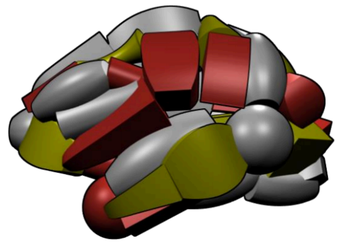
Vertices are Effective in Perceptual Grouping (and Ungrouping) in Object Recognition

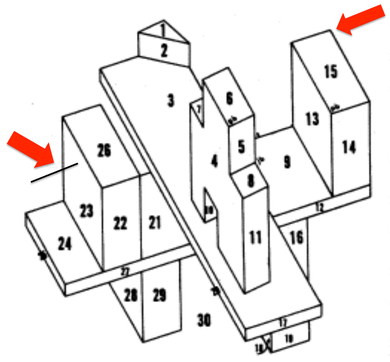
Isabel Irawan[2*], Eshed Margalit[2], Emily Meschke[2], Sarah B. Herald[2,4], & Irving Biederman[1,2]
[1]Department of Psychology, [2]Neuroscience Program, University of Southern California, [4]Yale University
*iirawan@usc.edu
http://geon.usc.edu

Image Understanding Lab



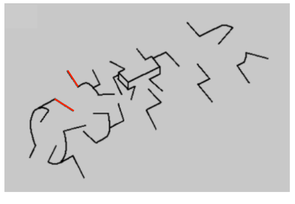
Geometric Background of Vertices

Guzman et. al (1968) established that the grouping of surfaces into volumes in complex scenes—as in the Blocks-World "Bridge" below– could be achieved largely on the basis of constraints emanating from vertices where two or three contours co-terminate to decode a given shape. An L-vertex, the point at which two contours co-terminate**,** shown by the upper right corner of surface 15, provides highly reliable evidence that a surface terminates at that vertex. In contrast, an X-vertex, shown over surface 23, where two contours cross without a change of direction at their crossing, provides no constraints for grouping and can be disregarded in shape-based models of vision.

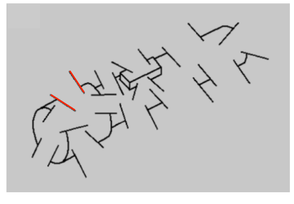


Previous Research

There is evidence that irrelevant L-vertices (those that interfere with the grouping of the contours of a single part) are more disruptive to object recognition than irrelevant T-vertices (Vessel et al., 2016).



Gaps bridged by L-vertices signal the termination of a surface, thereby inhibiting smooth continuation.

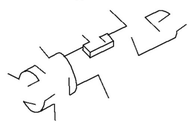
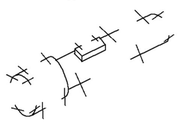


Same gaps and segments as on the left but slightly shifted to produce T-vertices, thereby allowing smooth continuation.

Purpose

Would the addition of extraneous line segments to gaps in an object's contours that created (inappropriate) L-vertices interfere with recognition of that object whereas the same contours positioned so as to create (inappropriate) X-vertices produce little or no interference with recognition?
Guzman's predictions:

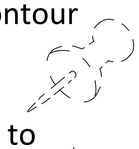
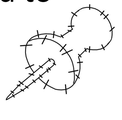
1. **Irrelevant L-vertices added to gaps should create greater difficulty in object recognition than X-vertices.**

Recognition of  should be more difficult than recognition of 

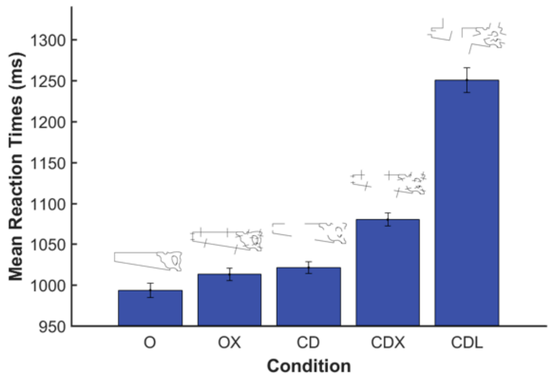
2. **Segments added to gaps to produce L-vertices should be more disruptive to recognition than when the same segments are translated to produce X-vertices.**

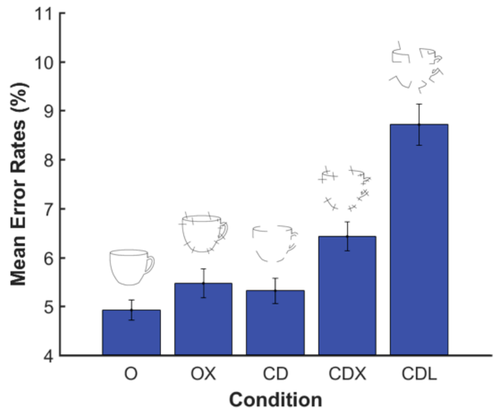
Experiment

TASK: Speeded naming of familiar objects in one of five conditions:

1. Original line drawings (O) 
2. Drawings with 50% of the middle of each contour deleted (CD) 
3. Original drawings with line segments added to produce X vertices (OX) 
4. Contour deleted drawings (CD) with line segments added to produce X vertices across the contours (CDX) 
5. Contour deleted drawings (CD) with line segments added to produce L vertices bridging the gaps (CDL) 

Results





Only the addition of the L-vertices to the gaps yielded a sizable and reliable increase in RTs and error rates compared to the original images. There were modest but not reliable decrements in performance (longer RTs and higher error rates) from the contour deletion and the addition of X-vertices.

Conclusion

Guzman was right on both counts:

1. L-vertices provide a strong signal for the termination of a surface. In the present context, where the Ls inappropriately signaled the termination of the surface of an object part, recognition performance was markedly degraded.
2. X-vertices have little or no effect on object recognition aside from the noise of their irrelevant contours. When the identical segments that produced L-vertices were translated so that they produced X-vertices, there was little or no interference in the recognition of the images.

References
Guzmán, A. (1968). Decomposition of a visual scene into three-dimensional bodies. In Proceedings of the December 9-11, 1968, fall joint computer conference, part I (pp. 291-304). ACM.
Vessel, E., Biederman, I., Subramaniam, S., Greene, M., (2016). Effective Signaling of Surface Boundaries by L-Vertices Reflect the Consistency of their Contrast in Natural Images. *Journal of Vision,* in press.
Acknowledgments
Supported by NSF BCS 0617699 and the Dornsife Research Fund