Estimating relative vehicle motions in traffic scenes

Zoran Duric[a,b], Roman Goldenberg[c], Ehud Rivlin[a,c], Azriel Rosenfeld[a,*]

[a]*Center for Automation Research, University of Maryland, College Park, MD 20742-3275, USA*
[b]*Computer Science Department, George Mason University, Fairfax, VA 22030-4444, USA*
[c]*Department of Computer Science, Technion, Israel Institute of Technology, Haifa 32000, Israel*

Abstract

Autonomous operation of a vehicle on a road calls for understanding of various events involving the motions of the vehicles in its vicinity. In this paper we show how a moving vehicle which is carrying a camera can estimate the relative motions of nearby vehicles. We show how to "smooth" the motion of the observing vehicle, i.e. to correct the image sequence so that transient motions (primarily rotations) resulting from bumps, etc. are removed and the sequence corresponds more closely to the sequence that would have been collected if the motion had been smooth. We also show how to detect the motions of nearby vehicles relative to the observing vehicle. We present results for several road image sequences which demonstrate the effectiveness of our approach. © 2002 Pattern Recognition Society. Published by Elsevier Science Ltd. All rights reserved.

Keywords: Traffic; Vehicle motion; Image stabilization; Darboux motion; Rate of approach

1. Introduction

Autonomous operation of a vehicle on a road calls for understanding of various events involving the motions of the vehicles in its vicinity. In normal traffic flow, most of the vehicles on a road move in the same direction without major changes in their distances and relative speeds. When a nearby vehicle deviates from this norm (e.g. when it passes or changes lanes), or when it is on a collision course, some action may need to be taken. In this paper we show how a vehicle carrying a camera can estimate the relative motions of nearby vehicles.

Understanding the relative motions of vehicles requires modeling both the motion of the observing vehicle and the motions of the other vehicles. In Ref. [1] we showed that the motions of vehicles could be represented using a *Darboux motion* model that corresponds to the motion of an object moving along a smooth curve that lies on a smooth surface. We showed that deviations from Darboux motion correspond primarily to small, rapid rotations around the roll and pitch axes of the vehicle. These rotations arise from the vehicle's suspension elements in response to unevenness of the road. We derived estimated bounds on both the smooth rotations due to Darboux motion (from highway design principles) and the non-smooth rotations due to the suspension, and showed that both types of rotational motion, as well as the non-smooth translational component of the motion (bounce), are small relative to the smooth (Darboux) translational motion of the vehicle.

By our analysis, both the rotational and translational velocity components of the observing vehicle are important. On the other hand, the rotational velocity components of an observed vehicle are negligible compared to its translational velocity. As a consequence we need to estimate the rotational velocity components only for the observing vehicle. This is the case even when an observed vehicle is changing its direction of motion relative to the observing vehicle (turning or changing lanes); the

* Corresponding author. Tel.: +1-301-405-4526; fax: +1-301-314-9115.
E-mail address: ar@cfar.umd.edu (A. Rosenfeld).

turn shows up as a gradual change in the direction of the relative translational velocity.

An important consequence of the Darboux motion model is that for a fixed forward-looking camera mounted on the observing vehicle the direction of translation (and therefore the position of the focus of expansion, FOE) remains the same in the images obtained by the camera. We use this fact to estimate the observing vehicle's rotational velocity components; this is done by finding the rotational flow which, when subtracted from the observed flow, leaves a radial flow pattern (radiating from the FOE) of minimal magnitude.

We describe the motion field using full perspective projection, estimate its rotational components, and derotate the field. The flow fields of nearby vehicles are then, under the Darboux motion model, pure translational fields. We analyze the motions of the other vehicles under weak perspective projection, and derive their motion parameters. We present results for several road image sequences obtained from cameras carried by moving vehicles. The results demonstrate the effectiveness of our approach.

In Section 2, we review research related to autonomous driving and analysis of road and traffic scenes, as well as selected references on independent motion detection and object motion understanding. In Section 3, we discuss the image motion field in images obtained by a vehicle-borne camera and describe a way to estimate the necessary derotation, as well as methods of estimating nearby vehicle motions from the normal flow field. Section 4 presents experimental results for several traffic scene sequences taken at different locations. More details about the Darboux motion model are provided in Appendix A.

2. Related work

There has been extensive research on vision tasks related to autonomous driving and traffic scenes. We will not attempt to review this literature here; we cite only a few recent references [2–5].

More relevant to this paper is work on vehicle detection and tracking by a moving observer. Early work on high-level descriptions of object/vehicle trajectories in terms of such concepts as stopping/starting, object interactions, and motion verbs has been described by Nagel et al. [6,7]. More recent work by Nagel et al. on vehicle tracking is described in Refs. [8–10]. Baker et al. studied algorithms for the detection, localization, pose estimation, and recognition of road vehicles [11–13].

Betke et al. [14] developed a real-time system for detection and tracking of multiple vehicles in a frame sequence taken on a highway from a moving vehicle. The system distinguishes between distant and passing vehicle detection. In case of a passing vehicle the recognition is performed by detecting large brightness differences over small numbers of frames. 2-D car models are used to create a gray-scale template of the detected vehicle for future tracking. Distant vehicles are detected by analyzing prominent horizontal and vertical edges. A square region bounded by such edges, which is strongly enough correlated with a vehicle template, is recognized as a vehicle. For each newly recognized vehicle a separate tracking process is allocated by a real-time operating system, which tracks the vehicle until it disappears and makes sure that no other process tracks the same vehicle. When one vehicle occludes another, one of the tracking processes terminates and the other tracks the occlusion region as a single moving object. Batavia et al. [15] describe an optical flow-based obstacle detection system for use in detecting vehicles approaching the blind spot of a car on a highway or city street. The system runs at near frame rate on PC hardware.

Aste et al. describe visual routines for real-time tracking of road lanes and detection of moving objects in those lanes [16]. Giachetti et al. discuss the use of optical flow for road navigation, including egomotion recovery and finding the relative motions of other moving vehicles [17]. They assume that the rotational motion due to the shocks can be estimated by the two-parameter motion in a 41×41 pixel window around the center of the image. Our work is similar, but we use the normal flow rather then full optical flow and we model the motion due to the shocks by three rotational parameters. Pei and Liou [18] describe methods of estimating vehicle motion, but demonstrate it using only toy vehicles. Krüger [19] describes a motion compensation model based on a planar dominant motion assumption, and uses it to detect obstacles. Stiller et al. [20] and Beauvais and Lakshmanan [21] describe methods of obstacle detection based on fusion of vision and (optical) radar.

Research on understanding object motion has almost always assumed a stationary viewpoint. Understanding object motion is based on extracting the object's motion parameters from an image sequence. Broida and Chellappa [22] proposed a framework for motion estimation of a vehicle using Kalman filtering. Weng et al. [23] assumed an object that possesses an axis of symmetry, and a constant angular momentum model which constrained the motion over a local frame subsequence to be a superposition of precession and translation. The trajectory of the center of rotation can be approximated by a vector polynomial. Changing the parameters of the model with time allows adaptation to long-term changes in the motion characteristics. Gil et al. [24] combined multiple motion estimators for vehicle tracking. Vehicle detection was performed using two features: the bounding rectangle of the moving vehicle, where the convex hull of the vehicle is computed for every frame and then translated according to the predicted motion parameters, and an updated 2-D pattern (gray-level mask) based on optimization of the correlation between the pattern and the image

using the motion parameters. These results were obtained using a stationary camera mounted above a highway under different road and illumination conditions.

In Ref. [25], Duric et al. tried to understand the motions of objects such as tools and vehicles, based on the fact that the natural axes of the object tend to remain aligned with the local trihedron defined by the object's trajectory. Based on this observation they used the Frenet–Serret motion model, and showed that knowing how the Frenet–Serret frame is changing relative to the observer gives essential information for understanding the object's motion. Our present work is a continuation of this work in a more realistic and complicated scenario, in which the camera is also moving, using the Darboux motion model.

Much work has been done on detection of independently moving objects by a moving observer; reported research on this subject includes [26–30]. However, this work has been related to detection, classification, and tracking of the motion, and has not paid much attention to motion estimation. Clarke and Zisserman in Ref. [31] addressed the problem of independently moving rigid object detection, assuming that all motions (including the camera motion) are pure translations. The idea is to track a set of feature points using correlation and movement assumptions derived from the previous frames, and, based on the feature point correspondences, determine the epipole (FOE) for the background points as the intersection of the features' image plane trajectories. The assumption is that a majority of the feature points are background points. Moving objects are found by fitting an epipole to those feature matches that are not consistent with the background epipole. The image plane extent of the moving object is defined by the smallest rectangle enclosing these features. The instability of the camera introduces strong rotational components into the relative motion; these are not dealt with in Ref. [31], but will be dealt with here.

A final topic related to this paper is the selective stabilization of image sequences obtained by a moving camera, particularly by a camera carried by a ground vehicle. We will not summarize this work in detail here; for significant references see Refs. [1,32–36].

3. Vehicle and image motion

3.1. Motion of the observing vehicle and the camera

Assume that a camera is mounted on the observing vehicle; let $\vec{d}_c$ be the position vector of the mass center of the vehicle relative to the nodal point of the camera. The orientation of the vehicle coordinate system $C\xi\eta\zeta$ relative to the camera is given by an orthogonal rotational matrix (a matrix of direction cosines) which we denote by R_c. The columns of R_c are the unit vectors of the vehicle coordinate system expressed in the camera coordinate

system. We will assume that the position and orientation of the vehicle relative to the camera coordinate system do not change as the vehicle moves. Thus, we will assume that R_c and $\vec{d}_c$ are constant and known.

Given the position $\vec{p}_e$ of a scene point E in the vehicle coordinate system $C\xi\eta\zeta$, its position $\vec{r}_e$ in the camera coordinate system is given by $\vec{r}_e = R_c\vec{p}_e + \vec{d}_c$. Since R_c and $\vec{d}_c$ are constant we have $\dot{\vec{r}}_e = R_c\dot{\vec{p}}_e$. The velocity of E is given by

$$\dot{\vec{r}}_e = -\vec{\omega} \times \vec{r}_e - \vec{T}. \tag{1}$$

It was shown in Refs. [1,37] that for normal vehicle speeds v ($v > 10$ m/s ≈ 22 min/h) the rotational velocity is $\vec{\omega} = R_c(vR_{v/d}^{\mathrm{T}}\vec{\omega}_d + \vec{\omega}_{v/d})$, where $\vec{\omega}_{v/d}$ is the non-smooth rotational velocity of the vehicle, $\vec{\omega}_d$ is the smooth (Darboux) rotational velocity of the vehicle (due mainly to the smooth road changes), and $R_{v/d}$ is the rotation of the vehicle frame relative to the Darboux frame; it was also shown that the largest term (by two orders of magnitude) of the translational velocity is $\vec{T} = vR_cR_{v/d}^{\mathrm{T}}\vec{t}$, where $\vec{t}$ is the translational velocity of the Darboux frame. Furthermore, it was shown that the smooth rotational and translational velocities can be estimated by subtracting the estimated non-smooth rotational velocity $R_c\vec{\omega}_{v/d}$ from $\vec{\omega}$, producing smooth rotational and translational velocities $\vec{\omega}_s = vR_c\vec{\omega}_d$ and $\vec{T}_s = vR_c\vec{t}$.

3.2. Motions of other vehicles

We are interested in other vehicles that are moving nearby. We assume, the other vehicles are all moving in the same direction. To facilitate the derivation of the motion equations of a rigid body $\mathscr{B}$ we use two rectangular coordinate frames, one ($Oxyz$) fixed in space, the other ($C_1x_1y_1z_1$) fixed in the body and moving with it. The position of the moving frame at any instant is given by the position $\vec{d}_1$ of the origin C_1, and by the nine direction cosines of the axes of the moving frame with respect to the fixed frame. For a given position $\vec{p}$ of P in $C_1x_1y_1z_1$ we have the position $\vec{r}_p$ of P in $Oxyz$

$$\vec{r}_p = R\vec{p} + \vec{d}_1, \tag{2}$$

where R is the matrix of the direction cosines (the frames are taken as right-handed so that $\det R = 1$). The velocity $\dot{\vec{r}}_p$ of P in $Oxyz$ is given by

$$\dot{\vec{r}}_p = \vec{\omega}_1 \times (\vec{r}_p - \vec{d}_1) + \dot{\vec{d}}_1. \tag{3}$$

where $\dot{\vec{d}}_1$ is the translational velocity vector and $\vec{\omega}_1 = (\omega_x \ \omega_y \ \omega_z)^{\mathrm{T}}$ is the rotational velocity vector.

It was shown in Refs. [1,37] that for a typical vehicle $||\vec{\omega}_1|| = \mathscr{O}(0.1)$ rad/s and $||\vec{r}_p - \vec{d}_1|| = \mathscr{O}(1)$ m. We thus have $||\vec{\omega}_1 \times (\vec{r}_p - \vec{d}_1)|| \leqslant ||\vec{\omega}_1||||\vec{r}_p - \vec{d}_1|| = \mathscr{O}(0.1)\mathscr{O}(1) = \mathscr{O}(0.1)$ m/s. For the translational velocity

we have $||\vec{d}_1|| = v$; for normal speeds of the vehicle $v > 10$ m/s ≈ 36 km/hr so that $||\vec{d}_1|| = \mathcal{O}(10)$ m/s. We can conclude that for any point P on the vehicle the translational velocity is two orders of magnitude larger than the rotational velocity.

If we make the fixed frame $Oxyz$ correspond to the camera frame at time t we have from Eqs. (1) and (3) that the velocity of a point P on the vehicle expressed in the camera frame is given by

$$\dot{\vec{r}}_p = -\vec{\omega} \times \vec{r}_p + \vec{\omega}_1 \times (\vec{r}_p - \vec{d}_1) - \vec{T} + \dot{\vec{d}}_1. \tag{4}$$

In Eq. (4) the vector $-\vec{T} + \dot{\vec{d}}_1$ corresponds to the relative translational velocity between the camera and the independently moving vehicle. Regarding the first and the second terms on the r.h.s. of Eq. (4) we can see that for comparable rotational velocities $\vec{\omega}$ and $\vec{\omega}_1$ the first term will dominate the second term since usually $||\vec{r}_p - \vec{d}_1|| \ll ||\vec{r}_p||$. We will use this observation in Section 3.5.

3.3. The imaging models

Let (X, Y, Z) denote the Cartesian coordinates of a scene point with respect to the fixed camera frame (see Fig. 1), and let (x, y) denote the corresponding coordinates in the image plane. The equation of the image plane is $Z = f$, where f is the focal length of the camera. The perspective projection onto this plane is given by

$$x = \frac{fX}{Z}, \quad y = \frac{fY}{Z}. \tag{5}$$

For weak perspective projection we need a reference point (X_c, Y_c, Z_c). A scene point (X, Y, Z) is first projected

Fig. 1. The plane perspective projection image of P is $F = f(X/Z, Y/Z, 1)$; the weak perspective projection image of P is obtained through the plane perspective projection of the intermediate point $P_1 = (X, Y, Z_c)$ and is given by $G = f(X/Z_c, Y/Z_c, 1)$.

onto the point (X, Y, Z_c); then, through plane perspective projection, the point (X, Y, Z_c) is projected onto the image point (x, y). The projection equations are then given by

$$x = \frac{X}{Z_c} f, \quad y = \frac{Y}{Z_c} f. \tag{6}$$

3.4. The image motion field and the optical flow field

The instantaneous velocity of the image point (x, y) under perspective projection is obtained by taking the derivatives of Eq. (5) and using (1):

$$\dot{x} = \frac{\dot{X}Z - X\dot{Z}}{Z^2}$$
$$= \frac{-Uf + xW}{Z} + \omega_x \frac{xy}{f} - \omega_y \left(\frac{x^2}{f} + f \right) + \omega_z y, \tag{7}$$

$$\dot{y} = \frac{\dot{Y}Z - Y\dot{Z}}{Z^2}$$
$$= \frac{-Vf + yW}{Z} + \omega_x \left(\frac{y^2}{f} + f \right) - \omega_y \frac{xy}{f} - \omega_z x. \tag{8}$$

The instantaneous velocity of the image point (x, y) under weak perspective projection can be obtained by taking derivatives of Eq. (6) with respect to time and using (1):

$$\dot{x} = f \frac{\dot{X}Z_c - X\dot{Z}_c}{Z_c^2} = \frac{-Uf + xW}{Z_c} - f\omega_y \frac{Z}{Z_c} + \omega_z y, \tag{9}$$

$$\dot{y} = f \frac{\dot{Y}Z_c - Y\dot{Z}_c}{Z_c^2} = \frac{-Vf + yW}{Z_c} + f\omega_x \frac{Z}{Z_c} - \omega_z x. \tag{10}$$

Let $\vec{i}$ and $\vec{j}$ be the unit vectors in the x and y directions, respectively; $\dot{\vec{r}} = \dot{x}\vec{i} + \dot{y}\vec{j}$ is the projected motion field at the point $\vec{r} = x\vec{i} + y\vec{j}$. If we choose a unit direction vector $\vec{n}_r$ at the image point $\vec{r}$ and call it the normal direction, then the *normal motion field* at $\vec{r}$ is $\dot{\vec{r}}_n = (\dot{\vec{r}} \cdot \vec{n}_r)\vec{n}_r$. $\vec{n}_r$ can be chosen in various ways; the usual choice (as we shall now see) is the direction of the image intensity gradient.

Let $I(x, y, t)$ be the image intensity function. The time derivative of I can be written as

$$\frac{dI}{dt} = \frac{\partial I}{\partial x} \frac{dx}{dt} + \frac{\partial I}{\partial y} \frac{dy}{dt} + \frac{\partial I}{\partial t}$$

$$= (I_x \vec{i} + I_y \vec{j}) \cdot (\dot{x}\vec{i} + \dot{y}\vec{j}) + I_t = \nabla I \cdot \dot{\vec{r}} + I_t,$$

where ∇I is the image gradient and the subscripts denote partial derivatives.

If we assume $dI/dt = 0$, i.e. that the image intensity does not vary with time, then we have $\nabla I \cdot \vec{u} + I_t = 0$. The vector field $\vec{u}$ in this expression is called the *optical flow*. If we choose the normal direction $\vec{n}_r$ to be the image
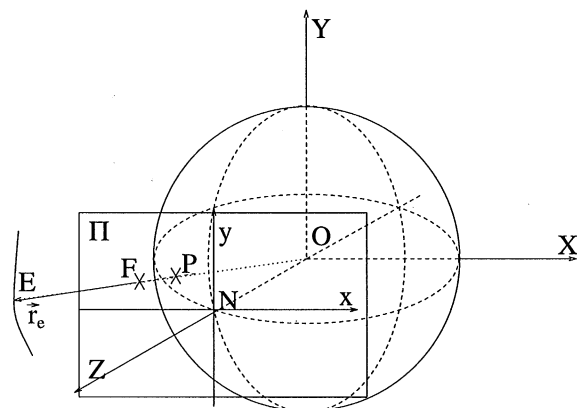
gradient direction, i.e. $\vec{n}_r \equiv \nabla I/||\nabla I||$, we then have

$$\vec{u}_n = (\vec{u} \cdot \vec{n}_r)\vec{n}_r = \frac{-I_t \nabla I}{||\nabla I||^2}, \tag{11}$$

where $\vec{u}_n$ is called the *normal flow*.

It was shown in Ref. [38] that the magnitude of the difference between $\vec{u}_n$ and the normal motion field $\vec{r}_n$ is inversely proportional to the magnitude of the image gradient. Hence, $\dot{\vec{r}}_n \approx \vec{u}_n$ when $||\nabla I||$ is large. Eq. (11) thus provides an approximate relationship between the 3-D motion and the image derivatives. We will use this approximation from now on.

3.5. Estimating camera rotation

We can now describe our algorithm for estimating the rotation of the camera (and the observing vehicle). We shall use the following notation: let I be the image intensity at $\vec{r}$, and let $\vec{n}_r = n_x \vec{i} + n_y \vec{j} = \nabla I/||\nabla I||$ be the direction of the image intensity gradient at $\vec{r}$. The *normal motion field* at $\vec{r}$ is the projection of the image motion field onto the gradient direction $\vec{n}_r$ and is given by $\dot{\vec{r}}_n = (\dot{\vec{r}} \cdot \vec{n}_r)\vec{n}_r$. From Eqs. (7) and (8) we have

$$\dot{\vec{r}}_n \cdot \vec{n}_r = n_x \dot{x} + n_y \dot{y}$$

$$= \frac{1}{Z}[n_x(-Uf + xW) + n_y(-Vf + yW)]$$

$$+ \left[n_x \frac{xy}{f} + n_y \left(\frac{y^2}{f} + f \right) \right] \omega_x - \left[n_x \left(\frac{x^2}{f} + f \right) \right.$$

$$\left. + n_y \frac{xy}{f} \right] \omega_y + (n_x y - n_y x)\omega_z. \tag{12}$$

The first term on the r.h.s. of Eq. (12) is the translational normal motion $\dot{\vec{r}}_t \cdot \vec{n}_r$ and the remaining three terms are the rotational normal motion $\dot{\vec{r}}_\omega \cdot \vec{n}_r$. From now on we will assume that the camera is forward-looking, i.e. that the *focus of expansion* (FOE) is in the image.

The *normal flow* at $\vec{r}$ is defined as $-I_t/||\nabla I||$. From Ref. [38] we know that the magnitude of the difference between the normal flow field and the normal motion field is inversely proportional to the gradient magnitude; we can thus write

$$\dot{\vec{r}}_n \cdot \vec{n}_r = \dot{\vec{r}}_t \cdot \vec{n}_r + \dot{\vec{r}}_\omega \cdot \vec{n}_r = -\frac{I_t}{||\nabla I||} + \mathcal{O}(||\nabla I||^{-1})$$

$$= \vec{u}_n \cdot \vec{n}_r + \mathcal{O}(||\nabla I||^{-1}). \tag{13}$$

If the camera motion is a pure translation the image motion field is a radial pattern; the magnitude of each image motion vector is proportional to the distance of the image point from the *focus of expansion* (FOE) and inversely proportional to the depth of the corresponding scene point. If the position $\vec{r}_0 = \vec{i}x_0 + \vec{j}y_0$ of the FOE is known the translational motion field can be obtained from the translational normal motion field by multiplying each of $\dot{\vec{r}}_t \cdot \vec{n}_r$ by a vector whose direction is $(\vec{r} - \vec{r}_0)$ and

whose magnitude is inversely proportional to the angle between the normal flow and the normal motion vector. The translational motion field is then given by

$$\dot{\vec{r}}_t = (\dot{\vec{r}}_t \cdot \vec{n}_r)\frac{\vec{r} - \vec{r}_0}{(\vec{r} - \vec{r}_0) \cdot \vec{n}_r}. \tag{14}$$

Note that if we knew $\vec{\omega}$ we could compute the rotational motion field $\dot{\vec{r}}_\omega$ and the rotational normal motion $\dot{\vec{r}}_\omega \cdot \vec{n}_r$ and use Eq. (13) to obtain

$$\dot{\vec{r}}_t \cdot \vec{n}_r \approx \vec{u} \cdot \vec{n}_r - \dot{\vec{r}}_\omega \cdot \vec{n}_r. \tag{15}$$

If we combine Eqs. (14) and (15) we have

$$\dot{\vec{r}}_t \approx (\vec{u} \cdot \vec{n}_r - \dot{\vec{r}}_\omega \cdot \vec{n}_r)\frac{(\vec{r} - \vec{r}_0)}{(\vec{r} - \vec{r}_0) \cdot \vec{n}_r}. \tag{16}$$

If the FOE is known or can be estimated (see Ref. [39]), we can use Eq. (16) to estimate the rotational velocity vector $\vec{\omega}$ by minimizing $\sum_{\vec{r}} ||\dot{\vec{r}}_t||^2$. Indeed, the image motion field in the neighborhood of the FOE is composed of the translational image motion and the rotational image motion. The roll component of the rotational image motion is orthogonal to the translational image motion, so that it increases the magnitude of the image motion field and the normal motion field. The yaw and the pitch components of the rotational image motion are approximately constant in the neighborhood of the FOE and just shift the position of the singular (zero) point of the flow field [40]. Furthermore, the rotational normal motion accounts for most of the image motion field at the distant image points [37]. Therefore, if we subtract the rotational image motion, the sum of the magnitudes of the resulting (translational) flow field will be minimal. Using Eqs. (12) and (16) we then have

$$\vec{\omega} = \arg\min_{\vec{\omega}} \sum_{\vec{r}} ||\vec{u} \cdot \vec{n}_r - \dot{\vec{r}}_\omega \cdot \vec{n}_r||^2 \frac{||(\vec{r} - \vec{r}_0)||^2}{||(\vec{r} - \vec{r}_0) \cdot \vec{n}_r||^2}. \tag{17}$$

In matrix form this problem corresponds to minimizing $||A\vec{\omega} - \vec{b}||$ (see Ref. [41]) where the rows $\vec{a}_i$ of A are given by

$$\vec{a}_i = \frac{||(\vec{r} - \vec{r}_0)||}{||(\vec{r} - \vec{r}_0) \cdot \vec{n}_r||} \left[n_x \frac{xy}{f} + n_y \left(\frac{y^2}{f} + f \right); \right.$$

$$\left. -n_x \left(\frac{x^2}{f} + f \right) - n_y \frac{xy}{f}; n_x y - n_y x \right] \tag{18}$$

and the elements b_i of $\vec{b}$ are given by

$$b_i = \vec{u} \cdot \vec{n}_r \frac{||(\vec{r} - \vec{r}_0)||}{||(\vec{r} - \vec{r}_0) \cdot \vec{n}_r||}.$$

The solution is given by

$$\vec{\omega} = A^+ \vec{b},$$

where A^+ is the generalized inverse of A (see Ref. [41]).

3.6. The motion of an observed vehicle

Finally, we show how to estimate the motion of an observed object (such as a vehicle). We assume that the FOE is known or has been estimated; that the camera rotation has been estimated, as discussed in Section 3.5; and that we have detected an observed vehicle (see Section 4.1). We then derotate the normal flow field to obtain the translational motion field (see Eqs. (14)–(16)).

From Eqs. (9) and (10) we obtain the (approximate) equations of projected relative motion for points on an observed vehicle under weak perspective:

$$\dot{x} = \frac{Uf - xW}{Z_c}, \tag{19}$$

$$\dot{y} = \frac{Vf - yW}{Z_c}. \tag{20}$$

Eqs. (19) and (20) relate the derotated image (projected) motion field to the scaled, smoothed translational velocity $Z_c^{-1}\vec{T} = Z_c^{-1}(U\ V\ W)^{\mathrm{T}}$.

Given the point $\vec{r} = x\vec{i} + y\vec{j}$ and the normal direction $n_x\vec{i} + n_y\vec{j}$, from Eqs. (19) and (20) the normal motion field $\dot{\vec{r}}_n \cdot \vec{n} = n_x\dot{x} + n_y\dot{y}$ is given by

$$\dot{\vec{r}}_n \cdot \vec{n} = n_x f U Z_c^{-1} + n_y f V Z_c^{-1} - (n_x x + n_y y) W Z_c^{-1}. \tag{21}$$

Let

$$\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \equiv \begin{pmatrix} n_x f \\ n_y f \\ -n_x x - n_y y \end{pmatrix},$$

$$\boldsymbol{c} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \equiv \begin{pmatrix} U Z_c^{-1} \\ V Z_c^{-1} \\ W Z_c^{-1} \end{pmatrix}. \tag{22}$$

Using Eq. (22) we can write Eq. (21) as $\dot{\vec{r}}_n \cdot \vec{n} = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{c}$. The column vector $\boldsymbol{a}$ is formed of observable quantities only, while each element of the column vector $\boldsymbol{c}$ contains quantities which are not directly observable from the images, but which describe the relative motion of the observed vehicle. To estimate $\boldsymbol{c}$ we need estimates of $\dot{\vec{r}}_n \cdot \vec{n}$ at three or more image points. We use linear least squares to estimate $\boldsymbol{c}$ from the normal flow.

The estimated $\boldsymbol{c}$ is our desired estimate of the motion of the observed vehicle. Note that the third component of $\boldsymbol{c}$ is the *rate of approach*. This quantity (measured in s^{-1}) is equivalent to the inverse of the *time to collision* and corresponds to the rate at which the observed vehicle is approaching the camera (or receding from it). (For example, $v = 0.1/\mathrm{s}$ means that every second the object travels 0.1 of the distance between the observer and its current position. A negative rate of approach means that the object is going away from the camera.)

4. Experiments

In Sections 4.1 and 4.2, we give examples illustrating road detection, stabilization, and vehicle detection. In Section 4.3, we present results for several sequences showing vehicles in motion.

4.1. Road and vehicle detection

We detect the road region by finding road edges and lane markers. A Canny edge detector is applied and Hough-like voting is used to detect dominant straight lines in the image. Each line is parameterized by its normal angle α and its displacement d from the center of the image. For all possible values of α and d the image is scanned along the corresponding line. The number of edge points that are found within a strip along the line, and whose gradient direction is orthogonal to the line direction, is taken as the vote for the corresponding point in the α–d plane. Among those lines, only the ones with a certain weight and orientation are chosen to be road line candidates.

If several lines with close α and d values are candidates, only the best representative of these lines is chosen. The other lines are eliminated by applying local maximum suppression in the α–d plane. Note that if the lines represent the two edges of a lane marker, the gradient directions will have opposite signs for the two edges.

Due to perspective, the road boundaries and lane marker lines should converge to a single point. Candidate lines that do not converge to a single point are not identified as a road or lane boundaries. The identified lines are the maximal subsets of candidate lines that all intersect at or close to one point (all the intersection points of every pair of the lines are located in a small region).

Our method of vehicle detection is based on the detection of intersecting vertical and horizontal lines; it is essentially the same as the method used in Ref. [14].

Fig. 2 presents some road detection and vehicle detection results for four different sequences (collected in three different countries).

4.2. Derotation

As was shown in Ref. [37] the impulsive effects introduce significant changes in the roll and pitch angular velocities (see Fig. 3). Fig. 4 shows three examples of image sequence derotation by compensating the rotational effects of road bumps. The estimated rotational normal flow component (column c) is subtracted from the total normal flow (column b), which yields the translational normal flow component (column d). We can see
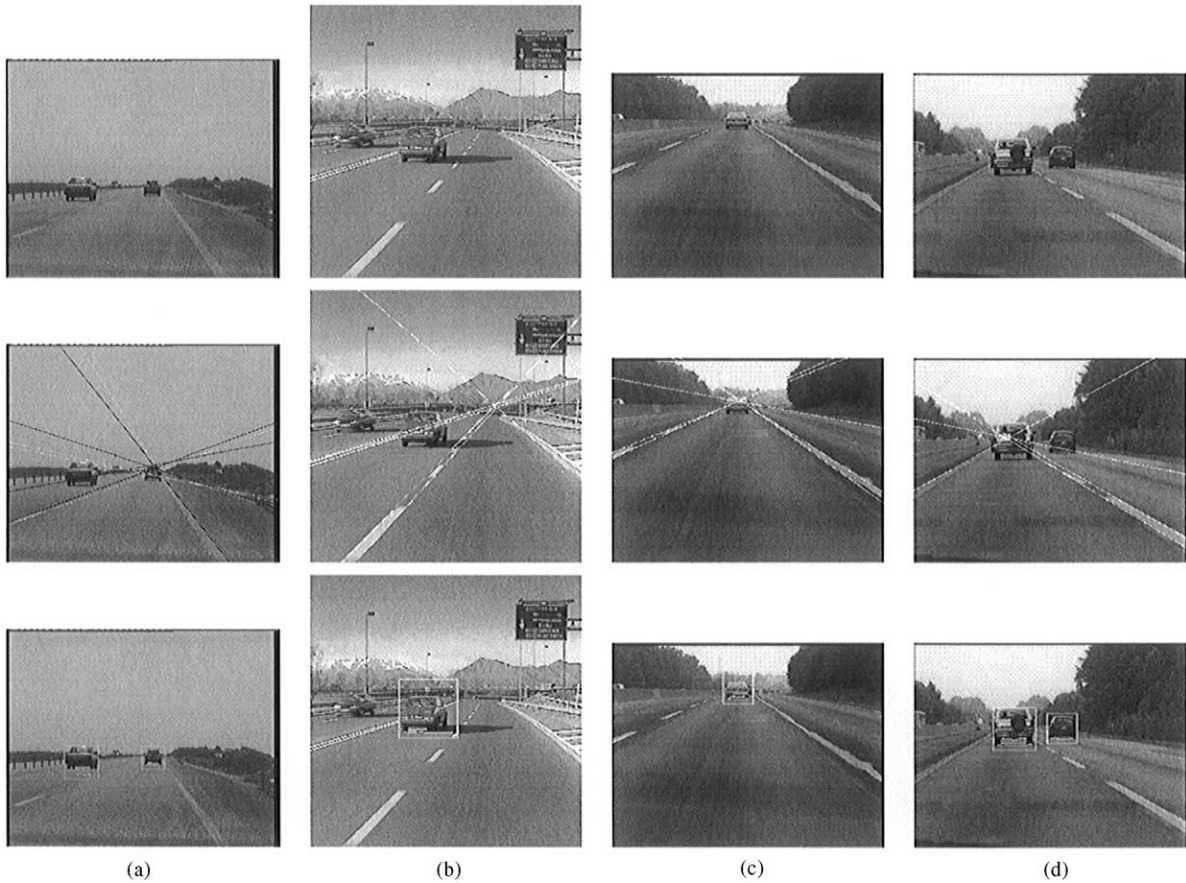
Fig. 2. (a–d) A selected frame from each of four sequences. Top: the input images. Middle: results of road detection. Bottom: results of vehicle detection.
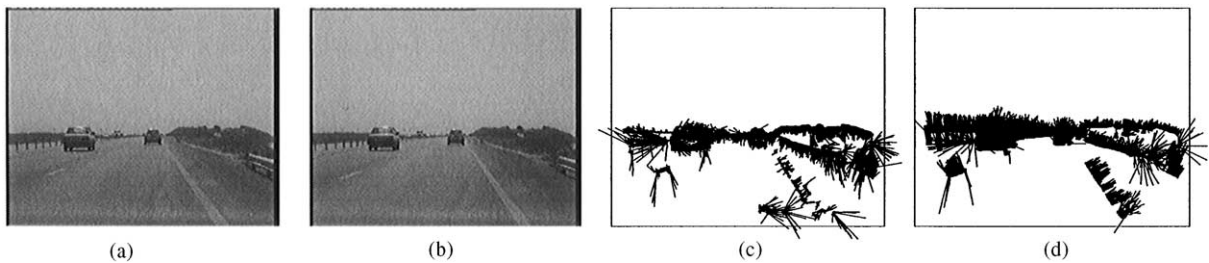

Fig. 3. (a) and (b) Two images taken 1/15th of a second apart; (c) and (d) their normal flows. One can see the effects of bumps. In the first frame, the flow vectors point downward; in the second, they point upward.

that the translational normal flow components at distant points are close to zero.

The rotation vector is estimated using the method based on FOE calculation, as described in Section 3.5; alternatively, it can be estimated from the apparent shifts of distant points, as described in Ref. [1,37]. Two examples of distant point identification, using horizon points, are shown in Fig. 5.

4.3. Relative motions of vehicles

After derotating and detecting moving vehicles, we can analyze their motions using the algorithm for motion estimation under weak perspective.

In the first experiment we used an image sequence taken in Haifa, Israel, from a vehicle following two other accelerating vehicles. The sequence consisted of
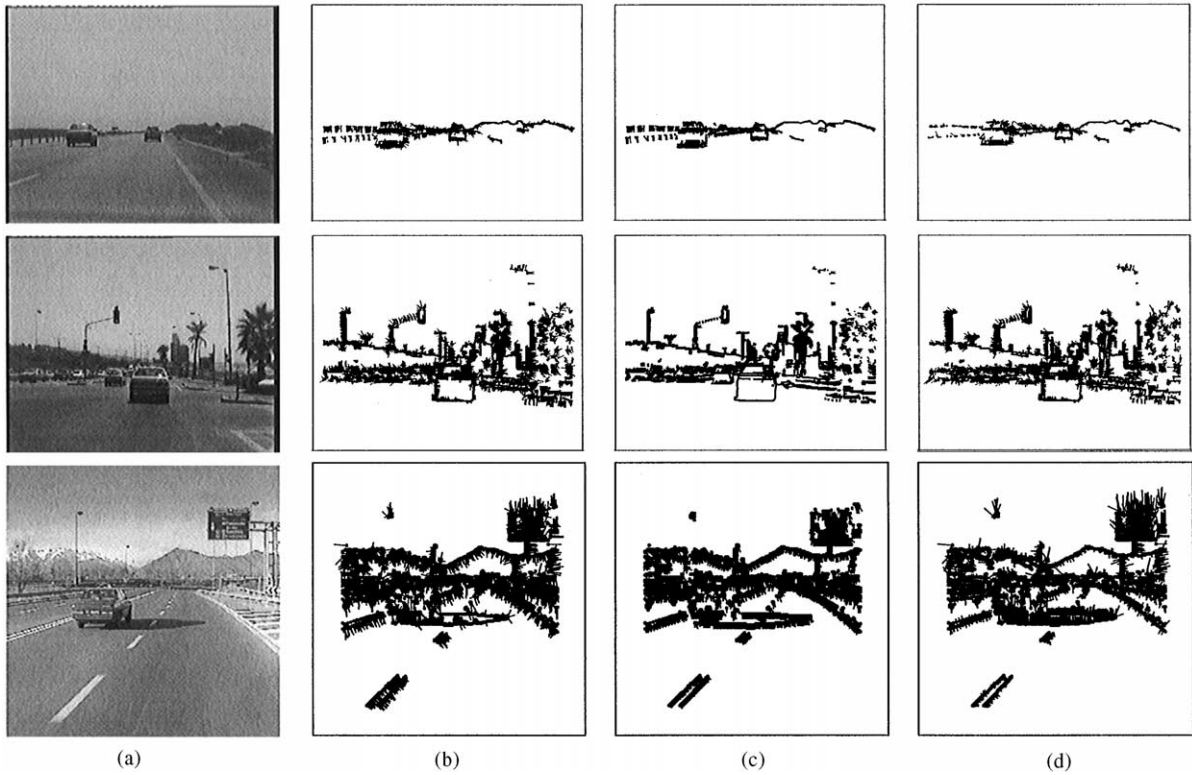
Fig. 4. Derotation results for one frame from each of three sequences: (a) input frame; (b) normal flow; (c) rotational normal flow; (d) translational normal flow.
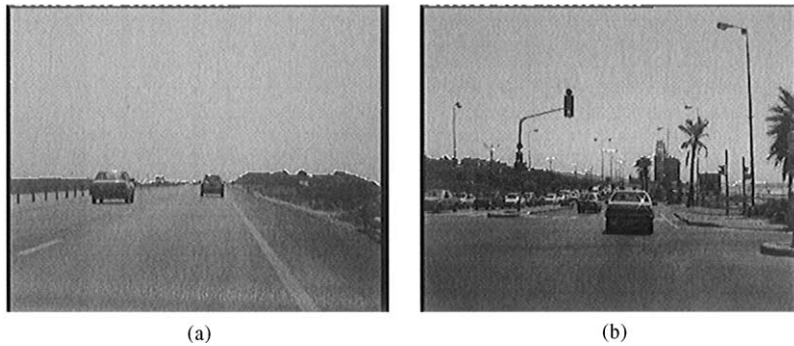

Fig. 5. Identification of distant image points in frames from two sequences.

90 frames (slightly less than 3 s). Fig. 6 shows frames 0, 30 and 60, and the corresponding normal flow on the vehicles. Fig. 7 shows estimated values of UZ_c^{-1}, VZ_c^{-1}, and WZ_c^{-1} for the central (closest) vehicle. These values correspond to the translations of the vehicles relative to the vehicle carrying the camera (i.e., in the observer coordinate system). Because of our choice of coordinate system the rate of approach v is equal to the negative of WZ_c^{-1}.

The graphs show that the motion components have a simple behavior; before they reach their extremal values they can be approximated by straight lines, indicating constant relative accelerations.

In the second experiment we used an image sequence of a van, taken in France, from another vehicle following the van [25,42]. The sequence consisted of 56 frames (slightly less than 2 s). Fig. 8 shows frames 5, 15, 25, and 35 as well as the corresponding normal flow. Fig. 9
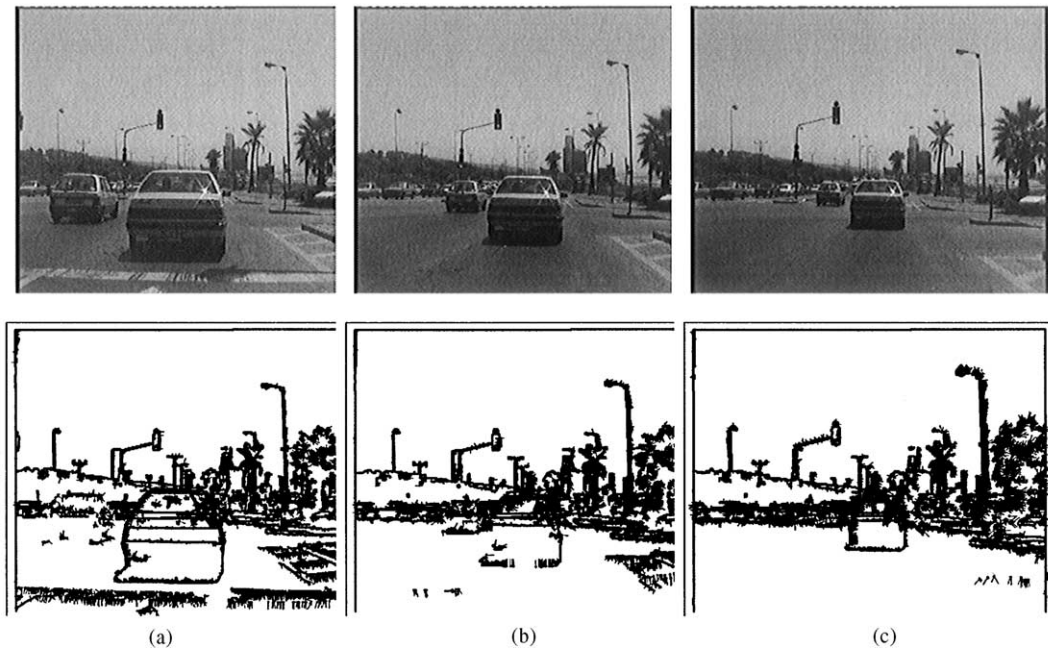
Fig. 6. Frames 0, 30 and 60 of a sequence showing two vehicles accelerating. The normal flow results are shown below the corresponding image frames.
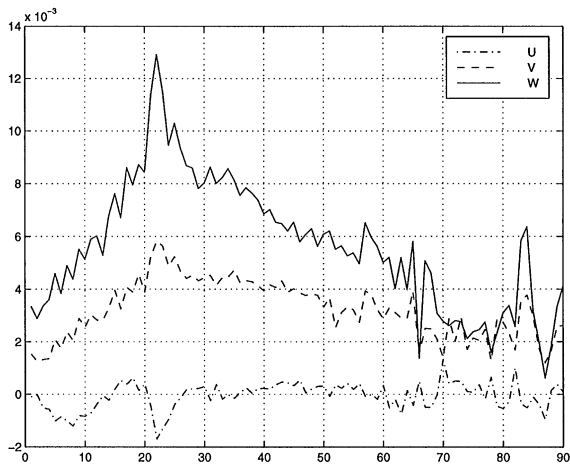

Fig. 7. Motion analysis results for the acceleration sequence. U, V, W are the scaled (by an unknown distance Z_c^{-1}) components of the relative translational velocity.

shows estimated values of UZ_c^{-1}, VZ_c^{-1}, and WZ_c^{-1}. The graph shows that there is an impending collision (rate of approach greater than $1\ \mathrm{s}^{-1}$). Around the 20th frame the rate of approach becomes zero (as do all the velocity components) and after that it becomes negative because the van starts pulling away from the vehicle carrying the camera. A similar image sequence was used in Ref. [42] in studies of vehicle convoy behavior.

The third sequence (taken from the IEN Galileo Ferrari Vision Image Library) consisted of 26 frames. Fig. 10 shows frames 1, 14 and 26, as well as the corresponding normal flow. Fig. 11 shows estimated values of UZ_c^{-1}, VZ_c^{-1}, and WZ_c^{-1}. The graph shows that the W component of the translational velocity is dominant over the U and V components, which is correct for a vehicle that overtakes the observer vehicle and does not change lanes; the two vehicles are moving on parallel courses.

Fig. 12 shows frames 1, 26 and 47 of another 48-frame sequence, taken in Haifa, Israel, as well as the corresponding normal flow. Fig. 13 shows UZ_c^{-1}, VZ_c^{-1}, and WZ_c^{-1} graphs for the left (overtaking) and central vehicles. One can see that the graphs differ mainly in the values of the W component, since the relative speed of approach for the left vehicle is greater than that for the central one. The U and V components are relatively small; all three vehicles are moving in the same direction.

5. Conclusions and plans for future work

Understanding the motions of vehicles from images taken by a moving camera requires a mathematical formulation of the relationships between the camera's motion and the image motion field, as well as a model for the other vehicles' trajectories and their contributions to
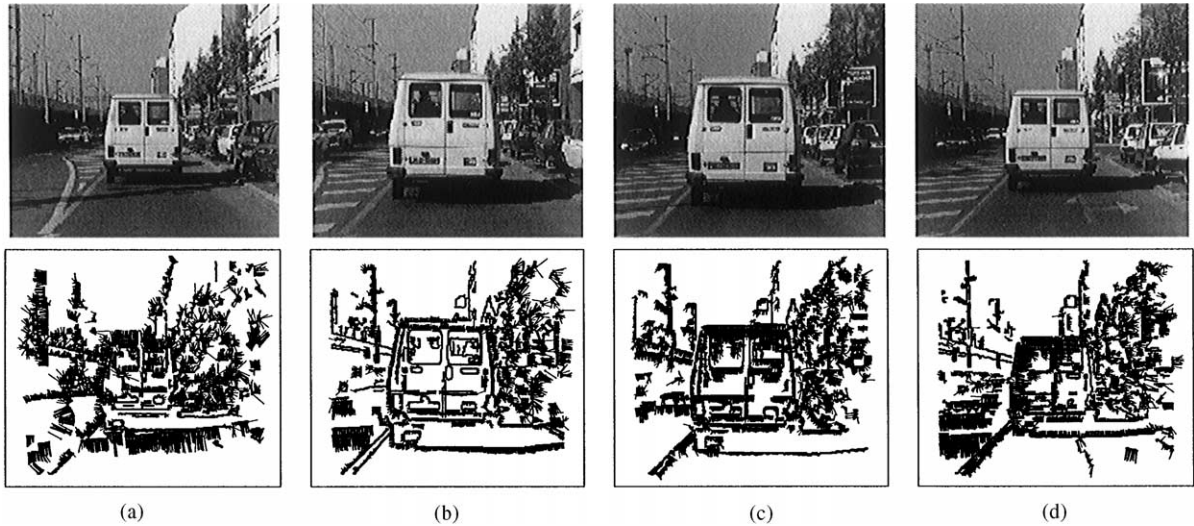
Fig. 8. Frames 5, 15, 25 and 35 of the van sequence. The normal flow results are shown below the corresponding image frames.
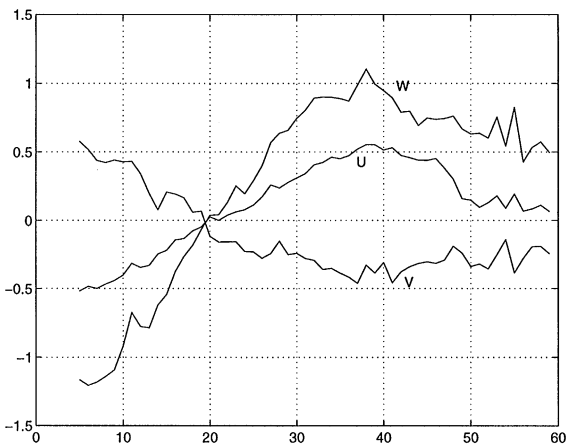

Fig. 9. Motion analysis results for the van sequence. U, V, W are the scaled (by an unknown distance Z_c^{-1}) components of the relative translational velocity.

the image motion field. The use of the Darboux frame provides a vocabulary appropriate for describing long motion sequences.

We have derived equations for understanding the relative motions of vehicles in traffic scenes from a sequence of images taken by a moving camera carried by an observing vehicle. We use the Darboux motion model for both the observing vehicle and the nearby moving vehicles. Using a full perspective imaging model we derotate the image sequence so that our model for the observed vehicles' motions can be applied. Using the weak perspective approximation we analyze the nearby vehicles' motions and apply this analysis to long image sequences.

Applying our analysis to various classes of traffic events [6] will be the subject of future research.

Acknowledgements

Appendix A. The Darboux motion model

The ideal motion of a camera moving along the ground can be described as motion along a smooth trajectory Γ lying on a smooth surface Σ. The *Darboux frame* defined by Γ and Σ has axes defined by the tangent $\vec{t}$ to Γ (and Σ), the second tangent $\vec{v}$ to Σ (orthogonal to $\vec{t}$), and the normal $\vec{s}$ to Σ (see Fig. 14). In this appendix we will present a mathematical description of motion that satisfies the Darboux frame assumption.

Consider a point O moving along a curve Γ which lies on a smooth surface Σ. There is a natural coordinate system $Otnb$ associated with Γ (even if it is a space curve), defined by the tangent $\vec{t}$, normal $\vec{n}$, and binormal $\vec{b}$ of Γ. The triple $(\vec{t}, \vec{n}, \vec{b})$ is called the *moving trihedron* or *Frenet–Serret coordinate frame*. We have the Frenet–Serret formulas [43]

$$\vec{t}\,' = \kappa \vec{n}, \quad \vec{n}\,' = -\kappa \vec{t} + \tau \vec{b}, \quad \vec{b}\,' = -\tau \vec{n}, \qquad (A.1)$$

where κ is the curvature and τ the torsion of Γ.

When the curve Γ lies on a smooth surface Σ, it is more appropriate to use the *Darboux frame* $(\vec{t}, \vec{v}, \vec{s})$ [43,44]. We take the first unit vector of the frame to be the tangent
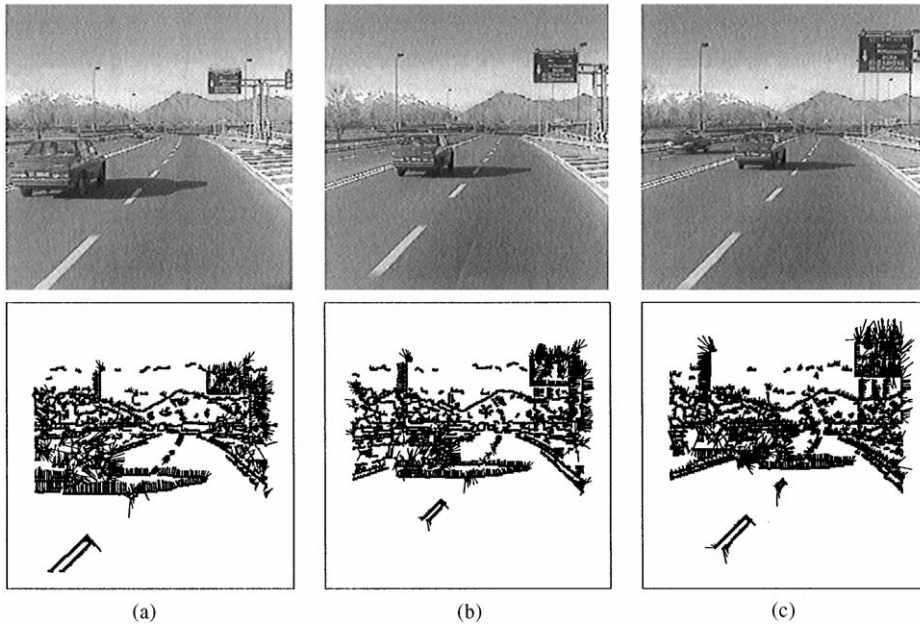
Fig. 10. Frames 1, 14 and 26 of the Italian sequence. The normal flow results are shown below the corresponding image frames.
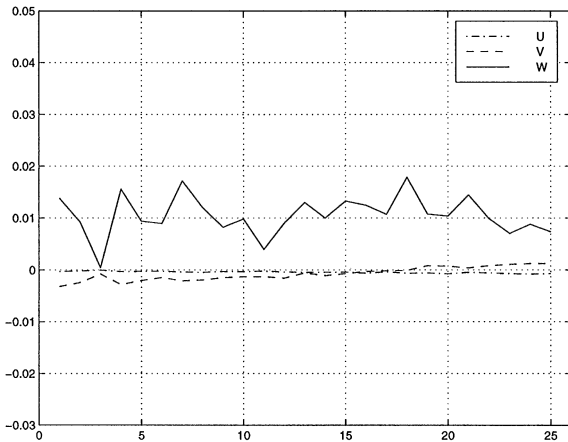

Fig. 11. Motion analysis results for the Italian sequence. U, V, W are the scaled (by an unknown distance Z_c^{-1}) components of the relative translational velocity.

$\vec{t}$ of Γ and the surface normal $\vec{s}$ to be the third frame vector; finally, we obtain the second frame vector as $\vec{v} = \vec{s} \times \vec{t}$ (see Fig. 14). Note that $\vec{t}$ and $\vec{v}$ lie in the tangent plane of Σ. Since the vector $\vec{t}$ belongs to both the $Otnb$ and $Otvs$ frames, they differ only by a rotation around $\vec{t}$, say through an angle $\psi \equiv \psi(s)$. We thus have

$$
\begin{pmatrix} \vec{v} \\ \vec{s} \end{pmatrix} = \begin{pmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{pmatrix} \begin{pmatrix} \vec{n} \\ \vec{b} \end{pmatrix}. \tag{A.2}
$$

The derivatives of $\vec{t}, \vec{v}, \vec{s}$ with respect to arc length along Γ can be found from (A.1) and (A.2):

$$
\vec{t}\,' = \kappa_g \vec{v} - \kappa_n \vec{s}, \quad \vec{v}\,' = -\kappa_g \vec{t} + \tau_g \vec{s}, \quad \vec{s}\,' = \kappa_n \vec{t} - \tau_g \vec{v}, \tag{A.3}
$$

where

$$
\kappa_g \equiv \kappa \cos\psi, \quad \kappa_n \equiv \kappa \sin\psi, \quad \tau_g \equiv \tau + \frac{d\psi}{ds},
$$

κ_g is called the *geodesic curvature*, κ_n is called the *normal curvature*, and τ_g is called the (*geodesic*) *twist*.

It is well known that the instantaneous motion of a moving frame is determined by its rotational velocity $\vec{\omega}$ and the translational velocity $\vec{T}$ of the reference point of the frame. The translational velocity $\vec{T}$ of O is just $\vec{t}$ and the rotational velocity of the $Otvs$ frame is given by the vector

$$
\vec{\omega}_d = \tau_g \vec{t} + \kappa_n \vec{v} + \kappa_g \vec{s}.
$$

Hence, the derivative of any vector in the $Otvs$ frame is given by the vector product of $\vec{\omega}_d$ and that vector. It can be seen that the rate of rotation around $\vec{t}$ is just τ_g, the rate of rotation around $\vec{v}$ is just κ_n, and the rate of rotation around $\vec{s}$ is just κ_g.

If, instead of using the arc length s as a parameter, the time t is used, the rotational velocity $\vec{\omega}_d$ and translational velocity $\vec{T}$ are scaled by the speed $v(t) = ds/dt$ of O along Γ. This speed and the three components of the rotational velocity of the Darboux frame define a rigid motion model which we call *smooth surface motion*.

We can use two coordinate frames to describe the motion of the platform carrying the camera. The "real"
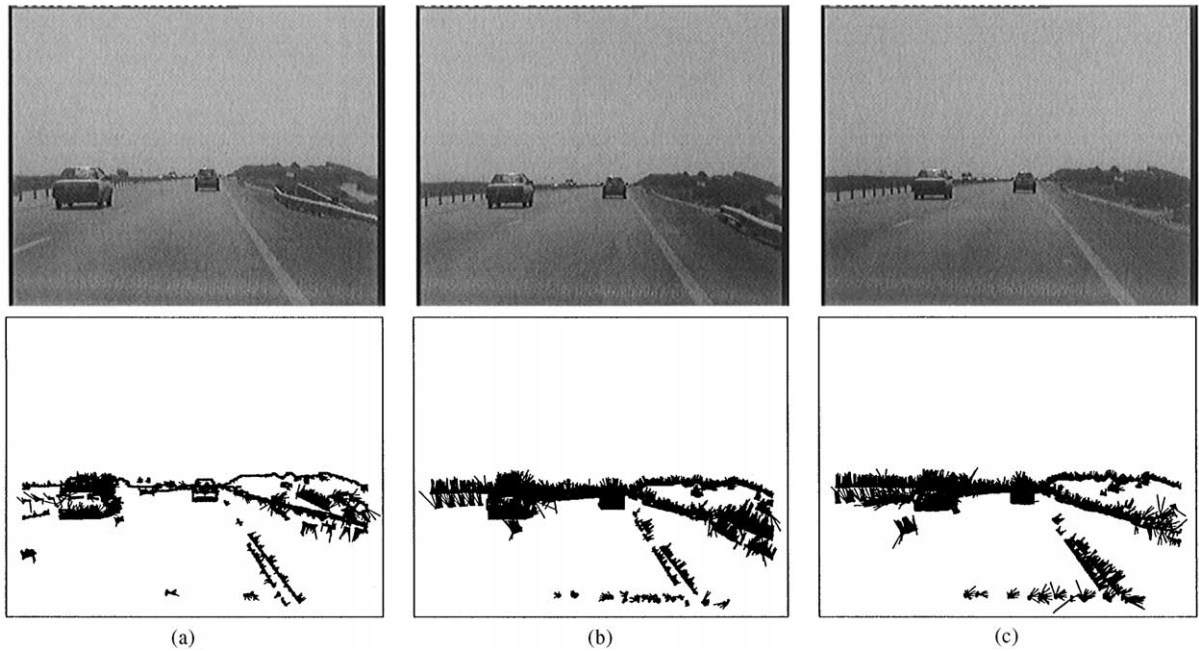
Fig. 12. Frames 1, 25 and 48 of the Haifa sequence. The normal flow results are shown below the corresponding image frames.
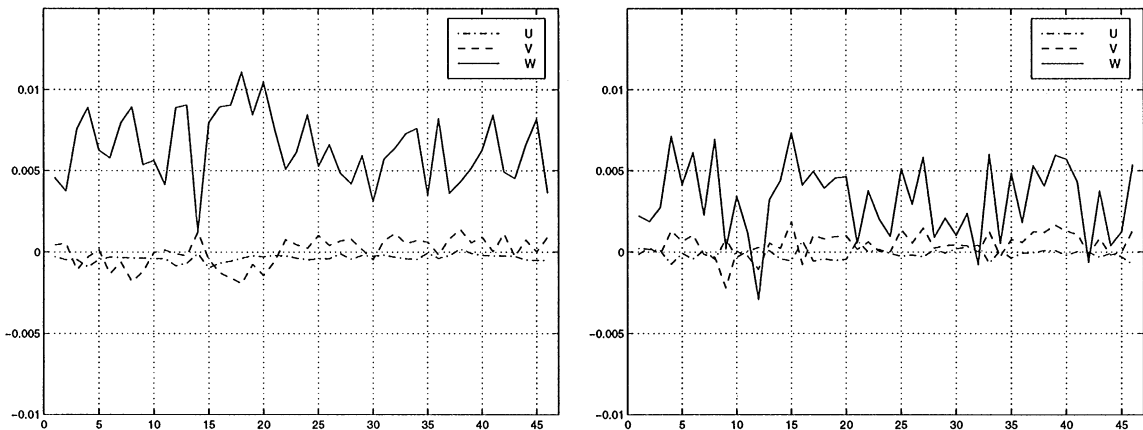

Fig. 13. Motion analysis results for the Haifa sequence. U, V, W are the scaled (by an unknown distance Z_c^{-1}) components of the relative translational velocity.

platform frame $C\xi\eta\zeta$ (which moves non-smoothly, in general) is defined by its origin C, which is at the center of mass of the platform, and its axes: $C\xi$ (fore/aft), $C\eta$ (crosswise), and $C\zeta$ (up/down); and the ideal platform frame $Otvs$ (the Darboux frame) corresponds to the smooth motion of the platform.

The motion of the platform can be decomposed into the motion of the $Otvs$ frame and the motion of the $C\xi\eta\zeta$ frame relative to the $Otvs$ frame. As we have just seen, the rotational velocity of the $Otvs$ (Darboux) frame is

$v\vec{\omega}_d = v(\tau_g\vec{t} + \kappa_n\vec{v} + \kappa_g\vec{s})$ and its translational velocity is $v\vec{t}$. We denote the rotational velocity of the $C\xi\eta\zeta$ frame by $\vec{\omega}_v$ and its translational velocity by $\vec{T}_v$.

The position of the $C\xi\eta\zeta$ frame relative to the $Otvs$ frame is given by the displacement vector $\vec{d}_{v/d}$ between C and O, and the relative orientation of the frames is given by an orthogonal rotational matrix (matrix of direction cosines) which we denote by $R_{v/d}$. The translational velocity of the platform (the velocity of C) is the sum of three terms: (i) the translational velocity of the Darboux
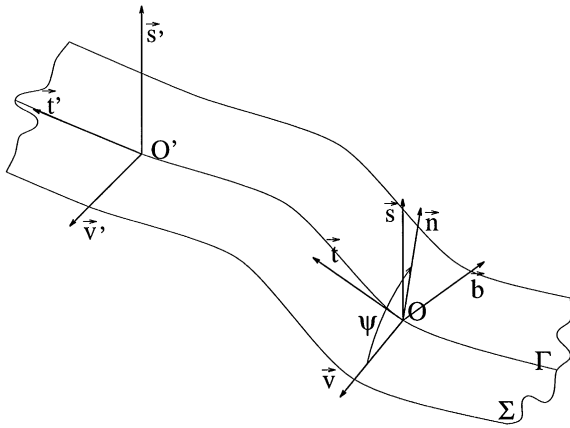
Fig. 14. The Darboux frame moves along the path Γ which lies on the surface Σ.

frame $v\vec{t}$, (ii) the translational velocity $\vec{T}_{v/d} \equiv \dot{\vec{d}}_{v/d}$, and (iii) the displacement $v\vec{\omega}_d \times \vec{d}_{v/d}$ due to rotation of C in the $Otvs$ frame. The translational velocity of the platform expressed in the $Otvs$ frame is thus $v\vec{\omega}_d \times \vec{d}_{v/d} + v\vec{t} + \dot{\vec{d}}_{v/d}$; its translational velocity in the $C\xi\eta\zeta$ frame is

$$\vec{T}v = R_{v/d}^{\mathrm{T}}(v\vec{\omega}_d \times \vec{d}_{v/d} + v\vec{t} + \dot{\vec{d}}_{v/d}). \tag{A.4}$$

Similarly, the rotational velocity of $C\xi\eta\zeta$ is the sum of two terms: (i) the rotational velocity $vR_{v/d}^{\mathrm{T}}\vec{\omega}_d$ of the $Otvs$ frame, and (ii) the rotational velocity $\vec{\omega}_{v/d}$, which corresponds to the skew matrix $\Omega_{v/d} = R_{v/d}^{\mathrm{T}}\dot{R}_{v/d}$. The rotational velocity of the $C\xi\eta\zeta$ frame expressed in the $Otvs$ frame is thus $v\vec{\omega}_d + R_{v/d}\vec{\omega}_{v/d}$; the corresponding expression in the $C\xi\eta\zeta$ frame is

$$\vec{\omega}_v = vR_{v/d}^{\mathrm{T}}\vec{\omega}_d + \vec{\omega}_{v/d}. \tag{A.5}$$

The rotational axis $\vec{c}_{v/d}$ which corresponds to the rotational matrix $R_{v/d}$ is specified by its three direction cosines c_x, c_y, c_z. The rotation around this axis through an angle δ is then given by the matrix

$$R_{v/d} = \cos\delta\ I + (1 - \cos\delta)\begin{pmatrix} c_x^2 & c_xc_y & c_xc_z \\ c_yc_x & c_y^2 & c_yc_z \\ c_zc_x & c_zc_y & c_z^2 \end{pmatrix}$$

$$+ \sin\delta \begin{pmatrix} 0 & -c_z & c_y \\ c_z & 0 & -c_x \\ -c_y & c_x & 0 \end{pmatrix}, \tag{A.6}$$

where I is the identity matrix. When δ is small (in Ref. [37] it is shown that $\delta = \mathcal{O}(0.05\,\mathrm{rad})$) we have $\cos\delta \approx 1$, $\sin\delta \approx \delta$, and

$$R_{v/d}^{\mathrm{T}} = I - \delta C_{v/d} + \mathcal{O}(\delta^2), \tag{A.7}$$

where the skew matrix $C_{v/d}$ is the matrix factor in the last term on the r.h.s. of Eq. (A.6). From Eq. (A.5) we

thus have

$$\begin{aligned}
\vec{\omega}_v &= vR_{v/d}^{\mathrm{T}}\vec{\omega}_d + \vec{\omega}_{v/d} \\
&= v\vec{\omega}_d - \delta C_{v/d}\vec{\omega}_d + v\mathcal{O}(\delta^2)\vec{\omega}_d + \vec{\omega}_{v/d} \\
&= v\vec{\omega}_d + \vec{\omega}_{v/d} - \delta v\vec{c}_{v/d} \times \vec{\omega}_d + \mathcal{O}(v\delta^2||\vec{\omega}_d||). \tag{A.8}
\end{aligned}$$

The significant terms on the r.h.s. of Eq. (A.8) are $v\vec{\omega}_d$, which is the smooth velocity of the platform; $\vec{\omega}_{v/d}$, which is the non-smooth velocity of the platform; and the cross-product term $\delta v\vec{c}_{v/d} \times \vec{\omega}_d$, which is also non-smooth. Since δ is small and $||\vec{c}_{v/d}|| = 1$, the cross-product term is small compared to $v\vec{\omega}_d$; indeed, it is $\mathcal{O}(v\delta||\vec{\omega}_d||)$. Hence,

$$\vec{\omega}_v = v\vec{\omega}_d + \vec{\omega}_{v/d} + \mathcal{O}(v\delta||\vec{\omega}_d||). \tag{A.9}$$

Rotations around the fore/aft, sideways, and up/down axes of a platform are called *roll*, *pitch*, and *yaw*, respectively. In terms of our choice of the platform coordinate system, these are rotations around the ξ-, η-, and ζ-axes.

References

[1] Z. Duric, A. Rosenfeld, Image sequence stabilization in real time, Real-Time Imag. 2 (1996) 271–284.

[2] J.-C. Hsu, W.-L. Chen, R.-H. Lin, C. Yeh, Estimations of previewed road curvatures and vehicular motion by a vision-based data fusion scheme, Mach. Vision Appl. 9 (1997) 179–192.

[3] W. Kasprzak, H. Niemann, Adaptive road recognition and ego-state tracking in the presence of obstacles, Int. J. Comput. Vision 28 (1998) 5–26.

[4] L. Dorst, Analyzing the behaviors of a car: a study in abstraction of goal-directed motions, IEEE Trans. Systems, Man, Cybernet.—Part A: Systems and Humans 28 (1998) 811–822.

[5] F. Paetzold, U. Franke, Road recognition in urban environment, Image Vision Comput. 18 (2000) 377–387.

[6] D. Koller, H. Heinze, H.-H. Nagel, Algorithmic characterization of vehicle trajectories from image sequences by motion verbs, in: Proceedings of the Conference on Computer Vision and Pattern Recognition, 1991, pp. 90–95.

[7] H. Kollnig, H.-H. Nagel, M. Otte, Association of motion verbs with vehicle movements extracted from dense optical flow fields, in: Proceedings of the European Conference on Computer Vision, Vol. 2, 1992, pp. 338–347.

[8] M. Haag, T. Frank, H. Kollnig, H.-H. Nagel, Influence of an explicitly modelled 3D scene on the tracking of partially occluded vehicles, Comput. Vision Image Understanding 65 (1997) 206–225.

[9] H. Kollnig, H.-H. Nagel, 3D pose estimation by directly matching polyhedral models to gray value gradients, Int. J. Comput. Vision 23 (1997) 283–302.

[10] M. Haag, H.-H. Nagel, Tracking of complex driving manoeuvres in traffic sequences, Image Vision Comput. 16 (1998) 517–527.

[11] G.D. Sullivan, K.D. Baker, A.D. Worrall, C.I. Attwood, P.M. Remagnino, Model-based vehicle detection and classification using orthographic approximations, Image Vision Comput. 15 (1997) 649–654.

[12] T.N. Tan, K.D. Baker, G.D. Sullivan, Model-independent recovery of object orientations, IEEE Trans. Robot. Automat. 13 (1997) 602–606.

[13] T.N. Tan, G.D. Sullivan, K.D. Baker, Model-based localization and recognition of road vehicles, Int. J. Comput. Vision 27 (1998) 5–25.

[14] M. Betke, E. Haritaoglu, L.S. Davis, Multiple vehicle detection and tracking in hard real time, Technical Report CS-TR-3667, Center for Automation Research, University of Maryland, College Park, 1996.

[15] P.H. Batavia, D.A. Pomerleau, C.E. Thorpe, Detecting overtaking vehicles with implicit optical flow, Technical Report CMU-RI-TR-97-28, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, 1998.

[16] M. Aste, M. Rossi, R. Cattoni, B. Caprile, Visual routines for real-time monitoring of vehicle behavior, Mach. Vision Appl. 11 (1998) 16–23.

[17] A. Giachetti, M. Campani, V. Torre, The use of optical flow for road navigation, IEEE Trans. Robot. Automat. 14 (1998) 34–48.

[18] S.-C. Pei, L.-G. Liou, Vehicle-type motion estimation by the fusion of image point and line features, Pattern Recognition 31 (1998) 333–344.

[19] W. Krüger, Robust real-time ground plane motion compensation from a moving vehicle, Mach. Vision Appl. 11 (1999) 203–212.

[20] C. Stiller, J. Hipp, C. Rössig, A. Ewald, Multisensor obstacle detection and tracking, Image Vision Comput. 18 (2000) 389–396.

[21] M. Beauvais, S. Lakshmanan, CLARK: A heterogeneous sensor fusion method for finding lanes and obstacles, Image Vision Comput. 18 (2000) 397–413.

[22] T.J. Broida, R. Chellappa, Estimation of object motion parameters from noisy images, IEEE Trans. Pattern Anal. Mach. Intell. 8 (1986) 90–99.

[23] J. Weng, T. Huang, N. Ahuja, 3-D motion estimation, understanding, and prediction from noisy image sequences, IEEE Trans. Pattern Anal. Mach. Intell. 9 (1987) 370–389.

[24] S. Gil, R. Milanese, T. Pun, Combining multiple motion estimates for vehicle tracking, in: Proceedings of the European Conference on Computer Vision, Vol. 2, 1996, pp. 307–320.

[25] Z. Duric, E. Rivlin, A. Rosenfeld, Understanding object motion, in: Proceedings of the International Conference on Computer Vision, 1998, pp. 925–932.

[26] R. Nelson, Qualitative detection of motion by a moving observer, Int. J. Comput. Vision 7 (1991) 33–46.

[27] M. Irani, B. Rousso, S. Peleg, Detecting and tracking multiple moving objects using temporal integration, in: Proceedings of the European Conference on Computer Vision, 1992, pp. 282–287.

[28] P. Torr, D. Murray, Statistical detection of independent movement from a moving camera, Image Vision Comput. 11 (1993) 180–187.

[29] R. Sharma, Y. Aloimonos, Early detection of independent motion from active control of normal image flow patterns, IEEE Trans. Systems, Man, Cybernet. B 26 (1996) 42–52.

[30] S. Fejes, L.S. Davis, Detection of independent motion using directional motion estimation, in: Proceedings of the International Conference on Computer Vision, 1998, pp. 979–986.

[31] J.C. Clarke, A. Zisserman, Detection and tracking of independent motion, Image Vision Comput. 14 (1996) 565–572.

[32] P.J. Burt, P. Anandan, Image stabilization by registration to a reference mosaic, in: Proceedings of the ARPA Image Understanding Workshop, 1994, pp. 425–434.

[33] M. Hansen, P. Anandan, K. Dana, G. van der Wal, P.J. Burt, Real-time scene stabilization and mosaic construction, in: Proceedings of the ARPA Image Understanding Workshop, 1994, pp. 457–465.

[34] M. Irani, B. Rousso, S. Peleg, Recovery of ego-motion using image stabilization, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1994, pp. 454–460.

[35] C. Morimoto, R. Chellappa, Fast 3D stabilization and mosaic construction, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1997, pp. 660–665.

[36] Y.S. Yao, R. Chellappa, Selective stabilization of images acquired by unmanned ground vehicles, IEEE Trans. Robot. Automat. 13 (1997) 693–708.

[37] Z. Duric, A. Rosenfeld, Stabilization of image sequences, Technical Report CAR-TR-778, Center for Automation Research, University of Maryland, College Park, 1995.

[38] A. Verri, T. Poggio, Against quantitative optical flow, in: Proceedings of the International Conference on Computer Vision, 1987, pp. 171–180.

[39] Y. Aloimonos, Z. Duric, Estimating the heading direction using normal flow, Int. J. Comput. Vision 13 (1994) 33–56.

[40] Z. Duric, A. Rosenfeld, L.S. Davis, Egomotion analysis based on the Frenet–Serret motion model, Int. J. Comput. Vision 15 (1995) 105–122.

[41] G.W. Stewart, Introduction to Matrix Computations, Academic Press, New York, 1973.

[42] P. Burlina, R. Chellappa, Time-to-X: analysis of motion through temporal parameters, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1994, pp. 461–468.

[43] E. Kreyszig, Differential Geometry, University of Toronto Press, Toronto, Canada, 1959.

[44] J.J. Koenderink, Solid Shape, MIT Press, Cambridge, MA, 1990.

About the Author—ZORAN DURIC received a Ph.D. in Computer Science from the University of Maryland at College Park in 1995. From 1995 to 1997 he was an Assistant Research Scientist at the Machine Learning and Inference Laboratory at George Mason University and at the Center for Automation Research at the University of Maryland. From 1996 to 1997 he was also a

Visiting Assistant Professor at the Computer Science Department of George Mason University. He joined the faculty of George Mason University in the Fall of 1997 as an Assistant Professor of Computer Science.

About the Author—ROMAN GOLDENBERG received the B.A. degree (with honors) in computer science from the Technion, Israel Institute of Technology, Haifa in 1995. From 1994 to 1996, 1999 he was with IBM Research Lab in Haifa. Currently he is a Ph.D. student at the Computer Science Department, Technion. His research interests include video analysis, tracking, motion based recognition, PDE methods for image processing, medical imaging, etc.

About the Author—EHUD RIVLIN received the B.Sc and M.Sc degrees in computer science and the M.B.A. degree from the Hebrew University in Jerusalem, and the Ph.D. from the University of Maryland.

Currently, he is an Associate Professor in the Computer Science Department at the Technion, Israel Institute of Technology. His current research interests are in machine vision and robot navigation.

About the Author—AZRIEL ROSENFELD is a tenured Research Professor, a Distinguished University Professor, and Director of the Center for Automation Research at the University of Maryland in College Park. He also holds affiliate professorships in the Departments of Computer Science, Electrical Engineering, and Psychology. He holds a Ph.D. in mathematics from Columbia University (1957), rabbinic ordination (1952) and a Doctor of Hebrew Literature degree (1955) from Yeshiva University, and honorary Doctor of Technology degrees from Linköping University, Sweden (1980) and Oulu University, Finland (1994) and an honorary Doctor of Humane Letters degree from Yeshiva University (2000).

Dr. Rosenfeld is widely regarded as the leading researcher in the world in the field of computer image analysis. Over a period of 35 years he has made many fundamental and pioneering contributions to nearly every area of that field. He wrote the first textbook in the field (1969); was founding editor of its first journal (1972); and was co-chairman of its first international conference (1987). He has published over 30 books and over 600 book chapters and journal articles, and has directed over 50 Ph.D. dissertations. In 1985, he served as chairman of a panel appointed by the National Research Council to brief the President's Science Advisor on the subject of computer vision; he has also served (1985–1988) as a member of the Vision Committee of the National Research Council. In honor of his 65th birthday, a book entitled "Advances in Image Understanding—A Festschrift for Azriel Rosenfeld", edited by Kevin Bowyer and Narendra Ahuja, was published by IEEE Computer Society Press in 1996.

He is a Fellow of the Institute of Electrical and Electronics Engineers (1971), won its Emanuel Piore Award in 1985, and received its Third Millenium Medal in 2000; he is a founding Fellow of the American Association for Artificial Intelligence (1990) and of the Association for Computing Machinery (1993); he is a Fellow of the Washington Academy of Sciences (1988), and won its Mathematics and Computer Science Award in 1988; he was a founding Director of the Machine Vision Association of the Society of Manufacturing Engineers (1985–1988), won its President's Award in 1987 and is a certified Manufacturing Engineer (1988); he was a founding member of the IEEE Computer Society's Technical Committee on Pattern Analysis and Machine Intelligence (1965), served as its Chairman (1985–1987), and received the Society's Meritorious Service Award in 1986, its Harry Goode Memorial Award in 1995, and became a Golden Core member of the Society in 1996; he received the IEEE Systems, Man, and Cybernetics Society's Norbert Wiener Award in 1995; he received an IEEE Standards Medallion in 1990, and the Electronic Imaging International Imager of the Year Award in 1991; he was a founding member of the Governing Board of the International Association for Pattern Recognition (1978–1985), served as its President (1980–1982), won its first K.S. Fu Award in 1988, and became one of its founding Fellows in 1994; he received the Information Science Award from the Association for Intelligent Machinery in 1998; he was a Foreign Member of the Academy of Science of the German Democratic Republic (1988–1992), and is a Corresponding Member of the National Academy of Engineering of Mexico (1982).