PAVEMENT CRACK DETECTION BASED ON SALIENCY AND STATISTICAL FEATURES

Wei Xu⋆ Zhenmin Tang⋆ Jun Zhou† Jundi Ding⋆
⋆ School of Computer Science and Engineering, Nanjing University of Science and Technology, China
† School of Information and Communication Technology, Griffith University, Nathan, Australia

ABSTRACT

Traditional pavement crack detection methods can not cope well with the complexity and diversity of noises in large image area. To solve this problem, we propose a novel unsupervised crack detection approach based on saliency and statistical features. The saliency is initially represented by a conspicuity map built from the intensity rarity and local contrast of image regions. Then spatial continuity of candidate crack pixels is measured based on the statistical features extracted in their neighborhood. This is followed by a Bayesian model to automatically update the saliency map. Finally, cracks are extracted after adaptive saliency map binarization. Experiments show that proposed method has generated consistent results as those by human visual inspection. The results have also proved the effectiveness of the proposed method in suppressing noises compared with several alternative methods.

Index Terms— Crack detection, saliency map, statistical feature, Bayesian model

1. INTRODUCTION

Cracks are the main distress of structured pavement surface. Traditionally, crack detection is performed by human visual inspection, which is dangerous, subjective, and time-consuming. To automate this process, some image-based crack detection and recognition methods have been reported in the past decade. These methods normally use curvilinear features to represent cracks, which can be extracted by Sobel-based algorithm [1], Canny edge detector [2], or Wavelet edge detection [3].

One problem in automatic crack detection is how to deal with noises introduced by shadows, water stains and oil spots. Some methods make assumption that there is clear distinction between noises and cracks. For example, iterative clipping method (ICM) [4] assumes that the intensity of crack is normally darker than that of its surroundings. It iteratively uses mean and standard deviation values to distinguish real crack regions from noises. Li et al., on the other hand, presented a neighboring difference histogram method (NDHM) [5] to segment a crack image using a globally optimized threshold. However, the above assumption is not always true, which makes these methods not robust in dealing with various noisy conditions. To improve the correctness and completeness of crack detection, NDHM is extended by an F*-seed-growing approach which has a strong anti-noise capability [5]. More recently, a CrackTree method was developed to address the influence from shadows [6]. It has shown outstanding precision and recall, but with high computation cost.

This research is supported by the National Natural Science Foundation of China (NSFC) under Grant No. 90820306 and Grant No. 61103058, and under Australian Research Council’s DECRA Projects funding scheme (project ID DE120102948).

In this paper, we propose a crack detection method based on saliency and statistical features. Visual saliency is often motivated by the low-level (bottom-up) image attributes, which generates the uniqueness, regions of interest, or surprise [7]. In the proposed method, saliency detection allows fast locating of possible crack regions, and statistical features that characterize the spatial context of cracks helps to distinguish cracks from noises. We commence from rarity and contrast based saliency measurement that produces an initial conspicuity map. Then a saliency map is constructed by incorporating the directional spatial features in the neighborhood of local regions, and updated by a Bayesian model. Finally, cracks are detected via adaptive binarization of the saliency map. The whole process is summarized in Fig. 1.

2. INITIAL CONSPICUITY MAP

The first step of our method is to construct the initial conspicuity map from rarity and contrast features. The rarity feature is measured through histogram based self-information, which takes the grayscale histogram entry as the probability that a pixel is chosen as rare [8]. The contrast is computed from the intensity difference between a pixel and the mean value of its neighborhood. Then, the initial conspicuity value $C_r(x, y)$ of pixel $(x, y)$ is defined as:

$$C_r(x, y) = -\log(H_r(x, y)) D(I_{r,m} - I(x, y))$$

(1)

where $I_{r,m}$ is the mean value of an $N \times N$ image patch $I_r$. $H_r$ is the histogram of $I_r$, and $H_r(x, y)$ is the histogram entry of corresponding intensity level of pixel $(x, y)$. $D(*)$ is the Euclidean distance.

Several problems may arise from this step. Firstly, some local rare pixels may be widely distributed over their neighboring blocks. Secondly, noise pixels with relatively high rarity may be misclassified as real crack pixels. To solve these two problems arising from local processing, we propose to explore statistical feature of spatial relationship between an image patch and its neighboring blocks. This allows tracking of the cracks by extending to neighborhood regions. Such tracking can be implemented by selecting a neighboring patch (or direction) with the least mean intensity difference to the center patch (Fig. 2). In this step, global compensation...
is applied to reduce the impact of uneven intensity along cracks, as done in [6] (Fig. 2 (b) and (c)).

Finally, the improved conspicuity map is given by:

\[
d = \arg \min_{\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}} D(I_{r,\mu} - I_{o,\mu})
\]

\[
C_r(x, y) = -\log(H_{r,d}(x, y)) + \frac{D(I_{r,\mu} - I(x, y))}{s} \quad (3)
\]

where \(I_{o,\mu}\) is the mean intensity value of neighbor blocks centered at \(r\) with direction \(\theta\). \(H_{r,d}(x, y)\) is the local histogram entry of corresponding intensity level of pixel \((x, y)\) in block \(r\) and its neighbor with direction \(d\). \(\sigma_g\) and \(\sigma_r\) are the global and local standard deviation.

3. CRACK SALIENCY MAP

The conspicuity values of dark particle textures and shadows are high due to their rarity and contrast features, which are not easily distinguishable from crack pixels. To reduce their impact, the spatial property of cracks shall be considered, i.e., the irregular continuity that satisfies the differentiable saliency measure [5, 6].

For a pixel \(x\) in a double-extending window \(W\) with an inner window \(W_l\) and an outer window \(W_o\) (as shown in Fig. 3), let its continuity feature be \(F(x)\). \(x\) can either be a real crack pixel (salient) \(C_1\) or noise interference (not salient) \(C_0\). Similar to [9], the saliency \(P(C_1|F(x))\) of pixel \(x\) is defined by the conditional probability

\[
S(x) = \frac{P(F(x)|C_1)P(C_1)}{P(F(x)|C_1)P(C_1) + P(F(x)|C_0)P(C_0)} \quad (4)
\]

\[
P(C_0) = 1 - P(C_1) \quad (5)
\]

where the prior probability \(P(C_1)\) and \(P(C_0)\) can be obtained from the initial conspicuity value. Let \(P(F(x)|C_1)\) be \(p_1\) and \(P(F(x)|C_0)\) be \(p_0\), then we propose a neighbor extending algorithm to compute the posterior probability \(p_1\) and \(p_0\). For each pixel \(x\) in conspicuity map, the algorithm is presented as follows:

**Initialization:** Let the minimum real crack length be \(L\) and initial extension step \(s = 1\).

1. If the conspicuity value of pixel \(x\) is less than twice of the average, \(S(x) = 0\), stop algorithm. Otherwise, take \(x\) as a candidate point.

**Repeat:**

2. Let \(x\) be the the center of inner window \(W_l\) with size \((2s - 1) \times (2s - 1)\), detect whether there is a candidate point with high conspicuity value in its \((2s + 1) \times (2s + 1)\) outer window \(W_o\).

3. If a point can’t be found, the posterior probabilities are set \(p_1 = (2s - 1)/L\) and \(p_0 = 1 - (2s - 1)/L\), respectively. This means that \(x\) is more likely to be particle noise or discontinuous crack fragment. Stop algorithm.

4. Otherwise, record the extending direction \(\theta\) (up, down, left or right) according to Fig. 3; if there are several candidate points, record them sequentially. Then, set \(s = s + 1\).

**Until** \(s > L/2\)

5. When \(s > L/2\), check whether the linear extending direction is consistent. Inconsistency suggests that a noisy pixel has been included. Thus, record the extension step \(s\) as \(t\) at this error location, modify \(p_1 = (2t - 1)/L\) and \(p_0 = 1 - (2t - 1)/L\). Stop algorithm.

6. When \(s > L/2\), count the candidate point number \(N\) in \(W_l\). Based on the rarity of crack area, we set a ratio threshold \(\alpha\). If \(N > \alpha(2s - 1)^2\), these points are believed to have "block effect" (experimentally \(\alpha < 0.4\)). Then the process can be extended until there is no candidate point between \(W_l\) and \(W_o\). Recount \(N\) of candidate point in \(W_l\) and modify \(p_0 = N/\alpha(2s - 1)^2\), \(p_1 = 1 - N/\alpha(2s - 1)^2\). Stop algorithm.

7. If both conditions of step 5 and 6 are not met when \(s > L/2\), set \(p_1 = (2s - 1)/L\) and \(p_0 = 1 - (2s - 1)/L\).

Finally \(S(x)\) can be written as:

\[
S(x) = \frac{p_1(x)C(x)}{p_1(x)C(x) + p_0(x)C_{1,\mu}(x)} \quad (6)
\]

where \(C(x)\) stands for the conspicuity value, \(C_{1,\mu}(x)\) stands for the mean conspicuity value in inner window \(W_l\) excluding pixel \(x\). Each direction in Fig. 3 has three angles according to the irregular extending direction of cracks.

Sometime, discontinuity of crack happens due to inconsistent illumination angle and inappropriate neighborhood processing. To connect broken cracks, we use the extending direction of candidate points in two neighbor local regions when they are within a tolerant distance \(d\). If their extending direction \(\theta\) is consistent (up-down, left-right), the saliency value of pixels between them are replaced by the average of two candidate points, as shown in Fig. 4.

In most cases, real cracks not merely exist in local region. Therefore, spatial information can be used to enhance the saliency map. For a pixel \(x\) in a window \(W_k\), let one of its linear neighbor window be \(W_k\), and the mean conspicuity value of candidate points in its neighbor \(W_k\) be \(p_{k,\mu}\), the spatial enhanced weight is defined as
\[ \omega_r = \max \left\{ \sum_{\theta, k} \exp \left( -\frac{D_\theta(W_r, W_k)}{N} \right) p_{k, \theta} \right\} \]

where \( D_\theta \) stands for the center distance between window \( W_r \) and \( W_k \) along the direction \( \theta \), and \( N \) is the number of neighbor windows.

The final saliency measure is

\[ S(x) = \frac{\omega_r p_1(x) C(x)}{\omega_r p_1(x) C(x) + p_0(x) C_{\mu, \sigma}(x)} \]

(8)

In our experiment, we have set the minimum crack length \( L = 9 \) and the tolerant distance \( d = 4 \) (pixels) empirically.

### 4. CRACK SALIENCY MAP BINARIZATION

We adopt a thresholding strategy to convert the saliency map into a binary image. Let the rarity parameter of crack area be \( \alpha \), \( i \) is the saliency value within the range \( [S_{min}, S_{max}] \) and \( \text{Sum}(S > S_i) \) stands for the number of salient pixels whose saliency values are higher than the average \( S_{\mu} \). Given a threshold \( l \), we compute

\[ \sum_{i=S_{\max}}^{1} \text{Num}_i \geq \alpha \times \text{Sum}(S > S_{\mu}) \]

where \( \text{Num}_i \) is the number of pixel with saliency value \( i \). The summation is accumulated from \( S_{\max} \) to \( S_{min} \) descendingly. Then the adaptive threshold \( T \) is determined as

\[ T = \max(l, T_{\text{otsu}} - \sigma) \]

(10)

where \( T_{\text{otsu}} \) is the threshold computed by Otsu method [10] and \( \sigma \) is the standard deviation of saliency value.

### 5. EXPERIMENTAL RESULTS

We have collected 261 pavement crack images taken by a line-scanning camera with the ground sample distance (GSD) of 0.92mm during a highway survey. The ground truth cracks (GT) are manually labelled. These images are of 512×512 pixels and contain various noise conditions such as particle textures, oil patches, water stains, and lane marks. Experiments were run on a Windows desktop with Intel(R) Core(TM) i5-2410M CPU and 3G RAM.

In order to validate the effectiveness of the proposed method, we compared it against several rarity or contrast based saliency approaches, which include Rare [8], LC [11], AC [12], FT [13] and MSSS [14]. These state-of-the-art methods are used to generate the conspicuity maps, which are followed by the algorithm in section 3 to get the final crack saliency maps. Fig. 5 shows the saliency maps of three representative crack images. All methods can highlight most of the real crack regions, but our method has generated the best saliency maps that are more consistent with the GT.

We used the precision-recall curve to compare different saliency detection methods by varying the threshold from 0 to 255 for binarization and computing the precision and recall values at each threshold with respect to the GT [13]. As can be observed from Fig. 6, our method has significantly outperformed the alternatives. All curves show a sudden rise on precision when the binarization
threshold changes in a small range, and then the precision become quite stable when the recall decreases. This implies that a major portion of real cracks can be detected and be successfully distinguished from most noises pixels if the binarization threshold is set to the appropriate range. This is partly due to the saliency value enhancement step as described in section 3, which has enhance the saliency of pixels with high conspicuity and suppressed those with low conspicuity.

We also compare our method with five alternative pavement crack detection methods in the literature, i.e., threshold based methods (ICM [4] and NDHM [6]), edge based methods (BEMD + Sobel [1] and Canny [2]), and the frequency domain based method (Wavelet [3]). Because cracks are elongated, which are about 1 to 3 pixels wide, if a detected crack pixel is no more than 1 pixel away from the GT, we consider it as a true positive [5].

Fig. 7 shows the results on three sample images. For fair comparison, we have applied grayscale correction [6] as a preprocessing step in all methods. Our approach has preserved most of the important real crack regions with less fracture, and suppressed the disturbance of particle textures. All alternatives are suffered from the interferences of various noises, discontinuous cracks, and missing part of the important crack fragments.

Fig. 8 shows the average performance of different methods over all 261 crack images. Our method has shown high precision ($P = 0.89$), recall ($R=0.83$) and F-measure ($F = 0.86$), which clearly outperforms all alternative methods. This is because real cracks and various noises with similar features are better distinguished. On the other hand, some noises caused by highly textured road surface are still difficult to exclude.

6. CONCLUSION

We have introduced the saliency concept into the challenging work of automatic pavement crack detection. Our method combines and improves the rarity and contrast based saliency measure. The proposed statistical feature extraction and Bayesian estimation method have greatly enhanced the saliency map. This implies that spatial property of cracks shall be considered during the feature extraction step. The experimental results on 261 images show that this approach has significantly outperformed several traditional methods.

7. REFERENCES


