Deformable Contours: Modeling and Extraction

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

This paper considers the problem of modeling and extracting arbitrary deformable contours from noisy images. We propose a global contour model based on a stable and regenerative shape matrix, which is invariant and unique under rigid motions. Combined with Markov random field to model local deformations, this yields prior distribution that exerts influence over a global model while allowing for deformations. We then cast the problem of extraction into posterior estimation and show its equivalence to energy minimization of a generalized active contour model. We discuss pertinent issues in shape training, minimax regularization and initialization by generalized Hough transform. Finally, we present experimental results and compare its performance to rigid template matching.

1 Introduction

Without a contour model, contour extraction from noisy image is an ill-posed problem [1]. Therefore, it is crucial for an extraction method to allow the incorporation of global contour models in its formulation. For generality, the model should be capable of representing any arbitrary shape.

A well known example is the generalized Hough transform [2] which combines modeling and extraction for *rigid* contours. Because GHT works by accumulating large number of votes, it is relatively insensitive to small vote fluctuation caused by noise and occlusion. In fact, if the superimposed noise is Gaussian and white, then GHT is simply an efficient implementation of matched template, which is optimum in terms of detection error.

Rigid templates, however, cannot account for *deformations* which frequently arise from diversity and irregularity of shape. See Figure 1 for examples. Since the degree of deformation is unknown in advance, a rigid template chosen *a priori* cannot produce satisfactory results for all cases. In fact, we will show that its performance degrades with deformation.

Figure 1: Handwritten characters with deformations.

Deformable templates, in contrast, employ weak models which deform in conformation to salient image features. Examples include the active contour model (snake) which uses smoothness constraints to restrict its solutions to controlled continuity splines [3]; parameterized templates for facial feature extraction [4]; elliptic Fourier decomposition for objects with shape irregularities [5]; implicit polynomials for curve and surface modeling [6]; and movable control points for hand-printed character recognition [7].

These methods typically consider only global [4, 5, 6] or local [3, 7] deformations. Global templates involve large structural interactions and contain less parameters to be optimized. However, these global parameters cannot exercise local control along the contour, and their physical meaning are sometimes obscure. In contrast, local models such as snakes contain more parameters and exert local control, but they are ill-suited for incorporation of global contour models.

This paper considers the problem of modeling and extracting arbitrary deformable contours from noisy images. Our method is general and capable of representing any arbitrary shape, accounts for global changes due to rigid motions, and retains ability for local control. The contour model is based on a stable and regenerative shape matrix which is invariant and unique under rigid motions. Combined with the local characteristics of the Markov random field to model local deformations, this yields prior distribution that

[*]This research was supported in part by the National Computer Board, Singapore, and in part by Sino Software Research Center, Hong Kong.

exerts influence over a global model while allowing for deformations. Under this framework, the problem of extracting contours with unknown deformation from noisy images turns into maximum *a posteriori* estimation. This is equivalent to minimizing the energy of a generalized active contour model (g-snake). We discuss pertinent issues in shape training, minimax regularization and initialization by generalized Hough transform, and present some experimental results.

2 Modeling Deformable Contours

We define a contour as the 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 $\mathbf{U} \in \mathbb{E}^n$ where each $\mathbf{u}_i = \mathbf{v}_i - \mathbf{g}$ represents the displacement from an arbitrary reference point $\mathbf{g}$. We say that $\mathbf{U}$ is *nontrivial* if $|x_{i-1}y_{i+1} - x_{i+1}y_{i-1}| > 0$ for $1 \leq i \leq n$. An example of a trivial contour is a straight line. More generally, the nontrivial condition ensures that $\mathbf{u}_{i-1}$ and $\mathbf{u}_{i+1}$ are linearly independent for all i.

Each $\mathbf{u}_i$ can be expressed as a linear combination of two linearly independent vectors. For modeling of highly variable but locally predictable contours, the two neighboring points are suitable as its basis:

$$\mathbf{u}_i = \alpha_i \mathbf{u}_{i_\alpha} + \beta_i \mathbf{u}_{i_\beta} \tag{1}$$

where the basis indices are given by:

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

Collecting and rearranging similar equations for all i, we obtain the *shape equation*

$$\mathbf{A}\mathbf{U}^T = \mathbf{0} \tag{3}$$

where $\mathbf{A}$ is an $n \times n$ *shape matrix* that contains the necessary information to describe a shape:

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

$\mathbf{A}$ is *regenerative*: the complete chain $\mathbf{U}$ can be generated from $\mathbf{A}$ by appropriate inversion if any two points on $\mathbf{U}$ are given. This is possible because $\text{rank}(\mathbf{A}) = n - 2$, and so an $(n-2) \times (n-2)$ invertible submatrix exists within $\mathbf{A}$. On the other hand, if a nontrivial $\mathbf{U}$ is given, the shape coefficients can be computed to generate $\mathbf{A}$.

Furthermore, $\mathbf{A}$ is *stable* for nontrivial contours because α, β and $\mathbf{U}$ are related by linear equations. Hence small changes in the contour imply small perturbations of the shape matrix and vice versa. Stability is extremely crucial when $\mathbf{A}$ must be estimated from sample contours which may be locally deformed.

To adequately represent objects with shape irregularities, one must account for both global and local deformations, which are discussed next.

2.1 Global Deformations

Global deformations correspond to effects of rigid motion such as scaling, rotation, stretching and dilation, as shown in Figure 2(a). These operations can be represented by affine transformations on $\mathbf{V}$. The following theorem states that $\mathbf{A}$ is invariant and unique under such transformations:

Theorem 1 *Two nontrivial contours satisfy the same shape equation if and only if they are related by a linear transformation.*

Proof Let $\mathbf{T}$ denote a 2×2 transformation matrix. First, note that if $\mathbf{A}\bar{\mathbf{U}}^T = \mathbf{0}$ and $\mathbf{U} = \mathbf{T}\bar{\mathbf{U}}$, then $\mathbf{A}\mathbf{U}^T = \mathbf{A}\bar{\mathbf{U}}^T\mathbf{T}^T = \mathbf{0}$.

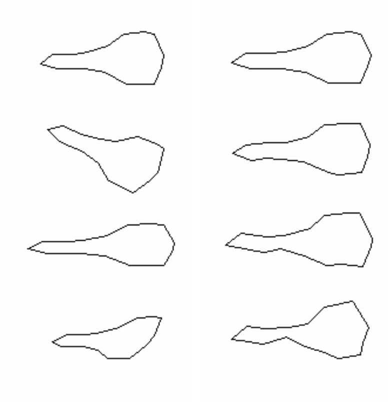
To prove necessity, recall that $\text{rank}(\mathbf{A}) = n - 2$. Equivalently, the dimension of the null space of $\mathbf{A}$ is 2. Write $\bar{\mathbf{U}}^T = [\bar{\mathbf{x}}, \bar{\mathbf{y}}]$, where $\bar{\mathbf{x}}, \bar{\mathbf{y}}$ are nx1 vectors. Because $\bar{\mathbf{U}}$ is nontrivial, $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ are linearly independent and thus span the null space of $\mathbf{A}$.

Now, since $\mathbf{A}\mathbf{U}^T = \mathbf{A}[\mathbf{x}, \mathbf{y}] = \mathbf{0}$, both $\mathbf{x}$ and $\mathbf{y}$ are also in the null space of $\mathbf{A}$. Hence they can be expressed as a linear combination of $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$, i.e.,

$$\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} t_{11}\bar{\mathbf{x}} + t_{12}\bar{\mathbf{y}} \\ t_{21}\bar{\mathbf{x}} + t_{22}\bar{\mathbf{y}} \end{bmatrix} \Rightarrow \mathbf{U} = \mathbf{T}\bar{\mathbf{U}} \tag{4}$$

Finally, note that contour displacement, $\mathbf{g} = \bar{\mathbf{g}} + \mathbf{d}$, has no effect on the shape equation. This completes the proof.

The invariance property implies that the shape coefficients of a contour are unaffected by affine transformations, while the uniqueness property makes recognition possible. For example, rectangles and parallograms have the same $\mathbf{A}_p$, and so are circles and ellipsoids with $\mathbf{A}_e$. However, $\mathbf{A}_p \neq \mathbf{A}_e$, i.e., they are unique, because families of parallograms and ellipsoids cannot be related by an affine transformation.



(a) global (b) local

Figure 2: Examples of global and local deformations. (a) Global deformations described by rotation, stretching and dilation of the first contour; (b) local deformations of successively higher degree.

2.2 Local Deformations

Suppose we are interested in a family of contours $\mathbf{U} \in \Omega \subseteq \mathbb{E}^n$ which are similar but exhibit small shape irregularities, as shown in Figure 2(b). To represent these random fluctuations in small localities, we define an internal energy which is induced by the shape matrix $\mathbf{A}$:

$$E_{int}(\mathbf{U}) = \frac{(\mathbf{A}\mathbf{U}^T)^T \mathbf{R}^{-1} (\mathbf{A}\mathbf{U}^T)}{l(\mathbf{U})} \tag{5}$$

where $l(\mathbf{U}) = \frac{1}{n}\sum_{i=1}^{n} \|\mathbf{u}_{i+1} - \mathbf{u}_i\|^2$ 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 dependent weightings on deformations.

Now, we may assign probabilities to $\mathbf{U}$ as follows :

$$p(\mathbf{U}) = \frac{1}{Z}\exp(-E_{int}(\mathbf{U})) \tag{6}$$

where $Z = \sum_{\mathbf{U}\in\Omega} \exp(-E_{int}(\mathbf{U}))$ is a normalizing constant. A probability measure of the form in (6) is called a *Gibbs measure* [8]. By equivalence, it also defines a *Markov random field*, i.e.,

$$p(\mathbf{u}_i|\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n) = p(\mathbf{u}_i|\mathbf{u}_{i_\alpha}, \mathbf{u}_{i_\beta}) \tag{7}$$

where the conditional probability of $\mathbf{u}_i$ given the entire chain $\mathbf{U}$ is completely specified by the conditional probability of $\mathbf{u}_i$ given its two basis points.

The following theorem on E_{int} implies that $p(\mathbf{U})$ is scale and rotation invariant:

Theorem 2 *The internal energy of a contour is scale and rotation invariant.*

Proof Let

$$\mathbf{T}(z) = \begin{bmatrix} z & 0 \\ 0 & z \end{bmatrix} \text{ and } \mathbf{T}(\theta) = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$$

represents scale change and rotation respectively. Now

$$\mathbf{u}^T\mathbf{T}^T(z)\mathbf{T}(z)\mathbf{u} = z^2\mathbf{u}^T\mathbf{u} \tag{8}$$

$$\mathbf{u}^T\mathbf{T}^T(\theta)\mathbf{T}(\theta)\mathbf{u} = \mathbf{u}^T\mathbf{u} \tag{9}$$

Applying (8) and (9) to numerator and denominator of E_{int} yields the appropriate results. This completes the proof.

In summary, we have combined a stable, invariant and unique contour model with the local characteristic of Markov random field to yield prior distributions for any arbitrary contour. We proceed next to present an image model which allows us to cast contour extraction from noisy images into posterior estimation.

3 Contour Extraction

3.1 Rigid Template

We define image as the vector function $\mathbf{F} : \mathbb{E} \rightarrow \mathbb{D}$. Depending on the data type, either $\mathbb{D} = \mathbb{R}$ (intensity image or edge magnitude) or $\mathbb{D} = \mathbb{R}^2$ (2×1 intensity gradient vector).

A *template* of a contour is the image $\mathbf{F} = \mathrm{B}_{\mathrm{U,g}}$:

$$\mathrm{B}_{\mathrm{U,g}}(\mathbf{r}) = \sum_{i=1}^{n} \mathbf{h}_i \delta(\mathbf{r} - \mathbf{u}_i - \mathbf{g}) \tag{10}$$

where $\mathbf{r} = (x, y) \in \mathbb{E}$, δ is the delta function, $\mathbf{h}_i \in \mathbb{D}$ and $|\mathbf{h}_i| = 1$. In other words, a template $\mathrm{B}_{\mathrm{U,g}}(\mathbf{r})$ is a special image with values equal to $\mathbf{h}_i$ if $\mathbf{r} = \mathbf{u}_i + \mathbf{g}$ but zero otherwise. For $\mathbb{D} = \mathbb{R}^2$, $\mathbf{h}_i$ can be adaptively determined by rotating the tangent direction of the contour by 90°, i.e., let $\mathbf{t}_i$ represents the *tangent* vector at $\mathbf{u}_i$:

$$\mathbf{t}_i = \frac{\mathbf{u}_i - \mathbf{u}_{i-1}}{\|\mathbf{u}_i - \mathbf{u}_{i-1}\|} + \frac{\mathbf{u}_{i+1} - \mathbf{u}_i}{\|\mathbf{u}_{i+1} - \mathbf{u}_i\|} \tag{11}$$

then $\mathbf{h}_i$ is a unit vector which is normal to $\mathbf{t}_i$.

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

$$\mathbf{f}(\mathbf{r})|\mathbf{U}, \mathbf{g} = \mathrm{B}_{\mathrm{U,g}}(\mathbf{r}) + \eta(\mathbf{r}) \tag{12}$$

where $\eta(\mathbf{r}) : N(0, \sigma_\eta^2 \mathbf{I})$ is Gaussian, $\mathbf{I} = 1$ if $\mathbb{D} = \mathbb{R}$; otherwise $\mathbf{I}$ is a 2×2 identity matrix. Consequently, $p(\mathbf{f}|\mathbf{U}, \mathbf{g})$ is Gaussian distributed and can be simplified to yield:

$$p(\mathbf{f}|\mathbf{U}, \mathbf{g}) = \frac{1}{C} \exp \left\{ \frac{-\mathbf{f}^T\mathbf{f} - n + 2\sum_{i=1}^n \mathbf{h}_i^T \mathbf{f}(\mathbf{u}_i + \mathbf{g})}{2\sigma_\eta^2} \right\} \tag{13}$$

where $C = (2\pi\sigma_\eta^2)^{M^2/2}$.

The displacement of a rigid contour can be estimated from the noisy image using the principle of maximum likelihood:

$$\bar{\mathbf{g}} = \arg\max_{\mathbf{g}} \sum_{i=1}^n \mathbf{h}_i^T \mathbf{f}(\mathbf{u}_i + \mathbf{g}) \tag{14}$$

In other words, the estimated reference point is the location that best correlates with the image. This forms the basis of *matched template* operation.

In the same token, contours that may be deterministically deformed have the following ML estimates:

$$\{\bar{\mathbf{g}}, \bar{\mathbf{T}}\} = \arg\max_{\mathbf{g}} \max_{\mathbf{T}} \sum_{i=1}^n \mathbf{h}_i^T \mathbf{f}(\mathbf{T}\mathbf{u}_i + \mathbf{g}) \tag{15}$$

where the estimated contour $\bar{\mathbf{U}} = \bar{\mathbf{T}}\mathbf{U}$. The correlation can be efficiently implemented using generalized Hough transforms [2]. Note that the computation cost required in (15) depends greatly on the range and resolution of $\mathbf{T}$. For example, if one expects the contours to be scaled and rotated, then the range of $\mathbf{T}$ should span the two operations. In many instances the range can be restricted through learning or prior knowledge. Otherwise $\mathbf{T}$ will cover all the transformations, thus increasing the cost of searching.

3.2 Deformable Template

We assert that any rigid template $\bar{\mathbf{U}}$ will perform poorly if the contours are deformable:

Assertion 1 *The expected correlation of a matched template decreases with deformation.*

Proof Let $\mathbf{E}[\cdots]$ denotes mathematical expectation. Observe that

$$\begin{aligned} \mathbf{E}[\mathrm{B}_{\bar{\mathrm{U}},\mathrm{g}}^T \mathbf{f}] &= \mathbf{E}[\mathbf{E}[\mathrm{B}_{\bar{\mathrm{U}},\mathrm{g}}^T \mathbf{f} | \mathbf{U}, \mathbf{g}]] \\ &= \mathbf{E}[\mathrm{B}_{\bar{\mathrm{U}},\mathrm{g}}^T \mathrm{B}_{\mathrm{U},\mathrm{g}}] \\ &= \mathbf{E}[\sum_{i=1}^n \sum_{j=1}^n \delta(\bar{\mathbf{u}}_i - \mathbf{u}_j)] \end{aligned}$$

For σ_i sufficiently small, $p(\bar{\mathbf{u}}_i = \mathbf{u}_j) \approx 0$ if $i \neq j$. Thus

$$\mathbf{E}[\mathrm{B}_{\bar{\mathrm{U}},\mathrm{g}}^T \mathbf{f}] \approx \mathbf{E}[\sum_{i=1}^n \delta(\bar{\mathbf{u}}_i - \mathbf{u}_i)] = p(\bar{\mathbf{U}})$$

Now, since $E_{int}(\bar{\mathbf{U}}) = 0$, from (6), we have

$$p(\bar{\mathbf{U}}) = 1/Z = \sum_{\mathrm{U} \in \Omega} \exp(E_{int}(\mathbf{U}))$$

which decreases if σ_i increases. This completes the proof.

The problem at hand is equivalent to extracting a contour with *unknown* deformation from a noisy image. Using Bayesian framework, this turns into maximum *a posteriori* (MAP) estimation. The estimates $\mathbf{U}_{map}$ and $\mathbf{g}_{map}$ are given by:

$$\begin{aligned} \{\mathbf{U}_{map}, \mathbf{g}_{map}\} &= \arg\max_{\mathrm{U},\mathrm{g}} p(\mathbf{U}, \mathbf{g}|\mathbf{f}) \\ &= \arg\max_{\mathrm{U},\mathrm{g}} p(\mathbf{U}) p(\mathbf{f}|\mathbf{U}, \mathbf{g}) \end{aligned} \tag{16}$$

noting that $p(\mathbf{U}, \mathbf{g}) = p(\mathbf{U})$. While this prior distribution creates a bias to a global model, it is revised into posterior distribution after observing the image. In the next section, we will show how one may formulate and solve these estimation problems using the generalized active contour models.

4 Generalized Active Contour Models

4.1 Formulation

Denote the constituent *internal* and *external* energy as follows:

$$E_{int}(\mathbf{u}_i) = \frac{\|\mathbf{u}_i - \alpha_i \mathbf{u}_{i_\alpha} - \beta_i \mathbf{u}_{i_\beta}\|^2}{l(\mathbf{U})} \tag{17}$$

$$E_{ext}(\mathbf{u}_i, \mathbf{g}) = 1 - \mathbf{h}_i^T \mathbf{f}(\mathbf{u}_i + \mathbf{g}) \tag{18}$$

We will substitute these equations for $p(\mathbf{U})$ in (6) and $p(\mathbf{f}|\mathbf{U}, \mathbf{g})$ in (14). Solving for MAP estimation and ignoring constants, we have

$$\begin{aligned} &\{\mathbf{U}_{map}, \mathbf{g}_{map}\} \\ &= \arg\min_{\mathrm{U},\mathrm{g}} \sum_{i=1}^n \left\{ \frac{E_{int}(\mathbf{u}_i)}{\sigma_i^2} + \frac{E_{ext}(\mathbf{u}_i, \mathbf{g})}{\sigma_\eta^2} \right\} \\ &= \arg\min_{\mathrm{U},\mathrm{g}} \sum_{i=1}^n \left\{ \frac{\lambda_i}{1 - \lambda_i} E_{int}(\mathbf{u}_i) + E_{ext}(\mathbf{u}_i, \mathbf{g}) \right\} \end{aligned}$$

where

$$\lambda_i = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_i^2} \in [0, 1] \tag{19}$$

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

The formulation in (19) is analogous to the *active contour models* (snakes) [3]. However, the original E_{int} only constrains the solution to the class of controlled continuity splines. Our formulation generalizes E_{int}, allowing for incorporation of prior models to create bias towards a particular type of contour. We call the resulting model *generalized active contour models*, or g-snakes.

From (19), the task of finding $\mathbf{U}_{map}$ and $\mathbf{g}_{map}$ turns into an energy minimization problem, where several algorithms [3, 9, 10, 11] have been proposed.

4.2 Regularization

The regularization parameters λ_i in (19) are derived from the variances σ_η and σ_i. In some instances these parameters may be learned from training samples. In cases where learning is impossible or unreliable, the minimax principle can be used.

The *local minimax criterion* [9] determines the optimal regularization by minimizing the worst case energy. We subject σ_η and σ_i to the constraint

$$\frac{1}{\sigma_i^2} + \frac{1}{\sigma_\eta^2} = 1 \tag{20}$$

and seek a solution for

$$\begin{aligned} \{\mathbf{U}_{lmc}, \mathbf{g}_{lmc}\} & \\ &= \arg\min_{\mathrm{U,g}} \sum_{i=1}^{n} \max_{\sigma_\eta^2, \sigma_i^2} \left\{ \frac{E_{int}(\mathbf{u}_i)}{\sigma_i^2} + \frac{E_{ext}(\mathbf{u}_i, \mathbf{g})}{\sigma_\eta^2} \right\} \\ &= \arg\min_{\mathrm{U,g}} \sum_{i=1}^{n} \max \{E_{int}(\mathbf{u}_i), E_{ext}(\mathbf{u}_i, \mathbf{g})\} \end{aligned} \tag{21}$$

The strategy solves the problem without requiring explicit computation of the regularization parameters. It has been shown to yield robust performance for various applications (see [9]).

4.3 Initialization

The objective of initialization is to place the g-snake in the proximity of the minima in (19) or (21) so as to facilitate speedy convergence in minimization. In other words, we want to pick initial $\{\bar{\mathbf{U}}, \bar{\mathbf{g}}\}$ that will, on average, converge to $\mathbf{U}_{map}$ or $\mathbf{U}_{lmc}$ more rapidly then all other choices.

Such criterion is difficult to specify in precise mathematical terms. However, we can make a reasonable assumption that a g-snake of lower energy is located closer to the minima compared to one of higher energy. Then the objective becomes that of finding a initial contour that will, on average, begin with lower energy than all other choices. Let

$$e(\mathbf{U}, \Lambda) = \sum_{i=1}^{n} \frac{\lambda_i}{1 - \lambda_i} E_{int}(\mathbf{u}_i) + E_{ext}(\mathbf{u}_i) \tag{22}$$

represents the total energy. We assert that

Assertion 2 *A contour $\bar{\mathbf{U}}$ that satisfies $E_{int}(\bar{\mathbf{U}}) = 0$ will, on average, begin with the lowest total energy. Moreover, this contour can be initialized through generalized Hough transform.*

Proof Since $E_{int}(\bar{\mathbf{U}}) = 0$ and $E_{int}(\mathbf{U}) \geq 0$, it immediately follows that $\mathbf{E}[E_{int}(\bar{\mathbf{U}})] \leq \mathbf{E}[E_{int}(\mathbf{U})]$. Also,

$$\begin{aligned} \mathbf{E}[E_{ext}(\mathbf{U})] &= 1 - \mathbf{E}[\mathrm{B}_{\mathrm{U,g}}^T \mathbf{f}] \\ &= 1 - p(\mathbf{U}) \\ &= 1 - \frac{1}{Z} \exp(-E_{int}(\mathbf{U})) \end{aligned}$$

$E_{int}(\bar{\mathbf{U}}) = 0$ implies that $p(\bar{\mathbf{U}}) \geq p(\mathbf{U})$ and hence $E[E_{ext}(\bar{\mathbf{U}})] \leq E[E_{ext}(\mathbf{U})]$. Combining the results for internal and external energy, we conclude that

$$\mathbf{E}[e(\bar{\mathbf{U}}, \Lambda)] \leq \mathbf{E}[e(\mathbf{U}, \Lambda)]$$

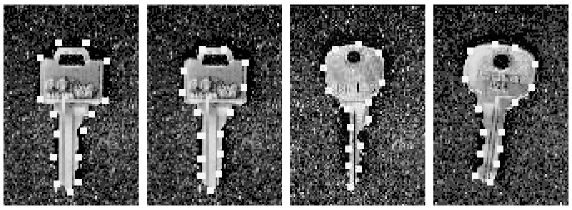
In other words, $\bar{\mathbf{U}}$ has the lowest total energy on average. One should therefore initialize the g-snake using a reference contour $\bar{\mathbf{U}}$ generated from tbe shape matrix, i.e., one can initialize the g-snake using generalized Hough transform.

The next section demonstrates by experiments how one may estimate the shape matrix from training samples, and subsequently apply the trained g-snakes in various applications. We also compare the matching performance of g-snakes and rigid templates.

5 Applications and Experiments

5.1 Training

Figure 3 shows the sequence of steps taken to train the shape matrix for key localization. In these examples the 2×1 intensity gradient vectors are computed by fitting planes in 2×2 windows using the method of least squares.



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

Figure 3: G-snake training. (a) Selected feature points; (b)(c)(d) $\hat{\mathbf{U}}_j$ used to generate $\hat{\mathbf{A}}_j$ for $j = 1, 2$ and 3. These $\hat{\mathbf{A}}_j$ are averaged for $\hat{\mathbf{A}}$ estimation.

Figure 3(a) shows the feature points that form the chain $\mathbf{U}$. The set generally include locations of high curvature which can be selected either manually or automatically through dominant point detection [12]. An initial estimate of the shape matrix are computed from $\mathbf{U}$. Using this shape matrix and minimax regularization, the total energy of the g-snake is minimized to yield $\hat{\mathbf{U}}_1$ as in Figure 3(b). The shape matrix is then updated to $\hat{\mathbf{A}}_1$.

We repeat the procedure for many key samples to obtain $\hat{\mathbf{A}}_1$, $\hat{\mathbf{A}}_2, \ldots,$ $\hat{\mathbf{A}}_m$. Using the principle of maximum likelihood, the estimated shape matrix is given as follows:

$$\hat{\mathbf{A}} = \frac{1}{m}\sum_{j=1}^{m}\hat{\mathbf{A}}_j \tag{23}$$

Similar method is employed to estimate the two reference points used to generate $\bar{\mathbf{U}}$ from $\hat{\mathbf{A}}$.

We may also estimate the local regularization parameters λ_i by learning the deformation and noise variance from the training samples. As insufficient training samples are available for this experiment, minimax regularization is used.

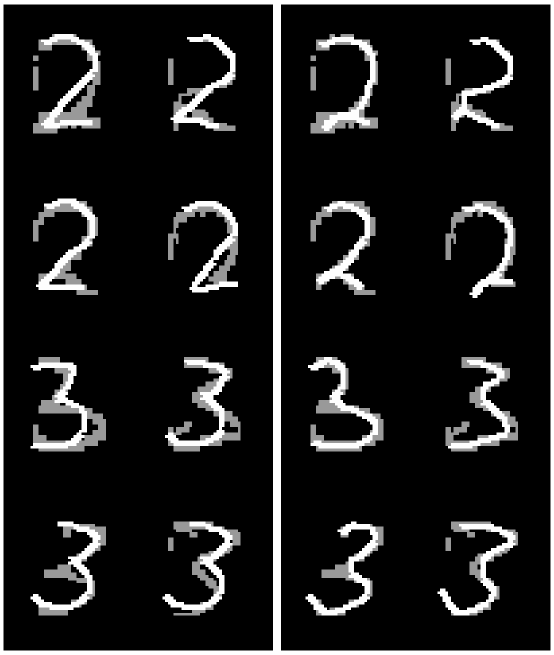
5.2 Initialization and Extraction

We use the generalized Hough transform to initialize the g-snakes, considering only rotation for $\mathbf{T}$ in equation (15). From the initialized contours, we minimize the energy on the g-snake and show the results in Figure 4. Beginning from the top left corner, each example shows the original image, edge magnitude map, final and initial boundaries in the clockwise order.

As evidenced in the edge maps, these scenes contain considerable clutter and occlusion. As a consequence, traditional methods based on edge linking are likely to produce poor results. It can be seen that the generalized Hough transform successfully approximate the locations of the keys, but cannot produce precise boundaries. The g-snakes refined these initial contours to yield accurate reproduction of the boundaries.

5.3 Handwritten Characters

Figure 5 shows applications of g-snake in extracting the medial axis of handwritten characters, using intensity image as the data type. Similar training steps are used to estimate two shape matrices for the two different characters. The results show that generalized Hough transform successfully estimated the necessary rotation, stretching and dilation to yield good initialization. After undergoing necessary deformations, the g-snake produces medial axes that match the underlying characters.

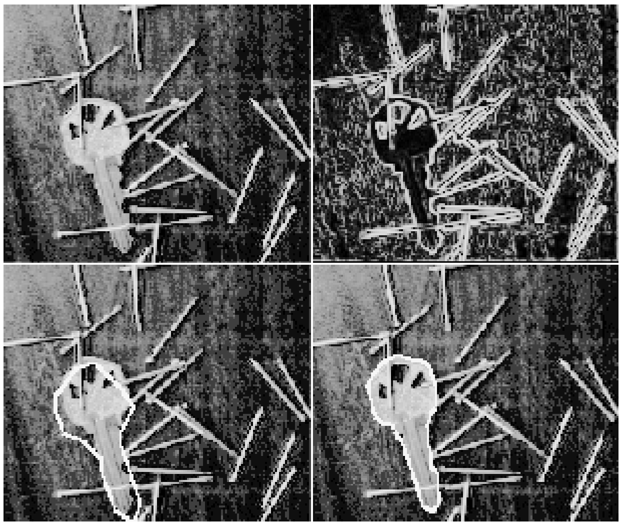


(a) initial (b) final

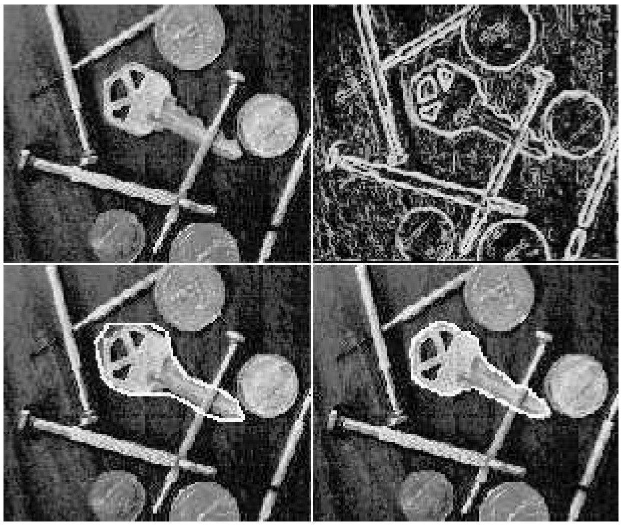
Figure 5: Extraction of handwritten characters.

5.4 Matching Performance

This section compares the matching performance of g-snake to rigid template under the effect of deformation and image noise. Matching performance is defined



(a)



(b)

Figure 4: Extraction of keys from noisy, occluded and cluttered images. In clockwise direction, each example shows original image, edge magnitude, final and initial boundary.

as the degree of correlation between a template and an image:

$$c(\mathbf{U}) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{h}_i^T \mathbf{f}(\mathbf{u}_i + \mathbf{g}) \tag{24}$$

$c(\mathbf{U})$ takes value between 0 for no correlation and 1 for perfect correlation. Using rigid template (generalized Hough transform) and g-snake with minimax regularization, we obtain the respective average correlation $\mathbf{E}[c(\mathbf{U})]$ and plot the results against deformation.

The experiment uses synthesized images containing deformed rectangles as shown in Figure 6(a). We deform the boundary on each of the 4 sides independently using a randomly generated cosine function. Let $b(s)$ represents an undeformed boundary indexed by $s \in [0, 1]$. The deformed boundary $d(s)$ is given by:

$$d(s) = b(s) + \nu \cos(\pi s + \phi) \tag{25}$$

where $\nu : N(0, \sigma_\nu^2)$ is a normally distributed boundary deformation process and ϕ is a random phase shift uniformly distributed in $[0, 2\pi)$. Large values of σ_ν induce large values of deformation variance σ_i, although the exact relationship cannot be determined. A plot of $\mathbf{E}[c(\mathbf{U})]$ versus σ_ν will therefore illustrate the effect of deformation on matching performance.

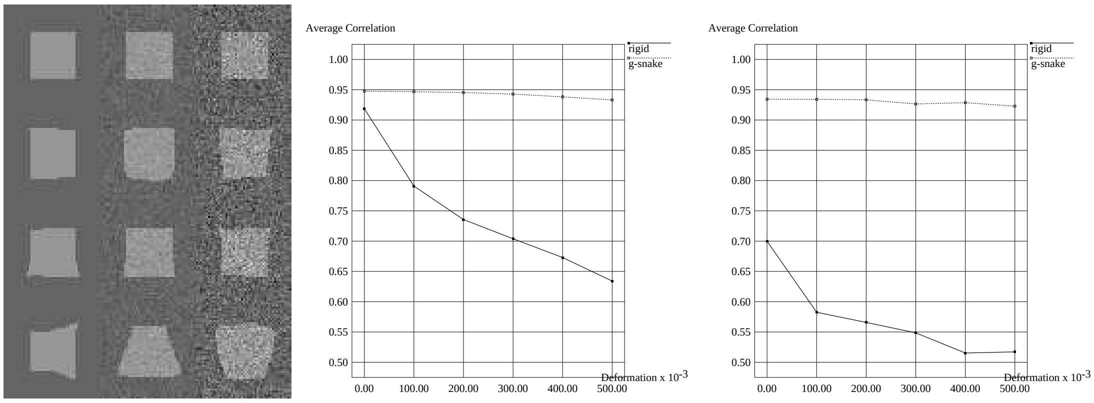
The intensity images are generated by setting the pixel value to 1 if it is enclosed by the boundaries, and 0 otherwise. Image noise is then introduced using zero-mean Gaussian white noise of variance σ_n.

Figure 6(a) to (d) show the plots of $\mathbf{E}[c(\mathbf{U})]$ versus deformation with $\sigma_n = 0.1$ and 0.3. These experiments confirm that matching performance of rigid template degrades with deformation. In contrast, the g-snake adapts well with deformation to yield high degree of correlation. Furthermore, by comparing the two plots, we observe that g-snake exhibits higher robustness to image noise.

6 Conclusions

We considered the problems of modeling and extracting arbitrary deformable contours from noisy images. Based on a *regenerative* shape matrix, our model encompasses both global and local deformations. In addition, it is *stable, invariant* and *unique.* Combined with the Markov random field to model local deformations, this yields invariant *a priori* distribution that exert influence over an arbitrary global model while allowing for deformation.

Using the Bayesian framework, the problem of extracting an object with unknown deformation from noisy images turns into MAP estimation. We showed that MAP estimation is equivalent to energy mini-



(a) (b) $\sigma_n = 0.10$ (c) $\sigma_n = 0.30$

Figure 6: (a) Synthetic images used to compare matching performance. Deformation variance σ_ν increases from top to bottom while image noise variance σ_n increases from left to right. (b)(c) Plots of average correlation $\mathbf{E}[c(\mathbf{U})]$ versus σ_ν with $\sigma_n = 0.1$ and 0.3.

mization in g-snake. Unlike snake, g-snake is capable of representing any arbitrary shape. We exploited the minimax principle to adaptively determine the optimal regularization when training samples are unavailable or unreliable. Furthermore, we may reliably and efficiently initialize the g-snakes using generalized Hough transform.

Finally, we demonstrated with experiments how one may apply the proposed g-snake in various applications. Quantitative measures obtained from the experiments confirm that g-snakes yield superior matching performance compared to rigid templates.

References

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

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

[3] Kass, A. Witkin & D. Terzopoulos, "Snakes: Active Contour Models," *Proceedings of First International Conference on Computer Vision*, 1987, pp. 259-269.

[4] A. L. Yuille, D. S. Cohen & P. W. Hallinan, "Feature Extraction from Faces using Deformable Templates," *Proc CVPR* 1989, pp. 104-109.

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

[6] 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.

[7] 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.

[8] R. Kindermann & J. L. Snell, *Markov Random Fields and their Applications*, American Mathematical Society, 1980.

[9] K. F. Lai & R. T. Chin, "On Regularization, Formulation and Initialization of the Active Contour Models (Snakes)," *Asian Conference on Computer Vision*, 1993, pp. 542-545.

[10] 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.

[11] D. J. Williams & M. Shah, "A Fast Algorithm for Active Contours and Curvature Estimation," *Computer Vision, Graphics, Image Processing*, vol. 55, 1992, pp. 14-26.

[12] A. Rattarangsi & R. T. Chin, "Scale-based Detection of Corners of Planar Curves," *IEEE Trans. Pat. Anal. Mach. Intell.*, vol. **PAMI**-14, 1992, pp. 430-449.