Proceedings of the 11th
KACC, October 1996

Target Identification for Visual Tracking

°Joon-Woong Lee*, Joo-Seop Yun**, In-So Kweon**

*KIA Technical Center, KIA Motors Co., 781-1, Soha-Dong, Kwang Myung-Shi, Kyunki-Do, KOREA
Tel: +82-2-801-2049; Fax: +82-2-801-2400
**Department of Automation and Design Eng., Korea Advanced Institute of Science and Technology,
207-43, Cheongryangri-Dong, Dongdaemoon-Ku, Seoul, KOREA
Tel: +82-2-958-3465; Fax: +82-2-968-1638 E-mail: kweon@design.kaist.ac.kr

Abstract In moving object tracking based on the visual sensory feedback, a prerequisite is to determine which feature or which object is to be tracked and then the feature or the object identification precedes the tracking. In this paper, we focus on the object identification not image feature identification. The target identification is realized by finding out corresponding line segments to the hypothesized model segments of the target. The key idea is the combination of the *Mahalanobis* distance with the geometrical relationship between model segments and extracted line segments. We demonstrate the robustness and feasibility of the proposed target identification algorithm by a moving vehicle identification and tracking in the video traffic surveillance system over images of a road scene.

Keywords Target identification, *Mahalanobis* distance, geometrical relationship, primitive matching

1. Introduction

Visual tracking of an object moving in a 3D space represents one of the central issues in dynamic scene analysis. In this paper, we mainly concerned on how to identify the moving target in the image domain to realize the visual tracking.

A great number of tracking algorithms have been proposed in the recent years. Deriche and Faugeras[1] tracked line segments along the image sequence in a stationary environment. They used the *Mahalanobis* distance as a similarity function for line matching between inter frame images. Koller *et al.*[2] used a model-based object tracking in monocular image sequences of road traffic scenes for gaze control. They defined five different 3-D generic-vehicle models, limited that motion takes place on a planar road and assumed a simple circular motion with constant velocity. Meyer and Bouthemy[3], and Bouthemy and Francois[4] proposed an elegant algorithm of motion-based region segmentation for the detection and discrimination of the target with a static camera for a dynamic scene and of tracking by pursuit of the extracted region based on the affine motion model. Smith and Brady[5] demonstrated moving objects tracking using feature-based optical flow computation in moving platform situation. They found corresponding corner points over several successive image frames, computed the optical flow between these corresponding points, and then clustered points with a similar optical flow. Kass *et al.*[6] presented a tracking method by active contour models, that is, snakes. Once the snake is interactively initialized on an object contour in the initial frame, it tracks the contour in subsequent frames by finding a local minimum of an energy function.

We establish a paradigm for a moving target identification by combining the *Mahalanobis* distance which analyzes geometric attributes statistically with the geometrical relationship between line segments of model and data. The former is based on sound statistical principles rather than on some *ad hoc* heuristics for image primitive matching problems, while the later well describes the topological relationships of matching primitives. Our method has several advantages such that it works irrespective of small or large motion, processes only single image frame, identifies the object of interest in the existence of same objects or multiple objects and works even under motion discontinuity and partial occlusion. To locate the identified target on an anticipated position in the image plane in the next frame is performed based on the motion model of the target. However, the motion parameter estimation is not within the scope of this paper. And we assume that the self-occlusion of the viewed surface of the object of interest is not occurred.

We demonstrate the robustness and feasibility of the proposed target identification algorithm by identifying and tracking a moving vehicle in the video traffic surveillance system over the road images.

2. Model-based target identification

In this section, we describe an algorithm identifying the object of interest among the line segments(data set) extracted by [7]. The object of interest is identified by selecting line segments best matched to model segments through the matching process. The matching process starts with the computation of the *Mahalanobis* distance[8] as the similarity measure between model segments and data segments. A matching candidate set(MCS) for each model segment by selecting line segments with the *Mahalanobis* distance smaller than a threshold value is then constructed. Finally, the data segment among line segments in the MCS well-matched to each model segment is selected by utilizing the geometrical relationship between the model segment and a data segment.

2.1 Matching function

Matching between extracted line segments and model line segments is carried out based on a statistical point of view. Here, we define two sets such as the data set of q data segments and the model set of n model segments:

$$D=\{d_j\}\ j=0,\dots,q-1,\quad M=\{m_i\}\ i=0,\dots,n-1. \tag{1}$$

Let $x=[x_1,x_2,\dots,x_n]^T$ be a random vector of n - Gaussian random variables: each with a mean μ_i and a variance σ_i^2. The probability density function of the random vector **x** is given by

$$f(x)=\frac{1}{(2\pi)^{n/2}\sqrt{|\Sigma|}}\exp\left[-\frac{1}{2}(x-\mu_x)^T\Sigma^{-1}(x-\mu_x)\right] \tag{2}$$

with the mean vector μ_x and the covariance matrix Σ. The quantity in the bracket is called the squared *Mahalanobis* distance[8] from **x** to μ_x and denoted as r^2. We can notice that minimizing r^2 would maximize $f(x)$. It is well known that r^2 follows χ_n^2 distribution [9]. We define a vector as

$$v=\tilde{y}_{m_i}-\tilde{y}_{d_j},\quad i\in M \text{ and } j\in D \tag{3}$$

where $\tilde{y}_m$ and $\tilde{y}_d$ are feature vectors of line segments in the model and data set, respectively. Assuming $\tilde{y}_m$ and $\tilde{y}_d$ are independent, then the matching function is formed as:

$$r_{ij}^2=v^T\Sigma_{vv}^{-1}v \\ \text{where } E[v]=0,\ \text{and } \Sigma_{vv}=\Sigma_{y_dy_d}+\Sigma_{y_my_m}. \tag{4}$$

The matching function r_{ij}^2 of a vector **v** has a χ_1^2 distribution.

2.2 Uncertainty modeling

2.2.1 Data segment

Uncertainty modeling of a data segment is done with the similar approach as in Deriche and Faugeras[1]. Let $\widetilde{y}_d = [\bar{x}\ \bar{y}\ \theta\ l]^T$ and $\widetilde{x}_d = [x_s y_s x_e y_e]^T$ be two random vectors. $\widetilde{y}_d$ is expressed as a vector function of $\widetilde{x}_d$ as follows;

$$\widetilde{y}_d = \begin{bmatrix} \frac{x_s + x_e}{2} \\ \frac{y_s + y_e}{2} \\ \tan^{-1}\frac{y_e - y_s}{x_e - x_s} \\ \sqrt{(x_e - x_s)^2 + (y_e - y_s)^2} \end{bmatrix} \tag{5}$$

Covariance of random vector $\widetilde{x}_d$ is

$$\Sigma_{x_d x_d} = \begin{bmatrix} \sigma^2_{x_s} & \sigma_{x_s y_s} & \sigma_{x_s x_e} & \sigma_{x_s y_e} \\ \sigma_{x_s y_s} & \sigma^2_{y_s} & \sigma_{y_s x_e} & \sigma_{y_s y_e} \\ \sigma_{x_s x_e} & \sigma_{y_s x_e} & \sigma^2_{x_e} & \sigma_{x_e y_e} \\ \sigma_{x_s y_e} & \sigma_{y_s y_e} & \sigma_{x_e y_e} & \sigma^2_{y_e} \end{bmatrix} \tag{6}$$

According to the error propagation[10], the covariance of $\widetilde{y}_d$ is defined as

$$\Sigma_{y_d y_d} = J_{y_d x_d} \Sigma_{x_d x_d} J^T_{y_d x_d} \tag{7}$$

where $J_{y_d x_d}$ is the Jacobian matrix of $\widetilde{y}_d$ for $\widetilde{x}_d$. Assuming no correlation between the end points then

$$\Sigma_{p_s p_s} = \begin{bmatrix} \sigma^2_{x_s} & \sigma_{x_s y_s} \\ \sigma_{x_s y_s} & \sigma^2_{y_s} \end{bmatrix} = \Sigma_{p_e p_e} = \begin{bmatrix} \sigma^2_{x_e} & \sigma_{x_e y_e} \\ \sigma_{x_e y_e} & \sigma^2_{y_e} \end{bmatrix}, \tag{8}$$

and the same covariance matrix for both end points, that is, $\Sigma_{p_s p_s} = \Sigma_{p_e p_e}$ leads to

$$\Sigma_{x_d x_d} = \begin{bmatrix} \Sigma_{p_s p_s} & 0 \\ 0 & \Sigma_{p_e p_e} \end{bmatrix}. \tag{9}$$

Hence, Eq.(7) becomes

$$\Sigma_{y_d y_d} = \begin{bmatrix} 0.5\sigma^2_x & 0.5\sigma_{xy} & 0 & 0 \\ 0.5\sigma_{xy} & 0.5\sigma^2_y & 0 & 0 \\ 0 & 0 & A & B \\ 0 & 0 & C & D \end{bmatrix} \tag{10}$$

where

$$A = \frac{2}{l^4}\left[\sigma^2_x (y_e - y_s)^2 + \sigma^2_y (x_s - x_e)^2 + 2\sigma_{xy}(x_s - x_e)(y_e - y_s)\right],$$

$$B = \frac{2}{l^3}\left[(x_s - x_e)(y_e - y_s)(\sigma^2_x - \sigma^2_y) + \sigma_{xy}((x_s - x_e)^2 - (y_e - y_s)^2)\right],$$

$$C = \frac{2}{l^2}\left[\sigma^2_x (x_s - x_e)^2 + \sigma^2_y (y_e - y_s)^2 - 2\sigma_{xy}(x_s - x_e)(y_e - y_s)\right].$$

2.2.2 Model segment

A random vector $\widetilde{x}_m$ is a vector for representing end point of a line segment in the model set. For a model segment we also assume no correlation between the end points. The error covariance of $\widetilde{x}_m$ is

$$\Sigma_{x_m x_m} = \begin{bmatrix} v_x & v_{xy} & 0 & 0 \\ v_{xy} & v_y & 0 & 0 \\ 0 & 0 & v_x & v_{xy} \\ 0 & 0 & v_{xy} & v_y \end{bmatrix} \tag{11}$$

where v_x and v_y are variances, and v_{xy} is the covariance of x and y coordinates. Actual error covariance **P** of the state estimation in the motion model of the target is used as these variances. Then, the uncertainty of $\widetilde{y}_m$ is derived as a same way of a data segment:

$$\Sigma_{y_m y_m} = J_{y_m x_m} \Sigma_{x_m x_m} J^T_{y_m x_m}. \tag{12}$$

2.3 Line matching

We compute the squared *Mahalanobis* distance r^2_{ij} of Eq.(4) based on two uncertainty models of Eqs.(7) and (12). For each model segment we establish an MCS provided that the r^2_{ij} is less than a threshold value d_{χ^2}. That is,

$$MCS_i = \{j \mid j \in D,\ r^2_{ij} < d_{\chi^2}\},\ i \in M \tag{13}$$

where the threshold d_{χ^2} is determined from χ^2 distribution table considering the 100(1-α)% confidence interval of statistics[9].

We propose two methods to find out the data segment matched to the *i*-th model segment.

2.3.1 Line matching by minimum value criterion

A best matching data segment *j* to the model segment *i* is determined as a data segment with minimum r^2_{ij} as

$$\forall i \in M,\ Z_i = \{j \mid j \in MCS_i,\ r^2_{ij} < r^2_{ik},\ \forall k \in MCS_i,\ k \neq j\}. \tag{14}$$

Even though this method is simple, it has two disadvantages. First, it can not prevent miss matching due to the motion discontinuity between the predicted motion and the real motion taken place on the way of tracking. Second, this method can not identify the occurrence of occlusion. Therefore, we do not use this method.

2.3.2 Line matching by geometrical relationship

We overcome limitations of the minimum value criterion by utilizing the geometrical relationship between line segments. The procedure to select the best matched data segment to a model segment m_i where $i \in M$ is as following.

(1) Model transformation

We transform the model to a data segment d_k where $k \in MCS_i$ based on the geometrical relationship between the d_k and the m_i.

$$P_1 = \mathrm{Trans}(-x_{s_i}, -y_{s_i})P,\quad P = \underbrace{\begin{bmatrix} x_s & x_e & \cdots \\ y_s & y_e & \cdots \\ 1 & 1 & \cdots \end{bmatrix}}_{2n},$$

$$P_2 = \mathrm{Rot}(z, \theta_{ik})P_1, \tag{15}$$

$$P_3 = \mathrm{Trans}(\Delta_x, \Delta_y)P_2,$$

where **P** is matrix formed with end points of model segments, (x_{s_i}, y_{s_i}) is the start point of m_i, θ_{ik} is a difference in orientation between d_k and m_i as shown in Fig. 1(a). Δ_x and Δ_y are differences in displacement between mid points of d_k and transformed m_i in P_2 as shown in Fig. 1(c).

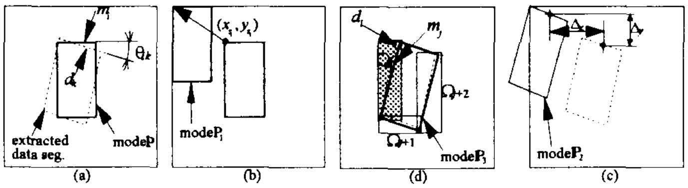


Fig. 1. Model transformation

(2) Building a search region for data segments

For a transformed model segment m_i, we build a data segment search region Ω_i as shown in Fig. 1(d) based on the estimation of 100(1-α)% confidence interval[9].

$$\Omega_i = (p^0_i, p^1_i),\ j \in M \text{ and } j \neq i,$$

$$p^0_i = \{\min(p^s_i(x), p^e_i(x)) - Z_{\alpha/2} \times \sqrt{v_x},\ \min(p^s_i(y), p^e_i(y)) - Z_{\alpha/2} \times \sqrt{v_y}\}, \tag{16}$$

$$p^1_i = \{\max(p^s_i(x), p^e_i(x)) + Z_{\alpha/2} \times \sqrt{v_x},\ \max(p^s_i(y), p^e_i(y)) + Z_{\alpha/2} \times \sqrt{v_y}\}$$

where $(p^s_i(x), p^s_i(y))$ and $(p^e_i(x), p^e_i(y))$ are start and end points of m_i in P_3, v_x and v_y are uncertainties of a model segment of Eq.(11), and $Z_{\alpha/2}$ is taken in standard normal distribution table.

(3) Matching quality and best matching candidate set

For a transformed model segment m_i, we check whether a data segment d_l where $l \in MCS_i$ satisfies the following four conditions:

- Existence condition

$$Exist(j,l) = \begin{cases} 1 & \text{if } MBR_{d_l} \cap \Omega_j \neq \phi, j \in M \text{ and } j \neq i \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

where MBR_{d_l} is the minimum building rectangle encompassing the d_l as shown in Fig. 2.

- Distance condition

$$Dist(j,l) = \begin{cases} 1 & \text{if } dist(m_j, p_{int}) \leq Z_{\alpha/2}\sqrt{v_x + v_y} \\ 0 & \text{otherwise} \end{cases} \tag{18}$$

where p_{int} is the intersection point between $x = \min(p_l^s(x), p_l^e(x)) - Z_{\alpha/2} \times \sqrt{v_x}$ defined in Eq.(16) and m_j and $dist(m_j, p_{int})$ is distance between m_j and p_{int} as shown in Fig. 2.

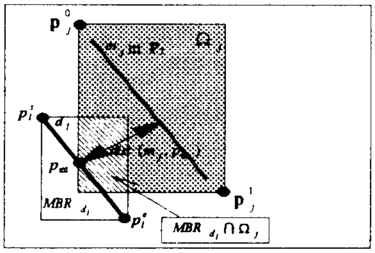


Fig. 2. The existence of a data segment, d_l in Ω_j

- Angle condition

$$Ang(j,l) = \begin{cases} 1 & \text{if } |\angle m_j - \angle d_l| < \delta \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

where δ is the minimum angle for which these two lines are corresponding.

- Length condition

$$Len(j,l) = \begin{cases} 1 & \text{if } |len(m_j) - len(d_l)| < \varepsilon \\ 0 & \text{otherwise} \end{cases} \tag{20}$$

where $len(m_j)$ and $len(d_l)$ are the line lengths of m_j and d_l, respectively. ε is considered as the minimum length which does not exceed if these two lines are corresponding.

We define a geometrical relationship between m_j and d_l as

$$Geo(j,l) = Exist(j,l) + Dist(j,l) + Ang(j,l) + Len(j,l). \tag{21}$$

When $Geo(j,l)$ equals to four, we consider the *matching quality* is satisfied. For all model segments except for m_i, we investigate the geometrical relationship of Eq.(21) and check for the existence of the data segment satisfying the matching quality. If there is a data segment satisfying the matching quality we formulate the best matching candidate set, $\Re_{ik}$ as;

$$\forall i \in M, \exists k \in MCS_i \quad \Re_{ik} = \{j \mid A[j], j \in M, j \neq i\}$$
$$\text{where } A[j] = \begin{cases} l & \text{if } Geo(j,l) = 4,\ l \in MCS_i \\ -1 & \text{elsewhere} \end{cases} \tag{22}$$

(4) Line matching

We define a magnitude of the matching candidate set as $mag(\Re_{ik})$ = number elements which $A[j] \neq -1$, $A[j] \in \Re_{ik}$. We search for the $\Re_{ik}$ with a maximum magnitude and let $\Re$.

2.4 Matching condition

If an $A[j]$ in $\Re$ makes a set $\Re_{jA[j]}, j \neq i$, and an element $A[j], j = j$ in $\Re_{jA[j]}$ equals to $A[j]$ of $\Re$, $\Re$ becomes the best matching set, $\Re_{pk}$.

When there exist multiple sets satisfying the matching condition we choose a set as the best matching set as following. Let one of them be $\Re_{pk}$ and compute the sum, $\wp$ of the *Mahalanobis* distance, r_{pk}^2 and the *Mahalanobis* distances of elements in $\Re_{pk}$, $r_{jA[j]}^2$ where $A[j] \neq -1$. That is

$$\wp = r_{pk}^2 + \sum_j r_{jA[j]}^2,\ A[j] \in \Re_{pk} \text{ and } A[j] \neq -1. \tag{23}$$

The set, $\Re_{pk}$ with the minimum $\wp$ becomes the best matching set. After the $\Re_{pk}$ is found out the line correspondence between model and data is realized as

$$\begin{aligned} & d_k \equiv m_p, \\ & d_{A[j]} \equiv m_j \text{ if } A[j] \in \Re_{pk} \text{ and } A[j] \neq -1 \end{aligned} \tag{24}$$

2.5 Line matching under occlusion

We analyze the following two cases for matching lines when occlusions exist.

1) Case 1

When the MCS of a model segment, m_i is null, that is, $MCS_i = \{\phi\}$ for a $i \in M$ or $A[i] = -1$ in the best matching set, $\Re_{pk}$ we conclude that a partial occlusion is occurred. If the partial occlusion is recognized, the model transformation is carried out as:

$$\begin{aligned} & P' = Rot(z, \theta)P \\ & P'' = Trans(\delta_x, \delta_y)P' \end{aligned} \tag{25}$$

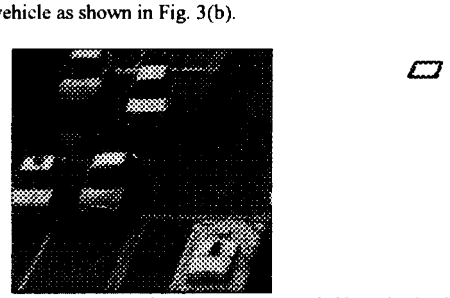
where **P** is the same matrix in Eq. (15), θ is average difference in orientation between model segments and their corresponding data segments of Eq. (24), and δ_x, δ_y are average differences in displacement between mid points of matched data segments and their corresponding model segments in P'. Then, the *i*-th transformed model segment in P'' is considered as the data segment corresponding to the *i*-th unmatched model segment, m_i.

2) Case 2

When $MCS_i = \{\phi\}$ for $\forall i \in M$ or there is no $\Re_{pk}$ we conclude that the full occlusion is occurred. In full occlusion, each segment in the predicted shape model is considered as the matched data segment.

3. Experimental results

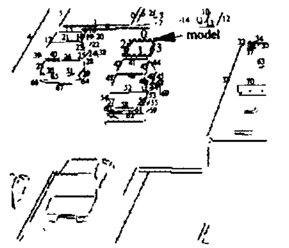
In this experiment, we want to track the vehicle surrounded by an ellipse shown in Fig. 3(a) without the motion model of the vehicle. Initially, we set a simple model composed of four line segments which represent the roof of the vehicle as shown in Fig. 3(b).



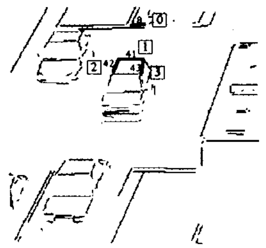
(a) Raw image (b)Hypothesized model

Fig. 3. Image of road vehicle and its model

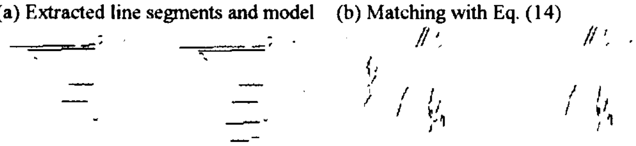
Target identification is carried out by finding corresponding line segments of given model segments and the model is updated by intersection points of corresponding line segments. This updated model becomes the new model to the next image frame. To explain the target identification process based on the proposed algorithm in detail, we take the image frame #2 as an example image. The extracted line segments and target model of the example image are shown in Fig. 4(a). The MCS corresponding to each model segment obtained using Eq. (12) are MCS_0 ={1,3,7,**8**,9,14,41,46,50}, MCS_1 = {1,3,7,8,9,14,16,19,**41**,46,50,52,56,58}, MCS_2 ={0,2,6,7,18,22,23,24,25, 28,29,**42**,43,45,47,49,50,53,68,69}, MCS_3 ={0,2,6,7,14,42,**43**,45,4,48,49, 53,68,69}. These MCSs are shown in Fig. 4 (c), (d), (e) and (f).



(a) Extracted line segments and model



(b) Matching with Eq. (14)



(c) MCS_0 (d) MCS_1 (e) MCS_2 (f) MCS_3

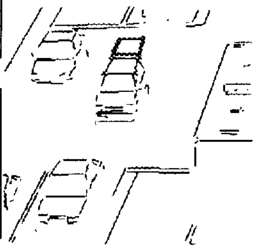
model	data	model	data
0	1	0	
		1	-1
		2	18
		3	-1
	3	0	
		1	-1
		2	0
		3	-1
	41	0	
		1	46
		2	42
		3	-1
	46	0	
		1	52
		2	-1
		3	-1
	50	0	
		1	-1
		2	-1
		3	69

model	data	model	data
1	8	0	-1
		1	
		2	0
		3	-1
	9	0	-1
		1	
		2	-1
		3	0
	46	0	41
		1	
		2	42
		3	43
	52	0	46
		1	
		2	47
		3	-1

model	data	model	data
2	42	0	41
		1	46
		2	
		3	-1
	47	0	46
		1	52
		2	
		3	-1

model	data	model	data
3	42	0	-1
		1	-1
		2	24
		3	
	43	0	41
		1	46
		2	42
		3	

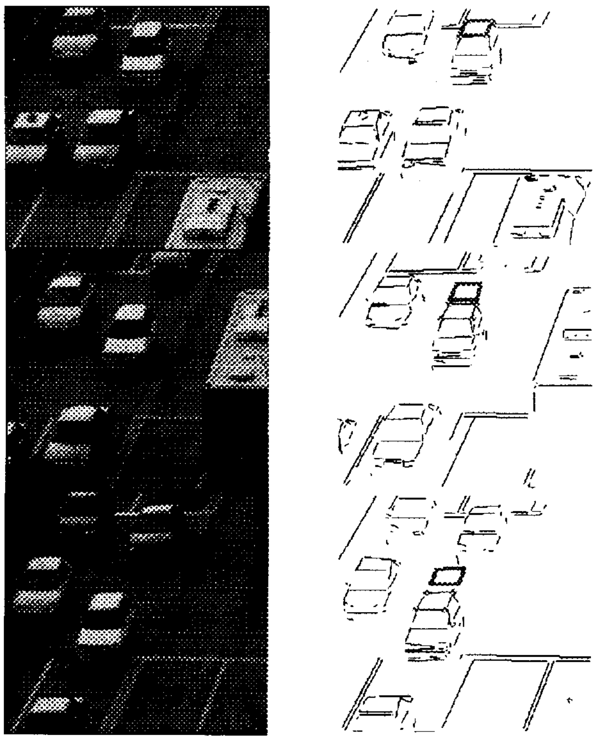
(g) Best matching candidate sets



(h) Matching result

Fig. 4. Target identification process in road vehicle image

The minimum *Mahalanobis* distance to each model segment is the data segment 8 for model segment 0, the data segment 41 for model segment 1, the data segment 42 for model segment 2, and the data segment 43 for model segment 3. This result is presented in Fig. 4(b) and can be easily notified that this result fails to identify the target. Fig. 4(g) is another representation of the best matching candidate set. $\Re_{3\ 43} = \{41, 46, 42\}$ is an example and this set satisfies the matching condition. Fig. 4(h) shows the exact matching result by the proposed method. We show the target identification for a series of images in Fig. 5.



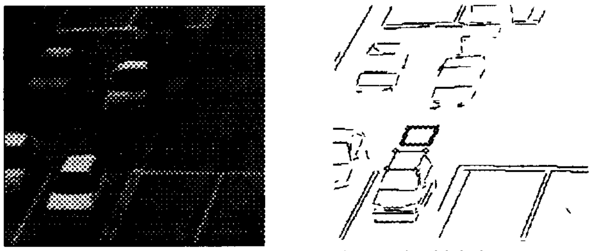


Fig. 5. Target identification in successive road vehicle images of frame 1 to frame 4

4. Conclusion

We developed a target identification algorithm to track a target moving in 3-D space. In fact the target identification results in the correspondence problem of finding out corresponding line segments among extracted line segments to the hypothesized model of the target. We establish the successful paradigm for a moving target identification according to the integration of the *Mahalanobis* distance which analyzes geometric properties statistically to the geometrical relationship between line segments of model and data. The *Mahalanobis* distance as the measure for the statistical analysis doesn't guarantee to provide the best matching result in the motion discontinuity, which normally takes place by an unpredictable motion change, or occlusion. We overcome these limitations by introducing the matching quality of geometrical relationship. The proposed algorithm has two advantages. First, it can prevent the miss matching or multiple matching due to the motion discontinuity between the predicted motion and the real motion. Second, this method can detect the occurrence of occlusion and solve the occlusion problem.

We showed the robustness and feasibility of the proposed target identification algorithm by experiments. The algorithm has successfully identified the target even though the partial occlusion or abrupt motion change exist. We expect this work will contribute in many areas of computer vision such as motion analysis, scene analysis, stereo vision and so on.

References

[1] R. Deriche and O. Faugeras, Tracking line segments, *Image and Vision Compt.*, Vol. 8, No.4, pp.261-270, 1990.

[2] D. Koller, K. Daniilidis and H.H. Nagel, Model-Based Object Tracking in Monocular Image Sequences of Road Traffic Scenes, *IJCV*, Vol. 10, No.3, pp.257-281, 1993.

[3] F.G. Meyer and P.Bouthemy, Region-Based Tracking Using Affine Motion Models in Long Image Sequences, *CVGIP:Image Understanding*, Vol.60, No. 2, pp.119-140, 1994.

[4] P.Bouthemy and E.Francois, Motion Segmentation and Qualitative Dynamic Scene Analysis from an Image Sequence, *IJCV*, Vol.10, No.2, pp.157-182, 1993.

[5] S.M. Smith and J.M. Brady, ASSET-2: Real time motion segmentation and shape tracking, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 17, No.8, pp.814-820,1995.

[6] M.Kass, A.Witkin, and D.Terzopoulos, Snakes: Active Contour Models, *IJCV*, pp.321-331,1988.

[7] I.S. Kweon and J.W. Lee, Extraction of Line Features in Noisy Images, *Pattern Recognition*, Currently under review.

[8] O. Duda and P.E. Hart, Pattern Classification and Scene Analysis, John Wiley & Sons, U.S.A. 1973.

[9] V.K. Rohatgi, An Introduction to Probability Theory and Mathematical Statistics, Wiley, 1976

[10] E.M. Mikhail, Observations and Least Squares, University press of America, U.S.A. 1976.