2010

DESIGN OF A USER INTERFACE FOR THE ANALYSIS OF MULTI-MODAL IMAGE REGISTRATION

Jean-Michel Filion

DESIGN OF A USER INTERFACE FOR THE ANALYSIS OF MULTI-MODAL IMAGE REGISTRATION

(Spine title: User Interface for Multi-Modal Image Registration)

(Thesis format: Monograph)

By

Jean-Michel Filion

Graduate Program in Engineering Science
Department of Electrical and Computer Engineering

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Engineering Science

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

THE UNIVERSITY OF WESTERN ONTARIO
SCHOOL OF GRADUATE AND POSTDOCTORAL STUDIES

CERTIFICATE OF EXAMINATION

Supervisor	Examiners
Dr. Roy Eagleson	Dr. Luiz Capretz
	Dr. Ken McIsaac
	Dr. Mark Daley

The thesis by

Jean-Michel Filion

entitled:

DESIGN OF A USER INTERFACE FOR THE ANALYSIS OF MULTI-MODAL IMAGE REGISTRATION

is accepted in partial fulfillment of the
requirements for the degree of
Master of Engineering Science

Date_____

_____
Chair of the Thesis Examination Board

Abstract

Image registration is the process of spatially aligning two or more images of a scene into a common coordinate system. Research in image registration has yielded a number of rigid and non-rigid image registration methods capable of registering images of a scene between modalities. In addition, techniques of information visualization have been applied to medical image registration research to produce an atlas based image registration method. This method is capable of registration medical images of a same modality between subjects for comparative studies.

This thesis aims to extend research in image registration by adding to it the visual encoding of time. The visual encoding of time furthers image registration research by enabling the simultaneous analysis of the spatial and temporal relationships that exist between images. The benefit of registering images with respect to both space and time is shown through the development of a software application capable of presenting a time-space narrative of x-ray images representing a patient's medical history. This time-space narrative is assembled by performing rigid atlas based image registration on a set of x-ray images and by visually encoding their timestamps to form of an interactive timeline. The atlas based image registration method was selected to ensure that images can be registered to a common coordinate system in cases where images do not overlap. Rigid image registration was assumed to be sufficient to provide the desired visual result.

Subsequent to its implementation, an analysis of the measured uncertainty of the image registration method was performed. The error in manual point pair correspondence selection was measured at more than +/- 1.08 pixels under ideal conditions and a method to calculate the unique standard error of each image registration was presented.

Keywords

Computer Vision, Image Registration, Visualization, Representation, Information Visualization, Scientific Visualization, Visual Analytics

Acknowledgements

I would like to thank Dr. Roy Eagleson from The University of Western Ontario for his encouragement and support. I would also like to thank my wife, Karen, for her patience and David Allison for his help.

Table of Contents

CERTIFICATE OF EXAMINATION ii
Abstract iii
Acknowledgements v
List of Tables viii
List of Figures ix
1 Introduction 1
1.1 Motivation 2
1.2 Thesis Goals and Scope 3
1.4 The Organization of the Thesis 4
2 Introduction to Visual Analytics 6
2.1 Information Visualization 6
2.1.1 Visual Representations 6
2.1.2 Interaction 9
2.2 Scientific Visualization 10
2.3 Visual Analytics 11
2.3.1 The Appropriateness Principle 11
2.3.2 The Naturalness Principle 12
2.3.3 The Matching Principle 12
2.3.4 The Principle of Congruence 12
2.3.5 The Principle of Apprehension 12
2.4 Summary 13
3 Literature Review 14
3.1 Panoramic Image Registration 14
3.2 Medical Image Registration 19
3.3 Atlas Based Medical Image Registration 23
3.4 Summary 26
4 An Interactive Time-Space Narratives of Imagery 27
4.1 Data 27
4.2 Visual Representation 29
4.2.1 Representing Space 29
4.2.2 Representing Time 32
4.3 Interaction 32

5 Method 34
5.1 Geometric Alignment 34
5.1.1 Computing Image Registrations Algebraically 36
5.1.2 The Least Squares Problem 36
5.1.3 Eigenvectors, Eigenvalues, and Solving the Least Squares Approximation 37
5.1.4 Singular Value Decomposition and the Least Squares Approximation 38
5.2 Classes of Transformations 39
5.3 Estimating Transformations 42
5.3.2 Estimating Translations 43
5.3.3 Estimating Similarity Transformations 44
5.3.4 Estimating Euclidean Transformations 47
5.3.5 Estimating Affine Transformations 49
5.3.6 Estimating Projective Transformations 52
5.4 Reconstructing the Timeline 56
5.5 Summary 57
6 Analysis 58
6.1 Error in Point Pair Correspondence Selection 58
6.2 Error in Homography Parameter Estimation 61
6.3 Error Propagation 62
6.4 Perturbation Analysis 64
6.4.1 Perturbation Analysis of a Rigid Image Registration 65
6.4.2 Perturbation Analysis of a Non-Rigid Image Registration 67
6.5 Summary 68
7 Conclusions and Future Work 69
7.1 Conclusions 69
7.2 Future Work 72
Bibliography 73
VITA 76

List of Tables

5.1 Number of point pair correspondences versus permissible transformations 42

6.1 Measured disparity based on 100 synthetic feature selections 60

6.2 Results from registration error analysis 63

6.3 Measured cumulative error in image registration 64

6.4 Number of feature pairs vs spectral matrix norm and condition number 66

6.5 Number of feature pairs vs spectral matrix norm and condition number 67

List of Figures

1.1 Image Registration Example 2

2.1 London Underground also known as Harry Beck's map [26] 7

2.2 Napoleon's advance on Moscow, also known as Minard's map [26] 8

2.3 Brown et al.'s panoramic image registration results without blending [3] 10

3.1 Tessellated spherical panorama constructed from 54 images [29] 16

3.2 Modern automatic panoramic image registration algorithm [3] 18

3.3 Brown et al.'s automatic panorama image registration results [3] 19

3.4 Image registration of CT and SPECT data [33] 21

3.5 One slice of a human head [25] 22

3.6 Average brain atlas of 25 patients [22] 24

4.1 X-ray imagery of a subject's neck 28

4.2 Proposed anatomical atlas 31

5.1 Point pair correspondences between an x-ray image and the atlas 34

5.2 Mapping between pixels related by a plane to plane homography [7] 35

5.3 Geometric representation of the classes of transformations [12] 39

5.4 Translation transformation relating point pair correspondences [12] 40

5.5 Similarity transformation relating point pair correspondences [15] 40

5.6 Euclidean transformation relating point pair correspondences [6] 41

5.7 Affine transformation relating point pair correspondences [12] 41

5.8 Projective transformation relating point pair correspondences [12] 41

5.9 Image registration based on the translation transformation group 44

5.10 Image registration based on the similarity transformation group 47

5.11 Image registration based on Euclidean transformation group 48

5.12 Image registration based on the affine transformation group 52

5.13 Image registration based on projective transformation group 55

5.14 Playback of stitched x-ray imagery in chronological order 56

6.1 Synthetic data generated to estimate feature selection uncertainty 59

6.2 Equation for the standard error in feature selection [31] 61

6.3 Standard error resulting from least squares approximation [31] 61

6.4 Measured propagation of error of a planar scene 64

6.5 Equation for the Spectral Matrix Norm [31] 65

6.6 Rigid image registration performed on two adjacent digital images 66

6.7 Rigid image registration performed on non-rigid MRI images 67

7.1 Software application presenting a time-space narrative of x-ray images 69

List of Acronyms

CT	Computed Tomography
DICOM	Digital Imaging and Communications in Medicine
MR	Magnetic Resonance
PET	Positron Emission Tomography
RANSAC	Random Number Consensus
SIFT	Scale Invariant Feature Transform
SPECT	Single Photon Emission Computed Tomography
SVD	Singular Value Decomposition

Chapter 1

Introduction

Scientists and researchers are often asked to analyse data and to derive from it some form of insight or understanding. This task can be difficult since the cognitive reasoning process is easily overwhelmed by large quantities of data [26]. Consequently, a popular approach to improve data analysis is to represent the data visually. Visual representations augment the cognitive reasoning process with perceptual reasoning which permits the analytical reasoning process to become faster and more focused [5]. This principle applies both to the broad scope of information visualization as well as to the specific scope of scientific visualization, which is generally concerned with the visual representation of physical things or scientific phenomenon [26]. To this end, visualization methods are continuously being developed to analyse scientific data, ranging from brain structure [22] to site surveys [1], in new and innovative ways.

One of today's most popular visualization techniques is that of image registration. Image registration, also known as image stitching, is the process of spatially aligning two or more images of a scene into a common coordinate system by geometrically transforming a target image to become spatially aligned with an unaltered source image [11]. It is often desirable to relate sets of images to a common coordinate system to provide high-level visualizations, to put detailed data into context, and to derive insight or understanding from multiple sets of data. Figure 1.1 shows a common image registration example, that of spatially relating two overlapping aerial photos.

Figure 1.1a: Source Image Figure 1.1b: Target Image Figure 1.1c: Registered Images

Figure 1.1: Image registration example

In Figure 1.1a, an unaltered source image is chosen to represent the desired coordinate system. The target image in Figure 1.1b is then geometrically transformed to fit within the source image's coordinate system. The resulting image registration, shown in Figure 1.1c, depicts the now spatially related information. While this method has been used to increase the resolution of digital maps and satellite photos for decades [20], it can also be used to reduce the cognitive effort required to analyse multiple sets of data. For instance, in Figure 1.1c, the analyst can now understand the geometric relationship that exists between the two original images using minimal cognitive effort. That is to say, the information contained within separate images can be combined to provide additional useful information. This thesis will take advantage of this same phenomenon to reduce the cognitive effort required to obtain insight on a patient's medical history based on spatially and temporally related x-ray images.

1.1 Motivation

As our technology improves, so too does the quantity of data that we have at our disposal. We see the evidence of this in every aspect of our lives; the availability of online papers, social networks, media, the exponential growth and demand of the storage

capacities of our servers and hard drives, and the increasing resolution of our imaging devices. However, our ability to collect data is growing much faster than our ability to analyse it [5]. As such, it becomes necessary to design new systems that augment our ability to navigate expanding information spaces and to derive from them the insights and understanding that we seek.

This need for new systems to analyse large quantities of data is especially true in the medical field. The modernization of our hospitals and health care systems requires the digital collection and availability of our medical records. With this new digital system in place, the available information space containing the medical records of each patient is available and is expanding. Doctors will soon be equipped with a complete digital history of the medical data collected for each patient. The analytical task becomes to derive from all this data some form of insight or understanding into the patient's medical history and to use it to contribute to the current diagnostic process. The method described in this thesis is designed to improve this analytical task by augmenting the analysts' cognitive reasoning process with perceptual reasoning about spatially and temporally encoded medical data. This is be accomplished by visually encoding spatial and temporal data and allowing the user to navigate the information space by playing back the sequence of available and registered images.

1.2 Thesis Goals and Scope

The primary goal of this thesis is to expand upon research in medical image registration by injecting concepts, principles, and techniques borrowed from the fields of information visualization, scientific visualization, and visual analytics. This will be accomplished through the development of a method capable of presenting a scientific visualization, based on principles of information visualization, to enable the visual

analysis of a patient's medical records over time. To this end, two methods borrowed from research literature will be combined; one to geometrically align images to a common coordinates system, and another to visually encode their relationships with respect to time. Consequently, a mechanism to embed medical images into a common time-space coordinate system will be devised. A secondary goal of this thesis is to design an intuitive user interface to navigate the information space with specific focus on augmenting visual analysis. Both of these goals will be described in detail in this thesis and, following their implementation, the quantifiable components of the system will be analysed.

1.4 The Organization of the Thesis

This chapter has provided an introduction to visual analytics through the method of image registration, discussed its implications, and defined the broad scope of the thesis.

- Chapter 2 provides and introduction to information visualization, scientific visualization, and visual analytics through the use of software tools. Its key concepts, components and techniques for representing space and time are discussed in order to better understand what is required when designing the visual analytic tool presented in this thesis.
- Chapter 3 includes a literature review of the methods developed to solve image registration problems. A look at image registration methods in academia with specific focus on medical applications are explored. An area of research literature that can benefit from further exploration is identified, and a visual analytic tool is briefly presented outlining how it can extend current research efforts.

- Chapter 4 will describe the essential use cases and requirements for the software application and introduces the image registration problem as it applies to this thesis.
- Chapter 5 presents an algebraic solution to the image registration problem and discusses its essential components with respect to the thesis implementation.
- Chapter 6 provides a statistical analysis of the measured accuracy of the algorithms implemented in this thesis.
- Chapter 7 presents the conclusions to the thesis and the potential for future work.

Chapter 2

Introduction to Visual Analytics

This chapter serves as an introduction to visual analytics by examining its principles, its components, and its challenges. An overview of information visualization, scientific visualization, and their relationship to visual analysis is presented. The principles examined in this section are exposed in order to better understand what requirements must be solicited to design a visual analytic tool.

2.1 Information Visualization

Information visualization is the use of computer-supported, interactive, visual representations of information to enhance human thinking [26]. This field has emerged from research in human-computer interaction, computer science, computer graphics, and psychology. Today, information visualization is being applied as a critical component of scientific research. Information visualization focuses on the creation of approaches for conveying abstract information in intuitive ways through the use of visual representations and interaction techniques.

2.1.1 Visual Representations

In order to understand how to design and use meaningful visual representations, we must first understand what is meant by the term *visualize*. To visualize something is a purely cognitive activity. It means to form a mental image or a mental model of something [26]. This cognitive activity is not performed by software or by a computer. However, the use of software and computers can help humans perform this activity immensely by providing a tool to present a visual representation of information sets. It is

essential to make the distinction between the term *presentation*, and the term *representation*. In information visualization, the term *presentation* refers to the act of displaying something while the term *representation* refers to the encoding of a set of information. Therefore, a visual representation refers to the visual encoding of data in such a way that its presentation assists the human cognitive activity of forming a mental model or mental image of something. An example of a visual representation is shown in Figure 2.1.

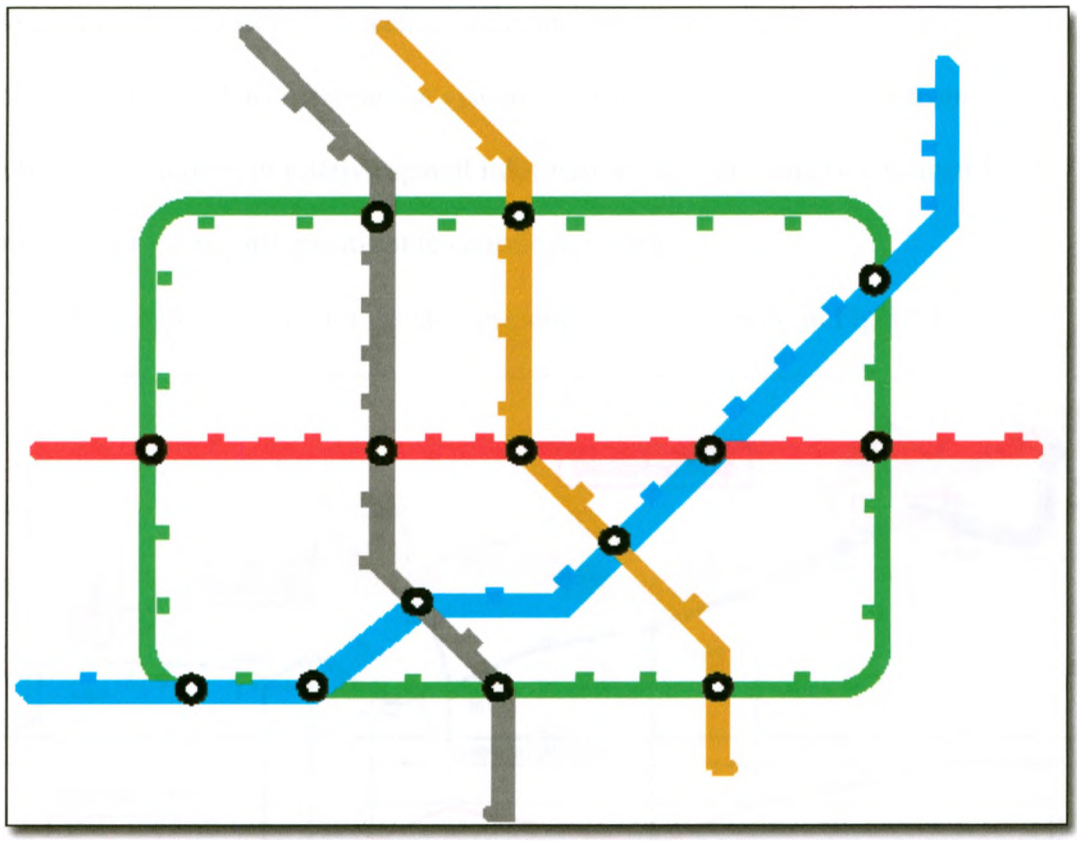

Figure 2.1: London Underground also known as Harry Beck's map [26]

Figure 2.1 depicts a classical example of a visual representation. This representation is in fact so effective at transmitting spatial information to the human conceptual system that it has been installed in almost every underground transit systems in the world. The representation uses colors to encode five different interconnected

subway lines. It uses circular icons to encode platforms where passengers may embark and disembark, and it encodes spatial information by the relative locations and orientations of the colored lines to each other. When a human user is confronted with this representation, his/her cognitive system is able to quickly analyse it to extract his/her current location, his/her desired destination, and the number of stops until he/she reaches it. Of further interest is the fact that the map is not to scale. In fact, the map has been purposefully distorted in both the x and y direction to represent spatial information only in terms of platform stops between relative positions rather than in some quantifiable coordinate system. This information visualization technique is called *xy-distortion* and can be used to represent relative spatial information without being constrained by the limitations of a standard quantifiable coordinate system.

A second example of a visual representation is presented in Figure 2.2.

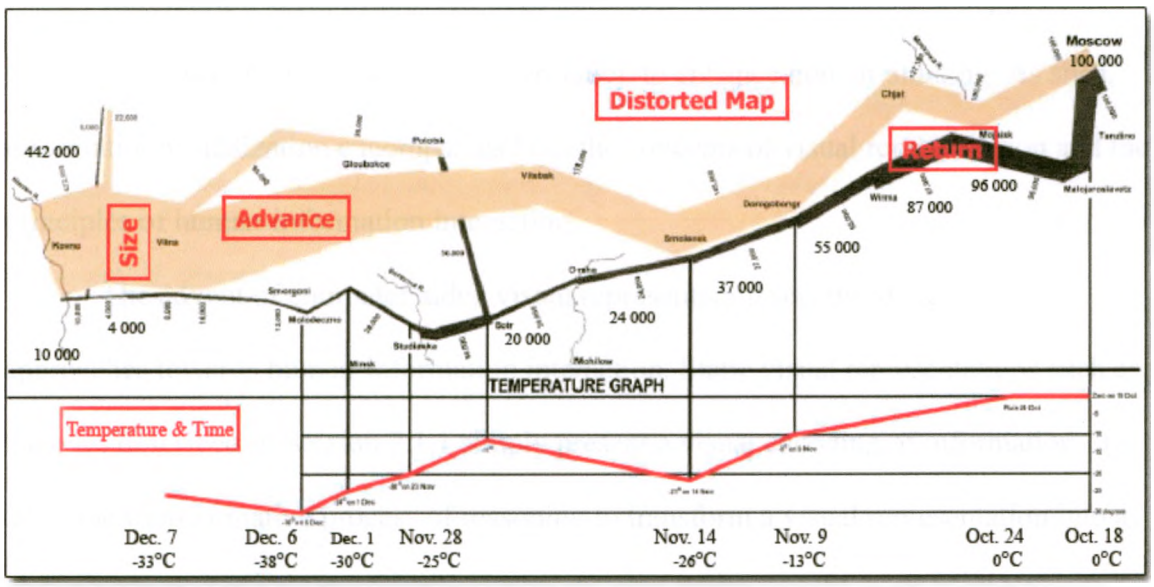

Figure 2.2: Napoleon's advance on Moscow, also known as Minard's map [26]

Figure 2.2 shows a second classical example of a visual representation, that of Minard's map. Charles Joseph Minard, Napoleon's map maker, successfully encoded multimodal

information into a single visual representation. Minard used color to encode the tactical modality of the army, beige for advancing and black for retreating. He used xy-distortion to encode the geographical information relating the relative position of the army to Moscow, Russia. Furthermore, he used the size of the colored bands representing the army to encode the number of living soldiers at the time annotated on x-axis of the temperature graph below. His map is a time-space narrative of Napoleon's advance and retreat from Russia relative to the measured temperature shown on the graph beneath it. Minard's map is an example of encoding time and space simultaneously. This information presentation technique is often called a *time-space narrative* and is frequently used in modern visual representations.

2.1.2 Interaction

Information visualization has been defined as the use of computer-supported, *interactive*, visual representations of information to enhance human thinking. As such, information visualization encompasses both the concepts of visual representation and the principles of human-information interaction.

The advent of computer aided visual representation sets the stage for novel approaches towards human-information interaction. Static visual representations such as the ones described in Section 2.1.1 simply present a visual encoding of information. It takes the transformative process of reasoning to transform a visual representation into a well constructed mental model from which we may derive insight or understanding. However, deriving well constructed mental models from static representations can require large cognitive efforts. One way to reduce the amount of effort required to transform a visual representation into an effective mental model is to allow the user to interact with the representation through an interactive visual interface. Visual interfaces

allow the user to explore their data in order to understand trends and anomalies, isolate and recognize information as pertinent, and to engage in the analytical reasoning process. The principles governing augmenting analytical reasoning through visual representations and the use of their visual interfaces will next be discussed.

2.2 Scientific Visualization

Contrary to the field of information visualization, which is primarily concerned with the representation of abstract data, the field of scientific visualization focuses on representing sets of data that possess a natural geometric structure. An example of a scientific visualization is Brown et al.'s [3] panoramic image registration research, shown in Figure 2.3.

Figure 2.3: Brown et al.'s panoramic image registration results without blending [3]

In Figure 2.3, several unique images are stitched together to form an overall visual representation of a panoramic scene. The resulting visualization can be classified as a scientific visualization because the data possesses a natural geometric structure. This visualization shows how image registration can help put detailed data, the individual images, into context. While the boundaries between the fields of scientific visualizations and information visualizations are not usually clear, the objective in scientific visualization is always to model real-world objects or phenomenon as naturally and as accurately as possible.

In research literature, medical researchers working in the field of scientific visualization have been applying concepts borrowed from the field of information visualization to further their

research efforts. An example of this is Rueckert et al.'s [22] research presenting a method that uses anatomical atlases to stitch medical images of the brain taken from different patients into an artificial, distorted, coordinate system.

2.3 Visual Analytics

Visual analytics is an outgrowth of the fields of information visualization and scientific visualization that focuses on facilitating analytical reasoning through interactive visual interfaces [32]. Section 2.1.1 explored how visual representations augment the cognitive reasoning process with perceptual reasoning which permits the analytical reasoning process to become faster and more focused [5]. However, most analytic activities often require the ability to interact with data in order to search for patterns or anomalies. As such, visual analysis can be described as a dialog between an analyst and data using a visual representation simply as a view or interface into the data. During an analysis dialog, the analyst observes, understands, questions, and then decides which new factors should be considered. The challenge of designing appropriate user interactions can therefore be described as the problem of determining how the analyst will request different perspectives on the data and how he/she will filter out unwanted details.

To this end, cognitive, perceptual and graphical design principles for visual analysis have been a hot topic of research. In 2005, Thomas et al. [5] presented five widely accepted principles for designing visual representations and their interactions such as to facilitate visual analysis. These principles are defined as follows:

2.3.1 The Appropriateness Principle

The appropriateness principle states that the visual representation should provide neither more, nor less information than is required to perform the current analytical task. The

presentation of any additional information will serve only as a distraction. Distractions will, in turn, make performing the desired analytical task more difficult.

2.3.2 The Naturalness Principle

Visual analysis is most effective when the properties of the visual representation closely resemble the information being presented. The naturalness principle supports the concept that visual metaphors are only useful when they match the users' cognitive model of the information. As such, purely artificial visual metaphors may hinder analysis and obstruct understanding.

2.3.3 The Matching Principle

The matching principle states that representations should be suggestive to the user performing the appropriate analytical task. Effective visual representations should present affordances to suggest to the user which actions are appropriate when deriving insights and understanding.

2.3.4 The Principle of Congruence

The principle of congruence indicates that the structure of the visual representation should correspond to the structure of the desired mental visualization. Furthermore, visual representation should focus and present the important concepts in the domain of interest.

2.3.5 The Principle of Apprehension

The principle of apprehension underlies the importance of research in perception. It states that the structure and content of the visual representation should be readily perceived and easily comprehended. If a visual representation is too complex, it may obstruct the analysis dialog between the analyst and the data.

The above mentioned principles for depicting information to facilitate visual analysis can be used to solicit the key requirements when designing an analytical tool.

2.4 Summary

In this chapter, an introduction to the concepts, principles and techniques of information visualization, scientific visualization, and visual analysis were presented. The main concepts of presentation, representation, interaction and analysis were examined. The currently accepted principles for depicting information to facilitate visual analysis were examined to provide the foundation for the requirements that will be solicited when developing the method proposed in this thesis.

Chapter 3

Literature Review

This chapter examines several approaches to solving the panoramic image registration problem as presented in research literature. Subsequently, some specific applications using image registration techniques to augment medical image analysis are discussed. Academic research efforts relating medical imagery to anatomical atlases are then examined. Finally, the open research questions that should be further explored are identified and become the main focus of the rest of this thesis.

3.1 Panoramic Image Registration

Panoramic image registration techniques have a sizable research literature. In this section, notable works and approaches to solving the image registration problem are examined.

In 1975, Milgram [19] introduced a technique that allowed overlapping images to be combined into a single panoramic image. The goal was to minimize the visual impact of the seams resulting from overlapping images. The author geometrically registered images one scan line at a time by choosing a seam point for each line. This created an artificial edge at each seam point which was then locally smoothed. The images were matched by user interaction.

In 1981, Fishler and Bolles [8] introduced the idea of automating the image registration process. The authors presented a major breakthrough in the field by presenting an algorithm for model fitting that was ideally suited for the automation of the image registration process. The algorithm was termed random sample consensus (RANSAC) and remains the backbone of many of the most modern image registration

algorithms [3]. RANSAC is an algorithm used to fit a model to a set of experimental data. It is essentially designed to smooth data that contains copious amounts of error. In feature based alignment, the number of mismatched features between images often exceeds the number of accurately selected feature pairs. Since RANSAC is designed to robustly validate such data, it is ideally suited to select accurate features from the large sample of erroneous candidates. Today, the set of algorithms that are used to locate common features between images remain prone to large quantities of error [3]. As such, the RANSAC algorithm is still used to filter their results.

Moving away from the feature based methods enabled by RANSAC, in 1995 Szeliski and Kang [28] presented a technique that was concerned with the registration of multiple images into a projective mapping. Their technique did not rely on special features to form a projective model. Instead, they directly solved a least-squares approximation of the motion parameters and the system of unknowns which lead to a statistically optimal estimation. Their method approximated a homography for each image that best mapped it to a common coordinate system. The authors were successful in their implementation but the solution was not suitable for situations where the data was collected by hand held cameras as their solution required carefully controlled camera motions.

Motivated by their aforementioned implementation, Szeliski and Shum [29] developed a novel approach to panoramic image registration which relied on a set of transforms that did not require carefully controlled camera motions. Instead of recovering the planar perspective transforms of each image, the authors' algorithm directly recovered 3D rotations using a 3-parameter motion model [29]. As a result, the authors

were able to stitch panoramic images without experiencing singularities at the poles of cylindrical and spherical maps.

Figure 3.1: Tessellated spherical panorama constructed from 54 images [29]

Figure 3.1 shows an example of Szeliski and Shum's solution of a tessellated spherical panorama. Their method, however, required prior knowledge of the camera's focal length [29].

Zoghlami et al. [35] noted that the problem of computing homographies to stitch images together falls mainly into two cases: the case where transformations are composed mainly of translations and the case where transformations are composed of large scaling and rotational angles. Zoghlami et al. [35] capitalized on the efficient methods to solve the first case, and presented a method to remove the user interaction necessary to solve the second case. This work removed the requirement for special user interaction for any camera rotation around the optical axis with fairly large zooming factors.

Similarly to Zoghlami et al.'s implementation, Capel and Zisserman [4] described a technique for the automatic image registration of images acquired by a camera rotating about the viewpoint. It built on the idea that the maximum likelihood estimation of the homographies that relate each image to a common coordinate system could be computed.

Their technique was novel in that it was automated, it did not accumulate errors over a sequence of images, and it introduced new concepts such as enhanced resolution.

In 1999 Lowe [14] presented an automated method in of image registration by introducing an algorithm capable of detecting a special class of image features. The features were invariant to image scaling, translation and rotation, while also tolerant to changes in intensity and perspective projection. The scale invariant feature transform (SIFT) algorithm created by the authors could identify stable points in scale space. The algorithm would create image keys that allowed for local geometric deformations by using blurred image gradients in multiple orientation planes. These keys were used as inputs to a nearest-neighbor matching pass that identifies candidate matches. The verification for each match was achieved by finding the least squares approximation for the unknown model parameters. The results showed that robust object recognition could be achieved in cluttered and somewhat occluded images [14].

In 2000, Shum and Szeliski [24] presented a new system for the construction of image registrations. Their system implemented patch based alignment algorithms to align pairs of images and global alignment algorithms to refine the alignment of the overall panoramic. The authors claimed to be able to improve the quality of their panoramic images significantly by combining global and local alignment methods.

Extending Shum and Szeliski's research efforts, in 2002 McLauchlan and Jaenicke [18] presented a journal that demonstrated that accurate panoramic registrations of planar scenes can be constructed by implementing photogrammetric bundle adjustment techniques. The author's method enabled the use of lines in camera self-calibration [18]. The technique allowed the computation of radial and other non-linear distortions. The

author concluded that implementing photogrammetric bundle adjustments could achieve better results than optimizations based on pairs of images.

In 2007, Brown and Lowe [3] presented a solution to the fully automated panoramic image registration problem. The solution used Lowe's earlier work [14] on the SIFT algorithm to extract scale invariant features, filtered those features using RANSAC [8] selecting only inliers, found connected components using bundle adjustments [18], and then rendered the final registration using multi-band blending. Their algorithm is shown in Figure 3.2.

Algorithm: Automatic Panorama Stitching

Input: n unordered images

I. Extract SIFT features from all n images

II. Find k nearest-neighbours for each feature using a k-d tree

III. For each image:
  (i) Select m candidate matching images that have the most feature matches to this image
  (ii) Find geometrically consistent feature matches using RANSAC to solve for the homography between pairs of images
  (iii) Verify image matches using a probabilistic model

IV. Find connected components of image matches

V. For each connected component:
  (i) Perform bundle adjustment to solve for the rotation $\theta_1, \theta_2, \theta_3$ and focal length f of all cameras
  (ii) Render panorama using multi-band blending

Output: Panoramic image(s)

Figure 3.2: Modern automatic panoramic image registration algorithm [3]

The authors went further to introduce blending and gain compensating techniques, as well as automatic straightening methods. As a result, a fully automated panoramic

image was assembled from a set of unordered source images. Their results are shown in Figure 3.3.

Figure 3.3: Brown and Lowe's automatic panorama image registration results [3]

Figure 3.3 shows an example of Brown and Lowe's image registration algorithm applied to a set of uncalibrated and unordered images. Their implementation is currently considered state of the art.

This section of the literature review has examined a subset of the extensive research and development effort in panoramic image registration. The next section of the literature review will examine how panoramic image registration must face new challenges in order to be applied within non-photographic imaging applications.

3.2 Medical Image Registration

The ever expanding domain of medical imaging provides researchers and clinicians with an increasingly multi-faceted view of human anatomy. Maurer et al. [17] remarked that the information provided by various image modalities is often both complimentary and synergetic. That is to say that the information collected using different techniques provides separate but useful information and that the combination of such information can provide additional useful information. Consequently, advances in image registration methods have contributed to medical image analysis by proving insight into the techniques required to register various image modalities to a common

coordinate space. In this section, recent works and novel approaches to medical image analysis using image registration are examined.

In 2001, Hill et al. [13] noted that radiological images from different sensing modalities are being used together with increasing frequency in medical research. The authors focused on the problem of accurately relating information from different images to each other for diagnosis, treatment and research. They reviewed the techniques used to solve image registration problems in the medical field and described their implications. The applications of image registration methods that the authors examined were those of combining images of a same subject from different modalities, aligning sequences of images to compensate for movements in the subject between scans, and aligning images from multiple subjects for cohort studies. The authors concluded that current registration algorithms can automatically register images that are related by affine transformations and that there has been substantial progress in non-rigid registration algorithms as well. They also noted that there remain a number of image registration problems such as automatic non-rigid image registration for deformable organs that are unsolved and that as a result image registration for medical applications will likely continue to be an active field of research.

In keeping with Hill et al.'s predictions, many researchers have been developing techniques to stitch together multimodal medical imagery. Woo et al. [33] for example, developed an automated method for the registration of cardiac computed tomography (CT) and single photon emission computed tomography (SPECT) data to help diagnose coronary artery disease [33]. The results of the authors multimodal image registration are shown in Figure 3.4.

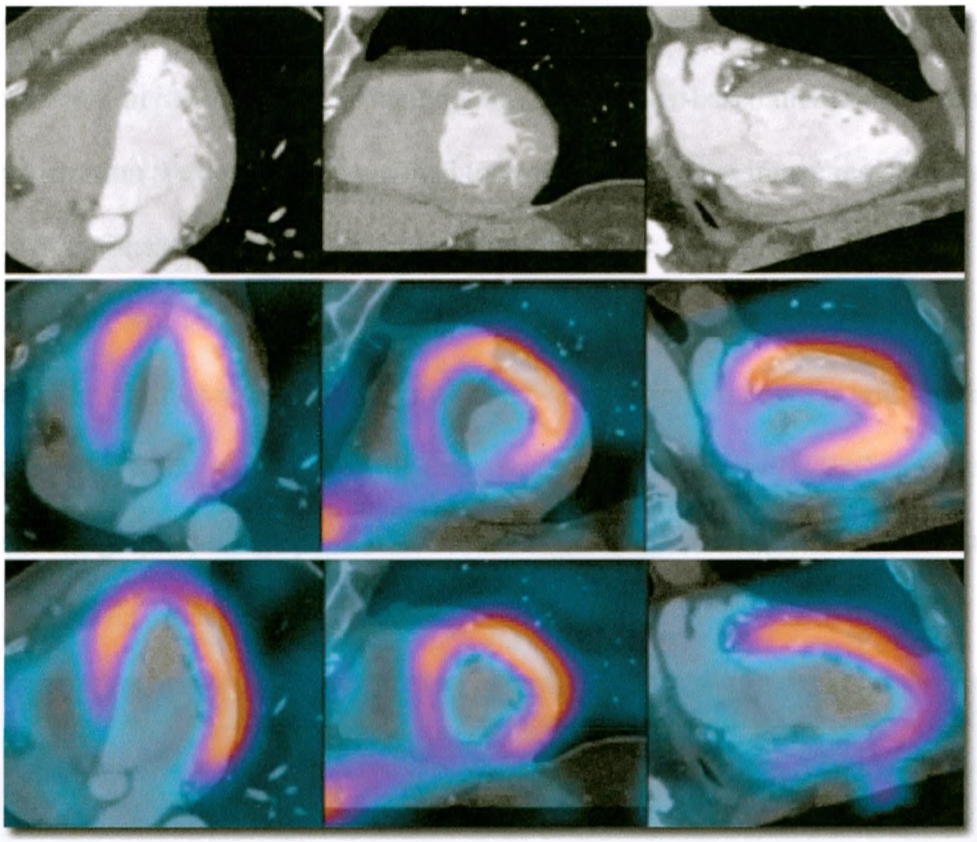

Figure 3.4: Image registration of CT and SPECT data [33]

Figure 3.4 shows Woo et al.'s original CT data as the top row, visually overlapped CT/SPECT data as the middle row, and their automatically stitched results as the bottom row. The authors' fully automated registration method used the geometric features from a segmentation of gated myocardial perfusion volumes to extract blood pools and used them as a mask to blend the disparity in overlapping intensities. The authors succeeded in automatically registration CT scans with SPECT data with an accuracy of less than 10mm in 87% of their cases [33].

Contributing to research efforts in medical image registration, Šerifović-Trbalić et al. [23] developed an image registration algorithm that combines anisotropic landmarks with rotational information. The authors examined the addition of rotational attributes in the thin-plate spline model that could enhance the estimation of the deformations that

warp one set of landmarks to another in sequences of breast images [23]. The authors were successful in integrating their method into an intensity-based hierarchical non-rigid image registration framework and claimed to have improved the image registration accuracy of deformed images of the breast significantly.

In 2007, Škerl et al. [25] presented a study that evaluated the effect that the sequence of modalities would have on the image registration process for multimodal brain scans. The study used 16 CT/MR (magnetic resonance) and PET/MR (positron emission tomography) image pairs with known registration standard errors. An example of their data is shown in Figure 3.5.

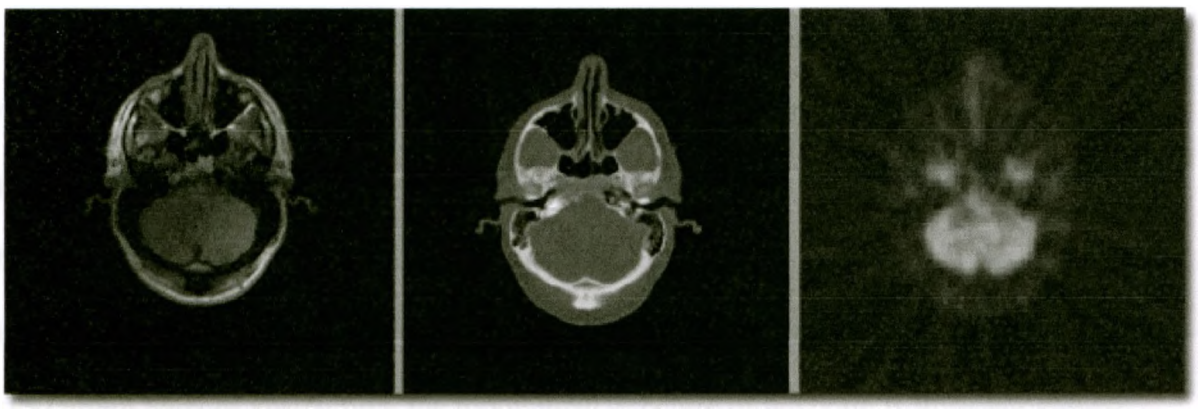

Figure 3.5a: MR data Figure 3.5b: CT data Figure 3.5c: PET data

Figure 3.5: One slice of a human head [25]

Figure 3.5 shows an example of the data used in Škerl et al.'s research. Figure 3.5a is of MR modality, Figure 3.5b is of CT modality, and Figure 3.5c is of PET modality. The study was restricted to rigid similarity matches between image modalities and investigated the effect that the order of selection of the source and target image modalities would have on the resulting registration. The study found that for PET/MR and CT/MR image registration, the behavior of the similarity measures depended significantly on which image is the source image, and which image is the target image.

Furthermore, the study concluded that for PET/MR and CT/MR image registration mutual information, normalized mutual information, and entropy correlation coefficients generated accurate similarity matches.

This section has provided an examination of a subset of research literature relating image registration methods to multimodal medical imaging within a subject. The next section will examine research literature supporting the natural extension of the image registration problem for medical applications; that of combining medical image registration methods to enable comparative studies between subjects, within modality.

3.3 Atlas Based Medical Image Registration

The logical extension of the medical image registration problem is that of registering images taken from different individuals for comparison. This idea originates from the observation that there is remarkable consistency between the structures of the human brain between subjects if the images are scaled and oriented relative to the deep internal structure of a common plane [17]. The assumption underlying this method is that at a certain level of abstraction, the topological structure of the brain is invariant among populations. This structure is called an anatomical atlas. Instead of registering images to each other, the objective becomes to find the transformation that will map the image to the anatomical atlas, while simultaneously accounting for local shape differences.
For example, Friston et al. [9] presented an automated solution to the image registration problem of an analyst wishing to match and compare images of a same modality taken from different subjects. The authors presented a spatial normalization technique and used it to match images of the brain to an ideal image, model, or template [9]. The anatomical atlas chosen was the space described in the atlas of Talairach and Tournoux [30]. They explained that their technique, which maps images to a common brain space, is practical

for inter-subject averaging of PET activation, change distribution analysis and statistical parametric mapping.

Similarly, D. Rueckert et al. [22] presented a method for determining the anatomical variability of the human brain across a sample population. However, instead of mapping the images to the brain atlas described by Talairach and Tournoux [30], the authors created their own anatomical brain atlas by building a model of the average anatomy of all subjects after non-rigid registration, and then applying the average deformation to the coordinate system of the reference subject. The authors were able to construct a statistical deformation model of the brain based on 25 MR images of different subject with schizophrenia. Their results are illustrated in Figure 3.6.

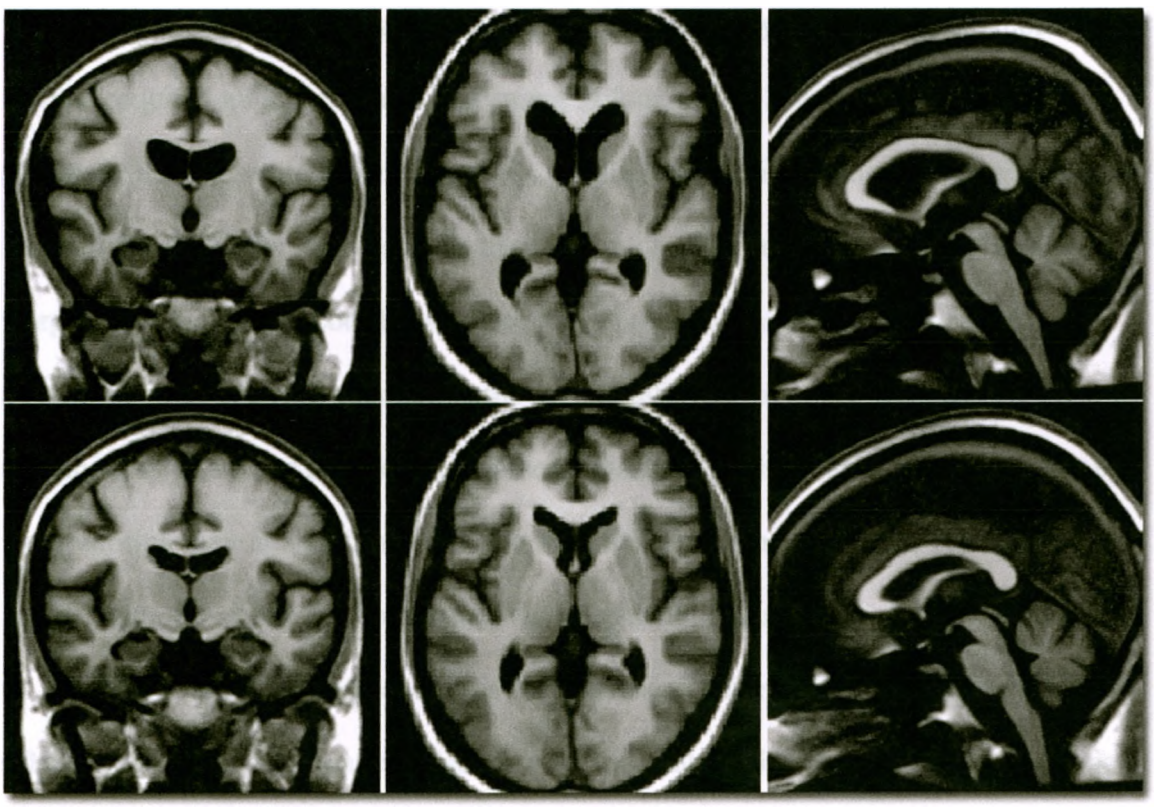

Figure 3.6: Average brain atlas of 25 patients [22]

Figure 3.6 shows Rueckert et al.'s results after registering the brain MR images from 25 patients suffering from schizophrenia. The top row shows the results embedded in the

coordinate system of the reference subject, and the bottom row shows the results embedded in a computed brain atlas coordinate system [22]. While there is a significant amount of research literature examining non-rigid inter-subject brain mapping or brain atlas matching [27], there is far less academic literature pertaining to using atlas matching techniques on other organs in the body.

Park et al. [21] provide one example of such research literature. Instead of constructing inter-subject brain mappings, the authors examined atlas matching techniques for organs and bone structures located inside the abdomen. In particular, the authors were interested in the liver, kidneys, and spinal cord. Using 32 CT scans, 31 of them were mapped into one individual representation using a thin plate spline as the warping technique, and mutual information as the similarity measure. The authors' method required an initial coarse placement of four control points to initialize the automatic atlas based registration process. Additional user interaction was required to manually segment the four organs in each of the 32 CT scans. The segments were then warped onto the patient coordinate space and a probabilistic atlas was calculated. The atlas was subsequently used to aid in the segmentation of low-contrast organs in an additional 20 CT scans not included in the original atlas. The authors concluded that by incorporating their atlas information into the Bayesian framework, their results showed improvements over the current unsupervised segmentation method.

This section has explored some current research literature pertaining to the inter-subject atlas method of image registration for medical applications. It has examined how brain atlas mapping has dominated the research area. It has also explored one of the few examples of using the atlas method for image registration of parts of the body other than the brain.

3.4 Summary

In this chapter, research papers focusing on panoramic image registration methods have been examined. Research papers presenting the natural extension of the panoramic image registration problem to medical image analysis have also been explored. Subsequently, the research literature describing the atlas based method of medical image registration for comparative studies has been explored. Since the goal of this thesis is to devise an image registration method by injecting concepts, principles, and techniques borrowed from the fields of information visualization, scientific visualization, and visual analytics, the development of a method capable of presenting a scientific visualization, based on principles of information visualization, to enable the visual analysis of a patient's medical records over time will be devised. To this end, a method of image registration will be combined with an information visualization technique to visually encode their relationships with respect to time. Consequently, a mechanism to embed medical images into a common time-space coordinate system will be devised. The research literature presented in this chapter will be used to present a method that permits the spatial and temporal registration of images into a common time-space coordinate system. This method will be the focus of the rest of this thesis.

Chapter 4

An Interactive Time-Space Narratives of Imagery

The purpose of this thesis is to generate an image registration method by combining information visualization techniques, concepts, and principles with medical image registration methods. This chapter will solicit the requirements for the proposed method capable of presenting an interactive visual representation taking the form of a time-space narrative based on the information visualization principles presented in Chapter 2 and the image registration methods presented in Chapter 3.

4.1 Data

The modernization of our medical institutions has given rise to digitally encoded medical records. These records contain data encoded in many forms including numerical data such as age, weight, and height, ordinal data such as sequences of symptoms and treatments, categorical data such as age ranges, sex, and ethnicity, temporal data such as important dates and times, and image data which can be collected from a number of imaging devices in two or more dimensions. Since the focus of this thesis is to extend the research in medical image registration by incorporating the visual encoding of time, a set of two-dimensional x-ray images is sufficient to represent a sample set of data.

The x-ray image modality is a popular form of medical imagery. These images are collected by casting radiation through a subject and recording its absorption characteristics. The absorption characteristics are represented as two dimensional arrays of spatially related intensity values and are used to depict internal structure such as bone structure, soft tissue, and/or metals. An example of an x-ray image is shown in Figure 4.1.

Figure 4.1: X-ray imagery of a subject's neck

Figure 4.1 shows an x-ray of a subject's neck from a profile perspective. In this image, the spatial relationship between pixels represents the relative locations of the x-rays at the time of data collection. The intensity value of each pixel represents the absorption characteristic of the material that was penetrated by the x-ray at each location. The relative intensities between groups of pixels are suggestive to the shape and structure of the subject's internal materials such as bones, teeth, soft tissue, and the metallic item lodged in his neck.

In addition to the data encoded in the pixel arrays of x-ray images, there is information embedded in the file header. The Digital Imaging and Communications in Medicine (DICOM) standard is an internationally accepted standard for the file format and distribution of medical images of many modalities, including x-rays. The pertinent information contained in a DICOM formatted header is the subject's name, the type of scan, the dimensions of the imagery, annotations, and the date and time at which the scan

was collected. As such, the sample set of x-ray images in this thesis will provide both imagery data to identify the absorption characteristics of internal structures, and temporal data to identify when the images were taken.

4.2 Visual Representation

Chapter 2 defined a visual representation as the visual encoding of data in such a way that its presentation assists the human cognitive activity of forming a mental model or mental image of something. As such, the goal of this visual representation is to encode the x-ray and temporal data described in the previous section in such a way that its presentation assists an analyst to form a mental model or mental image of a patient's medical records. Furthermore, the appropriateness principle, presented in Section 2.3.1, states that a visual representation should provide neither more, nor less information than is required to perform the current analytical task. Since the current analytical task is to reason about a patient's medical history, and that this reasoning requires both spatial and temporal information, it follows that an appropriate visual representation should simultaneously present the x-ray images, and their temporal relationships. As seen in Section 2.1.1, this type of visual representation is called a time-space narrative.

4.2.1 Representing Space

Since the x-ray data collection process requires that a patient to be subjected to doses of radiation and that exposure to radiation has been shown to be dangerous, it is typical to x-ray as small an area as possible when investigating an affliction. Consequently, one of the tasks that must be performed by an analyst when reading x-ray images is that of deriving from them a mental understanding of what small part of the subject's body the current image represents. This is a step of the analytical reasoning process that can benefit from a scientific visualization to reduce the cognitive effort

required to put the detailed data into context. As seen in Section 3.1, one of the most popular methods for putting detailed data into context is that of image registration. Image registration has been used to put adjacent satellite photos into context for decades but does require overlapping scan coverage. If the coverage of a subject's x-ray image is small, and if subsequent images are located far apart (such as an x-ray image of the head and an x-ray image of the shin) the images may not overlap. For this reason, the feature based image registration methods for planar panoramic scenes shown in section 3.1, and their analogous image registration methods for medical applications seen in Section 3.2, may not be suitable.

However, the principle of congruence (seen in Section 2.3.4) suggests that the structure of a visual representation should correspond to the structure of the desired mental visualization. Since the mental visualization that we seek is that of spatially aligned x-ray images, an atlas would present a congruent representation. As a result, the introduction of an atlas provides a common coordinate system in which non-overlapping x-ray images could be registered by atlas based image registration techniques (seen in Section 3.3). Furthermore, the naturalness principle indicates that visual analysis is most effective when the properties of the visual representation closely resemble the information being presented. As such, the atlas should take the form of a human body. However, registering two dimensional x-ray images to an accurate three-dimensional atlas of the human body, which grows and morphs with the lifecycle of the subject might not be the most effective representation for analysis. The apprehension principle states that if a visual representation is too complex, it may obstruct the analysis dialog. Since the x-ray images are two-dimensional images, embedding them into a third dimension could cause the cognitive cost of interaction (three-dimensional camera movements) to

burden the cognitive reasoning system. The apprehension principle also tells us that the structure and content of the visual representation should be readily perceived and easily comprehended. As a result, a simple and easy atlas to perceive is that of a two dimensional anatomical atlas representing the outline of a human body. Such an atlas is shown in Figure 4.2.

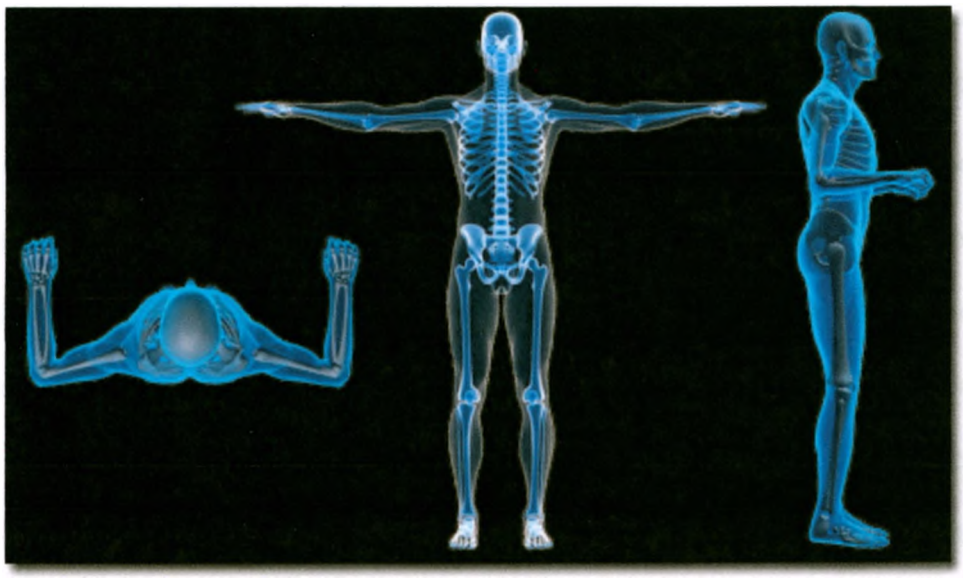

Figure 4.2 Proposed anatomical atlas

Figure 4.2 shows the proposed two-dimensional anatomical atlas to which x-ray images will be stitched to provide a spatially related visual representation of medical records. The atlas represents the three most commonly studied planes of the human body; the transverse, coronal, and sagittal planes. If the x-ray images are collected from drastically different planes than the depicted atlas planes, the atlas image simply needs to be changed or expanded. Furthermore, to compensate for the growth and morphing of the subject over time, the well known information visualization technique of xy-distortion should be implemented. As such, the x-ray images will be geometrically transformed to fit the fixed coordinate system of the anatomical atlas and any previously registered images will not be required to change as the subject grows.

The implementation of the proposed anatomical atlas and its corresponding atlas based image registration requirements will provide context for detailed x-ray data and consequently reduce the cognitive effort required to discern the geometric relationship between the x-ray images and their corresponding locations on a human subject.

4.2.2 Representing Time

Since the analytical task of the visual representation is to reason about a patient's medical history, and that a medical history by definition takes place over time, it follows that an effective visual representation should also present the data's specific temporal information. In the context of this thesis, the times that we wish to represent are the times at which medical images about a subject became available. These are the dates and times encoded in the image DICOM headers. These data/time stamps are of critical importance as they will be used to organize the x-ray images into a timeline sequence. This sequence will be used to see changes over time in the dimension and absorption characteristics of areas of interest in x-ray imagery. For example, given the right sequence and x-ray imagery, an analyst will be able to visualize the growth of a tumour or other affliction over time. The growth of such afflictions can be computed based on a timeline which must be readily perceived by the analyst. As such, the timeline must be an integral part of the representation and, in this thesis, must display the specific data/time at which the imagery was collected.

4.3 Interaction

The matching principle states that representations should be suggestive to the user performing the appropriate analytical task. In the context of this thesis, it means that it should be obvious to the user that he/she can navigate the information contained within the representation based on space or based on time. To accommodate such interactions,

the timeline should be represented as an interactive visual interface. The user must be able to automatically playback the timeline, and must be also be able to manually select the date and time. Furthermore, the user must be able to interact with the geometric information space. To accommodate such interaction, the user must be able to manipulate the geometry of the x-ray images, must be able to zoom to request different perspectives on the data, and must be able to pan the information space to investigate the relationship between the events unfolding in different areas of the body. Through its visual representation, the software application developed for this thesis will allow the analyst to interact with a subject's medical history, to explore it, and to derive insights or understandings that might be helpful for diagnosis.

Chapter 5

Method

This chapter presents the method used to geometrically align x-ray images to the anatomical atlas presented in Chapter 4. It also presents the method used to organize x-ray images into a visually encoded timeline. By combining these two methods, a time-space narrative of a subject's medical history can be achieved.

5.1 Geometric Alignment

In this thesis, image registrations are computed based on user defined point pair correspondences. Point pair correspondences identify common features between the x-ray images and the anatomical atlas. An example of which is shown in Figure 5.1.

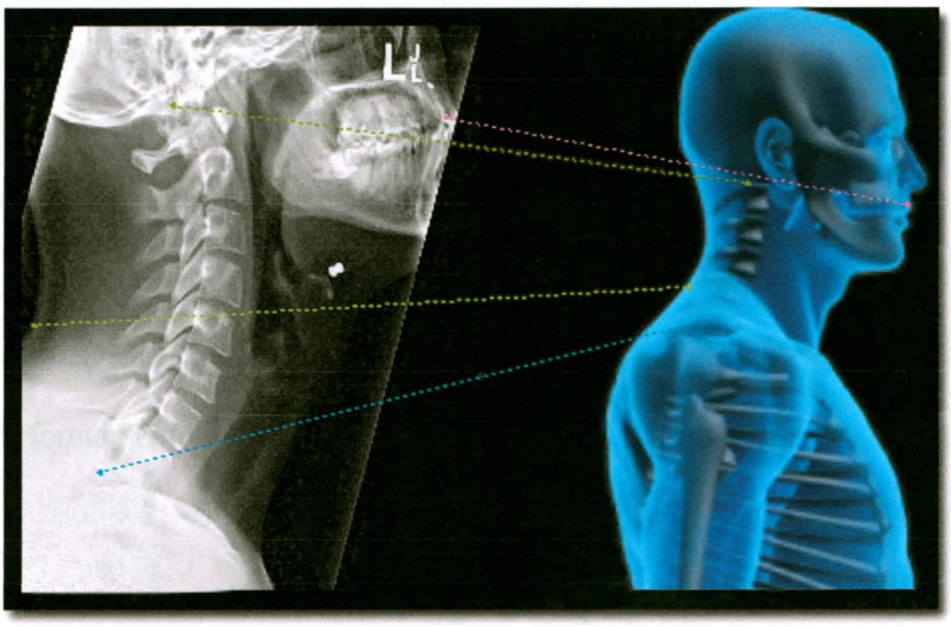

Figure 5.1: Point pair correspondences between an x-ray image and the atlas

Figure 5.1 shows four point pair correspondences identified between an x-ray image and the anatomical atlas. These correspondences identify four common features between the two geometric planes. The image registration method implemented in this thesis uses

these correspondences to embed the x-ray image into the anatomical atlas' coordinate system. To obtain accurate image registration results, it is important that the point pair correspondences be sparsely distributed and far from collinear [7].

Although there exists a number of algorithms to automatically select point pair correspondences, such as the scale invariant feature transform (SIFT) algorithm discussed in the literature review, this thesis implementation requires that the user will manually select each feature. This user-centric approach has the advantage of supporting imagery where features are difficult to automatically detect due to inaccurate geometry or unusual distortions and decreases the scope of the thesis by avoiding the open problems that exist today in automatic feature detection for various data modalities.

In a planar or planar panoramic geometric configuration, sets of x-ray images are related to each other, or to an anatomical atlas, by an invertible transformation matrix called a *homography* [12]. A homography is an invertible transformation from one projective plane to another which preserves straight lines. The geometric relationship between the pixels of planar panoramic images can be expressed algebraically as $p' = Hp$ [12] where p and p' represent pixels of a common feature between two images and H represents the homography that maps pixel p to p'. This equation is shown in its expanded form in Figure 5.2.

$$\begin{bmatrix} wx' \\ wy' \\ w \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Figure 5.2: Mapping between pixels related by a plane to plane homography [7]

It is assumed that the x-ray images and the anatomical atlas used in this thesis are planar in form and that they can be related to each other by a plane to plane homography.

5.1.1 Computing Image Registrations Algebraically

Since the x-ray images are related to the anatomical atlas by a plane to plane homography, image registration can be achieved by computing the homography H that satisfies the linear system of equations $p' = Hp$ for all feature pairs p to p'. Consequently, the x-ray image registration method implemented in this thesis consists of:

i. Choosing the anatomical atlas image as the source image.

ii. Choosing an x-ray image as the target image.

iii. Selecting common features between the target image and the source image.

iv. Computing the homography H which maps all target features to their corresponding source features.

v. Applying the homography H to the x-ray image's coordinates to embed it into the anatomical atlas' coordinate system.

5.1.2 The Least Squares Problem

The problem that presents itself in this thesis is that of relating point pair correspondences to each other by the linear mapping $p' = Hp$. Unfortunately, finding a single homography H that exactly satisfies the equation for all point pair correspondences is likely impossible as there will almost certainly be error in point pair correspondence selection. Therefore, a linear function for which all errors $p'_i - Hp_i$ are simultaneously minimized for all point pair correspondence is sought. However, it is generally impossible to find a linear function for which all the errors $p'_i - Hp_i$ are simultaneously minimized.

As such, a best fit linear solution is needed. This is accomplished by letting r denote the vector of residual errors $p'_i - Hp_i$ for all point pair correspondences. The function is then solved for which the vector norm $\| r \|_2$ is made as small as possible [31]. In practice, to minimize $\| r \|_2$ is the same as to minimize $\| r \|_2^2$ [31]. Therefore what is in fact minimized is the sum of squares of the residual errors $p'_i - Hp_i$. More specifically, $\| r \|_2^2 = \sum_{i=1}^{n} | p_i' - Hp_i |^2$. This is known as the *least squares problem.*

In this thesis, singular value decomposition is the method used to solve the least squares problem.

5.1.3 Eigenvectors, Eigenvalues, and Solving the Least Squares Approximation

In order to understand how to solve the least squares approximation of a set of linear equations, it must first be clear what is meant by eigenvector, what is meant by eigenvalue, and how they approximately solve a system of equations.

Generally, the action of a matrix on a vector changes both its magnitude and direction. However, there are cases where the action of a matrix on certain vectors changes only their magnitude while preserving their direction. Such vectors are called the eigenvectors of that matrix [31]. Each eigenvector must be multiplied by some scalar to affect only its magnitude while preserving its direction. Such scalars are called the eigenvalues of their corresponding eigenvectors. Algebraically, $Ax = \lambda x$ for some matrix A and eigenvector x with scalar eigenvalue λ [31].

Recall in section 5.1.2 that the minimization of $\| r \|_2^2 = \sum_{i=1}^{n} | p_i' - Hp_i |^2$ is sought and represents the least squares error between the expected coordinates and the actual transformed coordinates of the set of point pair correspondences. If matrix A denotes the

matrix of coefficients of the linear system of equations, and vector x denotes the vector of unknowns, then the solution to the equation $Ax = 0$ for some non-zero x is sought. However it is not usually possible to find a non-zero x that satisfies each equation in the system of equations simultaneously [31].

Instead, the right hand of the equation is substituted for the least squares residual error$\| r \|_2$. As such, $Ax = \| r \|_2$. To minimize $\| r \|_2$ means to choose the vector x which yields the smallest possible value$\| r \|_2$. Since what is desired is to preserve the direction of vector x affecting only its magnitude, the eigenvalue equation $Ax = \lambda x = \| r \|_2$ [31] can be of use. Therefore, to minimize $\| r \|_2$, one simply chooses the eigenvector x with the smallest possible eigenvalue λ. Fortunately, we know that singular value decomposition is a method to compute such an eigenvector.

5.1.4 Singular Value Decomposition and the Least Squares Approximation

Singular value decomposition (SVD) is a method for the factorization of a matrix [10] into a coordinate system where the covariance matrix is diagonal [16]. SVD takes as input a $n \times m$ matrix A in which n *rows* represents the equations and m *columns* represents the coefficients of an over-determined system of equations.

The SVD theorem states that:

$$A_{n \times m} = U_{n \times n} D_{n \times m} V_{m \times m}^{T}$$

Where:

U has left singular vectors as its columns.

D is diagonal with singular values arranged in descending order.

V^{T} has right singular vectors as its rows.

Calculating the SVD means finding the eigenvalues and eigenvectors of AA^T and A^TA [16]. The eigenvectors of A^TA compose the columns of V, and the eigenvectors of AA^T are the columns of U. Furthermore, the singular values in D are the square roots of eigenvalues from AA^T or from A^TA. Singular value decomposition yields singular values as the diagonal entries of the D matrix, which are arranged in descending order [16]. If the SVD of the *A* matrix is computed, a matrix of eigenvectors *V* arranged in descending order sorted by eigenvalue is obtained. If the right most eigenvector is taken, the values corresponding to the least squares approximation of the system of equations is obtained. These will be the values that most accurately describe the transformation from the set of domain image points to the set of range image points.

5.2 Classes of Transformations

There exists a hierarchy of transformation groupings with specific geometric properties that are useful when performing x-ray image registrations with respect to an anatomical atlas. In this section, these groupings are presented and their implications for x-ray image registration are discussed.

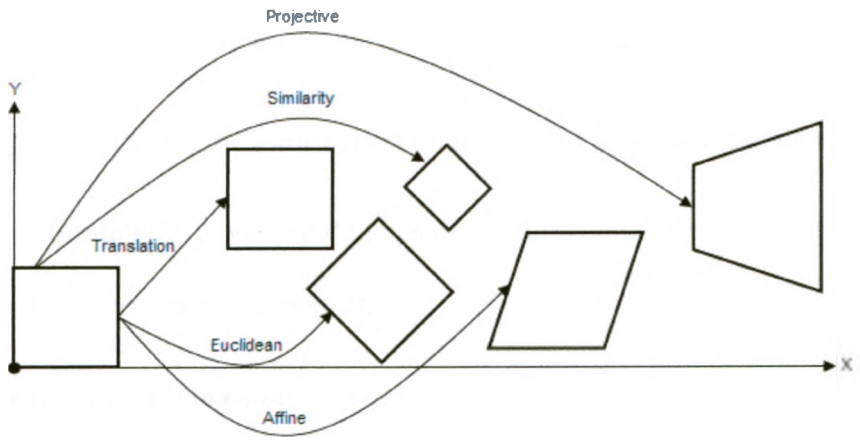

Figure 5.3: Geometric representation of the classes of transformations [12]

5.2.1 Translation

Translations are a class of transformation that permit only displacement [1]. Under this class of transformation, every pixel in the image moves parallel to and the same distance as every other pixel in the image [1]. Translations are described algebraically as:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Figure 5.4: Translation transformation relating point pair correspondences [12]

This class of transformation has two degrees of freedom and can be computed from 1 or more point pair correspondences [12].

5.2.2 Similarity Transformations

Similarity transformations are transformations that preserve shape [15]. They permit isotropic scaling, rotation, and translation and are described algebraically as:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} s\cos\alpha & -s\sin\alpha & t_x \\ s\sin\alpha & s\cos\alpha & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Figure 5.5: Similarity transformation relating point pair correspondences [15]

Similarity transformations have four degrees of freedom and can be computed from two or more point pair correspondences [12].

5.2.3 Euclidean Transformations

Euclidean transformations are transformations that preserve Euclidean distances and allow only rotation followed by translation [6]. They are in effect a model for rigid

body motion and are practical for relating images of a common scale. Euclidean transformations are described algebraically as:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha & t_x \\ \sin\alpha & \cos\alpha & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Figure 5.6: Euclidean transformation relating point pair correspondences [6]

This class of transformation has three degrees of freedom and can be computed from two point pair correspondences [12].

5.2.4 Affine Transformations

Affine transformations are non-singular linear transformations followed by a translation [1]. They permit isotropic scaling, rotation, translation, stretching, and sheering and can be described algebraically as:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Figure 5.7: Affine transformation relating point pair correspondences [12]

Affine transformations have six degrees of freedom and can be computed from three or more point pair correspondences [12].

5.2.5 Projective Transformations

Projective transformations are a transformation grouping that permits scaling, rotation, translation, sheering, stretching, and perspective warping. They are described algebraically as:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Figure 5.8: Projective transformation relating point pair correspondences [12]

Projective transformations have 8 degrees of freedom and can be computed from four or more point pair correspondences [12].

5.3 Estimating Transformations

This section describes how to estimate the transformation matrices that best satisfy the system of equations $p' = Hp$ for each transformation grouping presented in Section 5.2.

5.3.1 Choosing an Appropriate Transformation Class

Under-determined systems of equations cannot be solved by least squares approximation. As such, an appropriate transformation class based on the number of point pair correspondences that have been identified must be chosen. The following table illustrates the transformations that are permissible given the number of identified point pair correspondences.

Number of Point Pair Correspondences	Available Classes of Transformations				
	Translation	Similarity	Euclidean	Affine	Projective
0 pairs					
1 pair	•				
2 pairs	•	•	•		
3 pairs	•	•	•	•	
4 or more pairs	•	•	•	•	•

Table 5.1: Number of point pair correspondences versus permissible transformations

5.3.2 Estimating Translations

Consider the list of point pair correspondences with one or more pairs which map feature points on one image to corresponding feature points on a different image:

$$p = (x, y) \rightarrow p' = (x', y')$$

A translation which maps point p to point p' can be defined as:

$$p' = p + T$$

Solving for T:

$$T = p' - p$$

Expanding and writing it in homogenous matrix form:

$$\begin{bmatrix} wt_{x_i} \\ wt_{y_i} \\ w \end{bmatrix} = \begin{bmatrix} x_i' - x_i \\ y_i' - y_i \\ 1 \end{bmatrix}$$

In the exactly determined case, where there is only one point pair correspondence, the computed translation is the solution for $(t_{x_i}, t_{y_i})^T$. In the over-determined case, where there are two or more point pairs, one must solve $(t_{x_i}, t_{y_i})^T$ for each point pair. In the exactly determined case, the exact translation vector is computed. In the over-determined case, one must compute the least squares approximation of the set of all $(t_{x_i}, t_{y_i})^T$ computations.

The computed translation is written as a 3x3 matrix:

$$T = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

Multiplying the current image's transformation matrix H by the computed translation matrix T performs the desired translation transformation, $H' = HT$.

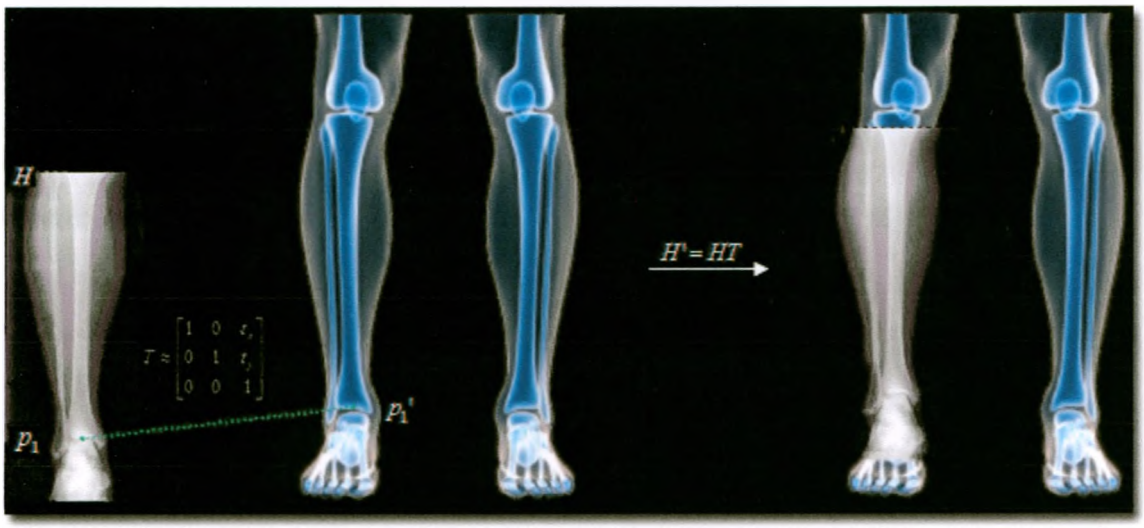

Figure 5.9: Image registration based on the translation transformation group

Figure 5.9 shows an example of an image registration computed using one point pair correspondence and the translation transformation group.

5.3.3 Estimating Similarity Transformations

Similarity transformations support scaling, rotation, and translation.

Consider the list of point pair correspondences with two or more pairs which map feature points on one image to corresponding feature points on a different image:

$$p = (x, y) \rightarrow p' = (x', y')$$

A similarity transformation which maps points p to p' can be described as:

$$p' = Sp \qquad (1)$$

Where S is a homogenous similarity transformation matrix of the form:

$$S = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} s\cos\alpha & -\sin\alpha & t_x \\ \sin\alpha & s\cos\alpha & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

In the exactly determined case, where there are two point pair correspondences, one can solve the equation:

$$X = A^{-1} * B$$

Where:

$$A = \begin{bmatrix} x_1 & -y_1 & 1 & 0 \\ y_1 & x_1 & 0 & 1 \\ x_2 & -y_2 & 1 & 0 \\ y_2 & x_2 & 0 & 1 \end{bmatrix} \quad B = \begin{bmatrix} x_1' \\ y_1' \\ x_2' \\ y_2' \end{bmatrix}$$

Solving $X = A^{-1} * B$ for X and rearranging into the transformation matrix:

$$S = \begin{bmatrix} X_{11} & -X_{21} & X_{31} \\ X_{21} & X_{11} & X_{41} \\ 0 & 0 & 1 \end{bmatrix}$$

In the over-determined case, where there are three or more point pair correspondences, equation (1) can be rewritten as:

$$S_i \bullet h = 0 \qquad (2)$$

Where:

$$S_i = \begin{bmatrix} x_i & -y_i & 1 & 0 & -x_i' \\ y_i & x_i & 0 & 1 & -y_i' \end{bmatrix} \quad h = \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix}$$

h being the 6×1 vector of unknown coefficients in S.

For example, if five point pair correspondences are considered:

$$S_5 = \begin{bmatrix} x_1 & -y_1 & 1 & 0 & -x_1' \\ y_1 & x_1 & 0 & 1 & -y_1' \\ x_2 & -y_2 & 1 & 0 & -x_2' \\ y_2 & x_2 & 0 & 1 & -y_2' \\ x_3 & -y_3 & 1 & 0 & -x_3' \\ y_3 & x_3 & 0 & 1 & -y_3' \\ x_4 & -y_4 & 1 & 0 & -x_4' \\ y_4 & x_4 & 0 & 1 & -y_4' \\ x_5 & -y_5 & 1 & 0 & -x_5' \\ y_5 & x_5 & 0 & 1 & -y_5' \end{bmatrix} \quad h = \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix}$$

To solve $S_t \bullet h = 0$ for h, one uses singular value decomposition (SVD) on matrix S_t and chooses the unit eigenvector of $S_t^T S_t$ with the least eigenvalue (this will be the last column of the matrix V generated by singular value decomposition).

The values taken from the last column in V are then divided by their scale factor w (the last value of the column taken from V) to return to 2D image coordinates. The values are then rearranged into matrix form:

$$V_{column} = \begin{bmatrix} aw \\ -bw \\ cw \\ fw \\ w \end{bmatrix} \qquad S = \begin{bmatrix} V_{11}/V_{51} & -V_{21}/V_{51} & V_{31}/V_{51} \\ V_{21}/V_{51} & V_{11}/V_{51} & V_{41}/V_{51} \\ 0 & 0 & 1 \end{bmatrix}$$

Multiplying the current image's transformation matrix H by the computed similarity transformation S performs the desired similarity transformation, $H' = HS$.

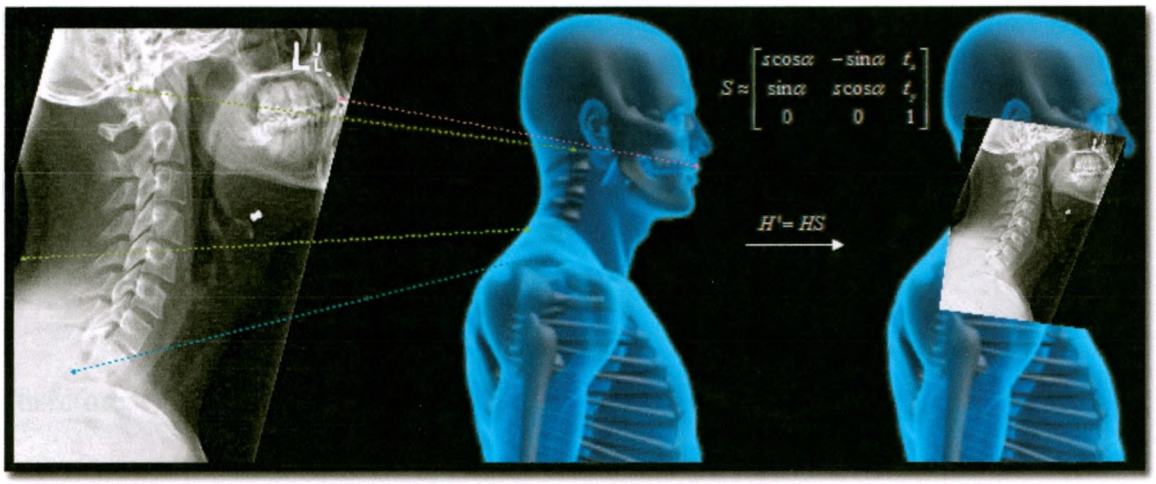

Figure 5.10: Image registration based on the similarity transformation group

Figure 5.10 shows an example of an image registration computed using four point pair correspondence and the similarity transformation group.

5.3.4 Estimating Euclidean Transformations

Euclidean transformations are a subset of similarity transformations that are restricted to rotation followed by translation. This class of transformation is surprisingly common when relating images to each other. Many surveying applications use a fixed range when collecting data. This facilitates the image registration process as all images are collected at a common scale. In this thesis implementation, to compute the Euclidean transformation E the application must first compute the similarity transformation S (see section 5.3.2). The scale factor embedded within the transformation is subsequently removed.

Let S be the computed similarity transformation based on the point pair correspondences $p = (x, y) \rightarrow p' = (x', y')$.

$$S = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} s\cos\alpha & -\sin\alpha & t_x \\ \sin\alpha & s\cos\alpha & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

Since: $\sin^2\alpha + \cos^2\alpha = 1$

$$S_{21}{}^2 + \cos^2\alpha = 1$$

Rearranging and solving for $\cos(\alpha)$:

$$\cos\alpha = \sqrt{1 - S_{21}{}^2}$$

Therefore:

$$E = \begin{bmatrix} \sqrt{1 - S_{21}{}^2} & S_{12} & S_{13} \\ S_{21} & \sqrt{1 - S_{21}{}^2} & S_{23} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha & t_x \\ \sin\alpha & \cos\alpha & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

Multiplying the current image's transformation matrix H by the computed Euclidean transformation matrix E performs the desired Euclidean transformation, $H' = HE$.

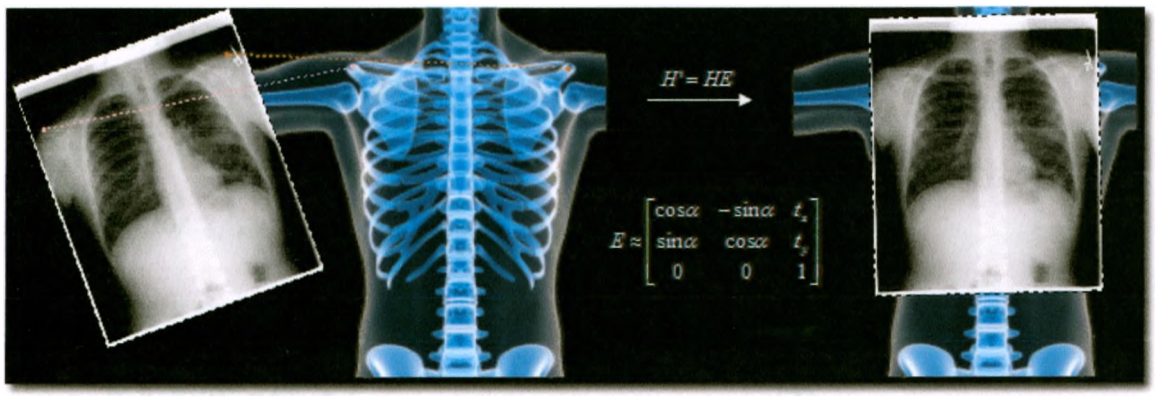

Figure 5.11: Image registration based on Euclidean transformation group

Figure 5.11 shows an example of an image registration computed using two point pair correspondence and the Euclidean transformation group.

5.3.5 Estimating Affine Transformations

Affine transformations support scaling, rotation, stretching, skew, and translation. Consider the list of point pair correspondences with three or more pairs which map feature points on one image to the corresponding feature points on a different image:

$$p = (x, y) \rightarrow p' = (x', y')$$

An affine transformation which maps points p to p' can be described as:

$$p' = Ap \qquad (1)$$

Where A is the 3x3 homogenous affine transformation matrix of the form:

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix}$$

Writing p and p' in homogenous coordinates and expanding (1) for all point pairs:

$$\begin{bmatrix} w_i x_i' \\ w_i y_i' \\ w_i \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \qquad (2)$$

Expanding for x_i', y_i', and w_i yields three parametric equations:

$$w_i x_i' = a x_i + b y_i + c \qquad (3)$$

$$w_i y_i' = d x_i + e y_i + f \qquad (4)$$

$$w_i = 1 \qquad (5)$$

Substituting (5) into (3) and (4):

$$x_i' = a x_i + b y_i + c \qquad (6)$$

$$y_i' = d x_i + e y_i + f \qquad (7)$$

Rearanging (6) and (7):

$$ax_i + by_i + c - x_i' = 0 \qquad (8)$$

$$dx_i + ey_i + f - y_i' = 0 \qquad (9)$$

In the exactly determined case, where there are three point pair correspondences, (8) and (9) can be used to form the equation:

$$X = A^{-1} * B$$

Where:

$$X = \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} \quad A = \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1 & y_1 & 1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_2 & y_2 & 1 \\ x_3 & y_3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_3 & y_3 & 1 \end{bmatrix} \quad B = \begin{bmatrix} x_1' \\ y_1' \\ x_2' \\ y_2' \\ x_3' \\ y_3' \end{bmatrix}$$

Solving $X = A^{-1} * B$ for X and rearranging its values to produce the affine matrix A :

$$A = \begin{bmatrix} X_{11} & X_{21} & X_{31} \\ X_{41} & X_{51} & X_{61} \\ 0 & 0 & 1 \end{bmatrix}$$

In the over-determined case, where there are four or more point pair correspondences, (8) and (9) can be written in the form:

$$A_i \bullet h = 0 \qquad (10)$$

where:

$$A_i = \begin{bmatrix} x_i & y_i & 1 & 0 & 0 & 0 & -x_i' \\ 0 & 0 & 0 & x_i & y_i & 1 & -y_i' \end{bmatrix} \qquad h = \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix}$$

h being the 6×1 vector of unknown coefficients in A.

For example, if five point pair correspondences are considered:

$$A_5 = \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1' \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -y_1' \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2' \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -y_2' \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x_3' \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -y_3' \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4' \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -y_4' \\ x_5 & y_5 & 1 & 0 & 0 & 0 & -x_5' \\ 0 & 0 & 0 & x_5 & y_5 & 1 & -y_5' \end{bmatrix} \qquad h = \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix}$$

To solve $A_i \bullet h = 0$ for h, singular value decomposition (SVD) is applied to matrix A_i and the unit eigenvector of $A_i^T A_i$ with the least eigenvalue (this will be the last column of the matrix V generated by singular value decomposition) is chosen.

The values taken from the column in V are subsequently divided by their scale w (the last value of the column taken from V) to convert back to 2D image coordinates. The values are rearrange to form the matrix A :

$$V_{>column} = \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ w \end{bmatrix} \qquad A = \begin{bmatrix} V_{11}/V_{71} & V_{21}/V_{71} & V_{31}/V_{71} \\ V_{41}/V_{71} & V_{51}/V_{71} & V_{61}/V_{71} \\ 0 & 0 & 1 \end{bmatrix}$$

Multiplying the current image's transformation matrix H by the computed affine transformation matrix A performs the desired affine transformation, $H' = HA$.

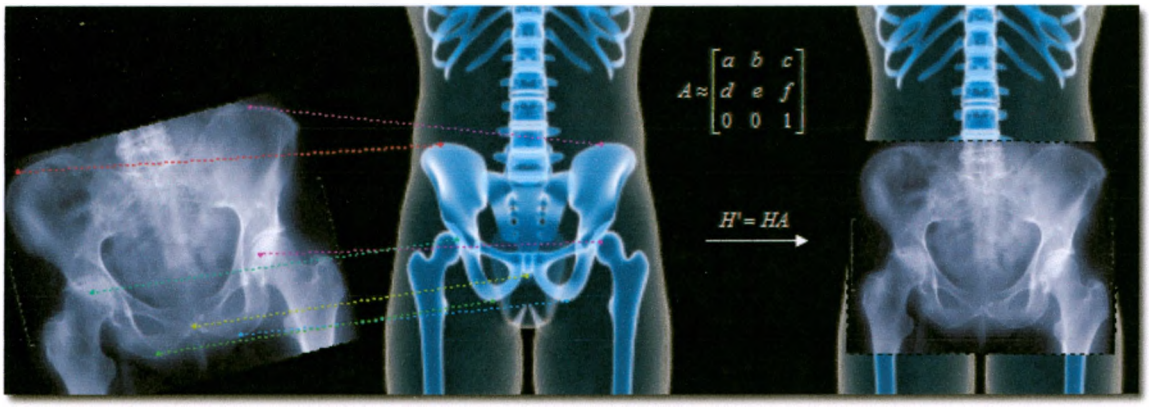

Figure 5.12: Image registration based on the affine transformation group

Figure 5.12 shows an example of an image registration computed using seven point pair correspondences and the affine transformation group.

5.3.6 Estimating Projective Transformations

Projective transformations support scaling, rotation, stretching, skew, translation, and image reprojection [12]. Projective transformations are typically useful when the data is taken from the same center of projection and rotated about an axis. Consider the list of point pair correspondences with four or more pairs which map feature points on one image to the corresponding feature points on a different image:

$$p = (x, y) \rightarrow p' = (x', y')$$

A homographic transformation H which maps points p to p' can be described as:

$$p' = Hp \qquad (1)$$

Where H is a 3x3 homography of the form:

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

Writing p and p' in homogenous coordinates and expanding (1) for all point pairs:

$$\begin{bmatrix} w_i x_i' \\ w_i y_i' \\ w_i \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \quad (2)$$

Expanding (2) for x_i', y_i', and w_i' yields three equations:

$$w_i x_i' = ax_i + by_i + c \quad (3)$$

$$w_i y_i' = dx_i + ey_i + f \quad (4)$$

$$w_i = gx_i + hy_i + i \quad (5)$$

Substituting (5) into (3) and (4):

$$(gx_i + hy_i + i)x_i' = ax_i + by_i + c \quad (6)$$

$$(gx_i + hy_i + i)y_i' = dx_i + ey_i + f \quad (7)$$

Rearanging (6) and (7):

$$ax_i + by_i + c - gx_i x_i' - hy_i x_i' - ix_i' = 0 \quad (8)$$

$$dx_i + ey_i + f - gx_i y_i' - hy_i y_i' - iy_i' = 0 \quad (9)$$

In the exactly determined case, where there are four point pair correspondences, (8) and (9) can be used to form the equation:

$$X = A^{-1} * B$$

$$X = \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \\ 1 \end{bmatrix} \quad A = \begin{bmatrix} x_i & y_i & 1 & 0 & 0 & 0 & -x_i x_i' & -y_i x_i' \\ 0 & 0 & 0 & x_i & y_i & 1 & -x_i y_i' & -y_i y_i' \end{bmatrix} \quad B = \begin{bmatrix} x_i' \\ y_i' \end{bmatrix}$$

Solving $X = A^{-1} * B$ for X and rearrange its values to produce the homography A:

$$A = \begin{bmatrix} X_{11} & X_{21} & X_{31} \\ X_{41} & X_{51} & X_{61} \\ X_{71} & X_{81} & 1 \end{bmatrix}$$

In the over-determined case, where there are four or more point pair correspondences, (8) and (9) can be written to form:

$$A_i \bullet h = 0 \qquad (10)$$

$$A_i = \begin{bmatrix} x_i & y_i & 1 & 0 & 0 & 0 & -x_i x_i' & -y_i x_i' & -x_i' \\ 0 & 0 & 0 & x_i & y_i & 1 & -x_i y_i' & -y_i y_i' & -y_i' \end{bmatrix} \qquad h = \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \\ i \end{bmatrix}$$

h being the 9x1 vector of unknown coefficients in A.

For example, if five point pair correspondences are considered:

$$A_5 = \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1 x_1' & -y_1 x_1' & -x_1' \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1 y_1' & -y_1 y_1' & -y_1' \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2 x_2' & -y_2 x_2' & -x_2' \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -x_2 y_2' & -y_2 y_2' & -y_2' \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x_3 x_3' & -y_3 x_3' & -x_3' \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -x_3 y_3' & -y_3 y_3' & -y_3' \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4 x_4' & -y_4 x_4' & -y_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -x_4 y_4' & -y_4 y_4' & -y_4' \\ x_5 & y_5 & 1 & 0 & 0 & 0 & -x_5 x_5' & -y_5 x_5' & -y_5 \\ 0 & 0 & 0 & x_5 & y_5 & 1 & -x_5 y_5' & -y_5 y_5' & -y_5' \end{bmatrix} \qquad h = \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \\ i \end{bmatrix}$$

To solve $A_i \bullet h = 0$ for h, singular value decomposition (SVD) is applied to matrix A_i and the unit eigenvector of $A_i^T A_i$ with the least eigenvalue (this will be the last column of the matrix V generated by singular value decomposition) is chosen [12]. The values

taken from the column in V are divided by their scale w (the last value of the column taken from V) to convert back to 2D image coordinates. The values are rearranged to form the homography:

$$V_{>column} = \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \\ w \end{bmatrix} \qquad A = \begin{bmatrix} V_{11}/V_{91} & V_{21}/V_{91} & V_{31}/V_{91} \\ V_{41}/V_{91} & V_{51}/V_{91} & V_{61}/V_{91} \\ V_{71}/V_{91} & V_{81}/V_{91} & 1 \end{bmatrix}$$

Multiplying the current image's transformation matrix H by the computed homography A performs the desired projective transformation, $H' = HA$.

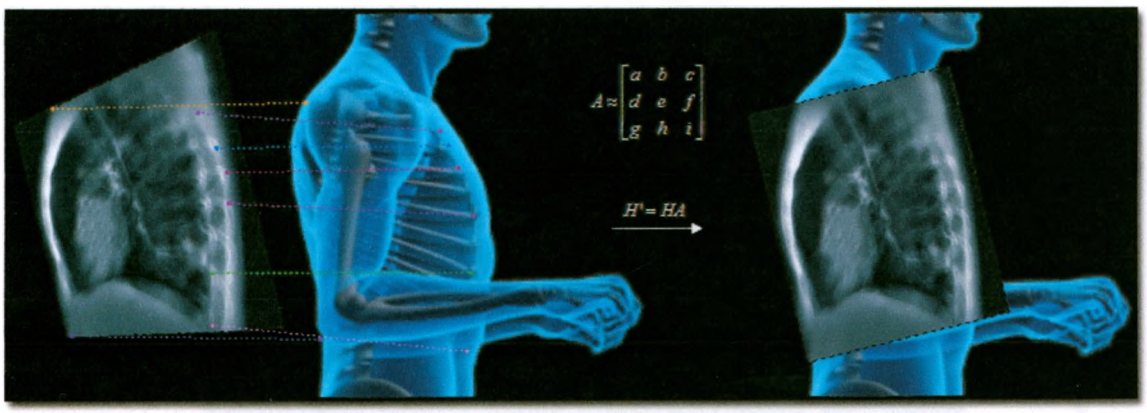

Figure 5.13: Image registration based on projective transformation group

Figure 5.13 shows an example of an image registration computed using seven point pair correspondence and the projective transformation group.

5.4 Reconstructing the Timeline

This thesis describes a system designed specifically to facilitate the analytical task of reasoning about the spatial and temporal relationships that exist between images. As such, a method to organize the images into an appropriate sequence and present their temporal information to the user must be devised.

In the context of this thesis, the temporal information that is of interest is the dates and time at which each of the subject's medical images became available. These dates and times are typically stored in the x-ray image's DICOM headers. Consequently, a software method was implemented to decode the DICOM headers and to store the date/time stamps into the newly geometrically aligned image objects. Subsequently, the list of images is sorted in chronological order.

Once a timeline is reconstructed, it must be intuitively displayed to the user. The techniques used to present the timeline to the user are shown below in Figure 5.14.

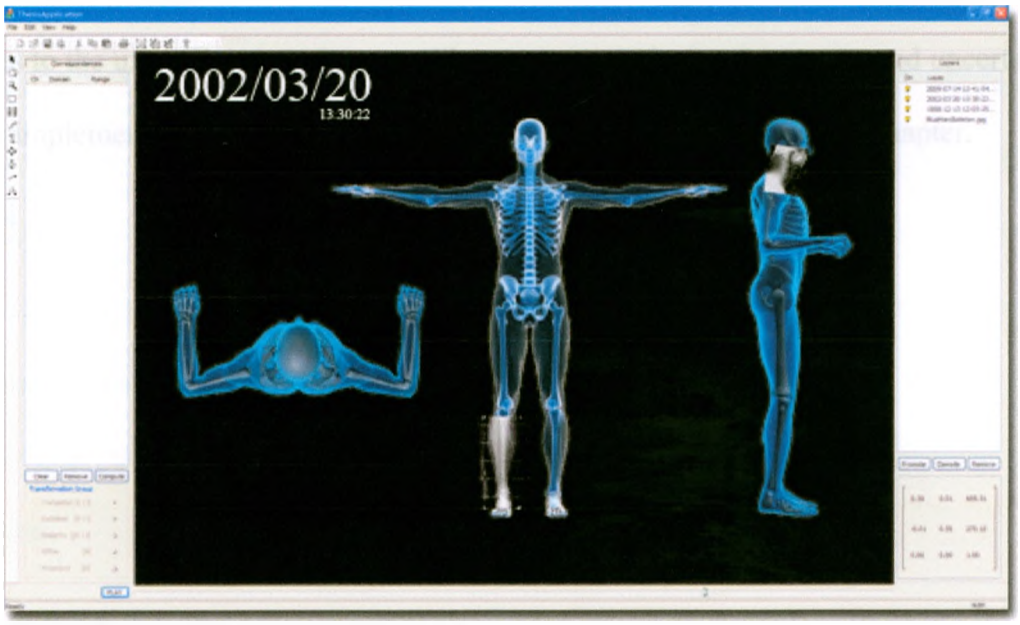

Figure 5.14: Playback of stitched x-ray imagery in chronological order

Figure 5.14 shows an example of the playback functionality of the thesis implementation. The example shows three x-ray images geometrically aligned with respect to the anatomical atlas, rendered over time. On the top left hand side of the application, the current date and time is displayed. Docked on the bottom of the application is the interactive timeline. The user is able to control the date and time by manipulating the slider and is able to animate the sequence of overlaying the spatially and temporally related images by clicking on the play button. On the right hand side of the application, the x-ray image's filenames are displayed in chronological order. The combination of these three presentation techniques allows the user to interact with and perceive the temporal relationships that exist between x-ray images.

5.5 Summary

This chapter has described the method used to stitch x-ray image to an anatomical atlas. Subsequently, it has presented the method used to reconstruct a timeline and the presentation technique used to communicate the temporally and spatially related x-ray images to the user. The next chapter will present an analysis of the measured uncertainty in the implementation of the image registration method presented in this chapter.

Chapter 6

Analysis

This chapter provides an analysis of the uncertainty in image registration resulting from the implementation of the image registration method presented in Chapter 5. The uncertainties in point pair correspondence selection, homography estimation, and the error propagation are measured and discussed. Ignoring the errors caused by floating point arithmetic (assuming that the arithmetic packages used in the implementation of this thesis are numerically stable), the image registration methods implemented in this thesis have two overlapping sources of uncertainty. These are the errors introduced by the manual selection of features, and the uncertainty in the approximation of homographies based on the erroneous features.

6.1 Error in Point Pair Correspondence Selection

The standard error in point pair correspondence selection is a measure of the disparity between the location of an intended feature, and where the user clicks in world coordinates to select it. To determine the standard error resulting from the user manually selecting feature points, an experimental dataset with known feature locations is created. The user is asked to select 100 feature points on the image with random ordering and the disparities between user clicks and feature locations are logged. Figure 6.1 shows the experimental dataset created to measure the uncertainty in point pair correspondence selection.

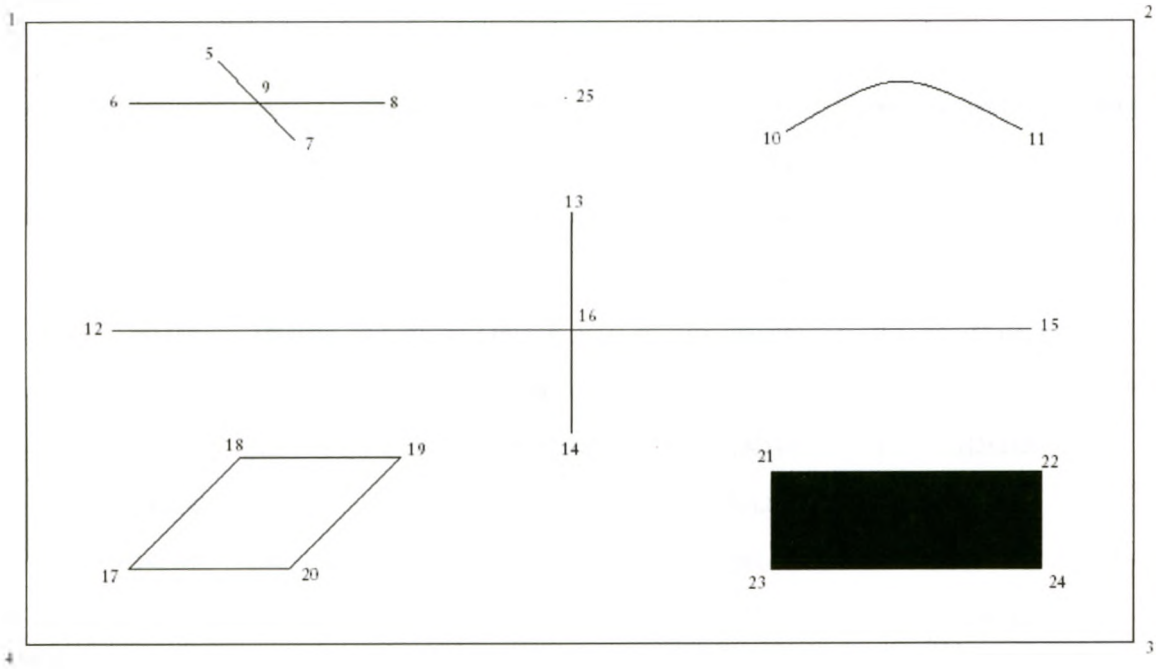

Figure 6.1: Synthetic data generated to estimate feature selection uncertainty

The synthetic data seen in Figure 6.1 is used to simulate common types of features seen in x-ray images. Features identified in the synthetic data include end-points, line intersections, and large changes in intensities. Table 6.1 is the sample data obtained by having a user select 25 features with random ordering and repeating the exercise 4 times for a total of 100 samples.

Feature	Coordinates	Attempt 1	Disparity^2	Attempt 2	Disparity^2	Attempt 3	Disparity^2	Attempt 4	Disparity^2
1	(57,59)	(57,59)	0	(58,60)	2	(57,59)	0	(57,59)	0
2	(776,59)	(776,60)	1	(776,60)	1	(776,60)	1	(776,60)	1
3	(776,464)	(776,464)	0	(777,465)	2	(777,465)	2	(777,465)	2
4	(57,464)	(58,465)	2	(57,464)	0	(58,465)	2	(58,465)	2
5	(181,85)	(182,86)	2	(182,86)	2	(182,86)	2	(181,85)	0
6	(124,112)	(125,113)	2	(125,113)	2	(124,113)	1	(125,112)	1
7	(230,136)	(231,137)	2	(231,137)	2	(231,137)	2	(231,137)	2
8	(288,112)	(288,112)	0	(289,113)	2	(289,113)	2	(289,113)	2
9	(207,112)	(208,113)	2	(207,113)	1	(207,112)	0	(208,113)	2
10	(550,131)	(550,132)	1	(550,131)	0	(550,132)	1	(550,132)	1
11	(703,130)	(704,131)	2	(703,131)	1	(703,131)	1	(705,132)	4
12	(113,260)	(114,261)	2	(114,262)	3	(114,261)	2	(114,261)	2
13	(410,327)	(411,184)	1	(410,185)	1	(410,185)	1	(410,185)	1
14	(410,327)	(410,327)	0	(411,327)	1	(411,328)	2	(410,328)	1
15	(709,260)	(709,261)	1	(709,260)	0	(710,260)	1	(709,260)	0
16	(410,260)	(410,260)	0	(410,260)	0	(410,261)	1	(410,261)	1
17	(123,415)	(123,415)	0	(124,415)	1	(124,415)	1	(123,415)	0
18	(195,343)	(195,344)	1	(196,343)	1	(195,344)	1	(196,343)	1
19	(299,343)	(299,344)	1	(300,344)	2	(299,343)	0	(299,344)	1
20	(227,415)	(228,415)	1	(228,416)	2	(228,415)	1	(228,415)	1
21	(540,353)	(540,353)	0	(540,353)	0	(540,353)	0	(540,353)	0
22	(716,353)	(717,353)	1	(716,353)	0	(716,353)	0	(717,354)	2
23	(540,416)	(540,417)	1	(541,416)	1	(540,417)	1	(540,416)	0
24	(716,416)	(717,418)	3	(717,417)	2	(717,417)	2	(717,416)	1
25	(406,109)	(407,110)	2	(407,110)	2	(407,110)	2	(406,110)	1

Table 6.1: Measured disparity based on 100 synthetic feature selections

The standard error for feature selection is computed as:

$$std.error = \sqrt{\frac{\sum_{i=1}^{n} disparity^2}{n}} = \sqrt{\frac{117}{100}} = 1.08 pixels$$

Figure 6.2: Equation for the standard error in feature selection [31]

By substituting the data from Table 6.1 into the equation for the standard error presented in Figure 6.2, we conclude that the user is able to approximately select a screen pixel +/- 1.08 pixels. As such, the user will likely introduce errors in the set of points used to calculate the approximate transformation that registers images into a common coordinate system. What follows is an analysis of how the errors in point pair correspondence selection affect homography estimation.

6.2 Error in Homography Parameter Estimation

In the context of this thesis, the standard error resulting from homography estimation is a measure of the disparity between where feature points are after the transformation Hp and where they are expected to be, p'.

The standard error for each image registration resulting from the least squares approximation of their transformation matrix is computed as:

$$std.error = \sqrt{\frac{\sum_{i=1}^{n} dist(p'_i, Hp_i)^2}{n}}$$

Figure 6.3: Standard error resulting from least squares approximation [31]

What follows is an analysis of the standard error of a set of six planar images registered to a common coordinate system based on selecting several point pair correspondences for each image. The results of the image registration are recorded in Table 6.2.

Image	Std. Error (pixels)
Source image	N/A
2	1.86
3	0.57
4	1.51
5	0.31
6	1.28

Table 6.2: Results from registration error analysis

Table 6.2 shows that the standard error for each image registration was less than two pixels. Furthermore, the sum of the standard errors is 5.53 pixels. What follows is an analysis of how the error propagates throughout the scene composed of six image registrations.

6.3 Error Propagation

If images are registered directly to the anatomical atlas, the errors in their homography estimations are fixed and do not vary across the scene. However, the user may also register images relative to each other. In such cases, the errors in image registration accumulate across the sequence of images registrations. To measure how the error is propagated throughout a scene, distance measurements are measured between an easily identifiable pixel in the source image, and various pixels in each subsequent image. These measurements are compared with their equivalent measurements in the ground truth image. Their disparities, representing the actual measured errors in the image registrations, are computed and recorded in Table 6.3.

Image	Measurement	Distance (registration)	Distance (truth)	Error (pixels)
1	1	390.03	390.23	0.20
	2	137.88	138.13	0.25
	3	187.29	187.11	0.18
	4	232.30	232.09	0.21
	5	333.65	333.45	0.20
2	1	456.97	456.20	0.77
	2	341.88	340.83	1.05
	3	365.77	366.48	0.71
	4	639.76	638.58	1.18
	5	985.08	983.85	1.23
3	1	688.05	686.21	1.84
	2	502.15	500.36	1.79
	3	959.38	958.22	1.16
	4	1174.01	1172.78	1.23
	5	1165.66	1163.98	1.68
4	1	1065.68	1062.97	2.71
	2	1088.00	1085.89	2.11
	3	1206.61	1203.70	2.91
	4	1295.17	1293.10	2.07
	5	1464.32	1461.75	2.57
5	1	2272.95	2269.72	3.23
	2	2563.01	2559.69	3.32
	3	2299.42	2296.61	2.81
	4	2578.56	2575.84	2.72
	5	2326.94	2324.29	2.65
6	1	3173.70	3170.28	3.42
	2	3171.12	3167.04	4.08
	3	2757.01	2753.85	3.16
	4	3099.56	3096.29	3.27
	5	3291.12	3287.74	3.38

Table 6.3: Measured cumulative error in image registration

Table 6.3 shows the measurements between easily identifiable pixels in the source image, and various pixels in subsequently registered images. The disparities between these measurements represent the actual measured errors in the image registrations and are visually represented by the graph in Figure 6.4.

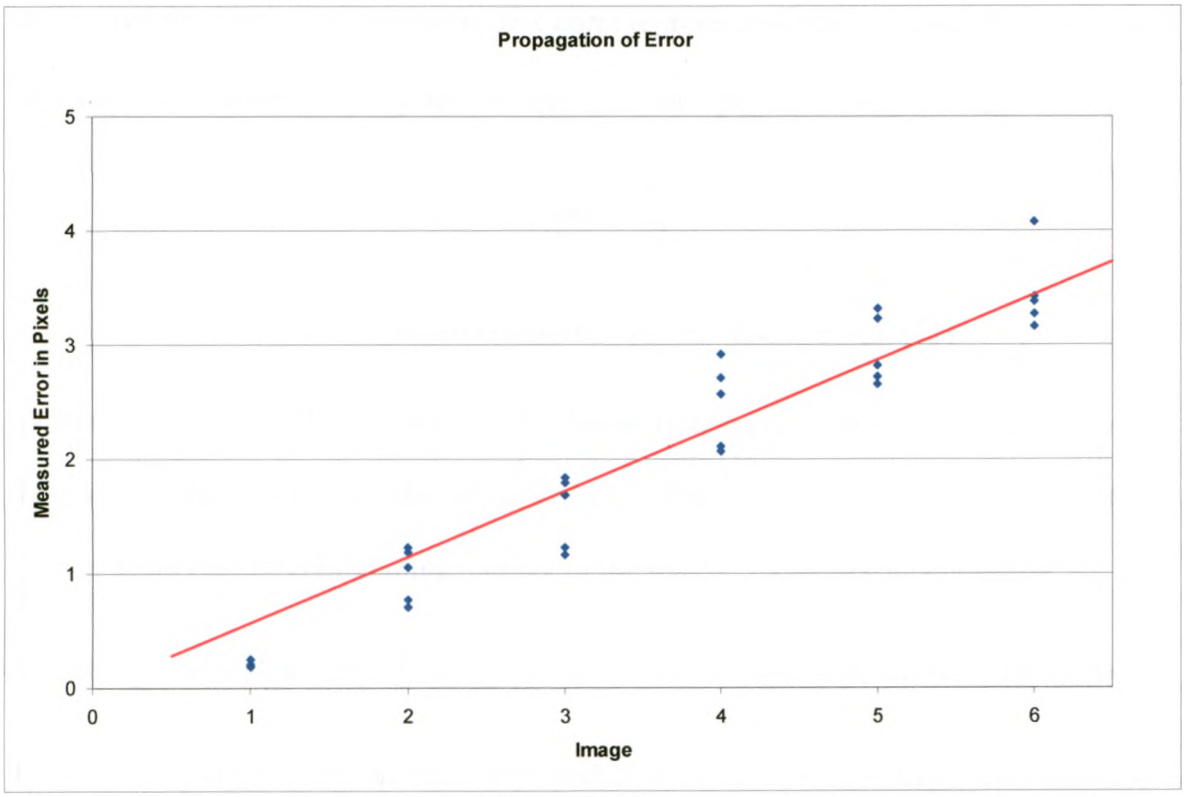

Figure 6.4: Measured propagation of error of a planar scene

Figure 6.4 shows that if images are registered relative to each other, and not directly to the anatomical atlas, the error grows linearly across the sequence of image registrations. This result is expected as image registrations are computed based on point pair correspondences between adjacent images. The error in one image registration will carry over into the next. The maximum error, measured at less than 4 pixels falls within the estimated sum of standard errors of +/-5.53 pixels.

6.4 Perturbation Analysis

In addition to estimating the standard errors caused by estimating the transformations, two sensitivity metrics are also computed: the spectral matrix norm, and the matrix conditioning number.

The spectral matrix norm, $\| A \|_2$, represents the maximum magnification that can be undergone by any vector h when acted on by A in the linear system $Ah = b$. In the

context of this thesis, the spectral matrix norm represents worst case magnification of our approximate solution to the system of equations $Ah = \|r\|_2$. It is computed as:

$$\| A \|_2 = \max_{h \neq 0} \frac{\| Ah \|_2}{\| h \|_2} = \sigma_1$$

Figure 6.5: Equation for the Spectral Matrix Norm [31]

A second measure of the sensitivity of the linear system of equations is the ratio of the largest to smallest singular value in the singular value decomposition of the matrix A. This is knows as the conditioning number of the matrix. Expressed algebraically, $\kappa(A) = \frac{\sigma_1}{\sigma_n}$ [31]. If $\kappa(A)$ is small, matrix A is said to be *well conditioned* – that is to say that the system is not very sensitive to perturbations in h. If $\kappa(A)$ is large, then matrix A is said to be *ill conditioned* and is potentially very sensitive to perturbations in h [31]. What follows are two experiments investigating the relationship between the number of point pair correspondences and the condition number of the image registration. The first experiment is performed on a rigid image registration problem, and the second on a non-rigid image registration.

6.4.1 Perturbation Analysis of a Rigid Image Registration

While the time-space image registration method presented in this thesis was demonstrated by registering x-ray images to an anatomical atlas, the method also supports many different types of imagery. The following analysis of the perturbations in point pair correspondence selection will be performed on the rigid image registration of two digital satellite images. Satellite imagery was selected to ensure a rigid image registration scenario.

Figure 6.6: Rigid image registration performed on two adjacent digital images

# Correspondences	Std. Error	Spectral Norm	Condition Number
5	0.83	3594	1 145 111
8	1.01	4122	937 490
10	1.11	4499	671 698
12	1.19	5834	601 303
24	1.22	9105	599 403

Table 6.4: Number of feature pairs vs spectral matrix norm and condition number

Table 6.4 shows the results for the rigid image registration of the two images shown in Figure 6.6. The image registration was repeated 5 times, each time increasing numbers of selected point pair correspondences. The data shows that as the number of point pair correspondences increases, so too does the spectral matrix norm. Meanwhile, the condition number decreases. This means that the worst case magnification of our solution to the linear system of equations increases with the number of point pair correspondences selected while its sensitivity to perturbations decreases. In other words, the more point pair correspondences that the user has to select the more potential there is to introduce error. However, the more point pair correspondences selected, the less the error in a single point pair correspondence will affect the final solution.

6.4.2 Perturbation Analysis of a Non-Rigid Image Registration

The following analysis of the perturbations in point pair correspondence selection will be performed on the rigid image registration of two non-rigid MRI brain images. The images were selected to ensure a non-rigid image registration scenario. The source image is pre-operative, while the target image is post-operative. As a result, the pixels in the images do not exactly map under a rigid image registration. However, as with the x-ray to anatomical atlas image registrations, rigid registration is sufficient to put the detailed data into context. The pixels that have not changed between scans are registered, while the pixels that have changed locations become apparent.

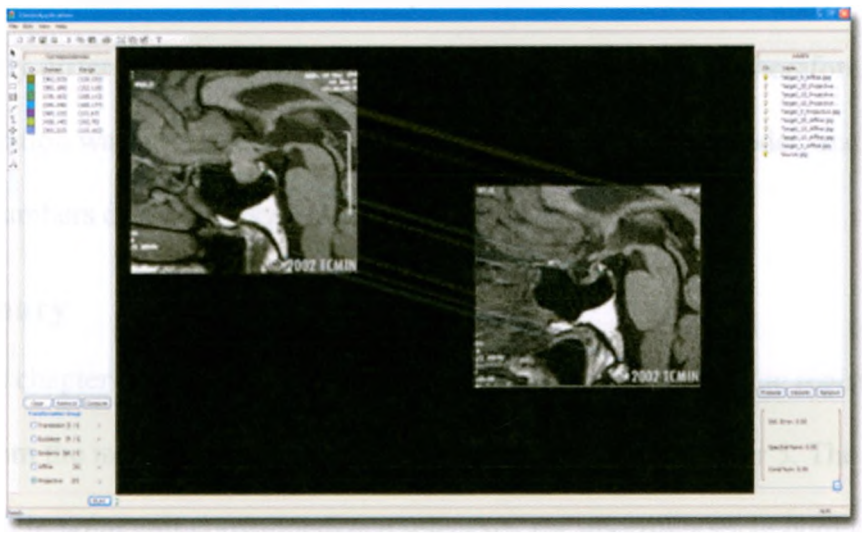

Figure 6.7: Rigid image registration performed on non-rigid MRI images

# Correspondences	Std. Error	Spectral Norm	Condition Number
5	0.82	882	77 126
10	1.54	1 257	24 304
15	1.25	1 554	34 253
20	1.23	1 798	35 558

Table 6.5: Number of correspondences vs spectral matrix norm and condition number

Table 6.5 shows the results for the rigid image registration of the two non-rigid MRI brain scan images shown in Figure 6.7. The image registration was repeated 4 times, each time increasing numbers of selected point pair correspondences. The data shows that, similarly to the rigid image scenario, as the number of point pair correspondences increases, so too does the spectral matrix norm. However, the condition number does not necessarily decrease. This means that the worst case magnification of our solution to the linear system of equations increases with the number of point pair correspondences selected while its sensitivity to perturbations is not necessarily dependant on the number of point pair correspondences selected. In other words, the registration is not only sensitive to the number of point pair correspondences chosen, but also to how geometrically different (distorted) the images are at those locations. Therefore, the non-rigid registration was not intended to be captured by the geometric model, and so the condition numbers do no vary systematically.

6.5 Summary

This chapter has provided an analysis of the uncertainty in image registration resulting from the implementation of the methods presented in Chapter 5. The uncertainties in point pair correspondence selection, the uncertainties in homography estimation, and the error propagation throughout the scene have been measured, presented, and discussed. The uncertainty in the manual selection of point pair correspondences was measured to be approximately +/- 1.08 pixels, the uncertainty in homography estimation was shown to be unique for each image registration, and the errors were shown to grow linearly throughout the scene if the images are registered relative to each other, and not directly to the anatomical atlas.

Chapter 7

Conclusions and Future Work

This chapter presents the conclusions to this thesis by describing what has been accomplished. Subsequently, the potential for future work is presented outlining areas of research that can expand upon this work.

7.1 Conclusions

This thesis has presented a time-space image registration method obtained by combining information visualization techniques, concepts, and principles with the method of atlas based image registration. The method is demonstrated by a software application capable of presenting a time-space narrative of a set of x-ray images which can be geometrically aligned to fit an anatomical atlas. The software application's interface is shown below in Figure 7.1.

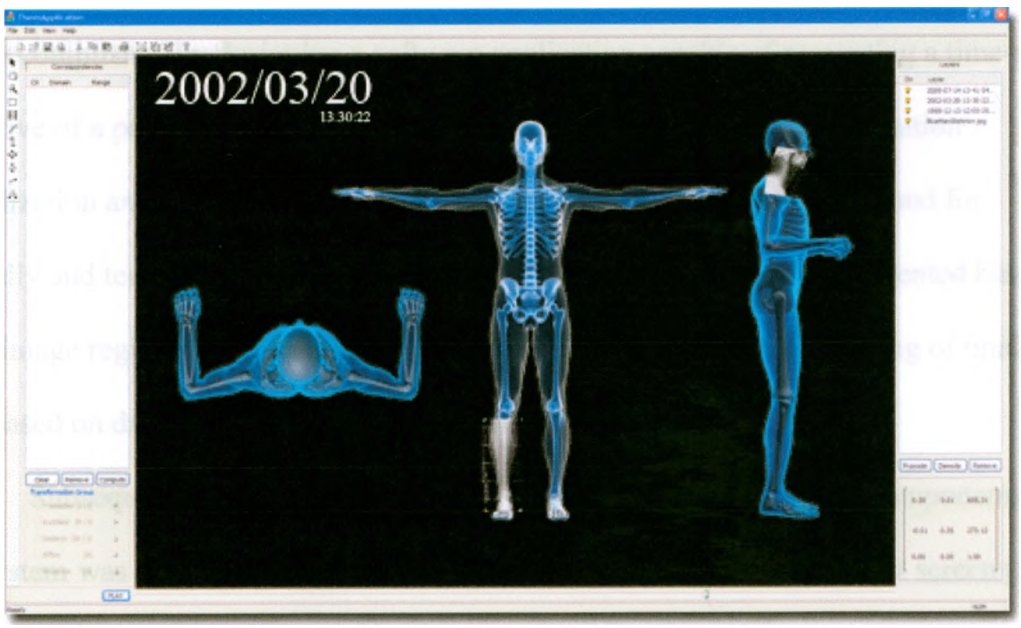

Figure 7.1: A time-space narrative of x-ray imagery

Figure 7.1 shows a visual representation of the time-space narrative achieved by the implementation of the method presented in this thesis. The tools on the right hand side of the application allow the user to select x-ray images, while the tools on the left hand side of the application allow the user to spatially align images based on feature matches or various manual interactions. A timeline is located on the bottom of the application. The user is able to interact with the timeline by choosing a date/time of interest or by playing back the patients medical history. When playing back the patients medical history, the images are overlaid on the anatomical atlas in chronological order. This animation gives the user visual queues about the sequence of the availability of information, while the geometrically aligned images give the user queues about the spatial relationship between the images and the anatomical atlas.

There were several goals that together formed the scope of this thesis. The primary goal was to expand upon research in the field of image registration by injecting techniques borrowed from the field of information visualization and visual analytics. This was accomplished by designing a software application capable of presenting a time-space narrative of a patient's medical records based on techniques used in information visualization and the principles used in visual analytics. To this end, a method for spatially and temporally relating medical images to each other was implemented based on rigid image registration using manually selected landmarks, and the ordering of images was based on data/time stamps embedded within their data.

Subsequent to its implementation, an analysis of the quantifiable components of the system was performed. The analysis showed that users are able to select screen pixels under ideal conditions +/- 1.08 pixels and that the error introduced by the user manually selecting point pair correspondences affects the standard error in the homography

estimation. Based on the point pair correspondences, each image registration will have a unique amount of error. Absolute errors allow us to compare adjacent pair-wise image registrations. As the sequence is extended, the relative errors provide the meaningful comparison. It was also shown that increasing the number of point pair correspondences decreases the sensitivity of the registration to perturbations in rigid image registration scenarios but was not intended to be captured for rigid registrations of non-rigid imagery.

Furthermore, an intuitive user interface to navigate the information space was designed specifically to augment visual analysis. This was accomplished by accommodating interactions based on spatial reasoning, and based on reasoning about time. Time based interactions were enabled by implementing an interactive timeline with automatic playback functionality. Spatial interactions were enabled by allowing the user to manipulate the geometry of individual x-ray images and to register them based on feature matches to the anatomical atlas or to each other. In addition, zooming and panning functionalities were implemented to request different perspectives on the data. Through its visual representation, the software application developed for this thesis allows an analyst to interact with a subject's medical history, to explore it, and to derive insights or understandings that may be helpful for diagnosis.

A drawback of the method presented in this thesis is that since the data is registered to an artificial coordinate system, an unknown amount of distortion can be introduced through image registration. This distortion can exaggerate or understate features in the data that might be pertinent for diagnosis.

7.2 Future Work

This thesis has explored research opportunities in atlas based image registration by incorporating principles of visual analysis and techniques of information visualization. While the implementation of this thesis was successful, there is potential for future work. The method can be extended to support three-dimensional image registrations allowing for three-dimensional anatomical atlases to be presented. Clinical trials can be conducted to determine the effectiveness of visual representations for augmenting the analysis of medical histories and compare them with the current method of analysis. Multimodal imagery could also be supported, expanding upon the software to support the intricacies of each individual modality. Automatic image registration algorithms can be added as they become available to remove the requirement for manual point pair correspondence selection when geometrically aligning images to the anatomical atlas. The method proposed in this thesis can be extended to provide a front-end user interface to navigate image and temporal data. The method in this thesis can also be extended to support the representation of time-space narratives of image encoded scientific data and its changes over time. Lastly, non-rigid image registration methods can expand on this work to provide image morphing capabilities for non-rigid geometric scenarios.

Bibliography

[1] A. Bovic, Handbook of Image and Video Processing, Elsevier Academic Press, Burlington, MA, USA, 2005.

[2] D. Brannan, M. Esplen, J. Gray, Geometry, University Press, Cambridge, United Kingdom, 2000.

[3] M. Brown, and D. Lowe, "Automatic Panoramic Image Registration using Invariant Features", The International Journal of Computer Vision, Vol. 74, No.1, 2007, pp. 59-73.

[4] D. Capel, and A. Zisserman, "Automated Registration with Super-Resolution Zoom", Proc. of the International Conference on Computer Vision and Pattern Recognition (CVPR98), California, USA, June 1998, pp. 885-891.

[5] K. Cook, J. Thomas, "Illuminating the Path: Creating the R&D Agenda for Visual Analytics", National Visualization and Analytics Center, USA, 2005, pp. 3-33.

[6] E. Davies, Machine Vision: Theory, Algorithms, Practicalities, Morgan Kaufmann Publishers, San Fransisco, California, USA, 2005.

[7] O. Faugeras, Q. Luong, The Geometry of Multiple Images: The Laws that Govern the Formation of Multiple Images of a Scene and Some of their Applications, The MIT Press, Cambridge, Massachusetts, London, England, 2001.

[8] M. Fischler, and R. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Application to Image Analysis and Automated Cartography", Communications of ACM, Vol. 24, No. 6, 1981, pp. 381-395.

[9] K. Friston, J. Ashburner, C. Frith, J. Poline, J. Heather, and R. Frackowiak, "Spatial registration and normalization of images", Human Brain Mapping, Vol. 2, 1995, pp. 165-189.

[10] G. Golub, C. Reinsch, "Singular Value Decomposition and Least Squares Solutions", Numerische Mathematik, Vol. 14, No. 5, 1970, pp. 403-420.

[11] A. Goshtasby, 2D and 3D Image Registration for Medical, Remote Sensing, and Industrial Applications, Hoboken, NJ : J. Wiley & Sons, USA, 2005.

[12] R. Hartley, A, Zimmerman, Multiple View Geometry in Computer Vision, Cambridge University Press, United Kingdom, 2000.

[13] D. Hill, P. Batchelor, M. Holden, and D. Hawkes, "Medical Image Registration", Physics in Medicine and Biology, Vol. 46, No. 3, 2001, pp. R1-R45.

[14] D. Lowe, "Object Recognition from Local Scale-Invariant Features", Proc. of the Seventh International Conference on Computer Vision (ICCV99), Corfu, Greece, Vol. 2, Sept. 1999, pp. 1150.

[15] Y. Ma, S. Soatto, J. Kosecka, S. Sastry, S, "An Invitation to 3D Vision: from Images to Geometric Models", Springer-Verlag, New York, USA, 2004.

[16] Massachusetts Institute of Technology, "Singular Value Decomposition (SVD) Tutorial", 2009. [Online] Available: http://web.mit.edu/be.400/www/SVD/Singular_Value_Decomposition.htm

[17] C. Maurer, and J. Fitzpatrick, "A Review of Medical Image Registration", Interactive Image-Guided Neurosurgery, American Association of Neurological Surgeons Publications Committee, pp. 1-49.

[18] P. McLauchlan, and A. Jaenicke, "Image Stitching using Sequential Bundle Adjustment", Image and Vision Computing, Vol. 20, No.10, 2002, pp. 751-759.

[19] D. Milgram, "Computer Methods for Creating Photoregistrations", IEEE Transactions on Computers, No. 11, 1975, pp. 1113-1119.

[20] N. Paragios, Y. Chen, O. Faugeras, Handbook of Mathematical Models in Computer Vision, Springer Science & Business Media Inc., USA, 2006.

[21] H. Park, P. Bland, and C. Meyer, "Construction of an Abdominal Probabilistic Atlas and its Application in Segmentation", IEEE Transactions on Medical Imaging, Vol. 22, No. 4, 2003, pp. 483-491.

[22] D. Rueckert, A. Frangi, and J. Schnbel, "Automatic Construction of 3D Statistical Deformation Models using Nonrigid Registration", in Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI'01), Vol. 2208, Utrecht, The Netherlands, Oct. 2001, pp. 77-94.

[23] A. Šerifoviæ-Trbaliæ, D. Demiroviæ, N. Prljaèa, G. Szekely, and C. Cattin, "Intensity Based Elastic Registration Incorporating Anisotropic Landmark Errors and Rotational Information", International Journal of Computer Assisted Radiology and Surgery, Vol. 4, No. 5, 2009, pp. 463-468.

[24] H. Shum, and R. Szeliski, "Construction of Panoramic Registrations with Global and Local Alignment", International Journal of Computer Vision, Vol. 36, No. 2, 2000, pp. 101-130.

[25] D. Škerl, B. Likar, J. Fitzpatrick, and F. Pernuš, "Comparative Evaluation of Similarity Measures for the Rigid Registration of Multi-Modal Head Images", Physics and Medical Biology, Vol. 57, No. 18, 2007, pp. 5587-5601.

[26] R. Spence, Information Visualization: Design for Interaction, Pearson/Prentice Hall, USA, 2007.

[27] C. Studholme, V. Cardenas, N. Schuff, H. Rosen, B. Miller, and M. Weiner, "Detecting Spatially Consistent Structural Differences in Alzheimer's and Fronto Temporal Dementia using Deformation Morphometry", in Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI'01), Vol. 2208, Utercht, The Netherlands, Oct. 2001, pp. 41-48.

[28] R. Szeliski, and S. Kang, "Direct Methods for Visual Scene Reconstruction", In IEEE Workshop on Representations of Visual Scenes, 1995, pp. 23-33.

[29] R. Szeliski, and H. Shum, "Creating Full View Panoramic Image Registrations and Environment Maps", Microsoft Research, 1997, pp. 1-8.

[30] J. Talairach, and P. Tournoux, Co-Planar Stereotaxic Atlas of the Human – An Approach to Cerebral Imaging, Stuttgart, Thieme, 1988.

[31] D. Watkins, Fundamentals of Matrix Computations, John Wiley & Sons Inc., USA, 1991.

[32] P. Wong, and J. Thomas, "Visual Analytics". IEEE Computer Graphics and Applications, Vol. 24, No. 5, Sept. 2004, pp. 20-21.

[33] J. Woo, J. Slomka, D. Dey, V. Cheng, B. Hong, A. Ramesh, D. Berman, R. Karlsberg, C. Kui, and G. Germano, "Geometric Feature-Based Multimodal Image Registration of Contrast-Enhanced Cardiac CT with Gated Myocardial Perfusion SPECT", Medical Physics, Vol. 26, No. 12, 2009, pp. 5467-5479.

[34] T. Wu, C. Liang, J. Wu, C. Lien, B. Yang, Y. Huang, and J. Lee, "Integration of PET-CT and Cone-Beam CT for Image-Guided Radiotherapy with High Image Quality and Registration Accuracy", Journal of Instrumentation, Vol. 4, No. 7, 2009, pp. 1-5.

[35] I. Zoghlami, O. Faugeras, and R. Deriche, "Using Geometric Corners to Build a 2D Registration from a Set of Images", Proc. of the International Conference on Computer Vision and Pattern Recognition, IEEE, Puerto Rico, 1997.