MetaMorphs: Deformable Shape and Texture Models

Xiaolei Huang, Dimitris Metaxas and Ting Chen

Division of Computer and Information Sciences, Rutgers University, Piscataway, NJ, USA

{xiaolei, dnm}@cs.rutgers.edu, chenting@graphics.cis.upenn.edu

Abstract

We present a new class of deformable models, MetaMorphs, that consist of both shape and interior texture. The model deformations are derived from both boundary and region information in a common variational framework. This framework represents a generalization of previous model-based segmentation approaches. The shapes of the new models are implicitly represented as an "image", which is the higher dimensional embedding space of distance transforms. The interior textures are captured using a nonparametric kernel-based approximation of the intensity distribution inside the models. The models shapes can undergo both global and local deformations. The local deformations are efficiently parameterized using the cubic B-spline based Free Form Deformations (FFD). When using the models for boundary finding in images, we derive the model dynamics from an energy functional consisting of both edge energy terms and intensity/texture energy terms. This way, the models deform under the influence of forces derived from both boundary and regional information. The proposed MetaMorph deformable models are efficient in convergence, have large attraction range, and are robust to image noise and inhomogeities. Various examples on noisy medical images with complex textures demonstrate the potential of the proposed technique.

1. Introduction

Computerized segmentation plays a fundamental role in medical image processing and analysis. Due to the common presence of cluttered objects, complex backgrounds, noise and intensity inhomogeities in medical images, the image segmentation still remains a difficult task. To address these difficulties, deformable models [6, 12, 3, 7] have been extensively studied and used as a model-based segmentation approach. Starting from an initial estimate, a deformable model evolves under the influence of both internal (e.g. smoothness) and external (e.g. image) forces to converge to the desired boundary of an image object. In traditional deformable models, image forces come primarily from edge (image gradient) information. Such reliance on edge information, however, makes the models sensitive to noise and highly dependent on the initial estimate.

In the past few years, there have been significant efforts to integrate region information into deformable models. In [10], local region analysis strategies are introduced for Active Contour Models. But the optimization of the integrated energy function is mostly heuristic. In [13], a generalized energy function that integrates region growing and boundary-based deformations was proposed. They proposed the idea of region competition to control the model deformation. In this formulation, however, the parameters of the regional intensity statistics can not be updated simultaneously with the boundary shape parameters so that the energy function has to be minimized in an iterative way. Hybrid segmentation frameworks are proposed in [5, 2]. In these methods, a region based segmentation module is used to get a rough binary mask of the object of interest. Then this rough estimation of the object can be used to initialize a deformable model, which will deform to fit edge features in the image using the gradient information. The region estimation and model deforming steps work recursively pushing each other out of local minima to achieve a better segmentation solution. However, in these frameworks, the region information and the boundary information are treated separately in different energy minimization processes so that the integration is still imperfect. As noted in [4] on incorporating region information in active contour models for tracking applications, the difficulty into coupling regional and boundary information is mostly due to the fact that the set of image regions does not have a structure of vector space, preventing us to use in a straightforward manner gradient descent methods, especially when statistical features of a region (such as mean and variance of intensity) are present. The authors turned to registration-like energy criterion in tracking, to circumvent this problem.

To address the limitations in previous efforts to incorporate region information in deformable models, we introduce in this paper a new class of deformable models, which we term the "MetaMorphs". The MetaMorph models possess both shape and interior texture (intensity distributions), and integrate boundary and region information coherently in a common variational framework. Instead of using traditional parametric curves/surfaces, the model shapes

in our framework are embedded in a higher dimensional space of distance transforms, thus represented by the distance map "images". The model deformation parameters are the global transformation parameters, and the local cubic B-spline based Free Form Deformations (FFD) parameters. The interior intensity statistics of the models are captured using the nonparametric kernel-based approximations. In our energy functional formulation, this derived region statistics are differentiable with respect to the model deformation parameters, which allows a unified gradient-descent based model parameter updating paradigm using both boundary and region information. When using the MetaMorph models for boundary finding in medical images, the model dynamics are derived from an energy functional consisting of a model interior shape term, a boundary data term and a model interior texture term. We are able to omit a boundary smoothness term due to our use of the B-spline based FFD. During model evolution, the three energy terms will have complementary effect, to help the model grow/shrink and out of local minima due to small spurious edges inside the object, to prevent the model from leaking at boundary gaps, and to enable the segmentation of objects with intensity inhomogeities and complex interior statistics. The weighting factors between the different energy terms can be dynamically adjusted after each iteration to achieve good results.

Our work is a generalization of previous explicit parametric deformable models to take into account model interior texture information. Another line of research on deformable models are the works on implicit geometric deformable models [8, 1, 9], which are implemented in the level set based curve evolution frameworks. The two lines of research have their own commonly known advantages and disadvantages.

The remainder of the paper is organized as follows. In section 2, we introduce the shape and texture representations of the MetaMorph models. In section 3, we derive the MetaMorph model dynamics from both boundary and region information. In section 4, the overall model fitting algorithm and experimental results are presented, and we conclude with discussions in section 5.

2. The MetaMorph Models

In this section, we present the shape and texture representations of the MetaMorph deformable models. The model shapes are implicitly represented as "images" in the space of distance transforms. The interior textures of the models are captured using the nonparametric kernel-based method, which can closely approximate complex multi-modal intensity distributions.

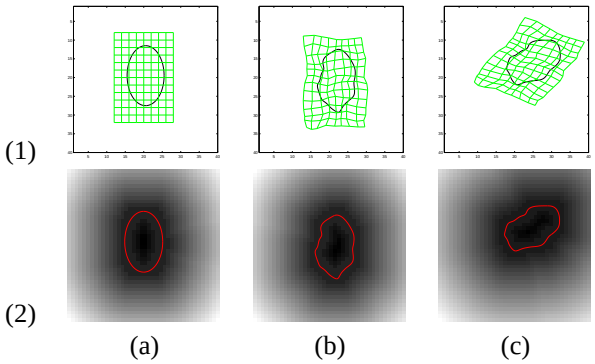


Figure 1: Transformations and deformations of the MetaMorph Models. (1) The model shape. (2) The implicit "image" representation of the model shape. (a) Initial model. (b) The model after local FFD deformations in its object-centered coordinate system. (c) Further global transformation of the model (anisotropic scaling, translation and rotation).

2.1. The Models Shape and Generalized Geometry

The models shape is embedded implicitly in a higher dimensional space of distance transforms. The Euclidean distance transform is used to embed the evolving model as the zero level set of a higher dimensional distance function. In order to facilitate notation, we consider the 2D case. Let $\Phi : \Omega \rightarrow R^+$ be a Lipschitz function that refers to the distance transform for the model shape $\mathcal{M}$. By definition Ω is bounded since it refers to the image domain. The shape defines a partition of the domain: the region that is enclosed by $\mathcal{M}$, $[\mathcal{R}_{\mathcal{M}}]$, the background $[\Omega - \mathcal{R}_{\mathcal{M}}]$, and on the model, $[\partial\mathcal{R}_{\mathcal{M}}]$. Given these definitions the following implicit shape representation is considered:

$$\Phi_{\mathcal{M}}(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in \partial\mathcal{R}_{\mathcal{M}} \\ +ED(\mathbf{x}, \mathcal{M}) > 0, & \mathbf{x} \in \mathcal{R}_{\mathcal{M}} \\ -ED(\mathbf{x}, \mathcal{M}) < 0, & \mathbf{x} \in [\Omega - \mathcal{R}_{\mathcal{M}}] \end{cases}$$

where $ED(\mathbf{x}, \mathcal{M})$ refers to the min Euclidean distance between the grid location $\mathbf{x} = (x, y)$ and the model $\mathcal{M}$.

Such treatment makes the model shape representation an "image", which greatly facilitates the integration of boundary and region information. It also provides a feature space in which objective functions that are optimized using a gradient descent method can be conveniently used. Since a sufficient condition for convergence of the gradient descent methods requires continuous first derivatives. The considered implicit representation satisfies such a condition. One can prove that the gradient of the distance function is a unit vector in the normal direction of the shape. This property will make our model evolution fast. Examples of this implicit representation can be found in [Fig. (1).2]. This shape representation in 3D is similarly defined in a volumetric space.

The model can undergo a hierarchy of both global and local deformations. In the 2D case, the global similarity transformations of the model are captured by the parameters:

$$\mathbf{q}_g = [t_x \;\; t_y \;\; \theta \;\; s_x \;\; s_y]^T$$

where t_x, t_y are the translations along x and y directions, θ is the rotation angle, and s_x, s_y are the scaling factors along x and y directions.

The local non-rigid deformations of the model is represented using a space deformation technique, the Free Form Deformations (FFD) [11], which is a popular approach in graphics, animation and rendering. The essence of FFD is to deform an object by manipulating a regular control lattice F overlaid on its volumetric embedding space. One of the main advantages of the FFD technique is that it imposes implicit smoothness constraints during deformation, since it guarantees C^1 continuity at control points and C^2 continuity everywhere else. Therefore there is no need for introducing computationally expensive regularization components on the local deformations of the model shape. Another advantage is that the extension of FFD to higher dimensions (3D) is straightforward and much easier than the traditional parametric representations. Furthermore, since the FFD is a space warping technique, it integrates naturally with the implicit model shape representation in a higher dimensional embedding space. In this paper, we consider an Incremental Free Form Deformations (IFFD) formulation using the cubic B-spline basis.

Let us consider a regular lattice of control points

$$F_{m,n} = (F^x_{m,n}, F^y_{m,n});\; m = 1, ..., M, \quad n = 1, ..., N$$

overlaid to a region $\Gamma_c = \{\mathbf{x}\} = \{(x, y) | 1 \leq x \leq X, 1 \leq y \leq Y\}$ in the embedding space that encloses the model in its object-centered coordinate system. Let us denote the initial configuration of the control lattice as F^0, and the deforming control lattice as $F = F^0 + \delta F$. Under these assumptions, the incremental FFD parameters, which are also the local deformation parameters for the model, are the deformations of the control points in both directions (x, y);

$$\mathbf{q}_l = \{(\delta F^x_{m,n}, \delta F^y_{m,n})\};\; (m, n) \in [1, M] \times [1, N]$$

The deformed position of a pixel $\mathbf{x} = (x, y)$ given the deformation of the control lattice from F^0 to F, is defined in terms of a tensor product of Cubic B-spline:

$$L(\mathbf{q}_l; \mathbf{x}) = \mathbf{x} + \delta L(\mathbf{q}_l; \mathbf{x}) = \sum_{k=0}^{3} \sum_{l=0}^{3} B_k(u) B_l(v) \; (F^0_{i+k,j+l} + \delta F_{i+k,j+l}) \qquad (1)$$

where $i = \lfloor \frac{x}{X} \cdot (M-1) \rfloor + 1,\; j = \lfloor \frac{y}{Y} \cdot (N-1) \rfloor + 1$. The terms of the deformation component refer to:

- $\delta F_{i+l,j+l}$, $(k, l) \in [0, 3] \times [0, 3]$ are the deformations of pixel $\mathbf{x}$'s (sixteen) adjacent control points,
- $B_k(u)$ is the k^{th} basis function of a Cubic B-spine, defined by:

  $$B_0(u) = (1-u)^3/6,\; B_1(u) = (3u^3 - 6u^2 + 4)/6$$
  $$B_2(u) = (-3u^3 + 3u^2 + 3u + 1)/6,\; B_3(u) = u^3/6$$

  with $u = \frac{x}{X} \cdot (M-1) - \lfloor \frac{x}{X} \cdot (M-1) \rfloor$. $B_l(v)$ is similarly defined.
- $\delta L(\mathbf{q}_l; \mathbf{x}) = \sum_{k=0}^{3} \sum_{l=0}^{3} B_k(u) B_l(v) \delta F_{i+k,j+l}$ is the incremental local deformation for pixel $\mathbf{x}$.

Now combining the global and local components, the deformation parameters for the model are:

$$\begin{aligned} \mathbf{q} &= [\mathbf{q}_g;\;\; \mathbf{q}_l] \\ &= \left[t_x \; t_y \; \theta \; s_x \; s_y;\; \{(\delta F^x_{m,n}, \delta F^y_{m,n})\}\right], (m, n) \in [1, M] \times [1, N] \end{aligned}$$

And the overall transformed position of a pixel $\mathbf{x}$ on the model can be written as:

$$D(\mathbf{q}; \mathbf{x}) = \mathbf{t}(t_x, t_y) + \mathbf{R}(\theta)\mathbf{S}(s_x, s_y) L(\mathbf{q}_l; \mathbf{x}) \qquad (2)$$

where $\mathbf{t}$ is the translation vector, $\mathbf{R}$ is the rotation matrix, $\mathbf{S}$ is the scaling matrix, and $L(\mathbf{q}_l; \mathbf{x})$ is the locally deformed coordinate given the FFD parameters $\mathbf{q}_l$.

One example for the model transformations is shown in [Fig. (1)]. An initial model is shown in [Fig. (1).a]. When its embedding space deforms due to the deformation of the FFD control lattice, as shown in [Fig. (1).b], the model undergoes local non-rigid deformations in its object-centered coordinate system. Then global transformations will further transform the model in the world coordinate system [Fig. (1).c], changing its scale (isotropic or anisotropic), position (translation) and orientation (rotation).

The extensions of the model to account for deformations in 3D is straightforward, by using 3D global transformations, and locally using control lattices in the 3D space and a 3D tensor product of B-spline polynomials.

2.2. The Models Texture

Rather than using traditional statistical parameters (such as mean and variance) to approximate the intensity distribution of the model interior, we approximate the distribution using a nonparametric kernel-based method. The nonparametric approximation is differentiable, more generic and can represent complex multi-modal intensity distributions.

Suppose the model is placed on an image I, the image region bounded by current model $\Phi_{\mathcal{M}}$ is $\mathcal{R}_{\mathcal{M}}$, then the probability of a pixel's intensity value i being consistent with the model interior intensity can be derived using a Gaussian kernel as:

$$\mathbf{P}(i | \Phi_{\mathcal{M}}) = \frac{1}{V(\mathcal{R}_{\mathcal{M}})} \iint_{\mathcal{R}_{\mathcal{M}}} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(i - I(\mathbf{y}))^2}{2\sigma^2}} d\mathbf{y}$$

where $V(\mathcal{R}_{\mathcal{M}})$ denotes the volume of $\mathcal{R}_{\mathcal{M}}$, and σ is a constant specifying the width of the gaussian kernel.

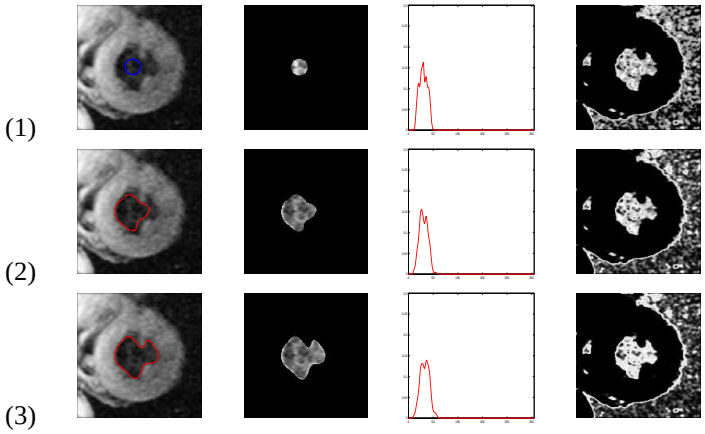


Figure 2: The Endocardium segmentation. (1) Initial model. (2) Intermediate result. (3) Final converged result. (a) The evolving model drawn in color (blue or red) on original image. (b) Interior of the evolving model. (c) The intensity p.d.f of the model interior. (d) The image probability map based on the p.d.f of the model interior.

Using this nonparametric approximation, the intensity distribution of the model interior gets updated automatically while the model deforms. The initialization of the model texture is flexible. We can either start with a small model close to the texture region to be segmented, or use supervised learning to specify the desired texture distribution a Priori. One example of the model interior intensity/texture representation can be seen in [Fig. (2)]. In the figure, we show the zero level set of the current model $\Phi_{\mathcal{M}}$ in color [Fig. (2).a], the model interior region $\mathcal{R}_{\mathcal{M}}$ [Fig. (2).b], the probability density function (p.d.f.) for the intensity of current model interior $\mathbf{P}(i|\Phi_{\mathcal{M}})$ for $i = 0, ...255$ [Fig. (2).c], and the probability map of every pixel's intensity in the image according to the model interior distribution [Fig. (2).d].

3. The MetaMorph Dynamics

We demonstrate the MetaMorph model fitting dynamics in the context of a segmentation problem.

In order to fit to the boundary of an object, the motion of the model is driven by both gradient (edge) energy terms and texture (intensity) energy terms derived from the image. The overall energy functional consists of three terms: the model interior shape data term E_S, the boundary data term E_B, and the intensity data term E_I. Formally, this can be written as:

$$E = aE_S + bE_B + cE_I \tag{3}$$

In our formulation, we are able to omit the boundary smoothness term in the traditional parametric or level-set based deformable models, since this smoothness is implicit by using the local Free Form Deformations model. Next, we derive the three energy terms respectively.

Figure 4: The effect of small spurious edges inside the object of interest on the "shape image". (a) The original MR image. (b) The edge map of the image. (c) The derived "shape image". Note the effect of the small spurious edges on the "shape image" inside the object.

3.1. The interior shape data term

The gradient information is a very important source of the image forces for a deformable model. We encode the gradient information of the image using a "shape image" Φ, which is derived from the un-signed distance transform of the edge map of the image. In [Fig. (3).c], we can see the "shape image" of an example MR heart image.

In the interior shape data term of the model, we aim to minimize the Sum-of-Squared-Differences between the implicit representation values in the model interior and the underlying "shape image" values at corresponding deformed positions. This can be written as:

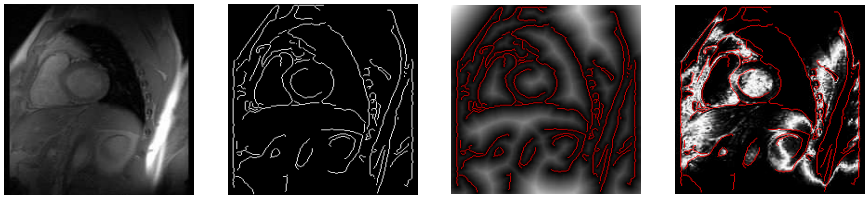
$$E_S = \frac{1}{V(\mathcal{R}_{\mathcal{M}})} \iint_{\mathcal{R}_{\mathcal{M}}} \left(\Phi_{\mathcal{M}}(\mathbf{x}) - \Phi(D(\mathbf{q};\mathbf{x}))\right)^2 d\mathbf{x}$$

During optimization, this term will deform the model along the gradient direction of the underlying "shape image". Thus it will expand or shrink the model accordingly, serving as a two-way balloon force without explicitly introducing such forces, and making the attraction range of the model large.

3.2. The boundary term

The previous interior shape term is good in attracting the model towards boundary structures from far-away locations. However, when there are small spurious edges detected within an object due to texture, the "shape image" inside the object could differ in the surrounding areas of those small edges. One such example can be seen in [Fig. (4).a-c]. We intend to segment the Endocardium of the Left ventricle in the MR image. During edge detection, some small spurious edges are picked up inside the object due to texture. And this changes the "shape image" of the object interior. In this situation, the forces due to the interior shape term may get stuck in local minima. To make the model deformation more robust to such situations, we consider a separated boundary data term, which allows higher weights for pixels in a narrow band around the model boundary $\partial\mathcal{R}_{\mathcal{M}}$.

$$E_B = \frac{1}{V(\partial\mathcal{R}_{\mathcal{M}})} \iint_{\partial\mathcal{R}_{\mathcal{M}}} \left(\Phi(D(\mathbf{q};\mathbf{x}))\right)^2 d\mathbf{x}$$



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

Figure 3: "Shape image" of an example MR heart image. (a) The original image. (b) The edge map of the image. (c) The "shape image", i.e. the distance transform of the edge map. The zero level set of the "shape image" is drawn in red. (d) The edges shown in red, superimposed on the intensity probability map derived based on the current model interior statistics.

Intuitively, this term will encourage the deformation that maps the model boundary to the image edge locations where the underlying "shape image" distance values are as small (or as close to zero) as possible. In the overall energy functional E, by setting the value for b larger than a, those model boundary pixels get higher weights.

One additional advantage of the boundary data term is that, at an edge with small gaps, this data term will constrain the model to go along the "geodesic" path, which coincides with the smooth shortest path connecting the two open ends of a gap. This behavior can be seen from [Fig. (5)]. Note that at a small gap of the edge map, the boundary term will favor a path with the smallest accumulative distance values to the edge points.

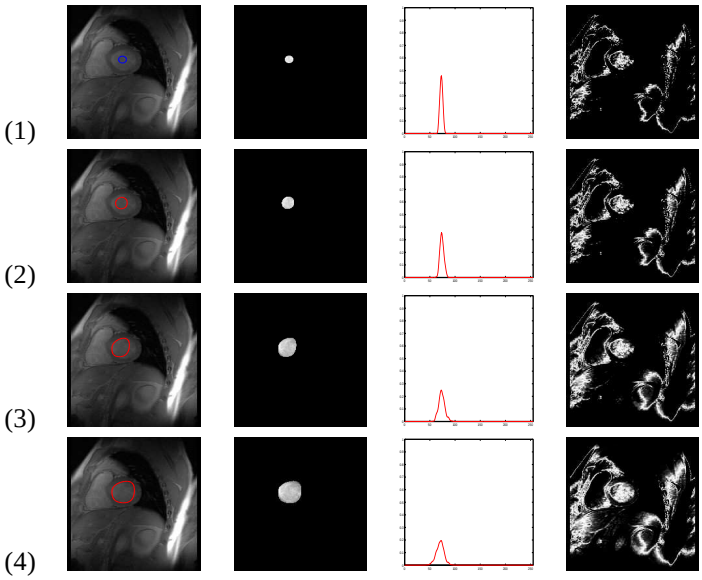
3.3. The interior texture term

One of the most attractive aspects of our MetaMorph deformable models is that they possess both shape and texture information, thus are natural to deform under the influence of forces derived from both edge and region information. We design the texture energy term such that it encourages the model to deform towards areas where the pixel probabilities of belonging to the intensity distribution learned from the past history of the model interiors are high. Suppose the image is I, the texture energy term is formalized using the log-likelihood of the intensities of the model interior after deformation:

$$
\begin{aligned}
E_I &= -\frac{1}{V(\mathcal{R}_\mathcal{M})} \iint_{\mathcal{R}_\mathcal{M}} log\mathbf{P}(I(D(\mathbf{q};\mathbf{x}))|\Phi_\mathcal{M}) d\mathbf{x} \\
&= -\frac{1}{V(\mathcal{R}_\mathcal{M})} \iint_{\mathcal{R}_\mathcal{M}} \Big[log\frac{1}{V(\mathcal{R}_\mathcal{M})} + log\frac{1}{\sqrt{2\pi}\sigma} \\
&+ log \iint_{\mathcal{R}_\mathcal{M}} e^{\frac{-(I(D(\mathbf{q};\mathbf{x}))-I(\mathbf{y}))^2}{2\sigma^2}} d\mathbf{y} \Big] d\mathbf{x}
\end{aligned}
$$

In this formulation, the interior texture term is differentiable with respect to the model parameters $\mathbf{q}$, thus a unified gradient-descent based model parameter updating scheme can be derived using both boundary and region information.

The texture information is very important in helping the model out of local minima, and converge to the optimal solutions. In the example shown in [Fig. (3)], a large portion of the object (Endocardium) boundary is missing due to errors in edge detection [Fig. (3).b]. This makes the interior shape term and the boundary term less reliable. Then the probability map [Fig. (3).d] computed based on the expected model interior statistics becomes the most reliable source for the model evolution. The model fitting results of this example are shown in [Fig. (6)].

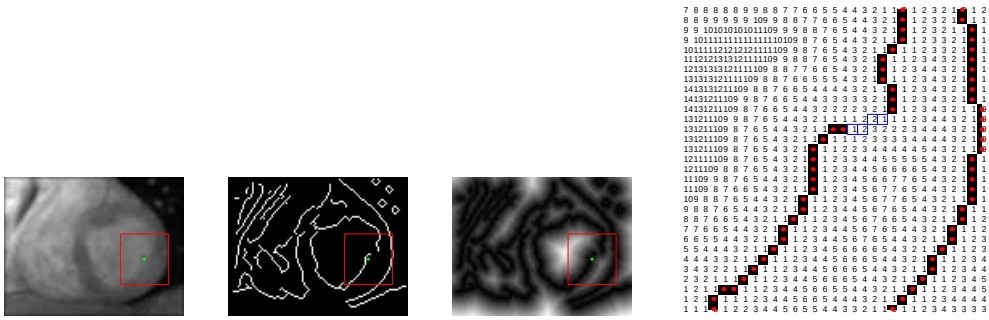


(1) (2) (3) (4)

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

Figure 6: Segmentation of the Endocardium of the Left Ventricle in a MR image with a large portion of the object boundary edge missing.(1) Initial model, with zero level set model shape shown in red. (2-3) Intermediate models. (4) converged model. (a) current model on the image. (b) model interiors. (c) the interior intensity p.d.f. (c) intensity probability map.

In [Fig. (4)], the spurious edges both inside and around the object boundary degrade the reliability of the shape terms, yet the texture information based probability map, as shown in [Fig. (2).d], gives a pretty clear indication of the rough boundary of the object. By combining the texture and shape information, our model is able to evolve close to the true boundary quickly using mainly the texture information and get refined to an optimum with the help of the boundary term. The model fitting results for this example can be seen in [Fig. (2)].



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

Figure 5: The boundary term constraints at small gaps in the edge map. (a) Original Image. (b) The edge map, note the small gap inside the red square region. (c) The "shape image" derived from the edge map. (d) Zoom-in view of the region inside the red square. The numbers are the "shape image" values at each pixel location. The red dots are edge points, the blue squares indicate a path favored by the boundary term for a MetaMorph model.

3.4. Model Evolution

Based on the above energy functionals, one can derive the following evolution equation for each element $\mathbf{q}_i$ in the model parameters $\mathbf{q}$, using the gradient descent method:

$$\frac{\partial E}{\partial \mathbf{q}_i} = a\frac{\partial E_S}{\partial \mathbf{q}_i} + b\frac{\partial E_B}{\partial \mathbf{q}_i} + c\frac{\partial E_I}{\partial \mathbf{q}_i}$$

where

- The motion due to the shape data term is:

$$\frac{\partial E_S}{\partial \mathbf{q}_i} = \frac{1}{V(\mathcal{R}_\mathcal{M})}\iint_{\mathcal{R}_\mathcal{M}} 2\big(\Phi_\mathcal{M}(\mathbf{x}) - \Phi(D(\mathbf{q};\mathbf{x}))\big)\cdot \big(-\nabla\Phi(D(\mathbf{q};\mathbf{x}))\cdot\frac{\partial}{\partial \mathbf{q}_i}D(\mathbf{q};\mathbf{x})\big)d\mathbf{x}$$

- The motion due to the boundary data term is:

$$\frac{\partial E_B}{\partial \mathbf{q}_i} = \frac{1}{V(\partial\mathcal{R}_\mathcal{M})}\iint_{\partial\mathcal{R}_\mathcal{M}} 2\Phi(D(\mathbf{q};\mathbf{x}))\cdot \big(\nabla\Phi(D(\mathbf{q};\mathbf{x}))\cdot\frac{\partial}{\partial \mathbf{q}_i}D(\mathbf{q};\mathbf{x})\big)d\mathbf{x}$$

- And the motion due to the intensity term is:

$$\frac{\partial E_I}{\partial \mathbf{q}_i} = -\frac{1}{V(\mathcal{R}_\mathcal{M})}\iint_{\mathcal{R}_\mathcal{M}}\Big[\Big(\iint_{\mathcal{R}_\mathcal{M}} e^{\frac{-(I(D(\mathbf{q};\mathbf{x}))-I(\mathbf{y}))^2}{2\sigma^2}}d\mathbf{y}\Big)^{-1} \iint_{\mathcal{R}_\mathcal{M}} e^{\frac{-(I(D(\mathbf{q};\mathbf{x}))-I(\mathbf{y}))^2}{2\sigma^2}}\cdot\Big(-\frac{(I(D(\mathbf{q};\mathbf{x}))-I(\mathbf{y}))}{\sigma^2}\cdot (\nabla I(D(\mathbf{q};\mathbf{x}))\cdot\frac{\partial}{\partial \mathbf{q}_i}D(\mathbf{q};\mathbf{x}))d\mathbf{y}\Big]d\mathbf{x}$$

In the above formula, the partial derivatives with respect to the global and local model parameters, $\frac{\partial}{\partial \mathbf{q}_i}D(\mathbf{q};\mathbf{x})$, can be easily derived from the model deformation formula for $D(\mathbf{q};\mathbf{x})$ [Eqn. (2)]. Since the derivation for the global parameters are trivial, we only give the derivation for the local FFD parameters in the Appendix.

4. The Model Fitting Algorithm and Experimental Results

The overall model fitting algorithm consists of the following steps:

1. Initialize the model parameters $\mathbf{q}$ to be $\mathbf{q}^0$, which indicates no deformation.
2. Compute $\frac{\partial E}{\partial \mathbf{q}_i}$ for each element $\mathbf{q}_i$ in the model parameters $\mathbf{q}$.
3. Update the model parameters $\mathbf{q}_i' = \mathbf{q}_i - \lambda\cdot\frac{\partial E}{\partial \mathbf{q}_i}$.
4. Using the new parameters, compute the new model $\mathcal{M}' = D(\mathbf{q}';\mathcal{M})$.
5. Update the model. Let $\mathcal{M} = \mathcal{M}'$, and re-compute the implicit representation of the model $\Phi_\mathcal{M}$, and the new partitions of the image domain by the new model: $[\mathcal{R}_\mathcal{M}]$, $[\Omega - \mathcal{R}_\mathcal{M}]$, and $[\partial\mathcal{R}_\mathcal{M}]$.
6. Repeat steps 1-5 until convergence.

In the algorithm, after each iteration, both the shape and interior intensity/texture of the model get updated based on the model dynamics, and the deformation parameters of the model get re-initialized for the new model. This allows continuous, both large-scale and small-scale deformations for the model to converge to the energy minimum.

Some examples of using this algorithm for segmentation are shown in [Fig. (2)] and [Fig. (6)].

When multiple models are initialized in an image, each model evolves based on its own dynamics. At the end of each iteration, a collision detection step is applied by checking whether the interiors of more than one models overlap. If a collision is detected, the models involved are tested based on their interior intensity statistics. Since the model textures are represented using nonparametric p.d.fs, the Kullback-Leibler Divergence can be used to measure the dissimilarity between two p.d.fs. Suppose a collision

is detected between model A with intensity probability density p_A and model B with probability density p_B, then the Kullback-Leibler Divergence between the two distributions is defined by:

$$D_{p_A \| p_B} = \int_U p_A(i) log \frac{p_A(i)}{p_B(i)} di$$

where U denotes the set of all intensity values. If this K-L distance is sufficiently small, the two models will be merged. Because we use the implicit model shape representation, the new model representation can be easily derived from the two merging model representations. At each pixel in the image domain, the new model embedding space (distance) value is the smaller one of the two embedding space values at that pixel in the two merging models. This smaller value is exactly the min distance to the union of those two models. And the new model shape is the zero level set of this new model implicit representation. The new model interior intensity statistics can be updated accordingly.

4.1. The weighting factor between the energy terms

Because of the ubiquitous presence of noise, intensity inhomogeneity, and texture in medical images, information from either gradient or intensity is not sufficient to achieve reliable segmentation. In our framework, the boundary and region information are naturally integrated in a common framework, and the three energy terms work together to overcome the local minima to some extent. The weighting factors a,b and c between the three terms, which imply two degrees of freedom, need to be assigned with care in order to achieve good performance. In the current protocol we use, the weight factors are automatically adjusted after each model evolution iteration. Without loss of generality, let $a = 1$, we always assign a higher weight for model boundary pixels, thus $b = ka, k > 1$. The relative weighting between the texture term and the shape terms is determined by a confidence measure, C_e, of the computed edge map. To decide this confidence value, we first compute the image intensity probability map, P_I, based on the model interior intensity statistics. Initially, a small threshold is applied on P_I to produce a binary image BP_I, in which pixels with intensity probability higher than the threshold have value 1. The connected component on this binary image around the model is computed. Then the confidence value C_e is determined based on the complexity of the original image gradient or edge map in this connected component area. The confidence value is low if there are high gradient and edges inside the region; the value is high, otherwise. Then we set the value for the weighting factor $c = \frac{1}{C_e}$.

In [Fig. (7)], we show the result of segmenting the Endocardium of the left ventricle in a noisy tagged MR heart image, using our framework and the weighting factors described above. Note that, due to the tagging lines, the intensity inhomogeneity, the detected boundary edges of the object are fragmented, and there are spurious small edges inside the region. In this case, the integration of both shape and texture information is critical in helping the model out of the local minima.

The model evolution is computationally efficient, due to our use of the global similarity and local FFD parameterization of the model. In the example shown in [Fig. (7)], the segmentation process takes less than $200ms$ to converge on a 2Ghz PC station.

5. Discussions and Conclusions

In this paper, we have presented a new class of deformable models, MetaMorphs, which possess both boundary shape and interior intensity statistics. In our framework, the boundary and region information are coupled coherently to drive the deformations of the models towards object boundaries.

The main novelty and contribution of the MetaMorph deformable models lie in several aspects. First, unlike traditional deformable models, MetaMorphs have both shape and texture statistics. Their shapes are implicitly represented in a higher dimensional space, and the interior statistics are approximated using a non-parametric kernel-based method, which allows flexible and differentiable distributions. Second, the models can undergo a hierarchy of global and local deformations. The local deformations are parameterized by the parameters of a cubic B-spline based Free Form Deformations model, which enables fast and smooth model evolution. Thirdly, the model interior intensity statistics impose model deformation constraints from the on-line learning process, without using any prior information from supervised learning before the segmentation. Lastly, the algorithm can be straightforwardly applied in 3D, and can handle efficiently the merging of multiple models that are evolving simultaneously.

In our future work, we will validate the segmentation results with our method, using ground truth data sets, and conduct more principled and quantitative study in assigning the weight factors between the energy function components.

Appendix

We can analytically derive the partial derivatives $\frac{\partial}{\partial \mathbf{q}_i} D(\mathbf{q}; \mathbf{x})$ for the local incremental B-spline FFD parameters in $\mathbf{q}_l$:

$$\delta P_{m,n} = (\delta P^x_{m,n}, \delta P^y_{m,n});\ m = 1, ..., M, \quad n = 1, ..., N$$

Without loss of generality, one can consider the $(m, n)th$ control point and its deformations in both directions. Then, from the definition for the deformations $D(\mathbf{q}; \mathbf{x})$, we have for each element $\mathbf{q}_i$ in $\mathbf{q}_l$, $\frac{\partial}{\partial \mathbf{q}_i} D(\mathbf{q}; \mathbf{x}) = \frac{\partial}{\partial \mathbf{q}_i} L(\mathbf{q}_l; \mathbf{x})$, and

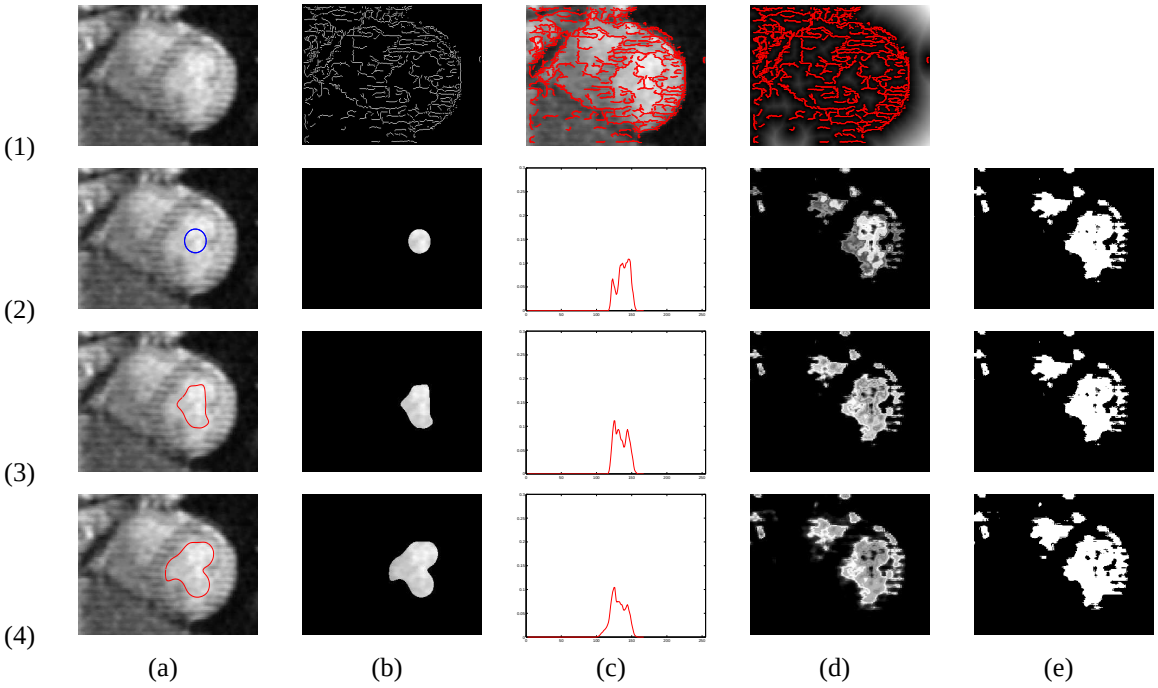


Figure 7: The tagged MR heart image. (1)(a) The original image. (1)(b) The edge map. (1)(c) The edge points overlaid on original image. (1)(d) The "shape image". (2) Initial model. (3) Intermediate result. (4) Final model (after 50 iterations). (2-4)(a) The evolving model. (2-4)(b) The model interior. (2-4)(c) The model interior intensity probability density. (2-4)(d) The intensity probability map of the image based on the p.d.f in (c). (2-4)(e) Thresholded intensity probability map. The threshold is the mean probability.

the following relations hold:

$$\frac{\partial\, \delta L(\mathbf{q}_l;\mathbf{x})}{\partial \delta P^x_{m,n}} = \begin{cases} \begin{bmatrix} B_{m-i}(u)\, B_{n-j}(v) \\ 0 \end{bmatrix}, & 0 \le m-i, n-j \le 3 \\ \mathbf{0}, & otherwise \end{cases}$$

$$\frac{\partial\, \delta L(\mathbf{q}_l;\mathbf{x})}{\partial \delta P^y_{m,n}} = \begin{cases} \begin{bmatrix} 0 \\ B_{m-i}(u)\, B_{n-j}(v) \end{bmatrix}, & 0 \le m-i, n-j \le 3 \\ \mathbf{0}, & otherwise \end{cases}$$

References

[1] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. In *IEEE International Conference in Computer Vision*, pages 694–699, 1995.

[2] T. Chen and D.N. Metaxas. Image segmentation based on the integration of markov random fields and deformable models. In *Medical Imaging Copmuting and Computer-Assisted Intervention*, pages 256–265, 2000.

[3] L. D. Cohen and I. Cohen. Finite-element methods for active contour models and balloons for 2-d and 3-d images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15:1131–1147, 1993.

[4] S. Jehan-Besson, M. Barlaud, and G. Aubert. Shape gradients for histogram segmentation using active contours. In *IEEE International Conference in Computer Vision*, pages 408–415, 2003.

[5] T.N. Jones and D.N. Metaxas. Automated 3d segmentation using deformable models and fuzzy affinity. In *Information Processing in Medical Imaging*, pages 113–126, 1997.

[6] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1:321–331, 1987.

[7] D. Metaxas. *Physics-Based Deformable Models*. Kluwer Academic Publishers, 1996.

[8] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42(5):577–685, 1989.

[9] N. Paragios and R. Deriche. Geodesic active regions for supervised texture segmentation. In *IEEE International Conference in Computer Vision*, pages 926–932, 1999.

[10] R. Ronfard. Region-based strategies for active contour models. *ijcv*, 13(2):229–251, 1994.

[11] T. W. Sederberg and S. R. Parry. Free-form deformation of solid geometric models. In *Proceedings of the 13th Annual Conference on Computer Graphics*, pages 151–160, 1986.

[12] L. H. Staib and J. S. Duncan. Boundary finding with parametrically deformable models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(11):1061–1075, 1992.

[13] S. Zhu and A. Yuille. Region competition: Unifying snakes, region growing, and bayes/mdl for multi-band image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(9):884–900, 1996.