LEARNING-BASED ADAPTIVE TONE MAPPING FOR KEYPOINT DETECTION

Aakanksha Rana

LTCI, Télécom ParisTech
Université Paris Saclay

Giuseppe Valenzise, Frederic Dufaux

Laboratoire des Signaux et Systèmes (L2S, UMR 8506)
CNRS - CentraleSupelec - Université Paris-Sud

ABSTRACT

The goal of tone mapping operators (TMOs) has traditionally been to display high dynamic range (HDR) pictures in a perceptually favorable way. However, when tone-mapped images are to be used for computer vision tasks such as keypoint detection, these design approaches are suboptimal. In this paper, we propose a new learning-based adaptive tone mapping framework which aims at enhancing keypoint stability under drastic illumination variations. To this end, we design a pixel-wise adaptive TMO which is modulated based on a model derived by Support Vector Regression (SVR) using local higher order characteristics. To circumvent the difficulty to train SVR in this context, we further propose a simple detection-similarity-maximization model to generate appropriate training samples using multiple images undergoing illumination transformations. We evaluate the performance of our proposed framework in terms of keypoint repeatability for state-of-the-art keypoint detectors. Experimental results show that our proposed learning-based adaptive TMO yields higher keypoint stability when compared to existing perceptually-driven state-of-the-art TMOs.

***Index Terms*—** High dynamic range, tone mapping operator, keypoint detection, stochastic gradient descent, machine learning.

1. INTRODUCTION

High Dynamic Range (HDR) imaging enables to capture and reproduce a wide range of radiance and luminance of the real world environments [1]. Such extended capability of HDR to store high contrast information from the very *dark* and *bright* regions of a scene has raised potential interest in solving challenging computer vision problems such as keypoint detection [2,3].

Keypoint detection is the primary and crucial stage of several computer vision algorithms *e.g.*, image retrieval, reconstruction, tracking, etc. Conventionally, keypoint detection algorithms [4] have been designed and optimized with respect to low dynamic range (LDR) imagery which is represented using gamma-corrected 8-bit integer representation and is approximately linear to human perception. Conversely, the real-valued HDR pixels are proportional to the physical luminance of the scene which is expressed in cd/m^2. As a result, HDR linear values are not appropriate when used with LDR-optimized keypoint detection algorithms. In such a scenario, a plausible solution investigated by recent studies [2,3,5] is to convert HDR into an LDR representation using a Tone Mapping Operator (TMO) [1].

TMOs have been traditionally designed to compress HDR content in a suitable 8-bit LDR representation with the aim of displaying HDR images on standard screens [6,7]. Consequently, their design has been tuned to preserve human-vision attributes, *e.g.*, image aesthetics, perceptual contrast, etc. However, differently from visual perception, keypoint detection is a computer vision algorithm designed for a machine, and thus follows different optimization criteria. It looks for stable locations (*e.g.*, corners, blobs etc.) in an image that can be repeatedly detected when the same scene is captured under different transformations such as illumination variations. Therefore, existing TMOs may be sub-optimal for keypoint dependent applications such as tracking or panorama stitching and need to be re-calibrated. Similar conclusions are drawn from the experimental validations of several recent studies [2,3,8].

In our recent work [9], we investigate the strategies for designing a keypoint-optimal TMO. To that end, we compare the optimization of existing TMO parameters with respect to: a) Repeatability Rate (performance measure criterion of keypoint detection) and b) correlation score between tone mapped images of the same scene with lighting variations. Our experimental study concludes that optimizing TMO parameters with respect to Repeatability Rate (RR) leads to higher keypoint stability over the per-pixel similarity between the tone-mapped images. Though this study points to the parametric sensitivity in TMOs, it does not provide any keypoint-detection-optimized TMO model. Therefore, the problem of designing an optimal TMO for the keypoint detection task remains open.

In this paper, we address this problem and develop a novel learning-based adaptive tone mapping operator (AdTMO). Our proposed framework aims at enhancing the repeated detection of sparse keypoint locations (*e.g.*, corners) in high-contrast areas of scenes undergoing complex real-world illumination transitions such as day/night change. To this end, we initially introduce an adaptive TMO which can be locally modulated, *i.e.*, its parameters can vary pixel-wise. We then propose to derive the per pixel modulation by means of a learned illumination invariant model. In this context, we train a Support Vector Regressor (SVR) to predict the desired pixel-wise modulation maps by using the linear HDR content from scenes captured with varying lighting conditions.

Our idea is mainly motivated by the conclusions of our previous study [9] where optimizing tone mapping parameters for keypoint detection is shown to yield significant gains in RR. However, in that work optimal TMO parameters are computed globally on the whole image using grid search and, more importantly, for a given scene. In this work, instead, we propose to *learn TMO parameters based on the local features* of the scene. Specifically, since keypoints are sparsely detected and depend on their neighborhood properties, we argue that local parametric modulations in TMOs can enhance the keypoint detection probability by adaptively mapping pixels based on their local higher-order characteristics. Moreover, to predict such optimal modulations in this context, we are inspired by the success of regression-based "task-optimization" models. In the literature, regression-based models have been explored for several image processing problems [10–12]. In this work, we employ SVR, which

The work presented in this document was supported by BPIFrance and Région Ile de France, in the framework of the FUI 18 Plein Phare project

978-1-5090-6067-2/17/$31.00 ©2017 European Union

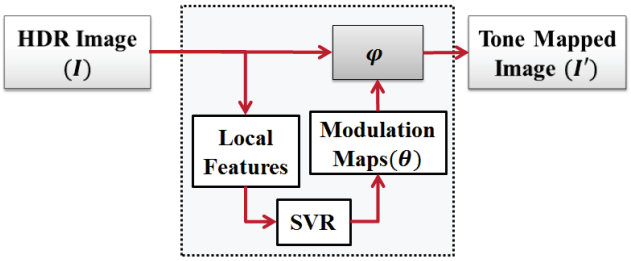


Fig. 1: *Learning based AdTMO.*

has been successfully used, *e.g.*, in image super-resolution [10], and which enables to cope with large variability in the input training samples compared with low-dimensionality approaches using explicit functions such as polynomial regression.

Learning-based models have been seldom pursued for designing keypoint-optimized TMOs. As a consequence, there is no standard dataset to train or test any model in this context. In this paper, we overcome this difficulty by proposing a simple detection-similarity-maximization model to generate appropriate training samples. We initially consider several HDR image pairs which are taken with the same viewpoint with different lighting variations. Then, we define our objective function to find optimal modulation maps for such scenes so that the similarity of the detector response maps is maximized. For the defined objectives, the optimization is carried out using stochastic gradient descent (SGD) [13] by deriving the required partial derivative architecture.

In this paper, we additionally propose an HDR dataset of 8 image scenes taken in indoor and outdoor locations with different lighting variations. We finally present the comparison of our approach with state-of-the-art TMOs using different keypoint detection schemes. Our results show consistent gains in term of overall keypoint stability across different illumination conditions with respect to popular tone mapping approaches proposed in the literature.

The paper is organized as follows. In Section 2, we provide the details of our learning-based adaptive TMO approach, similarity maximization model, the SVR training and proposed dataset. We present the experimental results and analysis in Section 3. Finally, the conclusions are drawn in Section 4, along with future research directions.

2. PROPOSED TONE MAPPING MODEL

2.1. General Framework

Let φ be a tone mapping function which maps the linear-valued HDR content of an image I to an output LDR I'. In general, for a pixel x TMO operates as: $I'(x) = \varphi(I(x), \boldsymbol{\theta})$, where $I(x) \in \Re$, $I'(x) \in [0, 255]$ and $\boldsymbol{\theta}$ represents a vector of parameters.

For several existing TMOs [6,7,14,15], parameters $\boldsymbol{\theta}$ serves diverse objective such as filter shape and size, brightness control, but all motivated for visual perception. Such parameters are often set as globally for an image and further chosen by trial and test procedures. For example, $\boldsymbol{\theta}$ serves as variance in ChiuTMO [15], sharpening constant in ReinhardTMO [6] and range and spatial variance in bilateral filtering based TMO [9].

Based on these observations, we assume function φ as an extension of existing tone mapping functions which can be modulated spatially by adapting their vector of parameters. The idea here is to facilitate the local adaption of function φ at sparse keypoint locations to further ease their identification and detection. In this paper, we call the corresponding parameter vectors as *modulation maps* so as to distinguish their purpose of modulating the TMO locally from global parametric tuning. The modulation maps are given as $\boldsymbol{\theta}(x) = \{\theta_1(x), \theta_2(x)..,\}$ and our proposed TMO operates as: $I'(x) = \varphi(I(x), \boldsymbol{\theta}(x))$.

To predict the *modulations maps*, we propose to learn a model by employing SVR [16] while complying with the following two constraints: (a) To distinguish the keypoint and its neighborhood locations, (b) To bring invariance (as much as possible) to the non-affine lighting variations in the physical world scenes.

By using the radial basis kernel mapping, our SVR minimizes the non-linear problem of predicting modulation maps $\boldsymbol{\theta}$ by linearly separating the input samples in high-dimensional space. We refer the reader to [16] for more details about kernel-based SVR optimization model. Fig. 1 outlines the general framework of our proposed keypoint optimal TMO.

2.1.1. Adaptive Tone Mapping Operator

Many tone mapping approaches aim at separating scene illumination, which can display large dynamic range variations, from the reflectance of objects, which instead has lower dynamic range characteristics [9,15]. Following this idea, our tone mapping function φ is expressed as: $\varphi = I \cdot L^{-1}$, where the illumination component L is estimated by an adaptive version of bilateral filtering [17] and is given as:

$$L(x, \boldsymbol{\theta}) = \frac{1}{W} \cdot \sum_{y \in \Omega} G_{\theta_1(x)}(\|x - y\|) \cdot G_{\theta_2(x)}(\|I(x) - I(y)\|) I(y), \tag{1}$$

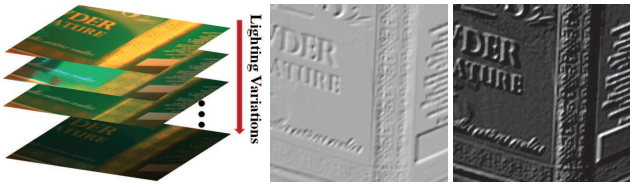
where G is a Gaussian kernel. Here, modulation vector $\boldsymbol{\theta}$ has two components: θ_1 and θ_2, also known as spatial and range variance. For each pixel location x, y is a pixel in neighborhood set Ω and the normalization factor $W = \sum_{y \in \Omega} G_{\theta_1(x)}(\|x-y\|) \cdot G_{\theta_2(x)}(\|I(x) - I(y)\|)$.

It is important to note that we have built our model using the bilateral filtering, mainly because its proposed adaptive formulation facilitates the integration of local modulation in the proposed TMO. Moreover, it has been previously studied in the context of keypoint detection in HDR imaging in varying lighting conditions [9]. However, any other tone mapping techniques with parametric formulations such as [6,15] could be used as well with our proposed framework.

2.2. Generation of Training Set: Detection Similarity Maximization Model

Suppose we are given a set of HDR scenes where each scene has images captured from the same viewpoint but with different lighting condition. To train the SVR for our proposed model, we need to compute the "ideal" modulation maps (θ_1, θ_2 in our case) for a scene which ensures high keypoint stability. In other words, for a scene undergoing lighting variations, we need to estimate the modulations ensuring maximum keypoint repeatability. To this end, one solution is to design an optimization model which maximizes the RR of multiple images of a given sequence.

RR is a measure of detector efficiency, defined as $\frac{r_{ref}}{\min(n_{ref}, n_{test})}$, where r_{ref} is the number of keypoints detected in the reference image which are repeated in the test image,



(a) HDR Images (b) Modulation (θ_1) (c) Modulation (θ_2)

Fig. 2: *Generation of training set.* The samples images undergoing different lighting variations shown in (a) are used to generate the θ_1 and θ_2 modulation maps in (b) and (c) respectively, using the detection similarity maximization model.

and n_{ref} and n_{test} are the number of detected keypoints in reference and test image, respectively. Since RR is a non-smooth and non-differentiable function, it cannot be directly used to define the similarity objective of our optimization model. Therefore, we instead propose an alternative solution to use *differentiable* detector response maps $\mathcal{R}$ and design a model that maximizes the similarity between these response maps of image pairs drawn from a given sequence. $\mathcal{R}$ is a score map which determines a pixel's strength to be a keypoint and it mainly depends on the choice of keypoint detection algorithm.

Our response map $\mathcal{R}$ is generated by a Harris corner detector [18]. It is based on the autocorrelation scores computed per pixel using the second-order moment matrix, and is given as:

$$\mathcal{R}\big(\varphi(x,\boldsymbol{\theta})\big) = det\{\boldsymbol{M}\big(\varphi(x,\boldsymbol{\theta})\big)\} - k \cdot tr\{\boldsymbol{M}\big(\varphi(x,\boldsymbol{\theta})\big)\}^2 \quad (2)$$

where $\boldsymbol{M}$ is the second order moment matrix as detailed in [18]. k is the sensitivity factor ($k = 0.04$). In this work, we have focused on the corner-based detectors as they are computationally inexpensive and highly used for real time applications, *e.g.*, tracking, wide-view panorama creations, etc. However, the model could be extended to region or blob-based detectors as well.

Objective: Let S be a scene consisting of N HDR images with lighting variations as shown in Fig. 2 (a). Let $P = \big\{(1,2),(2,3)..,\big\}$ be the set of $K = \binom{N}{2}$ pair combinations of N images. Our aim is to maximize the response similarity by minimizing the following objective function:

$$\mathcal{F}(\theta) = \frac{1}{K} \sum_{\{i,j\} \in P} \Phi(\mathcal{R}_i(\theta), \mathcal{R}_j(\theta)), \quad (3)$$

and obtain the resulting modulation maps $\boldsymbol{\theta} = \{\theta_1, \theta_2\}$ as shown in Fig. 2 (b) and (c). Inspired by max-margin formulations for image retrieval tasks [19], we define function Φ using the logistic penalty

$$\Phi(\mathcal{R}_i, \mathcal{R}_j) = \log(1 + \exp(\epsilon - \langle \mathcal{R}_i \cdot \mathcal{R}_j \rangle). \quad (4)$$

where ϵ is the penalty control factor, $\mathcal{R}_i$ and $\mathcal{R}_j$ are the response maps corresponding to the images $i, j \in S$, and $\langle \cdot \rangle$ denotes the scalar product.

Optimization using SGD. We optimize the objective function in Eq. (3) using Stochastic Gradient Descent (SGD) [13]. To do so, we build the partial derivative architecture required for the SGD implementation as follows.

To estimate $\boldsymbol{\theta}$ maps at each iteration t, SGD update rule is given as:

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t - \gamma_t \cdot \nabla \Phi_{\{i,j\}t}(\boldsymbol{\theta}_t), \quad (5)$$

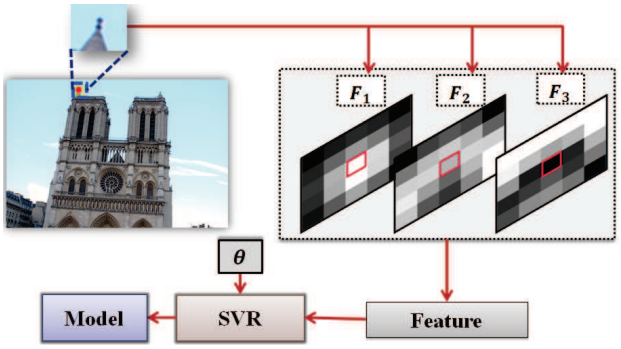


Fig. 3: *Training an SVR*. The sample pixel (red) with $s \times s$ neighborhood (blue) is chosen to extract the features maps (F_1, F_2, F_3) using response scores, gradients and intensity patterns respectively.

where γ_t is a learning rate that can be made to decay with t as $\gamma_t = \gamma_0/(t+1)$ and the gradient for the objective function in Eq. (3) is replaced (as detailed in [13]) with the gradient of a randomly chosen sample pair $\{i, j\}$ at time t, *i.e.*,

$$\nabla \Phi_{\{i,j\}}(\boldsymbol{\theta}_t) \triangleq \frac{\partial \Phi(\mathcal{R}_i, \mathcal{R}_j)}{\partial \boldsymbol{\theta}}\bigg|_{\boldsymbol{\theta}_t}. \quad (6)$$

We computed the gradient required in Eq. (5) using the chain rule as follows,

$$\nabla \Phi_{\{i,j\}}(\boldsymbol{\theta}) = \Big\{ \frac{\partial \Phi}{\partial \mathcal{R}_i} \cdot \frac{\partial \mathcal{R}}{\partial \varphi_i} \cdot \frac{\partial \varphi_i}{\partial \boldsymbol{\theta}}, \frac{\partial \Phi}{\partial \mathcal{R}_j} \cdot \frac{\partial \mathcal{R}}{\partial \varphi_j} \cdot \frac{\partial \varphi_j}{\partial \boldsymbol{\theta}} \Big\} \quad (7)$$

The simplified implementation procedure is summarized in Algorithm 1.

Algorithm 1 Generation of training set

Require: a scene S with N images and the set of possible image pairs P drawn from S.
K:= number of image pairs in P.
$epochs$:= number of passes over the set P.
Ensure:
for $iters = 1 : epochs$ **do**
Shuffle the order of n pairs in P
for $pair = 1 : K$ **do**
Compute $\nabla \Phi_{pair}$ (as in Eq. (7)).
Update $\boldsymbol{\theta}$ (as in Eq. (5)).
end for
end for

2.3. Support Vector Regressor Training for AdTMO

An illustration for SVR training is shown in Fig 3. Let's assume that a scene with multiple images captured under different lighting variations is given for training. Further, assume that the optimal modulation maps for the same scene are also given as described in Section 2.2.

To train an SVR model invariant to illumination variations, we first select random samples from keypoint and neighborhood locations across all the given images with varying lighting conditions. For each sample, we then consider a local patch of size $s \times s$ centered at that pixel. Next, we compute a feature vector which includes: a)



Poster Invalides Project-Room Light-Room Notre-Dame Camroom Grande-Arche Louvre

Fig. 4: Sample images from *HDR dataset*. The *HDR Dataset* is composed of 8 scene from different indoor/outdoor locations.

the second-order detector response scores F_1, b) the gradient magnitudes F_2 and c) the local intensity patterns F_3.

The second-order response scores are based on the choice of the keypoint detector. Therefore, our response score feature for each pixel x in patch $s \times s$ is given as: $F_1(x) = det\{\boldsymbol{M}(x)\} - k \cdot tr\{\boldsymbol{M}(x)\}^2$. The gradient magnitudes for each pixel in the local patch is computed as: $F_2(x) = \sqrt{G_x^2(x) + G_y^2(x)}$, where G_x and G_y are the gradients in horizontal and vertical directions. The local intensity patterns for each patch is recorded by subtracting the value of centered pixel from other pixels and given as: $F_3(x) = I(x) - I(c)$, where c is the pixel at center location.

These individual features are normalized and concatenated to form the final feature vector $\{F_1, F_2, F_3\}$ of dimension $3s^2$ representing a training sample.

3. RESULTS AND DISCUSSION

3.1. Dataset

In this paper, we propose an HDR dataset with 8 different HDR scenes as shown in Fig. 4. The *Light Room*, *Project Room* and *Poster* are the publicly available datasets and have been used for evaluating HDR for keypoint detection problems [3,20]. However, these 3 scenes have been captured in indoor locations and hence, they are less challenging in terms of physical-world illumination transformations such as day/light change. Therefore, we captured 5 additional scenes including 1 indoor *Camroom* and 4 famous outdoor locations in Paris: *Notre-Dame*, *Louvre*, *Invalides* and *Grande Arche*. The *Camroom* scene is shot with a Canon Mark III camera in the presence of powerful 2K Watt reflectors. All the other outdoor HDR scenes are captured with Canon 700D camera at different times of the day. To create the HDR images, LDR images have been fused using the algorithm in [21]. Note that all scenes are geometrically calibrated.

3.2. Experimental Setup

We test our proposed model for keypoint detection task on 8 HDR scenes. We initially compare our AdTMO with the non-adaptive bilateral filtering based tone mappings BTMO and its globally optimized version BTMO(opt) [9]. Similar to our tone mapping function, both these TMOs are based on illumination normalization where the luminance L is estimated using bilateral filter. However, both these TMOs use global range and spatial variances. Moreover, BTMO(opt) [9] is a variant of BTMO with an additional step of global parameter optimization, and approximates the maximum possible RR that can be achieved with BTMO model.

Then, we compare our model with state-of-the-art perception based TMOs: ChiuTMO [15], DragoTMO [14], ReinhardTMO [6] and MantiukTMO [7]. We considered these TMOs as they have been previously applied for HDR evaluation studies [5,9] for similar keypoint detection task.

We evaluate all these TMOs using popular and widely used corner detection schemes: Harris [18], Shi-Tom [22], FAST [23] BRISK [24]. In addition, even if our formulation is optimized for corner detection, we also test our TMO with respect to blob detectors such as SURF [25] and SIFT [26]. Since our model is designed for one image scale, we employed single-scale implementation for all keypoint detection schemes to ensure a fair comparison.

The detection performance is measured in terms of RR (as discussed in Section 2.2) with an error rate of 5 pixels. Namely, a keypoint is considered to be repeated in the test image if it lies in a circle of radius 5 centered on the projection of the reference keypoint onto the test image.

3.2.1. *Training and Implementation details*

For each test scene, we build the training set with 10,000 samples and use it to train and validate the SVR model. This training set is drawn from other scenes excluding the corresponding test scene. For instance, to test the Project Room scene, we build the training set by randomly selecting the samples from all other 7 scenes. For each training sample, we compute feature on a small patch size of 5×5 while following the feature extraction procedure from Section 2.3. Higher patch-size is not advisable as pixel correlation diminishes with increasing distance. Conversely smaller patch-size may extract insufficient information.

Implementation. We use the SVR implementation of LibSVM [27] using the Radial Basis Function (RBF) kernel. To obtain the optimal values of SVR parameters, the regularization cost and epsilon-SVR are tuned by 5-fold cross validation from the range of $[2^{-5}, 2^{15}]$ and = $[2^{-10}, 2^{5}]$, respectively.

We use the HDR Toolbox [28] for the implementation of the considered TMOs. Moreover, we use the Matlab's Computer Vision toolbox for Harris, Shi-Tom, FAST, BRISK and SURF, and Vlfeat for SIFT. Similar to previous keypoint evaluation studies [3,20], we selected the strongest 500 keypoints from each test image.

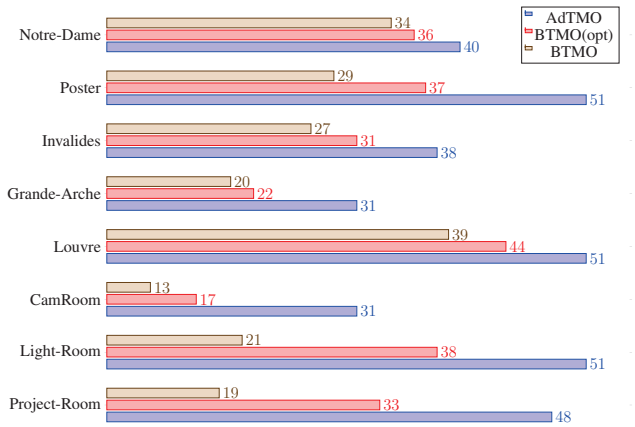


Fig. 5: *Quantitative Results I*: Repeatability Rates (RR) computed using AdTMO, BTMO(opt) and BTMO for each test scene using Harris keypoint detector. Note that while testing AdTMO for a particular scene we assured that the training for AdTMO is done on all other scenes.

3.3. Evaluation Results

Quantitative Results: We perform a thorough evaluation of our proposed AdTMO in quantitative terms as shown in Fig. 5 and Fig. 6. We basically evaluate the performance of our method over all test scenes using the Harris corner detector. In Fig. 5, we compare our model with the other variants of bilateral filtering based TMOs: BTMO and BTMO(opt) [9]. These results clearly show that local modulation of bilateral filtering based tone mapping function using the proposed learned model significantly improves the keypoint stability across both the indoor and outdoor scenes.

Comparison with popular TMOs. We evaluate the performance of our method across different keypoint detection schemes including both corner and blobs. In Fig. 6, we initially compute the RR for all scenes for each considered TMO and then average them to compute the Average Repeatability Rate (AvgRR). We observe that for either detector (corner or blob) our proposed model outperforms all the other TMOs (perception based or keypoint-based). Further, the lower standard deviations observed with our proposed TMO shows higher stability of keypoints than other perception-based TMOs. Although the algorithm presented in this paper has been optimized for corners, it gives comparable or better performance with respect to other methods on blob detectors. This is partially due to the single scale implementation of the blob detectors used in this evaluation. However, the performance may differ when the multi-scale blob detection is taken into account.

We compare our AdTMO with popular and visually pleasing Reinhard TMO [6] and MantiukTMO [7]. In Fig. 7, we show that the our method produces the highest number of repeated keypoints, even though both Reinhard TMO [6] and MantiukTMO [7] produce more visually appealing images.

4. CONCLUSIONS

We propose a new learning based adaptive tone mapping framework to detect stable keypoints under drastic changes of lighting conditions. To this end, we train a Support Vector Regressor using local characteristic features to learn a model which spatially modulates the pixel-wise adaptive TMO. Further, we introduce a simple and effective method for generating the training set to learn the SVR for the given problem. We evaluate our model on our proposed HDR benchmark dataset of indoor/outdoor scenes. Our model significantly outperforms state-of-the-art TMOs on the HDR dataset and also achieve state-of-the-art results across different keypoint detection algorithms. In the future, we plan to extend our model for scale space and explore its usability for real-time problems of tracking and panorama creation.



Fig. 7: *Repeated Keypoints*. Row I: 2 HDR images from *Invalides* scene taken at different day-time. HDR images are displayed after log scaling[14]. Row II: the repeated keypoints using our proposed AdTMO (66 repeated keypoints out of strongest 200 keypoints). Row III: the repeated keypoints using Reinhard TMO (7 repeated keypoints out of strongest 200 keypoints). Row IV: the repeated keypoints using MantiukTMO (5 repeated keypoints out of strongest 200 keypoints).

References

[1] F. Dufaux, P. Le Callet, R. Mantiuk, and M. Mrak, *High Dynamic Range Video: From Acquisition, to Display and Applications*, Academic Press, 2016.

[2] P. Bronislav, A. Chalmers, P. Zemcík, L. Hooberman, and M. Cadík, "Evaluation of feature point detection in high dynamic range imagery," *Journal of Visual Communication and Image Representation*, pp. 141 – 160, 2016.

[3] A. Rana, G. Valenzise, and F. Dufaux, "Evaluation of feature detection in HDR based imaging under changes in illumination conditions," in *IEEE International Symposium on Multimedia (ISM), Miami,USA, December, 2015*, 2015, pp. 289–294.

[4] C. Schmid, R. Mohr, and C. Bauckhage, "Evaluation of interest point detectors," *International Journal of Computer Vision*, pp. 151–172, June 2000.

[5] R. Suma, G. Stavropoulou, E. Stathopoulou, L. V. Gool, A. Georgopoulos, and A. Chalmers, "Evaluation of the effectiveness of HDR tone-mapping operators for photogrammetric applications," *Virtual Archaeology Review*, vol. 7, no. 15, 2016.

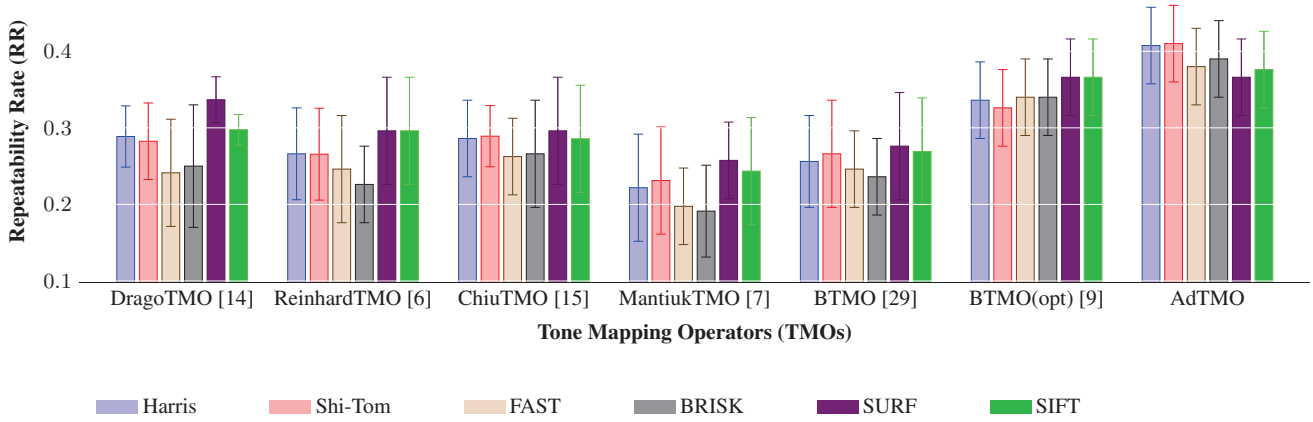


Fig. 6: *Quantitative Results II*: Average Repeatability Rates (AvgRR) computed on different TMOs using various keypoint detection schemes. The average is calculated over all test scenes.

[6] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Transactions on Graphics*, pp. 267–276, July 2002.

[7] R. Mantiuk, K. Myszkowski, and H. P. Seidel, "A perceptual framework for contrast processing of high dynamic range images," *ACM Transactions on Applied Perception*, vol. 3, no. 3, pp. 286–308, July 2006.

[8] A. Rana, G. Valenzise, and F. Dufaux, "An evaluation of HDR image matching under extreme illumination changes," in *The International Conference on Visual Communications and Image Processing (VCIP)*, Chengdu, China, Nov. 2016.

[9] A. Rana, G. Valenzise, and F. Dufaux, "Optimizing Tone Mapping Operators for Keypoint Detection under Illumination Changes," in *2016 IEEE Workshop on Multimedia Signal Processing (MMSP 2016)*, Montréal, Canada, Sept. 2016.

[10] K. S. Ni and T. Q. Nguyen, "Image superresolution using support vector regression," *IEEE Transaction on Image Processing*, vol. 16, no. 6, June 2007.

[11] K. I. Kim and Y. Kwon, "Single-image super-resolution using sparse regression and natural image prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 6, June 2010.

[12] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Transactions on Image Processing*, vol. 16, no. 2, pp. 349–366, 2007.

[13] L. Bottou, *Stochastic Gradient Descent Tricks*, pp. 421–436, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

[14] F. Drago, K. Myszkowski, T. Annen, and N. Chiba, "Adaptive logarithmic mapping for displaying high contrast scenes," *Computer Graphics Forum*, pp. 419–426, 2003.

[15] K. Chiu, M. Herf, P. Shirley, S. Swamy, C. Wang, and K. Zimmerman, "Spatially nonuniform scaling functions for high contrast images," in *Proceedings of Graphics Interface '93*, Toronto, Ontario, Canada, 1993, GI '93, pp. 245–253.

[16] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, no. 3, Aug. 2004.

[17] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Sixth IEEE International Conference on Computer Vision (ICCV)*, 1998, pp. 839–846.

[18] C. Harris and M. Stephens, "A combined corner and edge detector," in *In Proc. of Fourth Alvey Vision Conference*, 1988, pp. 147–151.

[19] A. Rana, J. Zepeda, and P. Pérez, "Feature learning for the image retrieval task," in *Computer Vision - FSLCV, ACCV 2014 - Singapore, November 1-2, 2014*, 2014, pp. 152–165.

[20] P. Bronislav, A. Chalmers, and P. Zemcík, "Feature point detection under extreme lighting conditions," in *Spring Conference on Computer Graphics*, 2012, pp. 156–163.

[21] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*, New York, USA, 1997, SIGGRAPH, pp. 369–378.

[22] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proceedings of the Sixth International Conference on Computer Vision*, Washington, DC, USA, 1998, ICCV '98, pp. 839–.

[23] E. Rosten and T. Drummond, *Machine Learning for High-Speed Corner Detection*, pp. 430–443, Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

[24] S. Leutenegger, M. Chli, and R. Y. Siegwart, "Brisk: Binary robust invariant scalable keypoints," in *Proceedings of the 2011 International Conference on Computer Vision*, Washington, DC, USA, 2011, ICCV '11, pp. 2548–2555.

[25] H. Bay, T. Tuytelaars, and L. V. Gool, "Surf: Speeded up robust features," in *9th European Conference on Computer Vision (ECCV)*, pp. 404–417. 2006.

[26] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[27] C.C. Chang and C.J. Lin, "Libsvm: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, May 2011.

[28] F. Banterle, A. Artusi, K. Debattista, and A. Chalmers, *Advanced High Dynamic Range Imaging: Theory and Practice*, Natick, MA, USA, 2011.

[29] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," in *Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques*, 2002, SIGGRAPH '02, pp. 257–266.