# Extracting contours by perceptual grouping

### Gul N Khan and Duncan F Gillies

A new contour extraction method is described in this paper. It is based on organization of the image data using perceptual grouping rules, and is therefore largely domain independent. The first step is to form an intermediate line segment representation of the contours by grouping edge points in parallel and at different resolutions. The main feature of the line segment extraction process is the identification, by perceptual criteria, of weak but significant edge points for participation in the grouping process. In the second stage the line segments are grouped into contours, again on the basis of perceptual criteria. This grouping stage is carried out hierarchically using a pyramid structure. In both cases, the use of perceptual grouping allows the filtering of noisy edges and line segments regardless of their strength.

Keywords: line and contour extraction, Hough transform, perceptual grouping, pyramid processing.

There is a widely held belief that humans isolate object boundaries in a scene before recognizing them. Analoogously, in machine vision it is frequently necessary to extract contours from an image since contours often form boundaries and so provide useful information for segmenting images into meaningful regions. There are two main approaches to the identification of object boundaries in an image. One approach is to extract uniform regions and then process them for their boundaries. The alternate approach assumes that a series of edge points define the boundary. The edge points are detected by sensing where the pixel value (intensity, colour or range) changes abruptly. These edge points are then linked into lines or contours. The approach to the extraction of curved contours and boundaries described in this paper is based on organization of the edge point data. In contrast to many current methods, no heuristics or prior knowledge is assumed. The organization uses only the information present in

Paper received: 2 August 1989; revised paper received: 22 April 1991

0262-8856/92/002077-12 (C) 1992 Butterworth-Heinemann Ltd

vol 10 no 2 march 1992

the data itself and a set of perceptual grouping rules which are domain independent.

In contrast to most contour extraction techniques, which link edge points into contours directly without an intermediate representation, we have adopted the strategy of forming an intermediate representation of short line segments. This has two advantages: firstly, it limits the use of domain specific knowledge at the early stages of visual data organization, and secondly it reduces the amount of data to be processed at higher levels. The lack of a simple signal to symbol paradigm gives further support to the idea that an intermediate representation should be formed as a first step in finding a boundary representation of an image. Generally, those intermediate representations which can be formed by local, independent and parallel processes. are preferred for early visual processing, and indeed line segment detectors have been discovered in mammalian vision<sup>1</sup> and, for this reason incorporated by Marr<sup>2</sup> in his well known raw primal sketch. Marr argued that line segments may be grouped together in two ways to form higher order features. In thetaaggregation, a set of similarly oriented line segments are grouped to form a curve in a direction which differs from the intrinsic orientation of line segments. In curvilinear aggregation, the line segments are grouped in such a way that the assembled contours preserve the orientation of segments.

Clearly, there are limitations to the use of line segments, since they form only a piecewise representation of any curve of second order or above. They will not be suitable for boundaries with sharp edges, or with fine detail. However, we contend that there are a large number of cases where either the boundaries are sufficiently smooth, or the detail required by the extraction process is not great, and consequently our method is applicable. For cases where fine detail is required, the method could still be applied at a higher resolution.

The line segment representation is formed at different resolutions by grouping edge points in different sized image windows. Each window is independent, and can therefore be processed in parallel. The grouping is directed by employing the laws of perceptual organization which provide an effective method of

Computer Vision and Graphics Group, Department of Computing, Imperial College of Science, Technology and Medicine, 180 Queen's Gate, London SW7 2BZ, UK

identifying weak but significant edge points for line extraction. It is easy to perform these processes on a local basis and they are highly effective in filtering out noisy edges irrespective of their strength. When the line segments have been obtained, they are grouped hierarchically using a pyramid structure to form curved contours.

## DIFFERENT APPROACHES TO CONTOUR EXTRACTION

In most cases, contour extraction methods incorporate some sort of knowledge into the grouping operations which map edge points into contours. This can be in the form of prior knowledge about the feature being extracted or, if little is known, by heuristics which are true for most domains. The use of both prior knowledge or heuristics limits the applicability of the method.

#### **Extraction of line segments**

The problem of line extraction has been worked on since the early days of scene analysis. Roberts<sup>3</sup> linked edge points into straight segments locally in a  $4 \times 4$  window. He eliminated edge points due to noise by ignoring those which did not have at least one neighbour with a direction within a fixed tolerance.

The Hough transform has been the most widely used method of straight line extraction. It was originally developed for grouping features into simple geometric lines and curves, but more recently has been generalized and can be implemented to group two or threedimensional features<sup>4</sup>. There are a number or parameterization schemes deveoped for implementing the Hough transform to identify straight lines. Most widely used in the angle-radius parameterization, where a line is defined by  $\rho$ , the shortest distance of line from the origin and  $\theta$ , the angle of its normal<sup>5</sup>. This method does, however, have some drawbacks; in particular, there is a difficulty in selecting an optimal quantization of the parameter space. In raster images, a linear quantization implies differing numbers of pixels in each division, and hence the resolution of lines for a given quantization changes with  $\rho$  and  $\theta$ .

Wallace<sup>6</sup> has proposed an alternative parametrization in which the lines are represented by two points on the opposite ends of image boundary. The main advantages of this parameter space are the constant line resolution throughout the image space and the ease with which feature points voting for a line within the parameter space can be determined. Furthermore, sidelobes<sup>7</sup> and bias<sup>8</sup> inherent in the original Hough transform can be reduced by deciding which feature point votes for which line within the parameter space. We have adopted the Wallace parameterization in our algorithm.

Hough techniques, in their basic form are domain independent, and this gives them great applicability. There have been cases where heuristics have been applied. For example, O'Gorman and Clowes<sup>9</sup> tried to improve the Hough transform for line detection by using the gradient direction to filter edge points for a line in a given direction. Their grouping process is based on collinearity but rejects the principle of thetaaggregation advocated by Marr<sup>2</sup>. Consequently, it will work well in the specific domain where straight lines are an intrinsic feature of the image, for example in engineering drawings. However, their transform will be less suitable for our domain, where line segments are a piecewise approximation to a curved boundary, or in shaded images, where lines can be formed displaying theta aggregation.

Weiss and Boldt<sup>10</sup> used geometric grouping to form long straight lines. They connected together short line segments, using grouping which is essentially based on collinearity, though they did also check that the segments were similar in contrast.

In most cases, the Hough transform uses the edge point magnitudes. As an alternative, the use of edge orientation the initial organization of image data has recently been advocated by Burns *et al*<sup>11</sup>. Their line extraction algorithm involves grouping of pixels into line support regions based on the similarity of edge orientation. The lines are extracted from each region by fitting a plane to the intensity surface of the region.

#### **Curve detection**

There are two main categories of curve detection, global and local. Many of the global curve detection techniques are based on graph search methods where the edges are viewed as the nodes of a graph, and a cost factor is associated with each link between nodes<sup>12</sup>. The minimum cost paths in the graph are taken to correspond to the desired boundaries. Heuristics have been applied to the search by Ramer<sup>13</sup> (and others). Although this was basically a global method, an intermediate organization based on streaks was used. The techniques using global criteria are difficult to implement in parallel. They also become computationally expensive when the amount of edge data is large.

Kelly<sup>14</sup> originated the concept of multi-resolution processing for contour extraction by employing coarse to fine planning. The edge pyramids are also used to extract boundaries of objects<sup>15</sup>. A pyramid is constructed by reducing the resolution of the image at successive levels, and then edge detectors are applied at each level. The edges between adjacent levels and at the same level are linked using proximity and orientation. The extraction of straight lines and smooth curves has also been carried out by using an overlapped pyramid structure<sup>16</sup>. In this method the curves are fed into the appropriate levels of the pyramid. The contours are approximated by line segments and at each level segments from the level below are combined using local position, curvature and direction. The main criterion used for detecting compatibility between contour ends is not general and it depends on the radius of a circle which they tried to fit between points. Princen *et al*<sup>17</sup> also have exploited of hierarchies in line extraction.

#### PERCEPTUAL GROUPING

One of the most obvious and interesting facts of human visual perception is that it is almost always organized.

The important part of this organization is the division of the visual field into two parts, figure and ground. The figure usually appears to be nearer than the ground which is extended uniformly behind the figure. The parts of a scene may correspond to objects already seen, but normally it is impossible to see the same object in the same configuration, illumination and from the same perspective in space more than once. This means that humans must be able to partition a scene into coherent, organized and independently recognizable entities without complete prior knowledge.

#### Laws of organization

The main question is how humans partition the scene into independent parts without knowing what might be present. The Gestalt psychologists argue that it reflects the basic and innate functioning of the human visual system. Their fundamental principles of perceptual organization are a set of generic criteria which underlie the procedures discovered by nature for partitioning the visual field. One of the earliest and intuitively most acceptable collection of such laws were proposed by Wertheimer<sup>18</sup>. These laws of organization have been formulated on the basis of their use in identifying ambiguous patterns similar to those shown in Figure 1. They will be seen to underlie the rules for perceptual grouping. These Gestalt laws include:



Figure 1. Laws of organization. (a) Proximity; (b) similarity; (c) continuity; (d) closure; (e) symmetry; (f) simplicity: from two different projections of cube, the simpler (3D) interpretation is perceived

- The Law of Proximity: the stimulus elements which are closer tend to be percieved as one entity. It will be observed that the closer elements in Figure 1a can be perceived as groups forming vertical columns.
- The Law of Similarity: similar elements of a stimulus tend to be part of a unit. This similarity may be in grey-level, colour, orientation or shape as shown in Figure 1b, which is perceived as vertical columns.
- The Law of Good Continuity: stimuli tend to form a group which minimizes a change or discontinuity, as demonstrated in Figure 1c, which is perceived as two lines with first order continuity.
- *The Law of Closure*: the stimulus elements tend to be grouped into complete figures which are most commonly known. The stimulus in Figure 1d will be generally perceived as a circle despite the fact that some part of it is missing.
- *The Law of Symmetry*: the regions which are surrounded by symmetrical borders are perceived as coherent figures in the scenes shown in Figure 1e.
- The Law of Simplicity: in the stimulus where more than one figure can be perceived, the ambiguity is resolved in favour of the simplest alternative. For example if a smaller number of different angles or lines are required to interpret a figure as threedimensional instead of two-dimensional, the observer will normally choose the three-dimensional alternative. This effect is shown in Figure 1f.
- The Law of Common Fate: if a group of elements are moving with a uniform velocity through a field of similar elements, the moving elements are perceived as part of a coherent group.

#### Perceptual grouping criteria

The laws of organization can be utilized to group image data. Among them, proximity, connectivity and similarity in orientation, contrast and intensity have been used both for aggregating edge points into significant line segments and linking line segments for form contours. These grouping criteria seem very simple and ordinary, but to date their strength has not been exploited effectively and completely for segmenting images. The likelihood of a line or contour structure, due to noise or some other accidental phenomenon, decreases with its consistency under the different laws of perceptual grouping. For this reason, these grouping criteria can be used to filter features due to noise and to identify weak but significant edge points and line segments. There are, however, two problems in using these aggregation methods. One is the lack of an effective implementation, and the other is the problem of combining the results when different criteria give different results.

The grouping based on spatial proximity has been used extensively in many clustering problems. The performance of proximity grouping depends on the accurate localization of edge points and line segments. Those edge points are grouped which are closer and lie on the straight line. In the case of line segments, proximity is explicitly defined by the distance between their end points for curvilinear aggregation and perpendicular distance between parallel lines for thetaaggregation. Connectivity is an important image relation since it is preserved over all possible viewpoints. Those edge points are grouped which are connected to each other on a straight line. When no distance tolerance is allowed for proximity grouping, the connectivity is automatically observed.

Similarity is another important principle for grouping edge points and line segments. Similarity in contrast can be used to filter out noisy edge points and line segments, irrespective of their strength, and to recover useful lines and contours. The probability of recovering a significant line or contour increases with the similarity in contrast within a small neighbourhood. Alternatively, similarity in intensity can be used rather than contrast. Theoretically, it should prevent the line segment from crossing over between the background and foreground sides of the boundaries.

Orientation based grouping is the most effective and widely used criterion for line and contour extraction. Since in this case we are aiming to extract straight line segments which form a good piece-wise approximation to a curve, it is not sufficient to group those edge points of common orientation<sup>9</sup>. Instead a slow drift in the orientation is allowed from one edge point to the neighbouring edge point on the line. This slow change in edge orientation, moving along a curved boundary is illustrated in Figure 2. The perceptual relationship formed by orientation grouping with slow drift proves to be more stable and regular for extracting contours due to curved surfaces. The results on artificial and endoscopic colon images, which are presented later, support this claim. In curvilinear aggregation of line segments, only those line segments whose orientation is roughly collinear with the contour directions are grouped. In the general case, the aggregation on the basis of similarity in edge orientation is allowed whether the change in intensity is normal to the line segment or not. This is equivalent to Marr's principle of theta-aggregation for line segment grouping. For contour extraction this principle groups bar-like features on the basis of parallelism and collinear displacements. Assuming that the edges are very short segments having directions, theta-aggregation can also be used to recover significant lines. It has been observed that it recovers those significant line segments which are generally missed. The theta-aggregation principle can also be combined with curvilinearity and orientation



Figure 2. Grouping edges into curved line segments



Figure 3. (a) Theta-Aggregation principle based stimulus; (b) combining theta-aggregation with curvilinearity and orientation drift

drift to extract useful line and contour structures. In Figure 3 some stimuli are shown to demonstrate that humans use these grouping criteria to recover curves and lines.

#### **CONTOUR EXTRACTION**

The laws of perceptual grouping are domain independent; however, their implementation will depend to some degree on the size of features that are being extracted. For example, the size of line segments used as an intermediate representation will vary, depending on the detail in the boundary contours that are to be represented. We now describe our procedure intended for one class of medical images, in which an image window size, of about  $8 \times 8$  pixels was appropriate for line extraction. This choice in turn influenced the other parameters of the algorithm. A full discussion on these parameters follows the algorithm. The two main components of the contour extraction method are the detection of line segments and the aggregation of line segments into contours.

#### Line segment detection

The extraction of line segments consists of edge detection and grouping of edge points. The edge data is processed to extract line segments through two different size window channels  $(4 \times 4 \text{ and } 8 \times 8)$ . The main assumption is that in each window there can only be one significant line segment. The different steps involved in the grouping process are as follows.

Edge points are detected using a simple edge detector (e.g. Sobel<sup>19</sup>). Since orientation is a significant

perceptual criterion for line extraction, accurate edge direction must be computed for the Sobel operator<sup>20</sup>, or as an alternative, an isotropic operator used. Only a few edge points, below a very low threshold, are removed on the basis that their orientation would not be sufficiently accurate.

Following this, the modified Hough transform for straight line extraction, introduced by Wallace<sup>6</sup>, is applied to each overlapping square window of the image. For each candidate line, the coordinates of its voting edges are also stored. From the candidate lines found, the best lines (in the sense of number of votes) are chosen. The number of votes must be larger than a minimum voting threshold  $V_{\rm min}$ , and the number of lines chosen is denoted  $L_{\rm max}$ . For each candidate line thus selected, the following perceptual filtering is performed.

Firstly, connectivity grouping is performed. The edge points which are not part of the largest cluster are dropped from the voting strength of the particular line being filtered. The connectivity of edge points is ensured by not allowing any gap between adjacent points on the line. If the number of edge points in the largest cluster drops below the threshold for minimum votes  $V_{\min}$ , it is assumed that the selected line segment is not significant, and further processing is aborted for that particular line. Connectivity grouping in this case is similar to proximity grouping which is based on the Euclidean distance between each voting edge. In the case where the window is small it is not necessary to have a separate test for proximity. Once the continuity filter has been applied, two further alternative grouping processes are applied.

The first is edge orientation grouping. A slow drift in orientation is allowed by computing the difference of orientation for each pair of adjacent edge points. This is equivalent to differentiating the orientation with respect to distance or calculating curvature along the curved line segment. The edges are partitioned in such a way that each partition consists of those edge points whose orientation change it within a tolerance  $\delta\theta$ . The edge count for the largest partition is then taken as the voting strength of that line and is compared with  $V_{\min}$  to test for the significance of line. If the count is less than  $V_{\min}$ , the selected line does not qualify as a significant line based on orientation grouping.

The other perceptual filter is based on contrast and intensity grouping. the edge points whose contrast and intensity are within their respective tolerances are grouped. The edge contrast tolerance  $\delta C$  compensates for errors in edge magnitude computation, while intensity tolerance  $\delta I$  helps to overcome errors due to quantization in grey-levels. Similar to the orientation grouping process, only edge points in the largest cluster are retained in the vote for a particular line.

If no significant line remains in a particular image window following either contrast and intensity or orientation filtering, it is assumed that no useful line segment is present in that window. When more than one line qualifies in one or both grouping processes, a choice is made as follows:

Preference is given to orientation grouping and if more than one line qualifies, the line with maximum number of votes is selected. Preference can also be given to that line whose direction is normal to the edge orientation. If orientation grouping fails to identify a line in a given window then the lines are tested solely on the basis of CI-Grouping.

The above steps only describe the line detection process in principle. In the detailed implementation, greater efficiency can be achieved. For a sequential implementation, contrast and intensity grouping is only required when orientation grouping fails. Alternatively, the above procedures can be implemented using parallel processing for the individual grouping criteria.

#### From line segments to contours

Perceptual criteria can again be applied as a filter in the next stage of the algorithm which groups line segments into contours.

A line pyramid is constructed in such a way that its leaf nodes store line segments extracted from each  $4 \times 4$  window, and the nodes above those store the line segments extracted from each  $8 \times 8$  window. The intermediate nodes store contours grouped from the lower nodes by a variety of processes. The grouping is mainly carried out on the line segments detected in each  $8 \times 8$  image window. Proximity, theta-aggregation, curvilinearity, continuity and similarity in line contrast are the main grouping principles used. We



Figure 4. 50% overlapped pyramid scheme where thick squares are the central  $2 \times 2$  children



Figure 5. Typical three-neighbours used to group line segments

have used a pyramid based on  $4 \times 4$  overlapping neighbourhoods, and each node has four parents and sixteen children, as shown in Figure 4.

For nodes at the first level above the line segment nodes, three neighbours are considered for linking lines to form a contour segment. One neighbour is in the prior grouping direction, and the other two are its left and right side neighbours. Initially, the grouping direction is approximated from line segment orientation. For example, as shown in Figure 5, if the grouping direction is NE, then three neighbours considered are in the N, NE and E directions, and for the N direction the three neighbours are NE, N and NW.

The grouping operation is carried out by each parent using the line data supplied by its 16 children as follows. If there is no line segment in the central  $2 \times 2$  block, then either no contour exists in the image area corresponding to that node, or some portion of the contour is missing. Support for the missing portion of contour segment is sought from the bottom level line data. If no contour segment is found, a no-contour flag is passed to its parents.

If one or more line segments exist in the central block then for every line segment, the possibility of grouping it with the neighbouring line segments is checked. Two line segments are grouped if they are parallel and located within a pre-defined distance  $D_{max}$ , or if their directions are at the most 45° apart and their end points are within the pre-defined distance threshold  $D_{max}$ . The central block line segments are kept as a contour segment along with its tangent directions at both ends. All the contour segments information is passed on to the parents for further grouping. This information also includes the identification tag for those neighbours which are in the outer block, and have been identified as part of that particular contour segment.

For all the nodes at higher levels, the contour segments from the central block children are tested for grouping among themselves or with the outer 12 children. This process proceeds as follows. For every central contour segment, the possibility of grouping it with neighbouring segments is checked. Two contour segments are grouped if their respective ends are within a pre-defined distance  $D_{max}$ , their end directions (tangents) are compatible and the neighbouring line segments of both contours follow at least one of the grouping criteria (theta-aggregation or curvilinearity). The same process is repeated for the resulting contours until there are no more grouping possibilities. From the resulting groups the central children contour segments are merged into longer contours by making a single group of line segments, and by modifying the end tangent directions. The information about the outer children which are part of the larger group is passed on to the parent along with the merged contour. The central contours which are not merged with any other contour segment, are flagged as completed and they are passed on to their parents. The higher level nodes just







Figure 6. (a) Colon image; (b) output of Sobel edge detector

pass on these completed contours to their parents until the root is reached.

At the root level, the length of all the contours is checked and only those with length longer than a threshold  $CL_{\min}$ , are retained.

During the grouping process, thinning can also be performed to eliminate those line segments which are parallel to the contour segments and located within a pre-defined distance. An additional step for filling the gaps can also be added, where the edge point data can be used for merging closer contours. At this stage, the knowledge about the shape and size of contours can also be exploited to achieve better results.

#### **EXPERIMENTAL RESULTS**

The performance of the technique was analysed by testing it on both endoscopic and computer generated images. In many applications of image analysis, noise is present in the digitized images, and medical images have typically a great degree of noise. This means that to obtain meaningful results, segmentation algorithms must be tailored for the application. Although the contour extraction algorithm was developed primarily as part of a contour detector for endoscopic colon images<sup>21</sup>, it is expected that it will prove to be of much wider applicability.

The colon images were digitized using 256 grey levels from a video tape of colonoscopy procedures. There



Figure 7. (a) Artificial image; (b) output of Sobel edge detector

are a number of sources of noise in them which include but are not limited to variable specular reflection, texture and different types of matter in the colon. The magnetic media left the additional noise in the digitized images due to the process of recording. The algorithms were also tested on artificial images which were constructed with a varying amount of known random noise, expressed as a percentage of the average signal level. The results on one colon and one of the artificial images containing  $\pm 22\%$  noise are presented here. The images along with their Sobel edge detector outputs are shown in Figures 6 and 7.

The outputs for the individual grouping operations are presented for two resolutions in Figures 8 to 12.



b

Figure 8. Line segments extracted using orientation grouping for the colon image. (a) Grouping based on orientation drift and theta-aggregation; (b) O'Gorman and Clowes method



b

Figure 9. Line segments extracted using orientation grouping for the artificial image. (a) Grouping based on orientation drift and theta-aggregation; (b) O'Gorman and Clowes method

The orientation grouping performs comfortably well for a large range of noise, but its performance deteriorates with the increase in noise. The edge point grouping based on edge contrast and pixel intensity provides better results for small amounts of noise, but the overall results become poorer as the noise increases.

The line extraction method based on the O'Gorman and Clowes method has also been implemented and applied to the same images. Their method only groups those edge points whose orientation is perpendicular to the line segment direction. The line segments of curved contours are difficult to extract in this way, and as high tolerance between the line normal and edge orientation is required for extracting all the useful line segments, which may in turn produce line segments due to noise. The method is optimized by increasing the tolerance between the edge orientation and the line normal direction from  $\pm 5^{\circ}$  to  $\pm 25^{\circ}$ . The maximum tolerance is that beyond which the line segments due to noise start appearing in the output.

Figures 8b and 9b shows the optimum results for line segments extracted by the O'Gorman and Clowes method. The method misses more than 50% of the useful contour segments compared to the orientation grouping based on slow drift and theta-aggregation (see Figures 8a and 9a). The amount of noise filtering achieved by the perceptual grouping can be estimated





Figure 10. Output line segments for the colon image. (a) When the perceptual grouping criteria are employed; (b) without using perceptual grouping

by comparing its output with the line segments extracted without applying perceptual grouping. In Figures 10a and 11a the line segments extracted by employing perceptual grouping for two resolutions are shown, while Figures 10b and 11b contain the line segments detected without applying perceptual grouping.

It is difficult to judge the performance of different line detection methods based on the total number of extracted line segments alone. Nevertheless, with the help of the output results presented already and the statistics of the output of different grouping processes, we can gain an insight into the ability of different grouping techniques in reducing the image data. The total number of line segments extracted under different





Figure 11. Output line segments for the artificial image. (a) When the perceptual grouping criteria are employed; (b) without using perceptual grouping

grouping methods and resolutions are shown in Table 1. It is easy to see that the perceptual grouping of edge points reduces the amout of line data almost to 50% without losing useful line structures. Another interesting feature is that orientation grouping covers more than 90% of the line structures. This is all due to the introduction of slow drift in orientation and theta-aggregation. Without this, the O'Gorman and Clowes technique is unable to extract comparable line segments.

The line segment representation formed already is fed to the pyramid structure for the formation of

 Table 1. Line segments extracted by employing different

 perceptual and other grouping processes

Test images		Figure 6 Colon image		Figure 7 Artificial image	
Grouping process	8×8	4×4	8×8	4×4	
Line segments extracted without applying perceptual grouping	48	784	721	1728	
Line segments detected by using perceptual grouping	146	400	164	468	
Line Segments found by O'Gorman and Clowes method of collinearity grouping	47	191	63	225	
Orientation grouping with slow drift and theta-aggregation	124	35	159	461	
Share of CI-Grouping in total line segments found by our algorithm	22	45	5	7	



Figure 12. Output line segments extracted using perceptual grouping for  $4 \times 4$  image window. (a) Colon image; (b) artificial image

contours. The grouped line segments for contours are shown in Figure 13. These groups have been formed in a single pass by travelling from the bottom level of the pyramid to the top. The longer lines which are based on an  $8 \times 8$  image window, are dominant in these groups because the  $4 \times 4$  window based lines (shown as thicker lines) are only used for filling the gaps. The groups of line segments are then replaced by contours which are shown in Figure 14.

#### Selection of thresholds

The grouping algorithms described above use a number of thresholds at various steps, and the reader may wonder about the selection of these thresholds and their sensitivity. In practice, they can be optimized by experimental results on a large number of images for a particular domain. However, experimentation demonstrated that the contour extraction method is robust to these choices and works well on widely different images without changing thresholds. The effects and the bases for these choices are now discussed.



h

Figure 13. Groups of line segments formed. (a) Colon image; (b) artificial image

The amplitude of response from an edge operator is a function of edge magnitude, its orientation and its distance from the centre of the operator. Edge removal on the basis of edge strength also removes the low amplitude, perceptually significant edges. The edge threshold used is intended only to remove low accuracy edge points. It can therefore be set for the whole image. The threshold is set at six for computer generated images (depending on the amount of noise), and 18 for endoscopic colon images.

The choice of maximum number of lines  $(L_{max})$  retained from the original Hough transform and the minimum number of votes  $(V_{min})$  required for each line both depend on the image window size and the amount of maximum background noise in the image. The maximum number of lines selected is taken as the image window side. The minimum number of votes  $V_{min}$  is also affected by the amount of overlapping of the windows and the tight control on the grouping principles. The minimum votes for qualifying as a significant line is chosen as the window side minus the overlap size (eight and four for  $12 \times 12$  and  $6 \times 6$  windows respectively).



Figure 14. Image contours overlaid on the images. (a) Colon image; (b) artificial image

The tolerance threshold for orientation  $\delta \Theta$  determines the limits on the curvature of contours to be extracted. The orientation tolerance is easy to estimate from the psychological studies which demonstrate how much tolerance humans allow in orientation when they group similar oriented patterns. The line data organization in the visual cortex can also be used as a clue in the selection or orientation tolerance. The cortical table<sup>1</sup> suggests that the maximum orientation resolution for line segments is 10°. The choice of 8° to 12° tolerance is influenced by all of these factors and has achieved good results.

The grouping thresholds  $\delta C$  and  $\delta I$  for edge contrast and pixel intensity produce stable results over a wide range of scene illumination. The stability improves when both grouping criteria are employed together. The pixel intensity threshold is affected considerably by the errors in edge localization, while for edge contrast a small variation changes the results considerably. For both  $\delta C$  and  $\delta I$  the tolerance is fixed at three.

The proximity grouping threshold for line segments  $(D_{\max})$ , is used as eight pixels for  $8 \times 8$  window line segments and four pixels for  $4 \times 4$  lines. When the line end points (for curvilinear aggregation) or lines themselves (for theta-aggregation) are separated more than  $D_{\max}$ , they are not grouped. Only those contours are

retained whose length is greater than a threshold  $CL_{\min}$ , which is chosen to be at least four line segments.

#### CONCLUSIONS

In this paper the capabilities of perceptual grouping are demonstrated at the early level of vision specifically in filtering noisy edges, irrespective of their strength. As the application of this research is for on-line image analysis<sup>21</sup>, the possibility of parallel implementation was kept in mind during the whole development process. Custom VLSI hardware for the Hough transform<sup>22</sup> encouraged its use as a part of the method. Previously, the Hough transform has been mostly used as a global method for contour detection. It is utilized here for detecting line segments locally. The bias in the Hough transform is prevalent for small image windows and precise detection of line segments cannot be guaranteed at different orientations. The image window can be rotated (e.g. 9°, 18°, 27° and 36°) to detect more accurate lines, but in our particular application the line and contour extraction accuracy is adequate. The Hough transform for a small window avoids long computation time and large memory requirements. It is possible to implement the whole process on a single VLSI chip. These individual processing elements can serve as a part of the pyramidal computer for the contour extraction process.

Normally, in existing contour extraction techniques the edge detection is assumed to be a local and parallel process, and edge linking is carried out as a global and sequential process. Our contour extraction method is implementable in parallel using pyramid computer architecture. The contours are constructed in a single pass, starting from the bottom level of the pyramid. The transition gap between signal to symbols (from edge points to contours) is traditionally considered as a source of discontinuity for the flow of information. Our approach which utilizes a multi-resolution representation offers the potential for eliminating this discontinuity.

The main problem with the resulting contours are their lack of smoothness and precise location. This has happened mainly because of the use of thetaaggregation which, though an important principle for the detection of significant contours, is imprecise in locating their position. In the results of Figures 13 and 14, the contours are placed in such a way that they pass through the mid-points of each line segment.

This is the first successful attempt to use perceptual grouping in detecting arbitrary shaped contours from noisy images. The results clearly demonstrate the viability of the method, though further improvements are possible.

#### ACKNOWLEDGEMENTS

The first author was financially supported by the Quaid-e-Azam Scholarship from the Ministry of Education Government of Pakistan, his parent department, and ORS Award from CVCP, UK. This work was also supported by Olympus Optical Co. Tokyo.

#### REFERENCES

- 1 Hubel, D H and Wiesel, T N 'FERRIER LEC-TURE: functional architechure of the Macaque monkey visual cortex', *Proc. Royal Soc. Lond. B*, Vol 198 (1977) pp 1–59
- 2 Marr, D 'Early processing of visual information', *Philosophical Trans. Royal Soc. Lond. B*, Vol 275 (1976) pp 483-524
- 3 Roberts, L G 'Machine perception of threedimensional solids' in Tippet, J et al (eds), Optical and Electro-Optical Information Processing, MIT Press, Cambridge, MA (1965) pp 159–197
- 4 **Ballard, D H** 'Generalizing the Hough transform to detect arbitrary shapes', *Patt. Recogn.*, Vol 13 No 2 (1981) pp 111–122
- 5 Duda, R O and Hart, P E 'Use of the Hough transformation to detect lines and curves in picture', *Commun. ACM*, Vol 15 No 1 (1972) pp 11– 15
- 6 Wallace, R S 'A modified Hough transform for lines', Proc. IEEE Comput. Soc. Conf. on Comput. Vision & Patt. Recogn., San Fransisco, CA (1985) pp 665–667
- 7 Brown, C M 'Inherent bias and noise in the Hough transform', *IEEE Trans. PAMI*, Vol 5 No 5 (1983) pp 493–505
- 8 Cohen, M and Toussaint, G T 'On the detection of structures in noisy pictures', *Patt. Recogn.*, Vol 9 No 2 (1977) pp 95–98
- 9 O'Gorman, F and Clowes, M B 'Finding picture edges through collinearity of feature points', *IEEE Trans. Comput.*, Vol C-25 No 4 (1976) pp 449–456
- 10 Weiss, R and Boldt M 'Geometric grouping applied to straight lines', Proc. IEEE Comput. Soc. Conf. on Comput. Vision & Patt. Recogn., Miami Beach, FL (1986) pp 489-495
- 11 Burns, J B, Hanson, A R and Riseman, E M 'Extracting straight lines', *IEEE Trans. PAMI*, Vol 8 No 4 (1986) pp 425–455
- 12 Martelli, A 'An application of heuristic search methods to edge and contour detection', *Commun. ACM*, Vol 19 No 2 (1976) pp 73–83
- 13 Ramer, U 'Extraction of line structures from photographs of curved objects', *Comput. Vision & Image Process.*, Vol 4 No 2 (1975) pp 81–103
- Kelly, M D 'Edge detection in pictures by computer using planning', in Meltzer, B and Michie, D (eds), *Machine Intelligence*, Vol 6, Edinburgh University Press, Edinburgh (1971) pp 397-409
- 15 Hong, T H, Shneier, M and Rosenfeld, A 'Border extraction using linked edge pyramids', *IEEE Trans. Sys., Man & Cybern.*. Vol 12 No 5 (1982) pp 660–668
- 16 Hong, T H, Shneier, M, Hartley, R L and Rosenfeld, A 'Using pyramids to detect good continuation', *IEEE Trans. Sys.*, *Man & Cybern.*, Vol 13 No 4 (1983) pp 631–635
- Princen, J, Illingworth, J and Kittler, J 'A hierarchical approach to line extraction', *Proc. CVPR* 1989 (1989) pp 92–97
- 18 Wertheimer, M 'Laws of organization in perceptual forms', *Psychologische Forschung*, Vol 4 (1923) pp 301–350. (Translation in A Sourcebook of Gestalt Psychology, Ellis, W (Ed) Harcourt Brace, New York, 1938.)

- 19 **Duda, R O and Hart, P E** Pattern Classification and Scene Analysis, Wiley, New York (1973)
- 20 Kittler, J 'On the accuracy of the Sobel edge detector', *Image & Vision Comput.*, Vol 1 No 1 (1983) pp 37–42
- 21 Khan, G N Machine Vision for Endoscope Control and Navigation, PhD Thesis, Department of Com-

puting, Imperial College, University of London (1989)

22 Rhodes, F, Dituri, J, Chapman, G, Emerson, B, Soares, A and Raffel, J 'A monolithic Hough transform processor based on restructurable VLSI', *IEEE Trans. PAMI*, Vol 10 No 1 (1988) pp 106–110