PAPER A Novel Content-Aware Stitching Algorithm for Real-Time Video Sequences

Kwang-Wook LEE[†], Seung-Won JUNG[†], Seung-Kyun KIM[†], Nonmembers, and Sung-Jea KO^{†a)}, Member

SUMMARY The panorama image obtained by image stitching can have visible artifacts due to the limitation of alignment accuracy and defects of the optical systems. Moreover, conventional image stitching algorithms cannot be directly applied to a real-time video stitching due to its complexity and waving artifacts. In this paper, we propose a real-time content-aware stitching algorithm which not only finds a seam by using path searching based on the greedy method, but also adaptively updates the seam by detecting objects moving toward the seam. Experimental results show that the proposed algorithm can successfully produce stitched video sequences without the waving and ghost artifacts commonly found in conventional stitching algorithms.

key words: video stitching, path-finding, mosaic, seam update

1. Introduction

Image stitching is used in many applications such as panorama image [1], [2], super resolution [3], and surveillance systems because it can overcome the resolution constraint and the small field-of-view of the camera. In particular, video stitching is a powerful technique to provide the panorama video for visual surveillance systems with high resolution. Such systems are used to monitor wide areas, e.g. long streets and supermarkets, that are not easily covered by distributed single cameras.

In order to obtain satisfactory results in image stitching, both structure and intensity should be aligned or matched within the overlapped region [4]. For aligning images, conventional algorithms usually exploit the scale invariant feature transform (SIFT) [5] and the random sample consensus (RANSAC) [6]. However, these algorithms often produce coarsely aligned images including a visible seam in the overlapped region due to the accuracy limitation of the alignment and unmodeled camera effects such as the vignetting effect, the difference of exposure times, and the white balance effect. To solve this problem, various image composite methods for the overlapped region have been proposed [7]-[12]. Some methods [7]–[9] minimize the intensity distortion in the overlapped region by using a blending algorithm. The others [10]–[12] search for a curve in the overlapped region on which the difference between images is minimal.

In this paper, we present a real-time content-aware video stitching algorithm. One of the main drawbacks

Manuscript received July 20, 2010.

Manuscript revised October 12, 2010.

[†]The authors are with the Department of Electronic Engineering, Korea University, Anam-Dong, Sungbuk-Gu, Seoul 136–701, Korea.

a) E-mail: sjko@korea.ac.kr

DOI: 10.1587/transinf.E94.D.357

of conventional algorithms for real-time video stitching is computational burden. Conventional algorithms produce satisfactory results in terms of visual quality but many of these methods are computationally too expensive to be applied to real-time applications. In addition, they do not deal with video characteristics related to objects moving across the seam. The proposed novel content-aware stitching algorithm can not only quickly find a seam in the overlapped region without visible artifacts by using the greedy method but also update the seam when the object moves across the seam.

The rest of our paper is organized as follows. Section 2 describes a proposed algorithm for a content-aware stitching algorithm. The experimental results are provided in Sect. 3 and conclusions are given in Sect. 4.

2. Proposed Algorithm

The proposed algorithm consists of two steps as shown in Fig. 1. First, for image alignment between the anchor and the covered image, we calculate homography by the SIFT followed by the RANSAC. For composite of the overlapped region without visible artifacts, we propose a computationally efficient content-aware path-finding algorithm. The proposed path-finding algorithm first decides the start and end positions of the path. Then, the path minimizing a new energy function is found based on the greedy method. Furthermore, to deal with the moving objects in video sequences, a seam update method is also presented. The following subsections describe the proposed algorithm in detail.



Fig. 1 Flowchart of the proposed system.

2.1 Start/End Point Decision

Conventional path-finding algorithms [10]–[12] for image stitching divide the overlapped region into two sub-regions as shown in Fig. 2 (a). However, in many cases, images are aligned arbitrarily as shown in Fig. 2 (b), and thus constraining the start/end locations of the seam is necessary to stitch images without visible seam artifacts. If conventional algorithms are applied straightforwardly, the unexpected visible seam can appear in the overlapped image. For example, if the end point of the seam reaches an arbitrary boundary of the image as shown in Fig. 2 (b), the image boundary region surrounded by the dashed circle can be perceived as a visible seam. To solve this problem, crossing points of two images are used for the start/end point of the seam as shown in Fig. 2 (c).

2.2 Fast Partitioning

In order to find a path satisfying the initialized start/end point, many energy minimization techniques, such as Dijkstra algorithm [10], min-cut/max-flow algorithm [11], and Dynamic programming [12], can be considered. However, conventional energy minimization methods in [10]–[12] are inadequate for real-time applications due to their large memory and time complexity requirements. Therefore, we introduce a novel path-searching algorithm based on the greedy method.

We define an energy function at pixel p in the overlapped region as a combination of three terms; the color vector difference $D_C(p)$, the magnitude of the gradient $D_G(p)$,



Fig. 2 Searching the seam. (a) Connecting the path in the aligned images. (b) Connecting the path from the arbitrarily start point to the end point. (c) Connecting the path from the crossing point to the other crossing point.

and the distance $D_L(p)$ from a current pixel to a line connecting the start point to the end point.

$$E(D_C(p), D_G(p), D_L(p))$$

= $(1 - \alpha) \cdot \frac{D_C(p)}{D_G(p)} + \alpha \cdot D_L(p),$ (1)

where $D_C(p)$ represents the Euclidian distance between the color vectors of the anchor image $I_A(p)$ and the covered image $I_C(p)$ at pixel p in the overlapped region,

$$D_C(p) = \|I_A(p) - I_C(p)\|_2.$$
⁽²⁾

In addition to the pixel intensity, the gradient also needs to be considered since the seam should not penetrate the object contour but flow along a high frequency region. Therefore, the normalized gradient value $D_G(p)$ is included in the proposed energy function as follows:

$$D_{G}(p) = \max\left(\frac{\max(|\nabla I_{A}(p)|, |\nabla I_{C}(p)|)}{\max\left(\sum_{q \in \Omega} \frac{1}{N} |\nabla I_{A}(q)|, 1\right)}, 1\right),$$
(3)

where Ω is a set of pixels in the overlapped region, *N* is the number of pixels in Ω , and $|\nabla I_{A(C)}|$ represents the magnitude of the gradient of the anchor(covered) image that can be calculated by the Sobel operator [15].

Lastly, the path obtained by using only aforementioned two elements of the energy function is not guaranteed to reach the predefined end point when the greedy method is utilized. Thus, we include a new element to the energy function, which is a distance factor $D_L(p)$ enforcing the connection of the path from the start point to the end point as follows:

$$D_L(p) = \|P_C(p) - P_M(p)\|_2,$$
(4)

where P_S , P_E , and P_C in Fig. 3 denote the start point, the end point, and the current pixel point, respectively, and P_M is the projection point of P_C onto the line $\overline{P_S P_E}$.

In addition, we utilize a weighting factor α , $0 \le \alpha \le 1$,



Fig. 3 Elements of the cost function.

to control the influence of $D_L(p)$ by

$$\alpha = 1 - \frac{2 \cdot \min\left(\overline{P_S P_M}, \overline{P_M P_E}\right)}{\overline{P_S P_E}},\tag{5}$$

where min $(\overline{P_S P_M}, \overline{P_M P_E})$ represents the minimum of Euclidian distance from the current projection pixel P_M to the end point P_E and to the start point P_S . Thus if P_C is located near the end point, the distance parameter has stronger effect than the other parameters. However, at the middle position of the overlapped region, it rarely affects the decision of the node energy. By including this distance term with weighting factor α , the path is guided into the end point. As a result of this adjustment, the seam can be reached to the predefined end point P_E even though the greedy method is used.

2.3 Search Strategy

In the greedy method, the path is locally determined based on the given grid-mask. Figure 4 shows new grid-masks adopted in our path searching algorithm. The use of these grid-masks enables the path to detour the image edges.

We define the grid-masks of four directions, GM_D , GM_U , GM_R , and GM_L , respectively. In order to obtain the reliable path, the grid-mask is determined locally based on relative positions of the start point and the end point and the gradient orientation calculated by the Sobel operator [15]. For instance, in Fig. 3, grid-masks GM_U and GM_R are chosen according to the relative position of the start and end points since the end point is located in the upper and right of the start point. Then, for each location, if the horizontal edge strength [15] is stronger than that of the vertical edge, GM_U is selected, otherwise GM_R is used. By deciding the direction of the grid-mask according to the local gradient, we can find a more visually pleasant path in the overlapped region.

2.4 Seam Update

In this subsection, we describe an adaptive seam update method for video adaptation. Although we find the seam using the proposed energy function, video contents are not considered in the proposed algorithm yet. As a result, the stitched sequence made by the proposed path-finding algorithm can yield perceivable artifacts around the transition.

If path-finding is performed for each frame independently, each frame can have visually plausible stitching result but the sequence combined frame by frame contains notable visual artifacts such as waving artifacts or flickering effects as shown in Fig. 5 (a). Alternatively, if the static seam is used to avoid the waving artifact, the visual quality can be significantly degraded when a moving object passes through the seam as shown in Fig. 5 (b). In this case, the shape of the moving object is deformed for a short period such that annoying artifacts are perceived. Moreover, the perceivable visual quality of the stitched image is poor since the observer responds more strongly to the moving object than to the background.

In summary, to avoid the waving artifact, the static seam is better but to prevent penetration of the seam to the moving object, the seam needs to be updated. To solve these problems, we propose an adaptive partial seam update method.

The seam is initialized by the proposed technique. Then, when an object moves toward the seam, the seam near the moving object is updated as shown in Fig. 6. Here, to detect and track the moving object, the real-time code book algorithm is used [13]. When the object approaches the initial seam, only the part of seam near the moving object is updated to protect the detected object. Since the seam is only partially updated, the computational complexity can be significantly reduced and the waving artifact can be avoided.



Fig. 4 Grid-masks for selecting the next node.



Fig. 5 Temporally dynamic seam determined frame by frame and static seam (a) Waving artifact occurs when the seam is searched frame by frame. (b) Visible artifact occurs when a fast moving object passes through the temporally connected static seam.



3. Experimental Results

To assess the performance of our algorithm, we compare the proposed algorithm with conventional algorithms proposed by Zomet [4] and Gu [14]. All the algorithms merge two images of size 640×480 captured by two different fixed cameras placed at the top of the building. Although we utilize two cameras for simple experiment, the proposed algorithm can be straightforwardly generalized to stitch more than three images. Figure 7 shows original images and their stitched image.

Figure 8 compares results in the overlapped region using GIST1 [4], Gu's algorithm [14], and the proposed algorithm, respectively. Figures 8 (a) and (b) illustrate the result of Gu's algorithm and GIST1. Compared to the conventional algorithms that produce smooth paths, our proposed algorithm can also find the path along the edge and the result shows a good stitched image as shown in Fig. 8 (c). Although the greedy algorithm is used, our result usually detours the object contour and does not produce the visible seam.

When the car moves through the overlapped region, a part of the seam is updated near the car. Figure 9 illustrates the seam update process. Figure 9 (a) shows the predefined seam before the object approaches. When the object crosses the predefined seam, the seam is updated to avoid penetrating moving objects as shown in Figs. 9 (b) and (c). Figure 9 (d) shows the recovered seam after the object passes. If the contour of the object does not contact the seam anymore, the seam update process is finished.

Note that the proposed algorithm is tested when only one object passes through the seam. Thus, when multiple objects rapidly pass across the seam, each partial seam adjacent to moving objects is updated. In other words, the proposed seam update can be used to refine the initial seam locally for each object region.





Fig.7 Original images and stitched result. (a), (b) Original images. (c) The stitched image.



Fig. 8 The magnified image for performance comparison of path-finding methods. (a) Gu's enhanced dynamic programming. (b) GIST1. (c) Proposed algorithm.

Since the proposed algorithm is designed for a real time video stitching, the complexity of the algorithm is also important. Table 1 gives a comparison of the computation complexity of the algorithms by measuring running time in seconds on a 2.4 GHz PC with a dual core processor with 4 GB memory. To compare the processing time, Figs. 7 (a) and (b) which have 539,774 pixels in the overlapped region is used. On average, our algorithm runs 2.7 times faster than





(c)



Fig. 9 Results of seam update process. (a) The predefined seam. (b)-(c) The updated seam following the object contour. (d) The recovered original seam after the object passes.

Table 1 Computational Complexity (UNIT : MS).

Test images	Gu's method	GIST1	Proposed
Fig. 7	47.9	46.8	17.6

the Gu's algorithm and 2.6 times faster than GIST1 during path finding since the proposed Greedy based method significantly reduces the number of computed nodes for path finding compared to the conventional method.

4. Conclusions

In this paper, we presented a fast video stitching algorithm that uses an adaptive seam update based on a greedy pathfinding algorithm. By comparing it with the conventional path searching approaches, the proposed algorithm is proven to be superior to conventional ones in terms of both the computational complexity and the visual quality of the stitching results. Therefore, the proposed image stitching algorithm can be adopted in surveillance systems or in real-time multicamera stitching systems.

Acknowledgments

This research was supported by Seoul Future Contents Convergence (SFCC) Cluster established by Seoul R&BD Program (No. 10570) and by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2010-0000449). This work was supported by a Korea University Grant.

References

- M. Brown and D. Lowe, "Recognising panoramas," Proc. Int. Conf. Computer Vision, vol.2, pp.1218–1225, Oct. 2003.
- [2] M. Uyttendaele, A. Eden, and R. Szeliski, "Eliminating ghosting and exposure artifacts in image mosaics," Proc. Conf. Computer Vision and Pattern Recognition, vol.2, pp.509–516, 2001.
- [3] W. Freeman, E. Pasztor, and O. Carmichael, "Learning low-level vision," Proc. Int. Conf. Computer Vision, vol.40, no.1, pp.1182– 1189, 1999.
- [4] A. Zomet, A. Levin, S. Peleg, and Y. Weiss, "Seamless image stitching by minimizing false edge," IEEE Trans. Image Process., vol.15, no.4, pp.969–977, 2006.
- [5] D. Lowe, "Distinctive image features from scale-invariant keypoints," Int. J. Comput. Vis., vol.60, no.2, pp.91–110, 2004.
- [6] M. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography," Commun. ACM, vol.24, pp.381–395, 1981.
- [7] M. Brown and D. Lowe, "Automatic panoramic image stitching using invariant features," Int. J. Comput. Vis., vol.74, no.1, pp.59–73, 2007.
- [8] P. Pérez, M. Gangnet, and A. Blake, "Poisson image editing," ACM Trans. Graph., vol.22, no.3, pp.313–318, 2003.
- [9] T. Porter and T. Duff, "Compositing digital images," Inter. Conf. on Computer Graphics and Interactive Techniques, vol.18, no.3, pp.253–259, 1984.
- [10] J. Davis, "Mosaics of scenes with moving objects," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.354–360, 1998.
- [11] Y. Boykov and V. Kolmogorov, "An experimental comparison of Min-Cut/Max-Flow algorithms for energy minimization in vision," IEEE Trans. Pattern Anal. Mach. Intell., vol.26, no.9, pp.1124– 1137, Sept. 2004.
- [12] T.D. Yeh, Y.P. Chen, and Z.Y. Liao, "An image stitching process using band-type optimal partition method," Asian Journal of Information Technology, vol.7, no.11, pp.498–509, 2008.
- [13] K. Kim, T.H. Chalidabhongse, D. Harwood, and L. Davis, "Realtime foreground-background segmentation using codebook model," Real-Time Imaging, vol.11, no.3, pp.172–185, 2005.
- [14] H. Gu, Y. Yu, and W. Sun, "A new optimal seam selection method for airborne image stitching," IEEE Int. Workshop on IST'09, pp.159– 163, 2005.
- [15] J. Canny, "A computational approach to edge detection," IEEE Trans. Pattern Anal. Mach. Intell., vol.8, no.6, pp.679–714, 1986.



Kwang-Wook Lee received the B.S. degree in Electronics Engineering from Korea University in 2004. He joined the Multimedia Communication Processing Lab in the Department of Electronic Engineering of Korea University, in 2004. He is now a Ph.D. candidate at Korea University. His interests are video signal processing, scalable video coding, image processing, and computer vision.



Seung-Won Jung received the B.S. degree in electronics engineering from the Department of Electrical Engineering, Korea University, Seoul, Korea, in 2005, where he is currently working toward the Ph.D. degree in electronics engineering. His research interests include image enhancement, image restoration, video compression, and computer vision.



Seung-Kyun Kim received the B.S. degree in Electronic Engineering from Chung-Ang University in 2005. He is currently pursuing the Ph.D. degree in Electronic Engineering at Korea University. He is interested in visual surveillance, image processing and video signal processing.



Sung-Jea Ko received the Ph.D. degree in 1988 and the M.S. degree in 1986, both in Electrical and Computer Engineering, from State University of New York at Buffalo, and the B.S. degree in Electronic Engineering at Korea University in 1980. In 1992, he joined the Department of Electronic Engineering at Korea University where he is currently a Professor. From 1988 to 1992, he was an Assistant Professor of the Department of Electrical and Computer Engineering at the University of

Michigan-Dearborn. He has published over 140 international journal articles. He also holds over 40 patents on video signal processing and multimedia communications. He is currently a Fellow in the IET and a Korean representative of IEEE Consumer Electronics society. He is the 1999 Recipient of the LG Research Award given to the Outstanding Information and Communication Researcher. He received the Hae-Dong best paper award from the IEEK (1997) and the best paper award from the IEEE Asia Pacific Conference on Circuits and Systems (1996), and the research excellence award from Korea University (2004).