Chapter 4
Image Mapping

4.1 Introduction

In the previous chapter, the feature points in two input images can be extracted and the matching pairs between them will be found by using the normalized cross-correlation method. Then, an image mapping method should be used to effectively map the source image $\boldsymbol{I}_1$ to the target image $\boldsymbol{I}_2$ based on the information obtained from the matching pairs.

However, there often exist some false matching pairs to make the mapping worse. Hence, a method called the planar image mosaics [13][14] has been proposed and used to determine a matching group that contains the most accurate four matching pairs and then form the mapping model based on the matching group. Although the mapping model is simple due to the use of only four matching pairs, it is still a problem to adopt the planar image mosaics since it is required to determine the matching group from a large amount of matching pairs obtained previously, usually more than ten thousands. To improve the drawback of time consuming, this thesis further employs the random sample consensus [4], which will highly shrink the number of matching groups, sometimes into half of the total amount. Finally, image compensation is applied to average the overlapped pixels and fill up the unmapped pixels within the reasonable image region caused by the image scaling or rotation.

4.2 Optimal Mapping

After getting the matching pairs each composed of two feature points, one in the source image and the other in the target image, it is required to determine a mapping model of all the matching pairs such that the feature points in the source image can be successfully mapped to the relative ones in the target image. However, different source and target images often result in different numbers of matching pairs with the use of the proposed robust feature extraction and feature matching methods and thus it is hard to find a general model to fit in with all the matching pairs. Furthermore, the matching pairs obtained by normalized cross-correlation method may still inevitably contain some false matching pairs. For this reason, if all the matching pairs are used as data input for modeling, the existence of false matching pairs must influence the correctness of model and the mapping would be failed. Consequently, the modeling method had better possess less and fixed number of matching pairs as data input to achieve efficient mapping between source and target images by translation, scaling and rotation. After that, the overlap region will be found and adopted for an image stitching technique to get well image stitching performance.

In 1994, Szeliski proposed a nonlinear planar transformation model [13] which uses only four matching pairs as a matching group to derive a set of mapping parameters. The proposed model can be expressed as

$$x' = \frac{m_0 x + m_1 y + m_2}{m_6 x + m_7 y + 1} \text{ and } y' = \frac{m_3 x + m_4 y + m_5}{m_6 x + m_7 y + 1} \tag{4.2-1}$$

where (x, y) and (x', y') respectively represent the original and mapped coordinates of the source image and $\boldsymbol{m} = [m_0, m_1, \ldots, m_7]^T$ that consists of mapping parameters, m_i, $i = 0,1,\ldots,7$, related to image translation, scaling and rotation. Rewrite (4.2-1) in a matrix form as

$$\begin{bmatrix} \boldsymbol{a}_x \\ \boldsymbol{a}_y \end{bmatrix} \cdot \boldsymbol{m} = \begin{bmatrix} b_x \\ b_y \end{bmatrix} \tag{4.2-2}$$

where $\boldsymbol{a}_x = [x \quad y \quad 1 \quad 0 \quad 0 \quad 0 \quad -xx' \quad -yx']$, $\boldsymbol{a}_y = [0 \quad 0 \quad 0 \quad x \quad y \quad 1 \quad -xy' \quad -yy']$, $b_x = x'$ and $b_y = y'$ are determined from different matching pairs. In order to solve the mapping parameter vector ***m***, randomly choose four matching pairs to obtain $\boldsymbol{a}_{xi}, \boldsymbol{a}_{yi}, b_{xi}, b_{yi}$, $i = 1,2,3,4$, respectively. Thus, (4.2-2) can be extended as

$$\boldsymbol{A} \cdot \boldsymbol{m} = \boldsymbol{b} \tag{4.2-3}$$

where

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_{x1} \\ \boldsymbol{a}_{x2} \\ \boldsymbol{a}_{x3} \\ \boldsymbol{a}_{x4} \\ \boldsymbol{a}_{y1} \\ \boldsymbol{a}_{y2} \\ \boldsymbol{a}_{y3} \\ \boldsymbol{a}_{y4} \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1x_1' & -y_1x_1' \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2x_2' & -y_2x_2' \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x_3x_3' & -y_3x_3' \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4x_4' & -y_4x_4' \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1y_1' & -y_1y_1' \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -x_2y_2' & -y_2y_2' \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -x_3y_3' & -y_3y_3' \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -x_4y_4' & -y_4y_4' \end{bmatrix} \text{ and } b = \begin{bmatrix} b_{x1} \\ b_{x2} \\ b_{x3} \\ b_{x4} \\ b_{y1} \\ b_{y2} \\ b_{y3} \\ b_{y4} \end{bmatrix} = \begin{bmatrix} x_1' \\ x_2' \\ x_3' \\ x_4' \\ y_1' \\ y_2' \\ y_3' \\ y_4' \end{bmatrix}.$$

If the inverse of ***A*** exists, then the mapping parameter vector ***m*** can be derived as

$$\boldsymbol{m} = \boldsymbol{A}^{-1} \cdot \boldsymbol{b} \tag{4.2-4}$$

which contains the parameters, m_i, $i = 0,1,\ldots,7,$ in the mapping model (4.2-1).

A proper way to get the mapping parameter vector ***m*** by a matching group with four matching pairs is described in (4.2-1) to (4.2-4). However, the number of matching pairs is usually more than four, so there are more than one matching groups to form their mapping parameter vectors corresponding to different mapping behaviors. Therefore, it is required to have a base to judge the performance of each mapping parameter vector and choose the one which can successfully represent the transformation of the source image.

Let N_m be the total number of the matching pairs derived in Section 3.3 and $\boldsymbol{m}^k = [m_0^k, m_1^k, \ldots, m_7^k]$ be the *k*th mapping parameter vector obtained by using (4.2-1) to (4.2-4) related to the *k*th matching group. In order to judge the performance of the vector $\boldsymbol{m}^k$, the matching counter $c(\boldsymbol{m}^k)$ is set as an index for the number of appropriate matching data. The $c(\boldsymbol{m}^k)$ will be increased by one while the Euclidean distance between $(x_1'^j, y_1'^j)$ and (x_2^j, y_2^j) is within the threshold T_m, where $(x_1'^j, y_1'^j)$ is obtained according to (4.2-1) and mapping parameter vector $\boldsymbol{m}^k$ of (x_1^j, y_1^j), one of the coordinates of the source image in $\boldsymbol{MP}_{\boldsymbol{I}_1,\boldsymbol{I}_2}(j),\ \ j = 1,2,\ldots,N_m$. The desired mapping parameter vector $\boldsymbol{m}_d$ is chosen to be the one of $\boldsymbol{m}^k$, whose matching counter $c(\boldsymbol{m}^k)$ is maximum. However, it is inevitable that different matching groups may have the same maximum matching counter value. In order to uniquely determine $\boldsymbol{m}_d$, the error function $e(\boldsymbol{m}^k)$ is further employed as

$$\begin{aligned} e(\boldsymbol{m}^k) &= \sum_{j=1}^{N_m} \sqrt{\left(x_2^j - x_1'^j\right)^2 + \left(y_2^j - y_1'^j\right)^2} \\ &= \sum_{j=1}^{N_m} \sqrt{\left(x_2^j - \frac{m_0^k x_1^j + m_1^k y_1^j + m_2^j}{m_6^k x_1^j + m_7^k y_1^j + 1}\right)^2 + \left(y_2^j - \frac{m_3^k x_1^j + m_4^k y_1^j + m_5^j}{m_6^k x_1^j + m_7^k y_1^j + 1}\right)^2} \end{aligned} \tag{4.2-5}$$

to accumulate the Euclidean distance between $(x_1'^j, y_1'^j)$ and (x_2^j, y_2^j) of all the matching pairs in $\boldsymbol{MP}_{\boldsymbol{I}_1,\boldsymbol{I}_2}(j),\ \ j = 1,2,\ldots,N_m$. Consequently, the desired mapping parameter vector $\boldsymbol{m}_d$ is chosen to be the one of $\boldsymbol{m}^k$, whose error function $e(\boldsymbol{m}^k)$ is minimum. In other words, the mapping parameter vector with maximum value of matching counter and minimum error function will be regarded as the best mapping model to represent the transformation of the source image.

4.3 Random Sample Consensus

The mapping model in (4.2-1) using a matching group with four matching pairs can avoid the defect of modeling derived by all the matching pairs, which usually consists of false matching pairs and leads to an imprecise modeling. However, it is a time consuming process to include all the possible matching groups and determine the $\boldsymbol{m}_d$ based on finding the maximum value of matching counter $c(\boldsymbol{m}^k)$ and minimum value of error function $e(\boldsymbol{m}^k)$. To reduce the calculation time, some investigators proposed several techniques to determine the $\boldsymbol{m}_d$ base on finding appropriate value of matching counter and error function. Although their results are not the best, their procedures indeed save the calculation time and their mapping models have been demonstrated to be acceptable and useful in the image stitching process. One of these beneficial techniques is called random sample consensus, which will be adopted in this thesis.

Random sample consensus is proposed in 1981 by Fischler and Bolles [4]. Their algorithm uses fixed number of data as a trial for modeling and finds the essential number of trials based on the sense of probability to obtain an appropriate mapping parameter vector. The method called planar image mosaics using the fixed data modeling [13][14], has been introduced in the previous section. In this section, the content will focus on how to obtain the essential number of trials based on probability to save the calculation time.

Let P_a be denoted as the probability of deriving appropriate matching pairs, whose error is within the tolerance of mapping model in (4.2-1), and expressed as

$$P_a = \frac{N_a}{N_m} \quad , N_a \le N_m \tag{4.3-1}$$

where N_m is the total number of the matching pairs and N_a is the number of appropriate matching pairs. Therefore, in order to obtain an appropriate mapping parameter vector, the four matching pairs as a trial for modeling input must all be the correct ones and the probability of finding an appropriate model can be described as

$$P_h = (P_a)^{N_g} \tag{4.3-2}$$

where N_g is the number of the correct input data and is set to be 4 in this thesis based on the mapping model in Section 4.2. Clearly, the higher P_a results in the higher P_h. Further, suppose that the probability of finding at least one proper model after N_t trials is P_o and can be obtained as

$$P_o = 1 - (1 - P_h)^{N_t} \tag{4.3-3}$$

then transpose (4.3-3) and take logarithm to the both side to get the number of trials N_t which can be expressed as

$$N_t = \frac{log(1 - P_o)}{log(1 - P_h)} \tag{4.3-4}$$

thus the value of N_t is determined according to the value of P_o which can be adjusted by the user and is set to be 0.99 in this thesis to achieve high accuracy.

Although it is a good measure to reduce the number of trials based on probability, there is still some problem that N_a can't be known before all the trials are tested. Hence it is necessary to choose N_a to be reasonable constant value. Because the optimal mapping method needs four accurate matching pairs to obtain a proper mapping parameter vector, it is suitable for choosing $N_a = 4$ and the goal of saving calculation time can be successfully achieved.

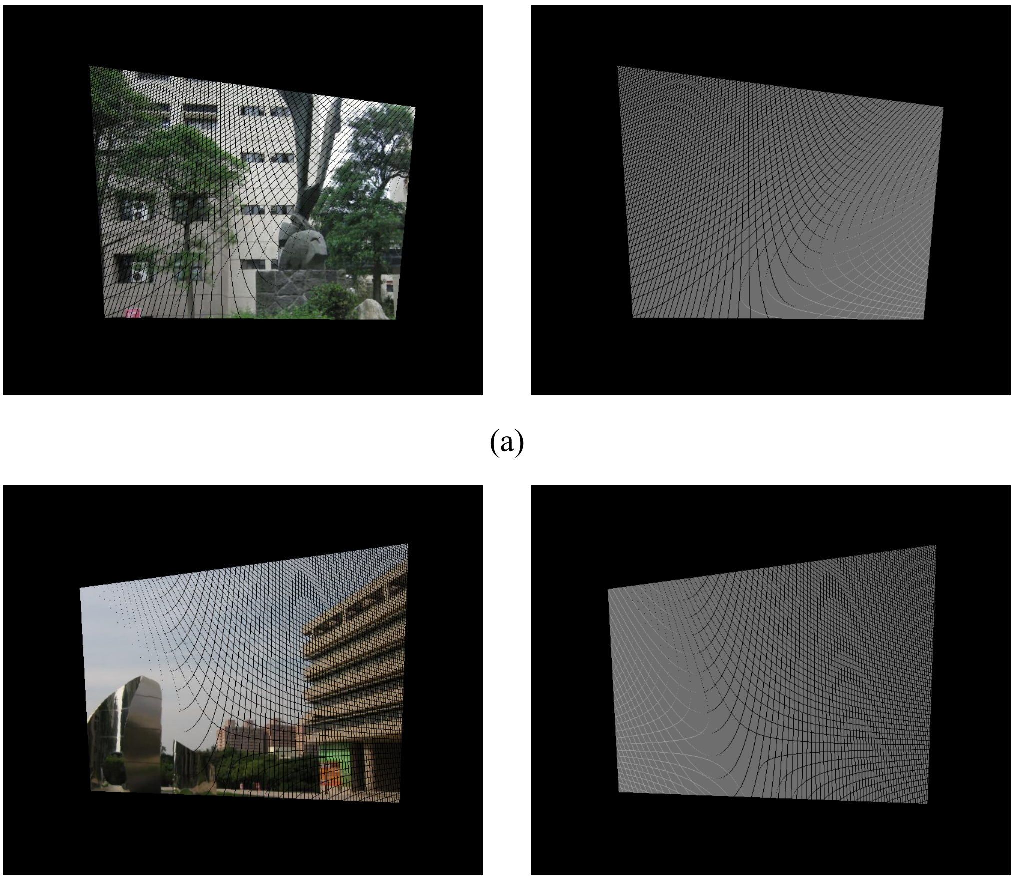
To show the advantage of the time-saving by random sample consensus method, the number of trials N_t in (4.3-4) and $N_{mg} = C_4^{N_m}$ are given in Table 4.1, where N_{mg} is related to the technique using all possible matching groups. Clearly, from Table 4.1 it can be seen that the more matching pairs, the superior using the random sample consensus technique. The experiments also confirm that this technique can indeed save a lot of calculation time and can choose an appropriate matching group to obtain a proper mapping parameter vector for source image.

Table 4.1 The number of trials N_{mg} and N_t

N_m / N_{mg} & N_t	20	30	40	50	60	70
$N_{mg} = C_4^{N_m}$	4845	27405	91390	230300	487635	916895
$N_t = \dfrac{\log(0.01)}{\log\left(1\text{-}\left(\dfrac{4}{N_m}\right)^4\right)}$	2876	14569	46049	112429	233134	431912
$\text{Time saving} = \left(1 - \dfrac{N_t}{N_{mg}}\right)$	40.64%	46.84%	49.61%	51.18%	52.19%	52.89%

4.4 Image Compensation

After getting the transformation of source image, it is necessary to put the source and the target image on big canvases according to their relative position. However, the way of mapping based on the Section 4.2 is a one-to-one mapping, it is inevitable that when the source image has to change its original shape to match up the target image, some pixels would be mapped to the same position and some positions within the reasonable image region would be unmapped. Consider the images shown in Figure 4.1, the left images are the mapping results after image mapping method without doing compensation. Moreover, the right images are the gray-level images with different gray-level values corresponding to different situations. The pixels in the right images which with gray color represent the region whose positions are mapped just once from the original source image. In other respects, the pixels with white color represent the region whose position are mapped more than once from the original source image and the black pixels represent the unmapped pixels. When the overlapped situation occurs, the former mapped pixel values would be replaced by the latter mapped pixel values in the same positions. Furthermore, it can be seen that it is required to fill some unmapped pixels with their reasonable colors because they are within the reasonable image region but without mapping from the original source image.



(a)

(b)

Fig 4.1 The mapped source images before compensative processes

For the overlapped problem, thanks to the good performance of mapping parameter vector chosen in the previous section, there is usually little color difference between each overlapping pixel in the same positions. Therefore, the easy and useful way to handle this problem is take average of these overlapping pixels. On the other hand, for the unmapped problem, it should first distinguish from the canvas pixels and the pixels needing to be compensated because both of them are unmapped ones. It is sensible that if the unmapped pixels within the reasonable image region, they can be remapped into this region by solving the simultaneous equation in (4.2-1) to get the reverse of the mapping model which can be obtained as

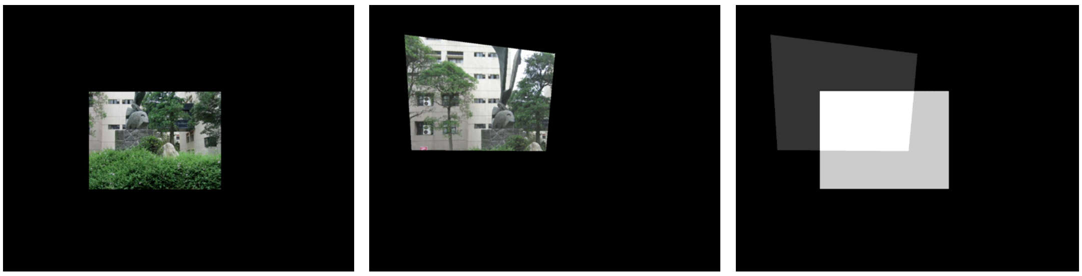
$$
\begin{aligned}
x_u &= -\frac{(m_6 m_8 - m_5)x'_u + (m_3 m_8 + m_2)y'_u + (m_3 m_5 - m_2 m_6)}{(m_4 m_8 - m_5 m_7)x'_u + (m_2 m_7 - m_1 m_8)y'_u + (m_1 m_5 - m_2 m_4)} \\
y_u &= -\frac{(-m_6 m_7 - m_4)x'_u + (m_3 m_7 - m_1)y'_u + (m_1 m_6 - m_3 m_4)}{(m_4 m_8 - m_5 m_7)x'_u + (m_2 m_7 - m_1 m_8)y'_u + (m_1 m_5 - m_2 m_4)}
\end{aligned}
\qquad (4.4\text{-}1)
$$

where (x'_u, y'_u) represents the coordinate of the unmapped pixel, and the mapping parameter vector $\boldsymbol{m} = [m_0, m_1, \ldots, m_7]^T$ can be derived appropriately in Section 4.2 and Section 4.3. Therefore, if the reverse-mapped coordinate (x_u, y_u) locates at the original region of the source image, it can be regarded as the pixel needing to be compensated and the pixel value of the coordinate (x'_u, y'_u) is set to be the same as that of (x_u, y_u). Figure 4.2 shows the images which are compensated by the method proposed above and the problems can be solved successfully without artifacts.

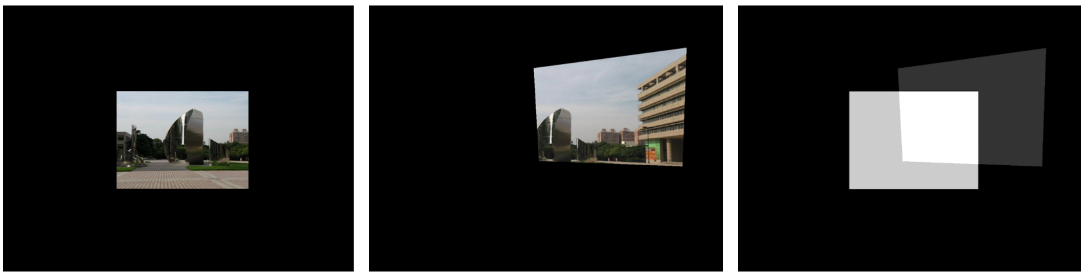
Fig 4.2 The mapped source images after compensative processes

In order to implement the image mapping technique proposed above by programs, it should first create three big canvases, $\boldsymbol{V}_{cb}$, $\boldsymbol{V}_{cs}$ and $\boldsymbol{V}_{ct}$, where $\boldsymbol{V}_{cb}$ is the whole rectangular canvas containing the information of the combined region and $\boldsymbol{V}_{cs}$ and $\boldsymbol{V}_{ct}$ are the canvases for the mapped source image and the target image respectively. As shown in Figure 4.3, put the target image on the center of $\boldsymbol{V}_{ct}$ with the coordinate offset (x_o, y_o) and set the pixel values of these coordinates to be 200 in $\boldsymbol{V}_{cb}$. Afterward, map the source image based on (4.2-1) and $\boldsymbol{m}_d$ with the same coordinate offset (x_o, y_o) and then compensate the overlapped and unmapped pixels proposed in the previous paragraph to form $\boldsymbol{V}_{cs}$. Moreover, add the pixel value 55 to these coordinates in V_{cb}. From the procedures stated above, the canvases V_{cs} and V_{ct} store the offset mapped source image and the offset target image, respectively, and canvas V_{cb} stores the information of the combined region which can easily recognizes the source region, target region, overlap region, and the unmapped pixels with different gray-level pixel values as 200, 55, 255 and 0, respectively.

In this chapter, the source image can be mapped to match up the target image by a useful mapping model, an appropriate mapping parameter vector and a time-saving method. Moreover, the defects of the mapping model are compensated and the overlap regions of two images are found. Therefore, in next chapters, the previous image stitching technique are introduced and the novel image stitching technique is proposed in this thesis to combine the mapped source image and the target image together successfully by processing the overlap regions with good performance.



(a)



(b)

Fig 4.3 Canvases V_{ct} (left), V_{cs} (middle) and V_{cb} (right)