Landmark Influence Visualization in Active Shape Models

Master’s Thesis

Korijn van Golen
Landmark Influence Visualization
in Active Shape Models

THESIS

submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE

in

COMPUTER SCIENCE

by

Korijn van Golen
born in Zwijndrecht, the Netherlands

TU Delft
Computer Graphics and Visualization Group
Department of Intelligent Systems
Faculty EEMCS, Delft University of Technology
Delft, the Netherlands
www.ewi.tudelft.nl

Clinical Graphics B.V.
Molengraafsingel 12-14
2629JD Delft
The Netherlands
www.clinicalgraphics.com
Landmark Influence Visualization in Active Shape Models

Author: Korijn van Golen
Student id: 1509152
Email: korijn@gmail.com

Abstract

Segmentation of shapes in Magnetic Resonance Image (MRI) data is a non-trivial problem. Active Shape Models (ASM) is an approach that attempts to fit a Statistical Shape Model to landmarks. One possible approach is to let a user fit the model onto the MRI data by placing landmarks for ASM. Usually, the user can only see the global fit of the Statistical Shape Model; adjustments to landmarks have a global effect. This hinders the user in time-efficiently constructing a set of landmarks that maximizes the accuracy of the segmented shape.

A novel landmark influence visualization is proposed that improves the usability of Active Shape Models. Landmark influence is defined as the change the model fit will undergo if a landmark is removed. The landmark influence is visualized using disk glyphs around landmarks. A region-of-interest filter can be used to inspect influence on local regions. Distance maps and animations visualize how the shape will change.

A user study was performed to evaluate the visualization. No correlation was found between segmentation accuracy and the visualization. Similarly, no correlation was found between the time required for segmentation and the visualization. Nevertheless, a questionnaire revealed that users found the disk glyphs and animations useful in their segmentation workflow.

Thesis Committee:

Chair: Prof. Dr. E. Eisemann, Faculty EEMCS, TU Delft
University supervisor: Dr. A. Vilanova, Faculty EEMCS, TU Delft
Company supervisor: Dr. Ir. P.R. Krekel, Clinical Graphics B.V.
Committee Member: Dr. D.M.J. Tax, Faculty EEMCS, TU Delft
Preface

I would like to thank Dr. Ir. P.R. Krekel of Clinical Graphics B.V. for the constructive feedback and key resources provided for the project. I would also like to thank Dr. A. Vilanova for her valuable suggestions and feedback during the planning and execution of this research. Their ongoing dedication has been essential to the completion of this work.

Korijn van Golen
Dordrecht, the Netherlands
June 26, 2014
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td>iii</td>
</tr>
<tr>
<td>Contents</td>
<td>v</td>
</tr>
<tr>
<td>List of figures</td>
<td>vii</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2 Related work</td>
<td>5</td>
</tr>
<tr>
<td>2.1 Active Shape Models</td>
<td>5</td>
</tr>
<tr>
<td>2.2 Comparative Visualization</td>
<td>6</td>
</tr>
<tr>
<td>3 Method</td>
<td>9</td>
</tr>
<tr>
<td>3.1 Overview</td>
<td>9</td>
</tr>
<tr>
<td>3.2 Background</td>
<td>9</td>
</tr>
<tr>
<td>3.3 Determination of Landmark Influence</td>
<td>12</td>
</tr>
<tr>
<td>3.4 Landmark Influence Visualization</td>
<td>13</td>
</tr>
<tr>
<td>4 Evaluation</td>
<td>21</td>
</tr>
<tr>
<td>4.1 Setup</td>
<td>21</td>
</tr>
<tr>
<td>4.2 Results and discussion</td>
<td>22</td>
</tr>
<tr>
<td>5 Conclusions and future work</td>
<td>27</td>
</tr>
<tr>
<td>Bibliography</td>
<td>29</td>
</tr>
<tr>
<td>Nomenclature</td>
<td>33</td>
</tr>
<tr>
<td>A Active Shape Models for the femur</td>
<td>35</td>
</tr>
<tr>
<td>A.1 Building a Statistical Shape Model</td>
<td>35</td>
</tr>
<tr>
<td>A.2 Applying the model</td>
<td>37</td>
</tr>
</tbody>
</table>
CONTENTS

B Principal Component Analysis 41
C Singular Value Decomposition 43
D User Study Design 45
E Questionnaire 49
List of figures

1.1 Active Shape Models fits a Statistical Shape Model (in this example a femur) onto a set of landmarks (blue sphere glyphs). (The text glyphs are unrelated, and are used in the global pose determination step of Active Shape Models, see Appendix A.2.) .................................................. 2

3.1 The landmark influence visualization pipeline. Active Shape Models (ASM) takes landmarks and a Statistical Shape Model (SSM) as inputs, and fits the model onto the landmarks. Landmark Influence is then determined. Finally, the model fit and landmark influence are visualized. .................. 10

3.2 Effect of varying each of the first three shape model parameters (vertically arranged) in turn between ±4 standard deviations (horizontally arranged), after performing PCA analysis on the shape space $X$. In this case, the modes determined by PCA respectively represent size, curvature and thickness. .... 11

3.3 The number of disk glyphs (0-3) and their colors (from blue to red) indicate a landmark’s influence. ................................................................. 14

3.4 The number of visible disks is determined by first mapping the influence $I_l$ to the range $[0 - 3]$ by scaling linearly and then rounding up to the nearest integer. 15

3.5 The color of each disk is determined by the landmark influence $I_l$. The color is evaluated independently for each disk. ............................... 15

3.6 The opacity of each disk is determined by the landmark influence $I_l$. The opacity is evaluated independently for each disk. ............................... 16

3.7 In order to reduce the visual load in scenes with large amounts of landmarks (top), blue disk glyphs are made transparent (bottom). (The text glyphs are unrelated, and are used in the global pose determination step of Active Shape Models, see Appendix A.2.) .................................................. 17

3.8 The distance map $D_l$ is visualized on the shape model instance using a color map. (The text glyphs are unrelated, and are used in the global pose determination step of Active Shape Models, see Appendix A.2.) .................................................. 18
3.9 The shape change if a landmark were to be removed is visualized by animating from shape $x$ to $x'$ and back. Since an animation is difficult to convey on paper, we resort to showing the difference between $x$ and $x'$, highlighted in white. All but the relevant landmark were removed in this image for clarity.

3.10 The region-of-interest filter allows the user to see the influence of landmarks on a specific region of the shape model instance. In this case, a box widget (in the left image) is used to determine the region of interest. Any kind of region can be used however. The right image shows the same situation without the ROI filter. Note that the influence values of landmarks are different, as can be seen by the disk glyphs. (The text glyphs are unrelated, and are used in the global pose determination step of Active Shape Models, see Appendix A.2.)

4.1 The boxplots show the measured accuracy (vertical axis; RMSE, see Equation (4.1)) of the users’ segmentations of four MRI datasets, using Active Shape Models, with Landmark Influence Visualization disabled (left) and enabled (right).

4.2 The boxplots show the time it took users to segment a dataset (vertical axis), using Active Shape Models, with Landmark Influence Visualization disabled (left) and enabled (right), for each of the 4 datasets.

4.3 Users were asked if they felt that the various features of landmark visualization were useful, on a 5-point Likert scale.

A.1 Suitable landmark positions (image from [Cootes et al., 2000]).

A.2 Effect of varying each of the first three model shape parameters (vertical axis) in turn between $\pm 4$ standard deviations (horizontal axis).

A.3 Landmarks on the image data are annotated with ‘zone labels’ to indicate which region of the shape model they belong to. In this case a landmark has been assigned to the region ‘trochanter minor’, abbreviated to ‘tm’.

A.4 Landmark zones on the upper section of the femur. The annotated landmark of Figure A.3 is assigned to the region labeled ‘trochanter minor’.

B.1 The principal components of a 2-dimensional set of observations drawn from a normal distribution.

B.2 The first two principal components of a 3-dimensional set of observations drawn from a normal distribution.
Chapter 1

Introduction

Pre-operative planning for surgery on the hip joint should preferably be done in 3D, because 3D shapes represent more information that can be used in analysis and simulations. Examples include determining the range-of-motion of joints before and after an operation [Krekel et al., 2006]. Clinics can usually supply either Computer Tomogram (CT) or Magnetic Resonance (MR) Imaging data to pre-operative planning systems in order to attain 3D bone shapes [OECD, 2012]. When performing CT scans, patients are exposed to radiation. In some cases, a patient will therefore undergo a radiation-free MRI scan instead. One of the first steps that pre-operative planning systems typically perform is segmentation; the shape of interest, e.g., the femur, is extracted from the image data. The segmented shape can then be used for subsequent tasks such as analysis and simulations [Krekel et al., 2006; Lee et al., 2003; Beller et al., 2008].

There are several approaches to this non-trivial segmentation problem. There are bottom-up approaches where features such as edges are recognized in the images and then grouped in order to reconstruct a shape. A top-down approach uses a prior model of what is expected in the data and attempts to find the best match of the model to the new imaging data.

When dealing with MR images bottom-up approaches are not effective. Bone structures are not easily identified in MR images. Therefore the segmentation based on local features is not possible.

It is possible to segment data in a top-down fashion by fitting a Statistical Shape Model (SSM) to the image data, a method known as Active Shape Models (ASM). This method uses external knowledge about what is expected in the data and attempts to find the best match of the model [Cootes et al., 2000; Rajamani et al., 2007; Heimann and Meinzer, 2009].

The ASM approach requires the placement of landmarks to register the SSM to, as in Figure 1.1. Placement of these landmarks is done manually, as automated landmark placement would require local features such as edges to be detected, and this is not possible for MR images as explained earlier.

Although model-based segmentation methods are constantly being improved, the segmented bone models are not guaranteed to be accurate. ASM attempts to fit a statistical model to a set of landmark points. The accuracy of the segmented shape in relation to the
Figure 1.1: Active Shape Models fits a Statistical Shape Model (in this example a femur) onto a set of landmarks (blue sphere glyphs). (The text glyphs are unrelated, and are used in the global pose determination step of Active Shape Models, see Appendix A.2.)
true shape depends on the accuracy of the manually placed landmarks and the quality of the SSM. The accuracy of a landmark’s position in relation to the underlying image data is in turn dependent on the quality of the MR images and ultimately on the user’s capability of interpreting and assessing the images and the model fit. Clinicians intend to prepare their surgery with the segmented shapes, and therefore it is essential to maximize accuracy by improving the usability of ASM.

Another important aspect is time; patients need to be treated quickly and effectively, not only to save expenses. Therefore, the duration of the pre-operative planning process needs to be minimized as well.

In practice, it can be difficult to place landmarks correctly as the SSM may not react as expected to the input landmarks. This is the case because a local change will have a global effect on the segmentation. The time required to properly set the landmarks and the accuracy of the segmented bone shapes may suffer under the usability issues that arise in the landmark placement step. This makes the approach less applicable to situations where time is scarce or where a high degree of accuracy is required, as is the case in pre-operative planning. Therefore, it is critical that the interactive process is improved.

Typically, it is unclear how a landmark point influences the global model fit. We hypothesize that if the user is able to assess or even predict the influence of a landmark on the model fit, he is able to develop an intuitive understanding of the model he is working with, without having to understand the multi-dimensional shape spaces that are involved. Then, if we define and measure a landmark’s influence on the model fit and visualize it, the user can gain a better understanding of how the global model fit reacts to the set of landmarks. We expect that this allows the user to determine if and how the set of landmarks should be modified in order to improve the global model fit, achieving a more accurate model fit in less time.

In this work a novel visualization technique is presented that visualizes landmark influence to support the user in time-efficiently setting up a set of landmark points that maximizes the reconstructed shape’s accuracy.

The main research questions of this project are:

1. Can the accuracy of reconstructed shapes using Active Shape Models be improved by visualizing Landmark Influence?

2. Can the time required to reconstruct a shape using Active Shape Models be reduced by visualizing Landmark Influence?

3. How much does Landmark Influence Visualization facilitate the usability of the Active Shape Models method?

The contributions of this project are the following:

- A definition of landmark influence.
Introduction

- A landmark influence visualization technique.
- An evaluation of the application of the landmark influence visualization technique to the interactive process of setting up landmarks for active shape models.

The rest of this article is structured as follows: Chapter 2 presents previous work on this topic. Chapter 3 goes into full detail on landmark influence visualization, starting with an overview of the visualization pipeline in Section 3.1. Background information required to develop a formal understanding of the Active Shape Models algorithm is given in Section 3.2. Chapter 4 presents the evaluation performed in order to assess the research questions. Section 4.2 discusses the results of the evaluation. Chapter 5 completes the exposition with conclusions and future work.
Chapter 2

Related work

We will now move on by discussing further research conducted on Active Shape Models. Afterwards, comparative visualization methods that are relevant to landmark influence visualization will be discussed.

2.1 Active Shape Models

Active Shape Models (ASM) is a segmentation method introduced by Cootes et al. [1995, 2000], making use of shape spaces by Kendall [1984]. A Statistical Shape Model (SSM) is built by performing statistical analysis on a set of labeled examples. The variance in shape is analyzed, resulting in a statistical model. This knowledge about plausible shape variance is used when interpreting a new image by finding the best match of the model to the new image’s data. In order to map to the data, ASM requires a set of landmarks to fit the SSM to. One approach is to let the user place the landmarks. The accuracy of the model fit depends on both the quality of the model itself and the landmarks that the user places on the image(s). ASM was extended by Cootes et al. [1998] with Active Appearance Models, wherein texture information is added to the model into a Statistical Appearance Model.

Heimann and Meinzer [2009] conclude that shape reconstruction software used in clinical settings should always offer possibilities for manual corrections.

Constrained Active Appearance Models (cAAM) was proposed by Cootes and Taylor [2001] in which the user can ‘pin’ certain model points to specific landmarks. Evaluations showed that pinning too many points can result in a deterioration of the model fit [Heimann and Meinzer, 2009].

van Ginneken et al. [2003] proposed interactive Active Shape Models (iASM) where the user can adjust the landmark positions by dragging them to their correct position in an interactive process.

Busking et al. [2010] created a shape space visualization technique which allows for visual exploration of the shape space and validation. Although the technique can grant the
user a deeper understanding of the statistical shape model itself, it is not designed to aid the user in the segmentation process using a Statistical Shape Model.

A recent effort by von Landesberger et al. [2013] attempts to give the user a deeper understanding of how the statistical shape model performs a given segmentation by combining interactive visualization and automatic data analysis. Their method enables the user to gain a deeper understanding of the segmentation process and in detecting problems in the selected algorithm and its configuration (e.g., the amount of parameters, iterations). Their tool consists of three linked views which show 1) the progress of segmentation quality, 2) a combined view of the segmentation process and static views of selected iterations, and 3) an overview of convergence patterns per landmark. Our approach differs in the respect that landmark influence visualization is intended to help the user place landmarks correctly, rather than analyze and assess the process itself. However, their method can be used in conjunction with landmark influence visualization in order to get the benefits of both.

### 2.2 Comparative Visualization

Since we intend to visualize the change that is caused in a shape model instance by a landmark, surface to surface comparison is an interesting field for this work.

Comparison tasks are relevant in many fields, such as biology, network analysis, organic chemistry and medical physiology, focusing on data objects of various types, including surfaces. Visualization research often focuses on comparison tasks in specific cases, while many similarities can be identified independent of the data being compared. Gleicher et al. [2011] addressed this situation by presenting a general taxonomy of visual designs for comparison. They found that all designs are combinations of juxtaposition, superposition and explicit encodings. The authors note that the reviewed works suggest that interaction, analytical & statistical tools, and animation offer further possibilities to augment visual comparison.

In work by Tversky et al. [2002] animation is presented as the natural approach for conveying concepts of change. Animations may be hard to perceive due to perceptual and cognitive limitations, such as visual overload. For the same reason, animations may be comprehended discretely, even when the motion is actually smooth and/or continuous. Interactivity can remedy these difficulties, however, by allowing the user to focus on specific parts of the animation by pausing, restarting, zooming, rotating, etc. Finally, Tversky et al. [2002] note that the most promising use of animations seems to be to convey real-time changes and re-orientations in time and space.

Heer and Robertson [2007] expand on the topic by noting that motion is highly effective at attracting attention to regions of interest, restating that animation is very effective at conveying changes and communicating cause-and-effect relationships. It is important to note that the ability to grab attention can also distract the user’s attention from less apparent regions of interest.
These works reinforce the idea that a landmark influence visualization technique could make use of animations and interactivity in order to convey the changes that are caused in the shape model instance by adjusting the set of landmarks.

Comparative surface visualization techniques rely on various measures to express the difference between shapes [Wilson et al., 1997; Di Gesú and Starovoitov, 1999; Veltkamp, 2001; Li et al., 2006]. They express differences at a global level however and this hides important local information. Only a small amount of methods address local differences [Gatzke et al., 2005; Masuda et al., 2003]. Typically, these measures require a correspondence mapping (e.g., a point correspondence) between the surfaces being compared. A color map is displayed on one of the surfaces to indicate the differences. This hides the shape information of the other surface.

In uncertainty visualization alternatives have been proposed, displaying either one or both of the nested surfaces transparently [Johnson and Sanderson, 2003]. This has the advantage that both shapes can be seen simultaneously, although it is harder to understand the shape of transparent surfaces, especially when they are overlaid. Some methods attempt to clarify the shape of transparent surfaces using textures [Interrante et al., 1997; Bair and House, 2007].

In a method by Weigle et al. [2005], two surfaces are compared visually by calculating the distance between corresponding points on the surfaces and displaying a color map on one of the two surfaces. They also propose an extension to the approach that replaces the color map with colored curvature glyphs.

Weigle et al. [2005] also proposes the use of constructive solid geometry (CSG). The intersecting regions of two surfaces are shown opaquely as a single opaque object, and the remaining parts of the two surfaces are shown transparently. This makes it apparent for the viewer what parts of the surfaces are intersecting.

Building on the previous work, Busking et al. [2011] show how curvature glyphs can hide small intersections and proposes a number of improvements. Both surfaces are shown simultaneously, intersection contours are added and local distance cues are given to improve perception, either in the form of fogging or distance fields. Another feature they added is thresholding; small differences can be removed from view to focus the attention to the larger differences. They also state that shape models are a special case of comparison, for which correspondence between points is available and relevant for the comparison. Point-correspondence glyphs (e.g., lines) are used to show the difference between surfaces when using a statistical shape model.
Chapter 3

Method

The proposed Landmark Influence Visualization technique will be described in the following sections. An overview of the algorithm will be given in Section 3.1. Background information on the Active Shape Models algorithm is provided in Section 3.2. The details of determining landmark influence will be explained in Section 3.3. Finally, in Section 3.4, we describe the visualization of landmark influence.

3.1 Overview

The visualization pipeline for this algorithm is an extension to Active Shape Models. See Figure 3.1 for a diagram of the visualization pipeline. Active Shape Models is performed first with its regular inputs, and Landmark Influence Visualization is performed afterwards, in two steps: landmark influence determination, an extra step added onto the end of Active Shape Models, and finally, landmark influence visualization. These two steps will be discussed in detail in Sections 3.3 and 3.4.

3.2 Background

We will now briefly go over the necessary details of the Active Shape Models algorithm. For a complete overview of Clinical Graphics’ implementation, see Appendix A. For an overview of all the variables used, the reader can refer to the Nomenclature section.

Overview The Active Shape Models algorithm takes as input a Statistical Shape Model (SSM) and a set of landmarks to match the SSM to. The SSM is used to find an optimal fit of the shape model to the set of landmarks.

Statistical Shape Model The SSM consists of \( n \) volumes that are annotated with \( m \) landmarks, or point coordinates. These landmarks are placed at suitable locations, or more specifically, at corners and/or junctions of object boundaries and at easily recognizable biological locations. The set of landmarks is augmented with equally spaced intermediate points along the object boundaries. All volumes have the same set of landmarks and all landmarks
3.2 Background

Figure 3.1: The landmark influence visualization pipeline. Active Shape Models (ASM) takes landmarks and a Statistical Shape Model (SSM) as inputs, and fits the model onto the landmarks. Landmark Influence is then determined. Finally, the model fit and landmark influence are visualized.

correspond to each other across volumes. This is important for the Analysis step later on.

The landmarks of a volume define a shape. The shapes are transformed to a common coordinate frame by translating each shape’s center of gravity to the origin and subsequently rotating the shapes into a common orientation using the iterative closest point (ICP) algorithm [Besl and McKay, 1992]. For every shape, its $m$ landmarks can then be written as a vector $x_j = (x_{j1}, y_{j1}, z_{j1}, x_{j2}, y_{j2}, z_{j2}, \ldots, x_{jm}, y_{jm}, z_{jm})$, where $j = 1, \ldots, n$ (recall that there are $n$ shapes). Then, putting all vectors $x_j$ side-by-side, $X = (x_1|\ldots|x_n)$ is the $n \times 3m$ shape space holding all of the landmark coordinates, where every column corresponds to a shape.

Analysis Principal Component Analysis (PCA) is performed on $X$. Matrix $X$ forms a cloud of points in $3m$-dimensional space. PCA computes the main axes of that point cloud, enabling approximations of the original points using a model with less than $3n$ parameters. See Figure 3.2 for a visualization of shape approximations on the main axes of variation calculated by PCA.

The PCA yields the eigenvectors $U_i$ and eigenvalues $\lambda_i$ of $X$. We define $P$ as the matrix containing the first $t$ eigenvectors, as in $P = (U_1|U_2|\ldots|U_t)$. Then, $P$ defines a subspace of the shape space $X$. $U_i$ is referred to as the $i$th mode of the model. Then, any shape $x$ can be approximated using

$$x \approx \bar{x} + Pb,$$

(3.1)
Figure 3.2: Effect of varying each of the first three shape model parameters (vertically arranged) in turn between ±4 standard deviations (horizontally arranged), after performing PCA analysis on the shape space $X$. In this case, the modes determined by PCA respectively represent size, curvature and thickness.

where $\bar{x}$ is the mean shape and $b$ is a $t$-dimensional set of parameters that defines the shape in the shape subspace defined by $P$.

The standard deviation of the $i$th mode across the training set is given by the eigenvalue $\lambda_i$. By varying the elements of $b$ we can vary the shape $x$ using Equation (3.1).

**Landmarks** A set of landmarks $L$ is created by the user in an interactive process. These landmarks describe the shape that the SSM should be fit to. Refer to Figure 1.1 for an example.

**Fitting** An iterative process is applied to fit the model to the landmarks. An initial model instance $x$ is selected. Typically this is the mean shape, e.g., $x := \bar{x}$. Next, an initial global transform is determined by applying the ICP algorithm [Besl and McKay, 1992]. The process then starts to alternate between optimizing the model fit to the landmarks
and updating the global pose of the model. Optimizing the model fit to the landmarks is done in three steps:

1. For every landmark in $L$, the closest point in the model instance $x$ is determined by minimizing distance, yielding point correspondence function $C$.

2. A least squares solver is used to solve for $b$ the equation

$$C(P)b = L - C(\bar{x}).$$ \hspace{1cm} (3.2)

3. The model instance $x$ is updated with the new parameters $b$ using

$$x := \bar{x} + Pb.$$ \hspace{1cm} (3.3)

The global pose of the model instance is then updated by applying the ICP algorithm [Besl and McKay, 1992] again. This iterative process is repeated $k$ times until convergence is achieved.

Convergence is determined with a user defined heuristic. Simple implementations execute a fixed number of iterations. More advanced implementations wait for an error metric such as the root mean square error to stabilize.

### 3.3 Determination of Landmark Influence

We are interested in the influence of a landmark on the shape model instance fit. We are not interested in the effect of a landmark on the global orientation of the shape model instance, since we consider the global orientation determination step of ASM to be a separate problem.

Therefore, we define the influence of a landmark $l$ as the euclidean distance between all $m$ pairs of points between two shape model instances $x$ and $x'$, where the $x'$ is the shape model instance $x$ fitted without taking landmark $l$ into account. In other words, a landmark’s influence can be seen as the change that the shape model instance would undergo if that landmark were to be removed.

After performing the ASM algorithm, we have the point correspondence function $C$ of the last iteration, shape space matrix $P$, the user-defined set of landmarks $L$ and a shape model instance $x$ fitted to $L$. For a summary of the ASM algorithm, see Section 3.2.

Let $I_l$ be the influence of a landmark $l$ on the shape model instance $x$. In order to determine the landmark influence of every landmark $l \in L$, an additional step is executed after the Active Shape Models algorithm has completed.

For every landmark $l \in L$: 

1. For every landmark in $L$, the closest point in the model instance $x$ is determined by minimizing distance, yielding point correspondence function $C$.

2. A least squares solver is used to solve for $b$ the equation

$$C(P)b = L - C(\bar{x}).$$ \hspace{1cm} (3.2)

3. The model instance $x$ is updated with the new parameters $b$ using

$$x := \bar{x} + Pb.$$ \hspace{1cm} (3.3)
1. The $i$th landmark is removed from landmarks $L$ and point correspondence function $C$, yielding respectively $L'$ and $C'$.

The function $C'$ selects the points from shape model instance $\bar{x}$ and shape space matrix $P$ that correspond to the landmarks $L'$.

2. In order to determine the shape model parameters $b'$ for the fit without $l$, Equation (3.2) is solved for $b'$ using landmarks $L'$ and point correspondence function $C'$.

$$C'(P)b' = L' - C'(<\bar{x}>). \quad (3.4)$$

3. The shape model instance $x'$ is determined using Equation (3.3):

$$x' := \bar{x} + P b'.$$  \hspace{1cm} (3.5)

4. The landmark influence $I_l$ is now determined by taking the euclidean norm of the euclidean distances between the $m$ pairs of points of the shape model instances $x$ and $x'$. Since they are both shape model instances, the $m$ points correspond to each other across instances by construction.

$$I_l = \sqrt{\sum_{i=1}^{m} \|x'_i - x_i\|^2} \quad (3.6)$$

Note that this metric is equivalent to the Frobenius norm of the vector $x' - x$.

Therefore, in our method, a landmark’s influence is defined as the euclidean distance between the shape model instances $x$ and $x'$. In other words, a landmark’s influence can be seen as the change that the shape model instance would undergo if that landmark were to be removed.

### 3.4 Landmark Influence Visualization

Several visual tools make up landmark influence visualization: First, in Section 3.4.1, the range of influence values will be visualized. Then, in Section 3.4.2, we will discuss disk glyphs that are used to visually represent a landmark’s influence value $I_l$. Next, the visualization of landmark distance maps will be explained in Section 3.4.3. Then, an animation that indicates how the shape model fit will change if a landmark is removed will be described in Section 3.4.4. Finally, in Section 3.4.5, we will describe a Region-Of-Interest (ROI) filter that can be used to analyze the influence of landmarks on a specific region of the shape model.
3.4 Landmark Influence Visualization Method

3.4.1 Influence range

The most basic visualization of landmark influence is to display the range of values found during the determination step in the pipeline (see Section 3.3). We propose displaying the minimum and maximum values of all $I_l \in I$ in the simple format $(\min I_l \in I, \max I_l \in I)$ on screen.

Note that the maximum value can be thought of as a stability metric. If all landmarks have low influence values, that means that the shape model instance will not change dramatically if any single landmark were to be removed.

3.4.2 Disk glyphs

In order to enable users to compare the influence of different landmarks, a visual representation of landmark influence is proposed. This visualization will be in the form of glyphs that decorate the existing landmark glyphs (in our implementation, spheres). The glyphs will be positioned in the direct neighborhood of the landmark so that it is obvious to the user that the glyphs and landmark are related to each other. See Figure 3.3.

To visualize individual landmarks’ influences we propose to display 0, 1, 2 or 3 disks around the landmarks. All influence values $I_l \in I$ are mapped to the range $[0 – 3]$ by scaling linearly. The mapped values can be rounded up to the nearest integer to determine the amount of disks that are visible, so that the landmark(s) with minimum influence will have 0 disks, and the landmark(s) with maximum influence will have 3 disks. See Figure 3.4.

Figure 3.3: The number of disk glyphs (0-3) and their colors (from blue to red) indicate a landmark’s influence.
Method

3.4 Landmark Influence Visualization

Figure 3.4: The number of visible disks is determined by first mapping the influence $I_l$ to the range $[0 - 3]$ by scaling linearly and then rounding up to the nearest integer.

The number of visible disks

![Number of Visible Disks](image)

Figure 3.5: The color of each disk is determined by the landmark influence $I_l$. The color is evaluated independently for each disk.

The color of the disks is used to indicate the fractional part of the mapped influence value. The first disk’s color is determined by mapping from blue to red on the range $[0 - 1]$. Similarly, the second and third disks’ colors are determined by mapping from blue to red on respectively the ranges $[1 - 2]$ and $[2 - 3]$. See Figure 3.5.

The color map from blue to red was chosen for two reasons: (1) there is no type of color blindness that has trouble differentiating between blue and red (except for total color blindness), and (2) the color map represents a linear, continuous scale with two endpoints; a minimum and a maximum.

In summary, the number of disk glyphs helps the user to categorize landmarks into three
3.4 Landmark Influence Visualization Method

3.4.2 Opacity

The opacity of each disk is determined by the landmark influence $I_l$. The opacity is evaluated independently for each disk. Groups that represent small, moderate and large influence values. When there are outliers they will likely be the only landmarks with 3 disk glyphs. They therefore take up significantly more screen space than other landmarks, and draw the user’s attention. The color of the disk glyphs enables the user to compare the influence of landmarks that share the same amount of disk glyphs.

When there are large amounts of landmarks, this visualization will occupy large parts of the screen space. In order to reduce the visual load, the approach can be extended with disk glyph opacity. The opacity will be determined in the same way as the color: see Figure 3.6. This means that disk glyphs are fully transparent when they are blue, and that they will become fully opaque as their color goes to red. More influential landmarks will then still draw attention while less influential landmarks will take up less screen space. See Figure 3.7.

3.4.3 Distance maps

For every landmark, there is a distance map $D_l$. It is a vector containing the pair-wise euclidean distances between all $m$ points of the two shape model instances $x$ and $x'$. Since both instances resulted from the same model, there is no need for determining a point correspondence. See Equation (3.7).

$$D_l = (\|x'_1 - x_1\|, \|x'_2 - x_2\|, \ldots, \|x'_m - x_m\|).$$

[Figure 3.6: The opacity of each disk is determined by the landmark influence $I_l$. The opacity is evaluated independently for each disk.]

The user can select a landmark and toggle the visualization of its distance map. It is visualized simply by using a color map. See Figure 3.8.
Figure 3.7: In order to reduce the visual load in scenes with large amounts of landmarks (top), blue disk glyphs are made transparent (bottom). (The text glyphs are unrelated, and are used in the global pose determination step of Active Shape Models, see Appendix A.2.)
Figure 3.8: The distance map \( D_l \) is visualized on the shape model instance using a color map. (The text glyphs are unrelated, and are used in the global pose determination step of Active Shape Models, see Appendix A.2.)
3.4 Landmark Influence Visualization

Figure 3.9: The shape change if a landmark were to be removed is visualized by animating from shape $x$ to $x'$ and back. Since an animation is difficult to convey on paper, we resort to showing the difference between $x$ and $x'$, highlighted in white. All but the relevant landmark were removed in this image for clarity.

3.4.4 Change animation

Although the distance map can give an indication of how much certain regions of the shape model instance will change when a landmark is removed, it is still unclear how the shape will change. In order to determine influence $I_l$ and distance map $D_l$, the shape $x'$ (the shape after removing a landmark) needed to be determined first. This shape can be used to animate from the current shape $x$ to $x'$ and back to $x$, on the user’s request, showing how the shape will change. The animation is performed by linearly interpolating the coordinates of $x$ towards $x'$ and then back to $x$. See Figure 3.9.

Since the influence a landmark has can sometimes be too small to perceive, it is useful to be able to exaggerate the animation. We therefore introduce an animation exaggeration parameter that will make the animation extrapolate past $x'$ by a user specified amount. Note that the animation’s main goal is to indicate how the shape will change. If the user wants to see the exact change, the exaggeration should be disabled.

3.4.5 Region-Of-Interest filter

So far, landmark influence visualization indicates influence on global fit only. In order to show local effects, a region-of-interest filter is introduced. When a region $R$ is selected by the user, the filtered landmark influence values $I_l'$ are determined and visualized immediately.
3.4 Landmark Influence Visualization Method

Figure 3.10: The region-of-interest filter allows the user to see the influence of landmarks on a specific region of the shape model instance. In this case, a box widget (in the left image) is used to determine the region of interest. Any kind of region can be used however. The right image shows the same situation without the ROI filter. Note that the influence values of landmarks are different, as can be seen by the disk glyphs. (The text glyphs are unrelated, and are used in the global pose determination step of Active Shape Models, see Appendix A.2.)

using Equation (3.8):

\[ I_l = \sqrt{\sum_{i \in R} \| x'_i - x_i \|^2}, \tag{3.8} \]

where \( R \) is a vector containing the indices of the points of the shape model instance that lie within the region-of-interest. Then, the filtered landmark influence values indicate the influence landmarks have on just the specified region of interest. See Figure 3.10.
Chapter 4

Evaluation

Recall that the research questions of this work are the following:

1. Can the accuracy of reconstructed shapes using Active Shape Models be improved by visualizing Landmark Influence?
2. Can the time required to reconstruct a shape using Active Shape Models be reduced by visualizing Landmark Influence?
3. How much does Landmark Influence Visualization facilitate the usability of the Active Shape Models method?

In order to evaluate the answers to these questions, a user study has been conducted. We will first present the setup and then move on to discuss the results.

4.1 Setup

An overview of the user study setup will be presented in this section. Refer to Appendix D for the full design.

The user study was set up to measure the accuracy of shape model fits, and the time required to construct those fits. We compare the case that users do not have landmark influence visualization enabled, with the case that users do have landmark influence visualization enabled.

Datasets were gathered beforehand, which include both CT and MRI volumes of the same subject. The participants segmented a shape from each of the MRI datasets. The CT data segmentations are used as ground truth shapes, to which the MRI data segmentations of the participants can be compared.

The in-person usability study was conducted on 10 users, who were asked to segment 4 different MRI datasets using Active Shape Models. For the first 2 datasets, they could not use landmark influence visualization, and for the remaining 2 datasets they could. The order
in which datasets were presented to the users was scrambled. Six of the users were experts, meaning that they were familiar with segmentation of MRI data using Active Shape Models and were familiar with the controls.

After the participants performed their segmentation tasks, the accuracy of their MRI data segmentations was determined. We only care for the accuracy of the participant’s segmented shape \( x \) compared to the ground truth shape \( g \), not its global orientation. Only the shape itself is important in subsequent processes (such as motion simulations). Therefore, in order to measure the accuracy of the fit, we use ICP [Besl and McKay, 1992] to optimize the position and global orientation of \( x \) with the ground truth shape \( g \). Afterwards, we determine the closest-point correspondence function \( C \) of the transformed shape model instance to the ground truth shape. Finally, we determine the Root Mean Square of the euclidean distances (RMSE) between the found points. See Equation (4.1).

\[
\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{m} ||x_t - C(g)_t||^2}{m}}
\]  

Additionally, a questionnaire was conducted among the users afterwards, to determine whether or not the usability of Active Shape Models has improved when using landmark influence visualization.

4.2 Results and discussion

The accuracy of the final shape model fit the users settled on for every dataset was determined by comparing their shape to a ”ground-truth” segmentation created using CT image data of the same volume, as described in Section 4.1, yielding the RMSE. See Figure 4.1 for a boxplot representation of the results.

A special case of the Pearson correlation analysis [Pearson, 1895], point-biserial correlation analysis [Tate, 1954], was performed on the continuous variable RMSE and the dichotomous variable ’landmark influence visualization enabled/disabled’. This revealed that a correlation of landmark influence visualization and the accuracy of the segmentations is only very weak (\(|\text{correlation}| < 0.10\)) at best. Inspecting just the participants who indicated that they had great experience with Active Shape Models does not change this matter. Additionally, the \( p \)-value is > 0.20, and therefore, the null hypothesis that there is no difference cannot be rejected. A different sample group may lead to different results, and therefore we cannot say anything in regard to research question 1.

The time required for the users to perform this segmentation was measured during the user study. See Figure 4.2 for a boxplot representation of the results. Again, point-biserial correlation analysis [Tate, 1954] of the results revealed that a correlation of landmark influence visualization enablement and the time required to perform the segmentations is only very weak (\(|\text{correlation}| < 0.10\)) at best. Also, the \( p \)-value is again > 0.20, meaning
Evaluation

4.2 Results and discussion

Figure 4.1: The boxplots show the measured accuracy (vertical axis; RMSE, see Equation (4.1)) of the users’ segmentations of four MRI datasets, using Active Shape Models, with Landmark Influence Visualization disabled (left) and enabled (right).

Figure 4.2: The boxplots show the time it took users to segment a dataset (vertical axis), using Active Shape Models, with Landmark Influence Visualization disabled (left) and enabled (right), for each of the 4 datasets.
that the null hypothesis that there is no difference cannot be rejected. A larger sample group may lead to different results, however, and therefore we cannot say anything in regard to research question 2 as well.

The participants were asked to indicate their attitude towards the usefulness of the different landmark influence visualization features (see Chapter 3), on a 5-point Likert scale. They were also asked to express their opinions. The results of the Likert scale questions are presented in Figure 4.3.

It is clear that the influence visualization (see Sections 3.4.1 and 3.4.2) and animations (see Section 3.4.4) were considered the most useful of all the features. Several reasons were given for the high ratings of the influence visualization: (1) the disk glyphs help to decide what landmarks should be investigated further, (2) the disk glyphs help find outliers (which were often found to be misplaced), and (3) the disk glyphs and influence indication give a sense of the stability of the fit.

The only user who remained undecided on the usefulness of influence visualization commented that “it is better to add more landmarks than to remove them, to improve accuracy”. This might be explained by noting that this user had no difficulties placing a set of landmarks that resulted in high accuracy, and therefore did not need to inspect the landmarks after running the Active Shape Models algorithm for the first time.

The users commented on the animations that they were insightful and helped them estimate the results of their actions. One user noted that it is still difficult to see the big picture.
Seven users noted that they were unsure of the usefulness of the distance maps. They commonly stated that the influence visualization was already enough to help them in their task. The users who did feel the distance map is useful stated that it helps them understand why a landmark’s influence value is very high or very low.

The region-of-interest (ROI) filter was found to be useful by half of the users, who commonly noted that it was useful to inspect the influence on a local region of the shape, specifically for cases where that region was the only part of the fit that was not accurate. Users who did not feel that the ROI filter was useful, noted that they either did not need it or that it was difficult to specify the region.

When asked what the most valuable feature was, influence visualization and animation tied with 4 votes each, with only a single vote for the ROI filter and a single vote for the distance maps. Half of the users who chose influence visualization and animation noted that the combination especially was useful in their workflow in order to achieve an accurate fit.

All six expert users indicated that they would like to integrate the visualization in their workflow.

In conclusion, the user study did not provide significant statistics, making it impossible at this point to say anything with regards to the first two research questions. Therefore, the user study should be repeated with a larger group of users, a higher percentage of experts, more datasets and perhaps even with multiple statistical shape models. Even so, the users appreciated the visualization and were positive about its usefulness.
Chapter 5

Conclusions and future work

In this work, a novel landmark influence visualization technique for use with the Active Shape Models segmentation algorithm has been presented. Active Shape Models functions as a black box that fits a Statistical Shape Model to a set of user-defined landmarks. Landmark influence visualization provides the user with a set of visual tools, which intend to help the user gain an intuitive understanding of the effect their set of landmarks has on the model fit.

A definition of landmark influence has been proposed: the amount of change introduced in the shape model instance if a landmark were to be removed. In our implementation, the euclidean distance between the current shape model instance and the changed shape model instance was used as a metric for the amount of change.

The visualization technique has been evaluated in terms of its usefulness in segmenting shapes using Active Shape Models. The accuracy of resulting segmented shapes and the time required to perform those segmentations has also been measured in a user study. No evidence was found to support the hypothesis that landmark influence visualization improved the accuracy or decreased the time required.

A questionnaire has been used to measure the users’ attitude towards the various features of landmark influence visualization.

Disk glyphs visualize the influence landmarks have on the shape model fit. They help the user decide what landmarks require further investigation, draw attention to outliers (which are sometimes misplaced landmarks) and give a sense of the stability of the fit.

Distance maps visualize the change that the shape will undergo if a landmark were to be removed. They can also be used to understand why a landmark is more or less influential than other landmarks.

Animations visualize how a shape will change if a landmark is to be removed. Users noted that the animations helped them estimate the results of their actions, and that they provide deeper insight into the landmark’s influence.

Finally, a region-of-interest filter can be used to inspect the influence of landmarks on a specific region of the shape, rather than the shape as a whole. Users found it useful for finding out what landmarks affect a local region of the shape, specifically in cases where the remainder of the shape was already accurate.
In conclusion, although the evaluation could not show whether or not the accuracy of segmented shapes increased and whether or not the time required for segmentation decreased, it became clear that users found the visualization useful in the segmentation process.

Future work should include repeating the user study with a larger group of experts, and using more datasets, with multiple statistical shape models. Then, a conclusion might be drawn from the results on the effects of using landmark influence visualization on the accuracy of and time required for Active Shape Models segmentations.

An interesting direction for future work would include taking into account second order landmark influence, or even further levels. It may be interesting to be able to see how other landmarks’ influence values will change when a landmark is to be removed.

Something that this approach did not take into account was the global orientation that is determined during Active Shape Models. It would be beneficial for the user if this component of the shape fitting procedure would also be taken into account when determining landmark influence, making this a suitable suggestion for future work.

Another suggestion would be to enable the user to drag a landmark across the data and presenting instant feedback about the influence on the shape.

Finally, it would also be interesting to display the distribution of influence values using a color map directly on the shape.


Nomenclature

$b$ The vector containing the shape model parameters of the Statistical Shape Model.

$C$ A point correspondence function.

$D$ The collection of all landmark distance maps.

$D_l$ A landmark’s distance map.

$I$ The collection of all landmark influence values.

$I_l$ A landmark’s influence value.

$L$ The user-defined set of landmarks.

$\lambda_i$ The $i$th eigenvalue yielded by Principal Component Analysis on $X$.

$m$ The amount of points in every shape in the Statistical Shape Model.

$n$ The amount of shapes in the Statistical Shape Model.

$P$ A column-wise subset of matrix $U$. The columns of this matrix represent the main axes of the Statistical Shape Model.

$R$ A Region-Of-Interest.

$t$ The amount of columns of matrix $P$.

$U$ A matrix containing column-wise the eigenvectors yielded by Principal Component Analysis on $X$.

$x$ The $n \times 3m$ matrix containing all the coordinates of all shapes in the Statistical Shape Model.

$x$ A shape.

$\bar{x}$ The mean shape of the Statistical Shape Model.
$x_j$ The column vector containing the coordinates of shape $j$ in the Statistical Shape Model.
Appendix A

Active Shape Models for the femur

In this section Clinical Graphics’ implementation of Active Shape Models, tailored towards the femur, will be described. There are two stages to the method. First, a model needs to be constructed. Second, when the model is complete, the model can be fitted to new image data.

A.1 Building a Statistical Shape Model

In the first stage, a Statistical Shape Model is built by annotating and subsequently performing analysis on a representative training data-set. In Clinical Graphics’ case the data-set consists of annotated femur shapes.

In the 3D case, each of the $n$ shapes needs to be annotated with $m$ landmarks, or point coordinates. These landmarks need to be placed at suitable locations, more specifically, at corners and/or junctions of object boundaries and at easily recognizable biological locations. The set of landmarks can be augmented with equally spaced intermediate points along the object boundaries. See Figure A.1.

![Figure A.1: Suitable landmark positions (image from [Cootes et al., 2000]).](image-url)
In order to annotate all shapes uniformly (e.g., all shapes have the same set of landmarks and all landmarks correspond to each other across shapes), a number of steps are taken. The segmentations are binary voxel shapes, of which one is manually annotated with landmarks. For all the other shapes, a deformation is found to register the annotated sample onto the new sample, in order to find the coordinates of the same set of landmarks on this new sample.

Before continuing to the analysis step, the shapes need to be transformed to a common coordinate frame. The translation is achieved by moving the center of gravity to the origin. The rotation step is performed by applying the iterative closest point algorithm [Besl and McKay, 1992], minimizing the distance to the first sample’s shape. Scaling can optionally be performed depending on whether or not scale is a feature that needs to be accounted for in the model. In the femur model of Clinical Graphics, scaling is not performed as it is considered a feature of the model.

For every sample the landmarks can then be written as a vector \( \mathbf{x}_j = (x_{j1}, y_{j1}, z_{j1}, x_{j2}, y_{j2}, z_{j2}, \ldots, x_{jm}, y_{jm}, z_{jm}) \) where \( j = 1, \ldots, n \). Then \( X = (x_1 | x_2 | \ldots | x_n) \) is the \( n \times 3m \) shape space holding all of the landmark coordinates, where every column corresponds to a sample.

In the analysis step, the data matrix \( x \) is normalized first. Then, Singular Value Decomposition (SVDs) is performed on \( x \). The data forms a cloud of points in \( 3m \)-dimensional space. SVD computes the main axes of that cloud, enabling approximations of the original points using less than \( n \) parameters. SVD is used instead of the more commonly applied Principal Component Analysis (PCA) since it is computationally less expensive. For more details on SVD, see Appendix C.

The SVD computes three matrices \( U, \Sigma, V^* = \text{svd}(X / \sqrt{n}) \) of which the first two, \( U \) and \( \Sigma \), are of our concern. The first \( t \) columns of \( U = (u_1 | u_2 | \ldots | u_t | \ldots | u_n) \) are the main axes of the data set. The values of matrix \( \Sigma \) are all zero, except for its \( t \) diagonals \( \sigma_{i,i} = 1, 2, \ldots, t \) which are equal to the standard deviations across the data along the corresponding axes in \( U \).

Then, any shape \( x \) can be approximated using

\[
\mathbf{x} \approx \bar{x} + Pb, \tag{A.1}
\]

where \( \bar{x} \) is the mean shape’s set of landmark points, \( b \) is a \( t \)-dimensional set of parameters of the distribution, and

\[
P = (u_1 | u_2 | \ldots | u_t). \tag{A.2}
\]

The standard deviation of the \( i \)th parameter \( b_i \) across the training set is given by \( \sigma_i \). The \( i \)th parameter of \( b \) is usually referred to as the \( i \)th mode of the model. By varying the elements of \( b \) we can vary the shape \( x \) using Equation (A.1). Figure A.2 shows the effect of varying each of the first three model shape parameters in a model of faces.

The model set up by Clinical Graphics currently contains \( n = 205 \) shapes and \( m = 21096 \) landmark points per sample. The landmark connectivity data makes up for 42188 cells (triangles). It is clear that \( n < 3m \) and that the points therefore do not span the
A.2 Applying the model

In the second stage, in order to interpret a new image, an iterative process is applied. Setting up the process involves selecting an initial model instance (typically the mean shape) and finding an initial global transform, which requires user interaction to initialize the process. After the initial global pose has been determined, the process starts to alternate between optimizing the model fit to the landmarks of the new image and refining the global pose of the model, until convergence is declared.

In order to find the initial global pose, the user places landmarks on the image data at suitable positions. In order to aid the process, the user can annotate landmarks with ‘zone labels’. These labels indicate that a landmark should be matched to a point in the specified

Figure A.2: Effect of varying each of the first three model shape parameters (vertical axis) in turn between ±4 standard deviations (horizontal axis).

3$^m$-dimensional space, however, this is not a requirement for the quality of the model, since we only need to cover the subspace of femur shapes.
Figure A.3: Landmarks on the image data are annotated with ‘zone labels’ to indicate which region of the shape model they belong to. In this case a landmark has been assigned to the region ‘trochanter minor’, abbreviated to ‘tm’.

Figure A.4: Landmark zones on the upper section of the femur. The annotated landmark of Figure A.3 is assigned to the region labeled ‘trochanter minor’.
Active Shape Models for the femur

A.2 Applying the model

region of the shape model. In Figure A.3 a landmark has been assigned to the ‘trochanter minor’ region. In Figure A.4 some of the regions are visualized by means of color coded surface areas overlaid on the femur.

Now that the setup has been completed, the iterative process begins. The global pose is determined by repeatedly applying the iterative closest point (ICP) algorithm [Besl and McKay, 1992], until convergence is declared. The point correspondence that is determined in ICP is constrained by the zone annotations on the landmarks and shape model. This means that landmark that have a zone annotation can only be matched to points in the shape model that are in the specified zone.

After the global pose has been updated, the second part of the iteration begins. The model fit to the landmarks of the femur is optimized. This is done in four steps.

First, for every landmark, the closest point on the model is determined by locating the nearest cell in the SSM and selecting the closest point of that cell. This point correspondence determines which rows from are selected from $P$ and the mean shape $\bar{x}$ in the subsequent steps. It should be noted that in general, although especially for concave regions, this may not always return the correct point correspondence. This depends greatly on the set of landmarks as a whole, the statistical shape model that is being used, and the global orientation is determined.

A least squares solver then minimizes $||Pb - (x - \bar{x})||$ in the equation $Pb = x - \bar{x}$, where $x$ is the set of manually placed landmark points. This determines the model parameters $b$ that are applied to Equation (A.1) to update the model instance $x$. The number of modes $u$ with $u \leq t$ (see Equation (A.2)) that are involved in the calculation is configurable (by default $u = 15$). The solver is optimal in the least squares sense. The reason a least squares solves is used instead of calculating $b$ directly is that in order to do so $P$ would have to be invertible, as in $b = P^{-1}(x - \bar{x})$. However, $P$ is not invertible, and therefore no exact solution exists.

The parameters are then each clamped to the range $(-v\sigma, v\sigma)$, where $v$ is a configurable parameter and $\sigma$ is the standard deviation (by default $v = 4.5$).

Finally, the model instance is updated with the parameters $b$ using Equation (A.1).

These iterative process is repeated $k$ times, where $k$ is a configurable parameter (by default $k = 20$). All intermediate results are stored, and finally, the result with the lowest error is selected to be the best match.
Appendix B

Principal Component Analysis

A key step in creating a Statistical Shape Model involves Principal Component Analysis (PCA), a statistical method that analyzes a set of observations of possibly correlated variables and transforms them into a new set of variables called principal components. The principal components are linearly uncorrelated and mutually orthogonal vectors that explain the variance in the data.

A 2-dimensional example set of observations is presented in Figure B.1. The arrows indicate the principal components found by applying PCA. The procedure is defined such that every component accounts for the largest possible amount of variance in the data, and each following component has the highest variance possible while being orthogonal to all

Figure B.1: The principal components of a 2-dimensional set of observations drawn from a normal distribution.
Principal Component Analysis example 2

Figure B.2: The first two principal components of a 3-dimensional set of observations drawn from a normal distribution.

previous components. In Figure B.2 a 3-dimensional example set is presented. It should be clear that the PCA approach can be extended to any number of dimensions $n$.

$$T = XW$$  \hspace{1cm} (B.1)

As in Equation (B.1), PCA is defined as a transformation $T$ that maps the original data $X$ to the coordinate system defined by the principal components of $X$, $W$.

The reason that PCA is essential to Active Shape Models is that it allows us to perform dimensionality reduction on the Statistical Shape Model, speeding up calculations significantly. Equation (B.1) can be adjusted to

$$T_t = XW_t$$ \hspace{1cm} (B.2)

to include only the first $t$ principal components. By construction, this minimizes the total squared reconstruction error $\|X - X_t\|_2^2$ and maximizes the variance that is preserved from the original data.
Appendix C

Singular Value Decomposition

Principal Component Analysis is computationally expensive. A more computationally-efficient alternative is Singular Value Decomposition, a procedure that yields similar analysis results to PCA. The procedure’s output is sufficient for the input the Active Shape Models algorithm requires.

The Singular Value Decomposition (SVD) of a matrix $X$ of size $m \times n$ is a factorization of the form

$$X = U \Sigma V^*$$  \hspace{1cm} (C.1)

where $U$ and $V$ are $m \times m$ and $n \times n$ orthonormal matrices and $\Sigma$ is an $m \times n$ diagonal matrix with the non-negative singular values $\sigma_i, i = 1, \ldots, \min(m, n)$ in decreasing order along the diagonal. The columns of $U$ and $V$ are denoted by the vectors $u_i, i = 1, \ldots, m$ and $v_i, i = 1, \ldots, n$.

Applying this to Active Shape Models, when the SVD is calculated of the matrix

$$X = \left( \frac{x_1 | x_2 | \ldots | x_n}{\sqrt{n}} \right)$$  \hspace{1cm} (C.2)

where $x_i, i = 1, \ldots, n$ are column vectors of height $m$ representing data points, the directions of maximum variation (a.k.a. the principal components, see Appendix B) are given by the first $t = \min(m, n)$ columns of $U$ and the variances by the squares of the singular values in $\Sigma$.

For more details, we refer the reader to the excellent document on SVD by Muller et al. [2004].
User Study Design

Goals of the study

This study needs to be performed in order to evaluate the landmark influence visualization technique’s value to end-users. The study will provide the author with behavioral observations and insight into the user experience. The observations will also allow for a comparison of the user’s performance with/without the aforementioned technique.

Research questions and issues

The intent of this user study is to collect observations with which the following research questions can be answered:

- Can the accuracy of reconstructed shapes using Active Shape Models be improved by visualizing Landmark Influence?
- Can the time required to reconstruct a shape using Active Shape Models be reduced by visualizing Landmark Influence?
- How much does Landmark Influence visualization facilitate the usability of the Active Shape Models method?

We will also collect more general feedback from the users.

Description of the target audience and key user behaviours we’ll recruit for

Users of the method want to segment shape(s) from volume(s), typically acquired using a CT or MRI scanner. They commonly require the segmented shape for other purposes, e.g. simulations. Segmentation is typically not performed by clinicians, but instead by engineers with in-depth knowledge of segmentation methods. Therefore, we identify researchers and engineers from the following fields as the target audience:
User Study Design

- medical visualization
- bioinformatics
- computer graphics

In the Clinical Graphics’ office there are currently 2 expert users, who use the software frequently (ranging from daily to weekly) for the aforementioned purpose.
At TU Delft and Yes!Delft there are a number of researchers (up to 10 domain experts) from the aforementioned fields who might be available as participants as well.

Test method

We will conduct an in-person usability study. The study will collect information such as time on task, insights, overall satisfaction, areas of concern, and unmet needs.

Project schedule and timeline

A pilot test will have to be run in order to work out how long it will take a single user to perform the tasks and to fill in the questionnaire. The practice task will probably take about 35 minutes and the actual tasks will take about 40 minutes. The questionnaire shouldn’t take longer than 10 minutes.

Tasks we’ll ask users to complete

- Practice task

In order to let the user get comfortable with the (software) environment he has to perform some practice tasks. These steps have to be performed on a single (practice) MRI dataset.

Once without the landmark influence visualization:

1. Annotate the femur in the MRI data with landmarks
2. Execute the Active Shape Models algorithm
3. Visually inspect the shape for errors
4. Repeat until satisfied with the result

And once with the landmark influence visualization (LIV):

1. Annotate the femur in the MRI data with landmarks, using LIV
2. Execute the Active Shape Models algorithm
3. Visually inspect the shape for errors
4. Repeat until satisfied with the result
- Actual tasks

We assume that the user is now comfortable with performing the tasks. The following steps have to be performed on four MRI datasets (different from the practice dataset), this time without getting feedback about their accuracy. The user is encouraged to achieve a high accuracy fit. The order in which datasets are selected for the actual tasks will be randomized over the different participants. Twice without the landmark influence visualization:

1. Annotate the femur in the MRI data with landmarks
2. Execute the Active Shape Models algorithm
3. Visually inspect the shape for errors
4. Repeat until satisfied with the result

And twice with the landmark influence visualization (LIV):

1. Annotate the femur in the MRI data with landmarks, using LIV
2. Execute the Active Shape Models algorithm
3. Visually inspect the shape for errors
4. Repeat until satisfied with the result

Questionnaire

Afterwards, the user will have to fill in a questionnaire about the technique after the user study. See Appendix E.

Data we’ll collect

The questionnaires and two metrics:

- Time on task
- Accuracy of segmented shapes [compared to the CT-based segmentation]
## Questionnaire

**Landmark Influence Visualization User Study**

Please answer the following questions.

**Age:**

**Gender:** male / female

**Current field:**

How much experience do you have in working with medical data in general (e.g. CT, MRI, X-ray)?
none / a little / moderate / above average / a lot

How much experience do you have in segmenting shapes from CT data?
none / a little / moderate / above average / a lot

How much experience do you have in segmenting shapes from MRI data?
none / a little / moderate / above average / a lot

How much experience do you have in working with Active Shape Models?
none / a little / moderate / above average / a lot

Please indicate your answer with an X, and use the white rows to explain your answers.

<table>
<thead>
<tr>
<th>I found the landmark influence visualization useful.</th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Don’t know</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Why?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| I found the distance maps useful.                |                   |         |            |       |                |
| Why?                                              |                   |         |            |       |                |

| I found the region-of-interest filter useful.    |                   |         |            |       |                |
| Why?                                              |                   |         |            |       |                |

| I found the landmark influence animations useful. |                   |         |            |       |                |
| Why?                                              |                   |         |            |       |                |

What was the most valuable feature, why?

Do you have any suggestions?

What features would you use in your work environment, why?