



Article citation information:

Pamuła, T. Classification of road traffic conditions based on texture features of traffic images using neural networks. *Scientific Journal of Silesian University of Technology. Series Transport*. 2016, **92**, 101-109. ISSN: 0209-3324. DOI: 10.20858/sjsutst.2016.92.10.

Teresa PAMUŁA¹

**CLASSIFICATION OF ROAD TRAFFIC CONDITIONS BASED ON
TEXTURE FEATURES OF TRAFFIC IMAGES USING NEURAL
NETWORKS**

Summary. The paper presents a method of classification of road traffic conditions based on the analysis of the content of images of the traffic flow. The view of the traffic lanes with vehicles is treated as a texture, while the change in the description of its characteristics is ascribed to the change in the density of traffic. Four classes of conditions are determined on the basis of the values of Haralick texture features. An MLP network is used for classification. Video data, which were registered by an UAV hanging over a traffic junction, are used for validation of the method.

Keywords: traffic conditions; textures features, neural networks.

1. INTRODUCTION

The problem of road traffic congestion heavily affects the performance of traffic control and management systems. The way in which congestion evolves and spreads in the traffic networks is important for devising effective counter measures for diminishing its destructive character. Automatic methods of describing and measuring the course of traffic density changes are especially desired, as control systems must act in real time. Many traffic control systems incorporate video monitoring subsystems for detecting traffic incidents. This

¹ Faculty of Transport, Silesian University of Technology, Krasińskiego 8 Street, 40-019 Katowice, Poland.
E-mail: teresa.pamula@polsl.pl.

provides an opportunity for monitoring the conditions of traffic and related parameters, such as traffic density, traffic flow, mean traffic speed and average delays.

Image processing methods are widely used for determining road traffic parameters [4]. Video-based detectors are utilized for counting and classifying vehicles. Surveillance systems are capable of tracking vehicles, measuring their speed and gathering their travel parameters. Most applications isolate individual vehicles in the sequence of images in the course of processing. Processing is focused on finding the description of the behaviour of the objects (vehicles). This approach is very demanding for the design of the computing parts of the systems.

Traffic conditions determine the way in which road users perceive their ability to travel along the road network. Traffic conditions are described with the use of the measure of this freedom of movement, known as levels of service (LOS). Depending on the country specifics, up to six LOS are defined. Although the American HCM defines six levels and the German HBS defines five [3], the following four levels are used to describe traffic conditions in Poland: LOS I = very good conditions; LOS II = good conditions; LOS III = medium; LOS IV = unfavourable. Average delays and queue lengths are the main factors taken into account in the classification of traffic conditions. Observed traffic density on the road is directly related to these parameters. It is clear that a method of classification of the course traffic density change is adequate to describe the conditions, while no precise measurement of the traffic parameters is necessary.

In effect, the task of classifying traffic conditions can be interpreted, in the field of image processing, as a task of classifying the scene content, where the scene depicts road lanes with traffic. The need for the isolation of objects becomes superfluous. Various ways of scene content classification are used in practice. The common approach is to use texture features.

The texture approach to the analysis of image content for determining traffic conditions is proposed by a number of research teams. Qiao and Shia [6] propose using edges and texture parameters of car bodies to isolate them. Movement parameters are determined on the basis of the video inter-frame difference analysis. A local binary pattern visual descriptor is used for texture description. Although the regions of interest in the image are defined, no attempt is made to compute descriptions for the entire regions. The regions are used to limit the area of vehicle detection and enable efficient calculation of traffic parameters, such as the stationary queue length and lane occupation.

In [1], vehicle corners and background characteristics are used in the representation of traffic state features. This approach does not exactly correspond with texture analysis, but is based on the idea of classification of the scene content without extracting individual vehicles. Classifiers are designed, which use background model histogram parameters and vehicle corners to recognize the following traffic states: non-blocked, slow, and jammed at different times of day.

Authors of the work presented in [2] use texture features, which are calculated using grey level co-occurrence matrices (GLCMs) for traffic congestion detection. Arbitrarily chosen features are energy and entropy, which are computed using four GLCMs representing horizontal, vertical and diagonal features in the image. The registered images are, at first, converted to 32 level grey images in order to diminish the necessary calculation burden. An experimentally established threshold enables the classification of the image content into jammed and free traffic states.

Neural networks (NNs) are proven to be efficient tools for classification tasks [8] when the input variables do not show a clear impact on the classification result. The results presented in [9] illustrate this approach. The traffic flow class is recognized using values from

a window in the time series of traffic flow. The size of the window is minimized simply to retain the recognition capability, thus reducing, at most, the signalling delay. Outputs of the NNs are to be used for signalling the need to change a traffic control or management strategy for use in intelligent transportation system applications.

The main contribution of the paper is the design of the method for using a set of Haralick texture features for classification of traffic conditions (LOS). An optimal set of texture features and a NN configuration is proposed to determine LOS. The small computing requirements will promote the use of this method for real-time applications in road traffic control systems.

The proposed method of classification consists of the following steps:

- acquisition and preprocessing of the video sequences of traffic scenes
- choice of texture features
- NN set-up, training and classification tuning

The solution consists of three stages, which are presented and discussed in the following sections of the paper. These are supplemented with the results of training and classification runs. The concluding section summarizes the properties of the solution and proposes further investigation topics.

2. ACQUISITION AND PREPROCESSING OF VIDEO SEQUENCES

The method is developed using sequences of video captured by a camera carried by a UAV. The image covers a rectangular area of about 100x200 m seen from the height of 100 m. It depicts a major junction with heavy traffic.

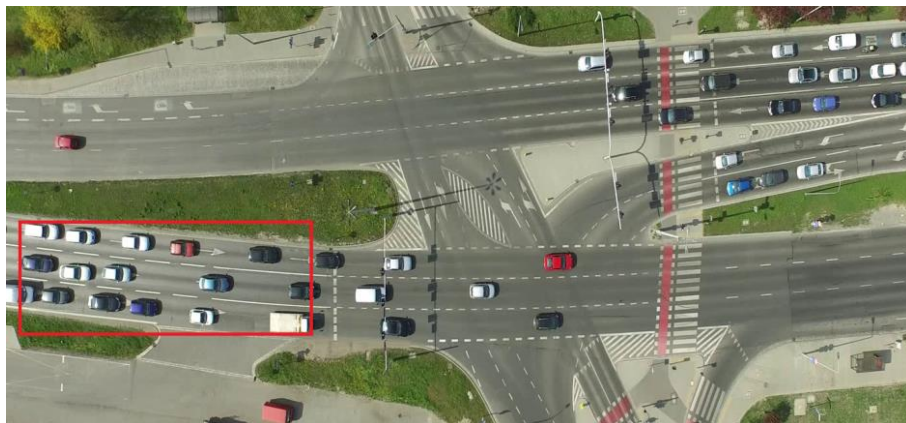


Fig. 1. Traffic scene image registered by a UAV hanging over a junction; red rectangle marks the analysed section of the road

A rectangular patch is cut out from the image (red rectangle in Fig. 1) and this is the focus of the texture analysis. This covers a stretch of the three-lane approach road to the junction.

Haralick texture features are used in the analysis, which require the determination of grey level co-occurrence matrices. A series of preliminary tests were carried out to determine the necessary number of grey levels for the effective texture representation of scenes with vehicles. Simple binning was used to reduce the number of grey levels.

$$P_{ij} = (\text{conv}(C_{ij}, \text{RGB_TO_256_GRAY}) \text{ div } 32) \cdot 32 \quad (1)$$

where P_{ij} represent eight level image pixels and C_{ij} represents the original pixels from the camera in the RGB colour space.

The tests prove that the reduction in the number of levels, from the usually used 256 to eight, does not significantly deteriorate the capabilities for describing textures in scenes with vehicles. This may be substantiated by the characteristics of the vehicles, which have regular shapes and mostly homogenously coloured bodies. Figure 2 gives an example of a converted image patch.

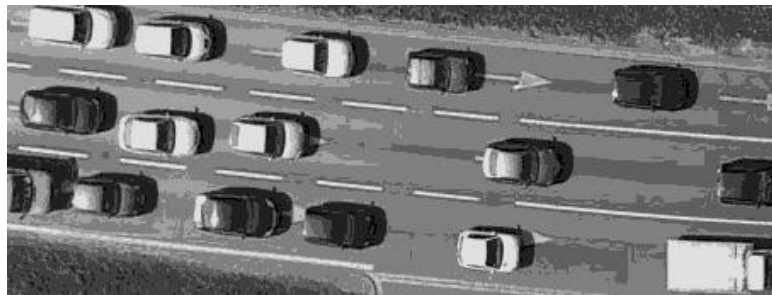


Fig. 2. Image patch converted to eight grey levels of representation

This reduction in the grey levels brings about a considerable time saving in computing the GLCMs

3. TEXTURE FEATURES

Haralick texture features [7] are used to describe the content in images of traffic scenes. These widely used features are based on statistics given by pairs of pixels. Although Haralick defined 14 texture features, other researchers subsequently extended this number to over 20.

3.1. Haralick texture features

Out of the original set of 14 features, six are used in this investigation. Table 1 summarizes the formulas for calculation of the features using GLCMs. N represents the number of grey levels, while $M_{i,j}$ represents the value of the GLCM entry.

Table 1. Texture features

Texture feature	Pixels statistics
Energy	$\sqrt{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} M_{i,j}^2}$
Entropy	$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} M_{i,j} (-\ln M_{i,j})$

Contrast	$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} M_{i,j} (i-j)^2$
Homogeneity	$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{M_{i,j}}{1+(i-j)^2}$
Dissimilarity	$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} M_{i,j} i-j $
Correlation	$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} M_{i,j} \left(\frac{(i-\mu_i)(j-\mu_j)}{\sigma_i \sigma_j} \right)$ μ_i, μ_j are the means and σ_i, σ_j are the standard deviations of the row and column sums

Energy and entropy are measures of the arrangement of objects, homogeneity describes the perception of “smoothness” of the image content, contrast indicates the amount of local variations present in the image, and correlation is a measure of linear dependencies in the image. Energy reaches high values for pixels of constant or repeating values. High values of entropy are obtained for complex pixel patterns in the image.

3.2 Grey level co-occurrence matrix

A GLCM is a square NxN matrix, where N is the number of pixel values, which contains frequencies of occurrence M_i , and j represents the pixel values that lie at set distances from each other. A matrix is generated for each displacement. For instance, for a displacement defined by the vector (1.0), the table contains entries, which are the number of pixels with the value i that lie to the right of pixels with the value j . Figure 3 illustrates the calculation procedure.

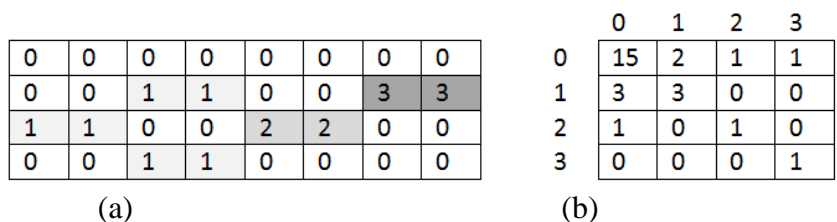


Fig. 3. (a) image pixel values; (b) GLCM

The size of the GLCM is determined by the square of N. In the case of 256 grey level images, there are 65,536 elements of the matrix. Eight level grey images require 64 element matrices, while matrices for binary images have just four elements. A texture description usually consists of several matrices, while four vectors are commonly used, which represent the texture “directions” 0°, 90°, 45°, 135°, that is, horizontal, vertical and two diagonals. This directional description is useful for classification of content with movement, where the change of values in the corresponding movement matrices will indicate the direction of change of the position.

As the test sequences in this investigation contain only horizontally moving vehicles, only the 0° matrix is used. The displacement vector is defined as a horizontal unit displacement.

2.3 Choice of texture features

Ranking is used to find a set of the most sensitive texture features for determining the LOS. Sixty image patches are processed and all texture features are calculated. In order to enable efficient comparison of their properties, the values are normalized. Figure 4 contains graphs of the feature values superimposed on the graph of LOS values, which were determined manually.

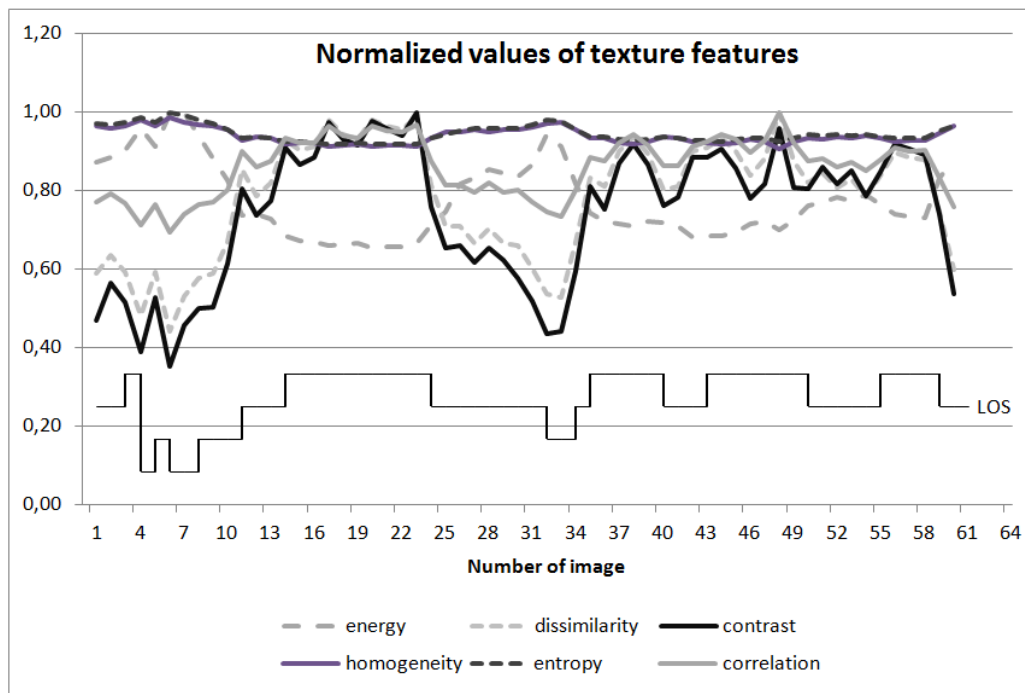


Fig. 4. Graphs of normalized texture features of images

The graph showing texture homogeneity is almost flat, which means that it is insensitive to LOS changes. The remaining texture features change with LOS. Most sensitive are contrast, energy and dissimilarity.

3. NEURAL NETWORK SET-UP

The validation of the performance of the proposed NN for classification of traffic conditions is carried out using a test sequence of 70 images. Sixty images are used for training and the remainder for testing the performance.

Table 2. LOS and corresponding content of the image patch

LOS	Number of vehicles
LOS I	0-5
LOS II	6-10
LOS III	11-14
LOS IV	>14

The LOS, defined in terms of vehicle densities, are scaled to the size of the analysed image patch; Table 2 summarizes the obtained values. The classification task becomes a task of distinguishing four textures defined by five ranges of texture features. Figure 5 contains examples of images with different LOS classes.

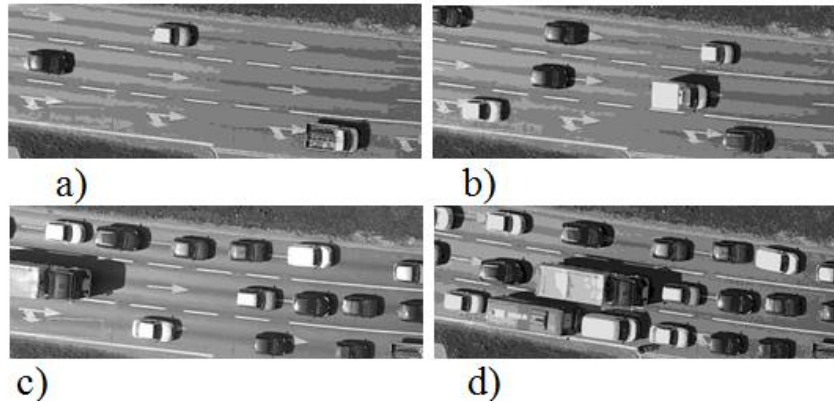


Fig. 5. Images of traffic conditions: (a) LOS I; (b) LOS II; (c) LOS III; (d) LOS IV

Images containing large vehicles, such as trucks, may pose a classification problem as they highly disrupt the texture.

3.1. Neural network configuration

A back-propagation NN, with five inputs and four outputs and one hidden layer, is proposed for investigations (Fig. 6). There are 10 neurons in the hidden layer. The number of neurons in the hidden layer is determined experimentally on the basis of authors' previous experience in constructing and classifying NNs. Normalized values of texture features constitute the inputs: C1 = energy; C2 = entropy; C3 = dissimilarity; C4 = contrast; C5 = correlation. The neurons use sigmoidal transfer functions.

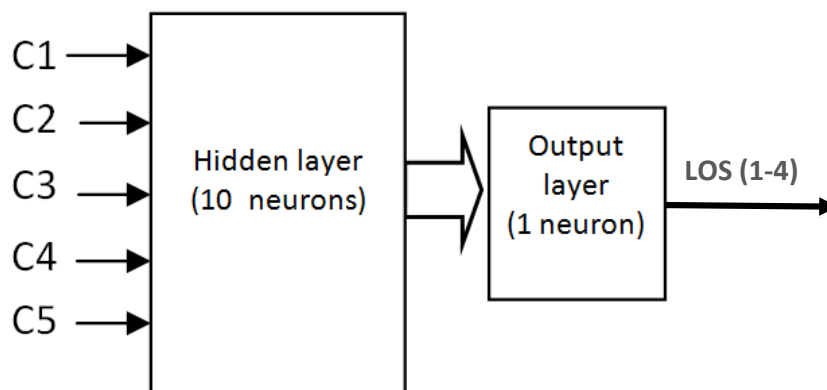


Fig. 6. NN structure

The training sequence consisted of sixty input vectors and the corresponding LOS classes. The training was stopped when the mean square error of less than 0.001 was attained.

3.2. Classification results

All 10 test images were correctly classified. The position of the cut from the image is important for the correct classification of the LOS. A rotation of the cut-out patch causes large changes in texture feature values as the GLCM uses a horizontal unit displacement vector, which is only sensitive when horizontal. The addition of GLCMs, which are calculated with other displacement vectors, into the classification procedure can improve the performance, but leads to a more complex solution.

4. CONCLUSION

The proposed method of traffic condition classification, based on texture features, correctly determines LOS for all the tested sequences. The devised configuration of the NN is adequate for the defined classification task. The low complexity of this solution is suitable for real-time evaluation of traffic conditions for implementation in intelligent transport systems. Its application for processing video from roadside cameras requires further research. Roadside cameras provide highly distorted views of the traffic lanes, especially when installed on low posts. The problem of occluding vehicles may also corrupt the interpretation of the image content. A perspective solution of the problem of roadside observation may include an expanded NN with additional input variables, which signal the position of the camera at the roadside.

References

1. Bi Song, Han Li-qun, Zhong Yi-xin, Wang Xiao-jie. 2011. "All-day traffic states recognition system without vehicle segmentation". *Journal of China Universities of Posts and Telecommunications* 18: 1-11.
2. Li Weia, Dai Hong-yingb. 2016. "Real-time road congestion detection based on image texture analysis". *Procedia Engineering* 137: 196-201.
3. Tracz Marian, Janusz Chodur. 2004. *Metoda obliczania przepustowości skrzyżowań z sygnalizacją świetlną*. Warsaw: GDDKiA. [In Polish: *The Method of Calculating the Capacity of Intersections with Traffic Lights*. Warsaw: GDDKiA].
4. Kalaitzakis Kostas, V. Kastrinaki, Michalis Zervakis. 2003. "A survey of video processing techniques for traffic applications". *Image and Vision Computing* 21: 359-381.
5. Xiying Lia, Yongye Shea, Donghua Luob, Zhi Yu. 2013. "A traffic state detection tool for freeway video surveillance system". In *13th COTA International Conference of Transportation Professionals. Procedia, Social and Behavioral Sciences*: 2453-2461.
6. Yu Qiaoa, Zhongke Shia. 2012. "Traffic parameters detection using edge and texture". *Procedia Engineering* 29: 3858-3862.
7. Haralick Robert M., K. Shanmugam, Its' Hak Dinstein. 1973. "Textural features for image classification". *IEEE Transaction on Systems, Man and Cybernetics* 3(6): 610-621.
8. Osowski Stanisław. 1996. *Sieci neuronowe w ujęci algorytmicznym*. Warsaw: WNT. [In Polish: *Neural networks in terms of algorithmic*. Warsaw: WNT].

9. Pamuła Teresa. 2011. "Road traffic parameters prediction in urban traffic management systems using neural networks". *Transport Problems* 6(3): 123-129.

Received 21.03.2016; accepted in revised form 28.07.2016



Scientific Journal of Silesian University of Technology. Series Transport is licensed under a Creative Commons Attribution 4.0 International License