SEGMENT MODEL BASED VEHICLE MOTION ANALYSIS

Pengfei Zhu, Weiming Hu, Xi Li, Li Li
National Laboratory of Pattern Recognition, Institute of Automation
Chinese Academy of Sciences
Beijing, China
Email: {pfzhu, wmhu, lixi, lli}@nlpr.ia.ac.cn

***Abstract*—Motion analysis is a very attractive research direction in computer vision field. In this paper, we propose a framework for analyzing real vehicle motion in visual traffic surveillance by using Segment Model (SM), which is a kind of probabilistic model. SM can grasp the underlying information of observation sequence by using segment distribution. It has been proved to be more precise than that of HMM. In the experiments, we compare our approach with the template matching method based on the Hausdorff distance and the state space method based on the Hidden Markov Model (HMM). The experimental results show the effectiveness of our approach.**

***Keywords*-Segment Model; Motion Analysis; Hausdorff Distance; HMM;**

I. INTRODUCTION

Visual motion analysis is a challenging research field in computer vision field, it has attracted great interests from computer vision researchers due to its promising applications in many areas, for example in visual surveillance.

Hidden Markov Model (HMM) [3] [2] is very popular in temporal sequence analysis and behavior recognition. In HMM, the signal non-stationarity is handled by a Markov chain which modulates the parameters of an output stochastic process. But HMM implementations rely on several hypotheses: the signal is assumed stationary within each state, the order of the Markov chain is usually set to one and observations are considered independent. However, these assumptions are far from reality. This limits the ability of HMM to mine the relations within a segment, and it is not enough to represent a non-stationary observation sequence by a piecewise state sequence.

Several attempts have been made in order to handle these problems. Segment Model (SM) is put forward in speech recognition field [9].

In SM, the observations at the state level are a sequence of observations, called segment, rather than a frame feature vector. Therefore, each segment in a SM defines a duration model that accounts for the segment length and an emission probability distribution of a sequence. SM can be considered as generalized HMM.

In this paper, a novel framework based on Segment Model is proposed to analyze vehicle motion trajectories.

Figure 1 shows the overview of our approach. Two parts are included in:

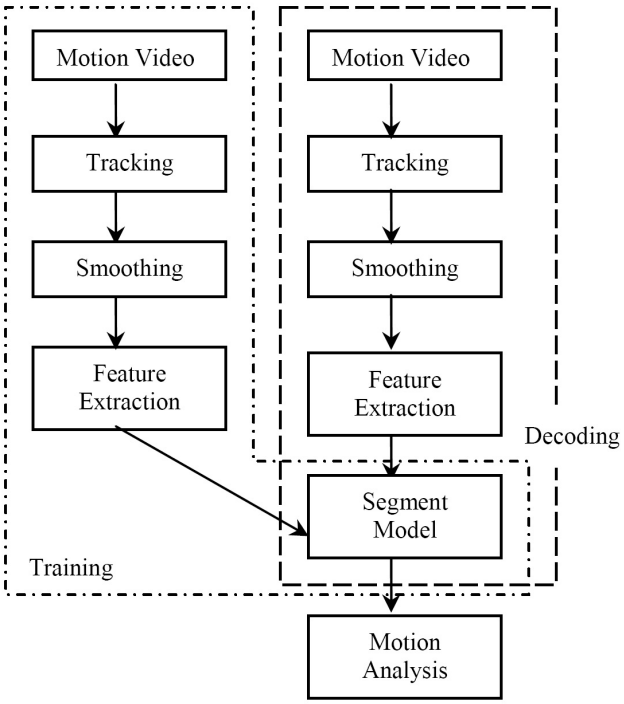

Figure 1. Overview of our approach.

1) Model Training: Given a vehicle motion video, the Kalman Filter based tracking algorithm [4] is used to extract vehicle motion trajectory. The Wavelet Multiresolution Analysis algorithm [6] is used to smooth the trajectory. Then we extract the features to train SM.

2) Model Decoding: In this part, the front steps are similar to 1), only the last step is using the Viterbi decoding algorithm [7] to decode.

To evaluate our framework, we compare our approach with the template matching method based on the Hausdorff distance and the state space method based on the Hidden Markov Model in the experiments.

This paper is organized as follows. In section 2, we introduce Segment Model in a nutshell. Smoothing for motion trajectories are shown in Section 3. Section 4 presents the approach for feature extraction. Section 5 shows the algorithms of training and recognition. Our experiments

978-0-7695-3718-4/09 $25.00 © 2009 IEEE
DOI 10.1109/AVSS.2009.9

Authorized licensed use limited to: INSTITUTE OF AUTOMATION CAS. Downloaded on December 3, 2009 at 03:41 from IEEE Xplore. Restrictions apply.

results are shown in section 6. Section 7 is the conclusion of our work.

II. Segment model

Segment Model [7] is a kind of probabilistic model, which is regarded as generalized HMM.

In SM, each segment generates a random length subsequence of observations:

$$p(o_1^d, d|s) = p(o_1^d|d, s)p(d|s) \tag{1}$$

where d is the possible observation length, s is the label of a segment, o_1^d describes the observation sequence with different length.

A segment is thus defined by the segment label and its duration instead of a state label s in a HMM.

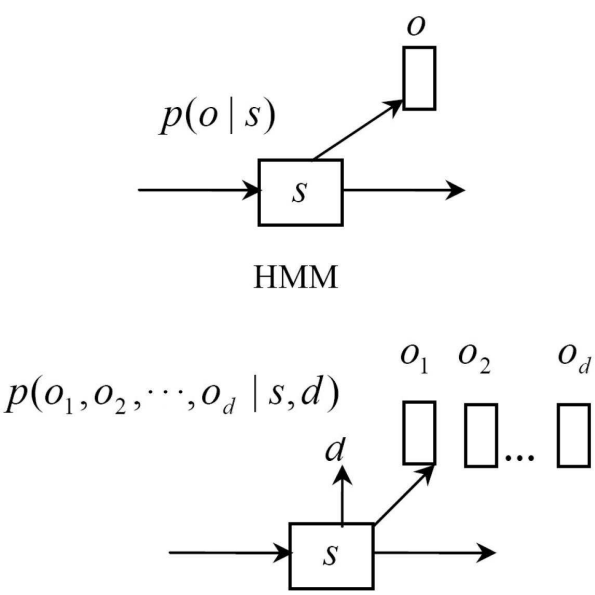

Figure 2. Generative processes of a HMM and a SM.

We can see the key idea of SM by comparison between HMM and Segment Model. SM models the signal at a segment level rather than at the observation level like HMM. An observation sequence is assumed to be generated by a succession of SM states or segments, each being responsible for a subsequence. We can see the difference from the perspective of generative models between HMM and SM in Figure 2. One frame is generated by an HMM state, while a variable-length frames sequence is generated by a segment associated with random length.

III. Trjectory smoothing

In this paper, the vehicle motion trajectories are obtained by the tracking algorithm based on the Kalman Filter [4]. This is not the focus of this paper, we don't discuss in detail. Before analyzing these motion trajectories, we have to smooth these trajectories. The Wavelet Multiresolution Analysis [6] [5] is used here.

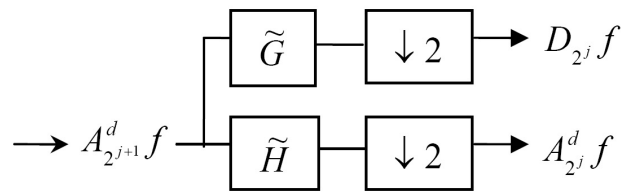

Figure 3. Wavelet Decomposition Algorithm.

Figure 3 shows the decomposition algorithm from the $j-th$ layer to the $j+1-th$ layer. $f(x)$ denotes the original signal, $A^d_{2^j} f$ denotes the approximation of $f(x)$ at 2^j resolution, $D_{2^j} f$ denotes the detail of signal $f(x)$. $\tilde{G}$ and $\tilde{H}$ are a band-pass filter and a low-pass filter respectively. $\downarrow 2$ denotes to keep one sample out two.

From figure 3, we can see that the orginal signal is decomposed to two signals: one is approximation part of signal, another is detail part of signal. The approximation part can be regarded as the low frequency part, the detail part can be regarded as the high frequency part. In our experiments, two-layer Wavelet Multiresolution algorithm is used, and we choose the low frequency part as our wanted smoothing data.

IV. Feature extraction

In this stage, we extract the features of smoothing motion trajectories. One trajectory can be expressed as a series of coordinates value:

$$Traj = \{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\} \tag{2}$$

where n denotes the $n-th$ frame of the motion.

Then calculate the δx, δy, s ,d of point (x, y) respectively:

$$\begin{cases} \delta x_i = x_{i+1} - x_i, \\ \delta y_i = y_{i+1} - y_i, \\ s_i = \sqrt{\delta x_i^2 + \delta y_i^2}, \\ d_i = tan^{-1}(\delta y_i / \delta x_i), \quad \text{if } \delta x_i < 0,\ d_i = d_i + \pi \end{cases} \tag{3}$$

Then the feature vector of point (x_i, y_i) can be expressed by

$$o_i = \{\delta x_i, \delta y_i, s_i, d_I\}. \tag{4}$$

One motion trajectory can be represented by

$$Traj = \{o_1, o_2, \cdots, o_n\}. \tag{5}$$

Authorized licensed use limited to: INSTITUTE OF AUTOMATION CAS. Downloaded on December 3, 2009 at 03:41 from IEEE Xplore. Restrictions apply.

V. Training and recognition

A. Segment model training

For training SM, two steps are included in:

1) Motion Feature Modeling

In HMM, the Baum-Welch algorithm is used to estimate parameters. However, because the transfer probability between states needn't be estimated, the classical EM algorithm [1] of GMM is directly used to estimate parameters of SM.

2) Segment Feature Modeling

Segment feature modeling is an very important part in SM. Duration information is used usually. A kind of probability distribution based modeling method has been used in duration modeling. Here we select the gamma ditribution.

B. Motion trajectory recognition

For recognition, we want to find the likelihood of a sequence of observations given a motion model. Two jobs are included in: searching the optimum segmentation of motion sequence and finding out the maximum likelihood matching to every supposed segments.

The Viterbi decoding algorithm [7] is used here. Segment-based recognition then involves finding

$$(\hat{T}, \hat{s}_1^{\hat{T}}) = \underset{T, s_1^T}{argmax}\{\underset{d_1^T}{max} p(o_1^N | d_1 T, s_1^T) p(d_1^T | s_1^T) p(s_1^T)\} \tag{6}$$

where N is the observation sequence length, T is the number of segments, s_1^T is the segment labels and d_1^N is the segment durations.

VI. Experiments

To evaluate the performance of our approach for analyzing motion trajectories, a database included 2000 vehicle motion trajectories extracted from real traffic video are used here.

A. Description of motions

In our experiments, 7 kinds of motions have been included in. The experiment scene and these motions have been showed in Figure 4.

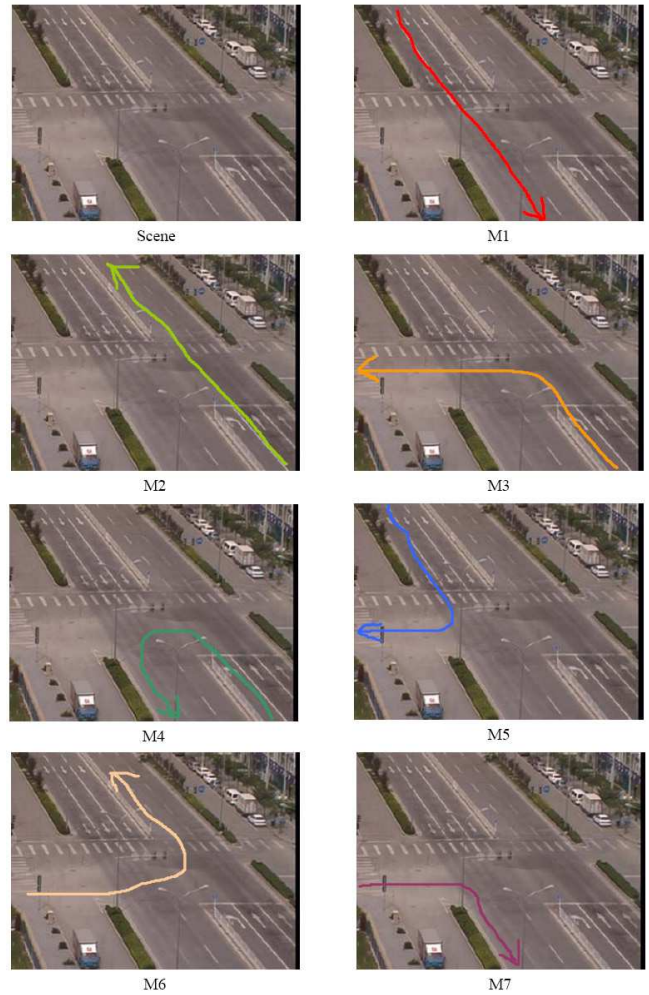

Figure 4. The experiment scene and the categories of motion trajectories.

B. Experimental results

To validate the effectiveness of our approach, we compare our method with the other methods:

(1) Template Matching Method

Here, we use the template matching based on the Hausdorff distance.

Assume that $T_1(t_1)$ and $T_2(t_2)$ are two motion trajectories, where t_1 and t_2 are the durations of the two motion trajectories respectively. So the Hausdorff distance is

$$d(T_1, T_2) = \underset{1 \le i \le t_1}{mean}(\underset{1 \le j \le t_2}{min}(|T_1(i) - T_2(j)|)) \tag{7}$$

$$\tilde{d}(T_1, T_2) = max\{d(T_1, T_2), d(T_2, T_1)\} \tag{8}$$

So the motion trajectory classification is

$$c = \underset{i}{argmin} \tilde{d}(T, M_i) \tag{9}$$

where T is a test trajectory, M_i is the $i-th$ class trajectory.

(2) State Space Method

Hidden Markov Model [8] , which is a very popular probabilistic model, is used here.

For the HMMs, two steps are included: parameter training of HMM and HMM-based recognition.

In the training stage, the parameters of the HMM are initialized to random values and the Baum-Welch algorithm is used to estimate the parameters iteratively using the forward-backward procedure.

In the recognition stage, the trajectory recognition is

Authorized licensed use limited to: INSTITUTE OF AUTOMATION CAS. Downloaded on December 3, 2009 at 03:41 from IEEE Xplore. Restrictions apply.

$$c = \underset{i}{argmax} P(T, X_i) \tag{10}$$

where T is the test trajectory and X_i is the $i-th$ class of HMMs.

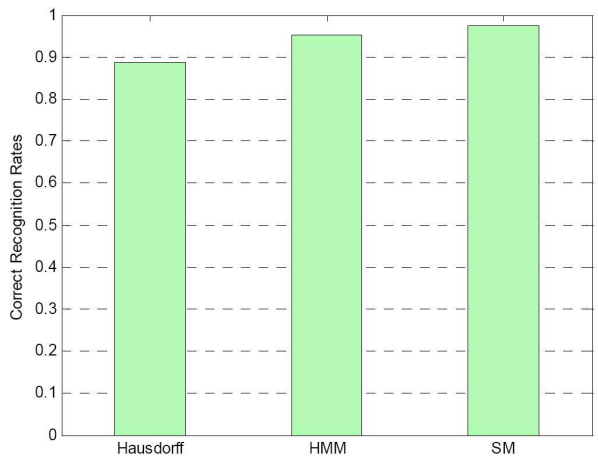

Figure 5. Correct Recognition Rates of Motion Analysis.

Table I
CONFUSION MATRIX.

	M1	M2	M3	M4	M5	M6	M7
M1	341	5	0	0	0	0	0
M2	2	291	0	1	0	0	0
M3	0	0	81	2	0	0	1
M4	0	0	3	84	0	5	0
M5	0	2	0	0	109	1	0
M6	0	0	0	0	5	112	1
M7	0	0	0	5	0	0	149

Figure 5 shows the correct recognition rates of three methods respectively. We can see that our approach performs better than the other two methods. This is mainly because of the strong modeling ability of SM. SM can grasp the underlying information of observation sequence by using segment distribution. It is more precise than that of HMM. Table I gives the confusion matrix, in which the element of each row shows the number of certain kind of motion is classified as other kinds of motions. We can see that most trajectories are correctly classfied. Only a few ones are wrongly recognized that mainly because of the poor quality of trajectory extracted from real vehicle motion.

Above all, from the experiments results we can see the effectiveness of our approach. The SM provides an alternative to HMM in motion analysis.

VII. CONCLUSIONS

In this paper, we have proposed a framework for analyzing real vehicle motion in visual traffic surveillance by using Segment Model (SM), which is a kind of probabilistic model. In the experiments, we have compared our approach with the template matching method based on the Hausdorff distance and the state space method based on the Hidden Markov Model (HMM). Experiment results have showed the effectiveness of our approach. Because of the strong modeling ability, using SM to analyze other object motion in computer vision such as human action analysis is our research direction in the future.

ACKNOWLEDGMENT

This work is partly supported by NSFC (Grant No. 60825204).

REFERENCES

[1] J. A. Bilmes, "A gentle tutorial of the em algorithm and its application to parameter estimation for gaussian mixture and hidden markov models," *Technical report, ICSI-TR-97-021, University of Berkeley*, 1998.

[2] L. Eikvil and R. B. Huseby, "Traffic surveillance in realtime using hidden markov models," *12th Scandinavian Conference on Image Analysis*, 2001.

[3] W. Hu, X. Xiao, Z. Fu, D. Xie, T. Tan, and S. Maybank, "A system for learning statistical motion patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 9, pp. 1450–1464, 2006.

[4] D. Koller, J. Weber, and J. Malik, "Robust multiple car tracking with occlusion reasoning," *Proc. of Europe Conference on Computer Vision*, 1994.

[5] X. Li, W. Hu, and W. Hu, "A coarse-to-fine strategy for vehicle motion trajectory clustering," *International Conference on Pattern Recognition*, pp. 591–594, 2006.

[6] S. G. Mallat, "A theory for multiresolution signal decomposition: A wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.

[7] M. Ostendorf, V. Digalakis, and O. Kinball, "From hmms to segment models: A unified view of stochastic modeling for speech recognition," *IEEE Transaction on Speech Audio Processing*, vol. 4, no. 5, pp. 360–378, 1996.

[8] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," pp. 257–286, 1989.

[9] Y. Tang, W. Liu, Y. Zhang, and B. Xu, "A framework for fast segment model by avoidance of redundant computation on segment," *International Symposium on Chinese Spoken Language Processing*, pp. 117–120, 2004.

Authorized licensed use limited to: INSTITUTE OF AUTOMATION CAS. Downloaded on December 3, 2009 at 03:41 from IEEE Xplore. Restrictions apply.