More Active Shape Model

Soren Klim, Stig Mortensen, Bjarni Bodvarsson, Lars Hyldstrup & Hans Henrik Thodberg

Institute of Informatics and Mathematical Modelling,

Technical University of Denmark, 2800 Lyngby, Denmark

s001851@student.dtu.dk

Abstract

This paper describes a comparison of the well known Active Shape Model (ASM) and the new More Active Shape Model (MASM). It is the *first* time MASM has been implemented and thus also the first time it has been compared to ASM. Both implementations were done in Matlab and the comparison were done on a real case of finding the metacarpals in x-ray images of the left hand. The paper also introduces the idea of using three different regression matrices in MASM, which improves both the search range and final convergence of the model.

Keywords: Active shape model, metacarpals, principal component analysis and regression

1 Introduction

This project deals with the problem of locating a shape in a digital image. Two methods are compared against each other:

- *The Active Shape Model* (ASM) [2]

- *The More Active Shape Model* (MASM) [1]

As their names indicate there are several similarities between the two methods. They are both based on a statistical shape model and on statistical information about the edge profiles. Both models are iterative and need a rough initial estimate of the desired image structure. However the algorithms for searching are the same but the information extracted from search is used in different ways to update the model parameters. The purpose of this project is thus to explore the limitations and robustness of the two methods and to compare them against each other in various tests.

1.1 Data

The comparison is conducted using a real case consisting of 24 x-ray images of left hands, where the four metacarpals in the hand are to be located.

The images have a resolution of 300 dpi, but have been blurred using a gaussian kernel. To avoid correlation between the pixels only every fourth pixel is used when sampling from the image. The resulting resolution is thus 75 dpi, but by using the 300 dpi image it is possible to quickly obtain image profiles by using the pixel values which are closest to the sample points.

The metacarpals in all 24 images have been annotated with 400 points (figure 1). These are used as references for the search results and also to build the models.

Figure 1: Example of an annotated x-ray image.

2 Active Shape Model

A short description of ASM will be given as it is compared to the MASM method.

The model consists of points which define the outline of the object. The model is based on the 24 x-ray images where the correct points of the metacarpals have been annotated. The shapes are then alligned using Procustes algorithm and the model shape and variation are described using *"Principle Component Analysis"* (PCA). The analysis results in a model which includes a statistical model of edge profiles and a statistical model of shape.

To find the object using ASM the mean shape is placed on the approximate correct position in the picture. The algorithm is then supposed to fit the shape to the object edges. This is done by finding the best Mahalanobis

fit within the edge profile of each point. The shape points are then moved to the best fit of the edge. This is done for all points in the shape and thereby generating a new shape. The model parameters of the new shape are constrained to be at most 3 standard deviations (see section 2.2). The search for the edge is done in a profile as seen in 2 [2].
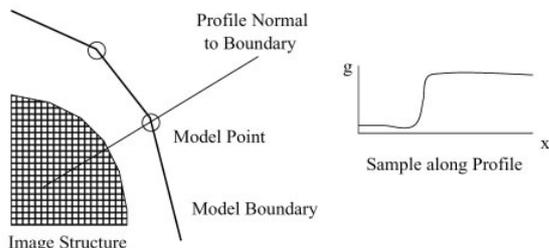

Figure 2: An edge profile is created at each model point

This iterative process is repeated until convergence or an certain number of times. In our implementation ASM iterates 21 times. One could also reduce the search area after each iteration which also damps the process. The shape model has four euclidian parameters and a number of PCA shape scores.
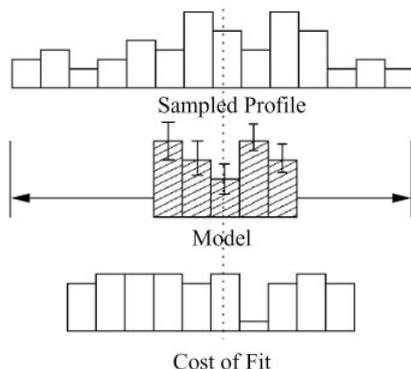
2.1 Mahalanobis Distance

Figure 3: The model is fitted to the sampled profile with the lowest cost of fit

Fitting the model to the edges is done by using Mahalanobis distances. The best fit is found within m pixels (100 px in our implementation) on either side of the shape point. A profile is created by creating a line perpendicular to the model point which has 2m + 1 samples which are put in a vector g. We sample the derivative along the profile, rather than the absolute grey-level values, this is done to reduce the effects of global intensity changes. The vector is then normalized by dividing with the sum of the absolute element values.

$$\mathbf{g}_i \rightarrow \frac{1}{\sum_j |g_{ij}|} \mathbf{g}_i$$

This is repeated for each image to get a set of normalized samples $\mathbf{g}_i$ for each model point. It is assumed that these are distributed as a multivariate gaussian, and the mean $\bar{\mathbf{g}}$ and the covariance $\mathbf{S}_g$ are estimated. This is done for each point and gives a statistical model of the profile about the point. There is constructed one model for each point. To evaluate a fit of a new sample, $\mathbf{g}_s$, to the model the following expression is given:

$$f(\mathbf{g}_s) = (\mathbf{g}_s - \bar{\mathbf{g}})^T \mathbf{S}_g^{-1} (\mathbf{g}_s - \bar{\mathbf{g}})$$

This is the Mahalanobis distance from the point $\mathbf{g}_s$ to the model mean. Minimizing this value is equivalent to maximizing the probability that $\mathbf{g}_s$ belongs to the distribution. The profile (width 2k+1) is fitted on the sample of m pixels of either side of the point. The best of the 2(m-k) possible fits is the one with the lowest f($\mathbf{g}_s$) see figure 3. This is done for all points, and the euclidian and shape parameters are updated.

2.2 Principal Component Analysis

A principal component analysis is conducted on the procrustes aligned shapes resulting in a series of principal components. In the implementation 17 principal components (PC) are used to describe 95% of the shape variation. Moreover 4 euclidian parameters place the shape in the image. In figure 4 it can be seen that the 1st PC describes the variation in length of the metacarpals and especially the left most. The 2nd PC describes the rotation of the left most metacarpus the 3 other metacarpals have almost no variation. In the 3rd PC the variation is very hard to describe as it has almost no visible variation.
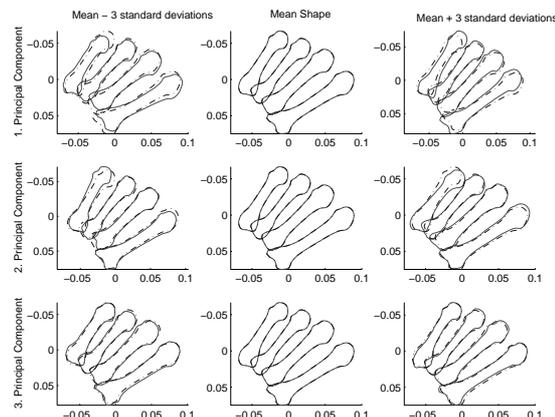

Figure 4: The first 3 principal components

2.3 Weaknesses

The points in ASM are only allowed to move perpendicular to the shape. This can be a problem if there for instance are few normals in a particular direction. Then the shape will have difficulties in moving in this direction. The problem is worsened if the boundary is

badly defined in the edge profile. This is in fact the case with the x-ray images in the vertical direction.

3 More Active Shape Model

The More Active Shape Model is an extension to the original *Active Shape Model*. MASM and ASM share the statistical shape model and the edge profile. The extension in MASM is the ability to recognize patterns in the overall movement.

3.1 ASM and more

MASM uses the same model parameters as ASM. The 17 shape parameters found by PCA and 4 euclidian parameters.

$$b = \begin{bmatrix} b_1 \\ \vdots \\ b_{17} \\ \theta \\ s \\ t_x \\ t_y \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_{17} \\ b_{18} \\ b_{19} \\ b_{20} \\ b_{21} \end{bmatrix}$$

MASM uses the same search algorithm as ASM. Search along the normal and find the best fit using Mahalanobis distance to the edge profile. The difference is that MASM doesn't move the shape to the best fit. It stores the movement in s and uses the overall movement to change the model parameters.

MASM uses the signed distances to the previous shape stored in (s) to evaluate its change in parameters. A way of explaining its strength is that it matches certain patterns of s with a specific change in model parameters.

Figure 5: The Signed Distances s - Taken from [**?**].

Our implementation of the MASM algorithm uses 21 iterations and has no measure of convergence. The algorithm will continue for the 21 iterations without the

ability to stop during the iterations. It is thus the same as in the implementation of ASM.

The training consists of teaching the model what a given sequence of signed distances correspond to in change of model parameters.

The change in parameters is evaluated using the following equation: $\delta b = B \cdot s$. The regression matrix B is the objective in the training.

3.2 Training

The training is done to determine the B matrix. The more experiments put into the evaluation of B - the more efficiently the model will act.

One experiment in the training consists of the mean model aligned to an annotated shape. The aligned mean shape is varied in each model parameter randomly from a uniform distribution and the signed distances are found using the Mahalanobis search.

$$\begin{bmatrix} b_1 & b_1 & \dots & b_1 \\ b_2 & b_2 & \dots & b_2 \\ \vdots & \vdots & \dots & \vdots \\ b_{21} & b_{21} & \dots & b_{21} \end{bmatrix} = B \cdot \begin{bmatrix} s_1 & s_1 & \dots & s_1 \\ s_2 & s_2 & \dots & s_2 \\ \vdots & \vdots & \dots & \vdots \\ s_{400} & s_{400} & \dots & s_{400} \end{bmatrix}$$

In our implementation we use different B matrices - One for large, one for normal and one for small variations of the model parameters. In the next section a more detailed explanation is found.

The matrix trained on large variations is used in the 7 first iterations. The algorithm then shifts to the normal and the last 7 iterations are done using the matrix based on small variations.

3.3 Principal Component Regression

The model parameters are varied in different ways to build three different regression matrices. The 3 different variation schemes are shown below.

	Wide	Normal	Narrow
$\pm Rotation$	$\pi/8$	$\pi/16$	$\pi/64$
$\pm Scale$	500	100	10
$\pm Trans_x$	60	30	5
$\pm Trans_y$	60	30	15

The shape parameters are varied $\pm$ 3 standard deviations at random.

For every scheme 50 experiments are performed in every image and the signed distances and the change in parameter are stored. As a result of this each B matrix is based on 1200 experiments.

In the Principal Component Regression [2] 300 eigenvectors are included in order to describe above 90% of the variation.

3.4 The regression matrices

Below is a plot of the correlation between the true variation in the 21 model parameters and the ones predicted by the matrices. Every column corresponds to one matrix and every row is a variation scheme.
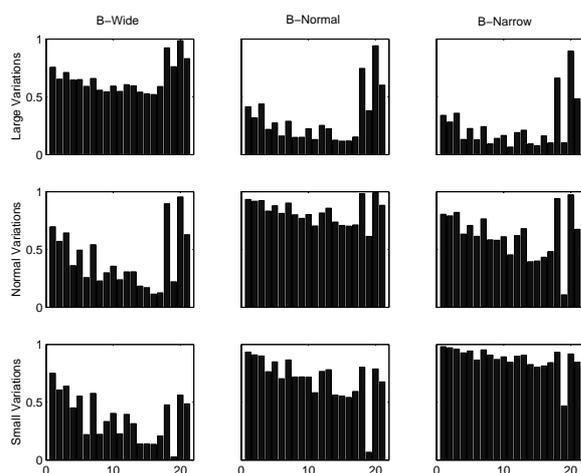

Figure 6: Comparison of the matrices.

In the first row the large variations are predicted by the 3 different matrices. It can be seen that the B_{Wide} matrix predicts the model parameters better and especially the euclidian parameters are predicted very well. The B_{Normal} is not able to handle rotation and y-translation that well.

In the second row the medium variations are predicted. Here the B_{Wide} matrix fails on the shape parameters and rotation. The B_{Narrow} predictions almost correspond to the B_{Normal} in shape parameters however in rotation B_{Narrow} is significantly lower.

In the third row the small variations are predicted. In this variation scheme the focus should be on the shape parameters. The B_{Narrow} predicts the shape parameters with high accuracy and it can be seen that the other matrices predicts worse.

The idea is to use the regression matrices where they are trained and where they also perform the best (corresponding to the diagonal in the figure 6).

The B_{Wide} is in the beginning to efficiently locate the metacarpals. However it will never converge using B_{Wide} but instead fluctuate around the metacarpals.

This is dealt with by shifting to the B_{Normal} regression matrix. The result is improved euclidian parameters without fluctuation and a better determination of the shape parameters. Finally the B_{Narrow} is used to fine tune the shape parameters.

4 Comparison of ASM vs. MASM

The comparison of the two methods will concentrate on robustness towards the initial guess of the shape. Moreover the speed and how they converge are also investigated.

The annotated shapes are used as reference for the search results. A measure of the quality of the search result is obtained by calculating the point to curve distance (PTC) between the search result and the annotated shape. The PTC is the distance from each point in the search result to the curve defining the annotated shape. The PTC can thus be looked upon as a cost of fit to be minimized by the two search methods.

Our experimentation with the images has shown that a PTC less than 100 indicates a good fit and a PTC less than 200 indicates a fair fit. If the PTC is greater than 200 the metacarpals were not found.

To make the comparison fair the two methods are given the same search conditions. The search range for each point is 100 pixels on both sides of the shape and the methods iterate 21 times, which should be more than enough to converge.

4.1 Euclidian transformations

The initial guess of the shape is created by an euclidian transformation of the mean shape. The optimal euclidian parameters for the initial guess can be found by an alignment of the mean shape to the annotated shape.

In order to investigate the robustness of the two methods towards the initial guess, the optimal initial guess is varied in scaling, rotation and translation. The variations are conducted separately for the tree transformations to be able to conclude on the robustness towards both scaling, rotation and translation.

- Scaling: app. ±40% with 40 sample points.

- Rotation: ±90° with 90 sample points.

- Translation: ±250px with 60 sample points.

All sample points are tested on 5 of the 24 x-ray images chosen at random each time. Only the mean curves are shown in the plots.

4.2 Scaling

The ability to handle variations in scale in the initial guess is shown in figure 7. It can be seen that there are no great differences between the search results from the two methods.
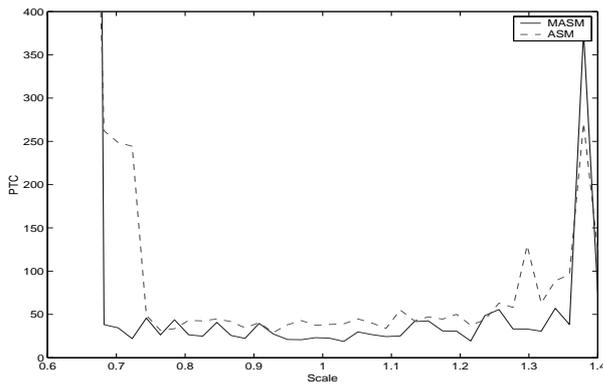
Figure 7: Plot of robustness towards scaling.

Figure 9: Plot of robustness towards hor. trans.

4.3 Rotation

From figure 8 is can be seen that MASM is much better at handling rotation than ASM. MASM can handle $\pm 50^\circ$ whereas ASM only handles $\pm 20^\circ$.
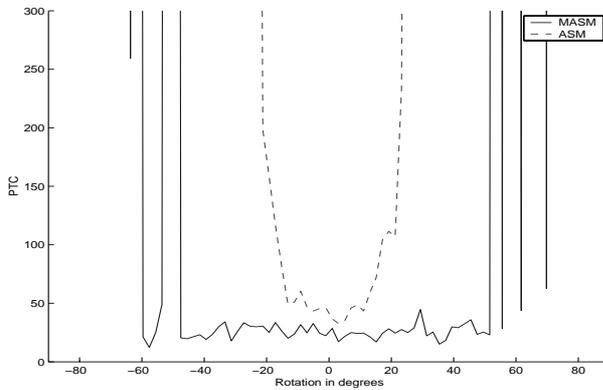

Figure 8: Plot of robustness towards rotation.

It can be seen that even rotation up to 70° (!) are in some cases handled successfully by MASM.

4.4 Translation

The comparison with regards to translation are split in a horizontal and vertical part. The vertical and horizontal translation could also be varied together yielding a 3D-plot, but since these are hard to visualize and compare on paper it has not been done here.

The test on horizontal translation (figure 9) shows almost equally good results for ASM and MASM. They are both able to find the shape within ± 100 pixels. This could also be expected since this is equal to the search range. One could think that increasing the search range would also increase the ability to handle horizontal translation, but this is not necessarily true, since if it is made just slightly larger it results in problems with points moving to the wrong bone.

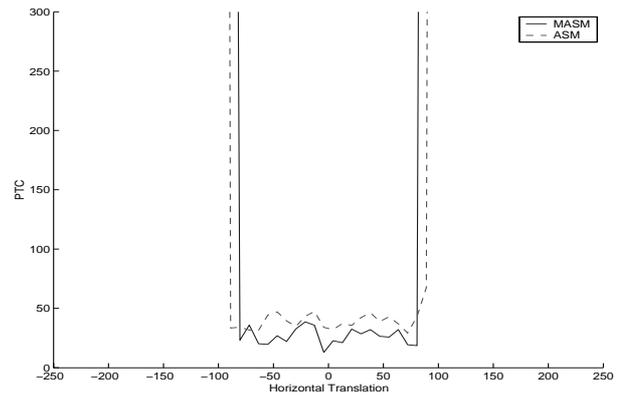The test on vertical translation (figure 10) shows a drastic difference. ASM starts having problems already around ± 50 whereas MASM doesn't have problems before around ± 200 pixels. This clearly shows the great advantage of MASM over ASM. ASM can only use the end points of the bone to move up and down. Unfortunately there are few end points and the edge is not very clear, which makes the ASM unable to make large vertical movements. However, MASM uses the information from from the residuals in all points in the shape to determine the vertical translations. The example here shows that MASM is a much more powerful method.
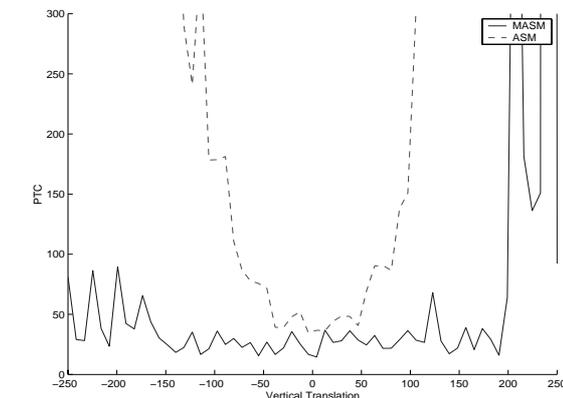

Figure 10: Plot of robustness towards vert. trans.

4.5 Convergence

The sections above have only dealt with analyzing the final result of the two methods. However, it is also interesting to look at how fast they converge.

To explore this a test using all 24 images is conducted. On each image a random variation of the euclidian parameters is chosen and then used as the initialization of the two methods. The random variation is limited in order to make it possible for both methods to find the shape. Figure 11 show the mean of these 24 experiments.

The ASM converges smoothly whereas it for the MASM is visible when the regression matrix (*B*)
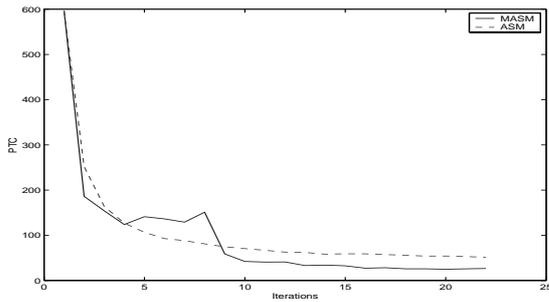
Figure 11: Plot of convergence.

changes at iteration 7 and 14. During the first 7 iterations it stabilizes around 150 with large fluctuations. During the next seven it rapidly decreases and during the last seven it is able to do a little more fine tuning of the search result.

This clearly illustrates the power of using several regression matrices. The first is trained and optimized to locate the shape and the next two to optimize the euclidian parameters and the shape parameters.

The plot of convergence also reveals information about the quality of the search results by ASM and MASM. When studying this together with the previous four figures in this chapter it can clearly be seen that MASM tends to converge at a lower PTC when finding the shape.

4.6 Speed

The algorithm for MASM requires fewer calculations than the ASM algorithm. The search in the image is directly transformed to an update of the 21 parameters by multiplying it with the regression matrix. ASM has to align the mean shape with the shape found by searching the image and then use the shape parameters to model the difference. However both methods look up the image profiles and calculate the Mahalanobis distances which requires the main part of the computational work, so the difference in speed between the two will only be small.

To test whether or not this hold true the mean time of 100 searches with ASM and MASM has been found.

Search times with 21 iter.	
ASM	32.457 sec.
MASM	31.567 sec.

This shows that MASM is slightly faster. It should be noted, that our implementation of ASM and MASM has been done in Matlab. Although it has been optimized for Matlab, the search time is not in any way comparable to the speed that can be achieved by making an optimized implementation in a low level language like C.

5 Conclusion

This paper documents the *first* time MASM has been implemented and thus also the first time it has been compared to ASM. Both implementations were done in Matlab. Furthermore, it also introduces the idea of using three different regression matrices, each trained on wide, normal and narrow movements of the shape.

Disadvantages of MASM

The MASM contains more information about the problem, which is stored in the regression matrix. The generation of *B* naturally requires a little more effort during the training of the model.

Advantages of MASM

During our work with the x-ray images of the metacarpals, MASM has proved to be superior to ASM. ASM is unable make large movements along the bone, but this is solved using MASM. On top of this MASM is also faster and gives a significantly better search result.

5.1 Acknowledgements

References

[1] H. H. Thodberg and A. Rosholm: *Application of the Active Shape Model in a commercial medical device for bone densitometry*, Proceedings of British Machine Vision Conference, 2001.

[2] T. F. Cootes and C. J. Taylor: *Statistical Models of Appearance for Computer Vision*, Wolfson Image Analyis Unit, Imaging Science and Biomedical Engineering, University of Manchester, 2000.

[3] Mikkel Bille Stegmann: *Active Appearance Models - Theory, Extensions & Cases*, Master Thesis, 2nd edition, IMM, Technical University of Denmark, 2000.

[4] Knut Conradsen: *En introduktion til statistik - Bind 2*, 6. udgave, IMM, Danmarks Tekniske Universitet, 2002.