An Integrated Approach to the Segmentation and Recognition of Objects using Thin Plate Spline Method

Xun Wang[*] (xun_wang_2001@yahoo.com), Feng Gao, Zhigang Peng, Lei He, William Wee
ECECS Dept., University of Cincinnati, Cincinnati, OH 45220, U.S.A.

Abstract

A novel approach for object segmentation and recognition is described. The aim of the approach is to select a proper shape model from a model set to guide object segmentation. The process of model selection, which is based on the shape similarity between the target object and shape models, is then used for object recognitions. The integrated process of object segmentation and recognition is formulated as a constrained contour energy minimization problem. The solution derived from this formulation produces an integrated searching process consisting of two iteratively alternating procedures of contour evolution and shape matching. The process stops at a final contour together with a shape distance measure to an object model for recognition. Successful illustrative results on both segmentation and recognition are reported.

1. Introductions

Object segmentation is an important, yet unfortunately, difficult issue in computer vision. Various challenges in object segmentation often require that segmentation methods take advantages of comprehensive information on the features of object boundary, interiors, and shape. Trying to integrate these features, model based segmentation methods [3] [4] are proposed and are rather successful in numerous applications. However, a limitation for these methods is that they often require a condition that a shape model [4] [16] that can well represent the shape of target object is given. The construction of a shape model that can well represent the target object often needs extensive training and is often unavailable in many situations. In the paper, we are trying to overcome the limitation by concerning with object segmentation using a model set, one of which represents the shape similar to that of target object. The aim of the method is to select a proper shape model from the given model set and use it to guide object segmentations. Since shape model selection is based on the shape similarity between objects and shape models, the method can be conveniently used for object recognition. In the following, we will list a few scenarios that illustrate the motivation of the proposed approach.

Consider we are segmenting a rather large number of objects, the target objects can be from different categories. (For example, the target objects can be bananas, apples, or cucumbers). Though we have the object shape models, we have to select the right model for every object segmentation application, which often requires human intervention. It is most desired that the method can recognize the object and select the corresponding shape model to guide segmentation.

The method also provides an alternative to other object recognition techniques [5] [6] [7]. Though object recognition techniques are rather effective in recognizing a large number of objects, they may have difficulties in *accurately* recognizing objects with reflectance, shadings, occlusions, and noises [7]. Since the approach is more robust to most segmentation difficulties, it has a potential of improving the recognition accuracy in these cases.

The paper formulates the integration of object segmentation and recognition as a constrained optimization problem. The contour energy function includes the snake energy function and the shape matching distance function between the contour and an object shape model using the thin plate spline method [2] [14], and the constraint includes contour interior features. Thin plate spline method provides a flexible class of coordinate transformations [2] [14] in shape matching and recognition. It needs only a set of corresponding contour points from the model shape and a target contour to establish the transformation coefficients. The point correspondence is established by shape context method [2], which is invariant under affine transform and robust to noises, missing parts, and local deformation. The solution derived from this formulation produces an integrated searching process consisting of two iteratively alternating procedures of contour evolution and shape matching. The process stops at a final contour together with a shape distance measure to an object model for recognition.

2. Related Works

In [16], an original parametric deformable model is proposed. The method applies a flexible shape constraint, in the form of probabilistic deformable model, to object segmentation, which is effective to represent natural objects with irregularity and diversity. This concept is extensively applied in later literatures [3] [4] [10] [11] [12]. [16] is further improved by incorporating more image features using game theoretical approach in [4] . The methods [4] are robust to noises, gaps, and are rather successful in many applications. However, they often have difficulties in dealing with poor initialization.

In an effort of dealing with poor initialization, the concepts of probabilistic deformable models are applied in the formu lations of level set [5]. In [10] [11] [12], object shape models are incorporated in level set methods. Typically, the methods are first to construct shape models by generating a distance map to characterize shape statistics, and then incorporate the constructed shape model into an objective function for segmentation. Level set based curve evolution techniques are then derived to minimize the objective functions.

In summary, model based object segmentation methods produces more favorable results in dealing with many segmentation difficulties including gaps, noises, and object interior inhomogenity than free form object segmentation methods [1] [9] if the shape model can well represent the shape of target object. However, in these methods, complicated model constructions are often needed for specific object segmentation applications. It is most desired that a common shape model set or library can be used in an extensive range of segmentation applications. Here, we note that deformable models have been proposed for the integrated problem of object segmentation and recognition [18]. However, these works are more suitable for the applications of document image segmentation and recognition and are unfavorable for segmenting and recognizing objects with shadings, occlusions, inhomogeneity, and reflectance.

3. Overview of Approach

The problem of object segmentation and recognition is composed of two components: 1) shape model set construction 2) shape model selection (object recognition) and object segmentation based on the selected shape model.

To build shape model set, we select an ideal object example as the shape model to simplify the process of shape model constructions. A shape similarity measure between the shape model and a given contour can then be defined by finding the corresponding point sets from the two contours and evaluating the matching distances of these two point sets using thin plate spline method [2]. The shape similarity measure is to be used for the later object segmentations.

When the shape model set is built, a model based object segmentation method guided by shape model is then applied for each model. The method is a model based deformable contour within a framework of constrained contour energy minimization [9]. The solution derived from the method produces two alternating processes of contour evolution and shape matching. The contour evolution is generally driven by the information of image gradient, object interior features, and estimated shape model. The contour shape matching dynamically combines the current contour with the shape model to generate an estimated shape model contour, which is then used in contour evolution. The algorithm stops at a resulting contour together with a distance measure to the selected model contour for object recognition. We select the resulting contour with the smallest distance to one of the shape models as the final result and the shape model as what target object belongs. The philosophy can be explained as “shape model competition”, in which each shape model in the model set competes to guide the object segmentation or shape formation, and the shape model that wins the competition, which is evaluated by the distance to the resulting contour, is declared as the winner, and its associated contour is the resulting contour.

4. Shape Model and a Given Contour Matching Using Thin Plate Spline Method

In this section, a shape matching distance measure is established between a close contour and a model close contour through a mapping for both continuous and discrete formulations. The coordinate transformations are derived to minimize the shape matching distance by using thin plate spline method [2][14]. The shape matching distance measure will be included in the energy minimization formulation.

Let $\Gamma(x(q), y(q))$ be a close contour in a Cartesian coordinate system C_t with x and y denoting x- and y-coordinates, where $q \in [0,1)$ is contour parameter and $(x(q), y(q))$ are contour points on $\Gamma(x(q), y(q))$. Let $M(u(p), v(p))$ be a model close contour in another Cartesian coordinate system C_m with u and v denoting u- and v-coordinates, where $p \in [0,1)$ is also the contour parameter and $(u(p), v(p)) \in R^2$ are contour points of $M(u(p), v(p))$. R^2 is a two dimensional open domain. It is easy to see that there exists an invertible continuous function $p = c(q)$ that maps point $(x(q), y(q))$ to its corresponding point $(u(p), v(p))$ on $M(u(p), v(p))$. Thus $M(u(p), v(p))$ can be denoted as $M(u(c(q)), v(c(q)))$. For simplicity, we will denote respectively $\Gamma(x(q), y(q))$, $M(u(c(q)), v(c(q)))$, and $(u(c(q)), v(c(q))$ as $\Gamma(q)$, $M(q)$, and $(u(q), v(q))$, respectively, from now on.

To match $\Gamma(q)$ and $M(q)$, we have to determine coordinate transformations $f_x(u(q),v(q))$ and $f_y(u(q),v(q))$ from $M(q)$ to $\Gamma(q)$, such that

$$e_{cs}(\Gamma(q), M(q), f_x, f_y) = \int_0^1 \mathrm{dist}^2((x(q), y(q)), \quad (1)$$
$$(f_x(u(q),v(q)), f_y(u(q),v(q))dq$$

is minimized, where

$$\mathrm{dist}^2((x(q), y(q)), (f_x(u(q),v(q)), f_y(u(q),v(q)))$$
$$= (x(q) - f_x(u(q),v(q)))^2 + (y(q) - f_y(u(q),v(q)))^2$$

In the paper, we use thin plate spline method [2], which is commonly used for flexible coordinate transformations, to minimize Eq. (1). Using thin plate spline method is: to first select two point sets from $M(q)$ and $\Gamma(q)$, and then evaluate the transformation f_x and f_y by minimizing the matching distance between the two point sets. Specifically, let

$$P_\Gamma = \left\{(x(q_i), y(q_i)) : q_i = \frac{i}{N}, i = 1,2 \cdots N\right\}$$

be an equally spaced point sequence set on $\Gamma(q)$ in a clockwise direction and $V_M = \{(u(q_i), v(q_i)) : i = 1,2,\cdots N\}$ be a point sequence set on $M(q)$ with each point $(u(q_i), v(q_i))$ corresponding to $(x(q_i), y(q_i))$. The corresponding point matching method is the shape context matching method described in Section 3.1 of [2].

We approximate Eq. (1) by $e_{ap}(P_\Gamma, V_M, f_x, f_y)$ defined as,

$$e_{ap}(P_\Gamma, V_M, f_x, f_y) =$$
$$\frac{1}{N}\sum_{i=1}^{N}\left\{\left[x(q_i) - f_x(u(q_i),v(q_i))\right]^2 + \left[y(q_i) - f_y(u(q_i),v(q_i))\right]^2\right\}. \quad (2)$$

Using thin plate spline method is to minimize

$$e_{ps}(P_\Gamma, V_M, f_x, f_y) = e_{ap}(P_\Gamma, V_M, f_x, f_y) + l(I_f^x + I_f^y) \quad (3)$$

where

$$f_x(u,v) = a_1^u + a_u^u u + a_u^v v + \sum_{i=1}^{N} w_i^u U(\|(u(q_i), v(q_i)) - (u,v)\|^2),$$

and

$$f_y(u,v) = a_1^v + a_v^u u + a_v^v v + \sum_{i=1}^{N} w_i^v U(\|(u(q_i), v(q_i)) - (u,v)\|^2)$$

are coordinate transformations that match V_M to P_Γ with

$$\sum_{i=1}^{N} w_i^u = 0, \sum_{i=1}^{N} w_i^u u(q_i) = \sum_{i=1}^{N} w_i^u v(q_i) = 0,$$
$$\sum_{i=1}^{N} w_i^v = 0, \text{ and } \sum_{i=1}^{N} w_i^v u(q_i) = \sum_{i=1}^{N} w_i^v v(q_i) = 0 \cdot$$

$a_1^u, a_u^u, a_u^v, a_1^v, a_v^u, a_v^v, w_i^u, w_i^v$ are coefficients to be determined.

$U(r)$ is defined as $U(r) = r^2 \log r^2$, and $U(0) = 0$.

$$I_f^x = \iint_{R^2}\left[\left(\frac{\partial^2 f_x}{\partial u^2}\right)^2 + \left(\frac{\partial^2 f_x}{\partial u \partial v}\right)^2 + \left(\frac{\partial^2 f_y}{\partial u^2}\right)^2\right] dudv,$$

and $I_f^y = \iint_{R^2}\left[\left(\frac{\partial^2 f_y}{\partial u^2}\right)^2 + \left(\frac{\partial^2 f_y}{\partial u \partial v}\right)^2 + \left(\frac{\partial^2 f_y}{\partial v^2}\right)^2\right] dudv$

are the bending energies, and $l > 0$ is the regularization parameter controlling the smoothing of the transformations of $f_x(u,v)$ and $f_y(u,v)$ [2] [14]. Eq. (3) can be minimized by solving a set of linear equations to be illustrated in the later section. According to [2] [14], bending energies an efficient measure indicating the shape similarity and in this paper, the total value of bending energies is used as an index for object recognitions.

5. Model Based Object Segmentation

In this section, model based object segmentation method is derived and applied for each shape model in the model set. Here, for ease of presentation, only a single model $M^g(q)$ is considered. As discussed in Section 4, the shape matching distance is used to define a distance between an input contour and model contour. This shape matching distance along with the features of object interior sand boundaries is integrated in a constrained contour energy minimization formulation for object segmentation. The bending energy of the shape matching is then used as the object recognition decision mechanism for selecting the class belonging of an extracted contour to one of the model set.

As shown in Fig. 5.1, $C(q,t)$ is a deforming close contour with interior $\Omega_c(t)$ at time t. Our problem then is to find $C(q,t)$ such that

$$E(C(q,t), M^g(q), f_x, f_y) = w \oint_{C(q,t)} g(|\nabla I(C(q,t))|)ds$$
$$+ (1-w)e_{cs}(C(q,t), M^g(q), f_x, f_y),$$
$$0 < w < 1 \quad (4)$$

is minimized subject to the region constraint [9] [13],

$$D(x,y) \geq T_V \quad \text{if } (x,y) \in \Omega_c(t) \quad (5)$$

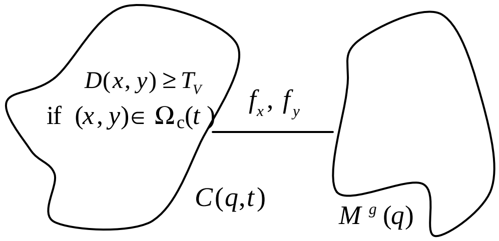

Fig. 5.1 An illustration of Eq. (4) and Eq. (5)

where s is the arc length, $I(x,y)$ is the image brightness at (x,y) and

$$g(|\nabla I(C(q,t))|) = \frac{1}{1+|\nabla I(C(q,t))|^2}.$$

$\nabla I(C(q,t))$ is the gradient of $I(x,y)$ with (x,y) on $C(q,t)$; T_V is a positive threshold. $D(x,y)$ can be any function characterizing the interior of expected target contour. As a special case, let

$$D(x,y) = \frac{1}{1+|\nabla G * I|^2} e^{-\frac{|I(x,y)-I_0|}{s}}$$

where $|\nabla G * I|$ is the absolute value of the gradient of $I(x, y)$ smoothed by a Gaussian filter $N(0, \sigma_1^2)$ and I_0 is the average intensity over target boundary interior Ω_T. Notice that the first term on the right side of Eq. (4) is the snake energy based on the image gradients, and the second term $e_{cs}(\cdot)$ is the total value of the matching squared distances between the evolving contour $C(q,t)$ and the model contour $M^g(q)$.

Using Lagrange approach, we have

$$\begin{aligned} L(C(q,t),M^g(q),f_x,f_y) &= w \oint_{C(q,t)} g(|\nabla I(C(q,t))|)ds \\ &+ (1-w)e_{cs}(C(q,t),M^g(q),f_x,f_y) \\ &- l_1 \iint_{\Omega_C} [D(x,y)-T_V]dxdy \end{aligned} \qquad (6)$$

where $l_1 > 0$ is a constant multiplier.

To minimize Eq. (6), our first step (Step A) is to deform $C(q,t)$ while keeping f_x and f_y unchanged. It can be shown that minimizing Eq. (6) is equivalent to minimizing

$$\begin{aligned} L(C(q,t),M^g(q),f_x,f_y) &= w \oint_{C(q,t)} g(|\nabla I(C(q,t))|)ds \\ &- (1-w)\iint_{\Omega_C} D_f(x,y)dxdy - l_1 \iint_{\Omega_C} [D(x,y)-T_V]dxdy \end{aligned} \qquad (7)$$

where $D_f(x,y) =$

$$\begin{cases} \text{dist}^2((x,y),T(M^g)) & \text{if } (x,y) \text{ is inside} T(M^g) \quad , \\ 0 & \text{if } (x,y) \in T(M^g) \\ -\text{dist}^2((x,y),T(M^g)) & \text{if } (x,y) \text{ is outside} T(M^g) \end{cases}$$

$$T(M^g) = \{(f_x(u(q),v(q)), f_y(u(q),v(q))) : (u(q),v(q)) \in M^g(q)\},$$

and $\text{dist}((x(q),y(q)),T(M^g))$ is the shortest distance from point $(x(q),y(q))$ to a point in $T(M^g)$.

Pattern after the derivation procedure of [13], the curve evolution formu la corresponding to Eq. (7) is

$$\begin{aligned} \frac{\partial C(q,t)}{\partial t} &= w[kg(|\nabla I|) - (\nabla g \cdot \vec{N})]\vec{N} \\ &+ [(1-w)D_f(x,y) + l_1[D(x,y)-T_V]]\vec{N} \end{aligned} \qquad (8)$$

Our next step (Step B) is to keep $C(q,t)$ constant, and then to minimize Eq. (6) or equivalently minimize e_{cs} by adjusting f_x and f_y. To facilitate our minimization process, the approximated discrete form e_{ps} is used as

$$e_{ps}(P_C,V_M,f_x,f_y) = l(I_f^x + I_f^y) + \frac{1}{N}\sum_{i=1}^{N}\{[x(q_i) - f_x(u(q_i),v(q_i))]^2 + [y(q_i) - f_y(u(q_i),v(q_{i)})]^2\} \qquad (9)$$

where $P_C = \{(x(q_i),y(q_i)) : i = 1,2,\cdots,N\}$ is an equally spaced point sequence set on $C(q,t)$ in clockwise direction, and V_M is the corresponding point sequence set as defined earlier.

According to [14], minimizing Eq. (9) is to solve,

$$\begin{cases} \begin{bmatrix} K' & P \\ P^T & 0 \end{bmatrix} \begin{bmatrix} W_u \\ A_u \end{bmatrix} = \begin{bmatrix} X \\ 0 \end{bmatrix} \\ \begin{bmatrix} K' & P \\ P^T & 0 \end{bmatrix} \begin{bmatrix} W_v \\ A_v \end{bmatrix} = \begin{bmatrix} Y \\ 0 \end{bmatrix} \end{cases} \qquad (10)$$

where $W_u = [w_1^u, w_2^u, \cdots w_N^u]^T$, $W_v = [w_1^v, w_2^v, \cdots w_N^v]^T$

$X = [x(q_1), x(q_2), \cdots x(q_N)]^T$,

$Y = [y(q_1), y(q_2), \cdots y(q_N)]^T$,

$A_u = [a_1^u, a_u^u, a_u^v]^T$, and $A_v = [a_1^v, a_v^u, a_v^v]^T$.

The ith row of P is $(1,u(q_i),v(q_i))$. K is a matrix composed of components,

$$K_{ij} = U(\|(u(q_i),v(q_i)) - (u(q_j),v(q_j))\|), \qquad (11)$$

$$K' = K + NlI_d,$$

where I_d is an identity matrix. We then solve Eq. (10) using a regular linear algebra method. The resulting

W_u, W_v, A_u, and A_v are then used to form new f_x and f_y transformations employed in Eq. (7) and Eq. (8) to guide contour deformations.

Discussion:

It is easy to see that the minimization of $L(C(q,t), M^g(q), f_x, f_y)$ is composed of two steps. The first step is the deformation of $C(q,t)$ driven by the curve evolution formula of Eq. (8). In this deformation, the contour point velocity is determined by the combined effect of region features $[D(x,y)-T_V]$, contour gradient function $\{wkg(|\nabla I|)-w(\nabla g\cdot\vec{N})\}$, and shape matching distance $\lfloor D_f(x,y)-T_f \rfloor$ between $C(q,t)$ and $M^g(q)$. The second step is the recalculation of f_x and f_y using Eq. (10). In here, the resulting f_x and f_y are to produce a new $D_f(x,y)$. In applications, the problem of the above solution is that it is easy to be trapped into local minima. More specifically, the computation of Eq. (10) can only accurately estimating f_x and f_y to guide contour evolution when the deforming contour is rather close to the target boundary. Since in the initial stage, the deforming contour $C(q,t)$ is rather arbitrarily placed inside the target object, the transformations f_x and f_y computed by matching the equally spaced point sets on $C(q,t)$ and shape model is unreliable in predicting the actual target object boundary to guide contour evolution.

To handle this problem, we apply Mean Field Annealing approach (See MFA based deformable contour method [17]) to parameter l_1 of Eq. (7) and set l_1 as a monotonically decreasing function of time.

$$l_1 = l_0 - c_1\sqrt{\frac{t}{p}} \qquad (12)$$

where l_0 and c_1 are positive constant.

With the introduction of Eq. (12), the contour deformation is mainly driven by the object interior features when the contour is small and outgrowing. As the contour approaches target boundary, the contour deformation driven by region features gradually decays and it is more and more affected by shape features. The whole contour deformation process can be viewed as a combination of a free form contour deformation driven by the image features and that driven by a parametric deformable model.

With the modification, the method gains improvements in overcoming local minima and segmentation difficulties of gaps and noises. However, it still have difficulties in segmenting objects with occlusions and heterogeneity, where the deforming contour mainly driven by object interior features is unable to approach a location nearby the target boundary. Here, only segments of target boundary instead of the entire target boundary can be approximately reached.

To deal with this problem, we first roughly estimate the size of the target object by a pre-segmentation method, such as K-mean, or by human intervention (proportional to the parameter, total iteration number N_T. see Section 7). We normalize the selected shape model to the estimated size and modify Step B of computing Eq. (10) as: iteratively search for a best match between the deforming contour segment and the model segment using shape context [2]. Select equally spaced point sequence set on the matched contour segments in clockwise direction and compute Eq. (10) for f_x and f_y. It should be pointed out that though in the method, only shape context [2] is used for matching contour segments, other shape matching method such as [8] can also be to be employed for the same purpose.

When the resulting contour is reached, a shape distances between the resulting contour and the selected contour is computed according to the bending energy E_b of the shape matching with,

$$E_b = C_k(W_u K W_u + W_v K W_v), \qquad (13)$$

The smaller E_b is, the higher shape similarity is. C_k is a constant and is set at 0.001 in our applications. The recognition decision is then based on having the minimum E_b among a class of models.

6. Algorithm Description

The algorithm can be illustrated as the follows:

(a) Select shape model $M^g(q), g=1$ from the model set $M^j(q), j=1,2,\cdots n$. Select an initial small seed region j (usually with size of 3 by 3 pixels) in Ω_Γ, the interior of the object.

(b) With $C(q,0)$ being the boundary of j at $t=0$, compute the mean brightness I_0 of j. Set $D_f(x,y)=0$.

(c) Obtain $V_M=\{(u(q_i),v(q_i)) : i=1,2,\cdots N\}$ of N equally spaced point sequence set of $M^g(q)$ in a clockwise manner.

(d) Evolve $C(q,t)$ according to Eq. (8) using the level set narrow band algorithm [5] for l iterations. If the maximum point velocity of $C(q,t)$ is less than a velocity threshold value h, or the total number of iterations N_T is reached, then stop the algorithm, compute the bending energy of the resulting contour and $M^g(q)$ according to Eq. (13) and go to Step (h). If the total number of contour points is less than N, repeat step (d).

(e) Link contour $C(q,t)$. Form $P_C = \{(x(q_i), y(q_i)) : i = 1,2,\cdots N\}$ an approximately equally spaced point sequence set of $C(q,t)$ in a clockwise manner. Select the best match point subsets in between P_C and V_M using the shape context method described in Section 3.1 of [2].

(f) Update l_1 according to Eq. (12), compute f_x, f_y according to Eq. (11) and $D_f(x,y)$ according to Eq. (7). Go to step (d).

(g) Save $C(q,t)$ as the resulting contour using shape model $M^g(q)$. Increase g by 1. If $g > N$, select the resulting contour with lowest bending energy as the segmentation results and the associated model class the recognition results.

Remark: The proposed algorithm has over all 7 parameters: velocity threshold h, total number of iteration N_T, factor C_k (Eq. (13)), l (iteration number per f_x and f_y updating), weight w (Eq. (8)), weight factors l_0 and c_1 (Eq. (12)), and multiplier l (Eq. (11)). Velocity threshold h and maximum iteration number N_T are common parameters for level set algorithms [9] and in all the following experiments, we keep h constant as 0.001. N_T is usually set at the range of 800~2000. As indicated above, C_k is set constant as 0.001. The iteration number l is set as 100. In most applications, w is set at 0.005 to emphasize the weight on the shape matching distance. l_1 is set as 1000. We set l as 5000.

7. Experiments

In this section, we would like to apply the derived algorithm to demonstrate both segmentation and recognition on Fig. 7.1(a) consisting of two apples, one banana, one cucumber, and one pear on a 144 by 108 image obtained from a web-site http://cs-www.bu.edu/groups/ivc/data.html from [15]. We first use the apple model on the derived algorithm discussed in Section 6 with an initial seed region inside each fruit object. The algorithm runs 5 times, one on each fruit. The resulting five contours are shown in Fig. 7.1(b), and the resulting E_b values are in column one of Table 1. As one can see that both Apple I and Apple II contours are rather nicely segmented even with Apple II touching the cucumber. Without the apple model, one will have some difficulty separating the two. The experiment is repeated for the banana model, the cucumber model and the pear model with the resulting contours shown in Figs. 7.1(c), 7.1(d) and 7.1(e), and E_b values in columns 2, 3 and 4, of Table 1, respectively. The banana in Fig. 7.1(c) has the best segmentation result noticing the dark banana skin cover on the upper right side of the banana. The cucumber in Fig. 7.1(d) has the best segmentation among all other models in Figs. 7.1(b), 7.1(c) and Fig. 7.1(e). Similar result can be reported on the pear in Fig. 7.1(e). Our recognition decision is based on the minimum E_b value on each object in using different model assumptions as shown in Table 1, i.e., selecting the row minimum value as indicated by (*). All are successfully recognized.

The algorithm has then been applied on Fig. 7.2(a) with a green pepper and a cucumber with Gaussian noises of variance 1000. With the noises, the touching parts of the cucumber and green pepper are rather unclear and besides the noises there are inhomogeneity inside cucumber and green pepper making the segmentation rather difficult. The algorithm overcomes the difficulties with the results demonstrated in Fig. 7.2(b).

Fig. 7.3 illustrates the process of the method extracting object boundary with occlusion. Fig. 7.3(a) is an image of size of 300 by 400. The aim is to segment the banana on the right side. As we can see from Fig. 7.3(b), 7.3(c), the method firstly segments the lower part of the banana. By correctly matching the segment to the shape model, the method is able to overcome the occlusions and reach the final result as illustrated in Fig. 7.3(d).

The algorithm has also been successfully applied on Fig. 7.4(a), Fig. 7.5(a), Fig. 7.6(a), and Fig. 7.7(a). Fig. 7.4(a) is an image of size 300 by 400 consisting of two bananas with partially occluded by a cucumber, Fig. 7.5(a) and Fig. 7.6(a) are images of size of 104 by 188 consisting of bananas and pear. Fig. 7.7(a) (c) are images of size of 240 by 320 consisting of multiple objects including cucumbers, egg plant, green and red peppers, and carrots. In all the images, reflection, shadow, and occlusions make the segmentations most difficult. With the help of the models, we successfully segmented all the objects as shown in Fig. 7.4(b), Fig. 7.5(b), Fig. 7.6(b), and Fig. 7.7(b) and Fig. 7.7(d). As we can see from the results, rather successful results are demonstrated.

8. Conclusion

We have introduced a framework of deformable contour methods that integrates edge, shape model, and object interior structure information into one unified optimization formulation for object segmentation and recognition. The framework can conveniently incorporate a very general class of region features including color and texture. The resulting algorithm is more robust to many segmentation difficulties including occlusion, gaps, and noises. It is also very straightforward to apply object interior features together with shape information for object recognition. Successful illustrative results on both segmentation and recognition, and occlusion recovery have been achieved using the derived algorithm.

Reference

[1] M. Kass, A. Witkin and D. Terzopoulos, "Snakes: Active Contour Models," *International Journal of Computer Vision*, Vol. 1, No. 4, pp. 321-331, 1988.

[2] S. Belongie, J. Malik, and J. Puzicha, "Shape Matching and Object Recognition Using Shape Contexts" *IEEE Trans. on PAMI*, Vol. 24, No. 4, pp. 509 -522, 2002.

[3] Y. Wang and L. Staib, "Elastic Model Based Non-Rigid Registration Incorporating Statistical Shape Information," pp. 1162-1173, *MICCAI* 1998.

[4] A. Chakraborty, and J. Duncan, "Game-Theoretic Integration for Image Segmentation" *IEEE Trans. On PAMI*, Vol. 21 No. 1, pp. 12 -30, Jan. 1999.

[5] R. Malladi, J. Sethian and B. Vemuri, "Shape Modeling with Front Propagation", *IEEE Trans. on PAMI*, Vol. 17, No.2, pp. 158-171, Feb. 1995.

[6] M. Blane, Z. Lei, H. Civi, and D. Cooper, "The 3L Algorithm for Fitting Implicit Polynomial Curves and Surfaces to Data" *IEEE Trans. on PAMI*, Vol. 22, No. 3, pp. 298 -313. Mar. 2000.

[7] T. Sebastian, P. Klein, and B. Kimia, "Recognition of Shapes by Editing Shock Graph,", pp. 755-762, *IEEE ICCV* 2001.

[8] E. Petrakis, A. Diplaros, and E. Milios, "Matching and retrieval of distorted and occluded shapes using dynamic programming" Volume: 24, Issue: 11, Nov. 2002, pp. 1501-1516, *IEEE Trans. PAMI*, 2002.

[9] X. Wang, L. He, and W. Wee, "Constrained Optimization: A Geodesic Approach", vol. II, pp.77-80, *IEEE International Conference of Image Processing*, 2002.

[10] D. Cremers, T. Kohlberger, and C. Schnorr, "Nonlinear Shape Statistics in Mumford-Shah Based Segmentation," pp. 93-109, *ECCV 2002.*

[11] A. Tsai, A. Yezzi, W. Wells, C. Tempany, D. Tucker, A. Fan, A. Grimson, and A. Willsky, "Model Badrf Curve Evolution Technique for Image Segmentation," pp. 463-468. *IEEE CVPR* 2001.

[12] N. Paragios and M. Rousson, "Shape Priors for Level Set Representations", *Proceedings of European Conference in Computer Vision*, (*ECCV* 2002) June 2002.

[13] X. Wang, L. He, and W. Wee, "A Constrained Optimization Approach to Deformable Contour Method", pp. 183-192, British Machine Vision Conference, 2002.

[14] F. Bookstein, "Principal Warps: Thin-Plate Splines and the Decomposition of Deformations", *IEEE Trans. on PAMI*, Vol. 11, No. 6, pp. 567-585, June 1989.

[15] S. Sclaroff and L. Liu, "Deformable Shape Detection and Description via Model-Based Region Grouping", *IEEE Trans. on PAMI*, Vol. 23, No. 5, pp. 475-489, May 2001.

[16] L. Staib and J. Duncan, "Boundary Finding with Parametrically Deformable Models," IEEE Transactions on Pattern Analysis and Machine Intelligence, 14(11), 1061 (1992).

[17] X. Wang, F. Gao, Z. Peng, L. He, W. Wee, "A Mean Field Annealing Based Deformable Contour Method", Submitted to *IEEE ICCV* 2003.

[18] K.W. Cheung, D.Y. Yeung, R.T. Chin. On deformable models for visual pattern recognition. *Pattern Recognition*, 35(7):1507-1526, July 2002.

Model / Object	Bending Energy			
	Apple	Banana	Cucumber	Pear
Apple I	**0.627***	1.361	1.276	1.773
Apple II	**1.699***	2.381	2.897	2.483
Banana	4.88	**0.498***	0.700	0.557
Cucumber	4.605	6.737	**1.919***	3.477
Pear	2.948	2.224	2.349	**1.954***

Table 1. Resulting Performance Matrix

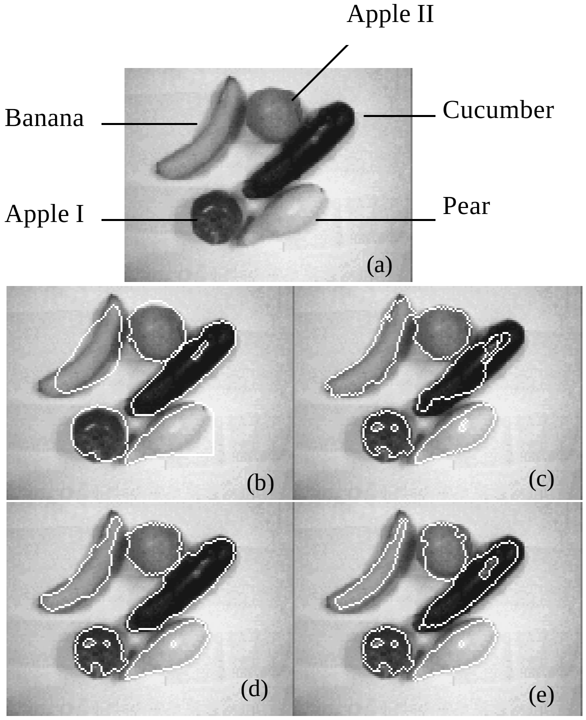

Fig. 7.1 (a) original image (b) results using the apple model (c) results using the banana model (d) results using the cucumber model (e) results using the pear model

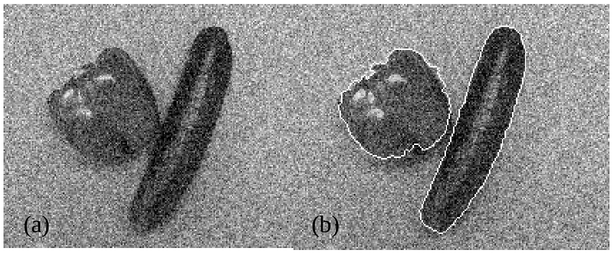

Fig. 7.2(a) Original image under Gaussian Noise (b) resulting contours

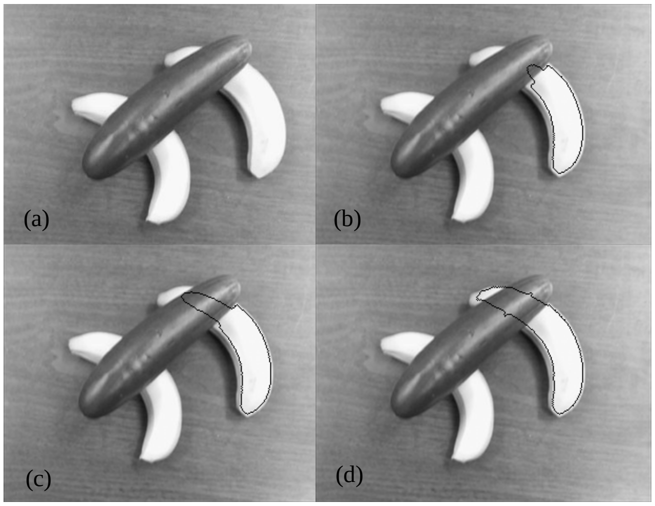

Fig. 7.3 (a) Original Image (b) resulting contour in iteration 150 (c) resulting contour in iteration 250 (d) final result

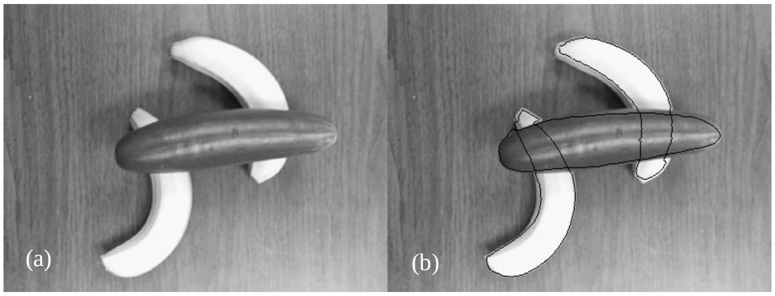

Fig. 7.4 (a) Original image (b) Resulting contours

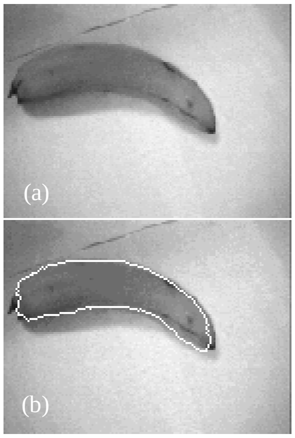

Fig. 7.5 (a) Original Image (b) resulting contour

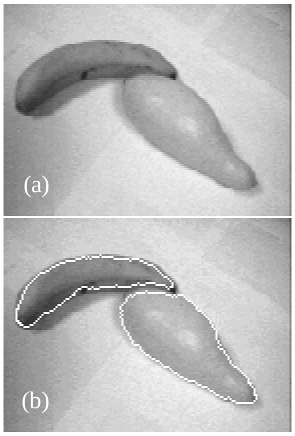

Fig. 7.6 (a) Original Image (b) resulting contours

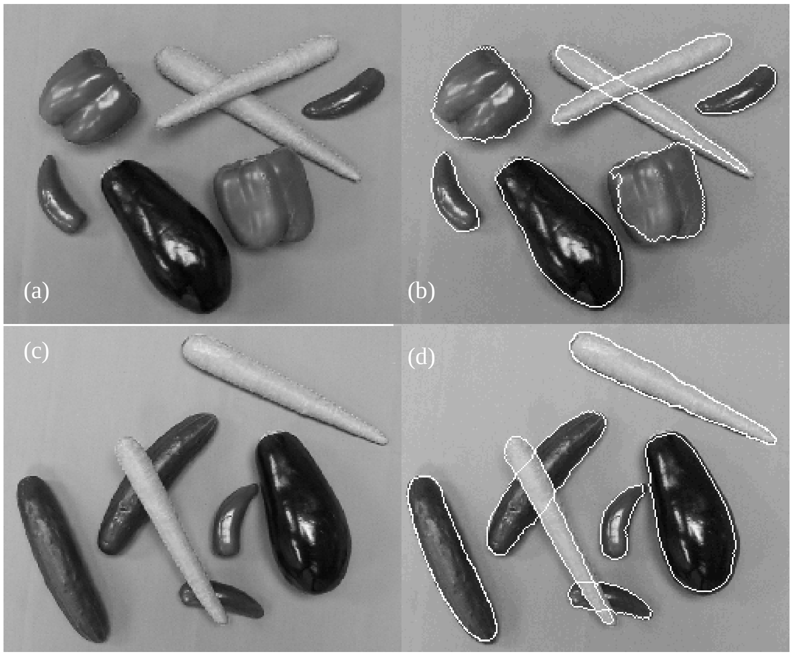

Fig. 7.7 (a) Original Image. (b) resulting contour of (a). (c) Another original image. (d) resulting contour of (c).