A Shock Grammar For Recognition

Kaleem Siddiqi
Department of Electrical Engineering
McGill University
Montreal, Canada H3A 1Y2

Benjamin B. Kimia
Division of Engineering
Brown University
Providence, RI 02912

Abstract

We confront the theoretical and practical difficulties of computing a representation for two-dimensional shape, based on shocks or singularities that arise as the shape's boundary is deformed. First, we develop subpixel local detectors for finding and classifying shocks. Second, we show that shock patterns are not arbitrary but obey the rules of a grammar, and in addition satisfy specific topological and geometric constraints. Shock hypotheses that violate the grammar or are topologically or geometrically invalid are pruned to enforce global consistency. Survivors are organized into a hierarchical graph of shock groups computed in the reaction-diffusion space, where diffusion plays a role of regularization to determine the significance of each shock group. The shock groups can be functionally related to the object's parts, protrusions and bends, and the representation is suited to recognition: several examples illustrate its stability with rotations, scale changes, occlusion and movement of parts, even at very low resolutions.

1 Introduction

What does it mean to recognize an object from its shape? Informally, this implies an identification of the shape with a familiar category or class of objects. Figure 1. This notion of categorization is crucial to many vision tasks, such as searching a database of shapes rapidly, reasoning about the attributes of new or unfamiliar shapes, etc. Curiously, whereas this ability to categorize appears to come naturally and effortlessly to humans, it has been extremely difficult to formalize for computers. In this paper, we address the computational aspects of this problem; specifically, we investigate the description of generic shape classes from the mathematical perspective of curve evolution.

Figure 1: These birds are effortlessly grouped into two categories, based on similarity in "form".

Existing proposals for shape representation emphasize properties of its region, e.g., symmetry and thickness [1], or of its boundary, e.g., curvature extrema [20] and inflection points, or of both [2]. An alternate classification is according to those where shape is viewed statically as a combination of primitives, e.g., generalized cylinders, versus those where shape is explained developmentally via a set of processes acting on a simpler shape [14]. Returning to the region-based symmetric axis transform (SAT) [1], this view has spawned a vast literature on the theoretical and computational aspects of skeletons. However, it is unfortunate that Blum's key insight that the SAT provides for qualitative shape descriptions in terms of "shape morphemes", e.g., disc, worm, wedge, flare, etc., is usually forgotten. Curiously, an evolutionary approach to shape description supports and complements this view, and gives it a sound mathematical foundation [8, 10]. To elaborate, Kimia et al. explore deformations of the shape's boundary, a special case of which is deformation by a linear function of curvature $\kappa$:

$$\frac{\partial C}{\partial t} = (\beta_0 - \beta_1 \kappa) \vec{N}$$

Here $C$ is the boundary vector of coordinates, $\vec{N}$ is the outward normal, $s$ is the path parameter, $t$ is the time duration (magnitude) of the deformation, and $\beta_0, \beta_1$ are constants. The space of all such deformations is spanned by the ratio $\beta_0/\beta_1$ and time $t$, constituting the two axes of the reaction-diffusion space. Underlying the representation of shape in this space are a set of shocks [11], or entropy-satisfying singularities, which develop during the evolution and are classified into four types, Figure 2 (left): 1) A FIRST-ORDER SHOCK is a discontinuity in orientation of the shape's boundary; 2) A SECOND-ORDER SHOCK is formed when two distinct non-neighboring boundary points collide, but none of their intermediate neighbors collapse together; 3) A THIRD-ORDER SHOCK is formed when two distinct non-neighboring boundary points collide, such that the neighboring boundary points also collapse together\(^1\); and 4) A FOURTH-ORDER SHOCK is formed when a closed boundary collapses onto a single

---

\(^1\) Whereas third-order shocks are not generic they merit a distinct classification because of their psychophysical relevance [9] and the abundance of biological and man-made objects with "bend-
shocks occur at "necks", third-order shock with infinite speed, but in opposite directions. They are like "components, e.g., fingers, limbs, legs of a table, etc. Also, they are simultaneously the limit of first-order shocks travelling tend to apply directly to images, Section 6.

we suggest how the shock-based framework might be ex-

tation shape classes' and algorithms for obtaining them, Sec-

sidering aspect of shape, leading to generic perceptual qual-


tistical representation via "cores", or regions of high medialness in intensity images [2]. Our work extends the above approaches in a number of ways, which are perhaps best understood in the context of the distinction between shocks and skeletons.

The set of shocks which form along the reaction axis reduces to the traditional skeleton when information regarding type, group, and salience is discarded [23]. However, first, the notion of type is essential to capture qualitative aspects of shape, leading to generic perceptual shape classes and algorithms for obtaining them, Section 2. Second, the grouping of shocks depends not only on their type but also on sequential, geometric and topological constraints obtained from a history of shocks, Section 3. This results in a hierarchical representation of shape by shock groups, as illustrated by numerous examples, Section 4. Third, the notion of salience connects "nearby" shapes, e.g., Figure 19, providing a foundation for a topology over shape for recognition. In conclusion, we suggest how the shock-based framework might be extended to apply directly to images, Section 6.

like" components, e.g., fingers, limbs, legs of a table, etc. Also, they are simultaneously the limit of first-order shocks travelling with infinite speed, but in opposite directions.

First-order shock groups describe "protrusions" second-order shocks occur at "necks", third-order shock groups describe "bends", and viewing the evolution in reverse, fourth-order shocks are seeds from which the shape is grown [9].

2 Shock Classification and Detection: Local Operators

In the design of shock detection operators we face two primary challenges: that of arriving at a complete shock classification scheme which leads to a computational algorithm for detection, and that of obtaining accurate geometric estimates without blurring across singularities. We discuss shock classification and detection in turn.

2.1 Classification of Shocks

An intuitive approach is to classify a shock based on properties of the boundary points which collide at it, Figure 3. Whereas this classification provides insight it is difficult to implement directly, e.g., the mapping of a shock to its associated bi-tangent points can become intractable in the presence of multiple nearby topological splits. Alternatively, one may rely on the differential properties of an embedding surface, an approach which proves to be computationally efficient and robust. For theoretical as well as numerical reasons, the original curve flow is embedded in the level set evolution of an evolving surface [3, 17], $z = \phi(x, y, t)$:

$$\frac{\partial \phi}{\partial t} + \beta(\kappa) |\nabla \phi| = 0,$$

with the correspondence that the evolving shape is represented at all times by its zero level set $\phi(x, y, t) = 0$. For convenience we take the initial surface $\phi_0$ to be the signed distance function to the shape's boundary (although any Lipschitz continuous function will suffice [3]). The classification of shocks based on differential properties of $\phi$ is summarized in Figure 4 and Table 1. A first-order shock corresponds to a discontinuity in the orientation of the tangent $T$ to the level curve, computed from $\phi$ as $\arctan(\frac{-\partial \phi}{\partial y})$. Since the colliding boundary points have normals pointing in opposite directions, $|\nabla \phi| = 0$ at second-, third- and fourth-order shocks. These shocks can be distinguished from one another by the Gaussian curvature, Table 1. Note that this classification is invariant to the choice of the embedding surface and that all

![Figure 2: LEFT: The four shock types. RIGHT: The sides of the shape triangle represent continua of shapes; the extremes correspond to the "parts", "bends" and "protrusions" nodes [9].](image1)

![Figure 3: A classification of shock types based on the tangents and the local neighborhood of the two shock generating boundary points. The curvature disparity is the sum of the two (signed) curvatures.](image2)
Figure 4: Shock classification based on properties of an embedding surface. **TOP LEFT**: First-order shocks occur at corners, corresponding to creases on the surface with $|\nabla \phi| > 0$. **TOP RIGHT**: A second-order shock corresponds to a hyperbolic point with $|\nabla \phi| = 0$. **BOTTOM LEFT**: Third-order shocks correspond to parabolic points with $|\nabla \phi| = 0$. **BOTTOM RIGHT**: A fourth-order shock corresponds to an elliptic point with $|\nabla \phi| = 0$.

<table>
<thead>
<tr>
<th>Shock Type</th>
<th>Orientation</th>
<th>Curvature</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>non-vanishing $\nabla \phi$</td>
<td>high $\kappa$</td>
</tr>
<tr>
<td>Second</td>
<td>isolated vanishing $\nabla \phi$</td>
<td>$\kappa_1 \kappa_2 &lt; 0$</td>
</tr>
<tr>
<td>Third</td>
<td>non-isolated vanishing $\nabla \phi$</td>
<td>$\kappa_1 \kappa_2 = 0$</td>
</tr>
<tr>
<td>Fourth</td>
<td>isolated vanishing $\nabla \phi$</td>
<td>$\kappa_1 \kappa_2 &gt; 0$</td>
</tr>
</tbody>
</table>

Table 1: Shock classification based on the gradient $|\nabla \phi|$, the level set curvature $\kappa$, and the principal curvatures $\kappa_1, \kappa_2$ of the surface.

the necessary quantities can be computed locally\(^3\).

### 2.2 Subpixel Shock Detection

We develop a subpixel implementation of the above ideas in order to obtain accurate geometric estimates in the vicinity of discontinuities and to localize shocks. Note that whereas the level set formulation supports subpixel curve evolution an algorithm that only attempts to locate shocks at grid points will suffer from discretization artifacts.

A class of techniques called *essentially non-oscillatory* (ENO) schemes have recently been introduced in the numerical analysis literature to address the problem of inaccurate differential estimates in the vicinity of discontinuities \([6]\). The basic idea is to select between two contiguous sets of data points for interpolation the one which gives the lower variation, such that at regions neighboring a discontinuity the smoothing is always from the side not containing it. By replacing polynomials with geometric interpolants: lines, circular arcs, etc., these ideas have been adapted to the 2D problem of locating level curves of an embedding surface while preserving and explicitly placing orientation discontinuities (first-order shocks) \([24]\). The method provides a subpixel contour tracer (for open and closed curves) which can be used to recover the shape's contour from the evolving embed-

\[ |\nabla \phi| = (\phi_x^2 + \phi_y^2)^{1/2}; \kappa_1 \kappa_2 = \frac{\phi_x \phi_y^{2} - \phi_y^{2}}{(1 + \phi_x^2 + \phi_y^2)^{3/2}}; \]

\[ \kappa_1 + \kappa_2 = \frac{[1 + \phi_x^2] \phi_x^2 - 2 \phi_x \phi_y \phi_y + [1 + \phi_y^2] \phi_y^2}{(1 + \phi_x^2 + \phi_y^2)^{3/2}}. \]

Figure 5: CLOCKWISE FROM TOP LEFT: The geometric ENO interpolation technique \([24]\) preserves discontinuities in the vicinity of first-, second-, third-, and fourth-order shocks; gridlines are overlayed and detected corners are marked.

Figure 6: LEFT: The zero crossing contours of $|\nabla \phi| - \epsilon$ demarcate regions around the putative shock points. RIGHT: Zero-crossing curves of $\phi_x$ and $\phi_y$ intersect at exactly three points, two of which are fourth-order shocks, and one of which is a second-order shock, as determined from the sign of $\kappa_1 \kappa_2$.

3 Shock Grouping: Global Interactions

The fact that the set of shocks formed under pure reaction ($\beta_1 = 0$) provides the SAT \([23]\) implies that geometric and topological properties that hold for skeletons, e.g., those studied in \([4, 22]\), must hold for shocks as well. We examine three types of constraints on shock formation in Figure 8: sequential, geometric and topolog-

\(^3\)Care must be taken to avoid regions where either $\phi_x$ or $\phi_y$ is identically zero over a neighborhood of grid points. Fortunately, $\phi_x$ and $\phi_y$ cannot both be identically zero over the same regions, since that would imply a 2D region of third-order shocks, which is an impossibility.
The symbols $S_1, S_2, S_4$ represent first-, second-, and fourth-order shocks. $S_f$ is a start symbol, $S_r$ is a terminal, and since third-order shocks never appear in isolation, a group of third-order shocks is an element of the alphabet, denoted by $S_3$. $E$ represents the end of a growing shock sequence and is used to enforce the requirement that shocks be added only to that end, making the grammar context dependent. Figure 9 illustrates the application of the grammar. Note that whereas the grammar suffices to describe the composition of a shock group, it does not reflect the geometric and topological constraints; this may be possible by embedding the grammar in a graph.

4 Examples

We illustrate the robustness of our two-stage numerical algorithm for shock detection and classification with several examples. The reconstructions are simulations of the "growth" of each shape from its shock-based representation, with linear interpolation of the radius function between successive shocks on the same branch.\footnote{The initial distance transform is blurred very slightly to combat discretization effects, hence the reconstructions have slightly rounded corners.} Figure 11 depicts the evolution of shocks for a
metric measurements, the examples the shock branches are smooth and the representation allows for precise reconstruction and accurate not affected and a qualitative description as a collection throughout. Finally, Figure

Figure 12: The shock branches remain smooth and no spurious branches are added under rotation or stretching. Further, the structural description of each triangle as “three protrusions converging onto a single seed” and of each rectangle as a “bend with two protrusions at each end” is preserved.

5 Structural Diffusion

A variety of approaches have been proposed to deal with the sensitivity of the SAT to boundary details, e.g., blurring to create a multiresolution SAT [18], the use of residual functions [16], and non-linear diffusion of the shape’s angle function [18]. Following the theoretical development of [8], the approach we suggest is to use curvature deformation (β) as a smoothing process to assign a significance to each shock group:

Remark 1 (Significance) The significance of a shock group is proportional to its survival with increasing amounts of curvature deformation.

6This choice enforces a number of desirable properties, e.g., in the case of β1/β2 → ∞, any embedded curve will evolve to a round point without developing self-intersections or singularities [5], and the number of extrema and inflection points is non-increasing, implying that no new shock branches can form.
Figure 13: The shock-based description and growth of a shape composed of trapezoids (TOP), and of an industrial shape (BOTTOM). The originals shapes are on the left, and the reconstructions on the right.

Figure 14: Shock detection under occlusion, and movement/bending of parts. LEFT: The original shapes. MIDDLE: The shock-based description. RIGHT: The reconstruction from shocks.

Figure 15: LEFT: The shock-based description of two handwritten letters. RIGHT: First-order shock speed and acceleration. The shock occurs at point B, and after one time step has moved to point C. With $AB = \kappa^{-1} \sin(\theta/2)$, the speed of the shock is obtained as: $s = AB = \beta_0/\sin(\theta/2)$. The acceleration is obtained by differentiating the speed as: $a = s(\beta_0^2 - s^2) \kappa/\beta_0$.

Figure 16: LEFT TO RIGHT: $\beta_0 = -0.2, 0.0, 0.25, 0.5$. Each column depicts the shock groups that have been detected up until the present time, with the evolved shape overlaid. Observe that branches are annihilated in order of the scale of the protrusion they represent.

We consider the effect of diffusion on each shock type; the detection of shocks with diffusion is coarse (not sub-pixel), and is only intended to provide a measure of significance for shocks obtained under pure reaction. When $\beta_0 \neq 0$ we interpret a first-order shock as a maxima of (sufficiently high) positive curvature. The survival of a first-order shock group with increasing diffusion reflects the "scale" of the corresponding protrusion, Figure 16.

The above notion of significance induces a hierarchical ordering of shock branches from fine to coarse, i.e., branches obtained under pure reaction are removed in the order that they annihilate under diffusion, and the structures that they represent are literally broken off, Figure 16. This brings out the coarse level similarity between shapes belonging to the same category, Figure 19, an essential requirement for recognition.

6 Shocks from Images

In conclusion, we suggest that the shock-based representation can be extended to apply to fragmented shapes as they typically arise in real imagery by allowing local edge hypotheses to interact via the evolution of a local embedding surface; recall that any Lipshitz continuous surface can be used. Such a surface can be constructed using the output of an edge operator, i.e., by first placing oriented receptive fields at each edge, Figure 20 (top), in analogy to the lifetime of a grey-level blob in scale space [15], when two protrusions are nearby the shock branches may merge.
Figure 17: LEFT TO RIGHT: $\beta_0 = -0.2$, $\beta_1 = 0.0, 1.0, 1.5$. Each column depicts shocks that have been detected up until the present time, with the evolved shape overlayed; for $\beta_1 \neq 0$ we focus on the higher-order shocks: of the two necks, the weaker one on the left is the first to annihilate with increased diffusion.

Figure 18: The significance hierarchy induced by the computation in Figure 16; in the reconstructions protrusion branches have been removed in the order that they are annihilated with increased diffusion.

Figure 19: COLUMN ONE: The original pear shapes, taken from [20]. COLUMN TWO: The shock-based description under pure reaction. COLUMN THREE: The reconstruction based on the pure reaction description (column two). COLUMN FOUR: Those branches of the pure reaction description (column two) that survive under diffusion. COLUMN FIVE: The reconstruction based on the description in the fourth column brings out the coarse level similarity between the shapes.

and then taking the union of all such receptive fields, Figure 20 (bottom left). By construction, the covering surface has the property that its zero-crossings pass through the original edge locations. Therefore, the evolution of the covering surface can allow for the detection, classification and grouping of shocks prior to obtaining a segmentation of the shape itself, Figure 20 (bottom right).

Acknowledgements
The support of NSERC, and NSF grant IRI-9305630, is gratefully acknowledged. We thank Jonas August, Michael Kelly, Eric Pauwels, and Steve Zucker for helpful discussions.

References


