CHAPTER 2.7
EXTRACTING PARTS OF 2D SHAPES USING LOCAL AND GLOBAL INTERACTIONS SIMULTANEOUSLY

Sibel Tari∗
Middle East Technical University, Department of Computer Engineering,
Ankara, Turkey 06531,
stari@metu.edu.tr

Perception research provides strong evidence in favor of part based representation of shapes in human visual system. Despite considerable differences among different theories in terms of how part boundaries are found, there is substantial agreement on that the process depends on many local and global geometric factors. This poses an important challenge from the computational point of view.

In the first part of the chapter, I present a novel decomposition method by taking both local and global interactions within the shape domain into account. At the top of the partitioning hierarchy, the shape gets split into two parts capturing, respectively, the gross structure and the peripheral structure. The gross structure may be conceived as the least deformable part of the shape which remains stable under visual transformations. The peripheral structure includes limbs, protrusions, and boundary texture. Such a separation is in accord with the behavior of the artists who start with a gross shape and enrich it with details. The method is particularly interesting from the computational point of view as it does not resort to any geometric notions (e.g. curvature, convexity) explicitly. In the second part of the chapter, I relate the new method to PDE based shape representation schemes.

1. Introduction

Perception research provides strong evidence in favor of part based representation of shapes in human visual system. Recent work using single cell recordings in area V4 in the primate visual cortex supports part based coding at intermediate levels. Many influential shape representation theories either explicitly or implicitly assume an organization in terms of constituent components. In Binford, Marr and Nishiara and Biederman, shape is represented via pre-defined simple shapes (primitives) and their spatial layout. Medial axis or local symmetry set (one of the influential ideas in shape representation) is closely connected with the notion of parts. Large number of computational methods utilize medial axis to infer part

∗Work is supported in part by TUBITAK 108E015.
In a recent work, Superpresents a quite successful part based recognition scheme.

There are powerful theories accompanied by computational implementations on what constitutes a good partition without resorting to predefined shapes or categorical units, e.g. Despite considerable differences among different theories in terms of how part boundaries are found, there is substantial agreement on that the process depends on many geometric factors both at the global and the local levels. Indeed, part of the difficulty in devising computational mechanisms for shape decomposition lies in the difficulty in integrating local boundary concavities with non-local shape descriptions. Recent works by Mi and DeCarlo and Zeng et al. address this challenging issue using contour curvature and local symmetry axis simultaneously. In another recent work, Xu, Liu and Tang combine effectiveness of both local and global features for matching shapes. One of the successful recognition schemes, the shape context by Belongie, Malik and Puzicha, is based on quantifying non-local interactions among boundary points. Moreover, some of the successful methods for partitioning images into pixel groups are based on non-local relations.

The rest of the chapter is organized as follows. In the new decomposition method is presented. The new method takes both local and global interactions within the shape domain into account. At the top of the partitioning hierarchy, the shape gets split into two parts capturing, respectively, the gross structure and the peripheral structure. The gross structure may be conceived as the least deformable part of the shape which remains stable under visual transformations. The peripheral structure includes limbs, protrusions, and boundary texture. Such a separation is in accord with the experimental studies suggesting that the global shape is processed before the details and with the behavior of the artists who start with a gross shape and enrich it with details.

In the new method formulated in a discrete setting is related to PDE based shape representation approach with particular emphasis given to a recent skeleton based scheme which I had previously developed with my student Cagri Aslan.

2. The New Method

The basic idea is to create a field within the shape domain with emergent structures capturing the parts automatically. This field is computed by minimizing an energy which captures both local and global; and both region and boundary based interactions among shape points.

Let us define a function $\omega$ defined on a discrete shape domain $\Omega$ as the minimizer of an energy $E(\omega)$. Let this energy be a sum of a region based energy $E_{\text{Reg}}(\omega)$ and a boundary based energy $E_{\text{Bdy}}(\omega)$; and the region based energy $E_{\text{Reg}}(\omega)$ is a sum of two energies which model the global $(G)$ and the local $(L)$ interactions within
the shape domain $\Omega$:

$$E(\omega) = E_{Reg}(\omega) + w_{Bdy}E_{Bdy}(\omega)$$

$$= E_{Reg}^G(\omega) + E_{Reg}^L(\omega) + w_{Bdy}E_{Bdy}(\omega) \quad (1)$$

Assuming that each of the three terms in (1) can be expressed as a sum of energies defined at each pixel $i,j$, the following form is obtained:

$$E(\omega) = \sum_{i,j \in \Omega} E_{Reg}(\omega_{i,j}) + E_{Reg}^L(\omega_{i,j}) + w_{Bdy}E_{Bdy}(\omega_{i,j}) \quad (2)$$

In the absence of any specific purpose or bias, equal importance can be given to both the region and the boundary by setting $w_{Bdy} = 2$. For computational reasons, it is preferable to choose a quadratic form for each energy. Let $E_{Reg}^G(\omega_{i,j})$ be:

$$E_{Reg}^G(\omega_{i,j}) = \frac{1}{|\Omega|} \left( \sum_{k,l \in \Omega} \omega_{k,l} \right)^2 \quad (3)$$

The minimizer $\omega_{i,j}$ of (3) satisfies:

$$\frac{1}{|\Omega|} \sum_{k,l \in \Omega} \omega_{k,l} = 0 \quad (4)$$

The condition satisfied by the minimizer of $E_{Reg}^G(\omega_{i,j})$ is independent of the pixel location $(i,j)$ and it explicitly states that the global average over the shape domain should be zero. It forces $\omega$ to attain both positive and negative values within the shape domain $\Omega$. This behavior when complemented with the behavior induced by the other terms will be shown to be quite instrumental in obtaining robust and parameter free separation of the gross structure and the peripheral structure.

The second term of the region based energy, $E_{Reg}^L$, has the following form:

$$E_{Reg}^L(\omega_{i,j}) = - (\omega_{i+1,j} \cdot \omega_{i-1,j} + \omega_{i,j+1} \cdot \omega_{i,j-1}) \quad (5)$$

Clearly, $E_{Reg}^L(\omega_{i})$ is minimized when the values of the neighboring pixels are similar. Thus the second term of the energy imposes smoothness on $\omega$. The condition for the minimizer of $E_{Reg}^L$ is not straightforward to calculate as that of $E_{Reg}^G$. First, a local continuous approximation at location $i,j$ is considered with the help of Taylor series. Second, the Gateaux derivative is calculated. Third, the Gateaux derivative is discretized and set to zero, to obtain the condition for the minimizer:

$$(-2 + 4) \ast \omega_{i,j} - \omega_{i-1,j} - \omega_{i+1,j} - \omega_{i,j-1} - \omega_{i,j+1} = 0 \quad (6)$$

The boundary based energy $E_{Bdy}(\omega)$ is chosen as a measure of pairwise interaction between two properly chosen boundary points such that the pairs indicate parts. One motivation to consider pairwise interaction between two boundary points comes from the well-established minima rule which is used in many computational procedures for shape decomposition. However, computationally, it is not an
easy task to model pairwise interactions among boundary points. These interactions are neither local nor global. A simple alternative is constructed by exploiting the connection between the concept of pairwise interaction among boundary points and the shape skeleton. With the help of this connection, $E_{Bdy}$ is expressed as an energy defined over the entire shape domain.

The connection can be explained with the help of the grass-fire model by Blum. Assume that one initiates fire fronts at time $t = 0$ along all the points of the shape boundary and lets these fronts propagate toward the center of the shape at a uniform speed. The locus of points where these fronts meet and extinguish defines the shape skeleton. Each skeleton point is formed as a result of interaction between two or more boundary points. During the course of the propagation, the time $t$ may be thought of as a function $t_{i,j}$ defined over the shape domain by setting the value to the time when the propagating fronts passes through the pixel $(i, j)$. The value of $t_{i,j}$ will be proportional to the minimum distance from $(i, j)$ to the nearest boundary points. Skeleton points are the ones which are equidistant from at least two boundary points (Fig. 1.)

This insight enables the expression of $E_{Bdy}(\omega)$ as a quadratic energy defined over the entire shape region $\Omega$ as the following form:

$$E_{Bdy}(\omega_{i,j}) = (\omega_{i,j} - t_{i,j})^2$$  \hspace{1cm} (7)

It is straightforward to calculate the condition for the minimizer of the $E_{Bdy}(\omega_{i,j})$ given in (7) as:

$$(\omega_{i,j} - t_{i,j}) = 0$$  \hspace{1cm} (8)

In the absence of other terms, (8) states that the field $\omega$ should be equal to the distance transform defined over $\Omega$. Putting together all three competing terms, (4,6,8), the first order derivative w.r.t. each unknown $\omega_{i,j}$ takes the following form:

$$\frac{\partial E}{\partial \omega_{i,j}} = \frac{1}{|\Omega|} \left( \sum_{k,l \in \Omega} \omega_{k,l} \right) + (w_{Bdy} - 2 + 4)\omega_{i,j} - w_{Bdy}t_{i,j}$$

$$- \omega_{i-1,j} - \omega_{i+1,j} - \omega_{i,j-1} - \omega_{i,j+1}$$  \hspace{1cm} (9)

Setting the derivative equal to zero yields that the minimizer of $E(\omega)$ satisfies the following condition at each pixel $(i, j)$:
Extracting Parts of 2D Shapes Using Local and Global Interactions Simultaneously

\[ w_{Bdy}t_{i,j} = (w_{Bdy} - 2 + 4)\omega_{i,j} - \omega_{i-1,j} - \omega_{i+1,j} - \omega_{i,j-1} - \omega_{i,j+1} + \frac{1}{|\Omega|} \left( \sum_{k,l \in \Omega} \omega_{k,l} \right) \]  

(10)

Recall that, in the absence of any specific purpose or bias, the weight \( w_{Bdy} \) is set to 2 to give equal importance to both the region and the boundary. Thus, \( \omega_{i,j} \) is computed by solving (11) given below, at all the pixels simultaneously, assuming that the values are zero at the boundary pixels.

\[ t_{i,j} = 4\omega_{i,j} - \omega_{i-1,j} - \omega_{i+1,j} - \omega_{i,j-1} - \omega_{i,j+1} + \frac{1}{|\Omega|} \left( \sum_{k,l \in \Omega} \omega_{k,l} \right) \]  

(11)

The field \( \omega \), computed by solving (11), is depicted in Fig. 2 for two sample shapes. It attains both negative and positive values. This behavior is dictated by the global region energy \( E_{\text{Reg}}^G \) which explicitly states that the global average over the shape domain should be small.

In Fig. 2 (a), the restriction of \( \omega \) where it is negative is displayed. This set is denoted by \( \Omega^- \). It captures the peripheral structure (protrusions, limbs, boundary texture). The darkest blue denotes zero; the darkest red denotes the lowest negative value. The removed inner part is the part on which \( \omega \) is positive; and is denoted by \( \Omega^+ \). This blob-like part captures the gross structure. The gross structure is the least deformable part of the shape which remains stable under a variety of visual changes as demonstrated in Fig. 3 using eight different instances of a hand silhouette.

In Fig. 2 (b), the absolute value of \( \omega \) is displayed on the entire shape domain. For visualization purposes, the negative values and the positive values are normalized, separately, to the \([0, 1]\) interval. The darkest blue denotes zero; the darkest red denotes one. Notice that various local maxima capture intuitive parts such as the body, the head, the tail, and the legs of the horse. These parts can be easily extracted by considering a growth starting from each local maxima. Separate growths from each pair of neighboring maxima meet at a saddle point.

2.1. Experimental Results and Discussion

The method is is discussed via a set of highly illustrative examples. These examples are silhouettes collected from various sources. Some of the original silhouettes are modified by the author to obtain shapes with holes, missing and/or extra parts.

In Figs. (4-7), some decomposition results are provided. These results are obtained by applying Matlab’s watershed command to \( \Omega^- \). (This is equivalent to considering a growth starting from each local minima of \( \omega \).)

In Fig. 4, the first column depicts the given shape. The second column depicts the normalized absolute value of \( \omega \). The third column depicts the parts. Bright
Fig. 2. The field $\omega$ computed by solving (11). For visualization purposes, the values are normalized. (a) The restriction of $\omega$ to areas where its values are negative. This part denotes the peripheral structure i.e. protrusions, limbs, and boundary texture. The removed inner part on which the values are positive is the gross structure. (b) The absolute value of $\omega$.

Fig. 3. The areas where $\omega < 0$ is shown in black. The gross structure (inner white blob) may be conceived as the least deformable part of the shape. It remains stable under a variety of changes.

colors are used for parts on which $\omega$ is negative; and gray is used for the part on which $\omega$ is positive.

The silhouette shown in the first row is a sampled down version of a human silhouette from its original resolution of $414 \times 459$ to $60 \times 60$. The silhouettes on the remaining rows are drawn by the author to introduce holes, missing portions, occlusions. In each case, semantically meaningful parts (the torso, the legs, the
The method is robust with respect to occlusion, missing data, extra objects, and holes. (a) Input silhouettes. (b) Absolute value of $w$. (c) Parts extracted by applying Matlab’s *watershed* command to $w$.

Arms, the head) are captured. Even though the formulation includes a global term, the local changes, no matter how significant they are, do not affect the detected parts.

In Fig. 5 the applicability of the method when the input consists of disconnected sets (multiple objects in a scene) is demonstrated.
Fig. 5. A scene with two silhouettes. The method is applicable to disconnected sets.

Fig. 6. The method is robust with respect to resolution changes. An artificial shape on (a) 220 × 220, (b) 60 × 60 lattices.
In Fig. 6, the robustness of the method with respect to changes in resolution is demonstrated. An artificial shape from Aslan is used in its original resolution in (a), and in a reduced resolution in (b).

In Fig. 7, the decomposition results for a variety of shapes are provided. The decompositions are consistent; similar shapes are partitioned similarly and the captured parts are compatible with our intuition.

In some cases, as in Fig. 8, the decomposition process starting from each and every local minima (i.e. ignoring saliency) may create un-intuitive parts. Normalized absolute value of $\omega$ for three sample turtle shapes are depicted in the first row. In all of the three cases, one can easily spot five local maxima in the peripheral part, and one local maxima in the central part. However, the peripheral structure of the first turtle shape is decomposed into seven pieces. For the second turtle, there are six pieces corresponding to the head, the two legs, the two arms, and the tail. For the third turtle, there are five pieces corresponding to the head, the two legs, the two arms. In the third turtle, due to smoother transition between the two legs, the tail part is missed. This produces inconsistencies among silhouettes from the same category. However, inconsistent parts are not as salient as the consistent parts.

![Sample decompositions. Similar shapes are partitioned similarly; and the parts are compatible with intuition.](image-url)
Fig. 8. Un-intuitive parts. See text.

Fig. 9. Saliency of a part. In each figure, the restriction of $\omega$ to locations, where its value is less than a given threshold, is depicted. The threshold is increased gradually. (a)-(h) $\omega < -4, -9, -14, -19, -24, -29, -34, -37$, respectively.

clarifying illustration is given in Fig. 9 with the help of the first turtle which is decomposed into eight pieces including the torso. A growth process is simulated starting from Fig. 9(h) and ending at Fig. 9(a). In each sub figure, the restriction of $\omega$ to the locations, where its value is less than a given threshold is depicted. The threshold is increased gradually. At the first threshold level in (h), only the head appears. At the third threshold level in (f), the arms appear. The head still continues as an individual blob. At the fourth threshold in (e) level the two legs appear. The five pieces remain separate. At the next threshold level in (d) there
are still five pieces. However, notice that in somewhere between (e) and (d), the tail piece comes to existence and then gets combined with the rightmost leg. It is appropriate to say that the saliency of the tail is quite low compared to the saliency of the other five pieces, due to its short life span as an individual entity. At the next threshold level in (c), the two legs combine to form a single piece corresponding to the lower body; and the two arms and the head combine to form the upper body. The separation of the peripheral structure into upper and lower body divides the elliptical gross structure from the minor axis. Finally in (a), the upper and lower parts combine to form a single peripheral structure.

In all of the previous examples, the gross structure is shown in gray, as a single part. There may be shapes such that the gross structure is composed of multiple parts, i.e. shapes with strong necks. The parts of the butterfly shape in Fig. 10(a) are shown in (b) and (c). In (b), the parts in the peripheral structure are shown in bright colors; the gross structure is in gray. In (c) the parts in the gross structure are shown in bright colors; the peripheral structure is in gray. For reference, the restriction of $\omega$ to the peripheral structure, i.e. where it is negative, is shown in (d). As the neck which connects the two lobes of the butterfly gets thinner, the gross structure may be split into two disjoint sets. (See Fig. 11)

2.2. Comparison to Recent Decomposition Methods by Mi and DeCarlo and Zeng et al.

In two recent papers, Mi and DeCarlo and Zeng et al. present decomposition methods which exploit both the skeleton and the boundary curvature information.

![Fig. 10. A shape whose gross structure is composed of two blobs. (a) A butterfly shape on a 60 x 60 lattice. (b-c) The parts. (d) The restriction of $\omega$ to areas where the values are negative.](image1)

![Fig. 11. When the neck gets thinner, the gross structure may split into two disjoint sets.](image2)
Neither of the methods consider a fully global context. It is worth comparing the three methods using an illustrative example: the prickly pear, which is shown in Fig. 12(a).

The decompositions obtained in Mi and DeCarlo\cite{7} and Zeng et al.\cite{6} using local symmetry axis and contour curvature are shown in (b) and (c), respectively. The new decomposition (using a reduced $80 \times 80$ resolution) is shown in (d) and (e). All of the three methods find two blobs. Similar to the new decomposition, the decomposition by Mi and DeCarlo\cite{7} separates the boundary texture (shown in gray in (b)) from the main structure. On the other hand, the decomposition by Zeng et al.\cite{6} does not separate boundary texture from the main structure leading to the interpretation of the shape as two prickly balls glued together.

As experimental studies on human subjects demonstrate\cite{16}, multiple (and mutually exclusive) parses of a given shape are possible. Thus, in a purely bottom-up process without considering a specific application or a context, one can not decide which partitioning scheme is the best.

I remark that the advantages of the new method are purely from the computational point of view. It does not involve any parameters or thresholds. It can work at very low resolutions as opposed to other methods which involve the computation of curvature or local symmetry axes, since their computation requires certain resolution.

Notice that the restriction of the field $\omega$ to where the values are less than a given threshold $-3$, in (f), indicates the first partitioning of the peripheral structure along the minor axis, similar to the turtle case in Fig. 8. This indication is consistent
with the result of the method of Zeng et al.\textsuperscript{6} which computes the partition line by sequentially eliminating the boundary detail using Discrete Curve Evolution.\textsuperscript{13}

The new method is essentially a parameter free method with the assumption that equal importance should be given to local and global terms. However, the other two methods do not use global features (note that non-local is not necessarily global). Thus, for a comparative evaluation purpose, it is worth trying to reduce the effect of the global term by imagining a constant weight $c < 1$ in front of $E_{Reg}^{G}$ in (2). In Fig. 13 (a-b), $c = 0.5$. One can notice the slight reduction of the peripheral region. The global term is responsible for the balance between the negative values and the positive values of $\omega$. As its importance decreases, more pixels tend to attain positive values. In (c), $c = 0.125$. As $c$ decreases, the peripheral structure shrinks further and the implied decomposition approaches to the one shown in Fig. 12 (c). In Fig. 14, the effect of reducing the importance of the global term is demonstrated using butterfly shapes.

![Fig. 13. The effect of reducing the importance of the global term. (a-b) $c = 0.5$. (c) $c = 0.125$.](image)

![Fig. 14. The effect of reducing the importance of the global term. $c = 0.25$.](image)

In Fig. 15, the importance of the global term is significantly reduced by setting $c = 0.025$. The decomposition result using the new method is shown in (a). For reference, the decomposition results by the previous methods are shown in (b-c).
Fig. 15. The effect of significantly reducing the importance of the global term. (a) The decomposition with the new method, $c = 0.025$. (b-c) The decompositions presented by Mi and DeCarlo\textsuperscript{6} and Zeng et al.\textsuperscript{5} respectively. [(b-c) taken from the original sources\textsuperscript{6,7}]

3. Connection to the methods of Tari, Shah and Pien,\textsuperscript{31,33} Aslan and Tari\textsuperscript{10,30,34} and Gorelick et al.\textsuperscript{29}

Recall that the field $\omega$ is the minimizer of the following energy:

$$E(\omega) = \sum_{i,j \in \Omega} E_{\text{Reg}}^G(\omega_{i,j}) + E_{\text{Reg}}^L(\omega_{i,j}) + w_{\text{Bdy}} E_{\text{Bdy}}(\omega_{i,j})$$

Let us omit the term $E_{\text{Reg}}^G$ that models the global interaction among the shape pixels to obtain a reduced energy:

$$\sum_{i,j \in \Omega} - (\omega_{i+1,j} \cdot \omega_{i-1,j} + \omega_{i,j+1} \cdot \omega_{i,j-1}) + w_{\text{Bdy}} (\omega_{i,j} - t_{i,j})^2 \quad (12)$$

By setting the first derivative of (12) with respect to $\omega_{i,j}$ equal to zero, the condition satisfied by the minimizer is obtained as:

$$(w_{\text{Bdy}} - 2) \omega_{i,j} - w_{\text{Bdy}} t_{i,j} + 4 \omega_{i,j} - \omega_{i-1,j} - \omega_{i+1,j} - \omega_{i,j-1} - \omega_{i,j+1} = 0 \quad (13)$$

Letting $(w_{\text{Bdy}} - 2) = \alpha > 0$ gives

$$(4 \omega_{i,j} - \omega_{i-1,j} - \omega_{i+1,j} - \omega_{i,j-1} - \omega_{i,j+1}) - \alpha \omega_{i,j} = (\alpha + 2) t_{i,j} \quad (14)$$

(14) is clearly the discretization using central difference approximation, of the PDE (15) given below:

$$(\Delta - \alpha) w(x, y) = f(x, y)$$

with $w(x) = 0$ for $x = (x, y) \in \partial \Omega$

where $\Delta$ denotes the Laplace operator, and the right hand side inhomogeneity $f(x, y)$ is a scaled distance transform. (15) is defined on a planar shape domain $\Omega$ which is a connected, bounded, open domain of $\mathbb{R}^2$.

Interestingly, when the right hand side inhomogeneity is replaced with a constant function $f(x, y) = 1$ and $\alpha$ is set to zero (i.e. $w_{\text{Bdy}} = 2$ as in the new method) one obtains the Poisson equation which has been recently proposed by Gorelick et al.\textsuperscript{29} as a shape representation tool. On the other hand, when $\alpha > 0$ and $f(x, y) = -\alpha$, one obtains the PDE:
\[
(\triangle - \alpha) v = -\alpha \\
\text{with } v(x) = 0 \text{ for } x = (x, y) \in \partial \Omega
\]

which has been proposed earlier by Tari, Shah and Pien.\citep{31,33} The qualitative behavior of the \( v \) function for small \( \alpha \) is essentially the same with that of the function obtained by solving the Poisson equation. Both of them are essentially weighted distance transforms\citep{35} where the local steps between neighboring points are given different costs. Tari, Shah and Pien\citep{31} have initially proposed \( v \) function as a linear and a computationally efficient alternative to the curve evolution scheme by Kimia, Tannenbaum and Zucker\citep{36} by showing that the successive level curves of \( v \) mimic the motion of curves with a curvature dependent speed in the direction of the inward normal.\citep{31} Furthermore, they demonstrated that the gradient of \( v \) along a level curve approximates the curvature of level curves, thus, suggesting a robust method for skeleton extraction by locating the extrema of the gradient of \( v \) along the level curves. The skeleton computation method of Tari, Shah and Pien\citep{31,33} exploits the connection among morphology, distance transforms and fronts propagating with curvature dependent speeds. Such connections have stimulated many interesting approaches in solving shape related problems.\citep{35} The importance of Tari, Shah and Pien approach is that it is the first attempt to unify segmentation and local symmetry computation into a single formulation by exploiting the connection between \citep{16} and the Mumford and Shah\citep{37} segmentation functional (via its Ambrosio and Tortorelli\citep{38} approximation). It naturally extends to shapes in arbitrary dimension.\citep{33} (In a related publication,\citep{39} the author introduces an additive normalization term to \citep{16} which forces the solution to oscillate, yielding the same boundary texture and gross structure separation. The proposed method is connected to a variety of morphological ideas as well as to the method of Tari, Shah and Pien in the variational calculus and PDE setting.)

In Fig.\citep{16} the basic method of Tari, Shah and Pien is illustrated using an example by C. Aslan.\citep{10,14} At the top row, the level curves of \( v \) function, mimicking the behavior of fronts propagating with curvature dependent speed, are shown. In each column, a different \( \alpha \) value is used when computing \( v \) via \citep{16}. As \( \alpha \) decreases from left to right, the relative speed of the high curvature points increases. Consequently, the inner level curves lose their concavities and become smoother earlier. Thus the value of \( \alpha \) determines the level of smoothing (or \textit{diffusion}). The arrows indicate the maxima and the saddle points. Topological interpretation of the shape varies with \( \alpha \). When, in the first column, \( \alpha = 1/4^2 \), the evolving shape boundary gets split into three curves, each of which shrinks into three distinct maxima separated by the two saddle points, indicating three parts. In (b) and (c) \( \alpha \) is reduced to \( 1/8^2 \) and \( 1/16^2 \), respectively. The level curves lose their concavities (which imply parts) earlier and evolving curves split into two parts instead of three.

Skeleton branches computed from the respective \( v \) functions (displayed at the
Tari, Shah and Pien method using three different $\alpha$ values. (a) $\alpha = 1/4^2$ (b) $\alpha = 1/8^2$ (c) $\alpha = 1/16^2$. The top row (level curves of $v$) illustrates different topological interpretations by varying $\alpha$. The arrows indicate the maxima and the saddle points. The bottom row displays the skeletons computed from the respective $v$ functions. [Figures by C. Aslan].

Un-intuitive skeleton branches (red color) in Tari, Shah and Pien. (a) $\alpha = 1/8^2$ (b) $\alpha = 1/16^2$. See text for discussion. [Figures by C. Aslan].

(bottom row) typically track the evolution of the indentations and protrusions of the shape. However, some of them exhibit a pathological behavior that frequently occurs in the Tari, Shah and Pien method when a limb is close to a neck. Notice that the skeletons contain branches that do not correspond to any protrusion or
indentation of the shape. Such branches are marked with red color in Fig. 17. Aslan and Tari\textsuperscript{10,30,34} claim that the reason of this pathology is \textit{insufficient diffusion}.

As seen in Fig. 17 increasing the amount of diffusion by decreasing $\alpha$ makes such branches disappear. In (a), the computation stopped while the shape was transforming from a shape with three major blobs to a shape with two major blobs. The circular branch colored with red is due to the interaction of the center of parts two and the neck between parts one and two. As shown in (b), increasing the amount of diffusion by decreasing $\alpha$ makes this branch disappear since the topological change is complete. This time, the shape is between the state with two blobs (parts one and two together and part three) and the state with one blob. The red branch is due to the interaction of the center point of part three and the neck between parts two and three.

Thus, as a remedy, they propose to increase the diffusion by gradually decreasing $\alpha$ so that almost every shape is forcefully interpreted as a single blob ignoring the part structure. Following this ad-hoc modification, the new $v$ function has been successfully applied in shape matching applications\textsuperscript{30,34,40,41}.

However, the method can not be applied to shapes which can not be reduced to a single blob. Such cases include:

- shapes with holes;
- thin and long shapes with constant width;
- shapes with more than one equally prominent parts.

The strategy adopted\textsuperscript{10,30,34} for shapes with two equally prominent parts (dumbbell-like shapes) is to retain their dumbbell-like character. This ad-hoc solution introduces a representational instability as the width of the neck that separates two prominent parts change.

Fig. 18. The level curves of $\omega$. The inner black level curve is the zero-level curve. In all of the three cases, $\Omega^+$ denotes the gross structure.
In Fig. 18-19 the level curves of $w$ are shown for three dumbbell-like shapes with varying neck thickness. In all of the three cases, $\Omega^+$ denotes the gross structure. Instead of relying on a single point to be the shape center as in the method of Aslan and Tari, the new method takes a different attitude. Robustness is obtained by replacing a point estimate for the center with an interval estimate. Parts of the peripheral structure are depicted in Fig. 19.

The method of Aslan and Tari implicitly codes the part structure via disconnected skeleton branches. Fig. 20 demonstrates the possibility of extracting parts from these disconnected branches. In (a) skeleton points detected with the method of Tari, Shah and Pien using the modified function is shown. In (b) the result of pruning and grouping procedure is shown. In (c) final representation (called the disconnected Aslan skeleton to distinguish it from the Tari, Shah and Pien skeleton) is depicted. In (d) parts are extracted, for each disconnected branch.

Fig. 19. Parts of the peripheral structure for dumbbell-like shapes of varying neck thickness.

Fig. 20. (a) Skeleton points detected with the method of Tari, Shah and Pien using the modified function. (b) After pruning and grouping skeleton points. (c) Disconnecting the major skeleton branch. (d) Parts obtained from disconnected branches. [Unpublished result by the author’s former student E. Baseski.]
Extracting Parts of 2D Shapes Using Local and Global Interactions Simultaneously

by fitting a spline that passes through the disconnection point and the nearest indentations. The parts captured by the new method is presented in Fig. 21 for the same cat shape (reduced resolution) for comparison.

References
1. E. Barenholtz and J. Feldman, Visual comparisons within and between object-parts: evidence for a single-part superiority effect., Vision Research. 43, 1655–1666, (2003).
2. A. Pasupathy and C. Connor, Population coding of shape in area V4, Nature Neuroscience. 5(2), 1332–1338, (2002).
3. T. Binford. Visual perception by computer. In IEEE Conference on Systems and Control, (1971).
4. D. Marr and H. K. Nishiara. Representation and recognition of spatial organization of three dimensional shapes. In Proceedings of the Royal Society of London, Series B, Biological Sciences, 200, pp. 269–294, (1978).
5. I. Biederman, Recognition-by-components: A theory of human image understanding., Psychological Review. 94(2), 115–117, (1987).
6. J. T. Zeng, R. Lakaemper, X. Wei, and X. Li. 2d shape decomposition based on combined skeleton-boundary features. In Advances in Visual Computing, 4th International Symposium, ISVC, pp. 682–691, (2008).
7. X. Mi and D. DeCarlo. Separating parts from 2d shapes using relatability. In ICCV, pp. 1–8, (2007).
8. S. Svensson and G. S. di Baja, Using distance transforms to decompose 3d discrete objects, Image Vision Comput. 20(8), 529–540, (2002).
9. C. Arcelli and L. Serino, From discs to parts of visual form, Image Vision Comput. 15(1), 1–10, (1997).
10. C. Aslan and S. Tari. An axis-based representation for recognition. In ICCV, pp. 1339–1346, (2005).
11. H. Rom and G. Medioni, Hierarchical decomposition and axial shape description, IEEE Trans. Pattern Anal. Mach. Intell. 15(10), 973–981, (1993).
12. B. J. Super, Knowledge-based part correspondence, Pattern Recognition. 40(10), 2818–2825, (2007).
13. E. H. Cohen and M. Singh, Geometric determinants of shape segmentation: Tests using segment identification., Vision Research. 47, 2825–2840, (2007).
14. D. Hofmann and W. Richards, Parts of recognition, Cognition. 18, 65–96, (1984).
15. L. J. Latecki and R. Lakaemper, Convexity rule for shape decomposition based on discrete contour evolution, Comput. Vis. Image Underst. 73(3), 441–454, (1999).
16. L. W. Remminger. Parts, Objects and Scenes: Psychophysics and Computational Models. PhD thesis, UC Berkeley, USA, (2003).
17. K. Siddiqi and B. Kimia, Parts of visual form: Computational aspects, *IEEE Trans. Pattern Anal. Mach. Intell.* 17, 239–251, (1995).
18. K. Siddiqi, K. J. Tresness, and B. Kimia, Parts of visual form: Ecological and psychophysical aspects, *Perception* 25, 399–424, (1996).
19. J. de Winter and J. Wagemans, Segmentation of object outlines into parts: A large-scale integrative study., *Cognition* 99, 275–325, (1999).
20. C. A. Burbeck and S. M. Pizer, Object representation by cores: Identifying and representing primitive spatial regions., *Vision Research* 35, 1917–1930, (1995).
21. D. Navon, Forest before trees: The precedence of global features in visual perception., *Cognitive Psychology* 9, 355–383, (1977).
22. C. Xu, J. Liu, and X. Tang, 2d shape matching by contour flexibility, *IEEE Trans. Pattern Anal. Mach. Intell.* 31(1), 180–186, (2009).
23. S. Belongie, J. Malik, and J. Puzicha. Shape context: A new descriptor for shape matching and object recognition. In *NIPS*, pp. 831–837, (2000).
24. J. Shi and J. Malik. Normalized cuts and image segmentation. In *CVPR*, pp. 731–737, (1997).
25. A. Buades, B. Coll, and J. M. Morel. A non-local algorithm for image denoising. In *CVPR*, pp. 60–65, (2005).
26. J. J. Koenderink and A. J. van Doorn, The shape of smooth objects and the way contours end., *Perception* 11, 129–137, (1982).
27. H. Blum, Biological shape and visual science., *Journal of Theoretical Biology* 38, 205–287, (1973).
28. [http://www.lems.brown.edu/vision/](http://www.lems.brown.edu/vision/)
29. L. Gorelick, M. Galun, E. Sharon, R. Basri, and A. Brandt, Shape representation and classification using the poisson equation, *IEEE Trans. Pattern Anal. Mach. Intell.* 28(12), 1991–2005, (2006).
30. C. Aslan. Disconnected skeletons for shape recognition. Master’s thesis, Department of Computer Engineering, Middle East Technical University (May, 2005).
31. S. Tari, J. Shah, and H. Pien. A computationally efficient shape analysis via level sets. In *MMBIA ’96: Proceedings of the 1996 Workshop on Mathematical Methods in Biomedical Image Analysis*, pp. 234–243, (1996).
32. S. Tari, J. Shah, and H. Pien, Extraction of shape skeletons from grayscale images., *CVIU* 66(2), 133–146, (1997).
33. S. Tari and J. Shah. Local symmetries of shapes in arbitrary dimension. In *ICCV*, pp. 1123–1128, (1998).
34. C. Aslan, A. Erdem, E. Erdem, and S. Tari, Disconnected skeleton: Shape at its absolute scale, *IEEE Trans. Pattern Anal. Mach. Intell.* 30(12), 2188–2203, (2008).
35. P. Maragos and M. A. Butt, Curve evolution, differential morphology and distance transforms as applied to multiscale and eikonal problems., *Fundamentae Informatica.* 41, 91–129, (2000).
36. B. B. Kimia, A. R. Tannenbaum, and S. W. Zucker, Shapes, shocks and deformations, *Int. J. Comput. Vision* 15(3), 189–224, (1995).
37. D. Mumford and J. Shah, Optimal approximations by piecewise smooth functions and associated variational problems., *Commun. Pure Appl. Math.* 42(5), 577–685, (1989).
38. L. Ambrosio and V. Tortorelli, On the approximation of functionals depending on jumps by elliptic functionals via Γ-convergence., *Commun. Pure Appl. Math.* 43(8), 999–1036, (1990).
39. S. Tari. Hierarchical shape decomposition via level sets. In *International Symposium on Mathematical Morphology*, pages 217–228, (2009).
40. E. Baseski, A. Erdem, and S. Tari, Dissimilarity between two skeletal trees in a con-
text, Pattern Recogn. 42(3), (2009).
41. A. Erdem and S. Tari. Coarse-to-fine matching of shapes using disconnected skeletons by learning class-specific boundary deformations. In Workshop on Graph-based Representations in Pattern Recognition, pp. 21–30, (2009).