

Casual Stereoscopic Panorama Stitching

Fan Zhang and Feng Liu
Department of Computer Science
Portland State University
{zhangfan, fliu}@cs.pdx.edu

Abstract

This paper presents a method for stitching stereoscopic panoramas from stereo images casually taken using a stereo camera. This method addresses three challenges of stereoscopic image stitching: how to handle parallax, how to stitch the left- and right-view panorama consistently, and how to take care of disparity during stitching. This method addresses these challenges using a three-step approach. First, we employ a state-of-the-art stitching algorithm that handles parallax well to stitch the left views of input stereo images and create the left view of the final stereoscopic panorama. Second, we stitch the input disparity maps to obtain the target disparity map for the stereoscopic panorama by solving a Poisson's equation. This target disparity map is optimized to avoid vertical disparities and preserve the original perceived depth distribution. Finally, we warp the right views of the input stereo images and stitch them into the right-view panorama according to the target disparity map. The stitching of the right views is formulated as a labeling problem that is constrained by the stitching of the left views to make the left- and right-view panorama consistent to avoid "retinal rivalry". Our experiments show that our method can effectively stitch casually taken stereo images and produce high-quality stereo panoramas that deliver a pleasant stereoscopic 3D viewing experience.

1. Introduction

Panorama stitching is a well studied topic and many software tools are available for users to create panoramas [26]. Most of these methods, however, are designed for monocular image stitching. Employing a monocular image stitching method to independently create the left and right view of a stereoscopic panorama is problematic as the left and right panorama may not be consistent. As shown in Figure 1 (a), the cat in the left panorama is different from that in the right panorama. This is because the input images are taken at different time and the cat appears different in the input images. The left and right panorama take the cat from different input images. The inconsistency will lead to "retinal rivalry" and

bring in "3D fatigue" to viewers [17]. Moreover, stereoscopic images have an extra dimension of disparity, which cannot be taken care of by independently stitching the two views. Figure 1 (a) shows that the resulting panorama has vertical disparities in the car headlight and tire area. This will also compromise the 3D viewing experience. Dedicated stereoscopic image stitching methods have been developed [8, 19, 21]. However, these methods require a user to densely sample the scene using a video camera and/or follow some specific rules to rotate the camera and cannot work well with a sparse set of casually taken input images.

The goal of this paper is to develop a technology that allows users to create stereoscopic panoramas as conveniently as monocular ones. As consumer stereo cameras now become more and more available to daily users, it becomes easy for them to take stereoscopic images. We therefore aim to develop a stereoscopic image stitching method that enables users to generate stereoscopic panoramas from casually taken stereo images. To achieve this goal, we need to address three challenges. First, our method needs to handle parallax well. No matter how a user moves a stereo camera, images from at least one of the left and right view have parallax. As we allow users to freely move the stereo camera, it is common that images from both views have parallax. Second, our method needs to stitch the left and right panorama consistently. Third, our method needs to take care of disparity to deliver a pleasant viewing experience.

This paper presents a three-step stereoscopic image stitching method to address the above challenges. First, we employ a state-of-the-art parallax-tolerant monocular image stitching method to create one of the two views of the stereoscopic panorama. Without loss of generality, we always select the left-view panorama to stitch first. Second, we stitch the disparity maps of the input stereoscopic images to create the target disparity map for the stereoscopic panorama by solving a Poisson's equation. This target disparity map is optimized to avoid vertical disparities and seamlessly merge the perceived depth field of the input stereoscopic images. Finally, we warp the right views of the input stereoscopic images and stitch them into the right-



(a) Independent stitching

(b) Our result

Figure 1. Stereoscopic panorama stitching. For each column, we show the left-view, right-view and red-cyan anaglyph of the stereo panorama. Stitching the left- and right-view panorama independently brings in inconsistency artifacts like monocular object (the cat) and vertical disparities (the car headlight and tire), which will cause “3D fatigue” to viewers. Our result is free from these artifacts.

view panorama according to the target disparity map. The stitching of the right views is formulated as a labeling problem that is constrained by the stitching of the left views to make the left- and right-view panorama consistent.

This paper contributes a stereoscopic image stitching method that allows users to generate stereoscopic panoramas as conveniently as they generate monocular ones. To develop this stereoscopic image stitching method, this paper also provides a novel algorithm to seamlessly stitch input disparity maps and a seam-cutting method to stitch the right panorama that is consistent with the stitching of the left panorama and respects the target disparity map. Our experiments show our method allows for easy production of stereoscopic panoramas that deliver a pleasant 3D panoramic viewing experience.

2. Related Work

Monocular image stitching is a well studied topic. A good survey can be found in [26]. This section focuses on stereoscopic image stitching and techniques for parallax handling which are most relevant to our work.

Stereoscopic image stitching. Stereoscopic panoramas require source images for the left and right panorama to be taken from different viewpoints. These images can be recorded using either a stereo camera or a moving monocular camera [4, 8, 10, 21, 19, 22, 24]. The early *PSI* system uses a stereo camera rig and rotates it horizontally around an axis passing through the optical center of the right camera

to collect a set of left images and a set of right images [8]. The left and right panorama are then created using a disparity warping technique and a hierarchical seaming algorithm. Couture *et al.* developed a stereoscopic panoramic video stitching method that captures input videos by rotating a stereo camera rig around an off camera center vertical axis [4]. Peleg *et al.* developed an *omnistereo* panorama system that mounts a monocular camera on a rotating arm to capture images from various viewpoints. The left and right panorama can then be synthesized by taking proper strips from input views [19]. Richardt *et al.* further improved this *omnistereo* method by correcting the deviations from the ideal capture setup and addressing the insufficient sampling problem using a flow-based ray upsampling algorithm [21]. All these existing methods require users to densely sample the scene using a camera and/or follow some specific rules to rotate it. In contrast, our work only requires sparse samples of the scene casually captured by a stereo camera. A recent method merges stereo images of different scenes into a panorama; however, it does not produce a regular panorama that conveys a wide field of view of the same scene [27].

Parallax handling. Traditional homography-based image stitching methods cannot handle parallax well [2, 26]. Thus, techniques like local warping [23], seam cutting [1, 5, 12], and blending [3, 20], are developed to reduce or eliminate artifacts caused by parallax. Spatially-varying warps are recently employed to align images for image stitching [14, 28, 29]. Since these methods are more flexible,

they can often better handle parallax than homography. Recent research shows that images do not need to be globally aligned to produce a good stitching result. A recent method, instead of estimating a best-fitting homography, searches for a good homography that enables optimal stitching to align input images [6]. A local stitching method further develops this idea and finds a local alignment that combines homography and spatially-varying warp to better handle parallax and allows for high-quality stitching [29]. Our method builds upon these existing methods to handle parallax. The first step of our approach uses the recent local stitching method [29] to stitch the left view of the final stereoscopic panorama. Our method also extends a spatially-varying warping method to transform the right views of the input stereoscopic images according to the target disparity map in a way that is robust against parallax.

3. Stereoscopic Panorama Stitching

Our method takes as input a sparse set of stereoscopic images casually captured using a stereoscopic camera and outputs a stereo panorama. We consider that a good stereoscopic panorama has the following properties. First, both the left and right panorama should be artifact-free. Second, the left and right panorama should be consistently stitched to avoid “retinal rivalry”. Third, the disparity map of the stereoscopic panorama should be carefully taken care of. Stitching should introduce no vertical disparities. Moreover, the horizontal disparity maps of input images should be seamlessly stitched to ensure proper depth perception.

In order to create such a good stereoscopic panorama, our method decomposes stereoscopic panorama stitching into three separate steps after a pre-processing step to estimate disparity maps of input stereoscopic images.

1. Stitch the left panorama from the left views of input stereoscopic images using a state-of-the-art monocular stitching algorithm.
2. Stitch the target disparity map of the output stereoscopic panorama from the disparity maps of input stereoscopic images.
3. Warp the right views of input stereoscopic images and stitch the right panorama according to the stitching of the left panorama and the target disparity map.

For simplicity, we consider the task of stitching two input stereoscopic images I_1 and I_2 . More images can be stitched similarly. Each stereoscopic image has a left and right image. For example, $I_{1,l}$ and $I_{1,r}$ are the left and right image of I_1 , respectively. We denote the two views of the output stereoscopic panorama as \hat{I}_l^p and \hat{I}_r^p .

Our method pre-processes input stereoscopic images to estimate their disparity maps. Like previous methods in stereoscopic image editing [13], we downsample each input

stereoscopic image, estimate dense correspondences from the downsampled images using an optical flow method [25], and scale up the resulting optical flow vectors as the disparities of the original image. We denote the disparity map for the input stereoscopic images I_1 and I_2 as D_1 and D_2 , respectively. We describe each step of our method below.

3.1. Left panorama stitching

Our method starts by stitching one of the two views of a stereoscopic panorama. Without loss of generality, our method selects to create the left panorama first. As our method allows a user to casually capture input stereoscopic images, there is parallax among the left input images. Actually, no matter how the input stereoscopic images are taken using a stereoscopic camera, parallax exists at least in one of the two views. Therefore, we choose to use monocular image stitching methods [6, 28, 29] that can handle parallax to create the left panorama.

Specifically, our method uses a recent parallax-tolerant monocular image stitching method [29]. This monocular stitching method first finds an optimal local alignment that allows for optimal stitching. The local alignment is a combination of homography-based warp and spatially-varying warp. Once input images are locally aligned, they are composed together using a seam-cutting algorithm [12] and a multi-band blending algorithm [3]. Please refer to the original paper for more details [29]. This step outputs the left panorama \hat{I}_l^p as well as the intermediate stitching information that will be used in later steps, including the warped left images and the seam where the warped images are merged.

3.2. Target panoramic disparity map estimation

A stereoscopic image has an extra dimension of disparity, which controls the perceived depth [17]. To generate a good stereoscopic panorama, we need to not only stitch the input images, but also seamlessly stitch the disparity maps of input images to ensure proper 3D depth perception. Like [16, 18], we stitch the disparity maps in the disparity gradient domain using a Poisson blending method [20] and obtain the target disparity map \hat{D}^p of the stereoscopic panorama. Specifically, we minimize the following energy function that aims to preserve the disparity gradients of the input stereoscopic images I_1 and I_2 .

$$\sum_{\hat{d}_i} \sum_{j \in N_i} \|(\hat{d}_i - \hat{d}_j) - dd_{i,j}\|^2, \quad (1)$$

$$\text{where } dd_{i,j} = \begin{cases} d_{1,i} - d_{1,j} & \text{if } l_i = 1 \\ d_{2,i} - d_{2,j} & \text{if } l_i = 2 \end{cases}$$

where \hat{d}_i and \hat{d}_j are the target disparities at neighboring pixels i and j of the left panorama \hat{I}_l^p and N_i is the four-connected neighborhood of pixel i . $dd_{i,j}$ is the disparity

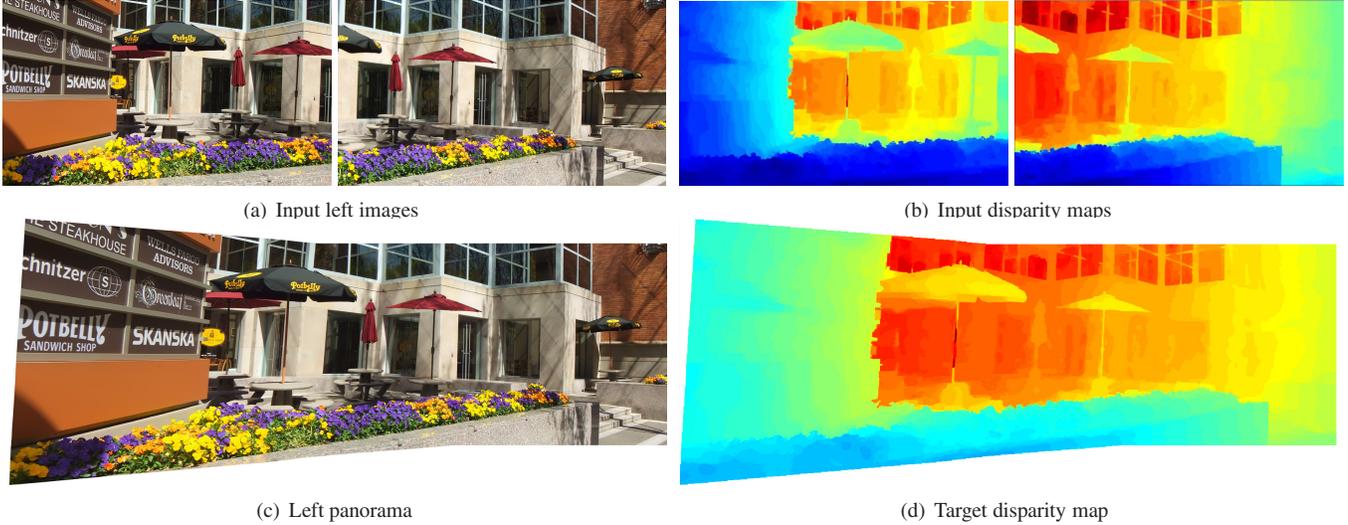


Figure 2. Target panoramic disparity estimation. On the left, we show the left images of input stereoscopic images (a) and the left panorama (c). On the right, we show the input disparity maps (b) and the target disparity map (d) stitched from the input disparity maps.

difference between pixel i and j in the proper input stereoscopic image. If pixel i in the left panorama comes from the input image $I_{1,l}$, which is indicated by its label $l_i = 1$, $dd_{i,j}$ takes the disparity difference in the input stereoscopic image I_1 . These labels come from the seam-cutting step in creating the left panorama. Similarly, if pixel i comes from the input image $I_{2,l}$, $dd_{i,j}$ takes the disparity difference in the input stereoscopic image I_2 . Here the input disparity values like $d_{1,i}$ can be obtained by finding the corresponding pixel in the input image according to the warping applied to create the left panorama and taking the corresponding disparity value. Finally, we set the boundary condition of the above energy minimization problem by keeping the original disparities of the pixels that originally come from $I_{1,l}$ and are out of the overlapping region. Figure 2 shows an example of target panoramic disparity estimation.

Since a panoramic image typically contains a large number of pixels, the above Poisson’s equation involves a large number of variables. To make this step efficient, we divide the left panorama into a uniform grid mesh and only compute the disparities for the grid vertices. Our experiments show that the mesh cell size of 5×5 pixels works well. This step outputs \hat{D}^p , the target disparity map of the final panorama (even before we create it). A user can further edit this target disparity map to manipulate the stereoscopic 3D viewing experience using tools like non-linear disparity mapping [13]. The disparity maps of all the results in this paper are directly from the above Poisson blending algorithm and they are not post-edited unless otherwise noted.

3.3. Right panorama stitching

After we obtain the target disparity map \hat{D}^p of the stereoscopic panorama, we first warp the right images $I_{1,r}$ and $I_{2,r}$ of the input stereo images according to the target dis-

parity map. This warping step aligns the right input images as they are warped accordingly to the same target disparity map. Compared to the common method that aligns images based on the feature correspondences between these images, this approach has an important advantage in that the alignment result better respects the target disparity map and avoids introducing vertical disparities. Once we warp the right input images, we stitch these warped images using an extended seam-cutting method guided by the left panorama to create the right panorama.

3.3.1 Right input image warping

All the right input images are warped in the same way. For simplicity and clarity, we omit the subscripts $\{1, 2\}$ here. For each grid vertex in the left panorama, we first find its corresponding point in the corresponding left input image by inverting the warping used to align left images to create the left panorama. We then find its corresponding point in the right image according to the input disparity map. In this way, we obtain a set of control points in the right image, denoted as $\{\mathbf{p}_{r,i}\}$. Their corresponding points in the left panorama are $\{\hat{\mathbf{p}}_{l,i}^p\}$. The disparities of these control points are known from \hat{D}^p . Our method uses these control points to guide the warping of each right image.

Various spatially-varying warp methods have been developed to warp an image guided by a set of control points [11, 13, 15, 28, 29]. We extend these methods to warp each right input image guided by the set of control points. Specifically, we divide each right image into a uniform grid mesh and formulate image warping as a mesh warping problem, where the unknowns are the coordinates of mesh vertices. The mesh warping problem is defined as a quadratic minimization problem that enforces the disparities of the control points and minimizes visual distortion.

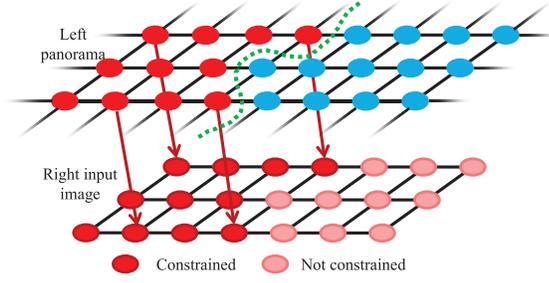


Figure 3. Control points. The control points only exist on one side of the stitching seam. Thus a part of the right-view image (with the light-red points) has no disparity constraint and will be distorted during warping if not taken care of.

We describe the energy terms below.

Disparity term. Our method encourages the control points to have the target disparities so that the stereoscopic panorama can deliver proper depth perception to viewers. Since each control point $\mathbf{p}_{r,i}$ in the right-view image is not necessarily a grid vertex, we first find the grid cell that encloses the control point in the right image and then represent it as a linear combination of the cell’s four vertices $\{\mathbf{v}_j\}$. The combination coefficients w_j are computed using the inverse bilinear interpolation method [7]. These coefficients are then used to combine the vertices $\hat{\mathbf{v}}_j$ in the output image to compute the location of the control point in the output image. We define the disparity energy term below.

$$E_d = \sum_{\mathbf{p}_{r,i}} \left\| \sum_j w_j \hat{\mathbf{v}}_j - \hat{\mathbf{p}}_{l,i}^p - \hat{\mathbf{d}}_i \right\|^2 \quad (2)$$

where $\hat{\mathbf{d}}_i$ is the target disparity vector of the control point $\mathbf{p}_{r,i}$, taking the form $\hat{\mathbf{d}}_i = [d_i \ 0]^T$, where d_i is the target (horizontal) disparity and the vertical disparity is set 0. $\hat{\mathbf{p}}_{l,i}^p$ is the corresponding point of $\mathbf{p}_{r,i}$ in the left panorama.

Global alignment term. The disparity term only directly constrains warping of the image region with control points. These control points, however, only exist on one side of the stitching seam in the left image that is finally selected to make the left panorama, as illustrated in Figure 3. For the regions with no control points, warping often distorts them. To solve this problem, we first estimate the best-fitting homography according to the control points and then employ this best-fitting homography to pre-warp the right input image. As the pre-warping result often provides a good approximation, our method encourages the regions without control points to be as close to the pre-warping result as possible. We define the global alignment term as follows

$$E_g = \sum_i \tau_i \|\hat{\mathbf{v}}_i - \bar{\mathbf{v}}_i\|^2 \quad (3)$$

where $\hat{\mathbf{v}}_i$ and $\bar{\mathbf{v}}_i$ are the corresponding vertices in the warping result and in the pre-warping result. τ_i is a binary value. We set it 0 if there is a control point in the neighborhood of $\hat{\mathbf{v}}_i$; otherwise it is 1.

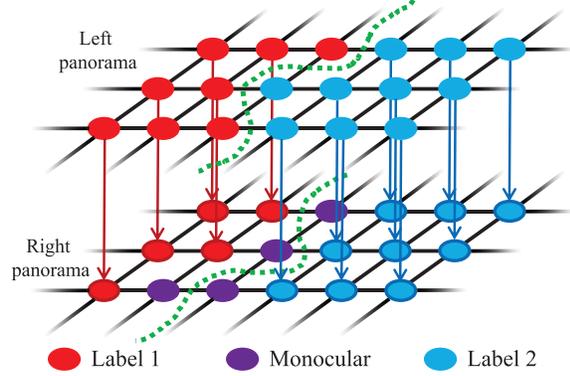


Figure 4. Seam-cutting for the right panorama. The labels, 1 or 2, from the left-view seam-finding result are propagated to the corresponding pixels in the right view (bottom). Pixels (in purple) in the monocular region in the right view will not have recommended labels from the left view. Propagated labels from the left panorama (top) are encoded as soft constraints in Equation 6 to tolerate left-right matching errors or make trade-off for the monocular stitching quality in Equation 7.

Smoothness term. To minimize visual distortion, our method encourages each grid cell to undergo a similarity transformation. We use the quadratic energy term from [9] to encode the similarity transformation constraint.

$$E_s = \sum_{\hat{\mathbf{v}}_i} w_i \|\hat{\mathbf{v}}_i - (\hat{\mathbf{v}}_j + u(\hat{\mathbf{v}}_k - \hat{\mathbf{v}}_j) + v\mathbf{R}(\hat{\mathbf{v}}_k - \hat{\mathbf{v}}_j))\|^2 \quad (4)$$

where $\hat{\mathbf{v}}_i$, $\hat{\mathbf{v}}_j$, and $\hat{\mathbf{v}}_k$ are every three vertices of a grid cell in the output mesh. w_i is the average saliency value inside the triangle defined by the three vertices and is computed using the same method as [15]. u and v are the coordinates of \mathbf{v}_i in the local coordinate system defined by \mathbf{v}_j and \mathbf{v}_k , where \mathbf{v}_i , \mathbf{v}_j , and \mathbf{v}_k are the corresponding vertices in the input mesh of the right image. $\mathbf{R} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

We combine the above energy terms and obtain the following linear least squares problem.

$$E = E_d + \lambda E_g + \gamma E_s \quad (5)$$

where λ and γ are weights with default values 0.7 and 0.4, respectively. We solve this energy minimization problem using a sparse linear solver. The outputs from this step are the warped right images $\hat{I}_{1,r}$ and $\hat{I}_{2,r}$ according to the target disparity map of the stereoscopic panorama.

3.3.2 Seam-cutting for right panorama stitching

We develop a seam-cutting method to stitch the warped right images $\hat{I}_{1,r}$ and $\hat{I}_{2,r}$ guided by the seam-cutting result in creating the left panorama. The goal is to create the right panorama such that it is consistent with the left panorama. We extend the seam-cutting method for monocular image stitching [12] with an extra energy term to handle the stereo



(a) Baseline result (independent stitching). From left to right: left-view, right-view and red-cyan anaglyph of the stereo panorama



(b) Our result. From left to right: left-view, right-view and red-cyan anaglyph of the stereo panorama

Figure 5. Comparison between independent stitching results and our results.

consistency problem. Specifically, we formulate this seam-cutting problem as a labeling problem. For each pixel in the overlapping region, we aim to assign it with a label either 1 or 2, indicating the pixel coming from $\hat{I}_{1,r}$ or $\hat{I}_{2,r}$. We now describe the energy terms for this labeling problem.

Stereo consistency term. Our method encourages pixels in the right panorama to take the same labels as their corresponding pixels in the left panorama, as shown in Figure 4. Therefore, for each pixel that a corresponding pixel in the left panorama can be found for, we encourage it to take the same label as the corresponding pixel in the left view.

$$E_{sc}(L) = \sum_{i \in S} \rho_i \delta(l_i \neq l_i^l) \quad (6)$$

where S is the set of pixels in the right panorama that we can find corresponding pixels in the left panorama for and is found in a similar way described in Section 3.3.1. L is the labeling map for pixels in the right panorama, l_i is the label for pixel i , and l_i^l is the label of the corresponding pixel in the left panorama. $\delta(l_i \neq l_i^l)$ is an indicator function that takes value 1 if $l_i \neq l_i^l$ and 0 otherwise. ρ_i is a weight that measures the confidence of matching pixel i between the left and right view, which is computed based on the color difference between pixels/patches or is available from the output of many optical flow and stereo matching algorithms.

Monocular color term. To create a seamless right panorama, our method aims to minimize the color difference between the overlapping regions of the $\hat{I}_{1,r}$ and $\hat{I}_{2,r}$ along the seam. Consider two adjacent pixels i and j in the overlapping region. If these two pixels take different labels, the color difference between $\hat{I}_{1,r}$ and $\hat{I}_{2,r}$ at pixel i and j should be as small as possible.

$$E_{mc}(L) = \sum_{i,j} (d(i, \hat{I}_{1,r}, \hat{I}_{2,r}) + d(j, \hat{I}_{1,r}, \hat{I}_{2,r})) \delta(l_i \neq l_j) \quad (7)$$

$$d(i, \hat{I}_{1,r}, \hat{I}_{2,r}) = \|\hat{I}_{1,r}(i) - \hat{I}_{2,r}(i)\|^2$$

where $d(i, \hat{I}_{1,r}, \hat{I}_{2,r})$ is the color difference at pixel i between $\hat{I}_{1,r}$ and $\hat{I}_{2,r}$, and $\delta(l_i \neq l_j)$ is an indicator function, taking value 1 if $l_i \neq l_j$ and 0 otherwise.

We combine the above terms and get the following minimization problem that aims to find an optimal labeling map.

$$\begin{aligned} E(L) &= \alpha E_{sc}(L) + E_{mc}(L) \quad (8) \\ \text{s.t. } l_k &= \begin{cases} 1 & \text{if } k \in \hat{I}_{1,r}^m \\ 2 & \text{if } k \in \hat{I}_{2,r}^m \end{cases} \end{aligned}$$

where α is a parameter with default value 0.5. $\hat{I}_{1,r}^m$ denotes the non-overlapping region in $\hat{I}_{1,r}$ where pixels take label 1. Similarly, pixels in the non-overlapping region $\hat{I}_{2,r}^m$ take label 2. We solve the above labeling problem using a standard graph-cut algorithm. After we find the seam, we use the seam and the multi-band blending algorithm [3] to compose the final right panorama.

4. Experiments

We experimented with our stereoscopic image stitching methods on a variety of images taken by stereo cameras *Fujifilm FinePix 3D W3* and *Panasonic HDC-Z10000*. These input stereo images were casually taken by these handheld cameras and therefore both the left images and right images exhibit large parallax. We compare our method to a baseline solution that employs a state-of-the-art monocular stitching method to stitch the left and right panorama independently [29]. Since the baseline method creates the left and right panorama independently, the disparity distribution is often problematic. For the panoramas from the baseline method, we manually shifted the left and right panorama vertically so that there are as small vertical disparities as possible in the main object. We also shifted them horizontally so that the horizontal disparities in the main object are as similar to the corresponding panoramas created by our method as possible. Our results were not adjusted.



(a) Baseline result (Independent stitching)

(b) Our result

Figure 6. Comparison between independent stitching results and our results. In each column, we show the left-view, right-view and red-cyan anaglyph of the stereo panorama.

Figure 5 shows the left, right, and red-cyan anaglyph versions of the stereoscopic panoramas stitched by the baseline solution and our method. The baseline method independently creates the left and right panorama and thus cannot ensure the consistency between the two panoramas. For example, a person in red T-shirt appears in the left panorama but disappears in the right one, as shown in Figure 5 (a). This inconsistency brings in the “monocular object violation” [17]. This is because different seams are used to stitch the left and right panorama. Our method stitches the right panorama constrained by the left panorama and is free from this monocular object violation, as shown in Figure 5 (b).

Figure 6 shows another example. Although we manually aligned the left and right panorama of the baseline result, significant vertical disparities still exist, as shown in Figure 6 (a), which will cause “3D fatigue” [17]. Since our method warps the right input images according to the target panoramic disparity map, our result is free from the vertical disparity artifacts, as shown in Figure 6 (b). Please refer to our project website for more results¹.

¹<http://graphics.cs.pdx.edu/project/stereostitch>

4.1. User Study

We conducted a user study to evaluate the user experience of viewing stereoscopic panoramas created by our method and the baseline method. Our study displayed stereoscopic panoramas on an *ASUS VG236H* 3D monitor with shuttered glasses. We selected 10 sets of input stereo images. For each set, we created two stereoscopic panoramas, one using our method and the other using the baseline solution. We obtained 20 stereoscopic panoramas in total.

There were 10 participants in our study, including 4 females and 6 males. They are students from various departments, including computer science, civil engineering, chemistry, biology, etc. They all have normal stereopsis perception. They do not know how each panorama was created. Before the study, we provided four panoramas, two from each method, for them to look at to get used to viewing stereoscopic panoramas. In our study, we showed the 20 stereoscopic panoramas mentioned above to each participant one by one in a random order. Participants can look at a panorama as long as they want. After a participant finishes looking at a panorama, we ask three questions.

1. Is it easy for you to perceive 3D?
2. Do you feel comfortable viewing the panorama?

	3D		Comfort		Quality	
	mean	std	mean	std	mean	std
Ours	4.10	0.52	4.19	0.61	3.84	0.45
Baseline	3.78	0.72	2.68	0.46	2.70	0.62

Table 1. User study results.

3. Are you satisfied with the quality of the panorama?

The participant rated each question using a Likert scale ranging from 1 to 5, with 5 being the most positive. We report the average scores (μ) and the standard deviations (σ) in Table 1. Our study confirms our hypothesis that independently creating the left and right view of a stereoscopic panorama is problematic and will damage the stereoscopic 3D viewing experience. The average comfort rate for the baseline results is 2.68. In contrast, our results deliver a more comfortable 3D viewing experience with the average rate 4.19. Similarly, users are more satisfied with our results ($\mu=3.84$, $\sigma=0.45$) than the baseline results ($\mu=2.70$, $\sigma=0.62$). The p -values of the paired two sample t-test between our results and the baseline results for both comparisons are smaller than 0.001, which shows that the difference between the two sets of results is very significant.

The stereoscopic panoramas from both our method and the baseline method can allow users to easily obtain 3D perception and our results ($\mu=4.10$, $\sigma=0.52$) are easier for users than the baseline result ($\mu=3.78$, $\sigma=0.72$). The p -value for the study is 0.123, which shows that the difference between the two sets of results is not statistically significant.

The post-study informal feedback shows that participants complained most that they cannot fuse the left and right view to obtain 3D perception for some regions in some panoramas. We found that this problem is due to the inconsistency between the left and right panorama. For example, a visually salient object only appears in one of the two views, which brings in “retinal rivalry”. The inconsistency problem only occurs in the baseline results.

4.2. Discussion

Compared to the monocular stitching method [29], our three-step approach adds two extra steps besides a standard pre-processing step to estimate disparity maps of input images: target panoramic disparity map stitching and right panorama stitching. Our method solves a Poisson’s equation to stitch the target disparity map. On two input images with size about 900×600 , this step takes less than 1 second using a desktop machine with *Intel i7* CPU and 16 GB memory. Right panorama stitching has two main computational steps: spatially-varying warping and seam-cutting using graph-cut. These two take less than 1 second in total.

Our method relies on the disparity maps of input stereo images to take care of disparity and consistency issues during stitching. The input disparity maps sometimes contain errors; however, we found that our method is robust against the disparity errors. For example, we use input disparities to

establish the left-right correspondences and propagate the labels from the left panorama to the right one. When a pixel in the right panorama is mapped onto a wrong pixel in the left panorama, the label recommended for it will mostly be corrected as neighboring pixels in the left panorama often share the same label except across the stitching seam. Moreover, our method encodes labeling propagation as a soft constraint. The errors can usually be corrected by the other energy term in the optimization. Similarly, while the warping of the right images is guided by the target disparity map, disparity errors in a few pixels will be corrected by the smoothness term of the warping energy function.

Our method could use sparse feature correspondences to replace the disparity map. As another baseline, we replaced the disparity map with the SIFT feature correspondences. Our tests showed that 1) this new baseline works better than the current baseline, and 2) it is not as robust as our current method. When feature points are not evenly distributed, the regions with no feature points are warped and stitched differently in the left and right panorama.

While this paper only shows examples of stitching two images, our method can be easily extended. For example, after creating the left view of a 360° panorama using cylindrical or spherical projection and parallax-tolerant stitching, we warp and stitch the right images guided by it and create a 360° stereo panorama. Please refer to our project website for more results.

The first step of our method uses a state-of-the-art monocular stitching method to stitch the left panorama [29]. While this monocular stitching method can handle parallax well in general, when the parallax among images from the same view is very large, it sometimes suffers from alignment artifacts. As our method builds upon this first step to produce the right panorama, the right panorama produced by our method also shares similar artifacts. Since our three-step approach is flexible, it can easily replace the monocular stitching method that is currently used with a more advanced monocular stitching method in future.

5. Conclusion

This paper described a stereoscopic image stitching method that allows users to create stereoscopic panoramas as conveniently as monocular panoramas. Our method consists of three steps. The first step stitches the left images of input stereo images into the left panorama using a state-of-the-art monocular stitching method. The second step stitches the disparity maps of input images into the target disparity map of the final stereoscopic panorama. The third step stitches the right input images into the right panorama guided by the target disparity map and the stitching process of the left panorama. Our experiments show that our method can create high-quality stereoscopic panoramas that deliver a pleasant 3D viewing experience to users.

Acknowledgements. This work was supported by NSF grants IIS-1321119, CNS-1205746, and CNS-1218589.

References

- [1] A. Agarwala, M. Dontcheva, M. Agrawala, S. Drucker, A. Colburn, B. Curless, D. Salesin, and M. Cohen. Interactive digital photomontage. *ACM Trans. Graph.*, 23(3):294–302, 2004. [2](#)
- [2] M. Brown and D. G. Lowe. Automatic panoramic image stitching using invariant features. *Int. J. Comput. Vision*, 74(1):59–73, 2007. [2](#)
- [3] P. J. Burt and E. H. Adelson. A multiresolution spline with application to image mosaics. *ACM Transactions on Graphics*, 2(4):217–236, 1983. [2](#), [3](#), [6](#)
- [4] V. Couture, M. S. Langer, and S. Roy. Panoramic stereo video textures. In *IEEE International Conference on Computer Vision*, pages 1251–1258, 2011. [2](#)
- [5] A. Eden, M. Uyttendaele, and R. Szeliski. Seamless image stitching of scenes with large motions and exposure differences. In *IEEE CVPR*, pages 2498–2505, 2006. [2](#)
- [6] J. Gao, Y. Li, T.-J. Chin, and M. S. Brown. Seam-driven image stitching. In *Eurographics 2013*, pages 45–48, 2013. [3](#)
- [7] P. S. Heckbert. Fundamentals of texture mapping and image warping. Technical Report UCB/CSD-89-516, EECS Department, University of California, Berkeley, Jun 1989. [5](#)
- [8] H.-C. Huang and Y.-P. Hung. Panoramic stereo imaging system with automatic disparity warping and seaming. *Graphical Models and Image Processing*, 60(3):196–208, 1998. [1](#), [2](#)
- [9] T. Igarashi, T. Moscovich, and J. F. Hughes. As-rigid-as-possible shape manipulation. *ACM Transactions on Graphics*, 24(3):1134–1141, 2005. [5](#)
- [10] H. Ishiguro, M. Yamamoto, and S. Tsuji. Omni-directional stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):257–262, 1992. [2](#)
- [11] P. Krähenbühl, M. Lang, A. Hornung, and M. Gross. A system for retargeting of streaming video. *ACM Trans. Graph.*, 28(5):126:1–126:10, 2009. [4](#)
- [12] V. Kwatra, A. Schödl, I. Essa, G. Turk, and A. Bobick. Graphcut textures: image and video synthesis using graph cuts. *ACM Trans. Graph.*, 22(3):277–286, 2003. [2](#), [3](#), [5](#)
- [13] M. Lang, A. Hornung, O. Wang, S. Poulakos, A. Smolic, and M. Gross. Nonlinear disparity mapping for stereoscopic 3d. *ACM Transactions on Graphics*, 29(4):75:1–75:10, 2010. [3](#), [4](#)
- [14] W.-Y. Lin, S. Liu, Y. Matsushita, T.-T. Ng, and L.-F. Cheong. Smoothly varying affine stitching. In *IEEE CVPR*, pages 345–352, 2011. [2](#)
- [15] F. Liu, M. Gleicher, H. Jin, and A. Agarwala. Content-preserving warps for 3d video stabilization. *ACM Transactions on Graphics*, 28(3):44:1–44:9, 2009. [4](#), [5](#)
- [16] S.-J. Luo, I.-C. Shen, B.-Y. Chen, W.-H. Cheng, and Y.-Y. Chuang. Perspective-aware warping for seamless stereoscopic image cloning. *ACM Trans. Graph.*, 31(6):182:1–182:8, 2012. [3](#)
- [17] B. Mendiburu. *3D movie making: stereoscopic digital cinema from script to screen*. CRC Press, 2009. [1](#), [3](#), [7](#)
- [18] Y. Niu, W.-C. Feng, and F. Liu. Enabling warping on stereoscopic images. *ACM Trans. Graph.*, 31(6):183:1–183:7, 2012. [3](#)
- [19] S. Peleg, M. Ben-Ezra, and Y. Pritch. Omnistere: Panoramic stereo imaging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(3):279–290, 2001. [1](#), [2](#)
- [20] P. Pérez, M. Gangnet, and A. Blake. Poisson image editing. *ACM Trans. Graph.*, 22(3):313–318, 2003. [2](#), [3](#)
- [21] C. Richardt, Y. Pritch, H. Zimmer, and A. Sorkine-Hornung. Megastereo: Constructing high-resolution stereo panoramas. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1256–1263, 2013. [1](#), [2](#)
- [22] S. M. Seitz, A. Kalai, and H.-Y. Shum. Omnivergent stereo. *International Journal of Computer Vision*, 48(3):159–172, 2002. [2](#)
- [23] H.-Y. Shum and R. Szeliski. Construction and refinement of panoramic mosaics with global and local alignment. In *IEEE ICCV*, pages 953–956, 1998. [2](#)
- [24] H.-Y. Shum and R. Szeliski. Stereo reconstruction from multiperspective panoramas. In *IEEE International Conference on Computer Vision*, pages 14–21, 1999. [2](#)
- [25] D. Sun, S. Roth, and M. J. Black. Secrets of optical flow estimation and their principles. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2432–2439, 2010. [3](#)
- [26] R. Szeliski. Image alignment and stitching: a tutorial. *Foundations and Trends in Computer Graphics and Vision*, 2(1):1–104, 2006. [1](#), [2](#)
- [27] T. Yan, Z. Huang, R. W. Lau, and Y. Xu. Seamless stitching of stereo images for generating infinite panoramas. In *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology*, pages 251–258, 2013. [2](#)
- [28] J. Zaragoza, T.-J. Chin, M. S. Brown, and D. Suter. As-projective-as-possible image stitching with moving DLT. In *IEEE CVPR*, 2013. [2](#), [3](#), [4](#)
- [29] F. Zhang and F. Liu. Parallax-tolerant image stitching. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 3262–3269, 2014. [2](#), [3](#), [4](#), [6](#), [8](#)