Transportation Informatics: An Image Analysis System for Managing Transportation Facilities – Phase II

The University of Toledo – University Transportation Center

Final Report

Eddie Y. Chou
Principal Investigator

Ezzatollah Salari
Co-Investigator

February 2012

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

Transportation Informatics: An Image Analysis System for Managing Transportation Facilities – Phase II

Executive Summary

One of the most important tasks in maintaining transportation facilities such as highways and streets is the evaluation of the existing condition. Visual evaluation by human inspectors is subjective in nature, therefore has issues of consistency and the speed and frequency of evaluation is limited due to the manual process. Automated evaluation using modern digital image processing and pattern recognition techniques can increase the efficiency and accuracy and decrease the costs of condition evaluation. Several automated condition evaluation systems have been developed, but these systems commonly require special devices such as strobe light, laser beams, etc, which increase the cost and limit the system to certain applications. In this study, a low cost automatic pavement distress evaluation approach is presented. This method can provide real-time pavement distress detection as well as evaluation results based on color images captured from a camera installed on a survey vehicle. The entire process consists of two main parts: pavement surface extraction followed by pavement distress detection and classification. In the first part, a novel color segmentation method based on a feed forward neural network is employed to separate the road surface from the background. In the second part, a segmentation technique based on probabilistic relaxation is used to separate distress areas from the road surface. The geometrical parameters obtained from the detected distresses are then fed to a neural network based pavement distress classifier in which the defects are classified into different types. Simulation results are given to show that the scheme presented in this report is both effective and reliable on a variety of pavement images.

I. Introduction

Condition evaluation is an important step in maintenance and repair planning. An automated pavement condition survey evaluates the condition of the existing roadway surface using modern imaging technologies. As costs of image collection and processing decrease while analysis techniques advance, various attempts have been made to develop automated condition evaluation system using imaging technology. Most of the existing systems record pavement surface images using a video camera or photographic camera mounted on a survey vehicle shown in Figure 1.

Figure 1: A van with pavement inspection device

In the late 1980s, the Japanese consortium Komatsu built an automated-pavement distress survey system [1], comprised of a survey vehicle and data processing system on board to simultaneously measure cracking, rutting, and longitudinal profiles. A maximum

resolution of 2048 x 2048 was obtained at the speed of 10 km/h. The Komatsu system worked only at night to control lighting conditions. The system represented the implementation of the most sophisticated hardware technology at that time. However, it did not output the types of cracking and only functioned during the night.

Earth Technology Corporation [2] created a research unit called Pavement Condition Evaluation Services (PCES). The automated system created by PCES was the first to use line-scan cameras at a 512-pixel resolution to collect pavement data. However, this effort was discontinued as the necessary technologies associated with the image capturing and processing was not advanced enough at the time.

In the early 1990s, Haas and Hendrickson [3] proposed a standard model to represent pavement surfaces that moved toward a unified and automated acquisition of key characteristics for improving data quality. Laser ranging was executed within a subset of the source image which designated an area of interest. The laser ranging would then complement or reject the vision data, so that dark areas which were not caused by pavement distress such as tire marks, oil spillings, and shadows, could be ruled out. Walker and Harris [4] also reported the development of a crack identification system using these laser ranging technologies. Velinsky and Kirschke [5] designed a machine vision system requiring laser ranging to overcome the shortcomings of an optical system.

Guralnick et al. [6] proposed a method using shadow moiré interferometer to measure coarse pavement distress, such as abnormal elevations and large sized potholes. The method allows detection of areas of the pavement that deviate from specified flatness criteria. The shadow moiré interferograms provide surface elevation variation measurements that cannot be obtained through ordinary videotaping. They can detect severe road elevation deformations caused by heavy loads, and potholes with undefined borders, which optical methods cannot detect.

Systems based on the Swedish PAVUE technology were used in the U.S. briefly in the

mid 1990s. The Swedish PAVUE data acquisition equipment includes four video cameras, a proprietary lighting system, and four S-VHS videocassette recorders [7]. The image collection subsystem is integrated into a Laser RST van. The off-line workstation is based on a set of custom designed processing boards in a cabinet to analyze continuous pavement data from the recorded video images. Surface images are stored on S-VHS tapes in analog format.

In the late 1990s, RoadWare Corporation was actively involved in using a new product, WiseCrax [8], for an automated survey of pavement surface. The data collection uses two analog cameras synchronized with a strobe illumination system, with each camera covering about half the width of a pavement lane. The image processing is done in an off-line office environment relying on the host CPUs to conduct image processing at a speed of two or three miles per hour with substantial operator assistance.

In the past few years, researchers at the University of Arkansas made substantial progress in automatically identifying and classifying pavement surface cracks at highway speed using a data collection system with one high-resolution digital camera and parallel processing of the data [9]. The current system can collect two-dimensional pavement surface images, identify and classify four types of cracks at a speed of over 60 MPH. The four types of cracks are longitudinal, transversal, alligator and block. The size of the cracks that can be identified and classified is about 2-millimeters. However, those existing automated systems commonly require special devices such as lights, laser, etc, which dramatically increases costs and limits the system to certain applications only. Thus, methods that are more economical, efficient, and practical for automatic pavement inspections are required.

This research aims to provide a reliable low cost automated pavement distress evaluation system capable of detecting cracks from complicated backgrounds while evaluating the severity of the damage. The proposed model consists of two major parts: pavement surface extraction and pavement distress detection as well as damage evaluation. In the

first part, a multistep color segmentation method is presented to separate the road surface from the background areas, such as houses, bushes, grass and trees. Following the road segmentation, a pavement distress detection algorithm based on probabilistic relaxation is described in the second part to further enhance the contrast between the cracks and the background. Based on the geometrical and topological parameters obtained from the crack structure, a neural network based pavement distress classifier is designed to assign the cracks into different types and severity groups. The overall procedure of the proposed system is illustrated in Figure 2.

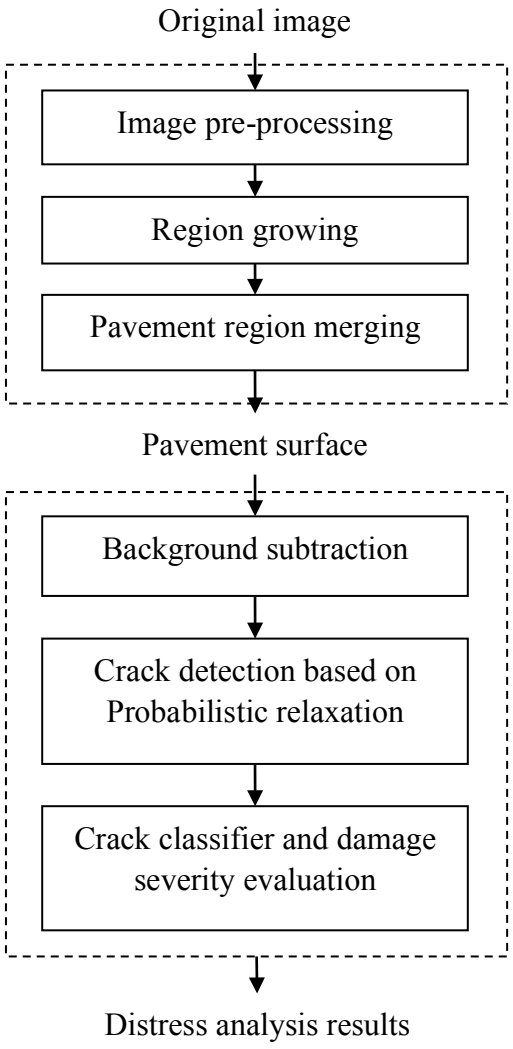


Figure 2: Overall diagram of the proposed iterative algorithm

II. Methodology

In this section, we present an automatic pavement surface extraction method that aims to separate the pavement surface from its complicated background. The input to the system is a series of forward view images continuously captured by a camera installed in front of the testing vehicle as it travels. A color segmentation approach was developed that consists of three successive steps. First, the input image is smoothed using a 3 x 3 median filter to remove the noise while preserving the edge sharpness. Then an initial region segmentation procedure based on region growing [10, 11] is applied to partition the image into small homogeneous regions. Finally, a novel region merging method based on a neural network is developed to differentiate the road surface from its background regions.

Figure 3: The input image to the system

2.1 Region growing

The RGB color space is commonly used in color image segmentation in which color is represented by triplet red, green and blue intensity values. Color distance is used as a measurement of color similarity where pixels/regions satisfying a certain degree of color homogeneity are grouped to form a cluster.

Let $I(i,j)$ be the color of the pixel located at image coordinates (i,j), in which $I(i,j)$ is defined in terms of color components(R_I, G_I, B_I). The color distance between pixel $I(i,j)$ and $I(i',j')$ is defined as,

$$\mathrm{D}\left(I(i,j), I(i',j')\right) = \sqrt{(R_I - R_{I'})^2 + (G_I - G_{I'})^2 + (B_I - B_{I'})^2} \qquad (1)$$

Assume(i_0, j_0) is the first pixel from the upper left corner to be examined for generating a new region. The region growing scheme operates from the pixel(i_0, j_0) in all directions to select the neighborhood pixels which are close enough to the pixel under study in terms of its color. If the neighborhood pixels satisfy the criterion,

$$\mathrm{D}\left(I(i,j), I(i',j')\right) < T_{grow} \qquad (2)$$

where T_{grow} is a threshold value for initial region growing, they are clusterized as a region and the $R_I\ G_I\ B_I$ values from equation (1) are replaced by the average RGB values of this region. This process is reiterated until no pixel meets the growing criterion. Then a new original pixel is chosen to start a new region growing process until every pixel in the image is assigned to a region.

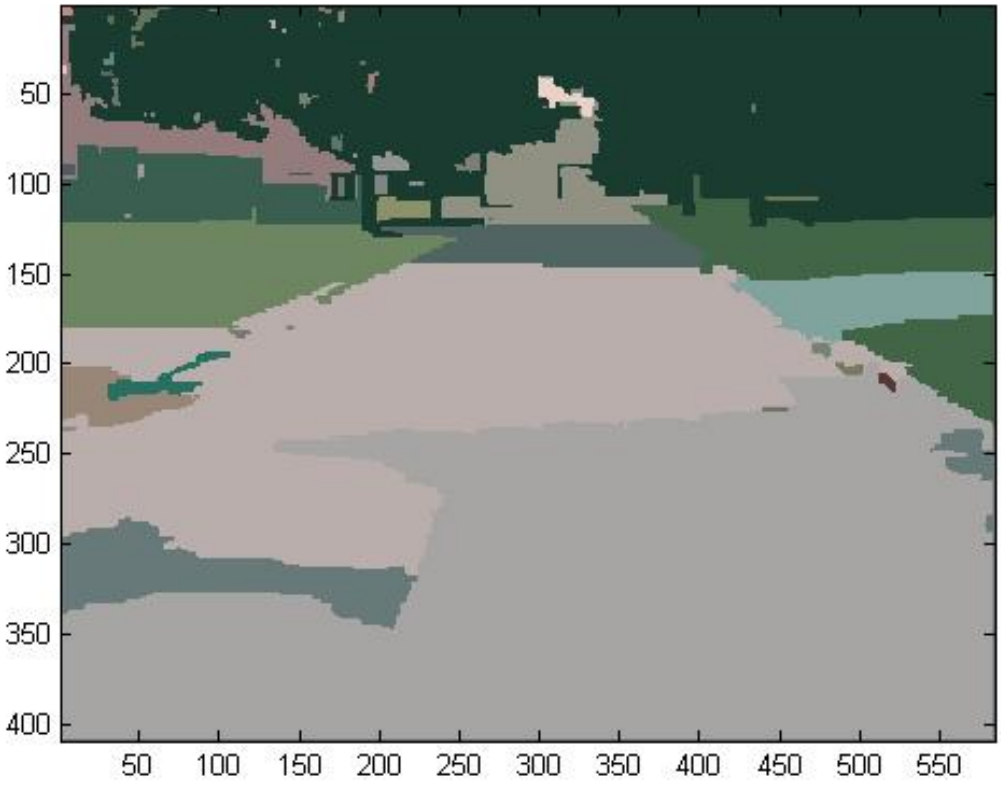


Figure 4: The initial segmentation by region growing

2.2 Pavement region merging

After initial segmentation, we have a number of undesired small regions in addition to the pavement region, as shown in Figure 4. To extract the pavement surface, we need to define a set of rules to guide the region merging process. In this section, a 3-layer feed forward artificial neural network is used to distinguish the pavement surface regions from the background regions. Figure 5 illustrates the basic framework of the neural network model. Each input node receives a feature descriptor that represents one aspect of the region. The output layer consists of 2 nodes that denote the pavement surface region and background region, respectively.

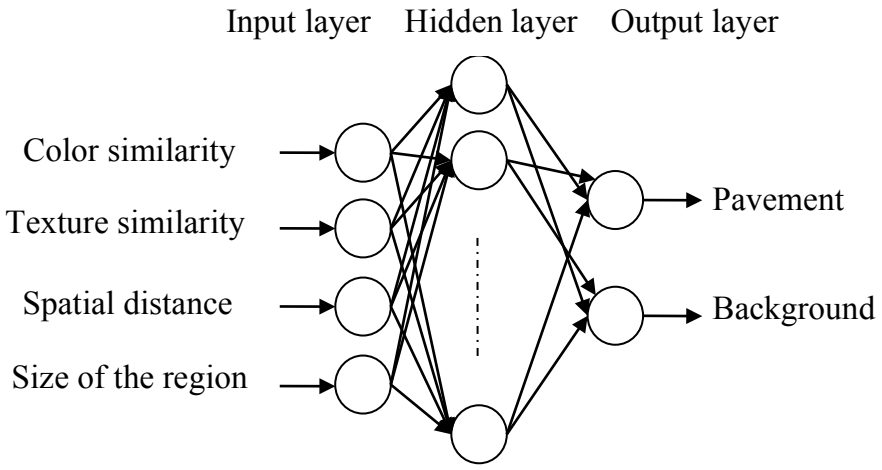


Figure 5: Neural network architecture for region merging

2.3 Feature selection

Appropriate selection of the input features from each region is very important to the success of the merging process. It is obvious that a region can be described in many respects, such as the color, texture, shape and size, etc. A detailed explanation of various features and their distinct characteristics are provided in the following.

2.3.1 Color feature

In the context of a pavement extraction application, color is considered to be more robust than other feature descriptors to represent the pavement surface. This is because the initially segmented small regions of the pavement surface often vary significantly in texture and shape due to the shadows and distresses, while their colors preserve a high degree of similarity. To represent the statistical characteristics of color distribution, color histograms in 3 different channels of the RGB color space are obtained as shown in Figure 6. It is evident that pixels that belong to the pavement surface take a dominant proportion of the whole image and mostly share a similar color. Therefore, an inference can be drawn that the pavement portion of the image corresponds to the most significant peak value in the color histograms.

To measure the similarity between a given region and the pavement surface, we only need to calculate the color distance from the average RGB value of the region to the peak value of each histogram using the following equation,

$$D_{color}(R, Pavement) = \sqrt{(\overline{R_r} - Peak_r)^2 + (\overline{R_g} - Peak_g)^2 + (\overline{R_b} - Peak_b)^2} \quad (3)$$

The measure of color similarity will serve as a main feature to the input of the neural network.

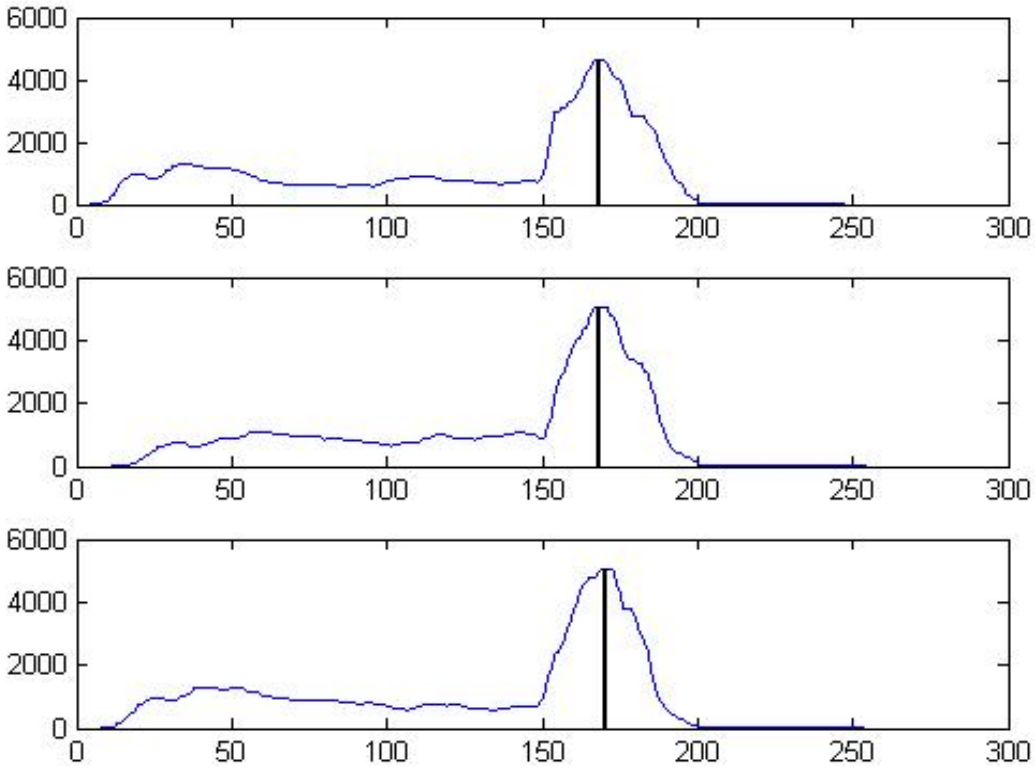


Figure 6: Histograms of the color image

2.3.2 Texture feature

The grayscale pavement image is used to derive the spatial texture features, which can then be combined with other features to produce the final segmentation. Like many of the existing algorithms for texture analysis and synthesis, our approach is based on multiscale frequency decomposition [12].

The key idea of the proposed method is that we are not interested in recognizing every little detail in the image. Rather, large areas with complex texture distribution such as grass and houses can be classified as "complex regions", while areas with relatively uniform texture distribution such as the sky and pavement surface can be classified as "smooth regions". In this way, only four major texture classes ("smooth," "horizontal," "vertical," and "complex") are defined.

In our texture analysis, we use the bi-orthogonal wavelet decomposition [13] which is a separable filter and considered to be computationally efficient. Since discarding the HH band does not result in significant loss of visual quality, only the HL and LH bands are used (note, H and L stand for the high-pass and low-pass bands in each of the horizontal and vertical orientations).

The most commonly used feature for texture analysis in a wavelet domain is the energy of the subband coefficients. Since the coefficients are quite sparse, it is necessary to perform some type of filtering operation to obtain a more uniform characterization of the texture. Here, we use a median energy filter. The advantage of the median filter is that it tends to remove the textures associated with transitions between regions. In such cases, the increase in wavelet coefficients due to the region boundary is concentrated along the edge and is not intensified by the median operator. The size of the window must be large enough to capture the local texture characteristics, but not too large to avoid border effects. We found that for the given image resolution and viewing distance, a 9 x 9 window size gives the best results.

Several clustering approaches are tested to obtain the texture segmentation. The simplest and most effective was to apply two-level K-means to each of the horizontal and vertical components separately. One of the cluster centers was always fixed at 0 (smooth texture) and the other was determined by the K-means algorithm. The added advantage of this approach is that we obtain four texture classes with obvious

interpretations: smooth texture, vertical texture, horizontal texture, and complex texture. Figure 7 shows the results, with smooth texture represented by white, vertical by light gray, horizontal by dark gray, and complex by black.

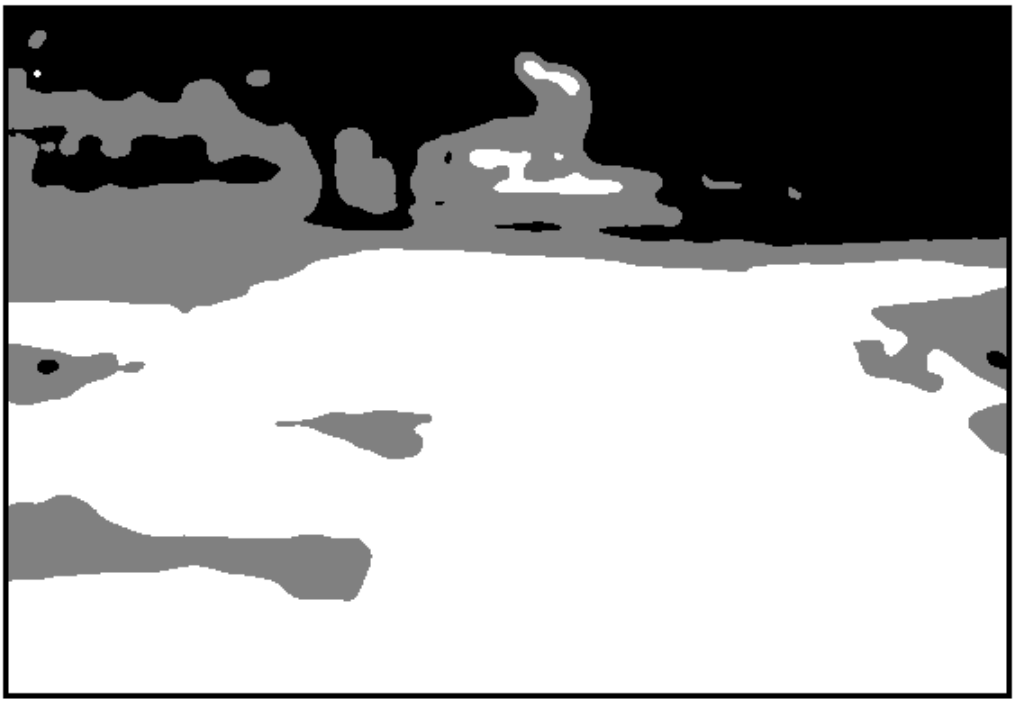


Figure 7: Image after texture segmentation

2.3.3 Other features

Other than the color and texture, size and location information of the regions are selected to be the other two descriptors as input into the neural network. Based on our experience, size and location can provide valuable information for the following reasons:

- Most of the small regions are more likely to be merged with the larger ones.
- Regions near the image boundary are more likely to be components of the background while regions located at the lower or middle part of the image have a higher probability of being parts of the pavement.

To keep track of the size of each region, we simply count the number of pixels in the region, and the number is recorded as an input to the neural network. As far as the location information is concerned, the centroid point of each region (highlighted by '*' in

Figure 8) is calculated to represent the location of the region by averaging the index of all pixels within the region. The last feature descriptor to the neural network is obtained by calculating the spatial distance between the centroid of each region and a pre-defined pavement surface marker (highlighted by 'o' in Figure 8). The distance between a region R and the pavement marker is the following Euclidian distance,

$$D_{spatial}(R, Pavement) = \sqrt{(\overline{X_R} - X_{marker})^2 + (\overline{Y_R} - Y_{marker})^2} \quad (4)$$

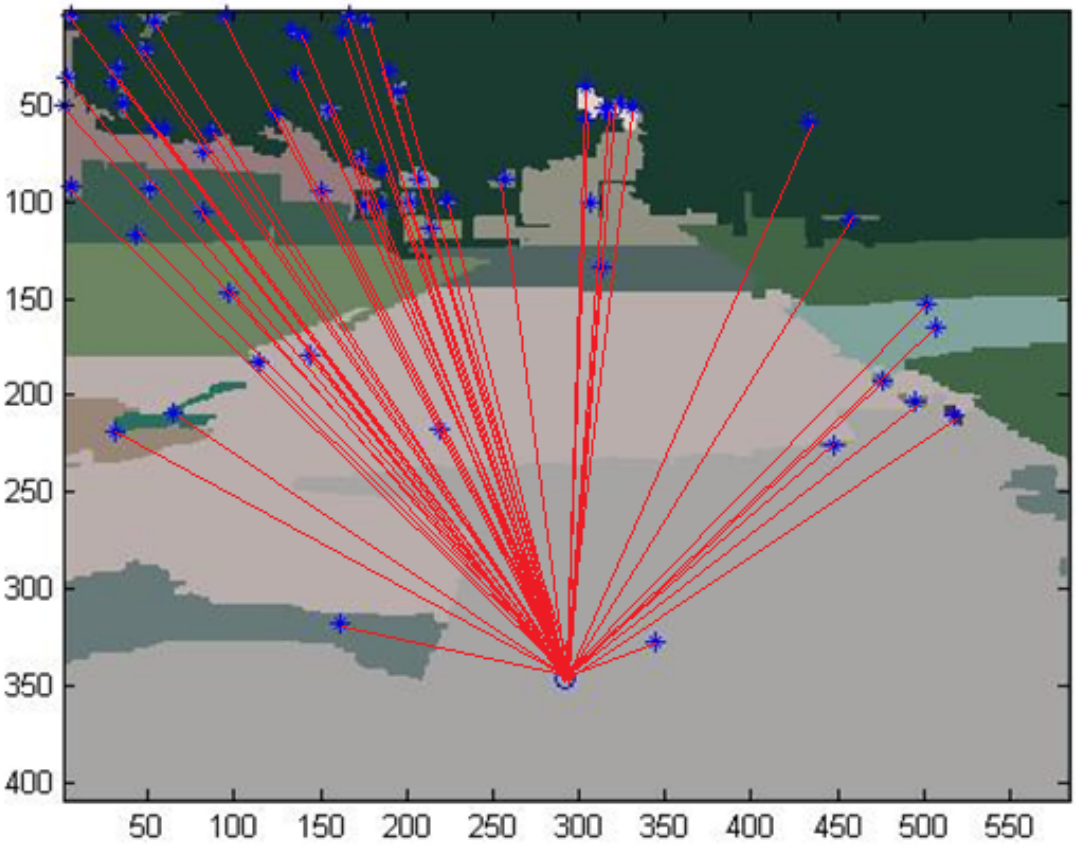


Figure 8: Spatial distance of each region from the road marker

These features will then be available to the neural network shown in Figure 5. After training the network with a set of artificial data, the final segmentation result is shown in Figure 9.

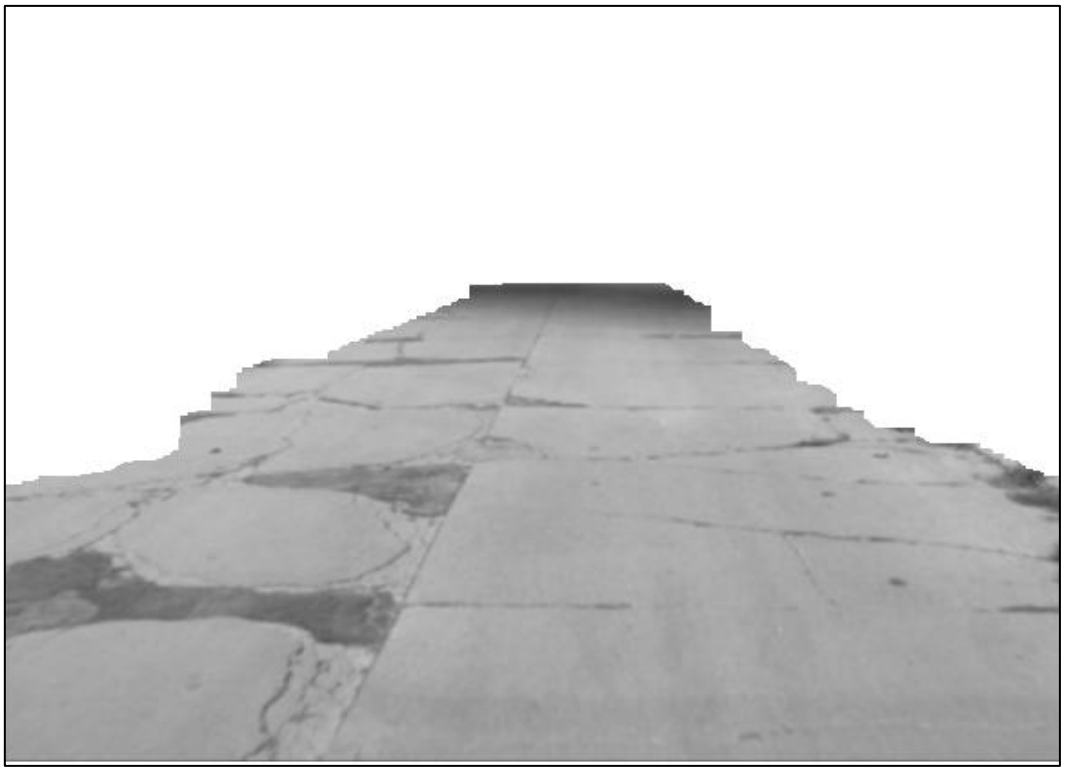


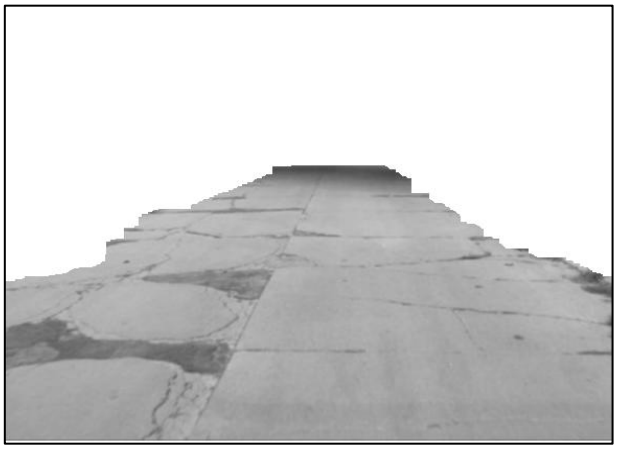
Figure 9: Pavement surface extracted by the proposed method

III. CRACK DETECTION AND EVALUATION

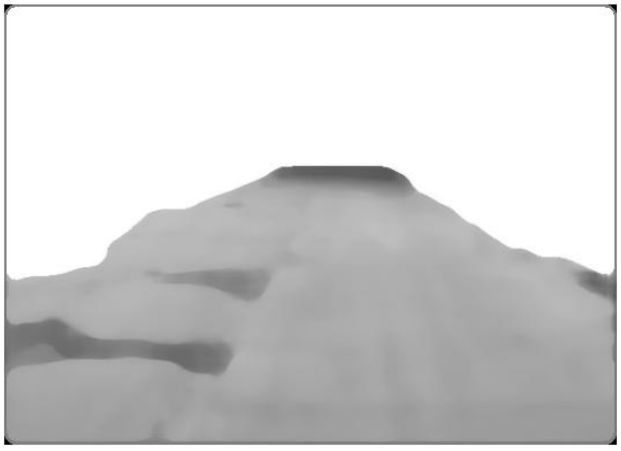
3.1 Non-uniform background removals

Crack detection on the pavement surface using image processing techniques is considered to be a difficult problem due to the presence of noise in the image of the pavement surfaces. To prevent the influence of the shadings and to provide a more uniform background, a background subtraction method is applied ensuring the same background lighting condition for all parts of the pavement images. To extract the background of the image, a relatively large size median filter is applied to the image to eliminate the detailed information on the pavement surface. In this way, the cracks with a thin structure are removed from the image, leaving the raw pavement background only. Then the extracted background is subtracted from the original image to obtain a subtracted image, of which the non-uniform illumination effect is fully removed. Figure 10 illustrates the process of removing light variations using a median filter. At first, we use a 15 x 15

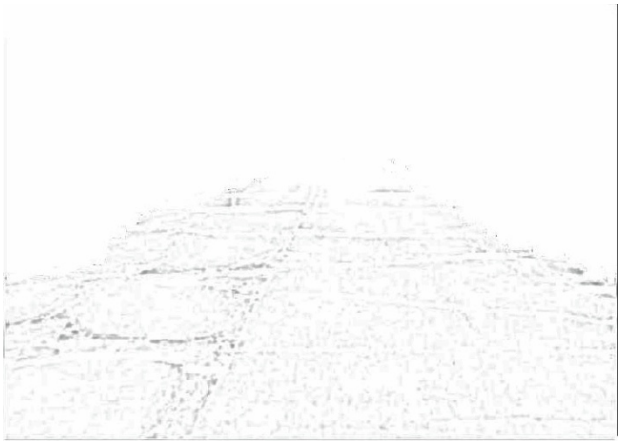
window size median filter to smooth the original image while blurring the image [14]. Then the original image is subtracted from the smoothed image in order to remove the non-uniform background illumination.



(a) Original image



(b) Smoothed image with the median filter



(c) Background subtracted image

Figure 10: Non-uniform illumination removal by median filtering

3.2 Crack detection

In this section, a probabilistic relaxation [14] technique is used to label the crack pixel from the noisy pavement images. For each pixel in the image, the initial probability $\mathrm{P_i}(\lambda_c)$ corresponding to a crack is assigned according to the intensity value of the output of pre-processing by,

$$\mathrm{P_i}(\lambda_c) = \frac{\log(I_{min}+1)}{\log(I_i+1)} \tag{5}$$

where I_i means the intensity value of pixel i, I_{min} is the minimum value in the image, and λ_c denotes the label for a crack pixel.

The probability of a pixel to be a crack is then updated by considering the label probabilities of its neighborhood. However, in crack detection, if the estimation of an isotropic neighborhood is used as a non-ambiguity, line structures like cracks in the image may be removed unexpectedly due to its structural property. Therefore, we divide the neighborhood into eight sub-regions according to the direction and then estimate the non-ambiguity in each sub-region, in order to remain line structures as cracks. This means that eight estimations of each pixel are calculated along different angles and structures passing through the pixel. The maximum value of eight estimations is used to update the probability. We can update probability $P_i'(\lambda_c)$ by,

$$P_i'(\lambda_c) = \mathrm{Min}\left(\frac{\mathrm{P_i}(\lambda_c)\max(\mathrm{Q_d}(\lambda_c))}{\frac{1}{N}\sum_{\mathrm{i}\in\mathrm{image}}(\mathrm{P_i}(\lambda_c))}, 1\right) \tag{6}$$

$$\mathrm{Q_d}(\lambda_c) = \frac{1}{N(\mathrm{R_d})}\sum_{\mathrm{i}\in R_d}\mathrm{P_i}(\lambda_c) \tag{7}$$

where $\mathrm{R_d}$ is one of the eight sub-regions illustrated in Figure 11. This probability updating process is repeated until it reaches a convergence. In the next step, an Otsu's

thresholding method based on the obtained probability instead of the original intensity value is applied for the purposes of the segmentation. Since the pixels located within the crack structures are of a higher probability to be a crack than the isolated noise pixels, this thresholding technique is able to remove undesirable disconnected pieces of noise components while filling in missing parts of cracks.

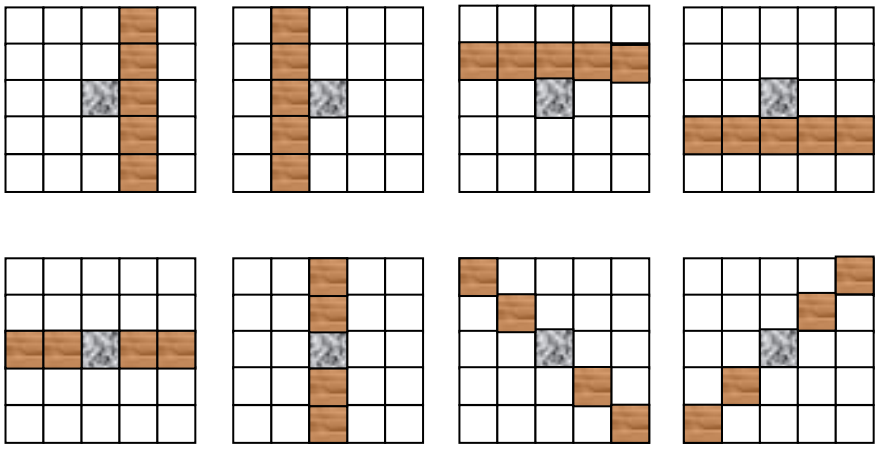


Figure 11: Eight sub-regions considered for updating the crack probability

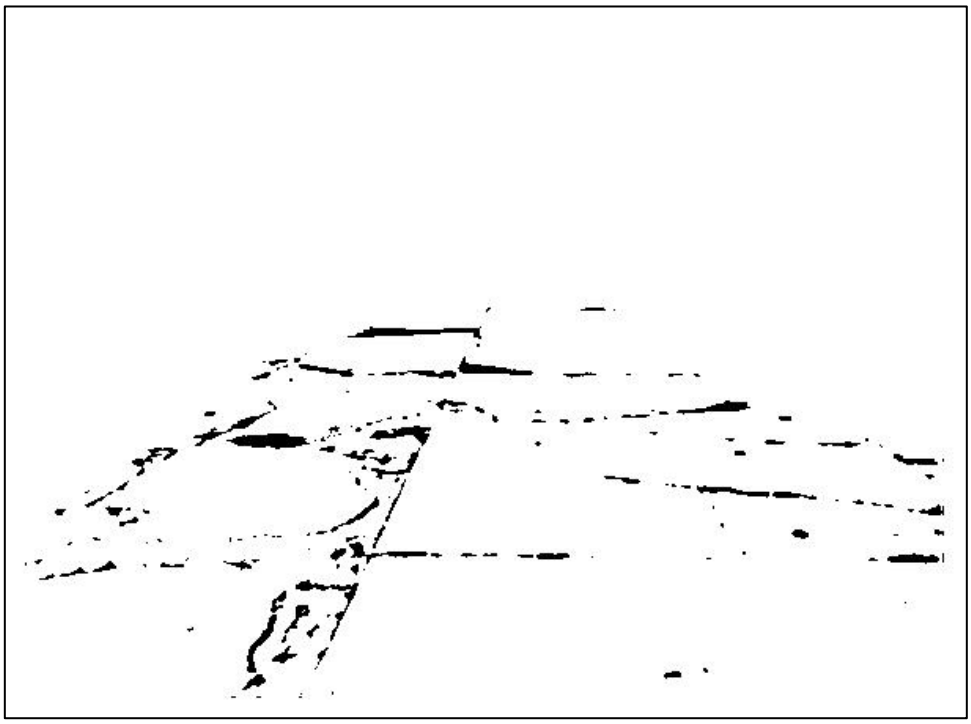


Figure 12: Crack detection results

3.3 Crack classification and damage severity evaluation

To reduce the computational complexity of the distress classification process, the binary image is partitioned into "tiles" [15]. Whether a tile is a crack tile or not is determined based on the percentage of crack pixels in a tile. If the percentage of crack pixels in a tile is greater than the predefined threshold, the tile is considered to be a crack tile. From the image of crack tiles, two kinds of histograms are computed: a vertical histogram and a horizontal histogram to represent the distribution of crack tiles in each column and row, respectively, as shown in the following equations,

$$H[i] = \sum_{j=1}^{M} crack\_tiles\,[i,j], i = 1,2 \ldots N \tag{8}$$

$$V[i] = \sum_{i=1}^{N} crack\_tiles\,[i,j], i = 1,2 \ldots M \tag{9}$$

where V and H represent the vertical and horizontal histograms, M and N denote the number of rows and columns, respectively.

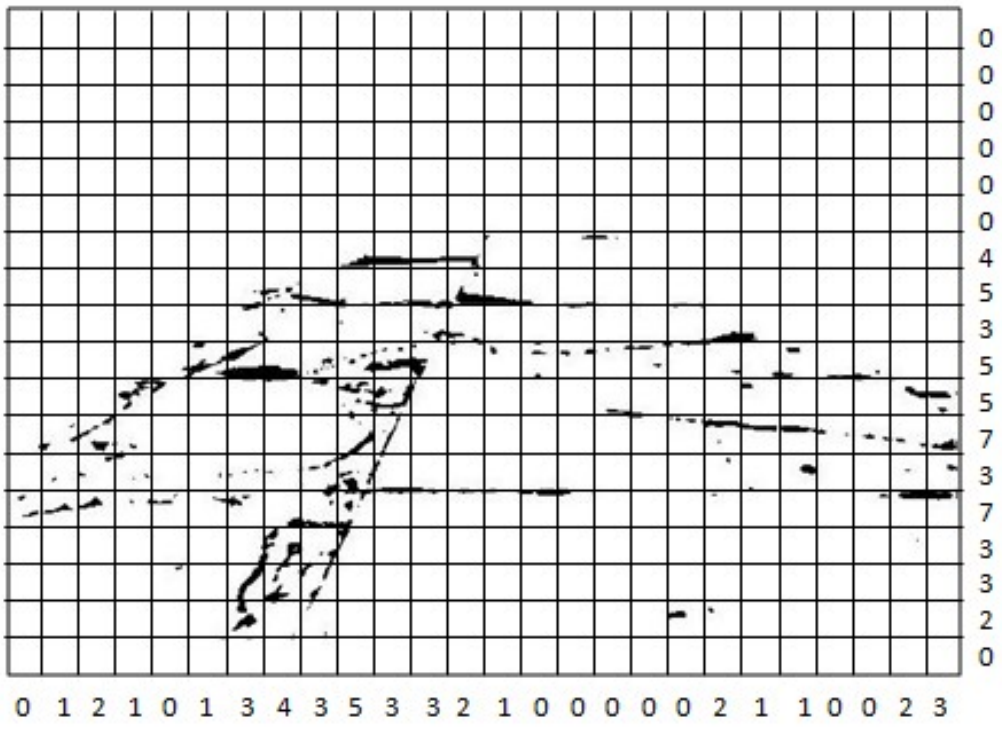


Figure 13: A tile image with vertical and horizontal histograms

Histograms show a clear pattern of a crack. If the crack is developed in a longitudinal direction, there would be a peak in the vertical histogram. On the other hand, if the crack is developed in the transversal direction, there would be a peak in the horizontal histogram. If the crack is of a block type, the peaks could be found in both vertical and horizontal directions. To represent the above observations, two accumulations of the differences between adjacent histogram values are calculated using Equations (10) and (11),

$$A_H = \sum_{i=1}^{N} H\,[i+1] - H[i], i = 1,2 \ldots N \tag{10}$$

$$A_V = \sum_{j=1}^{M} H\,[j+1] - H[j], j = 1,2 \ldots M \tag{11}$$

The values of these two accumulators will serve as inputs to a specially designed artificial neural network to begin the process of crack classification. The network is a 3 layer feed forward neural network. The connection weights of the neural network are obtained from the training process by using artificial tile images. Note, Num represents the number of crack tiles of an image.

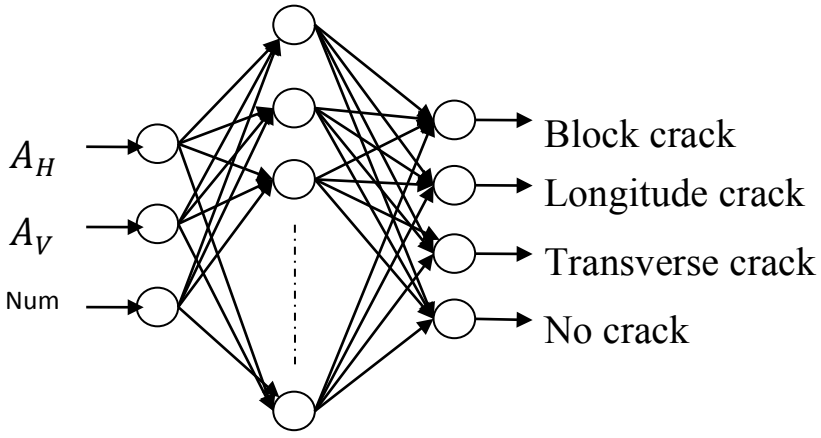


Figure 14: Architecture of the neural network for distress classification

IV. EXPERIMENTAL RESULTS

This section evaluates the performance of the proposed method by considering more than 100 actual pavement images. The images are of size 600 x 400 pixels captured from a camera generally installed in front of a testing vehicle on highway road surfaces. Figures 15 and 16 illustrate pavement surface segmentation along with crack detection results of the proposed system.

Tables 1 and 2 show the detailed training parameters for the region merging and distress classification neural networks, respectively. Thirty actual pavement images are utilized to train the network under manual supervision. With a learning coefficient equal to 0.01, the neural networks could achieve a high training accuracy of 95% and higher.

Table 1: Training process for the region merging neural network

Learning coefficient	Nodes for each layer	Epochs	Accuracy
0.01	3-6-2	500	97%

Table 2: Training process for the distress classification neural network

Learning coefficient	Nodes for each layer	Epochs	Accuracy
0.01	3-30-4	1500	95%

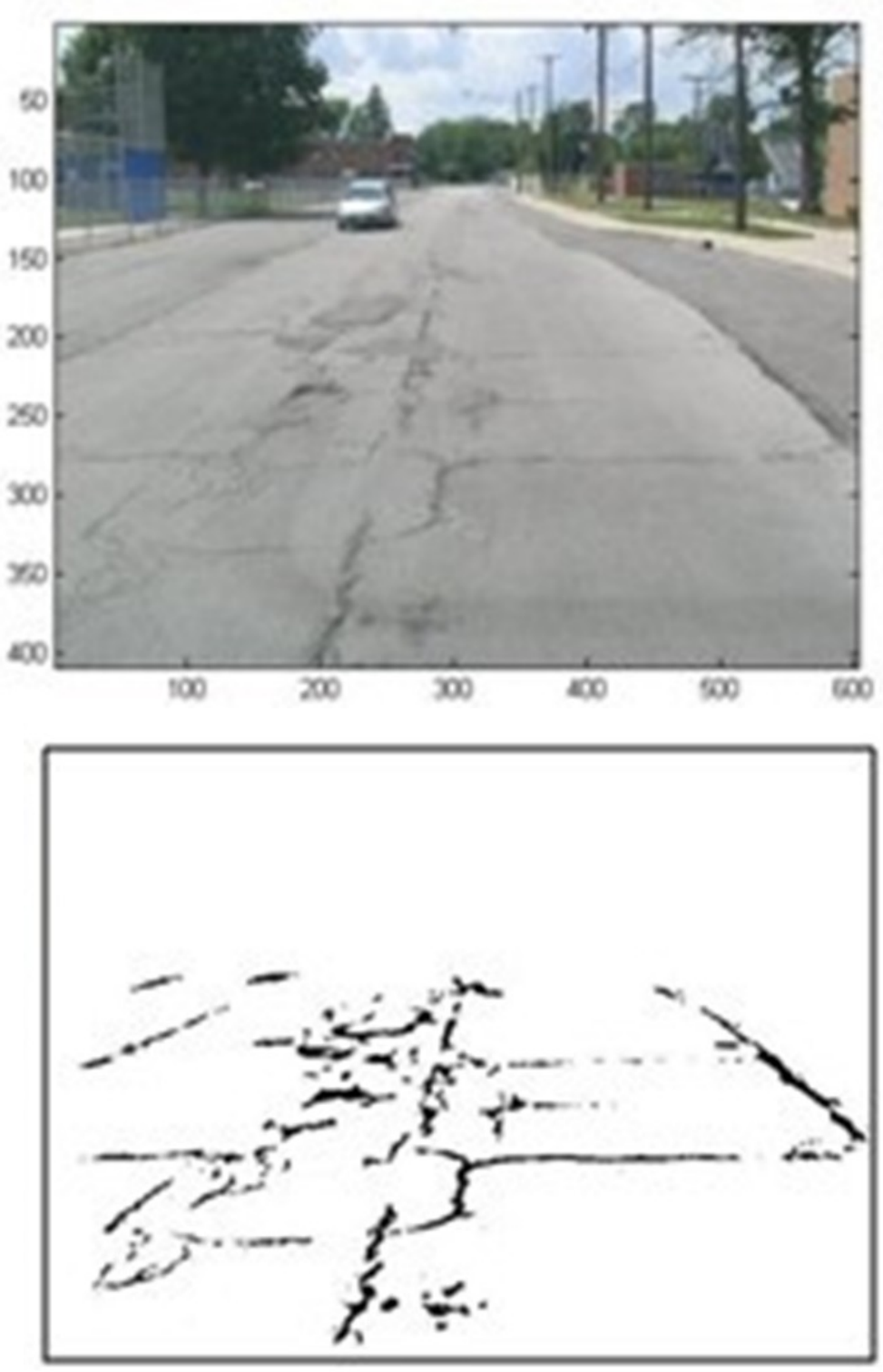


Figure 15: Pavement distress detection result No.1

Figure 16: Pavement distress detection result No.2

There are two main criteria used to assess the performance of the system: one is processing time and the other is classification accuracy. As shown in Table 3, the algorithm is able to correctly classify all cases with typical parameters like transversal

cracks and longitudinal cracks. For pavement with a combination of different types of cracks, the cracks are all grouped into the combination category. In this category, no specific crack type can be determined by the input parameters, but the damage severity level can still be determined by counting the number of crack tiles in the image.

Table 3: Performance of the proposed system

Crack type	Average processing time	Accuracy	Severity level	Num. of test images
No crack	1.9 s	100%	Low	16
Transversal	2.2 s	100%	Low	15
Longitudinal	2.1 s	100%	Low	22
Block crack (Combination of different types)	2.5 s	95%	High	23
			Median	19
			Low	5

V. CONCLUSION

This project further extends the scope of our earlier investigation in providing a more generalized solution in a less restricted environment. In the current study, we consider a road scene containing grass, trees, buildings and other objects in addition to the pavement itself. The presence of various objects and features in an acquired image poses a considerable challenge in detecting the desired structural failure patterns such as cracks and other surface anomalies. In other words, the presence of various features brings in complications for a recognition system to classify the desired patterns. Hence, this necessitates the application of a more sophisticated segmentation process involving texture features to extract and identify the pavement region from the rest of the aerial image prior to the extraction of the cracks from the pavement region.

In the segmentation process, we use both the color and texture information to extract the pavement regions. A probabilistic labeling scheme is utilized to extract the crack features from the pavement images. In addition, neural networks are designed to process the pavement images and are also used as a decision tool to provide a classification of various types of cracks. The proposed approach, in addition to producing competitive accuracy, has the positive attributes of design simplicity and computational efficiency. Our experimental results show that the method gives good crack detection results on different types of pavement images.

ACKNOWLEDGEMENT

This research study was funded by the University of Toledo's University Transportation Center (UT-UTC) and by the Michigan-Ohio University Transportation Center (MI-OH UTC), with both funding originated from the Research, Innovation and Technology Administration (RITA), US Department of Transportation. The researchers greatly appreciate the support.

REFERENCES

[1] T. Fukuhara, K. Terada, M. Nagao, A. Kasahara and S. Ichihashi. “Automatic Pavement Distress Survey System”. Journal of Transportation Engineering, Vol. 116, No. 3, pp. 280-286, May/June 1990.

[2] http://www.aecom.com/

[3] C. Haas and C. Hendrickson. “Computer Based Model of Pavement Surfaces”. Transportation Research Record 1260, pp. 91-98, 1990.

[4] R. S. Walker and R. L. Harris. “Noncontact Pavement Crack Detection System”. Transportation Research Record 1311, pp. 149-157, 1991.

[5] S. A. Velinsky and K. R. Kirschke. “Design Considerations for Automated Pavement Crack Sealing Machinery”. Proceedings of the Second International Conference on Applications of Advanced Technologies in Transportation Engineering, pp. 77-80, 18-21 August 1991.

[6] S. A. Guralnick, E. S. Suen and C. Smith. “Automating Inspection of Highway Pavement Surface”. Journal of Transportation Engineering, Vol. 119, pp. 1-12, 1993.

[7] K. C. P. Wang. "Design and Implementation of Automated Systems for Pavement Surface Distress Survey", ASCE Journal of Infrastructure Systems, Vol.6, No I, pp. 24-32, March, 2000.

[8] RoadWare, Online WiseCrax Product. Introduction, http://www.roadware.com/

[9] K. C. P. Wang and W. Gong. “Automated Real-Time Pavement Crack Detection and Classification”. Final Report for Highway IDEA Project 111, TRB-NCHRP-111, May, 2007.

[10] A. Tremeau and N. Borel. “A Region Growing and Merging Algorithm to Color Segmentation”. Patten Recognition, Vol.30, No.7, pp 1191-1203, 1997.

[11] R. Nock and F. Nielsen. “Statistical Region Merging”. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 26, No. 11, November 2004.

[12] J. Chen, T. Pappas, A. Mojsilovic and B. Rogowitz. “Adaptive Image Segmentation based on Color and Texture”, Proceedings on International Conference on Image Processing (ICIP), Vol. 3, pp.777-780, 2002.

[13] A. Cohen, I. Daubechies and J. C. Feauveau, “Bi-Orthogonal Bases of Compactly Supported Wavelets,” Communication Pure Appl. Math., Vol. 45, pp. 485–560, 1992.

[14] Y. Fujita and Y. Hamamoto. “A Robust Method for Automatically Detecting Cracks on Noisy Concrete Surfaces”. Lecture Notes in Computer Science, 2009, Vol.5579/2009, pp 76-85, 2009.

[15] B. J. Lee. “Position-invariant Neural Network for Digital Pavement Crack Analysis”. Computer-aided Civil and Infrastructure Engineering, Vol.19, pp 105-118, 2004