

An Improved Active Shape Model: Handling Occlusion and Outliers

Nicolae Duta¹, Milan Sonka²

¹ Department of Computer Science, Michigan State University
East Lansing, MI 48824

² Department of Electrical and Computer Engineering, The University of Iowa
Iowa City, IA 52242

Abstract. An improvement of the *Active Shape* procedure identifying new examples of previously learned shapes using the point distribution model is presented. The novel segmentation and interpretation approach incorporates *a priori* knowledge about the objects of interest and their specific structural relationships to provide robust segmentation and labeling.

The method was utilized to successfully identify 10 neuroanatomic structures in 19 individual MR images and 2 car classes (left-right and right-left oriented) in 400 perspective images of street scenes.

1 Introduction

There have been numerous attempts to build models describing shape and appearance of non-rigid objects, and employ them for automated object identification in the analyzed images [1–5]. Among them, the Point Distribution Model (PDM) representing the variation of a set of shapes around the average that was designed by Cootes and Taylor has many favorable shape representation properties [6, 7].

We report an improvement of the *Active Shape* procedure [7] designed to find new examples of previously learned shapes using the point distribution model. This approach is particularly useful if variations in shape and appearance are difficult to model as is often the case with non-rigid objects or when point-of-view perspective is involved. The method presented below is generally applicable to virtually any task involving deformable shape analysis.

2 Object Occlusion and Shape Outliers in PDM-based Image Interpretation: A New Approach

In many areas of image segmentation and interpretation, reliable *a priori* knowledge is available to help guide the image analysis process. In many cases, approximate positions of individual objects can be determined from context. Knowledge about object sizes, shapes, gray level appearance, etc. can be acquired from a training set of examples.

2.1 Knowledge-Based Point Distribution Model

In order to take advantage of the available a priori knowledge, three additional features were included in the model: Gray-level appearance, border strength, and average position. In the implementation described below, we also used the implicit knowledge about object context representing inter-relationships of several objects.

Gray-level appearance is calculated in neighborhoods around each of the shape model points. It is determined for every shape model point j of each training image along a profile g_j of a constant length, centered at the point j . Since the profiles vary with gray level scaling, derivatives of the gray levels along each profile are determined and normalized.

Border strength is determined for each border segment of the model. Every two consecutive model points that lie on the object boundary define a border segment. To compute its strength, a local filtering is applied to each clique on that border segment. The filter is based on a pair of close parallel profiles.

Average position of each shape model point that is calculated in the image coordinates is also incorporated in the model.

Our **knowledge-based shape model** combines generally applicable parameters of the point distribution model and the knowledge-specific parameters appropriate for the image segmentation task in question. As such, the complete model is composed of:

1. The eigenvectors corresponding to the largest eigenvalues of the covariance matrix describing the *Allowable Shape Domain* [6, 7].
2. The average gray level appearance values for each point of the model.
3. The average border strength for corresponding border segments and the parameters of the mask (width, length) for which the strength was computed.
4. The average position of the points of the average shape.
5. Connectivity information (the number of shapes, point ordering along contours).

2.2 Searching for Objects: Model Fitting

The searching procedure developed for our PDM approach to image segmentation and interpretation is based on a model fitting strategy that substantially differs from the *Active Shape Procedure* of Cootes and Taylor. The difference is twofold. First, our search is entirely model driven meaning that segmentation hypotheses are not influenced by possibly misleading image data and do not use any preprocessing. At each step of the fitting process, several model location hypotheses are considered and evaluated. Second, an outlier detection and replacement procedure has been developed to detect misplaced points and infer their new positions. The outlier detection improves robustness and accuracy of the shape model fitting process.

The searching procedure consists of the following steps: 1) Model fitting using linear transforms, 2) model fitting using piecewise linear transforms, 3) outlier removal, 4) final point adjustment, 5) final outlier removal.

Model Fitting Function As a result of the hypotheses generation processes, shape model locations are sequentially hypothesized. In order to evaluate the model location hypotheses, a *fitness function* is needed to assess the agreement between the image data and the particular model instance. We have designed a *fitness function* $F = F_B / (F_{GA})^2$ that consists of two components:

1. *Fitness of the gray level appearance* F_{GA} is determined as the average squared Euclidean distance between the actual gray level appearance and the mean gray level profile incorporated in the shape model.
2. *Fitness of the border* F_B is calculated as the ratio between the aggregate response of all four point cliques along the contour and the maximum possible response (twice the number of cliques).

Model fitting using linear transforms Shape instance hypotheses specify the locations of all model points within the analyzed image. The hypotheses are generated using affine transformations and are applied to the model average position. The parameters of the linear transforms that contribute to the hypotheses generation are application dependent and reflect the a priori knowledge one has about the scale and pose of the objects of interest. For the two applications reported here the parameters were: scaling in the range [0.9, 1.1], step 0.1 (both applications); rotation $[-8^\circ, 8^\circ]$, step 4° (brain model) and $[-4^\circ, 4^\circ]$ (car model); and translation $[-4, 4]$ pixels, step 1 pixel in both x, y directions (brain model) respectively $[-128, 128]$ pixels (car model). Each hypothesis represents a rigid transform of the average shape, and all generated hypotheses are sequentially evaluated using the model fitting function and the best fit is determined. If the prior knowledge includes the fact that the objects of interest are always present in the images processed (as is the case of the brain images) no thresholding is subsequently done, otherwise the best fit value is thresholded and no object reported if its value is low.

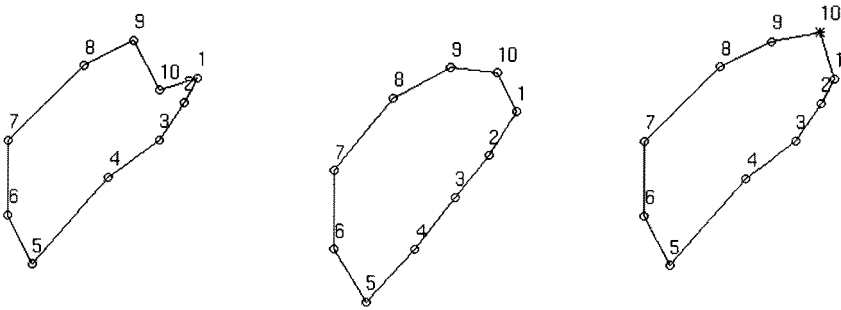
Model fitting using piecewise linear transforms Since non-rigid objects or object with inter-subject variability are discussed here, rigid linear transforms do not account for any potential deformations of the expected shape. Therefore, linear transforms (translation, rotation, and scaling) with a small range of parameters are applied to subsets of consecutive model points. Each model point is taken in turn as the center of such a subset. After the best position of the subset is obtained, a thresholding of the fitting value is performed and the center point of the set is considered to be occluded and discarded if the value is low. Otherwise, only the position of the center point is kept and the remaining points are discarded. The number of consecutive points that are considered for this transform is application dependent and is a function of local border strength and length. To preserve robustness, the center point is not moved if the two border segments adjacent to it are very weak.

Outlier removal: A new approach Under unfavorable circumstances, the previous step may introduce incorrectly determined vertices – **outliers**. This may happen if a subshape fitted by the previous step exhibits weak edges or if there exists another border of similar properties in the neighborhood. In the existing literature dealing with point distribution models, no outlier detection has

been introduced. Typically, when using PDM's, shapes that do not correspond to the allowed shape at any stage of the detection process are rejected (Fig. 1). In other words, the model parameters b_j (Eq. 2) must not exceed some maximum values. Cootes and Taylor propose to derive limits for b_j by examining the distributions of the parameter values required to generate the training set [7]. They recommend that b_j 's be chosen such that

$$-3\sqrt{\lambda_j} \leq b_j \leq 3\sqrt{\lambda_j} \tag{1}$$

since most of the population lies within three standard deviations of the mean. This approach can introduce two kinds of errors: 1) The shape/location hypothesis is completely rejected because one or two points are misplaced, such situation is documented in Fig. 1; or 2) the hypothesis is accepted even though one or two points are misplaced.



3 * sqrt (lambda) = [0.23 0.28 0.32 0.35 0.56 0.77]
 abs (b) = [0.54 0.05 0.24 0.35 0.21 0.28]
 Max. var. perc. = [2, 28, 28, 38, 16, 8, 21, 7, 5, 49,]

Fig. 1. Example of a shape hypothesis rejected by the Cootes' Active Shape procedure (left). Note an outlier responsible for rejection (marked by 10). The same shape hypothesis after our outlier removal and adjustment steps (right, adjusted vertex marked by *). The average shape is shown in the middle. The values given below the figure are specified by Eq. 1, 2, and 6.

To treat the problem of outliers in a systematic fashion, we have developed a new approach to outlier detection and position adjustment. The misplaced points are identified using the information about the relative positions of the shape model vertices that are implicitly included in the shape model. Let $z = (x', y')^T$ be the model point positions after the piecewise linear transforms were applied and the resulting shape was aligned with the shape average. According to the shape model, the hypothesized shape should satisfy

$$z = \bar{x} + Pb \tag{2}$$

where $\bar{\mathbf{x}}$ is the average shape, P is the matrix of the first t eigenvectors, and \mathbf{b} is a vector of weights. Therefore,

$$b_j = \sum_{i=1}^n P_{i,j}(z_i - \bar{\mathbf{x}}_i) = \sum_{i=1}^n v_{i,j} \quad (3)$$

where

$$|v_{i,j}| = |P_{i,j}(z_i - \bar{\mathbf{x}}_i)| \quad (4)$$

is the absolute variation induced by point i in parameter b_j . Let the percentage of variation induced by point i in parameter b_j be defined as

$$V_{i,j} = \frac{|v_{i,j}|}{\sum_{i=1}^n |v_{i,j}|} 100 \quad (5)$$

and let the maximum percentage of variation induced by point i in any of the parameters b_j be defined as

$$u_i = \max_{j=1..t} V_{i,j} \quad (6)$$

If all the points were to generate an equal amount of variation, then all the percentages u_i were approximately $100/n$, n being the number of object points. However, since outliers may be present, larger variation may be associated with some points – the outliers. A point is considered to be an outlier if the percentage of variance generated by its position in any of the parameters b_j of the model is more than $4\times$ greater than the average amount of variation. If several outliers are present, the variance is distributed among them and (perhaps) the well placed points. As a result, it is difficult to identify outliers if more than a few occur simultaneously. Once such misplaced points are detected, they must be moved to a new position that can be inferred from the alignment of the rest of the shape instance with the average shape.

Final point adjustment Some of the shape model points may have been declared outliers in the previous step. Consequently, their position may have been adjusted solely considering the average shape appearance and not considering the image data. Therefore, they must be subject to the position optimization step to better correspond with the image data.

Final outlier removal Resulting from the previous steps, or newly introduced during the final point adjustment, outliers may remain present in the shape model. Following the same outlier detection procedure as applied in the first outlier removal step, the outliers are identified and removed, no adjustment is attempted in this final step of model fitting.

3 Results

The method presented above was employed to design two PDM shape models. 1) Magnetic resonance images of human brain were segmented into neuroanatomic

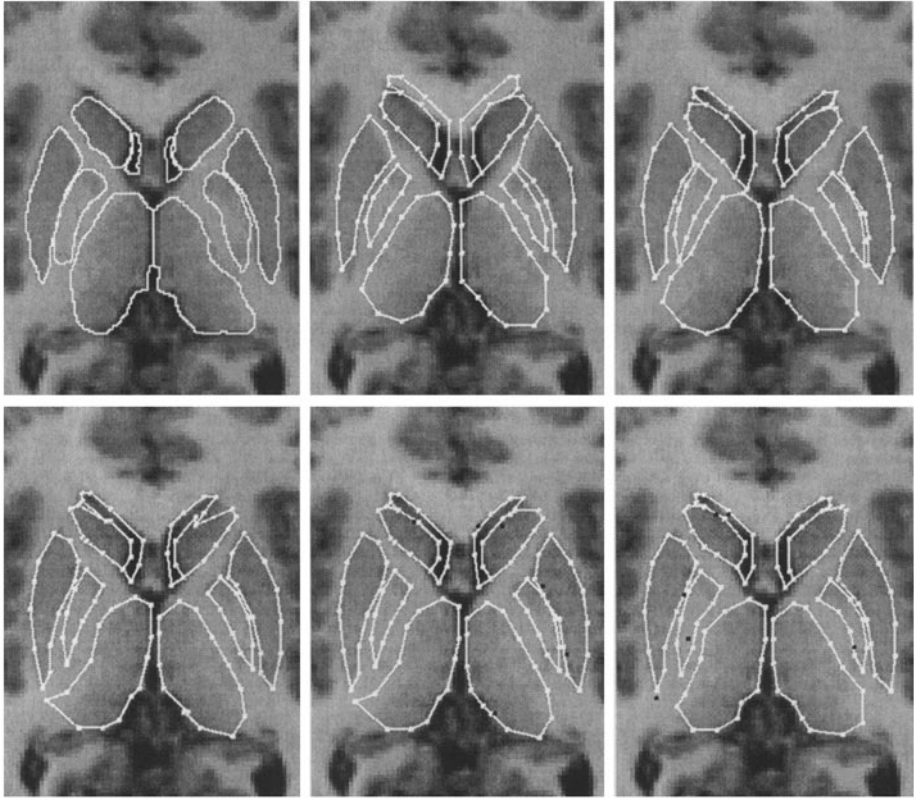


Fig. 2. Example of automated brain image segmentation and interpretation. Upper row from left to right: Manual tracings. Initial average position of the shape model. Optimal shape model position after linear transform step. Bottom row from left to right: Optimal shape model position after piecewise linear transform step. Outlier detection – outliers marked by dark dots). Final outlier detection – marked by dark dots and removed from consideration in the shape model.

brain structures. 2) Perspective images of moving cars on a street were used. Training image sets served for construction of shape models, and the method's performance was quantitatively assessed in separate testing image sets.

First, the PDM approach was trained in 8 images and tested in 19 MR brain images. Ten neuroanatomic structures were successfully segmented and interpreted in all images from the test set. Fig. 2 shows the observer-traced and computer-detected contours together with the intermediate results. In the test set, the neuroanatomic structures were identified with the labeling error of $7 \pm 3\%$, and the average border positioning error was 0.8 ± 0.1 pixels. The automatically-determined borders of the ten neuroanatomic structures exhibit a high level of accuracy and substantial speedup when compared to a previously reported genetic image interpretation approach [8].

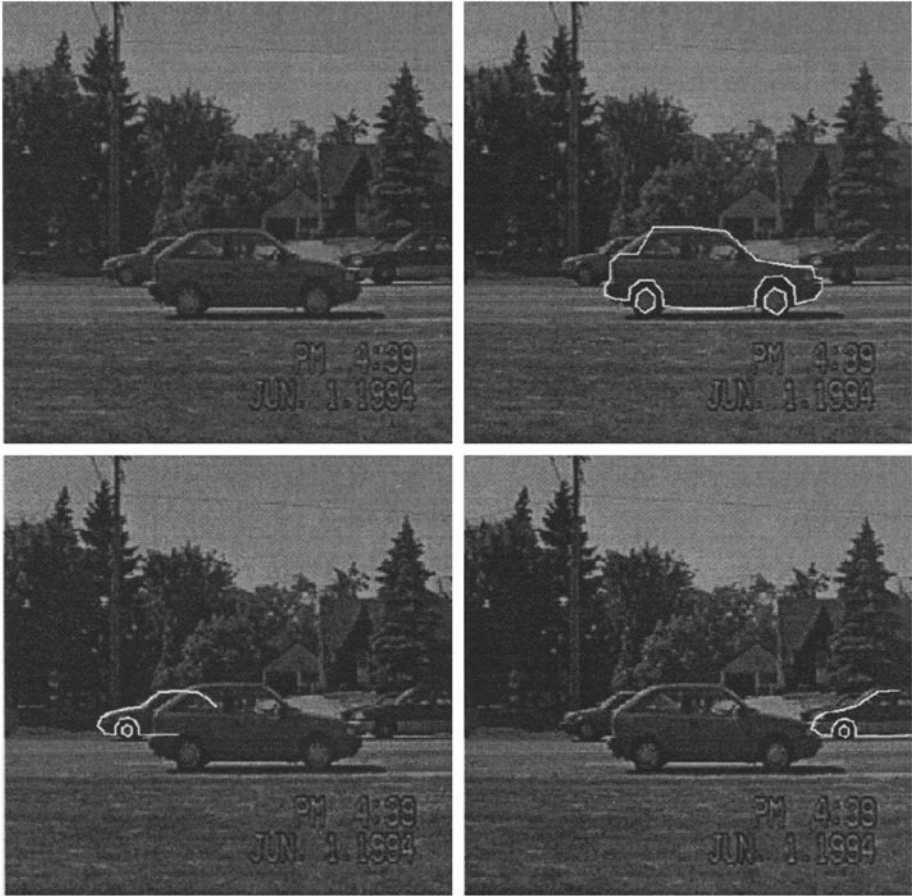


Fig. 3. Example of automated car segmentation. a) Original image. b) Optimal left-to-right car model position. c) Optimal right-to-left car model position. d) Second best right-to-left car model position.

Second, the PDM approach was employed in identification of two car classes (left-right and right-left orientation) in perspective images of street scenes. After training in 10 images, our approach correctly detected the cars present in 400 images from the test set. Fig. 3 shows an example of the computer-detected contours for the two car models. Note that, although the model was trained only on sedans, it was also able to accurately detect a hatchback car (except for the one contour point with an opposite convexity). Also note the good detection of the occluded cars in the background. As a matter of fact, all cars exhibiting less than 40% occlusion were successfully detected. The car segmentation results represent a substantial improvement over a recently published study [5].

4 Discussion

We expect further improvement of the method's performance by incorporating better structure positioning information. At this moment, no explicit statistical information is used concerning the relative positions of the objects. The only constraint that is enforced after each step is that no two objects overlap that belong to the same multi-object model. In case of object overlap, the faulting objects have to be grouped together. If an object is misplaced but does not overlap with another object, the hypothesis is accepted and a wrong detection may occur. While we did not experience such a behavior, the possibility remains present. Furthermore, no statistical information is used concerning the model points variation. Incorporation of such information may be useful for the outlier detection procedure.

5 Conclusion

A new fully automated segmentation and interpretation method has been presented. The method was utilized to identify 10 neuroanatomic structures in individual MR images and 2 car classes (left-right and right-left oriented) in perspective images of moving cars. In all cases, the method was independently trained and tested in separate images sets. The method correctly segmented and interpreted images from the two very dissimilar application areas.

Acknowledgments

This work was supported in part by the NSF grant IRI 96-16747.

References

1. M Kass, A Witkin, and D Terzopoulos. Snakes: Active contour models. In *Proceedings, First International Conference on Computer Vision, London, England*, pages 259–268, Piscataway, NJ, 1987. IEEE.
2. L H Staib and J S Duncan. Boundary finding with parametrically deformable models. *IEEE Trans. Pattern Anal. and Machine Intelligence*, 14(11):1061–1075, 1992.
3. A Jain, Y Zhong, and S Lakshmanan. Object matching using deformable templates. *IEEE Trans. Pattern Anal. and Machine Intelligence*, 18(3):267–277, 1996.
4. M Sonka, V Hlavac, and R Boyle. *Image Processing, Analysis, and Machine Vision*. Chapman and Hall, London, New York, 1993.
5. A Jain, M P Dubuisson, and S Lakshmanan. Vehicle segmentation using deformable templates. *IEEE Trans. Pattern Anal. and Machine Intelligence*, 18(3):293–308, 1996.
6. T F Cootes, A Hill, C J Taylor, and J Haslam. Use of active shape models for locating structures in medical images. *Image & Vision Computing*, 12(6):355–366, 1994.
7. T F Cootes, C J Taylor, D H Cooper, and J Graham. Active shape models – their training and application. *Computer Vision and Image Understanding*, 61:38–59, 1995.
8. M Sonka, S K Tadikonda, and S M Collins. Knowledge-based interpretation of MR brain images. *IEEE Trans. Med. Imaging*, 15:443–452, 1996.