The Scale Space Aspect Graph[1]

David W. Eggert * Kevin W. Bowyer *

Charles R. Dyer † Henrik I. Christensen ‡ Dmitry B. Goldgof *

* Department of Computer Science and Engineering
University of South Florida
Tampa, Florida 33620
eggertd, kwb or goldgof@csee.usf.edu

† Department of Computer Science
University of Wisconsin
Madison, Wisconsin 53706
dyer@cs.wisc.edu

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

Index terms: aspect graph, scale space, viewpoint space partition, image resolution, dynamic shape, Gaussian smoothing.

Abstract

Currently the *aspect graph* is computed from the theoretical standpoint of perfect resolution in object shape, the viewpoint and the projected image. This means that the aspect graph may include details that an observer could never see in practice. Introducing the notion of scale into the aspect graph framework provides a mechanism for selecting a level of detail that is "large enough" to merit explicit representation. This effectively allows control over the number of nodes retained in the aspect graph. This paper introduces the concept of the *scale space aspect graph*, defines three different interpretations of the scale dimension, and presents a detailed example for a simple class of objects, with scale defined in terms of the spatial extent of features in the image.

---

[1]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.

1 Introduction

The aspect graph [14] is considered important because it provides a complete view–centered representation of an object. Considerable research has been performed in recent years on algorithms that compute the aspect graph and its related representations [4, 7, 8, 11, 12, 18, 25, 26, 28, 30, 31, 32, 33, 36, 37]. However, the practical utility of the aspect graph has been questioned. A recent panel discussion on the theme "Why aspect graphs are not (yet) practical for computer vision" was held at the 1991 *IEEE Workshop on Directions in Automated CAD–Based Vision* [10]. One issue raised by the panel is that aspect graph research has not included any notion of *scale*. (Actually, the lack of knowledge of how to make use of the concept of scale was acknowledged to be a problem for computer vision in general rather than just aspect graphs in particular.) The use of scale can be seen as a method of making the representation more realistic by relaxing certain assumptions implicit in the aspect graph approach.

To date, the aspect graph has effectively been computed only for the ideal cases of perfect resolution in object shape, the viewpoint and the projected image, leading to the following set of practical difficulties:

- A very small change in the detail of the 3–D shape of an object may drastically affect the number of visual events and nodes in the aspect graph.
- A node in the aspect graph may represent a view of the object that is seen from such a small cell of viewpoint space that it is extremely unlikely to ever be witnessed, and
- The views represented by two neighboring nodes in the aspect graph may differ only in some small detail that is indistinguishable in a real image.

Each of these factors contributes to a rather large overall size of the aspect graph representation. (For example, the worst–case node complexity is $O(N^9)$ for an N–faced polyhedron assuming a 3–D viewpoint space [25, 33].) By introducing the concept of scale one hopes to reduce this large set of theoretical aspects to a smaller set of the "most important" aspects.

This problem relates to another issue raised about the aspect graph, namely, the problem of *indexing* during object recognition. This is an important topic, and has been addressed in part by other researchers using concepts such as equivalence [34] and probability [2, 8, 13, 36, 37]. While the work presented here will be of aid in addressing this difficulty, it is not proposed as a solution by itself. As such, this issue will not be addressed further in this paper.

We begin by briefly reviewing the aspect graph representation and the scale space concept. Then a general definition of the *scale space aspect graph* is given, followed by discussions of three different interpretations of the scale parameter that address the above mentioned difficulties. The most promising of these interpretations, in which image resolution is a function of scale, is supported by a complete example showing the scale space aspect graph for the case of a nonconvex polygon in the plane. We conclude with some suggestions of avenues for future research on these interpretations.

2 A Brief Review of the Aspect Graph Concept

A general definition of the aspect graph is that it is a graph structure in which:

- There is a node for each *general view* of the object as seen from some maximal connected cell of *viewpoint space*, and
- There is an arc for each possible transition across the boundary between the cells of two neighboring general views, called an *accidental view* or a *visual event*,

A general viewpoint is defined as one from which an infinitesimal movement in each possible direction in viewpoint space results in a view that is equivalent to the original. In contrast, an accidental viewpoint is one for which there is at least one direction in which an infinitesimal movement results in a view that is different from the original. Under this definition the aspect graph is complete, in that it provides an enumeration of the fundamentally different views of an object, yet is minimal in the sense that the cells of general viewpoint are disjoint.

The various algorithms that have been developed may be classified using three properties; the object domain, the view representation and the model of viewpoint space. Object domains have evolved from polygons [12], to polyhedra [11, 25, 30, 32, 33, 36, 37], to solids of revolution [7, 8, 18], to piecewise–smooth objects [4, 26, 27, 28, 29, 31], to articulated assemblies [3]. Almost without exception, a view of the object is represented using a qualitative description of the line drawing, such as the *image structure graph (ISG)* [21]. The actual labeling of *contours* and *junctions* varies slightly among researchers. Distinctions between general and accidental views are usually based on isomorphism of the ISG. Lastly, two viewpoint space models are commonly used. The first is the 2-D *viewing sphere*, on which each point defines a viewing direction for *orthographic projection* [7, 11, 18, 25, 26, 27, 28, 29, 31]. The other is 3–D space, in which each point is the focal point for a *perspective projection* [4, 8, 25, 3, 30, 32, 33, 36, 37]. (For greater detail on categorization and comparison of these algorithms, see [9].) Figure 1 depicts the aspect graph of a simple block, as found using each viewpoint space model.

Most algorithms for computing the aspect graph follow the same basic high–level approach:

Step 1. Use the geometric definition of the object to enumerate the set of accidental views that occur (usually derived by examining individual object regions alone and in pairs and triplets), yielding the corresponding bifurcation surfaces in viewpoint space from which they are seen.

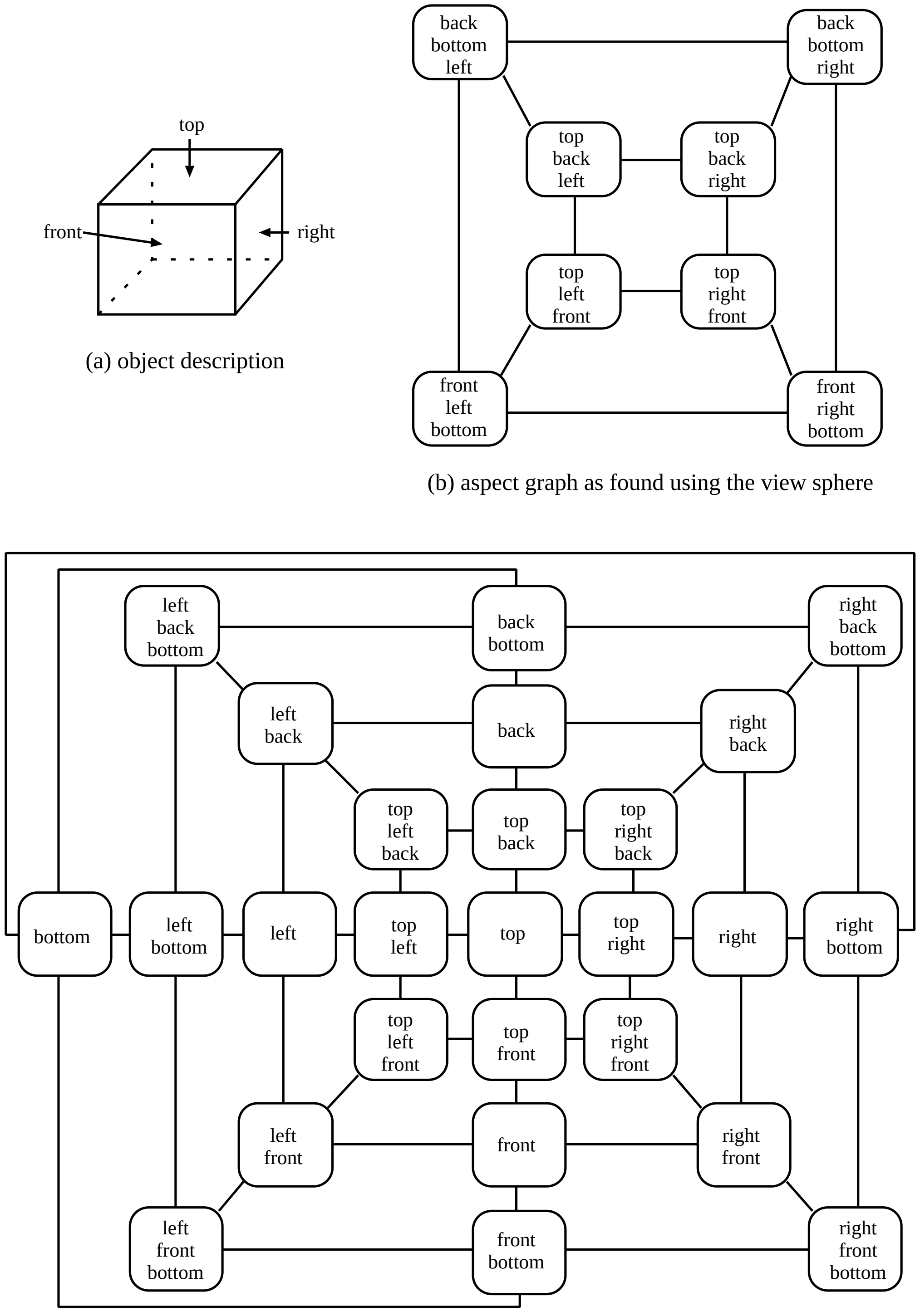
Step 2. Calculate the parcellation of viewpoint space defined by the meaningful portions of these accidental view surfaces.

Step 3. Traverse the parcellation of viewpoint space to build the aspect graph and attribute the nodes with descriptions of the representative views.

We concentrate on step one in this paper, by trying to develop a more proper set of visual event surfaces, as well as describe their nature in a more realistic manner.

3 A Brief Review of the Scale Space Concept

In its strictest sense, the phrase "scale space of X" is taken to mean a parameterized family of X in which the detail of features in X is monotonically decreasing with increasing scale and the qualitative



(a) object description

(b) aspect graph as found using the view sphere

(c) aspect graph as found using 3-D model of viewpoint space

Figure 1: Aspect graph of block using different models of viewpoint space.

features of X at a given scale can be traced back across all lower scales ("causality"). Research in this area was popularized by Witkin's scale space analysis of the inflections of a 1–D signal [39].

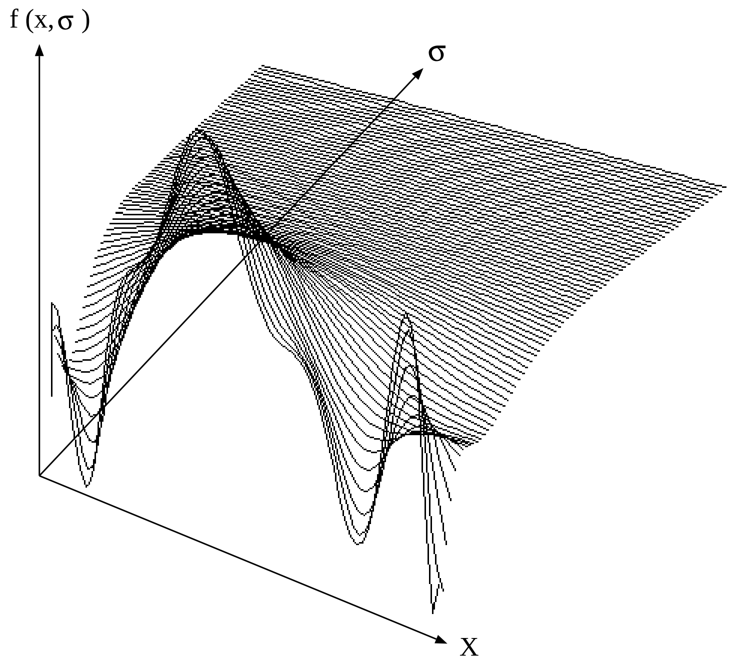


Figure 2: The scale space representation of an example 1–D signal.

Since that time the scale space concept has been applied to the curvature of 2–D curves [5, 22] and 3–D curves [23], the 2–D intensity map [1, 15, 20, 40] and 3–D object shape [16, 17]. In addition, a number of other researchers have described other "hierarchical" or "multi–resolution" representations, such as pyramids, that are similar to the scale space concept.

To better explain the approach we review the basics of Witkin's 1–D signal analysis. The qualitative (or symbolic) structure of a 1–D signal can be given in terms of the locations of its inflection points. The 2–D scale space of a 1–D signal is developed by introducing a second dimension, σ, representing the size of a Gaussian kernel used to smooth the original signal. In this one parameter family of 1–D signals a value of $\sigma = 0$ yields the original signal, while $\sigma = \infty$ means the signal is reduced to a flat line. The scale space of an example signal is shown in Figure 2. Here, a particular inflection that exists at one value of σ may be traced over increasing values of σ until it is eventually annihilated (merged with a neighboring inflection). In keeping with the monotonicity requirement, inflection points can only be annihilated as σ increases, never generated. Thus the value of scale at which an inflection ceases to exist is a measure of its strength or importance.

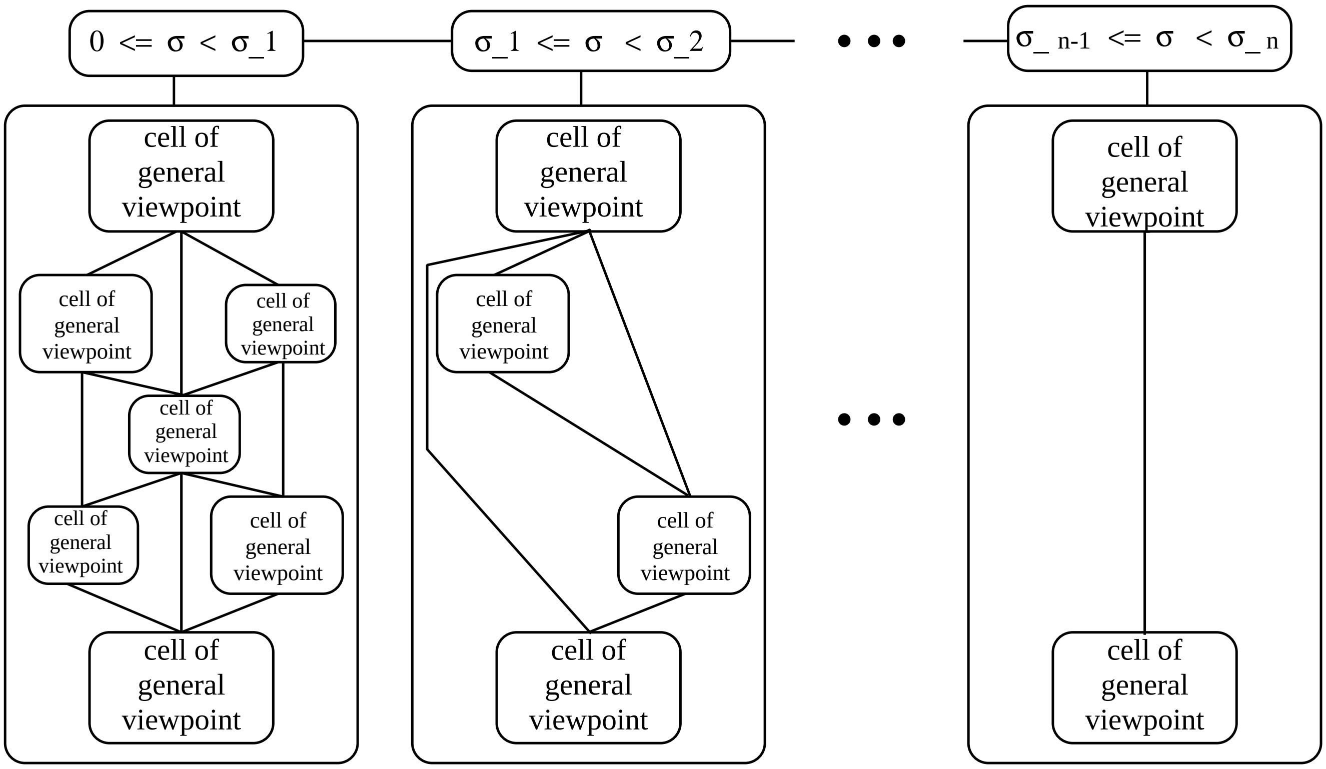
For 2–D and 3–D curves, similar analyses based on curvature and torsion of the signal create 3–D and 4–D spaces, respectively, in which the curve properties may be examined. When applying the concept to surfaces, such as 2–D intensity images and object surface descriptions, the change in topology as measured by some symbolic representation is examined. More will be said concerning these approaches in later sections on interpretations of scale.

4 The Scale Space Aspect Graph

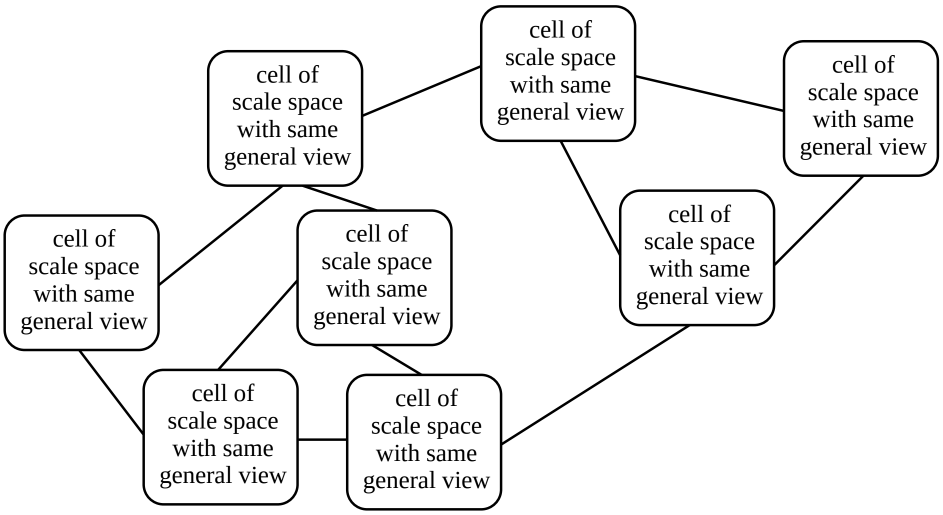
At this point, the high–level concept of a *scale space aspect graph* should be rather apparent. The essence of the scale space concept is that a one parameter family of instances of some entity be created, and that changes in the qualitative complexity of the entity are evident over changes in scale. The aspect graph is nothing more than a qualitative description of the underlying structure of the parcellation of viewpoint space into general views. Therefore it is appropriate to consider a parameterized family of these parcellations as the basis for the scale space aspect graph.

Given this basis, scale space is defined as a multi–dimensional space parameterized by both viewpoint location (direction) and scale value. This corresponds to a 4–D (x, y, z, σ) space when the perspective projection viewing model is assumed, and a 3–D (θ, ϕ, σ) space in the case of orthographic projection. Each visual event surface is now a function of viewpoint and scale. At $\sigma = 0$ the parcellation of the viewpoint space, and thus the aspect graph, is exactly as computed by some known algorithm. As σ increases, the parcellation of viewpoint space should deform in such a way that at certain discrete values of scale the aspect graph becomes simpler (has fewer nodes).

There are (at least) two alternative representations of the scale space aspect graph. In part (a) of Figure 3, it is depicted as an explicit sequence of conventional aspect graphs, each element representing a range of σ over which the aspect graph has a constant qualitative structure. This form bears considerable resemblance to the *visual potential* of articulated assemblies developed in [3]. In their representation separate instances of the aspect graph are recorded for varying articulation parameter values of an object (for instance, the angle of rotation on a hinge). At certain critical values of this configuration parameter the structure of the parcellation is changed.



(a) complete aspect graphs across discrete ranges of scale



(b) interconnected cells representing general views in scale space

Figure 3: Conceptual Depictions of the Scale Space Aspect Graph.

Here scale can be thought of in a similar manner. As an example, one might imagine the aspect graph of the block in Figure 1 appearing as the first in the series of Figure 3.a, and that later graphs may correspond to those after the disappearance of the one–face and two–face views, which are intuitively less important than the three–face corner views.

While the above form is more explicit and perhaps simpler conceptually, there is potentially a great deal of redundancy in the multiple instances of the aspect graph. Part (b) of Figure 3 depicts a more compact representation that is more directly analogous to the typical form of an aspect graph. Each node represents a "volume" of the scale space in which the same general view exists. Each arc again represents a visual event, but the underlying event surface is now parameterized by the scale dimension. This form corresponds most closely to the *asp* representation developed in [25]. In their representation features are represented in a higher–dimensional space according to their location on the image plane as a function of viewpoint position (direction). "Volumes" in this space represent particular feature configurations. The aspect graph is formed as the projection of these "volumes" into viewpoint space. This is essentially the conversion process used to generate a particular entry in the sequence of aspect graphs in the first representation in Figure 3.a from the second scale space form.

Other representations, such as extensions of the *interval tree* concept [20, 39] may exist depending on the interpretation of the scale parameter, which is the topic of the next section.

5 Interpretations of the Scale Parameter

Weaknesses in the aspect graph representation arise from certain implicit and explicit assumptions. Explicit assumptions, such as the use of the ISG as a view representation, will only be touched on lightly in the following sections. The effects of these choices alter the contents of the aspect graph, but the underlying approach is not changed. Instead we focus on the implicit assumptions, which have perhaps a more fundamental impact. These center around the qualitative nature of the representation, i.e., the lack of scale information. Three basic assumptions mentioned earlier are:

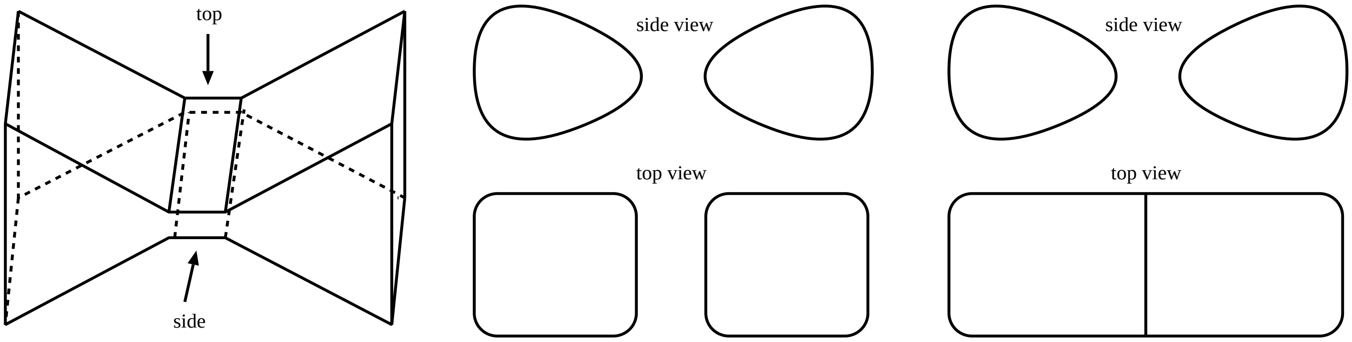
1. The object shape is known in exact detail. Visual events are generated through interactions of the various surface portions. Small bumps or indentations may generate several event surfaces, the visual changes of which might be considered insignificant. Also, certain events

may barely exist due to a fragile alignment. Thus small changes in the object shape can drastically alter the set of potential events, and thereby the parcellation of viewpoint space.

2. The camera is idealized as a point. This assumption manifests itself in the fact that each node in the aspect graph represents a view of equal significance. The underlying shape and size of the cell in the parcellation should have bearing on its importance. Since a camera does have a finite size, certain views are unlikely to ever be witnessed, because of their location in space and size compared to their neighbors.

3. There is infinite resolution in the projected image. In this case each feature in the ISG is accorded equal significance. This means that a given view may have a feature that is too small to detect from within its cell, and two views may differ by only such a feature and therefore be the same in practical terms. Also, each portion of the line drawing is distinguishable at an infinite viewing distance, a definite departure from reality. This leads to infinite–extent cells, when there should be a finite limit to meaningful viewing distance.

We have proceeded this far without assigning any particular meaning to the scale dimension and without saying how the scale parameter might be used to create a family of parcellations of the viewpoint space. This question does not have a unique answer, but the above visual phenomena will serve as areas of initial research, until a more complete understanding is achieved. Previous scale space representations have been applied to 1–D, 2–D and 3–D intensity functions by interpreting the scale parameter in terms of the solution to the diffusion equation [15] (or more specifically, as the variance of a Gaussian kernel used to blur the function). It has been proven that only under this interpretation will the qualitative features of the function disappear and not be created as the scale value is increased [15]. Unfortunately, the entities on which the aspect graph concept is based (such as visual events, projected line drawings, and 3–D shape) are not intensity functions. As such, it is hard to define what one means by "blurring" the parcellation of viewpoint space. Therefore, the requirement that the quantity of features monotonically decrease in size will likely have to be relaxed for most interpretations of scale.



(a) ''bowtie'' shaped object (b) views after ''dynamic shape'' smoothing (c) views after ''range'' smoothing

Figure 4: The effects of object smoothing on views of bowtie object.

5.1 Scale of features of object shape

The first interpretation involves examining the relative sizes of features of the object, by altering shape according to a scale parameter. This corresponds to smoothing the features of the object surface, such as knocking down small bumps and filling in small holes. One would hope that continual smoothing leads to effects that correspond to the loss of detail noticed while moving away from the object. Intuitively, one wants to smooth the object surface until a featureless blob is achieved, while examining the structure of the parcellation along the way. The question is how to perform the smoothing operation.

A technique which initially seems appealing is the "dynamic shape" concept [16, 17]. This is a form of 3–D volumetric blurring in which the surface is marked as the level set of the resulting distribution. The volume over which averaging occurs is a function of the scale parameter, while the determination of the surface level is a constant threshold. The effects of this technique, for a particular scale value, can be seen for the "bowtie" object in Figure 4. For a given level of smoothing the central portion of the object would cease to exist as the object splits in two. Subsequent views of it might appear as shown in Figure 4.b. While the view from the side might seem a logical consequence of smoothing the object, the view from the top does not. One would more likely expect to see the views in Figure 4.c, in which the smaller face no longer exists, but only because it has merged with neighboring faces. This is because the volumetric smoothing works upon *solid* shape, while that which is observable is *surface* shape.

Thus, since observed surface shape is really what we desire to smooth, and it is relative to the position of the viewpoint, we propose to smooth the object surface in a direction that is perpendicular to the viewing direction. Consider a 3–D shape $S(x, y, z)$ located at (0, 0, 0) of a coordinate system, and a distant viewpoint (x_v, y_v, z_v), from which an orthographic projection will be made. First, orient (rotate) the object so that the viewpoint is on the Z axis, i.e., $\hat{S} = R(x_v, y_v, z_v)S(x, y, z)$. Second, consider a certain visual scale, σ_0, defining a 2–D Gaussian kernel, which we convolve with each 2–D slice of $\hat{S}$ to get $\tilde{S}(x, y, z, x_v, y_v, z_v, \sigma_0)$ as:

$$\tilde{S} = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R(x_v, y_v, z_v)S(x', y', z')\frac{1}{2\pi\sigma_0^2}e^{-\frac{1}{2}\left[\frac{(x-x')^2+(y-y')^2}{\sigma_0^2}\right]}dx'dy'\right]dz'$$

The integrals can be exchanged to give:

$$\tilde{S} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} R(x_v, y_v, z_v)S(x', y', z')dz'\right]e^{-\frac{1}{2}\left[\frac{(x-x')^2+(y-y')^2}{\sigma_0^2}\right]}dx'dy'$$

$$\tilde{S} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} \hat{S}(x', y', z', x_v, y_v, z_v)dz'\right]e^{-\frac{1}{2}\left[\frac{(x-x')^2+(y-y')^2}{\sigma_0^2}\right]}dx'dy'$$

Now consider the interior integral. This is the projected thickness of the object along the viewing direction, which is independent of scale and could be precomputed (for each direction). Thus, by choosing a level set of $\tilde{S}$, we generate a visually similar surface at the σ_0 scale, which is effectively equivalent to smoothing the projected orthographic range image. This is similar to smoothing a projected image intensity function by the same 2–D Gaussian kernel for each viewpoint. (A topic to be mentioned again later.) A derivation assuming perspective projection leads to a similar result.

By using the above approach, the views in Figure 4.c could now be expected. The visual event surfaces generated by different portions of the object will now be highly dependent on viewpoint position for their existence. Two portions of the object that interact from one vantage may not have the same relative shape and position to do so from another. Eventually an interaction will no longer occur for any viewpoints as the smoothing level increases. Taken to the extreme, object shape should tend toward an ovoid with no visual event surfaces. Since the above computational process and the resulting representation are undoubtedly complex, we will not explore this interpretation any further in favor of another presented later that accounts for the same visual phenomena.

5.2 Scale of viewer relative to cell of viewpoint space

A second interpretation is to examine the relative sizes of cells in viewpoint space with respect to a finite–sized observer. In the past researchers have considered the probability of certain views based on relative cell volumes [2, 8, 13, 36, 37]. There are a few problems with this approach. First, volume of a cell, without regard to shape, is perhaps not a proper indication of the likelihood that a given view is seen. For two cells of the same volume, one blob–like and the other stick–like, if the observer is of any size at all then the blob–shaped cell becomes a more likely view than the stick–shaped cell. Second, the infinite–extent cells (those that exist for an orthographic model) are the only ones of consequence, since any finite–extent cell has effectively a zero probability. This is not proper, since in reality we cannot see the object from an infinite distance. It is possible to impose somewhat meaningful viewing limits based on environmental concerns, yielding a finite–sized viewpoint space, but we desire a more natural method of deciding these limits.

Thus we propose to generate a more extensive relation of viewer and cell. For this we relax the assumption that the viewer is idealized as a point. Instead, a finite–sized sphere, the radius of which is a function of scale, will model the volume of space in which light rays may be gathered and directed onto the image. (Imagine rotating the circular lens of a camera about the focal point to sweep out the volume of a sphere.) Any light impinging upon this sphere contributes to the composite image, as observable features from each point in the sphere are merged[2].

This interpretation can be explained in terms of changes in the parcellation as follows. For a given size sphere there will still be a region of viewpoints in a typical cell at which the sphere is fully contained within the cell. For those viewpoints from which the sphere overlaps the cell boundary, a composite view exists consisting of those views from the cell, the accidental boundary, and the neighboring cell. The nature of this composite view depends on the type of accidental boundary.

[2]Under the assumptions of *geometric optics* [38], a thin lens will focus all impinging light rays onto its optical axis. However, not all rays will be focused to the same point. Thus the view may be distorted or out of focus in regions corresponding to light rays that are far from the direction of the optical axis. Under the assumptions of *Gaussian optics*, all light rays are along or near the optical axis, resulting in a common focal point. In the following discussion the assumption of geometric optics is made, while ignoring any distortion. The constraints of Gaussian optics may then be used to impose intuitive limits on these results later.

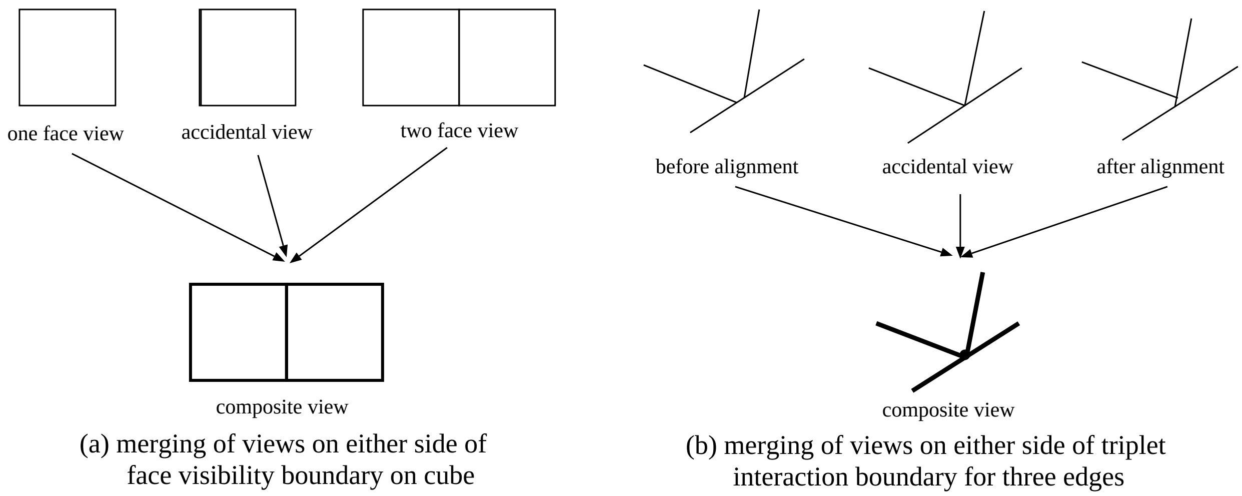


Figure 5: Blending of views across event surfaces for finite–sized viewer.

For example, consider a visual event surface that marks the visibility boundary for a face of the object. (See Figure 5.a for an example using a cube.) The composite view is the same as the one in which the new face is visible. This means the area of the two–face cell will impinge upon the region where the face is hidden, by a layer of thickness equal to the viewer sphere radius.

In other cases the composite view is really the accidental view itself. For example, consider an event surface representing a triple occlusion point in the image. (See Figure 5.b for an example of three edges on a polyhedron.) In the ideal case this alignment is only visible from the surface (a quadric surface for the example), but for a given size sphere, superimposing this view with those of the neighboring cells merely increases the apparent size of the triple point, as the nearby T junctions all merge together. Therefore, in this instance, the formerly accidental view can be seen from a volume of space and is now a "stable" view. This new cell of space is formed as the union of viewer sphere volumes centered at each point on the event surface.

Thus we can model the changes to the parcellation by extending the visual event surface positions by an amount equal to the radius of the current viewer sphere in one or two directions depending on the event type. When extensions occur in both directions a new general view is added to the aspect graph. In addition to these event surface extensions, the amount of available viewing space is reduced by a layer extended out from the object surface, since the camera can only get within a certain distance of the object.

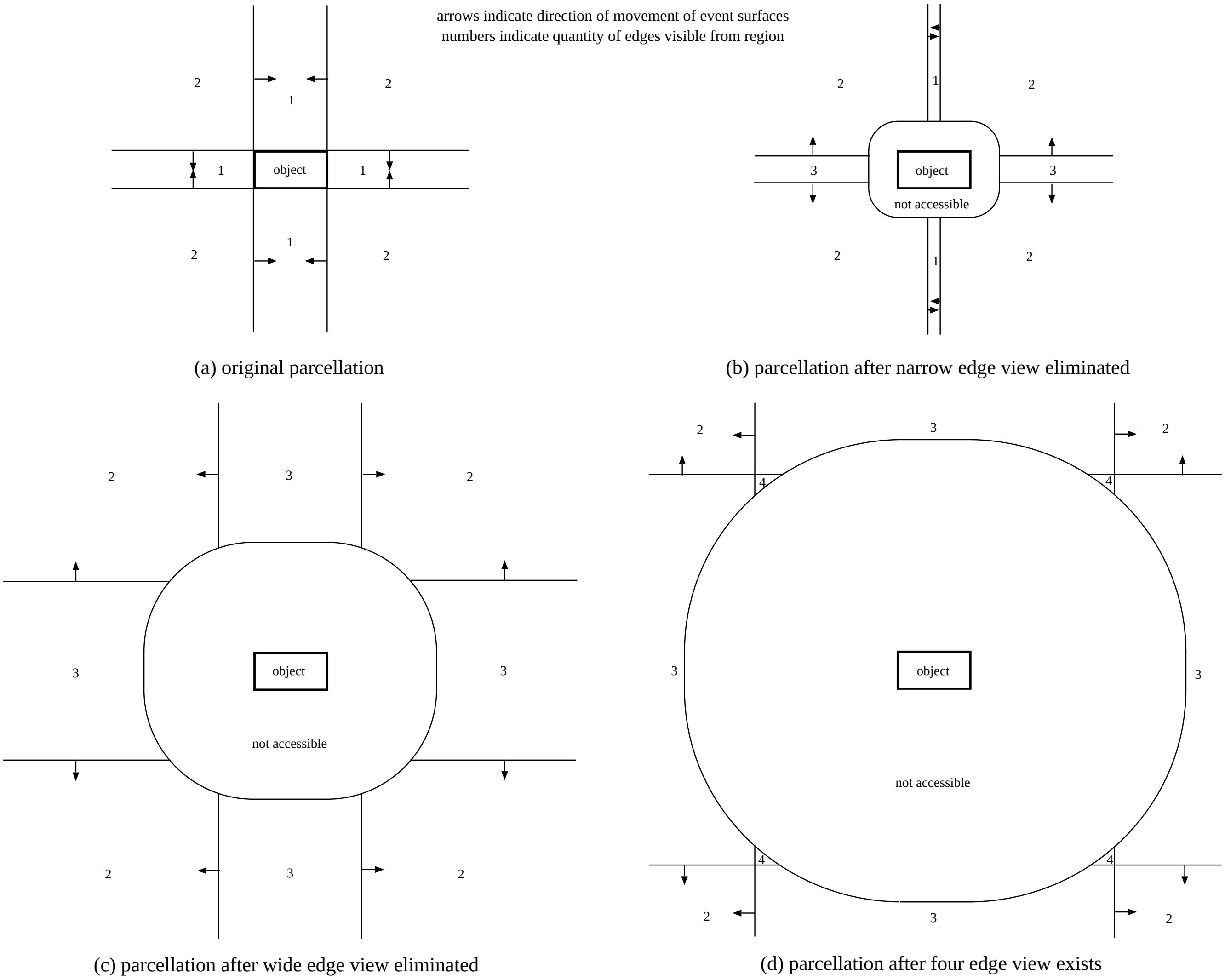


Figure 6: Changes to parcellation of rectangle based on expanding sphere radius.

As scale (sphere radius) changes certain cells are eliminated from the parcellation, while others come into existence. Those being shrunk on all sides will cease to exist at a scale that corresponds to the maximal size sphere at a point on the skeleton of the original cell as produced by a medial axis transform. Note that several such maxima may exist, resulting in different portions of the cell disappearing at different scales. (For example, an hour–glass–shaped cell would become two cells before it disappeared.) As these cells cease to exist other cells are generally created in the region of overlap of the expanding cells. In these areas a composition of the two expanding views is formed.

As a small example, we consider a rectangle in a plane as shown in Figure 6. For the plane, the viewer sphere becomes a circle. In part (a) the original parcellation is shown in which only one and two–edge views exist. The direction of expansion of each event line is indicated by the arrows. Initially the regions in which two edges are observed impinge on those regions seeing only one. At a scale equal to half the width, the cells corresponding to views of the narrower sides are eliminated and replaced by the overlap area in which that edge and the two longer edges are seen at once as shown in part (b). (This counterintuitive view, as with the four–edge view, would not exist under Gaussian optics constraints.) Notice also the expanding region about the polygon in which the camera no longer fits. In part (c) the one–edge views of the longer sides are removed in a similar manner at a scale equal to half the length. Finally, in part (d), the overlap of the regions viewing three edges (in which all four edges can potentially be seen) emerges from the region surrounding the object that the camera cannot enter. This occurs at a scale value of $l + w + \sqrt{2lw}$, where l and w are the dimensions of the polygon.

The importance of the various aspects generated by this process could be ranked according to the scale at which their cells disappear. But in this case the infinite ranging cells would be ranked equivalent. Perhaps a more accurate ranking is according to the "volume" of the scale space cell derived from the shape of the aspect's cell over all scales. However, this measure is still affected by the infinite–extent cells. Therefore we now discuss our final scale interpretation, which we believe is the most promising. It is supported by a detailed example.

5.3 Scale of features in the projected image

This last interpretation examines the relative sizes of the various features in the image. This corresponds to examining the view under varying levels of image resolution. The features in the image could be analyzed in at least two ways, according to their projected nature in the image intensity function, or in terms of their apparent size as a function of viewpoint position.

When analyzing an image intensity function more than one possibility exists. Given assumptions about object surface (say matte in texture) and light source placement (a point light source coincident with the viewpoint) an image intensity function can be constructed. Such a function can be subjected to 2–D Gaussian smoothing as a function of scale (much like the 1–D signal in

Figure 2), and the resulting features analyzed. In terms of the projected line drawing the features would be the edges detected in the smoothed image that are above a given magnitude threshold. Thus "weaker" edges would disappear first, meaning the strength of an edge defines its importance.

An alternative is to describe the image according to the surface topology of the intensity function, e.g., the "hills and dales" representation [15]. A study has been made of the changes that occur for a given image under Gaussian smoothing, such as the annihilation of saddle regions and merging of small hills or dales into one. Others are beginning to explore the types of visual events that exist for this view representation [35]. However, one difficulty is that much of the current theory that predicts changes in the ISG is not applicable under this representation. Also, since the image changes for each viewpoint, it is very difficult to conceive of the new nature of the visual event surfaces, as was noted for the object shape interpretations. Therefore we now concentrate on using scale as a measure of the size of features in the projected line drawing.

5.3.1 Measuring the size of features in an image

In current aspect graph analyses, a projected line drawing is constructed assuming an infinite resolution image plane. Also, the exact dimensions of the lines are lost as the drawing is abstracted to the image structure graph, where all arcs are treated equally. Both of these assumptions ease the computations that are made to calculate visual event boundaries, but limit the practical usefulness of the result. Thus it is proposed that the scale of the features in the view be reintroduced as a function of image resolution. Implicit in this method is the fact that we will account for changes in size due to viewing distance. These ideas are similar in nature to those used by researchers determining visibility constraints for automatic sensor placement [6].

First one must determine which features should be measured. In order to be measured, a feature must have some spatial extent in the image. This means that a junction, which occurs at a single point, should not be one of the features we now concentrate on. Alternatively, both edges (limbs) and object faces (portions of surface patches) generally have measurable extent in a view. How

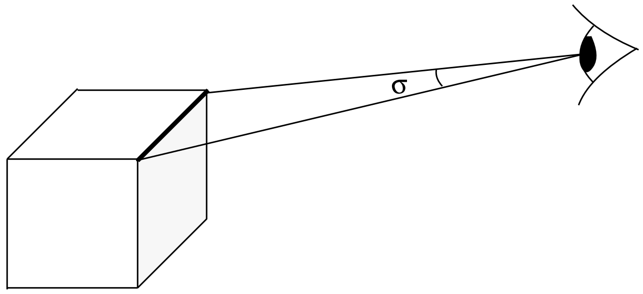


Figure 7: The angle of visual arc occupied by an edge feature.

does one quantify the size of a feature? It is not sufficient to measure the length of an edge or the area of a face on the object. It is the projection of these features that matters.

The first solution that comes to mind is to measure the dimensions of the features in an image coordinate system, the resolution of which is based on our scale parameter. Then the length along a projected edge, the perimeter or area of a face, or possibly the radius of a sphere that circumscribes the feature would be quantified in terms of a number of pixels. Unfortunately, this approach implicitly requires a more detailed camera model. Such parameters as focal distance, image plane size (field of view), the particular viewing direction and the viewing position must all be known. While such a sophisticated model would undoubtedly be more realistic, it is too complex to consider as a first step. (Of course, a multi–dimension scale space is imaginable, and may eventually prove necessary, but it is easier to first explore the concepts using a single–dimension scale space.)

An alternative measurement, used commonly by psychologists and biologists, is the angle of visual arc σ, or field of view, occupied by the feature. (See Figure 7.) Furthermore, if this is combined with the natural perspective viewing model [24] implicitly used by many aspect graph researchers, every feature's size can be described by only one parameter value in the range $0° - 360°$. Exactly how this value is measured depends on the feature. (Under this model points in the scene are projected onto a sphere of infinitesimal radius centered about the viewpoint, rather than onto an image plane at a fixed distance from the focal point, pointed in a particular direction. Each point's image coordinates are defined by the spherical angles of the ray from the focal point to it.)

For a straight edge, the distance between its projected endpoints will span a particular visual arc, as shown in Figure 7. For a curve, the maximum distance between any two projected points along its length indicates the visual extent. For a face, one must consider the maximum inscribable circle of the projected outline. For the square face of the cube seen in Figure 7, the size of the circle will decrease with distance and viewing directions off from the face normal. For a curved–surface patch the boundary curve includes both surface discontinuities and apparent contours, and thus it is more difficult to compute this maximal circle.

This scale model accounts for the quantitative size of a feature in both an absolute and relative sense. In the absolute sense, the actual effects of viewing distance are considered. Given a particular visual arc threshold, a feature cannot be distinguished after moving a certain distance away, as one would realistically expect. In a relative sense, a feature can be considered as insignificant if it is much smaller than the neighboring features, even when clearly visible. Again this is a function of viewpoint, as two features may change their relative size as the viewpoint moves. Unfortunately, one drawback of this model is that it does not directly correspond to Gaussian blurring of the features. Therefore, there may not be a strictly monotonic decrease in the size of the scale space aspect graph as a function of scale. This will become apparent in our later case study.

So how is the above interpretation used? It should be obvious that image resolution can be defined in terms of degree of visual arc. Pixel size in the image directly corresponds to the minimum visual arc necessary to distinguish a feature, while at $0°$ the camera has infinite resolution. At an angle of $360°$ there is only a single pixel in the image and everything projects to it. For a given scale, any feature mapping to a size smaller than one pixel is considered as not observable. Thus, for a given resolution, two views that were previously different may now appear equivalent. More exactly, the image resolution has a direct effect on the shape of the visual event boundaries. In the ideal case ($\sigma = 0°$), which corresponds to how aspect graphs have been computed in prior work, a visual event that denotes the alignment of some set of features is seen from a ruled surface. Once the scale becomes nonzero, features do not have to align exactly. As long as the points in question map to within one pixel of each other in the image, they can be considered as aligned. Thus the restriction of a ruled surface no longer holds, and the distortion from the ideal surface is increased as the scale parameter (visual arc required to distinguish the features) grows.

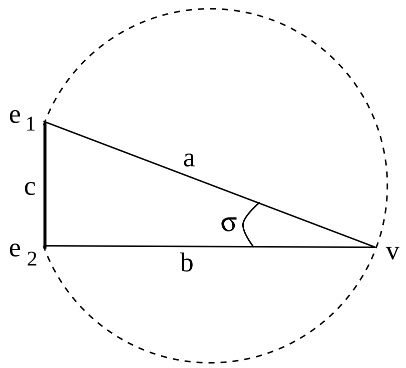


Figure 8: Visual event curve defined for an edge.

5.3.2 Case study : A nonconvex polygon in a plane

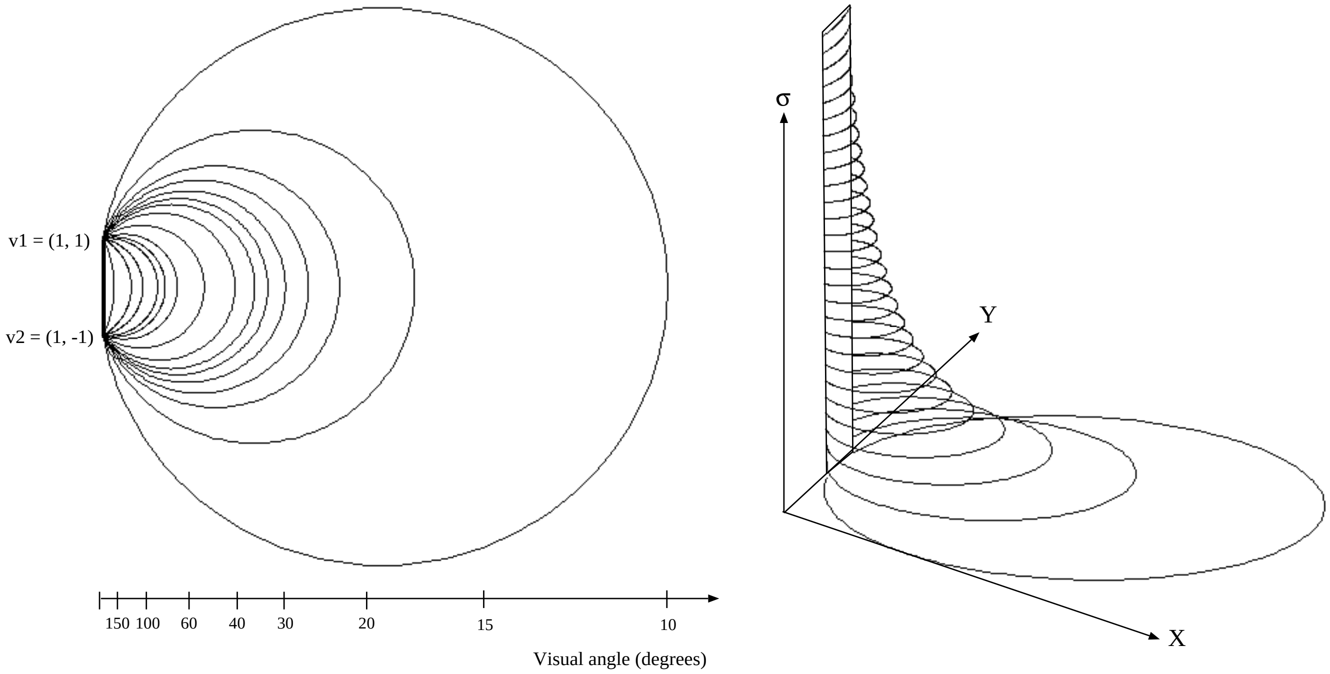
Since the interpretation of scale as spatial extent of features in the image seems to capture an important phenomenon in visual perception, and since this interpretation seems more computationally tractable in comparison to the others, we examine it more fully by presenting a case analysis of nonconvex polygons in a plane. This domain is chosen both for ease of presentation and computation. Extensions to broader domains are discussed in the conclusions.

For this problem there exists a 3–D scale space (x, y, σ). The only feature of nonzero spatial extent is an edge. An edge in isolation is considered visible when it is larger than the scale of resolution of the image. The degree of visual arc subtended by the projected line segment connecting the endpoints measures the feature. In order to construct the equation of the visual event surface denoting edge visibility, first assume the endpoints of the edge are given by $e_1 = (x_1, y_1)$ and $e_2 = (x_2, y_2)$. Then given a viewpoint defined by $v = (x_v, y_v)$, the visibility curve can be defined as the set of viewpoints such that the angle formed by the lines drawn from the viewpoint to each endpoint is constant. This is merely the set of vertex points of a set of triangles formed by the endpoints and the viewpoint, each triangle having the edge as one side and the chosen visual arc angle opposite it. See Figure 8. Using the law of cosines this curve's equation is:

$$c^2 - a^2 - b^2 + 2\,a\,b\,\cos\sigma \;=\; 0$$

where the lengths of the sides of the triangle shown in Figure 8 are given by:

$$c = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \qquad b = \sqrt{(x_2 - x_v)^2 + (y_2 - y_v)^2} \qquad a = \sqrt{(x_1 - x_v)^2 + (y_1 - y_v)^2}$$



(a) instances of visibility boundary curves in viewpoint space

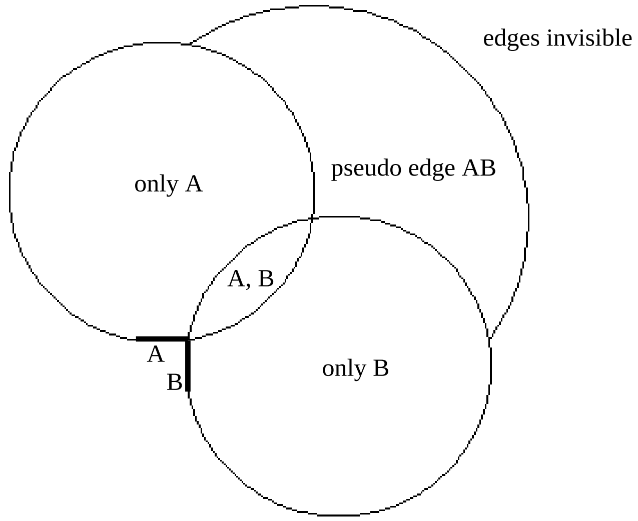
(b) visibility boundary surface for edge in scale space

Figure 9: Visibility ranges of an edge for varying visual arc angles.

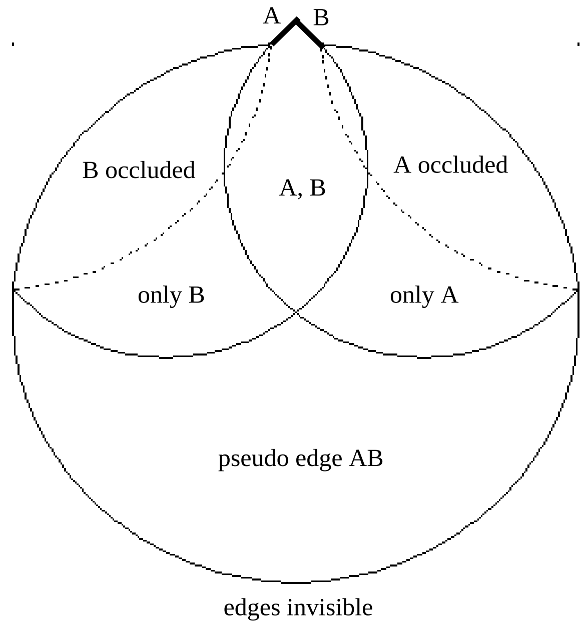
and σ is the visual arc angle. This curve is merely a circular arc segment [19], the size and center of which vary with σ for a particular edge as shown in Figure 9.a. The shape of the surface in scale space can be seen in Figure 9.b.

The nature of this surface is such that in the ideal case of infinite camera resolution ($\sigma = 0°$) the curve in the plane degenerates to the line containing the edge. As aspect graph theory predicts, the edge is visible from the infinite half–space on the proper side of this line. At any finite resolution ($\sigma > 0°$) the curve bounds a finite region of the plane, outside of which the projection of the edge is too small to distinguish. This distortion of the original line into the circular curve is an instance of relaxing the ruled event surface restriction mentioned earlier.

Having defined the visibility of an edge, we must next consider the visual interaction of a pair of edges, A and B, that are joined at a common vertex as shown in Figure 10. This particular interaction varies slightly based on whether the edges form a convex or concave angle. For a convex–angled pair (see Figure 10.a), a particular visual arc angle will define a visibility region for each



(a) joint visibility of convex-angled edge pair



(b) joint visibility of concave-angled edge pair

Figure 10: Joint visibility ranges of edge pairs.

edge bounded by the above event curve. If the given visual angle is not large there may be an area of overlap of these two regions in which both edges are visible and seen joined to one another. The intersection point between the curves falls along the line that bisects the angle between the edges.

Now consider a "pseudo edge" formed by chaining together the two edges. This pseudo edge has a wider defining angle than either of the others for certain vantage points, given by the distance between the noncommon endpoints. Thus it can be expected that from certain viewpoints neither of the edges is individually distinguishable, but the combined length of the pseudo edge is. This pseudo edge is seen as a blending of the two underlying edges, as the common vertex becomes indistinguishable. The meaningful portion of the curve representing the pseudo edge's visibility ends at intersections with the curves for each edge that occur along the lines containing the edges. These bounds are dictated by the fact that the pseudo edge only exists if both edges are potentially visible, but neither is distinguishable. This concept of a feature that is a blending of others is unique to the scale space approach, since an assumption of infinite resolution allows every edge to be perceived regardless of distance.

For a concave–angled edge pair a somewhat similar situation occurs, as given in Figure 10.b. Again there are three basic event curves defining regions from which either (1) a single edge is visible, (2) both edges are visible, or (3) the pseudo edge is visible. Here, the visibility curve encompasses the individual curves because it is not possible to see one of the edges alone if you cannot see the pair. Again the intersection points occur along the line extensions of the edges.

The final kind of interaction involves occlusion. Just as a change in visibility of an edge is marked by its endpoints appearing to coincide in the image, the limits of occlusion occur when the endpoints of two separate edges appear to coincide. Thus the same circular equation can be used to define this event curve. In Figure 10.b the merging of the noncommon endpoints marks the occlusion of one edge by another, and defines the same event curve as before. The change from occlusion limit to pseudo edge visibility limit occurs at the point of intersection with the individual edge event curve. The visibility limits of other neighboring edges that are the true occluders, shown by dotted curves in the figure, form the other limiting boundaries to the occlusion regions.

In order to consider an entire polygon, the set of event curves and their meaningful ranges is derived from the set of real and pseudo edges. Each pseudo edge can be modeled as derived from a convex or concave corner of the object, regardless of how many edges are actually involved. It some cases this corner is only implied as the junction of two smaller pseudo edges. Examples of this will be seen shortly.

We now examine how the scale space theory can be applied to compute the structure of the scale space aspect graph. In principle, the equations of each event surface can be used to calculate a data structure such as the *geometric incidence lattice*, used in previous aspect graph algorithms [33], to represent the subdivision of the 3–D scale space. However, since these surfaces are not ruled, the necessary intersection calculations can be very difficult. Instead, since the planar parcellation for a given scale can be more directly computed, we will explore the other form of the representation, namely, examining the structure of the aspect graph over the range of scale. For this we define a set of *scale events* that fall into three basic categories which will be illustrated shortly:

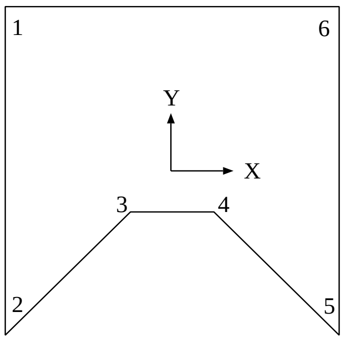
1. *Begin or end overlap of two curves.* Two curves may initially have their meaningful portions touch at a single point as they start to overlap, or their meaningful portions touch at a single point as they cease to overlap. If two event curves are involved, either a new region is created or one is deleted to represent the overlap area. If an event curve and an edge are involved, generally it is only the boundary of an existing cell that is altered.

2. *Triple point.* The meaningful portions of three curves intersect at a single point in the parcellation. If three event curves are involved, both before and after the critical scale these curves bound a region. The region that existed before will no longer be in the parcellation and it will be replaced by the new region. If two event curves and an edge interact, either a new region will be created or one will be removed. This is because the object exists on one side of the edge, and only one of the regions is meaningful. If one event curve and two edges are involved, a region may again be created or removed. In general this creation or deletion coincides with the insertion or removal of a meaningful segment of the event curve.

3. *Curve coincidence.* The meaningful portions of two curves coincide along their length in the parcellation. In the case of two event curves, before the critical scale they intersect in the parcellation and are part of the boundaries of a set of regions. At the critical scale the definitions of the two curves are the same, and any regions that existed between them no longer do. Beyond the critical scale the curves again intersect one another and bound a new set of regions. When an event curve and an edge are involved, the region between them is deleted from the parcellation. In addition the event curve is removed, since it will have passed into the object where it has no meaning.

It is of course possible for combinations of the above events to occur at the same scale. As such it is not necessary to consider larger groups of curves. For instance, if four curves were to intersect at a point, each of the four triplets of curves would be intersecting as well.

Finally, a note on the size complexity of this representation. In the ideal case, a polygon with N edges will generate $O(N^2)$ visual event lines, which when intersected in the plane will bound $O(N^4)$



(a) nonconvex polygon

Vertex	Coordinates (x,y)
1	(-1, 1)
2	(-1, -1)
3	(-0.25, -0.25)
4	(0.25, -0.25)
5	(1, -1)
6	(1, 1)

(b) vertex coordinates

Real edges	Vertex endpoints
A	6-1
B	5-6
C	4-5
D	3-4
E	2-3
F	1-2

(c) actual edge definitions

Pseudo edges	Vertex chain	Corner type	Subedges
G	2-3-4	concave	E D
H	3-4-5	concave	D C
I	2-3-4-5	concave	G H, G C, E H
J	1-2-3	convex	F E
K	4-5-6	convex	C B
L	5-6-1	convex	B A
M	6-1-2	convex	A F
N	1-2-3-4	convex	F G
O	1-2-3-4-5	convex	F I
P	3-4-5-6	convex	H B
Q	2-3-4-5-6	convex	I B

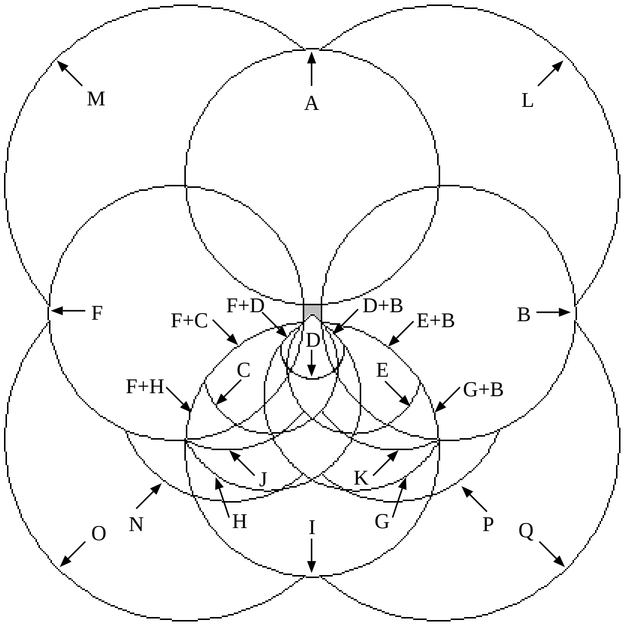
(d) pseudo edge definitions

Figure 11: Definition of example nonconvex polygon.

cells that correspond to aspects. Instead, now each of the N real edges generates an event curve. In addition, there are $O(N^2)$ pseudo edges, since there are N possible chains of edges of length 1 to N that comprise them. Each of these $O(N^2)$ pseudo edges also generates an event curve. In the 3–D scale space, the number of cells is bounded by the number of scale events that create or destroy them in the 2–D parcellation. Given S curves, the triplet scale event may generate $O(S^3)$ cells. Thus with $O(N^2)$ event curves, the number of cells in the scale space aspect graph is bounded by $O(N^6)$ in the worst case.

5.3.2.1 An example. Figure 11 defines a six-sided nonconvex polygon. In part (c) a list of the real edges is presented, while in part (d) the list of pseudo edges is defined. Each pseudo edge is composed of the string of actual edges it represents. Also given is its interpretation as a convex or concave corner, and the pair of edges that comprise the corner. Notice that pseudo edge I can be considered as a concave corner for three different edge pairs in the concavity.

There are seventeen event curves that serve as region boundaries somewhere in the scale space aspect graph. These are labeled in Figure 12 for a sample parcellation at $\sigma = 5°$. The curves A through F represent visibility limits of the actual edges, while J through Q represent the visibility limits of the pseudo edges across convex corners. The pseudo edges across concave corners, G



single letters indicate visibility boundary, + indicates occlusion boundary

Figure 12: Labeling of meaningful portions of event curves for polygon.

through I, generate their respective visibility limits, as well as six curve segments that mark the occlusion limits of the edges in the concavity by the side edges. Notice that the curve generated by vertices two and five serves as the termination limit for the majority of the event curves since the elements of the concavity are not visible if the two vertices appear merged.

If one constructs the parcellation of the 3–D scale space using a sampling of the scale parameter similar to that used for articulated assemblies [3], a total of fifty aspects is determined. These are listed in Table 1 and described according to the range of scale over which they exist and the view that is seen in terms of the edges and the vertices between them. (Due to the relative sizes and positions of the edges, pseudo edges J and K are never visible in isolation, and so they do not appear in the table.) The vertices have been broken down into four categories; T junctions, real vertices on the object, the implied vertex that is the merging of the endpoints of an indistinguishable edge, and an implied vertex in the overlap range of two pseudo edges. The structure of the scale space aspect graph, which corresponds to the parcellation of scale space, is given in Figure 13.

It is perhaps more instructive to examine the second form of the scale space aspect graph, which

Table 1: Index of polygon aspects with visual arc existence ranges.

ASPECT	EDGES IN VIEW	VISUAL ANGLE RANGE (degrees)	ASPECT	EDGES IN VIEW	VISUAL ANGLE RANGE (degrees)
1	A	0* - 180	26	F + H	0 - 45
2	B	0* - 180	27	G + B	0 - 45
3	C	45 - 180	28	G • C	0 - 127.98
4	D	36.87 - 180	29	G × H	0 - 216.87
5	E	45 - 180	30	G • K	0 - 45
6	F	0* - 180	31	G × P	0 - 45
7	G	0 - 225	32	I × P	0 - 59.04
8	H	0 - 225	33	J − C	0 - 30.96
9	I	0 - 255.96	34	J • H	0 - 45
10	L	0 - 90	35	N • C	6.1 - 24.75
11	M	0 - 90	36	N × H	0 - 45
12	N	0 - 59.04	37	N × I	0 - 59.04
13	O	0 - 90	38	E − C • B	0 - 22.5
14	P	0 - 59.04	39	E • D + B	0* - 14.04
15	Q	0 - 90	40	E • D • C	0* - 75.4
16	A • F	0* - 45	41	E • H • B	0 - 29.5
17	B • A	0* - 45	42	E • H × K	0 - 30.96
18	D • C	30.96 - 112.5	43	F + D • C	0* - 14.04
19	E + B	0* - 30.96	44	F • E − C	0 - 22.5
20	E − C	0 - 77.65	45	F • G • C	0 - 29.5
21	E • D	30.96 - 112.5	46	G × H × K	0 - 6.1, 24.75 - 45
22	E • H	0 - 127.98	47	J × G • C	0 - 30.96
23	E − K	0 - 30.96	48	J × G × H	0 - 6.1, 24.75 - 45
24	E • P	6.1 - 24.75	49	E • D • C • B	0* - 14.04
25	F + C	0* - 30.96	50	F • E • D • C	0* - 14.04

* indicates aspect exists for visual arc angle of 0°
\+ indicates edges joined by T junction vertex
− indicates edges joined by implied vertex that is collapsed edge
• indicates edges joined by actual vertex on object
× indicates edges joined by implied vertex due to overlap of edges

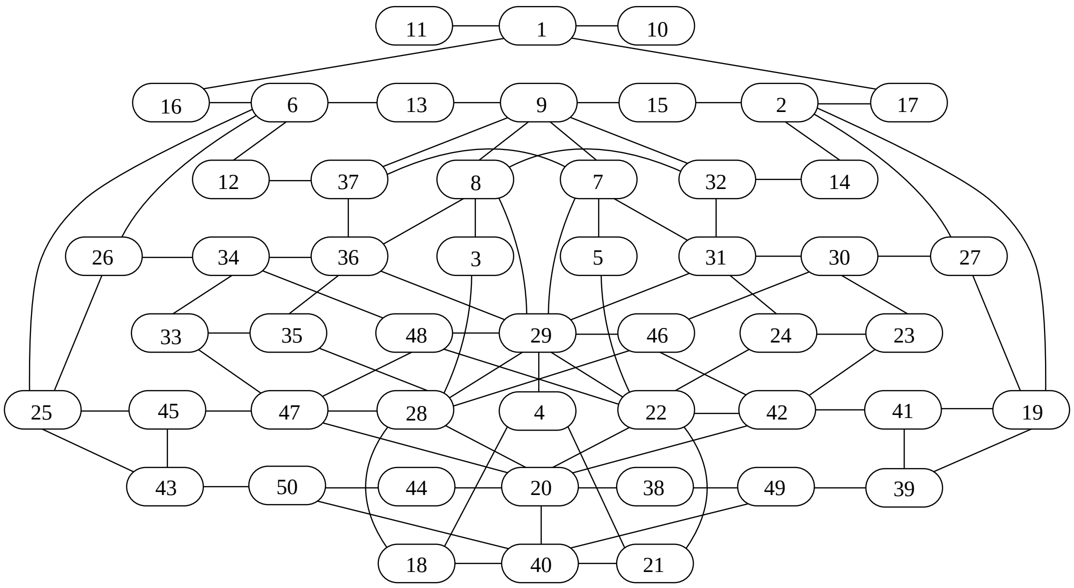


Figure 13: Scale space aspect graph of example polygon.

Table 2: List of changes in parcellation due to various scale events.

EVENT ANGLE	EVENT DESCRIPTION	INSERTED ASPECTS	REMOVED ASPECTS	TOTAL ASPECTS
0.0°	infinite lines become finite curves (cA - cC, cE - cH), new finite curves (cD, cI - cQ)	7-15, 20, 22, 23, 26-34, 36-38, 41, 42, 44-48	–	43
6.1°	triple points (cJ, cG, cC), (cK, cH, cE)	35, 24	48, 46	43
14.04°	triple points (cD, cE, cG), (cD, cE, cF), (cD, cC, cH), (cD, cC, cB) begin overlap (cD, eE), (cD, eC)	– 20^+, 40^+	43, 50, 39, 49 –	39
22.5°	end overlap (cE, cF), (cC, cB)	–	44, 38	37
24.75°	triple points (cJ, cG, cC), (cK, cH, cE)	48, 46	35, 24	37
29.5°	end overlap (cF, cG), (cB, cH)	–	45, 41	35
30.96°	triple points (cC, cF, cI), (cC, cF, cG), (cE, cB, cI), (cE, cB, cH) triple points (cC, eD, eE), (cE, eC, eD) coincide (cC, cE)	– 21, 18 22, 28	25, 33, 19, 23 – 22, 28, 42, 47	31
35.26°	triple points (cC, cD, eE), (cE, cD, eC)	22^+, 28^+	–	31
36.87°	triple point (cC, cE, eD) end overlap (cC, eE), (cE, eC)	4 –	– 22^+, 28^+	32
45.0°	triple points (cH, cE, eE), (cG, cC, eC) end overlap (cD, eE), (cD, eC), (cE, eD), (cC, eD) end overlap (cA, cF), (cA, cB), (cF, cI), (cB, cI) triple point (cJ, cF, cG), (cJ, cE, cG), (cJ, eE, eF) triple point (cK, cB, cH), (cK, cC, cH), (cK, eB, eC) coincide (cG, cH)	5, 3 – – – – 7, 8	– 21^+, 22^+, 18^+, 28^+, 4^+ 16, 17, 26, 27 34,48 30, 46 7, 8, 31, 36	24
59.04°	triple point (cN, cF, cI), (cN, cG, cI), (cN, eF, eG) triple point (cP, cB, cI), (cP, cH, cI), (cP, eB, eH)	– –	12, 37 14, 32	20
75.4°	triple point (cC, cD, cE)	29^+	40	19
75.96°	end overlap (cH, eE), (cG, eC)	–	5^+, 22^+, 3^+, 28^+	19
77.65°	end overlap (cC, cE)	–	20, 29^+	18
90.0°	triple points (cL, cA, cB), (cL, eA, eB), (cM, cA, cF), (cM, eA, eF) triple points (cO, cF, cI), (cO, eF, eI), (cQ, cB, cI), (cQ, eB, eI)	– –	10, 11 13, 15	14
112.5°	end overlap (cD, cE), (cD, cC)	–	21, 18	12
127.98°	end overlap (cE, cH), (cC, cG)	–	22, 28	10
180.0°	coincide (cA, eA), (cB, eB), (cC, eC), (cD, eD), (cE, eE), (cF, eF)	–	1, 2, 3, 4, 5, 6	4
194.04°	begin overlap (cG, eE), (cH, eC)	7^+, 8^+, 9^+	–	4
210.96°	begin overlap (cG, eD), (cH, eD)	7^+, 8^+	–	4
216.87°	triple point (cG, cH, eD)	–	29	3
225.0°	begin overlap (cI, eE), (cI, eC) triple points (cG, eD, eE), (cH, eD, eC)	9^+ –	– 7, 8	1
253.74°	begin overlap (cI, eD)	9^+	–	1
255.96°	triple points (cI, eD, eE), (cI, eD, eC)	–	9	0

+ indicates existence of aspect not changed, but boundary description is altered

cX indicates event curve X, eX indicates actual edge X

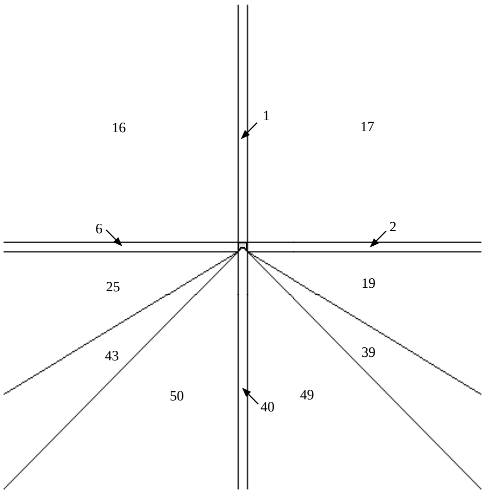
is the series of distinct parcellations existing over the changes in scale. There are 23 different angles at which scale events occur that modify the parcellation, excluding the initial transformation at $\sigma = 0°$. At several of the critical angles multiple events occur. Each of these events is described in Table 2 according to the interacting curves and edges, the newly created aspects, those aspects removed from the parcellation, and the resulting number of aspects after the events have occurred.

The representative parcellations at intermediate values of σ are drawn and labeled in Figure 14. We now discuss some of the basic scale events that occur for this object. Scale events are indicated by angle and the two parcellations of Figure 14 they bridge.

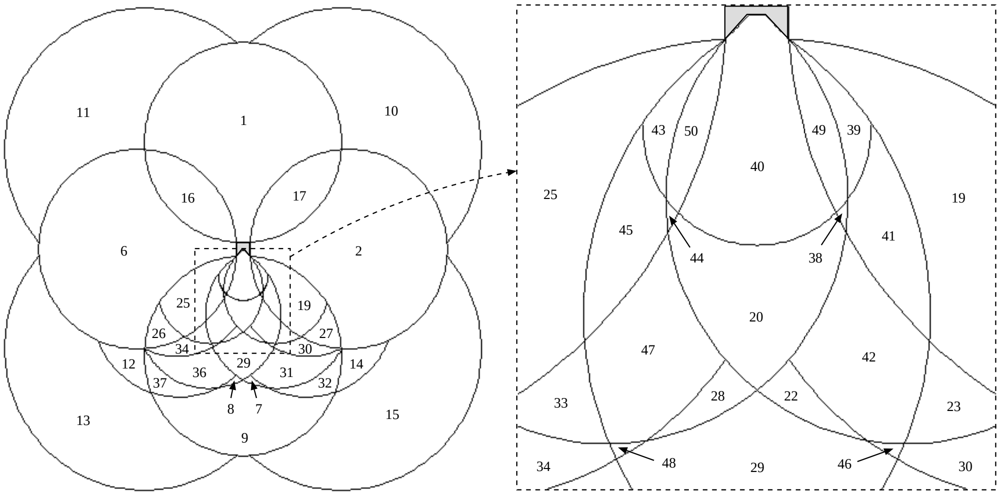
- *triple point between three event curves.* At angles $6.1°$ ((b) $\rightarrow$ (c)) and $24.75°$ ((e) $\rightarrow$ (f)) typical examples of this event occur, with one cell being created and another deleted. What is interesting about these cases is that they are exact opposites of one another. Cells 24 and 35 exist over the finite scale range between the events, while cells 46 and 48 exist before and after. The scale event angles were found using a search for the scale angle value at which the intersection points between curve pairs converged.

- *triple point between two event curves and one edge.* The types of events for this group vary. At angle $36.87°$ ((i) $\rightarrow$ (j)) cell 4 is created, while at $216.87°$ ((u) $\rightarrow$ (v)) the opposite occurs as cell 29 disappears. At $35.26°$ ((h) $\rightarrow$ (i)) an interesting event occurs in which a new cell in the parcellation appears (cell 22), but it is merely another disjoint area in which that aspect can be seen. These two areas are later merged as the result of another event. The angles at which these events occur can be calculated by solving for the location of a point on the participating edge from which the other two visible edges are seen to be the same size.

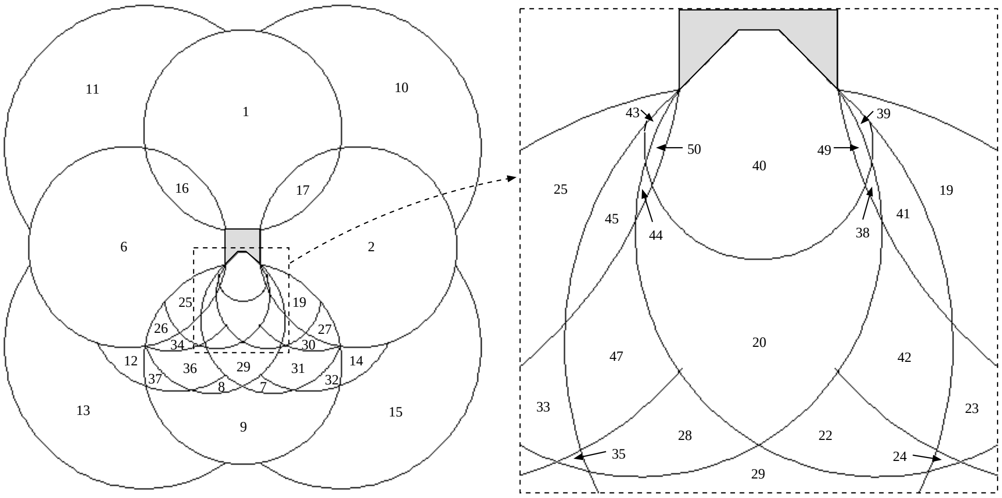
- *triple point between one event curve and two edges.* The creation of a cell can occur when a curve passes over a vertex and emerges from the object, as for cell 21 at $30.96°$ ((g) $\rightarrow$ (h)). This angle is found by solving for σ after substituting the vertex position into the event curve equation. It is also possible for cells to disappear as event curves are absorbed into the object after passing over the vertex. In these cases the event curves that represent a pseudo edge containing that vertex no longer exist. For a convex corner this coincides with a triple point of three event curves, as occurs at $90°$ ((o) $\rightarrow$ (p)) for pseudo edge AF. For a concave corner only the curve and edges are involved, as happens for pseudo edge G at $225°$ ((v) $\rightarrow$ (w)). These events occur at angles equal to that between the participating edges.



(a) visual angle = 0.0

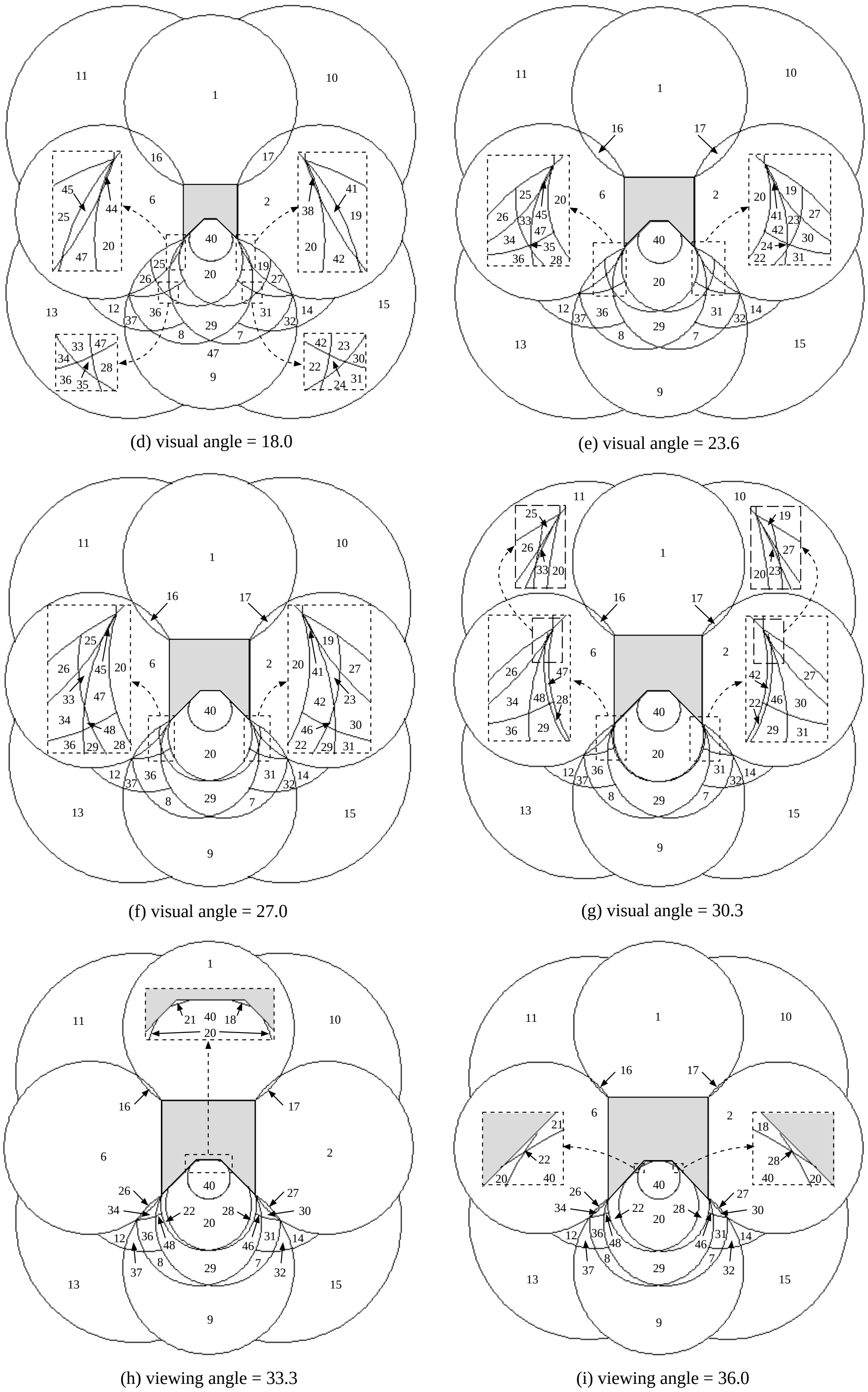


(b) visual angle = 4.0



(c) visual angle = 10.0

Figure 14: Representative aspect graphs over changes in scale.



(d) visual angle = 18.0

(e) visual angle = 23.6

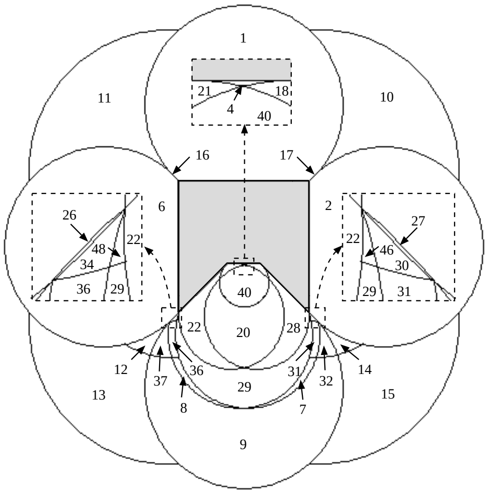
(f) visual angle = 27.0

(g) visual angle = 30.3

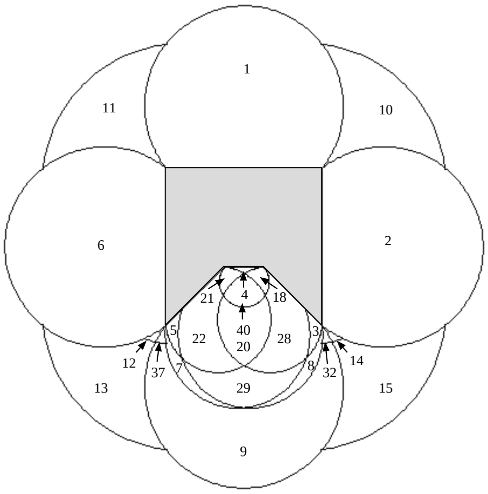
(h) viewing angle = 33.3

(i) viewing angle = 36.0

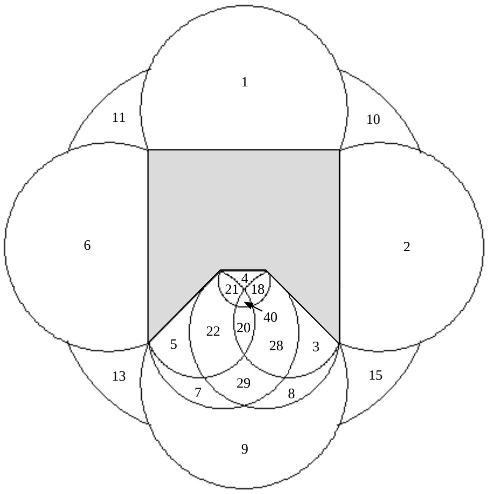
Figure 14: continued.



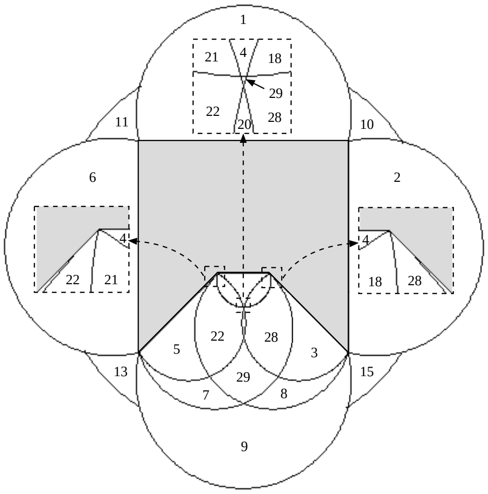
(j) viewing angle = 41.0



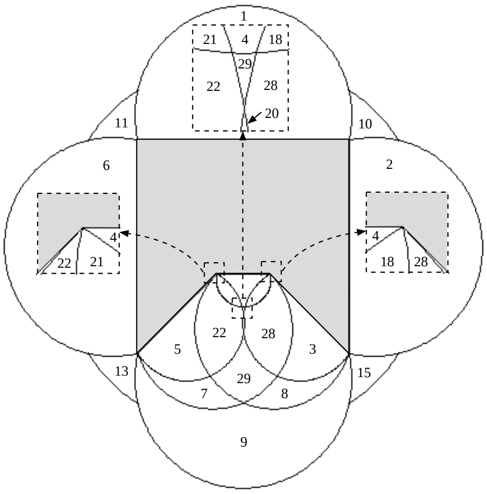
(k) viewing angle = 52.0



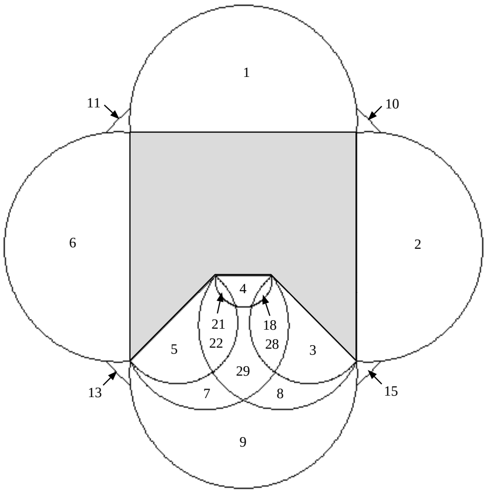
(l) viewing angle = 67.0



(m) viewing angle 75.7

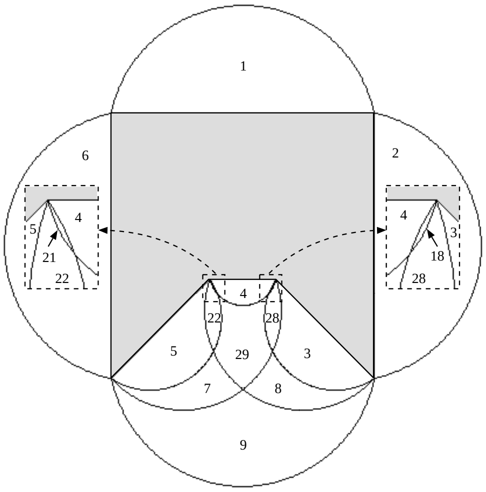


(n) viewing angle = 76.8

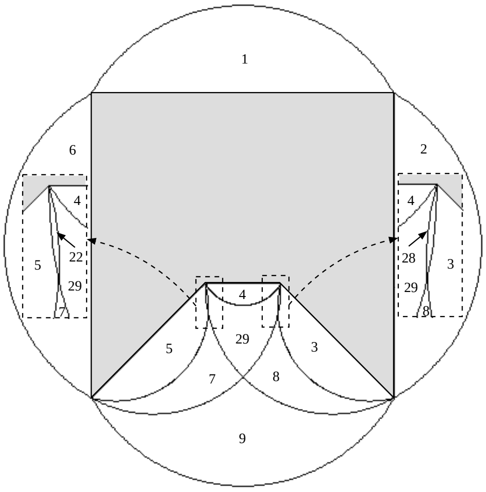


(o) viewing angle = 84.0

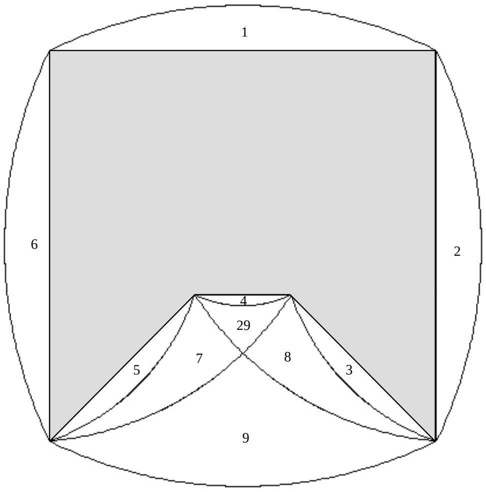
Figure 14: continued.



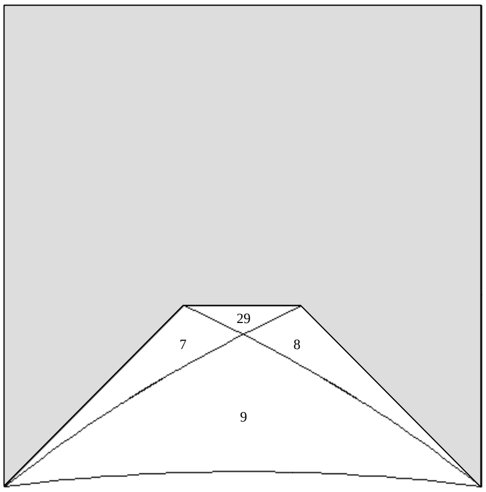
(p) viewing angle = 102.0



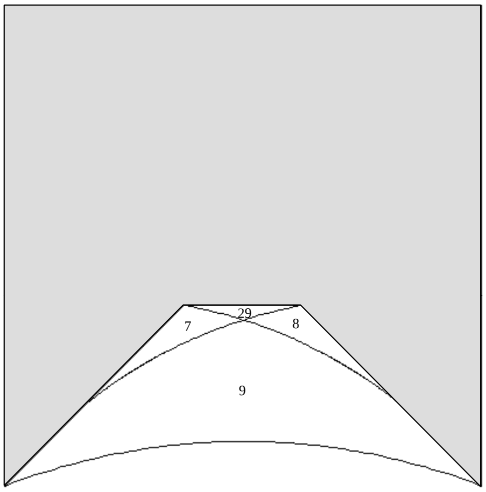
(q) visual angle = 120.0



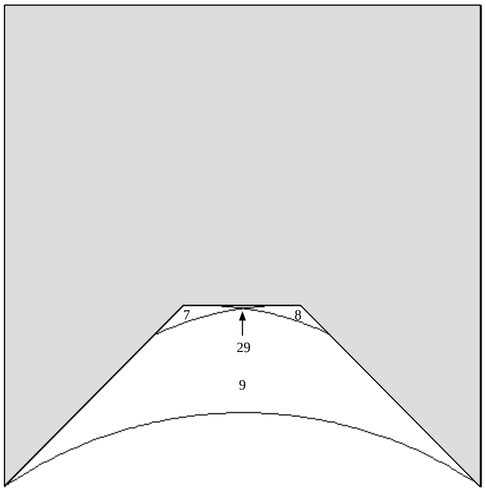
(r) visual angle = 154.0



(s) visual angle = 187.0



(t) visual angle = 201.0



(u) visual angle = 214.0

Figure 14: continued.

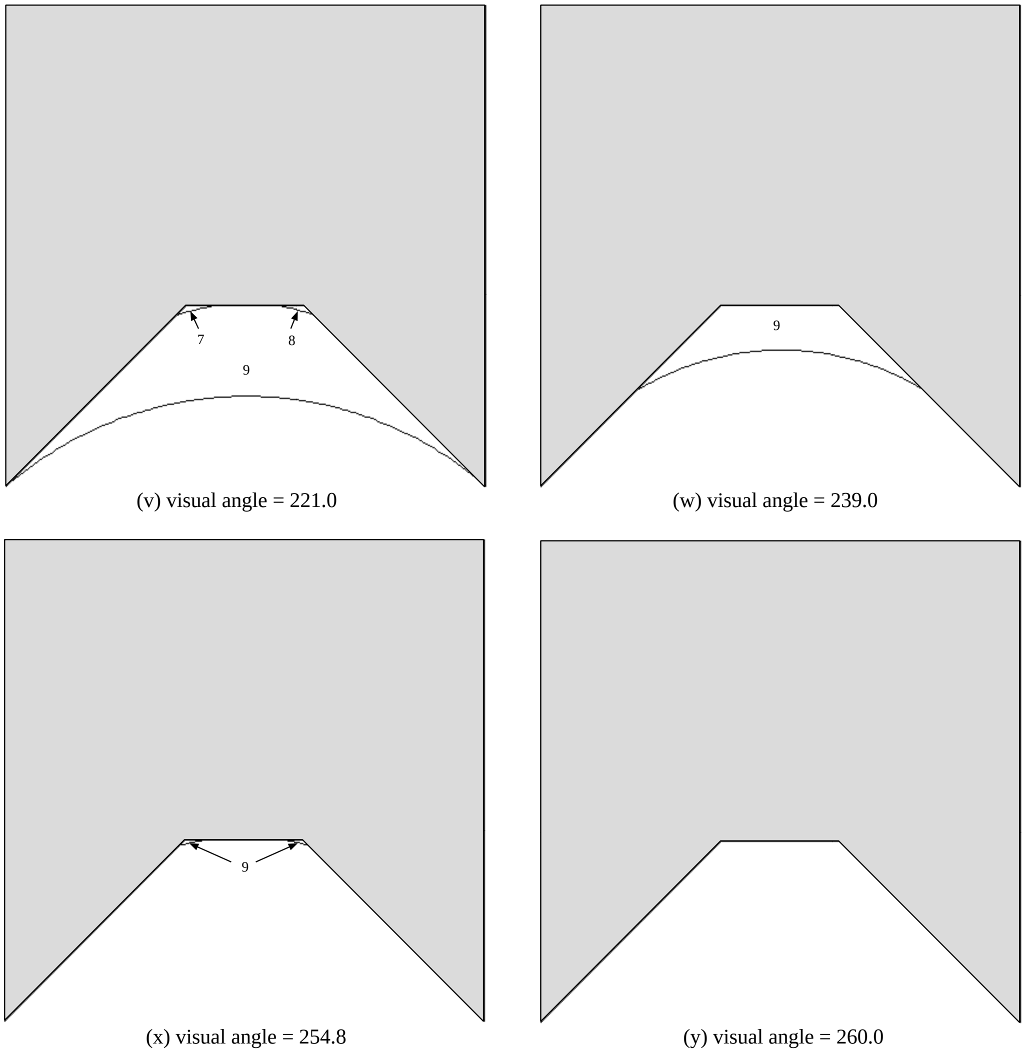


Figure 14: continued.

- *end overlap of two event curves.* This event marks the end of existence of a cell from which multiple edges are seen. It may happen that when overlap ends there is still a common point between the curves, for instance vertex 1 when cell 16 disappears at 45° ((j) $\rightarrow$ (k)), or they may be completely separated, as when cell 20 collapses at 77.65° ((n) $\rightarrow$ (o)). In the former the event angle corresponds to half the angle between the edges at the vertex, while in the latter one must search for the final common point's location to compute the angle.

- *end overlap of event curve and edge.* This event alters the boundary of a cell without changing its existence. Again there may be a single common point of contact after the event, as for curve D and edge E at $45°$ ((j) $\rightarrow$ (k)), or not, as when curve C leaves edge E at $36.87°$ ((i) $\rightarrow$ (k)). This latter event merges the two cells corresponding to aspect 22 generated as discussed earlier. The event angle in the first case is found as the angle between the event edge and the edge corresponding to the event curve, less $180°$ since the corner is concave. In the second case the angle is calculated after a search is done to find the final point of contact.

- *begin overlap of event curve and edge.* In this case the boundary of a cell is again altered without changing the existence of the cell, in essence performing the inverse operation of ending overlap. Before overlapping, the edge and curve may have a common point at a vertex. If not, they may begin intersecting at an endpoint of the edge or in the middle. Instances of these respective events are given for curve G and edge E at $194.04°$ ((s) $\rightarrow$ (t)), curve D and edge E at $14.04°$ ((c) $\rightarrow$ (d)), and curve I and edge D at $253.74°$ ((w) $\rightarrow$ (x)). These visual angles can be calculated in the same manners as for the ending of overlap.

- *coincidence of two event curves.* There are two occurrences of event curve coincidence, curves C and E at $30.96°$ ((g) $\rightarrow$ (h)) and curves G and H at $45.0°$ ((j) $\rightarrow$ (k)). In each instance four cells are removed and two inserted into the parcellation. What is interesting is that two of these cells correspond to the same aspect. For instance the aspect which sees only pseudo edge G exists between the curves G and H. Before coincidence this area is in the right half of the parcellation, while after coincidence it is in the left half. The angle at which this event occurs is found by solving for σ in the event curve equation for one edge with the location of the other curve's defining point inserted.

- *coincidence of event curve and edge.* For this event the curves merge with their respective edges, and both the curve and the cell between curve and edge no longer exist. This occurs at an angle of $180°$ ((r) $\rightarrow$ (s)) for all six real edges and their curves.

In this example a few other departures from the structure of an ideal aspect graph can be noticed. There is no longer a 1–to–1 correspondence between aspects and cells. This is because the cells in scale space are not "convex" and a given projection into viewpoint space may result in separate cells. It is also the case that the cell in scale space representing a single aspect need not be a single volume as is usually defined. For instance, there are two distinct cells for the view of pseudo edge G, which only touch at a single point in scale space.

As a final conclusion, one might wonder what the "important" aspects are for this polygon. If one uses range of scale as a measure, then aspect 9, which sees the view of the concavity, is the strongest. A complete volumetric analysis of the cells in scale space has not been performed. However, by examining the representative parcellations in Figure 14 it seems that those aspects corresponding to views of edges A, B and F as well as pseudo edges I, L, M, O and Q occupy the larger areas in each parcellation, and therefore might be considered as more important. The majority of the other views exist in near the concavity. It is interesting to note that the above views are **not** those corresponding to the larger sized cells of the ideal parcellation in Figure 14.a.

5.3.3 Extensions to 3–D objects

We next consider expanding the object domain. While polyhedra seem analyzable because the 3–D positions of the visible features remain constant, the viewpoint dependent nature of limbs on curved surface objects will likely render their solution computationally intractable. For polyhedra, the visibility surface for an edge feature is the toroidal shape generated by rotating the circular curves about the edge in question [19]. Similar to the pseudo edges in 2–D, the visibility of chains of edges around the boundary of a face must be examined, as the nature of the face's shape changes.

The visibility of a face as a whole is governed by the size of its projected area. If any two points along the boundary are closer than the visual limit, then that portion of the face projects to a dot. The visible portion of the face is the set of points for which the intersection of a cone of solid angle σ emanating from the viewpoint and the face's plane does not cross its boundary. Events will correspond to faces projecting as multiple smaller areas and finally to no visible area. Finally, the events involving self–occlusion can be related to those involving visibility, as was done for polygons, and analyzed in a similar manner. The continued formalization of these events and the definition of the corresponding event surfaces will be the subject of future research.

6 Discussion

In this paper we have presented the idea of incorporating a quantitative measure of scale into the qualitative representation of the aspect graph. The proper use of scale information in computer vision is far from a solved problem. Some elements of this work further emphasize this point.

We began by defining what a scale space aspect graph should be. After presenting several alternative scale interpretations, we now need to examine how closely they follow the original definition. Smoothing the visible surface shape of the object accounts for several intuitive visual effects. However, because of the viewpoint dependent nature of the calculations, the complexity of this method may be prohibitive. In addition, it seems that the desired visual effects could be generated through a combination of the other interpretations, and therefore leave this model redundant. While the use of a finite–sized observer is an aid in ranking cell importance, by itself it does not account for certain of the basic visual phenomena, as do the other interpretations. The approach is less complex than the first, but seems useful only when applied to finite–extent cells while exerting logical control on the maximum size of the viewer. Thus this interpretation seems best suited to a supplementary analysis of the results of other interpretations.

The most promising interpretation seems to be scale as visual feature extent. Both distance from viewer to object and relative sizes of features are taken into account, developing a much more realistic description of the viewpoint space parcellation. However, even this approach has certain minor shortcomings. First, the monotonic reduction in the number of cells is not strictly adhered to, as indicated by the final column in Table 2. The overall trend in size is towards zero, but at certain events new cells are introduced (for instance cell 4 at 36.87°). The large increase in the number of aspects from twelve to forty–three as a nonzero scale value is assumed is at first disconcerting. However, this does not seem so bad when one realizes that the ideal case (which could never happen in practice) is merely a highly degenerate instance of the more general parcellation. Lastly, the measurement of cell importance using volume could potentially benefit from the finite–sized viewer assumption. Whether such an interaction should be modeled as a two step process, or can be combined through a single scale parameter, is still a topic for research.

Finally, it seems necessary to evaluate the "correctness" of the importance labeling of the aspect as given by the different interpretations. In most cases the "important" aspects in the scale space version of an aspect graph do not necessarily correspond to those in the ideal version. Furthermore, some of the results do not agree with our intuition as to what a human would consider important. Future research, perhaps through psychophysical experiments, may lead to other useful scale interpretations or perhaps even a single unified scale parameter.

References

[1] Babaud, J., Witkin, A. P., Baudin, M. and Duda, R. O. 1986. "Uniqueness of the Gaussian kernel for scale-space filtering", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, pp. 26-33, 1986.

[2] Ben-Arie, J. "Probabilistic models of observed features and aspects with application to weighted aspect graphs", *Pattern Recognition Letters*, vol. 11, pp. 421-427, 1990.

[3] Bowyer, K. W., Sallam, M. Y., Eggert, D. W. and Stewman, J. S. "Computing the generalized aspect graph for objects with moving parts", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, to appear, 1993.

[4] Chen, S. and Freeman, H. "On the characteristic views of quadric-surfaced solids", *Proceedings of the IEEE Workshop on Directions in Automated CAD-Based Vision*, pp. 34-43, 1991.

[5] Clark, J. J. "Singularity theory and phantom edges in scale space", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, pp. 720-727, 1987.

[6] Cowan, C. K. "Automatic camera and light-source placement using CAD models", *Proceedings of the IEEE Workshop on Directions in Automated CAD-Based Vision*, pp. 22-31, 1991.

[7] Eggert, D. and Bowyer, K. "Computing the orthographic projection aspect graph for solids of revolution", *Pattern Recognition Letters*, vol. 11, pp. 751-763, 1990.

[8] Eggert, D. and Bowyer, K. "Computing the perspective projection aspect graph of solids of revolution", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, to appear, 1993. (See also Eggert, D. "Aspect graphs of solids of revolution", Doctoral Dissertation, Department of Computer Science and Engineering, University of South Florida, 1991.)

[9] Eggert, D. W., Bowyer, K. W., Dyer, C. R. "Aspect graphs: state-of-the-art and applications in digital photogrammetry", *Proceedings of the ISPRS 17th Congress: International Archives of Photogrammetry and Remote Sensing*, Part B5, pp. 633-645, 1992.

[10] Faugeras, O., Mundy, J., Ahuja, N., Dyer, C., Pentland, A., Jain, R., Ikeuchi, K. and Bowyer, K. "Panel theme: Why aspect graphs are not (yet) practical for computer vision", *Proceedings of the IEEE Workshop on Directions in Automated CAD-Based Vision*, pp. 98-104, 1991.

[11] Gigus, Z., Canny, J. and Seidel, R. "Efficiently computing and representing aspect graphs of polyhedral objects", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, pp. 542-551, 1991.

[12] Gualtieri, J. A., Baugher, S. and Werman, M. "The visual potential: One convex polygon", *Computer Vision, Graphics, and Image Processing*, vol. 46, pp. 96-130, 1989.

[13] Kender, J. R. and Freudenstein, D. G. "What is a 'degenerate' view?", *Proceedings of the ARPA Image Understanding Workshop*, pp. 589-598, 1987.

[14] Koenderink, J. J. and van Doorn, A. J. "The internal representation of solid shape with respect to vision", *Biological Cybernetics*, vol. 32, pp. 211-216, 1979.

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

[16] Koenderink, J. J. and van Doorn, A. J. "Dynamic shape", *Biological Cybernetics*, vol. 53, pp. 383-396, 1986.

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

[18] Kriegman, D. and Ponce, J. "Computing exact aspect graphs of curved objects: Solids of revolution", *International Journal of Computer Vision*, vol. 5, pp. 119-135, 1990.

[19] Levitt, T. S. and Lawton, D. T. "Qualitative navigation for mobile robots", *Artificial Intelligence*, vol. 44, pp. 305-360, 1990.

[20] Lindeberg, T. and Eklundh, J. "Scale detection and region extraction from a scale-space primal sketch", *Proceedings of the 3rd International Conference on Computer Vision*, pp. 416-426, 1990.

[21] Malik, J. "Interpreting line drawings of curved objects", *International Journal of Computer Vision*, vol. 1, pp. 73-103, 1987.

[22] Mokhtarian, F. and Mackworth, A. K. "Scale-based description and recognition of planar curves and two-dimensional shapes", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, pp. 34-43, 1986.

[23] Mokhtarian, F. and Mackworth, A. K. "A theory of multiscale, curvature-based shape representation for planar curves", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, pp. 789-805, 1992.

[24] Nalwa, V. S. "Line-drawing interpretation: A mathematical framework", *International Journal of Computer Vision*, vol. 2, pp. 103-124, 1988.

[25] Plantinga, H. and Dyer, C. R. "Visibility, occlusion and the aspect graph", *International Journal of Computer Vision*, vol. 5, pp. 137-160, 1990.

[26] Ponce, J. and Kriegman, D. "Computing exact aspect graphs of curved objects: Parametric surfaces", *Proceedings of the 8th National Conference on Artificial Intelligence*, pp. 340-350, 1987.

[27] Ponce, J., Petitjean, S. and Kriegman, D. "Computing exact aspect graphs of curved objects: Algebraic surfaces", *Proceedings of the European Conference on Computer Vision*, pp. 599-614, 1992.

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

[29] Rieger, J. "Global bifurcation sets and stable projections of nonsingular algebraic surfaces", *International Journal of Computer Vision*, vol. 7, pp. 171-194, 1992.

[30] Seales, W. B. and Dyer, C. R. "Modeling the Rim Appearance", *Proceedings of the 3rd International Conference on Computer Vision*, pp. 698-701, 1990.

[31] Sripradisvarakul, T. and Jain, R. "Generating aspect graphs for curved objects", *Proceedings of the IEEE Workshop on Interpretation of 3D Scenes*, pp. 109-115, 1989.

[32] Stewman, J. and Bowyer, K.W. "Direct construction of perspective projection aspect graphs for planar-face convex objects", *Computer Vision, Graphics and Image Processing*, vol. 51, pp. 20-37, 1990.

[33] Stewman, J. H. and Bowyer, K. W. "Creating the perspective projection aspect graph of convex polyhedra", *Proceedings of the 2nd International Conference on Computer Vision*, pp. 494-500, 1988.

[34] Stewman, J. H., Stark, L. and Bowyer, K. W. "Restructuring aspect graphs into aspect- and cell-equivalence classes for use in computer vision", *Proceedings of the 13th International Workshop on Graph-Theoretic Concepts in Computer Science*, pp. 230-241, 1987.

[35] Waldon, S. and Dyer, C. "Towards an Aspect Graph for Photometric Imaging", Technical Report, Computer Science Department, University of Wisconsin, 1991.

[36] Wang, R. and Freeman, H. "Object recognition based on characteristic view classes", *Proceedings of the 10th International Conference on Pattern Recognition*, pp. 8-12, 1990.

[37] Watts, N. "Calculating the principal views of a polyhedron", *Proceedings of the 9th International Conference on Pattern Recognition*, pp. 316-322, 1988.

[38] Welford, W. T. <u>Geometrical Optics: Optical Instrumentation</u>, (North Holland Pub., Amsterdam), 1962.

[39] Witkin, A. P. "Scale-space filtering", in <u>From Pixels to Predicates</u>, (Ablex Publishing Corp., Norwood, NJ), pp. 5-19, 1986.

[40] Yuille A. and Poggio, T. "Scaling theorems for zero crossings", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, pp. 15-25, 1986.