An Adaptive Potential for Robust Shape Estimation

Jacinto Nascimento and Jorge Marques
Instituto de Sistemas e Robótica - Instituto Superior Técnico
Av. Rovisco Pais, 1049-001 Lisboa, PORTUGAL
{pcjan@pop.,jsm@}isr.ist.utl.pt

Abstract

This paper describes an algorithm for shape estimation in cluttered scenes. A new image potential is defined based on strokes detected in the image. The motivation is simple. Feature detectors (e.g., edge points detectors) produce many outliers which hamper the performance of boundary extraction algorithms. To overcome this difficulty we organize edges in strokes and assign a confidence degree (weight) to each stroke. The confidence degrees depend on the distance of the stroke points to the boundary estimates and they are updated during the estimation process. A deformable model is used to estimate the object boundary, based on the minimization of an adaptive potential function which depends on the confidence degree assigned to each stroke. Therefore, the image potential changes during the estimation process. Both steps (weight update, energy minimization) are derived as the solution of a maximum likelihood estimation problem using EM algorithm.

Experimental tests are provided to illustrate the performance of the proposed algorithm.

1 Introduction

Active contours estimate the object boundary using a deformable curve. During the estimation process the model points move under the influence of image forces and internal forces. Image forces attract the model towards specific image features (e.g., edge points) and internal forces try to keep shape coherence during the convergence process [12].

The design of image forces has been thoroughly investigated (e.g., see [5, 6, 7, 15]). The main difficulty concerns the presence of invalid features (outliers) which are not located at the object boundary and attract the elastic model towards wrong shape configurations.

Several strategies have been proposed to improve the performance of active contours e.g., the use of a validation gate to reduce the search region [3], nonlinear filtering techniques with non-Gaussian distributions [10], the use of geometric and dynamic constraints to reduce shape and motion variability [4, 7] and robust estimation techniques which are able to reduce the influence of outliers on the final shape estimates [14].

A different approach is proposed here. The shape estimate is obtained by the minimization of a potential function as in the original snake algorithm. However a new adaptive potential is used which reduces the influence of outliers.

BMVC 2001 doi:10.5244/C.15.36

The method proposed in this paper is based on two key ideas. First, middle level features (strokes) are used instead of low level ones (edge points). Middle level features are more informative and reliable. Their use has been recently proposed by several authors in [9, 11, 14, 16]. Second, a confidence degree (weight) is assigned to each stroke. All strokes contribute to the image potential but with different weights. Weight assignment is not performed on a heuristic basis but it is obtained using a probabilistic model for the observed data and the EM algorithm.

The paper is organized as follows. Section 2 describes the estimation of active contours parameters, assuming that we know which features are valid and which are outliers (known labeling). Section 3 extends these ideas to the case in which such labeling information is unknown. Section 4 deals with the optimization issues and section 5 presents the experimental results. Section 6 concludes the paper.

2 Known Labeling

This section addresses shape estimation assuming that we know which features are valid and which are outliers. For the sake of simplicity no regularization forces will be considered in this section. Let y be the the set of all features detected in an image and let us assume that y is organized in strokes $y = \{y^1, \ldots, y^N\}$, y^j being the set of observations (edges points) belonging to the j-th stroke. Let x be a contour model defined by a sequence of 2D points $x_i, i = 1, \ldots, M$. The goal is to approximate the data contained in y by the contour model x. To accomplish this, we shall consider the potential function

$$P(x_i; y, k) = -\sum_{j,\, n} \mathbf{\Phi}(x_i; y_n^j, k^j) \tag{1}$$

x_i is the i-th model unit; y_n^j is the n-th observation of the j-th stroke; $k = \{k^1, \ldots, k^N\}$ is a set of stroke labels, ($k^j = 1$ if the j-th label is valid; $k^j = 0$ otherwise) and $\mathbf{\Phi}$ measures the influence of y_n^j on the image point x_i. The contribution of each feature y_n^j to the potential is defined by

$$\mathbf{\Phi}(x_i; y_n^j, k) = \begin{cases} \alpha\, G(|y_n^j - x_i|^2) & k^j = 1 \\ V & k^j = 0 \end{cases} \tag{2}$$

where $G(|y_n^j - x_i|^2)$ is a Gaussian kernel and V is a constant.

If y, k were known the contour model would be obtained by minimizing the contour energy

$$\hat{x} = \arg\min_x \sum_i P(x_i, y, k) \tag{3}$$

Equation (3) is equivalent to the snake algorithm with the Cohen potential [6] provided that we assume that all data is valid i.e., $k^j = 1, \forall_j$.

The problem may also be addressed in a probabilistic framework, by assuming the y, k are random variables with probability density function

$$p(y, k \mid x) = \alpha\, e^{-\sum_i P(x_i)} \tag{4}$$

The log likelihood function is

$$l(x;y,k) = \log p(y,k \mid x) = C - \sum_i P(x_i) \tag{5}$$

and the minimization of the log likelihood function leads to the same optimization problem defined in (3). In practice we do not know which features are valid and which are outliers. The labels k^j are therefore unknown. This problem is addressed in the next section.

3 Adaptive Potential

Since the stroke labels are unknown in practice, the object contour should be estimated by maximizing the likelihood function of the observed data

$$\log p(y \mid x) = \log \sum_k p(y,k \mid x) \tag{6}$$

This is however a difficult problem. One way to circumvent this difficulty is by using the EM algorithm [8] which optimizes the ML criteria by using an auxiliary function

$$U(x,\hat{x}) \triangleq E_k\{\log p(y,k \mid x) \mid y, \hat{x}\} \tag{7}$$

Using (4)

$$\begin{aligned} U &= \sum_j E_k\Big\{\log p(y^j,k^j \mid x)\Big\} \\ &= \sum_j \; w^j \log p(y^j,k^j=1 \mid x) \; + (1-w^j)\log p(y^j,k^j=0 \mid x) \end{aligned} \tag{8}$$

where $w^j = p(k^j = 1 \mid y^j, \hat{x})$.

The second term (8) (outlier potentials) can be discarded since it does not depend on x. Therefore the objective function becomes

$$U = C - \sum_i \mathcal{P}_a(x_i, y) \tag{9}$$

where

$$\mathcal{P}_a(x_i,y) = \sum_j \Big(-\sum_n G(|x_i - y_n^j|^2)\Big) w^j \tag{10}$$

will be denoted as an adaptive potential since it depends on the confidence degrees of the image strokes w^j which vary during the estimation process. The weights w^j are computed in the E step of the EM algorithm

$$\begin{aligned} w^j = p(k^j \mid y^j, \hat{x}) &= \alpha \; p(y^j \mid k^j, \hat{x}) \; p(k^j \mid \hat{x}) \\ &= \alpha \; c^j \prod_i \; e^{\sum_n G(|y_n^j - x_i|^2)} \end{aligned} \tag{11}$$

where c^j is a constant.

4 Contour Estimation

The cost function to be considered contains two terms: a regularization term based on strings with average length l_0, defined as in [13] and an image dependent term given by (10), i.e.

$$J = \sum_i (l_i - l_0)^2 + \mathcal{P}_a \tag{12}$$

$l_i = |x_{i+1} - x_i|$ is the distance between consecutive model points and l_0 is the average distance specified by the user.

The minimization of (12) is performed in the M-step, i.e.

$$\hat{x}^{t+1} = \arg\max_x J(x, \hat{x}^t) \tag{13}$$

for example using the gradient algorithm

$$x^{t+1} = x^t - \gamma \nabla_x J \tag{14}$$

where ∇_x is the gradient operator defined by $\nabla_x J = [\nabla_{x_1} J, \ldots, \nabla_{x_M} J]^T$.

Equation (14) can be rewritten as

$$x^{t+1} = x^t - \gamma_i f_{int} + \gamma_e f_{img} \tag{15}$$

where $f_{int}(x_i)$ and $f_{img}(x_i)$ are interpreted as internal and external forces given by

$$f_{int}(x_i) = -2\frac{l_i - l_0}{l_i}(x_{i+1} - x_i) \tag{16}$$

$$f_{img}(x_i) = \frac{1}{\sigma^2} \sum_j w^j \sum_n (y_n^j - x_i) G(|y_n^j - x_i|^2) \tag{17}$$

This result is related to other works. In [1] it is shown that several methods share the same structure and belong to a unified framework in which model points are attracted towards data centroids under the influence of external forces

$$f_{img}(x_i) = \mu_i(\xi_i - x_i) \tag{18}$$

After straightforward manipulation it is concluded that the algorithm developed in this paper also belongs to this framework and the external forces (17) can be rewritten as in (18) with

$$\mu_i = \sum_j w^j \sum_n \vartheta_i(y_n^j) \qquad\qquad \xi_i = \frac{\sum_j w^j \sum_n y_n^j \vartheta_i(y_n^j)}{\sum_j w^j \sum_n \vartheta_i(y_n^j)} \tag{19}$$

These expressions also suggest that other algorithms can be obtained by adopting other choices for $\vartheta_i(y)$. In this paper $\vartheta_i(y) = G(|y_n^j - x_i|^2)$. Similar expressions for edge point features were also derived by [2].

5 Experimental Results

This section presents experimental results obtained with synthetic and real images. We compare the proposed algorithm with the snakes algorithm obtained by assuming that all the data features are valid. In each iteration, the boundary model is resampled at equally spaced points. The gain γ_i used in (15) is chosen as proposed in [5], through the normalization of the internal forces. For the external forces we use independent gain factors acting on each model unit as in [1].

Example 1

Suppose that we want to estimate the contour of a square in the presence of a large number of outlier strokes. The data and the initial estimate of the contour are shown in the left column of Fig. 1. This figure shows the results obtained with Snake potential (upper row) and with the adaptive potential proposed in this paper (lower row). It is concluded that the adaptive potential reduces the influence of outliers and allows an accurate estimation of the object boundary.

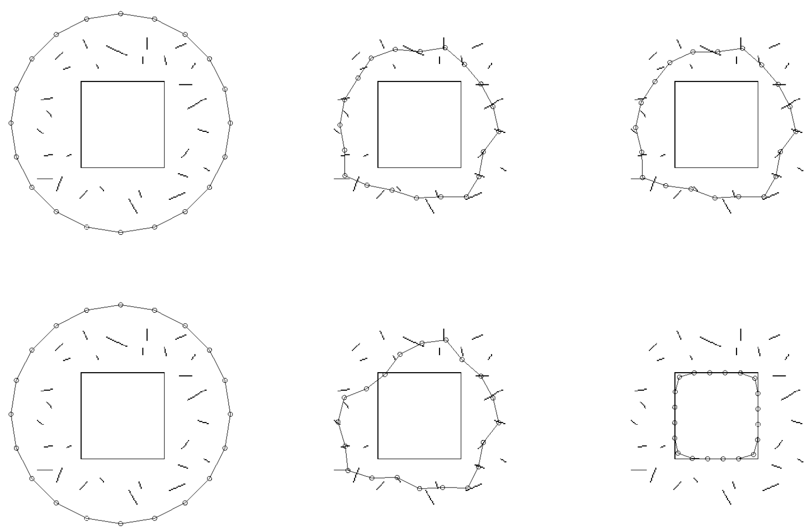

Figure 1: Results obtained with Snake potential (top row) and with the adaptive potential (bottom row). Each row shows initial, intermediate and final results.

Example 2

Another example is shown in Fig. 2. Two strokes are detected in the vicinity of the hand (left images): a valid stroke (hand boundary) and an outlier (white bar). Both methods were used to estimate the hand. It is observed that the adaptive potential manages to solve this problem as well while the classic potential leads to a local maximum. Fig. 3 shows the evolution of both potentials during the convergence process. The dark regions are the potential valleys which attract the model units.

Figure 2: Results obtained with Snake potential (top row) iterations 1, 7, 20 and with adaptive potential (bottom row) iterations 1, 7, 10.

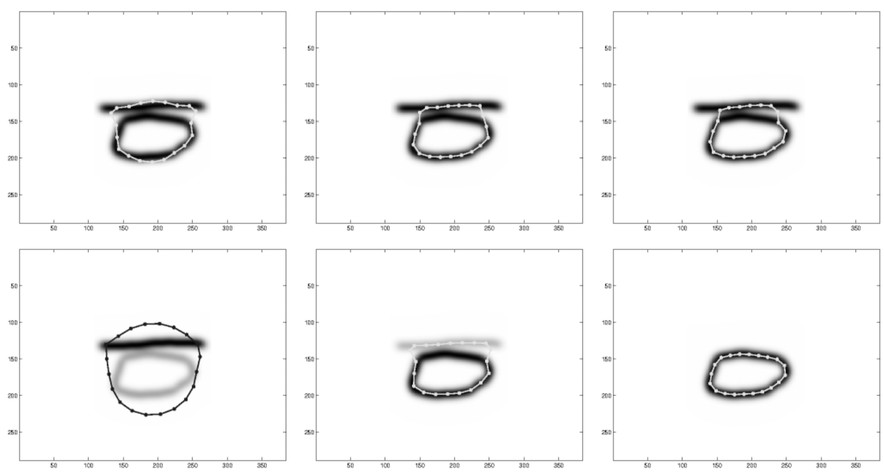

Figure 3: Image potential obtained with Snake potential (top row) iterations 4, 10, 20 and with adaptive potential (bottom row) iterations 2, 5, 10.

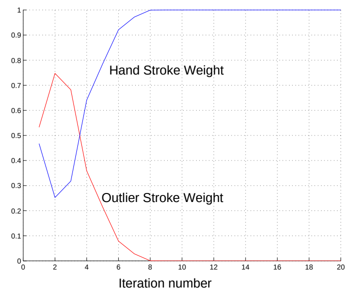

Figure 4: Evolution of the weights.

We can see that significant changes occur in the case of the *EM* based potential: the valley associated with the outlier stroke is filled during the estimation process. This is due to the variation of the weights illustrated in Fig. 4. At the first iterations both strokes had similar weights. However the weight of the hand stroke increases during the convergence process while the weight of the sweater stroke decreases.

Example 3

This example illustrates the performance of the proposed algorithm in the estimation of car boundaries. Figures 5,6 show the results obtained with the Snake potential and with the adaptive potential in two cases. These examples illustrates two typical situations. In the first example only a poor estimate of the object shape is available. The initial contour is therefore very far from the car boundary. In the second example, there is a good shape estimate but there is a significant shift with respect to the true boundary. The initial pose estimate is poor. The shape estimation with snake potential fails in both cases while the proposed algorithm manages to solve both problems well.

Figure 7 shows the strokes detected in both images.

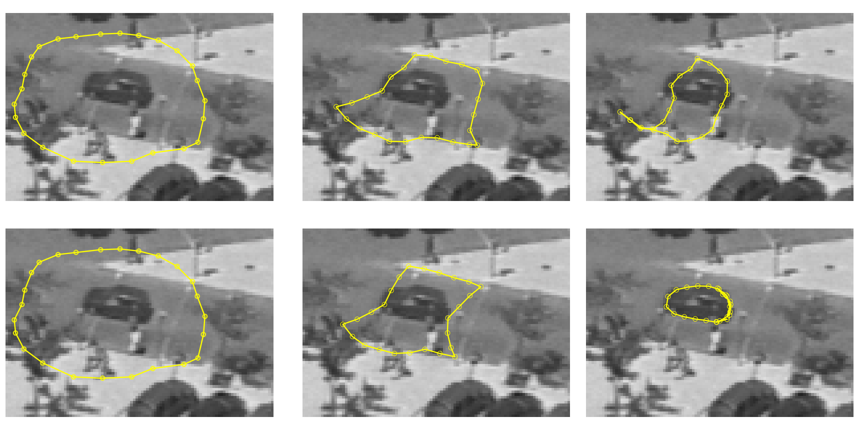

Figure 5: Results obtained with Snake potential (top row) and adaptive potential (bottom row) iterations 1, 7, 40.

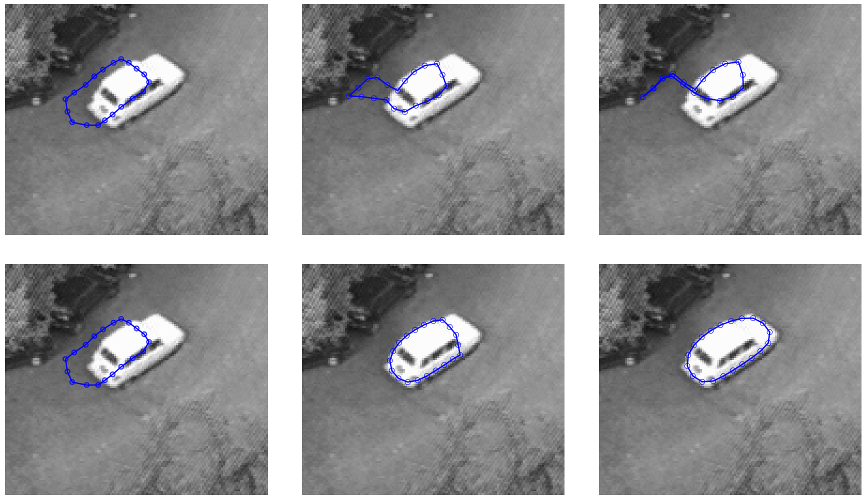

Figure 6: Results obtained with Snake potential (top row) iterations 1, 4, 30 and adaptive potential (bottom row) iterations 1, 4, 8.

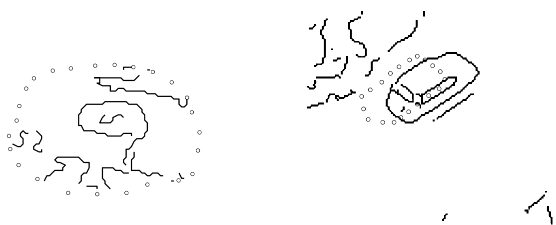

Figure 7: Strokes detected in the images (circles are the initial position).

6 Conclusions

This paper proposes a new algorithm for the estimation of object boundary in the presence of outliers. The object boundary is approximated by a deformable contour as in snakes. Model points are deformed by internal forces and by external forces computed using an image potential. However, instead of using the classic potential function which remain invariant during the convergence process, an adaptive potential is proposed which is able to discard the influence of outliers. This is achieved as follows. Image features (edges points) are organized in strokes and each stroke is either classified as valid or invalid (outlier). Since this information is not available a confidence degree is assigned to each stroke which is updated during the estimation process. Therefore all strokes contribute to the image potential function but with different weights. The image potential and the contour model are recursively estimated in a ML framework by the EM algorithm.

Experimental tests have shown that the proposed algorithm is robust and provides much better results than the snake algorithm in the presence of clutter.

References

[1] A. Abrantes and J. Marques, ”A Class of Constrained Clustering Algorithms for Object Boundary Detection”, *in IEEE Trans. Image Process.*, pp. 1507-1521, 1996.

[2] Abrantes, Private Communication, 2000.

[3] A. Blake and M. Isard, ”Active Contours”. Springer, 1998.

[4] A. Blake, R. Curwen and A. Zisserman, ”A framework for Spatiotemporal Control in the Tracking of Visual Contours”, *in Int. Journal of Computer Vision*, vol. 11(2), pp. 127-145, 1993.

[5] L. Cohen, ”On Active Contour Models and Ballons”, *in CV GIP: Image Understanding*, vol. 53(2), pp. 211-218, 1991.

[6] L. Cohen and I. Cohen, ”Finite Element Methods for Active Contour Models and Ballons for 2-D and 3-D Images”, *in IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15(11), pp. 1131-1147, 1993.

[7] T. Cootes, C. Taylor, D. Cooper and J. Graham, ”Active Shape Models - Their training and Application”, *in Computer Vision and Image Understanding*, vol. 61(1), pp. 38-59, 1995.

[8] A. Dempster, M. Laird and D. Rubin, ”Maximum Likelihood from incomplete data via the EM-Algorithm”, *in Journal of the Royal Statistical Society*, vol. B(39), pp. 1-38, 1977.

[9] H. Gu, Y. Shirai and M. Asada, ”MDL-Based Segmentation and Motion Modeling in a Long Sequence of Scene with Multiple Independently Moving Objects”, *in Pattern Analysis and Machine Intelligence*, vol. 18(1), pp. 58-64, 1996.

[10] M. Isard and A. Blake, ”A Mixed-State Condensation Tracker with Automatic Model-Switching”, *in Int. Conf. on Computer Vision*, pp. 107-112, 1998.

[11] S. Kalitzin, J. J. Staal, B. M. ter Haar Romeny amd M. A. Viergever ”Image Segmentation and Object Recognition by Bayesian Grouping”, *in IEEE Int. Conf. on Image Processing*, vol. 3, pp. 580-583, 2000.

[12] M. Kass, A. Witkin and D. Terzopoulos, ”Snakes: Active Contour Models”, *in International Journal of Computer Vision*, vol. 1(4), pp. 321-331, 1987.

[13] F. Leymarie and M. Levine, ”Tracking Deformable Objects in the Plane Using an Active Contour Model”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15(6), pp. 617-634, 1993.

[14] J. Nascimento and J. S. Marques, ”Robust Shape Tracking in the Presence of Cluttered Background”, *in IEEE Int. Conf. on Image Processing*, vol. 3, pp. 82-85, 2000.

[15] C. Xu and J. Prince, ”Snakes, Shapes, and Gradient Vector Flow”, *in IEEE Transactions on Image Processing*, vol. 7(3), pp. 359-369, 1998.

[16] X. Zhang and H. Burkhardt, ”Grouping Edge Points into Line Segments by Sequencial Hough Transformation”, *in Int. Conf. on Pattern Recognition*, vol 3, pp. 676-679, 2000.