Shape, Shocks and Wiggles [1]

Kaleem Siddiqi
Dept. of Computer Science
Yale University, P.O. Box 208285
New Haven, CT 06520-8285
siddiqi-kaleem@cs.yale.edu
fax: (203) 432 0593

Benjamin B. Kimia
Division of Engineering
Brown University, Box D
Providence, RI 02912
kimia@lems.brown.edu
fax: (401) 863 1157

Allen Tannenbaum
Dept. of Electrical Engineering
University of Minnesota
Minneapolis, MN 55455
tannenba@ee.umn.edu
fax: (612) 625 4583

Steven W. Zucker
Dept. of Computer Science
Yale University, P.O. Box 208285
New Haven, CT 06520-8285
steven-zucker@cs.yale.edu
fax: (203) 432 0593

Abstract We earlier introduced an approach to categorical shape description based on the singularities (shocks) of curve evolution equations. The approach relates to many techniques in computer vision, such as Blum's grassfire transform, but since the motivation was abstract, it is not clear that it should also relate to human perception. We now report that this shock-based computational model can account for recent psychophysical data collected by Christina Burbeck and Steve Pizer. In these experiments subjects were asked to estimate the local centers of stimuli consisting of rectangles with "wiggles" (sides modulated by sinusoids). Since the experiments were motivated by their "core" model, in which the scale at which boundary detail is represented is proportional to object width, we conclude that such properties are also implicit in shock-based shape descriptions.

Keywords: *shape, shocks, cores, wiggles, medial axis.*

Topics: *shape perception, shape description, shape representation.*

[1] We would be willing to prepare an extended version of this paper for the special issue of Image and Vision Computing.

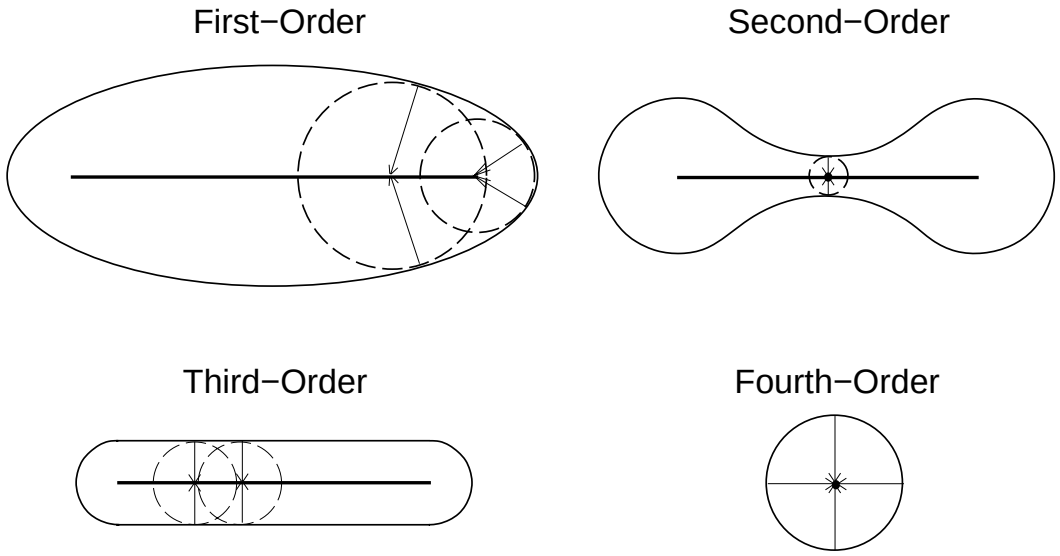


Figure 1: This figure illustrates the four types of shocks that arise in the reaction-diffusion space [6].

1 Shocks and the Shape Triangle

Observing that shapes that are slight deformations of one another appear similar, Kimia *et al.* propose the following evolution equation for visual shape analysis [5, 6]:

$$\boxed{\begin{cases} \mathcal{C}_t & = (\beta_0 - \beta_1 \kappa)\mathcal{N} \\ \mathcal{C}(s,0) & = \mathcal{C}_0(s). \end{cases}} \tag{1}$$

Here $\mathcal{C}$ is the vector of curve coordinates, $\mathcal{N}$ 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* [8], or entropy-satisfying singularities, which develop during the evolution and are classified into four types (Figure 1): 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

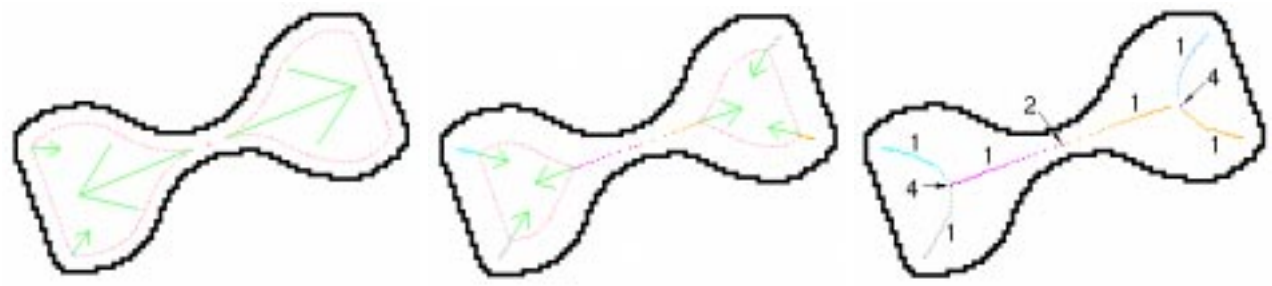


Figure 2: The detection of shocks for a dumbbell shape undergoing constant inward motion, taken from [10]. Each sub-figure is a snapshot of the evolution in time, with the outline of the original shape shown in black, the evolved curve overlayed within, and the arrows representing velocity vectors for the current first-order shocks.

neighboring boundary points also collapse together[2]; and 4) A FOURTH-ORDER SHOCK is formed when a closed boundary collapses onto a single point. To get an intuitive feel for the shock types consider the numerical simulation of a dumbbell shape evolving under constant inward motion (analogous to Blum's grassfire), Figure 2. Note the emergence of a qualitative description of the shape as that of two parts separated at a "neck" (second-order shock), with each part consisting of three "protrusions" (first-order shock groups) merging onto a "seed" (fourth-order shock).

Such generic perceptual shape classes have lead to the metaphor of a shape triangle, where three distinct processes compete with one another to explain shape [7], Figure 3 (left). The sides of the triangle reflect this competition and capture the tension between object composition (parts), boundary deformation (protrusions) and region deformation (bends). When one considers the *parts-protrusions* continuum and plots the time till formation of the second-order shock under constant inward motion, normalized by the time till annihilation of the shape, Figure 3 (right), a curious connection is unearthed. The ratio is a measure of local width versus global width, and provides a metric for the perceptual distance of the shape from the "parts" node along this axis.

[2]Whereas third-order shocks are not generic they merit a distinct classification for several computational reasons as well as their psychophysical relevance [7]. Consider the abundance of biological and man-made objects with symmetric "bend-like" components, *e.g.*, fingers of a hand, limbs, legs of a table, branches of a tree, *etc.*

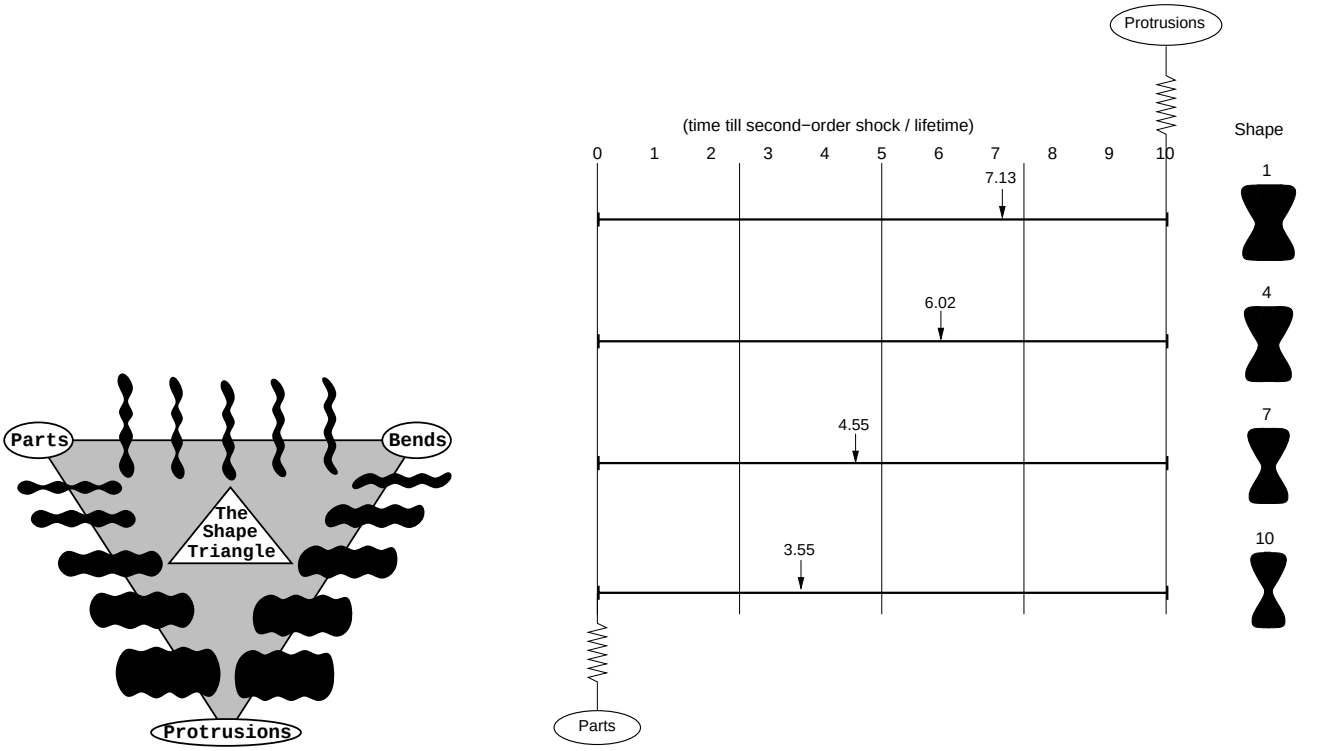


Figure 3: LEFT: The sides of the shape triangle represent continua of shapes; the extremes correspond to the "parts", "protrusions" and "bends" nodes [7]. RIGHT: Using the shock detection algorithm in [10], we plot the time till formation of the second-order shock (for constant inward motion), normalized by the lifetime of the shape, for samples of the "bow-tie" stimuli used in [7]. Observe the increase in this ratio in moving from the *parts* node (shape 10) to the *protrusions* node (shape 1).

Now, note that intermediate shapes along the *bends-protrusions* continuum closely resemble wiggles. The term "wiggle" appropriately describes not only the sinusoidal nature of the boundary modulation, but also its perceived effect on the object's central axis [4]. In the context of the shape triangle, for thin objects placed close to the "bends" node the axis is seen to wiggle, for thicker objects placed close to the "protrusions" node the axis is perceived to be straight. Within the shock-based framework, such effects are reflected in the geometry of the high-order shocks that arise, Figure 6. We are formally led to the following prediction:

Prediction 1 *The perceived center of a "wiggle" along a horizontal line in alignment with a sinusoidal peak coincides with a high-order shock* [3] *that forms under constant inward motion.*

[3] Type 3 or 4 for in-phase sinusoids, and type 2 for out-of-phase ones (mirror symmetric shapes).

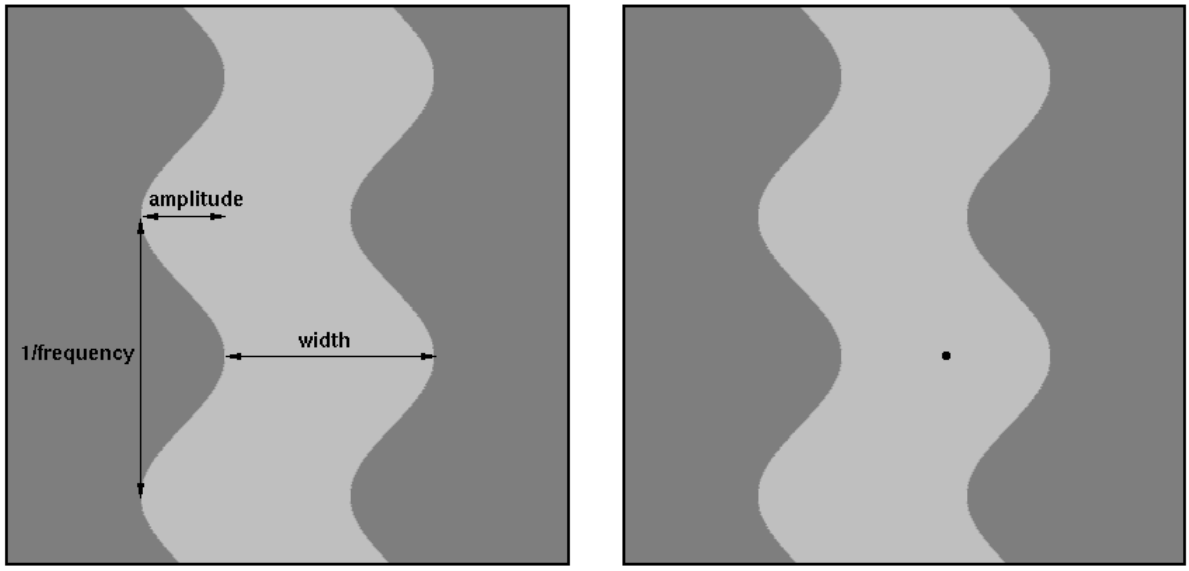


Figure 4: LEFT: The geometry of a "wiggle" stimulus. RIGHT: Is the dot to the left or to the right of the object's center?

The psychophysical experiments of Burbeck *et al.*, introduced next, involved precisely the above class of shapes.

2 Wiggles

Pizer *et al.* [9, 3, 4] have developed an alternative approach to visual shape analysis called the *core* model. Underlying the formulation of the core model is the hypothesis that the scale at which the human visual system integrates local boundary information towards the formation of more global object representations is proportional to object width. Psychophysical examinations of Weber's Law for separation discrimination support this proposal [2]. Arguing that the same mechanism explains the attenuation of edge modulation effects with width, Burbeck *et al.* have recently reported on an elegant set of psychophysical experiments where subjects were required to bisect elongated stimuli with wiggly sides. In the following we present their main findings. As predicted above, we will later demonstrate a strong correlation between these findings and computational results obtained from a shock-based description of the wiggles.

The stimuli consisted of rectangles subtending 4 degrees of visual arc in height, with sinusoidal edge modulation, Figure 4 (left). Two widths were considered (0.75° and 1.5°) and for each width there were 6 edge modulation frequencies ($0.25, 0.5, 1, 2, 4, 8$ cycles/$\circ$) and 2 edge modulation amplitudes (20% and 40% of object width). A black probe dot appeared near the center of each stimulus, in horizontal alignment with a sinusoidal peak. The subject was asked to indicate "whether the probe dot appeared to be left or right of the center of the object, as measured along a horizontal line through the dot."[4] As a sample experiment view the stimulus on the right of Figure 4 for a period of one second from a distance of 1.5 meters. You are likely to judge the dot to be to the right of the object's center. It may surprise you to find that it actually lies midway between the boundaries on either side, as can be verified by placing a ruler across the figure. In fact, despite instructions to make a local judgement your visual system is biased towards acquiring edge information across a more global spatial extent.

Burbeck and Pizer quantified this effect of edge modulation on the perceived center by varying the horizontal position of the probe dot and subjecting the data to probit analysis. The center of the object was inferred as the 50% point on the best-fitting probit function[5], and the *bisection threshold* was defined as the variance of this function. The *perceived central modulation* was then obtained as the horizontal displacement between the perceived centers in alignment with left and right sinusoidal peaks. The main findings were:

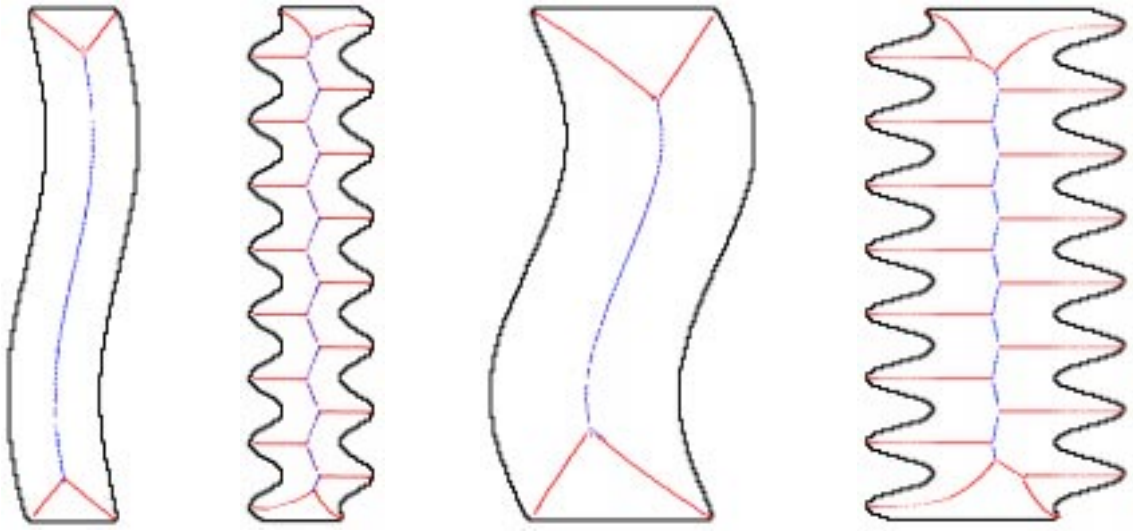
Result 1 *For a fixed edge modulation frequency the perceived central modulation decreases with increasing object width.*

Result 2 *For a fixed object width the perceived central modulation decreases with increasing edge modulation frequency.*

These results appear to be consistent with our earlier prediction. Specifically, if the perceived centers of the wiggle stimuli inferred by Burbeck *et al.* coincide with high-order shocks, *the central modulation computed as the horizontal displacement between fourth-order shocks in*

[4] See [4] for further details.

[5] The location at which a subject is statistically equally likely to judge the probe dot to be to the left or to the right of the object's center.



a. b. c. d.

Figure 5: The shock-based decription of selected 40% amplitude modulation stimuli used in [4]. a) 0.75 degree width, 0.25 cycles/degree edge modulation; b) 0.75 degree object, 2.0 cycles/degree edge modulation; c) 1.5 degree object, 0.25 cycles/degree edge modulation; d) 1.5 degree object, 2.0 cycles/degree edge modulation.

alignment with successive left and right sinusoidal peaks, Figure 6, should agree with the psychophysical data. Thus in the following section we compare computational results obtained using the shock detection algorithm of [10] with Burbeck *et al.*'s data.

3 Results and Discussion

We computed shock-based representations for all 24 wiggles using the algorithm in [10]. Results for selected stimuli are shown in Figure 5, with the geometry of the high-order shocks explained in Figure 6. As striking evidence in favor of our hypothesis, consider the computed central modulations overlayed as solid lines on the original observer data taken from [4], Figure 7. Note that any discrepancies between the computational and psychophysical data are well within the inter-subject discrepancies.

Whereas the core and the shock-based representation are motivated from quite different points of view, the strong overlap between computational and psychophysical results for each

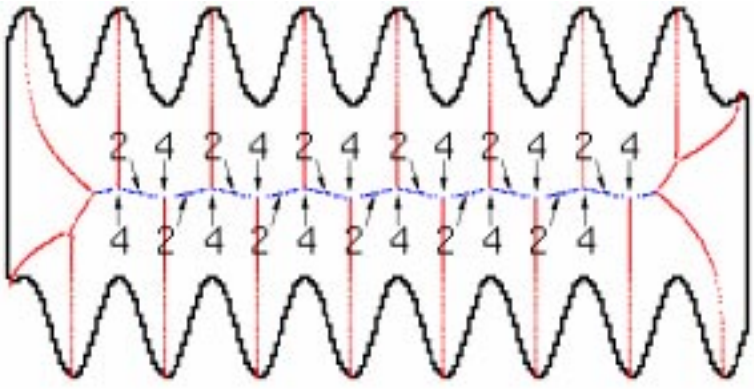


Figure 6: Shape (d) from Figure 5 is rotated and second-order and fourth-order shocks are labeled (all other shocks are first-order). Note that the fourth-order shocks are in alignment with the sinusoidal peaks.

model points to an intimate connection between the two. A precise mathematical comparison is beyond the scope of this paper. Instead, we identify the qualitative connections that have emerged.

First, the "fuzziness" of the core model, whereby the width of the core scales with object width, is paralleled by the ratio of a shock's formation time to the lifetime of the shape, a measure of local width/global width. This property is also reflected in the "bisection-threshold" or variance of the perceived centers in Burbeck *et al.*'s psychophysical experiments. Underlying this notion is the concept that the scale at which boundaries should interact to form more global object models is proportional to the spatial extent across which they communicate.

Second, the effects of edge modulation on perceived centers for the wiggle stimuli can be explained by both models. We submit that the shock-based predictions are clearer, the only relevant parameter being the location of the relevant high-order shocks. The core model is mathematically quite complex, and its simulation requires the proper choice of a number of parameters [9]. However, to put this comparison in proper context, we note that the algorithm for the core is designed to infer medialness directly from grey level images; a non-trivial task. On the other hand, the framework for shocks assumes that a foreground binary shape has already been segmented, an assumption which although not unreasonable for the synthetic wiggle stimuli, is certainly a limitation for more complicated images. Both models support medial-axis

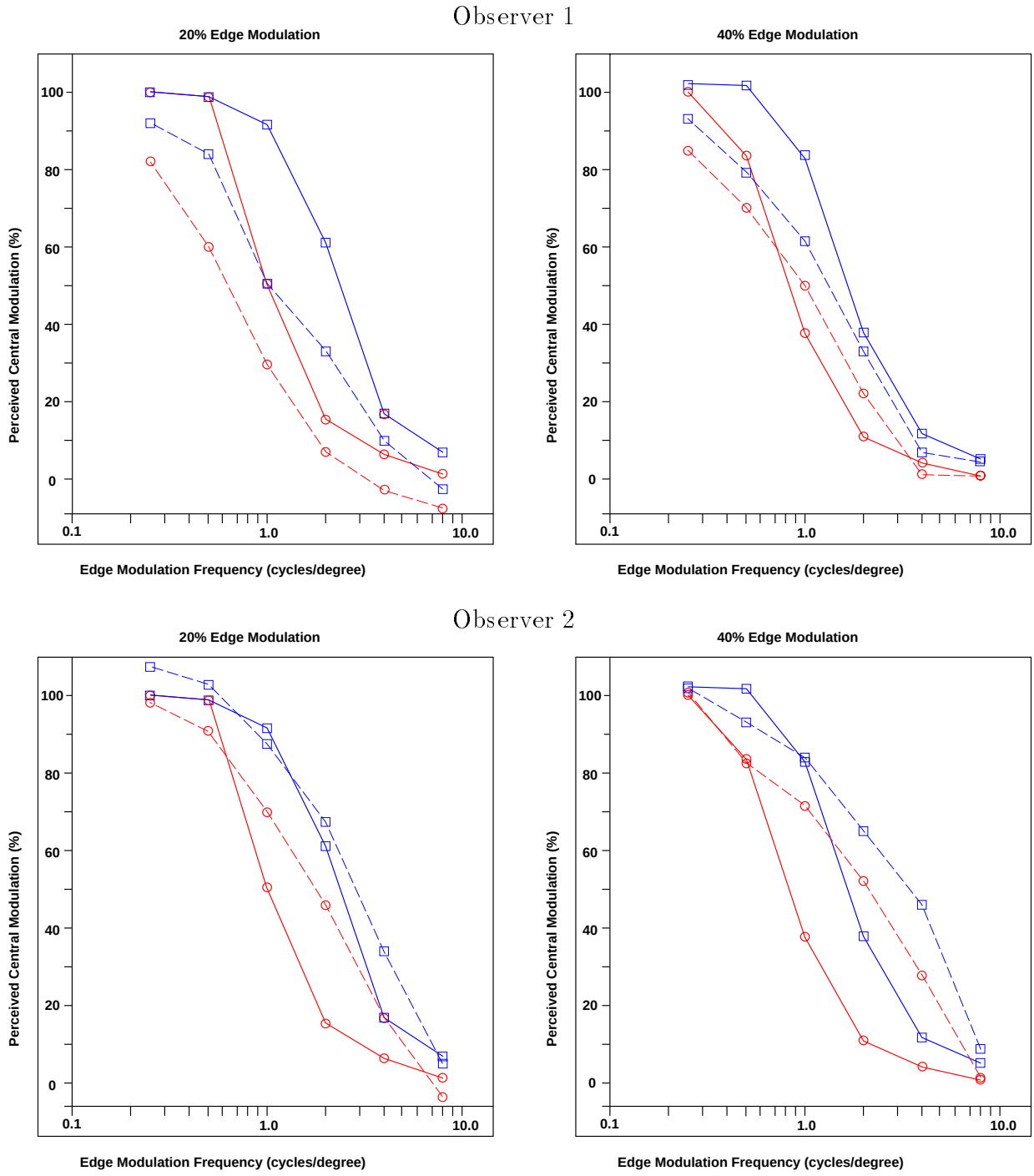


Figure 7: Central modulations computed from shock-based descriptions (solid lines) and are overlayed on the observer data (dashed lines) reproduced from [4]. The central modulations are expressed as a percentage of the edge modulation amplitude and are plotted against edge modulation frequency for amplitudes of 20% of the object width (left) and 40% of the object width (right). Results for the wider 1.5° object are depicted by the circles and for the narrower 0.75° object by the squares.

like representations for shape; see [1] for recent work on this topic. It is indeed gratifying that they are further quantifiably consistent with human performance involving shape.

Acknowledgements This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada, from the National Science Foundation, from the Air Force Office of Scientific Research and from the Army Research Office. We thank Christina Burbeck, Steve Pizer and Xiaofei Wang for fruitful discussions and for kindly supplying us with the wiggle stimuli.

References

[1] C. Arcelli, L. P. Cordella, and G. S. di Baja, editors. *Aspects of Visual Form Processing.* World Scientific, June 1994.

[2] C. Burbeck and S. Hadden. Scaled position integration areas: Accounting for weber's law for separation. *Journal of the Optical Society of America A*, 10(1):5–15, 1993.

[3] C. A. Burbeck and S. M. Pizer. Object representation by cores: Identifying and representing primitive spatial regions. *Vision Research*, 35:1917–1930, 1995.

[4] C. A. Burbeck, S. M. Pizer, B. S. Morse, D. Ariely, G. S. Zauberman, and J. Rolland. Linking object boundaries at scale: A common mechanism for size and shape judgements. *Vision Research*, In press, 1996.

[5] B. B. Kimia, A. Tannenbaum, and S. W. Zucker. Toward a computational theory of shape: An overview. *Lecture Notes in Computer Science*, 427:402–407, 1990.

[6] B. B. Kimia, A. Tannenbaum, and S. W. Zucker. Shape, shocks, and deformations I: The components of two-dimensional shape and the reaction-diffusion space. *Int. J. Computer Vision*, 15:189–224, 1995.

[7] B. B. Kimia, A. R. Tannenbaum, and S. W. Zucker. The shape triangle: Parts, protrusions, and bends. In *Aspects of Visual Form Processing*, pages 307–323. World Scientific, June 1994.

[8] P. D. Lax. Shock waves and entropy. In *Contributions to Nonlinear Functional Analysis*, pages 603–634, New York, 1971. Academic Press.

[9] B. S. Morse, S. M. Pizer, and C. A. Burbeck. General shape and specific detail: Context-dependent use of scale in determining visual form. In *Aspects of Visual Form Processing*, pages 374–383. World Scientific, June 1994.

[10] K. Siddiqi and B. B. Kimia. A shock grammar for recognition. In *CVPR'96 (IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, California, June 21–23, 1996)*, Washington, DC., June 1996. Computer Society Press.