Chapter 20
Fluctuating Distance Fields, Parts, Three-Partite Skeletons

Sibel Tari

> *So the nature of a visual experience can not be described in terms of inches of size and distance.*
> Rudolf Arnheim

Abstract Shapes are continuous objects, even when they are drawn on digital media or composed using finite number of elements. As such, they defy analytic approach; explicitization of their parts, hierarchies, skeletons, or even centroids is ill-posed. I describe a novel approach to perceptually organize shapes and explicate their features without being negligent of their continuous nature. The basic construct is an unusual phase field that can be conceived in a number of varying ways using varying mathematical machinery, so highlighting the field itself rather than how it is being computed. Connections among the field, Mumford-Shah and Tari-Shah-Pien models, and reaction-diffusion equation suggest that the field may bridge low-level and high-level visual processing.

20.1 Shapes Are Continuous

Shapes are continuous objects, even when they are drawn on digital media or composed using finite number of elements. In favor of the argument, Gestalt school produced numerous illustrations, e.g., (1) a finite set of isolated dots placed at a fixed distance from a fixed location creates a circle percept; (2) a peanut shape prevailingly reveals two discrete blobs or approaches to a blob, as its neck thins or thickens; (3) two squares juxtaposed to form an eight-pointed star gives way to new parts, namely eight triangles and an octagon.

S. Tari (✉)
Middle East Technical University, 06531, Ankara, Turkey
e-mail: stari@metu.edu.tr

In a similar vein but surpassing Gestalt, Stiny [50] speaks for ambiguity (i.e. multiple interpretations) in part perception, convincing the reader to place part boundaries wherever and whenever desired. Such a liberal view of shape topology certainly stands in strong contrast to prevalent theories on part perception exemplified by seminal works of Biederman [9] (imposing regularity on parts) and Hoffman and Richards [23] (imposing regularity on break points); nevertheless, contemporary work in neuroscience support the continuity of the process, too [37].

Additionally, shapes live in continua as opposed to standing in isolation. The basis of this argument is two-fold: (1) perception of a shape is influenced by nearby shapes (underlying category structures); indeed, several papers, e.g. [17,34], draw parallels between this contextual effect and experimentally observed non-metric characteristics (asymmetry and violation of triangle inequality) of pairwise dissimilarities; (2) shapes naturally deform into other shapes, say, a peanut changes to a blob-like shape or a pair of disjoint blobs as its neck thickens or thins, respectively.

To a great extent, it is the continuous nature of shapes that hinders their comprehension and description using analytic methods; thus, it is only natural that several chapters from this book present state of the art techniques for situating shapes in continua by constructing shape dynamics, shape spaces etc.

In pursuit of the goal, the presented approach is an attempt to perceptually organize shapes and explicate their features in the form of skeletons without being negligent of their continuous nature. At the core of constructions is an unusual distance-like, scalar field, ω, that implicitly codes interactions among shape elements, both over the region and along the boundary and both local and non-local.

Overview of the New Distance Field The most distinct feature of the new field is that it fluctuates, i.e., unlike previous ones [4, 22, 28, 42, 54], the new one is a non-monotone function of the distance to a set of input points. It is non-monotone because the value of the field at each point mimics the distance (a *diffuse* one [54, 55]) from the point to the union of (1) a set of input points and (2) *a yet to be determined curve.*

The set of input points specifies the shape boundary either incompletely (in cases of open contours, illusory contours or outcomes of edge detectors) or completely as the bounding curve of a connected bounded open set, say $\Omega \subset \mathbf{R}^2$. In either case, the shape boundary is indicated by the set of input points, denoted by $\partial \Omega$ and subjected to homogeneous Dirichlet type boundary condition.

The construction of the ω is valid regardless of the form the boundary is specified. If it is completely specified with a clear distinction between interior and exterior, the field can be computed only on the interior. If, however, the boundary is incompletely specified, the field needs to be computed on a wider scope R which is a simply connected subset of $\mathbf{R}^2$ containing the input set. Of course, a suitable boundary condition is needed on ∂R.

The new field has interesting emergent features; for example, it attains value *zero* not only on the shape boundary (which is due to homogeneous Dirichlet boundary condition on $\partial \Omega$) but also on a curve bounding a central cut. The latter locus is an emergent feature and divides the shape into central and peripheral regions.
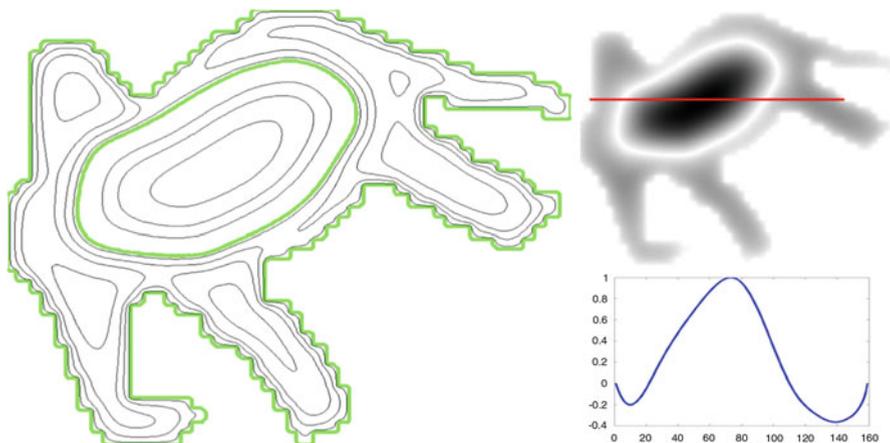
Fig. 20.1 The values of the new field fluctuate between positive and negative. The iso-level *curves* reflect the minimum distance to the union of the shape boundary and an emergent *curve* (shown in *green*)

Roughly speaking, the central region is the least deformable and coarse structure of the shape; moreover, the field inside it is a flexible indicator of the topology and centrality, not necessarily bound to a unique interpretation.

A sample field is depicted in Fig. 20.1. The leftmost figure displays some iso-level curves of the field and the top right one displays its absolute value. Finally, the bottom right figure is the plot of the values at the cross-section marked by the red line in the top right. Observe that the field exhibits interesting features in addition to the emergent zero-level curve: lower level sets in the peripheral region code appendages, the level curves neighboring the boundary resemble erosions of the shape, and the level curves neighboring the zero-level curve resemble erosions and dilations of the central region. All create a dynamic view of the shape.

What's more, the geometry of the level curves is intricately related to that of a well studied diffuse distance field [3, 4, 47, 54, 55] which is the solution of the screened Poisson PDE and has close connection to the Ambrosio-Tortorelli approx-imation [1] of the Mumford-Shah functional [35] (more on this in Sect. 20.2.3).

The intricate relation between the two fields facilitates an unorthodox skeleton abstraction that I call Three-Partite, due to its being composed of three distinct groups providing a separation into (1) gross structure; (2) limbs and (3) boundary texture or noise. Two examples are shown in Fig. 20.2 where the three skeleton groups are respectively shown in red, blue and green. The first group (red) consists of nested closed curves and abstracts the topology as well as the centrality of the coarse structure. It has two distinct layers: the inner (dark red) and the outer (light red). Both layers are coarser forms of the boundary. The outer layer, which is finer than the inner one, can be considered as the coarsest and least articulable structure still indicating the appendages. The inner layer is the emergent zero-level curve, bounding the central region. The second group (blue) consists of disjoint branches

Fig. 20.2 The three skeleton groups are respectively shown in *red*, *blue* and *green* colors
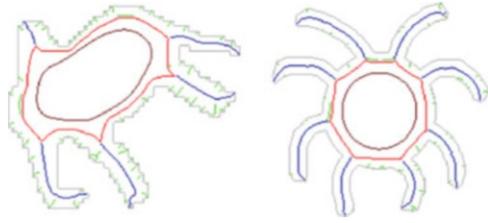
(or loops in case of holes in the periphery) that terminate upon reaching the outer layer of the first group. Each branch of the second group captures a ribbon-like section, a limb or protrusion. The third group (green) consists of disjoint branches, too; but these branches correspond to secondary details.

Incentives for the Three-Partite Skeleton are given in Sect. 20.3 after the new field is explored in Sect. 20.2.

20.2 Fluctuating Distance Field ω

The new field is formulated firstly in a discrete setting, and then re-defined in a continuous setting to better understand the geometry of its level curves by relating it to a previously studied one. In both settings, the field is denoted by ω.

In the discrete setting, ω is the minimizer of an energy which reflects both local and global interactions in the shape interior and along the shape boundary.

On one hand, the boundary and the interior are interchangeable indicators for a shape but each explicates a different set of features. Though it is customary to classify methods as either boundary or region based [21], complementarities between the two, as an experimental guess since Gestalt, are exploited in established mathematical models [33, 35]. Likewise, in the proposed computational model, boundary and interior terms are kept distinct.

As the complementarity between the region and the boundary, the complementarity between the local and the global is an experimental guess: On the computer vision side, several working techniques, e.g. [8,13,32,57,58], utilize local and global features simultaneously; on the human vision side, theories that differ in details agree on that the shape perception is guided by both local and global factors [7, 14–16, 36, 49].

I remark that the choice of terms in the proposed model are driven by experimental guess as well as a bias towards simple linear formulations, since precise physics of the problem is yet unclear.

20.2.1 Formulation

In the discrete setting, the field $\omega : \Omega \to \mathbf{R}$ is a real valued function on a discrete set of sites, $\Omega \subset \mathbf{Z} \times \mathbf{Z}$, equipped with a neighborhood system. The energy $E(\omega)$

is taken as a linear combination of a region based energy $E_{Reg}(\omega)$ defined on Ω and a boundary based energy $E_{Bdy}(\omega)$ defined on $\partial\Omega$, both of which is a sum of pixel-wise terms:

$$E(\omega) = \sum_{(i,j)\in\Omega} E_{Reg}(\omega_{i,j}) + \lambda \sum_{(i,j)\in\partial\Omega} E_{Bdy}(\omega_{i,j}) \qquad (20.1)$$

where $\lambda > 0$ is a constant. The additive form is for computational simplicity.

Region Energy The region energy composed of both local and global terms is expressed as a linear combination, too:

$$E_{Reg} = \underbrace{\sum_{(i,j)\in\Omega} E_{Reg}^{G}(\omega_{i,j})}_{E_{Reg}^{G}\,:\,\text{global}} + \beta \underbrace{\sum_{(i,j)\in\Omega} E_{Reg}^{L}(\omega_{i,j})}_{E_{Reg}^{L}\,:\,\text{local}} \qquad (20.2)$$

where $\beta > 0$ is a constant. In the absence of a specific bias towards local or global, a natural choice may be $\beta = 1$.

Global Component Consider the squared average, the simplest quadratic expression linking all the nodes, as the global energy:

$$E_{Reg}^{G}(\omega_{i,j}) := \left(\frac{1}{|\Omega|} \sum_{(k,l)\in\Omega} \omega_{k,l} \right)^{2}$$

Differentiating its sum over Ω w.r.t. $\omega_{i,j}$:

$$\frac{\partial}{\partial\omega_{i,j}} E_{Reg}^{G} = \frac{\partial}{\partial\omega_{i,j}} \sum_{(i,j)\in\Omega} \left(\frac{1}{|\Omega|} \sum_{(k,l)\in\Omega} \omega_{k,l} \right)^{2}$$

$$= \frac{2}{|\Omega|} \sum_{(k,l)\in\Omega} \omega_{k,l} \qquad (20.3)$$

Notice that the considered energy is minimized by either the zero function or a fluctuating function of which positive and negative values cancel each other.

Local Component The simplest local interaction model appears to be the sum of pairwise products of ω value at a site with ω values at the site's nearest neighbors. Such a component in the energy imposes regularity by favoring configurations in which each node agrees with its immediate neighbors. Assuming the usual four-connectivity, such that $\{(i+1,j),(i-1,j),(i,j+1),(i,j-1)\}$ defines the nearest neighbors of site (i,j), local interaction energy takes the form:

$$E_{Reg}^{L}(\omega_{i,j}) := -\left(\omega_{i,j}\cdot\omega_{i+1,j} + \omega_{i,j}\cdot\omega_{i-1,j} + \omega_{i,j}\cdot\omega_{i,j+1} + \omega_{i,j}\cdot\omega_{i,j-1} \right)$$

Differentiating w.r.t. $\omega_{i,j}$:

$$\frac{\partial E^L_{Reg}}{\partial \omega_{i,j}} = \frac{\partial}{\partial \omega_{i,j}} \sum_{(i,j)\in\Omega} -\left(\omega_{i,j}\cdot\omega_{i+1,j} + \omega_{i,j}\cdot\omega_{i-1,j} + \omega_{i,j}\cdot\omega_{i,j+1} + \omega_{i,j}\cdot\omega_{i,j-1}\right)$$

$$= -2\left(\omega_{i+1,j} + \omega_{i-1,j} + \omega_{i,j+1} + \omega_{i,j-1}\right) \tag{20.4a}$$

$$= -2\left[\mathbb{L}\left(\omega_{i,j}\right) + 4\omega_{i,j}\right] \tag{20.4b}$$

where $\mathbb{L}$ denotes the usual five-point discretization of the Laplace operator $\triangle := \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$, using step size *one*. Notice that (20.4a) is nothing but the sum of the four neighbors, and (20.4b) is obtained by re-arranging the terms after adding and subtracting $4\,\omega_{i,j}$.

The scope of the regularity induced by the local energy can be extended by considering pairwise interactions in a larger neighborhood $\mathscr{N}_{ij}$:

$$E^L_{Reg}(\omega_{i,j}) := -\sum_{(k,l)\in\mathscr{N}_{ij}} a_{i,j;k,l}\cdot\omega_{i,j}\cdot\omega_{k,l} \tag{20.5}$$

where $a_{i,j;k,l} = a_{k,l;i,j}$ are suitably chosen positive weights, say, inversely proportional to the squared distances between the pairs as defined by the neighborhood system. Then the derivative of E^L_{Reg} w.r.t. $\omega_{i,j}$ is nothing but a weighted sum of the neighbors:

$$\frac{\partial E^L_{Reg}}{\partial \omega_{i,j}} = \frac{\partial}{\partial \omega_{i,j}} \sum_{(i,j)\in\Omega} -\left(\sum_{(k,l)\in\mathscr{N}_{ij}} a_{i,j;k,l}\cdot\omega_{i,j}\cdot\omega_{k,l}\right)$$

$$= -2\sum_{(k,l)\in\mathscr{N}_{ij}} a_{i,j;k,l}\cdot\omega_{k,l} \tag{20.6}$$

As a result, similar to (20.4b), it can be be re-written as

$$-2\cdot\left[\mathbb{L}_{\mathscr{G}}\left(\omega_{i,j}\right) + \mu_{\mathscr{G}}\,\omega_{i,j}\right] \tag{20.7}$$

where $\mathbb{L}_{\mathscr{G}}(.)$ denotes the graph Laplacian and $\mu_{\mathscr{G}} > 0$ a constant that depends on the neighborhood system.

Of course, the scope of the local regularity may be further extended by including all of the pairwise interactions among the elements in $\mathscr{N}_{ij}$:

$$\sum_{\substack{(k,l)\in\mathscr{N}_{ij}\\(m,n)\in\mathscr{N}_{ij}\\(k,l)\neq(m,n)}} a_{k,l;m,n}\cdot\omega_{k,l}\cdot\omega_{m,n} \tag{20.8}$$

Fig. 20.3 The distance between a pair of nodes joined by a *blue arrow* is either 2 (the Manhattan) or $\sqrt{2}$ (the Euclidean)
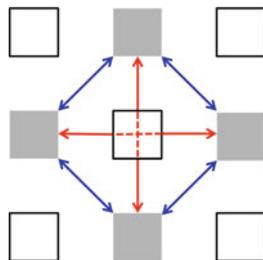
with the weights reflecting the distances (Fig. 20.3). Then the derivative w.r.t. $\omega_{i,j}$ is either of the following nine-point weighted averages in a 5×5 window centered at (i, j):

$$\frac{1}{4}\begin{bmatrix} \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & 2 & \cdot & 2 & \cdot \\ 1 & \cdot & \cdot & \cdot & 1 \\ \cdot & 2 & \cdot & 2 & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot \end{bmatrix} \qquad \frac{1}{4}\begin{bmatrix} \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & 4 & \cdot & 4 & \cdot \\ 1 & \cdot & \cdot & \cdot & 1 \\ \cdot & 4 & \cdot & 4 & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot \end{bmatrix}$$

The one the left is obtained if the weights are calculated according to the Manhattan, the natural metric for the 4-connected grid, and the one on the left according to the Euclidean. Upon subtracting a suitable multiple of $\omega_{i,j}$, both stencils become a realization of the nine-point discrete Laplacian on a redefined grid. That is, as long as local interactions are restricted to pairwise products (linear combinations of (20.5) and (20.8)), the derivative of the local energy w.r.t. $\omega_{i,j}$ is of the form:

$$\frac{\partial E_{Reg}^{L}}{\partial \omega_{i,j}} = -2\mu_1\Big[\mathbb{L}_*\big(\omega_{i,j}\big) + \mu_2\omega_{i,j}\Big] \tag{20.9}$$

where $\mathbb{L}_*$ is some approximation to the discrete Laplacian.

Adding Local and Global Components Combining (20.9) and (20.3) according to the additive model (20.2) yields

$$\frac{\partial E_{Reg}}{\partial \omega_{i,j}} = -2\Big[\mathbb{L}_*\big(\omega_{i,j}\big) + \bar{\mu}\,\omega_{i,j} - \frac{1}{O\big(|\Omega|\big)} \sum_{(k,l)\in\Omega} \omega_{k,l}\Big] \tag{20.10}$$

where $O\big(|\Omega|\big)$ denotes a positive constant on the order of the domain size and $\bar{\mu} > 0$ depends on assumed local connectivities as well as the parameter β in (20.2). Later on, we will see that the choice of $\bar{\mu}$ has no bearing on the result.

Boundary Energy Formulating interactions along a contour is not as simple as formulating them over a region. The task gets further complicated in the case of long-range interactions because critical ones occur among specific pairs of boundary points, e.g., two opposing indentation points on a neck region.

Fortunately, the usual distance transform provides a way out. At the outset, it is a transform that assigns to each point on a domain the minimum distance from the point to the domain boundary, yet it is at the same time a compact representation for boundary interactions. Indeed, its construction can be imagined as a dynamic process, a time-dependent evolution of the shape boundary such that each point on it moves with a unit speed in the direction of the unit normal; so the distance field t as a function over the digital plane is defined by setting $t_{i,j} =$ the time when the evolving curve passes through the point (i, j). Through the course of the evolution, singularities form as opposing pairs of boundary points meet and characterize the shape as its parts and symmetries are revealed [10].

The key point is that once a bridge between boundary interactions and the distance transform is established, E_{Bdy} is easily defined as a quadratic expression over the region Ω:

$$\sum_{(i,j)\in\partial\Omega} E_{Bdy}(\omega_{i,j}) := \sum_{(i,j)\in\Omega} \left(\omega_{i,j} - t_{i,j}\right)^2 \tag{20.11}$$

As a result,

$$\frac{\partial E_{Bdy}}{\partial \omega_{i,j}} = \frac{\partial}{\partial \omega_{i,j}} \sum_{(i,j)\in\partial\Omega} E_{Bdy}(\omega_{i,j})$$

$$= 2\left(\omega_{i,j} - t_{i,j}\right) \tag{20.12}$$

That is, $\omega = t$.

Complete Energy Combining (20.12) and (20.10) according to the additive model (20.1):

$$\frac{\partial E}{\partial \omega_{i,j}} = -\mathbb{L}_*(\omega_{i,j}) + \frac{1}{O(|\Omega|)} \sum_{(k,l)\in\Omega} \omega_{k,l} + (\lambda - \bar{\mu})\,\omega_{i,j} - \lambda t_{i,j} \tag{20.13}$$

Then the condition satisfied by the minimizer is obtained by setting the expression in (20.13) to *zero*, which is in matrix notation given by

$$\left(-\mathbf{L} + \frac{1}{O(|\Omega|)}\mathbf{J} + (\lambda - \bar{\mu})\,\mathbf{I}\right)\mathbf{w} = \lambda\,\mathbf{t} \tag{20.14}$$

where

- $\mathbf{w}$ and $\mathbf{t}$ are vector representations for the field ω and the distance transform t;
- $\mathbf{L}$ is the $|\Omega| \times |\Omega|$ matrix representation for a discrete Dirichlet Laplacian;
- $\mathbf{J}$ is the $|\Omega| \times |\Omega|$ matrix of ones;
- $\mathbf{I}$ is the $|\Omega| \times |\Omega|$ identity matrix.

Firstly, notice that the uniform scaling of the right hand side term with λ has no effect on the geometry of ω; it only scales the values, leaving the shape of the level curves intact. As a result, the geometric features of ω depends only on the choice of $(\lambda - \bar{\mu})$. Secondly, notice that $\left(-\mathbf{L} + \frac{1}{O(|\Omega|)} \cdot \mathbf{J} + (\lambda - \bar{\mu}) \cdot \mathbf{I}\right)$ is a positive definite matrix if $(\lambda - \bar{\mu}) \geq 0$. The difference between the strict positivity and zero, from the linear algebraic point of view, is that the former one has a better condition number so strict positivity of $(\lambda - \bar{\mu})$ may be preferable. But if $(\lambda - \bar{\mu})$ exceeds $\frac{1}{O(|\Omega|)}$, the non-spatial term dominates over the spatial ones. Such intuition suggest that $(\lambda - \bar{\mu})$, the only parameter, should be approximately $\frac{1}{|\Omega|}$. Later on this choice is supported in the continuous domain model as well.

Implementation Based on the above discussion the parameters $\bar{\mu} - \lambda$ and λ are eliminated: Firstly, $\bar{\mu} - \lambda$ is selected as inversely proportional to the domain size. Secondly, t is replaced by $\bar{t}$ denoting an arbitrary scaling of the distance transform because the right hand side scaling by λ does not affect the geometry of the solution.

Using the method of gradient descent, computing ω is cast as computing the steady state solution of the following differential-difference equation:

$$\frac{d\omega_{i,j}(\tau)}{d\tau} = -\frac{\partial E}{\partial \omega_{i,j}} \tag{20.15}$$

$$= \mathbb{L}_*(\omega_{i,j}) - \frac{1}{O(|\Omega|)} \sum_{(k,l) \in \Omega} \omega_{k,l} - \frac{1}{O(|\Omega|)}\omega_{i,j} + \bar{t}_{i,j}$$

Discretizing τ as

$$\frac{\omega_{i,j}^{n+1} - \omega_{i,j}^n}{\Delta \tau} = \left(\omega_{i,j+1}^n + \omega_{i,j-1}^n + \omega_{i+1,j}^n + \omega_{i-1j}^n - 4\omega_{i,j}^n\right) \tag{20.16}$$

$$- \frac{1}{|\Omega|} \sum_{(k,l) \in \Omega} \omega_{k,l}^n - \frac{1}{|\Omega|}\omega_{k,l}^{n+1} + \bar{t}_{i,j}$$

where $n = 0, 1, 2, \cdots$ is the iteration variable yields an iterative scheme once the terms in (20.16) are re-arranged. Note that $\mathbb{L}_*$ is simply replaced with the usual five point discrete Laplacian, and $O(|\Omega|)$ with $|\Omega|$.

Behavior of ω The solution satisfies three competing criteria:

$$- \text{having low expected value} \tag{20.17a}$$

$$- \text{being smooth} \tag{20.17b}$$

$$- \text{staying close to the usual distance transform} \tag{20.17c}$$

At a first glance, the first criterion seems to favor flatness by forcing the minimizer to attain values close to zero; yet this criterion is also satisfied when the values of

ω fluctuate between positive and negative. The other two criteria impose additional regularities so the fluctuations form a certain spatial pattern as opposed to being random localization of positive and negative values. Firstly, the locations of identical sign tend to form spatial proximity groups as defined by the graph neighborhood system. Secondly, the positive clusters tend to occur in the central region. This is due to the third criterion: Whenever ω is negative at a point (i, j), a penalty larger than $t_{i,j}^2$ is incurred; therefore, ω has a stronger tendency to be positive at locations where the value of distance transform is higher in order to minimize the total cost of deviating from the distance transform. Consequently, ω attains positive values in central locations where the value of the distance transform is higher and negative values at peripheral locations where the value of the distance transform is lower.

20.2.2 Illustrative Results

Few illustrative results are depicted in Figs. 20.4–20.6. In color illustrations, ω values are normalized to $[0, 1]$ within regions of identical sign for convenience of visualization and the zero-level is depicted in dark blue.

Central and Peripheral Separation In each case, the zero-level set partitions the shape domain into upper and lower zero level sets, Ω^+ and Ω^-, denoting respectively central and peripheral regions.

Ω^- contains all the detail: limbs, protrusions, and boundary texture or noise. In contrast, Ω^+ is coarse blob-like form, which can even be considered as an interval estimate of the center; most commonly, it is a simply connected set. Of course, it may also be either disconnected or multiply connected. For instance, it is disconnected for a dumbbell-like shape (two blobs of comparable radii combined through a thin neck) whereas it is multiply connected for an annulus formed by two concentric circles. Indeed, the annulus gets split into three concentric rings. Whereas the middle ring is the Ω^+, the union of the outermost and the innermost rings form a multiply-connected set corresponding to the Ω^-, separated by the Ω^+.

For quite a many shapes, however, Ω^+ is a simply connected set. Firstly, shapes obtained by protruding a blob as well as shapes whose peripheral parts are smaller or thinner than their main parts always have a simply connected Ω^+. This is expected: When the width of a part is small, the highest value of the distance transform t inside the part is small. That is, the local contribution to $(\omega - t)^2$ incurring due to negative values is less significant for such a part as compared to locations with higher positive values of t. Consequently, ω tends to attain negative values on narrow or small parts as well as on protrusions. Shapes with holes also have a simply connected Ω^+ as long as the holes are far from the center. Secondly, even a dumbbell-like shape may have a simply connected Ω^+. This happens if the join area, namely the neck, is wide enough, e.g., the top right in Fig. 20.4. Nevertheless, this does not cause any representational instability: Whereas the Ω^+ for a blob-like shape has a unique maximum located roughly at its centroid, the Ω^+ for a
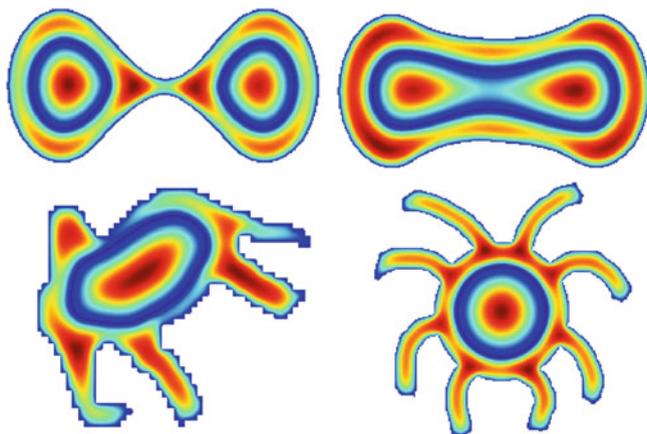
Fig. 20.4 Absolute value of ω. For ease of color visualization, values of ω are normalized within regions of identical sign. *Dark blue* color indicates zero

dumbbell-like shape has two local maxima indicating two bodies. Each body is captured by a connected component of an upper level set whose bounding curve passes through a saddle point. At a saddle point $\mathbf{p}$, such that $\omega(\mathbf{p}) = s$, the $s-level$ curve has a double point singularity, *i.e.* it forms a cross. As a result, the upper level set $\left\{\Omega^s = (x, y) \in \Omega^+ : \omega(x, y) > s\right\}$ yields two disjoint connected components capturing the two parts of the central region.

In contrast to Ω^+, the peripheral region, Ω^-, is multiply connected, even if the shape is simply connected. Most typically it is a ring-like domain whose holes are carved by Ω^+. It is also possible that Ω^- is disconnected. For instance, it is two concentric rings for an annulus. Additionally, Ω^- may be disconnected when there are several elongated limbs organized around a rather small central body, e.g., a palm tree. Ω^+, being small, is tolerated to grow and reach to the most concave parts of the shape boundary creating a split of Ω^- by the zero-level curve. The emergent zero-level curve bounds both Ω^+ and Ω^-.

Gap Closure An application in the case of incompletely specified boundaries is depicted in Fig. 20.5. The small figure on the left is the input, which consists of twelve isolated dots organized around a center point. The field ω computed in a rectangular region is depicted in the right. The gray is the central region; it is a circle. Notice that the level curves in the vicinity of the input points enclose firstly the individual dots and secondly the pairs of dots. Later the outer level curves enclose the entire set and the inner level curves become circles.

Restrictions of ω Each restriction of ω to either Ω^+ or Ω^- emulates a diffuse distance transform applied to the respective domain. The analogy is surely over the geometry of the level curves rather than actual ω values. More precisely, the level curves of ω in either domain resemble curvature dependent evolution of the boundary of the respective domain.

Fig. 20.5 Gap closure. (*left*)
The input. (*right*) The field
computed over a *rectangular*
region; the central region is
shown in *gray*
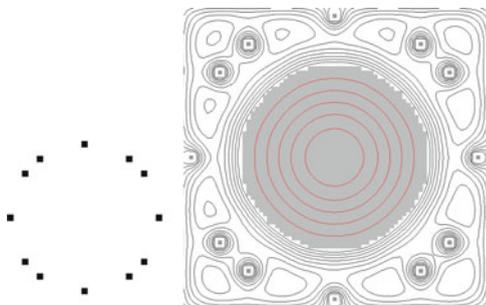
The qualitative resemblance of the level curves of ω to fronts evolving with curvature dependent speed triggers the observation that the condition satisfied by the minimizer of the discrete energy coding local/non-local and region/boundary interactions is indeed a central difference approximation for an integro-differential equation:

$$\triangle \, \omega(x, y) - \frac{1}{|\Omega|} \iint \omega(\alpha, \beta) \, d\alpha \, d\beta - \frac{1}{|\Omega|} \, \omega(x, y) \; = \; -\bar{t}(x, y) \quad (20.18)$$

where $\triangle$ stands for the Laplace operator. Recall that $\bar{t}$ denotes any uniform scaling of t. Interestingly, when the global component of the energy which is responsible for Ω^+/Ω^- separation is omitted, (20.18) reduces to the screened Poisson PDE:

$$\left(\triangle - \frac{1}{|\Omega|} \right) \omega(x, y) \; = \; -\bar{t}(x, y) \quad (20.19)$$

The screened Poisson PDE with a constant right hand side has been previously proposed by Tari et al. [54, 55] as the governing equation for a diffuse distance field, with a motivation to bridge the higher level process of shape abstraction and the lower level processes of image smoothing and boundary extraction. Tari, Shah and Pien (TSP) model is closely related to a famous by-product of variational segmentation models: the Ambrosio-Tortorelli (AT) phase field.

In Sect. 20.2.3, I relate (20.18) to a modification of the previous field which I call as the AT/TSP field. The connection between the two fields (AT/TSP and ω) improves our understanding of the new one. For example, the restriction of ω to either the central or peripheral region is geometrically equivalent to the AT/TSP field constructed for that region. Hence, the visual observation that the level curves in the vicinities of $\partial\Omega$ and the emergent zero-level curve are mimicking a curvature dependent motion is substantiated.

The curvature dependent behavior of the level curves, however, deteriorates away from the boundary; eventually, even initially parallel boundaries form a cross. This formation of saddle points is a consequence of increased interaction and it has been discussed previously by Tari et al. [55] (see Fig. 1 in [55]).

Fig. 20.6 Level *curves* of ω. For visualization purposes, the values are normalized within central and peripheral regions; the zero level is depicted in *dark blue*
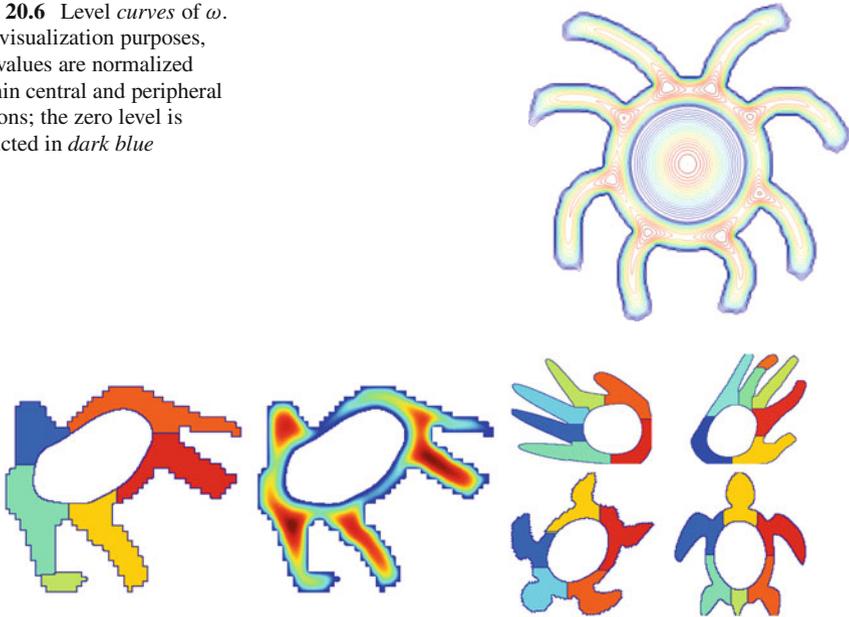
Fig. 20.7 Formation of watershed zones in the peripheral region is a consequence of the relation of the new field to the AT/TSP field

Fortunately, in the new framework, this behavioral deviation provides considerable convenience by assuring that the peripheral region is always partitioned, unless it is a perfect annulus. Figure 20.6 depicts level curves of ω for the octopus. Notice the formation of eight saddle points separating the arms of the octopus. Indeed the zones separated by the saddle points are easily extracted as watershed zones [31] of the peripheral region (Fig. 20.7). Of course, raw watershed partitions are subject to topologic instabilities. Obtaining topologically stable partitions is a separate issue which is tackled in a recent work by Tari and Genctav [51] by constructing a probabilistic part tree; nevertheless, watershed zones surrounding the central region are perfectly sufficient as rough part estimates.

20.2.3 *ω and the AT/TSP Field*

In vision applications, the AT (Ambrosio-Tortorelli) phase field appeared first as a technical device to apply gradient descent to the Mumford-Shah functional [35], where it served as a continuous indicator for boundary/not-boundary phases at each domain point [44]. The field is the minimizer of an energy composed of two competing terms: One term favors configurations that take values close to either 0 or 1 (separation into boundary/not-boundary phases) and the other term

encourages local interaction in the domain by penalizing spatial inhomogeneity. A parameter controls the relative influence of these two terms hence the interaction. As this "interaction" parameter tends to 0, the separation term is strongly emphasized; consequently, the field tends to the characteristic function $1 - \chi_S$ of the boundary set S and the respective energy tends (following the Γ convergence framework [11]) to the boundary length.

Over the years, the AT field has been used in numerous applications addressing a rich variety of visual problems, cf. Shah and colleagues [39, 44, 45, 48], March and Dozio [29], Proesman et al. [41], Teboul et al. [56], Bar et al. [6], Droske and Rumpf [40], Erdem and Tari [18], Patz and Preusser [38], and Jung and Vese [24]. In all these applications, because the field serves as an auxiliary variable to facilitate discontinuity-preserving smoothing and boundary detection, the interaction parameter ρ is chosen sufficiently small, i.e., $\rho \to 0$, to localize boundaries accurately.

Increased Interaction: the AT/TSP Field In the late 1990s, Tari, Shah and Pien (TSP) re-casted the AT field as a compact representation of the morphology of regions implied by putative edges [54, 55]. To such end, they weakened boundary/not-boundary separation by letting the interaction parameter be sufficiently large, in contrast to the prevalent applications of the AT field exemplified by the above referenced works. Furthermore, they examined the geometry of the level curves and then devised robust schemes that explicate curvature related criticalities such as skeletons.

On the technical side, they used standard coupled equations obtained by applying gradient descent to the AT approximations of available segmentation energies [44, 45]. For example, given an image $g(x, y)$ defined on a region $R \subset \mathbf{R}^2$, they computed a de-noised image $u(x, y)$ and the AT field $v(x, y)$ by solving

$$\frac{\partial u}{\partial \tau} = \nabla \cdot (v^2 \nabla u) - \frac{\beta}{\alpha}(u - g) \tag{20.20}$$

$$\frac{\partial v}{\partial \tau} = \triangle v - \frac{2\alpha |\nabla u|^2 v}{\rho} - \frac{(v-1)}{\rho^2} \tag{20.21}$$

subject to homogeneous Neumann conditions on the image boundary:

$$\left.\frac{\partial u}{\partial n}\right|_{\partial R} = 0, \quad \left.\frac{\partial v}{\partial n}\right|_{\partial R} = 0 \tag{20.22}$$

where ∂R is the image boundary and n is the outer unit normal vector to ∂R, while α and β are Mumford-Shah related parameters determining contrast and scale [45, 55], ρ is the interaction parameter. The key trick is to choose ρ sufficiently large, typically at the order of the radius of the maximal circle that can be fit inside putative shapes – regions implied by putative boundaries.

That is, in the TSP approach, boundary localization is traded with shape abstraction, thus, the role of AT field changed from being a fixed property detector

(boundary/no-boundary) to being an implicit shape coder. From now on, I refer to the v function as AT/TSP field if local interactions are emphasized during its computation.

When the computation of the AT/TSP field is restricted to shape interior (indicated by $\Omega \subset \mathbf{R}^2$, a connected bounded open set with a boundary $\partial\Omega$), the field satisfies on Ω a screened Poisson PDE with a constant source term:

$$\left(\triangle - \frac{1}{\rho^2} \right) v(x, y) = -\frac{1}{\rho^2} \tag{20.23}$$

subject to homogenous Dirichlet condition. The solution to (20.23) is the unique minimizer of the AT energy:

$$\iint\limits_{\Omega} \frac{1}{\rho} \underbrace{\left[v(x, y) - \chi_\Omega(x, y) \right]^2}_{\substack{\text{boundary/interior} \\ \text{separation}}} + \rho \underbrace{\left| \nabla v(x, y) \right|^2}_{\text{local interaction}} \, \mathrm{d}x \, \mathrm{d}y$$

$$\text{with } v(x, y) = 0 \text{ for } (x, y) \in \partial\Omega \tag{20.24}$$

where $\chi_\Omega(x, y)$ is the shape indicator that attains 1 in Ω and 0 on $\partial\Omega$.

There are two important practical implications of the AT/TSP model:

1. v is a smooth distance transform of which gradient at a point exponentially depends on the medialness of the point; this leads to a robust procedure for extracting curvature based criticalities [53, 55];
2. The level curves of v approximate evolution of the shape boundary such that each point on it moves with a curvature-dependent speed in the direction of the unit normal; that is, a nonlinear wave-like phenomena can be modeled using a linear diffusion operator; as a result, the reaction-diffusion scale space of Kimia et al. [25] can be easily constructed [54].

The right hand side of (20.23) may be scaled if the interest is purely on the geometry of the level curves. Then *infinite* interaction limit $\rho \to \infty$ can be considered, giving the Poisson PDE [22] or approximating the Aslan-Tari model [3, 4]; but then the exponential dependence on medialness [55] is lost. In the original paper by Tari, Shah and Pien, there are several experiments illustrating the role of ρ. In what follows, ρ is assumed to be at least on the order of the radius of the maximal circle that can be fit inside the shape for the diffusive effect of $|\nabla.|^2$ to influence the entire shape.

Now I modify (20.24) in a way that its minimizer is geometrically equivalent to the ω function defined earlier in Sect. 20.2.1 as a balance among different types of interactions. Take a note that the AT/TSP field which is the solution of the screened Poisson equation with a constant source term (20.23) is geometrically equivalent to the solution of a screened Poisson equation for which the source term is a monotonic function of the distance (20.19); the change of the source term affects the actual values while leaving the geometry of the level curves almost intact.

Fluctuating AT/TSP The key idea is to break down the monotone behavior of the AT/TSP to partition the shape into central and peripheral regions within which the field remains qualitatively similar to the AT/TSP. This can be achieved by additively augmenting the energy in (20.24) with a term minimized by a fluctuating function, e.g. a multiple of $\left(\iint \omega(x, y) \, dx \, dy \right)^2$, but one has to make sure that locations of identical sign form meaningful spatial proximity groups. The term $\left| \nabla v(x, y) \right|^2$ prevents wild fluctuations and ensures that nearby locations have similar values, but it can not influence where these groups form. An obvious way of obtaining central versus peripheral separation is to increase the residual $\left[v(x, y) - \chi_\Omega(x, y) \right]^2$ at central regions to penalize discrepancies more, hence, forcing the function to be positive in a central area while attaining the opposite sign in remaining locations to minimize penalty incurring due to the new term $\left[\iint \omega(x, y) \, dx \, dy \right]^2$.

There are two possible strategies: Either making the phase separation term space variant by multiplying it with a medialness dependent coefficient or using a medialness dependent indicator function. In both cases, medialness dependence can be modeled via any monotonic function of the distance. Using the second strategy and letting the distance transform $t(x, y)$ serve as the weighted indicator yields

$$\iint\limits_{\Omega} \frac{1}{\rho} \left[\omega(x, y) - t(x, y) \right]^2 + \rho \left[\left| \nabla \omega(x, y) \right|^2 + \left(\frac{1}{|\Omega|} \iint\limits_{\Omega} \omega(\alpha, \beta) \, d\alpha \, d\beta \right)^2 \right] dx \, dy,$$

$$\text{with } \omega(x, y) = 0 \text{ for } x, y = (x, y) \in \partial\Omega.$$

$$(20.25)$$

Similar to the AT/TSP field, the minimizer of (20.25) is a compromise between inhomogeneity and homogeneity though the inhomogeneity is forced both externally (by t) and internally (by the third term). Unlike the phases defined by the AT/TSP field (its level curves), those defined by ω do not necessarily mimic erosions of the shape boundary but they have a richer meaning that facilitates extraction of a novel skeleton as detailed in the next section.

To find the minimizer of (20.25), let us consider the Gâteaux variation

$$\delta E(\omega, v) := \lim_{\epsilon \to 0} \frac{E(\omega + \epsilon v) - E(\omega)}{\epsilon} = \left. \frac{\partial}{\partial \epsilon} E(\omega + \epsilon v) \right|_{\epsilon = 0} \qquad (20.26)$$

for all test functions v properly prescribed on the boundary. Setting $\delta E(\omega, v) = 0$ yields

$$\iint\limits_{\Omega} \left\{ \frac{1}{\rho^2} \left(\omega - t \right) - \triangle\omega \right\} v + \mathbb{E}[\omega]\mathbb{E}[v] = 0 \qquad (20.27)$$

where

$$\mathbb{E}[f] = \frac{1}{|\Omega|} \iint_{\Omega} f(x, y) \mathrm{d}x \, \mathrm{d}y.$$

The condition is not intrinsic, but it is is satisfied for all test functions if both of the following holds:

$$\frac{1}{|\Omega|} \iint \omega(\alpha, \beta) \, \mathrm{d}\alpha \, \mathrm{d}\beta = 0 \tag{20.28a}$$

$$\triangle \, \omega(x, y) - \frac{1}{\rho^2} \omega(x, y) = -\frac{1}{\rho^2} t(x, y) \tag{20.28b}$$

Even though the above pair of equations and (20.18) are not identical mathematically, they are equivalent conceptually, i.e., constraints induced by the pair (20.28a) and (20.28b) are equivalent to those in the discrete model:

Firstly, (20.28a) implies that the solution integrates to *zero* (compare to (20.17a) in the discrete case). Secondly, (20.28b) implies that the desired solution is a smooth approximation of the distance transform (compare to the pair (20.17b) and (20.17c) in the discrete case). Selecting the smoothing radius as consistent with the suggested practice in the AT/TSP model, i.e., $\rho = O\left(\sqrt{|\Omega|}\right)$, is in agreement with the choice of $\lambda - \bar{\mu}$ in the discrete model. (Recall that ρ needs to be approximately equal to the radius of the maximal circle that can be fit inside the shape for the diffusive effect of $|\nabla.|^2$ to influence the entire shape; see also [3]).

20.3   Three Partite Skeletons

The invention of skeletons is one of the most influential outcomes of the search for a mathematics of shape:

In the mid 1960s the school of mathematical morphology [43] brilliantly opted to understand a shape merely by observing how it interacts with template shapes called structuring elements using two basic processes called erosion and dilation. Eroded and dilated shapes define a *time* evolution in which the time is the size of the structuring element, hence, shapes are equipped with dynamics.

Independent from the school of mathematical morphology, Blum [10] proposed to code topologic and geometric properties implicitly using the Euclidean distance field $t(x, y)$. The relevant link is that successive level sets of the distance transform correspond to erosions and dilations of the shape with discs of increasing radius making the Euclidean distance field a compact representation of morphology. The level curves of the Euclidean distance transform are analogous to a time evolution of the shape boundary such that each point on it moves with a unit speed in the direction of the unit normal, so the distance field t as a function over the plane is defined by setting $t(x, y) = $ the time when the evolving curve passes through the

point (x, y). During the evolution opposing pairs of boundary points meet, each collapsing to a single point inside the shape. The locus of meeting points is called the skeleton. Each point on the skeleton is equidistant from at least two boundary points.

In the following years, several practitioners conceptualized skeletons strictly as thin versions of shapes with equidistance property; vast literature is dedicated to thinning and solving consequently arising problems; axioms that are necessitated by chosen algorithms rather than being induced by theory (e.g. the connectivity requirement) are introduced; holding back nonconforming variants.

Nevertheless, significant liberation is gained in the 1990s when vision researchers started to reveal connections among different computational models and algorithms to visual processing [12, 33, 45, 46]. These works contributed towards a better understanding of the theory level of visual processing emphasizing *what the computational problem is rather than how it is being solved* [30].

In particular, mathematical morphology, skeletons, curve evolution approaches, diffusion PDEs and variational methods are naturally tied by distance-like fields [28], reviving the Gestalt school's potential field theory. The Euclidean Distance Transform is given by the solution to the Eikonal equation $|\nabla t(x, y)| = c(x, y)$ with the shape boundary as the initial front and $c(x, y) = 1$ and new distance fields can be constructed by varying $c(x, y)$ resulting in weighted distance transforms or augmenting the initial front with seed points resulting in seeded distance transforms [28]. Of course a weighted distance transform of which level sets emulating curvature dependent motion is readily constructed via an elliptic PDE, namely the Poisson equation, be it screened [55] or not [22].

Indeed the ω function is equivalent to a seeded weighted distance transform: Consider the shape boundary depicted by the red curve in Fig. 20.8a, and imagine that the initial front is its augmentation with two seed points. During the course of the evolution, fronts growing out of each of the two seeds merge forming the blue curve in (a); then the blue curve propagating outward and the shape boundary (the red) propagating inward meet at the zero level curve (the pink). In Fig. 20.8b more level curves in the region where ω is negative, namely Ω^-, are depicted.

New distance fields give way to new skeleton models, one of which is the Three-Partite Skeleton.

The Three-Partite Skeleton consists of three groups. Recall Fig. 20.2. The first group S_G (shown in red) is an abstraction of the central or gross structure. It consists of two components: the inner component S_G^{in} (the dark red) coinciding with the zero-level curve and the outer component S_G^{out} (the light red). The second group S_R (shown in blue) abstracts ribbon-like sections corresponding to limbs and significant protrusions. Finally the third group S_T (shown in green) captures secondary details typically arising due to boundary texture or noise.

Conceiving the Three Partite Skeleton The relation between the level curves of ω and those of the AT/TSP field makes conceiving the Three-Partite Skeleton quite natural: With the exception of S_G^{in}, which is readily provided by the zero level curve of ω, all of the points of the Three-Partite Skeleton are in a generalized ridge set
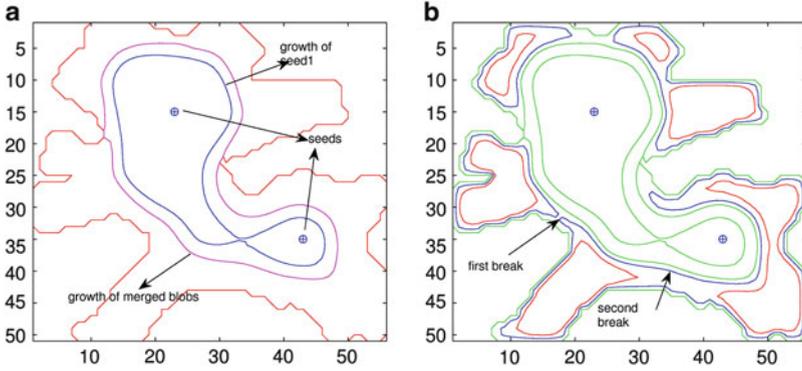
**Fig. 20.8** Seeded distance transform analogy for ω

defined by the zero-crossings of the flowline curvature, i.e., $\frac{d|\nabla\omega|}{ds} = 0$ where s is the arc-length along level curves. Moreover, the ridge points can be split into maximum and minimum convexity groups depending on the sign of the second derivative $\frac{d^2|\nabla\omega|}{ds^2}$ [55]. In Fig. 20.9, the points are depicted in black if $\frac{d^2|\nabla\omega|}{ds^2} < 0$ and in light gray if $\frac{d^2|\nabla\omega|}{ds^2} > 0$; skeleton points tend to be on the first group.

Branches tracking discretization artifacts are clearly visible in the form of disconnected short segments; they belong to S_T. Branches tracking main limbs are clearly visible too; they belong to S_R. Each S_R branch starts from a curvature maximum of the shape boundary $\partial\Omega$ and comes to an end at a point where the front initiated at $\partial\Omega$ meets the front initiated at the emergent zero level curve $\partial\Omega^+$, during the course of propagation. The termination points are in a sub-locus composed of a circular sequence of saddles and minima so defining a numerical realization of a parabolic locus [52]. This locus is S_G^{out}. All points in S_G^{out} are simultaneously reachable from $\partial\Omega$ and $\partial\Omega^+$; as such, S_G^{out} defines a *minimal* skeleton for Ω^-, the ring-like region bounded by the shape boundary $\partial\Omega$ and the emergent zero level curve $\partial\Omega^+$.

It is important that the level curves of the restriction of ω to Ω^- behave similar to the level curves of a highly diffused distance transform: Only if the tip of a protrusion moves sufficiently faster than the opposing point on $\partial\Omega^+$, the respective fronts meet at a point that is sufficiently away from the tip, leaving enough room for the respective branch to grow. As a result, a branch tracking a limb terminates when the limb connects to a body, making the branch length a good correspondent of the limb length.

The ridge locus takes an irregular appearance so becomes difficult to interpret as the level curves get rounder. Fortunately, the irregularities are to a large extent within the confines of the central region Ω^+, thus, can be cut out (Fig. 20.10).

Grouping the set of ridge points in the form of segments (branches) is an implementation issue, typically subject to several ad-hoc decisions. In the case of skeletons from ω, the implementation can be considerably simplified by confining
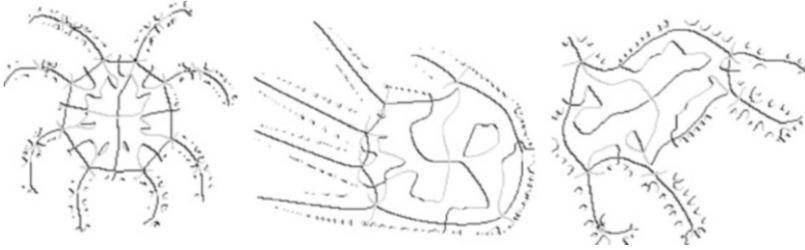
Fig. 20.9 Ridge points extracted by applying Tari et al. method [55] on ω

Fig. 20.10 Removing the ridge points that are inside Ω^+ (the *light gray* region) significantly reduces the set of candidate points. The *thick red curve* is S_G^{out}
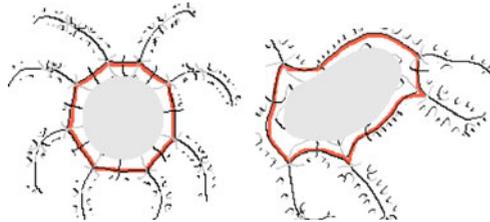

the grouping process to individual zones (recall that the peripheral region, unless it is a perfect annulus, always breaks into parts). Then grouping within each zone may be cast as a curve tracing task, performed by starting either from the boundary or the interior [2,20]. Unfortunately both cases necessitates too many ad-hoc decisions. A robust formulation via energy minimization is under development. In the mean time, leaving the grouping task as a technical issue that requires further attention, a sample set of branches is depicted in Fig. 20.11 in order to convey the concept. Typically the branches in S_R and S_T are linear disconnected elements, but there may be loop-like branches too, which occur if the shape has a hole in its peripheral region (Fig. 20.12).

Whether a branch belongs to S_R or S_T becomes unclear if limbs have limbs too (e.g. fingers of an arm). On one hand, it may be natural to put such branches in S_T since they track secondary details. On the other hand, it may be critical to distinguish them from noise, discretization artifacts and boundary texture. A generic strategy could be to assign fuzzy membership to such branches. This all depends on the task at hand. (The more important point is not to prune any branch.)

If the task at hand necessitates distinguishing the limbs of limbs from other secondary details, proper criteria need to be defined. Typically, within each zone, S_R branches are distinguished by the proximity of their termination point to a fixed locus, i.e. S_G^{out}. Likewise, a branch denoting a limb of a limb may be distinguished by the proximity of its termination point to a branch in S_R. A better strategy, however, is to consider the contribution of each branch to the reconstruction of the shape, i.e. the effective coverage of a branch, defined as the ratio of the area exclusively covered by the branch to its total coverage. Note that the region covered by a skeleton branch is the envelope of all circles such that each circle is centered at a skeleton point on

Fig. 20.11 Sample branches
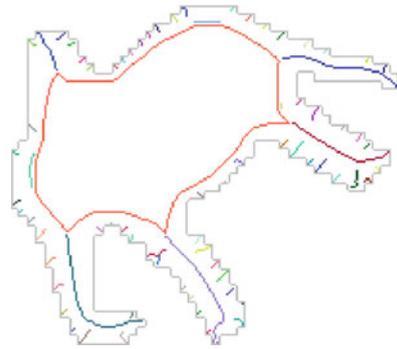
Fig. 20.12 Shape with a hole
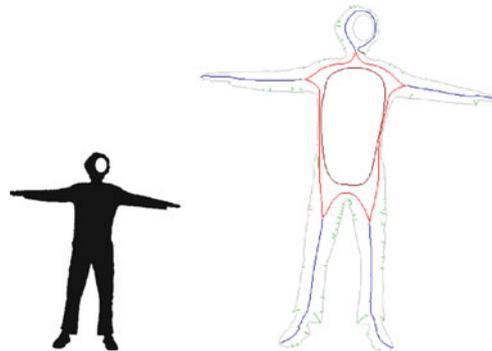
the branch, and has a radius given by the value of the distance transform at that point [10].

In general, computing the effective coverage is difficult [19] but becomes easy in our case because the branches can be put into a priority order, at least within a zone. This order is based on the branch length and proximity to a fixed locus, be it S_G^{out} or members of S_R.

Effective coverages of the branches for the elephant shape are depicted in Fig. 20.13. The gray toning of the covered regions (left) reflects the magnitudes of the effective coverage ratios of the corresponding branches. On the right, the five branches forming S_R covers almost the entire peripheral region.

20.3.1 Why Three Partite Skeletons?

I make four arguments in favor of the new skeleton model:

Boundary Details The most frequently reported problem of the classical skeleton is its sensitivity to boundary perturbations: small indentations or protrusions introduce spurious branches. A classical connectivity-preserving skeleton for the

Fig. 20.13 Coverage of
branches in S_R and S_T. The
gray toning of the covered
regions (*left*) reflects the
magnitudes of the effective
coverage ratios of the
corresponding branches
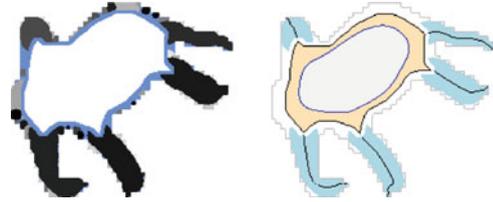
elephant shape is shown in Fig. 20.14; compare it to the Three-Partite one of the
same shape (Fig. 20.2).

Both skeletons contain spurious branches due to the wiggly boundary. Neverthe-
less, in the new model, they remain short and isolated, hence easily distinguishable,
from the main branches. In contrast, in the classical model, it is difficult to identify
spurious branches; consequently, there has been an extensive research on defining
criteria for branch pruning [5,19,27]. Pruning criteria and methods differ, but almost
all of the working skeleton methods reduce skeletons by eliminating those branches
that contribute little to the reconstruction of the shape.

On one hand, reduced skeletons decrease classification errors when experimented
using typical shape databases because secondary details most of the time act like
noise [4]. On the other hand, in a different context, secondary details may be
characterizing features: Observe that separation of the shapes shown in Fig. 20.15
into two perceptual proximity groups, as depicted in Fig. 20.16, is based on the
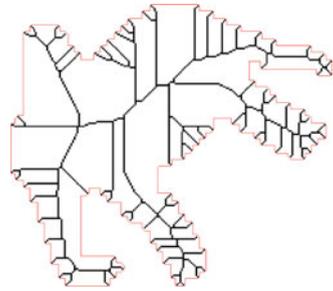boundary texture, a secondary detail subject to pruning.

Thus, it is essential not to throw away spurious branches but to separate them
from primary ones, keeping in mind that the purpose of perceptual organization
is to organize and gradually explicate relations. In the Three-Partite Skeleton the
desired separation is achieved by letting each branch to be disconnected as a result
of extensive smoothing. Disconnected skeletons may be uncommon but have always
been around: Earlier examples are Leyton [26], Burbeck and Pizer [14], and Tari
et al. [55], and more recent examples are Aslan and Tari [3,4] and Macrini et al. [27].

Perceived Topology Preserving topology is the common promise of skeleton
models, but it may cause instability because topologic equivalence is not perceptual
equivalence. Let us consider a set of peanuts and blobs (Fig. 20.17a). Where should
we place category boundaries? Is the fifth shape more similar to the second or to the
last one? How different are the first two shapes?

It seems natural to consider multiple abstractions of the topology. In the case of
Three-Partite skeletons, this is achieved by the layered structure of S_G considered
together with the critical points of ω inside the region enclosed by S_G^{in}. The
following is important: If a simple closed curve in S_G^{in} is blob-like, ω has a unique
maximum located roughly at the centroid of the region it encloses; but if it has
indentations, ω has at least two local maxima roughly located at the centroids of
implied blobs.

Let us consider the four Three-Partite Skeletons depicted in Fig. 20.17b, corre-
sponding to the four peanuts shown in the middle group in (a). As before, S_G^{out} and

Fig. 20.14 A connectivity preserving skeleton. The branch length is not a good indicator of significance


Fig. 20.15 A collection of shapes

Fig. 20.16 Depending on the context, a secondary detail (e.g. boundary texture) may be the key feature
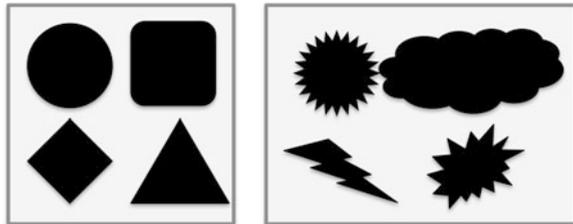
S_G^{in} are respectively shown as light and dark red; local maxima pairs are combined via a magenta line whenever applicable. The S_G of the leftmost shape consists of a pair of nested closed curves; the inner component S_G^{in} is blob-like so encloses the unique maximum of ω; signaling that the shape is a single blob. As the neck of the shape thins giving the second shape, S_G^{in} develops indentations, so the maximum becomes the saddle separating two local maxima, indicating that the shape is a composition of two blobs. As the neck thins further giving the third shape, S_G^{in} breaks into two curves, each of which is blob-like; the outer layer S_G^{out} still remains connected. On one hand, the second shape is similar to the first shape because its skeleton is a single pair of nested closed curves; on the other hand, S_G^{in} encloses two local maxima signaling that the shape is a composition of two blobs. That is, the second shape is a borderline shape for which two abstractions of the topology should be retained. Finally, as the neck sufficiently thins to yield the rightmost shape, the outer loop also splits; signaling even more strongly that the shape is a composition of two blobs. The fourth peanut (the fifth in Fig. 20.17a) is another borderline shape. The magenta line connecting the pair of local maxima passes though a saddle point region (light blue) which falls outside S_G. Hence the alternative abstraction connects the last shape to two discrete blobs.
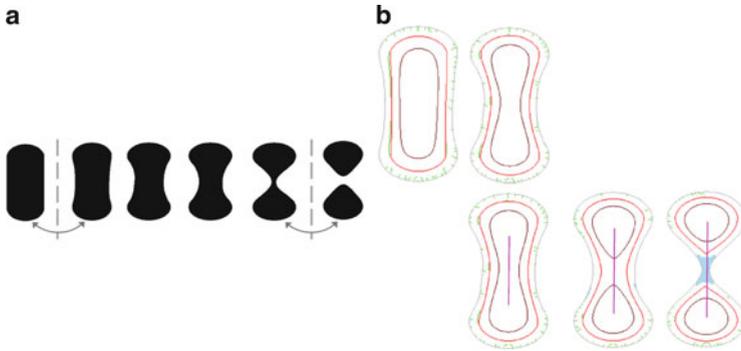
Fig. 20.17 (**a**): Peanuts of varying neck thickness and blobs. (**b**): The Three-Partite Skeletons of the four shapes depicted in the middle group in (**a**). The new model enables multiple interpretations of the shape topology. Transition shapes provide paths between blobs and peanuts

Fig. 20.18 Centrality
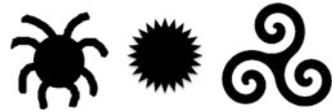
Centrality for Nearly Symmetric Shapes The classical skeleton of a circle is a single point located at its center reflecting the shape's symmetry. Likewise the branches of the skeleton of n-fold symmetric stars (or n-fold symmetric polygons) meet at the respective shape centroid; unfortunately, a slight deviation from the symmetry changes the connectivity pattern, e.g., the branches tracing the eight arms of the octopus in Fig. 20.18 do not all meet at a single point. This means that skeleton based abstractions fail to capture *centrality*, a natural consequence of fold symmetry (e.g., Figs. 20.18 and 20.5). This is rather ironic because the points of the classical skeleton are the medial ones.

In the new model, the notion of medialness is revised. S_G is a robust explication of the centrality. Indeed, the region bounded by S_G^{in}, namely Ω^+, can be considered as an interval estimate of the center, a robust alternative to a point estimate. The problem associated to connectivity pattern is solved by considering the branches outside S_G as they remain disjoint.

Notice that in Fig. 20.2, the eight disjoint branches in S_R (blue) tracing the eight arms of the octopus terminate without distorting the central part, which is a *perfect* circle. The S_G^{in} of the prickly ball is also a circle whereas the S_G^{in} of the triskelion is a triangle. For both shapes, S_R is empty; the centrality of the triskelion is emphasized (Fig. 20.19) consistent with the historic meaning of this shape.

Ligature Problem The fragility of the branch connectivity pattern is not specific to nearly symmetric shapes. Instabilities near junctions almost always occur; this is commonly addressed as the *ligature* problem [10]. One way to deal with the ligature problem is to restrict skeleton extraction to ribbon-like sections where the

Fig. 20.19 Consistent with
the historic meaning of the
triskelion, centrality is
emphasized
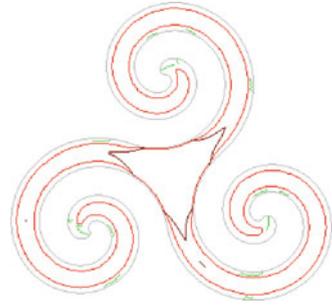
procedure is stable: While Shah [47] completed missing parts of the skeleton via
interpolation, Aslan and Tari [3] kept the skeleton disconnected. Recently Macrini
et al. [27] also proposed a disconnected skeleton by removing the ligature part of
the classical skeleton.

Ligature problem is all the more reason to cut out a central part.

20.4   Summary and Concluding Remarks

A distance-like field that attains both positive and negative values is described. It
emerges out of simple spatial interactions, both local and global. It is a feature
conscience field: A hierarchy of parts starting with peripheral versus gross structure
separation is readily provided by its level curves, which mimic curvature dependent
motion (so code curvature), and its generalized ridge points give rise to a generalized
medial locus that is organized in the form of an unconventional skeleton. The new
skeleton model suggests, among other things, (1) multiple interpretation of the shape
topology, and (2) *gross structure + texture + part* separation. Incentives for the new
skeleton model are discussed.

If the global interaction is ignored, the set of linear equations defining the field
reduces to a discrete implementation of a screened Poisson PDE, a computationally
advantageous alternative model for constructing reaction-diffusion scale space for
shapes. Exploiting a connection between the screened Poisson PDE and the famous
phase field of Ambrosio and Tortorelli, a connection between the new field and a
modification to the Ambrosio-Tortorelli model is established. Implication of this
connection is two fold: (1) a novel skeleton model, for which strong perceptual
incentives exist, is conceived, (2) the issue of bridging image segmentation and
shape analysis is raised, following former work by Tari, Shah and Pien.

References

1. Ambrosio, L., Tortorelli, V.: On the approximation of functionals depending on jumps by elliptic functionals via Γ-convergence. Commun. Pure Appl. Math. **43**(8), 999–1036 (1990)
2. Aslan, C.: Disconnected skeletons for shape recognition. Master's thesis, Department of Computer Engineering, Middle East Technical University (2005)
3. Aslan, C., Tari, S.: An axis-based representation for recognition. In: Proceedings of the ICCV, pp. 1339–1346. Springer, Berlin/New York (2005)
4. Aslan, C., Erdem, A., Erdem, E., Tari, S.: Disconnected skeleton: shape at its absolute scale. IEEE Trans. Pattern Anal. Mach. Intell. **30**(12), 2188–2203 (2008)
5. Bai, X., Latecki, L., Liu, W.Y.: Skeleton pruning by contour partitioning with discrete curve evolution. IEEE Trans. Pattern Anal. Mach. Intell. **29**, 449–462 (2007)
6. Bar, L., Sochen, N., Kiryati, N.: Image deblurring in the presence of impulsive noise. Int. J. Comput. Vis. **70**(3), 279–298 (2006)
7. Barenholtz, E., Feldman, J.: Visual comparisons within and between object-parts: evidence for a single- part superiority effect. Vis. Res. **43**, 1655–1666 (2003)
8. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. IEEE Trans. Pattern Anal. Mach. Intell. **24**, 509–522 (2002)
9. Biederman, I.: Recognition-by-components: a theory of human image understanding. Psychol. Rev. **94**(2), 115–117 (1987)
10. Blum, H.: Biological shape and visual science. J. Theor. Biol. **38**, 205–287 (1973)
11. Braides, A.: Approximation of Free-Discontinuity Problems. Lecture Notes in Mathematics, vol. 1694. Springer, Berlin/New York (1998)
12. Brockett, R., Maragos, P.: Evolution equations for continuous–scale morphology. In: Proceedings of the ICASSP, vol. 3, pp. 125–128. IEEE, Piscataway (1992)
13. Buades, A., Coll, B., Morel, J.M.: A non-local algorithm for image denoising. In: Proceedings of the CVPR, pp. 60–65. IEEE Computer Society, Los Alamitos (2005)
14. Burbeck, C.A., Pizer, S.M.: Object representation by cores: identifying and representing primitive spatial regions. Vis. Res. **35**, 1917–1930 (1995)
15. Cohen, E.H., Singh, M.: Geometric determinants of shape segmentation: tests using segment identification. Vis. Res. **47**, 2825–2840 (2007)
16. de Winter, J., Wagemans, J.: Segmentation of object outlines into parts: a large-scale integrative study. Cognition **99**, 275–325 (1999)
17. Edelman, S., Cutzu, F., Duvdevani-Bar, S.: Similarity to reference shapes as a basis for shape representation. In: Proceedings of the COGSCI, San Diego (1996)
18. Erdem, E., Tari, S.: Mumford-Shah regularizer with contextual feedback. JMIV **33**, 67–84 (2009)
19. Feldman, J., Singh, M.: Bayesian estimation of the shape skeleton. PNAS **103**(47), 18014–18019 (2006)
20. Genctav, M.: Matching global skeleton. Master's thesis, Department of Computer Engineering, Middle East Technical University (2010)
21. Gonzalez, R., Woods, R.: Digital Image Processing, 2nd edn. Addison-Wesley Longman Publishing Co., Inc., Boston (2001)
22. Gorelick, L., Galun, M., Sharon, E., Basri, R., Brandt, A.: Shape representation and classification using the poisson equation. IEEE Trans. Pattern Anal. Mach. Intell. **28**(12), 1991–2005 (2006)
23. Hofmann, D., Richards, W.: Parts of recognition. Cognition **18**, 65–96 (1984)
24. Jung, M., Vese, L.: Nonlocal variational image deblurring models in the presence of gaussian or impulse noise. In: Proceedings of the SSVM, pp. 401–412. Springer, Berlin/New York (2009)
25. Kimia, B., Tannenbaum, A., Zucker, S.: Shapes, shocks, and deformations I: the components of two-dimensional shape and the reaction-diffusion space. Int. J. Comput. Vis. **15**(3), 189–224 (1995)

26. Leyton, M.: A process-grammar for shape. Art. Intell. **34**(2), 213–247 (1988)
27. Macrini, D., Dickinson, S., Fleet, D., Siddiqi, K.: Bone graphs: medial shape parsing and abstraction. Comput. Vis. Image Underst. **115**, 1044–1061 (2011)
28. Maragos, P., Butt, M.A.: Curve evolution, differential morphology and distance transforms as applied to multiscale and eikonal problems. Fundam. Inf. **41**, 91–129 (2000)
29. March, R., Dozio, M.: A variational method for the recovery of smooth boundaries. Image Vis. Comput. **15**(9), 705–712 (1997)
30. Marr, D.: Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. W.H. Freeman, San Francisco (1982)
31. Meyer, F.: Topographic distance and watershed lines. Signal Process. **38**, 113–125 (1994)
32. Mi, X., DeCarlo, D.: Separating parts from 2d shapes using relatability. In: Proceedings of the ICCV, pp. 1–8. IEEE Computer Society, Los Alamitos (2007)
33. Morel, J.-M., Solimini, S.: Variational Methods in Image Segmentation. Birkhäuser, Boston (1995)
34. Mumford, D.: Mathematical theories of shape: do they model perception? In: Proceedings of the SPIE, vol. 1570. SPIE, Bellingham (1991)
35. Mumford, D., Shah, J.: Optimal approximations by piecewise smooth functions and associated variational problems. Commun. Pure Appl. Math. **42**, 577–685 (1989)
36. Navon, D.: Forest before trees: the precedence of global features in visual perception. Cogn. Psychol. **9**, 355–383 (1977)
37. Pasupathy, A., Connor, C.: Population coding of shape in area V4. Nat. Neurosci. **5**(2), 1332–1338 (2002)
38. Patz, T., Preusser, T.: Ambrosio-Tortorelli segmentation of stochastic images. In: Proceedings of the ECCV, pp. 254–267. Springer, Berlin (2010)
39. Pien, H.H., Desai, M., Shah, J.: Segmentation of mr images using curve evolution and prior information. IJPRAI **11**(8), 1233–1245 (1997)
40. Preußer, T., Droske, M., Garbe, C., Rumpf, M., Telea, A.: A phase field method for joint denoising, edge detection and motion estimation. SIAM J. Appl. Math. **68**(3), 599–618 (2007)
41. Proesman, M., Pauwels, E., van Gool, L.: Coupled geometry-driven diffusion equations for low-level vision. In: Romeny, B. (ed.) Geometry Driven Diffusion in Computer Vision. Lecture Notes in Computer Science. Kluwer, Dordrecht/Boston (1994)
42. Rosenfeld, A., Pfaltz, J.L.: Distance functions on digital pictures. Pattern Recognit. **1**, 33–61 (1968)
43. Serra, J.: Image Analysis and Mathematical Morphology. London Academic, Orlando (1982)
44. Shah, J.: Segmentation by nonlinear diffusion. In: Proceedings of the CVPR, pp. 202–207. IEEE Computer Society, Los Alamitos (1991)
45. Shah, J.: A common framework for curve evolution, segmentation and anisotropic diffusion. In: Proceedings of the CVPR, pp. 136–142 (1996)
46. Shah, J.: Riemannian drums, anisotropic curve evolution and segmentation. In: Proceedings of the Scale-Space, pp. 129–140. Springer, Berlin/New York (1999)
47. Shah, J.: Skeletons and segmentation of shapes. Technical report, Northeastern University. http://www.math.neu.edu/~shah/publications.html (2005)
48. Shah, J., Pien, H.H., Gauch, J.: Recovery of shapes of surfaces with discontinuities by fusion of shading and range data within a variational framework. IEEE Trans. Image Process. **5**(8), 1243–1251 (1996)
49. Siddiqi, K., Tresness, K.J., Kimia, B.: Parts of visual form: ecological and psychophysical aspects. Perception **25**, 399–424 (1996)
50. Stiny, G.: Shape: Talking about Seeing and Doing. MIT, Cambridge (2006)
51. Tari, S., Genctav, M.: From a modified Ambrosio-Tortorelli to a randomized part hierarchy tree. In: Bruckstein, A.M., ter Haar Romeny, B.M., Bronstein, A.M., Bronstein, M.M. (eds.) Scale Space and Variational Methods. Lecture Notes in Computer Science, vol. 6667, pp. 267–278. Springer, Berlin/Heidelberg (2011)
52. Tari, S., Shah, J.: Simultaneous segmentation of images and shapes. In: Proceedings of the SPIE, San Diego, vol. 3168, San Diego pp. 88–94 (1997)

53. Tari, S., Shah, J.: Local symmetries of shapes in arbitrary dimension. In: Proceedings of the ICCV, pp. 1123–1128. Narosa, New Delhi (1998)
54. Tari, S., Shah, J., Pien, H.: A computationally efficient shape analysis via level sets. In: Proceedings of the MMBIA, pp. 234–243. IEEE Computer Society, Los Alamitos (1996)
55. Tari, S., Shah, J., Pien, H.: Extraction of shape skeletons from grayscale images. CVIU **66**(2), 133–146 (1997)
56. Teboul, S., Blanc-Féraud, L., Aubert, G., Barlaud, M.: Variational approach for edge preserving regularization using coupled PDE's. IEEE Trans. Image. Process. **7**, 387–397 (1998)
57. Xu, C., Liu, J., Tang, X.: 2d shape matching by contour flexibility. IEEE Trans. Pattern Anal. Mach. Intell. **31**(1), 180–186 (2009)
58. Zeng, J.T., Lakaemper, R., Wei, X., Li, X.: 2d shape decomposition based on combined skeleton-boundary features. In: Proceedings of the Advances in Visual Computing, 4th International Symposium, ISVC, pp. 682–691. Springer, Berlin/New York (2008)