

Assignment 3

Medical Image Analysis

Azalea Alothmani

December 12, 2021

1 Introduction

One of the important goals of machine vision is the ability to recover an image, but also the ability to understand the image and what it represents. By definition, this is achieved by using models which describe and label the expected shape structure. In order to achieve that, model-based approaches have been developed over the past decade. However, the interpolation of complex and variable shapes such as internal human body organs has shown to be problematic due to the fact that these image structures are often noisy and incomplete. The variability in structure images needs to be organised using a model that is specific and capable of representing only specific patterns while allowing for the natural variability. Recent methods have developed to encounter this problem by using statistical models as image interpretation. [1]

In this project a shape analysis is going to be studied and a shape model will be created to segment kidneys in renal scintigraphy images. By developing the created shape model the possibility of improving the segmentation process is investigated by applying the shape model and a segmentation algorithm on images that have not been used during shape model construction. The first part of this project covers the development of a shape model based on a number of segmented kidneys. In the second part, the developed shape model algorithm is used to improve the segmentation of further kidneys.

2 Theory

The shape of an object can be represented by a continuous curve that corresponds to the contour of the object. The continuous curve is the result of a segmentation algorithm, and has been created from a sampled set of points/landmarks along the contour. The "smooth" curve is obtained by interpolation and re-sampling in order to create evenly distributed points.

2.1 Linear Interpolation

Linear interpolation is a curve fitting method, and the new estimated point is assumed to be on the line of the nearest points along the joining line between two known data points. The method is used to construct new data point by linearly interpolating between the known data points, x_i , and the known values in interval $x_{i-1} < x < x_{i+1}$. In image processing this method is used to perform transformation from a discrete to

a continuous image by linear or curve fitting. For a given discrete function f the task is to construct a continuous function by linear interpolation.

2.2 Shape Model Construction

In this section, the construction of a statistical shape model is explained and it can be previewed as a subsequent stage of image segmentation. Using statistical shape model is significant in representing object with complex shapes, since it is able to remove variations due to noise. A set of uniformly sampled landmarks using point distribution is used to describe the shape. The model is built using a set of training shapes that have been manually annotated with a set of landmarks given as complex numbers.

2.2.1 Building a Shape Model

In theory, each training shape is defined by a set of N landmarks/points in d dimensions, given the element vector created by concatenating the individual point position, x-and y-coordinates. Each point contains the coordinates of the corresponding landmark point along the shape boundary $\mathbf{X} = (x_1, x_2, \dots, x_N)$:

$$x_j = (x_j, y_j)$$

The order of the points in \mathbf{x} is not important as long as the consistency is preserved during the transforming between shape boundary vector \mathbf{X} and the coordinates (x_j, y_j)

The element vector \mathbf{X} can be defined as

$$\mathbf{X} = (x_1, \dots, x_N, y_1, \dots, y_N)$$

with the corresponding vector

$$\mathbf{X} = (x_1, \dots, x_N, y_1, \dots, y_N)^T$$

2.2.2 Aligning Training data

The training data needs to be aligned and evenly distributed, e.g. using arc-length. Even distribution can be obtained from interpolation and re-sampling.

For s given training images we generated s element vectors. Before performing any statistical analysis it is important to align the training set in order to ensure that the shapes represented in the same coordinate frame. This can be achieved by using Procrustes Analysis approach. Alignment can be made by finding the best similarity transformation (translation, rotation and scaling) that minimize the distance between corresponding points. For two point sets, x^1 and x^2 , this can be solved by minimizing the Euclidean distance between the two point sets.

Orthogonal Procrustes is statistical shape analysis method performed to find the optimal rotation R^* , transformation t^* and/or scaling s^* parameters that give the optimal orthogonal linear mapping. After performing similarity transformation, a comparison between the best transformation shape or model y_i achieved previously and the model achieved with the currently estimated parameters, R^* , t^* and s^* . The aim is to obtain a model that minimizing the Euclidean distance between the two compared models. In other words, the following condition need to be satisfied:

$$\min \sum_{i=1}^N \left\| \mathbf{x}_j^1 - (\mathbf{t} + s \mathbf{R} \mathbf{x}_j^2) \right\|^2$$

The alignment first involves computing the *centroid* or the mean of the coordinates of each chosen sample

$$(\bar{x}, \bar{y}) = \left(\frac{1}{n} \sum_{i=1}^n x_j, \frac{1}{n} \sum_{i=1}^n y_i \right)$$

The next step is to translate the the mean of the coordinates to the origin

$$\tilde{x}_i = x_i - \bar{x} \quad \text{and} \quad \tilde{y}_i = y_i - \bar{y}$$

Let $\mathbf{H} = \sum_{i=1}^N \tilde{y}_i \tilde{x}_i^T$ and use SVD to get homography matrix $\mathbf{H} = \mathbf{U} \mathbf{D} \mathbf{V}^T$.

Then in 2D

$$\begin{aligned} \mathbf{R}_* &= \mathbf{U} \text{diag}(1, \det(\mathbf{U} \mathbf{V}^T)) \mathbf{V}^T \\ s_* &= \frac{\sum_{i=1}^N \tilde{y}_i^T \mathbf{R}_* \tilde{x}_i}{\sum_{i=1}^N \|\tilde{x}_i\|^2}, \\ \mathbf{t}_* &= \bar{y} - \mathbf{R}_* \bar{x} \end{aligned}$$

Performing Procrustes analysis before constructed the model shape is important in order to remove the scale, translation and rotation. The following steps are followed to achieve this purpose:

1. Align each shape to the first
2. Calculate the mean of the transformed shapes (by calculating the mean value for each point)
3. Align the mean shape to the first (to guarantee convergence)
4. Align each shape to the mean shape
5. Update the mean shape
6. Iterate 3 to 5 until convergence (i.e. the mean shape does not change significantly from the previous iteration)

For M training shapes, the mean shape is given by

$$\bar{x}_j = \frac{1}{M} \sum_{i=1}^M x_j^i$$

where $i = 1, \dots, M$ denotes the shape number, and $j = 1, \dots, N$ denotes the point number.

2.2.3 Principal Variations

After performing the shapes aligning, the covariance function of the points and the corresponding eigen-values and eigen-vectors in order to obtain a shape model. To describe the mode of variations in the training shapes the eigen-vectors of the covariance matrix are calculated. Besides, the eigen-values are calculated to describe the significance of the variation in the training data. The modes of variation describe the

ways in which the points of the shape tend to move together. This can be found by applying principal component analysis to the deviations from the mean. The derivation of each shape with respect to the mean shape is given as:

$$dX^i = X^i - \bar{X}$$

The modes of variation are describes by the unit eigenvectors of the covaraince matrix, S :

$$SP^i = \lambda_i P^i$$

The 2Nx2N covariance matrix, S :

$$S = \frac{1}{M} \sum_i^M dX^i (dX^i)^T$$

Assuming that the eigen-values are ordered decreasingly, The eigen-vector corresponding to the highest eigenvalue describes the most likely variation in the training data.

2.3 Shape Model Construction

The point distribution model for every shape can be defined as

$$x = \bar{x} + Pb$$

\bar{x} is the mean landmark coordinates of the training shapes. The 2N eigenvectors represents a new basis $\mathbf{P} = (P^1 \ P^2 \dots P^{2N})$, where each point N has 2 coordinates. There are $K=2N$ eigen-vectors of the covariance matrix and each shape is associated with its eigenvalue λ_i . \mathbf{b} denotes the coordinates of the new basis \mathbf{P} , and can be interpreted as the amount of variation in the different variation modes. [2]

In order to avoid capturing the noise in the training data, a usual practice is to select the first t modes of the variation. The first t modes describe the relevant variations in the training dataset. The shape x can be approximated by

$$x \approx \bar{x} + P_t b_t$$

To select the number of modes look at the geometric appearance of each mode; (Usually the first modes describes the relevant variations in the training set and the last modes only capture noise. Another approach is to calculate the energy preserved in each shape. Each shape contributes to λ_i/E , where E is the total energy:

$$E = \sum_{i=1}^{2N} \lambda_i$$

By plotting the cumulative energy for the eigenvalues and selecting the the number of modes t such that $E/E_t = 95\%$. Each eigenvalue gives the variance of the data about the mean. The total variance is the sum of all eigenvalues.

2.4 Mode visualization

To visualize the different principle modes, the mean shape \bar{x} is plotted together with mean shape plus/minus a suitable multiple, k , of each mode in order to generate new examples of the shapes:

$$\bar{x} = \pm k p_i$$

where k is given as $k = 2 * \sqrt{(\lambda_i)}$. Besides, other alternatives can be plotted such as: $\bar{x} - \sqrt{\lambda_i} p_i$ and $\bar{x} + \sqrt{\lambda_i} p_i$.

2.5 Segmentation with Shape Model

The constructing shape model can be used to generate a representation of a general shape. After constructing the shape model, as described above, with the mean shape \bar{X} and shape modes \bar{P}_t , the shape parameter vector, \bar{b}_t and the pose parameters s , R and t can be estimated. The shape parameter vector is considered as the least-squares solution and estimated by multiplying $\mathbf{X} - \bar{x} \approx P_t b_t$ with the transpose of the basis matrix:

$$P_t^T (X - \bar{x} \approx P_t^T P_t b_t$$

\Rightarrow

$$b_t = (P_t^T P_t)^{-1} P_t^T (X - \bar{x} = P_t^T (X - \bar{X})$$

The shape could be written as:

$$Y = T_{s,R,t}(\bar{X} + P_t b_t)$$

where the transformation for each point $T_{s,R,t}$ is given by

$$y = T_{s,R,t}(X) = sR(x + t)$$

In order to fit the shape model to edge data, we start by initializing the pose parameters s , R , t and b_t by transform the landmarks to mean using the transformation algorithm described in sec.2.2.2. This gives the initial $\mathbf{X} = sR\mathbf{X} + t$. The procedure is:

- At each landmark, find the closest edge point orthogonal to the contour, and replacing with the closest point on boundary
- Calculate b_t and limit it to $\pm 3\sqrt{\lambda_i}$ for stability
- Find the new landmarks in the shape space based on the current estimate of b_t , and transform back to image space; $X = R^T(X - t)/s$
- Find closest point on original contour to the new landmarks and use these as new landmarks
- After initializing, updating the transformation parameters in image space and shape parameters in shape space, the iteration loop is finished as long as we reach convergence.

3 Methods

In the first part of this project, a shape model algorithm is developed. In order to create the shape model manual segmentation of kidneys is performed on some renal scintigraphy images.

3.1 Manual Segmentation

The manual segmentation of the provided images is performed by a physician and the coordinates of the created landmarks around one of the patients' kidneys are stored in **man_man_seg**. The landmarks are represented by complex numbers. The purpose of this part of the project is to re-sample these landmarks into 14 equally spaced landmarks around the kidney, where landmark number 1 coincides with the given landmark number 1 for the first patient. Linear interpolation is used to interpolate between the given landmarks.

In order to perform linear interpolation, a resampling function has been created. The function takes the 40 landmarks' x-and y-coordinates as inputs and create 14 new x-and y-coordinates as outputs after performing linear interpolation based on Euclidean distance between the landmarks. Matlab build-in function *interp1(sample_points, corresponding_values, xv)* is used which returns the interpolated values of a 1-D function at specific query points, xv, over the range of the corresponding values. First, the sample points/landmarks x-and y-coordinates are extracted and stored as a vector. The difference between the adjacent elements of x-coordinate and y-coordinates is calculated using the built-in function *diff()*. The Euclidean distances are then calculated according to:

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

The cumulative sum of the elements defined by *d* is calculated. These values represent the corresponding values that are used to return the interpolated values at the query points defined between 0 and the the last element in the cumulative sum vector. Finally, *interp1()* function is used to calculate the interpolated values, and the new equally spaced 14 landmarks x-and-y coordinates are calculated. Before using the resampling function, the x-and y-coordinates of the landmarks are extracted and sorted in descending order using *circshift()* to make the lowest point first in the list.

3.2 Shape Model

After performing linear interpolation, the next step is to construct the shape model using 20 manual segmented 20 kidney images. In summary, we have 14 (landmarks) x 40 (kidneys), where columns 1-20 are the right sided kidney of patient 1-20 and columns 21-40 are the mirrored left sided kidneys of patient 1-20.

The purpose is to only model the shape of a kidney, therefore the first step is to remove the scale, translation and rotation. Procrustes analysis is used to find similarity transformation as described in sec.2.2.2. We start by creating an algorithm for shape alignment according to following steps:

1. Align each shape to the first shape. The first shape is initialized and defined as `aligned_mean_shape`.
2. Store first shapes' 14 landmarks in a 1x1 cell
3. Align each shape to the first shape `'aligned_mean_shape'`
4. Perform transformation
5. store each transformed/aligned shape in a new cell

6. Calculate the mean of the transformed/aligned shapes, and the coordinates of the aligned shapes are stored.
7. Align the mean shape to the first
8. Align all shapes to the mean shape
9. Update the mean shape

Next, the aligned shapes are plotted in the same figure before proceeding with the shape model in order to verify that the translation, rotation and scale are eliminated.

After aligning the shapes calculate the covariance function of the points and the corresponding eigen-values and eigen-vectors in order to obtain a shape model. In order to calculate the covariance the steps described in sec.2.2.3 are followed. The corresponding eigenvalues and eigenvectors are calculated using Matlab built-in function *eig()*. The eigenvalues are sorted in decreasing order using built-in function *flip()* before plotting them for visualization. The eigenvectors are used as a new basis, N points with 2 coordinates each. The eigenvector corresponding to the highest eigenvalue describes the most likely variation in the training data.

In order to visualize the principal modes, a function is created by following the steps described in sec.2.4. lastly, the number of modes are selected by calculating the energy of each shape as described in sec.2.3.

3.3 Segmentation with a Shape Model

The constructed shape model is now used together with segmentation based on thresholding to segment kidneys in new images. The segmentation is performed on the right sided kidney since the model shape is right sided. Matlab built-in function *bwboundaries()* is used for the initial segmentation to find the coordinates of the edge of the segmented kidney. The boundaries and the landmarks of the last five images is performed. The lowest point of the boundary is sorted as the first element of the boundary to make sure that the first landmark is at the bottom of the kidney. The resampling method from the first part is used to resample the edge/boundary into 14 landmarks. The original and the resampled boundaries of the five last images are plotted.

The shape model for the last five images is estimated using the procedure described in sec.2.5. The true, mean and approximated shapes for the five last images are plotted.

4 Results

4.1 Manual Segmentation

After performing the steps mentioned in section 3.1, the resampled landmarks with the 14 equally spaced landmarks around the kidney are plotted together with the original landmarks for comparison. The result is shown in Fig.1.

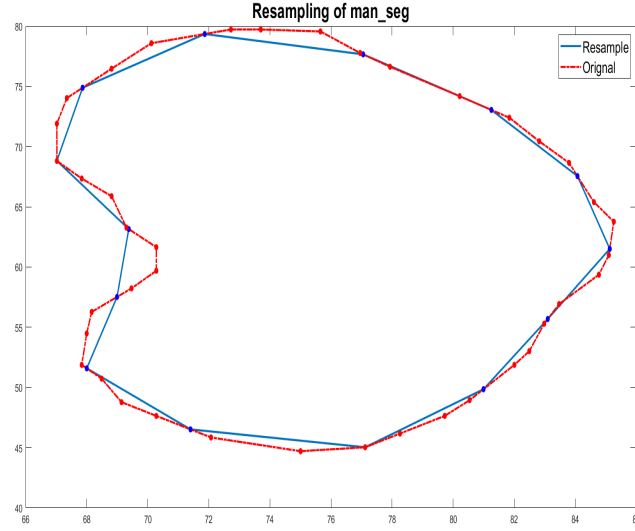


Figure 1: The resampled 14 landmarks together with original landmarks

Fig.2 shows a comparison between the original kidney image of patient 1 with the 40 landmarks in comparison to the 14 equally spaced resampled landmarks segmentation boundary.

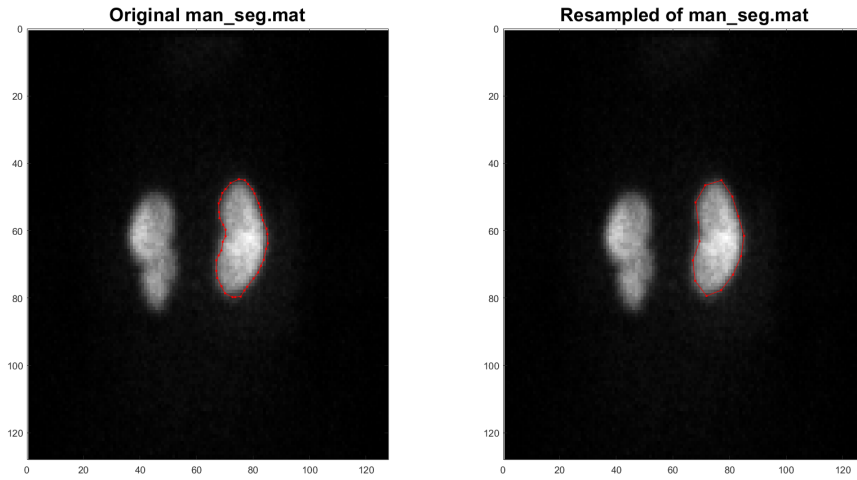


Figure 2: Kidney image of patient 1 with the segmentation boundary of the original landmarks in comparison to the boundary achieved after linear interpolation.

The left kidney is then mirrored to get it in the right position in the image. Fig.3 shows the right and left kidney for patient 1 after linear interpolation.

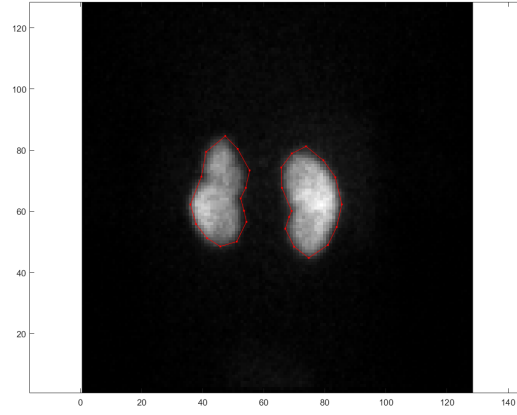


Figure 3: The left and right kidney with the segmentation boundary after linear interpolation.

4.2 Shape Model

The first step was to perform shape alignment and using generalized Procrustes analysis to eliminate translation, rotation and scale. The resulting aligned shapes can be seen in Fig.4

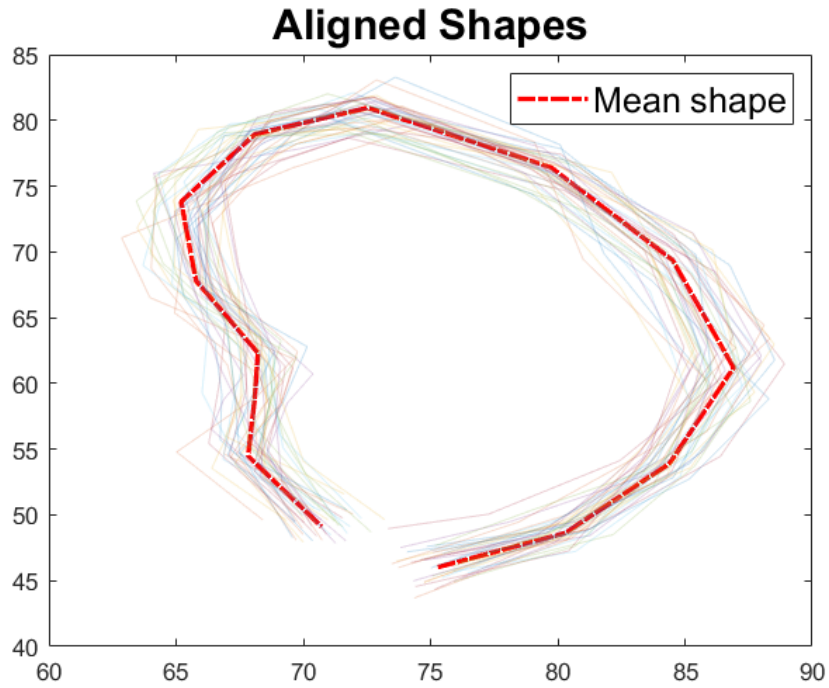


Figure 4: The aligned shapes together with the mean shape after performing Procrustes analysis to eliminate translation, rotation and scale.

After aligning the shapes calculate the covariance function of the points and the

corresponding eigenvalues and eigen-vectors are calculated. Fig.5.

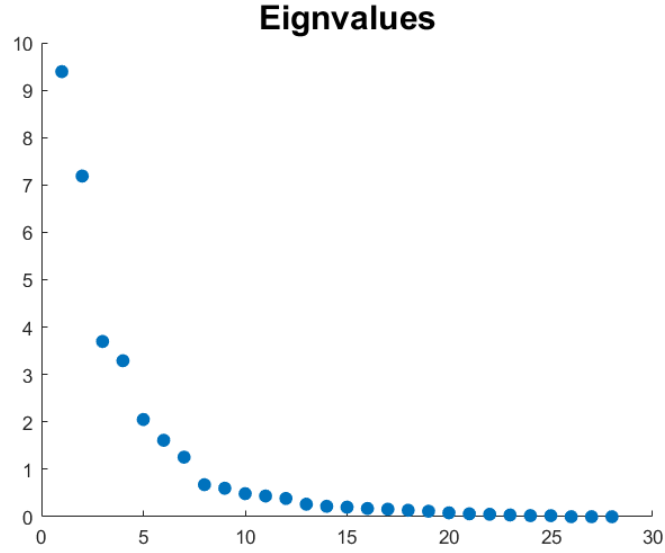


Figure 5: The eigenvalues achieved by first calculating the covariance function. The eigenvalues are sorted in descending order.

14 different principle modes in the shape model are then plotted and the energy contribution of each shape is calculated. The 14 principle modes are shown in Fig.6. As it can be observed, the mean shape does not change significantly from the previous iteration after 12 principle modes. Therefore, 12 modes can be suitable for the shape model.

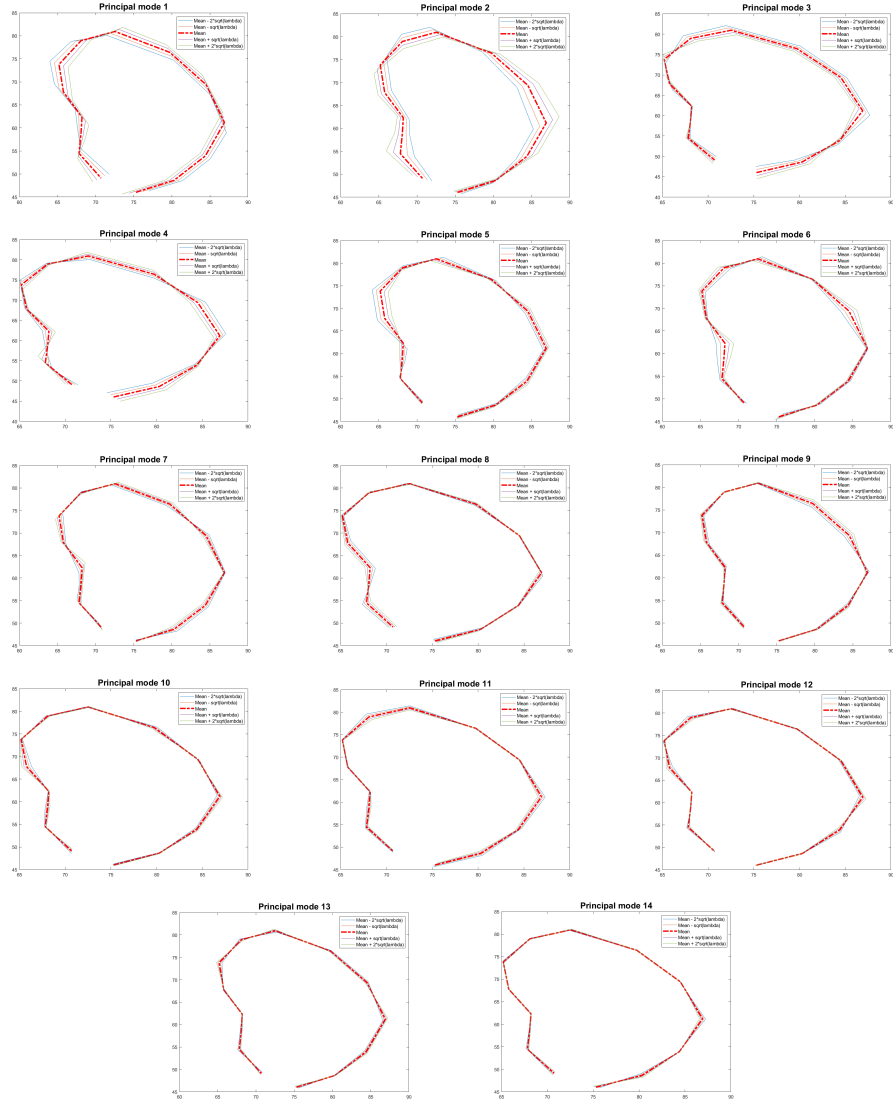


Figure 6: The 14 different principle modes in the shape model. The dashed curve corresponds to the mean mean shape and the other curve to the modes of variation with $k = \pm 1$ and ± 2 .

Only the first 7 modes that actually describe visible shape variations are shown and described further below.

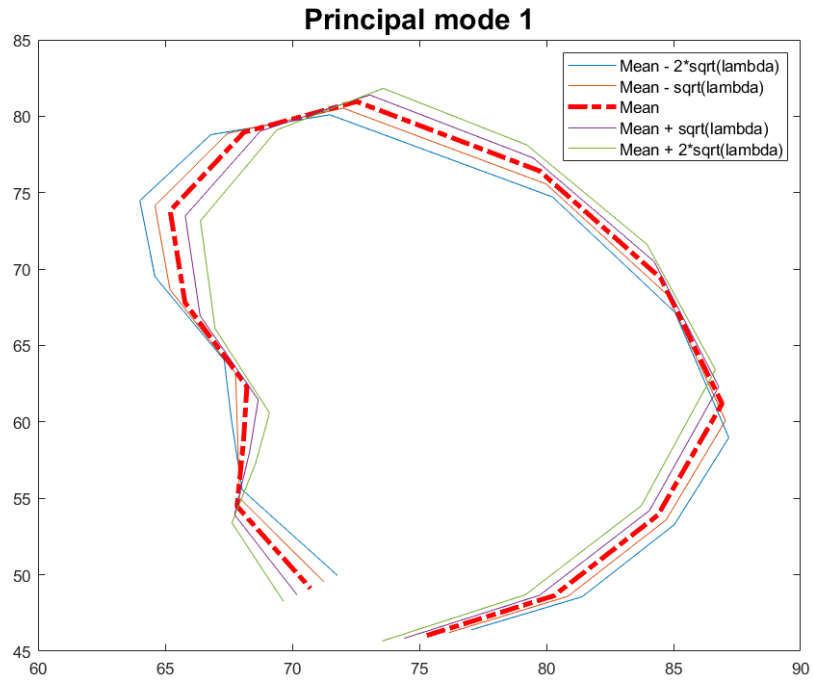


Figure 7: Mode 1. This mode seems to represent variations in the roundness of the kidney. Generally speaking, this mode could be interpreted as the roundness / compactness of the kidney.

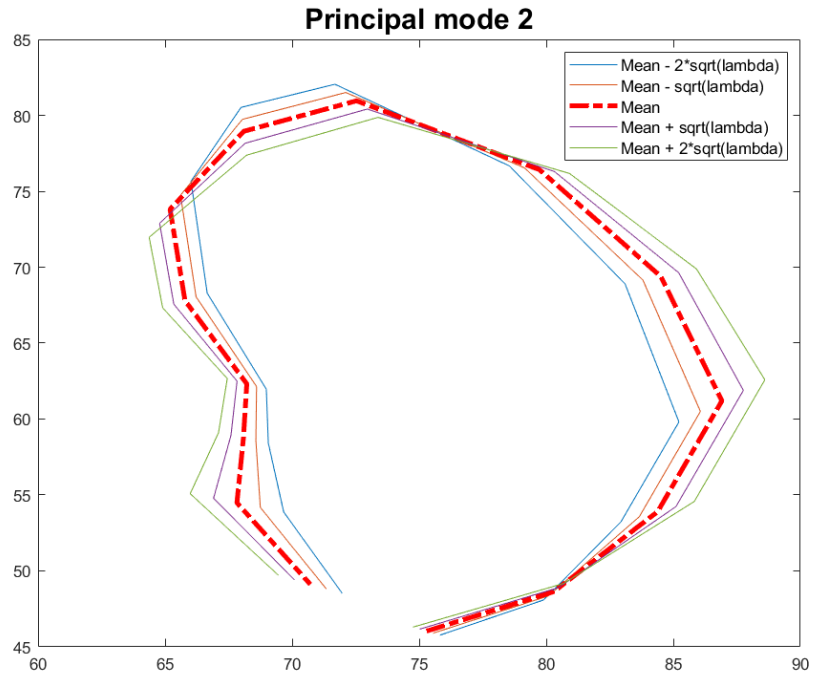


Figure 8: Mode 2. In general, this mode could be interpreted as the prominence of the lower left part of the kidney. The upper left part seems to come along with it.

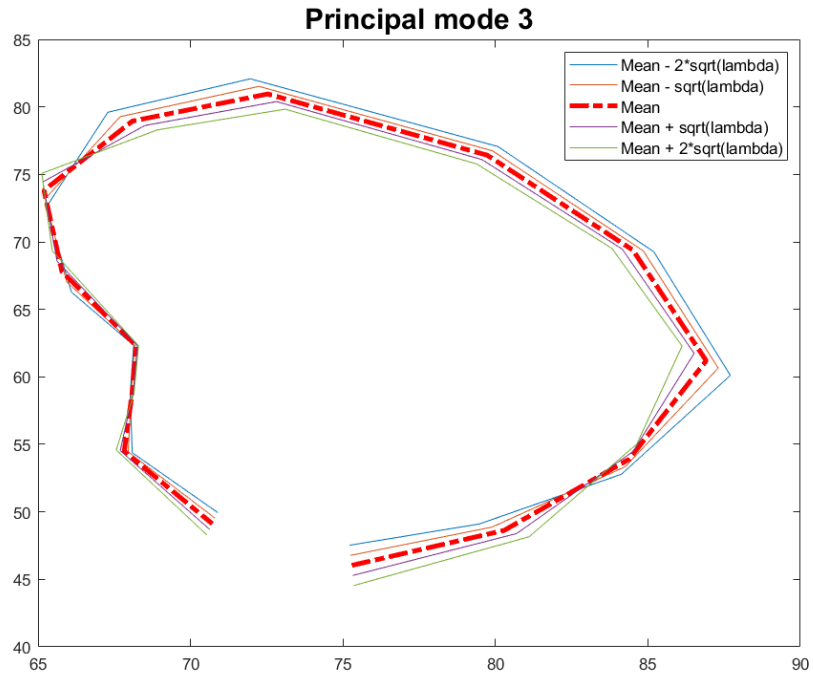


Figure 9: Mode 3. This mode seems to describe how straight / flat the left side of the kidney is. Generally, this mode could still be interpreted as how much the kidney resembles a cashew shape (on the left side).

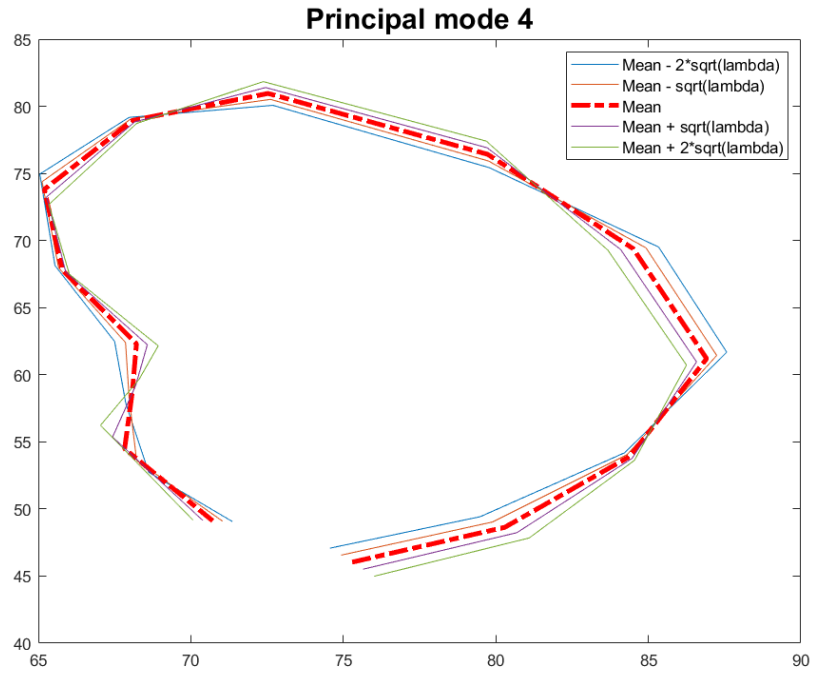


Figure 10: Mode 4. This mode is very similar to the mode 3 and also represents how cashew-like the kidney is shaped. Here, it however does not take into account as much of the left side as the previous mode.

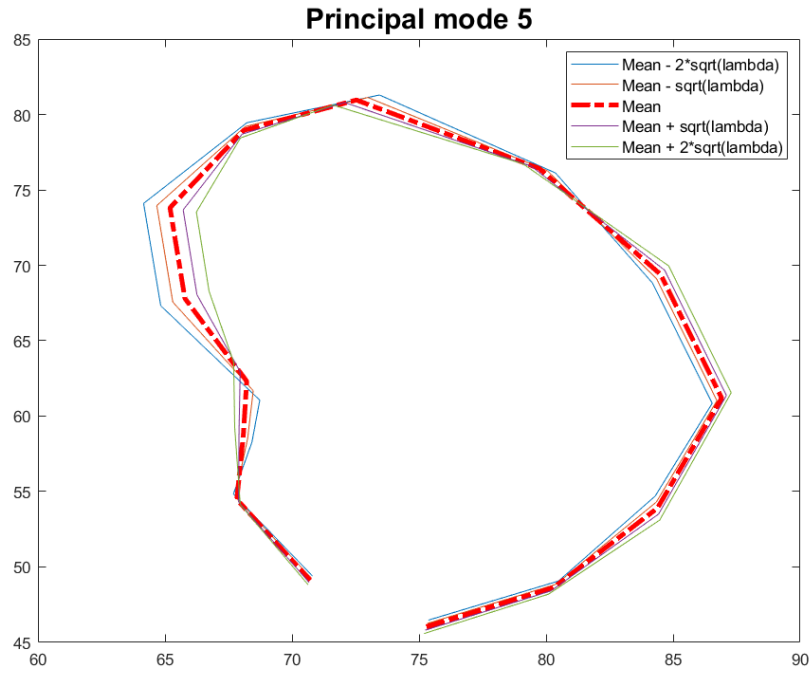


Figure 11: Mode 5. In all variations here, the kidney has a cashew shape on the left side.

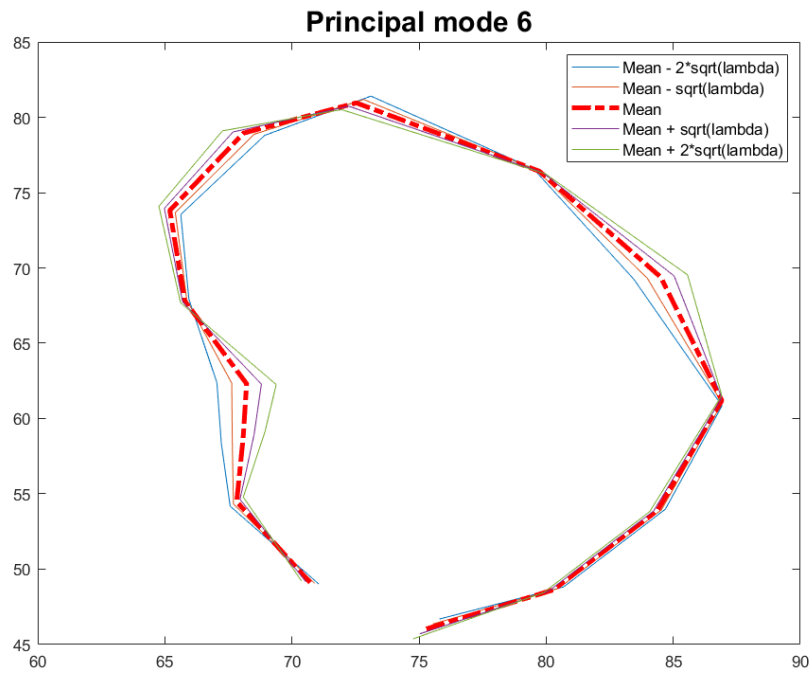


Figure 12: Mode 6. The shape of the kidney is very similar here for all cases.

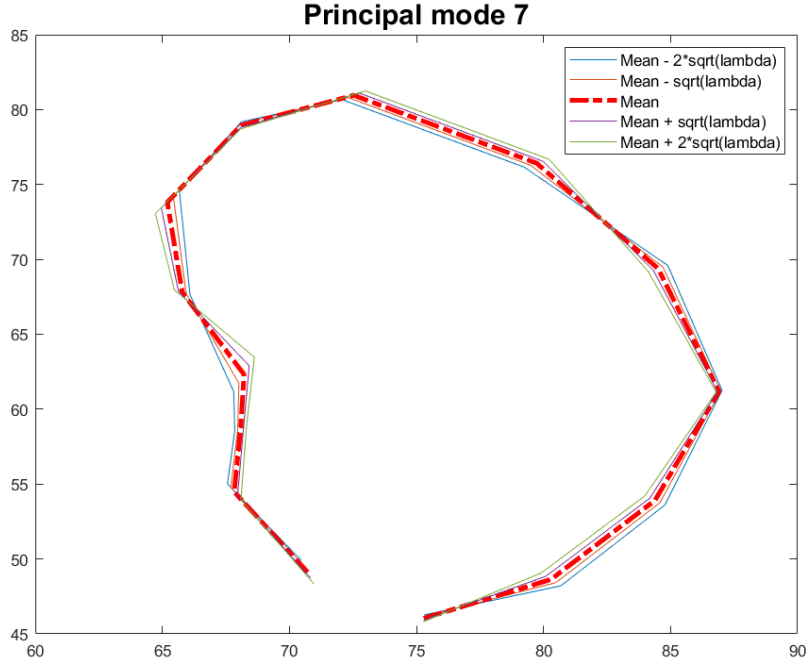


Figure 13: Mode 7. Here, there are some minor changes in all landmarks. However, it is hard to interpret this into a general meaning of general shape variation.

In order to determine the number of modes that are suitable for the constructed shape model, the relative contribution for each shape is considered by calculating the sum of all eigenvalues and plotting the contribution of the modes. Fig.14 shows the calculated cumulative energy over the eigenvalues. The goal is that the constructed shape model $x \approx \bar{x} + P_t b_t$ generates examples similar to the training set. This can be achieved by ranking the modes according to their energy content, and choosing the the number of modes with the a relative energy $> 95\%$. The number of modes suitable for the shape model is 12, according to Fig.14.

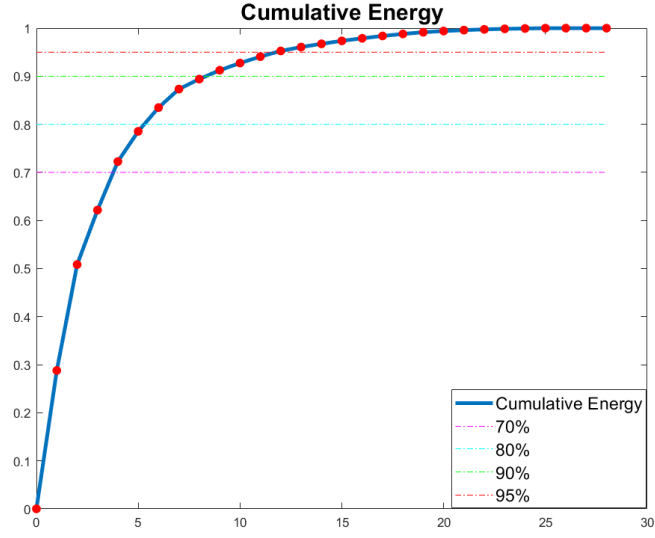


Figure 14

4.3 Segmentation with a Shape Model

After using the build-in function *bwboundaries()* to find the edge of the segmented kidney, the resampled algorithm described in sec.2.2.2 is used to visualize the original and the resampled boundary for the five last images. The result is shown in Fig.15.

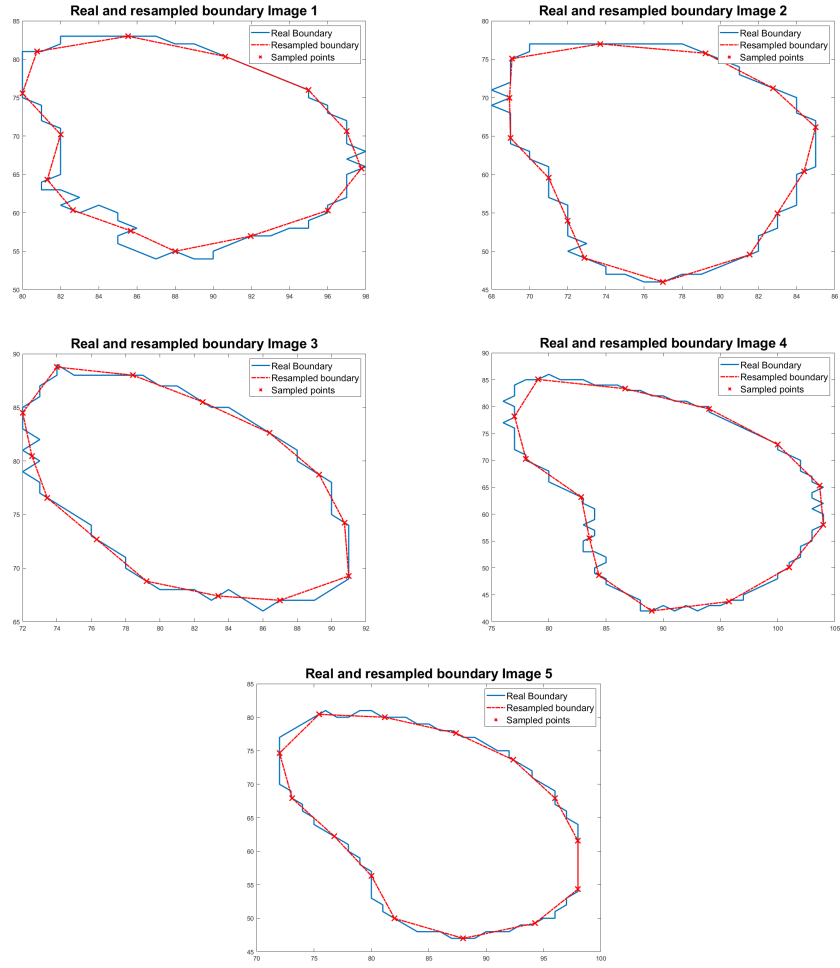
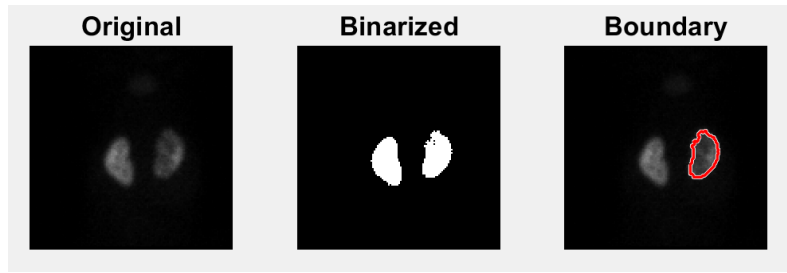
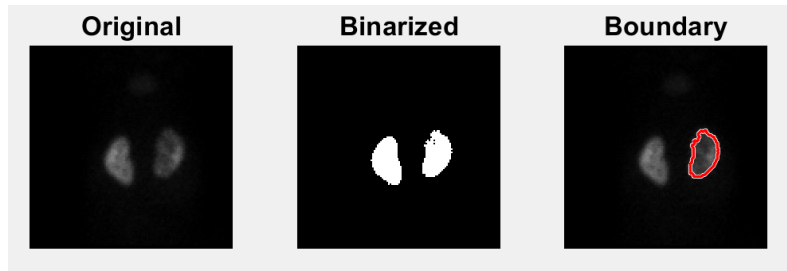


Figure 15: The original and the resampled edge boundary for the last five kidney images.

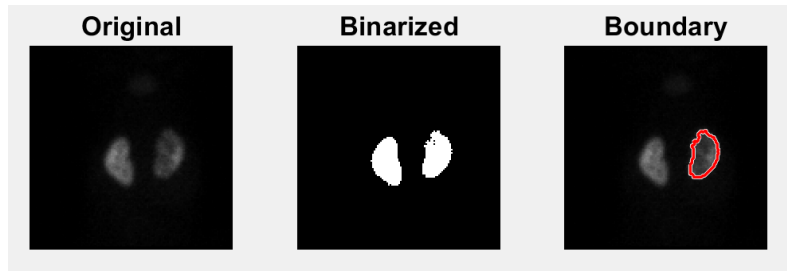
The resulting images after applying threshold segmentation, and the resampled boundary for the last five kidney images are shown below.



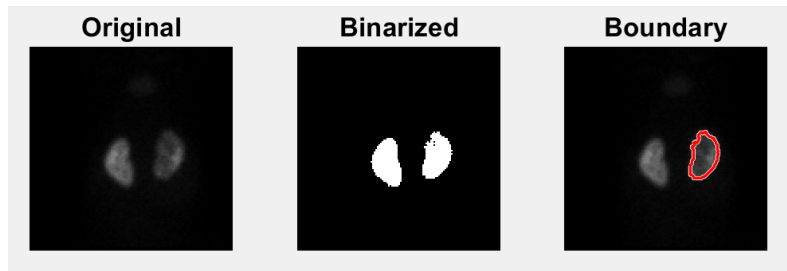
(a) Image 1



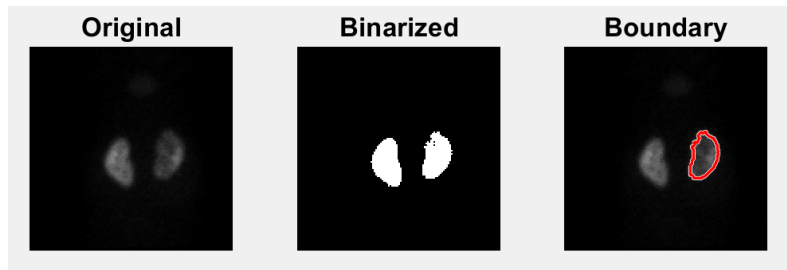
(b) Image 2



(c) Image 3



(d) Image 4



(e) Image 5

Figure 16: The five last kidney images are shown as original image, after thresholding and with the resampled boundary to segment the right side kidney.

After using the constructed shape model the approximated boundary together with the mean and true boundary is estimated. The segmented last five images using the shape model are shown in figures below. The approximated boundary together with the original and mean shape boundary for the five last kidney images. In each image, the segmentation boundary is plotted around the real kidney images, as viewed to the right of each image.

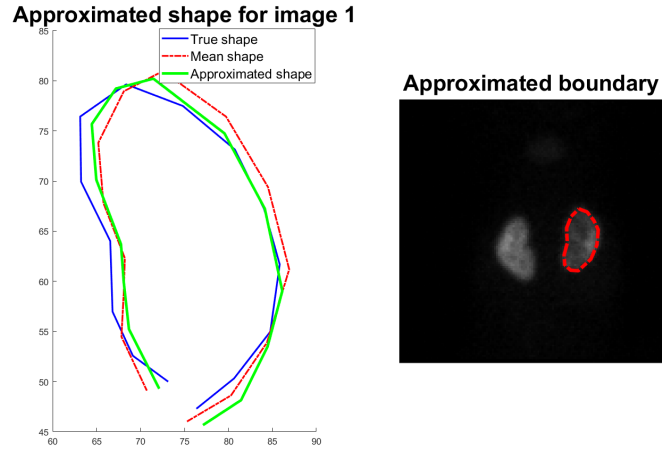


Figure 17: Image 1 with approximated boundary, true and mean shape.

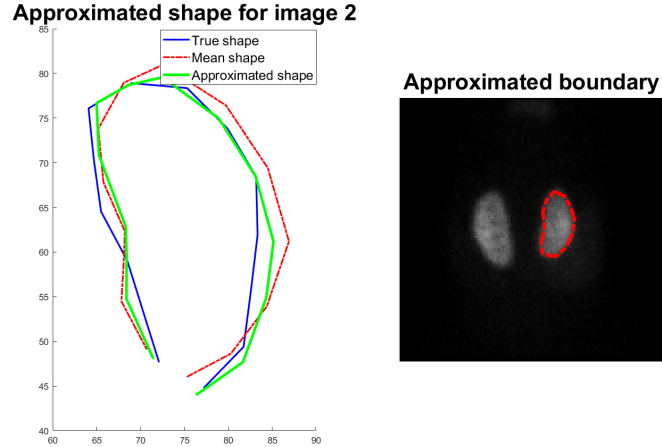
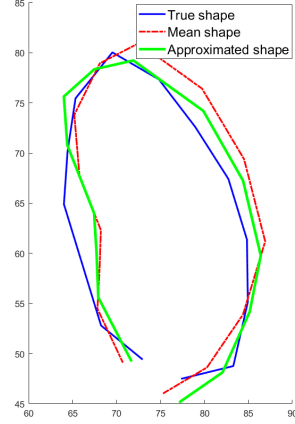


Figure 18: Image 2 with approximated boundary, true and mean shape.

Approximated shape for image 3



Approximated boundary

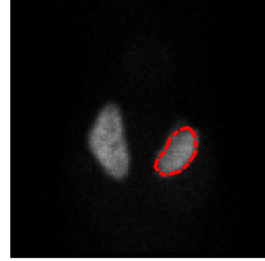
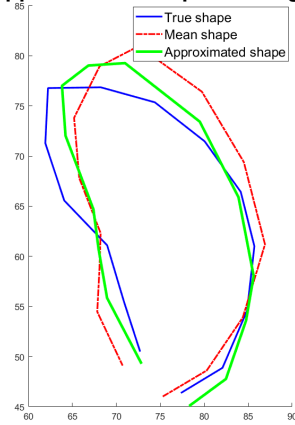


Figure 19: Image 3 with approximated boundary, true and mean shape.

Approximated shape for image 4



Approximated boundary

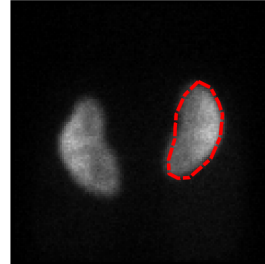


Figure 20: Image 4 with approximated boundary, true and mean shape.

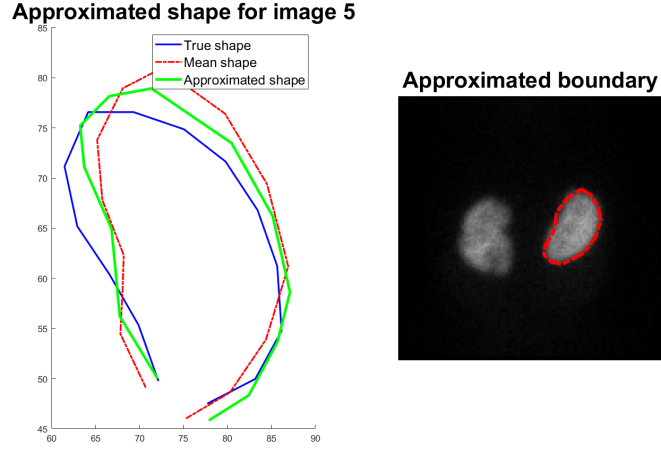


Figure 21: Image 5 with approximated boundary, true and mean shape.

5 Conclusion

In e.g. many medical applications, capturing the shape variability is important, and the constructed shape model should resemble a large portion of the shape instances which can be achieved by using statistical analysis. Medical image segmentation is a difficult task due to the high variability in medical images. The difficulty considered is due to the modes of variation in human anatomy and the many different modalities that used to create the medical images.

The modes of variation plays an important role in visualizing the variation in the data and usually used in principle component analysis (PCB). Modes of variation are a set of vectors centered at the mean of the training data and used to study the variation in the given data. Modes of variation illustrates the decomposition of the variation patterns in the data into descending order of eigenvalues. Therefore, the modes of variation is an efficient description of the variation around the mean sorted by the eigenvalues.

The constructed shape model was then used to segment unseen kidney images to investigate the performance of the constructed model. As it can be seen in Fig17-21, the approximated boundary differs slightly from the true boundary. As it has been considered, the shape model depends on the number of modes. Therefore, choosing the right number of mode is significant in order to obtain as correct shape model as possible. One possibility is to set constrains on the mode parameter vector b_t to be in the range $\pm 3\sqrt{\lambda_t}$ for stability. One can also use Kaiser's rule to decide the number of principal components, where only the principal components with variances exceed 1 should be retained. The principle components with variance less than 1 contains less information than the original variables. Another method to use is to examine so called scree plot, which is the principal component vs. proportion of variance. We look at the point at which the variance is dropped off, so called elbow.

References

- [1] T. F. Cootes (May 2004), Statistical models of appearance for computer vision (PDF)
- [2] Yuanwei Li et al (September 2017), Fully Automatic Myocardial Segmentation of Contrast Echocardiography Sequence Using Random Forests Guided by Shape Model