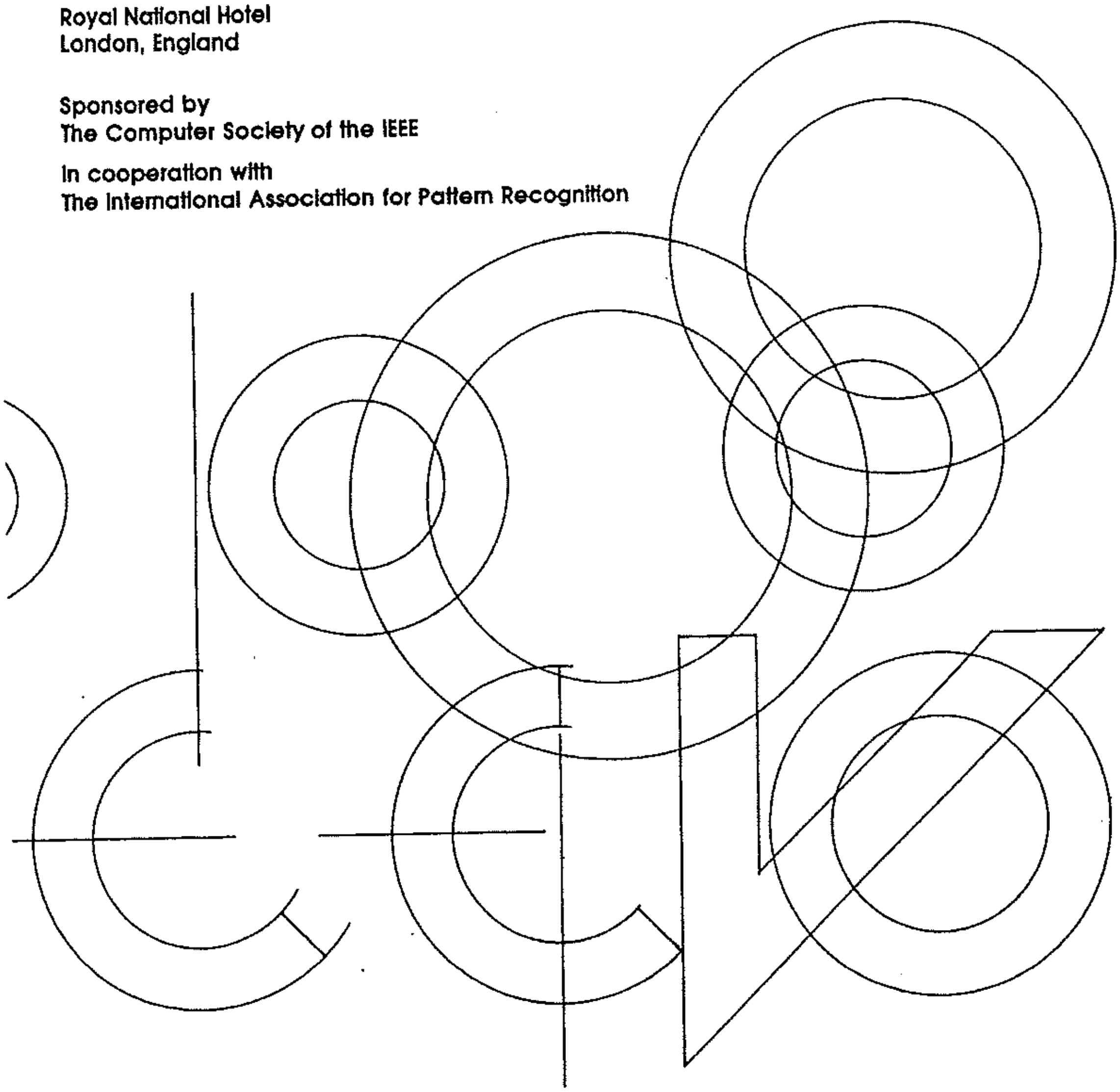
Proceedings

FIRST INTERNATIONAL CONFERENCE ON COMPUTER VISION

June 8–11, 1987
Royal National Hotel
London, England

Sponsored by
The Computer Society of the IEEE

In cooperation with
The International Association for Pattern Recognition



Computer Society Order Number 777
Library of Congress Number 87-45359
IEEE Catalog Number 87CH2465-3
ISBN 0-8186-0777-7
SAN 264-620X

THE COMPUTER SOCIETY OF THE IEEE

IEEE THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

COMPUTER SOCIETY PRESS

1st ICCV 1987 / pp 112-117

LINKING IMAGE-SPACE AND ACCUMULATOR-SPACE: A NEW APPROACH FOR OBJECT-RECOGNITION

G.Gerig
Institut für Kommunikationstechnik
ETH-Zentrum, CH-8092 Zürich, Switzerland

Abstract

Images of bad quality require robust, global techniques for object-recognition. The Hough technique for detection on curve-like image characteristics is well known (Hough [1], Duda [2]). Parallel hardware, alternative data structures (O' Rourke [3]) and projection techniques increase the efficiency of such methods, but the problem of interpreting the resulting accumulator remains. In this paper we discuss a new backmapping technique, which significantly reduces the complexity of evidence information in accumulator space, and which uniquely links specific points in image space to most evident locations in accumulator space. The interpretation is substantially simplified and the accuracy of recognition increased. Furthermore the global grouping of image points through common relations to evident accumulator cells offers the possibility of applying much more complex interpretation strategies.

The advantage of the new technique is illustrated by applying the method to three completely different recognition tasks in biomedical and medical imagery, where the link between points in image space and accumulated evidences is used in very different ways.

Keywords: Object-recognition, Hough techniques, matched spatial filtering, backmapping, medical image processing

1 Introduction

In real applications we are often confronted with images of poor quality or images where objects are embedded in a confusing context and therefore hard to detect. Such difficulties may be reduced by choosing optimum imaging and lighting conditions or by rearranging the scene content, such as in industrial scenes. For many image categories, however, this is not possible. Typical examples are biomedical images or medical tomographic data from both MR and X-ray CT, where well known segmentation techniques often fail to work due to coarse resolution, low signal to noise ratio, unsharp object boundaries, overlapping structures, and variations in background. Nevertheless there is an increasing need for cost effective automation of recognition tasks, even under difficult conditions to obtain results equal to or better than obtained by visual inspection.

Object recognition in images of bad quality requires robust, global techniques to group characteristic image points to desired structures. Well known techniques are template matching , Hough transform and extended Hough techniques. They all convert the global pattern detection problem into a local problem, i.e. finding local maxima in accumulator space or (equivalently) assigning best matching templates to sets of points in image space. If templates are represented by curves there is a strong relationship between template matching and Hough transform (Merlin,Farber [4], Sklansky [5], Turney et.al. [6]). Hough techniques show several advantages over simple matched filter techniques. They usually involve a larger set of parameters (position, size, orientation), and are applied to a limited set of characteristic image points, which represent the objects by their boundaries. The selection of significant boundary points preceding the matching process may be seen as a first introduction of a priori knowledge about the objects, and therefore leads to much more evident correlation maxima.

A previous paper (Gerig,Klein [7]) presented an efficient implementation of the Hough technique and proposed a new backmapping method to simplify the interpretation of the accumulator space, resulting in a link between points in image space and accumulator cells. Here, this new approach with its inherent ability to apply more complex interpretation strategies is discussed in more detail and extended to the detection of curves which are not given in analytic form. Additionally it is shown, with the new application of line-detection, that accumulated evidences do not always suffice to recognize and locate objects, but a feedback from accumulator cells to contributing boundary points can help overcome such basic limitations.

2 Efficient implementation of Hough techniques

Hough techniques allow the detection of predefined curves through evidence accumulation in parameter space. The set of all possible parameter vectors associated to each boundary point describes a hypersurface in accumulator space, and each cell intersected by the hypersurface is incremented. The transformation of each contour point into a hypersurface in accumulator space is a one-to-many mapping to a set of parameter vectors (accumulator cell locations), which represents the set of curves with possible membership of the contour point. The maxima in Hough space, i.e. the optimum parameter vectors, describe the best fitting curves in image space. Such a method of 'trying all the possibilities' has several important draw-backs:

- the dimension of the parameter space can become very high (i.e. number of parameters of analytic curves/ five-dimensional parameter space for general nonanalytic curves (reference point x,y, scaling x,y, orientation))

- the computational effort is proportional to the number of boundary points in image space (each boundary point transforms into a hypersurface)
- evident counts in accumulator space do not necessarily correspond to existing curves in image space, but can also be generated by accidentally arranged image points (each boundary point may contribute to several evident accumulator cells)

To substantially increase the efficiency, the Hough technique has been *implemented in parallel* (Merlin/Farber [4]). The overall computational expense thus becomes independent of the number of curve-points in image space. The *search range* on the parameter axes can be *constrained or quantized* by introducing a priori knowledge about parameter values. We have developed a projection technique to avoid maintaining a high dimensional parameter space (Gerig,Klein [7]). The simple assumption of excluding concentric objects allows us to *reduce the dimension of the parameter space,* because only one maximum accumulator count must be kept at each reference point location (representing the most likely occurrence of an object-curve related to that position). The 2-dimensional accumulator plane of possible reference point locations is frozen, while the other parameter axes (e.g. scaling, orientation) are condensed by applying the projection strategy (propagation of most evident parameter cell along the parameter axis). Two planes, one containing accumulated evidence and the other the corresponding coded parameter-vector, replace a higher dimensional accumulator space and allow more efficient storage and retrieval of evidence information.

When implementing a Hough technique, two further facts must be considered: Boundary points only represent fuzzy estimates of original object contours due to discreteness and insufficient localisation accuracy of boundary-detection algorithms. Shapiro [8] analyzed this effect on the mapping operation and proposed to use template curves of given width. In our parallel implementation, this can be done efficiently by smoothing along the 'size-parameter'-axis of the accumulator space. The choice of an appropriate incrementation strategy (increments to accumulator cells) greatly influences the result of the Hough transform, it allows local properties of boundary elements to be taken into account (Ballard [9]). Usual strategies are using increments proportional to the edge strength and/or increments inversly proportional to the perimeter of the template curve to quantify the degree of completeness for the match of image curves of varying size. These methods speed up the matching process and increase reliability of information in accumulator counts, but one major problem remains: Finding a good and reliable strategy for interpreting the accumulator space.

3 Backmapping

Cells of maximum counts represent a kind of probability for the occurrence of associated curves in image space, but counts are also generated by accidental arrangement of boundary points originating from several objects. The problem of distinguishing between true clusters and false clusters remains to be solved, a difficult problem since all information about contributing feature points has been lost. Usually simple thresholding, cluster detection or relative maximum detection algorithms are used, techniques that work exclusively in the accumulator space and therefore do not allow any feedback to contributing feature point data. They may be sufficient in simple cases but do not yield satisfactory results in more complex scenes. A grouping process using the Hough technique is robust and global,but the subsequent interpretation is often weak and fragile because of the large complexity of the information in the accumulator space.

To simplify the interpretation of the accumulator space and to solve more complex recognition tasks, it is often necessary to know *where* the contributions to evident accumulator counts come from. Here a new method is proposed to link the accumulated evidences and the contributing boundary point data. The idea is based on the simple assumption that *a given boundary point should be a member of only one curve* in image space, obviously of the most evident one (corners and crossings are not considered here). The most evident curve membership that may be assigned to each boundary point is defined by an associated accumulator cell of highest count and therefore by an optimum parameter set (a_j):

$$\forall x_i \quad (x_i \quad \in \quad f(x, a_{opt}) = 0\ ) \tag{1}$$

x	:	feature point
a	:	parameter vector
$f(x,a)$	:	parametric curve

The Hough procedure (2) is used a second time, but now instead of incrementing cells intersected by each hypersurface (3) the cell of highest count (A_{max}) lying on a hypersurface is found (4). The location of such a most evident cell (a_{opt}) will be related back to each contributing boundary point (5), and only this cell location will survive in the accumulator space. We now obtain a unique relation between each boundary point in image space and one optimum parameter set in accumulator space, relations are stored as pointer information assigned to each boundary point.

Hough transform:

$$\forall x_i \quad \left(x_i \quad \rightarrow \quad \{a \mid f(x_i, a) = 0\}_i\ \right) \tag{2}$$

$$\forall a_j \ \in \ \{a\}_i \quad : \quad A(a_j) = A(a_j) + inc \tag{3}$$

Backmapping:

$$A_{max,i}(a_{opt,i}) \quad = \quad max\{A(a) \mid f(x_i, a) = 0\}_i \tag{4}$$

$$\forall x_i \ (\ x_i \quad \leftarrow \quad a_{opt,i}\ ) \tag{5}$$

x_i	:	boundary point
$\{a\}_i$	:	hypersurface associated to x_i
$a_{opt,i}$	:	optimum parameter vector
$A_{max,i}$	:	accumulator cell of highest count

The nonlinear operation of assigning the location of maximum count to each of the boundary points reduces the one-to-many mapping of curve-points in image space to parameter space to a unique one-to-one mapping. This allows keeping a pointer information at each image point location as an additional attribute. Many boundary points are related to one parameter cell,

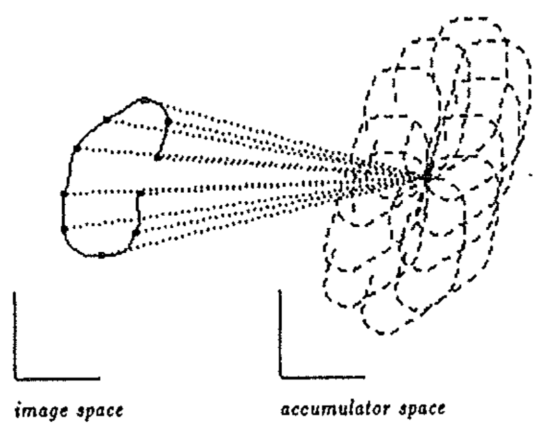


Figure 1: Backmapping of most evident counts

because maximum evidence is generated by the contribution of a set of points in image space. The accumulator space after backmapping therefore represents a much simpler description of the curve-structure in image space. Fig. 1 shows the advantages of linking image space and parameter space by looking at the first application of chapter 5, where nonanalytic curve templates are matched:

- If boundary points form a desired curve in image space, their hypersurfaces intersect at a common location in parameter space, where maximum evidence is accumulated. The backmapping process assigns this optimum cell location to each of them, resulting in a global connection of boundary points.
- The complexity of the accumulator space is significantly reduced, because only the cell location of maximum count of each hypersurface survives (black center), the remaining counts (dashed lines) are deleted.
- False maxima in accumulator space generally are less important than true maxima and are therefore suppressed (image points contributing to false maxima generally are also associated with more evident true maxima).

4 Interpretation

The backmapping results in a surviving accumulator space of significantly reduced complexity where each cell globally connects contributing boundary points by pointers. This structure represents a symbolic scene description and offers a variety of possibilities for the subsequent interpretation, e.g. it may act as a database for rule-based recognition tasks.

In simple situations the surviving accumulator counts will suffice to detecting objects, with the advantage that false maxima are suppressed and that surviving counts contain correct information about the number of contributing original image points. In more complex situations the pointers between contour points and parameter space locations may be used to apply more sophisticated interpretation strategies, because now a feedback to the original image space is possible. Examples of strategies are split/merge of nearby evident accumulator cells, or strategies for solving situations where different curves in image space accumulate exactly the same evidences. Fig. 2 shows examples of possible situations occurring in the case of circular boundary detection (second application of chapter 5). Ideal circles may be recognized by looking for surviving accumulator counts of high counts. In more complex situations this is not sufficient. First it is impossible to distinguish between overlapping objects and distorted objects, as they show exactly the same appearance in accumulator space (fig. 2 b,c). Secondly, it is not possible to decide whether the curve is noisy or partly visible, as they both accumulate the same evidences (fig. 2 d,e). However the additional use of the pointer information allows a correct classification, for example the two evident centers in case (2 c') describe one distorted curve and therefore can be merged, but both centers in (2 b') relate to different complete circles (uniform radial distribution of pointers).

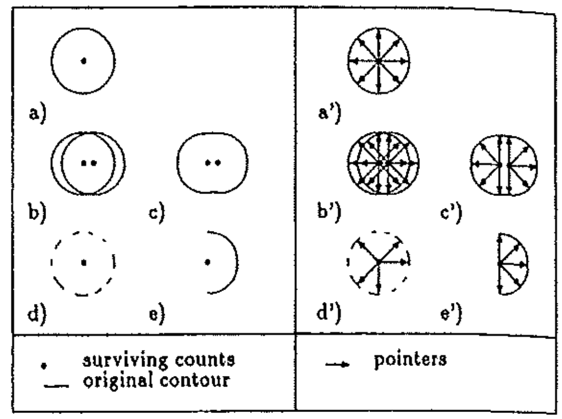


Figure 2: Interpretation using the pointer structure

5 Examples

The proposed method is illustrated by describing three applications. Two of them are typical examples of images that suggest the use of a robust technique for grouping curve-like image points according to a known object description, the link between accumulator space and image space is used for refining the match and for increasing the classification accuracy. In the third example the use of a Hough technique alone is not sufficient for detecting the objects, the matching must be followed by a feedback to contributing image points.

The first problem (fig. 3A) arises from analyzing time series of medical scenes. The objects that are to be tracked along a time sequence of MR images (32..64 frames) move and rotate between image frames due to breathing. To trace the effect of medicaments over a period of time,exactly the same subarea of the organ over the whole sequence of images must be compared. The templates of the left and the right kidneys are manually extracted from stable features of one typical frame, here represented as intensity valleys. They serve as curve-models for nonparametric matching. Assuming rigid body motion, the object size parameters may be frozen, but the orientation and translation parameters can vary within certain limits. The matching and backmapping results in motion and rotation parameters which are fed into a hardware warper to obtain a geometrically corrected series of images. As a further application surviving cell counts, together with their related feature points within a certain neighborhood of the absolute maxima, may be used to refine the warp coefficients allowing higher order geometrical fit (fig. 3B). The accumulator space (fig. 3B b,c) clearly shows that the evidence information is split into four separate maxima,corresponding to four subparts of the image curve. The pointers now allow a computation of the distortion of the image curve with respect to the template curve. Such a refining step becomes necessary when the assumption of rigid body motion is incorrect for organs distorted through stress. There is a great need for applying automated procedures as the number of sequences is growing rapidly and accurate analysis can no longer be efficiently obtained by interactive means.

The second type of scenes are electron micrographs of a thin section through plastic material (fig. 4a shows a subframe of a scene containing many difficulties). The recognition task is to find the number and size of the circular objects. Basic features are edges, but the recognition is impeded by gaps, partly invisible or distorted contours, overlapping objects and additional edge segments not belonging to the object boundaries. We have applied a three parametric Hough transform for detecting circular curves of unknown size, using parallel implementation, projection technique and backmapping. As most of the objects are not represented as ideal circular curves the surviving accumulator counts are used together with the pointer structure to recognize the objects hierarchically (using the strategies illustrated in fig. 2). Additional rules such as 'objects are darker than background' help to increase the classification accuracy, which reaches 96 percent even in the most difficult real scenes of that type and seems to surpass the results of human interpreters.

As a third example we applied our methods to the problem of line detection (fig. 5). It is often proposed to use a Hough transform for detecting lines (Hough [1], Duda [2]), but a well known problem is that lines go to infinity. Evidence is accumulated for linearly arranged image points, but there is no information about the position or topological appearance of contributing line-elements (e.g. contributions may come from one line, from several independent collinear lines (as in fig 5c) or from a dashed line). The proposed feedback from evident counts in accumulator space to feature points in image space (fig 5c illustrates the pointer structure) is a possible solution to obtain the additional knowledge . Fig. 5b shows the surviving counts after applying the backmapping (left) and the reconstruction result (right), where the pointer information is used for determining the line position. There are some falsely classified pieces due to the still imperfect strategy of breaking the contributions into continuous pieces. Again the example demonstrates the enormous reduction of complexity of the accumulator space using our new backmapping strategy and the advantages offered by the link between image space and accumulator space.

6 Conclusions

Hough techniques are widely used and have often been described in great detail, but overcoming the most difficult step - finding the correct number of true clusters - was still an open question, particularly as such robust methods are to be applied to real data where other methods fail. This paper shows that our new concept of backmapping and linking curve-points in image space to evident cells in accumulator space helps overcome many of the difficulties impeding successful interpretation, not only in one special case, but in the general use of different types of applications.

The usual interpretation strategy of recognizing objects by detecting local maxima in accumulator space is considerably simplified because of the reduced complexity of surviving accumulator space and the suppression of false maxima. In addition the feedback from evident cells in accumulator space to contributing feature points offers the possibility of applying more complex interpretation strategies, either for obtaining increased classification accuracy (by reconsidering early classification results) or even for recognizing structure configurations which could not be detected using simple maximum detection.

We obtained good results in detecting 2-D objects in real scenes, both with regard to cost effectiveness and recognition accuracy. In analyzing sequences of images or sets of biomedical scenes, we need automated procedures to extract object information out of a large number of frames, which cannot be efficiently obtained by visual inspection.

References

1. Hough P.V.C., **Method and means for recognizing complex patterns**, U.S. Patent 3069654, Dec. 18, 1962

2. Duda, R.O. and Hart, P.E., **Use of the Hough transformation to detect lines and curves in pictures**, Comm. ACM 15,1, pp. 11-15, Jan. 1972

3. O'Rourke, J., **Dynamically quantized Spaces for Focusing the Hough Transform**, Proc. of 7th IJCAI, pp. 737-739, 1981

4. Merlin, Ph.M. and Farber, D.J., **Parallel Mechanism for Detecting Curves in Pictures**, IEEE Trans. Comput., vol C-24, pp. 96-98, Jan. 1975

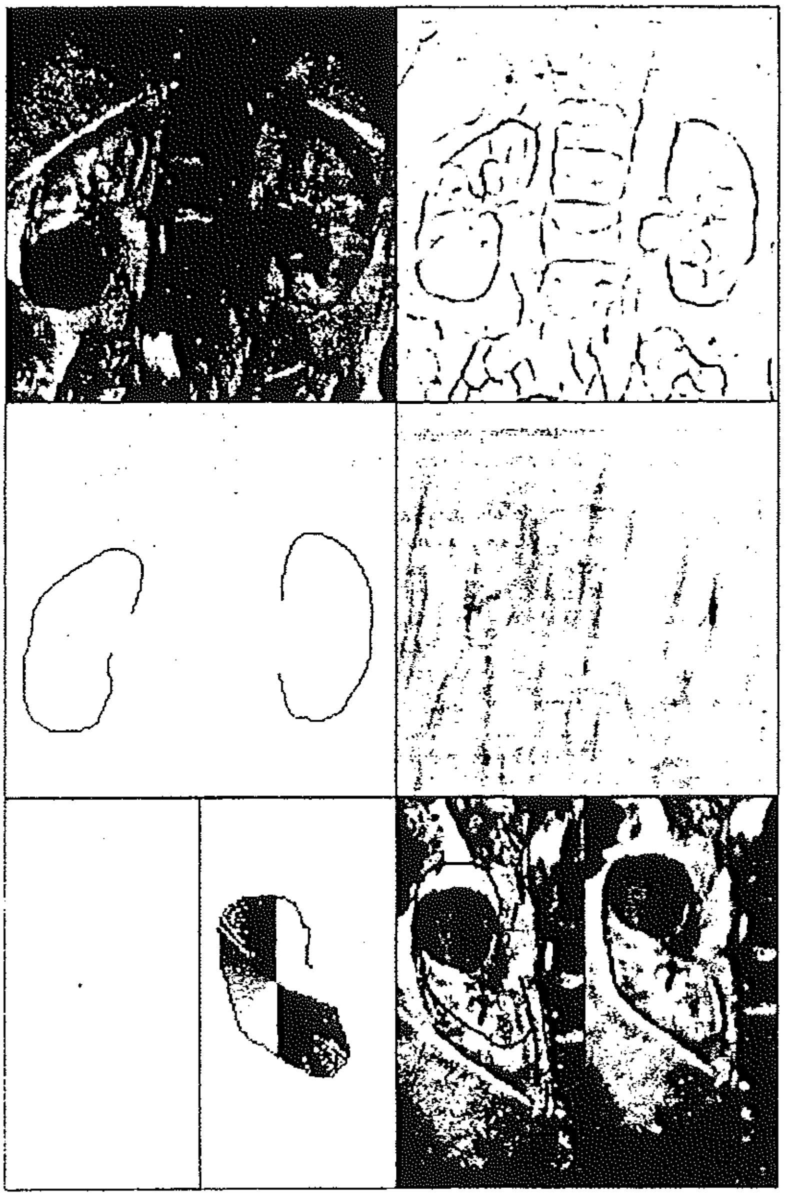
5. Sklansky, J., **On the Hough Technique for Curve Detection**, IEEE Trans. Comput., vol C-27, pp. 923-926, Oct. 1978

6. Turney, J.L., Trevor, N.M., Volz, R.A., **Recognizing Partially Occluded Parts**, IEEE Trans. Pattern Anal. Machine Intell., PAMI-7, pp. 410-421, 1986

7. Gerig, G. and Klein, F., **Fast Contour Identification through Efficient Hough Transform and Simplified Interpretation Strategy**, Proc. 8th Int. Conf. on Pattern Recognition, vol.1, pp. 498-500, Paris, France, Oct. 27-31, 1986

8. Shapiro, S.D., **Generalization of the Hough transform for curve detection in noisy digital images**, Proc. 4th Int. Joint Conf. Pattern Recognition, pp. 710-714, Kyoto, Japan, November 1978

9. Ballard, D.H., **Generalizing the Hough Transform to detect arbitrary shapes**, Pattern Recognition vol.13, No.2, pp. 111-122, 1981



a	b
c	d
e f	g h

fig. 3A: Detection of kidneys in MR images

a) original NMR-image of kidneys

b) extracted boundaries (valleys in intensity)

c) left and right template

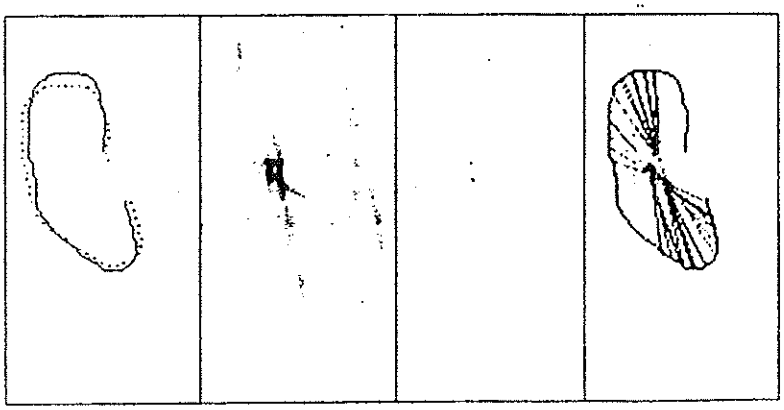
d) accumulator plane of evident reference point locations (orientation axis projected)

e) surviving accumulator counts after backmapping (left kidney)

f) pointer vectors directing from surviving counts to contributing boundary points (grey indicates pointer orientation)

g) overlay of original left image and left template

h) analogous to g), but image warped (translation x,y: -1,-13, rotation: -3 degree)



a	b	c	d

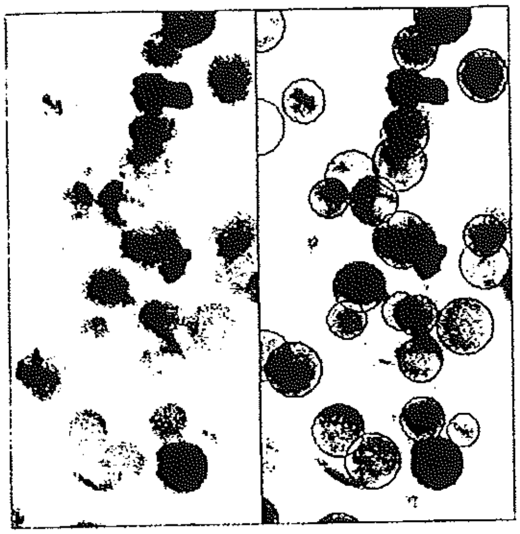
fig. 3B: Refinement of warp coefficients by using pointers

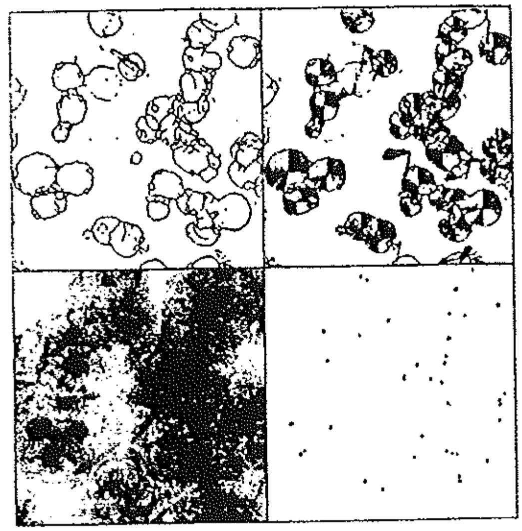
a) overlay of template function (dotted line) and distorted object (continuous line)

b) accumulated evidences of distorted curve

c) surviving accumulator counts after backmapping

d) (subset of) pointers linking surviving counts to generating boundary points (lead to additional scaling parameters scx,scy = 0.9,1.1)





a | b

c | d
e | f

fig. 4: Circular boundary detection

a) (most difficult) subframe of original image

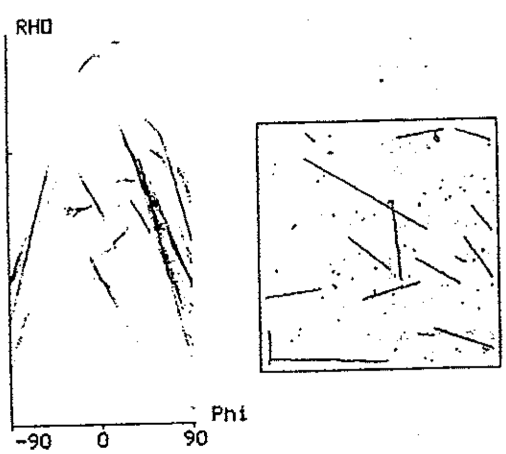
b) overlay of original image and classification result (black circles)

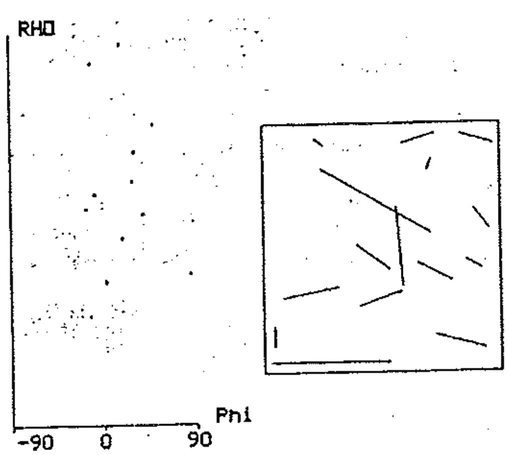
c) edge-features used for matching

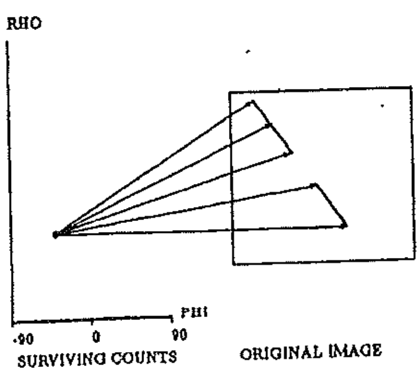
d) pointer vectors directing from surviving counts to contributing boundary points (grey indicates pointer orientation)

e) accumulator plane of evident center coordinates (scale-axis projected)

f) surviving accumulator counts after backmapping







a | b
c |

fig. 5: Detection of straight lines

a) original image (right) and corresponding accumulator space (left)

b) surviving counts after backtransform and reconstructed lines (minimum length 10 pixels)

c) example of link between surviving evident count and corresponding line-points