DETERMINATION OF PAVEMENT SURFACE CRACKS FROM VIDEO-IMAGES USING AN IMAGE SCALE-SPACE APPROACH

C. ARMENAKIS

Department of Earth and Space Science and Engineering, School of Engineering, Geomatics Engineering, York University, Toronto, Ontario, Canada
armenc@yorku.ca

L. NINGYUAN

Pavements and Foundations Section, Ontario Ministry of Transportation, Toronto, Ontario, Canada
Li.Ningyuan@ontario.ca

ABSTRACT

Assessment of pavement surface distress is an important component of the pavement management process. Pavement surface distresses characterize failures and distortions of the pavement surface structure. A large number of highway surface images have been collected through the application of a video image system. We present an automated approach that detects pavement surface cracks from a forward viewing video camera system. Initially the oblique imagery is transformed to a rectified one which supports quantitative measurements of the crack patterns. For the detection and extraction of pavement surface distress elements, we propose to use a scale-space image approach, where the image scale is defined based on the level of detail of the image structure to be detected. Finally, the detection of crack patterns is performed considering the image as a 3D intensity surface where the bright and dark lines are considered as ridges and valleys. The approach is based on determining the local directions of the image curvature along the curvilinear lines, and determining where along these directional profiles the second derivative of the line profile reaches its maximum absolute value.

1. INTRODUCTION

Assessment of pavement surface distress is an important component of the pavement management process. Pavement surface distresses characterize failures and distortions of the pavement surface structure. Common types of pavement surface distress are: potholes, various types of cracks (longitudinal, transverse, alligator), and rutting. Surface distress data are assessed for severity and quantified for their spatial extent. A systematic process needs to be in place for collecting pavement data and information, analysing them, and implementing cost-effective pavement maintenance and rehabilitation programs.

Conventional patrolling, visual inspection, and evaluation of pavement surface conditions face significant challenges, especially on high-speed and high-volume highways where road access and safety and operation are of concern for the transportation authorities. These field-based assessments are time-consuming, expensive, tedious, and labour intensive. In addition, they result in inconsistencies for the distress elements across evaluations due to subjectivity, high degrees of variability, and lack of quantitative data. The Ministry of Transportation of Ontario (MTO) has a low-cost and efficient high-speed image system that uses video techniques that has been installed on the Ministry's automatic pavement data collection equipment, and a large number of highway surface images have been collected through the application of the video image system. An automated pavement surface distress inspection (APSDI) system may reduce

disturbances to public traffic and road hazards to human inspectors during the field surveys. The digital imaging technology required to develop APSDI systems for pavement distress surveys has been used for some time now.

It has been reported in NCHRP [1], that almost all North American highway agencies use some form of automated means to both collect and process pavement condition data. Most of the data are collected in a single pass by an integrated system capable of capturing forward, lateral, and downward images as well as both longitudinal and transverse profiles. Independently of image capture systems, the images are usually "stamped" with the date, time, and a location-reference. After the pavement images have been collected, their analysis is performed either manually, by visual inspection, or with the use of specialized image analysis software. Manual processing of the collected images is obviously laborious and time consuming. Therefore the automated detection, classification and quantification of pavement surface distress features are highly desirable. Considering the complexity of pavement conditions, textures and illumination, developing and implementing a totally automated image based system is a challenging task. Over the last few years various methods for the evaluation of pavement surface conditions from image data have been developed with varying accuracies and acceptance levels.

For 2D approaches, many highway agencies use semi-automated survey systems. The post-processing requires substantial human interaction. Image processing based algorithms have been developed to perform recognition task for the crack type, as well as estimations of the severity and the calculation of its area. The processing algorithms for 2D images include histogram analysis (i.e., examinations of bimodal histograms), thresholding, spatial filters, gradient operators, edge detectors (i.e., Sobel, Canny), mathematical morphology, texture analysis, colour descriptors, object-specific analysis (use of mean and variance) and marker-controlled watershed segmentation, where the image digital numbers are treated as "elevations" and spectral segmentations. Following the crack detection, thinning, crack-tracing and the determination of the X, Y co-ordinates of the cracking on the image need to be performed. The first step in the image processing process is to distinguish any pavement surface distress from other non-distress noises. The classical method of thresholding is employed at the initial stage of image processing. It is used to segment the image into black-white images. The continuity, for example, of cracks in the process of removing noises, is maintained using a morphological closing to pre-process the image and fill small and thin holes in objects, connect nearby objects without changing their area significantly, and smoothing the boundaries of objects.

Median filtering, thresholding and statistical analysis of video images for the evaluation of pavement distress were reported by Kim [2]. Subirats et al., [3] have presented a continuous wavelet transform automated approach for crack detection on pavement surface images. The result is a binary image which indicates the presence or not of cracks on the pavement surface image. Oliveira and Correia [4] proposed an unsupervised system for automatic crack detection and classification using survey images acquired at high driving speeds. A multi-scale approach using Markov Random Fields for crack detection has been presented by Chambon et al., [5]. A Gaussian function is used to enhance pavement cracks following their detection using 2D matched filter. Li et al., [6] use edge detection for pavement distress assessments, while Qingquan and Xianglong, [7] used image segmentation to extract cracks from images. Interactive image segmentation for efficient road mapping through a continuous training of the detection algorithm is given by Barinova et al., [8].

While the pavement surface and the distress element are 3D in nature, most of the data acquisition and processing systems are 2D. To better detect, model and measure the 3D distress elements of the pavement, surface stereo-vision and laser scanning systems are emerging and are expected to become the main tools for monitoring pavement surface conditions (NCHRP, [1]; Ahmed and Haas, [1]).

In Section 2, the image geometry and the process of generating rectified images is presented. The scale-space image concept is discussed in Section 3 and the pavement surface crack detection approach is given in Section 4. Section 5 describes the steps of the method for processing an image and the preliminary results obtained, while Section 6 gives a summary of the approach and future steps.

2. IMAGE CAPTURE AND RECTIFICATION

Sequences of video images of road scenes have been captured from a forward looking video camera mounted inside a car (Figure 1). The camera optical axis Z is assumed to be horizontal and parallel to the road axis. The video camera is considered uncalibrated, that is the calibrated focal length is unknown or the nominal value sis used, the principal point location is set at the image centre and the lens distortions are not taken into account (calibration software is under development). Regarding the exterior orientation of the camera, the position is considered unknown, (Z can be determined as the height from the planar road surface), while the rotations ϕ and κ of the images with respect to a plane normal to the road axis are considered negligible.

The images show strong perspective geometry of the road lanes with respect to the road surface plane, mainly due to the angular tilt of approximately 90 degrees between the road and image planes, and they are characterised by large-scale variation between the low and the upper portions of the image (Figure 2). Considering that we need to extract quantitative

Figure 1 – Video camera system

measurements of the detected pavement surface distress elements detected in the video image as well as to further link to road location information (i.e., using GPS) it is therefore necessary to convert the perspective image geometry to an orthographic projection. Assuming that the elevation differences across the pavement are small enough, a very good approximation of the orthoimage will be the generation of a rectified image at the desirable scale. As the rotation ω around the image X axis pointing laterally to the optical / road axes is about 90 degrees, this is equivalent to eliminating the effect of the 90 degrees (approximately) tilt angle between the camera optical axis and the normal axis to the road plane surface. The result will be an image with parallel instead of converging road lanes and road edge lines (Figure 3).
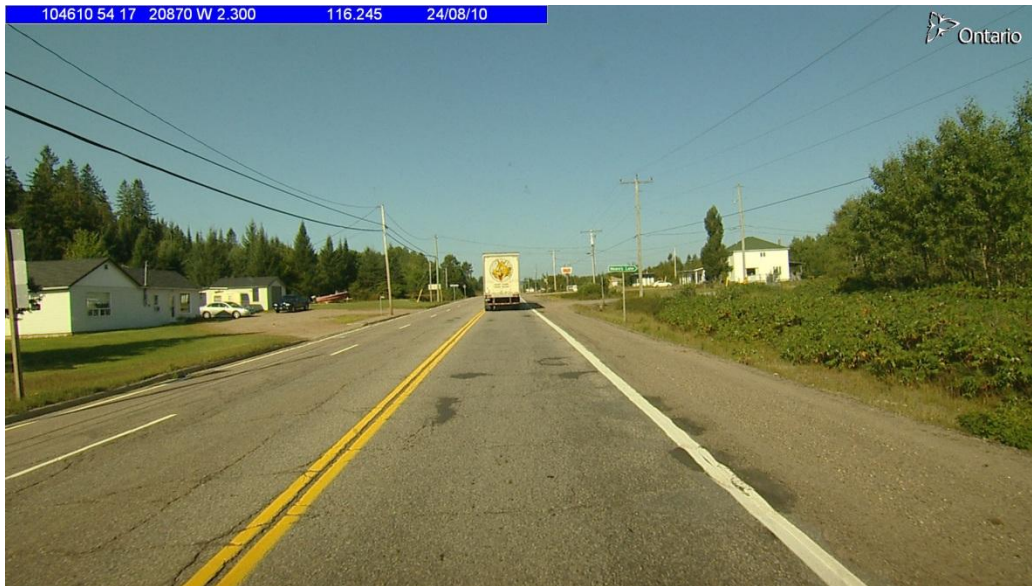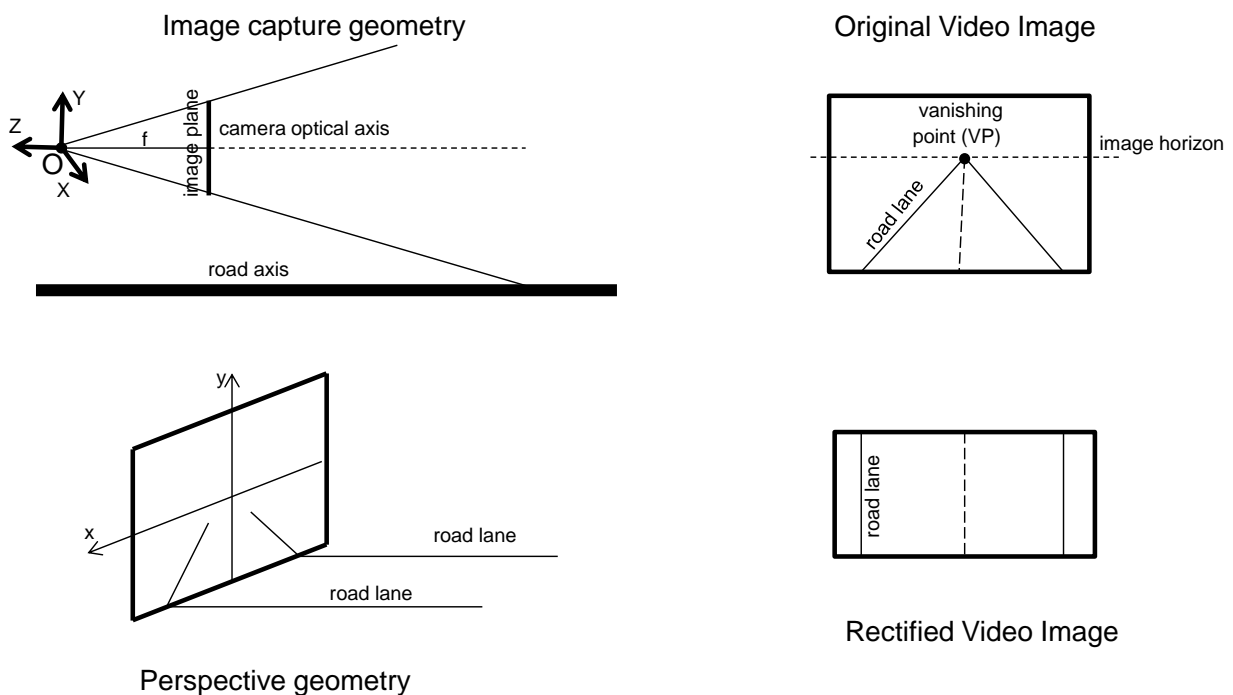

Figure 2 – Example of an image scene

Figure 3 – Image geometry and rectified imagery

The rotation ω was estimated using the vanishing point geometry, where the direction's parallel lines in the object space are mapped as convergent lines in the image. The point of convergence is the vanishing point VP (Figure 3). The image coordinate of the VP have been estimated as follows. Initially a combination of Canny and Sobel edge detectors were used to detect the lane markings. As several other feature edges were detected as well, a Hough transformation was applied to select lines based on the parametric notion of line distance from image origin and line orientation (ρ, θ). The image coordinates VP_x and VP_y of the VP were estimated as the intersection of the convergent lane marking lines using a least squares adjustment. The VP vector is then defined by VP_x, VP_y and VP_z where VP_z=-f.

Since rotations ϕ and κ are assumed to be negligible, the normalized vanishing vectors n(VP) and n(VP) are mutually orthogonal and equal to unit vectors. Therefore the rotation matrix R to be used for rotating the road axis to make it perpendicular to the image axis for the generation of the rectified imagery is estimated as:

$$R = \begin{bmatrix} n(VP_x) & 0 & 0 \\ n(VP_y) & 1 & 0 \\ n(VP_z) & 0 & 1 \end{bmatrix} \qquad (1)$$

The coordinates of the rectified image X_R, Y_R, Z_R with respect to the origin of the camera perspective centre O are computed as:

$$\begin{bmatrix} X_R \\ Y_R \\ Z_R \end{bmatrix} = sR^T \begin{bmatrix} x \\ y \\ -f \end{bmatrix} \qquad (2)$$

where x, y are image coordinates, f is the image focal length, and s is the scale. If Z_R=-f then the scale of the rectified image is the same as the original image.

The generated rectified images are used as the input images for the crack detection process and for the generation of image mosaics (Figure 4).
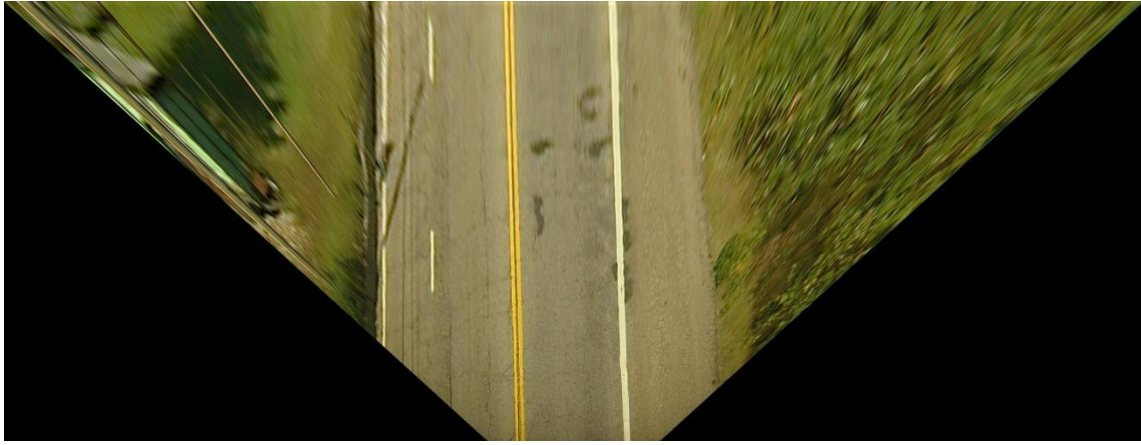
Figure 4 – Rectified image

3. SCALE-SPACE IMAGE REPRESENTATION

Due to the high resolution of the imagery, pixels-based classification algorithms, such as supervised and unsupervised classification algorithms and edge detectors may result in the extraction of noisy pavement surface distress elements using image segmentations. To avoid the extraction of noisy pavement surface distress elements, we propose to use a scale-space image approach (Lindeberg, [10]). This image scale-space approach consists of detecting multi-level image structures based on various image resolutions from fine-to-coarse.

Spatial scale (resolution) of an image depends on the levels of details we can distinguish. Key features in an image can be detected / located based on the scale of the image. It is analogous to being able to detect the trees versus detecting just the forest and it is similar to the abstraction process applied for map generalization, where the number of cartographic features shown depends on the various map scales. Thus, various features and, in general, various image details can be better detected at different image scales. Therefore, image structures such as pavement cracks can be detected at a certain image scale. Therefore, an image can be represented as a set of multi-scale images, where each image can support the detection of image structures based on their level of detail and their radiometric relationships to the neighbouring image details. Image pyramids are an example of multi-scale image representations.

This image scale-space approach consists of multi-level images of various resolutions from fine-to-coarse, usually based on a hierarchical image pyramid structure. It emulates the human cognitive perception of initially detecting a generalized form of patterns followed by an analysis of the details of the initially detected patterns. Thus, pavement surface distress patterns at coarser image scales are generalized forms of corresponding features at finer image scales.

The scale-space approach from high image resolutions to low image resolutions is a multi-level detection approach with a consecutive application of image segmentation algorithms only at image scales where distress patterns are salient enough and can be better detected. The scale-space approach involves successive segmentation of an image at several scales, where a specific scale corresponds to a window of perception.

An image I(x,y) is represented at a different resolution (scale) by applying a Gaussian spatial filter. The Gaussian filter is applied to smooth images and generates images at different scales from fine to coarser image scales.

$$g(x, y; \sigma) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{x^2 + y^2}{2\sigma^2} \right] \tag{3}$$

where g(x,y) is the 2D Gaussian filter at location x,y and σ is the standard deviation of the Gaussian function

The scale of the generated image depends on the σ factor of the Gaussian filter. The larger the value of σ, the smoother the new image will be. The scaled image is produced by a kernel convolution of the original image I(x,y) by the Gaussian filter g,

$$I_\sigma(x, y) = I(x, y) * g(x, y, \sigma) \tag{4}$$

A proper scale can be selected by, for example, examining the maximal responses of detecting an image feature at different scales. This of course relates to how prominent the feature is (e.g., how large) so it can be extracted from the background image values.

4. PAVEMENT CRACK DETECTION

To extract this linear crack feature, a scale-space representation of the rectified image is used. The 2D Gaussian spatial filter is applied to select the appropriate image scale determined by the width of the pavement crack and the initial image pixel size. This smoothing also serves to reduce the noise of the pavement texture. The image function I(x,y) is regarded as a 3D surface representation of the image intensity, where the very bright and the very dark areas are considered to be the ridges and valleys of the image function. Generally, the image features of the road cracks are elongated and thin features with a certain pixel width. Usually, they are dark lines against a brighter background of the pavement. To detect the pavement cracks we use Steger's detector of 2D curvilinear lines (Steger, [11]). This approach has been also used for the extraction of roads from high resolution aerial images (Auclair Fortier et al., [12]).

Initially the curvature direction of the linear structure is calculated. The maximum curvature direction is perpendicular to the crack line and is determined by the eigenvectors and the corresponding eigenvalues of the Hessian matrix H(x,y) of the scale-space smoothed image $I_\sigma(x,y)$:

$$H_\sigma(x, y) = \begin{bmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{bmatrix} \tag{4}$$

where I_{xx}, I_{xy} and I_{yy} are the second order derivatives of $I_\sigma(x,y)$. For their computations, convolutions of the image with derivatives of a Gaussian function with parameter σ are used.

The maximum curvature direction corresponds to the direction of the eigenvector that relates to the eigenvalue with the maximum absolute value λ_{max}. This corresponds to the second directional derivative of a bright line profile on a dark background that has a negative minimum and its first directional derivative has a zero crossing. Thus, the likelihood of detecting a line in the Hessian image can be defined using the λ_{max} points that belong to a ridge. For dark lines the line plausibility is λ_{max} and for bright lines the line plausibility is $-\lambda_{max}$. The scalar product of the gradient with the eigenvector of the maximum eigenvalue is computed to assess how close the selected pixel is to the centre of the ridge. Points in the centre of the ridge will have this scalar product very close to zero (Eberly et al., [13]).


5. EXPERIMENTAL RESULTS

The methodology presented above was implemented in MATLAB coding environment. Video image data were provided by the Pavements and Foundation Section, Ministry of Transportation of Ontario (MTO) for testing the proposed approach. The input image for pavement crack detection process is a rectified colour image. As the approach is based on the intensity and gradient values of the image, the removal of road lane markings is essential prior to the detection of the pavement surface cracks. This is to prevent possible erroneous crack detections due to the strong contrast between the pavement surface and the colour lane markings.

The K-means iterative unsupervised clustering method has been used for the image segmentation (Bishop, [14]) using the R, G, B intensity values. The number of K-classes has been set to K=3 to represent the three major segments classes on the rectified road images. The lane markings are typically brighter than the road or road shoulder surface. Two classes are set for two different shadings of the road surface and/or road side objects. The final representative class is for the lane markings. Segmented lane marking pixels shown in Figure 5a were then masked out of the RGB image (Fig. 5b).

Following the image segmentation and masking, the Hessian of the scale-space image was derived (Fig. 5c). As the boundaries of the masked road lane markings were apparent a masked Hessian image was generated for the detection of road cracks based on the computation of the maximum eigenvalues of the masked Hessian image (Fig. 5d). These maximum eigenvalues indicate dark line plausibility over brighter image background. A binary image of the strongest eigenvalues was produced, (Fig. 5e) followed by the application of an areal filter to remove noise (Fig. 5f). The detected cracks were then superimposed on the initial rectified imagery, (Fig. 5g) as well as on an image mosaic consisting of two consecutive rectified images (Fig. 5h).
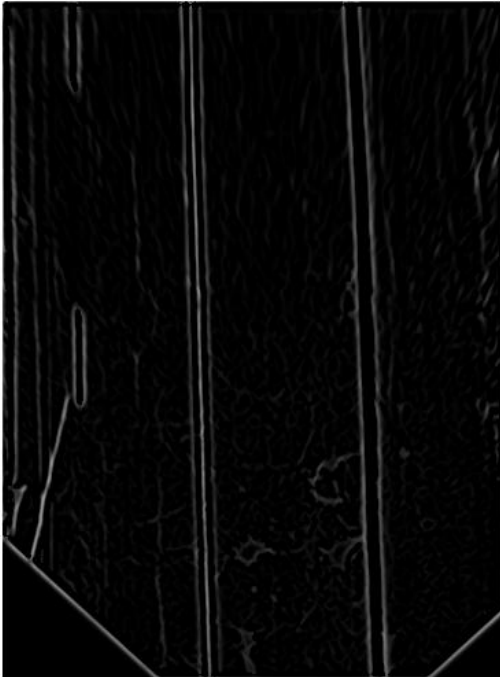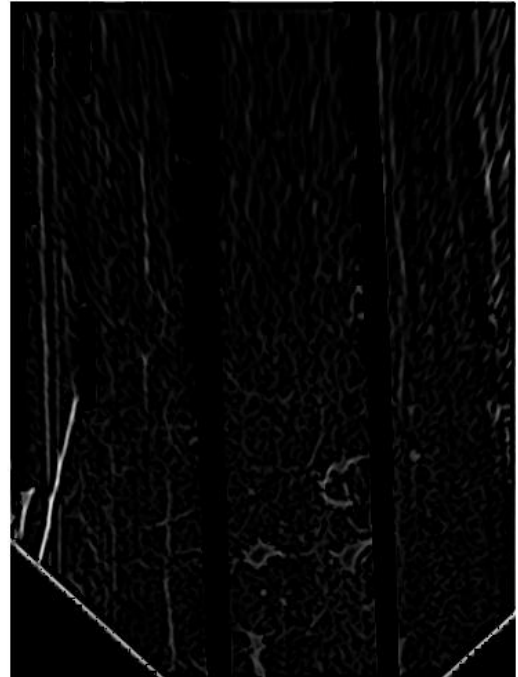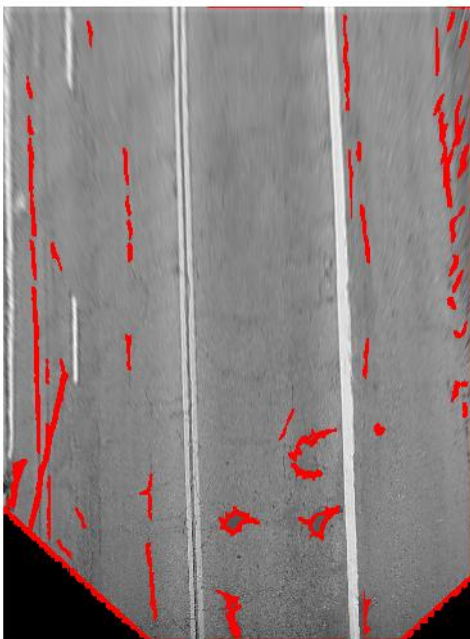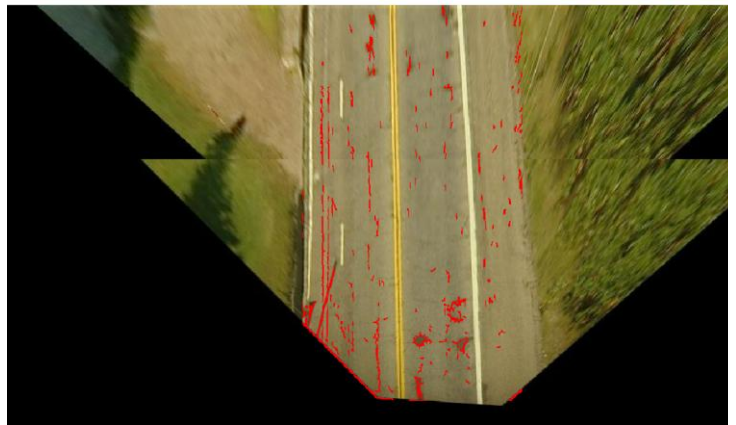
Segmentation Maps
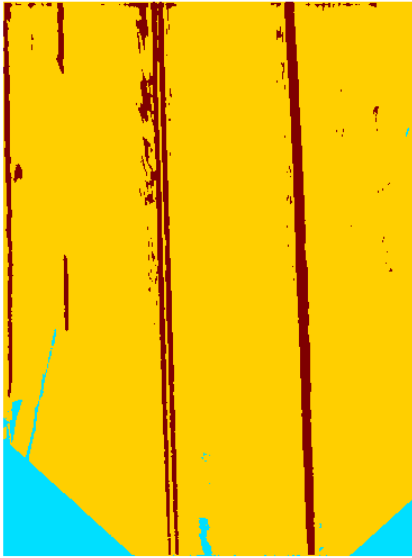
5a
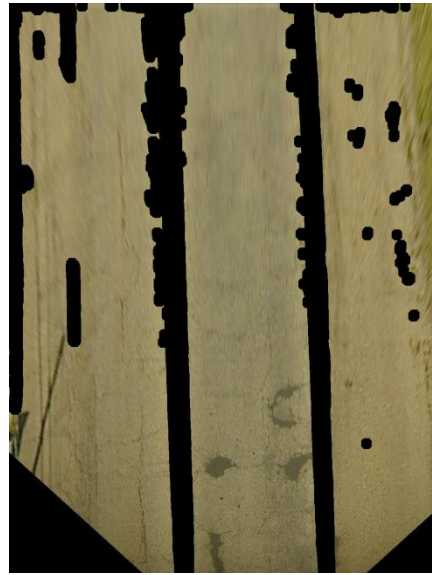
5b

5c

5d

5e   5f

5g   5h

Figure 5 – Intermediate and final results of the pavement crack detection process

6. CONCLUDING REMARKS

An approach for the automatic detection of pavement surface crack using images from a forward looking video camera has been presented. This road detection task is based on three main steps. First, a rectified image is produced from the forward looking images,

thus allowing for the planimetric quantification of the detected cracks. Second, a scale-space image is proposed that defines an appropriate image resolution for the detection of crack structures based on the level of detail of the road cracks. Third, the extraction of the crack patterns from the "smoothed" imagery is performed by adopting Steger's detector of curvilinear image structures, which precisely detects the crack line position using the eigenvectors and eigenvalues of the Hessian of the scale-space image. Future work will address the validation of the approach, the detection of areal surface distress elements, the classification of extracted cracks and the use of this method in production environments, where large numbers of highway surface images have been collected using video imaging systems.

7. ACKNOWLEDGEMENTS

REFERENCES

1. NCHRP (2004). Automated pavement distress collection techniques-A synthesis of highway practice. NCHRP Synthesis 334, Transportation Research Board.
2. Kim J. (2000). Statistical analysis of video images for evaluating pavement distress. KSCE Journal of Civil Engineering, 4(4):257-264.
3. Subirats P., J. Dumoulin, V. Legeay, D. Barba (2006). Automation of pavement surface crack detection using the continuous wavelet transform. IEEE ICIP 2006 p 3037-3040.
4. Oliveira H., P. L. Correia (2009). Automatic road crack segmentation using entropy and image dynamic thresholding, 17th European Signal Processing Conference (EUSIPCO 2009) pp.622-626.
5. Chambon, S., Gourraud, C., Moliard, J.-M., and Nicolle, P. (2010). Road crack extraction with adapted filtering and Markov model-based segmentation. In VISAPP.
6. Li, L., P. Chan, A. Rao, R.L. Lytton (1991). Flexible pavement distress evaluation using image analysis. Proceedings of the Second International Conference on Applications of Advanced Technologies in Transportation Engineering, pp. 473-477.
7. Qingquan and Xianglong (2008). Novel approach to pavement image segmentation based on neighoring difference histogram method. Congress on Image and Signal Processing, CISP'08, Vol. 2, pp. 792-796.
8. Barinova O., R. Shapovalov, S. Sudakov, A. Velizhev, A. Konushin (2009). Efficient road mapping via interactive image segmentation. CMRT09, ISPRS Com III, WG III/5.
9. Ahmed M.F.M, C.T. Haas (2009) The potential of low cost close range photogrammetry towards unified automatic pavement distress surveying, CD-ROM proceedings, Transportation Research Board (TRB) 2010 Annual Meeting
10. Lindeberg, T. (1993). Scale-space theory in computer vision. Kluwer Academic Publishers.
11. Steger, C. (1998). An unbiased detector of curvilinear structures. IEEE Transactions on Pattern Analysis and Machine Intelligence. 20(2):113–125.
12. Auclair Fortier, M.-F., D. Ziou, C. Armenakis, S. Wang (2001). Automated correction and updating of road databases from high-resolution imagery. Canadian Journal of Remote Sensing. 27(1):76-89.
13. Eberly. D., R. Gardner, B. Morse, S. Pizer, C. Scharlach (1994). Ridges for image analysis. Journal of Mathematical Imaging and Vision. 4:353-373.
14. Bishop, C.M. (2006). Pattern Recognition and Machine Learning. Springer.