A non-linear gray-level appearance model improves active shape model segmentation

Bram van Ginneken, Alejandro F. Frangi*, Joes J. Staal,
Bart M. ter Haar Romeny and Max A. Viergever

Image Sciences Institute, University Medical Center Utrecht, The Netherlands
*Departamento de Ingeniería Electrónica y Comunicaciones, Universidad de Zaragoza, Spain

Abstract

Active Shape Models (ASMs), a knowledge-based segmentation algorithm developed by Cootes and Taylor [1, 2], have become a standard and popular method for detecting structures in medical images. In ASMs – and various comparable approaches – the model of the object's shape and of its gray-level variations is based the assumption of linear distributions. In this work, we explore a new way to model the gray-level appearance of the objects, using a k-nearest-neighbors (kNN) classifier and a set of selected features for each location and resolution of the Active Shape Model. The construction of the kNN classifier and the selection of features from training images is fully automatic. We compare our approach with the standard ASMs on synthetic data and in four medical segmentation tasks. In all cases, the new method produces significantly better results ($p < 0.001$).

1 Introduction

Segmentation is an important topic in the field of biomedical image analysis. Dedicated solutions for the segmentation of a wide range of objects in a wide range of medical images have been proposed. Most of these algorithms are the result of identifying separate tasks and applying engineering principles to find appropriate solutions, as was noted by Duncan and Ayache in a recent review [3]. Although such approaches have been successful in several areas, there is a clear need for generic segmentation schemes that can be trained with examples. Such schemes acquire a model of the shape of the object to be segmented and the gray-level appearance of the object in the image and provide a mechanism to fit these models to new images, thus producing a segmentation.

A variety of such schemes have been proposed, see for instance [4, 5, 6, 7, 8, 9, 10]. One popular method is Active Shape Models[1] (ASMs) put forward by Cootes and Taylor [1] and popular in medical image segmentation [11, 12, 13, 6, 14, 15, 16].

The ASMs method uses a statistical model of the shape of the object, derived from landmark points[2] on the object contour, and a statistical model of the profiles perpendicular to the object contour, around each landmark. Both the shape and the appearance model are linear: they are derived from the covariance matrix. It is not unlikely that more sophisticated models will improve the overall performance of the scheme. This has been noted by several authors and more complex alternatives for the shape model have been proposed, for instance by Cootes and Taylor [17] and recently by Duta and Jain [18]. In this work we focus on the appearance model. Cootes and Taylor propose to construct the covariance matrix of the normalized first derivatives of the profile around each landmark, and to move landmarks in the image using the Mahalanobis distance. There is no particular reason why this should be an optimal choice for any application. Therefore we propose a new scheme that automatically selects local image features for each landmark that are optimally able to classify a position in the image as either inside or outside the object. As in the original ASMs method, this appearance model will be constructed for multiple resolutions to enable coarse-to-fine fitting.

In this paper the extension of ASMs will be tested and compared with the original scheme in four medical segmentation tasks: segmenting the left and right lung fields in chest radiographs and segmenting the cerebellum and corpus callosum in 2D slices from MRI brain studies.

[1] The term "active shape model" picked by Cootes and Taylor for their segmentation algorithm is a somewhat awkward choice. It is used in the literature in a general sense, as a shape model that includes shape variations. In our case it refers to the complete method (and not just the shape model) put forward in [1] and [2]. The latter reference is a technical report that describes the ASM method in detail; our implementation is based on this description.

[2] In the literature the term landmark often denotes anatomical landmark points, or points with recognizable geometric properties, but in this work it is a sample point along the object contour.

0-7695-1336-0/01 $17.00 © 2001 IEEE

2 Theory

In this section we first briefly review ASMs, and subsequently describe our extensions. In principle, the scheme can be used in nD, but in this work we give a 2D formulation.

2.1 Shape model

An object is described by n landmark points, (manually) determined in a set of s training images. The landmark points $(x_1, y_1), \ldots (x_n, y_n)$ are stacked in shape vectors

$$\mathbf{x} = (x_1, y_1, \ldots, x_n, y_n)^T. \quad (1)$$

These shapes can be aligned by translating, rotating and scaling them so as to minimize the sum of squared distances between the landmark points, using an iterative scheme known as Procrustes analysis [19, 2]. However, this alignment can also be omitted, in which case a shape model is built that can generate only shapes with a size, position and orientation consistent with the supplied examples. The mean shape is computed,

$$\bar{\mathbf{x}} = \frac{1}{s} \sum_{i=1}^{s} \mathbf{x}_i, \quad (2)$$

the covariance

$$\mathbf{S} = \frac{1}{s-1} \sum_{i=1}^{s} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T, \quad (3)$$

and the eigensystem of $\mathbf{S}$. The eigenvectors corresponding to the t largest eigenvalues λ_i (the principal components) are retained in a matrix $\mathbf{\Phi} = (\phi_1 | \phi_2 | \ldots | \phi_t)$. A shape can now be approximated by

$$\mathbf{x} \approx \bar{\mathbf{x}} + \mathbf{\Phi b} \quad (4)$$

where $\mathbf{b}$ is a vector of t elements containing the model parameters, computed by

$$\mathbf{b} = \mathbf{\Phi}^T (\mathbf{x} - \bar{\mathbf{x}}). \quad (5)$$

Such shape models are called point distribution models [20]. When fitting the model to a set of points, the values of $\mathbf{b}$ are constrained to lie within the range $\pm m\sqrt{\lambda_i}$, with m usually between 2 and 3.

The number t of eigenvalues to retain is chosen so as to explain a certain proportion f_v of the variance in the training shapes, usually between 0.90 and 0.995. The desired number of modes is given by the smallest t for which

$$\sum_{i=1}^{t} \lambda_i \geq f_v \sum_{i=1}^{2n} \lambda_i. \quad (6)$$

2.2 Gray-level appearance model

The gray-level appearance model that describes the typical image structure around each landmark is obtained from pixel profiles, sampled (using linear interpolation) around each landmark, perpendicular to the contour[3].

On either side k pixels are sampled using a fixed step size, which gives profiles of length $2k + 1$. Cootes and Taylor propose to use the normalized first derivatives of these profiles to build the gray-level appearance model. The derivatives are computed using finite differences between the $(j-1)$th and the $(j+1)$th point. The normalization is such that the sum of absolute values of the elements in the derivative profile is 1.

Denoting these normalized derivative profiles as $\mathbf{g}_1, \ldots, \mathbf{g}_s$, the mean profile $\bar{\mathbf{g}}$ and the covariance matrix $\mathbf{S}_g$ are computed for each landmark. This allows for the computation of the Mahalanobis distance between a new profile $\mathbf{g}_i$ and the profile model

$$f(\mathbf{g}_i) = (\mathbf{g}_i - \bar{\mathbf{g}})\, \mathbf{S}_g^{-1}\, (\mathbf{g}_i - \bar{\mathbf{g}}). \quad (7)$$

Minimizing the Mahalanobis distance $f(\mathbf{g}_i)$ is equivalent to maximizing the probability that $\mathbf{g}_i$ originates from the distribution $\{\mathbf{g}_1, \ldots, \mathbf{g}_s\}$.

2.3 Multi-resolution framework

These profile models, given by $\bar{\mathbf{g}}$ and $\mathbf{S}_g$, are constructed for multiple resolutions. The number of resolutions is denoted by L_{max}. The finest resolution uses the original image and a step size of 1 pixel when sampling the profiles. The next resolution is the image observed at scale $\sigma = 1$ and a step size of 2 pixels. For subsequent levels $\sigma \rightarrow 2\sigma$ and the step size is also doubled[4].

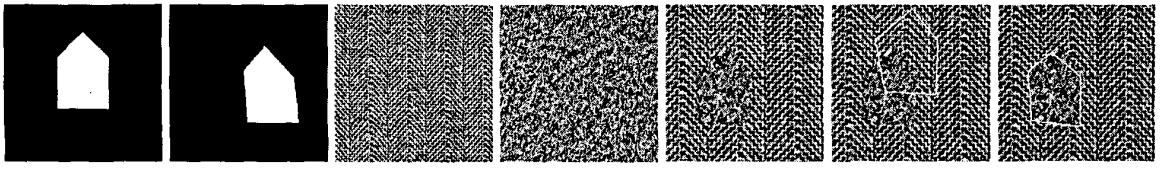
2.4 Optimization algorithm

Shapes are fitted in an iterative manner, starting from the mean shape. Each landmark is moved along the direction perpendicular to the contour to n_s positions on either side, evaluating a total of $2n_s + 1$ positions. The step size is $2^{(i-1)}$ pixels for the ith resolution level, the same as was used during construction of the model. The landmark is put at the position with the lowest Mahalanobis distance. After moving all landmarks, the shape model is fitted to the displaced points, yielding an updated segmentation. This is repeated N_{max} times at each resolution, in a coarse-to-fine fashion[5].

[3] This requires a notion of connectivity between the landmark points from which the perpendicular direction can be computed.

[4] We do not subsample the images, as proposed by Cootes and Taylor.

[5] We always perform N_{max} iterations, contrary to Cootes and Taylor who move to a finer resolution if a convergence criterion is reached before the N_{max}th iteration.



(a) (b) (c) (d) (e) (f) (g)

Figure 1: (a) A generic house image, resolution 200 × 200 pixels. New house images are generated by adding a displacement (dx, dy) to each corner point where dx and dy are randomly selected from the interval (-20,+20) pixels, and subsequently translating the distorted house by (dx, dy) where dx and dy are randomly selected from the interval (-35,+35) pixels. (b) An example of a generated house image. (c) Texture image (resolution 300 × 300 pixels) used to fill the outside of the house, using a random offset. (d) Texture image (resolution 300 × 300 pixels) used to fill the inside of the house, using a random offset. Both texture images are taken from the Brodatz set and histogram equalized. (e) An example of a generated textured house image. (f) Segmenting this image with the ASMs scheme leads to poor results; (g) but new approach with selected local features produces a reasonable result.

2.5 New gray-level appearance model based on selected local features

The aim of the appearance model is to be able to move the landmark points to better locations during optimization, along a profile perpendicular to the object contour. The best location is the one for which everything on one side of the profile is outside the object, and everything on the other side is inside of it[6]. We propose a method to estimate the probability that a location is inside/outside the object, optimized for the area around each landmark and each working resolution separately. We base this classification on local image features obtained by feature selection and a non-linear kNN-classifier, instead of using the fixed choice of the normalized first derivative profiles and the Mahalanobis distance.

Gaussian derivatives will be used as image filters. The rationale behind this choice is that these derivatives optimally describe local image structure since they make up the *local jet*, the Taylor expansion of the image at each location [21]. Features are extracted for each location by taking the first few moments of the local distribution of image intensities (the histogram) around each location. The most suitable choice for a window function to compute this histogram, is a Gaussian, since every other choice induces spurious resolution [22]. The size of this window function is characterized by a second scale parameter α. The construction of local histograms, extracted from a Gaussian aperture function, is called a *locally orderless image* and discussed in depth in [23]. The idea of using the moments of histograms of responses of an image to a bank of filters is a standard technique in texture analysis, see for instance [24].

This leaves us with several parameters to vary: the order of the Taylor expansion (*i.e.* the number of filters in the filter bank), the number of scales σ to consider, the number of scales α to use for the local window, and the number of moments m to extract from the local histograms. Our strategy is to compute an extensive set of features and use feature selection techniques to find the best set of features.

After a range of initial experiments we decided to use only first and second moments ($m = 1, 2$), all derivatives up to second order (L, L_x, L_y, L_{xx}, L_{yy}, L_{xy}), five inner scales ($\sigma = 0.5, 1, 2, 4, 8$ pixels), and a fixed relation between the inner scale σ and the histogram extent α of $\alpha = 2\sigma$. Hence the total number of feature images is $2 \times 6 \times 5 = 60$.

Obviously the image structure is different for each landmark, but the positions that are evaluated are also different for each resolution. Therefore we will select a distinct set of features for each landmark *and* for each resolution, amounting to nL_{max} feature sets[7].

From each training image and for each landmark a square grid of $N_{grid} \times N_{grid}$ points is defined with N_{grid} an odd integer and the landmark point at the center of the grid. The spacing is $2^{(i-1)}$ pixels for the ith resolution level.

For each landmark and for each resolution level, a feature vector with 60 elements is sampled at N_{grid}^2 points. The output of each feature vector is either inside (1) or outside (0) the object[8]. The set of training images is divided in two sets of equal size, a training set and a validation set. A kNN classifier with k_{NN} neighbors and weighted voting is used in which each neighbor votes with a weight of $\exp(-d^2)$,

[6]This assumes that the thickness of the object, in the direction perpendicular to a landmark, is larger than half the length of the profile. We will return to this point in the Discussion.

[7]In the original ASMs the same strategy is followed: nL_{max} mean profiles $\bar{\mathbf{g}}$ and covariance matrices $\mathbf{S}_g$ as they appear in Eq. (7) are computed: for each landmark, at each resolution.

[8]The landmark points themselves are (arbitrarily) considered to be inside the objects.

with d is the Euclidean distance from sample to neighbor in the feature space.

Sequential feature forward selection [25, 26] is used to find a feature set of at most f_{max} features. This set is subsequently trimmed by sequential feature backward selection, that is, features are removed if that improves performance[9]. The resulting set is the "optimal" set of features that will be used during segmentation.

When the model is fitted to an input image, the scheme starts by computing the 60 feature images. Instead of sampling the normalized derivative profiles, the optimal feature set at each position along the profile is fed into a kNN classifier to determine the probability that this pixel is inside the object. The objective function $f(\mathbf{g})$ to be minimized is the sum of absolute differences between the expected probability (0 or 1 for points outside or inside the object, respectively) and the predicted probability, for each point along the profile:

$$f(\mathbf{g}) = \sum_{i=-k}^{-1} g_i + \sum_{i=0}^{+k} (1 - g_i), \tag{8}$$

where the index along the profile $\mathbf{g}$, that is oriented from the outside to the inside of the object, runs from $-k$ to $+k$. This metric replaces the Mahalanobis distance from Eq. (7).

3 Example

Consider the segmentation of a simple object that is filled with a certain texture and placed on a background of a different texture. This is a particular example of a segmentation task for which the original ASMs scheme is not suited. Because of the randomness in the texture, the pixel profiles that cross the object border will not be very different from those that do not cross the border.

We created some sample data and segmented these with both schemes, the original method and the new method. This is described in Figure 1. Both methods were trained on 45 images, for parameter settings see the next section. The point to make is that the original ASMs cannot deal with such textural object boundaries, while the new method produces reasonable results.

4 Experiments

4.1 Materials

Four different segmentation experiments have been performed with two types of data. The images and objects used in the experiments, that are labelled A to D, are briefly described in Table 1.

The image data for Experiments A and B were standard PA chest radiographs selected from a tuberculosis screening program. The data contained both normal and abnormal cases of patients of 16 years and older. The images were taken with a mobile Electrodelca (Oldelft BV, Delft, The Netherlands). The tube voltage was 117 kV and the images were printed on 10 by 10 cm film and digitized with a Lumisys 100 scanner (Lumisys, Inc., Sunnyvale, CA) and subsampled to 256 $\times$ 256 pixels. Two observers have independently segmented the right and left lung field.

For Experiments C and D, a collection of 90 MRI slices of the brain has been used, in which the corpus callosum and the cerebellum have been segmented. The images and segmentations have been made available by the University of Iowa Hospitals and Clinics and have also been used by Brejl and Sonka in [10]. The resolution is 320 $\times$ 256 pixels, 0.7 mm per pixel, obtained by interpolating the original volumetric data acquired with 1.5 mm thick coronal slices with 0.7 by 0.7 mm resolution.

The objects in the images were annotated by a number of *fixed* landmarks and a closed contour between those fixed points from which a number of equidistant landmark points were sampled. Table 1 lists the number of fixed and total landmarks.

4.2 Methods

For each parameter of ASMs, a fixed setting was selected that yielded good performance, after initial pilot experiments. For the example images from the previous section and for the lung shapes, no shape alignment was performed (this improved performance) and a shape model was constructed in which 99.5% of the variance was explained ($f_v = 0.995$). For the brain structures, shape alignment was used (in this case better results were obtained with the use of alignment) and a shape model explaining 98% of the variance ($f_v = 0.98$) was constructed. The fact that f_v is lower for the aligned shapes is because alignment reduces the total amount shape variability and, thus, a higher proportion of the variation in the data can be attributed to noise.

The other settings were 4 levels of resolution ($L_{max} = 4$), 10 iterations per level ($N_{max} = 10$), profiles of length 5 ($k = 2$) and evaluation of 9 positions per iteration ($n_s = 4$). When fitting the shape model to the displaced landmarks, each mode was constrained within 2 times the standard deviation ($m = 2.0$). For the extended ASMs, at most 10 features were selected for each landmark and each resolution ($f_{max} = 10$). Training data were selected from 5 by 5 neighborhoods around each landmark ($N_{grid} = 5$). In the kNN classifier, 5 neighbors were used ($k_{NN} = 5$).

To compare different segmentations, the following

[9]This procedure of forward selection followed by backward selection is almost as effective as optimal 'floating' feature selection schemes [26]

experiment	object	image description	resolution	images		landmarks	
				test	training	fixed	total
A	right lung field	standard PA chest radiographs	256×256	115	115	3	40
B	left lung field	standard PA chest radiographs	256×256	115	115	3	40
C	corpus callosum	2D slices of MRI brain studies	320×256	45	45	3	50
D	cerebellum	2D slices of MRI brain studies	320×256	45	45	7	50

Table 1: Description of the objects and images used in the four segmentation experiments. The experiments are labelled A to D. There are two different image databases: chest radiographs used in Experiments A and B, and MRI brain data used in experiments C and D.

"overlap" measure Ω was used

$$\Omega = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP} + \mathrm{FN}}, \tag{9}$$

where TP stands for true positive (the area correctly classified as object), FP for false positive (area incorrectly classified as object), and FN for false negative (area incorrectly classified as background). When $\Omega = 1$ the overlap is complete and the result is perfect; for $\Omega = 0$ there is no overlap at all between the detected and true object. This measure more closely reflects the idea of a good segmentation than the average distance between the true and detected landmark location, because the latter is not sensitive to shifts of the landmarks along the contour.

In all experiments the performance when fitting the shape directly to the true landmarks (*cf.* Eq. (5)) was also computed because it indicates an upper bound for both the original method and the method with local features. For Experiments A and B manual segmentations by a second observer were available. Therefore Ω for the second observer can be compared with Ω for the automatic methods.

5 Results

The results of all experiments are given in Table 2. Note that fitting the shape model minimizes the distance between the predicted landmark position and the true landmark position; it does not necessarily optimize Ω. Therefore it is possible that an ASM scheme produces a set of model parameters $\mathbf{b}$ for which Ω is higher than Ω for fitting the shape model directly. This actually occurred in a few cases. Another practical measure of the optimal performance any automatic segmentation method that is trained with examples can achieve, is the variation between observers. This measure is given for Experiments A and B, where the median Ω of the ASM method with local features is close to median Ω of a second human observer.

In all cases ASMs with local features produced significantly higher Ω values than the original scheme ($p < 0.001$ in a paired t-Test for all experiments). This is also clear from Figure 2 which shows scatter plots for each experiment. In these plots, points which are above the diagonal line indicate images for which the segmentation with local features is better than the result of the original scheme. It is apparent that a substantial improvement is achieved for Experiment A, C, and D. Only for Experiment B, the left lung fields, there is a considerable number of cases where the original method has better performance. Figure 3 shows a typical result for each experiment.

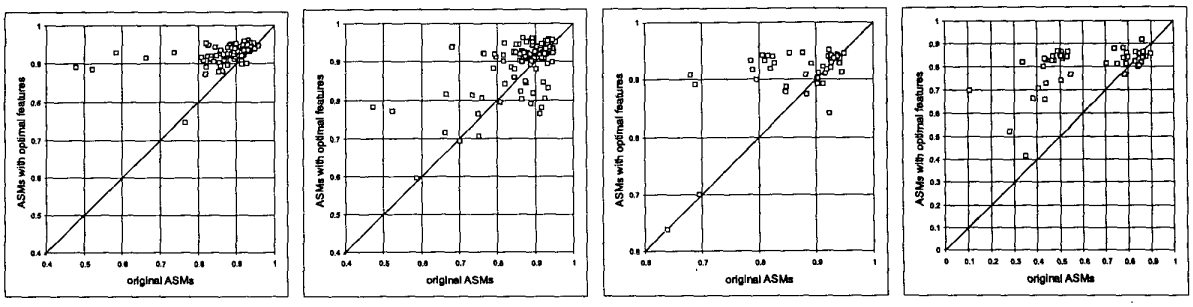
Experiment A: Right lung field	$\mu \pm \sigma$ (median)
ASMs	0.882 ± 0.074 (0.902)
ASMs with local features	0.929 ± 0.026 (0.933)
Fit of shape model	0.948 ± 0.030 (0.955)
Second observer	0.945 ± 0.017 (0.948)

Experiment B: Left lung field	$\mu \pm \sigma$ (median)
ASMs	0.861 ± 0.109 (0.891)
ASMs with local features	0.887 ± 0.114 (0.924)
Fit of shape model	0.942 ± 0.090 (0.955)
Second observer	0.934 ± 0.021 (0.938)

Experiment C: Corpus callosum	$\mu \pm \sigma$ (median)
ASMs	0.617 ± 0.206 (0.535)
ASMs with local features	0.805 ± 0.093 (0.837)
Fit of shape model	0.887 ± 0.052 (0.906)

Experiment D: Cerebellum	$\mu \pm \sigma$ (median)
ASMs	0.870 ± 0.078 (0.904)
ASMs with local features	0.910 ± 0.058 (0.927)
Fit of shape model	0.950 ± 0.014 (0.950)

Table 2: Mean, standard deviation and median results of the overlap measure Ω for all experiments. The original ASM scheme is compared with the method with the best local features and the result of directly fitting the shape model to the true landmark positions (*cf.* Eq. (5)). In Experiments A and B the mean, standard deviation and median results of a second human operator are also given.



Right lung fields | Left lung fields | Cerebellum | Corpus callosum

Figure 2: Scatter plots of each segmentation experiment. The overlap measure Ω for the original ASM scheme is plotted against Ω for the ASM method based on an optimal set of local features for each segmented image.

6 Discussion and conclusion

The fact that a set of features is selected for *each* landmark and *each* resolution separately is an important aspect of the new method. It turned out that the selected set of features varied considerably from experiment to experiment, landmark to landmark and resolution to resolution. More feature images can be used, by using higher moments, more (higher-order) derivatives and by relaxing the fixed relation between σ and α.

A more elaborate criterion for evaluating new landmark positions could be as follows. Currently landmarks are moved to those locations where the profile values are closest to 0 for points outside the object and closest to 1 for points inside the object. In practice the optimal profiles may be different. Especially if the object is very thin and the fitting occurs at a coarse resolution level, the innermost points of the profile may cross the border on the other side of the object! The actual profiles can be extracted from the training set and used to construct a model based on their mean and covariance matrices, that can steer the landmark displacement, in the same way as the original ASM scheme. Another enhancement would be to take into account the structure along the profile, instead of using local pixel classification for each position along the profile independently. This may improve performance. Consider a set of images, half show a black object on a white background, and the other half a white object on a black background. With local image features, it is impossible to classify locations as inside or outside the object. But a set of features measured along the profile can easily distinguish correct profiles, with an intensity jump at exactly the landmark location, from incorrect profiles, with no intensity jump or an intensity jump at a different location.

The example results from section 3 indicate that the proposed refined ASMs method may be especially useful to segment textured objects from textured backgrounds. An example could be segmentation in ultrasound images, but this remains to be tested. However, the method requires the texture of object and background to be different; otherwise - as in camouflaged objects - local classification is impossible and difficulties will arise. In that case an analysis of the complete profile would seem to be a preferable approach.

The original ASM scheme is an extremely fast segmentation method, yielding result in tenths of a second on regular PC hardware and 2D images. The new method is considerably more computationally expensive. However, all the feature selection can be done off-line. We used an optimized kNN classifier [27], and other, faster, classifiers could be used instead. Substantial speed improvements could be obtained by using pyramid-like schemes in which filters of large scale would be used on subsampled versions of the input images. Our current implementation of the ASMs method with local features requires twenty times as much computation time as the original ASMs scheme, and takes 4.1 seconds to complete on a 500 MHz Pentium III PC with the given parameter settings and images of 256 by 256 pixels.

In conclusion we have shown that active shape models, which provide a fast, effective, automatic, model-based method for segmentation problems in medical imaging, can be significantly improved through the use of an adaptive gray-level appearance model based on a non-linear classifier trained with an optimal set of local image features.

Acknowledgements

Marek Brejl, John Haller and Lisann Bolinger from the University of Iowa are gratefully acknowledged for allowing the use of their MR brain data and manual segmentations for this paper.

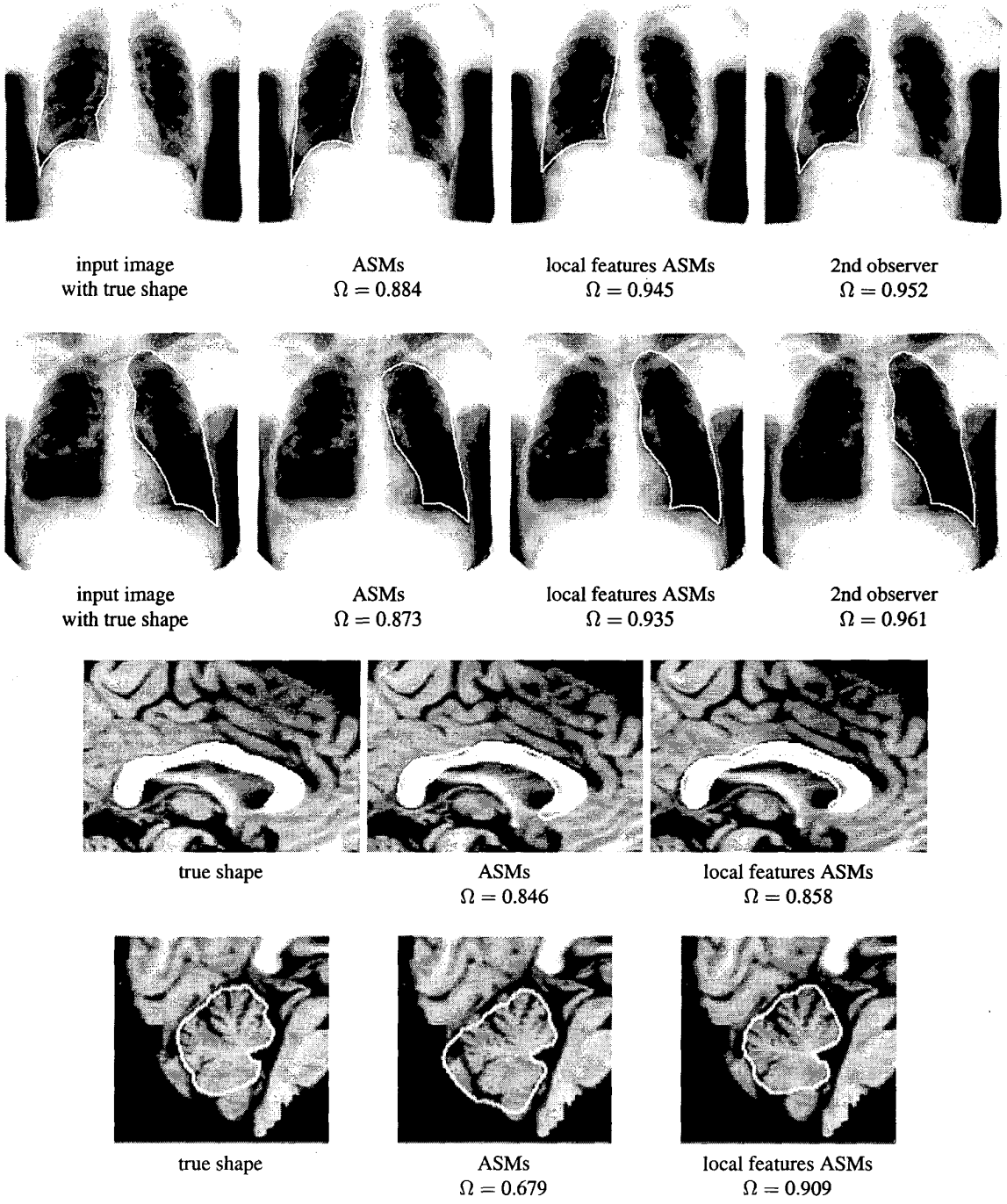


Figure 3: Example results for right lung field (Experiment A, top row), the left lung field (Experiment B, second row), the corpus callosum (Experiment C, third row) and the cerebellum (Experiment D, bottom row) segmentation. Segmentations are given by the thick white line. Below each segmentation Ω is given.

References

[1] T.F. Cootes, C.J. Taylor, D. Cooper, and J. Graham. Active shape models – their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, 1995.

[2] T.F. Cootes and C.J. Taylor. Statistical models of appearance for computer vision. Technical report, Wolfson Image Analysis Unit, University of Manchester, 1999.

[3] J.S. Duncan and N. Ayache. Medical image analysis: progress over two decades and the challenges ahead. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1):85–106, 2000.

[4] A. Blake and M Isard. *Active contours*. Springer, 1998.

[5] A.K. Jain, Y. Zhong, and M.P. Dubuisson-Jolly. Deformable template models: a review. *Signal Processing*, 71(2):109–129, 1998.

[6] N. Duta and M. Sonka. Segmentation and interpretation of mr brain images: An improved active shape model. *IEEE Transactions on Medical Imaging*, 17(6):1049–1067, 1998.

[7] A.K. Jain, Y. Zhong, and S. Lakshaman. Object matching using deformable templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(3):267–278, 1996.

[8] T.F. Cootes, G.J. Edwards, and C.J. Taylor. Active appearance models. In H. Burkhardt and B. Neumann, editors, *Proceedings of the European Conference on Computer Vision*, volume 2, pages 484–498, 1998.

[9] S.M. Pizer, D.S. Fritsch, P.A. Yushkevich, V.E. Johnson, and E.L. Chaney. Segmentation, registration, and measurement of shape variation via image object shape. *IEEE Transactions on Medical Imaging*, 18(10):851–865, 1999.

[10] M. Brejl and M. Sonka. Object localization and border detection criteria design in edge-based image segmentation: automated learning from examples. *IEEE Transactions on Medical Imaging*, 19:973–985, 2000.

[11] T.F. Cootes, A. Hill, C.J. Taylor, and J. Haslam. The use of active shape models for locating structures in medical images. *Image and Vision Computing*, 12(6):355–366, 1994.

[12] P.P. Smyth, C.J. Taylor, and J.E. Adams. Automatic measurement of vertebral shape using active shape models. In *Proceedings British Machine Vision Conference*, pages 705–714, 1996.

[13] S. Solloway, C.J. Taylor, C.E. Hutchinson, and J.C. Waterton. The use of active shape models for making thickness measurements from MR images. In *Proceedings of the 4th European Conference on Computer Vision*, pages 400–412, 1996.

[14] G. Behiels, D. Vandermeulen, F. Maes, P. Suetens, and P. Dewaele. Active shape model-based segmentation of digital X-ray images. In *MICCAI '99*, number 1679 in Lecture Notes in Computer Science, pages 128–137. Springer, Berlin, 1999.

[15] F. Vogelsang, M. Kohnen, J. Mahlke, F. Weiler, M.W. Kilbinger, B. Wein, and R.W. Günther. Model based analysis of chest radiographs. In *Proceedings SPIE*, volume 3979, pages 1040–1052, 2000.

[16] G. Hamarneh and T. Gustavsson. Combining snakes and active shape models for segmenting the human left ventricle in echocardiographic images. *IEEE Computers in Cardiology*, 27:115–118, 2000.

[17] T.F. Cootes and C.J. Taylor. A mixture model for representing shape variation. *Image and Vision Computing*, 17(8):567–574, 1999.

[18] N. Duta, A.K. Jain, and M. Dubuisson-Jolly. Automatic construction of 2-d shape models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(5):433–446, 2001.

[19] C. Goodall. Procrustes methods in the statistical analysis of shapes. *Journal of the Royal Statistical Society B*, 53(2):285–339, 1991.

[20] I. Dryden and K.V. Mardia. *The statistical analysis of shape*. Wiley, London, 1998.

[21] L.M.J. Florack, B.M. ter Haar Romeny, J.J. Koenderink, and M.A. Viergever. The Gaussian scale-space paradigm and the multiscale local jet. *International Journal of Computer Vision*, 18(1):61–75, 1996.

[22] J.J. Koenderink. The structure of images. *Biological Cybernetics*, 50:363–370, 1984.

[23] J.J. Koenderink and A.J. van Doorn. The structure of locally orderless images. *International Journal of Computer Vision*, 31(2/3):159–168, 1999.

[24] M. Unser. Local linear transforms for texture measurements. *Signal Processing*, 11:61–79, 1986.

[25] A. Whitney. A direct method of non parametric measurement selection. *IEEE Transactions on Computing*, 20:1100–1103, 1971.

[26] P. Somol, P. Pudil, J. Novovicova, and P. Paclik. Adaptive floating search methods in feature selection. *Pattern Recognition Letters*, 20(11–13):1157–1163, 1999.

[27] S. Arya and D.M. Mount. Approximate nearest neighbor queries in fixed dimensions. In *Proceedings of the 4th ACM-SIAM Symposium on Discrete Algorithms*, pages 271–280, 1993. `http://www.cs.umd.edu/~mount/ANN`.