

# Eliminating Structure and Intensity Misalignment in Image Stitching\*

Jiaya Jia<sup>§</sup>    Chi-Keung Tang<sup>†</sup>

<sup>§</sup>The Chinese University of Hong Kong

<sup>†</sup>The Hong Kong University of Science and Technology

## Abstract

*The aim of this paper is to achieve seamless image stitching for eliminating obvious visual artifact caused by severe intensity discrepancy, image distortion and structure misalignment, given that the input images are globally registered. Our approach is based on structure deformation and propagation while maintaining the overall appearance affinity of the result to the input images. This new approach is proven to be effective in solving the above problems, and has found applications in mosaic deghosting, image blending and intensity correction. Our new method consists of the following main processes. First, salient features or structures are robustly detected and aligned along the optimal partitioning boundary between the input images. From these features, we derive sparse deformation vectors to uniformly encode the underlying structure and intensity misalignment. These sparse deformation cues will then be propagated robustly and smoothly into the interior of the target image by solving the associated Laplace equations in the image gradient domain. We present convincing results to show that our method can handle significant structure and intensity misalignment in image stitching.*

## 1 Introduction

Image stitching generates a natural image composite given a set of globally registered images [19, 14] with limited overlapping area. Satisfactory results on image stitching requires a natural transition from one image to another, both in structure and intensity within and possibly beyond the overlapping area. In this paper, we address the problem of image stitching in the presence of severe structure and intensity discrepancy, and propose a novel technique to simultaneously and globally eliminate structure and intensity misalignment among images.

Previous techniques in image stitching [14, 19] optimize for an alignment function that minimizes the intensity difference in the vicinity of the overlapping area. There is, however, no guarantee that image features or structures are aligned after intensity alignment. Recently, non-parametric

patch-based techniques in texture synthesis based on texture deformation have been proposed in [20, 9]. To create a natural texture image of large size, the detected features are matched and deformed in the overlapping area between patches. Therefore, local structure across patch boundaries can be maintained after the synthesis. However, these techniques do not have sufficient ability to handle significant color or intensity inconsistencies in the input images. Moreover, complex global structures and detailed patterns inherent in many natural images may increase the ambiguity in their 2D patch matching process, where only local feature alignment is usually performed.

Our approach is based on 1D feature matching and deformation propagation in natural images. Unlike the previous techniques in image stitching (e.g., [19, 14]) or texture synthesis (e.g., [20, 9]), our method detects and matches the feature points by computing a novel optimal partitioning between the overlapping images. Feature matching along the optimal partition boundary is performed. We uniformly represent the global structure alignment by the feature deformation vectors, which are propagated robustly from the optimal partition toward the interior of the images. Our propagation is performed in the gradient domain, thus the intensity discrepancy among images can be effectively reduced during the deformation propagation.

The rest of our paper is organized as follows. Section 2 reviews related work. Then, in section 3, we present our optimization algorithm in detail. Main results in this paper and comparison with previous methods are presented in section 4. We conclude our paper in section 5.

## 2 Related work

Our image stitching aligns not only image intensity but also image structure. In this section, we review related work in maintaining smooth image transition and deriving structure deformation for image alignment.

### 2.1 Image stitching

Many image registration methods have been developed in recent years. When the input images have significant intensity difference, color blending with the use of a weighting mask over the overlapping area is commonly adopted for generating a smooth intensity transition. For instance,

---

\*This research is supported by the Chinese University of Hong Kong (CUHK) Direct Grant: 2050332, and the Research Grant Council of Hong Kong Special Administration Region, China: AOE/E-01/99.

the video mosaic proposed in [17] uses the Levenberg-Marquardt method to estimate the homography matrix for aligning two overlapping images. To reduce visible artifacts and local misalignment, the method blends the overlapping area by a bilinear weighting function. In [19], the feather-based algorithm was applied which uses averaging and interpolation functions to reduce intensity difference. Unnatural seam, however, is still inevitable since only local operations are performed in these methods. Burt et al. [4] used a multi-resolution spline to perform blending. Note that all these methods only locally blend images in the overlapping area to transit the images taken from one lighting environment to another.

The local alignment method proposed in [18] for performing deghosting works well in many situations. However, it requires the recovery of the true 3D ray directions, which makes it difficult to handle occlusion errors and restricts it to only real images. The color or intensity difference among images may also make the method trap into local minimum.

Methods using optimal seam were recently proposed to composite image or textures [8, 12, 1]. These methods first compute the color difference in the overlapping area between the two input textures or images. Then dynamic programming [8] or graph cut [12] is used to compute an optimal partition that produces the least color difference between the two textures/images. In [7], the partitions among different motions were also computed. However, when there exists large color difference in the overlapping area, satisfactory results cannot be produced.

The two methods described in [10, 11] combine image registration and intensity correction into a single optimization framework. The method in [10] requires the internal camera parameters be known before the optimization, whereas the tensor voting method described in [11] computes intensity alignment and corrects color globally at all the image pixels without any prior knowledge on the camera. However, structure misalignment may be resulted since structure was not explicitly considered in these methods.

Levin et al [14] proposed an image stitching algorithm that operates in the gradient domain, and introduced an optimization method based on the gradient strength in the overlapping area. This method produces good results in the presence of local or global intensity difference between the two input images. However, large structure misalignment cannot be handled by this method.

## 2.2 Structure deformation

Structure deformation and alignment have been topics of interest in medical image registration, especially in non-rigid transform for registering medical images of different modalities. Bajcsy et al. [2] first proposed to use the forces of external stretching and internal smoothness to register

medical images. A multiscale technique using a pyramidal representation is often applied. In [6], a two-step approach was proposed for the nonlinear registration of brain images. In the first step, an one-to-one mapping between corresponding boundaries are established. The second step deforms these boundaries subject to certain criteria. This method requires the corresponding boundaries be homothetical to each other, that is, they are related by a uniform scaling and an arbitrary length-preserving bending. Elastic registration [2] cannot precisely handle detailed and local features, which are common in natural images.

Fluid registration [3, 13] uses the non-rigid method called viscous fluid registration to align medical images. Unlike the elastic models where the desirable deformation may not be obtained because of the internal strain in the elastic continuum, these methods can achieve the desirable deformation since internal forces will disappear over time. However, they may easily introduce blurring and produce non-negligible distortion.

Recently, feature matching methods in texture synthesis were proposed in [20, 9]. In [20], the binary feature maps were first produced by using a two-pass Canny edge filter. Then a 2D feature matching process is applied. This method detects features in the overlapping area in multiple scales. However, their binary feature representation is unsuitable for feature matching in the presence of various feature types, which is common in natural images. These complications may lead to an incorrect warping function. Moreover, since [20] is designed for texture synthesis, it has insufficient ability to correct intensity misalignment. Our method, on the contrary, matches salient features on the partition boundary and avoid considering complex feature junctions. In [9], a deformation function was introduced to simultaneously maximize the color matches while minimizing the deformation distortion over the overlapping area. However, only local minima can be achieved, and the result may be easily affected by the intensity inconsistency between images. Our method addresses the intensity inconsistency by considering structure deformation in the gradient domain, and by smoothly propagating a set of sparse deformation vectors to the interior of images. Hence, our method achieves precise alignment and is free from the problem of breaking continuous structures.

## 3 Our method

For clarity, in this paper, we consider the fundamental case of stitching two images  $I_S$  and  $I_T$ , where the overlapping area is  $\Omega$ . Our technique can be readily generalized if more images are present. Also, we describe one case of structure deformation that deforms (a region of)  $I_T$  to match with  $I_S$  because the other cases are similar. Below,  $\nabla$  represents the gradient operator  $[\frac{\partial}{\partial x}, \frac{\partial}{\partial y}]$ , where  $\nabla_i$ ,  $i \in \{x, y\}$ , is the respective principal gradient components.

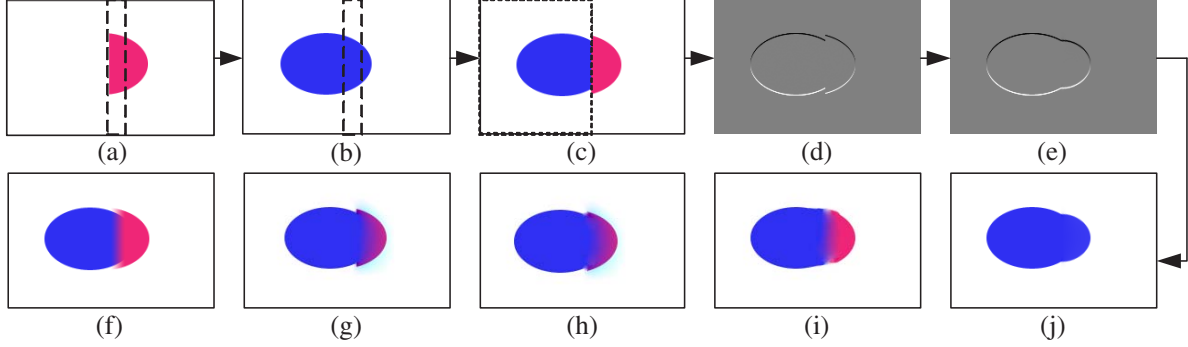


Figure 1: Structure deformation. (a) and (b) are two input images. The overlapping area is indicated by the dash lines. (c) Our method first computes an optimal partition in the gradient domain. (d) The y-component of the gradient map in the gray scale channel after deriving the optimal partition. (e) The features are matched along the optimal partition, and are further propagated inside the image. We show the y-component of the gradient map in the gray scale channel after propagation. (f) Result from feathering. (g) Result obtained using the optimal seam in the gradient domain. (h) Result by the GIST1 method proposed in [14], where artifact is also observed because features are not aligned in the method. (i) Result obtained using the method in [20]. Note that this method is designed for texture synthesis. The warping method cannot produce smooth transition in the presence of significant intensity difference. (j) Our final result. Note that not only features are properly aligned but also colors are globally propagated in the image as well.

Our general image stitching does not have any restriction on the shape of the underlying structures to be aligned. The overview of our method is as follows, where each step will be described in detail in this section.

1. Compute an optimal partition between the aligned images. This step generates a stitching boundary between the overlapping images that minimizes the cost function to be defined.
2. Detect and match features along the stitching boundary generated in the previous step.
3. The matched features are associated with their corresponding deformation vectors. They are propagated from the partition boundary to the interior of the images to maintain smooth transition in structure and intensity.
4. Construct a deformed gradient image based on the deformation vector computed at each pixel and reconstruct the final output image by solving the Poisson equations on the deformed gradient image.

Fig. 1 shows a simple example where the two shapes in (a) and (b) are to be stitched. The result from step 1 is shown in (c) where an optimal partitioning boundary is generated. (d) shows the y-component of the corresponding gradient map in the gray scale channel. Along the partition boundary, the features, namely, the edges, are matched automatically. The deformation corresponding to the matched features is smoothly propagated across the entire image. Fig. 1(e) shows the y-component of the resulting gradient map in the gray scale channel. Our final result is shown in (j), which clearly shows that the features are smoothly connected and the color difference between the two images is globally minimized. Here, we also compare our approach

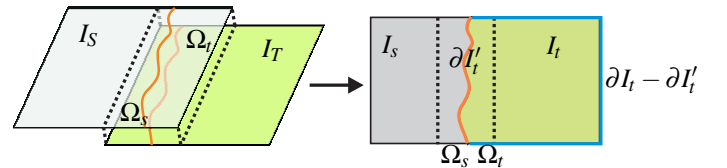


Figure 2: Optimal partition. Left: images  $I_S$  and  $I_T$  are overlapping in  $\Omega$ . The partition divides  $\Omega$  into  $\Omega_S$  and  $\Omega_T$ . Right: the new partitioned images are  $I_S$  and  $I_T$ . We define the intersection  $I_S \cap I_T$  as the optimal partition boundary  $\partial I'_S$ , which is indicated by the orange partition boundary. The blue boundary of  $I_T$  is  $\partial I_T - \partial I'_S$ .

with other techniques. (f) is the feathering result. (g) is the stitching result generated from the optimal seam in the gradient domain. Without explicit feature alignment, visual artifacts resulting from edge discontinuity and color inconsistency are observed. (h) is the result generated from GIST1 [14] which also operates in the image gradient domain. Structure misalignment can also be observed since structure alignment is not considered. (i) is the result produced by the texture deformation method [20] which cannot globally eliminate color inconsistency.

### 3.1 Optimal partition

In the first step, we compute an optimal partition in the overlapping area  $\Omega$  to minimize the structure misalignment. Here, we formulate the partitioning problem as one of labeling, and adopt the graph cuts method to find an optimal solution. We define the gradient alignment cost  $S(p, q)$  between any adjacent pixels  $p$  and  $q$  to be the sum of the com-

puted values in the *rgb* color channels:

$$S(p, q) = \sum_{r,g,b} ((1 - \beta)S_m + \beta S_d), \quad (1)$$

where  $S_m$  and  $S_d$  are two costs measuring the gradient smoothness and similarity between the neighboring pixels, which will be defined shortly.  $\beta$  is a weight used to balance the relative influence of the two costs, which is set to be 0.3 in our experiments. Before the above computation, we smooth both images by using a Gaussian filter with a diameter of 5. Denoting the smoothed output image as  $I^*$ .  $S_m$  is defined as follows:

$$S_m(p, q) = \|\nabla_x I_S^*(p)\| + \|\nabla_x I_S^*(q)\| + \|\nabla_y I_T^*(p)\| + \|\nabla_y I_T^*(q)\|, \quad (2)$$

where  $\|\nabla \cdot\|$  represents the norm of the gradient for each pixel. Thus  $S_m$  takes the gradient smoothness into account, which effectively avoids the partition from breaking object edges in both input images.  $S_d$  is defined as:

$$S_d(p, q) = \|\nabla_x I_S^*(p) - \nabla_x I_T^*(p)\| + \|\nabla_x I_S^*(q) - \nabla_x I_T^*(q)\| + \|\nabla_y I_S^*(p) - \nabla_y I_T^*(p)\| + \|\nabla_y I_S^*(q) - \nabla_y I_T^*(q)\| \quad (3)$$

where each term above represents the gradient-level similarity at the same pixel location in the overlapping area.  $S_d$  thus penalizes pixel dissimilarity in the gradient domain. As a result,  $S_d$  favors the partitioning curve to pass through the pixel locations where the neighboring pixels have similar gradient in the overlapping area.

Therefore,  $S(p, q)$ , combining  $S_m$  and  $S_d$ , enables the graph cuts method to produce a good initial partition with the maximum continuity in the gradient domain. Fig. 2 illustrates that the partition divides  $\Omega$  into  $\Omega_s$  and  $\Omega_t$ , which generates two new regions  $I_s$  and  $I_t$  (colored in gray and green respectively). The intersection  $I_s \cap I_t$  is the optimal partition  $\partial I'_t$ . The set of boundary pixels of  $I_t$  excluding  $\partial I'_t$  are denoted by  $\partial I_t - \partial I'_t$ .  $\partial I'_t$  and  $\partial I_t - \partial I'_t$  are indicated as orange and blue curves respectively on the right side of Fig. 2.

Comparing to the optimal seam methods proposed in [8, 12], our new cost function takes into account both the gradient smoothness and gradient similarity. Therefore, even the two input images have significant difference in intensity, our method can still optimize the gradient continuity. Moreover, our partition favors smooth area in both images, which effectively reduces the feature complexity along  $\partial I'_t$ , which will in turn reduce the matching ambiguity in the following steps. The comparisons of the partitioning results between the previous optimal seam methods and ours will be shown in section 4.

### 3.2 Feature detection

The optimal partition obtained above is subject to the continuity constraint given by  $S(\cdot)$ . However, structure discontinuity may still be inevitable in some cases, such as the

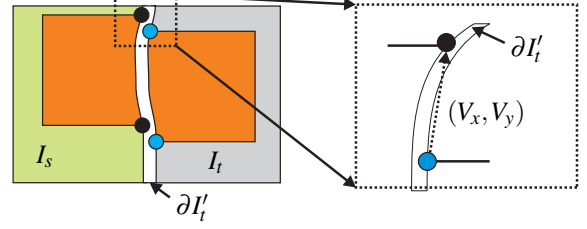


Figure 3: After optimal partitioning, there may still exist structure misalignment. Along the partition boundary, similar features (black and blue dots) in the two images should be matched. Hence, in the zoom-in view, features in  $I_t$  should be deformed in the direction  $(V_x, V_y)$  as indicated by the dashed arrow.

one illustrated in Fig 1(c). In this section, we propose a method based on feature matching to correct the misalignment along the optimal partition boundary.

Our observation is that in most situations if the structure is not aligned during the optimal partitioning, the features may have been broken at different locations along their respective stitching boundaries. Accordingly, a deformation vector can be defined, as shown in Fig. 3. For clarity of depiction, in the following, we describe our structure deformation method by matching features from  $I_t$  to  $I_s$ . Hence, only the partition boundary  $\partial I'_t$  will be considered. The analogous problems that  $I_s$  is deformed to match  $I_t$ , or both  $I_s$  and  $I_t$  are deformed to match their mean respective feature points, are similar.

**Feature localization** The human eye is more sensitive to edge discontinuity than to smooth color transition. Thus, in image stitching, structure misalignment present in a road, building, or the horizon will look more obvious to us than the intensity misalignment present in a large grassland, for instance.

Accordingly, we apply a Gaussian smoothing process, together with the anisotropic filter proposed in [16], on  $I_s$  and  $I_t$  along  $\partial I'_t$  respectively. We use two rotated kernels  $F_\theta^e$  and  $F_\theta^o$  defined in [16] to construct the square modulus  $W(x, y, \theta) = (F_\theta^e * I)^2 + (F_\theta^o * I)^2$ . The angle with the maximum  $W$  indicates the edge orientation at pixel  $(x, y)$ . Comparing with Canny edge detection [5], this method is more robust and precise for feature localization and detection, especially when junctions are present. In our experiments, we found that a sparse sampling of  $\theta$  in 10-degree increment is sufficient in all cases.

**Feature representation** We denote all features detected in image  $I_s$  and  $I_t$  along  $\partial I'_t$  (orange curve on the right of Fig. 2) as  $\{f_s^1, f_s^2, \dots, f_s^m\}$  and  $\{f_t^1, f_t^2, \dots, f_t^n\}$  respectively, where  $m$  and  $n$  are the total number of features along  $\partial I'_t$  in  $I_s$  and  $I_t$ . Note that the features are ordered along  $\partial I'_t$ , and each feature point  $f_s^k$  (or  $f_t^k$ ) is associated with a feature direction  $\theta_s(k)$  (or  $\theta_t(k)$ ) and the computed energy  $W_s(k)$

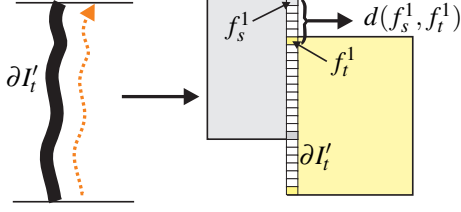


Figure 4: The partition boundary  $\partial I'_t$  (left) can be straightened into a line segment (right) by considering the relative distance of the sparse features along  $\partial I'_t$ . The path distance  $d(\cdot)$  between two features in  $I_s$  and  $I_t$  is measured by their distance in the line segment.

(or  $W_t(k)$ ).

The reason that only features at the boundary pixels are computed is that without taking into account all pixels in the overlapping area, it is unnecessary to consider complex structures such as T-junctions and W-junctions. The optimal partitioning introduced in the previous section effectively reduces the number of feature points by using a smoothness constraint.

### 3.3 Feature matching

We construct a boundary line segment as shown in Fig. 4. All pixels along  $\partial I'_t$  will be placed in it with a sequence according to their respective distances to the starting pixel of  $\partial I'_t$ , shown as the orange arrow in Fig. 4. We define  $d(f_t^{k_1}, f_s^{k_2})$  between any two features  $f_t^{k_1}$  and  $f_s^{k_2}$  as the distance along the 1D line. One example is shown on the right of Fig. 4.

We measure the similarity of two features  $f_t^{k_1}$  and  $f_s^{k_2}$  in  $I_s$  and  $I_t$  by considering the computed energy, feature angles and their path distances  $d(\cdot)$ :

$$e(f_t^{k_1}, f_s^{k_2}) = \alpha(W_t(k_1) - W_s(k_2))^2 + \gamma(\theta_t(k_1) - \theta_s(k_2))^2 + (1 - \alpha - \gamma)d(f_t^{k_1}, f_s^{k_2})^2, \quad (4)$$

where  $(W_t(k_1) - W_s(k_2))^2$  measures the energy distance between the two features, so a strong edge will not be matched with a relatively smooth region.  $(\theta_t(k_1) - \theta_s(k_2))^2$  constrains that the matched features should not differ too much in their directions.  $d(\cdot)$ , on the other hand, constrains that the two features to be matched should not be too far away from each other. The three terms above are normalized before the actual calculation. We set  $\alpha$  and  $\gamma$  as 0.4 and 0.2 in our experiments.

Given the feature measurement  $e(f_t^{k_1}, f_s^{k_2})$  between each feature pair in  $I_s$  and  $I_t$ , we enqueue all the feature points in  $I_t$  and apply the following algorithm to construct feature matches, by considering the case that  $I_t$  is deformed to align with  $I_s$ :

1. Select the feature  $f_t^{k_1}$  in the remaining queue with the largest energy  $W_t(k_1)$ , which represents the strongest feature point.
2. We search for a feature point  $f_s^{k_2}$  with the smallest similarity cost  $e(f_t^{k_1}, f_s^{k_2})$ . If the match does not violate the *containment rule* (below), these two points will be matched and  $f_t^{k_1}$  will be removed from the queue.

We define the containment rule as follows:

*Suppose  $k'_1 \leq k_1$ ,  $f_t^{k'_1}$  is matched to  $f_s^{k'_2}$  and  $f_t^{k_1}$  is matched to  $f_s^{k_2}$ . Then we must have  $k'_2 \geq k_2$ , and vice versa.* This constrains that no edge intersection will be produced by structure deformation to prevent unnecessary structure distortion.

3. If the queue is not empty, goto step 1 and continue the algorithm.

This algorithm agrees with our human perception that strong structures should be aligned properly and not be fragmented, as indicated by the fact that the strongest feature points are always picked first.

For each feature matching pair  $(f_t^{k_1}, f_s^{k_2})$ , we construct a *deformation vector*  $\mathbf{V}(f_t^{k_1}) = \{V_x(f_t^{k_1}), V_y(f_t^{k_1}), V_{\nabla_x}(f_t^{k_1}), V_{\nabla_y}(f_t^{k_1})\}$ .  $V_x$  and  $V_y$  are the  $x$  and  $y$  components of the vector pointing from  $f_t^{k_1}$  to  $f_s^{k_2}$  in the image space, as shown on the right of Fig. 3.  $V_{\nabla_x}$  and  $V_{\nabla_y}$  measure the difference in the gradient map between  $(f_t^{k_1}$  and  $f_s^{k_2})$ , and is calculated as:

$$V_{\nabla_i}(f_t^{k_1}) = \nabla_i I_t(f_t^{k_1}) - \nabla_i I_s(f_s^{k_2}), i \in \{x, y\}. \quad (5)$$

Therefore, our structure deformation not only considers the geometric difference in the image plane, but also the gradient difference in feature alignment.

### 3.4 Deformation propagation

For clarity, similar to section 3.2, we only describe the deformation from  $I_t$  to  $I_s$ . The generalization to handle other cases is similar. To smoothly propagate deformation from the sparse features along  $\partial I'_t$  into image  $I_t$ , we define the deformation area  $S_t$  to indicate the area that should be affected by the propagated deformation. By default, we allow the deformation to propagate into the entire image  $I_t$ , namely,  $S_t = I_t$ . However, it is often adequate to apply local structure deformation in the overlapping area, so another good choice is  $S_t = \Omega_t$ . In fact,  $S_t$  can be arbitrarily defined by the user to derive various image stitching effects.

In our boundary configuration shown in Fig. 5, the yellow region is  $S_t$ , within which deformation is performed. The blue ribbon is the boundary of  $S_t$ , that is,  $\partial S_t$ . The green region represents  $I_t - S_t$ . Since it is outside of  $S_t$ , this region should not be affected by any deformation.



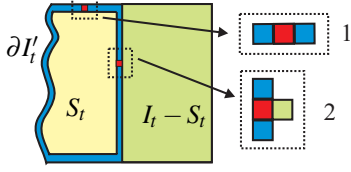


Figure 5: Boundary condition. In  $I_t$ , if a pixel  $p_i$  is not lying on  $\partial I_t'$ , and one of its neighborhood pixel is in  $I_t - S_t$ , we set the value as  $\mathbf{0}$ , as shown in block 2. Otherwise, if the pixel is not a feature point, its value is undetermined, as shown in block 1.

We consider the deformation vector assignment for a pixel  $p_i \in \partial S_t$  under the following three conditions. 1)  $p_i$  is a feature point. The corresponding deformation vector is already computed in the previous section. 2)  $p_i$  is adjacent to the pixels in  $I_t - S_t$ , e.g., the red pixel in box 2 in Fig. 5. To avoid producing structure discontinuity between  $S_t$  and  $I_t - S_t$ , the deformation of these pixels should be zero so that  $S_t$  and  $I_t - S_t$  can be aligned structurally. 3)  $p_i$  is not a feature point, and also not adjacent to  $I_t - S_t$ . In this case,  $p_i$ 's deformation is undetermined, as the red pixel of block 1 in Fig. 5. Accordingly, we define the transformation vector as follows:

$$\mathbf{V}(p_i) = \begin{cases} \mathbf{V}(f_t^k) & p_i = f_t^k \\ \mathbf{0} & \exists p_j, p_j \in N_{p_i} \& p_j \in I_t - S_t \end{cases} \quad (6)$$

where  $N_{p_i}$  is the 4-connected neighborhood of  $p_i$  and  $\mathbf{0} = \{0, 0, 0, 0\}$ . For all other boundary pixels not defined in Eqn. 6, we use the Neumann boundary condition to measure their deformation, that is,  $\nabla \mathbf{V}(p_i) = \mathbf{0}$ .

Therefore, to smoothly propagate the deformation inside image  $I_t$ , we propose to solve the minimization equation:

$$\mathbf{V}^* = \arg \min_{\mathbf{V}} \int \int_{p \in S_t} \|\nabla \mathbf{V}\|^2 dp \quad (7)$$

which is equivalent to solving the Laplace's equation:

$$\Delta \mathbf{V} = 0 \text{ with } \mathbf{V}|_{\partial S_t} = \mathbf{V}^*|_{\partial S_t} \quad (8)$$

where  $\Delta$  is the Laplacian operator. We solve the equation by using the Gauss-Seidel iterative method with a multi-grid under the boundary conditions, for each component in  $\mathbf{V}$  respectively. After the minimization, each pixel in  $S_t$  is associated with a deformation vector. Finally, using the deformation vectors at all the pixels in  $S_t$ , we perform an inverse mapping with bilinear interpolation in the gradient domain in  $S_t$  to construct the deformed gradient map (Fig. 1(e)). The final image is obtained by solving the Poisson equations on the deformed gradient map.

## 4 Results

We demonstrate our method using three difficult examples. Comparison with other methods using our implementation are also given.

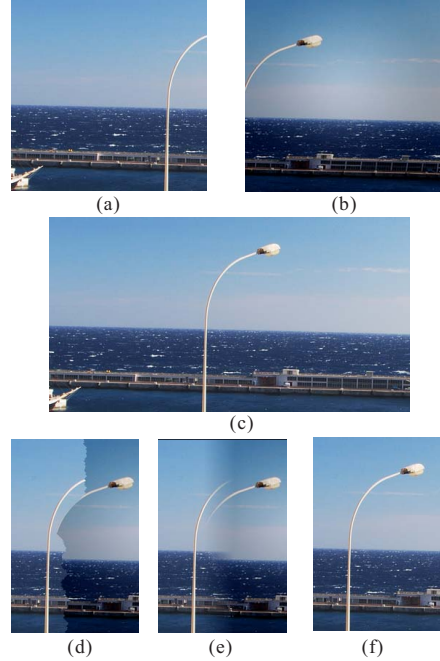


Figure 6: Lamp post. (a) and (b) are the two registered images. (c) is our image stitching result where the lamp post is seamlessly aligned. (d) and (e) are zoom-in views of the results generated by: (d) the optimal seam method and (e) feathering. (f) Our result with  $S_t = I_t$ , where structures are globally aligned with minimum distortion.

Fig. 6 shows one example where the two input images in (a) and (b) have large discrepancy both in color and structures. (c) shows our generated result. (d) is the result generated by the optimal seam method. (e) is the feathering result. (i) shows our result. Our method can handle color and structure deformation within the same framework, which can successfully solve the deformation problem for structure and intensity alignment between the two images.

In Fig. 7 (a) and (b), we show two images taken from a sunset scene. The alignment process is difficult because of the local displacement and the small overlapping areas, shown in the green boxes in the figure. Our result is shown in (c). Note that the structures are globally aligned and the intensities are also matched. From (d) to (k), we compare the results generated using previous methods. (d) is the result generated by the optimal seam method [8]. Since no pixels are similar in intensity, the result produces an obvious seam (indicated by the thin red curve). (e) is the result obtained by the optimal seam method operated in the gradient domain. Color difference is alleviated in the stitched result. Our method also performs optimal partitioning as described in section 3.1. Taking structure smoothness into consideration which makes the matching process more robust, our method will not destroy the boat as shown in (f). has also been significantly reduced. (g) is the feathering result where the ghosting effect is obvious. (h) shows the re-

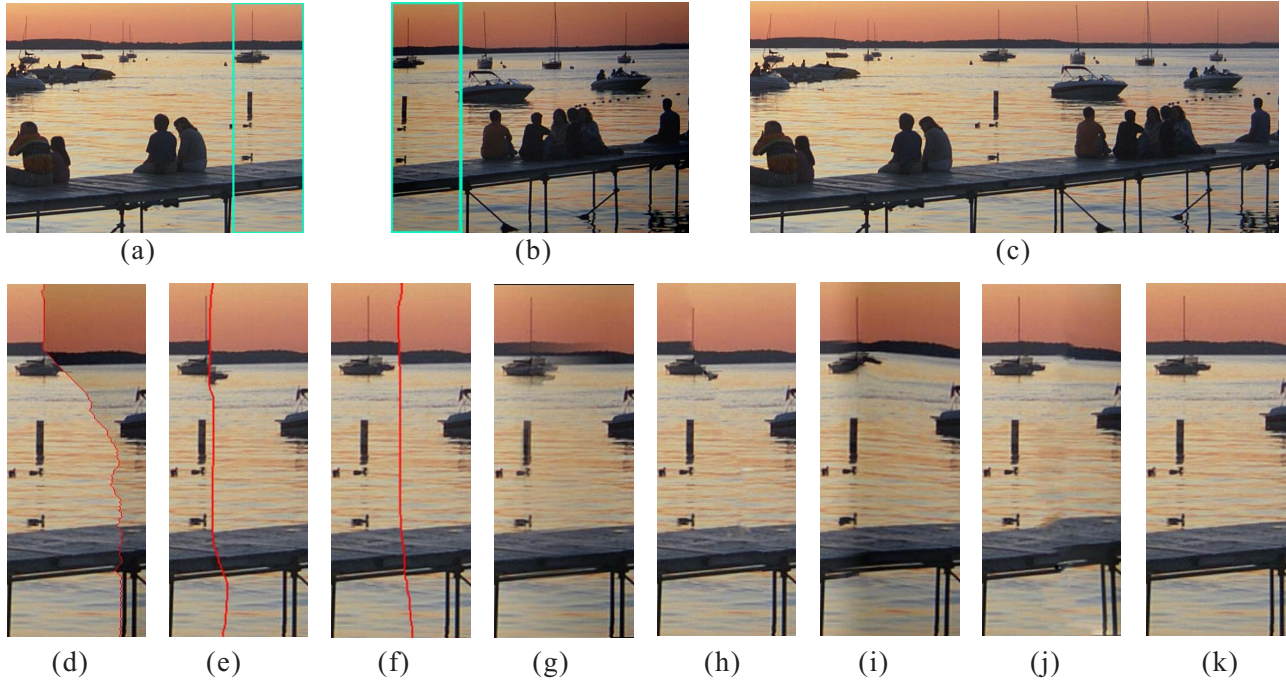


Figure 7: Sunset. (a) and (b) are the two registered images. The green boxes indicate the overlapping areas. (c) is our image stitching result, where the intensity and structures are globally aligned. (d)-(k) are zoom-in views of the results generated by: (d) The optimal seam method. The seam is highlighted as red. (e) The optimal seam method operated in the gradient domain. (f) Our initial optimal partition. Since smoothness is considered, the boat is not broken. (g) Feathering. The ghosting effect is evident. (h) GIST1 [14]. The structure misalignment cannot be eliminated. (i) Structure deformation by [20]. The 2D feature matching process cannot adequately warp the features in this complex case, such as the boat, thus causing the visual artifact. (j) Feature matching by [9]. The method is susceptible to local minima when the two images have different intensities. (k) Our result with  $S_t = I_t$ .

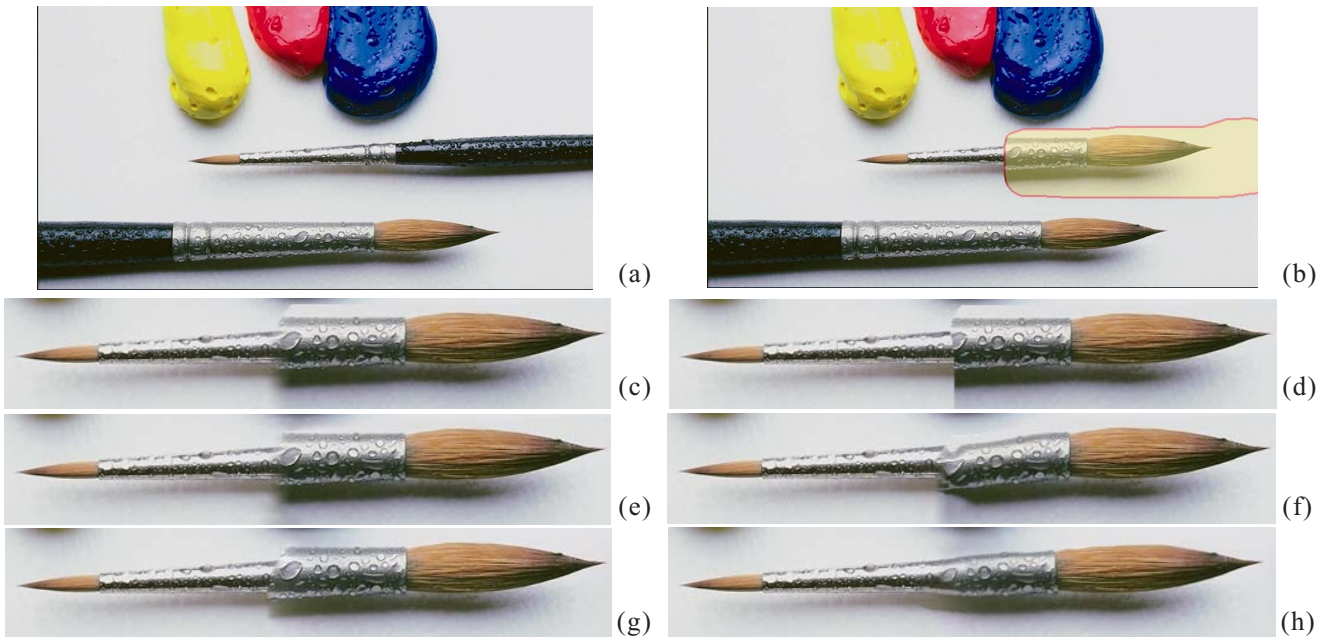


Figure 8: Brush. (a) Input image. (b) The lower brush is copied and pasted onto the upper one, as shown inside the transparent yellow region. (c) Feathering result. (d) Optimal seam result. (e) GIST1 [14] result. Misalignment is still obvious. (f) Result of structure deformation by [20]. The extracted feature has complex shapes and scales which easily confuse the warping process. (g) Result by direct Poisson blending [15]. (h) Our result. The complexity in feature matching is reduced to 1D, allowing for the sparse deformation vectors to robustly propagate into the interior of the image to enforce the necessary structure smoothness.

sult by GIST1 [14], where the imposed smoothness cannot eliminate the misalignment in structure. (i) is the structure deformation result obtained by [20]. 2D matching of the extracted features cannot adequately handle complex structures, such as the boat, and causes the incorrect warping result. The local distortion is also inevitable. (j) is the feature matching result by [9]. The significant difference in global image intensity makes their optimization method susceptible to local minima and thus causes inaccurate alignment. Starting from the optimal partition, our result shown in (k) is obtained by our local feature matching, which does not need to consider any complex shapes in the 2D space. Moreover, the necessary structure deformation is propagated properly and smoothly toward the interior of the images, which guarantees that no visual features will be broken while structure alignment is enforced during the propagation. In this example, we set  $S_t = I_t$  to globally reduce the distortion effort in the deformation process.

Our method can be readily applied to unconventional image stitching with arbitrary overlapping areas. In Fig. 8(a), we only use a single input image, where the lower brush is copied to align with the upper one. The user draws the mask as shown in the yellow region in (b). The feathering result is shown in (c), where misalignment is obvious. (d) shows the result obtained by the optimal seam method. The corresponding pixels along the boundary are mostly different in terms of intensity, thus resulting in the obvious seam as shown. (e) shows the result by the GIST1 method [14], which cannot eliminate misalignment in this example. (f) shows the result obtained using the deformation method in [20]. Since this example has complex features at multiple scales and in various scales, their warping function mixes up the interweaved features and causes significant alignment error. Shown in (g) is the result produced by direct Poisson blending [15]. Our result is shown in (h), where structures and intensities are aligned along the stitching boundary, and are smoothly propagated into the interior of the pasted region.

## 5 Conclusion

In this paper, we propose a novel approach of image deformation for image stitching where the overlapping regions may contain significant intensity inconsistency and structure misalignment. Instead of searching all possible pairs of features in the given images to derive the optimal pair for alignment, we propose to detect only the necessary features along the partitioning boundaries, and use them to construct the deformation vectors. By doing so, we can avoid the complex problem arisen in multiscale feature detection and matching, especially when the input images have different levels of details. From the sparse features detected along the partitioning boundary, we propagate the deformation into the target image smoothly and robustly by solving

the associated Laplace equation. Structure deformation and color correction are simultaneously achieved within the same framework operating in the image gradient domain. Hence, our method can uniformly handle difficult scenes with severe structure and intensity misalignment among images, as shown in the result section.

## References

- [1] Aseem Agarwala, Mira Dontcheva, Maneesh Agrawala, Steven Drucker, Alex Colburn, Brian Curless, David Salesin, and Michael Cohen. Interactive digital photomontage. In *Proc. of SIGGRAPH 2004*, 2004.
- [2] R. Bajcsy and S. Kovacic. Multiresolution elastic matching. *Computer Vision, Graphics and Image Processing*, (46):1–21, 1989.
- [3] Morten Bro-Nielsen and Claus Gramkow. Fast fluid registration of medical images. *Proceedings of Visualization in Biomedical Computing*, (1131):267–276, 1996.
- [4] P. J. Burt and E. H. Adelson. A multiresolution spline with applications to image mosaics. In *ACM Transaction on Graphics*, pages 2(4), 217–236, October 1983.
- [5] J. Canny. A computational approach to edge detection. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1986.
- [6] C. Davatzikos, J.L. Prince, and R.N. Bryan. Image registration based on boundary mapping. *IEEE Transactions on Medical Imaging*, (15):112–115, 1996.
- [7] James Davis. Mosaics of scenes with moving objects. In *CVPR 1998*, 1998.
- [8] Alexei A. Efros and William T. Freeman. Image quilting for texture synthesis and transfer. *Proceedings of SIGGRAPH 2001*, pages 341–346, August 2001.
- [9] Hui Fang and John C. Hart. Textureshop: Texture synthesis as a photograph editing tool. *Proceedings of SIGGRAPH 2004*, 2004.
- [10] D. Hasler and S. Susstrunk. Colour handling in panoramic photography. In *Proc. of SPIE*, Jan 2001.
- [11] J. Jia and C.-K. Tang. Image registration with global and local luminance alignment. In *ICCV03*, pages I: 156–163, 2003.
- [12] Vivek Kwatra, Arno Schodl, Irfan Essa, Greg Turk, and Aaron Bobick. Graphcut textures: Image and video synthesis using graph cuts. *ACM Transactions on Graphics, SIGGRAPH 2003*, 22(3):277–286, July 2003.
- [13] H. Lester and S.R. Arridge. Summarising fluid registration by thin-plate spline warps with many landmarks. *Proceedings of Medical Image Understanding and Analysis*, 1997.
- [14] Anat Levin, Assaf Zomet, Shmuel Peleg, and Yair Weiss. Seamless image stitching in the gradient domain. In *ECCV04*, May 2004.
- [15] Patrick Pérez, Michel Gangnet, and Andrew Blake. Poisson image editing. *ACM Transactions on Graphics (SIGGRAPH'03)*, 22(3):313–318, 2003.
- [16] P. Perona and J. Malik. Detecting and localizing edges composed of steps, peaks and roofs. *Proceedings of ICCV 1990*, 1990.
- [17] R. Szeliski. Video mosaics for virtual environments. In *IEEE Computer Graphics and Applications*, pages 22–30, March 1996.
- [18] Richard Szeliski and Heung-Yeung Shum. Construction of panoramic image mosaics with global and local alignment. *International Journal of Computer Vision*, 36(2):101–130, 2000.
- [19] M. Uyttendaele, A. Eden, and R. Szeliski. Eliminating ghosting and exposure artifacts in image mosaics. In *CVPR 2001*, 2001.
- [20] Qing Wu and Yizhou Yu. Feature matching and deformation for texture synthesis. *Proceedings of SIGGRAPH 2004*, 23(3):362–365, 2004.