CONTINUATION REPORT

First Year Summary of
Progression and Outcomes

THE UNIVERSITY of MANCHESTER

ISBE
Imaging Science and Biomedical Engineering

R. S. SCHESTOWITZ
IMAGING SCIENCE AND BIOMEDICAL ENGINEERING
VICTORIA UNIVERSITY OF MANCHESTER
UNITED KINGDOM

DEGREE PROGRAMME: **PH. D. MEDICAL BIOPHYSICS**
PROJECT SUPERVISOR: **C. J. TAYLOR**

2nd August 2004

Contents

List of Figures

List of Tables

Index

Abstract

Autonomous construction of robust statistical models has for some time been an arduous goal to achieve. This is particularly due to the need to establish dense correspondence across the training set. Numerous attempts have been made to automate the formation of good active appearance models, but none has yet been very successful. Potential exists, however, in the unification of model construction and image registration. The evaluation of non-rigid registration, which is based on non-linear warps, has been another subject exhibiting great difficulties and automatic selection of good warps is far from trivial.

In active appearance models, the main problem is the inability to select good landmark points without human judgement, as well as the difficulty in location and annotation of these landmarks using brute-force only. Non-rigid registration is a quickly emerging technique that can be used to warp multiple data instances and produce a group-wise optimal model. Contrariwise, past attempts sought a model which is derived from pair-wise registration and therefore depended on an arbitrary choice of a reference.

These arguments highlight the benefits summoned by the combination of these two techniques – active appearance model can aid the selection of good warps in non-rigid registration and the functionality of non-rigid registration can help obtain more compact and robust models of appearance or deformation, as well as diminish the necessity of manual annotation. This report outlines some previous work in the field and a summary of the current successful progress. It also explains some concepts that

bear potential or whose realisation can contribute to future endeavours and strengthening of the current algorithm.

Work throughout the year made registration using appearance model practical and powerful. The selection of warps is driven purely by the quality of a model and the warping space itself describes the range of legal deformations. It is safe to state that the goal of automatic construction of deformation models, based on registration, has been comfortably approached.

Acknowledgements

Special thanks must firstly go to my supervisor Chris Taylor for investing a great deal of time in the project and allowing me to attend relevant meetings across the country. Without this motivational boost, surely it would have been much more difficult to spend as many hours as I have. My advisor Steve Williams has supplied me with persistent support and occasional bits of advice for which I am particularly grateful. Always with a calm disposition and paternal backing, he was able to relieve my worries and alleviate my doubts. I would also like to thank Mark O'Leary for allowing me to proceed with all my academic duties while I was at work.

The project is funded by the EPSRC and it benefits greatly from its inclusion in the MIAS-IRC where I could at times get in contact with prominent people in the field. I am very thankful for having gotten the opportunity to work on this exciting project and gaining access to some very bright minds and supportive members of staff.

Endless gratitude cannot be neglected for support from the following people:

- Katherine Smith, Carole Twining and Stephen Marsland for putting the foundations for this work.
- Carole Twining, Tim Cootes [WWW-7] and many other individuals in ISBE without whom I would have faced impossible challenges.

- My mother Maya whose overwhelming care motivated me to take up this special opportunity to carry out research.
- My father Gabriel and my grandfather Avner Werner Max for promising and providing full financial support from my very first day in the University.
- Felix Goldberg who has been a distant friend and mathematical guru. He brought up some useful profound discussions at times and has been a true source of inspiration for several years.
- Dr. David Baxter who advised me to obtain a Ph.D. at Manchester University and later recommended Imaging Science and Biomedical Engineering, emphasising its promising vocation.
- Harvey Tobkes and his son Dr. Andrew Tobkes who stressed the potential of imaging science.
- Fred Kupersmith who suggested I pursue a higher level of qualification.
- Dr. Malcolm Taylor for giving me early advice on my Ph.D. in August and September 2003.
- Many tributary thanks to Prof. Tony Hegarty who strongly urged me to join Imaging Science and Biomedical Engineering once the offer was there.

Prologue

"A *representation* is a formal system for making explicit certain entities or types of information, together with a specification of how the system does this. And I shall call the result of using a representation to describe a given entity a description of the entity in that representation..."

"...This definition of a representation is quite general. For example, a representation for shape would be a formal scheme for describing some aspects of shape, together with rules that specify how the scheme is applied to any particular shape. A musical score provides a way of representing a symphony; the alphabet allows the construction of a written representation of words; and so forth. The phrase "formal scheme" is critical to the definition, but the reader should not be frightened by it. The reason is simply that we are dealing with the information-processing machines, and the way such machines work is by using symbols so stand for things– to represent things, in our terminology. To say that something is a formal scheme means only that it is a set of symbols with rules for putting them together – no more and no less..."

– David Marr [37].

Chapter 1

INTRODUCTION

"The pen is the tongue of the mind."
– *Cervantes.*

1.1 Project Background

A suitable way of describing this undertaken research is by briefly describing its aims in a simplistic form that requires limited understanding of its background. Context is key to the understanding of how current knowledge pertains to and contributes to the main hypothesis.

Given a collection of data objects (quite commonly in the form of two dimensional images) which are clearly different although they describe the same object, one wishes to *transform* them in some way so that they appear as identical as possible to one another. The solution to this task cannot be unique, meaning that there will be infinitely many solutions, i.e. transformations, that get similar results. For example, common sense may suggest that two such data objects should be selected each time and, subsequently, one of these objects should be transformed to fit the other. This raises the questions: Which objects should be selected? How should they be transformed? What conditions define a good fit?

One can use an existing technology to *model* all of these images and use this modelling process to minimise a term of complexity. The basic contention is that when this term is minimised, better identity across the set of data is granted and a single unique solution is always reached.

This process above is very beneficiary because one if its byproducts is a description of a group of transformations – the transformations that were used to manipulate data to attain identity. Such descriptions can be used, in a process of learning, to form knowledge about the observed differences in the data set. They can describe how to transform a single data object in a way which preserves prevalent, well-ground variations. They can be used to construct models which are capable of regenerating existing and yet unseen data that exhibits similar properties.

For raw data (as described above) to be modelled properly, knowledge about corresponding patterns and points in the data, must be gained. Thus far, human understanding of the data aided a process of annotation. That process involved mark-up of data regions or points that are homologous. However, once data is made merely identical, mark-up becomes trivial. This is because points tend to lie at identical position. Significantly enough, no manual mark-up of the data is necessary once the approach outlined above successfully works. The questions that remain are: Can data sets be transformed to reach a state of identity? Will the framework of models transcend the peril of data being changed?

1.2 Description of Task

The continuous research work invested in two separate yet related fields calls for a strategic merger which takes advantage of the best of both.

These fields are statistical models of appearance and non-rigid registration whose wide-spread use consequently made them independently usable and powerful. As they deal with problems that have a great deal in common, attempts have been made, and still are being made, to discover how one field is able benefit the other and to what capacity.

The broad field of image registration includes some important techniques that academic and clinical research groups have reasonable interest in [27, 64, 66, 67, 69]. An evident rise can be shown in the number of papers published in the field, medical context being a noticeable focus. It turns out that registration is in many respects highly-applicable to polymorphous bio-medical data as later discussions stress.

Registration is concerned with the assembly of data which is taken either at different points in time or at some arbitrary time instances where changes due to the passage of time can be ignored. Registration sees the most use in scenarios where multiple *different* objects or subjects[1] are being scanned or where the acquisition method varies. In the case of medical imaging, registration is commonly mentioned in one of three distinct circumstances: intra-subject registration, inter-subject registration and multi-modality imaging. This corresponds to the investigation of changes in one specific subject over time, the investigation and comparison between more than one subject and the fusion of data acquired from different modalities (e.g. CT, PET and MRI[2]) respectively. Most typically, however, only a single subject is involved.

The main problem that registration is determined to overcome is the *alignment* of several images with the aim of achieving better *correspondence* across the entire set of images to be dealt with. This quality of correspondence can be evaluated by similarity measures, examples of which are given later (3.2.2 in Chapter 3). With suitable overlap of some given

[1]From this point onwards, there will be a clear focus on 2-D imaging of the anatomical. This narrower view can be considered a case study for image registration and it allows descriptions to be easier to follow.

[2]See list of acronyms and abbreviation in the Appendix on page 188.

object[3] within a group of images, segmentation, analysis and comparison become significantly more straight-forward; these are almost impossible to guarantee in the absence of that overlap. Correspondence is not always simple to achieve algebraically since the object inspected or the aperture[4] may change position and angle over time or acquisition site. In reality, additional unwanted effects such as noise, distortion and change in form must be carefully accounted for. In some real-world applications, biological being an ideal exemplar, variability must be handled sensibly in order to understand the changing structures (as in soft tissue in the brain) that are present in an image. Therefore, the correspondence, as well as the permissible degree of freedom, must not be excessively rigid[5]. It is important to ensure that the chosen analysis mechanism caters for some level of flexibility to enable a rigourous registration process that is immune to high levels of misalignment.

The problem of registration would have been rather simple if it were not for the innate changes that are an integral part of any biological entity, e.g. brain [56], spine, etc. Simple alignment is therefore not necessarily sufficient to give good a solution – that is – plausible correspondence. As explained in Chapter 3 on page 49 of this report, registration methods can be further broken down into different classes, but their aims remain the same in essence. The methods aspire to find some correlation between two or more images, in which case a new entity is obtained that expresses the informative relations between the distinct images.

Image registration is said to be capable of positively affecting the performance of statistical models; possibly this holds the other way around too. More compact (and hence preferable) models of variability can be constructed if registration procedures are applied to its training data (see [39] for more details on learning and training and Section 2 on models).

[3] The word "object" will from here onwards refer to a structure of interest in n-dimensional space.

[4] In the case of medical imaging, there are even more factors to be considered, as opposed to a camera's aperture.

[5] "Rigid" refers to constrained variability and low model generalisability as explained later. It is significantly different from the term "rigid" in the actual context of registration.

This is obvious because registration clearly minimises the witnessed variability, that variability simply being change or difference in the data. The earlier parts of this report, and in particular the next two chapters, attempt to explain and show the commonality between the two techniques, whereas the latter parts explain in greater depth how the two techniques might (and possibly should) come together. It also insinuates that as soon as one can be incorporated within the other, detrimental issues that recur can finally be resolved.

In some previous work, the formation of appearance models, based on registered images, provided a fair indication of how desirable a prior process of registration was. However, the process was slow and therefrom emerged a need to find better ways of using the two techniques in a cunning and hence more efficient manner.

Quite broadly and even wishfully, some current research activities intend to bring together different phases of the handling of an image, from the moment when images are registered to the point where these are coupled with an appropriate statistical model (and even get segmented and measured). Research that this document describes can hopefully form a small part of such a large-scale goal. Arguably[6], it would not be venturous to state that model fitting, shape analysis, non-rigid registration, feature detection and segmentation can and should be put under one *single framework*. At least a few of these might become inseparable in the future.

In this way, by unifying image analysis phases, more compact and powerful representation of images can be used – images can be described by the parameters of non-rigid transforms that ought to generate them from a basal mean image. This is in fact what makes this unification of several methods quite appealing when compared with stand-alone active appearance models where construction is subjective and time-consuming.

[6] Ideas such as this are overly optimistic perhaps.

1.3 Principal Goals

As will be explicated in the later chapters, this project aims to discover a new way of registering data, i.e. aligning a set of images or volumes. Past work has motivated the belief that there are advanced criteria by which registration can be carried out. Not only will such registration be as powerful as desired, but also, as somewhat of a residue, one should be left with an entity expressing the data variation that was observed. What this means in simpler terms is that two disjunct contributions can be made by this work, assuming it ends up being successful. As it currently stands, not only were the goals better realised, but a significant step was made towards their establishment (Sections 7-11 reflect and support this argument).

Just as one would expect, research work involves a learning curve and the development of working relationships. A mentioning of this point-of-view can be found in this document although little emphasis has been put into unnecessarily or detailed bits of information. These are not pre-requisite formal goals, yet they are integrally used to serve the main goals which are purely scientific[7].

In summary, goals have been realised at a much earlier stage and have so far been reached and fulfilled in a way that is debatably beyond satisfactory. This report aspires to prove that this is indeed the case. It is not inclined to concentrate too much on certain aspects such as the daily contributory involvements, but rather show the progress made, the experiments performed and some results and conclusions, of which there are plenty.

[7]This text is goal-oriented and it embraces the technical, not much of the interpersonal and curricular.

Chapter 2

MODELS

"Being brilliant is no great feat if you respect nothing."
– *Johann Wolfgang von Goethe.*

2.1 Introduction

THIS chapter explains in some detail the history, motivation and nature of what is now called active appearance models. Some parts are less crucial for the understanding of this continuation report and they have been included as appendices. They are certainly worth an inspection.

Some of the concepts are better explained and visualised (by the *real* software) in the cited papers although these may be out-of-date. It is therefore worthwhile having a look at the relevant World Wide Web resources[1] which are gradually modified, e.g. the pages of Tim Cootes [WWW-7].

[1]All Web resources are listed at the end of this report.

2.2 The Approach

Image analysis is a general problem that can be tackled in various ways. This analysis is fundamental and essential to many processes such as industrial inspection, motion analysis [68], face recognition and medical image understanding. What makes this problem intrinsically laborious is the inability to take into account single pixels independently to infer the structure they form together, cohesively. The goal of such interpretation or analysis is not only to tackle the problem correctly, but also to do so efficiently, in a way that will not be overly affected by the size of the image, i.e. not reliant on the scale of the problem.

Analysis often involves *measurements* of meaningful structures in an image and possibly some explanation regarding the *form* of these structures. In order to derive any adjuvant information about a particular meaningful structure, image *segmentation* must first take place. Segmentation is concerned with the identification of certain regions of interest which may be characterised as belonging to the same object. By dividing the image into such regions, understanding of the nature of its constituent components can almost instantly be gained.

This report concentrates on a top-down approach to data[2] analysis. The approach relies on a high-level abstraction of the visual attributes of one structure. Alternatively, and often more usefully, this abstraction can represent a *collection* of structures that together form another aggregate structure. The reason why such an approach is referred to as a top-down approach is that it contains some existing information that it attempts to *fit* to the problem posed[3]. It makes assumptions about the problem and is in some sense taking a preliminary overview on the structures in an image as Figure 2.1 illustrates.

[2]The chapter considers images to be the default case. These methods are usually operable over an arbitrary number of dimensions, but 2-D proves to be easier for a reader (and the paper) to visualise.

[3]A bottom-up approach will look at low-level data and build up towards knowledge of higher complexity which has a meaning. Top-down is an opposite approach which 'knows' what it tries to find so it searches for a best lower-level match.

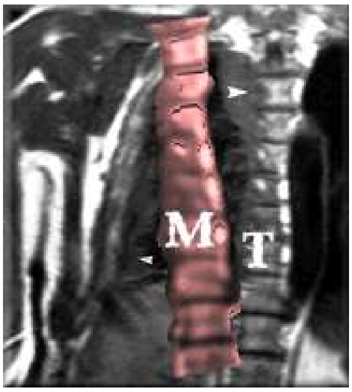

Figure 2.1: A target image T is being overlaid by a high-level representation (the model M) which seeks to find a good fit.

The rest of this chapter will describe popular methods of top-down image analysis, but will focus on active appearance models at the expense of other, less relevant methods.

2.3 Statistical Models

The next few sections explain in some depth the notion of *statistical* models and especially that of (statistical) appearance models. They move on to the description of active appearance models which are an extension to active shape models and a brief introduction to shape models may be worthwhile to begin with.

Given a collection of images depicting an object which possesses some innate properties, it is then possible to express the visual appearance or shape of that object in a way that discards subtle changes in view-point, object position, object size et cetera and is robust to some level of object deformation. That object which appears in the group of images need not even be the exact same one; it can be an object belonging to one common *class*. Some variation that is typical for that class can be handled (essentially be understood) reliably with the help of elementary transformations (to be described in 3.2.1), but their functionality is inevitably very limited and constrained.

There are statistical methods which allow the encoding of the variability which was *learned* during a so-called training process. That training process does not require far more than an exhaustive inspection of the set of images where objects (or shapes) appear. However, in order to interpret a large set of objects, some simplification steps are required. This results from the fact that most images where objects lie are expected to be of relatively large-scale in practice – certainly large enough to result in an exponential blow-up[4].

A method is sought which reduces the amount of information that is required to describe an object of interest and the different forms it can take. This is done by selecting points of interest which lie in the image – ones which will be a representative sub-set of the image contents[5]. Points must be picked so that they jointly preserve knowledge regarding the object of interest. That object is often well-hidden in that pool of image pixels. Such points are often chosen to become what is entitled *landmarks*. Landmarks are positions in the image which effectively distinguish one object from another in the set of images (see Figure 2.2 on the next page). They also have some interesting spatial traits which can form near-optimal curves (or contours) which together make up genuine *shapes*. The concatenation of the coordinates of these landmarks can then describe an image (or rather the object being focused on) in a concise and useful representation. In 2-D, for n landmarks, a vector of size $2n$ can infer the shape of the object present in an image. This lossy inference can be described as follows:

$$(x_1, y_1, x_2, y_2, ..., x_n, y_n) \Rightarrow \text{S} \tag{2.1}$$

where S is simply a discrete reconstruction of the shape in the image. It is *not* the actual image.

[4]Current model-based methods typically deal with only the order of tens of thousands of pixels. High-resolution medical images can contain millions of pixels.

[5]In most cases, edge detection is sufficient to capture regions or points of greater significance in the image. Edges and corners usually hold more information of use for subsequent analysis and aid segmentation. They lead to better identification of the different objects residing in the image.

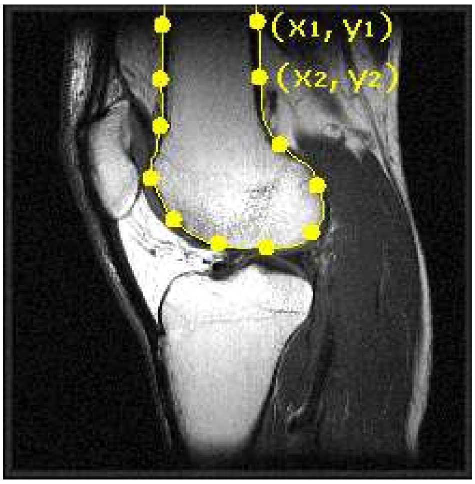

Figure 2.2: Landmark identification and mark-up in medical imaging.

It is worth pointing out that landmark points can be chosen *arbitrarily.* This turns out to be a serious issue as will be seen later along with possible solutions. Identification of objects is in most cases[6] done by drawing lines or selecting surfaces which surround these objects. Given continuous elements such as a lines or surfaces, by no criterion does it become obvious how to suitably sample them using points. The choice of points affects the quality of reconstruction as measured by the assigned errors.

With the concise landmark-based representation (described above in 2.1) set to be the convention and a collection of fair-sized vectors rather than a massive collection of images, it should be possible to express (in a feasible way) the legal range[7] of each one of the vector components. This in essence establishes the *model*. It is an entity that can be manipulated to reconstruct all the shapes (or as later explained – images) it originated from and far beyond that. This model encapsulates the variation which was learned from the data and it usually improves its performance as more legal examples are viewed and 'fed' to support some further training. Varying the parameters of the model can generate new (unseen)

[6]Contrarily, analysis of mammograms needs to account for texture as well as shape. The boundaries of a breast are not sufficient to characterise the distinguishable data that is of most interest.

[7]The legal range can be thought of as the values a parameters may take. In reality, a Gaussian distribution usually fits the observed range rather well.

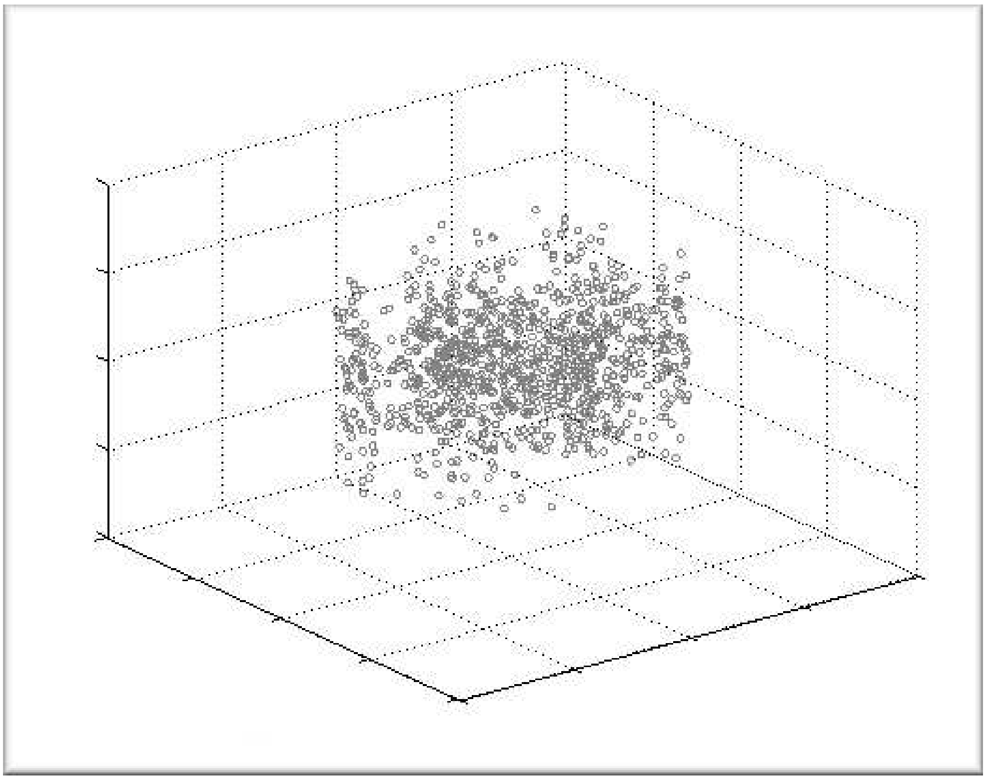

Figure 2.3: 3-D scatter of points.

examples as long as that value variation is restricted by the legal range, as learned from the training examples. The vector representation mentioned beforehand can be also looked at as a description of a fixed location in space that comprises d dimensions (see illustrative scatter in Figure 2.3). This turns out to be a useful demonstrative idea as will be seen later when dimensionality reduction is applied.

Shape models are "statistical information containers" which can be built from the images with overlaid landmark points identified and recorded. In order to make such a mechanism possible, it is vital to firstly achieve consistency amongst the coordinates of all landmarks. This means that all points need to be projected onto a common space – a process whose purpose is to ease collective analysis. That process can also be thought of as an alignment step which somehow links to the next chapter. More issues that are concerned with normalisation, projection and the like are described in slightly more detail later in this document.

A human expert usually performs annotation or landmarking of the images with the aid of some computerised special-purpose tools. In recent years, alternatives which are automatic showed great promise [10] and these extend to 3-D too [12]. The later chapter on page 67 is dedicated purely to that one piece of work which is so fundamental to this current new research.

Appearance models were later developed by Edwards *et al.* [17, 8] and the greatest advantage or essence of these was that they were able to sample grey-level[8] data (incorporation of full colour has been made possible by now, e.g. Stegmann *et al.* [52], [WWW-5]) from images rather than just points. Therefore, appearance models retained information about what an image *looks* like rather than just its *form* as visualised by contours (or surfaces in 3-D). Just as points in the image were earlier chosen, grey-level values (also referred to as *intensity* or *texture*) could be systematically extracted from a normalised image and preserved in an intensity vector for later analysis. This normalisation process and the representation of this intensity vector will be outlined later in this chapter.

What enables appearance models to exhibit quite an astonishing graphical resemblance to reality is that at the later stages of the process, a *combined* vector is made available. It incorporates *both* shape and intensity while keeping aware of how change in one affects the other (e.g. how expansion results in darkening and vice versa). Hence it has a notion of the *correlation* between the two – a notion that is dependent on the training data and Principal Component Analysis. Although appearance models are usually not as quick and accurate as shape models[9], they contain all the information that is held in the shape models and in that

[8] Colour can be simply thought of as an extension of the single grey-scale band being divided up into red, green and blue components. There are different possible colour schemes [41] which have no affect on the actual principle of intensity sampling.

[9] In principle, they (appearance models) can be made just as powerful, but in practice they suffer from requirements for high speed. As this text shall later explain, they can sometimes lead algorithms to getting trapped in local minima.

sense are a superset[10] of shape models. Also, some techniques have been developed and employed to speed up the matching of appearance models to image targets (see later in Section 2.7 and Appendix A). Tasks such as the matching of an appearance model to some target image are described later in this chapter and illustrated in [7].

2.4 Model Construction

An interesting and integral part of appearance model are their construction (or formation) step. The first step is concerned with the establishment of a model that not only describes a *mean* form of some object in an image (if not the image as a whole), but also the legal variation that can be applied to that mean in order to create new legal object instances. A model formulates the form which vectors can take and these vectors can easily be translated into a visual description. More desirable models will not be excessively data-permissive. They should allow recognition and acceptance of only reasonable variations of the object under investigation. There is a convenient mathematical way of expressing this variation and that is to assign a parameter to each mode of variation[11]. When change in these parameters occurs and the mean is deformed accordingly, there will be a direct effect on the appearance of the result. Rather usefully, each legal instance can always be uniquely and fully described by the parameters which were used to generate it from the model. The synthetic appearance and its vector representations are equivalent and inter-changeable. Visualising results is often convenient graphically, while logical operations are better thought of in terms of vectors.

[10] They can be thought of as a superset, simply being shape models which hold some additional information and the correlations between all encapsulated data.

[11] Offsets of standard deviation units from the mean of each mode then illustrate the effect each variation mode has.

2.4.1 Shape Model

To begin encoding the form of an object, landmarks need to be identified and statistical analysis applied so that it expresses these spatial shape properties, namely the landmark coordinates. From this analysis, a mean shape is obtained and it can be denoted by x_{mean} or $\overline{x}$. To obtain this mean, the procedure that is commonly used is Procrustes analysis. The generalised Procrustes procedure (or GPA for Generalised Procrustes Analysis) was developed by Gower in 1975 and has been adapted for shape analysis by Goodall in 1991. It processes each component of the vectors derived from the images and returns for each component a value that is said to be the mean. From here onwards, this vector which represents the mean of the data will be referred to as $\overline{x}$. Each shape x is then well-formulated by the following:

$$\mathbf{x} = \overline{\mathbf{x}} + \mathbf{P}_s\mathbf{b}_s. \tag{2.2}$$

The matrix P represents the Eigen-vectors of the covariance matrix (set of orthogonal modes of variation) and the parameters b_s control the variation of the shape by altering these modes of variation. The parameters essentially describe the magnitude of the covariance of each element in the matrix. These parameters and the range within which they must lie describe a level of freedom – that is – the freedom (or otherwise constraints) of the model.

Eigen-analysis is used quite extensively in the derivation of the expression above[12], but it will not be discussed in detail in the remainder of this report. Instead, a short explanation will be given on Principal Component[13] Analysis [26, 60] which from here onwards be referred to as

[12]The process which was proposed by Manfred Eigen allows the calculation of Eigen-vectors and Eigen-values. For a given matrix, Eigen-vectors describe directions in space that are derived from the matrix and the corresponding Eigen-values describe their magnitude.

[13]Plainly speaking, PCA only picks up n Eigen-values whose Eigen-values are the greatest.

PCA. What is worth emphasising is that the only variant in the model described above is b_s and as the values of $b_{1<i<s}$ are infinite ($b_{1<i<s} \in \mathbb{Z}$), the same must hold for x. There is an infinite number of shapes, each of which can be generated from one choice of value for each model parameter. One interesting alternative to PCA was presented in [25] By Jebara. It is explained at the end of Appendix A on page 163.

2.4.2 Intensity Model

The next stage involves the sampling of texture. In principle, having got the description of some shapes from a set with their given spatial correspondences, it is possible to estimate homologous points in between these correspondences. This essentially allows the prediction of the *denser* correspondence – that which involves larger sections of the image, rather than points only. Below lies a description of one special case; the descriptions are aimed to illustrate one possible way of sampling intensities. Construction of an intensity model is the more significant step which is carried out in the exact same way as was done for shapes (Equation 2.2).

At this stage, each of the images should be aligned to fit a common volume in space[14]. In practice, the properties of that space are implicitly defined by the mean shape[15]. Rigid (or Euclidean similarity) transformations, namely translation, scale and rotation, are not always sufficient to warp all images into that common space, e.g. in the ubiquitous case of human faces, different head sizes and facial expressions introduce difficulties. Nonetheless, it is crucial that good fitting is obtained before the sampling of grey-level commences. Following these basic transformations which align all images, the displaced control points of each image overlap and contain in between them shape-normalised patches. These patches are available for construction of texture vectors. Barycentric arithmetics, known for their frequent utility in computer graphics and stereo vision, are used to describe the location of all corresponding points within a

[14]Normalisation step as such is similar to the mapping onto a sphere, for instance.
[15]Oftentimes, the choice of the mean shape proves to be the least damaging choice.

patch[16]. This location of point is directly affected by the warps applied to shift a given shape onto the space of the mean shape.

Triangle meshes are subsequently created by stretching lines between neighbouring control points and intensity values are captured one by one (along a chosen grid of points to be sampled) and stored in a vector representative of texture. Each component in such a vector captures the intensity (or colour) of one single pixel as was learned from the examples. Statistical analysis, which is not different from the one above, results in the following formulation for texture:

$$\mathbf{g} = \overline{\mathbf{g}} + \mathbf{P}_g \mathbf{b}_g. \tag{2.3}$$

It is again worth the while to emphasise that the process if no different to dimensionality reduction in the case of shape. The use of the algorithm above implies that for short vectors and a low number of pixels sampled, noncontinuous appearances will be easy to spot[17]. In fact, objects will often appear to be nothing more than a collection of polygons that do not quite resemble realistic appearances[18]. To compensate for this, algorithms from the related field of computer graphics can be used, e.g. Phong and Gouraud shading. In practical use, geodesic interpolation [WWW-12] is used and the results can be quite astounding considering the low dimensionality of the available data. Compression here is dependent upon reconstruction strategy and assumptions about natural phenomena.

The models above (Equations 2.2, 2.3) are expressed linearly and quite compactly – a highly desirable and manageable form. This is due to PCA which reduces the length of the vectors describing shape and texture.

[16]It is helpful to think of two different triangles and the relationship between points within these triangles. Centre of gravity (COG) is used here to assign approximate correspondence.

[17]Analogically, in the case of shape, sharp-bended descriptors result from the low number of sample points.

[18]One of the main aims and great power of appearance models is full synthesised portrayal, so photo-realism is at a premium.

As earlier mentioned, although Eigen-analysis is involved in the process, its derivation, proofs, or characteristics are less than essential for the understanding of PCA which works as follows.

2.4.3 Principal Component Analysis

It is possible to visualise the data as points in a high-dimensional space as was earlier argued. By placing all images in that space, it is expected that some cloud of points will be present at a specific, though somewhat confined, region. The breadth of this region or the size of that cloud will depend on the variation amongst the images (or more generally data) that is being visualised. PCA relies on Eigen analysis to obtain the Eigen-vectors and Eigen-values of that cloud of points. The highest Eigen-value will correspond to the most significant Eigen-vector (see the single-headed arrow in Figure 2.4). It indicates the direction which best distinguishes the image data and is expected to be the longest one too – that is – the one whose magnitude is the greatest[19]. This is in fact what is considered to be the principal component which describes that data.

In a recursive manner, at each stage of the process, the current principal component is virtually saved and put aside until only negligible components remain present. The recursion will therefore deal with simpler, more uniform data. More and more principal components are set aside and leave a data of lower dimensionality that occupies a relatively low volume in space. A smaller number of components can then be used to express the variation up to a comparatively high level of fidelity. The process is lossy, but so are some other stages in model construction including the choice of a finite number of landmarks. That loss is controlled in the sense that one can choose the minimal amount of variation that must be

[19]If one thinks of the cloud in n dimensional space as a placement of characteristics $(c_1, c_2 ... c_n)$, the principal component is one characteristic which best separates instances of the data. It takes the largest range of variation. To simplify the concept, it can be assistive to think of a standard keyboard. The number of key will poorly distinguish one keyboard from another, but since the names and labels of manufacturer are diverse, this may as well be the one 'principal component'.

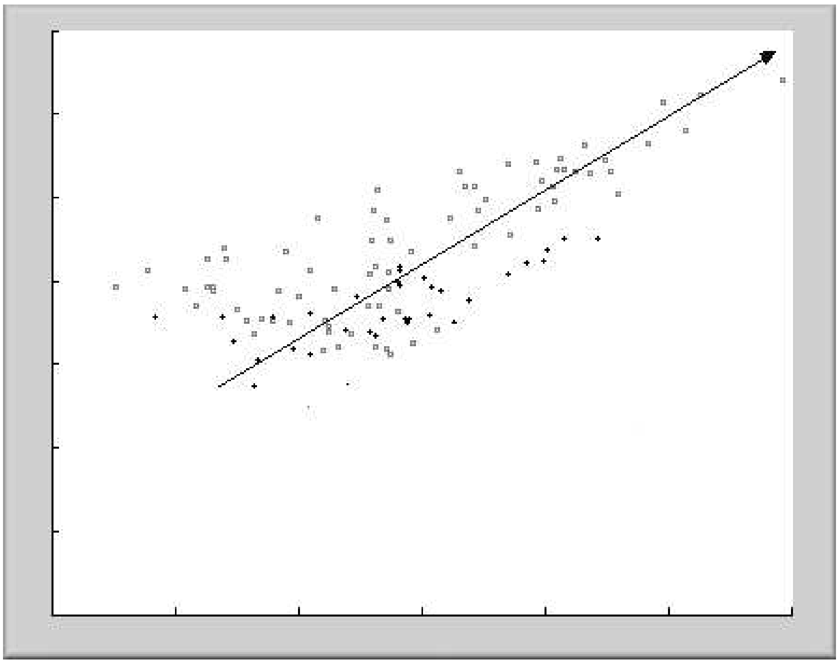

Figure 2.4: Principal component in 2-D is indicated by the arrow.

accounted for[20]. PCA is used to gain speed while retaining the best descriptors of variation or difference in shape and intensity. What this all comes down to is the acquisition of a model that is smaller in size and is easier to deal with. It is easier to deal with because: **(1)** it is smaller; **(2)** it is quicker to use and **(3)** some of its attributes are decomposed.

2.4.4 Combined Model

The two components x and g (the vectors above which are a function of generative models) need to be merged to establish a new model. That more expressive model accounts for both types of variability (shape and intensity) and holds within it the correlation between the two.

The parameters $b_s$and $b_g$are aggregated to form a single column vector

[20]A common choice is 98% of the observed variation which means 2% of the variation is not accounted for. This 2% of variation is usually the least informative though – being exactly what PCA is intended to accomplish.

$$\left\{ \begin{array}{c} \mathbf{b}_s \\ \mathbf{b}_g \end{array} \right\}. \tag{2.4}$$

It is in some sense, a simple concatenation of the two. However, since the values of intensity and shape can be quite different in their nature and granularity, some weighing is needed to attain equilibrium under which both shape and intensity reserve a noticeable effect. The danger is that if no weighing of any sort is applied, intensity values may supercede these of shape or vice versa. In less practical terms, if the extent of data values differs greatly, then the spread of the points in space is quite undesirable. The components to be identified by PCA are not as beneficial as they otherwise would have been. If some values are far greater than others, point vicinity takes a turn for the worse and the cloud might be elongated instead of nearly spherical (lending a 3-D analogy)[21]. For rather spherical spreads (or those of almost homogeneous variation), a greater number of large components will be available for selection. Consequently, the variation expressed by a fixed and constant number of principal components will be higher.

A weighing matrix that resolves the problem introduced above is by convention named $\mathbf{W}_s$ (the symbol s corresponds to *shape* as by default this matrix scales the shape parameters only. It gives logically equivalent results to these of applying the factor $\mathbf{W}_g = \frac{1}{\mathbf{W}_s}$ to intensities). The form in which coordinates are stored in x depends on the accuracy required (e.g. integers and floating-point numbers), the image size and the number of dimensions, whereas for grey-level values, this form is dependent on the number of allocated bits per pixel[22]. With weighing in place, the aggregation would take a form such as

[21]As an ad-hoc example, intensity frequently takes values in the range $0..255$ whereas normalised shape coordinates lies between 0 and 1 so fractions such as $\frac{1}{255}$ can be used as coefficients. The two should then scale almost indifferently.

[22]For colour it is common to use 24 bits and for grey-level just 8 bits. For more compact statistical appearance models, less than 8 bits (256 shades of grey) might suffice to achieve good results and in medical imaging 12 bits are nearly a standard in acquisition.

$$\left\{ \begin{array}{c} \mathbf{W_s b}_s \\ \mathbf{b}_g \end{array} \right\} \tag{2.5}$$

where $\mathbf{W}_s$ is chosen to minimise inconsistencies due to scale. Lastly, by applying a further PCA stage to the aggregated data, the following combined model is obtained:

$$\begin{array}{l} \mathbf{x}_i = \bar{\mathbf{x}} + \mathbf{Q}_s \mathbf{c}_i \\ \mathbf{g}_i = \bar{\mathbf{g}} + \mathbf{Q}_g \mathbf{c}_i \end{array}. \tag{2.6}$$

The appearance (shape and brightness levels) is now purely controlled by the parameters $\mathbf{c}_1, \mathbf{c}_2, ..., \mathbf{c}_n$ and there is no need to choose values for two families of distinct parameters as before. This combined model has the benefits of the dimensionality reduction performed, which is based on shape as well appearance. This means that it now encompasses all the variation learned and the correlation between these two distinct components. Since PCA was applied, the number n of parameters $\mathbf{c}_i$ is expected to be smaller than (or in extremity equal to) the number of parameters in $\mathbf{b}_s$ and $\mathbf{b}_g$ put together.

2.5 Model Training

This part of the the chapter is concerned with searching and fitting, also referred to as active model training. It explains how to *use* appearance model, though it can be discarded if no such functionality is necessary. Training will be dealt with first and fitting which is a closely-related subject will be explained in the next long section.

A descriptive statistical model is now available for utility and various analyses. That model is a type of flexible deformable [18] entity that can describe any instance of object or image[23] in the range of the training set[24]. Assuming that the training set was infinite in size or comprised all possible instances that the model might be presented with, it should then be considered a powerful, fully compatible, flawless model.

2.6 Model Fitting

To motivate model matching or fitting, one can argue that the previously constructed model involved a learning process which must somehow be exploited. For it is now known what objects of some type look like, it is possible to recognise and capture new objects of the same type.

It is still not trivial in any case how one should deform the model to achieve an appearance instance that is valid. It is now a completely opposite problem that users of this model can be faced with: how can one model generate new instances after similar existing instances generated that one model? In some sense, an inverted operation is needed so that the model can be used in the opposite way to the means by which it was created. Things are not very simple in reality and the alteration of model values needs to be guided by some minimisation (some of the next chapters elaborate on this with practical examples, e.g. Chapter 5) that obtains the matching which is being sought. Unfortunately, in an expectedly high-dimensional space as above, the process is almost endless unless extra knowledge about this minimisation problem is provided and in advance and utilised.

[23]The distinction here is hard because the model can describe more than one valid independent object and usually represents only a partial section of the entire image. In a medical context, the term *atlas* fits somewhat more nicely and it usually describes a single organ or anatomical structure.

[24]There is a subtlety which makes this phrasing a bit deceiving and inaccurate. The word "range" is a gross terminologically equivalent to the area that stretches in between the space of training set instances. It can be conceived as the space defined by a Gaussian distribution cloud that is deduced from the training set.

2.6.1 Learning the Correlations

The way in which this problem can be circumvented quickly involves learning how the parameters $\mathbf{c}_i$ affect the model[25] with respect to a typical target. Each parameter in $\mathbf{c}_i$has an unequalled effect on different regions in the model, e.g. its size, intensities and so on. By changing the value of each such parameter and recording the change that is perceived in an image (using pixel-based comparison of some kind), a type of deformation index can be maintained. This index indicates which parameters should be changed and if so in what way and to what degree in order to approach good overlap between a model and some target image.

More formally, the procedure works as follows:

For the model parameters $\mathbf{c}_i$ where $1 < i < n$, a parameter change δc (where one parameter value or more can be readjusted) is applied to generate some new shape and texture. δc expresses in a vector-based representation the offsets that each of the original parameters $\mathbf{c}_i$ is subjected to. The exhaustive pixel-wise difference in intensity[26] is calculated in accordance with:

$$\delta\mathbf{I} = \mathbf{I}_{model} - \mathbf{I}_{image} \tag{2.7}$$

to produce a new vector of intensities (the differences). This vector can also be visualised to display this difference to a human eye. A simple measure of difference is used although this need not necessarily be the case. Sum-of-squares of the pixel differences is then used because larger quadratic differences will have a greater effect on the final measure and summation then only consists of positive values. For example, observe the values derived in Equation 2.9 and in 2.10. The former shows how the

[25]There are some more complex considerations as the model needs to be aligned properly as well as change in its form.

[26]A simple raster scan that account for all pixels should clearly be fast under most contemporary computer architectures. This is indeed the case if simple mundane operations like subtractions are pipelined on the ALU.

values of the vector in 2.8, and particularly their summed difference, get accentuated, whereas in the later case makes them almost negligible[27].

$$\delta\mathbf{I} = sumofsquares(\{-1, 3, 5, 2, 6, -10, -1\}) \tag{2.8}$$

then becomes

$$\delta\mathbf{I} = sum(\{1, 9, 25, 4, 36, 100, 1\}) = 176 \tag{2.9}$$

as opposed to

$$\delta\mathbf{I} = sum(\{-1, 3, 5, 2, 6, -10, -1\}) = 4. \tag{2.10}$$

With this measure of intensity difference recorded, relational information can be expressed between the parameter change and this difference as it appears in image space where a model is superimposed on some target. That information (merely a correlation) can be learned by using a pseudo-target image which is the model in its mean form. It can be used for basic comparison that infers something about the model displacements and their corresponding effect[28].

This quantitative measure of difference obtained will however indicate solely the approximate "goodness" of the parameter change (as inferred from SSD or MSD) and not an overall focalised effect that it has on the given image. This means that it will not necessarily be obvious what parts in the two entities (model and target) remained similar and which ones did not[29]. A type of a sequential data such as a vector is hence

[27]This is reminiscent of the need for a median measure, where average is sensitive to erratic values or salt-and-pepper noise.

[28]It is possible to learn the properties of rotation, as an exemplar, by applying a rotation and looking at the difference between the resulting image and the original image. That is the main concept that this step is based upon, namely inferring $Transformation \iff Error$.

[29]The vector's distribution of values, i.e. positions with high absolute values, can answer this question quite grossly.

more useful as it retains the location of each computed difference value. Unsurprisingly, this also consumes far more space (and many vectors of this kind will in fact be necessary).

In either case, under the premise that space is more expendable than time complexity, a vector of difference is calculated and the correlation can be formulated as follows:

$$\mathbf{c}_i \rightarrow \mathbf{c}_i + \delta\mathbf{c} \rightarrow \delta\mathbf{I} \tag{2.11}$$

This type of offset δc that was applied to the collection of parameters $\mathbf{c}_i$ is accompanied by a global change in intensity values across the image frame. This correlation can now be stored aside and become accessible from an index as its size is proportional to the image size. The storage is dictated by the following (somewhat artificial) relation:

$$\delta\mathbf{c} = \mathbf{A}\delta\mathbf{I} \tag{2.12}$$

where A is a matrix[30] recording the change in intensities due to the parameter/s change δc. This is a type of matrix which is correspondent to an n-dimensional vector that expresses the change which was discovered off-line. It linearly defines (in a possibly high-dimensional space) the linear relation between change to the parameters and change to the intensities, or more precisely the *difference image*. It can be used to choose directions of change directly when performing a search and thereby avoid re-computation in a virtually recurring and almost identical problem[31].

The most fundamental (and perhaps even compact) procedure will carry out the steps above for each of the modes of variation, as well as the linear geometrical transformations. This can be a very laborious and

[30] The matrix **A** can be obtained using linear regression.

[31] The problem is nearly identical owing to one basic assumption – the assumption that similar objects are examined with some known pattern of model placement in the target image. The location of mismatches (indicated by high difference values, i.e. white) tend to show where supplemental model deformation is yet necessary.

cumbersome process although it depends on the robustness prescribed. As the next stage illustrates, models that are not rich enough will fail to converge in difficult scenarios, a classic example of which is inappropriate initialisation.

The matrix A holds real-valued numbers (preferably of limited accuracy to decrease space requirements and access speed). The values in this matrix form a 'path-finding' map that guides exploration for good parameter changes; this will be of great use when fitting the model to a target. In practice, such matrices are visualised by showing negative values as dark shades and positive one as increasingly brighter values.

2.6.2 Target Matching

The final stage, which is arguably the most fascinating one, involves the use of the model above, as well as the correlations learned and recorded for that model. It is possible to carry out a search which is driven by the calculated difference between the model and a given target image. In pragmatic terms, this means that fitting of the existing model will slowly be improved until the model approximately covers the target[32]. It is all done purely by changing the values of the model parameters. The model state, having explored many false states, then holds (in the form of parameter values) some information about the target image and this information can be further analysed. One parameter in a model of faces, for example, could describe the vertical angle of given faces. This is also where the power of a statistical model lies – being able to describe something compound in a very compact form.

The search for model match is reliant on error (or conversely similarity) measures which are repeatedly calculated after each attempted parameterisation of the model. Having applied some change to the parameters,

[32]This process of fitting strives to converge to the global minimum (of difference measure). Realistically speaking, the model and the target never reach complete equivalence, namely the difference value of absolute 0. Even if the target was used to train the model, PCA would corrupt the obscure the connection between the two.

a new estimate of difference is obtained. Each such change in parameter values is primarily guided by the matrices described on page 42. These express the correlation between variation modes (the similarity transformations as well as modes of appearance change) and the intensity values which describe difference (or match discrepancy).

The model, as shown in Figure 2.5 (or earlier on in Figure 2.1 on page 26), is initially placed somewhere inside the image frame, with reasonable proximity to its target. If the model is placed too far from its to-be target, there is a danger that it will be unable to converge to the target correctly. It will most likely get stuck in a local minimum (the global minimum being out of reach as Section 5.5 explains) and the outcome can be severe in a more crucial practice such as medical imaging (or perhaps more drastically, computer-guided or -aided surgery). The reason why good initialisation is essential is that significantly large displacements are rarely learned off-line and the difference between the target and the model is quite meaningless unless there is at least some partial overlap or commonality.

The algorithm which is used to perform the search quite rapidly has a general form that resembles the following:

- Place the appearance model **M** somewhere in the image, preferably at the centre where the target of interest (to be denoted by **I**) is likely to lie[33].
- For the appearance model in its current state and the static target, perform the following:
  - ✧ Calculate the differences between the model and the target. This can be done by synthesising **M** and calculating **M - I**.

[33] Advanced knowledge about the problem is highly conductive at this stage, otherwise some bottom-up image analysis is a must.

- ✧ Using the correlations learned off-line[34], set new values for the parameters $\mathbf{c}_i$ of $\mathbf{M}$.
- ✧ Compute the new difference measures between the model and the target (as previously).
  - ❏ Save the new state of the appearance model if the difference has been lowered, i.e. similarity is being approached.
  - ❏ If not, try re-adjusting the parameter change, potentially with inclusion of a scaling coefficient $k = 1.5, 0.5, 0.25$ and so forth. This often achieves good results, although it is a heuristics-driven technique.

- Iterate while no convergence has been reached and improvements are still observed at times.

More advanced methodologies and algorithms are used at present, but better clarity is achieved by adhering to simplicity.

The technique of matching an appearance model to a target image is well-depicted by a staged simulation, a video clip or a large sequence of images resembling the one in Figure 2.5. Somewhat remarkably, only a few dozens of iterations are required in order to get good matching outcomes. This of course depends on the algorithm, the magnitude of the problem and its innate involution.

As a superficial example, fitting of a perfectly round ball versus a human hand is an interesting problem. Assuming that there is a good contrast between the ball and the background, there should be few false alarms for good fits. An inspection of the difference image is then almost trivial for human appraisal in this case, while fingers become deceiving in the case of hands. In accordance with these very same arguments, the process of correlation-learning should often be custom-built. It should at least treat

[34] If these are not available, some guessing would be an alternative. It is important, however, to learn from the experience gained during this independent run of the program or else the optimisation would behave senselessly and lead to improvements being identified very slowly. General optimisers ought to make a good judgement as such.

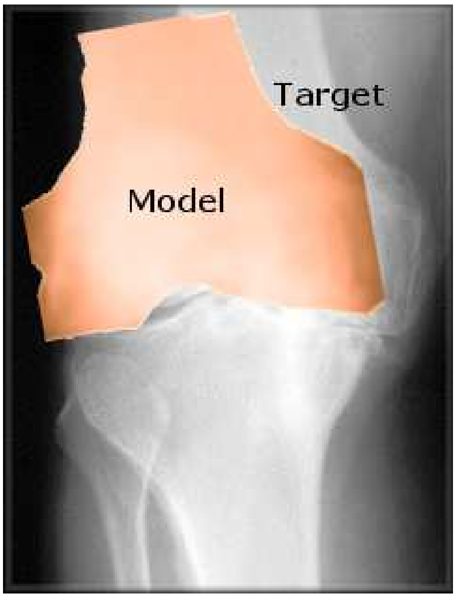

Figure 2.5: Model and target fitting.

the problems with respect to its complexity because sensitivity to change and matching (much like recovery) abilities vary greatly in reality.

2.7 Existing Extensions

The existing extensions to shape and appearance models are numerous and their purpose varies. Wavelet compression techniques are used to mitigate the troublesome space requirements (especially in 3-D, e.g. for analysis of brain volumes). These can also make active appearance models far more compact[35].

There are also some application-specific extensions such as the implementation of view-based models [13] and coupled-view models [14] for face recognition purposes. The principal idea is that 5 different models can express full appearance irrespective of the wide range of viewing an-

[35]A sparse collection of pixels (or voxels) can be encoded using a lossy function with an even smaller number of parameters.

gles around the head. Due to the symmetry of a human head, only 3 models are used in practice (two side views can be mirrored; frontal remains as is). The most appropriate model can then be chosen in a real-time dynamic sequence. The choice of the model to be used depends on the estimated rotation of the head and by estimating that rotation successfully, models do not break down[36] when introduced with a high degree of freedom (e.g. angular freedom). This idea can undoubtedly be exploited in applications other than faces, but it appears to have a limited demand in industry and it has not been pursued much lately. It is the switching between different models in real-time and the selection of the *most suitable* model that makes this study challenging and for medical imaging, where the viewing degree of freedom is very limited, this extension is merely irrelevant.

The effect that different facial expressions and aging factors have on statistical model was another intriguing aspect that was mainly pursued by Lanitis *et al.* [31, 32]. Lanitis has recently worked on synthesis of faces and the analysis of facial attributes [30].

For a greater level of detail, as well as information on further extensions and applications, see Appendix A on page 161.

2.8 Open Questions

Active appearance models are a powerful method of interpreting and synthesising[37] images. Nevertheless, they are heavy, complex and they require a long time to train. Active appearance models sometimes serve a purpose which is different from that of active shape models and often they require more time to reach good convergence, mainly due to their additional complexity. In that sense, some implementation issues in appearance models need to be addressed; this can hopefully make them very

[36]When models break down, fitting defaults to a local (and hence false) minimum.

[37]This can be considered as being a reversal of interpretation, in fact. This binds with the notable computer vision/graphics differentiation.

powerful in more aspects. Furthermore, the accuracy of appearance models is sometimes lower[38] than that which is offered by other methods. If synthesis of photo-realistic images is a pre-requisite of the model to be used, then AAM's are indeed a unique and sensational technology that does the job adequately.

An additional valid critique of AAM's speaks of its occasional failure to reach the global minimum when posed with the goal of fitting. It is still not immune to large initial displacements (and hence discrepancies) or target instances that deviate abnormally from the training set. Since AAM's still rely on a good initial placement in a given target, there are possibly pressing issues to be looked at.

It is yet hard to ignore the fact that results of an AAM fitting are sometimes less accurate than those of an ASM[39]. This brings up the doubts as for whether the extra complexity associated with texture is worthy of being considered. The investment of time and intensive effort, including the need for human intervention, raises some important doubts and scepticism.

A significant drawback that is associated with appearance models is that for automation of model construction, landmark selection [6, 24], or more fundamentally image correspondence [62], is necessary and yet somewhat difficult to achieve. It is not obvious how to choose landmarks sensibly and how to judge the optimality of an automatic choice of significant points. Since the efficiency of an appearance model depends greatly on the textures embedded in that model, it is not sufficient to use existing techniques to select landmarks and pseudo-landmarks (additional points between the original anatomical or mathematical landmarks), as quite recently suggested by Davies *et al.* [10]. A further explanation of this work is spread throughout some of the following chapters, but primarily Chapter 5 needs to be fully grasped for understanding of this undertaken project and its manifestation.

[38] Although some results support this claim, it is quite likely that better implementations and further improvements will prove otherwise.

[39] This is *not* necessarily so in the like-for-like comparison.

Chapter 3

NON-RIGID REGISTRATION

"An open mind is freer than a bird in flight."
– *Joseph Kung.*

3.1 Introduction

THE principles of registration, especially with respect to the approach that this project takes, will be dealt with in turn. The subject is very broad and for deeper understanding of alternatives, cited literature needs to be carefully read.

Approaches on which work around the world is based are distinct, but they all have commonalities and so-called components on which registration seems to be logically based. The chapter aims to identify these components and give some rough overview with an emphasis on methods which the author and his colleagues consider to be a way forward.

3.2 Registration

Image registration has become essential in several domains where reliable acquisition of aligned images cannot be assured [21] or turns out to be complex. Needful to mention, the significance of this problem is made most apparent when alignment of a *group* of images must be guaranteed and the images are rather different in nature although they describe the exact same thing.

Misalignment can be due to movement of the subject or objects of interest, change of view-point, change in general conditions at the scene and even morphic changes or ones due to change in mass and elasticity of organs [22]. Changes in form can be observed over time in the scene or within its constituent parts, e.g. the inevitable and involuntary change in the form of the lungs. In some circumstances, as later discussed, misalignment is due to profound changes in the form of objects (usually *subject*s and anatomy) being scanned. Alignment is a key step that must be completed before analysis of collections of images is safely embarked on. It allows better understanding of the contents of all images as a group.

3.2.1 Transformations

3.2.1.1 Overview

Image registration ordinarily involves the manipulation of image pixels according to some rules and under the imposition of several strict constraints. It is commonly desirable to obtain a maximum similarity estimation [47] or simply an overlap measure amongst a group of images with a minimal extent of distortion. Even a small level of distortion may induce wrong assumptions or violate some stern conditions which should otherwise be an unbreakable pre-requisite. It is possible to think of the transformations used as if they pertain to different levels of "interference" – the interference to the analysis and interference with the integrity of the data. A typical categorisation of transformation types is as follows (ordered by increasing interference)[1]:

Rigid Allows translation (relocation in space), rotation and scaling (uniformal size changes, i.e. shrinkage and enlargement)[2]. The normalised shape attributes are altogether preserved and the process is usually concerned merely with some common alignment. Such alignment usually aims to place all instances upright and centred in the space origin with a fixed size of maximum 1 unit. The instance is virtually confined to lie inside a bounding structure (a circle or sphere, in 2-D and 3-D respectively)[3]. In 3-D, for instance, there is a total of 6 degrees of freedom so a rigid transformation will be wholly characterised by a tuple of 6 values[4]. This *does* in fact fully describe a rigid transform.

Affine Allows the instance (image for example) to *stretch and skew* along at least one axis or dimension, but not necessarily all (so that homogeneous scaling can be broken). Despite the fact that previously essential constraints are broken, all lines that were parallel remain parallel after the transformation is applied[5]. Reconstruction is said to be possible so that this transformation is invertible. For all affine transformations $T_a(x)$ where x can be a vector representation of an image (or volume) and their inverse $T_a^{-1}(x)$, the expression $T(T_a^{-1}(x) = Id(x)$ must hold. This relation must always be calculable and retain simplicity which makes it easy to resolve. This will

[1]The names sometimes change in the literature despite standardisation. What is important is the description of transformations and not the names or mnemonics that wound up describing them.

[2]More strictly, the inclusion of scaling makes this a similarity transformation rather than rigid.

[3]Such a process is a very fundamental one in computer graphics modelling and various books cover shape-normalisation techniques and algorithms.

[4]1 value for scaling, 3 for *x, y* and *z* coordinates and 2 for rotation, e.g. the *xy* and *yz* angles θ_1 and θ_2.

[5]Popular transformations such as skew, shear and taper, on the contrary, are not parallelism-preserving. The importance of this rigorous constraint is that the distance between any two points remains proportional to the transformation.

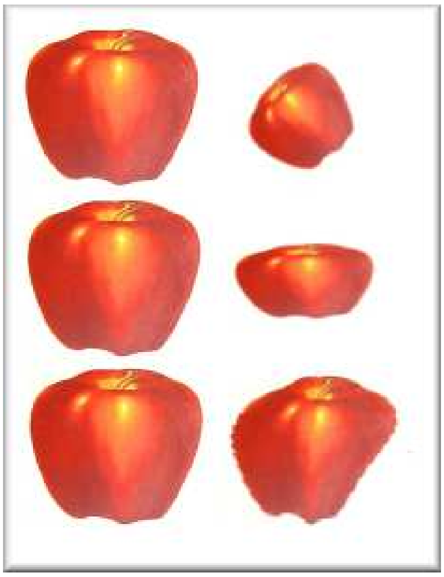

Figure 3.1: Registration examples; from top to bottom: rigid, affine and non-rigid transformations.

prove to be an important constraint when the practicability of warps is debated.

Non-rigid All other valid transformations fall into this category [15]. This includes tapering, spiral warps, pinching, etc. In principle, no inviolable constraints are in place, but quite clearly a non-rigid transformation attempts to preserve some of the primary structure[6] of the image while avoiding tearing and folding [49, 59]. This means that each pixel in the range must map to another and no pixel is left undefined. A bit more on this is to be explained later.

The images of an apple in Figure 3.1 illustrates the effect that each transformation type has on the image on the left.

As the figure suggests, the appearance of an object remains identical under rigid transformations. It is allowed strictly to grow, shrink, move, and rotate. Affine transformation allows an object to lose its original

[6]A random uncontrollable transformation will dispart basic structures in the image and can make valid interpretation impossible.

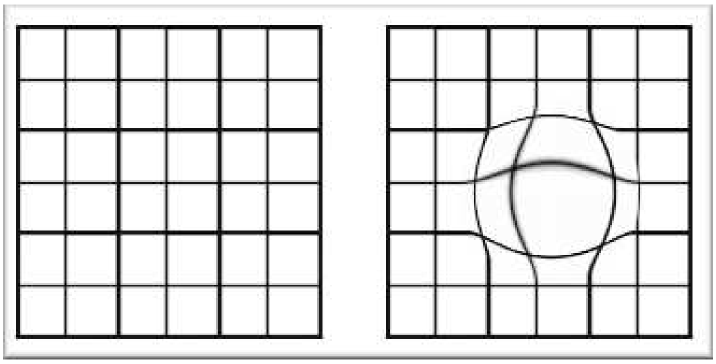

Figure 3.2: CPS non-rigid warp example. Warp is shown on the right-hand side.

form and non-rigid registration is far more permissive so the object can be subjected to rather obscure deformations.

What follows in this chapter briefly explains some of the main concepts, techniques and ideas currently employed. The actual key points, which describe non-rigid registration in the context of current investigative work on the issue[7], are as follows:

1. Warps
2. Similarity
3. Objective function

These three points will be explained in more detail with reference to current work, practical considerations and attempts already made. For now, a concise introduction would do. The approach often taken is that an image[8] needs to be warped (equivalent to transformation) until it matches another. The match is estimated with the assistance of similarity measures and this process of warping and similarity is sometimes wrapped

[7] This includes the Structure and Function Grand Challenge. The Grand Challenge aims to unify the different stages of analysis. It will be referred to yet again in Chapter 6 which deals with recent and on-going work, including that on non-rigid registration.

[8] More generally, arbitrary data of any complexity should be applicable.

up and put under one generic objective function. In that sense, the objective function bridges warps and similarity. Objective functions are then handled by an optimiser – a term which is further explained in 5.5 on page 71.

3.2.1.2 Diffeomorphism

The concepts and arguments introduced so far in this chapter show why there is an ever-increasing interest in non-rigid registration, based on non-rigid transformations[9]. The mathematics behind the required transformations and the theory that needs to be established in order to make them practical is constantly being explored and papers on the subject receive attention and recognition. Diffeomorphic [57] functions are *invertible, continuous* and *one-to-one* mappings for a given image[10]. These mappings are usually described by some local geometrical transformations that have an effect on pixels or the plane that pixels are embedded in.

Current diffeomorphic transformations that are used in Manchester University by Twining and Marsland [58] also benefit from having continuous derivatives at the boundaries, unlike for example, these of Lötjönen and Mäkelä [35] who suggested a similar transformation type. This, however, is a convenient property that is not a necessity. It is just a strategically good attribute to have in real-world applications.

What invertibility, continuity and one-to-one mappings mean in simpler terms is that for each transformation:

1. That transformation has an inverse so that any transformation (or *warp* as it will be later referred to as) can be reversed.

[9]This so-called mapping or transformation can be thought of as being a standard function, for example $f(x, y) = (x', y')$ in 2-D and it is applied to all the pixels within a predefined range.

[10]More generally, the functions are mappings defined over a matrix or a vector which is analogous to an image.

2. That transformation affects *all* data (pixels) within its boundaries so it has a spatially contained effect[11]. This means that every point must move as would be expected to give a continuous flow of intensities.

3. No two points should be mapped onto the same point as this would 'strip off' areas of the image.

3.2.1.3 Reparameterisation

Taking again an example from work on shapes[12], a shape can be described by a collection of landmarks as shown in Figure 2.2 earlier in this report. The landmarks are usually located at corners, T-junctions and edges that are easy to locate. Also, other additional points in between these landmarks can chosen to expand the representation of that shape and make it richer, though ideally curves should be continuous and the number of points that make them up arbitrary. To register multiple images, all corresponding landmarks and points must overlap in as accurate a way as possible. They must correspond to one another in one common spatial reference[13] so that image analysis can proceed. One way of doing this is to apply diffeomorphic warps to the space in which images will be embedded. That newly-defined plane is supposed to bring the collection of landmarks across the set of input images closer together. This ends up bringing a number of images to correspondence of some quality.

As shapes continue to be discussed, it is worth stressing that it is never obvious what choice of landmarks and intermediate points will result in an optimal overlap or even a good one. The quality varies depending on the pre-defined objective function. To automatically shift points and

[11]A pixel of course can be mapped onto the exact same original position, but the idea is that a continuous flow must prevail.

[12]The principles are better described by borrowing some concepts from work on landmark selection in shapes (to be further seen in Section 5). Similar methods as applied to images have not thoroughly been investigated yet.

[13]This can be thought of as a space which defines a common (non-linear) grid. In this grid, mappings between corresponding points become clearer, visually.

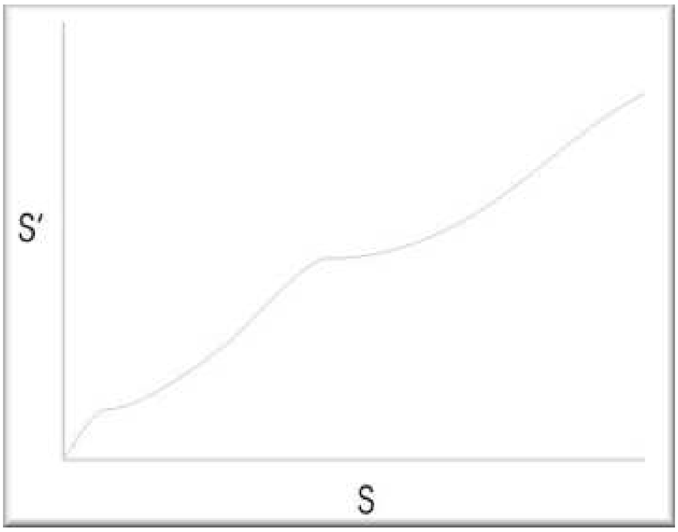

Figure 3.3: Monotonically-increasing function.

evaluate the subsequent global (or pair-wise) effects, *reparameterisation* of these points must take place in a way that preserves their order along the contour they form. A new spread of the points needs to be chosen iteratively and the results recorded. The spread of the points can be defined purely by a function and the reparameterisation alters this function to find preferable results. A monotonically-increasing function describes the distance of all points[14] from an arbitrary point on the curve in such a way that will not violate their sequential order.

Figure 3.3 shows what is meant by a monotonically-increasing function. The following must hold although its inverse may hold instead (a monotonically-decreasing function):

$$\forall(u \in S \wedge v \in S \wedge u < v) \rightarrow f_{mon}(u) < f_{mon}(v) \tag{3.1}$$

where f_{mon} is the monotonically-increasing function used and $f_{mon}(S) = S'$. More simply, the derivative at any point must be positive, i.e. $0 < \theta <$

[14]A continuous function is independent of the number of points. Therefore, the complexity can be increased progressively to obtain finer, more accurate results.

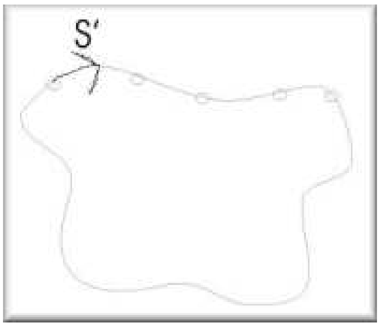

Figure 3.4: Reparameterisation example. A point moves along the curve a distance S' from the origin. All other points will do so as well to make this a continuous reparameterisation, often defined by a Cauchy.

90 so that $0 < tan(\theta) < 1$. In Figure 3.3, the meaning of reparameterisation as it is applied to points of a shape is made clearer. The distance or offset along the curve is guided by the value which was determined by the function above. In this particular way, all points which lie on the curve can be moved *simultaneously* without colliding with one another and new autonomous descriptors of shape become available. Instead of describing the movement of each individual point, an arbitrary number of points can be shifted according to one modifiable function. Davies *et al.* used this technique to optimise a shape model by evaluating the selection of landmark points. For each such reparameterisation, the specificity, generalisability and compactness were evaluated at some stage although minimum description length was ultimately chosen (to be discussed in Chapter 5 on page 67). The first and second of these terms were coined in the thesis published by Davies.

Prior to the invention of this technique, points were often chosen to be allocated a position on the curve so that they are equally-spaced. This approach was often a straight-forward and computationally inexpensive, but its results were unsatisfactory for more complex shape where the curve bends sharply. Some attempts were made at placing more points at regions of high-curvature, but these were still inferior to the aforementioned reparameterisation-based approach.

3.2.1.4 Warps

This short subsection adheres to a more local perspective – a perspective along the lines of which future research should move. The short part on diffeomorphism (Subsubsection 3.2.1.2) introduced functions that map a group of pixels to new positions. These functions will now be permanently referred to as *warps* plainly because this is the terminology that is invariantly used in the literature. Due to practical considerations, the warps used are chosen to be rather elementary and therefore computationally inexpensive. Some will argue that more sophisticated warps will produce better results in a smaller period of time because a smaller number of these is required to reach overlap as explained previously. However, they may also damage some structures in regions that are better left untouched, as well as interfere with previous warps that supposedly did the right thing.

The warps currently used are round (also extensible to spherical) and they can be described by their location, radius and possibly depth. As expected, these warps are parameterised by their horizontal and vertical location, magnitude and radius. Many such warps are applied at different scales to the image. Their position is quite random and good results are committed and carried on to later iterations while bad ones get discarded. Towards the later stages of the algorithm, only small local warps, much as in the case of reparameterisation, will entail constructive results.

The choice of warps is quite arbitrary and their most preferable complexity level is still an issue of active discussion[15]. The randomisation of the location at which warps are centred means that computation is saved on understanding the image and using any priorly-gained understanding. On the other hand, many warps are discarded in this way and wasted effort makes this method far less elegant.

[15]In Manchester University, Cootes and others are in favour of many small warps, but some are in favour of few rather more complex warps that are controlled by a larger number of parameters. More details on such issues appear in later discussions on current work on page 78.

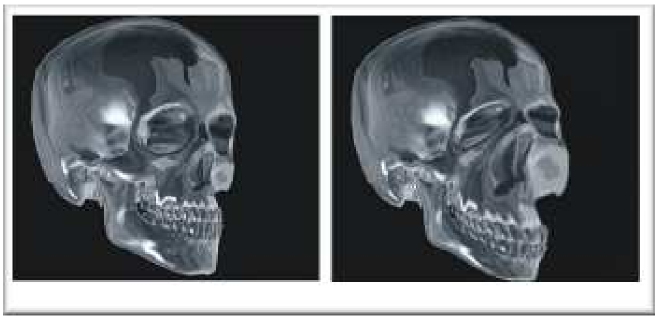

Figure 3.5: Warp applied to image. On the left: image before warp is applied; On the right: image after warping.

3.2.2 Measuring Similarity

There are various ways of measuring the conjectured similarity between two images. Mean-squared-differences or sum-of-squared-differences are rather poor methods of getting a useful measure of similarity if the positions where the images lie in space[16] are far away. Therefore, such measures are often better off used when convergence is foreseen. There are also measure that are immune to large spatial displacements or variability in form. Histograms of the intensity values in the images, where intensity values are accounted for globally (or locally, inside regions that require greater emphasis), would be far better measures under most circumstances. Extra strategic steps, such as the removal of empty bins in such histograms, are taken to make the histograms more powerful indicators of similarity. Active research repeatedly reveals better algorithms as will be described in brevity below.

Although the correlation ratio is still occasionally used to measure similarity, it is less relevant to this report and goes back over half a century ago [28]. Mutual information and normalised mutual information, as described by Studholme, give good measures that see high usage in existing non-rigid registration algorithms. Each one will be dealt with in turn.

[16]One can think of images as a vector of pixel values that define a position in a high-dimensional space

3.2.2.1 Mutual Information (MI)

Viola [61] has developed a way[17] of finding and measuring the similarity between two images or more by repeatedly comparing pairs of images. This ability to compare images is crucial for registration of images as it robustly and accurately returns an estimate of the beneficialness of the warps applied. Chapter 4 deals with the introduction of information theory and some of the basic measures which can explain the following in more detail. However, it is the principle that is worth understanding at this stage, rather than tedious technicality [42, 34].

Mutual information [WWW-13] computes *volumes of overlap* in images. If two images are matched, the joint histogram is then expected to give an indication of where *sharp* grey-value peaks are located and the sharpness value of these peaks. Under the converse case which is mis-registration, the joint histogram is then expected to show peaks of low sharpness and new peaks can emerge. The algorithms and advanced information theoretic expressions that take advantage of this observation are at this stage left out entirely. At this point, it is only worth defining a joint information (or entropy) to be $H(A, B)$ and state that MI calculates $H(A)+H(B)-H(A, B)$. This means that joint information is subtracted from the sum of information present in the two individual images.

3.2.2.2 Normalised Mutual Information (NMI)

Studholme [53] and Maes [36] suggested that some normalisation should be applied to mutual information as was described above. Quite a few steps are involved in this normalisation process and the full mathematical summary is left to the literature[18].

[17]The discovery of this mutual information is actually attributed to Maes as well. The thesis worked on by Viola in the mid-nineties received great recognition though and MI is ascribed to him.

[18]There is an additional distinction between symmetric and asymmetric normalised mutual information, but rationalè for this requires the full technical recipe. The dissertation at `http://www.lans.ece.utexas.edu/~strehl/diss/node107.html` summarises the way in which NMI evaluated.

The main difference is that the expression used for MI is significantly extended and divided by a normalisation term. The method is predominant in non-rigid registration as it yields good results

3.2.2.3 Mean Sum of Differences (MSD)

This measure was explained and illustrated in the context of active appearance models where difference needs to guide model fitting. Its idea is primitive, nevertheless it is effective, especially when faced with the simplest class of tasks. Pixels are compared in two images one by one, their squared grey-level difference is calculated and a sum[19] over all differences is returned – this obtains a sum of squared differences (SSD) in fact; MSD is simply normalised by the number of pixels which is a rational step to perform. This method is usually powerful if the two images compared are closely aligned and their intensity values are relatively continuous and low in contrast. In other words, MSD will tolerate only a low level of locally-situated difference, while contrariwise, MI and NMI rely on sparse dispersion of all pixels.

It is worth to consider suggestions on the issue of speeding up similarity measures. Some of the above measures depend heavily, from an efficiency point-of-view, on the dimensions of an image. As described in the context of active appearance models, a multi-resolution approach can be used to speed up the whole process. Blurring or averaging followed by re-sampling or sub-sampling allows for images of smaller size to be manipulated and complexity to be quadratically lessened. As the similarity measures are proportional to the images size, far better performance can be achieved by a transition from coarse to finer resolution. Pluim [43] identifies the effects that this approach will have on the measurement of similarity.

[19] One could suggest an extension to such a method and assign weights to differentiate regions of varying significance.

3.3 Summary

The main concepts that image registration involves are transformation and similarity. Care must be taken, however, when choosing a proper methods for each. By identifying some valuable properties that ought to be preserved, one can put together a sensible registration scheme.

There is yet little agreement on the power of each of the existing methods and comparisons are often biassed. Pointers will be provided to additional methods and approaches in later sections. Also, several possible results and comparisons will be proposed.

Chapter 4

INFORMATION THEORY

"The word 'thank' came from the word 'think."

– *Rev. Goh.*

4.1 Importance

THE arguments regarding the importance of information theory with respect to this project vary. Information theory is indeed valuable due to its relevance to past projects in the field, on which future projects will rely. Image analysis is often involves the passage and handling of large sets of data and extraction of the meaning of the data is a necessity. Compression becomes ever more crucial when voluminous models and entities are maintained in memory and, again, reasoning about compression goes back to the theory of information. New ways to encode data, avoid redundancy and describe objects succinctly are being sought as they often reduce the *complexity* of any system as well as its *size*. Measures of information are necessary to introduce and support learning capabilities which in turn form intelligent systems. Such systems can evaluate and judge improvement as illustrated thus far and as will be illustrated later.

4.2 Entropy

The term entropy (commonly also referred to as Shannon's entropy [WWW-15]) is used to denote a general measure of *uncertainty*. It is not a very sophisticated idea, yet a very fundamental one which was first introduced in 1948. Uncertainty is associated with the required amount of data so it can also be thought of as an information *measure* or quantifier. The value that quantifies uncertainty originally related to random variables which take different probabilities amongst a set of states (reminiscent of Markov chain models). Shannon's entropy has become a very useful way of evaluating structures and pattern in some data. The lower the entropy value, the more data is already inherent in that data. In a sense, the entropy indicates how much can be learned from the data and what is still unknown.

4.3 MDL

Minimum description length [46] provides a measure of the *minimal* amount of information necessary to encode some data. Any data can be transformed in a particular way so that it becomes a sequence of symbols (numbers or signals even, to be less general)[1]. The transition between one symbol to another can be encoded by some transition table which holds the probabilities of all possible transitions. For n symbols, up to n^2 transition need be defined. Markov chains are one such model type which is a convenient way of explaining the nature of MDL. Markov chains of a high

[1]Binary representation is quite complete in the sense that any data, e.g. programs and text, can be coded in a binary form. However, this representation might be very greedy of space and the issue of representation compactness then arises.

order can accommodate for data of more awkward and unpredictable variance. MDL infrequently defaults to higher-order models (see examples at [WWW-14]) which are superiorly expressive, although they require a far greater number of transitions to be specified.

With proper use of models as in the case above, data can and should be represented solely by all transitions and can then essentially replicated from these transitions. Unless the data is peculiar and shows no patterns, such a description would be compact for data large enough in bulk. MDL attempts to describe the extent to which some data is capable of diminishing in bulk (with or without loss being a separate issue) or rather the minimal amount of information that needs to be available to describe and reconstruct that data. In most cases, the more uncertainty present (i.e. higher entropy), the greater the minimal description length would be.

As an example, here is a vector representation of some arbitrary data: $v = \{3, 4, 3, 1, 1, 3, 4, 2, 2\}$. There is an alternative way of representing this data. By using a first-order transition table, e.g. $3 \rightarrow 4$:1, $3 \rightarrow 2 : 0$..., the likelihood or probability of transition from every element to its successor is revealed. Observable patterns are merely meaningless in this data example. Encoding of transitions will also be an inefficient approach as a result of the small vector size and the low sequential correlation[2]. On the sharp contrary, vectors such as $v = \{0\}$ or $v = \{1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1\}$ bear a very small measure of uncertainty. In the second of these[3], only one transition exists so it can be represented by a tiny model and the entropy is 0.

To summarise, MDL is a measure of the minimal amount of information that expresses a sequence[4]. By inspecting transitions it is possible to get an insight into the complexity of some model. A heart beat pattern,

[2]To make this appear more practical, one can think of a large (> 100000 pixels) image where patterns are present.

[3]This can be portrayed as a uniform plain-white image.

[4]General problem reducibility to a sequence is axiomatic as Turing Machines suggest.

for instance, is rather predictable and repetitive in comparison with the positions of a person's fingers over time. This means that the description length of the heart state should be shorter than that of the hand. Less information is required to capture the behaviour of the heart in motion (heart beats are not sporadic).

Issues of transitions and understanding of data patterns will later be explained in a different context. Rather than reconstructing vectors, images need to be reconstructed by the least number of bits. These bits are permitted to permute quantised natural numbers as the above arguments suggest.

Chapter 5

MDL MODELS

"History will be kind to me for I intend to write it."
– Sir Winston Churchill.

5.1 Background

AFTER understanding models based on statistics and with some additional background on registration, reasons for the work that follows become apparent. This section continues the argument in Section 2.3 – the argument which speaks of the possibly arbitrary choice of landmark points.

5.2 Landmark Selection

Past work by Kotcheff and Taylor [29] attempted to address the problem which was highlighted earlier on. This problem is that of selecting good corresponding points on a curve of shape. It did so by evaluating the choice of landmarks[1] via the model which resulted from that choice. The determinant of the covariance matrix of the model was said to be a good approximation for model quality. Poor models, it is yet again worth emphasising, imply that when PCA is applied to the data, the correctness of corresponding points and their distribution, will be invalid. Some of the technicalities will be described later on, along with mathematical notation.

5.3 MDL in Modelling

The concepts outlined in Chapter 4 have been applied to select preferable descriptors of shape [11]. Selection of points that describe a given shape, as explained in Chapter 2, was perpetually altered and evaluated to find shape models and examples that require a smaller set of data to be passed as an encoded message[2].

To express the process at a moderate pace, each time points on the curve that traces the shape are selected, a different model is ultimately constructed. A good and compact statistical model is one whose legal variations are relatively small and possibly so are the number of its control points. Such a model is found using a general optimisation regime under which points are reparameterised. MDL can be used as a replacement for similarity in an objective function that is iteratively evaluated for each such points reparameterisation. The minimisation process will described in reasonable detail in Section 5.5 on optimisation. The more genuine part of this seminal work is the use of an existing information theoretic measure, namely MDL, to guide an autonomous search for good models. This work will be explained with respect to current research in later chapters and especially in Subsection 6.1.1 on page 76 which takes a more focused scope. One alternative way of realising what this method is based on is to look at its objective function.

[1]Often the choice is random so that no assumption are made about the problem.
[2]An alternative method involving B-fitting was proposed by Thacker *et al.* [54].

5.4 Objective Function

5.4.1 Principles

The objective function is the actual function which needs to be minimised (fundamentally by finding a set of values for all parameters) in order for an optimal choice or a solution to be picked from the many alternatives offered. The function is most heavily based on similarity measures as was briefly explained earlier, but it allows this measure to be extended in some way. For example, it can be helpful to include the cost of the warps that are used. The reason why the cost of the warp is sometimes an integral part of the function is that long-winded warps are not nearly as desirable as uncomplicated ones that perform the task equally well or even better. This cost is often considered a *regularisation term* [9] which penalises a sequence of warps that form large trajectories in space. It will seek a solution that is simple rather than finding an odd trajectory in space that gives a similar solution. This is the case since the solutions are often not unique.

Objective functions are built to encapsulate in a concise and effective way everything that is repeatedly evaluated. They are therefore required to be a very efficient 'unit' (or black box) which will be invoked quite frequently. The *speed* of the registration will directly depend on the choice of an objective function that adds up results from warps, similarity calculations and possibly more components, as can be seen in current group-wise registration papers. The *quality* of the registration will of course depend on this function, too.

Let the two images $\mathbf{I}_m$ and $\mathbf{I}'_m$ be defined as the images before and after warping respectively. Let a warping function $f_w(x)$ also be defined to be $f_w(\mathbf{I}_m, < parameters >) = \mathbf{I}'_m$. For a similarity[3] function f_{sim}, the objective function can then take the form:

[3] It will temporarily be assumed that for an objective function that needs to be minimised, the similarity measure will return small values for good similarity and vice versa.

$$f_{objective} = f_{sim}(f_w(\mathbf{I}_m, < params >), \mathbf{I}_r) + < reg - terms > . \qquad (5.1)$$

The function then attempts to find a series of parameter values that will lead it to a globally minimal solution. More precisely, it attempts to find *assignments* for all parameters that describe the warps so that similarity is maximised (or difference minimised)[4].

The explanation on the objective function concludes the algorithmic approach that registration takes. Non-rigid registration algorithms can be assessed by methods such as the one described by Warfield [65].

5.4.2 The MDL-based Objective Function

As was explained in the previous subsection, objective functions define the means by which a solution is to be found. Efficiency is a reasonable concern so a sophisticated function that is prudent to construct the model more frequently than necessary must be employed. The function used in this context needs to drive the search for shape correspondences using a suitable parameterisation (in the case of image registration – transformations which increase similarity across *all* images). The different nature of the problem and the methods of solving it convey the ulterior goal somewhat differently than the vast majority of methods to date, resulting in the formulation below.

For the similar case of image registration, one can denote a transformation function $W(\bullet, params)$ and the construction of an appearance model to be $Model(\mathbf{x}_1, \mathbf{x}_2, .., \mathbf{x}_n)$ where $\mathbf{x}_i$ are the images used to train that model. One seeks a model that is more compact using the following (simplified) function

[4]Being slightly more specific, this function is said to minimise the sum of the difference between two images and another less significant term. The two images compared are the transformed image $\mathbf{I}'_m$ and the reference $\mathbf{I}_r$ in this case.

$F_{obj} = MDL(Model(\mathbf{x}_1..., \mathbf{x}_i.., \mathbf{x}_n)) - MDL(Model(\mathbf{x}_1..., W(\mathbf{x}_i, params).., \mathbf{x}_n))$

where $params$ should be found to minimise this expression for each image vector $\mathbf{x}_i$. A succinct description of this algorithm is as follows:

- Repeat
  - ✧ For each image vector $\mathbf{x}_i$,
    - ❑ Optimise F_{obj} by altering the values of $params$.
- Until convergence.

In practice, to indirectly and quickly evaluate MDL what will be obtained is $\sum_{i=1}^{n} log(\lambda_i)$ where $\lambda_{1<i<n}$ are the n Eigen-values of the covariance matrix whose magnitudes are the greatest. This is similar to the formulation of Kotcheff [29] where $\sum_{i=1}^{n} log(\lambda_i + \delta)$ is calculated to approximate

$$det(\mathbf{M} + \delta) \equiv \prod_{i=1}^{n} \lambda_i \propto \sum_{i=1}^{n} log(\lambda_i + \delta) \equiv log(det(\mathbf{M})) \quad (5.2)$$

where M is the covariance matrix under consideration.

5.5 Optimisation

5.5.1 Background

General optimisation is often used in the process of matching and its complexity can be relatively high[5]. This process is by convention concerned

[5]The behaviour of such a problem is not linear and it may cross over to the realms of quadratic programming (QP) where various parameters simultaneously control a func-

with the minimisation (the complement is used to generalise it to maximisation) of the value of a function and that function often comprises more than a single variable which makes it multi-dimensional. Many software products that act as general optimisers exist and the way they operate and perform varies. Some even switch between different algorithms depending on the stage of the optimisation and the changing granularity of the problem.

Optimisation over a function which varies in many dimensions is an expensive process. Often this optimisation requires some *a priori* knowledge of the problem domain so that performance winds up being satisfactory. In the case of image matching, advantages can be gained if the effect of variable alteration can be predicted in some way. An example of this was described in Section 2.6 on page 39 where pixel intensities have a dependency upon a group of parameters. Slightly less specifically, given the difference between two or more images, or even some generic data regarding a *change* caused by value changes in the function considered for optimisation, it should then be possible to determine paths that lead to quick convergence.

For the problems outlined and sometimes just alluded to so far in this document, common optimisation methods are gradient-descent and downhill simplex. However, many other methods exist[6] and whole books have been written on the subject [44]. The advocated strategy would sometimes be a utilisation of mixtures of different methods with rational choice of the most relevant one at each stage. That is plainly because the different characteristics of the methods make them advantageous at different states throughout the entire optimisation process.

tion and minimisation is therefore by no means trivial.

[6]To name several more methods: dynamic programming, genetic algorithms, Powell's, simulated annealing and steepest descent.

5.5.2 Problems

One of the main flaws of existing optimisation methods is their inability to find a global minimum (or minima) fairly quickly without some additional knowledge about the function under investigation. Rough assumptions about the behaviour of the curve along each of the axes[7] are otherwise made.

The pace of the optimisation process can be boosted at the expense of overall accuracy and error likelihood. Sometimes these cannot be jeopardised, health-care programs being an example of choice. It turns out that if no exhaustive search[8] is carried out, there is then a danger of convergence at some local minimum. In most applications, any stoppage at a local minimum would be highly undesirable although this may be better than a complete failure at identifying regionally lowest points. Local minima are a necessary evil for large and complex continuous functions.

In conclusion, there is a trade-off between speed and accuracy although accuracy can be achieved at a lower cost if more knowledge is acquired off-line, before the optimisation task actually begins. Quite expectedly, this also implies that many redundant computations will consume precious resources and time in order to train the optimiser.

5.6 Summary

This section has demonstrated some of the advantages gained by using an MDL approach for choosing landmark points in a set of shapes. The notion of an objective function function was explained, as well as an information-theoretic one. Once an objective function is in place, there are various issues that are concerned with the optimisation regime. The

[7]Optimisation is a multi-dimensional problem that searches along hyper-spaces, some of which are orthogonal to the many existing axes.

[8]Exhaustivity is impossible for continuous functions, but digital images are luckily discrete.

way by which good solution are sought is rather crucial and the later section on registration experiments (7.4) elaborates on current work relating to this subject.

Chapter 6

Project Plan

"Intellect is invisible to the man who has none."
– *Arthur Schopenhauer.*

6.1 Starting Point

It is worth starting off with a description of the some recent and relevant developments which were made before September 2003. The following few paragraphs summarise and shed light at some of the main principles. These principles consequently describe some methods which are still used in the existing system – systems that surely needed to be extended and their understanding was the most crucial.

Smith's work follows the work of Davies in a more-or-less obvious sense, but it explores a different domain with slightly different aims. Each of these two research efforts will be dealt with in turn.

6.1.1 Returning to Shape Models

Davies repeatedly performed a reparameterisation over a given series of shapes, or rather their defining points (although in principle he dealt with continuous curves where points are just implicitly defined). All these points were shifted in accordance with some displacements, as orchestrated by a monotonically increasing curve. This reparameterisation was applied to all examples, one reparameterisation for each example[1] in the training set to evaluate an optimal choice of point spreads, or more precisely, the favourable reparameterisations that act upon these points (in principle, that is defined with respect to a curve which will be represented by a finite number of sample points.

In current group-wise registration work, the elements that such reparameterisation affects are the points which control the *warps* applied to the data[2]. These chosen warps are then applied to all the examples (or data instances) and measures are used to describe the vague notion of modelability. One could argue that the model provides a good indicator of how "similar" the data is collectively*, en mass*e. Another way of explaining this process is to say that warps are being found that reveal data correspondences. Correspondences are found when points (or imaginary sample points of the curve) lie in analogous regions – that is – regions that are describing the same part of the logically equivalent class of objects. A warp implicitly defines an uneven plane for images to be embedded in and when all images get embedded in that plane, they should then be collectively similar. Interestingly, that similarity can be checked with the use of AAM's (reminiscent work can be found in [49, 33, 63]). Ways of evaluating an appearance model and ways of drawing conclusions about the data that was used to build it already exist. The algorithms developed for this work use a similarity measure such as MSD or MI to see

[1]There was also a further investigation into the optimisation scheme. All shapes can be optimised over simultaneously (also known as joint optimisation) or one can be optimised at any single iteration (known as sequential optimisation).

[2]The data type is irrelevant. It makes no difference whether it is an image of full appearance or just a 'brick-and-bump' as was repeatedly the case.

how similar images become during search[3], before a model is created. The model created from *all* the examples is the entity that defines the 'goodness' of the warps. A model can in some sense describe and measure of similarity across the entire set, as opposed to the pair-wise measures used previously. This construction of a model can in that unprecedented way guide the search for good warps. The system seeks control points that define good warps and it seeks such points using the idea of reparameterisation. The resulting warps must then produce good models for the *whole* data. For example, in the case of these specific experiments, all the data instances are warped to become quite similar so the model created from them has a low determinant.

6.1.2 Registration Based on Models

6.1.2.1 Summary

Non-rigid registration (NRR) and model-based image analysis were previously believed to possess some commonality – a premise on which current research is still based. There is a growing belief that the best of both can be exploited to construct a unified framework. This modified framework might be more robust and offer higher utility and functionality when compared with the other two approaches working solely.

It is claimed that warping of the images, as is already done in medical imaging in particular, can be used to find correspondences that are optimal in some respects. Pair-wise (and sometimes group-wise) image registration using non-rigid transformations was the way in which previous (local) research by Smith had planned to build good models of appearance. Furthermore, non-rigid registration made it possible to achieve better correspondence in images and maybe supercede other methods. By constructing models from the transformation parameters, one could also highlight and describe successful registration trajectories, i.e. ideal

[3]This similarity computation is incorporated in the objective function and it usually comprises a collection of pair-wise similarity measures.

warping sequences or a legal range of warps. As a result of the process in its entirety, active appearance models could be constructed automatically (for identification of correspondence no longer requires any human intervention) and non-rigid registration could guided by appearance models rather than similarity measures which are pair-wise.

6.1.2.2 Results

There was never evidence to indicate that the results of previous work were as successful as had been hoped. It was anticipated that since it dealt with group-wise registration (and globally optimal models), its results should clearly be better than those obtained from pair-wise registration under similar conditions. This approach came with an extra cost to make things even tougher. The process which offered little improvement was far slower and the results were not superior by any noticeable measure. In fact, it is possible that the results were partially biased – not positively, but rather negatively, due to few of the conditions set for the experiments.

Back at the start, circa October 2004, impending experiments aimed to disprove the claim that the conditions and choices made were the cause for any apparent advantage. They just as well aspired to increase apparent gains of group-wise registration – something which was not yet evident at the time. Answers as to why this was so will be given in Chapter 7.

6.2 On-going Work

6.2.1 A Critique

Work in the field seems to move in differentiable yet almost identical directions. On the one hand, speed is an issue that might not have a definite solution and parts of this document elaborate on imaginable hindrances and conceivable impediments of this kind. Algorithmic trade-offs and different choices of programming language or paradigm, operating system and platform is a matter worth pursuing. Since the process relies on wide global scope, i.e. investigation of various images simultaneously where change in one affects all, heuristics can perhaps be applied to decrease the number of iterations involved[4].

Claims of a similar nature can be made on the more intrinsic part of work being reviewed. For a start, values were often tweaked manually and no strong evidence was used to support such arbitrary selections which purportedly followed common sense. Another problem that has been realised is that much of the process comprised the simplistic joining of remotely-germane components whose nature is unique and autonomous. This means that components in the system often suffered from the undependable fusion which was in place – something which is a direct consequence of the knowledge that is still lacking in the field. There is much to be learned about how the numerous existing techniques, measurements and component should be merged effectively and, by all means, conveniently.

6.2.2 Parallel and Related Work

Rueckert *et al.* [48] describe statistical deformation models (SDM's) which are in essence surprisingly similar to appearance model. As it turns out, they are explicitly set to construct an appearance model using statistical analysis as described in Chapter 2.3. They do this analysis in a strategically different and indirect way though. To transform images, B-Splines are used which are quite powerful, well-understood and commonly used, e.g. in computer graphics rendering and curve fitting. However, they suffer from one main drawback which is deficiency of diffeomorphism. What

[4]Although knowledge of the problem is an integral part of most program optimisation steps, the more formal methods can be used to identify dependencies. A dependency graph can reliably indicate where re-evaluation is indeed necessary (e.g. Figure D.1 on page 183).

this practically means is that parts of the data can be torn or folded, i.e. structures can disappear. This cannot happen if the CPS-based warps are used, but a valid comparison is needed to discover if this attribute really is all about gains.

Similar concepts have been applied to segmentation in [16]. Much work has concentrated on using the knowledge and techniques from each one of these two to establish a more powerful framework of full appearance statistical models. The work is described in Chapter 6 on page 75 with reference to research that is associated with the GC (some of the cross-over papers are relevant in this context). An exclusive introduction to Rueckert's work will attempt to elucidate the current registration concepts which future research relies upon.

Non-rigid registration methods have been applied in several medical domains of expertise. Amongst these is the renowned brain analysis task, contrast-enhanced MR mammography and segmentation and tracking of the heart. The procedures currently employed are inclined to follow higher-order entropy measures that will not be delved any further. Rueckert's homepage [WWW-2] which is listed at the end gives the full details and references. Chapter 4 on information theory explained in brevity some of the basic ideas behind these so-called entropy measures.

The success of temporal non-rigid image registration method is dependent upon two factors:

1. **Search algorithm:** As earlier illustrated in the context of active appearance mode, good warps need to be searched to achieve good similarity.
2. **Similarity:** The performance relies on a suitable choice of similarity measures which guide the search until a sufficiently good fit is declared.

Learning the properties of similarity measures, the way they affects the search duration and the effect warps have on similarity are all important aspects of registration method development. This is reminiscent of the process in which correlation between parameter changes and intensity changes are learned in appearance models.

Statistical parametric mapping (SPM) are being used in University College London [WWW-16] in order to register bio-medical data. The term SPM refers to construction and assessment of spatially extended statistical process that can be used to test hypotheses about given medical data, especially in the domain of neurology. SPM spatially normalises images into a standard regular space and then applies some smoothing. Statistics which are then extracted from the registration of the data are addressed by theory of continuous random fields. None of this is arcane, though the concepts are rather unique to UCL.

Also in UCL, registration is performed which is based on fluid models. The rigid movement of objects does not usually impose problems as those introduced by soft tissue. Fluid registration is a matching technique which models these awkward morphological changes as compressible viscous fluid. The idea is presently applied in brain imaging where greater interest has existed for some time.

Change in organs due to resection (craniotomy being a banally-encountered scenario), expansion, movement etc. is often modelled using thin-plate splines [5] and the motion of organs can be handled using free-form deformation (FFD) which are based on B-splines. Prior to this embedment of high-order functions, the effects of rigid-body motion is annulled by Euclidean transformation. Similarity measures guide this process of rigid registration just as well. It is the technical description of the algorithms used that proves why these methods, which are used in Guy's Hospital, are extremely effective. As earlier mention, current work is done using the bi-harmonic [1] clamped-plate splines and possible investigation is considered for a model-based objective function that uses other morphometrical methods.

6.3 Goals

A main goal, which appertains to the big picture that is the GC, is the merger which involves (non-rigid) registration and statistical models. In both cases, some *dense correspondence* across some or all of the images is involved and must eventually be determined. Re-use of the information that is incorporated in each of this two techniques (which are believed to be inherently the same) would make the overall analysis task more powerful, flexible and well-integrated[5]. If even a moderate combination of the two is obtained, then new ways of building and using models will be open for investigation.

In NRR, lower-level inspection of image pixels identifies similarity using mutual information (or any other similarity measure for this argument's sake), whereas in statistical modelling, the correspondences are often marked by hand (as explained in previous chapters, this is no longer quite the case necessarily) or gathered in an ill-chosen fashion. It is imperative that effort is made to reuse the segmentation from NRR so that models can be constructed more quickly and fitted to targets before feature extraction takes over and does its part of the analysis job.

[5]The parallel development in both fields, especially the need to identify homologous structures, is what makes this GC suitably arranged and increases its potential of resulting in success.

Chapter 7

EXPERIMENTATION

"Civilisations can only be understood by those who are civilised."
– *Alfred North Whitehead.*

7.1 Overview

THROUGHOUT the past year, many results have been obtained and better understanding established. Some of these are more relevant to the preliminary research aims (see Form 2) than others. Although experiments were applied to 1-D data, principles that extend to a higher-dimensional data were learned and should form the theoretical grounds for future work concerning the model-based objective function. It has been agreed that shortly into the second year of this research, once sufficient understanding of the problem and confidence are gained, 2-D (and possibly later on, 3-D) application of the method will be investigated[1].

[1] Issues which yet cannot be ignored are related to efficiency. By scaling down the problem though, proof of method appropriateness is both possible and traceable.

Indication of prescribed tasks and some intermediate submissions of progress up-to-date can be found in the forms. They are available online[2] as Portable Document Format (PDF) and Word files at:

- `http://www.danielsorogon.com/Webmaster/Research/Progress/forms.htm`

All weekly progress reports can be found under:

- `http://www.danielsorogon.com/Webmaster/Research/Progress/`

There are clear advantages to the retaining of the records above. These will later allow chronological dissection of progress made, as well as the problems encountered on a day-to-day or week-to-week basis. They are yet not adhesive enough and this chapter attempts to provide a short summary of the large bulk of experiments by just listing some of the more important ones and explaining how they serve the raison d'être of the project. All experiments are documented to their finest detail at:

- `http://www2.cs.man.ac.uk/~schestr0/Experiments/`

They were generated using AART which is a newly-constructed tool that is shown in Figure 7.1 on page 86 and explained in Section 7.3 on page 86.

All results were obtained under MATLAB [WWW-1] which is the working environment on which AART operates[3].

[2]The author is aware that Web references and expansions through remote text are frowned upon. However, these are collectively an 'open door' to the large majority of work (some 2,000+ files); much of it is completely omitted from this report.

[3]On a separate note on knowledge- and code-sharing, GUI components and demos were made available at MATLAB Central. They received nearly 4,000 downloads, accredited to the author and his affiliation with ISBE and Manchester University. These contributions ranked him amongst the world's top 5 for popularity in July 2004. Confer [WWW-17] for more information.

7.2 Milestones

From the large pool of experiments which were performed, the next few sections, which include much of the 'meat' of this report, select and focus on a few which inferred important information and conclusions that ought to be highlighted.

The following experiments can be seen more or less as milestones. These are the catalysts for the more interesting and/or perplexing results which could be identified amongst the entire set of experiments. Each of these will be explicitly or implicitly mentioned later on as minor milestones are dealt with roughly chronologically in Section 7.4 (page 87), Section 7.5 (page 121) and Section 7.6 on page 124.

The milestone are:

1. Registration target and approach of the model-based evaluation to it.
2. Comparison and benchmark of different registration methods.
3. Finding the correlation between the size of the set and the performance of the model-based objective function.
4. Point insertion to compensate for the change in bump height.
5. Use of the residual of the model to better perform (4).
6. The optimisation refusing to improve steadily, fixed by dynamically changing the precision required from the optimiser.
7. Finding out that optimisation can go below target even when initialised at the correct solution described by a piece-wise linear warp.
8. Considerable speed-up of the algorithm.

7.3 Environment

From the point when the project commenced, a tool was needed to conduct experiments in a systematic manner – namely a flexible work environment for experimental studies. The problem which needed to solve was that of simple data registration so one dimensional data needed to be generated and then analysed in sensible ways, before, throughout and after registration.

The high-level procedural language MATLAB was used to develop a package that can be rapidly modified and tested. On top of this package, a graphical user interface was laid[4] and results were displayed in the form of hyper-text. As development of this package progressed, it was decided to name it Autonomous Appearance-based Registration Test-bed (AART). As the name implies, this is primarily a flexible environment in which registration tasks can be performed. Of particular interest it was registration which is based on appearance of images. This appearance can be described by the means of a model and the process of registration is intended to be free of user intervention, hence it is autonomous.

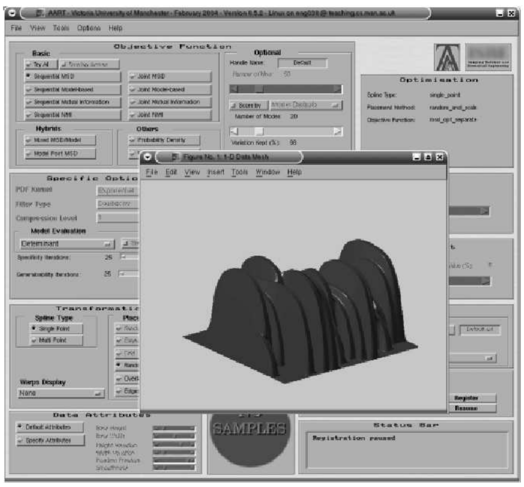

Figure 7.1: Autonomous Appearance-based Registration Test-bed in February 2004.

[4]The user interface is courtesy of Java and it runs over a Java virtual machine (this can be seen as either having pros or cons).

As it presently stands, AART is a stable tool that has a great number of run-time options. By setting these options, new experiments can be quickly conducted and results returned in visual form as well as in rich textual form.

7.4 Registration

This large section will provide explanation on work, experiments and some results pertaining to the main project goal. Most of these results are described in graphics and text and, at this stage, no registration video sequences are enclosed[5].

7.4.1 Initial Exploration

Experiments below will not be sorted purely chronologically and will not be listed according to streams of consciousness either. They will rather be explained in a logical way which builds coherently towards the inference of conclusions and the way in which experiments aided the understanding which was so necessary. A brief and incomplete list of milestones was included in the earlier part of this chapter and although this list is not expected to be complete in any sense, it should be able to cover much of the more important experiments in sufficient detail for them to be truly understood.

7.4.1.1 Generation of Data

It has been implicitly mentioned that presently only 1-D data is practically under consideration. In order to keep results consistent, much of the time was spent investigating a particular class of data – that which shall

[5]It seems likely that a CD-ROM will accompany later surveys (a la thesis). AART can generate several movie types with little user involvement.

be referred to as the bump[6]. The bump (a half-circle or ellipse) which was being generated varied in 3 separate ways:

1. Position
2. Height
3. Width

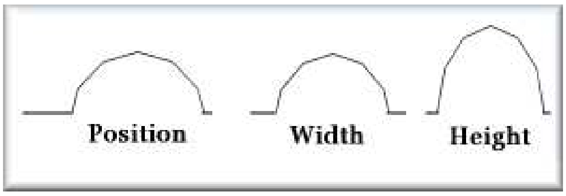

Figure 7.2: Illustration of the three variation modes.

Dealing with each of the above variation types in turn, position refers to the horizontal placement of the bump, height refers to the peak value (judged by its Y-component) and width refers to a relative width for the bump (see Figure 7.2). Later figures clarify what is meant by these properties visually. This synthetic data type was chosen due to few interesting and important attributes it possesses. It proves to be a difficult problem when treated as raw input for registration, but more importantly, it is in fact possible to know what one means by a *correct* answer to the problem. That correct solution is also feasible to identify[7]. As the notion of models is used here persistently, one could expect the modes of variation found to reflect on the three pre-defined modes being position, height and width.

Figure 7.3 shows what the vector representation of the data actually means.

[6]This is rather a different type of data than the one mentioned in past work (Appendix B where bumps are less composite) and that which has been tested in MDL shape optimisation (Chapter 5 where brick topped by a bump is looked at).

[7]In real-life circumstances, there will rarely be a correct solution for inter-subject registration. There may, however, be one for intra-subject registration, e.g. in the case of correction for movement.

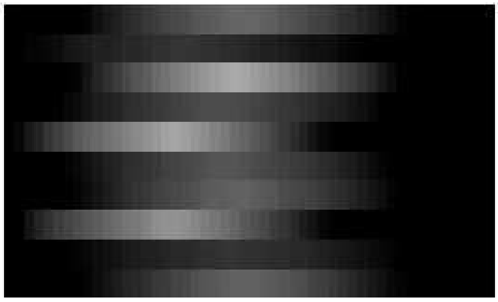

Figure 7.3: Data being registered. The registration process is visualised by an image composed of data vectors. The columns are 1-D vectors interpreted as grey-scale pixels.

Figure 7.4 shows the data in another more fascinating way which can dynamically illustrate the change due to registration. This representation has been used to form registration videos and it will be definitely returned to in future experimentation.

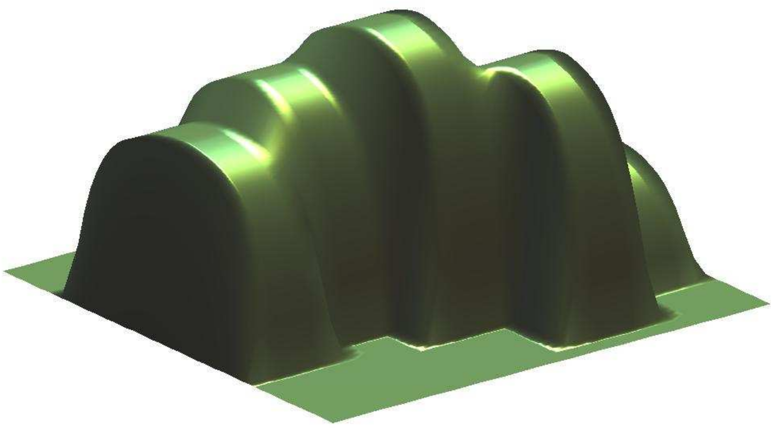

Figure 7.4: Original data set of size 5 before any application of warps.

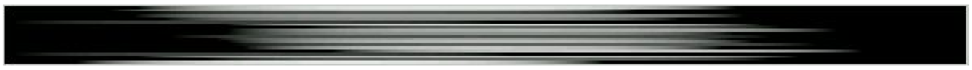

Figure 7.5: A larger example of pixel representation for 1-D bump data. This is somewhat of an enhancement to Figure 7.3 on the preceding page.

Later in this section, it will be shown what effect warps have on this data. Moreover, it is important to mention that the algorithm was applied to different synthetic data types although this was rare. Comparison is better performed over the same standardised dataset. Other data was usually used for reasoning about the correctness of algorithms and error detection via more trackable debugging tasks.

7.4.1.2 Analysis of Warps

Throughout the entire year, there was some general interest in how clamped-plate splines affect the data and how the model-based objective function affects the choice of warps. Several of the drawbacks of various families of warps and the problems concerned with diffeomorphism were identified, yet these were of greater interest to Marsland and Twining who posses knowledge of the more theoretical grounds. In Figure 7.6 lies a representation of a warp – that is – a reparameterisation curve that maps one point coordinate to another (and being a strict one-to-one mapping, it is a *bijection* as well). The idea was explained in some detail in 3.2.1.3 on page 55.

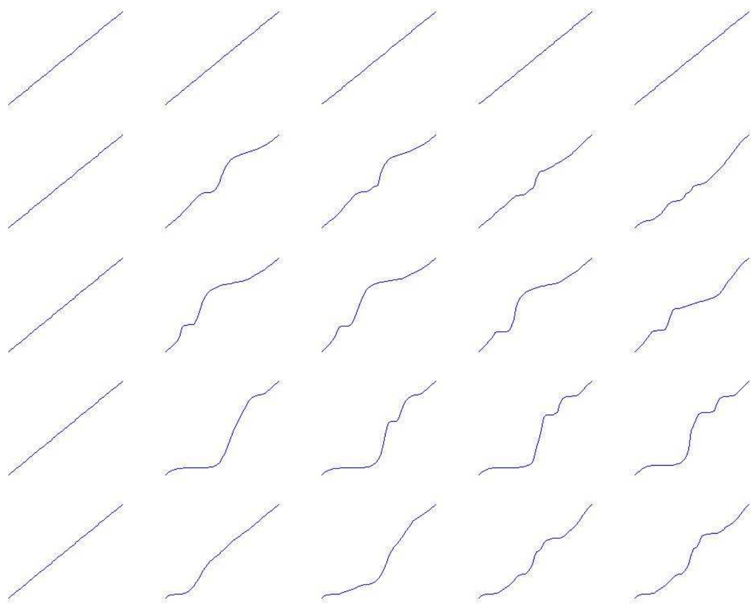

Figure 7.6: Warps shown as the MSD objective function runs. Each row shows the reparameterisation which is applied to one of the 5 images in the same row.

It should be noted that one data instance remains unchanged. That is in fact the reference which avoids the data from drifting away interminably. The left-hand side corresponds to the former iterations, the right-hand side – to the latter ones.

It was, at the earlier stages of this investigative study, vital to ensure that no cases of tearing and folding issues could arise. It turned out that one certain type of warp was problematic. A warp which was in essence made of a composition of knot-points (or control points in a more orthodox terminology for functions), also known as the multi-point warps, could produce unwanted effects and usage of that warp immediately ceased. Instead, a simpler single-point[8] warp has been used since, while the other was permanently conceded.

[8]This refers to the number of knot-points that are involved in the calculation of the transformation. A single point fully describes the Green's function which CPS builds upon.

7.4.1.3 Base-line Models

As a starting point for model construction and understanding, 10 regular data instances, which have been used in many of the experiments in upcoming sections, were generated and their statistical models built.

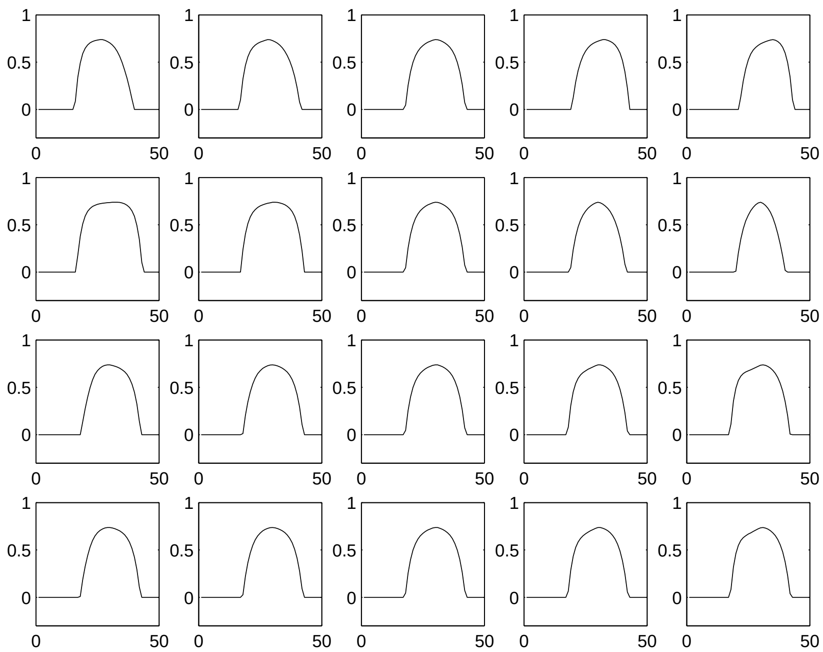

Figure 7.7: Shape model of 10 data instances at the start. The four principal modes are shown with up to ± 2 standard deviations away from the mean.

It can be seen in Figure 7.7 that shape is rather stable. This is because under these specific experiments, it was mainly (if not only) intensity that was used to create an appearance model[9] (more on this on page 109 in this very same chapter).

[9] The fact that shape component was chosen to be the reparameterisation curve has not been enlightened yet.

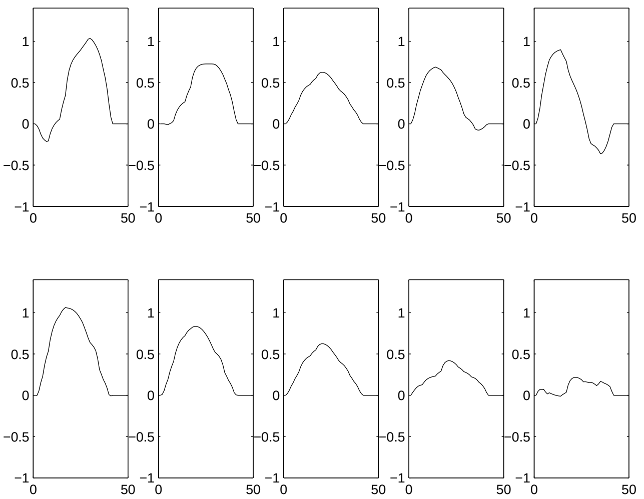

Figure 7.8: Intensity model of 10 data instances at the start. The two principal modes are shown with up to ± 2 standard deviations away from the mean.

Intensity models show some of the effects of height being changed, width varying and bump position moving from left to right. However, it is all rather fuzzy and the variation modes combine in a mysterious way. This is in fact why real correspondences need to be learned.

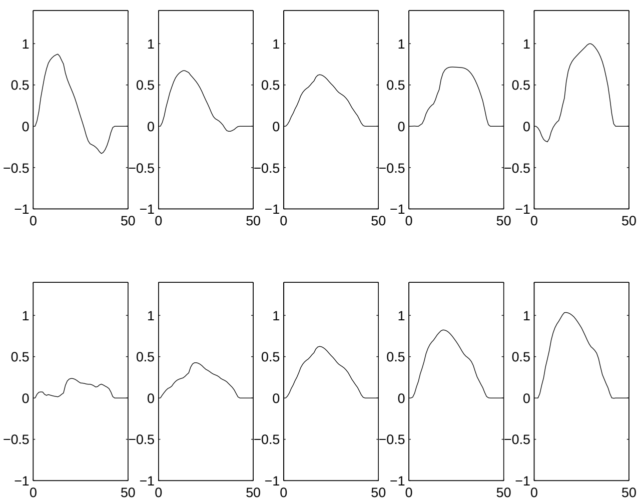

Figure 7.9: Combined (shape and intensity) model of 10 data instances at the start. The two principal modes are shown with up to± 2 standard deviations away from the mean.

The figure above shows the combination of intensity and shape. It is not yet too clear how to analyse it, but it resembles the intensity model which is the greater component that the combined model accounts for. Shape was quite static so it is expected to be merely invisible in the figure above.

7.4.1.4 Similarity Measures

Figure 7.10 shows one of the more useful measures for the model-based objective function. The MSD measure shows that the images become mutually similar as the model-based objective function proceeds, i.e. as the model of the data is minimised in its complexity. This confirms that a move is made in the right direction.

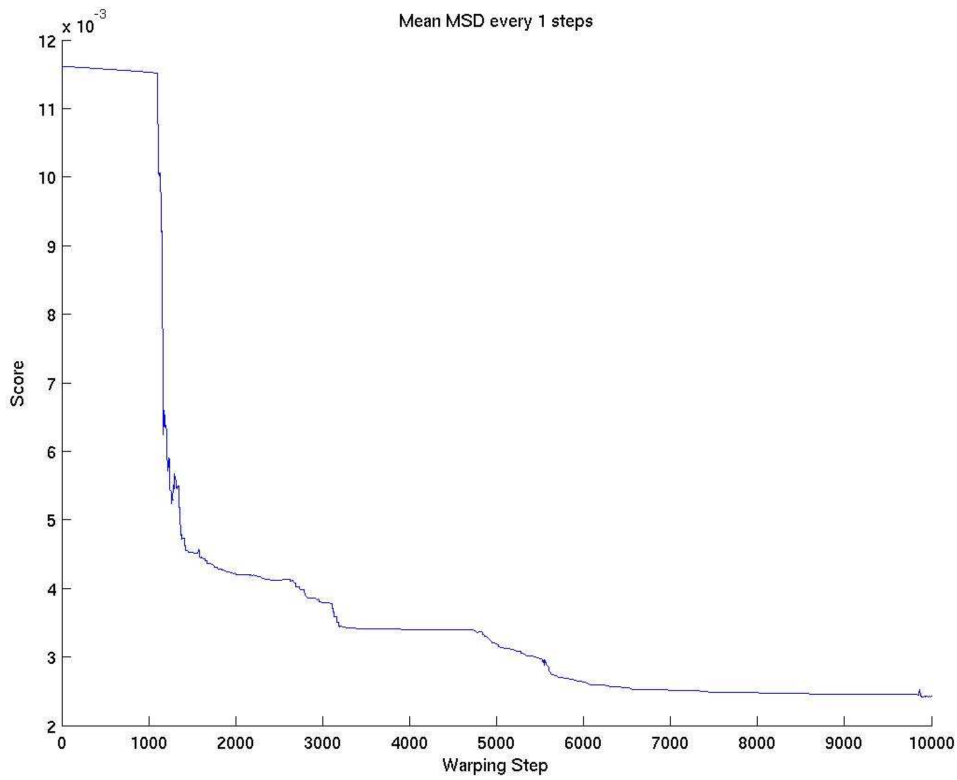

Figure 7.10: Mean MSD measures at each point during the model-based registration of 10 data instances.

Before investigating some of the different (and almost distinct) registration methods, it was worth looking at some measures and the way that these were affected as registration took place. Figure 7.11 shows the measures of MSD and MI for a group of 2-D synthetic data[10]. The large image shown in the figures is the reference image and the bottom plots show the measures for each of the 4 other images respectively.

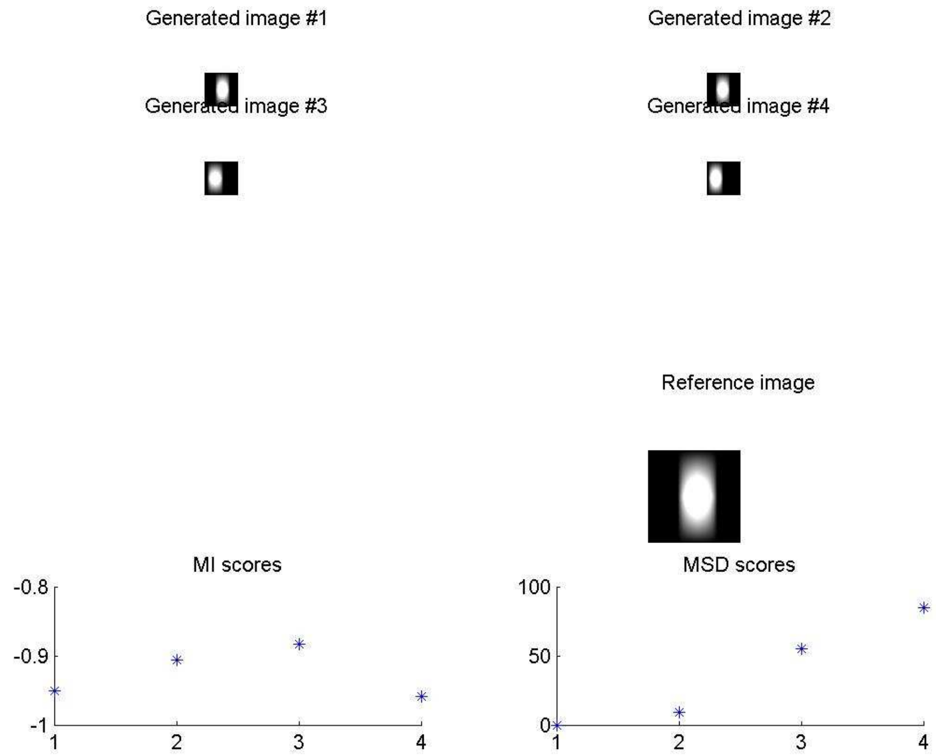

Figure 7.11: 2-D Synthetic data generated and evaluated for similarity against the reference.

7.4.1.5 Generalisation and Specificity

In the midst of experimentation, peculiar results were found when measurements of generalisation ability and specificity had been taken. The

[10] The data was generated by extending the 1-D bump data generator. It is *not* a Gaussian.

following is an extensive survey of these.

Specificity generates a number of random examples from the model and measures their distance with respect to the original set. Hence, it can be thought of as a measure of compactness. As the error bars in Figure 7.13 suggest, there is an improvement in the spread of these values (they spread shrinks in size) when the model-based objective function is employed, whereas it is not clear how the specificity rises when an MSD-minimising objective function is used.

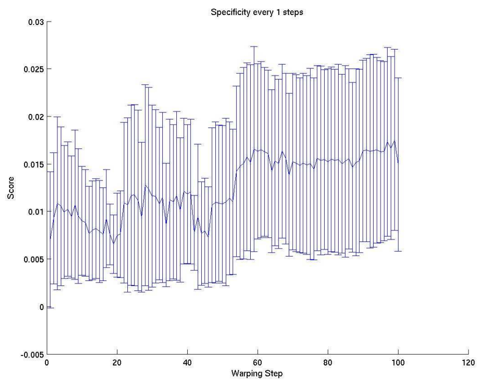

Figure 7.12: Specificity rising when MSD-based registration is performed.

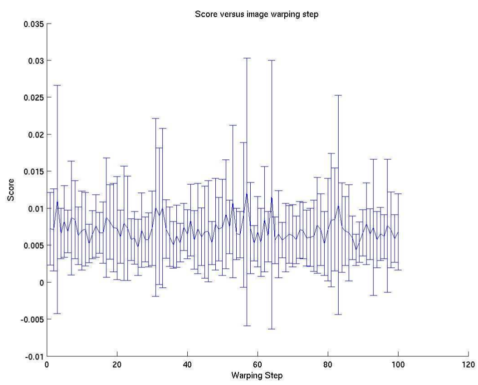

Figure 7.13: Specificity of model-based objective function.

Regarding the generalisability, there never appears to be a radical change in their range of values or mean values when the model-based objective function is applied. Likewise was the case for all other objective functions (e.g. Figure 7.18) so it appears as if it can be discarded as a measure of improvement. Only Figure 7.20 suggests that generalisability measures are of some use. However, generalisability measure very slowly changes.

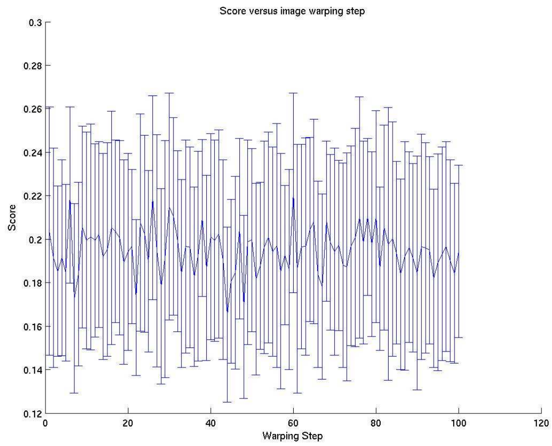

Figure 7.14: Generalisation ability of model-based objective function.

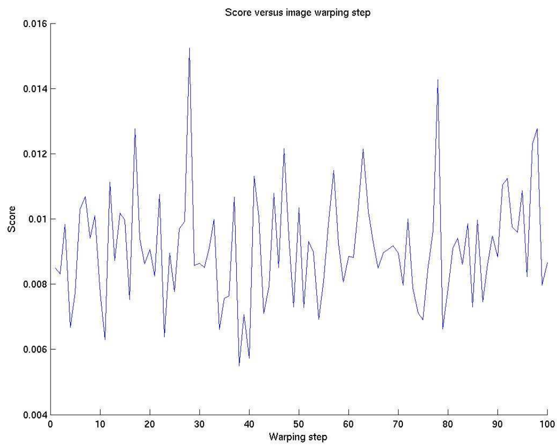

Figure 7.15: Specificity of the model-based objective function as registration proceeds.

A curious observation is that the model-based objective function may have its value decreasing at the start, yet no apparent improvement can be seen in the form of specificity. Figure 7.17 shows the steady value of specificity during the first 100 iterations of this model-driven algorithm. The model improves, but specificity does not. A similar story is said by generalisability (see Figure 7.16), but given the explanation above, it is not at all an unpredictable result.

More figures on generalisation ability are included for the realisation that it is a poor measure in the case of model-based registration. it should therefore not be pursued much further unless the algorithms alter.

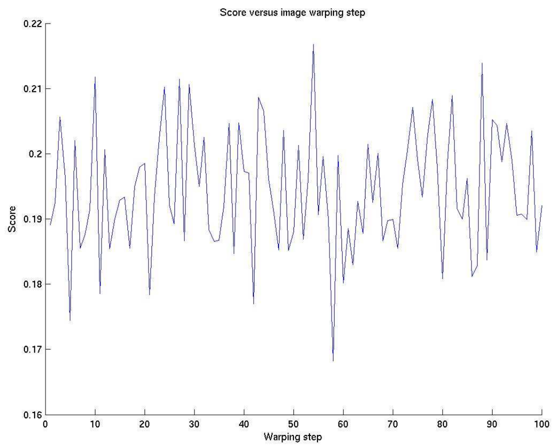

Figure 7.16: Generalisation ability of the model-based objective function as registration proceeds.

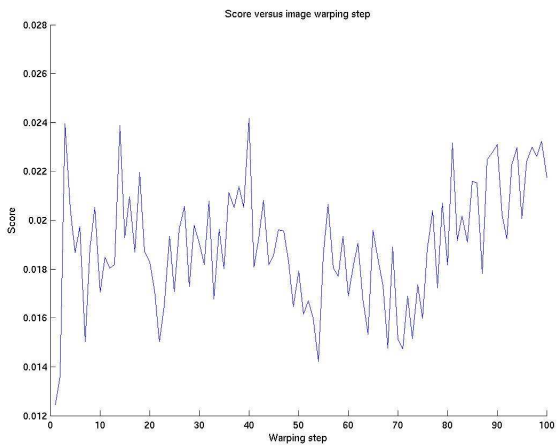

Figure 7.17: Specificity of the MSD objective function as registration proceeds.

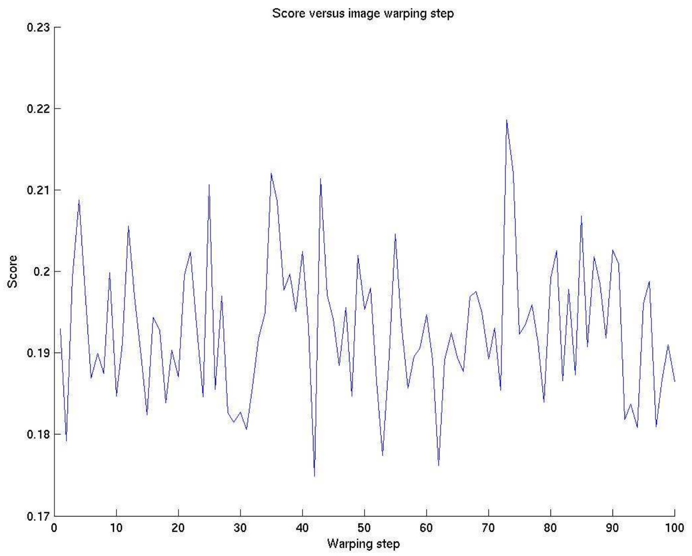

Figure 7.18: Generalisation ability of the MSD objective function as registration proceeds.

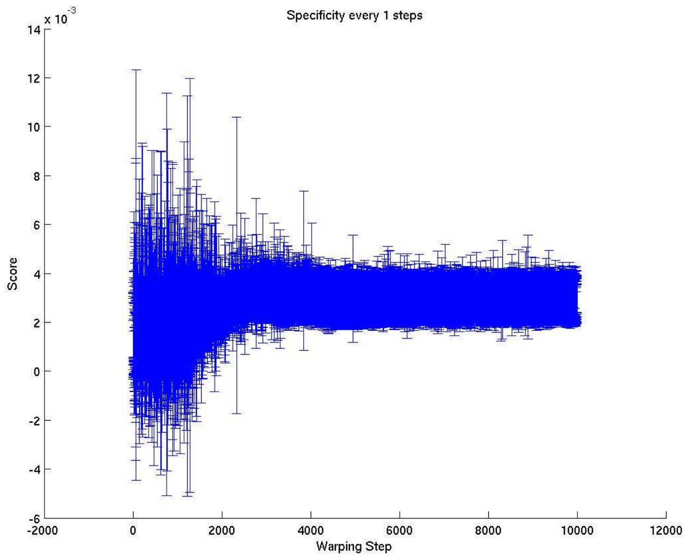

Figure 7.19: Specificity shown to be less erratic as the algorithm proceeds with registration.

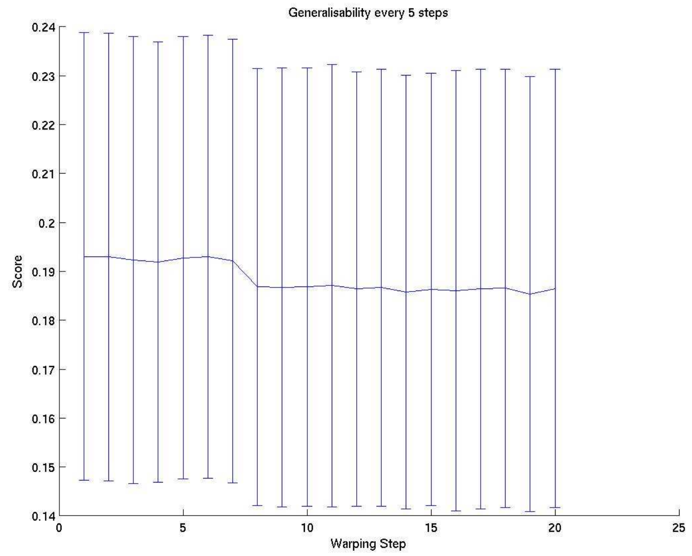

Figure 7.20: For MSD, generalisation slowly declines as shown for 2,000 iteration. Measurements are made every 100 iterations.

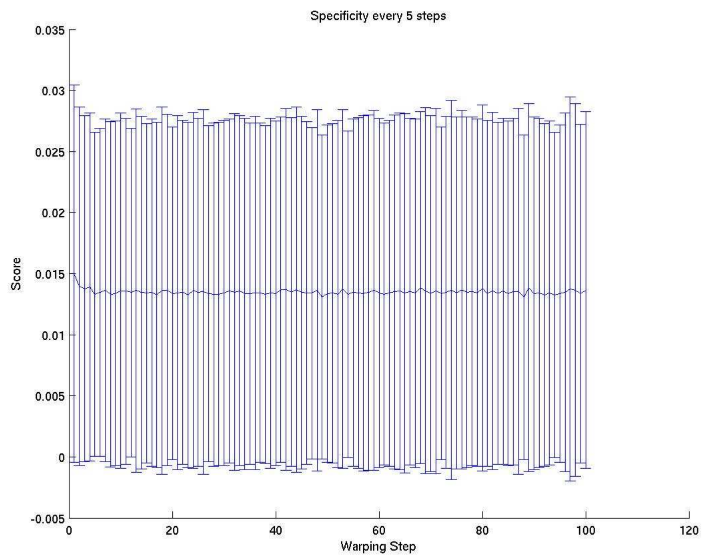

Figure 7.21: Specificity is merely unchanged as registration proceeds, unlike what is expected. It can be seen however, that there is a decline at the start where changes to the data are most radical.

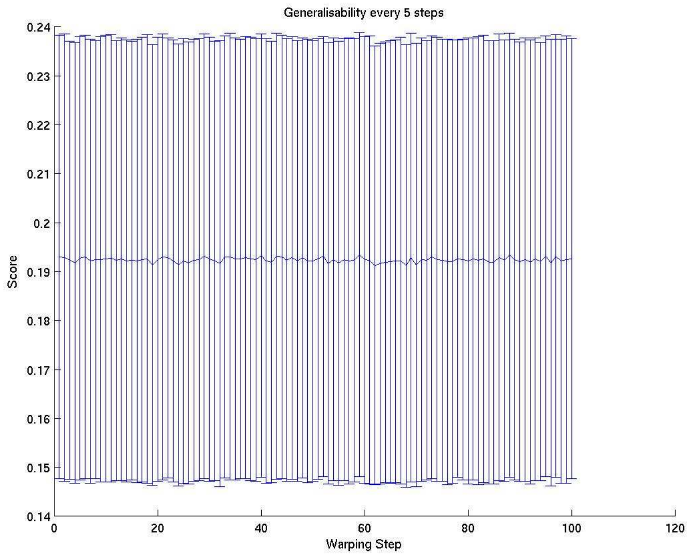

Figure 7.22: Generalisation ability measured every 100 warps. A total number of 10,000 iterations shows no substantial change to values while registration is performed.

7.4.2 Different Registration Approaches

The entire progress of this research began with some analysis and experimentation involving various registration schemes in a single dimension. These were quite naïve implementations[11] and MSD performed conspicuously well to result in the best ultimate similarity, whereas other methods barely had any positive effect.

One objective function that would of course be of significant interest was that which is based on models. As anticipated, its improvements were made smaller and smaller as time progressed and for several months it failed to reach a good solution (depicted by intermittent red in the figure below).

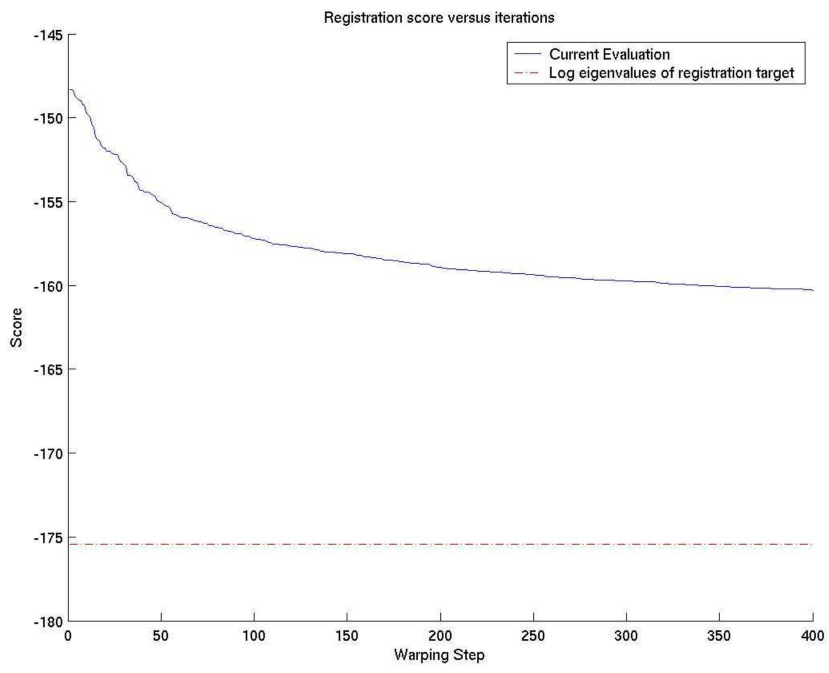

Figure 7.23: Multiple knot-point warps show that the curve is exponential when a model-based objective function is employed.

[11]More cunning implementations would have involved better 'dialogue' between the similarity measures and the warps chosen, for example.
Rather than that, each of the two components was treated as a black box, fully independent from the other.

7.4.2.1 Model-based Objective Function

A valid prototype for this registration method was already in place at the start. The way this function operates has been explained in earlier parts of the report. Figure 7.24 shows how measures of MSD, Generalisation ability and specificity change at each function evaluation step. The changes are expressed per cent with respect to the evaluation taken for the previous iteration. It appears rather noisy and reflects on the bad forms of this function before it was revised. As most lines remain flat, it is shown that values go in no particular direction[12].

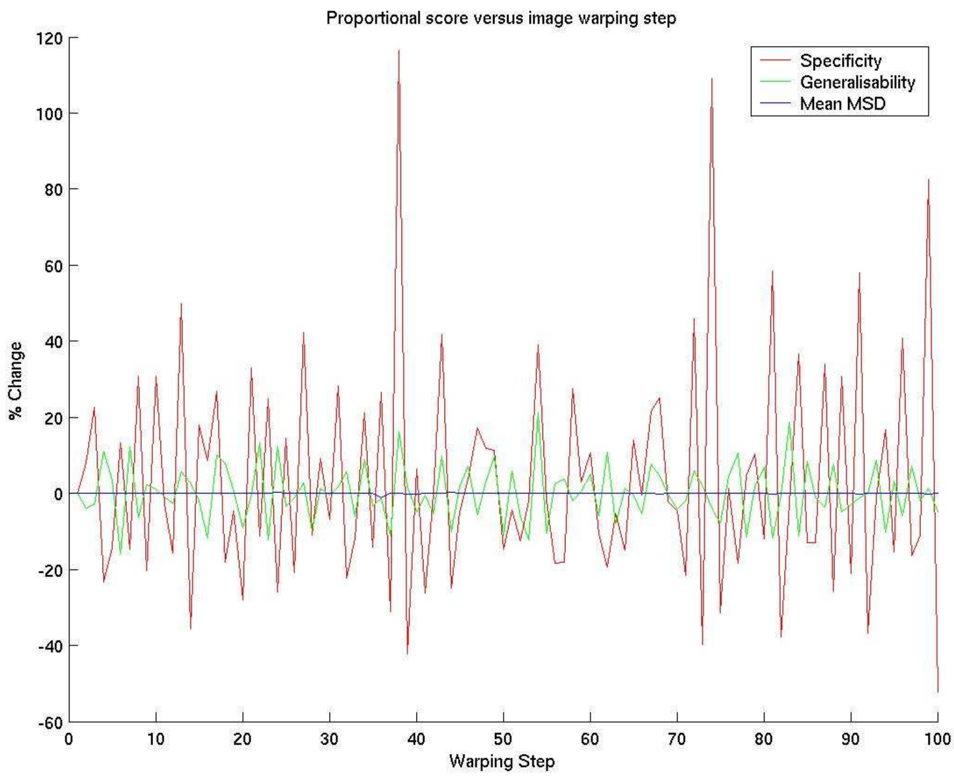

Figure 7.24: Various measures shown as the old model-based algorithm proceeds.

[12]This idea is borrowed from technical analysis in finance. It can be useful in science as well.

Due to this poor performance, rigourous work began to obliterate known issues and weaknesses, resulting in an improved model-based objective function.

7.4.2.2 Improved Model-based Objective Function

As already adumbrated in the text, towards the end of April 2004, many solutions were found which substantially improved the objective function and finally made it work. Further options also made it work relatively efficiently and obtain impressive results. Details on the changes which were applied will be explained later in this section.

Figure 7.25 below illustrates how registration practically operates upon the data. It can be observed that, in this case, while the model-based objective function guides transformation, data bumps align increasingly better. Therfore, good registration is finally achieved, driven purely by model complexity.

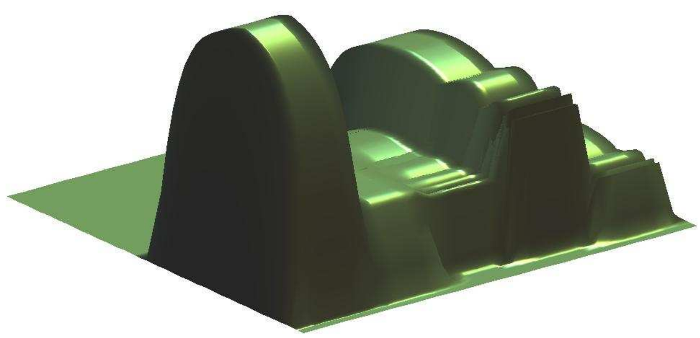

Figure 7.25: Data being visualised by AART. In this case, 5 bumps are shown at some arbitrary state during registration.

See this with reference to a previous figure that was shown on page 89. The model-based algorithm itself will be explained in Section 8.

7.4.2.3 Emergence of New Methods

New methods and extensions of existing ones soon began to emerge. As part of an evaluation of many techniques, with the aim of justifying the use of model-based functions, new ones were created with some logical backing. The majority of these are detailed below.

- **Pair-wise model based**:

  There have been several attempts to properly register data by creating models of the reference and each image in the set in turn. This is an interesting idea to look at because, in principle, models can be proven to be good pair-wise measures as well.

- **Probability Density Function (PDF):**

  Such functions describe the volume of data distributions. More uniform data, as one aspires to achieve across all images during registration, will result in lower such values. An exponential PDF was used in the experiments by default although over a dozen others are available in AART, including a Gaussian one. In line with Cootes' implementation for the ECCV 2004 paper [9], this PDF-based function was created and used amongst the different objective functions under evaluation.

- **Wavelets:**

  A personal suggestion was to use wavelets [45] as indicator of data complexity. It was inspired by Twining's mentioning of Fourier transforms. As compression is closely related to MDL, these can provide an accurate estimate of the complexity of data and abundance of patterns within that data. An extensive group of different wavelets are offered by the application and, by default, Daubechy was used in the experiments. Computationally cheaper alternatives to the wavelets are Fourier and Hough transforms, but these have not yet been incorporated into AART. All wavelet implementations were supplied by the MATLAB Wavelet Toolbox.

- **Mutual Information:**

  This strand of methods [61, 53] will analyse the peaks of image histograms. Normalised MI is currently one of the most robust and widely-used methods for 2-D data.

- **Hybrid Objective Functions:**

  Much earlier in the year, a combination of objective functions was investigated, mainly that of MSD and model-based. One such hybrid method performed an MSD-driven routine, followed by a model-based one. Under such approach, it is assumed that the model-based objective function is well-behaved near convergence. Other schemes combined and altered between MSD- and a model-based objective function every fixed number of iterations (the algorithms were operated in alternating cycles).

7.4.3 Comparisons and Benchmarks

7.4.3.1 A Comparative Analysis

As Form 2 (page 176) indicates, a quantitative analysis of different methods was needed to infer something about their behaviours. See Figure 7.26.

7.4.3.2 Comparative Analysis of Objective Functions

In late 2003 and in early 2004, a comparative analysis of different registration methods was conducted. Some results are shown in Figure 7.26.

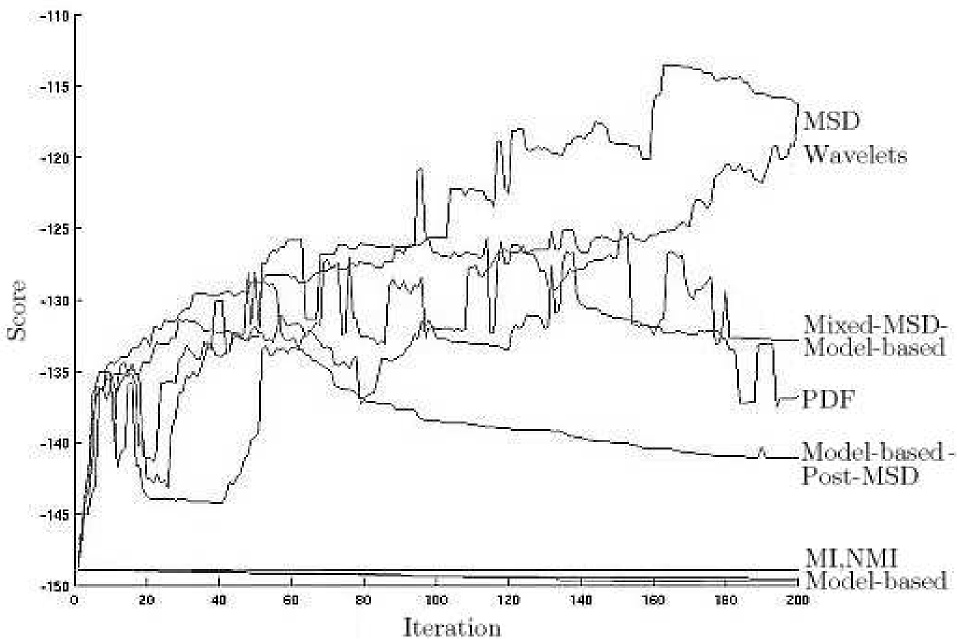

Figure 7.26: A comparative analysis of different objective functions. It illustrates that the model complexity decreases only for the newly-proposed objective functions. The Y-Axis value is an indicator of model compactness.

One of the primary aims was to benchmark different registration methods and come up with comparative results which highlight the up- and down-sides of each method. There had been a particular interest in the underlying behaviour of each method and the quality of registration as evaluated by a model of appearance.

Months later it was discovered that the registration method which had later been proposed had the potential of becoming much more successful. Up to a certain point in time, functions used for registration were simply unable to get decent results. It was revealed that the transformations applied were restricted to remain small in extent. The problem was resolved by changing this restriction term, whereupon larger, more radical transformation were permissibly applied and a good solution was shortly approached. The issue of speed (or efficiency) remained a worrying factor. It had to be addressed in order to make the registration method more practical in 2-D (and potentially an even greater number of dimensions).

7.4.3.3 Comparison Quantitatively

A comparison between most of the methods was conducted and the conditions were set to be impartial and well-scaled so that they evaluate a proper registration process.

For the results in Table 7.1, the number of iterations was set to 50. By another terminology[13], this equates to 1000 as each of the twenty data instances was subjected to up to 50 transformations. For single-point transformations, the placement of the control point was random (both in location and magnitude) and for multi-point transformations the positioning of points was made random to abstain from data-bias or advantageous *a priori* knowledge. The number of data instances was kept high at 20 in order to allow a substantial *group-wise* optimisation to be investigated. Objective functions based on mutual information remained flat simply due to the continuity of the data and the fact that it is one-dimensional. The table below shows the different values of $log \prod \lambda$.

[13] In AART, this definition of iteration is repeatedly referred to as *warping step/s*.

Objective Function	Single-point Warp	Multi-point Warp
PDF	-137.2658	-136.7145
Wavelets	-145.4988	-147.7877
Joint Model-based	-149.2192	-150.3197
Sequential Model-based	-148.7245	-149.9904
MSD	-143.0227	-149.0114
Joint MI	-142.3415	-136.0651
Sequential MI	-142.3712	-136.0651
Joint NMI	-142.3154	-142.3068
Sequential NMI	-142.3154	-142.3118
Model after MSD	-138.6823	-47.0961
Mixed Model/MSD	-129.3791	-105.3422

Table 7.1: Comparison of objective functions. The values indicate an approximation to model complexity (its determinant).

λ are the Eigen-values derived from the covariance matrix of the appearance model which had been constructed from all 20 data instances[14]. For completeness, differentiation is provided for optimisations which reparameterise over all dimensions at once (joint) or do so separately (sequential).

7.4.4 Problems Investigated

7.4.4.1 Varying Weights

In May 2004, a further investigation of the ratio between shape and intensity began. As a result, ways of stabilising the objective function at a *low* convergence point, may have been identified. Otherwise, with this ratio improperly set, convergence (though not a full one) continuously appeared well above the correct solution. Figure 7.27 shows what happens when the ratio $\mathbf{W}_S$ is inadequately picked.

[14]As 5.4.2 on page 70 explains, an extra term, epsilon, is used to refrain from multiplication by 0. Due to the finite precision of digital systems, Eigen-values may be assigned this zero measure.

In the past (and seldom at present) the value of $\mathbf{W}_S$ was chosen either to be a pre-set constant provided by the user, a value which is derived from the image derivatives, or a value that is proportional to the variances. In either case, it was always found to be overly high. When this value winds up taking shape into account, the objective function quickly fails to improve as shown below.

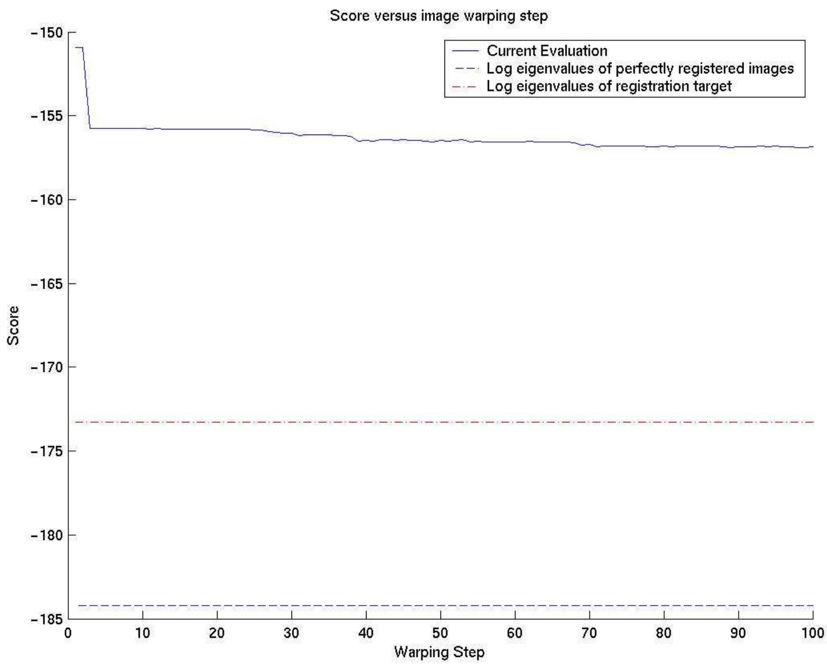

Figure 7.27: The old model-based objective function which gets stuck due to Ws inappropriately set.

Going back to Figure 7.6 on page 91, it can be seen what the *shape* is actually defined to be. The combined model is that which takes into account image intensity values along with the *warps* that accompany these newly-deformed values. The problem encountered resulted from the fact that all these curves were initially linear[15] and mutually identical. In other words, the shape defined for all data instances had no variance at

[15]The curves were all going from the bottom left to the top right corner, meaning that each point mapped onto itself and no changes were made.

all. It was therefore hard, using a proper combined model, to 'lure' the objective function to depart from that point of low variance. Warps were simply thrown away once they had been chosen.

7.4.4.2 The Curse of Set Size

At this stage, the model-based objective function could only cope well with set sizes that were rather small. It found it difficult to minimise a model by altering just one instance whose overall effect on that model was minute.

This problem was in no sense new. In a model-driven objective function, such as in the work of Kotcheff and Taylor, alteration of one single data instance does not affect the model considerably. The greater the set size becomes, the lower the effect which parameterisation (or in this case, image warps) have. The only exception to this is when a warp is applied uniformly to all data instances. In the case of registration though, it is impractical.

In order to deal with large enough problems, where for instance, dozens of images need be accounted for, resolutions need to be found that make convergence linearly proportional to the size of the set. The problem was also well-acquainted in the work on landmarks selection where sets remained 10 or 20 in size.

This fundamental problem suggests that a model-based approach is limited. It is not yet sufficiently well-behaved to study a *population*, only a smaller-scale case study.

7.4.4.3 The Hindrance of Speed

Speed was a worrisome issue, for multi-knot-point warps in particular. The calculation of a Green's function output for some given knot-points was not the real culprit. It was not the centre of slowness of propagation.

It was its application to an image which is followed by Eigen-analysis[16] which made the approach altogether slow. Ways were needed to be discovered to eliminate this as an issue. Some of the following solutions are inter-related to speed.

7.4.4.4 Optimisation Issues

The optimisation regime needed to be changed as it often got stuck without any constructive paths to a solution being found. It was at first unknown whether the objective function was malformed, or perhaps its handling by the general-purpose Nelder-Mead optimiser was not effective. It could be seen that there was no justification in believing that the function was truly stuck. Sharp drops in value could occasionally be observed, meaning that good warps were finally found and applied to the data. Figure 7.28 makes it rather crystal-clear.

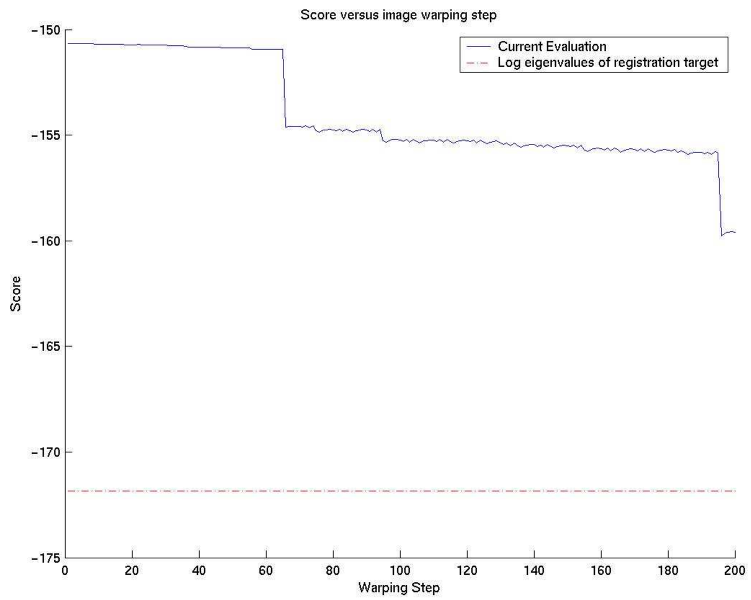

Figure 7.28: Drops which illustrate the problems with optimisation.

[16]There is a serious flaw present because this analysis is cubic in its complexity. Profiling is yet to be considered an option for improvements discovery.

Even at present some moderate drops are observed at times. With a less stochastic choice of warps, that unwanted effect might vanish.

7.4.4.5 Finding the Correct Solution

One essential step, which can be used to understand the algorithm's behaviour and the direction it takes, is that which infers a state which is optimal – that is — a state that the objective function must reach when it behaves correctly. Analysis was performed to discover which warps, when applied to the data available, give a sensible (or perfect) result, i.e. *align* the data. Figure 7.29 shows a set of correct warps with the corresponding set of results. These help in verifying that registration of data is finally obtained.

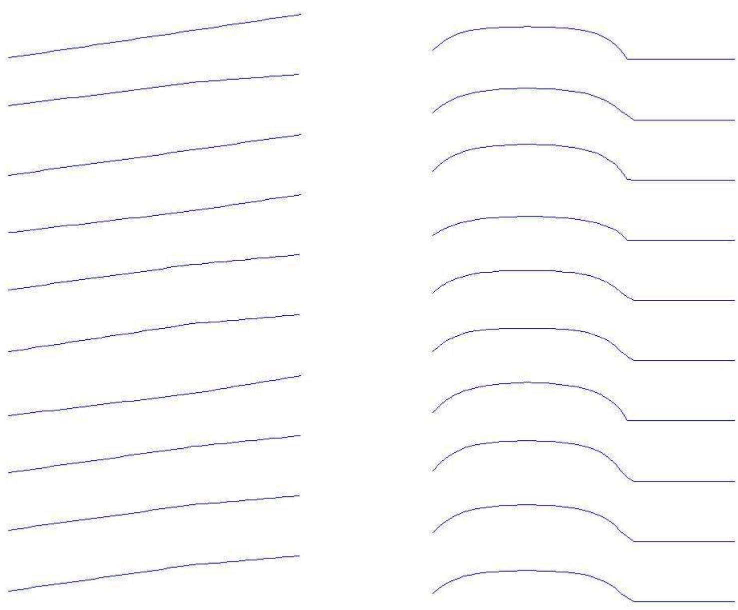

Figure 7.29: Data alignment to discover correct solution. On the left: piece-wise linear warps to be applied to original data; on the right: data after alignment.

The choice of a reference of course effects the 'correct' set of warps, but the notion of correctness is unchanged. In the case shown, reference is chosen to be the first of the data instances, meaning that it is assumed all data needs to be transformed based on that one image. It is warped to *fit* the first image. In practice, only the actual value is used to evaluate the performance of the model-based objective function. This means that this alignment is one amongst several possibilities, but its *value* provides an excellent estimate. Also worthy of mentioning is the fact that an immutable reference is maintained for the model-based case. This is why results are expected to be ultimately similar the the ones above, where that very same reference is chosen.

In AART, advanced caution is dispensed to avoid reference choice that damages data integrity. It is safer to choose as reference data that is representative of the whole set (essentially one which lies in the centre of the multi-dimensional cloud in vector scape). Reference is chosen which lies closest to the mean of the set. Definitions of distance are alterable. Currently, sum of squared distances[17] or fixed geometrical distances can be used in AART to locate a good reference.

7.4.4.6 Registration Target

As well as knowing how transformations behave and how they affect the data, a measure of model quality needed to be established and plotted against steps in the algorithm. To make this value more meaningful, the value that one aspires to reach was estimated and shown in the plot. Previous figures, as well as later ones, include this measure, which was calculated rather easily having got the correct solution, as described in 7.4.4.5 above.

[17]This measure is better immune to lage local misalignment. This is similar to arguments presented in Equation on page 40.

7.4.4.7 Solution and Perturbation

When the correct solution to registration was known, it was intriguing to see how the function would behave near that solution. Ideally, it should return to that correct solution quickly and quite controllably (in the sense that no abrupt data changes occur in the process). Two types of perturbations were attempted: random noise and a randomly-placed CPS warp. Results showed that the objective function failed to revert the data to its form in the correct solution. In fact, the objective function value dropped *below* that which was expected. This instability of the objective function is explained later in this subsection.

7.4.4.8 Interpolation Artefacts

Whenever a correct solution was calculated, an unwanted artefact appeared near the edges of the bump. This was later on realised to be a result of interpolation which could be resolved simply by increasing the sampling resolution[18]. This can possibly be seen (although it is rather subtle) on the right column in Figure 7.29.

7.4.4.9 Change in Data Generator

The process involving data generation of bumps was made more accurate so that unblemished half-ellipses would be invariantly created. This did not produce any better results, though reasoning about the correctness of data was no longer a worrying factor. Synthesis of data is now a richer component of the program with several more configurations to manipulate.

[18]The number of pixels that data comprises is an argument that may be changed.

7.4.4.10 Objective Function Instability

Once good results were found, certain similarity to the work of Davies (see 2002 thesis) was encountered. There was a linearly (almost logarithmic) flat drop in the objective function value. It was decided that such problems need to be resolved at once. For both images and shapes, this was now an important issue to address in a principled fashion.

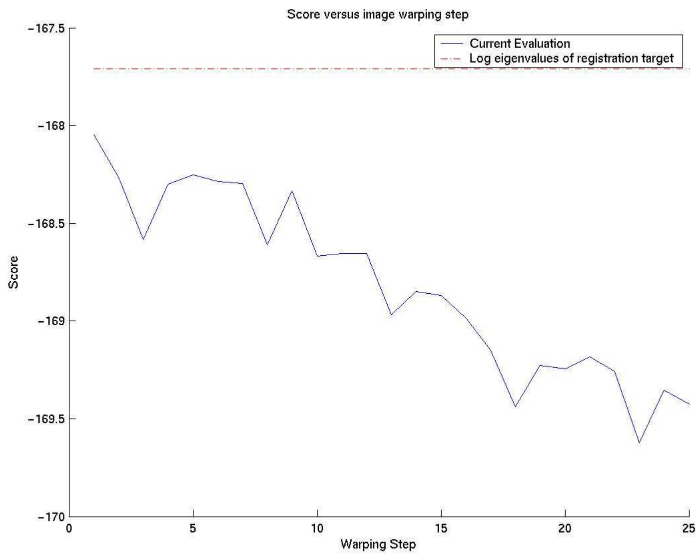

Figure 7.30: Evaluation going below target when initialised at the registration target. The target of registration is indicated by the straight horizontal line.

Twining in particular was one person who could suggest ideas or provide help on the matter. The problem can be overcome by using knowledge on models and comparison between models and their constituent reconstructed instances. This process of comparison allows the corresponding discrepancies to be determined.

Figure 7.30 shows that when the registration algorithm is initialised at the conceived correct solution, it can still slide below it. In fact, it always does. This suggests that the algorithm is not controlled properly and that the description length term can miss the correct solution.

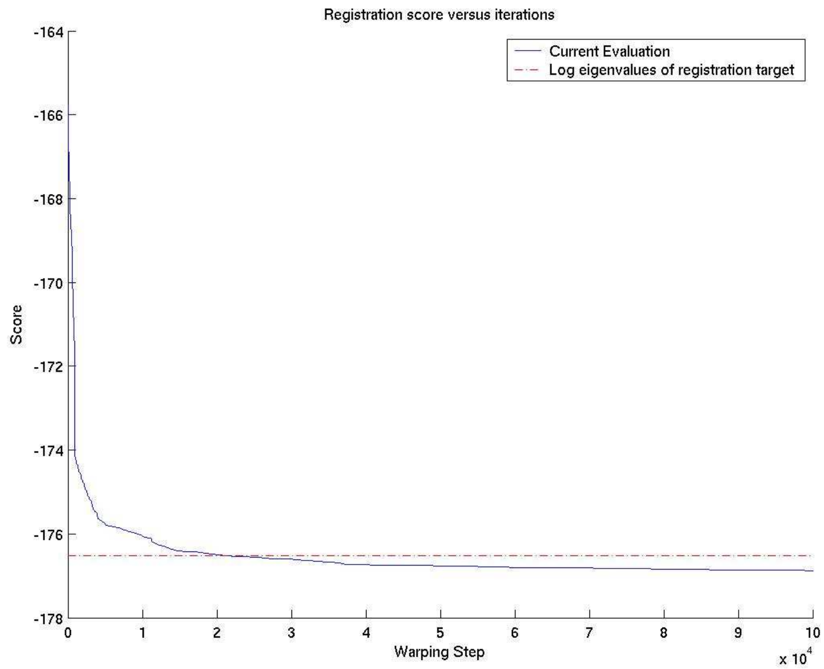

Figure 7.31: A long optimisation with the successful algorithm shows that it surpasses what is questionably the correct solution.

Figure 7.31 should make it clear that given a large enough number of iterations, no clear convergence is reached . Even more problematically, the value of the objective function slides below the point where it ought to have been optimal, by definition.

7.4.4.11 Other Synthetic Data under Inspection

During the process of objective functions investigation, while debugging in particular, simple sets of synthetic data were used. These are not worth any detailed exploration in this report, but results are wholly recorded in the WWW. Different data types can be selected from the Patron menu in AART and external data can also be imported directly, e.g. when real 1-D data becomes of use, if ever.

7.4.5 Attempted Solutions

Having identified many issues, this subsection presents a few ways of circumventing them.

7.4.5.1 The Successful Solution

A successful solution was found to almost the entire problem of registration or at least its major weaknesses. This was done simply by changing the optimisation so that it searches a *valid range* of possible warps (the space of warps in essence) and runs for a long enough period of time to benefit from computational savings. Also, an increasingly higher tolerance was sought for the optimiser, whereupon registration using statistical models finally become possible and even successful. More on these solutions and successful heuristics are to be listed in the remainder of this broad subsection.

7.4.5.2 Speeding up Convergence

The objective one should be after is the minimisation of some cost – a cost that is associated with the model and the images which are being transformed. Convergence of the algorithm is declared once that cost can no longer be reduced. It has been found to be true that very much computational effort is spent on refining existing transformations, even when the solution is yet far away. As it turns out, at early stages of the optimisation, coarse changes to the images should suffice.

Amongst the more important developments of this project, the author managed to find a way of substantially decreasing the amount of time which is required to register sets of 1-D data. This was done primarily by tweaking the optimiser which is involved in the process. It has been decided to aim for a different optimisation tolerance depending on the advancement towards the correct solution. By doing do, little time should be spent on obtaining lower costs at the early stages of registration. This rational observation motivated the re-implementation of a similar algorithm in a separate domain[19]

7.4.5.3 Height Being Forced

Coming back to Figure 7.30 on page 116, there was a serious flaw that is associated with the function's insufficiently constrained form. It was possible for the model to be improved by 'cheating' and concealing parts of the data.

There was a period where in order to avoid data breaking and bumps shrivelling, a height was being forced to remain the same at the pinnacle of the bump. Several different schemes were attempted. At times, the position for insertion was projected from the mean and later it was derived from the model discrepancy. The results were not of very high quality at that stage and Figure 7.32 shows one such case. When run over a long period of time, a sharp tip can be observed which is due to points being forced to lie in a single fixed position. On other occasions, parts of the image are still being hidden. This basically results in gradual darkening of images subjected to registration.

7.4.5.4 Point Insertion

An issue was latterly encountered where data drifted away quite slowly. Consequently, convergence could never truly be reached. The issue is

[19] This domain is the selection of landmarks in shapes for the construction of statistical models of shape. Modification of this will be described later.

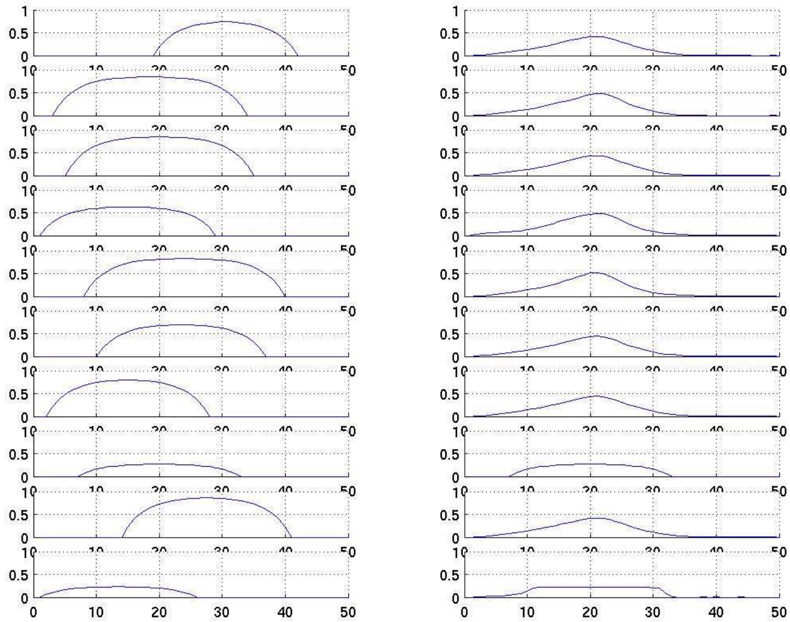

Figure 7.32: Highest peak being retained in registration.

worrying as it was found in Davies' work as well. This leads to the next point which is the model residuals. Present work places great emphasis on this matter.

7.4.5.5 A Parallel Discovery

Points are currently being inserted where the variation (among the set of data) is the greatest. There are different methods to do so (see menu layout in AART for better insight). It was later realised that this insertion of points is analogous to work done by Tomos Williams on shapes and this insertion must take into account the model and its discrepancies. Observations of this nature actuated the work described in Section 7.5.

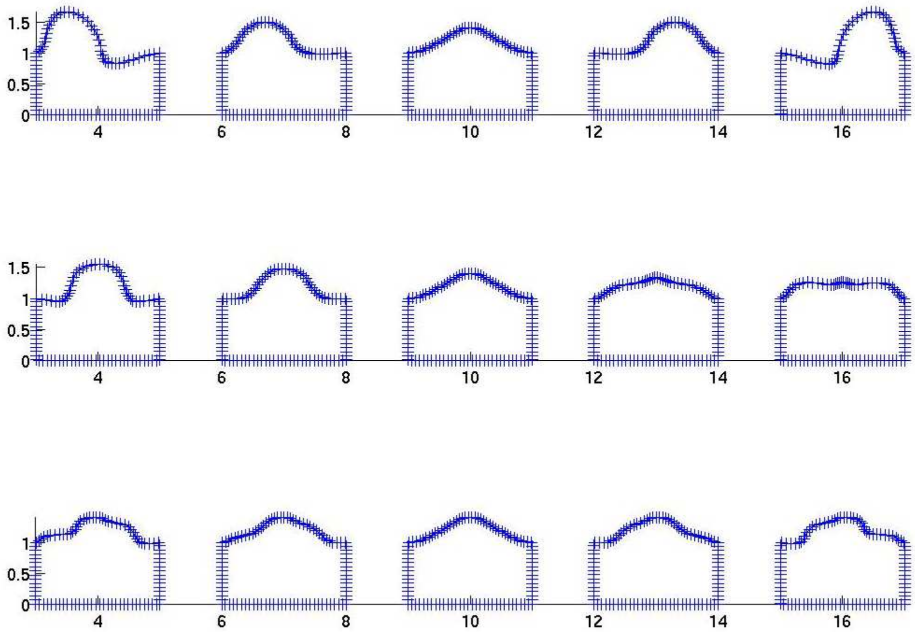

Figure 7.33: The unregistered bump data and its three principal modes of variation (± 2 standard deviations).

7.5 Shapes Revisited

Upon returning to the subject of shapes and automatic landmark selection, goals could be laid out perfectly well. There were the issues of speed and instability of the objective function. Also, it was vital to make the code work in the absence of its original developers and authors. These problems were all shortly embarked upon and some successful solution were found. Figure 7.33 shows the data which was typically handled[20], namely the brick-and-bump data. The 3 principal modes of variation are shown for the raw data at the start when points are spread at equally-spaced locations along the curve.

To ease the operation of the code, an interface (Figure 7.34) was built and used to bring the code back into workable order. The process otherwise

[20]Examples of human hands were later tested as well . They were an easier case that is quicker to reach convergence.

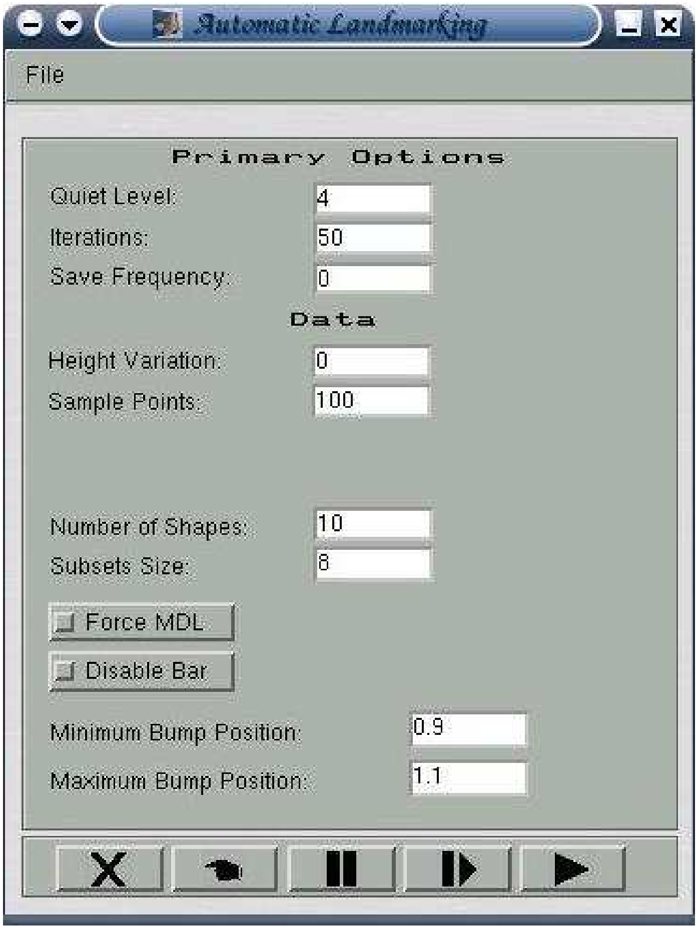

Figure 7.34: The graphical user interface for semi- automatic landmark selection as of June 2004.

involved manipulation of the code which requires the user to know the algorithm. This situation proved to be difficult and tolerable at best.

7.5.1 The Successful Use

After applying some changes to the code and running some over-night experiments, the code could finally be used in a way which made it usable and effective. At first, the original bump data was used with varying levels of variation that is inherent in the data. At a later stage, hand data was used as well to test a more realistic and interesting example.

Once the code was in full working order and an interface was in place, experiments could be set up and run very rapidly. The rest of this section describes some of these experiments and their corresponding and related equivalents. Many of the same ideas were parallelly applied to the code handling images and their model-based registration, namely that of AART.

7.5.2 Subsets Selection

This idea is concerned with sub-division of the problem and localisation of computation which otherwise becomes very demanding. Since the model constructed is conventionally build from the entire set under consideration, there is a clear correlation between the complexity of the entire problem and that size of the set. This correlation is not linearly proportional to the size of the set either. To get decent results (either in landmark selection or image correspondence), a very long optimisation is required for increasingly larger sets of data. Although the principles are genuine and sophisticated at first sight, they suffer from this unappealing relational complexity.

The next section explains this principle in the context of images. It was decided to try to apply the same concepts to shapes after it was arguably successful when images were under consideration.

7.5.3 Adaptive Precision

The source of this useful strategy came from the registration algorithm. It was found that coarse and hastily-chosen transformation sufficed at the former stages of the optimisation. It was premature to require a high precision from the optimiser at these stages. Putting this idea in different terms, the machine chooses to loop laboriously looking for very meticulously-chosen warps when, in fact, since all elements in the set is dynamic, a slack choice of a warp serves the overall aim. Sets are manipulated one instance at a time and the problem appears quite different every time a loop is completed. By lowering the precision demands, more iterations can fit within a tantamount time period (see Figure 7.36 on page 127).

The subject of precision and its impact on speed will be revisited in Section 8.4. It is an essential strategic choice which is suitable for a problem where its nature is dynamic[21].

7.6 Present Work

The writing-up of this report unsurprisingly interrupted some experiments in progress. There are partial results which establish a clearer path to more experiments of possible use. Future experiments in Chapter 10 describe some of these yet-to-be-performed experiments and their closely-related goals.

7.6.1 Model Residuals

Attempts were made at the identification of the point of convergence for the registration algorithm. A clear flaw with the cost that had been defined was discovered. It turned out that the way in which the algorithm presently evaluates models neglects to account for small artifacts. These artifacts *must* be encapsulated in this model. More crucially, these small artifacts which are left-out residuals need to form part of the model cost (description length term). In their absence, the objective function was able to drift away, thereby hiding vital structures in images. Not only

[21]As an example, the famous travelling salesman problem speaks of a pre-set value for each edge in a graph. This means that the problem never changes. What if the values of edges changed for each choice of a path? This renowned problem would then become less workable then it has become. Each choice then introduces a new, yet unknown, optimisation problem.

was the result of registration poorer due to an improper model cost, but also it was impossible to contend that one unique solution can ever be reached.

It is now realised that the objective function must account for the model errors in some way or another. This is why the residuals for each image, as reconstructed from the model, need to be calculated. All inverse warps needs to be calculated first to do so – a step that is not trivial (on-going and future work is scheduled to be done on this).

Before this incorporation of residuals goes on, it was suggested that the shapes problem is looked at again. It is believed that description length should have a term accounting for model discrepancy so that the optimisation can be made stable. Technicality concerning MDL is expected to be discussed with Carole Twining in the near future. Eventually, a term must be precisely defined to account for the residuals and, having solved the problem for the simpler and well-founded case of shapes, the less trivial case of image registration can be resumed. Several discussion have raised disagreements regarding the way in which residuals are defined in the context of appearance models, images and transformations.

To summarise, of current interest is the way in which model residuals can nullify erosion of data. They can be used to compose a proper description length for models and images. By resolving such a problem, better registration performance will be yielded.

7.6.2 Adaptive Precision for Images

This work is associated with the change in optimiser tolerance and the context is now different. This approach was never tried previously, but it seems to get good results and it incorporates a novel and elegant algorithm which seeks a multi-scale approach to choice of warps. Rather than scaling the data, it scales the level of warp quality and, in that sense, it is multi-scale in an unorthodox sense.

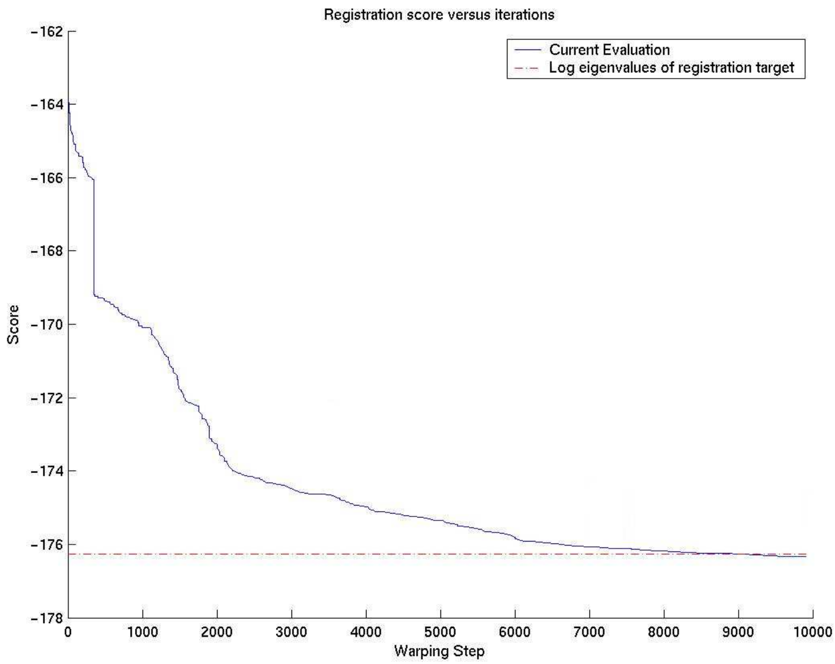

Figure 7.35: Automatic precision and the differing rates of convergence for image registration.

Figure 7.35 shows how the choice of tolerance (or precision homologously) affects the rate of convergence. What is less obvious to the mind, is the way this optimisation improves in terms of time. While this is a possible experiment to perform at present, registration versus time will be shown only later, when a different approach is discussed.

Figure 7.36 shows quite clearly the differing rates of objective function improvements in its evaluation. Different curvatures correspond to different choices of tolerance.

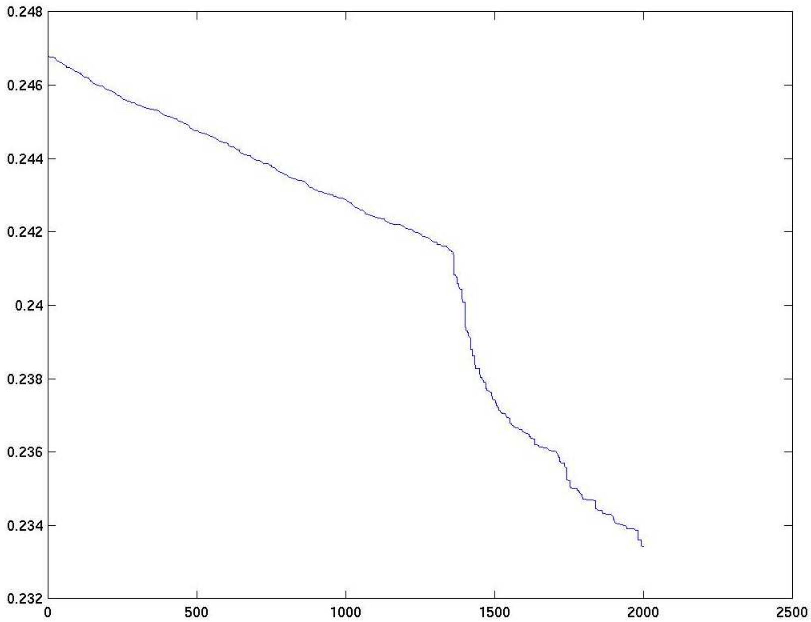

Figure 7.36: Adaptive precision requirement resulting in different rates of convergence. This curve is drawn in the context of shapes and selection of landmark point.

7.6.3 Varying Set Sizes

The idea has been discussed before in a lesser extent. This subsection shall provide some results along with basic analysis which abstains from drawing any final conclusions.

7.6.3.1 Shapes Subsets

The stochastic choice of subsets attempted to speed up the construction of models through simplifiction. The choice of subsets was at first made at every single iteration. At later stages, such a choice was only made once within a set cycle (e.g. 10 or 100 iterations). This intended to allow the objective function and the algorithm to stabilise and deal with a somewhat similar problem repeatedly. When subsets are not reshuffled often, the subset under consideration retains more commonalities.

This idea was expected to result in reduction in run-time. That is because Eigen-analysis is then being simplified and, along with it, the scope of the problem is reduced. When sets are smaller, they are strictly easier to handle.

Unfortunately, the approach taken above worked badly in terms of time. Its performance, as measured by the *entire* set of data, was worse as well. The first of these is almost a contradiction which is why further work must be considered. It ought to be discovered that handling of subsets should *at the least* reduce the complexity of model construction. Regarding the performance, results for images prove otherwise as explained below.

7.6.3.2 Images Subsets

Earlier experiments suggested similar conclusions to the ones above. It appeared as if the approach resulted in slower progress and worse results, as deduced from the full model of what was said to be registered data.

Later on, and quite recently in fact, it was shown that values go lower (i.e. registration is improved) by using the subset approach. Many iteration though were required to show this. The subset-driven function caught up with its full-set equivalent and sank well below it. It was not clear though what had happened to the data, which could as well drift away. It is indeed possible that it got eroded more quickly for reasons that were earlier explained.

To summarise, subset-driven function are yet to be investigated, but they do not seem as powerful as adaptive precision in problems of shapes and images, for example. They often showed to be worse in terms of time, as well as worse in terms of performance.

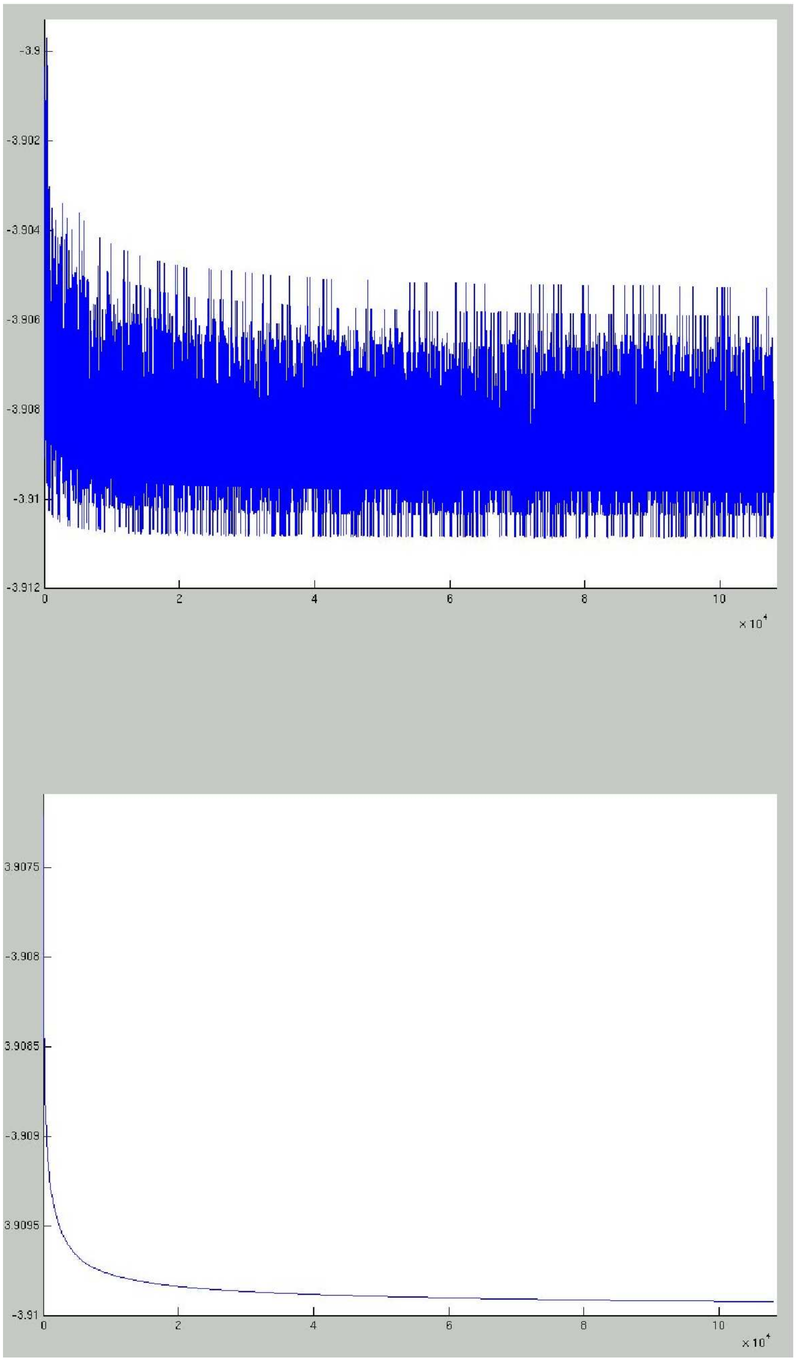

Figure 7.37: A comparison of performances in landmark selection. Shown above is an algorithm which is based on an entire set versus one which is based on a stochastic subset. The latter is quicker and it fluctuates due to the varying selection of a subset (3 shapes out of 10 in total).

7.7 Conclusions

Few conclusions were listed already, but what follows is a concise and generic summary.

Registration using statistical models is viable. It has clear drawbacks because it is (1) slow; (2) able to drift away (and destroy data) and (3) complex. When data registration is performed with the methods proposed, models of deformation are produced and correspondences identified.

Model-based algorithms result in appearance models whose determinant is by orders of magnitude lower than that which is measured at the start. The execution time they impose is inferior to MI, much as was expected all along, but might be superior to that of MSD.

The MDL term is improperly defined at present since it ignores the model discrepancies. When MDL is approximated by the determinant of the covariance matrix of the model, problems arise and registration (or landmark identification) is exacerbated past the stage when convergence should hold. Instead of convergence, a logarithmically decreasing 'tail' is observed and it indicates the need for objective functions being revised.

Chapter 8

ALGORITHMS

"Where all think alike, no one thinks very much."
– *Goethe Walter Lippmann.*

SOME methods and results have been discussed in the section about experiments. That section dealt with two distinct families of problems, but the main one was image registration and the relations it has to models of appearance. From this point onwards in this chapter, discussions will concentrate on that one later portion of the work. For the realisation of one successful approach and for completeness, this section will also explain the way in which registration is performed at present.

A later part of this chapter outlines the structure of the successful approach graphically. It is followed by the proposal of a two new approaches to be attempted in the near future if time permits. These more illustrative parts may be hard to follow, but references should clear up the way to full understanding. They should also serve as a starting point for greater exploration of this approach and fore-coming plans as they currently stand.

8.1 Registration Algorithm

This section presents the model-based objective function pseudo-code as it stands in June 2004. These explanations are intended to ease technical implementation being inter-communicated. Emphasise style is used to symbolise less significant I/O (input/output) steps which can be ignored. Many other steps are left out because they have little correlation to data registration itself.

The algorithm can be conceptually divided into three parts as follows:

8.1.1 Initialisation

- Generate data or retrieve it from file.
  - ✧ Apply data smoothing if required.
  - ✧ *Save data as images if requested by the user.*
- Choose data reference. By default, data instance which is closest to the mean is selected (see 7.4.4.5 on page 113).

8.1.2 Investigation and Preparation

- Find the target of registration where alignment is said to have been determined.
  - ✧ Align all data using a piece-wise linear warp if required.
  - ✧ Perturb the data if required, distancing it from the correct solution.
- *If necessary, save data to file.*
- Apply intensity offset if required. Intensity offset forces all peaks to equate in height.

8.1.3 Main Loop

- For all registration iterations:
  - ✧ Set the level of precision for the optimiser to reach. At present it increases as registration proceeds, but ideally it should increase when the advancements made are small.
  - ✧ For all data instances:
    - ❏ If the current data instance is not a reference:
      - ▷ Set up the positions of knot-points. Currently random placements with a sensible distribution are made.
      - ▷ Given the knot-points positions, apply warps to the current data instance and seek the warp parameters which minimise the cost $f(x)$, where x is the complexity of the model built from the entire set of data.
    - ❏ end if
  - ✧ end for
- end for
- *Statistics and registration logging take place.*

8.2 Algorithm Visualised

Figure 8.1 shows what is done in the algorithm above in a very loose and simplified form. A reference image, as shown at the top stays unaffected, while all other images are manipulated in the way described in Figure 8.2. These are used to construct a model which then infers a certain complexity measures, e.g. description length (as in Chapter 4). Based on that measures of complexity, subsequent warps are applied to the group of images.

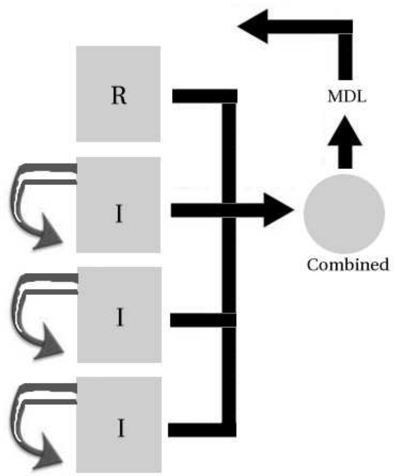

Figure 8.1: Schematic of the current registration algorithm. A reference image and the rest of the warped set form a combined model which is evaluated in an MDL-like manner.

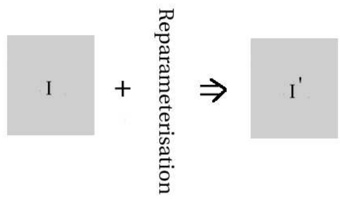

Figure 8.2: Current algorithm at a lower level. The idea of a reparameterisation is shown by emphasising that images are formed by aggregation of the previous image with some parameterisation.

For a full corresponding notation and further explanation on the figures, see the short presentation file at:

- http://www.danielsorogon.com/Webmaster/Research/Explanatory_Notes/2004/Group-wise_NRR_Strategy_and_Notation.pdf

It is certainly beyond the scope of this report and can be remitted. However, some of the figures should be thought-inspiring. A brief explanation of each has been included above for this reason.

8.3 New Algorithms

At the stage when results described in Chapter 7 resembled these which are described in Chapter 5 on MDL for shape models construction, it was decided to come up with new and more advanced ideas. In fact, parallel work by Twining on group-wise registration promoted collaboration and suggested that ideas should be exchanged in order to form a new, more powerful registration algorithm. As part of the discussions on issues of commonality and possibly duplicated effort, new schematics were drawn. This section is intended to explain them in some level of detail since they *might* at some point be implemented.

Two possible views on how the problem can be tackled are shown below. What is common to both is that they attempt to encapsulate the model discrepancy in one cunning way or another. They need to express a way in which discrepancies relate to the data and to the warps which are applied to that data.

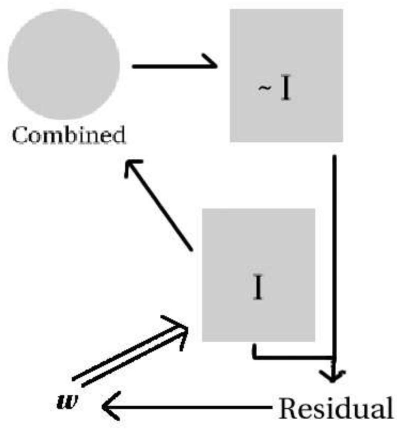

Figure 8.3: One possible proposal for the further development of the registration algorithm. The main idea to take is that residuals should drive warps that in turn affect the model.

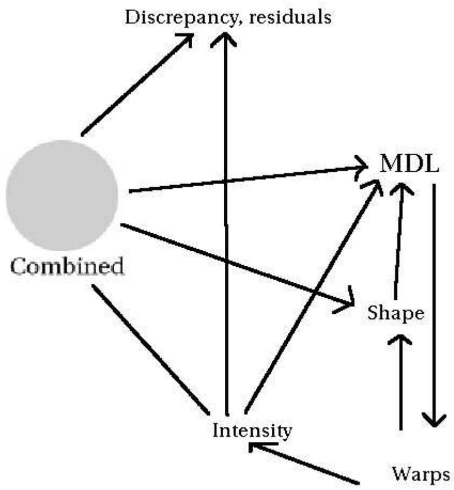

Figure 8.4: A second reasonable proposal for algorithm extension. MDL is the main driver of warps here.

For explanation of these, the aforementioned document needs to be looked at. These are rather long-winded to explain and the notation needs to be understood in advance. These notation is not final either and it is soon going to be changed as most recently agreed.

8.4 Extensions

Extensions were applied to the code which deals with MDL-based shape model construction (Chapter 5 on page 67).

As an example, below lies pseudo-code of what is referred to as adaptive precision. It simply selects a suitable tolerance for the optimiser. It does so in order to avoid excessive computations at early stages of the landmark selection procedure.

```
OUTPUT: precision_required
INPUTS: iterations_ratio, precision_automation_type
    % iterations_ratio is the ratio between the
    % current iteration and the total number
    % of iterations
switch precision_automation_type,
       case 'default'
            if (iterations_ratio < 0.1),
               return precision_required = 1e-1;
            elseif (iterations_ratio < 0.2),
               return precision_required = 1e-2;
            elseif (iterations_ratio < 0.3),
               return precision_required = 1e-3;
            ...
            else
               return precision_required = 1e-12;
            end
       case 'smart'
            % Look at the evaluation curve derivatives to
            % infer precision
end
```

8.5 Summary

The existing algorithm is well-behaved, but extensions to it, from which it can greatly benefit, are already foreseen. The problem is not thoroughly understood and some of the directions it may take depend on discussions and slow exploration of the effects that slight, or at times rather radical, changes make. Section 8.3 was aimed to show some of the more erratic ideas which either bear potential or possibly be disastrous. Nonetheless, these should motivate discussions and incompletely depict possible ways of going ahead.

Chapter 9

PROGRESS

"Procrastination is the thief of time."
– *Edward Young.*

9.1 General Progress

SEVERAL results have now been shown and discussed, but there are less technical matters that must be at least mentioned. The establishment of a naïve and simple model-based objective function was not the main implementation aspect that the author can take credit for. It was already in place at the beginning of the year, but it lacked many of the components that presently make it actually work and achieve good results which drive this research onwards. Results can now also be obtained rather quickly and flexibly since a front-end to the console-based functions was established. Technical details about the implementation can be found in the following sections on the WWW:

1. `http://www.danielsorogon.com/Webmaster/Research/NRR`
2. `http://www.danielsorogon.com/Webmaster/Research/Model_Based`
3. `http://www2.cs.man.ac.uk/~schestr0/Documentation`

(1) presents the project's technical aims, (2) explains the algorithms and (3) provides a very detailed overview on the application which is called AART (also see Figure 7.1 on page 86).

Experiments were often performed on several machines to collect different results of large registration processes simultaneously. Much of the computational power used resided in the Department of Computer Science where many of the strong and modern[1] computers laid idly.

There were some difficulties at getting persistent computational power at the very late stages of the year when long-standing clusters of computers were stored due to refurbishment work. However, that was the point when more theoretical ground was sought, as opposed to the analysis of large and cumbersome experiments. Such experiments were mainly used to provide apodictic proofs (as described in Chapter 7 on page 83). Also at that time, work on smaller experiments, which did not concern images, became much more appropriate.

9.2 Publicity

Some preparation and work on paper submission was considered at an early stage with the aim of obtaining feedback at the least. Furthermore, there was a slim chance of finding a place for the useful concepts and methods to be recognised and accepted as valid. There was not a high probability of acceptance because all results at the time were poor and

[1]Usually Pentium 4 processor with 256 Megabyte of random access memory (RAM). These were not computational servers in essence.

the text reflected on this. The results did not support the premise of the proposed methods of image registration. The work was still performed in 1-D over synthetic data, making its impact futile and the experiments uninteresting.

In Late February a paper was submitted to Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2004. It was cautiously recommended for acceptance by only one of the three reviewers. The feedback suggested that Guimond *et al.* [20] had performed similar experiments, but the paper suggested otherwise. It did not appear as if any groups used models and minimum description to guide registration – neither explicitly nor implicitly. Another main unavoidable flaw was the results being available for a 1-D case only; no practical medical results were displayed nor discussed (and illustration of either one is usually expected by the MICCAI community).

A poster presentation in the EPSRC summer school in Surrey (Figure 9.1) attracted a great deal of attention from both the organisers and the attendees. It appeared to be a close contender for the best poster prize.

9.3 Other Activities

9.3.1 Assorted Activities and Contributions

As this report can be perpetrated a show-case for progress and personal endeavour, this section provides an auxiliary note on progress. It is also an elaborative note on ways of research conduction from the point-of-view which excludes research. Below is a roughly random list of activities:

- EPSRC Summer School in Surrey attended in June 2004
- MIUA Summer School in Imperial College to be attended in September 2004

Figure 9.1: Down-sized poster.

- Attendance at plenary meetings of the IRC
- GSSEM PhD Workshop and other activities and lectures organised by the GSSEM.
- Contributions to the ISBE Internal Web site.
- MATLAB repository documentation and contributions to CVS.
- ISBE-All archiving system constructed.

Some of the later activities can also be seen as contributions. Nonetheless, these became merely by-products of documentation for personal use.

I have been working very long hours in the department, estimated at well over 60 hours a week. The large majority of the time was spent in the early morning and the weekends while very little time was spent at home.

9.3.2 Miscellaneous Meetings and Collaborations

A word on mutual and joint work has not been said yet. Many of the early experiments on shapes and landmark selection were performed in collaboration with Tomos Williams. At the later stages, this work was discussed with Rhodri Davies and some division of experiments and workload took place.

Collaboration with the structure and function (S&F) group meant that exchange of concepts with Carole Twining became mundane. Moreover, discussions on developments of the ideas and recent experiments (to be potentially work-inspiring) took place on a weekly basis.

For more accurate listing of various meetings including the Wednesday IRC meetings, also see:

- `http://www.danielsorogon.com/Webmaster/Research/Events/`
- `http://www.danielsorogon.com/Webmaster/Research/Meetings`

More compressed information and tables are appended on page 185.

Chapter 10

FUTURE WORK

"Reality is that which, when you stop believing in it, doesn't go away."
– *Philip K. Dick.*

10.1 Overview

THIS chapter draws a relatively precise picture of future work and possibly individual experiments that need to be performed in the course of the next year or two. It provides some rough guidelines for time, milestones and a loosely-defined line of operation that future work needs to take. In order to continue experimentation and collaborative chores in an organised and productive style, an intellectual plan and a detailed timescale need to be identified. Possibly, particular emphasis will be put on few specific research issues that need to be addressed. These issues will be the main lines along which research should move so that it is of *real practical use*.

There is no real micro-planning involved in this chapter, only a continued discussion and survey of assorted (and nonetheless related) items.

10.2 Aims

The aims of the project were formally specified in Form 2 (see Appendix C on page 176). The objectives set for the first year were slightly altered since the time of writing of the Literature Report[1]. While in a rough sense, all objectives set were eventually accomplished, more goals and intermediary experiments were proposed and later completed. It was known earlier in the year that aims listed in Form 2 cannot be *necessarily* the right way to go, but only a formal and inductive set of guidelines. This is clearly because of the dependency which one experiment has upon another. Also, development within the Structure and Function group required attention to different aspects of work, as soon as issues began to arise. With reference to aims listed in past documents, many of these were not at all definite. From a non-specific[2] perspective, it was expected that:

1. A full reproduction of past experiments should be trivial and possible extensions realised.
2. Development of existing code will commence to ultimately build genuine software.
3. Difficulties should be identified to avoid future impasse.
4. New practicable experiments should be agreed upon, performed and their results recorded.
5. Comparative figures will show the advancements of new methods.
6. Critical evaluation of existing work and proposition of new methods will hopefully emerge.

Dealing with each of the above in turn, all previous experiments, as implemented by Smith, can be performed within seconds. A software package which was constructed was in fact enabling such experiments to be

[1]The literature report is located at:
`http://www.danielsorogon.com/Webmaster/Research/Literature_Report`
[2]These requirements are intentionally very general. They can be applicable to most computer-scientific research.

analysed in more depth whenever one requires so (referring back to Section 7.3 on page 86).

Future difficulties were identified amongst the Structure and Function group and the project supervisor. These were frequently recorded and experiments that may suffer from such difficulties were avoided.

The new experiments were shown in previous chapters and comparative figures were an integral part of these. Not only was a benchmark for commonly-used methods established, but also a comparison – analytic and numeric – was made . Results amongst the different available alternatives justified the use of a model-based objective function as well.

Much more information about progress (and in finer level of detail) can be found in the weekly progress report and the sources that these reference. Experiment were very usefully put on an HTML-based database[3] and can be used as a supplementary resource for this document[4].

Proposition of new methods followed the analysis of different mutations of the model-based objective functions. However, nothing excitingly different or unexpected was discovered. It was merely the enhancement and modification of the basic model-based objective function that made it more powerful.

10.3 Impending Work

This section returns to the explanation of some of the work in progress. However, it concentrates on ways of moving forward, i.e. ways in which interrupted work is likely to develop. While it never will be obvious how

[3] Actually, these are hierarchical nested indices of experiments, sorted chronologically.

[4] These have faithfully and quickly served the need to produce figures for this report, as well as some of the results which were earlier outlined. In fact, the experiments pages can serve as a document of progress in their raw state. They might require some additional annotation to be legible.

experiments shall wind up, it is possible to at least propose them and list the expected outcomes, both pessimistically and optimistically. This is exactly what this section is set to achieve.

10.3.1 Further Registration Speed-up

There are clear intentions to reveal if any ways of speeding up the registration algorithm do exist. A little more work needs to be invested in reorganisation of the code and avoidance of unnecessary operations, especially within the MATLAB general optimiser. That general optimiser that to make senseless decisions at times[5]. Some reference to this kind of problem was covered in Section 7.6 on present work.

10.3.2 MDL and Models

At the time of writing, work is being put into the extension of automatic landmark selection for shapes. It is now realised that the model residuals must to be included in some form or another (e.g. description length) in the objective function for images and a good starting point is the simpler case which is , in fact, landmark selection[6]. When landmarks can be identified correctly *and* the objective function reaches stable convergence, application of the proven principles to images should resume. A detailed list of experiments and work to be done on images can be found in various personal memos and in the weekly progress reports (see previous sections for Web references).

As one example of the need for this issue to be resolved, see Figure 7.30 on page 116. The incomplete term for description length is described in [10].

[5]It was discovered, for instance, that values returned by the optimiser can be lower than the preliminary input values. That suggested that evaluations can be exacerbated as registration was performed and additional code was composed to resolve this.

[6]Most recently it turned out to have been complicated. Solution are agreed upon in present days and will soon be implemented and devised.

10.3.3 Automatic Landmark Selection

More considerable work has been put into shapes and the selection of landmarks to define point distribution models (PDM's)[7]. Work of this nature continues at the time of writing of this report, but clear goals are already known and their establishment is yet incomplete. It is inarguably desirable and even expected that improvement will be made not only to registration, but also to the correspondence problem in shapes. More importantly, revelations in one should affect work on the other. This way, neither of the two should ever lag behind the other in terms of performance or quality; implementation will not become overly forked either. Work on shapes can be sub-categorised as follows[8].

10.3.3.1 Subsets

The idea here is to speed up the algorithm by essentially pyramiding the whole set (see Figure 10.1) and building up towards a much quicker convergence.

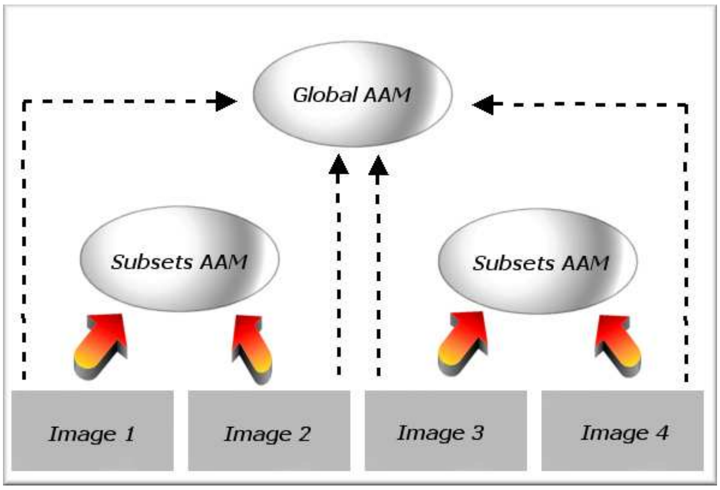

Figure 10.1: Illustration of the approach taken when registering using subsets.

[7]This has been work in progress since early June 2004.

[8]Note that where work has been done already, suitable explanations and examples were provided a previous chapter (Chapter 7) on experiments and results.

This nice hierarchy can allow larger sets to be dealt with, e.g. 50 or even hundreds, something which was thus far impractical. The figure shows how subsets are chosen in the context of image registration to create smaller AAM's. In practice the choice is stochastic although it is now realised that due to the internal intricacies of MATLAB, this arbitrariness results in reduced speed. By registering subsets, a globally good AAM can be constructed. Similar principles can be shown for shapes.

Instead of treating large sets and optimising over these, smaller sets can be handled, thereby reducing the burden of large Eigen analyses. Figures 10.3 and 10.2 illustrate that subsets appear to result in better and quicker descent[9]. The time required to optimise over subsets is surprisingly higher. This issue is a main one for future work.

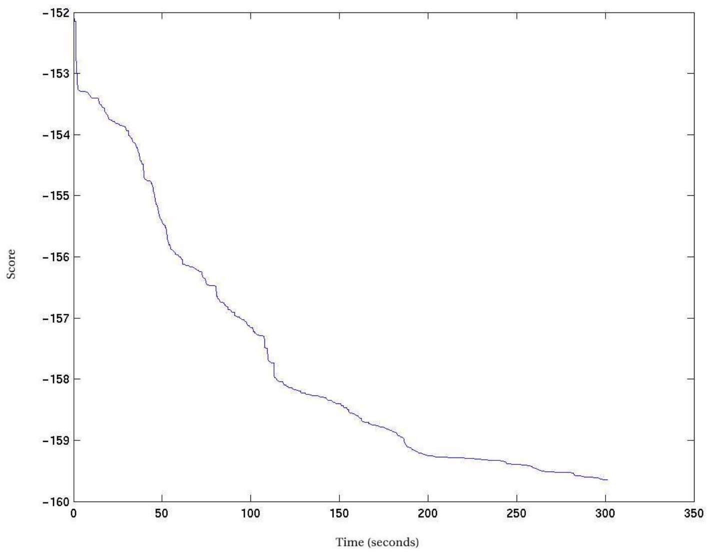

Figure 10.2: Images being registered according to the description length of the entire set of size 10. The X-axis indicates run-time time in seconds.

Figure 10.2 depicts one typical registration curve showing that the registration quality improves up to a point where betterment is low in extent.

[9]This excludes the start when subsets require time to stabilise by preliminary warps.

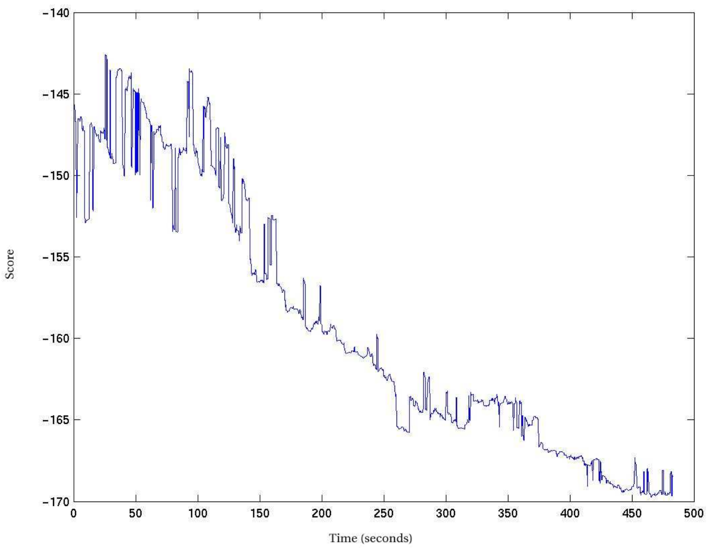

Figure 10.3: Images being registered according to the description length of random subsets of size 4. A choice of subset changes every 10 iterations. It can be seen that the score goes lower, but the time required is then greater.

In can be seen in Figure 10.3 that a subset-driven approach is slower though it is able to bring about some great improvements after an initial instability at the start. That slow start can be explained by pointing out that an insufficient number of different subset choices was cycled through. As a result, a rather localised optimisation is performed while the overall set benefits very little.

10.3.3.2 Varying Optimiser Tolerance

As part of speed-up through code modification, an adaptive precision approach and the like deserve to be looked into a little further. As the figures in the earlier mentioning of this issue show (e.g. Figure 7.35 on page 126), the rate of convergence is changed as the process goes on and so is the speed of the algorithm. There is more to be investigated to ensure the approach invariantly results in gains. It is also worthwhile to see if the choice of tolerance can be made more preferable, based on some empirical evidence. For instance, experiments with varying values for tolerance might be helpful.

10.3.3.3 Taboo Search (TS)

The issue was briefly investigated when better performance was sought for the landmark selection code. This appears to be a neglected method, but background reading investigated its potential for the clever selection of Cauchy's.

Cauchy's (essentially the means by which reparameterisation is guided) are chosen randomly and no sensible decision is made to avoid previous unsuccessful attempts to place a Cauchy. Taboo Search [19] is a technique of some rising interest in the 1990's. It retains a sparse data structure while optimising so that it can look up previous decisions and reach good solutions rather rapidly. It is similar to Simulated Annealing from a theoretic point-of-view.

10.3.3.4 Inference for Images

A few improvements to the registration algorithms were inspired by improvements antecedently made to the landmark selection algorithm. At times even some existing code was used to give ideas of how to improve registration, fix bugs and compensate for clutter.

One such example is the rollback for parameterisations which cause evaluation to exacerbate. The optimiser has a nature of returning an even greater value once bad warps have been picked. This needs to be fixed manually, by making some decisions outside the main optimisation routine.

10.3.3.5 Comparison of Optimisation Regimes

Different optimisation regimes were investigated with Tomos Williams, e.g. by changing optimiser parameters for the built in MATLAB *fmin()* function. More such comparisons can take place to (at least) regain confidence in the optimiser which is currently employed.

10.3.3.6 Unifying Shape and Image Code

This statement is concerned with the option of taking the two algorithms (one from AART and one from the code which optimises shapes) and trying to make them more flexible and powerful, jointly. At least several of their common functions can be shared by both, in which case one update affects all.

Perhaps some forging of the two would make sense and result in greater productivity and functionality. This latter suggestion is more far-fetched though.

10.3.4 Extension to 2-D

Since some of the gathered statistics indicate that the program achieved what it had been set to achieve, there are intentions of applying the same concepts to 2-D and data in the foreseeable future. The principles remain unchanged and the only required extension is that of CPS to a higher-dimensional space – something which has been developed already.

This extension step is expected to be trivial as CPS warps in 2-D have already been dealt with by Marsland and necessary 2-D data is available in the Division. Synthetic 2-D data can be generated as well if necessary and code exists for doing so (see Figure 7.11 on page 95). Once generation of data becomes possible and shapes for which the estimated solution is known can be created, e.g. triangles in 2-D space, then well-controlled 2-D tests can be performed under AART.

The only envisioned hindrances will then be the speed of execution and the diligent selection of knot-points for transformation. Since there is yet some lacking understand of the problem in 1-D, time is needed to improve the algorithm. Its requirements need to have a more responsive time-span.

10.3.5 Benchmarks using Flexible Platforms

When a sensible registration algorithm is available to be used with brain data, it can then be compared to other methods developed within the IRC. Comparative tests are performed using a regular, annotated and standardised data. Software of Crum *et al.* (confer the paper to appear in MICCAI 2004) is able to carry out such comparisons. It relies on the database from Boston which comprises 8 brain volumes with ground-truth landmarks (annotated by professional radiologists). Validation [50] of results with respect to other existing and to-be-established inter-operable software can be considered as well.

10.3.6 Application to 3-D Data

The possibility of 3-D registration is still on hold, awaiting the point where it becomes practical. If the algorithm is successful and fast enough, 3-D extensions remain a valid possibility. This will not, however, take place in the near future.

10.3.7 Creating Atlases of Deformation for Different Groups

Having got some methodology which derives average data with description of valid deformations, one can study the deformation of brains within

Experiment(s)	**Date to comply with (lenient)**
Speed-up	Perpetual
MDL in models	October 2004
Subsets versus entire set	September 2004
Optimiser investigation	October 2004
Comparison of optimisation regimes	November 2004
Extension to 2-D	January 2005
Application to 3-D	Indefinite
3-D benchmarking	Indefinite
Group-specific atlases	Hopefully 2005
Automatic appearance models construction	Indefinite

Table 10.1: Milestones for future experiments.

a schizophrenic group, for example. It is then possible to perform some classification tasks using model fitting.

10.3.8 Refining Appearance Models Construction

For reasons which were openly explained in earlier parts of this report, automating the selection of image correspondences is essential. This endemic problem in modelling can hopefully be solved at last. Hopefully, this refinement means *automation*.

10.4 Future Milestones

Milestones for future work on experiments seems a rather good idea. The table below does not specify any stringent requirements and does not need to be obeyed, yet it can be contributive as future reference.

Chapter 11

SUMMARY

"God does not play dice."
– *Albert Einstein.*

11.1 Brief Overview

SOME of the main relevant concepts and techniques in existence have been explained and numerous examples have been given, although their number was restricted to allow for a broader survey. Most such techniques directly relate to the problems which need to be tackled and their utilisation in past and present has been thoroughly explained. As future experimentation is expected to rely on recent research and is most likely to involve similar ideas, algorithms and paradigms, continuous reading of technical reports, alongside reproduction of the experiments, will be an essential portion of the research approached.

The problems with current techniques were found to vary from the interest in efficiency to possible flaws and gaps, a part of which being driven by insufficient correctness arguments and lacking ground-truth. Without a doubt, there are phases in current research where heuristics take over at the expense of valid implementation that can be reasoned about straight-forwardly. Many areas are still controversial and common assent is missing and might never be reached. As instances for the aforesaid claims, a group-wise brain analysis algorithm devised a wide range of domain-specific facts (see B.1.4 on page 169). Moreover, a major undecided issue is the most advantageous warp type and its corresponding complexity that strives to give ideal results per permanent time unit.

This project, much like other projects in this area, attempts to find some answers to the questions raised and extricate us from uncertainties and disagreements. It seeks a theoretical proof which can be backed by empirical evidence. Itprogressively implements a convenient tool for quickly evaluating and profiling different ideas and approaches. Whether it will be successful in the sense that it should provide inarguable answers and discover new techniques that are ingenuous, it is yet unknown. This project should draw conclusions regarding performance, feasibility and validate or invalidate some results of previous work. Preferably it should surpass previous work that it has built upon. No results will be taken for granted and a *critical* approach will be dispensed at all times.

Within the second year of the project, it is hoped that an implementation of a better warp and model test-bed will be available. It should achieve dense correspondence across a set of synthetic (and hopefully medical) images in 2-D and 3-D. Software should be capable of looking into the behaviour of warps regardless of the nature and scale of the data. It must also respond within a suitable time period, although the notion of "suitable time period" is loosely-defined. There is a growing belief that such tool can be of great interest to these who use and facilitate active appearance models.

Nonetheless, there is a real snag as the data under consideration should still be modified to approve the successful application of the techniques to data of higher dimensionality. The run-time and the results that can be retrieved within a limited time-frame is then the main impediment.

Although a partial time-line was specified for this project and its intermediate objectives, it is not yet clear where the project will turn and what it will eventually accomplish with success. It is known, however, what should *ideally* be accomplished. Semi-annual reports and documents will clarify the emerging plans and intentions as they become more concrete.

11.2 Conclusions

The appearance models currently used are not fast and cannot be argued to be ideal in any sense. A solution to these flaws would be highly desirable. While it is not clear how to optimise models or how to evaluate a model [29], there are measurable means for arguing about the quality of these models comparatively. Amongst the main problems that are ordinarily seen in appearance models is their inferior performance, although this depends on the functionality required. Automation could have a significant contribution to such a model, but correspondence needs to be achieved first. Luckily, issues of correspondence have been investigated largely in the past decade so this should not be a peril.

What is worth investigating even further is the ability of warps to improve models and at the same time encapsulate several analysis steps together. What is even more reassuring is the proven ability of models to improve warps selection and improve on existing group-wise registration methods. This improvement relies on the fact that a large-scale collective analysis replaces the weaker, yet computationally inexpensive, pair-wise scope.

11.3 Contributions

Taking a more positive stance, it has been illustrated not only that group-wise registration based on appearance model is possible, but also that it surpasses registration methods that are based purely on a reference image, as judged by the corresponding appearance model.

The current algorithms are being interpreted rather than compiled and no multi-scale approach is yet in use. The extension of the algorithms to 2-D and 3-D would require a long time to run, but remains practical. Compiled implementations might be available within months as well as heuristic optimisations that will make run-time more competitive with that of pair-wise approaches. Furthermore, the results have been shown to be better in a global sense and are not dependent on just one individual image.

Contributions of the work can be subdivided into three aspects:

1. It provides a benchmark environment and results for many methods, including several new ones.
2. Unprecedented model-based group-wise registration is introduced.
3. Automatic construction of increasingly better appearance model becomes practicable. Correspondences are obtained using techniques borrowed from image registration.

11.4 Final Discussion

As described in Chapter 10, the project is now expected to continue along the lines of implementation with some reliance on the work and results produced in year one. If difficulties often recur, there are various measures that can be taken to ensure productive alternatives are chosen. The next goal is to obtain dense correspondence across 3-D biomedical data using automatic self-instructing algorithms. As the project is intended to explore a scarcely known field, caution will be taken when time is spent without much potential on the horizon.

11.5 Vision

Given a collection of images describing the same objects, one should be able to build a good model autonomously. Given sets of images from normals and patients with a common pathology, different atlases will be constructed and diagnosis become solvable by computers.

"Cures were developed for which there were no known diseases."
– *Ronald Reagan,*
in memory of...

Appendices Index

"Nothing is so simple that it cannot be misunderstood"
– *Jr. Teague.*

Appendix A

Elaboration on Appearance Models

This short section describes some of the applications and extensions of shape and appearance models. These are all of little relevance, if any, to the project under consideration.

Real-time Active Appearance Models

Some machines are able to deal with small-sized fitting in real-time [2, 23]. It is possible to track faces in a video (frame rate should then typically be 24 frames/second and 15 at the minimum), but the resolution catered for is often relatively low (less than 100x100 pixel). Applications that respond so quickly were made far more practical owing to a multi-resolution (multi-scale) approach (see A.1 below). In order to decrease the total run-time, varying increasing image resolutions become available for selection at each search iteration. Finer resolution images are usually

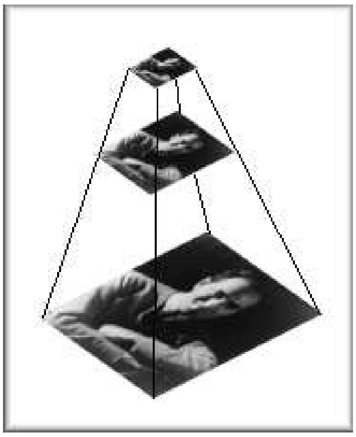

Figure A.1: A multi-resolution approach illustrated. Coarser representations are shown at the top levels and the original image lies at the bottom.

used at the later stages of the search, whereas low-resolution (coarse) ones at the very start. Since the similarity between the model and the target is poor ab ovo, the resolution (and hence the scale of the objects) will have little effect on the fitting. Some visual examples of AAM search are shown in [7].

An additional advantage that a multi-resolution approach offer is its notable improvement of structures. It even allows fitting to be more robust to large displacements from the target. This is due to its treatment of a large image as if it was a smaller one – one in which structures are represented by a smaller number of pixels.

Other Applications

Very common uses of AAM's are for medical image analysis and face recognition[1]. Active appearance models possess traits that make them robust and effective in the biological domain, whereas industrial inspection, for example, presents some inherently different problems. These problems are often solved more rapidly by other approaches that are

[1]Much of the popularity of this method has been imputed to face recognition tasks.

based on lower-level knowledge about the image contents. Since a broad range of tasks are performed in industrial inspection, however, it is also valid to assume that the suitability of top-down approach is irrespective of the problem. Nonetheless, there has been some successful application of these methods to the analysis and segmentation in printed boards. Kestra have been doing some successful work in this domain, though it is not a main-stream application of statistical models as yet.

In order to visualise biological shapes and full appearances, a model which handles anatomical variability and change needs to be used. It must account for natural or pathological changes such as the change in form of organs (atlases for different pathologies are suggested in [55]). Greater variability can be encountered when aligned images are obtained from different subjects in a population (inter-subject), the same subject at different time instances (or different sites) or when having to account for movement such as the that which occurs due to respiration, the cardiac cycle and so forth. A separate case to consider is multi-modal imaging which will not be explained in any detail although it is a quickly-developing area.

For an excellent overview on many of the different image analysis techniques, the book from Sonka *et al.* [51] is a valuable source. For a good review of model-based image analysis, papers from Cootes *et al.* are an even better source. Related and similar insights can be derived from work partially based on concepts such as snakes, bending-energy and active contours. Such conceptual approaches, are closely-related to active shape models. Sonka *et al.* discuss all of these in depth.

PCA Alternative

A recently published technique is said to be capable of finding good dense correspondence. It is described by Tony Jebara. Images are said to be better represented as sets of vectors for this specific purpose, as opposed

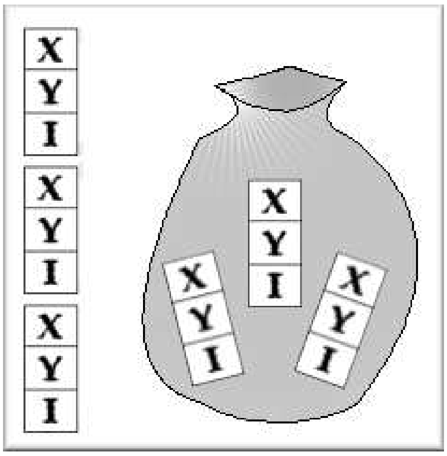

Figure A.2: Bag of pixels illustrated versus one conventional approach.

to vectorisation where fixed ordering is imposed by *concatenation* of the vectors. Pixels are represented by the common (X, Y, I) tuple and the ordering of these tuples is arbitrary (they are said to analogically be placed in a bag so an alternative notion would be *sets of pixel*). Ways exist in which good configurations for ordering these pixels can be found. This implies that vectorisation of the pixels is not the sole option for effective image representation. As the process of pixel ordering takes place, dimensionality reduction is indirectly performed which transforms the image into a volumetrically minimal subspace and this reduction outperforms principal component analysis by orders of magnitude. This is one of the points that make this idea so appealing, but it is still extremely slow[2].

Figure A.2 pictures the difference between a common approach of pixel ordering versus the alternative bag of pixels.

[2]The algorithms currently used for demonstration purposes take 3 days to run, but substantial speed-up is expected soon.

Appendix B

Project in Detail

THIS appendix describes in some finer detail the past work. Such work led to and supported current research which this report describes.

B.1 Overview

Active appearance models and non-rigid registration were attempted to be unified for a period of time before the author's undertaking of the project. The following is an overview on the work of Smith from which current developments stemmed.

B.1.1 The Data

The work of Katherine Smith investigated warps on a simple bump-like curve (see Figure B.1).

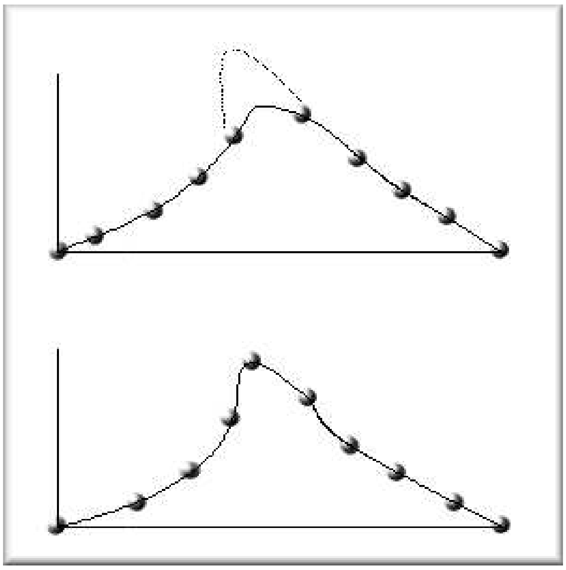

Figure B.1: Reparameterisation along the curve.

A real set of data in a clearer, more static form is shown in Figure B.2. Smith attempted to register the data using the principle of model complexity minimisation where the model is constructed from the entire set of simple 1-D bumps.

B.1.2 Description of the Approach

The first step taken by the application was the generation of some random bumps simpler than the ones described in Subsection 7.4.1.3. These bumps varied in their height and width; the step size of the bump (the steep ends of the flat pinnacle) was fixed, i.e. the bump was initially flat at the top.

Although the property of height was not intended to be ignored during registration, it was expected that it would remain unchanged due to CPS being perfectly diffeomorphic[1]. The bumps were all symmetric and the height was one of {*hi, low*} where $lo = 0$ and $0.7 < hi < 1$. The data was

[1]While the bump may have its form tweaked and manipulated, its highest peak should be preserved although it may move leftward or rightward.

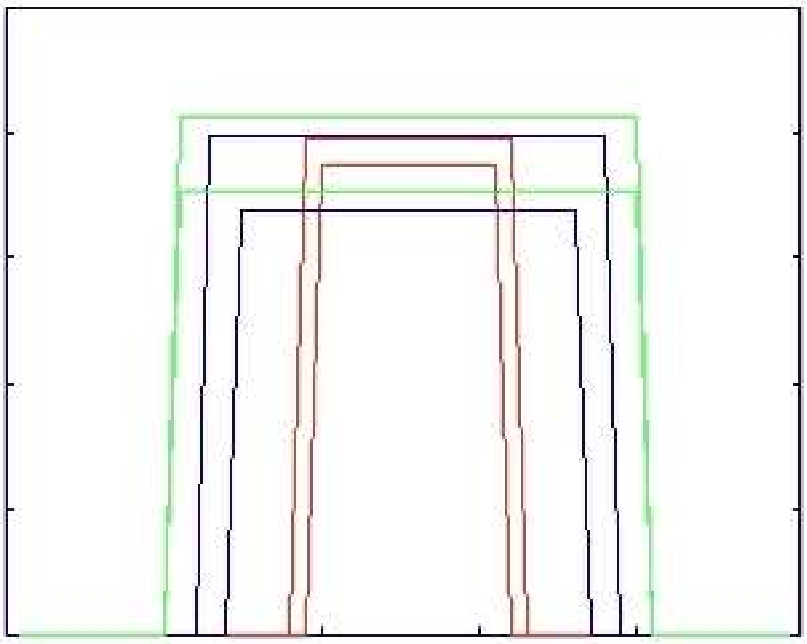

Figure B.2: An actual set of bump data. Different instances are indicated by distinct colours (or shades).

therefore far simpler than any 1-D data which is not constrained in any way. The height of the bump and the position at which the bump goes high could conjointly define that bump so two real numbers (a tuple) at the minimum would suffice to reconstruct each bump.

As images were being warped, the form of the bump quickly changed to give a smoother curve with more continuous derivatives. This of course depended on the type of warp which had been applied to the bump. At each iteration, new similarity with respect to some reference image or similarity with reference to the whole set of images was obtained using warps and measured using the methods outlined in Section 5.3 on page 68.

The similarity measures used in these experiments to evaluate similarity were mean-squared-difference (MSD) and mutual information (MI). The latter was more computationally expensive so although it gave better results, it needed to be used with caution. Likewise, the type of warp applied was often, but not alway,s a simple one which is controlled by a single allocated control point. In some cases, many control points were assembled to form an expensive warp of increased complexity. The choice of these points was often decided to be random as a successful rational choice would have required much more speed, consequently slowing down the whole process.

As explained to some extent beforehand in 3.2.1.3 on page 55, reparameterisation was used to perform points placements in the image of the bump. These points did not directly express the form of the bump, but rather controlled the *warps* that affected the bump point coordinates. Initially, the curve to be reparameterised was an ordinary linear function stretching from the origin to a point (n, n) where n is the number that is chosen to be the image width (the only dimension of the single-dimensional data). Points were later chosen according to the change imposed on the curve due to warping.

The experimentation Smith carried out allowed for many combinations of different options to be set, applied and appraised comparatively. The estimates of the "goodness" of warps were calculated using the creation of an appearance model from the group of images at present state, making this a group-wise optimisation methodology.

The images after warping had been applied were treated as training data for the creation of an appearance model. PCA reduced the complexity of that model as required. The compactness of the model which could be derived from the the sum of variances or the determinant of the covariance matrix[2] was then scoring the choice of warps after they had been applied. In this way, a better choice of warps could be made so that bad ones quickly get discarded and the state of all affected images reverted.

[2]This will indicate the *volume* of the model's scatter in space. The more compact a model appears, the lower this volume. More importantly, it is an approximation to the MDL objective function.

B.1.3 Synopsis

As the above descriptions tacitly suggest, this work was able to show how statistical models go hand-by-hand with non-rigid registration. In this case, they simple *evaluated* the (non-rigid) registration process and distinguished between the many alternatives offered by different families of warps, similarity measures and so forth. Needless to say, the run-time became a real difficulty when ill-chosen strategies were attempted. Smith took this into consideration in the final evaluation and comparison of all different experiments.

B.1.4 Concurrent Advancements

The work of Marsland, Twining and Taylor [38] went a step ahead and investigated a full 2-D model[3]. However, it concentrated on just a simple contour (defined by 12 control points) of the skull shape as pictured from an overhead perspective. Figure B.3 shows that warps can have an effect on the *whole* shape, but still lack some control over local structures such as the ventricles. Varying scale can solve problems like this and make the global non-rigid registration approach very robust. While this work produced elegant results, it did not explore many varying options as in the case of Smith.

B.1.5 Drawbacks

Drawbacks and gaps do exist and controversial implementation decisions are listed later in this appendix. There is much work to be done in order to find out which the better performing approaches are and the experiments and applications described above provide a substantial starting-point.

[3] It also used proper MDL terms rather than an approximation as Smith did.

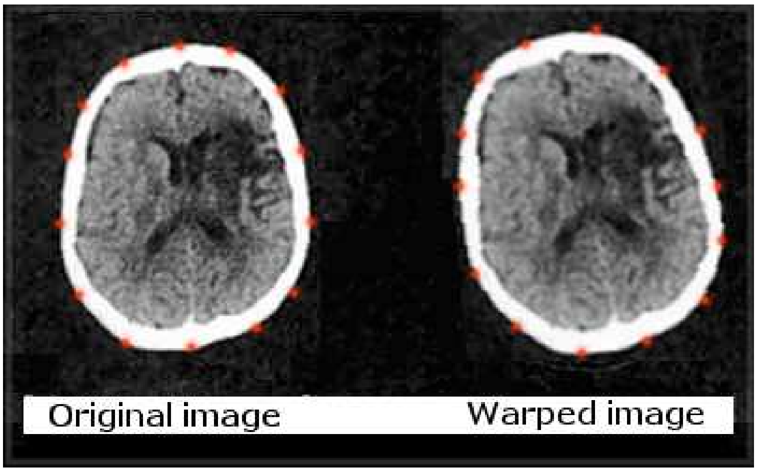

Figure B.3: Brain image warping. Points on the skull depict knot-points for the splines.

B.2 Alternatives

The main alternatives to non-rigid registration are rigid and affine equivalents, but these are merely impractical. In most real-world applications such as registration of brain-slice images, there is a very slim chance of getting satisfactory alignment of structures while preserving some continuity unless non-rigid transformations are applied. One may argue that affine registration should suffice, but what if parts of the brain expand beyond proportion? It therefore appears as if, from a registration point-of-view, no obvious alternatives are yet known. The ones mentioned above give the best performance yet and comparison with the closely-related active appearance models suggests that flexible deformation is mandatory, especially for bio-medical data. Nonetheless, one could argue that there should be more than just a single alternative to be looked at and many different aspects call for attention as the earlier parts explain. Here is a short summary that may help guide future endeavours[4]:

1. **Speed-up:** The methods operate very slowly for most globally-driven

[4]Although none of the points is far-fetched, not a single one of them proposes an unfamiliar approach; and yet, an open mind is the key to advancements.

approaches. A solution to this is desirable because not only would it stimulate more experiments and experiment feedback, but it would also make these methods usable and marketable.

2. **Data extension:** The simple existing bump which is generated in MATLAB needs to be extended, possibly by conversion to a smoother bump as the one described in the research of Davies and Taylor.

3. **Lambda coefficient**: In practice, when constructing an appearance model for registration's sake, an additional weight is assigned to one of two related components. The first component is associated with the reparameterisation curve and the second corresponds to data values, i.e. intensities. This weighting term, denoted by Lambda (symbolically λ) in the objective function, essentially weighs appearance against shape and its value is subjective and dependent upon the problem. Experiments can find (and have found before in Smith's work) alternative solutions or better assignments for lambda.

4. **Automation:** It would be desirable to create a (compilable) system that copes with the full cycle of analysis without outside intervention and without any pre-existent data annotation. This relates to the strands of artificial intelligence and autonomous systems.

5. **Generalisation:** Many *ad-hoc* algorithms are currently used for group-wise registration. An more impressive system would deal with arbitrary data without compromise to the quality of the results.

B.3 Relevance

Along the lines of past research, the possible developments and gains all appear to offer various advantages to this system of registration and analysis. Unlike many other such projects, this one is open-ended and is dependent on the outcomes, discussions and discoveries communicated or published in conjunction to one another. As a group effort is expected, a high level of interaction and communication will be involved. This has clearly been the case to date. It is important to note that this research is also of relevant to other on-going research that takes place locally.

B.4 Significance

If this research reaches and obtains its goals, more theoretical and practical grounds will be available for future applications of non-rigid registration in active appearance models and vice versa. By making real use of the intensity information that is available in appearance models, an exceptional practical strength can be exploited. Full image synthesis is an application where no other model type can yet be seen as a substitute. With more automation in place, higher accuracy, compatibility with other technologies, awareness and the like, ultimately, ubiquitous use of the technique can be (optimistically) anticipated.

It is worth pointing out that availability of active appearance models to individuals who deal with registration of images would take this technology one step ahead. This will introduce more concepts, metrics and studies which increase their functionality and flexibility. Reciprocally, previously "foreign" techniques can extend the functionality of active appearance (and maybe shape) models once they are put in the hands of groups with difference background and expertise.

As an instance for the first of the two contributions above, a radiologist could very comfortably view a highlighted model of an organ that is deformed in a natural and sound manner. Results and analysis can be managed rather quickly and neatly as automation and synthesis generation should be made operable and even interactive[5].

[5] A reasonable response time depends on the purpose of the system, the level of detail, etc.

B.5 Developments

This section summarises some of the previous preliminary developments; these are developments which were made before recent experiments to be listed in Chapter 7 on page 83 of this report.

The simple data used by Smith was proving slightly too cumbersome for responsive experimentation on a relatively strong machine (1.8 GHz, 512MB RAM), especially owing to the complex algorithms devised for group-wise registration. It was at that point advisable that evaluation via profiling toolkits was made to hasten the process as much as possible. Alternatively, coding of the algorithm in a compiled language as C++ was seriously looked at as a possibility. The complexity of the departmental VXL library was believed to make a step as such less than desirable and no such development has ever been made thus far.

Once speed-up had been taken care of or when it was at least known that a nearly flawless well-performing piece of software was at the user's disposal (and one which was under control), the simple 1-D data could see the addition of a few additional characteristics. That new composite[6] data had to retain some good commonality and similarity across the set of images and it could not be overly more complex and unpredictable in comparison with a simple bump. A double-humped curve, a round smooth line or even a contour of a a profile of a face could be sensible and more challenging choices[7]. In any case, whichever synthesis of data was eventually selected and experimented with, the choice of control points for the warping then became a more crucial issue[8]. A more localised control via warps then turned into a mandatory one because several separate structures exist in the data.

[6]It will be prematurely assumed that the new *synthetic* data type possesses several distinct morphological attributes.

[7]Some discussions also suggested that data should be similar to that used in Davies' thesis, i.e. a bump on top of a rectangular brick. Analysis based on current understanding then becomes viable too.

[8]Warps placement truly seemed tactless and poor at the time, but this needed to be confirmed by actual evidence.

The experiments of Marsland, Twining and Taylor had already shown the realistic application of warps to a medium-resolution two-dimensional data. Nonetheless, it was vital to point out that an elliptical shape was dealt with and *a priori* knowledge of the problem was used to increase the speed of the group-wise registration process. Control points that characterise the warps were initially placed on a circle whose centre was the image centre and radius corresponded to the typical position of the skull in standardised imaging. If the problem involved point selection for, let us say, knee cartilage and no knowledge about the object was available in advance, the results would have then taken far more than 10 hours to obtain (as was the case for the 12 points distribution around the skull's exterior). Edge detection is quite useful in an application of this kind. It was highly useful in the case of the skull data, but finding edges that form a circle (confer Hough transform) as in a skull is somewhat of a simplified problem. Subsequent developments should aim to address many issues exhibiting resemblance to aforementioned ones.

B.6 Challenging Issues

There are several issues that cannot be ignored and should therefore be systematically listed. Here is a brief unordered list of issues that appear to induce uncertainties and confusion:

1. A sequential series of warps is often an expensive step that results in poor productivity.
2. There is a wide range of warps and there is no consent on which the most effective ones are.
3. The existing algorithms are very slow and require long periods of waiting time until constructive feedback is received.
4. An existing system that sometimes struggles with one-dimensional data is required to be extended to 2-D and preferably 3-D too.

5. The data handled by the existing approach tends to be excessively simple. Feasibility of such an approach in complex applications is still unknown.

6. Medical imaging requires high fidelity and reliability. Unfortunately, the output from inner-body imaging has a significantly low SNR[9]; this conflicts with the fidelity requirement. The accuracy of this approach, e.g. the establishment of correspondence, is then unsatisfactory for some of the more critical procedures. This is exactly why true group-wise registration is important.

7. There is often little knowledge about the structures in an image and random warps are then the only reasonable choice, resulting in a slow process. Solutions might come up in the form of bottom-up analysis of an unknown image.

It is expected that many of the issues above will wind up being taken into consideration. They may affect the feasibility of the project, lead to failures or halt the pursuit for the original aims[10].

[9] The signal-to-noise ratio in medical images can be lower by orders of magnitude in comparison with visual images.

[10] Chapter 6 was bound to take a pessimistic point-of-view to describe worst-case scenarios. A more optimistic contemplation would have discussed the obtainable goals and the factors that make these goals arduous if not impossible to reach.

Appendix C

YEAR ONE PROGRESS

Form 2

Date of meeting: December 15th, 2003.

RESEARCH AIMS & OBJECTIVES

Synopsis of Research Project and overall aims:

- Fully automate obtaining a set of dense correspondences across a set of 3D medical images as a basis for building statistical models of shape and appearance.
- Develop a new approach with a rigorous theoretical basis and compare its performance with existing approaches to the problem.

- Apply the method(s) developed to demonstrate changes in morphology due to disease (or other causes) in a large dataset (eg brain, knee etc).

Objectives of research project for first year (full-time Students) or first two years (part-time Students), including literature searching:

- Establish benchmark results for correspondences obtained using existing non-rigid registration algorithms.
- Develop an in-depth understanding of the literature on: non-rigid registration, active shape/appearance models, and minimum description length methods (generally and as applied to shape correspondence).
- Develop a general understanding of current methods and problems in computer vision, with particular emphasis on medical image analysis.
- Carry out initial experiments using synthetic data to gain an insight into the problem of automatic image correspondence and an understanding of the key problem areas.
- Obtain and analyse initial results using both synthetic and real data (possibly only 2-D).
- Develop a plan for future work, based on the experience of the first year.

Key objectives for first 3 months:

- Complete machine learning module successfully.

- Establish a pattern of background reading.
- Undertake a detailed review of the literature in non-rigid registration, active shape/appearance models, and minimum description length methods (generally and as applied to shape correspondence).
- Gain good familiarity with using MATLAB to run computer vision experiments and to analyse results.
- Establish simple 1-D model building framework using MATLAB software from Kate Smith.
- Plan presentation for student seminar.

Key objectives for first year:

- See objectives for first year above.

VARIOUS COURSES

Course/Seminar Title	**Dates (if known)**
Introductory Course	22/09/03 - 26/09/03
Library Visits	ISBE Research Library: perpetual
Regulatory Core Courses	1/10/03 - 1/06/03
Computing Skills and Statistics	N/A
1st Year Workshop	N/A
Health and Safety Training (Compulsory)	N/A

Table C.1: The courses taken at the beginning of the academic year.

Deadlines

See Table C.2 for deadlines. The one shown further below (Table C.3) is for personal guidance only.

	Deadline date
Existing code mastered	December 20th, 2003
Submission of **Literature Report**	December 22nd, 2003
Extension to code completed	January 5th, 2004
Presentation	January - February 2004
Resolving Project Plan	(End of) January 2004
Progress Report Submission	March 24th, 2004
First implementation working	(Late) April 2004
Implementation entirely documented	July 2004
Experiments performed	August 2004
Continuation Report Viva completed	September 1st, 2004

Table C.2: Stricter deadlines.

Milestones

In line with Form 2 (page 176), but from a broader, more formal scope, here are some very rough estimates of expected milestones. The following chart summarises some of the milestones expected and the corresponding deadline dates.

	Recommended completion date
Literature Report Submission	December 15th, 2003
Literature Report Meeting	December 22nd, 2003
Extension to code completed	December 27th, 2003
Presentation	February 2004
Resolving Project Plan	January 2004
Progress Report Submission	March 1st, 2004
First implementation working	April 2004
Implementation entirely documented	July 2004
Experiments performed	August 2004
Continuation Report Viva completed	August 25th, 2004

Table C.3: General time guidelines.

The Gantt chart below attempts to assure compliance with the deadlines and guarantee that progress will be made as anticipated.

	Dec/03 **May/04**	**Jan/04** **Jun/04**	**Feb/04** **Jul/04**	**Mar/04** **Aug/04**	**Apr/04** **Sep/04**
Literature Report	√ 				
Presentation	√ 	√ 	√ 		
Work on existing code	√ 	√ 	√ 		
New implementation	 √	 √	√ √	√ 	√
Documentation	 √	 √	 √	 √	√
Experiments	 √	 √	 √	 √	
Continuation Report Viva			 √	 √	 √

Table C.4: Progress Gantt chart.

Work division that is project-specific and a better summary of the *technical* aspects will be entirely left out. More accurate aims were formulated in Form 2 for completeness.

Contingencies

As some feasibility considerations are yet to be resolved, it is vital that alternative directions for this research are realised and suggested. One facet of this issue is concerned with times at which *evaluation* of progress, development and achievement ought to take place and quality reviewed, apart from the formal evaluation in April 2004. By recognising dead-ends at the earlier stages of work, wasted effort can essentially be avoided. There are several types of problems that can come up:

1. A field is yet too poorly understood and there is a lack of basic knowledge to rely upon.
2. Effort is already invested in the exact same field or problems that the project poses are found to be resolved already.
3. The code dealt with is too hard to cope with.
4. Given algorithms or conventional methods are too slow to work with productively[1].
5. Alternative solutions with greater potential are identified, thereupon requesting all attention to be diverted to them exclusively.
6. Experiments fail to produce the results expected or hoped for.
7. Progress is held back by time restrictions.

Obstructions which are prone to happen more frequently would be 4, 5 and even quite frustratingly 6.

[1]Frequently it appears to be the case that in order to get reasonable results, high computational power is mandatory. In the absence of this power, experiments might fail or become impractical.

Appendix D

AART APPLICATION DOCUMENTATION

SOME comprehensive and detailed documentation is available on-line at:

`http://www2.cs.man.ac.uk/~schestr0/Documentation/`

In D.1 below lies an image which depicts the structure of the application. It should hopefully provide a reflection of the healthy nature of existing dependencies.

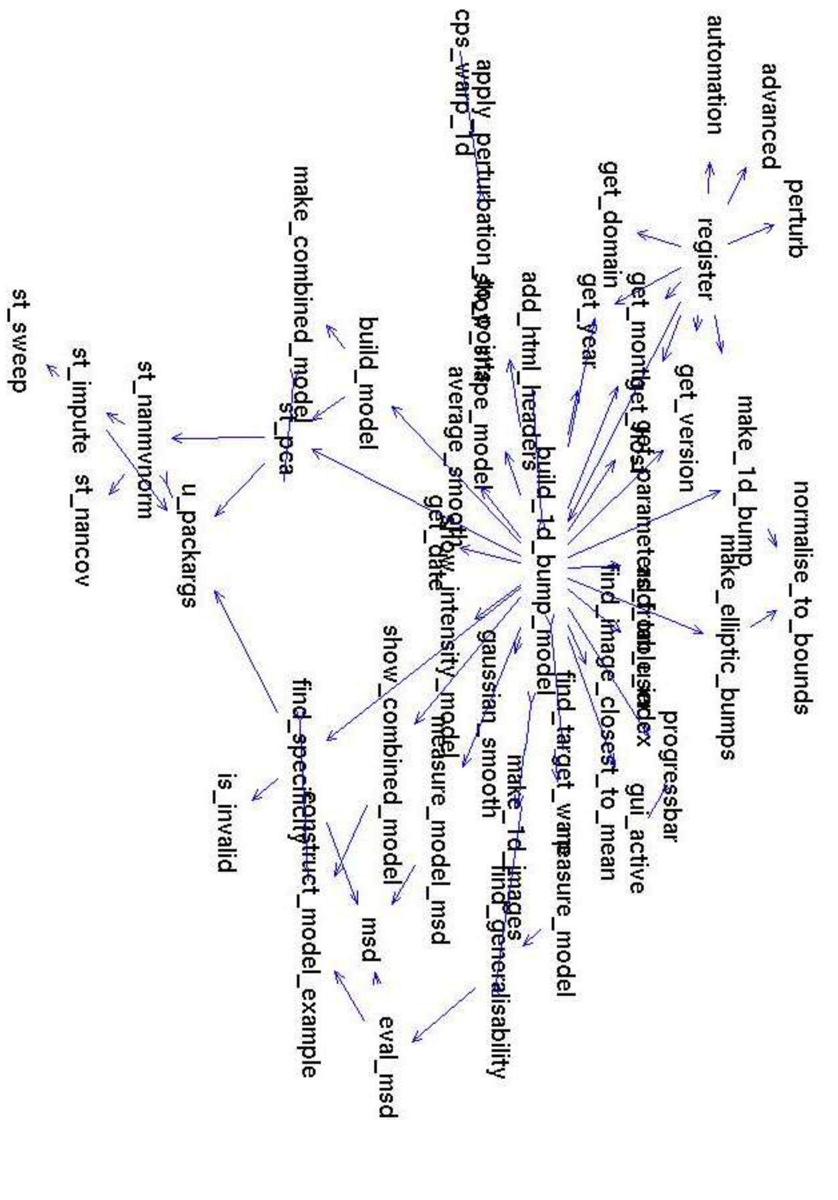

Figure D.1: The dependencies structure of AART.

Some statistics on AART:

```
Number of  files:         449
Number of directories:    68
Total size:               3099 KBytes
Estimated original LOC:   20,000
```

Corresponding statistics for work on landmark selection:

```
Number of  files:         11
Number of directories:    5
Total size:               126 KBytes
Estimated original LOC:   300
```

Also to see:

`http://www.danielsorogon.com/Webmaster/Projects/AART`

`http://www.danielsorogon.com/Webmaster/Projects/MDLGUI`

Appendix E

PROGRESS RECORDS

THIS appendix gives what is subjectively an abundant amount of information. Nevertheless, as part of the essential proof of progress, the following should not be left out.

Date	Description
22-26 September 2003	Introductory week
6-24 October 2003	Machine learning
20 January 2004	Oxford plenary meeting
10, 17 March 2004	Thesis writing seminar
2, 25 March 2004	Student presentations
30-31 March 2004	Manchester plenary meeting
April 2004	Mathematical methods
27 May 2004	UCL S&F meeting
2 June 2004	Ph.D. Workshop
21-25 June 2004	Surrey Summer School
13 July	Workshop, Guy's Hospital

Table E.1: Non-technical record of progress.

Also, a record of meetings, excluding the frequent ones with Carole Twining, completes the evidence of progress (see the next page for initials).

Date	Attendants	Description
24th September 2003	SRW	Issuing of letter
October 2003	SRW	×
October 2003	PB	Discussion about Advisor
26th November 2003	SRW	Advisor/Post-graduate Tutor
17th December 2003	SRW	×
8th January 2004	SM	×
January 2004	TFC	Registration, MATLAB, etc.
14th January 2004	SRW	Progress of work
16th January 2004	CJT, SRW, PB	Literature Report meeting
28th January 2004	CJT, TFC, SM, CJTw	S&F GC meeting
4th February 2004	CJT, CJTw	S&F GC meeting
9th February 2004	SRW	Presentation to be given
11th February 2004	CJT, TFC, CJTw	S&F GC meeting
18th February 2004	CJT, CJTw	S&F GC meeting
27th February 2004	CJT, TFC, CJTw	Registration Brainstorm
3rd March 2004	CJT, TFC, CJTw	S&F GC meeting
8th March 2004	SRW	Advanced Modules
12th March 2004	CJT, TFC, CJTw	S&F GC meeting
19th March 2004	CJT, SRW	Form 4
6th April 2004	SRW	Advice on giving a talk
28th April 2004	CJT, TFC, CJTw	S&F GC meeting
19th May 2004	CJT, CJTw	S&F GC meeting
24th May 2004	SRW	Current activities in ISBE
25th May 2004	CJT, TFC, CJTw	S&F GC meeting
2nd June 2004	CJT, CJTw	S&F GC meeting
7th June 2004	TW	×
9th June 2004	TW	×
10th June 2004	TW	×
11th June 2004	CJT, TFC, CJTw	S&F GC meeting
15th June 2004	CJT, TW	×
17th June 2004	CJT, CJTw	S&F GC meeting
17th June 2004	TW	×
28th June 2004	TW	×
12th July 2004	CJT, SRW	Form 5 Meeting

Table E.2: Miscellaneous work-related meetings.

CJT = Chris Taylor
TW = Tomos Williams
CJTw = Carole Twining
TFC = Tim Cootes
SM = Steve Marsland
PB = Paul Beatty

Meetings with the supervisor are also omitted (and can be found on the WWW).

Appendix F

LIST OF ABBREVIATIONS

SOME of the previously-mentioned acronyms are sorted alphabetically below. They should be used as a reference.

AAM: Active appearance model

AART: Autonomous appearance-based registration test-bed

ALU: Arithmetic and logic unit

ASM: Active shape model

COG: Centre of gravity

CPS: Clamped-plate spline

CT: Computed tomography

CVS: Concurrent versions system

EPSRC: Engineering and physical sciences research council

GC: Grand challenge

GPA: Generalised procrustes procedure

GSSEM: Graduate school of science, engineering and medicine

GUI: Graphical user interface

FFD: Free-form deformation

I/O: Input/output

IRC: Interdisciplinary research collaboration

ISBE: Imaging science and biomedical engineering

LOC: Lines of code

MDL: Minimum description length

MI: Mutual information

MIAS: Medical image and signal

MICCAI: Medical image computing and computer-assisted intervention

MIUA: Medical image understanding and analysis

MSD: Mean of squared differences

MRI: Magnetic resonance imaging

NMI: Normalised mutual information

NRR: Non-rigid registration

PCA: Principal component analysis

PDF: Probability density function; Portable document format

PDM: Point distribution model

PET: Positron emission tomography

RAM: Random access memory

S&F: Structure and function

SDM: Statistical deformation model

SNR: Signal-to-noise ratio

SPM: Statistical parametric mapping

SSD: Sum of squared differences

TS: Taboo search

UCL: University college london

WWW: World wide web

Appendix G

Additional Resources

All the existing project information is stored in a non-public domain. Please confer the index of all research material on-line at:

`http://www.danielsorogon.com/Webmaster/Research/resindex.htm`

This includes all organisational notes, documents, submissions, posters, forms, WWW links, project documentation, code documentation, experiments, outputs, logs, and more. Any data that has not been covered in this report can be located quite easily on that domain.

Appendix H

PRIMARY ON-LINE RESOURCES

World Wide Web Bibliography

AS much of the reading was based on educational and personal sites from the World Wide Web, several of the more dominant sources must be acknowledged.

- [WWW-1] `http://www.math.ufl.edu/help/matlab-tutorial/`

  An extensive MATLAB tutorial from the University of Florida.

- [WWW-2] `http://www.doc.ic.ac.uk/~dr/`

  Daniel Rueckert's academic pages.

- [WWW-3] `http://www-ipg.umds.ac.uk/d.hill/`

Derek Hill's abstracts and publications.

- [WWW-4] `http://www.dcs.gla.ac.uk/~mc/`

  Technical Reports of Matthew Cairn.

- [WWW-5] `http://www.imm.dtu.dk/image/research/`

  Related research in the Technical University of Denmark.

- [WWW-6] `http://www.ai.mit.edu/~viola/`

  Personal pages maintained by Paul Viola.

- [WWW-7] `http://www.isbe.man.ac.uk/~bim/`

  Publications and resources from Tim Cootes who ought to receive accolade for his clear explanations of statistical models.

- [WWW-8] `http://www.cs.jhu.edu/~wolff/course600.461/`

  Computer Vision at Johns Hopkins University.

- [WWW-9] `http://www.cs.wisc.edu/~dyer/cs766.html`

  Computer Vision at the University of Wisconsin.

- [WWW-10] `http://www.csse.monash.edu.au/~app/CSE5301/Lnts/L01.pdf`

  Neural networks material from Andrew P. Paplinskil.

- [WWW-11] `http://www.eg.org/EG/DL/Conf/EG91/papers/EUROGRAPHICS_91pp183_194_abstracthttp://www.eg.org/EG/DL/Conf/EG91/papers/EUROGRAPHICS_91pp183_194_abstract.pdf`

  Short and eloquent explanation on image discrepancy from Peter Shirley.

- [WWW-12] `http://www.ics.uci.edu/~eppstein/gina/interpolate.html`

  Explanation on interpolation from David Eppstein.

- [WWW-13] `http://www.lans.ece.utexas.edu/~strehl/diss`

  Alexander Strehl's dissertation with relation to mutual information.

- [WWW-14] `http://www.mdl-research.org/`

  An extensive resource on minimum description length with interactive examples.

- [WWW-15] `http://www-groups.dcs.st-and.ac.uk/~history/Mathematicians/Shannon.html`

  An ample resource for studying about Shannon and his work.

- [WWW-16] `http://www.fil.ion.ucl.ac.uk/spm/`

  The official UCL resource for SPM software, idealogy, etc.

- [WWW-17] `http://www.mathworks.nl/matlabcentral/fileexchange/loadAuthor.do?objectType=author&objectId=1094029`

  The general-purpose code shared by the author of this report.

Bibliography

[1] A. Abkar and H. Hedenmalm, "A Riesz Representation Formula for Super-Biharmonic Functions," Annales Academie Scientiarum Fennice Mathematica, vol. 26, pp. 305-324, 2001.

[2] J. Ahlberg and R. Forchheimer, "Face tracking for model-based coding and face animation," International Journal of Imaging Systems and Technology, vol. 13, pp. 8-22, 2003.

[3] C. J. Beeston and C. J. Taylor, "Automatic landmarking of cortical sulci," in Medical Image Computing and Computer-Assisted Intervention 2000, vol. 1935, pp. 125-133.

[4] K. K. Bhatia, J. V. Hajnal, B. K. Puri, A. D. Edwards, and D. Rueckert, "Consistent groupwise non-rigid registration for atlas construction," presented at ISBI, Arlington, VA, USA, 2004.

[5] F. L. Bookstein, "Principal Warps: Thin-Plate Spines and the Decomposition of Deformations," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 11, pp. 567-585, 1989.

[6] A. D. Brett and C. J. Taylor, "A method of automated landmark generation for automated 3D PDM construction," Image and Vision Computing, vol. 18, pp. 739-748, 2000.

[7] T. F. Cootes, C. Beeston, G. J. Edwards and C. J. Taylor, "A unified framework for atlas matching using Active Appearance Models," in Information Processing in Medical Imaging, Proceedings, vol. 1613, Lecture Notes in Computer Science, 1999, pp. 322-333.

[8] T. F. Cootes, G. J. Edwards and C. J. Taylor, "Active appearance models," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, pp. 681-685, 2001.

[9] T.F. Cootes, S. Marsland, C.J. Twining, K. Smith and C.J. Taylor, "Groupwise diffeomorphic non-rigid registration for automatic model building," Proceedings of the European Conference of Computer Vision 2004, pp. 316-32.

[10] R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton and C. J. Taylor, "3D statistical shape models using direct optimisation of description length," in Computer Vision - ECCV 2002 Pt Iii, vol. 2352, Lecture Notes in Computer Science, 2002, pp. 3-20.

[11] R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton and C. J. Taylor, "minimum description length approach to statistical shape modeling," IEEE Transactions on Medical Imaging, vol. 21, pp. 525-537, 2002.

[12] R. H. Davies, C. J. Twining, P. D. Allen, T. F. Cootes and C. J. Taylor, "Shape discrimination in the hippocampus using an MDL model," in Information Processing in Medical Imaging Proceedings, vol. 2732, Lecture Notes in Computer Science, 2003, pp. 38-50.

[13] T. F. Cootes, G. V. Wheeler, K. N. Walker and C. J. Taylor, "View-based active appearance models," Image and Vision Computing, vol. 20, pp. 657-664, 2002.

[14] T. F. Cootes, G. V. Wheeler, K. N. Walker and C. J. Taylor, "Coupled-view active appearance models," British Machine Vision Conference, vol. 1, pp. 52-61, 2000.

[15] B. M. Dawant, "Non-Rigid Registration of Medical Images: Purpose and Methods, A Short Survey," presented at Proceedings of the First IEEE International Symposium on Biomedical Imaging, 2002.

[16] S. Duchesne, J. C. Pruessner, and D. L. Collins, "Appearance-based segmentation of medial temporal lobe structures," NeuroImage, vol. 17, pp. 515-531, 2002.

[17] G. J. Edwards, C. J. T. Taylor and T. F. Cootes, "Interpreting face images using active appearance models," presented at Third IEEE International Conference on Automatic Face and Gesture Recognition, Nara, Japan, 1998.

[18] P. F. Felzenszwalb, "Representation and detection of deformable shapes," Computer Vision and Pattern Recognition, vol. 1, pp. 102-108, 2003.

[19] F. Glover, E. D. Taillard, and D. de Werra, "A user's guide to taboo search," Annals of Operations Search, pp. 3-28, 1993.

[20] A. Guimond, J. Meunier and J. Thirion, "Automatic computation of average brain models," presented at Medical Image Computing and Computer Assisted Intervention 1998, pp. 631-640, Cambridge, MA, USA.

[21] J. V. Hajnal, D. L. G. Hill and D. J. Hawkes, "Medical image registration," Boca Raton, Fla. ; London: CRC Press, 2001.

[22] S. Haker, A. Tannenbaum, and R. Kikinis, "Mass Preserving Mappings and Image Registration," presented at Fourth International Conference on Medical Image Computing and Computer Assisted Intervention, 2001.

[23] D. W. Hansen, M. Nielsen, J. P. Hansen, A. S. Johansen and M. B. Stegmann, "Tracking eyes using shape and appearance," IAPR Workshop on Machine Vision Applications, pp. 201-204, 2002.

[24] A. Hill, C. J. Taylor and A. D. Brett, "A framework for automatic landmark identification using a new method of nonrigid correspondence," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, pp. 241-251, 2000.

[25] T. Jebara, "Images as Bags of Pixels," presented at International Conference of Computer Vision, Proceedings, Nice, France, 2003.

[26] I.T. Joliffe, "Principal Component Analysis," Springer Series in Statistics, Springer, New York, 1986.

[27] S. C. Joshi and M. L. Miller, "Landmark Matching via Large Deformation Diffeomorphisms," IEEE Transaction on Image Processing, vol. 9, pp. 1357-1370, 2000.

[28] J. F. Kenney and E. S. Keeping, "Mathematics of Statistics," part 2, 2nd edition, Princeton, New Jersey: Van Nostrand, 1951.

[29] A. C. W. Kotcheff and C. J. Taylor, "Automatic construction of eigenshape models by genetic algorithm," in Information Processing in Medical Imaging, vol. 1230, pp. 1-14, Lecture Notes in Computer Science, 1997.

[30] A. Lanitis, "PROSOPO - A face image synthesis system," Advances in Informatics, Lecture Notes in Computer Science, vol. 2563, pp. 297-315, 2003.

[31] A. Lanitis, C. J. Taylor, and T. F. Cootes, "Automatic Face Identification System Using Flexible Appearance Models," Image and Vision Computing, vol. 13, pp. 393-401, 1995.

[32] A. Lanitis, C. J. Taylor, and T. F. Cootes, "Toward automatic simulation of aging effects on face images," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, pp. 442-455, 2002.

[33] L. LeBriquer and J. Gee, "Design of a statistical model of brain shape," presented at Proceedings of IPMI, 1997.

[34] B. Likar and F. Pernus, "A Hierarchical Approach to Elastic Registration Based on Mutual Information," Image and Vision Computing, vol. 19, pp. 33-44, 2001.

[35] J. Lötjönen and T. Mäkelä, "Elastic Matching Using a Deformation Sphere," presented at Fourth International Conference on Medical Image Computing and Computer Assisted Intervention, 2001.

[36] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal and P. Suetens, "Multimodality image registration by maximization of mutual information," IEEE Transactions on Medical Imaging, vol. 16, issue 2, pp. 187-198, 1997.

[37] D. Marr, "Understanding complex information processing systems," Vision, pp. 19-24, 1982.

[38] S. Marsland, C. J. Twining and C. J. Taylor, "Groupwise non-rigid registration using polyharmonic clamped-plate splines," presented at Medical Image Computing and Computer-Assisted Intervention, Montreal, Canada, 2003.

[39] T. Mitchell, "Machine Learning," New York ; London: McGraw-Hill, 1997.

[40] A. Najmi, R. A. Olshen and R. M. Gray, "A criterion for model selection using minimum description length," in Proceedings Compression and Complexity of Sequences, pp. 204-214, Salerno, Italy, 1997.S. C. Joshi and M. L. Miller, "Landmark Matching via Large Deformation Diffeomorphisms," IEEE Transaction on Image Processing, vol. 9, pp. 1357-1370, 2000.

[41] M. Petrou and P. Bosdogianni, "Image processing: the fundamentals," ISBN 0471 99883 4, 1999.

[42] J. P. W. Pluim, J. B. A. Maintz and M. A. Viergever, "Image Registration by Maximisation of Combined Mutual Information and Gradient Information," IEEE Transaction on Medical Imaging, vol. 19, pp. 809-814, 2000.

[43] J. P. W. Pluim, J. B. A. Maintz and M. A. Viergever, "Mutual information matching in multiresolution contexts," in Image and Vision Computing, vol. 19, issues 1-2, pp. 45-52, January 2001.

[44] W. H. Press, "Numerical recipes in C/C++ : the art of scientific computing. " Cambridge: Cambridge University Press, 2002.

[45] M. Rabbani and R. Joshi, "An overview of the JPEG 2000 still image compression standard," Signal Processing: Image Communication, vol. 17, pp. 3-48.

[46] J. R. Rissanen, "Stochastic complexity in statistical inquiry," presented at World Scientific Series in Computer Science, Singapore, 1989.

[47] P. Rogelj and S. Kovacic, "Similarity Measures for Non-Rigid Registration," presented at Medical Imaging 2001: Image Processing.

[48] D. Rueckert, A. F. Frangi and J. A. Schnabel, "Automatic construction of 3-D statistical deformation," IEEE Transactions on Medical Imaging, vol. 22, issue 8, pp. 1014-1025, 2003.

[49] D. Rueckert, A. F. Frangi and J. A. Schnabel, "Automatic construction of 3D statistical deformation models using non-rigid registration," presented at Medical Image Computing and Computer-Assisted Intervention 2001.

[50] J. A. Schnabel, C. Tanner, A. Castellano-Smith, A. Degenhard, M. O. Leach, D. R. Hose, D. L. G. Hill, and D. J. Hawkes, "Validation of non-rigid image registration using finite element methods: application to breast MR images," IEEE Transactions on Medical Imaging, vol. 22.

[51] M. Sonka, V. Hlavac and R. Boyle, "Image processing, analysis and machine vision", Pacific Grove, Calif. ; London: PWS Publishing, 1999.

[52] M. B. Stegmann, B. K. Ersboll and R. Larsen, "FAME - A flexible appearance modeling environment," IEEE Transactions on Medical Imaging, vol. 22, pp. 1319-1331, 2003.

[53] C. Studholme, D.L.G. Hill and D.J. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," Pattern Recognition, vol. 32, issue 1, pp. 71-86, 1999.

[54] N. A. Thacker, D. Prendergast, and P. I. Rockett., “B-fitting: an estimation technique with automatic parameter selection,” presented at British Machine Vision Conference 1996.

[55] P. M. Thompson, M. S. Mega, C. Vidal, J. L. Rapoport, and A. W. Toga, “Detecting Disease-Specific Patterns of Brain Structure using Cortical Pattern Matching and a Population-Based Probabilistic Brain Atlas,” presented at Proceedings of the 17th International Conference on Information Processing in Medical Imaging, 2001.

[56] A. Toga, “Brain Warping,” San Diego, CA, USA: Academic Press, 1999.

[57] A. Trouve, “Diffeomorphisms Groups and Pattern Matching in Image Analysis,” International Journal of Computer Vision, vol. 28, pp. 213-221, 1998.

[58] C. J. Twining and S. Marsland, “Constructing diffeomorphic representations of non-rigid registrations of medical images,” presented at Information Processing in Medical Imaging 2003.

[59] C. J. Twining, S. Marsland, and C. J. Taylor, “Groupwise non-rigid registration: the minimum description length approach,” British Machine Vision Conference 2004.

[60] C. J. Twining and C. J. Taylor, “The use of kernel principal component analysis to model data distributions,” Pattern Recognition, vol. 36, pp. 217-227, 2003.

[61] P. Viola and W. M. Wells, “Alignment by maximization of mutual information,” International Journal of Computer Vision, vol. 24, pp. 137-154, 1997.

[62] K. N. Walker, T. F. Cootes and C. J. Taylor, “Determining correspondences for statistical models of appearance,” in Computer Vision - ECCV 2000, Pt I, Proceedings, vol. 1842, Lecture Notes in Computer Science, 2000, pp. 829-843.

[63] Y. Wang and L. H. Staib, “Elastic model based non-rigid registration incorporating statistical shape information,” presented at Medical Image Computing and Computer-Assisted Intervention, 1998.

[64] Y. Wang and L. H. Staib, “Integrated approaches to non-rigid registration in medical images,” presented at Workshop on Applications of Computer Vision, 1998.

[65] S. K. Warfield, J. Rexilius, P. S. Huppi, T. E. Inder, E. G. Miller, W. M. Wells, III, G. P. Zientara, F. A. Jolesz and R. Kikinis, “An Entropy Measure to Assess Nonrigid Registration Algorithms for Statistical Atlas Construction,” presented at Medical Image Computing and Computer-Assisted Intervention, 2001.

[66] J. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dawant, J. Maurer, C. R., R. M. Kessler, R. J. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. A. van den Elsen, S. Napel, T. S. Sumanaweera, B. Harkness, P. F. Hemler, D. L. G. Hill, D. J. Hawkes, C. Studholme, J. B. A. Maintz, M. A. Viergever, G. Malandain, X. Pennec, M. E. Noz, J. Maguire, G. Q., M. Pollack, C. A. Pelizzari, R. A. Robb, D. Hanson and R. P. Woods, “Comparison and evaluation of retrospective intermodality brain image registration techniques,” Journal of Computer Assisted Tomography, vol. 21, pp. 554-566, 1997.

[67] M. Xu and W. L. Nowinski, “Talairach-Tournoux Brain Atlas Registration Using a Metalforming Principle-Based Finite Element Method,” Medical Image Analysis, vol. 5, pp. 271-279, 2001.

[68] C. Yam, M. S. Nixon, and J. N. Carter, “Automated person recognition by walking and running via model-based approaches,” Pattern Recognition, vol. 37, 2004.

[69] L. Younes, “Deformations, Warping and Object Comparison: A Tutorial,” 2000.

Also see the on-line resources listed in Appendix H of this report.

Document was compiled using LaTeX version 3.14159