

Positioning on Magnetic Resonance Tomography

Christian Wachinger, Andreas Keil, Nassir Navab

September 1, 2005

Two current developments in the latest models of magnetic resonance scanners are the reason for this project. The first one is that by using the new scanner the maximum exposure dose is already reached after taking a few images. The second one is that the new scanner have a movable bed on which the patient lies on. This led to the idea to improve the calculation of the underlying SAR model to better fit it to the patient. For the improvement, low resolution MR scans should be used which are taken while moving the patient to the target region. In the scope of the project we were mainly focusing on the positioning i.e. finding out if the patient lies there head or feet first and which part of body is currently scanned. Therefore different standard imaging algorithms and the principal component analysis (PCA) were applied.

A realtime decision based on multiple criteria whether the patient moves head or feet first into the scanner is possible. The detection of the head and feet was feasible with the PCA, except for the special cases. But nevertheless, no wrong decision was made. This was possible as a certain trust value was calculated that expressed the confidence in the decision. A criterion for the lung detection was developed that performed extremely well and robust. So, if the patient moves in head first and the head detection failed, when the lung is scanned the right decision is made. Also a criterion for the detection of the critical neck region was developed, based on the lung and head recognition. Moreover, the high potential of learning was shown, supporting the assumption that it may work in a final product if exhaustive training is performed.

Contents

1. Introduction	3
2. Standard Imaging Algorithms	3
2.1. Binary Threshold	4
2.2. Area	4
2.3. Perimeter	4
2.4. Region Growing on Background	5
2.5. Connected Components	6
3. Principal Component Analysis	6
3.1. The Algorithm	7
3.2. Classification of Slices	10
3.3. Evaluation Criterion for Head or Feet First	11
3.4. Conclusion for PCA	13
4. Results of Positioning	16
4.1. Head	16
4.2. Neck	16
4.3. Chest	16
4.4. Abdomen	16
4.5. Hip	17
4.6. Upper Leg	17
4.7. Feet	17
5. Further Ideas	17
5.1. Pattern Matching	17
5.2. Bounding Box around Patient	17
5.3. Independent Component Analysis	18
5.4. Support Vector Machines	18
6. Conclusion	18
A. Covariance Matrix	19
B. Mahalanobis Distance	20

1. Introduction

The reason for this project is, that with the latest generation of magnetic resonance scanners the maximum exposure dose is reached quite fast. This limits the amount of slices that can be taken and therefore the region that can be scanned. This is not optimal because one is interested in scanning large regions in a high resolution.

The number of slices that can be taken without exceeding the maximum exposure rate is calculated with a SAR model. The idea, how to solve the problem, is to get a more accurate SAR model, better fitting to the patient. This would allow to take more slices from less dangerous regions and prevent very sensitive ones like the neck. So, the question is how to get a more accurate model. The proposed approach is to make a low resolution scan during the patient is moved to the targeted region within the scanner. This is possible, as the latest generation of scanners has a movable bed where patient is lying on. This kind of pre-scan is completely transparent to the patient as he has to be moved in the scanner anyway. During this scan, additional information like the patient is lying head or feet first in the scanner or the position of the neck, if detected, is used to enhance the SAR model. So, to wrap it up, the subject of the project was to check whether it is possible or not to detect the different parts of the body like head, neck, chest etc. from a few low resolution MR slices.

Special focus was set on the detection whether the patient moves head or feet first into the scanner. The task might sound quite basic, but the difficulty lies within the robustness of the algorithm. There are at least eight different positions in which the patients can come into the scanner - lying on their back, on their left side, on their right side, on their stomach - each head or feet first. Next to these normal cases, a sheer infinite number of special cases exist like crossing legs, hand above the head etc.

To get the detection done, we tried a whole bundle of different methods that can be grouped into more or less standard imaging algorithms and a more complex approach, the principal component analysis. In the following sections firstly the methods are described and then the results from the positioning are listed. Also some further ideas promising good results were stated but it was not possible to cover them in detail within the project phase.

The project was done in cooperation between Siemens MR and the chair of Computer Aided Medical Procedures (CAMP) at the Technical University of Munich (TUM). For the project several test scans of human bodies were taken which were used to try out the different methods. The slices were taken in a resolution of 64*64 pixels.

2. Standard Imaging Algorithms

In this section, different standard imaging algorithms and their application are shortly described. The idea is to have several criteria in combination with the PCA to detect special cases and give more confident results. In addition, they can be used for recognizing special cases.

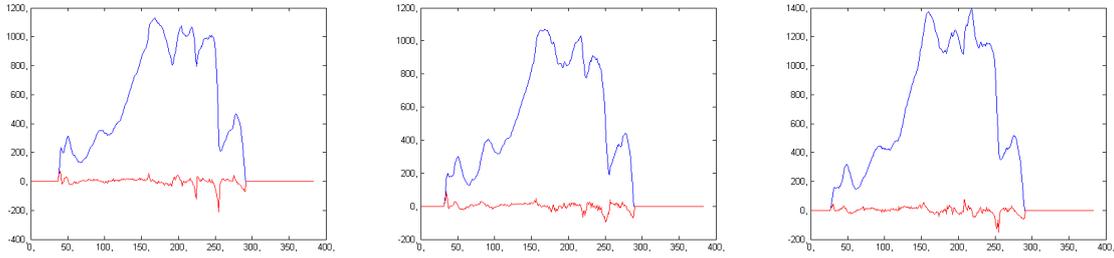


Figure 1: Plots of the Area (blue) and the Area Gradient (red). Feet: left, Head: right.

2.1. Binary Threshold

Binary Thresholding was mostly applied before using one of the others to segment out the volume. The threshold value was chosen with the help of the image histogram. As all volumes showed a very similar distribution we were using a pre-set threshold value of 26.0 that worked well.

2.2. Area

After applying the binary thresholding to the volume, the area of each slice is simply calculated by counting the pixels showing the patient. A few plots of the area along the volume can be seen in Figure 1. The area values are plotted along the y-axis and the z-axis through the body is plotted along the x-axis. All area curves have a very similar shape with a head peak, a neck valley, a high slope to the chest, three peaks in the trunk, a plateau in the knee section, a valley before the feet begin and a foot peak. The area is not suitable to distinguish between head and feet as they have similar values but it's good to distinguish head and feet from the trunk.

The disadvantage of the area criterion is it works with absolute area values so there may be problem with children. If a complete scan would be available a rescaling could be processed. But perhaps the entered weight of the patient could work as a scaling factor too.

Additionally to the area also the gradient of the area function can be used. A few plots can be seen in Figure 1. The highest area gradient magnitude is always in the neck area, so it possible to detect it. The problem is that the absolute values of the gradient differ from volume to volume, so a thresholding with a given value does not work. There are also high gradients at the beginning of the head and the end of the lung. Having a sequence of these three high gradients with their characteristic distance to each other, this could be used to detect head first.

2.3. Perimeter

On the segmented image, the perimeter was calculated for each connected component. The initial idea was to use the comparison between the area of the perimeter and the area of the patient to detect the lung. Unfortunately it has not worked well as can be seen in Figure 2. For the detection, the differences between the two curves should be low except in the the lung

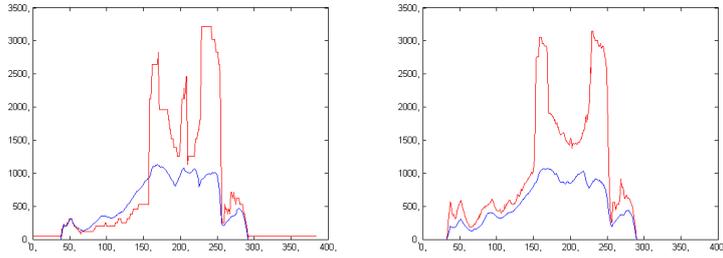


Figure 2: Plot of Area (blue) and Area of Perimeter (red). Feet: left, Head: right.

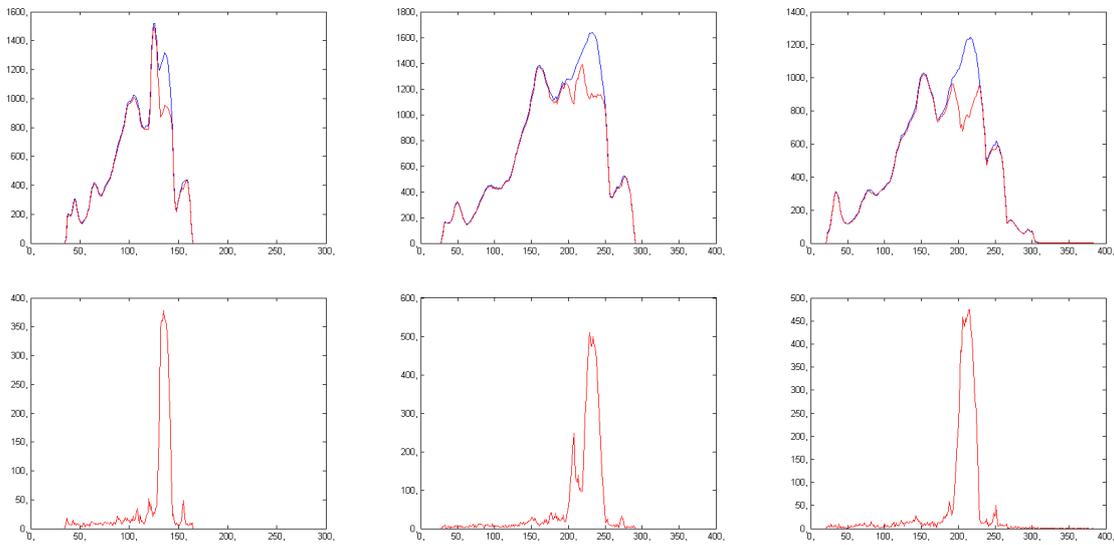


Figure 3: Plot of Area (red) and Area of Hull (blue) and the Difference between the two.

area where high differences are desired. One can see that the curves do not match very well. Although there is the desired large difference in the lung area, there is also a large difference in the hip area making a distinction impossible.

2.4. Region Growing on Background

As the perimeter was not working to detect the lung, we applied another method to better segment out the hull of the person. Therefore region growing was done on the background. Region growing can be done on the whole volume but it is better to do it slice wise as it is more robust to unexpected holes. After segmenting out the background an inversion shows the hull of the person. Then a comparison between the area of the hull and the normal area has to be done to find the lung. The detection works extremely well and is also very robust. Some plots of the areas and their differences can be seen in Figure 3.

It is also not difficult to develop to a decision criterion as the difference between the two

areas has a very nice peak in the lung area. A simple threshold with a value of 300 worked on all volumes. It also works for the special case that is plotted in the third column in Figure 3. With the help of the lung detection it is possible to find out if the patient moves head first into the scanner although the head detection may have failed. It can also be regarded as a warning signal before entering the critical neck region if the patient is scanned feet first.

2.5. Connected Components

A last standard method that was used is the calculation of the connected components. It is not possible to make a detection only depending on the connected components because it is not robust enough. It works as an auxiliary criterion supporting decisions and finding special cases. One criterion, for example, is to compare the area of the connected components on a slice. If there are two components having the same area, it supports the decision for feet. If there is just one component with the typical head area, it supports head. If there are two components with largely varying area, it supports special case.

3. Principal Component Analysis

During the project, the focus was set on the application of the PCA [6]. The way the PCA was applied is a bit uncommon but it is a standard technique in the area of face recognition see [7]. Normally, the PCA is used in medical imaging to find out the main body axes, having to deal with three dimensions. In our case we have number of pixels per slice many dimensions resulting in 256 ($16*16$) dimensions. One might wonder because of 256 dimensions and not 4,096 ($64*64$). The reason is that we sampled the images down before running the PCA. The reduced dimensionality is sufficient as we are not interested in fine structures. Moreover the computational complexity is lowered.

The reason for applying the PCA was to achieve a data reduction and therefore concentrating on the most important information in the slices. Regarding each slice as a vector with 256 dimensions it gets transformed to a lower dimensional space. The number of dimensions of the lower dimensional space depends on the data reduction one wants to achieve and has to be determined problem specific. Although one is interested in data reduction, the loss of information should be low. Therefore a coordinate system has to be found in which most information is conserved though just a few dimensions remain. The new coordinate system can be regarded as the data set intrinsic coordinate system. In Figure 4 you can see an example in 2D with the original coordinate system in black, the data set in red and the new coordinate system in green.

After mapping the slices to the low dimensional space a classification has to be done. Classes for each part of the body, one is interested in detecting, have to be defined. Having these pre-defined classes it is now possible to find out to which class a slice belongs to. We used a statistical distance measure - the Mahalanobis distance - to get the membership to a class.

Finally, to make a decision which part of the body could be below the scanner a decision criterion has to be found. An iterative one that processes not just one, but instead, several following slices works best. Also not just the distance to the class it belongs to is taken into

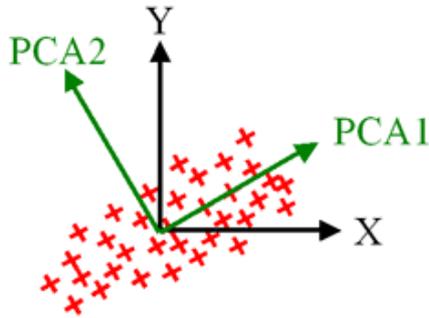


Figure 4: Principal Components for a given Data Set

consideration but also the distances to the other classes.

3.1. The Algorithm

In this section the PCA algorithm to get to the lower dimensional space, is described in detail. The PCA starts with a set of data points or vectors. The vector coordinates are related to the standard coordinate system. The aim is it now to find a coordinate system in which it is possible to erase some dimensions of the vector without losing to much information. So a coordinate system is searched where the first coordinates cover almost all the information and the last ones are less significant. This can also be seen as finding coordinate axes on which the vectors have a high variance, see also Figure 4.

To get the axes of then new coordinate system one has to do an Eigenvalue Decomposition of the covariance matrix of the vectors. For more information about the covariance matrix see Appendix A. The Eigenvectors from the decomposition form the axes of the new coordinate system and the Eigenvalues reveal information about the variance on each axis. So one is interested in choosing the Eigenvectors related to the highest Eigenvalues. Coming back to face recognition, the Eigenvectors are also titled as Eigenfaces because they look like faces again. Analogous the Eigenvectors in our case can be titled as Eigenslices. Some Eigenslices are displayed in Figure 5.

When the coordinate axes that span the new coordinate system are calculated, the matrix that maps each vector into that system has to be constructed. The construction of the projection matrix is quite basic as the Eigenvectors just have to be put in row-wise order into a matrix. Then a standard matrix-vector multiplication performs the projection.

After explaining the algorithm let's focus now on the most important question in this section, the number of principal components (Eigenvectors) to choose. We were analyzing the number of principal components in the range from 10 to 100. To evaluate the performance of each number of principal components we were using three different criteria: the percentage of the variance that is covered, a back projection of the slices to see their quality and the performance that is achieved by the classification.

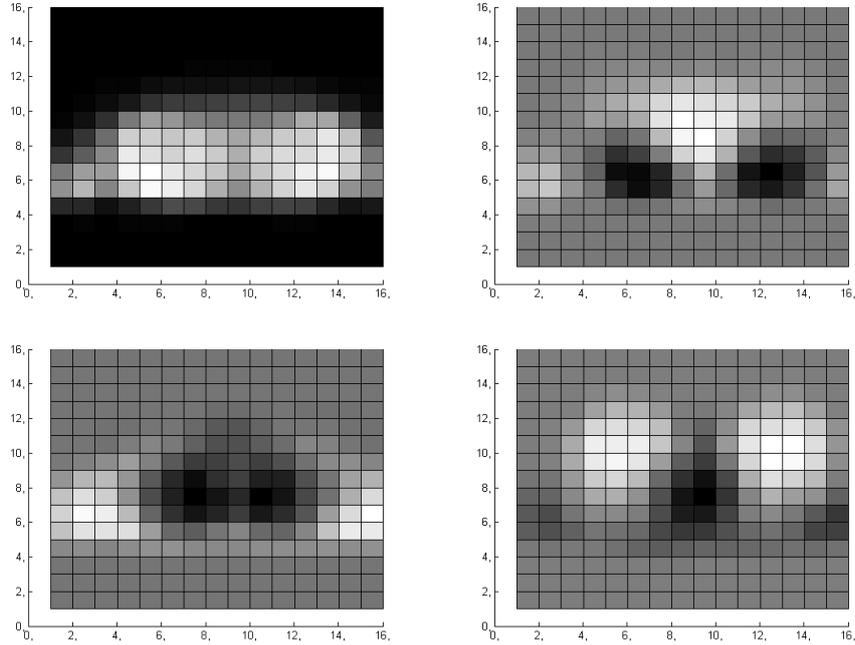


Figure 5: Four most significant Eigenslices

The percentage of the covered variances is calculated by summing up all the variances of the principal components that are used and dividing them by all the variances. In [4] was stated that most times so many components are chosen to get a covering of 70% to 90% of the variances. But in the case of quite similar structures in the images a cutoff above 90% would be appropriate. The covered percentages depending on the number of principal components chosen can be seen in Figure 6. It is remarkable that already with 10 principal components 87.3% of the variances are covered and with 50 principal components 98.1% are covered. So, the required 70% to 90% are covered for all the numbers in the investigated range, but as we are also interested in the finer structures for the classification, a value above 90% makes sense.

In Figure 7 you can see a back projection of the reduced vectors to the original vectors. Obviously, the quality gets better the more principal components are used. With 10, 20 and 30 principal components (PCs) it is hard to discover that the displayed slice is a head slice but with 50 PCs the shape gets more significant. With 100 PCs almost the original image quality is reached.

The last criterion that was used to evaluate how many PCs to choose is their performance in the classification. The classification is covered in detail in section 3.2 but some results are already presented right here because the performance in the classification is the most important criterion. During the analysis, using 15, 50 and 100 PCs was investigated. To make the results of the classification comparable, the learning process was always done with the same 5 persons. For the comparison, the distance values to the head and feet class of each slice of an untrained person were plotted, see Figure 8. In blue you can see the distance of each slice to the head class and in the green you can see the distance of each slice to the

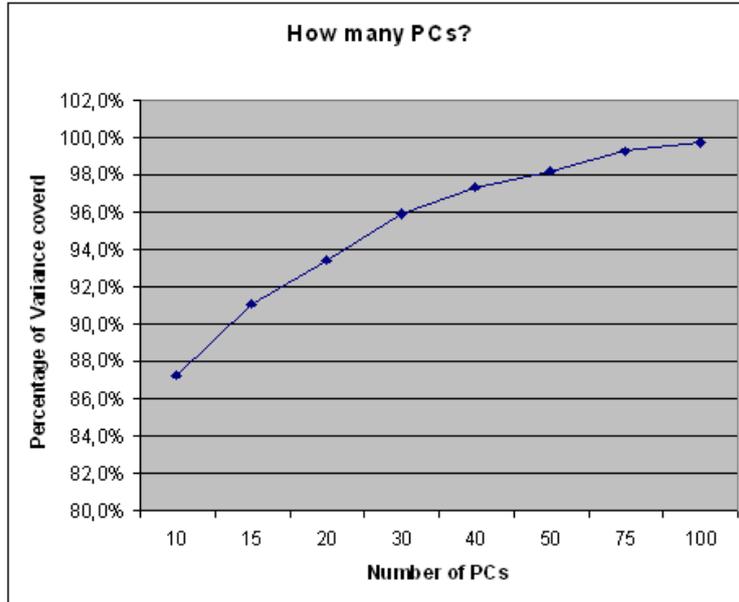


Figure 6: Plot of Covered Percentages depending on Number of PCs

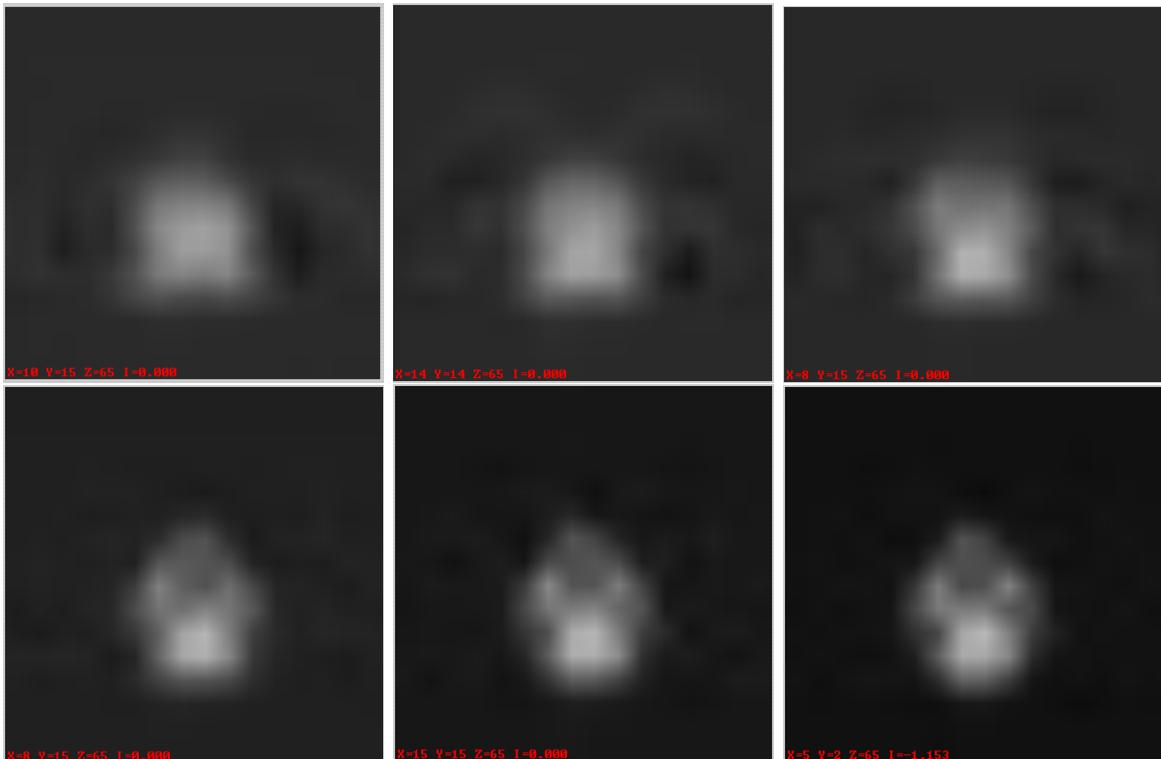


Figure 7: Back Projection of a Head Slice with 10, 20, 30, 50, 75 and 100 PCs

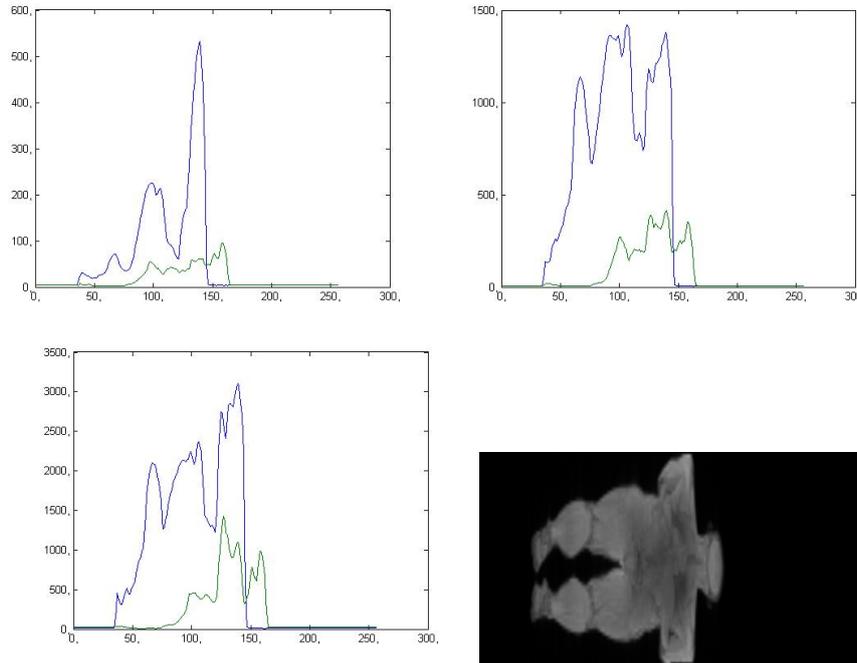


Figure 8: Mahalanobis Distances to the Head (blue) and Feet (green) Class for 15, 50 and 100 PCs

feet class. The x-axis in the plot is the z-axis through the body and the y-axis in the plot shows the Mahalanobis distances. In the Figure you can also see an image of the person the test were applied to. One can see that the absolute distance values are higher the higher the dimension but that is clear as more coordinates get summed up. In the comparison of the three different settings choosing 50 or 100 PCs shows much better results in the feet area. Overall the setting with 50 dimension performed best.

As a short conclusion, we were deciding to take 50 PCs because it seems to be the best compromise between not losing to much structure in the image, to make a nice classification possible, and also not to slow down the following processing too much because of a high dimensionality. Moreover the case with 50 PCs showed the best results in the classification.

3.2. Classification of Slices

Getting back to the initial question of making a decision to which part of the body a slice belongs to, one has to define these classes related to the body parts in before. In the terminology of machine learning we have to deal with supervised learning. The learning process is done by marking some slices belonging to one class and then telling the program which class it is. For instance, marking the head slices of a volume and then telling the program that the marked slices are head slices. Then these slices are transformed in the low dimensional space and their vectors are stored as representatives of these class.

Having the classes defined, it is now possible to evaluate the class membership of a new

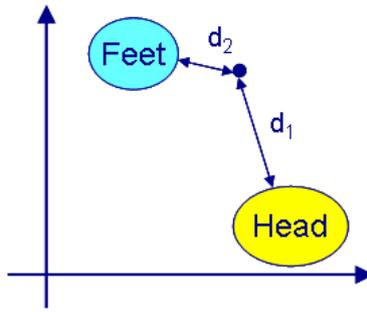


Figure 9: Distance of new Slice to Head and Feet Class.

slice. Firstly, you have to transform it to the low dimensional space and then calculate the distances to each class. This is illustrated in Figure 9. The distance measurement that was used is the Mahalanobis distance and described in detail in Appendix B. The lower the distance to one class, the more probable is it that this slice belongs to that class. A few plots of the Mahalanobis distance along a volume, having defined a head and feet class can be seen in Figure 8.

The question might come up if the classification can be done in realtime as a lot of things have to be calculated. Well, it can be done in realtime because most of the calculations can be done offline, including the PCA and the inverse covariance matrices for the Mahalanobis distance. Online, one just has to do a matrix-vector multiplication to map the image into the low dimensional space and a few further matrix-vector multiplications to get the Mahalanobis distances to the classes.

The classification was tried out with different parts of the body namely head, chest, abdomen, hip, upper leg and feet. The best classification results were obtained with the head, so each other class was compared to it, see Figure 10. Each plot shows in the blue the Mahalanobis distance of each slice to the head class and in green the Mahalanobis distance to the compared class. As stated before the head class performs very well as it delivers high values for the rest of the body and has a nice plateau near zero in the head region. It also delivers small values for black slices as the last head slices with almost no structure are quite similar to a black one. The chest and stomach class do not perform very well. The hip class is better because it has its minimal values in the hip region but still not near to zero. The upper leg class performs well because it has a pronounced minimum in the upper leg region. The best results next to the head class can be seen in the feet class. The learning process was always done with the same 5 randomly chosen persons that can be seen in Figure 11.

3.3. Evaluation Criterion for Head or Feet First

In the previous section the classification with different classes was described. In this section, the focus is set on the problem to decide if the person moves head or feet first into the scanner. Therefore the results from the classification are used but additionally a decision criterion has to be developed to evaluate the classification results. The idea is to use an iterative criterion to make a profound decision. It not just takes one, but several slices into consideration if a

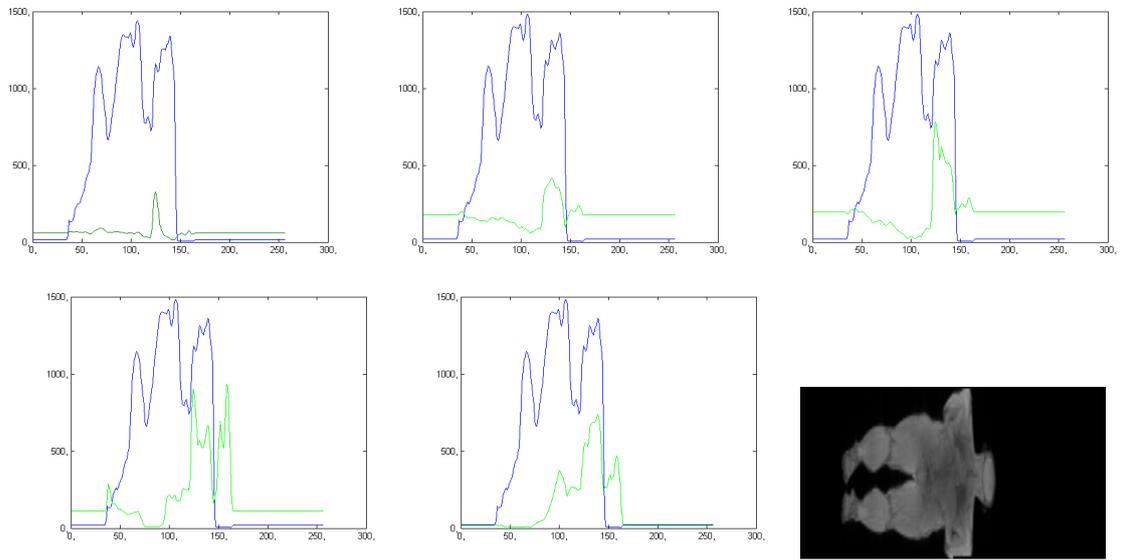


Figure 10: Chest, Abdomen, Hip, Upper Leg and Feet (green) compared to Head (blue)

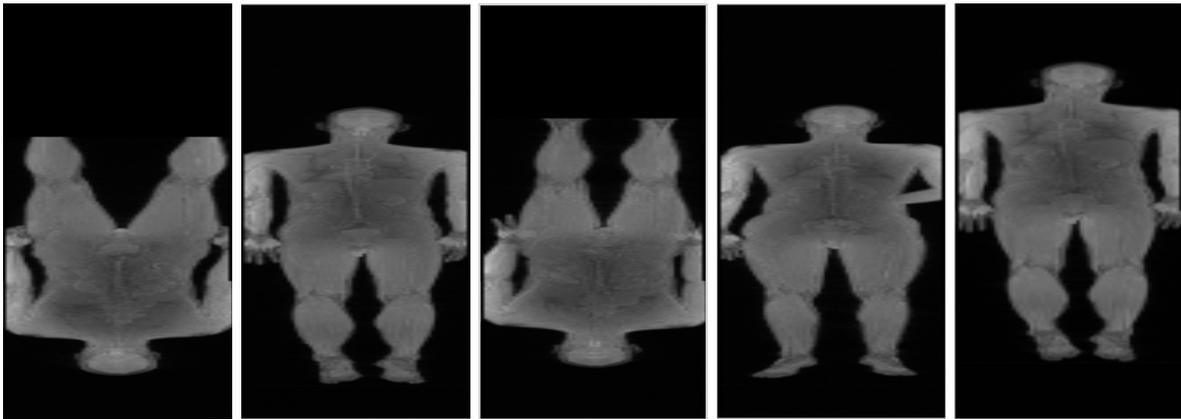


Figure 11: Persons used for Learning

```

trust_head = 0
While (trust_head < trust_threshold)
  If (area_of_slice > 0)    // no black slice
    If (dist_feet / av_dist_feet - dist_head / av_dist_head > distanceThreshold)
      p = av_dist_head / dist_head
      trust_head = trust_head + p

```

Figure 12: Pseudo-Code for Head Evaluation Criterion

certain trust value has not yet been reached. Moreover the criterion should have two properties. Firstly the distance to the nearest class and secondly the distances to the other classes should be considered.

The pseudo code of the criterion can be seen in Figure 12. In the criterion, the trust value is just increased if the difference between another class and the nearest class is greater than a certain *distance threshold*, in our case 20. In the criterion we haven't used the raw Mahalanobis distance values but instead applied a normalization on them. This was done by dividing through the average distance of that class. The average distance for one class is calculated from the slices representing that class.

In the case that the difference to the other class is large enough then the trust value is increased. The increase depends on the distance to the class. If the slice is q times the average distance of the class away, then we add the value $p = 1/q$ to the trust value. A decision for e.g. head is made if the head trust value exceeds a certain *trust threshold* that we have set to 800. In the code, black slices are filtered out by checking their area.

In Figure 13 plots of the head and feet trust values can be seen. A scanning procedure from the left and right is simulated. One can see that it nicely works in the normal case. The special case shown in the third row in Figure 13 makes some problems. As we get an increasing head trust although it is scanning feet slices. But the trust values are lower than in the normal case, so no wrong decision is made. Surprisingly, we also obtained bad results in the feet region for the person shown in the second column in Figure 13 although it is no special case. The reason was that he had a pillow below his feet during the scanning procedure and in our training set there was nobody with a pillow below his feet. So we were adding another person also having a pillow below his feet to the training set. The trust values became significantly better and can be seen in Figure 14. This shows the potential of the learning process and supports the assumption that after exhaustive learning it may work very robust.

3.4. Conclusion for PCA

We applied a principal component analysis on the image data with choosing 50 principal components. In this reduced space we applied supervised learning to get a classification of the slices. The best results were obtained with head and feet classes and the further analysis focused on these two. But also upper leg and hip classes showed good performance. A special decision criterion was developed to detect if the patient moves head or feet first into the scanner. The criterion works on all volumes except the special case. But the special case

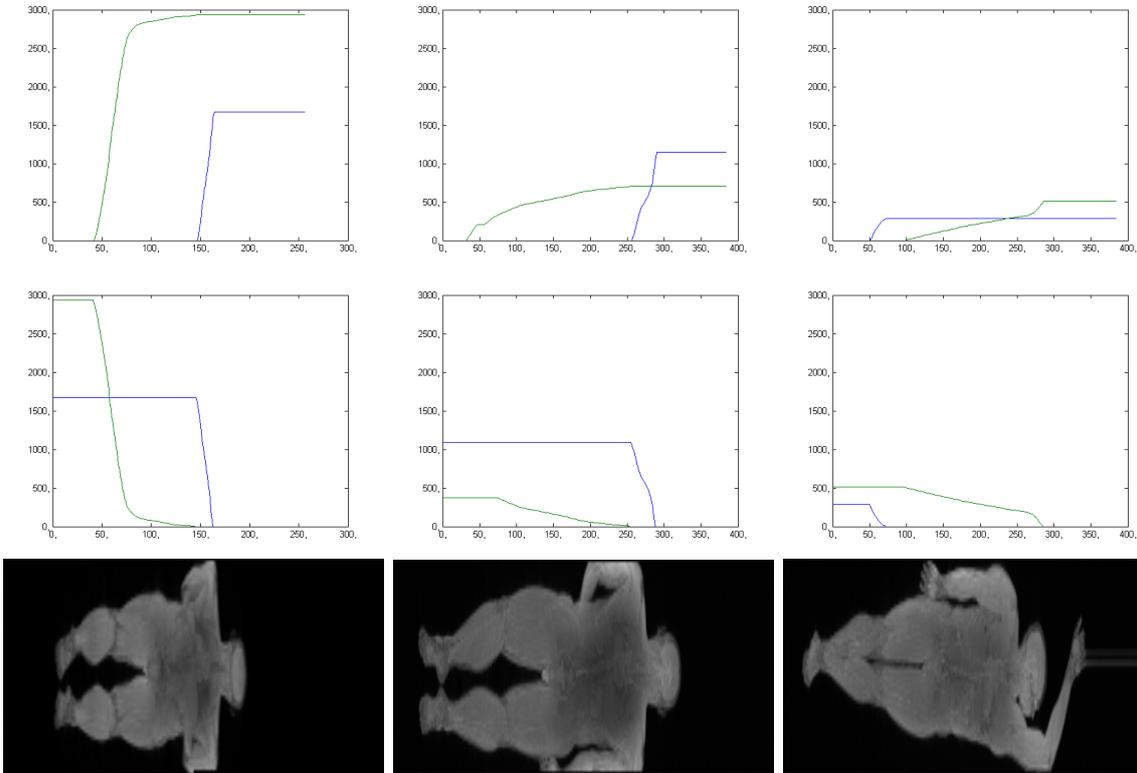


Figure 13: Plot of Head (blue) and Feet (green) trust values. A scan from left is simulated in the first row and from right in the second. The scanned persons are shown in the third row.

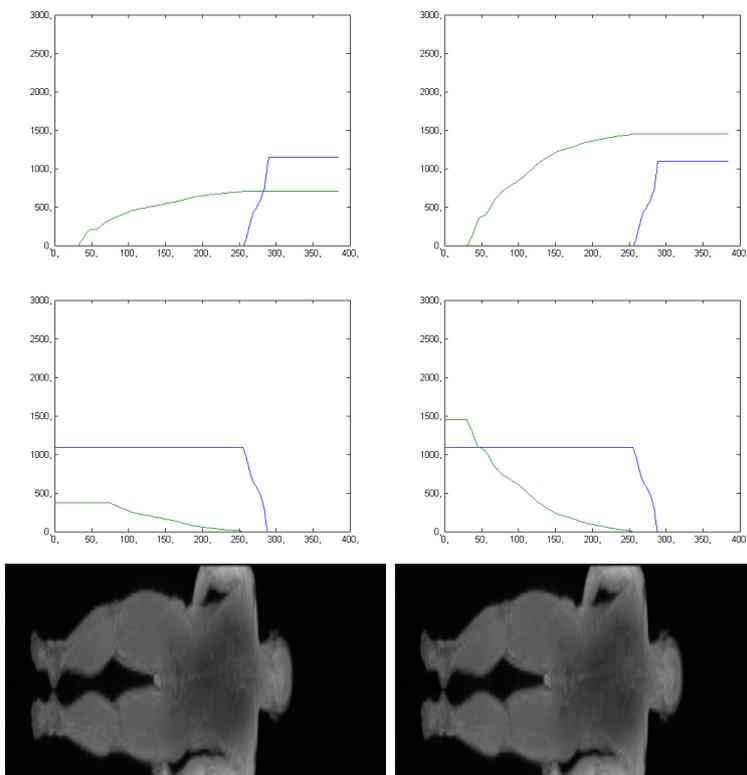


Figure 14: Left row: trust values before learning. Right row: trust values after learning with a different person also having a pillow below his feet. Significantly better results in the feet region after learning.

produces lower values so that at least no wrong decision is made. Moreover the performance can be improved by adding slices to the learning set. But the learning has to be done smart as adding untypical slices like that in the special case can also produce worse results for the normal case.

4. Results of Positioning

In the last two sections methods for the detection of the different parts of the body were listed. In this section each part of the body is considered in detail and methods for its detection are shown. It is desirable to have more than one method per body part to validate the results.

4.1. Head

The best way to detect the head is to use the PCA. There were some scans where the classification showed not so good results. The reason was that the head was a bit shifted to the left. The problem was removed after adding one of these to the learning set. To detect if there is perhaps an arm next to the head the area of the connected components can be used, because the arm has lower values than the head. To distinguish the head from the feet additionally the connected components may work as in the feet area there are two components with nearly the same area. The area can also be used to distinguish the trunk from the head as there are much higher values. Furthermore the area gradient can be used because there is a high gradient at the beginning of the head and at the end of the head (neck). Moreover, as all heads have a similar length the distance between the two gradients must be in this range. Problems may come up with children but they could be detected by entering the patient's age.

4.2. Neck

A criterion for the neck detection was developed that bases on the head first and lung detection. If the lung is found and the patient moves head first into the scanner, the neck can be detected by using the area function. The neck detection worked on all patients where the head was correctly detected. Next to the minimum in the area function there is also the highest area gradient magnitude in the neck section. Unfortunately the absolute values of the gradients vary from patient to patient so that no thresholding criterion is possible.

4.3. Chest

A criterion for the lung detection was developed that bases on the area difference between the area of the patient and the area of the hull. This criterion worked on all persons, also on the special case. The PCA was not showing good results.

4.4. Abdomen

For the abdomen no good criterion was found except using the area to distinguish it from the head and feet section. The PCA was not working well.

4.5. Hip

Like for the abdomen also for the hip no good criterion was found. The abdomen and hip can be detected by finding the trunk with the area function and excluding the chest.

4.6. Upper Leg

The PCA showed nice results for the upper leg class with a nice plateau near zero. Also the area of the connected components works as there are exactly two connected components with roughly the same area.

4.7. Feet

In the feet section the PCA performed very well. The detection for feet worked except for the special case. There were also some bad results for patients with a pillow below their feet but after adding one to the training set, the problem was removed, see section 3.3. Also the connected components in combination with area works as there are just two connected components with roughly the same area. A distinction from the trunk can also be done with the area function.

5. Further Ideas

In this section we will present further ideas for methods that might also work well for the positioning. During this project, it was not possible to test them but they may be topic of a following project.

5.1. Pattern Matching

Pattern matching could be applied to slices from all three views - sagittal, coronal and lateral. As there is no full scan of the body, the sagittal and lateral view is restricted to the begin or the end of the body. The area gradient can be used to determine how many slices to consider for these views. For instance, a cutoff could be done after the second high area gradient peak, so either the head or the feet would be on the image. Suitable patterns to match must be found to make a decision possible. Moreover, the Hough transformation for circle detection could be applied. It might be possible to distinguish between head and feet circles. But as the first head slices have a shape similar to that of legs, it might get complicated.

5.2. Bounding Box around Patient

In the scans that we were working on, all patients were nicely centered so no further adjustment had to be done. To make the algorithms more robust with respect to an usage in a final product the bounding box around the patient should be processed. Perhaps the calculation won't be too simple because in the pre-scan there is no full scan of the body. Centering each slice is not a good idea because the PCA is position dependent. So centering each slice would lower its discrimination power.

5.3. Independent Component Analysis

The Independent Component Analysis (ICA) is an alternative to the PCA. The ICA became known as a powerful solution for the problem of blind source separation. But it is also used in the field of face recognition [1]. In contrast to the PCA in which we are interested in finding a representation of the inputs based on uncorrelated variables, the ICA tries to find a representation based on statistically independent variables [3]. So, the ICA can be seen as a generalization of the PCA. In more detail, the PCA considers the second order moments only and it uncorrelates data, whereas ICA accounts for higher order statistics and it identifies the independent source components from their linear mixtures. Thus, ICA provides a more powerful data representation than PCA [5].

5.4. Support Vector Machines

The Support Vector Machines (SVM) present an alternative for classification that has to be done after the PCA was applied. The principle of the SVM is that an algorithm creates a hyperplane that separates the data into two classes with the maximum-margin. Given training examples labeled either “in” or “out”, a maximum-margin hyperplane splits the “in” and “out” training examples, such that the distance from the closest examples (the margin) to the hyperplane is maximized. The parameters of the maximum-margin hyperplane are derived by solving a quadratic programming (QP) optimization problem.

SVMs are also used in the field of face recognition. A comparison between using ICA/SVM and PCA/SVM for face recognition can be found in [1]. In the experiments, high recognition rates were achieved. The results using the combination PCA/SVM were not far from those using ICA/SVM.

6. Conclusion

A realtime decision based on multiple criteria whether the patient moves head or feet first into the scanner is possible. The detection of the head and feet was feasible with the PCA, except for the special cases. But nevertheless, no wrong decision was made. This was possible as a certain trust value was calculated that expressed the confidence in the decision. A criterion for the lung detection was developed that performed extremely well and robust. So, if the patient moves in head first and the head detection failed, when the lung is scanned the right decision is made. Also a criterion for the detection of the critical neck region was developed, based on the lung and head recognition. Moreover the high potential of learning was shown, supporting the assumption that it may work in a final product if exhaustive training is performed.

A. Covariance Matrix

Assuming that there are m random variables x^i with $i = 1, \dots, m$. For each random variable x^i we have N samples x_j^i with $j = 1, \dots, N$. Let \hat{x}^i be the mean value of the random variable x^i . The covariance matrix Σ of the random variables is defined as

$$\Sigma = \begin{pmatrix} Cov(x^1, x^1) & Cov(x^1, x^2) & \cdots & Cov(x^1, x^m) \\ Cov(x^2, x^1) & Cov(x^2, x^2) & \cdots & Cov(x^2, x^m) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(x^m, x^1) & Cov(x^m, x^2) & \cdots & Cov(x^m, x^m) \end{pmatrix}$$

with

$$Cov(x^k, x^l) = \frac{1}{N-1} \sum_{i=1}^N (x_i^k - \hat{x}^k)(x_i^l - \hat{x}^l).$$

Defining a vector \tilde{x}^i as

$$\tilde{x}^i = \begin{pmatrix} x_1^i - \hat{x}^i \\ x_2^i - \hat{x}^i \\ \vdots \\ x_N^i - \hat{x}^i \end{pmatrix}$$

then the calculation of $Cov(x^k, x^l)$ can be written as

$$Cov(x^k, x^l) = \frac{1}{N-1} (\tilde{x}^k)^T \tilde{x}^l.$$

resulting in a covariance matrix of the form

$$\Sigma = \frac{1}{N-1} \begin{pmatrix} (\tilde{x}^1)^T \tilde{x}^1 & (\tilde{x}^1)^T \tilde{x}^2 & \cdots & (\tilde{x}^1)^T \tilde{x}^m \\ (\tilde{x}^2)^T \tilde{x}^1 & (\tilde{x}^2)^T \tilde{x}^2 & \cdots & (\tilde{x}^2)^T \tilde{x}^m \\ \vdots & \vdots & \ddots & \vdots \\ (\tilde{x}^m)^T \tilde{x}^1 & (\tilde{x}^m)^T \tilde{x}^2 & \cdots & (\tilde{x}^m)^T \tilde{x}^m \end{pmatrix}.$$

Defining the matrix

$$X = (\tilde{x}^1 \quad \tilde{x}^2 \quad \cdots \quad \tilde{x}^m)$$

then the calculation of Σ can be written as

$$\Sigma = \frac{1}{N-1} X^T X.$$

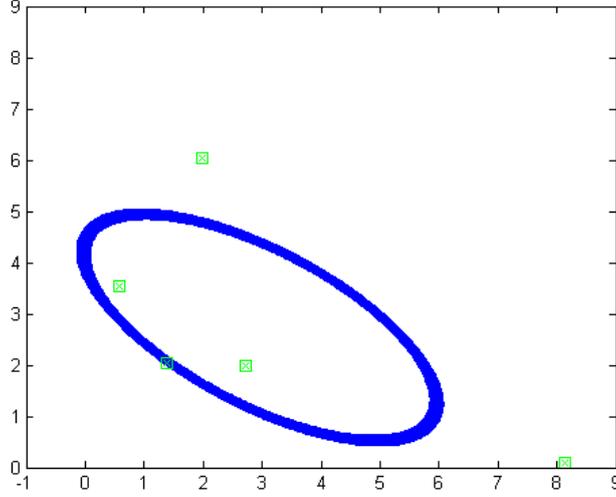


Figure 15: Ellipse with Mahalanobis Distance 1 to the Class (green)

B. Mahalanobis Distance

The Mahalanobis distance is a statistical distance measurement that is used for the classification of slices. It delivers better results than the normal euclidean distance because it takes the form of the class, the distance is measured to, into consideration. A definition can be found in [2].

For the Mahalanobis distance we assume that we have different classes each represented by several vectors and a new vector x we want to calculate the distances to. To make it more formally, each class i has the mean μ_i and the covariance matrix Σ_i . Then the Mahalanobis distance to class i is defined as:

$$\|x\|_M^i = \sqrt{(x - \mu_i)^T \Sigma_i^{-1} (x - \mu_i)}.$$

In Figure 15 you can see a plot of all points having Mahalanobis distance 1 to the class defined by the 5 green points. It has an elliptical form as the points have a higher variance along the north-west south-east axis than on the north-east south-west axis.

For the calculation of the inverse of the covariance matrix we were using the following simplification. According to Appendix A the covariance matrix can be written as

$$\Sigma = \frac{1}{N-1} X^T X.$$

Applying a singular-value decomposition to $X = UDV$ it results in

$$\Sigma = \frac{1}{N-1} (UDV)^T (UDV) = \frac{1}{N-1} V^T D^T U^T U D V = \frac{1}{N-1} V^T D^T D V$$

as U and V are orthogonal matrices. Inverting the covariance matrix it delivers

$$\Sigma^{-1} = (N-1)(V^T D^T D V)^{-1} = (N-1)V^T (D^T D)^{-1} V$$

as $V^{-1} = V^T$. Applying this to the Mahalanobis distance it gives

$$\begin{aligned}\|x\|_M &= \sqrt{(x - \mu)^T (N - 1) V^T (D^T D)^{-1} V (x - \mu)} \\ &= \sqrt{(N - 1) (V(x - \mu))^T (D^T D)^{-1} V(x - \mu)}.\end{aligned}$$

So the calculation is simplified as there is two times the same factor $V(x - \mu)$ and the inversion of the matrix $D^T D$ can be done fast as it is a diagonal matrix.

References

- [1] O. Deniz, M. Castrillon, M. Hernandez, *Face Recognition Using Independent Component Analysis and Support Vector Machines*, 2001
- [2] R. Hartley, A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2000
- [3] A. Hyvriinen, J. Karhunan, E. Oja, *Independent Component Analysis*, John Wiley & Sons, Inc, 2001
- [4] I.T. Jolliffe, *Principal Component Analysis*, Springer Series in Statistics, 2002
- [5] J. Karhunen, E. Oja, L. Wang, R. Vigario, and J. Joutsensalo, *A class of neural networks for independent component analysis*, IEEE Trans. Neural Networks, vol. 8, pp. 487–504, Apr. 1997
- [6] T. A. Runkler, *Information Mining*, Vieweg, 2000
- [7] M. Turk, A. Pentland, *Face recognition using Eigenfaces*, Proc. IEEE Conference on Computer Vision and Pattern Recognition, Maui, Hawaii, 1991