# Road crack extraction with adapted filtering and Markov model-based segmentation : introduction and validation

Sylvie Chambon, Christian Gourraud, Jean Marc Moliard, Philippe Nicolle

# ROAD CRACK EXTRACTION WITH ADAPTED FILTERING AND MARKOV MODEL-BASED SEGMENTATION

## *Introduction and validation*

S. CHAMBON, C. GOURRAUD, J.-M. MOLIARD, P. NICOLLE
*Laboratoire Central des Ponts et Chaussées, LCPC, Nantes*
*{chambon,gourraud,moliard,nicolle}@lcpc.fr*

Keywords: Fine structure extraction, evaluation and comparison protocol, matched filter, Markov model, segmentation.

Abstract: The automatic detection of road cracks is important in a lot of countries to quantify the quality of road surfaces and to determine the national roads that have to be improved. Many methods have been proposed to automatically detect the defects of road surface and, in particular, cracks: with tools of mathematical morphology, neuron networks or multiscale filter. These last methods are the most appropriate ones and our work concerns the validation of a wavelet decomposition which is used as the initialisation of a segmentation based on Markovian modelling. Nowadays, there is no tool to compare and to evaluate precisely the peformances and the advantages of all the existing methods and to qualify the efficiency of a method compared to the state of the art. In consequence, the aim of this work is to validate our method and to describe how to set the parameters.

## 1 INTRODUCTION

In many countries, the quality of roads is evaluated by taking into account numerous characteristics: adherence, texture and defects. Since 1980, many efforts have been spent for making this task more comfortable, less dangerous for employees and also, more efficient and less expensive by using an acquisition system of road images and by introducing semi-automatic defect detection, in particular crack detection. Nowadays, a lot of acquisition systems are proposed, see (Schmidt, 2003), but, as far as we are concerned, even if a lot of automatic crack detections had been proposed, no method had been widely applied. In fact, this problem is difficult because road cracks represent a small part of the images (less than 1.5% of the image) and is very low contrasted (the signal of the crack is mixed with the road texture). Actually, there is no evaluation protocol to compare and to highlight methods that are well dedicated to this task and we propose to introduce such a protocol. Moreover, our second aim is to introduce and to validate our contributions about a multi-scale approach based on two steps: a binarization with adapted filtering and a refinement of the binarization by a Markov model-based segmentation.

First, existing methods are resumed before describing our approach. Then, we introduce the evaluation protocol that enables us to validate the new approach and we give the details of the parameter settings.

## 2 DETECTION OF ROAD CRACKS

For this task, three steps can be distinguished: image acquisition, data storage and image processing. In this paper, even if the choices for the two first steps are important, we focus on the image processing step.
Four different kinds of methods exist, see Table 1. The *thresholding methods* are the oldest ones and also the most popular. They are based on histogram analysis (Acosta et al., 1992), on adaptive thresholding (Elbehiery et al., 2005), or on Gaussian modelling (Koutsopoulos and Downey, 1993). These techniques are simple but not very efficient: the results show a lot of false detections. Some methods are based on *morphological* tools (Elbehiery et al., 2005; Iyer and Sinha, 2005; Tanaka and Uematsu, 1998). The results contain less false detections but they are highly dependent on the choice of the parameters. *Neuron net-*

*works*-based methods have been proposed to alleviate the problems of the two first categories (Kaseko and Ritchie, 1993). However, they need a learning phase which is not well appropriate to the task. *Filtering* methods are the most recent. At the beginning, a contour detection has been used but the major drawbacks lie on using a constant scale, supposing that the width of the crack is constant and that we suppose that the background is not as noisy (or textured) as a road surface. This is not realistic and this is why, most of the filtering methods are based on a wavelet decomposition (Delagnes and Barba, 1995; Subirats et al., 2006; Wang et al., 2007; Zhou et al., 2005) or on partial differential equations (Augereau et al., 2001). We can also notice an auto-correlation method (Lee and Oshima, 1994) (the authors estimate a similarity coefficient between patterns that represent simulated cracks and patterns inside the images). Some methods also use texture decomposition (Petrou et al., 1996; Song et al., 1992) (the goal is to find a noise, i.e. the crack, inside a known texture).

Table 1: State of the art about road crack detection – In brackets, we give the years of publication.

| | |
|---|---|
| THRESHOLD (1992–1999) | (Acosta et al., 1992; Cheng et al., 1999; Koutsopoulos and Downey, 1993) |
| MORPHOLOGY (1998) | (Tanaka and Uematsu, 1998) |
| NEURAL NETWORK (1991–2006) | (Bray et al., 2006; Chou et al., 1995; Kaseko and Ritchie, 1993; Ritchie et al., 1991) |
| MULTI-SCALE (1990–2009) | (Chambon et al., 2009; Delagnes and Barba, 1995; Fukuhara et al., 1990; Subirats et al., 2006; Zhou et al., 2005) |

# 3 AFMM Method

First, we present the general algorithm of the method: Adapted Filtering and Markov Model-based segmentation (AFMM) and, second, our contributions in each of the parts of this algorithm.

## 3.1 Algorithm

The goal of this algorithm, presented in Figure 1, is to obtain, steps 1 to 3, a binarization (black pixels for background and white pixels for the crack) and a refinement of this detection by using a Markovian segmentation, step 4,. In the first part, a photometric hypothesis is used: a crack is darker than the background (the rest of the road pavement) whereas, in the second part, a geometric hypothesis is exploited: a crack is composed of a set of connected segments with different orientations. The number of scales for the adapted filtering has to be chosen and depends on the resolution of the image. By supposing a resolution of 1 mm per pixel, by choosing 5 scales, a crack with a width from 2 mm to 1 cm can be detected. Moreover, the number of directions (for the filtering) also has to be chosen and, it seems natural to take these four directions: $[0,\ \frac{\pi}{4},\ \frac{\pi}{2},\ \frac{3\pi}{4}]$ that correspond to the four usual directions used for crack classification. Adapted filtering is applied in each scale, each directions and then all the results are merged (mean of the coefficients). The results of this filtering is used to initialize the Markov model.

```
Input
Road images
---------------------------------------------
Initialization
Number of scales and angles
---------------------------------------------
Steps
1. For each scale do Estimate adapted
   filter (AF)
2. For each direction do Apply AF
3. Merge AF in all the directions
4. For each scale do
 (a) Initialize the sites (Markov)
 (b) While not (stop condition) do
     Update the sites
```

Figure 1: Algorithm with adapted filter and Markovian modelling (AFMM) – Steps 1 to 3 lead to a binary image using adapted filtering, while step 4 refines this result with a Markovian modelling.

## 3.2 Adapted filtering

The $\psi \in \mathcal{L}^2(\mathbb{R}^2)$ function is a wavelet if:

$$\int_{\mathbb{R}^2} \frac{|\Psi(\mathbf{x})|^2}{\|\mathbf{x}\|^2} d\mathbf{x} < \infty, \text{avec } \mathbf{x} = (i, j), \qquad (1)$$

where $\Psi$ is the Fourier transform of $\psi$. The equation (1) induces that $\int_{\mathbb{R}^2} \psi(\mathbf{x}) d\mathbf{x} = 0$. The wavelet family is defined for each scale $s$ and for each position $\mathbf{u}$, by :

$$\psi_{x,\hat{u},\theta}(\mathbf{x}) = \frac{1}{2}\psi(R^{\theta}((\mathbf{x}-\mathbf{u})/s)), \qquad (2)$$

where $\psi \in \mathbb{R}^2$ and $R^\theta$ is a rotation of angle $\theta$. One of the main difficulties to apply a wavelet decomposition is the choice of the mother wavelet $\psi$. Numerous functions are used in the literature: Haar wavelet, Gaussian derivatives, Mexican hat filter, Morlet wavelet. It is very hard to determine which one is the best for a given application. In the case of crack detection, two elements are present: the crack (if there is a crack) and the background (the road surface can be viewed as a repetitive texture). The goal of the crack detection is to recognize a signal, i.e. the shape is known up to a factor, mixed with a noise whose its characteristics are known. Consequently, adapted filtering is well designed for the problem: extracting singularities in coefficients estimated by a wavelet transform. If $s$ is a discrete and deterministic signal with values stored in the vector $\mathbf{s} = \begin{pmatrix} s_1 & \dots & s_N \end{pmatrix}$, $N$ the number of samples, and $\mathbf{z} = \begin{pmatrix} z_1 & \dots & z_N \end{pmatrix}$, is a noisy observation of $s$, supposed to be an additive noise: $\mathbf{z} = \mathbf{s} + \mathbf{b}$. The main hypothesis is that this second-order noise is centered and stationary, with auto-correlation function $\phi_{bb}$ of terms $\phi_{bb_{(i,j)}} = \phi_{bb_{|i-j|}}$, independent from the signal $s$. The adapted filter $\mathbf{h}$ of $\mathbf{s}$ is defined by:

$$\mathbf{h} = \phi_{bb}^{-1}\mathbf{s}. \tag{3}$$

The crack signal depends on the definition of the crack. In this paper, like most of the papers of this domain, crack pixels correspond to black pixels surrounded by background pixels (road pixels). This is why, in (Subirats et al., 2006), a crack is a piecewise constant function:

$$f(x) = \begin{cases} -a & \text{If } x \in [-\frac{T}{2}, \frac{T}{2}] \\ 0 & \text{Elsewhere,} \end{cases} \tag{4}$$

where the factor $a$ and the threshold $T$ have to be determined. It does not correspond to a realistic representation of the crack. Because of sub-sampling, lights, orientation of the camera, the signal is more like a Gaussian function with zero mean:

$$f(x) = -a\, e^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2}, \tag{5}$$

where $a$ is the size of the crack and depends on $\sigma$, the deviation of the Gaussian law, i.e. $a = \frac{1}{\sigma\sqrt{(2\pi)}}$. Consequently, the term $\sigma$ allows to fix the width of the crack (like threshold $T$ in equation (4)).

## 3.3 Contribution to segmentation

The goal of this part is to extract shapes, i.e. cracks, using the detection maps estimated at the first stage of the algorithm. The MRF (Markov Random Field) principle is introduced before the presentation of the improvements about the initialization of sites and the updating of the Markov Random Field.

For the first step of segmentation (initialization), in (Subirats et al., 2006), the sites are of size $3 \times 3$, consequently, a regular grid is considered in the image. The four configurations that are possible, are represented in Figure 2. The initialization of the sites is based on the configuration that maximizes the wavelet coefficients. More formally, if we denoted $\gamma_{2,0}$, $\gamma_{2,\frac{\pi}{4}}$, $\gamma_{2,\frac{\pi}{2}}$ and $\gamma_{2,\frac{3\pi}{4}}$, see in the bold rectangle in Figure 2, the four configurations, the best configuration $\gamma_{\text{best}}$ is:

$$\gamma_{\text{best}} = \underset{\alpha \in \left[0..\frac{3\pi}{4}\right]}{\operatorname{argmax}}\, m_{2,\alpha}, \tag{6}$$

where $m_{2,\alpha}$ is the mean of wavelet coefficients on the considered configuration $\gamma_{2,\alpha}$. These four configurations do not represent all the possible and are not realistic configurations. In fact, all these four configurations are centered, whereas, it is possible to have some non-centered configurations. Consequently, we use a set of configurations that includes this aspect and we employ a set of sixteen configurations illustrated in Figure 2. By modifying the number of configurations, we need to adapt the initialization of sites, i.e. equation (6).

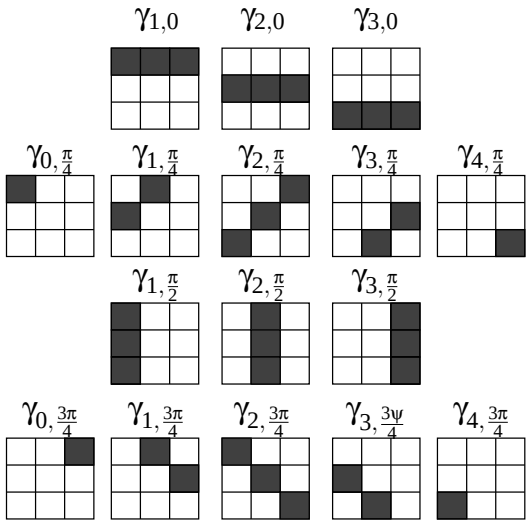



Figure 2: The sixteen configurations of the Markov model (initialization) – The sites are represented with dark gray levels.

The image is considered as a finite set of *sites* denoted $\mathcal{S} = \{s_1, \dots, s_N\}$. For each site, the *neighborhood* is defined by: $\mathcal{V}_s = \left\{ s' | s \notin \mathcal{V}_{s'} \;\&\; s' \in \mathcal{V}_s \Rightarrow s \in \mathcal{V}_{s'} \right\}$. A *clique* $c$ is defined as a subset of sites in $\mathcal{S}$ where every pair of distinct sites are neighbors.

The random fields considered are:

1. The observation field $Y = \{y_s\}$ with $s \in \mathcal{S}$. Here, $y_s$ is the mean of wavelet coefficients on the site.
2. The descriptor field $L = \{l_s\}$ with $s \in \mathcal{S}$. For this work, if there is a crack $l_s = 1$ elsewhere $l_s = 0$.

MRF model is well suited to take into account spatial dependencies between the variables. Each iteration, a global cost, or a sum of *potentials*, that depends on the values of the sites and the links between neighborhoods, is updated. This global cost takes into account the site coefficients (computed from the wavelet coefficients estimated during the first part of the algorithm: adaptive filtering) and the relation between each site and the sites in its neighborhood (in this paper, it corresponds to the eight neighbors). More formally, the global cost is the sum of all the potentials of the sites and contains two terms:

$$u_s(s) = \alpha_1 u_1(s) + \alpha_2 \sum_{s' \in \mathcal{V}_s} u_2(s, s'), \qquad (7)$$

where $\mathcal{V}_s$ is the neighborhood of site $s$. The choice of the values $\alpha_1$ and $\alpha_2$ depends on the importance of each part of the equation (7).

The function $u_1$ is given by:

$$u_1(y_s, l_s = 1) = \begin{cases} e^{\xi_1 (k - y_s)^2} & \text{If } y_s \geq k \\ 1 & \text{Elsewhere} \end{cases} \text{ and}$$
$$u_1(y_s, l_s = 0) = \begin{cases} e^{\xi_2 (y_s - k)^2} & \text{If } y_s < k \\ 1 & \text{Elsewhere,} \end{cases} \qquad (8)$$

The parameters $\xi_1$, $\xi_2$ and $k$ have to be fixed[1]. For the definition of $u_2$, we have to determine the number of cliques. In (Subirats et al., 2006), four cliques are possible. As there is four configurations in the previous approach, there is sixteen possibilities. A 8-neighborhood is considering but the potential function proposed in the precedent work only considers the difference of orientations between two neighborhoods and not the position between the two sites of the clique, see Table 2.

Some cases are not penalized with the old configuration. For example, these two unfavorable cases are not penalized: two sites with the same orientation but with no connection between them, two sites with the same orientation but their position makes them parallel. This is why, with the sixteen possible configurations presented in Figure 2, the new variant takes into account differences of orientations between two

[1]The choice of $k$ is related to the maximal number of pixels that belong to a crack (it depends on the resolution of images and hypothesis about the size and configuration of cracks). We have chosen $k$ in order to consider at most 5% of the image as a crack. Moreover, our experimentations have brought us to take $\xi_1$=$\xi_2$=100.

| | $\gamma_{2,0}$ | $\gamma_{2,\frac{\pi}{4}}$ | $\gamma_{2,\frac{\pi}{2}}$ | $\gamma_{2,\frac{3\pi}{4}}$ |
|---|---|---|---|---|
| $\gamma_{2,0}$ | $\beta_1$ | $\beta_2$ | $\beta_3$ | $\beta_2$ |
| $\gamma_{2,\frac{\pi}{4}}$ | $\beta_2$ | $\beta_1$ | $\beta_2$ | $\beta_3$ |
| $\gamma_{2,\frac{\pi}{2}}$ | $\beta_3$ | $\beta_2$ | $\beta_1$ | $\beta_2$ |
| $\gamma_{2,\frac{3\pi}{4}}$ | $\beta_2$ | $\beta_3$ | $\beta_2$ | $\beta_1$ |

Table 2: Function $u_2$ – This table represents the values of the function $u_2(s', s)$ for the four initial configurations of the sites. In the experiments we have taken the values of the parameters proposed by the authors (Subirats et al., 2006) and that give the best results: $\beta_1 = -2$, $\beta_2 = -1$ and $\beta_3 = 2$.

sites (there are $16 \times 16$ possibilities) and position of the two sites (there is eight possibilities because of 8-connexion). Consequently, the new potential function $u_2$ follows these two important rules:

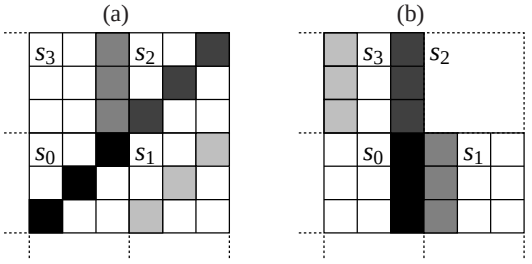
($R_1$) The lower the difference of orientations between two sites, the lower the potential.

($R_2$) The lower the distance between two sites, the lower the potential (in this case, distance means the minimal distance between the extremities of the two segments).

More formally, if $d$ denotes the Euclidean distance between the two closest extremities of the sites, with $d \in [0, d_{\max}]$ ($d_{\max} = 5\sqrt{2}$), $\theta_1$ and $\theta_2$, the orientations of respectively $s = \{s_i\}_{i=1..N_s}$, and $s' = \{s'_j\}_{j=1..N_{s'}}$ and $\theta_e$ the angle between the two sites, the $u_2$ function is defined by:

$$u_2(s', s) = \alpha_2 \left( \frac{|2\theta_e - \theta_1 - \theta_2|}{2\pi} \right) + (1 - \alpha_2) \left( \frac{J(\text{NbC}) \min_{i,j}(d(s_i, s'_j))}{d_{\max}} - \frac{\text{NbC}}{3} \right). \qquad (9)$$

where NbC indicates the number of elements of the two sites $s$ and $s'$ and $J(x)$ equals 1 if NbC $= 0$ and 0 elsewhere. The first term is induced by the rule about the orientations, rule ($R_1$), it is zero when the sites have the same orientation and this orientation is the same as the orientation between the sites, i.e. $\theta_e = \theta_2 = \theta_1$. It gives bad costs to configurations where the sites do not have the same orientation but also the particular case where they are parallel, see example (a) in Figure 3. The second and third terms express the rule ($R_2$) about the distances. Two aspects have to be distinguished: the number of connected pixels, when the sites are connected, and, on the contrary, the distance between the sites. It allows to give low influence to disconnected sites and also to increase the cost of sites that are parallel but connected, see example (b) in Figure 3. To study the influence of all these terms, the equation has been normalized and the different terms have been weighted.




$\theta_0 = \theta_1 = \theta_2 = \frac{\pi}{4}, \theta_3 = \frac{\pi}{2}$
$\theta_{e1} = 0, \theta_{e2} = \frac{\pi}{4}$ and
$\theta_{e3} = \frac{\pi}{2}$
For $s_1 : u_2 = \frac{\pi}{2} - 0 + 2$
For $s_2 : u_2 = 0 - 1 + 0$
For $s_3 : u_2 = \frac{\pi}{4} - 1 + 0$

$\theta_0 = \theta_1 = \theta_2 = \theta_3 = \frac{\pi}{2}$
$\theta_{e1} = 0, \theta_{e2} = \theta_{e3} = \frac{\pi}{2}$
For $s_1 : u_2 = \pi - 3 + 0$
For $s_2 : u_2 = 0 - 1 + 0$
For $s_3 : u_2 = 0 - 0 + 2$

Figure 3: Illustration of function $u_2$ – These two examples of sites with their respective neighborhoods show the behavior of potential $u_2$ with the two considered aspects: orientation and distance. In example (a), with the help of the orientation term, the configuration $s_1$ is penalized and $s_3$ is less penalized than $s_1$. In example (b), with the help of the two terms on distance, the site $s_3$ is penalized, compared to $s_2$. On the contrary, the particular case of $s_1$ is favorable and it equilibrates the penalty given by the orientations.

# 4 EVALUATION PROTOCOL

For the evaluation of automatic crack detection, there is no evaluation and comparison protocol proposed in the community. However, in all the countries, for estimating the quality of the road surface, it is important to know exactly the size and the width of the defects, i.e. to detect precisely the defect. This is why, it seems important to characterize quantitatively the performance of the methods. For building this kind of protocol, it is necessary, first, to choose the tested images, second, to choose how to build reference segmentations, and, third, to determine the criteria used for the quantitative analysis.

## 4.1 Tested images

The most difficult is to propose images with a reference segmentation or a "ground truth" segmentation. On the first hand, we create synthetic images with a simulated crack. As shown in figure 4, the result is not enough realistic and, on the second hand, we have taken a real image with no defect and we have added a simulated defect. The result is more realistic but the shape of the crack (which is randomly chosen) does not seem enough realistic. This is why, it appears important to propose a set of real images with manual segmentations that are enough reliable to be considered as reference segmentations or "ground truth" segmentations. To resume, the two first kinds of images allow to propose an exact evaluation and to well illustrate the behavior of the method in theoretical cases whereas the last kind of images allows to validate the work on real images with a "pseudo" ground truth.

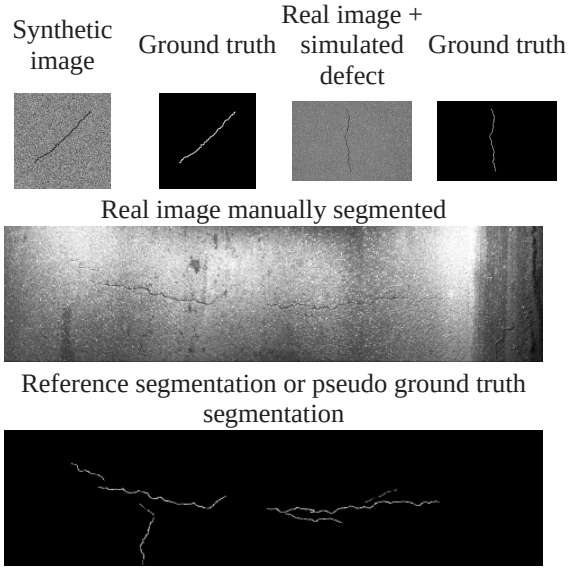



Figure 4: Tested images.

## 4.2 Reference segmentation

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

1. A pixel marked as a crack by more than two experts is considered as a crack pixel;
2. Every pixel marked as a crack and next to a pixel kept by step 1 or 2 is also considered as a crack.

The second rule is iterative and stops when no pixel is added. Then, the result is dilated with a structuring element of size $3 \times 3$.
In this part, we distinguish two datasets of real images. First, we have work with 10 images, in order to validate our manual segmentation and to determine how to fix the parameters of the proposed method. This first dataset is called *initial dataset*. The second one contains 32 images to complete the evaluation and we called it the *complementary dataset*.

**Initial dataset** To evaluate the reliability of the reference segmentations, we estimate, first, the percentage of covering between each operator, and, second,

the mean distance between each pixel (detected by only one expert and not kept in the reference image) and the reference segmentation. Table 3 compares the results for these 10 first images. We can notice that the first 4 images are the most reliable because the mean error is less than 2 pixels. On the contrary, the last 6 images are less reliable but they are also the most difficult to extract a segmentation.

**Complementary dataset** The same technique is used for establishing the reference segmentations with 32 images. By analyzing the results for criteria D., presented in Table 3, we decided to classify the 42 tested images in 3 categories:

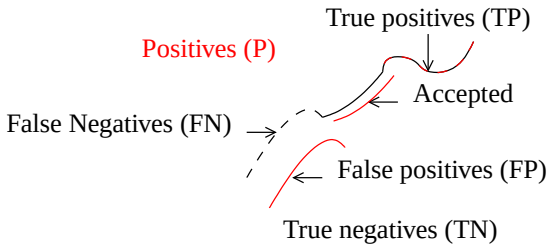
1. *Reliable:* All the images have obtained $D < 2$ and it means that all the operators have selected areas that are quite close to each other and the segmentation is reliable.
2. *Quite reliable:* All the images have obtained $2 \leq D < 8$, it means that some parts of the crack are not easy to segment and there are locally big errors.
3. *Ambiguous:* All the images have obtained $D \geq 8$. It clearly show that the images are really difficult to segment and in most of the cases, it means that some parts are detected as a crack whereas they are not a crack and reversely.

Table 3: Manual segmentation comparison for establishing the reference segmentations – For each image, we give the percentage of pixels that are preserved in the final reference segmentation compared to the size of the image (F), the percentage of covering between 2, 3 and 4 segmentations (over all the pixels marked as crack by the 4 experts) and the sum of them (S). For all the crack pixels not preserved in the final reference segmentation, the mean distance to this segmentation is given, noted D.

| Images | F (%) | 2 (%) | 3 (%) | 4 (%) | S (%) | D (pix) |
|---|---|---|---|---|---|---|
| 37  | 0.45 | 28.87 | 9.79 | 1.59 | 40.25 | 1.48 |
| 42  | 0.4 | 26.69 | 14.59 | 4.2 | 45.48 | 1.45 |
| 46  | 0.72 | 27.53 | 11.66 | 2.33 | 41.52 | 1.41 |
| 40  | 0.44 | 34.3 | 19.01 | 5.87 | 59.18 | 1.03 |
| 463  | 0.17 | 23.46 | 5.95 | 0.39 | 29.8 | 1.4 |
| 936  | 0.41 | 23.52 | 7.41 | 0.9 | 31.83 | 7.05 |
| 41  | 0.33 | 22.64 | 7.31 | 1.33 | 31.28 | 3.56 |
| 23  | 0.6 | 24.12 | 9.41 | 2.45 | 36.58 | 2.23 |
| 352  | 1.01 | 25.69 | 11.52 | 2.15 | 39.36 | 4.75 |
| 88  | 1.44 | 22.74 | 8.23 | 1.23 | 32.2 | 2.76 |

## 4.3 Evaluation Criteria

In this section, we introduce how the reference segmentation and the estimated segmentation are compared. In Figure 5, we present common evaluation criteria that are used for segmentation evaluation. We have added the principle of "accepted" detection that tolerates a small error on the localization of crack pixels. This criterion is needed because perfect detection seems, for the moment, difficult to reach, see the results in Table 3, that illustrate this aspect. In consequence, these "accepted" pixels have been included in the estimation of the similarity coefficient or DICE. The threshold for accepted pixels equals 0 for synthetic images whereas it depends on the mean distances, see D. in table 3, for the real images.




| | | |
|---|---|---|
| Sensitivity | $\frac{TP}{TP+FN}$ | Proportion of good detections |
| Specificity | $\frac{TN}{TN+FP}$ | Proportion of non-detected pixels |
| Similarity coefficient or Dice similarity | $\frac{2TP}{FN+TP+P}$ | Ratio between good detections and non-detections |

Figure 5: Evaluation criteria – In this figure, it corresponds to two simulated segmentations of the same crack: the black one is manual (or the reference) and the red one is estimated. The goal is to evaluate the quality of the estimated segmentation, that corresponds to the Positives (P). All the non-selected pixels that do not correspond to the crack are called the True Negatives (TN). Piece of crack with the two colors (red and black) are the correct detections or True Positives (TP). In the table, the criteria are resumed and in this work, we have used the DICE because this coefficient well represents what we want to measure: the quality of the detection against the percentage of the crack that is detected, or, how to reduce false detections and to increase the density.

# 5 Experimental results

For the proposed method, we want to determine, first, how to fix the different parameters, second, the preprocessing steps that are necessary, and, finally, which

variant is the most efficient. In consequence, we have tested different:

- *Parameter values* – The weights $\alpha_1$, equation (7), and $\alpha_2$, equation (9), are tested from 0 to 1 with a step of 0.1.
- *Pre-processings* – These pre-processings have been experimented to reduce noises induced by texture, to increase the contrast of the defect and, to reduce the light halo in some images (like the last six ones presented in Table 3):
  1. *Threshold* – Each pixel over a given threshold is replaced by the local average gray levels.
  2. *Smoothing* – A $3 \times 3$ mean filter is applied.
  3. *Erosion* – An erosion with a square structuring element of size $3 \times 3$ is applied.
  4. *Restoration* – It tries to combine the advantages of all the previous methods in three steps: histogram equalization, thresholding (like *Threshold*), and erosion (like *Erosion*).
- *Algorithm variants* – Four variants are compared:
  1. *Init* – This is the initial method proposed in (Subirats et al., 2006).
  2. *Gaus* – This variant considers the crack as a Gaussian function, see section 3.2.
  3. *InMM* – This is the initial version with an improvement of the Markov model (new definition of the sites and of the potential function), see section 3.3.
  4. *GaMM* – This is the *Gaus* version with the new Markov model.
- *Comparison* – We have compared this method with a method based on morphological tools and that is quite similar to (Tanaka and Uematsu, 1998) noted *Morph*.

## 5.1 Parameter influence

Among the results, two conclusions can be done. For each variant and each pre-processing, the weights between the term for adapted filtering and the term for the Markovian modelling should be the same, see equation (7), i.e. $\alpha_1 = 0.5$. However, when more weight is given to adapted filtering , the quality of the results is lower than when more weight is given to the Markovian segmentation. It means that in this kind of application, the geometric information is more reliable than the photometric one and it seems coherent with the difficulties of the acquisition.
For the new Markov model, we have noticed that the results are the best when the weights are the same between the orientation term and the distance term, see equation (9), i.e. $\alpha_2 = 0.5$. However, better results are obtained when the weight of the orientation is greater than the distance one instead of the reverse.

## 5.2 Pre-processing

These tests have been done with real images, because, a priori, synthetic images do not need pre-processings. The results are given by:

| *Init* | *Gaus* | *InMM* | *GaMM* |
|---|---|---|---|
| Restoration | Restoration | Threshold | Erosion |

However, for the four first images (acquired with lighting conditions more comfortable than the lighting conditions of the next 6 ones), the pre-processing is not significant for increasing the quality of the results. Moreover, with the new Markov model, the pre-processing step increases a little the quality.

## 5.3 Variants

The results are separated in two cases: with synthetic images and with real images. In Figure 6, the evolution of the similarity coefficient, or DICE, for the 11 synthetic images and the 10 real images is presented. With synthetic images, *GaMM* is clearly the best for most of the images. However, for one image (the fifth), the results are worse than with *Gaus* but they are still correct (DICE=0.72). On the contrary, for the most difficult images (the 3 first ones that contain a real road background), *GaMM* obtains acceptable results (DICE $> 0.5$) whereas the other methods are not efficient at all. An illustration is given in Figure 7: it shows how *GaMM* can reduce false detections.
The results with real images, see graph (b) in Figure 6, are coherent with those obtained with synthetic images. The new Markov model allows the best improvements, compared to the initial method *Init*. The method *GaMM* obtains the best results, except on few images, see an example in figure 8, where *InMM* is the best. However, *GaMM* gives results that are quite similar and the differences are not significant.

## 5.4 Complementary dataset and comparison

Finally, we have compared the result of *GaMM* on each of the complementary dataset (32 images) with the results obtained with a classic method based on morphological tools, like (Tanaka and Uematsu, 1998). The mean DICE is 0.6 with *GaMM* whereas it is 0.49 with the *Morphology* method, see Figure 9. It shows how this new method outperforms this kind

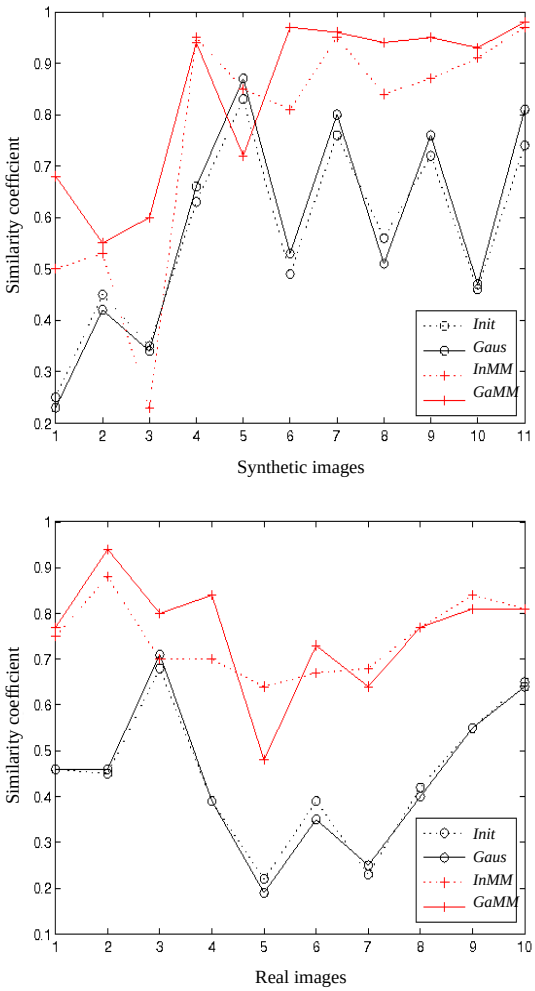



Figure 6: Variation of the similarity coefficient, see figure 5 – The first graph shows the results for synthetic images (the 3 first ones are obtained from real images with simulated defect) and the second graph presents the results with real images. Good performances of methods *InMM* and *GaMM* can be noticed.

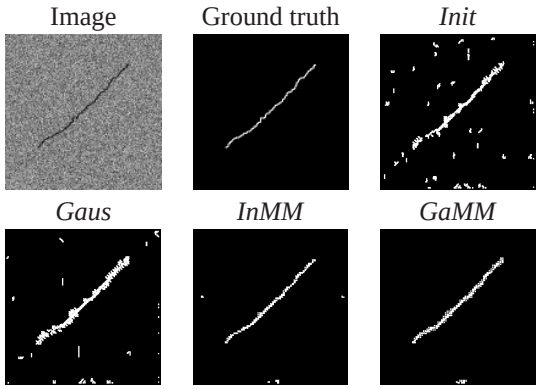



Figure 7: Segmentation results (synthetic images presented in Figure 4) – These are the results obtained with the four variants and it shows how *GaMM* gives the clearest result. We can also notice the good result of *InMM*.

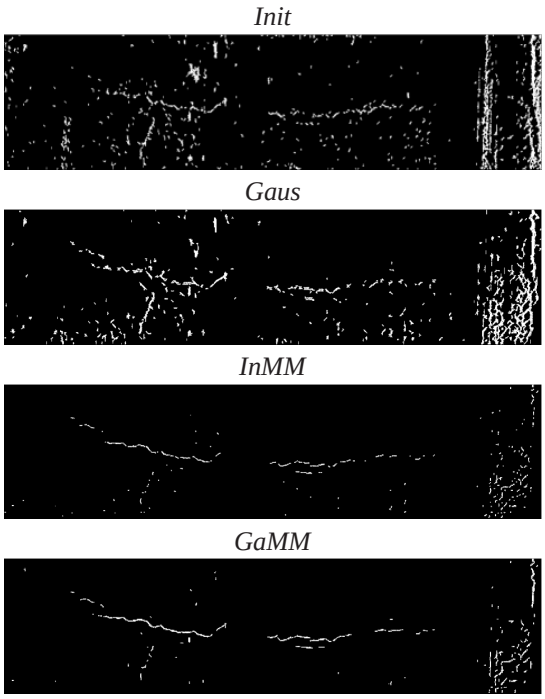



Figure 8: Segmentation results (real images) – These are the results obtained with the real images presented in figure 4. Method *InMM* obtains the clearest detection (i.e. with less false detection) but we can also noticed the good quality of the detection map with *GaMM*.

of method. However, if we compare image per image, the results show that in 50% of the cases *GaMM* is the best, see illustrations of these results in Figures 10 and 11. More precisely, it seems more performant with *Ambiguous* images, whereas the *Morphology* method is the best with *Reliable* images, see § "Complementary dataset".

# 6 CONCLUSIONS

In a first time, we have introduced new methods for the detection of road cracks. In a second time, we have presented a new evaluation and comparison protocol for automatic detection of road cracks and the

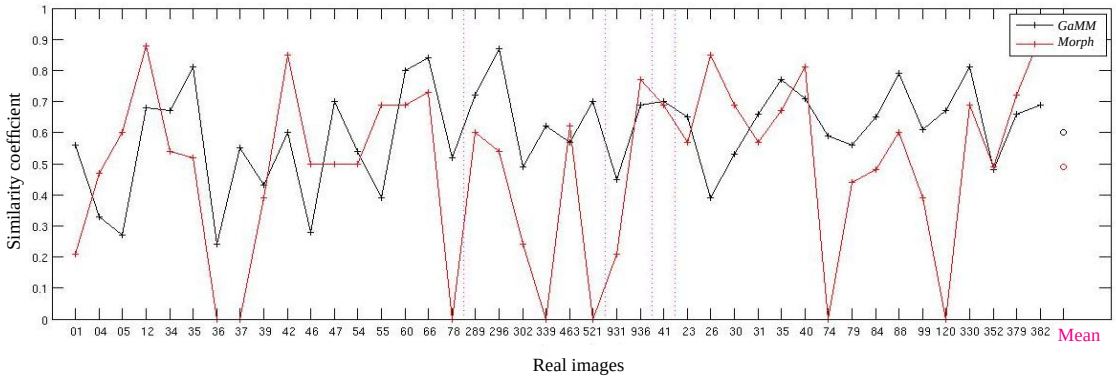



Figure 9: Comparison with *Morph* – The dotted lines illustrate the five different kinds of images tested. For the first one (that corresponds to real images with good illumination), the results are mixed whereas for the other ones, *GaMM* is the best.

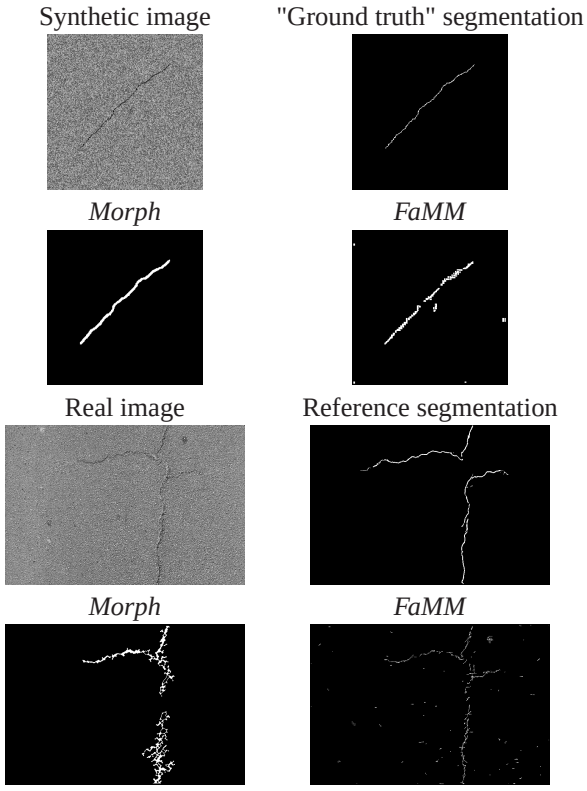



Figure 10: Differences between *Morph* and *GaMM* – Examples with synthetic and real images. *Morph* outperforms *GaMM* with simple synthetic images whereas *GaMM* gives better detection with real images.

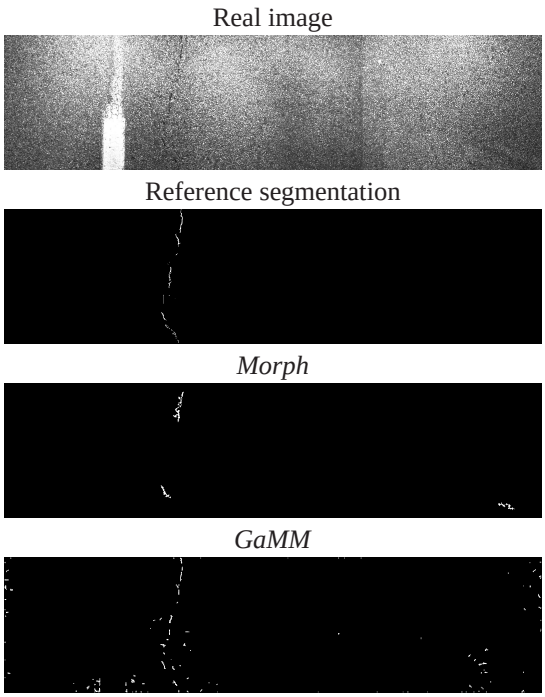



Figure 11: Differences between *Morph* and *GaMM* – Examples with real images acquired on a vehicule. The detection with *GaMM* is more complete than with *Morph*.

new methods are validated by this protocol. As far as we are concerned, we proposed real images with ground truth for the first time in the community. Our next work is to propose our ground truth to the community in order to have a larger comparison. Then, we want to increase our data set by taking into account the different qualities of road surface or road texture (because for the moment, each proposed method seems very dependent on the road texture). Finally, we want to refine our evaluation criteria by using (Arbelaez et al., 2009).

## REFERENCES


Acosta, J., Adolfo, L., and Mullen, R. (1992). Low-cost video image processing system for evaluating pavement surface distress. *TRR: Journal of the Transportation Research Board*, 1348:63–72.

Arbelaez, P., Fowlkes, C., and Malik, J. (2009). From contours to regions: An empirical evaluation. In *Computer Vision and Pattern Recognition*. To appear.

Augereau, B., Tremblais, B., Khoudeir, M., and Legeay, V. (2001). A differential approach for fissures detection on road surface images. In *International Conference on Quality Control by Artificial Vision*.

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., Subirats, P., and Dumoulin, J. (2009). 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*.

Cheng, H., Chen, J., Glazier, C., and Hu, Y. (1999). Novel approach to pavement cracking detection based on fuzzy set theory. *Journal of Computing in Civil Engineering*, 13(4):270–280.

Chou, J., O'Neill, W., and Cheng, H. (1995). Pavement distress evaluation using fuzzy logic and moment invariants. *TRR: Journal of the Transportation Research Board*, 1505:39–46.

Delagnes, P. and Barba, D. (1995). A markov random field for rectilinear structure extraction in pavement distress image analysis. In *International Conference on Image Processing*, volume 1, pages 446–449.

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.

Fukuhara, T., Terada, K., Nagao, M., Kasahara, A., and Ichihashi, S. (1990). Automatic pavement-distress-survey system. *Journal of Transportation Engineering*, 116(3):280–286.

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.

Kaseko, M. and Ritchie, S. (1993). A neural network-based methodology for pavement crack detection and classification. *Transportation Research Part C: Emerging Technologies information*, 1(1):275–291.

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

Lee, H. and Oshima, H. (1994). New Crack-Imaging Procedure Using Spatial Autocorrelation Function. *Journal of Transportation Engineering*, 120(2):206–228.

Petrou, M., Kittler, J., and Song, K. (1996). Automatic surface crack detection on textured materials. *Journal of Materials Processing Technology*, 56(1–4):158–167.

Ritchie, S., Kaseko, M., and Bavarian, B. (1991). Development of an intelligent system for automated pavement evaluation. *TRR: Journal of the Transportation Research Board*, 1311:112–119.

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).

Song, K., Petrou, M., and Kittler, J. (1992). Wigner based crack detection in textured images. In *International Conference on Image Processing and its Applications*, pages 315–318.

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, EUSIPCO*.

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

Wang, K., Li, G., and Gong, W. (2007). Wavelet-based pavement distress image edge detection with "À trous" algorithm. In *Transportation Research Record, Annual Meeting*, volume 2024, pages 73–81.

Zhou, J., Huang, P., and Chiang, F. (2005). Wavelet-based pavement distress classification. *TRR: Journal of the Transportation Research Board*, 1940:89–98.