

Extended Dynamic Range Imaging: A Spatial Down-Sampling Approach

Huei-Yung Lin and Jui-Wen Huang

Department of Electrical Engineering and Advanced Institute of Manufacturing with High-Tech Innovation
National Chung Cheng University
168 University Rd., Chiayi 621, Taiwan
Email: hylin@ccu.edu.tw, juiwenhuang@gmail.com

Abstract—In recent years, many works have addressed the issues of generating high dynamic range (HDR) images from the low dynamic range (LDR) counterparts. Since the HDR image contains a broader range of physical values which cannot be recorded by conventional sensors, the previous approaches use a sequence of images captured with different exposures to synthesize an HDR image. In this paper, we propose a spatial down-sampling technique to extend the dynamic range of an LDR image and generate an image with a broader dynamic range. The idea is to trade the large resolution of an image with a large brightness range for intensity quantization, and produce a so-called Extended Dynamic Range (EDR) image. Experimental results demonstrate that our approach is able to provide the better image quality than those derived from the existing LDR to HDR image conversion techniques.

I. INTRODUCTION

Nowadays, people's demands for the quality of life is keep increasing, and one particular aspect is related to the human visual perception of the real world. While the digital imaging devices are the primary source for the reproduction of visual information, the images taken by the conventional cameras do not generally reveal the exact scenes perceived by the human eyes. The main reason is that the commonly used technologies for cameras and image sensors is not able to acquire the full radiance range associated with a real scene. Thus, only a limited dynamic range of intensity values can be recorded in an image. Consequently, the cameras cannot reflect the true brightness of the real scene. On the other hand, a display device with the dynamic range comparable to the human perception is required for faithful scene reproduction. One needs a high dynamic range display to show the HDR images. The manufacture of HDR displays is pioneered by Seetzen [1]. Currently, many manufacturers such as LG, Philips, Samsung and AUO, etc. (e.g. [2]) have produced HDR displays.

In recent years, some techniques have been developed for the generation of high dynamic range images. The current HDR image acquisition methods are mainly divided to three categories— graphical image visualization, multi-exposure image capture, and specialized hardware. The first one is based on the physical analogy of light and illumination model for image composition [3]. It is a main technique to render the HDR images for computer graphics and visualization in early research. The second method takes multiple images with different exposures to synthesize an image of the real scene

with a broader brightness range [4], [5]. The third method adopts one or more image sensors integrated with specialized hardware design to directly capture the images with a larger dynamic range [6].

In the above methods, the first one required the ideal photometric modeling to render the artificial scenes and thus is not suitable for natural images. The third method requires the special equipment which is not yet mature and fairly expensive. Furthermore, the acquired images are not guaranteed to have the high dynamic brightness range perceivable by human eyes. In the second method, only a conventional digital camera is used for image acquisition, followed by software techniques to generate the HDR images. Thus, it is an inexpensive approach for practical applications. Nevertheless, an important issue for this approach is to correctly calibrate the camera response function (CRF). To assemble an HDR image from a series of multi-exposure images, Debevec and Malik proposed a simple method for recovering a CRF through a table which is minimized using a squared error function [4]. Mitsunaga and Nayar improved this algorithm with a more robust method based on a polynomial representation [5]. To avoid the camera response curve calibration, Mertens *et al.* presented a technique to simplify the acquisition pipeline by fusing a bracketed exposure sequence into a high quality image [7]. Note that the multiple exposure method assumes that the images are perfectly aligned, and there are no moving objects in the images. Furthermore, it does not take the sensor noise problem into consideration.

Except for the commonly used multi-exposure approaches, some researchers also investigated the possibility of generating an HDR image using its single LDR counterpart. It is substantially a convenient method to produce HDR images from existing LDR images. The so-called LDR2HDR techniques use the inverse tones mapping operator to extend the LDR image to an HDR image [8], [9], [10]. Banterle's method [8] is based on Reinhard's approach [11] and uses the proposed tone mapping technology to perform the inverse function to compute the extended dynamic range of the image. Meylan's approach [9] is based on his previous techniques [12] for inverse function derivation. Rempel [10] purposed a method for real-time dynamic image processing, such as the input signal of the DVD Players. Compared to the HDR generation using multiple image composition, the LDR2HDR techniques

offer a low-cost solution but only provide the limited quality in terms of the intensity dynamic range.

In this paper, we present a spatial down-sampling approach to extend the dynamic range of the input LDR image and synthesize an image with a greater dynamic range. The idea is to transform the high image resolution information of an input image to the high intensity range information of an output image. Since the images currently available are acquired with very high resolution (e.g. several millions of pixels) but usually displayed on lower resolution devices (monitor, TV, smartphone, etc.), it is reasonable to trade the high image resolution for the high intensity quantization. Thus, the proposed technique is considered as a framework of the LDR2HDR approaches. It takes an original image with high resolution but low dynamic range as input, and generates an output image with extended dynamic range by applying the trade-off between the image resolution and the dynamic range to the existing LDR2HDR approaches. Different from the general HDR image synthesis techniques, our extension for the dynamic range is adjustable via the down-sample factor. It has the advantage of showing the images on the devices with various HDR display capability since the additional tone mapping process can be avoided. The experimental results have demonstrated that, compared to the previous LDR2HDR algorithms, the EDR images generated using the proposed method provide better image quality in terms of visual perception evaluation.

II. EXTENDED IMAGE MAP

The conventional image (i.e., LDR image) is capable of storing 8-bit information per channel. It contains 256 levels of intensity value, which are recorded in the $[0, 255]$ interval. Compared the LDR image with the HDR image (commonly 24-bit or 32-bit per channel), the latter can store more details of the scene, provide more dedicated image quality, and even preserve more information close to the human perception. It has been a standard practice to generate HDR images from many LDR images captured with different exposure times. Recently, some researchers also propose the single LDR image extension for HDR image generation, which improves the convenience of HDR image production. These methods are usually based on the inverse tone mapping operation to generate the HDR images. With the nonlinear (or piecewise linear) expansion of highlights or brightness enhancement, the dynamic range of the image is increased and the intensity values are encoded using a standard HDR image format.

One major problem of the LDR2HDR (inverse tone reproduction) approach is its limitation on brightness enhancement. If the contrast stretching for the target dynamic range is set too large, then the fine detail on intensity quantization requires estimation and sophisticated interpolation techniques. Otherwise, an image with a larger dynamic range should be used as the input for the inverse tone reproduction process to mitigate the “interpolation for quantization” effect. Since the intensity range is fixed by 256 levels for LDR images, no matter how large is the range of the captured image irradiance,

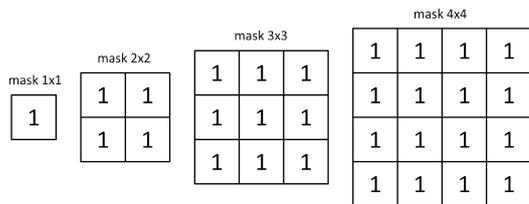


Fig. 1. Different mask size for image down-sampling.

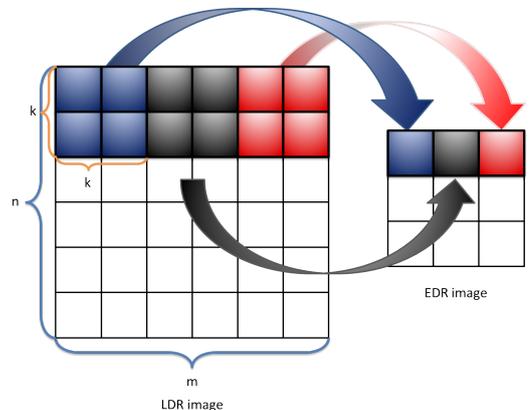


Fig. 2. Schematic diagram of the EDR image generation.

it is not possible to provide additional information beyond the quantized values. To make the intensity range greater than 256 levels, one can combine the current pixel with its $n - 1$ neighboring pixels to form a super-pixel with the intensity range approximately n times larger. Based on this idea, an image down-sampling approach can be carried out to generate a new image with a greater intensity range systematically, at the cost of image resolution reduction.

Without loss of generality, assume an n -bit LDR image I with the resolution of $m \times n$ pixels is given, and a $k \times k$ mask is used to combine the k^2 neighboring pixels for the down-sampling process.¹ For simplicity, let both m and n be divisible by k . Then the dimension of the resulting EDR image, denoted by I_k , is reduced to $(m/k) \times (n/k)$, and the intensity range is given by $k^2(2^n - 1) + 1$. The down-sample masks with different sizes and a schematic diagram of the LDR to EDR conversion are illustrated in Figs. 1 and 2, respectively. It should be noted that, different from the floating number representation for the HDR image format encoding, the EDR image is still represented by the unsigned integer. The LDR2EDR conversion merely generates an $(n + 2 \log_2 k)$ -bit conventional image (per color channel). It can either directly transform to a floating point number format for the HDR image representation, or pass to an LDR2HDR algorithm to generate an HDR image.

Fig. 3 shows the histograms of an LDR image and the EDR images generated by down-sampling with 2×2 , 3×3 and

¹For most digital images, there are 8 bits per channel (i.e. $n = 8$). However, n can be 12 or 14 for the RAW format used in many modern DSLR cameras.

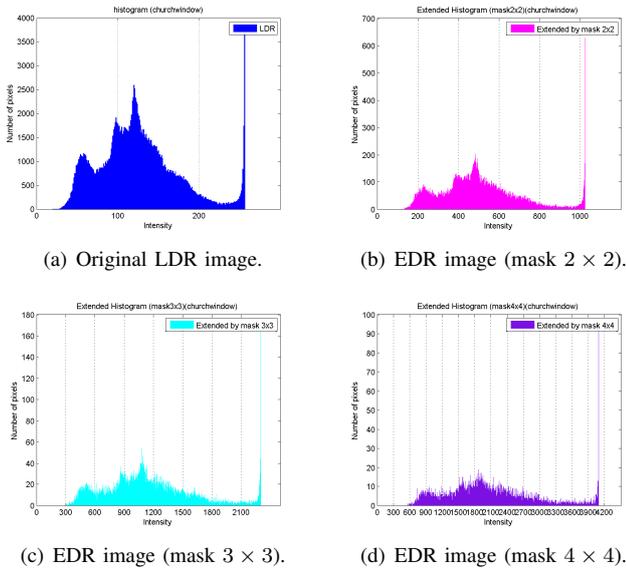


Fig. 3. The histograms of the original LDR image and the EDR images generated using various down-sampling masks.

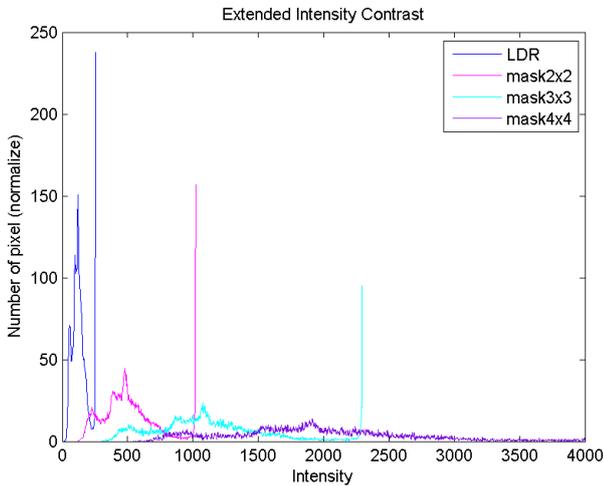


Fig. 4. Extended intensity contrast of the EDR images.

4×4 masking. The intensity range is extended from 256 to 1021, 2296 and 4081, respectively. Comparing Figs. 3(a), 3(b), 3(c) and 3(d), it is clear that the image histogram becomes flatter when the intensity range is increased by using a larger down-sampling mask. This process is similar to the image equalization technique for contrast enhancement, except that the histogram is usually not uniformly distributed due to the saturation and zero value of image pixels. Fig. 4 shows the extension of the intensity range from 256 to 4801 without normalization. After the dynamic range extension, the image intensity distribution moves to the right, which makes the intensity range 4, 9 and 16 times wider, respectively, and provides more dedicate quantization results.

Due to the masking process for down-sampling, the image resolution becomes smaller. For the mask size of $k \times k$, the

image resolution is reduced by a factor of $1/k^2$. Nowadays, commercially available smartphone embedded cameras commonly have ten million pixels, and consumer digital cameras can even support up to twenty million pixels. While the image with higher pixels is favorable for many uses, it has attendant disadvantages: the images require more memory capacity and more powerful processors for image processing tasks. Furthermore, the camera has been provided enough pixels for the $8'' \times 10''$ photos, but few people need to print such large photos. On the other hand, there was a variety size of sensors for an ideal number of pixels a few years ago. The corresponding optimal pixels for an 1/1.8 sensor was four million pixels. Now, even a sensor equipped with a 1/2.33 CCD contains 2 times more pixels. Thus, understandably, the sensor noise should get even worse. Our approach sacrifices the image resolution but enables a low dynamic range being widely extended.

III. HIGH DYNAMIC RANGE IMAGE SYNTHESIS

The final stage of our extended dynamic ranging imaging pipeline is to take the EDR image obtained previously as input to render a high dynamic range image output. To generate an HDR image with the dynamic range $[L, H]$ from an EDR image I_k , one simple approach is to apply a contrast stretching function

$$f_l(i) = L + \frac{H - L}{k^2(2^n - 1) + 1} \cdot i \quad (1)$$

for linear mapping between the intensity range $k^2(2^n - 1) + 1$ and the target radiance range $[L, H]$, where i is the intensity of the EDR image. Although an HDR image with the radiance map of order H/L can be derived, the mid-range of brightness tends to be compressed for human perception. Thus, it is not suitable for general HDR image synthesis.

In this work, we adopt the LDR2HDR methods to synthesize the HDR images. Unlike using a group of multiple exposure image sequence for HDR image synthesis, these methods use only a single image under normal exposure and the inverse tone mapping operator to expand the LDR image to an HDR image. In our implementation, a single LDR image is used to generate the EDR image, which is then passed to the LDR2HDR algorithm to synthesize an HDR image. The system flowchart, including the image quality evaluation, is shown in Fig. 5. The input LDR image (with the resolution of 750×1130) used in the experiments is shown in Fig. 6(a). An HDR image obtained from the multi-exposure image composition technique is shown in Fig. 6(b) for reference.² By applying the different sizes of the down-sample mask, 2×2 , 4×4 and 8×8 , we obtain the EDR images with the resolution of 375×565 , 187×282 and 94×142 , respectively.

To synthesize the HDR image from an EDR image using the LDR2HDR system, three algorithms, namely Banterle, Meylan and Rempel, are implemented in this work. In Banterle's

²Due to the limited dynamic range of most monitor and the current file format, the HDR image is tone-mapped for the display.

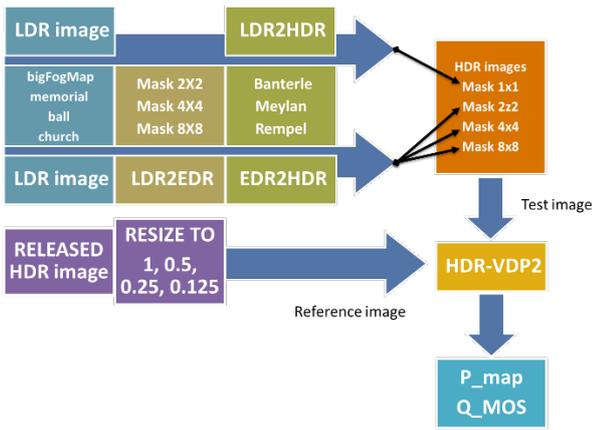


Fig. 5. The system flowchart of the proposed HDR image synthesis technique.



(a) LDR input image.

(b) HDR reference image.

Fig. 6. The input LDR image and the reference HDR image.

approach [8], they found a way to approximately invert Reinhard's photographic tone mapping operator [11]. An *expand-map* is then generated for the regions with high luminance using the median-cut based iTMO (inverse tone-mapping operator). Meylan *et al.* proposed a tone scale function that takes advantage of the increase in dynamic range of HDR monitors to recreate the brightness of specular highlights [9]. They were then clipped or compressed by the capturing and rendering process to a standard dynamic range. Rempel *et al.* described a method for boosting the dynamic range of legacy video and photographs for viewing on high dynamic range displays [10]. The nonlinear intensity encoding of LDR images is first compensated by an inverse gamma mapping, followed by the computation of a brightness enhancement function for the saturated image regions.

Fig. 7 shows the HDR image representation obtained using the proposed approach. Figs. 7(a), 7(b), 7(c) and 7(d) show the results derived from the EDR images down-sampled by the mask size of 1×1 (without down-sample), 2×2 , 4×4 and 8×8 , respectively. For each EDR resolution specification, the left, middle and right images are generated by the LDR2HDR

systems based on the techniques of Banterle, Meylan and Rempel, respectively. All images are tone-mapped using Reinhard's photographic tone mapping operator for display and resized to the same resolution for comparison.

IV. EXPERIMENTS AND EVALUATION

To evaluate the faithfulness of our HDR synthesis results, an image quality assessment method, HDR-VDP2, proposed by Mantiuk *et al.* is adopted [13]. It is designed for predicting the correlation of a pair of HDR images. The test image is calculated not only with respect to the reference image for the degree of similarity, HDR-VDP2 also estimates how the human eyes can distinguish two images in HDR displays. In our performance evaluation, the HDR radiance map released by Ward as the reference HDR image [14]. Because the use of HDR-VDP2 requires the same resolution of the reference and test images, when exploring the EDR data we resize the reference image by bicubic interpolation to create the same resolution of EDR images.

Fig. 8 shows the maps of probability of detection, which are defined by the probability of the differences between the images visible to an average observer. The evaluation results in Figs. 8(a) – 8(d) correspond to the HDR images generated from the down-sampled EDR images as shown in Figs. 7(a) – 7(d), respectively. The map of probability of detection tells us how likely we will notice the difference between the two images. The red color denotes the high probability, and the blue color denotes the low probability. For the HDR images generated using Banterle's and Meylan's LDR2HDR techniques, the results in Fig. 8(b) using 2×2 mask are better than those in Fig. 8(a) using 1×1 mask. Counterintuitively, the results in Figs. 8(c) using 4×4 mask and 8(d) using 8×8 mask are also worse than those in Fig. 8(b). That is, our EDR input derived from the 2×2 down-sample mask provides the best HDR image synthesis results when Banterle's and Meylan's methods are applied. If Rempel's LDR2HDR algorithm is used, the EDR image obtained from 4×4 down-sample mask provides the best HDR image synthesis results in terms of the probability of detection, as illustrated in the last column images of Fig. 8.

The above evaluation results demonstrate that, if the current LDR2HDR algorithms are adopted, the HDR image synthesized using our EDR data (either 2×2 or 4×4) is better than the result obtained using the original LDR image as input. Table I shows the quality degradation measured with respect to the reference image and expressed as a mean-opinion-score (MOS), Q_MOS . A high score indicates the high degree of correlation in terms of image quality assessment. In this experiment, we use six images ('bigFogMap' [14], 'memorial' [15], 'ball'³, 'church', 'churchwindow', 'lobby'⁴) for evaluation. The table shows that the bigFogMap image has the highest degree of similarity when the mask size of 8×8 is used. We can also see that Q_MOS is not necessarily

³Image source: <http://nada.cps.unizar.es/pub/>

⁴Image source: <http://pages.cs.wisc.edu/~csverma/>

TABLE I
IMAGE QUALITY CORRELATION

Image	mask	Banterle	Meylan	Rempel
bigFogMap	1 × 1	15.03	10.27	35.08
	2 × 2	12.33	10.76	42.89
	4 × 4	7.32	9.10	39.35
	8 × 8	15.93	17.11	63.69
memorial	1 × 1	11.54	7.18	22.96
	2 × 2	7.24	3.87	15.97
	4 × 4	6.95	4.32	17.87
	8 × 8	15.8	7.45	38.29
ball	1 × 1	9.77	16.00	46.63
	2 × 2	7.51	9.27	40.36
	4 × 4	12.17	12.71	61.41
	8 × 8	33.24	28.71	89.14
church	1 × 1	8.63	37.48	8.59
	2 × 2	8.40	35.45	5.40
	4 × 4	15.25	65.91	9.19
	8 × 8	30.92	93.80	23.11
churchwindow	1 × 1	2.59	73.33	17.26
	2 × 2	2.19	74.54	13.24
	4 × 4	3.90	93.44	28.73
	8 × 8	7.81	98.95	69.36
lobby	1 × 1	6.72	28.53	13.83
	2 × 2	8.13	32.17	14.84
	4 × 4	23.34	61.65	43.78
	8 × 8	31.77	90.73	81.65

improved when the mask size is increased. For example, in memorial case, Q_MOS reaches its maximum value for the 4×4 mask.

The test images ‘church’, ‘churchwindow’ and ‘lobby’ are from Verma’s dataset. Since the HDR radiance maps are not released, we build our own HDR image using the multiple exposure images in the dataset, and use it as the HDR-VDP2 reference image. When Banterle’s LDR2HDR technique is adopted, the results for these images are similar— they have the best Q_MOS for 2×2 masking. However, the maximum Q_MOS values for Meylan’s and Rempel’s methods are given by the largest mask size (8×8) in the experiment. This suggests that a better image quality might be achieved if a even bigger mask is used to generate the EDR data for HDR image synthesis. Nevertheless, if Rempel’s LDR2HDR method is used, the similarity has been improved with a great progress from 13.83 to 81.65.

In general, we consider the image resolution as an important factor of image quality. The smaller image gets a good quality score compared to a large image because the picture might be too small for the human eyes to distinguish the details. Table II shows the relationship between the image size and the evaluated image quality. The Q_MOS values for HDR-VDP2 are derived with respect to the reference image. In Table II, a single LDR image ‘bigFogMap’ is used to synthesize the HDR image with Banterle’s inverse tone mapping operator [8]. The first Q_MOS score 15.03 in the table is the same as the result in Table I using the 1×1 mask. After the test image is resized by interpolation using a bicubic algorithm, the Q_MOS values is increased. However, the ascending trend is

TABLE II
RELATIONSHIP BETWEEN THE IMAGE SIZE AND THE IMAGE QUALITY.

Image	resize ratio	Q_MOS
bigFogMap	1	15.03
	0.5	11.99
	0.25	7.24
	0.125	12.45

not as fast as those derived using the down-sampling masks as shown in Table I. This indicates that using the EDR data derived by the proposed technique can provide a better HDR image synthesis result.

V. CONCLUSIONS

One simple method to generate an HDR image is to combine multiple conventional LDR images acquired with different exposures. In practical uses such as the image sequence acquisition, it is not always possible to shoot the same scene with multiple image captures. In this paper, we present a technique to extend the dynamic range of the input LDR image, and use it to synthesize an HDR image. Using the EDR data derived from the down-sampled image, we are able to generate the HDR image with more delicate information compared to those directly obtained from the LDR2HDR algorithms. The experimental results and quantitative evaluation have demonstrated the effectiveness of the proposed EDR method for HDR image generation.

REFERENCES

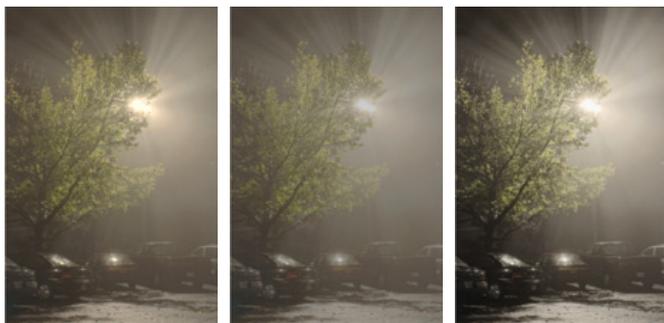
- [1] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh, and A. Vorozcovs, “High dynamic range display systems,” *ACM Transactions on Graphics (TOG)*, vol. 23, no. 3, pp. 760–768, 2004.
- [2] “Philips showcased latest lcd backlighting improvements at fpd 2006,” 2006. <http://preview.tinyurl.com/yhc4nb>.
- [3] G. W. Larson, R. Shakespeare, C. Ehrlich, J. Mardaljevic, E. Phillips, and P. Apian-Bennewitz, *Rendering with radiance: the art and science of lighting visualization*. Morgan Kaufmann San Francisco, CA, 1998.
- [4] P. E. Debevec and J. Malik, “Recovering high dynamic range radiance maps from photographs,” in *ACM SIGGRAPH 2008 classes*, p. 31, ACM, 2008.
- [5] T. Mitsunaga and S. K. Nayar, “Radiometric self calibration,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1999.*, vol. 1, IEEE, 1999.
- [6] S. K. Nayar and T. Mitsunaga, “High dynamic range imaging: Spatially varying pixel exposures,” in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition, 2000.*, vol. 1, pp. 472–479, IEEE, 2000.
- [7] T. Mertens, J. Kautz, and F. V. Reeth, “Exposure fusion,” in *Proceedings of the 15th Pacific Conference on Computer Graphics and Applications, PG ’07.* (Washington, DC, USA), pp. 382–390, IEEE Computer Society, 2007.
- [8] F. Banterle, P. Ledda, K. Debattista, and A. Chalmers, “Inverse tone mapping,” in *Proceedings of the 4th international conference on Computer graphics and interactive techniques in Australasia and Southeast Asia*, pp. 349–356, ACM, 2006.
- [9] L. Meylan, S. Daly, and S. Susstrunk, “The reproduction of specular highlights on high dynamic range displays,” *Color and Imaging Conference*, vol. 2006, no. 1, pp. 333–338, 2006-01-01T00:00:00.
- [10] A. G. Rempel, M. Trentacoste, H. Seetzen, H. D. Young, W. Heidrich, L. Whitehead, and G. Ward, “Ldr2hdr: on-the-fly reverse tone mapping of legacy video and photographs,” *ACM Transactions on Graphics (TOG)*, vol. 26, no. 3, p. 39, 2007.



(a) The results derived from 1×1 down-sampled EDR images.



(b) The results derived from 2×2 down-sampled EDR images.

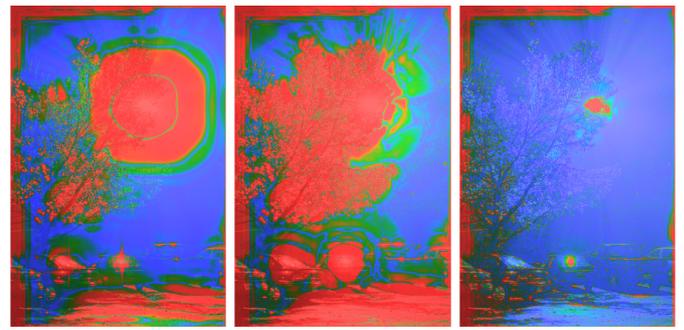


(c) The results derived from 4×4 down-sampled EDR images.

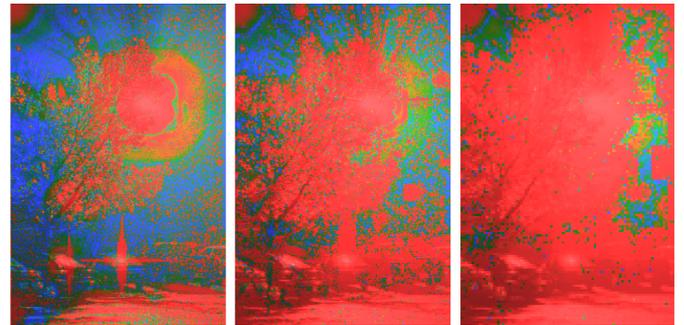


(d) The results derived from 8×8 down-sampled EDR images.

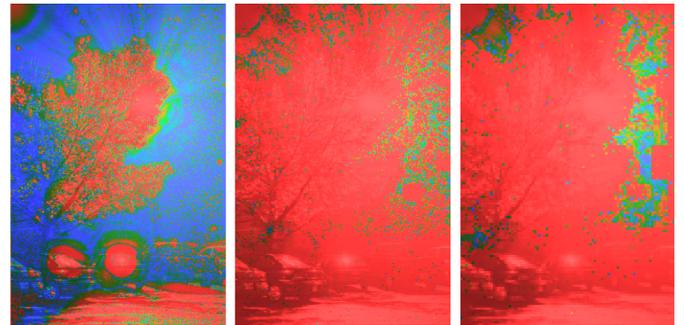
Fig. 7. The HDR images obtained using the proposed approach. From the left to the right: Banterle, Meylan and Rempel approaches for LDR2HDR.



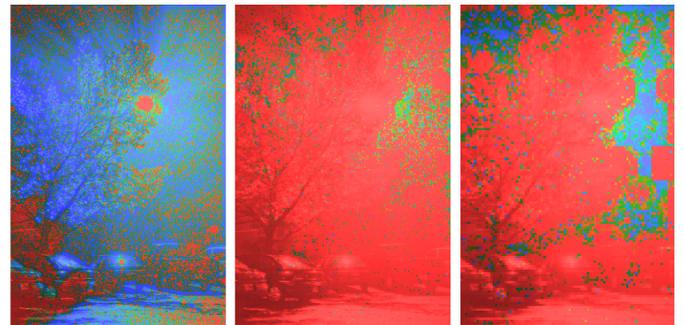
(a) The probability of detection associated with 1×1 down-sample.



(b) The probability of detection associated with 2×2 down-sample.



(c) The probability of detection associated with 4×4 down-sample.



(d) The probability of detection associated with 8×8 down-sample.

Fig. 8. HDR-VDP2 probability of detection for the images in Fig. 7. From the left to the right: Banterle, Meylan and Rempel approaches for LDR2HDR.

[11] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Trans. Graph.*, vol. 21, pp. 267–276, July 2002.

[12] L. Meylan, "Tone mapping for high dynamic range images," tech. rep., Ph. D. thesis, EPFL, Lausanne, Switzerland, 2006.

[13] R. Mantiuk, K. J. Kim, A. G. Rempel, and W. Heidrich, "Hdr-vdp-2: A calibrated visual metric for visibility and quality predictions in all

luminance conditions," *ACM Transactions on Graphics (TOG)*, vol. 30, no. 4, p. 40, 2011.

[14] G. Ward, "High dynamic range image encodings," *Anywhere Software*, 2005.

[15] R. Fattal, D. Lischinski, and M. Werman, "Gradient domain high dynamic range compression," *ACM Trans. Graph.*, vol. 21, pp. 249–256, July 2002.