Supplemental Material: Fast Volume Seam Carving with Multi-pass Dynamic Programming

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

I. Additional results

A. Video retargeting

Fig. 1 shows the results of video retargeting. The resolution (width, height, number of frames) of the original videos 6 - 8 is (136, 144, 150), (260, 240, 83), and (232, 176, 227), respectively. Videos 6 and 7 are the scenes in which foreground objects are moving and the background is stable. Noticeable deterioration is not created by continuous DP and graph cuts. Foreground object in video 6 and the background in video 7 are shaking between frames in the results of discontinuous DP. Video 8 is a scene in which both foreground and background are moving. Noticeable deterioration is not created by discontinuous DP and graph cuts. However, some players are distorted in the results of continuous DP and graph cuts.

The computation time and maximum memory consumption are shown in Table II(a) and (b), respectively. By replacing graph cuts with multi-pass DP, the computation time is reduced to 0.6%. The maximum memory consumption is reduced to 2.8% by replacing graph cuts with continuous DP, and 5.4% with discontinuous DP.

Table I(a) shows the results of the subjective evaluation tests of the three videos of Fig. 1. We observed that, on average, approximately half of the 163 subjects answered that there was no noticeable difference between the retargeted videos.

Table I(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 that were not fooled by the dummy questions. We observed that, on average, 67.1% of these 72 subjects answered that there were no noticeable differences between the retargeted videos.

TABLE I
NUMBER OF SUBJECTS WHO PREFERRED THE RETARGETED VIDEO OF EACH METHOD.

(a) All subjects

	Multi-pass DP (continuous /discontinuous)	Graph cuts	Cannot notice the difference	Total
Video 6	23 / 12	34	**94**	163
Video 7	22 / 11	64	**66**	163
Video 8	7 / 18	68	**70**	163
Average	17.3 / 13.7	55.3	**76.7**	163

(b) Subjects who were not fooled by the dummy questions

	Multi-pass DP (continuous /discontinuous)	Graph cuts	Cannot notice the difference	Total
Video 6	6 / 0	10	**56**	72
Video 7	4 / 2	21	**45**	72
Video 8	0 / 5	23	**44**	72
Average	3.3 / 2.3	18	**48.3**	72

B. Tone mapping

Fig. 3 shows the results of tone mapping. The resolution (width, height) of images 7 - 10 is (750, 1000), (1000, 664), (803, 535), and (401, 535), respectively. Fig. 3(a) and (b) look similar, but (a) is a bit brighter than (b). On the whole, the detail is clear in (c) and (d); however, the contrast in (a) and (b) is higher than in (c) and (d).

Fig. 2 shows the results when the parameter p is changed. p is a parameter which controls the subsampling rate of the cost volume. As p becomes larger, the subsampled cost volume becomes smaller. Bright region expands in the results with $p = 65$ compared to those with $p = 3$. The image quality depends on p.

The computation time and maximum memory consumption are shown in Table IV(a) and (b), respectively. By replacing graph cuts with continuous DP, the computation time is reduced to 1.2%, and the maximum memory consumption is reduced to 56%.

Table III(a) shows the results of the subjective evaluation tests of the four images of Fig. 3. We observed that, on average, 38.0% of 198 subjects noticed no difference between the two tone-mapped images.

Table III(b) shows the results of the 101 subjects who were not fooled by the three dummy questions. We observed that, on average, 57.7% of the 101 subjects noticed no difference between the two tone-mapped images.

C. Contrast enhancement

Fig. 4 shows the results of contrast enhancement. The resolution (width, height) of images 7 - 10 is (512, 512), (800, 530), (459, 700), and (400, 300), respectively. 8 bit luminance value was reduced to 7 bit, and then reverted to 8 bit for Fig. 4 (b) and (c). Fig. 4(b) and (c) look similar, but (b) is brighter than (c).

The computation time and maximum memory consumption are shown in Table V(a) and (b), respectively. By replacing graph cuts with continuous DP, the computation time is reduced to 0.4%, and the maximum memory consumption is reduced to 39%.

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, Bunkyo-ku 113-8656, Tokyo, 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, Hachioji-shi 192-0982, Tokyo, Japan e-mail: (i.tsuba@gmail.com).

Manuscript received **** **, 2016; revised **** **, 2016.

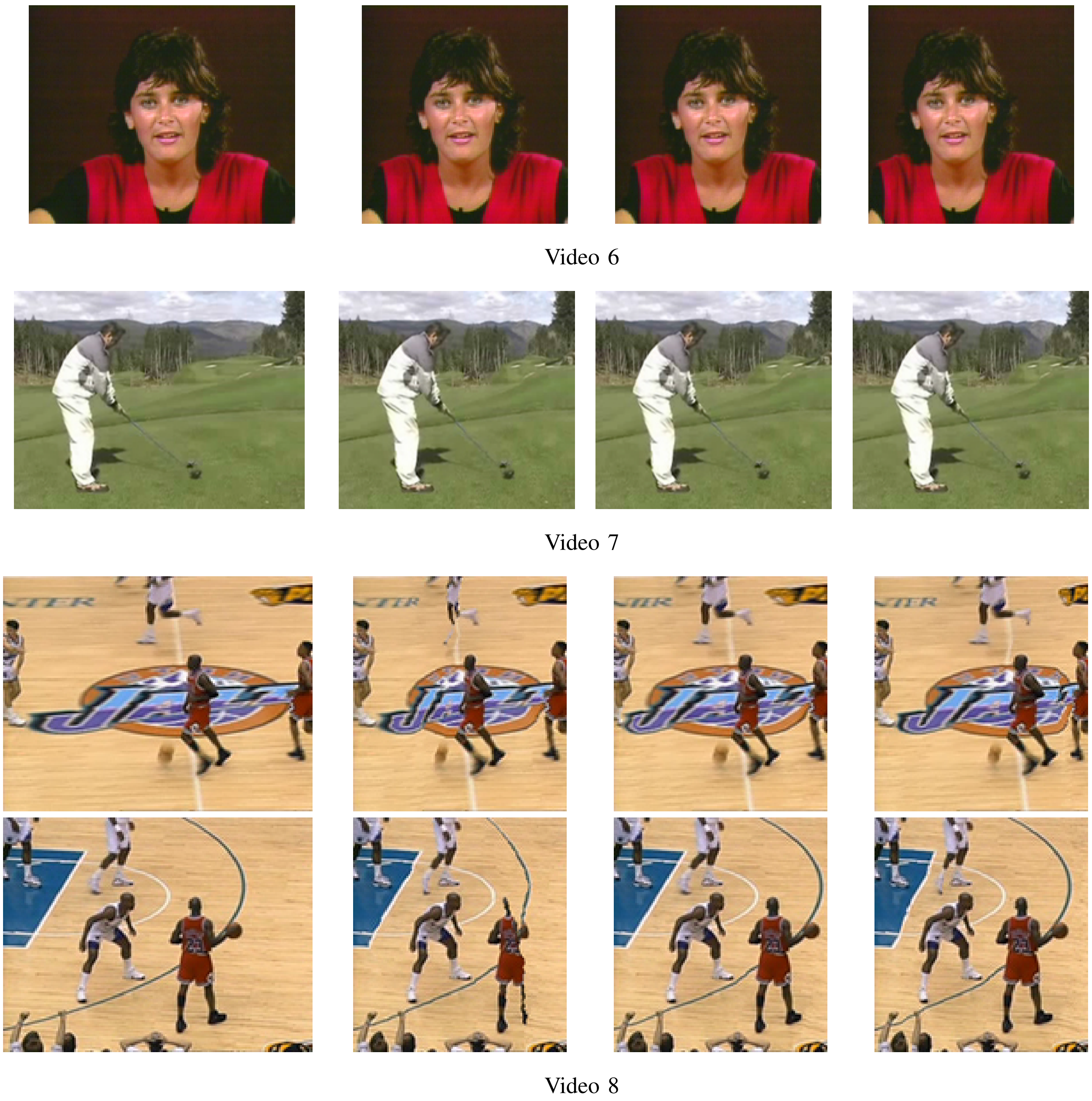
TABLE II
COMPUTATION TIME AND MEMORY CONSUMPTION FOR FIG. 1.

(a) Computation time [s]

	Continuous DP	Discontinuous DP	Graph cuts
Video 6	20	20	683
Video 7	49	51	1594
Video 8	83	84	25,083
Average	51	52	9120

(b) Maximum memory consumption [MB]

	Continuous DP	Discontinuous DP	Graph cuts
Video 6	17	32	373
Video 7	29	54	809
Video 8	39	75	1.7×10^3
Average	28	54	1.0×10^3



(a) Original (b) Multi-pass DP (continuous) (c) Multi-pass DP (discontinuous) (d) Graph cuts

Fig. 1. Results of video retargeting: Upper and lower raws are different frames in video 8.

TABLE IV
COMPUTATION TIME AND MEMORY CONSUMPTION FOR FIG. 3.

(a) Computation time [s]

	Mutli-pass DP	Graph cuts
Image 7	1.05	142.36
Image 8	0.91	34.11
Image 9	0.58	54.46
Image 10	0.28	11.74
Average	0.71	60.67

(b) Maximum memory consumption [MB]

	Mutli-pass DP	Graph cuts
Image 7	45	81
Image 8	40	71
Image 9	26	47
Image 10	14	24
Average	31	56

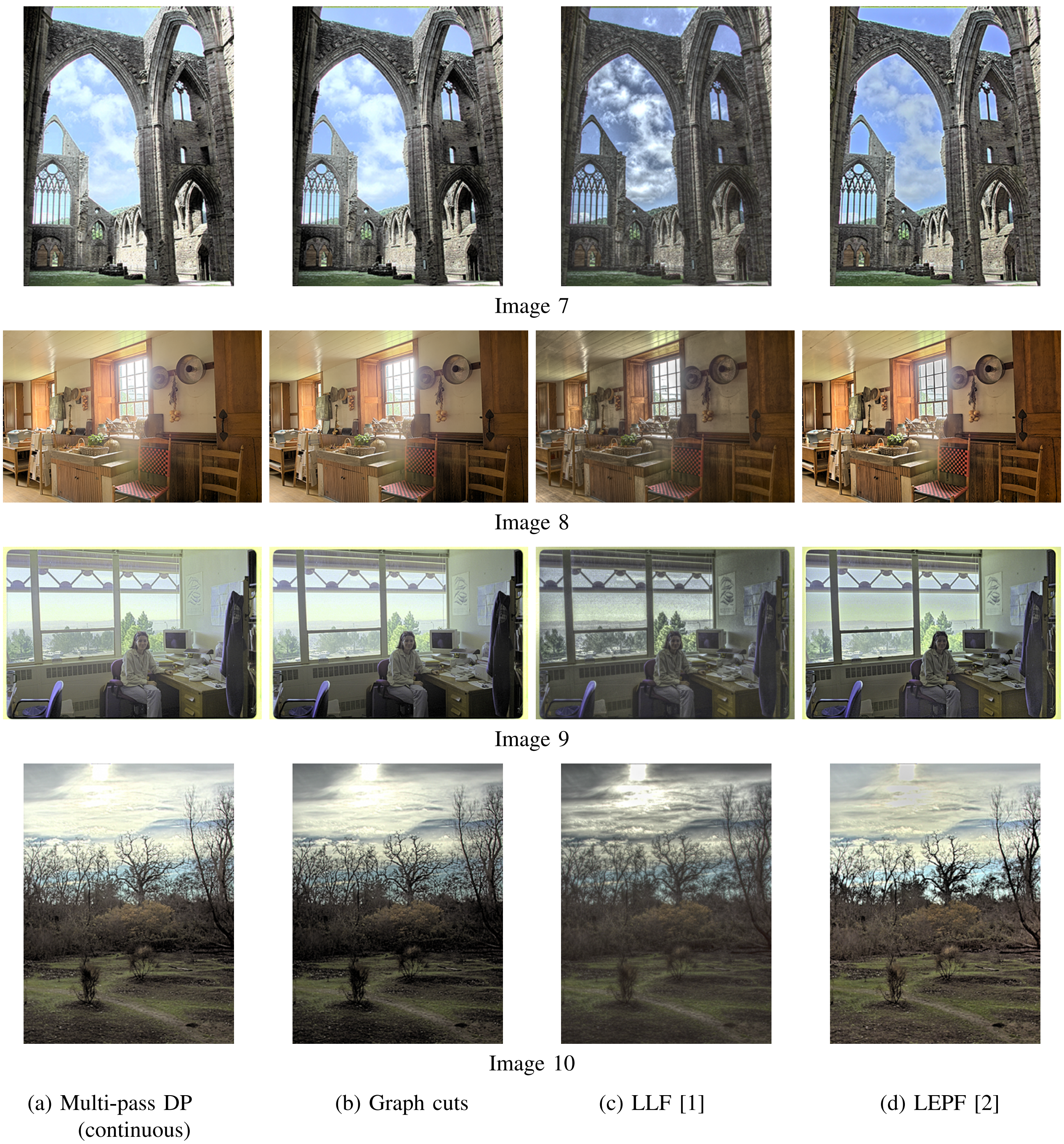


Fig. 3. Tone mapping results.

$p = 3$

$p = 65$

$p = 3$

$p = 65$

Fig. 2. Tone mapping results. p is a spatial subsampling parameter.

TABLE III
NUMBER OF SUBJECTS WHO PREFERRED THE TONE-MAPPED IMAGE OF EACH METHOD.

(a) All subjects

	Multi-pass DP (continuous)	Graph cuts	Cannot notice the difference	Total
Image 7	28	**91**	79	198
Image 8	33	52	**113**	198
Image 9	40	**107**	51	198
Image 10	**85**	55	58	198
Average	46.5	**76.3**	75.3	198

(b) Subjects who were not fooled by the dummy questions

	Multi-pass DP (continuous)	Graph cuts	Cannot notice the difference	Total
Image 7	10	31	**60**	101
Image 8	9	14	**78**	101
Image 9	16	39	**46**	101
Image 10	30	22	**49**	101
Average	16.3	26.5	**58.3**	101

Table VI(a) shows the results of the subjective evaluation tests of the four images of Fig. 4. We observed that, on average, 39.0% of 191 subjects noticed 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, 47.2% of the 90 subjects noticed no difference between the two tone-mapped images.

TABLE VI
NUMBER OF SUBJECTS WHO PREFERRED THE ENHANCED IMAGE OF EACH METHOD.

(a) All subjects

	Multi-pass DP (continuous)	Graph cuts	Cannot notice the difference	Total
Image 7	**93**	43	55	191
Image 8	41	**138**	12	191
Image 9	45	43	**103**	191
Image 10	36	27	**128**	191
Average	53.8	62.8	**74.5**	191

(b) Subjects who were not fooled by the dummy questions

	Multi-pass DP (continuous)	Graph cuts	Cannot notice the difference	Total
Image 7	**41**	18	31	90
Image 8	18	**64**	8	90
Image 9	18	11	**61**	90
Image 10	8	12	**70**	90
Average	21.3	26.3	**42.5**	90

TABLE V
COMPUTATION TIME AND MEMORY CONSUMPTION FOR FIG. 4.

(a) Computation time [s]

	Mutli-pass DP	Graph cuts
Image 7	0.45	175.21
Image 8	0.85	204.48
Image 9	0.29	15.52
Image 10	0.10	0.37
Average	0.42	98.90

(b) Maximum memory consumption [MB]

	Mutli-pass DP	Graph cuts
Image 7	17	45
Image 8	27	78
Image 9	21	43
Image 10	6	17
Average	18	46

Image 7

Image 8

Image 9

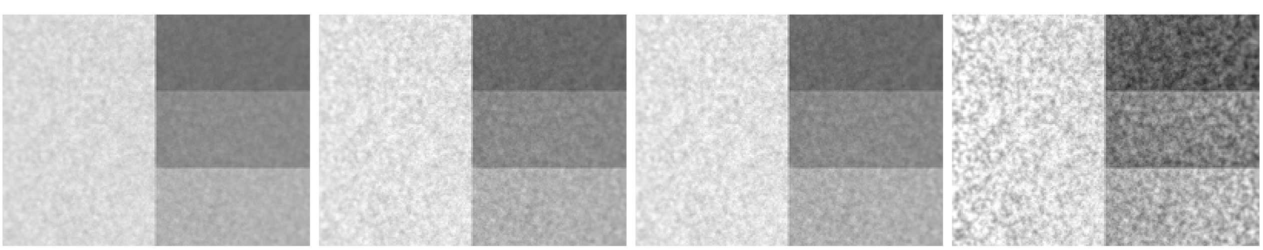


Image 10

(a) Original (b) Multi-pass DP (continuous) (c) Graph cuts (d) LLF [1]

Fig. 4. Contrast enhancement results.

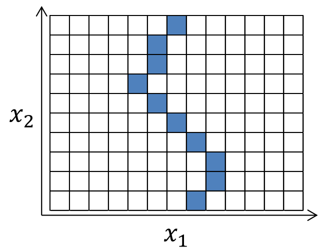


Fig. 5. Connected pass by DP in 2D plane.

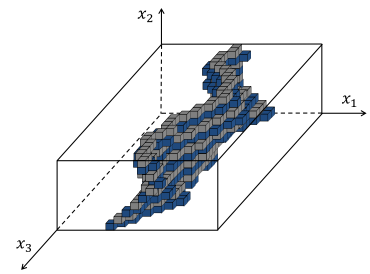


Fig. 6. Connected seam surface in 3D space.

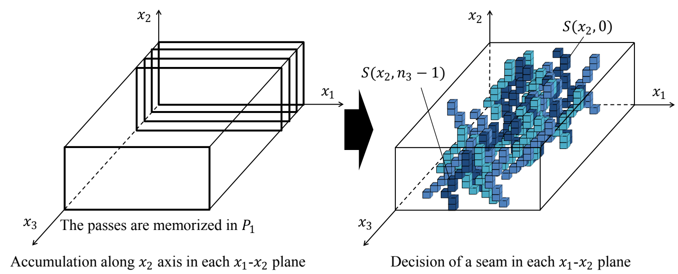


Fig. 7. DP in each $x_1 - x_2$ plane independently. Although the obtained seam is connected in the x_2 direction, not connected in the x_3 direction.

II. Derivation of the proposed multi-pass DP

In this section, we describe the derivation of the proposed multi-pass DP.

A. DP in 2D plane

First, we consider the process of dynamic programming (DP) to obtain an optimal path in a two dimensional plane as a simple example. As shown in Fig. 5, the objective is to obtain a path S that crosses the x_1 axis in the $x_1 - x_2$ plane $(0 \le x_1 < n_1,\ \ 0 \le x_2 < n_2)$. When the path passes coordinate (x_1, x_2), we describe it as $x_1 = S(x_2)$. We call the path "seam" in seam carving problems, and the seam must be connected in the x_2 direction. The cost $C(x_1, x_2)$ is assigned to each coordinate (x_1, x_2), and our objective is to obtain a seam such that the sum of the cost is minimum. This optimization problem is described as

$$\arg\min_{S} \sum_{x_2} C(S(x_2), x_2), \tag{1}$$

$$s.t.\ |S(x_2) - S(x_2+1)| \le 1. \tag{2}$$

The accumulation process of DP along the x_2 axis is described as

$$\begin{cases} A(x_1{,}0) = C(x_1{,}0), \\ A(x_1{,}x_2) = C(x_1{,}x_2) + \\ \qquad \min\limits_{j\in\{-1,0,1\}} A(x_1{+}j, x_2{-}1), \quad (x_2 > 0) \\ P(x_1{,}x_2) = \arg\min\limits_{j\in\{-1,0,1\}} A(x_1{+}j, x_2-1){+}x_1,\ (x_2 > 0) \end{cases} \tag{3}$$

where $A(x_1, x_2)$ is the accumulated cost function, and the value of j selected during the accumulation is recorded in paths P. The optimal path is obtained by tracking back P from $x_2 = n_2 - 1$ to 0.

$$\begin{cases} S(n_2-1) = \arg\min\limits_{x_1} A(x_1{,}n_2-1), \\ S(x_2) = P(S(x_2{+}1), x_2{+}1), \quad (x_2 < n_2{-}1) \end{cases} \tag{4}$$

Because j is selected from among $\{-1, 0, 1\}$, the obtained path $S(x_2)$ is guaranteed to be connected in the x_2 direction, in other words, Eq. (2) is satisfied.

The optimal path is obtained by the process described above. However, Fukushima et al. [3] stated that a suboptimal solution that is almost equal to the optimal one can be obtained by selecting the x_1 that has the minimum accumulated cost $A(x_1, x_2)$ at each x_2 without recording paths for tracking back (Fig. 7 in [3]). That process is described as

$$S(x_2) = \arg\min_{x_1} A(x_1, x_2). \tag{5}$$

However, the path obtained by Eq. (5) is not guaranteed to be connected in the x_2 direction. This method is reasonable for disparity estimation, in which (x_1, x_2) corresponds to $(disparity, xpixel)$, and the smoothness of the path (disparity) is required, but the connectivity is not. Therefore, for seam carving problems in which the connectivity is required, we vary Eq. (5) as

$$\begin{cases} S(n_2-1) = \arg\min\limits_{x_1} A(x_1{,}n_2-1), \\ S(x_2) = \arg\min\limits_{x_1\in\{S(x_2+1),S(x_2+1)\pm1\}} A(x_1, x_2). \quad (x_2 < n_2{-}1) \end{cases} \tag{6}$$

First, the x_1 which has the minimum accumulated cost $A(x_1, x_2)$ is selected at $x_2 = n_2 - 1$. Subsequently, x_2 is reduced by one, and the range of x_1 is restricted among $\{S(x_2+1) - 1, S(x_2+1), S(x_2+1) + 1\}$ where $S(x_2+1)$ is the previously selected x_1. In other words, the candidate of the next x_1 is restricted within the range of ± 1 of the previous result. The final seam that is connected in the x_2 direction is obtained by repeating the process from $x_2 = n_2 - 2$ to 0.

The seam obtained by Eq. (6) is equal to the one obtained by Eq. (7).

$$\begin{cases} S(x_2) = \arg\min\limits_{x_1} A(x_1{,}x_2), \\ A(x_1{,}x_2) = \begin{cases} A(x_1{,}x_2), & (|x_1{-}S(x_2{+}1)| \le 1) \\ \infty, & (otherwise). \end{cases} \end{cases} \tag{7}$$

When x_2 is reduced by one, the costs that are out of the range of ± 1 of the previous result is set to infinity. This setting makes the obtained seam satisfy Eq. (2).

B. DP in 3D volume

In this section, we consider the problem to obtain a seam surface that is connected in three dimensional space as shown in Fig. 6. When the seam surface passes coordinate (x_1, x_2, x_3), we describe it as $x_1 = S(x_2, x_3)$. This optimization problem is described as

$$\arg\min_{S} \sum_{x_2,x_3} C(S(x_2,x_3),x_2,x_3), \tag{8}$$

$$s.t.\ |S(x_2,x_3) - S(x_2+1,x_3)| \leq 1, \tag{9}$$

$$|S(x_2,x_3) - S(x_2,x_3+1)| \leq 1. \tag{10}$$

The globally optimal solution of this problem cannot be obtained by DP as Rubinstein et al.. [4] pointed out. Here, we consider the reason why DP cannot directly be applied to this volume seam carving problem. We first perform the accumulation process of DP along the x_2 axis in each $x_1 - x_2$ plane independently as Eq. (11).

$$\begin{cases} A_1(x_1,0,x_3) = C(x_1,0,x_3), \\ A_1(x_1,x_2,x_3) = C(x_1,x_2,x_3) + \\ \qquad \min\limits_{j\in\{-1,0,1\}} A_1(x_1+j, x_2-1, x_3). \quad (x_2>0) \\ P_1(x_1,x_2,x_3) = \arg\min\limits_{j\in\{-1,0,1\}} A_1(x_1+j, x_2-1, x_3)+x_1, \quad (x_2>0). \end{cases} \tag{11}$$

We obtain the seam surface by tracking back as Eq. (12).

$$\begin{cases} S(n_2-1, x_3) = \arg\min\limits_{x_1} A_1(x_1,n_2-1,x_3) \\ S(x_2,x_3) = P_1(S(x_2+1,x_3), x_2+1, x_3), \qquad (x_2<n_2-1) \end{cases} \tag{12}$$

The obtained seam surface is connected in the x_2 direction because j is selected from among $\{-1, 0, 1\}$. However, it is not connected in the x_3 direction as shown in Fig. 7 because the DP is performed in each $x_1 - x_2$ plane independently. In other words, although Eq. (9) is satisfied, Eq. (10) is not.

C. Multi-pass DP in 3D volume

1) Continuous method: The continuous method of the proposed multi-pass DP can obtain a suboptimal solution that is connected in both the x_2 and x_3 directions in 3D space as shown in Fig. 6. The flow chart is shown in Fig. 8. We derivate our continuous method by extending the suboptimal solution of DP in 2D plane to 3D space.

Definition

We define a vector whose elements are $x_1 = S(x_2, x_3)$ $(x_3 = 0, \cdots, n_3 - 1)$ as $\boldsymbol{X}_1$ when x_2 is fixed.

$$\boldsymbol{X}_1 := \boldsymbol{S}(x_2) = [S(x_2, 0), \cdots, S(x_2, n_3 - 1)]. \tag{13}$$

A $\boldsymbol{X}_1$ corresponds to a gray (or blue) seam in Fig. 6. We denote by $C(\boldsymbol{X}_1, x_2)$ the cost of the seam $\boldsymbol{X}_1$, in other words, $C(\boldsymbol{X}_1, x_2)$ is the sum of the costs at the coordinates where the seam $\boldsymbol{X}_1$ passes.

$$C(\boldsymbol{X}_1, x_2) := \sum_{x_3} C(S(x_2,x_3), x_2, x_3). \tag{14}$$

Extension of the 2D suboptimal to 3D

First, we independently accumulate the costs along the x_2 axis in each $x_1 - x_2$ plane, similarly to Eq. (11). This process is the step 1 in our main paper.

If simply tracking back P_1, the obtained seam surface is not guaranteed to be connected in the x_3 direction as described in Section II-B. As Rubinstein et al. [4] pointed out, we cannot obtain an optimal solution by DP in volume seam carving problems. In other words, we cannot extend the optimal solution obtained by Eq. (4) to 3D space. Therefore, we instead extend the suboptimal solution in Eq. (5) to 3D space in Eq. (16).

$$\boldsymbol{S}(x_2) = \arg\min_{\boldsymbol{X}_1} \sum_{x_3} A_1(S(x_2,x_3),x_2,x_3) \tag{15}$$

$$= \arg\min_{\boldsymbol{X}_1} A_1(\boldsymbol{X}_1, x_2). \tag{16}$$

In Eq. (5), we obtain a suboptimal seam by selecting the x_1 whose cost is the minimum at each x_2. Similarly, in Eq. (16), we obtain a suboptimal seam surface by selecting the seam $\boldsymbol{X}_1$ whose cost is minimum in each $x_1 - x_3$ plane. The solution $\boldsymbol{X}_1 = [S(x_2, 0), \cdots, S(x_2, n_3 - 1)]$ in Eq. (16) can be obtained by performing 2D DP in the $x_1 - x_3$ plane.

$$\begin{cases} A_2(x_1,x_2,0) = A_1(x_1,x_2,0). \\ A_2(x_1,x_2,x_3) = A_1(x_1,x_2,x_3) + \\ \qquad \min\limits_{j\in\{-1,0,1\}} A_2(x_1+j, x_2, x_3-1), \quad (x_3>0) \\ P_2(x_1,x_2,x_3) = \arg\min\limits_{j\in\{-1,0,1\}} A_2(x_1+j, x_2, x_3-1)+x_1, \ (x_3>0) \\ S(x_2,n_3-1) = \arg\min\limits_{x_1} A_2(x_1,x_2,n_3-1) \\ S(x_2,x_3) = P_2(S(x_2,x_3+1), x_2, x_3+1), \quad (x_3<n_3-1). \end{cases} \tag{17}$$

This is the step 2 in our main paper. Because j is selected from among $\{-1, 0, 1\}$, the obtained seam $\boldsymbol{X}_1$ is guaranteed to be connected in the x_3 direction, in other words, it satisfies Eq. (10).

However, the obtained seam $\boldsymbol{X}_1$ is not guaranteed to be connected in the x_2 direction ($\boldsymbol{S}(x_2)$ and $\boldsymbol{S}(x_2 + 1)$ are not connected), in other words, it does not satisfy Eq. (9). Therefore, similarly to Eq. (6), we restrict the range of candidate seam $\boldsymbol{X}_1$ within ± 1 of the previous seam $\boldsymbol{S}(x_2 + 1)$ when reducing x_2 by one. Subsequently, we obtain the optimal seam in Eq. (16). Similarly to Eq. (7), this process is redescribed as

$$A_2(x_1,x_2,x_3) = \begin{cases} A_2(x_1,x_2,x_3), & (|x_1-S(x_2+1,x_3)| \leq 1) \\ \infty. & (otherwise) \end{cases} \tag{18}$$

This is the step 3 in the continuous method in our main paper. From $x_2 = n_2 - 1$ to 0, by repeating the acquisition of seam in Eq. (17) and the restriction of the range in Eq. (18) alternately, we can obtain a suboptimal seam surface in both the x_2 and x_3 directions.

D. Discontinuous method

As shown in Fig. 9, the discontinuous method can obtain a seam surface that is connected in the $x_1 - x_3$ plane at $x_2 = n_2 - 1$ by simply tracking back P_1 after first step 2 in Eq. (17). This connectivity has the effect of making the seam surface prone to connecting in other $x_1 - x_3$ planes, but not guaranteed to be connected in other $x_1 - x_3$ planes. It is important because disconnectivity is sometimes required for seam carving depending on the applications (*e.g.*, retargeting for videos with extreme motion) as described in our main paper.

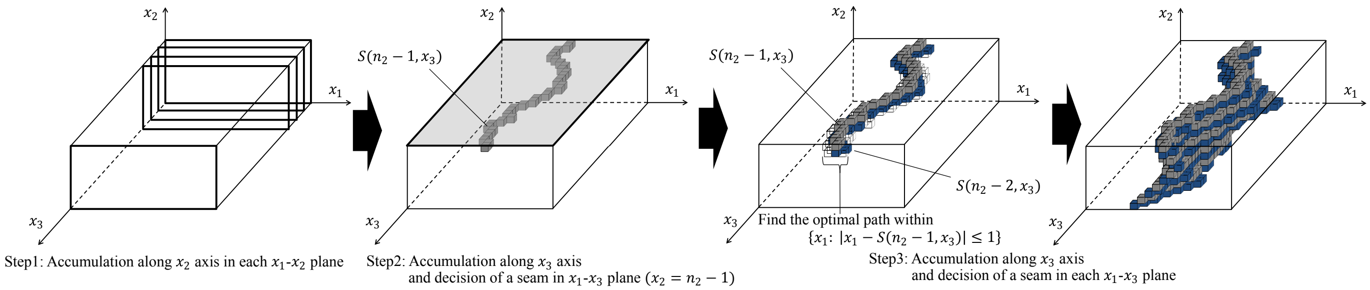


Fig. 8. The continuous DP process. The seam surface is connected in both the x_2 and the x_3 directions.

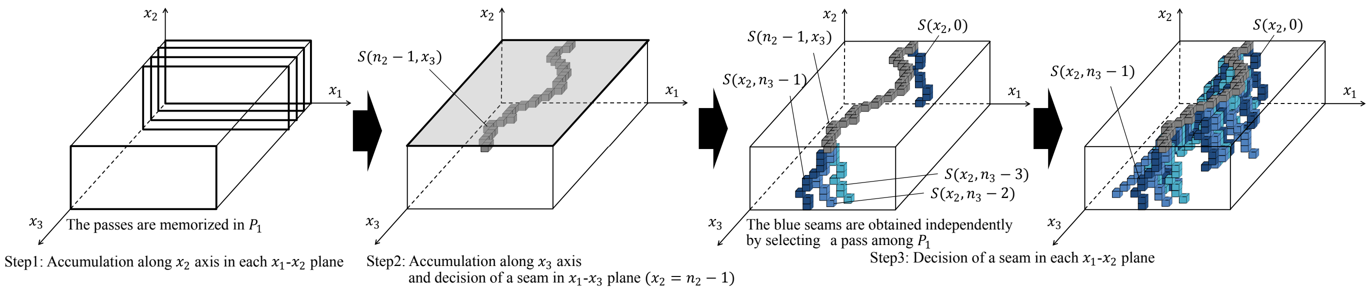


Fig. 9. The discontinuous DP process. The seam obtained as an intersection between the seam surface and the $x_1 - x_3$ plane at $x_2 = n_2 - 1$ is connected in the x_3 direction. The seams in different $x_1 - x_3$ planes are not necessarily connected.

REFERENCES

[1] S. Paris, S. Hasinoff, and J. Kautz, "Local Laplacian filters: edge-aware image processing with a Laplacian pyramid," *ACM Trans. Graph.*, vol. 30, no. 4, p. 68, 2011.

[2] 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, 2013.

[3] N. Fukushima, T. Yendo, T. Fujii, and M. Tanimoto, "Free viewpoint image generation using multi-pass dynamic programming," in *Electronic Imaging 2007*. International Society for Optics and Photonics, 2007, pp. 64 901F–64 901F.

[4] M. Rubinstein, A. Shamir, and S. Avidan, "Improved seam carving for video retargeting," *ACM Trans. Graph.*, vol. 27, no. 3, p. 16, 2008.