Ghost-Free High Dynamic Range Imaging via Moving Objects Detection and Extension

Benkang Zhang* and Qin Liu* and Takeshi IKENAGA†
* Nanjing University, Nanjing, China
E-mail: {zbk11, qinliu}@software.nju.edu.cn
† Graduate School of Information, Production and Systems, Waseda University, Fukuoka, Japan
E-mail: ikenaga@waseda.jp

***Abstract*—High dynamic range imaging(HDRI) techniques are proposed to synthesize high dynamic range (HDR) images from multi-exposure images. However, ghost artifacts may appear if images are synthesized directly when there are moving objects in the scene. This paper presents an algorithm to composite a HDR image from multi-exposure images without ghost artifacts. To remove the ghosts in the final image, the proposed algorithm firstly produces a 0-1 map based on a Markov Random Field(MRF) framework. The moving areas are detected and marked with 1. Then, moving areas are extended and used in the final exposure fusion step. The marked pixels are assigned zero weights to prevent ghost artifacts.**

***Index Terms*—High Dynamic Range Imaging, Ghost Free, Exposure Fusion, Moving Objects Detection, Image Segmentation**

I. Introduction

The dynamic range of an actual scene is much wider than the ranges that the most sensors or display devices can support. In order to get closer to the true image display, many high dynamic range imaging(HDRI) techniques have been proposed. Multi-exposure images are synthesized to overcome the limitations of sensors and display devices. There are two main approaches to do the synthesis. One[1] produces a HDR image firstly and then gets a LDR image with tone mapping. The other[2] synthesizes the LDR image directly from multi-exposure images. The proposed algorithm belongs to the second one.

One of the major problems of synthesizing multi-exposure images is the ghost artifacts which are caused by the movement of objects in the scene, which is also the main problem that the current multi-exposure images synthesis algorithms need to focus on.

Prior to this article, a lot of ghost-free HDR image synthesis algorithms have been proposed. Srikantha[3] introduced and compared many recent HDR image synthesis algorithms without ghost artifacts. One main idea is that the moving objects are detected, then the movement areas are excluded or assigned with small weights when synthesizing the images. Li[4] used IMF(Intensity Mapping Function) to detect moving areas. Gallo[5] used RANSAC[6]. Huang[7] used block matching to find moving areas. Recently there is a new moving areas detection idea presented by Zhou[8], this algorithm assumes that the background image is a low rank matrix and computes the optimal solution to detect moving areas. Lee[9] introduced this low-rank detection method into HDR synthesis. The algorithms mentioned above try to detect the moving areas and adjust the weights of pixels when synthesizing. However, it is difficult to detect moving areas accurately especially for multi-exposure images. Ghosts may still exist by using this kind of methods. The proposed algorithm also detects moving areas by a method based on Markov Random Field(MRF). However, the moving areas mask is extended before being used in final fusion step.

Section 2 describes the details of the proposed algorithm and results of evaluation experiments are shown in Section 3. Section 4 is the final conclusion.

II. Proposal

Fig. 1 shows the detailed flow of the proposed algorithm. Firstly, the CRF(Camera Response Function) is estimated using the method of Debevec[1], then we can adjust the reference image to different exposure values, the results are used to detect the moving areas. The areas are extended after some steps and used in the final exposure fusion.

The main idea of the proposed algorithm is to use a step to enhance the detection result. It is difficult to detect moving areas accurately especially for multi-exposure images. The underexposure and overexposure parts of images bring many problems. Furthermore, ghosts may still exist even if the moving areas were detected and marked. The left image of Fig. 2 shows a diagram of sequence with two images. The second image is selected as the reference image, and the black circle in the first image is marked as moving areas. When synthesizing the two images, the pixels in circle are excluded and assigned with zero weights, which causes different input values between the circle and its around regions. So a circular ghost may exist in the final image. In order to solve this problem the proposed algorithm extends the moving areas to make the areas that have small internal differences been synthesized with the same inputs. The right image of Fig. 2 shows the extended image.

A. Moving Areas Detection

$I(i,p)$ denotes the p pixel of the ith image. $T(i)$ denotes the exposure time of the ith image, and we assume that $T(i)$ is an increasing sequence.

Firstly, the CRF(Camera Response Function) is estimated using the method of Debevec[1]. Eq. 1 shows the relationships

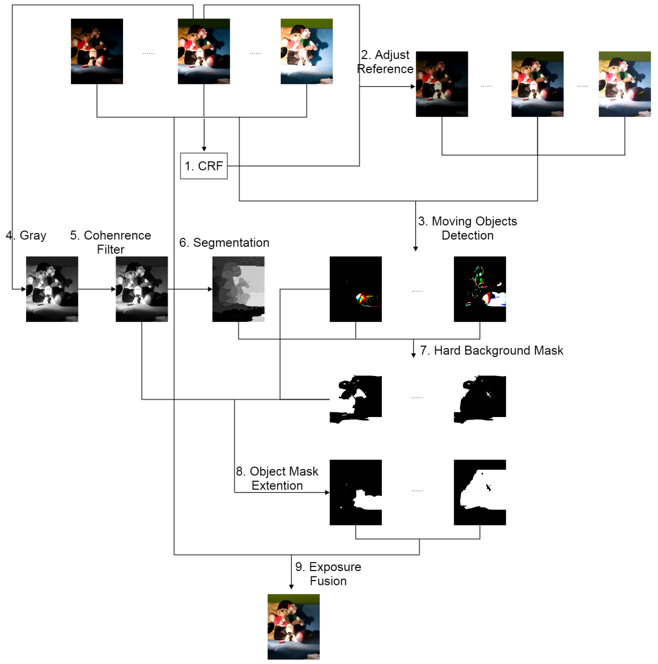
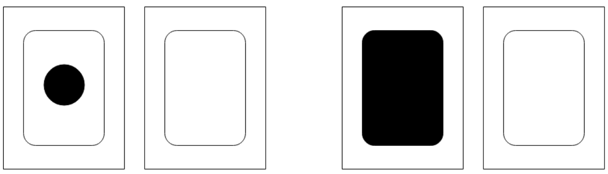


Fig. 1. Flow of proposed algorithm



(a) Conventional Pixel based Moving Area Detection

(b) Proposed Moving Area Extension Method

Fig. 2. Proposed Idea to remove the circle ghost

between scene radiance($E(p)$), exposure time($T(i)$) and image pixel value($I(i,p)$). g is the mapping function which can be computed using Debevec[1]'s method. $E(p)$ is the radiance value in the scene. We can adjust the reference image to different exposure times by replacing the $E(p)$ with the radiance of the reference image, the result is denoted by $R(i,p)$, i is the index of images and p is the index of pixels.

$$g(I(i,p)) = lnE(p) + lnT(i) \tag{1}$$

The result of Debevec[1] deviates the correct values when the pixels are under or over exposed. In order to avoid the incorrectness of the CRF, the reference image is divided into three parts: low, well and high corresponding to the under-exposed, well-exposed and over-exposed parts of the image and the three parts are processed differently. Eq. 2 shows the detailed definition, ref denotes the index of reference image.

$$A(i,p) = \begin{cases} l, & I(ref,p) \leq 1 \\ w, & 1 < I(ref,p) < 254 \\ h, & I(ref,p) \geq 254 \end{cases} \tag{2}$$

Then distance values are computed for three parts differently. Eq. 3.4.5.6 shows the details. The values show the distances between the images and the adjusted reference images, and in a way show the probabilities of being moving areas.

$$D(i,p) = \begin{cases} D_l(i,p), & A(i,p) = l \\ D_w(i,p), & A(i,p) = w \\ D_h(i,p), & A(i,p) = h \end{cases} \tag{3}$$

$$D_l(i,p) = \begin{cases} |I(i,p)|, & i < ref \\ max(I(i,p) - R(i,p), 0), & i > ref \end{cases} \tag{4}$$

$$D_w(i,p) = |R(i,p) - I(i,p)| \tag{5}$$

$$D_h(i,p) = \begin{cases} max(R(i,p) - I(i,p), 0), & i < ref \\ |255 - I(i,p)|, & i > ref \end{cases} \tag{6}$$

In the proposed algorithm, the moving areas are assumed to be continuous. A MRF based energy function is used to transform the labeling problem into a energy minimization problem. Eq. 7 shows the definition of the function. N denotes the neighbor pixels set. $L(p), L(q)$ denotes the label of pixel p and q. The label value equals 0 or 1, 0 means background and 1 means moving areas.

$$E = \sum_p E_D + \gamma \sum_{(p,q)\in N} E_S(L(p), L(q)) \tag{7}$$

Eq.8 shows the smooth cost of energy function

$$E_S(a,b) = \begin{cases} 0, & a = b \\ 1, & a \neq b \end{cases} \tag{8}$$

The data cost of the energy function is divided into three parts. Eq. 9.10.11.12 shows the details of the data cost. The th is a threshold parameter. $\sigma_l, \sigma_w, \sigma_h$ is the standard deviation of distance values in low, well, high of ith image. β is a parameter that can be tunned.

$$E_D(i,p) = \begin{cases} E_{D,L}(i,p), & A(i,p) = l \\ E_{D,W}(i,p), & A(i,p) = w \\ E_{D,H}(i,p), & A(i,p) = h \end{cases} \tag{9}$$

$$E_{D,L}(i,p) = \begin{cases} L(i,p)G(D(i,p),th)(2\gamma) \\ +(1-L(i,p))(1-G(D(i,p),th)) \\ (2\gamma + |D(i,p) - th|), \\ & i < ref \\ L(i,p)G(D(i,p),\beta\sigma_l)(2\gamma) \\ +(1-L(i,p))(1-G(D(i,p),\beta\sigma_l)) \\ (|D(i,p) - \beta\sigma_l|), \\ & i > ref \end{cases} \tag{10}$$

$$E_{D,W} = L(i,p)G(D(i,p),\beta\sigma_w)(2\gamma) + (1-L(i,p))(1-G(D(i,p),\beta\sigma_w))(|D(i,p)-\beta\sigma_w|) \quad (11)$$

$$E_{D,H}(i,p) = \begin{cases} L(i,p)G(D(i,p),\beta\sigma_h)(2\gamma) \\ +(1-L(i,p))(1-G(D(i,p),\beta\sigma_h)) \\ (|D(i,p)-\beta\sigma_h|), & i < ref \\ L(i,p)G(D(i,p),th)(2\gamma) \\ +(1-L(i,p))(1-G(D(i,p),th)) \\ (2\gamma+|D(i,p)-th|), & i > ref \end{cases} \quad (12)$$

The function G equals 1 if the value is less equal than some threshold value and 0 if greater. Eq. 13 shows the definition of G.

$$G(x,t) = \begin{cases} 1 & x \le t \\ 0 & x > t \end{cases} \quad (13)$$

The energy minimization problem can be solved by graph cuts. Boykov[10] shows the details of the algorithm. The energy function is applied to each image and pixels are labeled with 0 or 1 in the result. 0 is the background and 1 is the moving areas. The mask images are denoted by M.

B. Moving Areas Extension

The mask of moving areas is desired to be extended constrained by the edges. The extension algorithm is based on the method described in Boykov[11]. Two kinds of hard mask images are required. Hard mask means that the marked areas must be labeled with object or background.

Firstly, the images are transformed into gray images and filtered by the coherence filter which is described in Kroon[12]. This filter smooths the images and enhances the edges. The result images are denoted by R_f. Then R_f is segmented by the image segmentation algorithm described in Felzenszwalb[13]. The results are denoted by R_s which contains the segmentation label of each pixel.

The definition of hard masks of object and background is described in Eq. 14. H_O is the hard mask of objects and H_B is the hard mask of backgrounds. $erode$ is an erode function in order to reduce the noises.

The extension problem is also a labeling problem, which can be solved by constructing the energy function and finding the minimization of it. Eq. 15 shows the details of the energy function which is similar to the function of object detection. The main different part is $W'(i,p,q)$ which is a weight function that introduces the influence of edges into energy function. N is the set of all pair of neighboring pixels. The K is a big enough value which ensures that the areas marked

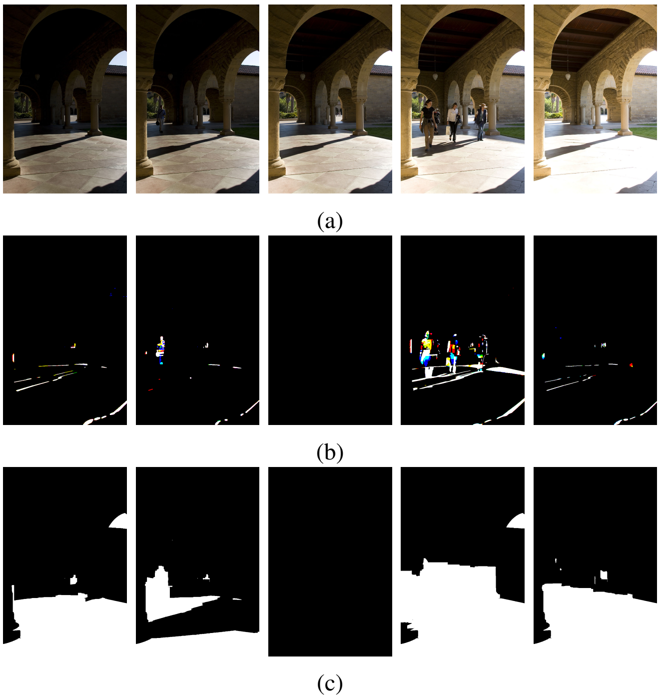


Fig. 3. Results of moving detection and extension. (a) Multi-exposure images (b) The moving area detection result (c) The extended mask

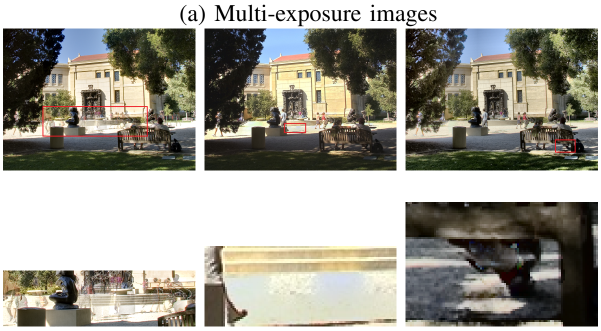
with object hard mask or background hard mask are labeled with 1(object) or 0(background). $std(R_f(i))$ is the standard deviation of ith R_f. The weight function penalizes a lot for similar intensities pixels when $|R_f(i,p) - R_f(i,q)| < \sigma$. And the σ which is controlled by the parameter λ can be considered as the threshold of the weight function.

$$\begin{aligned} &H_O = M \\ &B = \{l | \exists p, M(p) = 1 and R_s(p) = l\} \\ &H'_B(p) = \begin{cases} 1, & R_s(p) \in B^{\complement} \\ 0, & R_s(p) \in B \end{cases} \\ &H_B = erode(H'_B) \end{aligned} \quad (14)$$

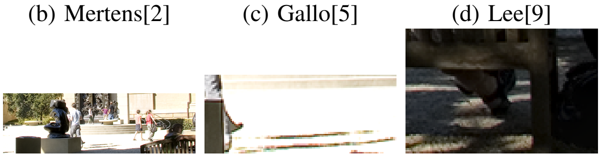
A sequence of 0-1 mask images can be acquired when the minimization of the energy function is estimated. The results are the mask images used in the final exposure fusion and denoted by M'.

$$\begin{aligned} &E' = \sum_p E'_D + \gamma' \sum_{(p,q)\in N} E'_S(L(p),L(q))W'(i,p,q) \\ &E'_D = L(p)H_B(p)K + (1-L(p))H_O(p)K \\ &E'_S = \begin{cases} 0, & a = b \\ 1, & a \ne b \end{cases} \\ &W'(i,p,q) = e^{-\dfrac{|R_f(i,p)-R_f(i,q)|^2}{2\sigma^2}}, \sigma = \lambda std(R_f(i)) \end{aligned} \quad (15)$$

(a) Multi-exposure images



(b) Mertens[2] (c) Gallo[5] (d) Lee[9]



(e) Proposed Method

Fig. 4. Comparison of results for a set of multi-exposure images.

C. Exposure Fusion

The proposed algorithm uses the exposure fusion to synthesize multi-exposure images. The details are described in Mertens[2].

The weight function is modified for ghost-free synthesis. Eq. 16 shows the modified edition of weight function. The pixels marked with label 1 are assigned to zero weights.

$$W(i,p) = (C(i,p))^{W_C} \times (S(i,p))^{W_S} \times (E(i,p))^{W_E} \times (1 - M'(i,p)) \tag{16}$$

III. Results

The proposed algorithm is tested with the parameters: $\gamma = 40, th = 5, \beta = 1.9, \gamma' = 10, \lambda = 50$. The test images can be found in http://alumni.soe.ucsc.edu/∼orazio/deghost.html.

Fig. 3 shows the results of the moving areas detection and extension. The mid image is selected as the reference image, so the mid mask image is all 0.

Fig. 4 shows the comparison between direct exposure fusion, Gallo[5], Lee[9] and the proposed algorithm. The problems in the results of the first three are marked with red rectangles. The results show that the proposed algorithm can remove the ghost artifacts efficiently. The methods of Gallo[5] and Lee[9] can remove most artifacts but some small ghosts may still exist.

IV. Conclusions

This paper has proposed a multi-exposure images synthesis algorithm to remove ghost artifacts. The moving areas are detected by a method based on MRF framework and extended by an algorithm based on image segmentation. The weights of the pixels are adjusted according to the extended mask images for the exposure fusion.

There are still some drawbacks of the proposed algorithm. Too many parameters in this algorithm need to be set, which influences the robustness of the algorithm. The marked areas are assigned to zero weights, which loses some information of the images and lowers the quality of the synthesized image. Some areas are extended too much and some overexposed or underexposed pixels are remained in the final image.

Acknowledgments

The Project was supported by the Natural Science Foundation of Jiangsu,China(Grant No. BK20130588).

References

[1] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *ACM SIGGRAPH 2008 classes*. ACM, 2008, p. 31.

[2] T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion," in *Computer Graphics and Applications, 2007. PG'07. 15th Pacific Conference on*. IEEE, 2007, pp. 382–390.

[3] A. Srikantha and D. Sidibé, "Ghost detection and removal for high dynamic range images: Recent advances," *Signal Processing: Image Communication*, vol. 27, no. 6, pp. 650–662, 2012.

[4] Z. Li, S. Rahardja, Z. Zhu, S. Xie, and S. Wu, "Movement detection for the synthesis of high dynamic range images," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*. IEEE, 2010, pp. 3133–3136.

[5] O. Gallo, N. Gelfand, W.-C. Chen, M. Tico, and K. Pulli, "Artifact-free high dynamic range imaging," in *Computational Photography (ICCP), 2009 IEEE International Conference on*. IEEE, 2009, pp. 1–7.

[6] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[7] S.-Y. Huang, Q. Liu, H. Wang, and T. Ikenaga, "Motion area based exposure fusion algorithm for ghost removal in high dynamic range video generation," in *Asia-Pacific Signal and Information Processing Association, 2014 Annual Summit and Conference (APSIPA)*. IEEE, 2014, pp. 1–4.

[8] X. Zhou, C. Yang, and W. Yu, "Moving object detection by detecting contiguous outliers in the low-rank representation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 35, no. 3, pp. 597–610, 2013.

[9] C. Lee, Y. Li, and V. Monga, "Ghost-free high dynamic range imaging via rank minimization," 2014.

[10] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, no. 11, pp. 1222–1239, 2001.

[11] Y. Boykov and G. Funka-Lea, "Graph cuts and efficient nd image segmentation," *International journal of computer vision*, vol. 70, no. 2, pp. 109–131, 2006.

[12] D. Kroon and C. Slump, "Coherence filtering to enhance the mandibular canal in cone-beam ct data," 2009.

[13] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *International Journal of Computer Vision*, vol. 59, no. 2, pp. 167–181, 2004.