Fusing complementary images for pavement cracking measurements

Ming Yao[1,2], Zuyun Zhao[1], Xun Yao[2] and Bugao Xu[1,2]

[1] School of Human Ecology, University of Texas at Austin, Austin, TX 78712, USA
[2] Center for Transportation Research, University of Texas at Austin, Austin, TX 78701, USA

E-mail: bxu@utexas.edu

CrossMark

Abstract

Cracking is a major pavement distress that jeopardizes road serviceability and traffic safety. Automated pavement distress survey (APDS) systems have been developed using digital imaging technology to replace human surveys for more timely and accurate inspections. Most APDS systems require special lighting devices to illuminate pavements and prevent shadows of roadside objects that distort cracks in the image. Most artificial lighting devices are laser based, and are either hazardous to unprotected people or require dedicated power supplies on the vehicle. This study was aimed to develop a new imaging system that can scan pavement surface at highway speed and determine the level of severity of pavement cracking without using any artificial lighting. The new system consists of dual line-scan cameras that are installed side by side to scan the same pavement area as the vehicle moves. Cameras are controlled with different exposure settings so that both sunlit and shadowed areas can be visible in two separate images. The paired images contain complementary details useful for reconstructing an image in which the shadows are eliminated. This paper intends to present (1) the design of the dual line-scan camera system, (2) a new calibration method for line-scan cameras to rectify and register paired images, (3) a customized image-fusion algorithm that merges the multi-exposure images into one shadow-free image for crack detection, and (4) the results of the field tests on a selected road over a long period.

Keywords: line-scan camera, camera calibration, image fusion, pavement cracking

(Some figures may appear in colour only in the online journal)

1. Introduction

Cracking is the most common pavement distress and it undermines a pavement's integrity and long-term performance. From the early 1970s, researchers have been dedicated to developing automatic pavement distress survey (APDS) systems to replace human visual rating so that the traffic disturbance and the risk to human inspectors could be reduced. Automatic systems can also provide more objective and prompt results for rehabilitation management [1–4]. Quality image acquisition methods and accurate algorithms for pavement distress detection are the two key factors to build up a successful APDS system. However, due to the complexity of field circumstances such as various lighting conditions and image capture at highway speed, implementing the APDS systems presents many challenges [5–7].

The image acquisition devices for an APDS system fall into two categories: 2D area-scan cameras and 1D line-scan cameras. In contrast to the 2D area-scan camera, the line-scan camera captures only one line at a time. Its image is generated by accumulating scanned lines over time. When used in pavement imaging, a line-scan camera can be mounted close to the image-capturing vehicle. It helps to keep its viewing direction perpendicular to the pavement surface, minimizing distortion caused by perspective projection at a tilt angle. Compared to the 2D camera, the line-scan camera usually has larger imaging sensor. This significantly improves the signal-to-noise ratio at each sensor element, and is a required feature for high-speed imaging during which the exposure time is usually of the order of tens of microseconds. Recent advancements in CCD and CMOS sensor technology also have dramatically
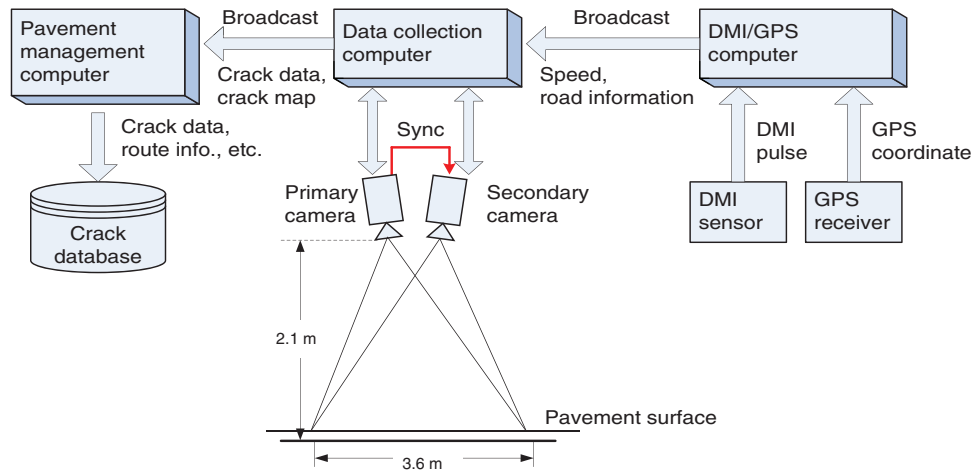
Figure 1. Schematic design of the complementary imaging system.

increased the cameras' resolution, sensitivity and frame/line rate, making it particularly suitable for pavement inspection.

To work with line-scan cameras, light-supporting devices were used by many researchers to project a transverse beam to illuminate the camera's view area so that the effects from shadow could be reduced. Halogen or fluorescent lamps, light-emitting diode (LED) arrays [8] and laser line-projectors [9, 10] are the primary light sources commonly used in APDS systems. The halogen or fluorescent lamps alleviate shadows by generating white light. However they require a high power supply installed on the vehicle, which is unsafe and inconvenient. They have been less used since a power-efficient camera filter was invented later for blocking sunlight. An LED-light-bar based APDS system [8] designed by our group was used by the Texas Department of Transportation (TxDOT). To block sunlight, a band-pass filter that matches the wavelength of the red LED was used. However, light arrays cannot project a focused beam at a distance of more than 35 cm, and high-power LEDs lose efficiency quickly if overheated, making the survey operation more difficult during summer. Laser line projection has become a dominant approach for line-scan cameras in APDS systems because of its high efficiency and compact size. However, the beam of the high power infrared lasers (class III or class IV) is potentially hazardous to unprotected human eyes. The non-uniform power distribution transversely across the lane also causes longitudinal streaks in images. Furthermore, the narrow beam (<5 mm) of the laser line may lose alignment with the camera's view line when vehicles undergo severe vibrations, yielding horizontal dark ripples in the image. These streaks and ripples are difficult to remove and are the main sources of false detection of cracks.

A safe, reliable and cost-effective APDS system for cracking inspection is still greatly needed for maintaining long-term performance of the highway network. This paper proposes a novel pavement imaging approach using dual line-scan cameras to build a new APDS system. The proposed method should allow pavement inspection in any precipitation-free climate during daytime, and with regular highway speeds on both asphalt and concrete pavements to determine the level of severity of cracking. In order to avoid problems related to safety, alignment and stability, the new system does not use any artificial lighting. The setup includes two cameras working under natural lighting to capture paired images. The paired images can complement each other to form a clear and shadow-free pavement image for automatic crack detection.

2. Complementary imaging system

2.1. System design

The basic idea of this complementary imaging method is to use two line-scan cameras to capture the same pavement surface simultaneously with different exposure settings, and to generate two distinct images that can complement each other. One camera is set in high-exposure mode to ensure pavement textures in shadows will be clear, while the other is set in low-exposure mode to acquire clear textures in sunlit areas. The cameras' exposure settings are adjusted dynamically according to the lighting situation, pavement condition and vehicle speed, so that the visible regions of the two images are kept at appropriate brightness and contrast levels. The ability to dynamically adjust image brightness level gives the system the ability to adapt to the transition from working in a shadowed region to a shadow-free region, and vice versa.

Our proposed system is composed of two line-scan cameras, which are placed side by side 2.1 m above the ground to cover a 3.6 m wide lane (figure 1). The cameras are Dalsa® Spyder III with 36 kHz scanning rate. The camera resolution is 2048 pixels/line, giving a spatial resolution of 1.78 mm/pixel at this height. A Nikon F-mount lens with 20 mm fixed focal length is attached to each camera. The lens aperture is set to F5.6 to achieve a sufficient depth of view. The cameras are synchronized by the same triggering pulse. The system needs a distance measurement instrument (DMI) and a Global Positioning System (GPS) receiver to generate traveling distance, driving speed and GPS coordinates. The above information is broadcast through a DMI/GPS computer to a data collection computer in order to create a tag for each image and make crack data traceable. The driving speed is needed to calculate the instant camera line rate to ensure a constant interval between two consecutive scan lines.
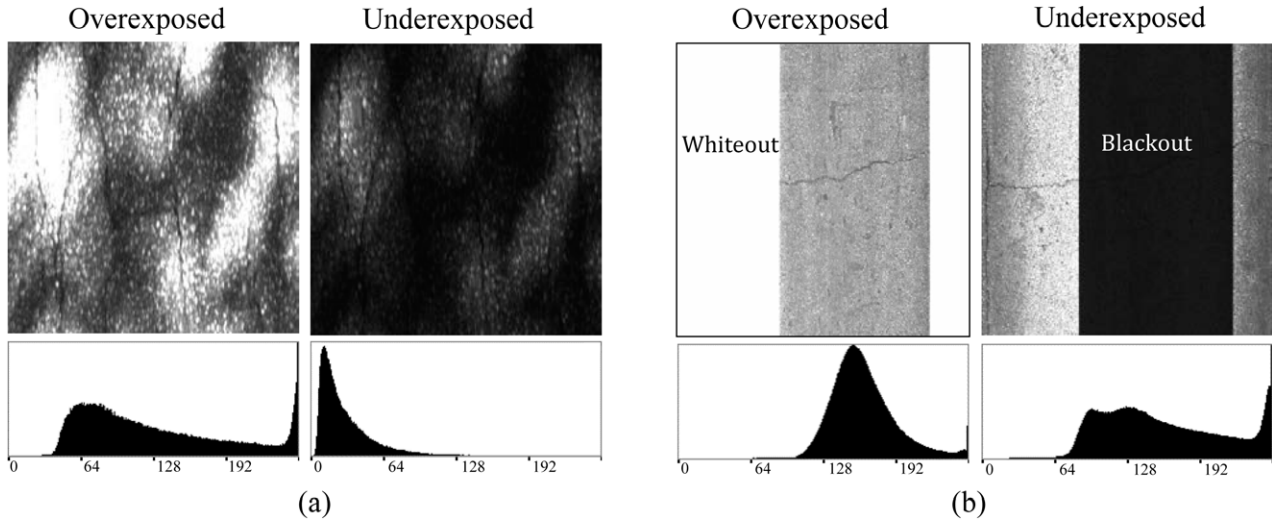
Figure 2. Complementary images and their grayscale histograms. Pixels with extreme values (0 and 255) are excluded.

The two cameras are connected to the data collection computer through a GigE interface. They are wired in series via a 15-pin GPIO (general-purpose input/output) connector for synchronization. The capturing mode of both cameras is configured as 'External Sync (ExSync)', in which line capturing is triggered by pulses received through the GPIO pins. An onboard programmable pulse generator on the primary camera receives a pulse generation command from the data collection computer, and updates the pulse frequency according to the vehicle's instant speed. Pulses are then sent to each camera's GPIO pin to trigger the capture.

2.2. Image brightness control

Image brightness is controlled dynamically by adjusting the exposure time and analog gain. Our image capturing program evaluates each incoming image, and sends an update to both cameras to adjust the exposure and gain settings. The primary camera is purposefully configured in overexposed mode, and is designated to enhance details of shadowed regions by letting sunlit regions become whited out in the image. On the other hand, the secondary camera is configured to be underexposed to make pavement details in the sunlit regions visible, letting the shadows be blacked out. For both cameras, the analog gain is always kept at the lowest possible level to avoid noise being amplified.

The exposure adjustment for the next frame is based on the evaluation of the histogram of the current frame. The histogram of a well-balanced image frame should cover a wide range of grayscales from black to white. In our system, the visible regions in pictures are considered as the regions of interest (ROI). As a result, pixels of whiteout and blackout are excluded from computing the histograms. To make a decision on whether the grayscale of the ROI is well balanced, we need to evaluate how far the overall brightness of the ROI deviates from the central gray level, i.e. 128. To do this, the accumulated histogram of the pixels in the ROI is calculated.

Let $H = \{H_i \mid i = 0, \ldots, 255\}$ be the histogram. H_i is the percentage of pixels at gray level i. The accumulated histogram

is $A = \{A_i \mid i = 0, \ldots, 255\}$, where $A_0 = H_0$, and $A_i = A_{i-1} + H_i$ $(i > 0)$. A_i is the percentage of the total pixels whose grayscale is lower than i. Assume α is a desired cutoff of A, and C_α is the grayscale where the accumulated histogram value is equal to α. We can assess the overall brightness G_o of the image by using

$$G_o = 0.5 \times (C_{0.5} + 0.5 \times (C_{0.3} + C_{0.7})). \tag{1}$$

If the histogram is similar to a normal distribution, G_o will be close to $C_{0.5}$. Otherwise, averaging $C_{0.3}$ and $C_{0.7}$ along with $C_{0.5}$ can lead to a more reasonable G_o. After the exposure is adjusted by using the difference between G_o and the central grayscale, the overall brightness of the incoming image will be expected to be brought toward the central level, preventing the ROI from being whited out or blacked out. Figure 2 shows two pairs of typical complementary images captured by the two cameras in our field test. Figure 2(*a*) contains tree shadows, and figure 2(*b*) has a vertical shadow caused by the vehicle driving against sunlight. Pavement cracks in ROIs are visible in both of the overexposed and underexposed images. When the image-capturing vehicle enters an area without any shadow, the overexposed camera outputs a complete whiteout image, while the underexposed camera outputs an image where the ROI covers the whole area. The exposure time and gain are reduced in the overexposed camera, but stay unchanged in the underexposed camera. Eventually, both cameras reach the same exposure time and gain settings, and output images with ROIs covering the whole image area.

2.3. Image acquisition

The data collection computer runs on multiple threads to handle image streams simultaneously. The image acquisition thread creates a frame buffer pool for its associated camera, and repeatedly checks the camera status to see whether one frame is completed. When it is done, the acquisition thread takes an available free buffer from the pool, and creates a queued job to save the frame. When the number of queued frames in the pool reaches a predefined value, the acquisition
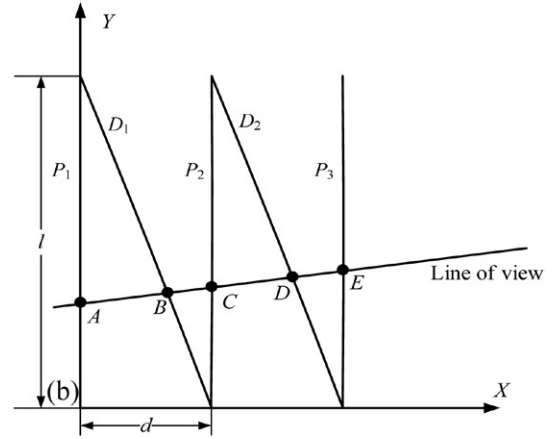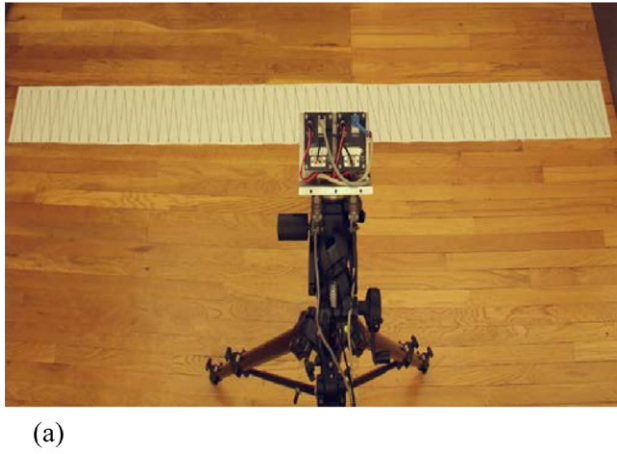
Figure 3. System calibration. (*a*) Camera mounted on a tripod to observe the target. (*b*) Calibration target.

thread will initiate an image-saving thread to store the queued images onto the hard disk at once. The saving thread runs in parallel to the acquisition thread. Therefore, image saving doesn't interrupt the image acquisition. The system captures and stores pavement images at a traveling speed up to $112\,\mathrm{km\,h^{-1}}$ (70 mph) without skipping.

3. Camera calibration

Image calibration is a process to determine the line-scan camera's parameters so that geometric distortions can be corrected and images can be registered properly for image fusion. Methods for calibrating 2D cameras have been widely available [11, 12], but the calibration of 1D line-scan cameras is substantially different from 2D cameras. Only a limited amount of research has been dedicated to line-scan camera calibration. These methods either require attaching the camera to a precision stage to scan a planar pattern [13, 14] or require 3D patterns [15]. Our goal is to design a flexible calibration method that works in a static setting, requires no mechanical devices, and uses only a planar pattern that is easy to carry. These are the important features for a calibration that is to be conducted in the work field.

Our approach requires a line-scan camera to view a planar pattern from at least two different angles and to define homography matrices through which feature points from the pattern coordinate space are mapped to image coordinate space. The next step is to estimate a homography matrix from each pattern observation, and to solve the intrinsic and extrinsic parameters with the known homography by applying constraints unique to the 1D imaging geometry.

3.1. Pattern design and geometry of feature points

Feature points in 2D camera calibration could be corners [11, 12], centers of circles [16, 17] or the intersection of lines [18]. Since the field of view of a line-scan camera is an invisible line, it is very difficult to detect the coordinates of feature points on the pattern. Our calibration involves a novel

pattern that allows us to estimate the scan line from the pattern geometry.

The pattern we use is a plane with two sets of mutually parallel lines interconnected in a zigzag configuration, as is shown in figure 3(*b*). The space between two vertical lines is 60 mm, and the height of each vertical line is 220 mm. A total of 58 vertical lines $\{P_n\}$ and 57 slanted lines $\{D_n\}$ are used. All these lines are defined on the $Z = 0$ plane. The origin of the pattern coordinate space is placed at the bottom end of P_1. The direction of the *x*-axis is perpendicular to the vertically parallel line set, and the direction of the *y*-axis is parallel to the vertical lines. Figure 3(*b*) shows two repeated pattern units to illustrate the pattern configuration.

For each repeated pattern unit, let *A*, *B*, *C* and *E* denote the four feature points seen by the camera, and *a*, *b*, *c* and *e* are their image coordinates. Since points *A*, *B*, *C* and *E* are collinear, the cross-ratio of these points is defined as

$$r_{ABCE} = (\overline{CA} \,/\, \overline{CB}) \,/\, (\overline{EA} \,/\, \overline{EB}) . \tag{2}$$

Because the cross-ratio of collinear points does not change under perspective projection, we have

$$r_{abce} = (\overline{ca} \,/\, cb) \,/\, (\overline{ea} \,/\, \overline{eb}) = r_{ABCE} . \tag{3}$$

Since the coordinates of *a*, *b*, *c* and *e* can be readily detected from a line-scanned image, one can compute the cross-ratio $r = r_{abce}$ from these coordinates, and further determine the coordinates of *A*, *B*, *C* and *E* as follows. Assume η is the ratio of $\overline{BA}$ to $\overline{CA}$, the relationship between η and *r* can be solved:

$$\eta = \frac{2r - 2}{2r - 1} . \tag{4}$$

Therefore, the *x*-coordinate of point *B* can be calculated as $x_B = \eta d$. Since point *B* is also on line D_1, its *y*-coordinate can also be calculated with the known line equation of D_1. Similarly, the intersection point *D* between the view line and D_2 can also be calculated by introducing the fourth vertical line P_4. With the coordinates of both *B* and *D* being determined, the equation of the view line section between P_1 and P_3 can be formulated, and the coordinates of points *A*, *C* and *E* can be determined as well.
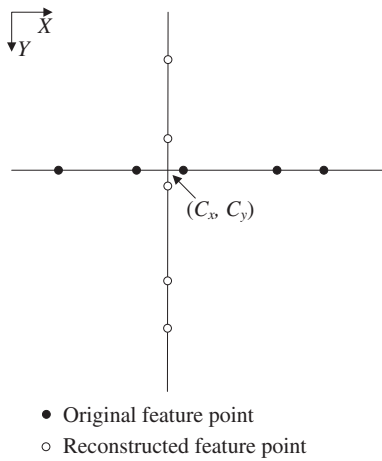
Figure 4. 1D to 2D feature point mapping.

3.2. 1D to 2D mapping

The imaging geometry of the 1D line-scan camera is a reduced form of the imaging geometry of the 2D camera in that only the imaging sensor is in 1D arrangement, but the other components, such as the lens distortion model and the perspective projection from the 3D world, remain unchanged. The classic 2D camera calibration methods based on a planar pattern cannot be specialized to the 1D problem. In line-scanned images, the feature points in both image coordinate and pattern coordinate are collinear, and they do not provide enough constraints in estimating a 3×3 homography. To make our 1D data fit the 2D scope, we applied a method to mirror the 1D feature points onto the second dimension so that the 3×3 homography can be estimated.

To create the projected feature points within the image coordinate space, we rotate the originally detected 1D feature points by 90 degrees around an estimated optical center C_x, which is initialized as half the width of the image. By doing this, it is equivalent to converting the 1D image into 2D in which the contrast in the second dimension is exactly duplicated from the first dimension. Thus, our 1D camera calibration problem becomes a 2D problem, on which some constraints on the intrinsic parameters can be applied: (1) the focal length must be the same in both vertical and horizontal directions, i.e. $f_x = f_y$ and (2) the optical center in both vertical and horizontal directions must be the same, i.e. $C_x = C_y$. Figure 4 illustrates the orthogonal projection to generate 2D feature point coordinates.

3.3. Camera calibration results

With the nonlinear mapping introduced by lens distortion, a solution of all camera parameters cannot be achieved directly in a least-squares fashion. Generally, given a set of feature point correspondences, the intrinsic and extrinsic parameters are optimized iteratively. The feature point estimation described in the previous section requires the intrinsic parameters, i.e. the calibration output, to estimate the feature points in the pattern coordinate space. This condition can be resolved

by nesting the iterative nonlinear optimization process within an iterative estimation of feature points. The combined calibration process thus consists of three steps: (1) estimate the initial camera parameters from a closed form solution. Feature point correspondences are generated by initializing the image center at the optical center; (2) apply nonlinear optimization of camera parameters based on initial estimates; (3) refine the positions of feature points based on optimized camera parameters. Steps 2 and 3 are repeated until convergence is reached. More details on solving the camera calibration parameters by applying the constraints that are unique in the geometry of line-scan cameras are reported in our previous work [19].

The accuracy and robustness of the calibration method were tested on a set of synthetic feature point correspondences first. In the perspective projection of a camera model, the location of an image feature point on the image plane is orientation-dependent. When generating the orthogonally reconstructed feature points with respect to the optical center, the distribution pattern of the original feature points should be preserved in the reconstructed points. Thus the simulation is helpful to discover if a rotation in 1D is significant to compromise the calibration results. Two observations are taken for each calibration, with the first being in the perpendicular pose and the second being looking at the calibration target with a small swivel angle varying from 2° to 20° in 2° intervals. Sixty independent trials were conducted for each orientation angle. It is discovered that the minimum error was observed at an orientation angle of about 5° [19]. This can be explained by the fact that the algorithm needs the variation of depths from the feature points to accurately estimate the focal length. The greater the orientation angle, the larger the variations of feature point depth. At around 5° the duplicated distribution of the reconstructed feature points is not significant enough to generate errors in calibration, while the depth variation from feature points is sufficient for the estimation of focal length.

The performance of the calibration method on line-scanned images was evaluated on twelve individual calibrations for each camera. In each individual calibration, two pictures were taken at two different poses. In the first pose, the camera looked at the pattern almost perpendicularly. In the second pose, a randomly selected +5° or −5° swivel angle was applied by turning the camera either to its left or to its right. The angle does not need to be precisely controlled. The means and standard deviations of the twelve calibrations are shown in table 1. The iterative process of refining the optical center takes three to five iterations to converge. In order to verify the correctness of the method, all feature points in the pattern are reprojected to the image coordinate space using the computed intrinsic and extrinsic parameters. The mean reprojection error was 0.46 pixels with standard deviation of 0.066 for the primary camera, and 0.49 pixels with standard deviation of 0.072 for the secondary camera.

3.4. Distortion correction

Figure 5 shows the preliminary results of lens distortion correction with the computed intrinsic camera parameters. The

Table 1. Calibration results and reprojection errors.

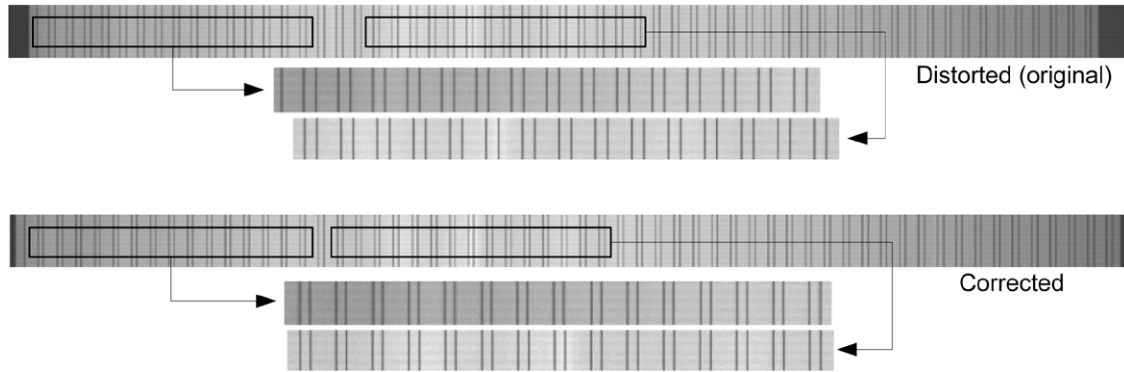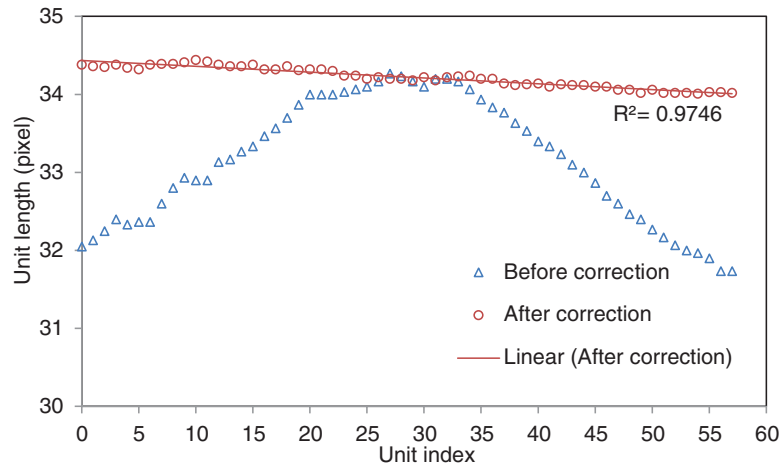|		f (pixel)	u_0 (pixel)	k_1	k_2	Reproj. errors
Primary	Mean	962.18	1021.46	−0.036	0.015	0.46
	STD	1.913	0.494	5.41×10^{-4}	3.25×10^{-4}	0.066
Secondary	Mean	978.27	1022.14	−0.032	0.023	0.49
	STD	1.741	0.499	6.26×10^{-4}	4.32×10^{-4}	0.072

Figure 5. Distortion correction on the pattern image.

Figure 6. Distance distribution between feature points before and after distortion correction.

image was captured with the camera observing our calibration pattern in a perpendicular orientation. Each vertical line segment in the image indicates a feature point on the pattern seen by the camera. If a perfect lens with zero radial distortion is considered, the line segments selected from the center and from the side of the image should be aligned due to the even distribution of feature points on the pattern. In figure 5, the upper image shows the original line-scanned image with radial distortion. As expected, the spaces between line segments on the sides are narrowed, and pattern lines are unevenly distributed. The lower one shows the result after radial distortion is removed so that all the line segments are evenly distributed and can be perfectly aligned.

To further verify the quality of distortion correction, we look at the linearity of the distances between the line segments on the image that are defined by feature points on pattern lines $\{Pi\}$. Since pattern lines $\{Pi\}$ are mutually parallel and are evenly spaced, when the camera observes these lines at an

angle, the distances between pairs of neighboring lines on the image are linearly distributed due to perspective projection. Figure 6 is a plot of the distribution before and after distortion correction. The R^2 value for the line fitting is 0.975. The two line-scan cameras were mounted on one piece of metal plate and were next to each other; therefore, the image planes of the two cameras are co-planar and the image registration only involves translations and rotations on the image plane once the image distortions are corrected.

4. Image fusion and shadow suppression

Image fusion is a process of combining information from two or more images into a single composite image that is more informative for visual perception or computer processing [20]. Many methods have been proposed to obtain a more informative image using multiple exposure techniques. Of these methods, the multi-scale decomposition (MSD) approach
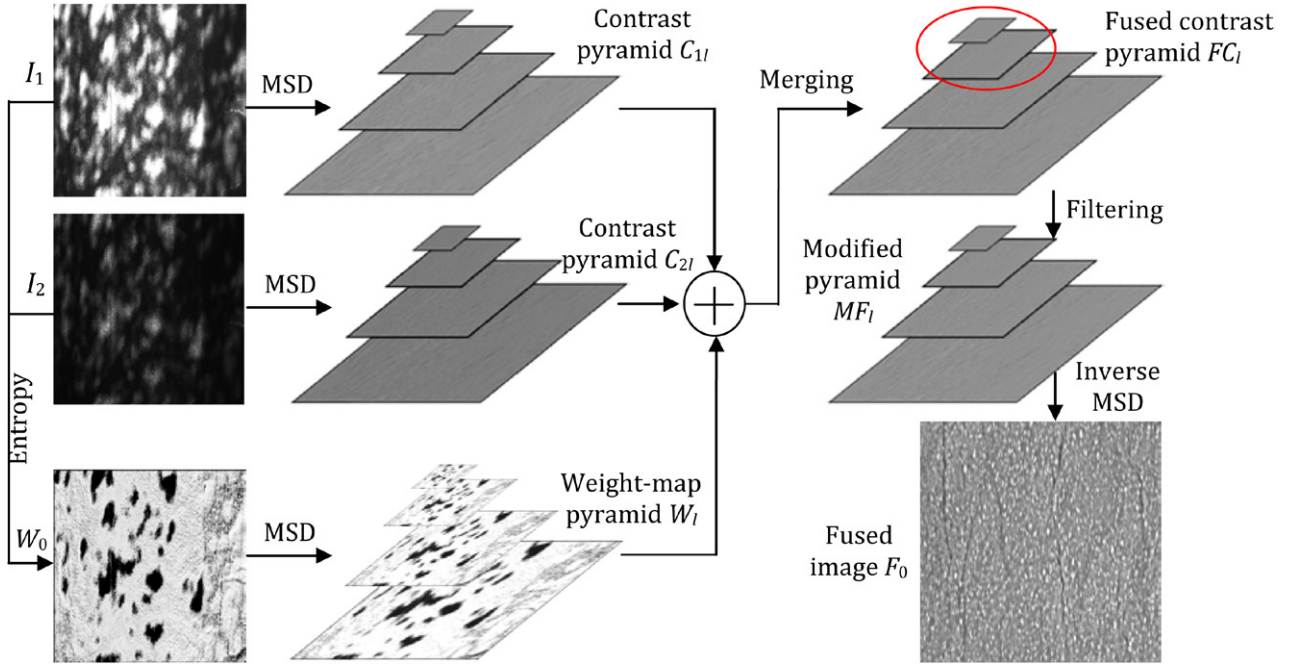
Figure 7. Scheme of MSD image fusion.

appears to be a particularly efficient and effective image fusion scheme [21–23]. In this scheme, the source images are transformed with MSD separately and merged at different scales when guided by a selected feature measurement. The composite image is then reconstructed by an inverse transform. Commonly used MSD methods for image fusion are Laplacian [24], contrast [25], gradient pyramid transforms [22, 26] and wavelet transform [27, 28]. The entropy, gradient, saturation and/or edge intensity are the features used to determine the weight map at different scales [21–23]. In the proposed system, we apply the contrast pyramid transform to merge the overexposed image and the underexposed image, and to create a new image in which all areas are well exposed and shadows are suppressed.

4.1. Contrast pyramids via MSD transform

A contrast pyramid is a representation of a sequence of contrast images decomposed from a source image via a MSD method. Let $G_l(i,j)$ be an image at the lth level in the pyramid for the source image G_0 ($0 \le l \le N$, N is the top level of the pyramid) and I be the original source image. Therefore, $G_0 = I$, and G_l can be obtained by convolving the image at the previous level G_{l-1} with a 5×5 weighting function $w(m,n)$,

$$G_l = \text{REDUCE}\,[G_{l-1}] = w \otimes G_{l-1},$$

$$\text{or} \quad G_l(i,j) = \sum_{m=-2}^{2} \sum_{n=-2}^{2} w(m,n)\,G_{l-1}(2i+m, 2j+n),$$

(5)

where $w(m,n)$ is a Gaussian-like function when it obeys the constraints: $w(m,n) = w'(m)\,w'(n)$, $w'(0) = 0.4$, and $w'(1) = w'(-1) = 0.5$. Let $G_{l,1}$ be the image obtained by expanding G_l one time, then

$$G_{l,1} = \text{EXPAND}\,[G_{l+1}].$$

(6)

or,

$$G_{l,1}(i,j) =$$
$$4 \sum_{m=-2}^{2} \sum_{n=-2}^{2} w(m,n)\,G_{l+1}((\text{floor})(i+m)/2, (\text{floor})(j+n)/2).$$

(7)

To highlight edge information (e.g. cracks) in the image, contrast pyramid C_l can be built by using the relative difference between two levels of G_l as follows:

$$C_l = \left(G_l - \text{EXPAND}\,[G_{l+1,1}]\right) / \text{EXPAND}\,[G_{l+1,1}].$$

(8)

The contrast pyramid can be iteratively generated by applying the REDUCE and EXPAND functions to the source image. Figure 7 displays the contrast pyramids of two source images I_1 (overexposed) and I_2 (underexposed) in which tree shadows complicate the appearance of pavement cracks.

4.2. Activity measure and weight-map pyramid

The activity measure reflects the amount of useful information carried in each source image; it is used to determine the contribution of the coefficients to the merged multi-scale representation. This attribute largely depends on the specific objective of the fusion application as well as the features of source images. The well-exposed areas need to be distinguished from overexposed or underexposed areas in each source image. Generally, an area that is overexposed or underexposed contains less texture information than a well-exposed area. Entropy is a measure of the texture information of an area or image [29] and can be computed by $E_g = \sum_{i=0}^{255} -p_i \log(p_i)$, where p_i is the
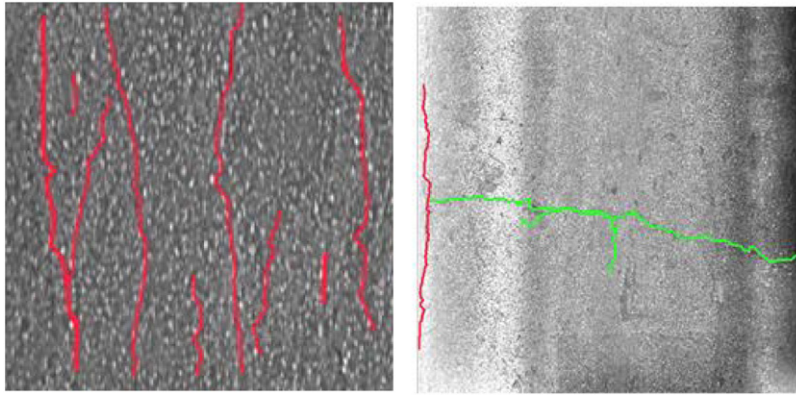
Figure 8. Detected cracks in a tree-shadowed image (left) and a vehicle-shadowed image (right).

probability of an arbitrary pixel that has grayscale i ($i = 0, \ldots,$ 255). When the number of pixels having grayscale i is n_i, and the image contains n pixels, $p_i = n_i / n$. In this application, the entropy will be calculated in a squared window (e.g. 5×5) for each pixel, and a weight map W_0 will be formed using the entropies of the two source images, I_1 and I_2. Each pixel in W_0 contains the larger entropy of I_1 and I_2 at the same location to select a higher activity (see W_0 in figure 7). Based on W_0, a weight-map pyramid W_l can be built via the MSD transform as defined in equation (7), as shown by W_l in figure 7. W_l provides weighting coefficients for merging two contrast pyramids at each level.

4.3. Merging strategy

The merging strategy refers to the approach to combine multiple multi-scale representations into one that contains more desirable information. For an application that involves only two source images, the weighted-average strategy is preferred. Based on the three pyramids, W_l, C_{1l} and C_{2l}, the fused contrast pyramid F_l can be derived as follows:

$$F_l(i,j) = W_l(i,j) \times C_{1l}(i,j) + [1 - W_l(i,j)] \times C_{2l}(i,j). \quad (9)$$

Through this blending, shadows in the images will be attenuated. To further suppress residual shadows, a high-pass filter can be applied to the contrast images at the top levels of the pyramid to form a modified pyramid (see MF_l in figure 7). Shadows are normally much wider than cracks and are present in top-level images in the pyramid. On the other hand, finer details, such as cracks, are only present in the lower-scale images, and they will not be suppressed by the filtering applied in the high levels. We will determine optimal levels at which the high-pass filter should be applied.

4.4. Image reconstruction

Image reconstruction is an inverse transform of the MSD to restore one image at the initial scale from the modified pyramid MF_l. This is a recursive procedure starting at the top level N of the pyramid:

$$F_N = MF_N \quad (10)$$

$$F_l = MF_l \cdot \text{EXPAND}\left[F_{l+1,1}\right] + \text{EXPAND}\left[F_{l+1,1}\right]. \quad (11)$$

As seen in figure 7, the fused image F_0 not only combines the two source images together with cracks being preserved, but eliminates unwanted information such as shadows. The proposed MSD image fusion scheme is summarized in figure 7 with the following major steps:

(a) Build the contrast pyramids C_{1l} and C_{2l} for source images I_1 and I_2.
(b) Calculate pixel-based entropy and build weight-map pyramid W_l providing merging coefficients.
(c) Merge C_{1l} and C_{2l} at each level using the weighted-average strategy to create a fused contrast pyramid FC_l.
(d) Apply a high-pass filter at top levels of FC_l to suppress shadows and get the modified pyramid MF_l.
(e) Reconstruct the fused image F_0 from MF_l using the inverse MSD transform.

4.5. Crack detection

Crack detection is the step to trace and measure cracks in the fused image by using a seed-growing algorithm developed in our previous project [30]. The procedures of the crack detection algorithm includes grid feature extraction, seed identification and seed connection. Within the algorithm framework, a pavement image is divided into a number of grids of a certain size (e.g. 8×8 pixels), and each grid is classified either as a non-crack or crack grid (called seed) using the texture information of the grid and the overall background. Then, seeds in the vicinity are connected based on geometrical and intensity constraints. The connected seeds serve as candidates for a crack, which is further verified by using the contrast to the pixels along its trace. The crack detection algorithm has been tested and evaluated with a section of public asphalt road. Pavement images were captured by the same line-scan camera but with laser illumination to eliminate shadows. Crack detection results were compared to crack maps that were generated on grid paper by field experts
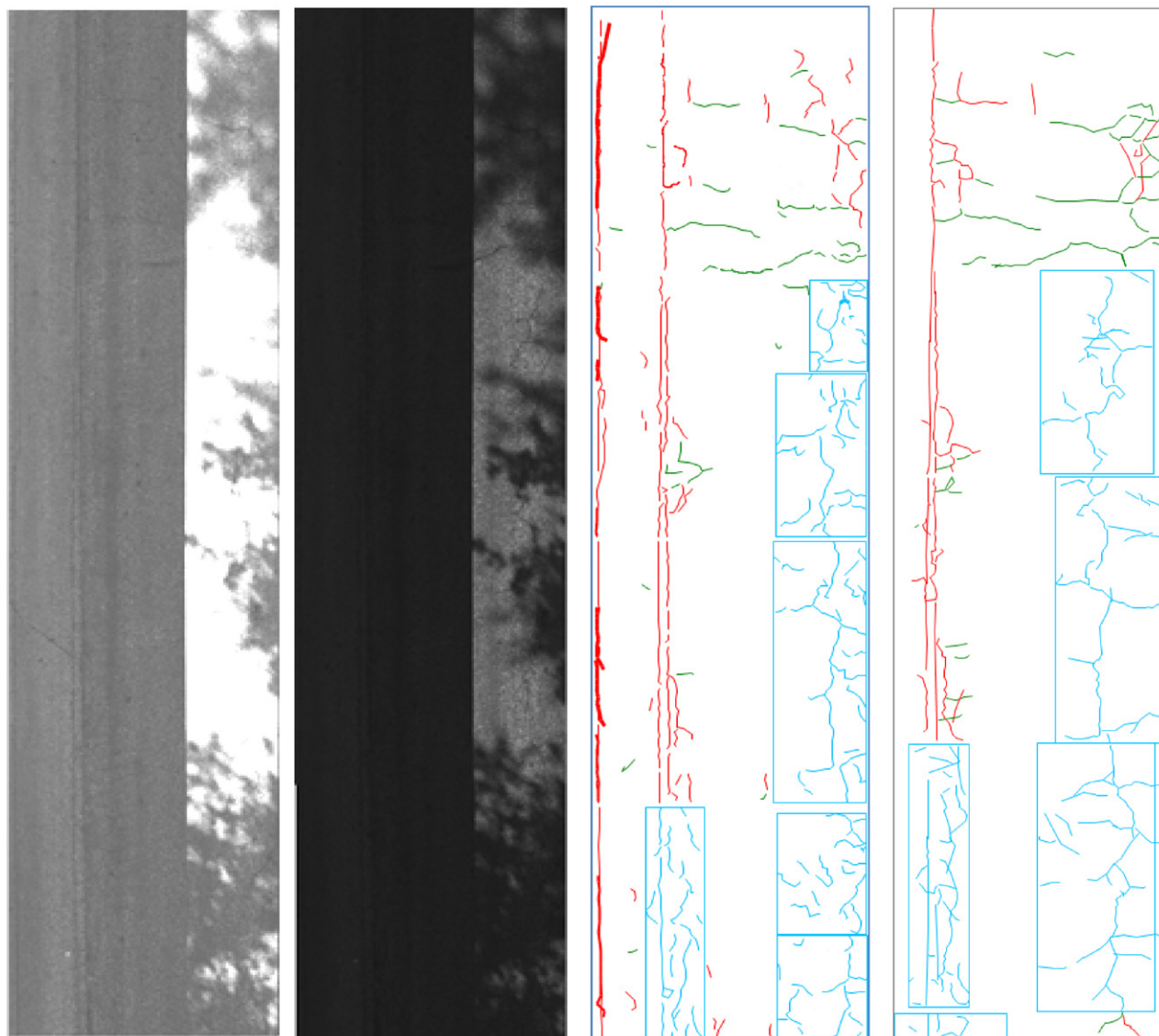
Figure 9. Cracks detected from complementary images are shown in crack maps. From left to right: overexposed pavement image, underexposed pavement image, results of automatic crack detection from complementary images, and results of manual crack identification from shadow-free image. Edge cracks may be different on the crack maps due to the difficulty of driving the vehicle along the same path across multiple scans.

Table 2. Crack measurements under different lighting conditions.

Shadow	Longitudinal (m)	Transverse (m)	Alligator A_a (%)	A_d (m/m^2)	Total (m)
With shadow auto	242.45	179.92	22.90	2.31	811.62
Without shadow auto	254.66	164.03	23.18	2.50	856.67
Without shadow manual	232.65	156.67	23.99	2.74	872.97

from TxDOT using visual evaluation. Results from automatic crack detection were mostly consistent with visual evaluation. The difference in crack length between the automatic and manual measurement was less than 10%, most of which comes from the fact that some shallow and narrow cracks are not captured with good contrast [30]. Figure 8 presents two examples of the detected cracks in the fused pavement images of the two pairs of complementary images in figure 2. Most of the cracks in these images are correctly marked,

indicating that the fused images preserve crack information after the shadows are eliminated.

5. Results and discussions

5.1. Crack detection accuracy

The leading protocol used in quantitatively evaluating pavement surface cracking in the United States is the standard

Table 3. Crack measurements in multiple runs.

	Longitudinal (m)			Transverse (m)			Alligator					
							Day 1		Day 2		Day 3	
Run	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	A_a (%)	A_d (m/m^2)	A_a (%)	A_d (m/m^2)	A_a (%)	A_d (m/m^2)
1	218.51	259.43	236.79	164.65	187.75	169.52	23.19	2.62	22.38	2.28	24.25	2.43
2	204.14	254.54	231.07	150.86	204.28	201.31	21.65	2.78	25.40	2.63	23.07	1.99
3	183.90	202.47	259.49	170.72	157.21	168.94	23.20	2.52	22.85	2.65	21.39	2.53
CV (%)	8.60	14.51	6.20	5.91	14.72	10.29	3.92	5.00	6.90	8.24	6.28	12.41

published by the American Association of State Highway and Transportation Officials (AASHTO) [31]. The protocol defines the procedures for quantifying cracks in asphalt pavement surfaces both in wheelpath and nonwheelpath areas. It also establishes a criterion for the minimum acceptable results from automatic crack detection: a crack map must include at least 30% of the level-1 cracks (whose width is under 3 mm), 50% of level-2 cracks (width from 3 to 5 mm) and 85% of the level-3 cracks (width above 5 mm). The AASHTO protocol is ideal for pavement with lane marks, so that wheelpaths areas can be estimated. Since our field test was conducted on a section of a local one-lane street and the wheelpath areas were ambiguous, we thus adopted the Texas Pavement Management Information System (PMIS) protocol [32] for our quantitative evaluation. The PMIS standard classifies cracks into different types, for example, longitudinal crack, transverse crack and alligator crack. Different crack types reveal possible reasons for pavement fractures. For instance, longitudinal cracking suggests overloaded traffic, while transverse cracking implies more likely a failure of subsurface structures.

We conducted field tests on an asphalt road on the campus of the University of Texas at Austin in a period which spanned over 1.5 years. The section of the road is about 220 m long with various shadows from building, trees and power lines. We chose the results from three test days in this period to report the performance of the system under different lighting conditions and road conditions. The first test day (Day 1) was about 15 and 17 months ahead of the second (Day 2) and third (Day 3) test days, respectively. On each test day, multiple runs were conducted to test data repeatability.

Figure 9 shows an example of the image series from the primary (overexposed) and secondary (underexposed) cameras and the corresponding crack map. The automatic crack detection results were compared to the results from a manual crack identification procedure on the shadow-free images of the same road captured by the system. The manual crack identification procedure was performed by an experienced researcher manually marking the cracks on the computer screen with editing tools that we developed together with the crack auto-detection software. The image series shown in figure 9 were formed by stitching four consecutive images, covering a 14.6 m long pavement section. The complementary images had both vehicle shadows and tree shadows. The crack map displays the locations and orientations of cracks useful for visual assessment. Longitudinal cracks were marked

in red, transverse cracks in green, and the areas of alligator cracking (short, dense and tightly connected) in blue rectangles. We found that this automatically detected crack map is consistent with the road cracking condition, particularly on these wide and deep cracks. The automatic detection results show minor differences with our manual markups, mostly on short cracks that are only one pixel in width.
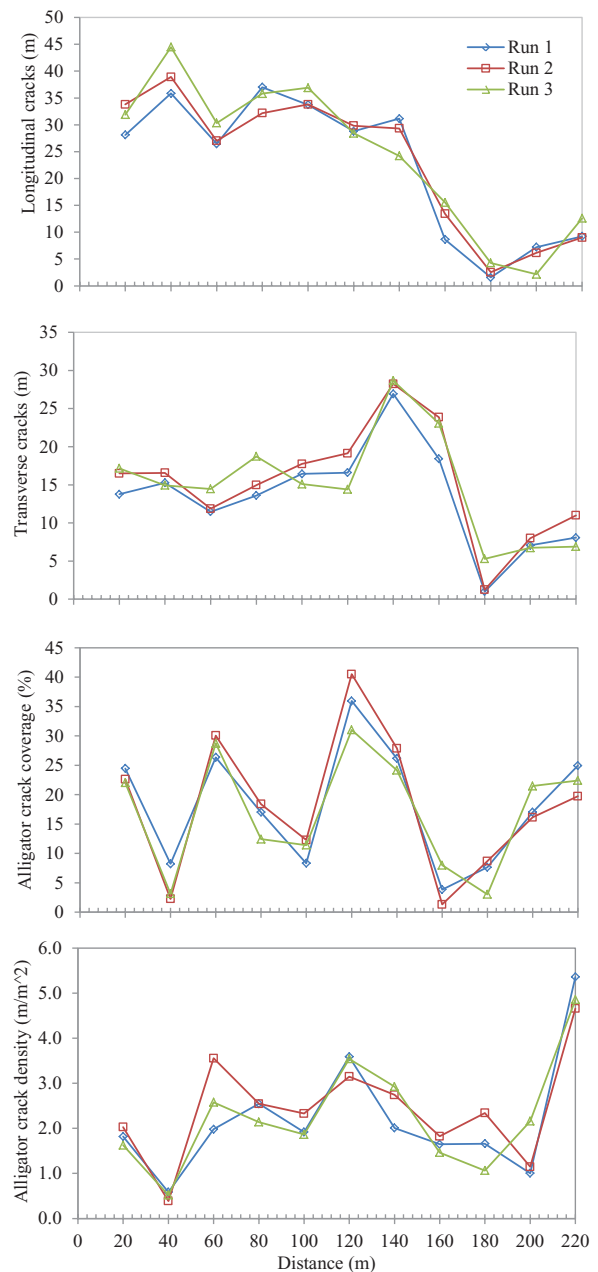
Table 2 presents the sums of crack data collected under two lighting conditions (sunny and cloudy) around noon, and the manual crack identification results. When it was sunny, there were both vehicle and tree shadows in the images. Longitudinal and transverse cracks are presented in total length. Alligator cracking is presented in the average percentage by area A_a, and the average density A_d. The last column in table 2 lists the total crack length by combining the length of all longitudinal, transverse and alligator cracks. Comparing to the results of the manual crack identification on shadow-free images (the third row), the system tends to overestimate the longitudinal cracks, but to underestimate transverse and alligator cracks. However, the total crack length of automatic crack detection is more consistent than each of the individual crack types. Part of the reason for this is the difficulty of driving the vehicle along exactly the same path during multiple scans. Thus, some cracks on the edges of the lane may or may not be captured, as evidenced in figure 9 in which a longitudinal crack on the left side of the crack map from the automatic crack detection from complementary images did not exist in the crack map from the manual crack identification. The difference may also indicate that our automatic crack detection algorithm is prone to designate dense cracks into a smaller alligator-cracking region. Some falsely detected transverse cracks in shadowed images were found to be fine shadows of overhead wires that were too similar to large cracks in the images.

A further comparison was made to show regional differences in the crack data between images with and without shadows by dividing the 220 m pavement into 55 equal segments, as shown in figure 10. Each data point in the figure represents the crack data summarized within one segment (a 4 m distance). The crack data collected in the two different lighting conditions demonstrated consistent trends, with R^2 being 0.87, 0.81, 0.93 and 0.85 in the longitudinal cracks, transverse cracks, and alligator crack coverage (A_a) and density (A_d), respectively. Minor disagreement between the automatic detection and the manual identification at certain distances can be observed, for example the longitudinal
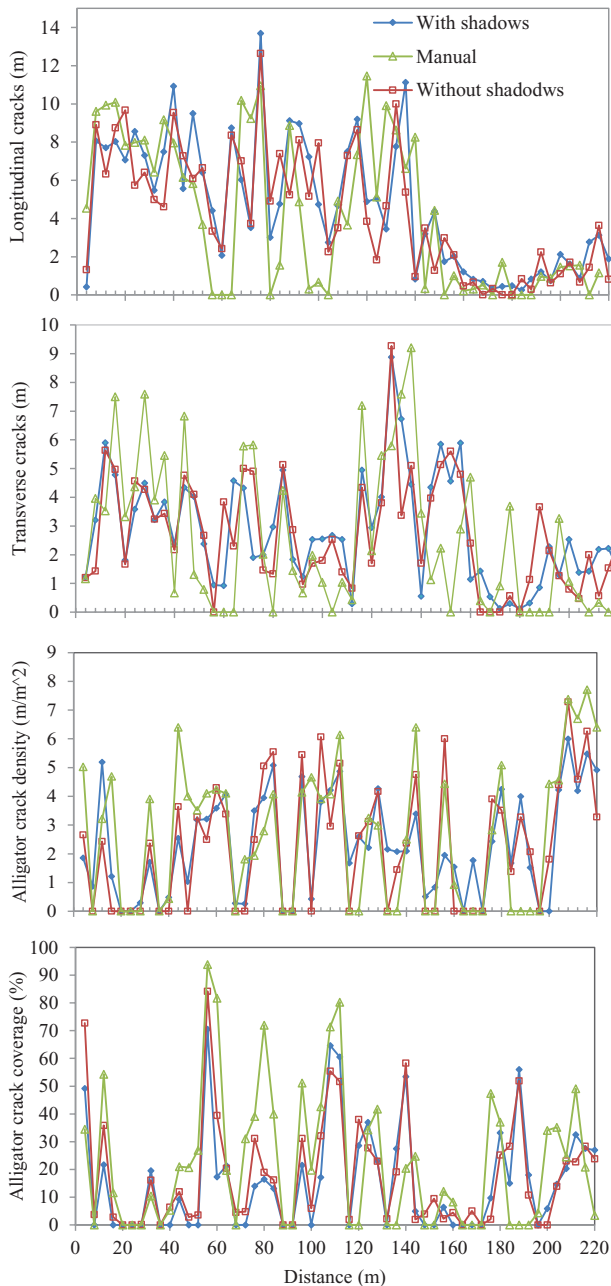
Figure 10. Regional crack data from images captured on sunny and cloudy days.

Figure 11. Repeatability of crack measurements in multiple runs.

cracks at distance marks 65 m and 85 m, and transverse cracks at marks 30 m and 40 m. These disagreements are caused by either false crack detection or counting cracks into an alligator block. An alligator cracking region represents surface fatigue where cracks appear in short segments and with smeared edges. Thus, it is more difficult to detect complete alligator cracks. This is one of the main reasons why the A_d (alligator density) is low. Because cracks cannot be counted repeatedly as different types, the manual detection showed lower longitudinal and transverse crack lengths than the two automatic detections. However, the total crack lengths are quite consistent between the automatic detections and the manual identification, with the highest difference being 7.5%.

5.2. Repeatability of multiple runs

Table 3 lists the measurements of longitudinal, transverse and alligator cracks in three runs on each test day. The graphs are shown in figure 11. Here, the coefficients of variation (CV) show the precision of the system among the three runs on 'Day 1', 'Day 2' and 'Day 3' respectively. Both longitudinal and transverse cracks were measured by lengths in meter, while alligator cracks were measured by alligator crack coverage (A_a) and alligator crack density (A_d). A_a indicates the percentage of the alligator cracking area over the whole pavement area, and A_d is equal to the sum of alligator crack lengths divided by the alligator cracking area. The CV values of all the crack categories in the three runs varied approximately from 4% to 15% on all test days. Since it was difficult to keep

Table 4. Crack measurements over a 1.5 year period.

Road test date		L_1 (m)	L_2 (m)	L_3 (m)	T_1 (m)	T_2 (m)	T_3 (m)	A_a (%)	A_d (m/m²)
Day 1		9.49 (1.88)	127.90 (4.72)	64.79 (12.04)	17.99 (2.51)	129.69 (13.65)	14.39 (3.65)	22.67 (0.89)	2.63 (0.13)
Day 2		20.85 (2.37)	149.50 (13.69)	68.46 (6.24)	17.17 (3.67)	147.84 (20.80)	18.07 (3.45)	23.54 (1.62)	2.52 (0.21)
(15 months later)									
Day 3		23.71 (1.55)	151.48 (7.94)	67.26 (5.81)	22.01 (1.9)	140.85 (18.33)	17.06 (1.73)	22.90 (1.44)	2.31 (0.29)
(17 months later)									
Change over	Δ_{21}	119.60	16.89	7.22	−4.87	13.98	26.22	2.79	8.81
time (%)	Δ_{31}	149.64	18.44	3.81	22.35	8.60	18.74	−1.04	13.62

the moving vehicle on the same driving path in different runs, some of these variations arose from the differences in the covered areas of the line-scan cameras in different runs.

In order to see the data consistency in different runs, we divided the 220 m pavement into 11 equal segments, and calculated the longitudinal, transverse and alligator crack data for each 20 m segment. Figure 10 displays the crack data in the three runs on Day 3. The three runs exhibit similar trends with the average correlation coefficients (R^2) of any two runs being 0.96, 0.94, 0.93 and 0.88 in the longitudinal cracks, transverse cracks, and alligator crack coverage (A_a) and density (A_d), respectively. These graphs also reveal the regional or localized cracking information, which is helpful for locating distressed pavements.

5.3. Pavement surface change over time

The crack data from the three test days can be used to check pavement surface changes over time since the testing period lasted 1.5 years. For longitudinal and transverse cracking, cracks are often categorized into three severity levels according to the crack width [31]. Table 4 lists the crack measurements at different severity levels where L_1, L_2, L_3, T_1, T_2 and T_3 are the longitudinal and transverse cracks at the three severity levels, Δ_{21} and Δ_{31} are the change from Day 1 to Day 2 and the change from Day 1 to Day 3. The values in parentheses are the standard deviations. A negative change indicates a decrease in crack length.

From table 4, we can see that longitudinal cracks increased at all three levels, particularly at level 1 (L_1), in both Δ_{21} and Δ_{31}, while transverse cracks had mixed change patterns. Transverse cracks increased at levels 2 and 3 (T_2 and T_3) in the first period (Δ_{21}) and then at levels 1 and 3 (T_1 and T_3) in the second period (Δ_{31}). Even though the alligator coverage (A_a) doesn't show significant changes over time, the alligator density (A_d) has increased obviously by 8.81% in Δ_{21} and 13.62% in Δ_{31}, meaning that the surveyed road became severely fatigued over time.

6. Conclusion

This new APDS system uses a dual line-scanning strategy to simultaneously acquire two images of the same pavement surface with different exposures, and applies an image fusion algorithm to merge the paired images into one composite image in which shadows of roadside objects and vehicles are suppressed. To generate complementary information, one camera is set in

a high-exposure mode to prevent shadowed regions from being blacked out in the image, and the second camera is set in a low-exposure mode to prevent sunlit regions from being whited out. Both high-exposure and low-exposure images are decomposed with the pyramid transform, and their contrast pyramids are merged at each level when guided by image entropies. The high frequency filtering is applied at the top level of the fused contrast pyramid to suppress shadows. Then, the inverse transform is employed to reconstruct a shadow-free image for crack detection. An effective calibration method for line-scan cameras is implemented prior to the image fusion to rectify camera distortions and align paired images for accurate pixel matching. The field tests on a selected road were performed over a 1.5 year period to evaluate the system's repeatability in multiple runs and its consistency in different lighting conditions. It was found that the CVs of longitudinal, transverse and alligator cracks measured on the three test days were all under 15%, and the R^2 of the three crack categories measured on sunny and cloudy days were all above 0.81. This APDS system not only eliminates the need for artificial lighting, but also reduces installation and maintenance costs and the consumption of energy in the vehicle.

Disclosure

An author of this publication owns Fabrate, LLC., an engineering services company that commercializes the technology licensed from the University of Texas at Austin. The terms of this arrangement have been reviewed and approved by the University of Texas at Austin in accordance with its policy on objectivity in research.

References

[1] Saraf C L 1998 Pavement condition rating system, review of PCR methodology *Report* FHWA/OH-99-001, Ohio Department of Transportation

[2] Cheng H D, Jiang X, Li J and Glazier C 1999 Automated real-time pavement distress analysis *Trans. Res. Rec.: J. Trans. Res. Board* **1655** 55–64

[3] Wang K C 1998 Automated system for pavement surface distress survey: a historical perspective on design and implementation *Proc. Transport Research Board 77th Annual Meeting (Washington, DC)* (Washington, DC: Transportation Research Board)

[4] Wang K C 1999 Use of digital cameras for pavement surface distress survey *Trans. Res. Record: J. Trans. Res. Board* **1675** 91–7

[5] NCHRP Synthesis 334 2004 *Automated Pavement Distress Collection Techniques* (Washington, DC: Transportation Research Board)

[6] Howe R and CleMena G G 1997 *An Assessment of The Feasibility of Developing and Implementing an Automated Pavement Distress Survey System Incorporating Digital Image Processing* (Richmond, VA: Virginia Transportation Research Council) No. VTRC 98-R1

[7] Wang K C P 2000 Designs and implementations of automated systems for pavement surface distress survey *J. Infrastruct. Syst.* **6** 24–32

[8] Xu B 2008 Summary of implementation of an artificial light system for automated visual distress rating system *Technical Report* FHWA/TX-08/5-4958-01-1 Center for Transportation Research, The University of Texas at Austin

[9] Pavemetrics System Inc. Laser Road Imaging System http://pavemetrics.com/en/lris.html (last accessed September 2014)

[10] Xu B, Yao X and Yao M 2007 Design of CrackScope (VCrack) *Technical Report* FHWA/TX-08/0-5708-1 Center for Transportation Research, The University of Texas at Austin

[11] Zhang Z 2000 A flexible new technique for camera calibration *IEEE Trans. Pattern Anal. Mach. Intell.* **22** 1330–44

[12] Tsai R 1987 A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses *IEEE J. Robot. Autom.* **4** 323–44

[13] Horaud R, Mohr R and Lorecki B 1993 On single-scanline camera calibration *IEEE Trans. Robot. Autom.* **9** 71–5

[14] Drareni J, Roy S and Sturm P 2011 Plane-based calibration for linear cameras *Int. J. Comput. Vis.* **91** 145–56

[15] Luna C A, Mazo M, Lazaro J L and Vazquez J F 2010 Calibration of line-scan cameras *IEEE Trans. Instrum. Meas.* **59** 2185–90

[16] Gurdjos G, Crouzil A and Payrissat R 2002 Another way of looking at plane-based calibration: the center circle constraint *Proc. European Conf. Computer Vision* vol 4 pp 252–66

[17] Fremont V and Chellali R 2002 Direct camera calibration using two concentric circles from a single view *Proc. Int. Conf. Artificial Reality and Telexistence (Tokyo)* pp 93–8

[18] Grosky W I and Tamburino L A 1990 A unified approach to the linear camera calibration problem *IEEE Trans. Pattern Anal. Mach. Intell.* **12** 663–71

[19] Yao M, Zhao Z and Xu B 2014 Geometric calibration of line-scan camera using a planar pattern *J. Electron. Imaging* **23** 013028

[20] Ardeshir Goshtasby A and Nikolov S 2007 Image fusion: advances in the state of the art *Inform. Fusion* **8** 114–8

[21] Mertens T, Kautz J and Van Reeth F 2007 Exposure fusion *15th Pacific Conf. on Computer Graphics and Applications, 2007. PG'07 (Maui, HI)*

[22] Block M *et al* 2009 Multi-exposure document fusion based on edge-intensities *10th Int. Conf. on Document Analysis and Recognition, 2009. ICDAR '09 (Barcelona)*

[23] Malik M H, Asif S and Gilani M 2008 Wavelet based exposure fusion *World Congress on Engineering 2008 (London)*

[24] Burt P and Adelson E 1983 The Laplacian pyramid as a compact image code *IEEE Trans. Commun.* **31** 532–40

[25] Toet A, van Ruyven L J and Valeton J M 1989 Merging thermal and visual images by a contrast pyramid *Opt. Eng.* **28** 789–92

[26] Burt P J and Kolczynski R J 1993 Enhanced image capture through fusion *Proc. Fourth Int. Conf. on Computer Vision (Berlin)*

[27] Mallat S G 1989 Multifrequency channel decompositions of images and wavelet models *IEEE Trans. Acoust. Speech Signal Process.* **37** 2091–110

[28] Li H, Manjunath B S and Mitra S K 1995 Multisensor image fusion using the wavelet transform *Graph. Models Image Process.* **57** 235–45

[29] Frieden B R 1972 Restoring with maximum likelihood and maximum entropy *J. Opt. Soc. Am.* **62** 511–8

[30] Yao X, Yao M and Xu B 2011 Automated measurements of road cracks using line-scan imaging *J. Test. Eval.* **39** 103331

[31] American Association of State Highway Transportation Officials 2000 *Protocol PP 44, Quantifying Cracks in Asphalt Pavement Surface* Washington, DC

[32] Texas Department of Transportation 2009 Pavement management information system rater's manual https://ftp.dot.state.tx.us/pub/txdot-info/cst/raters_manual.pdf (last accessed 11 February 2014)