

# Digital Photograph Stitching with Optimized Matching of Gradient and Curvature

Simon T.Y. Suen, Edmund Y. Lam and Kenneth K.Y. Wong

Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong

## ABSTRACT

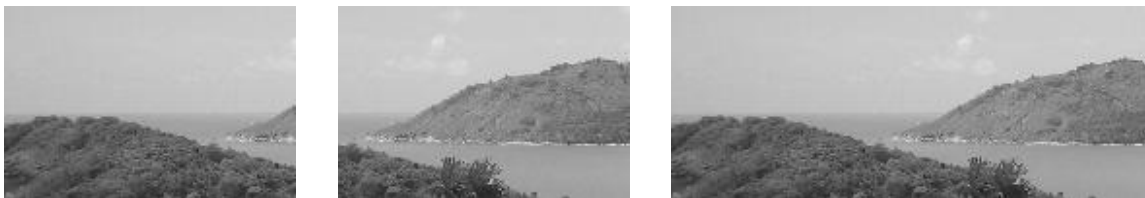
Digital photograph stitching blends multiple images to form a single one with a wide field of view. Sometimes, artifacts may arise, often due to photometric inconsistency and geometric misalignment among the images. Several existing techniques tackle this problem by methods such as pixel selection or pixel blending, which involve the matching of intensity, frequency, and gradient among the input images and adjust them to find the optimal match with the input images. However, our experience indicates that these methods have yet fully incorporated the mathematical properties of the photometric inconsistency. In this paper, we first introduce a general mathematical model describing the properties and effects of the photometric inconsistency. This model supports our claim that matching on the intensity and even the gradient domain is insufficient. Our method thus adds the extra requirement of an optimal matching of curvature. Simulations are carried out using our method, with input images suffering from different kinds of photometric inconsistency under the aligned and misaligned situations. We evaluate the results using both objective and subjective criteria, and we find that our method indeed shows an improvement for certain kinds of photometric inconsistency.

**Keywords:** Image stitching, panoramic mosaicing, photometric inconsistency, gradient, curvature

## 1. INTRODUCTION

Panoramic mosaicing is the process that combines multiple images to form a single one with a wide field of view. Figure 1 shows its simplest case in which only two input images are involved. This is particularly useful when the images taken have a relatively low pixel count, such as in camera phones or PDA cameras. Image alignment and stitching are the main procedures in this application. Image alignment establishes geometric correspondences among the images. Stitching blends the aligned images seamlessly.

Sometimes, artifacts in stitching may arise. It is often due to photometric inconsistency and geometric misalignment among the images. The former is caused by various physical reasons such as lens vignetting, exposure duration, scene illumination and camera gain.<sup>1</sup> They manifest in differences in pixel intensity values



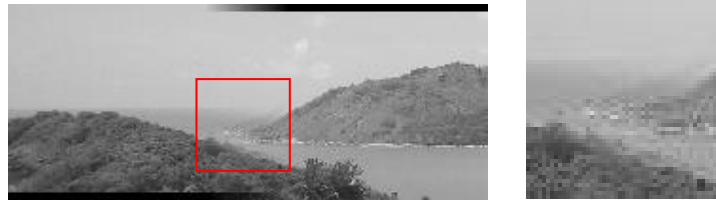
**Figure 1.** The simplest case of panoramic mosaicing in which only two input images are involved. The left and middle images are the input images, while the right one is the mosaic image.

---

Further author information: (Send correspondence to Simon T.Y. Suen)  
Simon T.Y. Suen: E-mail: tysuen@eee.hku.hk, Telephone: 852 2859 2696  
Edmund Y. Lam: E-mail: elam@eee.hku.hk, Telephone: 852 2241 5942  
Kenneth K.Y. Wong: E-mail: kywong@eee.hku.hk, Telephone: 852 2857 8483



**Figure 2.** Two input images with photometric inconsistency (left and middle) and their mosaic image (right) in which a visible cutting curve appears.



**Figure 3.** Stitching result under a misaligned situation (left) and its magnified overlap (right). Some blurred and double edges appear in the overlap.

even if they represent the same location in the object. Visible cutting curve is the most common artifact (see Fig. 2). The latter is inherited from the imperfection of the alignment process. Pixel intensity values at the same location may differ when they belong to different objects. Blurred edges and double edges are the most common artifacts in this case (see Fig. 3). These problems make the stitching difficult: we cannot simply take one of these values from the input images, as it would be inconsistent with the others.

## 2. EXISTING APPROACHES

The existing image stitching methods can be classified into four broad areas: optimal seam approach, smooth transition approach, color tuning approach and optimization approach.

### 2.1. Optimal Seam Approach

Optimal seam methods<sup>2-7</sup> search for the optimal curve in the overlap region where the differences between the two input images are minimal. Each image is then copied to the corresponding side of the curve. Seams on the curve are not noticeable when the differences are close to zero. However, such a seamless curve may not exist when there is photometric inconsistency throughout the area. Moreover, these methods are less appropriate to the cases with a narrow overlap region, because it may be very difficult to find a potential seamless curve.

### 2.2. Smooth Transition Approach

The second approach smoothes the transition between the input images. In feathering,<sup>8</sup> each overlapping pixel is computed as the weighted average of the intensity values from the inputs. The weighting coefficients are determined by the distance between the pixel and the corresponding image center. If the inputs are well-aligned, seams become mostly invisible, unless the photometric inconsistency is very severe. However, blurred or double edges often appear when they are misaligned. Laplacian pyramid blending<sup>9</sup> smooths the transition in each frequency band independently. It gives a wider transition zone to the lower frequency bands and a narrower zone to the higher ones. The blurred and double edges appear less noticeably under a misaligned situation. However, the misalignment could result in the mismatches among the frequency bands. Thus, the mosaic image cannot fully preserve the image contents from the inputs in the non-overlapping regions.

### 2.3. Color Tuning Approach

The third approach tunes the photometric properties of the input images before stitching. In color correction,<sup>10</sup> the photometric inconsistency is first formulated into a linear function by matching the intensity values between the inputs in the overlap region. All pixel intensity values in one of the inputs are then converted by the function in order to match another input before stitching. Exposure compensation<sup>8</sup> improves the feathering result when photometric inconsistency is very severe. It estimates the inconsistency between the inputs and the feathering result with a quadratic spline function. This function is then applied to the inputs before the next stitching. These methods can produce a very smooth transition at the expense of modifying the inputs. Furthermore, since the formulation of the inconsistency is highly sensitive to the geometric misalignment, a wrong formulation would lead to a wrong correction and thus an overexposed, underexposed or blurry result.

### 2.4. Optimization Approach

The fourth approach computes the mosaic image by an optimization process. Gradient-domain image stitching (GIST)<sup>1</sup> optimizes an energy function defined in terms of pixel gradient values. It aims to obtain a mosaic result whose pixel gradient values are close to that of the inputs. In the process, the creation of new edges are suppressed and the images could blend regardless of their differences in mean intensities.  $L_1$  optimization is recommended after comparing among its variations. It can stitch seamlessly even when both photometric inconsistency and geometric misalignment exist. However, the mean intensities of the inputs are modified in the result because the intensity values are not maintained in the process.

In summary, these approaches tackle the stitching problems by pixel selection or pixel blending, which involve the matching of intensity, frequency, and gradient among the input images and adjusting them to find the optimal match with the inputs. However, under some cases, they fail to meet the two major quality criteria for stitching which are the preservation of the image contents and the absence of noticeable seams.

In Sect. 3, we first introduce a general mathematical model describing the properties and effect of the photometric inconsistency. In Sect. 4, we show that this model supports our claim that matching on the intensity and even the gradient domain is insufficient. In Sects. 5 and 6, we describe our method which adds the extra requirement of an optimal matching of curvature and its implementation. In Sect. 7, simulations are carried out using our method, with input images suffering from different kinds of photometric inconsistency under the aligned and misaligned situations. We evaluate the results using both objective and subjective criteria, and we find that our method indeed shows an improvement for certain kinds of photometric inconsistency.

## 3. MATHEMATICAL MODEL

In this section, we describe the mathematical model of image capture and that of photometric inconsistency. We then analyze the properties of the latter model in the gradient and curvature domain in Sect. 4.

### 3.1. Image Capture

Let  $i(x, y)$  be a 2-D image representing a scene. We simulate the image capture geometrically by cropping out  $i(x, y)$  in the regions  $\Omega_1$  and  $\Omega_2$  to get the images  $g_1(x, y)$  and  $g_2(x, y)$ , respectively. The two captured images overlap in  $\Omega_0 = \Omega_1 \cap \Omega_2$ . We also assume that the intensity values of the captured images linearly map that of the scene image. We call this mapping the *linear photometric transfer function* (LPTF). Then, the image capture model is formed as:

$$g_n(x, y) = p_n(i(x, y)) = a_n(x, y)i(x, y) + b_n(x, y) \quad \forall (x, y) \in \Omega_n \quad \text{and} \quad n = 1, 2, \quad (1)$$

where  $p_n$  denotes the LPTF of  $g_n$ , and  $a_n(x, y)$  and  $b_n(x, y)$  are the coefficients of  $p_n$ .

Next, we describe the explicit form of the LPTF in a digital camera system. According to Ref. 11, in a typical image formation system, the recorded intensity value  $I$  is related to scene radiance  $L$  as:

$$I = L \frac{\pi t}{4} \left( \frac{d}{f} \right)^2 \cos^4 \theta + e, \quad (2)$$



**Figure 4.** Original image (left). Image captured with a shorter exposure duration (middle). Image captured with heavy lens vignetting (right). Note that exposure duration difference produces the change of the global brightness, while lens vignetting darkens the image gradually when getting close to the corner.

where  $f$  is the focal length of the imaging lens,  $d$  is the diameter of its aperture,  $\theta$  is the angle subtended by the principal ray from the optical axis,  $t$  is the exposure duration of the image detectors and  $e$  is the additive noise. Thus, in the LPTFs, we have  $a_n = \frac{\pi t_n}{4} \left(\frac{d_n}{f_n}\right)^2 \cos^4 \theta_n$  and  $b_n = e_n$ . In our discussion, we assume the magnitude of noise is small in the digital camera system and thus  $b_n = 0$ .

### 3.2. Photometric Inconsistency

We model the effect of the photometric inconsistency by using the LPTFs. We take those caused by exposure duration difference and lens vignetting as our examples. Figure 4 shows their visual effects on a photograph. Note that exposure duration difference produces the change of the global brightness, while lens vignetting darkens the image gradually when getting close to the corner.

For exposure duration difference, we assume that some variables are equal between the two LPTFs i.e.  $d_1 = d_2 = d$ ,  $f_1 = f_2 = f$  and  $\theta_1 = \theta_2 = 0$ . Then we have  $a_n = kt_n$ , where  $k = \frac{\pi}{4} \left(\frac{d}{f}\right)^2$  is a constant and  $t_n$  is the exposure duration for image  $g_n$ . The differences between the captured images in the overlap region are thus:

$$g_1(x, y) - g_2(x, y) = k(t_1 - t_2)i(x, y) \quad \forall (x, y) \in \Omega_0. \quad (3)$$

This illustrates that the inconsistency caused by exposure duration difference would appear as a constant gain of the intensities of the scene image.

For lens vignetting, with reference to Ref. 12, we assume that  $d_1 = d_2 = d$ ,  $f_1 = f_2 = f$ , and  $t_1 = t_2 = t$ , then  $a_n = \hat{k} \cos^4 \theta_n$  where  $\hat{k} = \frac{\pi t}{4} \left(\frac{d}{f}\right)^2$  is a constant and  $\theta_n = \tan^{-1} \frac{\sqrt{(x-w_n)^2 + (y-v_n)^2}}{f}$  is the angle subtended by the principal ray from the image center  $(w_n, v_n)$  of image  $g_n$ . The differences between the captured images in the overlap region are thus:

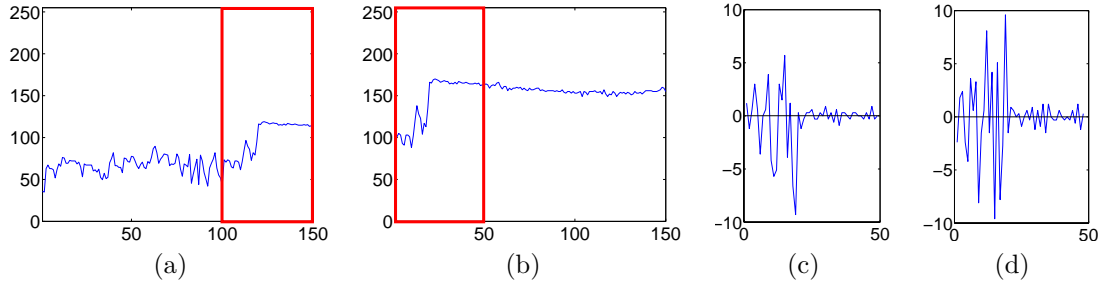
$$g_1(x, y) - g_2(x, y) = \hat{k}(\cos^4 \theta_1 - \cos^4 \theta_2)i(x, y) \quad \forall (x, y) \in \Omega_0. \quad (4)$$

This illustrates that the inconsistency caused by lens vignetting appears as a spatially varying gain of the intensities of the scene image.

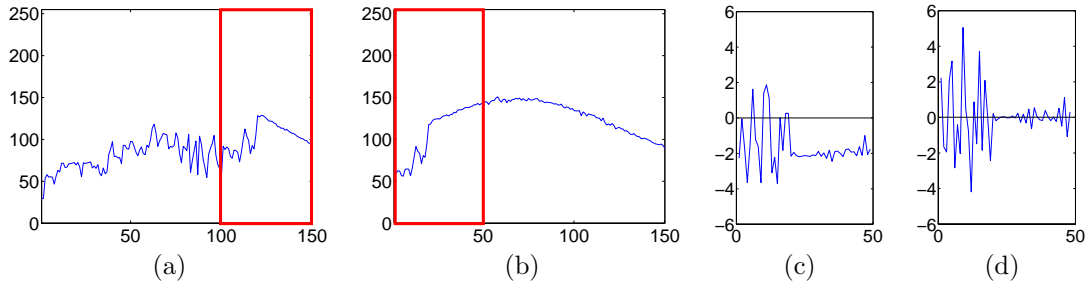
## 4. GRADIENT AND CURVATURE VALUES

In order to simplify our analysis, we now discuss in 1-D domain. By spline theory, each image can be approximately represented by a cubic spline which is a joint composite of several pieces of cubic polynomials. We now pick out a certain piece from them in the overlap region for further analysis. Furthermore, since the LPTFs of exposure duration difference and lens vignetting do not have a rapid spatial variation, we assume that a linear polynomial can approximately represent them within that piece. Then, we have the image capture model of each piece in 1-D domain as:

$$g_n(x) = p_n(i_n(x)) = (r_n x + s_n)(c_3 x^3 + c_2 x^2 + c_1 x + c_0) \quad \forall x \in \Omega_n \quad \text{and} \quad n = 1, 2, \quad (5)$$



**Figure 5.** (a), (b) Two 1-D image signals captured with different exposure durations. The boxes indicate their overlap region. (c) Differences between their gradient values within the region and (d) those between their curvature values. Both differences are approximately zero in the flat region on the right part of the overlap region.



**Figure 6.** (a), (b) Two 1-D image signals captured with lens vignetting. The boxes indicate their overlap region. (c) Differences between their gradient values within the region and (d) those between their curvature values. Only the latter are approximately zero in the flat region on the right part of the overlap region.

where  $r_n$  and  $s_n$  are the coefficients of  $a_n$  in  $p_n$  and  $c$  denotes the coefficients of the cubic polynomial of that piece from the scene image. We then look at their gradient and curvature values (i.e. the 1<sup>st</sup> and 2<sup>nd</sup> derivatives),

$$g'_n(x) = (r_n x + s_n)(3c_3 x^2 + 2c_2 x + c_1) + r_n(c_3 x^3 + c_2 x^2 + c_1 x + c_0), \quad (8)$$

$$g''_n(x) = (r_n x + s_n)(6c_3 x + 2c_2) + 2r_n(3c_3 x^2 + 2c_2 x + c_1). \quad (9)$$

Now, we consider that the piece is picked from an approximately flat region such that  $c_1 \approx 0$ ,  $c_2 \approx 0$  and  $c_3 \approx 0$ , then the respective differences between  $g_1$  and  $g_2$  are

$$g'_1(x) - g'_2(x) \approx (r_1 - r_2)c_0, \quad (8)$$

$$g''_1(x) - g''_2(x) \approx 0. \quad (9)$$

In case of exposure duration difference, it does not introduce a spatially varying gain to the captured image. We can assume that  $r_n$  is zero and thus the two differences become approximately zero. This implies that in a flat region, both gradient and curvature values can match each other and reject the effect of the inconsistency. Figure 5 shows an experiment verifying our analysis. The two 1-D signals are picked from the two 2-D images captured with different exposure durations. In the overlap region, their differences in gradient and curvature domains are approximately zero in the flat region on its right side.

However, in case of lens vignetting, we cannot assume that  $r_n$  is approximately zero as it introduces a spatially varying gain to the captured image. Thus, only the differences in the curvature domain are approximately zero. Figure 6 gives an experiment verifying this analysis. To conclude, for certain kinds of inconsistency which vary spatially such as those caused by lens vignetting, the curvature values between the captured images can better match each other. Thus, in addition to the intensity and gradient values, our method adds an extra requirement of matching the curvature values.

## 5. METHOD

Our method, the *Curvature Domain Image Stitching* (CIST), is an optimization scheme which minimizes an energy function defined in terms of pixel intensity, gradient and curvature values.

### 5.1. Energy Function

Let  $\hat{g}(x, y)$  be the resulting image,  $g_1(x, y)$  and  $g_2(x, y)$  be the two aligned input images. Let  $\alpha(x, y) \in [0, 1]$  be the weighting mask that determines the weighting between  $g_1$  and  $g_2$ . Let  $\beta_i(x, y) \in [0, 1]$  be another weighting mask that determines the weighting of the  $i^{\text{th}}$  derivative among all the considered derivatives i.e. the intensity, gradient and curvature values. Let  $\|\cdot\|_p$  be the  $L_p$  vector norm. The energy function is defined as:

$$E(\hat{g}) = \sum_{i=0}^2 \left( \sum_{(x,y) \in \Omega_1} \alpha(x, y) \beta_i(x, y) \|D^i(\hat{g}(x, y) - g_1(x, y))\|_p^p + \sum_{(x,y) \in \Omega_2} (1 - \alpha(x, y)) \beta_i(x, y) \|D^i(\hat{g}(x, y) - g_2(x, y))\|_p^p \right), \quad (10)$$

where  $\Omega_1$  and  $\Omega_2$  are the capturing regions of  $g_1$  and  $g_2$ , respectively.  $D^i f$  is the  $i^{\text{th}}$  derivative operator for the real function  $f$  such that  $D^0 f = f$ ,  $D^1 f = [\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}]$  and  $D^2 f = [\frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial y^2}]$ , representing the intensity, gradient and curvature values respectively. Next, we are going to discuss the three major variables (i.e.  $p$ ,  $\alpha$  and  $\beta_i$ ) in the function in the following subsections.

### 5.2. $L_p$ -norm

The variable  $p$  determines the type of vector norm measuring the distance between the result and one of the input images. The choice of norm determines the optimal solution and thus the mosaic image. Suppose the values of the two input images mismatch each other at a pixel location. For  $L_1$ -norm, any values bounded by these mismatched values can be the optimal solution. However, for  $L_2$ -norm, only the weighted average of the mismatched values is the optimal solution while the weighting is determined by the weighting mask  $\alpha$ . Their results are similar under a well-aligned situation. However, in Sect. 7, we will show that  $L_1$ -norm is better than  $L_2$ -norm under a misaligned situation.

### 5.3. Weighting Mask $\alpha$

The weighting mask  $\alpha$  determines the weighting between the inputs. As defined in Eq. 10, if  $\alpha = 1$ , the latter term in the big bracket becomes zero. The result  $\hat{g}$  is affected by the derivatives in  $g_1$  only. If  $\alpha = 0$ , the former term becomes zero. Then the result is affected by the values in  $g_2$  only. There are various ways to design the mask. We here introduce three possible types (for a vertical overlap) and test them in Sect. 7. Let  $c_x$  be the center of the overlap region, and  $d$  be a tunable parameter controlling the transition width. The uniform mask assigns a constant  $\mu \in [0, 1]$ , say 0.5, to the overlap region. Thus, we have

$$\alpha(x, y) = \begin{cases} 1 & \text{if } x < c_x - d; \\ \mu & \text{if } c_x - d \leq x \leq c_x + d; \\ 0 & \text{if } x > c_x + d. \end{cases} \quad (11)$$

The feathered mask assigns a gradual transition from 1 to 0 in the overlap region. Thus, we have

$$\alpha(x, y) = \begin{cases} 1 & \text{if } x < c_x - d; \\ \frac{c_x + d - x}{2d} & \text{if } c_x - d \leq x \leq c_x + d; \\ 0 & \text{if } x > c_x + d. \end{cases} \quad (12)$$

And, the optimal-cut mask assigns a step from 1 to 0 when crossing an optimal cutting curve searched by the optimal seam methods. Thus, let  $x = f(y)$  be the curve,

$$\alpha(x, y) = \begin{cases} 1 & \text{if } x \leq f(y); \\ 0 & \text{if } x > f(y). \end{cases} \quad (13)$$

In Sect. 7, we show that when comparing the three types of masks, only the optimal-cut mask does not produce artifacts such as blurred and double edges under a misaligned situation. Thus, we recommended it. Here, we introduce the searching scheme of the optimal-cut which can better suit our method. Normally, the curve is the minimal cost path through the error surface defined by the differences of pixel intensity values between the images. However, as analyzed in Sect. 4, the curvature values in a flat region can be less sensitive to the photometric inconsistency which varies spatially. Therefore, we define the error surface as

$$e(x, y) = |\Delta g_1(x, y) - \Delta g_2(x, y)| + |\Delta g_1(x, y) + \Delta g_2(x, y)|, \quad \forall (x, y) \in \Omega_0, \quad (14)$$

where  $\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$  is the Laplacian operator for the real function  $f$ . The first term is the differences of curvature values. The second term represents the smoothness at the point  $(x, y)$ . Then, the minimal cost path should pass through a smooth region with minimal differences in curvature values. The path can be easily found with dynamic programming described in Ref. 7.

#### 5.4. Weighting Mask $\beta_i$

The weighting mask  $\beta_i$  determines the weighting of the  $i^{th}$  derivative among all the considered derivatives. As explained in Sect. 4, we will emphasize the use of curvature values, especially in the overlap region. To achieve this, we design three possible types of masks (for a vertical overlap). First, the uniform mask assigns a constant to each  $\beta_i$  in the whole region of the mosaic image (denoted as  $\hat{\Omega} = \Omega_1 \cup \Omega_2$ ), such that  $\beta_0 < \beta_1 < \beta_2$ . Thus, we have for  $i = 0, 1, 2$ ,

$$\beta_i(x, y) = \mu_i \quad \forall (x, y) \in \hat{\Omega}, \quad (15)$$

where  $\mu_i$  is a constant,  $\mu_0 < \mu_1 < \mu_2$  and  $\sum_{i=0}^2 \mu_i = 1$ . The feathered mask divides the whole region into four vertical portions  $\epsilon_j$  of width  $\delta_j$ , where  $j = 1, 2, 3, 4$ . Let  $x_{\epsilon_j}$  be the leftmost  $x$ -coordinates of the portion  $\epsilon_j$ . Then, for  $i = 0, 1, 2$ ,

$$\beta_i(x, y) = \begin{cases} 1 - (x - x_{\epsilon_{i+1}})/\delta_i & \text{if } (x, y) \in \epsilon_{i+1}; \\ (x - x_{\epsilon_{4-i}})/\delta_i & \text{if } (x, y) \in \epsilon_{4-i}; \\ 0 & \text{otherwise.} \end{cases} \quad (16)$$

The sharp-cut mask divides the whole region into six vertical portions,  $\epsilon_k$ , where  $k = 1, 2, \dots, 6$ . For  $i = 0, 1, 2$ ,

$$\beta_i(x, y) = \begin{cases} 1 & \text{if } (x, y) \in \epsilon_{i+1} \cup \epsilon_{6-i}; \\ 0 & \text{otherwise.} \end{cases} \quad (17)$$

When using the uniform and the feathered mask, every resulting pixel in the mosaic image is the optimized result of the values in multiple derivative domains. When the values do not mismatch each others, unsatisfactory visual results will be produced. Section 7 shows a simulation to support this claim. Moreover, an optimization of multiple derivative values at a pixel location involves high computational power. Thus, we recommend to use the sharp-cut mask which can give a satisfactory result with the minimal computational effort.

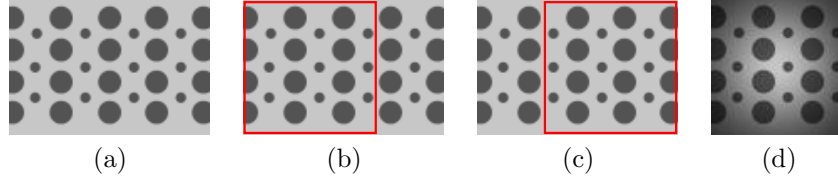
## 6. IMPLEMENTATION

We have implemented the optimization process for the energy function defined in Eq. 10 under  $L_1$  and  $L_2$ -norm. The  $1^{st}$  derivative operator we use is a forward differencing derivative filter  $[-1 \ 1]$  while the second one is thus  $[1 \ -2 \ 1]$ . With reference to Ref. 1, for the  $L_1$ -case, we can optimize the function shown below by linear programming with the `linprog` command in the Matlab Optimization Toolbox<sup>13</sup>:

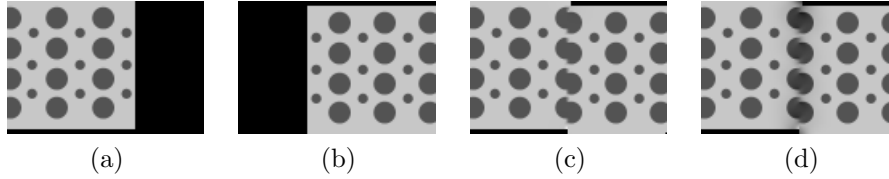
$$\min_x \sum_i (z_i^+ + z_i^-) \quad \text{subject to} \quad Ax + (z^+ - z^-) = b, \quad 0 \leq x \leq 255, \quad z^+ \geq 0, \quad z^- \geq 0. \quad (18)$$

where  $A$  is the sparse matrix containing the weighted derivative operators.  $b$  is the vector containing the known weighted derivatives of the input images.  $x$  is the vectorization of the resulting image. For the  $L_2$ -case, using the same  $A$  and  $b$ , we solve the constrained linear least squares problem defined below by using the `lsqlin` command in the Toolbox:

$$\min_x \|Ax - b\|_2^2 \quad \text{subject to} \quad 0 \leq x \leq 255. \quad (19)$$



**Figure 7.** (a) Pattern used as the scene image. (b), (c) Capturing regions for input image 1 and 2 respectively. (d) Input image 1 captured with heavy lens vignetting.



**Figure 8.** (a), (b) Misaligned arrangements for the input image 1 and 2 respectively. (c), (d) Optimization results under  $L_1$ - and  $L_2$ -norm respectively.

## 7. SIMULATION

### 7.1. Comparisons using Synthetically-captured Patterns

In the first stage, we use the pattern with  $80 \times 120$  pixels shown in Fig. 7(a), as the scene image to highlight the distinct results in the following comparisons: (1) Optimizing under  $L_1$ -norm vs.  $L_2$ -norm. (2) Applying different types of weighting masks for  $\alpha$ . (3) Applying different types of weighting masks for  $\beta_i$ . (4) Stitching in gradient domain vs. our method. (5) Stitching in intensity domain vs. our method. We simulate the image capture in the following ways. For the first two comparisons, we form the two input images by cropping directly from the regions shown in Figs. 7(b) and 7(c). For the other comparisons, we instead apply heavy lens vignetting to the image 1 after cropping. Figure 7(d) shows the resulting image 1. To do this, in Eq. 4, we choose  $k = 1$  and  $\cos^4 \theta_2 = 1$  while  $\cos^4 \theta_1$  decreases gradually to 0.2 at its image corner.

#### 7.1.1. $L_1$ -norm vs. $L_2$ -norm

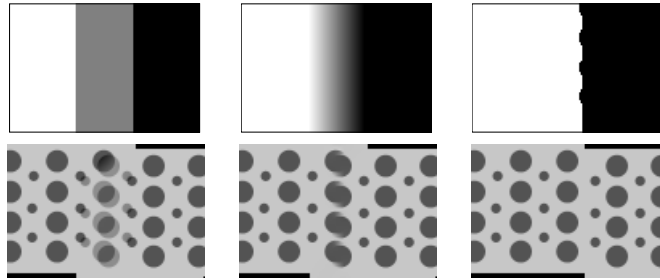
To highlight the distinct features between the optimization under  $L_1$ - and  $L_2$ -norm, we stitch the cropped images under the misaligned situation. Referring to Sect. 5.3, we use the optimal-cut mask for  $\alpha$  whose vertical cut lies in the middle of the overlap region. Also, with reference to Sect. 5.4, we use the sharp-cut mask with equally-wide vertical portions for  $\beta_i$ . Figure 8 shows the misaligned input images and the results under  $L_1$ - and  $L_2$ -norm.  $L_2$ -norm produces smearing artifacts on the vertical cut while these do not exist in the  $L_1$ -norm case. This is because at each mismatched pixel on the cut, only the weighted average of the curvature values is the optimal solution for  $L_2$ -norm. Thus, in the  $L_2$ -norm case, the curvature values at the pixel are mixed together while its neighboring values are severely modified. Therefore, the intensity values around the cut appear to be changed.

#### 7.1.2. Uniform, Feathered and Optimal-cut Mask for $\alpha$

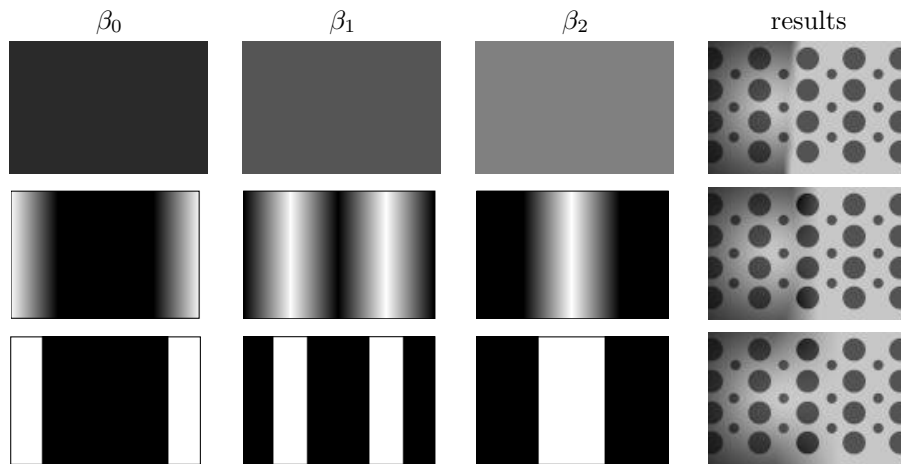
We use the same configuration as in Sect. 7.1.1 except the weighting mask for  $\alpha$  and we optimize under  $L_1$ -norm only. With reference to Sect. 5.3, we use  $\mu = 0.5$  for the uniform mask,  $d = 20$  for the feathered mask and the newly introduced searching scheme to design the sharp cut mask. Figure 9 shows the masks and their results. Only the optimal-cut mask does not produce double and blurred edges.

#### 7.1.3. Uniform, Feathered and Sharp-cut Mask for $\beta_i$

In this simulation, we stitch the captured images with photometric inconsistency, as stated in Sect. 7.1, under the well-aligned situation. Referring to Sect. 5.4, we use the uniform mask with  $\mu_0 = \frac{1}{6}$ ,  $\mu_1 = \frac{1}{3}$  and  $\mu_2 = \frac{1}{2}$ . For



**Figure 9.** Uniform mask for  $\alpha$  and its result (left column). Feathered mask and its result (middle column). Optimal-cut mask and its result (right column). Note that black means 0 and white means 1 in the masks.

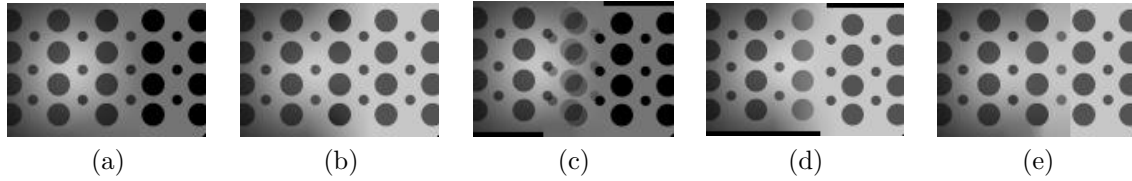


**Figure 10.** Uniform mask for  $\beta_i$  and its result (top row). Feathered mask and its result (middle row). Sharp-cut mask and its result (bottom row). Note that black means 0 and white means 1 in the masks.

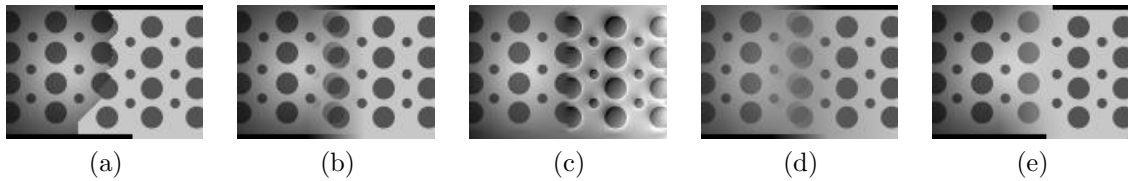
the feathered and the sharp-cut mask, we divide the whole mask into equally-wide vertical portions. Figure 10 shows the masks and their results. The uniform mask makes the cutting curve visible. The top and the bottom of the curve are also noticeable in the result of using the feathered mask. These artifacts are possibly caused by the mismatches of the derivative values at the pixels around the curve. However, the sharp-cut mask can produce a satisfactory result. By gathering with the results of the last two comparisons, we choose the optimization under  $L_1$ -norm with the optimal-cut mask for  $\alpha$  and the sharp-cut mask for  $\beta_i$  for the next few comparisons.

#### 7.1.4. Gradient Domain vs. Curvature Domain

First, we compare our method with the gradient stitching (GIST- $L_1$ ) recommended in Ref. 1. We stitch with the same configuration as in Sect. 7.1.3 under the aligned and misaligned situation. Figures 11(a) and 11(b) show that under the aligned situation, both method can stitch without a noticeable seam in the overlap. However, the right side of the GIST- $L_1$  result is darkened. Moreover, Figures 11(c) and 11(d) show that under the misaligned situation, double edges appear in the GIST- $L_1$  result. Secondly, we modify their energy function to maintain the intensity values near the left and the right boundaries but not in our result. Therefore, we modify the sharp-cut mask for  $\beta_i$  in GIST- $L_1$  as follows: the whole region is divided into four equally-wide portions, giving a full weighting to the intensity domain for the first and the fourth portions while giving it to the gradient domain for the middle two portions. Figure 11(d) shows the result of the modified GIST- $L_1$  under the well-aligned situation. It indicates that seams are noticeable on the top and bottom but not in the middle. These two tests support



**Figure 11.** (a), (b) Results of GIST- $L_1$  and our method under the well-aligned situation respectively. (c), (d) Their results under the misaligned situation respectively. (e) Result of the modified GIST- $L_1$  under the well-aligned situation.



**Figure 12.** Results under the misaligned situation: (a) optimal seam method (implemented as in Ref. 7), (b) feathering, (c) pyramid blending, (d) exposure compensation and (e) our method. Only our method can produce the result without double edges and a noticeable cutting curve.

our claim that the stitching in gradient domain is not sufficient to produce seamless results and maintain the brightness at the same time when photometric inconsistency varies spatially.

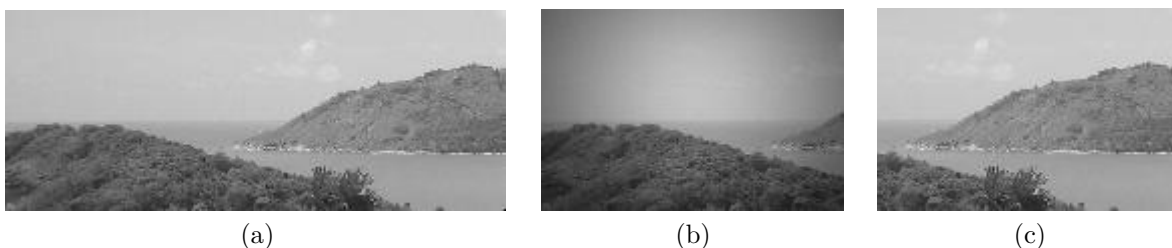
### 7.1.5. Intensity Domain vs. Curvature Domain

Next, we compare our method with the stitching methods in intensity domain under the misaligned situation. The input images are the same as those in Sect. 7.1.3. Figure 12 shows that our method outperforms the others as only we can produce the result without double edges and a noticeable cutting curve.

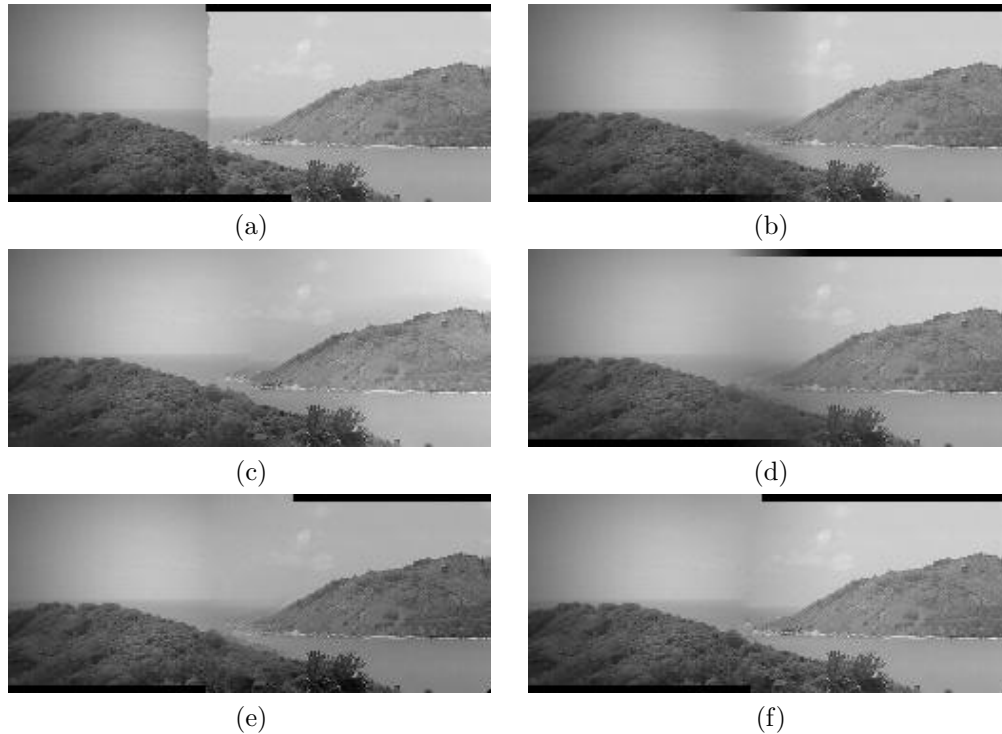
## 7.2. Comparisons using Synthetically-captured Pictures

In the second stage, we compare the performance of the existing methods by using a more realistic scene content. Other than showing their visual results, we also use the Mean Square Error (MSE) as an objective measure to evaluate their performance. We use the picture with  $100 \times 150$  pixels shown in Fig. 13(a), as the scene image. To simulate an image capture, we apply lens vignetting to the image 1 after cropping it from the scene image while the image 2 is directly cropped (see Fig. 13(b) and 13(c)). To do this, in Eq. 4, we choose  $k = 1$  and  $\cos^4 \theta_2 = 1$  while  $\cos^4 \theta_1$  decreases gradually to 0.5 at its image corner.

We stitch the captured images under a misaligned situation. Figure 14 shows the visual results of applying the methods. The cutting curve is noticeable in Fig. 14(a). Double or blurred edges are noticeable in the overlap



**Figure 13.** (a) Picture used as the scene image, (b) input image 1 captured with lens vignetting and (c) input image 2 cropped directly from the scene image.



**Figure 14.** Results under the misaligned situation: (a) optimal seam, (b) feathering, (c) pyramid blending, (d) exposure compensation, (e) GIST under  $L_1$ -norm, (f) our method.

**Table 1.**  $MSE_p$  and  $MSE_s$  of the results in Fig. 14: (a) optimal seam, (b) feathering, (c) pyramid blending, (d) exposure compensation, (e) GIST under  $L_1$ -norm and (f) our method.

	(a)	(b)	(c)	(d)	(e)	(f)
$MSE_p$	0	0	38.5120	38.7099	480.4857	1.1181
$MSE_s$	33.6764	532.2502	431.6775	973.1583	852.4017	383.8243

region in Figs. 14(b-e). Moreover, the brightness on the right side in Figs. 14(c) and 14(e) is modified. Only our method can produce a result satisfying the two major quality criteria: the preservation of image content in the non-overlapping regions and the absence of noticeable seam in the overlap region.

Next, we measure the two criteria objectively with the MSEs. The first criteria is measured by the MSE (denoted as  $MSE_p$ ) between the result and the two inputs under the non-overlapping regions. The second criteria is measured by the MSE (denoted as  $MSE_s$ ) between the result and the reference image under the overlap region. The reference one should produce the minimum amount of seams even under the misalignment situation. Figure 15 shows the reference we use which is the optimal seam result of stitching the inputs cropped directly from the scene image, under the same misalignment situation. Table 1 lists the  $MSE_p$  and  $MSE_s$  of the results in Fig. 14. For  $MSE_p$ , our method is approximately zero, indicating that it does not modify the input contents in the non-overlapping regions. For  $MSE_s$ , other than the optimal seam method, our method gets the lowest value. Although the optimal seam gets the lowest value, it shows a visible cutting curve in between the overlap region and the left non-overlapping region. Our designed  $MSE_s$  cannot take such kind of seams into account, but it is simple and the seams can be easily detected by human eyes. Hence, in addition to the preservation of image contents, our method works best to prevent the occurrence of seams.



**Figure 15.** Reference mosaic image used for calculating the  $MSE_s$ .

## 8. CONCLUSION

In this paper, we have introduced the mathematical model of photometric inconsistency. The model implies that during stitching, the gradient values are insufficient to reject the effects of certain kinds of photometric inconsistency which vary spatially. On the other hand, the curvature values can. Our method thus optimizes an energy function defined in terms of the pixel intensity, gradient and curvature values. We emphasize the importance of curvature values, especially in the overlap region, by applying some designed weighting masks. Our simulation results show that the optimization under  $L_1$ -norm with the optimal-cut mask for  $\alpha$  and the sharp-cut mask for  $\beta_i$  can yield seamless results under the aligned and misaligned situation. Our methods can outperform the existing approaches, especially when the inconsistency varies spatially, for instance, those caused by lens vignetting. Moreover, from our experience, the optimization involving only gradient values can perform faster than that involving only curvature values. In order to shorten the computation time, our future work will focus on designing a metric which measures the photometric inconsistency among the inputs and tells which values among the intensity, gradient or curvature values are necessary at that situation.

## REFERENCES

1. A. Levin, A. Zomet, S. Peleg, and Y. Weiss, "Seamless image stitching in the gradient domain," in *8th European Conference on Computer Vision*, **4**, pp. 377–389, May 2004.
2. D. L. Milgram, "Computer methods for creating photomosaics," *IEEE Transactions on Computers* **24**, pp. 1113–1119, November 1975.
3. S. Peleg, "Elimination of seams for photomosaics," in *Computer Graphics and Image Processing*, **16**, pp. 90–94, August 1981.
4. S. Yang, L. Li, and G. Peng, "Two dimensional seam-point searching in digital image mosaicking," in *Photogrammetric Engineering and Remote Sensing*, **55**, pp. 49–53, January 1989.
5. S. Hummer-Miller, "A digital mosaicking algorithm allowing for an irregular join line," in *Photogrammetric Engineering and Remote Sensing*, **55**, pp. 43–47, January 1989.
6. F. P. A. Junior and N. J. Leite, "A morphological algorithm for photomosaicking," in *8th European Signal Processing Conference*, **3**, pp. 1881–1884, September 1996.
7. A. A. Efros and W. T. Freeman, "Image quilting for texture synthesis and transfer," in *SIGGRAPH '01: Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pp. 341–346, August 2001.
8. M. Uyttendaele, A. Eden, and R. Szeliski, "Eliminating ghosting and exposure artifacts in image mosaics," in *Computer Vision and Pattern Recognition*, **2**, pp. 509–516, December 2001.
9. P. J. Burt and E. H. Adelson, "A multiresolution spline with application to image mosaics," *ACM Transactions on Graphics* **2**, pp. 217–236, October 1983.
10. G. Y. Tian, D. Gledhill, D. Taylor, and D. Clarke, "Colour correction for panoramic imaging," in *6th International Conference on Information Visualisation*, pp. 483–488, July 2002.
11. T. Mitsunaga and S. K. Nayar, "Radiometric self calibration," in *Computer Vision and Pattern Recognition*, **1**, pp. 1374–1380, June 1999.
12. D. Hasler and S. Süsstrunk, "Mapping colour in image stitching applications," *Journal of Visual Communication and Image Representation* **15**, pp. 65–90, March 2004.
13. *MATLAB Optimization Toolbox User's Guide*, The MathWorks, Inc., 2005.