Surface Classification for Road Distress Detection System Enhancement

M. Gavilán, D. Balcones, M. A. Sotelo, D. F. Llorca, O. Marcos, C. Fernández, I. García, and R. Quintero

Computer Engineering Department, University of Alcalá, Madrid, Spain
{miguel.gavilan,sotelo,llorca}@aut.uah.es

Abstract. This paper presents a vision-based road surface classification in the context of infrastructure inspection and maintenance, proposed as stage for improving the performance of a distress detection system. High resolution road images are processed to distinguish among surfaces arranged according to the different materials used to build roads and their grade of granulation and striation. A multi-class Support Vector Machine (SVM) classification system using mainly Local Binary Pattern (LBP), Gray-Level Co-occurrence Matrix (GLCM) and Maximally Stable Extremal Regions (MSER) derived features is described. The different texture analysis methods are compared based on accuracy and computational load. Experiments with real application images show a significant improvement on the the distress detection system performance by combining several feature extraction methods.

Keywords: Road surface classification, Multi-class SVM, Local Binary Pattern, Gray-Level Co-occurrence Matrix.

1 Introduction

Governments and institutions have been making a great effort to accomplish the objective of achieving a high quality road network along the last decades. They are, now more than ever, fully aware of the necessity of an adequate road inspection and maintenance. Nonetheless, authorities are not properly carrying out these actions because of the fact that there is still a lack of suitable methods to assess the road conditions and perform management programs both effectively and efficiently.

Distress measurement is an crucial factor when evaluating road quality. Manual road inspections are expensive and risky, so do not allow carry out a proper road maintenance. An automatic distress detection to quantify the quality of road surfaces and assist in prioritising and planning the maintenance of the road network become essential. This system must be robust and able to adapt itself to different road surfaces. However, available commercial systems share two mayors problems that strongly affect their consistence when a huge amount of road kilometers are analyzed. First, the presence of non-crack elements such as joints, repaired cracks, damaged white-lane marks and loops, which are usually

R. Moreno-Díaz et al. (Eds.): EUROCAST 2011, Part I, LNCS 6927, pp. 600–607, 2012.

incorrectly considered as road crack defects. Second, the strong dependency on the type of surface surveyed. Rough textures cause a considerably increment in the amount of reported false positives. Therefore, a surface classification is necessary so that the algorithm parameters can be adjusted to the challenges presented by the differences amongst the surface materials.

In this paper a road surface classifier to assist a distress detection algorithm is proposed. Texture classification is an old problem which has been studied during decades. A great amount of approaches can be found in the literature, therefore only a brief survey of the most important texture analysis methods is presented. They can be divided into four main groups: statistical, structural, filtering methods and model-based methods.

Statistical methods are one of the earliest methods proposed, they defend that qualities textures are based on the spatial distribution of gray values. One of the most known texture analysis methods is the Gray-Level Co-occurrence Matrix (GLCM) technique [1], which has been used with great successful in several applications, such as evaluation of rock images [2]. Finally autocorrelation methods, such in [3] where the usage local autocorrelation statistics of up to eighth order are proposed.

Structural methods extract texture primitives or texture elements, then placement rules are computed in order to describe texture, [4]. For example, in [5] a morphological pyramid is exploited through the computation of a multiresolution co-occurrence matrix, in order to distinguish among four types of road pavement.

Filtering methods are the biggest group due to the great number of different existing techniques. Besides of that, all of them follow a common procedure. A linear transform, filter, or filter bank is computed followed by some energy measurements that describe the texture image. A good review and comparative can be found in [6]. Some of the methods are based on Fast Fourier Transform (FFT), ring and wedge filters and other discrete transforms, such as discrete cosine transform (DCT) or Walsh/Hadamard transforms (WHT). But the two most popular are Gabor filters, firstly used to classify textures by [7], and those based on Wavelets Transformations, introduced by [8].

Last group, model-based methods, try to model the texture image according to a probability model or linear combinations of basic functions. The most popular of them are based on model random fields and fractal models.

The rest of this paper is organized as follows: Section II presents a brief overview of the distress detection system, Section III describes the feature extraction methods and Section IV presents the experimental results. Finally, conclusions and future work are discussed in Section V.

2 Distress Detection System Overview

The main goal of the automatic distress detection is to assess the road network quality. A global overview of the main components is shown in Figure 1.

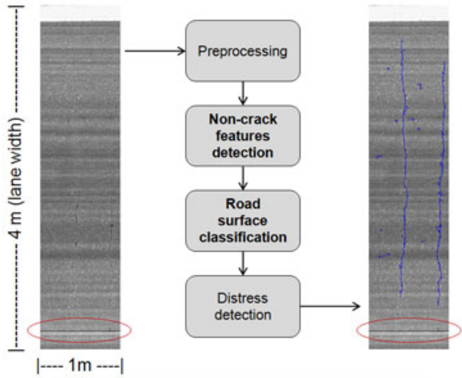


Fig. 1. Distress detection system. A global overview.

The inputs of the system are images of 4000x1000 pixels with 1 millimeter of resolution. A basic approach usually consists of a preprocessing stage together with the strictly speaking distress detection module. Nevertheless, these systems incorrectly report cracking at the edge of non-crack elements, mainly, joints, patches and road markings, so its necessary to identify and remove non-crack features them in order to reduce false positives. In addition, due to the high diversity of road surfaces, it is necessary to adjust the algorithm parameters by means of a classification stage.

The proposed distress detection detection systems also include both, the non-crack features detection and the road surface classification steps. Non-crack features detection module identifies the mentioned non-crack elements, reducing drastically the amount of false positives reported. Regarding the road surface classification module, its objective is to classify surfaces into different classes so the algorithm parameters can be adjusted for each type of surface.

The number of classes is chosen by visual inspection, according to the materials used to build the roads and the grade of granulation and striation. An example of each one of the 10 classes considered, of which 7 of them are bituminous (asphalt) materials and 3 of them concrete, is shown in Figure 2.

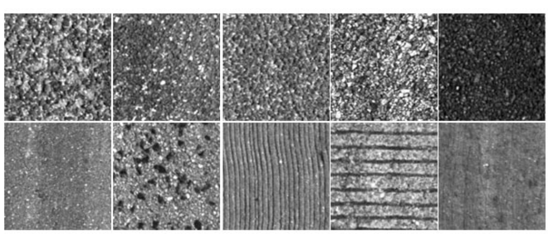


Fig. 2. Surface classes

The classification of the surfaces is carried out using a multi-class SVM [9], based on the standard one-versus-one approach to reduce the single multi-class problem into multiple binary classification problems. The maximum-margin hyperplane which best separates the 10 classes is constructed by solving an optimization problem. A linear kernel function has been chosen as it achieves the best accuracy when comparing with polynomial, radial basis and sigmoid kernel functions, keeping a low complexity.

A set of up to 20 squared subimages of 256x256 pixels are used to perform the prediction by majority voting decision. Subimages must be placed in the central columns of the image and avoiding the use of pixels masked as non-crack features. In order to obtain the feature vector different groups of calculations are computed from each subimage, as described in the next section.

3 Features Vector

Several groups of calculations are computed from subimages of 256x256 pixels to obtain the statistics which will form the feature vector. It is necessary to guarantee that processing times do not become too long, so that the application usage in a real exploitation scenery remains feasible. Different extraction method settings and feature group combinations are evaluated in order to choose the most suitable trade-off between classification prediction accuracy and computational load.

Some feature extraction methods, such as Gabor Filters and Wavelets, have been tested and discarded, as their inclusion does not provide a significant performance improvement. The main components of the feature vector are obtained using Gray-Level Co-Occurrence Matrix (GLCM), Maximally Stable Extremal Regions (MSER) and Local Binary Pattern (LBP). These three texture analysis techniques allow implementation variations contributing with different feature vector lengths and present open parameters. Therefore, each one of them will be presented separately in this section.

In order to take advantage of the different gray-level histogram distributions of each surface class, four statistics are also computed from the gray-level histogram: average, deviation, skewness and kurtosis. Finally, two components are obtained using the 1-D Fast Fourier Transform (FFT). These statistics are the averaged module over four equispaced rows and over four equispaced columns.

3.1 Gray-Level Co-occurrence Matrix (GLCM)

The Gray-Level Co-Occurrence Matrix, introduced by Haralick [1], stores how the often different combinations of pixel gray-levels occur in the image, separated by a particular distance in a specific direction. The most common setting considers 1 pixel of distance and four angles or pixel relationship directions (0^o, 45^o, 90^o and 135^o). In our approach, four angles are also considered but up to four pixel distances (1, 3, 5 and 7) are evaluated (GLCMM), resulting in 16 different matrixes. An alternative using a pyramid image reduction approach is also evaluated (GLCM-pyr).

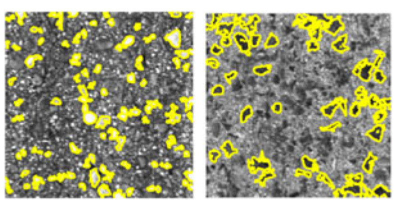
Haralick defines a set of 14 texture measurements based on this matrix of which only 4 have been considered relevant to the surface classification problem. They are homogeneity, entropy, contrast and energy.

The GLCM calculation requires an image quantification step in order to reduce the computational load. Different quantification levels have been tested, concluding that not a significant improvement is reached with more than 20 levels.

3.2 Maximally Stable Extremal Regions (MSER)

The different classes of surface present varied grades of granulation, varying in size and gray-level, as well as in their distribution, and resulting on either rougher or smoother surfaces. Therefore, blobs detection will help to improve the classification.

A solution based on Maximally Stable Extremal Regions (MSER) detection has been implemented. Although the MSER algorithm was originally proposed in [10] as an approach to solve the wide-baseline stereo problem, it has also been demonstrated to be a powerful blob detector, since it presents invariance to monotonic transformation of image intensities and allows multi-scale detection. An extremal region is a set of connected components so that maintains the intensity levels of them below a threshold. Maximally extremal regions are extremal region which satisfy a stability criterion. Minimal extremal regions also are detected inverting the original image, so dark and bright blobs are both detected. Regions are filtered according to their shape, so only regions that satisfy circular shape restrictions are considered. Two blob size ranges are distinguished. Three statistics are computed for each one of the ranges and gray-levels: the number of blobs, the averaged size and total area. Figure 3 shows white and black blobs detection result examples.



(a) (b)

Fig. 3. Maximally stable extremal regions detection. a) white blobs. b) black blobs.

3.3 Local Binary Patterns (LBP)

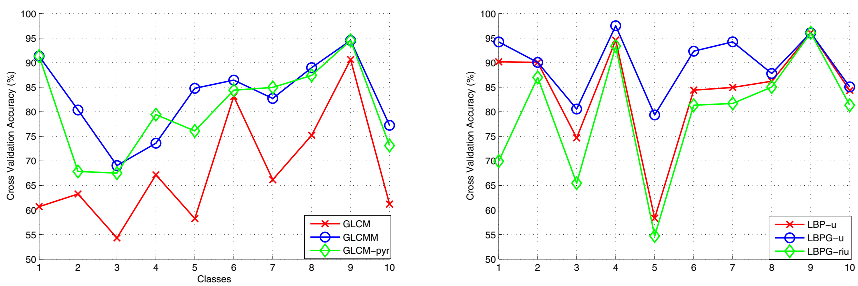
Local Binary Pattern (LPB) is a gray-scale and rotation invariant texture operator introduced by Ojala [11]. LBP is built by thresholding pixels in a 3x3 window with its central pixel. If the pixel value is greater or equal than the central pixel, it is set to 1 otherwise it is set to 0. LBP codifies the central pixel as the sum of the pixel values weighted according to their position around the central pixel.

Some extensions have been tested. For example, the variation of the distance of the samples (pixels around the central one) or the number of samples have not shown significant improvement in the performance.

Histograms formed only by uniform patterns (that contain at most two bitwise transitions both from 0 to 1 and vice versa), are usually computed (LBP-u). A common extension is to consider rotational invariant patterns together with the uniform ones (LBP-riu). Besides, LBP is also performed on the gradient image in both directions, horizontal and vertical (LBPG).

4 Experimental Results

A 10-fold cross validation algorithm over a training set of 9000 images of 256x256 pixels is used to assess the SVM performance for different feature extraction configurations and their combinations. In Figure 4(a) a comparison among implementations of the GLCM method is shown. The LBP configurations that reach the best performances are shown in 4(b).



(a) GLCM configurations (b) LBP configurations

Fig. 4. GLCM and LBP configurations comparison

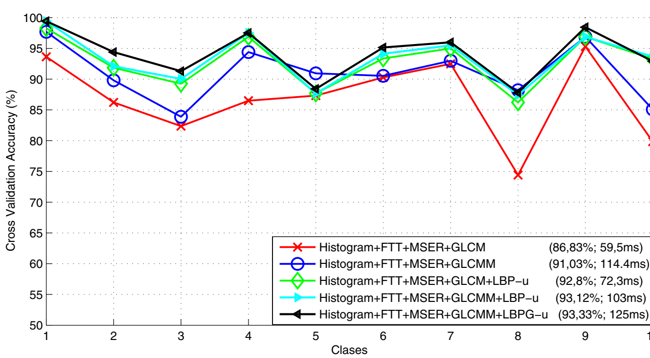


Fig. 5. Full feature vectors comparison

Finally, different combinations of the evaluated groups of statistics are shown in Figure 5. All of them contain MSER, Histogram and FFT derived statistics as well as features obtained from different LBP and GLCM implementations.

The feature vector chosen is the one which includes GLCM and LBP-u. This configuration achieves a weighted average accuracy of 92.2% with an acceptable computation time.

The confusion matrix for the selected configuration is shown in Table 1.

Table 1. SVM confusion matrix

	1	2	3	4	5	6	7	8	9	10
1	98.27	0.00	1.73	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00	91.84	2.04	1.53	0.00	4.59	0.00	0.00	0.00	0.00
3	0.77	0.26	89.26	0.51	7.93	1.28	0.00	0.00	0.00	0.00
4	0.00	1.04	1.04	86.88	0.83	0.21	0.00	0.00	0.00	0.00
5	0.00	2.17	4.35	0.72	87.68	5.07	0.00	0.00	0.00	0.00
6	0.00	3.07	1.79	0.00	1.02	93.35	0.26	0.26	0.00	0.26
7	0.75	0.50	1.75	0.75	0.00	1.25	94.99	0.00	0.00	0.00
8	0.00	0.00	0.00	0.0	0.00	1.18	0.00	86.22	0.00	12.06
9	0.00	0.00	0.00	0.00	0.00	0.78	0.78	0.00	96.88	1.56
10	0.00	0.37	0.00	0.37	0.00	1.49	0.00	4.48	0.00	93.28

Figure 6 shows the improvement in the percentage of detected cracked cells achieved by tuning the algorithm parameters according to the road surface classification. The experiment has been carried out with a set of 7250 images collected from 30 Spanish roads. The cracked cells reported without adjustment to the parameters present a lot of false positives. On the contrary, when using the road classification the automatic systems follows the Ground truth pretty accurately.

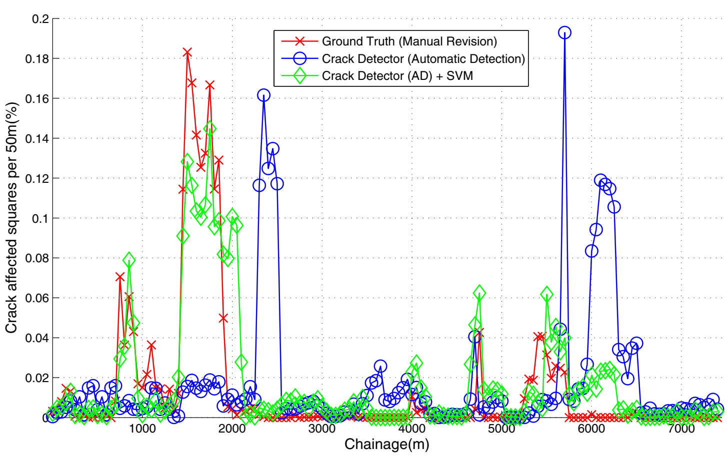


Fig. 6. Improvement achieved by the road surface classification step

5 Conclusions and Future Work

A high multi-class classification accuracy has been achieved by combining several texture analysis methods. Thus, different parameters can be used depending on the road surface, decreasing false positive rate in rough textures and detecting thinner cracks in soft textures, so that cracked length and cracked cells ratios are both improved. As future work, a hierarchical approach will be implemented, preforming firstly a concrete or non-concrete classification. As a future work a hierarchical approach will be implemented, preforming firstly a concrete or non-concrete classification.

Acknowledgement. This work has been funded by Acciona Ingeniería S.A.

References

1. Haralick, R.M., Bosley, R.: Texture features for image classification. In: Third ERTS, NASA Goddard Space Flight Center, Greenbelt, MD pp. 1929–1969 (1973)
2. Partio, M., Cramariuc, B., Gabbouj, M., Visa, A.: Rock texture retrieval using gray level co-occurrence matrix. Norsig (2002)
3. Toyoda, T., Hasegawa, O.: Texture Classification Using Extended Higher Order Local Autocorrelation Features. In: The 4th International Workshop on Texture Analysis and Synthesis, pp. 131–136 (2005)
4. Chen, J.-L., Kundu, A.: Rotation and Gray Scale Transform Invariant Identification Using Wavelet Decomposition and Hidden Markov Model. PAMI 16(2) (February 1994)
5. Paquis, S., Legeay, V., Konik, H.: Road Surface Classification by Thresholding Using Morphological Pyramid. In: ICPR (2010)
6. Randen, T., Husoy, J.: Filtering for texture classification: A comparative study. IEEE Trans. Pattern Analysis Machine Intelligence 21(4), 291–310 (1999)
7. Jain, A.K., Farrokhnia, F.: Unsupervised texture segmentation using Gabor filters. Pattern Recognition 24(12), 1167–1186 (1991)
8. Mallat, S.G.: A theory for multiresolution signal decomposition: The wavelet representation. IEEE Trans. PAMI. 11, 674–693 (1989)
9. Chang, C.C., Lin, C.-J.: LIBSVM: A Library for Support Vector Machines (2001), `http://www.csie.ntu.edu.tw/~cjlin/libsvm`
10. Matas, J., Chum, O., Urban, M., Pajdla, T.: Robust wide baseline stereo from maximally stable extremal regions. In: Proceedings of BMVC, pp. 384–393 (2002)
11. Ojala, T., Pietikaéinen, M., Harwood, D.: A Comparative Study of Texture Measures with Classification Based on Feature Distributions. PR 29, 51–59 (1996)