TRAFFIC LEVEL OF SERVICE GENERATION FROM VIDEO DETECTION SYSTEM USING CLUSTER ANALYSIS

Alexander Sohr, German Aerospace Center (DLR), alexander.sohr@dlr.de

Xiaoxu Bei, German Aerospace Center (DLR), xiaoxu.bei@dlr.de

Jiao Zou, The Research Center for Software Engineering Technology of Anhui Province (ASEC), zoujiao1986@sina.com

Jian Wu, The Research Center for Software Engineering Technology of Anhui Province (ASEC), hfwujian@163.com

Jiajie Wang, The Research Center for Software Engineering Technology of Anhui Province (ASEC), jiashengye@vip.qq.com

ABSTRACT

In general, traffic service quality can be described by different parameters and the traffic level of service (LOS) is mainly determined by one of them, such as traffic flow or traffic speed. However, the traffic phenomenon is extremely complex; its characteristics cannot be well represented only by one single parameter. The aim of this research is to determine a more representable LOS based on three measured parameters, derived from the local traffic video data collected in Hefei, China. The result of this work will be used for data fusion with floating car data (FCD) and other traffic information in order to more accurately capture the current traffic situation of the city.

Keywords: Traffic State Estimation, Level of Service, Cluster Analysis, Data Mining, Video Detection

INTRODUCTION

Rea-time urban road traffic state identification is an important premise of intelligent transportation systems and is the basis for dynamic navigation and traffic guidance. Furthermore it has an important role for traffic managers. The accuracy of the applied algorithm to identify the urban road traffic state directly affects the traffic condition of the entire city road network. The wrong estimation may lead to inaccurate route guidance as well as new congestions. The classic traffic state identification algorithms, such as California algorithm and Mcmaster algorithm, focus only on one single traffic parameter to set the thresholds for classifying the traffic state. Urban road traffic phenomenon is extremely complex, and cannot be simply described by one single parameter. To more accurately describe traffic conditions more traffic parameters should be considered and integrated in the LOS classification.

In Chinese cities, intelligent transportation systems are well equipped with stationary detection devices, for example: loop detectors, microwave detectors, geomagnetic detectors and video detection systems. Video detection systems are the most spread systems in Chinese cities. For the current research the data of one month collected by the Autoscop video detectors was used.

The collected data is first categorized into working days and holidays. Then, following the relevant provisions of the Highway Capacity Manual, the Cluster-analysis was used to classify the traffic states and identify the respective thresholds with the consideration of all available information (parameters) from the detector.

DATA ANALYSES

Autoscop video detector data and characteristic parameter extraction

The Autoscop video detectors are used in urban roads for stationary detection. They can detect multi-lane traffic parameters and also provide visual images as a good basis for model validation. Such a video detector mainly consists a multi-channel video-processing device, providing a virtual loop. Each continuous video image is converted to a discrete digital image, and then mathematically processed to extract the related traffic parameters. These traffic parameters include traffic volume, average speed, headway, occupancy and density.
Traffic state can be described by macroscopic parameters such as traffic volume, speed, density and queue length and microscopic parameters, e.g. time gap and distance between vehicles. Traffic parameter selection is the basis of the proposed “discriminant traffic state algorithm”, which will be explained below in detail.
After the analysis of the parameters and its variation, it was decided to use volume, average speed and occupancy as parameters for the algorithm to describe the traffic state, because these three parameters can describe most of the traffic state characteristics very well.

The traffic flow on working days often follows similar patterns, except incidents, and especially during typical morning and/or afternoon rush hours. On holidays traffic is not so highly concentrated during rush hours, but rather spread over the day. The respective trips tend to be shorter trips. In order to that two exemplary daily time series for a holiday and a working day were analyzed in this study. The video detector “camera013” in Hefei was used. Data from 2010 (Monday), September 6 and September 12, 2010 (Sunday) was used to analyze each parameter variation. Figures 1 to 3 show the comparison of the used three parameters (volume, speed and occupancy).

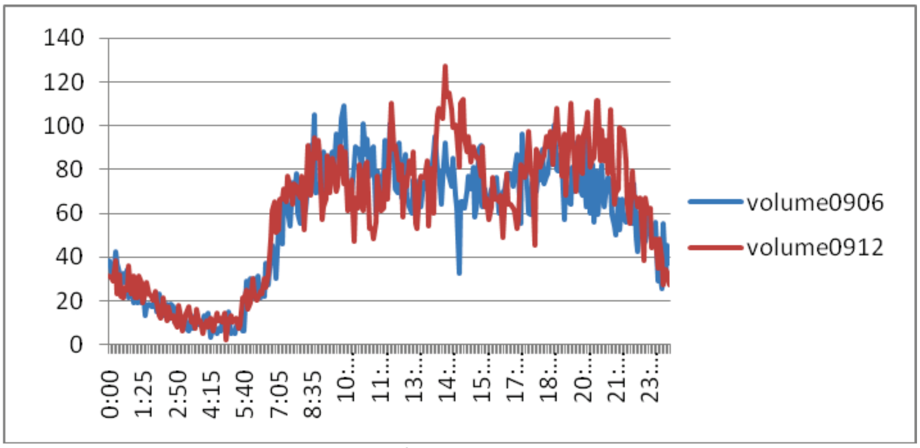

Figure 1 Daily course comparison of volume between a working day and a holiday

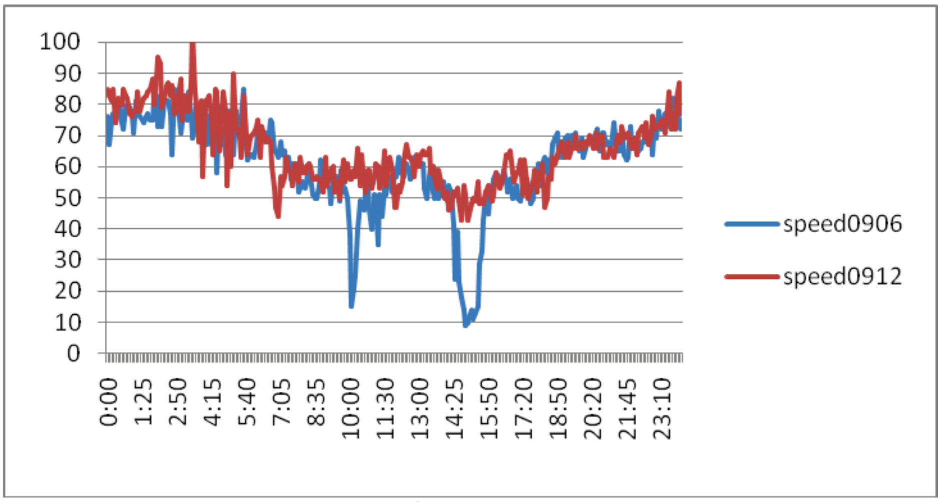

Figure 2 Daily course comparison of speeds between a working day and a holiday

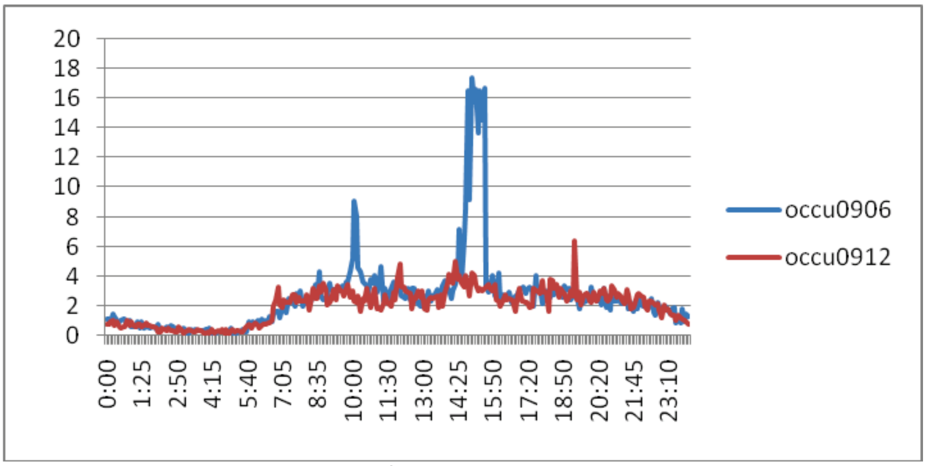

Figure 3 Daily course comparison of occupancy between a working day and a holiday

DISCRIMINANT ALGORITHM

In the following, the developed “discriminant traffic state algorithm” is described in detail.

Data preparation

In order to make the different data types comparable, the parameters are standardized to an interval between [0, 1], so a dimensionless processing is possible. Here the z-score standardized method is used. The z-score data standardization method is based on the raw data mean and standard deviation.
The video detector data of one month was used.

Assuming that the sample data is S , $S=\{(q_1,o_1,v_1),(q_2,o_2,v_2),\ldots,(q_n,o_n,v_n)\},\ (q_i,o_i,v_i)$ where each dataset contains q_i,o_i,v_i volume, occupancy and speed.
Then the mean and the standard deviation can be easily calculated by the following equations:

volume q: $\bar{q}=\sum_{i=1}^{n} q_i/n$ $\sigma_q=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(q_i-\bar{q})^2}$

occupancy o: $\bar{o}=\sum_{i=1}^{n} o_i/n$ $\sigma_o=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(o_i-\bar{o})^2}$

speed v: $\bar{v}=\sum_{i=1}^{n} v_i/n$ $\sigma_v=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(v_i-\bar{v})^2}$

(q_i,o_i,v_i) Normalized following format : $\left(\hat{q}=\frac{q_i-\bar{q}}{\sigma_q},\hat{o}=\frac{o_i-\bar{o}}{\sigma_o},\hat{v}=\frac{v_i-\bar{v}}{\sigma_v}\right)$

Cluster analysis Methods

Classic algorithms usually use fixed thresholds to classify traffic states. Thus they cannot fully reflect the traffic flow state and the correlation between the different parameters are seldom to be considered.

Therefore, cluster analysis is adopted here for data processing. The cluster analysis divides the dataset into cluster with different characteristics. All data points in one cluster have high similarity. Traditional cluster analysis algorithm can be divided into five categories: partitioning methods, hierarchical methods, density-based methods, grid-based methods and model-based methods. Here a partitioning method (PAM) is used, in particular the k-means algorithm with Euclidean distance as spacing function.
The K-means algorithm is one of the commonly used data mining algorithms. In detail an N sample collection is divided into K clusters (K < N), repeated through multiple cycles calculated. The distance is calculated for each point in the cluster to the cluster centre, so that the variance-sum within each cluster reaches a given threshold value. It has an objective function as follows:

$$J=\sum_{j=1}^{k}\sum_{i=1}^{n}\left\|x_i^{(j)}-c_j\right\|^2$$

$\left\| x_i^{(j)} - c_j \right\|^2$ is the distance between samples $x_i^{(j)}$ and centres of clusters c_j .

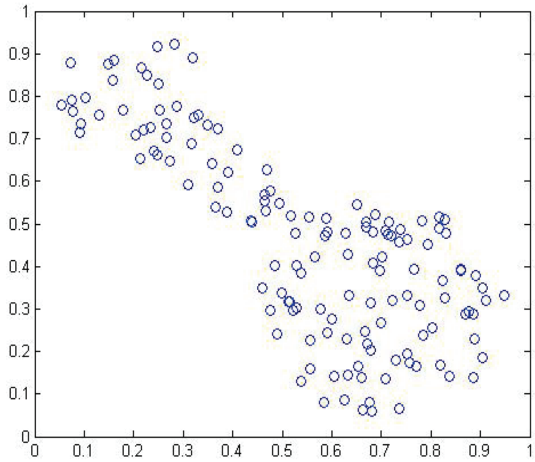

Figure 4 sample distribution before clustering

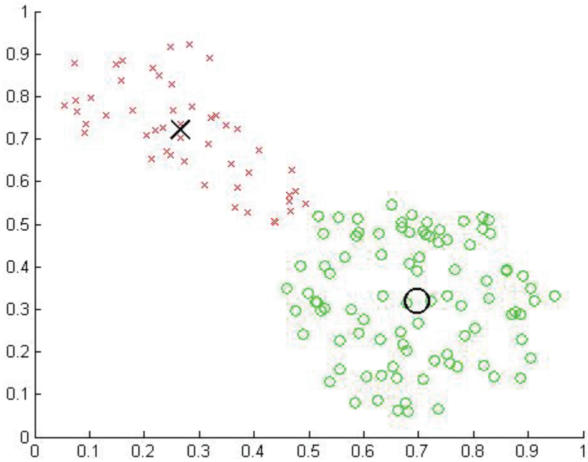

Figure 5 k-means clustering result

Model development

Every Detector (camera) is a single data source. The model uses the data of one month of each detector and divides it into two groups: working day and holiday.

Then for both groups the cluster centres are calculated. Accordingly to the Highway Capacity Manual four Clusters are generated, where each Cluster represents a certain Traffic state (free flow, unobstructed, stocked and jamed):

$(\tilde{q}_1, \tilde{o}_1, \tilde{v}_1)$, $(\tilde{q}_2, \tilde{o}_2, \tilde{v}_2)$, $(\tilde{q}_3, \tilde{o}_3, \tilde{v}_3)$, $(\tilde{q}_4, \tilde{o}_4, \tilde{v}_4)$ → One set of Cluster-centres for each group.

The Cluster centres are stored and for every online data point the Euclidean distance is calculated to each centre, to distinguish the traffic state. The steps are as follows:

(1) obtain a set of real-time data (q, o, v)
(2) Z-score standardization; $(\hat{q}, \hat{o}, \hat{v})$
(3) Calculating the Euclidean distance between $(\hat{q}, \hat{o}, \hat{v})$ and each cluster centre $(\tilde{q}_i, \tilde{o}_i, \tilde{v}_i)$;
(4) Pick the cluster with the shortest distance to assign the traffic state.

IMPLEMENTATION AND EVALUATION

The Implementation and Evaluation was done with the Autoscop video detector camera013 in Hefei City at the Changjiang Middle Road near Changfeng South Road. The camera detects the East-West direction (See Figure 4).

Figure 6 Position of “camera013” in Hefei

First of all, the Processing was implemented as described above and with one month of data the cluster centres –Matrix was calculated. For working days the Matrix is described as follows:

$$M(\hat{q},\hat{o},\hat{v}) = \begin{bmatrix} -1.23744 & 0.782024 & -0.86293 \\ 0.041053 & 0.471184 & -0.1843 \\ -0.793382 & -0.96536 & -0.586111 \\ 2.127761 & -2.69757 & 3.147024 \end{bmatrix}$$

Then taking the real time data (volume, speed and occupancy) and calculate the online traffic state following the “discriminant traffic state algorithm”.

The “real time” data is captured and processed in 5 minute intervals which results in 288 traffic states. The distribution of the traffic states for one video detector over the day are shown in Figure 7.

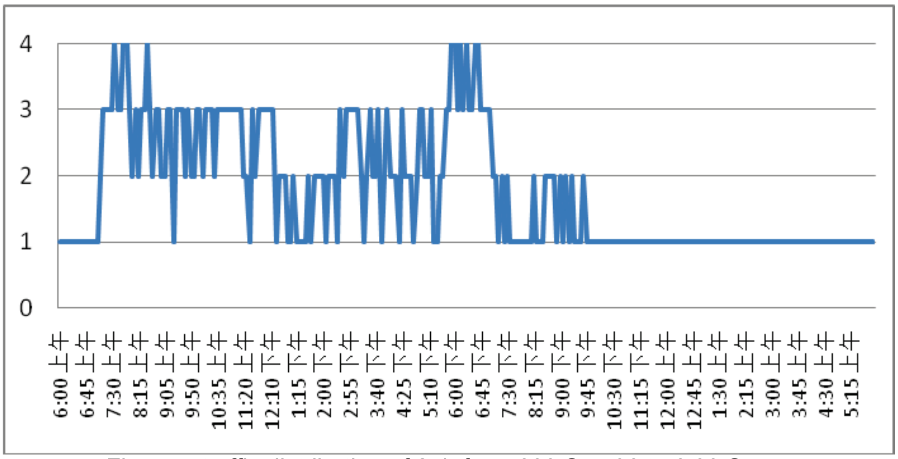

Figure 7 traffic distribution of 24h from 600 Sep.06 to 6:00 Sep.7

The plot for the workday in the example shows the morning and the evening rushours as well as the smooth traffic during night.

The program runs with high efficiency, usually in 0.025s all 120 Video-detectors data of Hefei are processed.

Evaluation

In order to evaluate the results of the algorithm, the estimated traffic states of one day (2010-09-06 6:00am to 2010-09-07 6:00am) were compared to visual observations of the Hefei traffic police. Figure 8 shows that for this day the results fit very well to the observations.

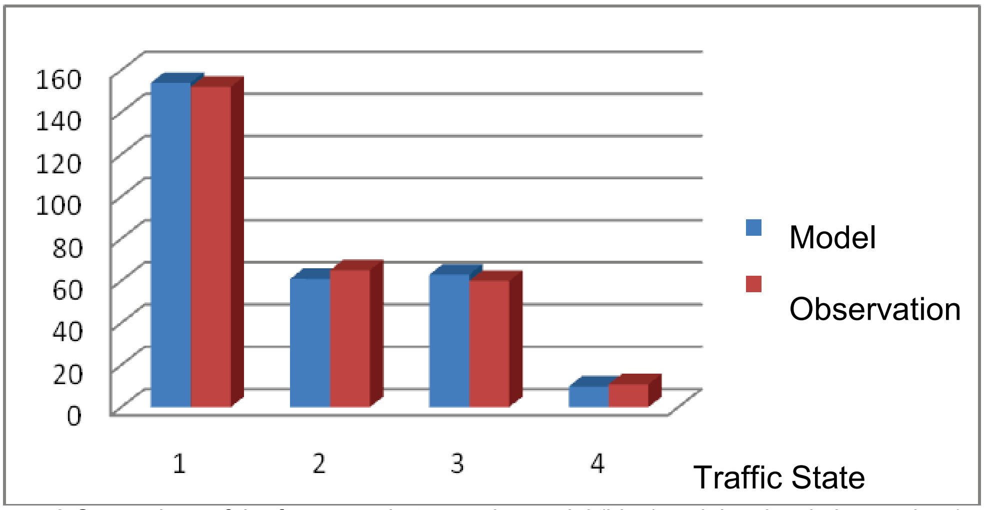

Figure 8 Comparison of the frequency between the model (blue) and the visual observation (red)

In order to more intuitive display the model discrimination results, screenshots of the video detection system were compared with the model results.

Figure 9 Comparison of traffic state between the model result and actual video

Figure 9 (top left) shows the video snapshot of 6:06am, the corresponding model discriminant traffic status value is 1 (free flow) It can be seen, that this matches. Figure 9 (top right) the video snapshot of 6:32pm is shown, where the corresponding model discriminant traffic state value is 4 (jammed). It is obvious, that the cars in the picture do not move fast and the traffic is congested.
In summary, the estimated traffic states processed by the developed system do match to the observations based on the Highway capacity manual and the algorithm can be used in production system at Hefei police department.

CONCLUSION

In Chinese urban intelligent transportation systems video detection is used for traffic control and dynamic real-time traffic status information is usually based on mobile detection like Floating Car Data (FCD). But FCD's accuracy depends on the number of connected vehicles. Small FCD-fleets lead to uncertainties and noisy data. therefore the developed algorithm can support the traffic information at this points where video detection is available. With that fusioned Traffic information a better Traffic Management and Information providing is possible.

REFERENCES

Martin, P.T. , J.P. Perrin, and P.B. Hansen, (2001) Incident Detection Algorithm Evaluation , pp.10.

Persaud, B.N. , F.L. Hall and L.M. Hall, (1990) Congestion Identification aspects of the McMaster Incident Detection Algorithm, in Transportation Research Record 1287, pp. 167-175.

Kanungo, T, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman and A. Y. Wu: An efficient k-means clustering algorithm: Analysis and implementation. In: IEEE Trans. Pattern Analysis and Machine Intelligence. 24, 2002, S. 881–892.

Sohr, A. and E. Brockfeld, (2011) Current and future use of Floating Car Data. 4th International Transportation Systems Performance Measurement Conference, 18.-20. Mai 2011, Irvine, U.S.A

Highway Capacity Manual. Transportation Research Board, Washington, D.C.. 2000. ISBN 0-309-06681-6