

Copyright
by
Zuyun Zhao
2014

The Thesis Committee for Zuyun Zhao
Certifies that this is the approved version of the following thesis:

Complementary Imaging for Pavement Cracking Measurements

APPROVED BY
SUPERVISING COMMITTEE:

Supervisor:

Bugao Xu

Jonathan Y. Chen

Complementary Imaging for Pavement Cracking Measurements

by

Zuyun Zhao, B.E.; M.E.; M.S.T.A.T.

Thesis

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science in Engineering

The University of Texas at Austin

December 2014

Abstract

Complementary Imaging for Pavement Cracking Measurements

Zuyun Zhao, M.S.E.

The University of Texas at Austin, 2014

Supervisor: Bugao Xu

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 of the artificial lighting devices are laser based, which are either hazardous to unprotected people, or require dedicated power supplies on the vehicle.

This study is aimed to develop a new imaging system that can scan pavement surface at highway speed and determine the severity level 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 presents (1) the design of the dual line-scan camera system for a high-speed pavement imaging system that does not require artificial lighting, (2) a new calibration method for line-scan cameras to rectify and register paired images, which does not need mechanical assistance for dynamical scan, (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.

Table of Contents

List of Tables	viii
List of Figures	x
Chapter 1: Introduction.....	1
1.1 Background of Pavement Distress Measurement	1
1.2 Methods for Pavement Image Acquisition	1
1.3 Methods of Line-scan Camera Calibration	4
1.4 Methods of Image Fusion	5
1.5 Motivation and Goals.....	6
1.6 Structures of the Thesis.....	7
Chapter 2: Framework of a Dual Camera Pavement Imaging System	9
2.1 Complementary Imaging	9
2.2 System Setup.....	10
2.3 Image Brightness Control	11
2.4 Image Acquisition	13
Chapter 3: Geometric Calibration of Line-scan Camera	15
3.1 Introduction.....	15
3.2 Camera Model and Imaging Geometry.....	17
3.2.1 Preliminary.....	18
3.2.2 Estimation of Homography for 1D Line-scan Camera	19
3.3 Calibration Pattern and Feature Point Correspondences	20
3.3.1 Calibration Pattern Design.....	20
3.3.2 Feature Point Estimation.....	21
3.3.3 Feature Point Correspondences.....	23
3.4 Solving the Camera Parameters	25
3.4.1 Closed-form Solution.....	25
3.4.2 Nonlinear Optimization	27
3.5 Experiments and Discussions	27

3.5.1 Synthetic Data.....	28
3.5.1.1 Noise test.....	28
3.5.1.2 Orientation test.....	29
3.5.2 Real Images.....	30
3.5.3 Calibration Errors.....	34
3.5.3.1 1D to 2D conversion	35
3.5.3.2 Radial distortion	35
3.5.3.3 Calibration target.....	36
3.6 Conclusion	36
Chapter 4: Image Fusion and Shadow Suppression	38
4.1 Introduction.....	38
4.2 Methodology	39
4.3 Pyramid MSD Transform	40
4.3.1 Contrast Pyramids via MSD Transform	40
4.3.2 Activity Measure and Weight-map Pyramid	42
4.3.3 Merging Strategy and Shadow Suppression	43
4.3.4 Image Reconstruction	44
4.4 Crack Detection	44
Chapter 5: Result Analysis and Discussions	46
5.1 Crack Detection Accuracy	46
5.2 Repeatability of Multiple Runs.....	49
5.3 Pavement Surface Change Over time	51
Chapter 6: Conclusions and Future Work	53
6.1 Summary of the Thesis	53
6.2 Suggested Future Work.....	54
Bibliography	55

List of Figures

Figure 2.1 Schematic design of the complementary imaging system.....	11
Figure 2.2 Complementary images and their grayscale histograms. Pixels with extreme values (0 and 255) are excluded.....	13
Figure 3.1 More than two feature points that are collinear in the pattern coordinate space do not add new constraints in solving the homography. P_2 and P_3 can be estimated by extending the ray from the focal point F to the line-of-view. ...	20
Figure 3.2 The calibration pattern being viewed by a line-scan camera. Only two repeated line structures are shown.....	21
Figure 3.3 Reconstruction of calibration feature points.....	24
Figure 3.4 Calibration errors from simulations.....	29
Figure 3.5 Examples of distortion correction on pattern images. Top part: original image with lens distortion. Pattern lines are unevenly spaced; Bottom part: distortion corrected. Pattern lines are evenly spaced and can be aligned.	32
Figure 3.6 Plots of unit lengths before and after radial distortion correction.	33
Figure 3.7 Fusion results from images captured by two identical line-scan cameras with two different exposure settings.	34
Figure 4.1 Scheme of MSD image fusion.....	42
Figure 4.2 Contrast pyramids of source images overexposed (left) and underexposed (right)	42
Figure 4.3 Detected cracks in a tree shadowed image (left) and a vehicle shadowed image (right).	45
Figure 5.1 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 in driving the vehicle along the same path across multiple scans.....	48
Figure 5.2 Regional crack data from images captured in sunny and cloudy days.....	51
Figure 5.3 Repeatability of crack measurements in multiple runs.....	51

List of Tables

Table 3.1 Calibration results for intrinsic parameters and reprojection errors	31
Table 5.1 Crack measurements under different lighting conditions.	49
Table 5.2 Crack measurements in multiple runs.....	50
Table 5.3 Crack measurements over a 1.5-year period.....	52

Chapter 1: Introduction

1.1 BACKGROUND OF PAVEMENT DISTRESS MEASUREMENT

The pavement condition is evaluated greatly upon the appearance of pavement distress. Cracking is one of the most common distress forms which limits the pavement performance and brings potential safety issue to drivers. Accurate measurement of the type, density and coverage of cracking is essential for the pavement maintenance officials to assess the pavement condition and come up with a proper pavement maintenance plan. Manual visual rating was the primary method of pavement inspection before early 1970s, and it was then replaced by the emerging automated pavement distress survey (APDS) systems due to its disadvantage at traffic disturbance, survey cost and risk to human inspectors. APDS systems aim to get rid of the limits in manual visual rating and provide more objective and prompt results for the pavement maintenance management [1]-[4]. Due to the complexity of real pavement circumstance and different natural lighting conditions, APDS still presents many challenges and issues [5]-[7].

In general, an APDS system contains image acquisition devices and image processing software that generates quantitative measurements and ratings. The technique for high quality pavement image acquisition and the design of robust, accurate, and efficient distress recognition algorithm remain as two active topics in related research fields. Among these two, the imaging quality is always an important factor which affects the subsequent pavement distress detection and recognition performance.

1.2 METHODS FOR PAVEMENT IMAGE ACQUISITION

APDS systems differ in their image acquisition devices. They can be generally divided into two categories, two-dimension (2D) area-scan camera and one-dimensional

(1D) line-scan camera. Unlike the area-scan camera, line-scan camera captures the object by one line at a time. The image is generated by moving the object, or the camera itself, and accumulating of the scanned lines over time. This property matches the nature of pavement imaging in which the camera is mounted to an image capturing vehicle and the whole system is travelling at a proper highway speed. The viewing direction is usually perpendicular to pavement surface. This helps decrease the distortion caused by perspective projection at a tilt angle. Line-scan camera also has a larger signal-to-noise ratio than regular 2D camera due to the larger imaging sensor size. This is an important feature for high-speed image capturing, because exposure time is usually as short as tens of microseconds. Recent advancements in CCD and CMOS sensor technology dramatically increases camera's resolution, sensitivity, and frame/line rate, making line-scan cameras particularly suitable for pavement inspection. A line-scan camera with 2048 pixels and a scanning rate up to 36 kHz is generally sufficient for fast, reliable and high-resolution pavement imaging [8].

Lighting devices are commonly used with line-scan camera to project a transverse beam, usually 12ft (3.65m) wide, so that the corresponding pavement can be illuminated and scanned by the line-scan camera. Halogen or florescent lamps [9][10], LED arrays [11], and laser line projectors [11][13] are three major lighting sources used by APDS systems. Halogen or florescent lamps generate white lights which can only alleviate shadows to a limited extent because a camera filter needs to be used to block the sunlight. Light assemblies with multiple halogen or florescent lamps also require devoted power generators to be installed on the vehicle due to high power consumption. The dimensions of the assemblies are often wider than the vehicle body, thus increase collision risks particularly in urban areas. We designed an LED light array in our lab for an APDS system used by the Texas Department of Transportation (TxDOT) [11]. The light array

contained 400 red LEDs and a cylindrical lens to create a focused beam pattern of size $25\text{mm} \times 3050\text{mm}$ on the pavement. To suppress road object shadows caused by sunlight, a band-pass optical filter that matches the wavelength of the LED was used. However, the light bar can only project a focused beam at a 1-foot distance from the ground, and it is two feet wider than the vehicle body. This increases difficulty in installation and requires caution in operation due to its size. In addition, those high-power LEDs lose efficiency quickly when overheated, making the survey operation much more difficult in summer time.

Recently, laser line projection has become a dominant means for line-scan cameras in APDS systems because of its high energy-efficiency and compact size. A laser projector is an off-the-shelf product that is easy to install and maintain. It can cut down the on-vehicle energy consumption from several kilowatts in incandescent lighting or several hundred watts in LED lighting to around 70 watts. It is also easy to provide adequate cooling through the vehicle air-conditioning system to the laser projector so that it can function reliably at any time. In 2008, we finished another TxDOT-sponsored project by designing a laser-based APDS system [13], which is now being used in a few major U.S. cities for pavement inspection. However, laser illumination has three major unresolved problems that hinder its widespread acceptance for routine operations.

(1) The beam of the high power infrared lasers (class III or class IV) is potentially hazardous to the unprotected human eyes. This is especially true when the survey is conducted in urban areas. Many states mandate that the users must have strict safety measures in order to get an operation license.

(2) The non-uniform power distribution across the lane always causes longitudinal streaks in the images generated by a line-scan camera.

(3) The narrow beam ($<5\text{mm}$) of the laser line frequently misses the alignment with the camera line when the vehicle undergoes severe vibrations, yielding horizontal dark ripples in the image. Streaks and ripples are often difficult to remove and are the major sources of false detections of cracks.

1.3 METHODS OF LINE-SCAN CAMERA CALIBRATION

Line-scan cameras have been commonly used in a wide range of machine vision applications, such as target detection [14][15], transportation inspection [16]-[18], and position and motion estimations [19][20], etc. For the applications which require quantitative measurement, camera parameters must be known through geometric camera calibration. Even though there is a plethora of research in the field of camera calibration, most of them are for 2D area-scan cameras [21]-[23].

There are two major challenges in calibrating a line-scan camera. Firstly, it is difficult to determine the location of a feature point in the pattern coordinate space. And secondly, a suitable mathematical model for line-scan camera is difficult to achieve for a static setup. When a static scene is captured by a line-scan camera, there is only one line of image information that is in the x -direction. The coordinates of feature points on a 2D pattern plan are hard to determine with the 1D information, unless the exact location of the view line can be found on the pattern plane. In order to be able to detect the feature points in the pattern coordinate system, methods have been reported to mount the camera onto a linear stage to allow the camera to scan the pattern surface [24]-[26]. A 2D image can be acquired from the scan motion, through which the feature points in the image coordinates system can be detected in the scanned image, and the point coordinates in the pattern coordinate space are readily available. Thus, the calibration problem can be solved with techniques similar to a 2D camera calibration. On the other hand, to avoid the

dynamic setup with mechanical devices, a 3D calibration pattern is needed [27]. With the increased dimension, more independent feature points that are not restricted on a single surface of the calibration pattern becomes available. With sufficient independent feature points, more constraints can be applied to the calibration problem and camera parameters can be solved. In this report, we present an innovative calibration method for line-scan cameras to estimate the intrinsic parameters, and show that the results are numerically sound. The calibration involves using a stationary planar pattern that consists of repeated vertical and slanted lines, and constructing a 2D calibration framework with 1D data. A feature point reconstruction method is applied to transform the 1D camera calibration problem into the 2D scope. Camera parameters are then solved by using a 2D camera model with constraints unique to 1D geometry.

1.4 METHODS OF IMAGE FUSION

With calibrated camera, geometric distortion in the image caused by camera optics can be corrected. This allows images captured from multiple line-scan cameras to be registered and merged to create new images that contain information from all of the original images, named as image fusion. Depending on the region of interest an image is processed, image fusion can be categorized into pixel, feature and symbol level [33]. According to the source image type, fusion can be done on multi-modality, multi-focus and multi-exposure images [28]-[31]. Pixel level fusion, which combines source images by directly fusing individual pixel value in linear or non-linear way, was used in many applications, due to the fact that it's time efficient and easy to implement, since it uses original source image information directly. However, this method inevitably brings undesired effects such as artifacts or reduced contrast to the processed image. An extensive study and research has been done to propose a new fusion method, multi scale

decomposition (MSD) image fusion [32][34], which can be used through various multi scale transforms to fuse multi-mode, multi-focus or multi-exposure images. One of the most important advantages of MSD image fusion is that its concept fits the real cases since objects in the real world usually present themselves in different scales. In our project, the tree shadow supposed to be eliminated or suppressed, appears in a coarser scale, while the cracks to be detected show in a finer scale. The brief concept of MSD based method is firstly to conduct multi-scale transform of each source image, then create a composite multiple scale representation by setting specific weight value on the decomposed images, and finally reconstruct to generate the fused image by performing the inversed transform on the composite multiple scale representation.

1.5 MOTIVATION AND GOALS

A safe, reliable, and cost-effective APDS system for cracking inspection is still highly desirable for maintaining long-term performance of the U.S. highway network. The specific aim of this study is to develop a novel pavement imaging method using dual line-scan cameras, and to design a new APDS system that can conduct pavement inspection to generate quantitative measurement of pavement cracking at highway speed and in any no-precipitation climate. In order to avoid problems related to safety, stability, and on-vehicle energy consumption requirement, the new system does not use any artificial lighting. Instead, it uses two cameras to capture paired images under natural lighting. The paired camera are configured at different exposure settings, and can complement each other to form clear, shadow-free pavement images. The research effort is divided into four major tasks:

(1) Design and construct a dual line-scan camera system on a survey vehicle that synchronizely outputs pairs of pavement images with complementary details at highway speed.

(2) Investigate new image calibration methods for line-scan cameras that matces geometric positions of the paired images.

(3) Develop a customized image fusion algorithm using the multi-scale decomposition scheme to create a shadow-free fused image out of the paired images.

(4) Create effective seed-tracing algorithms that detects and verifies cracks of various pavements, estimates cracking severity levels, and classifies them according to industrial standards.

The new APDS system will not only eliminate the need for special artificial lighting that is potentially harmful to the unprotected people, but will also substantially reduce installation and maintenance costs and the consumption of on-vehicle energy. The system allows a survey vehicle to drive in normal traffic, decreasing disturbance to the public as well as road hazards to human inspectors during the survey. It will also expedite data collection with high-speed imaging capacity, and improve the objectivity and accuracy of the survey data with high-quality images and enhanced image-processing algorithms.

1.6 STRUCTURES OF THE THESIS

This thesis is organized in four parts, each of which reports our contribution to the research on a new APDS development in detail.

Chapter 2 discusses the framework of a dual camera pavement imaging system. It describes the system setup and the image brightness control mechanism, which is

important for acquiring quality complementary images with highest entropy for effective image fusion.

Chapter 3 is dedicated to line-scan camera calibration. In this chapter, a geometric calibration algorithm of line-scan camera is presented and is discussed in detail in terms of mathematical background and performance evaluation. The calibration method includes four steps: the design of calibration pattern, the estimation of feature points, the construction of feature point correspondences, and the calculation of camera parameters.

Chapter 4 presents the procedures for image fusion and shadow suppression. A contrast pyramid constructed from multi-scale decomposition transform is applied, so that the transformed coefficients of the overexposed and the underexposed images at each scale level can be merged. A new image with all areas well exposed and shadows suppressed can be created by reverse transform of the pyramid. A novel crack detection algorithm which was developed in our previous project is applied to generate quantitative measurement of cracks. The results are briefly discussed in this chapter.

After a series of validation tests, Chapter 5 reports our evaluation results in terms of crack detection accuracy, multiple run repeatability. Our field test with the developed imaging system spans a period of 1.5 years. The evaluations results shows the conditions of the surveyed pavement at each of the time point, as well as the trend of the pavement condition over this time period.

Chapter 6 concludes the study and recommends possible future work for improvement of the system.

Chapter 2: Framework of a Dual Camera Pavement Imaging System

Complementary imaging is an important concept in our project for building a pavement imaging system that is free of artificial lighting. Following a brief review of basic principles underlying complementary imaging technique, we present the system setup, image brightness control mechanism and steps of image acquisition through camera-computer communication.

2.1 COMPLEMENTARY IMAGING

The basic ideas of this complementary imaging method are to use two line-scan cameras to scan the same pavement surface simultaneously with different exposure settings, and to generate two distinct images which can compensate each other. One camera is set in an over-exposure mode to ensure only the image of shadowed regions will be clear and sharp, while the second camera is set in an under-exposure mode for acquiring clear images of sunlit regions. The clear regions of the two images are complementary; together, they can form a complete, shadow-free picture of pavement if they are registered and synthesized properly. The cameras' exposure settings must be adjusted dynamically according to lighting situations, pavement conditions, and vehicle speeds, thus keeping visible regions in the two images always at appropriate brightness and contrast levels. When there are no shadows in the images, the exposures of the cameras are adjusted to a level at which both images are visible and can reinforce each other in crack detection. The large adjustment range of the exposure time and gain of the selected camera permits the system to work on any sunny or cloudy day. Since pavement survey is recommended to be conducted only in daytime (because ride-view imaging and

other inspection instruments require daylight), nighttime operation is not a real problem with this APDS system.

2.2 SYSTEM SETUP

Our proposed system is composed of two 2K line-scan cameras, which are placed side by side at a height of 7 feet from the ground to cover a 12-foot lane (Figure 2.1). The camera's resolution is 2048 pixel/line, giving a spatial resolution of 1.78 mm/pixel at this height. The cameras are synchronized by the same triggering pulse, but are set with different exposure times to target sunlit and shadowed regions, respectively. The system also needs a distance measurement instrument (DMI) and a GPS receiver to generate traveling distance, speed, and GPS coordinates. This information is broadcast to a data collection computer through a DMI/GPS computer in order to create a tag for each image and to make crack data traceable. The traveling speed is also required for calculating the instant line rate of the cameras to ensure a constant interval between two successive scan lines.

The two cameras are connected to the data collection computer through GigE interface. No dedicated image processing cards are needed in this system. The two cameras are wired in series via the 15-pin GPIO connectors for synchronization. Only one camera receives the line rate (trigger pulse) from the computer. This camera is called the primary camera, and the other is the secondary camera.

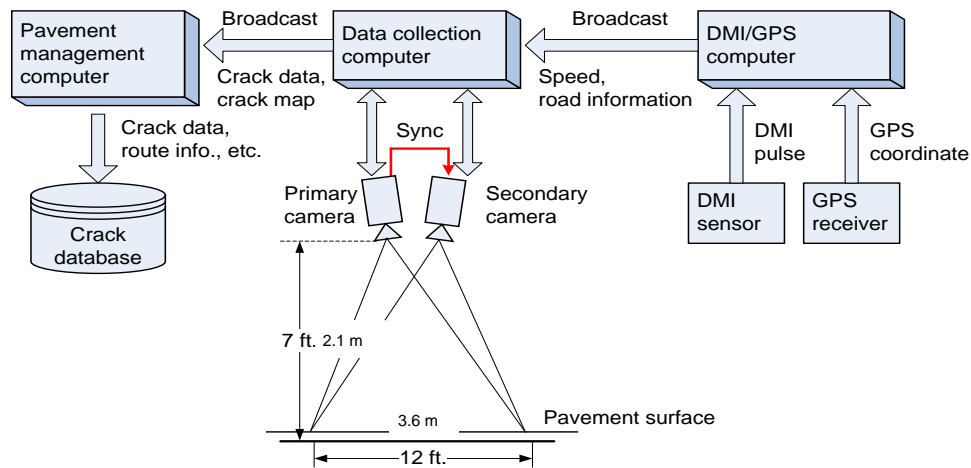


Figure 2.1 Schematic design of the complementary imaging system.

The two cameras should start/stop scanning simultaneously and use the same scan rate. The trigger mode of both cameras must be configured as “External Sync,” which means only the pulse received through the GPIO pin can trigger the camera to capture a line. Once the primary camera receives the new pulse frequency from the computer, its internal pulse generator sends corresponding pulses to the ExSync pins on the two GPIO ports. The pulse frequency is calculated depending on the instant vehicle speed and the pixel resolution.

2.3 IMAGE BRIGHTNESS CONTROL

Image brightness is controlled by dynamically adjusting the exposure time and analog gain. Our image capturing program evaluates each incoming image, and sends update to both cameras to adjust the exposure and gain settings. The primary camera is purposefully configured at overexposed mode, to enhance details of shadowed regions by letting sunlit regions become whiteout in the image. The secondary camera is configured to be underexposed to make pavement details in the sunlit regions visible, while letting the shadows to be blackout. For both of the cameras, the analog gain is always kept at the

lowest possible level to avoid noise being amplified. The key to the success of creating pairs of complementary images lies in the dynamic adjustment of the overexposure and underexposure time for the cameras.

The evaluation of the histogram of the current frame determines the exposure adjustment for the next frame. A well-balanced image frame is supposed to have a histogram which covers a wide range of grayscales from black to white. The visible regions in pictures in our system are regarded as Regions of Interest (ROI). Therefore, pixels of whiteout and blackout regions are excluded when generating the histograms. The offset between the overall brightness of ROI and the central gray level is used to determine whether the grayscale of the ROI is well exposed. The central gray level is set to be 128 in our project. To calculate and adjust the offset, the accumulated histogram of the pixels in the ROI is calculated.

Let $H=\{H_i \mid i=0,...,255\}$ be the histogram. H_i is the percentage of pixels at gray level i . Let $A=\{A_i \mid i=0,...,255\}$ be the accumulated histogram, where $A_0 = H_0$, and $A_i = A_{i-1} + H_i$ ($i > 0$). A_i is the percentage of the total number of pixels whose grayscale is lower than i . Assume α is a desired cutoff of A , and C_α is the grayscale where $A = \alpha$. We can calculate the overall brightness G_o of the image by using:

$$G_o = 0.5 \times (C_{0.5} + 0.5 \times (C_{0.3} + C_{0.7})). \quad (2.1)$$

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 reasonable G_o . Using the difference between G_o and the central grayscale, the exposure can be adjusted. The overall brightness of the image will then be expected to move toward the central level, so that ROI will less likely be whiteout or blackout.

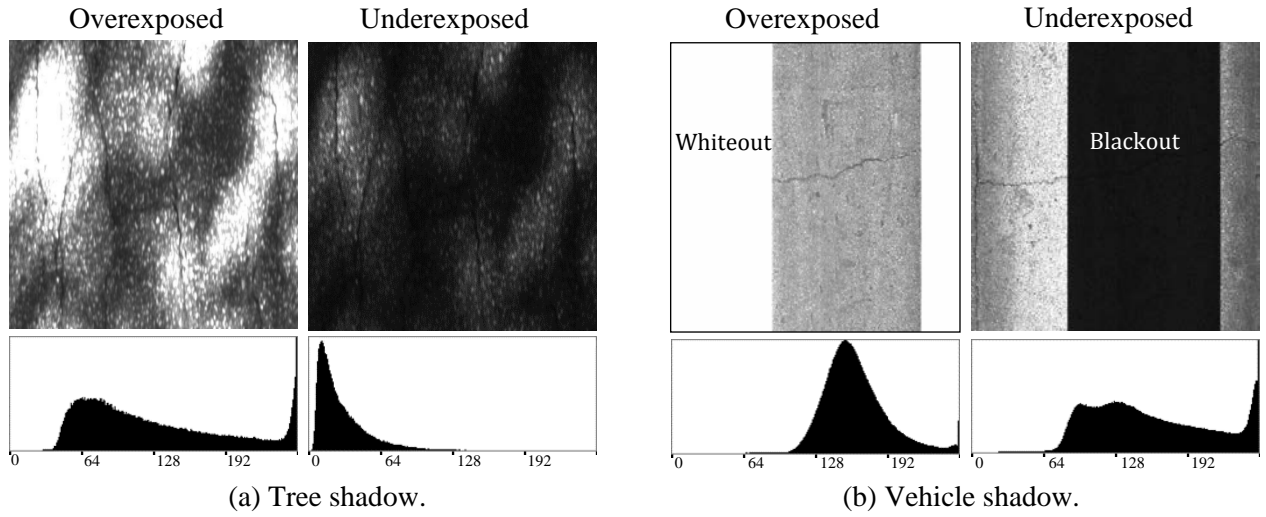


Figure 2.2 Complementary images and their grayscale histograms. Pixels with extreme values (0 and 255) are excluded.

Two pairs of complementary images are presented in Figure 2.2, which are captured by the dual line-scan cameras in the field test. Figure 2.2(a) contains tree shadows, and Figure 2.2(b) has vertical shadow caused by vehicle driving against sunlight. The ROI in overexposed (left) and underexposed (right) images makes cracks in shadowed and sunlit regions visible, respectively. The tree shadowed image in particular is used in this thesis as an example to explain the image-processing algorithm in the following chapters.

2.4 IMAGE ACQUISITION

Our image acquisition process is designed to be multi-threaded. Each camera is associated with separate threads to handle the streaming of images simultaneously. The DMI/GPS computer broadcasts the vehicle speed at a small time interval. If the data collection computer recognizes that there is a speed change, it will convert the new speed into the corresponding line rate, and send it to the primary camera to alter the pulse frequency. This will change the scanning rate of both cameras since they are

synchronized, and the change of scanning rate in the camera will be effective at the same time step.

The image acquisition thread maintains a frame buffer pool which is exclusively available to its associated camera. It repeatedly checks the camera status to see whether a frame, a given number of scan lines, is finished. When it is done, the acquisition thread copies the frame into a buffer at the end of the pool, creating a job queue to save image. When queued frames in the pool reaches a predefined limit, the image saving thread will be initiated to dump the queued images to the hard disk at once.

The saving thread runs in parallel to the acquisition thread, so that they won't interrupt each other, allowing the camera to scan the pavement without skipping. Therefore, consecutive images can be stitched seamlessly. Queuing images in a pool reduces the frequency of accessing the hard disk, and saves time for the camera to operate at high scan rates. The system will be able to scan and save real-time pavement images at traveling speeds up to 112 km/h (70 mph) without skipping.

Chapter 3: Geometric Calibration of Line-scan Camera

3.1 INTRODUCTION

A line-scan camera is an imaging device with only one row of photonic sensors. It captures a 2D image by accumulating scanned lines over time. Similar to the application of a 2D camera, a line-scan camera has to be calibrated in order to be used for quantitative measurement purpose. The camera calibration procedure determines the intrinsic and extrinsic parameters. Intrinsic parameters are the parameters that characterize the optical, geometric, and digital characteristics of the cameras. These include the focal length, the position of optical center, and the geometric distortion coefficients introduced by the optics. Extrinsic parameters are the parameters that uniquely identify the transformation between the unknown camera coordinate system and the known world coordinate system. These include parameters that define the location and orientation of the camera with respect to the world coordinate system.

In 2D camera calibration, the mapping of a feature point from its pattern coordinate space to the image coordinate space is related by the homography matrix, and the pair of the coordinates in both spaces is referred to as a correspondence. The numerical solution of a homography is achieved by applying numbers of feature point correspondences. Among almost all of the calibration methods, the estimation of the homography is the first step in solving the intrinsic and extrinsic camera parameters. The line-scan camera calibration problem can be approached by following the similar steps. However, there are two major challenges. The first one is the difficulty in determining the location of a feature point in the pattern coordinate space, and the second is the scarcity of mathematical models suitable for the line-scan camera in a static setup. When a static scene is captured by a line-scan camera, contrasts exist only in the x -direction. With 1D information, the corresponding feature points on a 2D pattern plane are hard to be

determined unless the exact location of the field-of-view can be found on the pattern plane.

To solve this problem, Horaud, *et al.* [24] proposed a pattern of four coplanar straight lines, three of which were parallel and the fourth intercepted the three with an arbitrary angle. The line-scan camera saw four points from this pattern. By looking at the ratio of the distances between each pair of adjacent points, the location of each point could be estimated. A linear stage was used to drive the pattern in both Y - and Z -axis, so that sufficient independent feature points can be found. Because the movements were in known directions with pre-determined speed, the shifts of the camera coordinate system can be translated into the offsets of coordinates of feature point in the pattern coordinate space, leaving the homography unchanged. Later, Luna, *et al.* [27], extended this technique by utilizing two sets of parallel lines (vertical lines and slanted lines) interconnected in a zig-zag configuration. To avoid the use of a mechanical device, a second tier of the same line sets were added to the pattern such that the pattern provided feature points in two different elevations. Sufficient independent feature points can be obtained in one shot. Although these patterns seem to be viable in line-scan camera calibration given that no additional device is needed to assist feature points estimation, they both rely on the assumption that camera lens distortion is small enough to be ignored, and constant within each repeated parallel line.

Besides these static setups, calibration based on the dynamic imaging model with 2D chessboard pattern were reported [25][26]. 2D images were captured by scanning the pattern with a line-scan camera attached to a linear stage. The feature points (chessboard corners) in the image coordinate space were detected from the scanned image. In this type of calibration, the scanning rate of the camera must be synchronized with the speed of the stage. A virtual 3D homography between the pattern coordinate space and the 2D

image coordinate space was adopted. This virtual homography was the result of the swapping motion of camera sensor array due to the scan.

In this chapter, we present a new calibration approach that only requires a line-scan camera to view a stationary planar pattern from two orientations. The multiple observations can be done by adjusting either the camera or the pattern. The relative positions between the camera and the pattern do not have to be known before the calibration. The homography matrix for each pattern observation is first estimated from feature point correspondences reconstructed from orthogonal projection in both the image plane and pattern plane, and then the initial camera parameters are obtained by solving the equations derived from the homography matrix under the orthonormal constraints among each column. Finally, these parameters are optimized in a non-linear fashion which minimizes the reprojection error. The calibration pattern in our method is essentially similar to the ones reported in [24] and [27]. However, our method differs from [24] in that repeated line patterns are used so that mechanical movement is not necessary to capture enough feature points for homograph estimation. Our method also differs from [27] in that a 2D calibration pattern instead of a 3D pattern is sufficient for the task. Compared with the approaches presented in [25] and [26], this technique is considerably easier to be implemented because there is no high precision motion or synchronization involved.

3.2 CAMERA MODEL AND IMAGING GEOMETRY

We start with the review of the classic geometry in 2D camera model, followed by the discussion of unique and important constraints in 1D imaging.

3.2.1 Preliminary

Without lens distortion, a 3D projection modeled by a pinhole camera is described by

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{M} [\mathbf{R} \quad \mathbf{t}] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3.1)$$

where $[X, Y, Z]^T$ are the coordinates of a 3D point in the pattern coordinate space, $[u, v]^T$ are the coordinates of the projection on the image in pixels, and s is an arbitrary scale factor. $[\mathbf{R} \quad \mathbf{t}]$ is a 3×4 matrix of extrinsic parameters, in which \mathbf{R} is the rotation matrix defined on Euler angles (α, β, γ) and constructed with Rodrigues's rotation formula, \mathbf{t} is the translation vector between the pattern coordinate system and the image coordinate system. \mathbf{M} is called the camera intrinsic matrix, and is defined by

$$\mathbf{M} = \begin{bmatrix} f_x & k & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (3.2)$$

in which f_x and f_y are the horizontal and vertical focal lengths in pixels, k is the skewness of the axes in the image plane, and (u_0, v_0) are the coordinates of the principal point, or optical center, on the image.

If a planar pattern is used in the calibration, then all the feature points can be defined on a single plane. Assume the Z -coordinate of each of these feature points to be 0, and let \mathbf{r}_i be the i th column of rotation matrix \mathbf{R} , equation (3.1) can be rewritten as

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{M} [\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{t}] \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}. \quad (3.3)$$

Equation (3.3) defines a homogeneous transformation of a point from the pattern coordinate space to the image coordinate space. Therefore, the point is related by a homography \mathbf{H} :

$$\mathbf{H} = \mathbf{M} [\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{t}]. \quad (3.4)$$

3.2.2 Estimation of Homography for 1D Line-scan Camera

A complete homography to map a world homogeneous coordinate $[X \ Y \ Z \ 1]^T$ to an image homogeneous coordinate $[u \ v \ 1]^T$ is a 3×4 matrix with 11 degrees of freedom. These degrees of freedom, or parameters, can be broken down into two categories: 5 intrinsic and 6 extrinsic parameters. The 6 extrinsic parameters relate the camera orientation to a world coordinate space (or the pattern coordinate space) and consist of 3 rotations and 3 translations. In order to estimate the homography with 11 degrees of freedom, theoretically 6 pairs of independent correspondences are needed, because each pair generates two formulas according to equation (3.3), and a total of 12 equations are sufficient to solve 11 parameters. Here, a correspondence can be presented in a vector form $[X \ Y \ Z \ u \ v]^T$, and an independent correspondence means that none of these vectors is a linear combination of any other two vectors. In the perspective model of the camera, a cross-ratio [37] in the world coordinate space also applies to the image coordinate space. Thus, if there is a point that is collinear with other two points in the pattern coordinate space, the equations based on this point correspondence are redundant, and will not add any new constraints to the solution. In the case of 1D line-scan camera, one can assume $f_y = 0$, $v_0 = 0$, and $k = 0$, then there are 8 parameters left to be solved, thus 4 pairs of independent point correspondences are needed to estimate the homography of a line-scan camera. However, since these 4 imaged feature points will be collinear on image plane due to the 1D geometry of the sensor array, their corresponded feature points in the world coordinate space must not be collinear to avoid redundancy. Otherwise, the correspondence defined by the third and fourth point is over constrained because their coordinates in pattern coordinate space can be estimated by back-projecting rays from the focal point,

passing through their image points, and intercepting the line-of-view, as is shown in **Figure 3.1**.

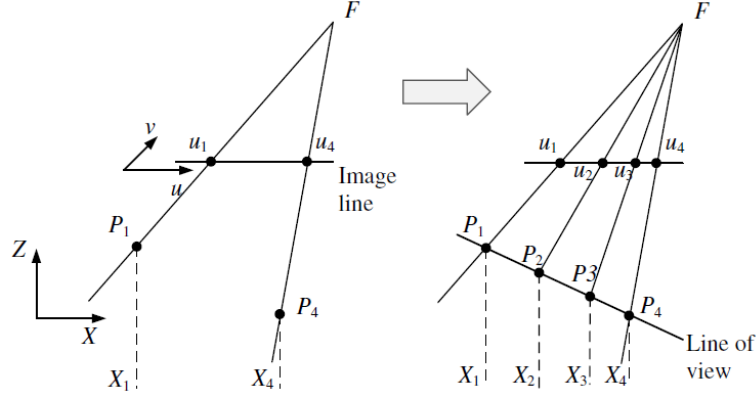


Figure 3.1 More than two feature points that are collinear in the pattern coordinate space do not add new constraints in solving the homography. P_2 and P_3 can be estimated by extending the ray from the focal point F to the line-of-view.

3.3 CALIBRATION PATTERN AND FEATURE POINT CORRESPONDENCES

3.3.1 Calibration Pattern Design

The calibration pattern we used is a planar pattern of two sets of mutually parallel lines. The first set consists of 58 evenly spaced vertical parallel lines $\{P_i\} = \{P_1, P_2, \dots, P_N\}$ ($N = 58$), and the second set consists of 57 slanted lines $\{D_i\} = \{D_1, D_2, \dots, D_{N-1}\}$, which are placed between the vertical lines. Figure 3.2 displays two of the repeated units to show how these lines are constructed.

All these lines are defined on $Z = 0$ plane. The origin of the pattern coordinate space is placed at the intersection of lines P_1 and D_1 . The direction of the x axis is perpendicular to the vertical parallel line set $\{P_i\}$, and the direction of Y axis is parallel to $\{P_i\}$. The pattern coordinate space is defined in such a way that the line sets $\{P_i\}$ and $\{D_i\}$ can be expressed in two simple formulas:

$$P_i: X = ih. \quad (3.5)$$

and

$$D_i: Y = \frac{l}{h}X - (i - 1)l. \quad (3.6)$$

in which l is the length of each vertical line in set $\{P_i\}$, and h is the distance between each adjacent vertical lines, i denotes the i -th line in the set.

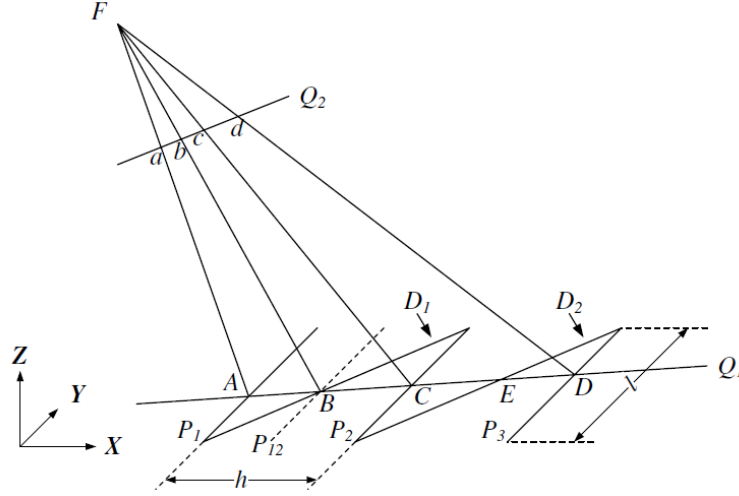


Figure 3.2 The calibration pattern being viewed by a line-scan camera. Only two repeated line structures are shown.

3.3.2 Feature Point Estimation

Figure 3.2 illustrates the geometrical configuration of the calibration pattern being viewed by a line-scan camera. In the figure, F is the focal point of the camera, line Q_2 is the imaging line. The plane determined by F and Q_2 is the viewing plane, which intersects the pattern at line Q_1 . Q_1 indicates the line-of-view of the camera. When the camera looks at this pattern, it sees a series of dots, which are the intersection points of the view line Q_1 and the pattern lines. The task of constructing point correspondences is to determine the coordinates of each feature point in both pattern coordinate space and image coordinate space.

For each repeated unit, let A, B, C and D denote the four feature points seen by the camera, *i.e.*, the intersections of the view line with $\{P_i\}$ and $\{D_i\}$, and a, b, c , and d are their images, respectively. Since points A, B, C and D are collinear and the rays FA, FB, FC and FD are concurrent, the cross-ratio of these points is thus defined as

$$r_{ABCD} = (\overline{CA}/\overline{CB})/(\overline{DA}/\overline{DB}). \quad (3.7)$$

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

$$r_{abcd} = (\overline{ca}/\overline{cb})/(\overline{da}/\overline{db}) = r_{ABCD}. \quad (3.8)$$

Since the coordinates of a, b, c , and d can be readily detected from a line-scanned image, one can compute the cross-ratio $r = r_{abcd}$ from these coordinates, and further determine the coordinates of A, B, C and D as follows. Assume α is the ratio of \overline{BA} to \overline{CA} . With r being available and $\overline{DA} = 2 \times \overline{CA}$, the relationship between α and r can be solved

$$\alpha = \frac{2r - 2}{2r - 1}. \quad (3.9)$$

Therefore, the x -coordinate of point B can be calculated as $X_B = \alpha h$. Since point B is also on the line D_1 , its y -coordinate can also be calculated by Equation (3.6). Similarly, the intersection point E of the view line to D_2 can also be calculated by introducing the fourth vertical line P_4 . With the coordinates of both B and E 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 D can be determined as well.

3.3.3 Feature Point Correspondences

From Section 3.2.2, we have concluded that to solve an 8 degrees of freedom 3D homography, 4 pairs of independent correspondences are needed. Geometrically, this rule requires that there must be 4 feature points that are not collinear in the pattern coordinate space. However, a planar pattern can only provide 2 independent feature points when the pattern is seen by a line-scan camera. In our calibration method, we solve the homography estimation by projecting the original feature points onto the orthogonal direction of the view line on the pattern plane.

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 u_0 , which is initialized as half the width of the image. By doing this, it is equivalent to convert the 1D image into 2D in which the contrast in the second dimension is exactly duplicated from the first dimension. At this point, our 1D camera calibration problem becomes a 2D problem, in which some constraints on the intrinsic parameters are applied: (i) the focal length must be the same in both vertical and horizontal direction, i.e., $f_x = f_y$; (ii) the optical center in both vertical and horizontal direction must be the same, i.e., $u_0 = v_0$. The camera we used in our experiment has a resolution of 2048 pixels. To convert the 1D feature points into 2D, we define these 2D feature points within a 2048×2048 image. Assuming the estimated optical center is at (u_0, v_0) , the i th feature point in the line-scanned image will be mapped to a point on a 2D image with coordinate (u_i, v_0) , in which u_i is the measured feature location from the line-scanned image. Then its projected counterpart can be computed as

$$\begin{bmatrix} u_i' \\ v_i \end{bmatrix} = \mathbf{R}_{\text{proj}} \left(\begin{bmatrix} u_i \\ v_0 \end{bmatrix} - \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} \right) + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix}, \quad (3.10)$$

in which $u'_i = v_0$, \mathbf{R}_{proj} is the rotation matrix defined by Rodrigues equation. Since rotation only exists on the xy -plane, it can be simplified to be a 2×2 matrix. When a rotation angle of $\pi/2$ is applied,

$$\mathbf{R}_{\text{proj}} = \begin{bmatrix} \cos \frac{\pi}{2} & -\sin \frac{\pi}{2} \\ \sin \frac{\pi}{2} & \cos \frac{\pi}{2} \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \quad (3.11)$$

The same technique is applied onto the feature points in pattern coordinate. Figure 3.3(a) shows an example of all the feature points plotted on the pattern coordinate space. In this figure, the horizontally arranged point array indicates the position of the view line, and the vertically arranged point array indicates the positions of the reconstructed feature points. The intersection of these two point arrays is determined by back-projecting the optical center (u_0, v_0) onto the pattern surface. The geometrical illustration of the feature points being viewed by a line-scan camera is shown in Figure 3.3(b).

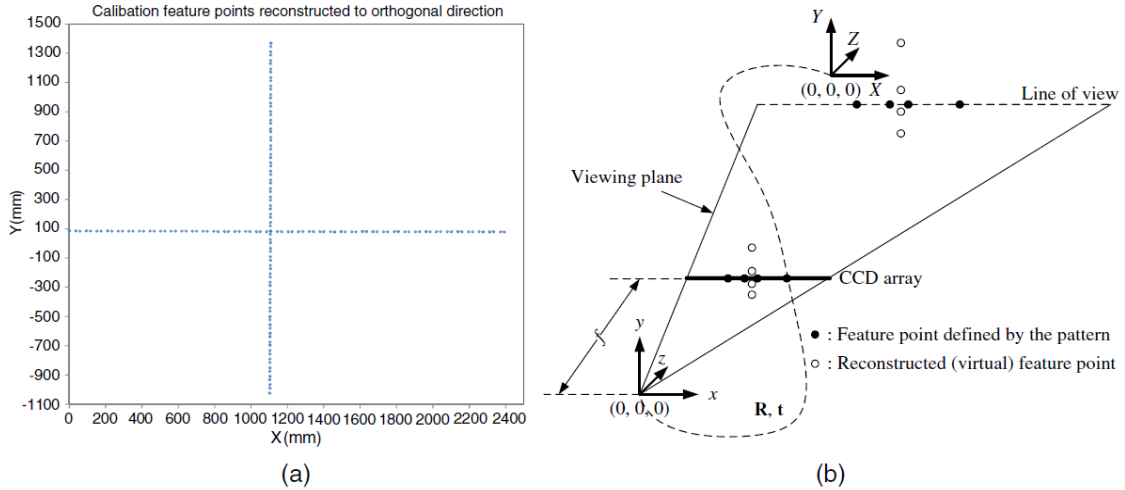


Figure 3.3 Reconstruction of calibration feature points. (a) Feature points in the pattern coordinate space with reconstructed points in the orthogonal (vertical) direction. (b) Geometrical illustration of the feature points viewed by a line-scan camera. Rotation and translation from the pattern coordinate space to the camera coordinate space are related by \mathbf{R} and \mathbf{t} .

3.4 SOLVING THE CAMERA PARAMETERS

Due to the nonlinear mapping introduced by lens distortion in the camera model, a solution of all camera parameters cannot be estimated directly in a least square fashion. Generally, given a set of feature point correspondences, camera parameters are optimized iteratively [21][22][25][26]. However, the feature point estimation described in Section 3.3.3 requires the intrinsic parameters, i.e. the calibration output, to construct point correspondences. This condition can be solved by nesting the iterative nonlinear optimization process within an iterative estimation of feature points. The combined calibration process thus consists of three steps: (1) initialize the optical center at the image center and estimate the camera parameters from a closed form solution; (2) apply nonlinear optimization of camera parameters based on the initial estimates from (1); and (3) refine the positions of feature points based on optimized camera parameters. Steps 2 and 3 are repeated until the convergence is reached.

3.4.1 Closed-form Solution

In the closed-form solution, the homography is estimated first. There are various algorithms proposed for homography estimation. We looked at the Direct Linear Transform (DLT) [38] which is simple yet effective when combined with geometric distance cost function. Once the homography is computed, the intrinsic and extrinsic parameters will be available by solving Equation (3.4).

Denote $\mathbf{H} = [\mathbf{h}_1 \ \mathbf{h}_2 \ \mathbf{h}_3]$ for homography that is estimated for each observation of the calibration pattern, \mathbf{h}_j is the j th column of \mathbf{H} . Equation (3.4) can be rewritten as

$$[\mathbf{h}_1 \ \mathbf{h}_2 \ \mathbf{h}_3] = \mathbf{M}[\mathbf{r}_1 \ \mathbf{r}_2 \ \mathbf{t}]. \quad (3.12)$$

Because \mathbf{r}_1 and \mathbf{r}_2 are orthonormal, we have

$$\begin{cases} \mathbf{r}_1 \cdot \mathbf{r}_2 = 0 \\ \|\mathbf{r}_1\| = \|\mathbf{r}_2\| \end{cases} \Rightarrow \begin{cases} \mathbf{h}_1^T \mathbf{M}^{-T} \mathbf{M}^{-1} \mathbf{h}_2 = 0 \\ \mathbf{h}_1^T \mathbf{M}^{-T} \mathbf{M}^{-1} \mathbf{h}_1 = \mathbf{h}_2^T \mathbf{M}^{-T} \mathbf{M}^{-1} \mathbf{h}_2 \end{cases} \quad (3.13)$$

Since the homography has eight degrees of freedom and there are six extrinsic parameters, two constraints in Equation (3.13) are left to be applied on intrinsic parameters. Let $\mathbf{B} = \mathbf{M}^{-T} \mathbf{M}^{-1}$, thus $\mathbf{B}^T = (\mathbf{M}^{-T} \mathbf{M}^{-1})^T \equiv \mathbf{B}$. So \mathbf{B} is symmetric and can be defined by a six-element vector $[B_{11} \ B_{12} \ B_{22} \ B_{13} \ B_{23} \ B_{33}]^T$, which represents the upper diagonal of \mathbf{B} [22]. Furthermore, since part of our feature point correspondences are reconstructed through orthogonal projection, we have $f_x=f_y$, $u_0=v_0$, and $k=0$. Thus, we obtain $B_{11}=B_{22}$, $B_{12}=0$, and $B_{13}=B_{23}$, only three numbers are required to uniquely define \mathbf{B} . Then we have $\mathbf{b} = [B_{11} \ B_{13} \ B_{33}]^T$ and,

$$\mathbf{h}_i^T \mathbf{B} \mathbf{h}_j = \mathbf{v}_{ij}^T \mathbf{b}, \quad (3.14)$$

with

$$\mathbf{v}_{ij} = [h_{i1}h_{j1} + h_{i2}h_{j2} \quad h_{i3}(h_{j1} + h_{j2}) + (h_{i1} + h_{i2})h_{j3} \quad h_{i3}h_{j3}]^T. \quad (3.55)$$

Therefore, two constraints defined in Equation (3.13) can be transformed into two linear equations of \mathbf{b} :

$$\begin{bmatrix} \mathbf{v}_{12}^T \\ (\mathbf{v}_{11} - \mathbf{v}_{22})^T \end{bmatrix} \mathbf{b} = \mathbf{0}. \quad (3.16)$$

When n images of the calibration pattern are taken, by stacking n equations of (3.16) on top of each other, a total of $2n$ equations are available to solve \mathbf{b} by applying the Levenberg-Marquardt algorithm. Once \mathbf{b} is estimated, we can compute the intrinsic parameters as

$$\begin{aligned} f_x = f_y &= \sqrt{\frac{\lambda}{B_{11}}} \\ u_0 = v_0 &= -\frac{B_{13}f_x^2}{\lambda} \\ k &= 0 \end{aligned}$$

with $\lambda = B_{33} - 2B_{13}^2/B_{11}$. Then camera matrix \mathbf{M} can be constructed according to Equation (3.2). Once \mathbf{M} is known, the extrinsic parameters for each observation can be readily computed as

$$\begin{aligned}\mathbf{r}_1 &= \lambda' \mathbf{M}^{-1} \mathbf{h}_1 \\ \mathbf{r}_2 &= \lambda' \mathbf{M}^{-1} \mathbf{h}_2 \\ \mathbf{r}_3 &= \mathbf{r}_1 \times \mathbf{r}_2 \\ \mathbf{t} &= \lambda' \mathbf{M}^{-1} \mathbf{h}_3\end{aligned}$$

with $\lambda' = 1/\|\mathbf{M}^{-1} \mathbf{h}_1\| = 1/\|\mathbf{M}^{-1} \mathbf{h}_2\|$.

3.4.2 Nonlinear Optimization

After the initial estimation of the intrinsic and extrinsic parameters has been worked out through closed-form solution, an optimization procedure can be applied to minimize the reprojection error in a maximum likelihood fashion.

To include the lens distortion, the projection of a feature point defined in the pattern coordinate space needs to be adjusted to reflect the geometrical distortion. Because the image captured by a line-scan camera only contains 1D information, tangential distortion actually does not exist, so lens distortion is essential radial distortion only. Based on literature [21][39], although the distortion can spread over more than 20 orders of magnitude, it is dominated by the first order. The non-linear optimization procedure follows Zhang's method [22] by applying Levenberg-Marquardt algorithm to minimize the reprojection error.

3.5 EXPERIMENTS AND DISCUSSIONS

The proposed calibration method is implemented within MATLAB®, and has been tested on both synthetic data and real images. The synthetic data are used to test the performance of calibration considering two factors: (1) the tolerance to the noise in points

detected from images, and (2) the effect of pattern orientations on the calibrated results. The calibration's practicability is verified by real images. Preliminary results of this work were originally published in [40].

3.5.1 Synthetic Data

The accuracy and robustness of the calibration method is tested on a set of synthetic feature point correspondences. The simulated intrinsic parameters are $f_x = f_y = 1260$ pixels, $k = 0$, and $u_0 = v_0 = 1020$ pixels. The image resolution is 2048×2048 pixels. The simulated pattern is an array of 41 points evenly distributed along a straight line. The distance between each neighboring points is set to be 60 mm, thus the point array covers a width of 2.4 m. The orientation and displacement of the pattern with respect to the camera are represented by a 3D rotation vector \mathbf{r} and a translation vector \mathbf{t} .

3.5.1.1 Noise test

In this test, Gaussian noise with 0 mean and σ standard deviation is added to the image points. Two observations are taken for each calibration with $\mathbf{r} = [180^\circ, 0, 0]$, $\mathbf{t} = [0, 0, 1800]$, and $\mathbf{r} = [180^\circ, 5^\circ, 0]$, $\mathbf{t} = [0, 0, 2000]$, respectively. In the first observation, the pattern plane is parallel with the imaging plane with the surface normal pointing to reverse direction. In the second observation, the pattern plane makes a 5° angle around the Y -axis with the imaging plane. The estimated camera parameters are compared with the ground truth. The level of noise varies from 0.1 to 1 pixel in 0.1-pixel increments. For each noise level, we conduct 60 independent trials and take the average residual errors as the results. As is shown in Figure 3.4(a)-(b), both the reprojection error and the errors of each intrinsic parameter increase when the noise level increases. While the reprojection error increases almost linearly, the errors of intrinsic parameters show slight fluctuations.

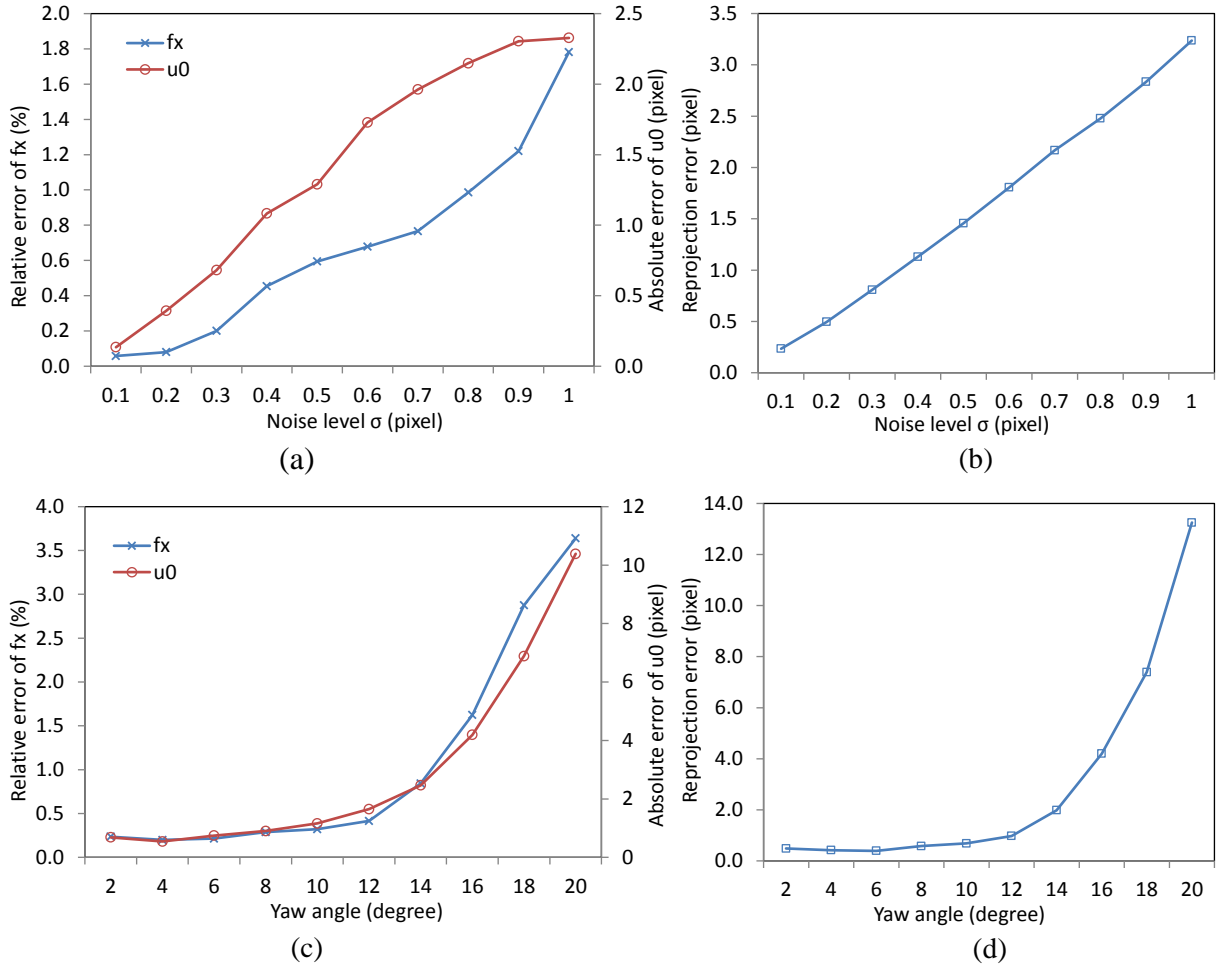


Figure 3.4 Calibration errors from simulations. (a) Errors in intrinsic parameters with respect to noise; (b) Reprojection error with respect to noise; (c) Errors in intrinsic parameters with respect to rotation angle; (d) Reprojection error with respect to rotation angle.

This is the result of global optimization, which may favor some parameters over the others at certain noise level.

3.5.1.2 Orientation test

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 is preserved in the reconstructed points. This simulation is to find out if a rotation in one dimension is significant to compromise the calibration results.

Two observations are taken for each calibration with $\mathbf{r} = [180^\circ, 0, 0]$, $\mathbf{t} = [0, 0, 1800]$, and $\mathbf{r} = [180^\circ, \theta^\circ, 0]$, $\mathbf{t} = [0, 0, 2000]$, respectively. The rotation angle varies from 2° to 20° at a 2° interval. Gaussian noise with 0 mean and 0.3 pixel standard deviation is added to image points. Sixty independent trails are conducted for each orientation angle. The results are shown in Figure 3.4(c)-(d). It is discovered that at orientation angle about 5° the minimum error is observed. This can be explained that the algorithm needs the variation of depths in the feature points to accurately estimate the focal length. The greater the orientation angle, the larger the variations of feather point depth along the Z-axis in the camera coordinate space. At around 5° the duplicated distribution of the reconstructed feature points are not significant enough to generate errors in calibration, while the depth variation in feature points are sufficient for the estimation of focal length. It can also be observed that when the rotation angle gets greater than 10° , the reprojection error increases almost exponentially. This indicates that rotation at this level or greater can make the calibration unstable. At 12° rotation, we start to observe calibration failures. At 20° rotation, about 65% of calibrations fail.

3.5.2 Real Images

To capture line-scanned image, a Dalsa[®] Spyder III line-scan camera is used. The camera is mounted on a heavy-duty tripod so that the orientation of the camera can be adjusted. A Nikon[®] 20 mm fixed focal length F-mount lens is attached to the camera though an F-mount adapter. The captured images are 2048 pixels wide. The calibration pattern is printed by a poster printer. Our pattern consists of 57 repeated units as described in Section 3.3.1, with $h = 60$ mm and $l = 220$ mm. The whole pattern covers a

width of about 3.4 m. The printed pattern is then attached to two laminated wood boxes that are connected together to create a uniformed flat surface. The camera and pattern are mounted at the same elevation, so that the camera has almost 0 pitch angle. A total of 12 calibrations are conducted. In each calibration, two pictures are taken at two different poses. In the first pose, the camera looks at the pattern almost perpendicularly. In the second pose, a randomly selected $+5^\circ$ or -5° yaw angle is applied by either turning the camera to its right or to its left. The yaw angle does not need to be precisely controlled. The mean and standard deviations of the 12 calibrations are shown in Table 3.1.

In the captured picture, the image of each feature point is a vertical line, and its horizontal position is calculated by averaging the left and right edges of the line. Edges are detected up to 0.1 pixel resolution using sub-pixel interpolation. Feature point correspondences are computed according to Section 3.3.3. The iterative process of refining the optical center to achieve an accurate estimation of feature points in the pattern coordinate space takes 3 to 5 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 of the 12 calibration is 0.46 pixels with standard deviation of 0.066.

Table 3.1 Calibration results for intrinsic parameters and reprojection errors

	f_x	u_0	k_1	k_2	reprojection err.
Mean	962.18	1021.46	-0.036	0.015	0.46
STD	1.913	0.494	5.41E-04	3.25E-04	0.066

Three-pose calibrations were also investigated in our experiments. Images used in these three-pose calibration are the perpendicular observation, $+5^\circ$ observation and -5° observation. Our existing 12 sets of calibration images contributed to 6 three-pose

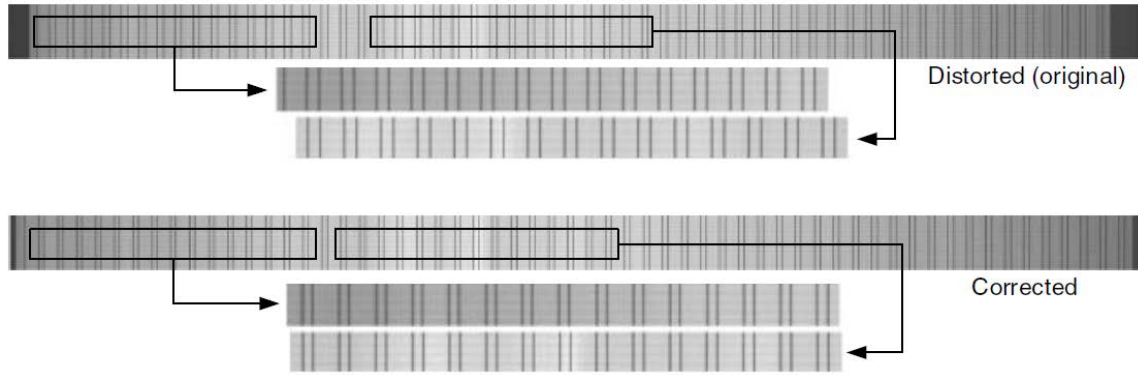


Figure 3.5 Examples of distortion correction on pattern images. Top part: original image with lens distortion. Pattern lines are unevenly spaced; Bottom part: distortion corrected. Pattern lines are evenly spaced and can be aligned.

calibrations. The mean reprojection error was 0.81 pixels with standard deviation of 0.064, compared to 0.46 pixels in two-pose calibration. This suggested a degradation in calibration when including more angled observations. The source of the degraded performance can be traced to the error in 1D to 2D conversion in constructing feature point correspondences, in which angled observation introduces more error than perpendicular observation. Adding a new pose into the calibration with more errors in feature point correspondences ultimately decreases the accuracy of the results, thus two observations with sufficient feature points in each are preferred in our calibration method.

To find out if the lens distortion coefficients are effective in correcting the radial distortion in the images, we performed distortion correction with obtained calibration results. Figure 3.5 shows the results of the correction. The image was captured with the camera looking at our calibration pattern in a perpendicular orientation. Each vertical line in the image indicates a feature point on the pattern seen by the camera. If a perfect lens is considered and radial distortion is zero, 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. The top part of Figure 3.5 shows the original line-scanned image with radial distortion, from which the spaces between line segments on the sides are narrowed.

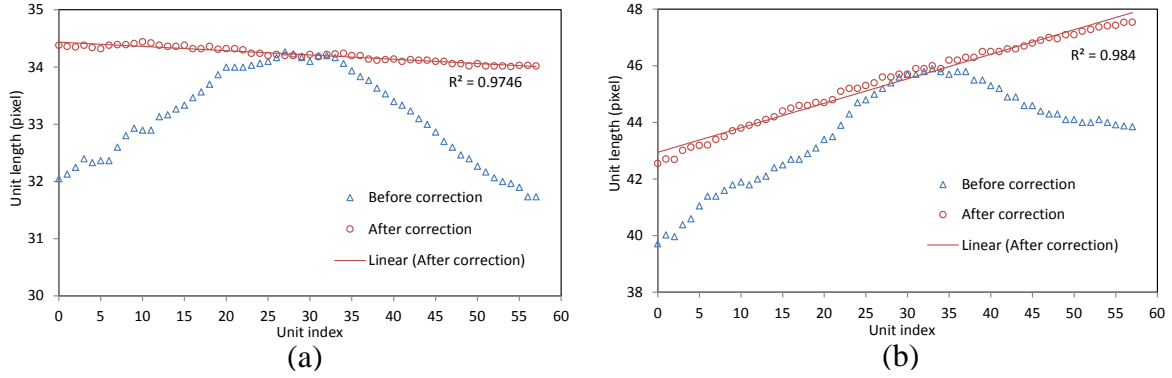


Figure 3.6 Plots of unit lengths before and after radial distortion correction. (a) Measurements taken from the observation at a perpendicular pose; (b) Measurements taken from the observation with yaw angle.

The bottom part of Figure 3.5 shows the result after distortion is removed. All the line segments can be perfectly aligned, regardless where in the image they are sampled.

To further verify the quality of distortion correction, we looked at the space within each repeated pattern unit, referred to as the unit length. Figure 3.6 shows the unit lengths measured before and after distortion correction. Figure 3.6(a) displays the correction results from a picture that was taken at a perpendicular camera orientation, while Figure 3.6(b) displays the results from a picture taken with a yaw angle. It is clear that radial distortion exists in both pictures, because the unit lengths in the middle of the pictures, where radial distortion is less, are quite consistent before and after correction. The effectiveness of the correction is evaluated by the linearity of the plots of unit lengths with respect to their locations on the image. We fitted a line to the graph after correction. The R^2 values of the two observations are found to be 0.975 and 0.985, respectively, which indicate good linearity for both cases.

The effectiveness of this calibration technique was demonstrated by one pair of our pavement images. The images were captured by two line-scan cameras configured in different exposure settings, and then were fused to generate images free from shadows of roadside objects, such as trees and traffic signs. The line-scan cameras were mounted on

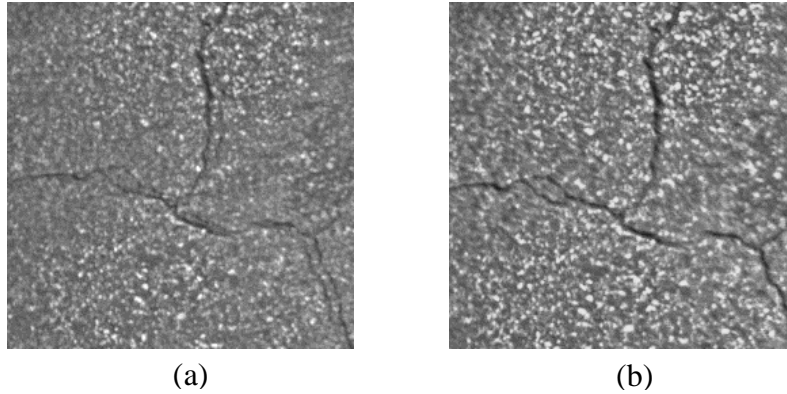


Figure 3.7 Fusion results from images captured by two identical line-scan cameras with two different exposure settings. (a) Fused from images without lens distortion correction; (b) Lens distortion corrected.

one piece of metal plate and were next to each other. Therefore, the image planes of the two cameras were co-planar and the image registration only involved shifts and rotations once the lens distortion was corrected. Figure 3.7 shows the results of the image fusion with and without the lens distortion correction. The image fusion from uncorrected images shows misalignment, because the lens distortion is non-linear and the images cannot be registered correctly. A perfect alignment was achieved when lens distortion was removed by applying the camera calibration method presented in this paper. The distortion free image ensures correct measurement of pavement distress and features, such as the length of a crack, as well as accurate feature localization.

3.5.3 Calibration Errors

As observed in the calibrations with real images, the mean reprojection error we have reached is 0.46 pixels, which indicate that there are other sources that may affect the performance of the calibration method. Here we briefly review some of the most significant error sources.

3.5.3.1 1D to 2D conversion

In our proposed line-scan camera calibration method, we superpose the 1D calibration problem into the 2D scope and treat the line-scan camera as a 2D matrix camera. The calibration problem is solved by an extensively studied camera model with added constraints derived from the unique geometry of a line-scan camera. The feature point correspondences are handled so that the coordinates on the second dimension are exactly duplicated from the first dimension. As a result, the calculated intrinsic parameters are symmetric, i.e., $f_x=f_y$ and $u_0=v_0$. In reality, it is common to have the same focal lengths in both X - and Y -direction for most of the lenses and cameras with square pixels. However, the optical center is usually not symmetric, because these values depend on the relative position between the photonic sensor and the lens. Given only the 1D locations of feature points in the image, it is technically difficult to estimate the optical center in 2D. Since the radial distortion is related to the distance from a current image point to the optical center, with one dimension missing, the amount of distortion cannot be accurately estimated. Although our experiments with real images show that the plot of unit lengths in distortion corrected images is quite linear, this is the result of small vertical offset between the optical center and the image center. The radial distortion along the horizontal direction is largely determined by the horizontal location of each feature point.

3.5.3.2 Radial distortion

Our procedure in estimating the intersection of the view line and the slanted pattern line within each repeated pattern unit is based on the measured image points. It is reasonable to accept that the camera sensor array is aligned in a straight line. However, one can argue that the view line is not necessarily a straight line when the optical center is not on the sensor array and lens distortion is present. As a result, the actual view line

may be slightly curved. Since the distance between each of two parallel lines are small (60 mm in our experiments), we consider the view line segment within each repeated pattern unit to be a straight line, and estimate the feature point coordinates individually for each repeated pattern unit. Furthermore, this strategy is based on the assumption that the amount of distortion is uniform within each pattern unit. This could also introduce error in the calculation of cross ratio within each pattern unit, and ultimately error in the estimated feature point location.

3.5.3.3 Calibration target

It is essential that the coordinates of feature points on the pattern are known with good precision so that their errors are not observable from the image. If this condition is not met, the feature point correspondences are not accurate, and the estimated camera parameters become inaccurate. It is commonly accepted that the relative accuracy in the pattern coordinate space should be better than the accuracy in the image coordinate space [23]. For example, in order to achieve an accuracy of 0.1 pixels as we did in feature point detection from captured pictures, the locations of a feature point in pattern coordinate space should be known to a resolution better than 0.16 mm, given the size of the pattern being about 3.4 m in lateral direction. This is still challenging because the accuracy can be affected by many factors, such as the precision of the printer, the flatness of the pattern mounting surface, any distortion on the printed paper due to the stretch when attaching the paper to the mounting surface, etc.

3.6 CONCLUSION

In this chapter we have presented a line-scan calibration method that is aimed to be applicable for implementation outside a laboratory setting. A planar pattern consisting

of repeated parallel lines is used in the calibration. This pattern allows the estimation of feature points in the pattern coordinate space without mechanical scanning, thus image capturing can be done in a static configuration. A typical calibration can be performed by taking two pictures of the pattern, with one being a perpendicular observation of the pattern and the other at a small angle (5°). The calibration method is tested on both synthetic data and real images. The average reprojection error in calibrating real images is 0.46 pixels. Our results also show that the distribution of unit length in a distortion corrected image is linear along the width of the image. This indicates the distortion coefficients together with the optical center are effective in correcting the radial distortion in line-scanned images.

Chapter 4: Image Fusion and Shadow Suppression

4.1 INTRODUCTION

The image fusion is the process that generates a single image from two or more source images. The fused image contains more information than any of the single source image and presents better expression and understanding of the scene for human vision or machine perception.

Many methods have been proposed to implement the fusion of multi-exposure images. They can be classified into two general categories. One category of the approaches requires determination of the camera response function by an initial camera illumination calibration procedure using multiple exposure images under known imaging condition. Once the response function for a camera is obtained, the high dynamic range radiance map can be calculated from several multiple exposure source images. The second approach is based on image feature measurements. The contrast, gradient, saturation, entropy or other image features may be used directly, or as activity criteria in multiple scale decomposition (MSD) method, to create a “visually pleasing” fused image.

Although these techniques can merge multi-exposure images into a more informative single image in different special application, they cannot be used directly in our system, because all of them fail to eliminate the shadow, which would cause considerable undesirable effect in our subsequent crack detection process if not suppressed or eliminated.

General shadow removal methods use either color information or inter frame context of image sequence to suppress or remove shadow [41]-[44]. In our research, the multi-exposure images are gray scale, and the shadow objects, such as tree, building, and telegraph pole, occurred randomly in an unpredictable mode. So there is neither color information nor inter-frame context to be exploited for removing shadow. However,

because the fused shadow-free image is used only for subsequent crack detection procedure, we need not to recovery the true reflectance image, but only an illumination-consistent image that keeps micro-scale texture information. This makes it possible to apply the image MSD method to implement image fusion and shadow suppression.

4.2 METHODOLOGY

Multi-exposure image fusion is one of major applications in image fusion research. It is used to form a single image by retrieving relevant information from each source image captured at different exposure. There is an abundance of research in this field. Debevec, et al. [28] and Robertson, et al. [29] proposed a response function of image processing from camera illumination calibration using multiple exposure images under certain giving condtions. With the response function known, a high dynamic range image could be generated with serveral source images at different exposure settings. The pixel values of the fused image were proportional to true radiance values of the real scene. Due to the fact that response function often varies with different cameras, the corresponding mothod has to be adjusted to fit each specific camera.

Some other reseachers assigned weight to pixels or regions of different source images according to their image entropy [30] or continuous edge intensity [31], and generated the fused image by calculating and combining the weight average of source images. However, this approach inevitably brings in some discontinuties or other visual artifacts in the fusion process, thus it needs extra steps for image smoothing.

Recently, approaches based on multiple-scale-decomposition (MSD) have been introduced to deal with the mult-exposure fusing problem [32]-[36]. Pyramid transforms [34]-[36] and wavlet transforms [36] are the two major types of MSD methods. The general procedures of the MSD fusion are firstly, to conduct a multiple scale transform of

each source image, and secondly, to create a composite multiple scale representation from the decomposed images with certain specific fusion rules. In the last step, a fused image is reconstructed by performing the inverse transform on the composite multiple scale representation [32]-[34]. This mechanism provides a flexible framework in which various transforms and fusion rules can be applied for applications with different requirements.

4.3 PYRAMID MSD TRANSFORM

4.3.1 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 l th level in the pyramid for the source image G_0 ($0 \leq l \leq N$, 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 $w(m, n)$

$$G_l = REDUCE[G_{l-1}] = w \otimes G_{l-1}, \quad \text{or} \quad G_l(i, j) = \sum_{m=-2}^2 \sum_{n=-2}^2 w(m, n) G_{l-1}(2i + m, 2j + n). \quad (4.1)$$

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,

$$G_{l,1} = EXPAND[G_{l+1}]. \quad (4.2)$$

or,

$$G_{l,1}(i, j) = 4 \sum_{m=-2}^2 \sum_{n=-2}^2 w(m, n) G_{l+1}((\text{floor})(i + m) / 2, (\text{floor})(j + n) / 2). \quad (4.3)$$

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:

$$C_l = (G_l - EXPAND[G_{l+1,1}]) / EXPAND[G_{l+1,1}]. \quad (4.4)$$

The contrast pyramid can be iteratively generated by applying the REDUCE and EXPAND functions to the source image. Figure 4.1 is the general scheme of MSD image fusion, and Figure 4.2 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.

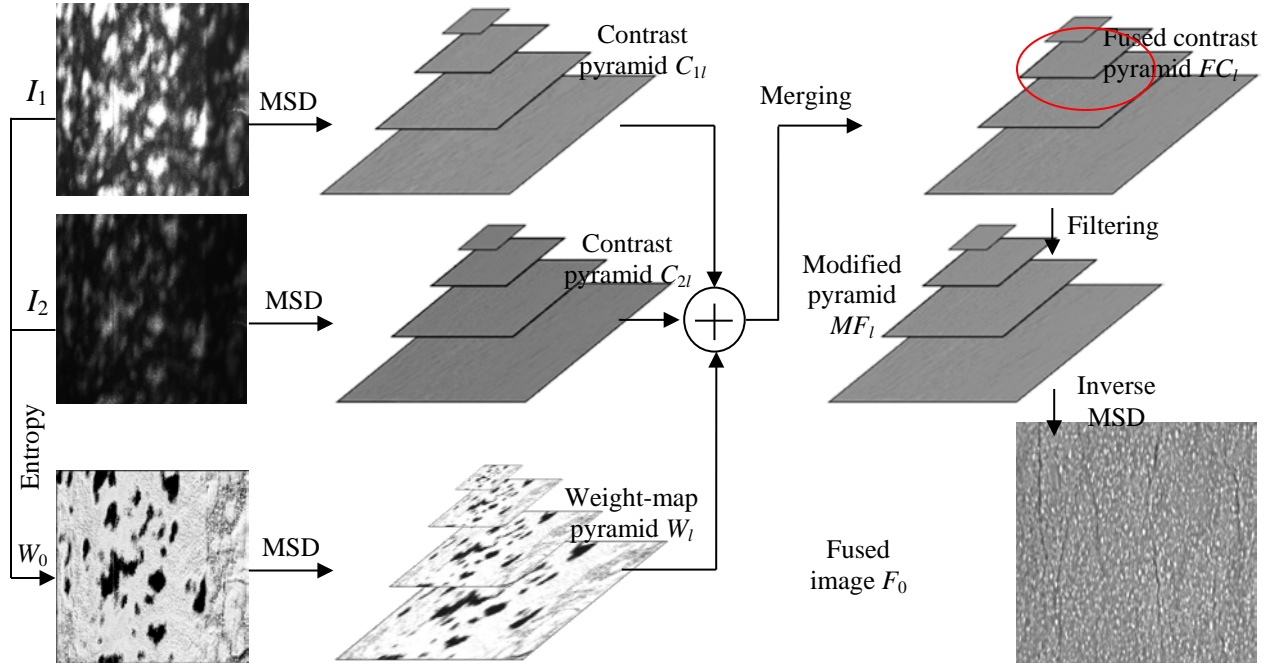


Figure 4.1 Scheme of MSD image fusion.

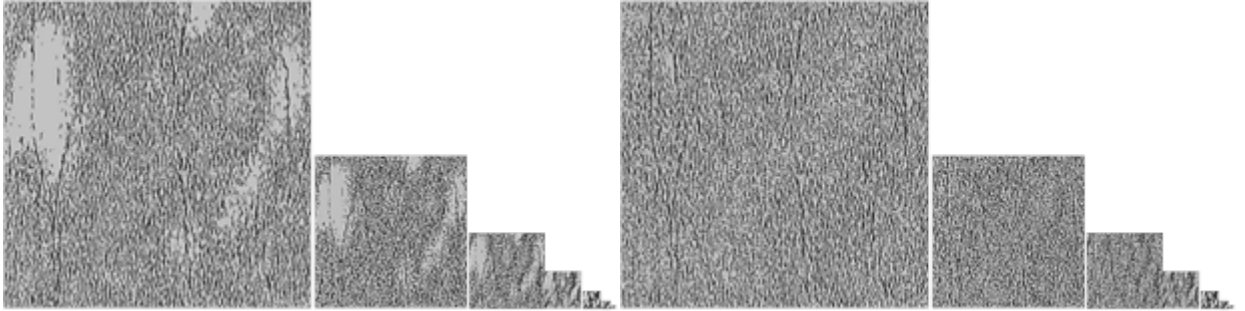


Figure 4.2 Contrast pyramids of source images overexposed (left) and underexposed (right)

4.3.2 Activity Measure and Weight-map Pyramid

The activity measure reflects the amount of useful information carried in each source image; and 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 and can be computed by $E_g = \sum_{i=0}^{255} -p_i \log(p_i)$, where p_i is the probability of an arbitrary pixel that has grayscale i ($i = 0, \dots, 255$). 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 4.1). Based on W_0 , a weight-map pyramid W_l can be built via the MSD transform as defined in Equation (4.3), as shown by W_l in Figure 4.1. W_l provides weighting coefficients for merging two contrast pyramids at each level.

4.3.3 Merging Strategy and Shadow Suppression

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:

$$F_l(i, j) = W_l(i, j) \times C_{1l}(i, j) + [1 - W_l(i, j)] \times C_{2l}(i, j). \quad (4.5)$$

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 4.1). Shadows normally have much larger width dimensions 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.

4.3.4 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:

$$F_N = MF_N. \quad (4.6)$$

$$F_l = MF_l \cdot EXPAND[F_{l+1,1}] + EXPAND[F_{l+1,1}]. \quad (4.7)$$

As seen in Figure 4.1, 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 4.1 with the following major steps:

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

4.4 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 [17]. The procedures of the crack detection algorithm include 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 a 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 constrains. The connected seeds serve as candidates for a crack, which are 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 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%, of which most of the difference comes from the fact that some shallow and narrow cracks are not captured with good contrast [17]. Figure 4.3 presents two examples of the detected cracks in the fused pavement images of the two pairs of complementary images in Figure 2.2. Most of the cracks in these images are correctly marked, indicating that the fused images preserve cracks information after the shadows are eliminated.

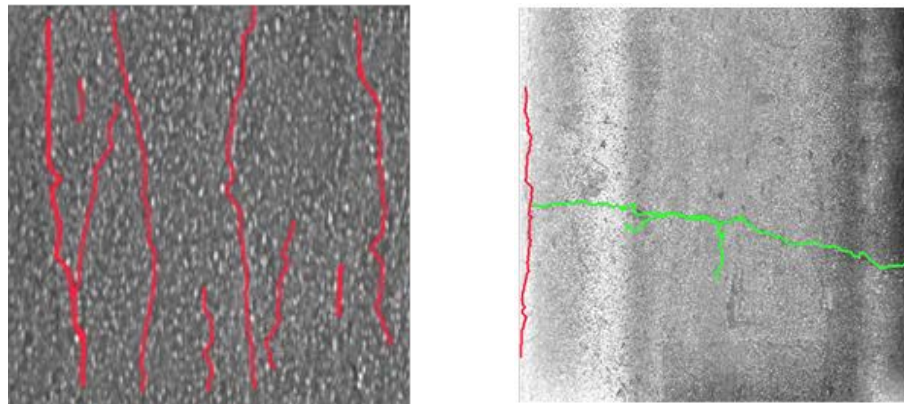


Figure 4.3 Detected cracks in a tree shadowed image (left) and a vehicle shadowed image (right).

Chapter 5: Result Analysis and Discussions

The purpose of validation test is to verify whether the proposed system can appropriately detect pavement cracks without lighting assistance, and whether the system shows relatively high consistence and reliance under multiple field tests. The crack results also shows the pavement surface change over time by comparing the road tests taken over around one and a half year span.

5.1 CRACK DETECTION ACCURACY

To validate the complementary imaging system for pavement crack detections, 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 5.1 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 are compared to the results from a manual crack identification procedure out of a set of shadow-free pavement images captured by the same system. The manual crack identification procedure was performed by an experienced researcher manually marking the cracks on images with editing tools that we developed together with the crack detection software. The image series shown in Figure 5.1 were formed by stitching four consecutive images, covering a 14.6-meter 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 rectangle. We found 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 difference with our manual markups mostly on short cracks that are only one pixel in width.

Table 5.1 presents the sums of crack data collected under two lighting conditions (sunny and cloudy) around a 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 of area A_a , and the average 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 last column in Table 5.1 lists the total crack length by combining the length of all longitudinal, transverse and alligator cracks. Comparing to the results of manual crack identification on shadow-free images (the third row), the system tends to overestimate both 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 type. Part of the reason for this is the difficulty to drive the vehicle along exactly the same path to cover the same pavement surface during multiple scans. Thus some cracks on the edges of the lane may or may not be captured. The difference may also indicate that our automatic crack detection algorithm is more likely to treat a crack as an alligator crack, due to the ambiguity to identify the boundary of alligator regions. Some falsely detected

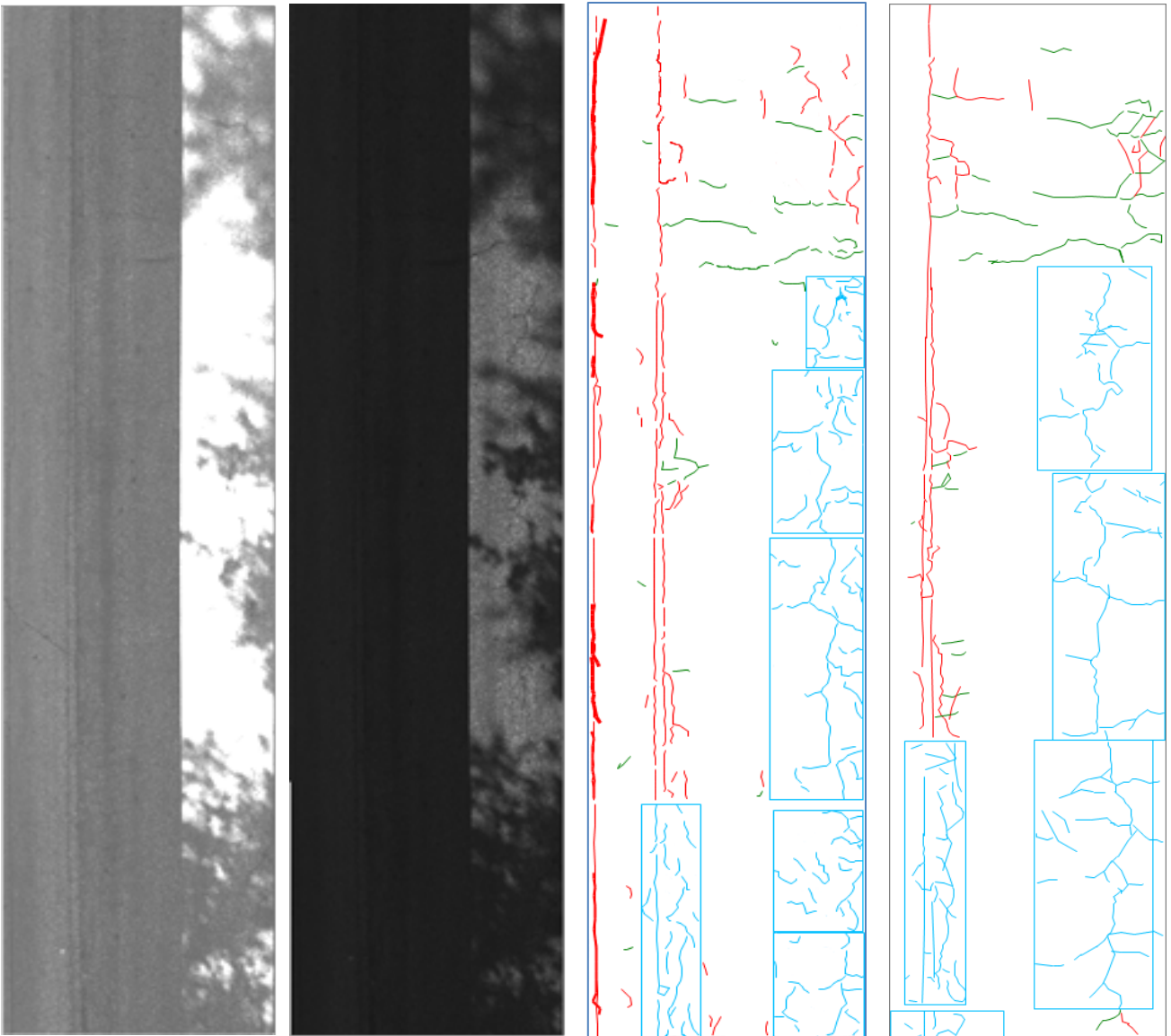


Figure 5.1 Cracks detected from complementary images are shown in crack maps. From left to right: overexposed pavement image, underexposed pavement image, results of automatic crack detection from complementary images, and results of manual crack identification from shadow-free image. Edge cracks may be different on the crack maps due to the difficulty in driving the vehicle along the same path across multiple scans.

transverse cracks from shadowed images turned out to be fine shadows of overpass wires which were like 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-meter pavement into 55 equal segments, as shown in Figure 5.2. Each data point in the figure represents the crack data summarized within one segment (a 4-meter distance). The crack data collected in the two different lighting conditions demonstrated the consistent trends with the R^2 s 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 distance can be observed, for example the longitudinal cracks at distance mark of 65 m and 85 m, and transverse cracks at mark 30 m and 40 m. These disagreements are caused by either false crack detection or counting cracks into an alligator block. However, the total crack lengths are quite consistent between automatic detection and manual identification, with the highest difference being 7.5%.

Table 5.1 Crack measurements under different lighting conditions.

Shadow	Longitudinal (m)	Transverse (m)	Alligator		Total (m)
			A_a (%)	A_d (m/m ²)	
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

5.2 REPEATABILITY OF MULTIPLE RUNS

Table 5.2 Crack measurements in multiple runs. lists the measurements of longitudinal, transverse, and alligator cracks in three runs on each test day. 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 A_a (%) and A_d (m/m^2). 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 maintain 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 lines-can cameras in different runs.

Table 5.2 Crack measurements in multiple runs.

Run	Longitudinal (m)			Transverse (m)			Alligator		Day 2		Day 3	
	Day 1			Day 1			Day 1		Day 2		Day 3	
	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

In order to see the data consistency in different runs, we divided the 220-meter pavement into 11 equal segments, and calculated the longitudinal, transverse, and alligator crack data for each 20-meter segment. Figure 5.3 Repeatability of crack measurements in multiple runs. displays the crack data in the three runs on Day 3. The three runs exhibit the 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.

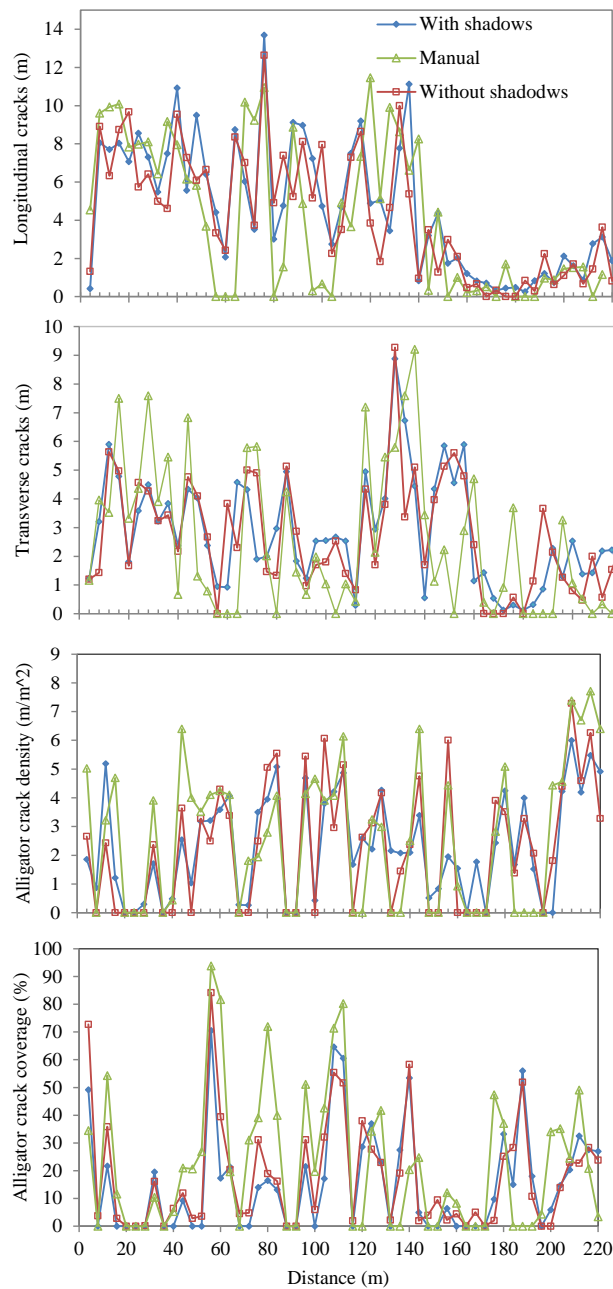


Figure 5.2 Regional crack data from images captured in sunny and cloudy days.

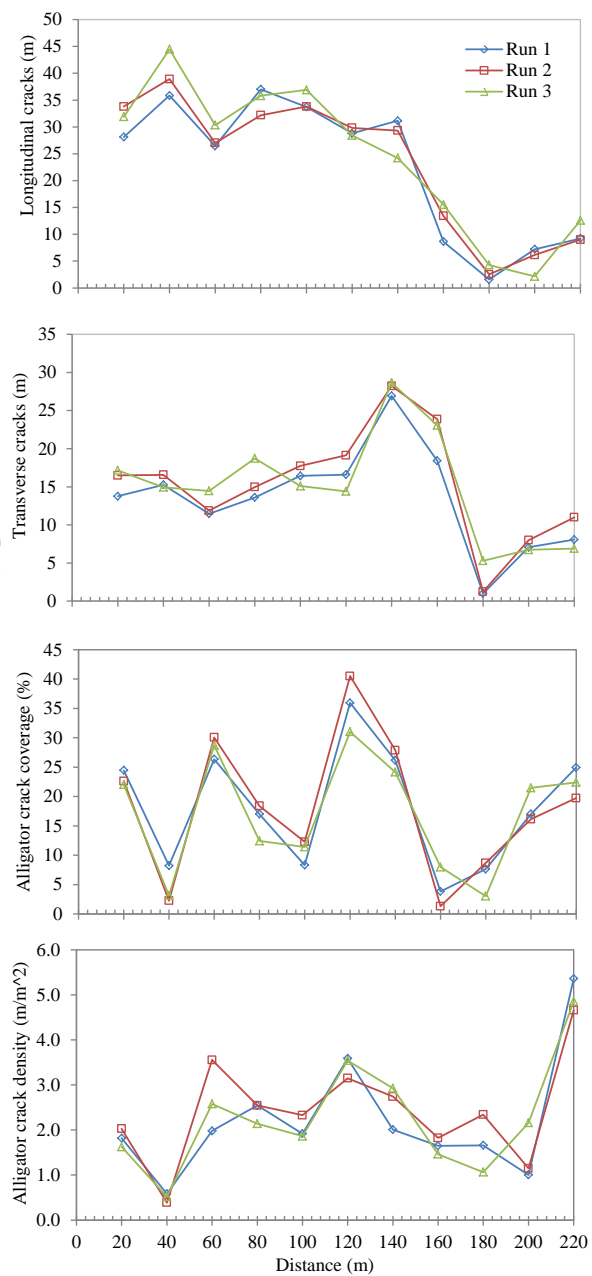


Figure 5.3 Repeatability of crack measurements in multiple runs.

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 (1—smaller than 3 mm, 2—between 3 and 6 mm, and 3—larger than 6 mm). Table 5.3 lists the crack measurements at different severity levels where L1, L2, L3, T1, T2 and T3 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 the decrease in crack length.

From Table 5.3, we can see that longitudinal cracks increased at all three levels, particularly at level 1 (L1), in both Δ_{21} and Δ_{31} , while transverse cracks had mixed change patterns. Transverse cracks increased at levels 2 and 3 (T2 and T3) in the first period (Δ_{21}) and then at levels 1 and 3 (T1 and T3) in the second period (Δ_{31}). Even though A_a doesn't show significant changes over time, 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 the time.

Table 5.3 Crack measurements over a 1.5-year period.

Road Test Date		$L_1(\text{m})$	$L_2(\text{m})$	$L_3(\text{m})$	$T_1(\text{m})$	$T_2(\text{m})$	$T_3(\text{m})$	$A_a(\%)$	$A_d(\text{m/m}^2)$
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 (15 months later)		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)
Day 3 (17 months later)		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)
Change over time (%)	Δ_{21}	119.60	16.89	7.22	-4.87	13.98	26.22	2.79	8.81
	Δ_{31}	149.64	18.44	3.81	22.35	8.60	18.74	-1.04	13.62

Chapter 6: Conclusions and Future Work

6.1 SUMMARY OF THE THESIS

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 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 blackout in the image, and the second camera is set in a low-exposure mode to prevent sunlit regions from being whiteout. 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 the 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 in the three test days were all under 15%, and the R^2 s of the three crack categories measured in 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 on-vehicle energy.

6.2 SUGGESTED FUTURE WORK

There is still room for improvement on the camera calibration algorithms developed in this study. Our line-scan camera calibration method is design to be easy to perform in the test field with a simple calibration pattern that is convenient to carry. Thus the calibration method reuses the existing well studied 2D camera calibration model with unique constraint that only applies to 1D geometry. Although the mathematical model is numerically accurate, it relies on restriction in camera pose when observing the calibration pattern. A better calibration model that is not constrained by any limits could potentially improve the calibration accuracy. This might be achieved through concepts of multiple view geometry.

Image fusion will also benefit from improved camera calibration. Our current calibration only computes the intrinsic camera parameters. Camera poses cannot be estimated from our current calibration. Thus image registration requires cameras to be mounted on a single plate so that the imaging planes of multiple cameras can be considered to be coplanar, and image registration only involves rotation and translation on the same plane. This can be improved by conducting full camera calibration which also computes the extrinsic camera parameters. Relative orientation and position between the cameras can be computed by the extrinsic parameters. As a result, images captured by arbitrarily positioned cameras can also be registered for fusion.

Bibliography

- [1]. Saraf, C. L., "Pavement condition rating system, review of PCR methodology," FHWA-OH-99-004, FHWA, U.S. Department of Transportation, 1998.
- [2]. Cheng, H. D., Jiang, X., Li, J., and Glazier, C., "Automated real-time pavement distress analysis," *Trans. Res. Rec.: J. Trans. Res. Board*, 1655(1), 55-64, 1999.
- [3]. Wang, K. C., "Automated system for pavement surface distress survey: A historical perspective on design and implementation," *Proc. Transport Research Board 77th Annual Meeting*, Transportation Research Board, Washington, D.C., 1998.
- [4]. Wang, K. C., "Use of digital cameras for pavement surface distress survey," *Transp. Res. Rec.* 1675, Transportation Research Board, Washington, D.C., 91-97, 1999.
- [5]. *NCHRP SYNTHESIS 334: Automated Pavement Distress Collection Techniques*. 2004, Transportation Research Board: Washington, D.C. p. 94.
- [6]. Howe, R. and G.G. CleMena, *An Assessment of The Feasibility of Developing and Implementing an Automated Pavement Distress Survey System Incorporating Digital Image Processing*, Virginia Transportation Research Council: Richmond, VA., No. VTRC 98-R1, 1997.
- [7]. Wang, K.C.P., *Designs and Implementations of Automated Systems for Pavement Surface Distress Survey*. *J. Infrastructure Sys.*, **6**(1), p. 24-32, 2000.
- [8]. WayLink Systems Co. Automated Distress Analyzer. Available from: <http://www.waylink.com/>.
- [9]. Inc., F.R. Automatic Road Analyzer. Available from: <http://www.roadware.com/>.
- [10]. CSIRO and Australia Roads and Traffic Authority. Pavement Inspection System. Available from: <http://www.csiro.au/solutions/RoadCrack.html>.
- [11]. Xu, B., Summary of Implementation of an Artificial Light System for Automated Visual Distress Rating System, Technical Report 5-4958-01-1, Center for Transportation Research, The University of Texas at Austin, August, 2008.

- [12]. Pavemetrics System Inc., Laser Road Imaging System,
<http://pavemetrics.com/en/lris.html>, last accessed September 2014.
- [13]. Xu, B., Yao, X. and Yao, M., Design of CrackScope (VCrack), Technical Report 0-5708-1, Center for Transportation Research, The University of Texas at Austin, 2007.
- [14]. A. Hirahara and K. Ikeuchi, "Detection of street-parking vehicles using line scan camera and scanning laser range sensor," in *Proceedings of IEEE in Intelligent Vehicles Symposium*, Piscataway, NJ, 2004.
- [15]. M. Lim and J. Limb, "Visual measurement of pile movements for the foundation work using a high-speed line-scan camera," *Pattern Recognition*, vol. 41, no. 6, pp. 2025-2033, 2008.
- [16]. K. C. P. Wang, Z. Hou, Q. B. Watkins and S. R. Kuchikulla, *Automated imaging technique for runway condition survey*, Federal Aviation Administration, 2007.
- [17]. X. Yao, M. Yao and B. Xu, "Automated measurements of road cracks using line-scan imaging," *Journal of Testing and Evaluation*, vol. 39, no. 4, 2011.
- [18]. C. Wang, A. Sha and Z. Sun, "Pavement core sample image acquisition system based on line-scan camera," in *2010 3rd International Congress on Image and Signal Processing (CISP)*, Yantai, China, 2010.
- [19]. R. S. Petty, M. Robinson and J. P. O. Evans, "3D measurement using rotating line-scan sensors," *Measurement Science and Technology*, vol. 9, pp. 339-346, 1998.
- [20]. Z. Zhao and G. Wen, "Model-based estimation of axisymmetric target's pose and speed using line-scan camera," in *2011 International Conference of System Science, Engineering Design and Manufacturing Informatization (ICSEM)*, Guiyang, China, 2011.
- [21]. R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," *IEEE Journal of Robotics and Automation*, vol. 3, no. 4, pp. 323-344, 1987.
- [22]. Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330-1334, 2000.

- [23]. J. Heikkila, "Geometric camera calibration using circular control points," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1066-1077, 2000.
- [24]. R. Horaud, R. Mohr and B. Lorecki, "On Single-Scanline Camera Calibration," *IEEE Transaction on Robotics and Automation*, vol. 9, no. 1, pp. 71-75, 1993.
- [25]. B. Hui, G. Wen, Z. Zhao and D. Li, "Line-scan Camera Calibration in Close-range Photogrammetry," *Optical Engineering*, vol. 51, no. 5, 2012.
- [26]. J. Drareni, S. Roy, and P. Sturm, "Plane-based calibration for linear cameras", *International Journal of Computer Vision*, 91(2), 145-156, 2011.
- [27]. C. A. Luna, M. Mazo, L. Lazaro and J. F. Vazquez, "Calibration of Line-Scan Camera," *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 8, pp. 2185-2190, 2010.
- [28]. Debevec, P.E. and J. Malik. Recovering high dynamic range radiance maps from photographs. *Computer Graphics Proceedings, Annual Conference Series*, 369-378, 1997.
- [29]. Robertson, M.A., S. Borman, and R.L. Stevenson. Dynamic range improvement through multiple exposures. in *Image Processing, 1999. IICIP 99. Proceedings. 1999 International Conference on*. 1999.
- [30]. Goshtasby, A.A., Fusion of multi-exposure images. *Image and Vision Computing*, 2005. 23(6): p. 611-618.
- [31]. Block, M., et al. Multi-Exposure Document Fusion Based on Edge-Intensities. in *Document Analysis and Recognition, 2009. ICDAR '09. 10th International Conference on*. 2009.
- [32]. Zhong, Z. and R.S. Blum, A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application. *Proceedings of the IEEE*, 1999. 87(8): p. 1315-1326.
- [33]. Piella, G., A general framework for multiresolution image fusion: from pixels to regions. *Information Fusion*, 2003. 4(4): p. 259-280.

- [34]. Burt, P.J. and R.J. Kolczynski. Enhanced image capture through fusion. 1993.
Berlin, Ger: Publ by IEEE.
- [35]. Mertens, T., J. Kautz, and F.V. Reeth, Exposure Fusion: A Simple and Practical
Alternative to High Dynamic Range Photography. *Computer Graphics Forum*, 2009.
28(1): p. 161-171.
- [36]. Wang, J., D. Xu, and B. Li, Exposure fusion based on steerable pyramid for
displaying high dynamic range scenes. *Optical Engineering*, 2009. 48(11): p. 117003-
10.
- [37]. J. N. Cederbery, "A course in modern geometries", 2nd ed., Springer, New York,
NY, 1995.
- [38]. R. Hartley and A. Zisserman, "Multiple view geometry in computer vision", 2nd ed.,
Cambridge University Press, Camgridge, UK, 2004.
- [39]. J. Weng, P. Cohen, and M. Herniou, "Camera calibrationwith distortion models and
accuracy evaluation," *IEEE Transactions on Pattern Analysis and Machine
Intelligence*, 14(10), 965-980, 1992.
- [40]. Yao, M., Zhao, Z., and Xu, B., Geometric calibration of line-scan camera with a
planar pattern, *Journal of Electrical Imaging*, 23(1), 013028-013028, 2014.
- [41]. Land, E.H. and J.J. McCann, Lightness and Retinex Theory. *J. Opt. Soc. Am.*,
1971. 61(1): p. 1-11.
- [42]. McCann, J. Lessons learned from Mondrians applied to real images and color
gamuts. in *Final Program and Proceedings of the 7th IS and T/SID Color Imaging
Conference: Color Science, Systems and Applications*, November 16, 1999 -
November 19, 1999. 1999. Scottsdale, AZ, United states: Society for Imaging Science
and Technology.
- [43]. Brainard, D.H. and B.A. Wandell, Analysis of the retinex theory of color vision.
JOSA A, 1986. 3(10): p. 1651-1661.
- [44]. Land, E.H., Recent advances in retinex theory and some implications for cortical
computations: color vision and the natural image. *Proceedings of the National
Academy of Sciences of the United States of America*, 1983. 80(16): p. 5163.