International Journal of Engineering Research in Computer Science and Engineering (IJERCSE) Vol 3, Issue 6, June 2016

# Classification of Cracks Old Digital Paintings using Fuzzy Entropy Based Clustering

Shruti Garg
BIT, Mesra/Dept. of CSE, Ranchi, INDIA

***Abstract:*** **Paintings are the representation of an artist thought. Famous paintings, like that of Mona Lisa, have a great value attached to them. Unfortunately, paintings cannot be preserved in their original state as they are adversely affected by atmospheric conditions. Cracks in old paintings are one such phenomenon observed very commonly. To identify cracks properly, they need classification from dark black brush strokes before filling. Then cracks are classified by unsupervised classification methods. And experimental results show that fuzzy entropy based clustering is suitable for classification of cracks.**

*Index Terms*—**Top hat transform, Fuzzy C means, K-means, Entropy**

## I. INTRODUCTION

Crack classification is required for removing cracks from old digital paintings. Cracks may be generated in paintings because of heat, humidity moisture etc or bad paint material. So there is a need for restoration of these paintings. For this purpose, three basic steps are required: Crack detection, Crack classification, and restoration. This paper is concern for crack classification only among three.

Cracks in paintings can be defined as a line which does not physically separate them. In the digital form of paintings the cracks are low eliminated areas also they are similar to dark black lines which can be extended in any direction. Thus top hat transform is applied to identify them in this work. These filters pass all low frequencies, which may consist of thin dark brush strokes. Thus to achieve a good quality of restoration separation of brush stokes and cracks is required, which is called classification of crack.

To classify the cracks four different methods had been applied. 1) K- Nearest Neighborhood 2) K-Means Clustering 3) Fuzzy c Means Clustering 4) Fuzzy Entropy Based Clustering

Cracks are dark lines in the bright background; thus initially, they are considered as ridge and valley structures[2][3][4][5] and segmentation methods had been applied to detect them such as thresholding, various line detector, etc. These methods wont work if the image has multiple luminances. Thus, the detection of cracks was done by applying top hat transform and median radial basis function network(MRBFN) had been applied for classifying them[1][13][7]. The MRBFN is a supervised classification method which requires desired output. Semiautomatic classification techniques were also be proposed in [14][1][6]. The semiautomatic approaches require user intervention.

There are five different sections present in this paper. Section II gives the process of crack detection. Various classification methods are discussed in section III. The results and discussions are described in Section IV. And finally, conclusions are stated in section V.

## II. CRACK DETECTION

Cracks are dark black lines in image. Hence the cracks can be considered as the local intensity minima that can be extended in any direction. A crack detection filter should be applied on the luminance component of an image; to identify such minima was presented in [1]. The morphological filter, called top-hat transformation is used to detect cracks in this work. The top-hat transform can be defined as[1]:

$$y(x) = f(x) - f_{nB}(x) \quad (1)$$

where $f_{nB}(x)$ is the opening of the function with the structuring set nB, defined as:

$$nB = B \oplus B \oplus B \ldots..B \ \ (n\ times) \quad (2)$$

The opening $f_{nB}(x)$ function erases all peaks (local maxima) in which the structuring element nB cannot fit because it is a nonlinear low pass filter. Thus, the image $f - f_{nB}$ contains only peaks without background. The cracks have very low pixel values generally. Thus, for identifying the cracks, the negation of luminance image is required before applying the top-hat transformation. The cracks can also be identified by applying closing filter on the original image *f(x)* with the structuring element nB and then subtracting f*(x)* from the result of closing $f^{nB}(x)$.

$$y(x) = f^{nB}(x) - f(x) \qquad (3)$$

However, the result of crack detection can vary by changing the given parameters:

- The size of structuring element B and its type such as it can be circular rather than square;
- Number of times dilation is repeated.

The change in the size of the final structuring element nB can change the thickness of the cracks to be detected. The structuring element taken here is square type and its size taken as 3× 3 and two times dilation operation is repeated because the thickness of cracks to be detected is targeted from 2 to 6 pixels. The output of top-hat transform is a grayscale image *t(k,l)* which consists pixels with a large grey value that is either a crack or crack-like structure. Further, a threshold operation is applied to *t(k,l)* to separate crack pixels from the remaining image. The threshold T for crack identification can be chosen by the histogram of top hat transform output. The threshold value is selected in such a way so that small percentage of the pixels *t(k,l)* remain above it. The selection of threshold by histogram is called global thresholding has given in [7]. The threshold value T is decided experimentally here. The value of threshold varies image to image based on their contrast. The output of the thresholding gives binary image b*(k,l)* which consist all crack pixel values as 1 that is termed as crack map in this paper.

## III. CRACK CLASSIFICATION

In many paintings, the thin dark brush strokes are having almost same thickness and luminance to cracks. For example hairs and beard of a person in a painting can be same as cracks appear in it. Thus, in order to avoid filling these thin dark brush strokes in the original paintings, it is important to separate brush strokes pixels from the actual crack pixels, before applying crack filling procedure. There are four different methods of classification are described in this paper. All of them are unsupervised as they don't require desired outputs.

### A. K-nearest neighborhood method

This is a supervised classification method. All pixels in the image were checked for all the neighbor, and all those points having Euclidean distance less than the threshold value 2 (obtained after trial and error method) from that data point will belong to class crack and the points whose distance is more than the threshold value will belong to brush stroke class.

Euclidean n-space, the distance from p to q is given by:

$$d(p,q) = \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2} \qquad (4)$$

### B. K-means clustering method

K-means is a simple, well known algorithm for grouping objects, called clustering. Here objects need to be represented as a set of numerical features. The requirement to execute the algorithm, two groups one for cracks and other for brushstrokes has to be specified earlier. Each object can be thought of as being represented by some feature vector in a one dimensional space that is intensity. The algorithm starts with two random points, these points serve as the initial centers of the clusters. After that all objects are assigned to each centre they are closest to. For each cluster, a new centre is computed by averaging the feature vectors of all objects assigned to it. This process is repeated until the process converges.

The objective function

$$J = \sum_{j=1}^{k}\sum_{i=1}^{n}||x_i^{(j)} - c_j||^2 \qquad (5)$$

$||x_i^{(j)} - c_j||^2$ is a chosen distance measure between a data point $x_i^{(j)}$ and the cluster $c_j$ centre, is an indicator of the distance of the n data points from their respective cluster centers.

### C. Fuzzy c means Algorithm

The fuzzy k-means algorithm (FKM) (which is also called the fuzzy c-means) [13] is an adaptation of the k-means (KM) [14] algorithm that uses soft membership function.

Unlike KM which assigns each data point to its closest centre, the FKM algorithm allows a data point to belong partly to all centers.

The objective function for FKM is:

$$FKM(X,C) = \sum_{i=1}^{n}\sum_{j=1}^{n}u_{ij}^{r}||x_i - c_j||^2 \qquad (6)$$

Where the fuzzy membership matrix $u_{ij}$ denotes the proportion of data point $x_i$ that is assigned to centre $c_j$, and is under the constraint $\sum_{j=1}^{k}u_{ij} = 1$ for all $i$ and $u_{ij} \geq 0$ and$r \geq 1$. The method is called "more fuzzy" if the value of r is large.

Bezdek [15] presented separate update function for $u_{ij}$ and $c_j$. The $u_{ij}$ update equation depends only on C and X, so its update will be the update of $c_j$. The membership function for FKM is:

$$m_{FKM}(c_j/x_i) = \frac{||x_i - c_j||^{\frac{-2}{r-1}}}{\sum_{i=1}^{k}||x_i - c_j||^{\frac{-2}{r-1}}} \qquad (7)$$

As $r$ tends to approach 1 from above, fuzziness decreases and the algorithm behaves more like standard KM. The centres share the data points less in this condition. As for the FKM technique, the centre points and the fuzzy membership matrix are the parameters which have to be initialized prior to the clustering process. The soft membership scheme introduced by the algorithm suits the nature of crack patterns which are quite subjective in terms of their class origin. Fuzziness serves the application well since we would like to expect perceptions towards crack pattern labeling to be described as a confidence measure instead of being an absolutely certain decision. A human observer for instance, might have divided perceptions about a particular pattern and so does a fuzzy classifier system.

### *D. Fuzzy Entropy based Clustering*

Entropy is a term that gives "degree of disorder". It is observed that there is similarity within the pixels of one object and these pixels must be different than pixels of another object. The same concept is used to calculate entropy of each pixel. Entropy of a pixel is given as:

$$E_{i,j} = -\left(S_{i,j} lg S_{i,j} + (1 - S_{i,j}) lg(1 - S_{i,j})\right) \quad (8)$$

Where

$$S_{i,j} = e^{-\alpha d_{i,j}} \quad (9)$$

Where $d_{i,j}$ is Euclidean distance of pixels from its 3x3 neighbors?

The value of entropy is generally very less or close to zero. In order to divide the result of top hat transform in to different cluster. The minimum entropy pixel has been taken and the values close to this minimum had been selected as seed point of cluster. The value of $\alpha$ is taken .5 here from [16].

***Algorithm:***

1. Calculate entropy of each $x_i$ in image.
2. Find first minimum entropy pixel $x_{min1}$ and second minimum entropy pixel $x_{min2}$ in image
3. Use $x_{min1}$ and $x_{min2}$ as seed points for fuzzy k mean clustering

## IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

Experiments are performed on 26 different images among which few results are shown here.

***Fig. 1 Original Image***

***Fig. 2 Threshold output***

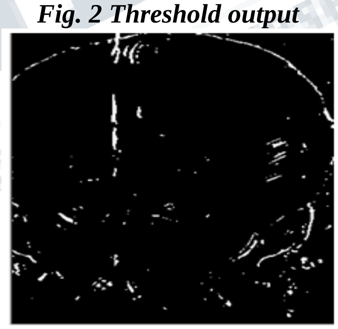

***Fig. 3 Cracks obtained by k-Nearest Neighbor***

***Fig. 4 Brush Strokes obtained by k-Nearest Neighbor***

*Fig. 5 Cracks obtained by K-Mean Clustering*

*Fig. 6 Brush Strokes obtained by K-Mean Clustering*

*Fig. 7 Cracks obtained by Fuzzy c Means Clustering*

*Fig. 8 Brush Strokes obtained by Fuzzy c Means Clustering*

*Fig. 9 Cracks obtained by Fuzzy Entropy based Clustering*

*Fig. 10 Brush Strokes obtained by Fuzzy Entropy based Clustering*

*Fig. 11 Original Image*

*Fig. 12 Threshold output*

*Fig. 13 Cracks obtained by k-Nearest Neighbor*

*Fig. 14 Brush Strokes obtained by k-Nearest Neighbor*

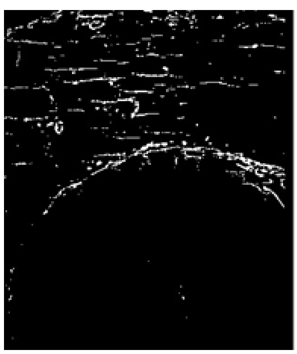

*Fig. 15 Cracks obtained by K-Mean Clustering*

*Fig. 16 Brush Strokes obtained by K-Mean Clustering*

*Fig. 17 Cracks obtained by Fuzzy c Mean Clustering*

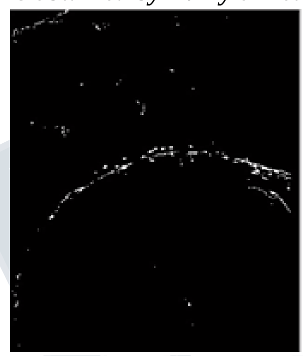

*Fig. 18 Brush Strokes obtained by Fuzzy c Means Clustering*

*Fig. 19 Cracks obtained by Fuzzy Entropy based Clustering*

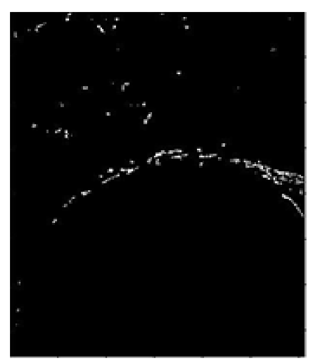

*Fig. 20 Brush Strokes obtained by Fuzzy Entropy based Clustering*

The above results shows that k-nearest neighborhood is giving good classification but the hairs, beard and eyes are classified in crack class and brush stroke class also consist the outline of cracks shown in fig. 3,4,13,14. Classification output by k-means clustering is not desirable shown in fig 5,6,15,16. The output of fuzzy c means clustering shown in fig 7,8,17,18 is good but it has more number of hair and beard pixels in crack class and outline if crack pixels is present in brush stroke class which is not there in fuzzy entropy based clustering shown in fig 9,10,19,20.

## V. CONCLUSIONS

This paper consist identification of cracks by top hat transform then the brush strokes and cracks are classified. For the classification purpose four different methods of classification has been applied to various images. The results of fuzzy c- means classifier is best for many of the paintings its drawback is, it picks random values which works initial cluster centers. Thus it is observed that the classes may interchange in different run. The last method described in this paper fuzzy entropy based clustering also give good results as it gives lesser brush strokes than fuzzy c means clustering also lesser number of crack pixels is misclassified as brush strokes. The conclusion is no automatic method can provide perfect classification. The least misclassification is found in fuzzy entropy based clustering. And there no random selection of initial clusters in fuzzy entropy based clustering so output will remain same in different execution whereas classes may interchange in fuzzy c-mean clustering.

## REFERENCES


1. I. Giakoumis, N. Nikolaidis, I. Pitas, "Digital image processing techniques for the detection and removal of cracks in digitized paintings," IEEE *Transactions on Image Processing* vol.15, No.1, pp. 178-188, 2006.

2. A. M. Lòpez, F. Lumbreras, "Evaluation of Methods for ridge and valley detection," *IEEE Transaction Pattern Analysis and Machine Intelligence, vol*. 21, pp. 327–335, 1999.

3. D. Elberly, R. Gardner, B. Morse, S. Pizer, C. Scharlach, "Ridges for image analysis," *Journal of Mathematical Imaging and Visualization*, vol. 4, pp. 353–373, 1994.

4. J. Gaugh, S. Pizer, "Multiresolution Analysis of Ridges and Valleys in Grey-Scale Images," *IEEE Transaction Pattern Analysis and Machine Intelligence*, vol. 15, No. 6, pp. 635–646, 1993.

5. R. Haralick, "Ridges and Valleys on Digital Images Computer Vision," *Graphics and Image Processing*, vol. 22, No.10, pp. 28–38, 1983.

6. M. Barni, F. Bartolini, V. Cappellini, "Image Processing for Virtual Restoration of Artworks," *IEEE. Multimedia*, vol. 7, no. 2, pp. 34-36, 2000.

7. A G Bors, I Pitas, "Median radial basis function neural network," *IEEE Transaction of Neural Network*, vol. 7, no 6, pp. 1351–1364, Nov 1996.

8. A. Kokaram, R. Morris, W. Fitzgerald, P. Rayner, "Interpolation of missing data in image sequences," *IEEE Transactions on Image Processing*, vol. 4, no. 11, pp. 1509-1519, 1995.

9. C . Ballester, M. Bertalmio, V. Caselles, G. Sapiro, J. Verdera, "Filling-In by Joint Interpolation of Vector Fields and Gray Levels," *IEEE Transactions on Image Processing*, vol. 10, no. 8, pp. 1200–1211, 2001.

10. T. F. Chan, J. Shen, "Non-texture inpaintings by curvature-driven diffusions," *Journal of Visual Communication and Image Representation*, vol. 12, no. 4, pp. 436-449, 2001.

11. S. Esedoglu, S. Shen, "Digital Inpainting Based on the Mumford-Shah-Euler Image Model," *European Journal of Applied Mathematics*, vol. 13, pp. 353-370, 2002.

12. I. Pitas, C. Kotropoulos, N. Nikolaidis, R. Yang, M. Gabbouj, "Order statistics learning vector quantizer," *IEEE Transaction of Image Processing*, vol. 5, no .6, pp. 1048–1053, 1996.

13. A. Gupta, V. Khandelwal, A. Gupta, M. Srivastava, "Image Processing Methods for the Restoration of Digitized Paintings," *Thammasat International Journal of Science and Technology*, vol. 13, no.3**,** pp. 66-72, July 2008**.**

14. S. V. Sachin, A. R. Mahajan, "Cracks Inspection and Interpolation in Digitized Artistic Picture using Image Processing Approach," *International Journal of Recent Trends in Engineering*, vol. 2, no. 2, pp. 97-99, May 2009.

15. S. Rajashekharan, G. A. Vijyalaxmi Pai, *Neural Network, Fuzzy Logic and Genetic Algorithms-Synthesis and Application*, 6rd ed., Prentice-Hall of India Pvt. Ltd, 2003, pp. 34-86.

16. J. Yao, M. Dash, S. T. Tan, H. Liu, “Entropy based Fuzzy Clustering and Fuzzy Modelling, ” Fuzzy Sets and Systems, 113, pp. 381-388, 2000