

Fast Volume Seam Carving With Multipass Dynamic Programming

Ryosuke Furuta, *Student Member, IEEE*, Ikuko Tsubaki, *Member, IEEE*, and Toshihiko Yamasaki, *Member, IEEE*

Abstract—In volume seam carving, i.e., seam carving for 3D cost volume, an optimal seam surface can be derived by graph cuts, resulting from sophisticated graph construction. To date, the graph-cut algorithm is the only solution for volume seam carving. However, it is not suitable for practical use because it incurs a heavy computational load. We propose a multipass dynamic programming (DP)-based approach for volume seam carving, which reduces computation time and memory consumption while maintaining a similar image quality as that of graph cuts. Our multipass DP scheme is achieved by conducting DP in two directions to accumulate the cost in a 3D volume and then tracing back to find the best seam. In our multipass DP, a suboptimal seam surface is created instead of a global optimal one, and it has been experimentally confirmed by more than 198 crowdsourced workers that such suboptimal seams are good enough for image processing. The proposed scheme offers two options: a continuous method that ensures the connectivity of seam surfaces and a discontinuous method that ensures the connectivity in only one direction. We applied the proposed volume seam carving method based on multipass DP to conventional video retargeting and tone mapping. These two applications are completely different; however, the volume seam carving method can be applied similarly by changing the axes of the cost volume. Even though the results obtained using our methods were similar to those obtained by graph cuts, our computation time was approximately 90 times faster than that of graph cuts and the memory usage was eight times smaller than that of graph cuts. We also extend the idea of tone mapping to the contrast enhancement method based on volume seam carving.

Index Terms—Contrast enhancement, multipass dynamic programming (DP), seam carving, tone mapping, video retargeting.

I. INTRODUCTION

THIS paper proposes a scheme for seam carving (finding seam surfaces) in 3D cost volume that features low computation time and memory consumption (Fig. 1). (Hereafter, we call seam carving for 3D cost volume as “volume

seam carving.”) Volume seam carving has been applied to various image processing tasks, such as video retargeting [1], video summarization [2], [3], and tone mapping [4]. The volume seam carving procedure is as follows: create a cost volume, find a seam surface that is less affected if removed, and then remove that seam surface. A seam surface is a 2D manifold in the cost volume and must be monotonic and connective.

To achieve the volume seam carving outlined above, to date, users can use only the graph-cut algorithm [1] because more efficient algorithms such as the dynamic programming (DP) algorithm used in still image seam carving [5] are not considered for use in this area. Rubinstein *et al.* [1] stated that volume seam carving cannot be solved with DP and deemed graph cuts the only choice. However, it is well known that the graph-cut algorithm utilizes a tremendous amount of computational time and memory when the number of nodes and edges increases. Consequently, to make problems solvable, Rubinstein *et al.* [1] proposed a multiresolution method and Chen and Sen [2] employed a video chunking method. However, the significant computation time and high memory usage of the graph-cut algorithm remain critical problems in the era of high-resolution images/videos.

In this paper, a multipass DP scheme for volume seam carving that realizes volume seam carving with low computation time and memory consumption is proposed. Here, we propose two options: a continuous mode and a discontinuous mode. The continuous method, which produces a connective seam surface, is presented in Section III-A. The discontinuous method, which produces a discontinuous seam surface, is presented in Section III-B. The discontinuous seam surface is connective in one direction and discontinuous in the other direction. We describe a monotonic manifold as a seam surface even if it is unconnected. In video retargeting, seams may be amenable to the discontinuous mode between frames, as discussed by Grundmann *et al.* [6], Chao *et al.* [7], and Yan *et al.* [8]. In contrast, in tone mapping, the connectivity of the seam surface is important. For these reasons, two different operational modes are presented here.

The proposed method is applied to video retargeting, tone mapping, and contrast enhancement to verify its efficacy. We present the results obtained, including the computation time and memory consumption, which, on average, decreased to approximately 1/90th and 1/8th, respectively, of those of

Manuscript received April 6, 2016; revised July 17, 2016 and September 14, 2016; accepted October 8, 2016. Date of current version May 4, 2018. This work was supported in part by Japan Society for the Promotion of Science (JSPS) under Grants 26700008 and 16J07267 and in part by Microsoft IJARC Core10. This paper was recommended by Associate Editor C. Zhu.

R. Furuta and T. Yamasaki are with the Department of Information and Communication Engineering, Graduate School of Information Science and Technology, The University of Tokyo, Tokyo 113-8656, Japan (e-mail: furuta@hal.t.u-tokyo.ac.jp; yamasaki@hal.t.u-tokyo.ac.jp).

I. Tsubaki is with the School of Media Science, Tokyo University of Technology, Tokyo 192-0982, Japan (e-mail: tsubakiik@stf.teu.ac.jp).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCSVT.2016.2620563

graph cut-based solutions. We also verify via large-scale subjective evaluation experiments that the suboptimal solution can produce an image quality roughly equivalent to that of graph cuts in volume seam carving applications. It is important to note that this paper presents a new and general optimization method (multipass DP) that is more efficient in terms of computation time and memory usage than the conventional graph cut-based methods for various types of volume-seam-carving-based applications, not for a specific application using the multipass DP.

The fundamental algorithm of our multipass DP and the experimental results of the video retargeting and tone mapping have already been presented in [9]. In this paper, we provide detailed explanations and present the results of additional experiments. In addition, we propose a new application of volume seam carving: contrast enhancement.

The remainder of this paper is organized as follows. Section II reviews related studies. Section III describes the details of the proposed multipass DP. Section IV presents the applications and experimental results. Finally, Section V summarizes our findings and concludes this paper.

II. RELATED WORK

A. DP for 3D Volume

DP has been utilized in various image processing applications other than seam carving. Let us take depth estimation as an example and compare it with seam carving. In depth estimation, a cost volume $C(d, x, y)$ (hereafter, for ease of understanding, the seam-carved axis is indicated as the first element) is created and a depth value d is estimated at each pixel so that the sum of the energy in all pixels is the minimum. This minimization is formulated in a Markov random field (MRF) framework to obtain a smooth depth map and can be solved using algorithms such as graph cuts and belief propagation [10]. DP is also used for this minimization when quick response is required. For example, scanline optimization is a well-known method in which energy minimization is solved in each line, considering the continuity only in the x -direction [11]. In other words, the continuity in the y -direction is not considered at all. Our multipass DP is different from this approach because 2D connectivity is guaranteed in its discontinuous mode and 3D connectivity is guaranteed in its continuous mode.

DP for 3D volume has been applied to depth estimation by Kim *et al.* [12], in which DP is performed in both the x - and the y -directions and discontinuous paths are penalized to maintain the smoothness of the depth map. Fukushima *et al.* [13] proposed a method similar to that of Kim *et al.* [12] for free viewpoint image rendering. Their method [13] performs DP in not only the forward x -direction but also the backward direction and then sums the two accumulated costs. After that, DP is performed in the y -direction similarly to [12]. Semiglobal matching (SGM) [14] is a DP-based technique for 3D volume in which cost accumulation is performed in 8 or 16 directions and a final disparity map is obtained in a winner-takes-all manner. Fukushima *et al.* [15] applied an idea similar to SGM [14] for real-time free viewpoint

image generation. Chen and Koltun [16] proposed a fast MRF optimization method based on DP for 3D volume, in which label assignments in even/odd lines in an image are optimally updated by DP fixing odd/even lines. However, unlike in our method, discontinuity is allowed in both the x - and the y -directions in these methods [12]–[16]. This is because, for example, discontinuous depth is reasonable at the boundary of two objects. On the other hand, continuity is sometimes required for seam carving, depending on the applications. The connectivity of a seam is assured by localizing the search range as done in [5]. If a seam is discontinuous, the image quality may be deteriorated by removing the pixels on the seam.

In [3], a DP-based video seam carving technique called ribbon carving is presented. However, the seam obtained in [3] is a chunk of connected rigid rectangle running vertically or horizontally. The technique is suitable for surveillance video summarization, but not for general-purpose volume seam carving. Moreover, the seam can be found by 2D DP rather than by multipass DP. From this point of view, the ribbon carving technique [3] is essentially different from our multipass DP.

B. Video Retargeting

Video retargeting is a nonlinear image resizing method for videos. Various methods have been presented for video retargeting, including cropping-and warping-based methods [17]–[23]. Seam carving has also been applied to video retargeting, with each frame image retargeted sequentially by seam carving [6]–[8]. In contrast to still image retargeting, temporal correlation is taken into account for video retargeting.

Rubinstein *et al.* [1] introduced volume seam carving for video retargeting. In their scheme, the video sequence is treated as a space–time volume and seam carving is extended from a seam on a 2D image to a seam surface in a 3D volume. The cost volume $C(x, y, t)$ is created by aligning the energy functions of all frames, and a seam surface is derived by energy minimization with graph cuts. Furthermore, all frame images are reduced in size by removing the pixels on the seam surface. Forward energy has also been proposed as an alternative to conventional backward energy to maintain the image quality. The forward energy criterion looks ahead to the resulting image after removing the seam, which takes into account the inserted energy due to the new edges created by previously nonadjacent pixels that become neighbors once the seam is removed. Conversely, the backward energy criterion looks behind at the image before removing the seam. This method can reduce temporal flicker because the removed seams are continuous in the temporal direction. However, it is time and memory consuming because it uses slow graph cuts to obtain an optimal seam surface, as discussed in Section I. This method can be significantly accelerated while maintaining the output quality by replacing graph cuts with our method.

Han *et al.* [24] proposed an algorithm that finds multiple seam surfaces simultaneously to improve the output quality. Although they use some techniques to accelerate the running time such as multiresolution strategy and searching for a small number of seam surfaces iteratively, their method is still slow

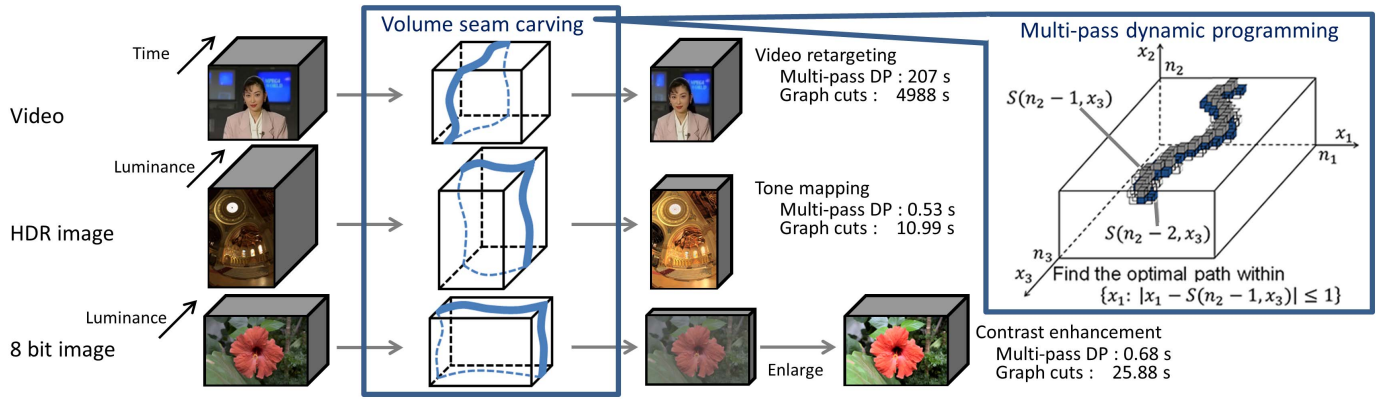


Fig. 1. Multipass DP for volume seam carving proposed to reduce computation time and memory consumption. Three example applications of multipass DP are presented.

because they employ graph cuts to obtain a globally optimal solution [24].

Recently, Jain *et al.* [25] have proposed a video retargeting method based on gaze tracking, in which a cropping window is determined by considering the gaze information. Katti *et al.* [26] utilized the gaze information to calculate the saliency costs for seam carving. Although these methods [25], [26] can obtain preferable results for users, they need to perform eye tracking to get the gaze information.

As discussed above, a variety of methods have been proposed for video retargeting. However, it is important to note that we are proposing a fast and memory-efficient optimization method for volume seam carving applications, not one specific to video retargeting.

C. Tone Mapping

Tone mapping is a technique for dynamic range compression that creates a low-dynamic-range (LDR) image from a high-dynamic-range (HDR) image. Many tone mapping methods have already been introduced. They can be roughly classified into global and local methods. Global methods have a uniform conversion characteristic in the whole image. Specifically, plural pixels that have the same luminance value in an input HDR image have the same luminance value in the output LDR image. Similar luminance values in the HDR image are grouped and replaced with a luminance value in the LDR image. Therefore, textures whose variation in luminance is small easily vanish in the global methods. A typical global method is histogram equalization, which converts the luminance distribution so that the histogram of the LDR image has a high variation. A high-contrast LDR image is obtained by histogram equalization; however, the contrast is sometimes too exaggerated. Larson *et al.* [27] subsequently presented an improved histogram equalization method that constrains the contrast.

On the other hand, local methods do not have a uniform conversion characteristic in the entire image and volume-seam-carving-based tone mapping is categorized as a local method. Recently, various methods that preserve detailed local

textures have been introduced. Fattal *et al.* [28] proposed a gradient domain method that manipulates the gradient field of the luminance image by attenuating the magnitudes of large gradients while preserving the fine details: the LDR image is obtained by solving a Poisson equation on the gradient field. Paris *et al.* [29] proposed local Laplacian filters (LLFs) that decompose an HDR luminance image into a Laplacian pyramid and compress the dynamic range in each level while keeping the local contrast. Aubry *et al.* [30] developed fast LLFs that improve LLFs in terms of computation time. Gu *et al.* [31] proposed local edge-preserving filters (LEPFs) that apply a multiscale decomposition to an HDR luminance image using an edge-preserving filter and compress the dynamic range in each level. He *et al.* [32] proposed a method based on a guided filter that decomposes an HDR image into a detail layer and a base layer and adds an enhanced detail layer to the compressed base layer. Recently, Ma *et al.* [33] proposed an iterative tone mapping method that optimizes the tone-mapped image quality index score. Because this method [33] requires a large number of iterations to converge, it is time consuming compared with the filter-based approaches [29], [31].

The histogram equalization method can create higher-contrast LDR images than the above methods, but it may delete the fine details. Local methods can retain fine details, but are prone to reducing the contrast. A method that uses seam surfaces was proposed by Tsubaki and Iwauchi [4] to satisfy both high-contrast and fine detail requirements by creating a cost volume $C(l, x, y)$ with luminance l and space axes. The cost volume is defined based on a local luminance histogram, and a graph is constructed similar to that by Rubinstein *et al.* [1]. Furthermore, a seam surface $S(x, y)$ is derived by graph cuts, and luminance values that are larger than the value in the seam surface are reduced by one. This process is iterated until the obtained image has the desired dynamic range. This method can obtain better or comparable results compared with filter-based state-of-the-art methods [29], [31]. However, this method needs subsampling in 3D volume to make processing time and memory usage reasonable because it employs graph cuts to obtain an optimal seam surface similar to [1]. This method can be significantly

accelerated while maintaining the output quality by replacing graph cuts with our method.

D. Contrast Enhancement

Contrast enhancement is a classical issue for which many algorithms have already been proposed. Histogram equalization and histogram modification are widely utilized and can create a very high contrast image [34]. However, the contrast of the image obtained is sometimes too exaggerated; furthermore, fine details are not preserved. These problems for contrast enhancement are the same as for tone mapping. Enhancement methods based on the retinex theory, which estimates the reflectance and the illumination at each pixel, are also utilized widely [35], [36]. They retain the fine details better than histogram modification; however, halo artifacts sometimes arise.

Image contrast is deteriorated by various factors, such as blur and haze. Some detail enhancement and dehazing approaches can also enhance the contrast. Yun *et al.* [37] adopted Laplacian pyramid decomposition to enhance the contrast of a low-pass filtered image by histogram equalization. An LLF has also been applied to detail enhancements other than tone mapping [29]. A dehazing method that utilizes a local color line model was more recently proposed by Fattal [38]. However, to the best of our knowledge, contrast enhancement using volume seam carving has not been reported to date. Furthermore, as stated above, contrast enhancement is similar to tone mapping. Thus, contrast enhancement using volume seam carving is proposed for the first time in this paper.

III. FAST VOLUME SEAM CARVING WITH MULTIPASS DP

We propose a new approach based on DP for volume seam carving. A cost volume is defined as $C(x_1, x_2, x_3)$, where x_1, x_2 , and x_3 are integers in the ranges $0 \leq x_1 < n_1$, $0 \leq x_2 < n_2$, and $0 \leq x_3 < n_3$, and n_1, n_2 , and n_3 are the resolutions of the cost volume in the x_1, x_2 , and x_3 directions, respectively. A seam surface that crosses the x_1 -axis is defined as $S(x_2, x_3)$. When the seam surface passes a coordinate (x_1, x_2, x_3) , we describe it as $S(x_2, x_3) = x_1$. $(x_1, x_2, x_3) = (x, y, t)$ for video retargeting, which reduces the x resolution at each frame t . (n_1, n_2) is the image size and n_3 is the number of frames. The value of a cost volume $C(x_1, x_2, x_3)$ at each coordinate is obtained in the same manner as conventional methods. For example, $C(x, y, t) = |(\partial/\partial x)I(x, y, t)| + |(\partial/\partial y)I(x, y, t)|$, which is defined as the backward energy for video retargeting [1] ($I(x, y, t)$ is the pixel value at the location (x, y) in the t th frame).

We present two methods: a continuous method and a discontinuous method. These two methods are proposed because different applications require different characteristics in the seam surface, as discussed in Section I.

A. Continuous Method to Obtain a Connective Seam Surface

1) *Step 1 (Accumulation Along the x_2 -Axis)*: In this step, cost values are accumulated in each x_1 - x_2 plane along the

x_2 -axis, from $x_2 = 0$ to $n_2 - 1$, with minimization just like in the original seam carving for image retargeting [5]. An accumulated cost function A_1 is obtained by

$$\begin{cases} A_1(x_1, x_2, x_3) = C(x_1, x_2, x_3) + \min_{j \in \{-1, 0, 1\}} \\ A_1(x_1 + j, x_2 - 1, x_3), \quad (x_2 > 0) \\ A_1(x_1, 0, x_3) = C(x_1, 0, x_3). \end{cases} \quad (1)$$

Here, C is the cost defined for the volume. Note that the minimum cost of $A_1(x_1, x_2, x_3)$ is not computed yet when arriving at the end point x_2 , unlike the original seam carving.

2) *Step 2 (Accumulation Along the x_3 -Axis and Determination of a Seam in the x_1 - x_3 Plane at $x_2 = n_2 - 1$)*: A seam is derived in the x_1 - x_3 plane at $x_2 = n_2 - 1$ in this step. The accumulated cost function $A_1(x_1, x_2, x_3)$ is further accumulated in the x_3 -direction, from $x_3 = 0$ to $n_3 - 1$, with minimization, and an accumulated cost function A_2 is obtained. The j value selected at the minimization is saved in a path $P_2(x_1, n_2 - 1, x_3)$

$$\begin{cases} A_2(x_1, x_2, x_3) = A_1(x_1, x_2, x_3) + \min_{j \in \{-1, 0, 1\}} \\ A_2(x_1 + j, x_2, x_3 - 1), \quad (x_3 > 0) \\ P_2(x_1, x_2, x_3) = \arg \min_{j \in \{-1, 0, 1\}} A_2(x_1 + j, x_2, x_3 - 1) + x_1, \\ (x_3 > 0) \\ A_2(x_1, x_2, 0) = A_1(x_1, x_2, 0). \end{cases} \quad (2)$$

Then, x_1 , which minimizes $A_2(x_1, n_2 - 1, n_3 - 1)$ in the current x_2 , is chosen and assigned to $S(n_2 - 1, n_3 - 1)$. A seam is obtained as the optimal path by following P_2 from $x_3 = n_3 - 1$ to 0:

$$\begin{cases} S(x_2, n_3 - 1) = \arg \min_{x_1} A_2(x_1, x_2, n_3 - 1) \\ S(x_2, x_3) = P_2(S(x_2, x_3 + 1), x_2, x_3 + 1), \quad (x_3 < n_3 - 1). \end{cases} \quad (3)$$

3) *Step 3 (Accumulation Along the x_3 -Axis and Determination of a Seam in Each x_1 - x_3 Plane)*: A seam in each x_1 - x_3 plane is derived in this step, starting from $x_2 = n_2 - 2$ and then reducing x_2 by one. First, the accumulated cost function A_2 is updated at each x_1 - x_3 plane

$$A_2(x_1, x_2, x_3) = \begin{cases} A_2(x_1, x_2, x_3), & (|x_1 - S(x_2 + 1, x_3)| \leq 1) \\ \infty, & (\text{otherwise}). \end{cases} \quad (4)$$

This update has the effect of making the seams between consecutive x_2 parameters connected.

Second, an accumulated cost function $A_2(x_1, x_2, x_3)$ and a path $P_2(x_1, x_2, x_3)$ are obtained by accumulating $A_1(x_1, x_2, x_3)$ in the x_3 -direction, from $x_3 = 0$ to $n_3 - 1$, in the same way as in (2). Then, x_1 , which minimizes $A_2(x_1, x_2, n_3 - 1)$, is chosen and assigned to $S(x_2, n_3 - 1)$. A seam is obtained as the optimal path by following P_2 from $x_3 = n_3 - 1$ to zero in the same way as in (3). Subsequently, x_2 is reduced by one and the derivation of a seam in the next x_2 is started from (2). The complete seam surface is obtained when the process ends at $x_2 = 0$.

Fig. 2 shows the continuous DP process. The gray path shows the seam obtained at $x_2 = n_2 - 1$. By creating a seam

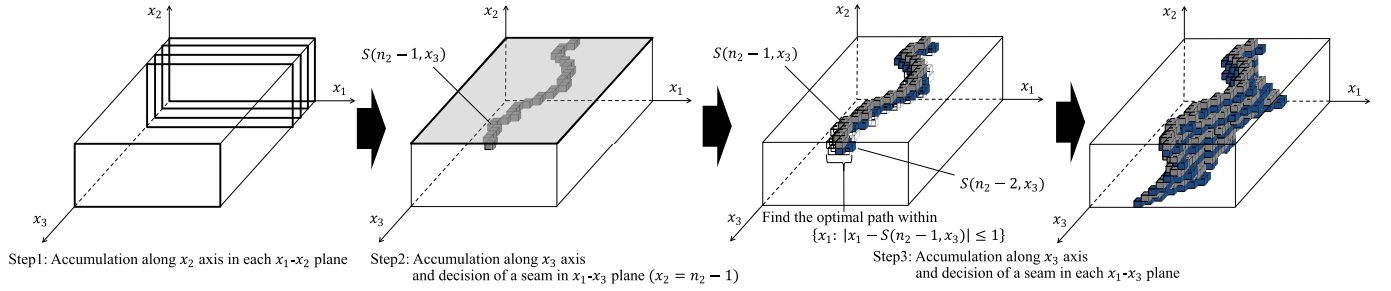


Fig. 2. Continuous DP process. The seam surface is connected in both the x_2 - and the x_3 -directions.

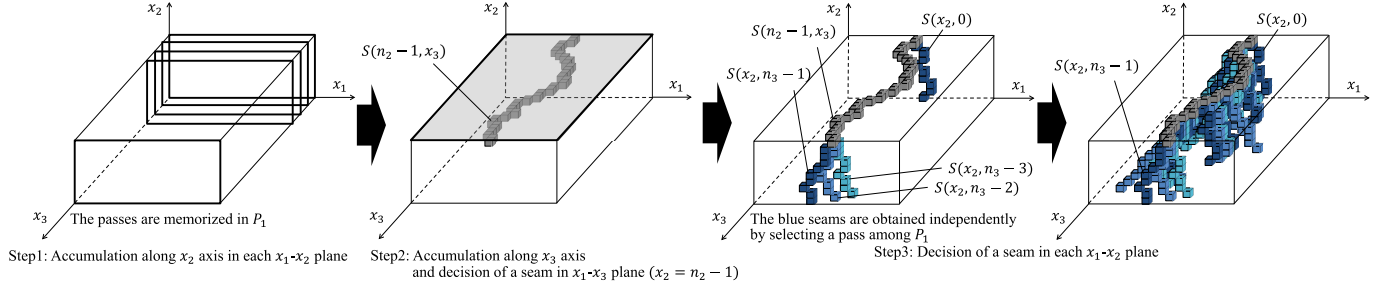


Fig. 3. Discontinuous DP process. The seam surface is connected in the x_2 -direction. The seam obtained as an intersection between the seam surface and the x_1 - x_3 plane at $x_2 = n_2 - 1$ is connected in the x_3 -direction. The seams in different x_1 - x_3 planes are not necessarily connected.

in each x_1 - x_3 plane successively so as to be connected to the seam in the previous x_2 , the continuous method enables the seam surface to become totally connected in the x_2 -direction. In addition, the seam obtained in each x_2 is connected in the x_3 -direction because j is selected from among $\{-1, 0, 1\}$ in (2). Hence, the obtained seam surface $S(x_2, x_3)$ is guaranteed to be connected.

B. Discontinuous Method to Obtain a Discontinuous Seam Surface

Steps 1 and 2 are virtually the same as in Section III-A. The only difference is that unlike the continuous method, a path P_1 in the x_2 direction is saved in (1) to give

$$P_1(x_1, x_2, x_3) = \arg \min_{j \in \{-1, 0, 1\}} A_1(x_1 + j, x_2 - 1, x_3) + x_1, \quad (x_2 > 0). \quad (5)$$

1) *Step 3 (Determination of a Seam in Each x_1 - x_2 Plane)*: A seam in each x_1 - x_2 plane is derived in this step. By selecting a path among those in P_1 that crosses seam $S(n_2 - 1, x_3)$ in the x_1 - x_3 plane at $x_2 = n_2 - 1$, the discontinuous method can obtain a seam independently in each x_1 - x_2 plane

$$S(x_2, x_3) = P_1(S(n_2 - 1, x_3), x_2 + 1, x_3), \quad (x_2 < n_2 - 1). \quad (6)$$

The complete seam surface is derived by obtaining a seam in every x_1 - x_2 plane.

Fig. 3 shows the process followed by the discontinuous method. The gray path shows the seam obtained at $x_2 = n_2 - 1$. The seam obtained in each x_3 is connected in the x_2 -direction because j is selected from among $\{-1, 0, 1\}$ in (5). Although the seam in $x_2 = n_2 - 1$ is connected in the x_3 -direction

because of (2), any seam obtained as an intersection of the seam surface and other x_1 - x_3 planes is not guaranteed to be connected because (3) is calculated independently at each x_3 . The connective seam in the x_1 - x_3 plane at $x_2 = n_2 - 1$, however, has an effect that makes the seam surface prone to connecting in other x_1 - x_3 planes.

C. Concrete Example

For ease of understanding, we explain the concrete process of the proposed method in video retargeting. Let (x, y, t) be the pixel location (x, y) in the t th frame in the video. For video retargeting to reduce the width of the image, DP is first conducted in each x - y plane to the direction $y = n - 1$ (where n is the image height). Specifically, the cost is accumulated at the x - t plane ($y = n - 1$). Then, the second DP is applied to the x - t plane, where $y = n - 1$. For the continuous mode, the best seam in the x - t plane ($y = n - 2$) is generated to ensure the connectivity with that in the x - t plane ($y = n - 1$), and this process is repeated until the last x - t plane ($y = 0$). In this manner, connectivity is ensured even though the generated seam surface is suboptimal, not global optimal. For the discontinuous mode, the seam is searched by tracing back in each x - y plane. In this mode, connectivity is not guaranteed, but a lower energy path than that in the continuous mode can be found.

IV. APPLICATION AND RESULTS

To show that the proposed method is general, we applied our proposed multipass DP to three different applications. All experiments were performed on a machine with an Intel Core i7-2600 3.4-GHz CPU and 16 GB of RAM, with the methods implemented in C++. In our experiments, we also

tried to use belief propagation [39], but it did not work well for our applications.

A. Video Retargeting

1) *Method*: Video retargeting is achieved by applying the proposed method to the energy functions presented by Rubinstein *et al.* [1]. When the backward energy is utilized, the procedure is as follows. First, the cost volume is calculated. Then, a seam surface is derived using the proposed method. Finally, the image size is reduced by one at each frame by removing the pixel on the seam surface. This process is iterated until the image size has reached the target size.

When the forward energy is utilized, the methods described in Sections III-A and III-B should be modified as follows. For the continuous method, (1) is replaced by

$$\begin{cases} A_1(x_1, x_2, x_3) = \min [A_1(x_1 - 1, x_2 - 1, x_3) + C_L(x_1, x_2, x_3), \\ \quad A_1(x_1, x_2 - 1, x_3) + C_U(x_1, x_2, x_3), \\ \quad A_1(x_1 + 1, x_2 - 1, x_3) + C_R(x_1, x_2, x_3)] \quad (x_2 > 0) \\ A_1(x_1, 0, x_3) = \\ \quad \min [C_L(x_1, x_2, x_3), C_U(x_1, x_2, x_3), C_R(x_1, x_2, x_3)], \end{cases} \quad (7)$$

$$\begin{cases} C_L(x_1, x_2, x_3) = \\ |I^t(x + 1, y) - I^t(x - 1, y)| + |I^t(x, y - 1) - I^t(x - 1, y)| \\ C_U(x_1, x_2, x_3) = |I^t(x + 1, y) - I^t(x - 1, y)| \\ C_R(x_1, x_2, x_3) = \\ |I^t(x + 1, y) - I^t(x - 1, y)| + |I^t(x, y - 1) - I^t(x + 1, y)| \end{cases} \quad (8)$$

where $(x, y, t) = (x_1, x_2, x_3)$ and $I^t(x, y)$ is a luminance image of the t th frame. For the discontinuous method, (5) is replaced in the same manner as in the continuous method. We adopted the forward energy for video retargeting in all the experiments.

2) *Experimental Results*: We tested the proposed method with eight videos. The results for five videos are shown in Fig. 4 (the others are shown in the Supplementary Material).¹ The resolutions (width, height, and the number of frames) of the original videos 1–5 in Fig. 4 are (352, 288, 300), (148, 144, 131), (282, 288, 91), (540, 280, 99), and (540, 280, 97), respectively. For comparison, the results for the seam surfaces derived by graph cuts are shown in Fig. 4(d). A simple multiresolution method was adopted for Fig. 4(d). More specifically, first, the images were downsampled by four, without any prefiltering, and a preliminary seam surface was derived by graph cuts. Then, the maximum and minimum values of the preliminary seam surface $x_{1\max}$ and $x_{1\min}$ were calculated. Finally, a final seam surface was derived from the full-resolution images by graph cuts within the range of x_1 from $(4x_{1\min} - 4)$ to $(4x_{1\max} + 4)$. No multiresolution or downsampling was adopted for Fig. 4(b) and (c) because it is not necessary.

Videos 1 and 2 are scenes in which foreground objects are moving and the background is stable. The other videos are scenes in which both the foreground and the background are moving. In videos 1 and 2, notable deterioration is not created in Fig. 4(b) and (d). In contrast, the foreground or background is shaking between frames in Fig. 4(c) because connectivity is not guaranteed. In video 3, visible deterioration is not created in Fig. 4(b) and (c). In videos 4 and 5, some players are distorted, as shown in Fig. 4(b) and (d). It is clear that the continuous method is suited for videos that have little motion and that the discontinuous method is suited for videos whose motion is strenuous. The discontinuous method may be suited for high-motion scenes compared with graph cuts. If a salient object is moving, a seam near the object should leap the object between adjacent frames to avoid distorting it. When graph cuts or the continuous DP is utilized, the seam cannot leap the object and then may cross it. On the other hand, our discontinuous multipass DP can avoid this problem by its discontinuous nature.

Fig. 5 shows the seam surfaces obtained for videos 1 and 5 in Fig. 4. Fig. 5(a) and (c) verifies that the seam surfaces are connective because the values change smoothly over the entire seam surface. The seam surfaces obtained by all methods for video 1 have horizontal stripes. This occurs because the object has moved slightly; therefore, the seam surface crosses at virtually the same position in every frame. In Fig. 5(b), video 1, some vertical stripes are shown, which cause background shaking. In Fig. 5(b), video 5, many vertical stripes are shown; however, the stripes vanish at the bottom of the seam surface because the seam surface is connective in the $x_2 = n_2 - 1$ plane.

Fig. 6 shows histograms of the discontinuous level in seam surfaces using the discontinuous method. Discontinuous level k is defined as $k = |S(y, t + 1) - S(y, t)|$. The vertical axis shows the frequency of k averaging in all seam surfaces. $k = 0$ and 1 indicate that the seam surface is connective at a specific point. Other values of k indicate that the seam surface is discontinuous at those points. In Fig. 6(a), because the sum of the frequency at $k = 0$ and 1 is approximately 60%, the seam surfaces are connective over approximately 60% of the area. In Fig. 6(b), because the sum of the frequency at $k = 0$ and 1 is approximately 30%, the seam surfaces are connective in only approximately 30% of the area. Fig. 6(b) clarifies why the seam surfaces for video 5 are more discontinuous than those for video 1.

The computation time and maximum memory consumption are shown in Table I(a) and (b), respectively. Continuous and discontinuous DP require virtually the same processing time in each video. By replacing graph cuts with multipass DP, we can reduce the computation time to 1.5%. The maximum memory consumption is reduced to 5.6% by replacing graph cuts with continuous DP and to 11% with discontinuous DP, even without the multiresolution or downsampling. Note that the computation time and memory usage of our multipass DP would increase by the order of $O(n)$, where n is the number of voxels in the cost volume. On the other hand, those of graph cut-based methods grow more rapidly with n .

¹The video results and the Supplementary Material are available on our Web page (<https://www.hal.t.u-tokyo.ac.jp/~furuta/>).



Fig. 4. Results of the video retargeting. The top and bottom rows are different frames in videos 3–5. (a) Original. (b) Multipass DP. (c) Multipass DP (continuous). (d) Graph cuts (discontinuous).

3) *Subjective Analysis:* We conducted subjective evaluation experiments to verify that the quality of the videos retargeted by our multipass DP is almost equal to that of graph cuts with respect to human perception. We used a crowdsourcing

service, and 163 participants took part in the test. The flow of the test was as follows. First, subjects were asked to watch four videos (the original and three videos retargeted by our continuous method, our discontinuous method, and graph cuts)

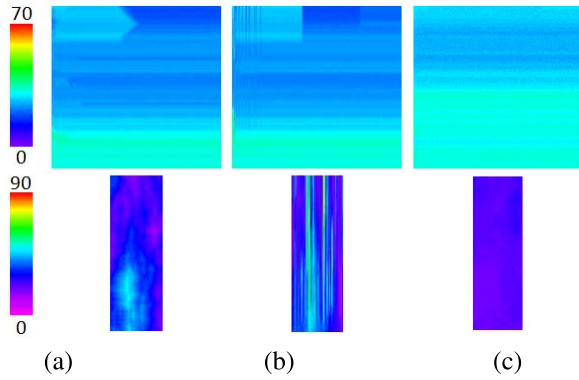


Fig. 5. Seam surfaces obtained in Fig. 4 in the first iteration. The top and bottom lines are for videos 1 and 5, respectively. The horizontal axis shows the time. (a) Multipass DP (continuous). (b) Multipass DP (discontinuous). (c) Graph cuts.

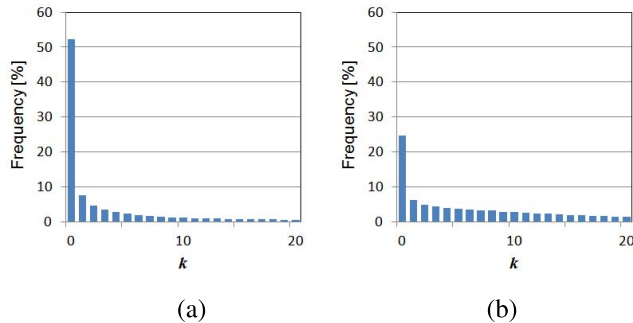


Fig. 6. Histograms of the discontinuous level in seam surfaces using the discontinuous method. (a) Video 1. (b) Video 5.

on our Web page, in which the original video was displayed at the center of the top row and the three retargeted videos were arranged horizontally on the row beneath. After that, subjects were asked to “choose the best video in terms of quality.” They could replay the videos any number of times and were able to choose one of four choices: “Video A is the best,” “Video B is the best,” “Video C is the best,” or “Cannot notice the difference.” The three retargeted videos and four choices were randomly ordered to remove any bias. Each video was scaled by a factor of $\min(\frac{320}{w}, \frac{240}{h})$, where w and h are the video width and height, respectively, to fit it in a 320×240 window. The aspect ratio was not changed. We used the eight videos mentioned in Section IV-A2 along with two other videos for dummy questions, in which three identical retargeted videos were displayed on the lower row. Therefore, the subjects were asked to answer ten questions including two dummy questions. The order of the ten questions was also random.

The result of the five videos in Fig. 4 is shown in Table II(a) (the others are shown in the Supplementary Material). We observed that, on average, approximately half of the 163 subjects answered that there was no notable difference between the retargeted videos. For video 1, which contains little movement in the scene, the graphcut method was the best of the three methods. However, our continuous method was the best for videos 2 and 3, which contain moderate motion. In addition, our discontinuous method was the best

TABLE I
COMPUTATION TIME AND MEMORY CONSUMPTION FOR FIG. 4.
(a) COMPUTATION TIME (s). (b) MAXIMUM MEMORY CONSUMPTION (MB)

	Multi-pass DP (continuous)	Multi-pass DP (discontinuous)	Graphcuts
Video 1	207	213	4988
Video 2	13	13	347
Video 3	83	85	4645
Video 4	445	449	43,419
Video 5	436	436	24,115
Average	237	239	15,503

	Multi-pass DP (continuous)	Multi-pass DP (discontinuous)	Graph cuts
Video 1	124	241	2.4×10^3
Video 2	15	27	358
Video 3	41	77	721
Video 4	65	123	1.2×10^3
Video 5	64	121	866
Average	62	118	1.1×10^3

TABLE II
NUMBER OF SUBJECTS WHO PREFERRED THE RETARGETED VIDEO OF EACH METHOD. (a) ALL SUBJECTS. (b) SUBJECTS WHO WERE NOT FOOLED BY THE DUMMY QUESTIONS

	Multi-pass DP (continuous /discontinuous)			Graph cuts	Cannot notice the difference	Total
Video 1	24	/	8	39	92	163
Video 2	42	/	13	28	80	163
Video 3	40	/	17	22	84	163
Video 4	14	/	50	25	74	163
Video 5	19	/	39	37	68	163
Average	27.8	/	25.4	30.2	79.6	163

	Multi-pass DP (continuous /discontinuous)			Graph cuts	Cannot notice the difference	Total
Video 1	6	/	2	11	53	72
Video 2	14	/	2	10	46	72
Video 3	11	/	2	5	54	72
Video 4	3	/	15	8	46	72
Video 5	4	/	12	11	45	72
Average	7.6	/	6.6	9.0	48.8	72

for videos 4 and 5, which contain extreme motion. On average, both of our methods obtained a score that was near that of graph cuts, and the sum of the two methods was significantly more than that of graph cuts. Therefore, in practical use, users can obtain an output that is equivalent to that of graph cuts by selecting the best result after applying both of our methods to the input video because our methods are very fast to calculate.

Table II(b) shows the results of the 72 subjects who correctly selected the choice “Cannot notice the difference” in the dummy questions, i.e., the subjects who were not fooled by the dummy questions. We can regard the result of Table II(b) as a more reliable result than that of Table II(a) because the noisy results of dishonest subjects were filtered out. We observed

that, on average, 67.8% of these 72 subjects answered that there were no notable differences between the retargeted videos.

As discussed above, the majority of subjects could perceive no differences. In addition, here, we excluded the subjects who selected the choice ‘‘Cannot notice the difference’’ and show that there was no significant difference between the number of subjects who preferred graph cuts and multipass DP. Following [40] and [41], we conducted a significance test of score differences. We regarded the score of each method as the sum of the numbers in each column in Table II(a) (e.g., the score of the continuous DP is $24 + 42 + 40 + 14 + 19 = 139$). The scores of the continuous DP, discontinuous DP, and graph cuts were 139, 127, and 151, respectively, and total score was 417. Unlike in [40] and [41], our subjective evaluation test was not a paired comparison because we asked the subjects to compare three videos at one time. Therefore, we regarded these scores as the result of a paired comparison where the number of subjects was $n = 417/3C_2 = 139$ and the number of objects (methods) was $t = 3$. We needed to find a value R so that there was no significant difference between the two methods with the confidence level α if the score difference was less than or equal to R . It can be found from (9) that

$$P\left(W_{t,\alpha} \geq \frac{2R - 0.5}{\sqrt{nt}}\right). \quad (9)$$

We set the confidence level $\alpha = 0.05$ and obtained $W_{3,0.05} = 3.31$ because the values of $W_{t,\alpha}$ were tabulated in [42]. We obtained $R = 34.05$ from (9) and observed that the results of the three methods (continuous DP, discontinuous DP, and graph cuts) were not significantly different because the score differences between any pairs of the three methods were less than R .

B. Tone Mapping

1) *Method*: Most tone mapping methods compress the Y (luminance) values of the YCbCr color space. We also compress the Y values, but we convert them using the logarithmic function in advance. Luminance image is defined by $L^0(x, y) = \log Y(x, y)$, where $Y(x, y)$ is the Y value of the input HDR image $I(x, y; c)$. First, a guided filter [32] is performed on the luminance image $L^0(x, y)$ and a high-frequency component $H(x, y)$ is calculated by

$$H(x, y) = \frac{L^0(x, y) - \bar{L}(x, y)}{L_{\max} - L_{\min}} \quad (10)$$

where $\bar{L}(x, y)$ is the image obtained by the guided filtering. L_{\min} and L_{\max} are the minimum and maximum values of $L^0(x, y)$, respectively. The initial luminance image $L(x, y)$ is defined as

$$L(x, y) = (r_0 - 1) \frac{\bar{L}(x, y) - L_{\min}}{L_{\max} - L_{\min}} \quad (11)$$

where r_0 is a constant that determines an initial luminance range.

Next, the cost volume is defined based on a local luminance histogram at each block. To reduce the memory consumption, we subsample the cost volume in both the space and

the luminance directions in advance. Subsampled coordinate and luminance values are expressed as $\tilde{x} = x/(2p + 1)$, $\tilde{y} = y/(2p + 1)$, and $\tilde{l} = l/q$ where $2p + 1, q$ is the subsampling parameter of the space and luminance, respectively. The cost volume is defined as

$$C(\tilde{l}, \tilde{x}, \tilde{y}) = \sum_{i=-w}^w \sum_{j=-w}^w \sum_{k=0}^{q-1} f(\tilde{l} + i, \tilde{x} + j, \tilde{y} + k) \quad (12)$$

$$f(l, x, y) = \begin{cases} 1, & l = \text{floor}(L(x, y)) \\ 0, & l \neq \text{floor}(L(x, y)) \end{cases}$$

where w is a window size and $\text{floor}()$ indicates rounding down to an integer.

Then, a subsampled seam surface $\tilde{S}(\tilde{x}, \tilde{y})$ is derived from the cost volume $C'(\tilde{l}, \tilde{x}, \tilde{y})$ using the proposed continuous method. A seam surface in full resolution $S(x, y)$ is obtained from $\tilde{S}(\tilde{x}, \tilde{y})$ by simple bilinear interpolation. Converting the luminance image using the obtained seam surface is expressed simply as

$$L'(x, y) = \begin{cases} L(x, y) - q, & L(x, y) > qS(x, y) \\ L(x, y), & L(x, y) \leq qS(x, y) \end{cases} \quad (13)$$

where $L'(x, y)$ is a converted luminance image. This equation has the effect of reducing the maximum luminance value by q . The meaning of this process is that the cuboid is divided into two parts by the seam surface: one part is pruned on the surface and both parts are again bonded to each other. As a result, the length of the luminance direction of the cuboid becomes shorter by q . The process from (12) to (13) is iterated until the dynamic range of $L'(x, y)$ becomes the target range r_T . The parameter q is calculated at every iteration using the current range r of $L(x, y)$ by $q = \text{ceil}(r/\rho)$, where $\text{ceil}()$ indicates rounding up to an integer. The parameter ρ controls the maximum size of the cost volume in the luminance direction.

Finally, the high-frequency component ($H(x, y)$ in (10)) is added to the obtained luminance image $L'(x, y)$ with enhancement and the LDR luminance image $L^L(x, y)$ is obtained. The final dynamic range is controlled to be r_L by linear transformation. r_L is the desired range of the LDR image, typically $r_L = 256$

$$L^L(x, y) = \frac{r_L}{r_T} L'(x, y) + \lambda(r_L - 1)H(x, y) \quad (14)$$

where λ is a constant that controls the level of enhancement. The LDR color image $I^L(x, y; c)$ is created from $L^L(x, y)$ by a method similar to that used by Fattal *et al.* [28]

$$I^L(x, y; c) = L^L(x, y) \left(\frac{I(x, y; c)}{Y(x, y)} \right)^{s_c} \quad (15)$$

where $c = \{R, G, B\}$ and the exponent s_c controls the color saturation of the resulting image. The method described above is almost the same as that of Tsubaki and Iwauchi [4]. The only difference is that graph cuts is replaced by the proposed multipass DP and the guided filter is utilized.

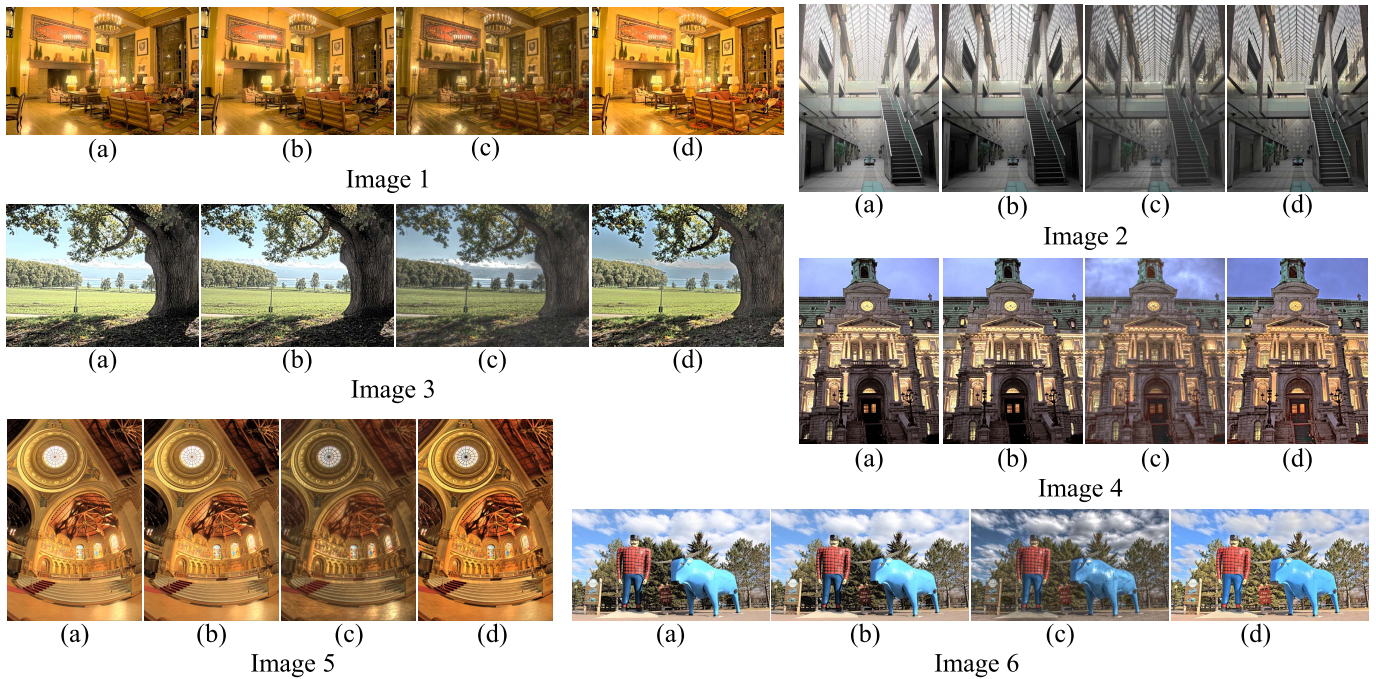


Fig. 7. Tone mapping results. (a) Multipass DP (continuous). (b) Graph cuts. (c) LLF [29]. (d) LEPF [31].



Fig. 8. Difference images for Fig. 7(a) and (b). The contrast was exaggerated by eight.

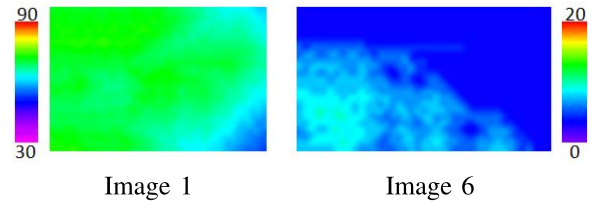


Fig. 9. Seam surfaces obtained by continuous DP for Fig. 7 in the last iteration.

2) *Experimental Results*: We tested the continuous method with ten HDR images obtained from [43]– [45]. The results for six images are shown in Fig. 7 (the others are shown in the Supplementary Material). The resolutions (width, height) of images 1–6 in Fig. 7 are (1000, 664), (760, 1016), (512, 381), (401, 535), (512, 768), and (1000, 563), respectively. The parameters $r_0 = 1024$, $r_T = 512$, $r_L = 256$, $p = 12$, $w = 12$, $\rho = 128$, $\lambda = 4$, and $s_c = 0.5$ were used. In the guided filtering, the window size was five and $\varepsilon = 0.01$. For comparison, the results for the seam surfaces derived by graph cuts are shown in Fig. 7(b). The images in Fig. 7(a) and (b) look very similar. Fig. 8 shows the difference images for Fig. 7(a) and (b), where the contrast was exaggerated by eight. The mean square errors between Fig. 7(a) and (b) were 9.0 and 20.5 in images 1 and 6, respectively. Because the peak signal-to-noise ratio (PSNR) was approximately 40, there were no visible differences between these images. Fig. 9 shows the seam surfaces obtained for Fig. 7 in the last iteration. The seam surface appears to be smooth and keeps the connectivity with the adjacent blocks.

Fig. 7(c) and (d) shows the results for LLF [29] and LEPF [31], respectively, for comparison. The parameters

$\sigma_r = \log(2.5)$, $\alpha = 0.5$, and $\beta = 0$ were used for LLF. On the whole, the details are clear in Fig. 7(c) and (d); however, the contrast in Fig. 7(a) and (b) is higher than that in Fig. 7(c) and (d). We evaluated these images using Q_{MOS} , sharpness, and variance. Q_{MOS} , the mean opinion score prediction from 0 to 100, where 100 is the highest quality, was proposed by Mantiuk *et al.* [46] for evaluating image quality. Sharpness is defined as $S = \frac{1}{N} \sum |I|$, where N is the number of pixels in image I [31]. Furthermore, we calculated the variance of the luminance image to express the image contrast. Fig. 10 shows the Q_{MOS} , sharpness, and variance of the images in Fig. 7. Multipass DP and graph cuts had similar values in all the indexes. The LLF had the highest Q_{MOS} , and the LEPF had a high sharpness. Multipass DP and graph cuts had high variance but had moderate values for Q_{MOS} and sharpness.

The computation time and maximum memory consumption are shown in Table III(a) and (b), respectively. By replacing graph cuts with continuous DP, we reduced the computation time to 2.2% and the maximum memory consumption to 51%. The number of iterations was 81 in Fig. 11(a) and (b) for each image.

Fig. 11 shows the computation time and maximum memory consumption when the subsampling parameter p was changed.

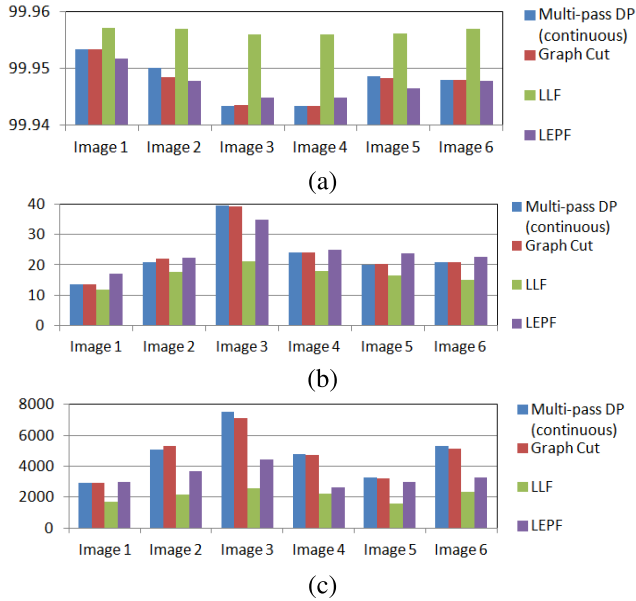


Fig. 10. Evaluation results for Fig. 7. (a) QMOS. (b) Sharpness. (c) Variance.

TABLE III
COMPUTATION TIME AND MEMORY CONSUMPTION FOR FIG. 7.
(a) COMPUTATION TIME (s). (b) MAXIMUM MEMORY CONSUMPTION (MB)

	(a)		(b)	
	Multi-pass DP (continuous)	Graph cuts	Multi-pass DP (continuous)	Graph cuts
Image 1	0.88	18.38	40	70
Image 2	4.17	245.29	46	82
Image 3	0.28	24.82	12	22
Image 4	0.28	3.69	14	24
Image 5	0.53	10.99	24	42
Image 6	0.76	26.09	34	90
Average	1.15	54.88	28	55

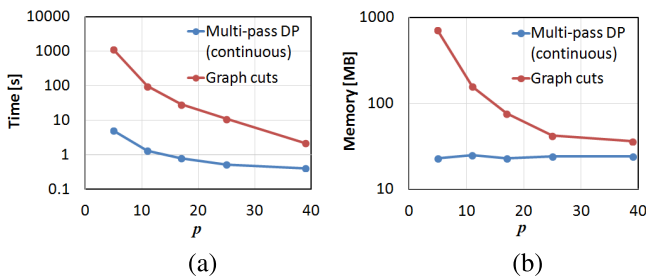


Fig. 11. Computation time and memory consumption for image 5 in Fig. 7. The horizontal axis shows the subsampling parameter of the space. (a) Computation time. (b) Memory consumption.

The ratio of the time for multipass DP to the time for graph cuts was smaller when p was smaller. The memory consumption for multipass DP did not vary by p .

3) *Subjective Analysis*: We conducted subjective evaluation tests for the tone mapping results, similar to those for video retargeting in Section IV-A3, to verify that the image quality of the results of our proposed method was almost equal to that of graph cuts with respect to human perception. The number of subjects was 198, and only two tone mapped images

TABLE IV
NUMBER OF SUBJECTS WHO PREFERRED THE TONE MAPPED IMAGE OF EACH METHOD. (a) ALL SUBJECTS. (b) SUBJECTS WHO WERE NOT FOOLED BY THE DUMMY QUESTIONS

	(a)			Total
	Multi-pass DP (continuous)	Graph cuts	Cannot notice the difference	
Image 1	23	46	129	198
Image 2	86	48	64	198
Image 3	28	58	112	198
Image 4	30	28	140	198
Image 5	36	35	127	198
Image 6	13	38	147	198
Average	36.0	42.2	119.8	198

	(b)			Total
	Multi-pass DP (continuous)	Graph cuts	Cannot notice the difference	
Image 1	3	11	87	101
Image 2	34	14	53	101
Image 3	7	17	77	101
Image 4	5	6	90	101
Image 5	7	9	85	101
Image 6	1	5	95	101
Average	9.5	10.3	81.2	101

were displayed side by side on the Web page. Subjects were asked to select one of three choices: “Image A is better,” “Image B is better,” or “Cannot notice the difference.” We used the ten HDR images mentioned in Section IV-B2 and three other images for the dummy questions, in which two identical images were displayed.

The results for the six images in Fig. 7 are shown in Table IV(a) (the others are shown in the Supplementary Material). We did not compare our method with the results of LLF [29] in Fig. 7(c) and LEPF [31] in Fig. 7(d). These methods are not volume-seam-carving-based methods, and the purpose of this experiment was to verify that the same image quality can be obtained by our multipass DP and graph cuts. We observed that, on average, 60.5% of the 198 subjects noted no difference between the two tone-mapped images.

Table IV(b) shows the results of the 101 subjects who were not fooled by the three dummy questions. We observed that, on average, 80.4% of the 101 subjects noted no difference between the two tone mapped images.

We conducted a significance test of score difference for Table IV(a), similar to that in video retargeting. The scores of continuous DP and graph cuts were 216 and 253, respectively, and the total score was 469. We regarded these scores as the results of a paired comparison where the number of participants was $n = 469$. We obtained $W_{2,0.05} = 3.64$ from [42] and calculated $R = 42.67$ from (9). Because the score difference was less than R , we observed that there was no significant difference between the two methods.

C. Contrast Enhancement

1) *Method*: We propose a contrast enhancement method based on luminance range compression. First, the dynamic luminance range is reduced via a process similar to tone mapping in Section IV-B1. The luminance values that have low frequency in the local histogram are then shrunk in this



Fig. 12. Contrast enhancement results. (a) Input. (b) Multipass DP (continuous). (c) Graph cuts. (d) LLF [29].

process. Next, the dynamic range is reverted to the original range using the linear enlargement method. The local contrast is enhanced by this process because the spatial gradient of the luminance image is relatively enlarged.

The compression process is the same as in (10)–(13), with the exception that luminance value $Y(x, y)$ in the YCbCr color space is used for $L^0(x, y)$ directly ($L^0(x, y) = Y(x, y)$) instead of using a logarithmic function, and parameter q is fixed at 1. The linear enlargement is performed using

$$L''(x, y) = \frac{L_{\max} - L_{\min}}{r_T} L'(x, y) + L_{\min}. \quad (16)$$

The high-frequency component is added by

$$L^L(x, y) = L''(x, y) + \lambda(L_{\max} - L_{\min} + 1)H(x, y). \quad (17)$$

Finally, a color image $I^L(x, y; c)$ is created from $L^L(x, y)$ using (15).

2) *Experimental Results*: We tested the continuous method with the ten images obtained from [29] and [47]. The results for six images are shown in Fig. 12 (the others are shown in the Supplementary Material). The resolutions (width, height) of images 1–6 in Fig. 12 are (800, 533), (800, 533), (512, 512), (876, 584), (795, 532), and (876, 584), respectively. The parameters $r_0 = L_{\max} - L_{\min} + 1$, $r_T = 128$, $p = 12$, $w = 12$, $\lambda = 2$, and $s_c = 1.4$ were used. An 8-b luminance value was reduced to 7 b and then reverted to 8 b. In the guided filtering, the window size was three and $\varepsilon = 0.01$. For comparison, the results for the seam surfaces derived by graph cuts are shown in Fig. 12(c). The images in Fig. 12(b) and (c) look very similar, with the exception that the blue sky is brighter in Fig. 12(b) than in Fig. 12(c). Fig. 13 shows the difference images for Fig. 12(b) and (c), where the contrast was exaggerated by eight. The mean square errors between Fig. 12(b) and (c) were 0.2 and 2.6 in images 2 and 4, respectively. Because the PSNR was larger than 40, there were no visible differences between these images. Fig. 14 shows the seam surfaces obtained for Fig. 12(b) in the final iteration.

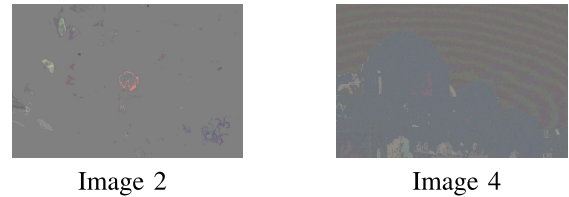


Fig. 13. Difference images between Fig. 12(b) and (c). The contrast was exaggerated by 8.

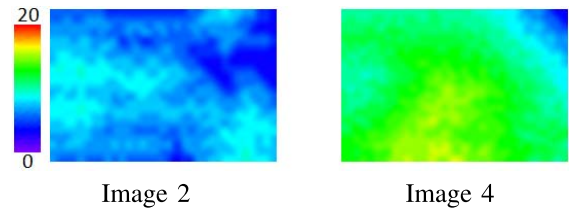


Fig. 14. Seam surfaces obtained by continuous DP for Fig. 12 in the last iteration.

The seam surface looks smooth and keeps the connectivity with the adjacent blocks.

Fig. 12(d) shows the results for the LLF method [29] for comparison. The parameters $\sigma_r = 0.4$, $\alpha = 0.25$, and $\beta = 1.0$ were used for the LLF. On the whole, the details were clear in Fig. 12(d); however, the contrast in Fig. 12(b) and (c) was higher than that in Fig. 12(d). We evaluated these images using sharpness and variance in the same way as for tone mapping. Fig. 15 shows the sharpness and variance of the images in Fig. 12. Multipass DP and graph cuts had similar values in both indexes. The results for the LLF were very sharp. The results for multipass DP and graph cuts had high variance, and the sharpness was higher than that in the original images.

The computation time and maximum memory consumption are shown in Table V(a) and (b), respectively. By replacing graph cuts with continuous DP, we reduced the computation time to 0.6% and the maximum memory consumption to 39%.

3) *Subjective Analysis*: We conducted subjective evaluation tests for the contrast enhancement results, similar to those

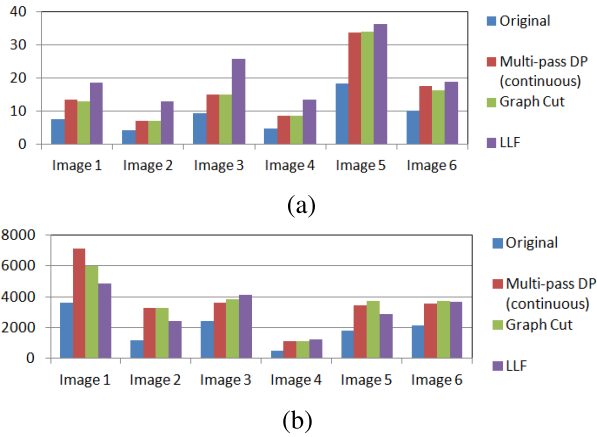


Fig. 15. Evaluation results for Fig. 12. (a) Sharpness. (b) Variance.

TABLE V
COMPUTATION TIME AND MEMORY CONSUMPTION FOR FIG. 12.
(a) COMPUTATION TIME (s) (b) MAXIMUM MEMORY CONSUMPTION (MB)

	(a)		(b)	
	Multi-pass DP (continuous)	Graph cuts	Multi-pass DP (continuous)	Graph cuts
Image 1	0.90	335.89	28	79
Image 2	0.68	25.88	27	68
Image 3	0.46	115.37	17	44
Image 4	0.62	24.12	30	71
Image 5	0.89	99.25	28	74
Image 6	0.92	200.16	33	86
Average	0.75	133.45	27	70

for tone mapping in Section IV-B3, to verify that the image quality of the results of our proposed method is almost equal to that of graph cuts with respect to human perception. The number of subjects was 191, and only two enhanced images were displayed side by side on the Web page. Subjects were asked to select one of three choices: “Image A is better,” “Image B is better,” or “Cannot notice the difference.” We used the ten images mentioned in Section IV-C2 and three other images for the dummy questions, in which two identical images were displayed.

The results for the six images in Fig. 12 are shown in Table VI(a) (the others are shown in the Supplementary Material). We did not compare our method with the results of the LLF [29] in Fig. 12(d). This latter method is not a volume-seam-carving-based method, and the purpose of this experiment is to verify that the same image quality can be obtained by our multipass DP as that of graph cuts. We observed that, on average, 34.7% of 191 subjects noted no difference between the two tone mapped images.

Table VI(b) shows the results of the 90 subjects who were not fooled by the three dummy questions. We observed that, on average, 46.3% of the 90 subjects noted no difference between the two enhanced images.

We conducted a significance test of score difference for Table VI(a), similar to that in tone mapping. The score of continuous DP and graph cuts was 476 and 272, respectively, and the total score was 748. We regarded these scores as the results of a paired comparison where the number of

TABLE VI
NUMBER OF SUBJECTS WHO PREFERRED THE ENHANCED IMAGE OF EACH METHOD. (a) ALL SUBJECTS. (b) SUBJECTS WHO WERE NOT FOOLED BY THE DUMMY QUESTIONS

	(a)			Total
	Multi-pass DP (continuous)	Graph cuts	Cannot notice the difference	
Image 1	74	75	42	191
Image 2	39	36	116	191
Image 3	37	102	52	191
Image 4	31	28	132	191
Image 5	16	123	52	191
Image 6	75	112	4	191
Average	45.3	79.3	66.3	191

	(b)			Total
	Multi-pass DP (continuous)	Graph cuts	Cannot notice the difference	
Image 1	30	36	24	90
Image 2	9	9	72	90
Image 3	12	42	36	90
Image 4	7	4	79	90
Image 5	1	53	36	90
Image 6	35	52	3	90
Average	15.7	32.7	41.7	90

participants was $n = 748$. We obtained $W_{2,0.05} = 3.64$ from [42] and calculated $R = 53.82$ from (9). Because the score difference was higher than R , we observed that there was a significant difference between the two methods. However, there were no visible differences between the results from continuous DP and graph cuts in some images, as discussed in Section IV-C2 (e.g., images 2 and 4 in Fig. 12). The quality improvement in other images is our future work.

V. CONCLUSION

In this paper, we have proposed a fast volume seam carving method based on multipass DP and applied it to video retargeting, tone mapping, and contrast enhancement. The seam surface obtained by the continuous method is monotonic and connective, as with the method by Rubinstein *et al.* [1]. Discontinuous DP creates a seam surface that is monotonic and connective in one direction. We verified that suboptimal seam surfaces obtained by our methods can be utilized sufficiently in those applications. We also showed that our methods create a seam surface that is approximately 90 times faster and consumes eight times less memory than conventional methods, which makes volume seam carving more practical.

Changing the axes enables video seam carving to be applied to entirely different applications. The characteristics of the cost volume and the optimal seam surfaces differ depending on the applications. In tone mapping and contrast enhancement, because a seam surface is utilized only to determine whether the value of the seam surface at a pixel is higher than the luminance value of the image, and many different seam surfaces can give the same result, in contrast to video retargeting. Furthermore, the density of the cost volume is different. Because the cost volume in tone mapping is very sparse, it is easy to find a seam surface whose energy is sufficiently low. However, because the cost volume in video retargeting is dense, the optimality of seam surfaces is more important.

REFERENCES

- [1] M. Rubinstein, A. Shamir, and S. Avidan, "Improved seam carving for video retargeting," *ACM Trans. Graph.*, vol. 27, no. 3, 2008, Art. no. 16.
- [2] B. Chen and P. Sen, "Video carving," in *Proc. Eurographics*, Apr. 2008, pp. 63–66.
- [3] Z. Li, P. Ishwar, and J. Konrad, "Video condensation by ribbon carving," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2572–2583, Nov. 2009.
- [4] I. Tsubaki and K. Iwauchi, "Tone mapping based on luminance compression using seam surfaces," *IEICE Trans. Inf. Syst.*, vol. 98-D, no. 4, pp. 606–615, 2015.
- [5] S. Avidan and A. Shamir, "Seam carving for content-aware image resizing," *ACM Trans. Graph.*, vol. 26, no. 3, 2007, Art. no. 10.
- [6] M. Grundmann, V. Kwatra, M. Han, and I. Essa, "Discontinuous seam-carving for video retargeting," in *Proc. IEEE CVPR*, Jun. 2010, pp. 569–576.
- [7] W.-L. Chao, H.-H. Su, S.-Y. Chien, W. Hsu, and J.-J. Ding, "Coarse-to-fine temporal optimization for video retargeting based on seam carving," in *Proc. IEEE ICME*, Jul. 2011, pp. 1–6.
- [8] B. Yan, K. Sun, and L. Liu, "Matching-area-based seam carving for video retargeting," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 2, pp. 302–310, Feb. 2013.
- [9] R. Furuta, I. Tsubaki, and T. Yamasaki, "Fast volume seam carving with multi-pass dynamic programming," in *Proc. IEEE ICIP*, Sep. 2016, pp. 1818–1822.
- [10] R. Szeliski *et al.*, "A comparative study of energy minimization methods for Markov random fields with smoothness-based priors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 6, pp. 1068–1080, Jun. 2008.
- [11] S. Birchfield and C. Tomasi, "Depth discontinuities by pixel-to-pixel stereo," *Int. J. Comput. Vis.*, vol. 35, no. 3, pp. 269–293, 1999.
- [12] J. C. Kim, K. M. Lee, B. T. Choi, and S. U. Lee, "A dense stereo matching using two-pass dynamic programming with generalized ground control points," in *Proc. IEEE CVPR*, Jun. 2005, pp. 1075–1082.
- [13] N. Fukushima, T. Yendo, T. Fujii, and M. Tanimoto, "Free viewpoint image generation using multi-pass dynamic programming," *Proc. SPIE*, vol. 6490, pp. 460–470, Mar. 2007.
- [14] H. Hirschmuller, "Accurate and efficient stereo processing by semi-global matching and mutual information," in *Proc. IEEE CVPR*, Jun. 2005, pp. 807–814.
- [15] N. Fukushima, T. Fujii, Y. Ishibashi, T. Yendo, and M. Tanimoto, "Real-time free viewpoint image rendering by using fast multi-pass dynamic programming," in *Proc. IEEE 3DTV-Conf.*, Jun. 2010, pp. 1–4.
- [16] Q. Chen and V. Koltun, "Fast MRF optimization with application to depth reconstruction," in *Proc. IEEE CVPR*, Jun. 2014, pp. 3914–3921.
- [17] F. Liu and M. Gleicher, "Video retargeting: Automating pan and scan," in *Proc. ACM Multimedia*, 2006, pp. 241–250.
- [18] L. Wolf, M. Guttmann, and D. Cohen-Or, "Non-homogeneous content-driven video-retargeting," in *Proc. IEEE ICCV*, Oct. 2007, pp. 1–6.
- [19] B. Li, L. Duan, J. Wang, R. Ji, C. Lin, and W. Gao, "Spatiotemporal grid flow for video retargeting," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1615–1628, Apr. 2014.
- [20] Y.-S. Wang, J.-H. Hsiao, O. Sorkine, and T.-Y. Lee, "Scalable and coherent video resizing with per-frame optimization," *ACM Trans. Graph.*, vol. 30, no. 4, 2011, Art. no. 88.
- [21] Y.-S. Wang, H.-C. Lin, O. Sorkine, and T.-Y. Lee, "Motion-based video retargeting with optimized crop-and-warp," *ACM Trans. Graph.*, vol. 29, no. 4, 2010, Art. no. 90.
- [22] Y. Niu, F. Liu, X. Li, and M. Gleicher, "Warp propagation for video resizing," in *Proc. IEEE CVPR*, Jun. 2010, pp. 537–544.
- [23] Y.-S. Wang, H. Fu, O. Sorkine, T.-Y. Lee, and H.-P. Seidel, "Motion-aware temporal coherence for video resizing," *ACM Trans. Graph.*, vol. 28, no. 5, 2009, Art. no. 127.
- [24] D. Han, X. Wu, and M. Sonka, "Optimal multiple surfaces searching for video/image resizing—a graph-theoretic approach," in *Proc. IEEE ICCV*, Sep. 2009, pp. 1026–1033.
- [25] E. Jain, Y. Sheikh, A. Shamir, and J. Hodgins, "Gaze-driven video re-editing," *ACM Trans. Graph.*, vol. 34, no. 2, 2015, Art. no. 21.
- [26] H. Katti, A. K. Rajagopal, M. Kankanhalli, and R. Kalpathi, "Online estimation of evolving human visual interest," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 11, no. 1, 2014, Art. no. 8.
- [27] G. W. Larson, H. Rushmeier, and C. Piatko, "A visibility matching tone reproduction operator for high dynamic range scenes," *IEEE Trans. Vis. Comput. Graphics*, vol. 3, no. 4, pp. 291–306, Oct./Dec. 1997.
- [28] R. Fattal, D. Lischinski, and M. Werman, "Gradient domain high dynamic range compression," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 249–256, Jul. 2002.
- [29] S. Paris, S. W. Hasinoff, and J. Kautz, "Local Laplacian filters: Edge-aware image processing with a Laplacian pyramid," *ACM Trans. Graph.*, vol. 30, no. 4, 2011, Art. no. 68.
- [30] M. Aubry, S. Paris, S. Hasinoff, J. Kautz, and F. Durand, "Fast local Laplacian filters: Theory and applications," *ACM Trans. Graph.*, vol. 33, no. 5, 2014, Art. no. 167.
- [31] B. Gu, W. Li, M. Zhu, and M. Wang, "Local edge-preserving multiscale decomposition for high dynamic range image tone mapping," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 70–79, Jan. 2013.
- [32] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 6, pp. 1397–1409, Jun. 2013.
- [33] K. Ma, H. Yeganeh, K. Zeng, and Z. Wang, "High dynamic range image compression by optimizing tone mapped image quality index," *IEEE Trans. Image Process.*, vol. 24, no. 10, pp. 3086–3097, Oct. 2015.
- [34] T. Arici, S. Dikbas, and Y. Altunbasak, "A histogram modification framework and its application for image contrast enhancement," *IEEE Trans. Image Process.*, vol. 18, no. 9, pp. 1921–1935, Sep. 2009.
- [35] D. J. Jobson, Z.-U. Rahman, and G. A. Woodell, "Properties and performance of a center/surround Retinex," *IEEE Trans. Image Process.*, vol. 6, no. 3, pp. 451–462, Mar. 1997.
- [36] M. Bertalmio, V. Caselles, and E. Provenzi, "Issues about Retinex theory and contrast enhancement," *Int. J. Comput. Vis.*, vol. 83, no. 1, pp. 101–119, 2009.
- [37] S.-H. Yun, J. H. Kim, and S. Kim, "Image enhancement using a fusion framework of histogram equalization and Laplacian pyramid," *IEEE Trans. Consum. Electron.*, vol. 56, no. 4, pp. 2763–2771, Nov. 2010.
- [38] R. Fattal, "Dehazing using color-lines," *ACM Trans. Graph.*, vol. 34, no. 1, 2014, Art. no. 13.
- [39] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient belief propagation for early vision," *Int. J. Comput. Vis.*, vol. 70, no. 1, pp. 41–54, Oct. 2006.
- [40] M. Rubinstein, D. Gutierrez, O. Sorkine, and A. Shamir, "A comparative study of image retargeting," *ACM Trans. Graph.*, vol. 29, no. 6, 2010, Art. no. 160.
- [41] I. Setyawan and R. L. Lagendijk, "Human perception of geometric distortions in images," *Proc. SPIE*, vol. 5306, p. 256–267, Jun. 2004.
- [42] E. S. Pearson and H. O. Hartley, *Biometrika Tables for Statisticians*, vol. 1. Cambridge, U.K.: Cambridge Univ. Press, 1966.
- [43] F. Drago, K. Myszkowski, T. Annen, and N. Chiba, "Adaptive logarithmic mapping for displaying high contrast scenes," *Comput. Graph. Forum*, vol. 22, no. 3, pp. 419–426, 2003.
- [44] M. D. Fairchild. (2008). *The HDR Photographic Survey*, MDF Publications. [Online]. Available: <http://www.rit-mcsl.org/fairchild/HDRPS/HDRthumbs.html>
- [45] H. Yeganeh and Z. Wang, "Objective quality assessment of tone-mapped images," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 657–667, Feb. 2013.
- [46] R. Mantiuk, K. J. Kim, A. G. Rempel, and W. Heidrich, "HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions," *ACM Trans. Graph.*, vol. 30, no. 4, 2011, Art. no. 40.
- [47] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski, "Edge-preserving decompositions for multi-scale tone and detail manipulation," *ACM Trans. Graph.*, vol. 27, no. 3, 2008, Art. no. 67.



Ryosuke Furuta (S'14) received the B.S. and M.S. degrees in information and communication engineering from The University of Tokyo, Tokyo, Japan, in 2014 and 2016, respectively, where he is currently pursuing the Ph.D. degree. His current research interests include computer vision, machine learning, and image processing, especially MRF optimization.

Mr. Furuta is a member of ACM and ITE.



Ikuko Tsubaki (M'04) received the B.S. degree in physics from the Tokyo Institute of Technology, Tokyo, Japan, in 1995 and the Ph.D. degree in informatics from The University of Tokyo, Tokyo, in 2004.

She was with Sharp Corporation, Chiba, Japan, from 2007 to 2016. Since 2016, she has been an Associate Professor with Tokyo University of Technology, Tokyo. Her research interests include image processing and computer vision.



Toshihiko Yamasaki (M'–) received the B.S. degree in electronic engineering, the M.S. degree in information and communication engineering, and the Ph.D. degree from The University of Tokyo, Tokyo, Japan, in 1999, 2001, and 2004, respectively.

From 2004 to 2006, he was an Assistant Professor with the Department of Frontier Informatics, Graduate School of Frontier Sciences, The University of Tokyo, where he is currently an Associate Professor at the Department of Information and Communication Engineering, Graduate School of Information Science and Technology. He was a JSPS Fellow for research abroad and a Visiting Scientist at Cornell University, Ithaca, NY, USA, from 2011 to 2013. His research interests include attractiveness computing based on multimedia big data analysis, pattern recognition, and machine learning. His publications include three book chapters, more than 60 journal papers, more than 160 international conference papers, and more than 500 domestic conference papers.

Dr. Yamasaki is a member of ACM, IEICE, ITE, and IPSJ. He has received around 50 awards.