Université de Montréal

3D reconstruction of a catheter path
from a single view X-ray sequence

par

Ji Yao Weng

Département d'informatique et de recherche opérationnelle

Faculté des arts et des sciences

Mémoire présenté à la faculté des études supérieures

En vue de l'obtention du grade de

Maître ès sciences (M.Sc.)

en informatique

août 2003

Université de Montréal

Université de Montréal

Faculté des études supérieures

Ce mémoire de maîtrise intitulé

3D reconstruction of a catheter path
from a single view X-ray sequence

Présenté par

Ji Yao Weng

A été évalué par un jury composé des personnes suivantes :

Dr. Sébastien Roy

Président-rapporteur

Dr. Jean Meunier

Directeur de recherche

Dr. Pierre Poulin

Membre du jury

Mémoire accepté le 10 septembre 2003

Sommaire

Plus de 1/3 des décès enregistrés au Canada au cours des dernières années sont causés par les maladies cardiovasculaires. L'un des problèmes cardiovasculaires le plus communs est l'athérosclérose coronaire qui consiste en l'accumulation de plaques sur les parois des artères. La recherche sur les niveaux de risque de l'athérosclérose coronaire est, par conséquent, de très grande importance pour le diagnostic et la stratégie thérapeutique à entreprendre ultérieurement. Cliniquement, l'imagerie IVUS (ultrasons intravasculaires) combinée avec l'angiographie est largement utilisée pour l'examen médical et le traitement des maladies cardiovasculaires.

Dans l'imagerie IVUS, les méthodes « pose estimation » sont employées pour déterminer la trajectoire du cathéter à partir d'une projection unique obtenue par angiographie. Des progrès remarquables ont été réalisés par les travaux de recherche sur la « pose estimation », à partir d'une vue unique en image, et par l'implémentation de méthodes de reconstruction 3D de la trajectoire du cathéter dans une imagerie IVUS. En dépit de ces améliorations significatives, l'exigence d'une connaissance antérieure de la configuration 3D des artères coronaires est un inconvénient majeur qui se pose lors de la construction 3D de la trajectoire du cathéter dans une imagerie IVUS. Malheureusement, il y a eu un manque d'exploration de nouvelles méthodes susceptibles de pallier cet inconvénient.

Cette thèse se focalise sur la « pose estimation » à partir d'une projection unique et ce, pour une reconstruction 3D de la trajectoire du cathéter dans une investigation IVUS. Premièrement, nous explorons l'état de l'art de la « pose estimation » projection unique en passant en revue et en simulant trois méthodes typiques choisies parmi d'autres. Les inconvénients des méthodes existantes nous ont motivés à étudier la possibilité d'une « pose estimation » à partir d'une séquence d'images. Nous proposons une nouvelle méthode de pose estimation à partir d'une séquence d'images de vue unique

visant la reconstruction du la trajectoire de cathéter dans imagerie IVUS. Ensuite, nous simulons avec un logiciel mathématique la méthode proposée par des courbes spirales et nous l'appliquons par la suite à une séquence d'images obtenues à partir d'une expérimentation fantôme. Les résultats obtenus montrent que les erreurs de reconstruction varient entre 5.81% et 6.68% pour les simulations et 1.25% à 1.36% (en termes de taille de fantôme reconstruit) pour l'étude fantôme.

Bien que ces chiffres pourraient sembler plutôt médiocres, ils sont réellement une mine d'or pour les médecins qui n'ont pas accès aux informations 3D (à moins d'utiliser un laboratoire complexe proposé par une autre méthodologie, ce qui est impraticable dans les cliniques). La méthode proposée a l'avantage, d'une part, d'une installation beaucoup plus facile dans les cliniques et, d'autre part, d'un coût inférieur à celui des méthodes existantes d'autant plus qu'elle conduira certainement, à l'avenir, à un outil de valeur.

Mots clés : cathéter, pose estimation, vue unique, séquence, IVUS

Abstract

Cardiovascular disease causes over 1/3 of all deaths in Canada in recent years. One of the most common cardiovascular problems is coronary atherosclerosis, the build up of plaque on artery walls. The investigation of the severity of coronary atherosclerosis is therefore very important for the diagnosis and therapeutic strategy that will be undertaken. Clinically, IVUS (Intravascular Ultrasound) imaging combined with angiography is widely used in examination and treatment of cardiovascular diseases.

In IVUS imaging, methods of pose estimation are used to determine the trajectory of the catheter from single projection images obtained by angiography. The research on pose estimation from single view images or one single view image has achieved remarkable progress, and methods have been implemented in the 3D reconstruction of catheter trajectory in IVUS imaging. Despite these significant improvements, the requirement of previous knowledge of the 3D configuration of the coronary arteries is a significant drawback that still exists in the implementation of 3D construction of a catheter trajectory in IVUS imaging. Unfortunately, there has been a lack of exploration of new methods that could overcome this drawback.

This thesis focuses on pose estimation from single projection aiming to the 3D reconstruction of the catheter trajectory in an IVUS investigation. Firstly, we investigate the state of art of pose estimation by single projection by reviewing and simulating three typical methods, which are chosen from a variety of published papers related to the topic. The drawbacks of existing methods motivate us to investigate the possibility of pose estimation from an image sequence. We propose a novel method of pose estimation from a single view image sequence targeted for the reconstruction of the catheter trajectory in IVUS imaging. Afterward, we simulate the proposed method with spiral curves in the Mathematica® environment, and apply the proposed method to an image sequence obtained from a phantom experiment. Typical results show reconstruction errors ranging from 5.81% to 6.68% for the simulations and 1.25% to 1.36% (in terms of the size of the

reconstructed phantom) for the phantom study. Although these figures could seem rather mediocre, they are actually a gold mine for the physician that has access to no 3D information at all in clinic today (unless he uses quite complex "laboratory" set-up proposed by other methodologies, which is unfeasible in clinical practice). The proposed method has advantages of much easier set-up in the clinical environment and lower cost than existing methods, thus it will certainly lead to a valuable tool in the future.

Key works: catheter, pose estimation, single projection, sequence, IVUS

Table of Contents

List of Tables

List of Figures

Acknowledgements

Montreal, QC   Ji Yao Weng

Chapter 1

Introduction

1.1 Introduction

Among all pathologies affecting the modern world, cardiovascular diseases are at the forefront. In Canada, they account for the death of more Canadians than any other disease. In 1999 according to Statistics Canada, cardiovascular diseases killed 78,942 Canadians, 36% of all deaths in Canada. They cost Canadian economy over $18 billion a year according to a 1994 study by the Heart and Stroke Foundation. One of the most common cardiovascular problems is coronary atherosclerosis, the build up of plaque (a combination of cholesterol, cellular waste, and other materials) on artery walls. Plaque can cause a heart attack by severely reducing or stopping the blood flow through a coronary artery. Moreover, the plaque can rupture and form blood clots capable of blocking arteries. The investigation of the severity of coronary atherosclerosis is therefore very important for the diagnosis and therapeutic strategy, such as medication, bypass surgery, angioplasty (dilation) with or without stent, which will be undertaken. For this purpose, two main imaging methods are used nowadays. Angiography (X-rays) consists of the injection with a catheter of a contrast product in the lumen of the arteries, making them opaque to X-rays. Using several views (projections) the physician can get an assessment of the position and geometrical severity of the stenosis (narrowing of the artery due to atherosclerosis). It is nevertheless important to remember that a single angiogram reveals only a 2-D "silhouette" of the true 3-D lumen. This has for consequence that an image under only one (or even a few) angle of sight can badly represent the extent of a complex stenosis. Although 3-D reconstruction algorithms exist in angiography based on two or more views, it remains that angiography visualizes only the inner lumen of the vessel and cannot determine directly if the vessel wall has

atherosclerosis; for instance it could miss diffuse (long) lesions with seemingly no stenosis. Intravascular ultrasound (IVUS) represents a complement (even an alternative) to angiography for the direct visualization of the arterial anatomy. A miniaturized ultrasonic transducer at the end of a catheter is inserted in the artery lumen and brought beyond the lesion of interest and then withdrawn gradually, manually or automatically. Contrary to angiography, which represents a silhouette of the arterial lumen, IVUS produces unique echographic images (Figure 1.1) showing the cross-section of coronary arteries. These images reveal clearly the lumen, walls and plaque, and offer a powerful tool for diagnostic purposes. Unfortunately, this sequence of images does not offer direct information about the 3-D geometry of the artery, an essential feature to reliably compute plaque volume or sheer stress.
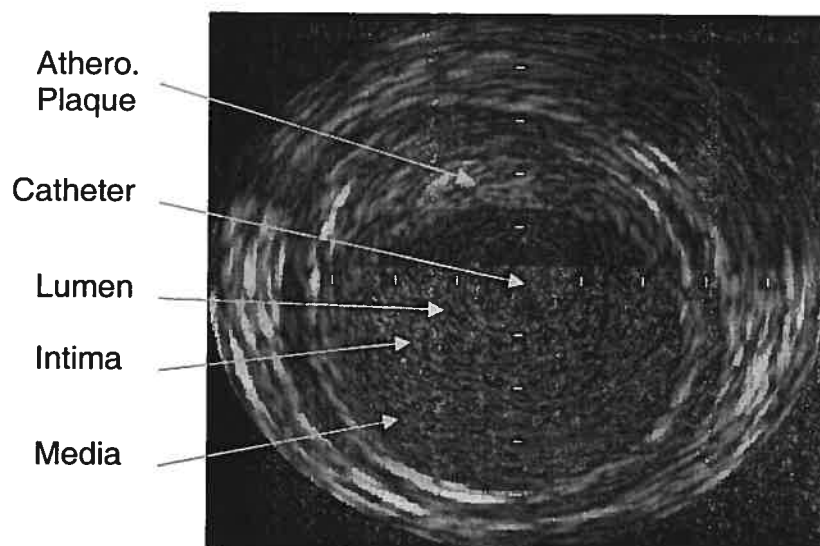
Figure 1.1   A single frame of IVUS image.

Because there is no information about the 3-D location and orientation associated with each frame of an IVUS image sequence while coronary angiography can be used to localize and even to reconstruct of coronary arteries in 3-D, IVUS and angiography together are widely utilized in assessment of coronary artery diseases in clinics (Figure 1.2).
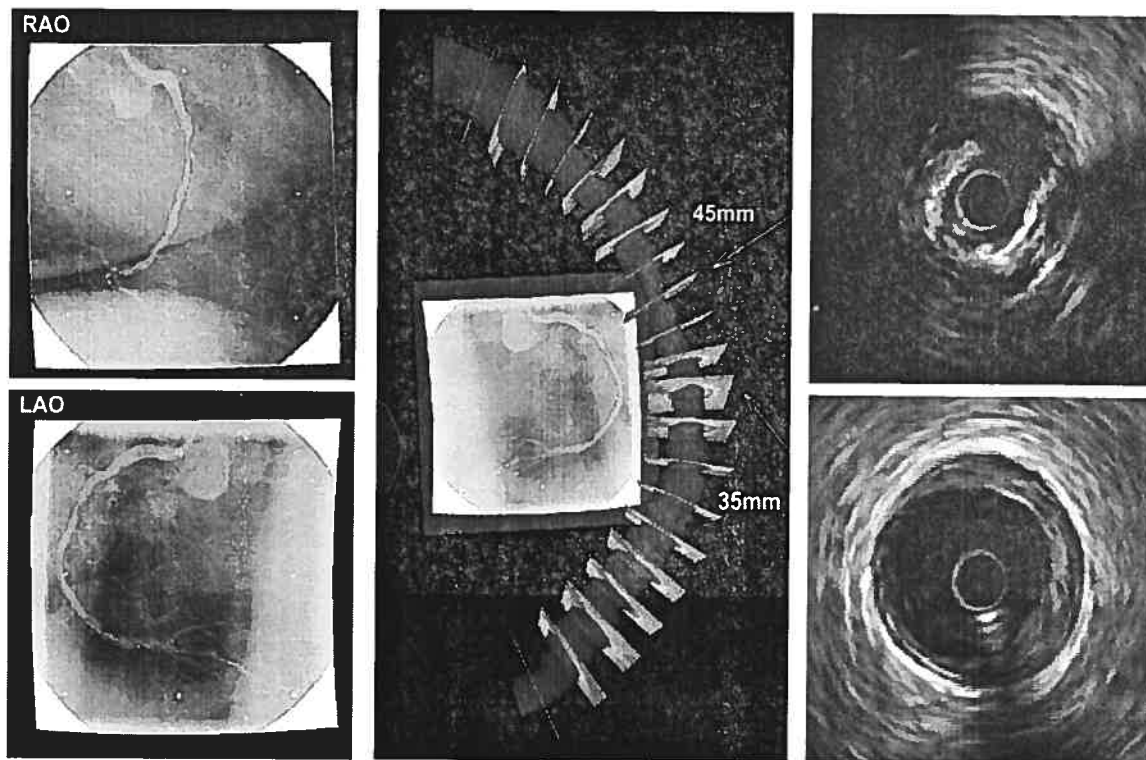
Figure 1.2   Angiograms (left) and two IVUS images taken from the sequence (right) with the corresponding 3-D reconstruction (middle). (ref: Whale et al. http://www.engineering.uiowa.edu/ ~awahle/WahIUS/Slides/virtual-angioscopy.pdf).

In a typical IVUS intervention, a few single projection angiographic sequences are taken to localize the IVUS transducer during its pullback. In addition, the pullback length can be measured. This implies that it could be possible to obtain the reconstruction of coronary arteries. The task of this research is to develop a novel pose estimation method in which an image sequence from one single projection is used to obtain the reconstruction of coronary arteries, instead of using a two-camera system such as in biplane angiography or additional CT or MR information in single projection angiography. As a result, the costs of clinical equipment and clinical processing could be reduced for reliable 3-D artery measurements.

1.2 Pose Estimation in Intravascular Ultrasound Modality

Pose estimation is a method for determining the location and orientation of a 3-D object with respect to a camera system from 3-D to 2-D point correspondences. Two kinds of pose estimation methods are used in angiography: biplane angiography [1, 2], in which images from two views are needed, and single projection angiography [3, 4, 5, 6, 10, 11, 14], in which only one view image and prior knowledge of 3-D configuration of the object are required. Research on biplane pose estimation has achieved significant progress. Unfortunately, the required biplane angiography system is more expensive, usually takes more physical space, and is much more complex to operate than single projection angiography. This partly explains why these systems are not available everywhere and even tend to disappear from clinical practice.

Pose estimation from single view needs a previous knowledge of the 3-D configuration of the object. In single projection angiography implementation, this knowledge of 3-D configuration of coronary arteries can be obtained by either computed tomography (CT) or magnetic resonance (MR). This requirement of an extra clinical processing is a drawback of single projection angiography. Pose estimation from single view has more challenges than biplane pose estimation due to its specific difficulties.

In this thesis we focus on pose estimation with a single angiographic projection with knowledge limited to the pullback distance traveled by the ultrasonic transducer in IVUS intervention.

1.3 Motivation

So far, researchers have focused on implementation of model-based pose estimation in single plane angiography ignoring the information that could be obtained in the specific situation of an IVUS intervention. In a practical clinical situation, an image sequence can be taken from single projection during pullback of the catheter during the IVUS intervention, and the pullback length of catheter in the interval between two successive images can be measured. Based on this specific situation in IVUS intervention, a novel method is proposed in this thesis to reconstruct the trajectory of the catheter tip. This trajectory represents the 3-D pose of the catheter that is estimated from

the 2-D projection on the image sequence.

The proposed pose estimation method overcomes the main drawbacks of biplane pose estimation methods and model-based single project pose estimation methods by reducing the clinical costs. Compared with biplane pose estimation and model-based pose estimation, the estimation of pose from a single plane angiography image sequence and the measurement to pullback length is rather difficult. A significant challenge is that, mathematically, there exist multiple pose solutions corresponding to a given image sequence with known pullback measurements, but there is no geometrical condition that can be used to determine the right solution from the multiple solutions. To solve this problem, we consider the physical characteristics of a catheter and assume that there is no sharp bend of the catheter.

In terms of accuracy, efficiency, and stability, the proposed method is not as good as the existing pose estimation methods. However, considering its much easier setup in the clinical environment the proposed method will certainly lead to a valuable tool in the future.

1.4 Organization of the Thesis

This thesis presents a study on pose estimation by single projection in the context of IVUS intervention. The thesis is organized as follows.

In Chapter 2, we begin with the definition of pose estimation from single projection. Then we review and simulate three typical methods of pose estimation chosen from various existing methods. These three typical methods are Newton's method, a classic method, SPT, a method developed intendedly for angiography implementation, and POSIT, a recently published method. The simulation results on several cases, and a comparison in terms of accuracy, efficiency, and stability are presented. In Chapter 3, we analyze the problem of pose estimation with a single plane image sequence and a measurement of the catheter pullback length, and the difficulties associated with this problem. To solve the problem, we propose a new method, which is presented in detail. Furthermore, the proposed method is simulated using several given cases and the simulation results are shown. In Chapter 4, the proposed method is applied to phantom (physical model) image sequences provided by the Montreal Heart. Chapter 5 concludes

the thesis by highlighting the findings of this investigation and by suggesting some possible future work.

Chapter 2

Existing Pose Estimation Methods

2.1 Introduction

Conventional 3-D IVUS uses an automated pullback device to get a stack of IVUS images providing additional information for volumetric measurements. Unfortunately this method assumes a straight vessel, which is a crude approximation for coronary arteries. In fact, due to the vessel curvature, the image planes are not parallel. Moreover the catheter twists when following a tortuous vessel generating a rotation artifact in the image plane. These problems must be considered to compute a true 3-D reconstruction of the vessel from IVUS images. A few groups [1, 2] have successfully implemented solutions to this problem using biplane angiography to infer the 3-D trajectory of the catheter. At the University of Iowa [1], they extract the 2-D catheter path in both biplane angiograms and then reconstruct the 3-D trajectory knowing the biplane imaging system geometry. Assuming a known constant pullback speed, the actual location of each IVUS frame along (and perpendicular to) the 3-D pullback path is obtained. To compute the catheter twist, they determine the relative rotation with a sequential triangulation method and the absolute orientation from the out-of-center position of the IVUS catheter used as a landmark in both angiograms and IVUS images. Another group [2] from the Cleveland Clinic Foundation has proposed a method that could be used with or without a pullback device. The 3-D trajectory of the IVUS transducer was computed as a function of time using biplane angiography similarly to the University of Iowa's group. Each IVUS frame was time-synchronized with the angiographic images and then correctly positioned perpendicular to the trajectory. Then, a 3-D segmentation method based on a 3-D extension of active contours (snakes) extracts the lumen surface. Finally the IVUS lumen is backprojected on the angiograms to find

the best twist. Unfortunately, the required biplane angiography system is more expensive, usually takes more physical space, and is much more complex to operate than its single plane parent. This partly explains why these systems are not available everywhere and even tend to disappear from clinical practice. In addition, for IVUS artery reconstruction, a calibration step (to assess the biplane imaging system geometry) is necessary and the X-ray radiation dose to the patient could typically be higher. All this would add a significant burden to the already quite complex IVUS intervention protocol. We believe that although the biplane approach gives spectacular results it will probably find limited use in clinical practice except in university hospital research laboratories. As for the use of 3-D position sensors (*e.g.*, six-degree-of-freedom magnetic sensor) for tracking the trajectory of the transducer, this is indeed an efficient method for 3-D conventional ultrasound transducer but this is certainly a big technological challenge for IVUS and is probably not reachable in the near future.

Because of the drawbacks existing in biplane angiography for IVUS artery reconstruction, there has been a rise of implementations of single projection angiography for IVUS artery reconstruction [4, 5, 11]. It overcomes the drawbacks existing in biplane angiography.

In this thesis, we focus on pose estimation from single projection methods, which are employed in single projection angiography. The rest of this chapter presents a review and simulations of three typical pose estimation methods that are chosen from many publications in the computer vision community and biomedical engineering community. These methods are Newton's method [3], SPT (Single Projection Technique) by Hoffman et al. [4, 5, 11], and POSIT (Pose from Orthography and Scaling with ITeration) by Demonthon et al. [6, 14]. In the next chapter, we will present our research on pose estimation by image sequence from single projection.

Newton's method is named after Newton's numerical method that is employed to solve systems of nonlinear equations, which are obtained from the relation of object points and their corresponding full perspective projections. Newton's method is simple and straightforward. The main drawback of Newton's method is that it has a limited suitable scope of location and orientation for the object.

SPT is a method of pose estimation by single projection that was developed for

angiography modality. The main idea of SPT is to align the object points with their corresponding projections by adjusting translation and rotation using projection-Procrustes technique [7, 8, 9]. The method has two main steps. The first step is to align the object points with their corresponding projections by adjusting the translation in the x-y plane iteratively. In the second step, projection-Procrustes technique is carried out to adjust the translation and rotation to optimize the alignment of the object points with their corresponding projections. SPT is tested in a phantom experiment by Hoffman et al., and the results are given in [5]. However, the phantom experiment is based on several dots that are fixed in a known configuration instead of a catheter. SPT is relatively complex. In terms of stability, SPT is better than Newton's method, but still has the drawback of a limited suitable scope of location and orientation for the object.

POSIT emerged recently. It consists of iteratively improving the pose computed with weak perspective projection (*i.e.*, scaled orthographic projection) camera model to converge to a pose estimation computed with a perspective projection camera model. Furthermore, Horaud et al. [14] derive the idea for POSIT and develop an improved version of POSIT by using paraperspective projection instead of weak perspective projection. According to our simulation in Mathematica, POSIT is the best method for pose estimation by single projection in terms of efficiency, accuracy, and stability among the three methods.

This chapter is organized as follows: Section 2 gives a geometrical definition for the problem of pose estimation from single projection. Section 3 reviews the Newton's method, the SPT method, and the POSIT method. A modification that we have made to reduce the slow convergence problem existing in SPT is also described in Section 3. In Section 4, a comparison among the three methods based on the simulation results is presented. Section 5 presents a summary for this chapter.
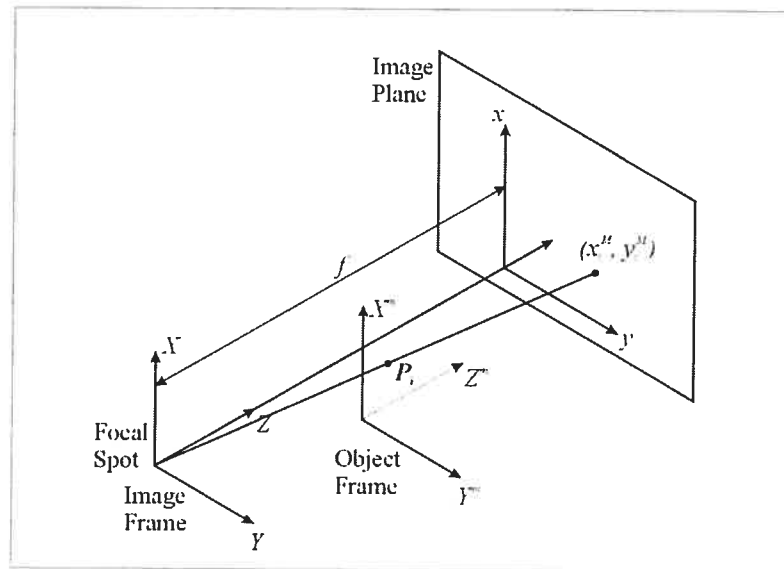
Figure 2.1 Geometry of pose estimation by single projection problem. The model points in the camera reference frame and the model reference frame are denoted by (X, Y, Z) and (X^m, Y^m, Z^m), respectively. The image of a model point on the image plane is denoted as (x^M, y^M).

2.2 The Problem of Pose Estimation by Single Projection

Pose estimation by single projection is also called model-based pose estimation, or model matching, or optical jigging. The geometry of pose estimation by single projection is depicted in Figure 2.1. The 3-D camera coordinate system, (X, Y, Z), is defined such that its origin is located at the focal spot. The image plane is located at a distance of focal length, f, from the focal spot, and the image coordinate system, (x, y), lying in the image plane, is parallel to X and Y. The superscript M in (x^M, y^M) is used to denote measured image coordinates. Independent of the camera reference frame, the 3-D model coordinate system, (X^m, Y^m, Z^m), is defined relative to the points in the object itself. The pose estimation by single projection problem can be stated as follows:

- Let P_1^m, ..., P_n^m, with $P_i^m = [X_i^m, Y_i^m, Z_i^m]^T$ and $n \geq 3$, expressed in the model reference frame, be n points of an object model.

- Let P_1, ..., P_n, with $P_i = [X_i, Y_i, Z_i]^T$, expressed in the camera reference frame, indicate the coordinates of the corresponding points on the object observed.

- Let p_1, ..., p_n, with $p_i = [x_i, y_i]^T$, be the n image points, expressed in the image reference frame, projections of the P_i.

The goal is to determine the rigid transformation, *i.e.*, rotation matrix R and translation vector T, aligning the camera reference frame and model reference frame:

$$P_i = R \cdot P_i^m + T.$$

The above equation can be written as

$$[X_i, Y_i, Z_i]^T = R \cdot [X_i^m, Y_i^m, Z_i^m]^T + T, \tag{2.1}$$

where T can be represented by T_1, T_2 and T_3, the translations along the X, Y, and Z axes, as follows:

$$T = [T_1, T_2, T_3]^T, \tag{2.2}$$

and R can be represented by Euler angles, ϕ_x, ϕ_y and ϕ_z, the rotation angles about the X, Y and Z axes of the camera reference frame as in the following equation:

$$R^T = \begin{bmatrix} \cos\phi_z & -\sin\phi_z & 0 \\ \sin\phi_z & \cos\phi_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\phi_y & 0 & \sin\phi_y \\ 0 & 1 & 0 \\ -\sin\phi_y & 0 & \cos\phi_y \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi_x & -\sin\phi_x \\ 0 & \sin\phi_x & \cos\phi_x \end{bmatrix}. \tag{2.3}$$

Inversely, Euler angles, ϕ_x, ϕ_y and ϕ_z, can also be computed from R by Euler angle decomposition:

$$\text{if } r_{31} \neq \pm 1 : \begin{cases} \phi_x = ArcSin\left(\dfrac{r_{31}}{\sqrt{1-(r_{31})^2}}\right) \\ \phi_y = ArcSin(r_{31}) \\ \phi_z = ArcSin\left(\dfrac{r_{21}}{\sqrt{1-(r_{31})^2}}\right) \end{cases}$$

$$\text{if } (r_{31}) = \pm 1 \begin{cases} \phi_x = 0 \\ \phi_y = ArcSin(-r_{31}) \\ \phi_z = ArcSin(-r_{23}), \end{cases} \tag{2.4}$$

with R being written as:

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}. \tag{2.5}$$

In the equations (2.4), we assume that $0 \leq \phi_x < 180°$, $0 \leq \phi_y < 90°$ and $0 \leq \phi_z < 180°$.

2.3 Review of Three Methods of Pose Estimation by Single Projection

In this section, we review three typical methods of pose estimation by single projection [3, 4, 5, 6, 10, 11, 14], which are chosen from many publications in the computer vision community and the biomedical engineering community. The methods are Newton's method, a classic method, SPT method, a method intended for implementation in angiography, and POSIT, a method that emerged recently.

2.3.1 Newton's Method

The algorithm employs Newton's iterative method to solve the system of nonlinear equations, which are obtained from the full perspective projection [3].

The relation between object points and their corresponding image points in camera coordinates is given by the full perspective projection equation:

$$[x_i, y_i]^T = \left[\frac{fX_i}{Z_i}, \frac{fY_i}{Z_i}\right]^T. \tag{2.6}$$

Plugging equations (2.1), (2.2), and (2.5) into equation (2.6), we can see that each image point correspondence generates two non-linear equations,

$$x_i = f\frac{r_{11}X_i^m + r_{12}Y_i^m + r_{13}Z_i^m + T_1}{r_{31}X_i^m + r_{32}Y_i^m + r_{33}Z_i^m + T_3}$$
$$y_i = f\frac{r_{21}X_i^m + r_{22}Y_i^m + r_{23}Z_i^m + T_2}{r_{31}X_i^m + r_{32}Y_i^m + r_{33}Z_i^m + T_3}. \tag{2.7}$$

The unknown components of R and T can be determined from a sufficient number of correspondences, each bringing two equations like equation (2.7). The resulting systems have six unknowns, ϕ_x, ϕ_y, ϕ_z, T_1, T_2, and T_3. Here R depends only on three free parameters ϕ_x, ϕ_y, and ϕ_z.

Newton's numerical method is now employed to solve the systems. Assuming that $(\overline{R}, \overline{T})$ is the true solution for the system, the method starts off with an initial guess for $\overline{R}$ and $\overline{T}$, say R^0 and T^0, and computes p_i through equation (2.7) with $R = R^k$ and $T = T^k$, $(k=0, 1, \ldots)$, until the residuals

$$\delta x_i = x_i(R^k, T^k) - \overline{x}_i$$
$$\delta y_i = y_i(R^k, T^k) - \overline{y}_i \tag{2.8}$$

are small enough. In equation (2.8), $\bar{x}_i$ and $\bar{y}_i$ are x and y coordinates of measured image points. The first-order expansions of residuals are

$$\sum_{j=1,2,3} \frac{\partial x_i}{\partial T_j} \Delta T + \sum_{k=x,y,z} \frac{\partial x_i}{\partial \phi_k} \Delta \phi_k = \delta x_i$$

$$\sum_{j=1,2,3} \frac{\partial y_i}{\partial T_j} \Delta T + \sum_{k=x,y,z} \frac{\partial y_i}{\partial \phi_k} \Delta \phi_k = \delta y_i, \tag{2.9}$$

where the partial derivatives with respect to T_1, T_2, and T_3 are

$$\frac{\partial x_i}{\partial T_1} = \frac{f}{Z_i}, \ \frac{\partial x_i}{\partial T_2} = 0, \ \frac{\partial x_i}{\partial T_3} = -f \frac{X_i}{Z_i^2}$$

and

$$\frac{\partial y_i}{\partial T_1} = 0, \ \frac{\partial y_i}{\partial T_2} = \frac{f}{Z_i}, \ \frac{\partial y_i}{\partial T_3} = -f \frac{Y_i}{Z_i^2}.$$

The partial derivatives with respect to the rotation angles are

$$\frac{\partial x_i}{\partial \phi_x} = -f \frac{X_i Y_i}{Z_i^2}, \ \frac{\partial x_i}{\partial \phi_y} = f \frac{Y_i^2 + Z_i^2}{Z_i^2}, \ \frac{\partial x_i}{\partial \phi_z} = f \frac{X_i}{Z_i^2}$$

and

$$\frac{\partial y_i}{\partial \phi_x} = -f \frac{Y_i^2 + Z_i^2}{Z_i^2}, \ \frac{\partial y_i}{\partial \phi_y} = f \frac{X_i Y_i}{Z_i^2}, \ \frac{\partial y_i}{\partial \phi_z} = f \frac{X_i}{Z_i^2}.$$

The six unknowns ΔT_1, ΔT_2, ΔT_3, $\Delta \phi_x$, $\Delta \phi_y$, and $\Delta \phi_z$ in equation (2.9) can be determined if at least three point correspondences are known.

In the iterative process, R and T are updated as follows:

$$T = T^{prev} + \Delta T \tag{2.10}$$

and

$$R = R^{prev} \cdot \Delta R, \tag{2.11}$$

where T^{prev} and R^{prev} are the values of T and R in the previous iteration, and ΔT and ΔR are the corrections of T and R, respectively. ΔR is calculated with $\Delta \phi_x$, $\Delta \phi_y$, and $\Delta \phi_z$ through equation (2.3).

The algorithm of Newton's method is summarized as follows:

The input is formed by n corresponding image and model points, with $n \geq 3$, and the initial estimates R^0 and T^0.

1) Using the current estimates of $\overline{R}$ and $\overline{T}$, compute P_i through equation (2.1).

2) Project P_i onto the image plane through equation (2.6).

3) Compute the residuals δx_i and δy_i $(i = 1,2,...,n)$.

4) Solve the linear system of n instances of equation (2.9) for the unknown corrections, ΔT_1, ΔT_2, ΔT_3, $\Delta\phi_x$, $\Delta\phi_y$, and $\Delta\phi_z$.

5) Update the current estimates of the translation vector T and rotation matrix R.

6) If the residuals are sufficiently small, exit; else go to step 1.

2.3.2 SPT Method

SPT (Single Projection Technique) [4, 5, 11] has been developed by Esthappan and Huffmann for implementing the orthogonal Procrustes algorithm [7, 8, 9]. The process of the method is composed of two parts described as follows:

The first part of the method uses the differences between the projected model points (x_i, y_i) and the measured image points (x_i^M, y_i^M) through the following equations to adjust T_1 and T_2 iteratively.

$$T_1 = T_1^{prev} - \left(\frac{1}{n}\sum_i [x_i - x_i^M]\right)\frac{T_3}{f}$$
$$T_2 = T_2^{prev} - \left(\frac{1}{n}\sum_i [y_i - y_i^M]\right)\frac{T_3}{f},$$

(2.12)

where T_1^{prev} and T_2^{prev} correspond to previous estimates of T_1 and T_2, respectively and n is the number of points in the model. Figure 2.2 shows the relation between (x_i, y_i) and (x_i^M, y_i^M).

In the second part of the method, R and T are optimized in an iterative manner by using the projection-Procrustes technique. Carrying out of the projection-Procrustes technique involves alignment of the point in the model with their respective projection lines, i.e., the ray trace from the origin, or the focal spot, of the camera reference frame to the measured image coordinates, $p_i^M = (x_i^M, y_i^M)$.
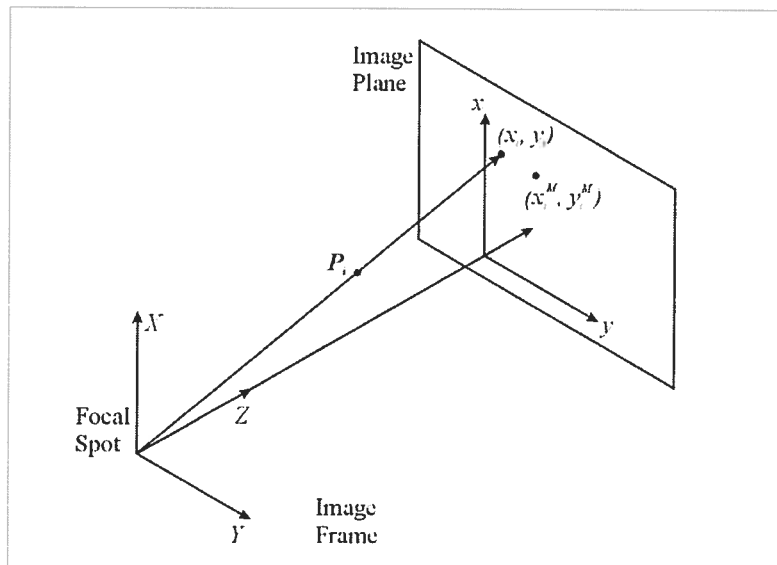
Figure 2. 2   Projection of a point P_i onto the image plane at $[x_i, y_i]^T$, and the corresponding measured image coordinates at $[x_i^M, y_i^M]^T$.

Firstly, the point positions, P_i, are projected onto their respective projection lines, forming the corresponding points, P_i^P; see Figure 2.3. P_i^P are given by:

$$P_i^P = [P_i \cdot \hat{h}]\hat{h},$$ (2.13)

where

$$\hat{h} = \frac{[x_i^M, y_i^M, f]}{\sqrt{(x_i^M)^2 + (y_i^M)^2 + f^2}}$$ (2.14)

is a unit vector directed from the focal spot to $p_i^M = (x_i^M, y_i^M)$, the measured image coordinates of the i-th point. The positions of the model points, P, are related to the points on their respective projection lines, P^P, in the camera reference frame according to the following equation:

$$P^P = sPA + \tau + E$$ (2.15)

where P and P^P are $n \times 3$ matrices representing the position of the model points in the camera reference frame and on the projection lines, respectively, A is a 3×3 rotation matrix, τ is a 1×3 vector representing the translation from the centroid of P to the centroid of P^P, s is a scalar, and E is an $n \times 3$ residual matrix. Equation (2.15) can be

rewritten as

$$E = (P^P - \tau) - sPA. \tag{2.16}$$
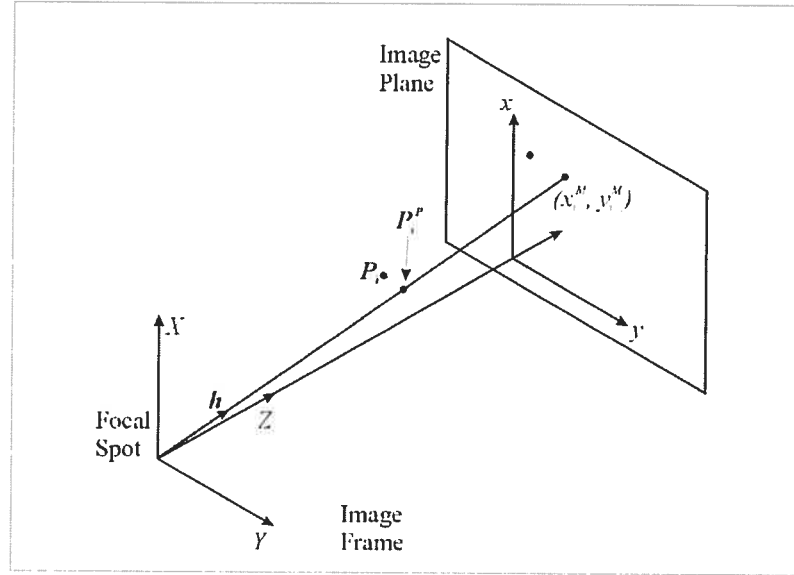
Figure 2.3   The point positions, P_i, are projected onto their repective perspective projection line, from the focal spot to the $[x_i^M, y_i^M]^T$, at the positions P_i^P.

Note that for some unknown reasons, our simulation shows that convergence of SPT was unacceptably slow which is different from the result given in [5]. We have made a modification to the formula given in [5] according to the Procrustes algorithm presented in [7, 8, 9]. We define a new coordinate system, named model^C, such that its origin is located at the centroid of P. Then we use P^C and $[P^P]^C$ to represent P and P^P in the model^C reference frame, and use P^C and $[P^P]^C$ to replace P and $P^P - \tau$ in equation (2.15), respectively. The translation from camera reference frame to the model^C reference frame is denoted as T^C. In the model^C reference frame, the equation (2.15) is rewritten as

$$E = [P^P]^C - sP^C A. \tag{2.17}$$

In [5], τ is not considered explicitly in the process of projection-Procrustes and the translation T^C is not mentioned. After the modification above, the process can converge to the solution as expected.

The Procrustes algorithm [7, 8, 9] is used to determine the transformation (*i.e.*, A and s) that aligns P^C and $[P^P]^C$ optimally, *i.e.*, such that the transformation minimizes the sums of squares of E, given by

$$Tr(E^T \cdot E), \tag{2.18}$$

In equation (2.18), $Tr(A) = \sum_{i=1}^{n} a_{ii}$ is called the trace of square matrix A. The solution for A is computed by using the orthogonal Procrustes algorithm according to the equation:

$$A = UV^T, \tag{2.19}$$

where

$$[P^C]^T [P^P]^C = U\Sigma V^T, \tag{2.20}$$

i.e., $U\Sigma V^T$ is the singular value decomposion of $[P^C]^T [P^P]^C$. In the iterative process of the algorithm, the rotation R is adjusted through the following equation:

$$R = R^{prev} A, \tag{2.21}$$

where R^{prev} corresponds to the estimate of R in the previous iteration.

The model points are oriented and positioned in the camera reference frame according to the refined estimate of R and the estimate of T through equation (2.12). Subsequently, $p_i = [x_i, y_i]$ and P^P are adjusted through equations (2.6) and (2.13), respectively.

The scale factor, s, is computed through the following equation:

$$s = \frac{Tr\{(P^C A)^T [P^P]^C\}}{Tr(P^T P)} \tag{2.22}$$

Subsequently, T_1 and T_2 are adjusted through equation (2.12), while T_3 is adjusted according to the following equation:

$$T_3 = \frac{T_3^{prev}}{s}. \tag{2.23}$$

The above iterative process is repeated until the difference between P and P^P is sufficiently small. SPT method returns T_1, T_2, T_3, ϕ_x, ϕ_y, and ϕ_z, where ϕ_x, ϕ_y, and ϕ_z are computed by applying Euler angle decomposition to R [5].

In order to accelerate the convergence of the process, the translation, T', we propose strategy of applying to the initial model reference frame such that the centroid of

object is located at the origin of the model reference frame and add it to the SPT algorithm in our simulation. The translation is indicated by the following equation:

$$P_i^m = [P_i^m]^{in} + T',$$
(2.24)

where $[P_i^m]^{in}$ represents object points in the initial model reference frame as the input of the algorithm. Inversely, in order to obtain the transformation, T^{in}, referring to the initial model reference frame, the transformation indicated in the following equation is applied:

$$T^{in} = T + T' \cdot R,$$
(2.25)

where T and R are the results obtained from the SPT method in the model reference frame transformed by T'. We name equations (2.24) and (2.25) as the preprocess and the postprocess, respectively.

The SPT algorithm can now be summarized as follows:

The input is formed by n corresponding image and model points, with $n \geq 3$, and the initial estimates R^0 and T^0. The algorithm is divided into two parts.

Preprocess according to equation (2.24).

The first part:

Adjust T_1 and T_2 through equations (2.12) iteratively. In my program, the iteration process is repeated 12 times.

The second part:

1) Compute the unit vector of projection lines through equation (2.14). Estimate P through equation (2.1), and project P onto projection lines, P^P, through equation (2.13).

2) Translate P and P^P in camera reference frame into P^C and $[P^P]^C$ in model^C reference frame.

3) Compute the SVD of $[P^C]^T[P^P]^C$, and then compute A through equation (2.19).

4) Refine R through equation (2.20), T_1 and T_2, through equation (2.12).

5) Optimize T_3 through equations (2.21) and (2.22).

6) If the difference between P and P^P is not sufficiently small, go to step 1.

The postprocess is taken according to equation (2.25), and then results of T_1, T_2, T_3, ϕ_x, ϕ_y and ϕ_z are returned.

2.3.3 POSIT Method

POSIT (Pose from Orthography and Scaling with ITeration) [6, 14] has been developed by DeMenthon and Davis. The method approximates the perspective projection with a scaled orthographic projection (SOP), also known as weak perspective projection, and estimates the pose by solving a linear system. The above process is iterated to achieve more accurate pose by optimized SOP.

The scheme of the POSIT method is depicted in Figure 2.4. The model reference frame is centered at P_0 and its coordinate system is (u, v, w). The origin of camera is at the focal spot and its coordinate system is (X, Y, Z), with $i, j,$ and k representing the unit vectors along X, Y, and Z axes, respectively. The object point P_i in the camera reference frame is represented as $P_i = [X_i, Y_i, Z_i]$, and its perspective projection and SOP are represented as $p_i = [x_i, y_i]^T$ and $p_i^{SOP} = [x_i', y_i']^T$, respectively. Plane K is through P_0 and parallel to the image plane.

The goal of the POSIT method is to compute the rotation matrix and translation vector of the object. The rotation R is the matrix whose rows are the coordinates of the unit vector i, j, k of the camera reference frame expressed in the object coordinate system (u, v, w) and is written as:

$$R = \begin{pmatrix} i_u & i_v & i_w \\ j_u & j_v & j_w \\ k_u & k_v & k_w \end{pmatrix},$$

where i_u, i_v, i_w are the coordinates of i in the coordinate system (u, v, w) of the object, and similarly with j_u, j_v, j_w and k_u, k_v, k_w.

To compute the rotation, we only need to compute i and j in the model reference frame. The vector k is then obtained by $i \times j$. The translation T equals to OP_0, and therefore the coordinates of the translation vector are X_0, Y_0, Z_0. The point p_0 is the projection of point P_0 on the image plane; and the translation T is aligned with vector Op_0 and is equal to $\dfrac{Z_0}{f} Om_0$. Therefore to compute the object translation, we only need to compute its z-coordinate Z_0. Thus the object pose if fully defined once we find $i, j,$ and Z_0.

Figure 2.4 Scheme of the POSIT method. The image p_i is the perspective projection of object point P_i and image p_i^{SOP} is the scaled orthographic project (SOP) of the object point P_i. K is the plane through P_0 and parallel to the image plane.

The POSIT method is based on using SOP to approximate the perspective projection. Here, we choose Z_0 as the depth of SOP and therefore SOP of $P_i = [X_i, Y_i, Z_i]^T$ is expressed as:

$$[x_i', y_i']^T = \left[\frac{fX_i}{Z_0}, \frac{fZ_i}{Z_0} \right]^T \tag{2.26}$$

where $[x_i', y_i']^T$ is the image of P_i with SOP.

The ratio $s = \dfrac{f}{Z_0}$ is the scaling factor of the SOP. The point $P_0 = [X_0, Y_0, Z_0]^T$ has the same image $p_0 = [x_0, y_0]^T$ in both SOP and perspective projection.

From Figure 2.4, we have

$$P_0 P_i \cdot \frac{f}{Z_0} i = x_i(1 + \varepsilon_i) - x_0$$
$$P_0 P_i \cdot \frac{f}{Z_0} j = y_i(1 + \varepsilon_i) - y_0$$

(2.27)

where ε_i is defined as

$$\varepsilon_i = \frac{1}{Z_0} P_0 P_i \cdot k$$

(2.28)

and k is

$$k = i \times j.$$

(2.29)

The proof of equation (2.26) is given in [4].

Equation (2.27) can be rewritten as

$$P_0 P_i \cdot I = x_i(1 + \varepsilon_i) - x_0$$
$$P_0 P_i \cdot J = y_i(1 + \varepsilon_i) - y_0$$

(2.30)

where

$$I = \frac{f}{Z_0} i$$
$$J = \frac{f}{Z_0} j.$$

(2.31)

Given an estimate of ε_i, equation (2.30) provides a linear system of equations in which the only unknowns are the coordinates of I and J. The linear systems of equation (2.31) can be solved by using Linear Least Square method [7].

The POSIT algorithm starts with an estimate of ε_i, say $\varepsilon_i = 0$, and solves equation (2.30). Then a more accurate ε_i is obtained by firstly computing i and j through equation (2.31), then k through equation (2.29), finally ε_i through equation (2.28). The above process adjusts ε_i until the change of ε_i between the current and previous iterations is sufficiently small.

In order to meet the algorithm requirement that the object point P_0 must be located at the origin of the model reference frame, the translation, T', applied to the initial model reference frame has been added to the SPT algorithm in our simulation. The translation is

noted as:

$$P_i^m = [P_i^m]^{in} + T',$$ (2.32)

where $[P_i^m]^{in}$ represents object points in the initial model reference frame as the input of the algorithm. Inversely, in order to obtain the transformation, T^{in}, in the initial model reference frame, we compute:

$$T^{in} = T + T' \cdot R.$$ (2.33)

The following is the summary of the algorithm.

Preprocess according to equation (2.32).

1) Translate the object with T^0 so that the P_0 is located at the origin of the transformed model reference frame through equation (2.32).

2) Let $\varepsilon_i = 0$, $(i = 1,2,...,n)$.

3) Compute $x_i(1 + \varepsilon_i^{prev}) - x_0$ and $y_i(1 + \varepsilon_i^{prev}) - y_0$; solve for vectors I and J using the Linear Least Square method; and compute $i = \dfrac{I}{s_1}$ and $j = \dfrac{J}{s_2}$.

4) Compute new ε_i: using $k = i \times j$; $Z_0 = \dfrac{f}{s}$; $\varepsilon_i = \dfrac{1}{Z_0} P_0 P \cdot k$.

5) If $|\varepsilon_i - \varepsilon_i^{prev}| \leq Threshold$, return the result; else go to step 2.

Postprocess according to equation (2.33). Then return T_1, T_2, T_3, ϕ_x, ϕ_y, and ϕ_z.

2.4 Comparison of the Three Methods

Six groups of test data were used; the last three groups are generated randomly by a Mathematica program while the first three groups are chosen manually, and used to test the pose methods presented above. In order to evaluate the noise robustness of the methods, further tests were carried out by adding random noise with maximum amplitude of ±5% to the measured images in the above test data. The test data without noise effect and the results are listed in Table 2.1 through Table 2.7, while the test data by adding noise and the results are listed in Table 2.8 through Table 2.14.

Following are the main observations that can be made regarding the comparison of these three pose estimation methods from the data in Table 2.1 through Table 2.7.

- The POSIT method is the most efficient and the most accurate method.
- The POSIT method does not need an initial estimate of rotation and translation, while both SPT and Newton's methods need initial estimates.
- The Newton's method is only suitable for a small range of rotation angles, typically smaller than 30° around the Z-axis, even when a proper initial estimate of rotation angles are given. The SPT method is better than the Newton's method, but is also limited to suitable rotation angles range due to the possibility of solutions falling into local minima. The drawbacks and a solving strategy are discussed in [11]. The POSIT, which does not need the initial estimates of R and T can be used with any rotation angles.

In conclusion, according to our simulation, the POSIT method demonstrates fairly good performances in terms of accuracy, efficiency, and is suitable for the whole rotation range.

Compare the test results between the data in Table 2.1 through Table 2.7 and the data in Table 2.1 through Table 2.7, we can see that when adding random noise to the measured images, the accuracy of the POSIT is affected significantly. For most of cases, for example: the data in Table 2.8, Table 2.9, Table 2.11, Table 2.12, and Table 2.13, the errors by the POSIT are even worse than the errors by SPT method.

Table 2.1 Within a small rotation around the Z-axis, Newton's method, SPT method, and POSIT obtain proper solutions. SPT demonstrates a good performance in terms of accuracy, while POSIT demonstrates a good performance in terms of efficiency.

$$Object = \begin{pmatrix} 0 & 0 & 0 \\ 10 & 0 & 0 \\ 10 & 10 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10 \\ 10 & 0 & 10 \\ 10 & 10 & 10 \\ 0 & 10 & 10 \end{pmatrix}, \phi_R = \begin{bmatrix} 10 & -15 & 20 \end{bmatrix}^T, T = \begin{bmatrix} 20 & 20 & 35 \end{bmatrix}^T, f = 100$$

$$image = \begin{pmatrix} 57.1 & 57.1 \\ 76.5 & 60.4 \\ 66.5 & 83.5 \\ 46.7 & 82.2 \\ 39.1 & 41.1 \\ 55.8 & 44.8 \\ 48.0 & 63.6 \\ 31.1 & 61.2 \end{pmatrix}$$

	Newton's Method	SPT	POSIT
Initial Estimates	$T_0=[0,0,f/2]^T$, $\phi R_0=[0,0,0]^T$	$T_0=[0,0,f/2]^T$, $\phi R_0=[0,0,0]^T$	
Time Cost (sec.)	0.14	0.09	0.04
Num of Iterations	30	20	20
Error of R (°)	0.0367	0.0000197	0.00155
Error of T	0.00673	0.0000149	0.00111

Table 2.2 Within a small rotation around the Z-axis, Newton's method, SPT method and POSIT obtain proper solutions. POSIT demonstrates good performances in terms of accuracy and efficiency.

$$Object = \begin{pmatrix} 0 & 0 & 0 \\ 10 & 0 & 0 \\ 10 & 10 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10 \\ 10 & 0 & 10 \\ 10 & 10 & 10 \\ 0 & 10 & 10 \end{pmatrix}, \phi_R = \begin{bmatrix} 30 & -40 & 50 \end{bmatrix}^T, T = \begin{bmatrix} 20 & -20 & 60 \end{bmatrix}^T, f = 100$$

$$image = \begin{pmatrix} 33.3 & -33.3 \\ 36.9 & -22.8 \\ 28.7 & -11.1 \\ 23.9 & -20.3 \\ 20.3 & -35.7 \\ 24.9 & -26.0 \\ 17.3 & -15.3 \\ 11.7 & -24.1 \end{pmatrix}$$

	Newton's Method	SPT	POSIT
Initial Estimates	$T_0=[0,0,f/2]^T$, $\phi R_0=[0,0,0]^T$	$T_0=[0,0,f/2]^T$, $\phi R_0=[0,0,0]^T$	
Time Cost (sec.)	0.18	0.1	0.04
Num of Iterations	36	24	20
Error of R (°)	0.386	2.49×10^{-6}	1.91×10^{-13}
Error of T	0.559	2.42×10^{-7}	8.70×10^{-15}

Table 2.3 With this test data, both Newton's method and SPT method cannot obtain a proper solution.

$$Object = \begin{pmatrix} 0 & 0 & 0 \\ 10 & 0 & 0 \\ 10 & 10 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10 \\ 10 & 0 & 10 \\ 10 & 10 & 10 \\ 0 & 10 & 10 \end{pmatrix}, \phi_R = \begin{bmatrix} 180 & -5 & 5 \end{bmatrix}^T, T = \begin{bmatrix} -10 & -10 & 35 \end{bmatrix}^T, f = 100$$

$$image = \begin{pmatrix} -28.6 & -28.6 \\ -0.222 & -31.8 \\ -2.76 & -60.9 \\ -30.9 & -56.9 \\ -43.4 & -39.9 \\ -3.92 & -44.9 \\ -7.49 & -85.9 \\ -46.7 & -79.4 \end{pmatrix}$$

	Newton's Method	SPT	POSIT
Initial Estimates	T_0=[0,0,f/2]^T, ϕR_0=[0,0,0]^T	T_0=[0,0,f/2]^T, ϕR_0=[0,0,0]^T	
Time Cost (sec.)	Cannot achieve result	Cannot achieve result	0.04
Num of Iterations			20
Error of R (°)			0.0000203
Error of T			4.35×10^{-6}

Table 2.4   The same test data as in Table 6.1(c1), but a proper set of initial estimates of rotation angles are given, SPT obtains a proper solution, while Newton's method still fails to find the solution.

$$Object = \begin{pmatrix} 0 & 0 & 0 \\ 10 & 0 & 0 \\ 10 & 10 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10 \\ 10 & 0 & 10 \\ 10 & 10 & 10 \\ 0 & 10 & 10 \end{pmatrix}, \phi_R = \begin{bmatrix} 180 & -5 & 5 \end{bmatrix}^T, T = \begin{bmatrix} -10 & -10 & 35 \end{bmatrix}^T, f = 100$$

$$image = \begin{pmatrix} -28.6 & -28.6 \\ -0.222 & -31.8 \\ -2.76 & -60.9 \\ -30.9 & -56.9 \\ -43.4 & -39.9 \\ -3.92 & -44.9 \\ -7.48 & -85.9 \\ -46.7 & -79.5 \end{pmatrix}$$

	Newton's Method	SPT	POSIT
Initial Estimates	$T_0=[0,0,f/2]^T, \phi R_0=[150,20,20]^T$	$T_0=[0,0,f/2]^T, \phi R_0=[150,20,20]^T$	
Time Cost (sec.)	Cannot achieve result	Cannot achieve result	0.04
Num of Iterations			20
Error of R (°)			0.0000203
Error of T			4.35×10^{-6}

Table 2.5   With this test data, Newton's method fails to find the solution. POSIT method demonstrates better performance in terms of efficiency and accuracy.

$$Object = \begin{pmatrix} 0 & 0 & 0 \\ 10 & 0 & 0 \\ 10 & 10 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10 \\ 10 & 0 & 10 \\ 10 & 10 & 10 \\ 0 & 10 & 10 \end{pmatrix}, \phi_R = \begin{bmatrix} 60 & -5 & 60 \end{bmatrix}^T, T = \begin{bmatrix} -10 & -10 & 70 \end{bmatrix}^T, f = 100$$

$$image = \begin{pmatrix} -14.3 & -14.3 \\ -6.45 & -7.78 \\ -16.7 & -3.54 \\ -25.1 & -9.25 \\ -14.4 & -24.8 \\ -7.12 & -17.7 \\ -16.7 & -13.2 \\ -24.7 & -19.6 \end{pmatrix}$$

	Newton's Method	SPT	POSIT
Initial Estimates	T_0=[0,0,f/2]^T, ϕR_0=[0,0,0]^T	T_0=[0,0,f/2]^T, ϕR_0=[0,0,0]^T	
Time Cost (sec.)	Cannot achieve result	0.1	0.04
Num of Iterations		20	20
Error of R (°)		6.55×10^{-5}	3.02×10^{-14}
Error of T		9.84×10^{-6}	2.94×10^{-14}

Table 2.6   With this test data, Newton's method failed to find the solution. The POSIT method demonstrates better performance in terms of efficiency and accuracy.

$$Object = \begin{pmatrix} -5 & -3 & 10 \\ -5 & 5 & -4 \\ -10 & 9 & -3 \\ -5 & -7 & -4 \\ -10 & 4 & 1 \end{pmatrix}, \phi_R = \begin{bmatrix} 207 & -49 & 22 \end{bmatrix}^T, T = \begin{bmatrix} -30 & 17 & 40 \end{bmatrix}^T, f = 100$$

$$image = \begin{pmatrix} -103 & 59.3 \\ -68.7 & 24.4 \\ -74.8 & 15.9 \\ -59.5 & 47.5 \\ -81.0 & 29.2 \end{pmatrix}$$

	Newton's Method	SPT	POSIT
Initial Estimates	$T_0=[0,0,f/2]^T$, $\phi R_0=[0,0,0]^T$	$T_0=[0,0,f/2]^T$, $\phi R_0=[0,0,0]^T$	
Time Cost (sec.)	Cannot achieve result	0.07	0.04
Num of Iterations		20	20
Error of R (°)		2.17	4.09×10^{-6}
Error of T		0.207	1.03×10^{-6}

Table 2.7 With the test data, Newton's method failed to find the solution. The POSIT method demonstrates better performance in terms of efficiency and accuracy.

$$Object = \begin{pmatrix} 4 & -9 & 7 \\ 10 & 5 & -9 \\ -2 & 5 & -2 \\ 7 & -8 & 1 \end{pmatrix}, \phi_R = \begin{bmatrix} 298 & -65 & 24 \end{bmatrix}^T, T = \begin{bmatrix} 9 & -4 & 41 \end{bmatrix}^T, f = 100$$

$$image = \begin{pmatrix} 11.1 & 2.66 \\ 58.2 & 6.87 \\ 25.7 & -17.0 \\ 24.7 & 4.06 \end{pmatrix}$$

	Newton's Method	SPT	POSIT
Initial Estimates	$T_0=[0,0,f/2]^T$, $\phi R_0=[0,0,0]^T$	$T_0=[0,0,f/2]^T$, $\phi R_0=[0,0,0]^T$	
Time Cost (sec.)	Cannot achieve result	0.07	0.04
Num of Iterations		20	20
Error of R (°)		0.153	3.05×10^{-8}
Error of T		0.0237	6.00×10^{-10}

Table 2.8   By adding random noise with maximum amplitude of ±5% to the measured images in the test data in Table 2.1, it is shown in this case that the errors of the result by the POSIT become larger than the errors of the result by the SPT. Newton's method still has the largest errors in its result.

$$Object = \begin{pmatrix} 0 & 0 & 0 \\ 10 & 0 & 0 \\ 10 & 10 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10 \\ 10 & 0 & 10 \\ 10 & 10 & 10 \\ 0 & 10 & 10 \end{pmatrix}, \phi_R = \begin{bmatrix} 10 & -15 & 20 \end{bmatrix}^T, T = \begin{bmatrix} 20 & 20 & 35 \end{bmatrix}^T, f = 100$$

$$image = \begin{pmatrix} 54.7 & 54.7 \\ 74.2 & 58.6 \\ 69.5 & 87.2 \\ 45.6 & 80.3 \\ 38.9 & 40.9 \\ 56.9 & 45.7 \\ 47.4 & 62.7 \\ 32.2 & 63.3 \end{pmatrix}$$

	Newton's Method	SPT	POSIT
Initial Estimates	$T_0=[0,0,f/2]^T$, $\phi R_0=[0,0,0]^T$	$T_0=[0,0,f/2]^T$, $\phi R_0=[0,0,0]^T$	
Time Cost (sec.)	0.18	0.09	0.05
Num of Iterations	36	20	20
Error of R (°)	6.88	5.08	5.80
Error of T	1.09	0.780	1.00

Table 2.9 By adding random noise with maximum amplitude of ±5% to the measured images in the test data in Table 2.2, it is shown in this case that the errors of the results by the three methods have no significant differences.

$$Object = \begin{pmatrix} 0 & 0 & 0 \\ 10 & 0 & 0 \\ 10 & 10 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10 \\ 10 & 0 & 10 \\ 10 & 10 & 10 \\ 0 & 10 & 10 \end{pmatrix}, \phi_R = \begin{bmatrix} 30 & -40 & 50 \end{bmatrix}^T, T = \begin{bmatrix} 20 & -20 & 60 \end{bmatrix}^T, f = 100$$

$$image = \begin{pmatrix} 33.4 & -33.4 \\ 36.1 & -22.3 \\ 28.9 & -11.2 \\ 24.0 & -20.3 \\ 19.9 & -35.1 \\ 24.2 & -25.2 \\ 17.0 & -15.1 \\ 11.8 & -24.2 \end{pmatrix}$$

	Newton's Method	SPT	POSIT
Initial Estimates	T_0=[0,0,f/2]^T, ϕR_0=[0,0,0]^T	T_0=[0,0,f/2]^T, ϕR_0=[0,0,0]^T	
Time Cost (sec.)	0.17	0.09	0.04
Num of Iterations	36	20	20
Error of R (°)	3.66	2.91	3.81
Error of T	1.24	1.25	0.670

Table 2.10 By adding random noise with maximum amplitude of ±5% to the measured images in the test data in Table 2.3, it is shown in this case that the errors of the result by the POSIT become significant large.

$$Object = \begin{pmatrix} 0 & 0 & 0 \\ 10 & 0 & 0 \\ 10 & 10 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10 \\ 10 & 0 & 10 \\ 10 & 10 & 10 \\ 0 & 10 & 10 \end{pmatrix}, \phi_R = \begin{bmatrix} 180 & -5 & 5 \end{bmatrix}^T, T = \begin{bmatrix} -10 & -10 & 35 \end{bmatrix}^T, f = 100$$

$$image = \begin{pmatrix} -27.1 & -27.1 \\ -0.217 & -31.1 \\ -2.88 & -63.6 \\ -29.7 & -54.6 \\ -43.8 & -40.2 \\ -3.97 & -45.5 \\ -7.65 & -87.8 \\ -46.0 & -78.3 \end{pmatrix}$$

	Newton's Method	SPT	POSIT
Initial Estimates	$T_0=[0,0,f/2]^T$, $\phi R_0=[0,0,0]^T$	$T_0=[0,0,f/2]^T$, $\phi R_0=[0,0,0]^T$	
Time Cost (sec.)	Cannot achieve result	Cannot achieve result	0.04
Num of Iterations			20
Error of R (°)			1.14
Error of T			0.849

Table 2.11 By adding random noise with maximum amplitude of ±5% to the measured images in the test data in Table 2.4, it is shown in this case that the errors of the result by the POSIT become much larger than the errors shown in Table 2.4.

$$Object = \begin{pmatrix} 0 & 0 & 0 \\ 10 & 0 & 0 \\ 10 & 10 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10 \\ 10 & 0 & 10 \\ 10 & 10 & 10 \\ 0 & 10 & 10 \end{pmatrix}, \phi_R = \begin{bmatrix} 180 & -5 & 5 \end{bmatrix}^T, T = \begin{bmatrix} -10 & -10 & 35 \end{bmatrix}^T, f = 100$$

$$image = \begin{pmatrix} -27.3 & -27.3 \\ -0.217 & -31.0 \\ -2.84 & -62.8 \\ -31.8 & -58.4 \\ -44.1 & -40.6 \\ -4.08 & -46.8 \\ -7.51 & -86.1 \\ -45.8 & -77.8 \end{pmatrix}$$

	Newton's Method	SPT	POSIT
Initial Estimates	$T_0=[0,0,f/2]^T, \phi R_0=[150,20,20]^T$	$T_0=[0,0,f/2]^T, \phi R_0=[150,20,20]^T$	
Time Cost (sec.)	Cannot achieve result	Cannot achieve result	0.04
Num of Iterations			20
Error of R (°)			3.19
Error of T			0.886

Table 2.12   By adding random noise with maximum amplitude of ±5% to the measured images in the test data in Table 2.5, it is shown in this case that the errors of the result by the POSIT become larger than the errors of the result by the SPT.

$$Object = \begin{pmatrix} 0 & 0 & 0 \\ 10 & 0 & 0 \\ 10 & 10 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10 \\ 10 & 0 & 10 \\ 10 & 10 & 10 \\ 0 & 10 & 10 \end{pmatrix}, \phi_R = \begin{bmatrix} 60 & -5 & 60 \end{bmatrix}^T, T = \begin{bmatrix} -10 & -10 & 70 \end{bmatrix}^T, f = 100$$

$$image = \begin{pmatrix} -14.2 & -14.2 \\ -6.70 & -8.08 \\ -17.0 & -3.61 \\ -26.3 & -9.67 \\ -14.0 & -24.1 \\ -7.13 & -17.7 \\ -17.2 & -13.6 \\ -25.4 & -20.2 \end{pmatrix}$$

	Newton's Method	SPT	POSIT
Initial Estimates	T_0=[0,0,f/2]^T, ϕR_0=[0,0,0]^T	T_0=[0,0,f/2]^T, ϕR_0=[0,0,0]^T	
Time Cost (sec.)	Cannot achieve result	0.1	0.04
Num of Iterations		20	20
Error of R (°)		0.648	1.58
Error of T		1.71	1.77

Table 2.13   By adding random noise with maximum amplitude of ±5% to the measured images in the test data in Table 2.6, it is shown in this case that the errors of the result by the POSIT become as large as the errors of the result by the SPT..

$$Object = \begin{pmatrix} -5 & -3 & 10 \\ -5 & 5 & -4 \\ -10 & 9 & -3 \\ -5 & -7 & -4 \\ -10 & 4 & 1 \end{pmatrix}, \phi_R = \begin{bmatrix} 207 & -49 & 22 \end{bmatrix}^T, T = \begin{bmatrix} -30 & 17 & 40 \end{bmatrix}^T, f = 100$$

$$image = \begin{pmatrix} -102 & 58.9 \\ -67.3 & 23.9 \\ -75.8 & 16.1 \\ -60.8 & 48.5 \\ -77.9 & 28.1 \end{pmatrix}$$

	Newton's Method	SPT	POSIT
Initial Estimates	$T_0=[0,0,f/2]^T$, $\phi R_0=[0,0,0]^T$	$T_0=[0,0,f/2]^T$, $\phi R_0=[0,0,0]^T$	
Time Cost (sec.)	Cannot achieve result	0.1	0.04
Num of Iterations		20	20
Error of R (°)		5.38	2.76
Error of T		0.402	0.704

Table 2.14   By adding random noise with maximum amplitude of ±5% to the measured images in the test data in Table 2.7, it is shown in this case that the errors of the result by the POSIT become larger than the errors of the result by the SPT.

$$Object = \begin{pmatrix} 4 & -9 & 7 \\ 10 & 5 & -9 \\ -2 & 5 & -2 \\ 7 & -8 & 1 \end{pmatrix}, \phi_R = \begin{bmatrix} 298 & -65 & 24 \end{bmatrix}^T, T = \begin{bmatrix} 9 & -4 & 41 \end{bmatrix}^T, f = 100$$

$$image = \begin{pmatrix} 11.5 & 2.75 \\ 59.7 & 7.05 \\ 24.5 & -16.2 \\ 23.7 & 3.90 \end{pmatrix}$$

	Newton's Method	SPT	POSIT
Initial Estimates	$T_0=[0,0,f/2]^T$, $\phi R_0=[0,0,0]^T$	$T_0=[0,0,f/2]^T$, $\phi R_0=[0,0,0]^T$	
Time Cost (sec.)	Cannot achieve result	0.07	0.04
Num of Iterations		20	20
Error of R (°)		1.44	13.1
Error of T		1.11	1.68

2.5 Summary

In the reconstruction of vessels from IVUS images, biplane angiography or single projection angiography is needed. Biplane angiography has the drawbacks of the expensive equipment, complex operation, and requirement of calibration. Because single projection angiography overcomes most of the drawbacks existing in biplane angiography, it has drawn increasing attention from researchers.

Single projection angiography employs the method of pose estimation by single projection. Among many publications on pose estimation by single projection, we have studied three methods. They are Newton's method, SPT, and POSIT. The simulation results have shown that the POSIT method is the best method in terms of accuracy, efficiency, and stability among the methods we have studied.

Although single projection angiography overcomes the drawbacks existing in biplane angiography, on the other hand, it requires a prior knowledge of the 3-D configuration of vessels. In the next chapter, we will investigate the possibility and propose a method of using an image sequence taken by single projection angiography to reconstruct a trajectory of catheter in IVUS intervention without a prior knowledge of configuration of vessels.

Chapter 3

A Method of Pose Estimation by an Image Sequence from Single Projection

3.1 Introduction

As presented in Chapter 2, in the reconstruction of catheter trajectory in 3-D IVUS artery intervention, either biplane angiography or single projection angiography is employed to provide the information of location and orientation for each frame of IVUS images. Single projection angiography has advantages of low equipment cost and is easier to operate over biplane angiography and therefore has drawn a lot of attention from researchers. However, for single projection angiography, there exists a limitation that a 3-D configuration of artery is required. That is due to the fact that the method of pose estimation by single projection, which is used in single projection angiography, requires prior knowledge of the 3-D configuration of the object. The 3-D configuration of artery can be obtained by computed tomography (CT) or magnetic resonance (MR). This consequently increases clinical cost.

In this chapter, we propose a novel method of pose estimation with image sequence taken from a single view. Compare with the methods discussed in Chapter 2, the proposed method only needs the measurement of pullback length of the catheter, instead of the requirement of the knowledge of the 3-D configuration of the object, instead. In a typical IVUS intervention, an angiographic sequence from single projection can be taken to localize the IVUS transducer during its pullback. Based on this fact, we propose a new method of pose estimation with an image sequence taken from a single view with no need of prior knowledge of the object 3-D configuration, targeting for an implementation in angiography. The proposed method avoids the additional CT or MR

imaging used in single projection angiography, and thus reduces the clinic cost and simplifies the IVUS investigation.

The proposed method is simulated in Mathematica by using a given spiral curve to represent the configuration of an artery. Several cases are implemented with various locations, orientations, and sizes of the given curve.

The chapter is organized as follows. In Section 2, we describe the configuration of an angiography system during an IVUS intervention. In Section 3, we present in detail our new method of pose estimation with an image sequence from single projection. In Section 4, we present the simulation of the proposed method in Mathematica. In Section 5, we give a summary of this chapter.

3.2 Angiography in IVUS intervention

In a typical 3-D IVUS intervention, angiography is used to determine the location and orientation of IVUS images. In this section, we describe the configuration of a typical angiography system geometrically. A set of typical measurements for the configuration is given, which will be used in the simulation of the proposed method.

Figure 3.1 shows the configuration of a coronary angiography system for 3-D IVUS. The focal length, f, is the distance from a focal spot (*i.e.*, X-ray source) to the image plane. The patient is located between the focal spot and the image plane. We use d to indicate the distance from the focal spot to the centre of the patient's heart. Notice that d cannot be measured accurately. In a 3-D IVUS intervention, a miniaturized ultrasonic transducer at the end of a catheter is inserted in the artery lumen, and then withdrawn gradually. An image sequence is taken during the pullback of the catheter. The pullback length of a catheter between grabbing images i and image $i+1$ is indicated as Δl_i, as shown in Figure 3.2. In the following discussion, we assume that $f=1000mm$ and $d=500mm$. The size of the heart is within a sphere of radius $r=100mm$. These assumptions for measurements are close to the real clinical situation.
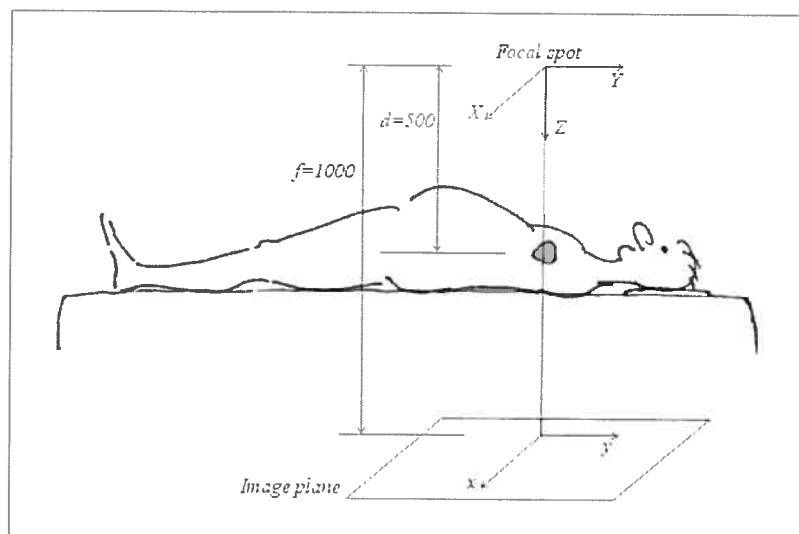
Figure 3.1 The configuration of a system of IVUS combined with coronary angiography

3.3 A Method of Pose Estimation with an Image Sequence from Single Projection

In this section, we propose a method of pose estimation with an image sequence from single projection. The method is intended to be implemented for the reconstruction of the trajectory of a catheter from an image sequence taken by single angiography. In the following discussion, the trajectory of a catheter is viewed as a 3-D curve.

This section is organized as follows. In subsection 1, we geometrically analyze the problem of pose estimation by an image sequence from single projection. In subsection 2, we discuss how to find multi-solutions for the trajectory curve. A strategy for eliminating invalid solutions using physical characteristics of the catheter path is presented in subsection 3. In subsection 4, we discuss the strategy that we use to reduce the accumulated error. In subsection 5, we discuss how to determine the location of the start point of the trajectory curve. In the last subsection, we summarize the proposed method.

3.3.1 A Geometrical Analysis of the Problem of Pose Estimation by an Image Sequence from Single Projection

We define the coordinate system for an angiography camera system as (X, Y, Z) with its origin located at the focal spot and image plane at $Z=f$; and define the coordinate

system for an image as (x, y), which lies in the image plane and is parallel to the X and Y axes. The coordinate systems (X, Y, Z) and (x, y) are illustrated in Figure 3.1. As shown in Figure 3.2, $P_i(X_i, Y_i, Z_i)$ and $p_i(x_i, y_i)$ respectively represent the 3-D location of a catheter tip and its corresponding projection on the image plane when the i-th coronary angiography image is captured. Since the size of the heart, within a sphere of radius $r=100mm$, is significantly smaller than the focal length, $f=1000mm$, we assume that catheter tip images are caught under a weak perspective projection. Furthermore, we assume that the length of the curve, Δl_i, is small enough so that each segment of the trajectory curve can be viewed as a straight line with its length approximately equal to Δl_i. Based on the above assumptions, the relation between a segment of the trajectory of the catheter tip, P_iP_{i+1}, and its weak perspective image, p_ip_{i+1}, as shown in Figure 3.2, can be represented as the following equations:

$$X_{i+1} - X_i = (x_{i+1} - x_i) \cdot s \tag{3.1a}$$

$$Y_{i+1} - Y_i = (y_{i+1} - y_i) \cdot s \tag{3.1b}$$

$$Z_{i+1} - Z_i = \pm\sqrt{\Delta l_i^2 - (x_{i+1} - x_i)^2 \cdot s^2 - (y_{i+1} - y_i)^2 \cdot s^2}, \tag{3.1c}$$

where $s = d/f$, a scale factor for weak perspective projection, and Δl_i is the pullback length of the catheter in the interval between capturing the i-th image and capturing the $i+1$-th image.

Based on the assumption of weak perspective projection, the X and Y-coordinates of the point on the trajectory corresponding to a point in the image can be estimated by following equations:

$$\begin{aligned} X_i &= sx_i \\ Y_i &= sy_i. \end{aligned} \tag{3.2}$$

Furthermore, we assume that the Z-coordinate of the start point of the trajectory curve, noted as Z_1, is known, and the length of pullback of catheter, Δl_i, is known, then Z_i $(i>0)$ can be computed by using equation (3.1c). Later in this chapter, we will present a strategy for estimation of Z_1.
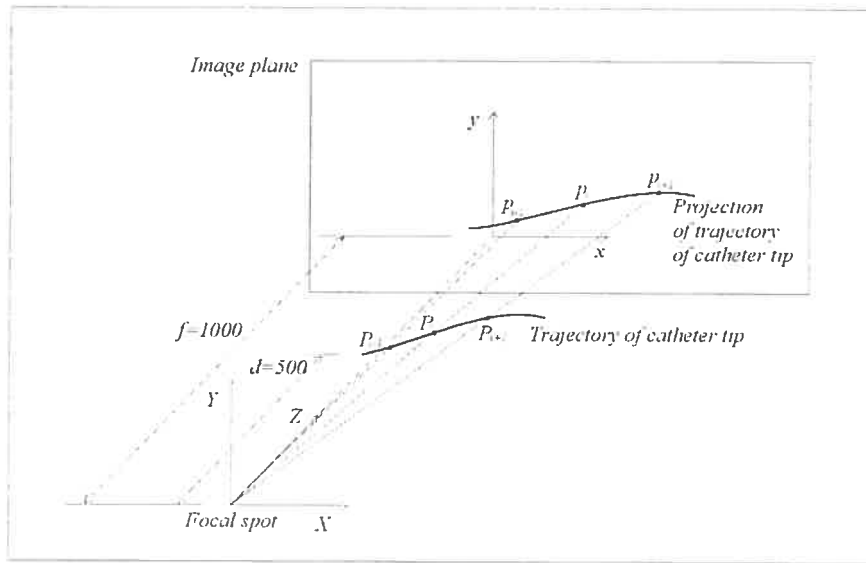
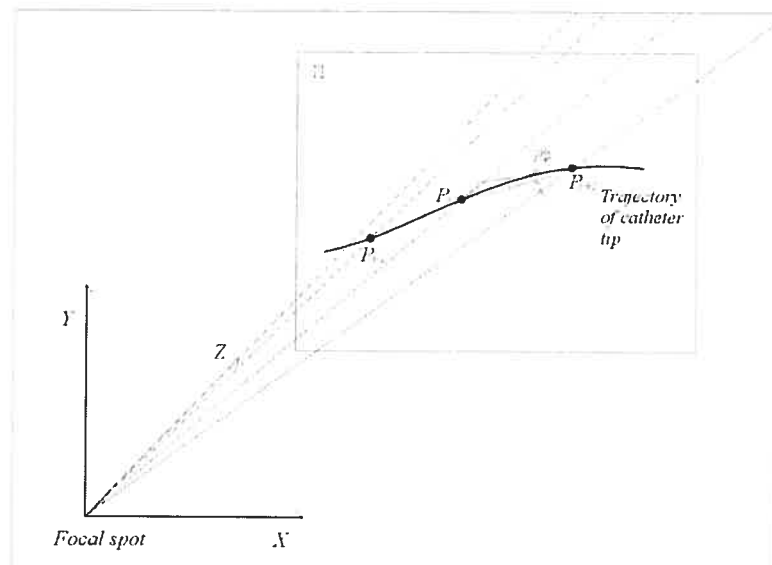Figure 3.2   A segment of trajectory of a catheter tip and its projection on the image plane

Figure 3.3 Π is the plane that passes through the point P_{i+1} and is parallel to the X and Y-axis, ϕ is the angle between the line of P_iP_{i+1} and the plane Π.

3.3.2 Finding Solutions for a Trajectory Curve

Obviously, X_i and Y_i can be easily computed through equation (3.2) and Z_i can be computed deductively through equation (3.1c) providing Z_l is previously known. However, there is an undetermined sign in equation (3.1c), which will lead to multiple solutions for Z_i. This subsection will focus on the determination of the sign in equation (3.1c).

The part of trajectory of the catheter tip in Figure 3.2 with respect to the object is enlarged in Figure 3.3, in order to depict the problem in detail. We use Π to denote the plane that passes through the point P_{i+1} and is parallel to the X and Y-axis, ϕ to denote the angle between the line of P_iP_{i+1} and the plane Π. In the following, we will consider two situations according to weather $\phi \leq$ *threshold* or $\phi >$ *threshold*. Geometrically, the difference between the two situations is that in the former one, the line P_iP_{i+1} is almost parallel to the plane Π; while in the later one, there is a significantly large angle between the line P_iP_{i+1} and the plane Π.

Firstly, we deal with the situation when ϕ is greater than the threshold. Due to the physical characteristic of a catheter and the configuration of an artery, the trajectory of the catheter tip must be a smooth curve, *i.e.*, there is no sharp bend on the trajectory. This characteristic can be represented mathematically by assuming that the value of the second derivative of the curve at the point P_i will be significantly small. Figure 3.2 shows the trajectory of a catheter tip that passes through two points, P_i and P_{i+1}, which are corresponding to the image points p_i, and p_{i+1}. Now we discuss the problem of solving P_{i+1} from known p_i, p_{i+1}, and P_i. In equation (3.1), Δl_i, the length of curve P_iP_{i+1}, can be approximated as the length of line P_iP_{i+1}, providing the length of curve P_iP_{i+1} is significantly small. By solving equation (3.1c), we have two possible solutions, denoted as $P_{i+1}^{(1)}(X_{i+1}^{(1)}, Y_{i+1}^{(1)}, Z_{i+1}^{(1)})$ and $P_{i+1}^{(2)}\left(X_{i+1}^{(2)}, Y_{i+1}^{(2)}, Z_{i+1}^{(2)}\right)$, for the point $P_{i+1}(X_{i+1}, Y_{i+1}, Z_{i+1})$, due to the options of sign in equation (3.1c). Now we need to determine the solution that best approximates P_{i+1} between $P_{i+1}^{(1)}$ and $P_{i+1}^{(2)}$. To achieve this, we use a strategy of comparing their second derivatives of two segments of curves constructed by possible solutions at point P_i. That is, the best solution represents a smooth curve, and therefore

has a smaller absolute value for the second derivative than the bad solution. We represent the trajectory of a catheter tip as the following equations:

$$\begin{aligned} X &= X(t) \\ Y &= Y(t) \\ Z &= Z(t). \end{aligned} \qquad (3.3)$$

The second derivatives of the X, Y, and Z at point P_i can be estimated as:

$$\begin{aligned} X'' &\approx \frac{\Delta X_i - \Delta X_{i-1}}{\Delta t_i - \Delta t_{i-1}} \\ Y'' &\approx \frac{\Delta Y_i - \Delta Y_{i-1}}{\Delta t_i - \Delta t_{i-1}} \\ Z'' &\approx \frac{\Delta Z_i - \Delta Z_{i-1}}{\Delta t_i - \Delta t_{i-1}} \end{aligned} \qquad (3.4)$$

where ΔX_i, ΔY_i, ΔZ_i, ΔX_{i-1}, ΔY_{i-1}, and ΔZ_{i-1} represent $X_{i+1} - X_i$, $Y_{i+1} - Y_i$, $Z_{i+1} - Z_i$, $X_i - X_{i-1}$, $Y_i - Y_{i-1}$, and $Z_i - Z_{i-1}$, respectively. Here $P_{i-1}(X_{i-1}, Y_{i-1}, Z_{i-1})$, the point previous to the point P_i, is assumed to be known. The best solution of P_{i+1} makes the curve pass through the points P_{i-1}, P_i, P_{i+1} smoother, and therefore has smaller value for X'', Y'', and Z'' computed by equations (3.4). We use the following formula to estimate the values for the second derivative of the curve at the point P_i:

$$\sqrt{\left(\frac{\Delta X_i - \Delta X_{i-1}}{\Delta t_i - \Delta t_{i-1}}\right)^2 + \left(\frac{\Delta Y_i - \Delta Y_{i-1}}{\Delta t_i - \Delta t_{i-1}}\right)^2 + \left(\frac{\Delta Z_i - \Delta Z_{i-1}}{\Delta t_i - \Delta t_{i-1}}\right)^2}. \qquad (3.5)$$

Based on the discussion above, the best solution for P_{i+1} will be the solution obtained from equation (3.2) and equation (3.1c) with smaller value given by formula (3.5).

Secondly, we deal with the situation of $\phi \leq threshold$. In this situation, the line $P_{i-1}P_i$ is almost parallel to the image plane, and values for the second derivative of the two possible solution curves at the point P_i are significantly close. Consequently, we are not able to determine the sign for equation (3.1c) by the assumption of the smooth characteristic of a trajectory curve. In this situation, we respectively compute the two possible solutions and, based on the two possible solutions, compute further solutions recursively until the reconstructed trajectory curves extend long enough. On the

trajectory, if there are n points where $\phi \leq$ *threshold*, then there are 2^n groups of possible solutions for the trajectory curve. The problem of choosing the right solution from multi-solutions is tackled in the next subsection.

3.3.3 Determination of the right solution from multiple solutions

This section discusses how to choose the right solution from multi-solutions obtained in the last subsection. We first project the points of each possible solution onto the image plane using the following equations:

$$X_i = \frac{Z_i}{f} x_i$$
$$Y_i = \frac{Z_i}{f} y_i. \tag{3.6}$$

For each possible solution, we compute the Euclidean distances between the projections of solution points and the measured images, and pick the solution associated with the smallest Euclidean distances as the right solution.

3.3.4 Reducing Accumulated Error

The assumption of weak perspective projection in equations (3.1) introduces errors into the solution for each segment of trajectory curve. Furthermore, deductively using equation (3.1c) produces a significant amount of accumulated error for the solution of trajectory curve. This subsection proposes a strategy to reduce the accumulated error.

As presented in subsection 3.2.1, we assume that the coronary angiography system is a weak perspective projection camera system. The main reason for this assumption is that the Z-coordinate of the trajectory curve cannot be measured in the clinical situation. Based on this assumption, we use d, depicted in Figure 3.1, as the average distance from the object to the focal spot. In the Chapter 2 of [3], it is suggested that a weak perspective projection becomes viable when $\frac{\overline{Z}}{\delta Z} > 20$. Here $\overline{Z}$ is the distance from an object to a focal spot in average, and δZ is the offset of an object point from $\overline{Z}$ along the Z-coordinate. In the situation we discuss, δZ is around $100mm$, half the

size of a human heart; and $\overline{Z}$ is equivalent to d, about $500mm$. Then we have $\dfrac{\overline{Z}}{\delta Z} = 5$, which is much smaller than the suggested value of 20. Although this may seem encouraging errors are nevertheless introduced in the result by solving equation (3.1c).

We now analyze the error in the worst-case in above configuration. Figure 3.4 depicts the perspective project and week perspective projection in the worst-case. H' and H are the week perspective projection and full perspective projection of object h on the image plane, respectively. We have

$$H' = \frac{fh}{\overline{Z} - \delta Z} = \frac{1000 \cdot 200}{500 - 100} = 500$$

and

$$H = \frac{fh}{\overline{Z}} = \frac{1000 \cdot 200}{500} = 400 \ .$$

The error for the worst-case is $\left| H' - H \right| / H' = \left| 500 - 400 \right| / 500 = 20\%$.
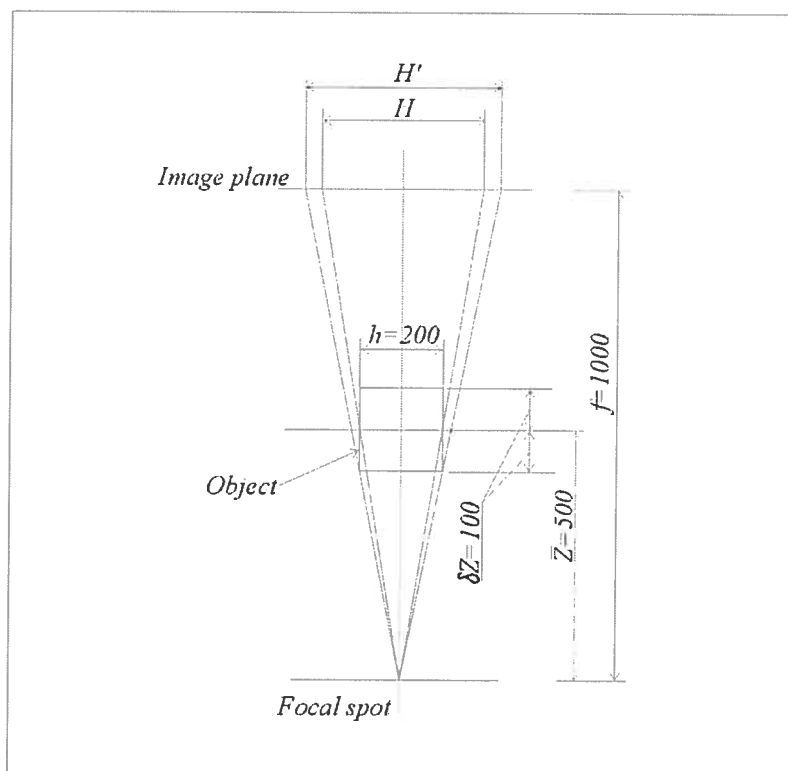

Figure 3.4   The error analysis of the week perspective project vs. the full perspective project.

Furthermore, another error is introduced by using the distance of each two subsequent points to approximate the curve length between the two points. Increasing the frame rate of the angiography image sequence can reduce this kind of error, but would cause an unacceptable increase of the X-ray dose to patients on the other hand.

In the deductive computation, abovementioned errors will accumulate. Even worse, a significant accumulated error may cause the method to become unstable.

To reduce the abovementioned errors, we propose a strategy that uses an iterative process to improve the weak perspective projection model, thus reducing the errors. At the first loop of the iteration, we use weak perspective projection and set $\overline{Z}$ to a given value. In our simulation, we have $\overline{Z} = 500\,mm$. Then in the following loops of the iteration, we use Z-coordinate values in the solution obtained in the previous loop of the iteration to replace the values for $\overline{Z}$. Our simulation results show that the strategy significantly reduces the errors in about ten loops.

3.3.5 Determination of the Location of the Start Point of the Trajectory Curve

In the previous discussion, we have assumed that Z_1, Z-coordinate for the first point of a trajectory curve, is known. However, in clinical situation, it cannot be measured. To solve the problem, we use a strategy that is similar with the one we have used for choosing the right solution from multi-solutions. That is, we compute all solutions using the method given in previous subsections with an assumed Z_1 from 400mm to 600mm by steps of 10mm. Among the solutions obtained, we compute the Euclidean distance between the projection of the points in each solution and the measured images, and pick Z_1 associated with the solution with the smallest Euclidean distance. Then repeat the above computation in a small range around the estimation for Z_1 obtained previously. For instance, if the estimation for Z_1 is 521mm, we compute a more accurate Z_1 and its associated solution by testing Z_1 from 511mm to 531mm by steps of 1mm. Obviously, the computation for Z_1 is rather time consuming. According to our previous analysis in Section 3.3.2, if there are n points where $\phi \leq threshold$, then the right solution has to be chosen from 2^n groups of possible solutions.

3.3.6 A Summary of Proposed Method

In this subsection, we summarize the method that has been described in the previous subsections.

The configuration of a single angiography camera system is shown in Figure 3.1. The object consists of m points corresponding to the trajectory of a catheter tip. Its centre is located near the Z-axis at about the distance of d from the focal spot. The input image sequence consists of m frames, and in each of them there is an image dot that is projected by the corresponding object point. In addition, the length of the curve between each two subsequent points can be measured as Δl_i (the pullback length). The following algorithm can be used for reconstruction of the object based on the above known conditions. The algorithm consists of four procedures.

Procedure Find_Solutions:

This procedure finds multi-solutions, which is returned as a *ListSolutions*.

Input: *numStartPoint, listSolutions=null, Z, solutionSet.*

Note that Z represents the estimated values for Z_1 through Z_m.

Return: *listSolutions*

1) Copy solution to *solutionSet1* and *solutionSet2*.

2) Compute $X_{numStartPoint}$ and $Y_{numStartPoint}$ through equation (3.2), and add them into *solutionSet1* and *solutionSet2*. Note: *solutionSet1* and *solutionSet2* are used to store the two possible solutions due to the sign in equation (3.1c).

3) $X_{numStartPoint+1}$, and $Y_{numStartPoint+1}$ through equation (3.2).

Compute $\Delta Z_{numStartPoint}$ through equation (3.1c). Then compute to two possible solutions for $Z_{numStartPoint+1}$ though $Z^1_{numStartPoint+1} = Z_{numStartPoint} + \Delta Z_{numStartPoint}$,

and $Z^2_{numStartPoint+1} = Z_{numStartPoint} - \Delta Z_{numStartPoint}$. Add solutions into *solutionSet1* and *solutionSet2*, respectively.

If (*numStartPoint==m*), return solutionSet1 and solutionSet2.

Repeat following steps 4) and 5) until $i \geq m$.

4) Let $\Delta Z_i^{temp} = Z_{i+1} - Z_i$, $\Delta X_i = X_{i+1} - X_i$, and $\Delta Y_i = Y_{i+1} - Y_i$. If $\Delta Z_i^{temp} \leq threshold$, goto step 7).

5) If $(\Delta X_i - \Delta X_{i-1})^2 + (\Delta Y_i - \Delta Y_{i-1})^2 + (\Delta Z_i^{temp} - \Delta Z_{i-1})^2 <$

$$(\Delta X_i - \Delta X_{i-1})^2 + (\Delta Y_i - \Delta Y_{i-1})^2 + (-\Delta Z_i^{temp} - \Delta Z_{i-1})^2,$$

$$Z_{i+1}^1 = Z_i^1 + \Delta Z_i^{temp} \text{ and } Z_{i+1}^2 = Z_i^2 - \Delta Z_i^{temp}$$

else

$$Z_{i+1}^1 = Z_i^1 - \Delta Z_i^{temp} \text{ and } Z_{i+1}^2 = Z_i^2 + \Delta Z_i^{temp}.$$

Compute X_{i+1} and Y_{i+1} through Equations (3.2).

Add solutions into *solutionSet1* and *solutionSet2*, respectively.

6) Return the solutions.

7) Recursively invoke Find_Solutions with *numStartPoint* being set to i and

$$Z_{i+1} = Z_i + \Delta Z_i^{temp}.$$

Add the returned solutions of the above invocation into the *listSolutions*.

Recursively invoke Procedure Find_Solutions with *numStartPoint* being set to

i. and $Z_{i+1} = Z_i - \Delta Z_i^{temp}$.

Add the returned solutions of the above invocation into the *listSolutions*.

Return the *listSolutions*.

When invoking Procedure Find_Solutions, *listSolutions* and *solutionSet* should be provided as empty lists. *numStartPoint* should be set to 1, and Z_1 is initialed to an estimated value.

Procedure Choose_Solution:

This procedure finds the best solution from multi-solutions obtained by the Procedure Find_Solutions.

Input: *listSolutions*

Return: *bestSolution*

For each solution in *listSolutions*

1) Project the solution onto the image plane through equation (3.6).

2) Compute the Euclidian distance between corresponding projected and measured points. The solution with the smallest Euclidian distance is the *bestSolution*.

3) Return *bestSolution*.

Procedure Optimization:

This procedure reduces the error.

Input: All previously obtained solutions

Return: *solution*

1) Set Z_l through Z_m to the estimated values.

Repeat steps 2) through 4) for twelve times

2) Invoke Procedure Find_Solutions.

3) Invoke Procedure Choose_Solution.

4) Modify the values for Z_l through Z_m according to the solution.

5) Return the solution.

Procedure Find_Z1:

This procedure finds the best estimated value for Z_l.

Input: Z_1^{Max} and Z_1^{Min}, which represent upper and lower boundaries of the object along the Z-axis, respectively. S is the step length used for looping.

Return: the best solution and its associated Z_0.

1) For Z_l in $[Z_1^{Min}, Z_1^{Max}]$, by step S

    Invoke Procedure Optimization and put the solutions into a list.

2) For each solution in the list

    Project the solution onto the image plane through equation (3.3).

    Compute the Euclidian distance between corresponding projected and measured points.

3) The solution with the smallest Euclidian distance is the best solution. Return the best solution and its associated Z_0.

3.4 Simulation of Proposed Method in Mathematica

In this section, we present the simulation of the proposed method in Mathematica and the analysis on the simulation results.

In the simulation of the proposed method, we set a camera system with its focal spot located at the origin of the camera reference frame, and image plane at $Z=1000$, *i.e.*,

focal length $f=1000$. Furthermore, we assume, in the object reference frame, a helicoidal trajectory of the catheter tip:

$$\begin{aligned} X &= r\sin(t) \\ Y &= r\cos(t) \\ Z &= h(t-\pi), \end{aligned} \qquad (3.7)$$

where $0 \le t \le 2\pi$. In order to simulate the proposed method with object located at various locations and orientations, we define a transformation from object reference frame to camera reference frame as follows:

$$[X_i, Y_i, Z_i]^T = R \cdot [X_i^m, Y_i^m, Z_i^m]^T + T \qquad (3.8)$$

where T can be represented by T_1, T_2, and T_3, the translations along X, Y and Z axes, as follows:

$$T = [T_1, T_2, T_3]^T, \qquad (3.9)$$

and R can be represented by Euler angles, ϕ_x, ϕ_y, and ϕ_z, the rotation angles about X, Y, and Z axes of the camera reference frame as following equation:

$$R^T = \begin{bmatrix} \cos\phi_z & -\sin\phi_z & 0 \\ \sin\phi_z & \cos\phi_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\phi_y & 0 & \sin\phi_y \\ 0 & 1 & 0 \\ -\sin\phi_y & 0 & \cos\phi_y \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi_x & -\sin\phi_x \\ 0 & \sin\phi_x & \cos\phi_x \end{bmatrix}. \qquad (3.10)$$

Since the image plane is perpendicular to the Z-axis, theoretically, the proposed method will demonstrate the same characteristic with changing ϕ_z. So in the following simulation, we set $T = [0, 0, 500]^T$, and set $\phi_z = 0°$, and assign ϕ_x, $\phi_y = 0°, 60°, 120°$.

In Procedure Find_Solutions of the proposed method, Δl_i, which represents the length of the curve between each two subsequent points, can be computed from the corresponding change of t in equation (3.7):

$$\Delta l = \int_{\Delta t} \sqrt{X'^2 + Y'^2 + Z'^2}\, dt, \qquad (3.11)$$

where, we assume that Δt_i is a constant. From Equation (3.7), we have

$$\begin{aligned} X' &= r\cos(t) \\ Y' &= r\sin(t) \\ Z' &= h. \end{aligned} \qquad (3.12)$$

Plugging equation (3.12) into equation (3.11), we have:

$$\Delta l = \sqrt{r^2 \cos^2(\Delta t) + r^2 \sin^2(\Delta t) + h^2} . \qquad (3.13)$$

Through equation (3.13), we can compute Δl from Δt, which are given in the following simulations.

3.4.1 Simulations based on known Z_I

In this subsection, the simulations are based on the assumption that Z_I, the Z-coordinate of the start point of the trajectory, is known *a priori*. Simulation for determining Z_I using Procedure Find_Z1 of the proposed method is presented in the next subsection.

Table 3.1 shows the simulation results when $r = 100$, $h = 32$ in equation (3.7), and $\Delta t = 0.3$. Among the nine situations tested, the results show that only in the case: $\phi_x = 0°$, $\phi_y = 60°$, $\phi_z = 0°$, the proposed method fails although in the cases: $\phi_x = 120°$, $\phi_y = 60°$, $\phi_z = 0°$ and $\phi_x = 120°$, $\phi_y = 120°$, $\phi_z = 0°$, there are significant errors (almost fails) and in the case: $\phi_x = 0°$, $\phi_y = 120°$, $\phi_z = 0°$, significant errors happen at the end of the trajectory. The average relative error is 8.69%, and the error is mainly contributed by the case of $\phi_x = 120°$, $\phi_y = 60°$, $\phi_z = 0°$. In this simulation, the proposed method almost fails.

Table 3.2 shows the simulation results when $r = 100$, $h = 32$ in equation (3.7), and $\Delta t = 0.2$. Compared with the results shown in Table 3.1, the results have been improved by reducing Δt for the case: $\phi_x = 0°$, $\phi_y = 60°$, $\phi_z = 0°$, though significant errors remain at the end of the trajectory. There is little improvement however for cases: $\phi_x = 120°$, $\phi_y = 60°$, $\phi_z = 0°$ and $\phi_x = 120°$, $\phi_y = 120°$, $\phi_z = 0°$ and little change with case: $\phi_x = 0°$, $\phi_y = 120°$, $\phi_z = 0°$. The average relative error is 6.68%.

Table 3.3 shows the simulation results when $r = 50$, $h = 16$ in equation (3.7), and $\Delta t = 0.25$. The results show improvement by reducing the size of object. This is due to the improvement of the condition of weak perspective projection, the camera model used in the proposed method. The average relative error for this simulation is 5.81%.

Table 3.1 Simulation of the proposed method when $r = 100$, $h = 32$ and $\Delta t = 0.3$.
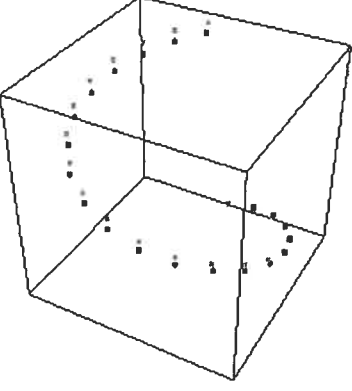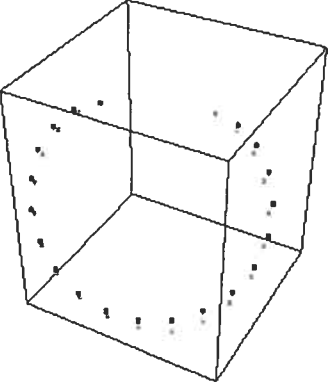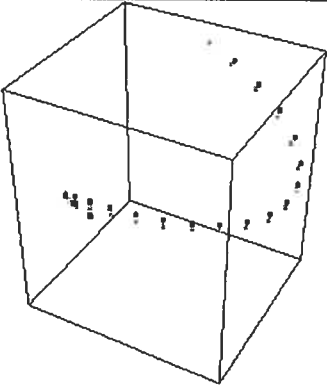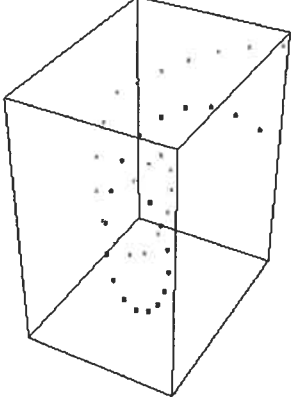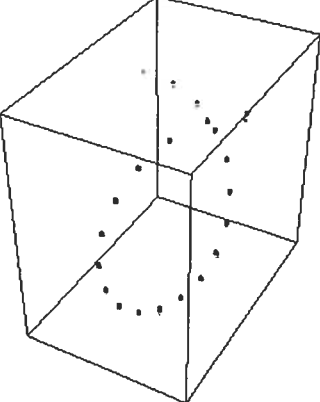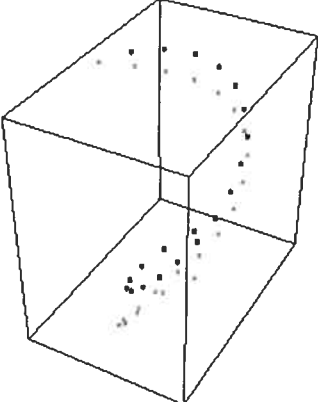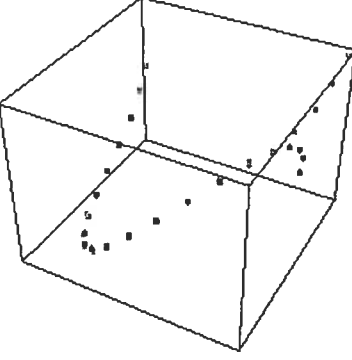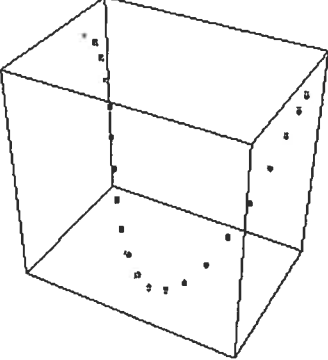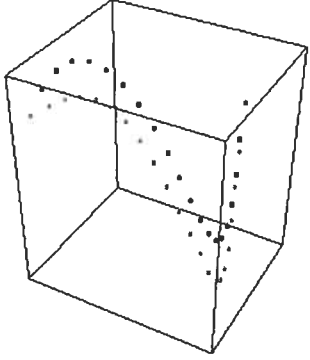
$\phi_x = 0°$, $\phi_y = 0°$, $\phi_z = 0°$ Average error $= 7.21773$ Maximum error $= 12.5507$	$\phi_x = 60°$, $\phi_y = 0°$, $\phi_z = 0°$ Average error $= 7.23549$ Maximum error $= 15.7908$	$\phi_x = 120°$, $\phi_y = 0°$, $\phi_z = 0°$ Average error $= 6.19131$ Maximum error $= 9.78618$
$\phi_x = 0°$, $\phi_y = 60°$, $\phi_z = 0°$ Average error $= 75.1197$ Maximum error $= 111.42$	$\phi_x = 60°$, $\phi_y = 60°$, $\phi_z = 0°$ Average error $= 2.74184$ Maximum error $= 6.62557$	$\phi_x = 120°$, $\phi_y = 60°$, $\phi_z = 0°$ Average error $= 25.0995$ Maximum error $= 41.0242$
$\phi_x = 0°$, $\phi_y = 120°$, $\phi_z = 0°$ Average error $= 14.2298$ Maximum error $= 133.691$	$\phi_x = 60°$, $\phi_y = 120°$, $\phi_z = 0°$ Average error $= 1.69273$ Maximum error $= 3.19318$	$\phi_x = 120°$, $\phi_y = 120°$, $\phi_z = 0°$ Average error $= 42.1034$ Maximum error $= 53.0031$

Table 3.2   Simulation of the proposed method when $r = 100$, $h = 32$ and $\Delta t = 0.2$.
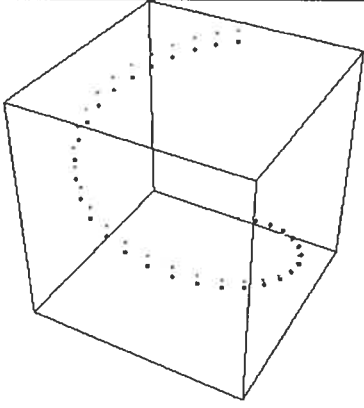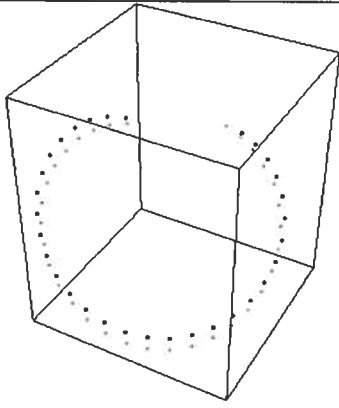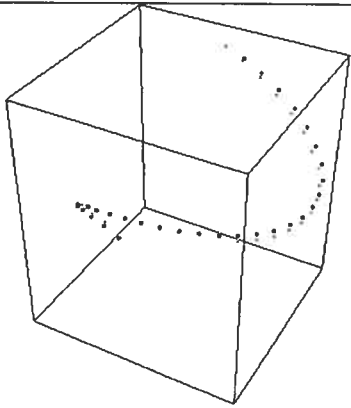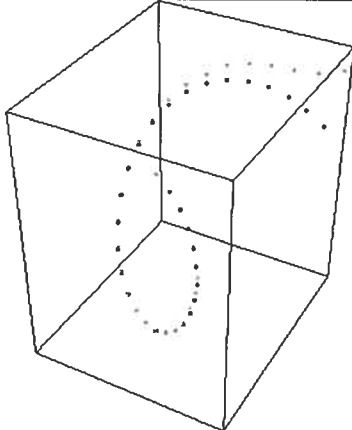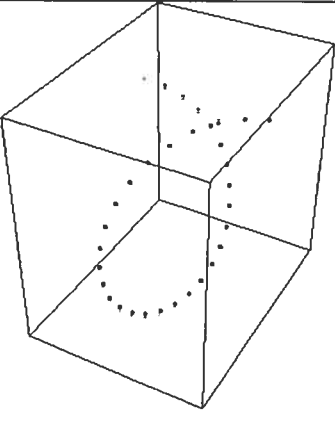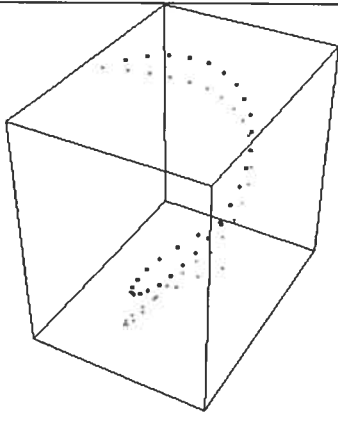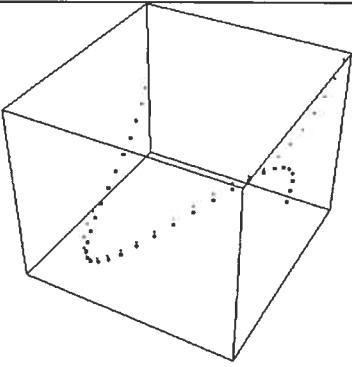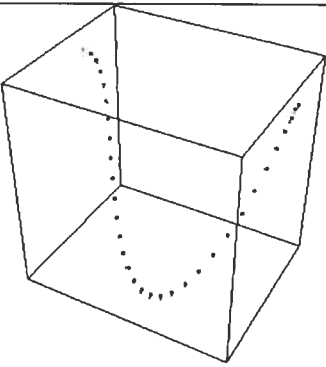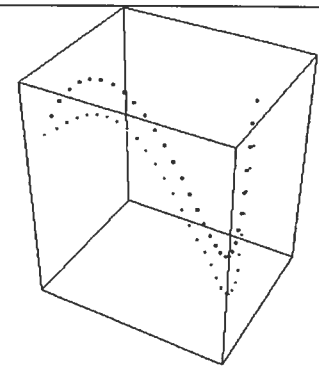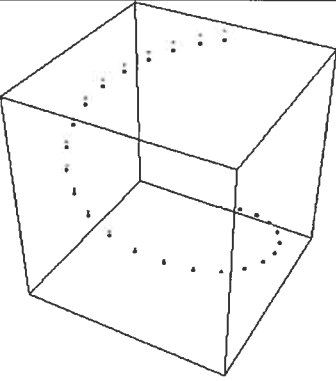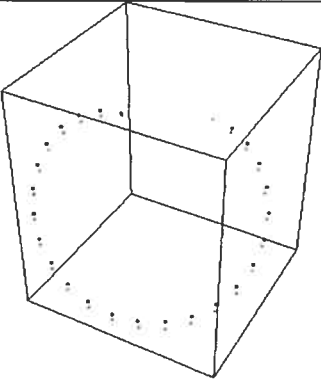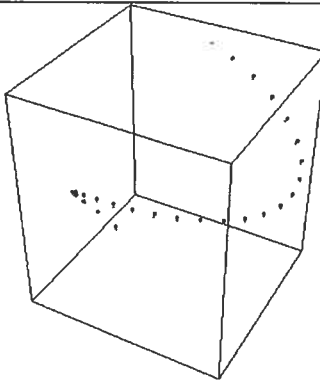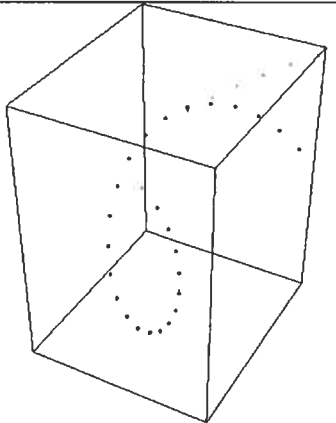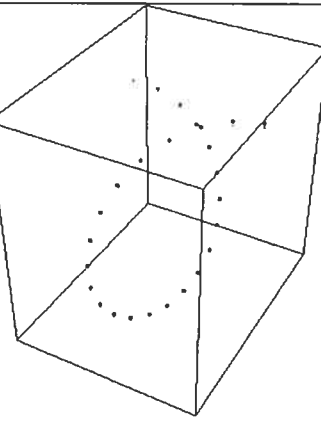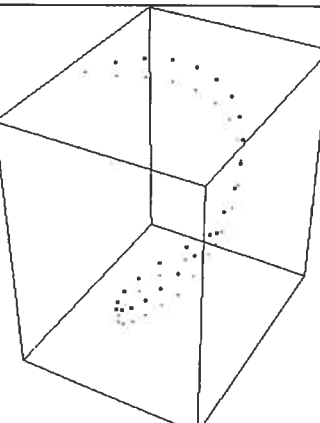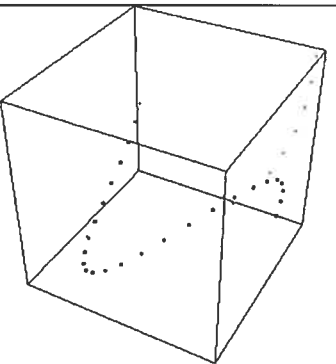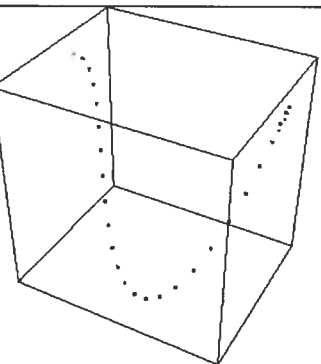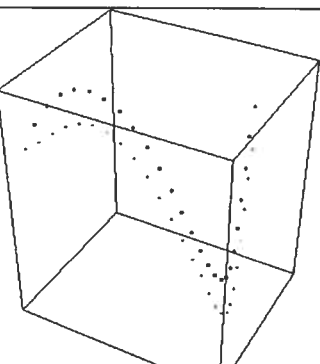

$\phi_x = 0°$, $\phi_y = 0°$, $\phi_z = 0°$ Average error $= 7.84649$ Maximum error $= 12.896$	$\phi_x = 60°$, $\phi_y = 0°$, $\phi_z = 0°$ Average error $= 10.4022$ Maximum error $= 14.9839$	$\phi_x = 120°$, $\phi_y = 0°$, $\phi_z = 0°$ Average error $= 3.70992$ Maximum error $= 7.38958$
$\phi_x = 0°$, $\phi_y = 60°$, $\phi_z = 0°$ Average error $= 7.25877$ Maximum error $= 52.8793$	$\phi_x = 60°$, $\phi_y = 60°$, $\phi_z = 0°$ Average error $= 2.3329$ Maximum error $= 3.44332$	$\phi_x = 120°$, $\phi_y = 60°$, $\phi_z = 0°$ Average error $= 24.9043$ Maximum error $= 40.6813$
$\phi_x = 0°$, $\phi_y = 120°$, $\phi_z = 0°$ Average error $= 20.6158$ Maximum error $= 173.845$	$\phi_x = 60°$, $\phi_y = 120°$, $\phi_z = 0°$ Average error $= 2.01517$ Maximum error $= 3.90757$	$\phi_x = 120°$, $\phi_y = 0°$, $\phi_z = 0°$ Average error $= 41.3752$ Maximum error $= 48.0218$

Table 3.3   Simulation of the proposed method when $r = 50$, $h = 16$ and $\Delta t = 0.25$.

$\phi_x = 0°$, $\phi_y = 0°$, $\phi_z = 0°$ Average error = 2.03839 Maximum error = 4.92998	$\phi_x = 60°$, $\phi_y = 0°$, $\phi_z = 0°$ Average error = 3.32538 Maximum error = 4.62476	$\phi_x = 120°$, $\phi_y = 0°$, $\phi_z = 0°$ Average error = 1.67105 Maximum error = 2.36308
$\phi_x = 0°$, $\phi_y = 60°$, $\phi_z = 0°$ Average error = 5.01847 Maximum error = 48.8279	$\phi_x = 60°$, $\phi_y = 60°$, $\phi_z = 0°$ Average error = 0.830369 Maximum error = 2.06972	$\phi_x = 120°$, $\phi_y = 60°$, $\phi_z = 0°$ Average error = 9.54844 Maximum error = 13.4792
$\phi_x = 0°$, $\phi_y = 120°$, $\phi_z = 0°$ Average error = 11.1837 Maximum error = 99.5178	$\phi_x = 60°$, $\phi_y = 120°$, $\phi_z = 0°$ Average error = 0.811139 Maximum error = 3.6921	$\phi_x = 120°$, $\phi_y = 0°$, $\phi_z = 0°$ Average error = 18.026 Maximum error = 20.6744

3.4.2 A Simulation for the determination of Z_l

In this subsection, we use Procedure Find_Z1 of the proposed method to determine Z_l. The simulation is based on the case: $r = 50$, $h = 16$, and $\Delta t = 0.25$. The simulation is accomplished in on two iterations. In the first iteration, we set $Z_1^{Min} = 450$, $Z_1^{Max} = 550$ and $S = 10$. The results in Table 3.4 show that Z_l is roughly estimated between 520 and 540. In the second iteration, we set $Z_1^{Min} = 520$, $Z_1^{Max} = 540$ and $S = 2$. The results in Table 3.5 show that Z_l is estimated as 530. This is close to the actual value for Z_l, $Z_l = 525.133$. The error item in the tables means the average difference between the images of the projection of solution and the measured images. The error values in bold face are the minimum value.

Table 3.4   In the first iteration, Z_l is roughly estimated between 520 and 540.

Z_1	450	460	470	480	490	500	510	520	530	540	550
error	0.547	1.377	1.118	0.970	0.868	0.833	4.471	0.633	**0.366**	0.598	0.557

Table 3.5   In the second iteration, Z_l is estimated as 530.

Z_1	520	522	524	526	528	530	532	534	536	538	540
error	0.633	0.606	0.432	0.398	0.376	**0.366**	0.411	0.616	0.602	0.596	0.598

As mentioned above, the computation of Z_l is rather time consuming due to its exhaustive search mechanism.

3.5 Summary

In this chapter, we have proposed a new method of pose estimation with image sequence taken from a single view. The proposed method has overcome the limitation existing in the method of pose estimation by single projection, employed in single angiography.

Simulation results for the proposed method have been presented in this chapter. It has been shown that for most cases, the proposed method has achieved acceptable results. In the view of these experiments, reducing the object size, which consequently improves the camera model of weak perspective projection, can improve the results. Reducing the object size is equivalent to increasing focal length, according to the camera model. Thus, in real clinical situations, a large focal length can be configured in angiography equipment in order to achieve improved results.

Chapter 4

Application of the Proposed Method to Synthetic Data

4.1 Introduction

In Chapter 3, we have presented our simulation of the proposed method on given curves with Mathematica. In the simulation, we did not consider the errors that might be introduced in measurement of image points, and assumed that the catheter trajectory was rigid. However, in clinical practice, these conditions are not necessarily met. These two unconsidered characteristics may affect the results of the proposed method. In this chapter, we will present the application of the proposed method to real data obtained from a phantom experiment, which simulates an IVUS intervention. The image sequence of the phantom experiment and the associated documents [15] used in our application were provided by the Montreal Heart Institute.

This chapter is organized as follows. In Section 2, we will describe the phantom experiment that was carried out at the Montreal Heart Institute, as well as an image sequence obtained from the experiment. In Section 3, we will test the proposed method on the image sequence of the phantom experiment and discuss the results. In the last section, we will summarize this chapter.

4.2 Phantom Experiment

A phantom experiment has been carried out at the Montreal Heart Institute, which has been designed for investigation of single projection angiography in 3-D IVUS. In the following, we describe the phantom experiment.
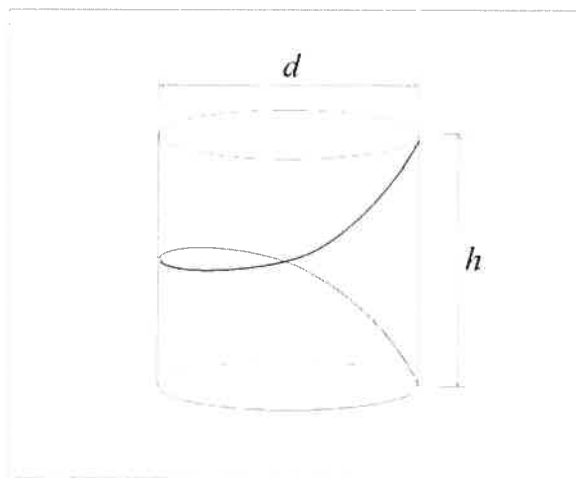
Figure 4.1 The phantom is built with a 1000*ml* beaker with a pipe spiraling on it about 360°. The measurement for *h* and *d* are: *h*= 125*mm* and *d*=109*mm*.
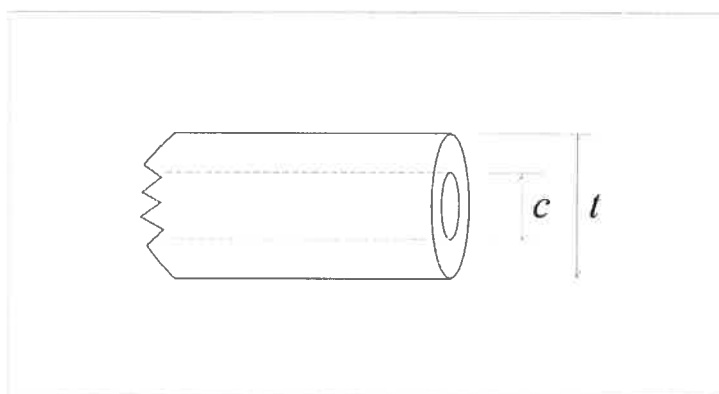
Figure 4.2 The measurements of the pipe: *c* = 3.176*mm* and *t* = 6.35*mm*.

The phantom, as shown in Figure 4.1, is built with a 1000*ml* beaker with a pipe spiraling on it about 370°. In Figure 4.1, *h* and *d* are measured with respect to the centre of the pipe. The measurements of the pipe are shown as Figure 4.2. The catheter used in the phantom experiment is a JOVUS Avanar F/X, as shown in Figure 4.3. It is an IVUS imaging catheter product from JOMED. The diameter of the catheter is 2.9*French* (equals to 1*mm*), which is smaller than the inner diameter of the pipe (3.176*mm*). Therefore while the catheter moves along the spiral pipe, a movement along the tube diameter direction may occur to the catheter. This can be observed in the image sequence. By this effect, the

trajectory of catheter is not a rigid curve.
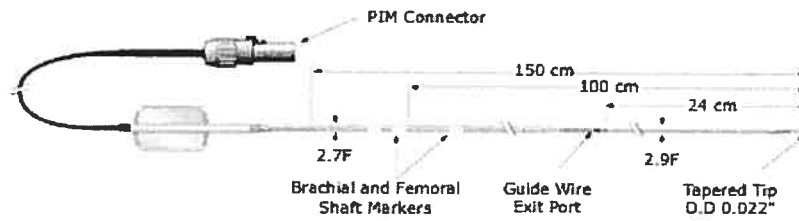
Figure 4.3   JOVUS Avanar F/X, an IVUS imaging catheter.

In the phantom experiment, a fluoroscopic camera system has been configured such that the focal length is $f = 945mm$, while the distance from the center of phantom to the focal spot is about $876mm$. The camera system is shown in Figure 4.4. The pullback of the catheter is performed by a Track Back II system, a catheter pullback device from JOMED. The velocity of pullback of the catheter is about $0.5mm/s$, which is measured with error according to [15], but no range of error is given. A sequence of fluoroscopic images has been captured by the fluoroscopic camera system during the pullback of the catheter. The sequence consists of 1109 frames of 1024×1024 pixels fluoroscopic images taken at a rate of 4 frames per second (fps), which is stored in 5 DICOM image files (DICOM is a standardized format for medical images). One frame of the image sequence is shown in Figure 4.5. Unfortunately, there is a gap with an unknown number of frames (about 200 frames), between frame 220 and frame 221 in the sequence. Therefore, only the frames from frame 221 to frame 1109, *i.e.*, 889 frames, can be used in our application. The trajectory to be reconstructed is about 67.9% of the whole trajectory.
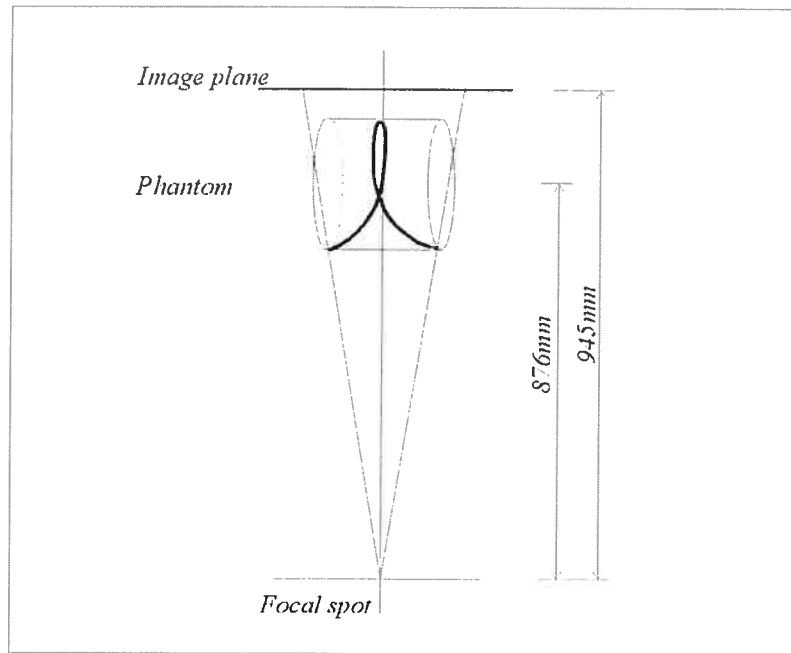
Figure 4.4   The configuration of the fluoroscopic camera system in the phantom experiment.
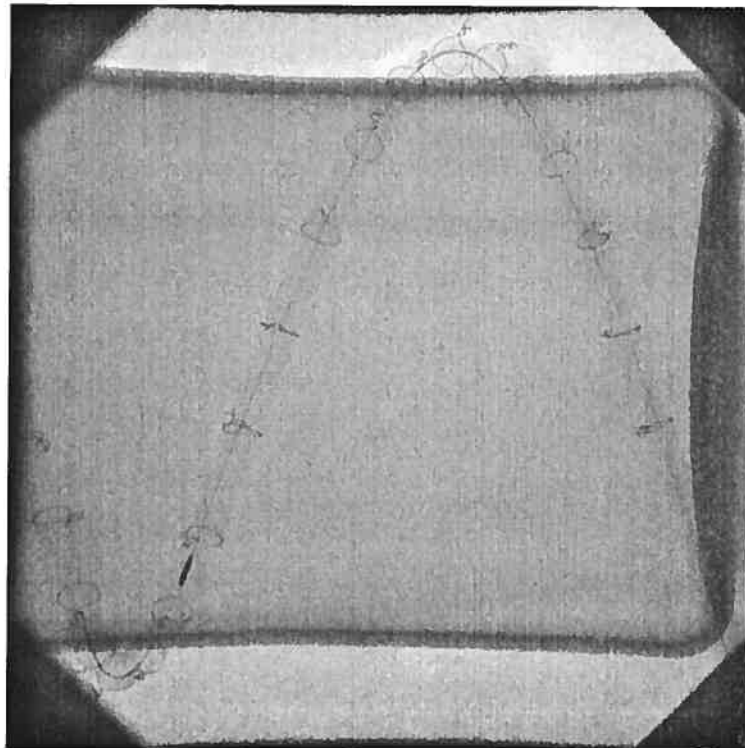
Figure 4.5   A frame of the fluoroscopic image sequence. The transducer along the catheter is visible in the lower left quadrant of the image. The circular object appearing along the tube are

simply twist tights attached at regular intervals along the pipe for reference purpose.

4.3 Application of the Proposed Method on the Phantom Experiment Data

The original image sequence consists of 889 frames with a rate of 4*fps*. It records the trajectory of the catheter that spirals around the beaker for about 250°. In our application, two image sequences with a specific number of frames are reconstructed by picking frames from the original image sequence with a certain number of frames interval. One reconstructed image sequence, called the Image Sequence I, has 20 frames, which are picked at an interval of 45 frames from the original sequence as shown in the first row of Table 4.1. Another reconstructed sequence, called the Image Sequence II, has 18 frames, which are picked at an interval of 52 frames as shown in the first row of Table 4.2.

The coordinates of the catheter (actually the IVUS transducer) tip on each frame is read in pixel unit manually using Osiris, a software system of medical image manipulation from Digital Imaging Unit, Center of Medical Informatics, University Hospital of Geneva. The coordinates of the images in *mm* unit can be calculated by multiplying the coordinates by the width of one pixel in the image plane, which is provided in [15] as 0.1465*mm*.

In the following subsections, we will present the application of the proposed method to the Image Sequence I and the Image Sequence II, respectively. These tests are carried out with Mathematica.

4.3.1 Application of the Proposed Method on Image Sequence I

In Table 4.1, the first row, named Frame, shows the frame numbers taken from the original image sequence. The next two rows, named x(*pix*) and y(*pix*), indicate the x and y-coordinates of the projection of the catheter tip measured in pixel unit. The fifth row and sixth row are the calculated values for above x and y-coordinates in the new coordinate system, in which the unit of *mm* is used and the origin is located at the centre of the image plane. The length of pullback of the catheter between each two subsequent frames is $0.05 \times 45/4 = 0.5625$ (*inch*).

Since the Z-coordinate of the start point of the trajectory of the catheter (denoted

as Z_l), which is associated with the image point in frame 221, is unknown, we use Procedure Find_Z1 presented in section 3.3.5, and assumer that Z_l is between 841*mm* and 906*mm*. Table 4.2 gives the data of the trial of Procedure Find_Z1 in the range above by step length of 5*mm*. From the computed results, it is obvious that the estimated Z_l is 846*mm,* shown in bold face. Figure 4.6 shows the reconstructed trajectory of the catheter. The last three rows in Table 4.1 give the *X, Y,* and *Z*-coordinates of the corresponding points in the reconstructed trajectory.

Table 4.1   The Image Sequence I, which is obtained from the original sequence with an interval of 45 frames.

Frame	221	266	311	356	401	446	491	536	581	626
x(*pix*)	215	244	277	310	348	393	430	471	514	556
y(*pix*)	850	780	695	603	497	404	320	229	155	102
x(*mm*)	-43.51	-39.26	-34.42	-29.59	-24.02	-17.43	-12.01	-6.006	0.239	6.446
y(*mm*)	49.51	39.26	26.81	13.33	-2.197	-15.82	-28.12	-41.45	-52.30	-60.06
X(*mm*)	-38.92	-34.96	-30.42	-25.98	-21.09	-15.38	-10.68	-5.370	0.264	5.876
Y(*mm*)	44.32	34.96	23.69	11.70	-1.92	-13.96	-25.01	-37.07	-47.16	-54.76
Z(*mm*)	846.0	836.2	828.7	822.3	822.3	827.5	835.2	840.7	849.1	859.9

Continuation of Table 4.1.

Frame	671	716	761	806	851	896	941	986	1031	1076
x(*pix*)	601	637	670	712	747	786	817	857	892	929
y(*pix*)	74	71	96	139	203	290	380	475	570	675
x(*mm*)	13.03	18.31	23.14	29.30	34.42	40.14	44.68	50.54	55.67	61.09
y(*mm*)	-64.16	-64.60	-60.94	-54.64	-45.26	-32.52	-19.33	-5.420	8.497	23.87
X(*mm*)	12.03	17.12	21.92	28.06	33.29	39.01	43.63	49.35	54.34	59.63
Y(*mm*)	-59.21	-60.41	-57.72	-52.34	-43.78	-31.61	-18.88	-5.293	8.294	23.31
Z(*mm*)	872.0	885.2	898.4	910.2	920.3	925.2	929.7	929.7	929.7	929.7

Table 4.2   The trail of Procedure Find_Z1 in the range of 838*mm* to 908*mm*.

Z_0	841	846	851	856	861	866	871	876	881	886	891	896	901	906
Err	1.76	**1.75**	1.79	1.87	1.98	2.13	2.31	2.52	2.74	5.05	4.49	2.89	2.50	2.17

Unfortunately, there is no accurate and precise data available that represents the trajectory of the catheter, with which we could compare our application results. Therefore, we are not able to estimate directly the errors of the results. However we can

assess indirectly our results by fitting an arbitrary spiral through our data set and look at the parameters h and d as well as the residual error. As shown in Figure 4.6, a spiral trajectory has been reconstructed. The average diameter of the constructed spiral trajectory is estimated as 108.3mm, which is close to 109mm, the known diameter of spiral pipe. The error is 0.642%. The height of the reconstructed spiral is estimated as 84.61mm. In our application, 68.4% of the spiral trajectory has been reconstructed. The height of h for the reconstructed of trajectory is 125 × 889 / 1299 = 85.54mm. The error is 1.08%. The residual error can be computed as 1.25%. It is observable from Figure 4.6 (d) that the reconstructed trajectory is slightly off the spiral curve at the last two points.

4.3.2 Application of the Proposed Method on the Image Sequence II

Table 4.3 is similar to Table 4.1 The length of pullback of the catheter between each two sequent frames is 0.05 × 52 / 4 = 0.650 (*inch*).

Again we use Procedure Find_Z1 presented in Section 3.3.5, and assume that Z_l is between 841mm and 906mm. Table 4.4 gives the data of the trial of Procedure Find_Z1 in the range above by step length of 5mm. From the computed results, Z_l is estimated to be 871mm, shown in bold face. Figure 4.7 shows the reconstructed trajectory of the catheter. The last three rows in Table 4.3 give the X, Y, and Z-coordinates of the corresponding points in the reconstructed trajectory.

The results are similar to the results obtained for the image sequence I. Again some observable errors happen at the last few points in the reconstructed trajectory.

Comparing the results for Z_l in Subsection 4.21 and the results for Z_1 in subsection 4.2.2, we can find a significant difference between the results that is produced by Procedure Find_Z1 of the proposed method and, it consequently will introduce differences in the Z-coordinates of the solutions.

Similar with the reconstruction of the catheter trajectory by the Image Sequence I presented in subsection 4.2.1, we can assess indirectly our results by fitting an arbitrary spiral through our data set and look at the parameters h and d as well as the residual error. As shown in Figure 4.6, a spiral trajectory has been reconstructed. The average diameter of the constructed spiral trajectory is estimated as 108.1mm, which is close to 109mm, the known diameter of spiral pipe. The error is 0.825%. The height of the reconstructed spiral

is estimated as 84.61mm. In our application, 68.4% spiral trajectory has been reconstructed. The height of h for the reconstructed of trajectory is $125 \times 889 / 1299 = 85.54mm$. The error is 1.08%. The residual error can be computed as 1.36%. It is observable from Figure 4.6 (d) that the reconstructed trajectory is slightly off the spiral curve at the last two points.

Table 4.3   The Image Sequence II, which is obtained from the original sequence with an interval of 52 frames.

Frame	221	273	325	377	429	481	533	585	637	689
X(pix)	215	250	288	330	373	422	469	516	566	616
Y(pix)	853	770	665	551	432	339	235	153	93	70
X(mm)	-43.51	-38.38	-32.81	-26.66	-20.36	-13.18	-6.299	0.5860	7.991	15.23
Y(mm)	49.51	37.79	22.41	5.713	-11.72	-25.34	-40.58	-52.59	-61.38	-64.75
X(mm)	-40.23	-35.03	-29.75	-24.10	-18.40	-12.02	-5.777	0.542	7.417	14.48
Y(mm)	46.04	34.49	20.32	5.164	-10.59	-23.10	-37.21	-48.72	-57.55	-61.57
Z(mm)	871.0	860.4	853.8	850.5	850.5	859.2	865.0	875.0	887.2	901.6

Continuation of Table 4.3.

Frame	741	793	845	897	949	1001	1053	1105
X(pix)	655	702	744	788	825	869	911	950
Y(pix)	84	127	195	295	400	506	622	748
X(mm)	20.95	27.83	33.98	40.43	45.85	52.30	58.45	64.16
Y(mm)	-62.70	-56.42	-46.44	-31.79	-16.40	-0.879	16.11	34.57
X(mm)	20.21	27.21	33.61	40.19	45.75	52.20	58.34	64.05
Y(mm)	-60.51	-55.15	-45.92	-31.60	-16.37	-0.877	16.08	34.51
Z(mm)	917.0	931.0	943.1	948.0	944.9	944.9	944.9	944.9

Table 4.4   The trail of Procedure Find_Z1 in the range of 838mm to 908mm

Z_0	841	846	851	856	861	866	871	876	881	886	891	896	901	906
Error	1.53	1.63	1.82	1.96	1.63	1.75	**1.00**	1.53	2.05	6.27	5.54	4.77	4.36	3.69

4.4 Summary

In this chapter, we have presented the application of the proposed method to the data of a phantom experiment. Firstly, we have reconstructed two image sequences from the original image sequence obtained from a phantom experiment. Then we have estimated the residual errors of geometrical size of reconstructed catheter trajectories

from two image sequences respectively as 1.25% and 1.36%.

(a)

(b)

(c)

(d)

Figure 4.6   The reconstructed trajectory of the catheter from the Image Sequence I. (a) top-front-right view (b) right view, (c) front view, (d) bottom view.
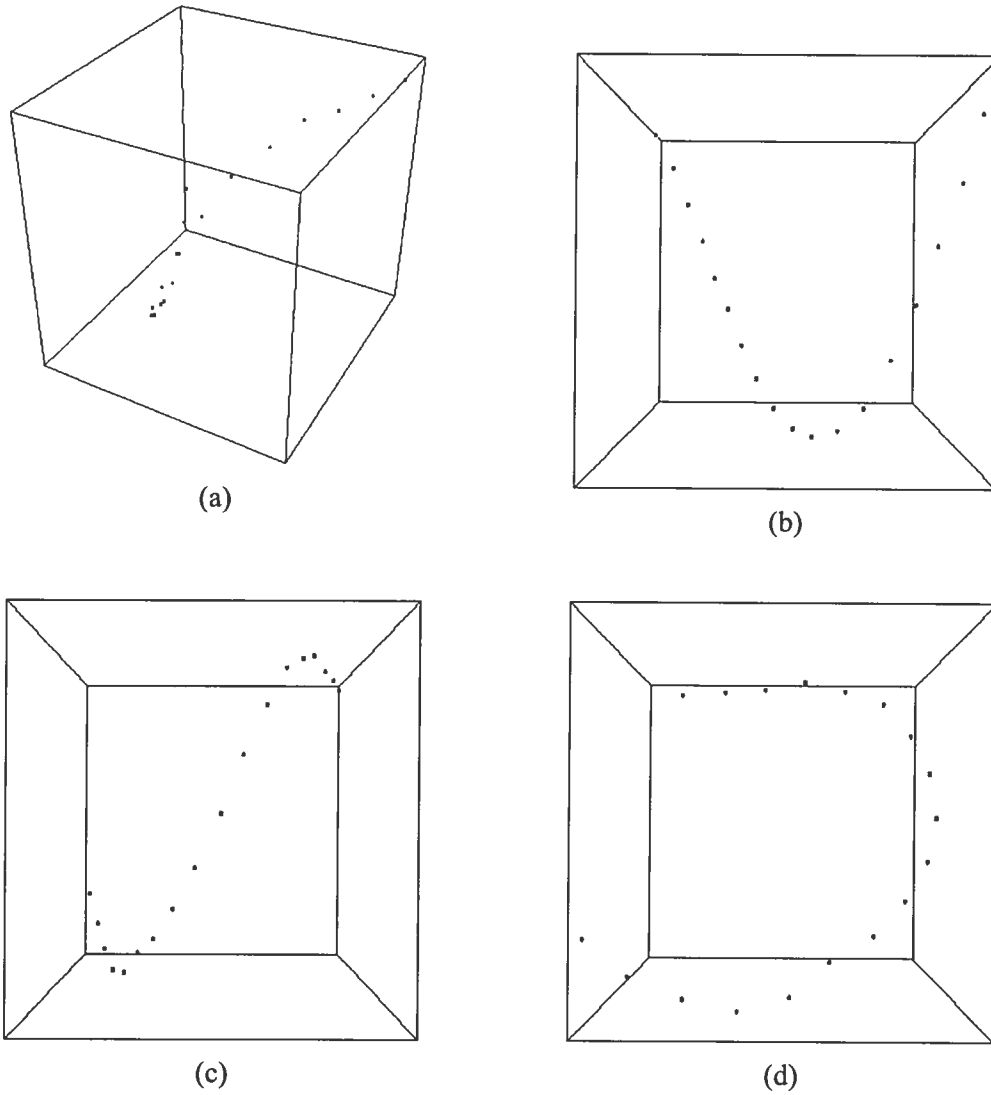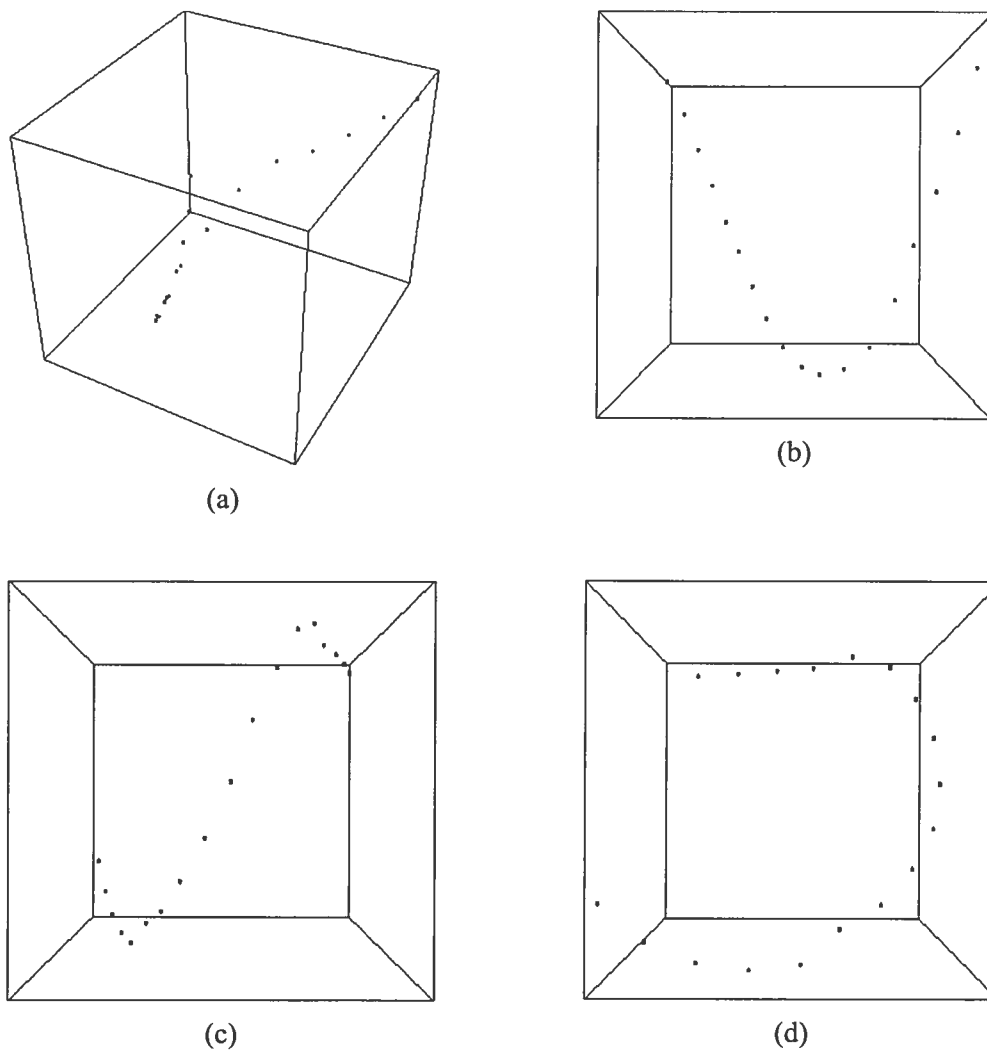
Figure 4.7 The reconstructed trajectory of catheter from the Image Sequence II. (a) top-front-right view (b) right view, (c) front view, (d) bottom view.

Chapter 5

Conclusion

5.1 Conclusion

This thesis research has investigated the problem of pose estimation from single projection in targeting the reconstruction of coronary arteries in 3-D. The research can be divided into two parts. In the first part of the research, we have studied existing methods of pose estimation by a still image from single projection published in the computer vision community and the biological engineering community. From these methods, we have chosen three typical methods to review and simulate with Mathematica in this thesis. From our research on the existing methods and our simulation results, we can draw the following conclusions: 1) All the existing methods require a previous knowledge of the 3-D configuration of the object. In the application of reconstructing the coronary arteries in 3-D, the requirement by the methods means that additional CT or MR images are involved, thus the clinical costs in terms of operation and equipment are increased. 2) There has been little research that toggles on pose estimation by image sequence from single projection. 3) According to the simulation results in Mathematica, the method of POSIT is the best method in terms of accuracy, efficiency, and stability among the reviewed methods.

In the second part of our research, we have investigated the possibility of achieving pose estimation from single projection without a previous knowledge of the 3-D configuration of the object. We have found that an image sequence from a single projection, which is easy to be obtained during an IVUS intervention, provides information that makes the pose estimation possible. Instead of the 3-D configuration of the artery, the known pullback distance of the catheter during the intervention is used. Furthermore, we have carried out simulations of the proposed method on given spiral

curves in Mathematica, and tested the proposed method with real data obtained from a phantom experiment. The following conclusions have been drawn from our research results with the new proposed method. 1) It is possible to carry out pose estimation during IVUS intervention from a single view X-ray image sequence without previously knowing the 3-D configuration of the object. 2) The proposed method overcomes the drawback of the other pose estimation methods by eliminating the requirement of previous knowledge of the 3-D configuration of the object. 3) In practical clinic, an X-ray image sequence from a single projection is easy to be obtained. 4) There still exist considerable errors in the construction of trajectory by the proposed method, compared with the reviewed methods of pose estimation. However, in actual clinical practice, the physician has absolutely no 3-D information and usually assumes that the artery (or object) is a straight tube, which is obviously untrue for coronary arteries (the physician cannot in clinical practice rely on the complex "laboratory" set-up proposed by other methodologies). Therefore, even the slightest bit of 3-D information is appreciated because vessel curvature affects the haemodynamics of blood flow in the arteries and the possible formation of clots and arteriosclerosis. From this point of view, the proposed method might provide useful information that meets the need of physicians in clinics.

5.2 Summary of Contributions

Though a variety of papers have been published on reconstruction of coronary arteries with pose estimation from a still image, there has still been a lack of research on the problem of reconstruction of coronary arteries from an image sequence. Our research has explored the possibility of solving this problem in the particular case of IVUS intervention and found a solution for the problem. The proposed method has many advantages over the other methods, among which its simplicity, much easier setup in the clinical environment and its lower cost. We thus believe that it might lead to a valuable tool in the future.

5.3 Future Research

The proposed method could be improved in following ways. 1) A more accurate result could be achieved by using NURBS instead of the straight lines we have used to

approximate the curved trajectory when calculating the length of curve. 2) A better condition for weak perspective projection can be obtained by adjusting the set-up of the camera system by increasing the distance between the focal spot and the image place. This could consequently improve the results of the proposed method in terms of stability and accuracy.

Since pose estimation from an image sequence has significant advantages over pose estimation from a still image in reconstruction of coronary arteries, it is worthy of exploration by researchers in medical imaging but also in other areas of computer vision in general. We believe according to our research that more solutions could be found regarding such problems.

References

[1]   Kenneth R. Hoffmann, Charles E. Metz, and Yang Chen, Determination of 3-D Imaging Geometry and Object Configurations from Two Biplane Views: An enhancement of the Mets-Fencil Technique, Med. Phys. 22(8), August 1995.

[2]   C. E. Metz, and L. E. Fencil, Determination of three-dimensional structure in biplane radiography without prior knowledge of the relationship between the two views, Med. Phys. 16, 45-51 (1989).

[3]   Emanuele Trucco and Alessandro Verri, *Introductory Techniques for 3-D Computer Vision*, Prentice Hall.

[4]   Kenneth R. Hoffman and Jacqueline Esthappan, Determination of Three-Dimensional Positions of Known Sparse Objects for a Single Projection, Med. Phys. 24 (4), April 1997.

[5]   Jacqueline Esthappan, Chapter 2 of Determination of Three-Dimensional Information by Use of A Three-Dimensional/Tow-Dimensional Matching Technique, A Ph. D. Dissertation, The University of Chicago.

[6]   Daniel F. DeMenthon and Larry S. Davis, Model-Based Object Pose in 25 Lines of Code, European Conference on Computer Vision, 1995.

[7]   Gene H. Golub and Charles F. Van Loan, *Matrix Computations,* The Johns Hopkins University Press, 1983.

[8]   Peter H. Schonemann, A Generalized Solution of the Orthogonal Procrustes Problem, Psychometrika 31, pp1-10, 1966.

[9]   J. C. Gower, Orthogonal and Projection Procrustes Analysis, *In Recent Advances in Descriptive multivariate analysis*, ed W.J. Krzanowski. Oxford University Press, Oxford. 113-134, 1995.

[10] David G. Lowe, Three-Dimensional Object Recognition from Single Two-Dimensional Images, Artificial Intelligence, 31, 3 (March 1987), pp 355-395.

[11] Jacqueline Esthappan and Kenneth R. Hoffmann, 3-D orientations of catheters from single projections. *Computer Aided Diagnosis in Medical Imaging*. K. Doi, H. MacMahon, ML Giger, KR Hoffmann, eds. (Elsevier Science, B. V., Amsterdam, 1999), pp. 357-61.

[12] Martin A. Fischler and Robert C. Bolles, Ramdom Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, Communications of the ACM, Vol 24 No.6, June 1981.

[13] K. Doi, H MacMahon, M.L. Giger, K.R. Hoffmann (eds), A Comprehensive Method for Geometrically Correct 3-D Reconstruction of Coronary Arteries by Fusion of Intravascular Ultrasound and Biplane Angiography, First International Workshop on Computer-Aided Diagnosis, Sept 20-23, 1998.

[14] Radu Horaud, Fadi Dornaika, Bart Lamiroy, and Stéphane Christy, Object Pose: The Link between Weak Perspective, Paraperspective, and Full Perspective, International Journal of Computer Vision, Vol 22, No. 2, 1997.

[15] Denis Sherknies, a Description of Images IVUS-RAD, http://www.iro.umontreal.ca/~sherknie/IVUS_RAD/description.pdf