On Classifying Deformable Contours Using the Generalized Active Contour Model

Kok F. Lai *†

†Electrical & Computer Engineering
University of Wisconsin
Madison WI 53706

Roland T. Chin †§

§Computer Science
Hong Kong University of Science & Technology
Clear Water Bay, Hong Kong

Abstract

Recently, we proposed the generalized active contour model (g-snake) to model and extract deformable contours from noisy images. This paper demonstrates the usefulness of g-snake in classifying among several candidate deformable contours. The g-snake is suitable for this task because its shape representation is unique, affine invariant and possesses metric properties. We derive the optimal classification test and show that this requires marginalization of the distribution. However, as the summation is peaked around the posterior estimate in most practical applications, only small regions need to be considered. Finally, we performed extensive experimentations and report significant improvement over matched template in handwritten numeral recognition.

1 Introduction

Recently, we proposed the *generalized active contour model* (g-snake) to simultaneously model and extract *deformable* contours from noisy images [1]. Unlike other deformable templates which consider only global [2, 3, 4] or local [5, 6] deformations, the g-snake is capable of representing any arbitrary shape, accounts for global changes due to rigid motions, and retains ability for local control. Moreover, incorporating an image noise model, the g-snake casts boundary extraction into Bayesian posterior estimation and thus exhibits high level of robustness and adaptability.

This paper demonstrates the usefulness of g-snake in classifying among several candidate deformable contours directly from noisy images. Although there already exist numerous methods [7, 8] that classify deformable contours, these typically assume that the contour has been previously extracted. However, without a contour model, contour extraction has been shown to be an ill-posed problem [9]. Thus the practicality of these methods is limited in applications involving noisy images.

If one assumes that the images contain *rigid* contours in i.i.d. Gaussian noise, then a *matched template* in the form of generalized Hough transforms [10] is optimum in terms of probability of error. However, this method is prone to generating false peaks in complex scenes [11], and we have shown [1] that its performance degrades rapidly with deformation. There have been relative few methods that classify deformable contours directly from noisy images, and these often rely on intuitive ideas [6] or heuristics [12].

We present here a formal strategy to classify deformable contours from noisy images based on g-snake. We will first present the contour and image models in g-snakes and then show that the optimal classification test requires marginalization of these distributions. However, in most practical applications, the summation is peaked around the posterior estimate and thus only small regions need to be considered. Finally, we perform extensive experimentations and report significant improvement over matched template in handwritten numeral recognition.

2 Contour and Image Models

A contour is a vector containing an ordered set of points, $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$. Each $\mathbf{v}_i$ is defined on the finite grid: $\mathbf{v} \in \mathbb{E} = \{(x, y) : x, y = 1, 2, \ldots, M\}$, thus $\mathbf{V} \in \mathbb{E}^n$. Denote $\mathbf{U}$ where each $\mathbf{u}_i = \mathbf{v}_i - \mathbf{g}$ represents the displacement from an arbitrary reference point $\mathbf{g}$. We can encode the shape of an arbitrary contour in the following *shape equation*:

$$\mathbf{A}\mathbf{U}^T = \mathbf{0} \tag{1}$$

where the $n \times n$ *shape matrix* is given by:

$$\mathbf{A} = \begin{bmatrix} 1 & -\beta_1 & -\alpha_1 & 0 & \ldots & 0 \\ -\alpha_2 & 1 & -\beta_2 & 0 & 0 & \ldots \\ 0 & -\alpha_3 & 1 & -\beta_3 & 0 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & -\alpha_{n-1} & 1 & -\beta_{n-1} \\ 0 & 0 & \ldots & -\beta_n & -\alpha_n & 1 \end{bmatrix}$$

*This research was supported in part by the National Computer Board, Republic of Singapore.

The shape matrix is *invariant* and *unique* under affine transformation. In other words, for a given $\mathbf{A}$, if $\bar{\mathbf{U}}$ satisfies $\mathbf{A}\bar{\mathbf{U}}^T = \mathbf{0}$, then

$$\mathbf{A}\mathbf{U}^T = \mathbf{0} \Leftrightarrow \mathbf{U} = \mathbf{T}\bar{\mathbf{U}} \tag{2}$$

where $\mathbf{T}$ is a 2x2 invertible transformation matrix. This implies that the shape formulation is *implicitly* invariant to effects of *rigid motions* such as scaling, rotation, stretching and dilation. Uniqueness is essential for recognition; it implies that two contours that cannot be related by an affine transformation cannot have similar shape matrix.

2.1 Modeling Deformation

Suppose the family of contours $\mathbf{U} \in \Omega \subseteq \mathbb{E}^n$ are similar but exhibit small shape irregularities. To represent these random fluctuations in small localities, we define an internal energy induced by $\mathbf{A}$:

$$E_{int}(\mathbf{U}) = \frac{(\mathbf{A}\mathbf{U}^T)^T\mathbf{R}^{-1}(\mathbf{A}\mathbf{U}^T)}{l(\mathbf{U})} \tag{3}$$

where $l(\mathbf{U}) = \sum_{i=1}^n \|\mathbf{u}_{i+1} - \mathbf{u}_i\|^2$ is a normalizing constant, and $\mathbf{R} = \mathrm{diag}\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2\}$ contains the deformation variances σ_i^2 that allows assignment of location weightings on deformations. We can decompose $E_{int}(\mathbf{U})$ into its n constituents:

$$E_{int}(\mathbf{U}) = \sum_{i=1}^n \frac{E_{int}(\mathbf{u}_i)}{\sigma_i^2} \tag{4}$$

where $E_{int}(\mathbf{u}_i) = \frac{1}{l(\mathrm{U})}\|\mathbf{u}_i - \alpha_i\mathbf{u}_{i_\alpha} - \beta_i\mathbf{u}_{i_\beta}\|^2$ and the basis indices are given by:

$$i_\alpha = \begin{cases} i-1; & i > 1 \\ 3; & i = 1 \end{cases} \qquad i_\beta = \begin{cases} i+1; & i < n \\ n-2; & i = n \end{cases} \tag{5}$$

The internal energy E_{int} possesses *metric properties* which corresponds to the intuitive notion of shape resemblence. Most significantly, it satisfies the triangular inequalility, i.e., ,

$$\begin{aligned} &|E_{int}(\mathbf{U}_p) - E_{int}(\mathbf{U}_q)| \leq \\ &\quad |E_{int}(\mathbf{U}_p) - E_{int}(\mathbf{U}_r)| + |E_{int}(\mathbf{U}_r) - E_{int}(\mathbf{U}_q)| \end{aligned} \tag{6}$$

which ensures that two contours that are highly dissimilar cannot both be similar to some third contour.

We can define a Gibbs measure to $\mathbf{U}$ using the induced energy function:

$$p(\mathbf{U}) = \frac{1}{Z}\exp(-E_{int}(\mathbf{U})) \tag{7}$$

By equivalence, the Gibbs measure also defines a Markov random field where the conditional probability of $\mathbf{u}_i$ is completely specified by its two basis points.

2.2 Image Model

We define an *image* as the vector function $\mathbf{F} : \mathbb{E} \to \mathbb{R}$. We also define the *template* of a contour as the image $\mathbf{F} = \mathrm{B}_{\mathrm{U,g}}$:

$$\mathrm{B}_{\mathrm{U,g}}(\mathbf{r}) = \sum_{i=1}^n \delta(\mathbf{r} - \mathbf{u}_i - \mathbf{g}) \tag{8}$$

where $\mathbf{r} = (x, y) \in \mathbb{E}$ and δ is the delta function. In other words, a template $\mathrm{B}_{\mathrm{U,g}}(\mathbf{r})$ is a special image with values equal to 1 if $\mathbf{r} = \mathbf{u}_i + \mathbf{g}$ but zero otherwise.

A *noisy* image $\mathbf{F} = \mathbf{f}$ containing a contour can be modeled as follows:

$$\mathbf{f}(\mathbf{r})|\mathbf{U}, \mathbf{g} = \mathrm{B}_{\mathrm{U,g}}(\mathbf{r}) + \eta(\mathbf{r}) \tag{9}$$

where $\eta(\mathbf{r})$: $N(0, \sigma_\eta^2)$ is Gaussian. Consequently, $p(\mathbf{f}|\mathbf{U}, \mathbf{g})$ and can be simplified to yield

$$p(\mathbf{f}|\mathbf{U}, \mathbf{g}) \propto \exp\left\{-\frac{1}{\sigma_\eta^2}\sum_{i=1}^n E_{ext}(\mathbf{u}_i, \mathbf{g})\right\} \tag{10}$$

where the external energy $E_{ext}(\mathbf{u}_i, \mathbf{g}) = 1 - \mathbf{f}(\mathbf{u}_i + \mathbf{g})$.

2.3 Generalized Active Contour Models

The problem of extracting a contour with *unknown* deformation from a noisy image is equivalent to maximum *a posteriori* estimation under the Bayesian framework, where

$$\{\mathbf{U}_{map}, \mathbf{g}_{map}\} = \arg\max_{\mathrm{U,g}} p(\mathbf{U})p(\mathbf{f}|\mathbf{U}, \mathbf{g}) \tag{11}$$

noting that $p(\mathbf{U}, \mathbf{g}) = p(\mathbf{U})$. Solving for the solution and ignoring constants, we have

$$\begin{aligned} &\{\mathbf{U}_{map}, \mathbf{g}_{map}\} \\ &\quad = \arg\min_{\mathrm{U,g}} \sum_{i=1}^n \left\{\frac{\lambda_i}{1-\lambda_i}E_{int}(\mathbf{u}_i) + E_{ext}(\mathbf{u}_i, \mathbf{g})\right\} \end{aligned}$$

where

$$\lambda_i = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_i^2} \in [0, 1] \tag{12}$$

are the local *regularization* parameters which control the amount of local template deformation.

We can exploit the local characteristics of the Markov random field to yield a minimization algorithm similar to [14] with complexity $O(nm^3)$, where m is the size of search region. Using line search strategy that concentrate initial search along the normal vectors as in Figure 1, we can further reduce the complexity to $O(nm^3/l^2)$, where l is the segment length in stratification. Furthermore, we have shown in [1] that the g-snakes can be efficiently and reliably initialized using the generalized Hough transform [10].

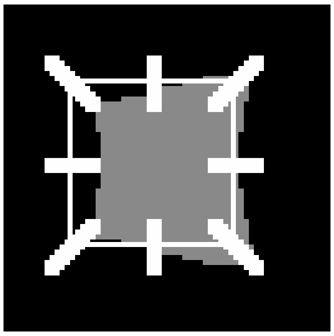


Figure 1: Regions of interest in line search. The white square is the initial contour, and the 8 normal vectors are the initial search directions.

3 Classifying Deformable Contours

Suppose there are two candidate models ω_0 and ω_1, which generate images $\mathbf{F} = \mathbf{f}$ that are distributed according to the density function $p_0(\mathbf{f})$ and $p_1(\mathbf{f})$ respectively. We identify these models with hypothesis H_0 and H_1 as follows:

$$\begin{array}{lll} H_0: & \mathbf{f}: & p_0(\mathbf{f}) \\ H_1: & \mathbf{f}: & p_1(\mathbf{f}) \end{array} \tag{13}$$

Assuming that the only difference between ω_1 and ω_0 is the contour distribution, and that $\mathbf{f}^T\mathbf{f}$ is constant, we can marginalize the distribution to yield

$$p_i(\mathbf{f}) = \sum_{\mathrm{U}\in\Omega_i} p_i(\mathbf{U})p(\mathbf{f}|\mathbf{U}) \tag{14}$$

If ω_i is rigid, i.e., $\Omega_i = \{\bar{\mathbf{U}}_i\}$, then (14) reduces to a simple form:

$$p_i(\mathbf{f}) = p(\mathbf{f}|\bar{\mathbf{U}}_i) \tag{15}$$

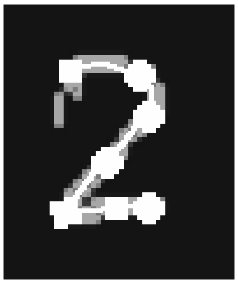


Figure 2: Regions that contribute significantly to summation. For clarity, some points (and regions) along the contour have been omitted.

Define the likelihood ratio $L_{ij}(\mathbf{f}) = p_i(\mathbf{f})/p_j(\mathbf{f})$. Using the Neyman-Pearson lemma, the following test will yield the smallest classification error:

$$\text{select} \left\{ \begin{array}{ll} H_1: & \text{if } L_{10}(\mathbf{f}) > k = \frac{p_0 C_{01}}{p_1 C_{10}} \\ H_0: & \qquad \text{otherwise} \end{array} \right. \tag{16}$$

where p_i are the prior probabilities of observing model ω_i, and C_{ij} is the *cost* associated with deciding H_j given that H_i is in force. Without loss of generality, we shall use $k = 1$ in subsequent discussions.

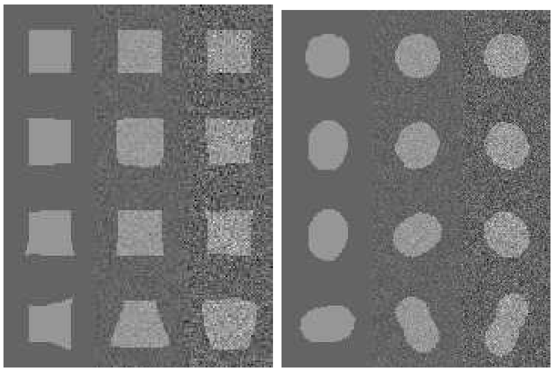
A classification test that involves one or more deformable models will require marginalization. We select H_1 if

$$\sum_{\mathrm{U}\in\Omega_1} p_1(\mathbf{U})p(\mathbf{f}|\mathbf{U}) > \sum_{\mathrm{U}\in\Omega_0} p_0(\mathbf{U})p(\mathbf{f}|\mathbf{U}) \tag{17}$$

The classification test can be easily generalized to the case of multiple classes $\omega_1, \omega_2, \ldots, \omega_q$. In this case, we select ω_γ if

$$\gamma = \arg \max_{1\le i\le q} \sum_{\mathrm{U}\in\Omega_i} p_i(\mathbf{U})p(\mathbf{f}|\mathbf{U}) \tag{18}$$

Again, we can exploit the local characteristics of the Markov random field to design a summation algorithm with complexity $O(nm^3)$. Moreover, in most practical applications, the summation in (14) is peaked around the posterior estimate $\mathbf{U}_{map}$ as shown in Figure 2. Consequently, we can significantly reduce the computation time by applying the summation algorithm in small rectangular regions (e.g. 3×3 windows) centered around $\mathbf{U}_{map}$. Hence the solution of g-snake can be rapidly post-processed for classification.



(a) rectangles (b) ellipsoids

Figure 3: Synthesized images for quantitative comparison. Deformation increases from top to bottom while image noise increases from left to right.

4 Experiments

4.1 Quantitative Comparisons

This section compares the classification performance of the proposed deformable template to rigid (matched) template using synthesized images containing deformed rectangles and ellipsoids as shown in Figure 3. The degree of deformation is controlled by deformation variance σ_ν in the generation process, while image noise is introduced using i.i.d. Gaussian noise with variance σ_n.

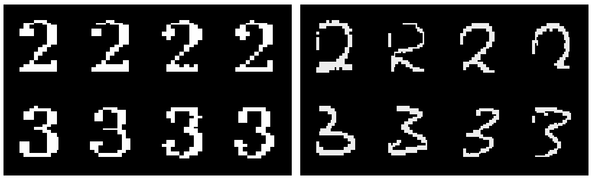
The rigid templates (generalized Hough transform) consist of 16 equally spaced points obtained from perfect rectangles and circles ($\sigma_\nu = \sigma_n = 0$). For each pair of σ_ν and σ_n, we generate training samples and use them to estimate the regularization parameters for deformable templates. Figure 4 show representative results for these experiments.

Figure 4(a) shows that when considerable deformation is present, deformable templates provide the necessary adaptation to yield smaller classification error. On the other hand, Figure 4(b) examines the effect of image noise and illustrates that deformable templates produce smaller error for all levels of image noise.

In summary, these rigorous experimentations have verified the superiority of deformable templates over rigid templates in the presence of deformation.

4.2 Numeral Recognition

This section compares the classification performance of deformable template versus rigid template, using printed and handwritten numerals (2 & 3) as shown in Figure 5. Table 1 shows the classification results using 190 character samples each.



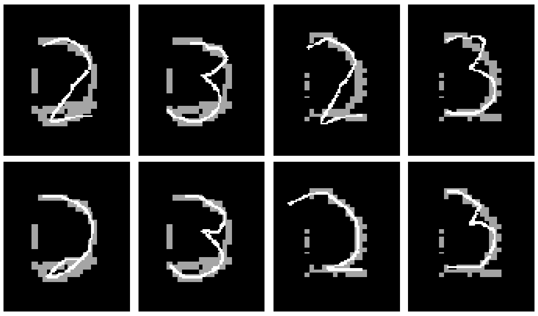
(a) printed (b) handwritten

Figure 5: Printed and handwritten numerals

	rigid	deformable
printed	0.00	0.00
handwritten	0.22	0.07

Table 1: Probability of error, P_e, for printed and handwritten numerals (2 & 3) classification using rigid and deformable templates.

As there are minimal deformation in printed numerals, rigid template is optimum and thus yields very good



(a) (b)

Figure 6: Illustrations of rigid and deformable templates. In both cases the rigid templates (top row) misclassified the characters, while the deformable templates (bottom row) produced correct classification.

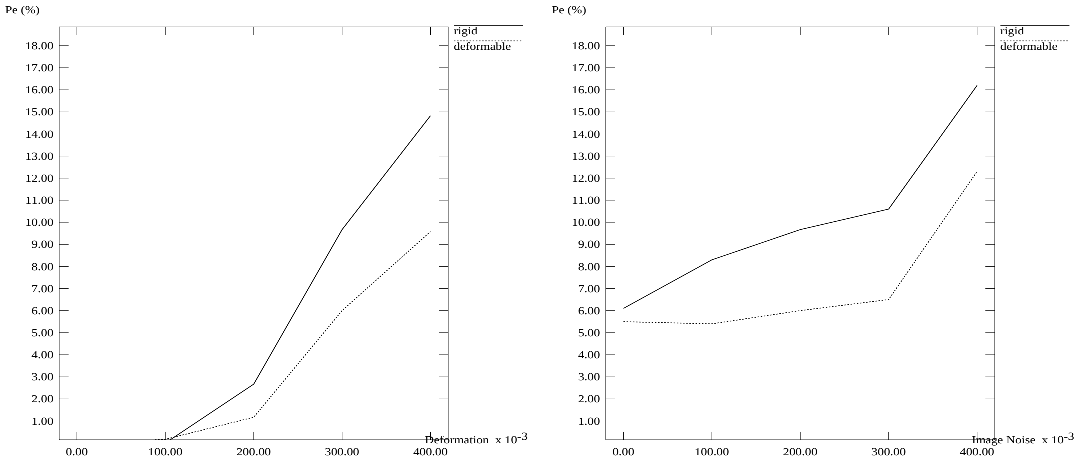
classification results. In this case it is unnecessary to employ deformable templates because very little improvement can be achieved.

Handwritten numerals, however, exhibit considerable deformation due to variation in writing style. Consequently, the rigid templates performed poorly and misclassified 22% of the samples presented. In contrast, deformable templates were successful in reducing the misclassification rate to 7%.

Figure 6 demonstrates how deformable template uses both the goodness of match and the degree of deformation to achieve correct classification. The top row of Figure 6 shows the best match provided by rigid templates, while the bottom row shows the posterior estimates of deformable templates. In both cases the rigid template corresponds to numeral "3" yielded higher correlation when compared to that of "2", and thus misclassified the character as "3". In contrast, the deformable templates in Figure 6(a) underwent the necessary deformation to yield higher correlation for "2", and thus correctly classified the character. In Figure 6(b), although the deformable templates yielded higher correlation for "3", the template corresponds to "3" underwent more severe deformation in order to achieve the match. Consequently, the deformable templates correctly classified the character as "2" using information derived from prior distributions.

5 Conclusions

We have presented a formal strategy for classifying deformable contours directly from noisy images. The formulation follows naturally from the generalized ac-



(a) $\sigma_n = 0.20$, varying σ_ν

(b) $\sigma_\nu = 0.30$, varying σ_n

Figure 4: Classification results. (a) Probability of error, P_e versus standard deviation of boundary deformation process, σ_ν, with standard deviation of image noise fixed at $\sigma_n = 0.20$ (b) P_e versus σ_n with $\sigma_\nu = 0.30$.

tive contour model, where the solution can be rapidly post-processed to yield classification results. Significant improvement over matched template has been reported in handwritten numeral recognition.

References

[1] K. F. Lai & R. T. Chin, "Deformable Contours: Modeling and Extraction," *Int. Conf. Comp. Vis. Patt. Recog.*, 1994, pp. 601-608.

[2] A. L. Yuille, D. S. Cohen & P. W. Hallinan, "Feature Extraction from Faces using Deformable Templates," *Int. Conf. Comp. Vis. Patt. Recog.* 1989, pp. 104-109.

[3] L. H. Staib & J. Duncan, "Boundary Finding with Parametrically Deformable Models," *IEEE Trans. Pat. Anal. Mach. Intell.*, vol. **PAMI**-14, 1992, pp. 1061-1075.

[4] J. Subrahmonia, J. D. Karen & D. B. Cooper, "Bayesian Method for the use of Implicit Polynomials and Algebraic Invariants in Practical Computer Vision," *Curves and Surfaces in Computer Vision & Graphics 3*, 1992, pp. 104-117.

[5] M. Kass, A. Witkin & D. Terzopoulos, "Snakes: Active Contour Models," *Procs of First Int. Conf. on Computer Vision*, 1987, pp. 259-269.

[6] G. E. Hinton, C. K. I. Williams & M. D. Revow, "Adaptive Elastic Models for Hand-Printed Character Recognition," *Neural Information Processing Systems*, vol 4, ed. J. E. Moody *et. al.*, 1992, pp. 512-519.

[7] Y. He & A. Kundu, "2-D Shape Classification using Hidden Markov Model," *IEEE Trans. Pat. Anal. Mach. Intell.*, vol. **PAMI**-13, 1991, pp. 1172-1184.

[8] S. Umeyama, "Parameterized Point Pattern Matching and Its Application to Recognition of Object Families," *IEEE Trans. Pat. Anal. Mach. Intell.*, vol. **PAMI**-15, 1993, pp. 136-144.

[9] T. Poggio & V. Torre, "Ill-posed Problems and Regularization Analysis in Early Vision, " *Proceedings of AARPA Image Understanding Workshop*, 1984, pp. 257-263.

[10] D. H. Ballard, "Generalizing the Hough Transform to Detect Arbitrary Shapes," *Pattern Recognition*, vol. 13, 1981, pp. 111-122.

[11] W. L. Grimson & D. P. Huttenlocher, "On the sensitivity of the Hough Transform for Object Recognition," *IEEE Trans. Pat. Anal. Mach. Intell.*, vol. **PAMI**-12, 1990, pp. 255-274.

[12] D. P. Huttenlocher, G. A. Klanderman & W. J. Rucklidge, "Comparing Images Using the Hausdorff Distance," *IEEE Trans. Pat. Anal. Mach. Intell.*, vol. **PAMI**-15, 1993, pp. 850-863.

[13] A. Blake, R. Curwen & A. Zisserman, "Affine-invariant Contour Tracking with Automatic Control of Spatiotemporal Scale", *Proc Int. Conf. on Comp. Vis.*, 1993, pp. 66-75.

[14] A. A. Amini, T.E. Weymouth & R. C. Jain, "Using Dynamic Programming for Solving Variational Problems in Vision," *IEEE Trans. Pat. Anal. Mach. Intell.*, vol. **PAMI**-12, no. 9, 1990, pp. 855-867.