Journal of Electronic Imaging

JElectronicImaging.org

Risk prediction algorithm based on image texture extraction using mobile vehicle road scanning system as support for autonomous driving

Nikola Slavkovic
Milan Bjelica

SPIE. IS&T imaging.org

Nikola Slavkovic, Milan Bjelica, "Risk prediction algorithm based on image texture extraction using mobile vehicle road scanning system as support for autonomous driving," *J. Electron. Imaging* **28**(3), 033034 (2019), doi: 10.1117/1.JEI.28.3.033034.

Risk prediction algorithm based on image texture extraction using mobile vehicle road scanning system as support for autonomous driving

Nikola Slavkovic[a,b,*] and Milan Bjelica[a]
[a]University of Belgrade, School of Electrical Engineering, Belgrade, Serbia
[b]ICT College Belgrade, Belgrade, Serbia

Abstract. We present an algorithm for risk prediction of road surface grip where skidding and sliding occur as main road surface problems. Prediction is done by defining a fine texture classification of the properties of road aggregate. In an experimental setup, data acquisition is performed with a supervised mobile vehicle scanning system, using a vehicle equipped with a camera and temperature sensor during movement along an arterial road. Image processing is performed by testing four texture feature extraction methods: Gabor filters, wavelet transform, gray level co-occurence matrix, and edge histogram descriptor, among which the Gabor transform shows the best results. The extraction of texture feature vectors follows by statistical algorithms for measuring feature vector similarity and reference vector selection, leading to image texture classification. The algorithm itself is upgraded by incorporating simultaneous surface temperature measurements in order to create and validate the final fine surface texture classification. The roads are classified and segmented into high-, medium-, and low-risk roads according to skid danger, enabling the creation of a map of high-risk zones. We validate our risk prediction algorithm by referring to crash rate data from the Road Traffic Safety Agency of Serbia database. This algorithm enables the location and mapping of high-risk zones and can be used as a support for autonomous driving and navigation. [DOI: 10.1117/1.JEI.28.3.033034]

Keywords: risk prediction mapping; vehicle road scanning system; fine image texture classification; temperature class validation; autonomous driving.

Paper 181071 received Jan. 10, 2019; accepted for publication Jun. 7, 2019; published online Jun. 25, 2019.

1 Introduction

Texture is one of the most relevant features for surface characterization. It can be defined as a repeated pattern of information or arrangement of the structure with regular and/or irregular intervals.[1] In a general sense, texture refers to surface characteristics described by size, shape, density, arrangement, and proportion of its elementary parts.[2] For the purpose of road assessment, road surface texture is the result of various types of asphalt and/or concrete mixtures.[3,4] Accordingly, commonly used methods for surface texture analysis, based on sophisticated contact or contactless surface scanning techniques such as two-dimensional (2-D) and three-dimensional (3-D) laser profilometer, line scan sensors, and other sensor devices, are recommended by related standards.[5,6]

However, the texture of a particular surface can also be analyzed and evaluated by processing an image captured from the surface or from a part of the surface. Texture feature extraction applied to the captured images is an important step in various image processing applications, e.g., remote sensing and diagnostics, biomedical imaging, image classification and retrieval, and industrial quality and production control.[7] Image processing techniques can be used for image texture characterization and classification. Numerous methods are used for image texture feature extraction, including transform-based, statistics-based, structure-based, or model-based. These methods are frequently used in research work for pattern recognition techniques, image retrieval, and image classification.[2]

As the main attribute of a surface, texture is commonly referred to road safety issues. It is used as a descriptor for the overall evaluation and classification[8] of road conditions. Road surface texture differs according to the various asphalt and/or concrete mixtures that are used as the top layer of the road surface, traffic density, surface age, weather conditions, etc. As a surface descriptor, its evaluation is very important for predicting and estimating the values of other relevant safety parameters such as skid resistance and friction coefficients, tire rolling resistance, and quantification values of the tire-road interaction.[9,10]

In addition to imaging and computer vision diagnostics, other relevant sensing techniques are used for the estimation and assessment of road conditions, e.g., air and surface temperature, global positioning system (GPS) locating and speed estimation, vibration, humidity, and luminescence.

Temperature measurement is usually performed at the same time as other sensing procedures.

Road assessment procedures are usually based on visual inspections provided by the observers (operators) who visually examine imaging data. Visually recognized surface changes are archived. The operators note the GPS positions on which they have visually noticed the changes. Subjective perception, based on the observer's visual inspection, is usually used for the classification of particular road segments. This is imprecise and time-consuming and largely depends on the observer's capabilities.[3,4] Therefore, the development

*Address all correspondence to Nikola Slavkovic, E-mail: nikola.slavkovic@ict.edu.rs

of automated road inspection methods is of high importance. They are significant for the faster and more precise classification of the overall road assessment and risk mapping process.

Road surface data acquisition can be provided using sophisticated equipment mounted on special, sophisticated vehicles moving along the road infrastructure. In this paper, an algorithm for road risk prediction, concerning skid danger, is presented. It is based on the minute classification of road surface texture using a mobile vehicle scanning system, which gathers images from a camera mounted on a moving vehicle. Surface temperature acquisition was also measured.

The benefit of the proposed algorithm is that it can be used in situations when there is a lot of disturbance caused by the movement of the camera along very uneven surfaces. An additional advantage of this method in comparison with others is that when there is a lot of disturbance, it can classify changes in the surface texture of very fine aggregate shapes and not only obvious changes such as cracks, potholes, and patches. The disturbances identified are caused by the speed of the vehicle relative to the recorded area, surface illumination (shadows), and/or by the vibrations transmitted to the vehicle and then to the camera during driving.

The data acquisition was carried out in daylight, at a constant vehicle speed. Frames were extracted from the video. Following this, regions of interest were cropped from each frame. These regions are processed with methods for texture feature extraction. Output feature vectors are compared with different similarity measures. The texture surface was identified into several classes by combining the measurement results. Surface temperature measurements obtained from the same moving vehicle equipped with a camera can act as a texture category validator, e.g., as a misclassifier for texture categories that do not represent the actual surface texture.

2 Related Research

This study presents the use of a mobile camera as a very useful technique for analyzing road surface texture when there is not necessarily obvious visible damage such as cracks, potholes, and patches, through analyzing the surface textures of the shapes of the aggregate. Road surface classification, especially supported by temperature validation, has not been explored in this way before.

The available literature on the automated detection of road surface distresses by the visual inspection of images and video data defines two concepts in surface image processing: single imaging and video imaging.[11]

Single imaging uses a static camera and can identify fine texture. The static image is analyzed more frequently than video imaging, because the changes in texture are more easily distinguishable and analysis of these features can determine the quality of the surface. On the other hand, the most up-to-date research literature suggests that video imaging analysis is only able to detect obvious surface distresses such as potholes, cracks, and patches.

A camera-equipped vehicle is a relatively new concept in road inspection. Vehicles are usually equipped with a line-scan camera and laser illumination, and a visual-based acquisition system is used to store the digital images that will be processed further to identify road distresses.[3] However, the methods based on 2-D video imaging have been focused mostly on crack, pothole, and patch detection. A method for the detection and assessment of potholes, cracks, and patches of Indian highways by collecting data with video cameras from which images were extracted was analyzed using various image processing techniques supported by heuristically derived decision logic.[12] This resulted in the classification of four different frame categories: frames with potholes, frames with cracks, frames with patches, and frames without any critical distress. The use of computer-vision techniques in the detection and analysis of cracks in concrete surfaces were captured and analyzed using high-quality images to build crack classification in bridge deck concrete material.[13] Noncrack features that are usually mistakenly reported as crack features, such as joints, sealed cracks, and white paint, are specifically detected and masked, as shown in Ref. 14. Crack detection was also performed by the use of low-cost screening systems, because they can detect significant texture changes in the crack area.[15,16] Pothole detection was performed using image segmenting into defective and nondefective regions using histogram shape-based thresholding.[17] The video images from parking cameras mounted on vehicles were used to develop methods for patch detection.[18] Dash cameras (also called black-box cameras) were used to develop methods for detecting defects such as potholes, patching, shoving, raveling, and all types of cracking, by calculating the entropy of each gray-scale frame taken from the videos.[19] The same types of cameras were used for the detection of potholes.[20]

Texture image processing is frequently based on the various methods used for texture feature extraction and enables assessment of the feature vectors. Analysis of the similarities between feature vectors is often performed, as in our study, using statistical data to characterize image textures using simplified descriptors.[21]

In this study, four methods were examined and applied for processing dynamically captured video frames relating to image texture feature extraction. The methods we used were the Gabor transform, often referred to as the Gabor filter bank or Gabor function;[22,23] the wavelet transform; the co-occurrence matrix; and the edge histogram descriptor. Depending on the stochastic nature of the road surfaces, these feature extraction approaches should provide adequate preprocessing for road surface texture classification. Methods based on multiresolution/multichannel analysis such as the Gabor transform and wavelet transform have been introduced as pyramid structured transforms for texture analysis by Mallat,[24] and later on, by Ma and Manjunath.[21] Wavelet tools were later developed to obtain more invariant image processing, especially for the purpose of texture detection, classification, and segmentation.[25]

Researchers fully describe the use of Gabor filters with various scale,[26] orientation, and kernel size,[23,27] for the purpose of texture segmentation and characterization.[28]

The Gabor filter bank or Gabor transform, applied to a texture image, is a very useful tool as it has optimal localization properties in both the spatial and frequency domains. Gabor filters provide the means for better spatial localization, but their usefulness is limited in practice because there is usually no single-filter resolution where one can localize a spatial structure in natural textures. These methods involve transforming original images using filters and calculating the energy of the transformed images. We use a mask a so-called "convolution kernel"[27] that represents the filter. Put simply,

this is a 2-D array (since 2-D images are involved here) of pixels, in which each pixel is assigned a value. In our study, the 2-D form of Gabor wavelet was used, which consists of a planer sinusoid multiplied by a 2-D Gaussian.[29] The Gabor wavelet used in 2-D extracts local features from an image, as described in Ref. 30.

The gray level co-occurrence texture matrix (GLCM) is a common method for texture feature extraction based on interpixel distance and orientation.[7] The GLCM method is a way of extracting second-order (involving two-pixel comparison analysis) statistical texture features. This is a very useful method for particular texture feature extraction.[28,31] The first step is to determine the co-occurring probabilities of all pairwise combinations of quantified gray levels in the fixed-size spatial window, given two parameters: interpixel distance and orientation. The algorithm implemented in this work is based on the description given by several authors.[32–34] The exact equations for frequent types of features, i.e., entropy, energy, contrast, correlation, and local homogeneity extracted from the GLCMs, are presented in Ref. 34.

The edge histogram texture descriptor resembles the color layout descriptor in its principle of capturing the spatial distribution of edges.[35] The distribution of edges is a good texture signature that is useful for image processing even when the underlying texture is not homogeneous. The edge histogram method using five filters is proposed in the MPEG-7 standard.[35]

The implementation of wavelet transform texture characterization was introduced and explained in detail.[32,33] The implementation was performed by obtaining the variances of the high-frequency sub-bands of the wavelet transform of each grid region.

Temperature is used as an important parameter in road surface quality assessment, as well as for the prediction of possible defects such as cracks and holes, and other parameters such as skid resistance, slipperiness levels, and friction coefficient.[36] The skid resistance of all surfaces consisting of various asphalt and concrete mixtures decreases as the surface temperature increases.[37] Therefore, in road assessment and categorization programs, surface temperature is usually measured from moving vehicles, most often simultaneously with the use of other sensors or devices for the diagnostics of the road infrastructure.[3]

There is little published scientific investigation into road risk prediction. Most studies use crash rates as the main input data for determining risk levels. An interesting algorithm is proposed in a paper[38] based on SVM algorithm and image segmentation processing technology, where the authors propose a method of video image processing technology for road surface state recognition (e.g., if the surface is wet or dry, or if there is snow, or ice on the road). However, there is a lack of investigation into risk prediction.

No studies have identified and described fine road surface texture changes obtained by a moving vehicle camera; they have only described obvious road surface changes, such as cracks, potholes, patches, or the simple presence of water, snow, ice, or other materials covering road surfaces.

3 Methodology

The research in this study was performed as shown in Fig. 1.

3.1 *Experimental Setup*

For the purpose of experimental analysis, a specific vehicle was equipped with a video camera, to enable video data acquisition. The monitoring and image capturing of the road surface was performed along a section of an arterial road with a camera recording video frames mounted on a vehicle moving, at a speed of 60 km/h. The camera was mounted on the back of the vehicle, above the offside wheel, oriented downward, in order to best capture the actual road surface conditions. The most discrepancies of the road surface texture are present in the wheel paths, not in the area between wheel paths.

The temperature of the road surface was also analyzed to validate the image texture processing algorithm. This was done by mounting a temperature sensor under the front bumper of the vehicle so that surface temperature data could be collected at the same time as recording the video frames. The measured values of temperature were averaged at a distance of 50 m, which is a common procedure in road assessment programs.[3,4] Each frame was referenced to GPS coordinates.

Each average temperature record was also referenced to GPS coordinates. This provided information about the precise position of the recorded frame or area of the measured temperature. There is, therefore, a correlation between the spatial distribution of frames and temperature measurements along the entire section of road used.

3.2 *Texture Characterization*

3.2.1 *Texture feature extraction*

In our research, for image texture analysis, we used the Gabor and wavelet transforms (as a wavelet-based and co-occurence matrix and edge histograms) for statistically based feature extraction methods. This enabled the computation of the feature vectors obtained from the captured frames.

Gabor function can be described as an image texture discriminator and is sensitive to different frequencies and scale information. These capabilities were of interest to us for our research. Gabor transform is a procedure that uses specific filters (functions) to apply to images in a similar way to conventional filters. A Gabor filter can be viewed as a sinusoidal plane of a particular frequency and orientation, modulated by a Gaussian envelope. A 2-D Gabor function $g(x, y)$ and its Fourier transform $G(u, v)$ should be written as described in Ref. 21:

$$g(x,y) = \left(\frac{1}{2\pi\sigma_x\sigma_y}\right)\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)+2\pi jWx\right], \quad (1)$$

where $j=\sqrt{-1}$ and W is the frequency of the modulated sinusoid

$$G(u,v) = \exp\left\{-\frac{1}{2}\left[\frac{(u-W)^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right]\right\}, \quad (2)$$

where $\sigma_u = 1/2\pi\sigma_x$ and $\sigma_v = 1/2\pi\sigma_y$.

If we assume that $g(x, y)$ is the mother wavelet, then Gabor function is formed as follows (a self-similar filter dictionary can be obtained by various transforms, i.e., dilatations and rotations of $g(x, y)$ through the generating function):

$$g_{mn}(x,y) = a^{-m}G(x',y'), \quad (3)$$

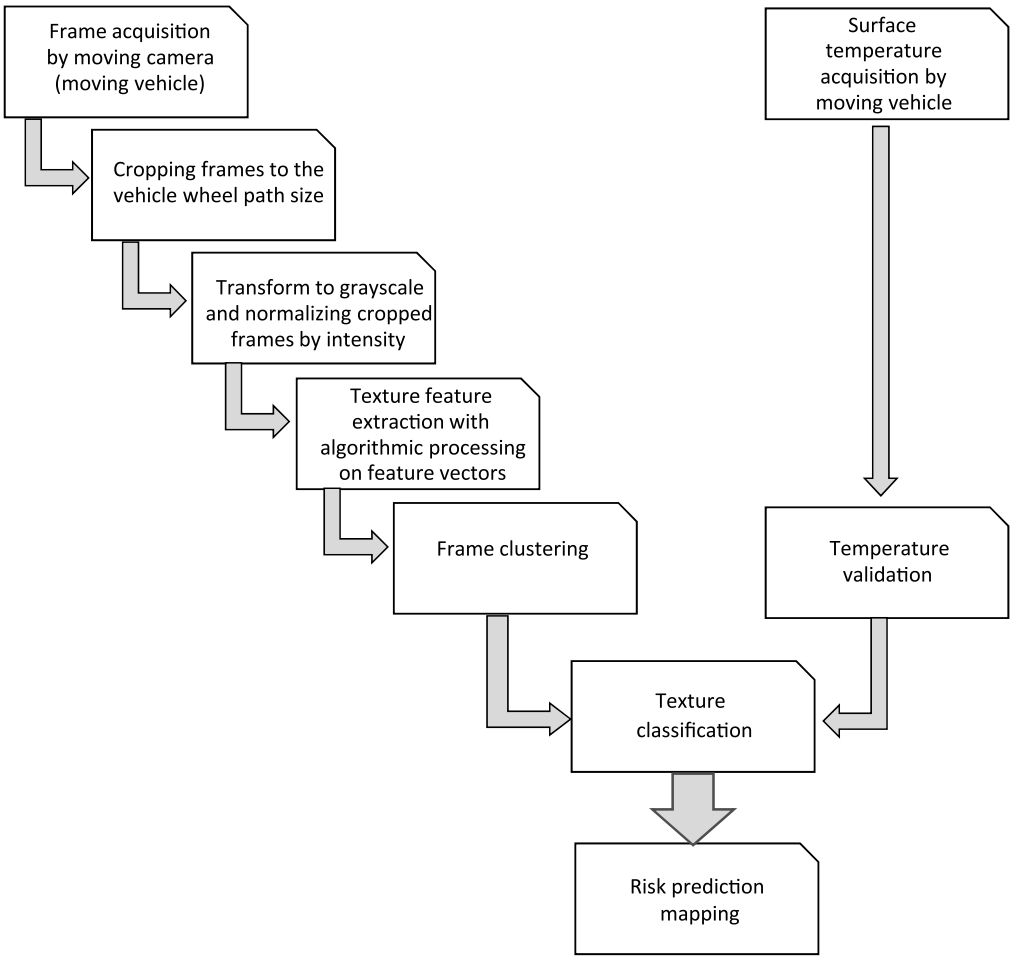

Fig. 1 Theoretical foundation illustration of algorithm for road risk mapping based on mobile vehicle data acquisition.

where $a > 1$; $m, n =$ integer and,

$$x' = a^{-m}(x \cos\theta + y \sin\theta), \tag{4}$$

$$y' = a^{-m}(-x \sin\theta + y \cos\theta), \tag{5}$$

where $\theta = 2\pi/K$, and K is the total number of orientations, and $a = (U_h/U_l)^{-\frac{1}{S-1}}$. The values of U_l and U_h represent the respective lower and upper center frequencies of interest.

Here S is the user-defined number of scales of analysis in the multiscale pyramid image decomposition. This results in the following equations for computing the filter parameters σ_u and σ_v, which enable the values of σ_x and σ_y to be computed:

$$\sigma_u = \frac{(a-1)U_h}{(a+1)\sqrt{2\ln 2}}, \tag{6}$$

$$\sigma_v = \tan\left(\frac{\pi}{2k}\right)\left[U_h - 2\ln\left(\frac{\sigma_u^2}{U_h}\right)\right]\left(2\ln 2 - \frac{2\ln 2^2\sigma_u^2}{U_h^2}\right)^{-\frac{1}{2}}, \tag{7}$$

where $W = a^{-m}U_l$, and $m = 0,1,2,\ldots,S-1$. Eliminating the filter response sensitivity to absolute intensity values, the real components of the 2-D Gabor filters are biased by adding a constant to make them zero mean.

Let $I(x,y)$ be the term of a particular image. Then its Gabor transform for representing a feature is given by

$$W_{mn}(x,y) = \int I(x_1,y_1)g^*_{mn}[(x-x_1),(y-y_1)]\mathrm{d}x_1\,\mathrm{d}y_1. \tag{8}$$

Operator * indicates the complex conjugate, and x_1, y_1 are the parameters of the convolution kernel. In this case, we assumed that local texture regions are spatially homogeneous, since the road surface texture tends to be homogeneous. By moving the convolution kernel by sliding a window over an image pixel by pixel, a convolution operation is conducted, producing a filtered image with new pixel level intensity responses as the new output. The mean μ_{mn} and standard deviation σ_{mn} of the magnitude of the transform coefficients are used to represent the region of interest to determine texture classification:

$$\mu_{mn} = \iint |W_{mn}(x,y)|\mathrm{d}x\,\mathrm{d}y, \tag{9}$$

and

$$\sigma_{mn} = \sqrt{\iint (|W_{mn}(x,y)| - \mu_{mn})^2\mathrm{d}x\,\mathrm{d}y}. \tag{10}$$

The feature vector can be extracted using the mean μ_{mn} and standard deviation σ_{mn} as feature components. In our study, the Gabor transform was based on a set of five scales, $S = \{0, 1, 2, 3, 4\}$, and six orientations, $K = \{0, 1, 2, 3, 4, 5\}$, (in our case, six orientations refer to the angles of 0 deg, 30 deg, 60 deg, 90 deg, 120 deg, and 150 deg, respectively), which create a bank of 30 filters. In addition, a convolution kernel of a size 21×21 pixels is applied. When all filtering parameters are taken into account, the feature vector is described by

$$X = (\mu_{0\_0}, \sigma_{0\_0}, \mu_{0\_1}, \sigma_{0\_1}, \ldots \mu_{s\_k}\sigma_{s\_k}, \ldots \mu_{4\_5}\sigma_{4\_5}). \quad (11)$$

The feature vector dimension is determined by five scales (S), six orientations (K), and two feature components (μ_{mn}, σ_{mn}), which makes 60 coordinates.

In this paper, the radial GLCM was used for feature extraction. Various types of features were extracted from 256 grayscale frames: entropy, energy, contrast, correlation, and local homogeneity (inverse difference moment). In this case, we used 24 directions chosen from the referent pixel, with each orientation representing a different pixel distance. Twenty-four different GLCM were formed, from which the above-mentioned features were extrapolated.

This means that the resulting feature vector was extracted from 24 GLC matrices, including four types of features. The feature vector has $24 \times 4 = 96$ coordinates.

For the edge histogram, a scale invariant nonhomogeneous texture descriptor, five filters were applied. It identifies the spatial distribution of five types of edges in five directions: vertical, horizontal, 45 deg, 135 deg, and nondirectional. The feature extraction was based on dividing the image into 16 (4×4) blocks, then applying five filters to each region in the five directions described. A 5-bin histogram for each block was generated.

The edge histogram was formulated by interlinking the results of the region edge histograms. The corresponding extracted feature vector has $16 \times 5 = 80$ coordinates.

The wavelet transform was applied based on Haar wavelets. Discrete wavelet features of the image were extracted and analyzed as presented in Ref. 24. The implementation was performed by obtaining the variances of the high-frequency sub-bands of the wavelet transform of each grid region. The frame was divided to $4 \times 4 = 16$ grid regions, and the four-level Haar wavelet transform was applied to obtain the variances of the high-frequency sub-band (in our case 12 sub-bands), of each region. The resulting feature vector has $4 \times 4 \times 12 = 192$ coordinates.

3.2.2 *Measurement of feature vector similarities*

For the purposes of texture characterization, it was necessary to apply certain similarity measurements of the feature vectors obtained by the methods specified by the texture feature extraction. The feature vectors obtained were compared further. Three measurements of similarity between the feature vectors were calculated. The measurements applied were the Euclidean distance (d), the correlation (R), and the normalized mean difference (D) between the average and median values. The combination of these three measures is giving us the more precise classification result since we are analyzing fine texture, and each measurement describes the feature vectors' similarity, but in a different way. The Euclidean distance gives a cumulative measurement of the vector components' difference. The correlation describes the similarity between the patterns that create the components of the individual feature vectors. The normalized mean difference describes the distribution of the component values for differences between mean and median values. On the basis of these measurements, further classification of the texture could be carried out using statistical algorithms.

To determine the similarity between feature vectors, it was necessary to define a reference value, that is, a reference vector in relation to which the similarity of all other vector features could be measured. The method of determining the reference vector is given at the end of this section.

Similarity measurements of feature vectors that are applied in this paper are:

1. *Euclidean distance (d)*

Distance between the feature vector of each frame and the reference vector. This measurement indicates the proximity of points in the hyper space, but it cannot identify the similarities in the structure of the components of the vector.

Let X_j be the feature vector of the current frame:

$$X_j = (X_{j1} \ldots X_{jk} \ldots X_{jN}),$$

where $j = 1, \ldots, F$, and X_{jk} represents the k'th coordinate of the feature vector, N is the number of the feature vector coordinates, and F represents the total number of analyzed frames.

The Euclidean distance d between the feature vectors can be expressed as follows:

$$d_{j,\mathrm{REF}} = \sqrt{\sum_{i=1}^{N} (X_{ji} - X_{\mathrm{REF},i})^2}, \quad (12)$$

where $j = 1, \ldots, F$, and X_{ji} is the coordinate of the feature vector of the current frame, and $X_{\mathrm{REF},i}$ is the coordinate of the reference feature vector. For the purpose of experimental analysis, a reference vector was estimated. The feature vector to which the highest Euclidean distance corresponded was selected as a reference. The measurements were implemented for further analysis according to the reference.

2. *Correlation (R)*

As a measurement of correlation, the Pearson correlation coefficient was calculated, which adequately described the similarity of the reference and feature vectors of other vectors to their component structure.

Pearson correlation coefficient R was calculated using the following equation:

$$R_{j,\mathrm{REF}} = abs\left\{ \frac{\sum_i X_{ji} X_{\mathrm{REF}} - \frac{(\sum X_{ji})(X_{\mathrm{REF}})}{N}}{\sqrt{\left[\sum_i X_{ji}^2 - \frac{(\sum_i X_{ji})^2}{N}\right]\left[\sum X_{\mathrm{REF}}^2 - \frac{(X_{\mathrm{REF}})^2}{N}\right]}} \right\}, \quad (13)$$

where $i = 1, \ldots, N$ and $j = 1, \ldots, R$.

3. *Normalized mean difference (D)*

Comparing the differences between the mean and the median, given in Eq. (14). This enabled the components of the feature vector to be distinguished by the distribution of the feature vector components.

Normalized mean values difference (D) is calculated as follows:

$$D_j = \frac{abs[\text{mean}(X_j) - \text{median}(X_j)]}{[\text{median}(X_j)]}, \tag{14}$$

$$D_{j,\text{REF}} = D_j - D_{\text{REF}}, \tag{15}$$

where $i = 1, \ldots, N$ and $j = 1, \ldots, R$ and the mean and median represent the standard statistical measurements of the central tendency for the set of data. In our case, the set of feature vectors related to all the frames was analyzed.

4. *Defining reference feature vector*

Defining the reference feature vector involves the selection of one feature vector in relation to which similarity is measured. One way of choosing this is to find a feature vector that is the furthest away from all other vectors, i.e., the vector that has the largest Euclidean distance from all other vectors. It is presented as follows:

$$S_j = \sum_k d_{j,k} \quad k = 1 \ldots R,$$

$$X_{\text{REF}} = (X_j | S_j = \max\{S_1, \ldots, S_R\}).$$

3.2.3 *Classification of feature vectors*

After extracting the feature vectors of texture, all three measurements (d, R, and D) were calculated for each individual feature vector relative to the reference vector. The values obtained were combined and presented in 3-D space. Based on mean-shift clustering,[39] numerical values of thresholds between certain clusters of points can be obtained, representing candidates for texture classifications. Certain areas of surfaces can form cluster points that further represent different texture classes.

3.3 *Temperature data acquisition*

As mentioned earlier, the measurement of road surface temperature from a moving vehicle is a common technique,[3,4] for the purpose of analyzing the effect of temperature on other parameters in road assessment programs.

The surface temperature was measured remotely, using an infrared sensor placed under the vehicle bumper.

The temperature was measured every 50 m and an average calculated. Measuring this way complies with road safety recommendations.[40]

Significant fluctuations in the temperature of asphalt or concrete do not usually occur during daylight, in constant sunny weather conditions (i.e., no rain, snow or ice on the road, and no obvious high wind speeds near the road surface).

In our study, surface temperature measurements were performed to validate texture classes.

Significant temperature changes were observed, where the road was in shade. This provided information for class validation and for final surface texture classification.

4 Results and Discussion

The moving camera produced 11,300 frames taken in daylight conditions. Each frame was cropped to a region of interest, i.e., the area relating to the width of the rear offside wheel. Furthermore, we converted each color frame into a gray scale and normalized the cropped frames by pixel intensity. This minimized some of the irrelevant properties of the surface aggregates.

The texture feature extraction was performed using the methods mentioned above: Gabor transform, co-occurrence transform, edge histogram descriptor, and the Haar wavelet transform. The 3-D coordinate system is defined by three metrics, d, R, and D. 3-D frame feature distributions for all of the methods mentioned are presented in Fig. 2.

The texture feature vectors of the captured frames obtained by the Gabor transform were statistically processed by the aforementioned algorithms. The texture feature vectors of the corresponding frames are represented by the points, spatially distributed in the 3-D diagram, presented in Fig. 2(a). The spatial points distribution that represents the texture feature vectors of the corresponding frames, extracted by the other three methods, are shown in Figs. 2(b)–2(d), respectively.

Application of the other three feature extraction methods such as co-occurrence matrix, edge histogram descriptor, and wavelet transform is not suitable for detecting separate clear clusters of road surface texture images, obtained by moving camera. Since the Gabor transform showed different variations of the points found in 3-D space, it was selected in the proposed methodology for differentiation between the frames, and further on for clusterization, Fig. 3.

Since the already known methodologies for the automatic process of clusterization cannot effectively lead to road classification for our needs, we applied a complex procedure of automatic clusterization and, further on, classification, which were implemented through two phases.

In phase 1, we modified mean shift clustering algorithm,[39] where the bar is used as kernel. The whole surface of the graphics in Fig. 3 was fragmented into vertical bars, widths of 0.01 values of correlation (R). Within each bar, the total Euclidean distance between all points within that bar was measured. Thus the histogram of the values of these sums of Euclidean distances was formed. The minimal value in the histogram determines the position of line 1 in Fig. 3. Thus we separated a set of dispersed points on the left side of line 1 as a separate cluster, further on class 4. All other points are positioned on the right side of line 1 [Fig. 4(a)].

In phase 2, we observed only the points on the right side of line 1 and collected the total number of points within the vertical bars. The surface of the graphics, on the right side of line 1, was fragmented into vertical bars, widths of 0.005 values of correlation (R).

Thus a new histogram is formed, shown in Fig. 4(b).

Detection of the interval with the highest density of points within which are the boundaries of the clusters forms line 2 and line 3, in Fig. 3.

The threshold for detection of interval of interest is defined as half the maximum value of histogram (max /2).

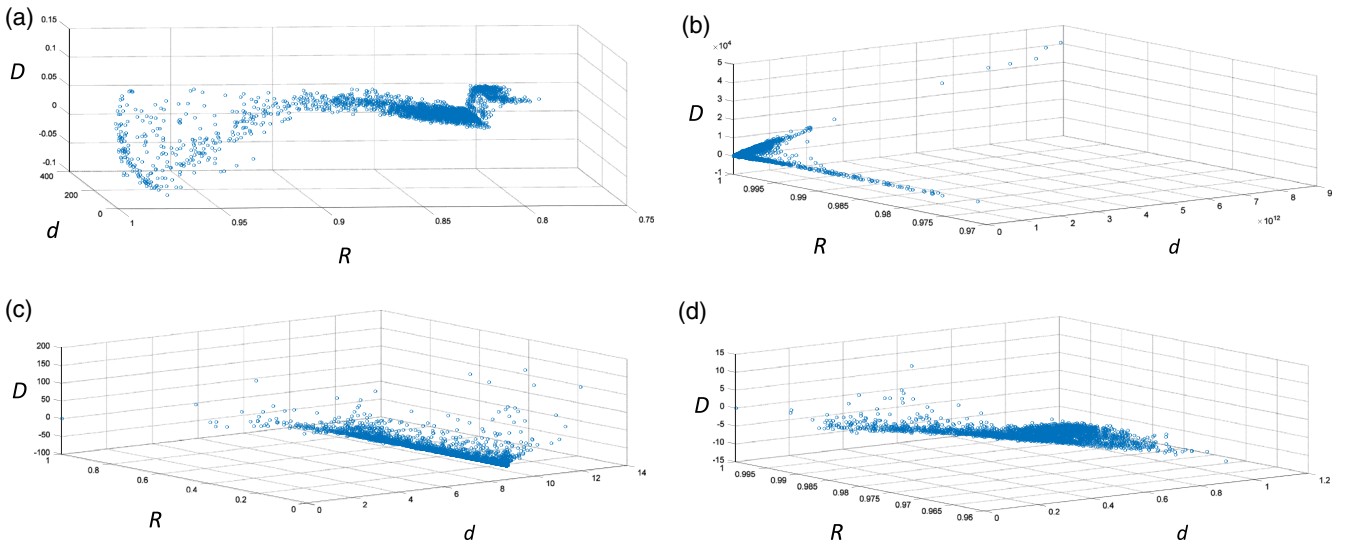

Fig. 2 3-D presentation of frame distributions based on texture feature vector extraction according to the texture analysis using: (a) Gabor transform, (b) co-occurrence transform, (c) edge histogram, and (d) DWT transform.

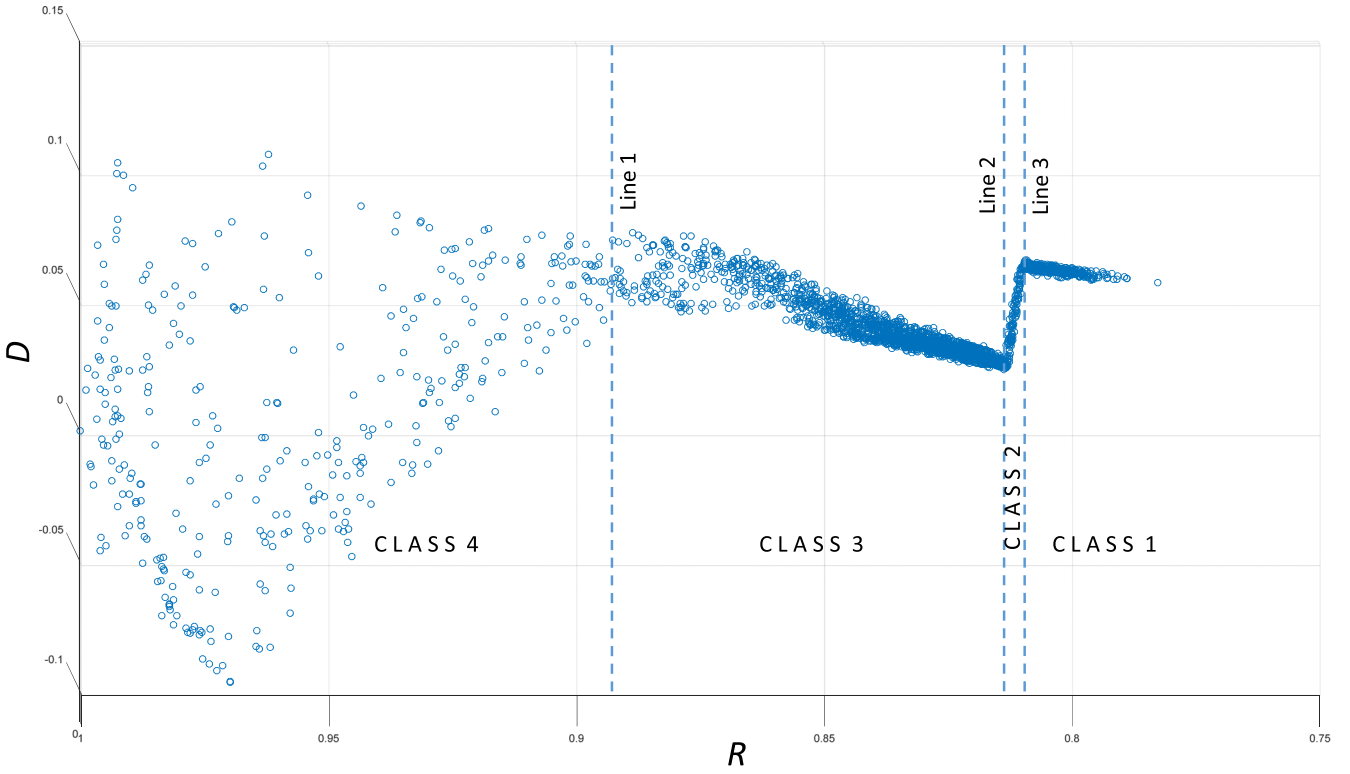

Fig. 3 The 2-D frame classification representation according to the Gabor transform-based texture feature extraction with defined class areas separated by threshold values (vertical dash lines).

The position of this threshold is shown in Fig. 4(b). All bars whose number of points exceed the defined threshold are the interval of interest. Within this correlation interval, the actual position of lines 2 and 3 are determined. Line 2 is determined by passing through a mean value of a minimum three points in the detected correlation interval. Line 3 is determined by passing through the mean value of 3 of the maximum points in the correlation interval.

Threshold values of correlation (R) are as follows: first class up to 0.8092, second class up to 0.8136, and third class up to 0.8925. Fourth class is over 0.8925.

We tested an automated system, which enabled us to define clear boundaries between group of points that correspond to the defined classes of risk, and by doing so, we identified four separate clusters. These clusters were candidates for image texture classes. For the purpose of surface

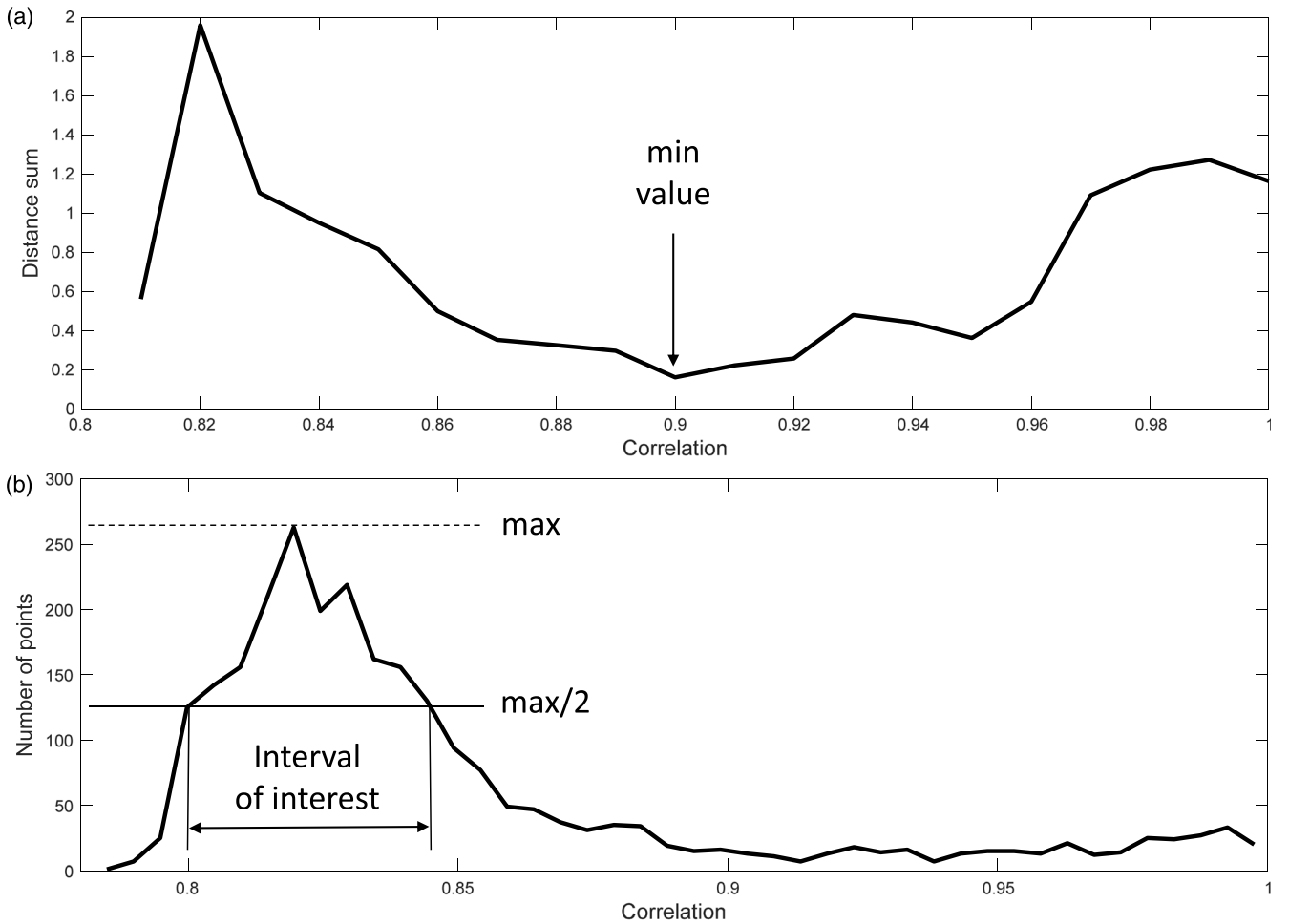

Fig. 4 Automatization of threshold setting for cluster recognition from Fig. 3 data: (a) the histogram determining the position of the line 1 and (b) the histogram determining the position of line 2 and line 3.

texture classification, we had to integrate image texture classification into surface texture classification. This enabled us to distinguish four classes of image texture in Fig. 3

In Fig. 5, four-class frame classification results based on the Gabor transform is illustrated. Here obvious diversity of frames belonging to class 4 is seen.

Samples gathered as the choice of representative frames that represents each class, including class 4, after final classification, are presented in Fig. 6.

It can clearly be seen in Fig. 6 that the images from class 4 show road surfaces obscured by static shade, as from trees. Analysis of the image texture reveals that class 4 does not

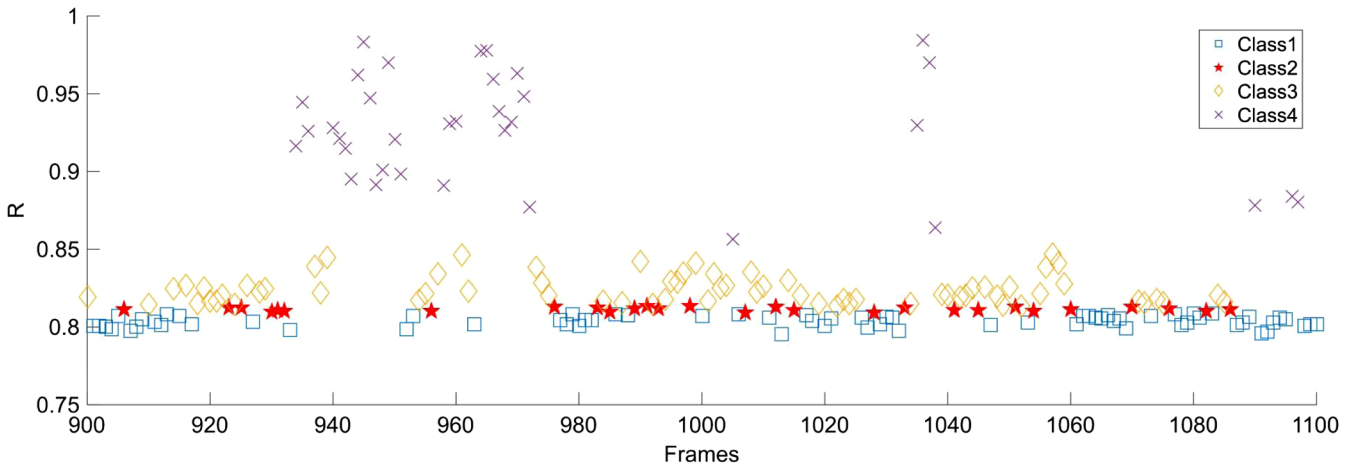

Fig. 5 Four-class texture classification, per frame, based on the Gabor transform.

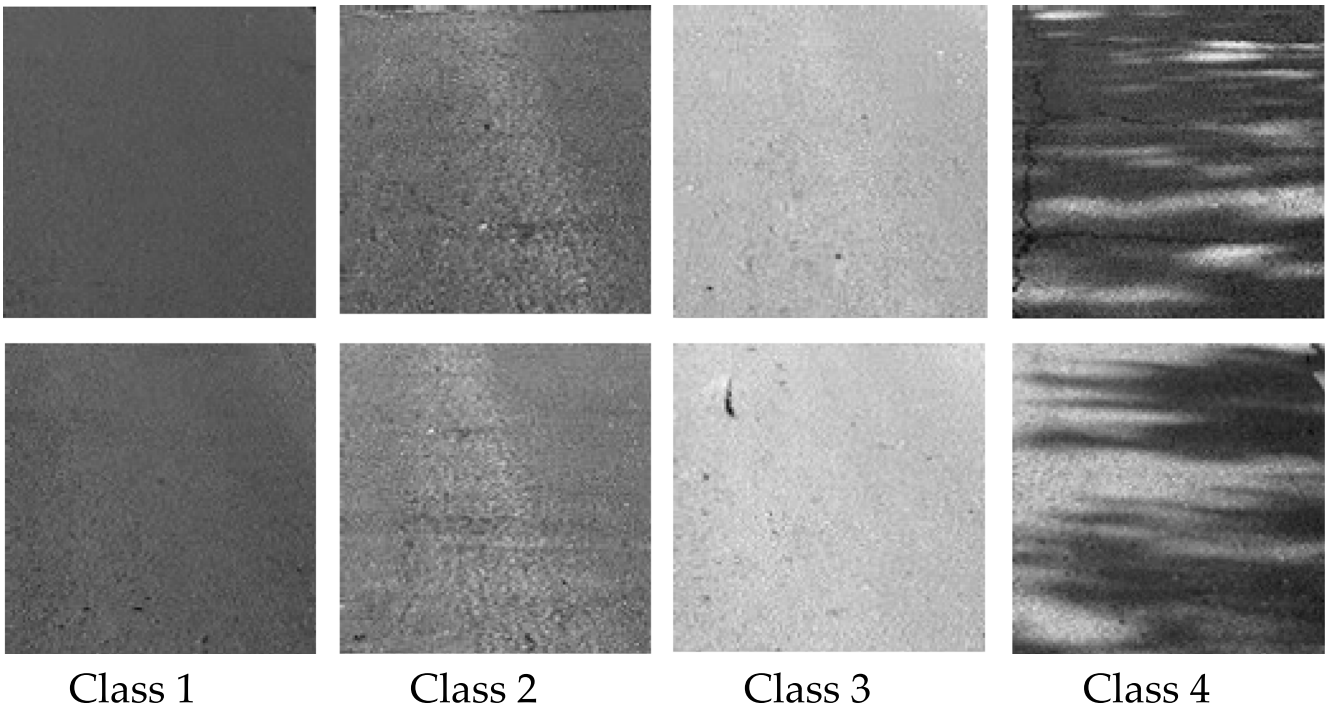

Fig. 6 Representative samples of classified frames. Classes 1 to 3 are validated and represent road surface texture. Class 4 is a misclassification and can be rejected.

represent the actual road texture surface. This is confirmed in the measurement of the surface temperature taken at the same time from the moving vehicle.

Surface temperature measurements related to image frames in class 4, reveal feature vectors distorted by static shadows from, e.g., tree branches and power pylons. Therefore, the second part of the experimental analysis of the road surface is the comparison between the averaged value of R, calculated on the sequence of the frames, and the collected temperature values.

The temperature of the road surface was measured and averaged at an interval of 50 m from the moving vehicle.[3,4] The time interval within which the temperature was averaged was the same time interval within which the sequence of frames was collected. The mean value of R in this group of frames was compared to the mean value of the temperature within the same time interval.

We introduced a temperature measurement for the purpose of improving the surface texture classification. Since both the temperature and video data are referenced by GPS coordinates, it was possible to spatially compare these two data sets. This made it possible to compare the correlation between the group of frames with the temperature averaged at 50 m.

A comparison of the average value of R and the average temperature is shown in Fig. 7. The total correlation coefficient value between these two curves is $R = 0.5651$, which indicates a medium degree of correlation between the curves. Nevertheless, it is obvious that significant peaks of low-temperature values and class correlation values correlate in counter-phase, which correspond to the dissipation of road surface temperature and image texture changes.

We excluded class 4, shown in Fig. 6. Peaks in Fig. 7 confirm that class 4 of image texture obviously does not represent the actual surface texture, since these are the areas of static shades, caused mostly by tree brunches, electricity power pylons, etc. We must emphasize that the changes in surface texture do not lead to temperature changes.[36] Our proposed methodology for using image processing from a moving camera to characterize road surface texture can, therefore, be confirmed as correct.

This means that road surface can be actually classified into three classes of road by its surface texture.

5 Road Risk Prediction Using Fine Texture Classification

It is well-established that vehicle road accidents and crashes occur more frequently when humidity is high and road surfaces are slippery, e.g., when there is rain or snow on the roads. We have used our new algorithm for classifying roads according to fine surface texture, to establish different road risk categories for roads in Serbia.

The texture classes differ according to their uniformity. If the texture is uniform and has uniform granularity distribution, this is positive and means it is safe for driving and is not slick in itself. The road friction is satisfactory enough and the road will not be very slippery.

If the texture of the road is not uniform, then it means that it is defective, too polished, or too slick in itself. Such sections of road are extremely dangerous, irrespective of the effects of weather or other incidental phenomena.

In our classification, as shown in Fig. 3, it is obvious that class 1 and class 2 are fairly uniform in texture, and that class 3 and class 4 are not.

(Class 4 looks very dispersed because of the effects of the static shadows, and as we said, it represents a misclassification.)

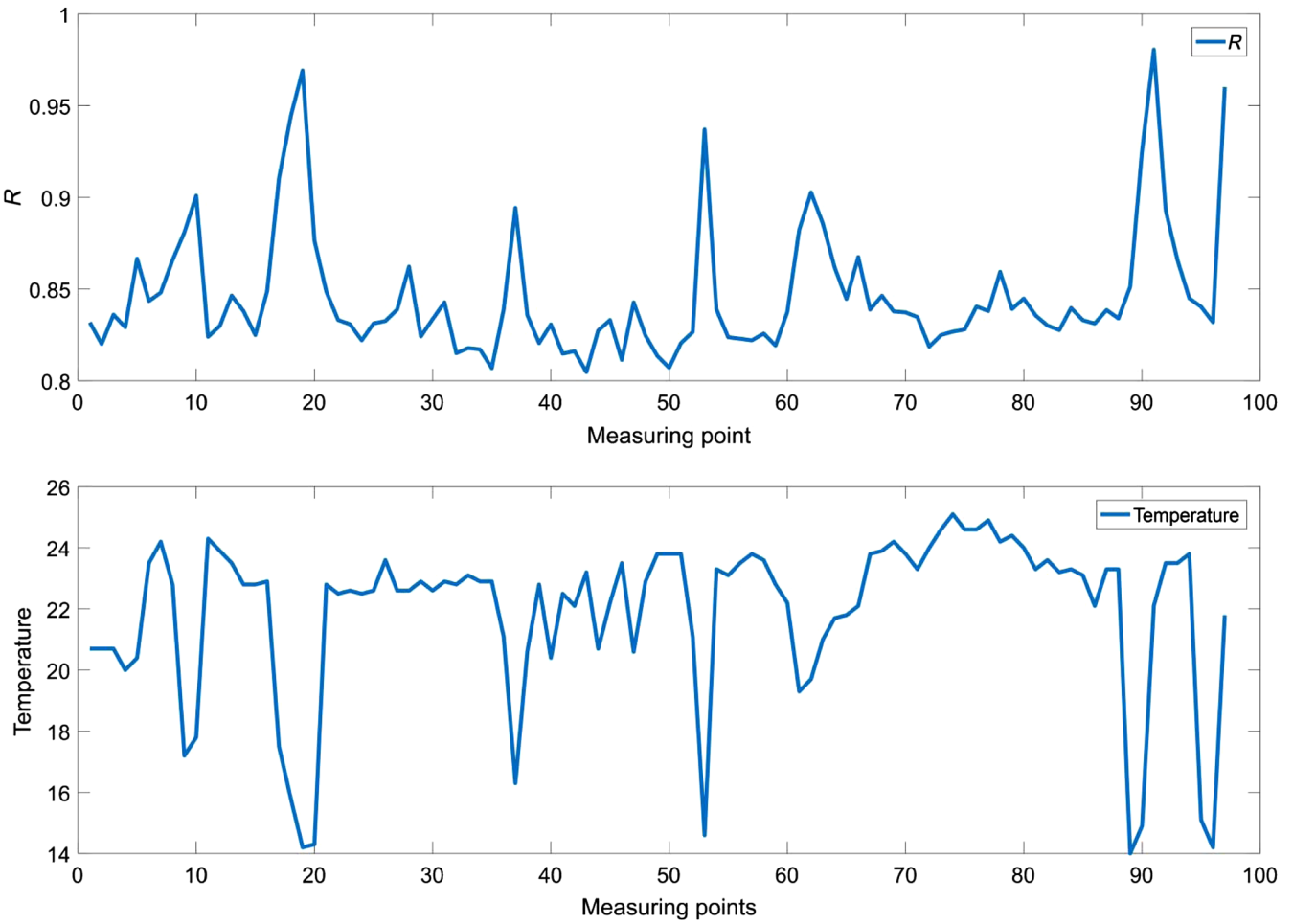

Fig. 7 Comparison of the average value of *R* and average temperature.

Class 1 shows the most uniformity. This means that the granularity of its texture is well preserved. It represents a nondangerous road surface and does not tend to be slippery.

Class 2 reveals so-called "bleeding," i.e., a loss of texture in certain places, which means that it can be dangerous in sections, but not all over. Class 2 is, therefore, a borderline surface category which can be slick in some parts. For such road surfaces, we propose further analysis using neural networks for detecting the exact location of dangerous spots. Another option would be using the Internet of Vehicles (IoV) network. This category of surface represented only 7% of the observed road.

Class 3 is the most polished and slick surface and is more critical in terms of road risk than class 2. Class 3 represents a dangerous, potentially slippery, and slick road.

According to these classifications, a map of the area can be highlighted to illustrate three classes of roads [Fig. 8(a)]:

1. High-risk, potentially slippery when wet, especially dangerous roads (class 3).
2. Medium-risk, borderline (class 2).
3. Low-risk, nonslippery roads (class 1).

In parts of the road where high-risk surfaces are detected, drivers should continue to take precautions and prepare for slick roads and very dangerous driving conditions.

We compared our road mapping—based on road risk prediction using fine texture classification—with existing traffic accident statistics from the same section of road. We used the "integrated database on security facilities of the Serbian roads" for years 2015, 2016, and 2017, provided by the Road Traffic Safety Agency of the Republic of Serbia, as presented in Ref. 41.

In Fig. 8(b), we present a map of actual traffic accidents that occurred in 2015, 2016, and 2017[41] on the same section of road used in our research, i.e., from Usce town to Brvenik town, in central Serbia. All measurements were conducted referencing GPS coordinates.

The dots in Fig. 8(b) show the points where traffic accidents occurred that were caused by skidding or slippery surface conditions (dark spots represent major traffic accidents and light spots represent minor traffic accidents). All traffic accidents, shown in Fig. 8(b), were obtained by the police department, municipality of Kraljevo, Raska, and Novi Pazar, and they refer to years 2015, 2016, and 2017. Our moving vehicle collected data from the same part of the road. Data filtering concerning the previous traffic accidents in Fig. 8(b) was performed through the filter "slippery carriageway due to weather conditions."

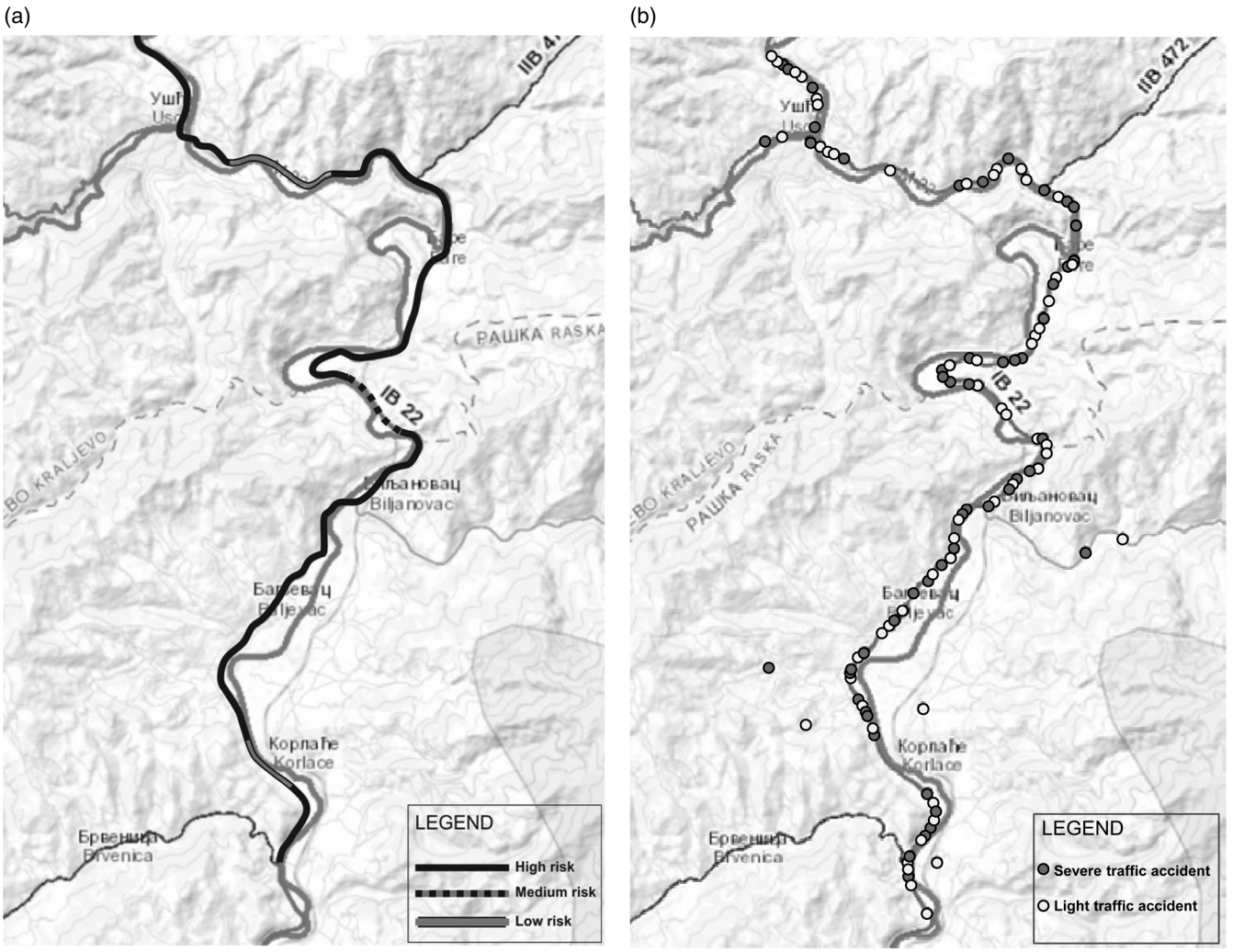

Fig. 8 (a) Proposed road mapping, according to road risk prediction and (b) map of real traffic accidents, according to the integrated database on security facilities of roads in the Republic of Serbia, for years 2015, 2016, and 2017.

By comparing the information represented in Figs. 8(a) and 8(b), we formulated our proposed risk map, by distinguishing roads into three classes: high, medium, or low risk, incorporating data relating to actual traffic accidents caused by slippery roads. We can conclude that the road segments that we have defined as high-risk zones [Fig. 8(a)] correspond entirely with the locations where traffic accidents actually occurred in years 2015, 2016, and 2017 [Fig. 8(b)]. Among 98 traffic accidents, only 1 had occurred in the section of the road we classified as "low risk," two accidents had occurred in the section of the road we classified as "medium risk," and the 95 accidents had occurred in the section of the road we classified as "high risk." This confirmed the validity of our risk prediction algorithm.

Improving road safety, especially on medium risk roads, can be done by providing connection technologies into the global intelligent transport network based on the communication system called IoV.[42] Vehicles in IoV can communicate with each other to determine and share the risk status of roads and enable drivers to adapt and react in real time. It is, therefore, a valuable addition to established risk zone mapping.

We suggest that our algorithm can also be implemented as a supportive tool for road risk prediction for autonomous driving decision making.[43]

The development of autonomous driving cars is recognized as one of the most challenging areas of automation ever undertaken. Highly detailed road mapping and road risk prediction must be developed and maintained to avoid accidents and crashes with autonomously driven cars. The issue does not only concern the engineering and technology for developing the vehicles themselves, but increasingly of concern is the question of how to prepare and maintain the roads for self-driving cars. We suggest that our algorithm can be used to support autonomous driving and navigation through its capability to locate and map high-risk road zones and predict dangerous driving conditions.

6 Conclusions

In this study, we have presented a risk prediction algorithm, based on road surface texture classification, using image texture processing, and a mobile vehicle scanning system. The algorithm enables the prediction of danger levels when driving on slippery and slick roads. Our aim is to produce a

highlighted road map classified according to the predicted risk of slippery road surfaces.

The road map of any area can be categorized according to skid risk, highlighting three classes of road: high risk (especially dangerous, slippery roads), medium risk, and low risk (nonslippery roads). In sections of the road where a high level of risk is detected, drivers should take precautions and prepare for slick roads and very dangerous driving conditions.

Surface texture analysis was performed by image processing of the video frames. The algorithm itself was upgraded with the correlation analysis based on the surface temperature data set obtained at the same time as the video frames, from the same moving vehicle.

Video frames were used as input images. A preprocessing technique, based on cropping frames to a suitable size, and gray level normalization, was performed. We then applied four methods of texture feature extraction: Gabor transform, co-occurrence matrix, edge histogram descriptor, and wavelet transform. The Gabor transform was the most effective. The other three methods were not suitable for texture classification of the frames captured by a moving camera that we used.

Three different measurements were used to compare the image frames for the purpose of texture characterization: Euclidean distance, Pearson correlation coefficient, and normalized mean value differences. The Pearson correlation was the most effective. Analysis of the correlation and the normalized difference of the mean values per particular frame, and compared to the reference frame, enabled us to classify the road surface into three categories.

We proved that it is possible to analyze the texture of dynamically captured images, despite the interference caused by the moving vehicle. We also proved that simultaneous acquisition of images and temperature data, referenced on the same measurement points by GPS coordinates, can validate surface texture classes and contribute to classification management. Of great significance was our comparison of our proposed road risk map based on our algorithm, with the actual database of traffic accidents caused by slippery roads. This confirmed that the high-risk zones that we identified with our algorithm corresponded to where accidents had actually occurred on the same sections of road. We were, therefore, able to conclude that we have identified a method for road risk prediction that can be used for navigation as well as an input and support mechanism for autonomous driving decision making.

Acknowledgments

This research was supported by the public enterprise "Roads of Serbia." Conflicts of interest: The authors declare no conflict of interest.

References

1. D. Lu et al., "The roles of textural images in improving land-cover classification in the Brazilian Amazon," *Int. J. Remote Sens.* **35**(24), 8188–8207 (2014).
2. R. Bala, "Survey on texture feature extraction methods," *Int. J. Eng. Sci. Comput.* **7**(4), 10375–10377 (2017).
3. "Road risk mapping," European Road Assessment Programme (EuroRap), 2018, http://www.eurorap.org/protocols/risk-mapping/.
4. "Road risk mapping design specification," The International Road Assessment Programme (iRAP), 2018, https://www.irap.org/2017/10/risk-mapping-design-specification/.
5. ASTM E1082-90, "Standard Test Method for measurement of vehicular response to traveled surface roughness," ASTM International, West Conshohocken, PA, http://www.astm.org (2012).
6. ISO 13473-1,2,3: 1997–2002. "Characterization of pavement texture by use of surface profiles—Part 1, Part2, Part 3," ISO Standard, https://www.iso.org/obp/ui/#iso:std:iso:13473:-1:ed-1:v1:en.
7. R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst. Man. Cybern.* **SMC-3**(6), 610–621 (1973).
8. T. Dahmen, "Modeling, simulation, and optimization of pacing strategies for road cycling on realistic tracks," Dissertation zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften an der Mathematisch-Naturwissenschaftliche Sektion, Fachbereich Informatik und Informationswissenschaft, Konstanzer Online-Publikations-System (KOPS) (2016).
9. P. Cairney and P. Bennett, *Surface Characteristics and Crash Occurrence on Selected Roads in Australia*, Austroads, Sydney, Australia (2009).
10. G. W. Flintsch et al., "Pavement surface macrotexture measurement and applications," in Transportation Research Record No. 1860, TRB, National Research Council, Washington, D.C., pp. 168–177 (2003).
11. V. Kastrinaki, M. Zervakis, and K. Kalaitzakis, "A survey of video processing techniques for traffic applications," *Image Vision Comput.* **21**, 359–381 (2003).
12. L. Huidrom, L. K. Das, and S. K. Sud, "Method for automated assessment of potholes, cracks and patches from road surface video clips," *Proc. Soc. Behav. Sci.* **104**, 312–321 (2013).
13. P. Prasanna et al., "Computer-vision based crack detection and analysis," *Proc. SPIE* **8345**, 834542 (2012).
14. M. Gavilán et al., "Adaptive road crack detection system by pavement classification," *Sensors* **11**(10), 9628–9657 (2011).
15. Y. X. Wai et al., "A low-cost video-based pavement distress screening system for low-volume roads," *J. Intell. Transp. Syst.* **22**(5), 376–389 (2018).
16. H. Oh, N. W. Garrick, and L. E. K. Achenie, "Segmentation algorithm using iterative clipping for processing noisy pavement images," in *Proc. 2nd Int. Conf. of Imaging Technol.: Tech. and Appl. Civil Eng.*, ASCE, Reston, VA, pp. 138–147 (1997).
17. C. Koch, G. M. Jog, and I. Brilakis, "Automated pothole distress assessment using asphalt pavement video data," *J. Comp. Civil Eng.* **27**(4), 370–378 (2013).
18. S. C. Radopoulou and I. Brilakis, "Patch detection for pavement assessment," *Autom. Constr.* **53**, 95–104 (2015).
19. G. Hadjidemetriou and S. Christodoulou, "Entropy-based automatic detection of multiple pavement defects," in *Conf.: Transp. Res. Board 97th Annu. Meeting*, Washington DC, United States (2018).
20. Y. Jo and S. Ryu, "Pothole detection system using a blackbox camera," *Sensors* **15**, 29316–29331 (2015).
21. B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. Mach. Intell.* **18**(8), 837–842 (1996).
22. D. Gabor, "Theory of communication. Part 1: the analysis of information," *J. Inst. Electr. Eng.—Art III: Radio Commun. Eng.* **93**(26) 429 (1946).
23. A. K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor Filters," *Pattern Recognit.* **24** (12) 1167–1186 (1991).
24. S. Mallat, "A theory for multi-resolution signal decomposition: the wavelet representation," *IEEE Trans. Pattern Recognit. Mach. Intell.* **11**, 674–693 (1989).
25. E. Zalama et al., "Road crack detection using visual features extracted by Gabor filters," *Comput.-Aided Civil Infrastruct. Eng.* **29**, 342–358 (2014).
26. A. C. Bovik, M. Clark, and W. S. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Trans. Pattern Anal. Mach. Intell.* **12**(1), 55–73 (1990)
27. P. Perona, "Deformable kernels for early vision," *IEEE Trans. Pattern Anal. Mach. Intell.* **17**(5), 488–499 (1995).
28. D. A. Clausi and M. E. Jernigan, "Designing Gabor filters for optimal texture separability," *Pattern Recognit.* **33**, 1835–1849 (2000).
29. M. Lyons et al., "Classifying facial attributes using A 2-D Gabor wavelet representation and discriminant analysis," in *Proc. 4th Int. Conf. Autom. Face and Gesture Recognit. (Cat. No. PR00580)*, Grenoble, France, pp. 202–207 (2000).
30. A. Sargano, P. Angelov, and Z. Habib, "A comprehensive review on handcrafted and learning-based action representation approaches for human activity recognition," *Appl. Sci.* **7**(1), 110 (2017).
31. D. A. Clausi, "An analysis of co-occurrence texture statistics as a function of grey level quantization," *Can. J. Remote Sens.* **28**(1), 45–62 (2002).
32. M. Campbell et al., "IBM Research TRECVID-2006 video retrieval system," in *NIST TRECVID Workshop*, Gaithersburg, USA (2006).
33. H. K. Ekenel et al., "Universitat Karlsruhe (TH) at TRECVID 2007," in *NIST TRECVID Workshop*, Gaithersburg, USA (2007).
34. M. Partio et al., "Rock texture retrieval using gray level co-occurrence matrix," in *Proc. 5th Nordic Signal Process. Symp., NORSIG 2002*, on board Hurtigruten, Norway, p. 5 s (2002).
35. B. S. Manjunath, P. Salembier, and T. Sikora, *Introduction to MPEG 7: Multimedia Content Description Interface*, Wiley & Sons, London (2002).

36. H. Ziari, A. T. Barakoohi, and A. Moniri, "Laboratory investigation of the effect of temperature on frictional properties of concrete pavements containing crushed glass," *Int. J. Pavement Res. Technol.* **10**(4), 297–303 (2017).
37. S. G. Jahromi et al., "Evaluation of pavement temperature on skid frictional of asphalt concrete surface," *Int. J. Pavement Eng.* **12**(1), 47–58 (2010).
38. Y. Lazarev et al., "Method of assessment and prediction of temperature conditions of roadway surfacing as a factor of the road safety," *Transp. Res. Proc.* **20**(17), 393–400 (2017).
39. K. Wu and M. Yang, "Mean shift-based clustering," *Pattern Recognit.* **40**(11), pp. 3035–3052 (2007).
40. J. Zhao, H. Wu, and L. Chen, "Road surface state recognition based on SVM optimization and image segmentation processing," *J. Adv. Transp.* **2017**, 6458495, 21 p (2017).
41. "The integrated database of characteristics of traffic safety," Road Traffic Safety Agency, Republic of Serbia, http://195.222.96.212/ibbsPublic/.
42. T. T. Dandala, V. Krishnamurthy, and R. Alwan, "Internet of Vehicles (IoV) for traffic management," in *Int. Conf. Comput. Commun. Signal Process. (ICCCSP)* (2017).
43. S. Al Dhahri et al., "A real-time perception system for autonomous cars targeting highways," *Proc. SPIE* **10752**, 107521B (2018).

Nikola Slavkovic received his BSc, MSc, and MPhil degrees in electrical engineering, all from the School of Electrical Engineering, University of Belgrade, Serbia. He is currently a PhD candidate in the School of Electrical Engineering, University of Belgrade and in the Communications Department of ICT College, Belgrade. He is a cofounder of FTTH Council Serbia. His research interests include signal processing, communication systems, sensor networks, communication optics, and modeling.

Milan Bjelica received his BSc, MSc, MPhil, and PhD degrees in electrical engineering from the School of Electrical Engineering, University of Belgrade, Serbia. He is currently in the Communications Department, working as an associate professor. His research interests include recommender systems, communication networks, multimedia, simulation, and modeling.