A ROBUST AND FAST ANTI-GHOSTING ALGORITHM FOR HIGH DYNAMIC RANGE IMAGING

Shiqian Wu, Shoulie Xie, Susanto Rahardja and Zhengguo Li

Institute for Infocomm Research, 1 Fusionopolis Way, #21-01 Connexis Singapore 138632
Email: {shiqian,slxie,rsusanto,ezgli}@i2r.a-star.edu.sg

ABSTRACT

This paper presents a robust and fast algorithm for automatically generating high dynamic range (HDR) images in presence of camera movement and moving objects. This scheme comprises five modules: 1) image alignment, 2) estimation of camera response function (CRF) in dynamic scenes, 3) moving object detection, 4) progressive image correction, and 5) construction of HDR images. The key advantage of the algorithm is the ability to generate HDR images without ghost artifact. The proposed algorithm is fast as it is a one-shot solution without iterative computation and post-processing or even manual operation. Experimental results demonstrate that the proposed method outperforms the existing commercial products.

Index Terms—HDR imaging, ghost artifact, camera response function, detection of moving objects, progressive image correction

1. INTRODUCTION

Many natural scenes have higher dynamic range than those recorded by a camera. Combining differently exposed low dynamic range (LDR) images of the same scene has been the most popular approach to generate an HDR image [1, 8]. However, these methods always result in two problems [2, 3, 4, 8]: 1) Ghost artifact appears due to dynamic scenes, such as moving people, vehicles and so on; 2) Blurry artifact occurs because of camera movements. While image registration has been employed to deal with the camera movement, ghost removal remains an open issue in HDR image generation.

Recently, statistical methods were employed to generate HDR images without ghost artifacts in [4] and [7]. These techniques do not rely on explicit object detection and motion estimation. But expensive computation is unacceptable for a large set of images due to its iterative strategy. It is also noted that the statistical methods cannot completely remove ghost artifacts unless the probability of a pixel belonging to the background is absolutely zero or one. Moreover, as indicated by these authors, their methods are exposed to getting halo effects in the generated HDR image, especially in image sequences including wavy motion like burning candle flames.

In [2], Grosch presented an alternative approach for movement detection based on the error map, where the error map was resulted from the predicted image via CRF in a specific exposure and the original image under the same exposure. This approach works well if moving objects are not in saturated or underexposed regions. Due to inaccuracy of the CRF, the error map is non-uniform, i.e., the predicted error is big for bright regions while small for dark region within two LDR images, and the predicted error increases along with the increase of exposure gaps. Accordingly, this method cannot thoroughly identify the moving objects so that the regions of the moving objects have to be corrected with a special clone brush. Moreover, the threshold is not robust as it depends on the amount of noise in the images. In [8], movement is detected based on the variation of variance. But this technique is sensitive to the estimated error of CRF, and this method is only effective to remove ghost artifacts which can be segmented. The concept of entropy was proposed to identify moving objects in [3]. Although entropy is independent of image intensities, this method can only cope with the scenes in which ghosts occur in regions with low dynamic range.

In this paper, a fast and robust algorithm is developed for automatic generation of HDR images in consideration of camera moment and dynamic scenes. Five criteria are first proposed to estimate the CRF in dynamic scenes. Three criteria are proposed to identify moving objects. An idea of progressively updating strategy is employed to identify the difference between moving objects and background in LDR images. A novel algorithm combining the CRF information and the original images is presented to correct the LDR images. To mitigate the errors from CRF and image alignment, the corrected images are further processed using image inpainting. The proposed method is non-iterative. Also, the approach is robust to different images (with different contents, dynamic scenes/still scenes) without tuning the predefined parameters, i.e., the thresholds are robust to varied scenes. The experimental results demonstrate that the proposed method is significantly superior to the existing commercial products [9–11].
2. ARCHITECTURE OF THE PROPOSED SCHEME

Most photos are captured without a tripod, and natural scenes usually contain moving objects. Such situations lead to blurry and ghosting problem for HDR generation using existing techniques. In our work, a generalized scheme for HDR generation has been developed shown in Fig. 1. The inputs are a set of LDR images with different exposure times. Image alignment is first conducted to eliminate camera movement. Then, moving objects are initially detected and appropriate samples are selected for CRF estimation. Using the CRF curve and the LDR images, a refined processing for detecting moving objects is performed and the original LDR images are progressively corrected. Finally, these corrected images are employed to generate the HDR image. In the following, we focus on CRF estimation in dynamic scenes, the movement detection and progressive image correction as shown in Fig. 1.

3. CRF ESTIMATION IN DYNAMIC SCENES

An important step for CRF estimation is to select appropriate samples. In [1], the samples are randomly selected for still scenes with fixed camera. This method leads to a problem for a dynamic scene as the co-located samples may locate on different objects. Here, five criteria are proposed to select the samples:

1) The exposure changes only affect the image intensities without color variation. Hence, the color error criterion shown in Section 4.1.3 is employed for approximately detecting moving objects;
2) Long exposure corresponds to big pixel values;
3) The samples are selected from smooth areas to eliminate the alignment error;
4) The samples are totally covered by the whole range;
5) The number of samples should be larger than $(Z_{\text{max}} - Z_{\text{min}})/(P - 1)$, where $Z_{\text{max}}$, $Z_{\text{min}}$ are the maximal and minimal pixel values respectively and $P$ is the number of the LDR images captured.

4. DETECTION OF MOVING OBJECTS AND IMAGE CORRECTION

4.1. Detection of Moving Objects

Assuming $P$ color images are captured, they are then converted to grayscale images $I_i \in \mathbb{R}^{M \times N}$ ($i = 1, 2 \cdots P$) and arranged from long exposure to short exposure. The following three criteria are proposed to detect the moving objects as accurate as possible.

4.1.1 Criterion 1: Monotonous criterion

Since CRF is monotonically increasing, we have the following criterion for grayscale images:

$$Z_{i,j}^{k-1} \geq Z_{i,j}^k \geq Z_{i,j}^{k+1}$$

where $Z_{i,j}^k$ is the pixel value in a grayscale image with exposure $k$ at position $(i, j) \in \mathbb{R}^{M \times N}$. Therefore, a pixel at the identical position which does not meet the above relation is classified as moving objects.

4.1.2 Criterion 2: Pixel error criterion

Suppose the CRF has been estimated, pixel $Z_{i,j}^l$ in grayscale image with exposure $\Delta t^l$ corresponding to $Z_{i,j}^k$ in the reference image $k$ can be predicted by

$$\hat{Z}_{i,j}^l = f(\frac{\Delta t^l}{\Delta t} f^{-1}(Z_{i,j}^k))$$

If

$$| \hat{Z}_{i,j}^l - Z_{i,j}^l | \leq \varepsilon$$

it implies homogeneous region, i.e., no moving object occurs at position $(i, j)$ in image $l$.

It should be noted that the above two criteria are not sufficient for moving object detection. For criterion 1, if the pixel values on the moving objects are within the difference of background in consecutive images, the moving objects cannot be tracked. For criterion 2, the discrimination is non-uniform. This is due to the nonlinearity of the CRF, which leads the large exposure difference yielding large predicted error. Also, for a fixed threshold, if the moving objects contain dark region, the dark region cannot be detected. Accordingly, the color criterion is proposed as follows.

Fig.1 Scheme of the proposed HDR generation

LDR images

Image alignment

Movement detection (coarse)

CRF estimation in dynamic scenes

Movement detection (refined)

Progressive image correction

HDR image generation

HDR image without ghost

398
4.1.3 Criterion 3: Color error criterion

The color photos captured by a camera normally consist of R, G, B three components. It is assumed that variations of exposures only change the intensity without changing the color, i.e.,

\[ C_{ij}^t(R, G, B) = K_{ij}C_{ij}^0(R, G, B) \]  

where \( C_{ij}^t(R, G, B) \) is the RGB color vector at position \((i, j)\), and \( K_{ij} \) is a positive scalar. This assumption is reasonable because a commonly used empirical law for camera’s CRF is Gamma function [8], and in this case \( K_{ij} = (\Delta t^l / \Delta t^l)^\gamma \) with \( \gamma \) being the camera’s contrast parameter given by manufacturer.

Thus, we deduce the following criterion from equation (4):

\[ \frac{< C_{ij}^t(R, G, B), C_{ij}^u(R, G, B) >}{|| C_{ij}^t(R, G, B) || \cdot || C_{ij}^u(R, G, B) ||} \geq \cos \alpha \]

then no moving object occurs at position \((i, j)\) in image \(l\), where \( \alpha \) is positive threshold and \(< \cdot >\) is the inner product.

Suppose \( S^1, S^2, S^3 \) are the pixel sets which are not satisfied with the aforementioned three criteria respectively, then, the set \( S \) covering the moving objects is as follows:

\[ S = S^1 \cup S^2 \cup S^3 \]

It is noted that a moving object is a continuous region in all cases. The above pixel-based processing may result in misclassification and yield noise or hole in the potential moving objects. The morphological close and open operators are accordingly performed to remove these effects on the object, while distorting the object as little as possible.

4.2 Progressive Image Correction

It is indicated in Section 3.1 that the CRF is nonlinear function, and the curve in the parts of dark and bright are not accurate. The predicted accuracy in equation (2) is dependent on the exposure difference due to the nonlinearity of the CRF. Large exposure difference yields large predicted error. Here, we propose a progressive image correction method to overcome these problems.

First, select the image \(k\) as the basis for the synthesis, for which all pixels are marked as valid, i.e. the elements of the mask \( W^k \) are set as 1s. The basis image only serves as of its two adjacent images \((k-1)\) and \((k+1)\). We use image \((k+1)\) as an example in the following analysis. By using the three criteria and morphological operations, we obtain

\[ W^{k+1}_m(i, j) = \begin{cases} 0 & (i, j) \in S \\ 1 & \text{otherwise} \end{cases} \]

Only the valid pixels are involved in the synthesis of HDR image. To improve the visual quality of the synthesized HDR image, it is very important to reproduce the pixels for those invalid pixels. Here, we use the CRF to do so and a new image \( T_{k+1} \) is produced as follows:

\[ T_{k+1} = \begin{bmatrix} Z^k_{i, j} & W^{k+1}_m(i, j) = 0 \\ Z^k_{i, j} & W^{k+1}_m(i, j) = 1 \end{bmatrix} \]

Then, the new image \( T_{k+1} \) is selected as reference to predict its consecutive image until all images are reconstructed. The idea to progressively update the reference image greatly mitigates the effects of the CRF inaccuracy and image noise.

4.3 Inpainting of Corrected Images

In the aforementioned processing, the regions with moving objects are compensated by predicted results using CRF curve. Due to the error of CRF estimation, boundaries occur between the predicted results and the original images, even the two parts are homogenous region. Therefore, it is necessary to smoothen such regions. In this scheme, the technique of image inpainting in [6] is employed in the following convolution:

\[ Z^k_{i, j} = Z^k_{i, j} \odot F \quad (i,j) \in \Omega \]

where \( F \) is the convolution kernel and \( \Omega \) represents the boundaries of the set \( S \) covering the moving objects. Then, the final corrected images after inpainting are as follows:

\[ Z^k_{i, j} = \begin{bmatrix} Z^k_{i, j} & (i,j) \in \Omega \\ Z^k_{i, j} & \text{otherwise} \end{bmatrix} \]

and the HDR image is composed using the new corrected images by

\[ \ln E_{i, j} = \frac{\sum_{k=2}^{p} W(Z^k_{i, j})(f^{-1}(Z^k_{i, j}) - \ln \Delta t^k)}{\sum_{k=2}^{p} W(Z^k_{i, j})} \]

The regions of moving objects are filled by the approximated pixels derived from the reference images so that the information from each pixel in all exposures is used. The inpainting of corrected images makes the resulting HDR image smoother in the boundaries of moving objects. Such compensation and inpainting greatly alleviates the CRF inaccuracy and the thresholds are robust to scene variations and the final HDR image is visually pleasing.

5. EXPERIMENTAL RESULTS

We conducted experiments using a variety of real images to verify the effectiveness of the proposed method. The images are captured by Canon EOS-1Ds Mark II. For the simplicity of processing, all image sizes are reduced to 1248×832. The middle image is selected as the first reference and the parameters predefined are as follows: the thresholds \( \varepsilon = 40 \), and \( \alpha = 20^\circ \). Our results are compared with those current commercial products [9-11]: FDRTools, Photomatix...
and Qtpfsgui. For comparison with the results by the commercial tools, all the generated HDR images in the following experiments are tone-mapped using the technique described in [5].

A sequence of images containing different types of moving objects, such as moving leaves, walking people, vehicles and so on, is exemplified in Fig. 2. Using the methods in [3, 4, 7, 8], it is very difficult to construct the HDR image without any ghost artifacts because the scene is not dominated by background. The composed HDR images by Qtpfsgui, FDRTools, Photomatix and our method are depicted in Figs. 3-6, respectively. It is observed that ghost artifacts are obvious in whole scene using Qtpfsgui and FDRTools, while the leaves appear ghosts using Photomatix, and the top of the truck is distorted. Our method produces the best HDR image which is ghost-free and visually pleasing.

It is highlighted that the proposed method is fully automatic, and the above results are implemented in the same parameters, i.e., the thresholds are robust to varying scenes. It is easily seen that our method significantly outperforms Qtpfsgui, FDRTools and Photomatix. The method is able to remove ghost artifacts while keeping the visual quality of the HDR image. In addition to dynamic scene, we also apply the proposed method to still scene for HDR construction. Due to the space limitation, we cannot show the results here. The experimental results demonstrate that our method also achieves best performance in both image details and color consistence in comparison with the three commercial products. The results by our method are comparable to that of HDRShop [12], which is specialized in HDR generation for still scenes. Therefore, our method can produce good HDR images in both still scenes and dynamic scenes.

6. CONCLUSION

This paper presents an automatic scheme for generating HDR images in consideration of camera movement and dynamic scenes. The system is robust for removing ghost artifacts due to the unique criteria of CRF estimation in dynamic scenes, detecting moving objects, as well as progressive image correction. The inpainting technique and morphological operations are further performed in corrected images to alleviate the inaccuracy resulted from image alignment and CRF estimation that the composed HDR image is smooth and visually pleasing. The system works fast as it is a non-iterative solution. The experimental results demonstrate that the proposed system is significantly superior to the existing commercial products.

7. REFERENCES

[12] HDRShop software: http://gl.ict.usc.edu/HDRShop/