Measurement of traffic parameters using video cameras

Kristian Kovačić, Edouard Ivanjko, Niko Jelušić and Hrvoje Gold
Department of Intelligent Transportation Systems
Faculty of Transport and Traffic Sciences
University of Zagreb
Email: kkovacic@fpz.hr, eivanjko@fpz.hr, njelusic@fpz.hr, hgold@fpz.hr

March 12, 2015

Technical report: UNIZG-FTTS-ZITS-VISTA-03-15

March 2015

Department of Intelligent Transportation Systems
Faculty of Transport and Traffic Sciences
University of Zagreb
Zagreb, Croatia

Contents

1 Introduction

The last century can be characterized with a significant increase in the number of road vehicles and road infrastructure build-up. At the same time various systems for a road traffic management have been developed in order to increase the road traffic safety, capacity and comfort. Mentioned management systems require a large amount of high quality measurement data in real time. This requirement is especially important for management systems in scope of intelligent transportation systems (ITS). Basic architecture of management systems in ITS is made of a component for logical decisions, an execution component and a measurement component. The component for logical decisions chooses actions which need to be executed with respect to measured traffic parameters. Chosen actions are passed to the execution component which acts upon the road traffic network to change its state. Possible actions can be traffic light control, virtual message signs control and information forwarding to the emergency services in incident situations, etc. After the actions have been executed, new state of the traffic system is observed with the measurement component and forwarded to the component for logical decisions. This report describes the architecture of the traffic parameter measurement component in the scope of ITS. The emphasis is on video cameras as a sensor for traffic parameters measurement. In this report basic approaches for vehicle detection in a video footage of a road traffic network are described, and the results of their quality analysis are given.

In traffic analysis of road traffic networks, various parameters can be monitored. They include traffic flow, distance between vehicles, vehicle velocity, vehicle class (motorcycle, light vehicle, heavy vehicle), etc. Measurement methods can be based on various sensors such as inductive loops, radars, video sensors, etc. Video camera is starting to gain more and more importance and frequency of use in road traffic monitoring. The main reason for this is the possibility to use only one camera for measuring multiple traffic parameters. Results of such a traffic analysis can be used for traffic planning and management in urban and rural areas including highways.

Increase of a computing power and decrease of CMOS (Complementary Metal Oxide Semiconductor) based video cameras prices have made the use of computer vision algorithms and video cameras more frequent in today's road traffic measurement and management systems. Cameras are used as a part of the vehicle or road infrastructure. They enable monitoring of a road infrastructure, detection of incident situations, vehicle tracking in the vicinity of a video camera, etc. In urban areas, video cameras can be used together with computer vision algorithms for traffic management in scope of ITS. Mentioned application includes adaptive control of light signalization, vehicle queue measurement, vehicle classification, etc. On highways, video cameras can be used in combination with license plate recognition (LPR) algorithms for statistical analysis such as vehicle country of origin distributions or Origin-Destination (OD) matrices [1]. Mentioned analysis is useful for transit traffic management and improvement of road traffic management in the touristic season.

2 Traffic flow parameters

Traffic flow is a simultaneous motion of multiple traffic entities (vehicles, trains, pedestrians) in a traffic infrastructure (road, railway, footpath) which behaves according to some laws. The traffic flow theory is a science which studies vehicle motion laws in a traffic flow. Vehicle motion in a traffic flow depends on various traffic parameters. Some of more important traffic parameters are: quantity of traffic flow, traffic flow characteristics, dynamic characteristics of a driving vehicle, psychological and physical characteristics of the driver and his motivation, characteristics of traffic control, and management system and environment conditions (visibility, road conditions, weather). conditions) [2].

In this report traffic parameters are divided into two groups:

- Individual vehicle motion parameters;
- Traffic flow parameters.

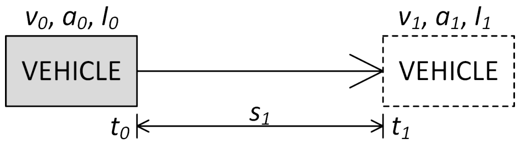

Figure 1: Individual vehicle motion parameters.

Systems for measuring traffic flow parameters based on processing an image from a video camera can measure both groups of traffic parameters using appropriate computer vision algorithms. Complex parameters of traffic flow can be obtained by computing basic motion parameters of an individual vehicle.

2.1 Individual vehicle motion parameters

The term *"individual vehicle motion parameters"* implies movement of a single vehicle on a route with the maximal safe velocity, no interference with other vehicles on the route and under the assumption that its movement depends only on road characteristics [3]. Parameters which can be used for defining a basic motion model of a individual vehicle are shown in Fig. 1 and are given below:

- Time t [s];
- Distance s [m];
- Motion direction ϕ [rad];
- Velocity v [${}^{m}/_{s}$];
- Acceleration a [${}^{m}/_{s^2}$];
- Impulse I [Ns].

By computing mentioned physical quantities for the specific time t_0, it is possible to predict a physical quantity in any future time t_n. Prediction can be made using a physical model of a vehicle which describes the change of a physical quantity in a time duration Δt defined by Eqs. 1-3.

$$I \quad = I(t_0) + \int_{t0}^{t} m \times a(u)du \tag{1}$$

$$v(t) = v(t_0) + \int_{t0}^{t} m \times a(u)du \tag{2}$$

$$s(t) = s(t_0) + \int_{t0}^{t} m \times v(u)du \tag{3}$$

If a vehicle location is defined in the Cartesian coordinate system, the vehicle motion model can contain the following physical quantities:

- X axis of coordinate system [m];
- Y axis of coordinate system [m];
- Tangential motion velocity v [${}^{m}/_{s}$];
- Tangential motion acceleration a [${}^{m}/_{s^2}$];

- Vehicle motion direction ϕ [rad or $^\circ$].

The existing model can be expanded with a change of motion direction $\Delta\phi$ by time t represented as ω [$^{rad}/_{s}$]. Mentioned model is defined by Eqs. 4 and 5, where X is vector which describes the model state (values of previously mentioned physical quantities) in a specific time, x and y are the vehicle location in the coordinate system [m], v is tangential motion velocity [$^{m}/_{s}$], a is tangential motion acceleration [$^{m}/_{s^2}$] in direction ϕ [rad or $^\circ$] which changes over time by ω [$^{rad}/_{s^2}$ or $^{\circ}/_{s^2}$], and t is time interval [s] between the time when last sample X was obtained and the current time for which a new state vector X will be computed [4].

$$X = \begin{bmatrix} X_x \\ X_y \\ X_v \\ X_a \\ X_\phi \\ X_\omega \end{bmatrix} \tag{4}$$

$$f(X,t) = \begin{bmatrix} X_x + X_v t cos(X_\phi) + X_a \frac{X_\omega t sin(X_\omega t + X_\phi) + cos(X_\omega t + X_\phi)}{X_\omega^2} \\ X_y - X_v t sin(X_\phi) + X_a \frac{X_\omega t cos(X_\omega t + X_\phi) - sin(X_\omega t + X_\phi)}{X_\omega^2} \\ X_v + X_a t \\ X_a \\ X_\phi + X_\omega t \\ X_\omega \end{bmatrix} \tag{5}$$

The system which can predict the location of a vehicle based on the Extended Kalman Filter (EKF) is proposed in [5]. EKF uses the vehicle location Λ_{ALP} in the update stage. Λ_{ALP} is a projection of the predicted vehicle location Λ_{Obj} in the previous stage of EKF on the virtual center line of a road lane obtained from a map. Vehicle motion direction ψ_{ALP} is interpolated by using map data also. For vehicle motion velocity, maximum allowed velocity values are used on a specific road segment. In Fig. 2, predicted vehicle location Λ_{Obj}, projected vehicle location Λ_{ALP}, predicted vehicle motion direction ψ_{ALP}, variance of projected vehicle location σ_η^2, variance of vehicle motion direction σ_τ^2, covariance matrix of vehicle motion location (projected), direction Σ_{ALP}, and vector for location (projected), and direction of vehicle motion X_{ALP} are given.

Song et al. in [6] have proposed a system which by using a road traffic video footage performs vehicle detection, vehicle trajectory monitoring and detection of irregular events such as forbidden lane change, full stop, aggressive acceleration or decelerations, etc. In order to detect mentioned events, vehicle velocity and

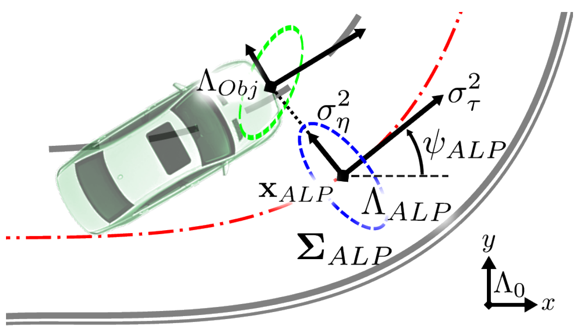

Figure 2: Projection of measured vehicle location on a virtual center line of a road lane.

Table 1: Measurement of accuracy for vehicle detection, tracking and trajectory analysing system.

Environment		Sunny	Cloudy	Rain	Night
Full stop	Detected / groundtruth $[vehicles]$	628 / 650	560 / 589	678 / 722	391 / 428
	Accuracy $[\%]$	96.6	95.1	93.9	91.4
Deceleration	Detected groundtruth $[vehicles]$	97 / 103	142 / 157	116 / 125	126 / 146
	Accuracy $[\%]$	94.2	90.4	92.8	88.4
Lane change	Detected / groundtruth $[vehicles]$	216 / 234	190 / 207	170 / 189	101 / 114
	Accuracy $[\%]$	92.3	91.8	90.5	88.6

Table 2: Accuracy comparison between the proposed system in [6] and commercial systems.

System	Manufacturer	Detection accuracy $[\%]$	False positive detections $[\%]$	Execution time $[ms]$
Citilog	Citilog Company of France	88.71	9.74	< 10
Times I	Beijing Aerospace Time Technology Development Co. LTD	85.23	10.56	< 10
Proposed system [6]		90.23	6.13	< 5

trajectory are needed. By further analysis of these parameters system can determine if previously mentioned events are true or false. Proposed system is robust and it can operate in difficult environment conditions (rain, mist, night, etc.).

Video cameras mounted above or nearby roads in several cities in China (Xi'an, Shanghai, Fuzhou) have been used for testing of the system proposed in [6]. Videos from cameras were processed in real time. Results shown in Tab. 1 describe the system accuracy in the interval between $88 - 96\%$. Accuracy results comparison between the proposed system in [6] and commercial systems is shown in Tab. 2. It can be concluded that the proposed system in [6] has better vehicle detection accuracy than commercial systems. All of tested systems can be run in real time.

2.2 Traffic flow parameters

Traffic flow is a simultaneous motion of multiple vehicles on a road in a certain order. To describe the motion of multiple vehicles by a mathematical model, indicators which describe the model need to be defined. Mentioned indicators in the field of traffic flow theory are called basic traffic flow parameters or basic traffic flow quantities.

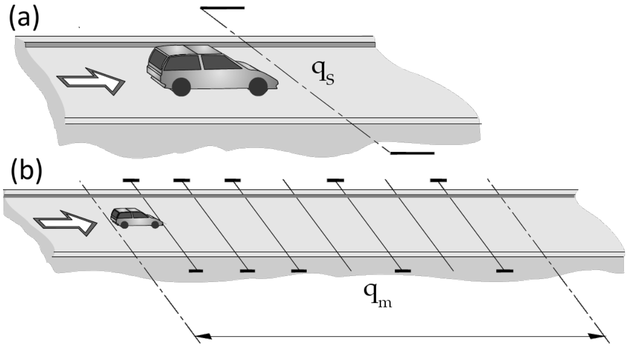

Figure 3: Measurement locations of two methods for computing vehicle traffic flow. q [3].

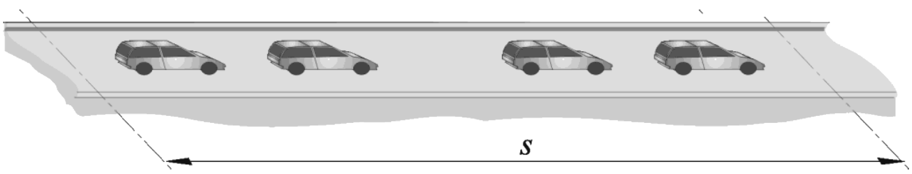

Figure 4: Traffic flow density g on a specific road segment [3].

Following indicators are used to describe the traffic flow:

- Vehicle traffic flow quantity q $[^{veh}/_{h}]$;
- Traffic flow density g $[^{veh}/_{km}]$;
- Traffic flow velocity v $[^{km}/_{h}]$;
- Travel time in traffic flow t $[h]$;
- Headway S $[m]$.

Vehicle traffic flow quantity represents the number of vehicles which passed the monitored segment of a road in a unit of time. Two methods for measuring vehicle traffic flow shown in Fig. 3 are: vehicle traffic flow measured on one road segment only (a); and mean value of vehicle traffic flows which are measured on multiple road segments (b). Traffic flow density represents the number of vehicles on a road lane segment in one or both directions depending of the road type as shown in Fig. 4. Traffic flow velocity can be separated into two quantities [2]:

- Mean spatial flow velocity which is the arithmetic mean of current vehicles velocities in a traffic flow on the observed segment of road defined by Eq. 6, where v_s is the mean spatial traffic flow velocity $[^{km}/_{h}]$, S is segment length $[km]$, t is travel time of i^{th} vehicle $[h]$ and N is number of observed vehicles $[veh]$;
- Mean time flow velocity which is the arithmetic mean of all vehicles velocities in a traffic flow at certain road intersection observed in certain time interval defined by Eq. 7, where v_t is mean time flow velocity $[^{km}/_{h}]$, N is number of observed vehicles $[veh]$, v_i is velocity of i^{th} vehicle $[^{km}/_{h}]$.

Table 3: Results of system proposed in [7].

Evaluation function		MRE	MMRE	RMSE
30 minutes	History profile approach	0.0674	0.6552	0.1351
	Kalman filter	0.0886	0.7841	0.1535
	System proposed in [7]	0.0456	0.3091	0.0690

$$v_s = \frac{S}{\sum\limits_{i=1}^{N} \frac{t_i}{N}} = \frac{NS}{\sum\limits_{i=1}^{N} t_i} \tag{6}$$

$$v_t = \frac{\sum\limits_{i=1}^{N} v_i}{N} \tag{7}$$

Vehicle headway distance S $[m]$ is defined as a spatial distance between two consecutive vehicles in a traffic flow. There are three methods for computing vehicle headway:

- Distance between individual vehicles in a traffic flow, S_{hi}, as shown in Fig. 5 which exist in a certain moment at a specific road segment;
- Mean value of current distances between all vehicles on a observed road segment S;
- Arithmetic mean of current distances on a observed segment S in a specific time period.

Wang and Ma [7] proposed a system which predicts a travel time for ITS based traffic management systems. Proposed system uses two sensors which can detect vehicles passing through detection points. After the vehicle passes through detection points, by knowing the distances between detection points the system computes traffic parameters (current motion velocity, traffic flow, mean travel time, etc.). Using a history data of model state $\hat{x}_t$, measured traffic parameters z_t and current model state $\hat{x}_t$, system predicts future model state $\hat{x}_{t+1}$ with particle filter. Experimental results with error values are given in Tab. 3, where MRE is the mean value of a relative error defined by Eq. 8, MMRE is the maximum mean value of a relative error defined by Eq. 9 and RMSE is the square root of mean value of square errors defined by Eq. 10. In Eq 10, $\tilde{R}_t$ is predicted travel time and R_t is the real travel time at time instance t.

$$MRE = \frac{1}{T} \sum_{t=1}^{T} \left| \frac{\tilde{R}_t - R_t}{R_t} \right| \tag{8}$$

$$MMRE = max \left| \frac{\tilde{R}_t - R_t}{R_t} \right| \tag{9}$$

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^{T} \left(\frac{\tilde{R}_t - R_t}{R_t} \right)^2} \tag{10}$$

System for measuring traffic parameters in urban areas has been proposed by Zou, Shi and Wang in [8]. Automatic computation of traffic parameters is based on image processing algorithms where images are obtained from a video camera as shown in Fig. 6. Each image from a video stream is first converted to the

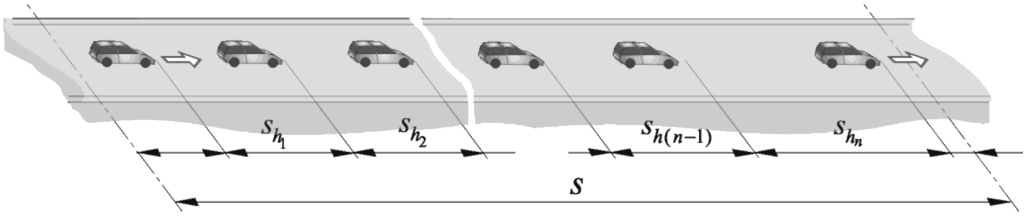

Figure 5: Vehicle headway distance S_h in a traffic flow [3].

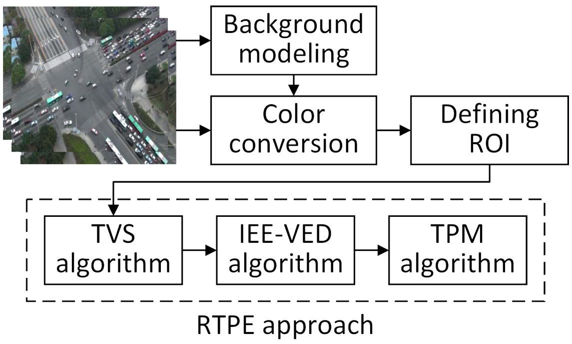

Figure 6: Flow chart of the system which measures traffic parameters with a video camera [8].

grayscale format. Then vehicle segmentation is performed based on texture (texture-based vehicle segmentation - TVS). Noise (pedestrians and other non-vehicle objects) are removed by checking the entropy in the image (image entropy-based vehicle exist detection - IEE-VED). Last step in the system is the computation of traffic parameters from the image (traffic parameters measurement - TPM). Traffic parameters which can be obtained as a system result are:

- Vehicle traffic flow;
- Tavel time;
- Mean spatial velocity of vehicle motion;
- Ratio between the number of vehicles detected by the camera which entered and leaved the observed area.

The accuracy of vehicle detection of the system described in [8] is 97.9%, where 375 of total 383 vehicles was successfully detected with the system.

3 Video cameras for road traffic monitoring

Basic camera types depending on the output image format which are used in ITS are grayscale, RGB and infra-red (IR) cameras. When grayscale cameras are used, each pixel in the image contains only information about pixel intensity. RGB cameras have an output image where each pixel contains information of red, green and blue color intensities. Benefit of a RGB camera is that RGB images can easily be converted to the grayscale format, however this conversion cannot be done in the opposite order. RGB images store more information of objects which are present inside them than grayscale images (color is the one of information). When night conditions are present in a scene, often grayscale and RGB cameras are not

Figure 7: GoPro HD - Hero2 RGB camera.

Table 4: Specification of RGB camera used with the proposed system.

Manufacturer	GoPro
Model	HD Hero2 Professional
Resolution $[px]$	1920×1080
FPS	30
FOV [°]	170

suitable for recording the scene. The main reason for this is insufficient sensitivity of CMOS or CCD (charged-coupled device) elements in the camera which are used for capturing the visible electromagnetic (EM) spectrum. In such cases IR cameras are used which can record the scene in the IR spectrum. They are useful for pedestrian detection in night conditions as during the night pedestrians are more noticeable in the IR than visible spectrum.

Road traffic videos used with the proposed system in this report are acquired by using the RGB camera shown in Fig. 7. The camera specifications are given in Tab. 4. Cameras were mounted above the road of the Zagreb city bypass in locations Ivanja Reka, Zaprešić and Lučko. They captured all incoming vehicles. Resolution of the cameras was high enough to allow detection and tracking of vehicles and automatic vehicle license plate recognition (LPR).

4 Vehicle detection and tracking

Basic problems which systems for automatic road traffic monitoring using video cameras should solve are vehicle detection and vehicle tracking. Vehicle detection represents a process which detects vehicle presence, and measures its location, size and other parameters of the vehicle in the image. Vehicle tracking is a process where the vehicle detected in the image is observed through time or series of consecutive images in the video footage. With this processes additional parameters can be computed such as vehicle motion velocity and direction.

4.1 Vehicle detection approaches and methods

Vehicle detection can be achieved by using one of the two main types of methods: background subtraction method and method based on machine learning. The background subtraction method is based on comparison between a background image and an image with moving objects (vehicles). All moving objects in the image can be detected by analysing variation between images. Main deficiency of this method is its

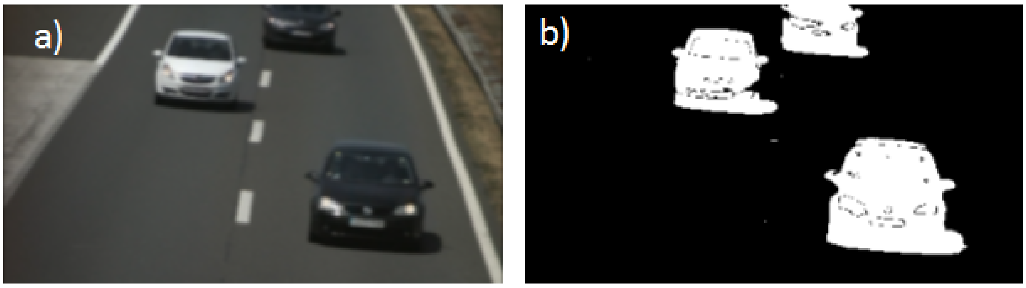

Figure 8: Image received from a video camera which records the road traffic (a) processed with background subtraction algorithm based on the mean value method (b).

incapability to detect static vehicles in the image (eg. parked vehicles). Method requires a static camera from which video is acquired (camera does not change its location and position under influence of wind, vehicle vibrations, etc.).

Machine learning methods are based on computing specific features in the image and performing classification to determine if computed features represent part of a vehicle or a background object. Before the classification process can be performed, an appropriate classifier needs to be learned using a learning dataset.

4.1.1 Background subtraction methods

Background subtraction methods are based on the assumption that on two consecutive images only moving objects can variate in their location, position and size in the image. The simplest and fastest object detection method is comparison of pixels in two consecutive images which have the same x, y location in the images. If variation between pixel intensities is larger than a specific threshold, pixel at this location will be classified as part of the foreground, or otherwise part of the background in the image. Mentioned procedure is defined by Eq. 11, where I_t is pixel value at x, y location in the image t, I_{t-1} is pixel value at location x, y in the previous image $t-1$ and th is minimum variance (threshold) which separates the moving and static objects. Variable P shows if pixel in the image I_t at specific location x, y is moving object (1) or static object (0).

$$P = \begin{cases} 1 & if\,|I_t - I_{t-1}| > th \\ 0 & otherwise \end{cases} \tag{11}$$

The median method is based on computing image background model which represents image with static objects (eg. road infrastructure) only. The background image model is then compared with the latest image in a set of consecutive image. Procedure is similar to the previously mentioned method described by Eq. 11 with exception that the latest image and the background image are compared. Median value of all pixels at specific location x, y in the set of consecutive images $I = \{I_n : n \geq 2\}$ is used for computation of the background image model. In the mentioned procedure arithmetic mean value or Gaussian mean value can also be used instead the median value [9]. Background and foreground objects separation with background subtraction method is shown in Fig. 8.

Method of Gauss mixture models is based on computing a probability that certain pixel intensity value at a specific location in an image is representing foreground or background object. If pixel in the image has probability to be a part of the background larger than specific threshold it is classified as background or otherwise as foreground [10].

4.1.2 Machine learning methods

Machine learning methods consist of algorithms which have the possibility to make predictions and decisions based on previously processed data with which a learning process has been performed. The continuation of the work describes use of image features and classifiers for object (vehicle) detection. In order to compute a specific type of features (information) in an image with which the system can perform vehicle detection by using computer vision, it is required to define what do the features represent or a procedure for their computation. Features can represent information such as gradient in the image or global and local structure in the image described in the following sections [11].

Local structures in the image can represent edges, lines, curves, contours, corners, cluster of points (blob) which separates from other clusters of points by some characteristic. In the field of computer vision, corners and similar group of points are often called interest points (key points). Global structure in the image is computed for the whole image or a fixed region of the image. They include various statistical information about the image such as histograms of pixel intensities, pixels mean values, variance, Fourier transforms and various moments [11].

In many cases it is difficult to define which features in the image should be used in a classification process. While computing features, many redundant features can occur. In such case it is difficult to choose which features are important and which are not important. Features need to be placed in a vector which is often called a feature descriptor or feature vector. If a feature vector dimensionality is not a fixed number and it is depending on various instances of feature classes, various methods for reducing the vector dimensionality can be performed such as principle component analysis [12]. The Feature vector is then processed with a classifier to determine a class to which the feature vector belongs. There are various classifiers which can be used in a classification process. Support vector machine (SVM) classifier tries to separate features by computing a multidimensional hyperplane which separate each component of the feature vector. Process of computing the hyperplane represents the learning process. The learning process is based on a learning dataset which consists of many feature vectors and class indices to which each vector belongs to. Mudoi and Kashyap proposed the system described in [13] which uses histogram of gradient (HOG) features and a SVM classifier for vehicle detection in a scene. Accuracy of the vehicle detection system is approximately 97%. Lan, Zhang, Lu and Guo have proposed system [14] which uses local binary patterns (LBP) features and GCI classifier. System accuracy is about 95% with LBP features and 98% with N-LBP features.

4.2 Vehicle tracking approaches and methods

To identify a vehicle on multiple consecutive images it is required to track its previous movement and predict possible future model state (eg. location and size). If detected object on the newest image has its current model state same as the predicted model state (the distance between model states values is below a certain threshold), it is considered that both objects represent the same vehicle in the scene. If an object from the previous image has no similar object in the newest image, it is considered that the object has left the scene. If an object from the newest image has no similar object in the previous image, it is considered that the object has entered the scene. There are various models which can be applied for predicting the movement of an object. The methods often used for modeling the movement are based on Markov chains. Markov filter is usually used for predicting the movement of an object because computation of its prediction depends only on a model state computed in the previous iteration (previous image) and current measurement values (eg. with and vehicle detector algorithm). The most often used filter based on Markov chains is the Kalman filter described in [5].

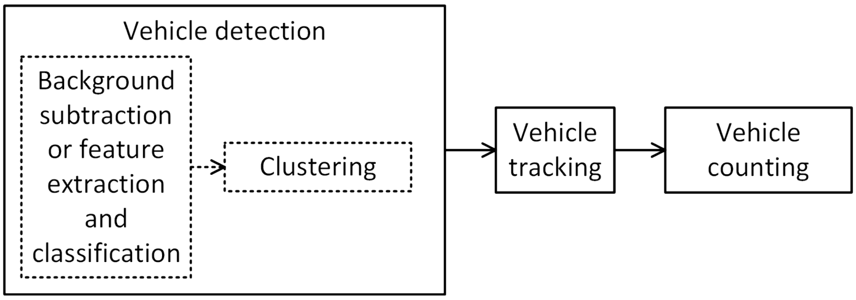

Figure 9: Work flow chart of proposed systems for vehicle detection and tracking.

4.3 Proposed vehicle detection and tracking systems

In the following sections of this report, a comparison between two proposed systems which are based on different algorithms for vehicle detection is described. Work flow chart of the proposed systems is given in Fig. 9. The systems use an image from a road traffic video as its input and process it with the implemented vehicle detection algorithm. In the first proposed system, the algorithm for vehicle detection is based on the background subtraction method where background is computed by the mean pixel value method on multiple consecutive images. The used method is described in [4]. The second system for vehicle detection is based on feature extraction from an image and uses a classifier for vehicle detection. Clustering is performed on detected objects, where objects are grouped in individual vehicles. Detected vehicles are processed with a vehicle tracking algorithm which is based on the EKF filter. The EKF performs prediction of the vehicle location in the image which is then compared with a vehicle location in the latest image received from the vehicle detection algorithm. This method allows vehicle tracking through time and computation of additional vehicle motion parameters. The last system component is the algorithm for counting vehicles which have passed in a scene.

4.3.1 Vehicle detection with background subtraction method

In the first proposed system, the background subtraction method is used for vehicle detection. Method consists of background model computation and background subtraction of the latest image. Background model computation is defined by Eq. 12, where BK_k is background model for the latest image I_k, BG_{k-1} is background model for the previous image $I_k - 1$, n is the number of images with which the background image model is computed. In the proposed system n is 105. Background subtraction from an image is defined by Eq. 13, where d is the class of an pixel in the image (1 if it is a foreground part, 0 if it is a background part), th is a threshold constant which represents the minimum difference between pixel intensities in order to classify pixel as a part of moving object.

$$BG_k = BG_{k-1} + \left\lfloor \frac{\sum_{i=1}^{n} sign\left(I_{k-i} - BG_{k-1}\right)}{n} \right\rfloor \tag{12}$$

$$d = \begin{cases} 1\ , & if\,|I_k - BG_{k-1}| \geq th \\ 0\ , & otherwise \end{cases} \tag{13}$$

4.3.2 Vehicle detection with a feature classification method

The second proposed system is based on a method which extract LBP features from an image as described in [15] and performs classification with the Gentle Adaboost (GAB) algorithm explained in [16] and [17]. Before vehicle detection can be performed in the image, a classifier needs to be learned to classify vehicle LBP features in the image. The learning process is performed with a learning dataset. The learning dataset consists of two subsets of images:

- Images which contains vehicles (positive samples);
- Images which do not contain vehicles (negative samples).

From the individual subsets, LBP features are extracted. The classifier is learned to separate specific features into two classes (with and without vehicles). In this report, three versions of datasets described in Tab. 5 are used to learn the classifier. The dataset created by the Faculty of Transport and Traffic Sciences is made from video footage of a road near the city of Zagreb obtained by the camera shown in Fig. 7, where its specifications are given in Tab. 4.

4.3.3 Vehicle tracking

Trajectory computation of a moving object in a set of consecutive images is performed by the EKF. A EKF state model is based on a vector x defined in Eq. 4, where x_x and x_y represent a location of the moving object on the x and y axis, x_v is the velocity and x_a is the acceleration of the moving object over an image plane, x_ϕ and x_ω are the object direction and change of direction by a time interval t. Prediction of a future model state is performed with Eqs. 5 and 14, where $x_{k||k-1}$ is a predicted state model vector in the step k based on a state model vector $x_{k-1|k-1}$ from the previous step $k-1$ and t is the time interval between the steps $k-1$ and k.

$$x_{k|k-1} = f(x_{k-1|k-1}, t) \tag{14}$$

After the prediction step has been finished, an update step is performed by using Eqs. 15-23, where y_k is the innovation vector, $P_{k|k-1}$ is the covariance matrix of the predicted state model, F_{k-1} is the state transition matrix, $P_{k-1|k-1}$ is the covariance matrix of the predicted state model from previous iteration $k-1$, Q_k is the covariance matrix of the process noise, S_k is the covariance innovation matrix, H_k is the measurement matrix, R_k is the covariance matrix of the measurement noise, W_k is the Kalman gain matrix and I is the identity matrix.

Table 5: Number of samples in datasets.

Dataset created by	Number of positive samples	Number of negative samples	Total number of samples
Faculty of Transport and Traffic Sciences	4298	1930	6228
Toyota Motor Europe Motorway dataset [18]	25155	26321	51476
Combination of the both datasets	29453	28251	57704

$$
\begin{aligned}
h(x) &= \begin{bmatrix} x_x \\ x_y \end{bmatrix} && (15)\\
y_k &= z_k - h(x_{k|k-1}) && (16)\\
F_{k-1} &= \left.\frac{\partial f}{\partial x}\right|_{x_{k-1|k-1}} && (17)\\
H_k &= \left.\frac{\partial h}{\partial x}\right|_{x_{k|k-1}} && (18)\\
P_{k|k-1} &= F_{k-1}P_{k-1|k-1}F_{k-1}^T + Q_{k-1} && (19)\\
S_k &= H_k P_{k|k-1} H_k^T + R_k && (20)\\
W_k &= P_{k|k-1} H_k^T S_k^{-1} && (21)\\
x_{k|k} &= x_{k|k-1} + W_k y_k && (22)\\
P_{k|k} &= (I - W_k H_k) P_{k|k-1} && (23)
\end{aligned}
$$

An important feature of the EKF is a requirement to choose initial values of the state model vector $x_{k-1|k-1}$ and matrix $P_{k-1|k-1}$ before EKF execution. In the proposed system, values of a vector $x_{k-1|k-1}$ are set by a histogram which contains previous values of the state vector x, while values of matrix $P_{k-1|k-1}$ are set to constant values in the step $k = 0$. The histogram is divided into $i \times x$ segments, where each segment covers a rectangular area of the image. Update of the histogram is performed in each EKF iteration where updated (corrected) values of the state model vector $x_{k|k}$ are added to the histogram. Average mean values can be used for choosing the initial values of the state model vector $x_{k-1|k-1}$ at step $k - 1 = 0$. Average mean values are computed using all the values added to the histogram.

4.3.4 Vehicle counting

The last step in the proposed system is vehicle counting. Vehicle counting is based on virtual markers M_1 and M_2. Virtual markers are located in the bottom part of the image in order to detect that an object is leaving the scene. Detection method is based on checking overlaps between each object and virtual marker. When the background subtraction method is used, a vehicle counter will be increased by 1 if following requirements are fulfilled:

- The total number of images in which an object was present in the scene is greater than *TFC*;
- When using the EKF for vehicle tracking, the direction of a vehicle movement during time vehicle is overlapping with a virtual marker is inside a specific interval $[\phi_{min}, \phi_{max}]$, where ϕ_{min} and ϕ_{max} are set by Eqs. 24-25.

$$
\begin{aligned}
\phi_{min} &= \left(\frac{1}{8} + 1\right)\pi[rad] && (24)\\
\phi_{max} &= \left(\frac{7}{8} + 1\right)\pi[rad] && (25)
\end{aligned}
$$

When the feature classification method is used, a vehicle counter will be increased by 1 if following requirements are fulfilled:

- The total number of images in which an object was present in the scene is greater than TFC;
- The total travelled distance of an object in the image plane is greater than $\frac{1}{5}max(w, h)$, where w is the image width $[px]$ and h is the image height $[px]$.

4.3.5 Experimental results

In this section experimental results are given which show comparison between previously mentioned methods for vehicle detection and tracking from aspects of accuracy and execution time. Used video footage for testings is recorded by a camera mounted above the bypass of the city of Zagreb at the location Ivanja Reka. Results of testing the background subtraction method are given in Tab. 6, where *W* and *H* are image resample ratios (horizontal and vertical) used in vehicle detection, *FP* is the number of false positive detections, *N* is the number of false negative detections and *GT* is the groundtruth data, *TFC* is the minimum number of images in which an object needs to be detected to get classified as vehicle and *EKF* shows if EKF filter has been used for vehicle tracking. Accuracy *A* is defined by Eq. 26. Execution time is given in minimum, average and maximum values *[FPS]*. Experimental results of the feature classification method are given in Tab. 7, where *S* is the number of stages in a cascade classifier, *MHR* (Minimum Hit Rate) is the minimum ratio between the number of detected objects and the total number of objects which have to pass into a next learning stage of a cascade classifier and *MCC* (Minimum Cluster Count) is the minimum number of individual clusters detected by the feature classifier which an object (vehicle) needs to contain. Execution time *[FPS]* is given when no vehicles, one vehicle and multiple vehicles are present in the scene.

$$\left(1 - \frac{FP + FN}{GT}\right) \times 100\% \tag{26}$$

The system based on the background subtraction method was tested on all combinations of method parameters where:

- The image resample ratio used in the detection process W, $H = \left\{\frac{1}{8}, \frac{1}{4}, \frac{1}{2}\right\}$;
- Threshold value $th = \{5, 8, 10, 15, 20, 25, 30, 35, 40, 45, 50\}$;
- $TFC = \{10, 15, 20, 30\}$;
- $EKF = \{used, notused\}$.

Table 6: Experimental results for the background subtraction method performed on a video footage of the Zagreb County bypass at the location Ivanja Reka.

EKF	Image		Method parameters		Execution time [FPS]			Counted vehicles			
	W [%]	H [%]	th	TFC	Min	Avg	Max	FP	FN	GT	A [%]
No	12.5	12.5	10	15	60	60	60	4	2	133	95
No	12.5	12.5	10	20	61	60	59	2	4	133	95
No	12.5	12.5	10	30	60	60	58	2	6	133	94
Yes	12.5	12.5	10	10	60	60	60	4	4	133	94
Yes	12.5	12.5	10	15	60	60	60	4	4	133	94
Yes	12.5	12.5	10	20	61	60	60	3	5	133	94
Yes	12.5	12.5	10	30	60	60	59	3	6	133	93
Yes	12.5	12.5	8	10	61	60	59	4	6	133	92
Yes	12.5	12.5	8	15	60	60	60	4	6	133	92
Yes	12.5	12.5	8	20	61	60	60	4	7	133	92
No	12.5	12.5	8	15	60	60	60	5	6	133	92
Yes	12.5	12.5	8	30	61	60	59	4	8	133	91
No	12.5	12.5	8	20	60	60	60	5	8	133	90

Table 7: Experimental results for the feature classification method performed on a video footage of the Zagreb County bypass at the location Ivanja Reka.

EKF	Image		Method parameters				Execution time			Counted vehicles			
	W [%]	H [%]	S	MHR	MCC	TFC	0	1	> 1	FP	FN	GT	A [%]
No	12.5	12.5	20	0.995	2	8	18	18	21	0	6	133	95
No	12.5	12.5	20	0.995	5	3	18	22	21	0	6	133	95
No	12.5	12.5	20	0.999875	5	8	15	13	9	0	6	133	95
No	12.5	12.5	20	0.999875	5	3	14	13	10	0	7	133	95
No	12.5	12.5	20	0.999875	5	15	14	13	11	0	7	133	95
No	12.5	12.5	20	0.995	2	3	18	17	21	3	5	133	94
No	12.5	12.5	20	0.995	2	15	17	19	21	0	8	133	94
No	12.5	12.5	20	0.995	5	8	16	19	21	0	8	133	94
No	25	25	20	0.995	5	8	8	7	7	1	7	133	94
No	25	25	20	0.995	5	3	8	7	7	3	6	133	93
Yes	12.5	12.5	20	0.995	2	3	18	17	21	0	9	133	93
Yes	12.5	12.5	20	0.995	2	8	18	18	21	0	9	133	93
Yes	12.5	12.5	20	0.995	2	15	17	19	21	0	9	133	93
Yes	12.5	12.5	20	0.999875	5	3	14	13	10	0	9	133	93
Yes	12.5	12.5	20	0.999875	5	8	15	13	10	0	9	133	93

Results obtained from the system tests of both methods are shown in Fig. 10. In Fig. 10, ratios between population of vehicle detection and tracking accuracy intervals and the total number of testings are shown. Values in red bars represent results obtained when the EKF for vehicle tracking is used and blue bars represent results when the EKF was not used. Results show that the maximum vehicle detection and tracking accuracy is 95%. Detection accuracy in the interval between $90 - 100\%$ is obtained in 6% of the total number of performed system testings. In testing of the system based on the feature classification method, all combinations of parameters values have been made:

- The Image resample ratio used in the detection process $W, H = \left\{\frac{1}{8}, \frac{1}{4}, \frac{1}{2}\right\}$;
- The number of stages $S = \{10, 20\}$;
- $MHR = \{0.99, 0.9925, 0.995, 0.9975, 0.999875\}$;
- $MCC = \{1, 2, 5, 15, 25, 30, 50\}$;
- $TFC = \{3, 8, 15\}$;
- $EKF = \{used, notused\}$.

Results of system testings based on the feature classification are shown in the left chart in Fig. 10. Detection accuracy in the interval between $90 - 100\%$ is obtained in 17% of the total number of performed system testings, where the maximum individual accuracy of all testings is 95%. Ratio between variations of accuracy results with and without the EKF is given in Fig. 11 for the background subtraction method and in Fig. 12 for the feature classification method. In charts given in Figs. 11-12, positive percentages represent increase and negative percentages represent decrease in accuracy when the EKF is used and not used in testings.

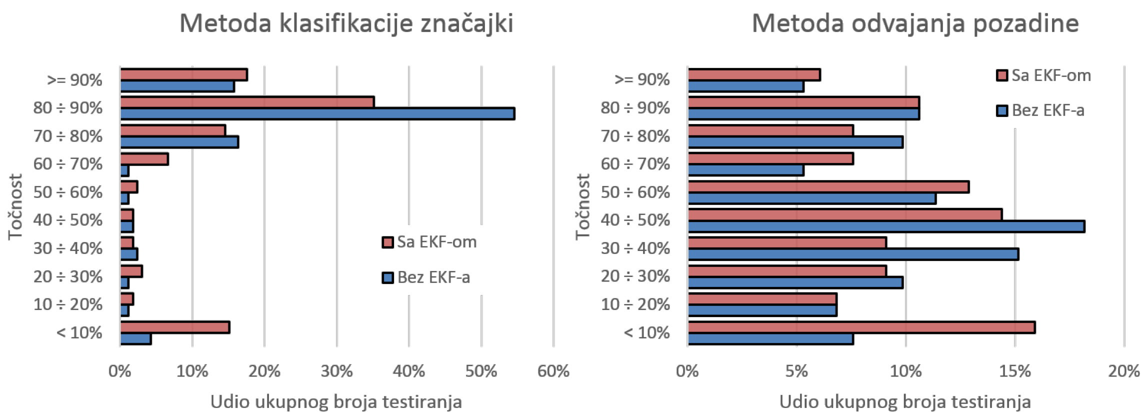

Figure 10: Ratio between obtained vehicle detection and tracking accuracy intervals and the total number of testings.

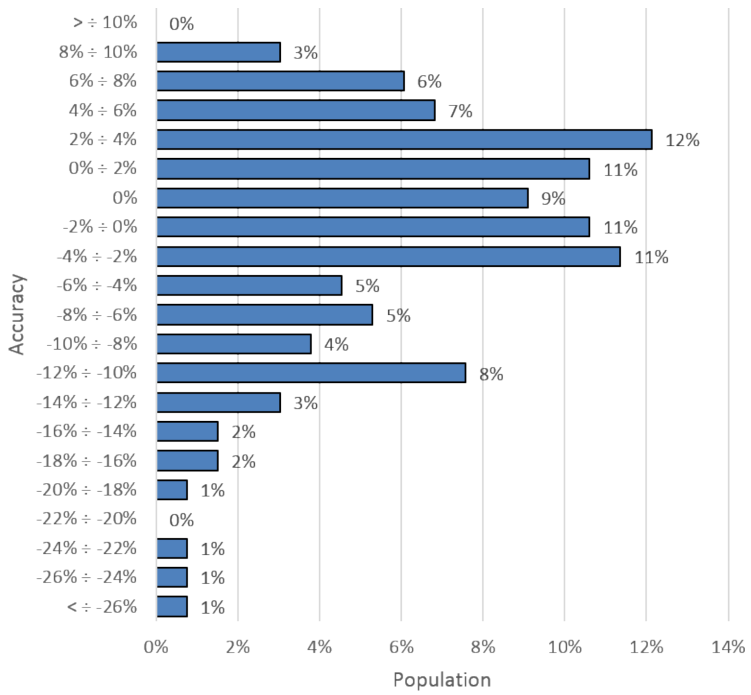

Figure 11: Ratio between variations of accuracy results for the background subtraction method performed with and without the EKF.

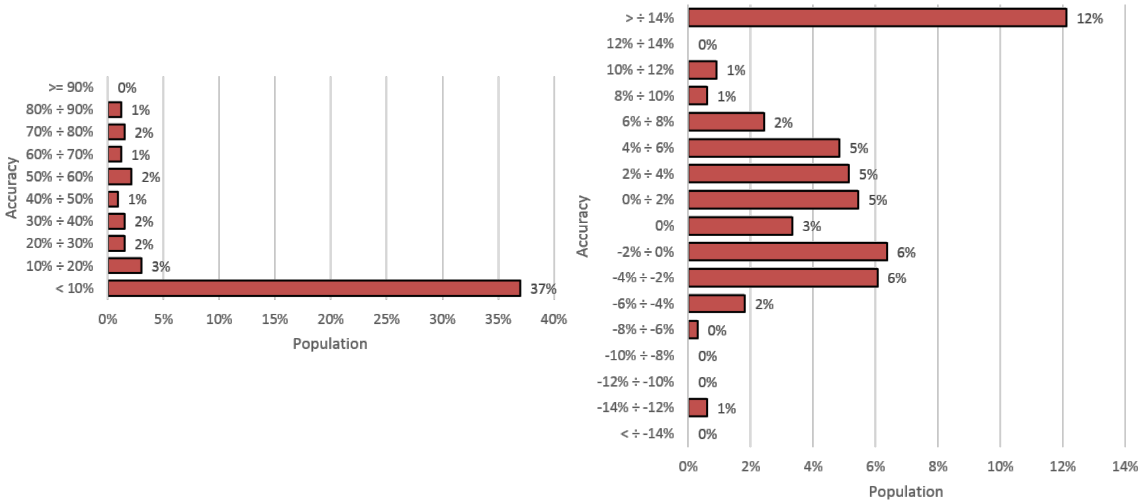

Figure 12: Ratio between variations of accuracy results for the feature classification method performed with and without the EKF.

5 Image quality requirements

One of factors which affects accuracy of vehicle detection and tracking using a video camera is quality of the image which the system receives. Today, there is no explicitly defined standard which describes image quality requirements in the of scope ITS. Analysis for obtaining mentioned requirements are described in this section. Before image quality requirements can be specified, the term *"image quality"* needs to be more precisely explained. Requirements for the image quality vary between various application areas as shown in Fig. 13. For example, NASA images require as less noise and distortion as possible (genuineness), medical images require that some specific regions in the images are noticeable (usefulness), etc. Driver assistance and active safety systems have image quality requirements defined as a combination of naturalness and usefulness [19].

There is a large number of video cameras which can be applied in the area of automotive industry.

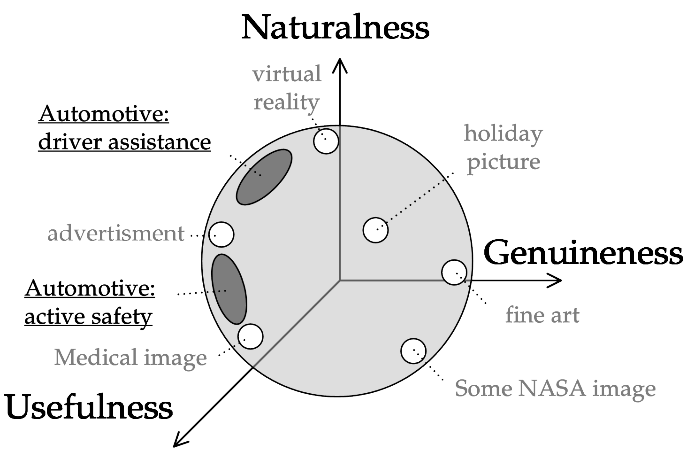

Figure 13: Distribution of image quality indicators between various application areas.

Available cameras differ in the output image quality and for this reason it is challenging to compare them. Hertel and Chang [19] have proposed the basic concept of objective and subjective image quality. In the objective image quality, three parameters are used in computation:

- Characteristic curve of tone reproduction (CCTR);
- Noise power spectrum (NPS);
- Modulation transfer function (MTF).

CCTR represents relationship between exposure and resulting image value (density or any other image units). Every image obtained from a video camera contains noise. NPS measures the local amplitude variance due of noisy sine waves as a function of spatial frequency. MTF is the frequency-dependence of the change in amplitude of sine waves after transmission by one of the imaging system components such as lens, pixel array or image signal processing algorithms. Ideal method for image quality assessment would be a method which for the result returns only one value which describes the image quality. Due to complexity of the image quality evaluation process, such method is still not defined. Possibility of an image processing system to detect small details with low contrast in an image can be defined by photographic signal to noise ratio (SNR) as shown in Eq. 27, where q is SNR, ΔD is the image density difference (singal *S*) and σ_D is the standard deviation of the density fluctuation (noise *N*).

$$q = \frac{\Delta D}{\sigma_D} \approx \sqrt{\frac{S}{N}} \tag{27}$$

Wang, Bovik, Sheikh and Simoncell [20] have proposed an objective method for determining the image quality which is based on quantitative determination of visible errors between the reference image and the image with distortion. Deficiency of the mentioned method is in requirements for the reference image which will be compared with the distorted image. Because in most cases it is hard to obtain the reference image, this method [20] and all previously mentioned methods [19] have a complex implementation. Pflugfelder, Bischof, Domínguez, Nölle and Schwabach [21] have defined 6 parameters or requirements of a video camera important in image processing desribed below:

- Progressive scan by which an effect of blending consecutive images (interlace effect) cause distortion of features in an image. Consequence is the decrease of accuracy in vehicle detection, tracking and license plate recognition. The mentioned problem is especially conspicuous in scenes where vehicles move with high velocities (highways);
- Motion blur of an object depends on the used camera geometry (distance from an object, image resolution, horizontal coverage area) and integration time;
- Camera sensitivity where *SNR* can be used as quantity for the determination of sensitivity;
- Blooming and smearing effects which appear in video cameras with CCD elements;
- Dynamic range of intensity values which a video camera covers;
- Resolution.

In mentioned parameters, algorithms in video cameras used for image enhancement are also included. Although the basic problem which algorithms try to solve is image enhancement, final result does not need to have a positive impact on the vehicle detection and tracking process. Because of the mentioned reason, it is required to determine if image enhancement algorithms are really necessary for vehicle detection and tracking.

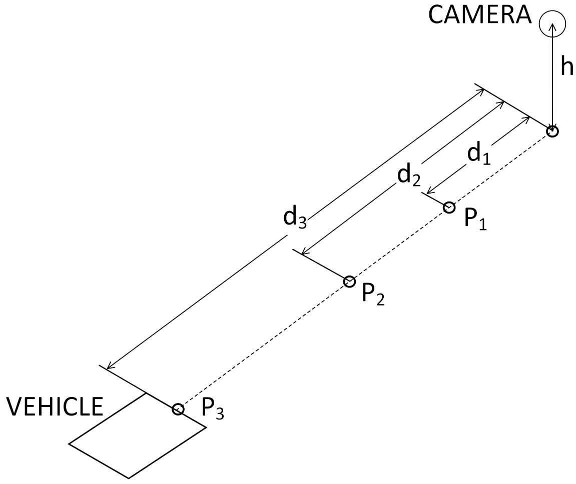

Figure 14: Required parameters used in measuring the image quality with the proposed method.

External factors which can have influence on vehicle detection and tracking consist of vibrations which happen due to passing of vehicles nearby the camera. Consequence can be decrease of a camera resolution. In some locations, the camera lens can be covered with various particles such as dust, snow, water which can have a negative impact on the vehicle detection and tracking system.

Method for measuring image quality requirements proposed in this report consists from two steps. First step requires creation of videos which will have various resolutions like for example: 960×540 $[px]$, 480×270 $[px]$ and 240×135 $[px]$. Videos need to have three detection points P_1, P_2 and P_3. Distances d_1, d_2 and d_3 between vehicle and camera as shown in Fig. 14 need to be known when a vehicle is passing through this detection points. Second step consists of processing videos with vehicle detection and tracking algorithms. When each vehicle is located at one of the detection points, information if proposed system has detected vehicle in that moment or not should be stored together with detection point index (P_1, P_2 or P_3) and a time stamp. With the proposed method relationship between video resolution, distance of a vehicle from the camera and detection accuracy can be computed. The ideal video resolution for maximum vehicle detection and tracking accuracy can be obtained by analysing the stored vehicle detection and tracking results.

6 Measurement of road traffic parameters on multiple nodes

Proposed architecture for measuring traffic parameters using multiple video cameras is given in Fig. 15. Video cameras forward a video stream to a local server where image processing is performed. Video stream is first processed by vehicle detection and tracking algorithms. By further processing of data computed by vehicle detection and tracking algorithms, it is possible to compute traffic parameters such as traffic flow velocity *v*, traffic flow *q*, traffic density *g*, vehicle headway *S* or travel time *t* for a specific segment of a road traffic network. From the video stream it is possible to read vehicle license plate (license plate recognition - LPR) which can be used for tracking an individual vehicle in the road traffic network. Mentioned parameters are forwarded to the central server. Based on mentioned parameters, the central server computes traffic flow parameters such as travel time t_p, mean vehicle velocity v_{sr} or traffic network OD matrix.

When measuring traffic parameters on multiple nodes of a traffic network, it is required to consider the number, location and position of every mounted camera by which data will be acquired. To make optimal distribution of camera locations various methods can be applied such as the point centrality measure, betweenness of a point, closeness of a point, etc. The point centrality measure is often used in traffic network

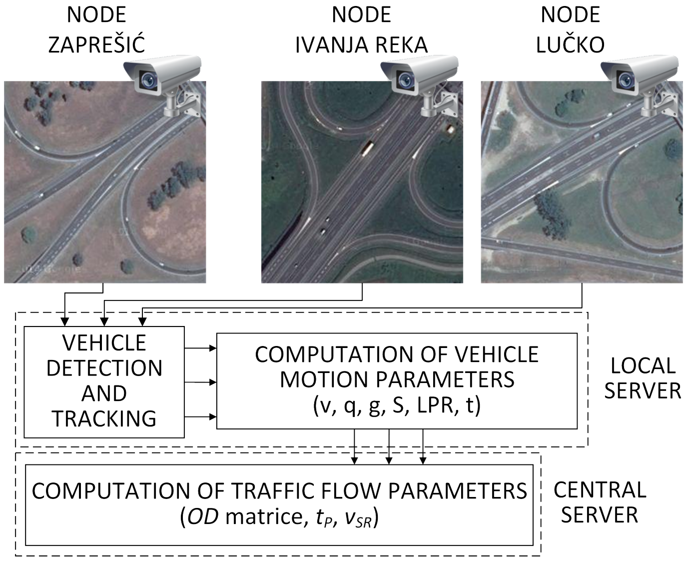

Figure 15: The system architecture for measuring traffic parameters on multiple nodes.

graph analysis to determine influence of individual points or regions in a traffic network depending on the road traffic network structure. One of methods to compute the point centrality measure is based on the degree of node (point). The degree of node represents the number of adjacent nodes which are connected to the observed node. It is defined by Eqs. 28-29, where C_i^D is the degree of node i. Increase of the node degree value cause the node influence to be increased also. Beside the mentioned method, betweenness of a point method or closeness of a point method can be used to compute the point centrality measure described in [22].

$$f(i,j) = \begin{cases} 1 & \text{if nodes i, j are adjacent (connected)} \\ 0 & \text{otherwise} \end{cases} \tag{28}$$

$$C_i^D = \sum_{j=1}^{n} f(i,j) \; i,j \in \{1,2,...,n\} \tag{29}$$

Paul, Malhotra, Dale and Qiang proposed the method described in [23] which allows optimal determination of point locations for data mining (sensors). The method is based on the modified point centrality measure method defined by Eq. 30, where $C_N(v)$ is the modified point centrality measure for node v, $N(v)$ is the set of nodes adjacent to node v, $D(v)$ is the degree of node v, D_{max} is the degree of the graph (traffic network), $P_{v,u}$ is the connection priority between nodes v and u, P_{max} is the maximum value of the connection priority between nodes in the traffic network, $D(u)$ is the degree of node u, $d_{v,u}$ is the distance between nodes v and u. The number of nodes for data mining is assumed to be 10% of the total number of nodes in a traffic network. Evaluation of the method proposed in [23] is based on a system error given in Fig. 16. The system error E is defined by Eq. 31, where B is a set of results of the point centrality measure defined by Eq. 30 and A is a set of values which represent frequencies of use for individual nodes in the traffic network based on vehicle motion trajectories in the traffic network.

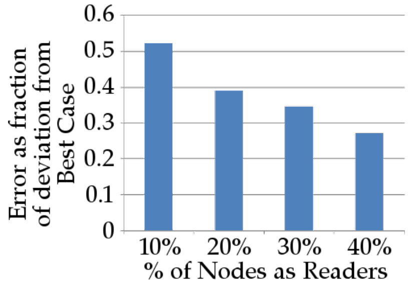

Figure 16: The error of modified point centrality measure method [23].

$$C_N(v) = \frac{D(v)}{D_{max}} + \sum_{u\in N(v)} \frac{P_{v,u}}{P_{max}} + \sum_{u\in N(v)} \frac{D(u)}{d_{v,u}} \tag{30}$$

$$E = \frac{|B \setminus A|}{|A|} \tag{31}$$

Point locations in the traffic network at which it is possible to install video cameras can be determined by using the previously mentioned method [23]. After installing cameras and additional equipments, it is required to process videos with algorithms based on computer vision. Vehicle motion velocity and road occupation can be computed by using vehicle detection and tracking algorithms. Automatic LPR software allows tracking of an individual vehicle in a traffic network. This feature has application in measurements of mean vehicle motion velocitiy, travelled distance and other information. In a larger scope, by knowing the number of vehicles which have appeared in a each video, it is possible to construct OD matrices which can be used for further studies of a road traffic network.

In continuation of this report, image processing of local traffic videos obtained from [24] is proposed in order to obtain OD matrix of a traffic network. Videos partially cover vehicle entrances and exits in the Zagreb city bypass as shown in Fig. 17 and specified in Tab. 9. Start time of recording for locations 1 and 2 begins in 11h. Start time of recording for other locations can be computed with Eq. 32, where t_{p_i} is the start time of recording at the location *i [min]*, S_a is the distance *[km]* which vehicle travels from entrance to exit locations on the highway, V_a is the maximum allowed vehicle velocity at the highway $[^{km}/_{h}]$, S_l is the distance *[km]* which vehicle travels from entrance to exit locations at two lanes road, V_l is the maximum allowed velocity at a local road $[^{km}/_{h}]$, t_n is the spent time at the toll station *[min]*, and n_s is the number of intersections with light signalization between entrance and exit points. Time is computed in minutes and it is shown as: $11h+$ the number of minutes. Computed time values are given in Tab. 8. By knowing the start time of recording it is possible to observe the traffic flow q which can be used for computation of OD matrices.

$$t_{p_i} = 11h + \left(\frac{S_a}{V_a} + \frac{S_l}{V_l}\right) 60 + t_n + n_s \; i = \{3, 4, ..., 10\} \tag{32}$$

Table 8: Start time values of recording at Zagreb County bypass [24].

i	S_a	V_a	S_l	V_l	t_n	n_s	t_{p_i}
3	2.86	130	2.93	50	0.5		11h + 5 min
4	12.13	130		50	0.5		11h + 6 min
5	16.91	130		50	0.5		11h + 8 min
6	23.55	130	2.38	50	0.5	2	11h + 16 min
7	38.05	130		50	0.5		11h + 18 min
8	40.09	130	4.26	50	0.5	4	11h + 28 min
9	42.96	130	3.36	50	0.5	2	11h + 26 min
10	48.09	130		50	0.5		11h + 22 min

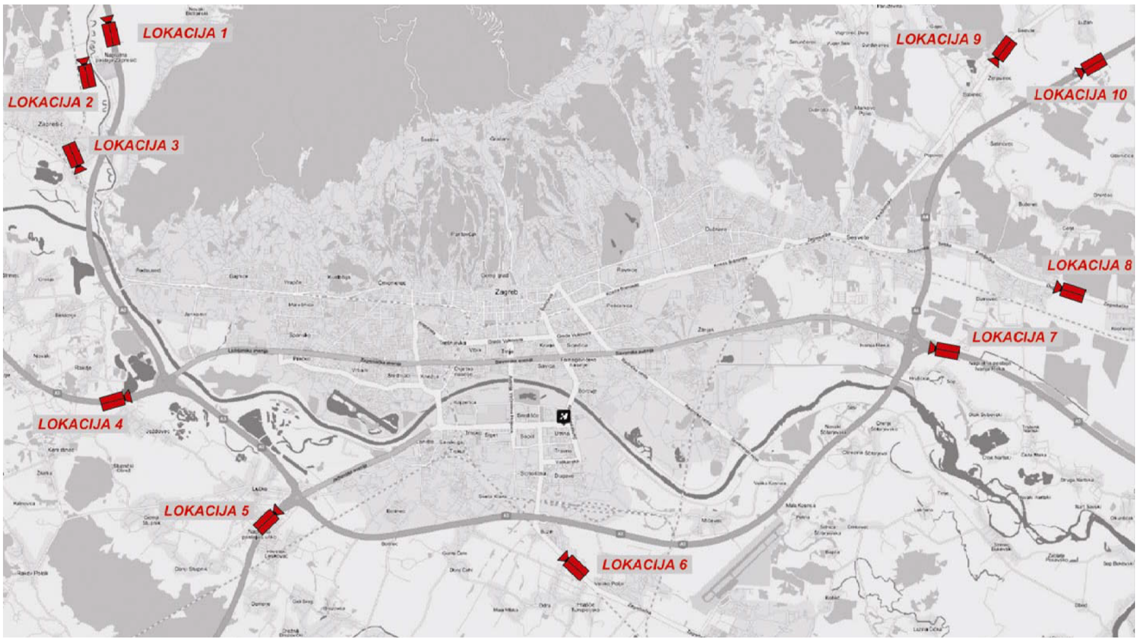

Figure 17: Locations of cameras on the Zagreb County bypass used in [24].

Table 9: Locations of cameras on the Zagreb County bypass used in master's thesis [24].

Id	Location	Position	Traffic flow
1	A2 - toll station Zaprešić, local road overpass in Jablanovac	Above highway	Entrance from highway, direction Krapine
2	D1 - north from entrance to West Gate shopping center, junction point with road toward Jablanovac	Near two lanes road	Entrance from local road, direction Krapine
3	East entrance to Zaprešić near gas station	Near four lanes road	Exit to Zaprešić
4	A3 - overpass at street "Zagrebačka ulica" in Rakitje, west from junction point Jankomir	Above highway	Exit to Ljubljana
5	A1 - overpass at street "Hojnikova ulica" in Lučko, south from junction point Lučko	Above highway	Exit to Karlovac
6	D30 - street "Zagrebačka ulica", overpass at the USA embassy, south from junction point Buzin	Above four lanes road	Exit to Velika Gorica and Sisak
7	A3 - overpass at street "Bilogorska ulica" between Dumovac and Hruščice	Above highway	Exit to Slavonski Brod
8	D41 - street "Dugoselska cesta" in Sesvete - near last bus station	Near two lanes road	Exit to Dugo Selo
9	D29 - street "Soblinečka ulica" in Žerjavinec	Near two lanes road	Exit to Varaždin on "stara cesta"
10	A4 - overpass of local road in Lužan highway	Above highway	Exit to Varaždin via highway

7 Conclusion

ITS management systems are an important element for road traffic optimization in urban and rural areas. The management system requires high quality and accurate real time traffic data for optimal traffic monitoring and management. Video cameras as sensors allow simultaneous computation of more traffic parameters from only one sensor, the video camera with acceptable accuracy.

In this report, experimental results of two systems for vehicle detection and tracking on a road traffic network using video cameras are compared. Testings are performed with various system parameter values where maximum vehicle counting accuracy is 95%. System based on the background subtraction method has an advantage by which it can detect any moving object in the video footage. As a consequence, determination of system parameters is a simple process. Main disadvantage of this method is sensitivity to camera vibrations and sudden lighting changes in an image which can cause increase in false positive and false negative detection. When a feature classification method is used, the system is robust to camera vibrations and sudden lighting changes in the image. Disadvantage of this method is requirement for a classifier learning process and a large dataset in order to perform accurate object detection. This requirement consumes a lot of time spent on preparation of a dataset and the classifier learning process. Mentioned method increase the number of system parameters, where determination of initial values of paramters is a more complex process which requires a lot of testings and system analysis.

Future work consists of measuring image quality requirements, where minimum video footage resolutions which give most accurate vehicle detection and tracking results would be determined. Besides mentioned, future work consists of performing analysis of vehicle trajectory accuracy. With obtained results described in this report, it can be concluded that use of the EKF can increase vehicle counting accuracy. Further analysis could show the amount of the EKF influence on vehicle trajectory prediction. Measurement of vehicle tracking accuracy can show if proposed system can efficiently extract complex information from the video footage which depends on vehicle trajectory. Complex information can consist of a driver aggressiveness factor or number of traffic violations (un-allowed road lanes changing, U-turns, etc.).

8 Acknowledgment

This project is co-financed by the European Union from the European Regional Development Fund.

The contents of this report is the sole responsibility of the University of Zagreb, Faculty of Electrical Engineering and Computing, Faculty of Transport and Traffic Sciences and do not necessary reflect the views of the European Union.

9 References

[1] V. Sergio K. Kovačić, E. Ivanjko. Real time vehicle country of origin classification based on computer vision. In *International Symposium on Electronics in Transport*, Ljubljana, Slovenia, March 2014.

[2] D. Cvitanić. *Teorija prometnog toka*. Faculty of Civil Enigneering, Architecture & Geodesy, University of Split, Split, Croatia, 2014.

[3] G. Kos I. Dadić. *Teorija i organizacija prometnih tokova*. Faculty of Transport and Traffic Sciences, University of Zagreb, Zagreb, Croatia, 2007.

[4] H. Gold K. Kovačić, E. Ivanjko. Real time vehicle trajectory estimation on multiple lanes. In *Third Croatian Computer Vision Workshop*, pages 21–26, Zagreb, Croatia, September 2014.

[5] D. Kasper G. Breuel D. Petrich, T. Dang and C. Stiller. Map-based long term motion prediction for vehicles in traffic environments. In *International IEEE Conference on Intelligent Transportation Systems*, pages 2166–2172, The Hague, Netherlands, October 2013.

[6] X. Ma Y. Yang X. Q. Liu H. S. Song, S. N. Lu and P. Zhang. Vehicle behavior analysis using target motion trajectories. *IEEE Transactions on Vehicular Technology*, 63(8):3580–3591, 2014.

[7] M. Wang and Q. Ma. Dynamic prediction method of route travel time based on interval velocity measurement system. In *International Conference on Service Operations and Logistics, and Informatics*, pages 172–176, Qingdao, China, October 2014.

[8] Y. Wang G. Shi Y. Zou, H. Shi and H. Zhao. A robust traffic parameter extraction approach for surveillance system at urban intersection. In *International Congress on Image and Signal Processing*, pages 1–5, Tianjin, China, October 2009.

[9] M. Piccardi. Background subtraction techniques: a review. In *International Conference on Systems, Man and Cybernetics*, volume 4, The Hague, Netherlands, October 2004.

[10] D. Arita T. Tanaka, A. Shimada and R. I. Taniguchi. A fast algorithm for adaptive background model construction using parzen density estimation. In *IEEE Conference on Advanced Video and Signal Based Surveillance*, pages 528–533, London, United Kingdom, September 2007.

[11] X. Jiang. Feature extraction for image recognition and computer vision. In *IEEE Conference on Computer Science and Information Technology*, pages 1–15, Beijing, China, August 2009.

[12] S. Zafeiriou G. Tzimiropoulos and M. Pantic. Principal component analysis of image gradient orientations for face recognition. In *IEEE Conference on Automatic Face and Gesture Recognition and Workshops*, pages 553–558, Santa Barbara, USA, March 2011.

[13] P.A. Kashyap D. Mudoi. Vision based data extraction of vehicles in traffic. In *International Conference on Signal Processing and Integrated Networks*, pages 202–208, Noida, India, February 2014.

[14] W. Lu K. Lan, H. Zhang and J. Guo. N-lbp based vehicle monitoring system. In *IEEE International Conference on Network Infrastructure and Digital Content*, pages 701–706, Beijing, China, November 2009.

[15] M. Kurt B. Kir and O. Urhan. Local binary pattern based fast digital image stabilization. *IEEE Signal Processing Letters*, 22(3):341–345, September 2014.

[16] G. Lemaître and M. Radojević. *Directed reading: Boosting algorithms*. Heriot-Watt University, Universitat de Girona, Universite de Bourgogne, December 2009.

[17] H. Ling E. Blasch G. Chen P. Liang, G. Teodoro and L. Bai. Multiple kernel learning for vehicle detection in wide area motion imagery. In *International Conference on Information Fusion*, pages 1629–1636, Singapore, July 2012.

[18] J. Trefny J. Sochman C. Caraffi, T. Vojir and J. Matas. A system for real-time detection and tracking of vehicles from a single car-mounted camera. In *International IEEE Conference on Intelligent Transportation Systems*, pages 975–982, Anchorage, USA, September 2012.

[19] D. W. Hertel and E. Chang. Image quality standards in automotive vision applications. In *IEEE Symposium on Intelligent Vehicles*, pages 404–409, Istanbul, Turkey, June 2007.

[20] H. R. Sheikh Z. Wang, A. C. Bovik and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, April 2004.

[21] G. F. Dominguez M. Nolle R. Pflugfelder, H. Bischof and H. Schwabach. Influence of camera properties on image analysis in visual tunnel surveillance. In *IEEE Conference on Intelligent Transportation Systems*, pages 868–873, Vienna, Austria, September 2005.

[22] P. Zeng Y. Zhang, X. Wang and X. Chen. Centrality characteristics of road network patterns of traffic analysis zones. *Transportation Research Record: Journal of the Transportation Research Board*, (2256):15–24, 2011.

[23] S. Dale J. Paul, B. Malhotra and M. Qiang. Rfid based vehicular networks for smart cities. In *IEEE Conference on Data Engineering Workshops*, pages 120–127, Brisbane, Australia, April 2013.

[24] M. Šoštarić. Model mjerenja tranzitnog prometa na gradskim obilaznicama. Master's thesis, Faculty of Transport and Traffic Sciences, University of Zagreb, Zagreb, Croatia, 2010.