RASTER-TO-VECTOR CONVERSION: A TREND LINE INTERSECTION APPROACH TO JUNCTION ENHANCEMENT

Peng Gao
Michael M. Minami

Environmental Systems Research Institute, Inc.
380 New York Street, Redlands, CA 92373

ABSTRACT

One of the problems with automatic or semi-automatic raster-to-vector conversion is dimpled junctions, which are caused by the underlying thinning algorithms. While useful for shape reconstruction from skeleton, dimpled junctions are undesirable for data conversion. Techniques, referred to as junction enhancement, have been developed to deal with, among others, the dimpled junction problem. However, existing junction enhancement techniques lack well-defined procedures and are only able to handle junctions on a case by case basis. This paper presents a trend line intersection approach to systematically enhance all types of junctions. This approach has a set of simple and well-defined procedures for estimating the trend lines of all arcs coming to a junction, and for calculating the new junction location as the intersection of all the estimated trend lines. It can handle junctions with any number of incoming arcs and any type of configuration using the single set of procedures. The experiments have shown that the new approach is very effective in dealing with dimpled junctions.

INTRODUCTION

Scan digitizing is becoming a practical alternative to table digitizing due to the wide availability of scanners (Becker 1992). Unlike table digitizing, however, scan digitizing requires the additional step of raster-to-vector conversion after a map is scanned in the raster form. There are still many unsolved problems with automatic or semi-automatic raster-to-vector conversion. One of them is dimpled intersections, or junctions, caused by the underlying thinning algorithms, as shown in Figure 1 (a). While useful for shape

Figure 1. (a) Dimpled junction; (b) Enhanced junction

reconstruction from skeleton, dimpled junctions are undesirable for data conversion, where a clean intersection is more satisfactory (Figure 1 (b)). Techniques, referred to as junction enhancement, have been developed to deal with, among others, the dimpled junction problem. Most current techniques start by grouping junctions into junction classes, such as, +-junction, T-junctions, X-junctions, Y-junction, etc. A template is pre-defined for each junction class, and used to match an input junction. If a match is found with one of the pre-defined classes, the input junction is enhanced with the rules associated with that junction class. However, the set of rules for each junction class is pre-defined, and if a junction deviates from its standard form, it will be either incorrectly adjusted or left dimpled. The number of junction classes is also pre-defined, and, if an input junction does not match one of the pre-defined classes, it will be left dimpled, too. It is clear that this approach lacks a systematic procedure for dealing with a specific junction, and lacks flexibility in handling different types of junctions. This paper presents a trend line intersection approach for dealing with the dimpled junctions. Given an input junction containing a node and arcs coming to it, this approach starts by estimating the trend line for each arc within a range close to the node and then, moves the junction to the intersection of trend lines of all the arcs and reconnects the arcs to the new node. This approach has a set of simple, well-defined procedures for estimating the trend line of an incoming arc, and for calculating the intersection of all the estimated trend lines. It can handle junctions with any number of incoming arcs and any type of configuration using the single set of procedures.

JUNCTION ENHANCEMENT

A junction is a structure consisting of a node, a set of arcs, and a relationship between the node and arcs. The location of a junction is the location of the node. Junction enhancement is a process of replacing the old node with a new one and reconnecting arcs to the new node. The trend line intersection approach is based on a simple observation and assumption that the true junction location is the intersection of the trend lines of all arcs coming into a junction. In this section, three important parameters, *range*, *angle tolerance*, and *distance tolerance*, and two concepts, trend line and extended trend line, are defined first, and followed by the description of the procedure for performing the junction enhancement by the trend line intersection.

Range and Trend Line

The *range* defines the segment of an arc within which the trend line is estimated. It is the straight line distance measured from the node to a point on the arc, see Figure 2. The

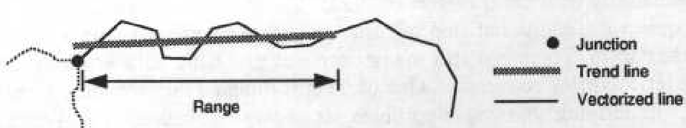

Figure 2. Range and Trend Line

trend line is a line in the form: $y = ax + b$, fitted to the arc within the range. The trend line

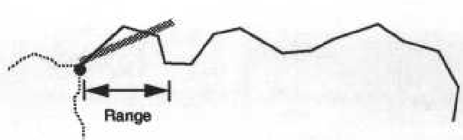

Figure 3. Smaller Range Value and Trend Line

represents the general direction of an arc within the range. The range value controls the sensitivity of the trend line to the local direction around the node. The smaller the range, the more sensitive the trend line will be to the local direction (Figure 3).

Angle Tolerance and Extended Trend Line

The *angle tolerance* specifies the maximum angle, α, between two trend lines such that they can be considered parallel and thus merged into an extended trend line (Figure 4). An *extended trend line* is a single line fitted to the two parallel arc segments from both sides. Figure 4 (b) shows the extended trend line which replaces the two parallel trend lines in Figure 4 (a).

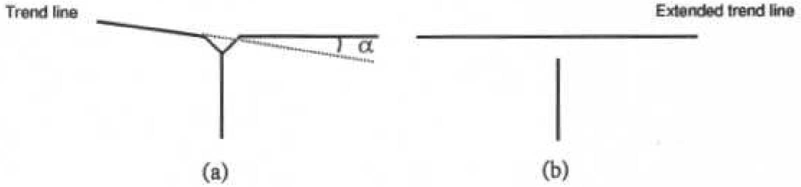

Figure 4. Angle Tolerance and Extended Trend Line

Distance Tolerance and Node

The *distance tolerance* specifies the maximum distance a node can move during junction enhancing (Figure 5). A junction is enhanced only when the distance between the

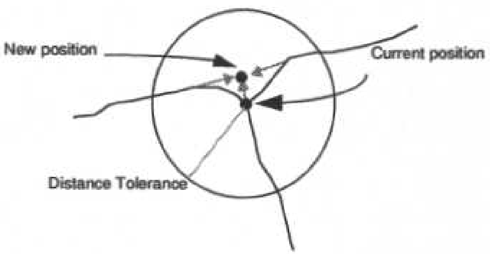

Figure 5. Distance Tolerance and Node

new node position and its current position is within the distance tolerance.

Junction Enhancement Procedure

In general, this procedure calculates a new node location and reconnects all related arcs to the new node to form a clean junction. The new node location is determined by first evaluating the trend of the vectorized lines, or arcs, coming into the junction. A trend line represents the general, or average, direction of the vectorized line based on a subset of the vertices within the line. The number of vertices used to determine the trend line is controlled by the *range* (Figure 2 and 3). Then, all trend lines are checked to determine if any two are parallel and thus should be merged. Two lines are assumed to be parallel if the angle between the trend lines falls within the specified *angle tolerance* (Figure 4). Finally, the remaining trend lines are intersected to create the new node, see Figure 6, and the

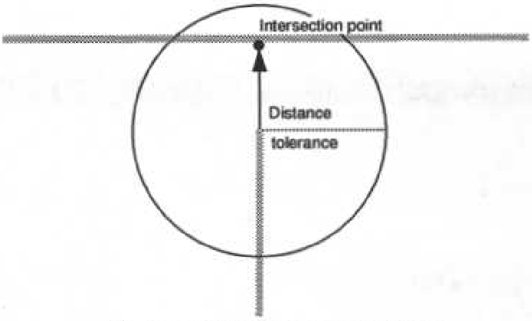

Figure 6. Trend Line Intersection

junction is enhanced if the distance between the current and new nodes is within the *distance tolerance*, as shown in Figure 5.

Assume that there exists a junction structure which contains a node, arcs and the topological relationship between the node and arcs connecting to it. The following is a more formal procedure for the trend line intersection approach.

1) *for each arc coming into a junction, sample the segment of the arc within the range, and fit a line:* $y = ax + b$*; to the sample points along the arc;*
2) *find all parallel trend lines whose angles are within the angle tolerance, refit an extended trend line for each parallel pair, and replace the two parallel ones with the extended trend line;*
3) *intersect all remaining trend lines (including the extended ones) to determine the node location:*
   a) *if number of extended trend lines is greater than 1, calculate the intersection points between all pairs of extended trend lines;*
   b) *if number of extended trend lines is 1, calculate the intersection points between the extended trend line and the rest of trend lines;*
   c) *if the number of extended trend lines is 0, then calculate the intersection points between the best fitted trend line and the rest of trend lines, where the best fitted trend line is selected based on a goodness of fit value;*
4) *calculate the new node location as the average of those intersection points that are within the distance tolerance to the old node location;*
5) *reconstruct the junction using the new node;*
6) *select another junction and repeat steps 1-5 or stop when there are no more junctions to process.*

It is easy to see that this procedure can be applied to junctions with any number of incoming arcs and with any kind of configuration. Given sample points along an arc, the trend line can be fitted using the Least Squares technique, and the χ^2 value can be used as the goodness of fit value. In general, the extended trend lines yield a more reliable new node location. Therefore, when there is more than one extended trend line, the new node location is simply determined by the intersection of the extended lines. This is the case of +-junctions. When there is only one extended trend line, the extended trend line is used as the base and the rest of the trend lines intersect with it. This is the case of T-junctions and K-junctions. When there is no extended trend line, the best fitted trend line is used as the base and the rest of the trend lines intersect with it. The new node location created in the absence of an extended trend line is less reliable. This is the case of Y-junctions.

RESULTS

The trend line intersection approach has been implemented and tested in a production system. This section illustrates the results of this approach when applied to artificially constructed junctions, and real junctions from a scanned soil/vegetation composite map. Figure 7 shows four different types of junctions, including +-junction, X-junctions, T-junctions, and Y-junctions, before enhancement., White lines are converted

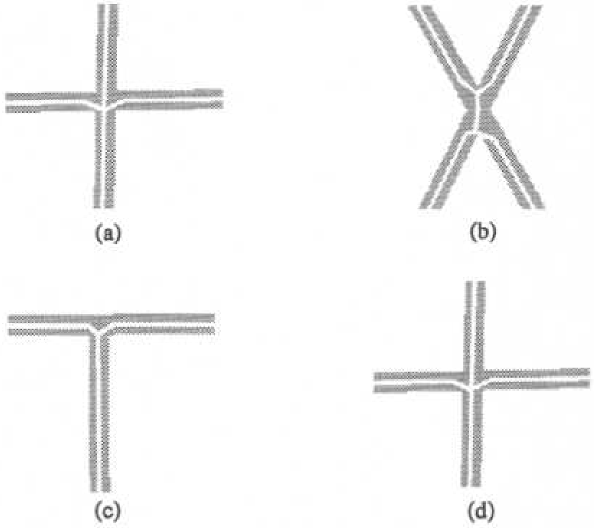

Figure 7. Artificial Junctions before Enhancement

vector lines displayed on top of raster lines in grey. Figure 8 shows the same junctions after enhancement. Figures 9 and 10 contain a small section of a scanned soil/vegetation composite map. Figure 9 shows vectorized junctions and arcs before enhancement, and Figure 10 shows the results after enhancement. Note that there are many new types of junctions in this map section, including K-junctions and 6-valent junctions (a junction with 6 incoming arcs), in addition to different forms of X-, T-, and Y-junctions.

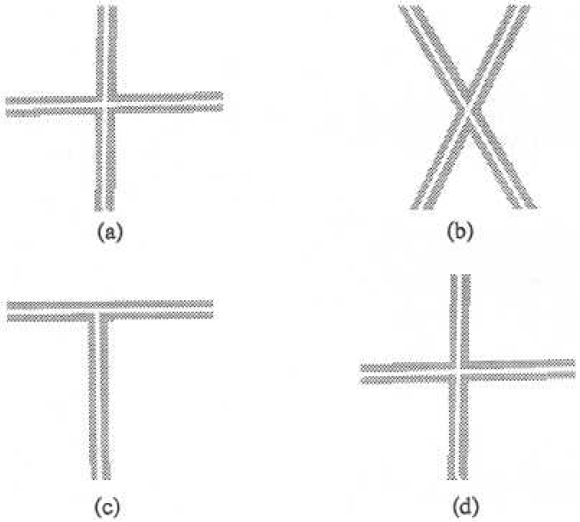

Figure 8. Artificial Junctions after Enhancement

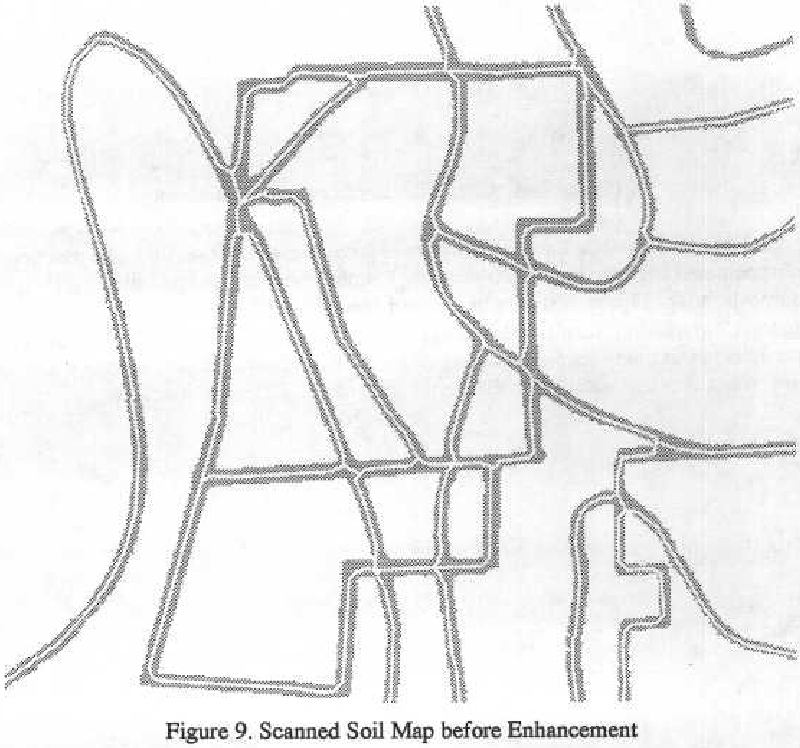

Figure 9. Scanned Soil Map before Enhancement

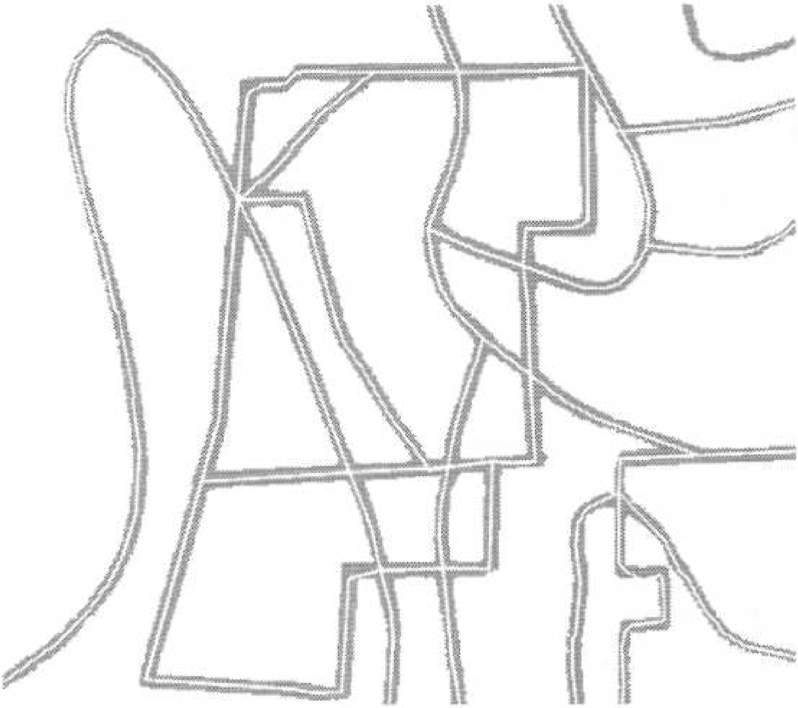

Figure 10. Scanned Soil Map after Enhancement

CONCLUSION

The trend line intersection approach provides a simple, yet effective, procedure for enhancing dimpled junctions, and is applicable to junctions with any number of incoming arcs and with any kind of local configuration. It eliminates the needs for ad hoc rules for handling different types of junctions.

REFERENCES

Becker, R.T, and Devine, H.A. 1992, Raster-To-Vector Conversion of Scanned Document fro Resource Management, *ASPRS/ACSM/RT 92 Technical Papers,* Washington, D.C.