GROUND-BASED AERIAL PHOTOGRAPHY TO SUPPORT TRAFFIC FLOW EXTRACTION

Nora Csanyi[1, 2]
Charles K. Toth[2]
Department of Civil and Environmental Engineering and Geodetic Science[1]
Center for Mapping[2]
The Ohio State University
1216 Kinnear Road, Columbus OH 43212-1154
nora@cfm.ohio-state.edu

Pitu Mirchandani
Department of Systems & Industrial Engineering
University of Arizona

ABSTRACT

Road traffic has been steadily increasing worldwide for several decades. With the rising number of vehicles using the transportation network, the effectiveness of traffic management is becoming more crucial, since new road construction does not keep up with the growth of the traffic volume. As a key for effective traffic management, monitoring traffic flow is of growing interest. Recently, new imaging-based systems have been introduced for traffic flow monitoring. These pole- or bridge-mounted cameras show encouraging performance, and the trend of switching toward imaging technologies is expected to increase. Vehicle (feature) extraction from the imagery is a computer vision task. Due to the oblique viewing angles of the pole-mounted cameras, the imagery is significantly distorted, and therefore to support the feature extraction process, it is important to rectify the imagery. This paper proposes a practical approach for rectifying the oblique imagery to the plane of the road surface. The method does not require an available digital elevation model for the rectification, it only requires measured distances from the camera to at least three points on the road surface well distributed in the image area. The paper provides a detailed description of the proposed method and an initial performance analysis.

INTRODUCTION

Road traffic, measured in various parameters of passenger and cargo volume, has been steadily increasing worldwide for several decades. For example, the total vehicle-miles traveled in a year, the most frequently quoted traffic parameter that includes traveled distances for all vehicle categories, has almost doubled in the US in the last 15 years (US Bureau of Transportation Statistics, http://www.bts.gov). However, the paved road network grew only a modest 10% for the same period; thus the increasing traffic on the practically stagnating road infrastructure results in more severe traffic congestion.

With the increasing number of vehicles entering the existing transportation network annually, the effectiveness of traffic management is becoming more crucial. Federal and local government transportation management services monitor and control the traffic over the urban road network and the nation's highway system. These agencies collect data for both long-term planning and real-time traffic control. Real-time information is usually gathered from many sources, such as electronic sensors in the pavement (loop detectors), road tubes, ramp meter sensors, and video and digital cameras (Toth and Brzezinska, 2006), which are sent to the traffic management centre at various times. Most of this information is only recorded; a small part of it is analyzed in real-time and used for immediate traffic control and decision making. Commonly, the density and flow of traffic are the two main parameters for describing the traffic stream. Namely, the density is the number of vehicles occupying a road lane per unit length at a given time, while traffic flow represents the amount of vehicles traveling over a road segment in a given time period (Pline, 1992).

The key to better traffic management is access to more accurate and more complete data and, of course, the capability for immediate processing of these data to provide a real-time response. Hence the interest in the new sensors that can provide large volumes of data in (near) real-time is steadily growing. Currently, loop detectors produce the largest amount of traffic data (Burns and Wendt, 2003). Although, this sensor technology is well established, the installation and maintenance are not simple, and the associated cost could be also high. Recently, more and more imaging-based systems have been introduced. These pole- or bridge-mounted cameras have shown significant performance improvements and, for example, they are used to calibrate loop-detectors. The trend of switching toward imaging technologies is expected to increase.

The use of pole-mounted camera systems obviously requires the rectification of the oblique viewing angle imagery for effective subsequent feature extraction process used for any traffic management task. Effective traffic management using these pole-mounted camera systems therefore calls for a near real-time rectification process. As opposed to the usual orthorectification process that requires available DEM of the image area and interior and exterior orientation parameters of the collected imagery, this paper proposes a simple method to rectify the oblique imagery to the road surface that in addition to the camera model only requires a minimum of three measured distances from the camera to well distributed ground points in the image area, thereby simplifying the rectification process. The first section of the paper describes the proposed method and then initial test results are shown to prove the effectiveness of the approach.

PROPOSED METHOD FOR RECTIFICATION OF IMAGERY COLLECTED BY FIXED LOCATION IMAGING SENSOR

The objective of this investigation is to develop a simple methodology for creating images rectified to the plane of the road surface from images captured by cameras with oblique viewing angles, mounted on poles next to the road, assuming no a priori information of the camera installation. In other words, the camera relative location with respect to the road surface is measured with the system at installation (at least three distances from the road surface must be known). In general, an orthorectification process would require the road surface DEM and the exterior orientation parameters of the collected imagery. However, assuming a simple planar road surface model, it is sufficient if the distance from the camera to at least three (but preferably more) points on the road surface is measured. Figure 1 shows the basic arrangement with four distances measured at the corners of the imaged road area. The distances can be directly measured by laser ranging or can be determined from absolute coordinate measurements, such as from GPS surveyed camera and road point coordinates.

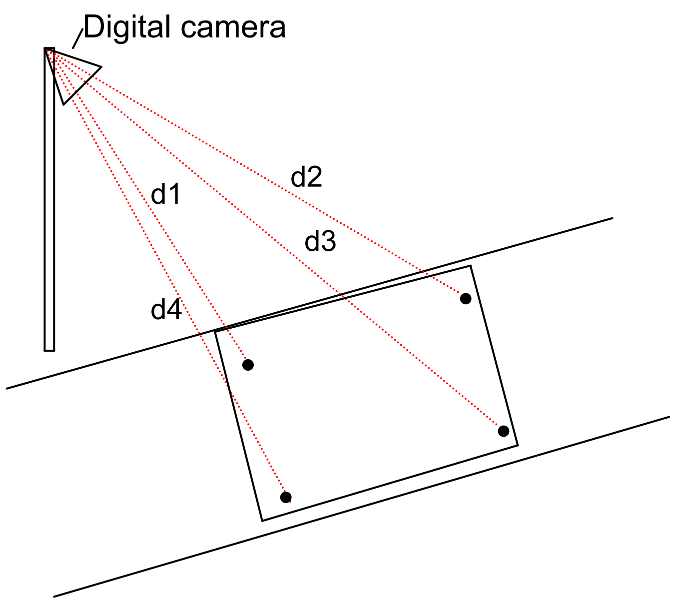


Figure 1. Imaging sensor and road surface spatial relationship.

The proposed method for image rectification described here assumes that the road surface is modeled by a plane; however, the method can easily be extended to more complex road surface models, such as piecewise planar patch-based surface modeling, or using a more complex DEM. The processing steps of the proposed method are as follows:

1. **Computing the measured road point coordinates in the image coordinate system**.
   To simplify the rectification process, all calculations are done in the image coordinate system. Therefore, the first step is to compute the coordinates of the ground points (to which distances from the camera were measured) in the image coordinate system. This can easily be done using the measured image coordinates of these ground points and the measured distances to the points from the camera perspective center as shown in Figure 2 and Equation 1.

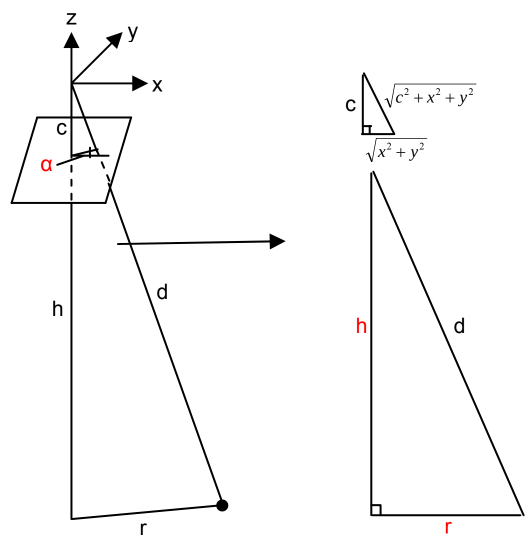


Figure 2. Imaging sensor model and spatial relationship for a point.

$$
\begin{aligned}
h &= \frac{c \cdot d}{\sqrt{c^2+x^2+y^2}} \\
r &= \frac{\sqrt{x^2+y^2} \cdot d}{\sqrt{c^2+x^2+y^2}} \\
\alpha &= \arctan \frac{y}{x} \\
X &= r \cdot \cos\alpha \\
Y &= r \cdot \sin\alpha \\
Z &= -h
\end{aligned}
\qquad \textbf{(1)}
$$

where

x	-	x image coordinate of measured ground point
y	-	y image coordinate of measured ground point
c	-	camera focal length
α	-	angle of image point direction from x axis
d	-	measured distance from perspective center to ground point
h	-	distance of ground point from perspective center along the camera axis
r	-	distance of ground point from perspective center projected to image plane
X	-	X coordinate of ground point in image coordinate system
Y	-	Y coordinate of ground point in image coordinate system
Z	-	Z coordinate of ground point in image coordinate system

2. **Plane fitting to road points in image coordinate system**
The road surface model for the image rectification is determined by fitting a plane to the measured road points. The fitted plane coefficients are computed by least squares adjustment and the plane fitting is done using the ground point coordinates computed in the image coordinate system in the previous step.

3. **Defining a local grid for the rectified image**
Although a simple plane model was considered for the road surface modeling, grid-based DEM has the advantage to work with generic surface models as well as to support the implementation of the actual image rectification. Therefore, the orientation and the spacing of the road grid should be defined. Using the convention of Figure 1, where the four corners of the target road area are measured, Figure 3 shows that the $\mathbf{x}_{\mathbf{grid}}$ axis is defined as a vector going through two ground points (3 and 4) and the $\mathbf{y}_{\mathbf{grid}}$ axis is defined to be perpendicular to the $\mathbf{x}_{\mathbf{grid}}$ axis and lying in the plane of the road surface determined in the previous step. The $\mathbf{y}_{\mathbf{grid}}$ vector is determined mathematically as the vector product of the $\mathbf{x}_{\mathbf{grid}}$ axis and the normal vector of the fitted plane. Finally, both $\mathbf{x}_{\mathbf{grid}}$ and $\mathbf{y}_{\mathbf{grid}}$ axes are normalized to be unit vectors for the generation of the rectified image; the unit for the vector length defined by the desired pixel size of the rectified image.

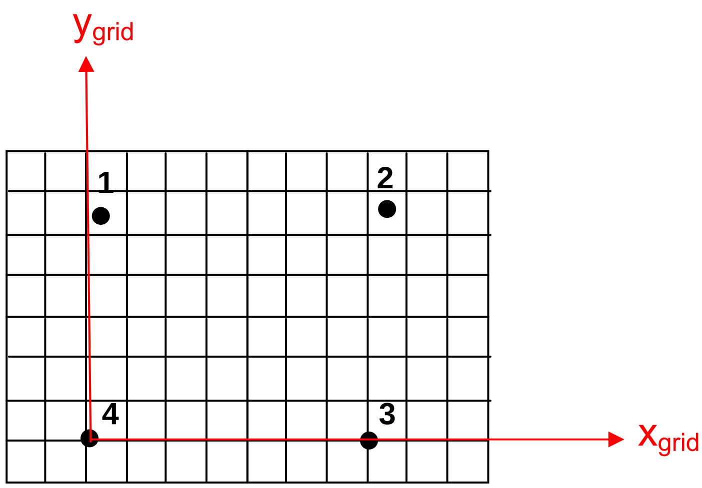


Figure 3. Definition of grid to model the road surface.

4. **Creating rectified image by backward projection**
The backward projection method to create orthophotos is a well-known photogrammetric process (Mikhail et. al., 2001). In simple terms, the image gray value of all the grid points of the DEM is determined to form the orthoimage; the XY object plane of the DEM forms the orthoimage plane. In a sequential process, see Figure 4, all the pixels of the DEM are visited and using the grid value Z (in our case the road elevation data), the X, Y, Z object space coordinates are projected into the source image to obtain the gray level value for the orthoimage pixel. Since the projected object space coordinates will not fall exactly at pixel centers in the source image, the intensity value for that position must be determined by interpolation. In this implementation, bilinear interpolation is used to determine the gray level at the projected location.

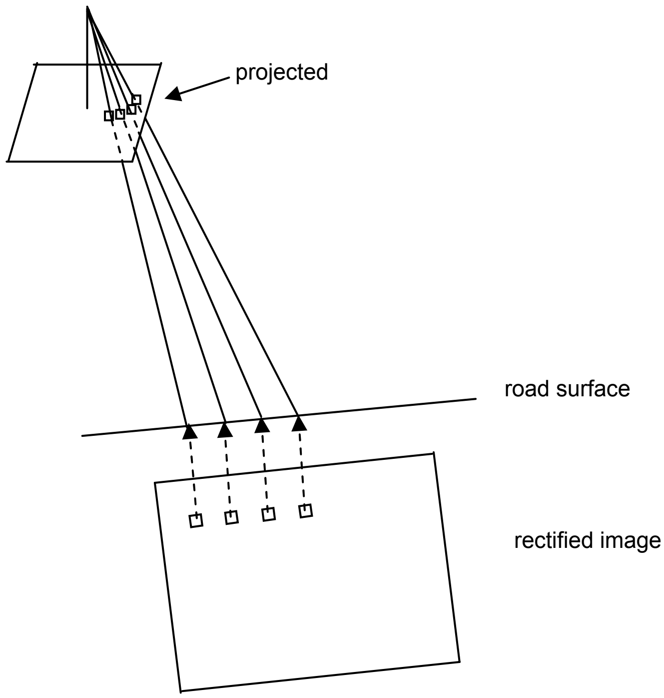


Figure 4. Orthoimage formation process.

Since in our basic case, the road surface (and the defined road grid unit vectors used for the orthorectification) is expressed in the image coordinate system, therefore the collinearity equations, used to describe the general case, for the back-projection are simplified to:

$$x = -c \cdot \frac{X}{Z}$$
$$y = -c \cdot \frac{Y}{Z} \qquad \textbf{(2)}$$

where

x	-	x image coordinate of DEM point
y	-	y image coordinate of DEM point
c	-	camera focal length
X	-	X coordinate of DEM point in image coordinate system
Y	-	Y coordinate of DEM point in image coordinate system
Z	-	Z coordinate of DEM point in image coordinate system

The solution presented here works in a relative coordinate system, as the image coordinate system is used to describe both the image measurements as well as the surface elevations. In this case when the road is modeled with a plane, the created rectified image is parallel to this plane and at the backward projection the Z coordinate of the ground point corresponding to the X, Y pixel position of the rectified image is readily given by the created unit vectors (in step 3) lying in the plane. The same technique can be applied for a piecewise modeling with planar patches. If a more complex model is used, such as airborne or terrestrial LiDAR data or current GIS data, then the general collinearity model should be applied. In addition, since the DEM data come in an absolute frame, the camera must be properly oriented in that coordinate system, too.

INITIAL TEST RESULTS

To perform an in situ test, a prototype system was assembled and a measurement was performed at The Ohio State University campus in summer 2006. The sensor suite included a Canon EOS 1Ds Mark II digital camera (16 Mpixel), Trimble 5700 and Novatel OEM4 GPS receivers. As there was no laser ranging device available, the distances were determined from the relative GPS point measurements.

Data Collection

For this test the camera was mounted on the top of a parking garage and images were captured from the intersection below at an oblique viewing angle. The camera parameters are listed in Table 1.

Table 1. Camera parameters

Focal length	51.387 mm
Pixel size	7.211 µm
Image size	2,496 x 1,664

To determine ranges to points on the road, six points along the two sides of the road were GPS-surveyed. For the measurements, the GPS antenna was mounted on a pole and each point was measured for approximately 5-6 minutes with a data acquisition rate of 1Hz. A GPS antenna and receiver placed near the camera was collecting static GPS data during the whole test. In addition, a second base station with a GPS antenna mounted on the top of the Center for Mapping, The Ohio State University was also running; this was about 1 km from the test area. The six points along the road are marked on an image, as shown in Figure 5.

Figure 5. Typical intersection image with control points marked.

Data Processing

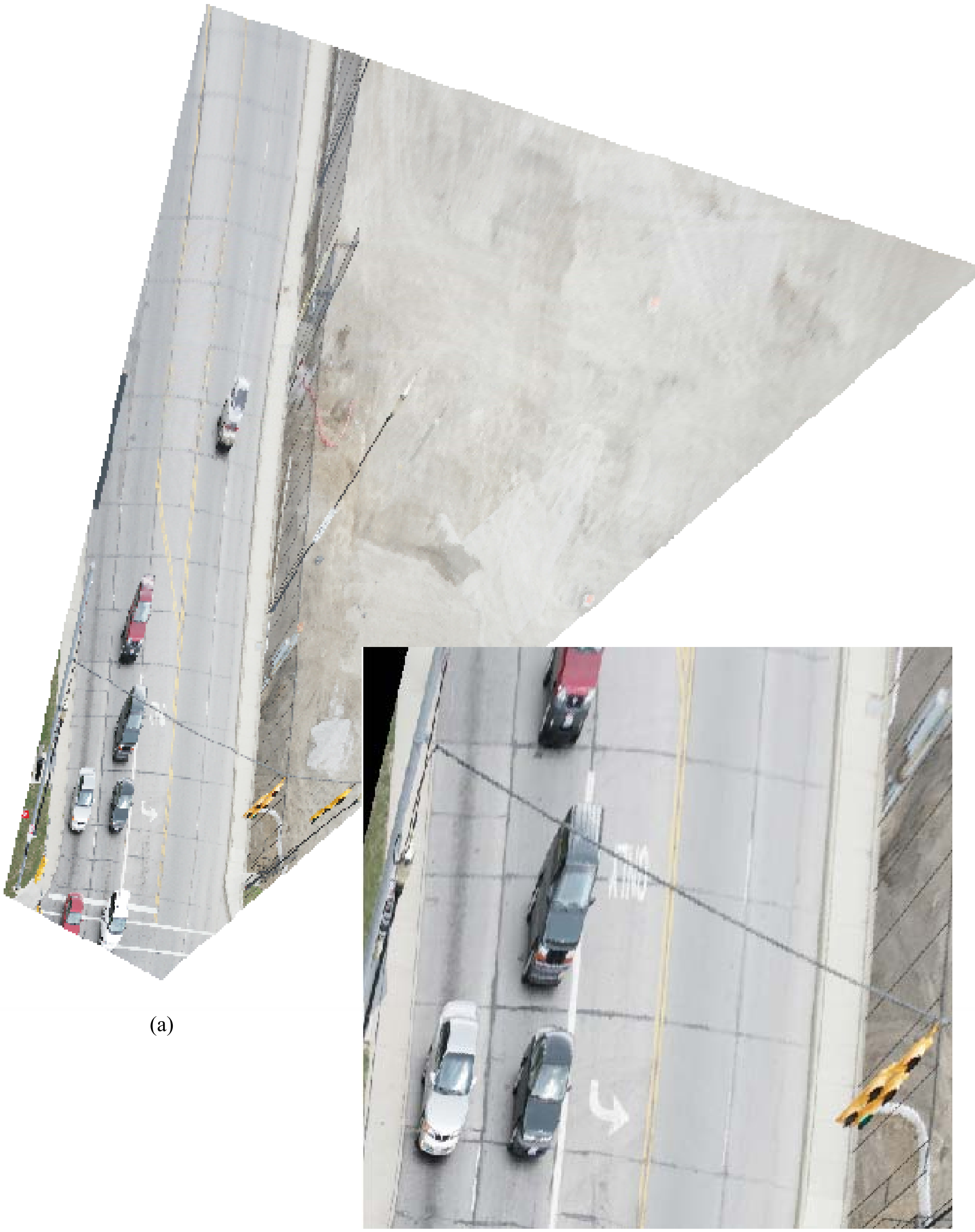
First, the GPS data were processed using the KARS software (Mader, 1992) in kinematic mode. The base station coordinates at the Center for Mapping were determined with cm accuracy using the NGS OPUS solution. Then using this base station, the approximate camera position was determined by KARS. During the processing, kinematic solution was calculated and then the solutions for each epoch were averaged for the final camera position. Once the camera position was determined, the measured points along the road were determined using the GPS antenna at the camera as base, and then averaging the L1 kinematic solutions for each epoch. Unfortunately, for one of the six measured road points (point #4) a good solution could not be determined and therefore only five points were used in the subsequent processing. After the approximate camera position and the road points were computed, the five ranges from the camera were calculated. The image rectification was done at a 5 cm object space resolution. The camera calibration results from an earlier calibration were used.

Rectification Results

A typical image with several vehicles is shown in Figure 6. The distortions and different vehicle sizes due to the oblique viewing angle of the camera is apparent.

Figure 6. Typical intersection image with several vehicles.

The rectified image is shown in Figure 7. Note the vehicle size close to the upper left corner of the image before and after rectification, the ground coverage in the horizontal plane, and the artifacts at the vehicles after rectification. Clearly, the true horizontal scale is beneficial for the subsequent vehicle extraction processes, although the distortion due to the vehicle height presents certain difficulties for vehicle tracking. This noticeable distortion of the vehicles can be significantly decreased in a more realistic camera installation, i.e., higher sensor height and unobstructed near vertical view of the road.



(a)

(b)

Figure 7. Rectified image (a) and detail close-up (b).

CONCLUSIONS

This paper proposed a simple image rectification method for imagery collected by pole-mounted cameras along transportation corridors to facilitate the success of subsequent feature extraction process for traffic flow monitoring tasks. As opposed to usual orthorectification, the proposed method does not require detailed elevation model of the imaged area and exterior orientation parameters of the acquired imagery. Besides the camera model, only a minimum of three distances from the camera to the road surface area have to be measured. The collected imagery is rectified to the road surface. The initial test results have proved that the rectification concept is valid and its implementation works well. Despite the rather low camera height and therefore rather oblique view at the initial test, the rectification resulted in good restoration of the vertical view of the road surfaces, including vehicles. The noticeable distortion of the vehicles can be significantly decreased in a more realistic camera installation, i.e., higher sensor height and unobstructed near vertical view of the road. Future tasks include data collection under more realistic sensor configuration and extending the concept for mobile platforms.

REFERENCES

Burns, S. G. – Wendt, J. J., (2003), Inductive Loop Detector Vehicle Signature Analysis, ITS Institute Advanced Transportation Technologies Seminar Series

Kraus, K., (1993). *Photogrammetry Volume I*, Dummler Books.

Mader, G.L., (1992). Rapid Static and Kinematic Global Positioning System Solutions Using the Ambiguity Function Technique, *Journal of Geophysical Research*, 97, 3271-3283.

Mikhail M. E., J. S. Bethel, J. C. McGlone, (2001). *Introduction to Modern Photogrammetry*, John Wiley & Sons, Inc.

Toth C., D. Grejner-Brzezinska, (2006). Extracting Dynamic Spatial Data from Airborne Imaging Sensors to Support Traffic Flow Estimation, *ISPRS Journal of Photogrammetry & Remote Sensing*, 61, 137-148.

Pline, J.L., (1992). *Transportation and Traffic Engineering Handbook, 4th ed*. Prentice Hall.

US Bureau of Transportation Statistics, http://www.bts.gov