

# Closing Gaps of Discontinuous Lines: A New Criterion for Choosing the Best Prolongation

Eugene Bodansky and Alexander Gribov

Environmental System Research Institute (ESRI), Inc.  
380 New York St., Redlands, CA, 92373-8100, USA  
{ebodansky, agribov}@esri.com

**Abstract.** Polyline that are or should be continuous can have gaps in them, either because of scanning or digital processing or because they are depicted with discontinuous symbology like dots or dashes. This paper presents a new criterion for finding the most likely prolongation of discontinuous polylines .

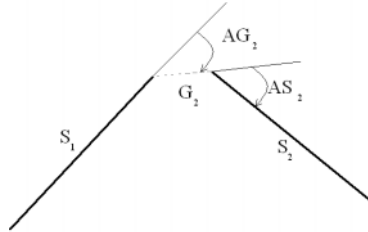
## 1 Introduction

Correctly recognizing discontinuous lines is an important part of raster-to-vector conversion or vectorization. Discontinuous lines can appear on raster representations of linear drawings because complex kinds of line were used (for example, dashed lines and dash-dotted lines) and because of scanning, binarization, and vectorization of source documents. For finding the most probable prolongation of the polyline, several parameters and thresholds have been used: the maximum value of gaps, the fan angle, and the range of line components for regular (for example, dashed) line symbology. Paper [1] presents a detection technique that uses a sequential stepwise recovery of components that meet certain continuity conditions. In [1], the maximum allowed length of a gap and the deviation angle between the gap of the new component and the previously detected component is used for choosing the next component of the discontinuous curve.

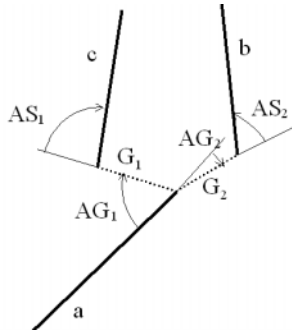
Let  $L$  be the maximum allowed length of a gap and  $T$  the maximum allowed deviation angle or fan angle. In compliance with [1] if  $G_i < L$ ,  $AG_i * AS_i > 0$ ,  $|AG_i| < T$ , and  $|AS_i| < T$ , then the candidate segment  $S_i$  is considered as the next component or a prolongation of the discontinuous curve (see Figure 1). These criteria have two disadvantages.

In the example shown in figure 2, the component “c” will be selected as the prolongation for the line “a” because for “c”  $AG_1 * AS_1 > 0$  and for “b”  $AG_2 * AS_2 < 0$ . We think that the component “b” is the more logical prolongation.

The criteria also do not allow selecting a prolongation from several candidates that all satisfy stated constraints.



**Fig. 1.** Two components of a discontinuous line  
 $S_i$  - a continuous segment,  $G_i$  - a gap,  $AS_i$  and  $AG_i$  - deviation angles



**Fig. 2.** Two candidate components ( $b$  and  $c$ ) for the line  $a$

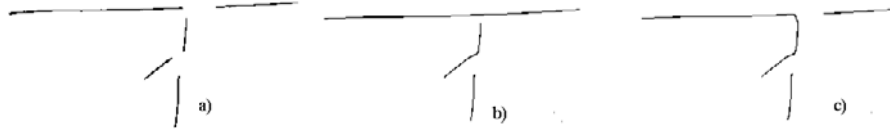
## 2 New Criterion for Choosing the Best Prolongation

We suggest the new method for selecting the linear component that is the best prolongation for a discontinuous curve. The method based on the mechanical analogy. Let a curve be the path of a particle moving at a constant speed. It is possible to calculate what job is needed to continue from the line to a candidate segment. A transfer from the line to the candidate segment will be defined as the segment connecting the closest ends of the centerline and the candidate segment. To compensate for the influence of noise, a part of a new virtual centerline containing the connecting segment will be smoothed with the local continuous smoothing algorithm described in article [2]. The derivatives of a transfer path will be continuous and so the radius of a curve at each point of the transfer path can be defined. The radius of smoothed curve can be calculated using the smoothing algorithm.

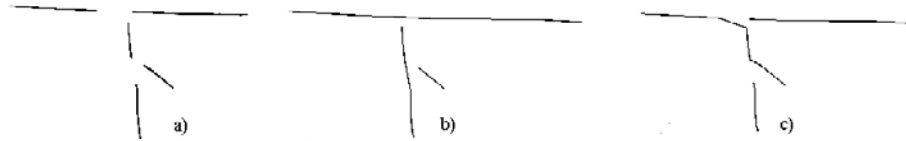
The job used for the transfer of the particle from the line to the candidate segment can be calculated as  $P1 + P2$ , where  $P1 = \int k_1 \frac{1}{R(l)} dl$  is a job for the direction changing,  $P2 = \int k_2 dl = k_2 L$  provides a constant speed.  $R(l)$  is a local radius of the curve,  $L$  is the length of the transfer path,  $k_1$  and  $k_2$  are constants. If  $P_i = k * P1_i + (1-k) * P2_i$  is a penalty where  $k$  is a constant, satisfying the restriction  $0 < k < 1$ , and  $i = 1, \dots, n$ ,  $n$  is the number of candidates for the prolongation of the line, then the candidate with the minimum value for the penalty will be chosen for the prolongation of the line. For large values of  $k$ , the criteria gives preference to more smooth result lines, for small values of  $k$  to lines with short transfer paths. Figure 3 shows the lines obtained with the gap closing technique with different values for  $k$ .

### 3 Tracing and Batch Mode

A sequential stepwise recovery technique can be used for semi automatic tracing with the operator control, but it could be not good for batch mode because closing gaps depends on the orientation of the document. In Figure 4, we show the result of closing gaps for source lines obtained by flipping the source lines shown in Figure 3.



**Fig. 3.** The source lines and the result of closing gaps. a) source lines, b)  $k = 1$ , c)  $k = 0$



**Fig. 4.** Flipped source lines and the result of closing gaps. a) flipped source lines, b)  $k = 1$ , c)  $k = 0$

## 4 Conclusions

Recognition of discontinuous polylines is difficult in documents that contain a lot of noise (speckles) and many different layers (for example, lines, letters, digits, and symbols). To increase the probability of making the correct solution, the recognition algorithm is needed that can be tuned with the operator. Sometimes it is necessary to find straight lines, sometimes smooth curves, sometimes polylines consisted of straight components, and so on. The suggested criterion of choosing prolongations of discontinuous polylines makes this possible.

## References

1. Dori, D., Wenyin, L., Peleg, M.: How to Win a Dashed Line Detection Contest. In: Kasturi, R., Tombre, K. (eds): *Methods and Applications. Lecture Notes in Computer Science*, Vol. 1072, First International Workshop, Springer-Verlag, Berlin Heidelberg New York (1997) 286–300
2. Bodansky, E., Gribov, A., Pilouk, M.: Post-processing of lines obtained by raster-to-vector conversion. In: *Vision (machine Vision Association of SME)*, Vol. 18, #1 ([www.sme.org/mva](http://www.sme.org/mva)), First Quarter, 2002.