Chapter 13
Shape as an Emergent Property

Ian H. Jermyn

13.1 Shape Inference

Changes in the properties of matter with position in space, on a scale small enough relative to our own that they can be treated as discontinuities, are ubiquitous in our world. The discontinuities define surfaces, which have geometric properties, which we call 'shape'. The omnipresence of such surfaces, and the distinctive physical properties of the matter that they surround, means that shape is frequently informative about matters of importance to us and to other biological systems, and so inferences involving shape become useful. In particular, because the reflective properties of matter often change along with the properties that define shape surfaces, measurements of light intensity, whether by retina or CCD, can be used to make inferences about shape. Image formation can be approximated in geometric terms, meaning that inferences about two-dimensional shape become relevant too.

What is required to make such inferences? To solve an inference problem, we should construct a probability distribution describing our knowledge of the unknown quantity of interest given the known information. For inferences involving shape, this will involve probability distributions $P(R|K)$, where $R \in \mathscr{R}$ is, in general, an element of a suitable set $\mathscr{R} \subset 2^{\mathscr{D}}$ of subsets of a space $\mathscr{D}$ possessing sufficient geometric structure to render the idea of shape meaningful. In any specific case, R will parameterize a set of propositions whose probabilities we wish to calculate, for example, "region R in the image domain $\mathscr{D}$ contains entity X" (where X is 'human being', 'Ian', 'road network', 'car', etc.). The quantity K denotes all the knowledge we have of the situation (or anyway all the information that we choose, or are able, to express). In particular, this will include all the information we have about the shape of R, perhaps arising from knowledge of X.

In order to make inferences involving shape, then, we need to understand how to construct distributions $P(R|K)$ for given knowledge K: that is, how to encode

I.H. Jermyn (✉)
Department of Mathematical Sciences, Durham University, Durham, UK
e-mail: i.h.jermyn@durham.ac.uk

S.J. Dickinson, Z. Pizlo (eds.), *Shape Perception in Human and Computer Vision*, Advances in Computer Vision and Pattern Recognition,
DOI 10.1007/978-1-4471-5195-1_13, © Springer-Verlag London 2013

Fig. 13.1 An example of an element of the space of 'regions', $\mathscr{R}$, showing the complicated nature of such elements

information about shape into the probability distribution. We will look at the properties probability distributions must have to be regarded as shape models, and then at how these properties have typically been implemented in machine vision. After looking at the drawbacks of this 'classical' approach, we then discuss an alternative, inspired by classes of shapes arising in certain image processing problems. In the resulting framework, shape becomes an emergent property of interactions in a network of simple nodes. It may therefore be of some biological relevance. We conclude with a discussion of this approach, and of what remains to be done to turn it into a complete shape modelling framework.

13.2 Modelling Shape

The key task, then, is to construct probability distributions $P(R|K)$ on a suitable space of shapes $\mathscr{R}$.[1] We will focus attention on a space $\mathscr{R} \subset 2^{\mathscr{D}}$, where $\mathscr{D} \subset \mathbb{R}^2$ is a relevant domain (often the support of an image), although much of the discussion applies to other spaces and other dimensions. 'Regions' $R \in \mathscr{R}$ are open sets, but we do not specify them further; we assume they possess whatever properties are needed to render the models well defined. Note that the space $\mathscr{R}$ is complicated: regions can have arbitrarily many connected components; connected components can contain holes; and these in their turn can contain connected components; and so on. Figure 13.1 shows an element of this space.

Now we need a category of mathematical objects to represent the elements of $\mathscr{R}$, a 'representation space' $\mathscr{S}$. Many such spaces have been used in the literature. Some are isomorphic to $\mathscr{R}$ (indicator functions, distance functions [16]). This seems good, but leaves the complexity of $\mathscr{R}$ intact: $\mathscr{S}$ still contains infinitely many connected components, for example. For others ('many-to-one'), there is a non-injective map from $\mathscr{R}$ to $\mathscr{S}$ (landmark points [6, 12], various Fourier descriptors [20], medial axis [2, 8]). Such representations are often low-dimensional, and can be intrinsically invariant to transformations. However, a region cannot be reconstructed from

[1] A continuum description of regions involves spaces of infinite dimension. The task of constructing probability measures on such spaces, once we move beyond Gaussians, is difficult, and we will not address it. The details are anyway usually irrelevant because they concern infinitely small distances; it is enough to imagine some kind of frequency cut-off imposed at a scale too fine to matter. There may be strong dependence on the scale of the cut-off unless the model parameters are made cut-off dependent, but in practice, since the parameters are determined experimentally for a known cut-off, this is of little importance.

Fig. 13.2 Predicting one part of a boundary from another. *Left*: a sample from the Ising model, where accurate prediction is not possible. *Right*: a case where accurate prediction is possible

its representation, and so there is no probability distribution on shapes; we do not consider them further. Finally, there are representations ('one-to-many') for which there is a non-injective map $\sigma : \mathscr{S} \to \mathscr{R}$ (parameterized closed curves, phase field). The advantage is that $\mathscr{S}$ may have a much simpler structure than $\mathscr{R}$; however, we have to think about the distribution induced on $\mathscr{R}$ by a distribution on $\mathscr{S}$:

$$P(R) = \int_{\mathscr{S}} \delta\big(R, \rho(S)\big) P(S) = \int_{\mathscr{S}_R} P(S), \tag{13.1}$$

where $\mathscr{S}_R = \{S \in \mathscr{S} : R = \sigma(S)\}$. Often one uses a saddle point approximation: $P(R) \propto P(S_R)$, where $S_R = \arg\max_{S \in \mathscr{S}_R} \rho(S)$, with $\rho(\cdot)$ a suitable density for $P(S)$.

Next, we have to construct probability distributions on $\mathscr{S}$ expressing shape information. Do all distributions on $\mathscr{S}$ count? The only precise answer is yes, but this is not very useful. The standard Ising model on $\mathbb{Z}^2$ can be viewed as a probability distribution on regions in $\mathbb{R}^2$ by associating the indicator function of a square with each vertex, but while this distribution undoubtedly contains information about region geometry, because regions with greater boundary length have lower probability, it cannot really be called a 'shape model'. This can be seen by looking at distant parts of the sample in Fig. 13.2: high probability regions do not have any properties in common that we would normally call 'shape'.

Why do we say that the Ising model does not contain shape information? The reason is that the set of high probability regions is too large, that is, the entropy is too high. In order to reduce the entropy, we have to create more dependence in the distribution. Indeed, it is clear that what we normally refer to as 'shape', involves the ability to make quite precise inferences about the overall region give only partial information about its boundary. For example, most people could make a good estimate of the part of the object boundary that is concealed in the image on the right-hand side of Fig. 13.2, given the part that is revealed: the conditional entropy is small. The same is not true of the left-hand side. In other words, the probability distribution induced by our knowledge of the object's identity and behavior contains strong, long-range dependencies between parts of the region boundary. Such dependencies, then, are key to the construction of non-trivial shape models. We now turn to how to build probability distributions that incorporate such dependencies.

13.3 The Classical Approach

We will first look at what we will call the 'classical approach' to shape modelling. The focus here is on defining a 'dissimilarity measure' $d : \mathscr{S} \times \mathscr{S} \to \mathbb{R}_{\geq 0}$ between points in $\mathscr{S}$. Usually, although not always, this measure is metric [3, 4, 10, 14, 15, 18, 19]. Once defined, the metric can be used to define a probability distributions on $\mathscr{S}$ as a function of distance to a 'template' shape S_0. The most common form of distribution is 'pseudo-Gaussian', taking the form:

$$P(S|S_0) \propto dS\, e^{-\frac{1}{2}d^2(S,S_0)}, \tag{13.2}$$

where S_0 is the 'template', and dS is an underlying measure on $\mathscr{S}$. (Mixtures of such distributions over a set of templates have also been used [5].)

The distribution (13.2) encourages S to be close to the region S_0, but this is rarely what is required. Typically, there will be uncertainty about the position, orientation, and perhaps scale of the shape. To incorporate such uncertainty in the classical approach, one must create mixture models over these transformations. Let G be the transformation group acting on $\mathscr{S}$, with the action denoted gS. Then the distribution one is really interested in is

$$P(S|S_0) = \int_G P(S|g, S_0)P(g|S_0) \propto dS \int_G dg\, \rho(g) e^{-\frac{1}{2}d^2(S,gS_0)}, \tag{13.3}$$

with $P(g|S_0) = dg\, \rho(g)$, where dg is an invariant measure on G. Often complete invariance to G is needed. This requires dS to be G-invariant, $\rho \equiv 1$, and G to act by isometries on $\mathscr{S}$: $d^2(gS, S_0) = d^2(S, g^{-1}S_0)$. In practice, the integral in Eq. (13.3) is rarely evaluated. Rather a saddle point approximation is made in which $g^* = \arg\min_{g\in G} d^2(S, gS_0)$ is substituted, giving $P(S|S_0) \propto dS\, e^{-\frac{1}{2}d^2(S,g^*S_0)}$, that is, pose is estimated. Although easier than performing the integral, this still requires significant computational effort.

How do the long-range dependencies necessary for nontrivial shape modelling arise in the classical approach based on templates? The answer is that the template itself, or rather its parameters, such as the group elements just discussed, act as hidden variables. Once they are integrated out, they introduce long-range dependencies between boundary points. A trivial example, in one dimension, is the following. In 1-d, regions are unions of intervals; we consider only connected regions. Let the template region be an interval of length 1, with center at c. Let the probability of a region $[x, y]$ be $P(x, y|c) = \delta(x - (c - \frac{1}{2}))\delta(y - (c + \frac{1}{2}))$. Thus, given c, x and y are independent. If we now add a uniform prior on c (suppose $c, x, y \in S^1$ so this is normalized), we can integrate out c to obtain $P(x, y) = \delta(y - x - 1)$. Thus not knowing c, x and y are dependent: in this singular case, x determines y completely. In the general case, integration over a group as above, or integrations over other unknown template parameters, play exactly the same role as in this simple example, introducing the long-range dependencies that contain nontrivial shape information.

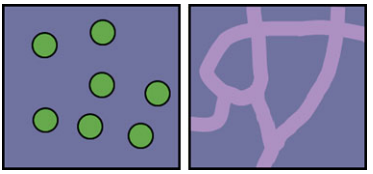

Fig. 13.3 *Left*: multiple instances of an entity ('gas of near-circles'). *Right*: a 'network' region

13.3.1 Drawbacks of the Classical Approach

While the classical approach to creating long-range dependencies using templates and a metric is useful and efficient in many applications, it does not apply, or is inefficient, in many important cases. In particular, the use of templates and metrics means that high probability shapes only occur 'close' to one or more points in the space of shapes. There are entities, however, for which our knowledge of their shape cannot be expressed in terms of small variations around a template shape or shapes. In particular, when the entity involved has an extent or a topology that is in some way unconstrained, the use of templates fails to allow sufficient variability.

Perhaps the most commonly occurring example is when multiple instances of an entity can be present: see Fig. 13.3 left. In this case, although each entity might be well described by a template and small variations, the whole may have any number of connected components, and hence is not amenable to a template/metric description. Although in principle this situation can be dealt with by using object point processes [13], in practice the large number of degrees of freedom per object, together with the necessity to estimate transformation group parameters for each instance, mean that such methods are very inefficient.

Another example is provided by 'network' regions: see Fig. 13.3 right. The set of network regions can be divided into topologically distinct subsets classified by the graph of which they are a fattened version. It is thus clear that such shapes cannot be described as variations around a finite number of templates.

To overcome these drawbacks, a new modelling framework is needed that allows the incorporation of strong constraints on region shape, without necessarily constraining region topology, and that provides intrinsic invariance. To achieve this, the long-range dependencies necessary for shape modelling will be encoded in the distribution in a new way. This turns out to be of interest in its own right, independently of the examples that inspired it.

13.4 Nonlocal Interactions

In this section, we will look at an alternative method for introducing the long-range dependencies needed in order to encode non-trivial shape information. Rather than using a template to introduce such dependencies, explicit nonlocal interactions between region boundary points will be introduced. These interactions generate long-range dependencies strong enough to constrain region shape, but because no template is used, they need not constrain region extent or topology. In addition, the

models will be intrinsically invariant to Euclidean transformations, meaning that these transformations do not have to be estimated for each instance of an entity. Multiple instances thus become easier to handle. We first describe these models in terms of the 'contour representation', but then go on to reformulate them in terms of nonlocal interactions in networks of simple real- and binary-valued nodes.

The contour representation represents $R \in \mathscr{R}$ by its boundary ∂R, which consists of a set of oriented, closed curves. (The dark lines bounding the region in Fig. 13.1 show an example of a region boundary; the boundary orientation is not shown in the figure.) The contour representation space $\mathscr{S} = \Gamma$, is thus the space of multiple, oriented, closed curves, subject to certain constraints, to which we return later. In fact, it is often convenient to write probability distributions in terms of circle embeddings $S^1 \to \mathbb{R}^2$: making a distribution invariant to the action of $\mathrm{Diff}(S^1)$ then ensures that it is well-defined on Γ, and hence on $\mathscr{R}$.

We now introduce a class of models, expressed in the contour representation and known as 'higher-order active contours' [17], that encode nontrivial shape information via explicit nonlocal interactions.[2]

13.4.1 Higher-Order Active Contours

The simplest Euclidean invariant model one can place on Γ is

$$E_{C,0}(\partial R) = \lambda_C L(\partial R) + \alpha_C A(\partial R), \tag{13.4}$$

where L and A are region boundary length and region area respectively; and $\lambda_C, \alpha_C \in \mathbb{R}_{\geq 0}$. This model, or minor variants of it, has been much used as a region model in the literature, starting with [11]. Indeed, it is essentially the Ising model in a constant external field, expressed in the contour representation. As such, although this model contains important information about the (low-resolution) smoothness of region boundaries, it contains no real shape information. Indeed, both L and A can be expressed as single integrals over ∂R involving only tangent vectors, meaning that only 'infinitesimally nearest neighbor' points on the boundary interact: the model does not contain the long-range dependencies necessary to incorporate nontrivial shape information.

To incorporate nonlocal interactions, and hence long-range dependencies, one must move from single integrals to multiple integrals, thereby incorporating information from more than one contour point at a time. Two integrals is the simplest case: pairs of points on the boundary then interact. The idea is illustrated in Fig. 13.4. One possibility among many for such a 'higher-order active contour' energy is

$$E_{C,\mathrm{NL}}(\partial R) = -\beta_C \iint_{S^1 \times S^1} dt\, dt'\, n \cdot \mathbf{G}\big(\gamma, \gamma'\big) \cdot n', \tag{13.5}$$

[2]We will not talk about probability distributions directly from now on, but rather about their energies E, defined, up to an additive constant, by $P(R|K) \propto \exp(-E(R|K))$.

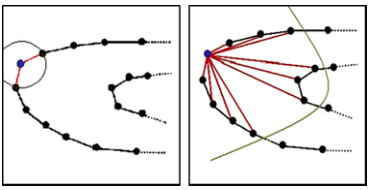

Fig. 13.4 *Left*: the 'nearest-neighbor' interactions induced by first derivatives. *Right*: nonlocal interactions

where γ is an embedding of a circle (or multiple circles, if the region has more than one connected component or is multiply-connected), whose image is ∂R; t, t' are coordinates on S^1; n indicates the un-normalized normal vector field to γ (i.e., $\dot{\gamma}$ rotated by $\pi/2$); (un)primed quantities are evaluated at $(t)t'$; and $\mathbf{G} : \mathbb{R}^2 \times \mathbb{R}^2 \to T^*\mathbb{R}^2 \boxtimes T^*\mathbb{R}^2$, where $\boxtimes$ indicates the outer tensor product, is a bitensor field. Note that it is easy to make $E_{C,\mathrm{NL}}$ intrinsically Euclidean invariant, for example by taking $\mathbf{G}(x, x') = \Psi(|x - x'|)\mathbb{I}$.

Summing the nonlocal term (13.5) with Eq. (13.4) gives an energy $E_C = E_{C,0} + E_{C,\mathrm{NL}}$ with interesting properties, documented in [9, 17]. In particular, for certain parameter ranges, calculable via stability analyses [7, 9], E_C has local minima corresponding to 'network' regions or to 'gas of near-circles' regions. Examples of such local minima of E_C, generated by gradient descent, are shown in Fig. 13.5. Network regions consist of a number of branches joining together at junctions, and can be thought of an 'fattened embedded graphs', as in Fig. 13.3. 'Gas of near-circles' regions consist of any number of connected components, each of which has infinitely many degrees of freedom, but which with high probability is 'close' to being a circle of a given radius. Note that 'gas of near-circles' regions represent multiple instances of a shape, with different interaction functions Ψ favoring different perturbations of the circle, and hence different shapes. Intrinsic Euclidean invariance of E_C means that no pose estimation is required.

Higher-order active contours demonstrate that non-trivial shape information can be encoded using explicit nonlocal interactions between boundary points, and they have been used successfully in a number of image processing applications [9, 17]. Nevertheless, the contour representation in which they are expressed suffers from a number of drawbacks arising from the fact that not all sets of oriented, closed curves are boundaries: constraints are needed to prevent (self-)intersections; curve orientations, which describe 'inside' and 'outside', have to be mutually consistent; and the space Γ is not connected, having one component for each topologically distinct set of regions (connected region components, holes, nested regions, ...). These difficulties can all be alleviated by changing the shape representation. In the

Fig. 13.5 Local minima under E_C. *Left*: network regions; *right*: gas of near-circles regions

process, we will see that shape information can be encoded in terms of nonlocal interactions in a network of simple binary- or real-valued nodes.

13.4.2 Reformulation as a Network of Nodes

A region can be represented by a binary-valued function $\phi : \mathscr{D} \to \{\pm 1\}$, with $R = \zeta(\phi) = \{x \in \mathscr{D} : \phi(x) > 0\}$. It turns out to be convenient to relax this to a real-valued 'phase field' function $\phi : \mathscr{D} \to \mathbb{R}$ (we will use the same symbol) controlled by an energy that encourages it to take on the values ± 1. Perhaps the simplest such energy is the Ginzburg-Landau energy:

$$E_0 = \int_{\mathscr{D}} dx \left\{ \frac{D}{2} |\partial \phi|^2 + \lambda \left(\frac{\phi(x)^4}{4} - \frac{\phi(x)^2}{2} \right) + \alpha \left(\phi(x) - \frac{\phi(x)^3}{3} \right) \right\}, \tag{13.6}$$

where all ϕ are evaluated at x. The ultralocal part of the integrand has minima at ± 1. This means that if $D = 0$, $\phi_R = \arg\min_{\phi : \zeta(\phi) = R} E_0(\phi)$ will take the value $+1$ inside, and -1 outside R, that is, will be binary. It is then easy to see that $E(\phi_R) = \frac{2}{3}\alpha A(R)$ up to an additive constant. Nonzero D has the effect of smoothing the discontinuity, and also measures the boundary length. Indeed it can be shown that

$$E_0(\phi_R) \simeq \lambda_C L(R) + \alpha_C A(R), \tag{13.7}$$

where λ_C and α_C are functions of λ, α, and D. The energy E_0 is thus $E_{C,0}$ reformulated in terms of the phase field representation.

It is now natural to ask whether it is possible to create a phase field energy that is equivalent to $E_C = E_{C,0} + E_{C,\mathrm{NL}}$. This is indeed the case. The equivalent of the nonlocal energy $E_{C,\mathrm{NL}}$ in Eq. (13.5) is

$$E_{\mathrm{NL}}(\phi) = -\frac{\beta}{2} \iint_{\mathscr{D}^2} dx\, dx'\, \partial\phi(x) \cdot \mathbf{G}\big(x, x'\big) \cdot \partial\phi\big(x'\big). \tag{13.8}$$

It can then be shown that

$$E(\phi_R) = E_0(\phi_R) + E_{\mathrm{NL}}(\phi_R) \simeq E_{C,0}(R) + E_{C,\mathrm{NL}}(R) = E_C(R). \tag{13.9}$$

The phase field representation $\mathscr{S} = \Phi$ is a one-to-many representation. Equation (13.9) shows that in a saddle-point approximation to Eq. (13.1), the phase field model E is equivalent to the contour model E_C. (The same is true in the Gaussian approximation also.) In particular, for different parameter ranges, the phase field model has local energy minima corresponding to networks and a gas of near-circles; the ranges can be found by translating the results of the stability analyses performed in the contour representation to the phase field representation, and are well verified numerically. This is useful because the phase field representation turns out to have multiple advantages. First, unlike the contour representation, there are no difficult constraints to implement: ϕ lives in a linear space Φ. Second, regions of arbitrary

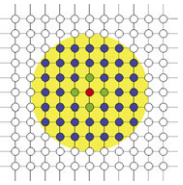

Fig. 13.6 Diagram of the network interactions in the spatially-discretized phase field model

topological complexity are all represented in a single, connected space, so that no special methods are needed to deal with multiple connected components, handles, etc. Coupled with the intrinsic Euclidean invariance of the energy, this means that multiple instances of an entity are modeled at essentially no extra cost. Third, in the contour representation, γ appears as an argument to $\mathbf{G}$, which makes the nonlocal term complicated. In the phase field representation, the nonlocal term is quadratic, and when translation invariant, is diagonal in the Fourier basis. This greatly simplifies any computations involving it.

The continuum form of the phase field energy is easy to manipulate, but computations, whether in a machine or biological visual system, will inevitably involve discretization of some sort. If we spatially discretize Eq. (13.6), the result is a Markov random field ψ, consisting of a set of real-valued nodes, interacting (see Fig. 13.6): with themselves via the potential (red); with their nearest neighbors via the derivative term (green); and also with the large number of nodes that lie within the support of the nonlocal interaction (blue). We thus see that nontrivial shape information can be encoded as nonlocal interactions in a network of real-valued nodes.

We can simplify things even further, at the cost of losing some geometric accuracy, as follows. On most of its domain, the phase field takes values very close to the set $\{-1, 1\}$. This suggests replacing ψ by a field taking values *only* in the set $\{-1, 1\}$, that is, by a binary-valued Markov random field ω. By definition, the distribution for ω is given in terms of that for ψ by $P(\omega) = \int_\Psi P(\omega|\psi)P(\psi)$. Binarization means that $P(\omega|\psi) = \delta(\omega, \mathrm{sgn}(\psi))$. In the saddle point approximation, the energy U of the binarized field is given by $U(\omega) = E(\psi_\omega)$, where $\psi_\omega = \arg\min_{\psi:\,\mathrm{sgn}\,\psi=\omega} E(\psi)$. Computing ψ_ω is a difficult task in itself. A crude but practically effective approximation gives rise to the energy [1]:

$$U(\omega) = \frac{D_b}{2} \sum_{i,j:i\sim j} (\omega_i - \omega_j)^2 + \alpha_b \sum_i \omega_i + \frac{\beta_b}{2} \sum_{i,j} \omega_i F_{ij} \omega_j, \tag{13.10}$$

where $\alpha_b = \frac{2\alpha}{3}$, $\beta_b = \beta$, $D_b = \frac{D}{4}$, and F is related to $\partial^2\mathbf{G}$. Gibbs sampling from this distribution, with appropriate temperature and parameter ranges (again derived from stability analyses performed in the contour representation) shows convergence to gas of near-circles or network regions. These then fluctuate, but remain stable, under further sampling. As with the other two representations, the probability distribution can have local maxima at these shape families. Thus, the same nontrivial shape information can be encoded as nonlocal interactions in a network of binary nodes.

13.4.3 Nonlocality via Local Interactions

Nonlocal interactions have so far been introduced explicitly. However, explicit nonlocality may not be plausible in some contexts, for example the biological, so it is important to note that nonlocality can arise from purely local energies. We introduce a vector field $v : \mathbb{R}^2 \to \mathbb{R}^2$, and define a joint distribution with Gibbs energy

$$\hat{E}(\phi, v) = E_0(\phi) + a\langle v|\partial\phi\rangle + \frac{1}{2}\langle v|F|v\rangle, \tag{13.11}$$

where $\langle|\rangle$ is the L^2 inner product on $\mathscr{D}$; F is a positive operator; and $a \in \mathbb{R}$. Notice that v couples to the gradient of ϕ, that is, to the boundary. At the same time v is spatially correlated (e.g., smoothed) by the interaction represented by F. It therefore induces an interaction between points on the boundary. To find this interaction, we marginalize over v. The resulting Gibbs energy is

$$\tilde{E}(\phi) = E_0(\phi) - \frac{1}{2}\beta\langle\partial\phi|F^{-1}|\partial\phi\rangle, \tag{13.12}$$

where $\beta = a^2$. This has the same form as the nonlocal phase field energy E defined in Eqs. (13.7), (13.8), and (13.9). A similar procedure works for the binary MRF model. Thus rather than encoding shape information via nonlocal interactions, it can instead be encoded by allowing the network to have several 'layers'.

13.5 Discussion

Thus, we reach the end of the story. Shape information can be encoded via nonlocal interactions in a network of binary- or real-valued nodes. In turn, these nonlocal interactions can be re-encoded as local interactions in a network with multiple 'layers'. Control of these interactions then allows different shape families to be modeled. Shape, therefore, does not have to be described by exogenous templates, or constructed from arbitrary building blocks. Instead, it can arise naturally, as an emergent property of the connections in a network of simple nodes.

The fact that shape information is encoded as interactions in a network means that shape processing can be inherently parallel. The domain $\mathscr{D}$ can be separated into subdomains that can be processed simultaneously, with some communication overheads. Note that to search for a shape in multiple subdomains using a template would be equivalent to searching for multiple instances, requiring separate pose estimations for each subdomain.

For image processing applications, the nodes are usually generated by spatial discretization onto a square lattice, but any discretization is possible, for example, using a hexagonal lattice. The nodes need not even correspond to spatial elements: one of the strengths of the phase field representation is that it can be written in any basis, and so a discretization could be generated by imposing a frequency cut-off

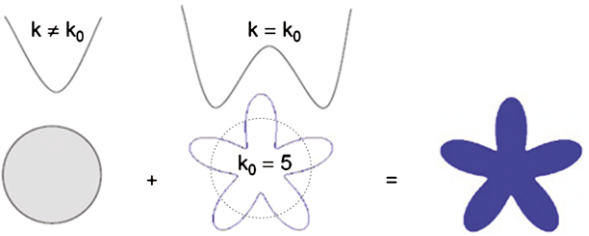

Fig. 13.7 A Fourier component, quadratically unstable at zero but stabilized at a finite value by quartic behavior, added to an otherwise stable circle

in Fourier space, or a scale cut-off in wavelet space. The local and non-local terms (but not the potential) are diagonal in the Fourier basis, while the potential term is diagonal in the spatial point basis. This suggests that a wavelet basis, which is intermediate between these two extremes, might simplify the interactions in the model. This would also have the advantage of providing a naturally multiscale representation of shape.

Despite all its promising aspects, however, the framework cannot yet be called a complete shape modelling method. Only simple shapes have been modeled so far, and the natural question is whether one can model more complex shapes.

One possible direction is suggested by observations in numerical experiments. Shapes have been seen that were neither circles nor bars, but instead were star-shaped: circles plus a sinusoidal perturbation of their radius. These appeared to be stable. While it is possible that this was an artefact of the numerical method, it is also possible that the chosen parameter values produced a new type of local minimum. A simple explanation is as follows. To second order in a small perturbation of a circle, only two behaviors are possible for each Fourier component of the perturbation: stable or unstable, corresponding to positive or negative second-order coefficient in the expansion. To fourth order, however, more complex behaviors can occur. In particular, if the second-order coefficient is negative but the fourth-order coefficient is positive, then although the zero amplitude state is unstable, there will be some finite amplitude that *is* stable. Now imagine that all Fourier components are stable quadratically except for one, which is unstable quadratically but stabilized by a fourth-order term. The circle itself is now a saddle-point of the energy, while a circle perturbed by a sinusoid of the correct frequency and amplitude is a local minimum. This is illustrated in Fig. 13.7.

This picture suggests that by adjusting the interaction function of the model, one might be able to assign different stable amplitudes to each Fourier component. Were this possible, it would be mean that any star domain could be modeled.

An alternative approach to the modelling of more complex shapes involves the introduction of higher-order interactions. There are two issues with such interactions. The first is learning the interactions necessary to model a given family of shapes. In order for an energy to model such a family, it should have local minima at the appropriate points in $\mathscr{R}$, and this involves difficult analysis. It could perhaps be achieved using standard statistical estimation techniques verified a posteriori for local minimality, or by placing constraints on the parameters during estimation. The first seems wasteful, while the second is complex, and its theoretical basis is not

clear. The second issue is algorithmic complexity. Simply evaluating higher-order terms is expensive, and although there are algorithms available for certain types of higher-order term in the binary MRF case, it is not likely that they would apply to the types of term needed. Nevertheless, some promising progress is being made in these areas, and there is good reason to hope that the picture of shape as an emergent property of interactions between network nodes can be fully realized.

Acknowledgements Most of the work described herein was carried out while the author was working in the Ariana group at INRIA. Many people contributed, and continue to contribute to this work. The author would like to acknowledge and thank: J. Zerubia, Z. Kato, M. Rochery, P. Horváth, A. El Ghoul, T. Blaskovics, J. Nemeth, C. Molnar, and I. Bechar.

References

1. Blaskovics T, Kato Z, Jermyn IH (2009) A Markov random field model for extracting near-circular shapes. In: Proc IEEE international conference on image processing (ICIP), Cairo, Egypt, November 2009
2. Blum H (1973) Biological shape and visual science. J Theor Biol 38:20–287
3. Bronstein A, Bronstein M, Kimmel R (2006) Efficient computation of isometry-invariant distances between surfaces. SIAM J Sci Comput 28(5):1812–1836
4. Cremers D, Soatto S (2003) A pseudo-distance for shape priors in level set segmentation. In: Paragios N (ed) Proc IEEE workshop variational, geometric and level set methods in computer vision, Nice, France, October 2003, pp 169–176
5. Cremers D, Osher SJ, Soatto S (2006) Kernel density estimation and intrinsic alignment for shape priors in level set segmentation. Int J Comput Vis 69(3):335–351
6. Dryden IL, Mardia KV (1998) Statistical shape analysis. Wiley series in probability and statistics: probability and statistics. Wiley, New York. ISBN 9780471958161
7. El Ghoul A, Jermyn IH, Zerubia J (2008) Phase diagram of a long bar under a higher-order active contour energy: application to hydrographic network extraction from VHR satellite images. In: Proc international conference on pattern recognition (ICPR), Tampa, Florida, December 2008
8. Geiger D, Liu T-L, Kohn RV (2003) Representation and self-similarity of shapes. IEEE Trans Pattern Anal Mach Intell 25(1):86–99
9. Horvath P, Jermyn IH, Kato Z, Zerubia J (2009) A higher-order active contour model of a 'gas of circles' and its application to tree crown extraction. Pattern Recognit 42(5):699–709
10. Joshi SH, Klassen E, Liu W, Jermyn IH, Srivastava A (2011) Shape analysis of elastic curves in Euclidean spaces. IEEE Trans Pattern Anal Mach Intell 33(7):1415–1428
11. Kass M, Witkin A, Terzopoulos D (1988) Snakes: active contour models. Int J Comput Vis 1(4):321–331
12. Kendall DG (1984) Shape manifolds, procrustean metrics, and complex projective spaces. Bull Lond Math Soc 16:81–121
13. Kulikova MS, Jermyn IH, Descombes X, Zhizhina E, Zerubia J (2011) A marked point process model including strong prior shape information applied to multiple object extraction from images. Int J Comput Vis Image Process 1(2):1–12
14. Leventon ME, Grimson WEL, Faugeras O (2000) Statistical shape influence in geodesic active contours. In: Proc IEEE computer vision and pattern recognition (CVPR), vol 1, Hilton Head Island, South Carolina, USA, June 2000, pp 316–322
15. Michor PW, Mumford D, Shah J, Younes L (2008) A metric on shape space with explicit geodesics. Rend Lincei, Mat Appl 9:25–57
16. Osher S, Sethian JA (1988) Fronts propagating with curvature dependent speed: algorithms based on Hamilton-Jacobi formulations. J Comput Phys 79(1):12–49

17. Rochery M, Jermyn IH, Zerubia J (2006) Higher-order active contours. Int J Comput Vis 69(1):27–42
18. Srivastava A, Joshi S, Mio W, Liu X (2005) Statistical shape analysis: clustering, learning, and testing. IEEE Trans Pattern Anal Mach Intell 27(4):590–602
19. van Kaick O, Zhang H, Hamarneh G, Cohen-Or D (2011) A survey on shape correspondence. Comput Graph Forum 30(6):1681–1707. ISSN 1467-8659
20. Zahn CT, Roskies RZ (1972) Fourier descriptors for plane close curves. IEEE Trans Comput, C 21(3):269–281