Deformable surfaces and parametric models to fit and track 3D data

Laurent D. COHEN
CEREMADE, URA CNRS 749
Université Paris 9-Dauphine
75775 Paris cedex 16, France
cohen@ceremade.dauphine.fr

ABSTRACT
We present an overview of part of our work on snakes, balloons and deformable models since many years. This includes 2D and 3D active contours, surface reconstruction with discontinuities and parametric deformable models that constrain the shape of the surface. We show examples of applications on medical images.

1. INTRODUCTION
Active contour models, introduced by Kass, Witkin and Terzopoulos [17], and many variations on these deformable models have been studied for almost a decade and used for many applications. Their use is particularly interesting for tracking nonrigid deformation of structures found in biomedical images. We began using snakes for detecting the contour of Left Ventricular cavity of the heart in MR Images [12]. Instead of the gradient potential, we defined an attraction potential from a binary image of already extracted edges using distance map or convolution. The main drawbacks of the model being its initialization and minimization, we defined the balloon model [12] to extract the closed contour of an object being less demanding on the initial curve. Starting from almost any curve inside the object permits to recover the whole boundary by inflating the curve like a balloon (see figure 1). This model was used as a first approach to make 3D reconstruction from cross sections (see figure 2). A simplified 3D model was introduced as a stack of 2D balloons deforming simultaneously [15]. These two approaches suit well in the case of a cylindric shape. To recover more general 3D surfaces in 3D medical images, the snake and balloon model was generalized to 3D and implemented with a finite element method in [15, 11] (see figure 3). The normal inflation force appeared later as the basic evolution equation for many models with moving front [7, 21, 29]. We also used recently a front evolution approach to solve the minimization in the snake problem [16].

Using deformable models and templates, the extraction of a shape is obtained through minimization of an energy composed of an internal regularization term and an external attraction potential (data fitting term), illustrated for example in [30, 32, 15, 11, 35, 31, 3, 2]. Since the relevant surfaces in medical images are usually smooth, the use of such models is often very efficient for locating surface boundaries of organs and structures, and for the subsequent tracking of these shapes in a time sequence.

Although previous approaches based on general deformable surfaces [15, 11] give satisfying results, they involve large linear systems to solve and heavy structures. This is particularly important when dealing with a huge amount of data like for object tracking in a sequence of 3D images. This is why a prior knowledge on the surface may be useful, by constraining its shape (section 5), or by decomposition in regions bounded by discontinuities (section 4). Parametric models are well adapted when to impose global constraints. However, further deformation of the parametric model is usually necessary like with deformable superquadrics [31] (see section 5.2).

2. RECONSTRUCTION WITH REGULARIZATION
Our work deal with the use of deformable models and parametric deformable templates for image analysis, in particular for segmentation and reconstruction with regularization. To extract and represent the shapes of objects in images, there is a simultaneous segmentation and reconstruction:

(1) localize “edges” that belong to the boundary of a same object (segmentation).

(2) Reconstruction, from a set of structured “edges” given in a known order, of a smooth curve or surface.

In classic methods, Step (2) is solved after Step (1) through explicit constraints. For deformable models, the two steps are merged in one through implicit constraints.

The general problem of reconstruction of a curve or surface $v$ from input data $d$ that are previously segmented can be formalized by minimization in $v$ of energy:

$$E(v, d) = \int R(v(s)) ds + \int V(v(s), d(s)) ds$$  \hspace{1cm} (1)
where

- \( R(v) \) measures the smoothness of the reconstruction \( v \). Smoothness of \( v \) is obtained
  - either through the term \( R(v) \) composed of first or second order partial derivatives of \( v \) in the case of a free curve or surface. For example \( R(v) = \|v''(s)\|^2 \) or \( \|\nabla v\|^2 \). These are the more general deformable models;
  - To introduce discontinuities, we can choose a non convex term \( R(v) = f(\|\nabla v\|) \).

- \( V \) measures fidelity of \( v \) to data \( d \). It can be
  - explicit (convex): \( V(v, d) = \|v - d\|^2 \)
  - implicit (non convex): \( V(v, d) = P_d(v) \) where \( P_d \) is an attraction Potential.

This formulation applies to many fields in Image Processing including edge extraction, segmentation, Image restoration and enhancement, and Matching [30, 23, 6, 17, 33].

3. DEFORMABLE MODELS

Active Contour Models or Snakes

In the case of plane curves, we are looking for a curve \( v(s) = (x(s), y(s)) \) minimizing energy:

\[
v \mapsto E(v) = \int_L w_1 \|v'(s)\|^2 + w_2 \|v''(s)\|^2 + P(v(s))ds
\]

(2)

This energy models mechanical properties that are between an elastic string (first order) and a more rigid rod or spline (second order).

Starting from an initial estimate, we solve the evolution equation with fixed or periodic boundary conditions:

\[
\frac{\partial v}{\partial t} = \frac{\partial}{\partial s}(w_1 \frac{\partial v}{\partial s} + \frac{\partial^2 v}{\partial s^2}) = F(v)
\]

(3)

where \( F = -\nabla P \), is the attraction force towards contours. This equation is equivalent to making a gradient descent of the energy to converge to a minimum of (2).

Minimizing Surfaces

For a deformable surface, the energy to minimize is:

\[
E(v) = \int_\Omega w_{01} \left( \frac{\partial v}{\partial r} \right)^2 + w_{01} \left( \frac{\partial v}{\partial s} \right)^2 + w_{20} \left( \frac{\partial^2 v}{\partial r^2} \right)^2 + w_{20} \left( \frac{\partial^2 v}{\partial s^2} \right)^2 + 2w_{11} \left( \frac{\partial^2 v}{\partial r \partial s} \right)^2 + \int_\Omega P(v)dx
\]

(4)

In this case, we model a surface that has physical properties between an elastic membrane (first order, like food wrapping plastic paper) and a thin plate (second order, for example projection transparencies).

The partial differential equation to solve writes:

\[
\frac{\partial v}{\partial t} - \frac{\partial}{\partial s}\left(w_{01} \frac{\partial v}{\partial s} - w_{01} \frac{\partial v}{\partial s} \right) = F(v)
\]

(5)

\[
+ \frac{\partial^2}{\partial r^2}(w_{20} \frac{\partial^2 v}{\partial r^2}) + \frac{\partial^2}{\partial s^2}(w_{20} \frac{\partial^2 v}{\partial s^2})
\]

\[
+ 2w_{11} \frac{\partial^2 v}{\partial r \partial s} = F(v)
\]

Resolution

The equations are solved using finite differences or finite elements. In any case, we can always write the iterative scheme as:

\[
(Id + \tau A)v^t = (v^{t-1} + \tau F(v^{t-1}))
\]

(6)

where \( v^t \) represents the vector of unknowns (nodes or degrees of freedom) of the discrete curve or surface at iteration \( t \).

Attraction Potential \( P \)

Instead of taking \( P(v) = -\|\nabla I(v)\|^2 \) introduced in the original model [17], it was often more interesting to use the attraction potential to a set of edge points that were obtained previously. Indeed, since in general the gradient is not constant along a contour, this permits to make the contours uniform by giving them equally weighted attraction along the boundary. This can also avoid a high concentration of nodes at higher gradient values and we obtain better results. Moreover, this edge image is usually easily made available together with the original image. Also, it permits to solve the problem when the data itself is a binary edge image. The attraction acts as a zero length string that links a point of the curve and a data point [15]. To define this attraction potential, two approaches were used. The first was convolution of the edge image with a Gaussian. The second, more popular, was a function of the Chamfer distance to edges. \( P(v) = g(\delta(v)) \) where \( \delta(v) \) is the distance between a point \( v \) and the closest edge point. Function \( f \) is increasing on \( R^+ \), for example, \( P(v) = e^{-\delta(v)^2} \) or \( P(v) = \text{Min}(\delta(v)^2, \beta) \).

In both cases, \( g \) is bounded, in order to allow the string linked to the data to break. This is related to robust statistics. Chamfer distance is a good approximation of this distance map. It can be obtained by a fast sequential algorithm or using front propagation (see [16]). We see in figure 1 an example of distance map.

Balloon Model: Expansion Pressure Force

Usually, to make the curve converge to the right solution, the user has to provide an initial guess that
Figure 1: MRI 3D Image of the heart. Top: edge detection and attraction potential. Middle and bottom: Evolution of the balloon on a slice.

is rather close to it. To make the curve or surface converge to the solution, even when it is not close to it, we introduced the “Balloon model” [12]. We add a pressure force pushing outwards like inflating a balloon. This gives the curve or surface a more dynamic behavior

\[ F_{\text{balloon}} = k \vec{n}(s) \]  \hspace{1cm} (7)

where \( \vec{n}(s) \) is the unit normal vector to the curve at point \( v(s) \). The curve behaves like a balloon which is inflated. It is stopped by a strong edge but avoids the curve being “trapped” by spurious isolated edge points. This makes the result much less sensitive to the initial conditions. Since by inflating the model, the size of the curve or surface increases, it may be necessary to increase the number of discretization nodes. This may be obtained by making reparameterization every few iterations, defining regularly spaced nodes.

Remark that this force derives from the inside area (or volume) energy

\[ E_{\text{area}} = -k \int_{\text{inside region}} dA \]  \hspace{1cm} (8)

that measures the area (or volume) inside the region bounded by the curve (or surface). Minimization of such energy corresponds to get a region as large as possible. This is obtained by a pressure force in the outward normal direction.

Figure 2: 3D reconstruction from cross sections (two views).

Figure 3: Evolution of the deformable surface to the energy minimum in a 3D MR image of a face (extracted from [15]).

Recently, many works have been made based on the evolution of a plane curve subject to a normal force. This was either in a purely mathematical framework or for many application in image processing [24, 1, 26, 7, 21, 8]. Remark also the similarity between the evolution of a plane curve submitted to a pressure force (7) with a dilatation in mathematical morphology.

The current trend to define a deformable curve or surface is by using an intrinsic geometric model [7, 21]. The surface is defined as the zero level set of a 3D function. This function is “deformed” in order to make its zero level set follow the minimization of the potential. These are called either geometric or geodesic deformable models [8, 9, 18], implicit deformable surface [94], or bubbles [29]. This is efficient to obtain an easy
way to change topology. Other models also permit the curve or surface to change topology [20, 28, 22].

The study of the evolution of a shape through a partial differential equation with invariants [1] or the presence of shocks in other cases (Reaction-diffusion space) [26, 19] permits the characterization of a shape.

4. RECONSTRUCTION WITH DISCONTINUITIES

When recovering a surface $u$ from explicit data $d$, the first order regularizing term $R(u) = \lambda^2 \int \|
abla u\|^2 dxdy$ has the same smoothing effect all over the surface. Large variations of elevation in the surface, corresponding to discontinuities between two regions, are blurred or removed by the smoothing. To avoid this problem, we can either make parameter $\lambda$ depend on position $(x, y)$ [30], or introduce discontinuities with penalty (as in [23, 6]). The advantage of the latter is that it does not assume the discontinuities as known in advance. The authors of [23, 6] introduce a penalty term at each point where a discontinuity is detected, instead of its contribution to $R(u)$. This set of discontinuities is also an unknown of the problem and an argument of the energy. The energy becomes:

$$E(u, B) = \int_R (u - d)^2 + \lambda^2 \int_{R-B} \|
abla u\|^2 + \alpha l(B) \quad (9)$$

where $R$ represents the image domain, and $B$ is a set of boundary points where discontinuities of $u$ are introduced. The last term $l(B)$ is the length of this set of contours. The constant $\alpha$ can be interpreted as a contrast detection factor since a point will choose to be in $B$ depending of the lower cost either of the gradient contribution in $R(u)$ or paying a penalty $\alpha$. In the discrete energy, the term $R(u)$ can also be formalized as $R(u) = f(\|
abla u\|)$ where $f(y) = \min \{y^2, \alpha\}$.

The regularizing effect operates only inside the regions bounded by $B$. Thus $u$ is piecewise regular and can be discontinuous along $B$.

In [14], we used an active contour to detect a curve of discontinuities in a surface. In figure 4, we see a digital terrain model of a mountain and lake. The surface has to be smoothed and at the same time, we wish to recover the lake as a constant elevation region bounded by discontinuities.

We thus defined the reconstruction of a surface $u$ composed of two different kinds of regions. One is a constant lake region $L$ at level $u_0$, the other is the background $(R-L)$. The boundary curve $B$ between them is obtained with an active contour model $v$. This combines the two problems of surface reconstruction with discontinuities and contour detection. We minimize an energy that is function of the couple of unknowns: (surface $u$, boundary curve $v$):

$$E_g(u, v) = E_{\text{snake}}(v) + E_{\text{lake}}(u, v) + E_{\text{outside}}(u, v)$$

$$= E_{\text{snake}}(v) + \int_L (u_0 - d(x, y))^2 dxdy + \int_{R-L} (u - d(x, y))^2 + \int_{R} (u - u_0)^2 \quad (10)$$

where $\Gamma$ corresponds to points on $B$ such that $(u - u_0) < 0$.

The minimization of a global energy allows direct interaction between surface terms and the active contour (see figure 4). Initialization of the contour is made easier (see [14]).

The algorithm successively minimizes energy $E_g$ with respect to each of the two variables $u$ and $v$. When $u$ is given, minimization in $v$ corresponds to active contour evolution where new external forces are added that derive from the surface terms. This allows the model to take into account the fact that the level inside the region has to be homogeneous. This avoids the curve being trapped by spurious edges. This property is not satisfied in classical active contours, where the curve “sees” only what happens locally along the curve and this may stop its evolution due to local minima.

Figure 4: Digital Terrain Model with a lake: noisy image and reconstruction. (from [14]).
5. PARAMETRIC DEFORMABLE MODELS

We have illustrated the capacity of 2D and 3D deformable models to define a local description of structures appearing in images. These local models do not allow constraints on the global shape of the object. The use of parametric curves and surfaces (circles, parabolas, segments in [35], Superquadrics [25, 27, 31], hyperquadrics [10] or B-Snakes [3, 13]) allows the definition of global constraints on the shape to find. The same kind of attraction potential is minimized over the curve or surface defined by a small number of parameters. These models make easier the retrieval of missing data (see figure 5). They reduce the number of unknowns and give a compact and stable representation.

5.1. Global Models

Using Superquadrics or hyperquadrics, a priori knowledge of the shape can be included in the model. Superquadrics are a generalization of quadrics with real exponents that can be defined by the implicit equation:

$$
\left( \left( \frac{x}{a_1} \right)^{2/\epsilon_2} + \left( \frac{y}{a_2} \right)^{2/\epsilon_2} \right)^{\epsilon_4/\epsilon_1} + \left( \frac{z}{a_3} \right)^{2/\epsilon_1} = 1.
$$

Hyperquadrics (see [10]) are a generalization of those:

$$
H(x, y, z) = \sum_{i=1}^{N} a_i x + b_i y + c_i z + d_i |x_i^\alpha| = 1.
$$

A shape $S_A$ being defined by a set of parameters $A$, a two step algorithm (see [10, 13]) can be used to minimize

$$
E(A) = \int_{S_A} P \approx \sum_i P(v_i(A))
$$

where $v_i(A)$ is a discretization of the shape $S_A$. Initialization is made by a model $A^0$ or a set of points $M^0$. Then the following two steps are iteratively applied:

**Global Fit.** An ordered set $M^k = (M^k_i)_{i \leq n}$ being given. New parameters $A^k$ are defined by minimization of $E^k_f(A)$:

$$
\sum_i \|v_i(A) - M_i^k\|^2 \approx \sum_i \left\{ \frac{1}{\|\nabla F_A\|} (F_A(M_i^k) - 1) \right\}^2
$$

**Local Deformation of the set of points.** $A^k$ being given, a new set $M_i^{k+1} = (M_i^{k+1})_{i \leq n}$ is defined by minimization of

$$
E_2(M) = \sum_i P(M_i)
$$

$$
M_i^{k+1} = v_i(A^k) - \alpha_i^{k+1} \nabla P(v_i(A^k))
$$

5.2. Local Refinement

**Hybrid hyperquadrics**

Although parametric models are able to describe complex shapes by a few parameters, this is usually not enough for recovery of complex shapes. It may be necessary to refine the global model by adding a local deformation [31]. We have studied such methods for each of the two previous models. The first model is using hybrid hyperquadrics [10] that is obtained by adding to the equation 5.1 exponential terms ($e^{-H}$). It enables to model local properties and introduce concavities:

$$
H(x, y, z) = \sum_{i=1}^{N} |K_i|^\gamma + \sum_{j=1}^{M} c_j e^{-\sum_{i=1}^{N} |K_i|^\gamma} = 1,
$$

(11)

**Free-Form Deformations**

To get a better reconstruction of data, we deform [4] the superquadric model using tri-dimensional B-splines called Free-Form Deformation (FFD) in computer graphics. A point $X$ is defined by its local coordinates $(s, t, u)$ and the position of control points $P_{ijk}$:

$$
\sum_{i,j,k=0}^{l,m,n} C_i^l C_j^m C_k^n (1-s)^l (1-t)^m (1-u)^n P_{ijk}
$$

(12)

A two-step iterative algorithm is used:

1: Compute the Displacement Field : 

$$
X_{n+1}^a = X_n + \delta X_n
$$

2: New Control points $P_{n+1}$ by minimization of 

$$
\|BP - X_{n+1}^a\|^2
$$

$$
X_{n+1} = BP_{n+1}
$$

Results that illustrate this approach are shown in figure 6. The model allows a very compact representation of the data. In the example, there are 6000 3D data points and the model is defined by 130 3D points (a 5x5x5 box). Since the FFD is volumic it can give the deformation between two surfaces, like in figure 6. Figure 5 shows how this model can retrieve a good reconstruction from sparse data. This applies to tracking (see figure 7) and permits to get estimates of the deformation. This provides a set of quantitative parameters and permits to detect pathologic areas [5].

**Acknowledgments.**

Some of the presented work were in collaboration with Nicholas Ayache, Isaac Cohen and Eric Bardinet. I would like to thank them and wishes that these fructuous collaborations go on.
6. REFERENCES


---

Figure 5: Reconstruction with sparse data. Above complete (right) and sparse (left) data; below the two reconstructions.

---

Figure 6: Steps of reconstruction with superquadric and FFD. On top: data and superquadric fit inside FFD box; in the middle: displacement field and final surface recovery. In the bottom, volumic deformation between the two epicardial and endocardial surfaces (extracted from [4]).

---