From Global to Local, a Continuum of Shape Models with Fractal Priors *

B. C. Vemuri[1] and A. Radisavljevic[2]
[1]Department of Computer & Information Sciences
[2]Department of Electrical Engineering
University of Florida, Gainesville, FL 32611

Abstract

3D Shape modeling has been a very prominent part of Computer Vision over the past decade. Several shape modeling techniques have been proposed in literature, some are local (distributed parameter) while others are global (lumped parameter) in terms of the parameters required to describe the shape. Hybrid models that combine both ends of this parameter spectrum have been in vogue only recently. However, they do not allow a smooth transition between the two extremes of this parameter spectrum.

In this paper, we introduce a *new shape modeling scheme* that can *transform smoothly from local to global* models or vice-versa. The modeling scheme utilizes a hybrid primitive called the deformable superquadrics *constructed in an orthonormal wavelet bases*. These multi-resolution bases provide the power to continuously transform from local to global shape deformations and thereby allow for a continuum of shape models – from those with local to those with global shape descriptive power – to be created. The characteristic of continuously transforming from local to global shape deformations allows us to generate fractal surfaces of arbitrary degree that can be useful in describing natural detail.

We embed these multi-resolution shape models in a probabilistic framework and use them for segmenting anatomical structures in the human brain from MRI data. A salient feature of our modeling scheme is that it can naturally allow for the incorporation of prior statistics of a rich variety of shapes. This stems from the fact that, unlike other modeling schemes, in our modeling, we require relatively few parameters to describe a large class of shapes completely.

*This work was supported in part by the University of Florida Brain Institute

1 Introduction

Modeling shapes is an important and integral part of Computer Vision as well as Computer Graphics. In Computer Vision, modeling shapes is important for shape reconstruction and shape recognition from sensed data, while in Computer Graphics, the application is for shape synthesis. Many shape modeling schemes have been proposed in the literature. In Computer Vision, the motivation for modeling shapes has been primarily driven by either shape reconstruction or shape recognition tasks. Shape reconstruction from sensed data requires a broad geometric coverage. The models must recover detailed structure from noisy data using only the weakest of the possible assumptions about the observed shapes. Generalized spline models with continuity constraints are well suited for fulfilling the goals of shape reconstruction (see [16, 4]). Generalized splines are the key ingredient of the dynamic shape modeling paradigm introduced by Terzopoulos et. al., [18]. Incorporating dynamics into shape modeling with generalized splines, allows for creating realistic animation for computer graphics applications and also for tracking moving objects in computer vision applications. With the advent of the dynamic shape modeling paradigm, there was a flurry of research activity in this area, with numerous application motivated modifications to the modeling primitives, and the external forces derived from data [20, 21, 22, 8, 5]. However, the distributed nature of these models – each point of the spline surface potentially contributes three spatial degrees of freedom – makes them unsuitable for shape recognition tasks. Shape recognition – a high-level vision task – requires that the shape models be characterized by a small set of parameters. Lumped parameter models are thus better suited for such tasks. Some examples of lumped parameter models include, superquadrics [3, 12, 2] and the parametrically deformable models [23, 9].

Recently, a hybrid modeling scheme dubbed, "deformable superquadrics" was introduced by Terzopoulos et al., [19]. These models have the advantage of combining the descriptive power of lumped and distributed parameter models and thus, simultaneously satisfying the requirements of shape reconstruction and shape recognition. However, these hybrid models do not exhibit a smooth transition in the number of parameters required for their description.

For example, the two extremes of the spectrum are occupied by lumped models at one end and distributed models at the other. The former requiring, few parameters and the later requiring a large number of parameters for their description. The modeling scheme described in [19] involves a superposition of the lumped and distributed parameters to completely characterize the plethora of generated shapes. Although, the hybrid models can be used to generate both local and global shape deformations, they *do not depict a smooth transition between these types of deformations.* In addition, there is *no smooth transition in the number of parameters* that can capture descriptions of the range of generated shapes.

In this paper we introduce a new multi-resolution dynamic modeling scheme that uses an orthonormal wavelet basis [11, 7]. Our models have the unique property of being able to smoothly scale up or down the range (from local to global) of possible deformations and the number of parameters required to characterize them. This is achieved by adding to the global (lumped) parameters of the deformable superquadrics, a set of wavelet coefficients of the membrane spline at any resolution. Earlier work reported in Pentland [13] on the use of wavelets in the context of shape modeling was limited to surface reconstruction. He used wavelets basis to approximate the Eigen vectors of the stiffness matrix assembled in the finite element solution to the surface reconstruction problem using the thin-plate-membrane splines. The main idea of his work was to speed up the solution to the surface reconstruction problem by a change of basis, from regular nodal basis to the natural basis represented by the Eigen vectors of the stiffness matrix. Wavelets basis were used as an approximation to the natural basis i.e., the Eigen vectors of the stiffness matrix. In the wavelet basis, the stiffness matrix is approximately diagonal and this approximate diagonalization leads to savings in storage space as well as computation time for the solution. In our research, the key focus is on development of a general shape modeling scheme which when embedded in a probabilistic framework can be utilized for easily incorporating a larger class of shapes as prior information. In addition, we also emphasize on the efficient computation of the model fitting process to the data.

The thin-plate-membrane regularizers were also used for a very different purpose by Szeliski et. al., [15] namely, generation of constrained fractals which are useful in representing natu-

ral images. They derive the fractal characteristic of a membrane and thin-plate regularizers and show that the membrane splines give raise to fractals of degree two while the thin-plate splines yield degree four fractals. They approximate a fractal of in-between degree by using a blend of the membrane and thin-plate models. The modeling scheme we introduce in this paper allows us to generate fractals of in-between degree without having to resort to approximations with a blend of the plate membrane models. We achieve this by modulating the frequency characteristics of the wavelet basis with a polynomial of appropriate degree. In order to generate stochastic fractal surfaces, we convert the energies into probability distributions using ideas from statistical mechanics [10] and employ a stochastic multi-resolution algorithm. We will present some examples of the synthesized fractal surfaces in a subsequent section.

The rest of the paper is organized as follows: In section 2 we briefly describe the geometry of the modeling scheme and the multi-resolution basis for expressing the geometry, followed by a new way to augment the global shape of the deformable superquadrics primitive for achieving the smooth transition from global to local shape descriptions. Section 3 contains discussion on how to embed our models in an probabilistic framework and the derivation of the mechanics of the modeling scheme. In section 4 we derive the internal energy of the prior model in the wavelet basis and present the fractal characteristics of this prior model with the aid of synthesized fractal surfaces. Section 5 contains a description of, the unsupervised learning technique that we use for incorporating specific a priori knowledge into our modeling framework and, implementation results on real MRI data followed by conclusions in section 6.

2 Multi-resolution Geometry of the Modeling Scheme

In this section, we will briefly describe the geometry and the construction of a multi-resolution wavelet basis for the modeling scheme. To dramatically increase the library of shapes that can be modeled and to achieve a smooth transition from global to local shape descriptions, we will present a *new method* that will augment the global parameters of the deformable

superquadrics.

2.1 Geometry of the Deformable Superquadrics

The deformable superquadrics are closed surfaces in space whose intrinsic coordinates $\mathbf{u} = (u, v)$ are defined on a domain Ω. The positions of points on the model in an inertial frame of reference are given by a vector-valued function $\mathbf{x}(\mathbf{u}) = (x_1(\mathbf{u}), x_2(\mathbf{u}), x_3(\mathbf{u}))^T$. In a model-centered coordinate frame, the position vector $\mathbf{x}$ becomes

$$\mathbf{x} = \mathbf{c} + \mathbf{R}\mathbf{p} \tag{1}$$

where $\mathbf{c}$ is the location of the center of the model, and the rotation matrix $\mathbf{R}$ specifies the orientation of the model-centered coordinate frame. Hence, $\mathbf{p}(\mathbf{u})$ denotes the canonical positions of points on the model relative to the model frame. We further express $\mathbf{p}$ as the sum of a reference shape $\mathbf{s}(\mathbf{u})$ and a displacement $\mathbf{d}(\mathbf{u})$ namely $\mathbf{p} = \mathbf{s} + \mathbf{d}$. For the parameterized reference shape $\mathbf{s}(\mathbf{u})$ we use the superquadrics with the bending deformation (as in [2]), due to it's attractive global shape characterization. The reference shape $\mathbf{s}$ is given by,

$$\mathbf{s} = \mathcal{B}(\tilde{\mathbf{s}}) \tag{2}$$

Where $\mathcal{B}$ is a bending deformation, a function of two parameters: the radius of curvature k and angle of the bending plane α (see [2] for details). The parametric equation of a superquadric $\tilde{\mathbf{s}}$ is given by

$$\tilde{\mathbf{s}} = \begin{pmatrix} a_1 C_u^{\epsilon_1} C_v^{\epsilon_2} \\ a_2 C_u^{\epsilon_1} S_v^{\epsilon_2} \\ a_3 S_u^{\epsilon_1} \end{pmatrix} \tag{3}$$

Where, $-\pi/2 \leq u \leq \pi/2$ $-\pi \leq v \leq \pi$, and $S_w^{\epsilon} = sgn(sin\, w)|sin\, w|^{\epsilon}$ and $C_w^{\epsilon} = sgn(cos\, w)|cos\, w|^{\epsilon}$. Here, $0 \leq a_1, a_2, a_3 \leq 1$ are aspect ratio parameters, $\epsilon_1, \epsilon_2 \geq 0$ are "squareness" parameters. This defines the geometry of our underlying reference shape namely, the superquadric. We collect all these shape parameters along with the two bending parameters into a vector $\mathbf{q}_s$. We will denote $\mathbf{c}$ by the vector $\mathbf{q}_c$ and use $\mathbf{q}_\theta$ to denote the vector of rotational coordinates of the model in a quaternion representation (see [19]).

An important aspect of our reference shape which differs from previous work of Terzopoulos et al., [19] and Bajscy et al., [2], is that we have constructed a *uniform tessalation* of

the superquadric reference shape. Superquadric shapes exhibit anomalous behavior when the squareness parameters ϵ_1, ϵ_2 depart from unity, causing grid lines to become unevenly spaced leading to distortion of the deformational properties of the superimposed membrane surface and inefficient use of computational resources. In addition, numerical instabilities may be caused in the computation of derivatives due to the large grid spacing. A Uniform tessalation of the superquadric can be used to overcome these problems. To achieve the uniform tessalation, we first construct uniformly spaced grid lines in the parameter direction u and apply the same technique to achieve uniform spacing along the parameter direction v. Thus, we seek a function of the parameter u (same argument applies to the parameter v) which when substituted into the equation of the superquadric i.e., equation 3, would yield a constant differential arc length increment for a constant differential increment in u. For a horizontal elliptical contour,

$$ds = \sqrt{(x\prime)^2 + (y\prime)^2}du \tag{4}$$

Setting $ds/du = constant = (circumference/2\pi)$ and using $(x = a_1 cos^{\epsilon}(f(u)),\ y = a_2 sin^{\epsilon}(f(u)),\ z = 0)$ results in an ordinary differential equation for $f(u)$

$$\frac{df}{du} = K\frac{\mid tan(f) \mid}{\epsilon\sqrt{a_1^2 \mid tan(f) \mid^4 \mid cos(f) \mid^{2\epsilon} + a_2^2 \mid sin(f) \mid^{2\epsilon}}} \tag{5}$$

$$f(0) = 0 \ \ as\ the\ initial\ condition \tag{6}$$

Where K is a constant. This ordinary differential equation can be solved numerically. In figure 1, a superquadric with its natural *nonuniform* tessalation is shown side by side to the same superquadric *uniformly* tessellated using the method described above.

In the following sections, we show that the displacement vector function $\mathbf{d} = \Phi S \mathbf{q}_d$ is completely determined by a set of wavelet coefficients collected into the vector $\mathbf{q}_d$, with S corresponding to the discrete wavelet transform and Φ containing nodal interpolation functions. Thus, the state of the model is fully expressed by the vector $\mathbf{q} = \left[\mathbf{q}_c^T, \mathbf{q}_\theta^T, \mathbf{q}_s^T, \mathbf{q}_d^T\right]^T$.

2.2 Multi-resolution Analysis and Wavelet Bases

Multi-resolution analysis has been quite popular in computer vision research for the past decade. Most of the research in the past has dealt with speeding up computations in solu-

Figure 1: Tessellations of a Superquadric, (a) nonuniform (b) uniform

tions of various low-level vision problems using multi-resolution techniques [17]. With the advent of the *wavelet transform (WT)*, multi-resolution analysis of functions has become more attractive due to the fact that wavelet decomposition of a function (signal) into different frequency components allows for studying each component at a resolution matched to its scale. In the following, we will very briefly describe the construction of a wavelet based multi-resolution approximation in 1D and its extension to a 2D function specifically the membrane displacement function $\mathbf{d}(u, v)$.

2.2.1 Detail Function from Multi-resolution Approximations

A multi-resolution approximation of a scalar valued function $d(u) \in L^2(\Re)$ at a resolution j is obtained by projecting $d(u)$ into a vector subspace $V_j \subset L^2(\mathcal{R})$ where integer (k) translates $\phi_{jk}(u) = \sqrt{2^{-j}}\phi(2^{-j}u - k)$ of a dilated scaling function $\phi(u)$ form the orthonormal basis of V_j. The approximation $\hat{d}^j(u)$ is expressed as :

$$\hat{d}^j(u) = \sum_k \alpha_{jk}\phi_{jk}(u) \tag{7}$$

Where the set of coefficients $\alpha_{jk} = \langle d(u), \phi_{jk} \rangle$ for all integer k corresponds to the discrete characterization of this approximation.

An important property of multi-resolution approximations is that all the necessary information to compute the function at a smaller resolution $j + 1$ is contained in the function at

Figure 2: Division of the frequency domain for a) multi-resolution and b) wavelet representation.

resolution j. In other words the subspaces V_j satisfy the containment hierarchy,

$$.....V_3 \subset V_2 \subset V_1 \subset V_0 \tag{8}$$

The scaling function ϕ has the effect of a low-pass filter with the cutoff frequency decreasing proportionally to 2^{-j} (see figure 2(a)). The key idea of the wavelet basis is to span the *difference spaces* W_{j+1} between the two approximations V_j and V_{j+1} such that $W_{j+1} \perp V_{j+1}$. A wavelet decomposition therefore consists of splitting the function at resolution j into the *detail function* in W_{j+1} and the next lower (j+1) resolution approximation in V_{j+1}. This leads directly to a scale wise, orthogonal decomposition of $L^2(\Re)$. To represent the function $d(u)$ of a finite resolution we can assume $d(u) \in V_0$, where V_0 is given by,

$$.....W_3 \oplus W_2 \oplus W_1 = V_0 \tag{9}$$

with $\oplus$ denoting the direct sum operator. The *hierarchical* orthonormal basis for d(u) is given by a family of dilations and translations of a wavelet function $\psi(u) : \psi_{jk}(u) = \sqrt{2^{-j}}\psi(2^{-j}u - k)$. Thus, $d(u)$ can be be written in this basis as,

$$d(u) = \sum_{j,k} \delta_{jk}\psi_{jk}(u) \tag{10}$$

Where the discrete wavelet coefficients are computed as $\delta_{jk} = \langle d(u), \psi_{jk} \rangle$.

As in the case of the multi-resolution approximation, wherein a low pass filtering interpretation was given to the scaling function, a filtering interpretation can also be given to the WT. At each scale j the WT corresponds to a logarithmically spaced band-pass filter (see Figure 2(b)). Our work in this paper is primarily based on this filter bank interpretation of the WT as discussed in Rioul et. al. [14].

2.2.2 Extension to 2-D and Fast Implementation

Extension to 2D of the multi-resolution approximation using wavelet basis can be easily achieved using separable wavelets obtained from the products of one-dimensional wavelets and scaling functions [11]. For the multi-resolution approximation of the membrane $\mathbf{d}(u,v)$ we use the tensor products of ϕ and ψ given by (see Mallat [11]),

$$\left(\begin{array}{lcl} \psi^{LH}_{jkl}(u,v) & = & \phi(u)_{jk}\,\psi(v)_{jl} \\ \psi^{HL}_{jkl}(u,v) & = & \psi(u)_{jk}\,\phi(v)_{jl} \\ \psi^{HH}_{jkl}(u,v) & = & \psi(u)_{jk}\,\psi(v)_{jl} \end{array} \right)_{(j,k,l)\in\mathcal{Z}} \tag{11}$$

Note that, since the displacement $\mathbf{d}$ is a vector valued function, we have to repeat the tensor product approximations shown above for each component of the displacement vector. The superscript H (L) represents high (low) pass filtering in each of the directions u and v in the parameter domain Ω, and the indices k and l represent the integer shifts (as in the 1D case) in u and v directions respectively. We collect the coefficients of the inner product of the basis and the function $\mathbf{d}(u,v)$ i.e., δ^{LH}_{jkl}, δ^{HL}_{jkl} and δ^{HH}_{jkl} into a vector q_d as illustrated in Figure 3. This displacement vector $\mathbf{q}_d$ contains a complete hierarchical description of the membrane surface. Note that $\mathbf{q}_d$ has the same number of elements as the original nodal discretization at $j=0$. For our WT implementation we adopted the specific $\phi(u)$ and $\psi(u)$ functions as given in Mallat [11], which are based on the cubic spline interpolants. If we denote the original nodal discretization at resolution $j=0$ by a vector $\boldsymbol{\alpha}_0$ then, we have,

$$\left. \begin{array}{lcll} \boldsymbol{\alpha}_0 & = & S\,\mathbf{q}_d & \text{Reconstruction} \\ \mathbf{q}_d & = & S^T\,\boldsymbol{\alpha}_0 & \text{Decomposition} \\ \mathbf{d}(u,v) & = & \Phi(u,v)\,\boldsymbol{\alpha}_0 & \text{Interpolation} \end{array} \right\} \tag{12}$$

where Φ represents a vector of interpolation functions (as in the 1D case, see equation 7) and multiplication by the *orthogonal* matrix S corresponds to performing the WT. A fast

Figure 3: a) Organization of the displacement vector $\mathbf{q_d}$, b) separation of the "global" and "local" parts of $\mathbf{q_d}$

implementation of the WT is achieved using Quadrature Mirror Filters (QMFs). A self explanatory block diagram of the filterbank interpretation of the QMF is presented in figure 4 (see Mallat [11] for details). Figure 4 shows an efficient way to implement the WT, with a computational cost of order $O(N)$ (where N is the number of grid nodes).

Finally, vector $\mathbf{q}_d$ is expressed in an orthonormal wavelet basis. Along with other parameters of the model it is subjected to the iterative minimization process as discussed subsequently.

2.3 From Global to Local Shapes with Wavelet Coefficients

In figure 3, we explicitly depict the hierarchical decomposition of the displacement vector $\mathbf{q}_d$ in an orthonormal wavelet basis. The coefficients δ_j toward the upper left corner are fewer then those toward the bottom-right corner, i.e. the corresponding grid nodes are spaced farther apart and the basis functions are coarser. Few parameters needed in the description and ability to depict the global shape gives the low frequency components of the wavelet representation a "semi-global" character. Depending on the application we can draw the boundary at an arbitrary j_0 to separate the global and local part of vector $\mathbf{q}_d$,

$$\left.\begin{array}{lcl} \mathbf{q}_d & = & [\mathbf{q}_{dg}\ ,\ \mathbf{q}_{dl}] \\ \mathbf{q}_{dg} & = & [\delta_n, \cdots, \delta_{j_0}] \\ \mathbf{q}_{dl} & = & [\delta_{j_0-1}, \cdots, \delta_1] \end{array}\right\} \tag{13}$$

Figure 4: a) Decomposition and b) reconstruction of a 2-dimensional array adopted from Mallat[89]

where $\mathbf{q}_{dg}$ contains the top (coarser) portion of the multi-resolution pyramid. One of the useful aspects of including $\mathbf{q}_{dg}$ into the global description along with superquadric parameters is the greater variety of shapes we can generate, such as *coarse dips, bumps* etc. (see figure 5). For a comparison of possible global deformations generated via tweaking of superquadric parameters and the coarse level wavelet coefficients of the displacement function, we illustrate the deformed shapes adjacent to one another. Figure 5 depicts the original reference shape on the top left hand corner, global deformation resulting from tweaking a superquadric parameter is shown on the top right hand corner. The bottom two shapes are generated via global deformations caused by tweaking the coarse level wavelet coefficients.

When we build our prior model with mean values and covariances on $\mathbf{q}_s$ and $\mathbf{q}_{dg}$ the model assumes this enhanced shape as it's new rest state and it deforms away from this shape under the influence of sensed data. The need for incorporating more specific information such as this, into the prior model is especially present in difficult segmentation problems encountered in Medical Imaging applications eg. segmentation of anatomical shapes from MR images. We will present two such examples in this paper.

Figure 5: a) Global shape deformations generated via tweaking global parameters of the superquadric model, b) global deformation generated by tweaking of coarse scale wavelet coefficients

3 Probabilistic Framework and Model Mechanics

In the last section, we have described our multi-resolution deformable superquadric model in a wavelet basis. The multi-resolution wavelet basis provides a computationally efficient solution to the problem of fitting the deformable superquadric model to data sets and also provides us a way to enhance the library of shapes that can be described using a few global parameters. In this section, we will present a way to incorporate specific prior knowledge about the shape to be extracted from the data and then derive the mechanics of our modeling scheme.

We cast the model fitting process in a probabilistic framework and incorporate prior distributions on the vector of geometric parameters that we are estimating. The combination of prior and sensor models results in a Bayesian model, since the posterior distribution of the parameters we are trying to estimate conditioned on the data can be computed using Baye's rule. In this paper, we use the multi-resolution physically-based shape models described in earlier sections as the prior models. We collect the statistics on the global parameters of the

model via a training process described in a subsequent section and for the local parameters, we convert the continuous strain energy that governs the deformation of the deformable superquadric model away from its natural shape into a probability distribution over expected shapes, with lower energy shapes being the more likely. This is achieved via a technique of statistical mechanics – conversion of energies into probabilities using Boltzmann, or Gibbs, distribution [10]. The link between physically-based (deformable models) models and priors is conveniently established using a Gibbs distribution of the form

$$p(\mathbf{q}) = \frac{1}{Z_p} exp(-E_p(\mathbf{q})), \tag{14}$$

where $E_p(\mathbf{q})$ is the discretized version of the internal smoothness energy of the model, and Z_p (called the *partition function*) is a normalizing constant. What was originally an elastic energy restoring a model towards a rest state now becomes a probability distribution over expected shapes, with lower energy shapes being more likely. To complete the formulation of the estimation problem, we combine this prior model with a simple sensor model based on linear measurements with Gaussian noise

$$p(\mathbf{D}/\mathbf{q}) = \frac{1}{Z_D} exp(-E_D(\mathbf{D}, \mathbf{q})), \tag{15}$$

where $E_D(\mathbf{D}, \mathbf{q})$ can be either an edge-based potential energy synthesized from the MR images or the potential energy in the springs attached from 3D data points to the surface of the deformable superquadric constraining it to conform to the observed data. Combining the prior and sensor models (from equations 14 and 15 respectively) using Baye's rule, we obtain the posterior distribution

$$p(\mathbf{q}/\mathbf{D}) = \frac{p(\mathbf{D}/\mathbf{q})p(\mathbf{q})}{p(\mathbf{D})} = \frac{1}{Z} exp(-E(\mathbf{q})), \tag{16}$$

where

$$E(\mathbf{q}) = E_p(\mathbf{q}) + E_D(\mathbf{D}, \mathbf{q}). \tag{17}$$

Computing the *Maximum A Posteriori* (MAP) estimate, i.e., the value of $\mathbf{q}$ that maximizes the conditional probability $p(\mathbf{q}/\mathbf{D})$, provides the same result as finding the minimum energy

configuration of the physically-based model. However, the advantage of the probabilistic framework being that, it allows for explicitly incorporating prior model and sensor model characteristics in addition to providing an uncertainty in the estimates.

The internal strain energy E_p of the model in our case is given by a quadratic form, $E_p = 1/2[(\mathbf{q} - \bar{\mathbf{q}})^T K_p (\mathbf{q} - \bar{\mathbf{q}})]$. Where, $\bar{\mathbf{q}}$ is the rest state of the model and the matrix K_p corresponds to the stiffness of deformation. We will discuss the internal energy E_p further in a subsequent section. The two types of sensor model energies for $E_D(\mathbf{D}, \mathbf{q})$ discussed earlier are expressed as

$$E_D(\mathbf{D}, \mathbf{q}) = \begin{cases} \frac{1}{2} \sum_i \beta (\mathbf{D}_i - \mathbf{x}(\mathbf{q}, \mathbf{u}_i)^2 & \text{spring energy} \\ \int_\Omega \beta \left| P(\mathbf{x}(\mathbf{q}, \mathbf{u})) \right| & \text{image potential} \end{cases} \tag{18}$$

In the above equations $\mathbf{u}_i$ determines the closest point $\mathbf{x}$ to the given measured point $\mathbf{D}_i$. We adjust the locations of these grid coordinates as the shape of the model evolves (see [19], for more on migrating point of influence). The magnitude of parameter β is related to the uncertainty of sensor measurements. $P(\mathbf{x}(\mathbf{q}, \mathbf{u}))$ denotes an edge based potential derived from the MR image data. Maximization of of the posteriori amounts to minimizing the total energy $E = E_p + E_D$ for which we employ a gradient descent method that amounts to iteratively solving,

$$\mathbf{q}^{k+1} = \mathbf{q}^k - \Lambda \frac{\partial E(\mathbf{q})}{\partial \mathbf{q}} \tag{19}$$

For the derivative of the energy on the right hand side in the above equation, we use the chain rule to get $\frac{\partial E}{\partial \mathbf{q}} = \frac{\partial E}{\partial \mathbf{x}} \frac{\partial \mathbf{x}}{\partial \mathbf{q}}$, where the contribution of each parameter group is

$$\begin{aligned} \frac{\partial \mathbf{x}}{\partial \mathbf{q}} &= \left[\frac{\partial \mathbf{c}}{\partial \mathbf{q}_c}, \frac{\partial (\mathbf{R}\mathbf{p})}{\partial \mathbf{q}_\theta}, \frac{\mathbf{R} \partial \mathbf{s}}{\partial \mathbf{q}_s}, \frac{\mathbf{R} \partial (\Phi \mathbf{S} \mathbf{q}_d)}{\partial \mathbf{q}_d} \right] && (20) \\ &= [\, \mathbf{I} \;\; \mathbf{B} \;\; \mathbf{RJ} \;\; \mathbf{\Phi S} \,] && (21) \\ &= \mathbf{L} && (22) \end{aligned}$$

Here, we adopt the notation as in [19]. Partial derivatives of the prior model and data energies defined above can be now written as

$$\frac{\partial E_p}{\partial \mathbf{q}} = K_p (\mathbf{q} - \bar{\mathbf{q}}) \tag{23}$$

and

$$\frac{\partial E_D}{\partial \mathbf{q}} = \mathbf{f_q} = \begin{cases} -\sum_i \beta(\mathbf{D}_i - \mathbf{x}(\mathbf{q}, \mathbf{u}_i)L & \text{spring energy} \\ \int_\Omega \beta \nabla \left| P(\mathbf{x}(\mathbf{q}, \mathbf{u})) \right| L & \text{image potential} \end{cases} \tag{24}$$

By substituting equations 23 and 24 into the equation for gradient descent 19 we obtain an expression for iterative updating of the state vector $\mathbf{q}$

$$\mathbf{q}^{(k+1)} = \mathbf{q}^{(k)} + \Lambda \left(\mathbf{f}_q^{(k)} - K(\mathbf{q}^{(k)} - \bar{\mathbf{q}} \right) \tag{25}$$

where Λ is a diagonal matrix containing step sizes and consequently effecting the speed of convergence. Close similarity of this evolution equation to the one derived in [19] reveals that equation 25 can be modified slightly and effectively used to generate motion sequences corresponding to first order dynamics $C\dot{\mathbf{q}} + K(\mathbf{q} - \bar{\mathbf{q}}) = \mathbf{f}_q$. In addition to producing the statistically most likely state $\hat{\mathbf{q}}$, this framework provides the uncertainty in the estimate via the associated covariance matrices. In our work, we employ the above gradient descent procedure in a multi-resolution framework for computational efficiency as well as alleviation from entrapment in local minima.

4 Prior Model in Wavelet Basis

In the previous section we described a probabilistic framework for embedding our model. We now derive the explicit form of the internal energy of the model in the wavelet basis. This energy is then converted to a Gibbsian probability distribution using the technique discussed in section 3.

4.1 Internal Energy of the Prior Model

In addition to generic smoothness assumptions, a prior model should incorporate the information about the possible shapes that an observed object can assume. This can be very important in applications involving segmentation of known natural shapes from "unfriendly" (low contrast and noisy) images. We express the probability of a given configuration $\mathbf{q}$ through a Gibbs distribution (see section 3), which assigns a low probability to configurations with large energy E_i and vice versa. The configuration with $E_i = 0$ represents the

rest state of the model. The internal energy of the membrane spline (which augments the reference superquadric shape s of the model), in the continuous form is given by,

$$E_i = \int w_0 \mathbf{d}^2 + w_1 \left(\left(\frac{\partial \mathbf{d}}{\partial u} \right)^2 + \left(\frac{\partial \mathbf{d}}{\partial v} \right)^2 \right) d\mathbf{u} \tag{26}$$

With no additional prior information about the shapes being modeled, the underlying global shape of the model namely, the superquadric, can be interpreted as the mean value or the rest state of the membrane surface – expressed by the displacement function $\mathbf{d}$. If we rewrite the deformation energy in terms of vector $\mathbf{p} = \mathbf{s} + \mathbf{d}$ we have an equivalent expression to equation 26,

$$E_i = \int w_0 (\mathbf{p} - \mathbf{s})^2 + w_1 \left(\left(\frac{\partial (\mathbf{p} - \mathbf{s})}{\partial u} \right)^2 + \left(\frac{\partial (\mathbf{p} - \mathbf{s})}{\partial v} \right)^2 \right) d\mathbf{u} \tag{27}$$

Where the role of the global (reference) shape $\mathbf{s}$ as the rest state becomes apparent. We extend our underlying global shape $\mathbf{s}$ with a low resolution mean displacement function $\bar{\mathbf{d}} = \hat{\mathbf{d}}^{j_0}$ corresponding to parameters $\mathbf{q_{dg}}$ as defined in equation 13 and shown in figure 3. This is done by altering the deformation energy equation 26 to,

$$E_i = \int w_0 (\mathbf{d} - \bar{\mathbf{d}})^2 + w_1 \left(\left(\frac{\partial (\mathbf{d} - \bar{\mathbf{d}})}{\partial u} \right)^2 + \left(\frac{\partial (\mathbf{d} - \bar{\mathbf{d}})}{\partial v} \right)^2 \right) d\mathbf{u} \tag{28}$$

The discrete version of this internal strain energy can be written as,

$$E_i = \frac{1}{2} (\mathbf{q}_d - \bar{\mathbf{q}}_d)^T \; K_d \; (\mathbf{q}_d - \bar{\mathbf{q}}_d) \tag{29}$$

Where, K_d is the stiffness matrix that determines the elastic properties of the model and $\bar{\mathbf{q}}_d{}^T = [\bar{\mathbf{q}}_{dg}{}^T, 0, \ldots, 0]$ incorporates the statistics of the possible shapes that can be encountered in the segmentation problem. These statistics are gathered through training process described in section 5. It is evident that the number of parameters in $\bar{\mathbf{q}}_{dg}$ is low, i.e. the representation of the model remains compact and thus making the training process feasible. We include other global shape parameters of the model in our prior statistics by extending the above energy as follows,

$$\begin{aligned} E_i \;\; = \;\; & \frac{1}{2} (\mathbf{q}_{\tilde{s}} - \bar{\mathbf{q}}_{\tilde{s}})^T \; C_{\tilde{s}}^{-1} \; (\mathbf{q}_{\tilde{s}} - \bar{\mathbf{q}}_{\tilde{s}}) + \\ & \frac{1}{2} (\mathbf{q}_b - \bar{\mathbf{q}}_b)^T \; C_b^{-1} \; (\mathbf{q}_b - \bar{\mathbf{q}}_b) + \\ & \frac{1}{2} (\mathbf{q}_d - \bar{\mathbf{q}}_d)^T \; K_d \; (\mathbf{q}_d - \bar{\mathbf{q}}_d) \end{aligned} \tag{30}$$

Where $\tilde{s}$ denotes the superquadric shape without the bending while, Cs represent the covariance matrices which are accumulated in the iterative training phase. In section 4.2 we show the derivation of stiffness matrix K in the wavelet basis and comment on its structure.

4.2 The Stiffness Matrix in Wavelet Basis

The stiffness matrix for the membrane operator is a very large, sparse and banded matrix in the ordinary nodal basis. By expressing this in an orthonormal wavelet basis, we can achieve near diagonalization of this matrix. We will show how the filter bank interpretation of WT (see figure 2) relates directly to the diagonal entries of the stiffness matrix K. This approach in turn gives us a straight-forward and *inexpensive* way to compute K.

The standard way to obtain stiffness matrix in wavelet basis is essentially a similarity transformation. By combining equations 12, 28 and 29 we get,

$$K^0 = w_0 S^T \left(\int \Phi \Phi^T du \right) S \tag{31}$$

$$K^1 = w_1 S^T \left(\int \left(\frac{\partial \Phi}{\partial u} \frac{\partial \Phi^T}{\partial u} \right) du \right) S \tag{32}$$

$$K = K^0 + K^1 \tag{33}$$

Here K^0 controls the local magnitude and K^1 the local variation of the (membrane) displacement function. Due to orthogonality of the interpolation functions in Φ and orthogonality of S, equation 31 simplifies to $K^0 = w_0 I$. However, the part K^1 of the stiffness matrix in equation 33 needs to be computed by a similarity transformation $K^1 = S^T \hat{K}^1 S$, where $\hat{K}^1$ represents the membrane stiffness matrix in nodal basis (see [13]). The drawback of computing K^1 using this similarity transformation is the large computational cost incurred ($O(N^2)$, where N is the number of nodal variables) which can be avoided by using the property that differential operators in wavelet basis remain sparse (see [1]) and have strong diagonal dominance (see [13]). This allows us to compute the diagonal elements of the stiffness matrix directly by taking inner products of the corresponding finite element basis functions (an $O(N)$ computational cost),

$$K^1_{jk} = w_1 \int_{-\infty}^{\infty} |j\omega|^2 |\Psi_{jk}(\omega)|^2 d\omega \tag{34}$$

Figure 6: (a) Scaling of the diagonal entries in the stiffness matrix K of the membrane model with spectral density, ω^{-2} and an arbitrary density $\omega^{-\beta}$, (b) Increasing energy with increasing resolution

which is done in frequency domain by employing Parseval's theorem [6]. Extension to 2-D case is straightforward and achieved by using the tensor product of the basis functions. The factor $j\omega$ results from the Fourier transform of the first derivatives representing the membrane regularizer. Note that both K^0 and K^1 have diagonal structure and that only the relative magnitude of their diagonal entries determine the smoothing effects of each matrix.

The filter bank interpretation of WT (Figure 2) explains how coefficients corresponding to lower resolutions j contribute less energy than those corresponding to higher resolutions. Figure 6 illustrates two aspects of the modeling, (a) the membrane displacement function which has a spectral density characterized by ω^{-2} and is a *fractal* process of degree two, can be used to modulate the ideal bandpass characteristics of the wavelet basis for creating the appropriate stiffness matrix. In addition, one can synthesize fractals of arbitrary degree by using a polynomial function $\omega^{-\beta}$ to modulate the bandpass characteristic of the wavelet basis for creation of the appropriate stiffness matrix; (b) The concept of increasing energy contribution with increasing resolution via a 1D example. For instance, if we assume an ideal band-pass characteristic for $\Psi_j(\omega)$ between ω_j and $2\omega_j$ we would have the following relationship between elements of K at j and $(j+1)$ resolutions.

$$K_{jk} = w_1 \int_{\omega_j}^{2\omega j} |j\omega|^2 d\omega = 8K_{(j+1)k} \tag{35}$$

Figure 7: Diagonal entries of K as an elevation map, for a $(32X32)$ grid computed via a similarity transformation

Therefore, for the membrane regularizer, we notice that the energy at a resolution $(j+1)$ is eight times that at the next lower resolution j.

For a verification of the above discussed sclaing behaviour, we computed the diagonal enteries of the K matrix for a (32 X 32) grid of nodes using a similarity transformation technique discussed earlier. Figure 7 depicts the diagonal enteries as an elevation map, where the elevation corresponds to the magnitute of the enteries. The enteries are organized as in a standard representation of images in wavelet basis (see figure 3 for the standard organization).

4.3 Prior Model of General Fractal Degree

Splines and stocahstic fractals are complementary techniques for modeling free-form shapes. Splines are well suited for modeling smooth man made objects and are easily controlled. Fractals on the other hand are well suited for modeling irregular shapes and are difficult to constrain. Szeliski et. al. [15] describe a way to constrain fractals by establishing a formal connection between splines and fractals using Fourier analysis of energy minimizing splines.

Our prior model corresponding to the membrane spline is a Markov random field with a Gibbsian distribution. Its spectrum can be determined using Fourier analysis techniques. In

Figure 8: Surface fitting result a) MAP Estimate and b) Representative sample from the a posteriori distribution, corresponding to a fractal surface (membrane model)

their work on constrained fractals Szeliski et. al. [15] show how prior models corresponding to a membrane and thin-plate regularizers posses a spectrum ω^{-2} and ω^{-4} respectively, which is fractal i.e., it is self-affine over scale. These spectra characterize special cases of fractals of general form given by,

$$S(\omega) \propto \frac{1}{\omega^{\beta}}. \tag{36}$$

In [15] fractals of an inbetween degree $2 \leq \beta \leq 4$ were approximated by blending the membrane and thin-plate terms. In contrast, our multi-resolution wavelet based modeling scheme discussed earlier allows us to construct fractals of any degree directly by utilizing the natural scale space characteristics of wavelet basis. Elements of K governing the prior distribution are computed using either equation 34 or 35 with ω^2 replaced by ω^{β}. Thus, yielding a stiffness matrix corresponding to regularizer whose spectral distribution is of order β. A random sample from the corresponding Gibbsian distribution of this prior model will be a fractal surface of degree β.

Figure 8 shows a surface fitting result to 3D point data set with (a) MAP estimate and (b) representative sample from the Gibbs distribution $p(\mathbf{q}|\mathbf{D})$ reflecting the underlying fractal prior. Fractal surfaces are very useful for modeling natural detail. In our application of Medical Imaging, our goal is to realistically model the shape of the anatomical structures

Figure 9: Supervised learning to incorporate information into the prior model

found in the human body, specifically the brain. Experiments toward achieving this goal are currently underway.

5 Supervised Learning and Implementation Results

In our work, prior knowledge about the object shapes that are to be segmented from the MRI data is collected via a supervised learning technique. A training sequence of several iterations provides samples of correctly segmented shapes for specific classes of objects – in our experiments a *gyrus* and the *hippocampus* from the human brain. Parameter sets $\mathbf{q}(i)$ corresponding to these shapes represent the training sequence for the global part $(\mathbf{q}_s + \mathbf{q}_{dg})$ of our prior model (see figure 9). With each iteration i we improve estimates of the mean vector $\bar{\mathbf{q}}$ and covariance matrix C through a simple equation for accumulating measurements as indicated in the Figure 9. Since we update only the statistics of the global parameters of the model, the local deformations remains constrained only by smoothness assumptions captured in the stiffness matrix K (refer to equation 30). Therefore, most of the flexibility of

the locally deformable surface remains unchanged. We recommend human visual supervision of the surface fitting in the training stage to assure that only good segmentation results enter the statistics of the prior model. An extension to multi-class case is straightforward, and can be achieved by defining multiple prior distributions $p_j(\mathbf{q})$ for each class j.

This method clearly demonstrates the useful aspect of global models which offer a compact description of the overall features of the object. Deformable superquadrics with the global shape extended by some offset deformation (as presented in previous sections) allow a much wider range of shapes to be representable in the prior model, all at the expense of few additional parameters. This makes our model very attractive for shape recognition as well as shape reconstruction tasks.

We now present two examples depicting segmentation of a gyrus and a hippocampus from MRI data. The data sets consisted of 64 slices each, with each slice of 1.25mm thickness. We applied a simple 3D edge detection technique to create a gradient force field which was used to attract the inialized model toward the edges of the gyrus and the hippocampus in the first and second examples respectively. Figure 10, illustrates our results displayed with *slices of the reconstructed shape* superimposed on the corresponding slices from the original data. The region of interest has been boxed in the original image and our results are displayed only for this region of interest. Left to right, the images are, a slice from the original data, a slice from the initialized 3D model, the corresponding slice from the final segmentation.

Further experimentation with our modeling scheme is currently under way and we expect to better our results by using more sophisticated 3D edge detectors [5] for defining the external forces to draw the model into the desired shape.

6 Conclusion

In this paper, we introduced a *new shape modeling scheme* that can *transform smoothly from local to global* models or vice-versa. The modeling scheme utilizes the hybrid primitive called the deformable superquadric *constructed in an orthonormal wavelet bases.* These bases provide the power to continuously transform from local to global shape deformations and

Figure 10: Segmentation results for (a) a gyrus and (b) a hippocampus from MRI data; Figure shows slices from 3D, of data, of the initialization, and the reconstruction respectively.

thereby allow for a continuum of shape models – from those with local to those with global shape descriptive power – to be created.

When embedded in a probabilistic framework, our modeling scheme allows us to incorporate more detailed prior information than when using lumped parameter models, while keeping the number of parameters required to describe the shape low. The same level of detail may also be incorporated by using some existing modeling techniques [9, 19] but, at the expense of increasing the number of parameters tremendously. From a computational point of view, our modeling scheme is very *attractive* for segmentation purposes as multi-resolution relaxation is a natural by product of implementing the gradient relaxation method in the wavelet basis. Also, entrapment in local minima during the minimization of the nonlinear energy function can be alleviated. Our future efforts will be directed toward testing with more MRI data and using better 3D edge detectors to define external forces.

References

[1] B. K. Alpert, [1992], "Wavelets and other bases for fast numerical linear algebra," in *Wavelets – A Tutorial in Theory and Applications*, C. K. Chui (ed.), pp. 181-216.

[2] R. Bajscy and R. Solina, [1987], "Three-dimensional object representation revisited," in *IEEE First Conference on Computer Vision*, London, England, pp. 231-240.

[3] A. H. Barr, "Superquadrics and angle-preserving transformations," *IEEE Comp. Graphics and Applications,* Jan. 1981, pp. 11-23.

[4] A. Blake and A. Zisserman [1987], *Visual Reconstruction,* MIT Press, Cambridge, MA.

[5] L. D. Cohen and I. Cohen, [1992], "Deformable models for 3D medical images using finite elements and balloons," *IEEE Conference on Computer Vision and Pattern Recognition*, June, Urbana, Champagne, Ill.

[6] R. N. Bracewell, [1978], *The Fourier Transform and its Applications*, McGraw-Hill Book Company, NY.

[7] I. Daubechies [1988], "Orthonormal bases of compactly supported wavelets," *Communications of Pure and Applied Mathematics*, XLI: pp. 909-996.

[8] H. Delingette, M. Hebert, and K. Ikeuchi, [1991], "Shape represnetation and image segmentation using deformable surfaces," *IEEE Conference on Computer Vision and Pattern Recognition*, Maui, Hawaii, pp. 467-472.

[9] L. H. Staib and J. S. Duncan, [1991], "Boundary finding with parametrically deformable contour models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 11, pp. 1061-1075.

[10] S. Geman, and D. Geman, [1984], "Stochastic relaxation, Gibbs distribution, and Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 6(6), pp. 721-724.

[11] S. G. Mallat, [1989], "A theory of multi-resolution signal decomposition: the wavelet representation," *IEEE Transcations on Pattern Analysis and Machine Intelligence*, Vol. 11(7), pp. 674-693.

[12] A. P. Pentland, [1986], "Parts: Structured descriptions of shape," in *Proceedings of AAAI*, Philadelphia, PA, pp. 695-701.

[13] A. P. Pentland, [1990], "Fast surface estimation using wavelet basis," M.I.T. Media lab Vision and Modeling Group, Technical Report No. 142, M.I.T., Cambridge, MA.

[14] O. Rioul and M. Vetterli, [1991], "Wavelets and signal processing," *IEEE SP magazine*, October, pp. 14–38.

[15] R. Szeliski and D. Terzopoulos, [1989], "Constrained fractals," in *Proceedings of ACM SIGGRAPH*, August, pp. 51-60.

[16] D. Terzopoulos [1986], "Regularization of inverse visual problems involving discontinuities," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **PAMI-8**, 413–424.

[17] D. Terzopoulos, [1986] "Image analysis using multigrid relaxation methods," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **PAMI-8**, No. 2, pp. 129-139.

[18] D. Terzopoulos, A. Witkin and M. Kass, [1988], "Constraints on deformable models: Recovering 3D shape and nonrigid motion," *Artificial Intelligence*, **36**, pp. 91-123.

[19] D. Terzopoulos and D. Metaxas, [1991], "Dynamic 3D models with local and global deformations: Deformable superquadrics," *IEEE Transactions on Pattern Analysis and machine Intelligence*, Vol. 13(7), pp. 703-714.

[20] B. C. Vemuri and R. Malladi [1991], "Deformable models: Canonical parameters for invariant surface representation and multiple-view integration," in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, Maui, Hawaii, June 3-6.

[21] B. C. Vemuri and R. Malladi, "Constructing intrinsic parameters with active models for invariant surface reconstruction," *IEEE Transactions on Pattern Analysis and Machine Intelligence, to appear.*

[22] Y. F. Wang, J. F. Wang, [1990], "Surface reconstruction using deformable models with interior and boundary constraints," in Proceedings of *IEEE Conference on Computer Vision*, Osaka, Japan, pp. 300-303.

[23] A. L. Yuille, D. Cohen, and P. W. Hallinan, [1989], "Feature extraction from faces using deformable templates," IEEE Conference on Computer Vision, San Diego, CA., pp. 104-109.