iii.) Validation of Centerline Segments

Every candidate segment is confirmed or rejected as a valid centerline segment based on two features:

- Intensity of the segment, evaluated by the geometric mean between the average and maximum intensity values of the segment. The reference values associated with direction x, I_x^{ref} , is defined in equation (2) as

$$I_x^{ref} = \max(I_x, I_g),\ for\ x = h, v, d_{45}, d_{135} \qquad (2)$$

where L_x, is the directional intensity value calculated from the subset of candidate segments with direction x and and. are obtained from four subsets.

- Length of the segment, measured by the number of points of the candidate segment. The length value is approximated using a scaling factor equal to (-1 2). The length reference values associated with direction x, L_x^{ref} is defined in equation (3) as

$$L_x^{ref} = \max(L_x, L_g)\ for x = h, v, d_{45}, d_{135} \qquad (3)$$

Comparing its intensity and length features with image dependent reference values that validates each centerline candidate segment. The classification of each individual segment s of a particular directional set x as a crack centerline segment is based on the inequality presented in equation (4). In this equation, where a product rule is used for combining the two feature values, I_x^s and L_x^s are, respectively, the intensity intensity and length features calculated intensity and the length features calculated for the segment s, and I_x^{ref}, L_x^{ref}, are the references obtained from (2) and (3) for the directional set containing the segment under analysis

$$I_x^s \sqrt{L_x^s} \geq I_x^{ref} \sqrt{L_x^{ref}}\ for\ x = h, v, d_{45}, d_{135} \qquad (4)$$

C. Crack Segmentation

a. Crack Enhancement

In our method, the background-normalized image is processed by top-hat operators using elements of increasing radius. The option for such a multiscale approach, instead of using a single monoscale operator, is justified by the expected dependence of the operator response from the vessel width. The range of the radius of the structuring elements varies from 1 to 8 pixels, covering the overall range of crack widths. For images with different scales, the set of structuring element radii should be adapted accordingly. The classical top-hat transform is defined as the difference between an image and its opened version. A problem associated with this classical implementation is sensitivity to noise, as a consequence of the fact that pixel values in an opened image are always less than or equal to the original ones; in such conditions, the difference image retains all small intensity fluctuations that can be found in the data. To overcome this inconvenience, a modification was adapted from [13], by considering two new steps in the top-hat definition: a closing precedes the opening result which is followed by a comparison, using a minimum operator, to get an image equal to the original one everywhere except for peaks and ridges. Equation (5) represents this modified top-hat transform, where I is the image to be processed, while S_c and S_o stand for the structuring elements for closing ($\bullet$) and opening ($\circ$) operators

$$TopHat_I = I - \min(I \bullet S_c) \circ S_o; I) \qquad (5)$$

As a consequence of the distinct size of the structuring elements, each image is adapted to highlight cracks

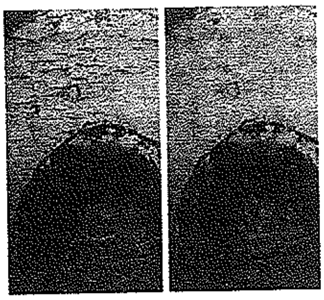


Figure 5. Output for Crack Filling

Figure 5-a Original Image **Figure 5-b** Crack filled (region growing method)

within a limited range of widths. Essentially, the images obtained with the smallest structuring elements just retain correct width data for small cracks, since only the central part of the other crack segments is represented. On the other hand, larger structuring elements are able to extract larger cracks, but thinner structures are naturally blurred.

b. Crack Filling Method

I) Crack Filling Based On Seed Points And Region Growing Technique

After identifying cracks and separating misclassified brush strokes, the final task is to restore the image using local image information to fill the cracks. We obtain a final image with the segmented cracks by iteratively combining the centerline image with the set of images that resulted from the crack segments reconstruction phase. In the first iteration, crack centerline pixels are used as seeds for a region-growing algorithm, which breed these points by aggregating the pixels in the reconstructed image derived from the top-hat operator with the smallest structuring element size. In each of the subsequent three iterations, the reconstructed images corresponding to the cracks with increasing width are in turn used for extending the output of the previous region-growing step.

The final crack segmentation is obtained after a cleaning operation aiming at removing all pixels completely surrounded by crack points, but not labeled as part of a crack. Considering that each pixel with at least six neighbors classified, as crack points must also belong to a crack does this. The final results of the iterative crack-filling phase are illustrated in Figure 5.

4. Controlled Anisotropic Diffusion

Anisotropic diffusion [12] is an image enhancement method that successfully combines smoothing of slowly varying intensity regions and edge enhancement. Smoothing is modeled as a diffusion that is allowed along homogeneous regions and inhibited by region boundaries. Anisotropic diffusion is described by the following equation (6)

$$\frac{\partial I(x,y,f)}{\partial f} = div \quad c(x,y,f \quad \nabla I(x,y,f)) =$$
$$d(x,y,f) \quad \Delta I(x,y,f) + \nabla c(x,y,f) \cdot \nabla I(x,y,f) \quad (6)$$

Where, div denotes the divergence operator and ∇, Δ the gradient and Laplacian operators with respect to the space variables x, y. At each position and iteration, diffusion is controlled by the conduction Coefficientc(x,y,f).Since diffusion should be inhibited across regions separated by discontinuities,the conduction coefficients should obtain small values in pixels with large intensity gradient magnitude.the iterative discrete solution to equation (6) is governed by the equation (7)

$$I^{t+1}_{i,j} = I^{t}_{i,j} + \lambda[C_N \cdot D_N I + C_S \cdot D_S I + C_E \cdot D_E I + C_W \cdot D_W I]^{t}_{i,j} \quad (7)$$

where 0 = = ¼ for the scheme to be stable, N, S, E, and W are the mnemonics for North, South, East, and West, and the symbol D indicates nearest-neighbor differences as in equation (8).

$$\begin{aligned} D_N I_{i,j} &\equiv I_{i-1,j} - I_{i,j} \\ D_S I_{i,j} &\equiv I_{i+1,j} - I_{i,j} \\ D_E I_{i,j} &\equiv I_{i,j+1} - I_{i,j} \\ D_W I_{i,j} &\equiv I_{i,j-1} - I_{i,j} \end{aligned} \tag{8}$$

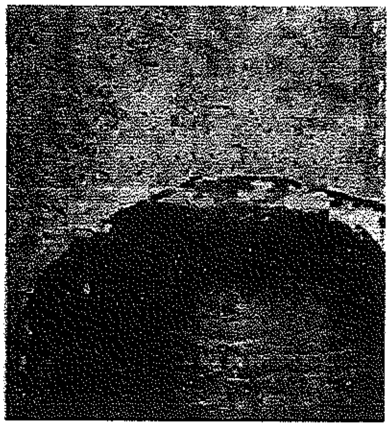
The conduction coefficients are evaluated at every iteration as a function of the magnitude of the intensity gradient. In our implementation, the following approximation was used as in equation (9).

$$\begin{aligned} C^t_{N_{i,j}} &= g\left(\left|D_N I^t_{i,j}\right|\right) \\ C^t_{S_{i,j}} &= g\left(\left|D_S I^t_{i,j}\right|\right) \\ C^t_{E_{i,j}} &= g\left(\left|D_E I^t_{i,j}\right|\right) \\ C^t_{W_{i,j}} &= g\left(\left|D_W I^t_{i,j}\right|\right) \end{aligned} \tag{9}$$

The following function equation (10), proposed in [12], and has been used in our case:

$$g(\nabla I) = \frac{1}{1 + \left(\frac{\|\nabla I\|}{K}\right)^2} \tag{10}$$

In our experiment, we fixed the value of K as 25. In order to fill the cracks, the anisotropic diffusion algorithm was applied selectively, in neighborhoods centered on crack pixels. All pixels within these neighborhoods participate in the diffusion process. However, only the values of the crack pixels are updated in the output image. Further improvements were obtained by taking into account crack orientation, i.e., by applying the operation only in a direction perpendicular to the crack direction. The final results of the iterative crack-filling phase by anisotropic Diffusion method are shown in Figure 6.

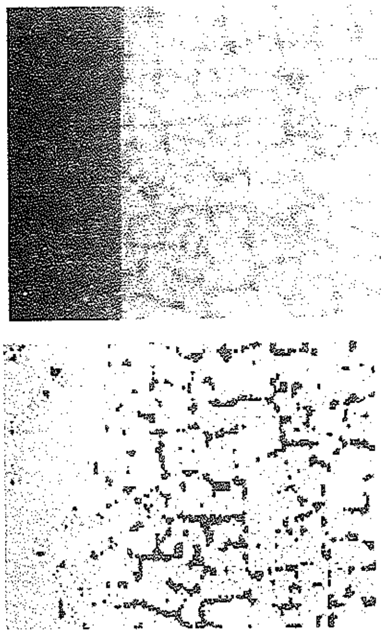


(a)

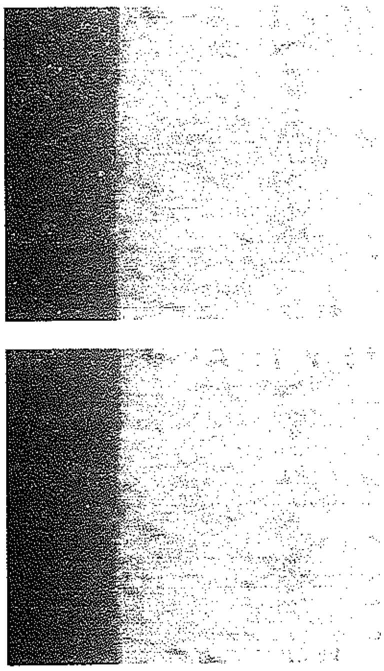
(b)

Figure 6. Output For Crack Filling
a) Original Image
b) Crack Filled (Controlled Anisotropic Diffusion)

The final results of the iterative crack-filling phase by region growing method and Anisotropic Diffusion method is shown in figure 7.



(a) and (b)



(c) and (d)

Figure 7. Output For Crack Filling
a) **Original** Image
b) **Detected** Crack Centerlines
C) **Crack** Filled (Region Growing Method)
d) **Crack filled** (Controlled Anisotropic **Diffusion**)

5. Discussion And Conclusion

In this paper, we have presented an integrated strategy for crack detection and filling in digitized paintings. Our method can be classified as a pixel processing-based approach. The initial step of crack centerline detection combines local information, used for early pixel selection, with structural features, as the crack length. Global intensity characteristics and local crack width information are adaptively exploited in the subsequent crack-filling phase.. A major feature of the method is its adaptability to particular image intensity properties, as most algorithm settings are based on threshold values computed from local or global image information. In our work Cracks are filled by using region growing technique in the seed points (Figure 7.c) and Anisotropic Diffusion technique(Figure 7.d). Region growing approach gives better results. When compared with crack segmentation approaches, the present proposal assigns a hard classification result to each image point.

References

[1] M. Barni, F. Bartolini and V. Cappellini, *"Image processing for virtual restoration of artworks"*, IEEE Multimedia, Vol. 7, No. 2, PP 34-37, Jun. 2000.

[2] F. Abas and K. Martinez, *"Craquelure analysis for content-based retrieval"*, in Proc. 14th Int. Conf. Digital Signal Processing, Vol. 1, PP 111-114, 2002

[3] L. Joyeux, O. Buisson, B. Besserer and S. Boukir, *"Detection and removal of line scratches in motion picture films"* in Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, PP 548-553, 1999

[4] A. Kokaram, R. Morris, W. Fitzgerald and P. Rayner, *"Detection of missing data in image sequences,"* IEEE Trans. Image Process., Vol. 4, No. 11, PP 1496-1508, Nov. 1995.

[5] *"Interpolation of missing data in image sequences"*, IEEE Trans. Image Process., Vol. 4, No. 11, PP 1509-1519, Nov. 1995.

[6] M. Bertalmio, G. Sapiro, V. Caselles and C. Ballester, *"Image inpainting"*, in Proc. SIGGRAPH, PP 417-424, 2000.

[7] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, and J. Verdera, *"Filling-in by joint interpolation of vector fields and gray levels"*, IEEE Trans. Image Process., Vol. 10, No. 8, PP 1200-1211, Aug. 2001.

[8] S. Masnou and J. M. Morel, *"Level lines based disocclusion,"* in Proc. IEEE Int. Conf. Image Process., Vol. III, 1998, PP 259-263, 1998

[9] T. Chan and J. Shen, *"Non-texture inpaintings by curvature-driven diffusions,"* J. Vis. Commun. Image Represen., Vol. 12, No. 4, PP 436-449, 2001.

[10] A. M. Mendonça, A. J. Campilho and J. R. Nunes, *"Analysis of perifoveolar region in eye fundus images"*, in Proc. 7th Scand. Conf. Image Anal., Vol. 1, PP 545-552, 1991.

[11] W. K. Pratt, *"Digital Image Processing"*, 3rd ed. New York: Wiley, 2001.

[12] P. Perona and J. Malik, *"Scale-space and edge detection using anisotropic diffusion,"* IEEE Trans. Pattern Anal. Mach. Intell., Vol. 12, No. 7, PP 629-639, Jul. 1990.

[13] P. Salembier, *"Comparison of some morphological segmentation algorithms based on contrast enhancement-Application to automatic efect detection,"* in Signal Process. V: Theories Appl., L. Torres, E.Masgrau and M. A. Lagunas, Eds. Amsterdamn, The Netherlands: Elsevier, 1990, PP 833-836, 1990

Author's Biography

G.Wiselin Jiji graduated her Engineering (CSE) from Manonmanium Sundarnar University in 1994 and completed her post graduation from Madurai Kamaraj University in 1998. She is currently submitted her research thesis in Anna University, Chennai. Her current research interests include remote sensing, image processing and texture analysis. He is working as Head of the Department in Computer Science and Engineering , Dr.Sivanthi Aditanar College Of Engineering, Tiruchendur. She Published 8 research papers in International Journals , one paper in National Journal and 25 papers in International /National Conferences.

Dr. L.Ganesan completed his B.E in Electronics and Communication Engg. from Thiagaraja College of Engineering, Madurai and M.E in Computer Science and Engg. from Government College of Technology, Coimbatore. He completed his Ph.D from Indian Institute of Technology, Kharapur in the area image processing. He has authored more than fourty publications in reputed International Journals and more than 150 National / International conferences. His area of interest includes image processing, multimedia and compressions. He is currently working as head of the department in Computer science and engineering, A.C. College of Engg. and Technology, Karaikudi.

Discovering Significant Rules For Constructing An Network Intrusion Detction Model

Mrutyunjaya Panda[1], Manas Ranjan Patra[2]

Abstract

Network intrusion detection systems (NIDSs) have become an important component in network security infrastructure. Currently, many NIDSs are rule-based systems whose performances highly depend on their rule sets.Unfortunately,due to the huge volume of network traffic, coding the rules by security experts becomes difficult and time consuming. Since, data mining techniques can build intrusion detection model adaptively, data mining based NIDSs have significant advantages over the rule-based system.Therefore, we apply Association rule mining for network intrusion detection. Its mining algorithms discover all item associations (or rules) in the data that satisfy the user-specified minimum support (minsup) and minimum confidence (minconf) constraints. In many applications, association rules will only be interesting if they represent non-trivial correlations between all constituent items. In this paper, Apriori and predictive Apriori methodologies are used for discovering the significant rules .We also report the experimental results over KDDCup'99 datasets. Finally, the result analysis shows that the proposed approach provides the most significant rules, thereby providing strict control over false discoveries.

Keywords: Intrusion Detection, Data Mining, Association Rule Mining, Network Security.

[1]Deaprtment of ECE, GIET, Gunupur, Orissa, India-765022, e-mail : panda_mrutyunjaya@yahoo.com

[2]Department of Computer Science, Berhampur, University, Orissa, India.
e-mail : panda_mrutyunjaya@yahoo.com,
mrpatra12@gmail.com

1. Introduction

With the tremendous growth of network-based services and sensitive information on networks, network security is getting more importance than ever. Although a wide range of security technologies such as information encryption, access control, and intrusion prevention can protect network -based systems, there are still many undetected intrusions.thus,network intrusion detection systems(NIDSs) play a vital role in network security. Network intrusion detection systems detect attacks by observing various network activities, while Host-based Intrusion Detection systems (HIDSs) detect intrusions in an individual host. An IDS usually does not affect the normal network operation of the targets.

There are two major approaches for detecting intrusions, signature -based and anomaly-based intrusion detection. In the first approach, attack patterns or the behaviour of the intruder is modelled (attack signature is modelled). Here, the system will signal the intrusion once a match is detected. However, in the second approach, normal behaviour of the network is modelled. In this approach, the system will raise the alarm once the behaviour of the network does not match with its normal behaviour. There is another Intrusion Detection (ID) approach that is called specification -based intrusion detection. In this, the normal behaviour (expected behaviour) of the host is specified and consequently modelled [1]. In this approach, as a direct price for the security, freedom of operation for the host is limited. Currently, many NIDSs such as Snort [2] are rule based systems, which employ misuse detection techniques and have limited extensibility for novel attacks. Their performances highly rely on the rules identified by the security experts. However, the amount of network traffic is huge, and it is very difficult to specify some intrusions using the rules. Therefore, the process of encoding rules is expensive and slow.

To overcome the rule based systems, a number of IDSs employ data mining techniques. Data mining is the analysis of (often large) observational data sets to find patterns or models that are both understandable and useful to the data owner [3]. Data mining can efficiently extract patterns of intrusions for misuse detection, establish profiles of normal network activities for anomaly detection, and build classifiers to detect attacks, especially for the vast amount of audit data. Data mining-based systems are more flexible and deployable. The security experts only need to label the audit data to indicate intrusions instead of hand-coding rules for intrusions.

Another major problem in this research area is the speed of detection. Computer networks have a dynamic nature in a sense that information and data within them are continuously changing. Therefore, detecting an intrusion accurately and promptly, the system has to operate in real time. Operating in real time is not just to perform the detection in real time, but also is to adapt to the new dynamics in the network. Real time operating IDS is an active research area perused by many researchers. Most of the research works are aimed to introduce the most time efficient methodologies. The goal is to make the implemented methods suitable for the real time implementation.

Association rule discovery [4] finds collections of items that co-occur frequently in data. In many applications, such rules will only be interesting if they represent non-trivial correlations between all constituent items. In this paper, such associations are described as significant rules and all remaining associations are treated as false discoveries [5, 6, and 7]. This paper builds upon this body of previous work in terms of arbitrary statistical hypothesis tests, while providing strict control over the risk of false discoveries. The proposed approach is evaluated using the KDDCup'99 datasets, which were used for the third International Knowledge Discovery and data Mining Tools Competition [8]. The contest involved building a classifier for detecting computer network intrusions from a very large database of network traffic. Our experimental results show that the detection performance is improved by employing various interestingness measures, using Apriori and Predictive Apriori algorithm. The rest of the paper is organised as follows. The section2, we describe about the problem statement of the proposed approach. The experiments and performance evaluations are presented in section3. In section4, some related work will be discussed. Finally, conclusion and future scope of the research will be outlined in section5.

2. Problem Statement

In this section, an overview of the proposed framework is discussed at first. Then, we will illustrate how to apply the association rule mining to build patterns with the most significant rules for detecting network intrusions.

2.1. Overview Of The Framework

The proposed framework applies association rule mining technique to build patterns for network intrusion detection. The working environment of the proposed