Int. J. Engng Ed. Vol. 22, No. 4, pp. 886–895, 2006
Printed in Great Britain.

0949-149X/91 $3.00+0.00

Traffic Monitoring Using a Three-Dimensional Object Tracking Approach*

R. JOE STANLEY, STEVE E. WATKINS, RANDY H. MOSS and ANAND GOPAL
Department of Electrical and Computer Engineering, University of Missouri-Rolla, 121 Emerson Electric Co. Hall, Rolla, MO 65409-0040, USA. E-mail: stanleyr@umr.edu

An orthographic projections-based algorithm was investigated for the traffic monitoring application of tracking a truck moving across a bridge through a sequence of video image frames. The two-dimensional position of a specific marker on the truck for each video frame is found using color thresholding-based segmentation. A coordinate transformation process provides the real-world, three-dimensional location of the marker position for tracking the truck. The procedure is experimentally illustrated with a field bridge. This paper presents the experimental bridge setup, the marker identification technique, the orthographic coordinate transformation algorithm, and web-accessible laboratory exercises developed for image processing and related curricula.

INTRODUCTION

ADVANCES IN VIDEO hardware and image processing are finding important roles in improving infrastructure, vehicular, and related transportation systems. Automated real-world outdoor scene analysis is a difficult problem that must be addressed in imaging-based systems for monitoring, surveillance, and video security applications. In particular, techniques that use a single camera have obvious economic and logistical advantages. One aspect of object tracking in a sequence of video frames is determining the object's location in the two-dimensional image frames using segmentation techniques. Another aspect of object tracking involves extracting real-world, three-dimensional attributes of that object for identification or classification. In an outdoor scene, there are numerous factors that complicate detecting and tracking objects from a sequence of video frames, including changes in lighting conditions, weather changes and background movement.

Numerous machine vision approaches have been applied to real-world video scene analysis for object recognition and tracking. Computer vision techniques have been applied: 1) to detect and track vehicles and pedestrians in traffic scenes [1–10], 2) to perform real-time analysis and segmentation of two-dimensional image sequences [11], 3) to identify and track multiple objects in complex scenes [12–15], 4) to interpret human behavior [16–19], 5) to track a number of persons in cluttered scenes using depth information [20], and 6) to detect and track people using multiple cameras and a computer network [21]. For many of these camera-based applications, two-dimensional images must be analyzed to extract and translate three-dimensional real-world scene information.

One commonly used approach for video frame analysis is interframe differencing. Interframe differencing is a useful technique to highlight moving objects between frames. These changes may be representative of moving background such as leaves, tree branches, etc., or moving people, trucks or vehicles. Interframe differencing approaches are often based on using three consecutive frames [1, 22]. Approaches such as edge detection and coincidence operators [1] have been used for highlighting potential moving objects.

This work describes an orthographic projections approach to transforming two-dimensional image coordinates to three-dimensional, real-world coordinates in the context of a traffic monitoring application. A single camera is the source of a sequence of two-dimensional video frames. Since an orthographic projections method is used, this technique requires that the depth of the portion of the scene in which measurements are made is small compared to the camera-to-scene distance [23]. The technique requires the real-world location of background marker positions and the identification of a marker position on the object vehicle. This paper discusses an experimental system for testing the approach, object marker identification using color thresholding, and the image processing algorithm. The experiment is based on identifying and tracking markers on a truck as it moves across an on-campus bridge. Sequential real-world images are taken from a dedicated web camera that views the Smart Composite Bridge (SCB) at the University of Missouri-Rolla (UMR). This bridge is an interdisciplinary field laboratory for education and research [24]. The image-processing concepts used for this traffic monitoring application include object segmentation, object labeling, and coordinate transformation.

* Accepted 21 July 2004.

TRAFFIC MONITORING EXPERIMENT

Traffic monitoring on a bridge was the desired operation for the experiment. The monitoring system consisted of a single camera located at a known position with respect to the bridge. The orientation of the bridge was known through a survey of the perimeter of the bridge deck. The perspective of the camera images was identified manually in baseline images; i.e. the pixel locations of selected survey points in the two-dimensional images were found. These pixel locations did not change. The images needed for analysis consisted of an image sequence captured with a truck at different positions on the bridge. A marker on the truck was selected for tracking. The relative height of the truck marker above the bridge deck was also known as well as its distance from one railing of the bridge. The image processing consisted of two main tasks. First, the pixel position of the truck marker was identified in each image. Second, the two-dimensional pixel location of the marker was converted to a three-dimensional real-world coordinate. Hence, the truck position was determined for each image frame and the truck velocity could be calculated from the frame rate. Note that positional information of a bridge load could be used to trigger other measurements, such as strain measurements from structural instrumentation.

Overview of bridge and camera resources

The Smart Composite Bridge at the University of Missouri-Rolla is the field laboratory for the image-processing experiment. This bridge is a short-span structure of length 9.1 m and width 2.8 m. It is located over a small creek on campus and is normally used by pedestrians and light vehicles. However, the bridge was designed and tested for highway loads as a cooperative product development and research effort among university, industry, and government partners [24]. Project documentation is available at http://campus.umr.edu/smarteng/bridge/index.html [25, 26].

A dedicated web camera provides an online imaging capability as part of the bridge research initiative. It is mounted on the side of the Emerson Electric Co. Hall (EECH) (Electrical and Computer Engineering Building) and views the bridge from an oblique angle at a distance of approximately 61.8 m. The camera is an Axis 200+ Web Camera (Axis Communications Model Number 0064-9) which contains a 768×582 24-bit digital color CCD camera and a self-contained web server (32-bit RISC processor). A zoom lens and an outdoor housing complete the system. Figure 1 shows the bridge and the camera, highlighted with the white box, mounted on EECH. The resulting real-world images have several interesting features. Trees and railings obscure part of the bridge; sunlight and shade change the background lighting; and seasonal coloration changes the overall background.

Experimental description

Figure 2 presents an example frame acquired from the mounted web camera and is representative of the frames captured in the total sequence of images. The sequence of image frames begins with the truck on the southwest end of the bridge and progresses as the truck moves right-to-left across the two-dimensional bridge scene. The truck travels in parallel with the sides of the bridge deck as it moves toward the northeast end of the bridge. For the parallel movement, the truck remains a fixed distance from the northwest and southeast sides of the bridge deck. For this initial experiment, the parking lights of the truck were turned on to serve as object markers. These

Fig. 1. The Smart Composite Bridge at the University of Missouri-Rolla. The location of the dedicated web camera is indicated on the building in the background.

Fig. 2. Frame acquired from the dedicated web camera of the Smart Composite Bridge. The traffic-monitoring experiment uses the right parking light of the truck as a marker for tracking.

markers or reference positions on the truck remain a fixed distance above the bridge deck. Twelve frames were captured for each image sequence. The truck moved in increments of 0.914 meters (3 feet) across the bridge.

The first image processing task was to identify the object marker in each frame. Histogram analysis and color thresholding techniques were used for segmenting the parking lights as an object marker. The centroid of the identified marker provided the two-dimensional pixel location in each frame. Object-labeling maintained the specific marker tracked through the sequence of frames. The second image-processing task was to convert the two-dimensional marker pixel locations to the three-dimensional coordinates. The orthographic projection technique used the bridge reference markers and the object marker to convert the image position coordinates to real-world spatial locations. The details of marker segmentation and the derivation of the coordinate transformation procedure are given in the next section. Both tasks were coded in Matlab® using the Imaging Processing Toolbox. Example image sequences and the code for the coordinate transformation are available for download at the SCB website [26]. An associated educational exercise was reported in a prior paper [27].

A traffic monitoring problem is presented for tracking a vehicle in a sequence of video frames. The orthographic projection-based approach can determine three-dimensional locations for a specified marker from a sequence of two-dimensional video frames using a single camera source. The two-dimension to three-dimension coordinate transformation technique requires the real-world location of bridge marker positions and a marker position on the object vehicle for a reference two-dimensional video frame. A simplified variation of orthographic projection [28] is used to perform the coordinate transformation.

Education and research for transportation are becoming more interdisciplinary. Optical metrology, smart materials and structures, computer technologies, etc., are finding important roles in improving infrastructure, vehicles, and related transportation systems. For instance, image processing can provide intelligent monitoring of traffic flows and control functions of other sensing systems. Consequently, people outside of traditional transportation fields with exposure to interdisciplinary transportation applications can contribute to needed research and development.

TRAFFIC-MONITORING IMAGE PROCESSING ALGORITHM DESCRIPTION

The traffic monitoring application involves tracking a moving vehicle across video frames for determining the position, requiring the coordinate transformation of two-dimensional images to three-dimensional, real-world positions of vehicle markers. The algorithm will be described in the context of the experiment (cf. Fig. 1). The truck in Fig. 2 maintains the parallel movement at a fixed distance from the bridge's deck as the truck moves across the bridge in the captured frames, moving from right to left.

Coordinates and procedure

Figure 3 shows a schematic of the bridge deck and three-dimensional coordinate system with origin at the camera lens of the camera mounted on EECH. In Fig. 3 the XYZ axes are drawn to show the orientation of the coordinate system. The real-world coordinates of the bridge deck corner positions (labeled 1–4) are surveyor measurements

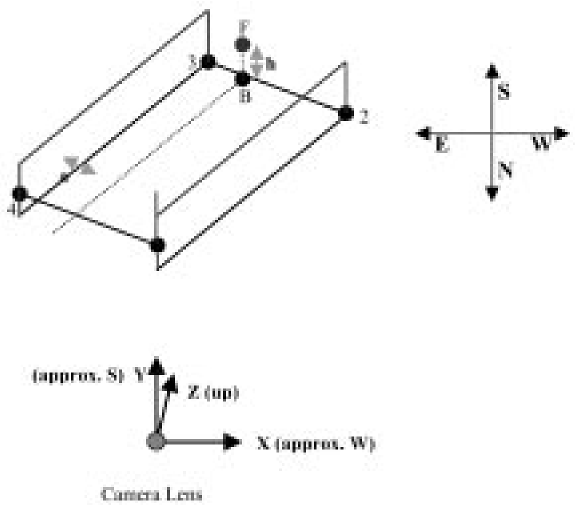

Fig. 3. Schematic of bridge deck with reference points labeled for use in two-dimensional to three-dimensional coordinate transformation. The coordinate axes and direction legend are shown. The dotted line within the bridge deck represents the path that the right wheels of the truck follow for crossing the bridge, starting at point B.

in meters from the camera lens origin. A directional legend is also provided in Fig. 3 to assist in referencing the labeled positions. All calculations required in the coordinate transformations presented are based on the positions labeled 1–4. In order to transform the two-dimensional coordinates within the video frames to three-dimensional spatial coordinates, we apply a variation of the technique of orthographic projections. The coordinate transformation procedure is to: 1) identify required real-world coordinate reference positions, including the labeled positions 1–4 in Fig. 3, the starting front right truck ground position (B) and the starting position of the truck's right front parking light (F), 2) derive real-world coordinates for reference points B and F, 3) determine the image coordinates for the centroid position of the truck's right front parking light at the truck's starting position, and 4) transform the segmented front right parking light centroid position to its real-world coordinate within the various image frames. The real-world position of the truck's right front parking light is used for tracking the truck's position from video frame to video frame. Two required parameters for performing the transformation calculations, labeled on Fig. 3, are: 1) the height of the truck's right parking light above the bridge deck, denoted as h and 2) the perpendicular distance of the right parking light from the bridge side 3–4, denoted as a. The parking light centroid image location transformation to its corresponding real-world position is described in the following sections.

Color thresholding

From the captured frames, color ranges for the markers, i.e. the parking lights, on the truck (as seen in Fig. 2) were manually determined from the image frame data. The parking lights provided global red, green and blue (RGB) color thresholds for segmentation. After thresholding and segmentation, the centroid pixel location is determined in the image. Other types of markers and more flexible segmentation approaches will be the subject of future research.

The transformation process provides the real-world coordinates of the truck's right parking light from the corresponding image coordinates found for the centroid image location of the right parking light. The centroid location, denoted (u_T, v_T), is obtained for each video frame as the output of the image processing operations.

Reference real-world positions

The bridge deck corner points are used in the transformation process. From Fig. 3, the real-world coordinates of the deck corner points with respect to the origin at the web camera lens are:

1. north corner of the bridge deck with coordinate (x_1, y_1, z_1);
2. west corner of the bridge deck with coordinate (x_2, y_2, z_2);
3. south corner of the bridge deck with coordinate (x_3, y_3, z_3); and
4. east corner of the bridge deck with coordinate (x_4, y_4, z_4).

Figure 2 shows the orientation of the bridge deck within the acquired images. Note that the north corner of the bridge deck is obscured by a tree. However, this pixel location can be extrapolated from other perimeter survey measurements. The labeled points 1–4 have known real-world locations with respect to the camera lens and are given in meters as $(x_1, y_1, z_1) = (0,\ 56.47,\ -8.94)$, $(x_2, y_2, z_2) = (3.90,\ 64.65,\ -8.92)$, $(x_3, y_3, z_3) = (1.52,\ 65.82,\ -8.90)$ and $(x_4, y_4, z_4) = (-2.37,\ 57.59,\ -8.92)$. The real-world coordinates for two additional reference points, B and F, are derived in the following subsection. B refers to the starting ground position beneath the truck's front right parking light on the bridge deck with coordinate (x_B, y_B, z_B), and F denotes the starting centroid position of the truck's right front parking light (x_F, y_F, z_F).

The truck starts at position B in Fig. 3 and moves in parallel with respect to side 3–4 (from Fig. 3) across the bridge, with the right front parking light maintaining a fixed perpendicular distance a from side 3–4. Both truck parking lights maintain a constant height h above the bridge deck. Thus, the reference front right parking light centroid point maintains a height h and distance a as the truck crosses the bridge. For the truck used in the acquired images (seen in Fig. 2), $\boldsymbol{h} = 0.91$ m and $\boldsymbol{a} = 0.44$ m. Because the bridge deck is not entirely level, there are residual angles of the deck sides with respect to the coordinate axes. These angles must be computed from the deck corner positions for projections calculations. These constraints provide the need to derive the real-world reference points B and F.

Real-world coordinate derivation for truck reference points

The next step in the coordinate transformation process is to determine the real-world coordinates for reference points B and F for the truck's starting ground position on the bridge deck and the truck's right front parking light centroid. The following calculations are made between the labeled points from Fig. 3. Using vector notation, let $\vec{V}_{34}$ and $\vec{V}_{32}$ denote the vectors from bridge deck positions 3 to 4 and 3 to 2, respectively, and are given as

$$\vec{V}_{34} = (x_4 - x_3)\vec{a}_x + (y_4 - y_3)\vec{a}_y + (z_4 - z_3)\vec{a}_z \quad (1)$$

$$\vec{V}_{32} = (x_2 - x_3)\vec{a}_x + (y_2 - y_3)\vec{a}_y + (z_2 - z_3)\vec{a}_z \quad (2)$$

where $\vec{a}_x$, $\vec{a}_y$, and $\vec{a}_z$ are the unit vectors in the X, Y, and Z directions, respectively. First, determine the vector from bridge deck positions B to F as the cross product-based expression

$$\vec{V}_{BF} = \left(\frac{\vec{V}_{34} \times \vec{V}_{32}}{|\vec{V}_{34} \times \vec{V}_{32}|}\right)\mathbf{h} \quad (3)$$

using a right-handed coordinate system. The denominator of the expression is given by

$$|\vec{V}_{34} \times \vec{V}_{32}| = \sqrt{A^2 + B^2 + C^2}, \text{ where} \quad (4)$$

$$A = (y_4 - y_3)(z_2 - z_3) - (y_2 - y_3)(z_4 - z_3), \quad (5)$$

$$B = -\left[(x_4 - x_3)(z_2 - z_3) - (x_2 - x_3)(z_4 - z_3)\right], \quad (6)$$

$$C = (x_4 - x_3)(y_2 - y_3) - (x_2 - x_3)(y_4 - y_3). \quad (7)$$

Second, find the length of the southwest side of the bridge's deck from deck corner point 2 to point 3 as

$$D_{23} = \sqrt{(x_3 - x_2)^2 + (y_3 - y_2)^2 + (z_3 - z_2)^2}. \quad (8)$$

Third, determine the projected length of D_{23} onto the XZ and XY planes, denoted as $P_{23,XZ}$ and $P_{23,XY}$, respectively, as given by the expressions

$$P_{23,XZ} = \sqrt{(x_3 - x_2)^2 + (z_3 - z_2)^2}, \quad (9)$$

$$P_{23,XY} = \sqrt{(x_3 - x_2)^2 + (y_3 - y_2)^2}. \quad (10)$$

Fourth, compute the distance between deck corner position 3 and the projection of the truck's right parking light ground position (B) in the XZ plane, which is given by

$$D_{3,B_{XZ}} = \mathbf{a}\left(\frac{P_{23,XZ}}{D_{23}}\right). \quad (11)$$

Fifth, find the distance between deck corner point 3 and the projection of the truck's right parking light ground position (B) in the XY plane as given by

$$D_{3,B_{XY}} = \mathbf{a}\left(\frac{P_{23,XY}}{D_{23}}\right). \quad (12)$$

Sixth, determine the angle (ϕ) of bridge deck side 2–3 projected onto the XZ plane with respect to the X-axis, where

$$\phi = \tan^{-1}\left(\frac{z_3 - z_2}{x_3 - x_2}\right). \quad (13)$$

Seventh, compute the angle (β) of bridge deck side 2–3 projected onto the XY plane with respect to the X-axis, where

$$\beta = \tan^{-1}\left(\frac{y_3 - y_2}{x_3 - x_2}\right). \quad (14)$$

Eighth, calculate the real-world coordinates of reference point B, given as

$$x_B = x_3 + D_{3,B_{XZ}} \cos(\phi), \quad (15)$$

$$y_B = y_3 - D_{3,B_{XY}} \sin(\beta), \quad (16)$$

$$z_B = z_3 - D_{3,B_{XZ}} \sin(\phi). \quad (17)$$

Ninth, determine the angle (α) of segment B–F with respect to the X-axis, where

$$\alpha = \cos^{-1}\left(\frac{A}{\sqrt{A^2+B^2+C^2}}\right). \quad (18)$$

Tenth, find the angle (λ) of segment B–F with respect to the Y-axis, where

$$\lambda = \cos^{-1}\left(\frac{B}{\sqrt{A^2 + B^2 + C^2}}\right). \quad (19)$$

Eleventh, compute the angle (σ) of segment B–F with respect to the Z-axis, where

$$\sigma = \cos^{-1}\left(\frac{C}{\sqrt{A^2 + B^2 + C^2}}\right). \quad (20)$$

Finally, the real-world coordinates of reference point F are given as

$$x_F = x_B + \mathbf{h}\cos(\alpha), \quad (21)$$

$$y_F = y_B + \mathbf{h}\cos(\lambda), \quad (22)$$

$$z_F = z_B + \mathbf{h}\cos(\sigma). \quad (23)$$

The real-world coordinates for reference positions B and F and the labeled deck corner points 1–4 are used in the transformation of the image frame front right parking light centroid position to its real-world coordinate location.

Initial parking light real-world coordinate transformation to image pixel location

From the initial image frame, the real-world coordinate of the right front parking light centroid position must be translated into a corresponding two-dimensional pixel coordinate. Let U and V denote the horizontal and vertical axes, respectively, for the M column $\times$ N row image frames. Figure 4 shows the bridge deck projected onto the image plane and the orientation of the axes U and V used for processing the image frames. For the U and V axes, the coordinate $(1, 1)$ is located in the upper left corner of the image, and (M, N) is located in the lower right corner of the image. The U and V axes from Fig. 4 provide parallel representations to the X and –Z axes from Fig. 3, respectively. The following calculations are performed to obtain the pixel coordinates of the truck's right front parking light (u_P, v_P) at the truck's starting position. Let (u_2, v_2) and (u_3, v_3) denote the image coordinates for bridge deck corner positions 2 and 3, respectively, from Fig. 4. These points are determined from manual image inspection. First, find the distance (C_{23}) between image coordinates corresponding to deck corner positions 2 and 3 from Fig. 3. Second, determine the number of pixels a unit of projected real-world distance occupies as

$$K = \frac{C_{23}}{P_{23,XZ}}. \tag{24}$$

Third, compute the distance in pixels between position 3 and the projection of position B onto the XZ plane as

$$C_{3,B_{XZ}} = KD_{3,B_{XZ}}. \tag{25}$$

Fourth, determine the angle (ω) of bridge deck side 2–3 within the image with respect to the U-axis, where

$$\omega = \tan^{-1}\left(\frac{v_2 - v_3}{u_2 - u_3}\right), \tag{26}$$

where ω is measured clockwise from the U axis. Fifth, find the image pixel coordinates of reference position B (u_B, v_B), which are given as

$$u_B = u_3 + C_{3,B_{XZ}} \cos(\omega), \tag{27}$$

$$v_B = v_3 + C_{3,B_{XZ}} \sin(\omega). \tag{28}$$

Sixth, determine the angle (θ) of reference points B–F with respect to the U-axis, which is given as

$$\theta = -\tan^{-1}\left(\frac{(x_4 - x_3)(y_2 - y_3) - (x_2 - x_3)(y_4 - y_3)}{(y_4 - y_3)(z_2 - z_3) - (y_2 - y_3)(z_4 - z_3)}\right). \tag{29}$$

Seventh, calculate the image pixel coordinates of reference position F (u_F, v_F), which are given as

$$u_F = u_B + K\,h \cos(\theta), \tag{30}$$

$$v_F = v_B - K\,h \sin(\theta). \tag{31}$$

The result obtained from these operations yields the image coordinates of the starting position of the truck's right front parking light (F) from the real-world coordinates of known positions on the bridge deck.

Image frame parking light centroid pixel to real-world coordinate transformation

Finally, the equations must be derived for translating the two-dimensional centroid locations of parking lights segmented within the remaining image frames to the corresponding three-dimensional real-world coordinates. The coordinate space from Fig. 3 is used to represent the real-world coordinates from the image coordinate transformation. As previously stated, the image coordinates of the centroid location for the front right parking light within an image frame is denoted as (u_T, v_T). The following operations must be performed to obtain the corresponding real-world coordinates (x_T, y_T, z_T). First, determine the distance in pixels between reference (u_F, v_F) and (u_T, v_T) as CFT. Second, compute the distance in pixels between positions 3 and 4 within the image as C_{34}. Third, convert the pixel distance into real-world distance in the XZ plane given as

$$P_{FT,XZ} = \frac{P_{34,XZ} C_{FT}}{C_{34}}. \tag{32}$$

This assumes that the distance from camera to object is constant for all objects on the bridge; i.e. magnification is constant. Fourth, compute the angle (ψ) between 3–4 and the X-axis in the XZ plane as

$$\psi = \sin^{-1}\left(\frac{z_3 - z_4}{P_{34,XZ}}\right). \tag{33}$$

Fifth, find x_T and z_T, which are given as

$$x_T = x_F - P_{FT,XZ} \cos(\psi), \tag{34}$$

$$z_T = z_F - P_{FT,XZ} \sin(\psi). \tag{35}$$

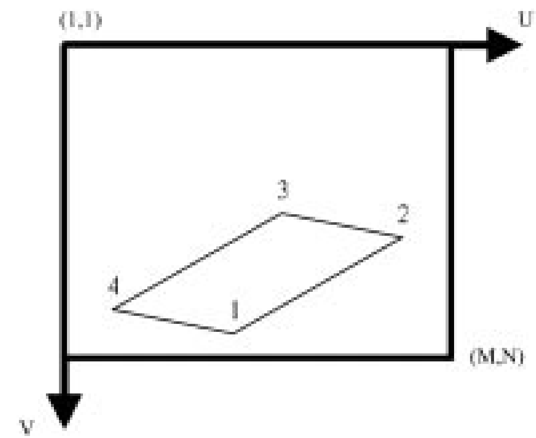

Fig. 4. Coordinate system used for image frames with axes U and V.

Sixth, find the angle (τ) between 3–4 and the X-axis in the XY plane, where

$$\tau = \sin^{-1}\left(\frac{y_3 - y_4}{P_{34,XY}}\right). \quad (36)$$

Seventh, convert the pixel distance into real-world distance in the XY plane given as

$$P_{FT,XY} = \frac{P_{FT,XZ}P_{34,XY}}{P_{34,XZ}}. \quad (37)$$

Eighth, determine y_T, where

$$y_T = y_F - P_{FT,XY}\sin(\tau). \quad (38)$$

This derivation is for the right parking light. The derivation can be easily extended to the left parking light.

MARKER POSITION ERROR DETERMINATION

In this section, the procedure for determining the real-world position of the left and right front parking lights is applied and evaluated from the individual image frames. Specifically, a laboratory exercise was given to 16 students of a Machine Vision course at UMR during the fall semester 2002. The exercise required the students to manually determine red, green, and blue color thresholds to segment the left and right parking lights of the truck for three image frames acquired in the springtime using the camera setup described in section *II. A. Overview of Bridge and Camera Resources*, including the image frame shown in Fig. 2. The centroids of the resulting thresholded objects were computed to represent the marker position. The absolute difference in pixels and real-world distance between the threshold-based marker positions and reference marker positions were found in order to evaluate the impact of image segmentation of the marker position. The boundaries of the left and right parking lights in the three image frames were manually found and filled. The centroids of the resulting filled regions were used as the reference marker positions of the parking lights in the image frames. Using the image-coordinate to real-world transformation procedure, the centroid marker positions were converted into their corresponding real-world coordinates. For the image frames examined, the distances from the rail **a** are 0.44 m and 2.16 m for the right and left parking lights, respectively. Table 1 shows the average and standard deviation absolute difference results in pixels and in real-world distance (in meters) between the student thresholds selected for automatic segmentation to determine the centroid locations for the left and right parking lights and the reference positions for each image frame.

From Table 1, the average absolute pixel difference between the student-determined marker position centroid and the corresponding reference position centroid locations ranged from 0.283 to 0.371 pixels with a corresponding average real-world distance range of 0.009 to 0.012 meters. Note that the image coordinate to real-world transformation is nonlinear and that pixel variations result in small real-world position variations, as seen in the average absolute difference results in Table 1. As seen in Fig. 2, the orientation of the truck with respect to the camera position hampers the identification of the 'true' left and right parking light centroid locations. Overall, there are minor differences between the absolute difference pixel and real-world centroid positions for the left and right parking lights for the three image frames examined. Depending on the precision required for tracking the truck, the average absolute difference distance results are relatively low, facilitating a reasonable tracking of the truck's marker position across image frames at sampling rates such as 1/30 second.

The traffic monitoring application presented is oversimplified compared to typical, realistic three-dimensional computer vision tasks. For this application, reference measurements were chosen, which were easy to make, that simplified and improved the accuracy of the coordinate conversion process, which also simplified the application for teaching students coordinate transformation techniques. In order to apply the algorithm presented to an actual traffic-monitoring situation, the height of the marker (h) (centroid locations of the parking lights used in our approach) and the distance of the marker from a bridge railing ($\boldsymbol{a}$) must be known. If the geometry of the imaging situation is such that the road is planar and the camera angle allows for sufficient accuracy, then

Table 1. Average and standard deviation absolute difference in pixels and meters between the calculated marker position and the actual marker position for three image frames obtained from 16 students in a Machine Vision course.

	Pixel distance difference between color threshold-based centroid and reference centroid		Real-world absolute distance difference between color threshold-based centroid and reference centroid (meters)	
Frame Number	Average	Standard Deviation	Average	Standard Deviation
Frame 1	0.283	0.247	0.011	0.012
Frame 2	0.371	0.350	0.012	0.011
Frame 3	0.283	0.321	0.009	0.012

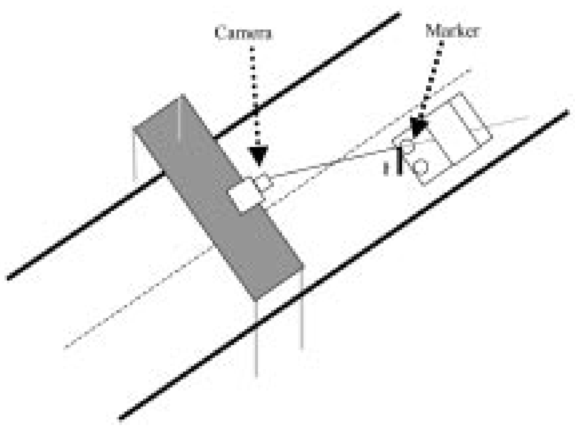

Fig. 5. Illustration of the camera and road configuration that could be used with markers on the vehicles to apply the projections-based approach for vehicle tracking.

the real-world position of the marker could be found if only the height of the marker were known. Markers could conceivably be placed on vehicles at a given height just for this purpose. Fig. 5 shows an illustration of the camera and road configuration that could be used with markers on the vehicles to apply the projections-based approach for vehicle tracking. From Fig. 5, the 3-D geometry of camera and road shows that, if the angle of the ray from the camera is known (from the pixel position) and the height of the marker above the road is known, then the real-world position can be found.

CONCLUSIONS

A traffic monitoring problem is presented for tracking a vehicle in a sequence of video frames. The orthographic projection-based approach can determine three-dimensional locations for a specified marker from a sequence of two-dimensional video frames using a single camera source. The two-dimension to three-dimension coordinate transformation technique requires the real-world location of bridge marker positions and a marker position on the object vehicle for a reference two-dimensional video frame. The coordinate transformation process uses these parameters to track the position of the vehicle. The movement of a truck across the Smart Composite Bridge at the University of Missouri-Rolla illustrated the technique. Processing of each image must discriminate markers in the presence of differing lighting conditions, weather, etc. The truck example used color thresholding to detect the parking lights as markers and a dedicated web camera as the image source

The orthographic projection approach and image sequences from the Smart Composite Bridge have been incorporated into a Matlab®-based student exercise [27]. Two downloadable versions of the traffic monitoring exercise are available in the 'Educational Resources' section of the Smart Composite Bridge website (http://campus.umr.edu/smarteng/bridge/index.html) [26] to facilitate student learning at other institutions. The site includes assignment guidelines, image files, sample Matlab® code, and different image sets. One version of the sample code implements the variation of the orthographic projections technique presented in this paper for translating the two-dimensional bridge scene image coordinates to the three-dimensional bridge scene coordinates. The more challenging version provides the bridge survey information and the image sequences needed to derive an image to real-world coordinate transformation. Future research will address object identification, people-tracking applications, alternative methods for identifying object markers, and coordinated triggering of smart strain instrumentation.

Acknowledgements—This monitoring development was partially supported by the National Science Foundation (NSF) through Combined Research-Curriculum Development Grant # EEC-9872546 with Mary Poats as technical contact. The Smart Composite Bridge was primarily sponsored by the same NSF grant. It is documented at http://campus.umr.edu/smarteng/bridge/index.html.

REFERENCES

1. A. T. Ali and E. L. Dagless, Computer vision for security surveillance and movement control, *Proceedings of the IEE Colloquium on Electronic Images and Image Processing in Security and Forensic Science*, **6** (1990), pp. 1–7.
2. C. J. Radford, Vehicle detection in open-world scenes using a hough transform technique, *Proceedings of the. IEE 3rd International Conference on Image Processing and Its Applications* (1989), pp. 78–82.
3. G. S. Hobson, H. Roe, J. Hawley, A. D. Houghton, M. Seed and N. L. Seed, Traffic tracker, *Proceedings of the IEE Colloquium on Image Analysis for Transport Applications*, **4** (1990), pp. 1–4.
4. D. Gao and J. Zhou, Adaptive background estimation for real-time traffic monitoring, *Proceedings of the IEEE Conference on Intelligent Transportation Systems* (2001), pp. 330–333.
5. R. Cucchiara, M. Piccardi and P. Mello, Image analysis and rule-based reasoning for a traffic monitoring system, *Proceedings of the IEEE Conference on Intelligent Transportation Systems*, **1**(2) (2000), pp. 119–130.
6. J. M. Manendez, L. Salgado, E. Rendon and N. Garcia, Motorway surveillance through stereo computer vision, *Proceedings of the International Carnahan Conference on Security Technology* (1999), pp. 197–202.
7. B. Gloyer, H. K. Aghajan, S. Kai-Yeung and T. Kailath, Vehicle detection and tracking for freeway monitoring, *Proceedings of the 28th Asilomar Conference on Signals, Systems and Computers*, **2** (1994), pp. 970–974.

8. O. Masoud and N. P. Papanikolopoulos, A novel method for tracking and counting pedestrians in real-time single camera, *IEEE Transactions on Vehicular Technology*, **50**(5) (2000), pp. 1267–1278.
9. C. Curio, J. Edelbrunner, T. Kalinke, C. Tzomakas and W. von Seelen, Walking pedestrian recognition, *IEEE Transactions on Intelligent Transportation Systems*, **1** (2000), pp. 155–163.
10. L. Zhao and C. E. Thorpe, Stereo- and neural network-based pedestrian detection, *IEEE Transactions on Intelligent Transportation Systems*, **1**(3) (2000), pp. 148–154.
11. S. M. Smith, ASSET-2: Real-time motion segmentation and object tracking, *Real Time Imaging*, **4** (1998), pp. 21–40.
12. G. L. Foresti, Object detection and tracking in time-varying and badly illuminated outdoor environments, *Opt. Eng.*, **37**(9) (1998), pp. 2550–2564.
13. F. Meyer and P. Bouthemy, Region-based tracking using affine motion models in long image sequences, *Computer Vision, Graphics and Image Processing: Image Understanding*, **60**(2) (1994), pp. 119–40.
14. D. B. Gennery, Visual tracking of known 3D object, *Int. J. Comput. Vis.*, **7**(3) (1992), pp. 243–270.
15. Z. Li and H. Wang H, Real-time 3-D motion tracking with known geometric models, *Real Time Imaging*, **5** (1999), pp. 167–187.
16. I. Haritaoglu, D. Harwood and L. S. Davis, W4: Real-time surveillance of people and their activities, *IEEE Trans. Pattern Anal. Mach. Intell.*, **22** (2000), pp. 809–830.
17. N. M. Oliver, B. Rosario and A. P. Pentland, A Bayesian computer vision system for modeling human interactions, *IEEE Trans. Pattern Anal. Mach. Intell.*, **22**(8) (2000), pp. 831–843.
18. Y. Ricquebourg and P. Bouthemy, Real-time tracking of moving persons by exploring spatio-temporal image slices, IEEE Trans. *Pattern Anal. Mach. Intell.*, **22**(8) (2000), pp. 797–780.
19. F. Bremond and M. Thonnat, Tracking multiple nonrigid objects in video sequences, *IEEE Trans. Circuits Syst. Video Technol.*, **8** (1998), pp. 585–591.
20. R. Luo and Y. Guo, Real-time stereo tracking of multiple moving heads, *Proceedings of the IEEE ICCV Works on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems* (2001), pp. 55–60.
21. T. Fukuda, T. Suzuki, F. Kobayashi, F. Arai, Y. Hasegawa and M. Negi, Seamless tracking system with multiple cameras, *Proceedings of the 26th Annual Conference of the IEEE Industrial Electronics Society*, **2** (2000), pp. 1249–1254.
22. K. Yoshinari and M. Michihito, A human motion estimation method using 3-successive video frames, *Proceedings of the International Conference on Virtual Systems and Multimedia Gifu* (1996), pp. 135–140.
23. B. K. P. Horn, *Robot Vision*, MIT Press, Cambridge, MA (1992), pp. 21–22.
24. S. E. Watkins, J. F. Unser, A. Nanni, K. Chandrashekhara and A. Belarbi, Instrumentation and manufacture of a smart composite bridge for short-span applications, *Proceedings of the SPIE 4330, Smart Structures and Materials 2001: Smart Systems for Bridges, Structures, and Highways* (2001), pp. 147–157.
25. S. E. Watkins, V. M. Eller and R. H. Hall, WWW instructional documentation for the development of a smart composite bridge, *Int. J. Continuing Engineering Education and Lifelong Learning*, **13**(1/2) (2003), pp. 180–189.
26. S. E. Watkins and R. H. Hall. Smart composite bridge and educational resources, *Smart Engineering Project*. Available at http://campus.umr.edu/smarteng/ (2004).
27. R. J. Stanley, S. E. Watkins, A. Gopal and R. H. Moss, A web-sharable real-world imaging problem for enhancing an image processing curriculum, *IEEE Transactions on Education*, **47**(2) (2004), pp. 211–219.
28. Y. Lamdan, J. T. Schwartz and H. J. Wolfson, On recognition of 3-d objects from 2-d images, *Proceedings of the IEEE Int. Conf. On Robotics and Automation* (1998), pp. 1407–1413.

R. Joe Stanley received the B.S.E.E. and M.S.E.E. degrees in Electrical Engineering and a Ph.D. degree in Computer Engineering and Computer Science from the University of Missouri-Columbia. As a graduate student at the University of Missouri-Columbia, he worked under training grants from the National Library of Medicine and the National Cancer Institute. Upon completing his doctoral study, he served as Principal Investigator for the Image Recognition program at Systems & Electronics, Inc., in St. Louis, MO. He is currently an Assistant Professor in the Department of Electrical and Computer Engineering at the University of Missouri-Rolla. His research interests include signal and image processing, pattern recognition and automation.

Steve E. Watkins is Director of the Applied Optics Laboratory and Professor of Electrical and Computer Engineering at the University of Missouri-Rolla. He directs the Smart Engineering Project and the Smart Composite Bridge development. His research includes several interdisciplinary projects addressing the development of engineering curricula and web-based resources and the application of smart-structures and fiber-optic sensor systems. He is a member of SPIE, senior member of the Institute of Electrical and Electronics Engineers (IEEE), the 2000 recipient of the IEEE Region 5 Outstanding Engineering Educator Award, a 1993 finalist in the Outstanding Young Engineer Award Program from the Eta Kappa Nu Association, and a past National Science Foundation Graduate Fellow. He has been a visiting physicist at Kirtland Air Force Base and a visiting scholar at NTT in

Japan. He received his Ph.D. in electrical engineering from the University of Texas at Austin in 1989.

Randy H. Moss received the B.S.E.E. and M.S.E.E. degrees in Electrical Engineering from the University of Arkansas, where he was a National Merit Scholar, and the Ph.D. degree from the University of Illinois, where he was an NSF Graduate Fellow. He is currently a Professor of Electrical Engineering at the University of Missouri-Rolla. His research interests are in the areas of image processing, pattern recognition and computer vision. He is especially interested in medical and industrial applications of machine vision techniques. He serves as an Associate Editor of Pattern Recognition and Computerized Medical Imaging and Graphics. Dr. Moss is a member of Sigma Xi, the Pattern Recognition Society, Eta Kappa Nu, Tau Beta Pi and Phi Kappa Phi.

Anand Gopal is completing his M.Sc. degree in electrical engineering at the University of Missouri-Rolla. He received the Bachelor of Engineering degree in electrical and electronics engineering from Crescent Engineering College in 2001.