ROBUST GENERATION OF HIGH DYNAMIC RANGE IMAGES

Zhengguo Li, Zijian Zhu, Shoulie Xie, Shiqian Wu and Susanto Rahardja

Signal Processing Department, Institute for Infocomm Research, Singapore

ABSTRACT

A robust scheme is proposed to generate an anti-ghosting high dynamic range (HDR) image from a set of low dynamic range (LDR) images with different exposure times. Three major contributions of this paper are 1) a bi-directional prediction method; 2) an adaptive threshold for the classification of pixels; 3) Bayes estimator based methods for the on-line updating of predicted values and the synthesis of pixels to fill in the regions of moving objects to preserve their dynamic ranges. The proposed scheme is suitable for both static and dynamic scenes.

Index Terms— High dynamic range, collaborative image processing, ghosting artifacts, similarity

1. INTRODUCTION

A high dynamic range (HDR) image can be produced by sequentially capturing multiple low dynamic range (LDR) images using different exposures [1]. An image with large exposure time is saturated in the bright scene regions but captures the dark areas well. In contrast, an image with small exposure time is under-exposed in the dark scene regions but captures the bright areas well. Clearly, one image is not able to represent the whole scene but a set of images is able to do so. A HDR image can be synthesized by properly integrating all these LDR images. This method could be called collaborative image processing and the essence is " a single is impossible but a group possible". This method works well for static scenes. However, ghosting artifacts can be produced when there are moving objects in the scene [2, 3, 4, 5, 6].

The ghost removal methods using camera response function (CRF) approximation [6] and variance measurement [3] are highly sensitive to the estimation error of CRF. An entropy based method was proposed to be independent of the CRF by using the feature of local entropy [4]. However, as realized by the authors, this is unachievable when two image regions share the same structure but with different intensities. Another serious drawback of the methods in [3, 4, 6] is the loss of dynamic range for moving objects in the synthesized HDR image when their dynamic ranges are inherently high. In addition, Khan et al [2] generates HDR image by iteratively weighting the probability of the pixels belonging to the background or the moving objects with the cost of high complexity.

In this paper, we propose a robust scheme to produce anti-ghosting HDR images by introducing three novel components to improve the quality of the final HDR image. They are 1) a bi-directional prediction method; 2) an adaptive threshold for the classification of pixels; and 3) Bayes estimator based methods for on-line updating of predicted values and synthesis of pixels to fill in invalid regions for recovering the full dynamic range of the moving object. The process of the proposed scheme is illustrated in Fig. 1. A bi-directional prediction with an adaptive threshold is proposed to classify all pixels into valid and invalid for each input image. The CRFs are estimated by using all valid pixels. They are adopted to design a refined classification scheme such that ghosting artifacts are well removed. The initial values of intensity mapping functions (IMFs) are computed by using the CRFs. To reduce the sensitivity with respect to the estimated error of the CRFs, the IMFs are on-line updated by using a Bayes estimator based method. Both spatial correlation inside the image and co-location correlation between this image as well as its reference image are used to compute invalid pixels. As such, the dynamic ranges of moving objects are well preserved. After all images are processed, all updated reference images are used to generate the final HDR image. Experimental results show that our method is suitable for both static scenes and scenes with moving object.

2. THE ESTIMATION OF CAMERA RESPONSE FUNCTIONS WITH MOVING OBJECTS

Let $E_k(i,j)$ denote the image irradiance at position (i,j) when the kth LDR image is captured, i.e. (i,j) is a spatial position, and k indexes over exposure time Δt_k. $Z_k(i,j)$ is the corresponding intensity measurement that is recorded by the kth LDR image. The relationship among $Z_k(i,j)$, $E_k(i,j)$ and Δt_k is given as [7]

$$Z_k(i,j) = f(E_k(i,j)\Delta t_k + n_{s,k}(i,j)) + n_{f,k}(i,j), \quad (1)$$

where $f(\cdot)$ is a CRF, $n_{s,k}(i,j)$ is sensor noise and $n_{f,k}(i,j)$ is noise introduced by the electronics.

With a set of pictures that are identical except for their different exposure times, the values of $E_k(i,j)$ can be estimated accurately by using the assumption "their values are almost

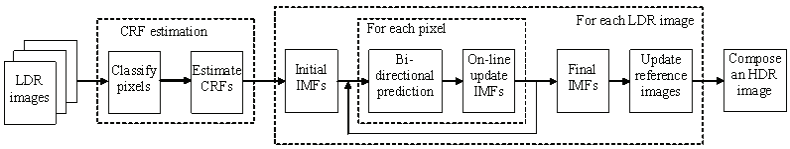

Fig. 1. The process of robust generation of HDR image.

the same for all co-located pixels ". However, this is not true when there are moving objects in the scene. It is necessary to properly classify all co-located pixels into valid and invalid as in [2, 3, 4, 5, 6] and only valid pixels are involved in the estimate of CRFs.

2.1. The Rough Classification of Pixels

Since the estimation of CRF is not very sensitive to the classification of pixels, a rough scheme is adopted to classify all pixels into valid and invalid. A basis image is selected with relative high reliability, in which all pixels are marked as valid. Let $\Omega(Z_{k_0}(i,j))$ denote the set of valid pixels, and it is initialized as $\{Z_{k_0}(i,j)\}$. Pixel $Z_k(i,j)$ is added into the set $\Omega(Z_{k_0}(i,j))$ if

$$\Gamma(Z_k, \hat{Z}_k) < \xi_k. \tag{2}$$

The details on the similarity measurement $\Gamma(Z_k, \hat{Z}_k)$ and the threshold ξ_k are given as follows:

1) The similarity measurement $\Gamma(Z_k, \hat{Z}_k)$ is defined as

$$\begin{cases} |\hat{Z}_k - \Lambda_{k,\pi(k)}(Z_k)|; & \text{if } W(Z_k) > W(\hat{Z}_k) \\ |Z_k - \Lambda_{\pi(k),k}(\hat{Z}_k)|; & \text{otherwise} \end{cases}, \tag{3}$$

where $\pi(k)$ denotes the reference image of the kth image, Z_k and $\hat{Z}_k$ represent the kth LDR image and its co-located reference image pixel value, respectively, and $W(z)$ is a triangle weighting function [1].

The intensity mapping function (IMF) from image k to its reference image $\pi(k)$ is calculated by [8]

$$\Lambda_{k,\pi(k)}(z) = \arg \min_{l \in [\Lambda_{k,\pi(k)}(z-1), 256]} \{|H_k(z) - H_{\pi(k)}(l)|\} \tag{4}$$

$$H_k(z) = \sum_{l=1}^{z} |\Theta_k(l)|, \Theta_k(z) = \{(i,j)|Z_k(i,j) = z\}, \tag{5}$$

where the value of $\Lambda_{k,\pi(k)}(z)$ is 1, and $|\Theta_k(z)|$ is the cardinality of the set $\Theta_k(z)$. The IMF $\Lambda_{k,\pi(k)}(z)(\cdot)$ is defined as a forward IMF, and the IMF $\Lambda_{\pi(k),k}(z)(\cdot)$ as a backward IMF. Since the two IMFs are computed in the same way, we illustrate only the forward IMF in this paper. It is shown from Equation (3) that the reference pixel is calculated using only the more reliable pixel value from either the original image or the reference image. We thus call the proposed prediction method as a bi-directional method.

2) The threshold ξ_k is adaptive to both the values of Z_k and $\hat{Z}_k$ as

$$\xi_k = \begin{cases} a\hat{Z}_k + (\epsilon(Z_k) + \epsilon(\hat{Z}_k))\hat{Z}_k; & \text{if } W(Z_k) > W(\hat{Z}_k) \\ aZ_k + (\epsilon(Z_k) + \epsilon(\hat{Z}_k))Z_k; & \text{otherwise} \end{cases}, \tag{6}$$

where a is a constant with typical value 0.0935, and the scale factor $\epsilon(z)$ is defined as

$$\epsilon(z) = \begin{cases} \frac{1}{64}(1 - \frac{2z}{255})^{10}; & \text{if } z > 128 \\ \frac{5}{4}(1 - \frac{2z}{255})^{10}; & \text{otherwise} \end{cases}. \tag{7}$$

The second item in Equation (6) is to address the possible effects of noises $n_{s,k}$ and $n_{f,k}$, especially the sensor noise $n_{s,k}$. Since the sensor noise is more serious when the images are under exposed, the scale factor is larger in this area.

2.2. The Estimation of Camera Response Functions

Similar to [1], the CRF $f(\cdot)$ is obtained by solving an optimization problem over the set $\Omega(Z_{k_0}(i,j))$. Due to the space limitation, the details are omitted.

3. THE GENERATION OF REFERENCE IMAGES

The final HDR image is very sensitive to the classification of pixels. Thus, a refined classification scheme is designed by using the following two criteria:

(c1) *Inter Pixel Criterion*: For the pixels that are neither saturated nor under-exposed in valid region, if any two points have equal pixel value, their co-located pixels in the reference image will have the same value, and vice versa.

(c2) *Co-located Pixel Criterion*: A pixel is valid only when the irradiance values in the image and its reference image are almost the same for at least two color channels. All valid pixels are firstly determined by Equation (4) for each pixel independently. To reduce the effects of noise, the valid pixels are smoothened by using an average filter in the neighborhood as

$$B_\rho(i,j) = \{(x,y)||x-i| \le \rho, |y-j| \le \rho\}, \tag{8}$$

where ρ is a constant and its typical value is 1.

3.1. The Processing Order of All Images

The basis image k_0 is the image with all valid pixels, it serves as the reference image of images (k_0-1) and (k_0+1). As two

successive images have the strongest correlation, the updated image (k_0-1) is the reference image of image (k_0-2). Thus, all images are processed in the order of k_0, (k_0-1), $\cdots$, 1, (k_0+1), $\cdots$, n_0. The value of $\pi(k)$ is updated as

$$\pi(k) = \begin{cases} k+1; & \text{if } k < k_0 \\ k-1; & \text{if } k > k_0 \end{cases} . \tag{9}$$

3.2. The Values of IMFs

The initial values, on-line updated values and final values of the IMFs $\Lambda_{k,\pi(k)}(\cdot)$ and $\Lambda_{\pi(k),k}(\cdot)$ are discussed as below.

3.2.1. The Initial Values of IMFs

If the reference image is the basis image, the initial values of the forward IMF is computed by using the estimated CRFs as

$$\Lambda^{(0)}_{k,\pi(k)}(z) = g^{-1}(g(z) + \log(\frac{\Delta t_{\pi(k)}}{\Delta t_k})), \tag{10}$$

where $g(z)$ is $\ln(f^{-1}(z))$, Δt_k and $\Delta t_{\pi(k)}$ are the exposure times. If the reference image is not the basis image, the inter pixel criterion (c1) can be applied, i.e. the initial values of IMFs are computed by using the final values of IMFs of the previous processed image.

3.2.2. The On-line Updating of IMFs

To reduce the sensitivity to the estimated error of CRFs, the IMFs are updated on-line by using the Bayes estimator with the risk function as the mean square error. After $Z_k(i,j)$ is marked as valid, exponential smoothing is adopted to update the forward IMF as

$$\Lambda^{c}_{k,\pi(k)}(Z_k) = \Lambda^{p}_{k,\pi(k)}(Z_k) + \frac{\hat{Z}_k - \Lambda^{p}_{k,\pi(k)}(Z_k) + 2^{\alpha-1}}{2^{\alpha}} \tag{11}$$

where 'c' and 'p' mean the current and previous values, is a constant and its typical value is 3. The initial value of $\Lambda^{p}_{k,\pi(k)}(z)$ is set at $\Lambda^{(0)}_{k,\pi(k)}(z)$. The values of IMFs are stored in a look-up table for fast processing. To avoid the possible error accumulation and propagation, each image is divided into blocks with size as 32x32. Within each block, the values of IMF are updated by Equation (10), and the start values are computed by

$$\Lambda^{(s)}_{k,\pi(k)}(Z_k) = \Lambda^{(s)}_{k,\pi(k)}(z) + \frac{\Lambda^{(f)}_{k,\pi(k)}(z) - \Lambda^{(s)}_{k,\pi(k)}(z) + 2^{\alpha-1}}{2^{\alpha}}, \tag{12}$$

where $\Lambda^{(f)}_{k,\pi(k)}(\cdot)$ is the final values of IMF for the previous block. The updating process can be applied to improve the robustness of the proposed scheme in the sense that 1) the sensitivity of the proposed scheme with respect to the accuracy of CRFs is reduced; and 2) the effect of noise is also reduced because the mean value is used.

3.2.3. The Final Values of IMFs

After the prediction of the kth LDR image, all pixels in the kth LDR image are classified into two groups, valid pixels and invalid pixels . The final values of IMFs are computed by using the Bayes estimator over all pixels in the valid set.

3.3. The Updating of Reference Images

In [8, 4, 6], all invalid regions are filled in by pixels from one single selected image. Even though there are different methods to select such an image, the dynamic ranges of moving objects will be reduced if the dynamic ranges of moving objects are high. Controversially, our method is able to synthesize pixels by using the inter pixel criterion (c1) to fill in the invalid region. To further reduce the noises, the invalid pixel value is updated by the final value of IMFs.

4. THE SYNTHESIS OF THE FINAL HDR IMAGE

With the prediction given in the above section, the final HDR image is composed by using all the updated reference images $\hat{Z}_k(i,j)$. The invalid regions of these images are filled in by using pixels from more than one original LDR images instead of one single image as in [8, 4, 6]. The dynamic ranges of moving objects could be thus preserved better in the synthesized HDR image.

5. EXPERIMENTAL RESULTS

Extensive experiments are carried out to verify the efficiency of the proposed method. We first compare the proposed bi-directional prediction method with the conventional uni-directional prediction method. We test an image sequence that is composed of seven images and both the head and body of man move in the scene. The experimental results are illustrated in Fig. 2 and it can be shown that the bi-directional method can be used to significantly improve the quality of the darkest part (i.e. the hair and body of man and the television) in the synthesized HDR image.

Fig. 2. Bi-(left) vs Uni-(right) directional prediction methods.

We now verify our proposed Bayes estimator based methods for the updating of predicted values and reference images. We study an image sequence that consists of nine images and the scene is static. It is shown in Fig. 3 that the proposed

Bayes estimator based methods can be used to smoothen both inside and outside walls. It can be demonstrated from Figs 2 and 3 that the proposed method is suitable for both dynamic and static scenes.

Fig. 3. Smoothing with Bayes estimator (left) and no Bayes estimator (right).

We also evaluate the efficiency of our adaptive threshold by comparing it with the case that the threshold is fixed as 32. We test an image sequence that is composed of eight images and there are several moving objects in the scene. It is illustrated in Fig. 4 that the proposed adaptive threshold can be applied to remove ghosting artifacts while a fixed threshold cannot do so. Finally, we compare the proposed scheme with

Fig. 4. Adaptive (left) and fixed (right) threshold.

commercial softwares FDRTools, Photomatrix and Otpfsgui [9], by testing an image sequence that is composed of 11 images. It is shown in Fig. 5 that there are ghosting artifacts by using these commercial software, especially by Qtpfsgui while the ghosting artifacts are removed via our method.

6. CONCLUSIONS

In this paper, we propose a robust scheme for the synthesis of anti-ghosting high dynamic range (HDR) images. Our three contributions are a bi-directional prediction method; an adaptive threshold; and Bayes estimator based methods for the updating of predicted values and the synthesis of invalid pixels. The proposed scheme is suitable for both static scenes and scenes with moving objects.

7. REFERENCES

[1] P. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photograph," In *Proceedings SIGGRAPH 1997*, pp.369-378, 1997.

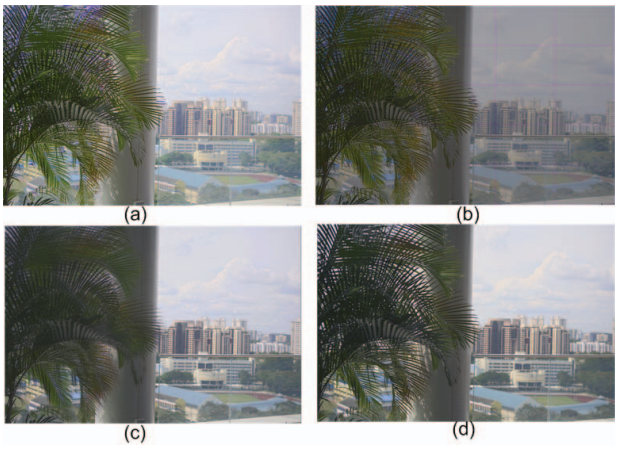

Fig. 5. Different schemes for ghost removal. (a) FDRTools, (b) Photomatix, (c) Qtpfsgui, and (d) Our method.

[2] E. A. Khan, A. O. Akyuz, and E. Reinhard, "Ghost removal in high dynamic range images," in *Proc. IEEE International Conference on Image Processing*, Atlanta, GA, pp.2005-22008, 2006.

[3] E. Reinhard et al., High dynamic range imaging: acquisition, display and image-based lighting, Morgan Kaufmann, 2005.

[4] K. Jacobs, C. Loscos, and G. Ward, "Automatic high dynamic range image generation for dynamic scenes," IEEE Computer Graphics and Applications, Vol. 128, No.2, pp.84-93, Feb. 2008.

[5] S. Kang, M. Uttendaele, S. Winder, and R. Szeliski, "High dynamic range video," In ACM Trans. On Graphics, pp. 319-325, 2003.

[6] T. Grosch, "Fast and robust high dynamic range image generation with camera and object movement," In *Vision, Modeling and Visualization*, L. Kobbelt, T. Kuhlen, T. Acch and R. Westermann Eds, pp.277-284, 2006.

[7] S. Mann, "Comparametric equations with practical applications in quantigraphic image processing," IEEE Transaction on Image Processing, Vol. 9, No. 8, pp.1389-1406. August 2000.

[8] M. Grossberg and S. Nayar, "Determining the camera response from images: what is knowable?" IEEE Transactions on pattern analysis and machine intelligence, pp. 1455-1467, Vol. 25, No. 11, Nov. 2003.

[9] FDRTOOLS, PHOTOMATIX, and QTPFSGUI, http://fdrtools.com/front_e.php, http://www.hdrsoft.com/, http://qtpfsgui.sourceforge.net/, 2009.