Road Surface Crack Detection using a Light Field Camera

David Fernandes
Instituto de Telecomunicações
Instituto Superior Técnico
Lisboa, Portugal
davidluisdias@tecnico.ulisboa.com

Paulo Lobato Correia
Instituto de Telecomunicações
Instituto Superior Técnico
Lisboa, Portugal
plc@lx.it.pt

Henrique Oliveira
Instituto de Telecomunicações
Instituto Politécnico de Beja
Portugal
hjmo @lx.it.pt

Abstract—During traditional road surveys, inspectors capture images of pavement surface using cameras that produce 2D images, which can then be automatically processed to get a road surface condition assessment. This paper proposes a novel crack detection system that uses a light field imaging sensor, notably the Lytro Illum camera, instead of a conventional 2D camera, to capture road surface light field images. Light field images capture the light rays originating from different directions, thus providing a richer representation of the observed scene. The proposed system explores the disparity information, which can be computed from the light field, to obtain information about cracks observable in the pavement images. A simple processing system is considered, to show the potential use of this type of sensors for crack detection. Encouraging experimental crack detection results are presented based on a set of road pavement light field images captured over different pavement surface textures. A performance comparison with a state-of-the-art 2D image crack detection system is included, confirming the potential of using this type of sensors.

Keywords—Light field imaging, road crack detection, image processing.

I. INTRODUCTION

Due to their constant usage, road pavement surfaces exhibit distresses and periodical road surveys are scheduled, to collect data about pavement condition that support the decision for the type and scheduling of appropriate maintenance activities. During road surveys, inspectors usually collect pavement surface images using 2D cameras, which can be automatically processed to detect distresses, with cracks being the most commonly found pavement surface defects.

Road crack detection methods, that mainly process conventional 2D images, can be grouped into the following categories in terms of the image characteristics: (i) photometric; (ii) geometric; (iii) a combination of both [1]. For instance, techniques based on wavelets [2] or autocorrelation filtering [3] explore some of the photometric or geometric characteristics of road surface images, while other model analysis strategies [4] use the combination of both characteristics. More recent methods typically employ conventional machine learning techniques [5] [6].

As an alternative to conventional 2D cameras, light field cameras are emerging as powerful imaging sensors, capturing not only the average light intensity at a pixel, but rather the contributions of light rays emanating from different directions that reach the camera sensor. The resulting images are thus richer and open new possibilities for image processing and analysis [7].

This paper proposes a new crack detection strategy, by exploring light field images captured with a Lytro Illum, to improve and complement the automatic detection of cracks in images of road pavement surface, notably by exploring the disparity between the different viewpoints available from the captured imagery. Obtained results favorably compare with those achieved by the analysis of conventional 2D images.

This paper is organized as follows. Section II introduces light field imaging and section III describes the proposed automatic crack detection system. Experimental results are presented in section IV, including the comparison with a system that processes conventional 2D images. Section V presents conclusions and discusses some hints for future work.

II. LIGHT FIELD IMAGING

Light field imaging allows capturing high-dimensional data sets, including information about light rays coming from different directions, thus providing richer scene representations than conventional 2D images.

The plenoptic function has been introduced in 1991 by Adelson and Berger, to describe the information carried by light rays at every point in 3D space [8]. A light ray of wavelength λ will reach a point (x, y, z) in space, travelling from a direction described by an azimuth / orientation pair (Θ, Φ), at a given instant *t*. The resulting plenoptic function is therefore a 7D function of $(x, y, z, \Theta, \phi, \lambda, t)$, as illustrated in Figure 1 (left).

Since the processing of this 7D function is computationally heavy, some simplifications can be considered to reduce its dimensionality. Of particular interest is the 4D light field parametrization of the plenoptic function. The light rays can then be described as emanating from a point *(u,v)* on a U-V plane, closer to the observed scene, to a point *(s,t)* in a S-T plane, closer to the camera, at a given time instant and with each ray keeping its RGB radiance. This leads to a static 4D light field representation *L*(*u,v,s,t*), as illustrated in Figure 1 (right).

ISBN 978-90-827970-1-5 © EURASIP 2018

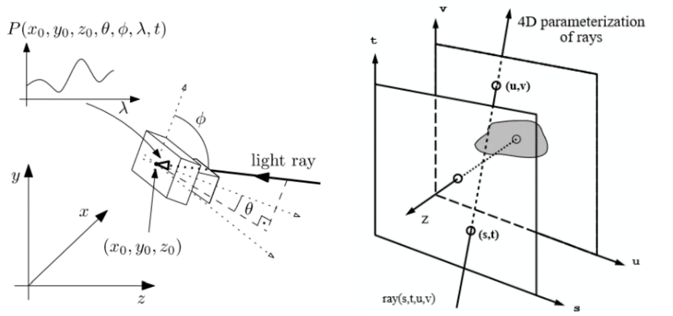


Figure 1 – Visual representation of the plenoptic function [9] (left) and the representation of the two-plane parametrization [10](right).

This paper considers a light field acquired by a lenslet camera, which includes an array of microlenses, each one playing the role of a small camera and allowing the acquisition of a *micro-image* that corresponds to a portion of the available visual information. From these micro-images it is possible to reconstruct 2D sub-aperture images, corresponding to different viewpoints of the scene, thus forming a multi-view array (of 2D images) that can be explored for the detection of cracks in road pavement surfaces.

III. Proposed System

The architecture of the proposed road pavement surface crack detection system using a light field camera (here designated by the ***Light Field Crack Detection*** system – LFCD) is presented in Figure 2.

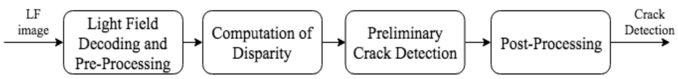


Figure 2 – Proposed LFCD system architecture.

Since the light field (LF) camera captures light rays travelling in different directions, a richer visual scene representation is obtained, which provides additional cues, useful for the automatic detection of cracks in images of pavement surface. More specifically, the proposed LFCD system allows recovering a multi-view array of 2D sub-aperture images, from which the disparity can be exploited for improving crack detection. The main components of the proposed system are described in the following.

A. Light Field Decoding and Pre-Processing

After the acquisition of a light field image, it can be decoded from the Bayer pattern into a 15×15 array of 2D sub-aperture RGB images $M(i,j)$, for instance using the Matlab Light Field Toolbox v 0.4 [11]. Each of the resulting 2D images has a dimension of 435×625 pixels and it represents a slightly different viewpoint of a scene. An example of a recovered 2D sub-aperture pavement surface image is shown in Figure 3.

The subsequent pre-processing converts each 2D RGB image to gray scale, as the color information is not particularly useful in the crack detection problem. Additionally, since cracks are composed of darker pixels, those pixels with intensity values higher than a predefined threshold *th* see their value replaced by the *th* value, thus leading to a reduction of image intensities variance. For the road images dataset considered in this paper, *th* was empirically set to 50, after exhaustive testing.

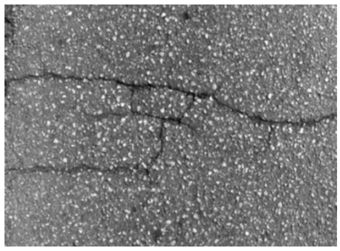


Figure 3 – Example of a sub-aperture image, in this case the central sub-aperture image $M(8,8)$, whose position in the multi-view array is depicted in blue in Figure 4.

B. Computation of Disparity

To explore the disparity captured by the LF images, a selection of 2D sub-aperture images is considered, and the disparity between image pairs, at symmetrical positions regarding the central sub-aperture image, is computed.

Figure 4 illustrates how the horizontal ($dimh$) and vertical ($dimv$) disparity images are computed, namely by subtracting the multi-view array images of column 3 with those of column 13, represented in green in Figure 4. The individual differences are added up according to equation (1). A similar procedure is applied to the two sets of sub-aperture images represented in red in Figure 4, according to equation (2). The obtained disparity images highlight the positions where road cracks are more likely to be found, making them more salient and easier to detect.

$$dimh = \sum_{i=3}^{13}[M(i,3) - M(i,13)] \quad (1)$$

$$dimv = \sum_{j=3}^{13}[M(3,j) - M(13,j)] \quad (2)$$

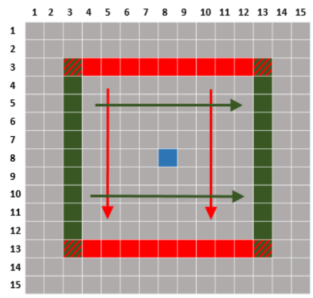


Figure 4 – Illustration of sub-aperture image selection for the computation of disparity images.

C. Preliminary Crack Detection

In the preliminary crack detection module, a simple and fast algorithm is applied to enhance the visibility of crack areas in images. For this purpose, an edge detector is applied to the horizontal and vertical disparity images, $dimh$ and $dimv$ respectively, and the results are added to include the edges found in both orientations. Along with edge information due to presence of road cracks, also other small pavement irregularities (typically corresponding to false detections) are enhanced and appear as undesired detection results in the edge image. To reduce this effect, a 5×5 median filter is applied. A sample result is illustrated in Figure 5.

ISBN 978-90-827970-1-5 © EURASIP 2018

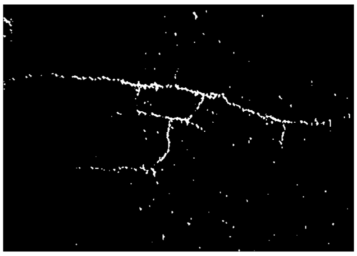


Figure 5 – Example of a preliminary crack detection result.

D. Post Processing

The goal of the post-processing module is to reduce possible false detections resulting from the preliminary crack detection, by applying a region filter, which identifies image areas more likely to correspond to cracks. After that, a connected component analysis is performed, aiming the reduction of those possible false detections, as represented in the block diagram of Figure 6.

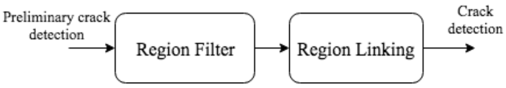
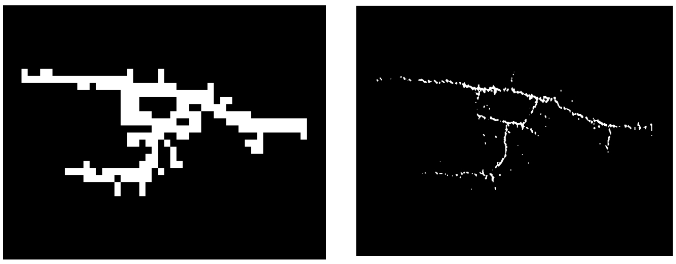


Figure 6 – Detail of the post-processing module.

The region filter considers a set of small circular regions with 16 pixel diameter (empirically chosen after exhaustive testing), covering the edge image in such a way to ensure a 25% overlap with neighbors. It then assigns label '1' to circles with more than 50% of edge pixels, and label '0' otherwise. The overlapping is useful to maintain the continuity of existing crack regions.

Connected components are identified in the region filter output image, and very short connected components are excluded as they are very likely to correspond to false crack detections.

These identified connected components form a mask that covers the areas where relevant cracks are located, as illustrated in Figure 7 (a). That mask can then be used to exclude false crack detections obtained in the preliminary crack detection, leading to better detection results at pixel level, as illustrated in Figure 7 (b).



(a) *(b)*

Figure 7 – Region filter: (a) mask used to select crack areas, (b) pixel-based crack detection results after applying the mask.

To further improve the crack detection, a morphological closing filtering is applied, to remove small holes.

Then, candidate crack regions are found by grouping crack pixels using a region linking algorithm [12]. Relevant connected components are identified by fulfilling a set of three geometric requirements: (i) more than 70% of eccentricity for an ellipse fitted to it; (ii) ellipse major axis longer than 15 pixels; (iii) width of at least 1 mm. The remaining connected components (the non-relevant candidates) are considered for removal, unless they are spatially linked to relevant components, in which case the region linking algorithm keeps them labeled as cracks, thus improving the crack detection result. An example of the region linking module output is shown in Figure 8.

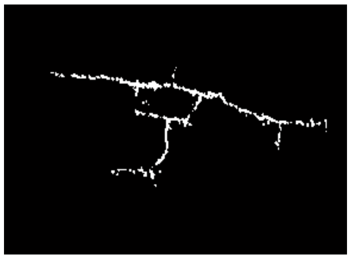


Figure 8 – Region linking results for the image of Figure 7 (b).

IV. EXPERIMENTAL EVALUATION

This section presents the strategy used to evaluate the LFCD system performance on detecting cracks, starting by briefly introducing a benchmarking system that explores the usage of 2D images. After that, a description of the dataset used, as well as the metrics considered to perform the evaluation procedure, are presented.

A. 2D Benchmark Crack Detection – The CrackIT System

As a benchmark for comparing the performance of the proposed LFCD system, an existing 2D crack detection system, known as CrackIT, which fully incorporates the methodology described in [12], was considered. This system was selected as it is a good representative of the state-of-the-art, and it has an implementation that is publicly available at http://amalia.img.lx.it.pt/CrackIT/ [13].

The architecture of the considered benchmarking solution is illustrated in Figure 9, and briefly described in the following.

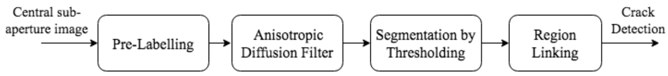


Figure 9 – CrackIT system architecture.

The CrackIT system uses a 2D pavement surface image as input and the processing starts by assigning a 'crack' or 'non-crack' label to each non-overlapping image block, based on the mean and standard deviation of pixel intensity values in each block – the pre-labelling, as detailed in [5]. After that, an anisotropic diffusion filter [12] is applied to deal with random textures present in pavement surface images, followed by an image intensity normalization technique to reduce the non-uniform background illumination, with these two aspects being typically observed in images captured by traditional 2D cameras. Then, a segmentation by thresholding [12] is applied

ISBN 978-90-827970-1-5 © EURASIP 2018

to obtain the initial crack detection result. As a last step, the same region linking procedure mentioned in section III-D is applied to obtain the final crack detection result.

B. Evaluation Dataset

As mentioned before, the Lytro Illum lenslet light field camera has been used to acquire the image dataset considered in this work. The acquisition setup considered the camera positioned at 1 meter from the pavement surface, with its optical axis perpendicular to the road pavement surface. All the images were collected during a manual survey. The dataset for testing is composed of 13 light field images, captured from pavements with different surface texture characteristics. For each image, a multi-view array was computed and each of the recovered 2D sub-aperture images presented a spatial resolution of 1 mm^2 per pixel.

To be able to compare the proposed system against the 2D benchmarking, the 2D central sub-aperture image (corresponding to the blue square in Figure 4) is used as input to the CrackIT system, as it corresponds to what a traditional 2D camera would capture.

C. Evaluation Metrics

To evaluate the crack detection results, a block-based ground truth was created by an expert, who manually labeled each 25×25 pixel block of each image, with the help of a graphical user interface developed for that purpose, as containing cracks or not.

Since for some blocks only a very small number of crack pixels were included, the expert sometimes labeled those blocks as not containing cracks. As a consequence, and in order to have a fair evaluation of the automatic results produced by the algorithms, the block-based evaluation done in this paper assumes that a block will only be classified as containing a crack if the number of detected crack pixels in that block is at least 4% of the block size (25 pixels).

While comparing the detected crack blocks with the ground truth, the following block-based evaluation metrics were taken in to account:

- True Positives (TP) – Blocks that belong to cracks;
- False Positives (FP) – Blocks falsely detected as containing cracks;
- False Negatives (FN) – Missed blocks that belong to cracks.

This evaluation is summarized using the well-known metrics: recall (re), precision (pr) and f-measure (Fm), according to the following equations:

$$re = \frac{TP}{TP + FN} \tag{3}$$

$$pr = \frac{TP}{TP + FP} \tag{4}$$

$$Fm = \frac{2 \times TP}{(2 \times TP)+FN+FP} \tag{5}$$

V. Experimental Results

The evaluation of the proposed LFCD system is presented in the following, in comparison with the 2D system considered for benchmarking, the CrackIT system [5].

Table 1 and Figure 10 summarize the results achieved using the considered performance metrics for all the test images. From these results it is possible to observe that the usage of light field imaging, even when only applying simple image processing techniques, seems to provide a better crack detection result than using the more conventional 2D crack detection system CrackIT – the *Fm* metric results increase, in average, from 79% to 85%. Additionally, when doing a visual evaluation of the results (a subjective analysis), by looking at the identified cracks, it seems that by exploring the light field disparity information, it was possible to obtain a better definition of the cracks present in the pavement surface test images, as can be seen by the quality of the results shown in Figure 11.

Table 1 – Precision, recall and Fm results for the LFCD and CrackIT systems.

Image	*re* (LFCD)	*re* (CrackIT)	*pr* (LFCD)	*pr* (CrackIT)	*Fm* (LFCD)	*Fm* (CrackIT)
1	100%	98%	100%	100%	100%	99%
2	100%	100%	100%	100%	100%	100%
3	93%	100%	97%	73%	95%	85%
4	86%	86%	93%	42%	89%	56%
5	95%	100%	75%	73%	84%	85%
6	96%	96%	74%	61%	84%	74%
7	91%	69%	77%	94%	83%	80%
8	75%	90%	94%	47%	83%	62%
9	71%	97%	91%	90%	80%	93%
10	70%	100%	93%	69%	80%	82%
11	82%	90%	74%	67%	78%	77%
12	81%	72%	72%	80%	77%	76%
13	80%	76%	61%	46%	69%	58%
Average	85%	90%	85%	72%	**85%**	**79%**

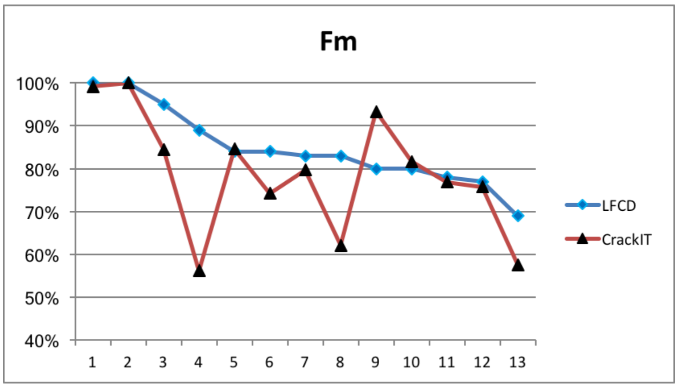
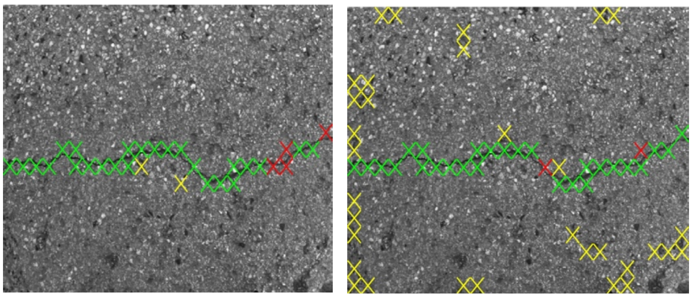


Figure 10 – Comparison of LFCD and CrackIT performance using the f-measure.

Figure 11 shows block detection sample results obtained from the proposed LFCD and the CrackIT systems, where blocks marked in green represent the ones correctly classified as containing cracks, the yellow ones denote those blocks detected by the systems but not manually classified as cracks, and the red ones correspond to blocks manually classified as cracks but not detected by the systems.



(a) (b)

Figure 11 – Block detection results for image 4 of the dataset: (a) proposed light field system, $re = 86\%$, $pr = 93\%, Fm = 89\%$. *(b) CrackIT system,* $re = 86\%$, $pr = 46\%, Fm = 56\%$.

For the sample result presented in Figure 11, the proposed LFCD system showed a greater capacity to deal with the effects of the pavement surface texture, discarding most of the false positive detections presented in the CrackIT results, while detecting most of the blocks containing cracks, hence achieving a higher value of the precision metric. In fact, the LFCD system seems to be more robust to deal with the different pavement surface textures found in images, as shown by the smaller variations of the *Fm* metric in Figure 10 thus outperforming the CrackIT system in this respect. By looking at the precision metric results listed in Table 1, the LFCD system seems to perform significantly better for the entire dataset, meaning that the proposed system produces less false positives.

Looking at the recall metric the behavior of the two systems, it is more similar, with both systems being able to capture the more relevant crack information observable from the images.

Overall, the *Fm* metric (see Table 1 and Figure 10) shows the advantage of the proposed LFCD system.

As an alternative to the consideration of disparity information, as used by LFCD, one might consider computing a depth image of the pavement, computed using the light field information. However, the initial experiments using the Lytro software suit to create a depth map, did not provide very helpful information for crack detection.

VI. Conclusions

This paper presented a novel crack detection system, exploring for the first time a light field imaging sensor. Although the crack detection system presented has room for improvements, it allowed to show that exploring the disparity information available from the light fields can bring benefits for crack detection. In the experiment, it was noted that this disparity information is more abundant in the regions close to the center of each processed image.

The proposed LFCD system performed well when compared to the 2D-based CrackIT system, considered for benchmarking purposes. The experimental work has shown that the additional information available in light field images can be favorably explored for crack detection.

Future work involves considering a more sophisticated image processing solution, as well as the acquisition of a larger image dataset for testing.

Acknowledgment

This work has been partially supported by Instituto de Telecomunicações under Fundação para a Ciência e Tecnologia, Grant UID/EEA/ 50008/2013.

VII. Refrences

[1] S. Chambon and J.-M. Moliard, "Automatic Road Pavement Assessment with Image Processing: Review and Comparison," *International Journal of Geophysics,* vol. 2011, 2011.

[2] K. Wang, Q. Li and W. Gong, "Wavelet-Based Pavement Distress Image Edge Detection with À Trous Algorithm," *Transportation Research Record: Journal of the Transportation Research Board,* vol. 2024, pp. 73-81, 2007.

[3] J. Shirataky and T. Tomikawa, "A Study of Road Crack Detection by Image Processing," *Part B. Science and Technology,* vol. 24, pp. 67-71, 2000.

[4] S. Chambon, "Detection of Points of Interest for Geodesic Contours : Application on Road Images for Crack Detection," *International Joint Conference on Computer Vision Theory and Applications VISAPP,* 2011.

[5] H. Oliveira and P. Correia, "Automatic road Crack Detection and Characterization," *IEEE Transactions on Intelligent Transportations Systems.,* vol. 14, pp. 155-168, 2013.

[6] Y. Shi, L. Cui, Z. Qi, F. Meng and Z. Chen, "Automatic Road Crack Detection Using Random Structured Forests," *IEEE Transactions On Intelligent Transportation Systems,* vol. 17, pp. 3434-3444, 12 December 2016.

[7] J. Restrepo and L. Mignard-Debise, "Principles of Light Field Imaging," *IEEE Signal Processing Magazine,* 2 September 2016.

[8] E. Adelson and J. Bergen, "The Plenoptic Function and the Elements of Early Vision," *Computation Models of Visual Processing,* pp. 3-20, 1991.

[9] F. Pereira, E. A. da Silva and G. Lafruit, Plenoptic Imaging: Representation and Processing, Academy Press Library in Signal Processing, 2014.

[10] S. J. Gortler, R. Grzeszczuk, R. Szeliski and M. F.Owen, "The Lumigraph," Microsoft Research, 1996.

[11] D. G. Dansereau, "Light Field Toolbox for Matlab v0.4 [Online].," 2016. [Online]. Available: http://www.mathworks.com/matlabcentral/fileexchange/49683-light-field-toolbox-v0-4. [Accessed on 2017].

[12] H. Oliveira and P. Correia, "Automatic Crack Detection on Road Imagery Using Anisotropic Diffusion and Region Linkage," in *Proc. of 18th European Signal Processing Conference*, 2010.

[13] H. Oliveira and P. Correia, "CrackIT – An Image Processing Toolbox For Crack Detection and Characterization," in *IEEE International Conf. on Image Processing*, Paris, 2014.