research. The kinematics of a multi-module vehicle was investigated using a numerical approach to solve the combined holonomic-nonholonomic problem. Simulations indicated that significant change in axle lengths can be present on uneven ground. In the absence of a variable-length axle (VLA), this translates to increased slipping which results in motion that is difficult to predict.

VSMC's with (unactuated or actuated) VLA's have the potential to improve navigation based on dead-reckoning in the presence of bumps and terrain variations. For example, it has been noted in the literature [8] that unpredictable slipping significantly increases the burden on sensor-based navigation even on relatively easy uneven ground. In the presence of decreased wheel-ground slip, along with the additional information of the axle-length variation, it is expected that more accurate dead-reckoning can be achieved as compared to the case of vehicles with fixed length axles.

This article was directed toward kinematics of vehicle systems on uneven ground. Ongoing and future research involves investigating enhanced dead-reckoning using VLA's. The force distribution, dynamics and traction capabilities of wheeled vehicles with active or passive VLA's are also being studied.

ACKNOWLEDGMENT

The authors would like to thank the reviewers for their constructive comments.

REFERENCES

[1] J. C. Alexander and J. H. Maddocks, "On the kinematics of wheeled mobile robots," *Int. J. Robot. Res.*, vol. 8, no. 5, pp. 15–27, 1989.
[2] F. Ben Amar, Ph. Bidaud, and F. Ben Ouezdou, "On modeling and motion planning of planetary vehicles," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 1993, pp. 1381–1386, vol. 2.
[3] D. B. Bickler, "Roving over mars," *Mech. Eng.*, vol. 120, no. 4, pp. 74–77, 1998.
[4] J. Borenstein, "Control and kinematic design of multi-degree-of-freedom mobile robots with compliant linkage," *IEEE Trans. Robot. Automat.*, vol. 11, pp. 21–35, Feb. 1995.
[5] G. B. Campion, G. Bastin, and d'Andrea-Novel, "Structural properties and classification of kinematic models of wheeled mobile robots," in *Proc. IEEE Int. Conf. Robot. Automat.*, 1993, pp. 462–469.
[6] B. J. Choi and S. V. Sreenivasan, "Motion planning of a wheeled mobile robot with slip-free motion capability on a smooth uneven surface," in *Proc. IEEE Int. Conf. Robot. Automat.*, Leuven, Belgium, May 1998.
[7] P. W. Davis, S. V. Sreenivasan, and B. J. Choi, "Kinematics of two wheels joined by a variable-length axle on uneven surface," in *Proc. ASME Design Tech. Conf., DETC97/DAC-3857,*, Sacramento, CA, 1997.
[8] E. Krotkov, M. Hebert, and R. Simmons, "Stereo perception and dead reckoning for a prototype lunar rover," *Auton. Robots*, vol. 2, no. 4, pp. 313–331, 1995.
[9] P. F. Muir and C. P. Neuman, "Kinematic modeling of wheeled mobile robots," *J. Robot. Syst.*, vol. 4, no. 2, pp. 281–329, 1987.
[10] S. V. Sreenivasan and P. Nanua, "Kinematic geometry of wheeled vehicle systems," in *Proc. 24th ASME Mech. Conf. 96-DETC-MECH-1137*, Irvine, CA, 1996.
[11] S. V. Sreenivasan and K. J. Waldron, "Displacement analysis of an actively articulated wheeled vehicle configuration with extensions to motion planning on uneven surface," *J. Mech. Design*, vol. 118, pp. 312–317, 1996.
[12] S. V. Sreenivasan and B. H. Wilcox, "Stability and traction control of an actively actuated micro-rover," *J. Robot. Syst.*, vol. 11, no. 6, pp. 487–502, 1994.
[13] X. Yun and N. Sarkar, "Unified formulation of robotic systems with holonomic and nonholonomic constraints," *IEEE Trans. Robot. Automat.*, vol. 14, pp. 640–650, Aug. 1998.
[14] B. H. Wilcox and D. B. Gennery, "A mars rover for the 1990's," *J. Brit. Interplanetary Soc.*, Oct. 1987.

LANA: A Lane Extraction Algorithm that Uses Frequency Domain Features

Chris Kreucher and Sridhar Lakshmanan

***Abstract*—This paper introduces a new algorithm called Lane-finding in ANother domAin (LANA) for detecting lane markers in images acquired from a forward-looking vehicle-mounted camera. The method is based on a novel set of frequency domain features that capture relevant information concerning the strength and orientation of spatial edges. The frequency domain features are combined with a deformable template prior, in order to detect the lane markers of interest. Experimental results that illustrate the performance of this algorithm on images with varying lighting and environmental conditions, shadowing, lane occlusion(s), solid and dashed lines, etc. are presented. LANA seems to detect lane markers remarkably well under a very large and varied collection of roadway images. A comparison is drawn between this frequency feature-based LANA algorithm and the spatial feature-based LOIS lane detection algorithm. This comparison is made from experimental, computational, and methodological standpoints.**

***Index Terms*— Bayesian estimation, denoising, discrete cosine transform, global shape model, intelligent vehicles, likelihood ratio, number of operations.**

I. INTRODUCTION

Lane detection, the process of locating lanes in an image with no prior estimate to aid the search, is an important enabling or enhancing technology in a number of intelligent vehicle applications, including lane excursion detection and warning, intelligent cruise control, lateral control, and autonomous driving. Studies such as [1]–[3] contain a detailed discussion of these applications and their overall impact on the economy, environment, and driver safety.

The first generation of lane detection systems were all edge-based. They relied on thresholding the image intensity to detect potential lane edges, followed by a perceptual grouping of the edge points to detect the lane markers of interest. Also, often times the lanes to be detected were assumed to be straight. See [4]–[6] and the references therein. The problem with thresholding the intensity is that, in many road scenes, it isn't possible to select a threshold which eliminates the detection of noise edges without also eliminating the detection of true lane edge points. Therefore, these first generation lane detection systems suffered when the images contained extraneous edges due to vehicles, on-off ramps, puddles, cracks, shadows, oil stains, and other imperfections in the road surface. The same deficiency also applied when the lanes were of low contrast, broken, occluded, or totally absent.[1]

The second generation of systems sought to overcome this problem by directly working with the image intensity array, as opposed to separately detected edge points, and using a global model of lane

Manuscript received April 20, 1998; revised January 13, 1999. This work was supported by Grants and Contracts DoD-DAAH04-95, DoD-DAAE07-96-C150, NSF-CDA9413862, and NSF-EEC9531589. This paper was recommended for publication by Associate Editor G. Hager and Editor V. Lumelsky upon evaluation of the reviewers' comments.

The authors are with the Department of Electrical and Computer Engineering, Vehicle Electronics Laboratory, University of Michigan, Dearborn, MI 48128-1491 USA (e-mail: ckreuche@umich.edu; lakshman@umich.edu).

Publisher Item Identifier S 1042-296X(99)03390-X.

[1] As would be the case when the road has no lane markers, but only pavement edges.

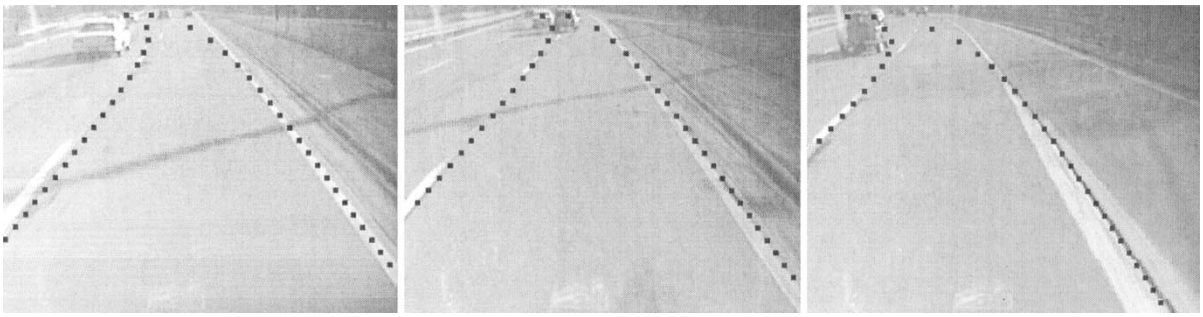
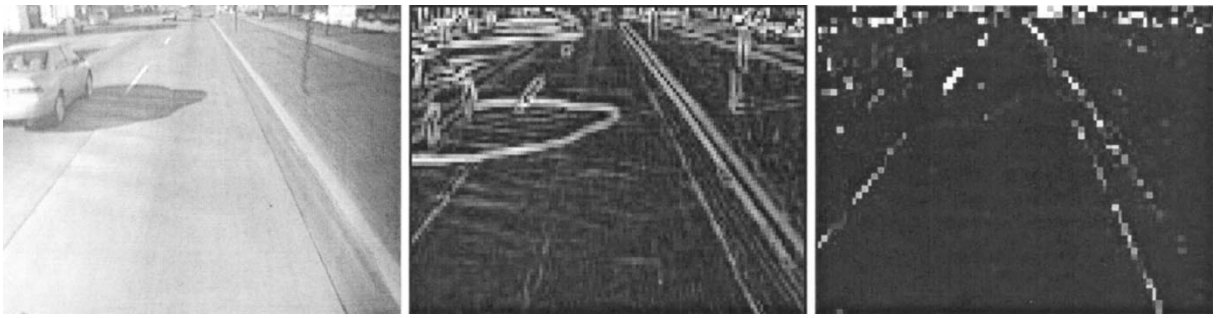


Fig. 1. Examples of the true lane shape hypothesis having a likelihood value less than a hypothesis that includes a vehicle.



(a) (b) (c)

Fig. 2. (a) A typical road image. (b) The image's gradient field. (c) The image's new feature vector field.

shape. For example, ARCADE [6] uses global road shape constraints derived from an explicit model of how the features defining a road appear in the image plane. A simple one-dimensional edge detection is followed by a least median squares technique for determining the curvature and orientation of the road. Individual lane markers are then directly determined by a segmentation of the row-averaged image intensity values. ARCADE, unlike its predecessors, does not require any perceptual grouping of the extracted edge points into individual lane edges. The RALPH system [7] is another example of a second generation lane detection system. Like ARCADE, it too uses global road shape constraints. The crux of RALPH is a matching technique that adaptively adjusts and aligns a template to the averaged scanline intensity profile in order to determine the lane's curvature and lateral offsets. The LOIS lane detector [8], yet another example of a second generation lane detection system, uses template matching as well. However, unlike RALPH, LOIS' match is over the entire image and not just an averaged scan line. At the heart of LOIS is a likelihood function that encodes the knowledge that the edges of the lane should be near intensity gradients whose orientation are perpendicular to the lane edge. This allows strong magnitude gradients to be discounted if they are improperly oriented and weak magnitude gradients to be boosted if they are properly oriented. There are several other such second generation systems; the reader is referred to [9] for a description of those. Many of these have been subject to several hours of testing, which involved the processing of extremely large and varied data sets, and it suffices to say that the second generation lane detection systems perform significantly better in comparison to the first generation ones.

However, not all of the problems associated with the first generation systems have been overcome. In particular, a number of second generation lane detection systems still have a tendency to be "distracted" or "pulled" away from the true lane markers by the presence of strong and structured edges such as those created by a vehicle outline.[2] In portions of the image whose distance from the camera is large, vehicle outlines have a much higher contrast compared to the true lane markers. In such cases, hypothesis that include the vehicle outline as part of the template are more (or at least equally) favored than those that do not include them. This "distraction problem" is illustrated in Fig. 1, which shows example images where the LOIS lane detection algorithm determines the best hypothesis to be the one that includes the vehicle outline as part of the template. The net result is that, although second generation lane detection systems provide a fairly accurate estimate of the vehicle's offset and perhaps even orientation, relative to the true lane markers, their curvature estimates are not reliable.

One way to overcome this problem, as [10], [11] point out, is to provide information about obstacles ahead of the vehicle to the lane sensing system to avoid corrupting the gradient field data used to estimate the lane shape parameters. A second way to overcome this problem is by tracking lanes from frame-to-frame given an existing model of the road geometry [23]–[25]. Prior knowledge of the road geometry permits lane tracking algorithms (unlike pure detection algorithms) to impose strong constraints on the likely location and orientation of the lanes in a new image. Yet another way to overcome the curvature estimation problem is to find image features that include the same amount of information about the true lane markers as the image intensity gradient field, which are not as sensitive to extraneous edges. The primary intent of this paper is to report the discovery of such a desirable set of image features (see Fig. 2) and to make a systematic comparison of those features to the image intensity gradient field.

[2] Lane detection algorithms get pulled away only when the distracting edges are outside of the current lane. When a vehicle is present in the current lane, the vehicle outline helps reinforce the correct lane hypothesis (see [8], [11], and [23]).

The method proceeds as follows: A given image is broken up into 8×8 pixel blocks. For each block, a frequency-domain-based feature vector is computed. This feature vector reflects the amount of "diagonally dominant edge energy" that is contained in that 8×8 block. The block feature vectors are then collectively used in combination with a deformable template shape model of the desired lane markers. This combination is accomplished in a Bayesian setting, where the deformable template model plays the role of a prior probability, and the feature vectors are used to compute a likelihood probability. The lane detection problem is reduced to finding the global maximum of a four-dimensional posterior probability density function, and an exhaustive search is employed to find the global maximum. Since the algorithm relies on frequency domain features and not the more familiar spatial domain ones, it has the name Lane-finding in ANother domAin (LANA).

LANA was applied to a widely varying set of roadway images-this set includes images obtained under a variety of lighting and environmental conditions, shadowing, lane occlusion(s), solid and dashed lines, etc. LANA was also tested on a number of images where the LOIS lane detector has problems finding the true lane markers. These results seem to indicate that lane detection in the frequency domain has some inherent advantages over detection in the spatial domain, and so LANA has much promise, in terms of being used to design a real-time lane detection and tracking system.

To gain additional insight into its performance, a systematic comparison between the frequency feature-based LANA algorithm, and the spatial feature-based LOIS lane detection algorithm was undertaken. This comparison was made from three standpoints: experimental, computational, and methodological. At this juncture, LANA seems to be better than LOIS.

The rest of this paper is organized as follows: a detailed description of the new frequency domain lane edge feature is presented in Section II; the deformable template shape model is described in Section III; Section IV encompasses the Bayesian combination of the shape model and the frequency domain features; detailed experimental results using LANA are contained in Section V; Section VI includes a systematic comparison between LANA and LOIS; and, finally, Section VII concludes the paper with some relevant remarks.

II. Frequency Domain Features of Lane Edges

There are many previously published papers that deal with frequency domain counterparts to spatial domain features [12]–[21]. Some of these papers [12]–[15] deal with the problems of texture image restoration, segmentation, and classification using frequency domain features. While others [16]–[20] use frequency domain features to extract edges and also to achieve edge-preserving image coding/compression [16]–[20]. References [16], [20], [21] are the most relevant to this paper. Especially, [21] deals with a problem similar to this paper: Curve extraction using a multidimensional Fourier transform. Curves are represented in a piecewise linear fashion and linked together via a quad tree. At any given node, the image's intensity profile is used to determine whether or not an edge is present at that node. This is accomplished by using a multiresolution Fourier transform (MFT). Large regions of the image are first examined in the MFT domain for the presence/absence of edge-like features. If an edge-like feature is deemed present in a certain region, then the region is further subdivided by using the quad tree, and a similar presence/absence decision is made at the lower nodes of the tree. This process is repeated until every pixel in the image has a classification in terms of whether or not it lies on an edge. The MFT is convenient for detecting edge features at multiresolutions and has been used to detect globally relevant edges in a variety of images. The approach in this paper has some commonality with the one in [21], especially in the use of frequency domain to detect edge-like features and the interpretation of these features' significance in a global context. However, the methods and models employed are vastly different between this paper and [21].

Lane edges are the objects of interest in this work. Recall that the features of interest are those that discriminate between lane markings and extraneous (nonlane) edges. An examination of roadway scenes obtained from a forward-looking vehicle-mounted camera easily reveals that lane markers tend to have "diagonally dominant" orientations in the image plane due to the perspective transformation inherent in the ground plane imaging process, whereas the extraneous edges have no such preferred orientations. This paper finds the frequency domain to be a convenient vehicle to discriminate between edges that are diagonally dominant and those that are randomly oriented. Details follow.

A given image is first divided into 8×8 blocks of pixels. Each of the 8×8 pixel blocks are then orthogonally decomposed in terms of a set of 64 discrete cosine transform (DCT) basis elements. Each of these elements, as seen in Fig. 3, correspond to spatial domain edges of a certain strength and orientation. Out of these 64 elements, "diagonally dominant" edges are best represented by a set of 12. The matrix in Fig. 3 indicates which 12 of the 64 they are. Fig. 4 shows several examples of the "value" of these 12 from the standpoint of lane detection. For each original image in Fig. 4, the corresponding feature image is obtained by summing the squares of its 12 special DCT decompositions. As one can see, despite the original image having features/edges of various strengths and orientations, the corresponding DCT feature images contain only information about those edges that are diagonally dominant. The rest of this paper explains how such DCT-based features can be exploited for precisely locating the lane markers.

Note that the frequency domain features adopted in this paper are similar to the ones presented in [20] and [22]. In [20], these frequency features were used for code-book optimization. Whereas in [22], the objective was to detect faces using these frequency domain features. The principal contributions of this paper are:

1) identification of specific DCT basis elements that capture lane features and nothing else;
2) use of these features in a Bayesian paradigm to robustly detect the lane markers;
3) comparison of the frequency domain and spatial domain features for lane detection.

III. Deformable Template Shape Model

As mentioned in the introduction, the algorithm presented in this paper (LANA) uses a global shape model to predict the manner in which lane markers appear in images. As commonly done [8], this paper also assumes that lane markers are circular arcs on a flat ground plane. For small-to-moderate curvatures, a circular arc with curvature k can be closely approximated by a parabola of the form

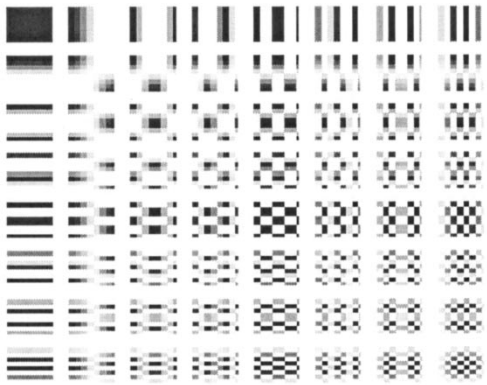
$$x = 0.5 * k * y^2 + m * y + b. \tag{1}$$

The derivation of the class of corresponding curves in the image plane is given for the case of an untilted camera, but it can be shown that the same family of curves results when the camera is tilted. Assuming perspective projection, a pixel (r, c) in the image plane projects onto the point (x, y) on the ground plane according to

$$x = c * cf * y \tag{2}$$

and

$$y = \frac{H}{r * rf} \tag{3}$$



$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(a) (b)

Fig. 3. (a) DCT basis elements. (b) Matrix that represents which 12 of the 64 capture diagonally dominant edges.

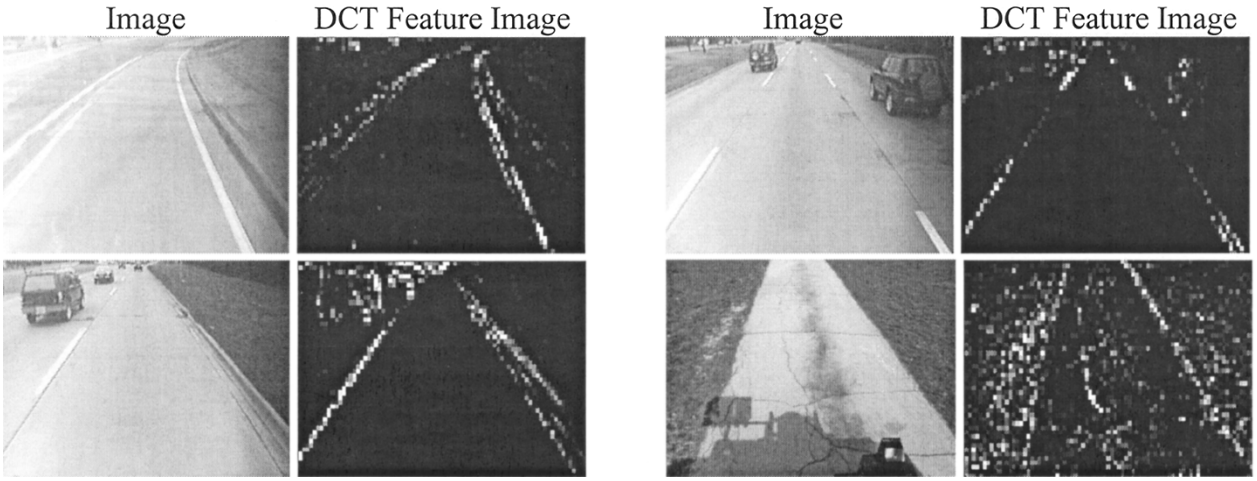


Fig. 4. DCT features of typical roadway scenery.

where H is the camera height, rf is the height of a pixel on the focal plane divided by the focal length, and cf is the width of a pixel on the focal plane divided by the focal length. Substituting (2) and (3) into (1) and performing some simple algebraic manipulation results in the image plane curve

$$c = \frac{0.5 * k * H}{rf * cf * r} + \frac{b * rf * r}{H * cf} + \frac{m}{cf} \tag{4}$$

or, combining the ground plane and camera calibration parameters together

$$c = k'/r + b' * r + vp. \tag{5}$$

In the case of a tilted camera, the same family of curves results if the image coordinate system is defined so that row 0 is the horizon row. For left and right lane edges defined by concentric arcs, the approximation is made that the arcs have equal curvature and equal tangential orientation where they intersect the X axis, so k' and vp will be equal for the left and right lane edges. While the radius of curvature and tangent orientation of the left and right lane edges will differ slightly, constraining the left and right lane edges to have the same k' and vp parameters closely approximates the actual lane edge shapes for all but very small radii of curvature. As a result, the lane shape in an image can be defined by the four parameters k', b'_{LEFT}, b'_{RIGHT}, and vp. In summary, the k' parameter is linearly proportional to the curvature of the arc on the ground plane. The vp parameter is a function of the tangential orientation of the arc on the ground plane, with some coupling to the arc curvature as well (depending on the amount of camera tilt). The b'_{LEFT} and b'_{RIGHT} parameters are functions of the offset of the arc from the camera on the ground plane, with couplings to arc curvature and tangential orientation (again, the relative contributions of these couplings depends on the camera tilt) [6]. The flat ground plane assumption is occasionally violated due to a vertical curvature in the road ahead, and also the camera height and tilt changes by small increments due to the suspension rock of the host vehicle. However, accounting for these variations is difficult and also their effect on the accuracy of the detected lanes is not very apparent.

In the next section, this deformable template shape model is used in conjunction with the DCT-based lane features to detect the precise location of the lane markers.

IV. Bayesian Lane Detection

It is assumed that the values of the lane shape parameters k', b'_{LEFT}, b'_{RIGHT}, and vp are influenced by a prior probability density function (pdf)

$$\begin{aligned} P(k', b'_{LEFT}, b'_{RIGHT}, vp) \\ \propto \left(\operatorname{atan}\alpha\left(\left(b'_{RIGHT} - b'_{LEFT}\right) - 1\right)\right. \\ \left. - \operatorname{atan}\alpha\left(\left(b'_{RIGHT} - b'_{LEFT}\right) - 3\right)\right) \times \left(1 - \beta\left(\frac{k'}{\chi}\right)^2\right) \end{aligned} \tag{6}$$

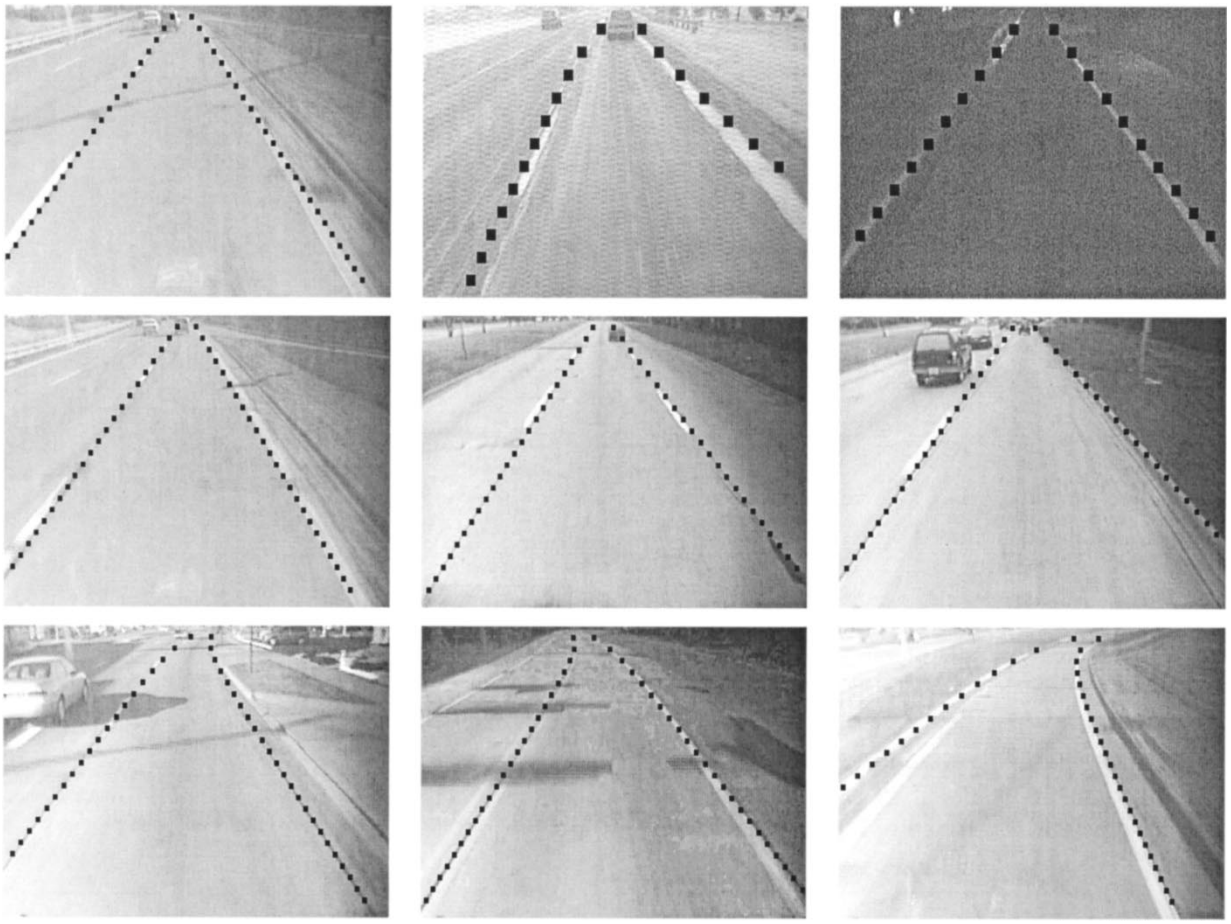


Fig. 5. Example results of processing a varied set of images using LANA.

where $\alpha = 10$, $\beta = 0.01$, and $\chi = 600$. This prior pdf embodies two types of *a priori* knowledge about roadways: First, roadways tend to have a certain range of widths, and this is enforced by the term involving atans; Second, roadways tend to have a certain range of curvatures, and this is enforced by the parabolic weighting function on k'.

It is also assumed that given the values of k', b'_{LEFT}, b'_{RIGHT}, and vp, the probability of the observed image having the DCT feature values (the ones described in Section II) is given by the likelihood pdf[3]

$$P(\text{DCT feature values}|k',\, b'_{LEFT},\, b'_{RIGHT},\, vp) \propto \sum_{i,j} \sum_{k,l \in C_{i,j}} (dct\_coeff(k,l))^2 \tag{7}$$

where the sum over (i,j) covers those 8 × 8 pixel blocks through which the left and right lanes (as dictated by k', b'_{LEFT}, b'_{RIGHT}, and vp) pass, $C_{i,j}$ denotes the set of 12 DCT basis elements that capture diagonally dominant edge features (see Section II for details), and $dct\_coeff(k,l)$ denote the (k,l)th DCT coefficient of the (i,j)th block of 8× 8 pixels. This likelihood pdf encodes the knowledge that the true lane markers lie along portions of the image that uniformly have a high amount of perspective (diagonally-dominant) edge energy.

[3]Ideally, the likelihood pdf should contain a normalizing factor that depends on the lane shape parameters. However, calculating this factor is very difficult, as it involves an integration of the RHS of (7) over all the DCT feature values. Omission of such a factor is ubiquitous to Bayesian methods in image analysis.

These two pdfs are combined using Bayes' rule, and the lane detection problem is reduced to one of finding a maximum *a posteriori* (MAP) estimate

$$\begin{aligned} &\underset{k',\, b'_{LEFT},\, b'_{RIGHT},\, vp}{\text{argmax}} \\ &\quad \cdot P(k',\, b'_{LEFT},\, b'_{RIGHT},\, vp|\text{DCT feature values}) \\ &= \underset{k',\, b'_{LEFT},\, b'_{RIGHT},\, vp}{\text{argmax}} P(k',\, b'_{LEFT},\, b'_{RIGHT},\, vp) \\ &\quad \times P(\text{DCT feature values}|k',\, b'_{LEFT},\, b'_{RIGHT},\, vp) \\ &= \underset{k',\, b'_{LEFT},\, b'_{RIGHT},\, vp}{\text{argmax}} (\operatorname{atan}\alpha((b'_{RIGHT} - b'_{LEFT}) - 1) \\ &\quad - \operatorname{atan}\alpha((b'_{RIGHT} - b'_{LEFT}) - 3) \times \left(1 - \beta\left(\frac{k'}{\chi}\right)^2\right) \\ &\quad \times \sum_{i,j} \sum_{k,l \in C_{i,j}} (dct\_coeff(k,l))^2. \end{aligned} \tag{8}$$

The MAP estimate is eventually found by a straightforward exhaustive search over the four-dimensional parameter space of k', b'_{LEFT}, b'_{RIGHT}, and vp. The next section contains several experimental results that illustrate the efficacy of this Bayesian lane detection procedure. Optimization methods such as the Metropolis algorithm [8] and constrained local searches based on recent road geometry [23]–[25] could also be used, instead of exhaustive search.

V. Experimental Results

The lane extraction procedure LANA, described in the previous sections, was applied to a varied set of images. The images include

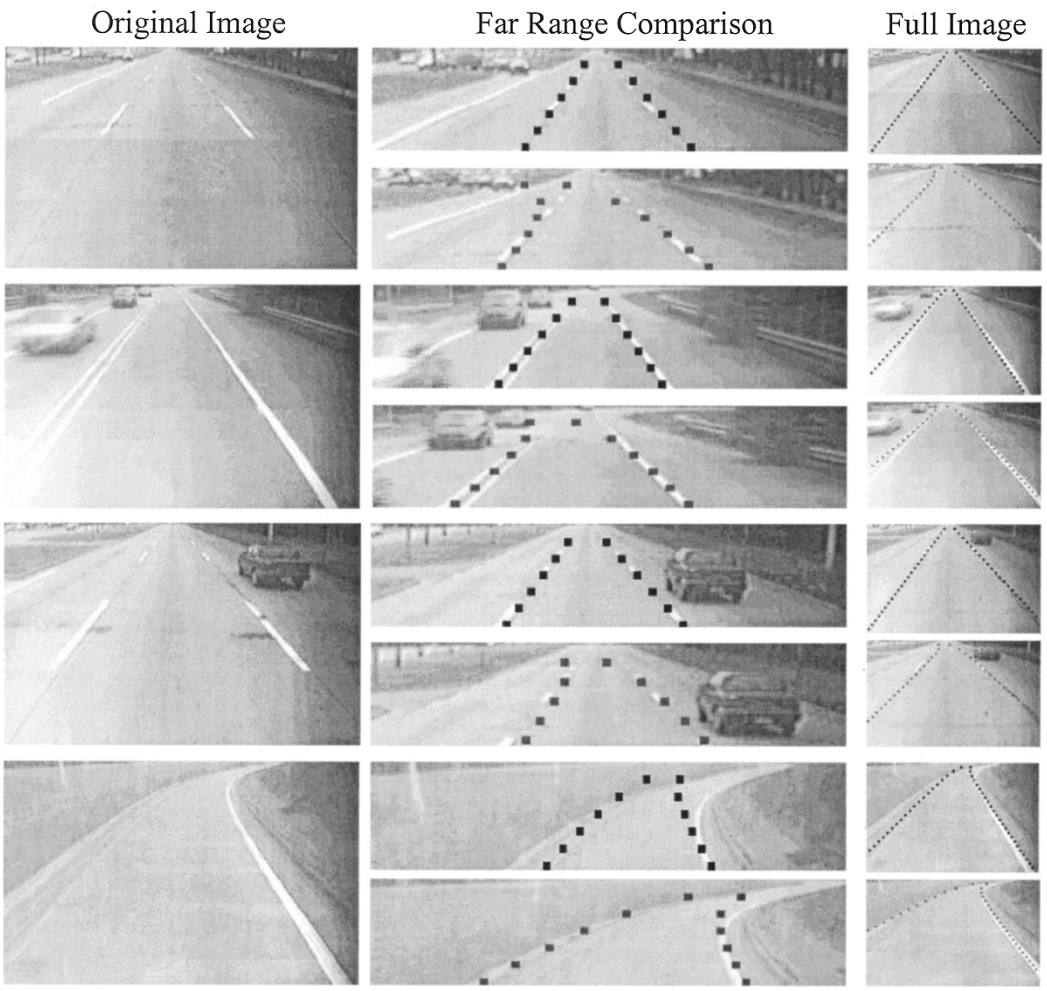


Fig. 6. Experimental comparison between LANA and LOIS. For each row, the LANA results are on top and the LOIS results are on the bottom of the middle and right columns.

those that were obtained under a variety of lighting and environmental conditions, shadowing, lane occlusion(s), solid and dashed lines, etc. (see Fig. 5).

LANA provides a parametric description of the lane on the image plane via the parameters k', b'_{LEFT}, b'_{RIGHT}, and vp. Following the discussion in Section III, these parameters can be easily transformed onto the corresponding ground plane parameters k, b_{LEFT}, b_{RIGHT}, and m.

VI. Comparison of LANA and LOIS

In this section, we compare two lane detection algorithms that use two very different features to detect the lane markers. The two algorithms are:

1) The LOIS lane detection algorithm that uses a spatial domain image intensity gradient field.
2) The LANA lane detection algorithm that uses frequency domain oriented-edge features.

This comparison between LANA and LOIS is undertaken from three standpoints:

1) experimental;
2) computational;
3) methodological.

The experimental comparison seems to indicate that LANA has some advantages over LOIS. Especially, LANA does not seem to be distracted by strong non-lane edges in far range. While this is a desirable characteristic more often than not (see the first three rows of images in Fig. 6), sometimes it prevents LANA from detecting lanes with sharp curvature correctly (see the last row of images in Fig. 6).

A computational comparison of LANA and LOIS follows. Given a hypothesized lane shape over a 640 $\times$ 480 image, the number of operations involved in the computation of the LANA and LOIS objective functions are tabulated in Table I. Both the LANA and LOIS lane detection algorithms incur a computational overhead every time a new image is processed. For LANA this overhead is the DCT computation, whereas for LOIS it is the gradient field computation. Table I also provides a comparison of the number of operations involved in computing the LANA and LOIS overheads.

The size of the images in Fig. 6 were all 640$\times$ 480. Both LANA and LOIS searched over (the same) approximately 400 000 possible lane shapes, comprised of nine different curvature values, 50 different orientations, and 30 different locations each for the left and right lanes. On a Pentium-266MHz 96MB-RAM Desktop-PC, for each image, LANA took approximately 30 s to search over the 400 000 possible lane shapes, in comparison LOIS took approximately 2 h for the same task. Potential speedups for the LANA algorithm include improved optimization techniques such as those mentioned in Section IV, and hardware implementations that allow direct acquisition of DCT feature images.

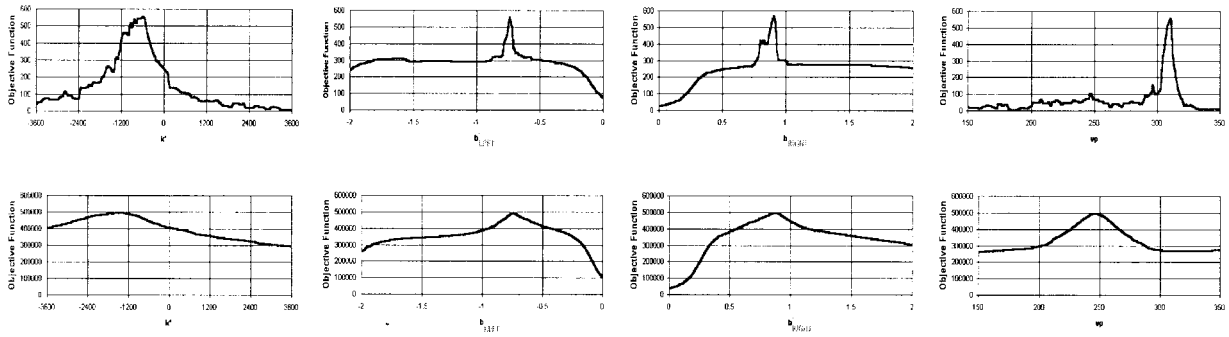


Fig. 7. Cross sections w.r.t. k', b'_{LEFT}, $b'_{RIGHT,}$ and vp. Top—for the LANA objective function. Bottom—for the LOIS objective function.

TABLE I
COMPARISON OF THE NUMBER OF OPERATIONS FOR LANA AND LOIS. (a) OBJECTIVE FUNCTION COMPUTATION AND (b) OVERHEAD COMPUTATION

Type of Operation	# of Ops for LANA	# of Ops for LOIS
Addition	600	112,000
Multiplication	240	133,000
Division	360	2,400
Math Library Call	0	800
Table Look-Up	0	66,000

(a)

Type of Operation	# of Ops for LANA	# of Ops for LOIS
Addition	4,900,000	5,300,000
Multiplication	4,900,000	614,000
Math Library Call	0	614,000

(b)

Finally, LANA and LOIS were compared from a methodological standpoint. An assessment was made as to which of them is inherently more reliable for lane detection. Shown in Fig. 7 are a set of graphs that represent the cross-sections of the LANA and LOIS objective functions along k', b'_{LEFT}, b'_{RIGHT}, and vp axes for the same image:

The graphs in the top row of Fig. 7 are more relatively peaked, with respect to the floor, for all four parameters k', b'_{LEFT}, b'_{RIGHT}, and vp, than the ones in the bottom row. Hence, between LANA and LOIS, LANA is expected to discriminate better between the globally "correct" and "incorrect" hypothesis. This was verified by a comparison of the ratio between the global maxima and the average (floor) value of the LANA and LOIS objective functions. For LANA, this ratio was found to average around 15, whereas for LOIS, the ratio was around 2.5.

VII. CONCLUSION

This paper introduced a new Bayesian algorithm called LANA (lane-finding in another domain) for detecting lane markers in images acquired from a forward-looking vehicle-mounted camera. The method was based on a novel set of frequency domain features, and it was shown to consistently detect the unknown lane markers correctly. This was true, even in situations where a more traditional (spatial domain feature-based) algorithm such as LOIS fails. A systematic assessment of LANA, and especially a comparison to LOIS, was undertaken. LANA was shown to have advantages over LOIS from three different standpoints—experimental, computational, and methodological. Future work includes a real-time realization of LANA and more extensive testing aboard an automobile.

ACKNOWLEDGMENT

The authors wish to thank M. Beauvais, Dr. K. Kluge, and Dr. V. Prasad for their helpful discussions.

REFERENCES

[1] R. J. Bishop, "Precursor systems analysis of an automated highway system," in *Proc. IEEE Veh. Technol. Conf.*, 1993, pp. 364–367.
[2] S. A. Lund, "Intelligent vehicle/highway systems in the US—In need of a vision," *SAE Trans.*, vol. 100, pp. 1258–1279, 1991.
[3] J. Tsao, R. W. Hall, and S. E. Shladover, "Design options for operating automated highway systems," in *Proc. IEEE-IEE Veh. Navigation Inform. Syst. Conf.*, 1993, pp. 494–500.
[4] S. K. Kenue, "LANELOK: Detection of lane boundaries and vehicle tracking using image-processing techniques—Parts I and II," *SPIE Mobile Robots IV*, 1989.
[5] K. C. Kluge, "YARF: An open-ended framework for robot road following," Ph.D. dissertation, Carnegie Mellon Univ., Pittsburgh, PA, 1993.
[6] ——, "Extracting road curvature and orientation from image edge points without perceptual grouping into features," in *Proc. Intell. Veh. Symp.*, 1994, pp. 109–114.
[7] D. Pomerleau and T. Jochem, "Rapidly adapting machine vision for automated vehicle steering," *IEEE Expert*, vol. 11, no. 2, pp. 19–27, 1996.
[8] S. Lakshmanan and K. Kluge, "LOIS: A real-time lane detection algorithm," in *Proc. 30th Annu. Conf. Inform. Sci. Syst.*, 1996, pp. 1007–1012.
[9] K. C. Kluge, "Performance evaluation of vision-based lane sensing: Some preliminary tools, metrics and results," in *IEEE Conf. Intell. Transport. Syst.*, 1997.
[10] S. Kenue, "Vision-based algorithms for near-host object detection and multilane sensing," in *Proc. SPIE Intell. Veh. Highway Syst. Conf.*, 1995, pp. 88–104.
[11] M. Beauvais and S. Lakshmanan, "CLARK: A heterogeneous sensor fusion method for finding lanes and obstacles," in *Proc. Intell. Veh. Symp.*, 1998.
[12] T. Iso, Y. Watanabe, and K. Shimohara, "Human face classification for security system," in *Proc. IEEE Int. Conf. Image Process.*, 1996, pp. 479–482.
[13] R. Muzzolini, Y.-H. Yang, and R. Pierson, "Texture characterization using robust statistics," *Pattern Recognit.*, vol. 27, pp. 119–134, 1994.
[14] ——, "Local frequency features for texture classification," *Pattern Recognit.*, vol. 27, pp. 1397–1406, 1994.
[15] Z. Q. Gu, C. N. Duncan, P. M. Grant, C. F. N. Cowan, E. Renshaw, and M. A. Mugglestone, "Textural and spectral features as an aid to cloud classification," *Int. J. Remote Sensing*, vol. 12, pp. 953–968, 1991.
[16] A. Kubrick and T. J. Ellis, "Perceptually based directional classified gain-shape vector quantization," *IEEE Trans. Circuits. Syst. Video Technol.*, vol. 5, pp. 96–108, 1995.
[17] P. J. Tadrous, "Simple and sensitive method for directional edge detection in noisy images," *Pattern Recognit.*, vol. 28, pp. 1575–1586, 1995.
[18] Y. J. Han, Y. Feng, and C. L. Weller, "Frequency domain image analysis for detecting stress cracks in corn kernels," *Appl. Eng. Agric.*, vol. 12, pp. 487–491, 1996.

[19] W.-L. Wang, G. Jin, Y. Yan, and M. Wu, "Image feature extraction with the optical Haar wavelet," *Opt. Eng.*, vol. 34, pp. 1238–1242, 1995.
[20] Y. S. Ho and A. Gersho, "Classified transform coding of images using vector quantization," in *IEEE Int. Conf. ASSP*, 1989, pp. 1890–1893.
[21] A. D. Calway and R. Wilson, "Curve extraction in images using a multiresolution framework," *CVGIP: Image Understanding*, vol. 59, pp. 359–366, 1994.
[22] H. Wang and S.-F. Chang, "A highly efficient system for automatic face region detection in MPEG video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, pp. 615–628, 1997.
[23] K. C. Kluge, C. Kreucher, and S. Lakshmanan, "Tracking lane and pavement edges using deformable templates," in *Proc. SPIE Intell. Veh. Highway Syst. Conf.*, 1998.
[24] J. Kosecka, R. Blasi, C. J. Taylor, and J. Malik, "Vision-based lateral control of vehicles," in *Proc. IEEE Conf. Intell. Transportation Syst.*, 1997.
[25] E. Dickmanns and B. D. Mysliwetz, "Recursive 3-D road and relative ego-state recognition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, pp. 199–213, 1992.

On the Use of Linear Camera-Object Interaction Models in Visual Servoing

Benedetto Allotta and Carlo Colombo

***Abstract*—We investigate the exploitation of linear models of camera-object interaction for an efficient modeling and control of image-based visual servoing systems. The approach includes a method for coping with those representation ambiguities typical of linear interaction models which may affect both planning and control. The implementation of an eye-in-hand servoing system based on affine camera models and using image contours as relevant visual features is described and discussed. The system, including an image planner, a two-dimensional/three-dimensional (2-D/3-D) controller, and a visual analysis module, allows an intuitive specification and execution of relative positioning tasks w.r.t. still or moving rigid objects. Results of real-time experiments with a robotic platform featuring a PUMA manipulator provide a further insight into characteristics and performance of affine visual servoing systems.**

***Index Terms*— Affine interaction models, eye-in-hand systems, visual servoing.**

I. Introduction

The use of visual sensors in the exteroceptive feedback loop of a robot system, referred to as *visual servoing*, appears to be a natural approach to developing flexible positioning strategies, with applications including robot grasping, manipulation, and navigation (conveyor belt management, part-placement, assembly, etc.) [1], [2].

Several approaches to visual servoing were experimented in the recent past. Three-dimensional (3-D)-based approaches rely on closing the loop in the Cartesian space [3], [4]. These are less robust than image-based approaches [5], [6], where error signal measurement and loop closure are performed directly at the image level, thus bypassing any inaccuracies in camera calibration and kinematic models [7].

Notwithstanding the improvements of the last few years, much work remains to be done on both the control and computer vision aspects of visual servoing. Design issues with a key impact on the overall characteristics and performance of a visual servoing system are the models of visual interaction and the type of image features used for object tracking. The tracking strategies proposed so far have been based mainly on realistic models of camera projection (e.g., perspective) but rather simple primitives such as points or lines [8]–[10]. Yet, on the one hand tracking such primitives can be infeasible and/or lead to unreliable results for some robotic contexts, and on the other hand simpler camera models are accurate enough to carry out visual analysis in a number of real contexts [11], [12].

In this paper, we investigate using linear approximations in the modeling of an image-based visual servoing system with the purpose of devising control strategies. In Section II, we show that embedding an affine camera-object interaction model into the system model, the linear mapping between appearance evolution and 3-D motion assumes a form which is particularly suitable for real-time servoing, since it is independent of the number and type of features being tracked over time. We also expound a method to cope with the intrinsic limitations of linear models while keeping the advantages of the proposed framework.

The implementation of an affine visual servoing system featuring a manipulator-mounted camera and using active, affine-deformable contours as image features is described and motivated in Section III. The system includes a planner which permits unambiguous and safe task completion, an image-based servo controller, and an estimation module for on-line system parameter identification. Experimental results obtained with a setup featuring a PUMA manipulator are discussed in Section IV, providing an insight into system performance in terms of robustness and application perspectives.

II. Theoretical Framework

A. Preliminaries and Notation

Let us constrain the positioning problem to the geometric interaction between a camera and a single rigid object, in motion one w.r.t. the other. If $\boldsymbol{P}$ is a generic point of the object, its coordinates are expressed in a generic reference frame $\{w\} = \{\boldsymbol{P}_w, {}^wX, {}^wY, {}^wZ\}$ as ${}^w\boldsymbol{P} \in \mathbb{R}^3$. In the following, we use a camera-centered frame $\{c\}$ fixed to the camera and with ${}^cZ$ parallel to the optical axis, and a frame $\{o\}$ fixed to the object. Relative motion of camera and object is described by means of the relative twist screw $\Delta\boldsymbol{V} = \boldsymbol{V}_c - \boldsymbol{V}_o$, where $\boldsymbol{V}_c = [\boldsymbol{T}_c^T \quad \boldsymbol{\Omega}_c^T]^T$ and $\boldsymbol{V}_o = [\boldsymbol{T}_o^T \quad \boldsymbol{\Omega}_o^T]^T$ are, respectively, the camera and object twist screws, $\boldsymbol{T}$ and $\boldsymbol{\Omega}$ indicating translational and angular velocities.

Let $\boldsymbol{p} = [x \quad y]^T \in \mathbb{R}^2$ be the image projection of point $\boldsymbol{P}$. The generic camera projection model is expressed by $\boldsymbol{p} = \pi({}^o\boldsymbol{P}, \boldsymbol{\gamma})$, where $\boldsymbol{\gamma} \in \mathbb{R}^g$ is a vector of camera parameters. The image velocity of $\boldsymbol{p}$ can be expressed as

$$\dot{\boldsymbol{p}} = B(\boldsymbol{p}, \zeta, \boldsymbol{\gamma})\, {}^c\Delta\boldsymbol{V} \tag{1}$$

where matrix B is a function of the image point, its associated depth $\zeta\colon \mathbb{R}^2 \times \mathbb{R}^g \to \mathbb{R}$ s.t. $\zeta(\boldsymbol{p}, \boldsymbol{\gamma}) = {}^cZ$, and camera parameters. By choosing the $2N$-vector $\boldsymbol{p} = [\boldsymbol{p}_1^T \quad \boldsymbol{p}_2^T \quad \cdots \quad \boldsymbol{p}_N^T]^T$ of image point coordinates as the visual state, and the relative twist ${}^c\Delta\boldsymbol{V}$ as the input vector, the dynamics of image appearance is driftless, time-varying

Manuscript received May 26, 1997; revised October 26, 1998. This paper was recommended for publication by Associate Editor S. Hutchinson and Editor V. Lumelsky upon evaluation of the reviewers' comments.

B. Allotta is with PERCRO Lab, PERCeptual RObotics Laboratory, Scuola Superiore Sant'Anna, Pisa I-56127, Italy (e-mail: ben@sssup.it).

C. Colombo is with the Dipartimento di Sistemi e Informatica, Firenze I-50139, Italy (e-mail: columbus@dsi.unifi.it).

Publisher Item Identifier S 1042-296X(99)01211-2.