Road quality evaluation and road material recognition using 3D textons

Eng. Delia A. Mitrea
Computer-Science Department
Technical University of Cluj-Napoca
Baritiu 26-28, Cluj-Napoca
Romania
Delia.mitrea@cs.utcluj.ro

Prof. Phd. Eng. Sergiu Nedevschi
Computer-Science Department
Technical University of Cluj-Napoca
Baritiu 26-28, Cluj-Napoca
Romania
Sergiu.nedevschi@cs.utcluj.ro

Abstract – **This paper aims for finding an efficient and accurate texture-based method for road material recognition and road quality evaluation, in order to tune the vehicle speed accordingly and to avoid accidents. The application could be used in both autonomous and non-autonomous driving. The adopted method uses 3D textons for the detection of the microstructures that could be meet in the asphalt.**

I. INTRODUCTION

Texture is a very important feature, essential in human visual perception, suitable for material recognition and material quality evaluation. Many image processing methods try to characterize texture as a spatial arrangement of gray levels in the image (Gray Level Co-occurrence Matrix - GLCM, Difference Histogram [6]), by statistics related to the spatial distribution of primitives like edges (edge based first order statistics, edge based Generalized Co-occurrence Matrix, GCM [7]), by models like fractals or Markov Random Field Models (MRF), or calculating a transform (Fourier, Gabor, Wavelet) [6].

A more accurate method is detecting the types of microstructures in the texture - edges, ridges, spots, waves - and calculate a statistic that characterizes the spatial arrangement of these microstructures, called textons [1].

The goal is to characterize the 3D texture in order to recognize it under various illumination, orientation or scale conditions. The classical methods for texture analysis are more suitable for 2D texture and fail in these situations. The method based on textons could be very easily adapted to do this, by considering the appearance of texture under all possible orientation and illumination conditions and using a Markov-Chain-Monte-Carlo method for material characterization.

Particularizing to road quality analysis, we look for the frequency of microstructures like ridges, waves, edges or spots in order to characterize rough, waved or damaged surfaces. A more efficient analysis refers to defect detection and localization . This could be done by splitting the image into regions and comparing the texture in a certain region with the texture in the neighboring regions.

II. 3D TEXTONS IN 3D TEXTURE ANALYSIS AND RECOGNITION

A. Motivation

3D texture analysis could mean evaluating the structure of the intensity surface of the image. This is the main purpose of some fractal-based methods, like those in [2] or in [3].

The first one counts the number of boxes which are contained in the structure of 3D objects (box-counting method). This number represents the fractal dimension associated with that object and is defined in the following way:

$$\dim_B F = \lim_{\delta \to 0^+} \frac{\log N_\delta(F)}{-\log \delta} \quad (1)$$

where F is a set of R^n, and $N_{\delta}(F)$ is the minimum number of subsets of ray δ that are contained in F. This number could be approximated through the slope of the graph ***($\log N_{\delta}(F)$, $\log \delta$)*** .

The structure of the 3D intensity surface can be also characterized by the directions of the normal vectors in each point, or by calculating the image shape spectrum (ISS) in each point p:

$$S(p) = \frac{1}{2} - \frac{1}{\pi} \cdot \arctan \frac{k_1(p) + k_2(p)}{k_1(p) - k_2(p)} \quad (2)$$

where $k_1(p)$ and $k_2(p)$ - the principal curvatures, maximum and minimum, of the intensity surface in that point, are computed using the first and second partial derivatives of the intensity function I [13].

For making the analysis more accurate we can combine these 3D specific features with classical methods like GLCM, GCM or MRF and we can define microstructures as groups of pixels having the same values of the features.

The problem is that the structure of the intensity surface, as well as the intensity values in the image and the appearance of microstructures change when the orientation and illumination conditions change. So, we must consider the texture under all possible conditions of orientation and illumination and define the microstructures accordingly. This is done by using the method of 3D textons .

B. *2D and 3D textons*

The texton is the elementary structure of the texture which corresponds to the texture cell from the classical definition. Each texton is characterised by the similarity of the intensity values of its pixels and the texture is as a spatial arrangement the textons. Each texton is defined by a feature vector that expresses the common characteristics of the pixels in that texton, and which is the centre of the cluster that represents that texton.

In the 2D case, each pixel has associated a feature vector that characterise it from a certain point of view, considering its relation with the neighbours. There are many pixels that have approximatively the same feature vectors, so they can be grouped in a single class or cluster. After defining the classes, a single representative vector is chosen for each class; this will be the center of the class, or the appearance vector of that texton [1].

In the 3D case there are the following problems: different microstructures generate the same appearance in various conditions of orientation and illumination and the same microstructure can have different appearances when the orientation and illumination conditions varies.

The algorithm that uses 2D textons will be insensitive to these problems and will fail, so we must consider more images, for different orientation and illumination conditions. In this case, the feature vectors for the 3D textons will be the result of chaining the feature vectors of the correspondent pixels in images representing the same material under various orientation and illumination conditions (will be f size $N_{fil}N_{cond}$, where N_{fil} is the number of filters and N_{cond} is the number of conditions).

C. *Recognition of the material*

First, a representation of the visual appearance of the material is due. This is done by constructing the histogram of the 3D textons in the image. Then, we need a histogram matching algorithm in the recognition phase. We can use the chi-squared distance to compute the similarity of two histograms [1]:

$$\chi^2(h_1,h_2)=\frac{1}{2}\sum_{n=1}^{\#bins}\frac{(h_1(n)-h_2(n))^2}{h_1(n)+h_2(n)} \quad (3)$$

Where #bins is the number of textons in the texton vocabulary.

Then, we must decide on the number of images used for the recognition phase. If we use multiple images, under the same conditions of orientation and illumination as the training set, we are in the classical case of recognition by histogram matching.

In practical cases, especially in real-time applications, it is more suitable to consider a single test image , taken in the given conditions of orientation and illuminations. Because, as stated before, various microstructures may generate the same appearance, we can assign more possible texton labels to the same pixel. The cyclic problem of establishing the real configuration of textons in the material and, simultaneously, recognizing the material according to the training set, is solved using a Markov-Chain-Monte-Carlo Algorithm [12].

III. FEATURE EXTRACTION

In order to identify in the texture microstructures like ridges, edges, waves or spots we have chosen the Laws energy measures and the corresponding convolution mask for level, edge, spot, wave, ripple [6].

So, for each pixel is calculated the response to the following filters:

The level - used in order to represent a region of higher intensity level than the neighbouring regions.

$$L_5=\begin{bmatrix}1 & 1 & 1\\ 1 & 2 & 1\\ 1 & 1 & 1\end{bmatrix}$$

The edge - in order to represent a transition between two regions of various intensity levels, or some lesions of the surface :

$$E_5=\begin{bmatrix}0 & 0 & -1 & 0 & 0\\ 0 & 0 & -2 & 0 & 0\\ -1 & -2 & 0 & 2 & 1\\ 0 & 0 & 2 & 0 & 0\\ 0 & 0 & 1 & 0 & 0\end{bmatrix}$$

The spot - in order to identfy small regions of image that contain pixels with the same properties and that differ from the neighboring regions, having a greater or smaller intensity level, representing pits or 2D structures that characterize the texture:

$$S_5=\begin{bmatrix}0 & 0 & -1 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\\ -1 & 0 & 4 & 0 & -1\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & -1 & 0 & 0\end{bmatrix}$$

The wave - for finding regionscharacterized by repeated variations of the intensity - from low towards high values, then again towards low values, and again towards high values, in order to represent the waves on the road surface :

$$W_5 = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ -1 & -2 & 0 & 2 & 1 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

We also used the Image Shape Spectrum (2) in order to characterize the local curvature of the intensity surface in the current point, and the Laplacian of Gaussian convolution mask for different dimensions(11x11, 5x5,3x3) [10].

IV. MATERIAL RECOGNITION

A. *The training phase*

Multiple images taken from different regions of the same road material are used in order to learn the system about a large number of possible instances of that material. In the same time, we use multiple images of the same region of material under various conditions of orientation and illumination. First, the features of the material are extracted, from all the instances of the material under the considered conditions of orientation and illumination, then the 3D textons are formed using a k-means clustering method [11]. The number of centers, k, will be chosen according to the image dimension, in order to avoid the possible loss of information .

After labeling each pixel with the label of the corresponding texton, we build histograms of textons for each material and memorize these histograms on the disk.

B. *The recognition phase*

The recognition is done using a single image, representing an unknown material in a certain condition of orientation and illumination.

We build the texton vocabulary, but we assign, to each pixel many labels, representing the possible textons to which the pixel belongs. These labels are assigned considering the distance to the appearance vectors, for those textons corresponding to the same distance between the appearance vector of the texton and the feature vector of the pixel.

Then, the Markov-Chain-Monte-Carlo method is used in the following way:

1. Establish an initial configuration of the textons by chosing, for each pixel, one of the possible labels at random.
2. For each material in the training set:
   - 2.1. Read the histogram memorized during the training set for that material
   - 2.2. For each quarter of the test image
     - 2.2.1. Repeat:
       - 2.2.1.1. Change the pixels labels chosing, by random, another label from the possible labels.
       - 2.2.1.2. Calculate the histogram of textons on the entire image, in the new configuration
       - 2.2.1.3. Calculate the probability of the new state.
       - 2.2.1.4. If the probabilty of the new state is greater than the probability of the current state, then accept the new state
       - 2.2.1.5. Otherwise, reject the new state and come back to the previous configuration of labels
       - 2.2.1.6. Termination condition: the assignment of labels remains the same for many iterations, so the convergention is reached.
   - 2.3. If the probability of the state obtained for the current material is greater than the maximum probability obtained for the previous materials, assign the image to the current material.

The probability of the current state can be estimated like being in reverse proportion with the chi-squared distance of the histograms , or we can directly compute the probability that two histograms of the same material have a distance larger than χ [1].

V. DEFECT DETECTION AND LOCALIZATION

In order to detect and localize the defects in the texture, we split the image in multiple regions . After defining the textons for the entire image, we calculate the histogram of textons for each region. We compare the histogram of each region with the histograms of the neighboring regions. If there are significant differences between these histograms, then we can consider that we have found a region with defect. We try to extend the region with defect using a region growing technique based on the similarity of histograms. The algorithm is the following:

1.Define textons on the material represented by the current image

2.Split the image in square zone, of dimension 10x10 pixels.

3. Compute the histogram on each zone.

4.Compare the histograms of the neighbouring regions. Compute and memorize the differences between the histograms of neighbouring regions to the first histogram (hi).

5. Determine the biggest difference between the histograms. The region corresponding to the second histogram is the region with defect and will be marked.

6. Analyse the histograms of the neighbouring regions of the defect region and extend the marked region to these zones as well, if the difference between histograms is small enough.

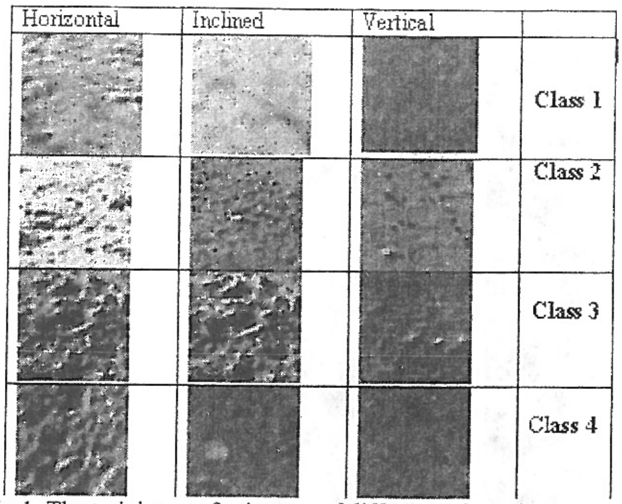
VI. EXPERIMENTAL RESULTS

In another experiment we used 4 classes of bitmap images of dimensions 80x80, in number 7x3, each class representing the same type of asphalt under three orientations: horizontal, vertical, inclined by an angle of approximately 45 degrees. A part of the training set and part of the results are illustrated bellow.

In another experiment, we used 6 classes of different types of asphalt. Each class contained bitmap images of dimensions 80x80, in number 8x3, representing various regions of the same type of road, in three different conditions of illumination - the images were taken in the morning, at noon and in the afternoon. In the recognition phase we used one single image, representing a certain region of an unknown material, under one of the considered conditions of orientation and illumination.

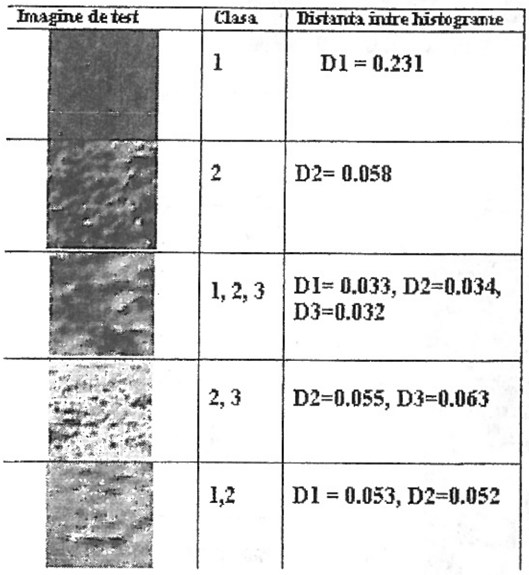
One of our future goals is using a single image of the unknown material, under unknown conditions of orientation and illumination and predict the texture configuration of each material in the training set, under intermediary conditions of orientation or illumination reported to the images in the training set. A similar method is used in [5], where the authors predict the intensity values in the image and the new directions of the normal vectors, for various distances from the camera, using only two images taken at two distances – maximum and minimum .

The training set and the results are illustrated bellow:



Horizontal	Inclined	Vertical	
			Class 1
			Class 2
			Class 3
			Class 4

Fig 1. The training set for images of different orientations



Imagine de test	Clasa	Distanta intre histograme
	1	D1 = 0.231
	2	D2= 0.058
	1, 2, 3	D1= 0.033, D2=0.034, D3=0.032
	2, 3	D2=0.055, D3=0.063
	1,2	D1 = 0.053, D2=0.052

Fig 2. The first result set

In our first experiment, we used images taken from pieces of road, which we rotate under different angles. In the second experiment, we used images take directly from the road, keeping the camera at a constant distance and repeating the same experiment in the morning, at noon and in the afternoon.

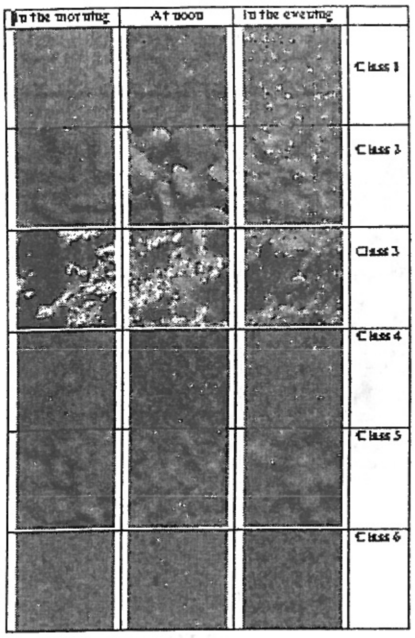


Fig 3. The training set for images under different conditions of illumination

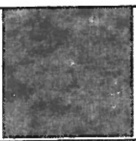
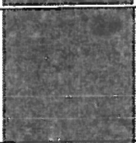
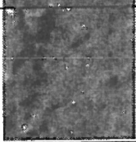
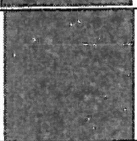
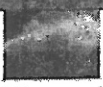
Test image	Class	Distance between histograms
	5	D5=0.096
	1, 4	D1= 0.180 D4 = 0.173
	1,4,5	D1= 0.183 D4 = 0.141 D5 = 0.123
	1	D1 = 0.375
	2	D2 = 0.187

Fig 4. The second result set

Evaluating the results of all the experiments of material recognition, we can say that the results were good in 70% of cases.

For defect detection and localization we used bitmap images of size 60x60. Some of the results are the following:

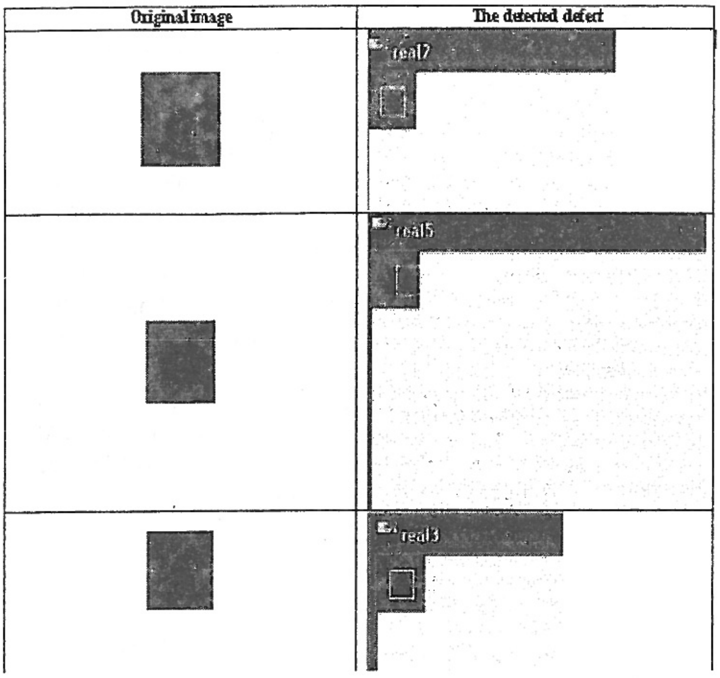


Fig 5. Defect detection and localization

VII. CONCLUSIONS

In conclusion, we can say that we obtained good results in material recognition and also in defect detection and localization, thanks to the Laws convolution masks that detect microstructures like ridges, waves, ripples, to the Laplacian of Gaussian convolution masks, to the image shape spectrum (ISS) and to the method, based on textons, for texture analysis and recognition. The applied method is viable regardless the orientation and illumination conditions.

The method can be developed to be more sensitive to textons spatial arrangement by using Textons Co-occurrence Matrices instead of histograms , but this will appreciably slow the algorithm and increase the memory requirements.

Regarding the insensitivity to the scale, there are two things to be done: 1. use multiple images in the training set, taken for various distances to the observer; 2. use Laws and LoG convolution masks of various dimensions in order to detect smaller or bigger microstructures.

In order to improve speed, the image can be split into regions and the texton setting-up algorithm can be executed in parallel on each region. Also, the feature extraction phase can be speed-up using hardware convolution methods.

VIII. REFERENCES

[1] T.Leung, J.Malik, Computer Science Division, University of California at Berkley, "Representing and Recognising the Visual Appearance of Materials using Three-dimensional Textons", *International Journal of Computer Vision*, Dec. 1999

[2] Mircea Olteanu Politehnica University Bucharest, Math. Dept. 2, Mihai Tanase - Computer Science Faculty, Politehnica University Bucharest, Cluj-Napoca, Romania: "An algorithm for the analysis of fractal-like structures and miscellaneous applications", *A&QT-R 2004 (THETA 14) 2004 IEEE-TTTC International Conference on Automation, Quality and Testing, Robotics May 13 – 15, 2004*

[3] Shmuel Peleg, Joseph Naor, Ralph Hartley, David Avnir- IEEE Transaction on Pattern Analysis and Machine Intelligence, *"Multiple Resolution Texture Analysis and Classification"* Vol. PAMI-6, No.4, July 1984

[4] S. Barsky and M. Petrou, "The 4-source photometric stereo technique for 3-dimensional surfaces in the presence of highlights and shadows. Analysis and Machine Intelligence", (2003), *IEEE Transactions on Pattern*

[5] X. Lladó, V. Ila, J. Freixenet and J. Martí, Computer Vision and Robotics Group.University of Girona, Spain,"Evaluating surface texture prediction when seen from different distances", *A&QT-R 2004 (THETA 14) 2004 IEEE-TTTC- Quality and Testing, robotics May 13-15, 2004 Cluj-Napoca, Romania, International conference on Automation*

[6] Larrry S. Davis, Department of Computer Sciences, University of Texas at Austin, Austin, Texas 78712: *"Image Texture Analysis Techniques – A Survey"*

[7] Larry S. Davis, M. Clearman, J.K. Aggarwal: *"A Comparative Texture Classification Study Based on Generalised Co-occurrence Matrix"*

[8] M. Hassner and J. Sklansky, Image Modelling, A. Rosenfeld (Ed.), *"The Use of Markov Random Fields as Models of Texture"*, Academic Press, 1981, 185-198

[9] Andrzej Materka and Michael Strzelecki, Technical University of Lodz *"Texture Analysis Methods – A Review"*, Institute of Electronics ul. Stefanowskiego 18, 90-924 Lodz, Poland

[10] Sergiu Nedevschi,*"Prelucrarea imaginilor si Recunoasterea Formelor"*, Editura Albastra, Cluj-Napoca, 1998,

[11] R. O. Duda, P. E. Hart and D. G. Stork, John Wiley & Sons, *"Pattern Classification "* 2nd ed, 2000,

[12] Stuart Russel, Peter Norvig, *"Artificial Intelligence. A modern Approach"*,

[13] Gh. D. Ionescu, *"Teoria diferentiala a curbelor si suprafetelor cu aplicatii tehnice"*, editura Dacia, Cluj-Napoca, 1984