Selection and Use of Image Features for Segmentation of Boundary Images*

Deborah Walters

Department of Computer Science, 226 Bell Hall
University at Buffalo (SUNY), Buffalo, New York, 14260, USA

Abstract

An algorithm is developed to segment arbitrary boundary images into sets of boundaries which represent a single object, and to group together lines which correspond to a single object or object part. The algorithm is based on features which were found to used by humans in the early stages of visual processing, and which have a high correlation with perceptually significant aspects of images. In addition, the data structure used is based on the image representation used in the primate visual cortex.

By using perceptually valid features, the algorithm is able to enhance the perceptually significant edges in an image using simple, local, parallel computations. It demonstrates that selective processing can occur in the parallel stages of early visual processing, without domain specific knowledge, iterative processing, or top-down control of some mechanism to shift attention.

KEYWORDS: image segmentation, boundary images, visual psychophysics

*This project was supported by grant IST 8409827 from NSF.

1. Introduction

Images contain too much information for humans or machines to process all of it in detail. The human visual system solves this problem by performing a initial, cursory analysis of the entire image which allows it to pick out automatically what is important in the image (Treisman, 1986), and then to selectively process that information in preference to the rest of the information. That is, a rapid, parallel analvsis of the entire image indicates which regions are likely to contain the most useful information. Then the following stages of analysis, which require more focused, serial processing can concentrate primarily on the preselected regions. This can also be a useful approach for a computer vision system, and in fact the parallel computation of intrinsic images can be viewed as an example of the first stage (Barrow and Tenenbaum, 1981). This paper describes another type of processing which enables the early visual processing to indicate which regions of an image are likely to contain the most useful information, and to selectively process such regions in parallel. This is accomplished by selectively processing image features which have a high correlation with perceptually significant aspects of an image. This is not a new approach. For example, edge detection techniques pick out object boundaries and other edges which are more perceptually significant than more uniform image regions (Marr, 1982). However, the research in this paper presents a new set of features which can enable strong inferences about which regions of an image contain the most perceptually relevant information.

Determining which aspects or features of an image contain the most useful information, and should therefore be preferentially processed is difficult because there are nearly an infinite number of potential features. The dimensions of physics cannot necessarily be used to determine which features are relevant, as perceptual features may lie along some other dimensions. For example, the perceived color of a region depends not only on the wavelength and intensity of the light reflected from it, but also on the relative contrast between it and neighboring regions. So how can perceptually relevant features be found?

The question is further complicated as one set of features may be ideal for one task, but useless for another. One basic machine vision task is to segment an image into different regions which correspond to different objects, or object parts. This may be possible based on the color, shading, texture and shape information. But, are the color and shading information necessary for image segmentation? Not always, as humans can readily segment simple line drawings or boundary images which lack that information. So one way to study image segmentation is to study line drawing perception, and as line drawings are much simpler than natural images, this should make the selection of features easier. Once features are found from line drawings, then it is possible to test them in the analysis of natural images.

But even for simple line drawings, it is not obvious which features should be used. As the goal is to find perceptually significant aspects of an image, and then to determine which features correlate with those aspects, it is desirable to determine what aspects of an image have perceptual significance for humans. It is not possible to just introspect about possible features, as the relevant preattentive stages of human visual processing are not available for conscious introspection (Julesz & Schumer, 1981). The approach taken in this paper is to use psychophysical experiments to explore preattentive vision and to discover image features used by humans. Once potential features are found, their usefulness is tested by developing a computational algorithm based on them, and then testing the algorithm. The algorithm developed here can segment arbitrary boundary images containing both straight lines and curves. It is a simple, data-driven, bottom-up approach, which requires no domain specific knowledge, and demonstrates the importance of using perceptually valid features.

2. Psychophysical Experiments

The psychophysical experiments are based on the perceived contrast of lines phenomenon (Walters and Weistein, 1982a). The patterns in Fig. 1 can be used to illustrate this phenomenon. When viewed at low contrast the lines in the cube (Fig. 1a) appear to have higher contrast than the lines in Fig. 1b. If these differences in perceived contrast can be correlated with the presence of particular image features, it would suggest that stimuli with those features are processed differently from stimuli lacking the features. In particular, stimuli having features associated with high perceived contrast may be preferentially processed. The aim of the psychophysical experiments was to isolate such features. The experiments have been reported elsewhere (Walters and Weistein, 1982b; Walters, 1984, 1985), so only a brief description is included here.

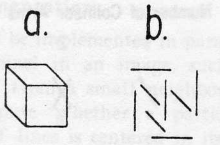

Figure 1

By looking at lots of pairs of patterns designed to differ in terms of various global and local properties, it was found that the difference in perceived contrast did not correlate with any of the global features. For example, closure, global connectivity, perceived 3-dimensionality and objectness did not correlate with perceived contrast. Local features were also explored, and the presence of angles, and the number of free line ends were ruled out. The only two features that did correlate with perceived contrast were line length, and the local connections between the ends of the line segments. For lines which subtended less than one degree of visual arc, perceived contrast was a positive function of line length. For longer lines, there was no correlation between perceived contrast and length. The other local property is the way in which line ends are connected, and experiments show that there is actually a hierarchy of end connections. Figure 2 shows the results of one such experiment. The brightness of various patterns formed of 30 minute line segments was measured relative to a line which subtended 60 minute of visual arc. Some of these patterns could be referred to as the "L", "Fork", and "T" junctions from the Huffman-Clowes tradition (Huffman, 1971; Clowes, 1971; Waltz, 1975). But it turns out that that is not the most useful classification. As section 3.2 explains, it is better to classify these patterns in terms of the spatial relations between the ends of the lines.

From the results in Fig. 2 we can see that line segments joined at their ends have higher contrast than segments where one end abuts the middle of the other segment. And these abutting lines have higher contrast than lines that intersect, while intersecting lines have higher contrast than unconnected lines.

Further experiments found one additional pattern in the hierarchy, as shown in Fig. 3. Two lines connected end-to-end (pattern A) have higher contrast than three lines connected end-to-end (pattern B), which have higher contrast than lines which connect end-to-middle (pattern C), which have higher contrast than the lines which intersect, which in turn have higher contrast than parallel lines.

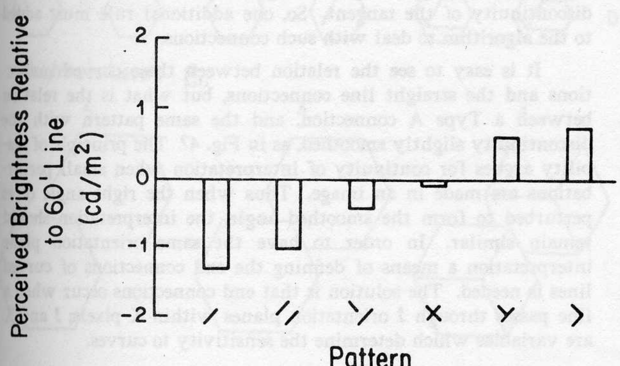

Figure 2

3. Computer Model of Contrast Enhancement

The psychophysical experiments provide evidence that the length of lines, and the connections between the ends of lines, are basic features for human vision. This hypothesis was further tested by implementing it as a computer model. The model receives a boundary image as its input, and outputs the "perceived" contrast of the pattern.

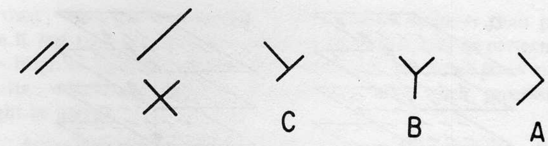

Figure 3

3.1. Enhancement Rules

The model uses the presence of the different types of end connections to implement the following enhancement rules. 1) Each section of line with length L is enhanced by amount "l". 2) Lines terminating at Type A connections are enhanced by amount "a". 3) Lines terminating at Type B connections are enhanced by amount "b". 4) Lines terminating at Type C connections are enhanced by amount "c". 5) Amount "a" > amount "b" > amount "c".

3.2. Detection of Features

The presence of the features can be detected in an image by defining the different end connections in terms of a discrete geometry (Rosenfeld, 1979). The first step is to determine whether each point in the image is part of a line. Thus points can be labeled as either non-line or line points. The line points can then be further broken down into end points and non-end (middle) points. But consider the intersection of two lines. The point that lies on the intersection can be considered to be a middle point of one or the other line, or can even be considered an end point of each of four shorter line segments. Thus some way of defining the point is needed which avoids these ambiguities. An edge detection technique could be used to label each point in the image with the orientation and amplitude of the best line or edge centered on that point. But at the intersection point, it is not so clear what the best line would be. Some edge detection techniques give the orientation of either one or the other line, while others give an average of the two orientations. So, just at a point that provides lots of information about the scene, the edge detection methods don't give sensible answers.

Another problem in edge detection arises because many of the popular approaches to edge detection in computer vision are based on the use of the mathematics of continuous functions. This creates problems in detecting the Type A connection, which is defined in terms of a tangent discontinuity, as in the mathematics of continuous functions, discontinuities are problematical. Poggio et al. (1985) have suggested that the solution to this problem is to regularize the computation. For example, get rid of the discontinuities by convolving the image with a gaussian, and then look for edges in the blurred image. The advantage of this technique is that patterns can then be represented as smooth continuous functions, but it is rather unfortunate from the contrast enhancement point of view, as it gets rid of the the tangent discontinuities, which appear to be such important features for early vision. So edge representation methods based on continuous functions are not very useful for this model.

The solution to these edge representation problems can be found by looking at how edges are represented in the primate visual cortex. If an amplitude/orientation scheme were used, there would need to be two "edge" neurones in the primary visual cortex for each retinal ganglion cell: one to signal the amplitude of the edge at that point, and another to signal the orientation of the edge. But the cortex does not have that organization, instead there is a whole column of edge neurones for each spatial location, and each neurone is sensitive to edges with a narrow range of orientations (Hubel and Weisel, 1968). So, instead of just signally the "best" amplitude and orientation, the orientation column signals the amplitude at many orientations. This representation has several advantages over the amplitude/orientation scheme. For example, at the intersection of two lines, both orientations can be represented, which would disambiguate the pattern.

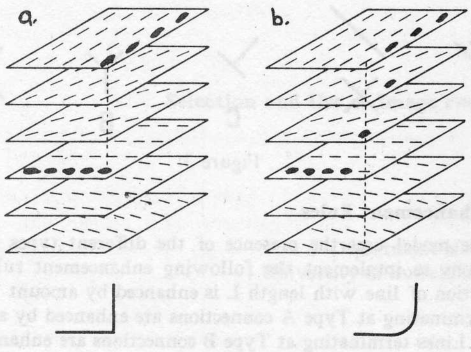

Figure 4

The data structure used in the general enhancement algorithm, is based on the orientation representation of the mammalian visual system. Figure 4(a) shows the basic form of the data structure. It is an orientation plane representation. It is a 3-d space where each point represents a short piece of line having a specific orientation and located at a specific x-y location. An image is transformed into this representation by convolving the image with a separate oriented edge kernel for each orientation plane. It is possible to construct any number of separate orientation planes for this representation. In the current implementation 8 or 16 orientation planes are used. In addition, if boundaries are present at different scales in the image, then a separate orientation plane structure is needed for each scale. This would be required for grey-level images, though not for line drawings containing a single width of line.

One advantage of the orientation plane representation is that it makes it easy to define lines and find tangent discontinuities. A line is defined to be a set of dark pixels such that each pixel is connected to neighboring pixels of similar orientation. The specific definition of 'connected' is orientation dependent. Line pixels can only be connected to other line pixels which lie within a certain x,y distance, and a certain orientation distance, and the greater the x,y distance, the greater the possible orientation distance which can yield a connection. These definitions can be used to label all the pixels in an image as either nonline, end 'e' or middle 'm' points.

Connections can be defined in terms of these 'e' and 'm' labels. For example, for a Type A connection located at point (x1,y1), examining the x1,y1 position in each plane would yield exactly two 'e' labels, and no others. (Actually, the examination may involve a small neighborhood around the (x1,y1) point.) This suggests how to define the different connections in terms of the 'e' and 'm' labels.

3.3. Completeness of the Feature Set

Another question that needs answering is, are these features geometrically complete? Does it cover the space of all possible connections? This question can be answered in terms of all the possible combinations of 'e' and 'm' labels. Figure 5 shows all the possible connections for straight lines of just three possible orientations, in terms of the number of 'e' and 'm' labels at the center of the connections. The upper left three are not connections. The upper right two are both intersections, which receive no contrast enhancement. Three of the connections were used in the psychophysical experiments, and there are two additional connections needed to cover the space. The new connections are hypothesized to belong to the classes as labeled.

There are only a limited number of orientations in Fig. 5. For more orientations, the set would be extended, and the labels can also be extended. Everything out to infinity in the top row is a Type D connection. Everything out to infinity in the second row is a Type C connection. And everything out to infinity in the other rows is a Type B connection. It is important to be able

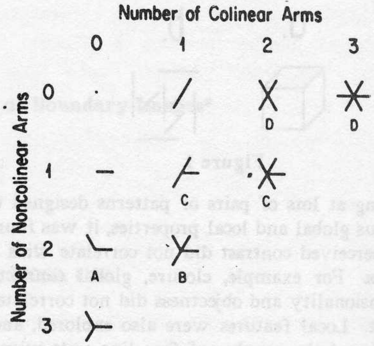

Figure 5

to label all possible types of end connections, but at the same time, the probability of any of the higher order types occurring in a natural scene are very small, thus they are not as important as the few seen in Fig. 5. It can now be seen that the set of connections is complete, and that there is a means of detecting the presence of the features as all connections can be classifyed into the four perceptually valid classes using these rules. 1) All connections with exactly two 'e' labels are of type A. 2) All connections with two 'e' labels, and at least one additional 'e' or 'm' label are of type B. 3) All connections with exactly one 'e' label and one or more 'm' labels are of type C. 4) All connections with no 'e' labels and two or more 'm' labels are of type D.

3.4. Dealing with Curved Boundaries

The examples thus far have dealt only with line drawings containing straight lines. To be useful the algorithm should be able to deal with curves as well. Does the perceived contrast of curves also vary with the type of end connections present? Further psychophysical experiments confirmed this hypothesis. Figure 6 shows the hierarchy of end connections for curved lines, with the lines on the right having the highest perceived contrast, and the lines on the left having the lowest. The results for the curved lines are identical to the results for the straight line segments, although curved lines have a possibility of one further type of end connection, as shown in the A' pattern. With curved segments, two segments can merge or join into one, without a discontinuity of the tangent. So, one additional rule must added to the algorithm to deal with such connections.

It is easy to see the relation between these curved connections and the straight line connections, but what is the relation between a Type A connection, and the same pattern with the discontinuity slightly smoothed, as in Fig. 4? The principle of stability argues for continuity of interpretation when small perturbations are made in an image. Thus when the right angle of is perturbed to form the smoothed angle, the interpretation should remain similar. In order to have the same orientation plane interpretation a means of defining the end connections of curved lines is needed. The solution is that end connections occur when a line passes through J orientation planes within K pixels; J and K are variables which determine the sensitivity to curves.

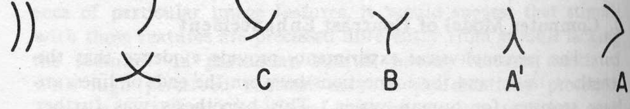

Figure 6

3.5. Model Implementation

The model can be implemented in parallel. Imagine a simple processor at each pixel in an image, such that each processor receives input only from a small neighborhood of pixels. Each processor can compute whether a particular spatial relation between the ends of lines is centered in its neighborhood, and if so, can send the appropriate enhancement out along the appropriate pixels. Note that it's not the mechanism of contrast enhancement that is being modeled - it's the overall computation that is of concern.

4. Model Results

Figure 7a shows the output of the computer model for four of the psychophysical patterns. The results are displayed in terms of a threshold. The highest threshold - that is the highest contrast lines are at the top. The threshold becomes lower in each subsequent line. At the bottom is the lowest threshold where all of the lines which were present in the patterns appear.

The model results agree with the experimental results for all of the patterns used in the psychophysical experiments. Thus the hypothesis that perceived contrast is a function of line length and the type of connections between the ends of lines, is further supported.

4.1. Uses of Features Suggested by the Model

A further use of the computer model is to go beyond the psychophysical results. One limitation with the human experiments is that subjects are only able to make global judgements of contrast - they can say pattern A was overall brighter than pattern B, but they cannot say whether a particular line in pattern A was brighter than the others. But the computer model gives such results, which can help in determining why such processing might be useful.

Actually, some of the possible uses can be seen with these simple stimuli of Fig 7a. Lines which are part of object contours are enhanced relative to lines which form texture, or are perceived as noise. And, the outer contours of objects are enhanced relative to the inner contours.

Looking at another example (Fig. 7b) shows how the model goes one step further. Figure 7b contains two distinct objects, one partially occluding the other. At the highest threshold the lines composing the two objects are not spatially continuous. The the model selectively enhances objects in the foreground, and helps to group lines into two sets which correspond to the two objects.

Figure 7c shows the model results for an impossible object. For the possible object, the model results at an early level (level 2) show the main properties of the object - a blob with a hole in it. That is, the model is giving the topological structure of the object at a very early level. But with the impossible object, at the first level it is represented as a single object, then as one object with two sub-objects, or object components. The model again agrees with our perception of one object if we look at one corner of the drawing, and another if we look at the diagonally opposite corner, and neither we nor the model can get the perceptions to merge.

5. Uses of Image Features

The results of the computer model give more support to the hypothesis that the length of lines, and the spatial relations between the ends of lines, are perceptually valid features. But how should these features be USED?

5.1. Current Approaches

Various theories concerning the use of features have been proposed in computer vision. One conceptually simple use of features is to represent objects in terms of a list of features (Feldman, 1985). A model of an object can be expressed in terms of features and the relations between them, and then portions of an image can be compared to the model to see if the object is present. This is similar to the way line drawing junctions were first used by Roberts (1965). But this use involves domain specific knowledge, which is a major drawback as it is thus not easily extendible to deal with arbitrary images.

Guzman(1979), Kanade (1981), Draper (1981), and Lee et al.(1985) have used very similar line drawing features in their boundary image interpretation algorithms. The features are used in various constraint satisfaction systems. This paper presents another, related use of end connection features, which is not limited to trihedral vertices, and accomplishes a somewhat different task.

5.2. Selective Enhancement

A different use of features is suggested by the psychophysical experiments and computer model. Lines appear to be selectively enhanced based on the presense of a few basic features. (The potential usefulness of this enhancement is described in the next section.) It appears that selective enhancement is possible, even in the automatic parallel stages of processing. This requires no top-down processing, no domain-specific knowledge, and no iterative processing.

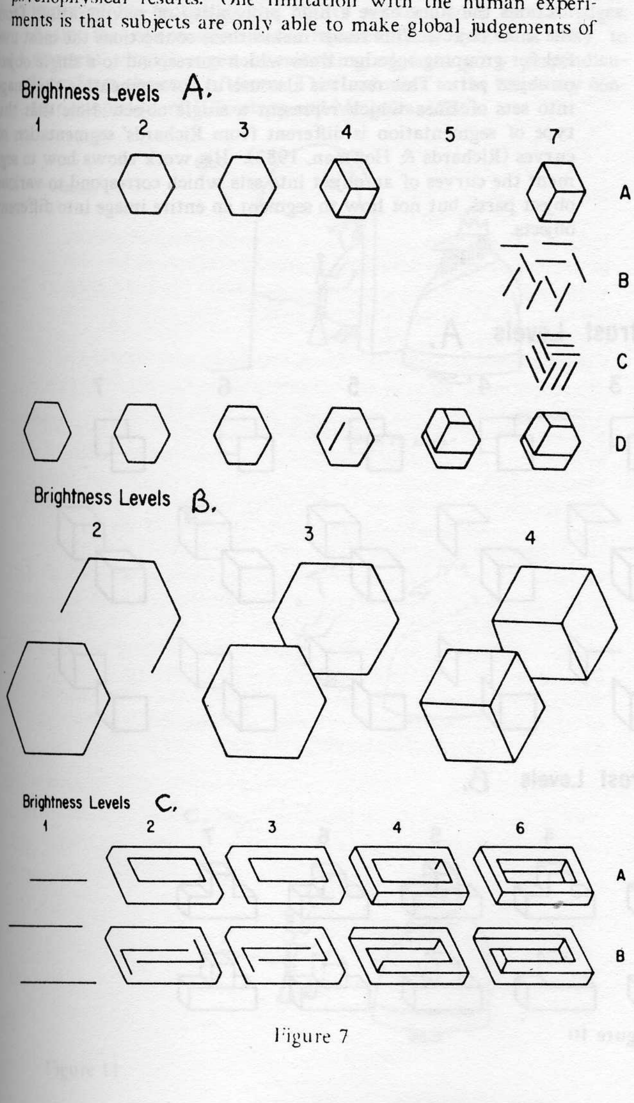

Figure 7

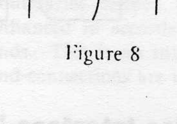

Figure 8

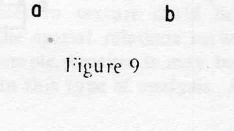

Figure 9

6. Perceptual Significance of Selective Enhancement

Why should such parallel selective enhancement be useful? The computer model provided some hints. The outer contours of objects were enhanced more than the inner contours, and object contours were enhanced more than lines interpreted as texture or noise, and the highest contrast lines were correctly segmented. But why would these results be helpful?

The selective enhancement of outer contours is important for object recognition. An object can usually be recognized just from it's silhouette, which is simply it's outer boundary - outer contours have a special perceptual significance. And the edges of a silhouette can only contain type A connections. This may be the reason that end-to-end connections appear to receive the most contrast enhancement. And this supposed correlation between type A connections and the outer contours of objects makes it possible to infer that the most enhanced lines in an image have a high probability of having arisen from the occluding contours of objects.

The type B connections can arise from either an inner or outer contour of an object, and thus do not have as strong a correlation with outer object boundaries. Even when we divide the type B connections into forks and arrows as Chakraverty and others do (Chakraverty, 1979; Lee et al., 1985), they can still both arise from both types of object contours.

Yet, if two simple assumptions are made, both type A and type B junctions have a high correlation with object contours in general.

The first assumption is:

Assumption 1: Viewing position is representative.

(This means we assume we are looking at an object along a viewing direction which is not one of the few viewing directions which results in the accidental alignment of object boundaries or wires in a scene. (Binford, 1981; Cowie, 1982))

Result 1: Two or more lines meeting at a junction should be interpreted as two or more wires or object boundaries that meet.
Assumption 2: Object position is representative.

(This means we assume objects or wires in a scene are not accidentally aligned. The first assumption concerns looking at objects in such a way as to make them appear to be accidentally aligned. The second assumption concerns cases where the objects are in some form of accidental alignment with each other, independent of the viewing position.)

Result 2a: Line ends meeting at a point should all be interpreted as having arisen from the same object.

Result 2b: Two or more line middles falling on a point should be interpreted as wires or texture boundaries.

Result 2c: Connections containing both ends and middles are most generally interpreted as object boundaries that either occlude or meet other object boundaries.

Result 2d: The end line in a connection should be interpreted as arising from a different object from the middle lines.

The assumptions about nonaccidental alignment do not mean that images with accidental alignment cannot be enhanced or segmented using this algorithm. It just means that the most general interpretation for a basic feature will be utilized. Thus in the majority of cases, the correct interpretation will arise, while a few cases may exist where the algorithm gives an incorrect interpretation.

From these assumptions we see that in Type A and B connections the lines have a high probability of having arisen from the same object. This result makes these connections the most useful for grouping together lines which correspond to a single object or object part. This result is also useful for segmenting the image into sets of lines which represent a single object. Note that this type of segmentation is different from Richards' segmentation of curves (Richards & Hoffman, 1983). His work shows how to segment the curves of an object into sets which correspond to various object parts, but not how to segment an entire image into different objects.

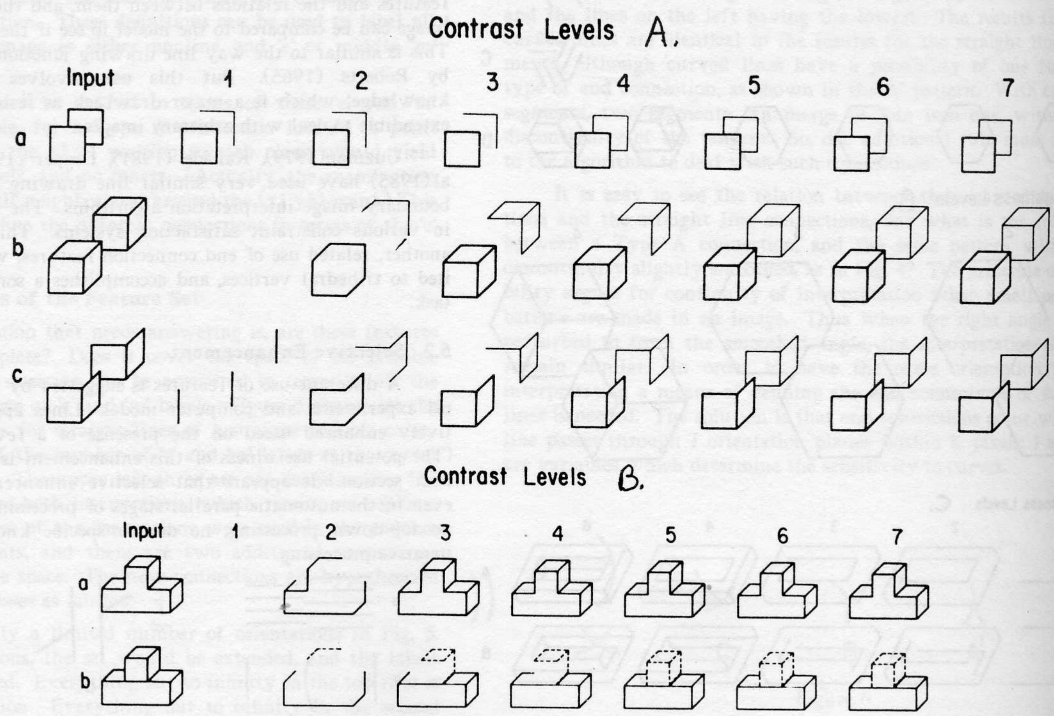

Figure 10

So, from the contrast enhanced lines, certain inferences about the line drawing can be made. Why would these inferences be useful for a visual system? Well, as previously mentioned, a major problem for a visual system is that there is too much information in a visual image to process all of it in detail. One solution is to have some automatic preprocessing system which determines which lines or areas contain the most important information, and then to concentrate the serial processing on those areas, while ignoring other potentially less fruitful areas. This model automatically enhances those lines which have a high probability of being part of object contours, rather than just part of texture or noise. If the next stage of processing has to be selective, it can "attend" only to the enhanced lines and thus not waste resources processing spurious edges. But note that some stages of the selective processing can be done in parallel, and they do not require the top-down control of some mechanism to shift attention (Ullman, 1986).

7. Additional Perceptual Effects

Before describing the results of implementing the enhancement algorithm for curves and straight lines, a couple of other perceptual effects incorporated into the algorithm need to be described. The connections between the ends of straight and curved lines are basic features for the human visual system. The human visual system is also sensitive to virtual edges such as the ones seen in Fig. 8a. Thus the enhancement algorithm may be improved by including the ability to deal with virtual edges. Again the orientation plane representation makes it easy to represent virtual edges, by allowing edges to grow perpendicularly from the ends of lines. This is similar to the boundary contour completion of (Grossberg and Mingolla, 1986).

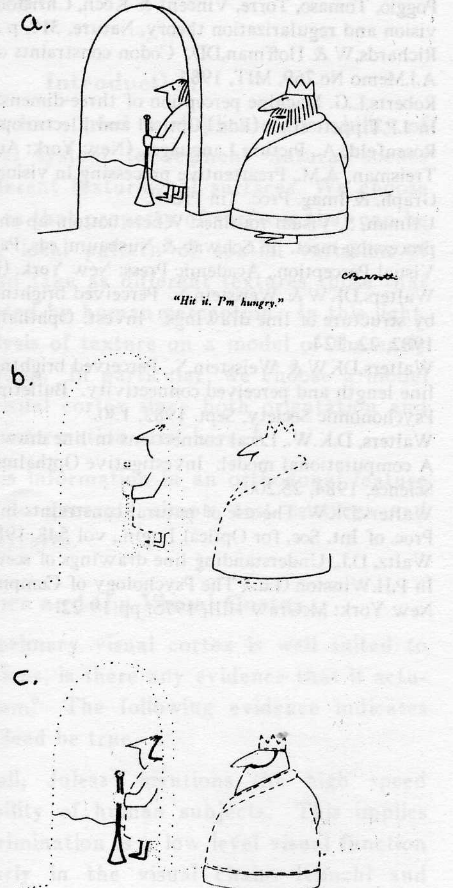

Figure 11

Figure 9 demonstrates one further trick of the human visual system, which is incorporated into the algorithm. The figure is perceived as representing a single diagonal line, which appears to pass behind the surface represented by the vertical lines. Thus pairs of Type C connections, when aligned and having a particular symmetry can be interpreted as occlusion of a single line. A related approach ahs been used by (Lee et al., 1985) to find hidden vertices in line drawings.

8. Segmentation Examples

The contrast enhancement algorithm can be implemented using the orientation plane representation. Figure 10a shows some examples of applying the segmentation algorithm to 2-D , origami, and 3-D objects. The drawings are correctly segmented in all three cases, as indicated by the different line styles for the different objects. Again the outer contours of the objects are enhanced more than the inner contours, and objects in the foreground are enhanced relative to occluded objects. Note that the one set of rules can deal with the three separate domains.

Another example is seen in Fig. 10b, where the first and second images differ only by a single line segment, yet the second alone is represented as two separate objects. Note that this segmentation is indicated early in the process - ie. at Level 2. Thus even at this early stage the segmentation is correct. Note that the segmentation is performed without using either implicit or explicit models of objects, and the top-down processing that model matching requires. This is different from most current algorithms, and enables any boundary image to be processed.

9. Grouping Performance on an Arbitrary Line Drawing

To demonstrate the ability of the enhancement and grouping algorithm, to deal with an arbitrary line drawing, a cartoon from the New Yorker was processed in accord with the algorithm. Figure 11a shows the original cartoon. In Fig. 11b, only the most enhanced groups of lines are displayed - those involved in Type A and A' connections. Only 61 of the total of 86 lines are present, yet object recognition is possible. If just the remaining 25 lines are displayed, object perception is not possible - which is weak evidence that the algorithm picks out the most perceptually salient lines.

The grouping at this stage is depicted by the different line styles. Sixteen of the twenty-three separate objects or object parts are represented at this stage. (Due to reproduction limitations, only four line styles are used in the figure, however each instance of line style indicates a separate set of lines.) Again the algorithm is effective in reducing the complexity of the drawing in terms of the number of lines, without diminishing the grouping capabilities.

Figure 11c shows the final grouping of the cartoon. The sets all correspond to object or object parts that are readily named by humans: ie 'crown', 'robe', 'cuff', 'sleeve', 'foot', etc. There are no groups which would have to be described as "the upper right hand portion of object x", which again suggests that the grouping has perceptual significance. The algorithm could be used as a powerful preprocessor for a scene analysis system, as it accomplishes a lot, given just a handful of simple rules. Later stages of analysis could use the enhanced sets of lines as input to an object recognition algorithm (Pentland, 1985; Biederman, 1985).

10. Texture Boundaries

Up to this point the spatial relations between the ends of curves and lines have been discussed, which are one type of boundary found in images. But the algorithm could equally well be applied to other boundaries - for example, texture boundaries. Instead of using lines or edges as the input to the algorithm, the boundaries defined by differences in texture could be contrast enhanced in accordance with the spatial relations between their ends. This is an interesting example, because it may be that the end-connections are used twice in this type of analysis. Julesz has many impressive experiments designed to uncover the features used by humans in texture segregation, which he calls textons (Julesz & Bergen, 1983). At last count, the texture features include color, elongated blobs (which includes lines), their free terminators, and crossings. Julesz' rejection of global properties such as closure, and global connectivity for preattentive texture discrimination agrees with my findings that these properties are not relevant to changes in perceived contrast. And as there is a strong correlation between the number of free terminators, and the type of end-connections, it may be that the hierarchy of end connections found to alter perceived contrast, can equally well explain texture segregation of patterns composed of lines and curves.

11. Natural Images

A technique for finding texture boundaries in natural images is necessary before the selective enhancement algorithm can deal with natural images which contain regions defined by texture edges rather than intensity edges. But can the selective enhancement algorithm work for natural images which do not contain texture boundaries? We are currently addressing this question by implementing the algorithm for natural images.

There are two basic problems when applying line drawing algorithms to natural images. First is the problem of extracting the edges from the image, leaving the connection information intact. We are developing edge detection techniques, similar to those of Canny (1984), to alleviate this problem. A related problem is that many spurious edges are found with many edge detection techniques. The lack of contrast enhancement for short edges, and the selective enhancement of end connected edges both reduce such problems.

The other problem is that extracted edges are often noisy and contain large gaps. But a similar problem is found in many cartoons - end connections or lines are implied but not explicitly present. The selection enhancement algorithm uses the creation of virtual edges to solve this type of problem, and it may be possible to extend this solution to natural images.

12. Conclusion

The end connections of lines and curves appear to be basic features which allow bottom-up processing of boundary images using a single set of simple rules. The contrast enhancement algorithm suggests that certain selective processing can be performed in the parallel stages of preattentive processing. It is a data-driven approach which accomplishes tasks previously thought to require domain specific knowledge. Object models are obviously necessary for some stages of object recognition, but the contrast enhancement algorithm demonstrates that some steps which were previously thought to require model matching, do not. This shows how productive bottom-up processing can be when psychophysically valid features are used in perceptually valid ways.

13. References

Barrow, H.G. & Tennenbaum, J.M., Interpreting line drawings as three- dimensional surfaces. Artificial Intelligence, 1981, 17, 75- 116.

Biederman, Irving, Human image understanding: Recent research and a theory, Computer Vision, Graphics and Image Processing, (in press).

Binford, T.O., Inferring surfaces from images. Artificial Intelligence, 1981, 17, 205-244.

Canny, John F., Finding Edges and Lines in Images. Master's thesis, MIT, June, 1983.

Chakravarty,I. A generalized line and junction labeling scheme with applications to scene analysis, IEEE Trans. Pattern Anal. Machine Intell 1, 1979, pp202-205.

Clowes, M.B., On seeing things. Artificial Intelligence, 1971, 2, 79-116.

Cowie, R., Modelling people's interpretation of line drawings. D. Phil. Thesis, University of Sussex, 1982.

Draper,S.W., The use of gradient and dual space in line-drawing interpretation. Artificial Intelligence, 1981, 17, 461-508.

Feldman, J.A., Four frames suffice: A provisional model of vision and space. Behav. and Brain Sci. 8, 2, 1985.

Grossberg, S., and Mingolla, E., Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. Psych. Rev. (in press)

Guzman, A., Decomposition of a visual scene into three-dimensional bodies. In A. Grasseli (Ed.), Automatic Interpretation and Classification of Images. New York: Academic Press, 1969.

Hubel,D.H. & Weisel,T.N., Receptive fields and functional architecture of monkey striate cortex. J. of Physiology, 1968.

Huffman, D.A., Impossible objects as nonsense sentences. In B. Meltzer & D. Mitchie (Eds.) Machine Intelligence, 6 pp 295-323, Edinburgh: Edinburgh University Press, 1971.

Julesz, B. & Schumer, R.A., Early visual perception. Annual Review of Psychology, 1981, 32, 575-627.

Julesz, B. & Bergen, J.R., Textons, the fundamental elements in preattentive vision and perception of textures. Bell System Technical Journal, 62,1619-1646, 1983.

Kanade, T., Recovery of the three-dimensional shape of an object from a single view. Artificial Intelligence, 1981, 17, 409-46.

Lee,SH, Haralick,RM & Zhang,MC. Understanding objects with curved surfaces from a single perspective view of boundaries. AI, 26, 1985, pp 145-169.

Marr,David., Vision. San Francisco: W.H.Freeman, 1982.

Pentland, Alex. P., Perceptual organization and the representation of natural form, SRI Technical Note 357, 1985.

Poggio, Tomaso, Torre, Vincent & Koch, Christof, Computational vision and regularization theory, Nature, 317, p 314 - 319, 1985.

Richards,W & Hoffman,DD. Codon constraints on closed 2d shapes. A.I.Memo No.769, MIT, 1984.

Roberts,L.G. Machine perception of three-dimensional objects. In J.P.Tippett et al. (Eds.) Optical and Electro-optical Informa

Rosenfeld, A., Picture Languages, (New York: Academic Press), 1979.

Treisman, A.M., Preattentive processing in vision. Comp. Vis. Graph. & Imag. Proc. (in Press).

Ullman, S. Visual routines: Where bottom-up and top-down processing meet. in Schwab & Nusbaum, eds, Pattern Recognition: Visual Perception., Academic Press: New York, (in press).

Walters,DKW & Weisstein,N. Perceived brightness is influenced by structure of line drawings? Invest. Ophthalm. & Vis. Sci., 1982, 22, 124.

Walters,DKW & Weisstein,N. Perceived brightness is a function of line length and perceived connectivity. Bulletin of the Psychonomic Society, Sept. 1982, 130.

Walters, D.K.W., Local connections in line drawing perception: A computational model. Investigative Opthalmology and Visual Science, 1984, 25,200.

Walters,DKW. The use of natural constraints in image segmentation. Proc. of Int. Soc. for Optical Engin., vol 548, 1985, pp 27-34.

Waltz, D.I., Understanding line drawings of scenes with shadows. In P.H.Winston (Ed.), The Psychology of Computer Vision. New York: McGraw Hill, 1975, pp 19-22.