AN INTEGRATED APPROACH TO ROAD CENTERLINE RECONSTRUCTION USING STEREO IMAGE SEQUENCES FROM A MOBILE MAPPING SYSTEM

Chuang Tao

The University of Calgary, Department of Geomatics Engineering
2500 University Drive, NW, Calgary, AB, Canada T2N 1N4,
IWG V/III

KEY WORDS: Vision, automation, integration, modeling, object reconstruction, image sequence analysis, image matching.

ABSTRACT:

The reconstruction of 3D road centerlines becomes a physical problem of solving an energy-minimizing 3D B-splines shape model based on "*shape from sequences*". The reconstruction is described as a process whereby a 3D road centerline shape model is deformed gradually, driven by forces arising from object space (internal energy) and image space (external energy). The integration of multiple constraints from a mobile mapping system is implemented. Recent test results demonstrate that this approach functions reliably even in situations where the road condition is far from ideal.

1. INTRODUCTION

The management of vehicles and infrastructure requires a high quality and up-to-date highway related spatial information system. Road centerline information is very important for generating road network information systems. It can be used to compute road inspection parameters such as the longitudinal profile and the surface deformation. The acquisition of up-to-date road centerline data by conventional field survey is prohibitive in cost and in actual environment reasons. Since 1992, The University of Calgary jointly with GEOFIT Inc. has been developing a mobile mapping system, VISAT™, for fast spatial information collection, especially for road inventory (Schwarz et al., 1993; Li et al., 1994). In the system, CCD digital cameras, mounted on the top of the van, have been employed to collect stereo and sequential images of road centerlines. The integration of the GPS and the INS has been applied in vehicle location and image sequences georeferencing. The information of road centerlines can be extracted from images during the post-mission processing. The first system demonstration was made in July 1993. The prototype system was tested between October 1993 - October 1995. The production system is available since 1995.

The extraction of road centerline information from images has been researched in vision-based vehicle navigation (Thorpe et al., 1988; Schneiderman and Nashman, 1994). But the automated reconstruction of 3D road centerlines from mobile mapping systems to generate a road network information system is still a new project (He and Novak, 1992). In our system, not only is the centerline information required to be extracted in an automatic way, but also the accuracy and robustness of the extracted data should meet the requirements of mapping applications. This greatly differs from the previous research.

An integrated approach to road centerline reconstruction from image sequences is proposed. In this approach, the reconstruction of road centerlines from image sequences has been considered as a problem on *"shape from sequences."* The problem is to synthesize road centerline information available from successive images into a 3D shape model. Firstly, a 3D physically-based shape model of road centerlines is set up. In order to synthesize the constraint information coming from object assumptions and image sequences into the model, the model is defined as an active and deformable 3D curve. The physically-based deformation mechanism has been endowed with the model such that the model can be progressively deformed under the action of internal and external constraint forces. The extracted data of road centerlines from image sequences can act as external energy which enforces the model to deform towards its desired position. Internal energy arises from smoothness constraints representing the natural characteristic of the shape of road centerlines. It maintains the a priori assumptions about the shape of the model. Under a combination of the actions of internal and external forces, the model will be deformed incrementally towards the final state at which forces from different sources are balanced. The model resulting at the end of an input sequence represents a 3D road centerline shape.

This new approach leads a solution of the integration of multiple constraints in the mobile mapping system: ego-motion constraints (GPS/INS based vehicle trajectory determination and image sequence georeferencing), object space constraints (road model based reconstruction and model driven feature extraction/matching), and image space constraints (stereo and motion image geometry). To implement the approach, three key problems are involved:

- How to set up an approximate 3D shape model of road centerlines?
- How to synthesize multiple constraints into a shape model?
- How to obtain the reliable constraint information (external energy) from image sequences?

The above problems are addressed in section 2, 3 and 4, respectively. The test results of real image sequences and computational aspects of the approach are evaluated in section 5. Concluding remarks and future work are given in section 6.

2. VEHICLE TRAJECTORY DETERMINATION FOR MODEL INITIALIZATION

The approximate 3D shape model of road centerlines forms the basis of the implementation of *"shape from sequences."* A novel method for obtaining such an approximate shape model is proposed. The kinematic vehicle trajectory is utilized to generate an approximate shape model. In VISAT™, a three-

sensor system configuration is applied: GPS, INS and cameras. The GPS and the INS have been integrated to determine the vehicle trajectory (Schwarz et al. 1993). After post-mission Kalman filtering, the trajectory is obtained with high accuracy, and represented in a global coordinate system. The velocity of vehicle is about 60 km/hr and the image capture rate is one image pair per 0.4 second.

A high accuracy method (El-Shemy and Schwarz, 1993) has been used to calibrate parameters of the geometric transformation relationships between the vehicle coordinate system and the various sensor systems such as the GPS, the INS and the camera. As a result, the image sequences are georeferenced in the same coordinate system.

3. SYNTHESIS OF MULTIPLE CONSTRAINTS INTO A PHYSICALLY-BASED 3D SHAPE MODEL

It has been of considerable interest in using deformable models to address image segmentation. The active contour model (2D snake) has been shown successful in interactive boundary extraction (Kass et al., 1988; Terzopoulos et al., 1988; and Menet et al., 1990). The main advantage of deformable models is that both geometric and physical constraints can be incorporated into the model, comparing the conventional methods in which only geometric constraints are considered. For this reason, we proposed an approach based on an idea of "*shape from sequences*", in which a physically-based deformable curve model in space is employed to accommodate the combinations of multiple constraints for the reconstruction of 3D road centerlines.

3.1. 3D Road Shape Modeling by B-splines

In this road model, 3D cubic B-splines are introduced in defining a 3D active curve model (a 3D B-snake):

$$Q(u_j) = (x(u_j), y(u_j), z(u_j)) = (\sum_{i=0}^{m-1} V_i B_i(u_j)), \quad j=0,1,...,n-1 \quad (1)$$

where B_i are the basis functions of B-splines, $V_i=(X_i,Y_i,Z_i)$ are a set of control vertex to this B-spline curve, n and m are the numbers of sampling points on the curve and control vertices, respectively. There are following reasons of choosing B-splines to represent the model: (a) the property of local control: only a small part of the curve needs to change if a control vertices is modified; (b) the possibility of including features (e.g., corners and straight lines): corners or straight lines can be imposed on the curve if multiple knots are posed; (c) invariant characteristics: the control vertices of B-splines are invariant under affine and projective transformation; (d) numeric advantages: the number of unknowns in (1) of determining a curve is m, the number of control vertices, instead of n, the number of sampling points on the curve.

3.2. Deformation Mechanism

By use of the least action principle (Hamilton's principle), the deformable dynamics of the model is governed by an equation of motion:

$$E = e_1 E_{int}(Q(u)) + e_2 E_{ext}(Q(u)) \rightarrow min \quad (2)$$

where E_{ext} represents the external energy acting on the model (E_{ext} is defined in section 3.5), and E_{int} is the internal energy resisting deformable away from the natural state and maintaining the local continuity and connectivity, e_1 and e_2 are constants for weighting these two terms. E_{int} is described as the smoothness energy consisting of two terms:

$$E_{int} = \alpha \left(\int_{\Omega} |Q'(u)|^2 du \right) + \beta \left(\int_{\Omega} |Q''(u)|^2 du \right) \quad (3)$$

The first term is a measure of the distance discontinuities (stretching effectiveness) while the second term is a measure of the orientation discontinuities (bending effectiveness). The constants α and β control the relative balance of stretching and bending force. The total energy E is a function of the set of control vertices. Therefore, minimizing the energy of the curve results in a final shape of the model.

3.3. Solution for the Minimization of Energy

Substituting (1) and (3) into (2), the discrete version we get is

$$E = \sum_{j=0}^{n-1} \{ e_1 [\alpha(u_j) (\sum_{i=0}^{m-1} V_i B_i'(u_j))^2 + \beta(u_j) (\sum_{i=0}^{m-1} V_i B_i''(u_j))^2] + e_2 F_{ext}[Q(u_j)] \} \quad (4)$$

To obtain the extreme minimum of E, Apply

$$\partial E / \partial V_k = 0, \qquad k=0,1,2,...,m \quad (5)$$

Finally the resulting equations are

$$\sum_{i=0}^{m-1} V_i [\sum_{j=0}^{n-1} \alpha(u_j) B_k'(u_j) B_i'(u_j) + \beta(u_j) B_k''(u_j) B_i''(u_j)] = -(e_2/e_1) \sum_{j=0}^{n-1} B_k(u_j) F_{ext}'(Q(u_j)), \quad k=0,1,2,...,m \quad (6)$$

An iterative method (Kass et al., 1988) can be used to solve the above equations:

$$V = (A+\gamma I)^{-1}(\gamma V - F) \quad (7)$$

where γ is an Euler step size, and A is a banded matrix (m×m). It should be noted that the dimension of the inversion is m instead of n (the dimension of the classic snake).

3.4. Reconstruction Procedure

Given an image sequence of road centerlines, the following steps of processing are performed:

1. The vehicle trajectory corresponding to the desired road centerlines is obtained by a post-mission Kalman filter using the GPS/INS navigation data (figure 1). T_i indicates the position of the exposure camera station along the trajectory.

2. The initial line segments of road centerlines in the first stereo image pair of the image sequence are delineated by the operator. After photogrammetric intersection, a corresponding 3D line segment is determined. This segment is denoted by $\overline{S_0 S_1}$ in figure 1.

3. Calculate the vector differences $\overline{T_i T_{i+1}}$ between the positions of station T_i and T_{i+1}. Shift the differential vector $\overline{T_i T_{i+1}}$ from the trajectory to the desired centerline position one by one. A sequence of points S_i on the approximate centerline is obtained.

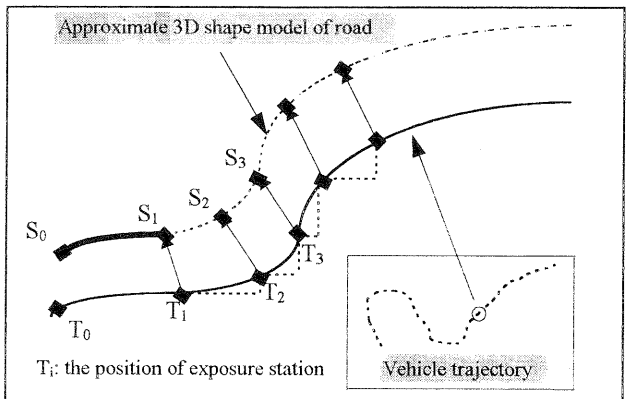

Figure 1. Generation of a 3D approximate shape model

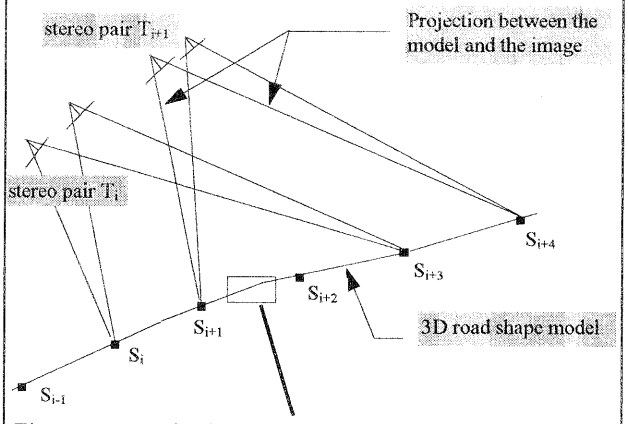

Figure 2(a). Projection between the model and the image

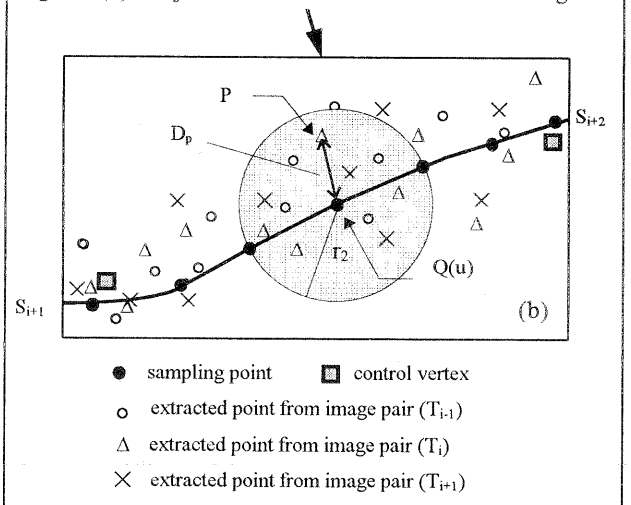

Figure 2(b). External force field

4. Apply an algorithm of B-splines approximation to these data points S_i *(i=0,1,2,...)* and then an approximate 3D shape model of road centerlines is set up (denoting in dotted line, figure 1).

5. Since the orientation parameters of the camera stations are known, the shape model can be back projected on the sequential image (figure 2a). At least two consecutive stereo image pairs covering the same segment along the model are available.

6. The extraction and matching of road centerline features on these images are conducted based on the image projection of the model. Section 4 will describe an algorithm on model driven extraction and matching of road centerline features. At the end of this step, 3D points associated with image features of road centerlines are obtained. In figure 2b, The symbols marked by "o", "Δ" and "×" represent the 3D points obtained by the feature extraction and matching from images. As can be seen, it looks like a deformable curve placed in a force field. Each point in the field acts a force on the curve and deforms the curve. We will define these forces (external forces) in section 3.5.

7. The internal and external energy for the curve segments from S_i to S_{i+3} are calculated. The final shape of these segments is determined by solving the motion equations described in section 3.3.

8. Move to the next station (e.g., $S_i \rightarrow S_{i+1}$). Repeat the process from the step 5 to 7. The shape model is thus deformed incrementally, driven by the successive images. The final result of the deformation of the model is a 3D centerline shape when the last sequential image pair is processed.

3.5. Definition of External Energy

The external energy plays a key role in pushing the model into the desired position. As shown in figure 2b (2D illustration), there are forces between extracted points P and the curve points $Q(u)$. In order to quantify this kind of energy, we apply a gravity-type field to describe it. The reason is that the closer the distance between the points, the greater the force. To avoid the singularity when the distance is approaching zero, we employ the following function in quantifying the external energy:

$$E_{ext} = f(D_p / r_1) \tag{8}$$

and the above $f(x)$ is defined as:

$$f(x) = \begin{cases} x^2, & r_1 > D_p \\ x^{-1}, & r_1 \le D_p \le r_2 \\ 0, & r_2 < D_p \end{cases} \tag{9}$$

where $D_p = ||Q(u)P||$, the distance between $Q(u)$ and P, r_1 and r_2 are the coefficients representing the range of the influence of the point to the curve. $r_1 = 0.1$(meter) and $r_2/r_1 = 3$ are chosen according to the system accuracy. Thus $f(x)$ acts like a spring when the point is close to the curve ($D_p < r_1$) and makes no effect when $D_p > r_2$. The corresponding external force can be derived:

$$F_{ext} = \nabla E_{ext} = f'(D_p / r_1) / r_1 \tag{10}$$

Considering the historic effect of forces, the total F_{ext} becomes

$$F_{ext} = \omega_1 F_{ext}(T_{i-1}) + \omega_2 F_{ext}(T_i) + \omega_3 F_{ext}(T_{i+1}) \tag{11}$$

Again, ω_1, ω_2 and ω_3 are weighting coefficients. During the process of the iterative solution for equations (7), the F_{ext} of each point $Q(u)$ along the model should be calculated repeatedly until the motion equations reach a converge state.

4. MODEL DRIVEN FEATURE EXTRACTION AND MATCHING ON SEQUENTIAL IMAGES

This is a critical part of acquiring the reliable external constraints from image sequences. The approximate positions of road centerlines on the image can be predicted by the projection of the model. The processing window can be constrained in a local area.

4.1 Model Driven Feature Extraction

Firstly, A new oriented edge detector is introduced to detect edge candidates and compute the edge magnitude and orientation (Tao et al., 1996). Secondly, a feature filtering algorithm is proposed to filter the undesired feature points based on the following three constraint components represented by a group of rules:

•**Single edge constraint**: *Rule(1)* The orientation of the desired edge should be perpendicular to the predicted line with an acceptable tolerance of ±20°. *Rule(2)* The magnitude of the edge should surpass a certain threshold.

•**Dual-edge constraint**: road centerlines are painted by white or yellow markings with a certain width. Based on this knowledge, a dual-edge constraint is designed. A dual-edge is composed of a pair of edges locating on the same row of the image. *Rule(3)* The orientations of these two edges must be opposite to each other. *Rule(4)* The distance between these two edges should be within a few pixels (2-5 pixels). *Rule(5)* The average gray value of the zone between these two edges must surpass a threshold and must be brighter than that of the surrounding areas.

•**Shape constraint**: an extracted edge should be located along a smoothed line and may not have big difference from the predicted line. *Rule(6)* If the distance between a dual-edge with its nearest one is beyond a range, this dual-edge will be removed as a blunder (continuity constraint). *Rule(7)* The position of the center point of the dual-edge should locate along a smoothed line (smoothness constraint).

Rule (7) has been implemented by a modified Hough transform (Tao et al., 1996). The distance between the center point of dual-edge and the predicted line is projected onto the Hough space. After the accumulation of the number of the distances, the majority of the distance range is determined and then the blunders are recognized. After feature filtering, the output feature point is the center point of the dual-edge.

4.2 Constrained Matching Range

Because the orientation parameters and the height parameter of each camera station are known, the corresponding point on the image B of a given point on the image A can be estimated. Along with the eppipolar line constraint, the searching range for matching on the image B can be located (Tao et al., 1996). In figure 3, the dotted straight line means the eppipolar line. Three eppipolar lines corresponding to three image pairs pass through the point P_0 . The box shows the constrained matching range. This method provides an effective way to address the enduring problem of matching images with large geometric discrepancies.

4.3 Spatio-temporal Consistency Matching

In order to make use of sequential image information, a scheme of spatio-temporal consistency matching is proposed. The scheme is designed as follows (in figure 3, assuming the current left image is the master image):

1. For each extracted point on the master image, perform matching on the other three images within the constrained window. Kanade's SSD correlation criteria (Kanade et al., 1992) are used in our matching.

2. For the matched point on the next left image and the current right image, locate their best matches on the next right image in their constrained matching ranges, respectively (the arrow indicates its own constrained matching range in figure 3).

3. If two of three matching points on the next right image are closer than 1.5 pixels, the matching point pair is determined. In figure 3 case, P_0 and P_3 are considered as a matching point pair because the distance between the point "Δ" and the point "•" is closer than 1.5 pixels.

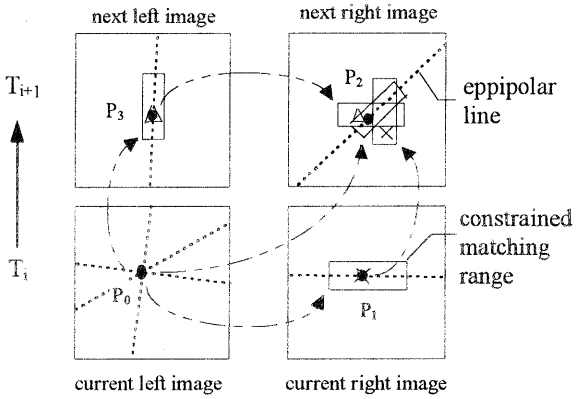

Figure 3. Spatio-temporal matching

Each image will be taken as the master image, and the above process will be executed in a loop. Finally, using photogrammetric intersection, the corresponding 3D point of a matching point pair can be computed and will act as an external energy source (as shown in figure 2b).

5. EVALUATION OF THE APPROACH

5.1 Computational Aspects

Several parameters must be set: a desirable trade-off can be achieved based on the trajectory accuracy outputting from the Kalman filtering algorithm. If the accuracy of trajectory is low, then $e_1 = 0.3$ ($e_2 = 1 - e_1$) is chosen, otherwise $e_1 = 0.5$. To keep a smooth centerline shape, $\alpha = 0.7$ and $\beta = 0.5$ are used. $\omega_1 = 0.5$, $\omega_2 = 1.0$ and $\omega_3 = 0.8$ are selected for representing historical effects of the external energy.

Two main parts of the approach involving time-consuming aspects are: (a) The computation of the inversion in equation (7) by LU decomposition needs O(m) time. In our implementation, control vertices are seeded every 2.0~2.5 meters along the model. The dimension (m) of the inversion is

10-20 according to the vehicle velocity and image capture rate. Without using B-splines curve representation, the dimension (n) would be 80-160 (assuming sampling interval on the curve is 0.3 meters); (b) Due to the known orientation parameters of cameras and a 3D road shape model, the processing window constraint for feature extraction and the constrained matching range for feature matching are available. The computation problem of the highly time-consuming part is greatly reduced.

5.2 Test Results

The software based on this approach has been implemented and installed in the ImagExpert system which is a software package for automatic processing of VISAT™ images. This package is developed on a SUN SPARC workstation with an X-Window and MOTIF environment. A high level object-oriented development toolkit WNDX was used to design the Graphic User Interface. The package has been also interfaced with AUTOCAD V12. The built-in communications between our package and the AUTOCAD core was implemented. The process of the reconstruction can be displayed in AUTOCAD environment simultaneously.

The present approach has been widely tested on real image sequences captured by VISAT™ system. Tests have been conducted to evaluate the approach with images under different road conditions, such as roads with missing, intermittent and faded centerline markings caused by shadows, cracks and wet roads, solid and dashed line patterns, and sharp curves and gaps. The typical examples are given in figure 4-6 (dotted black lines represent the results). On these complex cases, the approach performed successfully and reliably. Guided by our road model, no failure problems occurred even in big gaps and turns. The accuracy of the reconstruction result will be discussed in a separate paper.

6. CONCLUDING REMARKS AND FUTURE WORK

An integrated approach has been proposed to synthesize multiple constraints from image sequences, in which the reconstruction of road centerlines has been considered as a *"shape form sequences"* problem. Based on our tests, the approach has the following characteristics:

- *Physically-based mechanism*: The synthesis of multiple constraints has been implemented by the least action principle, a generalized principle of physical motion. The advantage of the method is that constraints both from object assumptions and image sequences can be incorporated into a model.
- *Model initialization*: The highly accurate vehicle trajectory determined by a multi-sensor (GPS/INS) integration technique is used for generating an approximate 3D shape model of road centerlines.
- *Model based reconstruction*: The whole process of the reconstruction is designed based on a 3D shape model. This design leads to a robust result because the shape model is deformed incrementally under the combinations of the actions of the internal and external energy.
- *Model driven feature extraction and matching*: Effective geometric constraints are available for feature extraction and matching on images with large geometric discrepancies. The combined processing of stereo and motion image can also be applied.
- *B-Splines modeling (3D B-snake)*: Not only geometric characteristics, but also numeric advantages can be achieved in 3D curve reconstruction.

Due to the robustness of the approach, in future work we intend to combine this approach with a Kalman filter based navigation algorithm. In this sense, accurate navigation using multiple sensors and 3D line feature reconstruction are integrated and solved in a combined case. This scheme is used to take both navigation error and reconstruction error into account and to further improve the global accuracy of both results.

7. ACKNOWLEDGMENTS

I would like to give special thanks to Dr. R. Li and Dr. M. A. Chapman who introduced me to this research and gave me invaluable comments. I would like to thank Dr. K. P. Schwarz who led the project from an original idea to a production system. I would also like to thank my colleagues, Mr. L. Qian and Mr. Y. Xu for their great help on implementing the system. Mr. Derek Lichti's help on language correction is appreciated. Finally, the support of Canadian Natual Science and Engineering Research Council (NSERC) and GEOFIT Inc., Laval, Quebec is appreciated.

8. REFERENCES

H. Baker, R. Bolles, and J. Woodfill, 1994. Real Stereo and Motion Integration For Navigation, Int Archives of ISPRS Com. III symposium, Munich, SPIE-2357, pp. 17-24.

N. El-Shemy, and K. P. Schwarz, 1993. Kinematic Positioning in Three Dimensions Using CCD Technology, Vehicle Navigation and Information System'93 Conf., Ottawa.

G. He, and K. Novak, 1992. Automatic Analysis of Highway Features from Digital Stereo Images, Int. Archives of ISPRS Com. III, part B3, pp.119-124.

T. Kanade, M. Okutomi, and T. Nakahara, 1992. A Multiple-baseline Stereo Method, Proc. of ARPRA Image understanding Workshop, San Diego, CA, pp. 409-426.

M. Kass, A. Witkin, and D. Terzopoulos, 1988. Snakes: Active Contour Models, Int. J. of Comp. Vision, pp. 321-331.

R. Li, K. P. Schwarz, M. A. Chapman, and G. Marcel, 1994. Integrated GPS and Related Technologies for Rapid Data Acquisition, GIS World, Vol.7, No.4, 41-43.

S. Menet, P. Saint-Marc, and G. Medioni, 1990. Active contour models: overview, implementation and applications, Int. Conf. Syst. Man Cybernet. pp. 194-199, Los Angeles.

H. Schneiderman, and M. Nashman, 1994. A Discriminating Feature Tracker for Vision-Based Autonomous Driving, IEEE Trans. on Robotics and Automation. Vol. 10, No. 6, pp. 769-775.

K. P. Schwarz, H. Martell, N. El-Sheimy, R. Li, M. Chapman, and D. Cosandier, 1993. VISAT - A Mobile Highway Survey System of High Accuracy, Vehicle Navigation and Information System conference'93, Ottawa, October 12-15.

C. Tao, 1996. Road Centerline Reconstruction from Seuqential Images Based on Shape from Sequences, SPIE proceedings of Visual Communciations and Image Processing'96, Orlando, Florida, USA, March 17-20.

C. Tao, R. Li, and M. A. Chapman. 1996. A Model Driven Approach for Extraction of Road Line Feature Using Stereo Image Sequences From A Mobile Mapping System, to be published in the proceedings of ASPRS/ACSM Annual Convention, April, Baltimore, Maryland.

D. Terzopoulos, A. Witkin, and M. Kass, 1988. Constrains on deformable models: Recovering 3D shape and nonrigid motion, Artificial Intelligence, Vol. 36, pp. 91-123.

C. Thorpe, et al., 1988. Vision and Navigation for the Canegie-Mellon Navlab, IEEE Trans. on PAMI, Vol. 10, No. 3, pp. 362-372.

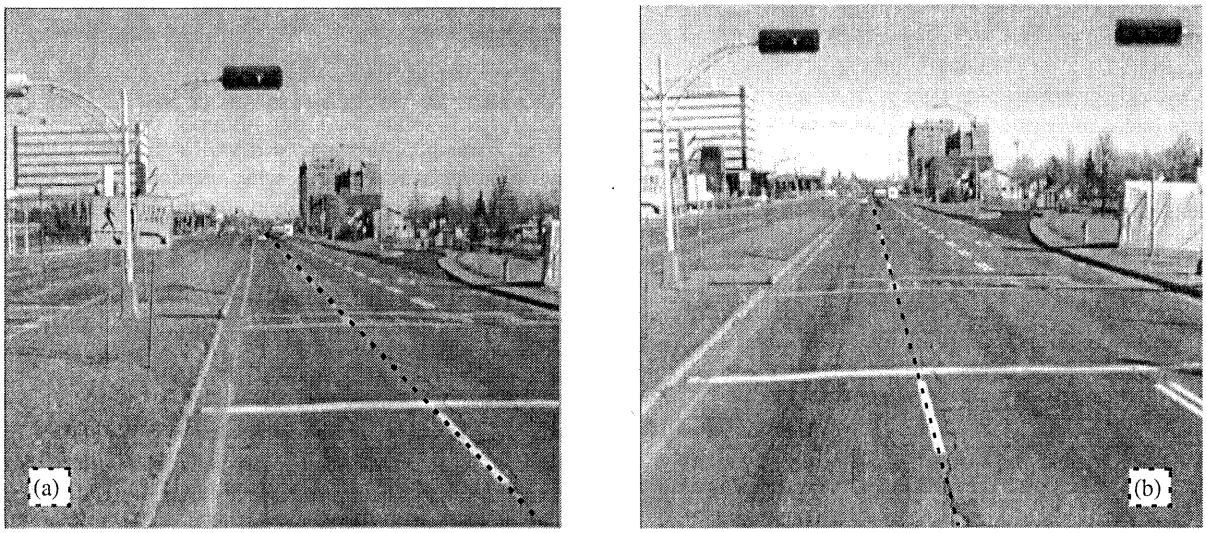

Figure 4. Reconstruction result (dotted black line); *case*: *gaps and noisy markings*, (a) left image (b) right image

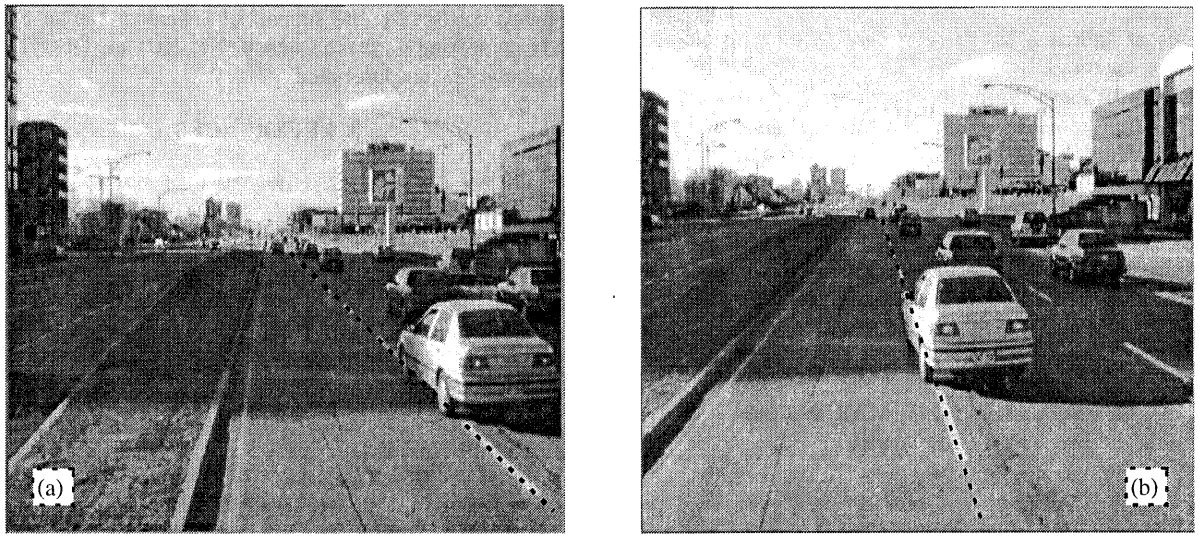

Figure 5. Reconstruction result (dotted black line); *case*: *shadows, cracks and blocks (car),* (a) left image (b) right image

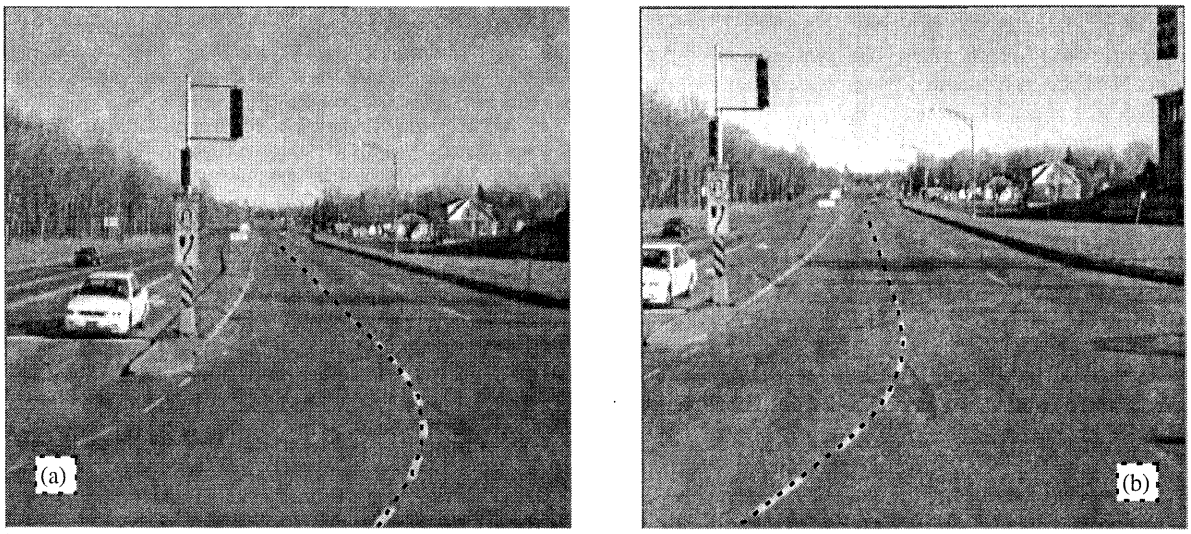

Figure 6. Reconstruction result (dotted black line); *case*: *sharp curve and dashed line pattern,* (a) left image (b) right image