HTF: A NOVEL FEATURE FOR GENERAL CRACK DETECTION

Han Hu, Quanquan Gu and Jie Zhou

State Key Laboratory on Intelligent Technology and Systems
Tsinghua National Laboratory for Information Science and Technology (TNList)
Department of Automation, Tsinghua University, Beijing, China, 100084
{huh04, gqq03}@mails.tsinghua.edu.cn, jzhou@tsinghua.edu.cn

ABSTRACT

Recent years, image-based crack detection has attracted more and more attentions for its potential applications on the inspection, diagnosis, and maintenance of various products, e.g. metal workpiece, concrete structure, asphalt road and etc. Generally, the applications are inevitably confronted with noises such as non-uniform illuminated conditions, shadings, stains and nature textures. To ease the problem, traditional methods usually focused on a special application and are carried out with strictly controlled image acquisition environments or an adhoc preprocessing procedure. In this paper, we propose a general crack detection method which can deal with various products as well as the noises in a unified fashion, and even with the same parameters. The method partitions an image into overlapped small grid cells and determines whether the cells contain cracks using a well designed feature descriptor. Experiments using real images of various products show the effectiveness of the proposed feature as well as the method.

***Index Terms*—** crack detection, grid cell, Hough transform, Support Vector Machine

1. INTRODUCTION

For many architectural and manufacturing products, it is a very important task to inspect the surfaces and repair the damages or degradations. Traditionally, this task is done by skilled human beings. For example, the road menders are spending months of time to find and repair the degradations in asphalt pavements. However, if we can automatically inspect the surfaces by captured images, then a road mender just needs to drive a survey lorry with camera underneath, and the machine can automatically inspect the road surfaces and stop to seal the damaged area if needed, which saves a lot of manual work and time cost.

Among the various defects, crack is one of the most common and fundamental type. In the last decade, several methods for automatic crack detection have been proposed. These methods can be roughly grouped into two categories: global based and local grid based. Global based methods search the crack points over the whole image by using anisotropy measures [1], frequency spectrum [2], efficient preprocessing [3] or percolation model [4, 5]. Because these methods usually need to inspect all pixels and trace them in the neighborhoods, they are much time-consuming. Another drawback is revealed by their strong assumptions on the images, i.e. with few noises [1, 2], uniform illuminations [1, 4, 5], vertical or horizontal cracks[1], relative brighter background [3, 4, 5] and etc.

Other methods consider to partition the image into small grid cells, and find the potential crack cells [6, 7]. These methods limit the searching space to those crack cell candidates and so they save a lot of time. Our work also follows this manner. We design a novel feature which well describes the characteristics of cracks and is invariant to rotation and illumination, and use a Support Vector Machine (SVM) classifier [8] to judge whether a cell is a crack. In this way, we can ease the problems of noises, non-uniform illuminations, and empirical parameter tuning. In the experiments on 50 real images of various products, we compare the proposed feature with the features used in [6, 7] and other popular features [9, 10, 11] used for texture classification or object detection. The results show that the proposed feature performs much better than the others.

This work was supported in part by Mitsubishi Heavy Industries (MHI), in part by the Natural Science Foundation of China under Grants 60721003,60673106, 60875017, in part by the National 863 Hi-Tech Development Program of China under Grant 2008AA01Z123, and in part by the National Key Project of China under Grant 2009BAH40B03

2. THE PROPOSED METHOD

2.1. Method Overview

The flowchart of the proposed approach is shown in Fig.1. We first train a linear SVM classifier which determines whether a cell is a crack or not. This procedure can be done off-line and the trained classifier is saved in the storage or memory. Then we start the online crack detection procedure. We partition the image into 20×20 pixel grid cells and extract a feature vector for each cell. Then we determine whether the cell is a crack or not by inputting its feature into the trained SVM classifier. Since we detect cracks on local cells, we can do the detection at the same time as the acquisition of the surface images.

Different from [7], in our method, the cells are overlapped. That is because there may be ambiguities between a cell with crack on its border and a cell on the joint area of different background types (See Fig.3). Therefore, to avoid wrongly detecting the multi-background cells as cracks, we have to also reject the border-crack cells. Thus using nonoverlapped cells as in [7], those crack areas are missing. However, using overlapped cells, such cracks can be identified by examining the overlapped adjacent cells (see the blue cell in Fig.3).

2.2. Hough Transform Based Feature (HTF)

The most key issue of the method is feature. The feature should be able to describe the physical characteristics of cracks. Carefully examining the defects, we can find that the pixels of a crack in a block have three significant properties: 1) they have much different gray scale compared with the pixels of good area; 2) they together approximately locate on a line; 3) the gray scale changes around the two sides of cracks are approximately symmetric. To describe

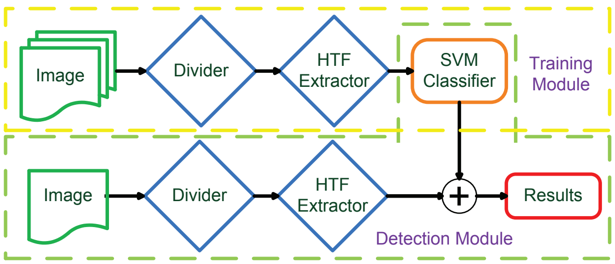


Fig. 1. Flowchart of the proposed approach.

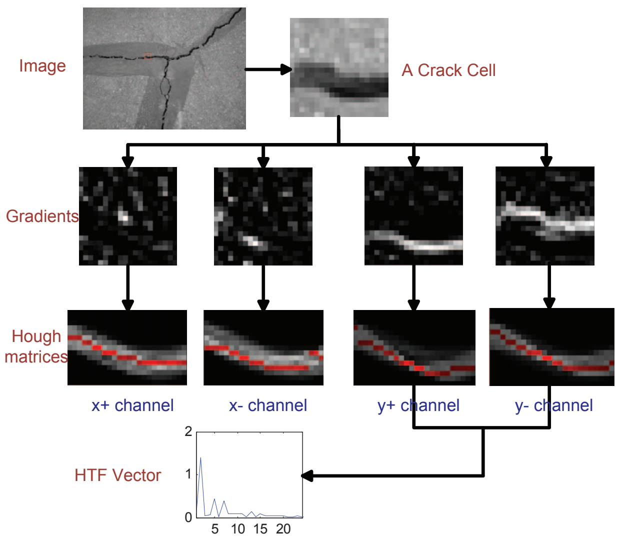


Fig. 2. Sketch map for HTF calculation. All the gradients and Hough matrices are normalized for visualization that the maximum value is 255. The red points in the Hough matrices represent the biggest value of each column, and the brightness indicates the relative magnitude. For the selected cell in this figure, we use the vertical gradients to form the HTF vector.

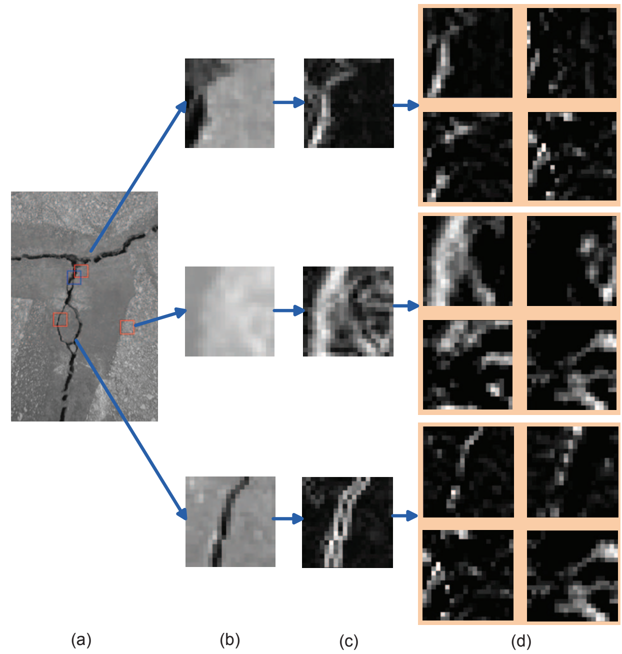


Fig. 3. Several kinds of cells. In the figure, they are boundary-crack cell, joint-backgrounds cell and crack cell. (a) is an image of real chapped asphalt road marked with the selected red grid cells on it; (b) are the zoomed grid cells; (c) is the amplitude gradient response; (d) are the 4-channel gradient responses, and from up-left to down-right, they are "x+","x-","y+" and "y-", respectively.

these physical characteristics, we design a novel feature based on Hough transform, referred as HTF. Fig.2 shows the sketch map for HTF construction. For each cell, we first calculate the multi-channel gradient responses. Then, we calculate the soft Hough matrix for each channel. And finally, we form the feature vector by extracting meaningful properties from the Hough matrices.

2.2.1. Multi-channel Gradient Responses

Since the method is designed generally for various products, the crack area can be either black, white, or in-between, and so is the regular area. Therefore, directly using original gray scale images, we can hardly deal with them in a unified fashion. Nevertheless, we can solve it by using the nonnegative gradient versions of the original image, which are invariant to the absolute gray scale.

The usual practice is to obtain the amplitude gradient response. However, this may lose some key information. For example, as shown in Fig.3, the third cell contains a black crack, while the second one is on the joint of non-uniform backgrounds. Using amplitude gradient responses (see Fig.3(c)), they are similar in shape. To avoid this problem, instead of using amplitude gradient response, we calculate 4-channel edge responses, referred as "x+", "x-", "y+", and "y-", respectively. "x" represents the gradients of horizontal direction, and "y" represents the gradients of vertical direction. "+" and "-" represent the gradient responses only preserving positive and negative elements, respectively. We calculate the 4 channels by using corresponding masks and only maintaining the positive values (see Fig.4). See Fig.3(d), the crack cell has large responses on both "+", "-" channels, while the multi-background cell only has large responses on one of "+" or "-". So we can tell them apart.

2.2.2. Soft Hough Transform

The pixels of cracks in a cell together approximately locate on a line. However, it is hard to directly extract numeric values from original cells to describe these properties. To overcome this difficulty, we adopt the Hough transform, which is a popular technique to identify positions of arbitrary shapes in an image [12]. Here, we use the Hough transform mainly for obtaining the position-angle information of an original cell as,

$$r = x \cos\theta + y \sin\theta, \tag{1}$$

where (x, y) is the coordinates of a white pixel, and (r, θ) is the corresponding position-angle parameter curve. In tradition, Hough transform is asserted on the binary image. Here we attempt to extend it to the real gradients image by using the numeric values as weights,

$$\frac{1}{3}\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad \frac{1}{3}\begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} \quad \frac{1}{3}\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$$

Fig. 4. The 3×3 masks for the 4-channel gradient responses and only the non-negative responses are remained. From the left to right, they are "x+", "x-","y+" and "y-".

referred as soft Hough Transform. We can then extract several meaningful numeric properties from the four Hough matrices $H_{\{x,y\}\pm}$ to form HTF feature.

2.2.3. *Feature Calculation*

In our default settings, the (r, θ) parameter space is divided into 15×24 bins. Notice that the summation of bins for each θ are constant. So for each θ, we use the maximal bin value to the constant summation as the measure of r concentricity on the angle. The crack is also highly concentrated in θ domain, which means that only a few θs have high r concentricity. To be rotation invariant, we sort the r concentricity measures, and form a 24-elements vector $\mathbf{v}$ for a H. The descent velocity of $\mathbf{v}$ can be a measure of the θ concentricity.

The crack may trace along arbitrary angles. Nevertheless, it must have significant gradient responses in at least one direction of "x" and "y". And we use the one with the larger r and θ concentricity as the main direction to calculate HTF, which is measured by,

$$\begin{aligned} &\max(\mathbf{v}_+) \cdot \sum_i ((\mathbf{v}_+(i+1) - \mathbf{v}_+(i)) \cdot \alpha^i) \\ &\quad + \max(\mathbf{v}_-) \cdot \sum_i ((\mathbf{v}_-(i+1) - \mathbf{v}_-(i)) \cdot \alpha^i), \end{aligned} \tag{2}$$

where $\mathbf{v}_+, \mathbf{v}_-$ are the r concentricity measure vectors for all θs of "+" and "-" Hough matrices, and α is an empirically chosen attenuation factor (in our implementation, we set it as 0.9). Hough matrices in either "x" or "y" direction, $\max(\mathbf{v}_\pm)$ indicates the largest r concentricity, and $(\mathbf{v}_\pm(i+1) - \mathbf{v}_\pm(i)) \cdot \alpha^i$ measures the θ concentricity. In the following, we only use the Hough matrices of the main direction for calculation. And for convenience we do not mark the variables with their direction ("x" or "y").

We calculate the following formulas as the elements of HTF:

1. the summation of $H_\pm$'s column, which measures the total strength of gradient responses;

2. the largest r concentricity
$$\max(\mathbf{v}_+) + \max(\mathbf{v}_-); \tag{3}$$

3. the angle concentricity $(i = 1, \cdots, 23)$
$$\frac{(\mathbf{v}_+(i+1) - \mathbf{v}_+(i))}{\max(\mathbf{v}_+)} + \frac{(\mathbf{v}_-(i+1) - \mathbf{v}_-(i))}{\max(\mathbf{v}_-)}; \tag{4}$$

4. the difference measures of "+" and "-" responses
$$\frac{||\mathbf{v}_+ - \mathbf{v}_-||_1}{||\mathbf{v}_+ + \mathbf{v}_-||_1}, \tag{5}$$
$$\frac{||\mathbf{v}_+^0 - \mathbf{v}_-^0||_1}{||\mathbf{v}_+^0 + \mathbf{v}_-^0||_1}, \tag{6}$$
$$\min f_r(G_+, G_-), \tag{7}$$

where $|| \cdot ||$ is is the ℓ_1 norm operator; $\mathbf{v}_\pm^0$ are the original vectors of $\mathbf{v}_\pm$ before sorting; $G_\pm$ are the "+" and "-" channel gradient responses of the original cell. $f_r(G_+, G_-)$ represents the correlation function of G_+ and G_-. Eq.(5) and Eq.(6) indicate the overall difference of the "+" and "-" channels' r concentricity at all the θs. Eq.(7) measures the inconsistency of shifts between the "+" and "-" gradient responses. Generally, the shifts of the "+" and "-" gradient responses are consistent, which means that choosing a suitable shift, a crack cell has a high correlation between G_+ and G_-. And for non-crack cells, the shifts are usually inconsistent.

3. EXPERIMENTAL RESULTS

We now verify the accuracy of the proposed method. In the experiments, we use 50 real images of various products for the test, including metal workpieces, concrete structures, asphalt road surface and etc (see Fig.5). It is a challenging task as the images are with various resolutions, illuminated conditions and textures and we should handle them with the same parameters. To verify the effectiveness of the proposed HTF feature, we compare it with other five features:

1. Cell Templates (CT) [6]: it uses Principal Component Analysis (PCA) to calculate several orthonormal cell templates, and the reconstruction coefficients form the feature vector.

2. Border Gray Scale Chain (BGC) [7]: The feature is based on the observation that a crack should go through the border two times. It uses the brightness profiles of border pixels as the feature vector. To be rotational invariant, the first element is set to be the brightness of the darkest pixel.

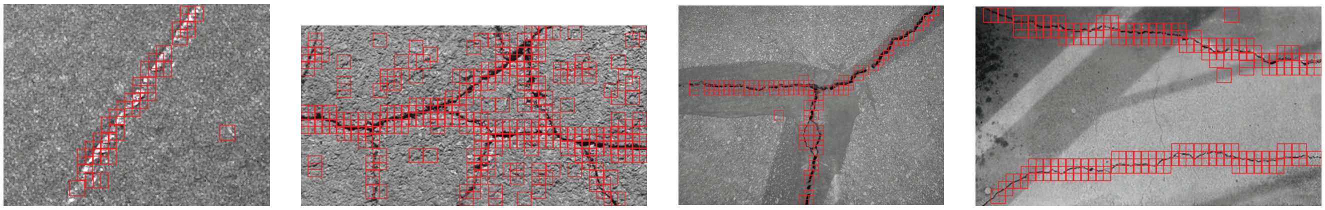
3. Moment Feature (MF) [9]: it uses several lower order central moments as the feature.

4. Histogram of Oriented Gradients (HOG) [10]: it calculates the cell's histogram of oriented gradients as the feature. In the experiment, we use 2×2 "cell"s (Notice that the "cell" here is a parameter of HOG in [10]) for a block and 5×5 pixels for a "cell".

5. Local Binary Patterns (LBP) [11]: it is defined as a gray-scale invariant texture measure, derived from a general definition of texture in a local neighborhood.

We partition the images into overlapped 20×20 pixel grid cells and label them as crack or non-crack. Then we randomly select 1000 crack and 1000 non-crack cells to train a SVM classifier, and the remaining ones for testing. The features are evaluated by the receiver operating characteristic (ROC) analysis [13] (TPR versus FPR), where TPR (True Positive Rate) means the proportion of crack cells that are correctly identified, and FPR (False Positive Rate) means the proportion of non-crack cells that are incorrectly identified as crack ones. We obtain the operating points on the ROC curve by shifting the decision hyperplane of the SVM classifier.

We conduct 20 trails of training and testing sets segmentation for each feature, and we use the average (FPR, TPR) point at each shifting parameter to form the ROC curve. Fig.6 shows the ROC curves by using various features. The closer to the upper left corner the curve is, the better the performance is. So we can see that the proposed HTF performs much better than other features. Noticing that for crack detection, there exist ambiguities when judging a cell



(a) (b) (c) (d)

Fig. 5. Crack detection results using HTF.

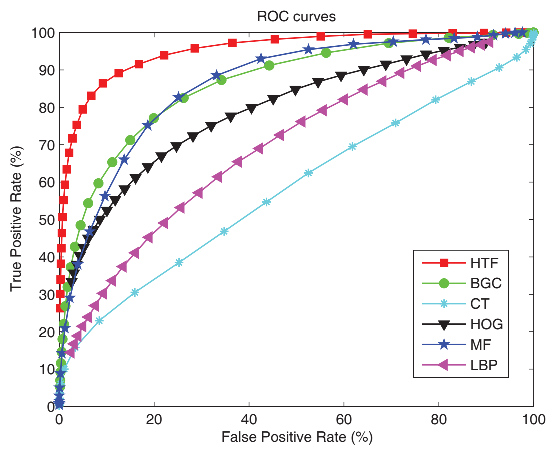


Fig. 6. ROC curves using various features.

as a crack or not. In our experiments, only if a cell locates on long cracks, we label it as a crack. However, many cells with line textures or tiny blemish may also look like a crack when we only focus on the very cell, which contributes to a large portion of classification errors. So the ROC performances of the proposed HTF feature is wonderful and much convincing.

Fig.5 shows the crack detection results of some typical images in a test. The cells marked by red rectangles are the detected crack ones. Fig.5(a) is a metal workpiece, with a relative brighter crack. Fig.5(b) and (c) are asphalt road surfaces, while (b) with irregular textures, and (c) with non-uniform backgrounds due to disabled sealing material. Fig.5(d) is an evaluation image used in [3]. The image captures a concrete structure with complex illuminated conditions. We can see that all these images are well handled by the proposed HTF feature.

4. CONCLUSIONS

We have presented a local grid based method for general crack detection based on a novel feature, HTF. Based on several observations, we designed a novel feature HTF using 4-channel gradient responses and Hough Transform. We showed that the proposed HTF feature is invariant to rotation and illumination. And so it can deal with various products as well as the noises in a unified fashion and even with the same parameters. Experiments on real surface images of different products demonstrate the effectiveness of the proposed HTF feature and the crack detection method.

5. REFERENCES

[1] F. Roli, "Measure of texture anisotropy for crack detection on textured surfaces," *Electronics Letters*, vol. 32, 1996.

[2] P. Subirats, J. Dumoulin, V. Legeay, and D. Barba, "Automation of pavement surface crack detection using the continuous wavelet transform," in *International Conference on Image Processing*. IEEE, 2006, pp. 3037–3040.

[3] Y. Fujita, Y. Mitani, and Y. Hamamoto, "A method for crack detection on a concrete structure," in *International Conference on Pattern Recognition*. IEEE, 2006, vol. 3, pp. 901–904.

[4] T. Yamaguchi and S. Hashimoto, "Improved percolation-based method for crack detection in concrete surface images," in *International Conference on Pattern Recognition*. IEEE, 2008, pp. 1–4.

[5] T. Yamaguchi, S. Nakamura, R. Saegusa, and S. Hashimoto, "Image-based crack detection for real concrete surfaces," *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 3, pp. 128–135, 2008.

[6] T. Tomikawa, "A study of road crack detection by the meta-genetic algorithm," in *AFRICON*. IEEE, 1999, vol. 1, pp. 543–548.

[7] Y. Huang and B. Xu, "Automatic inspection of pavement cracking distress," *Journal of Electronic Imaging*, vol. 15, 2006.

[8] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *5th Annual ACM Workshop on COLT*. ACM, 1992, pp. 144–152.

[9] J. Prokop and Anthony P. Reeves, "A survey of moment-based techniques for unoccluded object representation and recognition," *CVGIP: Graphical Models and Image Processing*, pp. 438–460, 1992.

[10] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition*. IEEE, 2005, pp. 886–893.

[11] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, July 2002.

[12] R.O. Duda and P.E. Hart, "Use of the hough transformation to detect lines and curves in pictures," *Comm. ACM*, vol. 15, pp. 11–15, January 1972.

[13] A. R. van Erkel and P. M. T. Pattynama, "Receiver operating characteristic (roc) analysis: Basic principles and applications in radiology.," *European Journal of Radiology*, vol. 27, no. 2, pp. 88–94, May 1998.