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.

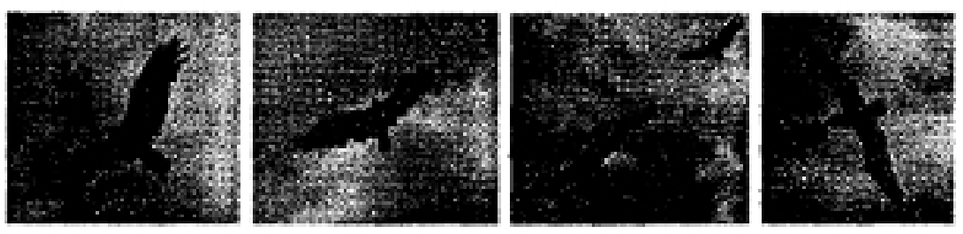

Figure 1: These birds are effortlessly grouped into two categories, based on similarity in "form".

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 κ:

$$\boxed{\begin{cases} \frac{\partial C}{\partial t} & = (\beta_0 - \beta_1 \kappa)\vec{N} \\ C(s,0) & = C_0(s). \end{cases}} \qquad (1)$$

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 β_0, β_1 are constants. The space of all such deformations is spanned by the ratio β_0/β_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 immediate 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-

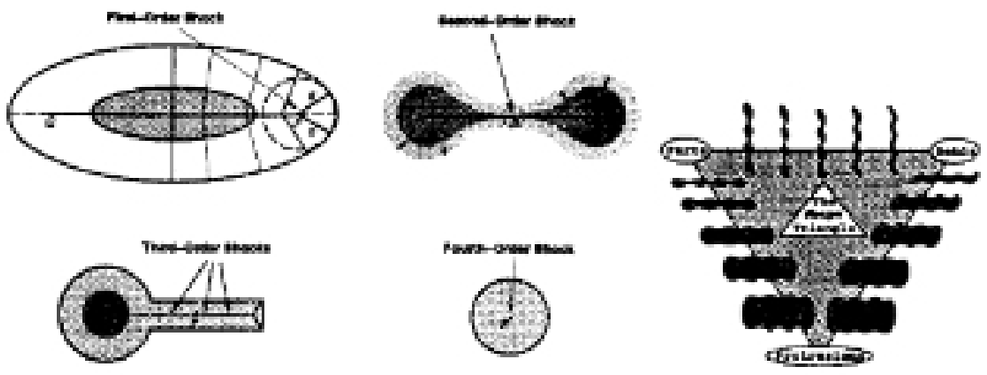

Figure 2: LEFT: The four shock types. RIGHT: The sides of the shape triangle represent continua of shapes; the extremes correspond to the "parts", "bends" and "protrusions" nodes [9].

point. While these definitions are intuitive, they do not easily lend themselves to algorithms for shock detection. A key idea of this paper is that shock computations can be made robust by relying not only on better (subpixel) *local* detectors and classifiers, but also on *global* interactions between shocks, through a shock grammar. In related work, Leymarie and Levine have simulated the grassfire transform using active contours [13]; Scott *et al.* have suggested the use of wave propagation to obtain the full symmetry set [21]; Kelly and Levine have demonstrated the use of annular operators in obtaining coarse object descriptions from real imagery [7]; and Pizer *et al.* have proposed a computational model for object 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[2] 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.

[2]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].

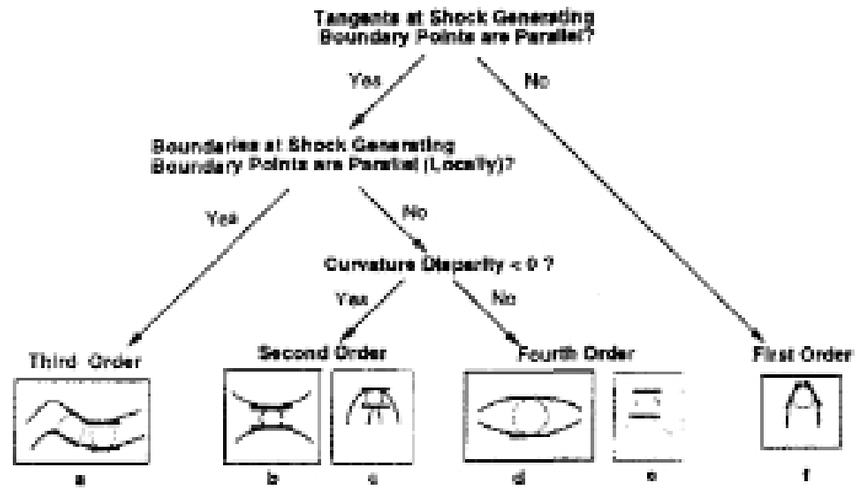

Figure 3: A classification of shock types based on the tangents and the local neighborhood of the two shock generating boundary points. The curvature disparity is the sum of the two (signed) curvatures.

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], $z = \phi(x, y, t)$:

$$\phi_t + \beta(\kappa)|\nabla\phi| = 0, \tag{2}$$

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 ϕ_0 to be the signed distance function to the shape's boundary (although any Lipshitz continuous function will suffice [3]). The classification of shocks based on differential properties of ϕ is summarized in Figure 4 and Table 1. A **first-order** shock corresponds to a discontinuity in the orientation of the tangent $\vec{T}$ to the level curve, computed from ϕ as $\arctan(\frac{-\partial\phi/\partial x}{\partial\phi/\partial y})$. 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

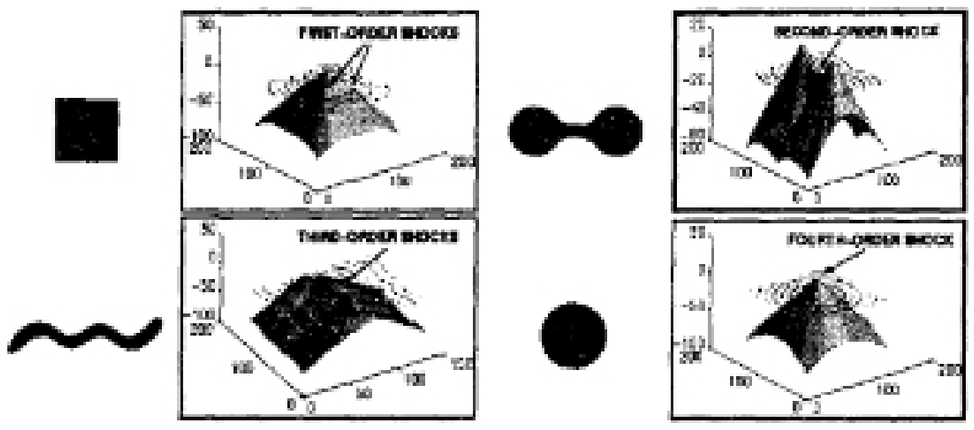

Figure 4: Shock classification based on properties of an embedding surface. TOP LEFT: First-order shocks occur at corners, corresponding to creases on the surface with $|\nabla\phi| > 0$. TOP RIGHT: A second-order shock corresponds to a hyperbolic point with $|\nabla\phi| = 0$. BOTTOM LEFT: Third-order shocks correspond to parabolic points with $|\nabla\phi| = 0$. BOTTOM RIGHT: A fourth-order shock corresponds to an elliptic point with $|\nabla\phi| = 0$.

Shock Type	Orientation	Curvature
First	non-vanishing $\nabla\phi$	high κ
Second	isolated vanishing $\nabla\phi$	$\kappa_1\kappa_2 < 0$
Third	non-isolated vanishing $\nabla\phi$	$\kappa_1\kappa_2 = 0$
Fourth	isolated vanishing $\nabla\phi$	$\kappa_1\kappa_2 > 0$

Table 1: Shock classification based on the gradient $|\nabla\phi|$, the level set curvature κ, and the principal curvatures κ_1, κ_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 embedding surface, Figure 5, and can be extended to higher order shock detection as follows. Recall that $|\nabla\phi| = 0$ at higher order shocks. Therefore, the geometric interpolation method may be used to find ϵ crossings of $|\nabla\phi|$, Figure 6 (left). However, this approximation always yields 2D *regions* surrounding the putative shock points. As a solution, since *ϕ_x and ϕ_y must each go to zero independently* for $|\nabla\phi|$ to go to zero, the problem can be reduced to two 1D problems by considering zero-crossing curves of ϕ_x and ϕ_y[4], and finding overlaps, Figure 6 (right). This suggests the algorithm for higher order shock detection outlined in Figure 7; further details appear in [23].

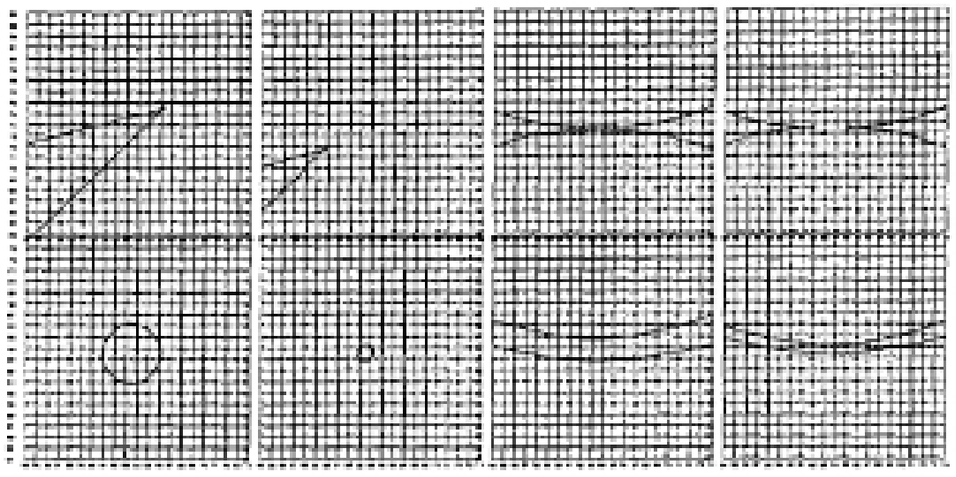

Figure 5: CLOCKWISE FROM TOP LEFT: The *geometric* ENO interpolation technique [24] preserves discontinuities in the vicinity of first- , second-, third-, and fourth-order shocks; gridlines are overlayed and detected corners are marked.

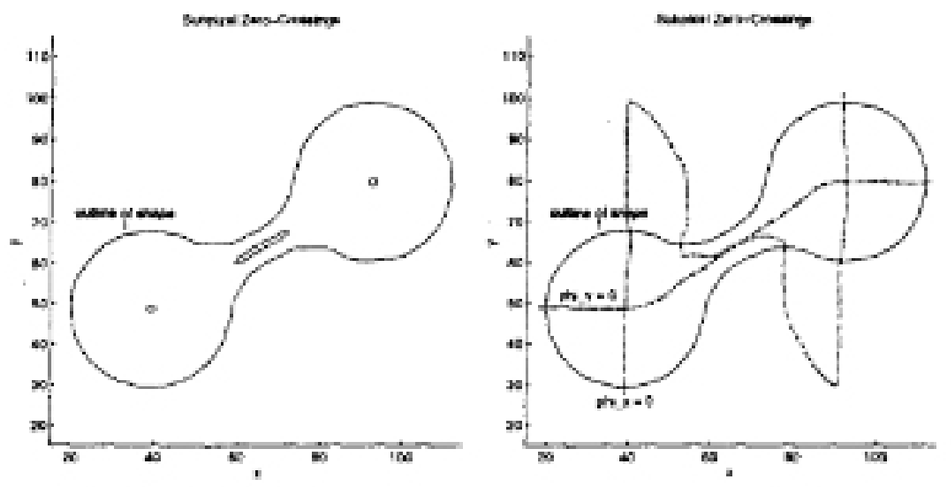

Figure 6: LEFT: The zero crossing contours of $(|\nabla\phi| - \epsilon)$ demarcate regions around the putative shock points. RIGHT: Zero-crossing curves of ϕ_x and ϕ_y intersect at *exactly* three points, two of which are fourth-order shocks, and one of which is a second-order shock, as determined from 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] $|\nabla\phi| = (\phi_x^2 + \phi_y^2)^{1/2}$; $\kappa_1\kappa_2 = \frac{\phi_{xx}\phi_{yy} - \phi_{xy}^2}{(1+\phi_x^2+\phi_y^2)^2}$;
$\kappa_1 + \kappa_2 = \frac{(1+\phi_x^2)\phi_{yy} - 2\phi_x\phi_y\phi_{xy} + (1+\phi_y^2)\phi_{xx}}{(1+\phi_x^2+\phi_y^2)^{3/2}}$.

[4] Care must be taken to avoid regions where either ϕ_x or ϕ_y is identically zero over a neighborhood of grid points. Fortunately, ϕ_x and ϕ_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.