Robust Active Shape Models: A Robust, Generic and Simple Automatic Segmentation Tool

Julien Abi-Nahed[1,2], Marie-Pierre Jolly[1], and Guang-Zhong Yang[2]

[1] Imaging and Visualization Department, Siemens Corporate Research
Princeton, New Jersey, USA
[2] Royal Society/Wolfson Foundation Medical Image Computing Laboratory
Imperial College London, London, UK
{julien.abi.nahed, marie-pierre.jolly}@siemens.com,
{julien.nahed, g.z.yang}@imperial.ac.uk

Abstract. This paper presents a new segmentation algorithm which combines active shape model and robust point matching techniques. It can use any simple feature detector to extract a large number of feature points in the image. Robust point matching is then used to search for the correspondences between feature and model points while the model is being deformed along the modes of variation of the active shape model. Although the algorithm is generic, it is particularly suited for medical imaging applications where prior knowledge is available. The value of the proposed method is examined with two different medical imaging modalities (Ultrasound, MRI) and in both 2D and 3D. The experiments have shown that the proposed algorithm is immune to missing feature points and noise. It has demonstrated significant improvements when compared to RPM-TPS and ASM alone.

1 Introduction

The Active Shape Model (ASM) [1] paradigm is a popular method for image segmentation when *a priori* information about the shape of the object of interest is available. The effectiveness of ASM depends on the ability to extract the correct feature points from the data. In fact, the model is represented by a collection of points, each trying to find a corresponding feature point in its vicinity. The ambiguity occurs when multiple feature points are possible candidates for a single model point. In this case, the model points do not have enough information to choose their correct partners especially when the initialization is far away from the correct solution. In other words poor segmentation results are often introduced when features are chosen locally for each model point. In practice, it is easy to extract a range of candidate feature points for each model point, but the process of model/feature matching while rejecting outliers is not trivial. There is a rich literature dealing with the robustness of ASM. Early methods tried to include intensity information leading to different versions of Active Appearance Models (AAM) [2]. From our experience, however, training and fitting AAM is more tedious, and over-fitting due to the higher dimensionality of the model

R. Larsen, M. Nielsen, and J. Sporring (Eds.): MICCAI 2006, LNCS 4191, pp. 1–8, 2006.

can be a significant issue for practical applications. Realizing the dramatic effect of the outliers on ASM's local search, Rogers *et al.* [3] tried to estimate ASM parameters using robust estimators instead of the traditional least squares. The idea is very interesting since least squares is not the optimal estimator when the distribution of the residuals between model and feature points is not Gaussian. Despite of increasing robustness achieved, the technique is still based on a local search for detecting feature points, thus leading to local minima. On the other hand, the original Robust Point Matching (RPM) algorithm [4] is a non-rigid point matching algorithm which allows a global to local search for the correspondences between two point sets. An important characteristic of RPM framework is that it allows the incorporation of deformations using any method. In [4] thin-plate splines were used to introduce deformations leading to the RPM-TPS algorithm. Also, RPM has the ability to match point sets of arbitrary size by rejecting outliers from both sets. Motivated by these observations, we combined the two complementary techniques, namely RPM and ASM into Robust Active Shape Models (RASM) where the deformations are controlled by ASM. We demonstrate its practical value by applying the proposed method in 2D and 3D, using *in vivo* data acquired with different medical scanners: Ultrasound (US) and Magnetic Resonance Imaging (MRI). The results demonstrate the practical value of the algorithm in automatic image segmentation. The rest of the paper is organized as follows. In Section 2 we describe the RASM algorithm in details. Section 3 presents experiments and results. Finally, in Section 4 we conclude and discuss future work.

2 Methodology

Following from the active shape model algorithm [1], we encode the template using a *Point Distribution Model* consisting of a mean shape $\overline{\mathbf{S}}$ which deforms along the main *modes of variation* recovered after applying Principle Component Analysis (PCA) on the shapes in the training set. An instance $\mathbf{S}$ of the model is represented by $\mathbf{S} = \overline{\mathbf{S}} + \mathbf{Pb}$ where the columns of matrix $\mathbf{P}$ are the t *modes of variation* we choose to retain, $\mathbf{b}$ is a vector representing the contribution of each *mode of variation* in a given instance (the b's are also called the shape parameters). Our goal is to match the template $\mathbf{S}$ with the features extracted from the scene. The problem can therefore be formulated as: given two sets of points $\mathbf{S} = \{S_i = S_{ix}, S_{iy}, S_{iz}, i = 1, ..., M\}$ and $\mathbf{Y} = \{Y_j = Y_{j_x}, Y_{j_y}, Y_{j_z}, j = 1, ..., F\}$, the goal is to align the two sets of points by recovering the correspondences and the deformations that would superimpose the two point sets in the same reference frame and reject outliers in both data sets (outliers in the features correspond to noise, and outliers in the template correspond to missing features).

2.1 Energy Formulation

For RASM, we choose to formulate the matching task in an optimization approach that is similar to [4], it is also possible to think of the problem in probabilistic terms, in fact the two approaches are equivalent, more details can be

found in [5]. RASM jointly estimates correspondences, an affine transformation and deformations. The system maintains a matrix $\mathcal{M}$ of size $(M+1)$ by $(F+1)$ to store both the correspondences and the outliers (as in RPM). The elements m_{ij} of $\mathcal{M}$ are defined as follow($i = 1, ..., M+1$ and $j = 1, ..., F+1$):

$$m_{ij} = \begin{cases}]0,1] & \text{if } S_i \text{ matches } Y_j \\ 0 & \text{otherwise} \end{cases}$$

$$m_{M+1,j} = \begin{cases}]0,1] & \text{if } Y_j \text{ is an outlier} \\ 0 & \text{otherwise} \end{cases} \qquad m_{i,F+1} = \begin{cases}]0,1] & \text{if } S_i \text{ is an outlier} \\ 0 & \text{otherwise} \end{cases}$$

The segmentation/matching problem is solved by minimizing the following objective energy function:

$$\begin{aligned} E(M,f,b) = & \sum_{i=1}^{M+1}\sum_{j=1}^{F+1} m_{ij}||Y_j - f(\overbrace{\overline{S_i} + \sum_{k=1}^{t} P_{i,k}b_k}^{S_i})||^2 + T\sum_{i=1}^{M}\sum_{j=1}^{F} m_{ij}\log m_{ij} \\ & + T_0\sum_{i=1}^{M} m_{i,F+1}\log m_{i,F+1} + T_0\sum_{j=1}^{F} m_{M+1,j}\log m_{M+1,j} \\ & + T trace[(f-I)^T(f-I)] \end{aligned} \tag{1}$$

The first term corresponds to the geometrical alignment of the two sets where m_{ij} are the fuzzy correspondence variables. They take a value between 0 and 1 following the soft-assign technique [4]. Y_j are the feature points and S_i the template points. ASM is being incorporated into RPM by expressing S_i as a PDM model. The function f represents an affine transformation. The second term is an entropy barrier function, T is a parameter used for deterministic annealing to control the degree of fuzziness of the correspondences m_{ij}(T_0 is the initial value of T). As the temperature T is decreased, the correspondences harden to get closer to the binary values of 0 and 1. The last term is a regularization term to prevent the affine transformation from unphysical behavior by penalizing the residual part of f different from the identity matrix I. This penalization is important at the begining of the annealing but vanishes as T approaches zero. Unlike the RPM formulation, we no longer need a regularization parameter for the deformations because ASM guarantee a valid final shape. Starting with an affine transformation we activate ASM local deformations when the correspondences are meaningful. Notice how the rejection of outliers is incorporated into the energy function without the need to estimate any additional parameters. In fact, the outliers are modeled by clusters whose centers (S_{M+1} and Y_{F+1}) coincide with the center of mass of the point sets. The clusters temperature which can be viewed as a search range [5] is kept high and does not follow any annealing schedule. This is done to allow the rejection of any point that does not show any strong evidence of being a valid match throughout the entire annealing process. The last two terms in the geometric alignment part of the energy, along with the third and fourth terms in the energy, represent the outliers incorporated in the

objective function. We found this formulation of the energy to give better results, it implies updating the outliers in $\mathcal{M}$ at each temperature. In order to guarantee that the derived correspondences are one to one, the following constraints need to be satisfied $\sum_{i=1}^{M+1} m_{ij} = 1 \quad \forall j = 1, ...F \quad$ and $\quad \sum_{j=1}^{F+1} m_{ij} = 1 \quad \forall i = 1, ...M$. This can be enforced explicitly via the iterative row and column Sinkhorn balancing normalization [4] of matrix $\mathcal{M}$.

2.2 RASM Algorithm

Given the two sets of points $\mathbf{S} = \{S_i, i = 1, ..., M\}$ and $\mathbf{Y} = \{Y_j, j = 1, ..., F\}$ where $\mathbf{Y}$ is the feature set and $\mathbf{S}$ the template. We start by setting the ASM shape parameters to zeros so that $\mathbf{S} = \overline{\mathbf{S}}$. The transformation is set to identity, and the initial/final temperatures with the annealing rate are also set beforehand as in [4]. The initial temperature T_0 is estimated as the largest square distance between all the points S_i and Y_j. The annealing rate (AR) is fixed at 0.95 (but could be anywhere between 0.9 and 0.99). T_f is set to the average square distance between the points of $\overline{\mathbf{S}}$. Like RPM, RASM is a two step algorithm. In the first step, we evaluate the correspondences by differentiating the energy function w.r.t m_{ij} and setting the result to zero, we get the following update equations:

$$\begin{aligned}
m_{ij} &= e\left(\frac{-||Y_j - f(\overline{S_i} + \sum_{k=1}^{t} P_{i,k} b_k)||^2}{T}\right) \forall i = 1, ..., M \text{ and } \forall j = 1, ..., F \\
m_{M+1,j} &= e\left(\frac{-||Y_j - f(S_{M+1})||^2}{T_0}\right) \forall j = 1, ..., F \\
m_{i,F+1} &= e\left(\frac{-||Y_{F+1} - f(\overline{S_i} + \sum_{k=1}^{t} P_{i,k} b_k)||^2}{T_0}\right) \forall i = 1, ..., M
\end{aligned} \tag{2}$$

In the second step the affine transformation as well as deformations are estimated given the correspondences m_{ij}. First we create an intermediate point set $\mathbf{V} = \{V_i = \frac{\sum_{j=1}^{F} m_{ij} Y_j}{\sum_{j=1}^{F} m_{ij}}, i = 1, ..., M\}$ and minimize the equation:

$$E(f, b) = \sum_{i=1}^{M} W_i ||V_i - f(\overbrace{\overline{S_i} + \sum_{k=1}^{t} P_{i,k} b_k}^{S_i})||^2 \tag{3}$$

using weighted least squares. We use binary weights to account for outliers in the template. If $W_i = \sum_{j=1}^{F} m_{ij}$ is less than a threshold (e.g. 0.1), the weight is set to zero and the corresponding point is not taken into account at this particular iteration. At the beginning, we only solve for f as an affine transformation and assume $\mathbf{b} = 0$. Once the temperature is cool enough, more precisely when $T = T_f + 0.2(T_0 - T_f)$ (i.e. for the last 20% of the process), we introduce ASM deformations. At this point, we jointly solve for f as an affine transformation using weighted least squares and solve for $\mathbf{b}$ using the classical ASM iterations [1].

Fig. 1. A typical evolution of RASM algorithm when applied to the left ventricle feature points extracted from a B-mode 2-chamber ultrasound image. Crosses are model points, dots are feature points, gray points are outliers. (a) initialization; (b) alignment after an affine transformation. (c) final shape recovered by RASM; and (d) final shape recovered by RPM-TPS.

3 Experimental Results

3.1 Left Ventricle in 2D Cardiac Ultrasound

B-mode 2-chamber view echocardiography images were acquired with an Acuson Sequoia 256 Scanner (Mountain View, CA). Using 33 patients we trained two ASM models of the left ventricle (at ED and ES) using [6]. The derived models were subsequently applied to 20 patients with known cardiac dysfunction. A typical RASM example is shown in Figure 1 where we follow the evolution of the ED model to align to the feature points. The feature detection was performed using the dedicated technique proposed in [7]. Figures 1(a)-(c) illustrate the initialization of the algorithm, intermediate result with affine transformation, and final result with affine transformation and ASM deformations. Figure 1(d) shows the results obtained when applying RPM-TPS. This example illustrates a major strength of the algorithm. On the left side of the endocardial border (the anterior wall), two major subsets of feature points can be observed. The ASM search algorithm yields feature points from a mixture of these two subsets and the resultant shape is a compromise between the true and outlier feature points. With RASM, the outliers can be correctly identified, resulting in a more faithful depiction of the endocardial border. Moreover the global shape is preserved and the missing feature points (near the apex for example) are accounted for, through the outliers in the template. The RPM-TPS algorithm (Matlab implementation available on the web) results in a bad shape that does not look like a left ventricle. In general, we observed that RPM-TPS did not allow control of the deformations and resulted in contours that did not resemble the desired shape. On the other hand, even though ASM segmented contours that resembled left ventricles, it was not able to recover the correct feature points and the contours did not align with image features. RASM however was able to recover left ventricles shapes aligned with image features. Figure 2 shows more examples of RASM performance for the segmentation of the left ventricle in cardiac ultrasound images.

To assess the robustness of RASM, and the effect of the density of the model points on the performance, we performed several tests on synthetic and *in vivo*

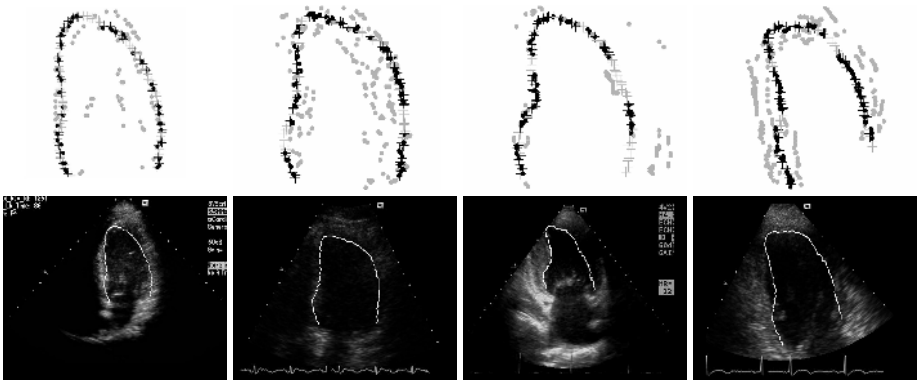


Fig. 2. Point sets and RASM segmentation overlayed on the images

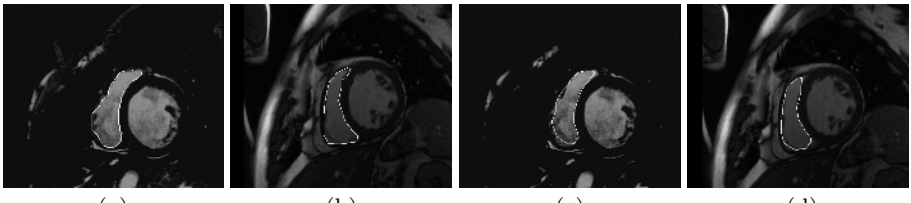
2D data in [8], where we showed that the algorithm could tolerate at most 50% of feature outliers. The robustness was shown to depend on the nature of the noise and on the number of missing model points missing in the data. We also compared the recovered contours with the ground truths drawn by a medical expert. We computed the distance between every point on a segmented contour and the closest point on a ground truth contour (and vice versa). We found an average distance of 0.3mm and a maximum distance of 1.9mm which is totally acceptable to compute volumes and ejection fraction.

3.2 Right Ventricle in 3D Magnetic Resonance Imaging

A right ventricle (RV) model at ED was built from 13 hand segmented CT and MR volumes. The CT volumes were acquired using a Siemens SENSATION 16 scanner and segmented using the multi-label segmentation method proposed in [9]. The MR slices were acquired using a Siemens SONATA 1.5T scanner and contours were manually drawn by experts on all slices. After constructing meshes from the segmentation using marching cubes, we applied the harmonic embedding technique proposed by [10] to parameterize these shapes. This powerful technique provides a statistically and geometrically optimal 3D ASM model. To test the RASM algorithm, we acquired four breath-hold retrospectively ECG-gated trueFISP cine MRI sequences on a Siemens SONATA 1.5T scanner. The short axis images were situated between the valve plane and the apex of the RV. The region of interest was automatically determined by first locating the left ventricle using the technique proposed in [11] and then identifying the right ventricle blood pool next to it. Then, feature points were extracted with a simple Canny edge detector. Figure 3 shows the evolution of the RASM algorithm recovering the RV. Figure 4 shows an example of RASM segmentation. We also show the ground truth manually drawn by experts for comparison. In order to validate the segmentation results we compared the recovered contours (for the four subjects) from RASM with the ground truth contour as in 2D. We found an average error distance of 1.1mm and a maximum of 3.2mm. Furthermore, we

(a) (b) (c)

Fig. 3. A typical evolution of RASM algorithm when applied to the right ventricle feature points extracted from MR images. Crosses are model points, dots are feature points, gray points represent outliers. (a) initialization; (b) alignment after an affine transformation. (c) final shape recovered by RASM.



(a) (b) (c) (d)

Fig. 4. Ground truths (a)/(b) and segmented contours (c)/(d) recovered after 3D RASM fitting of the RV model on MR data

studied the evolution of the energy while changing the time when ASM deformations are introduced in the algorithm. On average, we found that [10%-30%] of the total time is a valid range for this parameter. Notice here a powerful property of RASM in 3D: we do not need to extract the feature points in 3D. Instead we perform several 2D feature detection, and RASM will fit the model points to the extracted feature points. This is particularly useful for any modality that acquires multiple slices to analyze a 3D structure.

4 Discussion and Conclusion

Non rigid point matching is very important in computer vision. But how much to deform and how to deform the model remain critical issues. RPM-TPS matches point sets using thin-plate splines for deformations without an explicit prior on the final shape [4]. As it was first designed for registration, RPM-TPS is rather ill-suited for the task of segmentation. We preferred to incorporate *a priori* knowledge using the simple and fast ASM which despite being very useful for cluttered scenes, is seriously limited by the ability of the feature extractor to detect the correct points. The strength of RASM is that given two datasets, it is able to establish as many correspondences as possible between them while rejecting outliers. The size of the two datasets is irrelevant to the performance of the technique. These are the strengths inherited from the RPM framework. RASM

also inherited the major strengths in ASM namely the model specificity, generalization ability and compactness that allow only legal shapes to be detected. In this paper, we have shown that by the incorporation of outliers rejection, RASM is able to match point sets of arbitrary size more robustly than ASM or RPM-TPS alone. The results obtained have demonstrated the strength and potential clinical value of the technique. In several cases, RASM succeeded when both RPM-TPS and ASM failed. The quantification of the segmentation results are very encouraging and could be improved by making the feature extraction more application-dependent. More specifically, we plan to introduce weights on the features similar to the work in [12]. The new information added would increase the robustness and accuracy of RASM. In the long term, we are also looking into extending the RASM framework to 4D.

References

1. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models — their training and application. Computer Vision and Image Understanding **61** (1995) 38–59
2. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. IEEE Trans. Pattern Anal. Mach. Intell. **23** (2001) 681–685
3. Rogers, M., Graham, J.: Robust active shape model search. In: ECCV (4). (2002) 517–530
4. Chui, H., Rangarajan, A.: A new algorithm for non-rigid point matching. In: CVPR. (2000) 2044–2051
5. Chui, H., Rangarajan, A.: A feature registration framework using mixture models. In: MMBIA. (2000) 190
6. Huang, X., Paragios, N., Metaxas, D.N.: Establishing local correspondences towards compact representations of anatomical structures. In: MICCAI(2). (2003) 926–934
7. Jolly, M.P.: Assisted ejection fraction in b-mode and contrast echocardiography. In: ISBI. (2006) 97–100
8. Abi-Nahed, J., Jolly, M.P., Yang, G.Z.: Robust active shape models. In: MIUA. (2005) 115–118
9. Grady, L., Funka-Lea, G.: Multi-label image segmentation for medical applications based on graph-theoretic electrical potentials. In: ECCV Workshops CVAMIA and MMBIA. (2004) 230–245
10. Horkaew, P., Yang, G.Z.: Construction of 3d dynamic statistical deformable models for complex topological shapes. In: MICCAI (1). (2004) 217–224
11. Jolly, M.P.: Combining edge, region, and shape information to segment the left ventricle in cardiac mr images. In: MICCAI. (2001) 482–490
12. Lin, N., Papademetris, X., Sinusas, A.J., Duncan, J.S.: Analysis of left ventricular motion using a general robust point matching algorithm. In: MICCAI(1). (2003) 556–563