Road Traffic Model using distributed Camera Network

S Indu*
Delhi Technological University
Bawana Road
Delhi 110042
s.indu@rediffmail.com

Sankalp Arora
Delhi Technological University
Bawana Road
Delhi 110042
ar.sankalp@gmail.com

Santanu Chaudhury
Electrical Engg. Deptt.
Indian Institute of Technology
Delhi
schaudhury@gmail.com

Asok Bhattacharyya
Delhi Technological University
Bawana Road
Delhi 110042

ABSTRACT

Traffic monitoring/prediction using a distributed camera network is presented in this paper. The activities on each road link are monitored and features are derived to identify the pattern. Then it is learnt, classified, predicted and communicated to neighboring road links. We used GMM-EM based classification and HMM based prediction. Optimum path is determined by assigning proportional weights to the predicted states of the connected road links. The proposed method is neither based on tracking nor on vehicle detection. Apart from this the method is flexible, adaptive, robust and computationally light . Unlike the existing methods it does not assume or draws analogies of traffic moving as particles, neither does it impose restriction on road conditions or road tributaries and distributaries. The model is validated using traffic simulator and tested on real road network.

Keywords

traffic surveillance, spatial interest points, spatial-temporal interest points, traffic classification, hidden Markov model, traffic prediction

1. INTRODUCTION

In this paper we are presenting a distributed camera network based traffic model. The traffic model we developed extracts video based features and processes the same to classify the road condition as open, slight congestion, heavy congestion or traffic jam. It can also be used for predicting the traffic conditions along the road which is extended to find the best route to the destination. Currently, the dominant technology for this purpose is magnetic loop detectors, which are buried underneath highways to count vehicles passing over them. Video camera based systems present some advantages over the current technology. First, a much larger set of traffic parameters can be estimated in addition to vehicle counts and speeds. These include vehicle classifications, link travel times, lane changes, rapid accelerations or decelerations, queue lengths at urban intersections, etc. Second, cameras are less disruptive and less costly to install than loop detectors, which require digging up the road surface. Therefore, video cameras are becoming more popular in traffic monitoring and control systems.

Sensor nodes which can communicate to the nearby nodes are placed in all road links. A sensor node is equipped with, a camera with accessories to process the data and communication devices. The features (SIP and STIP) are derived from the video frames generated by the camera. Spatial interest points are points (SIP) in spatial domain with significant variation in local intensities where as spatio-temporal interest points (STIP) are points in space time domain with significant variation in local intensities. It is observed that vehicles on a road generate SIP and the moving vehicles generate STIP. Hence, the number of SIP is indicative of number of vehicles on a road and ratio of STIP to number of SIP is suggestive of percentage of moving vehicles. These points are classified using GMM and then traffic state prediction is carried out using HMM. Once the next state of each road link is available, we can obtain the optimal path by assigning weights (in accordance with the predicted state) to each link.

2. RELATED WORK

Traffic modeling has gained significant interest among researchers and lot of work has been carried out already. Even then it is still a challenging work due to its nonlinear nature. The existing research works can be divided into two groups (i) Traffic classification and (ii) classification and prediction.

Most of the existing road classification research works are based on image segmentation and tracking vehicle, which cannot be used for on line learning due to the computational complexity[10] [5] [14] and [7]. Peng Cheng and et. al. [4] developed a particle filter based traffic estimation utilizing hand off data. In this case every vehicle should contain a cell phone in it, which is not practical always. Ren C Boel and et. al. have developed a hybrid stochastic

*Corresponding Auther

ICVGIP '10, December 12-15, 2010, Chennai, India

traffic model[2]. The whole network is divided into sections and the traffic is predicted by using sending and receiving functions. Later they have developed a particle filter based modeling which is faster and can even be used for online prediction. But the main disadvantage was to detect vehicle in every section which makes complex computation. X.Li and et. al.[9] in their work "A Hidden Markov Model framework for traffic event detection using video features" used DCT in spatial and temporal domain as features and used HMM for event detection. This method is computationally complex due to the need of image processing which makes online learning difficult. Y. Zou and et.al [18] have developed HMM based traffic incident detection. The method is location specific and also needs vehicle detection. In general traffic modeling is done with image processing and tracking. We are presenting a model which do not require tracking and hence suitable for online prediction. Yan Qi in his Ph.D. [12] thesis "Probabilistic Models For Short Term Traffic Conditions Prediction" compared the traffic prediction using HMM based model and one step stichastic model. He derived traffic features from embeeded magnetic loops on the road. Embeeding magnetic loop on the road will be a tedious effort. In general all the above methods belong to group (i) and were doing traffic modeling on an individual road link and not extended for real time navigation.

Recently, Wireless Sensor Network (WSN) is playing an important role in traffic modeling / navigation. Vipin Kumar Verma and et. al.[16] developed an ITS using WSN. This method need a sensor in each vehicle, distributed servers on each crossing, and embedded tracking sensors on the road. This will make the system complex and costly. Ben-Jye Chang and et. al in their work "Wireless Sensor Network-based Adaptive Vehicle Navigation in Multihop-Relay WiMAX Networks"[3] explains a traffic navigation model by analyzing the data collected from each moving vehicle through WSN. Here distance between the vehicles and average speed of the road is considered and cost of each road link is calculated. The main disadvantage is that every vehicle should have a WSN in it. Malik Tubaishat and et. al (Adaptive Traffic Light Control with Wireless Sensor Networks) [15] have developed a traffic light control system using embedded magnetic loop sensors on each lane of the road. Mainly this method concentrates on crossings and trying to minimize the weighting period. The proposed work presents an autonomous distributed camera network (instead of embedded sensor network) with traffic state prediction service in terms of traffic conditions of the road. The macroscopic view based traffic model so achieved is computationally light (as we are not using tracking or vehicle identification), camera setup independent and fast (simplification of the video processing at sensor node). Hence the model is suitable for real time implementation. Since the proposed work is based on distributed processing, the camera nodes in the network are not subordinate to any other node and are able to collect, process and communicate information at their own will. As we are communicating the traffic state to the neighboring nodes only, the network is energy efficient too.

3. PROBLEM FORMULATION

The main aim of the proposed work is to develop an autonomous distributed camera network for traffic prediction. A sample road network is shown in fig.1. R_1 - R_{10} are different road links and C_1 - C_{10} are cameras for recording

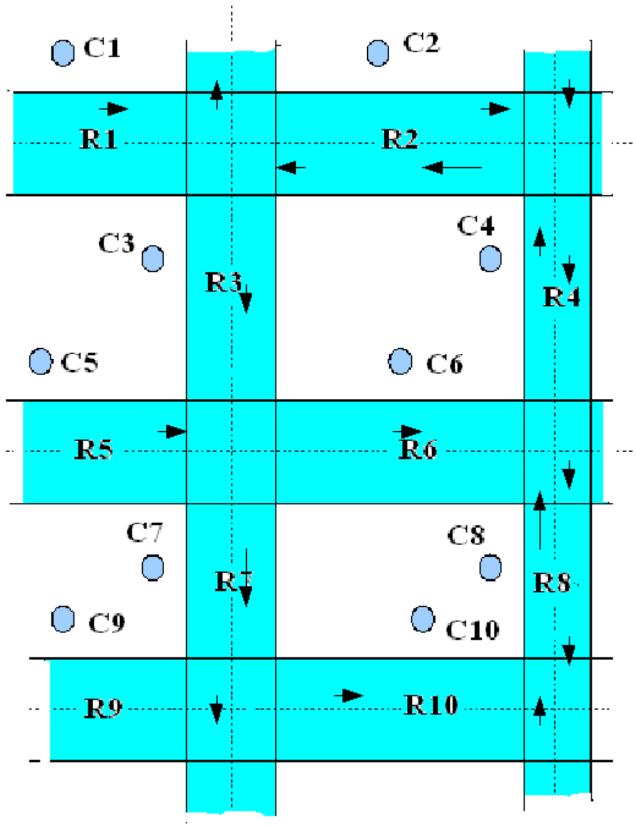


Figure 1: Road Network

the activities on the road links respectively. The video features of each road link will be extracted, processed, predicted and communicated to the nearby nodes periodically by each node. Each node will be sending the predicted states of the road and neighboring road links along with the ID of the respective road link for updating the current status.

Traffic on any road can be completely defined by the number of moving vehicles and their average velocity. But these two features depend on each other. Therefore we classify the road states by comparing the no. of vehicles with the number of moving vehicles. If the no. of vehicles in a road link is η and the no. of moving vehicle is γ, we can classify the road state $\eta\gamma$ domain. The major classification considered is: Stopped (S), Heavy congestion (HC), Mild congestion (MC), Slight Traffic (ST), and Open (O) We can extract SIP and STIP from the video frames recorded by the camera. The classification of the traffic state can be done based on the ratio of STIP to SIP as the ratio will give the indication of the state of the road. This is possible only when we get a dense feature set of correlated SIP and STIP, whereas the existing operators are providing sparse feature set of uncorrelated SIP and STIP. Hence we made a novel spatial interest point detector which provides dense feature set of correlated points by modifying the Harris corner detector.

The parametric model of the saliency features distribution is learned by fitting a Gaussian Mixture Model (GMM) using the Expectation-Maximization (EM) algorithm on a hand labelled training data set. This approach is used because it is suitable for fast data processing with only three features, weight, mean and covariance.

Using the spatal (η_{sp}) and spatio-temporal interest (γ_{pt}) points we can classify the road state by using Gaussian mixture model given by equation 1

Let the feature vector be $S = (\eta_{sp}, \gamma_{pt})$, no. of classes be k and no. of Gaussian mixtures be n

The conditional probability that belongs to Kth class is given by

$$p(S/M^k) = \sum_{i=0}^{n^k} \frac{w_i^k}{2\Pi^{d/2}|\ \Sigma_i\ |^{1/2}} e^{(S-\mu_i^k)^T \Sigma_i^k (S-\mu_i^k)} \quad (1)$$

The traffic on road links can be universally mapped into any of the above traffic states on a road link. Hence the traffic on a road link can be considered as a finite state machine transiting from one traffic state to another with time. The problem with such a model for traffic prediction is that it is too rigid and allows only small changes in traffic state distribution and brings up the need of learning 2 Dimensional finite state machines, with states transiting with variation in space and time. In order to make the system self evolving and avoid the complications of 2 dimensional finite state machine learning, we turn to Hidden Markov Models for traffic modeling. The HMM so learnt, does soft classification after adjusting to lower minima for that road link and also evolve through time while learning online and can predict the future state. The flowchart of the proposed Traffic model is given in fig.2.

Once we get the future state of neighboring road links, we can assign weight to each state and the optimal route can be determined. The system is energy efficient as the sensor nodes have to communicate only the traffic state periodically. Driver of the vehicle can communicate to the nearby sensor node and can select his route.

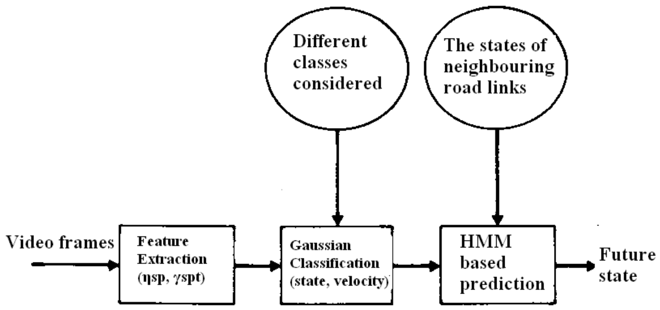


Figure 2: Traffic prediction Model

4. TRAFFIC MODEL

4.1 Road Traffic States

Traffic on a road segment can be completely defined on the basis of:

- Number of vehicles on the road segment(η).
- Number of moving vehicles (γ).
- Average velocity of the vehicles (v).

We can define each and every road state as an area in domain of η and v by mapping them into a certain fixed average number of vehicles and a certain average velocity of vehicles with a given fuzziness. These road states, if mapped accordingly can correspond to road states as we define cognitively like open, slight traffic, mild congestion, heavy congestion, jam etc. The states into which a road link can be classified vary with different road links as the two properties of a road segment, ie. η and v are not completely independent. The number of vehicles on a given road link affects the average time taken to traverse it; similarly the average time taken to traverse a road link affects the number of vehicles on it. We take advantage of the above fact to minimize our feature vector by defining the road state, by comparing the number of vehicles to the number of moving vehicles on the road. These two parameters are estimated by extracting spatial interest points (SIP) and spatio-temporal interest points (STIP) as the number of spatial interest points indicates the number of vehicles on the road link and number of spatio-temporal interest points indicates the number of moving vehicles on the road. Using SIP and STIP, the road states are classified in the $\eta\gamma$ domain using Gaussian mixture models and Maximum Likelihood.

4.1.1 *SIP and STIP extraction*

The currently available $STIP$ detectors [17] [8] [11] like Harris corner detector, determines the second moment matrix over image derivatives used to to achieve rotational invariance. As it is not required in the proposed method, the computational burden of calculating second moment matrix over image derivatives is avoided and direct second moment matrix over intensities is obtained. This helps in making the system computationally light and the STIP so generated is a very crude form but optimized to work in real time on low computational power systems. Therefore we found spatial interest points by recording the intensity values along each dimension in a local neighborhood and then found the Eigen values of recorded 2 dimensional data. This concept is then further extended to temporal domain. This way we obtained spatial and corresponding spatio-temporal interest points with detection of minor changes in intensity over temporal domain. The fig. 3 shows the SIP and STIP detected by our algorithm

Let us define the image as $f^{sp}(x, y)$. To remove high frequency variation over the spatial domain we convolve the image with a 2D Gaussian kernal with variance σ_1^2 to obtain $F^{sp}(x, y)$.

$$F^{sp}(x, y) = f^{sp}(x, y) * g^s p(x, y, \sigma_1) \quad (2)$$

The spatial interest points are obtained by listing the intensities at points $F^{sp}(x-1, y)$, $F^{sp}(x, y)$, $F^{sp}(x+1, y)$ along a dimension I_X and points $F^{sp}(x, y-1)$, $F^{sp}(x, y)$, $F^{sp}(x, y+1)$ along a dimension I_y.

$$I_x = \begin{bmatrix} F^{sp}(x-1, y) \\ F^{sp}(x, y) \\ F^{sp}(x+1, y) \end{bmatrix} \quad (3)$$

$$I_Y = \begin{bmatrix} F^{sp}(x, y-1) \\ F^{sp}(x, y) \\ F^{sp}(, y+1) \end{bmatrix} \quad (4)$$

$$l_{sp} = Cov(I_x, I_Y) \quad (5)$$

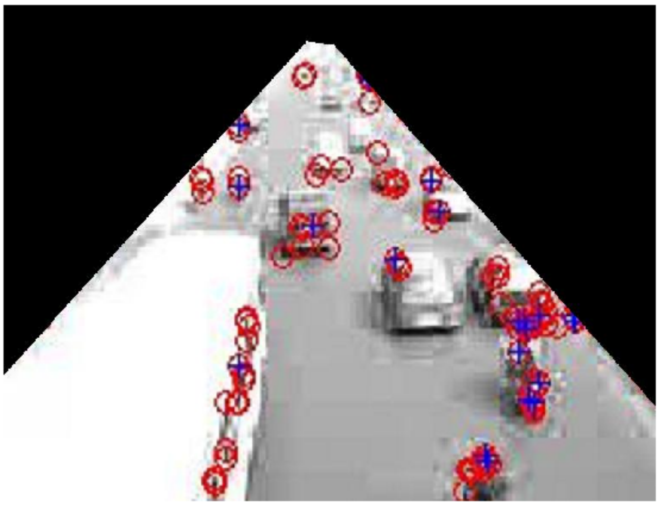


Figure 3: The red circle shows SIP and blue cross shows STIP

$$| l_{sp} |> th_{sp} \quad (6)$$

Now, determinant of l_{sp} will give us the product of Eigen values, this product is indicative of variance along the two principal directions, hence in our operator the $\det(l_sp)$ denotes the strength of spatial interest point. In our case we define a strength threshold (th_{sp}), above which all the points are considered spatial interest points

For spatio-temporal domain the consecutive smoothened frames are stacked to form $F^{sp}(x, y, t)$. And I_x, I_y and I_t are expressed as

$$I_x = \begin{bmatrix} F^{sp}(x-1, y, t) \\ F^{sp}(x, y), t \\ F^{sp}(x+1, y, t) \end{bmatrix} \quad (7)$$

$$I_Y = \begin{bmatrix} F^{sp}(x, y-1, t) \\ F^{sp}(x, y, t) \\ F^{sp}(x, y+1, t) \end{bmatrix} \quad (8)$$

$$I_t = \begin{bmatrix} F^{sp}(x, y, t-1) \\ F^{sp}(x, y, t) \\ F^{sp}(x, y, t+1) \end{bmatrix} \quad (9)$$

$$l_{spt} = Cov(I_x, I_Y, I_t) \quad (10)$$

$$| l_{spt} |> th_{spt} \quad (11)$$

A spatial temporal interest point should satisfy both the equations. The resulting operator gives a dense representation of a video with strength of each point defined crisply in both spatial and temporal domain independently. Such points can be used to identify moving spatial interest points and hence detect moving objects in the field of view. The next section concentrates on using the above defined spatial and spatial interest points as features for road traffic state classification.

4.1.2 Classification

The normal traffic situation can be roughly categorized into two states, open and congestion. But we observe that such a classification is not enough to describe the traffic situation. Thus, we use traffic patterns similar to what humans define; Stopped (S), Heavy congestion (HC), Mild congestion (MC), Slight Traffic (ST), and Open (O). They are defined as follows. **Stopped**: there are a large number of vehicles and almost all the vehicles run very slowly or are completely stopped. **Heavy congestion**: there are a large number of vehicles and most vehicles run slowly, **Mild congestion**: most of the vehicles run at half speed. **Slight Traffic**: vehicles run at normal speed. **Open**: there is no vehicle or minimum number of vehicles in the region of interest. It is important to note that for different road links, the above defined traffic state will map to different region in the $\eta\gamma$ domain. Hence for every road link the corresponding parameters and their feature set mapping has to be learnt for every class or traffic pattern. To avoid this problem we normalized the feature set by considering the no. of vehicles per unit area. The normalized feature set will be $(\eta_{sp}\gamma_{spt})$ where η_{sp} and γ_{spt} will be no. of SIP and no. of STIP per unit area. The classified set will give an indication of average velocity (v). By doing so we can do offline learning once, which can be further used for real time application.

Classifier.

A parametric model of the saliency features distribution is learned by fitting a Gaussian Mixture Model (GMM) using the Expectation-Maximization (EM) algorithm on a hand labeled training data set. This approach is used because its properties are well-known and it is suitable for fast data processing with only three features. See [1] for practical details on the EM algorithm and [6] for classification using GMM. The resulting density probability model for each class is the sum of n Gaussians with weight, mean and covariance matrices $(w_i^k, \mu_i^k, \Sigma_i^k)$i=1...n.

Let $M^k = (w_i^k, \mu_i^k, \Sigma_i^k)$i=1...$n^k$ be the GMM of the kth class. Let$S = (\eta_{sp}\gamma_{spt})$ be the saliency features of a road state to be classified. The conditional probability that the point belongs to the kth class is given by equation(1) and

$$k^{max} = argmax(p(S/M^k)) \quad (12)$$

Using equation(1) and equation(12), the class or the road state and velocity can be suggested.

Training Model.

In order to capture the variability of the traffic states, the training set must contain data from all the traffic states defined. The labeling process is performed using a graphical interface developed in matlab which allows the selection of frames individually or in groups. We enforce a balanced labeled dataset between the different traffic states. In order to keep the high frequency change in traffic from influencing the state classification, averaged features are computed over a fixed number of frames. This labeling and model fitting is performed off-line and only once. Once a model is obtained, it can be used to classify road states in real time.

4.2 Traffic Prediction

4.2.1 Traffic as finite state machines

The traffic on road links can be universally mapped into any of the above traffic states on a road link. Hence it can be considered as a finite state machine transiting from one traffic state to another with time. As the states are transiting with variation in space and time, we need to learn two dimensional finite state machines for traffic prediction. To avoid this complication and to make the system self evolving, we use Hidden Markov Model for traffic monitoring / prediction. The HMM so learnt, does soft classification after adjusting to lower minima for that road link and also evolve through time while learning online. A hidden Markov model (HMM) is a statistical model in which the system being modeled is assumed to be a Markov process with unobserved state and it can be considered as the simplest dynamic Bayesian network. In a regular Markov model, the state is directly visible to the observer, and therefore the state transition probabilities are the only parameters. In a hidden Markov model, the state is not directly visible, but output, dependent on the state, is visible. Each state has a probability distribution over the possible output tokens. Hence while expectation maximization or learning, the HMM trains itself to optimize the probability distribution of output over states to a local minima and also learning the transition probabilities. [13] and [1].

4.2.2 Hidden Markov Models for traffic prediction

Let $R-1$, R,$R+1$ be the three neighboring road links, linked together, T_{R-1}, T_R, T_{R+1} be the traffic states defined at $R-1$, R,$R+1$ with $T \equiv (T^1, T^2, T^3 ... M)$ by GMM based classification and T_{R_n} denote the traffic state at n^{th} time.

For learning the 2 dimensional relation (spatial and temporal) of the traffic states, we use Hidden Markov Model with an assumption that for very short interval of time (refresh time of our system), the traffic condition at a given road link change only due to the traffic conditions on neighboring road links.

Therefore, we define an HMM state for road link to be $T_{R-1_n}, T_{R_n}, T_{R-1_n}$

Hence there are M3 states for a HMM for a road link. Using expectation maximization, we learn the transition matrix of dimension $M^3 \mathrm{X} M^3$ given by

$$TM = \begin{bmatrix} P(T_0/T_0) \ ... \ P(T_0/T_M) \\ .... \ .... \ ..... \\ .... \ .... \ .... \\ P(T_M/T_0) \ ... \ P(T_M/T_M) \end{bmatrix} \tag{13}$$

learning.

We take the learnt GMMs as our class definitions, or class distribution initially. Then the road is kept under observation by the system for sufficient time to generate data, for learning transitions and optimizing state distributions. The temporal flow of states on the road is recorded while it is under observation using state definitions described earlier. The data so generated is used to learn HMM parameters through Balm - Welch algorithm. A compact description of the learning procedure is defined below:

- The initial state distributions represented as $B \equiv (b_j(.))$ are given by GMMS described in the classification section.
- These state distributions are used to observe and record the temporal flow of observed features on a road, using cameras and features described in section 4.1. Let this data or flow of features be represented by $O \equiv (O_t)$.
- The transition matrix for state distributions is initialised randomly and is represented by A.
- The initial state distribution is again given by GMMs learnt previously and is represented by $\pi \equiv (P(Q_(x_0)^i))$.
- $\lambda^* = argmax(P(O/\lambda^*))$ is found by Baum-Welch algorithm, where $\lambda = (A, B, \pi)$. The λ^*, so obtained has tuned parameters as per the information available.

The same process can be followed periodically, making the system defined as self evolving and changing as the road conditions change. The transition matrix and state distributions are hence learnt on periodically making the system adaptive.

Prediction.

Using the transition matrix and from the learnt HMM we sample the next state for the road link
$[T_{R-1_{n+1}}, T_{R_{n+1}}, T_{R+1_{n+1}}]$

The priors can be calculated by marginalizing any two road links; we can compute the probability of future states, using the formula-

$$P_R(T^i_{R_n}) = \sum_{j=1}^{M} \sum_{k=1}^{M} P_R(T^j_{R-1_n} T^i_{R_n} T^k_{R+1_n}) \tag{14}$$

and

$$P_R(T^i_{R_n}) = \sum_{g=R-1}^{R+1} P_g(T^i_{R_n}) \tag{15}$$

After calculating the probabilities of future traffic states of road links, we calculate the priors for next HMM sampling

$$P_R(T^j_{R-1_{n+1}} T^i_{R_{n+1}} T^k_{R+1_{n+1}}) = P(T^j_{R-1_{n+1}} T^i_{R_{n+1}} T^k_{R+1_{n+1}}) \tag{16}$$

Using the priors calculated the next state is sampled using Monte Carlo method. The GMMS learnt during classification are used to initialize the HMM states or provide the HMM states an initial belief as the system runs , it learns online using expectation maximization to tune the GMMs learnt for each to local minima , hence making the system self evolving and adaptive through online learning.

cost of the road links.

The cost of the road link can be calculated as

$$C_i = \Sigma_{i=x-1}^{x+1} P_i(T^i_{R_n}) R_i w_i \tag{17}$$

Where R_i and w_i are the road length and weight of the road class specified by the predicted state $P_i(T^i_{x_n})$ of the corresponding road link respectively. Time taken by a vehicle to traverse a road path can be predicted like this and hence the best path towards a destination can easily be suggested using cost based optimal path algorithms.

5. RESULTS

Figure 4: Each column illustrates a road traffic state with the region of interest bounded in yellow, spatial and spatial temporal points in red and blues respectively. From left to right each column indicates open, slight traffic, mild congestion and heavy congestion road states.

5.1 Classification

We evaluated the method suggested using the data collected from real traffic scenes. The data set includes various illumination conditions, e.g. sunny, overcast and night time. The video data is low resolution (172X180) and is taken at 4fps in grayscale. All testing clips are hand-labeled to make a comparison with a ground truth. The training video data is chosen such that there is no overlap with the testing data. The total length of the training data is about 120 minutes. Figure 4 shows the four traffic states, that is open, slight traffic, mild congestion and heavy congestion at a road link. The Region of Interest is marked by yellow boundary. The red points indicate spatial interest points and the blue points indicate spatial temporal interest points generated using the STIP operator. Interestingly, as the traffic state changes from open to heavy congestion, the number of spatial interest points become more as compared to number of corresponding spatial temporal interest points, hence further validating our feature selection .It is visible that all of the existing traffic states are successfully detected. We compared our results with the hand labeled ground-truth. When we examine the 'false' classifications given by the technique, an interesting fact is found that the system is more receptive to traffic changes and suggests a continuous state change, which is more intuitive or common to observe in real traffic than abrupt changes as suggested by human operator. Even if we consider all the states defined by the operator to be true, the classifier shows around 84% accuracy in determining the correct states. The optimum number of Gaussians used for modeling the data is learnt by testing for the classification rates for different number of Gaussians and most accurate results reaching saturation are obtained when 2 Gaussians used to represent a single traffic state. Table 1 and Table 2 shows the confusion matrix of classification results of road at Rajouri Garden and Moti bagh, Delhi.

The corresponding road parameters that is the number of vehicles and average velocity, for every traffic state for a given road link under observation are learnt by training a GMM over the labeled data with number of vehicles per unit area and average velocity over average velocity of road link as inputs. The most probable number of vehicles and average velocity for a given state is designated as the road

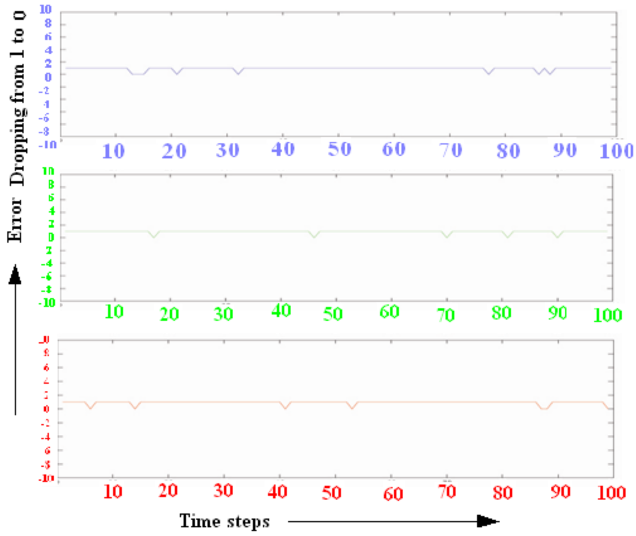


Figure 5: graph of predicted road states. Red , green and blue lines depict the prediction error plot with respect to time for different road links. The drop to value to zero indicates error in state prediction. For hundred time steps.

parameters when it is belonging to a certain traffic state.

5.2 prediction

The prediction technique was tested in real road conditions as well as simulation. The cameras were installed at foot over bridges at Motibagh to get time synchronised videos. Videos at Rajouri Garden , Delhi were used to form prior GMMs. These GMMs were then successfully tested at Motibagh and then were given as priors to HMM. It was observed that 89.891% of the predicted states were correct and predicted time taken to cross the road stretch under observation normalised over actual time taken was centered at 0.96373 with a variance of 0.1052. Even better results were produced in simulation where 94.026% of the predicted

Table 1: confusion matrix on the road at Rajouri Garden, Delhi

	Open	Slight Traffic	Mild Congestion	Heavy Congestion
Open	86.11521	3.148855	12.64479	3.374447
Slight Traffic	2.215657	81.10687	0.22522	0.698162
Mild Congestion	5.317578	9.446565	73.9704	4.212241
Heavy Congestion	6.351551	6.29771	13.15959	91.71515

Table 2: confusion matrix on the road at MotiBagh, Delhi

	Open	Slight Traffic	Mild Congestion	Heavy Congestion
Open	80.93	9.49	0	2.3
Slight Traffic	8.53	76.54	0.18	0
Mild Congestion	9.01	16.04	99.82	1.73
Heavy Congestion	1.52	1.08	0	95.97

states were correct and predicted time taken to cross the road stretch under observation normalised over actual time taken was centered at 1.0238 with variance of 0.0711. Error graph of predicted road states is shown in figure 5

6. CONCLUSION

We presented a novel method for monitoring / prediction using the camera network. The method is simple, real time and computationally light. We also present detailed results on road state classification. The method presented is robust to light changes, adaptive to varying road topologies under observation, camera set up independent and requires no pre-processing; the requirements of the method presented are very low with successful testing on resolution as low as 172X180 and frame rate of 4fps. The low demands on video quality of data, suggests that the method can also be implemented on relatively low quality satellite imagery. The selection of states makes it possible to learn a 2 dimensional relation of road states using HMM. The HMM so used makes the system self evolving or self adjusting which could not have been the case with Finite state machine model. The prediction results are satisfactory and indicative of the suitability of the model used. It is worth noting that the model used in this paper does not assume or draws analogies of traffic moving as particles, neither does it impose restriction on road conditions or road tributaries and distributaries, resulting in a robust and adaptive framework of real time traffic classification and prediction.

7. 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. *TR-97-021, International Computer Science Institute Berkeley CA, 94704*, April 1998.

[2] R. Boel and L. Mihaylova. A particle filter for freeway traffic estimation. *IEEE Symposium*, June 2004.

[3] B.-J. Chang1, B.-J. Huang, and Y.-H. Liang. Wireless sensor network-based adaptive vehicle navigation in multihop-relay wimax networks. *22nd International Conference on Advanced Information Networking and Applications*, pages 56–63, June 2008.

[4] P. Cheng, Z. Qiu, W. Qiu, and B. Ran. A particle filter based model for traffic state estimation using simple wireless network data. *IEEE International Conference on Service operation and logistic, and informatics*, 2005.

[5] R. Cucchiara, M. Piccardi, and P. Mello. Image analysis and rule-based reasoning for a traffic monitoring system. *IEEE Transaction on Intelligent Transportation System*, 1(2):119–129, June 2000.

[6] R. Duda, P. Hart, and D. Stork. *Pattern Classification.* Wiley Interscience, 2000.

[7] Y.-K. Jung, K.-W. Lee, and Y.-S. Ho. Content-based event retrieval using semantic scene interpretation for automated traffic surveillance. *IEEE Transaction on Intelligent Transportation System*, 2(3):151–163, September 2001.

[8] W. Kienzle1, B. Schĺolkopf1, F. A. Wichmann, and M. O. Franz1. How to find interesting locations in video a spatiotemporal interest point detector learned from human eye movements. *22nd International Conference on Advanced Information Networking and Applications*, pages 56–63, June 2005.

[9] X. Li and F. M. Porikli. A hidden markov model framework for traffic event detection using video features. *IEEE International Conference on Image Processing*, 2004.

[10] B. Maurin, O. Masoud, and N. Papanikolopoulos. Monitoring crowded traffic scenes. *IEEE Int. Conference on Intelligent TransportationSystems*, pages 19–24, June 2002.

[11] K. Mikolajczyk and C. Schmid. Scale and affine invariant interest point detectors. *International Journal of Computer Vision*, 60(1):63–86, 2004.

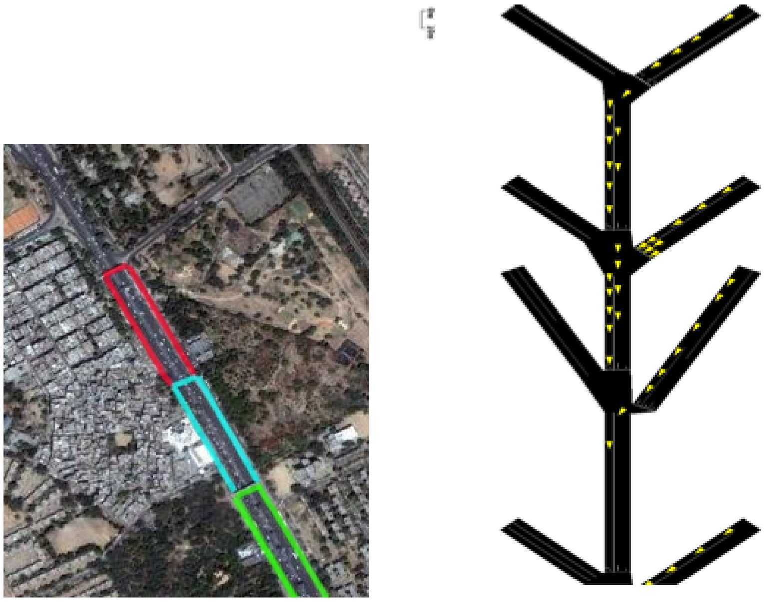
[12] Y. Qi. Probabilistic models for short term traffic condition prediction. *Ph.D. Thesis, M.Ap.Stat., Louisiana State University,*, May 2010.

[13] L. R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *In Proceedings of the IEEE*, 77:257–286, 1989.

[14] D. Terzopoulos and C. Brown. Model-based vision for traffic scenes using the ground-plane constraint. *Real-time Computer Vision Cambridge University Press*, 1994.

[15] M. Tubaishat, Y. Shang, and H. Shi. Adaptive traffic light control with wireless sensor networks. *1-4244-0667-6/07/ IEEE*, pages 187–191, June 2007.

[16] V. K. Verma, R. Choudhari, S. K. Singh, T. M. Abhai Pratap Singh, and P. Srivastava. Intelligent



(a) (b)

Figure 6: (a) satellite image of Mahatma Gandhi Road near Motibagh, Delhi (b) partial image of a simulated network generated in Simulation of Urban Mobility software (SUMO) developed by German Aerospace Centre (DLR).

transport management system using sensor networks. *IEEE Intelligent Vehicles Symposium Eindhoven University of Technology Eindhoven*, pages 4–6, June 2008.

[17] G. Willems, T. Tuytelaars, and L. V. Gool1. An efficient dense and scale invariant spatio temporal interest point detector. *22nd International Conference on Advanced Information Networking and Applications*, pages 56–63, June 2008.

[18] Y. Zou, G. Shi, H. Shi, and H. Zhao. Traffic incident classification at intersections based on image sequences by hmm/svm classifiers. *IEEE International Conference on Hybrid Intelligent Systems*, August 2009.