

# Robust Grouping of Intensity Changes in Outdoor Scenes by a Histogram and Graph based Method

Axel Korn

Fraunhofer-Institut für Informations- und Datenverarbeitung (IITB)  
D-7500 Karlsruhe 1, Federal Republic of Germany  
Phone +49 721 6091-250, FAX +49 721 6091-413

---

*The problem of grouping the gray value variations at the border of complex structured surfaces in outdoor scenes is considered in order to get a simple symbolic representation of such boundaries. Our approach is based on a combination of histogram und minimal spanning tree (a graph-theoretical method) evaluation. In this way the direction and magnitude of the gray value gradient and additionally the spatial distance of contour points are considered in a common framework. After cutting the tree at some edge joining "inconsistent" clusters, which can be selected by combining the mean and standard deviation of all tree edges, the subtrees are symbolically described by their principal axes of inertia. From many successful applications, the examples presented in this contribution demonstrate that the largest subtree is appropriate in order to localize, e.g., road boundaries.*

---

The issue to be addressed in this paper relates to the grouping of data in low-level-vision for the detection of object boundaries in outdoor scenes. The difficulty of this task depends mainly on the structure of the scene, where, e.g., texture, occlusions, and poor contrast may disturb relevant contours.

It is generally accepted that the detection of common properties of picture domain cues, such as, e.g., orientation or contrast of contours is insensitive to problems associated with a local analysis. Numerous papers published in the last few years suggest methods to detect texture boundaries (e.g., Vistnes 89, [1] and Voorhees and Poggio 88 [2]) or to group fragmentary contour elements to object boundaries (e.g. Mohan and Nevatia 89 [3], Fan, Medioni, and Nevatia 89 [4]). However, the applications shown in many of these papers do not refer to outdoor scenes and, additionally, most approaches seem to be very time consuming. One of our goals is to show that a very efficient technique for boundary detection in outdoor scenes could be implemented by combining some relatively simple methods.

## METHOD

The strategy in our approach is to look for a global feature, namely the orientation of contour elements, by maximum detection in a histogram of gradient directions. After that, all points with directions of the gray value gradient in a certain interval around a significant maximum are considered as nodes of a minimum spanning tree (MST). For the weights of the graph edges the differences of the position and contrast of contour points are considered. The contrast is given by the magnitude of the gray value gradient. First experiments have yielded acceptable results even if an Euclidean distance is used in a (x,y, gradient magnitude)-space.

The decomposition of the MST into subtrees associated with particular contour segments is certainly a problem. It has been approached empirically by defining the weight of an "inconsistent" graph edge as a linear superposition of the mean weight of all edges and its standard deviation. The definition of an "inconsistent" edge is demonstrated in Fig. 1 which is taken from Zahn 71 [5].

Fig. 1 b) shows the MST of the weighted linear graph in a) and c) illustrates the "inconsistent" cluster-joining edge (A, B) between clusters  $C_1$  and  $C_2$ .

In a last step the nodes of a subtree are considered physically as mass points with masses equal to the magnitudes of the corresponding gray value gradients in the image domain where they form a cluster. This cluster may define an object boundary. The position of the cluster is the point of gravity and its orientation is given by the principal axis of inertia.

The advantage of the MST method compared to some other graph-theoretical methods for describing Gestalt clusters has been discussed by Zahn 71 [5]. But in his approach only the Euclidean distance in (x,y)-space determines the weights of the MST edges. As mentioned above the nodes of the MST in our approach have the attributes position (x, y), magnitude, and direction of the gray value gradient which is

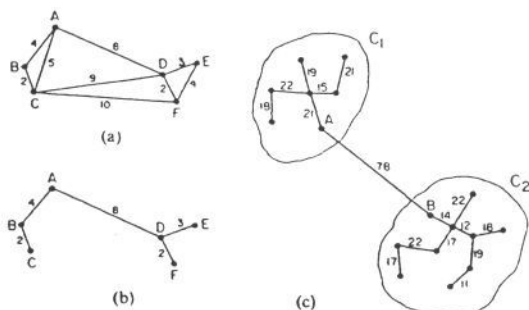


Figure 1. Example of a graph in a), the MST (Minimal Spanning Tree) of a) in b), and illustration of an "inconsistent" cluster-joining edge (A, B) in c) from Zahn 71 [5].

which is calculated according to the method described in Korn 85 [6], Korn 88 [7]. Furthermore, the nodes of the MST are contour elements which are the maxima of the magnitude of the gray-value gradient in the direction of the gradient. Strictly speaking, these points are only candidates for contour elements because we don't compare the result for different Gaussian masks in a filter bank to find the best filter parameter as proposed in Korn 88 [7].

## IMPLEMENTATION AND RESULTS

Fig. 2 shows an outdoor scene which is picked up by a CCD-camera in a moving car. In order to facilitate autonomous drivings, the border of the road should be detected in rather large image domains. Three of such possible domains are marked by the white rectangular frames in Fig. 2. They are referred in the following as window 1 (left), window 2 (middle), and window 3 (right). Each window has a size of 150 columns and 120 rows.

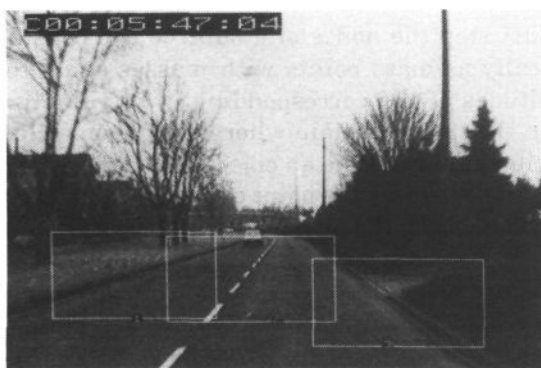


Figure 2. Image of an outdoor scene. In window 1, 2, and 3 borders of the road should be detected.

The first processing stage is the convolution with the normalized gradient of Gaussians which is done in

real time up to a 8 x 8 convolution kernel in our image processing system VISTA (Visual Interpretation System for Technical Applications, Paul et al. 88 [8]). The output of VISTA is based on a procedure which approximates the gray value gradient  $\mathbf{n}$  with components  $n_1$  and  $n_2$  [7]

$$\mathbf{n} = (n_1, n_2)^T = -\sigma\sqrt{2\pi}\nabla(G*f).$$

Here  $G = G(x,y,\sigma)$  is the 2-dimensional Gaussian with standard deviation  $\sigma$ ,  $f = f(x,y)$  the image function, and  $*$  the symbol for convolution. With the standard deviation  $\sigma = 1.0$  of the Gaussian one gets a good compromise between noise reduction and sufficiently high resolution. The output of VISTA consists of four images: direction and magnitude both of the gray value gradient and of its maxima in gradient direction. The latter are called contour points in the subsequent discussion.

The next step is to generate the histogram of gradient directions for all contour points in a window. The VISTA system delivers in real-time an estimate of the gradient direction covering the full 360 degree angular range quantized into 180 intervals of 2 degrees each. The histogram of these data for all contour points in a window is generally very noisy. It must be smoothed by a suitable low pass filter which is in our approach a 1-dimensional Gaussian. Such a smoothed histogram for the directions of the 2610 contour points in window 1 is shown in Fig. 3. The standard deviation of the 1-dimensional Gaussian was  $\sigma = 4$  degrees. The maxima are automatically detected. Gradient direction 66 degree has the largest amplitude and the opposite direction 242 degree yields the second significant maximum. Here the following angle convention is used: the gradient shows always from the darker to the brighter part at intensity changes, the direction 0 degree = 360 degree is the horizontal x-axis, and the angles are increasing clockwise.

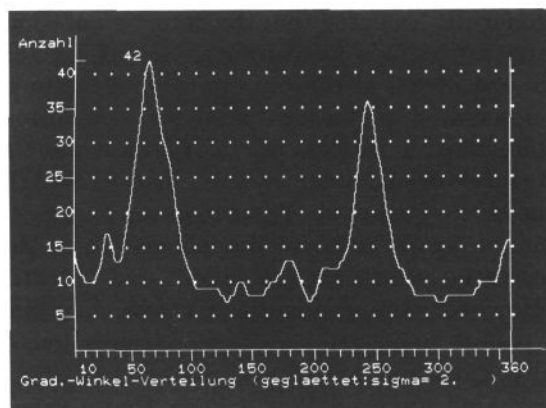


Figure 3. Smoothed histogram of the direction of the gray value gradient for all CP (Contour Points) in window 1 (see Fig. 4). Significant maxima are found at 66 and 242 deg.

In Fig. 4 all contour points of Fig. 2 are depicted. Pre-processing with the normalized gradient of Gaussian was performed with standard deviation  $\sigma = 1.0$ . All maxima with gradient magnitude larger than 10 are shown. If all contour points in window 1 are considered as mass points where each mass is given by the gradient magnitude, then the inner "rectangle of inertia" can be calculated by a method well known in physics. The axes of this rectangle are derived from the ellipse of inertia by taking the two principal axis after adjusting the larger axis approximately to the size of the point cluster leaving the ratio of the axes constant. In this way one gets easily robust parameters for the extent and the orientation of the cluster. The points of gravity of the rectangle and of the considered contour points coincide. The angle  $\delta$  of the larger axis with the (horizontal) x-axis is measured clockwise following our conventions for the gradient direction. This yields 184 degrees for the inner rectangle shown within the window 1 of Fig.4. In order to detect the direction of the road border some further steps are apparently necessary.



Figure 4. The gradient magnitude of all CP of Fig. 2 above the threshold 10 for a gradient operator with  $\sigma = 1.0$ . In window 1 the RI (rectangle of Inertia) of these CP (see text) is shown.

In Fig. 5 the RI (rectangle of inertia) is calculated for all contour points in window 1 with gradient directions in a  $\pm 16$  degree interval around the histogram maximum at 66 degree. The size of the interval around the histogram maximum is not a critical parameter. It determines only the number of contour points considered for the construction of the MST. For our purposes the width at half maximum amplitude of the most prominent histogram peak yields a suitable size. The direction of the larger axis is  $\delta = 168$  degree. As expected one gets a better approximation of the orientation of the road border. All points which are considered for the computation of a RI are depicted in the corresponding windows of Figs. 4, 5, 7, 9, 10. The contour points in the upper left part of the window shown in Fig. 5 are responsible for the shift of the center of the RI in this direction. The ratio 4.7 : 1 of the axes of the RI in Fig. 5 also indicates



Figure 5. Approximation of the left road side by the RI which is calculated out of all CP in window 1 with gradient directions  $66 \pm 16$  deg. (see Fig. 3).

a still rather large scatter of contour point characteristics contributing to the maximum in the direction histogram of Fig. 3. In order to select a subsample from this maximum such that the selected contour points belong to a cluster of straight line segments, additional criteria have to be considered.

For this purpose the distance of contour points with similar gradient directions and their difference of gradient magnitudes have to be taken into account. Among a lot of cluster analysis methods the MST method looks most promising to get some results comparable with perceptual grouping (see e.g. Zahn 71 [5]). The MST clustering approach is known to emphasize chain like structures as against e.g., sphere-like structures. For the 392 contour points with gradient directions  $66 \pm 16$  degrees a MST has been constructed. For our examples the weight  $w$  of the tree edges is given by

$$w = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (g_i - g_j)^2}$$

where  $x_i, x_j, y_i, y_j$  are the cartesian coordinates and  $g_i, g_j$  the gradient magnitudes of two contour points. The resulting MST has a mean edge weight  $m = 4.2$  with the standard deviation  $sd = \pm 5.2$ . Cutting at an "inconsistent" cluster joining-edge (see Fig. 1) whose weight is defined by  $c = m + sd = 9.4$  yields 30 subtrees. The position of the 166 nodes of the largest subtree are shown as bright structure along the road in Fig. 6. The corresponding RI is depicted in Fig. 7. The long axis with  $\delta = 160$  degree approximates very closely the road border. The ratio of the axes of inertia yielding a value 79, clearly indicates a line structure. All contour points except those on the road border have been eliminated by this approach.

The same procedure has been applied to window 3. The MST has been constructed from 494 contour points with gradient directions  $\pm 16$  degrees around 138 degrees which was the significant histogram maximum. The mean edge weight of the MST was  $m$

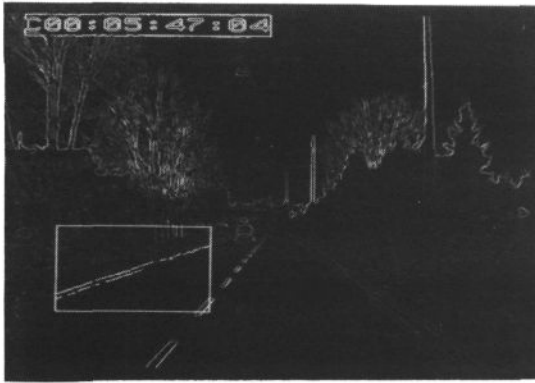


Figure 6. The bright structure along the left road side is the representation in the image domain of all nodes of the largest MST subtree of window 1 (see text).

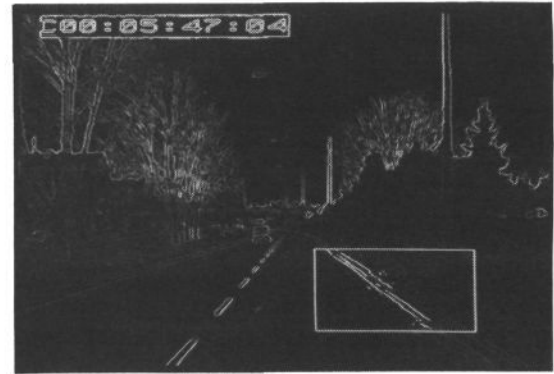


Figure 8. The bright structure along the right road side is the representation in the image domain of all nodes of the largest subtree of window 3.

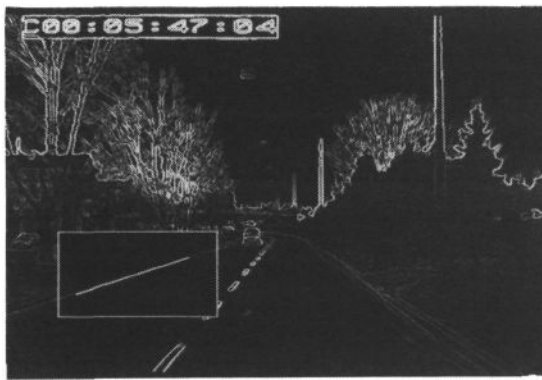


Figure 7. The RI in window 1 corresponding to the nodes in Fig. 6 is a close approximation of the road side.

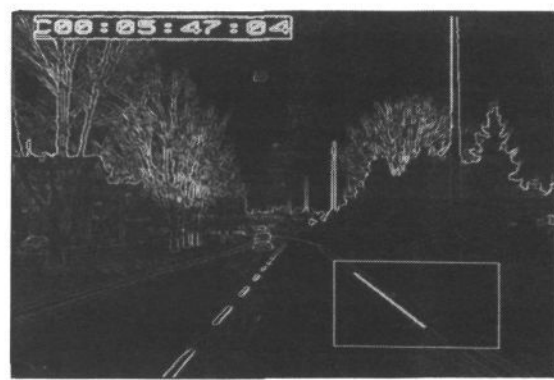


Figure 9. The RI in window 3 corresponding to the nodes in Fig. 8 is a close approximation of the right road side.

= 4.2 with  $sd = \pm 3.5$ . Cutting at  $c = 7.7$  yields 31 subtrees. The 249 nodes of the largest subtree are depicted in window 3 in Fig. 8 as bright structures. The RI with an axis ratio 81 and an orientation  $\delta = 48$  degree is a good approximation of the road border in window 3. More difficult is the representation of the dividing strips which are included in window 2 in Fig. 2. The individual markings can only be described by a single subtree if the cutting distance for an "inconsistent edge" is chosen much larger than the default values described above. In addition, out of the 1683 contour points in window 2 the 323 points in the intervals  $24 \pm 16$  degrees and  $210 \pm 16$  degrees around the two significant maxima at 24 and 210 degree of the histogram of gradient directions are selected for the construction of the MST. The tree has a mean edge weight 6.2 with a standard deviation 4.7. Applying a very large cutting distance, e.g. 30, one gets only two subtrees. The larger one has 287 nodes. Its corresponding RI with  $\delta = 119$  degree describes the location and direction of the dividing strips rather good as can be seen in Fig. 10.



Figure 10. The RI in window 2 corresponds to a MST subtree which contains 90 % of the CP with gradient directions  $24 \pm 16$  and  $210 \pm 16$  deg.

## DISCUSSION

In addition to the efficacy of our approach for the symbolic representation of rather complex structures in images, we think that the simplicity is another important aspect. We have already realized a hardware module in order to compute gray value gradients, contour points and their histograms. Using po-

sition and gradient magnitude of contour points, the construction of the MST is remarkably simple.

Besides these advantages the perceptual distance, i.e. the MST edge weightings of contour points whose direction can be selected arbitrarily out of one or more maxima of the corresponding histogram, can be controlled by an appropriate weighting function. The function proposed in this paper is only one of a variety of possible functions which should be adapted to some suitable perceptual model. The method is not restricted to the evaluation of straight lines but as well to grouping of chain like structures which may be curves. Dependent on the applications only the interpretation of the histogram of gradient directions must be adapted. These aspects are to our opinion important advantages compared with widespread clustering methods in image processing as for example the Hough transformation.

## ACKNOWLEDGEMENTS

The author thanks H.-H. Nagel for valuable comments on a draft version of this contribution. B. Engler and L. Berger have been excellent research assistants. The work has been supported by the German Ministry of Defense.

## REFERENCES

1. **Vistnes, R.:** "Texture Models and Image Measures for Texture Discrimination", *Int. J. Computer Vision*, vol. 3, 1989, 313-336
2. **Voorhees, H. and Poggio, T.:** "Computing texture boundaries from images", *Nature*, vol. 333, 1988, 364-367
3. **Mohan, R. and Nevatia, R.:** "Perceptual Organization for Segmentation and Description", in *Proc. DARPA Image Understanding Workshop*, Palo Alto, California, May 1989, 415-424
4. **Fan, T.-J., Medioni, G., Nevatia, R.:** "Recognizing 3-D Objects Using Surface Descriptions", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-11, 1989, 1140-1157
5. **Zahn, Ch.:** "Graph-Theoretical Methods for Detecting and Describing Gestalt Clusters", *IEEE Trans. Computers*, vol. C-20, 1971, 68-86
6. **Korn, A.:** "Combination of different spatial frequency filters for modelling edges and surfaces in gray-value pictures", in *Proc. Comput. Vision for Robots*, Cannes, France, Dec. 2-6, 1985, SPIE, vol. 595, 22-30
7. **Korn, A.:** "Toward a Symbolic Representation of Intensity Changes in Images", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-10, 1988, 610-625
8. **Paul, P., Hättich, W., Nill, W., Tatari, S., Winkler, G.:** "VISTA: Visual Interpretation System for Technical Applications - Architecture and Use", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-10, 1988, 399-407

