Perceptual Support of Diagram Creation and Editing

Eric Saund and James Mahoney

Palo Alto Research Center
Palo Alto, CA 94304, USA
{saund,mahoney}@parc.com

Abstract. Diagrams mediate thinking and understanding largely through the human visual system's innate ability to perceive visuo-spatial structure. Tools for working with diagrams will benefit from the ability of machines to identify visual structure in concert with their human users. This poster and its companion summarize recent progress in *perceptually-supported* diagram creation and editing. In particular, our research group has deployed a document image editing program realizing some measure of Gestalt Perceptual Organization for sketches and diagrams.

1 Background: Diagrams as Input versus Diagrams as Intrinsic Representations

Over the past decade many investigators have identified the goal of building "intelligent" computer programs for human work with sketches and diagrams. In their envisionments, many of which have experimental implementations, the user draws and annotates in freehand strokes using pen and paper or digital stylus on instrumented whiteboards or tablets. The machine parses and recognizes these markings, then responds by cleaning up or "beautifying" drawings, accessing databases, performing simulations, or invoking reasoning engines [15, 8, 9, 5, 1, 16, 14, 7, 13].

By and large, most work in this area is focused on providing naturalistic pen interfaces as *input* mechanisms—input to graphics formatting and presentation programs, input to database search queries, input to reasoning systems, etc. The value-added smarts of the computer is viewed as residing in domain knowledge and competence, while the greatest barrier to actualizing this competence remains the difficulty of achieving robust and accurate machine interpretation of diagrammatic input.

To mitigate this problem, a common and sensible strategy has been to place strong limits on what *may* be drawn by the user, and apply correspondingly strong prior constraints about what *can* be recognized by the system [2]. For example, one group has demonstrated the ability to interpret closed polygonal paths as two dimensional-projections of solid objects [1], while another has shown that arrows drawn in idiosyncratic ways can be recognized in the symbology of military diagrams [5]. In systems of these designs it is important that ambiguity

in the interpretation of any given input marking be eliminated quickly. Otherwise is to risk combinatorial explosion in the mapping of input primitives (eg. pen strokes) to the components of domain model shapes.

We observe however that another stance is possible with regard to building smart, perceptually-enabled diagram manipulation tools. This approach focuses less directly on mapping from a diagram to domain content, but instead views the diagram as an object of interest to manipulate in its own right, i.e., as a form of external working memory. The process of drawing and elaborating and re-working a diagram engages the perceptual system to form multiple overlapping interpretations consisting of different segmentations and groupings of the primitive markings. On this hypothesis, these motor and perceptual processes in turn engage cognitive processes to conceptualize domain content in different ways, corresponding to alternative, sometimes partial, sometimes only partially coherent, "parses" of the diagram.

2 Perceptual Organization and Diagram Analysis

Perceptual segmentation and grouping processes are associated with the Gestalt principles of Psychology, dating originally to the early 20th century but receiving attention in the contemporary study of Computer Vision [17, 12, 6, 11, 18, 3, 4, 10].

While human vision is remarkably adept at recognizing known objects or object categories in complex scenes, it is equally capable at finding patterns and creating sense out of utterly unfamiliar imagery. Aspects of visual scene analysis occurring apart from object recognition per se include figure/ground segmentation; segmentation of regions into coherent objects; assigning relative depths to surfaces; detection of potentially interesting or novel events; factoring shadows and other lighting effects from geometrical properties; tracking moving objects; and detecting coherent motion among disparate motion cues. The ecological rationale for visual systems possessing these abilities has been discussed at length [18]. The Gestalt psychologists set out to understand perception via simple figures that distill salient visual properties or pattern qualities that in natural scenes are found in confluent abundance. As markings on paper, many of these figures bear notable similarity to the representational constituents of diagrams.

The Gestalt principles of primary concern in diagram analysis include Smooth Continuation, Figural Closure, Spatial Proximity, Symmetry, and Feature Similarity. See Figure 1. While these principles offer explanatory power in accounting for human perceptual capabilities, they have proven extremely difficult to formalize as algorithms for computer programs. Although each of these principles is intuitively associated with formal geometrical properties (e.g. smoothness $\sim$ continuity in the derivative of contour tangent direction; closure $\sim$ a topological donut) the perceptually relevant phenomena are "fuzzy," or tolerant to deviations from straightforward mathematical formulations. Moreover, these principles interact and trade off with one another. Computer Vision lacks any adequate

formulation for uniting the various Gestalt phenomena under a common theoretical or algorithmic framework.

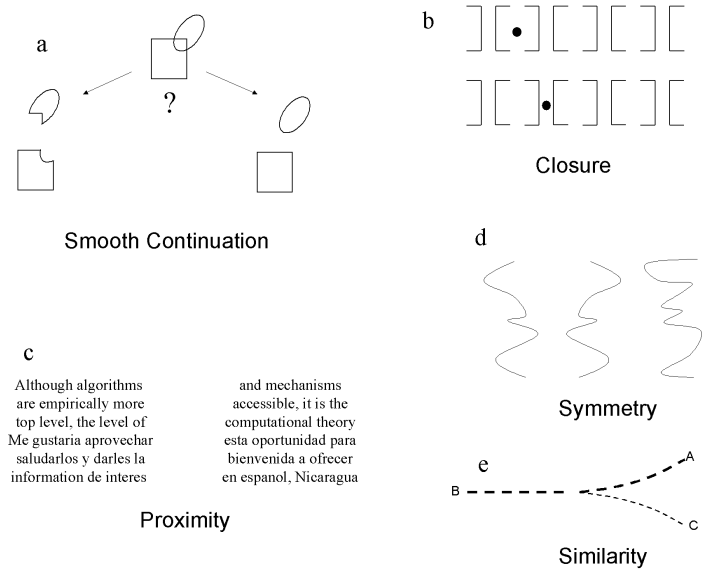

Fig. 1. Illustrations of five Gestalt principles of visual perceptual organization. a. The central figure appears to be a combination of parts having smooth boundaries. b. The top dot appears to fall on a foreground object defined by a nearly closed boundary contour, while the bottom dot appears to lie on the background. c. The apparent visual partitioning of the text based on proximity overrides a partitioning based on semantic content. d. The curves that appear to go together are the pair forming a bilateral symmetry. e. The curve appears to continue from point B on the basis of similarity of its local properties.

Nonetheless, work in our laboratory suggests that tools for working with diagrams can benefit from the ability to identify visual structure in accordance with the Gestalt laws, even at today's relatively primitive level of technological development. Perceptual grouping is useful in at least two ways. First, as machines become capable of perceiving visual structure corresponding to that readily identified by human users, user interface techniques can be devised that permit people to select and manipulate salient collections of visual markings at will, and thereby reconfigure diagrams according to their imaginations. Second, successful perceptual organization offers a stable foundation for recognition of symbols, shapes, notations, and domain objects. Not only do visually salient segmentation and grouping raise significant structure above noise, clutter, and imprecision in drawing and imaging, but these processes inherently make available alternative interpretations of locally ambiguous image data, which simplifies

processes for matching configurations of lines and symbols to libraries of known shapes and notations.

The companion poster to this one presents a prototype application, called *ScanScribe* that demonstrates how visual structure at the level of Perceptual Organization can be exploited to facilitate selection and editing of visually salient figural objects in diagrams.

References

1. Alvarado, C., and Davis, R.; [2001]; "Resolving ambiguities to create a natural computer-based sketching environment," *Proc IJCAI*, Seattle, Vol.2: 1365-1371.
2. Blostein, D., Lank, E., and Zanibbi, R.;[2000]; "Treatment of Diagrams in Document Image Analysis," in M. Anderson, P. Cheng, and V. Haarslev, eds, *"Theory and Application of Diagrams*, LNCS 1889, 330-344.
3. Boyer, K., and Sarkar, S., eds.;[1999]; *Computer Vision and Image Understanding Special Issue on Perceptual Organization*, 76(1).
4. Boyer, K., and Sarkar, S., eds.;[2000]; *Perceptual Organization for Artificial Vision Systems.* Kluwer.
5. Forbus, K., Furguson, R.W., and Usher, J.M.; [2001]; "Towards a Computational Model of Sketching," *Proc. IUI '01*, Santa Fe.
6. Green, C.; [2000]; "Introduction to: 'Perception: An introduction to the Gestalt-theorie'
   by Kurt Kaffka (1922)", "http://psychclassics.yorku.ca/Koffka/Perception/intro.htm".
7. Corel Grafigo Pen-Based Business Graphics Software; "http://www.corel.com".
8. Gross, M.; [1996]; "The electronic cocktail napkin - computer support for working with diagrams," *Design Studies*, 17(1):53-70.
9. Igarashi, T., Matsuoka, S., Kawachiya, S., and Tanaka, H.; [1997]; "Interactive Beautification: A Technique for Rapid Geometric Design", *Proc ACM UIST*, Banff, 105-114.
10. Jacobs, D., and Lindenbaum, M.; eds.;[2001]; *POCV 2001: The Third Workshop on Perceptual Organization in Computer Vision*, CIS Report CIS-2001-05, Technion, Haifa, Israel.
11. Kanizsa, G.; [1979]; *Organization in Vision: Essays on Gestalt Perception*, Praeger, New York.
12. Koffka, K.; [1922]; "Perception: An introduction to Gestalt-theorie. *Psychological Bulletin*, 19: 531-585.
13. Leszynski InShape Software Development Toolkit; "http://www.leszynski.com/tabletpc/tpcinShape.htm".
14. Landay, J.A., and Myers, B.A.;[2001]; "Sketching Interfaces: Toward More Human Interface Design," *IEEE Computer* V. 34. No. 3, March 2001, 56-64.
15. Pavlidis, T., and Van Wik, C.;[1985]; "An Automatic Beautifier for Drawings and Illustrations," *Proc. ACM SIGGRAPH*, 225-234.
16. Stahovich, T., Davis, R., and Shrobe, H.;[1998]; "Generating Multiple New Designs from a Sketch," *Artificial Intelligence*, Vol 104, 211-264.
17. Wertheimer, M.; [1923]; "Laws of Organization in Perceptual Forms", in Ellis, W., ed, *A source book of Gestalt psychology*, Routledge & Kegan Paul, London, 1938.
18. Witkin, A., and Tenenbaum, J.; [1983]; "On the role of structure in vision," in J. Beck, B. Hope, and A. Rosenfeld, eds., *Human and Machine Vision.* Academic Press.