HIGH DYNAMIC RANGE IMAGING FOR STEREOSCOPIC SCENE REPRESENTATION

Huei-Yung Lin and Wei-Zhe Chang

Department of Electrical Engineering
National Chung Cheng University
Chia-Yi 621, Taiwan
E-mail: hylin@ccu.edu.tw, ilct0308@hotmail.com

ABSTRACT

This paper presents a method for generating high dynamic range and disparity images by simultaneously capturing the high and low exposure images using a pair of cameras. The proposed stereoscopic high dynamic range imaging technique is able to record multiple exposures without any time delay, and thus suitable for high dynamic range video synthesis. We have demonstrated that it is possible to construct the camera response function using a pair of images with different amount of exposure. The intensities of the stereo images can then be normalized for correspondence matching. Experiments using the Middlebury stereo datasets are presented.

Index Terms— high dynamic range image, stereoscopic

1. INTRODUCTION

High dynamic range imaging (HDRI) is an emerging research topic, especially in computer vision, computer graphics and imaging science communities. The objective is to capture, represent and display the significant amount of luminance variation of the real-world scenes. This problem commonly arises from the limited dynamic range capability of current imaging devices, compared to the much wider perception range of the human visual system. Since the HDRI is aimed to span a large luminance range for vivid image representation and reproduction, it has many applications on, for example, cinematography, photography, computer graphics and visualization, etc.

The most intuitive way of capturing HDR images is to adopt special light detecting sensors capable of recording images with extended intensity range. However, commercially available HDRI systems are usually very expensive and still only accessible to the professional users. For this reason, many researchers have proposed various HDRI techniques using low-cost imaging devices [1]. The most popular approach is to synthesize an HDR image using several image captures under different amounts of exposure [2, 3]. Although easy to implement, one major drawback of this method is that the ghost phenomenon may be present in the synthesized HDR image. This is usually inevitable especially for fusing multiple exposures of a dynamic scene [4].

In addition to the increasing demand of high quality reproduction of real-world scenes, there is always a need for stereoscopic image representation for the binocular human visual system. The objective is to provide us the depth perception ability using stereo image disparities. Thus, there are many advantages and a large application domain available from combining HDRI with stereo vision. For example, this so-called *stereoscopic HDRI* or *SHDRI* can be used to generate the content of 3D TV, video games, augmented reality, or used for advanced video surveillance.

In this work, we propose an SHDRI technique by simultaneously capturing the high and low exposure images using a pair of cameras. Since the stereo image pair can be acquired without any time delay in this approach, the ghost phenomenon introduced by dynamic scene HDR image synthesis is mitigated. Furthermore, the disparity maps associated with the HDR images can be obtained by some readily accessible stereo matching algorithms. To facilitate the correspondence searching with comparable stereo image brightness and display the HDR images on low dynamic range (LDR) devices, a tone-mapping algorithm based on the estimated camera response function is presented.

Although there is a fairly large amount of literature related to the HDRI research, to the authors' best knowledge no previous work on stereoscopic HDRI has been addressed. We demonstrate in this research that it is possible to construct the camera response function using a stereo image pair with different amount of exposure. The HDR and disparity images can then be obtained based on the derived camera response function. Experimental results using the Middlebury stereo datasets [5] have illustrated the feasibility of our approach.

2. HDRI BACKGROUND

Due to the limited dynamic range of conventional imaging sensors, it is not possible to fully capture a high contrast scene

The support of this work in part by the National Science Council of Taiwan, under Grant NSC-96-2221-E-194-016-MY2 is gratefully acknowledged.

with a single exposure using standard cameras. One simple way to create an HDR image is combining several LDR exposures based on the knowledge of the camera response function. In general, the camera response characteristics are not provided by the manufacturers. Thus, we adopt the classic method proposed by Debevec and Malik [2] to derive the intensity response curve, with some modifications on sample point selection as described in the following section. The unknown scene radiance can then be recovered using the intensities of the acquired images and the associated camera exposure settings.

The HDR image generated from multiple LDR exposures usually suffers from noise, blurring, and ghost phenomenon, etc. In the previous work, Grosch proposed a ghost removal algorithm by eliminating the moving object in the scene based on the difference of median threshold bitmaps of two consecutive exposures [6]. Since the noise can easily affect the threshold result, especially for the images with low exposure, the method usually requires to set the threshold manually. Akyuz and Reinhard presented a noise reduction method by replacing the low exposure images with multiple weighted high exposure ones [7]. Different from the previous approaches, except for the issue on noisy image captures under low illumination conditions, we will also have to deal with the ghost image problem of SHDRI due to stereo mismatches.

Although the HDR radiance map can be used to represent an HDR image, most of the existing display devices and media are not capable of covering the full dynamic range. Compressing the HDR images for a LDR display is known as the tone-reproduction or tone-mapping problem, and there is a number of techniques with remarkable performance. In this work, the algorithm proposed by Drago [8] is used to generate the tone-mapped images for display and correlation-based stereo matching.

3. THE ALGORITHM

This section presents the proposed technique for stereoscopic high dynamic ranging imaging. Two cameras placed side-by-side as a conventional stereo configuration are used for image acquisition. Each stereo image pair acquired with different shutter speed is used for HDR image synthesis and stereo disparity computation.

3.1. Deriving the Camera Response Function

To derive the camera response function from a single viewpoint with multiple exposures, it is necessary to obtain the intensities associated with the same image pixel. In the stereoscopic high dynamic range imaging, however, a scene point is not necessarily projected to the same pixel for two images. Thus, it is mandatory to find the feature correspondences between the images and then used as the sample points for camera response function derivation. In this work, the SIFT descriptor [9] is used to select the sample points for camera response curve modeling. An example of the correspondence feature extraction is shown in Fig. 1.

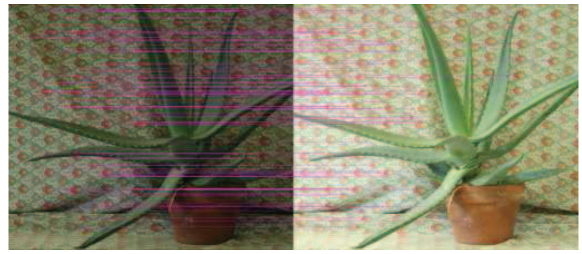


Fig. 1. Correspondence feature extraction using SIFT.

Ideally the camera response curve can be derived using the obtained feature correspondences and the associated image intensities. However, there might exist incorrect correspondences due to noise or stereo mismatch. Based on the SHDRI setting with a conventional stereo configuration, the epipolar, ordering and exposure constraints are used to increase the robustness of correspondence matching. Moreover, the feature points with less intensity variation in the neighborhood are selected as sample points for deriving the camera response curve. A window size of 11×11 pixels is used to calculate the variance associated with each feature correspondence. Only those feature points with variance smaller than a given threshold are qualified as sample points.

To ensure that the camera response curve is well sampled for the full brightness range, the above feature points are further divided into 10 equally-spaced groups based on their intensities, and a fixed number of samples is selected from each group.

3.2. Image Normalization and Correspondence Matching

In this work we adopt the stereo matching algorithm based on belief propagation to derive the disparity map [10]. Since the correlation should be evaluated on the images captured with similar illumination conditions, the stereo image pair captured under different exposures should be normalized prior to the correspondence matching process. Using the camera response function derived from the previous section, one image can be transformed to another with a different exposure setting. Thus, by transforming the image pair to the ones with the same exposure, the stereo matching algorithm can be carried out straightforwardly.

Fig. 2 (left two images) shows the normalized stereo image pairs corresponding to the low, median and high exposures (top, middle and bottom figures). The disparity images obtained from stereo matching using belief propagation are shown in Fig. 2 (right figures) for all three cases. It can be seen that there are no significant differences among the disparity images. This also verifies the correctness of the derived camera response function. Since the disparity of each pixel

(a) Normalized low exposure image pair and computed disparity map.

(b) Normalized median exposure image pair and computed disparity map.

(c) Normalized high exposure image pair and computed disparity map.

Fig. 2. Normalized stereo image pairs and the resulting disparity images.

between the image pair is provided by the disparity map, it is possible to transform one image to align with the other. An HDR image can then be obtained by calculating the radiance map using a weighted sum of these two images.

3.3. Ghost Removal and Tone-Mapping

Due to noise or stereo mismatches, the disparity image does not in general provide correct correspondences for all pixels. Using wrong correspondences to generate the HDR image will cause misalignment of the pixels and result in the ghost phenomenon as shown in Fig. 3 (left figures). Note that this is different from the ghost phenomenon appeared in the conventional HDRI with single viewpoint multiple exposures [11, 12], and therefore moving object detection is not required for ghost removal.

To remove the ghost phenomenon, we first transform the image with the viewpoint for HDR image synthesis to an image under the same exposure with the other image using the camera response curve. If the intensity difference between the corresponding pixels of the two images with the same exposure is larger than a threshold, then the correspondence is classified as a stereo mismatch. In this case, the intensity of the mismatch pixel is given by the newly created image instead of the weighted sum of the stereo image pair. Fig. 3 (right images) illustrates the synthesized HDR image using the proposed ghost removal method.

The final step of HDRI for LDR display is the tone-reproduction or tone-mapping process. Since the tone-mapping stage takes the HDR image as input and is irrelevant to how the image is created, the proposed SHDRI uses the existing techniques for tone-reproduction. The tone-mapped SHDR images in this work are generated using the method presented by Drago [8].

4. RESULTS

The proposed SHDRI technique has been tested using the Middlebury stereo datasets. There are six parallel views with three different exposures in the "art" dataset, and we select two different views with different exposures for SHDRI. Figs. 4(a) and 4(b) show the left and right image with exposures of 1 second and 4 seconds, respectively. The camera response curve derived from the stereo image pair is illustrated in Fig. 4(c). Red, blue and green colored curves correspond to the individual color channels.

Figs. 5(a) and 5(b) show the HDR and disparity images obtained using the proposed methods, respectively. For comparison, the HDR generated using the conventional method with single viewpoint multiple exposures and the disparity image obtained from stereo image pair with same exposure are shown in Fig. 6(a) and 6(b), respectively. The PSNR for the disparity image is 44.65 dB, which indicates the difference between the results obtained from these two implementations is fairly limited.

5. CONCLUSION

In this work, we have presented a method to generate high dynamic range and disparity images using a stereo image pair with different exposures. Although there is a fairly large num-

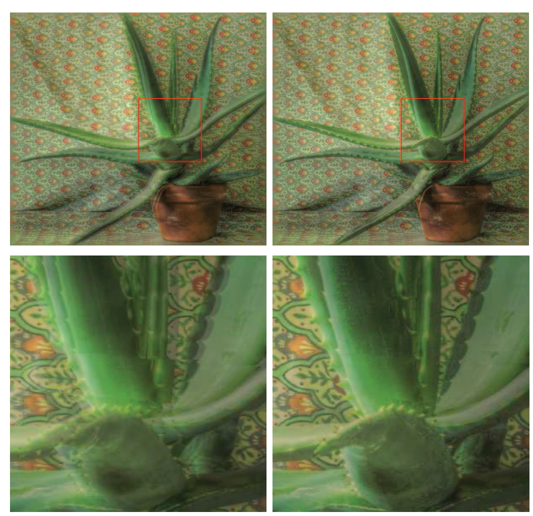
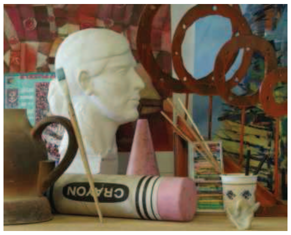
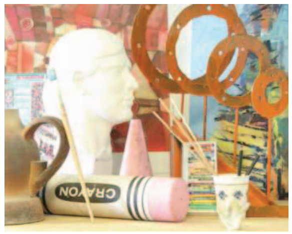


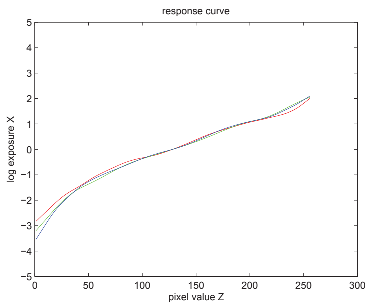
Fig. 3. Ghost phenomenon due to stereo mismatch (left images) and the ghost removal results (right images).



(a) Left image.

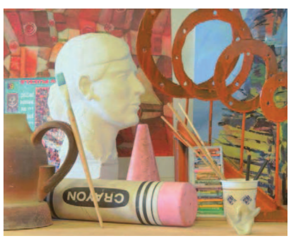


(b) Right image.



(c) Camera response curve.

Fig. 4. Experiment using the "art" images.

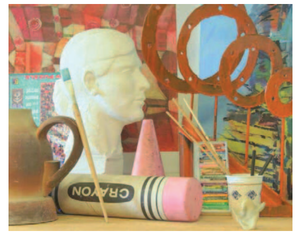


(a) Synthesized HDR image.

(b) Disparity image.

Fig. 5. The results obtained by the proposed SHDRI. Two images with different viewpoints, different exposures are used.

ber of literature related to HDRI research, no previous work on stereoscopic HDRI has been addressed. The camera response function is derived based on the stereo image pair, which is then used to normalize the images for correspondence matching. We have demonstrated the feasibility of the proposed technique using standard stereo image datasets.



(a) Synthesized HDR image.

(b) Disparity image.

Fig. 6. The results obtained by conventional methods. (a) From two images of the same viewpoint with different exposures. (b) From a stereo image pair with the same exposure.

References

[1] S.K. Nayar and T. Mitsunaga, "High dynamic range imaging: Spatially varying pixel exposures," in *IEEE Computer Vision and Pattern Recognition*, 2000, pp. I: 472–479.

[2] Paul E. Debevec and Jitendra Malik, "Recovering high dynamic range radiance maps from photographs," in *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques*. 1997, pp. 369–378, ACM Press.

[3] Sing Bing Kang, Matthew Uyttendaele, Simon Winder, and Richard Szeliski, "High dynamic range video," in *SIGGRAPH '03: ACM SIGGRAPH 2003 Papers*, New York, NY, USA, 2003, pp. 319–325, ACM.

[4] Katrien Jacobs, Céline Loscos, and Greg Ward, "Automatic high-dynamic range image generation for dynamic scenes," *IEEE Computer Graphics and Applications*, vol. 28, no. 2, pp. 84–93, 2008.

[5] D. Scharstein and R. Szeliski, "Middlebury stereo vision page," *http://www.middlebury.edu/stereo*, 2002.

[6] T. Grosch, "Fast and robust high dynamic range image generation with camera and object movement," in *Vision, Modeling and Visualization, RWTH Aachen*, 2006, pp. 277–284.

[7] A.O. Akyuz and E. Reinhard, "Noise reduction in high dynamic range imaging," *Journal of Visual Communication and Image Representation*, vol. 18, no. 5, pp. 366–376, October 2007.

[8] F. Drago, W.L. Martens, K. Myszkowski, and N. Chiba, "Design of a tone mapping operator for high dynamic range images based upon psychophysical evaluation and preference mapping," *IS&T/SPIE Electronic Imaging*, 2003.

[9] David G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[10] J. Sun, N.N. Zheng, and H.Y. Shum, "Stereo matching using belief propagation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 7, pp. 787–800, July 2003.

[11] K. Jacobs, C. Loscos, and G. Ward, "Automatic high-dynamic range image generation for dynamic scenes," *Computer Graphics and Applications, IEEE*, vol. 28, no. 2, pp. 84–93, March-April 2008.

[12] E.A. Khan, A.O. Akyuz, and E. Reinhard, "Ghost removal in high dynamic range images," in *Image Processing, 2006 IEEE International Conference on*, Oct. 2006, pp. 2005–2008.