Extending and Applying Active Appearance Models for Automated, High Precision Segmentation*

Mikkel B. Stegmann, Rune Fisker, Bjarne K. Ersbøll

Informatics and Mathematical Modelling, Technical University of Denmark,
Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, Denmark

*Full title: *Extending and Applying Active Appearance Models for Automated, High precision Segmentation in Different Image Modalities*

Introduction

The deformable template model Active Appearance Model (AAM) by Cootes et al. has proven to be a very fast and flexible method to perform segmentation of shape varying objects. In this study we propose a set of extensions aiming at automating the usage by an initialization scheme and enhance the accuracy of the AAM.

Active Appearance Models

AAMs are elegantly capable of learning both shape and texture (appearance) variability from examples simultaneously. These examples constitute a training set, which can be viewed upon as a set of representative solutions to the segmentation problem the model should solve.
The major steps in the AAM are listed below:

AAM Model Training

I. A set of representative images is chosen and annotated by experts.

II. The resulting shapes are spatially aligned using a Procrustes Analysis.

III. Appearance variation is collected in a consistent manner, by establishing a warp function between the prototype and each training example.

IV. In order to derive a specific and compact representation of the variation of shape (landmarks) and appearance (pixels), a principal component analysis (PCA) is performed on the aligned training set.

V. The compact parameterisation is then used to generate synthetic images of the object class.

AAM Segmentation

I. The model is automatically placed in an initial configuration over the unseen image.

II. Using a principal component multivariate linear regression model, new images are generated to fit the unseen image in the best possible way.

Neighborhood AAMs

Dealing with objects surrounded by a relatively consistent neighborhood over the training set, specificity can be markedly increased by including this neighboring region. This can be accomplished by placing additional landmarks outside the original shape. To preserve the eigenvalue distribution of the shape PCA these landmarks must be linear combinations of the original landmarks. An example is given in figure 1.

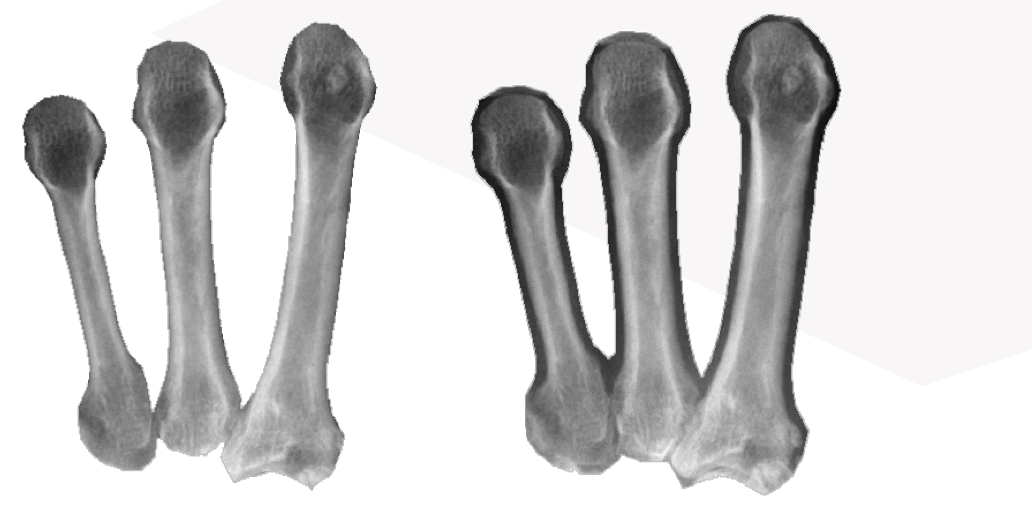

Figure 1. Left: Metacarpal 2,3,4 annotated using 150 landmarks. Right: Corresponding shape with neighborhood region added using 2x150 = 300 landmarks.

Border AAMs

For largely heterogeneous objects w.r.t. texture it is suggested to obtain texture samples only from structured regions – i.e. regions with consistent texture over the training set. In figure 2 a symmetric neighborhood around the object border is modeled to avoid the unstructured interior of the pork chops – which would be regarded noise by the texture PCA.

Figure 2. Left: Shape annotated using 83 landmarks. Right: Corresponding border representation using 3x83 = 249 landmarks.

Summary

This study presents a set of extensions to the deformable template model: Active Appearance Model (AAM). All proposed extensions lead to a higher segmentation accuracy when assessed on radiographs, MRI and perspective images. Further, an initialization method is proposed, which rendered the AAM fully automated. In two of three cases sub pixel accuracy was obtained.

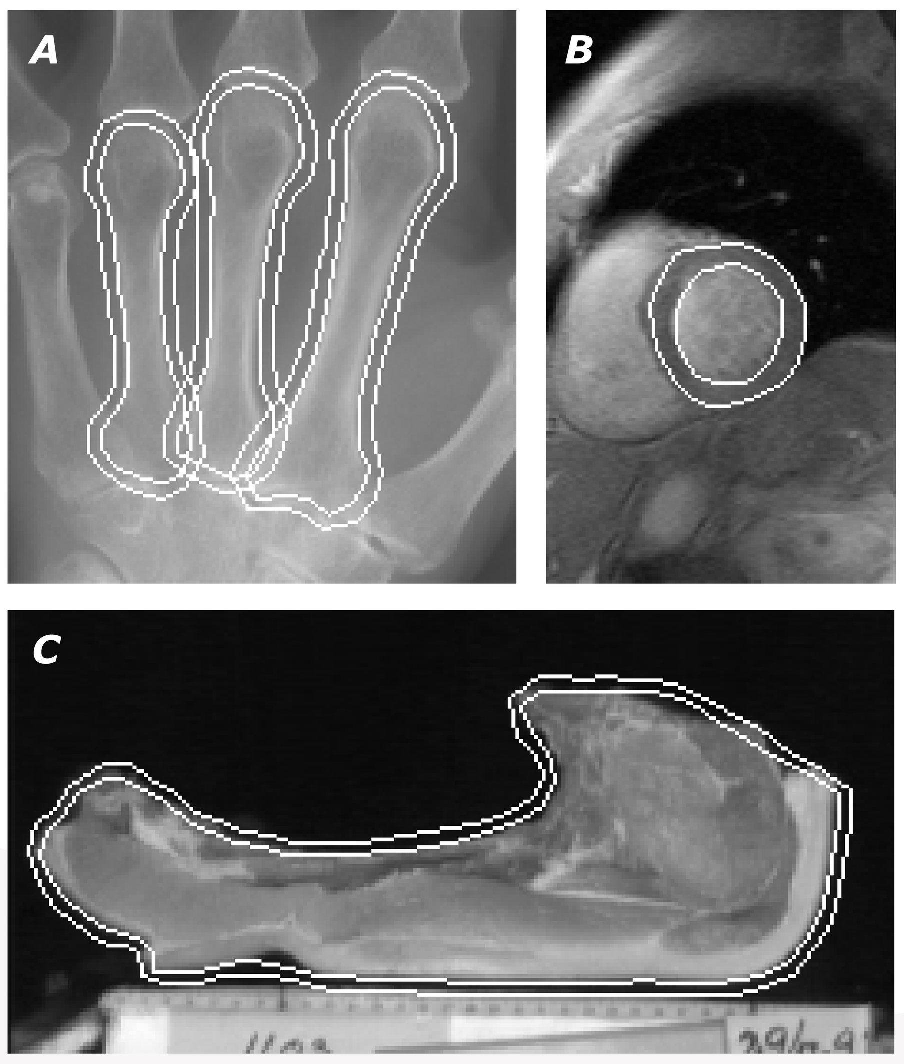

Figure 4. Collage of cases with overlaid segmentation results. Data for the corresponding AAMs are:

A - Radiographs of Metacarpals
Training set: 23 images (240x275 pixels)
Shape model: 150 landmarks
Texture model: ~13.000 pixels
95% variation explained using: 18 parameters

B - Cardiac Magnetic Resonance Images
Training set: 13 images (256x256 pixels)
Shape model: 66 landmarks
Texture model: ~2.200 pixels
95% variation explained using: 11 parameters

C - Perspective images of Pork Chops
Training set: 13 images (256x191 pixels)
Shape model: 83 landmarks
Texture model: ~15.000 pixels
95% variation explained using: 10 parameters

#	Type	Point to curve deviation	Mean intensity deviation	Init fail.
A – Metacarpals				
1	Basic AAM	0.88	4.9	1
2	1+Neighborhood	0.84	5.2	0
3	2+SA	0.82	5.0	0
4	3+Lorentzian	0.83	5.0	0
B – Cardiac MRIs				
1	Basic AAM	1.18	7.1	0
2	1+Neighborhood	1.73	7.5	0
3	1+SA	1.06	5.9	0
4	3+Lorentzian	1.13	6.0	0
C – Pork Chops				
1	Basic AAM	1.12	13.2	0
2	1+Neighborhood	0.91	13.9	0
3	2+SA	0.89	13.6	0
4	3+Lorentzian	0.91	13.6	0
5	Border AAM	0.86	23.5	0

Table 1. Leave-one-out test results. Point to curve measure has units of pixel distances. Mean intensity deviation is measured on 8 bit pixels.

Contact Information
Mikkel B. Stegmann: mbs@imm.dtu.dk,
http://www.imm.dtu.dk/~aam/

Fine-tuning the Model Fit

The AAM Search provides a fast way of optimizing the AAM using prior knowledge. However, prior assumptions regarding the appearance of the search space is only true up a certain accuracy. We propose to use a general purpose random-sampling optimization scheme Simulated Annealing (SA) to further fine-tune the AAM fit.

Robust Similarity Measure

Traditional AAMs uses the L_2-norm as similarity measure. If the i^{th} image measure is denoted g_i, the model parameters $\mathbf{c}$, and the L_2-norm $\rho(e_i, \sigma_s) = e_i^2$, the model fitting is carried out by a minimization of the summed ρ:

$$E = \sum_{i=1}^{m} \rho(g_i - u(i, \mathbf{c}), \sigma_s) = \sum_{i=1}^{m} \rho(e_i, \sigma_s)$$

Here σ_s is a scale parameter that is unused in the L_2-norm. It is easily seen that outliers (measurements with large residuals) will dominate the similarity measure above due to the rapid growth of the quadratic function. We propose to use the robust Lorentzian error norm, which is less sensitive to large residuals:

$$\rho(e_i, \sigma_s) = \log(1 + \frac{e_i^2}{2\sigma_s^2})$$

Here the scale parameter σ_s determines what should be deemed outliers. Refer to figure 3 for an example where a robust similarity measure is needed to obtain a correct fit.

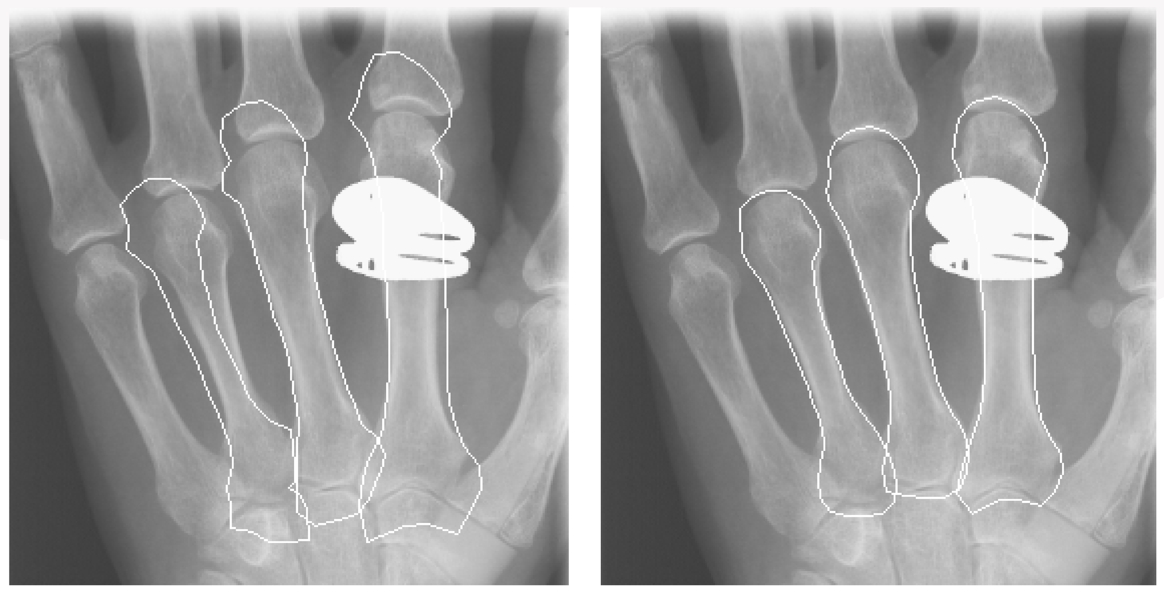

Figure 3. Example of AAM search and Simulated Annealing fine-tuning, without (left) and with (right) the use of a robust Lorentzian norm. The landmark error decreased from 7.0 to 2.4 pixels (pt.crv.).

Initialization

As pointed out by Cootes et al. AAMs are highly dependent on good initialization. To accommodate this, we propose a search-based initialization scheme that exploits an inherent property of the AAM search; namely convergence within some range from the optimum. Thereby a deterministic search in the hyperspace spanned by model- and pose-parameters is narrowed down from *dense* to *sparse*. Inspired by *Genetic Algorithms* we let the best *n* guesses from the initial search population survive. Upon these a further investigation is carried out and the *fittest* element is denoted the initial configuration.

Experimental Results

In order to asses the initialization scheme and the proposed extensions AAMs have been applied to the three image modalities shown in the center column (figure 4).
In each case a leave-one-out evaluation was performed on the training set. Results are enumerated in table 1.

Conclusion

We have presented a set of extensions which all yield higher landmark accuracy when applied to the type of situations they address.
The performance has been assessed on three different image modalities, reaching a mean landmark accuracy of 0.82, 1.06 and 0.86 pixels (pt.crv.). All experiments were carried out without manual interaction.
We conclude that the AAM approach with the proposed extensions is a fully automated, robust and accurate segmentation method that captures domain knowledge through observation and applies to very different cases.