Wieslaw PAMULA[1], Marcin Jacek KŁOS[2]

ON SITE PROCESSING OF VIDEO STREAM FOR MAPPING TRAFFIC PARAMETERS

Summary. Traffic surveillance provides crucial data for the operation of intelligent transportation systems. The growing number of cameras in the transport system poses a problem for the efficient processing of surveillance data. Processing of video data for extracting traffic parameters is usually done using image processing methods and requires substantial processing resources. An alternative way is to transform the video stream and map the traffic parameters using the obtained transform coefficients. Spatiotemporal wavelet transform of the video stream contents, using filter banks, is proposed for mapping traffic parameters. Performed tests prove good resilience to illumination changes of road scenes. Mapping errors are smaller than in the case of the commonly used video detectors at sites on multilane roads with low to moderate traffic load.

Keywords: video surveillance, discrete wavelet transforms, traffic flow, traffic density, intelligent transportation systems

---

[1]Faculty of Transportand Aviation Engineering, The Silesian University of Technology, Krasińskiego 8 Street, 40-019 Katowice, Poland. Email: wieslaw.pamula@polsl.pl. ORCID: https://orcid.org/0000-0001-9792-6528
[2] Faculty of Transportand Aviation Engineering, The Silesian University of Technology, Krasińskiego 8 Street, 40-019 Katowice, Poland. Email: marcin.j.klos@polsl.pl. ORCID: https://orcid.org/0000-0002-4990-1593

1. INTRODUCTION

Intelligent transportation systems integrate information and communication technologies to improve the functioning of road networks and increase the efficiency of moving people and goods [1]. Determining the state of the transport system is decisive for the development of traffic control decisions [2]. Modelling of state changes contributes to establishing traffic management strategies [3,4]. Traditional approaches to measuring road traffic parameters, which constitute the bulk of information on the transport systems state, incorporate inductive loop detectors, magnetic detectors, ultrasonic detectors, radar detectors, and laser detectors [5,6]. Most of these devices can provide only spot data, which means data collected at defined areas of traffic lanes.

The development of intelligent transportation systems is manifested by the growing number of cameras distributed on the road network [7,8]. Hence, it is essential to integrate them into a functional entity to provide traffic data. The large volume of raw image data can pose an acute problem for the efficient processing of the content. Video cameras provide rich contextual information on the course of road traffic. Image processing-based methods are mostly used for extracting traffic parameters [9–11]. The accuracy and reliability of measurements, in this case, are coincident with the complexity of applied image processing algorithms and necessary computing resources. Detection of individual vehicles is a preliminary stage of determining road traffic parameters. The variable parameters of the observation, especially changes in illumination, pose problems for correct detection [12,13]. Background modelling is used to determine the empty road template, which is subtracted from the current video frames. The result depicts moving objects, by default, vehicles [14,15]. Several models developed for different observation environments require modest processing power for implementation and achieve proper vehicle detection mainly for limited changes of the road view [16,17].

The extraction of low-level features of images increases the processing complexity for detecting objects. This approach involves the application of filters and the clustering of filter results. Detection results surpass background subtraction methods but still fall short of expectations in the case of highly changing illumination of the traffic scene. Local feature descriptors, such as Histograms of Gradients (HOG)[11], Scale Invariant Feature Transform (SIFT) [18], Speed-Up Robust Feature (SURF), and Gradient Location Orientation Histogram (GLOH) [19] improve the detection capabilities but impose still higher processing requirements [20,21].

Vehicle motion models once again introduce a higher requirement on processing power. The models are based on the calculation of optical flow and connected region analysis. Horn–Schunck optical flow estimation algorithm [22] is the starting point of modifications in the course of increasing the robustness of vehicle detection. For instance, Peng et al. proposed the use of inter-frame differences for triggering calculations, which significantly reduces the computation burden for updating the optical flow field [23].

Wavelet-based transforms are readily used for the analysis of road traffic parameters. The ability of localizing features and multi-resolution representation of parameter changes are arguments for the application of these transforms. The input data for analysis are predominantly a series of measurements collected at sites of the road network. These are time instants of detections of individual vehicles, vehicle speed values, distances between vehicles, or aggregated quantities. The registered data points are indexed in time order and can be regarded as time series data. Time series analysis comprises frequency domain methods and time domain methods. Wavelet analysis belongs to the frequency domain group of methods. Two main problems are studied: designation of the wavelet basis function of the transform and

determination of the level of decomposition, for effective description of the traffic data. The "ability" of a wavelet to represent different features of traffic data is characterised by the size of the support and the number of vanishing moments, whereas the decomposition level delimits the resolution for the extraction of data attributes.

Early works [7,24] concentrated their efforts on the analysis of traffic flow patterns. The statistical autocorrelation function (ACF) is used for the selection of the decomposition level. This function is usually used to detect trends and seasonality in a time series. In this case, ACF is calculated for the original dataset and wavelet decompositions at different levels of the dataset. Equal ACF values signify the correct choice of the decomposition level. In [25], was proposed the wavelet transform of loop detector data for revealing bottlenecks, transient traffic, and traffic oscillations. Wavelet-based energy peaks from vehicle to vehicle are traced. The duration and intensity of the peaks are processed to obtain traffic features and calculate traffic parameters [25,26]. The task of detecting singularities in noisy traffic data is studied, singularities in traffic data may indicate bottlenecks or traffic incidents.

The problem of video-based traffic surveillance is addressed in [27]. The authors use a two-dimensional discrete wavelet transform for extracting features describing vehicles from the images. Haar wavelet is used as the basis, and the decomposition is done in the space domain. Tests using highway traffic images prove good resilience to shadows on the traffic lanes.

In this paper, spatiotemporal wavelet transform of the video stream from the observation camera is proposed for mapping road traffic parameters instead of applying image processing. The change of contents of the stream represented by transform coefficients, instead of detecting and tracking vehicles, is the basis of the mapping. The mapping of video for representation to traffic parameters is not reported. Reported methods for crowd analysis using video from surveillance cameras share components of this approach. Crowd density is determined using direct processing of video content [28].

The digital form of video data imposes the use of discrete wavelet transform (DWT) versions of the transform. Preliminary tests show that the use of wavelet-based transform of video data retains the characteristics of traffic parameters. A set of detection fields is defined on the image of an observed traffic lane. Passing vehicles are recorded entering these areas. The weighed sum of coefficients of the wavelet transform of a vehicle detection field content corresponds to the traffic density observed on the traffic lane; a similar correspondence is observed for traffic flow.

The primary objective of this paper is to present the idea of the method for the application of the spatiotemporal wavelet transform to map road traffic parameters such as traffic flow and traffic density.

2. MAPPING ROAD TRAFFIC PARAMETERS

The problem of mapping road traffic parameters using transform coefficients of a video stream is the goal of this study. What parameters of a wavelet transform give a good estimation of the road traffic parameters such as traffic flow and traffic density?

The domain of wavelet transforms is chosen as the basis for this study. The literature review gives examples of feature extraction, especially space features, for finding relations with traffic parameters. The proposed idea focuses on the temporal features of the video stream. Decomposition of the video stream in time enables the extraction of features that can describe road traffic. Additional space decomposition reduces the data stream for processing. The task is to determine the levels of decomposition and choose the coefficients that are significant for

mapping road traffic. The mapping results should not diverge substantially from the mappings obtained by commonly used video-based devices - video detectors. The video stream to be processed comes from surveillance cameras mounted above roads leading to town centres. No special cameras were used for the video data collection.

2.1. Road scene model

Traffic flow and traffic density values carry the most important information useful for controlling and managing road traffic. These parameters are the objects of mapping. The input to the mapping is a video stream depicting the changing traffic scene. CCTV cameras are the usual source of video data. Important parameters of the stream are resolution, range of observation, and speed of registration. Video stream F is represented as a sequence of images $I(x, y)$ registered at consecutive moments of time t. It is a discrete entity:

$$F(t, x, y) = \{t, I(x, y): t = 1 \dots K, x = 1 \dots M, y = 1 \dots N\} \tag{1}$$

The values of N, M indicate spatial resolutions – number of rows, columns of the images, moments of time are defined by the speed of registering video data. Fixed lengths of video streams are transformed – K images.

Road traffic combines the movement of vehicles of different sizes and with various dynamic properties. To capture these characteristics, a multi-resolution representation is proposed as the basis for mapping traffic parameters. This approach is related to finding description keys at distinct scales of observation of the traffic. Such descriptions can be nested [29]. Techniques to compute nested sequences of multi-resolution representations are closely related to wavelets. Multiscale representation using wavelets was introduced by S. Mallat in [30].

2.2. Description of video stream contents

Description of image contents is done using a two-dimensional spatial discrete wavelet transforms. To capture changes of the video stream in time, the transforms are extended to include processing in time. The video stream is represented using wavelet coefficients $W_\varphi(j_0, k, m, n)$ and $W_\psi^i(j, k, m, n)$:

$$F(t, x, y) = \frac{1}{\sqrt{KMN}} \sum_k \sum_m \sum_n W_\varphi \varphi_{j_0, k, m, n}(t, x, y)$$

$$+ \frac{1}{\sqrt{KMN}} \sum_{i=1}^{7} \sum_{j=j_0}^{j_r} \sum_k \sum_m \sum_n W_\psi^i \psi_{j, k, m, n}^i(t, x, y) \tag{2}$$

The dyadic scale is used, the scaling function $\varphi_{j_0, k, m, n}(t, x, y)$ and wavelet functions $\psi_{j, k, m, n}^i(t, x, y)$ are defined:

$$\varphi_{j_0, k, m, n} = 2^{j/2} \varphi(2^j t - k, 2^j x - m, 2^j y - n)$$
$$\psi_{j, k, m, n}^i = 2^{j/2} \psi^i(2^j t - k, 2^j x - m, 2^j y - n) \tag{3}$$

Where j – scale, k, m, n – shifts, all integers, the mother wavelet is shifted and scaled by powers of 2.

Separable wavelet functions are used for transforming the video stream, in this case, the functions can be rewritten, exposing the 1D components. There are seven combinations of φ and ψ:

$$
\begin{aligned}
&\psi^1 = \varphi(t)\varphi(x)\psi(y), \quad \psi^2 = \varphi(t)\psi(x)\varphi(y) \\
&\psi^3 = \varphi(t)\psi(x)\psi(y), \quad \psi^4 = \psi(t)\varphi(x)\varphi(y) \\
&\psi^5 = \psi(t)\varphi(x)\psi(y), \quad \psi^6 = \psi(t)\psi(x)\varphi(y) \\
&\psi^7 = \psi(t)\psi(x)\psi(y)
\end{aligned}
\tag{4}
$$

Efficient computing of coefficients is carried out using filters. Transform functions are substituted by filters defined by sets of weights corresponding to the characteristics of the functions. Filter $g(k)$ represents scaling, while $h(k)$ represents wavelets. By applying the filters recursively [30], the coefficients are obtained as:

- approximation

$$
S_{j+1}(k,m,n) = \sum_{t=1}^{K}\sum_{x=1}^{M}\sum_{y=1}^{N} g(t)g(x)g(y)S_j
\tag{5}
$$

- details

$$
D_{j+1}^i(k,m,n) = \sum_{t=1}^{K}\sum_{x=1}^{M}\sum_{y=1}^{N} K^i(t,x,y)S_j
\tag{6}
$$

where K^i are combinations of filters corresponding to (4).

The approximation coefficients are decomposed with combinations of filters and then down sampled. This is represented as a filter bank in Figure 1.

Tab. 1

Prediction and update functions for Deslauriers-Dubuc wavelets

Wavelet	Prediction and update functions
DD(1,1)	$D_{1,i} = S_{0,2i+1} - S_{0,2i}$
	$S_{1,i} = S_{0,2i} + \lfloor D_{1,2i}/2 \rfloor$
DD(2,2)	$D_{1,i} = S_{0,2i+1} - \lfloor (S_{0,2i} + S_{0,2i+2})/2 + 1/2 \rfloor$
	$S_{1,i} = S_{0,2i} + \lfloor (D_{1,i-1} + D_{1,i})/4 + 1/2 \rfloor$
DD(4,4)	$D_{1,i} = S_{0,2i+1} - \lfloor 9(S_{0,2i} + S_{0,2i+2})/16 - (S_{0,2i-2} + S_{0,2i+4})/16 + 1/2 \rfloor$
	$S_{1,i} = S_{0,2i} + \lfloor 9(D_{1,i-1} + D_{1,i})/32 - (D_{1,i-2} + D_{1,i+1})/32 + 1/2 \rfloor$

Table 1 lists the wavelet transforms used in this study. The least demanding computationally wavelet DD(1,1) corresponds to the Haar wavelet. Deslauriers-Dubuc interpolating scaling functions, also known as Interpolets, are good candidates for such applications. To streamline calculations, the lifting scheme is used. The lifting step consists of $P(k)$- prediction, $U(k)$ - update functions (mappings):

$$
\begin{aligned}
D_{n,i} &= D_{n-1,i} + \sum_k P(k)S_{n-1,i-k} \\
S_{n,i} &= S_{n-1,i} + \sum_k U(k)D_{n-1,i-k}
\end{aligned}
\tag{7}
$$

The choice of wavelet basis functions for the transforms is conditioned by the complexity of the calculation. Effective solutions, for instance, incorporating logic-based processing, suitable for on-site designs [9], call for integer based calculations.

2.3. Method for mapping of road traffic parameters

The calculation of spatiotemporal wavelet transform coefficients according to Mallat's scheme results in a set of detail coefficients for every decomposition level and one set of approximation coefficients for the last decomposition level. Decomposition level labels the consecutive step of application of the set of filters to S_j in the course of calculation of the transform coefficients. The filters are applied to pixel values in 3 dimensions (t, x, y). The notation (l_t, l_x, l_y) means that changes in time of pixel values are filtered l_t times, whereas changes at space positions are filtered l_x and l_y times.

The lower the decomposition level, the larger the number of coefficients in the set. The question arises, which coefficients carry significant clues for describing traffic parameters and which can be discarded. If the length of the transformed stream is $2k$ at the k decomposition level in time, there are seven sets of detail coefficients and one set of approximation coefficients. Sets have the same number of elements defined by the level of space decomposition and the size of the input image. Input image of the size 512×512 pixels, decomposed at the third level, is represented by sets of $(512/2^3) \times (512/2^3) = 4096$ coefficients.

Only the coefficients at the highest level of decomposition are used for mapping. These carry the synthetic description of the image contents changes at patches of the size $(512/2^3)$ and in a period of $2k$ frames of the video stream. A sum of weighted coefficients is proposed for mapping traffic parameters q:

$$q = w_0 S_j + \sum_{i=1}^{7} w_i D_j^i \tag{8}$$

Weight values $W = [w_0, w_1, \ldots, w_7]^T$ depend on the way the traffic is observed. Camera position relative to traffic lanes defines the angle of observation, deformation of observed vehicles, light, and contrast of the traffic scene. A calibration procedure is required to obtain the values of the weights. Weights are calculated based on previous observations of the traffic.

The observed traffic lane is represented by sums of coefficients c, calculated for $2k$ frames of the video stream that cover the area of the traffic lane.

$$q = w_0 S_j + \sum_{i=1}^{7} w_i D_j^i \quad c_0 = \sum_a S_j \qquad c_0 = \sum_a D_j^i \tag{9}$$

These are calculated for a number of periods s and used as input data for calculating weights:

$$C = \begin{bmatrix} c_{10}, c_{11}, \ldots, c_{17} \\ c_{20}, c_{21}, \ldots, c_{27} \\ \ldots \\ c_{s0}, c_{s7}, \ldots, c_{s7} \end{bmatrix} \tag{10}$$

For s previous measurements of the traffic parameter $Q_p = [q_1, q_2, \ldots, q_s]^{\mathrm{T}}$ the over-determined system $Q_p = CW$ defines the calculation task. The goal is to find the best weights W in the sense of solving the quadratic minimization problem with the objective function:

$$S(W) = \sum_{i=1}^{s} \left| q_i - \sum_{j=0}^{7} c_{ij} w_j \right|^2 = \| Q_p - CW \| \tag{11}$$

which gives:

$$W = (C^T C)^{-1} C^T Q_p \tag{12}$$

The derived set of weights is specific for a measurement site.

3. RESULTS AND COMPARISON WITH PERFORMANCE OF VIDEO DETECTORS

Mapping of traffic flow and traffic density using wavelet transform coefficients is examined in this section. Road traffic data collected at several sites are used for calculating the weights of the representations (equation 12). Figure 1 shows examples of camera sites where multilane roads with high to low traffic loads are observed.

Fig. 1. Camera sites views: a) high traffic, b) medium traffic, c) low traffic

The range of observation is limited by the acceptable sizes of vehicles expressed in the pixels of the image. Image resolution and noise level impose the condition that the smallest vehicle size should be a few hundred pixels. This defines a field of view not longer than 150 metres when a standard CCTV camera is used. The highest level of wavelet transform decomposition in the space of the image is determined by the need to preserve vehicle representations. A CCTV image of the size 720×568 split into 25×25 patches satisfies the considered limits. This limits the level of decomposition in space to 5.

The level of wavelet transform decomposition in time k is established by the requirements of updating traffic parameters. The update period is defined by the duration of $2k$ frames of the video stream. Weights for mapping road traffic parameters are calculated for the collected traffic data, and the errors of mappings are evaluated. Mean absolute percentage error (MAPE) and root mean square percentage error (RMSPE) values are used for comparing the results.

3.1. Measurement sites

Data from three measuring sites localized on multilane roads are used for the tests – Figure 1. The sites have highly illuminated and shadowed lanes. Vehicles travelling on the illuminated road lane cast shadows on the parallel lane, causing it to be shadowed. The parallel lane contains combined shadows from both traffic lanes, and this is a source of errors. The moving sun may change the proportion of shadows on the lanes if the lanes are north- or south-bound; in this case, the proportions are constant. The lane with more shadows is named shadowed.
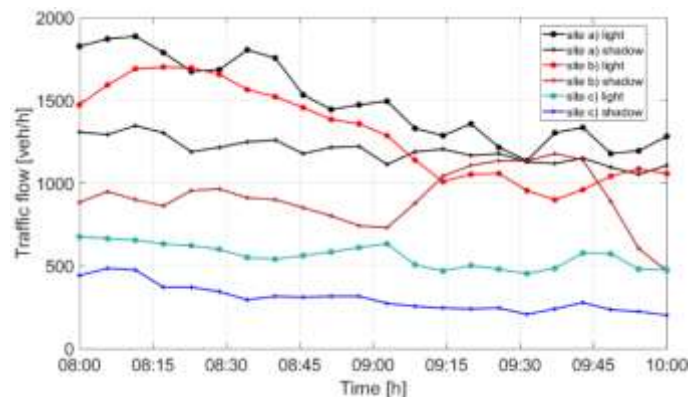

Fig. 2. Graphs of traffic flow values recorded at the measurement sites

This shadowing phenomenon is of particular interest as the source of measurement errors. In the case of video detectors, it usually causes extra vehicle detections. Morning traffic parameters are measured. Three measurement sites differing in the size of the traffic flow are selected: high, medium, and low traffic flow. Figure 2 illustrates the changes in values of traffic flow at the measurement sites during the measurement time period. High traffic flow surpasses 1800 veh./h per lane, whereas the site with low traffic has flow values below 500 veh./h per lane. The largest flow differences between these two road lanes are noted at site (b) with the medium volume of traffic.

3.2. Comparison of results of the proposed method

The proposed approach for mapping traffic parameters is compared with the performance of video-based measurement devices present at traffic sites. Two devices are chosen to represent the current state of video-based vehicle detection technology. Both devices use basic image processing techniques for determining the presence of objects in the detection fields predefined on observed images of the road. The objects assumed to be vehicles are counted, and the times of their entry and duration of presence in the detection fields are used to calculate traffic parameters such as flow and density.

The algorithm of detection of the first device, A, tracks the content of the detection field and when it substantially differs from the model of the observed background an object's presence is signalled. The background is modelled statistically using one probability distribution. An example of the implementation of this principle of operation is protected by a patent [29]. The second device – B uses a more complex detection principle in which both the background and the detection fields are modelled using a fuzzy-based feature update algorithm, and when the two models differ an object's presence is signalled.

The carried out mappings using DD(2,2) and DD(4,4) show no significant performance advantage over DD(1,1). The calculation of wavelet coefficients requires, in these instances, much more processing resources, which impairs on-site implementations. Representative results of mapping traffic parameters using DD(1,1) are further discussed. Traffic flow and traffic density are mapped using coefficients of DD(1,1) transforms of video streams of road scenes.

3.3. Traffic flow

DD(1,1) with varying decomposition parameters is used for mapping traffic flow. Second, third and fourth space decomposition levels and 13th and 14th temporal decomposition levels are investigated. The best mapping results are obtained for the set (14,3,3). None of the transform coefficients explicitly outweighs the others; this indicates that the camera observation parameters decisively determine the weights. Larger values of weights are noted for mapping most of the flow values on highly illuminated road lanes than on corresponding shadowed lanes.

The video database consists of non-compressed films of road lanes recorded at the measurement sites. This material is inputted to the vehicle detection devices – video detectors, in real-time, and the detection results are recorded. Standard detection settings were used. The obtained values are matched with the reference sets of traffic parameter values.

Table 2 summarizes the results of measuring traffic flow. MAPE error values for all measurement sites are lower than RMSPE values. MAPE values are less sensitive to outliers in comparison to RMSE values. The difference does not exceed ¼ of the MAPE, indicating a few outliers.

Tab. 2

Values of mapping errors for traffic flow

Mapping errors [%]		Measurement sites					
		High traffic		Medium traffic		Low traffic	
		Light	Shadow	Light	Shadow	Light	Shadow
Video detector A	RMSPE	25	26	27	13	15	16
	MAPE	24	22	24	10	14	15
Video detector B	RMSPE	10	4,9	19	6,7	19	35
	MAPE	8,9	3,8	14	5,0	16	33
Proposed method	RMSPE	5,9	7,8	21	16	9,9	20
	MAPE	3,8	4,1	16	12	8,0	15

Traffic on illuminated lanes is more accurately mapped than on the shadowed lanes, except in the case of medium traffic. The graph in Figure 3 shows that at the medium traffic site, traffic flow changes are more volatile than at the other sites. Examination of the video shows that large errors arise when vehicles temporarily slow down or stop due to abrupt changes in traffic density (traffic jams), and this is not captured by the transform. Some errors are caused by container trucks travelling in bunches. Higher placement of the observation camera can remedy this weakness.
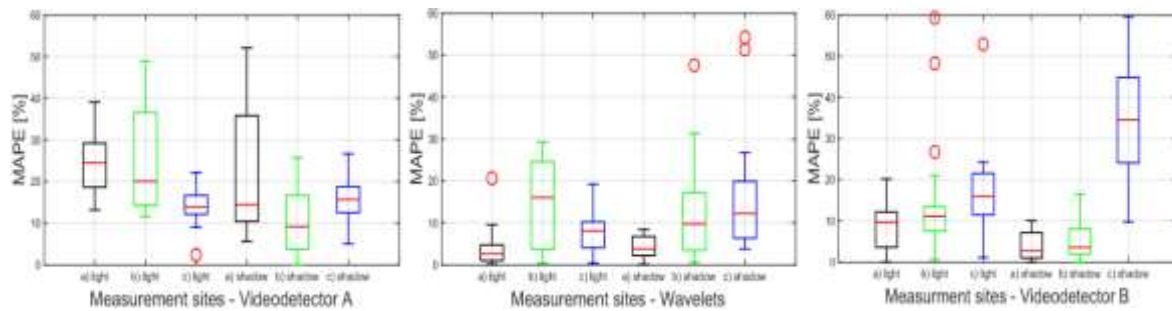
Fig. 3. Mapping errors of traffic flow values

Differences in RMSPE and MAPE error values are small, although video detector B shows a larger number of outliers. Video detector A copes better with low traffic, while video detector B with high traffic. The proposed transform-based processing performs better than the video detectors. There are no outstanding error values. Box plots presented in Figure 3 illustrate the error statistics in detail.

Video detector B gives smaller errors than the mappings, but there are numerous outliers. Detailed inspection of the video material shows that these are the consequences of stopped vehicles, as it is in the case of mappings but the results generate much higher error values.

3.4. Traffic density

Again wavelet DD(1,1) with varying decomposition parameters is used for mapping traffic density. The same range of decomposition parameters is applied. The best mapping results are obtained for the set (14,3,3). Table 3 presents the errors in mapping traffic density. In comparison to flow mapping weights, the density mapping weights are substantially different. Some weights have very small values for all examined measurement sites. This can be of use in optimizing processing operations for calculating traffic density.

Tab. 3

Values of mapping errors for traffic density

Mapping errors [%]		Measurement sites					
		a) High traffic		b) Medium traffic		c) Low traffic	
		Light	Shadow	Light	Shadow	Light	Shadow
Video detector A	RMSPE	10	14	9,5	9,5	17	38
	MAPE	7,7	12	7,0	7,9	14	32
Video detector B	RMSPE	11	6,0	22	8,1	17	38
	MAPE	10	5,2	17	6,6	12	26
Proposed method	RMSPE	9,0	19	14	14	12	22
	MAPE	7,2	15	12	11	10	15

Mapping errors follow the same pattern as in the case of traffic flow. In all, errors are larger, especially at the high traffic site. Box plots presented in Figure 4 illustrate the error statistics in detail.
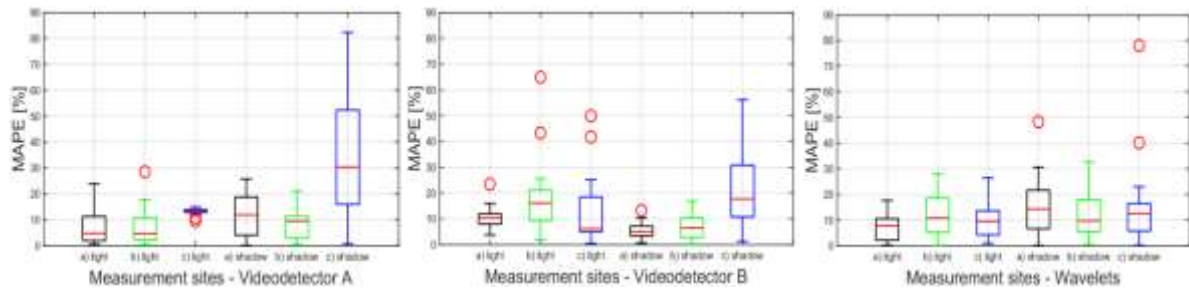
Fig. 4. Mapping errors of traffic density values

Differences in RMSPE and MAPE error values are small for the proposed method. Traffic density at low traffic sites on shadowed lanes is determined with the largest errors by both video detectors. This poor performance may be linked to losing infrequently passing vehicles due to inadequate detection ability of objects partially covered by the shadows, which disrupt the object's view.

4. DISCUSSION

Table 4 summarises the comparison of the performance of video detectors and wavelet mappings. The advantage of the proposed method is not significant but the consistency of the mapping - there are no outliers, is important for traffic control and management systems.

The processing algorithm of video detector B presumably loses vehicles due to poor sensitivity to infrequently passing objects on the image. This may be caused by the parameters of updating the background model in the device. Similarly, several outliers in the case of medium traffic at an illuminated site also suggest that such conditions pose momentary difficulties in discerning and tracking features.

Tab. 4

Average errors in mapping and measuring traffic parameters for all sites

Average errors [%]	Video detector A		Video detector B		Proposed method	
	flow	density	flow	density	flow	density
RMSPE	20	16	16	17	13	15
MAPE	18	13	13	13	10	12

The proposed method maps traffic flow and traffic density more accurately than commonly used video-based vehicle detection devices. In the case of high and low traffic, the ranges of errors are substantially lower. High momentary errors are recorded when the video contents reveal a stopped vehicle, which caused numerous lane changes by vehicles approaching this obstacle. High errors are also caused by queues of container trucks. These situations are less effectively represented by the transform coefficients especially related to temporal changes. The coefficient values indicate the scale of changes in time at different time resolutions. A high level of decomposition diminishes the sensitivity to high speed changes of contents, which are induced by such traffic situations.

Large vehicles present in the traffic lanes cause error fluctuations. Another level of decomposition can be chosen to alleviate the deficiency of different size object mapping in the course of transforming the video data. This approach should consider the characteristics of the observed road, that is, whether it is a transit road with heavy vehicles or an urban road mainly with cars.

The errors in measuring traffic density are higher than traffic flow; it can be attributed to the higher influence of illumination changes in deriving the results. Modelling background as well as calculating wavelet coefficients is susceptible to noise. Changing illumination values can be regarded as a noise factor with highly volatile probability distribution parameters.

The advantage of the transform-based approach lies in the reduction of computing operations for obtaining the mapping of traffic parameters. For instance, background subtraction requires background modelling involving statistical calculations using image pixel neighbourhoods that are hundreds of calculations per image pixel. Observation cameras provide video streams with a resolution of 720×576 at 25 frames a second, which amounts to over 10 MB/s when a pixel is represented using a single byte. The commonly implemented background model uses a mixture of Gaussians, usually 3, updated every video frame requiring at least 1-2 Gips.

Transform calculations may be done in a processing pipeline using a non-processor based device. Implementation of calculations in all requires tens of operations per pixel, which are performed in parallel, at the speed of the incoming pixels.

5. CONCLUSIONS

The proposed method enables the mapping of road traffic parameters on multilane roads with smaller errors than the solutions currently implemented in video detecting devices. The video detecting devices perform poorly, especially when the road image is corrupted by shadows of vehicles travelling on adjacent traffic lanes.

The spatiotemporal wavelet transform, by selecting different decomposition parameters, allows for the representation of features at different resolutions in time and space. It represents the features of objects at different scales - by choosing a decomposition level, it is possible to "filter out" vehicles with different sizes or characteristic details of appearance. This makes it possible to identify the position of individual vehicles in the video stream. The time transformation describes the dynamics of changes in the movement of the vehicles. This information is useful for mapping the changes in traffic parameters.

The discrete wavelet transform can be implemented using the lifting scheme, significantly reducing the required computation budget. The application of an embedded processing system comprising of a field programmable gate array can efficiently calculate transform coefficients in real-time. Such one chip solutions can be integrated with traffic monitoring cameras and function as traffic data collection subsystem nodes in intelligent transportation systems.

Finally, the proposed method of mapping road traffic parameters proves that a set of weighed coefficients of a wavelet transform give a credible estimation of road traffic parameters, such as traffic flow and traffic density. Hence, the proposed method requires further studies in the optimization of the processing algorithms suitable for available hardware resources.

References

1. Bok Jinjoo, Youngsang Kwon. 2016. „Comparable Measures of Accessibility to Public Transport Using the General Transit Feed Specification." *Sustainability* 8(3). DOI: 10.3390/su8030224.
2. Rith Monorom, Alexis Fillone, Jose Bienvenido M. Biona. 2019. „The Impact of Socioeconomic Characteristics and Land Use Patterns on Household Vehicle Ownership and Energy Consumption in an Urban Area with Insufficient Public Transport Service – A Case Study of Metro Manila." *Journal of Transport Geography* 79: 102484. DOI: 10.1016/j.jtrangeo.2019.102484.
3. Zhang Tingru, Alan H.S. Chan, Hongjun Xue, Xiaoyan Zhang, Da Tao. 2019. "Driving Anger, Aberrant Driving Behaviors, and Road Crash Risk: Testing of a Mediated Model."*International Journal of Environmental Research and Public Health* 16(3): 1-13. DOI: 10.3390/ijerph16030297.
4. Ortega Jairo, János Tóth, Tamás Péter. 2021. „Planning a Park and Ride System: A Literature Review."*Future Transportation* 1(1): 82-98. DOI: 10.3390/futuretransp1010006.
5. Federal Highway Administration. 2016. „Traffic Monitoring Guide FHWA." Fhwa. Available at: http://www.fhwa.dot.gov/policyinformation/tmguide/.
6. Klein Lawrence A. 2017. *ITS Sensors and Architectures for Traffic Management and Connected Vehicles*. Boca Raton : Taylor & Francis, CRC Press. DOI: 10.1201/9781315206905.
7. Jiang Xiaomo, Hojjat Adeli. 2004. „Wavelet Packet-Autocorrelation Function Method for Traffic Flow Pattern Analysis." *Computer-Aided Civil and Infrastructure Engineering* 19(5): 324-37. DOI: 10.1111/j.1467-8667.2004.00360.x.
8. Mandellos Nicholas A., Iphigenia Keramitsoglou, Chris T. Kiranoudis. 2011. „A Background Subtraction Algorithm for Detecting and Tracking Vehicles." *Expert Systems with Applications* 38(3): 1619-31. DOI: 10.1016/j.eswa.2010.07.083.
9. Tasgaonkar Pankaj P., Rahul Dev Garg, Pradeep Kumar Garg. 2020. „Vehicle Detection and Traffic Estimation with Sensors Technologies for Intelligent Transportation Systems." *Sensing and Imaging* 21(1). DOI: 10.1007/s11220-020-00295-2.
10. Singleton Patrick A., Keunhyun Park, Doo Hong Lee. 2021. „Varying Influences of the Built Environment on Daily and Hourly Pedestrian Crossing Volumes at Signalized Intersections Estimated from Traffic Signal Controller Event Data." *Journal of Transport Geography* 93: 103067. DOI: 10.1016/j.jtrangeo.2021.103067.
11. Buch Norbert, Sergio A. Velastin, James Orwell. 2011. „A Review of Computer Vision Techniques for the Analysis of Urban Traffic." *IEEE Transactions on Intelligent Transportation Systems* 12(3): 920-39. DOI: 10.1109/TITS.2011.2119372.
12. Semertzidis T., K. Dimitropoulos, A. Koutsia, N. Grammalidis. 2010. „Video Sensor Network for Real-Time Traffic Monitoring and Surveillance." *IET Intelligent Transport Systems* 4(2): 103-12. DOI: 10.1049/iet-its.2008.0092.
13. Roy Arunesh, Nicholas Gale, Lang Hong. 2011. „Automated Traffic Surveillance Using Fusion of Doppler Radar and Video Information." *Mathematical and Computer Modelling* 54(1-2): 531-43. DOI: 10.1016/j.mcm.2011.02.043.
14. Xu Yong, Jixiang Dong, Bob Zhang, Daoyun Xu. 2016. „Background Modeling Methods in Video Analysis: A Review and Comparative Evaluation." *CAAI Transactions on Intelligence Technology* 1(1): 43-60. DOI: 10.1016/j.trit.2016.03.005.

15. Garcia-Garcia Belmar, Thierry Bouwmans, Alberto Jorge Rosales Silva. 2020. „Background Subtraction in Real Applications: Challenges, Current Models and Future Directions." *Computer Science Review* 35: 100204. DOI: 10.1016/j.cosrev.2019.100204.
16. Mitrović Dejan. 2005. „Reliable Method for Driving Events Recognition." *IEEE Transactions on Intelligent Transportation Systems* 6(2): 198-205. DOI: 10.1109/TITS.2005.848367.
17. Lee Uichin, Mario Gerla. 2010. „A Survey of Urban Vehicular Sensing Platforms." *Computer Networks* 54(4): 527-44. DOI: 10.1016/j.comnet.2009.07.011.
18. Salih Yasir, Aamir Saeed Malik. 2011. „Comparison of Stochastic Filtering Methods for 3D Tracking." *Pattern Recognition* 44(10-11): 2711-37. DOI: 10.1016/j.patcog.2011.03.027.
19. Karasulu Bahadir, Serdar Korukoglu. 2012. „Moving Object Detection and Tracking by Using Annealed Background Subtraction Method in Videos: Performance Optimization." *Expert Systems with Applications* 39(1): 33-43. DOI: 10.1016/j.eswa.2011.06.040.
20. Guo Yulan, Mohammed Bennamoun, Ferdous Sohel, Min Lu, Jianwei Wan, Ngai Ming Kwok. 2016. „A Comprehensive Performance Evaluation of 3D Local Feature Descriptors." *International Journal of Computer Vision* 116(1): 66-89. DOI: 10.1007/s11263-015-0824-y.
21. Kihl Olivier, David Picard, Philippe-Henri Gosselin. 2015. „A Unified Framework for Local Visual Descriptors Evaluation." *Pattern Recognition* 48(4): 1174-84. DOI: 10.1016/j.patcog.2014.11.013.
22. Horn Berthold K.P., Brian G. Schunck. 1981. „Determining Optical Flow." *Artificial Intelligence* 17(1-3): 185-203. DOI: 10.1016/0004-3702(81)90024-2.
23. Peng Yanan, Zhenxue Chen, Q.M. Jonathan Wu, Chengyun Liu. 2018. „Traffic Flow Detection and Statistics via Improved Optical Flow and Connected Region Analysis." *Signal, Image and Video Processing* 12(1): 99-105. DOI: 10.1007/s11760-017-1135-2.
24. Adeli Hojjat, Samanwoy Ghosh-Dastidar. 2004. „Mesoscopic-Wavelet Freeway Work Zone Flow and Congestion Feature Extraction Model." *Journal of Transportation Engineering* 130(1): 94-103. DOI: 10.1061/(ASCE)0733-947X(2004)130:1(94).
25. Zheng Zuduo, Soyoung Ahn, Danjue Chen, Jorge Laval. 2011. „Applications of Wavelet Transform for Analysis of Freeway Traffic: Bottlenecks, Transient Traffic, and Traffic Oscillations." *Transportation Research Part B: Methodological* 45(2): 372-84. DOI: 10.1016/j.trb.2010.08.002.
26. Zheng Zuduo, Soyoung Ahn, Danjue Chen, and Jorge Laval. 2011. „Freeway Traffic Oscillations: Microscopic Analysis of Formations and Propagations Using Wavelet Transform." *Procedia - Social and Behavioral Sciences* 17: 702-16. DOI: 10.1016/j.sbspro.2011.04.540.
27. Ibtissam Slimani, Abdelmoghit Zaarane, Abdellatif Hamdoun, Issam Atouf. 2018. „Traffic Surveillance System For Vehicle Detection Using Discrete Wavelet Transfor." *Journal of Theoretical and Applied Information Technology* 96(17). DOI: 10.13140/RG.2.2.34426.08649.
28. Bendali-Braham Mounir, Jonathan Weber, Germain Forestier, Lhassane Idoumghar, Pierre-Alain Muller. 2021. „Recent Trends in Crowd Analysis: A Review." *Machine Learning with Applications* 4: 100023. DOI: 10.1016/j.mlwa.2021.100023.
29. US4847772A. Michalopoulos Panos G., A. Richard Fundakowski, Meletios Geokezas, Robert C. Fitch. 1989. "Vehicle detection through image processing for traffic surveillance and control". Patent number 4,847,772. Available at: https://patents.google.com/patent/US4847772A/en.

30. Mallat Stephane G. 2009. „Multiresolution Approximations and Wavelet Orthonormal Bases of L2(R).” *Fundamental Papers in Wavelet Theory* 315(1): 524-42. DOI: 10.1090/s0002-9947-1989-1008470-5.