Estimating Roadway Horizontal Alignment Information Using Machine Learning

Bekir Bartin[1] and Kaan Ozbay[2]

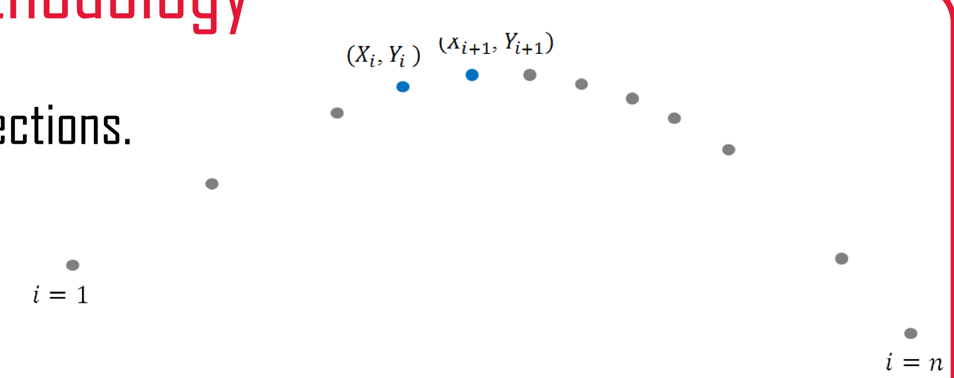
Motivation

- Roadway horizontal alignment data are crucial and necessary information for various safety analyses.
- However, this information is not readily available in states' roadway inventory databases.
- The necessity of horizontal alignment dataset and the difficulties in obtaining it have become more apparent since the publication of HSM.
- As a result many researchers resort to manually extract this information using satellite images. However, extracting this information, especially for a multitude of roadways demands a significant amount of time, and the quality of the extracted data is often questionable.
- Among the available methods, the Geographic Information Systems(GIS)-based ones appear to be an ideal method in terms of cost and time, since it utilizes GIS roadway centerline shapefiles, already available in state DOT's data repository.

Objectives

- To introduce an *artificial neural network* (ANN) based road horizontal alignment estimation method, which uses GIS roadway centerline data, and to evaluate its accuracy.
- The proposed ANN method is first trained using available GIS road alignment data and then applied to a new roadway, of which the actual alignment information is known *a priori*.

Problem Definition & Methodology



- A roadway is a smooth function consisting of several **tangent** and **curved** sections.
- GIS nodes (vertices) are discrete points located along this function.
- If how many distinct tangent and curved sections there are and the start and end of these sections were known one could use curve fitting to estimate the horizontal alignment information. However, the lack of such information makes the horizontal alignment estimation problem complicated.
- The proposed approach includes a two-step process:

  Step 1 - Segmentation: Estimation of the number and boundary of tangent and curved sections.
  Step 2 - Circular Curve Fitting: The piecewise least squares estimation of circular curve sections using Taubin fit method.

- ANN is used in Step 1. The use of ANN for the estimation of road horizontal alignment stems from the idea that distinct road sections, either curved or tangent, have similar explanatory variables.
- Seven explanatory variables are identified that are likely to indicate whether a point on road is part of a curved or a tangent section. These are: Heading angle (θ), curvature (κ), change in heading angle (dθ), distance between consecutive vertices (dL), milepost (MP), radius, radius of a circumscribed circle (RCC), cumulative change in heading angles at three consecutive points(Σ), the difference in cumulative heading angles around each vertex (λ).
- Therefore, each GIS vertex has a vector of explanatory variables, denoted by $\boldsymbol{\alpha}$, where,

  $$\boldsymbol{\alpha} = (\kappa_i, \theta_i, d\theta_i,\ dL_i, RCC_i, \Sigma_i, \lambda_i)$$

- ANN method uses a training dataset of GIS vertices with $\boldsymbol{\alpha}$ information and whether or not each vertex belongs to a tangent or a curved section.

Dataset

Training, Validation & Test Datasets

- The ANN model is trained using two different datasets.
- The first dataset is the actual horizontal alignment data from a 28-mile section of a roadway, Route AH76 located in Herat Afghanistan, which was built in 2020. The roadway's horizontal dataset was available in CAD format. It was converted to ArcMap shapefile and a total of 3,304 vertices were extracted automatically.
- The second dataset includes 2,199 vertices that were extracted manually from 10 rural roads in New Jersey. The decision whether a vertex belongs to a curved section or not was conducted subjectively.

Evaluation Dataset

- The evaluation of the ANN methods is performed by using a freeway segment on I-80, between mileposts 15 and 25.72, in Eureka County, Nevada (NV) in the eastbound direction.
- The alignment information for I-80 is obtained the literature in which the locations of curved sections and respective radii values are tabulated. The GIS roadway centerline shape file for I-80 is obtained from the NDOT website.

Results

ANN Training and Test Results

- The NN model has 2 hidden layers with 16 and 8 nodes, respectively.
- The last node is the sigmoid activation function which mapped all values between 0 and 1 into the form of a sigmoid curve.
- Once the neural network is created, the model is trained for 200 epochs with a batch size of 32. After training was completed, the model accuracy prediction on the test was 90.01 %.

Evaluation Results

- Following table is the comparison of the actual horizontal alignment and the ones estimated by the ANN method for I-80 in NV.
- I-80 segment is 10.72 miles long with 10 curves.
- The total curved and tangent section lengths are 4.02 and 6.7 miles, respectively.
- All curved and tangent sections are detected by the ANN method.
- The overlap ratio of the estimated curve lengths with the actual ones is 97%. The total undetected curve length is 0.11 miles (57.4 feet on average).
- Similar overlap ratio for the tangent sections is 92%. the total undetected tangent length is 0.545 miles (261.5 feet on average).
- Such discrepancy is expected as the vertex resolution for I-80 is 183 feet (0.03 miles) between two vertices on average.

Actual Alignment Data				ANN Method*			
k	SMP	EMP	$R_k(ft.)$	$\hat{k}$	$\widehat{SMP}$	$\widehat{EMP}$	$\hat{R}_{\hat{k}}(ft.)$
1	15.00	15.48	-	1	15.00	15.47	-
2	15.48	15.91	5,500	2	15.47	15.96	5,499
3	15.91	15.99	-	3	15.96	16.00	-
4	15.99	16.38	5,000	4	16.00	16.42	5,021
5	16.38	16.66	-	5	16.42	16.62	-
6	16.66	16.95	3,000	6	16.62	17.02	3,172
7	16.95	17.05	-	7	17.02	17.06	-
8†	17.05	17.24	3,078	8	17.06	17.27	3,126
9†	17.24	18.75	-	9	17.27	18.71	-
10	18.75	19.19	6,000	10	18.71	19.19	6,135
11	19.19	19.31	-	11	19.19	19.29	-
12	19.31	19.57	3,584	12	19.29	19.60	3,744
13	19.57	20.12	-	13	19.60	20.15	-
14	20.12	20.47	6,000	14	20.15	20.48	6,032
15	20.47	22.46	-	15	20.48	21.01	-
16	22.46	22.72	5,000	16†	21.01	21.16	11,406
17	22.72	23.35	-	17	21.16	22.47	-
18	23.35	24.17	5,108	18	22.47	22.77	5,492
19	24.17	24.54	-	19	22.77	23.27	-
20	24.54	25.13	5,909	20	23.27	24.12	5,115
21	25.13	25.70	-	21	24.12	24.48	-
				22	24.48	25.13	6,027
				23	25.13	25.72	-

Conclusions

- This study presents a novel approach for estimating roadway horizontal alignment data that utilizes ANN model trained using available roadway horizontal alignment data.
- As extracting this data is time consuming and costly, an efficient and fast method is required as this information is crucial in various traffic safety related analyses.
- The proposed approach takes advantage of the available GIS centerline shapefiles available for nearly all roadways in the U.S.

[1] Ozyegin University, Istanbul TURKEY [2] Professor and Director, NYU C[2]SMART Center
This study was funded by New Jersey Department of Transportation