captures (i.e. images with bright and shadowed areas) in both LDR and HDR formats (Section 4.1). We performed the FP detection and the description of the images in the LDR and HDR datasets using an adapted version of SIFT (Section 4.2). Finally, we matched FPs from pairs of images and evaluated the matching performance (Section 4.3). Figure 1 illustrates the pipeline of the experiment.

4.1 Dataset

The datasets by Rana et al. (Rana et al., 2015) and Přibyl (Přibyl et al., 2012) were the only datasets found that explore complex lighting configurations and are publicly available (Section 2.1). Those datasets are composed of a series of captures, each one made with a different shutter speed. The faster is the shutter, the darker the obtained photo is, as the camera sensor is less stimulated. The algorithm by Debevec and Malik (Debevec and Malik, 2008) uses this collection of captures with different shutter speeds to generate the HDR image of the scene. In each collection, the LDR sample image is acquired with automatic exposure.

We chose to use the dataset by Přibyl et al. (Přibyl et al., 2012), which consists of two scenes: a planar

scene (2D) containing three posters in A4 sheet format; and a volumetric (3D) scene containing several non-planar rigid objects. Both scenes were placed in a dark room with controlled lighting. As we are interested in investigating the effect of lighting changes in images with different dynamic ranges, in this study, we use the lighting changing sequence of both 2D and 3D datasets by Přibyl (Přibyl et al., 2012).

4.2 FP Detection and Description

We used as FP detector the SIFT detection step, known as DoG. Minor changes were made in DoG to support 32-bit floating point pixel values. In this study, we adopt the approach used by Rana et al. (Rana et al., 2015) and Melo et al. (Melo et al., 2018), i.e., to select the 500 FPs with strongest responses of each detector.

To describe the FPs, the SIFT description step was adapted to receive both LDR and HDR images. HDR images were converted to a common interval of $[0.0, 255.0]$, using *floating-point* instead of *unsigned char*, and the description of the best 500 FPs was generated. The implementation is available at GitHub[1].
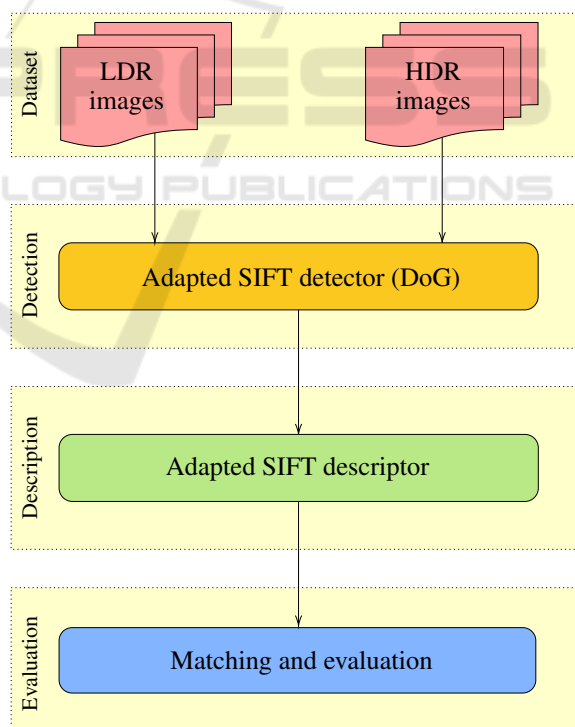

Figure 1: Flowchart representing the executed pipeline.

---

[1] https://github.com/welersonMelo/TCC_UFS/

4.3 Matching and Evaluation

Our aim is to evaluate if the description using HDR images is better than the description generated using conventional LDR images. The most common way to do that is by matching description vectors from different captures of the same scene. To do that, the 500 FPs with strongest responses from each detector were selected. This approach was used in other studies (Rana et al., 2016a; Melo et al., 2018).

To evaluate the description matching, we computed a precision-recall (PR) curve (Mikolajczyk and Schmid, 2005; Rana et al., 2016a). The PR curve is based on the true and false matches between FPs from a pair of images. The description of a FP is usually represented by a vector that encapsulate its characteristics (description vector). Given a description vector, we have a match if the nearest neighbor distance ratio (NNDR) criteria is satisfied. In this criteria, a good match occurs when the ratio between its distance from the first best match (d_{BM1}) and the second best match (d_{BM2}) is less than a given threshold th (Equation 1). The threshold is defined by testing the value that best fits the dataset used (Lowe, 2004). We use 0.7 as threshold and the Euclidean distance to compute the distance between two description vectors.

$$\text{NNDR} = \frac{d_{BM1}}{d_{BM2}} < th \qquad (1)$$

Two description vectors are considered to have a true positive match if they correspond to two FPs that are repeated in the reference and test images. A test FP is considered repeated if it lies in a circle of radius r centered on the projection of the reference FP onto the test image. In our case, we considered $r = 5$. Analogously, a match is considered a false positive if the corresponding FPs are not repeated.

To compute the PR curve, we need the *recall* and *precision* values. Recall is defined by Equation 3 and precision by Equation 2, where *tp*, *fp*, and *fn* refer to numbers of true positives, false positives, and false negatives respectively. By varying the threshold *th*, we draw a PR curve and measure the area under this curve. The value of the area under the curve is also known as average precision (AP). The mean of APs for all possible image pairs from a dataset is called mean average precision (mAP) (Rana et al., 2016a).

$$P = \frac{tp}{tp + fp} \qquad (2)$$

$$R = \frac{tp}{tp + fn} \qquad (3)$$

5 RESULTS

The first step was to detect the FPs of the LDR and HDR databases using the DoG detector. Then, the description vectors of the FPs was calculated using the adapted SIFT descriptor. For each dataset, FPs from every possible pair of images were matched. As each dataset has seven captures, 21 pairs of images were considered for each dataset.

Afterwards, we obtained the PR curve and AP values of each pair of images. Table 1 shows the AP values of the DoG detector applied to LDR and HDR images from the 2D lighting dataset. Table 2 shows the AP values of the adapted SIFT algorithm applied to LDR and HDR images from the 3D lighting dataset.

With the AP values in hand, we calculated the mAP of the FP matching. The last lines of Tables 1 and 2 show the calculated mAP values from description vectors generated by the adapted SIFT algorithm. The mAP values are in the interval [0.0, 1.0]. Higher values are better. Matching results show an improvement of 61.50% in 2D lighting dataset and 81.80% in 3D lighting dataset when using HDR instead of LDR images.

Figures 2, 3, 4 and 5 show examples of FP matching using LDR and HDR images respectively from the 2D lighting dataset. Green lines are correct matches and red lines are incorrect matches.

6 CONCLUSIONS

In this study, we evaluated the use of HDR images as input to an FP extraction pipeline. Although previous studies have explored FP detectors in HDR images, none have presented FP descriptors designed for HDR images. This study compares the FP matching performance of description vectors generated from LDR and HDR images. The code of the proposed SIFT detector and descriptor are available at GitHub. Using Přibyl et al. dataset (Přibyl et al., 2012), the SIFT algorithm adapted to support HDR images, and the mAP metric, we evaluated if the usage of HDR images improves the performance of FP matching.

Using HDR images to describe FPs significantly improved the FP matching. The adapted SIFT algorithm, when applied to HDR instead of LDR images, increased the mAP by 61.50% in the 2D lighting dataset and by 81.80% in the 3D lighting dataset. This advocates that using HDR images can improve description performance.

The usage of HDR images as input in FP description algorithms is an area yet to be explored, as there are few studies about this topic in the literature. Fu-

Table 1: Average precision in 2D lighting dataset using pairs of LDR and HDR images.

Image Label	2D lighting LDR average Precision						2D lighting HDR average precision					
	2	3	4	5	6	7	2	3	4	5	6	7
1	0.38	0.14	0.39	0.74	0.53	0.72	0.80	0.21	0.75	0.88	0.88	0.87
2	–	0.10	0.16	0.43	0.60	0.41	–	0.19	0.65	0.73	0.88	0.77
3	–	–	0.39	0.29	0.12	0.28	–	–	0.31	0.24	0.36	0.25
4	–	–	–	0.40	0.22	0.38	–	–	–	0.75	0.70	0.66
5	–	–	–	–	0.42	0.72	–	–	–	–	0.82	0.89
6	–	–	–	–	–	0.45	–	–	–	–	–	0.79
mAP	**0.39**						**0.63**					

Table 2: Average precision in 3D lighting dataset using pairs of LDR and HDR images.

Image Label	3D lighting LDR average Precision						3D lighting HDR average precision					
	2	3	4	5	6	7	2	3	4	5	6	7
1	0.16	0.17	0.04	0.07	0.18	0.25	0.58	0.67	0.03	0.16	0.55	0.68
2	–	0.48	0.03	0.05	0.58	0.53	–	0.84	0.03	0.17	0.85	0.80
3	–	–	0.01	0.04	0.59	0.75	–	–	0.03	0.17	0.77	0.87
4	–	–	–	0.06	0.04	0.08	–	–	–	0.03	0.12	0.08
5	–	–	–	–	0.06	0.02	–	–	–	–	0.10	0.12
6	–	–	–	–	–	0.61	–	–	–	–	–	0.80
mAP	**0.22**						**0.40**					

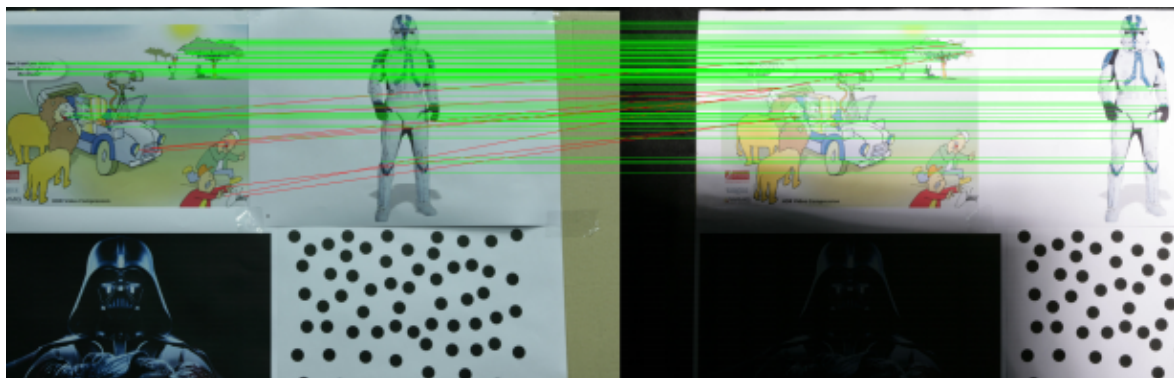Figure 2: Example of FP matching using 2D lighting dataset and LDR images.

Figure 3: Example of FP matching using 2D lighting dataset and HDR images.
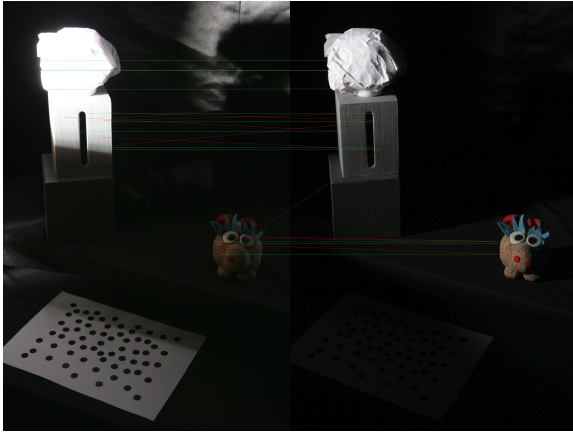
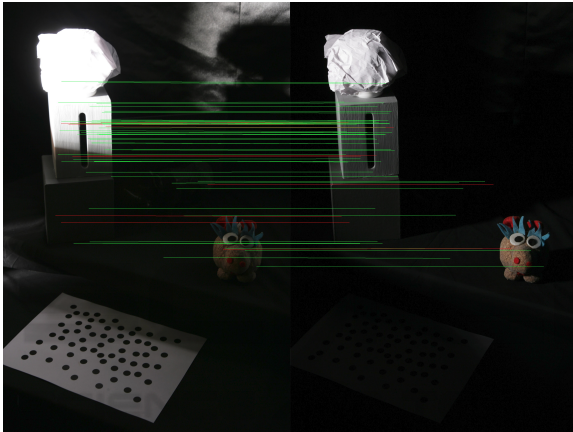Figure 4: Example of FP matching using 3D lighting dataset and LDR images.

Figure 5: Example of FP matching using 3D lighting dataset and HDR images.

ture works may include testing and evaluating the performance of alternative detectors and descriptors applied to HDR images and comparing the adapted SIFT algorithm with methods that use tone mapping. Other metrics may also be used, such as repeatability rate and uniformity.

REFERENCES

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

Andrade, B. T., Mendes, C. M., de Oliveira Santos Jr, J., Bellon, O. R. P., and Silva, L. (2012). 3d preserving xviii century barroque masterpiece: Challenges and results on the digital preservation of Aleijadinho's sculpture of the prophet joel. *Journal of Cultural Heritage*, 13(2):210–214.

Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L. (2008).

Speeded-up robust features (surf). *Computer Vision and Image Understanding*, 110(3):346–359. Similarity Matching in Computer Vision and Multimedia.

Debevec, P. E. and Malik, J. (2008). Recovering high dynamic range radiance maps from photographs. In *ACM SIGGRAPH 2008 Classes*, SIGGRAPH '08, New York, NY, USA. Association for Computing Machinery.

Fattal, R., Lischinski, D., and Werman, M. (2002). Gradient domain high dynamic range compression. In *Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, pages 249–256.

Ige, E. O., Debattista, K., and Chalmers, A. (2016). Towards HDR based facial expression recognition under complex lighting. In *Proceedings of the 33rd Computer Graphics International*, CGI '16, page 49–52, New York, NY, USA. Association for Computing Machinery.

Khwildi, R. and Zaid, A. O. (2018). A new retrieval system based on low dynamic range expansion and sift descriptor. In *2018 IEEE 20th International Workshop on Multimedia Signal Processing (MMSP)*, pages 1–6. IEEE.

Leutenegger, S., Chli, M., and Siegwart, R. Y. (2011). Brisk: Binary robust invariant scalable keypoints. In *2011 International Conference on Computer Vision*, pages 2548–2555.

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

Mantiuk, R., Myszkowski, K., and 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.

Melo, W. A. L. J., de Tavares, J. A. O., Dantas, D. O., and Andrade, B. T. (2018). Improving feature point detection in high dynamic range images. In *2018 IEEE Symposium on Computers and Communications (ISCC)*, pages 00091–00096. IEEE.

Mikolajczyk, K. and Schmid, C. (2005). A performance evaluation of local descriptors. *IEEE transactions on pattern analysis and machine intelligence*, 27(10):1615–1630.

Ostiak, P. (2006). Implementation of HDR panorama stitching algorithm. In *Proceedings of the 10th CESCG Conference*, pages 24–26. Citeseer.

Přibyl, B., Chalmers, A., and Zemčík, P. (2012). Feature point detection under extreme lighting conditions. In *Proceedings of the 28th Spring Conference on Computer Graphics*, pages 143–150.

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*, 38:141–160.

Rana, A., Valenzise, G., and Dufaux, F. (2015). Evaluation of feature detection in HDR based imaging under changes in illumination conditions. In *2015 IEEE International Symposium on Multimedia (ISM)*, pages 289–294.

Rana, A., Valenzise, G., and Dufaux, F. (2016a). An evaluation of HDR image matching under extreme illumination changes. In *2016 Visual Communications and Image Processing (VCIP)*, pages 1–4.

Rana, A., Valenzise, G., and Dufaux, F. (2016b). Optimizing tone mapping operators for keypoint detection under illumination changes. In *2016 IEEE 18th International Workshop on Multimedia Signal Processing (MMSP)*, pages 1–6.

Rana, A., Valenzise, G., and Dufaux, F. (2017a). Learning-based adaptive tone mapping for keypoint detection. In *2017 IEEE International Conference on Multimedia and Expo (ICME)*, pages 337–342.

Rana, A., Valenzise, G., and Dufaux, F. (2017b). Learning-based tone mapping operator for image matching. In *2017 IEEE International Conference on Image Processing (ICIP)*, pages 2374–2378.

Rana, A., Valenzise, G., and Dufaux, F. (2019). Learning-based tone mapping operator for efficient image matching. *IEEE Transactions on Multimedia*, 21(1):256–268.

Rey Otero, I. and Delbracio, M. (2014). Anatomy of the SIFT Method. *Image Processing On Line*, 4:370–396. https://doi.org/10.5201/ipol.2014.82.

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

Se, S., Lowe, D., and Little, J. (2002). Global localization using distinctive visual features. In *IEEE/RSJ international conference on intelligent robots and systems*, volume 1, pages 226–231. IEEE.