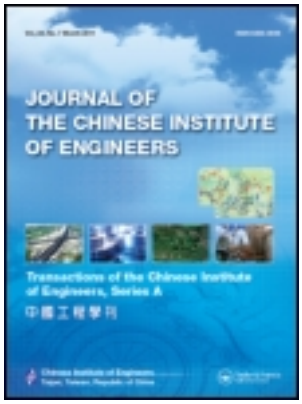

Journal of the Chinese Institute of Engineers
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tcie20

Image tracking of motorcycles and vehicles on urban roads and its application to traffic monitoring and enforcement

Jen-Chao Tai [a] & Kai-Tai Song [b]

[a] Department of Mechanical Engineering , Minghsin University of Science and Technology , Hsinchu Country 304, Taiwan, R.O.C.

[b] Department of Electrical Engineering , National Chiao Tung University , Hsinchu 300, Taiwan, R.O.C. Phone: 886-3-5731865 Fax: 886-3-5731865 E-mail:

Published online: 04 Mar 2011.

To cite this article: Jen-Chao Tai & Kai-Tai Song (2010) Image tracking of motorcycles and vehicles on urban roads and its application to traffic monitoring and enforcement, Journal of the Chinese Institute of Engineers, 33:6, 923-933, DOI: 10.1080/02533839.2010.9671681

To link to this article: http://dx.doi.org/10.1080/02533839.2010.9671681

IMAGE TRACKING OF MOTORCYCLES AND VEHICLES ON URBAN ROADS AND ITS APPLICATION TO TRAFFIC MONITORING AND ENFORCEMENT

Jen-Chao Tai and Kai-Tai Song*

ABSTRACT

Image tracking has increasingly gained attention for use in vision-based traffic monitoring and surveillance applications. For many cities in Asia countries, it is desirable to detect multiple motorcycles as well as cars for urban traffic monitoring and enforcement. In this paper, a novel contour initialization and tracking algorithm is presented to track multiple motorcycles and vehicles at any position on the roadway. This method has the capability to detect moving vehicles of various sizes and to generate their initial contours for image tracking. The proposed method is not constrained by lane boundaries or vehicle size. To track vehicles on roadways, dynamic models are designed to predict the horizontal and vertical positions of vehicle contours. A Kalman filter is designed to update the prediction based on real-time image measurement. Practical experimental studies using video clips are presented to evaluate the performance of the proposed method. Traffic parameters such as traffic flow, vehicle speeds and traffic density are obtained with satisfactory accuracy.

Key Words: traffic monitoring, active contour model, Kalman filter, image tracking.

I. INTRODUCTION

In recent years, CCTV cameras have been widely used on urban roads for surveillance and traffic monitoring applications. Many machine vision techniques have been developed to derive useful information from the acquired images. The information obtained from image sequences allows precise vehicle tracking and classification. Useful traffic parameters including vehicle speeds, traffic flow, etc. can be obtained (Faro *et al*., 2008; Wu *et al*., 2007; Lin *et al*., 2006; Lan *et al*., 2003; Hsu *et al*., 2004; Tai *et al*., 2004; Pece and Worrall, 2002; Lim *et al*., 2002). In practical applications, various types of vehicles on a multi-lane road need to be segmented and tracked simultaneously. Further, for many cities in Asia countries, it is desirable to detect multiple motorcycles as well as cars in urban traffic monitoring and enforcement applications. This paper aims to study a timely and precise tracking initialization procedure in image tracking of multiple vehicles on a roadway. After contour initialization, a robust tracker can be applied for vehicle tracking.

Many powerful tools for real-time contour initialization and tracking have been reported for image-based traffic monitoring and enforcement. Pece and Worrall (2002) proposed an expectation-maximization (EM) contour algorithm to track vehicles. Their method used cluster analysis of image difference to accomplish tracking initialization. Lim *et al*. (2002) presented a feature-based algorithm to obtain vehicle states. Loop detectors were used to initialize image tracking for vehicle speed estimation in their design. Masoud *et al*. (2001) employed sets of blobs and rectangular patches to track vehicles. Their method established the correspondence among blobs and tracked vehicles for tracking initialization. Kamijo *et al*. (2000) proposed a spatio-temporal Markov random field model

*Corresponding author. (Tel: 886-3-5731865; Fax: 886-3-5715998; Email: ktsong@mail.nctu.edu.tw)

J. C. Tai is with the Department of Mechanical Engineering, Minghsin University of Science and Technology, Hsinchu Country 304, Taiwan, R.O.C.

K. T. Song is with the Department of Electrical Engineering, National Chiao Tung University, Hsinchu 300, Taiwan, R.O.C.

to obtain the state of each pixel for tracking purposes. They employed a slit at each entrance to examine entering vehicles. Hsu *et al*. (2004) used the concept of detection zone and entropy to monitor similar-sized cars. Lai *et al*. (2000) proposed the idea of virtual loop and direction-based motion estimation to classify and track vehicles, assigning virtual loops to each lane for tracking. Tai *et al*. (2004) presented an initialization method exploiting the concept of detection line and contour growing. Their algorithm achieves reliable vehicle tracking at road intersections. However, one drawback of their method is that it simply uses fixed-size models for all vehicles. Such a method does not have the capacity to detect and to track vehicles of various dimensions simultaneously. Thus, it is desirable to develop a tracking algorithm for general traffic imagery with vehicles of various dimensions, including cars and motorcycles. Further, previous methods cannot handle vehicles that travel across lane boundaries in the initialization stage of image tracking. Urgent attention is required to develop an initialization and tracking algorithm for all moving vehicles in any position on a multi-lane road.

In this study, a vision-based traffic monitoring system (VTMS) is developed to automatically detect and track multiple vehicles on a multi-lane road. The image size and shape of a moving vehicle often vary in an image sequence due to vehicle motion in the camera's field of view. The VTMS needs to segment and exactly recognize the same vehicle under such conditions. To do so, an active contour model is adopted to represent a vehicle and cope with the robustness problems of vehicle tracking (Vard *et al*., 2008; Yilmaz *et al*., 2006; Koschan *et al*., 2002; Iannizzotto and Vita, 2000).

The rest of this paper is organized as follows. Section II gives an overview of the VTMS. An image measurement algorithm for active contour representation will be described. Section III presents the proposed contour initialization method and the image tracking system. Experimental results of traffic parameter estimation of the proposed method will be presented in Section IV. Section V summarizes the contribution of this work.

II. SYSTEM OVERVIEW

Figure 1 shows the block diagram of the proposed vehicle contour initialization and tracking system for traffic monitoring and enforcement. This image tracking system consists of four parts: foreground segmentation, contour initialization, vehicle tracking and traffic parameter estimation. In foreground segmentation the binary images of moving vehicles are determined from image sequences. The contour initialization part detects a vehicle using a specially designed detection window for generating initial vehicle contours. Once initialized, the vehicle contour will be tracked and updated in the image sequence. The vehicle tracking module employs a dynamic model to predict the vehicle contour from its previous states. The contour of a targeted vehicle is iteratively obtained by using image measurement and Kalman filtering. The proposed contour initialization and detailed tracking algorithms will be described in Section III. In the traffic parameter estimation module, useful parameters are calculated and transmitted to the traffic management center. The procedure of finding a vehicle binary image, including foreground segmentation, active contour model, and image measurement, is presented below.

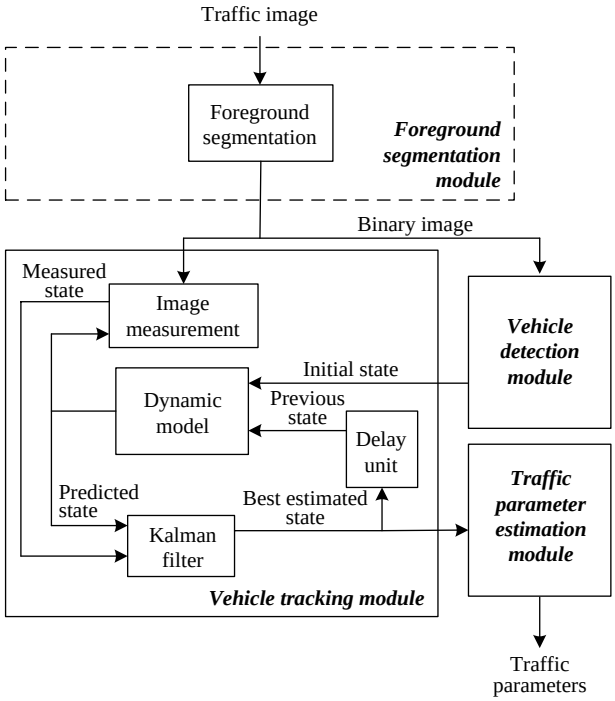

Fig. 1 Block diagram of the system architecture

1. Foreground Segmentation

For an image sequence captured by a static camera, pixel values may have complex distributions. There exist various noisy fluctuations and shadows in the imagery (Song and Tai, 2007), However, for most cases, the intensity of a background pixel dominates the largest Gaussian. For the foreground moving vehicle segmentation, we first apply Gaussian mixture models (GMMs) to generate the background image. GMM approaches to obtaining reliable background images have been widely adopted for many applications (Stauffer and Grimson, 1999). GMM models provide effective background estimation under environmental variations through a mixture of Gaussians for each pixel in an

image sequence. To reduce the computation time, we represent each pixel by using three Gaussian models: one Gaussian for background intensity, one for moving foreground and the other Gaussian for noise. Incoming intensity is checked against the existing Gaussian distributions, until a matching is found. If none of the Gaussian distributions matches the current intensity, the least probable distribution is replaced by a distribution with the current value as its mean, with an initially high variance, and low prior weight.

If a Gaussian matches the current pixel value x_t at time t, the mean μ_t and the variance σ_t^2 of Gaussian is updated such that

$$\mu_t = \mu_{t-1} + \beta(x_t - \mu_{t-1}), \quad (1)$$

$$\sigma_t^2 = (1-\beta)\sigma_t^2 + \beta(x_t - \mu_{t-1})^T(x_t - \mu_{t-1}), \quad (2)$$

where β is the learning rate. The mean and the variance parameters for unmatched distribution remain the same. Further, all weights are updated by

$$\omega_t^* = \omega_{t-1} + \beta(M_t - \omega_{t-1}), \quad (3)$$

where M_t is 1 for a matched Gaussian, otherwise, M_t is 0. All weights are reprocessed by $\omega_t = \frac{\omega_t^*}{S_\omega}$ for normalization, where S_ω is $\sum\omega_t^*$ and the result ω_t is equivalent to the probability of intensity based on past values. The intensity of a background object is most frequently recorded in the image sequence. The Gaussian that has the largest weight is considered as the background model. The Gaussian generated by moving objects has the second largest weight. Foreground pixels can be found if they match the distribution of the foreground Gaussian. Accordingly, a binary image of moving vehicles can be effectively generated using the GMMs model.

Windshields of vehicles might be erroneously recognized as background, because both gray intensities might be very similar. This situation can be improved in foreground segmentation by using morphological hole-filling. Favorable results are obtained as the image of a windshield usually lies inside of a vehicle image after closing operations. The foreground pixels are then grouped into different regions using connected components labeling algorithm for image measurement (Bovik *et al*., 2001). In traffic imagery, shadows attached to their respective moving vehicles introduce distortions and cause problems in image segmentation. Interested readers refer to (Song and Tai, 2007) for a discussion and a statistical method for shadow suppression in traffic imagery.

2. Active Contour Model

Active contour modeling is a powerful tool for model-based image segmentation and tracking (Yilmaz *et al*., 2006; Koschan *et al*., 2002; Lim *et al*., 2002). Based on the active contour concept, an image measurement method for obtaining the best-fit vehicle contour curve for image tracking is presented below. In this work, *B*-spline functions are adopted to represent vehicle contours in image frames. The vehicle contour $(x(s), y(s))$ is represented using N_B *B*-spline functions:

$$x(s) = \sum_{n=0}^{N_B-1} B_n(s)q_n^x = \boldsymbol{B}(s)\boldsymbol{Q}^x \quad \text{for} \quad 0 \le s \le N_B \quad (4)$$

and

$$y(s) = \sum_{n=0}^{N_B-1} B_n(s)q_n^y = \boldsymbol{B}(s)\boldsymbol{Q}^y \quad \text{for} \quad 0 \le s \le N_B, \quad (5)$$

where

$$\boldsymbol{B}(s) = (B_0(s), B_1(s), \cdots, B_{N_B-1}(s))^T,$$

$$B_0(s) = \begin{cases} s^2/2 & \text{if } 0 \le s < 1 \\ \frac{3}{4} - (s-\frac{3}{2})^2 & \text{if } 1 \le s < 2 \\ (s-3)^2/2 & \text{if } 2 \le s < 3 \\ 0 & \text{if otherwise} \end{cases},$$

$$B_n(s) = B_0(s-n),$$

$$\boldsymbol{Q}^x = \begin{pmatrix} q_0^x \\ q_1^x \\ \vdots \\ q_{N_B-1}^x \end{pmatrix} \text{ and } \boldsymbol{Q}^y = \begin{pmatrix} q_0^y \\ q_1^y \\ \vdots \\ q_{N_B-1}^y \end{pmatrix}.$$

The contour $r(s)$ is represented by a vector $\boldsymbol{Q}$ $(=\begin{pmatrix}\boldsymbol{Q}^x \\ \boldsymbol{Q}^y\end{pmatrix})$ and $\boldsymbol{Q}$ contains the X-Y coordinates of the control points of the B-spline curve, such that

$$\boldsymbol{r}(s) = \begin{pmatrix} x(s) \\ y(s) \end{pmatrix} = \boldsymbol{I}_2 \otimes \boldsymbol{B}(s)^T\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{B}(s)^T & \boldsymbol{0}' \\ \boldsymbol{0}' & \boldsymbol{B}(s)^T \end{pmatrix}\boldsymbol{Q}$$

$$= \boldsymbol{U}(s)\boldsymbol{Q}, \quad (6)$$

where $\boldsymbol{I}_2$ denotes a 2×2 identity matrix, $\otimes$ is the Kronecker product denotation of two matrices and $\boldsymbol{0}'$ is $(0, 0, 0, 0, 0, 0, 0, 0)^T$. In this design, a control point vector $\boldsymbol{Q}$ containing eight control points is used to represent the vehicle contour, as shown in Fig. 2. The control points are indicated by circles in the figure.

3. Shape Space Transformation

In a traffic image sequence, the contour of a moving vehicle changes due to its motion and the projective effect of camera view angle. Within a reasonable view

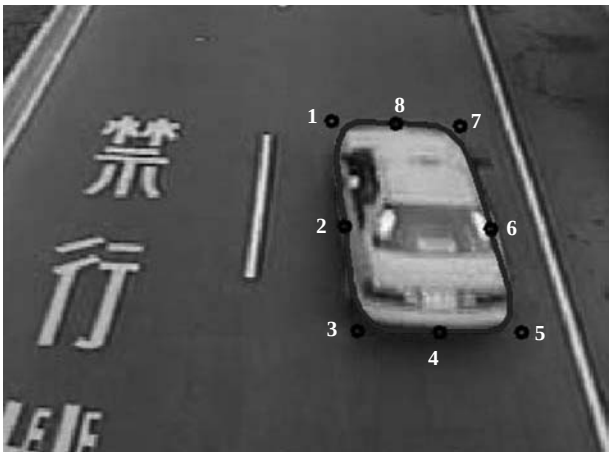

Fig. 2 Active contour of a vehicle

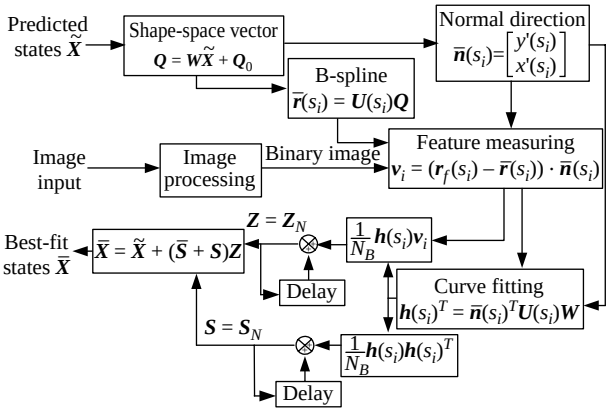

Fig. 3 Flowchart of the procedure for finding shape-space vector of the best fitting curve

angle, it can be assumed that the variation of vehicle contour is linear in traffic imagery. The vehicle contour can then be described by a shape-space planar affine transformation in the image plane. The boundary curve $\boldsymbol{r}(s)$ of each vehicle is expressed using a template curve $\boldsymbol{r}_0(s)$ (Iannizzotto and Vita, 2000).

$$\boldsymbol{r}(s) = \boldsymbol{u}_t + \boldsymbol{M}\boldsymbol{r}_0(s), \tag{7}$$

where $\boldsymbol{u}_t = (u_x, u_y)^T$ is a two-dimensional translation vector and $\boldsymbol{M}$ is a 2 ×2 affine-matrix comprising one rotation and three deformation (horizontal, vertical, and diagonal) elements. Subtracting $\boldsymbol{r}_0(s)$ from Eq. (7), one obtains:

$$\begin{aligned} \boldsymbol{r}(s) - \boldsymbol{r}_0(s) &= \boldsymbol{u}_t + (\boldsymbol{M} - \boldsymbol{I})\boldsymbol{U}(s)\boldsymbol{Q}_0 \\ &= \boldsymbol{U}(s)\begin{pmatrix} \boldsymbol{1}' & \boldsymbol{0}' & \boldsymbol{Q}_0^x & \boldsymbol{0}' & \boldsymbol{0}' & \boldsymbol{Q}_0^y \\ \boldsymbol{0}' & \boldsymbol{1}' & \boldsymbol{0}' & \boldsymbol{Q}_0^y & \boldsymbol{Q}_0^x & \boldsymbol{0}' \end{pmatrix}\boldsymbol{X} = U(s)\boldsymbol{W}\boldsymbol{X}, \end{aligned} \tag{8}$$

where is $\boldsymbol{1}'$ is $(1, 1, 1, 1, 1, 1, 1, 1)^T$, $\boldsymbol{Q}_0^x$, $\boldsymbol{Q}_0^y$ are X-Y coordinates of the control points of the template curve $\boldsymbol{Q}_0$ and the shape-space vector is

$$\boldsymbol{X} = (u_x \;\; u_y \;\; M_{11} - 1 \;\; M_{12} \;\; M_{21} \;\; M_{22} - 1)^T.$$

Subtracting $\boldsymbol{r}_0(s)$ from Eq. (6), Eq. (8) can be rewritten as

$$\boldsymbol{r}(s) - \boldsymbol{r}_0(s) = \boldsymbol{U}(s)(\boldsymbol{Q} - \boldsymbol{Q}_0). \tag{9}$$

Comparing Eq. (8) and Eq. (9), one obtains a linear transformation:

$$\boldsymbol{Q} = \boldsymbol{W}\boldsymbol{X} + \boldsymbol{Q}_0. \tag{10}$$

Using Eq. (10), one can transform a vehicle contour to a shape-space vector $\boldsymbol{X}$. This simplifies the post-processing of contour tracking in the image plane and the vehicle contour will be restricted to varying steadily by the shape-space vector.

4. Image Measurement

The image measurement procedure is responsible for obtaining the best-fit curve of the vehicle contour in an image according to a predicted vehicle contour generated from a predicted shape-space vector $\tilde{\boldsymbol{X}}$, a template curve $\boldsymbol{Q}_0$ and its shape matrix $\boldsymbol{W}$. The binary image of a traveling vehicle is segmented from traffic imagery by using GMM (see Section III.1). The contour feature $\boldsymbol{r}_f(s)$ is obtained by applying one-dimensional (1-D) image processing along the normal direction of a predicted curve. Curve-fitting method of the detected features is employed to obtain the best-fit curve of the vehicle contour $\bar{\boldsymbol{r}}(s)$ (Blake and Isard, 1998). In carrying out the curve fitting of contour features, one has to increase the tolerance for image disturbance and eliminate possible interference from features of other objects in the background. A contour shape-space vector $\tilde{\boldsymbol{X}}$ and a regularization constant α are used to stand for the relative effect of the shape in the curve fitting and meet the criteria mentioned above.

Figure 3 depicts the flowchart of the procedure for finding the shape-space vector of the best fitting curve. Introducing the concepts of information matrix $\boldsymbol{S}_i$ and information weight sum $\boldsymbol{Z}_i$, the algorithm for finding shape-space vector of the best fitting curve can be summarized as follows:

1) Select N regularly equal-spaced samples s_i, i = 1, 2, 3, ⋯, N and $s_1 = h$, $s_{i+1} = s_i + h$, $s_N = Nh = N_B$.
2) For each i, find the position of $\boldsymbol{r}_f(s_i)$ by applying one-dimensional (1-D) image processing along the

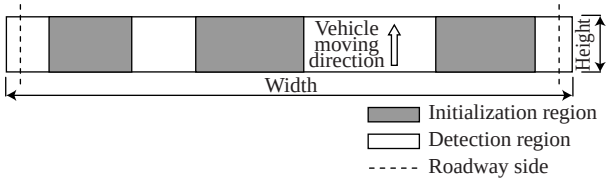

Fig. 4 A detection window consists of initialization regions and detection regions

normal line passing through $\bar{\boldsymbol{r}}(s)$ ($\bar{\boldsymbol{r}}(s)$ is the contour of $\tilde{\boldsymbol{X}}$) at $s = s_i$.

3) Initialize

$$\boldsymbol{Z}_0 = 0,\ \boldsymbol{S}_0 = 0.$$

Iterate, for $i = 1, 2, 3, \cdots, N$

$$\boldsymbol{v}_i = (\boldsymbol{r}_f(\boldsymbol{s}_i) - \bar{\boldsymbol{r}}(s_i)) \cdot \bar{\boldsymbol{n}}(s_i) \tag{11}$$

$$\boldsymbol{h}(\boldsymbol{s}_i)^T = \bar{\boldsymbol{n}}(s_i)^T \boldsymbol{U}(\boldsymbol{s}_i)\boldsymbol{W} \tag{12}$$

$$\boldsymbol{S}_i = \boldsymbol{S}_{i-1} + \frac{1}{\sigma_i^2}\boldsymbol{h}(s_i)\boldsymbol{h}(s_i)^T \tag{13}$$

$$\boldsymbol{Z}_i = \boldsymbol{Z}_{i-1} + \frac{1}{\sigma_i^2}\boldsymbol{h}(s_i)\boldsymbol{v}_i \tag{14}$$

where $\bar{\boldsymbol{n}}(s_i)$ is the normal unit vector of curve $\bar{\boldsymbol{r}}(s)$ at $s = s_i$ and $\sigma_i^2 = N_B$.

4) The aggregated observation vector is

$\boldsymbol{Z} = \boldsymbol{Z}_N$ with the associated statistical information

$\boldsymbol{S} = \boldsymbol{S}_N$

5) The best-fit curve is expressed as a shape-space vector (Blake and Isard 1998)

$$\bar{\boldsymbol{X}} = \tilde{\boldsymbol{X}} + (\bar{\boldsymbol{S}} + \boldsymbol{S})^{-1}\boldsymbol{Z}, \tag{15}$$

where $\bar{\boldsymbol{S}} = \alpha \boldsymbol{W}^T(\frac{1}{N_B}\int_0^{N_B}(\boldsymbol{I}_2 \otimes \boldsymbol{B}(s)^T)^T(\boldsymbol{I}_2 \otimes \boldsymbol{B}(s)^T)ds)\boldsymbol{W}$.

III. THE PROPOSED INITIALIZATION ALGORITHM

The contour initialization step detects the moving vehicle and generates an initial contour for tracking. This step is important for successfully tracking a vehicle in the image sequence. To track multiple vehicles of various sizes on a multi-lane road, we propose a contour initialization algorithm to generate initial contours for image tracking by using a detection window.

1. Contour Initialization

A novel concept of detection window is proposed in this work to handle contour initialization. As shown in Fig. 4, depending on current traffic imagery, there can be multiple detection regions and initialization regions in the detection window. In the beginning, the entire detection window is categorized as a detection region. The system works to check whether there is any vehicle entering the detection region. As a vehicle is detected, the related detection region will change into an initialization region. The rest of the detection region remains unchanged. If the detected vehicle leaves the initialization region, this region will be released and become a detection region again. Thus, the detection region and the initialization region are automatically adjusted according to the current traffic image sequence.

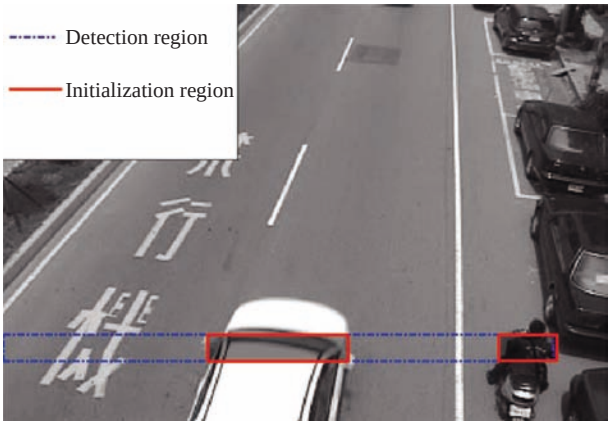

Fig. 5 Detection of a car and a motorcycle

(i) Detection Window

To facilitate vehicle detection, we divide the detection region into several 1-pixel-width sub-regions. An empty sub-region transfers to a filled sub-region if a moving object appears in this sub-region. When the front part of a vehicle, such as a vehicle's bumper, enters the detection window, a cluster of filled sub-regions appears and it grows gradually as the vehicle moves forward. Finally the number of filled sub-regions will be fixed and can be used to find the width of the vehicle in the image plane. In our design, if a related detection region contains enough cluster filled sub-regions (according to an assigned threshold), it will change into an initialization region as mentioned in the previous paragraph. Fig. 5 shows a test example where a car and a motorcycle appear in the detection window simultaneously. Both the car and the motorcycle are detected; two initialization regions are automatically generated in this case.

(ii) Contour Generation

The result of contour initialization is the

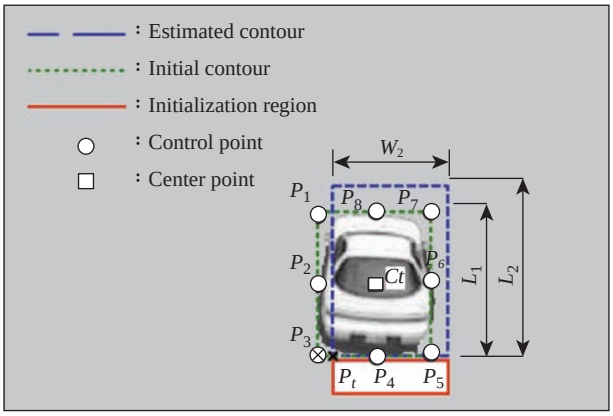

Fig. 6 Generation of an initial contour

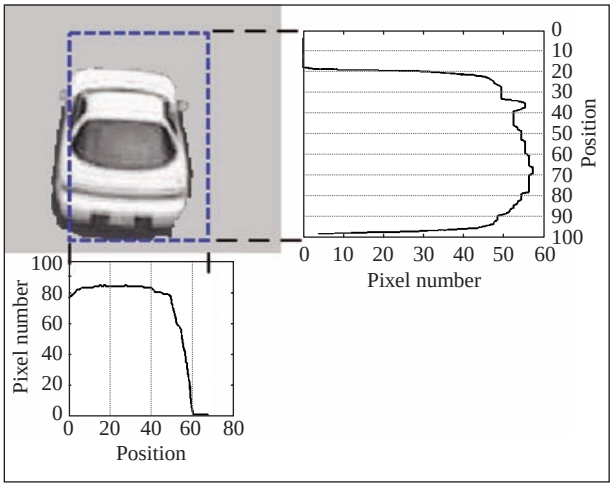

Fig. 7 The projection profile of an estimated contour

generation of an initial contour of the detected vehicle as it leaves the detection window. Fig. 6 depicts the design of initial contour generation. First, an estimated contour is automatically generated by using geometrical information obtained from the initialization region. The width of the estimated contour is the width w of the initialization region. The length L_2 of the estimated contour is assigned to Rw, R is an empirical ratio of the length to the width of vehicle in the captured image frame. The location of the estimated contour is assigned at the exit of the initialization region, as shown in Fig. 6. It is clear that the estimated contour is generated via simple geometrical relationships in the initialization region; it may not perfectly match the actual situation of the vehicle image. However, the estimated contour will be corrected by subsequent image measurement, as described below.

The dimension and location of the estimated contour are corrected to derive an estimation contour by analyzing the actual vehicle image. The length L_2 and the location of the estimated vehicle contour can deviate from the generated initial contour. The width w of the vehicle image was previously estimated by the detection window when the vehicle entered the detection window. The length and the position of vehicle image, however, should be corrected to actual values. In the current design, we analyze the projection information to estimate the size and position of the vehicle image. The binary image is projected to two one-dimensional arrays and the projection is used to measure the occupancy of the vehicle image for obtaining the size and position of the vehicle in the image frame. The projection values are the sum of vehicle pixels along vertical and horizontal directions, respectively. Fig. 7 illustrates the method of estimating the size and position of a vehicle in the image plane. In this case, the vertical projection reveals that the corner point Pt of the initial contour should shift to the point P_3, and the horizontal projection reveals that the length L_2 should be corrected to L_1, as shown in Fig. 6. The control points and center point of the initial contour are then generated according to w, L_1, P_3. The initial contour is employed to obtain the template $\boldsymbol{Q}_0$ and the shape-space vector $\boldsymbol{X}$ for image tracking:

$$\boldsymbol{Q}_0^x = \begin{pmatrix} u_1 - u_c \\ u_2 - u_c \\ \vdots \\ u_8 - u_c \end{pmatrix}, \quad \boldsymbol{Q}_0^y = \begin{pmatrix} v_1 - v_c \\ v_2 - v_c \\ \vdots \\ v_8 - v_c \end{pmatrix} \text{ and } \boldsymbol{X} = \begin{pmatrix} u_c \\ v_c \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

where (u_1, v_1), (u_2, v_2), $\cdots$, (u_8, v_8) and (u_c, v_c) are pixel-based coordinates of points P_1, P_2, $\cdots$, P_8 and Ct.

The sizes and positions of moving vehicles can be detected by using the proposed detection window. Accordingly, a proper contour of each vehicle can be initialized to represent individual vehicles on an unconfined roadway.

2. Kalman Filtering

The vehicle contour is represented by a shape-space vector $\boldsymbol{X}$ with six elements. The first two elements of $\boldsymbol{X}$ are position coordinates of the template curve and the rest of the elements are shape scaling elements, as described in Section II.3. The vehicle tracking module employs two dynamic models (see later) to predict the horizontal and vertical position from their historical position states. The predicted states are provided to the information fusion stage for tracking a vehicle. As for the shape scaling elements, because the change of vehicle contour is very small within two consecutive image frames; it is not necessary to employ complex dynamic models to predict the shape scaling elements. Thus the predicted states of these elements can be simply predicted using the

previous measured state $\bar{\boldsymbol{X}}$ obtained from the image measurement, as described in Section II.4.

Below is the design of our dynamic model for contour prediction. The state of horizontal or vertical position can be governed by

$$x_{k+1} = x_k + (x_k - x_{k-1}) + \varepsilon_k, \tag{16}$$

where x_{k+1} is the position state and ε_k is the system noise. Let $\boldsymbol{X}_k = \begin{pmatrix} x_{k-1} \\ x_k \end{pmatrix}$, then the dynamical model is

$$\boldsymbol{X}_{k+1} = \boldsymbol{A}\boldsymbol{X}_k + \boldsymbol{\Gamma}\boldsymbol{\varepsilon}_k, \tag{17}$$

where $\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix}$ and $\boldsymbol{\Gamma} = \begin{bmatrix} 0 \\ m \end{bmatrix}$. The observation model is

$$\boldsymbol{o}_{k+1} = \boldsymbol{C}\boldsymbol{X}_k + \boldsymbol{\eta}_k, \tag{18}$$

where $\boldsymbol{o}_{k+1}$ is the measurement state, $\boldsymbol{C} = [0 \;\; 1]$ and $\boldsymbol{\eta}_k$ is observation noise. Let $\{\varepsilon_k\}$ and $\{\boldsymbol{\eta}_k\}$ be the zero-mean Gaussian white noises such that $Var(\varepsilon_k) = \boldsymbol{Y}_k$ and $Var(\eta_k) = \boldsymbol{R}_k$ are positive definite matrices and $E(\boldsymbol{\varepsilon}_k\boldsymbol{\eta}_l) = 0$ for all k and l. Eqs. (17) and (18) are the state space description of a linear stochastic system. A Kalman filter is designed to combine the information from the predicted states and the best-fit states obtained from Eq. (15) (Bozic, 1994). The tracking procedure over one time-step is summarized as follows:

1) Predict the *a priori* estimate $\hat{\boldsymbol{X}}_{k+1}^-$ according to the previous state $\hat{\boldsymbol{X}}_k$:

$$\hat{\boldsymbol{X}}_{k+1}^- = \boldsymbol{A}\hat{\boldsymbol{X}}_k. \tag{19}$$

2) Determine the *a priori* estimate error covariance ahead,

$$\boldsymbol{P}_{k+1}^- = \boldsymbol{A}\boldsymbol{P}_k\boldsymbol{A}^T + \boldsymbol{\Gamma}\boldsymbol{Y}\boldsymbol{\Gamma}^T. \tag{20}$$

3) Compute the Kalman gain,

$$\boldsymbol{K}_{k+1} = \boldsymbol{P}_{k+1}^-\boldsymbol{C}^T(\boldsymbol{C}\boldsymbol{P}_{k+1}^-\boldsymbol{C}^T + \boldsymbol{R})^{-1}. \tag{21}$$

4) Use a Kalman filter to obtain the *a posteriori* estimate by incorporating the measurement state $\boldsymbol{o}_{k+1}$:

$$\hat{\boldsymbol{X}}_{k+1} = \hat{\boldsymbol{X}}_{k+1}^- + \boldsymbol{K}_{k+1}(\boldsymbol{o}_{k+1} - \boldsymbol{C}\hat{\boldsymbol{X}}_{k+1}^-), \tag{22}$$

5) Update the *a posteriori* error covariance

$$\boldsymbol{P}_{k+1} = (\boldsymbol{I} - \boldsymbol{K}_{k+1}\boldsymbol{C})\boldsymbol{P}_{k+1}^-; \tag{23}$$

then go to step 1 for next iteration.

At each procedure, the process is repeated with the previous *a posteriori* estimates used to predict the new *a priori* estimates.

3. Vehicle Speed Estimation

As vehicles in an image sequence are successfully tracked, traffic parameters such as traffic flow, vehicle speeds and traffic density can be obtained. The traffic flow can be derived as the ratio of detected vehicle numbers to elapsed time. Traffic density D (car/km) is calculated as follows

$$D = \frac{q}{V_{avg}}, \tag{24}$$

where q is the traffic flow (car/hr) and V_{avg} is the average travel speed (km/hr).

Vehicle speed can be obtained from two recorded positions of vehicle image and the elapsed time between these two positions. The center of the bottom edge of a vehicle image is taken as the reference point of the vehicle position. This point can be easily obtained from the control vectors of the vehicle contour. Assuming the vehicles lie on a flat plane and the camera has been calibrated, we can transform the image coordinate (u_I, v_I) into the world (X_w, Y_w) coordinate as shown below (Schoepflin and Dailey, 2003):

$$X_w = \frac{h}{f\sin\phi}\frac{uv_0}{(v_0 - v_I)} \tag{25}$$

and

$$Y_w = \frac{h}{f\sin^2\phi}\frac{v_0 v_I}{(v_0 - v_I)}, \tag{26}$$

where f is the focal length of the camera, ϕ is the tilt angle of the camera, h is the height of the camera and (u_0, v_0) is the vanishing point of parallel lanes.

From the tracking result of a tracked vehicle, the reference point P_a (tracking operation is initialized) at time t_a and P_b (as the vehicle attains a predefined region, tracking operation terminates) at time t_b are recorded. Using Eqs. (25) and (26), positions P_a and P_b are transformed to world coordinates to calculate the traveling distance L between P_a and P_b. The vehicle speed V_s can be calculated by

$$V_s = \frac{L}{(t_b - t_a)}. \tag{27}$$

IV. EXPERIMENTAL RESULTS

Practical experiments of traffic parameter extraction have been conducted to evaluate the tracking performance of the proposed method by using two video clips of traffic taken in Hsinchu city. The frame rate adopted in both experiments is 15 frame/s. The pixel resolution of each test frame is 352×240 pixels. Fig. 8 illustrates an example of image tracking of cars

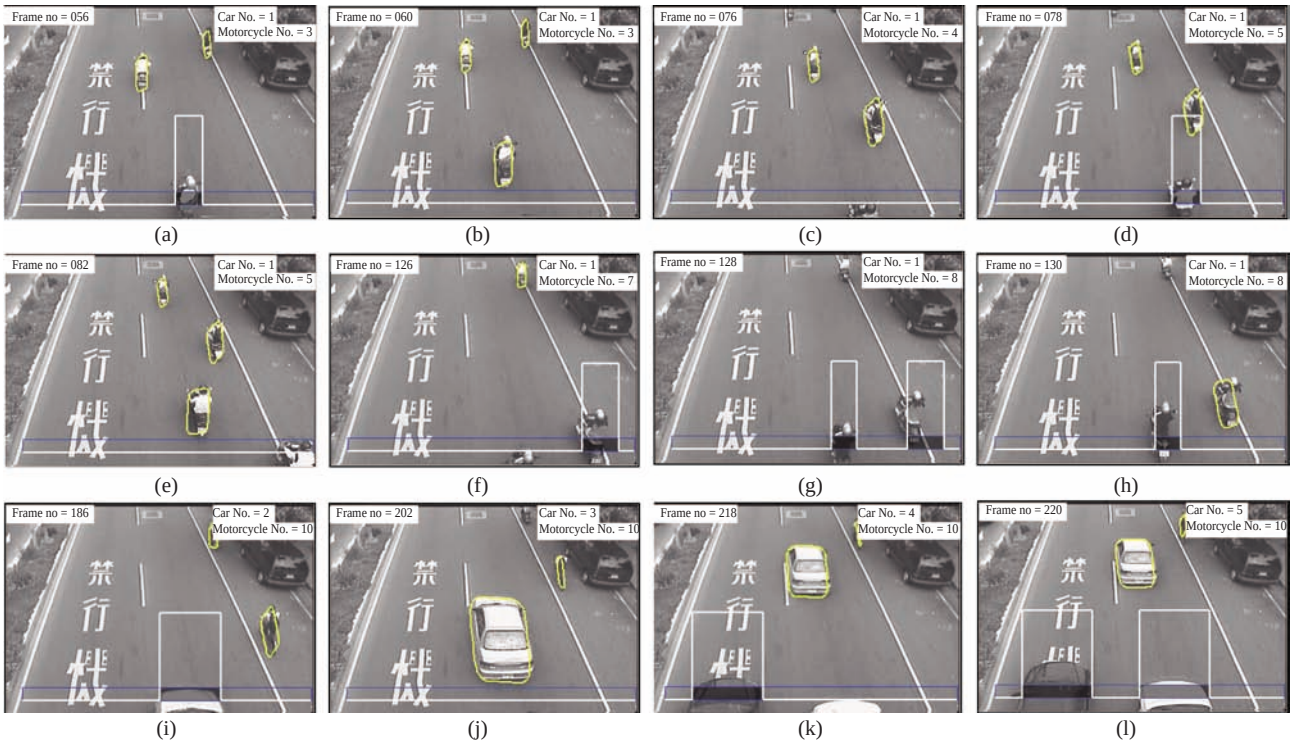

Fig. 8 Experimental results of image tracking of cars and motorcycles

and motorcycles. In Fig. 8(a), two motorcycles are tracked and another motorcycle is detected in the detection window. Fig. 8(b) shows that the motorcycle leaves the detection window and an initial contour is generated accordingly in the initialization region. In Fig. 8(c), two motorcycles are tracked and another motorcycle has not yet entered the detection window. In Fig. 8(d) a motorcycle is detected in the detection window. The motorcycle leaves the detection window and an initial contour is generated accordingly as shown in Fig. 8(e). Another motorcycle is detected in Fig. 8(f). Two motorcycles are simultaneously detected in Fig. 8(g). One motorcycle leaves the detection window and an initial contour is generated in Fig. 8(h). A car is detected in Fig. 8(i). The car leaves the detection window and an initial contour is generated for tracking in Fig. 8(j). The car is tracked while another car is detected in Fig. 8(k). The car is tracked as expected and the other two cars are detected simultaneously in Fig. 8(l). The experimental results demonstrate that the proposed contour initialization procedure successfully provides initial contours for vehicle contour tracking. Moreover, this experiment indicates that cars and motorcycles are detected and tracked simultaneously in any position of a multi-lane road by the proposed method. A video clip of experimental results can be found at http://isci.cn.nctu.edu.tw/video/ItsTracking/Tracking1.wmv.

The second experiment was conducted using traffic recorded by a surveillance camera installed at the main gate of the Hsinchu Science Park, Taiwan, where a traffic monitoring system has been installed for evaluation. Fig. 9 shows the experimental results of vehicle tracking for traffic parameters estimation. In Fig. 9(a), a car is detected in the detection window. The first detected car leaves the detection window and an initial contour is generated accordingly in Fig. 9(b). In Fig. 9(c), two cars are tracked and another detected car passes through the detection window. Fig. 9(d) shows that the detection window detects a vehicle in the left lane. Two cars simultaneously cross the detection window and are detected in Fig. 9(e). Cars leave the detection window and are tracked in Fig. 9(f). It is interesting to note that two tracked cars are close to each other and still have been tracked precisely in Fig. 9(g). In Fig. 9(h), a car traveling between two lanes is detected in the detection window.

In this example, useful traffic parameters are estimated using the proposed method. In a time duration of 20 minutes, a total of 335 vehicles are detected (the ground truth is 334 vehicles). The accuracy of vehicle number estimation is quite satisfactory. This is mainly due to the detection window being able to handle vehicles that travel across lane boundaries. Table 1 shows the experimental results of traffic parameter estimation. In the table, ground truth was manually measured from image sequences. The estimated parameters include average speed:

Table 1 Experimental results of traffic parameter estimation

Time index (min)	Vehicle number (car)		Traffic flow rate (car/hr)		Average speed (km/hr)			Traffic density (car/km)	
	Ground truth	Estimated	Ground truth	Estimated	Ground truth	Estimated	Error	Ground truth	Estimated
1	13	14	780	840	36.9	35.6	3.51%	21.1	23.6
2	38	38	2280	2280	44.0	42.2	4.00%	51.9	54.0
3	6	6	360	360	53.3	51.1	4.17%	6.8	7.0
4	2	2	120	120	43.1	42.9	0.47%	2.8	2.8
5	38	37	2280	2220	42.5	41.7	2.03%	53.6	53.3
6	12	12	720	720	43.5	43.2	0.82%	16.5	16.7
7	15	15	900	900	41.2	41.8	-1.47%	21.9	21.6
8	21	21	1260	1260	43.0	43.8	-1.84%	29.3	28.8
9	5	5	300	300	49.8	50.8	-2.16%	6.0	5.9
10	23	23	1380	1380	36.9	35.6	3.51%	37.4	38.8
11	4	4	240	240	45.6	46.0	-0.80%	5.3	5.2
12	23	23	1380	1380	38.7	38.4	0.77%	35.7	36.0
13	26	27	1560	1620	40.2	40.2	0.01%	38.8	40.3
14	5	5	300	300	48.8	49.6	-1.67%	6.2	6.1
15	30	30	1800	1800	41.6	42.1	-1.28%	43.3	42.7
16	22	22	1320	1320	46.2	47.0	-1.59%	28.6	28.1
17	1	1	60	60	44.6	44.5	0.24%	1.3	1.3
18	27	28	1620	1680	39.3	39.9	-1.48%	41.2	42.1
19	12	12	720	720	45.7	46.8	-2.27%	15.7	15.4
20	11	10	660	600	32.8	34.4	-4.69%	20.1	17.5

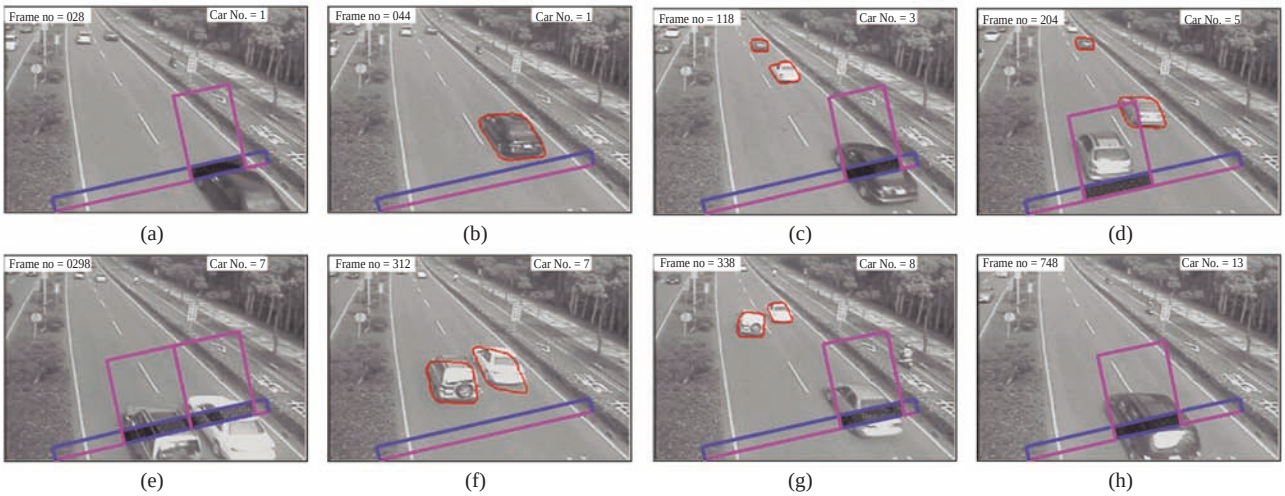

Fig. 9 Traffic monitoring results in Hsinchu Science Park

42.87 km/hr (the ground truth is 42.88 km/hr), the traffic flow rate : 1002 car/hr (the ground truth is 1005 car/hr) and the density is 23.37 car/km (the ground truth is 23.44 car/km). The error of average speed estimation is within 5%. A video clip of experimental results can be found at http://isci.cn.nctu.edu.tw/video/ItsTracking/Tracking2.wmv.

V. CONCLUSIONS

An automatic contour initialization and tracking method have been developed for image tracking of multiple vehicles based on active contour and image measurement. A novel detection window image processing scheme has been proposed to detect moving vehicles of various dimensions and generate their initial contours for image tracking on a multi-lane road. The proposed contour initialization and tracking schemes have been tested for traffic monitoring and enforcement applications. Experimental results show that the proposed method successfully tracks motorcycles as well as multiple cars on an urban multi-lane road. Traffic

parameters are successfully extracted by using the developed method. The error of average vehicle speed estimation is less than 5%.

Several directions are interesting for further study. One interesting topic is to investigate the issue of occlusion in vehicle images. Image occlusion greatly influences the accuracy of image measurement. Methods need to be developed to identify individual vehicles for traffic monitoring and enforcement applications. Color information of individual tracked cars can be beneficial for solving this problem (Hu *et al*., 2004). On the other hand, in order to increase the accuracy of foreground segmentation, it will be interesting to select adaptive thresholds for handling the change of environment illumination on the road.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Council, Taiwan, R.O.C., under Grant NSC 95-2221-E-009-178 and by the Ministry of Education, Taiwan, R.O.C., under Grant EX-91-E-FA06-4-4.

NOMENCLATURE

$B_0(s)$, $B_1(s)$, ⋯, $B_{N_B-1}(s)$	*B*-spline functions
$\boldsymbol{B}(s)$	matrix of *B*-spline functions
D	traffic density
f	focal length of camera,
h	height of camera
$\boldsymbol{I}_2$	2 ×2 identity matrix
L_1	length of estimated vehicle contour
L_2	predicted length of estimated vehicle contour
$\boldsymbol{M}$	2 ×2 affine-matrix
M_t	matching index
N	number of regularly equal-spaced samples
N_B	number of *B*-spline functions
$\overline{\boldsymbol{n}}(s_i)$	normal unit vector of curve $\overline{\boldsymbol{r}}(s)$
$\boldsymbol{o}_{k+1}$	measurement state
$\boldsymbol{P}_{k+1}$	a posteriori error covariance
$\boldsymbol{P}_{k+1}^{-}$	a priori estimate error covariance
$\boldsymbol{Q}$	control point vector
$\boldsymbol{Q}_0$	control point vector of template curve
$\boldsymbol{Q}^x$	X coordinates vector of control points
$\boldsymbol{Q}^y$	Y coordinates vector of control points
$\boldsymbol{Q}_0^x$	X coordinates vector of the control points
$\boldsymbol{Q}_0^y$	Y coordinates vector of the control points
q_n^x	X coordinate of n*th* control point
q_n^y	Y coordinate of n*th* control point
$\boldsymbol{R}_k$	variance of η_k
$\boldsymbol{r}(s)$	coordinates vector of the control points
$\boldsymbol{r}_0(s)$	coordinates vector of template-curve control points
$\overline{\boldsymbol{r}}(s)$	best-fit curve of the vehicle contour
$\overline{\boldsymbol{r}}(s)$	contour of $\widetilde{\boldsymbol{X}}$
$\boldsymbol{r}_f(s)$	contour feature
$\boldsymbol{S}$	associated statistical information
$\boldsymbol{S}_i$	information matrix
S_ω	summation of ω_t^*
t	time
$\boldsymbol{u}_t$	translation vector
(u_0, v_0)	vanishing point of parallel lanes in image plane
(u_I, v_I)	image coordinate
V_{avg}	average travel speed
V_s	vehicle speed
$\boldsymbol{W}$	shape matrix
w	width of vehicle image
$\boldsymbol{X}$	shape-space vector
$\widetilde{\boldsymbol{X}}$	contour shape-space vector
$\widetilde{\boldsymbol{X}}$	predicted shape-space vector
$\overline{\boldsymbol{X}}$	shape-space vector of best-fit curve
$\boldsymbol{X}_k$	vector of position state
$\hat{\boldsymbol{X}}_k$	previous state
$\hat{\boldsymbol{X}}_{k+1}$	a posteriori estimate
x_t	current pixel value
x_{k+1}	position state
(X_w, Y_w)	world coordinate
$(x(s), y(s))$	vehicle contour
$\boldsymbol{Y}_k$	variance of ε_k
$\boldsymbol{Z}$	aggregated observation vector
$\boldsymbol{Z}_i$	information weight sum
α	regularization constant
β	learning rate
ε_k	system noise
$\boldsymbol{\eta}_k$	observation noise
ϕ	tilt angle of camera
μ_t	mean of Gaussian
σ_t	standard deviation of Gaussian
ω_t	weight of Gaussian
ω_t^*	updated weight of Gaussian

REFERENCES

Blake, A., and Isard, M., 1998, *Active Contours*, Springer, London, UK.

Bovik, A. C., Aggarwal, S. J., Merchant, F. A., Kim, N. H., and Diller, K. R., 2001, "Automatic Area and Volume Measurement from Digital Biomedical Images," D. P. Hader, ed, *Image Analysis: Methods and Applications*, 2nd ed., CRC, BocaRaton, FL, USA, pp. 23-64.

Bozic, S. M., 1994, *Digital and Kalman filtering*, Edward Arnold, London, UK.

Faro, A., Giordano, D., and Spampinato, C., 2008, "Evaluation of the Traffic Parameters in a Metropolitan Area by Fusing Visual Perceptions and CNN Processing of Webcam Images," *IEEE Transactions on Neural Networks*, Vol.19, No.6, pp.1108-1129.

Hsu, W. L., Liao, H. Y. M., Jeng, B. S., and Fan,

K. C., 2004, "Real-Time Traffic Parameter Extraction Using Entropy," IEE *Proceedings Vision Image and Signal Processing*, Vol. 151, No. 3, pp. 194-202.

Hu, W., Xiao, X., Xie, D., Tan, T., and Maybank, S. J., 2004, "Traffic Accident Prediction Using 3D Model Based Vehicle Tracking," *IEEE Transactions on Vehicular Technology*, Vol. 53, No. 3, pp. 677-694.

Iannizzotto, G., and Vita, L., 2000, "Real-Time Object Tracking with Movels and Affine Transformations," *International Conference on Image Processing*, Vancouver, BC, Canada, pp. 316-318.

Kamijo, S., Matsushita, Y., Ikeuchi, K., and Sakauchi, M., 2000, "Traffic Monitoring and Accident Detection at Intersections," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 1, No. 2, 2000, pp. 108-118.

Koschan, A., Kang, S. K., Paik, J. K., Abidi, B. R., and Abidi, M. A., 2002, "Video Object Tracking Based on Extended Active Shape Models with Color Information," *European Conference on Color Graphics Imaging Vision*, University of Poitiers, France, pp. 126-131.

Lai, A. H. S., and Yung, N. H. C., 2000, "Vehicle-Type Identification Through Automated Virtual Loop Assignment and Block-Based Direction-Biased Motion Estimation," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 1, No. 2, pp. 86-97.

Lan, L. W., Kuo, A. Y., and Huang, Y. C., 2003, "Color Image Vehicular Detection Systems with and without Fuzzy Neural Network: A Comparison," *Journal of the Chinese Institute of Engineers*, Vol. 26, No. 5, pp. 659-670.

Lim, D. W., Choi, S. H., and Jun, J. S., 2002, "Automated Detection of All Kinds of Violations at Street Intersection Using Real Time Individual Vehicle Tracking," *IEEE Southwest Symposium on Image Analysis and Interpretation*, Sante Fe, NM, USA, pp. 126-129.

Lin, C. T., Huang, Y. C., Mei, T. W., and Pu, H. C., 2006, "Multi-Objects Tracking System Using Adaptive Background Reconstruction Technique and Its Application to Traffic Parameters Extraction," *IEEE Conference on Systems*, Man, and Cybernetics, Taipei, Taiwan, pp. 2057-2062.

Masoud, O. N. P., and Kwon, E., 2001, "The Use of Computer Vision in Monitoring Weaving Sections," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 2, No. 1, 2001, pp. 18-25.

Pece, A. E. C., and Worrall, A. D., 2002, "Tracking with the EM Contour Algorithm," *Proceedings of the 7th European Conference on Computer Vision*, Copenhagen, Denmark, pp. 28-31.

Schoepflin, T. N., and Dailey, D. J., 2003, "Dynamic Camera Calibration of Roadside Traffic Management Cameras for Vehicle Speed Estimation," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 4, No. 2, pp. 90-98.

Song, K. T., and Tai, J. C., 2007, "Image-Based Traffic Monitoring with Shadow Suppression," *Proceedings of IEEE*, Vol. 95, No. 2, pp. 413-426.

Stauffer, C., and Grimson, W. E. L., 1999, "Adaptive Background Mixture Models for Real-Time Tracking," *IEEE Conference on Computer Vision and Pattern Recognition*, Fort Collins, CO, USA, pp. 246-252.

Tai, J. C., Tseng, S. T., Lin, C. P., and Song, K. T., 2004, "Real-Time Image Tracking for Automatic Traffic Monitoring and Enforcement Applications," *Image and Vision Computing*, Vol. 22, No. 6, pp. 485-501.

Vard, A. R., Nilchi, A. R. N., and Moallem, P., 2008, "Object Detetion and Image Segmentation Using Texture Pressure Energy in Parametric Active Contour Models," *Journal of the Chinese Institute of Engineers*, Vol. 31, No. 4, pp. 649-657.

Wu, J., Yang, Z., Wu, J., and Liu, A., 2007, "Virtual Line Group Based Video Vehicle Detection Algorithm Utilizing Both Luminance and Chrominance," *IEEE conference on Industrial Electronics and Applications*, Harbin, China, pp. 2854-2858.

Yilmaz, A., Javed, O., and Shah, M., 2006, "Object Tracking: a Survey," *ACM Computing Surveys*, Vol. 38, No. 4, Article 13. pp. 1-45.

Manuscript Received: Sep. 23, 2008
Revision Received: Jun. 27, 2009
and Accepted: Jul. 27, 2009