Université du Québec à Chicoutimi

Mémoire présenté à
L'Université du Québec à Chicoutimi
comme exigence partielle
de la maîtrise en informatique

offerte à

l'Université du Québec à Chicoutimi

par

Wang Xin

The Research on Algorithm of Image Mosaic

Septembre 2008

Abstract

Image based rendering (IBR) has been the most important and rapid developed techniques in the computer graphics and virtual reality fields these years. Image mosaic which is one of the hot topics of IBR is also becoming research interests of many researchers in the image processing and computer vision fields. Its application covers the areas of virtual scene construction, remote sensing, medical image and military affairs etc. However, some difficult issues need to be studied further, including new optimization methods for image registration, new accelerating methods for image stitching etc, which are the main topics of this thesis.

First, as the precision and automatic degree of image mosaic suffers from the algorithm of image registration, a new image stitching optimization method based on maximum mutual information is presented in this thesis. The main idea of the new method is to combine PSO algorithm with wavelet multiresolution strategy and parameters of PSO are adapted along with the resolution of the images.The experiments show that this method can avoid registration process to get stuck into local extremes in image interpolation operations and finds the optimal exchange through limited iterations computation, and obtain subpixel registration accuracy in process of image stitching.

Secondly, to solve the problem of image blur stitching when the geometric deformation and the changes of the scale factor among images are serious, a new method based robust features matching is proposed in this thesis. Firest, it searches overlap area between two images by using phase correlation, and then detects Harris corner points in the overlap areas which are reduced to different scale images by adopting multi resolution pyramid construction. This method can solve the Harris arithmetic operator robust characteristics inspection algorithm for the scale effects. In order to improve the running performance of image feature matching, a dimension reduction method of high dimension feature vector is proposed based on PCA. And last the globe parameters are optimized by Lmeds to realize images mosaic. The experiments prove that the methods proposed by this thesis can reduce the computation cost with guarantee of image mosaic quality.

KEYWORDS： Image mosaic, Registration, Mutual information, PSO, Dimension reduction

Acknowledgement

I wish to express my sincere gratitude to my academic advisor, Prof. Cao Zuo Liang who helped and guided me towards my academic as well as my professional success. This thesis would also not be successful without my advisor. He provided me with a substantial amount of help and suggestions throughout my work. And same time I got much help from Paul Girard Ph.D.who is my co-supervisor from UQAC, I appreciate his assistance. I would also like to thank my friends who help me with positive affirmation. Finally, I would like to thank my parents for their unending support and eagerness to love.

Table of Contents

List of Figures

List of Tables

Abbreviations

IBR	Image based rendering
PCA	Principle Components Analysis
PSO	Particle Swarm Optimization
PV	Partial Volume
LMedS	Least Median of squares
VE	Virtual Environment
VR	Virtual Reality

Chapter 1 Introduction

This chapter will mainly introduce background and significance of the research, analyze the current researching situation of the image mosaic technique which is virtual scene oriented, expound main problems this thesis will solve and some major research results, and explain the overall structure of the thesis.

1.1 Research Background, Aim and Significance

The mosaic of images is an important topic in Image Base Rending (IBR) area[1] as well as a major technique in rebuilding the Virtual Environment (VE)[2]. The problems it is to solve are how to mosaic pictures in small visual area in to a bigger one, to satisfy people's need to observe and browse scene in broad area. The mosaic of scene image is started in accompany with the rapid development of Virtual Reality (VR) technique and the need to real-time draw in VE.

VE is an advanced human-computer interface system to do real-time stimulation and interaction through human beings' several perceptions (sight, hearing, touch, olfaction, and taste). It has so-called 3I characteristics, they are, Immersion, Interaction and Imagination[3]. In this virtual world showed by VE, various complicated entities in nature are changed into a wonderful world in which human beings can control and change those entities. And in this world people can feel the same way as the real world. With the development of VR technique, there are higher demands on the truthfulness of virtual scene. And the sense of reality is an important research topic in computer graphics area. For a long time scholars from different countries have devoted themselves to express various fine structure of objective world with geometrical modal, and have achieved fruitful results[4]. However, with rapidly changing and complicated modeling requirements in objective world, the contradiction between the sense of reality and computing scale is always a tough problem[5][6]. Therefore, scholars analyze the sense of reality' showing problems with the combination of image processing technique and computer graphics. They obtain many effective drawing effects through analyzing images and techniques such as processing and mixing. In this way, they can express details geometrical modals can't and make up for the deficiency of geometrical drawing. Since the later half of the last century, image-oriented drawing technique has always been a hot topic in the world, and it's important part—image mosaic technique has arouse broad attention around the world, and had may valuable results [7][8].

The basic difference between IBR' drawing technique and rebuilding scene in traditional computer graphics lies in: it does not depend on geometrical modal. Instead, it codes the environment with a group of image sequence obtained before; and through properly changing, it generates new view graphs of different view points; and finally, it realizes complete environment walkthrough. Because IBR technique fully makes use of the truthfulness of image and its advantage of quick showing, it has become an important way to strengthen the truthfulness of VE and improve the efficiency of scene drawing.

Image-oriented drawing methods mainly involve three basic problems [9]:

(1) Given scene images of different view points in a real scene, how to generate new scene images of view points. Now Image Morphing technique [10][11] is based on pixel position and color interpolation among original images. However, there is no standard to decide whether the obtained interpolation image is the image of a certain new view point in the same scene or object. The research results show that current Image Morphing technique can not guarantee this. Therefore, how to acquire needed view graphic from the image sequence which is never nominated.

(2) Given a group of segmented image in certain scene, how to generate new view graphic which contains this group of segmented image, that is, the problem of image mosaic. Its core is to find out a change which trains the overlapping parts of images and "sew" them into a new view graphic with a broader visual area. This involves how to realize the matching of images. Image registration is still a classical problem in image processing.

(3) How to automatically mosaic several real images into a panorama. Basically a panorama has cylinder, sphere, and cube, etc. and cylinder is the easiest and the most popular one now [12][13]. However, the most famous commercial panorama generating system such as QuickTime need some conditions to collect images [14][15]. So, under the condition of guaranteeing the mosaic quality, it is an important research issue to improve the automation of panorama's generation and reduce the mosaic mistakes.

From above we can see that in image-oriented drawing technique, the image mosaic is the key connection in the process of building virtual scene. And the research on this technique will definitely influence a lot on IBR method as well as the whole virtual reality technique.

Each image reflects partial information of certain objective or scene in space and time. Image mosaic technique will provide an effective and complete scene expression method to make people know things and understand real world more objectively and vividly. Therefore, image mosaic technique is important research content of subjects such as photogrammetry, computer graphics, and image processing and computer vision. And it is widely used.

Application areas are as follows [16][17]

(1) The mosaic of large airscape and satellite remote sensing image;

(2) Meteorological and environmental monitoring;

(3) Sea-bottom survey and geological survey;

(4) Combination pf medicine and scientific micro-fragment image;

(5) The 3D rebuilding of objects;

(6) The building of virtual scene and virtual walkthrough;

(7) Video compression, video search browse, and video edit, etc;

(8) The digitized saving of file;

(9) Military reconnaissance and taking evidence.

Image mosaic which involves several subjects and areas such as computer graphics, image processing, computer vision, modal recognition, and artificial intelligence technology, etc, is a pioneering and interdisciplinary research. Therefore, research on rapid, highly accurate, and high automation image mosaic technique has theoretical significance and utility value to expand application area of image mosaic and promote deep development of related subjects.

1.2 The Research Situation and Major Problems of Image Mosaic Technique

At present, image mosaic, as a hot topic in computer graphic, image and computer vision areas, has attracted broad and deep research by many scholars home and abroad. However, because it involves many interdisciplines' theory and technique [18][19], image mosaic is still in the research and developing stage, and has not formed complete theory and technique, which restricts the rapid promotion and application of image mosaic.

In recent ten years, in order to increase the quality and robustness of image mosaic, scholars home and abroad have made various foundational research and development of actual system, and have achieved fruitful results.

First, as the key part of image-oriented drawing technique, image mosaic is developing with deeper and deeper application of computer graphics. In graphics area, people are always devoted to generation virtual scene which is quite similar with the real scene. The traditional way is to imitate the real scene by building three-dimensional geometrical modal. Whether virtual scene can be built by images reflecting real scene ? To answer this question, people start application research on image-oriented drawing technique. In the year 1980, Lippman and Miller first raised the concept of IBR, it is a kind of technique which takes several realistic images which are photographed in real situation with camera as example, and makes use of image processing technique and visual calculation method, to build three-dimensional virtual scene [20]. Now there is no detailed technical definition for IBR technique. But we could consider it as the main information when drawing images. We should try to remove the geometric model part to obtain scene which is similar to true photos. [21]. The best advantage of IBR lies in the generated environment is the real situation images reflect. Therefore, it is especially suitable for simulation study which is based on natural scene.

At present, IBR technique is mainly divided into four types: Plenoptic Function, View Interpolation, Lumigraph, and Full-View-Mosaic or Panorama [22] [23].

With all the scene data of current view points, the following question is how to show segmented scene which is related to view points quickly. Now the method of re-projection is mainly used. The projection process is three-dimensional world is formed an image on two-dimensional surface through an ideal perspective camera. The common re-projection mapping surfacing include: plane, cylinder, and sphere and cube surface[24] [25].

In conclusion, panorama generation technique, one of the most representative techniques of IBR research area, its key part is to form virtual scene which is highly real with collected real scene images and mosaic processing.

However, the panorama technique still has some restrictions. First, it is only suitable when the camera is in a fixed position. And it cannot work when camera is moving. Next, for the situation of hand-hold camera, because complete plane is hard to achieve in shooting process, or the pictures are taken at different times, and there are exposure differences, therefore the mosaic of two adjacent images' projection needs to accurate position the overlapping place, and then register the images, to get rid of the seam after mosaic. Li [26] adopts multi resolution response modal to solve the image indistinct problem when view point is moved forward. However, they need qualified original data. And at the same time they should consider complicated edge matching problem. In this way, the complication of arithmetic will increase dramatically. Therefore, how to combine TIP with image mosaic technique, and to solve jumping and indistinct phenomenon of scene walkthrough, these are problems to be solved urgently.

1.3 The Main Research Work and Creativity

Based on the analysis on section 1.2 and problems existed in current image mosaic technique, this thesis will use synentropy technique, multi resolution response technique, and principal component analysis technique, and combined intelligence compute technique, on the premise that the quality of image mosaic is guaranteed, to analyze related robust mosaic method from the perspective of optimizing strategy and reducing complication of calculating. This thesis will study the solution to indistinct problems of scene image generated by virtual scene and walkthrough in depth, and analyze and validate through experiment.

The main research contents of this thesis are:

1、 Image registration is a basic part in mosaic process. This thesis use maximum mutual information with good robustness and high automation in image registration process. To solve the problem that calculus of interpolation leads to local minimum in optimization process, based on multi resolution response analysis and particle group optimization algorithm, this thesis analyzes optimization algorithm of multiscale and self-adaptation of image mosaic technique. The optimization algorithm raised in this thesis make its parameters can adapt themselves according to mutual information and multi resolution response progression. Experiments prove that this method can avoid the influence of local extremum effectively, and find out optimized change through finite iteration. Compared to current algorithm, this method increases the computed speed of image registration, and offers subpixel registration promise for the subsequent image synthesis.

2、 For problems such as image have obvious differences in rotation and scale, big spending in image registration and mosaic calculation according to gamma and mosaic effect is not good, etc, this thesis analyzes character-oriented image mosaic technique. First fix on the overlapping area of images to reduce the range of character measurement; next make multi resolution analysis for this area; and with Harris arithmetic operators, extract characteristic corner in different scale image and build descriptor with fixity. And then make dimensionality reduction to eigenvector with principal component analysis, to reduce calculation spending. And test reliability compliance through experiments.

1.4 Structure of Thesis

Chapter 2 introduces background information related to this thesis and main technique foundation. It includes basic process of image mosaic, camera imaging modal and tion modal, image geometry changing foundation. And at last, it analyzes the key technique of image mosaic.

Chapter 3 discusses maximum mutual information's image mosaic technique. It discusses characteristics and problems of image registration which is based on maximum mutual information, raises optimization strategy and realization arithmetic through the combination of particle group arithmetic and multi resolution response. And analyze and validate the arithmetic through experiments.

Chapter 4 analyzes characteristics-oriented image mosaic technique. This chapter emphasizes the adjusted strategy of testing and registrating corner characteristics which is based on Harris operator, and raises the method to reduce dimension of eigenvector to reduce calculation spending. And last analyzes the effectiveness of the method through experiments.

Chapter 5 concludes the research work of the whole thesis, and analyze and outlooks future research.

Chapter 2 Research Foundation of the Image Mosaic

Image mosaic is an interdisciplinary problem. It involves consolidated application of various areas such as computer graphics, modal reorganization, and image processing. Several mathematic tools are also brought in this technique, and good effect is achieved. This chapter will enclose the main process, outline analysis, and some necessary background information of image mosaic.

2.1 The Definition of Image Mosaic

As is shown in figure 2-1, image mosaic is a kind of technique that means mosaic two or more than two images which have partial scene overlapping, and generate a panorama of wide vision or 360 ° visual angles.

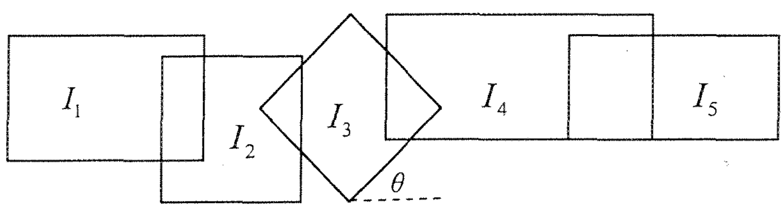


Figure 2-1 A Sketch Map of Image Mosaic

Suppose $I_1(x)$ and $I_2(x)$ are two images waiting to be mosaiced, and its dimensionality is *d*. suppose the coordinate of the image is $x=(x_1,x_2,....x_d)^T$, f represents space changing relationship of two images' corresponding coordinate points, that is, $x^{'}=f(x)$, in which *x* and $x^{'}$ are two corresponding points of image $I_1(x)$ and $I_2(x)$ respectively. In this way, we call $I_2(x)$ which stays stable as norm image, and changing image $I_1(x)$ as moving image.

2.2 The Process of Image Mosaic

Image mosaic process generally contains image preprocessing, image registration and image mosaic:

1、preprocessing

The aim of preprocessing is to simplify the difficulties of image registration, that is, to make some replication change and transformation of coordinates on original image to locate adjoining image roughly, and find out the overlapping area, reduce registration range, and increase the speed of image mosaic processing.

2、Image registration

The quality of image mosaic mainly depends on the registration of image. The core problem of image registration is to find out a space variation and registrate the coordinate points of overlapping parts between images. Registering arithmetic should make sure the accuracy, and should control the scale of arithmetic. This is the key step of image mosaic technique as well as the emphases of this thesis.

3、Image mosaic

When the image is registered, and space variation between images is calculated, we should composite several original images into one big image or a panorama. The main process of image mosaic includes: mosaic adjoining images' registering areas, get rid of accumulated mistakes made in the process of overall mosaic and distortion phenomenon in image overlapping area, and draw and output panorama mosaic image.

2.3 Imaging and Moving Modal of Camera

Camera is the main device to obtain images. And analyzing camera's imaging modal is a basic precondition of image mosaic. Only after us mastering camera's geometry imaging modal, we could make sure a point's projection changing relationship from a three-dimensional space to a two-dimensional surface. The imaging modal of camera is usually expressed by perspective projection change. This section will offer basic knowledge of imaging modal and imaging change from the concept of projective geometry.

2.3.1 Change Basis of Space Geometry

Each point in space can be expressed by a vector, and vector space made up of these vectors can describe point space in real world [27]. In the construction of virtual scene, description of various structures in three-dimensional world with Euclid space is always restricted. For example, two rail ways cross in infinite distance, and Euclid space can not describe point in infinite distance. And we need the descriptions in projective space.

Definition 2-1 Projective space

Suppose V is n+1 dimensional vector space , V's one-dimensional subspace's set is called n-dimensional projective space, which can be expressed as p^n. And one point x in projective space p^n's corresponds to a n+1 dimensional-vector $x = [x_{0,} x_{1,} .....x_n] \neq 0$ in vector space V= R^{n+1}.

Definition 2-2 Homogeneous coordinates

When and only when there is a real number λ which makes $[x_{0,} x_{1,} .....x_n] = \lambda [x_0^{'}, x_1, ,,, x_n^{'}]$, we say these two points are the same, and it can be expressed as $[x_{0,} x_{1,} .....x_n] \cong [x_0^{'}, x_1, ,,, x_n^{'}]$. And this n+1 dimensional vector is called the point x's homogeneous coordinates, which is expressed as $x = [x_{0,} x_{1,} .....x_n]$. In this way, points in two-dimensional projective space p^2 and three-dimensional projective space p^3 are expressed as $[x, y, z]$ and $[X, Y, Z, W]$ respectively.

Usually, when points in p^2 and p^3 are expressed with homogeneous coordinates, we suppose $z = 1$, $W = 1$, that is, $P(x, y)$'s homogeneous coordinates is expressed as $P(x, y, 1)$, and $P(x, y, z)$'s homogeneous coordinate is expressed as $P(x, y, z, 1)$.

Definition 2-3 Affine space

Suppose $0 \le i \le 1$, projective space's subset $A_i^n = \{x = [x_0, x_1, ... x_n] \in p^n\}$ is called a affine part of p^n, and it is defined as the affine space. Under homogeneous coordinate, every element x in A_i^n can be expressed as: $x = [x_{0,} x_{1,} ... x_{i-1,} 1, x_{i+1,} x_{n+1}]$. p^n 's hyperplane Pn $\{x^n \in p^n \mid x_i = 0\}$ is called plane at infinity.

Definition 2-4 Euclidean space

For any vector v in vector space V upon which affine space A^n depends, we define its distance as $|v| = \sqrt{(v.v)}$, in which $(v.v) = v^T v$ is vector inner-product, which defines the affine space of this distance, and it is called Euclidean space. The nonsingularity change from Euclidean space E^n to itself is called Euclidean change.

If Euclidean change is T^k, then T^k has the following form:

$$T^k = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \tag{2-1}$$

R is orthogonal matrix $R^T R = I$, and I is unit matrix. Usually we use R to express a space's rotation transformation, and t is a translation vector.

The imaging process of is a process of projection change; the coordinate systems it involves are the world coordinate system, camera coordinate system, and image plane coordinate system.

(1) The world coordinate system is three-dimensional coordinate system, which expresses the absolute coordinate of scene point in objective world. This system can define a point's exact position in real world through coordinate (x, y, z).

(2)Camera coordinate system is a three-dimensional coordinate system which takes camera as original point. It usually takes plane shaft of camera as z shaft.

(3) The image plane coordinate system is a two-dimensional image plane coordinate system formed in the camera equipment, which expresses scene point's projection on the image plane. Usually, the image plane parallel with xy plane of the camera coordinate system, the original point lies in the plane shaft of camera, which is superposition with the original point in camera coordinate. This coordinate system is similar to the plane coordinate system in which the film is .

The relationship between camera coordinate system and world coordinate system can be described with rotation matrix R and translation vector t. If the homogeneous coordinates in the world coordinate system and the camera coordinate system of a certain point P in the space p^3 are $[X\ Y\ Z\ 1]^T$ and $[x\ y\ z\ 1]^T$ respectively, then there are the following relationships:

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \qquad (2\text{-}2)$$

2.3.2 Imaging Modal

Imaging means mapping scenes in three-dimensional space into images in two-dimensional coordinates the common mapping transformation is perspective projection transformation. Perspective projection is also called pin-hole imaging modal, its characteristic is that all the lights come from scene go through a projective center (photocenter), and the photocenter is corresponding to the center of the lens. The lines that go through projective center and are perpendicular to image plane are called axis of projection or plane shaft, such as expressed in figure 2-2. *xyz* is a rectangular coordinate system in camera, and it follows the right-handed rule. Its original point lies in the projective center *O*. the axis z is superposition with the axis of projection and points to the scene. Axis x and axis y parallel with coordinate axis x' and y' in plane, and the distance between plane xy and image plane is

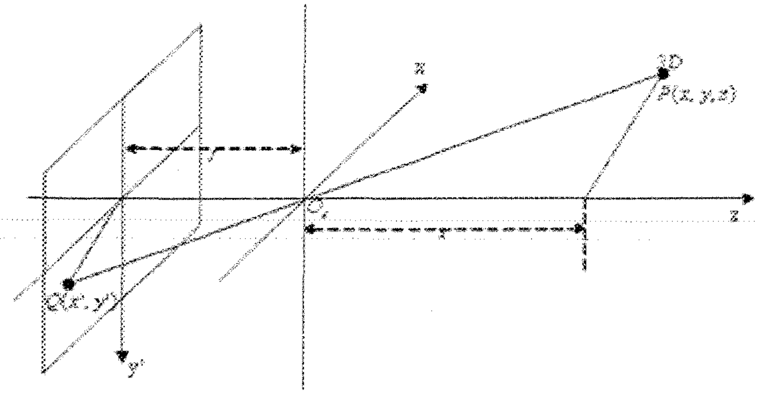


Figure 2-2 Perspective Projection Transformation Diagram

h, which is generally called the focus of the camera. In actual camera, image plane locates at f distance behind projective center, and the projective image is a reversed image.

Suppose (x, y, z) is the scene point P's coordinate of camera coordinate system, (x', y') is the point Q's coordinate in plane. The following is the expression of above projection relation in homogenous coordinates:

$$z\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & f & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \qquad (2\text{-}3)$$

Substitute formula (2-2) into (2-3), we can get the relation between 3D scene point P's coordinate in the world coordinate system and its projective point Q in computer image coordinate system:

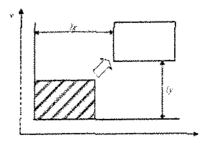
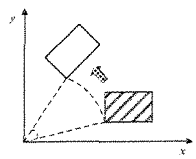
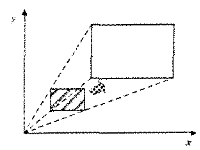
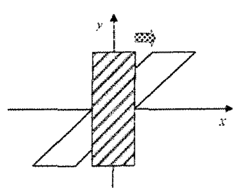
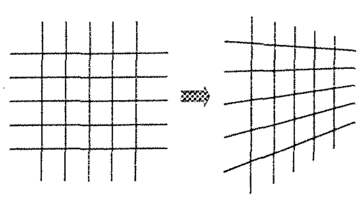
$$z\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & f & 0 \end{bmatrix}\begin{bmatrix} R & t \\ 0^{\mathrm{T}} & 1 \end{bmatrix}\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = M\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \qquad (2\text{-}4)$$

2.3.3 Image Coordinate Transformation Modal

The transformation of image is a process to map a point $P(x, y)$ in original image to object image $P'(x', y')$. The transformation of image can be defined by transformation matrix M:

$$P' = MP \Leftrightarrow \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = M \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \qquad (2\text{-}5)$$

Table 2-1 The Transformation of Image

name		Transformation form	Transformation matrix M		Degree of freedom	
Affine transformation	Parallel move		$\begin{bmatrix} a & c & e \\ b & d & f \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}$	6	2
	rotation			$\begin{bmatrix} \cos\Theta & -\sin\Theta & 0 \\ \sin\Theta & \cos\Theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$		2
	scale			$\begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$		2
	shear			$\begin{bmatrix} 1 & sh_x & 0 \\ sh_y & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$		2
Projection transformation			$\begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix}$		8	

The general image coordinate transformation ways are mainly affine transformation and projection transformation. Affine transformation includes parallel move, rotation, scaling, and shear, etc. the definition of each transformational type is shown in table 2-1.

The transformation of many different images can be superpositioned through transforming the dot product of the matrix:

$$P' = T(t_x, t_y) \cdot R(\Theta) \cdot S(s_x, s_y)$$

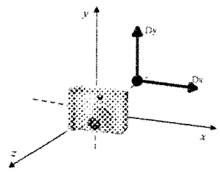
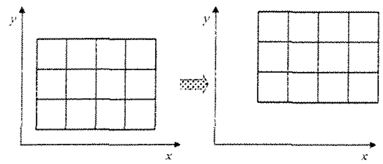
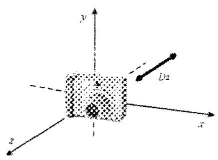
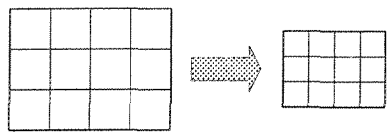
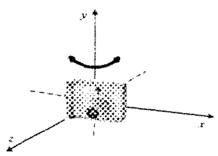
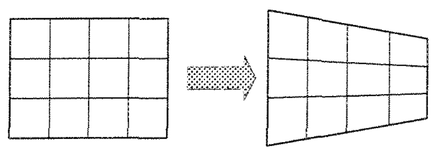
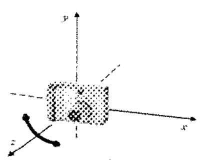
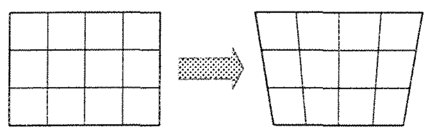
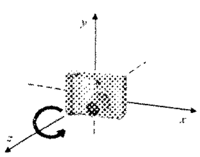
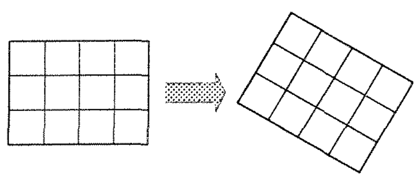
$$\sim \begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = \left(\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\Theta & -\sin\Theta & 0 \\ \sin\Theta & \cos\Theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x \\ y \\ w \end{bmatrix} \qquad (2\text{-}6)$$

Mosaic the adjoining two images, its transformational relation can be expressed by the projection transformation with 8 degree of freedom.

2.3.4 The Tion Modal of Camera

When collecting images with camera, different moving ways may generate different effects to the scene. The relationship between moving way of camera and imaging result [28] is shown in table 2-2

Table 2-2 The Relationship between Cameras's Moving Way and The Imaging Result

name	Camera movement	The result of image transformation	Image transformation
Parallel movement			Translation transformation
Changing the focus			Shrinking transformation
Level rotation			affine transformation
Vertical rotation			Projection transformation
circuition			Rotation transformation

Among camera's basic moving ways, imaging plane does not parallel with the scene plane in parallel transformation and vertical transformation, and this will generate the phenomenon of keystone distortion and linear frequency modulation [29]. The so-called keystone distortion is the phenomenon that the original parallel lines intersect with the photoed image. And the linear frequency modulation is the phenomenon that the scene images, which is taken in the process of camera movement, its space frequency will enlarge and reduce with the change of the space position. Parallel movement, camera scaling, and rotation transformation will not have the above two distortion phenomenon because its imaging plane parallel with scene plane. These two distortions brings the effect of perspective projection, therefore the coordinate transformation of plane vertical lens rotation is projection coordinate transformation.

When collecting images in fixed place, the movement of camera only includes parallel rotation and vertical rotation. It is necessary to find out the kinematical parameter of the camera when you project panorama on cylinder or sphere. Only when we know the parallel and vertical rotation angle between images, we can map images into the system of polar coordinates of cylinder or sphere.

The rotation angle of the camera is the lens's rotation angles, and the image point collected by the axis of lens is also the center pixel point of the image. Therefore, as long as we obtain the movement angle of center pixel, we could know the rotation angle of the camera. As shown in figure 2-3, the image P_2 is mapped onto the surface of image P_1, and the midpoint of P_2 is in the O_2 coordinate of the image P_1 after mapping, and the midpoint of the image P_1 is O_1. Therefore, the moving distance of P_2's axes is $D = |O_2 - O_1|$, and then we can obtain the rotation angle of P_2's axes

$$\varphi = \arctan\frac{D}{R} = \arctan\frac{|O_2 - O_1|}{R} \tag{2-7}$$

And R is the pixel focus.

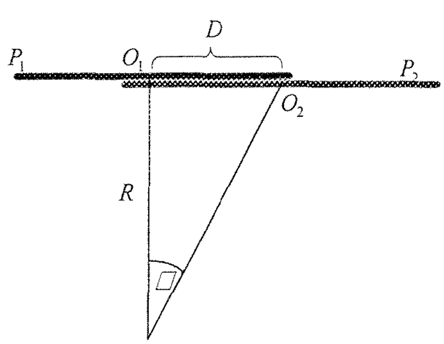


Figure 2-3 The Calculation of Rotation Angle of Camera

2.4 Key Technique of Image Mosaic

2.4.1 Image Registration

When we need to mosaic several images, their image registration can be divided into two steps: local registration and overall registration. Local registration is to solve the transformational relation between adjoining images, that is, to solve the transformational matrix of two-dimensional movement relation, and then recollect the sample of images according to transformation matrix and do interpolation calculation to realize mosaic of adjoining images. Overall registration is to get rid of accumulated error when mosaicing many images to realize the exact mosaic of panorama.

Because image registration technique is widely used in medical image analysis, virtual environment construction, remote sensing image registration, and environment monitoring, scholars home and abroad have done a lot of researching work in this area. They have raised different image registration methods according to different application areas. Although arithmetic are different from each other, they can be regarded as different combinations of feature space, measure of similarity, searching space, and searching algorithm[30].

1、 Feature space

The chose of feature space determines which feature of the image will take part in the registration, and which feature can be ignored. Choosing reasonable space can increase the capability of registration, and reduce the calculation of it. It can be generally divided into registration based on gamma and registration based on features according to different feature space;

(1) Registration based on gamma

This method can also be called registration based on area, correlation method, and template matching. It uses gamma directly to match without any structural analysis.

The simplest method in registration is that first user appoints some original matching points in turn, and then minimizes the sum of squares of energy difference among these matching points according to the moving relations.

$$E_{ssd}(u)=\sum_i [I_1(X_i+u)-I_0(x_i)]^2=\sum_i e_i^2 \qquad (2\text{-}8)$$

In this formula, u means the offset between two images, $e_i=I_1(X_i+u)-I_0(x_i)$ expresses the residual error between matching points of these two images, that is error matching point. The typical registration which is based on gamma also includes cross-correlation [31][32] and phase correlation which is based on Fourier transformation theory [33]. Their mutual feature is that they can match images with translation registration exactly, or images with slight rotation and scaling. Its shortcomings are large calculation and it needs long time.

(2) Registration based on features

In image mosaic, image registration algorithm based on features is to solve transformation modal's optimized parameter space through choosing features between two images and the principle of coherence. Usually we could adopt edge feature, area feature, and corner feature to complete the matching process of mosaic images [34]. The key of image registration which is based on features lies in the extraction of features of images and the corresponding relations between features of two images. Its shortcomings are that large amount of image information may be lost during feature extracting process, and it suit badly to general scenes.

Registration based on features will mainly go through search, match, and transform relation three process. In features registration algorithm, features test based on Harris corner is the most stable one [35]. On the basis of Harris detector, we develop features detector such as Harris-Affine [36]. Harris-Affine can detect automatically a series of image features in affine transformation, and it has the feature of affine invariance, and has strong robustness. This thesis will analyze the features detecting and matching method which are based on Harris corner.

2、Measure of similarity

Measure of similarity is the yardstick to know whether two images have achieved registration. For features based registration algorithm, we usually adopt various distance functions as measure of similarity of features, such as Euclidean distance, and Hausdorff, etc. as for the region correlation algorithm, we usually adopt correlation as measure of similarity, such as cross correlation, related coefficient, and phase coherence, etc. recently, mutual information method raised in literature [37] which is used in image registration has aroused broad attention. Mutual information, derives from informationism, is a statistic correlated measure in two information sets. Mutual information $MI(X,Y)$ of two extraneous variables X and Y:

$$MI(X,Y) = H(X) + H(Y) - H(X,Y) \tag{2-9}$$

In this formula, $H(X)$, $H(Y)$, $H(X,Y)$ represent extraneous variables X and Y's entropy and combination entropy respectively. The maximum mutual information registration need not make any assumption on the relations among image pixel gamma. And it needs no pretreatment of images. This thesis will take mutual information as measure of similarity, and analyze corresponding algorithm of image mosaic.

3、Searching space

The image registration is an optimum estimates of a parameter, and the space made up of these parameters is called searching space. The transformation of images can be divided into three categories: overall, local and play ground. Overall transformation is usually based on matrix algebra theory. It describes the transformation of the whole image with a parameter matrix. A typical overall geometry transformation includes the following one or several: parallel move, rotation, various scaling of isotropy and anisotropy, transformation of quadratic polynomial or cubic polynomial; local transformation is sometimes called stretching mapping, which allows parameters have position dependency, that is to say, different positions have different transformational parameters modal. Transformation parameters are always defined at certain key points, and do interpolation between areas; playground registration usually has two kinds: direct registration and registration based on features.

4、Search algorithm

Search algorithm is a method to find out an optimum transformation in search space, and maximize the measure of similarity. Search algorithm is very significant to reduce calculation. The more complicated search algorithm is, the more important it is to select sound algorithm, and higher the demand is. The general search algorithm is usually an optimum process, and it can be divided into the following ones: optimization in which parameters can be calculated directly, and optimization in which parameter needs to be optimized. This thesis will center on optimizing search algorithm to analyze.

In image mosaic which is based on features. , we could do feature matching directly after feature detection, for example, compares the similarity of surrounded pixels directly [38]. But this method is not surely stable, especially in the situation where exists a lot of redundant features. The better way is to give each feature a certain descriptor word, which makes feature matching an index process to find similar feature descriptor words. There are various ways to express descriptor word, foe example, Mikoaljcyzk raised extraction feature with directional column diagram statistics in feature neighborhood as description of feature. And Lowe [39] got a high dimensional SFIT descriptor word from a complicated processing to describe the unchanging feature. This thesis will give a solution to how to reduce calculation spending generated by high dimensional feature descriptor.

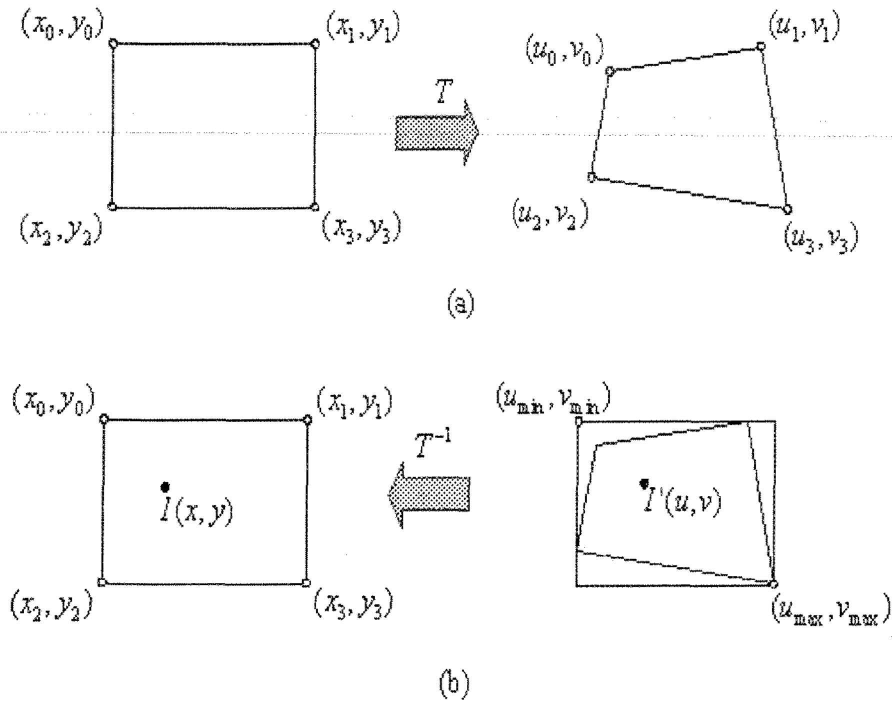
After finding out feature matching pair, we could solve tion transformation relation between images. This tion relation is always corresponding matrix, because it is hard to calculate complicated tion relation with only two images. Therefore, tion relation can be realized through iteration, and in each iteration, we should choose some feature matching pairs randomly to solve an initial tion relation. Such as LMedS(Least Median of squares) [40], which takes the corresponding matrix in least iteration method as the final corresponding matrix.

The first step of overall registration is to remove the accumulated error problem when several images superposition together. Usually we adopt binding regulation method [41] to renew the movement parameter, which will minimize the energy difference between new image and original one. Overall registration might be further optimization of transformation matrix, or it might be the scaling solution of the camera tion relation. In corresponding matrix optimization, first we should appoint a reference image, and then transform the transformation relation obtained in local registration into transformation relation between images and reference image. And at last do optimization-compute based on this kind of transformation relation, to realize the goal to transform images into base plane.

After overall registration, we should do interpolation calculation to finish the mosaic of images. Because after images being scaled, and rotated, these transformations can solve the matching coordinate of the corresponding pixel points between transformation images and reference image. The pixel coordinate in images are all integral values, and after transformation, there might be nonintegral coordinate. In this way, we should use interpolation to obtain the pixel data of nonintegral coordinate. Interpolation methods in common use now include interpolation of the nearest point, bilinear interpolation, and PV interpolation [42].

The image coordinate transformation method described in 2.3.3 all transform the moving images into the space where reference image is according to image coordinate transformational parameters. This transformation form is called positive-going transition [43]. Positive-going transition is equal to pull and twist the image with fixed number of pixel from the direction of horizontal and vertical. Because the expansion of the image, the image generated might have network cavitation. For this kind of problem, scholars raise some solutions such as enthesis, Splatting method, and gridding-sharing method, etc [44]. Another way to avoid cavitation is transformation by reciprocal direction [45]. Its theory is to find out the position of objective image's each pixel in original image, and then do interpolation to obtain color grey value , this method is from the perspective of object image instead of pulling and twisting original image, in this way, cavitation in positive-going transition will not appear during the transformation.

The steps of transformation by reciprocal direction are shown in figure 2-4:



(a) solve the peak point's coordinates of the original image in the objective image plane

(b) transformation by reciprocal direct and interpolation

Figure 2-4 The Process of Interpolation in Image coordinate transformation

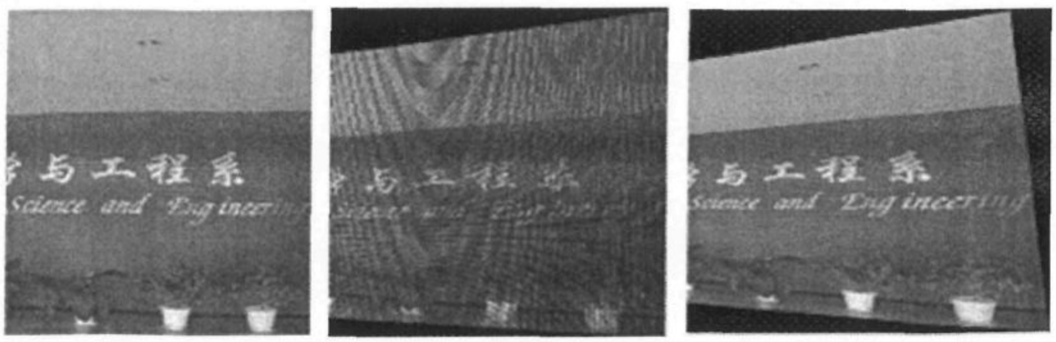
The first step, suppose the transformation from moving images to reference image is $(u_i, v_i) = T(x_i, y_i)$, then transformation function can get four peaks in moving image $(x_m, y_m), m = 0,..,3$, and the coordinate on the plane of reference image is $(u_m, v_m), m = 0,..,3$.

The second step, find out a minimum rectangle area which can completely contain reference image's covering area. And the two peaks on the cross of the rectangle area are:

$$\begin{aligned} u_{\min} &= \min\{u_m\}, v_{\min} = \min\{v_m\} \\ u_{\max} &= \max\{u_m\}, v_{\max} = \max\{v_m\}, m = 0,..,3 \end{aligned} \qquad (2\text{-}10)$$

The third step, for each pixel point in rectangle areas of the objective plane, we can get its corresponding pixel in its moving image through inverse transformation function $(x_i, y_i) = T^{-1}(u_i, v_i)$. If the corresponding pixel is in the covering area of moving image, then we can get the color grey value of this pixel point through image interpolation, and then give obtained color value to the pixel point in reference image. The same operation is done to all pixel points in the objective area, and we get transformation image after interpolation.

Figure2-5 shows the contrast result of positive-going transition and transformation by reciprocal direction of images



(a) moving image (b) results of positive-going transition (c) results of transformation by reciprocal direction

Figure 2-5 Contrast Result of Positive-Going Transition and Transformation by Reciprocal Direction of Images

2.4.2 Image Composite

The main aim of image composite is to guarantee the continuity of superposition area between images. And there are no visible composite lines. At present, there are many methods to composite images' superposition area. They are [46]:

1、 Direct average method

Superposition and Averaging the gray value of corresponding pixel points in superposition area between images after registration. And do low-pass filtering to images. In this way, there will be obvious composite lines in the final mosaic images. If there are moving object in the scene, then the final mosaic image will generate "ghost-like" phenomenon along the trace of moving object.

2、 Median filtering

This method do median filtering to registrated image's superposition area, and in certain conditions, it can overcome the phenomenon of detail blurry in images in direct average method. It can keep image content with higher space frequency. But there are still obvious marks in the final mosaic image.

3、 weighted average method

Weighted average method is similar to direct average method, but the superposition area is not added simply, but adds and average after weighing. For each image, the pixel of center area of the image has higher weighing, and the edge area of the image has lower pixel weighing. In this way, we can realize smooth transition of the contents of image, and remove the mosaic mark in image effectively.

4、 Multi-resolution analysis

This method first divide the image into a series of subband image with different resolution responses, frequency characteristics, and directional feature, and them mosaic in each subspace, and at last compose image data of superposition area in original resolution responses with reconstructing algorithm. In this way, we can get mosaic image. This method can realize the smooth transition of the contents of image effectively. However, the calculation is complicated.

Chapter 3 Self-Adapting Image Mosaic Based on Maximum Mutual Information

Image registration is a key part to guarantee the quality of image mosaic. This chapter comes up with an optimized mosaic method which uses maximum mutual information as image registration principle, builds image pyramid through small wave resolution technique, combines with particle swarm optimization algorithm, which makes its parameters self-adapt according to mutual information and multi-resolution series. Experiments prove that this method can avoid the influences of local extremum effectively. It finds out the optimized registration transformation through finite optimization iteration, and to increase the quality of image mosaic.

3.1 Introduction

Image mosaic mainly includes two steps: registration and composite. The core problem of registration is how to make use of the similarity principle of adjoining image superposition area to solve the space transformational relation between two images. Previous study divided image registration into two methods: full automatic and semi-automatic[47]. The automatic method is registration based on gamma, which mainly include: average error, maximum cross correlation, invariant moments, frequency domain correlation, small wave transformation, etc. the common shortcoming of these methods is that they are sensitive to the change of gray value, especially the change of light will reduce the function of registration.. In recent years, maximum mutual information registration method raised by Mase and Collignon has aroused broad attention in the world. Gamma can be viewed as extraneous variable, and mutual information is the similarity measurement of two extraneous variables. When the corresponding pixel points are consistent between images, its gamma mutual information will be maximized. This method does not need any manual intervention, and has good robustness. Image registration process is a function optimization process, current research shows that because objective function's unsmooth can easily leads to local extremum, which will make it hard to find overall optimization in the process, and leads to wrong matching[48].

In the process of function optimization, the independent variable is the space transformational parameter of image rotation and parallel movement, etc, and induced variable is the similarity of the superposition part of the image. Therefore, there are three problems to solve to realize image mosaic: the selection of similarity standard, the selection of registration area, and the design of optimization algorithm.

The degree of image registration should be measured by similarity standard. Gray value-based cross correlation and mutual information principle are obtained through calculating the statistical characteristics quantity of image. When in itself these two images are the same, then cross correlation and mutual information will reach their max. Literature [49] compares these two similarity standards, and find out that mutual information's antimierophonic ability is stronger than cross correlation, and mutual information's wave crest near the optimum point is sharper, which is helpful for the rapid constringency of optimum algorithm. Therefore, this thesis will take mutual information as the principle to evaluate similarity, and analyze how to use maximum mutual information and optimization to realize self-adapting image mosaic.

3.2 General Introduction of Mutual Information

Mutual information is a basic concept in information theory, which is used to measure two extraneous variables' statistical correlation or the amount of information one variable contained in another[50]. The mutual information between two extraneous variables depends on entropy and the combination entropy between two extraneous variables. Mutual information is used in 1995 in image registration, and now most scholars think maximum mutual information is a good image registration principle.

3.2.1 Entropy

Entropy is also a concept in information theory. In 1948, when analyzing R.V. L.Hartley's method on measure information[51], Shannon used probability to evaluate a system's complexity and uncertainty with axiomatization[52].

As for the extraneous variable of probability distribution function $P(a)$, and the entropy $H(A)$ is defined by the following formula:

$$H(A) = -\sum_a P_A(a) \log P_A(a) \qquad (3\text{-}1)$$

And its physical meaning can be expressed as:

- Entropy means receive information amount a signal provides;
- Entropy means before receiving the signal, receiver has uncertainty with source of information;
- Entropy is the extraneous measurement to extraneous variables, and is directly proportional with randomicity.

To gamma, the more the gray scale is, the more scattered the gamma value is, the larger its entropy is; entropy is also a measurement of grey scale column diagram shape. And when the image column diagram has one or more peaks, the entropy value is generally smaller; if the column diagram is plain, its entropy is larger.

3.2.2. Combination Entropy

The combination entropy $H(A,B)$ is the statistics to test the relativity of extraneous variables A and B. for these two extraneous variables A and B, if their joint probability distribution function is $P_{AB}(a,b)$, then their combination entropy $H(A,B)$ can be defined as:

$$H(A,B) = -\sum_{a,b} P_{AB}(a,b) \log P_{AB}(a,b) \qquad a \in A, b \in B \qquad (3\text{-}2)$$

When misregistration happens to these two images, then their combination gray column diagram becomes dispersed. We could use combination entropy as a measurement of dispersion, and registrate two images through minimum combination entropy [53].

3.2.3 Mutual Information

When $H(A), H(B)$, and $H(A,B)$ represent extraneous variables A and B and their combination entropy respectively, if:

$$H(A,B) = H(A) + H(B) \qquad (3\text{-}3)$$

Then A and B are independent from each other. Therefore, we could use the generalized distance between two variables' joint probability distributions and complete probability distributions as similarity measurement between variables. And it is called mutual information.

Mutual information $I(A,B)$ of extraneous variables A and B can be defined as:

$$I(A,B) = H(A) + H(B) - H(A,B) \qquad (3\text{-}4)$$

$$= H(A) - H(A \mid B) \qquad (3\text{-}5)$$

$$= H(B) - H(B \mid A) \qquad (3\text{-}6)$$

In the formula, $H(B \mid A)$ and $H(A \mid B)$ express B's conditional entropy when A is given and A's conditional entropy when B is given respectively:

$$H(B \mid A) = -\sum_{a,b} P_{AB}(a,b) \log P_{A|B}(a \mid b) \qquad (3\text{-}7)$$

$$H(A \mid B) = -\sum_{a,b} P_{AB}(a,b) \log P_{A|B}(a \mid b) \qquad (3\text{-}8)$$

Their relation is shown in figure 3-1, from formula (3-1), (3-7), (3-8), we can get:

$$I(A,B) = \sum_{a,b} P_{AB}(a,b) \log \frac{P_{AB}(a,b)}{P_A(a) P_B(b)} \qquad (3\text{-}9)$$

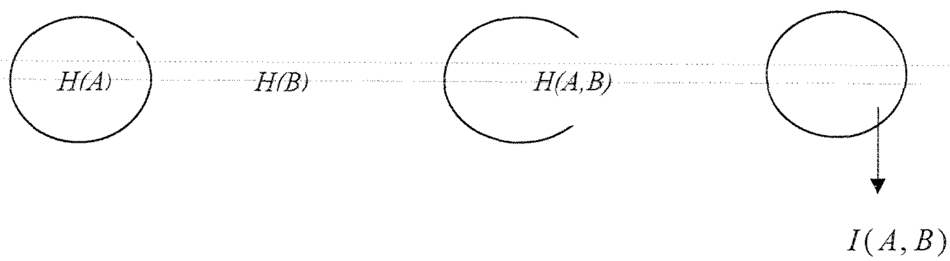


Figure 3-1 Two Extraneous Variables' Mutual Information

Formula (3-5), (3-6), (3-9) can be viewed as the best three definitions of mutual information, and the first two can explain the essence of mutual information in mathematics. For example, $H(B)-H(B\mid A)$ shows when given image *A*, the reduced amount of image *B*'s uncertainty or image *B*'s information amount contained in image *A*. in definition, the positions of *A* and *B* can change. It can also express image *A*'s information amount contained in image *B*.

3.3 Image Mosaic Based on Maximum Mutual Information

Because the method of maximum mutual information registration need no assumption on relation between image pixel grey value, and any preprocess to images. Therefore, this section will discuss the selection of adjoining registration area, the optimization of image transformation parameter, and seamless image mosaic, etc, on the basis of mutual information registration-based evaluation principle

3.3.1 The Evaluation Principle of Mutual Information Registration

Definition 3-1: suppose $T(.)$ is a kind of space transformation, $s(x,y,z)$ is reference image, and $f(x,y,z)$ is moving image, for all the coordinates (x,y,z), if $s(x,y,z)=\ f(T(x,y,z))$, then $T(.)$ is the optimum transformation.

Suppose the photocenter of camera is fixed, and shooting images from different directions rotary, that is, there are only parallel movement $(t_x,t_y,0)$ and rotation *(θ)*, and

$T(.)$ transformation model can be expressed as:

$$T(x,y)=\begin{pmatrix}\cos(\theta) & \sin(\theta) & t_x \\ -\sin(\theta) & \cos(\theta) & t_y \\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}x \\ y \\ 1\end{pmatrix} \qquad (3\text{-}10)$$

Therefore, to determine optimum transformation T is to solve the optimum parameters t_x, t_y and θ. And in order to mosaic images $f(x,y)$ and $s(x,y)$, we should solve space transformation T according to given similarity measurement, which will maximize the similarities between two images. The space mapping relation can be expressed as:

$\alpha_T = \alpha(s(x,y), f(T(x,y)))$, in which α is similarity measurement.

For pixel p whose gray value is k in $s(x,y)$, and its corresponding pixel $T(p)$ in $f(x,y)$ whose grey value is q, k and q's relativity in statistics can be evaluated with mutual information formula (3-9), in which probability distribution is obtained by combination grey value column diagram of the superposition part of two images, therefore, the calculation of mutual information $MI(f(T(x,y)),\ s(x,y))$ is dependent on α_T in essence.

The theory of registration which uses mutual information as similarity measurement is: when two images with the same structure have optimum registration, their corresponding image feature's mutual information will be maximized [54], that is:

$$\begin{aligned}\alpha^* &= \arg\max MI(s(x,y), f(T(x,y))) \\ &= \arg\max \alpha(s(x,y), f(T(x,y)))\end{aligned} \qquad (3\text{-}11)$$

The registration evaluation standard of two images can be regarded as an optimum transformation T. Therefore, registration process is transformed into a process to solve T's 3 parameter's optimum value, that is, image registration based on maximum mutual information is the parameter optimum process. For image registration, mutual information is directly proportional with the size of superposition area of two images [55]. The superposition changes with the change of optimum value of transformation T. Therefore normalized mutual information can best reflect the change of registration function α^*, literature [56] presents the form of normalized mutual information.

$$NMI = \frac{H(a)+H(b)}{H(a,b)} \tag{3-12}$$

In this formula, a and b represent two images to be registrated.

3.3.2 The Process of Mutual Information Image Mosaic Algorithm

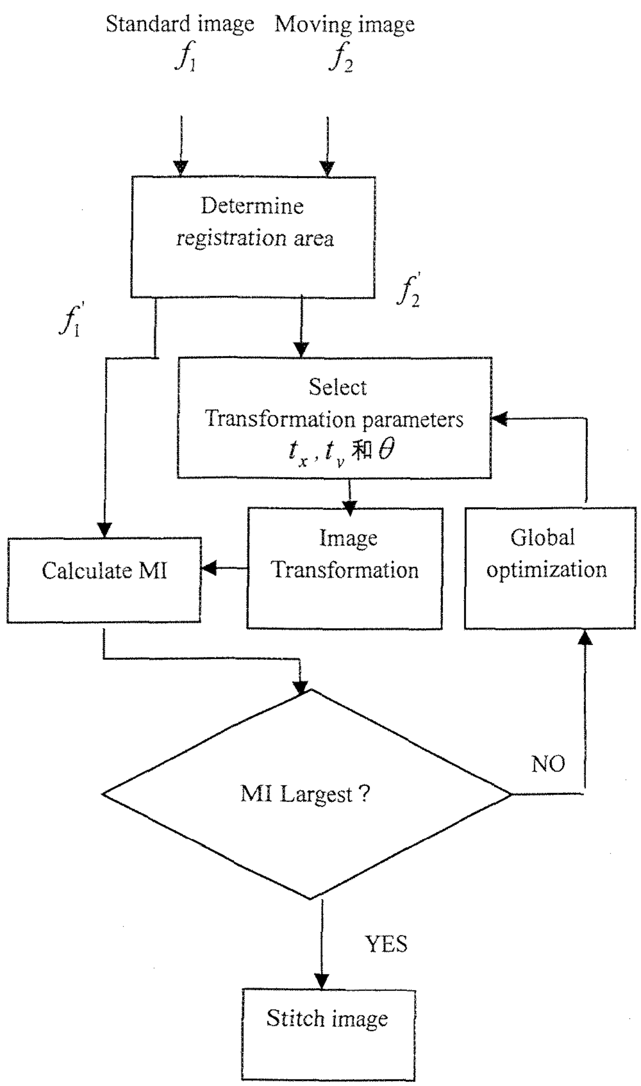


Figure 3-2 The Process of Mutual Information Image Mosaic Algorithm

The process of mutual information image mosaic algorithm raised in this thesis is show in figure 3-2. f_1 represents the basic image, and f_2 represents the moving image. First choose the registration area of adjoining images automatically to reduce the calculation of similarity function; and then through overall optimization algorithm and local search algorithm, optimize 3 transformation parameters t_x, t_y and θ in image registration process by calculating mutual information, until optimum solution T is obtained; and at last do image composite to registrated images.

3.3.3 The Confirming to Registration Area

Image mosaic method based on grey value usually compares all superposition parts of adjoining images to get corresponding transformation parameters when their similarity is maximized. In order to increase the speed of image registration in practical application, we should choose registration area under the condition of making sure the registration precision.

Suppose f_2 and f_1 is the image obtained after parallel moving (t_x,t_y) and rotating angle θ, then we have:

$$f_2(x,y) = f_1(x\cos\theta + y\sin\theta - t_x, -x\sin\theta + y\cos\theta - t_y) \qquad (3\text{-}13)$$

Because image's parallel movement and rotation can all be expressed in Fourier transformation [65], we use FET method to quickly determine the registration area of images . The algorithms are:

1) Transform Fourier of two images

$$F_2(u,v) = \exp[-j2\pi(ut_x + vt_y)]F_1(u\cos\theta + v\sin\theta, -u\sin\theta + v\cos\theta) \qquad (3\text{-}14)$$

If M_1 and M_1 represent value respectively, then:

$$M_2(u,v) = M_1(u\cos\theta + v\sin\theta, -u\sin\theta + v\cos\theta) \qquad (3\text{-}15)$$

2) Transformation of polar coordinates

According to the transformation relation between a point (x,y) in rectangular coordinate system and its corresponding point (ρ,θ) in polar coordinates:

$$x = \rho\cos\theta, \qquad \rho = \sqrt{x^2 + y^2}$$

$$y = \rho\sin\theta, \qquad \theta = \arctan(\frac{y}{x})$$

Transform M_1 and M_2 into polar coordinates :

$$M_2(\rho,\theta) = M_1(\rho,\theta+\theta_0) \qquad (3\text{-}16)$$

Analysis: in polar coordinates, there is only a displacement of angle θ_0 between M_1 and M_2. Use $l_1 \times l_2$ as template in the center of M_1 along the direction of θ with three-step research method[57] to find corresponding part, and solve the evaluation value of θ_0. Because search is done in only one direction, there isn't much calculation.

3)rotate f_2 for angle -θ_0, sampling to get middle image f_3. Because there is only slight rotation between f_3 and f_1, it can be neglected, and then the parallel movement T can be evaluated by f_3 and f_1, that is: $f_1(x,y) \approx f_3(x+t_x, y+t_y)$.

4) define the cross correlation function of two images

$$c(f_3, f_1) = f_3(x,y) ** f_1(-x,-y) \qquad (3\text{-}17)$$

In this formula, ** means two-dimensional convolution. Substitute

$f_1(x,y) \approx f_3(x+t_x, y+t_y)$, and use two-dimensional Fourier transformation and normalized it, then we get normalized cross-power spectrum formula:

$$c(f_3, f_1) = \delta(x-t_x, y-t_y) \qquad (3\text{-}18)$$

This function is 0 in all places except for (t_x, t_y). In practical application, because of the influence of noise, etc, there may appear several points whose value are not 0. we use the peak value and obtain the evaluation value T_0 of T. In this way, we could determine the registration area of two images.

3.3.4 Optimum Algorithm of Mutual Information

The process of image registration based on maximum mutual information is in essence searching 6 space transformation parameters when mutual information is maximized. Therefore the calculation method of mutual information $I(A,B)$ directly relates to the refinement of image transformation parameters and image registration. There are many key technique including interpolation and optimization in calculation mutual information $I(A,B)$. Common interpolation algorithms include nearest neighbor method, linear interpolation method, PV (Partial Volume Distribution) method, etc.

Standard deviation brought by nearest neighbor method is half of sample interval and scanning beam interval. It has less calculation, a certain precision. It is a simple but effective method. And its shortcoming is its obvious discontinuity after gamma processing. Linear interpolation method avoids the shortcoming of gamma discontinuity of nearest neighbor method, but is can add new grey value, and the slight change in space transformation parameters will leads to unprophetable transformation in marginal probability, which leads to miscalculation of mutual information. PV interpolation is good for optimum search. And smaller space transformation increment make the change of mutual information more flattened. However, PV interpolation may lead to local extremum of mutual information function[58].

For PV interpolation may have local extremum and disturbing registration precision, this thesis raises an optimum strategy in mutual information registration method. This method combines improved particle swarm algorithm and multi resolution response based on small wave resolution. It is easy to get rid of local extremum with particle swarm algorithm. And it is good for having overall optimum solution. And with image's multi resolution response which can increase the speed of registration effectively, we can solve the optimum solution of transformation parameters, and research ideal registration effect.

1、 The generation of image with multi resolution response

Multi-resolution analysis is a signal and image processing technique which is popular in recent years. Because the feature of frequency demultiplication in transformation, we can divide the image into similar signal (image) and detail signal (image) in different scales. Therefore multi-resolution analysis can reflect local transformation features of original image, so that it provides advantages for image analysis and process. These similar signals (image) in different scales are called multi-resolution image. Because their shapes are like pyramids', they are called pyramids. There are many structures of multi-resolution image, and this thesis adopts two of them.

■ Gaussian pyramid

Suppose original image is f_0, and we use it as the bottom of Gaussian pyramid. And the constructing process of lth floor of the Gaussian pyramid is: convolve the $l-1$ image and a window function $w(u,v)$ with lowpass feature, and then sample interlacing and intercolumn to the result of convolution. E.g.:

$$f_l(i,j)=\sum_{u=-2}^{2}\sum_{v=-2}^{2}w(u,v)f_{l-1}(2i+u,2j+v) \qquad (3\text{-}19)$$

$1<l\le N$, $0\le i<C_l$, $0\le j<R_l$, in which N is the decomposition layers of Gaussian pyramid, C_l is the $l$the layer image's number of columns, and R_l is the lth layer image's number of rows. Window function $w(u,v)$ is also called weighting function or

product nucleus, and its amount is:

$$w(u,v) = \frac{1}{256}\begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix} \tag{3-20}$$

From top to bottom, from left to right, the linenum u, and columnnum v of the matrix are: $-2,-1,0,1,2$, because the shape of window function is similar to Gaussian distribution, the image pyramid obtained is called Gaussian pyramid. The forming process of Gaussian pyramid implement series of low-pass filtering, that is why with the increasing of decomposition layers, the resolution responses are reducing, and the image becomes blurrier. Figure 3-3 offers the constructing result of this method.

The kth layer(one pixel)

The k-1layer

The k-2 layer

The bottom (original image)

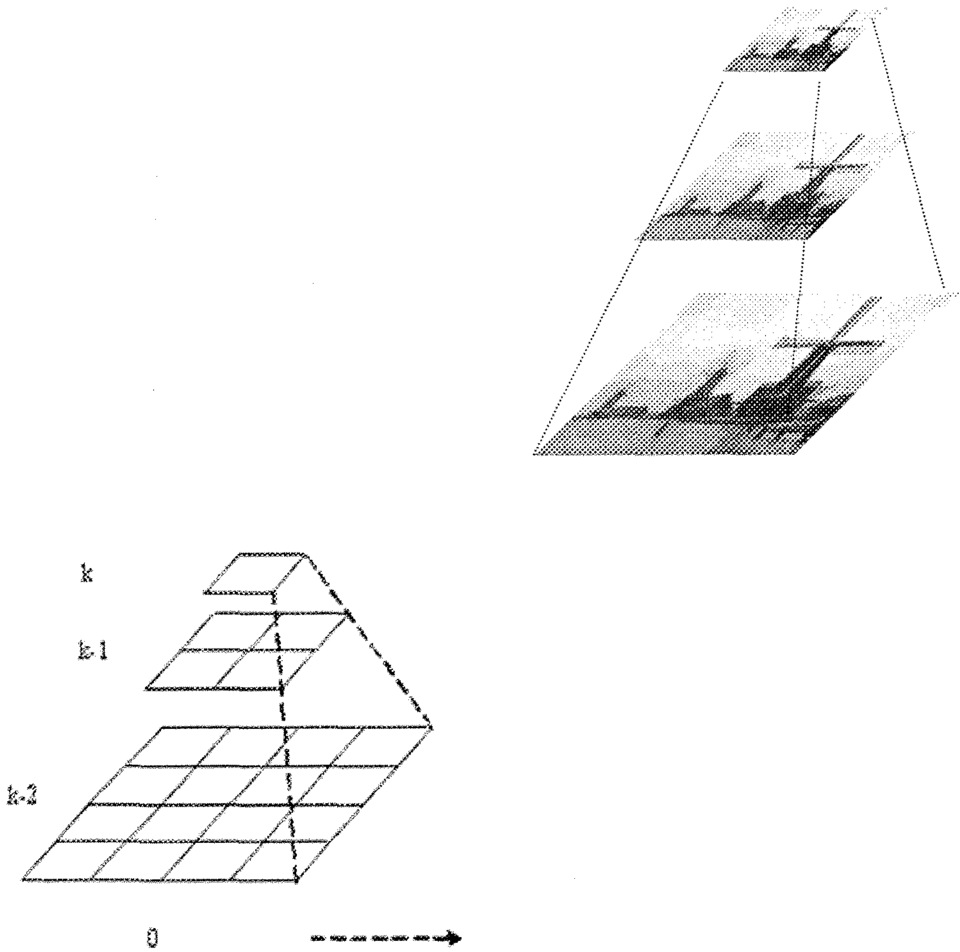


Figure 3-3 The Structure of Image Gaussian Pyramid

- Image pyramid structure based on small wave transformation

Suppose $\{C_0(m,n)\}$ is a two-dimensional image, then for given scaling function and small wave function, we can use the following formulas to do small wave decomposition [59]:

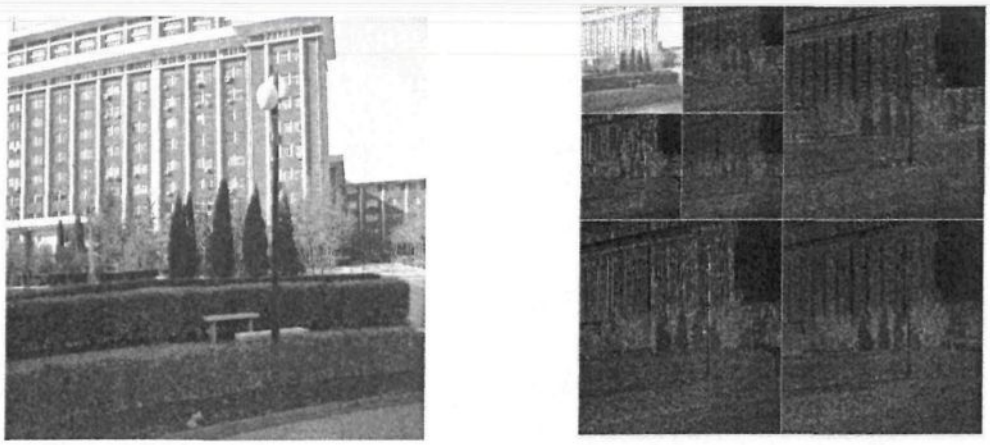
$$c_k(i,j)=\sum_m\sum_n c_{k-1}(m,n)h(2m-i)h(2n-j) \qquad (3\text{-}21)$$

$$d_k^h(i,j)=\sum_m\sum_n c_{k-1}(m,n)h(2m-i)g(2n-j) \qquad (3\text{-}22)$$

$$d_k^v(i,j)=\sum_m\sum_n c_{k-1}(m,n)h(2m-i)h(2n-j) \qquad (3\text{-}23)$$

$$d_k^d(i,j)=\sum_m\sum_n c_{k-1}(m,n)g(2m-i)g(2n-j) \qquad (3\text{-}24)$$

In these formulas, g and h are double-scaling equation's coefficients of scaling function and small wave function respectively, and they are low-pass filter and high-pass filter separately. And k is the measurement. The subimage c obtained from formula (3-21) reflects the smoothness of original image, and it is called low frequency image. And subimage d generated from formulas (3-22),(3-23), and (3-24)reflect the detail features of original image, and are called horizontal, vertical, and declinate high-frequency image, whose corresponding superscript are h , v , and d respectively. Therefore, The image's multi-resolution response analysis is small wave decomposition through a group of low-pass and high-pass filter, and then decomposite the next scale to low frequency image obtained in each layer, thus image pyramid is built. Figure 3-4 offers a real image whose resolution response class is m, and figure 3-4 (a) uses Daubechie4's scaling function and small wave function, and formula (3-21)-(3-24) to realize two decompositions, and obtains decomposed image figure 3-4 (b) .



(a) original image (b) two-layer small wave decomposition

Figure 3-4 The Small Wave Multi-Resolution Response Decomposition of Image

2. Improved particle swarm optimization

Particle Swarm Optimization (PSO) is raised by Kennedy and Eberhart in 1995. it originated from the imitation of simple social system. And at first it imitate the process of bird flock looking for food, and then it is discovered as a good optimization tool, that is, search for optimum solution through the cooperation of individuals in a group[60].

PSO is an evolved computing technique based on swarm intelligence. Swarm intelligence means "features of subject with no intelligence expressing intelligent behavior through cooperation", on the premise that there is no centralized control and overall modal; it provides the foundation to complicated distributed problems' solution [61]. The particle swarm optimization's optimum solution searching mainly depends on its "memory" ability and information sharing mechanism among particles. The main advantage of this algorithm lies in its simple parameters, easy program realization, and guaranteed and quick algorithm convergence。 And it conforms to our need to select search strategy. This algorithm has been used in many areas [62].

The basic idea of PSO is: image each potential solution to optimization problem as a particulate with no volume and quality in d-dimensional searching space, and we call them particle, and each one's position is a solution to the problem. Every particle has a fitness value, which is determined by optimization object to evaluate the searching capability of the particle, and direct the searching process of particle group. When the iterative algorithm stops, the optimum solution variable of sufficiency function is the optimum solution of optimization search. Each particle will fly in a certain speed in searchspace, the speed of particle shows its position change in unit iterations is particle (representing variable)'s movement in d-dimensional space, and search for new solution through adjusting it's position. Each particle can remember the best solution it has searched, which shows single particle's corresponding sufficiency optimization solution from the beginning of searching to present iteration. And the best position of the whole particle swarm has experienced, that is, the global best it has searched. It shows the whole particle swarm's corresponding sufficiency optimization solution from the beginning of searching to present iteration. And its value is the most adaptive one among each particle's present best solutions. Each particle uses the following information to change its position:

1) present position ;

2)present speed ;

3) the distance between present position and the its best position ;

4) the distance between present position and the swarm's best position。

Optimization search is going with the form of iteration in the particle swarm which is formed by random initialization.

The mathematics description of solving transformation parameter between registration images with PSO is:

Suppose the searchspace is *d*-dimensional. And regard potential optimum solution of image transformation parameter as a "particle" in searchspace, and 6 transformation parameters (the parallel moving amount of three coordinate axes and the rotation angles of three coordinates axes) construct object search space of *d*=6-dimensional. Suppose the original state is a swarm $X=\{x_1,\cdots x_i,\cdots x_m\}$ composed by m particles, in which the position of the ith particle is $x_i=(x_{i1},x_{i2},\cdots x_{i6}),i=1,2\cdots m$; we could calculate its adaptive value according to x_i, which is now the value of mutual information of two images, and evaluate the excellence of x_i according to adaptive value. And then suppose the speed of the ith particle is $v_i = (v_{i1}, v_{i2}, \cdots v_{i6})$, and the optimum position it has searched now is $p_{id} = (p_{i1}, p_{i2}, \cdots p_{i6})$, and the overall optimum position the swarm has searched is $p_{gd} = (p_{g1}, p_{g2}, \cdots p_{g6})$. Each particle x_i will gradually change its speed and position according to (3-25) and (3-26) [63].

$$v_{id}^{(t+1)} = v_{id}^{(t)} + c_1 r_1 (p_{id}^{(t)} - x_{id}^{(t)}) + c_2 r_2 (p_{gd}^{(t)} - x_{id}^{(t)}) \tag{3-25}$$

$$x_{id}^{(t+1)} = x_{id}^{(t)} + v_{id}^{(t+1)} \tag{3-26}$$

In these formulas, $d = 1, 2, ..., 6$ and, $i = 1, 2, ..., m$, t is the current revolution algebra, r_1 and r_2 are random numbers between [0 , 1]; c_1 and c_2 are acceleration constants, which are also called learning factors. The moving direction of particles are determined by three parts: its own speed v_{id}, the distance ($p_{id} - x_{id}$) from its best experience, and the distance ($p_{gd} - x_{id}$) from the swarm's best experience, their relative importance is determined by accelerating factor c_1 and c_2 respectively.

Some scholars including Shi[64] improved the formula (3-25) in the following ways：

$$v_{id}^{(t+1)} = w v_{id}^{(t)} + c_1 r_1 (p_{id}^{(t)} - x_{id}^{(t)}) + c_2 r_2 (p_{gd}^{(t)} - x_{id}^{(t)}) \tag{3-27}$$

In this formula, w is nonnegative number, which is called inertia weight, controls the influence of particle's previous speed to the current one, and enables it the trend to expand the space. It can adjust PSO's algorithm's overall and local optimization ability. Research shows that when w is in bigger value, it is easy to skip out of the local search, and strengthen the overall optimization ability; on the contrast, the smaller w value is good for the algorithm constringency, that is, strengthened local optimization ability and reduced overall optimization ability.

Because it takes a lot of time to do mutual information algorithm, we take increasing searching speed as objective, combine multi-resolution responses analysis with PSO algorithm, and improve PSO algorithm according to w value's influence on optimization.

Literature [63] [64]show that w has great influences on optimization performance, it is often a fixed value of 1.4 or less than 1.4. Suppose w=1.4-at, and a is a mediating factor, which guarantee the value of w reduce linearly according to iteration t, experimental results prove that w=1.4-a can get suitable seed in the early stage of searching with relatively strong overall searching ability. After each iteration, we do local optimization combined with multi-resolution responses to make sure the solving speed and increase the precision;

The basic steps of improved PSO algorithm are:

(1) Originalize particle swarm, that is, set randomly the original position x_i and original speed v_i of each particle;

(2) Calculate each particle's sufficiency value (the mutual information of two images);

(3) For each particle, compare its sufficiency value and the sufficiency value of its best position p_{id} it has experienced, and renew p_{id};

(4) For each particle, compare its sufficiency value and the sufficiency value of the whole particle swarm's best position p_{gd} it has experienced, and renew p_{gd};

(5) Adjust the position and speed of particle according to formula (3-25) and (3-26);

(6) If final condition is reached (good position or biggest iteration)), it ends; or it will transform to step (3-26) (2).

The algorithm to solve space transformation parameters with improved PSO and multi-resolution responses is:

(1) Make sure the registration area of two images;

(2) do small wave decomposition of images;

(3) Regard the docomposited current hierarchy images as inputted images;

(4) carry out improved PSO algorithm to do optimization search;

(5) If search to the bottom of image decomposition, then output optimization transformation parameter; if not, turn to (3).

3.3.5 The Seamless Synthesis of Images

After solving the space transformation of adjoining images, we need to mosaic the whole image sequences into a wide-angle image. It includes two steps.

1、 The calculation of image sequences kinematical parameter

When synthesis the image sequence, we could choose a certain image in the sequence as a reference image, and other image do corresponding space kinematical parameter transformation with this image their reference coordinates.

If H_{ij} means the transformation parameter matrix from image I_i to image I_j, then reference image's different positions in the mosaic sequence determines three calculation forms in transformation matrix [65]:

1. If any image I_i takes its previous one I_{i-1} as reference image in image sequence, then it describes a continuous transformation relation in the process of image mosaic, that is, continuous transformation parameter matrix among images are: $H_{21}, H_{32}, H_{43}, \cdots H_{n(n-1)}$;

2. If we take the first image in image sequence as reference image, then it describes each image's accumulated transformation relation compared to the first image in image sequence. If the accumulated transformation parameter matrix of the image is H_{k1}, then its calculation formula is:

$$H_{k1} = \frac{1}{H_{k1}(3,3)} H_{k(k-1)} H_{(k-1)(k-2)} \cdots H_{21} H_{11} \tag{3-28}$$

In this formula, H_{11} is a 3×3 unit matrix, which expresses image's transformation compared to itself. The right side of the formula dividing $H_{k1}(3,3)$ is the regularization to H_{k1}, to make sure that H_{k1}=3.

$$H_{mn} = \frac{1}{H_{mn}(3,3)} H_{m,1} H_{n,1}^{-1} (1 \le m, n \le N) \tag{3-29}$$

3. If takes any image in an image sequence as reference image, them we use absolute transformation parameter matrix H_{ij} to represent each image's transformation relation compared to the selected reference image.

2、 The synthesis of images

Because of a series of transformation among images, if we simply superposition images when mosaic, there will be obvious seam. In order to remove the seam of mosaic, we could synthesis the pixel value of two images according to a certain weight value to a new image in the superposition area. Suppose on a certain scanning beam, images f and f' are superposition in the space interval $[x_1, x_2]$, then the value of a pixel point in this space interval in the synthesized new image f'' is:

$$f''(x) = f(x)w(x) + f'(x)w'(x) \tag{3-30}$$

In which

$$w(x) = \begin{cases} 1 & 0 \le x \le x_1 \\ -x-1 & x_1 < x \le x_2 \\ 0 & x > x_2 \end{cases} \tag{3-31}$$

$$w(x) = \begin{cases} 0 & 0 \le x \le x_1 \\ x-1 & x_1 < x \le x_2 \\ 1 & x > x_2 \end{cases} \tag{3-32}$$

The above calculation is done according to scanning seam, therefore this algorithm is suitable for the synthesis of two images with any shape.

3.4 Experiments and Results Analysis

In order to test the effectiveness of maximum mutual information image mosaic method based on multi-resolution responses and PSO method raised in this thesis, we design the following experiments respectively:

1) Imitating mosaic experiment: intercept two images as images to be mosaiced by hand, record their actual rotation angles and parallel movement deviationvalue between them, and then do Gaussian noise(typical value is 0, and variance is 0.01) to two images separately to imitate images' noise in different situation.

2) Actual mosaic experiment: choose three groups of photos; each adjoining photo's superposition is larger than 10%.

Use mosaic algorithm into the above two experiments, and obtained calculated angle, parallel deviation value, errors in mosaic, calculation time, mutual information, and evolved iteration, etc. and compares it to other mosaic algorithm. All the calculations are done by MATLAB program, and realized on PC.

3.4.1 Experiment Data and Results

1) Imitating mosaic experiment

Intercept 8 groups of image respectively to three photos(figure 3-5): campus of science and technology university(256×256), the Neva (256×256) , and residence (256×256) . Do 24 experiments according to different movements and rotation angles. And its angle variance is between the degree of -8 to +8, and its parallel movement amount is between -100 to +100.

(a) Campus of science and technology university (b) the Neva (c) residence

Figure 3-5 Original Images of Imitate Mosaic Experiment

Table 3-1 The Result of Imitating Mosaic Experiment

(a) the mosaic result of campus of science and technology university

Original image		Multi-resolution response			Imitating anneal[65]			Method in this thesis		
Parallel movement	rotation	Parallel movement registration error	rotation registration error	calculation time	Parallel movement registration error	rotation registration error	calculation time	Parallel movement registration error	rotation registration error	calculation time
[-3 , -7]	-7	3.8	0.61	23.2	2.15	0	25.78	3.98	0	20.1
[-10,-70]	5	2.71	3.39	22.3	0.92	0.2	32.36	0.87	0.2	30.2
[-20,-70]	-3	3.46	0.25	20.56	0.31	0	39.41	0.29	0	33.21
[20,-80]	7	3.11	3.09	23.45	2.40	3.0	38.13	2.46	3.21	35.78
[-14, 76]	3	7.20	0.25	23.40	2.36	0.21	37.58	3.01	0.	32.6
[5, 50]	-5	9.89	4.91	24.86	2.28	0	38.87	3.68	0	35.97
[-30, -80]	1	0.04	0.69	23.67	0.29	0.20	25.6	0.28	0.2	23.8
[-25,-75]	-1	0.69	0.12	22.08	3.82	0.20	24.87	0.54	0.2	20.43
Mean value		3.61	3.16	22.19	3.57	0.23	32.83	3.14	0.23	28.76

(b) the result of "the Neva"

Original image		Multi-resolution response			Imitating anneal[65]			Method in this thesis		
Parallel movement	rotation	Parallel movement registration error	rotation registration error	calculation time	Parallel movement registration error	rotation registration error	calculation time	Parallel movement registration error	rotation registration error	calculation time
[-20,50]	-8	2.67	0.89	20.56	3.78	0.20	23.87	3.54	0	23.56
[8, 30]	-5	0.12	0.86	19.65	3.54	0.6	28.13	3.06	0.2	26.09
[56,-49]	-3	3.45	0.81	23.22	2.88	0	25.34	2.21	0	24.32

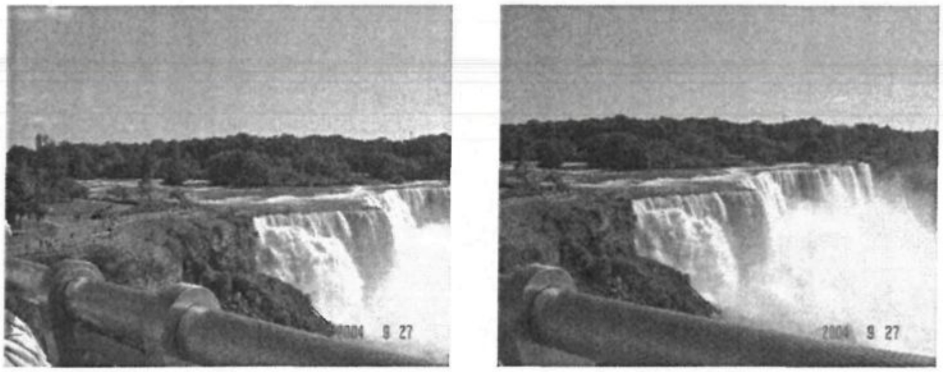
[5, 50]	-1	2.65	0.38	20.32	3.8	0.4	24.81	3.65	0.2	22.23
[30, 10]	3	2.31	0.89	22.76	4.02	0.6	26.76	2.68	0.6	26.06
[20,-80]	1	3.27	0.85	22.45	2.43	0.2	24.21	2.08	0.2	23.33
[3, 78]	5	3.4	0.2	22.89	0.41	0	28.94	3.2	0	24.31
[10,15]	8	8.98	4.2	23.26	0.97	0	27.41	0.6	0	24.06
meanvalue		2.61	3.14	23.89	3.98	0.25	26.18	3.63	0.15	23.99

(c) the result of "residence"

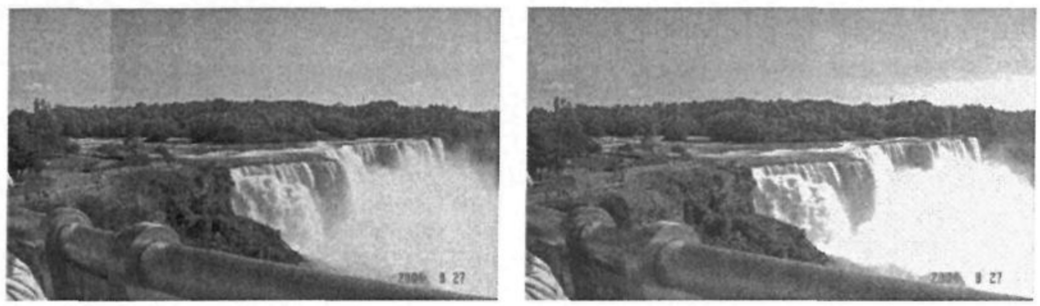
Original image		Multi-resolution response			Imitating anneal[65]			Method in this thesis		
Parallel movement	rotation	Parallel movement registration error	rotation registration error	calculation time	Parallel movement registration error	rotation registration error	calculation time	Parallel movement registration error	rotation registration error	calculation time
[-5,12]	-7	2.98	3.87	19.32	3.68	0	20.02	3.2	0	19.54
[10,8]	-5	3.34	0.67	18.32	3.65	0.4	23.65	3.6	0.2	23.88
[-9,-3]	1	3.21	0.26	16.98	3.75	0.4	20.78	3.58	0.2	19.86
[5, 7]	-3	3.51	3.5	16.09	2.46	0.2	23.34	3.85	0.2	18.88
[-18,-10]	7	3.66	0.8	17.06	0.64	0.2	22.18	0.6	0.2	19.4
[4,-9]	3	3.67	0.41	18.44	3.6	0.2	23.85	3.0	0.2	20.4
[-20,-5]	5	6.8	0.2	17.11	3.6	0.4	28.12	3.6	0.2	23.47
[-18,-12]	8	3.86	0.71	17.23	0.54	0.2	24.23	0.32	0.2	20.33
平均值		3.13	3.05	17.57	3.74	0.25	23.02	3.47	0.18	20.22

2) Actual mosaic experiment

(1) the photo of "Niagara" (256×256)



(a) reference image (b) reference diagram



(c) The mosaic effect of multi-resolution response direct searching method (d) mosaic effect of algorithm raised by this thesis

Figure 3-6 Actual Mosaic Experiment

	Average error		Maximum mutual information	Calculation time(s)
	Parallel movement error	Rotation error		
Multi-resolution response	2.68	3.15	3.706	23.44
Imitating anneal	2.01	0.2	3.713	29.57
Algorithm in this thesis	3.89	0.2	3.7421	23.86

Table 3-2 The Experiment Result of Mosaic of "Niagara"

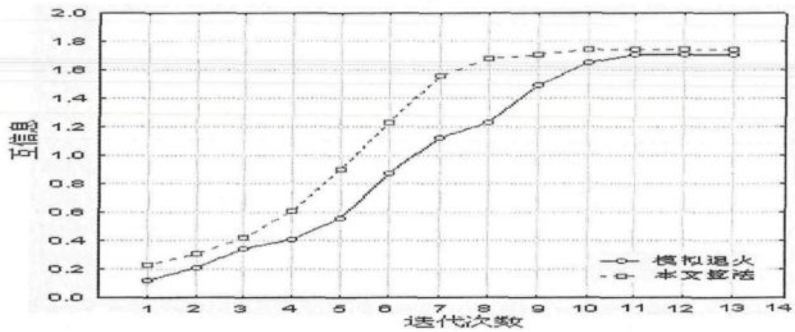
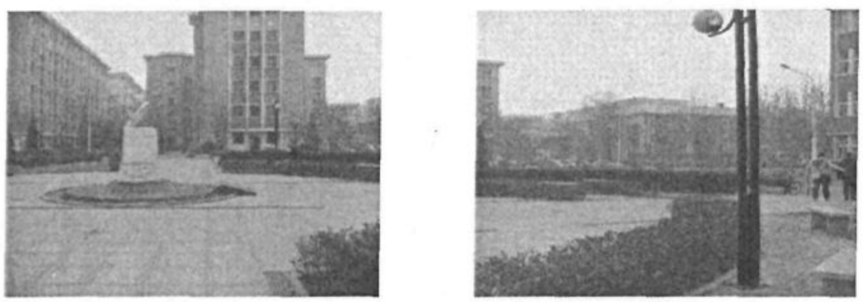
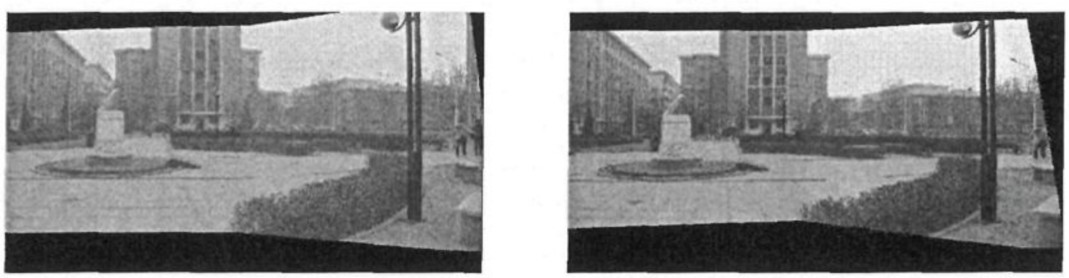


Figure 3-7 Relation of Two Optimization Algorithm Mutual Information's Increasing According to Evolved Algebra

(2) actual photo of the "campus" (256×256)



(a) reference image (b) reference diagram (superposition25%)



(c) Mosaic effect of multi-resolution response direst searching method (d) mosaic effect of algorithm in this thesis

Figure 3-8 Image Mosaic of Campus

	average error		Maximum Mutual information	Calculation time(s)
	Parallel movement error	Rotation error		
Multi-resolution response	3.09	3.22	0.89	28.44
Imitating anneal	2.08	0.6	0.93	35.68
Algorithm of this thesis	3.98	0.6	0.94	33.78.

Table 3-3 The Result of "Campus" Image Mosaic

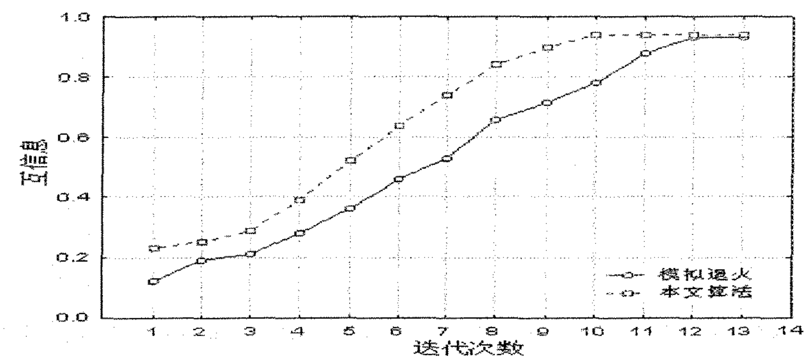
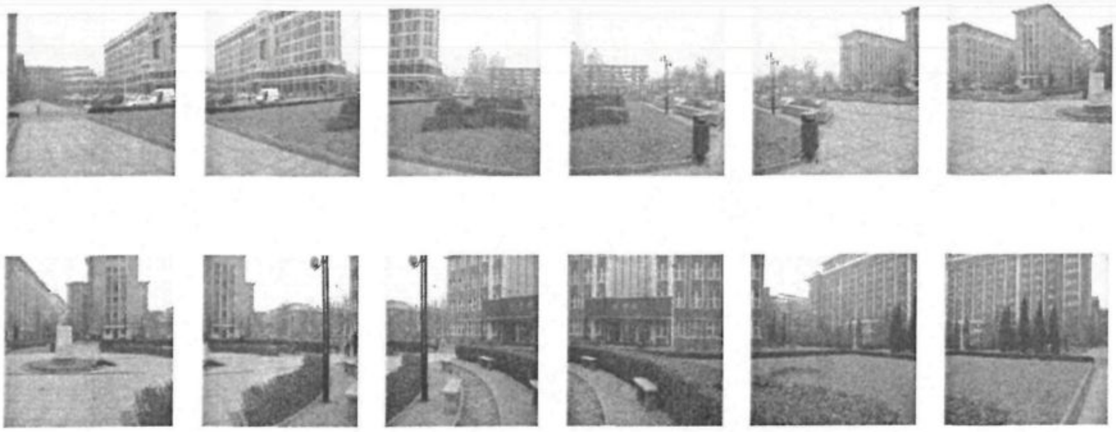


Figure 3-9 Relation of Two Optimization Algorithm Mutual Information's Increasing According to Evolved Algebra

(3) Mosaic of campus panorama



(a) Sequence images of science and technology university

(b) Panorama of science and technology university

Figure 3-10 Panorama Mosaic of Sequence Images

3.4.2 Analysis of Experimental Result

In imitating experiment, this thesis uses multi resolution response directs searching method, imitating anneal in literature [63], and mosaic algorithm offered in this thesis. We do mosaic experiment to these 24 groups of images, table 3-1(a) , (b) , (c) shows the results of the experiments. The unit of errors in this table is pixel, and the unit of time is second.The unit of errors in this table is second, and the unit of time is second.

In experiment, we adopt image with the resolution response of 640*480. We can figure that the mosaic with more than 5 pixels error in parallel movement registration, or more than 3 degrees rotation registration errors are failure.

Form data in table 3-1 we can see that : in image registration precision, no matter in angle error or parallel movement error, the method offered in this thesis is smaller than errors brought by adopting multi-resolution response direct searching method and literature[65]; in the aspect of success ration, there are totally 3 mosaic failure in 24 experiments of multi-resolution response direct searching method; in the aspect of computational efficiency, method offered by this thesis is better than that in literature [63], but a little worse than multi-resolution response direct searching method. Because the mosaic precision and success efficiency is improved greatly, image mosaic offered in this thesis is far better than traditional multi-resolution response direct searching method, and it improved the calculation efficiency of imitation anneal optimization.

When use this image mosaic algorithm to scene image mosaic experiment which is collected in actual situation, the result shows that it is better than previous algorithms. In order to compare the algorithm capacity, figure 3-6 and 3-8 offers mosaic effect of two images with different superposition areas with multi-resolution response direct searching method and method raised in this thesis respectively. Because the pollution of image noise, multi-resolution response direct searching method easily to be misled to local extremum when it is registrated at the first time at the bottom, therefore from the wrong start, multi-resolution response is impossible to find a real overall optimum solution. On the other hand, mosaic algorithm based on the combination of PSO and multi-resolution response can converge nearly to overall optimum solution quickly, and then through local searching, it can get precise solution quickly. The experiments' results in figure 3-7 and 3-9 show that, the method in this thesis can get maximum mutual information and converge to optimum solution in less than 10 times iteration algorithm, and its capacity is 5% better than method offered in literature [66].

The experiments result of panorama mosaic through actual photo shows that, the algorithm offered by this thesis realizes the even transition of image contents without obvious seam or "ghost". But in the experiments we found that, the images need to be mosaiced in algorithm in this thesis have some superposition areas, and situations which have superposition with more than 50 pixels can reach good effect.

Chapter 4 Feature-Based Robust Image Mosaicing Algorithm

Researches in this chapter mainly focus on feature-based image mosaicing algorithm. First, identify similar regions between the two images by using the phase correlation approach offered in Chapter 3, and thus narrow the scope of feature detection; second, put the similar regions under Multi-resolution decomposition, extract Harris corners to solve the problem of Harris operator's easy being impacted by the scale. Then on the basis of principle component analysis, reduce dimension of the High-dimensional feature description vector to enhance the operational image mosaicing efficiency; finally, bring forward an approach, which could automatically identify image mosaicing order according to the number of effective matching points, and could optimize parameters by making use of the Lmeds to mosaic images precisely.

4.1 Introduction

According to researches and experiments in Chapter 3 we could see that to solve the problem of optimizing transformation parameter by using the gray value directly can make full use of the original information of the image, figure out highly precise transformation parameters via calculation, and finally realize image mosaicing. But this process involves too much calculation work, so it takes too much time. Feature-based image mosaicing is an approach in which we could first extract features from the image, and then figure out parameters of space transformation by establishing features-matching-based relationship between images. Compared with direct connection approach based on gray value, feature-based approach can reduce cost of calculation, but how to extract features and build the relationship between images remains to be a hard nut to crack.

Corner being one of the features of an image, is the point which intensely transforms in gray value in the edging areas. It differs violently from its surrounding points, so it shows features of the image well. By detecting corners of the image, we could get the general description of the whole picture.

There are many methods for corner detection of an image. The very first approach, geometry-based corner detector is to extract the edge of the image as the chain code by dividing and cutting, then find some meaningful turning points in the edging area[67]. Sometimes, if the image is too complicated, detection effect would not be very satisfying. So scholars at home and abroad conducted extensive research on this subject, and then brought forward many corner detecting approaches in which we could operate gray-scale images much more directly. But this kind of approaches is lack of common adaptability, and is hard to resort to under general circumstances.

In 1982, Kitchen, for the first time proposed to take gray-scale changes of the neighborhood of the image pixels into consideration, and to detect corners in the image by calculating curvature and gradient of the points. Kitchen found that if local gradients are multiplied, gradient direction would change, and then it would become easier to identify corners[68]. Wang and Brady found that all the curvatures of gray-scale image are in direct proportion to normal edge, and are inversely proportional to edge strength [69]. Moarvec brought forward interest point detector—one local window in an image moves in parallel in four directions then get the average change rate of the gray-scale [67]. Harris and Stephens modified this algorithm, by estimating the self-related differential of the gray-scale of the pixel to detect corners. As the most classical corner detection approach, it has been proved by experiments that feature points extracted by Harris operator have the characteristics of sound image rotation, parallel movement and illumination invariance[70].

4.2 Feature Point Detection of Images

This thesis extracts feature points of images based on Harris operator, introduces in methods of different scale images detecting and ways of dimension-reducing of feature descriptors, purposes to enhance the effectiveness of image feature mosaicing and reduce the expense of calculation.

This thesis proposes that before feature corner detection we should firstly make use of the given phase correlation to figure out translation initial transformation and rotation angle of the images which are going to be mosaic. And thus, identify the overlapping regions of the two images, and then detect corners in this region with Harris operator.

4.2.1 Basic Principles of Harris Operator

Harris corner detection algorithm was first brought forward by Harris and Stephens in 1988. The Principle of Harris operator [71] is:

Assuming the image brightness value is $f(x,y)$, after a slight shift (u,v) of template window $w(x,y)$, take the changed brightness value of pixel (x,y) as $E(x,y)$, then:

$$E(u,v)=\sum_{x,y} w(x,y)[f(x+u,y+v)-f(x,y)]^2 \quad (4\text{-}1)$$

$$E(u,v)\cong[u\ v]M\begin{bmatrix}u\\v\end{bmatrix} \quad (4\text{-}2)$$

Where

$$M=\begin{bmatrix}A & C\\ C & B\end{bmatrix}=w(x,y)\otimes\begin{bmatrix}f_x^2 & f_x f_y\\ f_x f_y & f_y^2\end{bmatrix} \quad (4\text{-}3)$$

f_x、 f_y are respectively the first derivatives of the pixel along x、 y, $w(x,y)$ takes Gaussian function in order to better deal with the noise.

If two Eigen values of Matrix M have local maximums, then this point would be a corner. Figure 4-1 gives the mobile test on the template window w in the image.

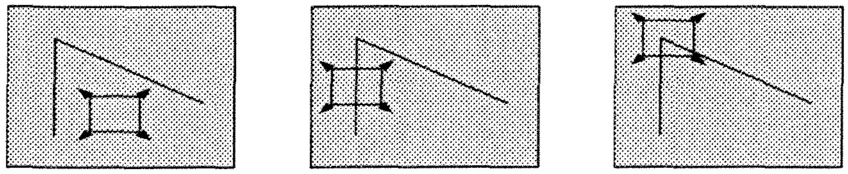
In order to avoid calculation on Eigen value, we take Evaluation function:

$$R=M_d-kM_t^2 \quad (4\text{-}4)$$

In this function, M_d stands for determinant value of Matrix M, M_t stands for trace of Matrix M; k is a fixed parameter, general admission k=0.6.

Corner is defined as R more than a threshold corresponding to the location of the point, the size of the threshold is decided by the quantity of required feature points.

Quantity of corners is decided by the reliable probability of robust algorithm [71].



(a) In the area covered by the sliding window, when the flat window slides in all directions gray density value E rarely changes	(b) The window locates on an edge, when it slides along the edge, E change is small; when it slides perpendicular to the edge, E change is big.	(c) The window covers one corner, gray density value E changes greatly whichever direction it slides in

Figure 4-1 Illustration of Harris Operator Testing Principle

According to figure 4-1, firstly we could see that Harris is the operator of the isotropic. So when the image rotates, corner testing can not be influenced. Harris operator's rotation invariance is very important to detections of feature points.

Second, Harris corner detection makes use of the first derivatives of every pixel along the horizontal and vertical direction to detect changes of the gray value. So the enhance of light of the overall image can not affect the detection of the corner.

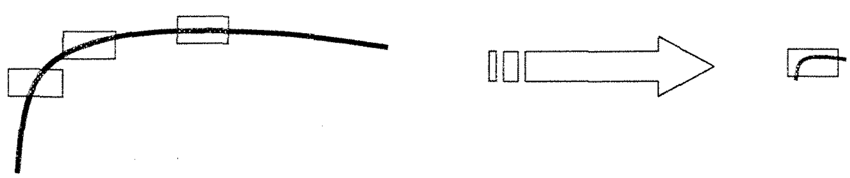


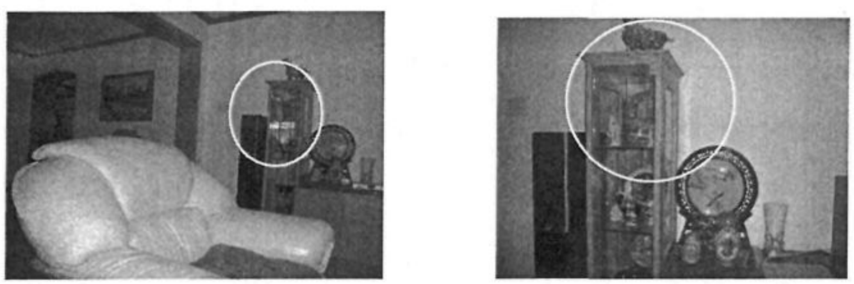
Figure 4-2 Impacts of Scale on Harris Operator

But, Harris operator is easily affected by image scale. It depends on Harris corner detection window whose size is fixed. As shown in figure 4-2, radian of one curve in the source image is moderate, so Harris operator has detected and recorded three corners; when the scale is reduced, radian of the curve changes quickly, and is covered by the Harris detection window, so Harris operator only gets one corner.

When scene photos are taken, there might be some changes of the focal length of the camera settings and changes of shooting distance between the camera and the object, because of time difference. And that leads to different resolutions in shooting of the same scene. Assuming the distance between the camera and the object is d, and the focal length of the camera lens is F. Image resolution increases along with F's increase, and reduces with the decline of d, so $r = F / d$ illustrates a close resemblance to the image resolution [72].

Look at the two images in Figure 4-3. Focal length of the left one is 10mm, while the right one is 15mm. We could notice that, there is an area in both the high-resolution image and the low-resolution image which coincide with each other, and the low-resolution image (the left one) covers a larger scope of the scene.

When mosaicing images with different resolutions, Harris operator can not be used directly for feature point detection, because changes of resolution equal to changes of scale, and Harris operator has the problem of space-scale instability [72]. To this end, we will offer feature detection approach of images with different resolutions through discussion in the following part.



(a) low-resolution image (b) high-resolution image

Figure 4-3 Scene Image with Different Resolutions

4.2.2 Feature Point Detection Algorithm of Images with Different Scales

The main defect of Harris operator is it is greatly affected by the change of the scale. As for this, this thesis offers another algorithm with which we could avoid using space scale, but still use Harris operator to detect feature points of images with different resolutions. This algorithm combines classical Harris operator corner detection with Gauss theorem[73], making use of Gaussian kernel image convolution to get the different scale spaces of the image, and then carry out feature point detection in different scale spaces. The algorithm is as follows:

(1) Calculate the scale factors of the two mosaicing images $s = r_1 / r_2$, r_1 , r_2 respectively stands for the resolution of image1 and image2, set $r_1 < r_2$.

(2) Construct Gaussian pyramid structures of different scales containing $i = 1, 2, ..., s$ for image 2.

(3) Extract Harris corner of each image scale in the Gaussian pyramid. Every corner shows its image-scale and its coordinate in this image scale.

(4) Gather all the corners extracted from different image-scales into a corner collection, and take it as S2—collection of all the input feature points of image 2.

(5) Extract corners from image 1 to build a feature point collection of image 1, S1.

Figure 4-4 shows the Gaussian pyramid structure of Figure 4-3(b), and the situation of corner extraction of each image scale. We could see that as scale becomes smaller, Harris operator extracts less feature points gradually.

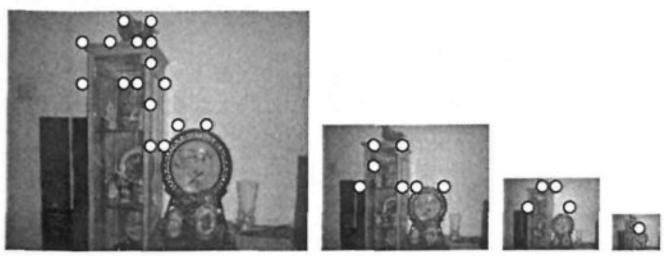


Figure 4-4 Corner Detection of Different Image Scales of the Gaussian Pyramid Structure

After corner collections S1 and S2 have been extracted from the two images with the above mentioned algorithm, it is necessary to build descriptors with certain data structure for the feature points in order to search and compare feature points from different collections effectively and finally figure out the largest corresponding point collection among images.

4.2.3 Construction of PCA (Principle Components Analysis)-based Feature Description Vector

4.2.3.1 Basic Approach of Descriptor Construction

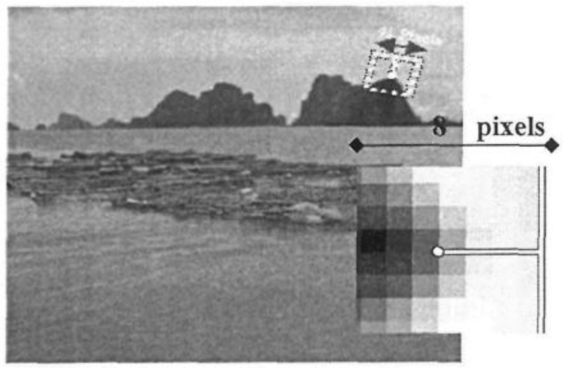


Figure 4-5 Rotate the Sample Coordinate System According to the Direction of the Feature Point, Extract Sample from the 41×41 Region after Rotation by Factor 5, and Get a Description Sample of 8×8

Set a feature point (x, y, θ) in scale space l($l = 1,2,...,s$), θ stands for direction vector of the feature point. Construction approach of the descriptor is as follows [74]:

In the 41×41 region centered with (x, y), take a feature point sample every six points in θ's direction, and get an 8×8 sample, as shown in Figure 4-5. Take it as a feature description block of the feature point, call it feature description vector. As for the solution of θ and sampling method, please consult to literature [74].

Then normalize feature description vectors extracted from each feature point, and make their element mean value 0, standard deviation 1. Feature description vectors could get rid of influence of the intensity of illumination through normalization, because feature points have illumination intensity robustness under comparison.

$$d_i = (d_i' - \mu) / \sigma \qquad (4\text{-}5)$$

Where $d_i, i \in \{1, ..., d^2\}$ is a element of description vector, and

$$\mu = \frac{1}{d^2}\sum_{i=1}^{d^2} d_i$$

$$\sigma = \sqrt{\frac{1}{d^2}\sum_{i=1}^{d^2}(d_i - \mu)^2} \qquad (4\text{-}6)$$

Take one of the 8×8 normalized feature description vectors as the data structure of the comparison between corresponding feature points of the images, and save it in EMS memory or the files.

Since every feature description vector shows shape feature of the image. Usually, low-dimensional feature vectors can only describe the general frame of an image, and only high-dimensional feature vectors can offer the details of the shape of an image. Feature point descriptor constructed in the approach mentioned in the literature [75] can reach a vector dimension of 8×8=64. And these high-dimensional feature vectors may cost much in calculation in the following steps of feature points searching, comparison, and estimation of the transformation parameters.

So in order to reduce the complexity of the image processing algorithm, and to improve the accuracy of the results, in this thesis, we introduce principle component analysis approach into image corner feature vector, analyze and dimension-reduce feature vector with 64 dimensions, hoping to reduce the cost of calculation by reducing high-dimension feature of the image, under the precondition of preserving corner feature information.

4.2.3.2 PCA-Based Dimension Reducing Processing of Feature Vector

PCA (Principle Components Analysis) is a statistical analysis method, which can turn a number of indicators into a composite indicator of the few [75]. By projection, PCA projects high-dimensional data to low-dimensional space with minimal information loss. And thus, reduce dimension of the data, to achieve the goal of simplifying the structure of the data. It is also an approach to turn a number of correlated variables into a few irrelevant variables through comprehensive simplification and with the minimal information loss.

Assuming there are n feature points from an image, and we get p indicator (factor) data from each point, so we have totally np data, which form a feature sample matrix. And those p indicators often have mutual impact on each other. PCA is to detect a few indicators, which could comprehensively represent all the information about the original p indicators (factors), at the same time irrelevant to each other. According to the idea of PCA, dimension reducing method is as follows:

Set p as the quantity of elements (dimension) of each corner feature vector, and n as the quantity of corners detected in the image, thus we have the whole collection X of all the p dimensions and its sample with volume $n\{x_1, x_2, \cdots, x_n\}$, all of which form a sample matrix of $n \times p$ steps.

(1) Calculate the covariance matrix Σ of the sample matrix:

$$\Sigma = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x})$$

(2) Calculate Eigen value and feature vector of Σ: set the obtained Eigen value as $\lambda_i (i = 1, 2, \cdots, p)$, and $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p \geq 0$, set the corresponding feature vector as $e_1, e_2, \cdots, e_p$, and $e_i' = (e_{i1}, e_{i2}, \cdots, e_{ip})$, $e_i' e_i = 1$, $(i = 1, 2, \cdots, n)$.

(3) Calculate the i th principle component: $Y_i = e_i' X = e_{i1} X_{(1)} + e_{i2} X_{(2)} + \cdots + e_{ip} X_{(p)}$, $i = 1, 2, \cdots, p$

(4) Calculate the information contribution rate of the ith principle component: $\lambda_i / \sum_{i=1}^{p} \lambda_i$.

(5) If the cumulative information contribution rate of the first r principle component $\sum_{i=1}^{r} \lambda_i / \sum_{i=1}^{p} \lambda_i \geq 85\%$, then choose $Y_1, Y_2, \cdots, Y_r$ to be feature factors.

According to the above mentioned dimension reducing algorithm, the thesis selects

5 scene images as are shown in Figure 4-6 to carry out the dimension reducing experiment. In the experiment, first gray the color images, and extract no less than 20 corners from each image using Harris operator. Then construct feature description vector for each corner using the method offered in Section 4.2.3.1; finally, take the neighborhood pixels in an area of 7×7 and 13×13 centered each corner of the images as sample factors to analyze the principle component.

(a) (b) (c) (d) (e)

Figure 4-6 Sample Images of Dimension Reducing Experiment

Table 4-1 and Table 4-2 show the results of PCA experiment of Figure 4-6(a)

Table 4-1 Result of Experiment of 7×7 Data 49 Dimension

(a) Eigen value & Cumulative contribution rate

	Eigen value	**% Total**	**Cumulative**	**Cumulative**
1	33.92742	69.23963	33.92742	69.2396
2	5.03909	10.28385	38.96651	79.5235
3	4.94102	10.08372	43.90753	89.6072
4	1.47505	3.01031	45.38258	92.6175
5	0.98879	2.01794	46.37137	94.6355
6	0.64871	1.32390	47.02008	95.9593
7	0.43365	0.88499	47.45373	96.8443
8	0.30573	0.62394	47.75946	97.4683
9	0.27281	0.55676	48.03227	98.0250
10	0.22022	0.44944	48.25249	98.4745
11	0.20120	0.41061	48.45369	98.8851
12	0.15754	0.32152	48.61123	99.2066
13	0.11966	0.24420	48.73089	99.4508

14	0.09466	0.19318	48.82555	99.6440
15	0.06194	0.12642	48.88749	99.7704
16	0.04227	0.08626	48.92976	99.8566
17	0.03721	0.07594	48.96697	99.9326
18	0.02527	0.05157	48.99224	99.9842
19	0.00776	0.01584	49.00000	100.0000

Carry out principle component analysis on 49 (dimensions) feature variables using statistical analysis software SPSS13.0, and get the variation cumulative contribution rate of the first 6 principle components, 95.96%. Therefore, the first 6 principle components have sufficiently extracted all the information offered by the 49 feature variables. Variation contribution of the first 6 principle components is as listed in Table 4-1(b), the score of the first 6 principle components are shown in Table 4-1 (c). We could take the score of the first 6 principle components as feature variables, and then go on analyzing them. In this way, with minimum information loss, it greatly reduces feature dimension, and enhances the efficiency of data analysis.

(b) Variation contribution of the first 6 principle components

Principle component serial number	Eigen value λ	Variation contribution rate (%)	Variation cumulative contribution rate (%)
1	33.92742	69.2396	69.2396
2	5.03909	10.28385	79.5235
3	4.94102	10.08372	89.6072
4	1.47505	3.01031	92.6175
5	0.98879	2.01794	94.6355
6	0.64871	1.32390	95.9593

(c) Score of the first 6 principle components

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
1	1.44839	0.57627	-0.26674	-0.95503	0.44847	-0.30441
2	1.30802	0.40181	-0.22726	0.51428	0.49272	-0.46134

3	1.16752	-0.93129	-0.08859	1.85866	0.76392	-0.91044
4	1.03820	0.57653	1.69462	-0.10600	-1.58486	-0.36582
5	0.18625	0.33665	-0.01038	-1.93102	0.17284	-0.40293
6	-0.30356	1.04766	0.95427	-0.17317	1.25441	0.01095
7	-1.42504	2.20357	-2.60729	0.02205	0.77132	-0.81727
8	-1.80412	-1.18817	1.49207	-0.91602	1.58432	0.04627
9	-1.20606	-0.56103	0.87288	0.77845	-0.19659	-1.98126
10	-1.28116	1.23871	0.62979	2.47741	-1.29375	1.08456
11	0.68699	-0.20932	0.14128	-0.33015	0.01854	1.21747
12	1.15718	0.14057	-0.33507	0.34234	-0.10516	1.09705
13	-0.31573	-0.28404	-0.69029	-1.35360	-2.47821	-0.37235
14	0.54581	-0.42467	0.15971	0.16716	0.57537	-1.03574
15	-0.08183	-2.59069	-2.03488	0.56272	-0.36643	0.16961
16	0.16829	0.02426	-0.11438	0.04885	1.13196	2.49394
17	0.17435	0.41121	0.04685	-0.30033	-0.85404	0.09179
18	0.04267	-0.17430	0.06587	0.23535	0.29275	-0.59762
19	-1.39798	-0.82915	-0.01313	-0.57340	-0.56707	1.13037
20	-0.10819	0.23542	0.33066	-0.36855	-0.06052	-0.09281

Similarly, Table 4-2 lists results of PCA carried out on 169 (dimension) feature variables in a region of13×13 using the statistical analysis software SPSS13.0, and gets the variation cumulative contribution rate of the first 9 principle components, 95.32%. Therefore, the first 9 principle components have sufficiently extracted all the information of the 169 feature variables. Variation contributions of the first 9 principle components are as shown in Table 4-2 (b), and the scores of the first 9 principle components are as shown in Table 4-2 (c).

Table 4-2 13×13 **Data, 169 Dimensions**

(a) Eigen value & Cumulative contribution rate

	Eigen value	% Total	Cumulative	Cumulative
1	100.8321	59.66394	100.8321	59.6639

2	28.6887	16.97558	129.5208	76.6395
3	10.1617	6.01282	139.6825	82.6523
4	6.9348	4.10344	146.6173	86.7558
5	4.8251	2.85506	151.4423	89.6108
6	3.5456	2.09800	154.9880	91.7088
7	2.3939	1.41648	157.3818	93.1253
8	2.0697	1.22468	159.4515	94.3500
9	1.6447	0.97317	161.0962	95.3232
10	1.5227	0.90102	162.6189	96.2242
11	1.3894	0.82215	164.0083	97.0463
12	1.0171	0.60184	165.0254	97.6482
13	0.8947	0.52939	165.9201	98.1776
14	0.8159	0.48277	166.7360	98.6603
15	0.7189	0.42538	167.4549	99.0857
16	0.5043	0.29839	167.9591	99.3841
17	0.4063	0.24042	168.3655	99.6245
18	0.3791	0.22430	168.7445	99.8488
19	0.2555	0.15117	169.0000	100.0000

(b) Variation contribution of the first 9 principle components

Principle component serial number	Eigen value λ	Variation contribution rate (%)	Variation cumulative contribution rate (%)
1	100.8321	59.66394	59.6639
2	28.6887	16.97558	76.6395
3	10.1617	6.01282	82.6523
4	6.9348	4.10344	86.7558
5	4.8251	2.85506	89.6108
6	3.5456	2.09800	91.7088
7	2.3939	1.41648	93.1253
8	2.0697	1.22468	94.3500
9	1.6447	0.97317	95.3232

(c) Scores of the first 9 principle components

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9
1	1.53990	-0.13779	-0.65619	0.24192	0.25361	-0.81278	-0.16647	0.42018	1.31319
2	1.36708	0.29140	-0.54189	0.23101	-0.14893	0.01917	-0.21579	-0.38006	-0.02771
3	1.11814	-0.22595	0.26717	0.47680	0.83559	-1.07314	-0.72375	-1.17777	0.69585
4	0.83440	0.87337	1.42690	0.50480	-0.65750	-0.06403	0.77765	1.05627	-0.18285
5	0.26069	0.40105	0.55430	-2.55805	-0.38493	-0.47791	1.37201	0.36124	-0.07813
6	-0.03438	0.66469	-0.19884	-1.00759	0.03066	-0.98787	-0.06568	-1.16595	-2.69024
7	-1.11166	0.86820	-3.55585	-0.30451	-0.19933	-0.16391	-0.26540	0.30183	0.53094
8	-1.54066	-0.58356	0.88770	-2.26393	1.14279	0.09226	0.05838	-0.35857	1.21975
9	-1.51112	1.08220	0.44583	1.65324	2.69754	0.07966	1.04016	1.00979	-0.15395
10	-1.61258	0.49211	0.67258	1.45476	-1.98721	-2.10957	0.71882	-1.15706	0.28375
11	0.56978	0.08236	0.30820	0.32600	-0.42655	0.86670	-0.40634	0.92654	-1.24194
12	1.13949	-0.05999	0.16390	0.02716	0.00573	-0.34176	0.60050	0.65593	1.65807
13	-0.39585	0.32207	0.00592	0.34073	-1.52043	3.02449	1.00098	-0.92541	0.56416
14	0.45455	0.14594	0.43705	0.12366	0.68861	0.64298	-1.55631	-0.51251	-0.71913
15	0.15434	-3.39740	-0.73453	0.65458	0.45710	0.07396	1.82629	-0.54708	-0.84140
16	0.22418	0.01294	-0.45727	0.04937	0.28166	0.25989	-0.39445	0.52048	-1.08718
17	0.31041	0.10097	-0.14308	-0.20635	-0.64531	-0.15862	-0.38535	0.61608	0.15172
18	-0.18711	0.31285	0.48414	0.22667	0.61084	0.84046	-1.15210	-2.26767	0.71925
19	-1.42678	-1.72055	0.64937	0.10350	-1.00800	-0.18019	-2.29512	1.60016	0.13771
20	-0.15281	0.47511	-0.01540	-0.07375	-0.02594	0.47018	0.23197	1.02360	-0.25186

4.3 Robust Match of Images

To look for the match points between feature point collections of two images is the key to get the solution of the image transformation parameters. There are many ways to search match points [76]. After comprehensive analysis on merits and short-comings of each method, this thesis offers a k-d tree method based on Harr Small wave coefficients index, which makes fast search for feature matching points possible.

4.3.1 Harr Small wave Coefficients Index Based k-NN Search Algorithm

Harr small wave [77] decompose vector X according to the following formula:

$$X_{s_{j,n}} = \frac{1}{\sqrt{2}}[X_{s_{j-1,2n}} + X_{s_{j-1,2n+1}}]$$
$$X_{d_{j,n}} = \frac{1}{\sqrt{2}}[X_{s_{j-1,2n}} - X_{s_{j-1,2n+1}}] \qquad (4\text{-}7)$$

Where $X_{s_{j,n}}$ and $X_{d_{j,n}}$ are respectively approximate coefficients and detail coefficients of the scale, when scale is j. When dimension of X meets the condition $2^{n-1} < N < 2^{n}$, after n times small wave decomposition, X can be completely decomposed. By then, the one and only approximate factor is C_x. The approximate coefficients and detail coefficients composes the vector small wave coefficients.

Since Harr small wave has the quality of orthogonality, so we could find k same neighbor points in the time domain as well as in the small wave domain. So make use of Harr small wave to decompose feature description vector with n dimensions of the extracted feature point completely. Then build feature description vector index between the approximate coefficient C_x and the last detail coefficient D_x. Finally, use 2-d (tow dimensional) index approximately to search for the matching points.

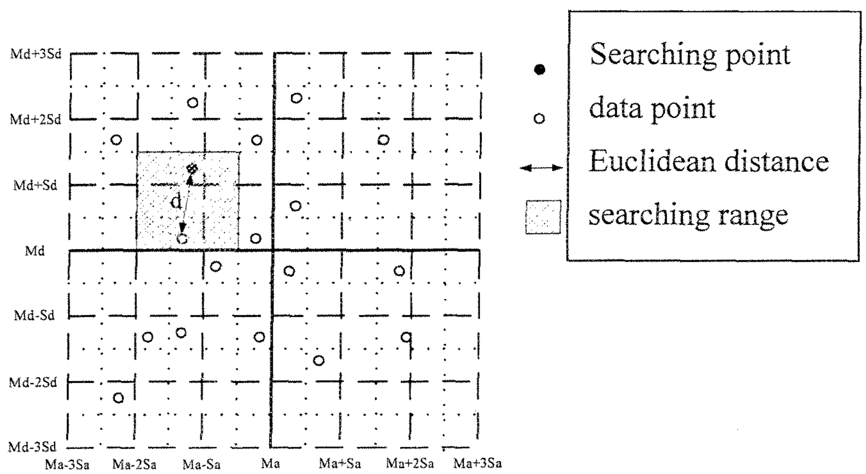


Figure 4-7 Harr Small wave Coefficients Index Based Feature Point Searching

Set the mean value of each feature point descriptor $x=\{x_0,x_1,\ldots,x_{N-1}\}$ as $M_x=\frac{1}{N}\sum_{i=0}^{N-1}x_i$, approximate coefficients $C_x=N\cdot M_x/\sqrt{2^n}$ [72] , $n=\log_2 N$, as for the feature descriptor $x=\{x_0,x_1,\ldots,x_{63}\}$, $C_x=\frac{1}{8}\sum_{i=0}^{63}x_i$, when $N=64$, we could get its detail coefficients $D_x=\frac{1}{8}(\sum_{i=0}^{31}x_i-\sum_{i=32}^{63}x_i)$.

After we get all feature point descriptors of each image—approximate coefficients average M_a, standard deviation S_a, detail coefficients average M_d, and standard deviation S_d, build each image a 2-d index of feature vectors, each dimension of the table contains 12 buckets, covering a range of standard deviation from dimension average -3 to +3. There are $12\times12=144$ buckets in total, each bucket covering a range of $\left[\frac{1}{2}S_a\times\frac{1}{2}S_d\right]$, as is shown in Figure 4-7.

In order to look for the corresponding point in the second image of feature point q in the first image, we could first search for the corresponding bucket in the feature point index of the second image by using the small wave approximate coefficient and detail coefficient of point q, then find the point with the shortest Euclidean distance to this bucket in an area of $3\times3=9$ buckets centered with this bucket, and the final result (the already found point) is the matching point.

4.3.2 Robust Transformation Parameters Estimation of Adjacent Images

For two images taken from the same viewpoint and meet the condition of rigid transform, feature matching point pair (x',y') and (x,y) meet transformation [78]

$$\begin{pmatrix}x'\\y'\\s'\end{pmatrix}=\begin{pmatrix}m_0&m_1&m_2\\m_3&m_4&m_5\\m_6&m_7&m_8\end{pmatrix}\begin{pmatrix}x\\y\\1\end{pmatrix}=M\begin{pmatrix}x\\y\\1\end{pmatrix}\tag{4-8}$$

And there

$$x' = \frac{m_0 x + m_1 y + m_2}{m_6 x + m_7 y + 1}, \quad y' = \frac{m_3 x + m_4 y + m_5}{m_6 + m_7 + 1} \tag{4-9}$$

Where m_i ($i = 0,1,..,8$) is the projection transformation parameters. Therefore, we can turn the registration problems of two adjacent images into a question of finding the solution for 9 transformation parameters m_i, that is, estimation of 9 parameters.

In the course of image mosaicing two kinds of error usually happen, random noise and outside point. Random noise usually caused by noise of the image itself and precision of feature extraction algorithm. The error it causes is limited within one pixel in general; while outside point is caused by false match in feature points matching process. This kind of error happens rarely, but seriously affects the accuracy of the final estimation. Thus, we have to take robust estimation method into consideration.

In statistics, there are many mature robust estimation methods[79], including Least Median Squares method and M estimation method. Least Median Squares, which is also named LmedS method, is to seek the minimum non-linear optimization solutions of a problem. It can figure out all the possible residual least median squares for all the data, in which way rejects outside points, and manages to estimate the optimal transformation parameters $\hat{a}$:

$$\hat{a} = \arg\min_a Median(r_{i,a}^2) \tag{4-10}$$

Where r is the residual. This method is an ideal way to deal with robustness of both noise and incorrect data.

Based on the calculation result in Section 4.3.1—collection of initial matching points of two adjacent images, as long as four feature point pairs are offered, any three of which are not in a line, there will be only one initial transformation matrix M. This thesis tries to get the solution of each parameter in transformation matrix M adopting LmedS method. Details are as follows:

(1) Randomly extract 4 initial matching point pairs, work out M according to figure (4-9), and calculate the median of the sum of error of all the matching point pairs.

$$E_{mcd} = \underset{i=1,2,\ldots,n}{med}[d^2(x_{i,2}, Mx_{i,1}) + d^2(x_{i,1}, M^T x_{i,2}] \qquad (4\text{-}11)$$

Where $x_{i,1}, x_{i,2}$ is the corresponding feature point; d is Euclidean distance;

(2) Repeat 1), take down M corresponding to minimum E_{med};

(3)In order to reduce the influence of noise, calculate variance of robust estimation:

$$\hat{\sigma} = 1.482 \times (1 + \frac{5}{\text{n-4}})\sqrt{E_{med}} \qquad (4\text{-}12)$$

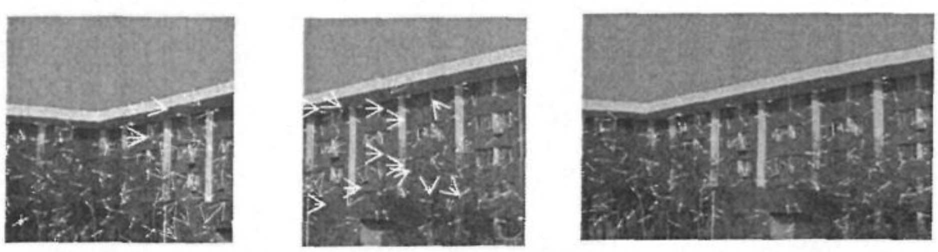
$$w_i = \begin{cases} 1 & r_i^2 \le (2.5\hat{\sigma})^2 \\ 0 & \text{else} \end{cases} \qquad (4\text{-}13)$$

Where $r_i = d^2(x_{i,2}, Mx_{i,1}) + d^2(x_{i.1}, M^T x_{i,2})$

(4) According to w_i in step (3), figure out value of the new M' adopting weighted least square method:

$$M' = \arg\min \sum_i w_i r_i^2 \qquad (4\text{-}14)$$

Figure 4-8 shows results of robust image mosaicing algorithm. We could see that all the outside points pointed out in Figure 4-8 (a) and (b) by the large amount of yellow arrows have been rejected in (c).



(a) source image (b) source image (c) result of stitching image

Figure 4-8 Results of Robust Algorithm, a Large Amount of outside Points Have Been Rejected

4.4 Image Stitching

In order to mosaic many images of the same scene into one wide viewing angle scene image, it is necessary to set a reference image, then figure out the mapping relation between other images and this reference image—this is the global optimization problem of scene image transformation. And in what order should the images be stitched in order to revive the real scene faithfully? As for the two problems above, on the basis of literature [80] study, this thesis puts forward the method to fix the image stitching order, and then work out the global mapping transformation of each image according to Leveberg-Marguardt (L-M) algorithm.

4.4.1 Processing of Image Stitching Order

Take any image f_i as reference image, $f_i \in G = \{f_j : j = 1,2,...,n\}$, looking for the best matching image f_j ($i \neq j$) in the image collection G, then put f_j into the global matching collection M; delete f_j from G, then go on searching for the best matching image of f_j; and so on, until G turns out to be null. By then, M has already figured out the image stitching order.

After the input image has gone through the processing in Section 4.3, every image could find its best matching image. Use search algorithm of stitching order mentioned above to fix the order of serial of images in order to further optimize the global transformation parameters.

4.4.2 Optimization of the Global Transformation Parameters of the Image

According to the rigid image space transformation principle, assuming M_{ij} stands for transformation matrix from best matching image f_i to f_i. If the transformation matrix from the given f_i to the reference image is M_j, then the corresponding matrix from f_i to the reference image is $M_i = M_{i,j}M_j$. Thus, we could figure out the global transformation parameter from every single image to the reference image.

In order to stitch image serial precisely by using global transformation matrix, it is indispensable to optimize global transformation parameters of the image, and therefore get the stable transformation relation between scene images.

Set point p' and q' respectively as corresponding point of point p and q in its own estimated image, then we define the geometric distance from the real matching point of the point in the image to its estimated matching point as:

$$d(p,q') = d(p,M^{-1}q) = \|p - M^{-1}q\|, d'(q,p') = d(q,Mp) = \|q - Mp\| \quad (4\text{-}15)$$

Where $\|.\|$ stands for Euclidean distance.

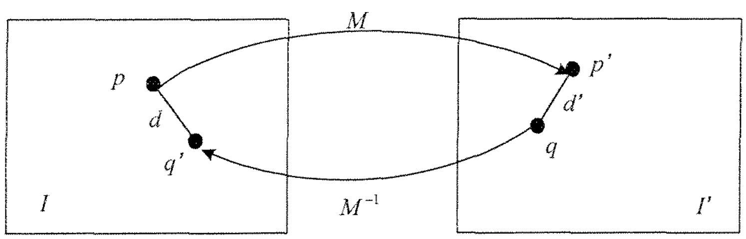


Figure 4-9 Euclidean Distance between the Matching Point and Estimated Point

Local optimization of the matching relationship of two images is to minimize the geometric distance between the feature matching points and the points which have gone through parameter calculation in the two images, that is:

$$dis = \min(\sum_{i=1}^{n}\left(d_i(p_i,q_i')^2 + d_i'(q_i,p_i')^2\right)) = \min(\sum_{i=1}^{n}\left(\|p_i - M^{-1}q_i\|^2 + \|q_i - Mp_i\|^2\right)) \quad (4\text{-}16)$$

Get the solution M which minimizes *dis*, and it is the transformation parameter of the tow images after optimization.

Take global optimization of the image serial into consideration. Set the optimizing parameter image as f, optimized image collection as G; images in G which are also related to f composes collection $F=\{f_1,...,f_k\}$. Set global transformation parameter of f as M_f, global transformation parameter of f_k as M_k. dis_k stands for the sum of geometric distance between searching matching points and parameter calculation matching points Therefore, the global optimization figure should be:

$$D=\sum_{m=1}^{k} dis_m = \sum_{m=1}^{k}\sum_{i=1}^{n}\left(| p_{mi} - M_m M_f^{-1} q_{mi}|^2 + | q_{mi} - M_m^{-1} M_f p_{mi}|^2\right) \quad (4\text{-}17)$$

When intending to get a stable global transformation parameter for f, that is, to get the solution when D reaches its minimum M_f, it turns into the least squares problem. Based on the processing of the errors of image feature points in Seciton 4.3.2, global optimization could adopt Leveberg-Marguardt (L-M) [81] non-linear algorithm to achieve fast solution. Its basic idea is:

Set there is functional relation among y, independent variable x and parameter $\alpha=(\alpha_1,\ldots,\alpha_m)^T$, $y=y(x;\alpha)$, in which $(x_i,y_i)(i=1,\ldots,n)$ is the observed value of (x,y) to n.

$$\chi^2(\alpha)=\sum_{i=1}^{n}\left[\frac{y_i - y(x_i;\alpha)}{\sigma_i}\right]^2 \quad (4\text{-}18)$$

Gets its minimum, where σ_i is the standard deviation of the observation error of the ith point $(x_i,y_i)(i=1,\ldots,n)$. This question equals to solving the nonlinear equation

$$\nabla\chi^2(\alpha)=0 \quad (4\text{-}19)$$

Here ∇ stands for gradient operator, Solve this equation 4.9 with Newton method is to:
Give α an initial value $\alpha^{(0)}$,

$$\begin{aligned}
&\alpha^{(k+1)} = \alpha^{(k)} + \Delta\alpha^{(k)} \\
&\nabla^2\chi^2(\alpha^{(k)})\Delta\alpha^{(k)} = -\nabla\chi^2(\alpha^{(k)}) \\
&k = 0,1,2,...
\end{aligned} \quad (4\text{-}20)$$

While the steepest descent method to solve the original problem is:

$$\Delta\alpha^{(k)} = \mu\left[-\nabla\chi^2(\alpha^{(k)})\right] \tag{4-21}$$

As for the fixed k take it as:

$$\begin{aligned}&[\alpha] = \frac{1}{2}\nabla^2\chi^2(\alpha^{(k)}), \beta = -\frac{1}{2}\nabla\chi^2(\alpha^{(k)})\\&\Delta\alpha = \Delta\alpha^{(k)}\end{aligned} \tag{4-22}$$

According to figure 4-21 and 4-22 we get

$$\begin{aligned}[\alpha]\Delta\alpha &= \beta\\ \Delta\alpha &= \mu\beta\end{aligned} \tag{4-23}$$

Via figure 4-23 get L-M method

$$([\alpha] + \lambda I)\Delta\alpha = \beta \tag{4-24}$$

Modify figure 4-24 into

$$([\alpha] + \lambda D)\Delta\alpha = \beta \tag{4-25}$$

In which $D = diag(\alpha_{11}, \alpha_{22}, \ldots, \alpha_{mm})$, by

$$\frac{\partial\chi^2}{\partial\alpha_j} = -2\sum_{i=1}^{n}\frac{[y_i - y(x_i;\alpha)]}{\alpha_i^2}\frac{\partial y(x_i;\alpha)}{\partial\alpha_j}, j = 1,\ldots,m$$

$$\frac{\partial^2\chi^2}{\partial\alpha_l\partial\alpha_j} = 2\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\left[\frac{\partial y(x_i;\alpha)}{\partial\alpha_l}\frac{\partial y(x_i;\alpha)}{\partial\alpha_j} - [y_i - y(x_i;\alpha)]\left[\frac{\partial^2 y(x_i;\alpha)}{\partial\alpha_l\partial\alpha_j}\right]\right]$$

So

$$\alpha_{lj} \approx \sum_{i=1}^{n}\frac{1}{\alpha_i^2}\left[\frac{\partial y(x_i;\alpha)}{\partial\alpha_l}\frac{\partial y(x_i;\alpha)}{\partial\alpha_j}\right], l = 1,\ldots,m \tag{4-26}$$

Solution for non-linear least squares problem is:

(1) Set the initial value of fixed parameter as α, calculate $\chi^2(\alpha)$;

(2) Pick a value for λ, e.g. $\lambda = 0.001$, make $k = 0$;

(3) Solve the linear equation (4-25) and get $\Delta\alpha$, calculate $\chi^2(\alpha + \Delta\alpha)$;

(4) If $\lambda = 0$, then end the iteration, otherwise turn to (5);

(5) If $\chi^2(\alpha + \Delta\alpha) \geq \chi^2(\alpha)$, then take $\alpha \Leftarrow \alpha + \Delta\alpha, \lambda = 10\lambda$, turn to (3);

If $\chi^2(\alpha + \Delta\alpha) < \chi^2(\alpha)$,

If $k \leq 2$, then take $\alpha \Leftarrow \alpha + \Delta\alpha, \lambda \Leftarrow \lambda / 10, k \Leftarrow k+1$, turn to (3);

If $k > 2$, then take $\lambda \Leftarrow 0, \alpha \Leftarrow \alpha + \Delta\alpha$, turn to (3).

To image transformation parameter optimization problem studied in this thesis, solution process of the algorithm is:

As for the transformation parameter

$$M = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix}$$

Set the coordinate of a pair of matching points in the source image as (u, v), the coordinate in the target image as (x, y); By transformation matrix M calculation, (u, v) transforms into (x_0, y_0), and the relation between them is :

$$x_0 = \frac{au + bv + c}{gu + hv + 1}, y_0 = \frac{du + ev + f}{gu + hv + 1} \qquad (4\text{-}27)$$

Make

$$D = \sum\left[\left(\frac{au + bv + c}{gu + hv + 1} - x\right)^2 + \left(\frac{du + ev + f}{gu + hv + 1} - y\right)^2\right]$$

the minimum. Figure out M's parameter $a \sim h$ when D is the minimum, that is, to figure out all the parameter when $D' = 0$.

Solve the equation according to $\chi^2(\alpha) = D$.

$$\nabla \chi^2(\alpha) = D' = \begin{bmatrix} \frac{\partial D}{\partial a} & \frac{\partial D}{\partial b} & \cdots & \frac{\partial D}{\partial h} \end{bmatrix}^T = [0] \tag{4-28}$$

In the calculation process

$$[\alpha] = \frac{1}{2}\nabla^2 \chi^2(\alpha^{(k)}) = Hessian(D) = \frac{1}{2}\begin{bmatrix} \frac{\partial^2 D}{\partial a^2} & \frac{\partial^2 D}{\partial a \partial b} & \cdots & \frac{\partial^2 D}{\partial a \partial h} \\ \frac{\partial^2 D}{\partial b \partial a} & \frac{\partial^2 D}{\partial b^2} & \cdots & \frac{\partial^2 D}{\partial b \partial h} \\ \vdots & \ddots & & \vdots \\ \frac{\partial^2 D}{\partial h \partial a} & \frac{\partial^2 D}{\partial h \partial b} & \cdots & \frac{\partial^2 D}{\partial h^2} \end{bmatrix} \tag{4-29}$$

$$\beta = -\frac{1}{2}\nabla \chi^2(\alpha^{(k)}) = -\frac{1}{2}\begin{bmatrix} \frac{\partial D}{\partial a} & \frac{\partial D}{\partial b} & \cdots & \frac{\partial D}{\partial h} \end{bmatrix}^T \tag{4-30}$$

By optimization of the global transformation parameter via L-M algorithm, we could make image stitching more natural and the stitching effect much closer to perfection.

4.5 Results of the Experiments & Analysis

Mosaicing two groups of images adopting the approach put forward in this thesis. Figure 4-10 (a) and (b) are outer door scene images of the library in University of Science and Technology taken with common digital camera. Size of both images is 256 ×256, with 279initial matching points. Transformation matrix is as shown is (4-31).

(a) original image

(b) original image

(c) Mosaicing result based on feature vector with 49 dimensions

(d) Mosaicing result based on feature vector with 6 dimensions

Figure 4-10 Mosaicing of the Outer Scene Image of the Library

$$H = \begin{pmatrix} 1.0052 & 0.2471 & 70.3823 \\ 0.5132 & 1.0162 & 2.9437 \\ 0.0001 & 0.0001 & 1 \end{pmatrix} \tag{4-31}$$

Observe the transformation matrix in figure 4-31, by H(1,3) and H(2,3) we estimate the general amout of movement should be(70.3823 ,2.9437); when H(1,1) and H(2,2) do not equal to 1, it indicates the scale changes of the images, but the changed scale could be omitted; Judging from the value of H(l,2) and H(2,1) we could tell that there is certain rotation angle of the images.

In this experiment, carry out feature matching search on all the 279 detected feature points in different ways—one is to use small wave index matching method directly, the other is to carry out dimension reducing processing first and then goes on to feature matching search. Table 4-3 shows experimental data relationship of feature points matching when using different dimensions. We could see that the accuracy rate of mosaicing using feature vectors with 49 dimensions is 2.1 percentage points higher than that of the dimension-reduced feature vector with 6 dimensions, but the searching time seems to be 8 times longer. This is because the less dimension feature description vector is with, the less storage space it needs, thus registration would become much faster. There is a balance between operating time and dimension of the Eigen value vector. According to Figure 4-10 image (c) and (d), we could see, change of dimension did not affect the accuracy of mosaicing. Therefore, adopting the approach brought forward in this thesis of mosaicing images with rigid transformation could reduce the storage cost as well as the operating time under the condition of guaranteeing the mosaicing quality.

Table 4-3 Comparison between the Results of Feature Processing with Different Dimensions

Dimension of feature vector n=49				Dimension of feature vector n=6			
Searching time	Quantity of matching points	Quantity of Correct matching	Correct matching proportion	Searching time	Quantity of matching points	Quantity of Correct matching	Correct matching proportion
9.96	279	195	69.8	1.2188	279	189	67.7

Figure 4-11 shows mosaic effect of six digital photos of a residential area using the automatic image processing order identification algorithm offered in this thesis and optimization of the parameters of the whole bundle.

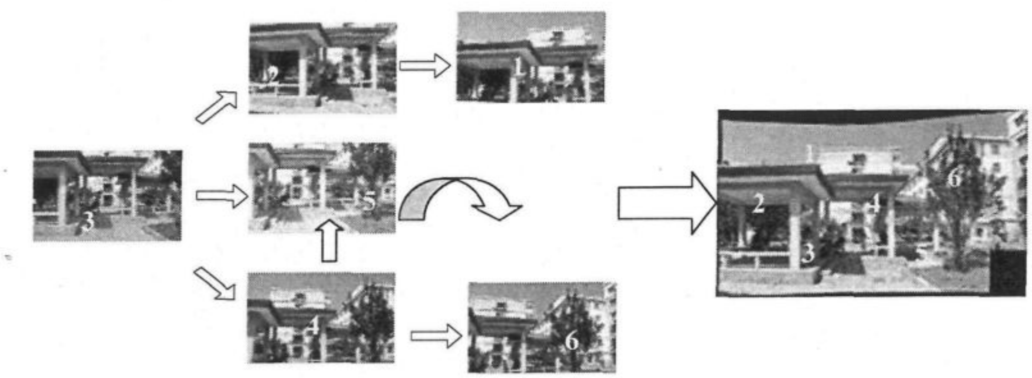
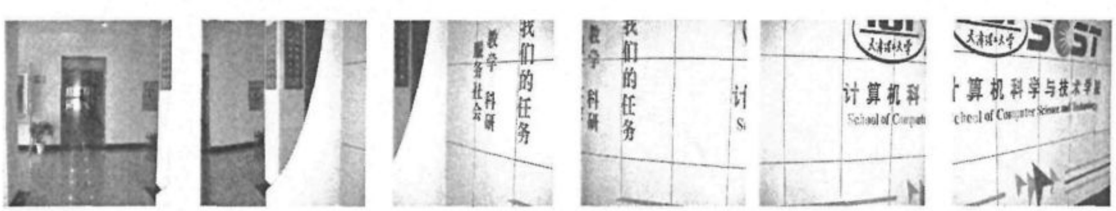


Figure 4-11 Mosaic Effect Image of a Corner of a Residential Area

We could see that the mosaic image reflect the real scene of a corner of the residential area.

Figure 4-12 shows the cylindrical panoramic mosaic effect of 12 images taken in the College Hall of Micro-computer.



(a) photos taken in the college hall

(b) panoramic image of the college hall

Figure 4-12 Cylindrical Panoramic Image of the College Hall

Chapter 5 Summary and Forecast

5.1 Summary

For the problems of image mosaicing, this thesis has done technical analysis from two perspectives, maximum mutual information's image mosaic technique in chapter 3, and characteristics-oriented image mosaic technique in chapter 4. The analyzing methods in each chapter are different. Based on previous two chapters, chapter 5 discusses the combination of image mosaic technique and TIP.

Based on the analysis of the existing image mosaicing approaches, this thesis studies deeply into some of the key issues in the mosaicing process that needs to be solved urgently, and has achieved phased and practical research results in the following areas:

(1) This thesis on the basis of small wave multi-resolution strategy, using particle swarm optimization, parameter's adaptive adjustment optimization along with the mutual information of the images and series of multi-resolution to solve the error matching problem. In gray-scale image registration process, because of the interpolation calculation, optimization process is easy to be led into local extremum, and thus cause error matching problem. Experiment proves that using this method could avoid the influence of the local extremum effectively. By finite order optimization iterative, it could find the optimal transformation quickly, therefore, enhances the computing speed of image registration, and achieves registration precision on sub-pixel level.

(2) To solve the problems of low quality of stitching and high computing cost caused by the obvious rotation, zooming, and light difference of the image, this thesis proposes robust fast image mosaicing algorithm. Firstly, make use of phase correlation method to identify overlay areas of the images; then, by using multi-resolution pyramid structure, extract Harris corners from the overlay areas in every layer of the image from top to

bottom from images of different scales; finally, to solve the problem of high computing cost of high-dimensional feature description vector, this thesis brings forward principle component analysis method for dimension reducing, and realizes accurate stitching of the images by using robust global parameter optimization. Experiment shows that, by using the method put forward in this thesis, we on the one hand ensured the image mosaicing quality, on the other hand reduced computing cost effectively.

(3) Another hard nut is to reconstruct virtual scene by using the scene image with unknown viewpoints. This thesis offers seamless connection strategy based on cobweb grid model and geometric principle of projection to construct multi-scene pseudo three dimensional models, and gives fast search algorithm in registration regional to achieve three dimensional mapping and registration of scene images. And on this basis, improve stitching precision of scene images by using small wave multi-resolution fusion in stitching regions of three dimensional models. Experiment shows that passage type virtual scene constructed on the basis of this method diminishes cavity phenomena in the process of roaming, and realizes the unlimited and smooth roaming of virtual scene.

5.2 Forecast

This thesis did some significative exploration in image mosaicing, and achieved certain results. But image mosaicing technology is still under consistent development. Considering the contents of this thesis, we think the following aspects would still be hot points and difficult problems of research:

Firstly, it remains to be an undone question about how to carry out feature matching precisely when mosaicing images with noise signal or even incomplete two-dimension information. For example, when mosaicing several intensity images, we need to registration feature points of different images. Concerning that lens of every image taking devices has interference factors, more or less, such as radial distortion and tangential distortion, and in the real world there are also complicated interactions between objects and light source, and all these factors greatly influence precision of feature points matching. At present, feature matching algorithm can get stable effects only when the images are offered in the camera's consistent dynamic shooting order, and still there have to be large overlay areas between the images. As for images taken from a relatively distant viewpoint, effects of existing algorithm are still not very satisfying. So it remains a chief task for researchers to do in a rather long period hence after to find how to achieve a stable and effective reconstruction result of virtual scenes with given images.

Secondly, there are many complicated data processing algorithms in image mosaicing process, such as decomposition of matrix, solve equations and nonlinear optimization problems. All these operations can only be operated in the CPU, so it takes much time to run, and fails the real-time requirement of scene construction. This thesis studies on how to reduce the computing cost, but still can not reach the anticipated execution speed under certain circumstances with real-time interactive requirement. Along with the appearance and development of programmable graphic hardware, data processing mentioned above, could be transformed to be executed into the coprocessor in the graphics acceleration card, which could lighten the operation burden of the CPU. And therefore, to improve execution speed of image mosaicing algorithm by using newly designed algorithm which could make full use of programmable graphic hardware's processing capacity is another research task for the coming researchers.

Finally, present image mosaicing relies on specific images, and different mosaicing algorithms can be only used in registration and fusion of certain types of images. Future studies should focus on the improvement of the precision, effectiveness and robustness of mosaicing algorithms, and invent a reasonable and accurate evaluation method of image mosaicing on this basis.

References

[1] M.Irani and S.Peleg, Improving Resolution by Image Registration. Graphical models and Image processing, 1991, 53(3):231~239.

[2] Szeliski R. Image mosaicing for tele-reality applications. IEEE Computer Society on Applications of Computer Vision(WACV'94), Sarasota Florida, 1994, 44~53.

[3] Burdea, G., Coiffet.P. , Virtual Reality Technology, John Wiley&Sons, 1994.

[4] Wu An Hua, Liu Xue Hui, Virtual reality and third dimension graph production, Chinese Journal of Image and Graphics, 1997, 2 （4）：205~212.

[5] S.E.Chen ,QuickTimeVR An image-based approach to virtual environment navigation. In: Proceedings of the ACM SIGGRAPH Conference on Computer Graphics, 1995: 29~38.

[6] R. Szeliski and H Y. Shum , Creating full view panoramic image mosaics and environment maps. Proc of ACM SIGGRAPH97 , Los Angeles , 1997:251~258.

[7] H..Y Shmu ,S.B.Kang ,A Review of Imgae-based Rendering Techniques ,IEEE/SPIP Visua Communictions and Image Proeessing(VCIP)2000 , Perth , Australia, 2000 , 2~13.

[8] Irani M , Anandan P, Hsu S , Mosaic based representations of video sequences and their applications. The Fifth International Conference on Computer Vision, Cambridge, Massachusetts, 1995, 605~611.

[9] Seitz, S. M., Dyer, C. R. ,View Morphing ,SIGGRAPH'96 Proceedings, 1996 ,21~30

[10] Shum , Heung Yeung , Kan , SingBing , Review of image-based render technology, Proceedings of SPIE-The International Society for Optical Engineering , 2000, 4067(I) : 2~13

[11] Shum ,Heung Yeung ,Kan , SingBing ,Chan ,Shing –Chow, Survey of image based representations and compression techniques. IEEE Transactions on Circuits and Systems for VideoTechnology , 2003 , 13(11):1020~1037

[12] Meehan J., Panoramic Photography, Watson-Guptill , 1990

[13] Werner, T., Hersch R.D., Hlavac,V., Rendering Real World Object Using View Interpolation, Proceeding of the International Conferences on Computer Vision, 1995.

[14] Hoffert,E., Mighdoll,L., Kreuger,M., at el, Quicktime:An Extensible Standard for Digital Multimedia, Proceedings of the IEEE Computer Conference, 1992.

[15] Narayanan,P.J., Virtualized Reality: Concept and Early Result, Proc IEEE Workshop on the Representation of Visual Scenes, IEEE, 1995.

[16] Szeliski R, Image mosaicing for tele-reality applieations. IEEE Computer society On Applications of Computer Vision(WACV' 94) , Sarasota, Florida, 1994, 44~53.

[17] Miller,G., Hoffert,E., Chen,S.E, etc. The Visual Museum: Interactive 3D navigation of a Multimedia Database. The Journal of Visualization and Computer Animation, 1992.

[18] Rossignac, J., Borrel,P., Multiresolution 3D Approximation for Rendering Complex Scene, Second Conference on Geometric Modeling in Computer Graphics, 1993.

[19] McMillan,L., Bishop,G., "Head Tracked Stereosopic Display Using Image Warping", Stereoscopic Display and Virtual Reality System II, Proc , SPIE, 1995.

[20] Lippman A, Movie maps: An application of the optical vodeodisc to computer graphics. Proceedings of the 7th Annual Conference on Computer Graphics and Interactive Technology, 1980: 32~42.

[21] Maciel,P.W.C.,Shierley,P., Visual Navigation of Large Environment using Textured clusters, Symposium on Interactive 3D Graphics, 1995.

[22] McMillan,L. ,Bishop,G., "Plenoptic Modeling: An Image based Rendering System" , *SIGGRAPH'95 Proceedings* , 1995

[23] Chen, S. E. and Willams, "View Interpolation for Image Synthesis" , SIGGRAPH'93,Proceeding,1993.

[24] R.Szeliski , Imgae Alignment and Stitching:A tutorial , Microsotf Reseacrh , 2004.

[25] Quddus A , FahmyM M. An improved small wave based corner detection technique. Proceedings of IEEE International Conference Acoustics , Speech and Signal. Processing, 1999:3213~3216.

[26] L. Z.Huang, Tour into the picture revisited. WSVG2001, 41~48.

[27]. Donald Hearn,M.Pauline Haker, Computer Graphics with OpenGL, Third Edition. Publishing House of Electronic Industry,2005

[28] Sun Jia Guang, Computer Graphics, Qing Hua Press, 1998

[29] Richard Hartley, Andrew Zisseman , Multiple View Geometry in Computer Vision . Cambridge University Press. 2003.

[30] L.G.Brown. A Survey of image registration techniques.ACM Computering Surveys, 1992, 24(4): 325~376

[31] W.K. Pratt. *Digital Image Processing* (2nd ed ed.),, Wiley, New York (1991).

[32] Polana R and Nelson R. Low level recognition of human motion. In: Proc IEEE Workshop on Motion of Non-Rigid and Articulated Objects, Austin, TX, 1994 , 77~82

[33] An FFT-based technique for translation, rotation and scale-invariant image registration. *IEEE Transactions on Image Processing* **5** (1996), pp. 1266–1271.

[34] F. Murtagh, A feature based O(N^2) approach to point pattern matching, Proceedings of the Internatinal Confernce on Pattern Recognition ICPR'92, Hague, The Netherlands, 1992, pp. 174–177.

[35] Harris C, Stephens M. A Combined Corner and Edge Detector[R]. In: Proceedings Fourth Alvey Vision Conference, Manchester, UK, 1988: 147-151

[36] B.S. Manjunath, C. Shekhar and R. Chellapa, A new approach to image feature detection with applications. *Pattern Recognition* **29** (1996), pp. 627–640I.

[37] J.P.W. Pluim, J.B.A. Maintz and M.A. Viergever, Mutual information matching in multiresolution contexts. *Image and Vision Computing* 19 (2001), pp. 45–52.

[38] A. Roche, G. Malandain and N. Ayache, Unifying maximum likelihood approaches in medical image registration. *International Journal of Imaging Systems and Technology* 11 (2000), pp. 71–80

[39] D. Lowe, "Distinctive image features from scale-invariant interest points," *International Journal of Computer Vision* **60** (2), 91–110 (2004).

[40] Sam Roweis. Levenberg-Marquardt Optimization[J].

[41] M. Brown and D. Lowe, “Recognising panorama,” *Proceedings of the 9th International Conference on Computer Vision*, pp. 1218–1225, Nice, France (October 2003)..

[42] Paul S. Heckbert. Fundamentals of Texture Mapping and Image Warping[R]. Technical Report: CSD-89-516. 1989.6

[43] Catmull E, Smith A R, 3D transformations of images in scanline order[J]. Computer Graphics,1980, 14(3):279-285

[44] F. Remondino, A. Roditakis. 3D Reconstruction of Human Skeleton from Single Images or Monocular Video Sequences, the 25th Pattern Recognition Symposium, Lecture Notes in Computer Science, Springer, Magdeburg, Germany, DAGM 2003, 100~107.

[45] Weinhaus F, Wallerman M. A ,lexible approach to image warping[J]. SPIE on Image Processing Algorithms and Techniques, 1990, Vo1.1244:108-122

[46] B. Triggs, P. McLauchlan, R. Hartley and A. Fitzgibbon. Bundle Adjustment: A Modern Synthesis[J]. In Vision Algorithms: Theory and Practice, number 1883 in LNCS, pages 298-373. Springer-Verlag, Corfu, Greece, September 1999..

[47] Maintz, J.B.A., Viergever, M.A. A Survey of Medical Image Registration. Med. Image Analysis. 1998, 2(1):1~36

[48] Egnal, “Mutual information as a stereo correspondence measure”, Technical report, Dept. of computer and information science, University of Pennsylvania, 2000.

[49] Colerhodes A A , Johnson K L , L Emoigne J , *et al* . Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient [J] *IEEE Trans. on Image Processing* , 2003 , 12 *(12)* : 1495-1511.

[50] VIOLA P , WELLS W M . Alignment by maximization of mutual information[J] . International Journal of Computer Vision , 1997 , 24(2) : 137-154 . .

[51] C.E.Shannon. A mathematical theory of communication. Bell Sys. Tech. J., 1948,27:379A. Pentland, Looking at people: sensing for ubiquitous and wearable computing. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22 (1): 107~119

[52] T Moeslund. E. Granum. A survey of computer vision-based human motion capture. Computer Vision and Image Understanding, 2001, 81(3): 231~268

[53] Buzug T.,Weese J. Improving DSA images with an automatic algorithm based on template matching and an entropy measure. Computer Assisted Radiology.1996,1124:145-150

[54]P.Thevenaz,M.Unser. Optimization of mutual information for multiresolution image registration.IEEE Trans.on Image Processing,2000,9(12):2083-2098

[55] H.Chen and P.K.Varshney. Registration of multimodel brain images:some experimental results.SPIE Conference Sensor Fusion:Architecture,Algorithms and Applications VI,2002,4371(4):122-133

[56]J.P.W.Pluim , J.B.A.Maintz , M.A.Viergever. Image registration by maximization of combined mutual information and gradient information. IEEE Trans.on Medical Image, 2000,19(8):809-814

[57] R.Li,B.Zeng,M.L.Liou,A New Three-Step Search Algorithm for Block Motion Estimation.Ieee Transaction on Circuits and Systems for Video Technology,Vol.4 August 1994:438-442.

[58] J.P.W.Pluim, J.B.A.Maintz, M.A.Viegever. Interpolation artifacts in mutual information based image registration. Computer Vision Image Understanding, 2000, 77(2): 211~232

[59] Yang Fu Sheng, Small wave transformation project analysis and application, Science Press, 1999

[60] Peng Xi Yuan, Peng Yu, Dai Men Feng, Group intelligence theory and application, Electronic journal, 2003,31 (12A).1982~1988.

[61] Zheng Yi, Wu Bin, Think by the bird group and the ant - based on the main body simulation and the aggregated intelligent research, Microcomputer world, 2001,1.7~13.

[62] Kennedy J.,Eberhart R C.,Particle swarm optimization.In: Proceedings of IEEE International Conference Onneural networks. Vol.IV. IEEE. 1995. 1942~1948

[63] Parsopulos K.E.,Vrahatis M.N..Recent approaches to global optimization problems through particle swarm optimization. Nature Computing.Kluwer Academic Publishers.2002, 235~306.

[64]Shi Y.,Eberhart R.C..A modified particle swarm optimizer. Proceedings of the International Joint Conference on Evolutionary Computation.1998.69~73.

[65] Lertrattanapanich S. Image registration for viedo mosaic. Master's thesis,The Pennsylvania State University,1999

[66]Wang Yuan Yuan, Application of adaptive simulated annealing and the search for the realization of multi-resolution image mosaic of Medicine, Precision Optical Engineering, 2006,6(14):1106-1106

[67] Lee J S,Sun Y N,Chen C H. Multiscale corner detection by using small wave transform.IEEE Transactions on Image Processing,1995,4(1):100-104

[68]Les Kitchen , Azrel Rosenfeld,Gray-level corner detection, Pattern Recognition Letters 1982,(1):95-102

[69]Wang H, Brady M, Real-time coner detection algorithm for motion estimation.Image and Vision Computing, 1995 , 13 (9) : 695-703

[70] C. Schmid, R. Mohr and C. Bauckhage, Evaluation of interest point detectors. *Internat. J. Comput. Vis.* **37** 2 (2000), pp. 151–172

[71] Harris C, Stephens M. A Combined Corner and Edge Detector[R]. In: Proceedings Fourth Alvey Vision Conference, Manchester, UK, 1988: 147-151

[72] Y. Dufournaud, C. Schmid and R. Horaud, Matching images with different resolutions. In: *IEEE Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society Press, Hilton Head Island, SC (2000), pp. 612–618

[73] Linderberg, T. Scale-space theory: A basic tool for analyzing structures at different scales[J]. Journal of Applied Statistics. 1994. 21(2): 224-270.

[74] M. Brown, R. Szeliski and S. Winder. Multi-Image Matching using Multi-Scale Oriented Patches[R]. Technical Report MSR-TR-2004-133. 2004

[75] Fang Kai Tai, Multi-dimensional statistical analysis, Science Press, 2002

[76] Caixc,Hu Yh,Sto Xh, et al. A mixture feature selection method for remote sensing image[A]. Proceeding of SPIE International Conference on Space Information Technology, 2006

[77] Wu Jun, Based on small wave transformation image registration method research, Hunan University journal, 2005

[78] Zhang Zhengyou. Determining the epipolar geometry and its uncertainty: a review[R].France:INRIA SophiaAntipolis,1996

[79] Meer P.Mintz D. Kim D. et al. Robust regression methods in computer vision:a review[J] International Journal of Computer Vision.1991,6(1):59-70

[80]. Triggs,B . et al . Bundle adjustment-a modern synthesis . International Workshop on Vision Algorithms . 1999:298-372, Springer,Kerkyra,Greece.

[81] Sam Roweis. Levenberg-Marquardt Optimization[J].