Multi-resolution approach for fine structure extraction : Application and validation on road images

Nicolas Coudray, Argyro Karathanou, Sylvie Chambon

HAL Id: hal-00614116

https://hal.archives-ouvertes.fr/hal-00614116

Submitted on 9 Aug 2011

MULTI-RESOLUTION APPROACH FOR FINE STRUCTURE EXTRACTION

Application and validation on road images

Nicolas Coudray, Argyro Karathanou

Laboratoire MIPS, Université de Haute-Alsace, Mulhouse, France

nicolas.coudray@uha.fr,e0602532@uha.fr

Sylvie Chambon

Laboratoire Central des Ponts et Chaussées, LCPC, Nnates, France

chambon@lcpc.fr

Keywords: Segmentation, multi-resolution, road images, cracks.

Abstract: In the context of fine structure extraction, this paper presents a new method based on multi-resolution segmentation applied for the detection of road cracks. A method already developed to detected low-contrasted biological membranes has been adapted to detect cracks on images: crack features are defined as heterogeneities rather than transitions of closed regions characterizing the membranes. This new methodology is quantitatively validated on reference segmentations and compared to an adapted filtering and Markovian modelling algorithm.

1 INTRODUCTION

Detecting fine structures is very helpful in a lot of domains: to extract ceramic damages (Elbehiery et al., 2005), to find cracks in underground pipes (Iyer and Sinha, 2005), to detect road network in satellite images (Geman and Jedynak, 1996), to follow vessels in medical images (Frangi et al., 1998). Since 1990, a lot of algorithms have been proposed in the domain of crack detection on road pavement surface. In fact, every country needs to evaluate, periodically, the quality of roads and most of this work is done manually, which is expensive, non reproducible, dangerous and not very efficient. In consequence, a lot of efforts have been made in the field of research on automatic or semi-automatic procedures for detecting deterioration on the roads and, in particular detecting cracks. In 2003, the report of Schmidt (Schmidt, 2003) gives a good summary of the technologies developed in this field. The main difficulty of this task is that the default is not well contrasted compared to the texture of the road, and that it represents a very small part of the images (about 1.5% of the image).

The most recent algorithms proposed to tackle this problematic are based on multi-resolution approaches. Even if they are the most efficient, they have shown their limits by giving a sparse detection. In the field of detection of biological membranes, recent methods based on multi-resolution watershed segmentation provided great performances, even if the objects of interest are low-contrasted (Coudray et al., 2007). By analogy, the problem of biological membranes detection is similar to crack detection. The goal of this paper is to adapt this watershed method to road crack detection, and compare its performances with a Markovian modeling-based algorithm developed specifically for this purpose (Chambon et al., 2010).

Firstly, a brief state of the art of fine structure extraction is exposed. Secondly, the studied crack algorithms are presented: the Markovian modeling algorithm and the multi-resolution segmentation. Thirdly, the protocol used to analyze the performances is given. And finally, before concluding, results are summarized and discussed.

2 FINE STRUCTURE EXTRACTION

We distinguish four methods: those based on a THRESHOLD (Koutsopoulos and Downey, 1993), methods combining a threshold with mathematical morphology tools (MORPHOLOGY) (Tanaka and Uematsu, 1998), NEURAL NETWORK-based methods (Bray et al., 2006), and MULTI-SCALE filter-

ing (Subirats et al., 2006). A detailed state of the art can be found in (Chambon et al., 2010).

THRESHOLD methods are popular and simple but inefficient (the amount of false positives and of false negatives is important). Methods based on MOR-PHOLOGY reduce the number of false positives but they are strongly dependent on the choice of the parameters. NEURAL NETWORK techniques are efficient but they need a learning step that is quite expensive and non-automatic. Finally, MULTI-SCALE methods seem to be the most efficient: unlike other methods, they lead to less false positives, but more false negatives. In conclusion, we can use methods that detect most of the crack but give a lot of errors or methods that obtain only a part of the crack with low errors.

In the field of microscopy images, a watershed segmentation has been proposed to extract biological membranes (Coudray et al., 2007). The problem of this task is comparable to the problem of road crack detection: edges of the membranes (here the default) are low and heterogeneously contrasted compared to the background (the road). Consequently, we propose to adapt this method to the road application.

3 CRACK DETECTION ALGORITHMS

In this part, we briefly present a recent crack detection algorithm based on an adapted filter combined with Markovian modeling segmentation and, then, we introduce a new method, based on multi-scale analysis with watershed and we describe all the contributions in order to adapt this method to a road crack detection.

3.1 Adapted Filtering and Markovian Modeling (AFMM)

The method proposed in (Chambon et al., 2010) uses these two hypotheses, well known in the field of road crack detection:

(H_1) Pixels of a crack are darker than the background (the road);

(H_2) A crack is a set of connected segments with different orientations.

Consequently, the principle of the method is, first, to binarize the image (by adapted filtering) and, second, to refine the binarization by a segmentation based on Markovian modeling. At the beginning, we have to choose the number of scales, which depends on the resolution of the image (the size of the pixels): if we suppose that we have 1 mm/pixel, 5 scales are sufficient (we can detect cracks from 2 mm to 1 cm).

Then, we also have to choose the number of directions for the adapted filtering and, considering the aspect of the road cracks, it is realistic to use these directions: $[0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}]$. Interested readers can find more details about this work in (Chambon et al., 2010). The goal of this paper is to compare this method with the new one presented below.

3.2 Multi-Resolution Detection of Heterogeneities (MRDH)

This method is inspired from the 3-step segmentation method developed for the segmentation of low and heterogeneously contrasted biological membranes (Coudray et al., 2007). The aim was to partition the image by detecting the closed contours of membranes. In this initial algorithm, the edges are firstly thresholded at different scales, and, secondly, the resulting binary images are combined in a Reconstructed Gradient-Like (RGL) image. Finally, the watershed algorithm is applied on the RGL image to obtain the searched partition. For the cracks, the third step has been adapted and a fourth step has been added to discard spurious detections. The steps are detailed below.

A multi-resolution transform is first applied to the initial image $I(x,y,1)$ using a pyramidal transform: images $I(x,y,s)$ are obtained at scale s using:

$$I(x,y,s) = \sum_{m=-k}^{+k} \sum_{n=-k}^{+k} G(m,n,s)I(sx+m,sy+n,1), \quad (1)$$

where $G(m,n,s)$ is the smoothing average filter. Dyadic pyramids reduce the size of the image by a factor of 2 between each scale. Here, non-dyadic pyramids are used to better represent and identify the features, with $s \in V_s = \{1,2,\ldots,10\}$ (at scales coarser than 10, the images become too small to be analyzed). Heterogeneities are detected and thresholded at each scale, leading to binary images $B(x,y,s)$. In this study, features to be identified, i.e. the cracks, are better described as local heterogeneities than local transitions (e.g. Figure 1), making a local standard deviation filter more appropriate than a gradient filter like Sobel, for instance.

Figure 1: Initial image example ($I(x,y,1)$) – The crack corresponds to a local low-contrasted heterogeneity.

In the next step, binary images are combined to form the RGL image $I_{RGL}(x,y)$:

$$I_{RGL}(x,y) = \max_{s \in V_s} \left((11-s) \times B(x,y,s) \right). \quad (2)$$

With the $(11-s)$ weighting, the gray-level on the RGL image corresponds to the finer scale at which the feature has been identified. Details identified with a better precision have therefore a higher value (e.g. Figure 2).
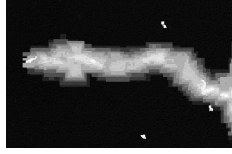

Figure 2: RGL image of Figure 1 ($I_{RGL}(x,y)$) – The brighter the pixels, the higher the precision of the detection.

The next step aims to thin the RGL image features to obtain 1-pixel wide objects, and has been adapted from our initial work to this road application. It needed to be modified since, contrary to membrane edges, cracks do no form closed regions: therefore, we can either apply the watershed algorithm after considering the background of the RGL image as seeds (i.e. pixels which value is 0 and which touch $non-0$ value pixels), or by applying a gray-scale thinning algorithm (Redding, 1996). For its simplicity and lower computational time, the watershed approach has been used in this study. The resulting image I_{pc} (e.g. Figure 3) points to the potential crack pixels (PCP).
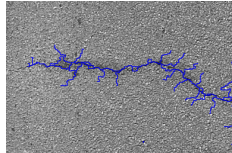

Figure 3: Potential crack pixels from Figure 2 (I_{pc} displayed in blue on the original image).

A last step has been added to validate I_{pc} and reduce the false detections: a PCP is validated if its contrast is above a given threshold C_{Tr}. To make the measure more robust to the noise, instead of analyzing each PCP independently, the contrast is averaged on segments (a segment is a set of contiguous pixels having at most two neighboring PCP, see Figure 4 for an illustration of this vocabulary). The average contrast $\overline{C_{seg}}$ of a segment is the difference between the average gray-level of the segment ($\overline{GL_{seg}}$), and the average gray-level of the neighboring background ($\overline{GL_{Bkg}}$):

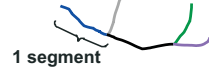$$\overline{C_{seg}} = \overline{GL_{Bkg}} - \overline{GL_{seg}}. \quad (3)$$

Figure 4: A potential crack composed of 5 segments (each having a different color).

$\overline{GL_{Bkg}}$ (e.g. Figure 5) is measured using a large filter averaging the neighboring values of the background (pixels not labeled as PCP). The filter kernel size did not seem critical, but it should be large enough to reduce the influence of noise, and small enough to consider the irregularities of the road: a 25×25 pixels kernel was experimentally chosen.
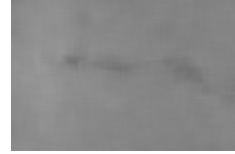

Figure 5: Estimated background of Figure 1.

A segment is considered as targeting a crack if $\overline{C_{seg}} > C_{Tr}$ (e.g. Figure 6). Measures have been realized to set the main parameters of the algorithms, i.e. the kernel size of the standard deviation filter and the threshold C_{Tr}; we identified that the best results were obtained with a 5×5 pixels kernel and $C_{Tr} = 50$. We will describe the analysis methodology, and we will restrain the presentation of the influence of C_{Tr} on the results.
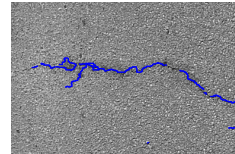

Figure 6: Final crack detected on Figure 1 (using $C_{Tr} = 40$).

4 ASSESSMENT METHODOLOGY

To evaluate and compare the crack detection methods, we had to choose: a/ the tested images, b/ how to determine the "ground truth" segmentation or reference segmentation and c/ the criteria in order to quantitatively evaluate the results.

4.1 Images

The algorithms were evaluated on 3 sets of test images: on 14 synthetic crack images, on 10 real images, and on 10 pre-processed real images. For synthetic images, it is easy and reliable to give a ground truth segmentation. For real images, estimate a "ground truth" or a reference segmentation is more complicated but the images are more realistic than the synthetic images.

The 14 synthetic images were built using different kinds of backgrounds: cracks were added on 8 images built with a random dot texture (sample S_1), and on 6 images of road containing no cracks, 2 of them acquired with a static camera (sample S_2), and 4 of them dynamically acquired with a camera embedded on a vehicle (sample S_3). For the last four ones, controlled lights were added. For all the 14 images, the cracks were randomly added, with a random shape and a random gray-level (Figure 7).

On the 10 real images, 4 were acquired using a static camera and 6 using the dynamic system (Figure 7).

In the third set of test images, the 10 real images were pre-processed using:

1. *Threshold –* In order to reduce the light halo in some images (the last six ones presented in Figure 9), each pixel over a given threshold is replaced by the local average gray levels.

2. *Smoothing –* A mean filter of size 3×3 is applied.

3. *Erosion –* An erosion with a square structuring element of size 3×3 is applied.

4. *Restoration –* This last pre-processing tries to combine the advantages of all the previous methods in three steps: histogram equalization, thresholding (like *Threshold*), and erosion (like *Erosion*).

4.2 Reference Segmentation

For real images, we briefly explain how the manual segmentation is validated. Four experts manually segmented the images with the same tools and in the same conditions. Then, the four segmentations were merged, following these rules:

1. A pixel labeled as crack by almost two experts was kept;

2. Every pixel near to a pixel kept by step 1 was also kept.

The second rule is iterative and stops when no pixel is added.
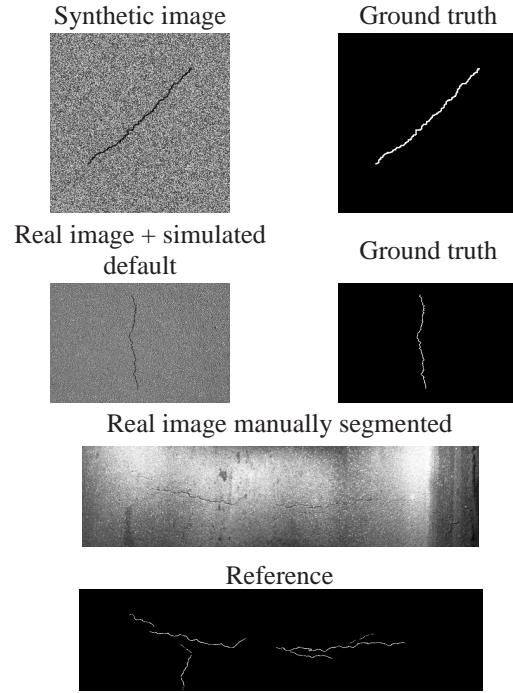

Figure 7: Tested images.

Then, the result is dilated with a structuring element of size 3×3. Results of the manual segmentations are presented in Figure 9 (for more details, the study of the reliability of these reference segmentations can be found in (Chambon et al., 2010)).

4.3 Quantitative Assessment

In Figure 8, the evaluation criteria are presented. They are the most used in the literature, except "accepted" pixels. In consequence, we have included accepted pixels in the computation of the similarity coefficient. For estimating accepted pixels, a threshold is applied on the distance between the detected pixel and the reference segmentation. This threshold is estimated using the mean distance between each of the four segmentations used for establishing the reference.

The DICE values lie on the $[0; 1]$ interval, but to simplify the visualization of the results, we display DICE$\times 100$ in all figures and tables of section 5.

5 RESULTS AND DISCUSSION

5.1 Synthetic Images

Figure 10 compares results for the different methods with synthetic images. It shows that the multi-resolution with $C_{Tr} = 50$ is the best to identify most of
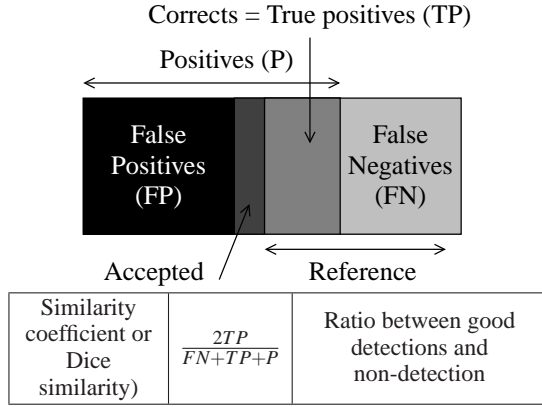
Figure 8: Evaluation criteria – Representation of the recovery between an estimated segmentation (positives, P) and a reference segmentation.

Similarity coefficient or Dice similarity)	$\frac{2TP}{FN+TP+P}$

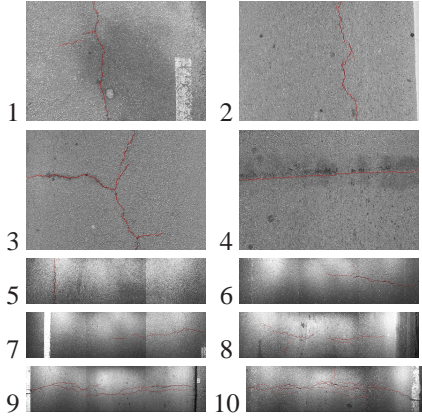(Ratio between good detections and non-detection)

Figure 9: Manual ground truth segmentations.

the edges, while the AFMM method is more adapted to keep the proportion of FN low. By comparing DICE values, the multi-resolution method appears to be the most adapted.
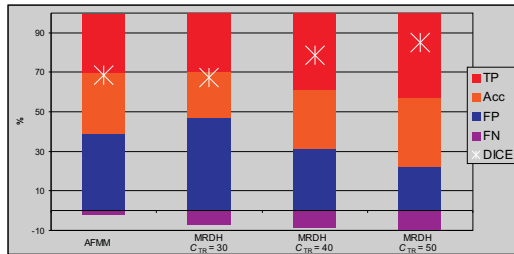

Figure 10: Summary of the algorithm performances on synthetic images.

Table 1 details the results for the different types of images (samples S_1 to S_3 presented in section 4.1). On fully synthetic images, both algorithms perform very well. It appears that the images acquired with the dynamic system are the most difficult to process, whatever the method used.

Table 1: Results on Synthetic Images – The 3 image categories are: S_1, fully synthetic images, S_2, background acquired with a camera and S_3, background acquired with the dynamic system. Results in bold indicate best values and it illustrates how the MRDH approach outperforms the AFMM, in particular, with the most difficult images (sample S_3) where DICE is correct (greater than 50) whereas it is very low with the other methods.

	AFMM			MRDH, $C_{Tr}=30$			MRDH, $C_{Tr}=40$			MRDH, $C_{Tr}=50$		
	S_1	S_2	S_3	S_1	S_2	S_3	S_1	S_2	S_3	S_1	S_2	S_3
TP+Acc	87	46	15	68	41	30	86	53	43	**91**	**66**	**59**
FN	**1.5**	**3.7**	**2.6**	2.8	4.4	17	5.3	6.9	18	6.7	6.9	18
DICE	92	63	24	81	59	44	92	69	57	**94**	**80**	**70**

5.2 Real Images

Figure 11 displays the results obtained on the real images. We see that the amount of FN is higher than with synthetic images, but we note that these values are probably over-estimated, the human-reference images tending to select only the major cracks. The AFMM algorithm identified almost half of the cracks selected by the human, and a bit more than half were identified with the multi-resolution method. Considering the DICE measure, the multi-resolution tests gave the best overall results; though, on camera acquired images, DICE values can be considered as equivalent for both methods.
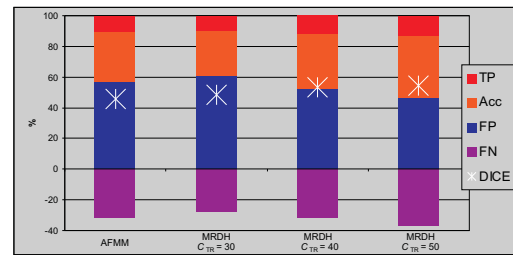

Figure 11: Summary of the algorithm performances on real images.

5.3 Real Pre-Processed Images

Here, we compare the AFMM method with the MRDH method (only the one with the more efficient

parameter, i.e. $C_{Tr} = 50$, is discussed here) when the real images are pre-processed. In particular, we evaluate the influence of the 4 pre-processing described in section 4.1.

Table 2: DICE of pre-treated images: the three numbers show the average values: **with the 10 images** (with the 4 images acquired in static way, with the 6 images acquired in a dynamic manner).

	AFMM	MRDH, $C_{Tr} = 50$
None	**46** (49,44)	**55** (52,57)
Threshold	**64** (68,61)	**55** (60,50)
Smoothing	**66** (77,59)	**52** (63,42)
Erosion	**59** (58,60)	**57** (64,51)
Restoration	**52** (49,55)	**57** (59,55)

Table 2 shows the DICE measures obtained. It can be seen that the four pre-processings enhance the overall performances.

When pre-processed with the *threshold*, *smoothing*, *erosion* and *restoration*, the AFMM's algorithm is better, with the best results being obtained after a *smoothing*. Also, *smoothing* and *threshold* have an even more beneficial effect when static images are used.

According to the DICE measure, the performances of MRDH applied after pre-treatment are quite disparate. In fact, it illustrates how this new method is globally quite robust against the acquisition conditions. Images do not need to be pre-processed in order to increase the performances and this is an important superiority compared to the AFMM method which is highly dependent on the acquisition conditions. Analyzed in details, we notice that for static images, DICE values seem to be enhanced by each pre-processing, while DICE values seem to decrease for "dynamic" images.

6 CONCLUSIONS

A methodology for adapting a multi-resolution-based segmentation to the field of road crack detection has been introduced. The experimental results demonstrate the efficient results obtained on 14 synthetic images and 10 significant real images. Moreover, it shows how it outperforms a previous method based on adapted filtering, in particular with the most difficult images acquired dynamically and that present illumination defaults.

The next step of this work will be to validate this work on more data (32 more images will be available soon). It has been shown that MRDH improves the percentage of true detections (this percentage is higher than with AFMM), but the AFMM presents less false negatives. The next improvement will be to reduce this phenomenon by introducing more constraints on the final step, like for example, an active contour constraint instead of using only the constraint on the contrast.

REFERENCES

Bray, J., Verma, B., Li, X., and He, W. (2006). A neural nework based technique for automatic classification of road cracks. In *International Joint Conference on Neural Networks*, pages 907–912.

Chambon, S., Gourraud, C., Moliard, J.-M., and Nicolle, P. (2010). Road crack extraction with adapted filtering and markov model-based segmentation. In *VISAPP*.

Coudray, N., Buessler, J.-L., Kihl, H., and Urban, J.-P. (2007). Tem images of membranes: a multiresolution edge-detection approach for watershed segmentation. In *Physics in Signal and Image Processing (PSIP 2007)*.

Elbehiery, H., Hefnawy, A., and Elewa, M. (2005). Surface defects detection for ceramic tiles using image processing and morphological techniques. *Proceedings of World Academy of Science, Engineering and Technology (PWASET)*, 5:158–162.

Frangi, A. F., Niessen, W. J., Vincken, K. L., and Viergever, M. A. (1998). Muliscale vessel enhancement filtering. In *MICCAI*, pages 130–137.

Geman, D. and Jedynak, B. (1996). An active testing model for tracking roads in satellite images. *IEEE Pattern Analysis and Machine Intelligence*, 18(1):1–14.

Iyer, S. and Sinha, S. (2005). A robust approach for automatic detection and segmentation of cracks in underground pipeline images. *Image and Vision Computing*, 23(10):921–933.

Koutsopoulos, H. and Downey, A. (1993). Primitive-based classification of pavement cracking images. *ASCE, Journal of Transportation Engineering*, 119(3):402–418.

Redding, N. (1996). The autoscaling of oblique ionograms. Technical report, Defence science and technology organization canberra.

Schmidt, B. (2003). Automated pavement cracking assessment equipment – state of the art. Technical Report 320, Surface Characteristics Technical Committee of the World Road Association (PIARC).

Subirats, P., Fabre, O., Dumoulin, J., Legeay, V., and Barba, D. (2006). Automation of pavement surface crack detection with a matched filtering to define the mother wavelet function used. In *European Signal Processing Conference*.

Tanaka, N. and Uematsu, K. (1998). A crack detection method in road surface images using morphology. In *Workshop on Machine Vision Applications*, pages 154–157.