The Scale Space Aspect Graph*

David W. Eggert[1] Kevin W. Bowyer[1] Charles R. Dyer[2] Henrik I. Christensen[3] Dmitry B. Goldgof[1]

[1]Department of Computer Science and Eng.
University of South Florida
Tampa, Florida 33620
eggertd, kwb or goldgof@csee.usf.edu

[2]Department of Computer Science
University of Wisconsin
Madison, Wisconsin 53706
dyer@cs.wisc.edu

[3]Institute of Electronic Systems
Aalborg University
DK-9220 Aalborg East
hic@vision.auc.dk

Abstract

Currently the *aspect graph* is computed under the assumption of perfect resolution in the viewpoint, the projected image and the object shape. Visual detail is represented that an observer might never see in practice. By introducing scale into this framework a mechanism is provided for selecting levels of detail that are "large enough" to merit explicit representation, effectively allowing control over the size of the aspect graph. This paper introduces the *scale space aspect graph*, and focuses on an interpretation of the scale dimension in terms of the spatial extent of image features. A brief example is given for polygons in a plane.

1 Introduction

The aspect graph is considered an important view–centered representation and has been the subject of much recent research. However, one issue raised at a recent panel discussion [7] is the lack of use of *scale* information (a problem not unique to aspect graphs). It is hoped that the use of scale can be seen as a method of making the representation more realistic.

To date, the aspect graph has been computed only under the ideal assumptions of perfect resolution in the viewpoint, the projected image and the object shape, leading to the following practical difficulties: (1) a node in the aspect graph may represent a view of the object that is seen only from an extremely small cell of viewpoint space, (2) the views represented by two neighboring nodes may differ only in some realistically indistinguishable detail, and (3) very small changes in the detail of 3–D object shape may drastically affect the number of aspect graph nodes. The net result is a representation of large overall size, containing potentially unimportant object views. Using scale information should reduce this large set of theoretical aspects to a smaller set of the "most important" aspects.

*This work was supported at the University of South Florida by Air Force Office of Scientific Research grant AFOSR-89-0036 and by National Science Foundation grant IRI-8817776.

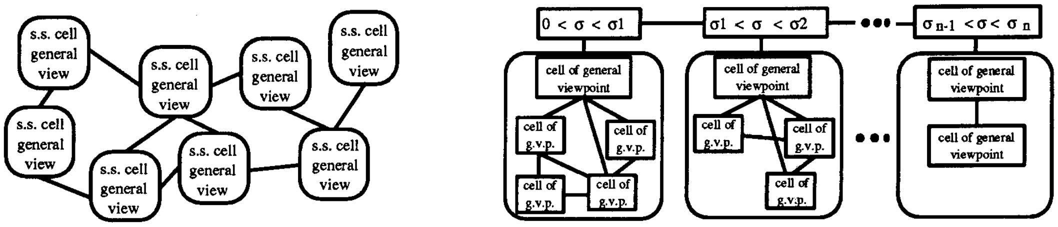
2 The scale–space aspect graph

The conventional aspect graph is defined such that (1) there is a node for each *general view* of the object as seen from some maximal connected cell of *viewpoint space*, and (2) there is an arc for each possible transition, called an *accidental view* or a *visual event*, between two neighboring general views. A general (accidental) viewpoint is defined as one from which an infinitesimal movement in every (some) possible direction in viewpoint space results in a view that is equivalent (not equivalent) to the original. Algorithms developed to date compute aspect graphs for different domains (including polyhedra [8, 15, 20, 21], various curved–surface objects [3, 5, 12, 16, 17, 19] and articulated assemblies [18]), view representations (usually a qualitative description of the line drawing) and models of viewpoint space (2–D view sphere under *orthographic projection* and 3–D space under *perspective projection*). A review of recent work can be found in [2].

When the phrase "scale space of X" is used, it is taken to mean a parameterized family of X in which the detail of features in X monotonically decreases with increasing scale. Under some parameterized transformation, usually Gaussian, the nature of X is changed in such a way that its qualitative features at a given scale can be traced back across all lower scales ("causality"). Original popularized by Witkin's analysis of the inflections of a 1-D signal [22], the scale space concept has been applied to the curvature of 2-D curves [14], 2-D intensity maps [10, 13] and 3-D object shape [11]. In this paper, X is an aspect graph, or more precisely, the corresponding parcellation of viewpoint space underlying the qualitative description of the aspect graph. At the scale value $\sigma = 0$, the parcellation of viewpoint space, and therefore the aspect graph, is exactly as computed by some known algorithm. Ideally, as σ increases, this parcellation should deform so that at certain discrete values of scale the aspect graph becomes simpler.

The qualitative structure of scale space, the multi–dimensional space parameterized by viewpoint location

0-8186-2855-3/92 $3.00 ©1992 IEEE



(a) interconnected cells representing general views in scale space

(b) complete aspect graphs across discrete ranges of scale

Figure 1: Conceptual Depictions of the Scale Space Aspect Graph.

and scale value (for example, a 4-D (x, y, z, σ) space under perspective projection), can be represented in (at least) two forms. The first, shown in Figure 1.a, resembles the *asp* of Plantinga and Dyer [15]. Each boundary (visual event surface) in the parcellation is parameterized by the viewpoint location and scale, and each cell or "volume" in the space corresponds to a general view that exists over many scales. The second, shown in Figure 1.b, is analogous to the *visual potential* of articulated assemblies of Sallam *et. al.* [18]. This form presents an explicit sequence of aspect graphs, each constant in structure over a range of σ. While this second form is more explicit and perhaps simpler conceptually, there is potential redundancy in the multiple instances of the aspect graph.

2.1 Interpretations of scale

Previous scale space representations when applied to intensity functions interpret the scale parameter in terms of the solution to the diffusion equation [10] (or more specifically, as the variance of a Gaussian kernel used for blurring). Only under this interpretation will the qualitative features disappear monotonically with increasing scale [10]. Unfortunately, the entities on which the aspect graph concept is based (visual events, projected line drawings, 3-D shape, etc.) are not intensity functions, so we will relax this restriction.

In an effort to define "blurring" of the representation, we return to the three areas of concern mentioned earlier: (1) Blurring the parcellation most closely corresponds to examining relative cell size, which previous researchers have looked at as view probability [1, 9, 21]. This interpretation can be realized by changing the relative sizes of viewer and object as a function of scale, making smaller cells no longer accessible. (2) A more realistic set of visual event boundaries is generated if a minimum detectable feature size is given as a function of image resolution. As image resolution is reduced (blurring the image), fewer features are distinguishable, and neighboring views may no longer be distinct. (3) If the object surface is blurred as a function of scale, fewer features that visually interact will exist, thereby reducing the number of event surfaces and subsequently the cells in the parcellation.

Each of the above approaches offers some reduction in aspect graph size as a function of scale, while corresponding to slightly different visual phenomena. Ideally all of these effects would be combined in one solution. However, this paper only focuses on what initially seems the most promising of the three ideas – the use of image resolution as a scale measure. (See [6] for more detail on the other two approaches.)

2.2 Scale as image resolution

Visual event calculation is simplified by assuming an infinite resolution image plane and only considering qualitative changes in a line drawing. But this ignores practical effects, such as when a feature (or even an object) becomes too small to distinguish in a real image as one moves away from it. Reintroducing a quantitative measure such as finite image resolution, as done by others in determining visibility constraints for automatic sensor placement [4], steps toward reality. However, a method must be chosen for specifying resolution as a function of scale. The use of a camera's image coordinate system, which seems intuitive at first, unfortunately is complicated by the need for the camera's focal length, field of view, as well as viewing direction and position. A perhaps more natural alternative, commonly used by psychologists and biologists, is the degree of visual arc occupied by the feature (see Figure 2). Under this model a feature's size can be described by a single angular value, in the range $0° - 360°$.

In order to measure a feature, it must have some spatial extent in the image. Junctions only occur as individual points. Contours and object surfaces generally project to some measurable extent in an image. It is not sufficient to measure the size of the entities on the object. It is the projection of these features that is measured by the visual arc. A straight edge's visual arc is defined by the distance between its projected endpoints, as shown in Figure 2. For a curve, the

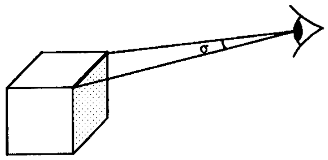


Figure 2: Visual arc occupied by an edge feature.

maximum distance between any two projected points along its length could indicate visual extent. The maximum distance between any two projected points on the boundary of a face could describe its overall image area, as between the corners of the cube face in Figure 2. Alternatively, the radius of a projected circle that encompasses the feature could be used.

This model can be understood as follows. Image resolution, or pixel size, directly corresponds to the minimum value of σ necessary to distinguish a feature. At a value of 0° the camera has infinite resolution, and current aspect graph assumptions hold. At a value of 360° all features project to a single pixel. At values in between, a given feature may project to several,or only one pixel, depending on viewpoint position, becoming indistinguishable at greater distance. Similar effects are obtained by maintaining a fixed viewpoint, and varying the pixel size with greater values of σ. These effects are indicative of changes in the visual event surfaces as a function of scale. Deviations from the ideal case of a ruled surface occur because two features no longer need align exactly to project to the same pixel. This change in nature of event surfaces is elaborated upon more in the following case study.

2.2.1 Nonconvex polygons in a plane

Some of these general ideas can be clarified by examining in more detail a particular class of objects, nonconvex polygons in a plane, which generate a 3–D scale space (x, y, σ). Visual events involve the only feature of interest, an edge. Each visual event, be it visibility of a single edge or occlusion of two edges, occurs when two vertices on the polygon align. Under ideal assumptions, this alignment is seen from the line in the plane containing the two vertices, and the appearance of the object is different on either side of this line. For a given scale σ, this line becomes a circular arc as shown in Figure 3. The edge endpoints, e_1 and e_2, project to the same pixel for each viewpoint, v, that is the vertex of a triangle with angle σ. The viewpoint position is governed by the law of cosines as applied to the triangle in the figure. This curve is actually part of a surface in scale space as shown in Figure 4. The edge is only visible for those points inside the curve (surface) from which each endpoint is distinguishable.

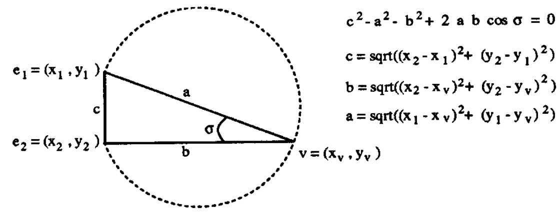


Figure 3: Visual event curve defined for an edge.

In order to discuss interactions of edges, the concept of a "pseudo edge" must be defined. A pseudo edge is a chain of edges on the polygon boundary. It is said to exist only when each edge in the chain is indistinguishable, but the chain as a whole is of sufficient size to be detected. Such a concept does not occur in normal aspect graph theory since every feature is visible from any distance. For instance, consider a red edge and a blue edge that are connected to each other. In this new framework, there are viewpoints at which each edge itself may be too small to see, but the visual angle between the separated ends may be sufficient to detect a purple or blurred combination of the edges.

The visual appearance of a pseudo edge can be categorized based on whether it forms a convex or concave corner on the object, as shown in Figure 5. For the convex–angled edge pair shown in part (a), regions exist in which each edge by itself is visible, both edges are visible, or only the pseudo edge is detectable. Outside of these regions all parts of the edges project to a single pixel. For the concave–angled edge pair in part (b), there are additional visibility regions related to self–occlusion between the edges. The boundaries of these regions are of two kinds. One is the visibility limit (the dashed curve) of the occluding edge (not shown in the figure), for example, that which would be connected

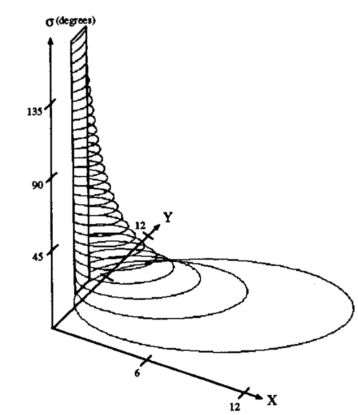
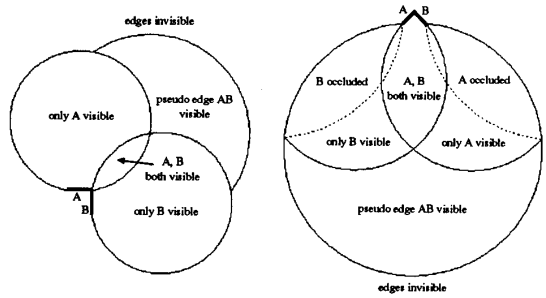


Figure 4: Visibility boundary surface for edge.



(a) visibility of convex-angled edge pair (b) visibility of concave-angled edge pair

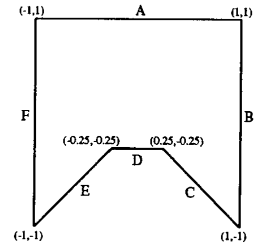
Figure 5: Visibility ranges of pseudo edge pairs.

to the unattached end of A. The other limit, marking complete occlusion, occurs when the unattached ends of the edges appear to coincide. This is the same curve that marks the visibility of the pseudo edge combining A and B. The visual event associated with this curve changes at the intersection of it and the visibility curves for the individual edges.

Calculating the scale space aspect graph representation can proceed in several ways. The easiest conceptually is to directly calculate the parcellation of scale space using the parameterized event surface equations. Unfortunately, even with existing surface intersection techniques, this is not a trivial task. Perhaps a more straightforward approach is to construct the viewpoint space parcellation for a starting scale, and incrementally adjust it for changes in scale that cause an alteration in the parcellation structure. For any given scale, the set of visual event curves can be found using the equations for the visibility limits of the actual edges and all possible pseudo edge combinations. Using the model interactions of Figure 5, meaningful extents of event curves can be found by calculating intersection points amongst themselves and with the polygon.

The types of changes in the viewpoint parcellation, termed *scale events*, are limited. The initial change from the ideal representation to one corresponding to a nonzero scale involves deforming the existing event lines into circular event arcs, and adding event curves relating to the pseudo edges as indicated above. The remaining events, which can occur in combination, fall into three basic categories:

(1) *Begin (end) overlap of two curves.* Two curves, either visual event curves or polygon edges, may make initial (final) contact with their meaningful portions at a single point in the parcellation. When two event curves are involved, a cell representing the region of overlap of the areas bounded by the curves is created (deleted). If an event curve and edge are involved, only the boundary of an existing cell is usually altered.



(a) nonconvex polygon

Pseudo edges	Corner type	Subedges
G	concave	E D
H	concave	D C
I	concave	G H, G C, E H
J	convex	F E
K	convex	C B
L	convex	B A
M	convex	A F
N	convex	F G
O	convex	F I
P	convex	H B
Q	convex	I B

(b) pseudo edge definitions

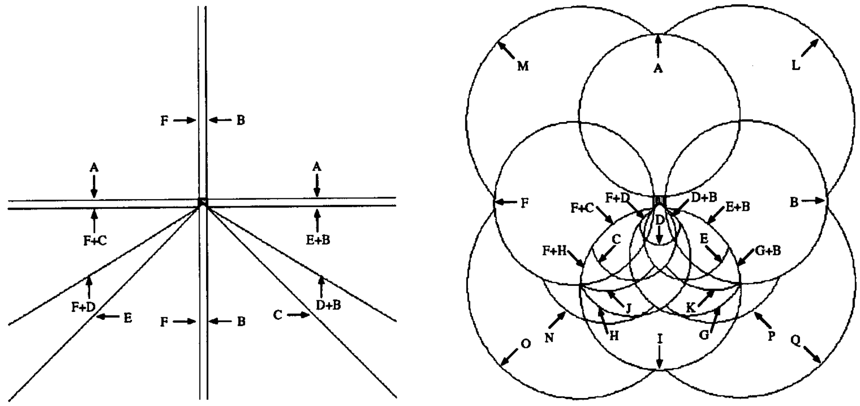
Figure 6: Definition of example nonconvex polygon.

(2) *Triple point.* The meaningful portions of three curves, either event curves or edges, intersect at a single point in the parcellation. When three event curves are involved, a different region is bounded by only these curves before and after the critical scale, one replacing the other in the parcellation. If two event curves and an edge interact, a single new region is created (removed), because the object exists on one side of the edge, leaving only one meaningful cell. For one event curve and two edges a single region may again be created (removed). However, this change also usually coincides with the insertion (removal) of a meaningful segment of the event curve.

(3) *Curve coincidence.* The meaningful portions of two curves, either event curves or edges, coincide along their length in the parcellation. In the case of two event curves, they intersect in a point in the parcellation and are part of the boundaries of two distinct sets of regions existing before and after the critical scale. At the critical scale the definitions of the two curves are the same. When an event curve and an edge are involved, the region bounded by them is deleted from the parcellation along with the event curve, since it has passed completely into the interior of the object.

2.2.2 An example

The object analyzed is the six–sided polygon shown in Figure 6. In addition to the six actual edges, eleven pseudo edges can be defined, three of which fit the concave corner model and eight the convex model. The structure of the scale space aspect graph was constructed using the incremental technique described earlier. The viewpoint space parcellation was calculated (event curves generated automatically, meaningful portions segmented by hand) for evenly sampled values of scale to isolate the scale events. Certain of the event scale values could then be directly calculated once the participating curves were determined, while others required an iterative search. Due to space restrictions, only a summary of the entire solution is given here (for complete results see [6]).



single letters indicate visibility boundary, + indicates occlusion boundary

(a) parcellation at visual angle of zero degrees (b) parcellation at visual angle of five degrees

Figure 7: Labeled viewpoint parcellations of polygon for initial transformation.

There are seventeen visual event surfaces (curves), one each based on the real and pseudo edges. In the 3–D scale space there are fifty total aspects. These cells exist over varying scale ranges; the largest range being 256° for the view of only pseudo edge I, which is the final cell to disappear, and the smallest range being only 14°, for which one of the views seen is edges $B-E$ connected to one another. There are 24 different scale events, including the initial transformation to a nonzero scale. This change is shown in Figure 7. The event curves are labeled according to the edge visibility or pair occlusion they represent. Instances of most of the different types of scale events are shown in Figure 8. Only the interactive portion of the parcellation is displayed, with the curves labeled as in Figure 7 and the affected cells shaded.

3 Discussion

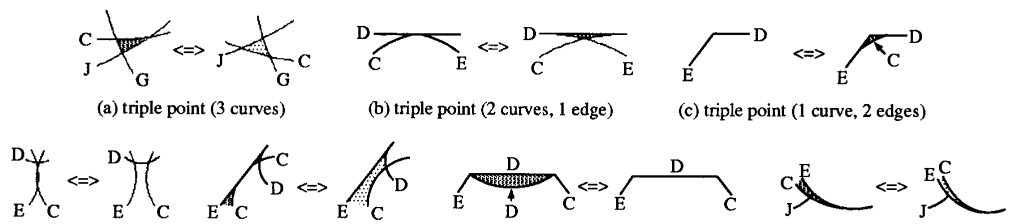
Having presented a method of incorporating quantitative information into the aspect graph representation, we must evaluate its properties. One property of a scale space representation is that the features of X should monotonically decrease in complexity. Although not shown in the previous section, the number of cells in the viewpoint parcellation does not monotonically decrease. There are certain scale events in which only new cells are created (even ignoring the large increase as a nonzero scale is used). However, the eventual result at the final scale is a graph in which no views are visible since everything projects to one pixel. Thus while local increases in the parcellation complexity may occur, the global trend is a decrease in size.

The actual aspect graph for a particular scale is also different from the ideal. First, because a cell corresponding to a view in scale space is not necessarily "convex", its projection into viewpoint space may yield multiple cells. Thus the normal 1–to–1 correspondence of cells to aspects is no longer preserved. Perhaps more important is that the parcellation of viewpoint space is now finite in size (for nonzero scale), and the event boundaries are much more realistic. This greater correspondence to our intuitive notions is important.

One claim is that a smaller set of the "most important" aspects could be produced using this technique. Measuring importance can be done in different ways. In keeping with Witkin's analysis [22], one measure could be the range of scale values over which a view is seen. But since cells have size, a better measure may be the "volume" of the cell in scale space. For the example object, the view of the concavity existed over the greatest scale range. The views exhibiting the largest volume were those seeing either a single edge not in the concavity, or one of the pseudo edges formed from these edges. It is interesting to note that these views are not necessarily the largest in the ideal parcellation.

Measuring the volume of a cell can be done for all of scale space or just some range of scale values. If only a single scale value is used the view probability concept results. This suggests that one of the other scale interpretations could be combined with the current to develop a ranking of the important aspects. One area of future research is to explore the relationship of the visual arc and cell size interpretations, determining if the effects are truly independent and how to best combine them. Continuing to explore other scale interpretations is also important.

One other avenue of research is extending the current interpretation to broader domains. We are cur-



(a) triple point (3 curves) (b) triple point (2 curves, 1 edge) (c) triple point (1 curve, 2 edges)

(d) overlap (2 curves) (e) overlap (1 curve, 1 edge) (f) coincidence (1 curve, 1 edge) (g) coincidence (2 curves)

Figure 8: Examples of scale events in viewpoint parcellations of the polygon.

rently considering polyhedra. For convex objects, visual events include changes in face visibility, as well as changes in apparent face shape. For nonconvex objects, occlusion between faces also occurs. The equations of event surfaces in 3–D space are fairly direct extensions of those in 2–D space, but the difficulty lies in determining which points on a boundary interact to form the defining visual arc angle for arbitrary viewpoints.

References

[1] Ben-Arie, J. "Probabilistic models of observed features and aspects with application to weighted aspect graphs", *Patt. Rec. Letters*, **11**, 421-427, 1990.

[2] Bowyer, K., Dyer, C. "Aspect graphs: An introduction and survey of recent results", *Int. J. Imaging Systems and Technology*, **2**, 315-328, 1990.

[3] Chen, S., Freeman, H. "On the characteristic views of quadric-surfaced solids", *IEEE Work. Directions in Automated CAD-Based Vision*, 34-43, 1991.

[4] Cowan, C. "Automatic camera and light-source placement using CAD models", *IEEE Work. Directions in Automated CAD-Based Vision*, 22-31, 1991.

[5] Eggert, D. "Aspect Graphs of Solids of Revolution", Ph.D. Diss., CS&E Dept., USF, 1991. (See also Eggert, D., Bowyer, K. "Computing the perspective projection aspect graph of solids of revolution", *IEEE Trans. PAMI*, to appear, 1992.

[6] Eggert, D., Bowyer, K., Dyer, C., Christensen, H.,Goldgof, D. "The scale space aspect graph", Tech. Report, CS&E Dept., USF, 1992.

[7] Faugeras, O., Mundy, J., Ahuja, N., Dyer, C., Pentland, A., Jain, R., Ikeuchi, K., Bowyer, K. "Panel theme: Why aspect graphs are not (yet) practical for computer vision", *CVGIP: Image Understanding*, **55**, 212-218, 1992.

[8] Gigus, Z., Malik, J. "Efficiently computing and representing aspect graphs of polyhedral objects", *IEEE Trans. PAMI*, **13**, 542-551, 1991.

[9] Kender, J., Freudenstein, D. "What is a 'degenerate' view?", *ARPA IUW*, 589-598, 1987.

[10] Koenderink, J. "The structure of images", *Biological Cybernetics*, **50**, 363-370, 1984.

[11] Koenderink, J. Solid Shape, (MIT Press, Cambridge, Mass.), 1990.

[12] Kriegman, D., Ponce, J. "Computing exact aspect graphs of curved objects: Solids of revolution", *IJCV*, **5**, 119-135, 1990.

[13] Lindeberg, T., Eklundh, J. "Scale detection and region extraction from a scale-space primal sketch", *ICCV*, 416-426, 1990.

[14] Mokhtarian, F., Mackworth, A. "Scale based description and recognition of planar curves and two dimensional shapes", *IEEE Trans. PAMI*, **8**, 34-43, 1986.

[15] Plantinga, H., Dyer, C. "Visibility, occlusion and the aspect graph", *IJCV*, **5**, 137-160, 1990.

[16] Ponce, J., Kriegman, D. "Computing exact aspect graphs of curved objects: Parametric surfaces", *8th Nat. Conf. on AI*, 340-350, 1987.

[17] Rieger, J. "The geometry of view space of opaque objects bounded by smooth surfaces", *Artificial Intelligence*, **44**, 1-40, 1990.

[18] Sallam, M., Stewman, J., Bowyer, K. "Computing the visual potential of articulated assemblies", *ICCV*, 636-643, 1990.

[19] Sripradisvarakul, T., Jain, R. "Generating aspect graphs for curved objects", *IEEE Work. Interpretation of 3D Scenes*, 109-115, 1989.

[20] Stewman, J. H. "Viewer centered representations for polyhedral objects", Ph.D. Diss., CS&E Dept., USF, 1991. (See also Stewman, J., Bowyer, K. "Creating the perspective projection aspect graph of convex polyhedra", *ICCV*, 494-500, 1988).

[21] Wang, R., Freeman, H. "Object recognition based on characteristic view classes", *ICPR*, 8-12, 1990.

[22] Witkin, A. "Scale-space filtering", in From Pixels to Predicates, (Ablex Pub., Norwood, NJ), 5-19, 1986.