밝기 변화에 강인한 특징 기술자를 이용한 고품질 HDR 동영상 합성

Tu Vo Van, 이철
부경대학교
tuvovan@pukyong.ac.kr, chullee@pknu.ac.kr

Robust HDR Video Synthesis Using Illumination Invariant Descriptor

Tu Vo Van, Chul Lee
Pukyong National University

요 약

We propose a novel high dynamic range (HDR) video synthesis algorithm from alternatively exposed low dynamic range (LDR) videos. We first estimate correspondences between input fames using an illumination invariant descriptor. Then, we synthesize an HDR frame with the weights computed to maximize detail preservation in the output HDR frame. Experimental results demonstrate that the proposed algorithm provides high-quality HDR videos without noticeable artifacts.

1. 서론

Recent advances in digital imaging technology have enabled a variety of devices, including smartphones, to capture high-resolution images. However, common digital cameras can only capture images with significantly lower dynamic range than real scenes have [1]. While specially designed cameras can capture such high dynamic range (HDR) image, they are too expensive to be employed in consumer devices. Instead, software-based algorithms have been developed to synthesize an HDR image from multiple images taken with different exposure times [2].

In spite of recent progress in HDR still imaging, it still remains a significant challenge to capture HDR videos from conventional devices. This is because the motions between successive frames bring severe artifacts, degrading the qualities of the synthesized HDR frames. Several attempts have been made to address such problems in HDR video synthesis. The most practical technique is to capture alternating (long and short) exposures for every frame and then merge them using a correspondence estimation algorithm. For instance, Kang *et al.* [3] employed optical flow to estimate the correspondences between frames. Kalantari *et al.* [4] employed a patch-based correspondence estimation as well as optical flow, but its computational complexity is too high to be used in practical applications.

In this work, we propose a novel HDR video synthesis algorithm by addressing the motion problem between two adjacent frames by employing the dense adaptive self-correlation (DASC) [6] as feature descriptor, which it is robust to exposure variations. Experimental results show that the proposed algorithm provides higher synthesis qualities than the conventional algorithms [3]-[5].

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. NRF-2016R1C1B2010319).

2. 제안하는 기법

The most common approach to HDR video synthesis is to capture LDR frames with alternating exposures and then merge them to obtain the final HDR frame [3]-[5]. However, the quality of the synthesized HDR frame is affected by the accuracy of correspondence estimation between consecutive frames. In this work, we employ the DASC for feature descriptor [6]. The DASC $\mathrm{D_n(x,y)}$ at pixel location (x,y) in the nth frame is defined as a series of an adaptive self-correlation similarity for patches and a randomized receptive field pooling within a local support window.

Note that the input frames contain over- and/or under-exposed regions, where each pixel contains unreliable information. We first measure the reliability of the pixel value $I_n(x,y)$ in the nth frame using the contrast weight

$$c_n(x,y) = \begin{cases} 1, & \text{if } I_{\mathrm{th}} < I_n(x,y) < 255 - I_{\mathrm{th}} \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

where I_{th} is the threshold value.

To synthesize the nth HDR frame, accurate motions between the $(n-1)$th and $(n+1)$th frames to the nth frame should be estimated. However, when a pixel in the nth frame is poorly exposed, *i.e.*, $c_n(x,y) = 0$, we fail to estimate an accurate correspondence. To address this problem, we propose an accurate motion estimation algorithm for poorly exposed regions using the $(n-1)$th and $(n+1)$th frames. Specifically, let $D(x,y)=\|D_{n-1}(x-v_x,y-v_y)-D_{n+1}(x+v_x,y+v_y)\|$ and $\nabla I(x,y)=\|\nabla I_{n-1}(x-v_x,y-v_y)-\nabla I_{n+1}(x+v_x,y+v_y)\|$ denote the Euclidean distances between two descriptors and gradients, respectively, for corresponding pixels in frames $(n-1)$ and $(n+1)$ described by the motion vector (v_x,v_y). Then, we find the optimal motion vector (v_x,v_y) by minimizing

$$E(v_x,v_y)=\sum_{(x,y)\in W}\alpha\times D(x,y)^2+\beta\times\nabla I(x,y)^2 \qquad (2)$$

where α and β are weighting coefficients, and W denotes a set of the neighboring pixels.

Although we estimate correspondences that minimize (2), they may not be accurate. Therefore, in addition to contrast weight in (1), we measure the reliability of a motion vectors for the nth frame at each pixel location. Let $D'(x,y)=\|D_{n-1}(x-v_x,y-v_y)-D_n(x,y)\|$ denote the Euclidean distance between two descriptors for corresponding pixels in frame n and $n-1$. Then, the pixel reliability is defined as

$$r_n(x,y)=\exp\left(-\frac{D'(x,y)^2}{p(I_n(x,y))}\right) \qquad (1)$$

where $p(\cdot)$ is the weighting function that has the peak value at the middle of pixel value range [2].

Then, the weight for each pixel in the nth frame is computed by multiplying two weight terms, given by

$$w_n(x,y)=c_n(x,y)\times r_n(x,y). \qquad (2)$$

Finally, we synthesize the HDR frame by weighted averaging the irradiance map for each frame by

$$h(x,y)=\sum_{n=1}^{N}w_n(x,y)\times f(I_n(x,y)) \qquad (3)$$

where $h(x,y)$ is the HDR radiance value at pixel (x,y) and f denotes the camera response function [2].

3. 실험 결과

We evaluate the performance of the proposed HDR video synthesis algorithm on an alternatively exposed LDR dataset in Fig. 1. The parameter I_{th} in (1) is fixed to 25, and α and β in (2) are set to 0.4 and 0.6, respectively.

Fig. 1 compares the synthesized results of the *ThrowingTowel* sequence obtained by Kang *et al.*'s algorithm [3], Kalantari *et al.*'s algorithm [4], and Mangiat and Gibson's algorithm [5]. In Fig. 1(a), since Kang *et al.*'s algorithm fails to estimate accurate correspondences in poorly-exposed regions, it causes missing pixels on the fast moving objects, *e.g.*, the towel. Mangiat and Gibson's algorithm in Fig. 1(b) synthesizes an HDR image without missing pixel, but provides blurring artifacts. In Fig. 1(c), Kalantari *et al.*'s algorithm provides higher synthesis quality, but it still yields missing pixels on the moving objects. On the contrary, the proposed algorithm in Fig. 1(d) provides the best synthesis quality, since we find accurate motions between input frames, especially in poorly-exposed regions.

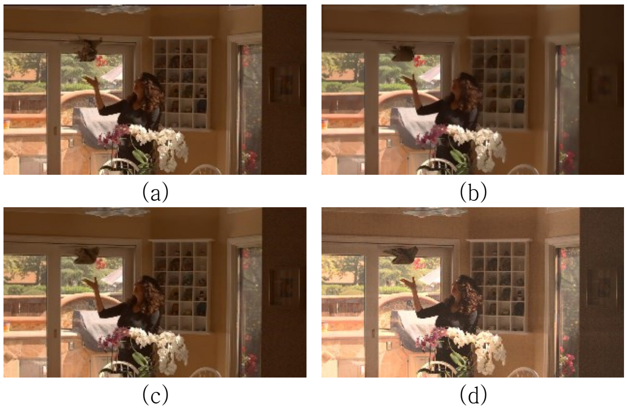

Fig. 1. HDR video synthesis results of the *ThrowingTowel* sequence by (a) Kang *et al.*'s algorithm [3], (b) Mangiat and Gibson's algorithm[5], (c) Kalantari *et al.*'s algorithm[4], and (d) the proposed algorithm.

4. 결론

We proposed a novel algorithm to synthesize HDR videos from alternatively exposed LDR frames. We estimated motion vectors between two input frames with different exposures by employing the DASC feature descriptor, and then warped them to the reference frame. Then, we synthesized high-quality HDR frames by merging with the weights computed from two reliability measures. Experimental results showed that the proposed algorithm can obtain high-quality HDR video sequences.

5. 참고문헌

[1] E. Reinhard, G. Ward, S. Pattanaik, P. Debevec, W. Heidrich, and K. Myszkowski, *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting*, 2nd ed. Morgan Kaufmann Publishers, 2010.

[2] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *Proc. ACM SIGGRAPH*, Aug. 1997, pp. 369-378.

[3] S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski, "High dynamic range video," *ACM Trans. Graphics*, vol. 22, no. 3, pp. 319-325, Jul. 2003.

[4] N. K. Kalantari, E. Shechtman, C. Barnes, S. Darabi, D. B. Goldman, and P. Sen, "Patch-based high dynamic range video," *ACM Trans. Graphics*, vol. 32, no. 6, pp. 202:1-8, Nov. 2013.

[5] S. Mangiat and J. Gibson, "High dynamic range video with ghost removal," in *Proc. SPIE Applications of Digital Image Processing XXXIII*, vol. 7798, Aug. 2010, pp. 779812:1-8.

[6] S. Kim, D. Min, B. Ham, S. Ryu, M. N. Do, and K. Sohn, "DASC: Dense adaptive self-correlation descriptor for multimodal and multi-spectral correspondence," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 2103-2112.