Asian Journal of Information Technology 7 (11): 498-509, 2008
ISSN: 1682-3915

An Image Stitching Process Using Band-Type Optimal Partition Method

Tien-Der Yeh, Yon-Ping Chen and Zih-Yang Liao
Electrical and Control Engineering, National Chiao Tung University,
1001 Ta Hsueh Rd., Hsinchu, Taiwan 300, ROC

Abstract: A practical image stitching method, named band-type optimum partition method or BOP method in short, is introduced in this study. By the use of multiple cut lines and band-type borders, this method achieves 2 major advantages. First, it reduces the discontinuities appearing on the boundaries of the overlap region. Second, it eliminates the undesired ghost image effect, which is caused by misalignment of the input images. Besides, the BOP method is implemented by evaluating the sum of similarity and smoothness costs on a group of pixels instead of a single pixel. Finally, experimental results are presented to demonstrate the success of the BOP method.

Key words: Image stitching, optimum partition, band-type, cut line, seam

INTRODUCTION

Recently, many investigators have paid attention to the technology of image stitching and applied it to several situations, such as moving object detection and tracking (Hsieh, 2004), texture synthesis and transfer (Efros and Freeman, 2001), object insertion (Jia and Tang, 2005) and panoramas construction (Brown and Lowe, 2003). This study focuses on the use of image stitching for panoramas construction to enhance the resolution of a scene or combine individual images into a high quality panoramic image.

The entire process of the image stitching is shown in Fig. 1, including camera calibration, image registration, image mapping, brightness normalization and the proposed Band-type Optimum Partition method (BOP method). At the beginning, the camera calibration corrects the distortion caused by the lens. Next, the image registration finds 2 sets of feature points, one for the target image and the other one for the source image and then determines the matching pairs between these 2 sets. There are many methods presented to do image registration (Zitova and Flusser, 2003; Han and Lin, 2006; Lowe, 2004), wherein Lowe's Scale Invariant Feature Transform (SIFT) is used in this study, due to its invariant properties under typical image geometrical transformations such as rotation and scaling. After image registration, the image mapping employs Random Sample Consensus (RANSAC) (Fischler and Bolles, 1981) to save computation time and finds a transformation parameter vector from the matching pairs. Based on the transformation parameter vectors, the source image is transformed and mapped onto the target image. Moreover,

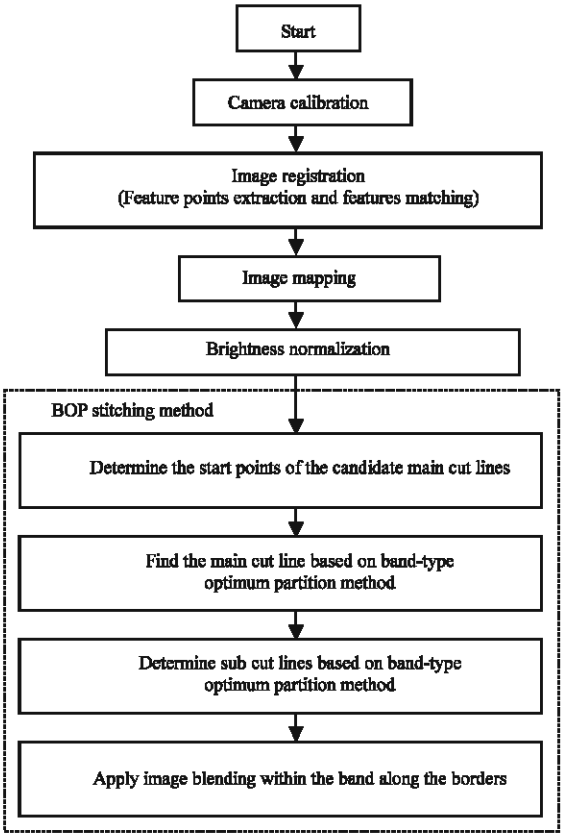


Fig. 1: Flowchart of the proposed image stitching process

Corresponding Author: Tien-Der Yeh, Electrical and Control Engineering, National Chiao Tung University, 1001 Ta Hsueh Rd., Hsinchu, Taiwan 300, ROC

since the mapping often produces unmapped points, also known as black holes (Kaczmarek *et al*., 2003) a reverse transformation model is used for compensation to make the transformation source image without any unmapped points. After the relation of the source and target images is obtained and the geometric information of the overlap region is collected, the proposed BOP method is employed to stitch the source and target images into a single image.

There are 2 major types of image stitching methods presented in literature. The methods of the 1st type, such as gradient blending (Vladan *et al*., 2005) and intensity adjustment (Hsieh, 2004), do not require any cut lines when dealing with the input images in the overlap regions. Although, these methods are not limited to the shape of the overlap region, they often result in the ghost image effect (Ashley and Richard, 2001), if any misalignment exists in the overlap region. The methods of the 2nd type, such as optimum partition (Jia and Tang, 2005) and minimum error boundary cut (Efros and Freeman, 2001), need to find a single cut line within the overlap region and process the image stitching based on it. Although, these methods can reduce the undesirable ghost image effect due to the use of the single cut line, it may generate significant discontinuities on the boundaries of the overlap regions if the color difference between the input images is large. The idea of the proposed method is derived from the above 2 statements and targets to solve the disadvantages in general image stitching.

Unlike the above 2 types of image stitching methods, the BOP method is proposed not only to minimize the ghost image effect but also to overcome the difficulty of processing the boundaries. By evaluating the smoothness and similarity costs on the source and target images, the BOP method determines multiple cut lines in the overlap region. It then applies the image blending along the cut lines to complete the image stitching process. Finally, experimental results are used to show that the proposed BOP method is indeed applicable to unpredictable overlap regions and no significant ghost image effect is observed.

MATERIALS AND METHODS

The whole process includes the Band-type Optimum Partition (BOP) stitching method and four pre-processings before it. As shown in Fig. 1, they are respectively camera calibration, feature points extraction and feature matching, image mapping and brightness normalization.

Camera calibration: The first step of the proposed process in Fig. 1 is restoring the distortion caused by camera lens. There are many methods in literature discussing camera calibration. In this study, the problem is solved by using the algorithm proposed by Heikkilä and Silvén (1997). It constructs the model of camera lens successfully and corrects the images from distortion according to the model.

Feature points extraction and features matching: After camera calibration, the next step is to detect and register the feature points, or say keypoints, of the source and target images by Lowe (2004) SIFT technique and then find the matching pairs by comparing the distance of the feature vectors. Note that the dimension of the feature vector is set to be 4×4×8 = 128 as same as lowe proposed.

Let the feature points of the source image detected by SIFT method form the set FP_S, where the i-th point, p_i, is located at coordinate (x_{pi}, y_{pi}) with associated feature vector V (p_i). The feature points of the target image form the set FP_T, where the kth point, q_* is located at coordinate (x_{pk}, y_{pk}) with associated feature vector V (q_k). To calculate the similarity between the elements of FP_S and FP_T, the Euclidean distance between points p_i and q_k is represented as,

$$L(p_i, q_k) = \left\| V(p_i) - V(q_k) \right\| \tag{2.1}$$

By fixing the s-th feature point $p_s \epsilon FP_s$, the corresponding matching point q_t is found by $q_t \epsilon FP_T$ and satisfies the following condition,

$$L(p_s, q_t) \le L(p_s, q_k),\ \forall q_k \in FP_T \tag{2.2}$$

Further define the matching cost U (s, t) as,

$$U(s,t) = L(p_s, q_t) \tag{2.3}$$

By setting a threshold value T_c for U (s, t), the set of matching pairs $MP_{S,T}$ can be obtained as

$$MP_{S,T} = \left\{ (p_s, q_t) \mid U(s,t) < T_c, p_s \in FP_S, p_t \in FP_T \right\} \tag{2.4}$$

The matching pair set $MP_{S,T}$ will be used for deriving the transformation parameters in next subsection.

Image mapping: To combine the source and target images into a single image, it is required to find their geometrical relationship and transform them onto the same coordinate. In this study, the source image is transformed to fit the target image according to a nonlinear planar transformation model proposed by Szeliski (1994). The

model uses only 4 matching pairs as a matching group to derive a set of transformation parameters, expressed as:

$$x'_i = \frac{m_0 x_i + m_1 y_i + m_2}{m_6 x_i + m_7 y_i + 1} \quad \text{and} \quad y'_i = \frac{m_3 x_i + m_4 y_i + m_5}{m_6 x_i + m_7 y_i + 1} \tag{3.1}$$

where (x_i, y_i) and (x'_i, y'_i), respectively represent the ith original and transformed coordinates of the image and m_j, $j = 0, 1, 2, \ldots, 7$, are the transformation parameters related to image translation, scaling and rotation. Rewrite Eq. 3.1 in a matrix form as

$$A.\, m = b \tag{3.2}$$

where, $m = [m_0, m_1, \ldots, m_7]^T$ is the transformation parameter vector and

$$A = \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1x'_1 & -y_1x'_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2x'_2 & -y_2x'_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x_3x'_3 & -y_3x'_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4x'_4 & -y_4x'_4 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1y'_1 & -y_1y'_1 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -x_2y'_2 & -y_2y'_2 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -x_3y'_3 & -y_3y'_3 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -x_4y'_4 & -y_4y'_4 \end{bmatrix}, \quad b = \begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \\ x'_4 \\ y'_1 \\ y'_2 \\ y'_3 \\ y'_4 \end{bmatrix} \tag{3.3}$$

If the inverse of A exists, the transformation parameter vector m can be derived as

$$m = A^{-1}.\, b \tag{3.4}$$

In other words, if a matching group (4 matching pairs) is selected, a corresponding transformation parameter vector m can be determined by Eq. 3.4.

Let N_m be the total number of matching pairs $MP_{S,T}$ in Eq. 2.4, thus, the number of possible combinations to form the matching groups is $N_g = C^{nm}_4$. Further let $m^k = [m^k_0, m^k_1, \ldots, m^k_7]$, k = 1 to N_g, be the kth transformation parameter vector obtained from the kth matching group by Eq. 3-4. In order to find, the most appropriate m^k, a function D (q_j, p'_j) is defined to measure the distance between the pixels of the transformed source image and the target image, expressed as

$$\begin{aligned} D(q_j, p'_j) &= \|q_j - p'_j\| \\ &= \sqrt{\left(x_{q_j} - x'_{p_j}\right)^2 + \left(y_{q_j} - y'_{p_j}\right)^2} \\ &= \sqrt{\left(x_{q_j} - \frac{m^k_0 x_{p_j} + m^k_1 y_{p_j} + m^k_2}{m^k_6 x_{p_j} + m^k_7 y_{p_j} + 1}\right)^2 + \left(y_{q_j} - \frac{m^k_3 x_{p_j} + m^k_4 y_{p_j} + m^k_5}{m^k_6 x_{p_j} + m^k_7 y_{p_j} + 1}\right)^2} \end{aligned} \tag{3.5}$$

where, q_j at coordinate (x_{qj}, y_{qj}) is the jth feature point in FP_T and p'_j at coordinate (x'_{pj}, y'_{pj}) is the transformed point from the jth feature point $p_j \in FP_S$, at coordinate (x'_{pj}, y'_{pj}) by Eq. 3.1. Further, by giving a threshold value T_m, a mapping counter c (m^k) is set to count the number of correct mapping pairs, expressed as

$$\begin{aligned} c(m^k) &= \sum_{j=1}^{N_m} G(q_j, p'_j), \\ G(q_j, p'_j) &= \begin{cases} 1, & \text{if } D(q_j, p'_j) \le T_m \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{3.6}$$

The desired transformation parameter vector m_d is chosen to be the m^k whose matching counter c $(m)^k$ is maximum. However, if >1 m^k results in the same maximum c (m^k), a mapping error function is further employed as

$$e(m^k) = \sum_{j=1}^{N_m} D(q_j, p'_j) \tag{3.7}$$

Then the desired transformation parameter vector m_d is chosen to be the m^k with minimum e (m^k).

Since the number of matching pairs is often far >4, it is very time-consuming to calculate all possible m^k and the corresponding c (m^k) and e (m^k). Instead, a methodology named Random Sample Consensus (RANSAC), which has been proposed by Fischler and Bolles (1981) is employed to reduce computation time in this study. The methodology indicates that if the number of total trials is N_t, then the probability P_o of finding at least one proper transformation model (4 pairs) after N_t trials is

$$P_o = 1 - \left(1 - \left(\frac{4}{N_m}\right)^4\right)^{N_t} \tag{3.8}$$

In other words,

$$N_t = \frac{\log(1 - P_o)}{\log\left(1 - (4/N_m)^4\right)} \tag{3.9}$$

By assigning a high accuracy rate $P_o = 0.99$, it can be shown that N_t in Eq. 3.9 is smaller than half of N_g, the total number of possible combinations when N_m>50.

Based on the obtained transformation parameter vector m_d, the source image can be transformed by the forward mapping model Eq. 3.1 pixel by pixel onto the coordinate of the target image. However, it might happen that some pixels on the transformed image are not determined and appear as black holes (Kaczmarek *et al*., 2003). In order to compensate these pixels, the reverse model derived from Eq. 3.1 is used and expressed as

$$x_u = -\frac{(m_5m_7 - m_4)x'_u + (m_2m_7 + m_1)y'_u + (m_2m_4 - m_1m_5)}{(m_3m_7 - m_4m_6)x'_u + (m_1m_6 - m_0m_7)y'_u + (m_0m_4 - m_1m_3)}$$
$$y_u = -\frac{(-m_5m_6 - m_3)x'_u + (m_2m_6 - m_0)y'_u + (m_0m_5 - m_2m_3)}{(m_3m_7 - m_4m_6)x'_u + (m_1m_6 - m_0m_7)y'_u + (m_0m_4 - m_1m_3)} \quad (3.10)$$

where (x'_u, y'_u) is the coordinate of the source image after transformation, while (x_u, y_u) is the coordinate before transformation.

Brightness normalization: In general, it is impossible to acquire the source and target images in the same color due to the changing lighting conditions or the unfixed camera parameters. This generates a significant border between two neighboring pictures in image stitching and degrades the quality of the combined image. In this study, the transformed source and target images are processed with the following brightness normalization. Let $Y_S(p_k)$ and $Y_T(p_k)$, k =1, 2, …, N_x, respectively be the brightness on kth pixel p_k of the transformed source image and the target image within the overlap region and N_x be the total number of pixels within the overlap region. Then the averaged brightness $\bar{Y}_S$ and $\bar{Y}_T$ for the transformed source and target images are calculated by

$$\begin{cases} \bar{Y}_S = \frac{1}{N_x}\sum_{k=1}^{N_x} Y_S(p_k) \\ \bar{Y}_T = \frac{1}{N_x}\sum_{k=1}^{N_x} Y_T(p_k) \end{cases} \quad (4.1)$$

Based on $\bar{Y}_S$ and $\bar{Y}_T$, the brightness difference $\tilde{Y}_S$ and $\tilde{Y}_T$ are

$$\begin{cases} \tilde{Y}_S = \bar{Y}_S - \frac{\bar{Y}_S + \bar{Y}_T}{2} = \frac{\bar{Y}_S - \bar{Y}_T}{2} \\ \tilde{Y}_T = \bar{Y}_T - \frac{\bar{Y}_S + \bar{Y}_T}{2} = \frac{\bar{Y}_T - \bar{Y}_S}{2} \end{cases} \quad (4.2)$$

and the normalized Y components are

$$\begin{cases} \hat{Y}_S(p_k) = Y_S(p_k) - \tilde{Y}_S \\ \hat{Y}_T(p_k) = Y_T(p_k) - \tilde{Y}_T \end{cases} \quad (4.3)$$

where, $\hat{Y}_S(p_k)$ and $\hat{Y}_T(p_k)$ are the normalized Y components of the kth pixel p_k in the 2 images after brightness normalization.

Image stitching: To reduce undesirable artifacts, some image stitching methods, such as gradient blending (Vladan *et al*., 2005) and intensity adjustment (Heieh, 2004), are proposed without any cut line to deal with overlap regions not necessarily in rectangular shape. However, they often result in undesired artifacts in form of ghost images or blurred images if some misalignments take place in the overlap region. Some other stitching methods, such as optimum partition (Jia and Tang, 2005) and minimum error boundary cut (Efros and Freeman, 2001), are proposed to find a single cut line within the overlap region and process the image stitching based on it. Although, these methods can reduce the ghost image effect due to the use of the single cut line, they are still limited to rectangular overlap regions and special alignment of the source and target images, such as Fig. 2a and b. The case that t he irregular shapes of overlap regions, such as Fig. 2c, generate additional edges after region assignment, such as Fig. 2d, is not suitable for these methods. However, these shapes frequently appear in general image stitching.

In order to reduce the ghost image effect and process unpredictable overlap regions, this study proposes a stitching method, called the band-type optimal partition method or BOP method in brief. Unlike the Jia and Tang (2005) optimum partition method in which finds a zero-thickness cut line to combine 2 input images together, the BOP method modifies the cost evaluation method by finding cut lines with thicknesses, which have the best similarity and smoothness within. This property minimizes the ghost image effect when the image blending is applied within the thickness of the cut lines. Furthermore, the BOP method introduces a new concept using multiple cut lines so that it considers more boundaries then traditional image blending when

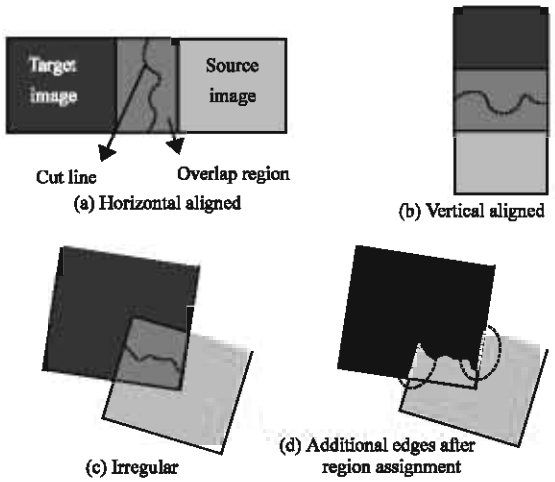


Fig. 2: The relationship of the source and target images

the overlap region is not in rectangular. Note that the first one determined from the multiple cut lines is called the main cut line.

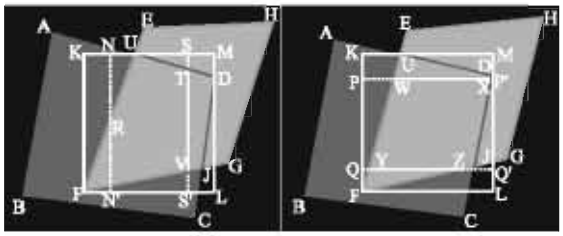
The procedure of the BOP method is shown in the dashed rectangle in Fig. 1. In the 1st step, the direction of the main cut line is determined according to horizontal and vertical edges within the overlap region. In the 2nd step, the multiple cut lines based on band-type optimal partition are determined one by one and become band-type borders between the transformed source image and the target image. Finally, the image blending method is employed within the band-type border to avoid ghost image effect within the large blending areas.

Starting points of candidate main cut lines: In the optimum partition method (Jia and Tang, 2005), the cut line is determined after calculating the costs of all possible paths within the overlap region. This leads to a large amount of calculation. In order to save the calculation time, the proposed BOP method finds four candidate main cut lines within the overlap region and chooses the one with lowest cost as the main cut line.

The 1st step of the BOP method is to determine 4 starting points g_m, m = 1-4, each of a candidate main cut line. To explain the procedure, let's consider a general case shown in Fig. 3 as an example, where the quadrangle UFJD is the overlap region and KFLM is the smallest rectangle covering the entire overlap region. Let the image J be the sum of the source and target images, i.e., $J = I_S+I_T$. With the well-known sobel masks

$$S_1=\begin{bmatrix}-1 & -2 & -1\\ 0 & 0 & 0\\ 1 & 2 & 1\end{bmatrix},\ S_2=\begin{bmatrix}-1 & 0 & 1\\ -2 & 0 & 2\\ -1 & 0 & 1\end{bmatrix},$$
$$S_3=\begin{bmatrix}-2 & -1 & 0\\ -1 & 0 & 1\\ 0 & 1 & 2\end{bmatrix}\text{ and }S_4=\begin{bmatrix}0 & -1 & -2\\ 1 & 0 & -1\\ 2 & 1 & 0\end{bmatrix} \quad (5.1)$$

the gradient images $\nabla_x J$ and $\nabla_y J$ with respect to x-axis and y-axis within the overlap region are, respectively $\nabla_x J = J*S_1$ and $\nabla_y J = J*S_2$. Further, locate the pentagon URIVT in Fig. 3a, which is the intersection of the overlap region and rectangle NSS'N', the middle part in vertical direction of the rectangle KFLM with $\overline{NS} = 2/3\overline{KM}$. Similarly, locate the pentagon WYZJX in Fig. 3b, which is the intersection of the overlap region and rectangle PQQ'P', the middle part in horizontal direction of the rectangle KFLM with $\overline{PQ} = 2/3\overline{KF}$. Moreover, let $\nabla_x J^a$ and $\nabla_y J^a$, parts of $\nabla_x J$ and $\nabla_y J$, be the gradient images with respect to the pentagon URIVT in Fig. 3a and $\nabla_x J^b$ and $\nabla_y J^b$, also parts of $\nabla_x J$ and $\nabla_y J$, be the gradient images with respect to the pentagon



(a) The vertical main cut line **(b) The horizontal main cut line**

Fig. 3: The possible region for the path

WYZJX in Fig. 3b. Then, by setting threshold T, define N^A_x and N^A_y, $A \in \{a, b\}$, as the number of the pixels of $\nabla_x J^A$ and $\nabla_y J^A$, whose values are larger than T. After N^a_x, N^a_y, N^b_x and N^b_y are obtained, it is necessary to determine whether the starting points g_m is chosen from the case of Fig. 3a or b. If $N^a_x < N^a_y$ and $N^b_x < N^b_y$, which means the number corresponding to the vertical edge detector S_2 is larger, then Fig. 3a is chosen since the main cut line tends to move vertically. If $N^a_x > N^a_y$ and $N^b_x > N^b_y$, which means the number corresponding to the horizontal edge detector S_1 is larger, then Fig. 3b is chosen since the main cut line tends to move horizontally. For the other conditions, the decision is made by comparing $N^a_x/N^a_x+N^a_y$ and $N^b_y/N^b_x+N^b_y$. If $N^a_x/N^a_x+N^a_y \geq N^b_y/N^b_x+N^b_y$, then Fig. 3b is chosen; otherwise Fig. 3a is chosen. When the case of Fig. 3a is selected, the starting points g_m are located at an upper edge of pentagon URIVT, i.e., $\overline{UR}$ or $\overline{UT}$ in this example. When the case of Fig. 3b is selected, the starting points g_m is located at a left edge of pentagon WYZJX, i.e., $\overline{WY}$ in this example.

Let us assume the 4 starting points are on the upper edges, $\overline{UR}$ and $\overline{UT}$, in Fig. 3a and redraw it in Fig. 4a. We denote the coordinates of N, U and S as (x_N, y_N), (x_U, y_U) and (x_S, y_S). Clearly, their y-coordinates are the same. Define λ = round $(x_S + x_N/8)$ and the coordinates of the points p_n on $\overline{NS}$ are (x_{pn}, y_{pn}), where $x_{pn} = x_N + 2\lambda + n$ for $n = 0\text{-}4\lambda$ and $y_{pn} = y_U$. Let the point q_n with coordinate (x_{qn}, y_{qn}) be the projection of p_n on the edges $\overline{UR}$ and $\overline{UT}$. It is obvious that $x_{qn} = x_{pn}$. Now, similar to Jia and Tang (2005) method, we define the gradient-level alignment costs of q_n as

$$C(q_n) = (1-\beta)C_g(q_n) + \beta C_d(q_n),\ 0 \leq \beta \leq 1 \quad (5.2)$$

β is chosen as 0.4 in this thesis. The gradient-level smoothness cost $C_g(q_n)$ is defined as

$$C_g(q_n) = \sum_{i=-1}^{1}\sum_{j=-1}^{1} \nabla I_S\left(x_{q_n}+i, y_{q_n}+j\right) + \sum_{i=-1}^{1}\sum_{j=-1}^{1} \nabla I_T\left(x_{q_n}+i, y_{q_n}+j\right)$$

$$= \sum_{i=-1}^{1}\sum_{j=-1}^{1}\sqrt{\sum_{p=1}^{4}\left|I_S\left(x_{q_n}+i, y_{q_n}+j\right)*S_p\right|^2} + \sum_{i=-1}^{1}\sum_{j=-1}^{1}\sqrt{\sum_{p=1}^{4}\left|I_T\left(x_{q_n}+i, y_{q_n}+j\right)*S_p\right|^2} \quad (5.3)$$

where, * is the convolution operator and S_1 to S_4 are the Sobel masks as in Eq. 5.1. The gradient-level similarity cost $C_d(q_n)$ is defined as

$$C_d(q_n) = \sum_{i=-1}^{1}\sum_{j=-1}^{1}\left\|\left(I_S\left(x_{q_n}+i, y_{q_n}+j\right) - I_T\left(x_{q_n}+i, y_{q_n}+j\right)\right)*S_1\right\| + \sum_{i=-1}^{1}\sum_{j=-1}^{1}\left\|\left(I_S\left(x_{q_n}+i, y_{q_n}+j\right) - I_T\left(x_{q_n}+i, y_{q_n}+j\right)\right)*S_2\right\| \quad (5.4)$$

The 4 starting points g_m in Fig. 4b of candidate main cut lines are, respectively selected from the set

$$\Psi_m = \{q_n,\ n=(m-1)\lambda+1,\ (m-1)\lambda+2,\ \ldots,\ m\lambda\}$$

i.e., $g_m \in \Psi_m$ for m =1-4. By calculating all the gradient alignment costs C (q_n) in Ψ_m, the starting points g_m is determined to possess the minimum cost C (g_m).

Main cut line based on BOP method: The main cut line will be chosen from 4 candidates main cut lines each starting from one of the starting points, g_1, g_2, g_3 and g_4 in Fig. 4b. Each candidate main cut line is determined point-by-point according to the gradient-level alignment costs Eq. 5.2 in the proposed BOP method. The point-by-point process is concerned with an iteration procedure from current point V_c with coordinate (x_c, y_c) to next point V_n with coordinate (x_n, y_n), where, V_n is chosen from three candidates V_{n1}, V_{n2} and V_{n3}. Figure 5a shows the 3 candidates of a vertical main cut line, where the coordinates of V_{n1}, V_{n2} and V_{n3} are (x_c-1, y_c+1), (x_c, y_c+1) and (x_c+1, y_c+1). For a horizontal main cut line, the coordinates of V_{n1}, V_{n2} and V_{n3} are (x_c+1, y_c-1), (x_c+1, y_c) and (x_c+1, y_c+1), as shown in Fig. 5b.

To determine the next point V_n in the point-by-point process, it is required to calculate the accumulated gradient-level alignment costs of V_{n1}, V_{n2} and V_{n3}, respectively denoted as A (V_{n1}), A (V_{n2}) and A (V_{n3}). By the use of Eq. 5.2, the accumulated costs of A (V_{n1}), A (V_{n2}) and A (V_{n3}) are defined as

$$A(V_{n1}) = \sum_{j=1}^{7} C(V_{ej}),\ A(V_{n2}) = \sum_{j=2}^{8} C(V_{ej}),\ A(V_{n3}) = \sum_{j=3}^{9} C(V_{ej}) \quad (5.5)$$

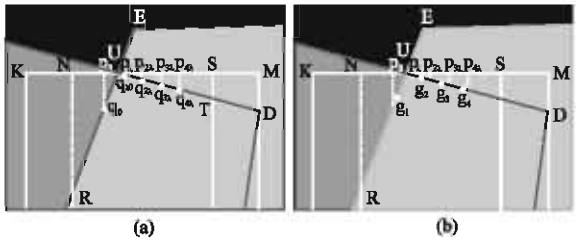


Fig. 4: Zoomed view of Fig. 3 (a) Segmentation of upper edges (b) location of starting points

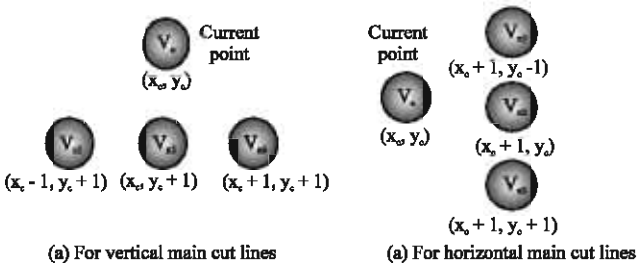


Fig. 5: Candidates of next point

where, V_{ej}, j = 1-9, are neighboring points of the three candidates as shown in Fig. 6. Note that V_{n1}, V_{n2} and V_{n3}, are equivalent to the points V_{e4}, V_{e5} and V_{e6}. For a vertical cut line, the coordinates of V_{ej} are (x_c-4+j, y_c+1) and for a horizontal cut line, the coordinates of V_{ej} are (x_c+1, y_c-4+j). Then the point V_n is chosen from the three candidates V_{n1}, V_{n2} and V_{n3} by comparing their accumulated costs A (V_{n1}), A (V_{n2}) and A (V_{n3}). The determination algorithm is given as

$$V_n = \begin{cases} V_{n2}, & \text{if } A(V_{n2}) \le \min\{A(V_{n1}), A(V_{n3})\} \\ V_{n3}, & \text{if } A(V_{n3}) < \min\{A(V_{n1}), A(V_{n2})\} \\ V_{n1}, & \text{otherwise,} \end{cases} \quad (5.6)$$

which means the point V_n possesses the smallest accumulated cost. Besides, if two or more candidates have the same accumulated costs, the priority to be selected is in the order of V_{n2}, V_{n1} and V_{n3}.

Since the candidate main cut lines are determined in an iterative procedure, the determination algorithm (5.6) is further modified as

$$V_n(m,j) = \begin{cases} V_{n2}(m,j), & \text{if } A(V_{n2}(m,j)) \\ & \le \min\{A(V_{n1}(m,j)), A(V_{n3}(m,j))\} \\ V_{n3}(m,j), & \text{if } A(V_{n3}(m,j)) \\ & < \min\{A(V_{n1}(m,j)), A(V_{n2}(m,j))\} \\ V_{n1}(m,j), & \text{otherwise} \end{cases} \quad (5.7)$$

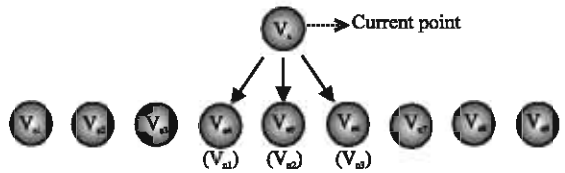


Fig. 6: Evaluate of accumulated costs of next points

where, m indicates the mth candidate main cut line and j denotes the jth iteration. In other words, the point-by-point process in the j-th iteration of the m-th candidate main cut line is applied to the current point V_c (m, j) to choose its next point V_n (m, j). When the procedure starts, the initial current point V_c (m, 1) is assigned to be g_m, m = 1-4, in Fig. 4b and its next point V_n (m, 1) is determined according to Eq. 5.7. When j≥2, the point V_n (m, j-1) is assigned as the current point V_c (m, j) of the j-th iteration. The iteration procedure stops when the V_n (m, j) reaches the boundary of the overlap region. Following the procedure, the average gradient-level alignment cost of the mth candidate main cut line is obtained as

$$L(m) = \frac{\sum_{j=1}^{J_m} A\left(V_n(m,j)\right)}{J_m} \qquad (5.8)$$

where, J_m is the total number of iterations; A (V_n (m, j)) is the accumulated gradient-level alignment cost determined from Eq. 5.5 and 5.7. Finally, the main cut line will be chosen as the candidate main cut line with the lowest average gradient-level alignment cost L (m).

Note that the main feature of the proposed BOP method is the cost evaluation in Eq. 5.5 of its 3 candidates of next point. The cost of each candidate is accumulated from itself and its 6 neighboring points. The way to include neighboring points is helpful for image blending quality, which will be demonstrated in the experimental results later.

Determination of sub cut lines: In addition to the main cut line, there are 2 sets of sub cut lines needed to reduce the discontinuities in the overlap region and achieve better image blending quality. Each set is composed of 2 sub cut lines and also determined by the proposed BOP method.

Consider Fig. 7 as an example with a horizontal main cut line Ψ_0, which has been determined by the BOP method. Besides, the overlap region EKCI is divided by the main cut line into 2 parts, one for the transformed source image and the other for the target image. It is apparent that $\overline{MK}$, $\overline{NI}$ and the main cut line Ψ_0 form the boundary between the source and target images.

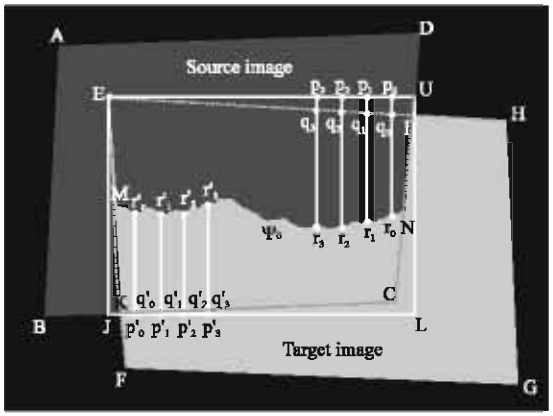


Fig. 7: The discontinuities after determination of the main cut line

However, the 2 edges $\overline{MK}$ and $\overline{NI}$, unlike the main cut line possessing best similarity and smoothness, may degrade image blending quality due to serious discontinuities. In order to avoid the defect, the edges $\overline{MK}$ and $\overline{NI}$ will be replaced by 2 sub cut lines, which are also achieved with best similarity and smoothness and serve as new borders between the transformed source and target images for better image blending quality.

Similar to the determination of main cut line, it is required to determine the starting points of candidate sub cut lines in the beginning. At 1st, the starting points of the first set candidate sub cut lines in Fig. 7 are located in the intervals $\overline{q_0q_1}$, $\overline{q_1q_2}$ and $\overline{q_2q_3}$, where q_i, i = 0-3, are the points on the upper edge $\overline{EI}$ of the overlap region. Note that q_i is chosen to have the same x-coordinate with p_i, where, p_i is located on the rectangle EJLU, the smallest rectangle covering the entire overlap region. Besides, the 4 points p_i, i = 0-3, satisfy

$$\overline{Up_0} = \overline{p_0p_1} = \overline{p_1p_2} = \overline{p_2p_3} = \frac{1}{13}\overline{EU}$$

Similarly, the starting points of the second set candidate sub cut lines are located in the intervals

$$\overline{q'_0q'_1},\ \overline{q'_1q'_2} \text{ and } \overline{q'_1q'_3}$$

and p'$_i$, i = 0-3, satisfy

$$\overline{Jp'_0} = \overline{p'_0p'_1} = \overline{p'_1p'_2} = \overline{p'_2p'_3} = \frac{1}{13}\overline{EU}$$

Let g_m and g'$_m$, m = 1-3, be the starting points within the regions $\overline{q_{m-1}q_m}$ and $\overline{q'_{m-1}q'_m}$ and found based on the lowest gradient-level alignment cost according to Eq. 5.2.

After the starting points g_m and g'_m, m = 1-3, are determined, the two set of candidate sub cut lines can be obtained also by the same determination procedure of main cut line. Figure 8a shows two sets of candidate sub cut lines $\{\Psi_{11}, \Psi_{12}, \Psi_{13}\}$ and $\{\Psi_{21}, \Psi_{22}, \Psi_{23}\}$, as an example. Each candidate sub cut line starts from a starting point and stops at the main cut line Ψ_0, such as Ψ_{11} and Ψ_{12}, or the boundary of the overlap region, such as Ψ_{13}. Sometimes, some candidate sub cut lines may merge together, which can be seen from Ψ_{21} and Ψ_{22}. Finally, the first sub cut line Ψ_1 and the second sub cut line Ψ_2 are respectively chosen from $\{\Psi_{11}, \Psi_{12}, \Psi_{13}\}$ and $\{\Psi_{21}, \Psi_{22}, \Psi_{23}\}$ for possessing lowest average gradient-level alignment cost Eq. 5.7. Figure 8b shows an example where Ψ_{12} and Ψ_{23} are chosen as Ψ_1 and Ψ_2, respectively.

In Fig. 8b, after Ψ_1 and Ψ_2 are selected, it can be found that 1 discontinuous edges $\overline{Iq_2}$ and $\overline{Kq'_3}$ still exist and may degrade image blending quality. To further improve the quality, the same procedure of determining Ψ_1 and Ψ_2 can be repeated to find two horizontal sub cut lines. Since the discontinuous edges become smaller, the horizontal sub cut lines can be directly determined without candidate sub lines. The starting point of the right sub cut line Ψ_3 is located in the interval $\overline{v_0v_1}$, where v_0 and v_1 have the same y-coordinates with u_0 and u_1. Note that u_0 and u_1 are located on the smallest rectangle EJLU and satisfy

$$\overline{Uu_0} = \overline{u_0u_1} = \frac{1}{13}\overline{UL}$$

The same procedure can be applied to determine the left sub cut line Ψ_4. Once the starting points are determined, the sub cut lines Ψ_3 and Ψ_4 can be obtained by the proposed BOP method accordingly. Figure 9a shows the determination result of the main cut line and 4 sub cut lines and Fig. 9b shows the selected borders between the source and target images.

Image blending: Recap that the Band-type Optimum Partition (BOP) method results in the best smoothness and color similarity within the band of cut lines. This characteristic is particularly helpful for the image blending quality of the band along the cut lines. Especially, apply the image blending within the defined width of the band significantly reduces ghost image effect due to the blending area is relatively small in the output image. In this subsection, different implementation methods for image blending will be discussed for different types of borders. The difference of the borders can be found in 3 aspects; 1st, a border can be either cut lines or boundaries of the overlap region; 2nd, a border can be in either horizontal direction or vertical direction. Let's see an example in Fig. 9b, where the borders $h_3'h_2'$, $w_1'z_1'$ and z_1w_1

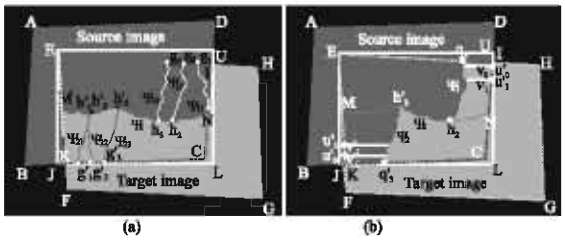


Fig. 8a: The candidate sub cut lines (b) The selected sub cut lines

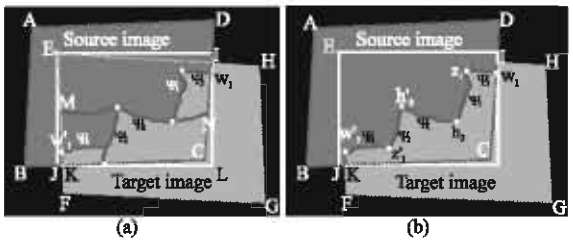


Fig. 9: The cut lines and borders between source and target images (a) The main cut line and four sub cut lines, (b) Borders after remove useless cut lines

are parts of the horizontal cut lines, the borders $h_3'z_1$ and z_1h_2 are parts of the vertical cut lines and the borders $w_1'K$, Iw_1 are parts of vertical boundaries of the overlap region.

Before image blending, it is worth introducing the blending coefficient η (b) for pixel b in the overlap region. Let (R_S, G_S, B_S) be the color components of pixel b in the source image and (R_T, G_T, B_T) be the color components of pixel b in the target image. The resulted color components (R_b, G_b, B_b) of pixel b after image blending is then expressed as

$$\begin{aligned} R_b &= 0.125\left(\eta(b)\times R_T + \left(8-\eta(b)\right)\times R_S\right) \\ G_b &= 0.125\left(\eta(b)\times G_T + \left(8-\eta(b)\right)\times G_S\right) \\ B_b &= 0.125\left(\eta(b)\times B_T + \left(8-\eta(b)\right)\times B_S\right) \end{aligned} \quad (5.9)$$

Besides, the blending coefficient η (b) ranges from 0-8, which is determined by λ (b), the distance from pixel b to the border. Next, let's focus on the discussion of λ (b).

The method to find all λ (b) for b in the band along the borders is first to define the type of each pixel on the borders and then determine the distances of the pixels in its neighborhood. The first type of the pixels on the borders is the corner points, such as the pixels w_1', z_1', h_3', h_2, z_1 and w_1 in Fig. 9b. For simplicity, let the center pixel n at (x_n, y_n) in Fig. 10 be a corner point and assume pixel b is a neighboring pixel at $(x_n + i, y_n + j)$, $-3 \le i \le 3$, $-3 \le j \le 3$. Then, the distance from pixel b to pixel n is defined as

$\lambda(b) = |i| + |j|$ and a pixel b with $\lambda(b) > 3$ will be ignored as shown in Fig. 10. Note that the obtained $\lambda(b)$ related to the corner point n may be overwritten and then changed later.

Besides the corner points, the pixels on the borders can be classified into four different types, which are horizontal cut lines, vertical cut lines, horizontal boundaries of the overlap region and vertical boundaries of the overlap region. For the borders formed by horizontal cut lines, such as $h_3'h_2$, $w_1'z_1'$ and z_1w_1 in Fig. 9b, let the pixel on the border be pixel h at (x_h, y_h) as depicted in Fig. 11a and then the distance of pixel b at $(x_h, y_h + j)$, $-3 \le j \le 3$, is determined as $\lambda(b) = |j|$. Similarly, for the borders formed by vertical cut lines, such as $h_3'z_1$ and z_1h_2 in Fig. 9b, let the pixel on the border be pixel v at (x_v, y_v) as depicted in Fig. 11b and then the distance of pixel b at (x_v+i, y_v), $-3 \le i \le 3$, is $\lambda(b) = |i|$. For the vertical boundaries of the overlap region, such as $w_1'K$ and Iw_1 in Fig. 9b, let the pixel on the boundary be pixel u at (x_u, y_u) as depicted in Fig. 11c whose overlap region is on the left and then the distance of pixel b at (x_u+i, y_u), $-3 \le i \le 0$, is $\lambda(b) = |i|$. Figure 11d gives an example of horizontal boundary with upper overlap region. Let the pixel on the boundary be pixel g at (x_g, y_g), then the distance of pixel b at (x_g, y_g+j), $-3 \le j \le 0$, is $\lambda(b) = |j|$.

An example to demonstrate the process of determining $\lambda(b)$ is given in Fig. 12, where the points from a_0 to a_{27} are on the borders. Clearly, there are 4 corner points, a_3 at (0, 4), a_6 at (5, 4), a_{10} at (5, 8) and a_{22} at (17, 4). The sets of $\{a_4, a_5\}$ and $\{a_{11}, a_{12}, a_{13}, \ldots, a_{21}\}$ are 2 horizontal cut lines, the sets $\{a_7, a_8, a_9\}$ and $\{a_{23}, a_{24}, a_{25}, \ldots, a_{27}\}$ are 2 vertical cut lines and the set $\{a_0, a_1, a_2\}$ is a boundary of the overlap region. Let's focus on the pixel b at (4, 5), a point related to four border points a_3, a_6, a_5, a_7, to see how $\lambda(b)$ is determined during the process. First, the process starts from the corner point a_6, the distance between pixel b at (4, 5) and a_6 is $\lambda(b) = 2$. Then the same process goes to the corner point a_3, which should lead to $\lambda(b) = 3$. However, this distance is longer than the previous $\lambda(b) = 2$ related to a_6. Therefore, the distance $\lambda(b) = 2$ is kept and no overwriting happens. After finishing the corner points, the process will apply to the vertical cut lines, point by point. When a_7 on the vertical cut line is processed, the distance from pixel b to a_7 is $\lambda(b) = 1$, which is shorter than the previous $\lambda(b) = 2$. As a result, overwriting happens and the distance is changed into $\lambda(b) = 1$. Then, the process will further apply to the horizontal cut lines, point by point. When a_5 on the horizontal cut line is processed, the distance is $\lambda(b) = 1$, same as the previous one and $\lambda(b)$ is unchanged. Finally, the distance of pixel b at (4, 5) is set as $\lambda(b) = 1$. The above processes demonstrate the determination of $\lambda(b)$ when pixel b related to more than one border points.

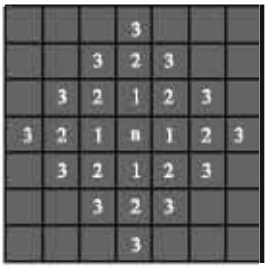


Fig. 10: Distances from a corner point n to its neighboring points

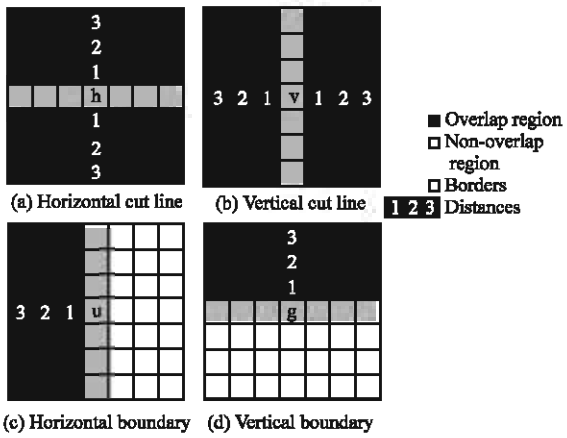


Fig. 11: $\lambda(b)$ for different types of borders

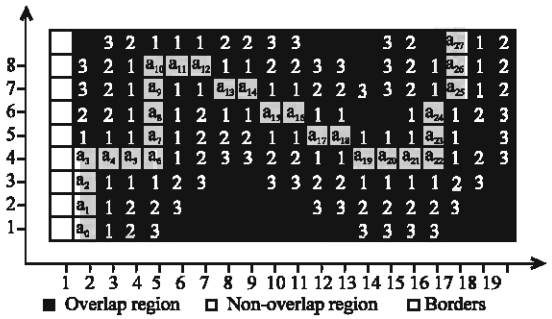


Fig.12: The border and distances $\lambda(b)$ of its neighboring pixels

Once all the distances are obtained within the band of the border, the blending coefficients in Eq. (5.9) can be determined as

$$\eta(b) = 4 + k \cdot \lambda(b) \tag{5.10}$$

where, k =1 or -1. If pixel b is assigned to the target image after region assignment, then k=-1, otherwise k = 1. Note that $\lambda(b) = 0$ for b on the borders and thus $\eta(b) = 4$, i.e., the blending coefficients of all the pixels on the borders uniformly equal to 4.

RESULTS AND DISCUSSION

There are many image stitching methods presented for different application requirements. As one of them, the proposed Band-type Optimum Partition (BOP) method is different from others because of the two properties it has. First, it ensures the combined image quality on the boundaries of the overlap region due to the original boundaries are replaced by the multiple cut lines which own the best similarity and smoothness in the overlap region. Second, the image blending within the band along the cut lines provides good quality for image blending and eliminates ghost-image effect.

One of the applications using image stitching is scenery resolution enhancement. Due to the limited resolution of a camera, a single picture captured from it suffers from the sacrifice of either a wide scene or the contained details. In other words, it is impossible to capture a wide scene into a single picture without loss of its details. Fortunately, the image stitching technique can breakthrough this limitation and combine the individual pictures with details into a single panorama.

Figure 13 gives an example of scenery enhancement by the proposed method. Due to the page size limitation, the source images and the aligned image are shown in downscale Fig. 13a-c, while the magnified scales of the overlap regions are shown in Fig. 13d-e. Figure 14 compares the stitching result of the proposed BOP method with one of the most common used method, the scene. multi-band blending. It can be seen that almost no artifact can be observed on the proposed BOP method because of it blends the 2 source images only in the band-type area of best similarity and smoothness. The image quality outside the band-type area is kept unchanged which might avoid artifacts caused by source images' mismatches or registration error. On the other hand, for the result of multi-band blending, it can be observed that the edges of the trees are somewhat blurred due to its seamless blending method.

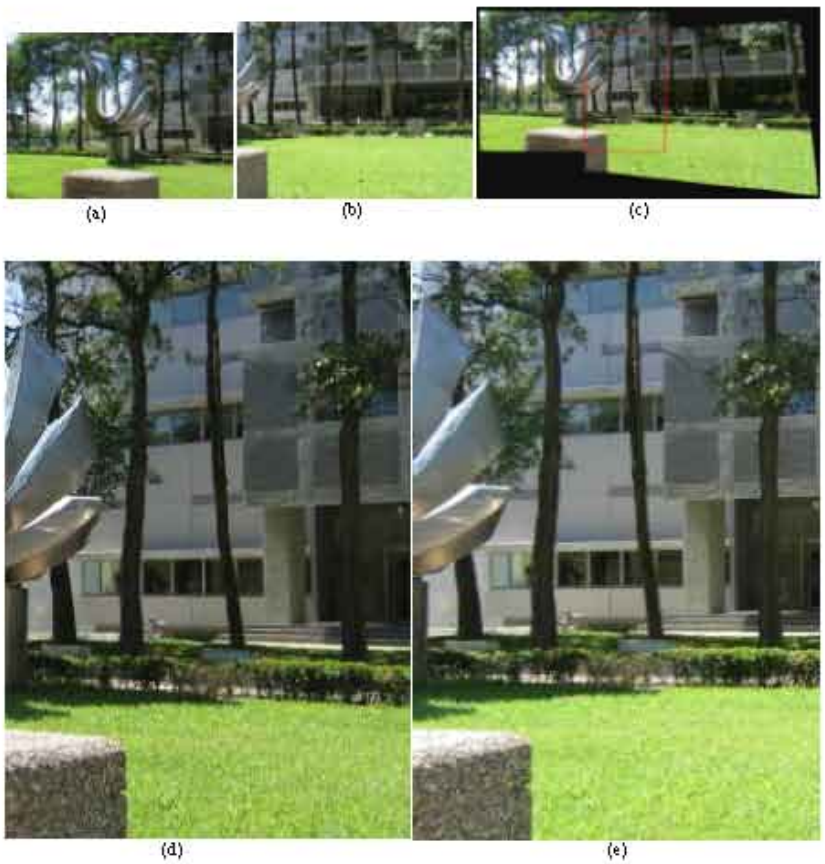


Fig. 13a: The first source image (b) The second source image (c) The aligned image with overlap region enclosed in red quadrangle (d) Magnified view of the first image around the overlap region (e) Magnified view of the second image around the overlap region

(a) (b)

Fig. 14a: The stitching result of the proposed method (b) The stitching result using multi-band blending

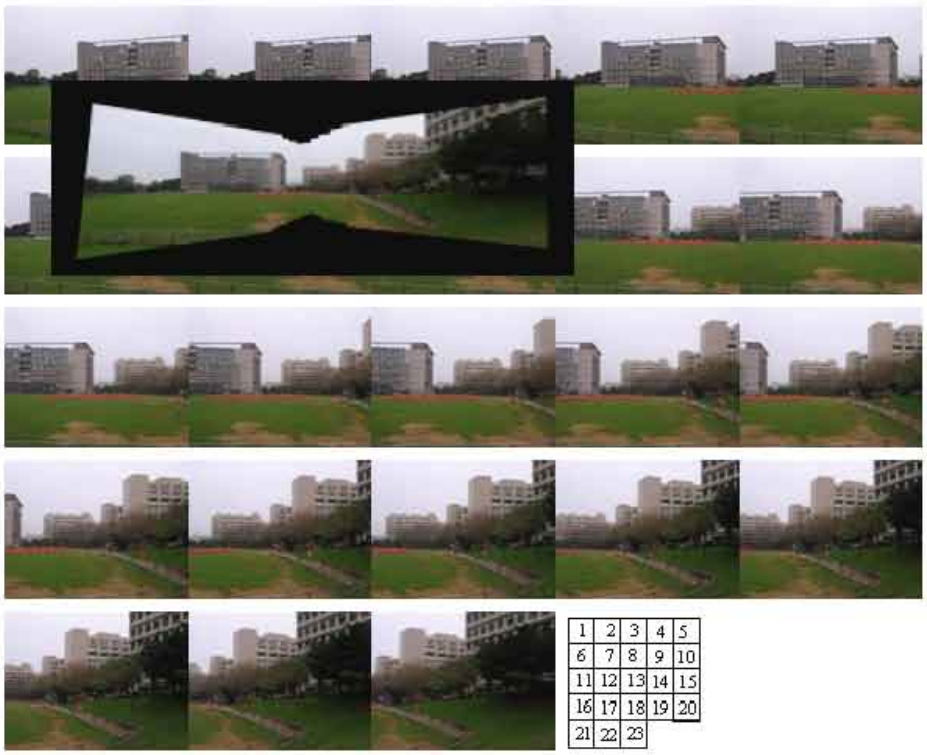


Fig. 15: The snapshots from a video clip. The numbers at the bottom row is the stitching sequence of the corresponding snapshot

The other one application is to rebuild the original scene from a series of video snapshot. Figure 15 is an example of this where the input video is taken in our campus. Figure 16 shows the stitching result. It can be seen that the proposed BOP method successfully rebuilds the scene.

Fig. 16: The stitching results from snapshots in Fig.15

CONCLUSION

A complete systematic process of image stitching has been presented in this study, including camera calibration, image registration, image mapping, brightness normalization and the proposed BOP method. This study mainly focuses on the BOP method, which is developed to deal with unpredictable overlap regions and ghost image effect. The BOP method adopts the concept of multiple cut lines, including one main cut line and 4 sub cut lines, for overlap regions in unpredictable shapes. Each cut line is determined according to the lowest smoothness and similarity costs. These cut lines are served as the boundaries to optimally partition the source and target images. Most importantly, the BOP method is processed in a band, not just along a cut line, which improves the image blending and reduces ghost image effect. The experimental results show that the BOP method can be applied to the overlap regions not in rectangular shapes and no ghost image effect exists on the final panorama.

REFERENCES

Ashley, M.U.E. and S. Richard, 2001. Eliminating Ghosting and Exposure Artifacts in Image Mosaics. In: Proc. Conf. Computer Vision and Pattern Recognition, pp: II:509-II:516. 11-13 Dec. DOI: 10.1109/CVPR.2001.991005. http://www.cs.berkeley.edu/~eden/Deghosting.pdf.

Brown, M. and D.G. Lowe, 2003. Recognising panoramas. Proc. IEEE. Int. Conf. Comput. Vision, Vancouver, BC, Canada, 2: 1218-1225. DOI: 10.1109/ICCV.2003.1238630. http://www.cs.ubc.ca/~mbrown/papers/iccv2003.pdf.

Efros, A.A. and W.T. Freeman, 2001. Image quilting for texture synthesis and transfer. International Conference on Computer Graphics and Interactive Techniques, ACM Press. New York, pp: 341-346. 12-17. http://graphics.cs.cmu.edu/people/efros/research/quilting/quilting.pdf.

Fischler, M.A. and R.C. Bolles, 1981. Random Sample Consensus: A Paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM., 24 (6): 381-395.

Han, B. and X. Lin, 2006. A Novel Hybrid Color Registration Algorithm for Image Stitching Applications. IEEE Trans. Consumer Electronics, 52 (3): 1129-1134. DOI: 10.1109/TCE.2006.1706518. INSPEC: 9103380. http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1706518&isnumber=36021.

Heikkilä, J. and O. Silvén, 1997. A 4-step camera calibration procedure with implicit image correction. Proceedings of IEEE Computer Society Conference on Computer Vision and Patten Recognition, San Juan, Puerto Rico, pp: 1106-1112. 17-19 June 1997. DOI:10.1109/CVPR.1997.609468.INSPECT:5644297, http://www.vision.caltech.edu/bouguetj/calib_doc/papers/heikkila97.pdf.

Hsieh, J.W., 2004. Fast stitching algorithm for moving object detection and mosaic construction. Image and Vision Comput., 22 (4): 291-306. DOI: 10.1016/j.imavis.2003.09.018. http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6V09-4BDCC1F-1&_user=1194694&_rdoc=1&_fmt=&_orig=search&_sort=d&view=c&_acct=C000051941&_version=1&_urlVersion=0&_userid=1194694&md5=f8b1ac808945d4ea8edec7de3088e592.

Jia, J. and C.K. Tang, 2005. Eliminating structure and intensity misalignment in image stitching. Proceedings of the 10th IEEE Int. Conference on Computer Vision, Beijing, China, IEEE, 2: 1651-1658. 17-21 Oct. DOI: 10.1109/ICCV.2005.87. INSPEC: 8824363. http://www.cse.cuhk.edu.hk/~leojia/all_final_papers/image_stitching_iccv05.pdf.

Kaczmarek, K., B. Walczak, S. de Jong and B.G.M. Vandeginste, Comparison of Image-Transformation Methods Used In Matching 2D Gel Electrophoresis Images. Acta Chromatographica, No. 13, 2003. http://www.us.edu.pl/uniwersytet/jednostki/wydzialy/chemia/acta/acl3/zrodla/01_AC13.pdf.

Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision, 60(2):91-110. DOI: 10.1023/B:VISI.0000029664.99615.94. http://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf.

Szeliski, R., 1994. Image mosaicing for tele-reality applications. IEEE. Workshop on Applications of Computer Vision, Sarasota Florida, pp: 44-53. 5-7 Dec. DOI: 10.1109/ACV.1994.341287. INSPEC: 4852561. http://www.hpl.hp.com/techreports/Compaq-DEC/CRL-94-2.pdf.

Zitova, B. and J. Flusser, 2003. Image registration methods: A survey. Image and Vision Computing, 21(11):977-1000.DOI:10.1016/S0262-8856(03)00137-9. http://library.utia.cas.cz/prace/20030125.pdf.