A Fully Automated Segmentation of Knee Bones and Cartilage Using Shape Context and Active Shape Models

Behnaz Pirzamanbin

May 16, 2012

Abstract

In this master's thesis a fully automated method is presented for segmenting bones and cartilage in magnetic resonance imaging (MRI) of the knee. The knee joint is the most complex joint in the human body and supports the weight of the whole body. This complexity and acute task of the knee joint leads to a disabling disease called Osteoarthritis among the adult population. The disease leads to loss of cartilage and torn cartilage cannot be repaired unless surgical techniques are used. Therefore, one of the important parts of finding the disease and planning the knee surgery is to segment bones and cartilages in MRI.

The segmentation method is based on Statistical Shape Model (SSM) and Active Shape Model (ASM) built from a MICCAI 2010 Grand challenge training database. First, all the data are represented by points and faces. A Shape context algorithm is applied on 60 data sets to obtain consistent landmarks. The mentioned consistent landmarks and Principal Component Analysis are used to build a Statistical Shape Model. The resulting model is used to automatically segment femur and tibia bones and femur and tibia cartilages with Active Shape model. The algorithm is tested on the remaining 40 MRI data sets provided by the Grand challenge 2010, and compared with six other submitted papers.

Acknowledgments

First and foremost I offer my sincerest gratitude to my supervisors, Petter Strandmark and Johannes Ulen, who have supported me throughout my thesis with their patience and knowledge. I attribute the level of my Master's degree to their encouragement and effort.

I would like to thank Sebastian Hanner for his useful comments on this report. And my special thanks to Martin Millnert for all the scientific discussions and encouragements during all the procedures of my master thesis.

Finally, I would like to thank my parents, Hamideh and Anoosh, for their support throughout all my studies at University and my entire life.

Contents

1 Introduction

Osteoarthritis (OA) is one of the major diseases and causes of disability in developed countries and affects mostly the adult population [1, 3]. Knee bone and cartilage are the main parts of the body which are severely affected by OA. OA leads to loss of articular cartilage and pain, see Figure 1. This happens because one of the main responsibilities of the knee is to support the whole weight of the human body.

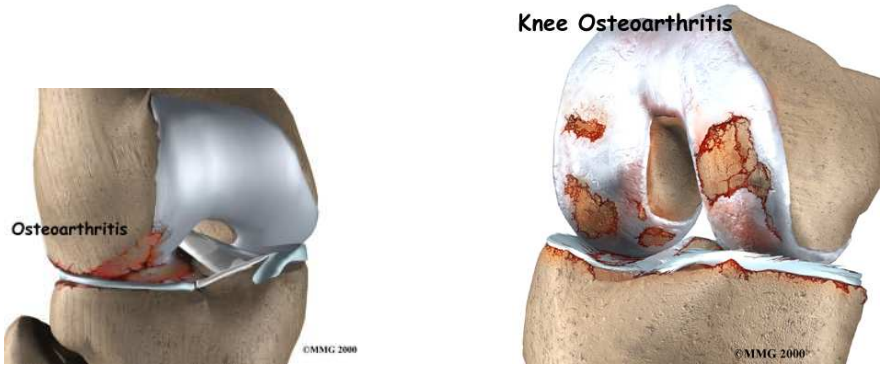


Figure 1: Osteoarthritis in knee joint [6]. Both figures show the knee joint with femur and tibia bone and their cartilage. On the right side the reader can clearly see the tear in the femur cartilage.

The knee joint is the largest and most complicated joint in the body. It joins the thigh with the lower leg and has two articulations. One of these articulations is between the femur and tibia and the other one is between the patella and femur, see Figure 3. Wide variety of the movement such as, flexion, extension and medial and lateral rotation can develop acute injury and OA [5].

The main part of the knee joint consists of two bones. The upper bone is Femur and the lower bone is Tibia (Figure 2). The femur bone is the longest and strongest bone of the human body. Furthermore, the tibia bone supports most of the body weight and is an important part of both knee and ankle joints [5].

Articular cartilage is a white, elastic and smooth tissue that covers the end of the bone in the joint, see Figure 3. It makes the bones move and slide easily over each other. The movement of the joint over years can wear off the cartilage and because of the lack of self-restoration and self-healing, damaged cartilage needs to be repaired by surgical techniques [5].

Several studies have proved the effectiveness of magnetic resonance imaging (MRI) in finding the disease and planning better surgical operations [3]. MRI is a medical imaging technique that uses a magnetic field and pulses of radio frequency waves to visualize the detailed structure of the knee. MRI gives information with good contrast between the different tissues and organs, such as cartilage, ligaments and muscles. The contrast is better than computed tomography (CT) scan, ultra sound and X-ray. Another benefit is that MRI uses no ionizing radiation which makes the procedure safe. The human body consists of $70 - 80\%$ water. When a knee is inside the MRI scanner, a strong magnetic field makes the hydrogen inside water molecules align with the direction of the

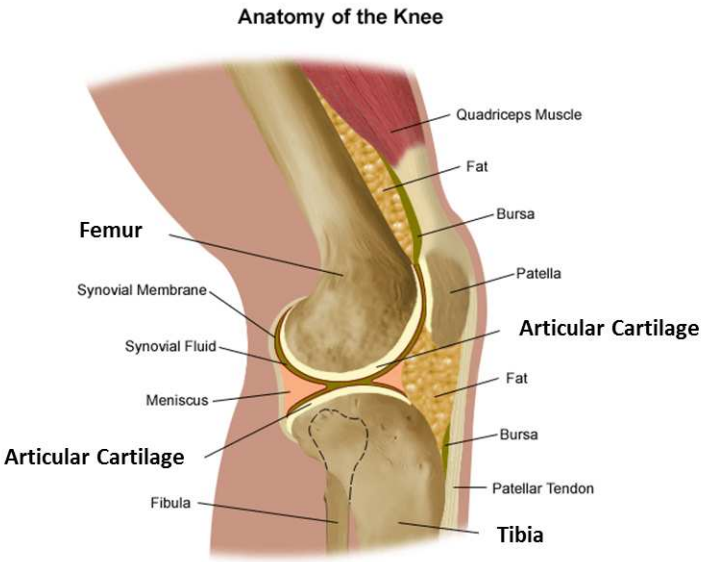


Figure 2: Anatomy of a knee, bones, cartilage, soft tissues etc. The upper bone of the knee joint is Femur. The lower part of knee contains two bones Tibia and fibula. Tibia supports the weight of the body.

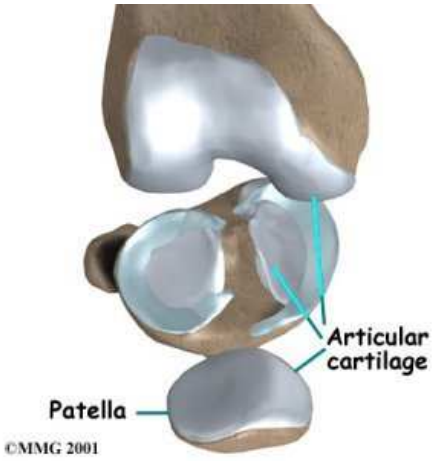


Figure 3: Main bones and articular cartilage of the knee joint [5].

field. A radio transmitter produces a varying electromagnetic field of about $1 - 10$T depending on the scanner. Then the photons of this field which have the right energy or resonance frequency (RF) can be absorbed by aligned water hydrogen protons of a body and flip their spin. These changes make hydrogen protons release different energy as photons when the field is removed. These photons can be detected by the scanner as an electromagnetic signal like radio waves.

An MRI image can be made because the photons in different tissues return to their equilibrium state at different rates. Spin density, T1 and T2 relaxation times, flow and spectral shifts are five different tissue variables which can construct the images. By changing the setting of these variables in the scanner, it can create contrast between different types of body tissue.

T1 and T2 relaxation time measure how long time it takes for a tissue to get back to equilibrium after the RF signal. The basic MRI scans are T1-weighted and T2-weighted scans which in particular differentiates fat from water. On T1 weighted scans, tissues containing water and fluid are darker and fat containing tissues are brighter and they use a short echo time (TE) and short repetition time (TR). TE and TR are basic parameters of image acquisition, TE is the time when the signal is detected and TR is the repetition time. In contrast with T1-weighted images are the T2-weighted images whose watery parts are bright and fat is dark and also they use a long TE and long TR [7],[8].

The MRI scanner creates 3D objects by producing several 2D slices and adding them together. These slices are defined relative to the human body with three main directions; sagittal, coronal and transverse (Figure 4).

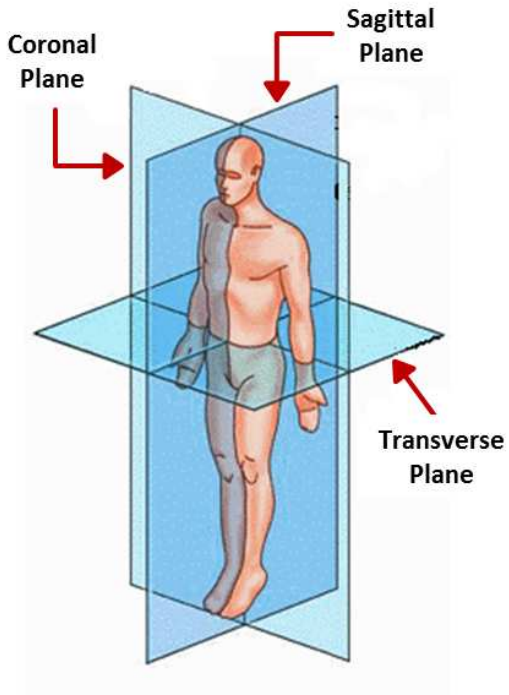


Figure 4: Human body plane and three main cuts of the body, sagittal, coronal and transverse.

Segmentation of bones and cartilage in the MRI image is one of the important parts of finding the disease and planning the knee implants or surgery [1]. The purpose of this master's thesis is to automatically segment the bones and cartilages of the knee joint. The data set is from Segmentation of Knee Images 2010 (SKI10). Ski10 is one of the three competitions of the Grand Challenge Workshop 2010, organized in conjunction with the MICCAI 2010 conference. It is part of a series of challenges for segmentation and registration of medical image data [1]. Different results are submitted to the website which makes a comparison of the different results from different models on the same data set. One of the problems of medical image processing is that people usually apply their models on different data sets so making the models sensitive to the changes of the data set so then, it becomes hard to compare the methods.

In Table 1 on pages 9 and 10, the summary of the six submitted papers are given, sorted by top ranked results. The results contain Average Bone score, Average Cartilage score and Average Total score. These different measures are computed using software from the challenge organizer [1].

The Avg Bone Score is computed from the average (AvgD) and the root mean square (RMSD) surface distances. To calculate these distances, for each boundary voxel in the segmentation the closest boundary voxel in the reference image is searched and the Euclidean distance between them is computed. The average of all the Euclidean distances is the average surface distance. Moreover, the root mean square surface distance is calculated in the same way but at the end instead of taking the average, first all the distances are squared and then the root is taken of the average value [1]. Values close to zero are perfect segmentations and yield good scores close to 100.

The Avg Cart Score is computed from the volumetric overlap error (VOE) and volumetric difference (VD). The volumetric overlap error is calculated between the set of voxels of the segmentation S and the reference R as $VOE = 100(1 - |S \cap R|/|S \cup R|)$, and the volumetric difference is calculated between R and S as, $VD = 100(|S| - |R|)/|R|$. These are measured in percent and values close to 0% are of perfect segmentation. Finally the Avg Total Score is the average of the bone and cartilage score [1].

Previous work				
RMSD Bone	**AvgD Bone**	**VOE Cart**	**VD Cart**	**Method**
1) Fully Automatic Segmentation of the Knee Joint using Active Appearance Models [12]				
1.35	0.81	35.45	-17.35	• Active Appearance Models • Minimum Description Length Groupwise Image Registration method • Hierarchical modeling scheme
Summary	Submitted as the last paper. But they didn't use the training data set from Grand Challenge (GC). They used the Osteoarthritis Initiative (OAI) data set and applied the model on test data. They argue that the short shaft in OAI model did not segment all shaft in GC. Hence they applied an additional step which used long shaft shape model.			
2) Model-based Auto-Segmentation of Knee Bones and Cartilage in MRI Data [13]				
1.39	0.93	31.6	2.5	• Statistical Shape Models • Graph based optimization • Multi object technique using generalized Hough Transform
Summary	Their model has a good conceptual system. They start with bones segmentation and find the bone surface and used this information plus cartilage thickness map to segment the cartilage.			
3) Hierarchical Decision Framework with a Priori Shape Models for Knee Joint Cartilage Segmentation - MICCAI Grand Challenge [14]				
3.67	2.13	32.35	1.85	• Optimal graph phase • Pattern recognition functionality • A detection using Haar Wavelet features and AdaBoost classifier
Summary	This method detects patella in the search procedure of bone. They first segment bone-cartilage as a single object and then they split the regions with different properties such as texture difference and thickness, identifying cross-object interacting regions.			

<table>
<tr><td colspan="5">4) Learning Local Shape and Appearance for Segmentation of Knee Cartilage in 3D MRI [15]</td></tr>
<tr><td>1.98</td><td>1.07</td><td>32.3</td><td>6.15</td><td>• Local shape and appearance
• Using branch-and-mincut
• Bone-cartilage interface classification using Markov random field</td></tr>
<tr><td>Summary</td><td colspan="4">Start with bone segmentation follow by classifying Bone-cartilage interface and finally cartilage segmentation using region of interest patches and probability of cartilage's properties, shape and appearance and boundary.</td></tr>
<tr><td colspan="5">5) Fully Automated Segmentation of the Knee using Local Deformation-Model Fitting [16]</td></tr>
<tr><td>2.66</td><td>1.56</td><td>67</td><td>-33.1</td><td>• Statistical deformation models
• Localization fitting of deformation models</td></tr>
<tr><td>Summary</td><td colspan="4">In this paper, they build the deformation model from the training image. And they use the reference image and perform a deformation model fitting.</td></tr>
<tr><td colspan="5">6) On Nonparametric Markov Random Field Estimation for Fast Automatic Segmentation of MRI Knee Data [17]</td></tr>
<tr><td>4.35</td><td>2.96</td><td>65.45</td><td>9.1</td><td>• Nonparametric Markov random field model
• Monte Carlo simulation
• Variational inference
• Using atlas based prior information</td></tr>
<tr><td>Summary</td><td colspan="4">They used the global statistical model and intensity values of features, every procedure of this method is fully automated but the results are not good in comparison with the other methods.</td></tr>
</table>

Table 1: Comparison of previous works in Grand Challenge. These papers work on different methods but on the same data set from [1]. The table shows the results of each method in the way that the best result comes first and therefore the worst is at the end.

For more details see [1].
After the introduction, in the second chapter different information about the

data set are given and the first step of preparing the data set using isosorface is explained. The third chapter contains the shape context algorithm which provides the consistent landmarks from the different parts of the knee in all the MRI images.

Chapter four consists of main part, segmentation of the knee bones and cartilages. In this chapter the reader can understand the theoretical and practical parts of the model, Statistical Shape Models (SSM) and Active Shape Models (ASM). In Chapter five, the application of the algorithm and the results provided and the results compared to the previous works. Finally, the discussion about the results and some new ideas to improve the model are given in Chapter six. Figure 5 shows the conceptual model of different chapters.

Conceptual model:

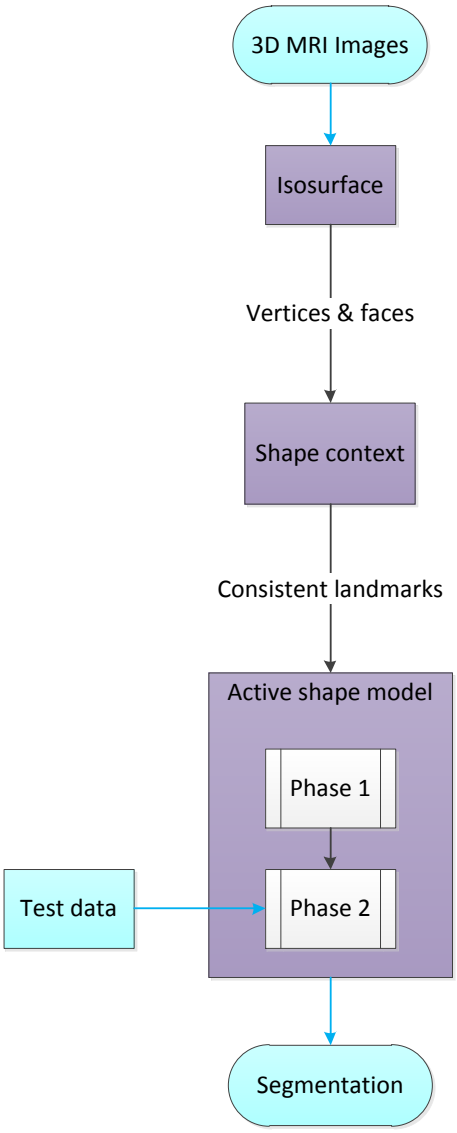


Figure 5: Conceptual model of this master's thesis method.

2 Data

The full data set contains 250 MRI images, training data and test data which are available in [1]. The data originate from the surgical planning program of Biomet, Inc. and was obtained from over 80 different centers in the USA, using machines from General Electric, Siemens, Philips, Toshiba, and Hitachi [1]. All these data are available in raw image data files along with an ASCII readable MHD header which is Insight Segmentation and Registration Toolkit (ITK) compatible [1]. ITK is an open source system which provides many different software tools for image analysis [9]. The mhd format can be read using different software such as, MITK-3M3 [10] and Amira [11]. Left and right knees are equally distributed in the data set.

The dimensions of each file and more information about the images are stored in the header file same as following example,

ObjectType	=	Image
NDims	=	3
BinaryData	=	True
BinaryDataByteOrderMSB	=	False
CompressedData	=	False
TransformMatrix	=	1 0 0 0 1 0 0 0 1
Offset	=	44.1378 33.5916 45.1903
CenterOfRotation	=	0 0 0
AnatomicalOrientation	=	RAI
ElementSpacing	=	0.3906 0.3906 1
DimSize	=	302 335 104
ElementType	=	MET-SHORT
ElementDataFile	=	image-002.raw

The image data, voxels, are three dimensional pixels and in x, y and z order. The segmentations are multi-label images of unsigned char pixel type as following:

0	=	background,
1	=	femur bone,
2	=	femur cartilage,
3	=	tibia bone,
4	=	tibia cartilage.

All images were obtained in the sagittal plane with a pixel spacing of 0.4×0.4mm and a slice distance of 1mm without contrast agents. About 90% of these images captured in machines with field strength varying between $1 - 3$T. The vast majority of images used T1-weighting, but some used T2-weighting. The model in this master's thesis is based on the T1 weighting images. These images were segmented by Biomet, Inc. experts. They selected the 100 most accurate segmentations from the set of 250 images. The available data set of 100 images were randomly divided into a training set of 60 images and a test set of 40 images. For more information check [1].

An example for a knee MRI from the data set is shown in Figure 6.

By studying the data sets one faces a problem with the intensity range. The intensity range vary between $0 - 250$ and $0 - 4500$ in different images. The reason is the use of different scanners. This problem leads to inaccurate calculation for optimal movement which is based on the intensity appearance model. Therefore, when the algorithm tries to find the new movement from big range intensity in comparison with the appearance model it is confused because the edge profiles are different and hence it loses track of the optimal movement. To solve this, each data set is normalized to a reference. The reference is chosen to be the median intensity of the femur bone which contains the most amounts of voxels in comparison to other parts of the knee.

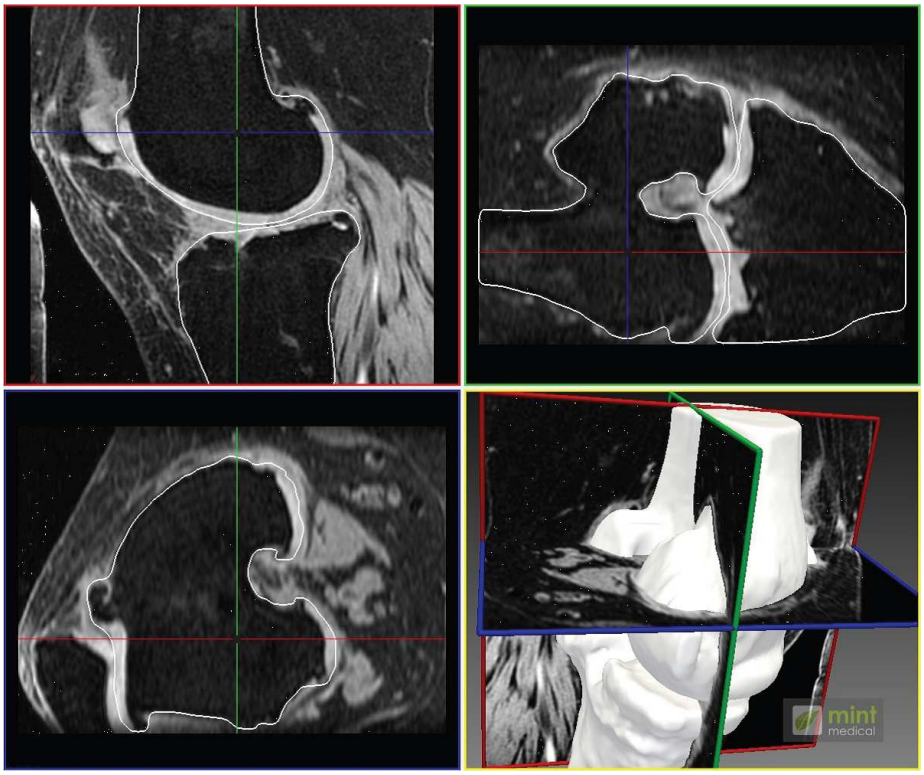


Figure 6: Left top is a sagittal view, right top is coronal view, left bottom transverse view and the last is 3D image of a knee. The white shape is the segmentation of the bones and cartilage [10].

Moreover, each image consists of two parts, a binary mask and an intensity matrix. In this master's thesis, the model is based on the surface of the shape. Therefore one needs to extract the surface from the binary image. Isosurface is used to extract the surface from different parts of the knee (tibia cartilage and bone, femur cartilage and bone).

A computed isosurface will give the vertices and the faces of each shape in the data set. The function used to compute isosurfaces connects points that have identical values same as contour lines which connect the points of equal elevation [28]. In other words, after using the isosurface function on a binary volume, the surface of the binary shape will be represented by triangles also

known as faces. Each of these triangles are made by connecting three vertices. This means the volumetric data set is reduced to three dimensional matrices of vertices and faces. In each case, an isosurface was computed on a data set with two different scalings in the image size, $(1/2, 1/4)$ to end up with a proper amount of points on the surface of different objects. These two different scaling will be used in iterative shape context in two levels coarse to fine. The shape context provides consistent landmarks of different shapes.

3 Shape context

The idea behind the shape context is to find the similarity of two different shapes. For each point p_i on the model shape, shape context algorithm finds the best match of that point to q_j on the target shape. Shape context algorithm considers the set of vectors from a point to all other points on the shape. These vectors correspond to the relation of the whole shape to the reference point. In this case, for a point p_i a coarse histogram h_i given in (1) is computed. Shape context, h_i , contains the relative coordinate of the $n-1$ points to p_i:

$$h_i(k) = \#\left\{q \neq p_i : (q - p_i) \in bin(k)\right\}, \tag{1}$$

where bins are uniform in log scale. These bins make a clear difference between closer points and farther ones. As shape contexts' distribution are histograms, the cost of matching two points is denoted with cost matrix, using the χ^2 test statistic. The χ^2 test statistic gives a comparison between expected and observed values [29],

$$\sum_{k=1}^{K} \frac{(\mathrm{O}_k - \mathrm{E}_k)^2}{\mathrm{E}_k}$$

in this case the observation is $h_i(k)$ and the expected value is $\frac{h_i(k)+h_j(k)}{2}$ which leads to

$$C_{ij} = C(p_i, q_j) = \frac{1}{2}\sum_{k=1}^{K} \frac{(h_i(k) - h_j(k))^2}{(h_i(k) + h_j(k))}, \tag{2}$$

where $h_i(k)$ is the K-bin normalized histogram at p_i.

The cost matrix can include additional terms such as, tangent angle difference, texture difference, and etc. based on the appearance similarity at two different points.

The main goal of the shape context algorithm is to minimize the total cost of matching matrix between all pairs of points of two shapes. The matching should be one-to-one. Total cost of matching points p_i in model shape to points q_j of target shape is denoted by

$$H(\pi) = \sum_i C(p_i, q_{\pi(i)}) \tag{3}$$

where the $q_{\pi(i)}$ is the permutation of q_j after applying the constraint of being one-to-one. This problem can be solved with Hungarian method in $O(n^3)$ [23]. A simple example of matrix form using Hungarian method is given in the following, for more details on the method and a proof of time complexity of the algorithm see [24].

3.1 Example: Hungarian method

Consider five different jobs $1, 2, 3, 4$ and 5 and five different contractors A, B, C, D and E, the cost of doing each of these jobs by each of the contractors

are given in the matrix below,

$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 7 & 2 & 1 & 9 & 4 \\ 9 & 6 & 9 & 5 & 5 \\ 3 & 8 & 3 & 1 & 8 \\ 7 & 9 & 4 & 2 & 2 \\ 8 & 4 & 7 & 4 & 8 \end{pmatrix} \end{array}$$

The task is to find the total lowest cost of doing all the jobs when each contractor does only one job. In the first step, the lowest cost in each row is subtracted from the rest in that row. Hence, there will be at least one zero in each row,

$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 6 & 1 & \mathbf{0} & 8 & 3 \\ 4 & 1 & 4 & 0 & 0 \\ 2 & 7 & 2 & \mathbf{0} & 7 \\ 5 & 7 & 2 & 0 & 0 \\ 4 & \mathbf{0} & 3 & 0 & 4 \end{pmatrix} \end{array}$$

The bold zeros show which contractors do only one job with lowest cost. However, it seems there is no contractor for job 1 and at the same time two contractors B and D do job 5 with same cost. To solve this, in the second step the lowest cost in each column subtracted from the rest cost in that column.

$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 4 & 1 & \mathbf{0} & 8 & 3 \\ 2 & 1 & 4 & 0 & \mathbf{0} \\ \mathbf{0} & 7 & 2 & 0 & 7 \\ 3 & 7 & 2 & \mathbf{0} & 0 \\ 2 & \mathbf{0} & 3 & 0 & 4 \end{pmatrix} \end{array}$$

Now each job can be done by one contractor and the total lowest cost is 15.

Summary of shape context:

1. Calculate the distance of a point to all other points.
2. Normalize the distance by mean distance.
3. Create logarithmic distance scale for normalized distances.
4. Create distance histogram by iterating for each bin, $\left[\frac{1}{8}, \frac{1}{4}, \frac{1}{2}, 1, 2\right]$ as in equation 1.
5. Calculate angle between all points.
6. Bin angles.
7. Matching - Cost Matrix: Calculate cost of matching each point to every other point as in equation 2.
8. Matching - Additional Cost Terms:
    - Surrounding Texture Difference,

- Tangent Angle Difference.

9. Matching: Find pairing of points that leads to least total cost according to equation 3.
    - Hungarian Method $\left(\mathcal{O}(n^3)\right)$.

3.2 Example: Shape context

Considering the following happy face with 6 landmarks, the aim is to find the corresponding landmarks in the target image.

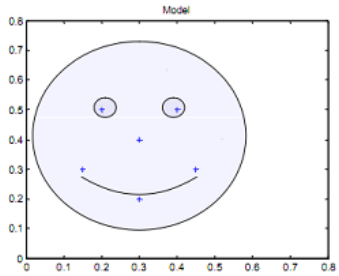


Figure 7: Model

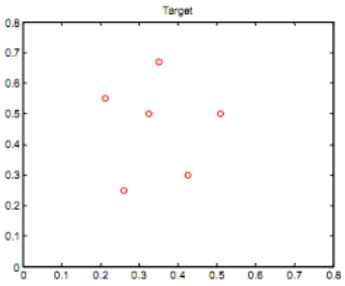


Figure 8: Target

Given the coordinate of the landmarks, the distance between all points is calculated, see Table 2.

Coordinates on shape: **Compute Euclidean distance from each point to all others:**

$$\begin{array}{ll}(1)\\(2)\\(3)\\(4)\\(5)\\(6)\end{array}\begin{pmatrix}0.2 & 0.5\\0.4 & 0.5\\0.3 & 0.4\\0.15 & 0.3\\0.3 & 0.2\\0.45 & 0.3\end{pmatrix}$$

$$\begin{pmatrix}0 & 0.2000 & 0.1414 & 0.2062 & 0.3162 & 0.3202\\0.2000 & 0 & 0.1414 & 0.3202 & 0.3162 & 0.2062\\0.1414 & 0.1414 & 0 & 0.1803 & 0.2000 & 0.1803\\0.2062 & 0.3202 & 0.1803 & 0 & 0.1803 & 0.3000\\0.3162 & 0.3162 & 0.2000 & 0.1803 & 0 & 0.1803\\0.3202 & 0.2062 & 0.1803 & 0.3000 & 0.1803 & 0\end{pmatrix}$$

Table 2: Coordinate and distances between the landmarks.

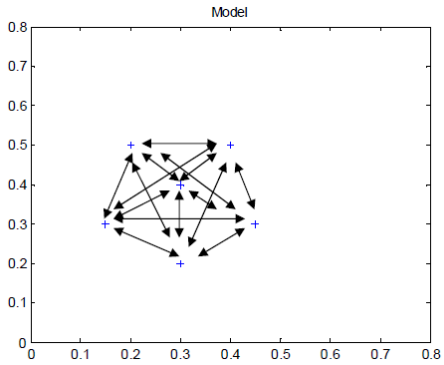


Figure 9: The distance between each point to other landmarks is shown with the double arrow. The results are in Table 2.

Second step is to normalize distances between each point and create log distance scale for normalized distances,

$$0.1250\ 0.2500\ 0.5000\ 1.0000\ 2.0000$$

Right side of Table 3 shows the log distance scale of the normalized distances of the left side where bins with higher numbers describe points closer together.

$$\begin{pmatrix} 0 & 1.0623 & 0.7511 & 1.0949 & 1.6796 & 1.7004 \\ 1.0623 & 0 & 0.7511 & 1.7004 & 1.6796 & 1.0949 \\ 0.7511 & 0.7511 & 0 & 0.9575 & 1.0623 & 0.9575 \\ 1.0949 & 1.7004 & 0.9575 & 0 & 0.9575 & 1.5934 \\ 1.6796 & 1.6796 & 1.0623 & 0.9575 & 0 & 0.9575 \\ 1.7004 & 1.0949 & 0.9575 & 1.5934 & 0.9575 & 0 \end{pmatrix} \quad \begin{pmatrix} 5 & 1 & 2 & 1 & 1 & 1 \\ 1 & 5 & 2 & 1 & 1 & 1 \\ 2 & 2 & 5 & 2 & 1 & 2 \\ 1 & 1 & 2 & 5 & 2 & 1 \\ 1 & 1 & 1 & 2 & 5 & 2 \\ 1 & 1 & 2 & 1 & 2 & 5 \end{pmatrix}$$

Table 3: Difference distance matrix on the left and bins matrix of log distance scale on the right.

Compute angle between all points (0 *to* 2π):

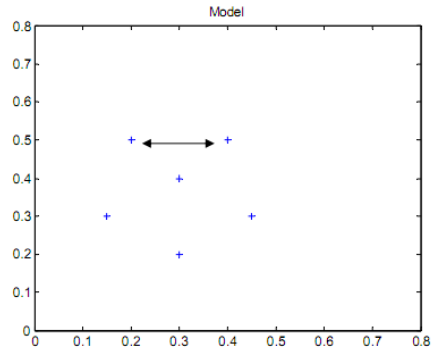


Figure 10: The angles between point one and two is shown with double arrow.

Computing the angle scale can be done in different ways, in this example the following equation is used to find the differences,

$$Scale = 1 + floor(angle/(2 * pi/n - bins)).$$

0	***0***	5.4978	4.4674	5.0341	5.6084		1	1	6	5	5	6
3.1416	0	3.9270	3.8163	4.3906	4.9574		4	1	4	4	5	5
2.3562	0.7854	0	3.7296	4.7124	5.6952		3	1	1	4	5	6
1.3258	0.6747	0.5880	0	5.6952	0		2	1	1	1	6	1
1.8925	1.2490	1.5708	2.5536	0	0.5880		2	2	2	3	1	1
2.4669	1.8158	2.5536	3.1416	3.7296	0		3	2	3	4	4	1

Table 4: Angles left and bins corresponding to angles on right. The gray italic numbers are the angles between point one and two and other way around.

Next step is to find the cost matrix. After applying the Hungarian method to pair points to find the least total cost, the following pairing in Figure 11 is suggested [25].

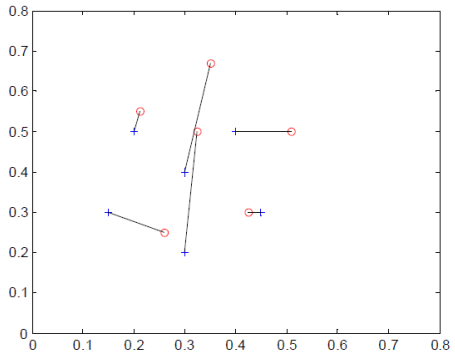


Figure 11: The suggested pairing is shown with lines between the points from model to target. These lines are according to minimum cost.

The iterative shape context provides consistent landmarks between the model shape and the target shapes. This gives same numbers of vertices in all shapes which can be fitted by the faces from the model. In the following chapters, it will be explained how these consistent landmarks can make a shape model by means of principal component analysis.

4 Segmentation

4.1 PCA

Principal component analysis (PCA) is a mathematical method that reduces the dimensionality of the variables. A principal component (PC) is the direction of the variance of the variables. The PCs are ordered by the highest value. The first PC accounts for the most variability in the data set and the rest accounts for the remaining variability. By this the less important PCs can be removed from the variables.

The main tasks of PCA are:

1. To discover or to reduce the dimensionality of the data set.
2. To identify the most meaningful directions of the variance of the variables [20].

Principal component analysis is based on the statistical representation of a m-element random vector $\mathbf{x}$

$$\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)^T, \tag{4}$$

where $\mathbf{x}_i,\ i = 1, \ldots, n$ are n observations. Hence, the vector's mean is denoted by

$$\mu_j = \frac{1}{n} \sum_{i=1}^{n} x_{i,j}, \quad j = 1, \cdots, m \tag{5}$$

and the (l, k)th element of the covariance matrix by

$$\mathbf{C}_{l,k} = \frac{1}{n-1} \sum_{i=1}^{n} \left(x_{l,i} - \mu_l\right) \left(x_{k,i} - \mu_k\right), \quad l, k = 1, \cdots, m \tag{6}$$

or by considering matrix $\mathbf{B} = \mathbf{x} - \mu$,

$$\mathbf{C} = \frac{1}{n-1} \mathbf{B}\mathbf{B}^T. \tag{7}$$

PCs are the eigenvectors of a symmetric covariance matrix [21].The corresponding eigenvectors, PCs, obtain by sorting the eigenvalues in descending order from the largest value to the lowest. Hence the first PC is the direction of the largest eigenvalue or the largest variance of the data.

Suppose X is a data vector and μ_x be the mean by equation 5. Let A be a matrix consisting of eigenvectors, by transforming a data vector X, one gets

$$Y = A(X - I\mu_x), \tag{8}$$

which are points in the orthogonal coordinate system defined by the eigenvectors. Components of Y can be seen as the coordinates in the new orthogonal basis. One can reconstruct the original data vector X from Y by

$$X = A^T Y + I\mu_x, \tag{9}$$

Instead of using all the eigenvectors of the covariance matrix, one may represent the data in terms of only a few PCs. For example, Y is constructed by taking the K first eigenvalues and their corresponding PCs, eigenvectors A_K ,

$$Y = A_K(X - \mu_x) \tag{10}$$

and

$$X \approx A_K^T Y + \mu_x. \tag{11}$$

This means that the original data is projected on the coordinate axes having the dimension K and then transform the data back by a linear combination of the basis vectors. This minimizes the error between the data and this reduced representation [22].

By picking the eigenvectors corresponding to the largest eigenvalues one loses as little information as possible about the whole data set. Moreover, by varying PCs and corresponding eigenvalues one can construct a set of data using fewer dimensions than in the original data. It is not straightforward to choose suitable number of components in PCA methods. Choosing too few components, one might lose important information of the variables and by choosing too many one might be overwhelmed by the number of dimensions. There are two methods which help to choose the number of components:

1. Plot the eigenvalues, the eigenvalues which are different from zero are corresponding to the main PCs. By drawing points on the graph, the eigenvalues make a curve shaped as an elbow. All the points before the flat line are the main eigenvalues, see Figure 16.

2. Find the cumulative sum of eigenvalues and choose all eigenvalues above the threshold M which is usually chosen between $90\% - 99.5\%$. This means to find
$$\frac{\sum_{i=1}^{K} \lambda_i}{\sum_{i=1}^{n} \lambda_i} \geq M. \tag{12}$$

The smallest value of K which fit to equation 12 is the desired number of components [4]. In this master thesis, M is 90% for Hand example and 98% for knee segmentation.

4.2 Example: PCA

A simple example of PCA is the 2 dimensions random variables: Figure 12 shows the two main PCs of a point cloud. The point cloud is a two dimension normal distribution with $\mu = (2, 4)$ and $\Sigma = (1\ 1.5; 1.5\ 3)$. Notice that the most variance of the data occurs along the first PC.

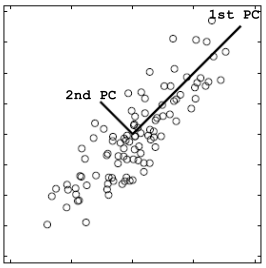


Figure 12: **o**s are the observations in two dimension and two lines are along the first and second PC.

Consider the large amount of data in this master's thesis which can lead to a heavy computation of covariance matrix. For this reason, Singular value decomposition (SVD) is used to calculate PCs as opposed to using the covariance matrix.

Singular value decomposition :
To speed up the algorithm and have a numerically more stable method instead of using eigenvalues and eigenvectors on data set by computing the cross product BB^T equation 7, singular value decomposition is suggested [21].

The singular value decomposition (SVD) of a matrix is obtained by

$$B = U\Sigma V^T, \tag{13}$$

where the squared diagonal elements of Σ contain the singular values, and the matrix V contains the corresponding singular vectors or PC's of the covariance matrix of B [21].

PCA algorithm :

1. Calculate $B = X - \mu$.
2. Calculate $SVD(B) = U\Sigma V^T$.
3. Sort eigenvalues, the squared diagonal elements of Σ, in descending order
4. Calculate the cumulative sum of eigenvalues or plot them to procure the efficient number of PC's, K.
5. Save to Kth of eigenvalues and corresponding PC's as main directions of the data.

In the next chapter, the important role of PCA method to make a shape model of images or shapes will be explained.

4.3 Statistical shape model

Given a set of training data e.g. images of a shape, one can build a statistical shape model. Each image is represented by a set of M landmark points. These landmarks must be consistent in different training data sets. Each shape can then be represented by a $2M$ element vector

$$X_i = [x_{i1}, \ldots, x_{iM}, y_{i1}, \ldots y_{iM}]\,. \tag{14}$$

The aligned training set forms a point cloud in the $2M$ dimensional space by means of Iterative Closest Points (ICP, Appendix A). PCA is used to obtain the main variance of the data which one can model in lower dimension. The shape model is then

$$X = \mu_X + \mathbf{Pb}, \tag{15}$$

where μ_X is the mean of the aligned training data, $\mathbf{P}$ is a $2M \times k$ matrix whose columns are eigenvectors along the k principal axes of the cloud, and $\mathbf{b}$ is a k element vector of shape parameters with limitation of $\pm m\sqrt{\lambda_k}$ where λ_ks are the variance of the b_ks and m is usually between 2 and 3 [4]. By varying the shape parameters within limits learnt from the training set, we can generate new examples. This limitation is obtained from the fact that most of the data lies between three square root of variance (standard deviation of the mean) [2].

4.4 Example: Statistical shape model on hand

One of the popular examples of PCA in two-dimensional case is the variation of a hand. The data set which is used in this example contains 18 different hands which are constructed on 72 landmarks in x, y coordinates with lines drawn between points.

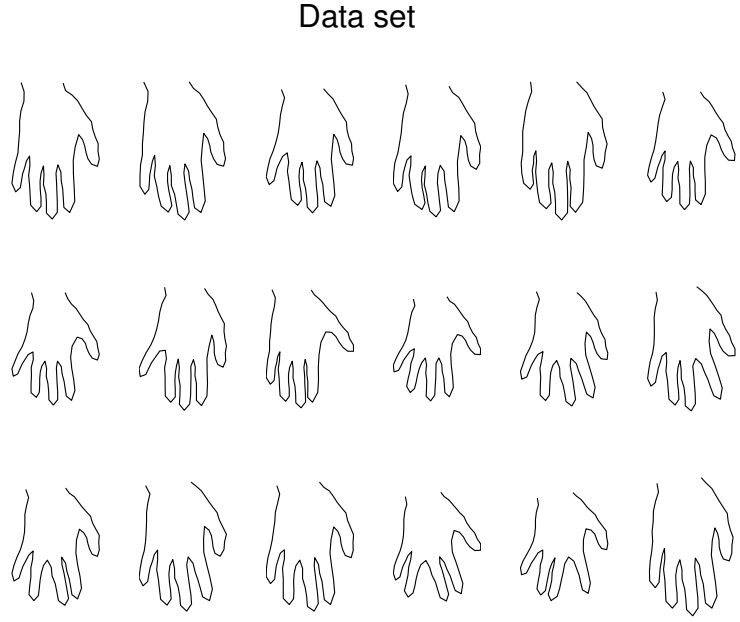


Figure 13: 18 different variations of a hand.

The first step is to reshape the matrices to vectors and then construct the X in the following way,

$X_{M\times N}$: matrix with M the training vector length and N the number of training data sets.

Secondly, one of the data set is chosen to be the model then iterative closest points (ICP) method (appendix A) is applied on the rest of data to find a best matching of different hands to the model according to translation and rotation.

The variations of the data sets are shown in Figure 14 and the matching data sets in Figure 15.

Figure 14: Each color corresponds to different variation of the hand.

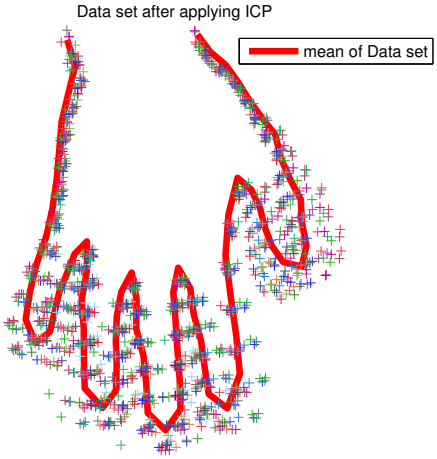


Figure 15: The red line shows the mean shape of data set and + shows the point cloud after applying ICP and aligning them to the model.

Figure 15 shows the points cloud of data set after applying ICP. Now one can use the PCA to find the principal directions of the data. To find the number of components mentioned methods on page 22 are used. From Figure 16 and

the cumulative sum, one can conclude the data set can be presented by the four largest eigenvalues and PCs which contain 90% of the variance of the data.

$$\frac{\sum_{i=1}^{4} \lambda_i}{\sum_{i=1}^{18} \lambda_i} = 0.9098 \tag{16}$$

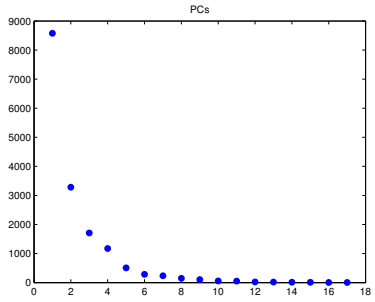


Figure 16: Sorted eigenvalues of the hands by descending order. The plot looks like an elbow e.g. becomes small after four points and close to zero.

In the next figures the reader can see the effect of each four PC's and also the constructed shape model.

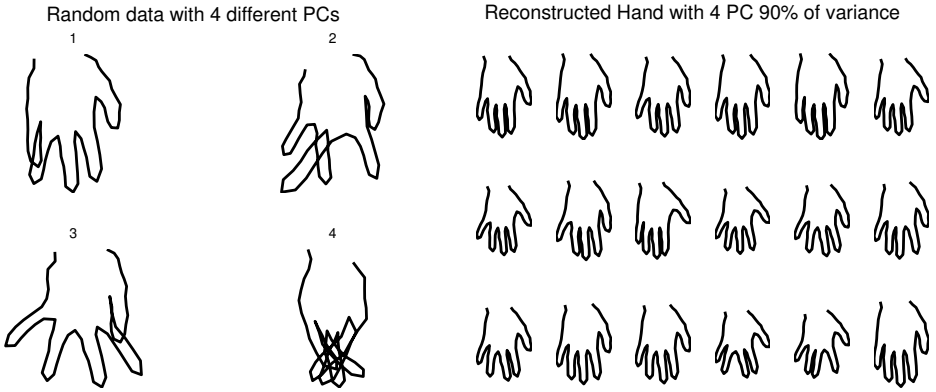


Figure 17: The left figure shows the effect of obtained main four PC's on a set of random data to reonstruct the hand and the right figure shows the reconstructed Hand with four PC's. Right figure is similar to the Figure 13 which shows that as little information lost during PCA method.

Figure 18 shows that by varying the four PCs, one can obtain the most variation of any hands.

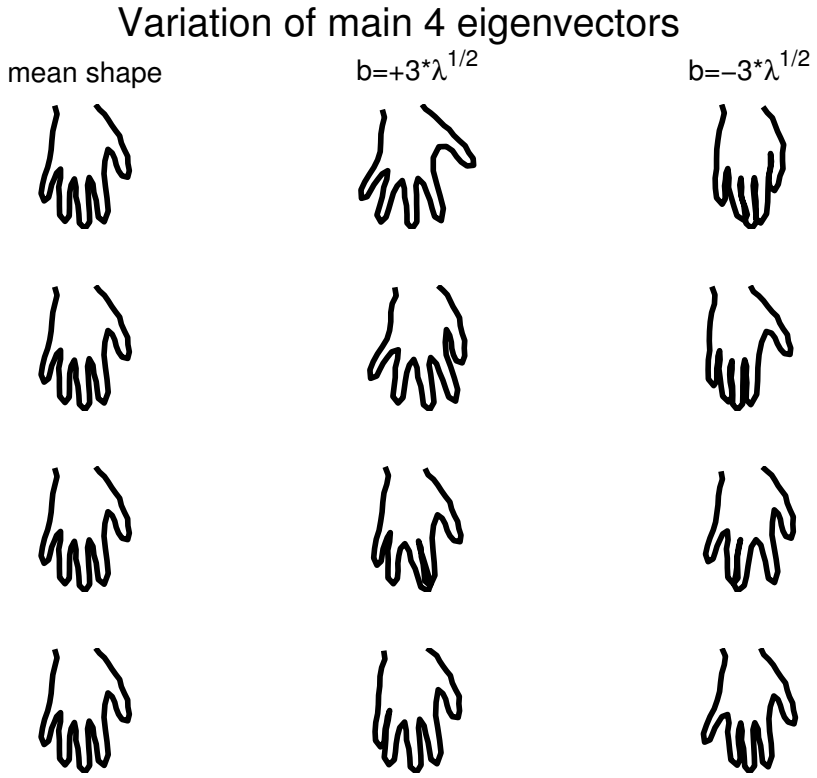


Figure 18: The variation of four different PCs around the limit of the variance of the data.

The PCA method can be extended into three dimensions. In this case, data sets have three dimensions x, y and z which should be treated exactly in the same way as in the two dimensions case.

For example, suppose image X_i is one of the training data, one can construct the original data $\Omega = [X_1, X_2, \ldots, X_N]$, N is the number of training data and

$$X_i = [x_{i1}, \ldots, x_{iM}, y_{i1}, \ldots, y_{iM}, \ldots, z_{i1}, \ldots, z_{iM}].$$

M is the length of each image. From the above vectors; one can obtain the matrix of the data set,

$$\Omega = \begin{bmatrix} x_{11} & & x_{N1} \\ \vdots & \cdots & \vdots \\ x_{1M} & & x_{NM} \\ y_{11} & & y_{N1} \\ \vdots & \ddots & \vdots \\ y_{1M} & & y_{NM} \\ z_{11} & & z_{N1} \\ \vdots & \cdots & \vdots \\ z_{1M} & & z_{NM} \end{bmatrix}$$

4.5 Resolution

To get better results and speed up the algorithm a multi resolution approach is used. In this master thesis three different resolutions are used $\frac{1}{4}$, $\frac{1}{2}$ and 1.

Generally, an algorithm starts with low resolution where the calculations are faster because the algorithm wants to handle less data and secondly because the amount of noise is less in low resolution image which ends up to easier way to find the right object in the image. In the high resolution level, the algorithm will find the small structure and more details in the object. The algorithm stops searching in one level after specific amount of iteration in that level. The number of iterations are obtained from where the algorithm converges and stops improving [18].

In Figures 19-21 three different resolution levels are shown.

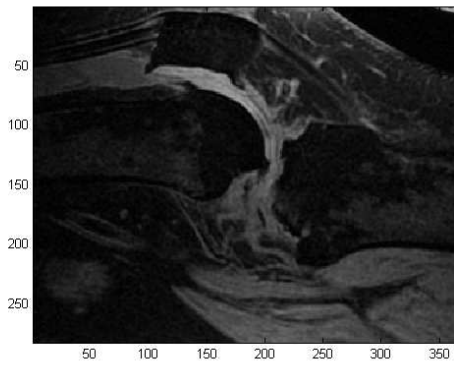


Figure 19: Normal resolution.

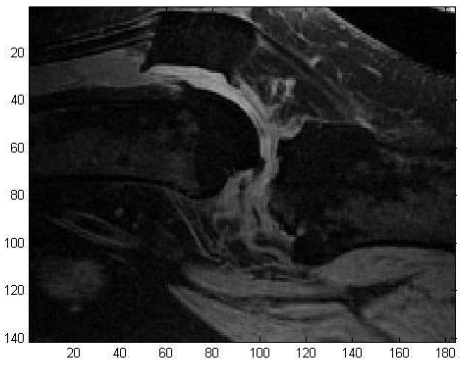


Figure 20: $\frac{1}{2}$ resolution.

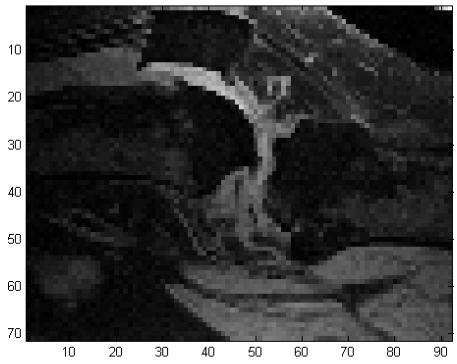


Figure 21: $\frac{1}{4}$ resolution.

4.6 Active shape model

The active shape model is a statistical model of shapes which iteratively deform to fit a new image. The shapes are constrained by a statistical shape model to vary only among the PCs of the landmarks of the training data. The mean shape is the initial estimate for shape parameters. Moreover, the rotation and the translation parameter initialize from the model shape [18],[19].

The active shape model is updated iteratively as follows:

1. Compute a suggested movement for each point by looking along normal of the surface through each landmarks to find the best match point for the model of the image appearance at that landmark.

2. Update translation and shape parameters to best fit the model according to the found points.

3. Repeat until convergence.

To get better results the above algorithm will be done in three different resolutions of the image. Searching for the results in different resolutions helps the algorithm to improve the robustness and the speed. The procedure above is a summary of the active shape model, more explanation about each step is explained further on in this chapter. Figure 22 shows the conceptual model of the active shape model.

The active shape model has two main phases:

1. Train the model parameters using the data sets

2. Apply the shape model

In the first part, the model is trained using the data set to obtain the shape model and appearance model of each landmark. The first phase will be followed by the second phase where the active shape model will be applied on the test data to segment the relative object in the new points' set.

In Figure 22 the conceptual model of active shape model is shown in a flowchart. The algorithm stops when there is no better movement or when it converges to the local optimal points.

Phase 1) Train on the data set
In this phase, after obtaining all landmarks on the surface of all the images by means of shape context algorithm, one needs to find the shape model and appearance models around surface. All data will be aligned e.g. all the translation and rotation will be removed. To remove translation one needs to centralize the data by subtracting the mean from data set and then correct the rotation for each new position.

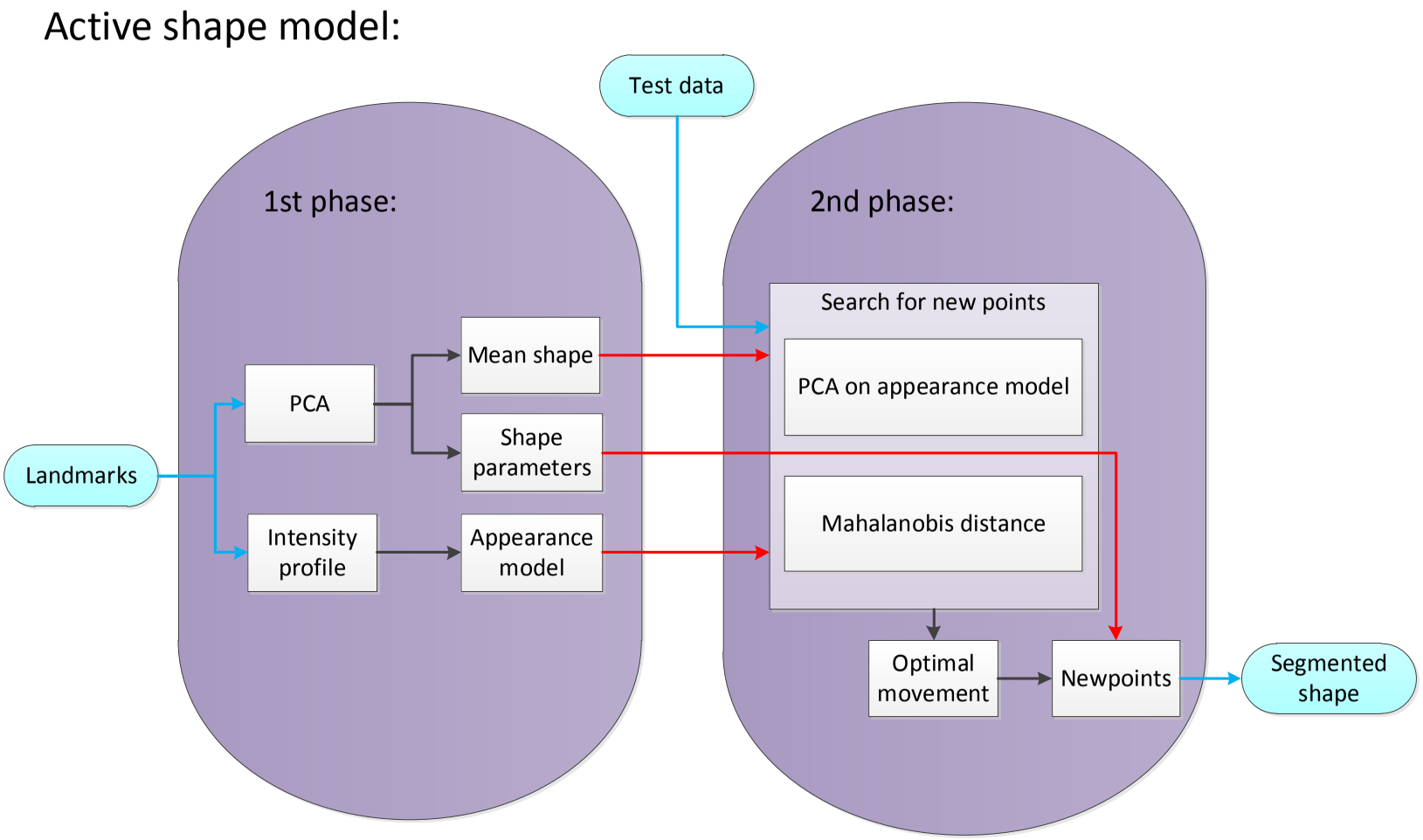


Figure 22: Conceptual model of Active shape model

Shape model:
For the shape model the PCA method is used. The mean shape will be the initial value of the shape parameters in phase 2).

Gray-level appearance model:
Appearance model of each landmark is used to model the variations of intensity around each edge point. For a specific landmark i, $2i+1$ equidistant points are sampled in the normal direction of the surface by fitting cubic splines. These points are stored in a vector g_i. The variations of these profiles are big; to minimize this effect, [18] proposed to use the normalized derivative of the intensity, dg_i. The derivatives of the profiles are calculated using the finite differences between $i-1th$ and $i+1th$ pixel. After normalizing the derivative,

$$Normalize(dg_i) = Ndg_i = \frac{dg_i}{\sum_j |dg_{ij}|}, \tag{17}$$

mean $\overline{Ndg}$ and covariance Σ_{Ndg} of these derivatives are also computed.

Phase 2) Apply model on new data
The mean shape is the initial estimate for shape parameters. The position of the mean shape, for example the femur bone, is at the position of femur bone in the MRI image.

Suggest movement and get new point :
To get better results it is convenient to start the calculation from rough image scale to fine image scale. In order to find the movement of a landmark, one needs to search n pixels in direction of the normal to the surface. In this part $n > i$ pixels are sampled along the normal direction of each side of all landmarks. The reason of choosing $n = i + s$ pixels is to let the algorithm search among more pixels for the optimal movement. Here i is the number of samples along the normal direction of the model and s is the number of extra search pixels, see Figure 23 and Figure 25. Hence, intensity profile on the new data set is created. From the intensity profile one can obtains the derivative. At this point the optimal movement is calculated by minimizing two functions:

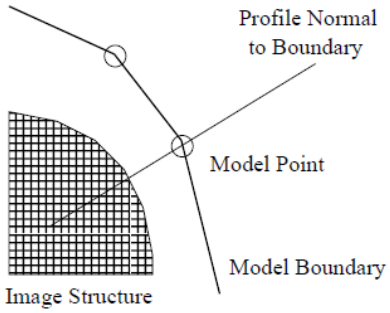


Figure 23: The concept of searching for a new point along the normal to the object's edge [19].

1. Mahalanobis distance,

$$f(g_s) = (g_s - \overline{Ndg})^T \Sigma_{Ndg}^{-1} (g_s - \overline{Ndg}), \tag{18}$$

where g_s is the sample made around the new point candidate. To find the movement one needs to take the lowest value of the $f(g_s)$. Minimizing the $f(g_s)$ is equivalent to maximizing the probability that g_s originates from a multidimensional Gaussian distribution. Mahalanobis distance assumes a normal distribution of profiles. However in some cases, two or more gray level structures describe the edge in different images, see Figure 24. Therefore, the distributions of the profiles are not normal. Since Mahalanobis distance assumes a normal distribution of profiles, using the shape parameters of appearance model is suggested in this case [4].

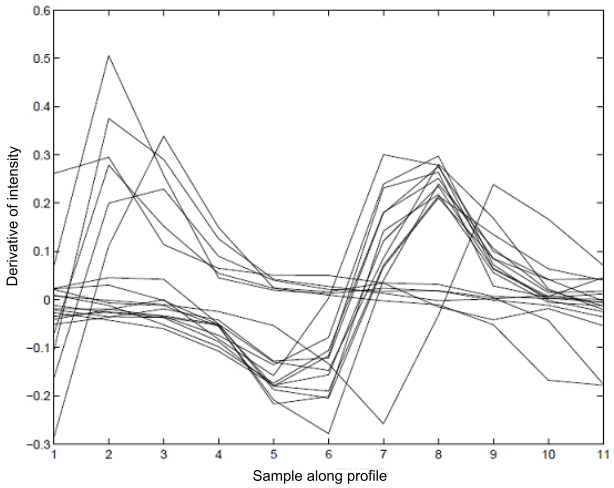


Figure 24: Two different gray level structures describe the edge.

2. Shape parameters of appearance model

If a profile intensity g_i is considered as an approximation of the appearance where

$$g_i = \overline{g} + \mathbf{Q}\mathbf{c}, \tag{19}$$

according to equation 15 then

$$\mathbf{c} = \mathbf{Q}^T (g_i - \overline{g}), \tag{20}$$

and

$$f(g_s) = \sum \mathbf{c}^2 = \mathbf{c}^T \mathbf{c} = (g_s - \overline{g})^T \mathbf{Q}\mathbf{Q}^T (g_s - \overline{g}), \tag{21}$$

then the movement is obtained by minimizing the equation 21.

When a new point is being searched the above functions are calculated to find the optimal movement and to fit the best point. Figure 25 shows this procedure.

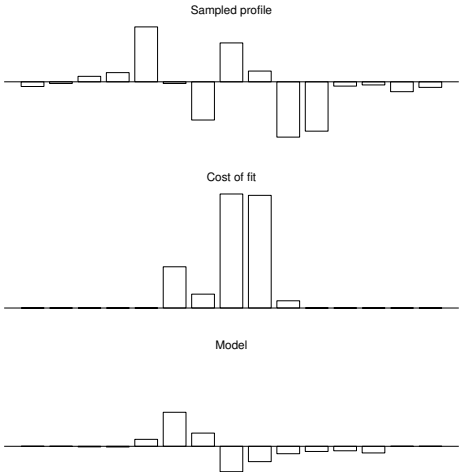


Figure 25: One iteration of phase 2 on one specific landmark. The top figure shows $2 \cdot (5 + 2) + 1$ sampled profile. These 7 points sampled along the normal direction of each landmark to search for a new point. The middle bar in each plot represents the corresponding landmark. The second figure correspond to the optimal movement by minimizing the cost function, $2 \cdot (7 - 5) + 1$. The last figure shows the model's profile $2 \cdot 5 + 1$. According to the figures the best movement is to move two pixels to right in this iteration.

After locating the point the last step is to apply constraints on the shape model parameter and updating the translation and rotation parameters [18],[19].

Updating the parameters and constraining the restrictions :
To find the shape parameter of the new landmark points, one needs to solve the same equation as in equation 20,

$$\mathbf{b} = \mathbf{P}^T(Y - \mu_X), \tag{22}$$

where Y is the new points set. Then constrains are applied to the b parameters:

$$-3\sqrt{\lambda_i} \leq b_i \leq +3\sqrt{\lambda_i}, \tag{23}$$

where λ_i is the variance of shape mode i.

Finally, the new landmarks are checked against the old points, and if there is no improvement the algorithm stops.[2].

5 Applications, results and comparison

After studying the methods and constructing the algorithm, in this section the performance of the algorithm is investigated using the data sets. The application starts with segmenting the bones and continues with the cartilages.

5.1 Applications

Femur Bone: As in Figure 5, after extracting the points from different images, the shape context algorithm is applied to obtain the consistent landmarks in the training data set. In Figure 26, the reader can see the results of the shape context on the femur bone after five iterations. Figure 26 also shows in each iteration the pair matching is getting better.

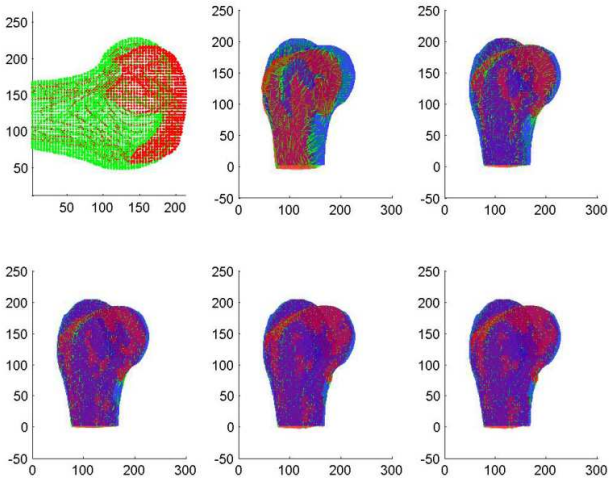


Figure 26: The green shape in first subplot in top left is target and the red shape is the model. Each subplot after the left top is the matching pair which gets better after each iteration of the algorithm. The results are the consistent landmarks between model and target.

After finding the corresponding points in each data set according to the model (see chaper 3), the data sets are ready to use for the active shape model. Active shape model algorithm has two phases. The first phase is to train the model using data sets.

Using the data sets from the shape context, one can obtain the mean shape and constraints for the shape parameters. Moreover, the variation of the gray-level appearance model of each landmark is obtained in this phase from appearance model.

After training the model parameters using the data set, the model can be applied onto a test data set. Example below obtained from applying the algorithm on one of the training data sets.

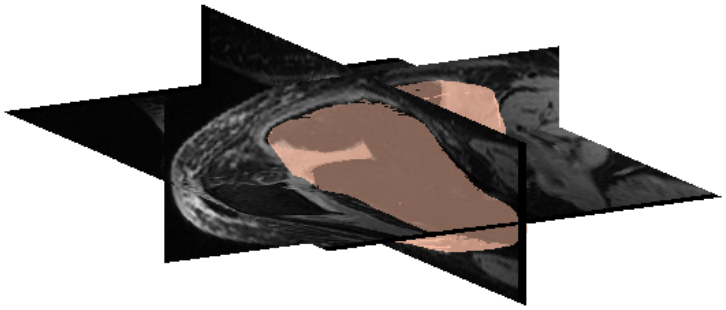


Figure 27: Three dimensional MRI image and the segmented femur bone in pink.

The colored shape in figure 27 is the result of segmentation of femur bone using the algorithm on the MRI image of the test data set.
Figure 28 shows the obtained femur bone from the above segmentation.

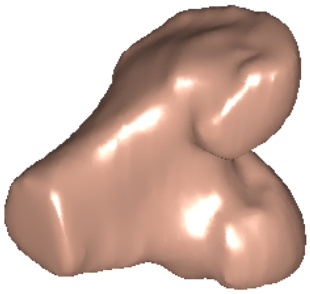


Figure 28: Segmented bone.

After segmenting the femur bone of the test data, the results are transferred to a three-dimensional binary image. In Figure 29 the result of the binary image is shown in sagittal slices.

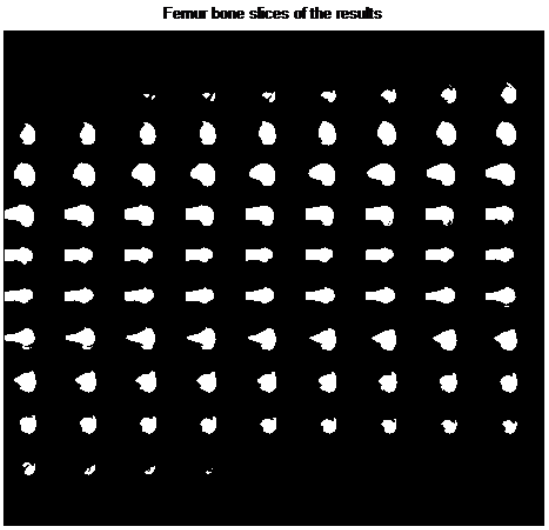


Figure 29: Different slice of segmented femur bone.

For more detailed result, one slice of segmentation is shown on the MRI raw pictures in Figure 30.

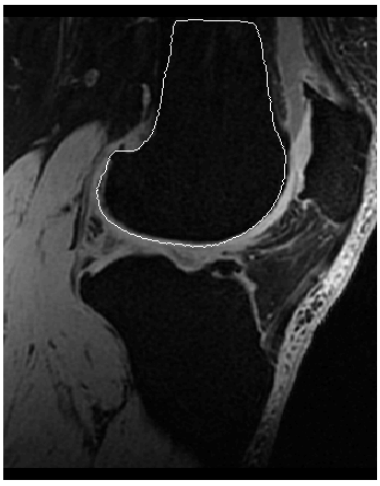


Figure 30: One slice of femur bone in a MRI's slice. The white line is the segmented femur bone on the original MRI picture. The image shows how good the model on one of the training data set is.

Tibia bone : Same procedure applied on the tibia bone (Figure 31) and the result are shown in Figure 32.

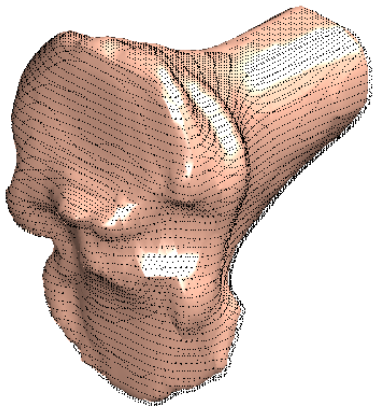


Figure 31: Original Tibia Bone. The black points on the bone are the vertices which are choosen to bulid the model of tibia bone.

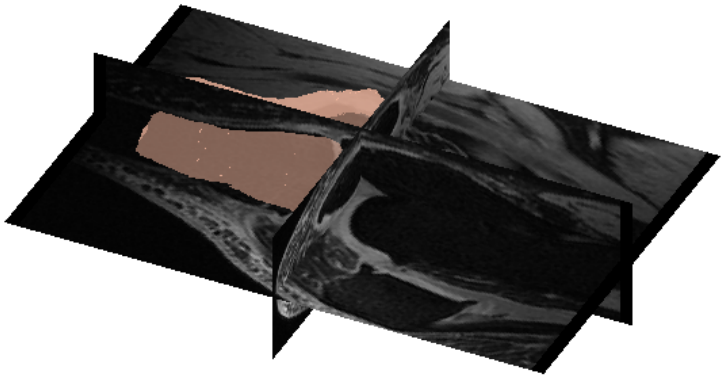


Figure 32: Three dimensional MRI image and the segmented tibia bone in pink.

The algorithm applied on both femur cartilage (Figure 33) and tibia cartilage (Figure 34). Due to low thickness of the cartilage the search points through the normal direction are choosen to be fewer than the bone search points.

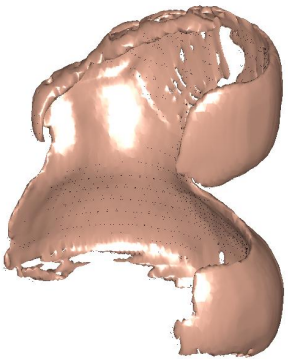


Figure 33: Femur Cartilage and the chosen landmarks (black points).

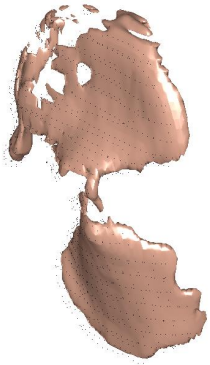


Figure 34: Tibia Cartilage and the chosen landmarks (black points).

5.2 Results

Table 5 shows the results of the model on twenty of the test data sets between 61 and 100. This results obtained from the software provided by the organizer [1]. As it was explained in the introductional chapter on page 8, the Avg Bone Score is computed from the average (AvgD) and the root mean square (RMSD) surface distances. The Avg Cart Score is computed from the volumetric overlap error (VOE) and volumetric difference (VD). These are measured in percent. Finally the Avg Total Score is the average of the bone and cartilage score [1].

Img	Femur bone		Tibia Bone		Fem. Cartilage		Tib. Cartilage	
	AvgD [mm]	RMSD [mm]	AvgD [mm]	RMSD [mm]	VOE [%]	VD [%]	VOE [%]	VD [%]
61	2.63	3.55	6.31	7.90	90.04	-66.98	91.25	-0.17
63	2.68	3.70	2.76	3.46	91.78	-78.08	92.65	-50.25
65	7.45	8.81	10.52	12.12	95.52	-68.87	100.00	-88.73
67	3.39	4.36	3.78	5.13	88.72	-56.35	97.28	-70.66
69	3.29	4.18	6.04	7.38	97.33	-83.34	100.00	-100.00
70	3.09	4.17	4.01	5.01	95.56	-84.86	89.61	-66.32
72	7.21	8.32	7.40	8.64	96.03	-66.82	100.00	-80.13
74	8.84	10.12	10.99	11.75	84.29	-69.82	100.00	-98.50
78	2.71	3.89	2.73	3.70	93.99	-75.27	87.38	-34.91
80	5.92	7.46	6.75	8.21	99.79	-83.15	100.00	-95.75
87	9.67	10.97	11.64	12.30	95.80	-79.43	100.00	-96.96
89	3.48	4.76	8.86	11.01	83.55	-53.78	97.71	-82.93
90	9.88	11.91	10.17	11.65	99.36	-93.68	100.00	-100.00
92	8.84	10.23	10.96	11.99	91.82	-70.12	100.00	-72.89
94	3.13	4.28	2.54	3.34	93.66	-78.22	84.53	13.86
96	6.55	8.85	4.67	6.34	87.51	-73.92	87.11	-37.85
98	7.98	10.18	10.31	12.34	99.17	-97.05	90.61	-70.89

Table 5: The results on the different test data set. The measures of the bone are in millimeter and the measures for the cartilage are in percentage.

According to Table 5 the average femur bone measures are $AvgD = 7.07$ and $RMSD = 8.92$ millimeter and the average tibia bone measures are $AvgD = 6.6$ and $RMSD = 8.1404$ millimeter. The average femur cartilage's volumetric overlap error is 93.08% and volumetric difference is -50.25%. Moreover, the average tibia cartilage measures are $VOE = 88.11\%$ and $VD = -18.64\%$.

6 Discussion

Table 5 of the results shows the active shape model is a good model to segment the bones. The minimum distances for femur bone are 2.63mm and 3.55mm from image 61 and for tibia bone are 2.54mm and 3.34mm from image 94. The maximum distances are 17.96mm and 23.75mm for femur bone and are 24.74mm and 33.12mm for tibia bone from image 76. The results of cartilages as it was explained are in percentage. These results show that the ASM is not a good model to segment the cartilage.

One of the reasons that the ASM did not work for cartilage is the variation of the cartilage shape. One may consider the fact that the bone get damage later than cartilage hence, the mean shape can be so different from each training data set corresponding to range of damage. Another reason is cartilage thickness which makes it harder to be detected in the model.

The data sets provided by the organizers lacked DICOM files. The DICOM file gives more information about the MRI images, such as, orientation of the knee (left or right). Therefore, the ASM based on both left and right. This probably adds more variation to the model. Moreover, as it was explained in chapter "Data" page 14, MRI images are obtained from different scanners which lead to different ranges in intensity.

There are additional works to this master's thesis which one could consider as a new study, such as:

1. identifying the knee orientation by testing on a known side,
2. using another model different than ASM model to segment the cartilage,
3. restricting the movement after finding bone segmentation and adding cartilage thickness map to the model,

Due to the time restriction for this master's thesis, there was no more time to further the work; however, interested students and researchers could continue and further develop this work to achieve a better result.

A Iterative Closest Point (ICP)

ICP is an iterative algorithm of finding a best match between two points by minimizing the root mean squared distance.

Consider two sets of points $X, Y \subseteq \mathbb{R}^d$ where d is the dimension, $|X| = n$ and $|Y| = m$. We want to find a one-to-one matching function $\theta : X \to Y$ that minimizes the root mean squared distance (D) between X and Y,

$$D(X, Y, \theta) = \sqrt{\frac{1}{n} \sum_{x \in X} \|x - \theta(x)\|^2}. \tag{24}$$

And considering rotation and translation, we obtain following minimization:

$$min_{\theta: X \to Y, t \in \mathbb{R}^d, R \in SO(d)} \sum_{x \in X} \|Rx - t - \theta(x)\|^2, \tag{25}$$

where R is the rotation matrix, t is the translation vector and $SO(d)$ is the set of special orthogonal matrices (rotation matrices) in d dimensions [26].

ICP algorithm alternate between its two steps :

1. A matching step, where a fixed given translation and rotation, the optimal matching is computed by minimizing the D, this step is the slowest part of the algorithm. For this reason, a k-d tree is used to find the nearest neighbor more quickly in the algorithm. For an example on k-d tree see [30]. For instance if there is a two-dimensional case, the algorithm takes $O(m \log m)$ time for preprocessing and $O(\log m)$ time per point in X where $m = |Y|$. $\forall x \in X$, find closest $y \in Y$ in the following manner:

   - Construct a Voronoi diagram on Y. For an example on Voronoi diagrams see [31]. A voronoi diagram contains points, $S = \{s_i\}_{i=1}^{\infty}$ in the plane. For each point, $s_i, i = 1, 2, \ldots$ there exists a cell which is also called a Dirichlet cell. All points in this cell are closer to s_i than to any s_j ,$i \neq j$. Defining a Voronoi diagram with its dual might be easier: Delaunay triangulation. A Delaunay triangulation for a set P is a triangulation such that no points p_i in P is inside of any circumcircle of any triangles. A circumcircle is a circle which is constructed from three points or a triangle where the points are on the edge of the circle and the center of the circle has a same distance from each three points.
   - $\forall x \in X$, do point-location in Y.

2. A transformation step, where a given matching the optimal translation and rotation are computed. Assume there is a function θ such that $\theta(x_i) = y_i$, where $x_i \in X$ and $y_i \in Y$. We want to find a matrix R (rotation matrix) and a vector t (translation) which minimizes the following equation:

$$\sum_{i=1}^{n} \|Rx_i - t - y_i\|^2 \tag{26}$$

The mean of X and Y is defined as

$$\bar{x} = \frac{1}{|X|}\sum_i x_i \ , \bar{y} = \frac{1}{|Y|}\sum_i y_i, \tag{27}$$

This definition will help simplify the minimization problem as follows:

$$\acute{x}_i = x_i - \bar{x} \Rightarrow x_i = \acute{x}_i + \bar{x} \ \acute{y}_i = y_i - \bar{y} \Rightarrow y_i = \acute{y}_i + \bar{y}. \tag{28}$$

Hence,

$$\begin{aligned}\sum_{i=1}^{n} \|Rx_i - t - y_i\|^2 &= \sum_{i=1}^{n} \|R(\acute{x}_i + \bar{x}) - t - (\acute{y}_i + \bar{y})\|^2 \\ &= \sum_{i=1}^{n} \|R\acute{x}_i - \acute{y}_i + (R\bar{x} - \bar{y} - t)\|^2\end{aligned}$$

Now if $t = R\bar{x} - \bar{y}$, It obtains

$$\sum_{i=1}^{n} \|R\acute{x}_i - \acute{y}_i\|^2 . \tag{29}$$

Equation 29 can be expanded as follows:

$$\begin{aligned}\sum_{i=1}^{n} \|R\acute{x}_i - \acute{y}_i\|^2 &= (R\acute{x}_i - \acute{y}_i)^T(R\acute{x}_i - \acute{y}_i) \\ &= RR^T \sum_{i=1}^{n} \|\acute{x}_i\|^2 - 2tr(R\sum_{i=1}^{n} \acute{x}_i\acute{y}_i^T) + \sum_{i=1}^{n} \|\acute{y}_i\|^2 \\ &= \sum_{i=1}^{n} \|\acute{x}_i\|^2 - 2tr(R\sum_{i=1}^{n} \acute{x}_i\acute{y}_i^T) + \sum_{i=1}^{n} \|\acute{y}_i\|^2 ,\end{aligned}$$

since $RR^T = I$.

Let $N = \sum_{i=1}^{n} \acute{x}_i\acute{y}_i^T$, hence, the problem of minimizing equation 29 is reduced to that of maximizing the following:

$$tr(RN)$$

Taking the singular value decomposition (SVD) of matrix N,

$$N = U\Sigma V^T$$

where U and V are orthogonal matrices and $\Sigma = diag(\sigma_1, \sigma_2, , \sigma_d)$ such that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_d \geq 0$. From this, it can be proven [32] that to minimize equation 26 the rotation matrix should be defined as:

$$R = VU^T$$

.

Finally the algorithm stops its iterations when the matching stays unchanged [26].

The following is a brief description of the ICP algorithm:

ICP(A,B)

1. Initialize the rotation, $R = I$(the identity matrix) and the translation $t = 0$.
2. Matching Step: Given R and t, compute optimal μ by finding
$$\min_{\mu} D(A, B, \mu)$$
3. Transformation Step: Given μ, compute optimal R and t by finding
$$\min_{R,t} D(RA - t, B, \mu)$$
4. Go to step 2 unless μ is unchanged.

B Testing the error of converting a surface to a binary mask

After the segmentation of the different parts of the knee joint, Femur bone, Tibia bone, Femur cartilage and Tibia cartilage, is done one need to convert the results to a binary mask. Then this binary mask can be used to calculate the error of the model in comparison with the original data sets and to visualize the results. To convert a surface to a binary mask, the vertices and faces are written to a file format *.stl*. Standard Tessellation Language (STL) is a file format which contain a unit normal and vertices of the triangles using a three dimensional Cartesian coordinate system [33]. The test start with some simple shapes such as, cube and sphere and it continues with a cylinder and then femur bone and tibia bone, as a complex shape. All the shapes start with 50 points in each dimension and then the number of points increases to investigate the boundary of the error.

Example cube :
The first test is made on a simple shape cube, see Figure 35.

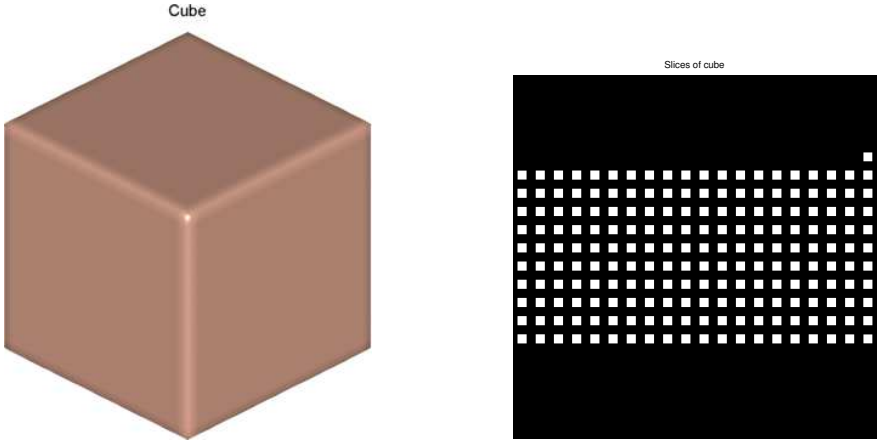


Figure 35: The left figure is a simple cube and the right figure is the extraction of the binary mask of a cube.

After extracting a binary mask from the volumetric cube, 50 slices obtained, see Figure 36.

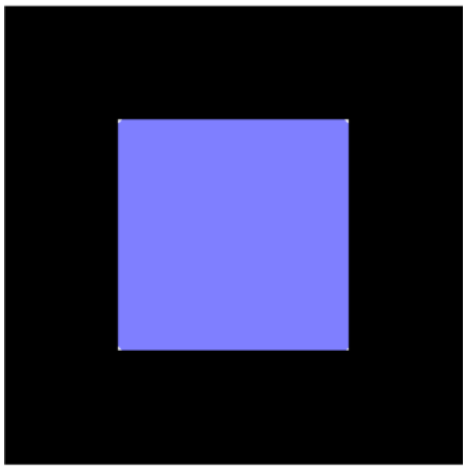


Figure 36: The white shape is the original cube and blue shape on the white shape is the binary mask obtained from stl file. The white line around the blue cube is the error.

Figure 36 shows the difference from the original cube (white) and the binary mask (blue) obtained from the stl file convertor. The white line along the boundary of the cube is same in all slices which lead to some errors. The error of a $50 \times 50 \times 50$ cube of the stl convertor is equal to 0.1104. The same procedure applied on 100, 200, 300 and 400 dimension and errors are as following,

Number of points	Error
50	0.1104
100	0.0305
200	0.0080
300	0.0027
400	0.0015

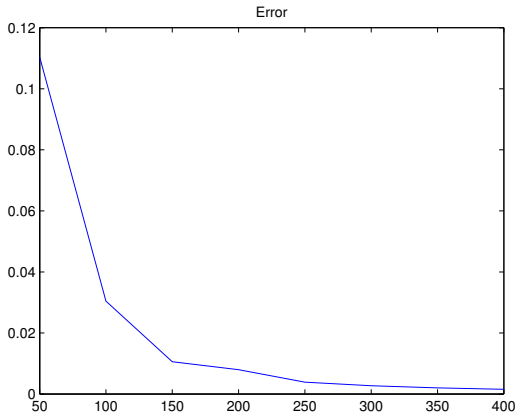


Table above shows that the errors are between %1 to 10% for a simple shape such as cube.

Example sphere :
Second test is made on a sphere, see Figure 37 left side.

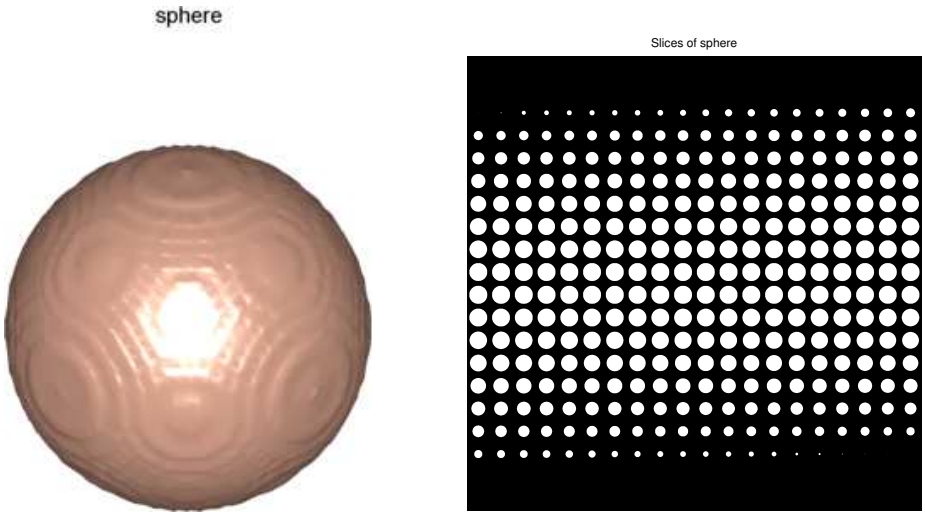


Figure 37: The left figure is a simple sphere. The right figure is a extraction of a binary sphere.

Figure 38 shows the error of converting a sphere surface to a binary mask. These errors mostly happen in the edges.

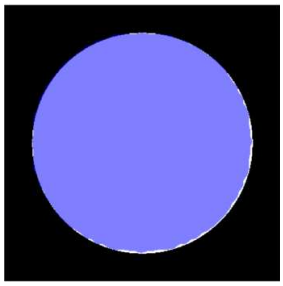


Figure 38: The blue shape is the binary mask obtained from the stl convertor and the white pixels around the sphere are the difference between the original shape (white) and the blue shape. The error of a stl convertor is happened around the edges. Moreover, this error is much bigger than a cube.

Number of points	Error
50	0.1998
100	0.0615
200	0.0227
300	0.0571
400	0.0249

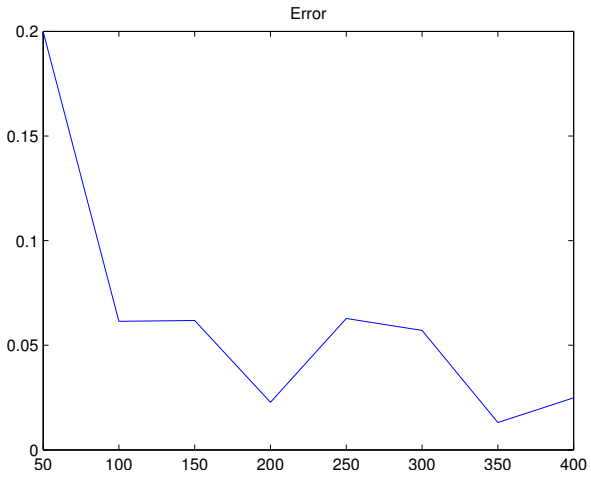


Figure and table above show the errors are between 2% to 20%.

Example cylinder:
Third test is made on a cylinder as the shape of a cylinder is close to a shape of a bone (Figure 39).

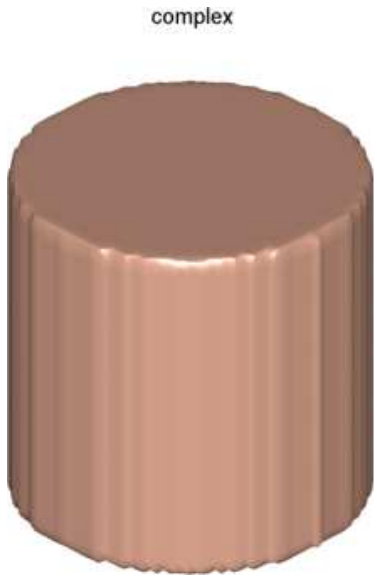


Figure 39: A cylinder as a complex shape close to a bone shape.

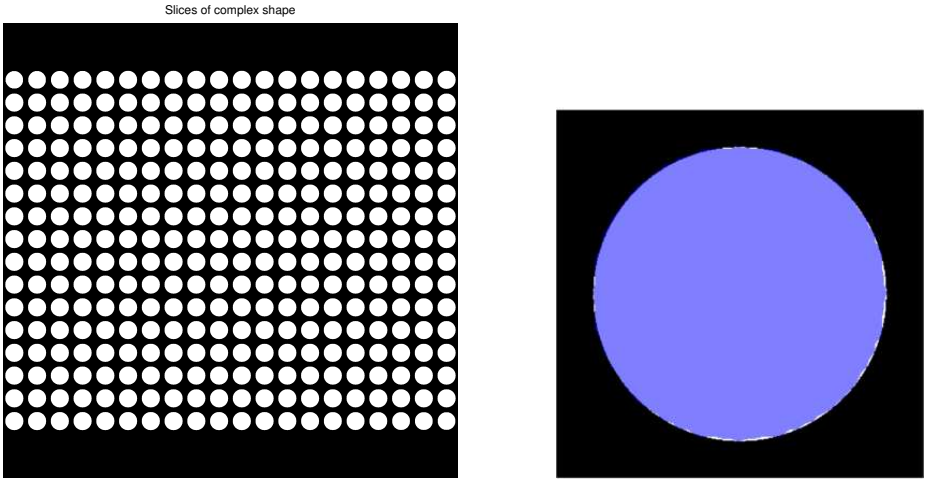
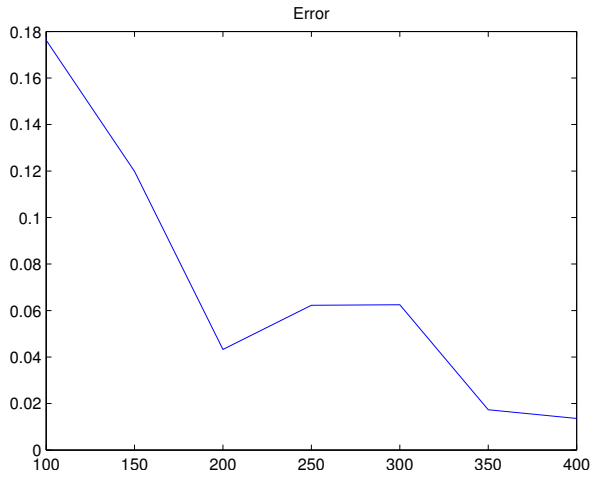


Figure 40: The left side figure is the slices of a binary volume. The right figure is one of the slices with a blue layer of stl convertor binary mask. The white boundary is the error between the blue mask and the original volume.

From the blue layer in Figure 40, it is clear that the error happened in the boundary of a cylinder.
Table below shows that the error of a cylinder is between 1% and 10%.

Number of points	Error
100	0.1762
200	0.0433
300	0.0625
400	0.0136



Example bone:
The test is made on the both bones, Femur (Figure 44) and Tibia (Figure 45). First five different data sets are selected to compare the results, see Table 6. Secondly, different amount of vertices are chosen to investigate the error, see Table 7.

Figure 41: Femur bone

Figure 42: Tibia bone

Tables 6 and 7 show the results on both bones. The errors are roughly between $4\% - 8\%$.

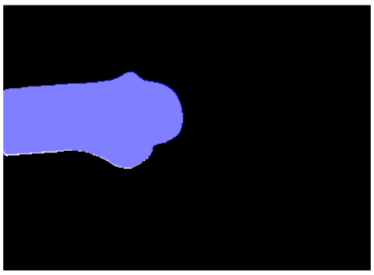
Femur bone		**Tibia bone**	
Data	**Error**	**Data**	**Error**
1	0.0497	1	0.0551
2	0.0491	2	0.0459
3	0.0427	3	0.0707
4	0.0634	4	0.0658
5	0.0550	5	0.0529

Table 6: Two first column shows the error of Femur bone in five different data sets and the rest shows the error of Tibia bone.

Femur bone			**Tibia bone**		
Number of		**Error**	**Number of**		**Error**
Vertices	Faces		Vertices	Faces	
112536	225068	0.0497	88216	176428	0.0551
28464	56924	0.0513	22196	44388	0.0458
7036	14068	0.0440	5460	10916	0.0428
1712	3420	0.0785	1348	2692	0.0742

Table 7: shows the error of both bones by decreasing the number of vertices.

The reader can see the results of converting in Figure 43 where the blue mask is the converted shape and white one is the original data set. It is clear that the error occurred in edges.



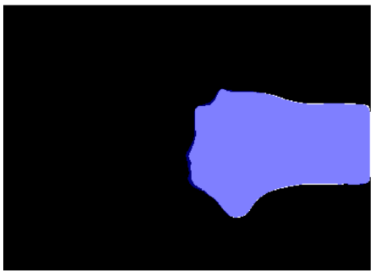


Figure 43: shows the converted shape (blue mask) and the original data set (white shape). It is clear that the error occurred in edges.

Figure 46 shows the extraction of femur and tibia bone after converting the surface into binary mask.

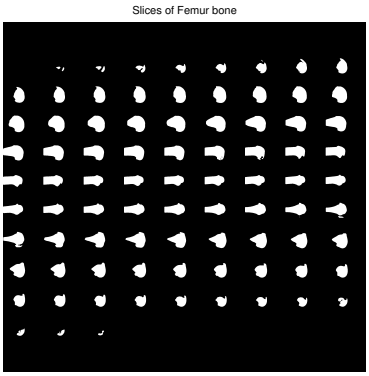


Figure 44: Femur bone

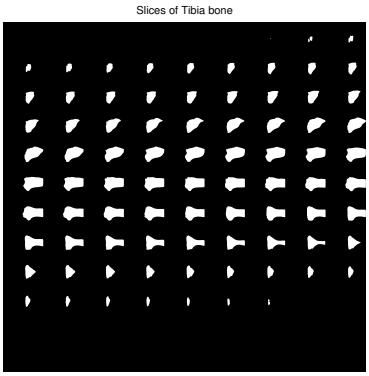


Figure 45: Tibia bone

Figure 46: shows the extraction of femur and tibia bone after converting the surface into binary mask.

In conclusion, the test shows that the converting procedure from surface to binary mask using stl format has an error between 4% to 8%. This means that $4 - 8\%$ of the errors in the final results can be attributed to this error.

References

[1] Heimann, T., Morrison, B., Warfield, S., Styner, M., Niethammer, M. *Medical Image Analysis for the clinic - A Grand Challenge: Segmentation of Knee Images 2010* (2010), `http://www.ski10.org`.

[2] Cootes, T. F., Taylor, C. J., Cooper, D. H., Graham, J.*Active shape models-their training and application, computer vision and image understanding* , vol. 61, No. 1, January, pp. 38-59, 1995.

[3] Williams, T. G., Holmes, A. P., Waterton, J. C.,Maciewicz, R. A., Hutchinson, C. E., Moots, J. R., Nash, A. F. P., Taylor, C. J. *Anatomically corresponded regional analysis of cartilage in asymptomatic and osteoarthritic knees by statistical shape modeling of the bone* , IEEE transactions on medical imaging, vol. 29, No. 8, August 2010.

[4] Ginneken, B. V., Frangi, A. F., Staal, J. J., Romeny, B. M. T. H., Viergever, M. A. *Active shape model segmentation with optimal features* , IEEE transaction on medical imaging , Vol. 21, No. 8, August 2002.

[5] American Academy of Orthopaedic Surgeons.
`http://orthoinfo.aaos.org/topic.cfm?topic=a00422` ,
Copyright©2006.

[6] Image©Medical Multimedia Group.
`http://www.orthogate.org/patient-education/knee/osteoarthritis-of-the-knee.html,`

[7] Wikipedia, Magnetic resonance imaging.
`http://en.wikipedia.org/wiki/Magnetic_resonance_imaging` (last visited Dec. 15, 2011).

[8] Copyright©2009, WebMD, LLC.
`http://bit.ly/z3DWrg`
`http://www.webmd.com/pain-management/knee-pain/magnetic-resonance-imaging-mri-of-the-knee`

[9] `http://www.itk.org/`

[10] `http://www.mint-medical.de/imprint/`

[11] `http://www.amira.com/licensing.html`

[12] Graham, V., Wolstenhholme, C., Scott, L., Bowes, M. *Fully automatic segmentation of the knee joint using active appearance models* , MICCAI grand challenge 2010.

[13] Seim, H., Kainmueller, D., Lamecker, H., Bindernagel, M., Malinowski, J., Zachow, S. *Model-based auto-segmentation of knee bones and cartilage in MRI data* , MICCAI grand challenge 2010.

[14] Yin, Y., Williams, R., Anderson, D. D., Sonka, M. *Hierarchical decision framework with a priori shape models for knee joint cartilage segmentation* , MICCAI grand challenge 2010.

[15] Lee, S., Shim, H., Park, S. H., Yun, II. D., Lee, S. U, *Learning loca shape and Appearance for segmentation of knee cartilage in 3D MRI* , MICCAI grand challenge 2010.

[16] Amberg, M., Luthi, M., Vetter, T., *Fully automated segmentation of the knee using local deformation-model fitting* , MICCAI grand challenge 2010.

[17] Korc, F., Schneider, D., Forstner, W. *On nonparametric Markov random field estimation for fast automatic segmentation of MRI knee data* , MICCAI grand challenge 2010.

[18] Cootes, T. F., Taylor, C. J. *Statistical Models of Appearance for computer vision, Imaging science and biomedical engineering* , university of Manchester, March 2004.

[19] Cootes T.F., Taylor C.J. *Chater 7: "Model-Based Methods in Analysis of Biomedical Images" in "Image Processing and Analysis"* , Ed.R.Baldock and J.Graham,Oxford University Press, 2000, pp223-248.

[20] `http://www.fon.hum.uva.nl/praat/manual/Principal_component_analysis.html`
©djmw, November 10, 2010.

[21] Jolliffe I, T. *Principal Component Analysis* , Series: Springer Series in Statistics, 2nd ed., Springer, NY, 2002, XXIX, 487 p. 28 illus. ISBN: 978-0-387-95442-4.

[22] Hollmn, J. *Process Modeling Using the Self-Organizing Map.* Master's thesis, Helsinki University of Technology, 1996. `http://users.ics.tkk.fi/jhollmen/dippa/node30.html`.

[23] Belongie, S., Malik, J., Puzicha, J. *Shape matching and object recognition using shape contexts* , IEEE transactions on pattern analysis and machine intelligence, vol. 24, No. 24, April 2002.

[24] Papadimitriou, C., Stieglitz, K. *Chapter 11: Combinatorial Optimization, Algorithms and complexity.* Prentice Hall, 1982.

[25] Petersen, N. *shape context* , 2008.
`http://www.cs.utexas.edu/~grauman/courses/spring2008/slides/ShapeContexts425.pdf`

[26] Philips, J. M., guest lecturer ICP,
`http://www.cs.duke.edu/courses/spring07/cps296.2/scribe_notes/lecture24.pdf`.

[27] Wikipedia, Delaunay triangulation,
`http://en.wikipedia.org/wiki/Delaunay_triangulation` (last visited Dec. 15, 2011, 14:41 GMT).

[28] `http://www.mathworks.se/help/techdoc/ref/isosurface.html`
Copyright 1984-2004 The MathWorks, Inc. (Revision: 1.8.4.5 Date: 2005/06/21 19:37:46)

[29] Fornasini, P. *The uncertainty in physical measurements. Chapter 11: The chi Square Test.* Springer New York, pp. 187-198, 2009, ISBN: 978-0-387-78650-6.

[30] K-d tree example: `http://upload.wikimedia.org/wikipedia/commons/9/9c/KDTree-animation.gif`.

[31] Voronoi diagram example: `http://www.cs.cornell.edu/home/chew/Delaunay.html`.

[32] Arun, K. S., Huang, T. S., Blostein, S. D. *Least-Squares Fitting of Two 3-D Point Sets* , IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 9, pp. 698-700, 1987.

[33] Wikipedia, STL (file format), `http://en.wikipedia.org/wiki/STL_(file_format)(LastvisitedMar.11,2012,18:56GMT).`