Classification of Traffic Events based on the Spatio-Temporal MRF Model and the Bayesian Network

Shunsuke KAMIJO Masao SAKAUCHI

Institute of Industrial Science, University of Tokyo
4-6-1 Komaba, Meguro-ku, Tokyo JAPAN 153-8505
Tel:+81-3-5452-6272 Fax:+81-3-5452-6274 E-mail:kamijo@iis.u-tokyo.ac.jp

Abstract

In order to support safe and efficient driving, it is important to classify the behaviors of vehicles and to understand what is going on the traffic situations. For that purpose, our system employed a vision sensor rather than spot sensors because of its rich information. We then have developed a dedicated vehicle tracking algorithm based on the Spatio-Temporal MRF model which is robust against heavy occlusions even in low-angle images. This algorithm outputs maps that represent distributions of vehicle regions and motions by every image frames. By analyzing time-series variations of these distribution maps, our system extracts observation sequences for traffic events classification. We then constructed a classification network based on the Bayesian Network which consists of a fusion of deterministic sub-networks and probabilistic sub-networks. Since vehicle behaviors in the traffic should be restricted by traffic rules, such behaviors would be classified by deterministic network based on explicitly defined traffic rules. However, it is difficult to classify and to understand behaviors such as accidents, near misses, and some other ambiguous behaviors by deterministic classification networks. In order to classify such ambiguous behavior sequences, it is effective to employ probabilistic sub-networks such as Hidden Markov Model with its learning methods. Consequently, by integrating those deterministic networks and probabilistic networks into the Bayesian Network, our system were able to successfully classify behaviors of rule violations, accidents, near miss situations distinguishing from ordinary behaviors in traffic images.

1 Introduction

For the past many years, one of the most important objectives on ITS researches has been the development of systems that reduces incidents and traffic jams in urban traffics. It is said that half of the incidents can be avoided if drivers can aware of dangerous situation in 0.5 seconds before the moments of incidents. And if drivers can aware of the occurred incidents, they will be able to avoid secondary incidents. For that purpose, some researches on collision avoidance at crossroads have been performed by such as Maeoka[1] and Watanabe[2]. They employed multiple video camera, where one camera is watching in-coming vehicles and the other camera is watching vehicles inside the crossroad waiting for turns.

In order to collect rich information from road traffics, vision sensors are very useful. Although range sensors are successful in some limited scope of usage[3] for traffic monitoring, vision sensors have remarkable advantages for detailed traffic analyses. Algorithms for object tracking have a long history in Computer Vision research. However, the occlusion problem had impeded object tracking from being put into practical use for many years. To resolve the problem, some previous works employed stereo vision method[4], and some other works employed shape models of objects to estimate texture matching with images[8]. However, stereo systems would require a huge amount of calculation time and complicated system architectures, and shape models would suffer from objects of unexpected shapes. We then decided to resolve the problem by the segmentation of spatio-temporal images. For that purpose, the Spatio-Temporal Markov Random Field model(S-T MRF) was defined[15][16]. The S-T MRF was define by extending the above spatial MRF model so that it would be effective for segmentation of spatio-temporal images.

Finally, by using such a precise tracking algorithm, we then developed a dedicated algorithm for traffic scene understanding based on the Bayesian Network. Although there are successful previous works on traffic rule reasoning[9] or violation detection[10], they can be applied to simple situation such as small crossroad with a single lane. Compared to such simple reasoning methods, the Bayesian Network[11][12] is a generalized model of causal networks that are able to describe causal effects among various events in a network model. For example, Heckerman et.al[13] successfuly applied the Bayesian Network to a expert system for Medicine. Since our purpose is to analyze complicated behaviors at a large crossroad, the Bayesian Network would provide an integrated method for events classification and traffic scene understanding.

In this paper, vehicle tracking based on S-T MRF will be described in Section.2, and events classification based on the Bayesian Network will be described in Section.3.

2 Spatio-Temporal MRF Model

2.1 Basic Idea

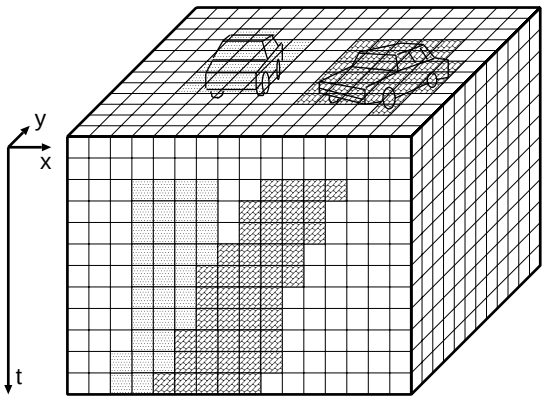
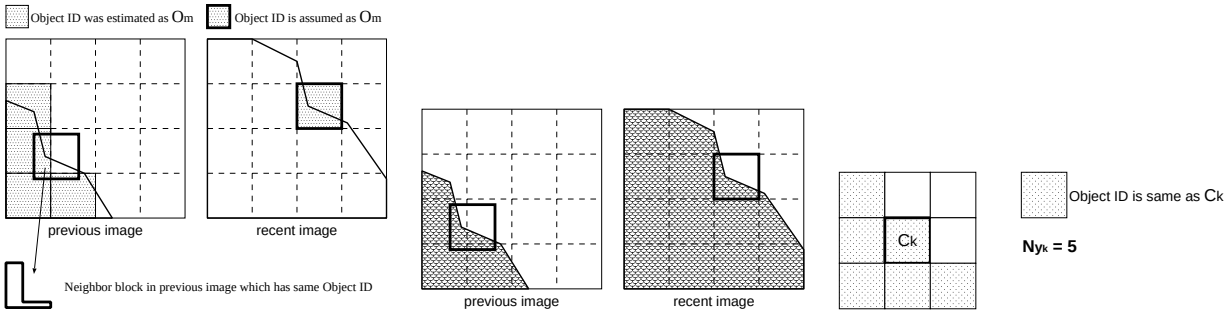


Figure 1: Segmentation of Spatio-Temporal Images

Usually, the spatial MRF segments an image by each pixel. However, since the usual video cameras do not have such high frame rates, objects typically move for ten or twenty pixels among consecutive image frames. Therefore, neighbor pixels within a cubic clique will never have correlations of either intensities or labeling. Consequently, we defined our Spatio-Temporal Markov Random Field model(S-T MRF)[15][16] as to divide an image into blocks as a group of pixels, and to optimize labeling of such blocks by referring to texture and labeling correlations among them, in combination with their motion vectors. Combined with employing stochastic relaxation method, our S-T MRF optimizes object boundaries precisely, even when serious occlusions occur.

Here, a block corresponds to a site in the S-T MRF, and only the blocks that have different textures from the background image are labeled as one of the object regions. In this paper, an image has 640x480 pixels and a block has 8x8 pixels; such a distribution of labels on blocks is referred to as an Object-Map. S-T MRF estimates current Object-map $X(t) = y$; given previous Object-map$X(t-1) = x$, previous image $G(t-1; i, j) = g(i, j)$, and current image $G(t, i, j) = h(i, j)$.

2.2 Parameters for Optimization



(a)Neighbor condition between Consecutive Images (b)Texture Matching (c)8 neighbor blocks

Figure 2: Parameters for S-T MRF

In our previous works[15], three energy function were defined in order to solve the segmentation problem by S-T MRF. Since details of the functions can be found in the previous paper, summary of the idea will be explained here. Following energy functions were deduced from the Boltzmann distribution which represents exponential value of Gaussian function about parameters M_{xy_k} and D_{xy_k} as described in the previous paper[15].

At first, the motion vector of each block is estimated between the previous image and current image by block matching technique. By referring to the motion vector, two S-T MRF energy will be evaluated as shown in Function(1):

$$U_{pre}(D_{xy_k}, M_{xy_k}) = b(M_{xy_k} - \mu_{M_{xy}})^2 + c(D_{xy_k} - \mu_{D_{xy}})^2 \quad (1)$$

$$D_{xy_k} = \sum_{0 \le di < 8, 0 \le dj < 8} |G(t; i+di, j+dj) - G(t-1; i+di-v_{mi}, j+dj-v_{mj})| \quad (2)$$

M_{xy_k} is a goodness measure of the previous Object-map $X(t-1) = x$ under a currently assumed Object-map $X(t) = y$. Assume that a block C_k has a object label O_m in the current object map $X(t)$, and C_k is shifted backward in the amount of estimated motion vector, $-\overrightarrow{V_{O_m}} = (-v_{mi}, -v_{mj})$ of the object O_m, in the previous image (Figure.2(a)). Then the degree of overlapping is estimated as M_{xy_k}: the number of overlapping pixels of the blocks labeled as the same object. The more the overlapping pixels are, the more likely a block C_k belongs to the object. The maximum number is $\mu_{M_{xy}} = 64$, and the energy function $U_M(M_{xy_k})$ takes a minimum value at $M_{xy_k} = 64$ and a maximum value at

$M_{xy_k} = 0$. D_{xy_k} represents texture correlation between $G(t-1)$ and $G(t)$. Suppose that C_k is translated backward in the image $G(t-1)$ referring to the estimate motion vector $-\overrightarrow{V_{O_m}} = (-v_{mi}, -v_{mj})$. The texture correlation at the block C_k is evaluated as(See Figure.2(b)): $U_D(D_{xy_k})$ takes maximum value at $D_{xy_k} = 0$. The smaller D_{xy_k} is, the more likely C_k belong to the object. That is, the smaller $U_D(D_{xy_k})$ is, the more likely C_k belong to the object.

The last S-T MRF energy is of neighbor condition within a current Object-Map as shown in Function(3).

$$U_N(N_{y_k}) = a(N_{y_k} - \mu_{N_y})^2 \tag{3}$$

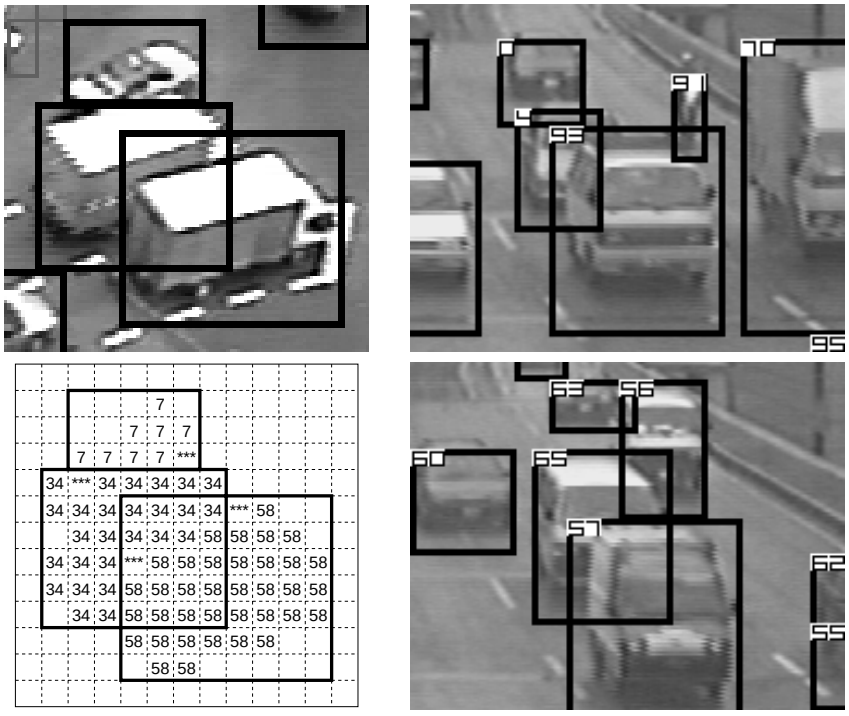
Here, N_{y_k} is the number of neighbor blocks of a block C_k that belong to the same object as C_k as shown in Figure.2(c). Namely, the more neighbor blocks that have the same object label, the more likely the block is to have the object label. Currently, it is assumed that $\mu_{N_y} = 8$, because $U_N(N_{y_k})$ should have minimum value when block C_k and all its neighbors have the same object label.

In addition, some of such estimated motion vectors would have errors. Such errors frequently occur on the boundaries of occluded objects, because the appearances of object boundaries vary along sequence of images. Since those errors should lead to segmentation errors, it is necessary to correct errors of motion vectors themselves. For that purpose, it would be effective to optimize motion vectors themselves by referring to motion vectors of their neighbor blocks. This condition can be integrated as the following energy function(4):

$$\begin{aligned} U(y_k(t)) \quad + \quad & fU_{mv}(C_k(t-1)) = \\ & a(N_{y_k} - \mu_{N_y})^2 + b(M_{xy_k} - \mu_{M_{xy}})^2 + cD^2_{xy_k} + f\sum_{B_k} |\overrightarrow{V_{C_k(t-1)}} - \overrightarrow{V_{B_k(t-1)}}|^2 / N_{x_k} \end{aligned} \tag{4}$$

Energy terms of $U_M(M_{xy_k})$ and $U_D(D_{xy_k})$ will be evaluated by referring to respective motion vectors of blocks belonging to the object. $U_{mv}(C_k(t-1))$ will be estimated by using motion vectors at $T = t-1$; $C_k(t-1)$ represents the original block of $C_k(t)$, N_{x_k} represents the number of neighbor blocks that have same label as $C_k(t-1)$. Thus, motion vectors of blocks at $T = t-1$ and Object-Map at $T = t$ will be optimized simultaneously by considering both similarities in motion vectors among neighbor blocks and in texture correlations between consecutive images.

2.3 Tracking Results by S-T MRF



(a) Crossroad Image & Object-Map (b) Low-angle Images at Merge Traffic

Figure 3: Tracking results by S-T MRF

Figure.3(a) shows the tracking result image and the Object-Map by applying our S-T MRF to images at a crossroad, and Figure.3(b) shows the tracking result images of low-angle images at highway merge traffic. Parameters were decided by trial and error as: $a = 1/2, b = 1/256, c = 32/1000000, f = 1/4$. 25 minute traffic images were examined for a large crossroad which has three lanes for each direction. During the, 3214 vehicles went through the crossroad. As a result, the method was able to segment and track vehicles at about 95% success rate against occlusions. On the other hand, 40 minute images were also examined for the merge traffic on the Tokyo Metropolitan Expressway. During then, 2,381 vehicles have passed this junction, and the method achieved 91.3% success rates in tracking against occlusions.

3 Events Classification based on the Bayesian Network

As can be seen in Figure.10 in Section.5, there simultaneously exist ten or twenty vehicles at such a large crossroad, and those vehicles run through in so many behaviors. Accordingly, it seams very difficult to classify their behaviors by such simple methods as [9][10]. The Bayesian Network is very useful to resolve so complicated reasoning problems including deterministic problems and probabilistic problems.

3.1 Building a Network Model for Traffic Events Classification

By using such a precise tracking results, we then developed a dedicated algorithm for traffic events classification based on the Bayesian Network. The Bayesian Network[11][12] is a generalized model of causal networks that are able to describe causal effects among various events in a network model. This Bayesian Network would provide an integrated method for events classification and traffic scene understanding.

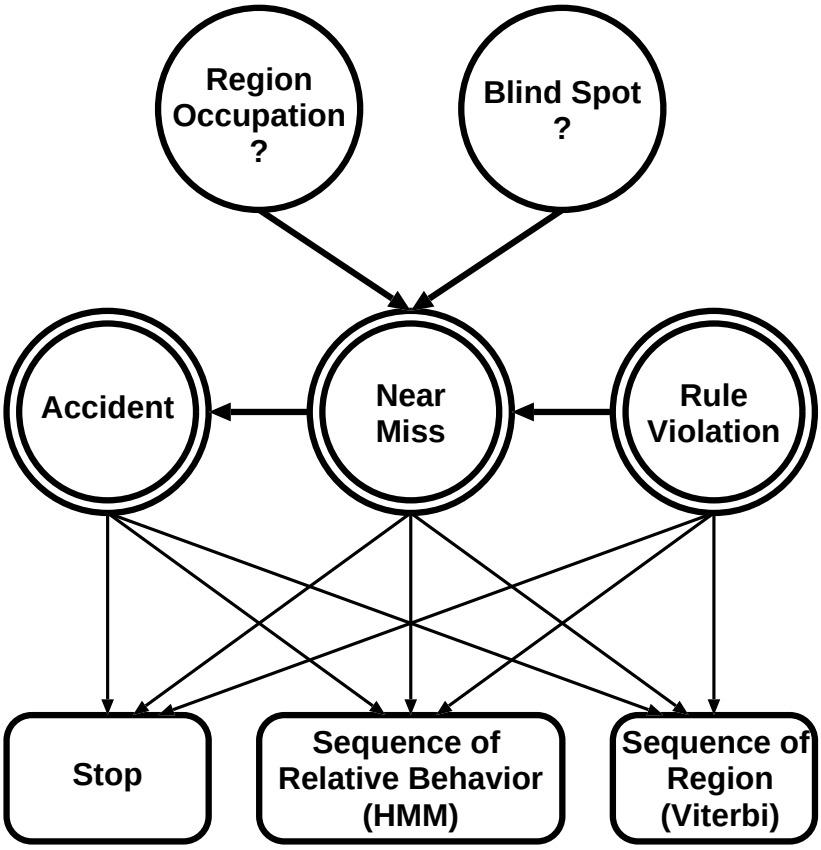


Figure 4: The Bayesian Network for Traffic Behaviors

Generally, nodes and causal links of the Bayesian network are to be built by learning methods and/or deterministic methods. Since the traffic rules can be described in deterministic ways, it seams appropriate to build the node and causal links by deterministic methods. In addition, nodes and causal links of the network do not necessarily have significant meanings. However, it also seams appropriate to build them with some significant meaning. It then will becomes easy to analyze causality and to understand situations with significant meanings. Consequently the Bayesian Network for traffic events classification can be modeled as shown in Figure.4 by considering explicitly defined traffic rules.

3.2 Nodes and their Variables

In Figure.4, we assumed three traffic events as described in Table.1 and three test functions as described in Table.2. Those nodes are connected by causal links with each others. By testing observations from traffic images by three test functions, an event which caused such observations will be determined. When the test result cannot be classified into none of three events, it will be regarded as an ordinary situation.

Node	Variable
Rule Violation(RV)	violate(VIO) safe(SF)
Near Miss(NM)	near miss(NM) safe(SF)
Accident(ACC)	accident(ACC) safe(SF)

Table 1: Nodes and Variables for Events

Node	Variable
Stop(STP)	run(RN) stop at permitted region(PM) stop at prohibited region(PH)
Sequence of Relative Behavior (RB)	PAS-00 PAS-01 PAS-02 PAS-13 TDM-01
Sequence of Region (RG)	permitted(PM) prohibited(PH)

Table 2: Nodes and Variables for Test Functions

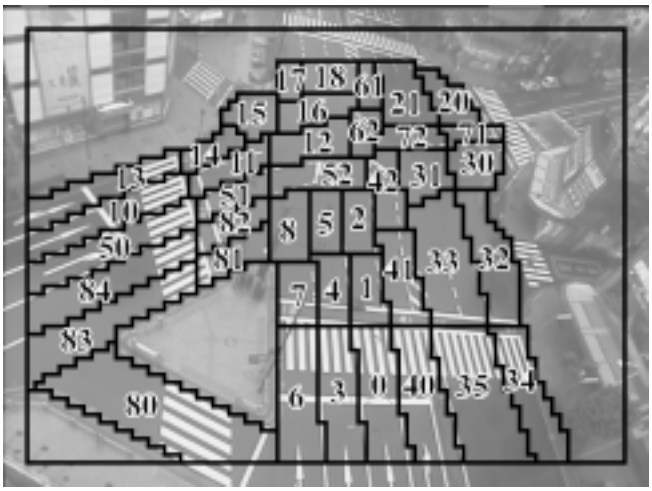


Figure 5: Labeling of Region

In Table.1, nodes of three traffic events and their variables are described. In Table.2, nodes of three test functions and their variables are described.

For example, the node of 'Stop' has three variables of 'run(RN)', 'stop at a permitted region(PM)', and 'stop at prohibited region(PH)'. Those variables can be determined by simple evaluations as follows: When vehicle tracking results were obtained, Object-Map represent labeling of blocks into vehicle IDs. Therefore, a region where each vehicle exists can be estimated by checking blocks which is labeled into the vehicle. If a vehicle traveling from bottom to top of the image exists at region-4 while the signal is GREEN for the vehicle, the test result is determined as STP=PH.

For another example, the node of 'Sequence of Region ' has two variables of 'permitted(PM)' and 'prohibited(PH)'. Those variables can be determined by testing a sequence of such vehicle regions. For example, when a sequence of regions are '0-1-2', test result is determined as RS=PM. On the other hand, test result is determined as RS=PH when a sequence of regions are '40-1-2'.

However, 'Sequence of Relative Behaviors' with respect to a pair of vehicle such as passing, tandem, or bumping cannot not be determined by such deterministic test functions. Therefore, we employed Hidden Markov Model(HMM) in order to test such relative behaviors as described in Section.4. HMM is a stochastic method, and it can be regarded as a variation of Bayesian network(See Figure.9). Therefore, this test function for classification of relative behaviors is a sub-network of the entire network for traffic events classification. Since HMM has no significant meaning of nodes, nodes will be built by some evaluation of performance for classification.

3.3 Causal Links and Probabilities on Causal Links

		RV	
		VIO	SF
STP	RN	90%	98%
	RM	2%	1%
	RH	8%	1%

(a)$P(STP|RV)$

		ACC	
		ACC	SF
STP	RN	2%	98%
	RM	8%	1%
	RH	90%	1%

(b)$P(STP|ACC)$

		RV	
		VIO	SF
RG	PM	2%	98%
	PH	98%	2%

(c)$P(RG|RV)$

Table 3: Examples of Conditional Probabilities on Causal Links

Although causal links of the network do not necessarily have significant meanings as described previously, it seams effective to defined such causal links by referring to explicit traffic rules. And conditional probabilities are to be attached on causal links respectively. Here in generally, such

probabilities are estimated in learning methods and/or deterministic methods. However in such cases of traffic behaviors, most of the probabilities on causal links can be defined in deterministic way by considering traffic rules. Table.3 shows examples of such conditional probabilities on the causal links. On the other hand, the test function of 'Sequence of Relative Behavior' is built by HMM as a variation of the Bayesian network as described previously. Since HMM is a probabilistic method with no significant meaning on causal links, and probabilities on causal links will be estimated by learning process[14].

In addition, causal links from 'Rule Violation' to 'Near Miss' represents that 'Rule Violation would cause 'Near Miss' in some probabilities. As the same, 'Near Miss' would cause 'Accident' in some probabilities.

In order to classify traffic events among RV, NM, and ACC, all the three test functions of STP, RB, and RG should be evaluated at first. The most probable event then will be determined by considering such evaluation results of the three test and conditional probabilities from the three events[12][13].

4 Classification of Relative Behaviors by HMM

Motion sequences of vehicles that are involved in abnormal traffic events would be very different from motion sequences in ordinary situations. Therefore, by observing such motion sequences, it would be possible to forecast accident and near miss situations as well as to detect them.

4.1 Feature Extraction about Relative Motion and Relative Location

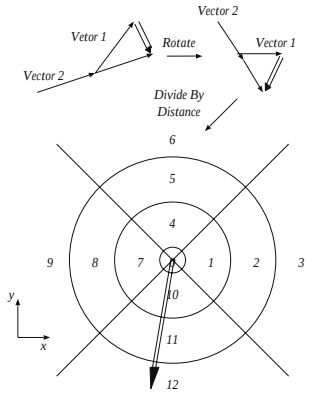


Figure 6: Feature Extraction about Relative Motion Vector

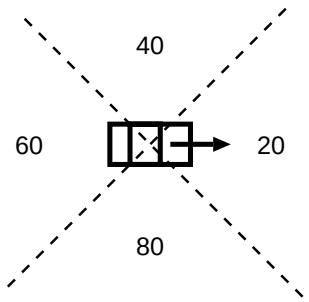


Figure 7: Feature Extraction about Relative Location

The algorithm of behavior classification should be independent of geometric factors, such as geometry of the intersection, angle of video camera and position where the accident occurred. Therefore, we focused on the relative behavior of a pair of vehicles.

At first, the relative motion of a pair of vehicles seems to be important. Features of relative motion can be obtained as follows(Figure.6):

$$\overrightarrow{V_n} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} (\overrightarrow{V_2} - \overrightarrow{V_1})/d_{12} \tag{5}$$

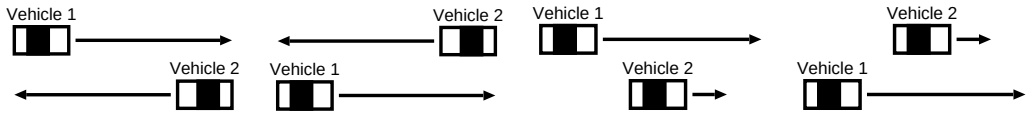
In addition, in order to classify situations with respect to a pair of vehicles, relations between the locations of the vehicles are necessary as well as relative motion vectors of the vehicles. Here, we call such relative locations between a pair of vehicles as 'relative locations'.

4.2 Combined Observations for Relative Behaviors

Even when the relative motion vectors are equal, situations differ depending on difference in relative locations. Therefore, in order to classify situations with respect to a pair of vehicles, it is necessary to consider features obtained from both relative motion vectors and relative locations. It seems to be appropriate to define a combined observation as summation of an observation number obtained from relative motion vectors and an observation number obtained from relative locations.

Ordinary behaviors between a pair of vehicles can be classified into following four classes. Figure.8(a) represents two vehicles on the opposite side passing each other, and such states are observed in cases of through traffics. Figure.8(b) represents two vehicles on the opposite side passing each

other, but the locations of the two vehicles are different from Figure.8(a). Such states are observed in cases when both vehicles make right turns simultaneously. Figure.8(c) represents a vehicle passing by the other vehicle from the left lane of the same side. Such states are observed while the other vehicle is waiting in order to make right turns. Figure.8(d) represents a vehicle passing by the other vehicle from the right lane of the same side. Such states are observed while the other vehicle is waiting for pedestrians in order to make left turns.



(a) PAS-00 (b) PAS-11 (c) PAS-02 (d) PAS-13

Figure 8: Relations between the two Vehicles of Normal Traffics

For example as in Figure.8(a), an observation number of Vehicle-2 relative to Vehicle-1 according to relative motion vector between them is '7', '8' or '9', and an observation number according to relative locations will be '20'. Thus an observation number of Vehicle-2 relative to Vehicle-1 will be '27', '28' or '29'. In this case an observation of Vehicle-1 relative to Vehicle-2 is equal to an observation of Vehicle-2 relative to Vehicle-1. However in the case of 8(c), an observation number of Vehicle-2 relative to Vehicle-1 as '27' etc differs from that of Vehicle-1 relative to Vehicle-2 as '61' etc.

Hereby, combined sequences will be categorized into following four types:

[pas-0] 27 27 88 88 89 87 87 87 87 67 67 67
27 28 88 88 88 88 87 87 87 67 67
27 27 20 27 28 87 88 88 88 88 87 87 67 67
[pas-1] 27 27 47 47 47 47 47 47 47 67 67 67
27 27 27 28 47 48 48 48 48 47 47 67 67
27 27 27 48 48 48 49 68 68 68 68
[pas-2] 61 61 42 42 43 41 41 41 41 21 21 21
61 60 60 61 62 41 42 42 42 42 41 41 21 20 21
62 62 62 43 43 43 42 22 22 22 22
[pas-3] 61 61 82 82 83 81 81 81 81 21 21
61 60 60 61 62 81 82 82 82 82 81 81 21 21
62 62 62 83 83 83 82 22 22 22 22 21

By using above four types of observation sequences, pas-0, pas-1, pas-2, and pas-3, relative behaviors between a pair of vehicles can be categorized as follows:

[PAS-00] Figure.8(a) : The sequence of Vehicle-2 observed from Vehicle-1 is as pas-0, and the sequence of Vehicle-2 observed from Vehicle-1 is as pas-0.

[PAS-11] Figure.8(b) : The sequence of Vehicle-2 observed from Vehicle-1 is as pas-1, and the sequence of Vehicle-2 observed from Vehicle-1 is as pas-1.

[PAS-02] Figure.8(c) : The sequence of Vehicle-2 observed from Vehicle-1 is as pas-0, and the sequence of Vehicle-2 observed from Vehicle-1 is as pas-2.

[PAS-13] Figure.8(d) : The sequence of Vehicle-2 observed from Vehicle-1 is as pas-1, and the sequence of Vehicle-2 observed from Vehicle-1 is as pas-3.

In order to classify those ordinary behaviors, observation sequences of every pairs of vehicles will be examined by a classification method to determine into which class out of PAS-00, PAS-11, PAS-02 and PAS-13 they most likely to classified, as described in subsection.4.4.

4.3 Observation Sequences for Bumping Accidents and Tandem Behaviors

Since observation sequences of vehicles that are involved in bumping accidents will be very different from those of ordinary situations the system is expected to be able to detect bumping accidents by distinguishing obtained observation sequences of bumping accidents from those of ordinary situations by using HMM. However, observation sequences of a bumping accident and a tandem behaviors would be very similar to each other. In addition, it is impossible to distinguish between contacting and not contacting about overlapping vehicles in low-angle images. Therefore, it would be very difficult to distinguish between a bumping behavior and a tandem behavior. Consequently, it seams appropriate to categorize a bumping behavior and a tandem behavior into the same class.

Typical observation sequences of bumping and tandem behaviors are expected to be as follows. For examples of bumping, observation numbers as 61, 62, 63 in tdm-0 and 27, 28, 29 in tdm-1 means that a vehicle is getting closer to the other vehicle, observation numbers as 60 in tdm-0 and 20 in tdm-1 means just the moment of an impact, and observation numbers as 69, 68, 67 in tdm-0 and 23, 22, 21 means reaction after the impact.

[tdm-0] 61 62 62 60 60 60 67 68 67
61 62 62 63 60 60 60 60 60 69 68 67 67
61 60 61 62 60 60 60 68 68 67
[tdm-1] 27 27 28 28 20 20 20 21 22 21
27 28 28 29 20 20 20 20 22 23 22 21
27 28 27 28 20 20 20 20 23 23 22 20 21

Considering the case that Vehicle-2 bumped Vehicle-1 from then behind, an observation sequence of Vehicle-2 relative to Vehicle-1 appears to be as examples of tdm-0, and an observation sequence of Vehicle-1 relative to Vehicle-2 appears to be as examples of tdm-1. Here this class of sequence is defined as TDM-01.

4.4 Classification utilizing HMM

Classifying sequences is a class of recognition problems to classify time series observations. There are various techniques for time-series event-recognition, such as DP-Matching, Neural Network and Hidden Markov Model. At first, we prefer methods based on stochastic models to those based on concrete models, because an accident sequence consists of a large number of random processes and those can be described by using stochastic models. DP-Matching is a kind of strict recognition method which is not robust against disturbance in observations. On that point, stochastic methods utilizing Hidden Markov Model(HMM) or Neural Network(NN) are superior to DP-Matching. In addition, we prefer methods that are robust against disturbance in length of observation sequences. HMM is considered to be superior to NN on that point. Thus, we decided to apply HMM to accident detection. We then applied simple left-to-right HMM for accident detection, as shown in Figure.9. A HMM model has parameters as follows.

- a_{ij} : Transition probability from state i to state j. Where, $a_{ij} = 0(j \neq i, i+1)$.
- $b_{ij}(k)$: Probability to output observation k when transition from state i to state j occurred. Where, $b_{ij}(k) = 0(j \neq i, i+1)$.

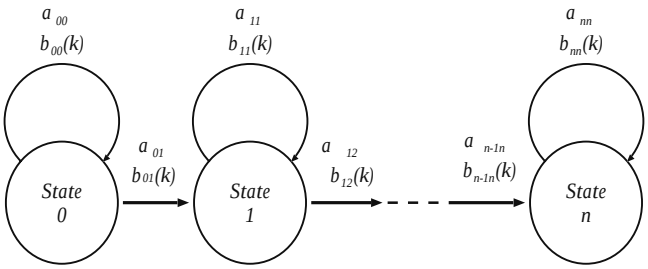
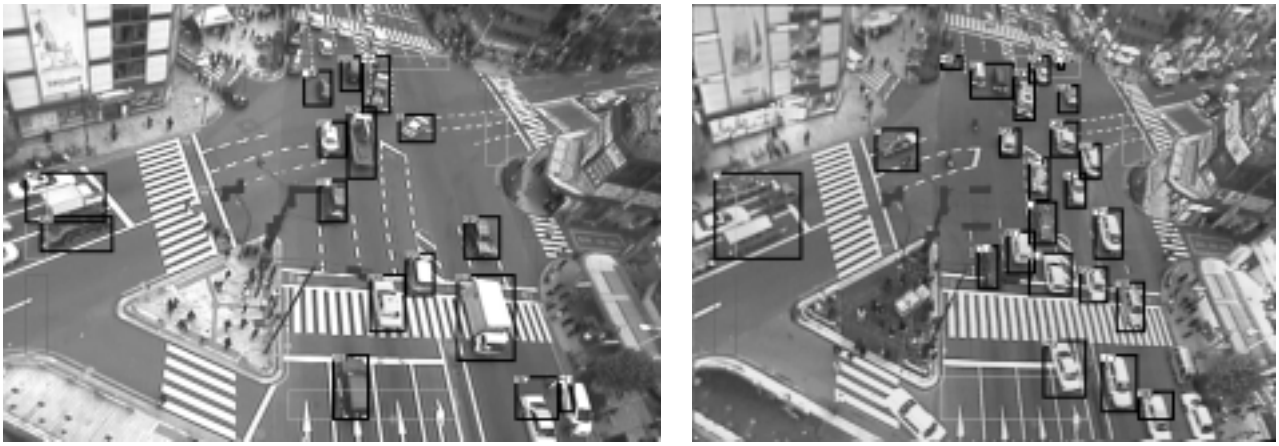


Figure 9: Left-to-Right HMM

Probability parameters of $a_{ij}, b_{ij}(k)$ are trained by Baum-Welch algorithm[14]. For recognition, Viterbi variables will be calculated inductively. The model λ is provided with each category to be recognized. And the model which has the most likely probability with a test sequence is determined.

5 Experimental Results

5.1 Detection of Rule Violation and Near Miss



frame 5805 frame 190

frame 5820 frame 204

frame 5835 frame 230

(a) violation (b) Violation caused Near Miss

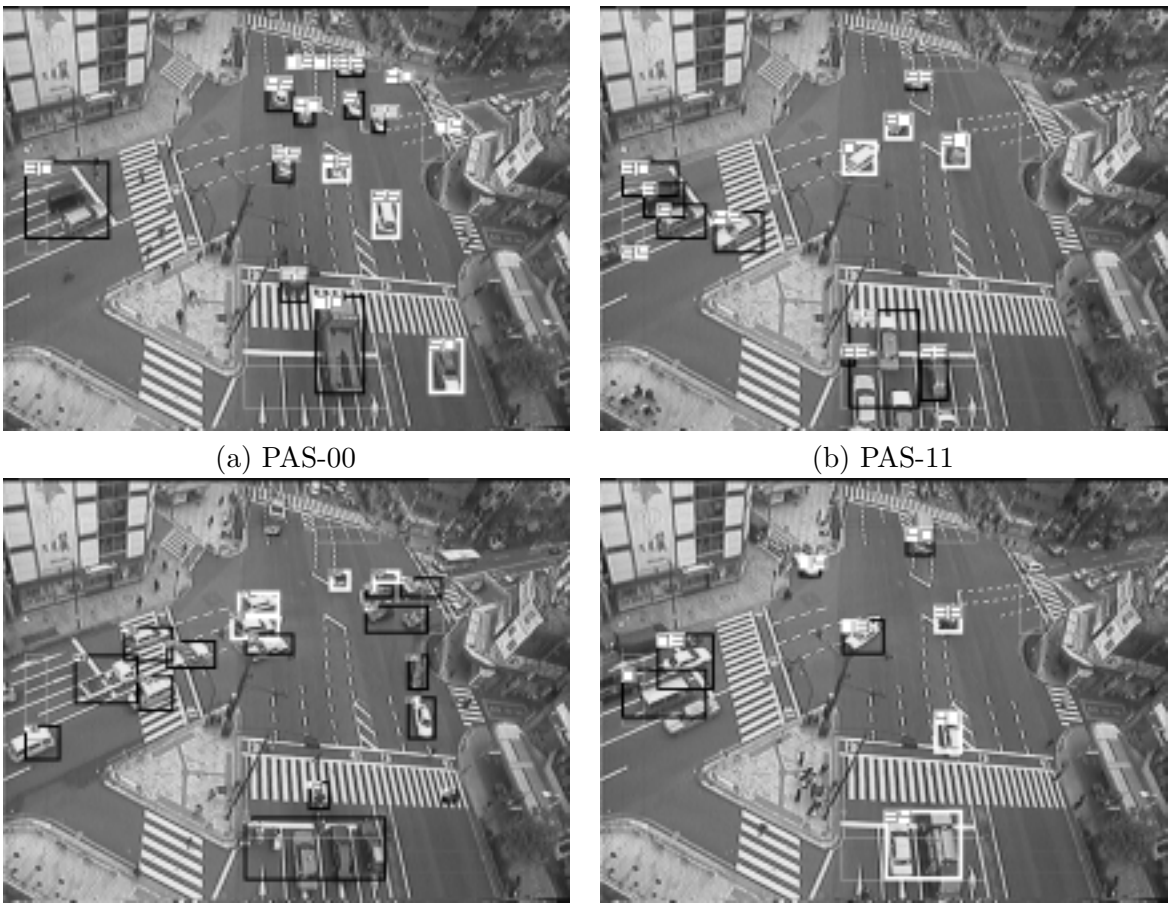
Figure 10: The Statistics System and Q-V curve

Figure.10 represents a scene of a rule violation detection. As described in Figure.5, the system has a map of regions at the crossroad, where regions are signified according to traffic rules such as lanes, directions, stopping-lines, and so on. By observing a sequence of regions where a vehicle traveled, the system is able to find a rule violation. For example in Figure5, although driving from region-41 to region-42 is permitted(RG=PM), driving from region-41 to region-33 is prohibited(RG=PH). As the

same, although driving from region-40 to region-41 is permitted(RG=PM), driving from region-40 to region-0 or deriving from region-40 to region-1 are prohibited(RG=PH).

In addition, it is important to forecast the situation which would lead to near miss situation in advance. In the Figure10(b), the violating vehicle should drive from region-1 via region-2 to region-52, because traffic signal of between the bottom and the top is GREEN now. On the other hand, vehicle waiting for the right turn should drive from region-62 to region-52 to make the right turn. In such the situation, it is difficult for the vehicle waiting for the right turn to notice that the violating vehicle might come close to him behind the other vehicles. Consequently, the system should have the knowledge to find such a situation, and it is important to supervise to avoid a collision for both the violating vehicle and the waiting vehicle. For that purpose, in Figure.4, nodes of 'Region Occupation and 'Blind Spot' was defined.

5.2 Classification of Relative Behaviors



(a) PAS-00 (b) PAS-11

(c) PAS-02 (d) PAS-13

Figure 11: Recognition Results for PAS-00, PAS-11, PAS02, PAS13

Four figures in Figure.11 show results of classifying combined observation sequences of a lot of pairs of vehicles into classes of PAS-00, PAS-11, PAS-02, PAS-13 by using HMM. In Figure.11(a), vehicles of ID-50,56 were detected as PAS-00 with respect to the vehicle of ID-46 and they were indicated in white rectangle. Here, vehicles other than ID-46 were out of examination for simplicity of the figure. In Figure.11(b), vehicles of ID-0,87 were detected as PAS-11 with respect to the vehicle of ID-91. In Figure.11(c), vehicles of ID-55,65,77 were detected as PAS-02 with respect to the vehicle of ID-23. In Figure.11(d), vehicles of ID-38,41 were detected as PAS-13 with respect to the vehicle of ID-34.

To verify reliability of this algorithm, about tree hundred pairs of vehicles were investigated with respect to each class. As results, success rate of classifications were 93.1% for PAS-00, 85.2% for PAS-11, 92.0% for PAS-02, and 88.9% for PAS-03. The classification result of PAS-11 is 85.2% and it is lower than other classes of PAS-00, PAS-02, PAS-03. Here, since vehicles run in the left side in

Japanese traffic rules, it is important to remind that PAS-11 situations are observed when pairs vehicles make right turns simultaneously. Since vehicles stop for long periods at the right turns lane in such cases, observation sequences tend to be short than other situations. Therefore, at beginning periods of passing situations, vehicles tend to move quite slowly, and this then caused lacks of beginning parts in observation sequences about PAS-11.

5.3 Classification of an Accident

Since there are very few data of accidents, the HMM model for accidents and tandem situations was trained by fifty sequences of usual tandem situations including disturbance in observation numbers and length of sequences. And Figure.12 shows results of accident detections.

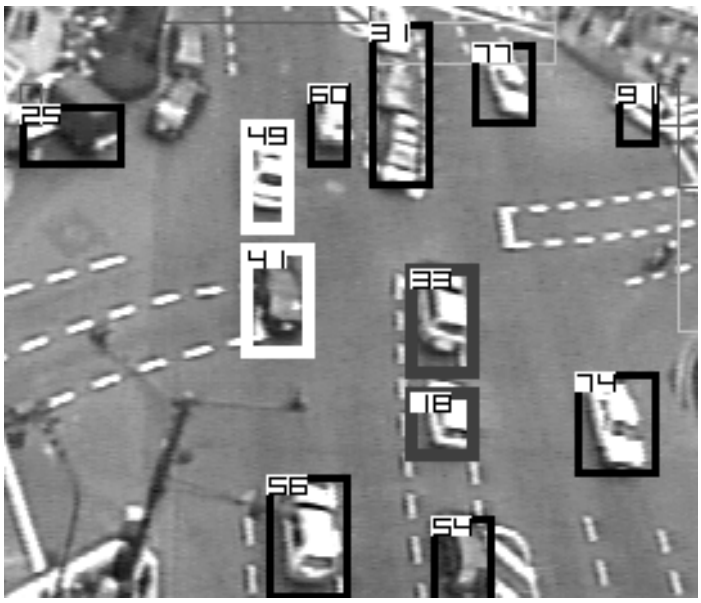


Figure 12: Bumping Accident

In this figure, the algorithm detected two pairs of vehicles that relative behavior were classified as RB=TDM-11 by the test function of 'Sequence of Relative Behavior'. However, a pair of Vehicle-18 and Vehicle-33 was not classified into an accident(ACC=ACC), because the two vehicle stopped at permitted region to wait right turns. On the other hand, a pair of Vehicle-41 and Vehicle-49 were also classified into RB=TDM-11, and they were classified into an accident(ACC=ACC). In this situation, a pair of Vehicle-41 and Vehicle-49 stopped at prohibited region.

6 Conclusions

We developed a dedicated algorithm for traffic events classification based on the Bayesian network. Our vehicle tracking algorithm based on the Spatio-Temporal MRF model which is robust against heavy occlusions even in low-angle images, and this algorithm outputs maps that represent distributions of vehicle regions and motions by every image frames. In order to classify time-series variations of those distribution maps, we constructed a classification network based on the Bayesian Network. The network consists of a fusion of deterministic sub-networks based on explicitly defined traffic rules and probabilistic sub-networks based on learning methods. Consequently, by integrating those deterministic networks and probabilistic networks into the Bayesian Network, our system were able to successfully detect rule violations, accidents, near miss situations, and so on.

References

[1] A.Maeoka, M.Iwasaki, T.Tajima, "Supporting Safe Driving at Signalized Intersection", IV2001, Tokyo, May. 2001, pp.227-232.

[2] H.Watanabe et.al., "Development of Intersection Monitoring System", 8th World Congress on ITS, Sydney Oct.2001, CD-ROM.

[3] K.C.Dietmayer, J.Sparbert, and D.Stellerm, "Model-Based Object Classification and Object Tracking in Traffic Scenes from Range-Images" IV2001, Tokyo, May. 2001, pp.25-30.

[4] Y.Sugaya, Y.Ohta, "Stereo by Integration of Two Algorithms with/without Occlusion Handling", ICPR'00, Barcelona, Sep. 2000, Vol.1, pp.109-113.

[5] Natan Peterfreund, "Robust Tracking of Position and Velocity With Kalman Snakes" IEEE Trans. Pattern Analysis and Machine Intelligence(PAMI), Vol.21 No.6, 1999, pp.564-569.

[6] M.Kass,A.Witkin,and D.Terzopoulos, "Snakes: Active contour models" Int'l J.Computer Vision, Vol.1, 1988, pp.321-331.

[7] Holger Leuck and Hans-Hellmut Nagel, "Automatic Differentiation Facilitates OF-Integration into Steering-Angle-Based Road Vehicle Tracking", IEEE CVPR'99, pp.360-365.

[8] Warren F.Gardner and Daryl T.Lawton "Interactive Model-Based Vehicle Tracking", IEEE Trans. PAMI, Vol.18 No.11, 1996, pp.1115-1121.

[9] R.Cucchara, M.Piccardi, P.Mello, "Image Analysis and Rule-Base Reasoning for a Traffic Monitoring System.", IEEE/IEEJ/JSAI ITSconf., Tokyo, Oct. 1999, pp.758-763.

[10] D.Lim, S.Choi, J.Jun, "Automated Detection of All Kinds of Violations at a Street Intersection using Real Time Individual Vehicle Tracking.", 8th World Congress on ITS, Sydney Oct.2001, CD-ROM.

[11] J.Pearl, "Reverend Bayes on Inference Engines: A Distributed Hierarchical Approach.", National Conference on Artificial Intelligence, Pittsburgh PA, 1982, pp.133-136. AAAI Press.

[12] J.Breese, "Construction of Belief and Decision Networks", Journal of Computational Intelligence, Vol.8, No.4, pp.624-647.

[13] D.Heckerman, E.Horviz, and B.Nathwani, "Towards normative expert systems: Part I, the Pathfinder project.", Methods for Information in Medicine 31, 1992, pp90-105.

[14] Lawrence R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, " Proc. of the IEEE, Vol.77, No.2, Feb 1989, pp257-286.

[15] S.Kamijo, Y.Matsushita, K.Ikeuchi, M.Sakauchi, "Occlusion Robust Vehicle Tracking utilizing Spatio-Temporal Markov Random Field Model", 7th World Congress on ITS, Torino, Nov. 2000.

[16] S.Kamijo, K.Ikeuchi, M.Sakauchi, "Vehicle Tracking in Low-angle and Front-View Images based on Spatio-Temporal Markov Random Field Model" 8th World Congress on ITS, Sydney Oct.2001, CD-ROM.

[17] S.Kamijo, Y.Matsushita, K.Ikeuchi, M.Sakauchi, "Traffic Monitoring and Accident Detection at Intersections", IEEE trans. ITS, Vol.1 No.2, June. 2000, pp.108-118.