Automatic Road Pavement Assessment with Image Processing: Review and Comparison

Sylvie Chambon, Jean Marc Moliard

Research Article

Automatic Road Pavement Assessment with Image Processing: Review and Comparison

Sylvie Chambon, Jean-Marc Moliard
Institut Français des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux (IFSTTAR), France
{*chambon,moliard*}*@ifsttar.fr*

In the field of noninvasive sensing techniques for civil infrastructures monitoring, this paper addresses the problem of automatic crack detection by automatic analysis of optical images of the surface of the French national roads. The first contribution is a state of the art of the image processing tools applied to civil engineering. Second, we describe the proposed method to detect fine defects in pavement surface. This approach is based on a multi-scale extraction and a Markovian segmentation. Third, an evaluation and comparison protocol which has been designed for evaluating this difficult task – the road pavement crack detection – is introduced. Finally, the proposed method is validated, analysed and compared to a detection approach based on morphological tools.

1 Introduction

The evaluation of road quality is an important task in many countries, like, for example, in France, where the national roads are inspected each three years in order to estimate the needed reparations and constructions. To estimate the quality, these aspects can be taken into account: the adherence, the micro-texture, the macro-texture and the surface degradations. Before 1980, all these inspections were accomplished manually, but, it can be automated with noninvasive techniques, like, image processing. To be more comfortable, less dangerous for employees and users of the road but also more efficient and less expensive, many systems have been proposed, based on ground penetrating radar [27] or laser system [53]. However, for noninvasive evaluation of surface degradations, the recent research results seem more promising with optical image processing approaches for these reasons [60]:

(1) The acquisition systems based on optical devices are easier to design and to use than other kinds of systems (they are less sensitive to mouvement or vibrations).

(2) They also allow a dense acquisition (each millimiter), i.e. the acquisition can be realized for the entire road surface, whereas for the other systems, like laser, the measurements are available every 4 millimeters at normal speed (90 km/h)[1].

(3) As the acquisition system is more dense, the measurement of the defects is more precise than with other systems.

(4) Even if the images are not always well contrasted, they are more contrasted than the images/signals that can be given by other devices, i.e. the ratio between noise and signal is greater with optical sensor than with other kind of sensors.

Nowadays, many acquisition systems are available [4, 60], see Table 1 (interested readers can find details about the evaluation of such systems in [46, 62]). Moreover, to the best of our knowledge, many semi-automatic detection of road defects can be found in the literature but only one is commercialized (by INO[2]). From all the approaches proposed, it is difficult to know which ones is the most adapted to the task and what are the actual methods that are favoured, this is why, the first goal of this paper is to present a state-of-the-art of existing methods in non-invasive control based on image processing.

This task of crack detection is difficult for the special case of road crack detection because it needs to detect a signal weakly represented (1.5% of the whole image) and weakly contrasted (the road possesses a texture that hides the crack). Recent methods have shown their limits: the detection contains a lot of false detections (induced by the particular texture of the road), the detection is not enough precise (we have

[1] It has been determined from the most recent systems.

[2] http://www.ino.ca/fr-CA/Realisations/Description/project-p/systeme-laser-mesure-fissures.html

a region of detection and not the skeleton plus the width of the crack). The main default of the existing methods is that they do not take into account the specific geometry of the crack: it is a thin and linear object. In consequence, the second aim of this work is to introduce a new method that take into account some geometric properties of the cracks.

Even if this problem is hard, and very important in the field of civil engineering, as far as we are concerned, there is no protocol for evaluating and comparing existing methods and it is difficult to known what kind of methods has to be chosen for this task. So, we think that, with the multiple methods proposed in the literature, it is important to evaluate and to compare the various methods in order to validate previous work and to identify the approaches that can be employed and/or the methods that need improvements. So, the third aspect discussed is the introduction of such a protocol.

In consequence, the objectives of this publication are, first, to give a state of the art of existing methods in noninvasive control based on image processing for estimating the quality of the road surface, second, to present our method and, third, to introduce a protocol of evaluation and comparison that allows to highlight the advantages and drawbacks of each method.

2 Automatic Road Crack Detection

In the literature, many papers have introduced approaches to detect thin objects in textured images, like in medical imagery, for the detection of blood vessels [13], satellite imagery, for road network detection [29]. Since 1990, algorithms have been proposed for semi-automatic detection of road cracks (interested readers can see [59] for details about road imaging system and their limits in 1999). For the detection of cracks, three components have to be taken into account:

(1) **Acquisition** (see. Table 2 for details);

(2) **Storage** and

(3) **Image processing**.

In this paper, only the last step is studied but the choices for the two first steps are important for the success of the image treatment. Moreover, most of the references are given in the field of road quality assessment, but, some of them come from different applications, like cracks and defects in concrete (for bridges or pipelines), on ceramics or on metallic surfaces (for industrial applications). For road cracks, most of the time, these hypotheses can be exploited:

(1) *Photometric hypotheses*

- (H_{p_1}) The crack pixels are darker than the road pixels.
- (H_{p_2}) The gray-level distributions of road crack and road surface are distinguishable.

(2) *Geometric hypotheses*

- (H_{g_1}) A crack is a fine continuous object.
- (H_{g_2}) A crack is a set of connected segments that have different orientations.
- (H_{g_3}) A crack does not have a constant width on the whole length.

(3) *Photometric and geometric hypotheses*

- (H_{pg_1}) The points inside a crack can be considered as points of interest, from a photometric and/or geometric point of view.

These different hypotheses can be complementary, like (H_{p_1}) and (H_{p_2}) or (H_{g_1}) and (H_{g_3}), but some of them are opposite, like (H_{g_1}) and (H_{g_2}). The hypothesis (H_{pg_1}) combines two kinds of constraint because, the definition of a point of interest (POI), that is a significant point in a scene, can be expressed both with photometric constraints (the distribution of gray levels near POI has some particularities) and geometric constraints (the point of interest can be a corner, an edge, or any kind of geometric structure).

For image processing, we enumerate semi-automatic and automatic detection approaches, and, five families can be considered, see Table 3:

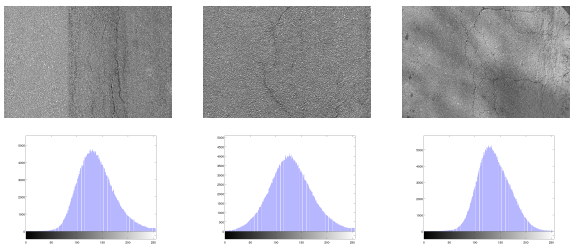


Figure 1: Examples of histograms (second line) of images (first line) with cracks – These images present only one mode in their histograms and it is impossible to separate the gray level distribution of cracks from the gray level distribution of the road pavement.

(1) based on **histogram** analysis (hyptotheses H_{p_1}, H_{p_2} and H_{p_2}): the most ancient and the most popular ones. These methods use a thresholding based on histogram analysis [3, 44, 67], with Gaussian hypotheses [40] and/or adaptive or local thresholding [23, 26]. These approaches are simple and not time consuming, but they also give many false detections. In fact, these methods assume that the two gray level distributions (the road pavement distribution and the crack distribution) can be separated

NAME	COUNTRY	L	CLS	LC	IPT
ADDA, Automated Distress Data Acquisition (1991)	USA			-	
ACM, Automated Crack Monitor (1991)	USA			-	
SIRANO (1991)	France			-	
HARRIS, Highways Agency Road Research Information System (1999)[a]	UK	-		-	
ADA, Automated Distress Analyser (2007)[b]	USA	-		-	
AIGLE RN (2008)[c]	France		-		
AMAC (2004)[d]	France	-	-	-	-
Profilograph and laser (2007)[e]	Danemark	-	-	-	
REAL, Road Excellent Automatic Logging[f] (1992)	Japan	-	-	-	-
RoadCrack (2008)[g]	Australia		-	-	-
ARAN, Automatic Road Analyser (1993) and then FuGro ADVantage (2003)[h]	Canada	-	-		-
PAVUE[j]	Sweden		-		-

[a] http://www.trl.co.uk/facilities/mobile_test_equipment/highways_agency_road_research_information_system.htm

[b] http://www.waylink.com/

[c] http://www.cete-normandie-centre.equipement.gouv.fr/IMG/pdf/15-AigleRNpress_cle243947.pdf

[d] http://www.vectra.fr/suite.php?page=suivante&newsid=253

[e] http://www.greenwood.dk/road.php

[f] http://www.pasco.co.jp/eng/solutions/geospatial/ground/

[g] http://www.csiro.au/solutions/psaa.html

[h] http://www.roadware.com/

[i] http://rst.ramboll.se/en/produkter/ /media/BF0A615A1878494E88DAE424BAA0A144.ashx

Table 1: The quality evaluation systems of the road pavement based on optical devices – L, CLS, LC and IPT mean respectively, that the described system involves a Laser, a Controlled Lighting System, a Linear Camera and/or some Image Processing Tools. The first set corresponds to systems without image processing tools. While the second group of systems are able to provide non-dense measurement, the third group of systems can give a dense evaluation of the quality of the road.

based on a global level statistics (histogram)[3]. In Figure 1, we can see that most of the time, this hypothesis is not valid.

(2) based on **mathematical morphological tools** [6, 21, 26, 34, 35, 52, 66] (hypotheses H_{p_1} and H_{g_1}). An initial thresholding is needed and the results contain less false detections than methods based on histogram analysis. However, the major drawback of this kind of techniques is the strong dependence to the parameter choices.

(3) based on a **learning** phase in order to alleviate the problems of the two first groups of methods [51, 50] (hypotheses H_{p_1} and H_{p_2}). Most of them are based on neural networks [17, 32, 37]. The drawback is the learning step that can not allow a fast and fully automatic analysis.

(4) based on **filtering**, the most recent ones (hypotheses H_{p_1}, H_{g_1} and H_{g_3}). Using edge detections by fixed scale filtering is not adapted to the task of the detection of road cracks because the width of the crack is not constant and this is why many methods are based on wavelet detections [2, 7, 71, 73] with adaptive filtering [12, 13, 65] (these approaches will be detailed in the § 4), contourlets [43], Gabor filters [1], Finite Impulse Response filter (FIR) [24] and methods using models based on partial differential equations (PDE) [5, 49]. Some techniques also use auto-correlation filtering [42, 61] (similarity measures are estimated between the targets that simulate cracks and targets of the original image). An other kind of algorithms is based on texture analysis [54, 63] (the crack is considered as a noise inside a texture).

(5) based on an analysis of a **model** [48, 10] (hypotheses H_{p_1}, $H_{g_1-g_3}$ and H_{pg_1}). These approaches are based on a local analysis versus a global analysis in order to take into account the local constraints and the global constraints of a crack. By multi-scale analysis of texture combined with an algorithm of minimal path [48] or by local detection of point of interest combined with geodesic contours [10].

In conclusion, we can notice that:

- Many methods have been proposed but the problem is not still solved. The result contains many false positives and the detections are incomplete. Moreover, most of the existing techniques can give good results for a specific class of road pavement,

[3] This separation would be possible based on local statistics around the crack.

O.	SENSOR	LASER	R.	CITATIONS
$\perp$	1 (2D sensor)			[3, 37, 40, 18, 30, 36, 45, 16, 66, 68, 73]
$\perp$	1 (2D sensor)	-		[28, 31]
$\perp$	more than one (2D sensors)			[19, 58]
$\perp$	more than one (1D sensors)			[55, 56]
$\perp$	more than one (2D sensors)		-	[70, 71]
$\not\perp$	more than one (2D sensors)		-	[24]

Table 2: The imaging systems – For each kind of configuration, the given details are: the orientation of the sensor (O.) that can be perpendicular or not to the road plane, the number of sensors, the presence (-) or not of a laser, the presence (-) or not of a overlapping images (R.) between the different acquisition. For the sensors, we distinguish 2D sensors (CCD, Charged Couple Device) from 1D sensor (Linear camera coupled with laser).

HISTOGRAM (H_{p_1}, H_{p_2})	LEARNING (H_{p_1}, H_{p_2})	MORPHOLOGY (H_{p_1}, H_{g_1})	FILTERING (H_{p_1}, $H_{g_1-g_3}$)	MODEL-BASED (H_{p_1}, $H_{g_1-g_3}$, H_{pg_1})
[3, 14, 16, 19, 21, 23, 30, 33, 36, 38, 39, 40, 41, 44, 55, 67, 72]	[8, 17, 18, 20, 32, 37, 45, 50, 51, 57, 58]	[6, 15, 25, 26, 34, 35, 47, 56, 64, 66]	[1, 2, 5, 7, 13, 22, 24, 29, 42, 43, 49, 52, 54, 61, 63, 65, 69, 71, 73]	[9, 10, 11, 12, 48, 68, 70]

Table 3: The classification of crack detection (or thin objects in textured images) into five different families. For each family, the hypotheses that are employed are specified.

i.e. the performance of the method is dependent on the road texture.

- Old methods based on histogram studies, even those that are local, do not design correctly the problem, i.e. they do not take into account geometric characteristics of the cracks and photometric characteristics of the road pavement.

- Learning methods are efficient but the learning step is expensive (the time and the investments from the users that are not expert in image processing).

For all these reasons, even if learning methods have been used in our previous work, this paper focuses on the presentation of two methods that try to alleviate the limits of the old ones: we try to obtain a denser detection with a low rate of false detections.

3 Proposed and Compared Methods

Some preliminary works about methods adapted to this task (the detection of road cracks), have included experiences on a learning method. A neuron-based method has been tested [45], on the real images of size 768×512 presented in § 4.2. Results are interesting but learning methods are not easy to use for non-specialist in image processing, and, they cost a lot of time to the users. The main goal is to propose a system that facilitates the work of users and not a system that induces a lost of time by including a learning phase, and a maintenance each year in order to maintain the performances of the system[4]. In consequence, we have now focused our work on methods that allow automatic processing and, in particular, we present two approaches:

(1) The first, *Morph*, belongs to the families (1) and (2) because it combines thresholding and refinement by morphological analysis.

(2) The second, *GaMM*, of families (4) and (5), is based of the advantage of multi-scale analysis and local modelling of the crack.

Morph has been proposed before *GaMM* and is quite near the method presented in [66]. The contributions of this section are about *GaMM*, we propose a new model for the sites and the potentials used in the Markovian model. The advantages of this new method will be illustrated with qualitative and quantitative results in § 5.

3.1 Morphological method (*Morph*)

The chosen approach is based on hypotheses H_{p_1}, H_{g_1} and H_{g_3} and it follows these steps:

(1) Pre-processing of the images: to reduce the texture and increase the contrast between the road pavement and the crack;

(2) Binarization by thresholding (the threshold is different in the various variant and a local threshold can be used);

[4] This maintenance is necessary because conditions and systems of acquisition can changed every year, and the road pavement also evolves.

(3) Refinement by closing;

(4) Segmentation with shape analysis;

(5) Extraction of the crack characteristics.

For step 1, three variants are introduced by combining these local tools: an erosion in gray levels, a conditional median filtering, a histogram equalization, a mean filtering (these pre-processings are detailed in § 5.1). The step 4 is realised in two passes: first a connected components labeling is realized and, second, the size and the shape of each component is determined in order to remove components where shape is not similar to a crack: the shape of a crack has to look like a thin object, the width, w, and the height, h, are used for this task. More precisely, from an expert point of view, a crack is not significant if $h < 50cm$ but, we can suppose that we manage to detect only a small part of the crack and this constraint becomes $h < 7.5mm$. Moreover, the mean width $w_{min} < 3mm$ and the maximal width $w_{max} < 6.5mm$. All these thresholds are empirically set. In Figure 2, we illustrate the kind of results obtained at each step for the 3 variants. The final proposed method, named *Morph*, merges the 3 results (with a weighted sum and the weights are chosen with a learning phase) and refines the result by computing the closing in gray levels of the fusion result.

3.2 Adaptive Filtering and Markovian Modelling (*GaMM*)

More recently, our work focused on the field of wavelet decomposition. As it is difficult to chose the mother wavelet (useful for generating the wavelet family for multi-scale analysis) well adapted to the detection of road cracks, the adaptive filter theory seems convenient and, in particular, it allows to build a mother wavelet adapted to our task. First, we present the first step of the algorithm based on adaptive filtering (hypotheses H_{p_1} and H_{g_3}) and the second on Markovian segmentation that can take into account the particular geometry of the crack (H_{g_2} and H_{g_3}).

3.2.1 Algorithm

The goal of this algorithm, presented in Figure 3, is to obtain, step 1, a binarization (black pixels for background and white pixels for the crack) and a refinement of this detection by using a Markovian segmentation, step 2. Using adaptive filtering is important in order to allow the detection of non-constant width of the crack (which is realistic) (hypothesis H_{g_3}). The number of scales for the adaptive 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. The adaptive filtering is applied in each scale, each directions and then all the results are merged on each scale (mean of the coefficients). The results of this filtering is used to initialize the Markovian modelling used for the segmentation step.

```
Input
Road images
-----------------------------------------
Initialization
Number of scales and angles
-----------------------------------------
Steps
1) For each scale do
1a) For each direction do
      Estimate Adaptive Filter (AF)
1b) Merge AF in all the directions
2) For each scale do
2a) Initialization of the sites (Markov)
2b) While not (stop condition) do
      Updating of the sites
3) Fusion of the results on each scale
```

Figure 3: The studied algorithm for the method based on adaptive filtering and segmentation by Markovian modelling – The step 1 leads to a binary image using adaptive filtering, while step 2 refines this result with a Markovian modelling.

3.2.2 Adaptive filtering

Some details are provided in order to realize the step 1a and 1b in Figure 3. The $\psi \in \mathcal{L}^2(\mathbb{R}^2)$[5] function is a wavelet if:

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

where Ψ is the Fourier transform of ψ. 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 :

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

where R^θ is a rotation of angle θ.

One of the main difficulties to apply a wavelet decomposition is the choice of the mother wavelet ψ.

[5] $\mathcal{L}^2$ is the square integrable space.

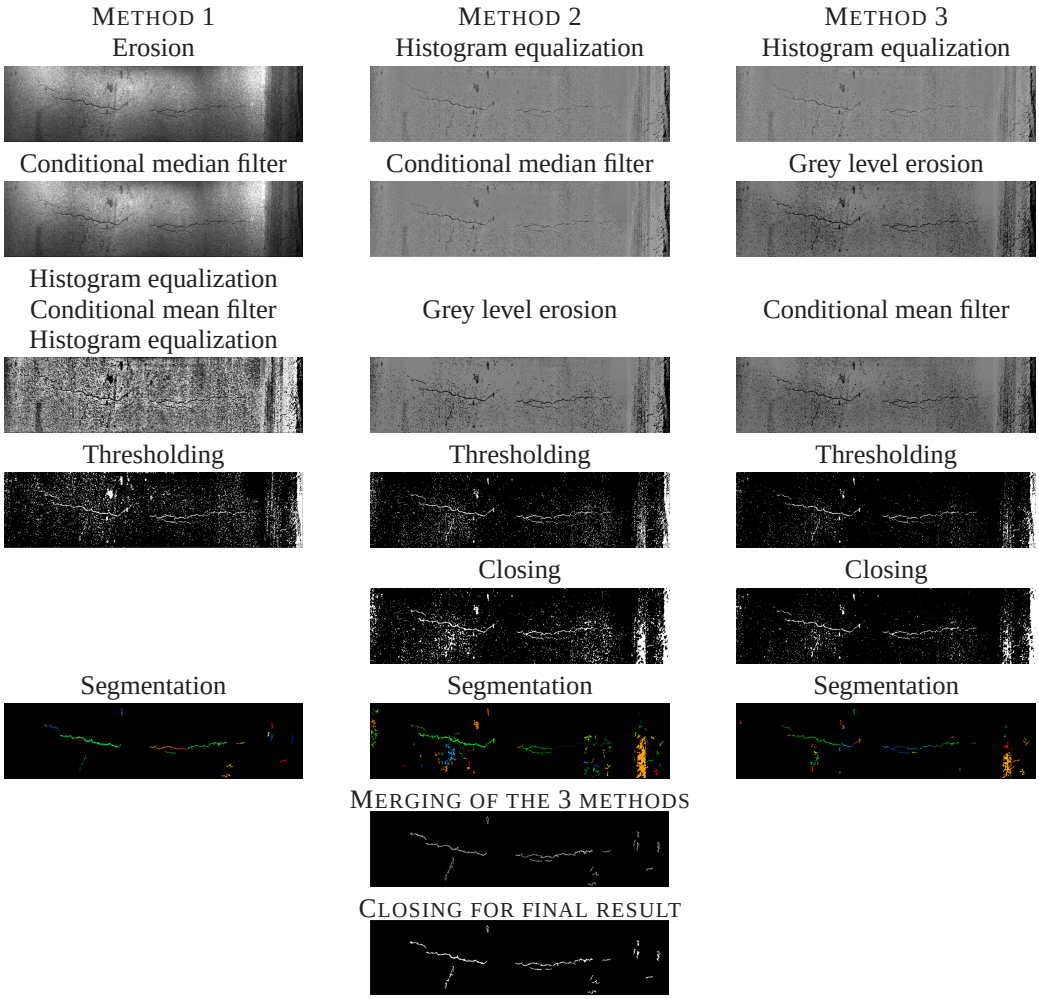


Figure 2: The different steps of the method *morph* – The conditional filtering is applied when the gray level is higher than 40 (to prevent the removal of the crack). The last step is proposed in order to reduce false detections and to complete the detection.

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 (its shape is known up to a factor) mixed with a noise whose characteristics are known. Consequently, adaptive filtering is well designed for the problem: extracting singularities in coefficients estimated by a wavelet transform. If $\mathbf{s}$ is a discrete and deterministic signal with $\mathbf{s} = \begin{pmatrix} s_1 & \dots & s_N \end{pmatrix}$, N the number of samples, and $\mathbf{z} = \begin{pmatrix} z_1 & \dots & z_N \end{pmatrix}$, is a noisy observation of $\mathbf{s}$, $\mathbf{b}$ is 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 the auto-correlation function ϕ_{bb} of terms $\phi_{bb_{(i,j)}} = \phi_{bb_{|i-j|}}$, independent of the signal $\mathbf{s}$. The adaptive 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 in most of the papers of this domain, crack pixels correspond to black pixels surrounded by background pixels (road pixels). This is why, in [65], a crack is a piecewise constant function f, defined for each position $x \in \mathbb{R}$ by:

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

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 σ, the deviation of the Gaussian law, i.e. $a = \frac{1}{\sigma\sqrt{(2\pi)}}$. Consequently, the term σ allows to fix the width of the crack (like threshold T in equation (4)). Finally, for the step 1, **h** is estimated for each size of signal (determined by σ) for the 5 scales, as explaines in the beginning of § 3.2.1, and ϕ_{bb} is interpolated in order to have the same size. Then the filter is transformed by rotation in order to cover the 4 orientations.

3.2.3 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 (step 2a of the algorithm of the Figure 3). For the first step of segmentation (initialization), the sites are of size 3×3, consequently, a regular grid is considered in the image. In [65], four configurations are possible and represented in Figure 4 (the part inside the rectangle with low gray levels). The initialization of the sites is based on the configuration that maximizes the coefficients obtained with the adaptive filtering. More formally, if we denoted $\gamma_{2,0}$, $\gamma_{2,\frac{\pi}{4}}$, $\gamma_{2,\frac{\pi}{2}}$ and $\gamma_{2,\frac{3\pi}{4}}$, the four configurations, the best configuration γ_{best} is:

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

where $m_{2,\alpha}$ is the mean of the coefficients on the considered configuration $\gamma_{2,\alpha}$. These four configurations do not represent all the possibilities 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 the set of sixteen configurations illustrated in Figure 4 (all the presented sites). By modifying the number of configurations, we need to adapt the initialization of sites and equation (6) becomes:

$$\gamma_{\text{best}} = \underset{i \in [0;4],\quad \alpha \in \left[0,\ldots,\frac{3\pi}{4}\right]}{\operatorname{argmax}}\ m_{i,\alpha}, \tag{7}$$

where $m_{i,\alpha}$ is the mean of the coefficients on the considered configuration $\gamma_{i,\alpha}$.

The image is considered as a finite set of **sites** denoted $\mathcal{S} = \{s_1, \ldots, s_N\}$. 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 $\mathcal{S}$ whose every pair of distinct sites are neighbors.

These random fields are considered:

(1) The observation field $Y = \{y_s\}$ with $s \in \mathcal{S}$. Here, y_s is the mean of the coefficients on the site.

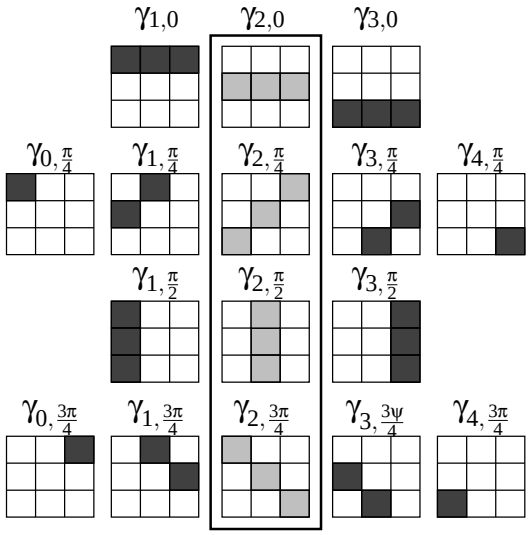


Figure 4: The sixteen configurations in order to improve the modeling of sites – The four initial configurations proposed in [65] are in the bold rectangle, the sites are represented by the clearer gray levels, and for the proposed configurations the sites are represented by the darker gray levels.

(2) The descriptor field $L = \{l_s\}$ with $s \in \mathcal{S}$. If there is a crack $l_s = 1$ elsewhere $l_s = 0$.

At 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 coefficients of the sites (computed from the coefficients estimated during the first part of the algorithm: adaptive filtering) and the configurations of each site and its neighbor sites (the 8 neighbors). More formally, the global cost is the sum of all the potential functions of the sites. This potential function contains two terms:

$$u_s = \alpha_1 u_1(s) + (1 - \alpha_1) \sum_{s' \in \mathcal{V}_s} u_2(s, s'), \tag{8}$$

The first term, u_1, corresponds to the **data term**, and it evaluates how a site is similar to a crack from a photometric point of view (hypotheses H_{p_1} and H_{p_2}). This term is based on the results given by the adaptive filtering. The second term, u_2, represents the **constraints** induced by the neighbors of the site. More precisely, it estimates the consistency between a site and each neighbor site and it takes into account the geometric hypotheses H_{g_2} and H_{g_3}. The choice of the value α_1 depends on the importance of each part of the equation (8) and it will be discussed in § 5.1.1.

The function u_1 is given by:

$$u_1(y_s, l_s = 1) = \begin{cases} e^{\xi_1 (k - y_s)^2} & \text{If } y_s \geq k \\ 1 & \text{elsewhere} \end{cases} \text{ and}$$
$$u_1(y_s, l_s = 0) = \begin{cases} e^{\xi_2 (y_s - k)^2} & \text{If } y_s < k \\ 1 & \text{elsewhere,} \end{cases} \quad (9)$$

The parameters ξ_1, ξ_2 and k have to be fixed[6]. For the definition of u_2, we have to determine the number of cliques. In [65], 4 cliques are possible and the 8-connexity is considered. 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 4. Some cases are not penalized with the

	$\gamma_{2,0}$	$\gamma_{2,\frac{\pi}{4}}$	$\gamma_{2,\frac{\pi}{2}}$	$\gamma_{2,\frac{3\pi}{4}}$
$\gamma_{2,0}$	β_1	β_2	β_3	β_2
$\gamma_{2,\frac{\pi}{4}}$	β_2	β_1	β_2	β_3
$\gamma_{2,\frac{\pi}{2}}$	β_3	β_2	β_1	β_2
$\gamma_{2,\frac{3\pi}{4}}$	β_2	β_3	β_2	β_1

Table 4: The function u_2 used in [65] – This table presents the values $u_2(s', s)$ for the sites in low gray levels in Figure 4. In our experiments, like the authors, we have chosen $\beta_1 = -2$, $\beta_2 = -1$ and $\beta_3 = 2$.

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 configurations that are presented in Figure 4, the potential has to take into account the differences of orientations between two sites (there are 16×16 possibilities) and the position of the two sites (there is 8 possibilities because we consider the 8 neighbors). 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, the distance implies the minimal distance between the extremities of the two segments).

[6]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 experiments have brought us to take ξ_1=ξ_2=100.

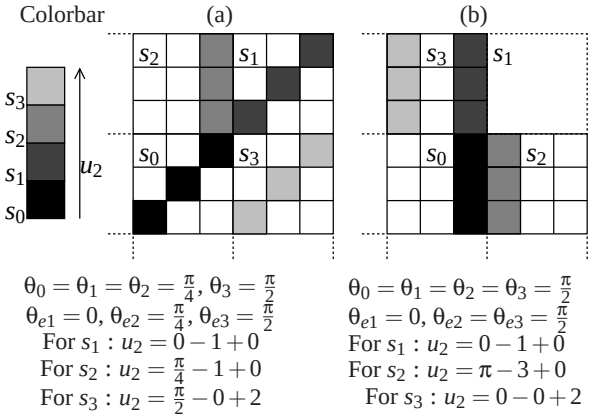


Figure 5: Examples of the function u_2 – These two examples of sites with their respective neighbors show the behavior of the potential u_2 with the two considered aspects: orientation and distance. In example (a), with the help of the orientation term, the configuration s_3 is penalized and s_2 is less penalized than s_3. In example (b), with the help of the two terms about the distance, the site s_3 is penalized, compared to s_1. On the contrary, the particular case of s_2 is favorable and it compensates the penalty given by the orientations.

More formally, if:

- d denotes the Euclidean distance between the two closest extremities of the sites, with $d \in [0, d_{\max}]$[7];
- θ_1 and θ_2 are the orientations of respectively $s = \{p_i\}_{i=1..N_s}$, and $s' = \{s'_j\}_{j=1..N_{p'}}$ where, p_i, respectively p'_i, is the pixel i of the N_s, respectively N'_s, pixels that composes the site s, respectively s';
- θ_e is the angle between the two sites;

the u_2 function is defined by:

$$u_2(s', s) = \alpha_2 \left(\frac{|2\theta_e - \theta_1 - \theta_2|}{2\pi} \right) + (1 - \alpha_2) \left(\frac{J(\text{NbC}) \min_{i,j}(d(s_i, s'_j))}{d_{\max}} - \frac{\text{NbC}}{3} \right). \quad (10)$$

where NbC indicates the number of connected pixels between the two sites s and s' and $J(x)$ equals 1 if $x = 0$ and 0 elsewhere. The first term is induced by the rule about the orientations, (R_1). This term equals zero when the sites have the same orientation and this orientation is the same as the orientation between the sites, i.e. $\theta_e = \theta_2 = \theta_1$. This first term penalizes the configurations where the sites do not have the same orientation but also the particular case where they are parallel, see example (a) in Figure 5. The second term and the third term express the rule (R_2) about the distances. Two aspects have to be distinguished: the

[7]As the sites are of size 3×3, $d_{\max} = 5\sqrt{2}$)

number of connected pixels, when the sites are connected, and, on the contrary, i.e. when the sites are not connected, the distance between the sites. It allows to give low influence at disconnected sites and also to increase the cost of sites that are parallel but connected, see example (b) in Figure 5. To study the influence of all these terms, the equation has been normalized and the different terms have been weighted (using α_2, the choices for α_2 will be discussed in § 5.1.1).

4 Evaluation protocol

For the evaluation of image processing methods, nowadays, there is no protocol used by the community of road pavement analysis. However, all the approaches for characterizing the quality of the road advise to take into account the severity of the defect. For this, we need to know precisely the size, the width and the location of the cracks. This is why, now, it seems important to precisely evaluate the performances of all these automatic existing methods. For designing this protocol, we have to determine which kind of images have to be tested and which kind of criteria of efficiency can be used. In the community, to the best of our knowledge, there is no reference images and most of the time, the evaluation of the performances is qualitative.

For the evaluation of automatic crack detection, to the best of our knowledge, no evaluation and comparison protocol has been 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 defects, i.e. to detect precisely the defect. This is why, it seems important to characterize quantitatively the performances 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. For the reference segmentations, two approaches can be used to estimate them:

(1) *To compute synthetic images with synthetic defects* – In this case, the exact position of the defects is known and these reference segmentations can be considered as ground truth.

(2) *To propose reliable segmentations of real images* – It supposes that we are able to provide a segmentation that is reliable enough to be employed as a reference. For evaluation, these segmentation can be called “pseudo-ground truth”.

The two solutions are studied, and, this is why we have to explain how the manual segmentations (that are our references) are done. Before, we briefly describe the acquisition system.

4.1 Acquisition

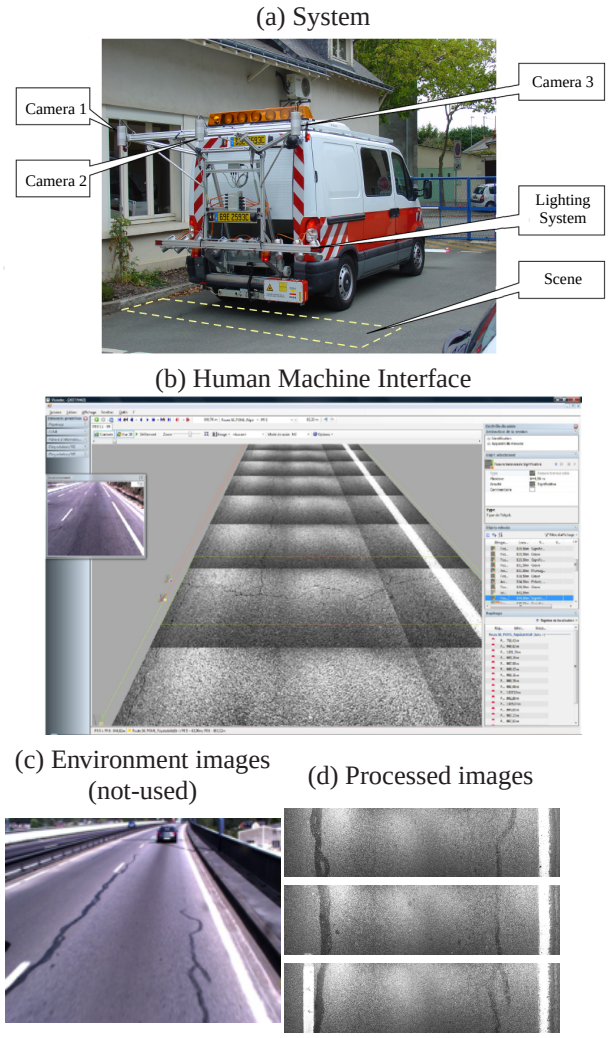


Figure 6: The acquisition system used for the evaluation – In (a), The acquisition system is illustrated. In (c) and (d), an example of the final images is given. In (d), we can see the road that is visible in (c). The processing is done 1 meter by 1 meter, i.e. independently on each image presented in (d). The surface contains two reparations of vertical cracks. In some cases, the sensors do not have the same settings and the global illuminations are different, so, it can generate some “false cracks”. This aspect has been easily taken into account in a pre-processing step by eliminating the junction area in the region of interest.

The acquisition system used for the dataset of our experiments is described in Figure 6. It contains 4 video cameras with 3 sensors in gray levels in the backside of the car and 1 color one in front of the car. The first camera is needed to determine the environment conditions (weather, location, traffic) whereas the three other ones are used for the crack detection.

The resolution of this one is smaller than the 3 others and moreover, the optical axis is not perpendicular to the road surfaces, on the contrary of the 3 others. The 3 cameras have been physically synchronized directly during the acquisition. To be independent of the illumination problems, nine stroboscopic lights have been added. The position of the lights is perpendicular to the road plane and distant of 1 meter from the surface. The light power has been chosen in order to not deteriorate the visualisation of the road pavement and of the defects.

4.2 Reference images

The most difficult is to propose images with a reference segmentation. On the first hand, we introduce synthetic images with a simulated crack (the size of these images are 256×256)[8] As shown in Figure 7, the result is not realistic enough. It does not seem realistic because the contrast is too important between the road and the crack. Moreover, the interuptions of the crack, the changes in the direction, the presence of many paths, etc, in the default, are no simulated. In order to be more realistic, it seems that we have to design and to implement a complex heuristic to simulate the crack and it is too much effort for finally having only a synthetic default. This is why, on the second hand, we have simulated different defects on real images that previously contain no defect (the size of these images are 768×512 and 1920×480). The result is more realistic but the shape and the photometric aspect of the cracks (which are randomly chosen) does not seem realistic enough. This is why, it appears important to propose a set of real images (size 768×512 and 1920×480) with manual segmentations that are reliable enough to be considered as reference segmentations. To summarize, the two first kinds of images allow to propose an exact evaluation and to illustrate theoretically the behavior of the method whereas the last kind of images allows to validate the work on real images with a pseudo-ground truth.

4.3 Reference segmentations

For real images, we briefly explain how the manual segmentations are validated. Four experts have manually segmented the images with the same tools[9] 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 squared structuring element of size 3×3. To evaluate the reliability of the reference segmentations, we estimate, first, the percentage of covering between each operator, and, second, the mean distance, D, between each pixel (detected by only one expert and not kept in the

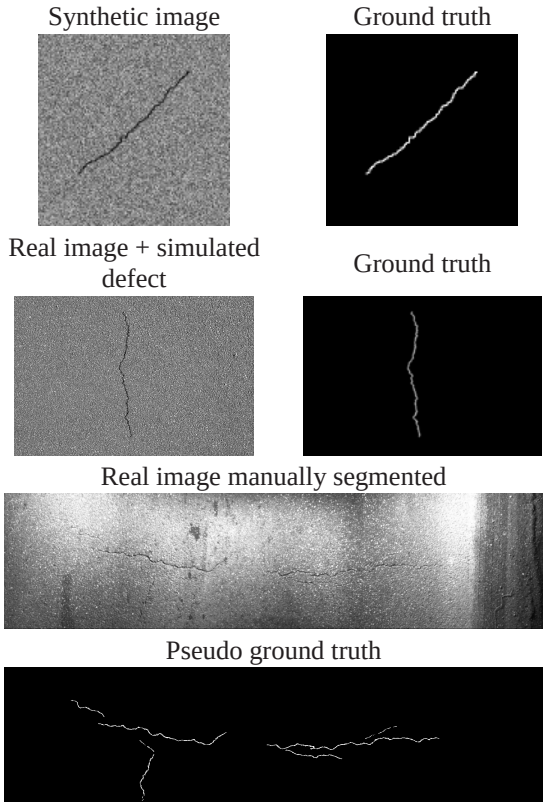


Figure 7: Examples of the images tested.

[8] The road is a random texture, i.e. each intensity is randomly chosen by supposing a uniform distribution of intensities in $[0;255]$. Then, the user gives the position of the beginning, the length and the orientation (vertical, horizontal or oblique) of the crack. The crack points are built by randomly selecting the next point in the neighborhood and the intensity in $[0;100]$.

[9] We use a "home-made" software that proposes an interface that helps the person to segment the default. The principle is that the user has to select points on the crack. These points have to be close enough (from 5 to 20 pixels of distance). Then, the path between two close points is automatically detected by using a simple heuristic: the path that minimizes the mean intensity is selected. The interface is complete enough to allow the displacement of the points, the removing of some points, the removing of some cracks. The user can also select the width of the path (crack). Some filters are also proposed to improve the contrast between the crack and the road in order to help the user.

reference image) and the reference segmentation.

Table 5 shows some results for 5 of the 42 images manually segmented. We have distinguished 5 families: the first one contains images acquired in static whereas the four other ones are acquired in dynamic. Moreover, we have 4 different kinds of road pavement acquired in dynamic. The 10 images have been taken in order to show results on each of these families. We can notice that the first 4 images are the most reliable because the mean error is less than 2 pixels. The precision of these results is satisfactory. On the contrary, the last 6 images show the important variabilities between operators and how it is difficult to extract a segmentation for these images, and, in particular, in the image 936, where the error is due to a bad interpretation of one of the four operators who finds a defect that does not exist.

IMAGES	F (%)	2 (%)	3 (%)	4 (%)	S (%)	D (PIX)
42 	0.4	26.69	14.59	4.2	45.48	1.45
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
88 	1.44	22.74	8.23	1.23	32.2	2.76

Table 5: The comparison of the 4 manual segmentations for estimating the final reference segmentations – For each image, are presented: the percentage of pixels in the whole image that are preserved as crack pixels in the final reference segmentation (F), the percentage of recovering between 2, 3 and 4 manual segmentations and the sum of these 3 percentages (S). For all the crack pixels that are not preserved in the final reference segmentation, the mean distance to this segmentation is given (D).

By analyzing the results for the criterion D, presented in Table 5, we can classify the 42 tested images in 3 categories, i.e. images with:

(1) A **reliable** segmentation: The criterion $D < T_r$. It means that all the operators have built segmentations that are quite near to each other.

(2) A segmentation that is **moderately reliable** : The criterion $T_r \leq D < T_a$. It means that some parts of the crack are not easy to segment and there are local errors.

(3) An **ambiguous segmentation:** The criterion $D \geq T_r$. It clearly shows that the images are difficult to segment and in most of the cases, it means that some parts are detected as a crack whereas they are not and reversely.

The threshold have been empirically chosen and $T_r = 2$, $T_a = 4$. In Figure 8, first, we present the mean distance, D_i, $i \in \{1,2,3,4\}$, between the final reference segmentation, S_r, and each manual segmentation, S_i (obtained for the four operators) and, second, the criterion D for each real images of our protocol. The first graph illustrates how it is important to combine the four manual segmentations instead of using only one manual segmentation. Indeed, we can notice that each operator, alternately, gives an interpretation that is different from the three others. The second graph explains how the thresholds are chosen for determining the detections that are "accepted" for the evaluation, see § 4.4 for explanations about accepted detections.

All these remarks are illustrated in Figure 9. Overall, the four segmentations are near to each other and if the segmentations are combined it permits to detect the width of the crack. However, these examples also present the difficulties of this task: areas where the cracks are less visible, the texture elements that have the same size and/or the same gray levels as the crack pixels. Thus, in some cases, one operator extends the crack or gives a different shape. In some extreme cases, the operator can even confound a crack and an other object of the scene (a piece of wood, for example). In another way, these last examples highlight the interest in combining different segmentations in order to obtain reference segmentations as reliable as possible.

4.4 Criteria of efficiency

In this section, we introduce how the reference segmentation and the estimated segmentation are compared. In Figure 10, we present common evaluation criteria that are used for segmentation evaluation:

(1) the percentage of correct detections (true positives) (TP) ;

(2) the percentage of false positives (FP) ;

(3) the percentage of false negatives (FN) and

(4) the similarity coefficient (DICE).

This last criterion seems to be the most significant because it evaluates the ratio between the FP and the FN, and it well resumes the results of all the criteria. Moreover, it directly expresses what is important to evaluate: how the evaluated method can reduce errors of detection whilst increasing the density of good detection.

For real images, the detections that are "accepted" have been added in order to tolerate 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 5, 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

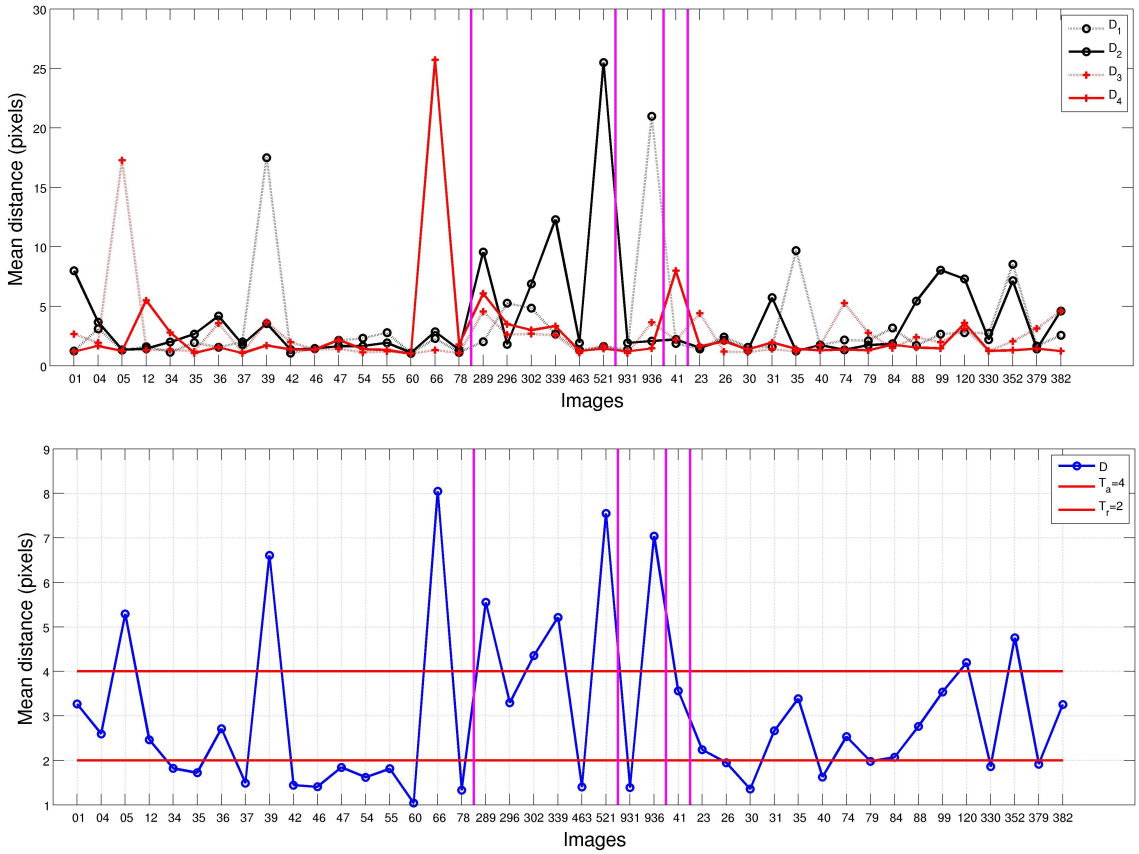


Figure 8: The variations between each manual segmentations that are used for building pseudo-ground truth – The first graph represents for each operator (one curve for one operator) and each image (x-coordinates), cf. Figure 17 for the corresponding images, the mean distance, D_i, $i \in \{1,2,3,4\}$, whereas, the second graph, presents D, cf. s§ 4.3 and Table 5. This graph allows us to distinguish the different categories of images (red axes): **reliable** ($D < T_r$), **moderately reliable** ($T_r \leq D < T_a$), **ambiguous** ($T_a \leq D$). In the two graphs, the purple axes represent the five different samples of images (each sample corresponds to a kind of road pavement). The first ones were acquired with a static system whereas the four others were acquired with a dynamic system.

pixels equals 0 for synthetic images whereas it depends on the mean distances, see D in table 5, for the real images.

5 Experimental results

In this section, two aspects are studied:

(1) the evaluation of the method based on an adaptive filtering and a Markovian modelling in order to characterize its behavior, to estimate the best parameters and to determine the best variant ;

(2) the comparison to the *Morph* method.

5.1 Adaptive filtering and Markovian Modelling

We want to determine, first, how to fix the different parameters, second, the pre-processing steps that are necessary, and, finally, which variant is the most efficient. In consequence, these points have been studied:

- *Parameter values* – The weights α_1, equation (8), and α_2, equation (10), 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:

(1) *Threshold* – This pre-processing has been proposed in order to reduce the light halo in the six last images presented in Figure 7 and in all the im-

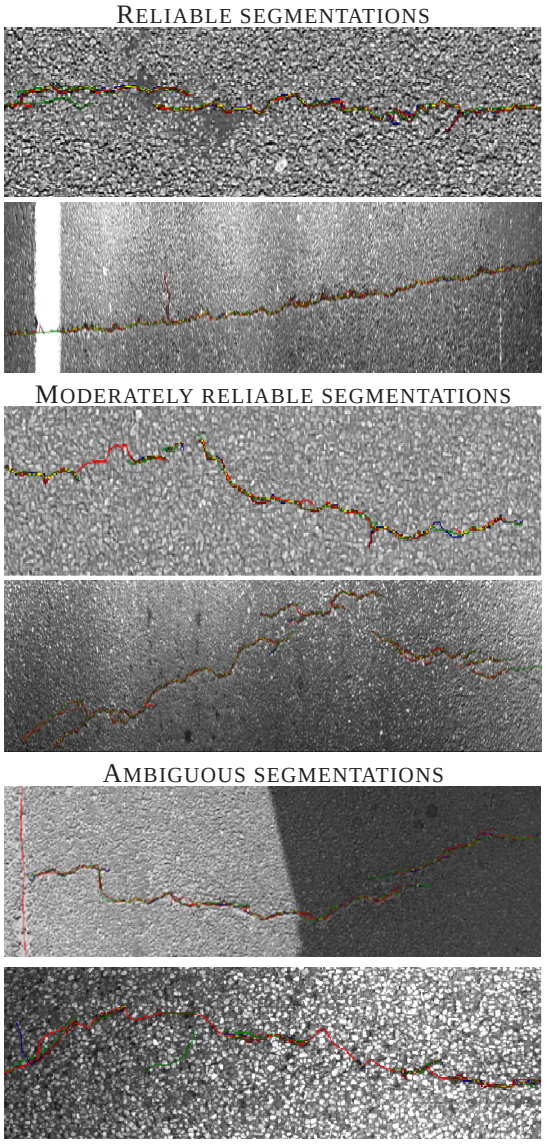
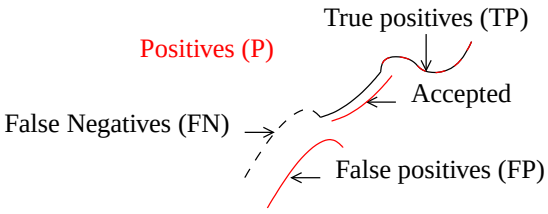


Figure 9: The levels of difficulties of the tested images – On these images, we present the four examples of manual segmentations. The codes are: red (light and dark), blue, green, for each of the four operators. The parts in yellow correspond to the parts of the cracks detected by more than one operator. There are two examples per category of segmentations. To better visualize, only a part of the images is shown. In the ambiguous images, we clearly see the mistake of the operator in dark red.

ages of the four last categories in Figure 8. Each pixel lower than a given threshold is replaced by the local average of the gray levels.

(2) *Smoothing* – A mean filter of size 3×3 is applied to reduce the granularity of the texture.

(3) *Erosion* – An erosion (in gray level) with a square structuring element of size 3×3 is also applied to reduce the granularity of the texture.



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

Figure 10: The evaluation criteria – In this Figure, two segmentations of the same crack are represented: the black line corresponds to a manual (or reference) segmentation and the red line is an estimated segmentation. 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). The piece of line with the two colors (red and black) are the correct detections or True Positives (TP). In this table, different criteria are introduced but, 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, in order to determine how to reduce false detections whilst increasing the density.

(4) *Restoration* – It combines the advantages of all the previous methods in three steps: a histogram equalization, a thresholding (like *Threshold*), and an erosion (like *Erosion*).

In order to preserve the crack signal, each pixel under a given threshold is not filtered[10].

• *Algorithm variants* – Four variants are compared:

(1) *Init* – This is the initial method proposed in [65].

(2) *Gaus* – This variant supposes that the distribution of the gray levels inside a crack follows a Gaussian function, see § 3.2.2.

(3) *InMM* – This is the initial version with an improvement of the Markovian modelling (new definition of the sites and of the potential function), see § 3.2.3.

(4) *GaMM* – This is the method *Gaus* with the new Markovian modelling.

• *Comparison* – We have compared this method with the method based on morphological tools and that is quite similar to [66], *Morph*.

[10] Experimentally, this threshold equals 40.

5.1.1 Influence of parameters

Among the results, two conclusions can be done:

(1) For each variant and each pre-processing, the weights between the term for adaptive filtering and the term for the Markovian modelling should be the same, see equation (8), i.e. $\alpha_1 = 0.5$. However, when more weight is given to adaptive 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 information. It seems coherent with the difficulties of the acquisition.

(2) For the Markovian modelling, 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 (10), 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. It implies that the orientation characteristics are more reliable than the distance ones and this remark is coherent with the fact that cracks present strong spatial constraints. Moreover, it is also linked with the difficulties induced by the acquisition (the lighting system makes the photometric information less reliable).

5.1.2 Pre-processing

These tests have been done with real images, because, the 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 Markovian modelling, the pre-processing step does not significantly increase the quality of the results.

5.1.3 Variants

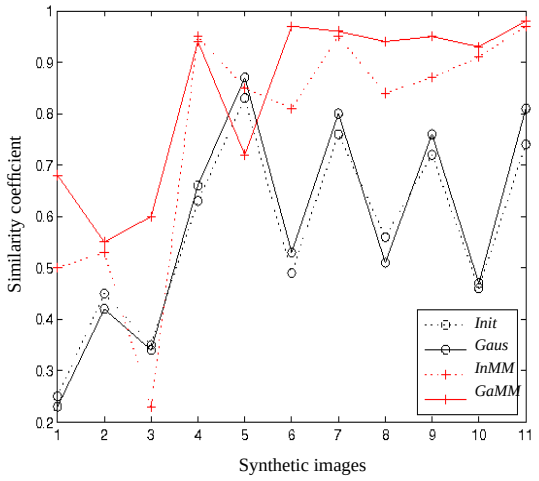
The results are presented for two different cases:

(1) with synthetic images and

(2) with real images.

For the first category, the ground truth is available whereas for the second category, a pseudo-ground truth is used and the detections which are accepted are taken into account in the evaluation, i.e. a threshold is applied to the distance between the segmentation estimated by the evaluated method and the pseudo-ground truth segmentation. The thresholds applied on the distance for the accepted detections are determined with the results given in Table 5, column D.

(a)



(b)

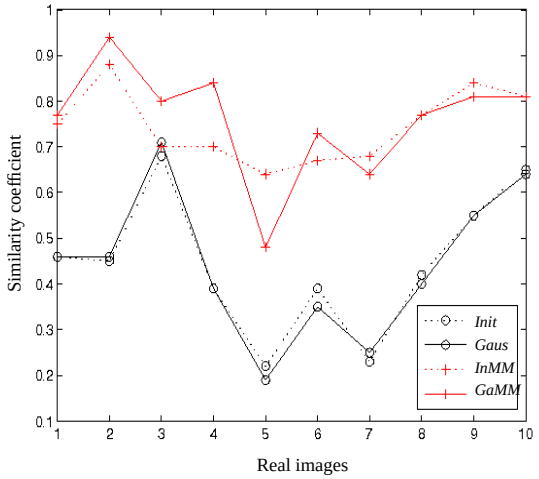


Figure 11: The variations of the similarity coefficients, see Figure 10, for the 4 variants – The first graph shows the results for synthetic images (the 3 first ones are obtained from real images with simulated defect) and the second graph presents the results with real images. The good performances of the methods *InMM* and *GaMM* can be noticed.

In Figure 11, the evolution of the similarity coefficients, or DICE, for the 11 synthetic images, 11.(a) and 10 of the real images is presented, 11.(b). With synthetic images, the method *GaMM* is clearly the best for most of the images. However, for one image (the fifth), the results are worse than the results of the method *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), the method *GaMM* obtains acceptable results (DICE $>$ 0.5) whereas the other methods are not efficient at all. Illustrations are given in Figures 12 and 13: they show how the method *GaMM* can reduce false detections.

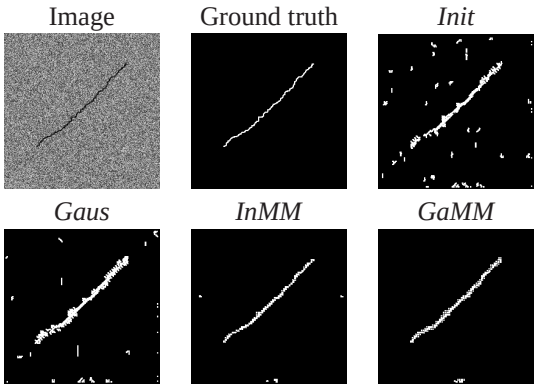


Figure 12: The segmentation results on some synthetic images presented in Figure 7 – These are the results obtained with the four variants and these illustrations show how the method *GaMM* gives the clearest result. We can also notice the good results of the method *InMM*.

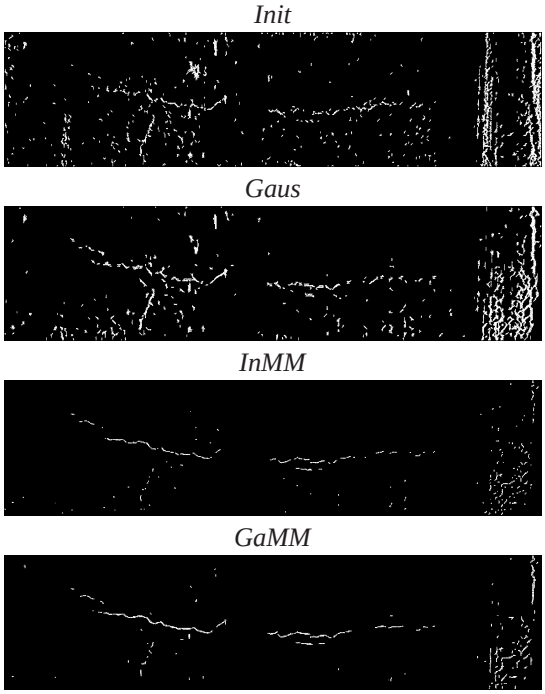


Figure 13: The segmentation results on some real images – These are the results obtained with the real images presented in Figure 7. The method *InMM* obtains the clearest detection (i.e. with less false detection) but we can also noticed the good quality of the detection map with the method *GaMM*.

5.2 Results and comparison with *Morph*

Finally, we have compared the results of *GaMM* on each of the complementary dataset (32 images) with *Morph*. The mean DICE is 0.6 with *GaMM* whereas it is 0.49 with *Morph*, see Figure 14. It shows how *GaMM* can outperform *Morph*. 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 15 and 16. More precisely, *GaMM* seems more efficient with *ambiguous* images, whereas *Morph* is the best with *reliable* images. Finally, we can also precise the execution time for the two methods: about 1 minute for *GaMM* and 5 seconds for *Morph* with a processor intel®core^{TM}2 duo of 2 GHz). These execution times give only some indications because the implementation, in particular for *GaMM*, has not been optimized.

6 Conclusions

In conclusion, this paper gives a review about image processing methods for the crack detection of road pavement. It can help the researchers who want to choose and to adapt an auscultation method to the constraints of the transport structure that is studied (it depends on the quality of the surface, the needs of the auscultation). Moreover, a new method for the detection of road cracks has been introduced and we have presented a new evaluation and comparison protocol for automatic detection of road cracks. As far as we are concerned, we proposed real images with ground truth for the first time in the community. The new method, *GaMM*, has been validated by the proposed protocol and compared to a previous one, *Morph*. This evaluation shows the complementarity of the two methods: the *Morph* method obtains more true positives than the *GaMM* method whereas this ones reduces the percentage of false positives.

Our first improvements of this work will focus on the evaluation and comparison protocol. 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 quality of the road texture). In a second step, our future work will include new experiments about the acquisition system. Indeed, the acquisitions and the results obtained with the acquisition system presented have shown its limits, for example, in Figure 7, some parts of the crack are not ”visible”. It comes from the fact that to highlight the crack, it depends on the orientation of the lights and of the sen-

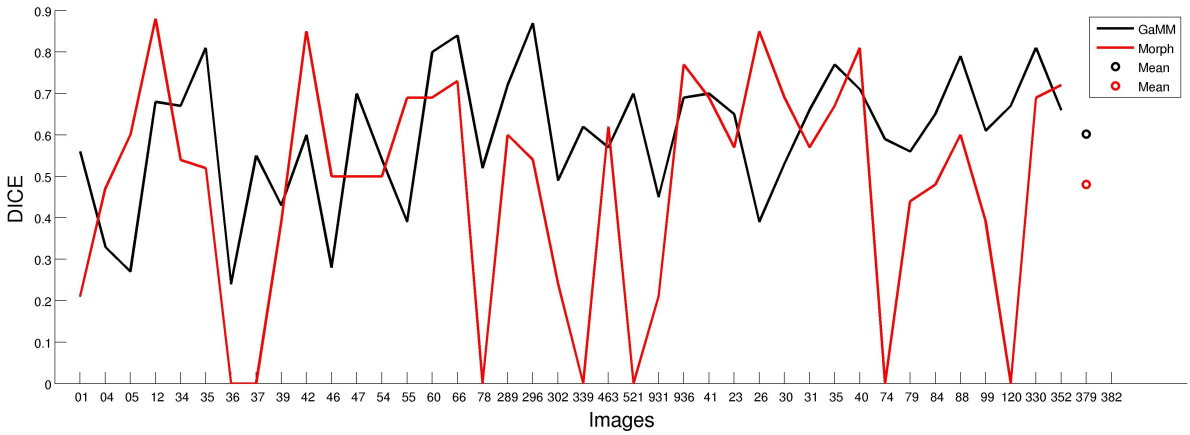


Figure 14: The comparison of the similarity coefficients between *GaMM* and *Morph* – The dotted lines illustrate the five sets of tested images. For the first set that corresponds to real images with no illumination problems, the results are mixed whereas for the four other sets, *GaMM* is the best. The mean of this criterion is 0.6 (variance = 0.0257) for *GaMM* whereas it is 0.49 (variance = 0.0750) for *Morph*. However, this method has one step of characterization of the cracks (not introduced in *GaMM*) and this step can remove cracks that do not respect the characteristics of a cracks (in length, size and shape). This step contributes to reduce errors, but, in some difficult cases, it decreases the performances of the detection, compared to *GaMM*.

sors. Using one single sensor and one light always in the same position/orientation, we can sometimes miss some defaults in the acquisition. So, it seems important to study other kinds of system to improve the quality of the automatic treatments.

Third, we want to improve the *GaMM* method, to begin, by adding the extraction of the crack characteristics, like in *Morph*.

REFERENCES

[1] F. S. Abas and K. Martinez. Classification of painting cracks for content-based analysis. In *SPIE, Annual Symposium Electronic Imaging : Machine Vision Applications in Industrial Inspection XI*, volume 5011, pages 149–160, 2003.

[2] I. Abdel-Qader, O. Abudayyeh, and M. E. Kelly. Analysis of Edge-Detection Techniques for Crack Identification in Bridges. *Journal of Transportation Engineering*, 17(4):255–263, 2003.

[3] J. J. Acosta, L. Adolfo, and R. L. Mullen. Low-Cost Video Image Processing System for Evaluating Pavement Surface Distress. *Journal of the Transportation Research Board (Transportation Research Record)*, 1348:63–72, 1992.

[4] Y. Adachi. *Sensing Issues in Civil Structural Health Monitoring*, chapter Monitoring Technologies for Maintenance and Management of Urban Highways in Japan – A Case of Hanshin Expressway Public Corporation, pages 13–22. springer, 2005.

[5] B. Augereau, B. Tremblais, M. Khoudeir, and V. Legeay. A Differential Approach for Fissures Detection on Road Surface Images. In *International Conference on Quality Control by Artificial Vision, Le Creusot, France*, 2001.

[6] C. Bhagvati, M. M. Skolnick, and D. A. Brivas. Gaussian normalization of morphological size distribution for increasing sensitivity to texture variations and its application to pavement distress classification. In *Computer Vision and Pattern Recognition*, pages 700–703, 1994.

[7] C. Boukouvalas, F. De Natale, G. De Toni, J. Kittler, R. Marik, M. Mirmehdi, M. Petrou, P. Le Roy, R. Salgari, and G. Vernazza. ASSIST: automatic system for surface inspection and sorting of tiles. *Journal of Materials Processing Technology*, 82(1):179–188, 1998.

[8] J. Bray, B. Verma, X. Li, and W. He. A Neural Nework based Technique for automatic Classification of Road Cracks. In *International Joint Conference on Neural Networks*, pages 907–912, 2006.

[9] S. Chambon. Detection of road cracks with multiple images. In *International Conference on Computer Vision Theory and Applications, VISAPP*, 2010.

[10] S. Chambon. Detection of points of interest for geodesic contours: Application on road images for crack detection. In *International Conference*

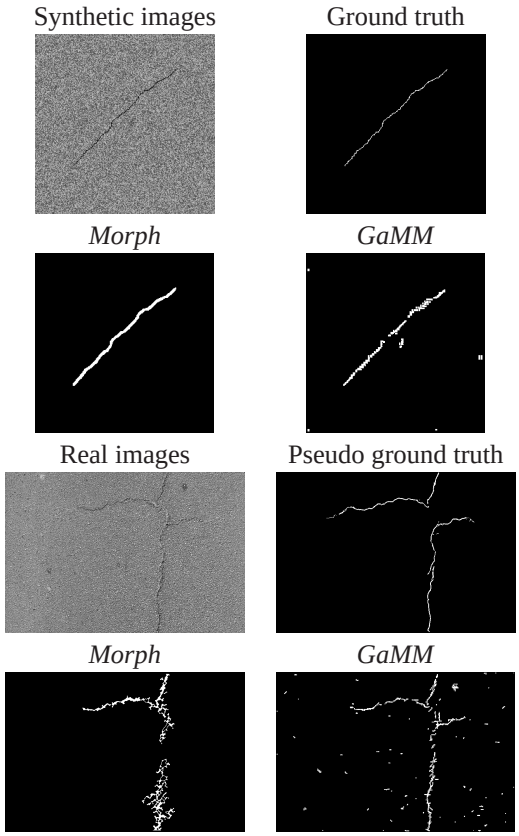


Figure 15: The differences between *Morph* and *GaMM* – Examples with synthetic images and real ones. *Morph* is more efficient than *GaMM* with simple synthetic images whereas *GaMM* has a better behavior with real images.

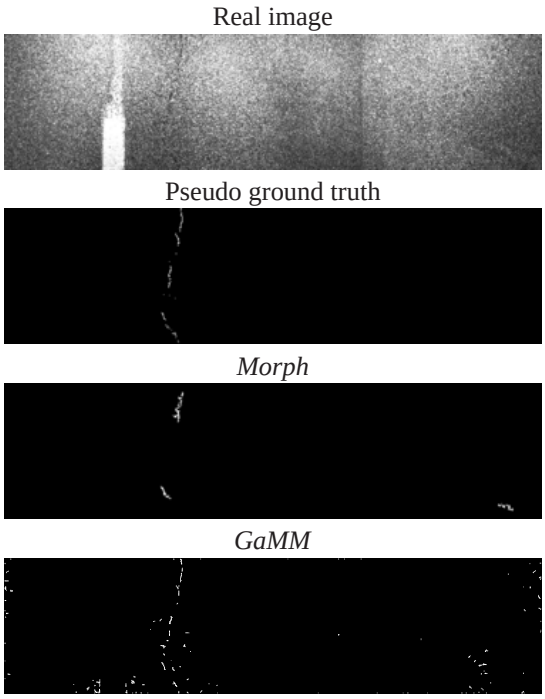


Figure 16: The differences between *Morph* and *GaMM* – Examples with real images acquired on a vehicle. The detection with *GaMM* is more complete than the detection with *Morph*.

on Computer Vision Theory and Applications, VISAPP, Algarve, Portugal, 2011.

[11] S. Chambon, C. Gourraud, J.-M. Moliard, and P. Nicolle. Road crack extraction with adapted filtering and markov model-based segmentation – introduction and validation. In *International Conference on Computer Vision Theory and Applications, VISAPP*, 2010.

[12] 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*, 2009.

[13] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum. Detection of blood vessels in retinal images using two-dimensional matched filters. *IEEE Transactions on Medical Imaging*, 8(3):263–269, 1989.

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

[15] H. D. Cheng, X. Jiang, J. Li, and C. Glazier. Automated Real-Time Pavement Distress Analysis. *Journal of the Transportation Research Board (Transportation Research Record)*, 1655:55–64, 1999.

[16] H. D. Cheng and M. Miyojim. Novel System for Automatic Pavement Distress Detection. *Journal of Computing in Civil Engineering*, 12(3):145–152, 1998.

[17] J. Chou and H. D. Cheng. Pavement distress classification using neural networks. In *IEEE Transactions on Systems, Man and Cybernetics*, volume 1, pages 397–401, 1994.

[18] J. Chou, W. A. O'Neill, and H. Cheng. Pavement Distress Evaluation Using Fuzzy Logic and Moment Invariants. *Journal of the Transportation Research Board (Transportation Research Record)*, 1505:39–46, 1995.

[19] K.-M. Chua and L. Xu. Simple Procedure for Identifying Pavement Distress from Video Im-

ages. *Journal of Transportation Engineering*, 120(3):412–431, 1994.

[20] D. Corso, R. Fioravanti, and S. Fioravanti. Morphological analysis of textured images for identification of thin structures. In *International Conference on Acoustics, Speech, and Signal Processing*, volume 4, pages 2359–2362, 1995.

[21] M. Coster and J.-L. Chermant. Image analysis and mathematical morphology for civil engineering materials. *Cement and Concrete Composites*, 23(2):133–151, 2001.

[22] N. Coudray, A. Karathanou, and S. Chambon. Multi-resolution approach for fine structure extraction – application and validation on road images. In *International Conference on Computer Vision Theory and Applications, VISAPP*, 2010.

[23] D. Darwin, M. N. Abou-Zeid, and K. W. Ketcham. Automated crack indentification for cement paste. *Cement and Concrete Research*, 25(3):605–616, 1995.

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

[25] J. Dumoulin, P. Subirats, V. Legay, D. Meignen, C. Gouraud, J.-P. Delmulle, and M. Garnaud. Progressive Automation Of Pavement Surface Distress By Imaging Techniques. In *Final seminar of the 11A025 research project*, 2005.

[26] H. Elbehiery, A. Hefnawy, and M. Elewa. 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, 2005.

[27] R. Forest and V. Utsi. Non destructive crack depth measurements with ground penetrating radar. In *International Conference on Ground Penetrating Radar*, pages 799–802, 2004.

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

[29] D. Geman and B. Jedynak. An Active Testing Model for Tracking Roads in Satellite Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(1):1–14, 1996.

[30] A. Georgopoulos, A. Loizos, and A. Flouda. Digital image processing as a tool for pavement distress evaluation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 50(1):23–33, 1995.

[31] S. A. Guralnick, E. S. Suen, and C. Smith. Automatic inspection of highway pavement surfaces. *Journal of Transportation Engineering*, 119(1):1–12, 1993.

[32] C.-J. Hsu, C.-F. Chen, C. Lee, and S.-M. Huang. Airport pavement distress image classification using moment invariant neural network. In *Asian Conference on Remote Sensing*, volume 1, pages 216–220, 2001.

[33] Y. Huang and B. Xu. Automatic inspection of pavement cracking distress. *Journal of Electronic Imaging*, 15(1), 2006.

[34] A. Ito, Y. Aoki, and S. Hashimoto. Accurate extraction and measurement of fine cracks from concrete block surface image. In *Annual Conference of the Industrial Electronics Society*, volume 3, pages 2202–2207, 2002.

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

[36] S. Jitprasithsiri and H. Lee. Development of a Digital Image Processing Algorithm to Compute a Unified Crack Index for Salt Lake City. In *Annual Meeting of Transportation Research Record*, 1996. paper 960889.

[37] M. S. Kaseko and S. G. Ritchie. A neural network-based methodology for pavement crack detection and classification. *Annual Meeting of Transportation Research Record*, 1(1):275–291, 1993.

[38] D. H. Kil and F. B. Shin. Automatic road-distress classification and identification using acombination of hierarchical classifiers and expert systems-subimage and object processing. In *International Conference on Image Processing*, volume 2, pages 414–417, 1997.

[39] K. R. Kirschke and S. A. Velinsky. Histogram-Based Approach for Automated Pavement-Crack Sensing. *Journal of Transportation Engineering*, 118(5):700–710, 1992.

[40] H. N. Koutsopoulos and A. B. Downey. Primitive-Based Classification of Pavement Cracking Images. *Journal of Transportation Engineering*, 119(3):402–418, 1993.

[41] H. N. Koutsopoulos and I. El Sanhouri. Methods and Algorithms for Automated Analysis of Pavement Images. *Journal of the Transportation Research Board (Transportation Research Record)*, 1311:103–111, 1991.

[42] H. Lee and H. Oshima. New Crack-Imaging Procedure Using Spatial Autocorrelation Func-

tion. *Journal of Transportation Engineering*, 120(2):206–228, 1994.

[43] C.-X. Ma. Pavement distress detection based on nonsubsampled contourlet transform. In *International Conference on Computer Science and Software Engineering*, pages 28–31, 2008.

[44] D. S. Mahler, Z. B. Kharoufa, E. K. Wong, and L. G. Shaw. Pavement Distress Analysis Using Image Processing Techniques. *Microcomputers in Civil Engineering*, 6(1):1–14, 1991.

[45] D. Meignen, M. Bernadet, and H. Briand. One Application of Neural Networks for Detection of Defects Using Video Data Bases: Identification of Road Distresses. In *International Workshop on Database and Expert Systems Applications*, pages 459–464, 1997.

[46] A. Mraz, M. Gunaratne, A. Nazef, and B. Choubane. Experimental evaluation of a pavement imaging system : Florida department of transportation's multipurpose survey vehicle. *Journal of the Transportation Research Board (Transportation Research Record)*, 1974:97–106, 2006.

[47] M. Mustaffara, T. C. Ling, and O. C. Puan. Automated Pavement Imaging Program (APIP) for Pavement Cracks Classification and Quantification – a Photogrammetric Approach. In *The Congress of the International Society for Photogrammetry and Remote Sensing*, volume WG IV/3, 2008.

[48] T. S. Nguyen, M. Avila, S. Begot, F. Duculty, and J.-C. Bardet. Automatic detection and classification of defect on raod pavement using anisotropy measure. In *European Signal Processing Conference*, pages 617–621, 2009.

[49] N. Nishimura and S. Kobayashi. A boundary integral equation for an inverse problem related to crack detection. *International Journal for numerical methods in engineering*, 32(7):1371–1387, 1991.

[50] H. Oliveira and P. L. Correia. Identifying and retrieving distress images from road pavement surveys. In *First Workshop on Multimedia Information Retrieval: new trends and challenges, International Conference on Image Processing*, 2008.

[51] H. Oliveira and P. L. Correia. Supervised strategies for cracks detection in images of road pavement flexible surfaces. In *European Signal Processing Conference*, 2008.

[52] H. Oliveira and P. L. Correia. Automatic road crack segmentation using entropy and image dynamic thresholding. In *European Signal Processing Conference*, 2009.

[53] L. D. Payne. Automating road surface analysis. In *ACM/SIGAPP symposium on Applied computing*, pages 944–950, 1992.

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

[55] J. Pynn, A. Wright, and R. Lodge. Automatic Identification of cracks in road surfaces. In *International Conference on Image Processing and its Applications*, volume 2, pages 671–675, 1999.

[56] C. Rasse, V. Leemans, M.-F. Destain, and J.-C. Verbrugge. *Bearing Capacity of Roads, Railways and Airfiels*, chapter Application of Image Analysis to the Identification and Rating of Road Surface Distress, pages 61–68. ISBN 90 5809 396 4. Correira and Branco, Swets and Zeitlinger, 2002.

[57] S. G. Ritchie, M. Kaseko, and B. Bavarian. Development of an Intelligent System for Automated Pavement Evaluation. *Journal of the Transportation Research Board (Transportation Research Record)*, 1311:112–119, 1991.

[58] H. C. Rughooputh, S. D. Rughooputh, and Jason J. M. Kinser. Automatic inspection of road surfaces. In SPIE, editor, *Machine Vision Applications in Industrial Inspection VIII*, volume 3966, pages 349–356, 2000.

[59] C. Scheffy and E. Diaz. Asphalt concrete fatigue crack monitoring and analysis using digital image analysis techniques. In *Accelerated Pavement Testing International Conference*, 1999.

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

[61] J. Shirataki and T. Tomikawa. A study of road crack detection by image processing. Part B. Science and Technology 24, p. 67–71, Kanagawa Institute of Technology, 2000.

[62] I. Sokolic. Criteria to evaluate the quality of pavement camera systems in automated evaluation vehicules. Rapport de master, Departement of Civil and Environnemental Engineering, College of Engineering, 2003.

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

[64] S. Sorncharean. Crack Detection on Asphalt Surface Image Using Enhanced Gril Cell Analysis. In *IEEE International Symposium on Elec-*

tronic Design, Test and Applications, pages 49–54, 2008.

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

[66] N. Tanaka and K. Uematsu. A Crack Detection Method in Road Surface Images Using Morphology. In *Workshop on Machine Vision Applications*, pages 154–157, 1998.

[67] E. Teomete, V. R. Amin, H. Ceylan, and O. Smadi. Digital image processing for pavement distress analyses. In *Proceedings of the 2005 Mid-Continent Transportation Research Symposium*, page 13, 2005.

[68] T. Tomikawa. A study of road crack detection by the meta-genetic algorithm. In *IEEE conference on Africa, AFRICON*, volume 1, pages 543–548, 1999.

[69] B. Tremblais and B. Augereau. A fast multi-scale edge detection algorithm. *Pattern Recognition Letters*, 25(6):603–618, 2004.

[70] K. C. P. Wang and W. Gong. Automated Pavement Distress Survey: A Review and A New Direction. In *Pavement Evaluation Conference*, pages 21–25, 2002.

[71] K. C. P. Wang, G. Li, and W. Gong. Wavelet-based pavement distress image edge detection with ”À trous” algorithm. In *Annual Meeting of Transportation Research Record*, volume 2024, pages 73–81, 2007.

[72] H. G. Zhang and Q. Wang. Use of Artificial Living System for Pavement Distress Survey. In *Annual Conference of the IEEE Industrial Electronics Society*, pages 2486–2490, 2004.

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

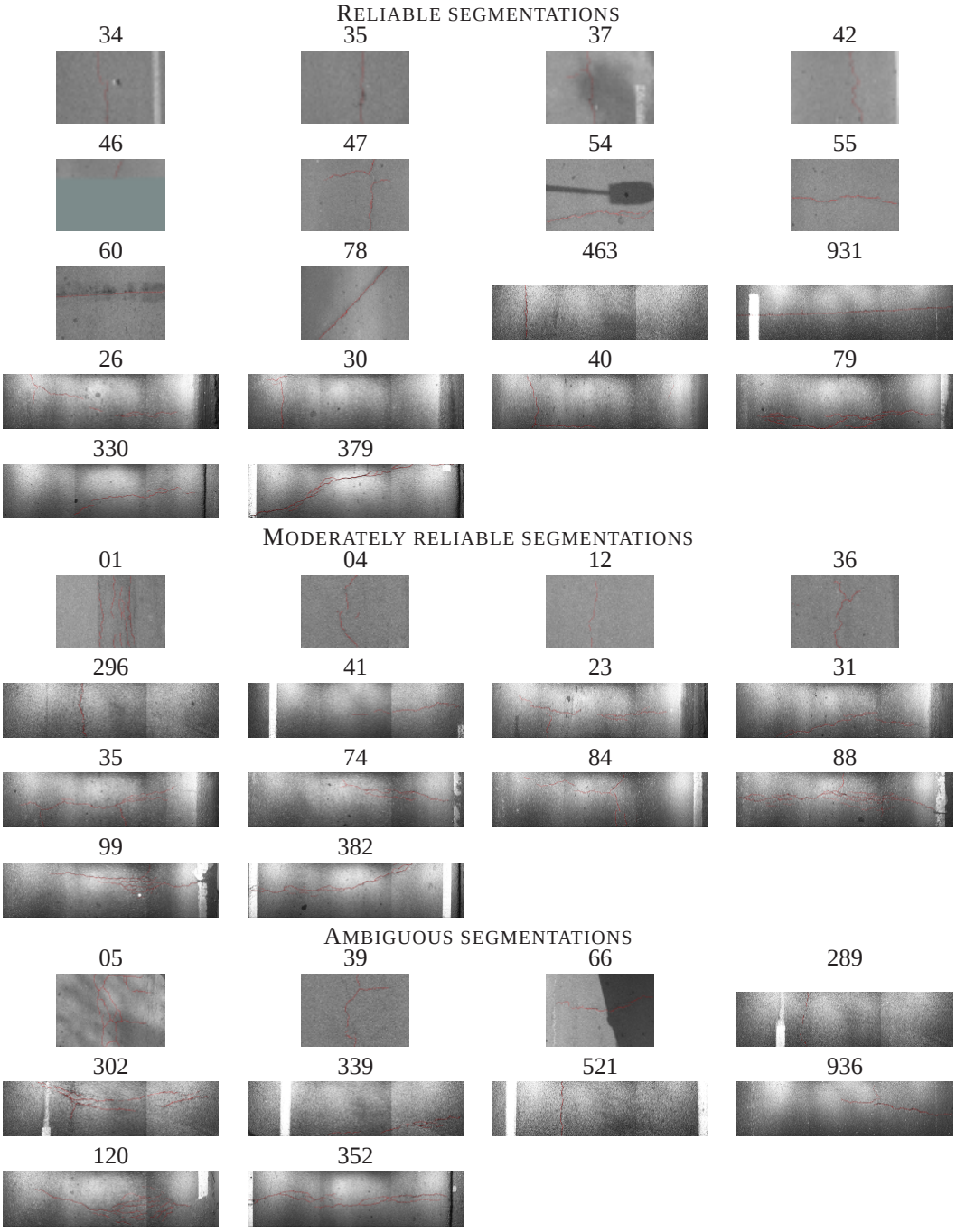


Figure 17: The reference segmentations (or pseudo-ground truth) used in our evaluation protocol for real images – In each original image, the reference segmentation is highlighted in red. Visually, these manual segmentations seem correct and useful for being the reference segmentations in our comparison protocol.