AUTOMATIC ROAD CRACK SEGMENTATION USING ENTROPY AND IMAGE DYNAMIC THRESHOLDING

Henrique Oliveira[1,2], Paulo Lobato Correia[1]

[1]Instituto de Telecomunicações - Instituto Superior Técnico, Universidade Técnica de Lisboa
Av. Rovisco Pais, 1, 1049-001 Lisboa, Portugal
[2]Escola Superior de Tecnologia e Gestão – Instituto Politécnico de Beja
Rua Afonso III, 1, 7800-050 Beja, Portugal
hjmo@lx.it.pt, plc@lx.it.pt

ABSTRACT

Human observation is commonly used to collect pavement surface distress data, during periodic road surveys. This method is labour-intensive, subjective and potentially hazardous for both inspectors and road users. This paper presents a novel framework for automatic crack detection and classification using survey images acquired at high driving speeds. The resulting images are pre-processed using morphological filters for reducing pixel intensity variance. Then, a dynamic thresholding is applied to identify dark pixels in images, as these correspond to potential crack pixels. Thresholded images are divided into non-overlapping blocks for entropy computation. A second dynamic thresholding is applied to the resulting entropy blocks matrix, used as the basis for identification of image blocks containing crack pixels. The classification system then labels images as containing horizontal, vertical, miscellaneous or no cracks. Two image databases are used for test purposes, to infer about the method's robustness, one of which acquired using professional high speed equipment.

1. INTRODUCTION

Roads are important infrastructures that exhibit distresses due to their constant usage. These distresses, usually in the form of cracks in pavement surface, reduce pavement performance, implying loss of asset value, poor quality of service and constraining the access to remote areas. To avoid such problems, good road maintenance policies are required, relying on the establishment of adequate rehabilitation management procedures.

Road surveys provide the necessary data collection tools about the pavement surface condition. To achieve this, inspectors typically travel along the surveyed roads collecting data (including images) about surface distress types and locations. This type of survey, based on human observation, requires driving slowly along the road (usually less than 10 km per day), prone to subjectivity, since two inspectors can produce different analysis results over similar distress situations, and raise some security concerns for both the inspectors and the other road users, especially in high speed roads like highways [1].

To successfully implement adequate rehabilitation actions, automatic systems for fast and reliable pavement surface defects acquisition and analysis are being developed, instead of relying solely on the more conventional, slow, labour-intensive and subjective, human inspection procedures [2]. Automation of the procedures leads to cost reductions, as well as more objective and standardized rehabilitation decisions.

An image based automatic pavement surface distress survey system poses some challenges. The hardware for data acquisition can be complex and expensive [3] [4], and also complex data processing techniques are needed due to the variability of pavement conditions and textures. Neural networks [1][5], Markov random fields [6], artificial living systems approaches [7], among others [2], have been reported in the literature.

In this paper, a simple system architecture for the analysis of images acquired during road surveys is proposed. Two distinct databases are used for testing purposes: the first one is acquired using the high speed acquisition system; the second was collected during a human observation survey. A different analysis technique for crack detection had been previously implemented over this second database, allowing a comparative study to be conducted, by using well-known metrics and a ground truth manually created for evaluation purposes.

The rest of the paper is organized as follows. Section 2 briefly presents the image acquisition procedures. The proposed unsupervised crack detection and classification system architecture is described in section 3. Section 4 presents experimental results and section 5 draws some conclusions and presents some hints for future work.

2. HIGH SPEED IMAGE ACQUISITION: THE LRIS SYSTEM

Automatic systems for road pavement surface distress data acquisition and processing is an active research field. Despite the performance improvements of recent equipments, some problems still remain, for instance related with implementation costs, processing speed or accuracy [3]. In [8], INO presented a system, LRIS, capable of acquiring pavement surfaces images during road surveys at speeds that can surpass 100 km/h.

The LRIS system is composed by two high speed/high resolution linescan cameras (each one acquiring half road lane images) in conjunction with high power lasers, see Figure 1.

The cameras and the projectors are aligned in the same plane in a symmetrically crossed optical configuration. This configuration increases the visibility of very small cracks since

the incident illumination angle of the laser causes the projection of shadows in crack areas.

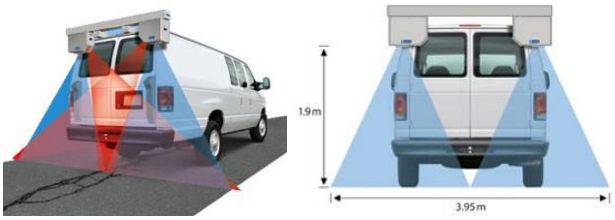

Figure 1: Schematics of LRIS system [8].

A very good contrast is provided between crack's and no crack's areas and, due to the high power lasers, the system can operate in full daylight, being immune to variations in outside lighting conditions and to shadows cast by road side objects, viaducts and the inspection vehicle itself [8]. Two samples of the LRIS image database (DB1), acquired during a road survey while driving at 70km/h, are shown in Figure 2. The images acquired by each sensor have a dimension of 4096×2048 pixels.

As an alternative, human observation is commonly used to gather information about pavement surface distresses, during road surveys made by inspectors. Usually, digital photos of defects are also taken during such surveys. Two samples of the human observation image database (DB2) considered in the scope of this paper, are shown in Figure 3. These images were captured along Portuguese roads, using a digital camera with its optical axis perpendicular to the road pavement surface and have a dimension of 2048×1536 pixels.

The images of both databases are processed in grayscale.

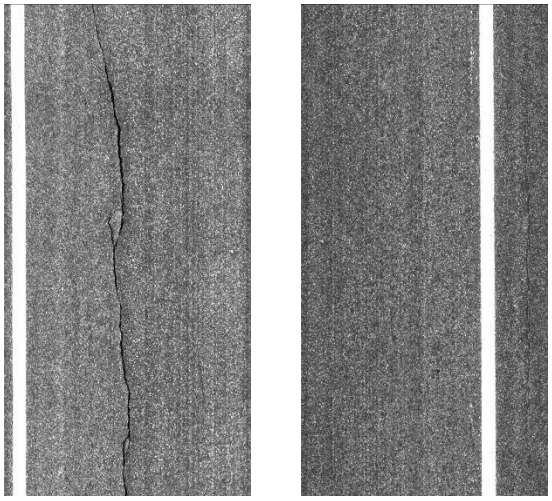

Figure 2: Two sample LRIS database original images (DB1) [8]. The left sensor image reveals a relevant longitudinal crack (left), while the right sensor image reveals the absence of defects (right).

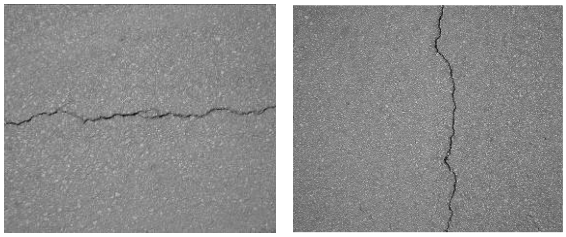

Figure 3: Two sample original images from the database acquired during a human observation survey (DB2).

3. SYSTEM ARCHITECTURE

In this paper, a simple unsupervised system for the automatic detection of cracks in images acquired during road surveys, and their classification into a predefined set of crack classes (defined according to the Portuguese Distress Catalogue [9]) is proposed, following the system architecture show in Figure 4. For each database (DB1 and DB2), different pre-processing strategies are adopted as explained in the next section. The remaining modules use the same strategy for both image databases.

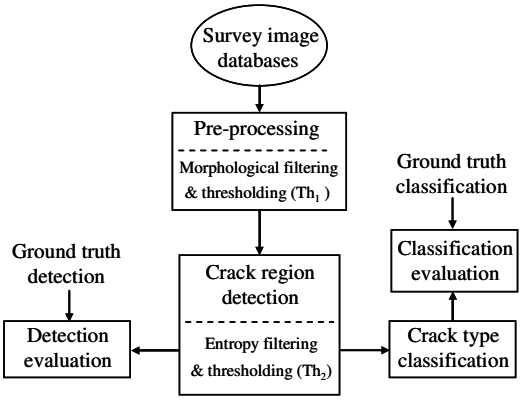

Figure 4: System architecture for crack detection and classification.

3.1 Pre-processing

Due to the different characteristics of DB1 and DB2 images, the pre-processing operations required in each case, to achieve the best possible crack detection results, are also different.

Analysing DB1, the image histograms are bimodal when containing road pavement cracks (crack pixels appear darker than non crack pixels when using INO hardware [8]), presenting one large mode at the middle and a smaller one near the origin, as cracks correspond to darker image areas – see the left histogram of Figure 5. This reveals that a threshold, can be used to separate the two modes (at approximately Th_1), segmenting the image into crack pixels and non crack pixels.

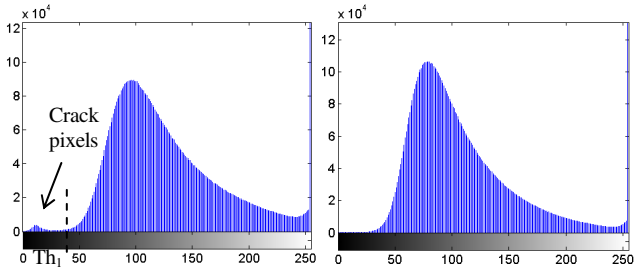

Figure 5: Histograms of the images show in Figure 2.

For images without the presence of crack pixels, the histogram shows a single mode, which can be modelled by a Rayleigh probability density function (*pdf*) – see the right side of Figure 5.

DB1 images present a high variance in pixel intensities, as can be observed in the plot shown in the left part of Figure 6, for row 945 of the left image of Figure 2. Nevertheless, this plots shows that a threshold could be used to segment almost all crack pixels. DB2 images present a lower pixel

intensity variance, as can be observed in the right plot of Figure 6. Again, a threshold can be used for image segmentation into crack pixels and non crack pixels, even if in this case the threshold value needs to be higher than the one used for DB1.

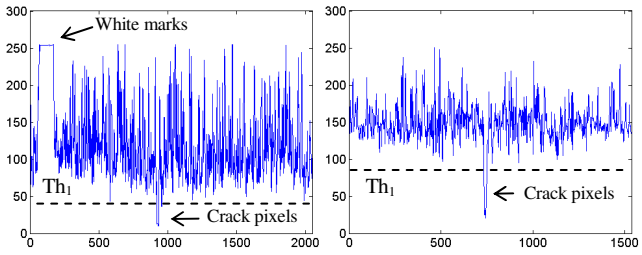

Figure 6: Pixel intensities for row 945 of Figure 2 left image (left) and for column 740 of Figure 3 left image (right).

Additionally, DB2 images present a non-uniform illumination (in opposition to what was observed for DB1), due to the digital camera falloff effect. To correct this, the pre-processing strategy described in [10] is applied to DB2 images. Images are divided into non-overlapping blocks of dimension 75 × 75 pixels, being normalized based on the block's average pixel intensities, without losing information about crack pixels. The effect of this normalization procedure is illustrated in Figure 7, where image blocks are adjusted to the same mean intensity value, except for those blocks possibly containing crack pixels.

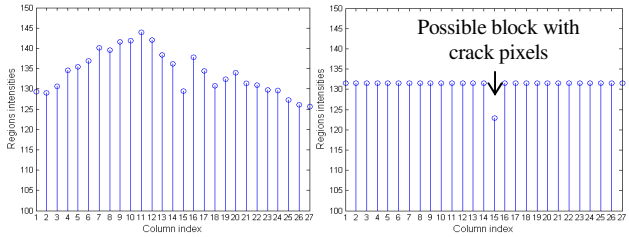

Figure 7: DB2 normalization procedure, reducing non-uniform background illumination effects. The contrast to blocks potentially containing crack pixels is enhanced.

After normalization of DB2 images, a common pre-processing stage for all images (of DB1 and DB2) is applied, to reduce the intensity variance observed in Figure 6:

$$\text{Img} \circ se = (\text{Img} \ominus se) \oplus se. \qquad (1)$$

where the symbol ∘ represents the morphological opening operation, ⊖ denotes the morphological erosion while ⊕ stands for the morphological dilation. *se* represents a disk-shaped structuring element with a radius of 5 pixels [11], empirically chosen as it provides good experimental results.

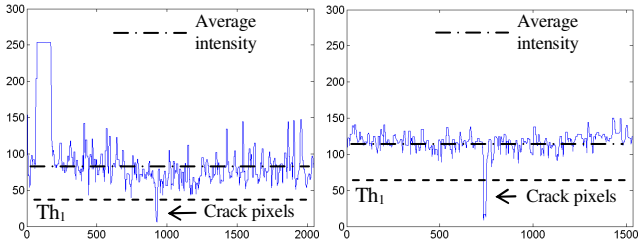

Figure 8: Pixel intensities corresponding to the case shown in Figure 6, after morphological opening operation.

Morphological opening results, show that the variance of pixel intensities is lower when compared with the original ones (Figure 6), as show in Figure 8. After applying morphological filtering to reduce pixel intensities variance, the threshold value (Th_1) needs to be computed.

Since DB1 and DB2 images usually present different mean pixel intensities, a further normalization step is applied. The aim is to change pixel intensities in each image, so that its average becomes equal to a predefined mean reference intensity value. From the research previously conducted by the authors, a gray level value of 125 seems to be a good average reference intensity value, being adopted here (this gray level value represents the average of pixel intensities computed using both image databases).

A dynamic threshold value, Th_1, unique for each image, is then computed according to the expression:

$$Th_1 = Th(Ot) - 0.5 \times std(\text{Img}) \qquad (2)$$

where *Th(Ot)* is the threshold value computed according to a modified Otsu method [12], using only the intensity levels lower than the mean intensity level for each image. This provides increased immunity to noise. *std(*Img*)* is the standard deviation of all image pixel intensities. The output of the thresholding operation assigns label '0' to pixels whose value is above the threshold Th_1, and '1' to potential crack pixels, those with intensity below Th_1.

3.2 Crack Detection

Each binary image obtained after applying the threshold Th_1 (for both database images), is divided into non-overlapping blocks of dimension 75 × 75 pixels. This dimension, empirically chosen, provides a good compromise between complexity and accuracy.

For each binary image block, its entropy $E_{binblock}$ [11] is computed according to the expression

$$E_{binblock} = \left| f_0 \times \log_2(f_0) + f_1 \times \log_2(f_1) \right|, \qquad (3)$$

where f_0 and f_1 are respectively the frequency of pixels labelled with '0' and '1'. The variation of the entropy function in terms of the number of block pixels labelled '1', is represented by the plot shown in Figure 9, for a generic binary image block of dimension 75 × 75 pixels. This plot was obtained considering the presence of a square of size *dimension* × *dimension* inside the block, to illustrate the entropy evolution. This square side *dimension* is represented in the horizontal axis of the plot.

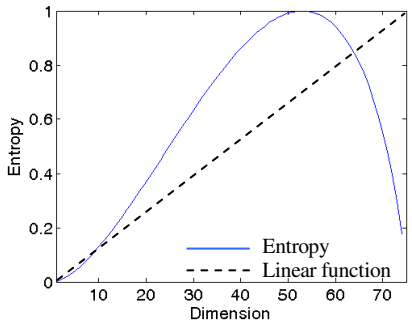

Figure 9: Entropy function, for a generic binary image block of dimension 75 × 75 pixels. The horizontal axis represents the side dimension of a square of ones completely inside the block.

As can be observed, entropy provides a fast measurement of crack pixels presence in an image block, being invariant to the position of crack pixels inside the block. Additionally, the presence of only a few cracks pixels in the block is made more evident than if a linear function would be used. The entropy is maximized when half the block pixels are cracks (i.e., when *dimension* takes value 53).
In order to classify image blocks as containing crack pixels or not, another thresholding operation (Th_2) is applied, now to the entropy blocks matrix:

$$Th_2 = 0.5 \times Th(Ot)_{binblock}, \quad (4)$$

where $Th(Ot)_{binblock}$ is a threshold computed using a modified Otsu method [12], as done for Th_1. Histograms of entropy blocks matrix computed for the left images show in Figure 2 and Figure 3 are show in Figure 10.

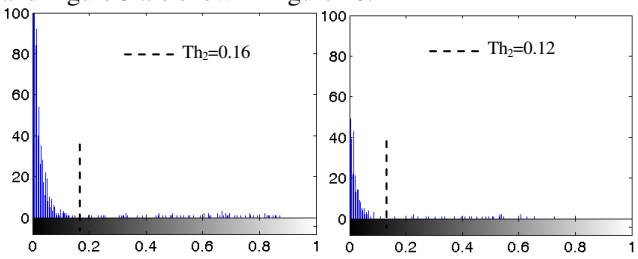

Figure 10: Histograms for the entropy blocks matrix computed for the left side images of Figure 2 and Figure 3. The computed Th_2 values are marked in both histograms.

After block entropy thresholding, the resulting isolated blocks labelled with '1' are removed. These blocks typically correspond to noise due, e.g., to oil stains in the pavement.

3.3 Crack Type Classification

After detection of regions with crack pixels, they can be classified into types. The crack types considered here follow the specifications of the Portuguese Distress Catalogue [9]: longitudinal (L), transversal (T) or miscellaneous (M). Crack type classification uses a classification system exploiting another 2D feature space according to the technique developed in [13]. A classification is assigned to each connected crack region identified during the detection phase.
The 2D feature space used is show in Figure 11. Crack classification uses the standard deviation of the column (feature one) and row (feature two) coordinates, of the detected crack regions, i.e., image regions labelled with '1' in the detection results matrix.

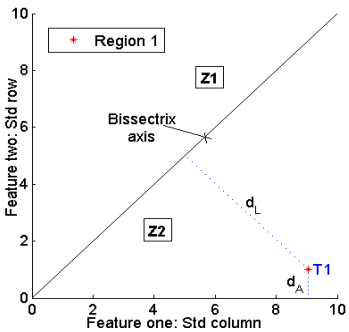

Figure 11: 2D feature space used for crack classification. T1 represents the crack shown in Figure 2, being classified as transversal.

Crack classification is performed by computing two distances: d_L and d_A, for each connected crack region, which is represented by a point in the 2D feature space.
A probability of a crack belonging to longitudinal or transversal classes is computed, according to:

$$P(y_i = c_{cr} \mid x_i) = 1 - \frac{d_{Ai}}{d_{Ai} + d_{Li}}, \quad (5)$$

and the probability of a crack belonging to the miscellaneous class is computed according to:

$$P(y_i = c_M \mid x_i) = 1 - \frac{d_{Li}}{d_{Ai} + d_{Li}}, \quad (6)$$

where *cr* is one of the class indexes T or L, d_{Ai} is the distance from point i to the nearest axis, d_{Li} is the distance from point i to the bisectrix and x_i is the observation i. The decision for a given crack type is made when the corresponding probability is greater than 0.5.

4. EXPERIMENTAL RESULTS

The proposed system for automatic crack detection and classification is evaluated using DB1 and DB2. For each test image the corresponding ground truth was manually obtained. The test subsets are composed by 20 and 56 gray-scale images for DB1 and DB2, respectively.
For the pre-processing module, a disk-shape structuring element with a radius of 5 pixels was used for all DB1 and DB2 images. The left images of Figure 12 show segmentation results obtained for the left images of Figure 2 and Figure 3. The right images of Figure 12 represent the corresponding, manually obtained, ground truth results.
The ground truth information is used to evaluate the system performance (see Table 1), by computing a global error-rate (classification error for the detection of regions with and without crack pixels), a crack error-rate (1 minus the Recall value), Precision (*pr*), Recall (*re*) and a Performance Criterion (PC) metric, reflecting the overall classifier performance:

$$pr = \frac{\text{Number of regions correctly classified as cracks}}{\text{Total number of crack regions detected}}. \quad (7)$$

$$re = \frac{\text{Number of regions correctly classified as cracks}}{\text{Total number of crack regions (ground truth)}}. \quad (8)$$

$$\text{PC} = \frac{2 \times pr \times re}{pr + re}. \quad (9)$$

For comparison purposes, the two bottom lines of Table 1 show the results of processing both image databases using the crack detection methodology described in [10], where a pattern recognition system based on a Gaussian quadratic classifier was considered. DB1* values were computed in the scope of this work, but using the methodology of [10], while DB2* results are the ones reported in Table 1 of [10].
The evaluation results of Table 1 show that the proposed methodology achieves better Precision results than the technique reported in [10] for both image databases (84.0% against 59.3% for DB1 and 95.1% against 92.5% for DB2). Also the overall system perform (PC) is better, using the proposed methodology, for both image databases (89% and 95% against 60.5% and 94.7% for DB1 and DB2, respectively).
In terms of Recall (viewed as the most important metric for

this type of application, where missing crack areas must be more penalised), the proposed methodology achieves a significantly better value for DB1 (94.8% against 61.7%). Although the results for DB2 are not better than those described in [10] (95.6% against 97.0%), the gain in system robustness leads to the conclusion that the proposed methodology's global performance is quite good.

In terms of crack classification, 100% recall and precision were obtained for all classes of detected cracks, which reveals a very good overall system performance.

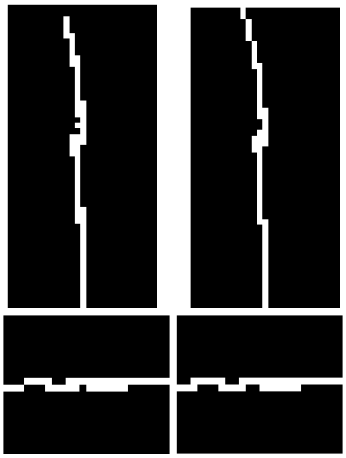

Figure 12: Detection results (left) for the left side images of Figure 2 and Figure 3 and the corresponding ground truth (right).

Table 1: Crack detection results for the proposed method (top two rows) and for the method proposed in [10] (bottom two rows).

Image DB	Global Error-rate	Crack Error-rate	*pr*	*re*	PC
DB1	0.26%	5.3%	84.0%	94.8%	89.0%
DB2	0.02%	4.5%	95.1%	95.6%	95.0%
DB1*	2.95%	38.3%	59.3%	61.7%	60.5%
DB2*	0.64%	3.0%	92.5%	97.0%	94.7%

Moreover, the proposed methodology presents faster processing times, when compared to those reported in [10]. Using the same hardware and software platforms, the proposed system takes 5 seconds/image against the 31 seconds reported in [10], for DB1 images, and 3 seconds/image against 18 seconds for DB2 images.

5. CONCLUSIONS AND FUTURE WORK

In this paper a simple unsupervised system for crack detection and classification into several classes is proposed, achieving promising performance results. It was demonstrated that the present proposal is able to deal with images acquired using distinct imaging sensors, showing an interesting robustness, which is reflected in the similar recall values obtained (94.8% and 95.6%, for DB1 and DB2, respectively). Moreover, a faster processing was also achieved with the proposed methodology when compared to the one developed in [10].

Future developments will consider the usage of additional filtering techniques to further reduce the variance of pixel intensities in road surveys image databases. Also additional work towards a more accurate dynamic thresholding method will be conducted, to further improve the crack detection results. Additionally, the different features used in both methodologies could be combined, also envisaging the improvement of crack detection results.

6. ACKNOWLEDGEMENT

The authors acknowledge the support from Fundação para a Ciência e Tecnologia (FCT) and also from INO [8], represented by Mr. John Laurent, for allowing the usage of the image database here labelled as DB1, for non profit research work

REFERENCES

[1] D. Meignen, M. Bernadet and H. Briand, "One application of neural networks for detection of defects using video data bases: identification of road distresses", *Proc. 8th Int. Workshop on Database and Expert Systems Applications*, 1-2 September, pp. 459 – 464, 1997.

[2] H. D. Cheng and M. Miyojim, "Automatic Pavement Distress Detection System", *Journal of Information Sciences 108*, ELSEVIER, pp. 219-240, 1998.

[3] Y. Huang and B. Xu, "Automatic Inspection of Pavement Cracking Distress". *Journal of Electronic Imaging. Society for Imaging Science and Technology*, 2006.

[4] K. Wang, "Design and Implementations of Automated Systems for Pavement Surface Distress Survey", *Journal of Infrastructure Systems*, ASCE, Vol. 6, 2000.

[5] J. Chou and W. A. O'Neill and H. Cheng, Pavement distress classification using neural networks, in *IEEE International Conference on Systems, Man, and Cybernetics*, p. 397-401, vol.1, 1994.

[6] S. Chambon, P. Subirats and J. Dumoulin. Introduction of a wavelet transform based on 2D matched filter in a Markov Random Field for fine structure extraction: Application on road crack detection, in *IS&T/SPIE Electronic Imaging - Image Processing: Machine Vision Applications II*, San Jose, USA, 2009.

[7] H. G.Zhang and Q. Wang, "Use of Artificial Living System for Pavement Distress Survey", *Proc. 30th Annual Conference of the IEEE Industrial Electronics Society*, Korea, Vol.3, pp. 2486-2490, 2004.

[8] http://www.ino.ca

[9] JAE*, Catálogo de Degradações dos Pavimentos Rodoviários Flexíveis – 2ª Versão,* Ex-Junta Autónoma das Estradas, Portugal, 1997.

[10] H. Oliveira and P. L. Correia, "Supervised Strategies for Crack Detection in Images of Road Pavement Flexible Surfaces", *Proc. 16th European Signal Processing Conference (EUSIPCO)*, Lausanne, Switzerland, 25-29, 2008.

[11] R. Gonzales and R. Woods, *Digital Image Processing 3th edition*, Pearson International Edition, USA, 2008.

[12] L. Dong, G. Yu, P. Ogunbona and W. Li, "An efficient iterative algorithm for image thresholding", *Pattern Recognition Letter*, Vol. 29, issue 9, pp. 1311-1316, 2008.

[13] H. Oliveira and P. L. Correia, "Identifying and retrieving Distress images from road pavement surveys", Proc. 1st ICIP Workshop on Multimedia Information Retrieval: new trends and challenges, ICIP 2008, Las Vegas, USA, 2008.