

Identification of Curves and Straight Sections on Road Networks from Digital Vector Data

R. Andrášik*, M. Bíl, Z. Janoška, V. Valentová

Transport Research Centre, Brno, Czech Republic

**Corresponding author: richard.andrasik@cdv.cz*

DOI: 10.2478/v10158-012-0033-0

ABSTRACT: Identification of curves and straight sections on roads is important from a traffic safety point of view. We present a method for the automatic identification of curves based on available geographic data – sequence of points which determines a line. The Douglas-Peucker algorithm is used to generalize the lines. At every point of the line, the coefficient of curvature is calculated. According to the coefficient of curvature it is possible to decide which line segments are curves and which are straight sections. For making the decision, other characteristics of lines were also studied, but the obtained results support that the coefficient of curvature is the most suitable of them. The method was applied on two sections from the district of Olomouc.

KEY WORDS: Horizontal curves, classification, osculating circle, coefficient of curvature.

1 INTRODUCTION

The main reasons why the identification of curves and straight sections is important are the increased risk of traffic accidents in curves on the one hand and the calming effect of curves on the other hand. Curves are places where the risk of a traffic accident increases as a result of getting out of the lane, or due to poor driver's visibility. According to the synthesis of the World Road Association (PIARC, 2003) the accident rate in horizontal curves is 1.5 to 4 times higher than on straight sections. In addition to safety risks, curves, to some extent, represent also a calming element, because drivers slow down when entering a curve. They are forced to pay more attention in areas with frequent curves, while in areas with prevailing monotonous straight sections there is a risk of a loss of concentration of the driver. For this reason, it appears that the most dangerous are those curves which follow after a long straight section of the road. Identification of curves and their characteristics – particularly the size of horizontal curves and length of curves, is therefore desirable from both safety experts and for example from road administrators – curves are often equipped with crash barriers or other technical elements and a tool for their automatic identification can facilitate road administrations to perform their management tasks.

Identification of curves and straight sections is an issue which has already been studied in the past, including several recent publications on the use of a GIS to identify curves (Bogenreif et al., 2012; Findley et al., 2012; Rasdorf et al., 2012).

2 AREA OF INTEREST AND DATA SOURCES

The aim of this work was to propose a method for the automatic identification of curves on roads based on available geographic data. The proposed method is based on the principle of determination of the radius of an osculating circle (Fuchs & Tabachnikov, 2007; Stewart, 2011) which adjoins the road. By classifying these radii we can then determine whether a given segment of the road is part of a curve or a straight section.

This method was compared with four other methods based either on identification of the radius of the circle circumscribed to a road segment, or on the calculation of the angle formed by break points of a road line.

The most accessible, while also official, source of data on the road network in the Czech Republic is the recording of road equipment and facilities from the database of the Road and Motorway Directorate of the Czech Republic (RMD CR). Data can also be sourced from OpenStreetMap, however, their relevance is difficult to verify. Geographic data can also be sourced from local authorities, whose data sets are likely to originate from the database of RMD CR. Last but not least, there are commercial products, such as the reason, they have the possibility to acquire the data by using GPS devices or by digitizing maps. In this work we used RMD CR data as they are acquired directly by the administrator of the roads in question, and it is a source accessible to state administration and local governments, the authorities which are the most active both in the field of traffic safety and road administration.

Two test sites were selected (see Figure 1). The first location is a segment of road II/446 between the villages Chomoutov and Pňovice with a total length of 8.7 km. This section is situated in a flat terrain with several shallow curves. The second section is the section of road II/444 located between the village of Lipina and the junction near the mill Těšíkovský mlýn with a length of about 5.6 km. This section contains frequent horizontal and vertical curves, and runs through an urban area as well as a densely forested valley. Both locations are situated near the town Šternberk in the district of Olomouc.

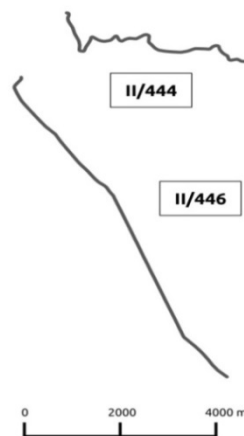


Figure 1: Road sections II/444 (up) and II/446 (down).

3 DATA PRE-PROCESSING

Before using the method itself for identifying the curves, it is necessary to properly prepare the data. It often happens that in some places, the points defining a road segment are too close to each other, contain redundant information, or contain defective information incurred during data collection that could result in the distortion of the identification of curves. This can be avoided by using the geometric line generalization.

3.1 Geometric line generalization

Generalization is used to reduce and simplify the volume of data while maintaining a sufficient amount of information contained in such data. Lines can be generalized with the use of several algorithms:

- Leaving out/maintaining every k-th point of a line;
- Length test (removal of nearby points);
- Angle test (eliminating points with small angle difference);
- Lang's algorithm;
- Visvalingam-Whyatt algorithm;
- Reumann-Witkam algorithm;
- Douglas-Peucker algorithm.

A comprehensive list and description of methods for line generalization is presented in Clayton (1985). In this case, we selected the last mentioned method of line generalization that is described below.

3.2 Douglas-Peucker algorithm

We select the tolerance of accuracy (e.g. 0.1 m). The Douglas-Peucker algorithm works as follows:

- First, connect the start and end point of the section;
- Find a point whose distance from the line segments formed by all the connected points is the largest (we consider the classical Euclidean distance, i.e. we draw a perpendicular line from the point);
- Compare this distance with the selected tolerance and if the distance is lower than the accuracy tolerance, terminate the calculation;
- Include a point whose distance is greatest among the connected points. Return to step 2.

Through this process we obtain a simplified line. This method is illustrated in the example in Section 5.

4 METHOD FOR IDENTIFICATION OF CURVES

Typically, a curve is constructed as a segment of a circle of a given radius and is connected to the straight sections of the road by transition curves. Due to this, it is suitable for the identification of curves to try and locally approximate the line given by points $X_i = (x_{i1}, x_{i2})$, for $i = 1, \dots, n$, by a circle.

The first idea that comes to mind is to draw a circle to every three neighbouring points and then assign its radius to the middle point of these three points as road curvature characteristics at this point.

If we had virtually ideal data available, i.e. break points are approximately equally apart from each other, there are neither “dense” nor “sparse” places in the data and there are only very small measurement errors (up to 0.1 m) that can be removed using the Douglas-Peucker algorithm; we could use the radius of a circumscribed circle to identify curves without much difficulty.

If no precise data are available, the calculation of the radius of a circumscribed circle may lead to the erroneous determination of curves because of the assignment of relatively small radii. For example, in a situation with a straight section, on which all points defining this segment lie

on a line except one point, which is deflected from the straight line by only 0.5 m, the radius of a circumscribed circle will be relatively small and therefore there will be identified a false curve.

As we can see, the determination of curves using a circumscribed circle with the approximation of the line between the points it circumscribes has some weaknesses, and therefore a better method is to use the so-called osculating circle, which, if combined with the Douglas-Peucker algorithm, has none of the above-mentioned drawbacks. Its definition implies that it does not approximate three neighbouring points, but tries to approximate the curvature of a line at one specific point. Its radius is dependent on the change in slope of the tangent to the curve and therefore it tells us how much the road is curved at a given point.

Both above-mentioned methods have been tried in order to identify curves and were supplemented with two alternative procedures. The first determines the curves based on the angle formed by three consecutive points, the second method uses the cumulative angle of three consecutive break points.

4.1 Calculation of radius of osculating circle

Let us denote the curve $c : I \subset \mathbb{R} \rightarrow \mathbb{R}^2$. Based on the geometry of curves (Fuks & Tabachnikov, 2007; Stewart, 2012) we know the formula for calculating the coefficient of curvature at a given point:

$$\kappa(t_0) = \frac{\|\dot{T}(t_0)\|}{\|\dot{c}(t_0)\|}, \quad t_0 \in I,$$

where $T(t)$ denotes the tangent vector of the curve and $\dot{T}(t_0)$ its derivative with respect to variable t at point t_0 . Similarly $\dot{c}(t_0)$ denotes a derivative of curve c with respect to t at point t_0 . The number $R(t_0) = 1/\kappa(t_0)$ denotes the radius of an osculating circle at point t_0 . The larger is the radius, the smaller is the curvature of the curve.

Since we have no analytical specification of the curve (our road section), we have to approximate the depiction of c and T at each point by finite differences. For $I = 2, \dots, n-1$ the calculation is as follows:

$$\begin{aligned} T(X_i) &= (X_{i+1} - X_{i-1}) / \|X_{i+1} - X_{i-1}\| \\ \dot{T}(X_i) &= (T(X_{i+1}) - T(X_{i-1})) / (\Delta t_1 + \Delta t_2) \\ \dot{c}(X_i) &= (X_{i+1} - X_{i-1}) / (\Delta t_1 + \Delta t_2) \end{aligned}$$

Here we deduce the formula for calculating the coefficient of curvature easily:

$$\kappa(X_i) = \frac{\|\dot{T}(X_i)\|}{\|\dot{c}(X_i)\|} = \frac{\|T(X_{i+1}) - T(X_{i-1})\|}{\|X_{i+1} - X_{i-1}\|}$$

The use of this method is demonstrated in the examples in Section 5.

4.2 Distinguishing curves and straight sections

When describing a road network (e.g. for statistical purposes), it is strictly necessary to distinguish which part of a road is a curve and which is not. This distinction can be made on the basis of:

- Expert recommendation or legislature (it is stated what the radius of a road section must be so that we consider it to be a curve);
- A histogram of radii of osculating circles (e.g. the case of bimodal distribution would be indicative of two groups of data in the database);

- Manual identification of curves and straight sections by an expert on selected data and finding a classification rule which, after entering the characteristics (osculating circle radius, distance to the nearest break point, angle between three neighbouring points, etc.) would say whether a certain section is a curve or a straight section (this would be the so-called classification with a teacher).

5 RESULTS AND DISCUSSION

This section describes the results achieved by applying the methodology described above to data from Section 2.

First, we used the Douglas-Peucker algorithm for line generalization. As can be seen in Table 1, regarding section II/446, which is not very curved (see Figure 1), a relatively large amount of points were eliminated in contrast to section II/444. This is in accordance with expectations.

Table 1: Number of points defining a road section before and after using the Douglas-Peucker algorithm.

	Before	After
II/446	125	113
II/444	212	210

It is interesting that the data from RMD CR before and after generalization differ very slightly. It is possible that these data have already been processed using some generalization algorithm.

In the second step, we calculated for each point:

- The radius of an osculating circle that touches the curve (our road section) at this point;
- Radius of a circle circumscribing three consecutive points;
- Angle between three consecutive points;
- Cumulative angle of three points.

For both road sections, a traffic engineering expert defined curves and straight sections. For each of the analyzed break points it was then determined whether after this classification it is a curve or straight section. For each of the four monitored variables, histograms were prepared comparing the distribution of values in both categories.

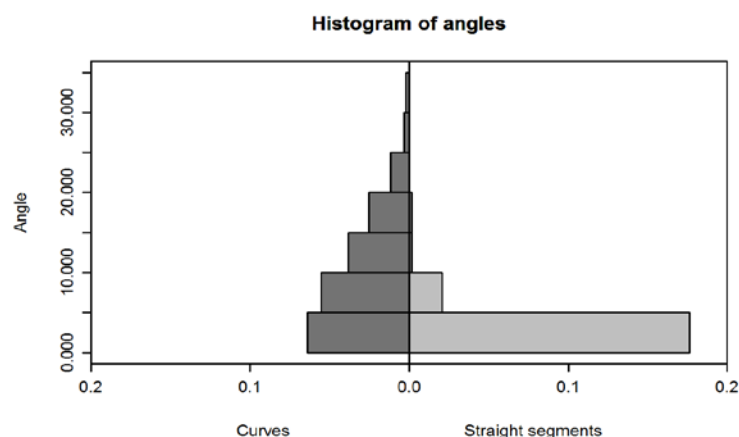


Figure 2: Histogram of angles.

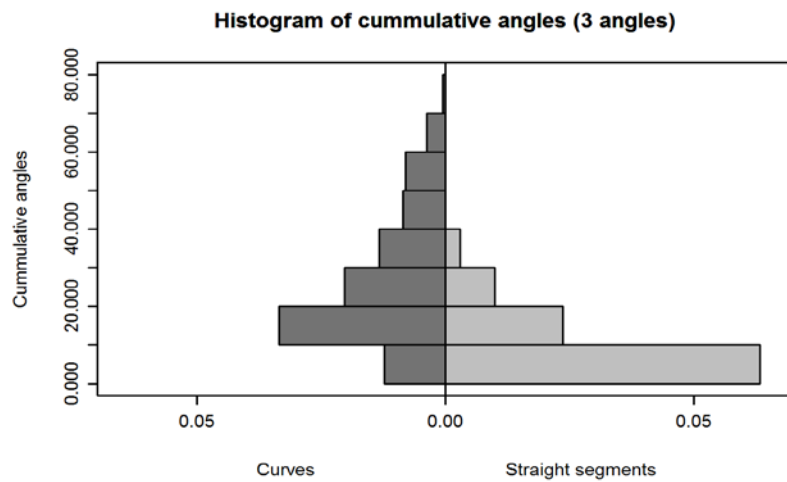


Figure 3: Histogram of cumulative angles (sum of three angles).

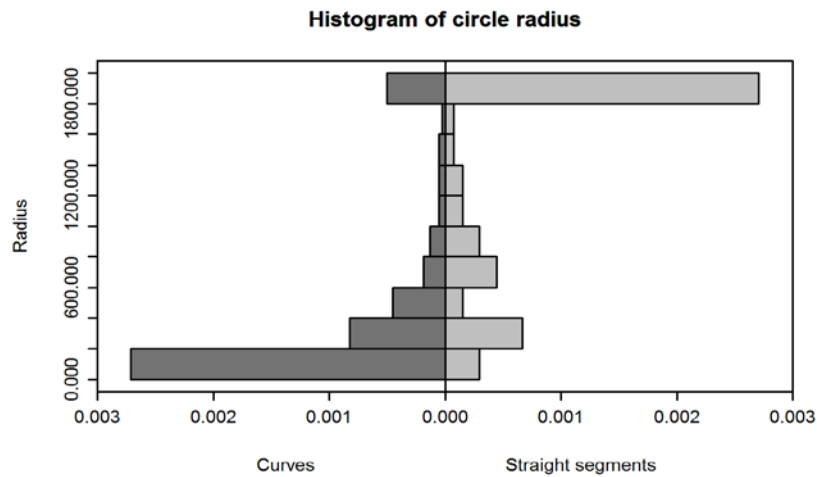


Figure 4: Histogram of radii of circumscribed circles.

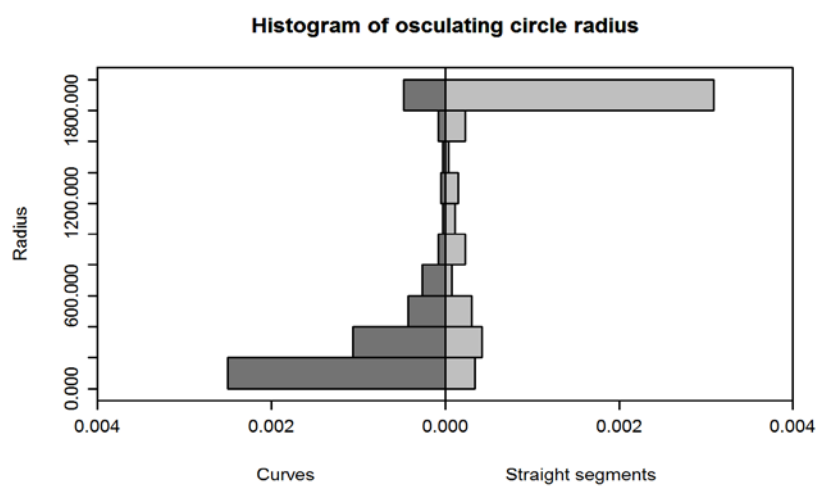


Figure 5: Histogram of radii of osculating circles.

Figures (2 – 5) clearly show that it is not possible to find a value that would clearly define straight sections and curves when using the characteristics derived from the size of angles. Histograms of the radii of circles are strongly asymmetrical and more suitable for this purpose. Table 2 shows a comparison of the osculating circle method and circumscribed circle method for different classification levels.

Table 2: Classification success rate for the given boundary values of radii. The values indicate the proportion of correctly classified break points.

Threshold value of radius (m)	II/446		II/444	
	Circumscribed circle	Osculating circle	Circumscribed circle	Osculating circle
100	65%	65%	54%	56%
300	73%	72%	74%	76%
500	74%	73%	80%	80%
1000	70%	78%	80%	81%
1500	70%	74%	78%	80%
2000	72%	75%	77%	78%

The success rate for the section II/444 is specified in more detail in Table 3. The specificity value indicates the probability that if the point is determined as a part of a curve, this determination is correct. Sensitiveness is an analogy to specificity for straight sections. The value of PPV (positive predictive value) indicates what percentage of curves was correctly identified, and NPV (negative predictive value) indicates what percentage of straight sections was correctly identified. As shown in table 3, the difference between these two methods lies in the specificity, where the osculating circle appears to be a better solution. Identification of straight sections is therefore equally successful in both methods, for the identification of curves the osculating circle method is more successful.

Table 3: Specificity and sensitivity of the circumscribed circle method and osculating circle method when used on section II/444. Threshold value of radius is 1000 m.

	Circumscribed circle	Osculating circle
Sensitivity	80.1%	80.3%
Specificity	79.5%	83.8%
PPV	94.5%	95.9%
NPV	47.7%	47.7%

6 CONCLUSION

A method was proposed for the automatic classification of curves and straight sections based on the geometry of roads. This method is based on calculating the radius of an osculating circle and was compared with three alternative methods. The methods based on the calculation of curve radius appeared significantly better than those based on the calculation of break angles of lines. Both methods based on the calculation of the radii of curves are comparable in identifying straight

sections, while for the identification of curves, the osculating circle method produces better results. Determination of margins for identifying curves is a matter for further research and was not further addressed in this work. Overall, the proposed method of automatic identification is successful and easy to implement into a GIS environment, and thus suitable for broader use.

ACKNOWLEDGEMENT

This paper was prepared with the help of project No. TA01031581 funded by the Technology Agency of the Czech Republic. Further acknowledgements are extended to the employees of the Police Presidium of the Czech Republic and the Road and Motorway Directorate of the Czech Republic for providing the data.

REFERENCES

- Bogenreif, C., Souleyrette, R. R., Hans, Z., 2012. Identifying and Measuring Horizontal Curves and Related Effects on Highway Safety. *Journal of Transportation Safety & Security*, 4 (3).
- Clayton, V. H., 1985. *Cartographic Generalization: A Review of Feature Simplification and Systematic Point Elimination Algorithms*. NOAA Technical Report NOS, 112. Rockville, Md.: National Oceanic and Atmospheric Administration, National Ocean Service, Office of Charting and Geodetic Service.
- Findley, D. J., Zegeer, CH. V., Sundstrom, C. A., Hummer, J. E., Rasdorf, W., Fowler, T. J., 2012. Finding and Measuring Horizontal Curves in a Large Highway Network: A GIS Approach. *Public Works Management & Policy*, 17 (2).
- Fuchs, D. B., Tabachnikov, S., 2007. *Mathematical Omnibus: Thirty Lectures on Classic Mathematics*. Providence (USA): American Mathematical Society. ISBN 0821843168.
- PIARC, 2003. *Road Safety Manual*. Paris: PIARC. ISBN 2-84060-158-3.
- Rasdorf, W., Findley, D. L., Zegeer, C.V., Sundstrom, C.A., Hummer, J. E., 2012. Evaluation of GIS Applications for Horizontal Curve Data Collection. *Journal of Computing in Civil Engineering*, 25 (2).
- Stewart, J., 2011. *Calculus*. Brooks Cole Publishing, USA.

**The article was produced under the support of the project
Transport R&D Centre (CZ.1.05/2.1.00/03.0064)**