Segmentation by Minimum Length Encoding

Daniel Keren and Ruth Marcus and Michael Werman and Shmuel Peleg*

Department of Computer Science
The Hebrew University of Jerusalem
91904 Jerusalem, ISRAEL

Abstract

A digitized waveform is approximated by segments whose total description length is minimal for a given error bound. This approximation can be computed efficiently, and can be used for segmentation. The method is also shown to be applicable for segmentation and edge detection in gray level and range images.

1 Introduction

Minimal length encoding as a method for describing and explaining information is not new. William of Occam in the 14'th century already stated the principle known as *Occam's Razor Principle*: *Entities should not be multiplied beyond necessity*. Of all possible descriptions of an object the shortest description is preferred. Description length is addressed by Kolmogorov complexity [4], and also appears in Valiant's learning theory and Rissansen's minimum description length principle [8].

We use this principle to segment waveforms into simple segments. These segments can be encoded by a short description, and also provide a good and meaningful segmentation of the waveform. We also show some applications using one-dimensional methods to segment two-dimensional gray scale and range images.

Waveform segmentation has been addressed in [6] and a specific type of image segmentation using minimal description length is described in [3]. A complete description of the proposed approach appears in [5].

2 Segments

Waveform approximations will be built from *segments*, which are members of a parametric family of functions $F = \{f_t\}$. In our studies we used polynomials or trigonometric functions as the family of functions.

A digitized waveform is given as a sequence of data points, $\{x_i, y_i\}_{i=1}^n$. We wish to describe the waveform with a segment, which is a function in F restricted to the interval $[x_1, x_n]$. There is usually no function f_t in F such that $f_t(x_i) = y_i$ for all i. A segment is therefore defined

---

acceptable if it approximates (in some norm) the given set of points.

The norm used in this paper is based on the χ^2 distribution: a maximal likelihood estimator, which represents the probability that the data originated from a function f in F. Segments are acceptable when the incomplete Gamma function is above a threshold. Acceptability becomes easier as the threshold is lowered Other possible norms are maximum absolute deviation, sums of absolute deviations, fixed tolerances or robust statistics [7, 9].

3 Waveforms

A *waveform* is defined as a concatenation of acceptable segments. There is more than one possible waveform representing a given sequence. Indeed, one can choose every single point (x_i, y_i) as a segment, or in the other extreme fit one member of F to all points. However, it is desirable to choose as few segments as possible, each having a simple form. Segments are therefore weighted according to the minimal information needed to encode them.

Given a sequence of data points, each possible continuous subsequence is being represented by a segment which is the simplest acceptable function approximating this subsequence. In case of polynomials, for example, this will be the lowest-order acceptable polynomial. A directed graph $G = (V, E)$ is then defined, V being the set of acceptable segments and $(v_1, v_2) \in E$ whenever subsequence v_2 immediately follows v_1. The vertices V are weighted by the cost of encoding the corresponding segment. A minimum length encoding of the entire waveform is then equivalent to a minimum weight path in G, with total encoding cost equal to the sum of weights of the segments in this path.

Figure 1 is an example of the segmentation process where F is a family of polynomials. Polynomial segments are described by the length of the segment, the degree of the polynomial and the polynomial coefficients. Total cost of a polynomial segment is therefore its degree plus two.

Computing χ^2 and recovering the best fitting segment for each of the $\binom{n}{2}$ intervals takes no more than $O(n^2)$ time because they are both functions of the smaller moments of the data points, $\sum x_i, \sum y_i, \sum x_i^2, \sum y_i^2, \sum x_i y_i, \cdots$. It takes only a constant time to compute the smaller moments for the points from i to $j + 1$ given the moments for the points from i to j. Computing the lowest weight path takes
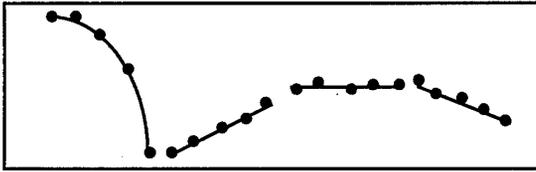
Figure 1: Segmentation with a family of polynomials.

Data points were generated from four polynomial segments with added Gaussian noise. The segments from left to right are: second order, first order, constant and first order. The costs of encoding each segment are 5, 4, 3 and 4 respectively, giving total cost of 16 for the entire waveform.

$O(n^2) \log n$ computation time. This is the worst case, but if there are not many long acceptable segments the computation complexity can approach $O(n)$ by not trying to merge two barely acceptable segments into a longer one.

3.1 Comparison with Split-Merge

A popular approach for waveform segmentation is binary sub-division [6], which attempts to fit a function in F to all data points. If this fails, the data points are divided, and the function fitting attempt is continued recursively on each half. After splitting stops, similar adjacent segments are merged according to some acceptability criteria.

Results of applying this process at two different acceptance thresholds are shown in Figure 2. Because the splitting process is usually not consistent with the structure of the segments, and because the process is not guaranteed to find a representation which is *minimal* in any sense, the minimal length segmentation achieves better results.

4   2D Segmentation

Minimum length encoding cannot be computed in two dimensions as its computational complexity is of the class NP-Complete [2]. We extend the minimal length encoding to higher dimensions by working on one dimensional projections. Since the intersection of a d'th degree bivariate polynomial surface with a plane is a d'th degree univariate polynomial, one dimensional polynomials are usually used. 2D polynomial surfaces are suitable for describing optical flow [1], luminance functions, and range data of man made objects.

Two-dimensional segmentation is obtained from one-dimensional segmentations of cuts in two orthogonal directions. Taking at least two different directions is necessary to detect all discontinuities in the image. Whenever an edge becomes almost parallel to one of the directions, it can be detected by segmentation of a cut in the other direction; see Figure 3 for clarification. Cuts were made in two types of orthogonal grids: a dense grid and a sparse grid. In both grids the cuts were taken parallel to the x and y axes.
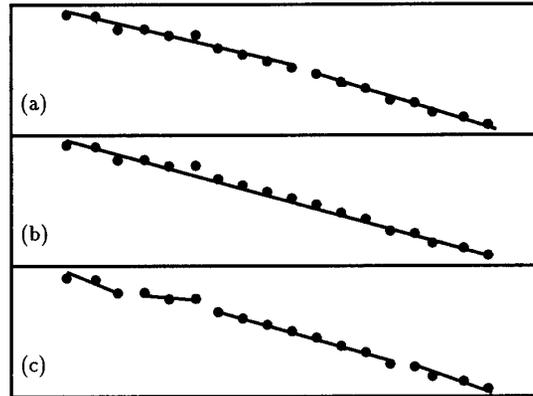

Figure 2: Comparison of minimum length encoding to binary sub-division: Data points were generated from two linear segments with added Gaussian noise.
a) Successful segmentation of minimal length encoding.
b) Binary sub-division: at split-merge threshold of 0.35, the two segments were not separated.
c) Binary sub-division: at split-merge threshold of 0.3 too many segments are obtained.

A dense grid includes all rows and columns of an image. Combining all the discontinuities found by the segmentation of each row and each column yields an edge-map. These edges can further be linked together to produce a partition of the image.

A sparse grid includes only part of the rows and columns of an image. Discontinuities found in the one-dimensional segmentations can be used to insert break-points in the two-dimensional grid at the locations of the discontinuities. This defines a set of connected components on the grid. Each connected component corresponds to a different region in the image which can be described by a smooth surface, because the borders between segments in the one-dimensional cuts correspond to edge-pixels in the two-dimensional data. These connected components can serve as seed-regions in a region-growing method [10]. Figure 4 demonstrates the use of a sparse grid. Using a sparse grid is faster than using a dense grid as not all rows and columns of the picture are segmented. On the other hand, some edges may not be detected using a sparse grid.

5   Experimental Results

The input images are assumed to be noisy, quantized samples of a piecewise-smooth function. The first examples use a synthetic image to demonstrate the ability of the algorithm to detect very low contrast edges.

Figure 3 shows the application of a dense grid. This synthetic image has a square object with a background composed of two constant grey-levels. The grey-levels of the square change linearly from one background grey-level
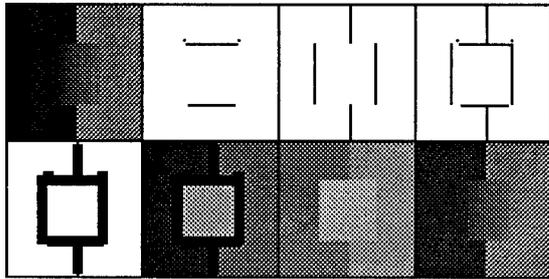
Figure 3: Dense-grid segmentation.
At the top from left to right: The original grey level images; Discontinuities at all columns; Discontinuities at all rows; Sum of discontinuities in both directions.
At the bottom from left to right: Edge linking; Edges superimposed on the image; Image segmentation; Reconstruction of the image from the segmentation and computed surfaces.
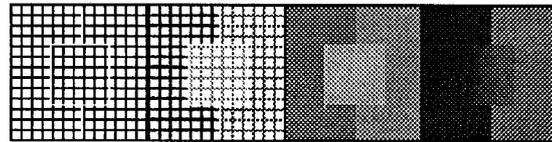

Figure 4: Sparse-grid segmentation.
Segmentation along every fifth grid point. From left to right: The sparse grid including breakpoints; The three connected components of the sparse grid; The segmented image; The reconstructed image.
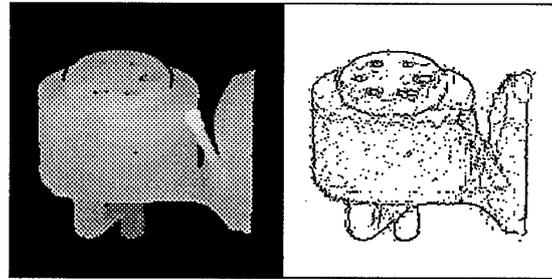

Figure 5: Range data segmentation: sharpener

to the other, producing two very low contrast edges on both sides of the square.

Figure 4 shows the results of applying a sparse grid. The one-dimensional segmentations were performed only on every fifth row and column in the image.

Figure 5 shows the edge-map of a range image. The segmentation was done using second degree polynomials. Since the object in this example is not piecewise second-degree polynomial there are many redundant edges in the edge map. Note the detection of the edge between the cylinder-wall and the top disk in the sharpener, even though it has a very low-contrast and is noisy.

6 Concluding Remarks

We have presented a tractable minimal length encoding algorithm for segmenting and compressing waveforms, and have demonstrated applications for two-dimensional data.

There are many fundamental questions concerning the use of one-dimensional cuts to achieve a meaningful two-dimensional segmentation:

a. How many cuts are needed?
b. At which directions should the cuts be taken?
c. How should the one-dimensional results be combined?

Answers to the first two questions are usually data-domain dependent. Solutions can be determined either prior to the segmentation or actively during the segmentation process, taking into account former one-dimensional results.

References

[1] G. Adiv. Determining three-dimensional motion and structure from optical flow generated by several moving objects. *IEEE Trans. on Pattren Analysis and Machine Intelligence*, 7:384–402, July 1985.

[2] B. Chazelle and D.P. Dobkin. Decomposing a polygon into its convex parts. In *11'th Annual ACM STOC*, pages 38–48, Atlanta, 1979.

[3] Y.G. Leclerc. Constructing simple stable descriptions for image partitioning. In *Proceedings: Image Understanding Workshop*, pages 365–382, Cambridge, MA, April 1988.

[4] M. Li and P.M.B. Vitanyi. Inductive reasoning and Kolmogorov complexity. In *Proc. 4th IEEE Structure in Complexity Theory Conference*, 1989.

[5] R. Marcus. Functional approximation based segmentation using the minimal description length principle. Master's thesis, The Hebrew University of Jerusalem, December 1989.

[6] T. Pavlidis. Algorithms for shape analysis of contours and waveforms. *IEEE Trans. on Pattren Analysis and Machine Intelligence*, 2:301–312, July 1980.

[7] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. *Numerical Recipes*. Cambridge University Press, 1986.

[8] J. Rissanen. Minimum-description-length principal. In *Encyclopedia of Statistical Sciences*, volume 5, pages 523–527. J. Wiley, New York, 1987.

[9] M. Werman, A.Y. Wu, and R.A. Melter. Recognition and characterization of digitized curves. *Pattern Recognition Letters*, 5:207–213, 1987.

[10] S.W. Zucker. Region growing: Childhood and adolescence. *Computer Graphics and Image Processing*, 5:382–399, 1976.