A UNIFIED GRADIENT DOMAIN METHOD FOR SEAMLESS IMAGE PROCESSING

Shiming Ge[1,2]*, Kongqiao Wang*[2]*, Hao Wang*[2] *, and Feng Zhao*[1,2]

[1]School of Information Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China
[2]Nokia Research Center, Beijing 100176, China
{ext-shiming.1.ge|kongqiao.wang|hao.ui.wang|ext-feng.1.zhao}@nokia.com

ABSTRACT

Seamless image processing concerns stitching parts of images in a visually natural manner. In this paper, we present a seamless processing method which unifies Featuring, Optimal seam method [1] and some recent gradient domain methods including Poisson image editing [2], Drag-and-drop pasting [3] and GIST [4]. To evaluate the processing quality visually, a variational energy function is proposed to compute both the similarity of the stitched image to each of the input images and the visibility of the seam between the stitched images. The minimum of the energy function gives a globally consistent composition in both geometrical and photometric structures. Based on 3 propositions, we study the energy function and compare it with other methods theoretically. The experimental results show the benefits of our stitching method.

Index Terms—Image composition, image stitching, seam elimination, gradient domain method, total variation

1. INTRODUCTION

Seamless image processing has been widely recognized as a common practice in various computer graphics and image processing applications such as image mosaics [4][5][6], image composition or editing [2][3], texture synthesis [1][7], scene completion [8], etc. The aim of a seamless processing algorithm is to eliminate seams by smoothing the transition between images while retaining the similarity of the resulting mosaics to each of the input images both geometrically and photometrically.

Feathering [5], weighted averaging with the use of a distance map, is commonly adopted for generating a smooth transition. However, feathering may cause blurring of image edges due to smoothing out high-frequency exposure variations. One solution to this problem is multi-resolution blending, such as Laplacian pyramid blending [6]. Note that both spatial domain blending and multi-resolution blending only locally blend images within the overlapping area to smooth the images taken from different environments.

Methods using optimal seam were proposed to composite images or textures [1][7]. These methods search for a curve on which the color difference between the two overlapped images/textures is minimal. However, the seam is visible when there exists large color differences in the overlapping area so that no such curve can be found.

Recently, gradient domain methods have been used to compose image seamlessly. The idea is to minimize the gradient difference between the resulting image and each of the input images. Perez et al. [2] proposed Poisson image editing to do seamless object insertion. The actual pixel values for the copied region are computed by solving Poisson equations that locally matches the gradients while obeying the fixed Dirichlet condition at the seam boundary. To make this idea more practical and easy to use, Jia et al. [3] proposed a cost function to compute an optimized boundary condition, which can reduce blurring. Zomet et al. [4] proposed an image stitching algorithm by optimizing gradient strength in the overlapping area. This method produces good results in the presence of local or global color difference between the two input images.

In this paper, we present a gradient domain-based image stitching method. In our method, an energy function is defined by introducing a guidance gradient field into the general total variation (TV) model. A good stitching will minimize the energy function, eliminate photometric inconsistencies, and correct geometric misalignments. We study the energy function and compare it with some existing method theoretically based on three propositions.

2. OUR GRADIENT DOMAIN METHOD

2.1. Variational energy function

Due to the advantage of bounded total variation in the estimation of discontinuities, Rudin conjectured that TV norm is more proper to model images than L2 norm [9]. The general TV model for an image I is expressed through its TV norm:

$$\| I \|_{TV} = \int_{\Psi} \| \nabla I(p) \| \, dp \qquad (1)$$

where, p is the pixel, $\nabla = [\partial/\partial x, \partial/\partial y]$ is the gradient operator, $\|\bullet\|$ denotes L2 norm, and Ψ is the image domain to be processed.

Directly reconstructing I via minimizing general TV energy function (1) produces an unsatisfactory, blurred interpolation [10]. To overcome this problem, we introduce further constraints in the form of a guided gradient field $\mathbf{V}$, which gives an extended version of (1):

$$E(I) = \int_{\Psi} \| \nabla I(p) - \mathbf{V}(p) \| \, dp \tag{2}$$

As it is enough to solve the problem for each color channels separately, we consider only scalar image functions. Assume I_i and Ψ_i ($i = 1,2,\cdots,n$) be the input aligned image and its corresponding image region, respectively. Then the entire image region is $\Psi = \cup \Psi_i$. Let $\mathbb{W} = \{w_i, i = 1,2,\cdots,n\}$ be a collection of the weighting masks all defined over Ψ. Given two images I and J, denote

$$e(w, I, J, p) = w(p) \| \nabla I(p) - \nabla J(p) \| \tag{3}$$

Then, discretizing Eq.(2), the resulting image I is defined as the argument that minimizes the variational energy function $E(I, \mathbb{W})$ which cumulates the weighted differences:

$$E(I, \mathbb{W}) = \sum_{i=1}^{n} \sum_{p \in \Psi_i} e(w_i, I, I_i, p), \text{ with } \sum_{i=1}^{n} w_i(p) = 1 \tag{4}$$

where, $\mathbf{V}$ is derived directly from the source images and is computed with backward difference for the non-overlapping region, thus $w_i(p)$ is 1 for $p \in \Psi \setminus \bigcup_{j \neq i} \Psi_j$ and 0 for $p \in \bigcup_{j \neq i} \Psi_j$.

The dissimilarity $E(I, \mathbb{W})$ between the resulting image and the input images is defined by a weighted distance between their derivatives (gradients). On the overlapping area, the energy function $E(I, \mathbb{W})$ penalizes for derivatives which are inconsistent with any of the input images. In the locations where all I_i have low gradients, $E(I, \mathbb{W})$ penalizes for high gradient values in the resulting image. This property is good for eliminating false stitching edges.

2.2. Numerical Implementation

The Euler-Lagrange equation for Eq.(4) is

$$\sum_{i \in J_O} w_i(O) \nabla \bullet \frac{\nabla I(O) - \nabla I_i(O)}{\| \nabla I(O) - \nabla I_i(O) \|} = 0, \text{ for } O \in \bigcap_{i \in J_O} \Psi_i \tag{5}$$

where $\nabla \bullet$ is divergence operator, J_O denotes the index set of regions which a target pixel O belongs to. Similar to [10], we use mid-point central differencing to approximate Eq.(5). As shown in Figure 1, at a given target pixel O, let E, N, W, S denote its four adjacent available points, and e, n, w, s be the corresponding four midpoints (not directly available). Write $\Lambda_O = \{E, N, W, S\}$ and $\mathrm{V}_O = \{e, n, w, s\}$. Adopting the Gauss-Jacobi iteration scheme at each step t:

$$I^{(t)}(O) = \frac{\sum_{Q \in \Lambda_O} \sum_{i \in J_O} w_i(Q) \left(I^{(t-1)}(Q) - I_i(Q) + I_i(O) \right)}{\sum_{Q \in \Lambda_O} \sum_{i \in J_O} w_i(Q)} \tag{6}$$

$$w_i(Q) = w_i(O) / \| \nabla I(O) - \nabla I_i(q) \|, \quad Q \in \Lambda_O, q \in \mathrm{V}_O \tag{7}$$

where, the gradient in midpoint q in Eq.(7) is approximated with its adjacent available points, such as $\nabla I_i(s) = \left(I_i(S) - I_i(O), [I_i(SE) - I_i(SW) + I_i(E) - I_i(W)]/4 \right)$.

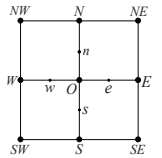

Fig.1. Neighbors

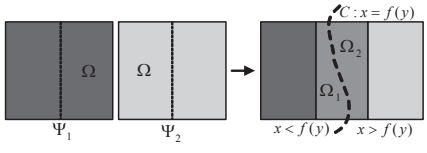

Fig.2. Image stitching

3. THEORETICAL ANALYSIS

As shown in Figure 2, we consider the fundamental case of stitching two aligned images I_1 and I_2 overlapping with Ω. Our technique can be readily generalized if more images are present. Denote $w = w_1$ and $\mathbb{W} = \{w, 1-w\}$, then

$$E(I, \mathbb{W}) = \sum_{p \in \Psi_1} e(w, I, I_1, p) + \sum_{p \in \Psi_2} e(1-w, I, I_2, p) \tag{8}$$

In the following, we study Eq.(8) and compare it with some stitching methods based on three propositions.

3.1. Comparison with feathering

Proposition 1 shows an equivalence between our method and Feathering of image derivatives (gradients).

***Proposition 1*:** Let w be a feathering mask defined over the entire image, without loss of generality, with $w(p) = 1$ for $p \in \Psi_1 \setminus \Omega$, $w(p) = 0$ for $p \in \Psi_2 \setminus \Omega$, and $w(p) = 0.5$ for $p \in \Omega$. A global minimum of $E(I, \mathbb{W})$ defined by Eq.(8) can reach, with respect to the weighting gradient field $\mathbf{V}$:

$$\mathbf{V}(p) = w(p) \nabla I_1(p) + (1 - w(p)) \nabla I_2(p) \tag{9}$$

***Proof*:** When the gradient field for the resulting image is $\nabla I = \mathbf{V}$, the value of $E(I, \mathbb{W})$ on the image I is:

$$\begin{aligned} E(I, \mathbb{W}) = E(\mathbf{V}) &= \sum_{p \in \Psi_1} e(w(1-w), I_2, I_1, p) + \sum_{p \in \Psi_2} e(w(1-w), I_1, I_2, p) \\ &= 0.5 \sum_{p \in \Omega} \| \nabla I_1(p) - \nabla I_2(p) \| \end{aligned}$$

Let $\hat{\mathbf{V}}$ be another gradient field with $\nabla \hat{I} = \hat{\mathbf{V}}$, then

$$\begin{aligned} E(\hat{\mathbf{V}}) &\geq \sum_{p \in \Omega} [e(w, \hat{I}, I_1, p) + e(1-w, \hat{I}, I_2, p)] \\ &= \sum_{p \in \Omega} [e(0.5, \hat{I}, I_1, p) + e(0.5, \hat{I}, I_2, p)] \\ &\geq 0.5 \sum_{p \in \Omega} \| \nabla I_1(p) - \nabla I_2(p) \| = E(\mathbf{V}) \end{aligned}$$

where the triangle inequation for L2 norm is used. Hence the minimum is obtained using the guidance field $\mathbf{V}$.

The resulting mosaic generated by Feathering of images is affected by the size of overlapping area. When the size is too small, the transition between the input images may be unnatural, especially when the color difference is large. Our method tends to mix image gradients, and can smooth out the small structure mismatch.

3.2. Comparison with optimal seam method

In proposition 2 we prove that our method is as good as the optimal seam method [1] when a perfect seam (in which the color difference is zero) exists.

***Proposition 2*:** Assume there is a curve $C: x = f(y)$, such that for each $p \in \{(f(y), y)\}$, $I_1(p) = I_2(p)$. As shown in Figure 2, the curve partitions the mosaics region (also overlapping area) into two regions, $\Psi_1 \setminus \Omega_2$ and $\Psi_2 \setminus \Omega_1$ where $x < f(y)$ and $x \geq f(y)$, respectively. Let w be a weighting mask defined as $0.5 \leq w(p) \leq 1$ for all $p \in \Psi_1 \setminus \Omega_2$, and $0 \leq w(p) \leq 0.5$ for all $p \in \Psi_2 \setminus \Omega_1$. Then the optimal seam solution I is a global minimum of $E(I, \mathbb{W})$ in Eq.(8).

$$I(x,y) = \begin{cases} I_2(x,y), & x < f(y) \\ I_1(x,y), & x \geq f(y) \end{cases} \tag{10}$$

***Proof*:** Note that when I is defined by Eq.(10), we have $\nabla I = \nabla I_1$ for all $p \in \Psi_1 \setminus \Omega_2$, and $\nabla I = \nabla I_2$ for all $p \in \Psi_2 \setminus \Omega_1$. Then the value of the energy $E(I, \mathbb{W})$ is:

$$E(I, \mathbb{W}) = \sum_{p \in \Omega_2} e(w, I_2, I_1, p) + \sum_{p \in \Omega_1} e(1-w, I_1, I_2, p)$$

Let $\hat{I}$ be another solution, then

$$\begin{aligned} E(\hat{I}, \mathbb{W}) &\geq \sum_{p \in \Omega} e(w, \hat{I}, I_1, p) + \sum_{p \in \Omega} e(1-w, \hat{I}, I_2, p) \\ &\geq \sum_{p \in \Omega_2} (e(w, \hat{I}, I_1, p) + e(w, \hat{I}, I_2, p)) \\ &\quad + \sum_{p \in \Omega_1} (e(1-w, \hat{I}, I_1, p) + e(1-w, \hat{I}, I_2, p)) \\ &\geq \sum_{p \in \Omega_2} e(w, I_2, I_1, p) + \sum_{p \in \Omega_1} e(1-w, I_1, I_2, p) = E(I, w) \end{aligned}$$

Thus, I defined by Eq.(10) is a global minimum of $E(I, \mathbb{W})$.

If the optimal seam methods cannot find an ideal seam on the overlapping area, artifacts may still exist because the optimal seam methods do not modify the color values of input images. In contrast, our method modifies the color values by interpolating the mosaics using a guidance vector field, which produces an intensity consistent mosaic.

3.3. Comparison with other gradient domain methods

Besides, we compare our method with some recent gradient domain methods, including Poisson image editing [2], Drag-and-drop pasting [3] and GIST [4].

The energy in Poisson image editing and Eq.(2) can be written in a unified form:

$$E(I) = \sum_{p \in \Psi} \| \nabla I(p) - \mathbf{V}(p) \|^{\alpha} \tag{11}$$

where power α is $\alpha = 2$ for Poisson image editing and $\alpha = 1$ for our model. If the region Ψ contains step edges where $\mathbf{V}$ is a 1-D delta function δ, to handle step edges requires Eq.(11) to be finite. This implies that if $\sum \delta^{\alpha} < \infty$ for α, then $\alpha \leq 1$. To ensure convexity, $\alpha = 1$ is the ideal choice, which leads to TV estimation.

Poisson image editing may produce unnatural blurring when salient structures in source and target images do not match. To tackle this problem practically, Drag-and-drop pasting [3] performs Poisson editing following by finding an optimal boundary. However, when the perfect boundary can not be found, misalignments still exist. Proposition 3 shows that our method can obtain the same result as Drag-and-drop pasting when there exists a perfect boundary. Rather than Drag-and-drop pasting, our method can correct misalignments even when the optimal boundary is not ideal.

***Proposition 3*:** Assume there exists a curve $C: x = f(y)$, such that for each $p \in \{(f(y), y)\}$, $I_1(p) = I_2(p) + k$ where k is a constant offset. As shown in Figure 2, the curve partitions the mosaics region (also overlapping area) into two regions, $\Psi_1 \setminus \Omega_2$ and $\Psi_2 \setminus \Omega_1$, where $x < f(y)$ and $x \geq f(y)$, respectively. Let w be a weighting mask defined as $0.5 \leq w(p) \leq 1$ for all $p \in \Psi_1 \setminus \Omega_2$, and $0 \leq w(p) \leq 0.5$ for all $p \in \Psi_2 \setminus \Omega_1$. Then the optimal partition solution I is a global minimum of $E(I, \mathbb{W})$ in Eq.(8).

$$I(x,y) = \begin{cases} I_1(x,y) + k_1, & x < f(y) \\ I_2(x,y) + k_2, & x \geq f(y) \end{cases} \tag{12}$$

where k_1 and k_2 are two constants with $k_2 - k_1 = k$.

Proposition 3 can be considered as a generalized version of proposition 2. The proof is similar. Proposition 1 and 2 show our method unifies GIST stitching method [4].

4. EXPERIMENTAL RESULTS

In this section, we demonstrate the results of our stitching method in image stitching and editing applications. Comparisons with some methods including Featuring, Laplacian pyramid blending (LPB) [6], Optimal seam method (OSM) [1], Poisson image editing (PIE) [2], Drag-and-Drop pasting (DDP) [3] and GIST [4] are also given. We initialize the stitching image by PIE with gradient field defined by Eq.(9), and then iterate Eq.(6) 100 times.

Figure 3 gives an example of stitching two input images. Our weighting masks are defined with proposition 1. The two input images are aligned with a small overlapping area shown in the red box. The two images contain color discrepancy and local structure misalignments. Figure 3(c) is our stitching result where the structure misalignments are globally aligned and the colors are also matched. From Figure 3(d) to Figure 3(i) are the local views of the results generated by: (d) LPB. The transition is sharp. (e) OSM. The result produces a visible seam. (f) Feathering of images. The structures are not corrected. (g) PIE. (h) GIST. The structure misalignments still exist. (i) Our method.

Figure 4 shows an example for stitching object parts (marked in boxes) where object parts from different images are combined to generate the final image. We can see that our method is better in eliminating structure misalignments.

Our stitching method can be used for some image editing tasks, such as seamless object insertion. As shown in Figure 5, a rainbow is dropped in the destination image. The colors and structures between source and destination images do not matched (see Figure 5(c)). Our weighting masks are computed according to the distances from the boundary of the inserted region. It shows that our method is more powerful in overcoming geometric seams.

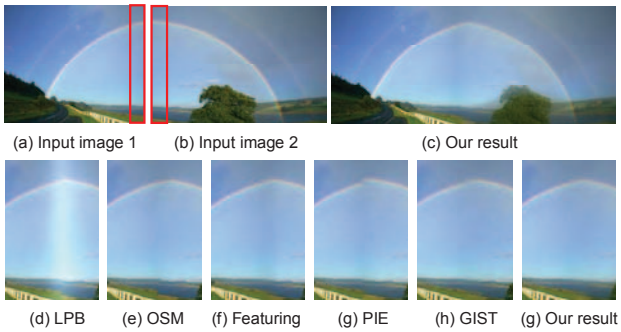

Fig.3. Stitching two images

Top row: the source images, where the object parts are indicated in red zones. Bottom row: from left to right are the results generated using directly cloning, PIE, DDP and our method, respectively.

Fig.4. Stitching object parts

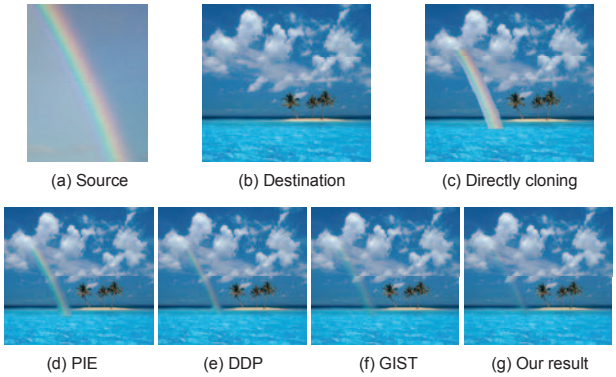

Fig.5. Object insertion

To provide a quantitative evaluation, we perform a perceptual study on a set of 4 stitching results by voting. Each stitching result includes 3 versions generated with PIE, GIST and our method, respectively. Each of the 10 voters views the 3 versions of a stitching result and gives votes with "high", "medium", or "low" according to their evaluations on visual quality. Table 1 gives the voting result. 55% of the "high" votes cast to our method and the remaining cast to GIST. Most of the "low" votes cast to PIE. It reflects that our method is slight better than GIST.

Table 1. Voting result for quantitative evaluation

Method	High	Medium	Low
PIE	0	5	35
GIST	18	20	2
Our method	22	15	3

5. CONCLUSION

In this paper, we propose an image stitching method in gradient domain. Our stitching approach optimizes a variational energy function defined by introducing a guidance vector field into total variation (TV) model. Based on three propositions, we study the relations between the proposed approach and other methods. Due to accurate estimation of discontinuities of TV model, our proposed stitching method is especially valuable in overcoming small geometric seams. Note that optimizing the energy function is computational intensive because of its nonlinearity. Therefore a fast implementation of the algorithm will be studied in further research.

6. REFERENCES

[1] A.A. Efros and W.T. Freeman, "Image quilting for texture synthesis and transfer," in *ACM SIGGRAPH*, pp.341-346, 2001.

[2] P. Perez, M. Gangnet, and A. Blake, "Poisson image editing," *ACM Transaction on Graphics*, vol.22, no.2, pp.313-318, 2003.

[3] J. Jia, J. Sun, C.K. Tang, and H.Y. Shum, "Drag-and-drop pasting," *ACM Transactions on Graphics*,vol.25, no.3, pp.631-636, 2006.

[4] A. Zomet, A. Levin, S. Peleg, and Y.Weiss, "Seamless image stitching by minimizing false edges," *IEEE Transactions on Image Processing*, vol.15, no.4, pp.969-977, 2006.

[5] M. Uyttendaele, A. Eden, and R. Szeliski, "Eliminating ghosting and exposure artifacts in image mosaics," in *IEEE Conference on Computer Vision and Pattern Recognition*, Hawaii, pp. 509-516, 2001.

[6] J. Burt and H.Adelson, "A multiresolution spline with applications to image mosaics," *ACM Transactions on Graphics*, vol.2, no.4, pp. 217-236, 1983.

[7] Kwatra V, Schodl A, Essa I, G. Turk and A. Bobick, "Graphcut textures: Image and video synthesis using graph cuts," *ACM Transactions on Graphics*, vol.22, no.3, pp.277-286, 2003.

[8] J. Hays and A.A. Efros, "Scene completion using millions of photographs," *ACM Transactions on Graphics*, vo1.26, no.3, 2007.

[9] L. Rudin, "Images,numerical analysis of singular and shock filters," *Technical Reports*, California: California Institute of Technology, 1987.

[10] T. Chan and J. Shen, "Mathematical models for local non-texture inpaintings," *SIAM Journal on Applied Mathematics*, vol.62, no.3, pp.1019-1043, 2002.