

# CRAQUELURE ANALYSIS FOR CONTENT-BASED RETRIEVAL

*F.S. Abas<sup>1</sup> and K. Martinez<sup>2</sup>*

Department of Electronics and Computer Science,  
University of Southampton,  
Southampton, S017 1BJ, United Kingdom.

<sup>1</sup>fsa00r@ecs.soton.ac.uk, <sup>2</sup>km@ecs.soton.ac.uk

**Abstract:** In this paper, we describe a method for the extraction of distinguishable features from crack patterns, particularly those in paintings. First, we filter the selected crack image using 8 differently oriented Gabor filters. Then we thin the image to 1 pixel wide using a morphological thinning algorithm. Next we implement a crack following algorithm and generate statistical structure of global and local features from a chain code based representation. We describe an orientation-based feature extraction method to represent a crack network from sets of local orientation features. The resultant features are used as a guide towards classifying crack network patterns into several predefined classes, i.e circular, rectangular, spider-web, unidirectional and random. A simple classification experiment is presented to describe the significance of those extracted features towards classifying craquelure patterns.

## 1. INTRODUCTION

The Artiste project aims to provide access across museum collections using metadata as well as content-based retrieval of image data. One of the image-based requirements, which came from the Uffizi Gallery in Florence, is to automatically classify the craquelure (cracks) in paintings to aid in damage assessment. Craquelure in paintings can also be used for other research [1]. It can be a very important element in judging authenticity, use of material or environmental and physical impact because these can lead to different craquelure patterns. Although most conservation of fine artwork relies on manual inspection of deterioration signs, the ability to screen the whole collection semi-automatically is believed to be a useful contribution to preservation. Crack formations are influenced by factors including aging and physical impacts which also relate to the wooden framework of the paintings. It is hoped that the mass screening of craquelure patterns will help to establish a better platform for conservators to identify the cause of damage.

In this paper, we present the detection and statistical feature extraction of simple crack networks as a basis to a more efficient algorithm in the future. The algorithm consists of: (1) detection of a crack structure, (2) extraction of global and local features, and (3) an experiment demonstrating the potential use of the features for successful classification.

To fulfill the purpose of this research, we are using X-radiographs of paintings taken by the research labs of the Louvre Museum in Paris and the Hamilton Kerr Institute, Cambridge. These tend to show cracks very well, as they change the transmission of X-rays and suppress details in the paint layer. They provide clearer information compared to visible images of paintings as far as detecting cracks is concern and they are also a common documentation medium in museums.

## 2. CRACK NETWORK DETECTION

In most cases, cracks can be considered as being local minima with a rather elongated structure [2]. In [3], cracks are identified by taking into account the fact that they have considerably darker grey levels compared to the background and are characterized by a uniform grey level. From a local point of view they also have a strong orientation tendency.

Gabor filters are band-pass filters which have both orientation-selective and frequency-selective properties as well as optimal joint-resolution in both spatial and frequency domains [4]. A two dimensional Gabor filter can be viewed as a sinusoidal plane with a particular frequency and orientation modulated by a Gaussian envelope (Fig. 1). The equation below represents a Gabor filter in spatial domain:

$$h(x, y; f, \theta) = \exp \left[ -\frac{1}{2} \left( \frac{x'^2}{\delta_x^2} + \frac{y'^2}{\delta_y^2} \right) \right] \cos(2\pi f x') \quad (1)$$

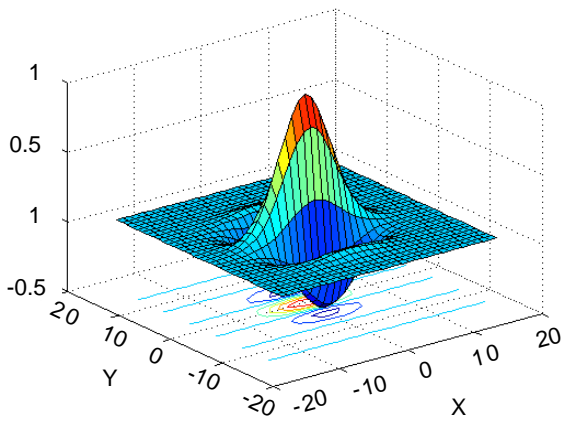
$$x' = x \sin \theta + y \cos \theta \quad (2)$$

$$y' = x \cos \theta - y \sin \theta \quad (3)$$

where  $\delta_x$  and  $\delta_y$  represent the space constants of the Gaussian envelope along the  $x$  and  $y$  axes respectively,  $f$  denotes the frequency of the sinusoidal plane wave along the  $x$ -axis and  $\theta$  is the orientation. The orientation tendency of the Gabor filter (Fig. 1) serves as a great tool for the detection of crack patterns.

The method we used is adopted from the method used by Jain et al. [4] for fingerprint detection. We convolve the input image with 8 differently oriented Gabor filters using angles  $0^\circ$ ,  $22.5^\circ$ ,  $45^\circ$ ,  $67.5^\circ$ ,  $90^\circ$ ,  $112.5^\circ$ ,  $135^\circ$  and  $157.5^\circ$ . However, the filter parameters, namely  $\delta_y$ ,  $\delta_x$

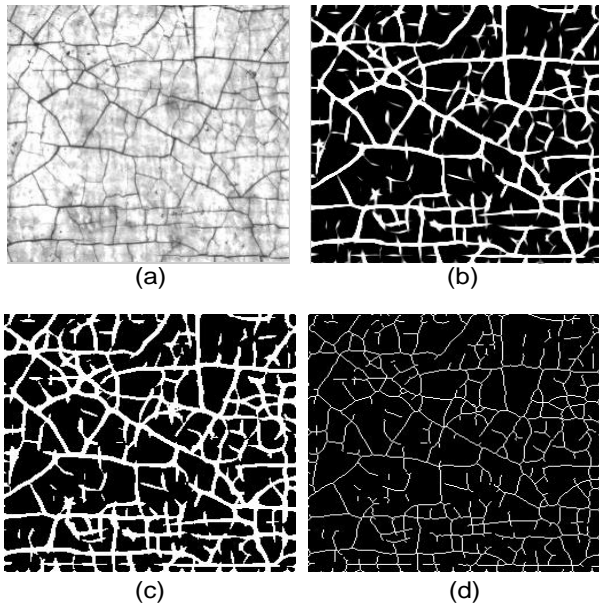
### 3. FEATURE EXTRACTION



**Fig. 1.** Gabor filter with dimension =  $33 \times 33$ ,  $\delta_x = \delta_y = 4.0$ ,  $f = 0.1$  and  $\theta = 90^\circ$ .

and  $f$  have to be properly selected in order to obtain a reasonably good output image. In scanned X-ray films the cracks vary in size between around 30 to 1 pixels, so one size of detector is not optimal and two or three sets of filters may eventually be used.

The maximum of the outputs of the filters are then combined into a result shown in Fig. 2(b). To extract the statistical features of a particular crack network, it is more convenient and less time consuming to work with 1 pixel wide cracks rather than variable sized ones. In order to do that, we apply a morphological thinning algorithm with 10 iterations to produce a thinned version of the detected cracks. Fig. 2 illustrates the results of the process. The output is used as a source for a statistical analysis of the crack network in order to create the higher level features.



**Fig. 2.** Results of the crack detection process. (a) Original. (b) Filtered. (c) Binarized. (d) Thinned.

The feature extraction stage takes into account the fact that cracks are made up of line segments. A combination of line segments produces a network which in this paper we call the crack network. Crack networks vary in multiple ways, so it is very important that the features that best distinguish them are identified. For our algorithm, we have predefined several types of crack pattern classes which have obvious distinctive features and they are circular, rectangular, spider-web, unidirectional and random. It is worth stressing that some of the relevant features may overlap with each other in terms of their unique representation of a particular crack pattern thus, revealing the importance of multiple feature combinations to classify a particular crack pattern.

#### 3.1. Statistical feature structure

A crack following algorithm is applied on a crack network and features are collected as it 'runs' along the lines. This feature extraction approach collates statistical information from the lines as a global representation and a separate process is implemented later to extract local features from the crack network structural representation. Junctions are also detected during the crack following process where their location and number of branches are stored. A junction is defined here as a point where line segments meet. A Freeman chain code based representation [5] is generated for 2 main reasons: (1) post-detection filtering and reconstruction such as thin-line code filtering [6] and (2) post-detection feature extraction. For the remaining sections, we address local features as notations with a subscript  $l$  and global features with notations with a subscript  $g$ .

#### 3.2. Local and global directionality measure

Through observation, one of the most discriminating features in a crack network is its orientation spread. Obviously, circular cracks and rectangular cracks differ significantly in terms of their orientation spread. Our approach takes this factor into account by extracting local orientation features which are then combined to produce a global feature representing a particular crack network. The algorithm explained below is used to compute a feature we call directionality:

1. Generate orientation histograms for each line segment in a single crack network and normalize the histograms such that the values are within the interval  $[0,1]$  as shown by the equations below:

$$O_l = [O_0 \dots O_N], \quad N = 7 \quad (4)$$

$$t = \sum_{i=0}^N O_i \quad (5)$$

$$O'_i = \frac{O_i}{t} \quad \forall \quad i = 0 \dots N \quad (6)$$

where  $O_i$  is the initial orientation histogram,  $t$  is the total accumulation for a single line and  $O'_i$  is the normalized orientation histogram.

2. We define 4 simple directionality histogram models where each bin represents a chain code index of 0 to 7 respectively:

$$m^1 = 0.125[11111111] \quad (7)$$

$$m^2 = 0.25[11110000] \quad (8)$$

$$m^3 = 0.5[11000000] \quad (9)$$

$$m^4 = [10000000] \quad (10)$$

$m^1$ ,  $m^2$ ,  $m^3$  and  $m^4$  represent ideal histograms as far as orientation spread is concerned where they roughly indicate circular, semicircular, bidirectional and unidirectional spread respectively.

3. Measure dissimilarity between each normalized histogram and all 4 ideal histograms using Euclidean distance measure. Dissimilarity corresponding to  $m^1$  is as shown below:

$$e^1 = \sqrt{\sum_{i=0}^7 (O_i^1 - m_i^1)^2} \quad (11)$$

where  $e_1$  represents error for a single histogram dissimilarity measure. A somewhat different approach is taken to compute dissimilarity for  $m^2$ ,  $m^3$  and  $m^4$  where each of them are rotated and dissimilarity,  $e^j$  is calculated for each rotation. The minimum dissimilarity after a complete rotation is chosen as the dissimilarity corresponding to  $m^j$  where  $j = 2, 3, 4$ .

4. Construct a similarity histogram,  $w_l = [(1 - e_1)(1 - e_2)(1 - e_3)(1 - e_4)]$ . Let  $w_l = [f_1 f_2 f_3 f_4]$ , the histogram indicates the score obtained by a particular line corresponding to each class defined in step 2.
5. Compute the directionality value by using the following equation:

$$D_l = \begin{cases} 0.25 - \frac{f_1}{4} & \text{if } \max(w_l) = f_1 \\ 0.5 - \frac{f_2}{4} & \text{if } \max(w_l) = f_2 \\ 0.5 + \frac{f_3}{4} & \text{if } \max(w_l) = f_3 \\ 0.75 + \frac{f_4}{4} & \text{if } \max(w_l) = f_4 \end{cases} \quad (12)$$

where  $D_l$  is used as one of the orientation-based local features and its significance in our classification approach will be described later. Each line in a particular network has now been assigned a directionality measure which is normalized to the interval  $[0,1]$ .

The global directionality histogram is then generated taking into account the significance of each line in the crack network. Significance measure,  $s$  is computed by taking the ratio between a line segment length and total crack network length. Computation of the line length is explained in further detail in section 3.3. The ratio is used as a weighting to indicate how much local directionality contributes to the global directionality.

Global directionality is computed using the algorithm below:

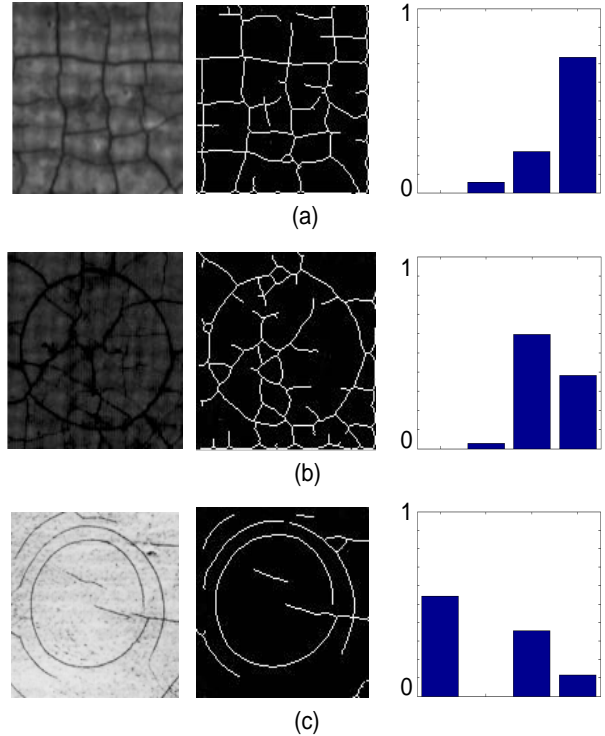
1. For all existing lines in the crack network, find the index,  $i$  ( $i \in [1,4]$ ) of the maximum similarity,  $w_l$ .

2. Multiply  $\max(w_l)$  with the significance measure,  $s$ .

3. Accumulate the resultant values in a global directionality histogram,  $w_g$  by matching it with the corresponding index,  $i$ .

4. Compute the directionality value,  $D_g$  similar to the one explained in equation (12)

Fig. 3 illustrates three examples of crack detected images and their respective directionality histograms. The algorithm is tested on real X-ray images of paintings with size 128x128. Note that the Gabor filter parameters are not automatically determined, but instead they are tuned separately for each image.



**Fig. 3.** Algorithm tested on 3 sample crack images. Figure shows original crack images, detected cracks and their corresponding directionality histograms,  $w_g$ . (a)  $D_g = 0.9340$ . (b)  $D_g = 0.6484$ . (c)  $D_g = 0.1146$ .

### 3.3. Straight line to actual length ratio

Another feature which we find useful in distinguishing different type of cracks is the straight line to actual length ratio. Straight line length is defined as the direct distance between 2 connected junctions while actual length is given as the distance through the crack path. Let us denote this particular feature as  $L$ . We employ a method used by Vossepoel and Smeulders [7] to calculate the length of the actual line by using the formula

$$p = 0.980n_e + 1.406n_o - 0.091n_c \quad (13)$$

where  $n_e$  and  $n_o$  is the number of even and odd chain elements respectively while  $n_c$  is defined as the number of occurrences of consecutive unequal chain elements in

the Freeman chain code string. The length of the straight line representation is approximated by the equation

$$q = \sqrt{(|x_1 - x_2|)^2 + (|y_1 - y_2|)^2} \quad (14)$$

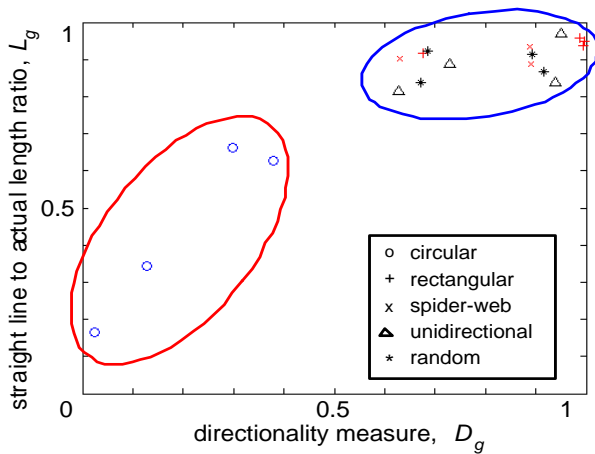
where  $|x_1 - x_2|$  represents the horizontal distance and  $|y_1 - y_2|$  represents the vertical distance. We compute  $L_i$  for each line segment in a crack network and calculate the global value,  $L_g$  by using the formula

$$L_g = \sum_i^M s_i L_i \quad (15)$$

where  $M$  is the total number of line segments in a crack network,  $s_i$  is ratio between the length of a line segment and the overall length of a crack network.

#### 4. EXPERIMENT

An experiment was conducted in order to test the potential use of the features extracted to classify crack patterns. We used 19 test images for this purpose. Each of them were initially manually classified based on our perception of its type. Features were extracted from each of the crack patterns and each of them represented by 2 features with values between the interval  $[0,1]$ . The features used were directionality measure,  $D_g$  and straight line to actual length ratio,  $L_g$ . We constructed a 2-dimensional plot representing each crack network in order to show their position in the feature space (Fig. 4). It can be seen



**Fig. 4.** Figure shows location of different crack types in feature space where 2 clusters are clearly visible as shown by the circles.

that there are two main groups clearly separating the circular from the other types, so the first classification stage has been successful.

#### 5. CONCLUSION AND FUTURE WORK

In this paper, we have discussed the detection of crack patterns as well as the extraction of useful features that might be very useful to classify cracks into different classes. Experimentally it is shown that multi-oriented Gabor filters can be used to extract suspected crack patterns. Although there are various other methods to detect ridge

and valley structures, Gabor filter has a good potential as far as detecting cracks are concerned since it provides the ability for scale-specific and multi-orientation detection.

Extraction of features corresponding to orientation has also been discussed. The Freeman chain code representation serves two main purposes: 1) post-detection filtering and reconstruction 2) post-detection feature extraction. We have also experimented on the effectiveness of features corresponding to orientation especially in its capability to separate circular cracks with the rest of the crack classes in the feature space.

As for future direction, we believe that successful classification of cracks really depends on inter-relation between features of different nature and that is what we are trying to achieve.

#### REFERENCES

- [1] S. Bucklow, "The Description of Craquelure Patterns," *Studies in Conservation*, vol. 42, 1997.
- [2] I. Giakoumis, I. Pitas, "Digital Restoration of Painting Cracks," in *ISCAS '98, Proceedings of the IEEE International Symposium on Circuits and Signals*, 31 May-3 June 1998, pp. 269-272.
- [3] M. Barni, F. Bartolini, V. Cappellini, "Image Processing for Virtual Restoration of Artworks," *IEEE Multimedia*, vol. 7, no. 2, pp. 34-37, April-June 2000.
- [4] A.K. Jain, S. Prabhakar, L. Hong, "A Multichannel Approach to Fingerprint Classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 4, pp. 348-359, April 1999.
- [5] H. Freeman, "Computer Processing of Line Drawing Images," *Computer Surveys*, vol. 6, no. 1, pp. 57-98, 1974.
- [6] L. O'Gorman, "Primitives Chain Code," in *ICASSP-88, International Conference on Acoustics, Speech, and Signal Processing*, 11-14 April 1988, pp. 792-795.
- [7] A.M. Vossepoel, A.W.M. Smeulders, "Vector Code Probability and Metrication Error in the Representation of Straight Lines of Finite Length," *Computer Graphics and Image Processing*, vol. 20, 1982.

#### Acknowledgements

We would like to thank Centre of Research and Restoration of the museums of France (C2RMF), Victoria and Albert Museum, National Gallery (London), Uffizi Gallery (Florence), NCR and Giunti Interactive Labs as well as Dr. Spike Bucklow (Hamilton Kerr Institute, Cambridge) and Dr. Paul H. Lewis (University of Southampton) for their support.