Determination of Global Properties in Document Processing

Zhangzheng Yu
Institut für Angewandte Mathematik
Universität Hamburg
Bundesstraße 55
D-2000 Hamburg 13
Germany

Gerd Maderlechner
Siemens AG
Zentrale Aufgaben Informationstechnik
Otto-Hahn-Ring 6
D-8000 München 83
Germany

Abstract. The aim of this paper is to concisely describe two effective and efficient algorithms for Hough transform and medial axis transform. New aspects of applying the transforms in document processing are also considered.

1. Introduction

Generally speaking, a picture processing system consists of a series of algorithms for corresponding different tasks. Some of them are applied to original images (preprocessing) to extract essential features which are then analysed and interpreted by some other algorithms strongly depending on the concrete context.

Algorithms for picture preprocessing are usually controlled by a few parameters. There are many demands on selection of the parameters in application of the algorithms. One of them is approxibility of an operator through another operator which is *simpler* but sufficient for application. On the one hand, the parameters must be so chosen as to reduce the great amount of the original image data. On the other hand, the parameters have to be so determined for one to get sufficient information for facilitating interpretation of an image. A mathematical formulation of the requisition reads, that an operator should be approximated by another operator which filters out unnecessary information of original images for the purpose of not overwhelming the interpretation system, but maintains sufficient information of the images in order to make distinquishing different classes of images through different features possible. Another yardstick to messure the goodness of an algorithm is its stability. An algorithm is stable if its application to a slightly different image yields a result which only slightly differs from the result of its application to the original image. Obviously, if a given algorithm is not stable, the interpretation of its results can become difficult and not reliable. Unstable algorithms are usually regularized in the sense of Tikhnov. A further important aspect of an algorithm is its computational efficiency. Even if an algorithm satisfies the other two criteria mentioned earlier, its inefficiency heavily limits its application.

Among a lot of algorithms often used in a picture processing system, we intend only to investigate Hough transform and medial axis transform in this paper. There is a great plenty of literature on the two transforms. It is well known, that the Hough transform is invertible under some suitale conditions. But the instability of the reverse transform hinders an effective interpretation of a transformed image. In addition, conventional implementations of the transform are very time-consuming. One of our intentions in this paper is to give a detailed description of a new implementation of the transform, through which the time consumption for its execution is reduced approximately to one eighth. The feasibility of establishing a theory on stability and regularization of its inverse transform is mentioned here but a comprehensive consideration has to be sacrificed because of limitation of this paper. The instability of medial axis transform can be easily verified. A suggestion for regularizing the transform is made and stability of the regularized transform is considered here. A new algorithm, which can be very easily parallelized, is also given.

2. Hough transform

For a given image, the Hough transform, sometimes also called Radon transform, assigns to each line the integral over the image along this line. There exist several variants of the transform depending on the parametrization of lines and on the approximation of the line integrals. There are also generalizations of the transform for the purpose of detecting some geometrical patters with given description through matching methods.

The transform has been studied by some authors. It is well known, that the Hough transform can be inverted, indeeed, an explicit formula for the inverse is known (see e.g. Deans, Natterer). It is auch known, that this inverse, if considered in a functional space, is not bounded. Actually, the formula for the inverse transform contains the operation of derivation which is obviously not bounded or stable. Nevertheless, the transform has found applications in diverse fields. Especially in binary picture processing, the transform yields a stable and robust means for interpretation of image contents. The apparent contradiction was more closely investigated since the question was considered to be of importance and should be settled before practical implementation of the method.

A binary image can be interpreted as a set in the plane (the set of all black points in the image). Any sufficiently *regular* set is uniquely determined by its boundary. In the image space of the Hough transform a boundary curve is represented by the set of all tangents of it (see Eckhardt, Scherl, Yu). Using Hausdorff metrics, we can establish a theory on the invertibility and stability not only of the transform itself but also its inverse transform, if the boundary of a planar compact set can be approximated through convex curves. Conventional versions of the transform were theoretically analysed on continuous images but computatioally implemented on discrete images. The discrepancy arising in this precarious situation is surmounted by the new approach. The theoretical details of this approach will be reported elsewhere.

Deriving from the theory, we can easily construct a corresponding algorithm whose application is effectively and efficiently. As an asumption, we are given a discrete binary image. The steps of the algorithm are:

a. Extract the boundary of the image using any known algorithm depending on the underlying topological concept;

b. Run along the boundary curve and store the boundary points according to a given direction;

c. Starting from an arbitary boundary point, run along the boundary if the Hausdorff distance of the line segment joining the starting point and the moving point to the boundary part delimited by the two points not greater than a prescribed nonnegative parameter characterizing the goodness of the approximation. Choose the last possible point satisfying the condition as the ending point. Using the ending point as starting point for the next approximating line segment, repeat the preceding process until the whole boundary is approximated by a series of line segments.

d. Ascribe the length of each line segment to the point in Hough space whose coordinates characterize the line overlapping the considered line segment.

The experience made by us has shown that we can get a time reduction of not less than factor eight compared with known conventional algorithms. The Hough space produced through the algorithm is in addition sparsely occupied. This fact makes the serch for maximal values in Hough space very easily and effectively.

3. Medial Axis Transform

The plane is assumed to be equipped with Euclidean metrics. We denote with $K(x,r)$ a circle of radius r and centre point x. Let G be a compact set on the plane. A point x belongs to the medial axis of G, if there exists a positive r such that $K(x,r)$ lies entirely in G and has at least two common points with the boundary of G. The medial axis of G will be denoted with $M(G)$.

It is obvious that the set G can be reconstucted from $M(G)$ under some *regular* conditions if for each point of $M(G)$ the radius of the corresponding maximal circle inscribed in G is stored simultaneously. The so produced medial axis has useful properties. For example, the medial axis represents the connectedness property of the transformed set in form of a graph. By means of this transform, one can reduce data valume of a binary image to its *essence* which facilitates interpretation of the image. The medial axis of a planar set can be further tuned (for instance, by pruning end points of the graph) for application in image segmentation.

As one can read in literature, the construction of medial axis of a planar set is ill-posed. That means, little deviations of a set cause great change of its medial axis. This instability is naturally not desirable. So there is necessity for regularizing medial axis transform.

A formal regularization theory on medial axis transform and a concrete realization on a computer are envisioned in the present paragraph.

In order to make statements on the continuousness of the regularized medial axis transform, we introduce a metrics being known as Hausdorff metrics on compact subsets in the plane. For a compact set G and a positive number t, we define the t-parallel-set $G(t)$ as the set of all points x in the plane for which there exists a y such that the distance between x and y is not greater than t in the sense of Euclidean metrics. The Hausdorff metrics between two compact sets G_1 and G_2 is then defined as

$$d_H(G_1, G_2) = inf\{t : G_1 \subset G_2(t)\, and\, G_2 \subset G_1(t)\}.$$

Given two nonnegative numbers p and q. A point x belongs to the regularized medial axis if there exists a r not less than p such that $K(x,r)$ lies entirely in G and has at least two common points with the boundary of G whose Euclidean distance is not less than $2r \cdot sin\frac{q}{2}$, that means, the angle between the two rays joining x with each of the two common points respectively is not less than q. The set of all points of the regularized medial axis of G will be denoted with $M_{p,q}(G)$.

The regularized medial axis transform has a lot of useful properties some of which are only cited here.

Monotony: If p_1 is not greater than p_2, and q_1 not greater than q_2, then $M_{p_2,q_2}(G)$ lies in $M_{p_1,q_1}(G)$.

Approximation: $M_{0,0}(G) = M(G)$ and for each x out of $M(G)$ there is nonnegative p and q such that x lies in $M_{p,q}(G)$.

Especially significant is the *stability* of the regularized medial axis transform.

Theorem: Given two compact sets G_1 and G_2 in the plane whose Hausdorff distance and the Hausdorff distance between the two boundaries of the two sets are not greater than a positive number δ. For predetermined parameters p and q, the regularized medial axis $M_{p,q}(G_1)$ lies in the $\frac{1+e(\frac{\delta}{p})}{2sin^2\frac{q}{4}}\delta$-parallel-set $M(G_2)(\frac{1+e(\frac{\delta}{p})}{2sin^2\frac{q}{4}}\delta)$ with $|e(\epsilon)| \leq \epsilon + o(\epsilon)$. Furthermore, there are numbers a, b und c, such that the regularized medial axis $M_{p,q}(G_1)$ lies in the $c\delta$-parallel-set $M_{ap,bq}(G_2)(c\delta)$ of the medial axis of G_2 regularized with the parameters ap and bq.

A proof of that theorem has to be omitted here because of limitation of the paper.

The two parameters p and q in the regularization can be manipulated. If we need only rough information of a binary image, greater parameters p and q should be used. If finer information about the original image is needed, one can make the parameters smaller depending on the context. So there arises a natural hierarchy of local as well as global information about the processed image.

The regularized medial axis transform can be very easily and relatively efficiently (in sense of sequential processing) executed. For implementation of the transform, all pixels in the image, starting from the original point of the coordinate system, are stored clockwise in a one-dimensional list corresponding to increasing radius. If one has evaluated the distance between one point of the set to be transformed and the boundary of the set, then the distance between one neighbour of this point and the boundary of the set can be easily estimated and exactly calculated by means of trying out not more than three radii.

It is emphasized here that the algorithm has a structure adapting sequential as well as parallel processing, because the transform is executed line by line (or row by row) independently. In the picture displayed at the end of this paper, the regularized medial axis of the backgrund is drawn in.

4. Conclusion

We have hitherto discussed feasibility of regularizing Hough transform and medial axis transform. The two transforms can be combined in natural ways. For instance, one can calculate at first the medial axis of the backgrund of a document and then use the regularized Hough transform which delivers the line structure of the document. The transforms were programmed on $VAX7800$ of Siemens Reserch Center in München-Perlach. Considering only the backgrund of a binary image with limited support, it is interesting for us to observe that medial axes produced according to different metrics and whose end points having been pruned have the same connectedness property. That is also easy to verify. This topological invariance of medial axes can be utilized if one is only interested in segmenting a document by means of topological instruments.

5. Acknowledgements

One of the authors (Yu) expresses his special thanks to Prof. Dr. U. Eckhardt who has supporting him continuously as well as generously, and is also very grateful to Dr. E. Hundt for his giving him the opportunity to utilize technical facilities in Siemens Reserch Center in München-Perlach.

5. Literature

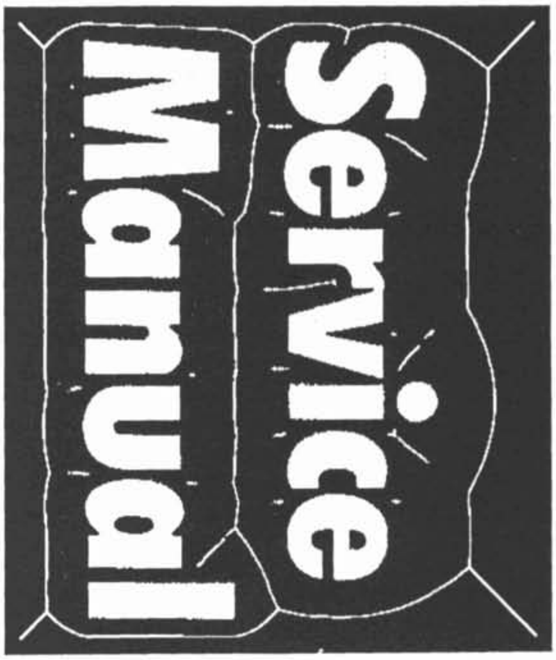
Serra J: Image Analysis and Mathematical Morphology. Academic Press. 1982

Eckhardt U, Scherl W, Yu Z: Representation of Plane Curves by Means of Descriptors in Hough Space. Rechenzentrum der Universität Hamburg. 1987

Deans S R: The Radon Transform and Some of its Applications. Jone Wiley & Sons. 1983

Natterer F: The Mathematics of Computerized Tomography. John Wiley & Sons. 1986

Tikhnov A N, Arsenin V Y: Methods for Solving Ill-posed Problems. John Wiley & Sons. 1977



$p = 10, sin\frac{q}{2} = 0.8$