Object Matching Using Deformable Templates

Anil K. Jain, *Fellow, IEEE*, Yu Zhong, and Sridhar Lakshmanan

Abstract—We propose a *general* object localization and retrieval scheme based on object shape using deformable templates. Prior knowledge of an object shape is described by a prototype template which consists of the representative contour/edges, and a set of probabilistic deformation transformations on the template. A Bayesian scheme, which is based on this prior knowledge and the edge information in the input image, is employed to find a match between the deformed template and objects in the image. Computational efficiency is achieved via a coarse-to-fine implementation of the matching algorithm. Our method has been applied to retrieve objects with a variety of shapes from images with complex background. The proposed scheme is invariant to location, rotation, and moderate scale changes of the template.

Index Terms—Object matching, deformable templates, image database, image segmentation, Bayesian optimization, multiresolution algorithm.

1 Introduction

This paper addresses the problem of locating and retrieving an object from a complex image using its 2D shape/boundary information. This problem has wide applications in image processing and computer vision, including image database retrieval, object recognition, and image segmentation. In all of these applications a priori shape information is available in the form of an inexact model of the object which needs to be matched to the objects present in the input image. For example, in an image database retrieval system, the user may have some clues about the object of interest in terms of its shape, color, texture, etc. An automatic image retrieval system [9], [16], [20], [32] should be able to search the database for the images which contain objects with similar characteristics as specified by the user.

We approach the problem of object localization and identification as a process of matching a deformable template to the object boundary in an input image. The prior shape information of the object of interest is specified as a sketch or binary template. This prototype template is not parameterized, but it contains the edge/boundary information in the form of a bitmap. Deformed templates are obtained by applying parametric transforms to the prototype, and the variability in the template shape is achieved by imposing a probability distribution on the admissible mappings. Among all such admissible transformations, the one that minimizes a Bayesian objective function is selected.

The objective function we try to minimize consists of two terms. The first term plays the role of a Bayesian data likelihood. This likelihood term is a potential energy linking the edge positions and gradient directions in the input image to the object boundary specified by the deformed template. The second term corresponds to a Bayesian prior. This prior term penalizes the various deformations of the template— large deviations from the prototype result in a large penalty.

The deformable template minimizes the objective function by iteratively updating the transformation parameters to alter the shape of the template so that the best match between the deformed template and the edges in the image is obtained. The objective function is non convex, and in order to find its minima efficiently we employ a multiresolution algorithm that uses deformed templates at coarser resolutions to initiate matchings at finer resolutions.

We note that certain elements of our work bear a close resemblance to existing studies. For example, the representation of the deformation as probabilistic transformations on the prototype template is akin to the work of Grenander and his colleagues [1], [5], [28], where such transformations are used to derive a set of object images from the "ideal" one. Also, the use of potential functions to influence template deformations towards salient image features is akin to the work in [23], although the potentials are constructed differently based on the edge positions and directions.

In our opinion, the primary contribution of this paper is that it sensibly combines existing ideas along with new ones to provide a systematic paradigm for *general* object matching. Our experimental results show that under this new paradigm:

1) we can match objects that are curved or polygonal, closed or open, simply-connected or multiply-connected;
2) we can retrieve objects based on boundary information alone, even in complex images;
3) we can localize objects independent of their location, orientation, size, and number in the image; and
4) we can achieve such a *general* object matching in a computationally efficient coarse-to-fine manner.

The rest of the paper is organized as follows. In Sec-

• *A.K. Jain and Y. Zhong are with the Department of Computer Science, Michigan State University, East Lansing, MI 48824. E-mail: jain@cps.msu.edu, zhongyu@cps.msu.edu.*
• *S. Lakshmanan is with the Department of Electrical and Computer Science Engineering, University of Michigan, Dearborn, MI 48128. E-mail: lakshman@umich.edu.*

Manuscript received Jan. 13, 1995; revised Nov. 9, 1995. Recommended for acceptance by B. Dom.
For information on obtaining reprints of this article, please send e-mail to: transactions@computer.org, and reference IEEECS Log Number P95174.

tion 2, we review some well-known approaches to template matching, and, in particular, the deformable template matching approach. Section 3 describes our general deformable template model. Section 4 presents the Bayesian approach for the deformable template matching and the coarse-to-fine matching algorithm. The experimental results of our approach are presented in Section 5. We conclude the paper in Section 6 with a discussion and an outline of future work.

2 Related Work and Literature Review

There is a rich collection of publications on shape matching using either rigid or deformable templates. An elegant and versatile technique to detect parameterized shapes (of object boundaries) was first proposed by Hough [21]. It was later generalized to detect any shape represented in a tabular form by Ballard [2]. Basically, the Hough method transforms points in the spatial feature space into a parameter space, and the specified shape is detected by finding the peak(s) in the parameter space. In this way, a global evidence accumulation process of shape detection is transformed into a search for local peaks. Furthermore, the Hough Transform (HT) method is relatively insensitive to noise and occlusion. However, its application is limited because of its excessive requirement for memory and computation especially as the number of parameters increases. Complete surveys on different variants of the HT technique and its applications can be found in [22] and [25].

The HT method can be viewed as template matching. However, it is a rigid scheme in that it is not capable of detecting a shape which is *different* from the template by transformations other than translation, rotation or scaling. A deformable template, on the other hand, is able to "deform" itself to fit the data, by transformations that are possibly more complex than translation, rotation, and scaling.

The deformable models that have been proposed in the literature can be partitioned into two classes:

1) free-form, and
2) parametric.

In free-form deformable models, there is no global structure of the template; the template is constrained by only local continuity and smoothness constraints. Such a template can be deformed to match salient image features like lines and edges using potential fields (energy functions) produced by those features. Since there is no global structure for the template, it can represent an arbitrary shape as long as the continuity and smoothness constraints are satisfied. One example of free-form deformable model is the elastic deformable model [3], [29]. This method establishes an elastic model for one of the two images to be matched. Then this image is "warped" iteratively towards the other one by some local forces. The applications of this model include handwritten numeral recognition, cartoon frame filling, motion detection, and volume matching. Another example is the active contour model proposed by Kass et al. [23] and Terzopoulos et al. [33]. In this approach, an energy-minimizing spline, called a "snake," is driven by a mixture of

1) the internal spline force which enforces the smoothness,
2) the image force which attracts the spline to the desired features, and
3) the external constraint force.

Each force creates its own potential field and the spline actively adjusts its position and shape until it reaches a local minimum of the potential energy. This idea of active contour has been successfully applied to edge and subjective contour detection, motion tracking, stereo matching and image segmentation [6], [23], [26], [34].

A parametric deformable template is used, on the contrary, when some prior information of the geometrical shape is available, which can be encoded using a small number of parameters. In general, a parametric deformable template can be represented as either

1) a collection of parameterized curves, or
2) the image of a prototype template under a parametric mapping.

In the first scheme, the template is represented by a set of curves which is uniquely described by some parameters. The specific analytical form incorporates the prior knowledge of the shape of the objects under analysis. The geometrical shape of the template can be changed by using different values of the parameters. Similar to the free-form deformable model, a potential field can be established based on the salient image features. The deformable template evolves to minimize its energy by updating the parameters.

Lakshmanan and Grimmer. [24], for example, have used a parametric template model to locate the road boundary in radar images, where the two straight, parallel edges of a road are parameterized and the edge detection problem is formulated as a Bayesian estimate using a physics-based model of the radar imaging process. Yuille et al. [35] have drawn eye and mouth templates using circles and parabolic curves. The parameters which control the shape of a template are the center and the radius of the circle, and the characteristic parameters of the parabola. The image energy term is defined in terms of edges, peaks, and valleys in the input intensity image. Yuille et al. were able to accurately locate eyes and mouths in real images when the initial positions of the templates is close enough to the desired objects. Boundary templates with more degrees of freedom were proposed by Staib and Duncan [30] to detect objects in medical images. They used elliptic Fourier descriptors to represent open or closed boundary templates, where the parameters of the deformable templates are the Fourier coefficients. A distribution on the Fourier coefficients is specified to favor particular shapes. A Bayesian decision rule is then used to obtain the optimal estimate of the boundary, where the likelihood function is based on the correlation between the template and the boundary strength in the input image. A similar scheme is employed by Chakraborty et al. [4] where the likelihood function is linked to both the region homogeneity and the boundary strength. The applicability of parametric deformable model is limited because the shapes under investigation have to be well-defined so that they can be represented by a set of

curves with preferably a small number of parameters.

The pattern theory proposed by Grenander [17] described a systematic framework to represent shape classes of a characteristic structure, which can accommodate certain variability. Their shape model consists of

1) a model template which describes the overall architecture of the shape, and
2) a parametric statistical mapping which governs the random variations in the building blocks of the shape [1], [5], [18], [19], [28].

Usually, the prototype template is based on the prior knowledge of the objects, which is often obtained from training samples. The parametric statistical mapping is chosen to reflect the particular deformations allowed in the application domain.

The shape classes described by Grenander's pattern theory can be very versatile because of different choices of the prototype template and the deformation process [18]. For example, in their work on human hands [5], Chow et al. used a polygon to represent the contour of a human hand, where the shape building blocks are the polygonal edges. Variations in different hands are described by Markov processes on the edges. In another paper on restoration of human hands from noisy images, Amit et al. [1] used an intensity image to represent a typical human hand. All instances of the class of hands are obtained by applying a number of admissible transformations to the "ideal" hand image.

A similar scheme is used by Cootes et al. [7] in their work on line-drawing type of "active shape models". They compute the "mean" shape of a class of correctly annotated training objects as the prototype template. The deformations, given the standard shape, are modeled using linear combinations of the eigenvectors of the variations from the mean shape. This scheme is capable of representing a class of similar shapes with a specific variation. Their active shape model is able to learn the characteristic pattern of a shape class and can deform in a way which reflects the variations in the training set.

Our deformation model falls into the second class of parametric deformation models. It shares some of the characteristics of the work by Grenander et al. and by Cootes et al., but has its own characteristics which are appropriate for the specific application domain of interest to us. We represent the prototype template in the form of a bitmap which describes the characteristic contour/edges of an object shape. It is then deformed to fit salient edges in the input image by applying a probabilistic transformation on the prototype contour which maintains smoothness and connectedness. The matching is carried out by maximizing the a posteriori probability which combines both the prior shape information and the image information. Bayesian frameworks have been previously adopted for contour estimation [10], [31] where the prior is used to impose local smoothness and the likelihood is calculated based on edge positions. In our case, it is natural to choose a prior which reflects the global shape of the object of interest. The likelihood model that we use to fit the deformed template to the salient edges in the input image is similar to the ones used in [10], [23], [31], but the exact functional form of our likelihood is different because it incorporates both the edge position and direction information to give a better edge localization. Details of the deformation and likelihood models are given in the subsequent sections.

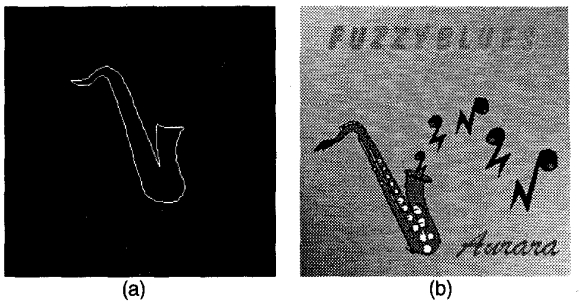
3 A General Deformation Model

The proposed deformable model is appropriate in situations where an inexact knowledge about the shape of the object of interest is available and when this shape information can be represented in the form of a hand-drawn sketch. In content-based image database retrieval systems, queries often include the shape of the object. The user may describe the shape of an object using a sketch and ask to retrieve all the images in the database that contain such a shape. The sketch used to describe the shape does not need to match the object boundaries in the image exactly. Fig. 1a shows the sketch of an object and Fig. 1b shows an image which contains a deformed version of this template. It is important that a retrieval system be robust to position, orientation, scale differences, and most importantly, moderate deformations of the object shape.

In our matching scheme, the deformation model consists of

1) a prototype template which describes a representative shape of a class of objects in terms of a bitmap sketch,
2) a set of parametric transformations which deform the template, and
3) a probability distribution defined on the set of deformation mappings which biases the choice of possible deformed templates.

Each of the three components is discussed in more detail in the following subsections.



(a) (b)

Fig. 1. Deformable template matching: (a) a prototype template (bitmap of contour) of a saxophone, (b) an image containing a saxophone to be matched with the template in (a).

3.1 Representation of the Prototype Template

The prototype template consists of a set of points on the object contour, which is not necessarily closed, and can consist of several connected components. We represent such a template as a bitmap, with bright pixels (grey level of 255) lying on the contour and dark pixels (grey level of 0) elsewhere (Fig. 1a). Such a scheme captures the global structure of a shape without specifying a parametric form for each class of shapes. This model is appropriate for general shape

matching since the same approach can be applied to objects of different shapes by drawing different prototype templates.

3.2 Deformation Transformations

The prototype template describes only one of the possible (though most likely) instances of the object shape. Therefore, it has to be deformed to match objects in images. We introduce a set of deformation transformations with associated distributions, so that the template can deform its shape to match the objects in the database. A prerequisite for these transformations is that if the prototype template is connected, then the transformations should preserve the connectedness. We perform a deformation of the template by introducing a displacement field in the 2D template image.

Imagine that the template is drawn on a 2D planar rubber sheet which has a fixed boundary, but it can be deformed by stretching, squeezing, and twisting locally in the interior. As the rubber sheet deforms, the template drawn on it also changes its shape. The deformed rubber sheet can be obtained by applying a continuous mapping which maps the domain of the 2D image onto itself. The resulting 2D displacement field is represented as a continuous 2D vector function with certain boundary conditions. Without a loss of generality, we assume that the template is drawn on a unit square $S = [0, 1]^2$. The points in the square are mapped by the function $(x, y) \mapsto (x, y) + (\mathcal{D}^x(x, y), \mathcal{D}^y(x, y))$, where the displacement functions $\mathcal{D}^x(x, y)$ and $\mathcal{D}^y(x, y)$ are continuous and satisfy the following boundary conditions: $\mathcal{D}^x(0, y) \equiv \mathcal{D}^x(1, y) \equiv \mathcal{D}^y(x, 0) \equiv \mathcal{D}^y(x, 1) \equiv 0$. The space of such displacement functions is spanned by the following orthogonal bases [1]:

$$\begin{aligned} e^x_{mn}(x,y) &= \big(2\sin(\pi nx)\cos(\pi my), 0\big) \\ e^y_{mn}(x,y) &= \big(0, 2\cos s(\pi mx)\sin(\pi ny)\big), \end{aligned} \tag{1}$$

where $m, n = 1, 2, \ldots$. These basis functions, which consist of trigonometric functions of different frequencies, vary from global and smooth to local and "coarse" as m and n increase. Specifically, the displacement function is chosen as follows:

$$\mathcal{D}(x,y) = \big(\mathcal{D}^x(x,y), \mathcal{D}^y(x,y)\big) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \frac{\xi^x_{mn}\cdot \mathbf{e}^x_{mn} + \xi^y_{mn}\cdot \mathbf{e}^y_{mn}}{\lambda_{mn}} \tag{2}$$

where $\lambda_{mn} = \alpha\pi^2(n^2 + m^2)$, $m, n = 1, 2, \ldots$ are the normalizing constants. The parameters $\underline{\xi} = \big\{\big(\xi^x_{mn}, \xi^y_{mn}\big), m, n = 1,2,\ldots\big\}$, which are the projections of the displacement function on the orthogonal basis, uniquely define the displacement field, and hence the deformation. We use a finite number of terms in the infinite series in (2) as the displacement function for the deformation mapping:

$$\mathcal{D}_{\underline{\xi}}(x,y) = \sum_{m=1}^{M}\sum_{n=1}^{N} \frac{\xi^x_{mn}\cdot \mathbf{e}^x_{mn} + \xi^y_{mn}\cdot \mathbf{e}^y_{mn}}{\lambda_{mn}}. \tag{3}$$

Note that the dependence of the displacement $\mathcal{D}$ on the deformation parameter vector $\underline{\xi}$ is made explicit in (3). This continuous function preserves the connectedness of the prototype template. It also preserves the smoothness of the template when M and N are not too large (only low frequency components are used). It should be noted that the length of the deformable template varies depending on the deformation. Note also that we are only concerned with the displacements of the points on the prototype template. Fig. 2 illustrates the deformations of a saxophone template sketched on a grid using the displacement functions defined in (3), as the order (M, N) increases; the deformation coefficients corresponding to the lower order terms remain the same as (M, N) increases. One can see that the deformation becomes more complex as higher frequency components are added to the displacement function.

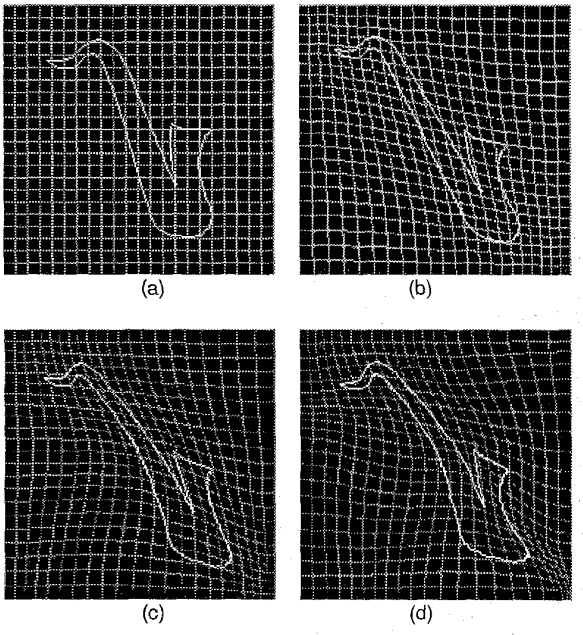


Fig. 2. Deformed saxophone templates with different transform interpolation levels: (a) no deformation; (b) order 1 ($N = M = 1$); (c) order 2 ($N = M = 2$); (d) order 3 ($N = M = 3$).

3.3 A Probabilistic Model of Deformation

The family of functions defined in (3) can represent complex deformations by choosing different representation coefficients ξ_{mn} and different values of M and N. Not all the transformations result in a deformed template that visually resembles the prototype template. Usually, large values of the ξ_{mn} result in a large deformation. As all the prior shape information is represented in the prototype template, it is natural to assume that the prototype template exemplifies the most likely a priori shape of the object. Also, small deformations that leave the template similar to its original shape are more likely than large displacements. We impose a probability density on the family of functions in (3) to bias the possible deformed templates. Specifically, the ξ_{mn}'s are assumed to be independent of each other, independent along x and y directions, and identically Gaussian distributed with mean zero and variance σ^2:

$$\Pr(\underline{\xi}) = \prod_{m,n=1}^{M,N} \Pr(\xi_{mn}) = \prod_{m,n=1}^{M,N} \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{{\xi^x_{mn}}^2 + {\xi^y_{mn}}^2}{2\sigma^2}\right\}$$
$$= \frac{1}{(2\pi\sigma^2)^{MN}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{m,n}\left({\xi^x_{mn}}^2 + {\xi^y_{mn}}^2\right)\right\} \quad (4)$$

The value of σ^2 reflects the confidence about the prototype template, with large values of σ^2 allowing more deformation. This model is chosen for its simplicity, although the *independent and identically distributed* assumption of the parameters may not be appropriate in some situations.

4 Bayesian Formulation and Objective Function

A Bayesian inference scheme is employed to integrate the prior shape knowledge of the template and the observed object in the input image. The prior information of the object shape can be presented by a combination of

1) a prototype template,
2) a set of deformation transformations of the template, and
3) a probability density on the set of deformation transformations.

We propose an energy function based on the image boundary strength and the deformed template in order to arrive at the likelihood. This likelihood is then combined with the prior using Bayes' rule to obtain the a posteriori probability density of the deformations of the template given the input image. The object is located by deforming the template so that the a posteriori probability density is maximized. The final shape and position of the deformed template gives a description of the object in the image.

4.1 Prior Distribution

We use the prior distribution to bias the global transformations (rotation, translation, and scale change) and local deformations that can be applied to a prototype template. Let $\mathcal{T}_0$ denote the prototype template, and $\mathcal{T}_{s,\Theta,\underline{\xi},\underline{d}}$ a deformation of the prototype. This deformation is realized by rotating the prototype template by an angle Θ, locally deforming the rotation by a set of parameters $\underline{\xi}$, scaling the local deformation by a factor of s, and translating the scaled version along the x and y directions by an amount $\underline{d} = (d^x, d^y)$:

$$\mathcal{T}_{s,\Theta,\underline{\xi},\underline{d}}(x,y) = \mathcal{T}_o\left(s \cdot \left[(x,y) + \mathcal{D}_{\underline{\xi}}(\mathcal{R}_\Theta(x,y))\right] + (d^x, d^y)\right), \quad (5)$$

where $\mathcal{D}_{\underline{\xi}}$ denotes the displacement function given in (3), and $\mathcal{R}_\Theta(x, y)$ is the rotation of a point (x, y) by an angle Θ. Assuming that $s, \Theta, \underline{\xi}$, and $\underline{d}$ are all independent of each other, we have:

$$\Pr(s,\Theta,\underline{\xi},\underline{d}) = \Pr(s) \times \Pr(\Theta) \times \Pr(\underline{\xi}) \times \Pr(\underline{d}). \quad (6)$$

Suppose all translations, rotations, and scale sizes are equally possible as long as the transformed template falls in the input image, (4) reduces to:

$$\Pr(s,\Theta,\underline{\xi},\underline{d}) =$$
$$\kappa \frac{1}{(2\pi\sigma^2)^{MN}} \exp\left\{-\frac{1}{2\sigma^2}\left[\sum_m\sum_n\left({\xi^x_{mn}}^2 + {\xi^y_{mn}}^2\right)\right]\right\} \quad (7)$$

which is essentially the same as (4), except for the normalizing constant κ. This constitutes our Bayesian prior. Intuitively, a deformed template with a geometric shape similar to the prototype template is favored, regardless of its size, orientation, or location in the image.

4.2 Likelihood

The likelihood specifies the probability of observing the input image, given a deformed template at a specific position, orientation and scale. It is a measurement of the similarity between the deformed template and the object(s) present in the image. The likelihood we propose only uses the edge information in the input image. Other definitions of the likelihood can incorporate texture, grey-scale homogeneity, or color information in more complex situations.

The deformable template is attracted and aligned to the salient edges in the input image via a directional edge potential field. This field is determined by the positions and directions of the edges in the input image. For a pixel (x, y) in the input image its edge potential is defined as:

$$\Phi(x,y) = -\exp\left\{-\rho\left(\delta_x^2 + \delta_y^2\right)^{1/2}\right\}, \quad (8)$$

where (δ_x, δ_y) is the displacement to the nearest edge point in the image, and ρ is a smoothing factor which controls the degree of smoothness of the potential field. We modify this edge potential by adding to it a directional component. This new edge potential induces an energy function that relates a deformed template $\mathcal{T}_{s,\Theta,\underline{\xi},\underline{d}}$ to the edges in the input image Y:

$$\mathcal{E}(\mathcal{T}_{s,\Theta,\underline{\xi},\underline{d}}, Y) = \frac{1}{n_{\mathcal{T}}}\sum\left(1 + \Phi(x,y)\left|\cos(\beta(x,y))\right|\right), \quad (9)$$

where the summation is over all the pixels on the deformed template, n_T is the number of pixels on the template, $\beta(x, y)$ is the angle between the tangent of the nearest edge and the tangent direction of the template at (x, y), and the constant 1 is added so that the potentials are positive and take values between 0 and 1. This definition requires that the template boundary agrees with the image edges not only in position, but also in the tangent direction. This feature is particularly useful in the presence of noisy edges. The lower this energy function the better the deformed template matches the edges in the input image. Using this energy function we now define a probability density that specifies the likelihood of observing the input image, given the deformations of the template:

$$\mathcal{P}_r(Y \mid s,\Theta,\underline{\xi},\underline{d}) = \alpha \exp\left\{-\mathcal{E}(\mathcal{T}_{s,\Theta,\underline{\xi},\underline{d}}, Y)\right\}, \quad (10)$$

where α is a normalizing constant to ensure that the above

function integrates to 1. The maximum likelihood is achieved when $\mathcal{E}\left(\mathcal{T}_{s,\Theta,\underline{\xi},\underline{d}}, Y\right) = 0$, i.e., when the deformed template $\mathcal{T}_{s,\Theta,\underline{\xi},\underline{d}}$ exactly matches the edges in the input image Y.

We note that the normalizing constant α is not necessarily independent of the deformed template [12]. Exact computation of this constant involves an intractable summation. If this constant can be computed, then it removes the inherent bias of the energy function (and hence the likelihood) towards deformations of the template that increases its size. Since we cannot compute it, we have chosen to incorporate its effects by normalizing the energy function directly with respect to the template size—see (9). We acknowledge that this represents a deviation from proper Bayesian inference, but many of the existing studies also suffer from the same drawback.

4.3 Posterior Probability Density

Using Bayes rule, the prior probability of the deformed templates in (7) and the likelihood of the input image given the deformed template in (10) can be combined to obtain the a posteriori probability density of the deformed template given the input image:

$$\left.\begin{aligned}
&\mathcal{P}_r\left(s,\Theta,\underline{\xi},\underline{d}\,|\,Y\right) = \\
&\mathcal{P}_r\left(Y\,|\,s,\Theta,\underline{\xi},\underline{d}\right)\mathcal{P}_r\left(s,\Theta,\underline{\xi},\underline{d}\right)/\mathcal{P}_r(Y) \\
&= C_1 \exp\left\{-\mathcal{E}\left(\mathcal{T}_{s,\Theta,\underline{\xi},\underline{d}}, Y\right)\right\} \prod_{m=1}^{M}\prod_{n=1}^{N} \frac{1}{2\pi\sigma^2}\exp\left\{-\frac{1}{2\sigma^2}\left(\xi_{mn}^{x\,2} + \xi_{mn}^{y\,2}\right)\right\} \\
&= C_2 \exp\left\{-\mathcal{E}\left(\mathcal{T}_{s,\Theta,\underline{\xi},\underline{d}}, Y\right) - \sum_{m=1}^{M}\sum_{n=1}^{N}\frac{1}{2\sigma^2}\left(\xi_{mn}^{x\,2} + \xi_{mn}^{y\,2}\right)\right\}
\end{aligned}\right\} \quad (11)$$

where C_1 and C_2 are normalizing constants.

Our objective is to maximize the a posteriori density in (11) with respect to $s, \Theta, \underline{\xi}, \underline{d}$. Taking the natural logarithms on both sides of (11) results in:

$$\begin{aligned}
&\log \mathcal{P}r\left(s,\Theta,\underline{\xi},\underline{d}|Y\right) = \\
&\log C_2 - \mathcal{E}\left(\mathcal{T}_{s,\Theta,\underline{\xi},\underline{d}}, Y\right) - \sum_{m=1}^{M}\sum_{n=1}^{N}\left(\xi_{mn}^{x\,2} + \xi_{mn}^{y\,2}\right)\Big/2\sigma^2.
\end{aligned} \quad (12)$$

Equivalently, we seek to minimize the following objective function with respect to $s, \Theta, \underline{\xi}, \underline{d}$:

$$\begin{aligned}
&\mathcal{L}\left(\mathcal{T}_{s,\Theta,\underline{\xi},\underline{d}}, Y\right) = \\
&\mathcal{E}\left(\mathcal{T}_{s,\Theta,\underline{\xi},\underline{d}}, Y\right) + \gamma\sum_{m=1}^{M}\sum_{n=1}^{N}\left(\xi_{mn}^{x\,2} + \xi_{mn}^{y\,2}\right)
\end{aligned} \quad (13)$$

where $\gamma = 1/2\sigma^2$. This objective function consists of two parts:

1) a model-based term which measures the deviation of the deformed template from the prototype template, and
2) a data-driven term which describes the fitness of the deformed template contour to the boundary in the image.

The parameter γ can be interpreted as providing a relative weighting of the two penalty measures; a larger value of γ implies a lower variance of the deformation parameters, and as a result, a more rigid template.

4.4 Multiresolution Algorithm

The objective function to be minimized in (13) is not unimodal. In fact, it is a very complex function over the deformation parameter space. The minima for this function can, in principle, be obtained by using Monte Carlo relaxation algorithms such as the Gibbs sampler [13], the Metropolis algorithm [11], [27], or the stochastic diffusion algorithms [15], [14], [19]. In all such algorithms, the minima are obtained by constructing an ergodic Markov chain whose limiting (stationary) distribution has support over only the modes of the a posteriori probability density function. However, because of the characteristics of stochastic sampling, these approaches achieve the optimal solution at the cost of excessive computing.

We have employed a multiresolution approach to quickly find the good solutions. At the coarsest stage, a smooth potential field is used with a large value of ρ in (8). This smooth potential field has fewer spurious local minima, which helps the deformable template to find the global valleys. This stage attempts to roughly locate the global optima efficiently without regard to accuracy. Therefore, it can be performed at a reduced resolution. A subsampled deformable template is placed at a set of regular positions, and a set of discretized orientations, in the input image using the smooth edge potential field. The spacing between the template positions is chosen to be one fourth the size of the template so that all the significant local minima of the energy surface are covered. This computation can be done efficiently because

1) we work only with a subsampled template,
2) we use coarse step-sizes for the deformation parameters,
3) initial positions with considerably high energy can be discarded immediately, and
4) fewer numbers of deformation parameters and iteration steps are needed since we are only interested in an approximate match.

Finer level matchings are initialized using the good candidates screened from the previous stage. At finer stages, finer energy potential fields are used because more accurate localization is desired. The deformed templates obtained from the previous stage with low energy (below a threshold) are used as initial templates for the matching. A larger number of deformation parameters and finer step sizes are used to obtain better matches. We have used three stages altogether. In all the stages, the local minimization is performed by using a deterministic gradient descent algorithm. The hierarchy of step sizes from coarse to fine allows the multi-stage process to escape local minima in the deformation parameter space. This coarse-to-fine matching can automatically find the acceptable solutions to the minimization problem at an affordable computational cost.

The algorithm for object localization is summarized as follows:

```
1. Calculate the edge map and gradient direction of
   the input image using the Canny edge detector;
2. Compute the directional edge potential images at
   three resolution levels (coarse to fine) according
   to the edge map;
```

3. Perform the coarsest-level ($M = N = 1$) matching, initialized at evenly spaced positions, and over a discretized set of orientations, using a coarse step size and a smooth edge potential field. The matching process is:
   Loop: Calculate the objective function and the partial derivatives[1] with respect to parameters $s, \Theta, \underline{\xi}, \underline{d}$
   - If the objective function is less than a threshold, goto Step 4 (finer-level matching).
   - Else if the number of iterations exceeds a threshold, report no object of interest and stop;
   - Else use the gradient descent method to update the deformation parameters. Go to *Loop*.
4. Perform finer-level ($M = N = 2$) matching initialized by the candidate templates generated by the coarsest-level search. If the objective function is less than a threshold, goto Step 5 (finest-level matching). The deformation procedure is the same as described in Step 3, but a finer step size and a less smooth edge potential field are used.
5. Perform finest-level ($M = N = 3$) matching initialized by the candidate templates generated by the finer-level ($M = N = 2$) search. If the objective function is less than a threshold, output the retrieved object. Here, an even finer step-size and a less smooth edge potential field than Step 4 is used.

5 Experimental Results

The proposed deformable template-based matching scheme has been applied to different objects in several real images. Our experimental results have been divided into six sets. Before we describe those results, we discuss some pertinent issues.

Given an input image, we start out by sketching a prototype template which resembles an object of interest in the image. We have used the Canny edge detector ($\sigma = 1$, mask size is 9×9, and the lower and upper thresholds are 0.5 and 0.9, respectively) to calculate the edge map of the input image as well as the gradient direction at each edge point. We discretized the template orientation into a number of different directions (currently, we use 12 different directions) which uniformly cover the interval $[0^\circ, 360^\circ]$. The deformed template is initialized at each orientation. Templates with orientations within each of these direction intervals are expected to be recovered by the deformation process itself. The M and N values in the displacement function (3) were set to 1, 2, and 3, respectively, for the three coarse to fine resolution levels. Unlike other parameters, we do not perform a global minimization of the objective function over all the values of s. Instead, we settle for a minimum of the objective function when s is in the local neighborhood of 1.0. In terms of the $\underline{d}$ parameters, we move the template around only at the coarsest resolution, $M = N = 1$. We initialize the coarse template at different positions. The spacing between these positions is about one fourth the size of the template. We have used three different values (12, 8, and 3) of the smoothness parameter ρ in (8) when calculating the directional edge potential fields at different resolutions. The parameter γ in (13) which reflects the relative weighting between the prior (template) and the likelihood (data) was set to 0.1. In our experiments, we found that the results were not very sensitive to the the choice of γ as long as it is in the range 0.1 to 1.0. Finally, in order to increase the speed of the likelihood-energy calculations, we precompute some images, namely, the Canny edge image and three potential field images corresponding to the edge potential fields at the three coarse to fine resolutions. An example of these four images is given in Fig. 3.

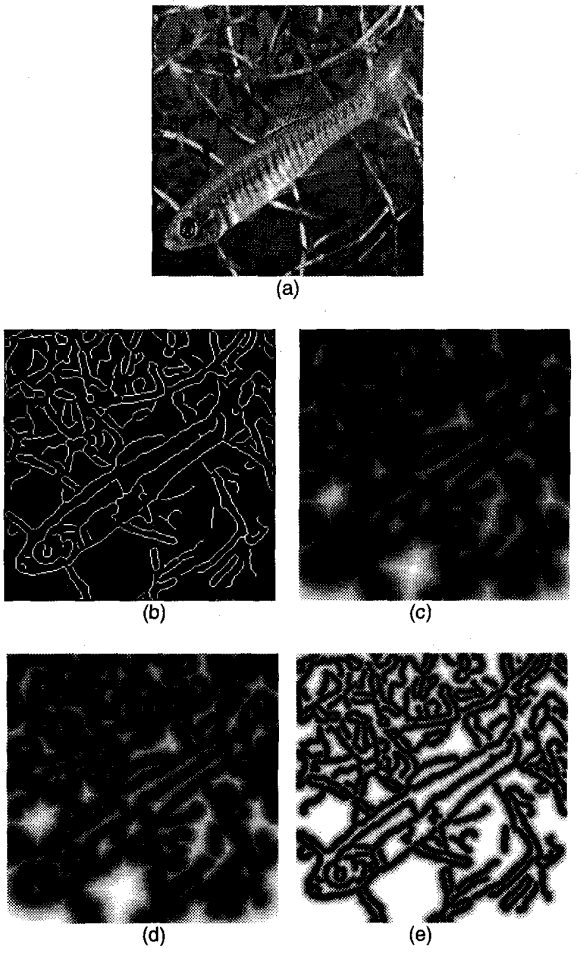


Fig. 3. Preprocessing: (a) input image (256×256); (b) edge map using a Canny edge detector ($\sigma = 3$); (c) magnitude of the coarsest-level edge potential field ($\sigma = 12$); (d) magnitude of the second-level edge potential field ($\sigma = 8$); (e) magnitude of the finest-level edge potential field ($\sigma = 3$).

As mentioned before, our experimental results have been divided into six sets. With each set, the objective is to illustrate one or more salient aspects of our object matching

1. The partial derivatives with respect to the d^x and d^y are approximated by finite differences. The partial derivatives with respect to the rest of the parameters are obtained by the chain rule, and by using the partial derivatives of d^x and d^y with respect to those other parameters.

method. The first set of experiments demonstrates how the given prototype templates deform themselves locally to match the object contours in the input images when placed in a neighborhood of the desired objects. The given templates are initially placed manually in the neighborhood of the object of interest, and the deformable matching is performed directly (and only) at the second resolution level. Then, the gradient descent algorithm is used to find the local minimum of the objective function in (10). The final deformed template as well as the intermediate sequences of the deformable template are presented to show how the deformable template evolves to match the salient edges in the input image.

Fig. 4 shows the localization process when we apply a saxophone template to an image containing a saxophone, where Fig. 4a shows the manually chosen initial position of the template, and Figs. 4b and 4c are the intermediate snapshots of the deformation process. The final template is shown in Fig. 4d. In Fig. 5 we show the same localization sequences when we apply a fish template to an image containing a fish. In both the cases the deformable templates are able to converge to the object contour. The objects of interest are retrieved correctly. It should be noted that the objective function value $\mathcal{L}$ is calculated based on the second resolution level energy field, and the same template will have a higher value of $\mathcal{L}$ if the finest-level (third level) energy field is used in calculation.

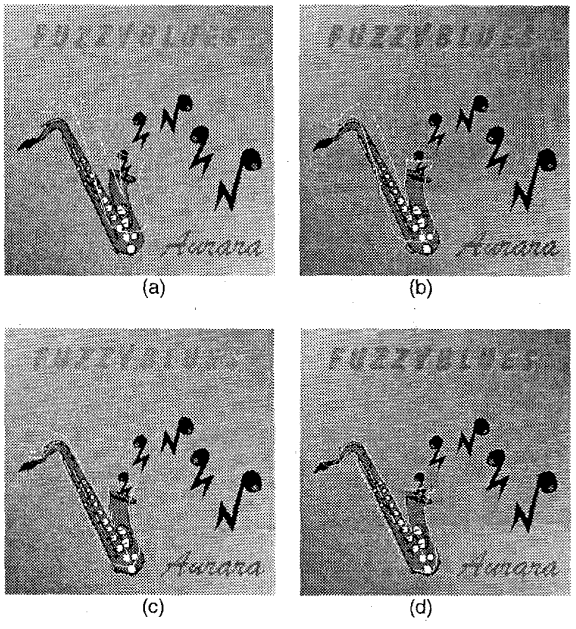


Fig. 4. Localization of a saxophone using manually chosen initial template position. Image size is 285×286. (a) Initial position, $\mathcal{L} = 0.603$; (b) 10 iterations, $\mathcal{L} = 0.327$; (c) 16 iterations, $\mathcal{L} = 0.186$; (d) 30 iterations, $\mathcal{L} = 0.123$.

The second set of experiments uses the three-stage coarse-to-fine deformable template matching scheme to *automatically* search the input image for the desired shapes. We illustrate the fact that our matching scheme can localize objects independent of their location, and orientation in the image. Objects of different shapes are retrieved using different prototype templates. In Fig. 6 we show the retrievals of a guitar and a star separately. Each of the retrieval is illustrated by the initial hand-drawn template, the input image, the resulting template at the coarsest level (initial template for the finer level), and the retrieved object image.

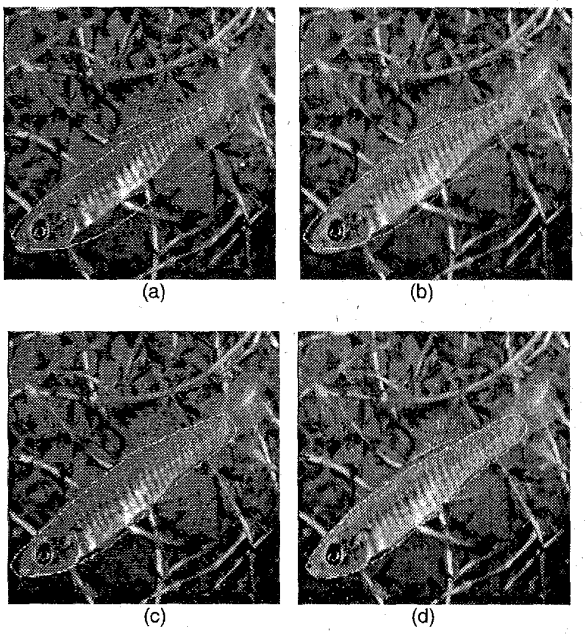


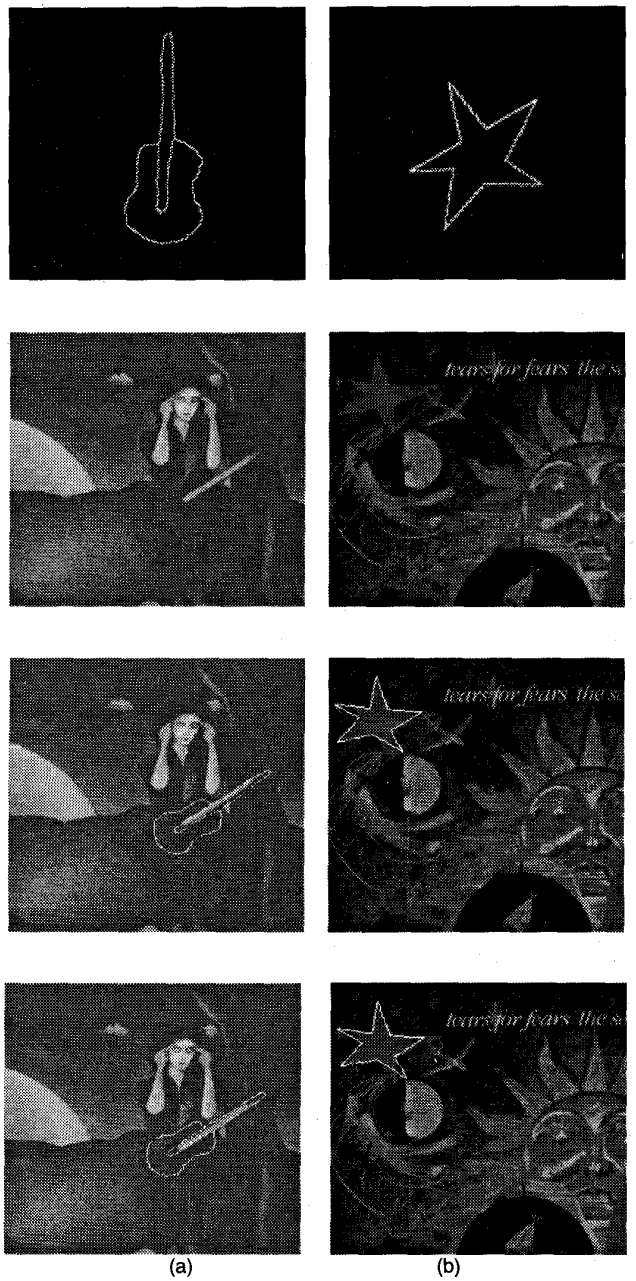
Fig. 5. Localization of a fish using manually chosen initial template position. Image size is 256×256. (a) initial position, $\mathcal{L} = 0.432$; (b) four iterations, $\mathcal{L} = 0.308$; (c) seven iterations, $\mathcal{L} = 0.221$; (d) Forty iterations, $\mathcal{L} = 0.157$.

In the third set of experiments, we demonstrate that our matching scheme is able to retrieve all the objects in an input image that resemble the prototype template. Note that these objects may be of different sizes, different orientations, and may have local variations. In [19], multiple objects resembling the same prototype template were localized by using jump-diffusion algorithms. The addition of jumps further slows down the convergence of computationally demanding diffusion algorithms. In contrast, the multi-resolution minimization method adopted in this paper localizes multiple objects without any additional overhead.[2] In Fig. 7, we have applied a "seed" template to the image of the cross-section of an orange using our multi-stage deformation algorithm. Fig. 7a shows the "seed" template. Fig. 7b shows the input image. All the resulting templates with objective function values below a threshold are presented in Fig. 7c. All the six "seeds" of different orientations are retrieved correctly, and there is no false retrieval.

In the fourth set of experiments, we illustrate the fact that our matching scheme can even handle prototype templates that are not closed. We have applied a prototype hand template to three different hand images, and this template con-

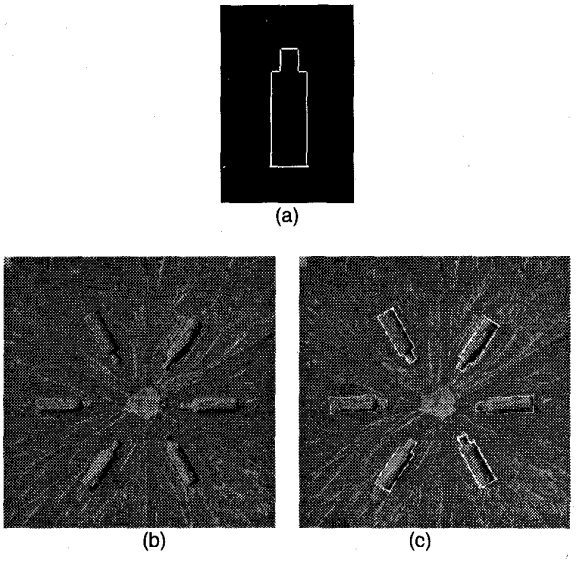
2. Each object resembling the prototype template corresponds to a different minimum of the objective function. The objective function value is below the threshold at each of those minima, and hence all these objects are retrieved.

sists of an open curve. Note that the hand in each of the three input images is slightly different from the others both in shape and orientation. The retrieval process is shown in Fig. 8, where Fig. 8a shows the hand-drawn template, Fig. 8b shows the three different input images, and 8c shows the corresponding retrieval and localization. Note that although the hand instances have different shapes, the deformable template is able to accommodate these variations.



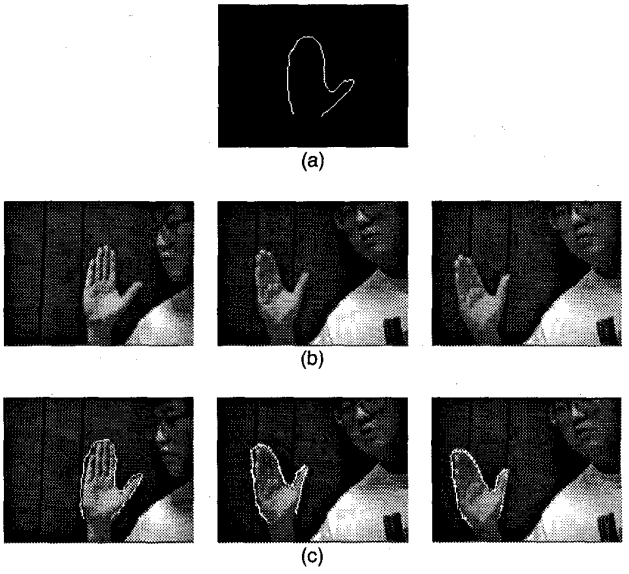
(a) (b)

Fig. 6. Automatic retrieval and localization of desired objects using the proposed three-stage coarse-to-fine matching. (a) Retrieval of a guitar using multiresolution deformable template matching (320 × 304), $\mathcal{L} = 0.186$; (b) retrieval of a star using multiresolution deformable template matching (256 × 256), $\mathcal{L} = 0.157$. (From top to bottom: hand-drawn template, input image, retrieved deformed template at the coarsest level, retrieved deformed template at the finest level.



(a) (b) (c)

Fig. 7. Automatic localization of "seeds" using the proposed three-stage coarse-to-fine matching (a) A "seed" template; (b) the input image of the cross-section of an orange (453 × 436); (c) retrieved objects when the objective function is thresholded at 0.160.



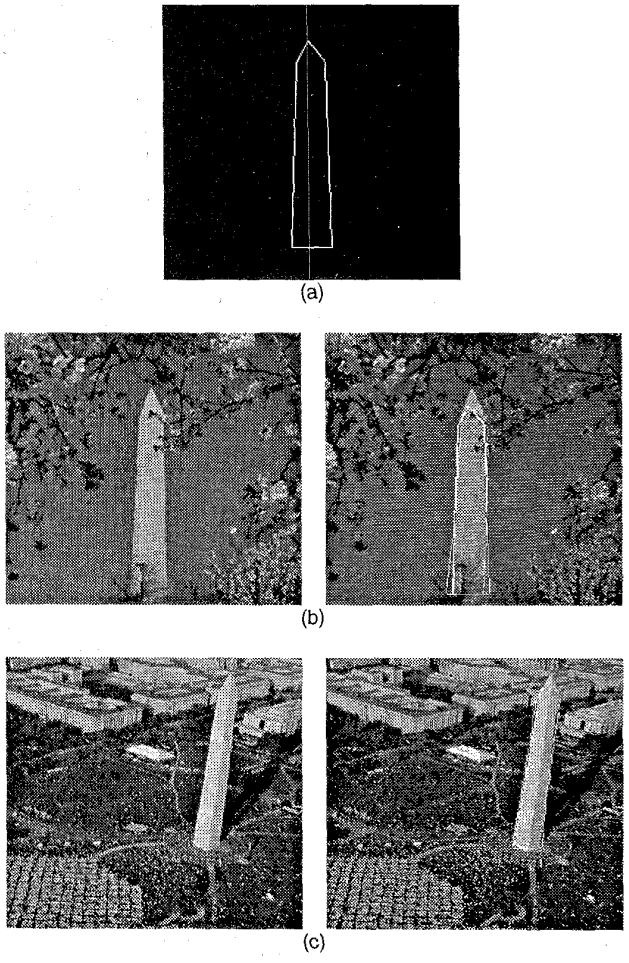
(a) (b) (c)

Fig. 8. Automatic retrieval of human hand using coarse-to-fine matching. (a) The hand template; (b) input images which contain a hand (121 × 160); (c) retrieved hand.

The fifth set of experiments illustrates the scale invariance aspect of our matching scheme. In Fig. 9, we show the retrieval of two tower images using the same prototype tower template. Note the significant scale difference between the template (Fig. 9a) and the corresponding objects in the two images (Figs. 9b and 9c).

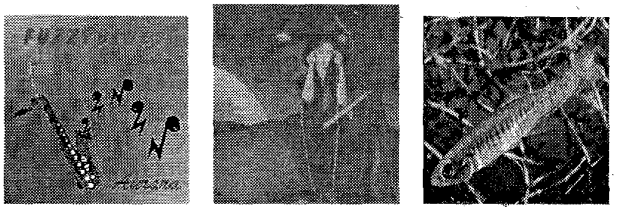
Our objective function can not only be used to retrieve objects in an input image that resemble a prototype template, but the same function can also be used to reject the hypothesis that a certain object is present in an image. We illustrate this in our sixth and final set of experiments. Figs. 4-9 show that when the image contains an object re-

sembling the prototype template, then it can be retrieved and localized accurately by our approach. In the following experiments, we demonstrate that if we apply a deformable template to an image which does not contain an object of similar shape, then the resulting objective function value will be sufficiently large so that we can reject the presence of desired objects in the image. In Fig. 10a, a fish template is applied to the image containing a saxophone, in Fig. 10b we apply the same template to a CD cover image, and in Fig. 10c we apply a saxophone template to the fish image. Table 1 shows the best objective function values when we search different images using different templates. Note that the final objective function values are significantly larger when a wrong template is applied to the image compared to the cases when a correct match between the deformable template and the object contour occurs. The computation time for the multiresolution matching algorithm depends both on the size and complexity of the input image and the template. Excluding the computation of the edge map and the directional edge potential field, it takes about five seconds to retrieve the hand template and 12 seconds to retrieve the saxophone on a Sun Sparc 20 workstation.



(a) (b) (c)

Fig. 9. Applying a tower template. (a) The tower template; (b) retrieval of tower 1 (280 × 280), $\mathcal{L} = 0.227$; (c) retrieval of tower 2 (280 × 280), $\mathcal{L} = 0.243$.



(a) (b) (c)

Fig. 10. What if the template is not present in the image? (a) Applying a fish template to a saxophone image, $\mathcal{L} = 0.587$ ($\mathcal{L} = 0.142$ for the saxophone template); (b) applying the fish template to a guitar image, $\mathcal{L} = 0.230$ ($\mathcal{L} = 0.142$ for the guitar template); (c) applying a saxophone template to a fish image, $\mathcal{L} = 0.430$ ($\mathcal{L} = 0.170$ for the fish template.

TABLE 1
FINAL OBJECTIVE FUNCTION ($\mathcal{L}$) VALUES WHEN THREE DIFFERENT TEMPLATES ARE APPLIED TO THREE DIFFERENT IMAGES. SMALL OBJECTIVE FUNCTION VALUES ARE PREFERRED.

template \ image	saxophone	fish	hand
saxophone	0.162	0.530	0.667
fish	0.647	0.194	0.462
hand	0.407	0.375	0.219

6 Discussions and Future Work

This work presents a systematic approach for object localization and retrieval using deformable templates. The deformation model consists of a hand-drawn prototype template (a bit map) and probabilistic transformations on it. This model is used as the a priori probability in a Bayesian scheme to reflect the prior knowledge of the shape. The likelihood function is based on the edge map of the input image. The object is localized by maximizing the a posteriori probability. The retrieval scheme is general since it can be applied to objects with arbitrary shapes as specified in the bitmap of the prototype template. As the global optimization is computationally expensive, we have used a multiresolution approach: deformed templates at a very low resolution are used to initiate more accurate matching at a higher resolution. This three-stage coarse-to-fine implementation retrieves objects in a short amount of time. Experimental results on real images show that this approach can correctly localize objects of varying shapes at different locations, orientation, and scale regardless of the number of occurrences in the image. The presence or absence of an object within an input image can also be determined by using our scheme.

The proposed multiresolution scheme of automatic object localization and matching is promising for large image database retrieval based on object shape. We are in the process of testing this approach on a large image database. In addition to the edge magnitude and direction which we have used in the likelihood function, we also plan to consider other attributes such as the interior and exterior inten-

sity, homogeneity, texture, and color to further speed up the retrieval.

REFERENCES

[1] Y. Amit and U. Grenander and M. Piccioni, "Structural image restoration through deformable template," *J. Am. Statistical Assn.*, vol. 86, no. 414, pp. 376-387, June 1991.

[2] D.H. Ballard, "Generalizing the Hough transform to detect arbitrary shapes, *Pattern Recognition*, vol. 13, no. 2, pp. 111-122, 1981.

[3] D.J. Burr, "Elastic matching of line drawings," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 3, no. 6, pp. 708-713, Nov. 1981.

[4] A. Chakraborty, L.H. Staib, and J.S. Duncan, "Deformable boundary finding influenced by region homogeneity," *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* (CVPR), Seattle, pp. 624-627, June 1994.

[5] Y.S. Chow, U. Grenander and D.M. Keenan, *HANDS. A Pattern-Theoretic Study Of Biological Shapes*. New York: Springer-Verlag, 1991.

[6] L.D. Cohen and I. Cohen, "Finite-element methods for active contour models and balloons for 2D and 3D images, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, pp. 1,131-1,147, Nov. 1993

[7] T.F. Cootes, C.J. Taylor, and A. Lanitis, "Active shape models: Evaluation of a multi-resolution method for improving image search," *Proc. British Machine Vision Conf.*," vol 1, pp. 327-336, 1994

[8] R.O. Duda and P.E. Hart, "Use of the Hough transforms to detect lines and curves in pictures," *Comm. ACM*, vol. 15, no. 1, pp. 11-15, 1972

[9] C. Faloutsos, R. Barber, M. Flicker, J. Hafner, W. Niblack, and W. Equitz, "Efficient and effective querying by image content," *J. Intell. Information Systems*," vol. 3, pp. 231-262, 1994.

[10] M. Figueiredo and J. Leitao, "Bayesian estimation of ventricular contours in angiographic images, *IEEE Trans. Medical Imaging*, vol. 11, no. 3, pp. 416-429, Sept. 1992

[11] S.B. Gelfand and S.K. Mitter, "Metropolis-type annealing algorithms for global optimization in $\mathcal{R}$," *Siam J. Control and Optimization*, vol. 31, no. 1, pp. 111-131, Jan. 1993.

[12] S. Geman, private communication, Apr. 1995.

[13] S. Geman, D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721-742, Nov. 1984.

[14] S. Geman, C. Huang, "Diffusions for global optimazation," *SIAM J. Control and Optimization*, vol. 24, pp 1,031-1,043, Sept. 1986.

[15] B. Gidas, "Global optimization via the Langevin equation," *Proc. IEEE Conf. Decision and Control*, Fort Lauderdale, Fla., pp. 774-778, 1985.

[16] M.M. Gorkani and R.W. Picard, "Texture orientation for sorting photos at a glance," *Proc. 12th Intl Conf. Pattern Recognition*, Jerusalem, vol. 67, no. 5, pp. A459-A464, Oct. 1994.

[17] U. Grenander, "Pattern synthesis: Lectures in pattern theory," *Applied Mathematical Sciences*. Springer-Verlag, vol. 18, 1976.

[18] U. Grenander and D.M. Keenan, "Towards automated image understanding, *Advances in Applied Statistics: Statistics and Images* vol. 1," K.V. Mardia and G.K. Kanji, eds., Chap. 6, pp. 89-103. Carfax Publishing Company, 1993.

[19] U. Grenander and M. I. Miller, "Representations of knowledge in complex systems," *J. Royal Statistical Society*, vol. 56, no. 3, pp. 1-33, 1994.

[20] B. Holt and L. Hartwick, "Visual image retrieval for applications in art and art history," *Proc. SPIE*, vol. 2,185, pp. 70-81, Feb. 1994.

[21] P.V.C. Hough, "Method and means for recognizing complex patterns," U.S. Patent 3069654, 1962.

[22] J. Illingworth and J. Kitter, "A survey of Hough transform," CVGIP, vol. 44, pp. 87-116, 1988.

[23] M. Kass, A. Witkin and D. Terzopoulos, "Snakes: active contour models, IJCV, vol. 1, no. 4, pp. 321-331, 1988.

[24] S. Lakshmanan and H. Grimmer, "Detecting straight edges in radar images using deformable templates," to appear, *IEEE Trans. Pattern Analysis and Machine Intelligence*.

[25] V. F. Leavers, "Survey: Which Hough transform?," *CVGIP: Image Understanding*, vol. 58, no. 2, pp. 250-264, Sept. 1993.

[26] 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, no. 6, pp. 617-634, June 1993.

[27] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Reller, and E. Teller, "Equations of state calculations by fast computing machines," *J. Chem. Phys.*, vol. 21, pp. 1,087-1,092, 1953.

[28] M.I. Miller, G.E. Christensen, Y. Amit, and U. Grenander, "Mathematical textbook of deformable neuroanatomies," *Proc. Natl. Academy of Science*, vol. 90, pp. 11,944-11,948, Dec. 1993.

[29] M. Moshfeghi, S. Ranganath, and K. Nawyn, "Three-dimensional elastic matching of volumes," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 3, no. 2, pp. 128-138, Mar. 1994.

[30] L.H. Staib and J.S. Duncan, "Boundary finding with parametrically deformable models," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 14, no. 11, pp. 1,061-1,075, Nov. 1992.

[31] G. Storvik, "A Bayesian approach to dynamic contours through stochastic sampling and simulated annealing," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 16, no. 10, pp. 976-986, Oct. 1994

[32] R. Szeliski and J. Coughlan, "Hierarchical spline-based registration," *Proc. IEEE Conf. Computer Vision and Pattern Recognition* (CVPR), "Seattle, pp. 194-201, June 1994.

[33] D. Terzopolous, A. Witkin, and M. Kass, "Constraints on deformable models: Recovering 3D shape and nonrigid motion, *AI*, no. 36, pp. 91-123, 1988.

[34] A. Witkin, D. Terzopoulos, and M. Kass, "Signal matching through scale space," *Int'l J. Computer Vision*, pp. 133-144, 1987.

[35] A.L. Yuille, P.W. Hallinan and D.S. Cohen, "Feature extraction from faces using deformable templates," *Int'l J. Computer Vision*, vol. 8, no. 2, 133-144, 1992.

Anil Jain is a University Distinguished Professor and chairman of the Department of Computer Science at Michigan State University. His current research interests are computer vision, image processing, neural networks, and pattern recognition. He has made significant contributions and published a large number of papers on the following topics: statistical pattern recognition, exploratory pattern analysis, neural networks, Markov random fields, texture analysis, interpretation of range images, and 3D object recognition. Several of his papers have been reprinted in edited volumes on image processing and pattern recognition. He received the best paper awards in 1987 and 1991, and received certificates for outstanding contributions in 1976, 1979 and 1992 from the Pattern Recognition Society. Dr. Jain served as the editor-in-chief of the *IEEE Transactions on Pattern Analysis and Machine Intelligence* (1991-1994), and currently serves on the editorial boards of *Pattern Recognition Journal*, *Pattern Recognition Letters*, *Journal of Mathematical Imaging*, *Journal of Applied Intelligence*, and *IEEE Transactions on Neural Networks*. He is the co-author of the book *Algorithms for Clustering Data* (Prentice-Hall, 1988), has edited the book *Real-Time Object Measurement and Classification* (Springer-Verlag, 1988), and has co-edited the books, *Analysis and Interpretation of Range Images* (Springer-Verlag, 1989), *Neural Networks and Statistical Pattern Recognition* (North-Holland, 1991), *Markov Random Fields: Theory and Applications* (Academic Press, 1993), and *3D Object Recognition* (Elsevier, 1993). Dr. Jain is a Fellow of the IEEE. He was the Co-General Chairman of the 11th International Conf. on Pattern Recognition, Hague (1992), General Chairman of the IEEE Workshop on Interpretation of 3D Scenes, Austin (1989), Director of the NATO Advanced Research Workshop on Real-time Object Measurement and Classification, Maratea (1987), and codirected National Science Foundation supported workshops on Future Research Directions in Computer Vision, Maui (1991), Theory and Applications of Markov Random Fields, San Diego (1989), and Range Image Understanding, East Lansing (1988). Dr. Jain was a member of the IEEE Publications Board (1988-90) and served as the distinguished visitor of the IEEE Computer Society (1988-90). He is currently a distinguished lecturer of the IEEE Computer Society's Asia-Pacific Lectureship Program.

Yu Zhong received the BS and MS degrees in computer science and engineering from Zhejiang University, Hangzhou, China in 1988 and 1991, and the MS degree in statistics from Simon Fraser University, Burnaby, Canada, in 1993. She is currently a doctoral student at Michigan State University. Her research interests include image processing and machine vision.

Sridhar Lakshmanan received his BS degree in electronics and communications engineering from the Birla Institute of Technology, Mesra (India) in 1985, and his MS and PhD in electrical and computer engineering from the University of Massachusetts (Amherst) in 1987 and 1991, respectively. Since 1991 he has been an Assistant Professor in the Electrical and Computer Engineering Department at the University of Michigan–Dearborn. Presently, he is also an adjunct assistant professor in the Department of Computer Science at Michigan State University. His research interests include object detection and identification, and image modeling.