Digitizing Device to Capture Track Impression
Grant Number 2010-DN-BX-K145

Mihran Tuceryan
Jiang Yu Zheng

Abstract

In crime scene investigations it is necessary to capture images of impression evidence such as tire track or shoe impressions. Currently, such evidence is captured by taking two-dimensional (2D) color photographs or making a physical cast of the impression in order to capture the three-dimensional (3D) structure of the information. This project aims to build a digitizing device that scans the impression evidence and generates (i) a high resolution three-dimensional (3D) surface image, and (ii) a co-registered two-dimensional (2D) color image. The method is based on active structured lighting methods in order to extract 3D shape information of a surface. A prototype device was built that uses an assembly of two line laser lights and a high-definition video camera that is moved at a precisely controlled and constant speed along a mechanical actuator rail in order to scan the evidence. A prototype software was also developed which implements the image processing, calibration, and surface depth calculations. The methods developed in this project for extracting the digitized 3D surface shape and 2D color images include (i) a self-contained calibration method that eliminates the need for pre-calibration of the device; (ii) the use of two colored line laser lights projected from two different angles to eliminate problems due to occlusions; (iii) the extraction of high resolution color image of the impression evidence with minimal distortion; and (iv) a super resolution method that takes advantage of the model of the rail motion in order to further increase the resolution of the extracted color image. The resulting digitized 3D image has a resolution of 0.0438mm in the Y-axis (along the rail motion direction), 0.2369 mm along the X-axis (limited by the HD video camera resolution) and 0.5mm along the Z-axis (depth). The resulting digitized 2D color image has a resolution of 1900 pixels in the X-axis direction and up to 4500 pixels in the Y-axis direction (depending on the speed of the rail motion and length of scan). The Y-axis resolution can be further increased up to 9000 pixels using super resolution techniques. The scan of a long tire track (about 1.75m) at the lowest scan speed takes about 20 minutes at the slowest rail speed and processing the resulting video to generate the 3D image and the 2D color image takes less than 1 hour on a regular PC.

Table of Contents

Executive Summary

In crime scene investigations it is sometimes necessary to capture images of tire track or footprint impression evidence. Currently, such evidence is captured by taking two-dimensional (2D) color photographs or making a physical cast of the impression in order to capture the three-dimensional (3D) structure of the information. On this project we have developed new methods for recording impressions of tire tracks and footwear and we have built a prototype device to implement these methods. The device scans the impression evidence using imaging techniques to capture (i) a high-resolution three-dimensional (3D) image of the surface with metric depth measurements, and (ii) a co-registered high-resolution two-dimensional (2D) color image. The method is based on active structured lighting methods to extract 3D shape of a surface. The prototype device uses an assembly of two line laser lights and a high-definition video camera that is moved at a precisely controlled and constant speed along a mechanical actuator rail in order to scan the evidence and capture the 3D image. Prototype software was also developed which implements the image processing algorithms, the necessary calibration methods, and the surface depth calculations. The methods developed in this project for extracting the digitized 3D surface shape and 2D color images include (i) a self-contained calibration method that eliminates the need for pre-calibration of the device before scanning; (ii) the use of two colored line laser lights projected from two different angles to eliminate problems due to occlusions; (iii) the extraction of a high resolution color image of the impression evidence with minimal distortion; and (iv) a "super resolution" method that takes advantage of the rail motion model in order to further improve the resolution of the extracted color image.

Obtaining the 3D impression images are based on two steps: (i) the actual scanning of the impressions in the field which results in a high definition digital video on a standard memory card, and (ii) the computational reconstruction of the 3D image and the 2D color image from the video. The scanning is done in the field where the evidence is. During this step, the field technician will set up the device physically over the evidence to be scanned and place a small calibration object in the scene. The device has built-in control panel and buttons that will reset the camera assembly position on the rail and perform a short or long scan depending on the type of evidence (long tire track impression or short impression such as footprint or partial tire tracks). The device does not have to be calibrated ahead of time and does not have the issues of calibration being lost during the transportation process. The analysis is done in the crime lab using the video data captured in the field. The computation of the 3D image of the surface is performed by analyzing each video frame, extracting the laser lines in the frame, and from their deformations, estimating the surface heights at each point along the detected laser light. This is done using the triangulation geometry between the laser light source and the camera capturing the scene. The calculation of the configuration of this geometry is crucial for accurate computations of the surface depths. Normally, in devices used in laboratories, this is done by an offline calibration process and this information is used to do the 3D calculations assuming that the configuration does not change from scan to scan. We cannot make this assumption with our device because the device is supposed to be transported to the field for the scans, the integrity of this configuration during this transportation process cannot be guaranteed after the calibration is performed. For this reason, we developed a calibration method which does not rely on pre-calibrated configuration for the scans to be useful. We integrated the calibration method into the capturing of the video so that the calibration of the geometry can be done as part of the analysis after the video is captured. We accomplished this by embedding a calibration object whose exact dimensions are known in the scene during the capture. We also developed a method to use the images of the calibration object in the video to compute the exact camera/laser light geometry during the analysis phase of the data.

The extraction of the laser lights is performed by detecting pixels of certain hues in each video frame.

Since we use two different colored laser lights (red and green), we use this hue information for detecting the pixels in the image that belong to the laser lights. We utilized two different methods and fused their results in order to improve the localization accuracy of the laser lights in the image. The first method is based on the detection of the peak in the laser light within a band of pixels detected based on the hues. The peak is typically at the center of this band. However, in certain situations the laser lights saturate the camera sensors and result difficulty locating the peaks. In this case, we do an edge-detection with the aim of detecting the leading and trailing edges of the laser light and thus localizing the center point more accurately. We have further enhanced the hue based detection of the laser light by using an adaptive method. This method uses randomly selected frames from the video and examining the color information of the pixels that belong to background regions in order to build a threshold surface in the color parameter space which then is used to detect pixels belonging to the laser light. The resulting detection of laser pixels is much more robust and cleaner. Stacking the detected laser lights in each video frame results in an image we call "disparity image" which is an image that contains the information about the deformations of the laser light due to surface geometry. We use this disparity image and the camera/laser light configuration to calculate a depth map for each of the pixels in the disparity image.

In addition to the calculation of the 3D surface image, we extract from each video frame one line of video data which then are stacked to form a single high resolution color image. As the camera moves in a linear motion along the rail, each captured video frame corresponds to a slightly shifted image of the scene. Thus, extracting one line from each frame is analogous to a line camera which is then shifted in time to form the complete image. The linear motion of the camera and stacking of the lines from the video frames corresponds to an orthographic projection along this axis (in the direction of motion) and thus does not suffer from the distortions due to perspective projection. The perspective projection along the axis perpendicular to the motion direction still holds because this is governed by the camera optics. The resolution of the image in the direction of motion can be increased by decreasing the speed of the motion or the resolution can be decreased by increasing the camera motion, resulting in a tradeoff of image resolution and the time to scan the evidence.

We have also developed a method of increasing the resolution along the camera motion by using image estimation techniques known as "super resolution." This method takes advantage of sub-pixel overlap due to motion of the camera of successive lower resolution images in order to estimate a higher resolution image. Because our camera motion is well known and precisely controlled, we are able to take advantage of this a priori knowledge to implement this further improvement in resolution. The resulting digitized 3D image has a resolution of 0.0438mm in the Y-axis (along the rail motion direction), 0.2369 mm along the X-axis (limited by the HD video camera resolution) and 0.5mm along the Z-axis (depth). The resulting digitized 2D color image has a resolution of 1900 pixels in the X-axis direction and up to 4500 pixels in the Y-axis direction (depending on the speed of the rail motion and length of scan). The Y-axis resolution can be further increased up to 9000 pixels using super resolution techniques. The scan of a long tire track (about 1.75m) takes about 20 minutes at the slowest rail speed and processing the resulting video to generate the 3D image and the 2D color image takes about 1 hour on a regular PC. This method speeds up the process of recording track impressions and improves the quality of the recordings made. As a portable device, our laser head is small and battery driven, as compared to current non-portable laser devices used in the lab environment that is fully supported by power. With a complementary device for use in the laboratory, it may be possible to collect comparable renderings from confiscated tires or shoes. This will make it possible to have the computer, using pattern recognition and comparison techniques, help the investigator expedite the process of making comparisons and matches.

We have also worked on developing techniques to utilize the captured 3D images for matching against

similar digital 3D images in databases to quickly narrow down the suspect candidates. The matching is meant to narrow down the number of possible hits and, therefore, help the examiner with the match on a smaller number of candidate images.

We have tried digitizing impressions in different materials such as stone, mud, soil, sand, and snow. Some of these materials result in excellent quality 3D images. Some materials, such as snow, are very difficult media, in which the laser light tends to attenuate and becomes difficult to detect. In yet other materials, such as coarse sand, one color laser appears fine but the other color is hard to detect. Scanning impressions in daylight causes similar problems with the ambient bright light making it harder to detect the laser in the images. Some of these problems can be alleviated by using more powerful laser lights or using some sort of shading (such as a tent or cover) to lessen the effects of the bright ambient light.

In order to evaluate the accuracy of the system, we have designed and conducted a systematic method of testing in collaboration with Indiana Forensic Institute. The conclusion of this testing is that the device is capable of 0.5mm resolution in depth and higher resolutions in X and Y dimensions of the image.

The method developed in this project is a non-contact and non-destructive process and it will often render the making of casts optional. The device is portable with minimal weight. The device costs about as much as an accessorized digital camera with an associated laptop computer. With the 3D depth data, investigators can further have the depth of the impression on soft ground, in addition to the traditional impression that has usually been thresholded to binary shape marks that indicate contact points on the ground. The information upon which comparisons are based becomes richer. The scanning times and computation times compare favorably to existing methods of taking casts of the impressions.

The staff who worked on this project include the following: Mihran Tuceryan (Faculty member), Jiang Yu Zheng (Faculty member), Ruwan Janapriya Egoda Gamage (graduate student), Abhishek Joshi (graduate student), Mechanical Engineering seniors (undergraduate students), Jafar Alobaid (undergraduate student), Hussain Alhudud (undergraduate student), Honyuan Cai (graduate student), Chiang Tsunghan (undergraduate student). The independent testing of accuracy over the Summer and Fall of 2012 was done by a consulting services purchased from Mr. Herb Blitzer of the Indiana Forensic Institute, Indianapolis.

1 INTRODUCTION

In crime scene investigations it is necessary to capture images of impression evidence such as tire track or shoe impressions. Currently, such evidence is captured by taking two-dimensional (2D) color photographs or making a physical cast of the impression in order to capture the three-dimensional (3D) structure of the information [2, 8, 10]. The 2D photographs, under the right illumination conditions, may highlight feature details in the evidence, but do not provide metric depth measurement information for such features. Obtaining a 3D physical cast of the impression may destroy the evidence in the process. Therefore, the use of a 3D imaging device which can capture the details of such impression evidence can be a useful addition to the toolkit of the crime scene investigators (CSI). In this report, we present the design of such an impression imaging device which includes a calibration method for obtaining the 3D image with the proper metric information. We have built the prototype hardware and the software to scan and calculate 3D digital images of such impression evidence. We had the goal of sub-millimeter accuracy (0.5mm) in depth resolution. The spatial resolution in the other two dimensions could be achieved using the appropriate imaging device and settings.

2 LITERATURE REVIEW

The normal process of imaging impression evidence requires that the camera's optical axis be perpendicular to the ground at the site of the track. Also, there is a requirement for proper oblique lighting in order to see details created by varying depths of impression as intensity variations in the photographs. In the case of tire tracks, where the impression may cover several linear feet, the ground may not be level and camera distances may lead to confusing readings [2, 8, 10]. The guidelines for imaging such evidence are specified by the Scientific Working Group for Shoe print and Tire Tread Evidence (SWGTREAD) guidelines [1]. The guidelines in the literature to image impression evidence suggest that sufficient resolution be used to capture as fine details as possible. Smith cites the range of the ridge spacing of the Schallamach patterns to vary from 0.05mm to 0.5 mm [11]. However, the size of the features that would transfer to materials such as mud, sand, and snow would be limited due to the range of granule sizes and angle of repose in these materials [14, page 315]. For our device described in this report, we had the goal of sub-millimeter accuracy (0.5mm) in depth resolution. The spatial resolution in the other two dimensions could be achieved using the appropriate imaging device and settings.

A device based on similar principles has been built before in order to scan relics excavated from archeological sites and construct their 3D computer models [18, 9]. The device was constructed for measuring tall sculptures and has been brought into a UNESCO World Heritage Site. Calibration of such devices is not trivial [6, 5].

There has been a lot of previous work on structured lighting methods to extract 3D surface shape [17]. Some of these methods rely on solving the correspondence problems [13]. The accuracy characteristics of others do not meet the forensic requirements of impression evidence scanning [4]. In most outdoor applications it is important to implement on-site calibration to avoid miss-alignments during transportation [12].

The current work modifies it to satisfy the special requirements in the field of forensic imaging. Buck et al. present the use of existing commercial 3D imaging devices for footwear impressions [16]. The existing devices do not satisfy some of the imaging requirements (e.g., resolution in depth) in forensics applications. They do not work very well outdoors on long track impressions with a single scan. They usually require multiple short scans which need to be stitched together.

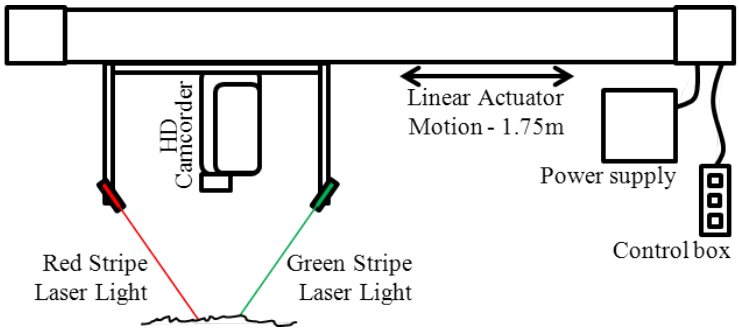


Figure 1: The design for the device prototype built.

3 METHODOLOGY

The device to capture 3D images of impression evidence is based upon using triangulation geometry of a projected laser light on a surface and observing its image from a slightly offset position. Typically, in the general case, this method suffers from the difficulty of resolving correspondences of points in one image with the points in a second image. We resolve this by performing a sequential scanning of the surface with a line laser. This eliminates the correspondence ambiguities and lets us calculate the 3D surface geometry directly for each point on the laser line. The detection of the line is done based on its color (hue) and is repeated for each step in the scan. For each frame in which we extract a single laser line, we also extract a color image of the scene. This results in a high resolution color image of the evidence along with the metric 3D surface information. In addition, we have implemented super resolution (SR) algorithms which result in an even higher resolution color image of the evidence exploiting the specific camera motion of the device during the scanning process. In the rest of this section, we describe the hardware configuration of the device, the various calibration algorithms, the image processing and 3D surface estimation algorithms, and the super resolution color image computation algorithms.

3.1 HARDWARE SETUP

The device for digitizing the impression evidence consists of a motorized rail (actuator) with a HD video camera and two line laser lights, each with a different color as shown in Figure 1.

A physical prototype of the device is shown in Figures 2 and 3. To build the device, we used the following components:

1. Linear actuator rail: Motion Dynamic belt drive linear actuator MSA-PSC, 1771mm long.
2. Myostat Cool Muscle (CM1 Series) Integrated Servo motor with a planetary gearbox 8:1.
3. 24VDC Power supply, USB interface cable, and programming interface.
4. Canon VIXIA HF M30, HD camcorder.
5. Laser stripe lights.
   (a) 5mWatt, 532nm green module.
   (b) 3mWatt, 650nm red module.

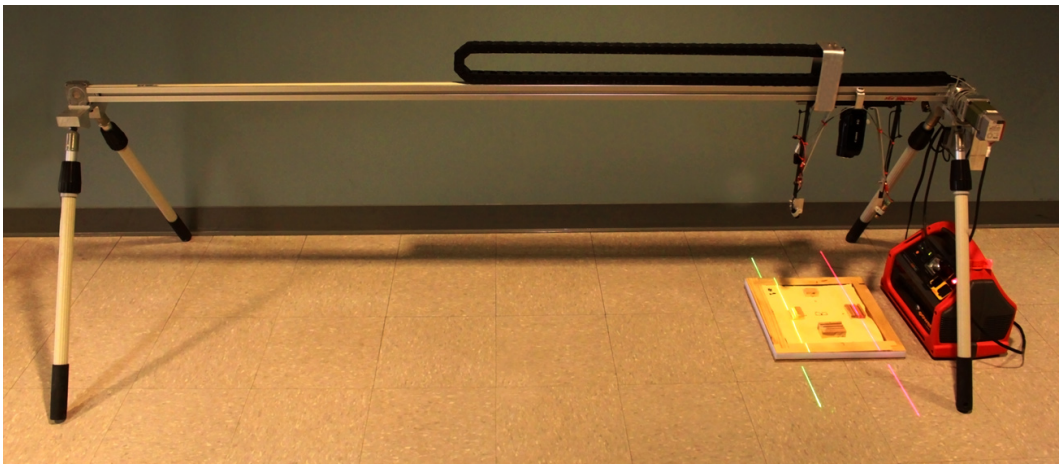


Figure 2: The prototype of the device built.

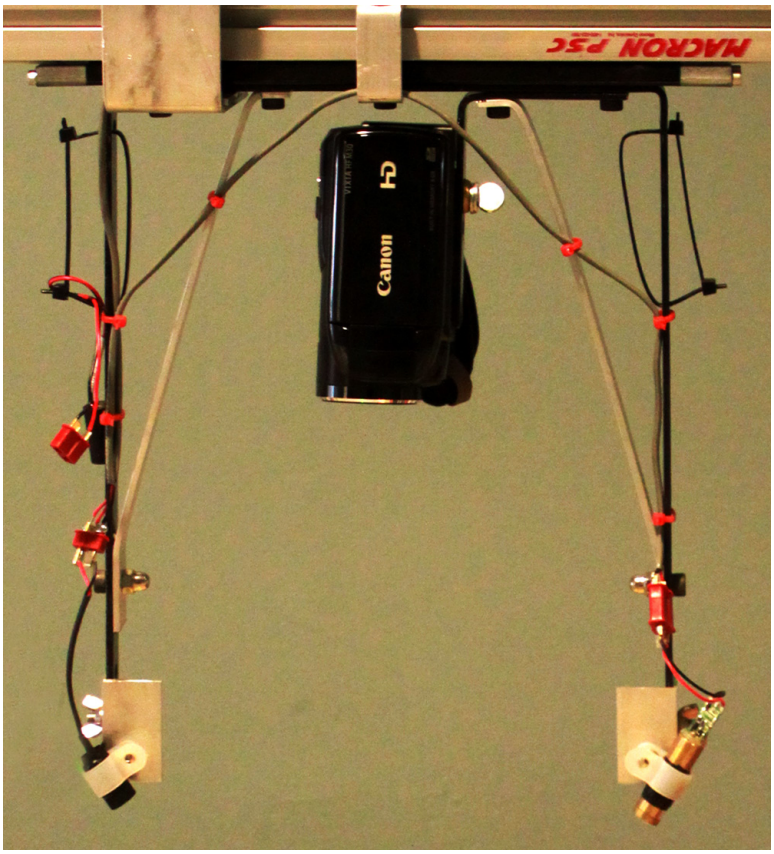


Figure 3: A close-up of the laser and camcorder assembly.

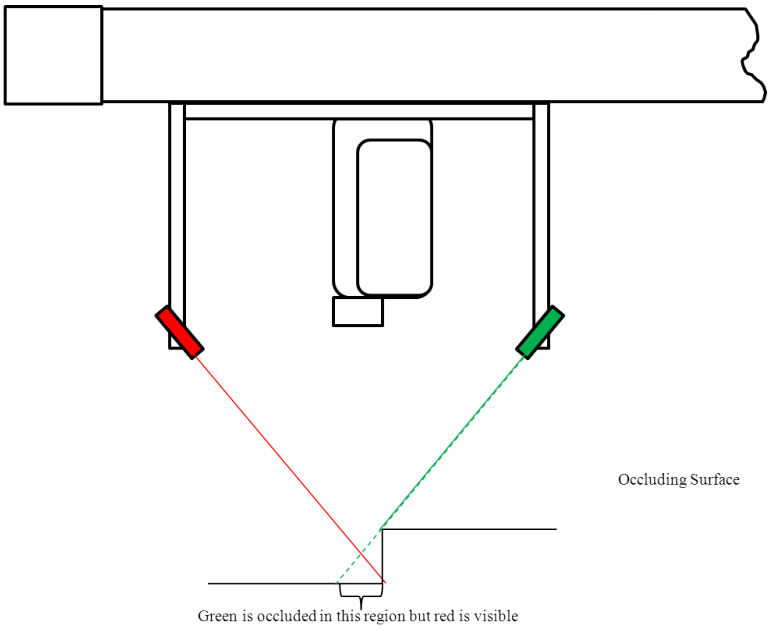


Figure 4: Two lasers are used to eliminate the blind spots due to occlusion. The different laser colors solve the correspondence problems for calculating the triangulation.

We designed and built a bracket and a leg assembly on the two sides of the actuator rail that can be taken off in order to package and carry the device to the field. We also designed and built a bracket to mount the camcorder and the laser lights onto the rail with a fixed geometric configuration. The final physical prototype of the device is shown in Figures 2 and 3.

The two different colored lasers are used to eliminate the blind spots due to occluding surfaces as shown in Figure 4. Also, the two different colored lasers are used in order to distinguish the two different light sources, and, thus, solve the correspondence problem more easily.

The device has a total weight of (11 lb. excluding the legs) so that it is portable by one person and can be loaded into a normal van or a pick-up truck. Its legs can also easily be dismantled for easy transportation. By setting the device at the measuring site properly, a user can push one of the buttons that are short scan and long scan to move the laser-camera set in one of the preset distances (500mm, 1700mm). This will provide an appropriate distance along the rail direction according to the user's requirement. A fast homing button is also provided to reset the laser-camera assembly at a preset home position. Further, a stop button allows the emergency stop and selecting a new starting position for scanning short length ground scenes.

During an outdoor tire and footprint scan, the power can be drawn from a nearby vehicle or from a charged portable battery unit. The camera focus and exposure are set to manual mode and fixed after autofocus and auto exposure are obtained at the beginning of the scanning. This will eliminate possible camera optical parameters changing during the scanning process due to changing heights in the scanned surface.

A user interface for controlling the device has been developed on a PC as well as on a mobile tablet using Android platform. This interface allows the field technician to control scan parameters such as rail speed, start and stop positions, and/or run length in a more customized way than the preset buttons. The

more advanced user interface will also allow the capability to change the preset values of these buttons.

The video of two laser stripes are recorded into a flash disc. After the scan of the entire site is done, the disc is moved to a PC in the crime lab and the video clips are copied to the hard disc of PC for image processing, 3D reconstruction, and further analysis.

3.1.1 Hardware Costs

Over the life of the project, we purchased and tried various hardware components to improve the performance of the device. The final set of components we used in the prototype and the PC for processing as reported here and their costs is given in Table 1.

Component used	**quantity**	**cost**
PSC Series Belt Drive Linear Actuator, 1826mm Stroke	1	$414
60mm single stage inline planetary gearbox 8:1	1	$547
CM1 Series Integrated Servo, C Type, NEMA23, Long	1	$495
CM1 Series Power + I/O Cable, 1m	1	$32
1.8m USB Power/Communications "Y" Cable for CM1 Series Motors	1	$65
IDEC PS5R Series Slim Line 24VDC Power Supply, 90 Watt [3.75A Rated], AC or DC Input, 1 x IP20 Screw Terminal Output	1	$120
Leg assembly (designed and built by mechanical engineering seniors)	1	$500
Canon VIXIA HF M30 HD Camcorder	1	$390
Memory Cards (32 GB, Class 4, SD/SDHC)	2	$60
Line laser modules	2	$100
Android tablet	1	$372
Electronic parts and components (various)		$300
Portable power supply	1	$150
Power inverter for automobile	1	$38
Cable carrier for the actuator rail	1	$140
PC computer for processing (custom assembled in-house)	1	$2500
Total		**$6223**

Table 1: The total cost and the component costs used in the prototype.

3.2 IMAGE PROCESSING ALGORITHMS

3.2.1 Laser stripe detection based on hues

The detection of the laser stripe in the image is based on the color properties of the pixels. Because we have laser stripes of two colors (red and green), we detect these laser stripes by detecting pixels that fall within the color range of these hues present in the laser lights. Figure 5 shows a typical video frame and the intermediate steps during the laser stripe detection for the red laser line.

(a) (b) (c)

Figure 5: Color based laser stripe detection steps (a) original video frame (b) region of interest of the original image and laser pixels detected (c) Results are super imposed and region of interest is highlighted in yellow.

The detection of the laser stripes is done by restricting the search to a smaller region of interest (ROI) in the video frame which is determined by the camera/laser configuration and a limit on surface variations. Set up a region of interest (ROI) for red stripe. The region of interest is determined by the geometry of the camera and laser stripe light. It is an approach to reduce the search region and thus reduce the computational time for searching the laser light in the image. The search region is determined by the amount of variation in surface height, which in turn has a direct effect on the deformation of the laser line in the image. If the surface variation is not large as in typical impression evidence surfaces, the deformation of the laser line will be small. Therefore, we can restrict the region in which we need to perform the search for detecting the laser light to a small subarea of the image. This search ROI is also determined by the angle with which the laser light is projected on the surface with respect to the camera capturing the images. One laser light will be projected somewhere in the upper half of the image and the second laser light will be projected somewhere in the lower half of the image. Thus for each laser light, we can define the search ROI based on this a priori knowledge of the device configuration. The yellow lines in Figure 5c depict the ROI for the red laser light in this example frame.

Once the ROI is defined, the laser stripe is detected by using a Hue-Saturation-Value (HSV) representation of the image. This identification is done if the pixel's hue value falls within a specified range $[\text{hue}_{\text{min}}, \text{hue}_{\text{max}}]$, which is determined by the color of the particular laser light. A dilate morphological image operation is performed in the vertical direction. The pixels resulting from this operation constitute a set of candidate points for the laser pixels. This set of pixels typically forms a thick (5-10 pixels wide) band and the location of the laser stripe needs to be refined for better accuracy Figure 5b shows the candidate points so detected in the example video frame of Figure 5a.

In the mask generated based on the hue field, we search for the peak intensity values in each vertical line. These peaks are the pixels with the maximum Value (brightness) for the pixel in the HSV representation and correspond to the peak of the laser stripe. Figure 5c shows the identified laser pixels superimposed on to the original image.

We combined laser edge information to improve accuracy of peak intensity method. The edge detection is carried out only in a region (a band) closer to detected peak location of laser intensity. We use a vertical edge operator to detect horizontal edges. Resulting leading and trailing edges of the laser stripe are used for further processing. The midpoint of the leading and trailing edge is calculated and if it is within a certain distance of the peak location, the midpoint is marked as the laser location.

Figure 6 shows disparity images created using the methods explained above. The black pixels correspond to failures of laser detection. As shown, failures are minimal in the combined approach.

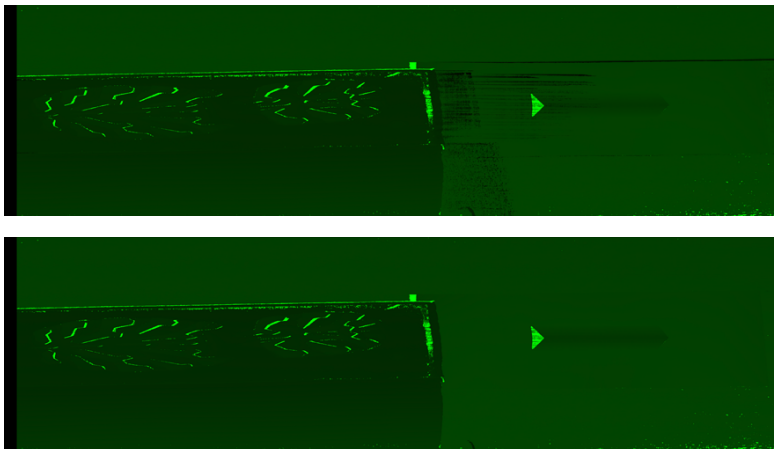


Figure 6: Disparity image created using peak intensity method (top) and disparity image created using the combination of peak and edge detection method (bottom)

3.2.2 Laser stripe detection based on RGB lookup table

The peak detection and edge based method described above was still resulting in a lot of false positives for laser pixels (i.e., pixels were labeled as laser but were actually non-laser pixels). To improve the robustness of the laser stripe detection, we have implemented a slightly modified method based on an adaptively constructed RGB lookup table [15]. This approach contains four image processing steps.

- *Stripe edge detection:* Edge detection is done using a Sobel filter. The red channel of the image is used in edge detection for red laser. Similarly, green channel is used for green laser image. The resultant image has pixels highlighting horizontal edges of the laser stripe. This goes through another step with a high and low thresholds to eliminate noise and low response edges. All pixels in between a negative and a positive edge are considered as pixels belong to the laser stripe.

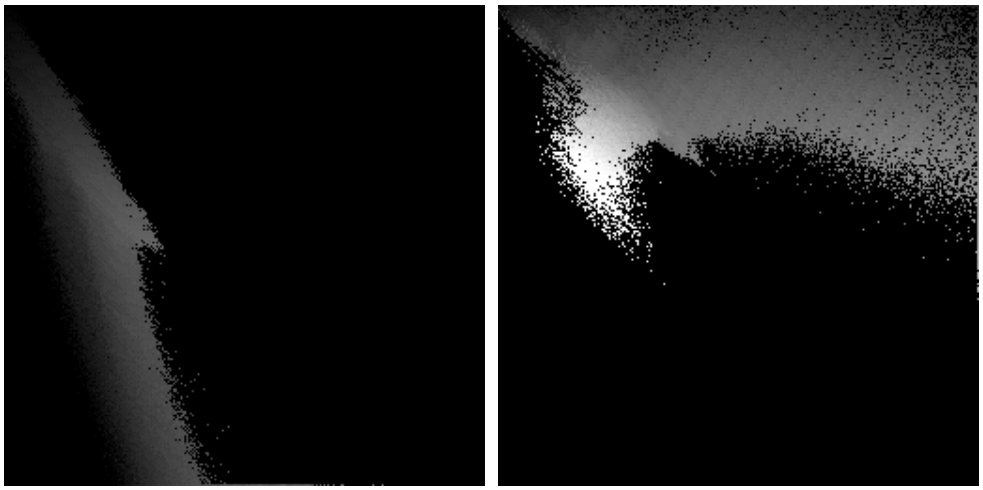
- *RGB lookup table based color validation:* To validate the pixels in a laser stripe, two lookup tables — one for each of the green and red laser images — are generated from every 100th frame of the video. The lookup table consists of color values belonging to the *background color subspace*. These values are generated from pixels c_i outside the region of interest(ROI) in which laser stripe is searched in order to capture the color characteristics of the background colors. Here $c_i = (R, G, B) \in ([0, 255], [0, 255], [0, 255])$, where the square brackets indicate ranges in the three color bands. Let's consider the lookup table: bcs_{red} for red laser. First, we construct a complement lookup table, $\overline{bcs}_{red}(R, G, B)$ as follows:

$$\overline{bcs}_{red}(R, G, B) = \begin{cases} 1 & \text{if } \exists i : (R, G, B) \in c_i \text{ or } (R = 0) \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

  Then, we use this to construct the actual lookup table $bcs_{red}(R, G, B)$ that will be used for detecting the laser pixels:

$$bcs_{red}(R, G, B) = \begin{cases} 1 & \text{if } \exists R' \geq R : \overline{bcs}_{red}(R', G, B) = 1 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

  The lookup table bcs is constructed with the assumption that laser stripe color values are suppose to have higher red components than the background (Figure 7). A given color pixel $x_i = (r, g, b)$ is



(a) (b)

Figure 7: RGB lookup tables for an indoor scan (a) Green lookup table. A 255x255 matrix, red and blue channels are rows and columns respectively. Intensity represents green channel.(b) Red lookup table. A 255x255 matrix, green and blue channels are rows and columns respectively. Intensity represents red channel.

classified as a red laser pixel, if $bcs_{red}(r, g, b) = 1$. The lookup table $bcs_{green}(R, G, B)$ is constructed in a similar manner to the red one and used in the same way for the green laser pixel detection. If a stripe from the edge detection step contains more than 95% of laser pixels, it is considered a laser stripe.

- *Stripe width validation:* Each stripe is validated again whether it is between a higher and a lower threshold. The lower threshold is set to 2 pixels and higher threshold is set to 20 pixels. These numbers provided more signal to noise ratio during testing.

- *Peak estimation:* To determine the peak position of stripe with sub-pixel accuracy, we used center of mass method [7]:

$$y_{peak} = \frac{\sum_{i=m}^{m+w} iI(i)}{\sum_{i=m}^{m+w} I(i)} \tag{3}$$

where i is the starting row of the laser stripe and w is the width of the detected laser stripe and $I(i)$ is the intensity of that particular color channel at pixel i within each column.

3.2.3 Color texture image extraction

Simultaneously with the laser stripe detection, we can also extract a high resolution color image of the impression evidence. This is done by extracting the image line perpendicular to the scanner motion direction in the middle of each video frame. The color image captured by this process has the following properties:

- The spatial resolution of the image along the scan direction is dependent on the speed with which the linear actuator moves. Therefore, the resolution along this direction can be increased by scanning the impression using a slower motion of the actuator. Since the video is being captured at a fixed rate of

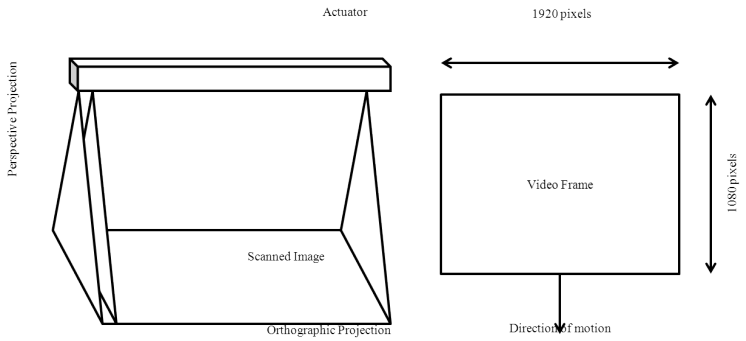


Figure 8: Properties of scanned image.

30 frames per second, the distance between the scan lines of successive video frames will be smaller as the actuator moves more slowly. The image formed along the scan direction is an orthographic projection determined by the scan motion.

- In the direction perpendicular to the scan motion, the resolution is a perspective projection with the spatial resolution determined by the highest resolution of the HD camcorder. The size of a highest resolution video is 1920×1088 pixels. In order to maximize the resolution of the resulting image, the camcorder is oriented such that the highest resolution dimension of the video frames (i.e., 1920 pixels) is perpendicular to the actuator motion.

These are summarized in Figure 8.

3.3 CALIBRATION OF THE TRIANGULATION CONFIGURATION

We are using an L-shaped calibration object (shown in Figure 10) with known dimensions to calibrate the geometric configuration of the laser beams and the camera in order to compute the height map image. The system captures the calibration object in at least one on-site scanning task, i.e., when it's first taken off the truck and set up. This eliminates the risk that a pre-calibrated configuration can be invalidated during the transportation and set-up of the device. Instead, the scanning of the evidence and calibration of the configuration is self-contained and the only requirement in the field is that the criminalist places the calibration object in the scene for the first scan.

One of the parameters of the device that is used in the calibration and that also affects the resolution of the scanned image is the speed with which the camera-laser configuration is moved along the rail. Because the camera is capturing 30 frames of video per second, the faster the camera moves, the less the resolution of the computed image is in that dimension. We obtain the highest resolution height map image when the rail is moving at the slowest speed.

We measured the speed (v mm/s) of the rail when it was moving. The camera calibration gives the focal length (i.e. f) in pixel units. Currently we operate the system in its slowest speed of 1.3138 mm/s.

We performed calibration of the video camera in order to correct for lens distortion at widest zoom angle settings, and found that f is 2110 in pixel units. Moreover, we apply this radial distortion correction to every frame we capture from the camera and this corrected image is used for all subsequent depth calculations.

We use a vanishing point method to find the pose of the camera with respect to the rail. Then we use the calibration object as seen in the image to calculate the orientations of the red and green laser planes.

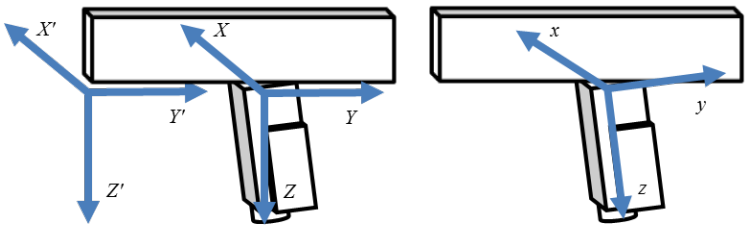


Figure 9: Figure showing various coordinate system definitions.

Everything, including the points on the evidence surface is eventually calculated in a single coordinate system which is defined by the rail. The details of each of these calculations are described below.

3.3.1 Coordinate System Definitions

We use three main coordinate systems (see Figure 9).

1. o-xyz: original image camera coordinate system. Coordinates are measured in pixel units after calibration of camera intrinsic parameters.

2. O-XYZ: ideal image coordinate system. The coordinate system after correcting roll and tilt. Y axis is parallel to rail. Coordinates are in pixel units.

3. O'-$X'Y'Z'$: rail-based 3D world coordinate system. The coordinate system sharing the same axes of ideal image coordinate system. Coordinates are in metric units.

3.3.2 One Time Calibration of the System

We first perform an off-line calibration to estimate the optical properties and radial distortion of the camera. Radial distortion refers to the variation in magnification of the image from the center of the optical field to the periphery. Typical examples of this are barrel and pincushion distortions and distortions in images obtained using fish-eye lenses. This calibration is done for the video camera once at the beginning to obtain the intrinsic camera parameters such as focal length (in pixels), camera center position, and radial distortion. This assumes that during the scan the camera zoom is fixed. This calibration can also be performed in the lab in advance if the zoom factor is determined to be fixed at widest angle and confirmed at the camera buttons before the scanning takes place. This calibration can even be done only once if the same video camera is used for a long period. This is different than calibrating for the camera and laser light geometric configuration which is done as an integrated part of the scans and is described in more detail in Sections 3.3.3 to 3.3.9.

The camera calibration uses a standard function in the OpenCV software package. A checkerboard grid pattern is shown to the video camera at different angles. Several frames with different poses of the board are captured by the video camera. The software detects the corners and computes the camera focal length, optical center in the image frame, and the radial distortion parameters of the lens. With these parameters, an image can be transformed to the o-xyz coordinate system.

The motor translation speeds are also measured in the lab. Using the video camera to shoot the moving assembly through two specified positions on the rail, we can calculate the precise translation velocity of the

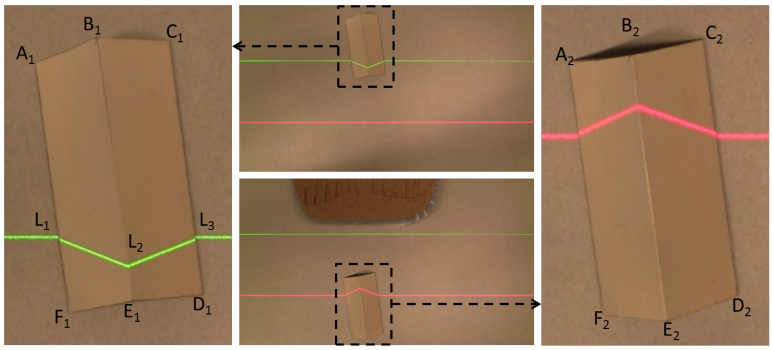


Figure 10: A frame (first frame) is captured when green laser scan over the calibration object at time t_1 (top-center). Another frame (second frame) is captured when red laser scan over the calibration object at t_2 (bottom-center).

cart (excluding the accelerating and stopping periods) by counting the video frames (each frame takes 1/30 sec).

3.3.3 Calculating Vanishing Point of the Linear Motion

Ideally, the camera should be connected to the rail as shown in Figure 4, looking downward and the y-axis of the image aligned perfectly with the rail's direction of motion. Achieving this physically, however, is not realistic. Therefore, we assume the camera is put on the rail roughly as described above and the small misalignments are accounted for via a calibration procedure. This calibration calculates the exact pose of the camera with respect to the rail coordinate system. It uses a method based on computing the vanishing point in order to correct the misalignment of the coordinate system. When two parallel lines are projected in an image, their projections intersect at a finite point unless the lines are exactly parallel to the image plane or to the optical axis. This point of intersection in the image plane is called the "vanishing point." This is the point we typically see when we look at the converging railroad tracks in an image.

To calibrate our system, the calibration object is placed in the scene roughly pointing in the direction of the rail motion. Assume the camera is perfectly aligned with the rail coordinate system. If we translate the camera along the rail and take two images, one before and one after the translation, corresponding points on the calibration object in the two images will form parallel lines. This will result in a vanishing point formed by the lines at infinity. In reality, the camera is not perfectly aligned with the rail system resulting in the vanishing point to be finite. We use this fact to estimate the camera pose with respect to the rail coordinate system from the calculated vanishing point.

We capture two frames as seen in Figure 10, one at $t = t_1$ when the green laser is projected onto the calibration object and the second at $t = t_2$ when the red laser is projected onto the calibration object. We mark the corners of the calibration object $(A_i, B_i, C_i, D_i, E_i, F_i, \text{ for } i = 1, 2)$. This is done via an interactive interface developed in the software that lets the person doing the computations in the crime lab pick these points. The following are the steps for achieving this calibration.

3.3.4 Calculating the Vanishing Point

The lines with which the vanishing points are calculated are defined by two sets of corresponding points between the two frames captured during the camera motion. One set of corresponding points are A_1 and

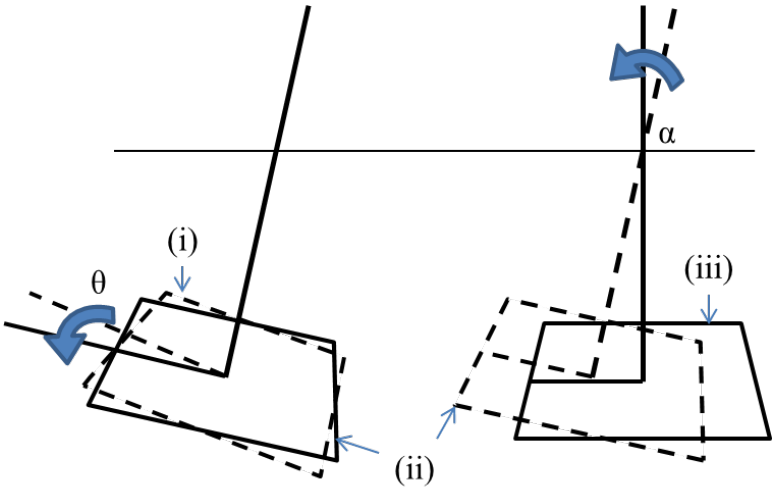


Figure 11: Image Coordinate System (i), Roll Corrected Image Coordinate System (ii) and Ideal Image Coordinate System (iii).

A_2, and the other set is C_1 and C_2. Let the line A_1A_2 be $y = m_1x + c_1$ and the line C_1C_2 be $y = m_2x + c_2$. We consider only the x and y pixel locations of A_1, A_2, C_1, and C_2, and by computing the intersection we obtain the vanishing point (x_v, y_v).

$$\begin{pmatrix} 1 & -m_1 \\ 1 & -m_2 \end{pmatrix} \begin{pmatrix} x_v \\ y_v \end{pmatrix} = \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}$$

From which we obtain

$$\begin{pmatrix} x_v \\ y_v \end{pmatrix} = \begin{pmatrix} 1 & -m_1 \\ 1 & -m_2 \end{pmatrix}^{-1} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}$$

3.3.5 Computing the pose of the camera from the vanishing point (for O-XYZ transformation)

Camera roll (θ) around its optical axis is given by,

$$\theta = \tan^{-1}\left(\frac{x_v}{y_v}\right) \tag{4}$$

Camera tilt (α) between optical axis and the line goes through the optical center and also perpendicular to linear motion.

$$\alpha = \frac{y_v}{|y_v|} \tan^{-1}\left(\frac{f}{\sqrt{x_v^2 + y_v^2}}\right) \tag{5}$$

3.3.6 Coordinate System Transformation (from o-xyz to O-XYZ)

Let the vector x be the coordinates of a point in o-xyz and the vector X be the coordinates of a point in O-XYZ.

The goal is to transform all quantities into the ideal coordinate system as shown in Figure 11. This is given by the following equations:

$$\mathbf{X} = \mathbf{R}_x(\alpha)\mathbf{R}_z(\theta)\mathbf{x} \tag{6}$$

Equation (4), gives the roll angle in CW direction. We need to rotate the coordinate system CCW by this angle. Hence,

$$\mathbf{R}_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Equation (5), gives the tilt angle in CCW direction. We need to rotate the coordinate system CW by this angle. Therefore,

$$\mathbf{R}_x(\alpha) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & \cos\alpha & \sin\alpha & 0 \\ 0 & -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Therefore the total transformation matrix, $\mathbf{T}$, for correcting the roll and tilt, is given by

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & \cos\alpha & \sin\alpha & 0 \\ 0 & -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

3.3.7 Computing Rail Coordinates of a Point on the Calibration Object

Transforming A_1 and A_2 to corresponding ideal image coordinate systems when $t = t_1$ and $t = t_2$, respectively.

$$\begin{pmatrix} X_{a_1} \\ Y_{a_1} \\ Z_{a_1} \\ 1 \end{pmatrix} = \mathbf{T} \begin{pmatrix} x_{a_1} \\ y_{a_1} \\ z_{a_1} \\ 1 \end{pmatrix} \quad \begin{pmatrix} X_{a_2} \\ Y_{a_2} \\ Z_{a_2} \\ 1 \end{pmatrix} = \mathbf{T} \begin{pmatrix} x_{a_2} \\ y_{a_2} \\ z_{a_2} \\ 1 \end{pmatrix}$$

After we apply the roll and tilt correction, the transformed coordinates of the points A_i may not be parallel to the motion direction. Therefore, we project these points to a plane parallel to the motion direction. We use the $Z = f$ plane. This is done by projecting a ray from the origin (camera's optical center) via the point in the ideal image coordinate system.

Considering the first frame, i.e., when $t = t_1$, we obtain

$$\frac{Z_{a_1}}{f} = \frac{Y_{a_1}}{Y}$$

and

$$\frac{Z_{a_1}}{f} = \frac{X_{a_1}}{X}$$

Therefore, the projection of A_1 onto the $Z = f$ plane, denoted by A_{1f}, is

$$A_{1f} \equiv \left(\frac{X_{a_1}}{Z_{a_1}} f, \frac{Y_{a_1}}{Z_{a_1}} f, f \right) \tag{7}$$

Similarly, the projection of A_2 onto the $Z = f$ plane, denoted by A_{2f}, is

$$A_{2f} \equiv \left(\frac{X_{a_2}}{Z_{a_2}} f, \frac{Y_{a_2}}{Z_{a_2}} f, f \right) \tag{8}$$

3.3.8 Calculating the 3D Coordinates in Rail Coordinate System of the Points on the Calibration Object

By using two frames with a certain interval on the rail, we can estimate 3D position of a sample point covered by both frames. Assume the distance travelled between the first frame (at time $t = t_1$) and the second frame (at time $t = t_2$) is d (in metric units). Then by similar triangles, we obtain the Z value for point A in the rail coordinate system as:

$$\frac{Z'}{f} = \frac{d}{(\frac{Y_{a_2}}{Z_{a_2}} f - \frac{Y_{a_1}}{Z_{a_1}} f)}$$

and, therefore,

$$Z' = \frac{d}{(\frac{Y_{a_2}}{Z_{a_2}} - \frac{Y_{a_1}}{Z_{a_1}})} \tag{9}$$

Now considering edges of a triangle we can calculate the X and Y coordinates of the points,

$$\frac{Z_{a_1}}{Z'} = \frac{X_{a_1}}{X'}$$

Also,

$$\frac{Z_{a_1}}{Z'} = \frac{Y_{a_1}}{Y'}$$

Therefore, the coordinates of the point A_1 at time $t = t_1$ are given by:

$$A_1' \equiv \left(\frac{X_{a_1}}{Z_{a_1}} Z', \frac{Y_{a_1}}{Z_{a_1}} Z', Z' \right)$$

Points corresponding to B_1, C_1, D_1, E_1, and F_1 are similarly computed in the rail coordinate system. The position and pose of the calibration object is completely known. The accuracy of the positions is further improved using the physical dimensions of the calibration object.

3.3.9 Finding the laser plane

Let's consider the green laser plane. First, we transform the pixel locations of L, M, and N (in Figure 10) to the ideal image coordinate system using Equation (6).

We project rays, starting from optical center, through all transformed points. Next we compute the intersection of the ray and calibration object edges to find the points where laser touches the calibration object. Since edges are in rail coordinate system, we are getting laser points in rail coordinate system. Finally, using these points, green laser plane is computed.

We find the red laser plane using the same method with the red laser points analogous to L, M, and N. We perform all the steps above when $t = t_2$ coordinates systems.

Assume a point on the laser plane is $\mathbf{P}_a$, and its surface normal is $\mathbf{N}$, using the vectors $\overrightarrow{LM}$ and $\overrightarrow{MN}$ we have the equation of the laser plane normal given as the cross-product:

$$\mathbf{N} = \overrightarrow{LM} \times \overrightarrow{MN}$$

And the equation of the laser plane, for any point $\mathbf{P}$ on it, is given by the dot-product:

$$\mathbf{N} \cdot (\mathbf{P} - \mathbf{P}_a) = 0$$

3.4 CALCULATING THE HEIGHT MAP

Each detected laser pixel (x, y) in a frame, is transformed to the ideal image coordinate system. Through that point we project a ray starting from the optical center. The ideal image coordinate system and the rail coordinate system share the same axes and origin, but they may be at a different scale. Therefore, finding the intersection of the ray and a laser plane gives a rail coordinate of the laser pixel directly.

By applying the offset between red and green laser coordinate systems, i.e. the translation transformation is along y axis — d as in Equation (9), we bring the red laser plane to $t = t_1$ rail coordinate system. This way, a ray and laser plane intersections always provide registered results for both red and green lasers points. This makes the later fusion of the two height maps easier.

For any point $\mathbf{P}$ on the laser plane (lit by the laser stripe), its 3D position satisfies:

$$\mathbf{N} \cdot (\mathbf{P} - \mathbf{P}_a) = 0$$

$$\mathbf{N} \cdot \left(\left(\frac{X(t)Z'}{f}, \frac{Y(t)Z'}{f}, Z' \right) - \mathbf{P}_a \right) = 0$$

where $\mathbf{P}_a$ is a point on the laser plane. From this, Z' can be computed as

$$Z' = f \frac{\mathbf{N} \cdot \mathbf{P}_a}{\mathbf{N} \cdot (X(t), Y(t), f)}$$

and

$$X' = X(t) \frac{\mathbf{N} \cdot \mathbf{P}_a}{\mathbf{N} \cdot (X(t), Y(t), f)}$$

$$Y' = Y(t) \frac{\mathbf{N} \cdot \mathbf{P}_a}{\mathbf{N} \cdot (X(t), Y(t), f)} + tV$$

where V is the camera translation speed and t is time. The depth calculation is performed in a lookup table so that the 3D transformation from the image coordinates can be performed much faster.

Finally the height map is constructed by choosing the z value as the intensity. In the results section, the z value is mapped into 0-255 range of 8-bit gray level intensity in order to allow viewing the results in common image viewing software such as Adobe Photoshop or Gimp. The depth map has a resolution of (7500×7500) where each pixel corresponds to 0.1mm in size. One intensity level in the depth map corresponds to 0.1mm, and the coarsest level for looking at the global elevation changes starts from 10mm.

The prototype software, however, stores the real heights as real numbers for each pixel (in double data type) in a binary data file. These data files can be used for further processing, such as pattern matching, using the actual metric information.

The color texture map of the scanned part has also the same resolution, stored in another image with points exactly corresponding to the ones in the depth map. The user can compare depth map and the color texture map images to find the shapes of the impression unrevealed in the texture map due to the color and lighting on the ground, and confirm the depth on strange shapes by examining the color texture image.

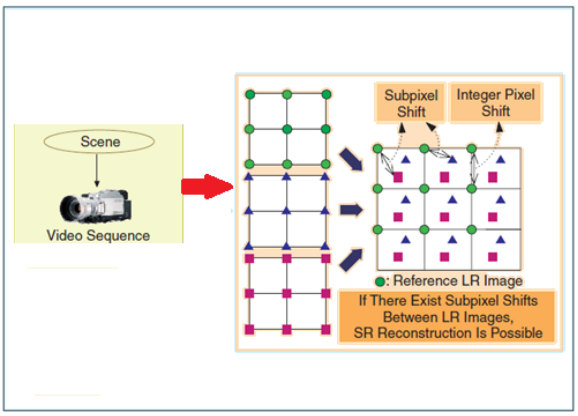


Figure 12: Image Acquisition System.

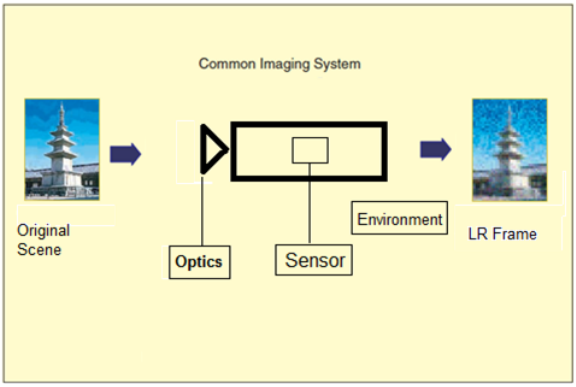


Figure 13: Basics of Super Resolution.

3.5 COMPUTING SUPER RESOLUTION COLOR IMAGE

A further improvement on the resolution of the color image of the evidence digitized can be obtained by using the so-called "Super Resolution" (SR) techniques. Since obtaining the super resolution image will be computationally intensive, we view this as an optional processing that can be turned off if not desired. However, the capability will be there because, as the discussion below indicates, the necessary information would be in the data collected.

This class of techniques relies on multiple, subpixel overlapping lower resolution images in order to estimate a higher resolution image. The major advantage of using signal processing to achieve high resolution is its cost effectiveness because existing low resolution imaging systems can still be utilized. Moreover, the existing highest-resolution hardware available may not be sufficient to satisfy the requirements of a particular application such as forensic imaging. SR image reconstruction is very effective in applications where multiple frames of the same scene can be obtained, and the motion to obtain such images is not very big or is very constrained.

In super resolution, the lower resolution images represent different looks of the same scene. The technique relies on the fact that if the motion of the camera is sufficiently constrained and there is overlap in the pixels of the images of different views of the scene, this overlapped information can be used to recover sub-pixel level image information to compute a higher resolution image. Thus, the low resolution images are subsampled and shifted with sub-pixel precision. Multiple scenes can be obtained from one camera with several captures or from multiple cameras located in different positions. The scene motions can be obtained using controlled motion in imaging systems, e.g., video sequence obtained from a camera mounted on a linear actuator. If the low resolution images are shifted by integer multiples of pixel units, no new information can be obtained to reconstruct the high resolution image because there would be no overlap between pixels of the low resolution images.

During the process of recording a digital image, there is inherent loss of spatial resolution due to optical distortions (out of focus, diffraction limit, etc.), motion blur due to limited shutter speed, noise that occurs within the sensor or during transmission, and insufficient sensor density as shown in Figure 13. Super Resolution also covers image restoration techniques that produce high quality images from noisy and blurred images.

Super Resolution reconstruction is an example of an ill-posed inverse problem [3] as a number of possible solutions exist for a given set of observed images. A common model for SR is stated in the following

way: The low resolution input images are the result of projection of a high resolution image onto the image plane, followed by sampling. The main goal is to find the high resolution image which best fits this model given the observed low resolutions images.

3.5.1 Super resolution adapted for 3D impression digitizing device

Most Super Resolution reconstruction algorithms assume general motion and try to estimate the motion as well as the higher resolution image from the set of low resolution images. However, in our case we have a priori knowledge about the precise motion of the imaging device. We utilize this knowledge to simplify the process of estimating the higher resolution image. In our design of the impression scan device, the motion is linear and the successive frames of the video will possibly have pixel overlaps. Depending on the speed with which the camera assembly is moved along the rail and the height of the camera with respect to the surface being scanned, the amount of this overlap in pixels may vary. Moreover, this overlap in pixels will be only along the axis aligned with the direction of the camera motion. Unlike general Super Resolution algorithms, we do not need to estimate the motion because we know precisely the parameters of this motion. However, as discussed earlier in Section 3.3.5, the camera may not be completely aligned with the coordinate axes defined. The result of this roll and tilt calibration can be used to correct the image in each frame so that the images are properly aligned and then the pixel overlap along the direction of the motion can be used to improve the image resolution in that direction.

3.5.2 Pre-Processing for Roll and Tilt Correction

The next step after extracting successive frames is to compensate for the distortions introduced due to roll and tilt of the camera. Before any SR reconstruction can take place it is necessary to compensate for the roll and tilt of the camera as described in Section 3.3.5 and bring the frame into an ideal image plane. From this point forward all input images are assumed to have been passed through such a correction for roll and tilt.

Since the images have been compensated for roll and tilt it is now possible to treat each column as an independent vector aligned with the direction of the linear motion. The problem of SR reconstruction of the whole image can now be treated as a set of independent 1D SR reconstructions of the image columns along the motion.

3.5.3 Least Squares Formulation and Estimation of Higher Resolution Image

Given K lower resolution images $\{\mathbf{X}_L^{(n)}\}_{n=1}^K$ of size $M_1 \times 1$ pixels find the higher resolution image $\mathbf{X}_H$ of size $N_1 \times 1$ pixels (where $N_1 > M_1$) which minimizes the squared error function:

$$E(\mathbf{X}_H) = \sum_{n=1}^{K} \|P_n(\mathbf{X}_H) - \mathbf{X}_L^{(n)}\|^2 + L\delta\|C\mathbf{X}_H\|^2 \tag{10}$$

where

1. $\|\cdot\|$ can be any norm, but in our case is L_2.
2. $P_n(\mathbf{X}_H)$ is the projection of $\mathbf{X}_H$ onto the coordinate system and the sampling grid of image $\mathbf{X}_L^{(n)}$.
3. $\mathbf{C}\mathbf{X}_H$ is called the edge preserving term. It is a penalty term to ensure we do not smooth out the edges. It penalizes the energy function if we try to smooth over the edges. $\mathbf{C}$ is a simple $[1, -1]$ kernel designed to get an edge response.

4. δ is a scalar constant which is used to assign weight to the penalty term. It is usually less than 0.1.

5. $\mathbf{L}$ is called a line process. It is a boolean array which turns off the penalty term for certain areas. $l_i = 1$ indicates that there is a discontinuity in the sub-interval $x = [i-1, i]$. The diagonal entries of $\mathbf{L}$ are set to $(1 - l_i)$ to indicate if there is a discontinuity at that location and if the penalty term needs to be turned off so the estimated HR image is not smoothed over the discontinuity. $\mathbf{L}$ ensures that the penalty term is applied only in areas having edges which is determined by an edge detection operator such as the Sobel edge detector.

The projection $P_n(\mathbf{X}_H)$ is modeled by four stages:

1. Geometric Transformation,

2. Blurring,

3. Subsampling,

4. Additive noise.

Most modern algorithms differ on the optimization technique used for solving the equation, the constraints on $\mathbf{X}_H$ added to the system and the modeling of the geometric transformation, blur and noise.

3.5.4 Image Registration Model

Geometric Transformation Coordinate system of $\mathbf{X}_H$ has to be determined in order to have a unique $\mathbf{X}_H$. Usually the coordinate system of one of the input image is multiplied by the factor q. The geometric transformation of $\mathbf{X}_H$ to the coordinate system of the input images is computed by finding the motion between the input images. High accuracy of motion estimation is crucial for the success of super resolution.

The motion between input images is governed by the linear translation of the actuator. Since the camera moves only along the Y-axis it is possible to accurately model the geometric transformation. Depending on the amount of sub-pixel overlap depending upon the speed of the actuator and height of the object from camera.

Blur Image blur is usually modeled using convolution. A low pass kernel is used for convolution. We have to take into account blur caused by optics as well as sensors while modeling the blur. In order for the super resolution problem to be uniquely solvable we need to have an accurate kernel.

Subsampling The subsampling matrix maps the high resolution image $\mathbf{X}_H$ to the input low resolution images. Consider a high resolution image $\mathbf{X}_H$ of size $N_1 \times 1$. $\mathbf{X}$ in this case is the ideal un-degraded image of a continuous scene. It is sampled at or above the Nyquist rate.

Additive noise It is assumed that the noise is additive and normally distributed with zero mean. It thus becomes possible to find the maximum likelihood solution by minimizing the error function. The assumption of normal distribution of the noise is not accurate in most of the cases, as most of the noise in the imaging process is non-Gaussian (quantization, camera noise, etc.), but modeling it in a more realistic way would end in a very large and complex optimization problem which is usually hard to solve [19].

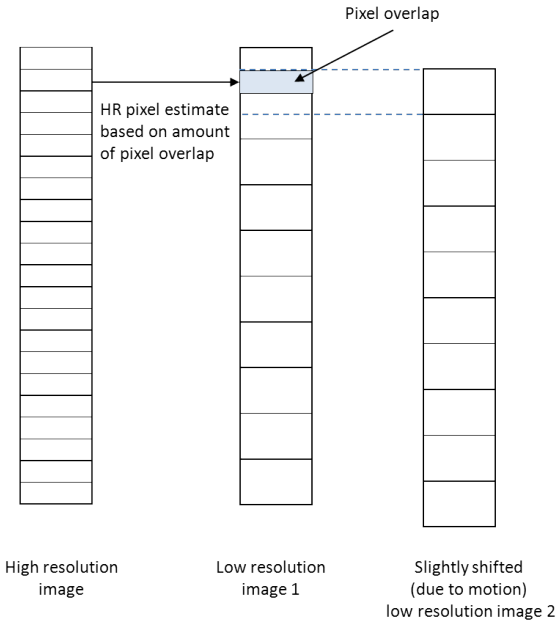


Figure 14: High resolution Geometric Transformation from multiple low resolution images.

3.5.5 Sub-pixel Overlap Estimation

In order to accurately model the geometric transformation it is necessary to accurately compute the amount of sub-pixel overlap. A computed spatiotemporal profile computed from the video can be used to help with this analysis. A spatiotemporal profile of a video is an image composed by cutting the video along the time axis. This is done by extracting one column at the center of each frame and stacking them along the time axis. This results in a collection of one-dimensional (1D) image data accumulated along the time axis, which can be viewed as a two-dimensional spatiotemporal profile image (see Figure 15). This profile can be analyzed to estimate the amount of pixel overlap in the lower resolution images due to the camera motion.

The lines in the profile represent the edges of the objects in the scene during the camera motion. These object edges change position in the image based on the speed with which the camera is moving. The angles of these edges in the spatiotemporal image are correlated to the amount of sub-pixel overlap in successive video frames. The magnitude and direction of the vectors orthogonal to these edges can be used to calculate the amount of sub-pixel overlap.

The profile of a video is passed through filters that give the partial derivatives along the X and Y axes. Let $\mathbf{I}$ be the profile image of a video. Then $\mathbf{I}_x = \partial\mathbf{I}/\partial x$ and $\mathbf{I}_y = \partial\mathbf{I}/\partial y$ represent the partial derivatives of $\mathbf{I}$ along the X and Y axes, respectively.

Magnitude E_{mag} and direction θ can be calculated using the following equations.

$$E_{mag} = \sqrt{\mathbf{I}_x^2 + \mathbf{I}_y^2} \text{ and } \theta = \tan^{-1}\left(\frac{\mathbf{I}_y}{\mathbf{I}_x}\right)$$

$\theta = \pi/4$ represents a sampling rate equal to the speed of translation and implies no sub-pixel overlap. Thus, $\theta \leq \pi/4$ implies there is no sub-pixel overlap and thus no SR reconstruction is possible. On the other hand, $\theta = \pi/2$ implies that the pixels overlap and hence there is no sub-pixel overlap between successive frames. Thus for SR reconstruction, we need $\pi/2 > \theta > \pi/4$. Based on various heights and linear speeds,

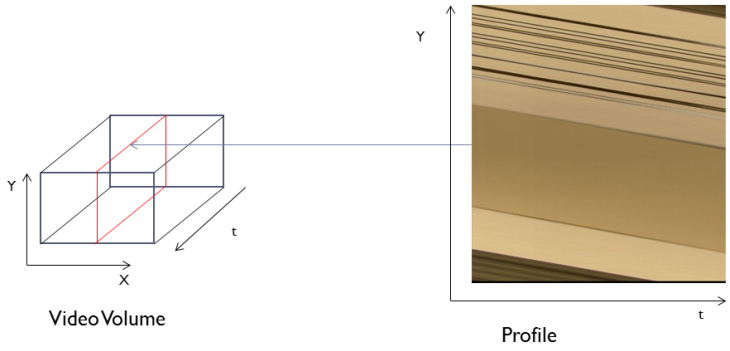


Figure 15: Generating the spatiotemporal profile of the captured video.

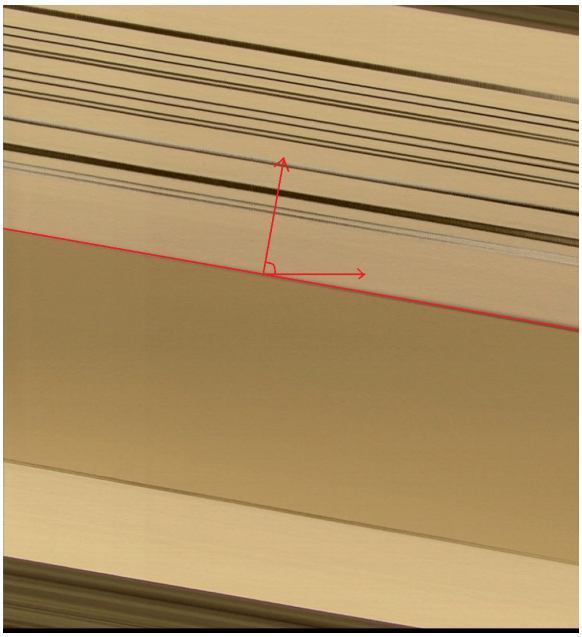


Figure 16: Profile of a video with distance from camera $= 20.5''$ and speed $= 1$. In this example, $\theta = 84.1^{\circ}$

Distance from camera (inches)	Speed (pulses/sec)	Threshold (top 5%)	θ	percentage overlap
20.5	1	122	84.1657	87.03
	2	110	63.84	41.87
	3	51	52.2	16.00
	4	39	42	-6.67
18	1	74	64.4	42.11
	2	51	55.5	23.33
	3	40	49.4	9.78
	4	43	39.2	-12.89
15.5	1	49	51	13.33
	2	26	44	-2.22
	3	30	40	-11.11
	4	23	37	-17.78

Table 2: Sub-Pixel overlap based on Speed and Distance from Camera

Table 2 shows the percentage sub-pixel overlap between successive frames. It is evident that we have no overlap or negligible sub-pixel overlap for heights greater than 18".

3.5.6 Error Minimization

It is critical that the initial estimate be as close to the actual solution as possible to ensure that the Energy function finds the correct solution. One of the important steps to create the initial estimate is to calculate the percentage of sub-pixel overlap between two successive frames. Based on this percentage of sub-pixel overlap the higher resolution image is estimated. An overlap of approximately 75% implies that three low resolution pixels contribute to one higher resolution pixel and thus we can have 3 times improvement in resolution. On the other hand an overlap of approximately 50% implies two low resolution pixels contribute to one higher resolution pixel and thus we can have 2 times improvement in resolution. Initial estimate of each high resolution pixel consists of average of contributing low resolution pixels. Random noise, having varying standard deviation, is added to the initial estimate.

After the initial estimate of the HR image is generated, we refine the estimated solution by minimizing the edge-preserving error function given in Equation (10) using a standard optimization method. The solution for SR reconstruction equation is to find the $\mathbf{X}_H$ which minimizes the scalar energy function E while not smoothing over the discontinuities. An example intermediate case is shown in Figure 17 to illustrate the calculation.

3.6 USER INTERFACE

In order for the field practitioner to operate the device and perform a scan, we have implemented two different methods of user interface. The first method is based on four physical buttons on the device whose functions are pre-defined. These functions and the number of buttons is limited by the hardware interface of the actuator rail but makes the device and interaction with it self-contained. The second method is a wireless control through a tablet PC with which the field practitioner can interact with the device. This method provides more flexibility and allows the practitioner to change parameters (e.g., rail motion speed)

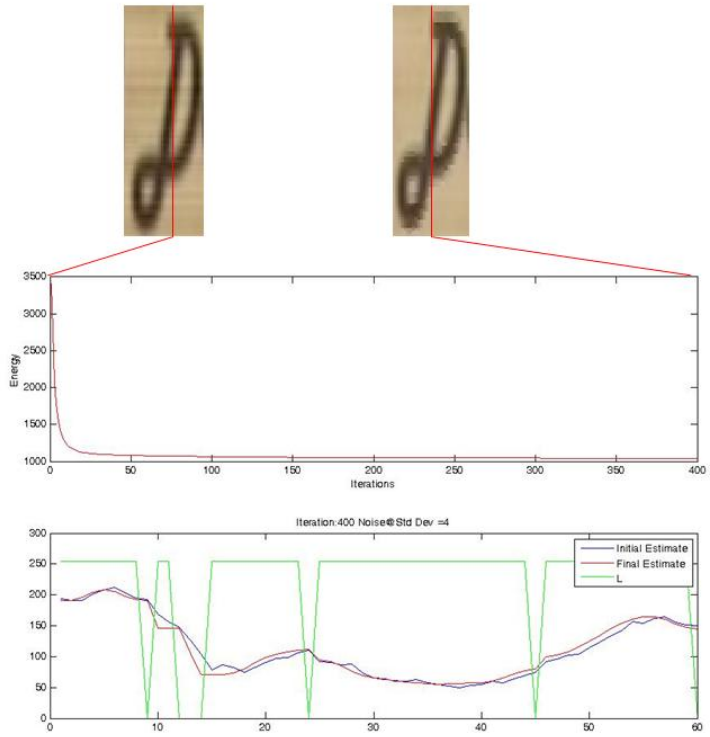


Figure 17: Error Function minimization for Additive Noise with Standard Deviation $= 4$. From the above figure we can see that the smoothing is not applied over edges. We illustrate in details the estimation of pixel values along one column (the red column of the top images) in the example images. The top left image is the initial HR estimate and the top right image is the estimate after 400 iterations. The top plot shows the change in energy over 400 iterations. The bottom plot shows 1D profile of the pixel values and discontinuities (green plot) along the column. As can be seen the HR is refined while the discontinuities are not smoothed over.

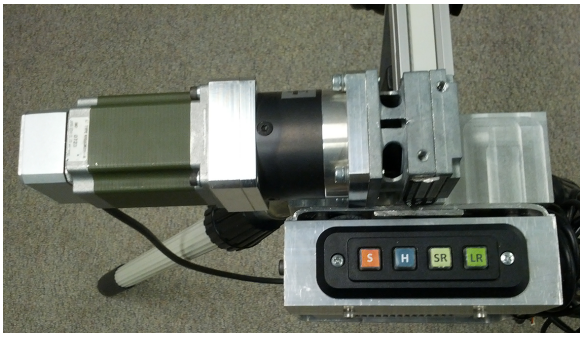


Figure 18: Four button control pad of the device.

through a software interface.

3.6.1 Physical Button User Interface

The Myostat Cool Muscle (CM1 Series) Integrated Servo motor supports 4 inputs and 2 outputs that are user definable. Each I/O port can be assigned multiple functions by setting motor parameters. To make the device self-contained, we have programmed the motor to accept four button control pad as shown in Figure 18. Four buttons are as follows.

- S: Emergency Stop button.
- H: Home button.
- SR: Short Run button. SR button starts from the home position and moves 1/3 of the total length of the device.
- LR: Long Run button. LR button starts from the home position and moves the total length of the device.

3.6.2 Android Tablet based Wireless User Interface

We have designed a wireless user interface using a Android tablet. The tablet connects to the device using a Bluetooth connection. The Android application as shown in Figure 19 has more control over the device compared to the physical button interface. The block diagram of the system is shown in Figure 20. Using the tablet, operator can perform following.

- Stop the device.
- Home the device.
- Short Run: Scan 1/3 of the total length of the device.
- Medium Run: Scan 2/3 of the total length of the device.
- Long Run: Scan the total length of the device.
- Set 5 pre-defined speed values.

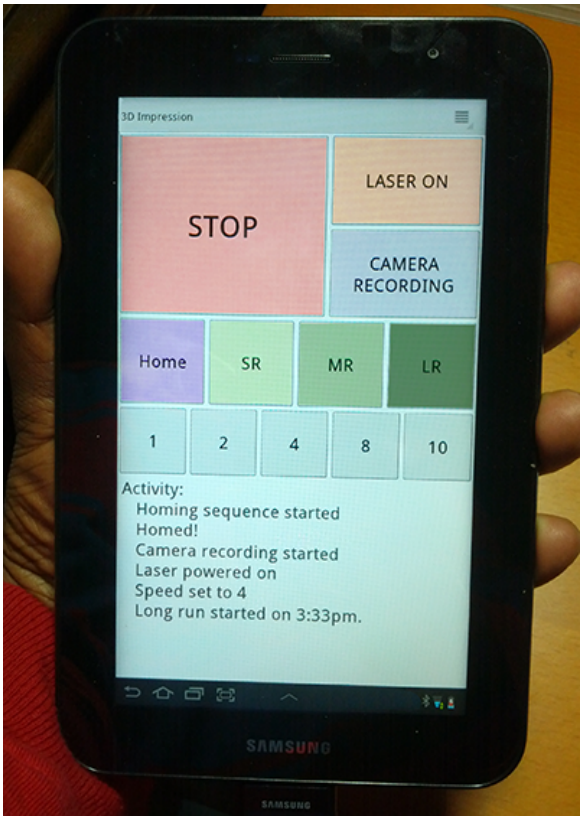


Figure 19: The user interface of Android 7" tablet

- Turn on/off laser modules.
- Start & Stop the camera.

3.6.3 Software, Graphical User Interface

We have developed a software application to calibrate device, to process video and to measure 3D data. We have used technologies, Visual C# .NET and Visual C++. Furthermore, following libraries are also used for the implementation.

- OpenCV : For video and image processing routines.
- Intel Threading Building Blocks : For multi-core programming.
- Point Cloud Library : For processing and visualizing 3D point cloud data.

The software consists of easy to use tools to view and to take 3D measurements as shown in Figure 21.

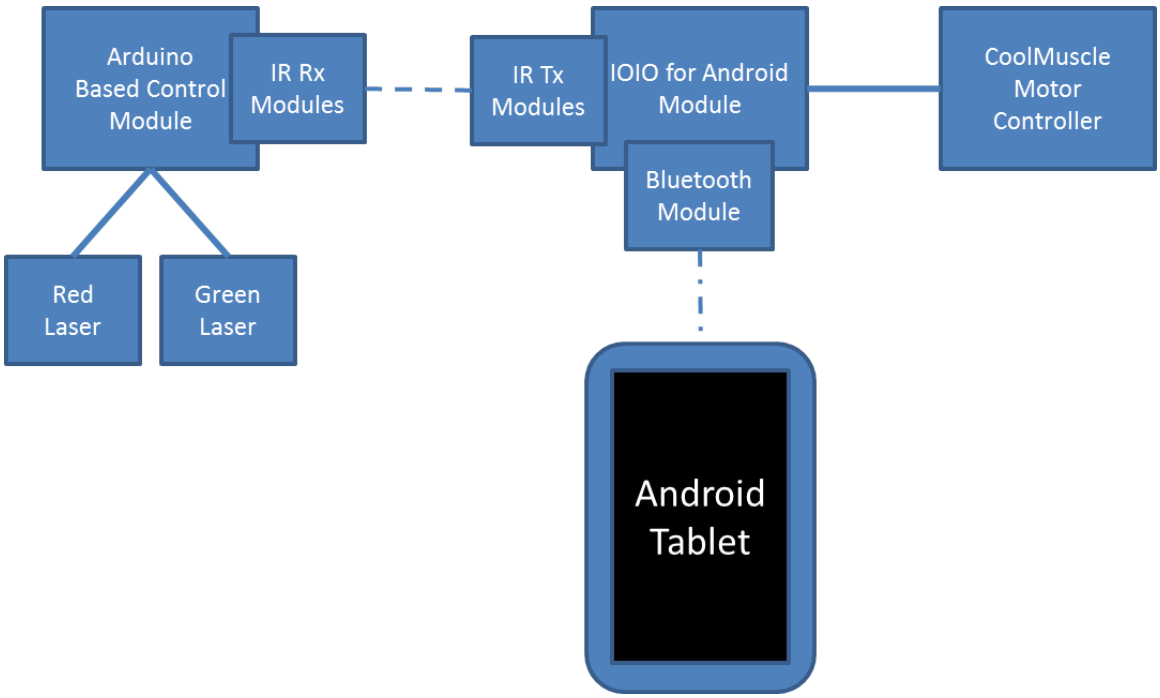


Figure 20: The block diagram of the wireless Android tablet controller design

4 RESULTS

4.1 Results of 3D Scans of Impression Data

We have scanned a tire cast, with the calibration object in the scene. The HD video file from which these height map images were computed consists of a total of 11,146 frames captured in 371.533 seconds. The total size of the video is approximately 1.03 GB and each frame in the video has a resolution of 1920×1088 pixels. The results of processing this video are given below.

First, Figure 22 gives the computed red and green laser displacement (disparity) images. For each laser pixel the displacement is measured using the Equation 11. The black regions in these images are the missing laser data due to occlusions of the laser stripes. Notice the offset because of the different configurations of the two lasers and the complementary locations of the occluded regions from the two lasers.

$$H(x,t) = y - y_{\text{offset}} \tag{11}$$

where y_{offset} is a value defining a window for each laser. Laser pixels are searched only within that window.

From these disparity images and the results of the calibration described above, we can compute the height map at each pixel in the image for the two lasers separately. These results are shown in Figure 23. The size of the final computed height images (given in Figure 23) is 2559×2607 pixels.

Finally, these separately computed height maps are fused into one single height map image. This results in filling the missing height data in the occluded regions using the alternate height map. Since, all coordinates are in a single coordinate system, the estimation of the fused height data is much easier. The final fused results are shown in Figure 24. In order to see the effectiveness of using two lasers in completing the missing data, we show below, in Figure 25, the difference between the two registered height maps from the

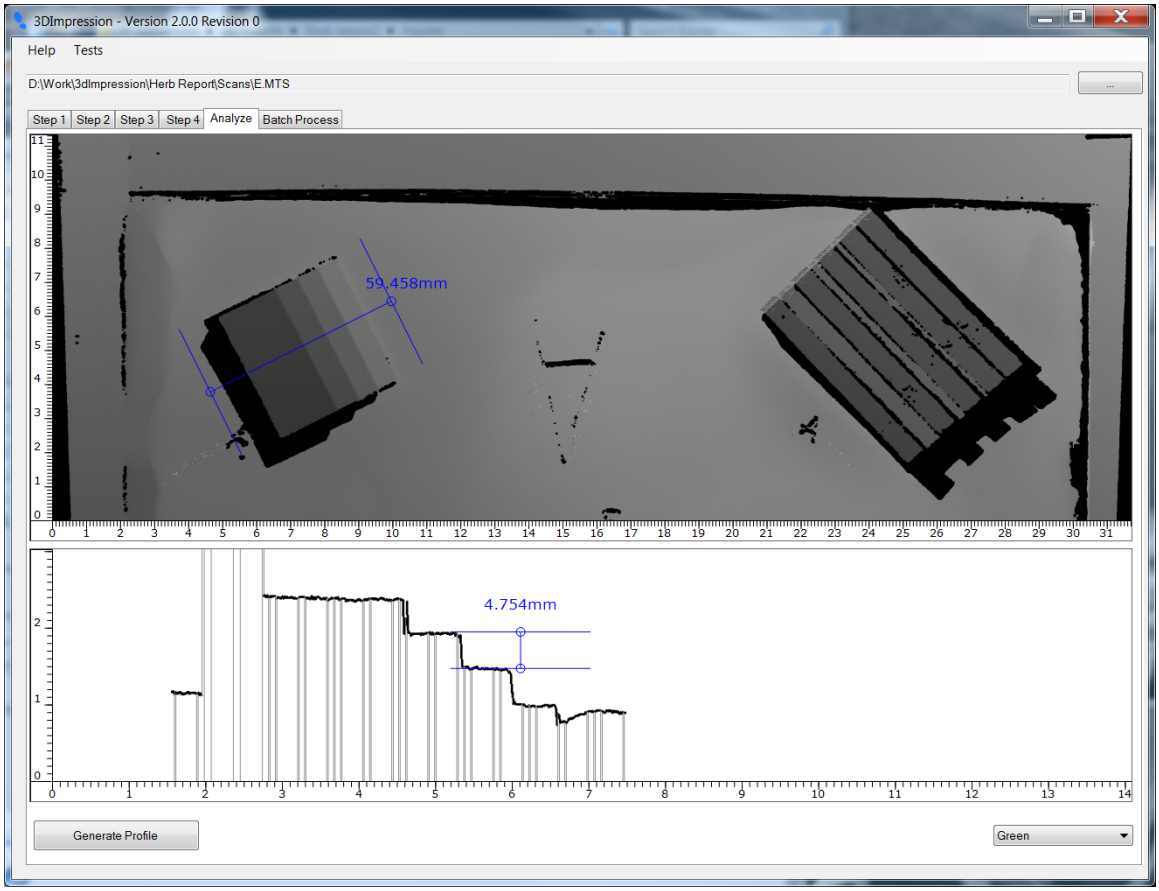


Figure 21: The graphical user interface to capture 3D measurements. Using the blue ruler, one can measure metric distances between two points in the scanned image. In addition, a 1D profile of the depth values along this blue line is plotted along with a tool for measuring depth/height values precisely (bottom figure).

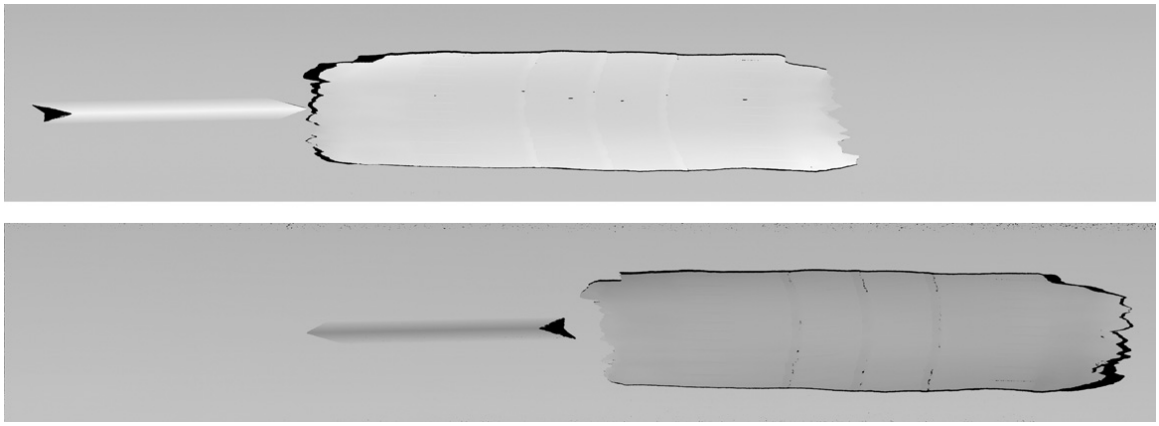
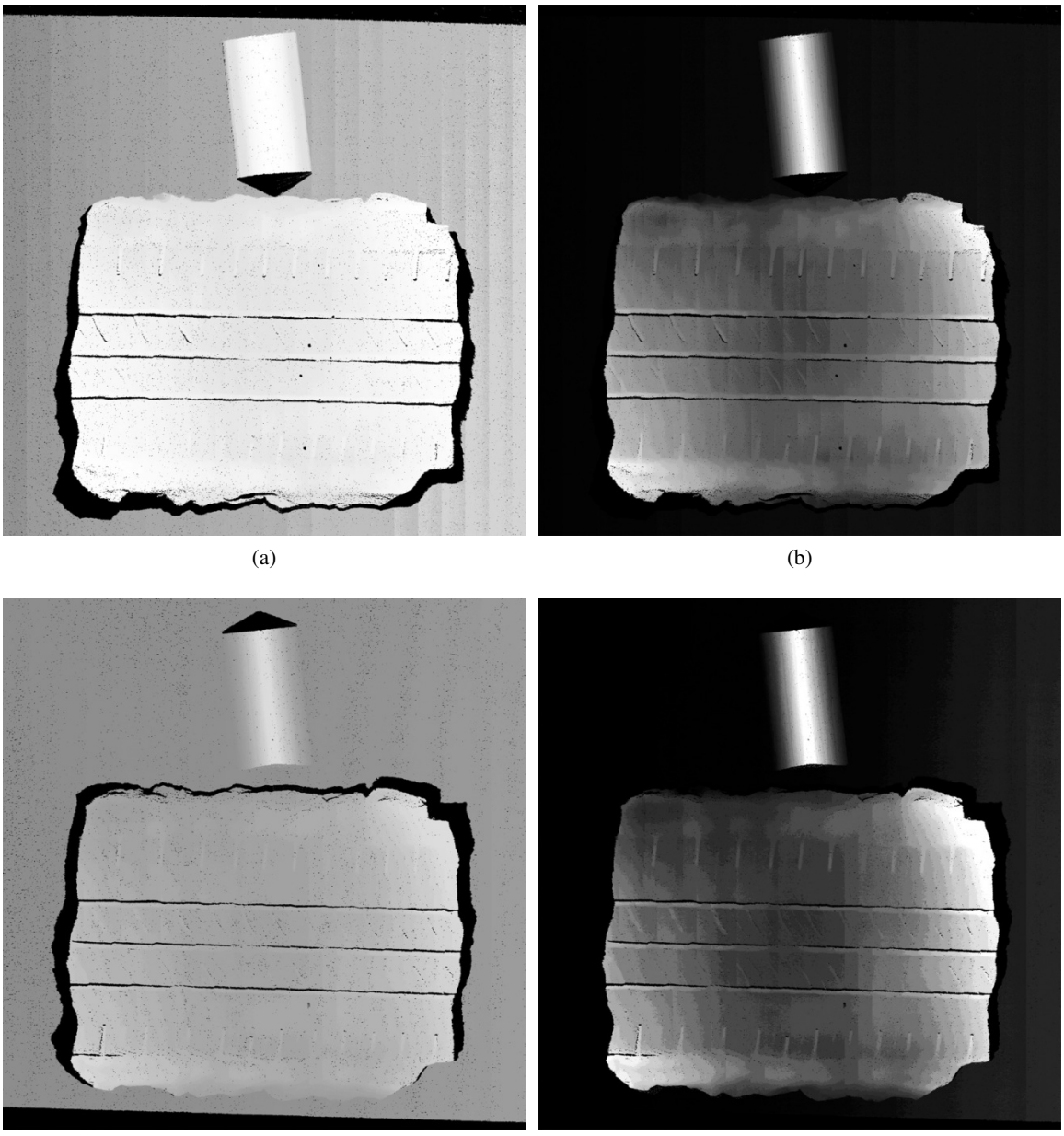


Figure 22: Top: Disparity image calculated from the green laser. Bottom: Disparity image calculated from the red laser



(a) (b)

(c) (d)

Figure 23: Computed three-dimensional height map images from the two lasers. (a) & (b) are height maps computed from the green laser. (c) & (d) are height maps computed from the red laser. (b) & (d) are enhanced versions to make the computed results more visible to the human eye. The height values are shown as brightness values. The brighter pixel value means the higher the height at that pixel.

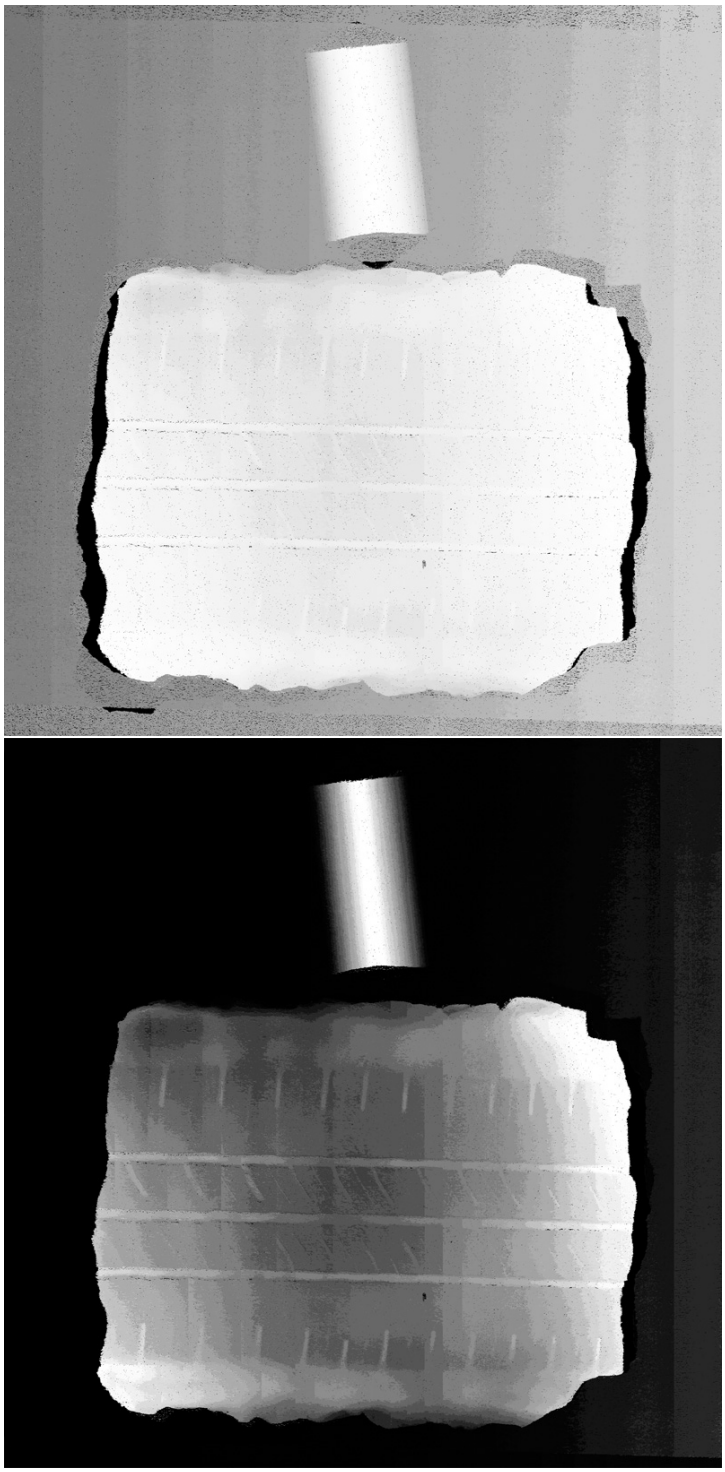


Figure 24: Height map generated by fusing red and green laser height maps. (a) Original fused height map; (b) Enhanced height map for visual clarity

green and red lasers. In most of the pixels, the two height maps agree (therefore, the difference in those pixels is 0). Where one image has data and the other is missing height data due to occlusions, the difference is big, and the fusion fills in these pixels with the appropriate value.

Finally, from the scanned video, we extract a high resolution texture image which is correctly aligned with the computed, fused height map. This result is shown in Figure 26 below.

We also scanned some other images in more natural settings at a speed of 13.138mm/s. For example, Figure 27 shows a shoe print scanned on a lawn. The color texture images are shown in Figure 28. Some more 3D scan results are shown in Figures 31-46. In these pictures, the height range is adjusted to the 8-bit gray level range in order to make the features visible. Figure 49 shows some color texture images that were captured.

4.2 Results of Super Resolution Color Image Scans of Impression Data

A number of experiments were performed to validate the SR reconstruction and to study how various factors affected the performance and resolution of SR reconstruction. Many videos were shot for experimentation. In order to study the improvement is resolution we shot 12 videos with different height and speed settings. Below is a summary of all the different settings used:

- Speed settings: 1, 2, 3 and 4 pps (1 pps = 1.32mm/sec)
- We obtained data for 3 different distances from the camera: $20.5''$, $18''$, and $15.5''$.
- The camera was completely zoomed out with fixed exposure and focus. This was done to ensure that the auto focus and auto exposure do not introduce any noise.
- To each data set we added random Gaussian noise to study the performance of SR reconstruction. In each case the mean of the noise was 0 with varying amount of standard deviation. We used noise with standard deviation as 0, 1 and 4.

4.3 Scan Results of a Shoe with Known Characteristics

Recently, we have been in communication with Mr. Alan Kainuma of SWGTREAD and Honolulu Police Department regarding the capabilities of our impression scan device. As part of this interaction, he has graciously helped us with what features criminalists would like to detect and has sent us a sample shoe with all the features he would like to see in the impression scan (see Figure 54). We scanned the shoe sole with our device as shown in Figure 53. In this figure the heights are mapped to pseudocolor in order to make the height changes more visible. The smaller heights in this picture are closer to the camera. In this test, we specifically checked whether our algorithm can detect known characteristics found on a shoe as shown in Figure 53. Ten known characteristics of the shoe print is shown in the top image of Figure 54. Five out of ten characteristics are visible in the middle image of Figure 54. Most of the remaining characteristics are visible in the texture image found at the bottom of Figure 54. We also made an impression of this shoe sole in a Bubber and some of the very fine details in the photograph of the shoe sole did not transfer to the Bubber. We did scan the impression image and we could detect the larger features (features 1 and 2 as well as the long cut on the left labeled as feature 4.)

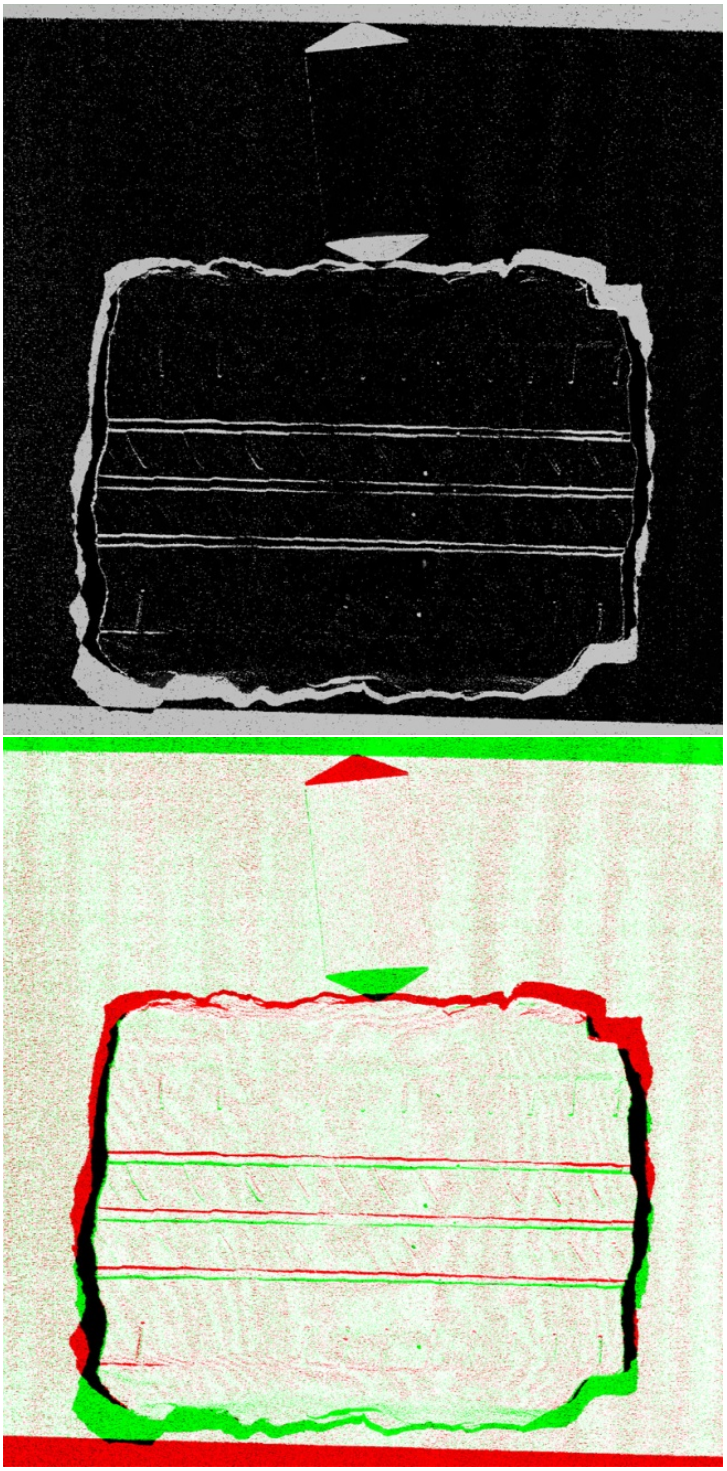


Figure 25: Height map differences and contribution to the fusion. Per pixel differences between red and laser generated height maps (top). Per pixel contributions of each laser labeled (bottom). Green and red pixels denote the contributions of each laser light to the final fused height map where only one laser color was visible. For example, a red pixel indicates that the red laser was detected at that pixel, but not the green one. White pixels are pixels where both green and red contributions are averaged. Black pixels are pixels where no laser readings were found

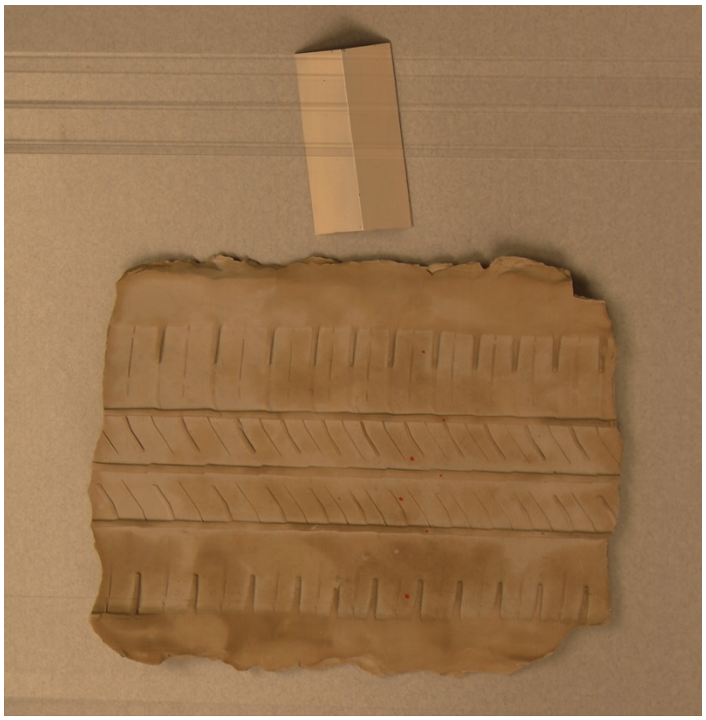


Figure 26: This color texture image is generated by stacking frame stripes (1 pixel thick, horizontal stripes). Furthermore, it is registered with the computed height maps.

4.4 Accuracy and Resolution Measurements for the System

The test setup we have used to determine the accuracy is shown in Figure 55. We created wooden blocks with known dimensions. Next, we have embedded these blocks on to two dental stone plates: plate A and plate B. Targets 1, 5, 7 & 9 have a series of slots that vary in width from about $1/4$ inch to $1/32$ inch. There is $1/4$ inch between each slot. The other targets are each a series of steps, with each step $1/4$ inch wide and $1/4$ inch high.

We have used following test scenarios.

- Plates are positioned in the sequence of A followed by B.
- Plates are positioned in the sequence of A followed by B, but each rotated by 90 degrees. This is in order to test sensitivity of the scan to orientation.
- Plates are positioned in the sequence of B followed by A.
- Plates are positioned in the sequence of B followed by A, but each rotated by 90 degrees.
- Plates are positioned in the sequence of A followed by B, but B is lifted 40mm above the surface of A. This is to test sensitivity of the scan to features at different depths.
- Plates are positioned in the sequence of A followed by B, but both plated tilted by 5 degrees. This is in order to test sensitivity of the scan to features that are on an inclined surface where the depth changes gradually.

Figure 27: Shoe prints scanned in a natural setting. Subimage containing the shoe print is enhanced for visual clarity.

Figure 28: Texture image of shoe prints scanned in the natural setting. The image was acquired during night time.

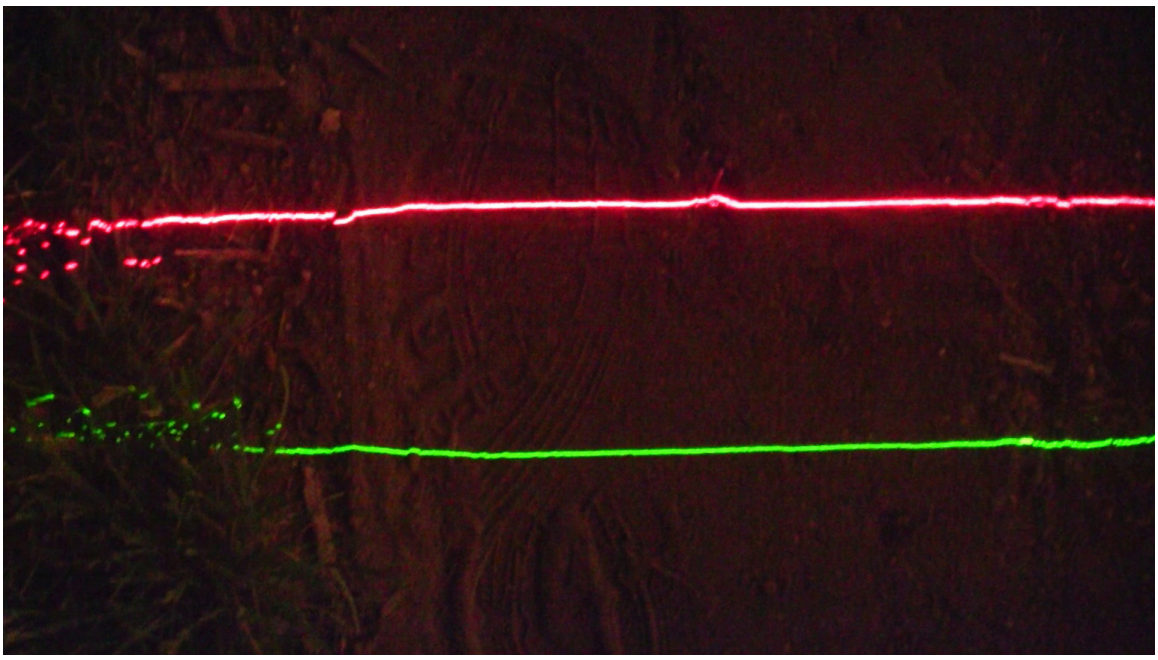


Figure 29: A sample image frame from the scan of Figure 27.

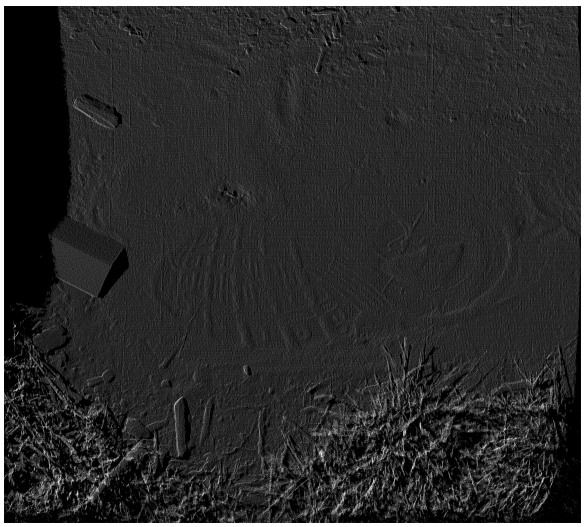


Figure 30: Height map of the shoe print region in the scan of Figure 27.

Figure 31: z range: 305.5mm, 256.0mm. The red pixels indicate locations where we do not have laser data to compute depth values. Brightness indicates height values–the brighter the value the closer the point on the surface is to the cameras optical center (i.e., the larger the height).

Figure 32: A sample frame from the scan of Figure 31

In each of the above scenarios, we computed 4 measurements, using the software we developed, as shown in Figure 56 and Figure 57 for each wooden block. Notice some measurements are approximately along xy plane, and other measurements approximately along z axis. Then we compared them with physical dimensions of wooden blocks, which we measured using a vernier caliper. The resulting absolute error amounts between the measured physical dimensions and those computed from the scanned 3D model are used to calculate means and standard deviations of these errors. Table 3 gives the error statistics (means and standard deviations) of the 208 readings as well as separate statistics for the different test case scenarios. We have calculated these error statistics separately for measurements along the xy plane and along the z axis (depth). Figures 58, 59, and 60 show the 3D point cloud results of the test setup.

We have determined the following accuracy measurements for the system:

- Video camera: frames per second (fps) = 30, focal length (f) = 2110 pixels.
- Rail Speeds in pulses per second (pps) as indicated by motor controller. The pps speed setting corresponds to the following metric speeds:
- For the widest angle of lens corresponding to the lowest image resolution, we have observed empirically that we can resolve 0.5mm in Z' values as reflected in detectable differences of at least 1 pixel in the disparity image computed. We are currently working on obtaining theoretical expressions of the resolution in Z'. With the increasing zoom value (determined by f), the accuracy can be increased to a much higher level because the camera can have up to 10 times zoom. However, this increased zooming may narrow down the field of view. In this report, we give the lowest resolution results.
- At the lowest speed, the device takes approximately 20 minutes to scan a 1.75m long surface (shorter scans such as shoe impressions take shorter amount of time). Processing takes around 1 hour in a eight core Intel® Core™ i7-2700 3.5GHz / 16 GB RAM machine.

Table 4 gives the processing times based on the speeds with which the camera moves on the rail. The slowest speeds result in highest resolutions along the scan direction with the ability to detect smaller features

Figure 33: z range: 292.7mm, 278.4mm. Brighter points are closer to camera.

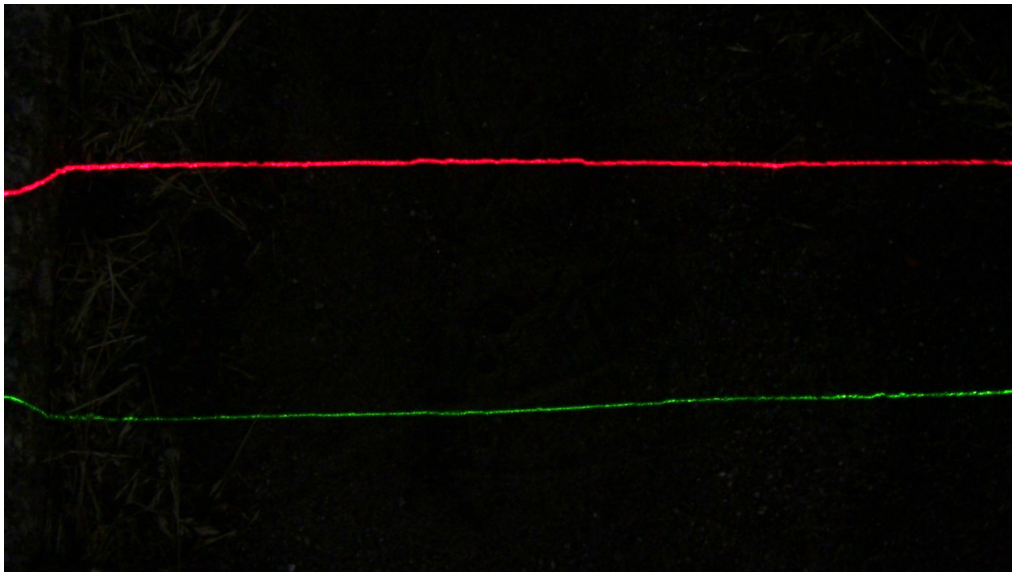


Figure 34: A sample frame from the scan of Figure 33

	Error mean xy (mm)	Error Std Dev xy (mm)	Error Mean z (mm)	Error Std Dev z (mm)
Order AB	0.54	0.34	1.45	0.47
Order AB, Each Rotated 90 degrees	0.74	0.36	1.55	0.34
Order AB, B lifted by 40mm	0.57	0.48	0.78	0.31
Order AB, Each Tilted by 5 degrees	0.56	0.36	0.78	0.32
Order BA	0.63	0.36	0.52	0.23
OrderBA, Each Rotated 90 degrees	0.60	0.40	0.67	0.31
Overall	**0.60**	**0.39**	**0.94**	**0.51**

Table 3: Statistics of absolute errors between measured physical dimensions of the test object and those computed from the scanned 3D model. Means and standard deviations of these errors for the different test cases (rows) have been separated for the measurements along xy plane (columns 2 and 3) and along the z axis (columns 4 and 5).

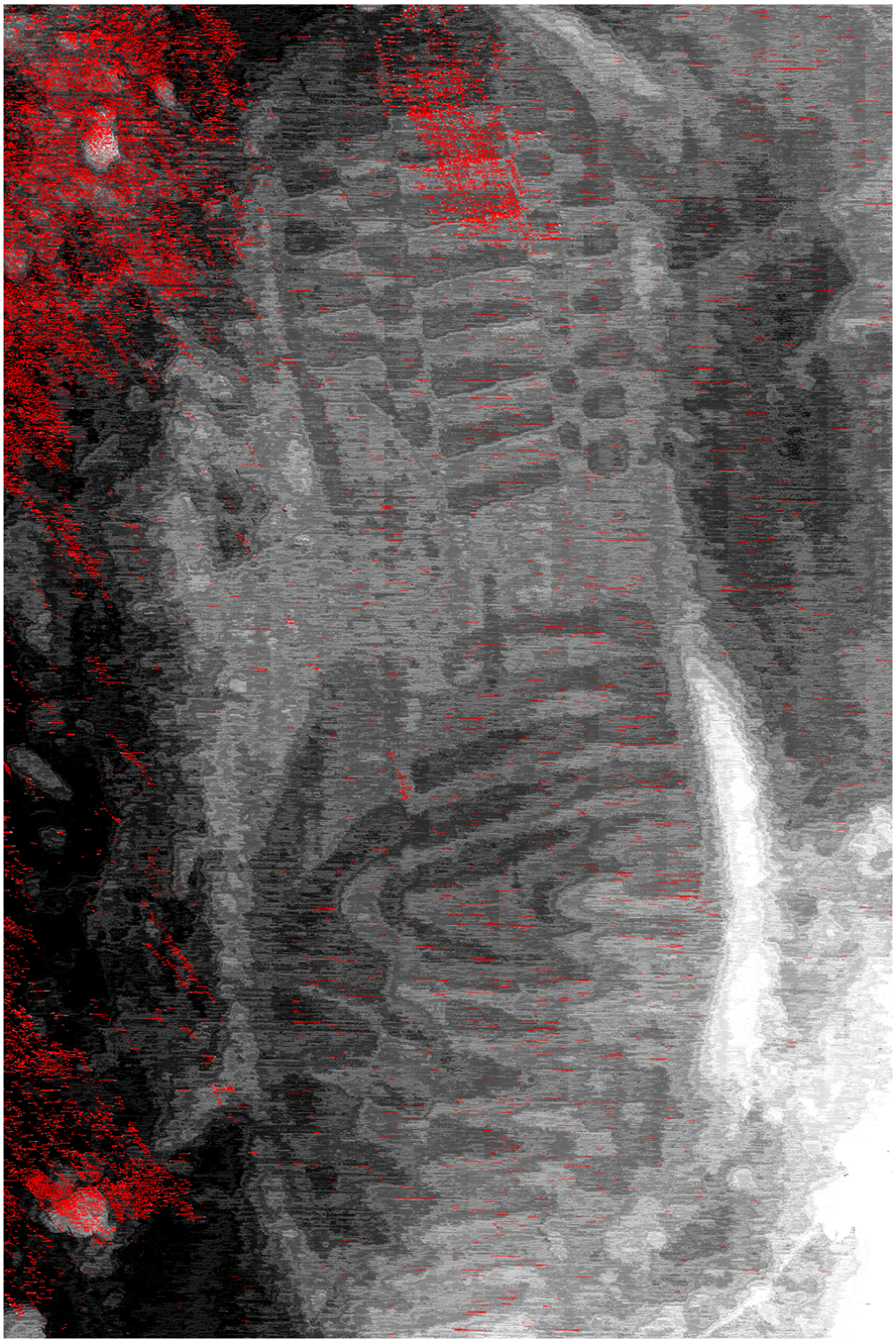


Figure 35: z range: 299.1mm, 265.6mm. Brighter points are closer to camera.

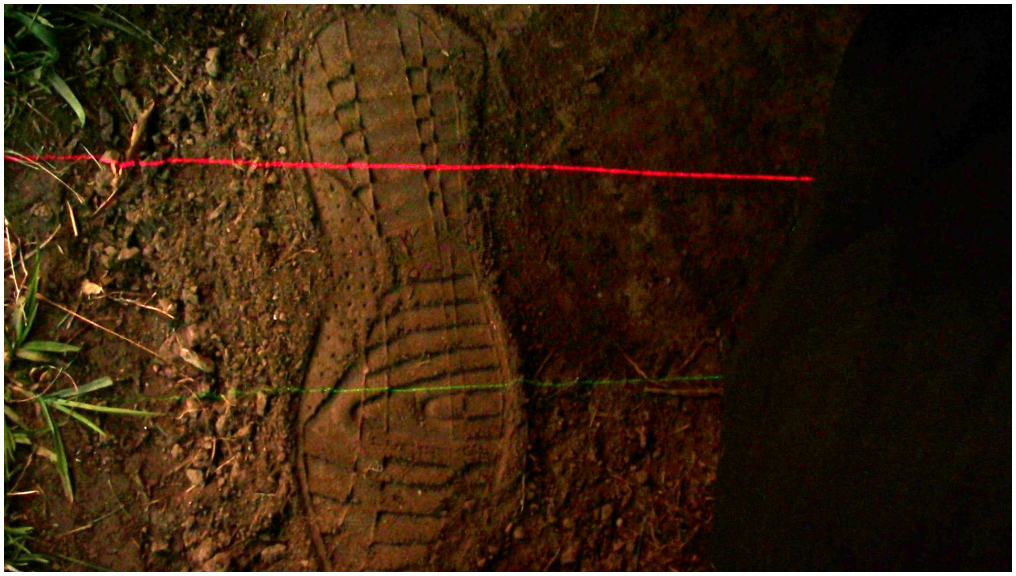


Figure 36: A sample frame from the scan of Figure 35

Scan Speed (pulses/second)	Scan Speed (mm/s)	Processing time (min)
1	1.3138	20
4	5.2552	5
10	13.138	2

Table 4: Processing time for different speeds (Short Scans 500mm)

in the impression. At this speed, however, the computational times will be larger to process the data in order to generate the 3D image. Conversely, the fastest rail speeds will have less resolution but faster computational times.

5 CONCLUSIONS

In conclusion, we have developed an inexpensive high resolution 3D impression device for digitizing footwear and tire impressions in crime scenes. Currently CSI personnel must expend considerable effort to capture proper track information to facilitate comparisons between footwear or tires and track impressions found at crime scenes. The photographic technique requires careful attention to many special issues and the casting of impressions is a science of its own—especially in difficult media such as loose sand or snow. The method developed in this project is a non-contact and non-destructive process and it will often (but not necessarily always) render the making of casts optional. The device is portable with minimal weight.

The device achieves sub-millimeter resolution in depth (0.5mm) and, depending on the imaging geometry and camera motion speed, even higher resolution in the other two spatial dimensions. For some of the scans given in this report, the device achieved a spatial resolution of 0.0438mm in the Y-dimension (the direction of rail motion, at the slowest speed) and 0.2369mm in the X-dimension (perpendicular to the rail

Figure 37: Brighter points are closer to camera.

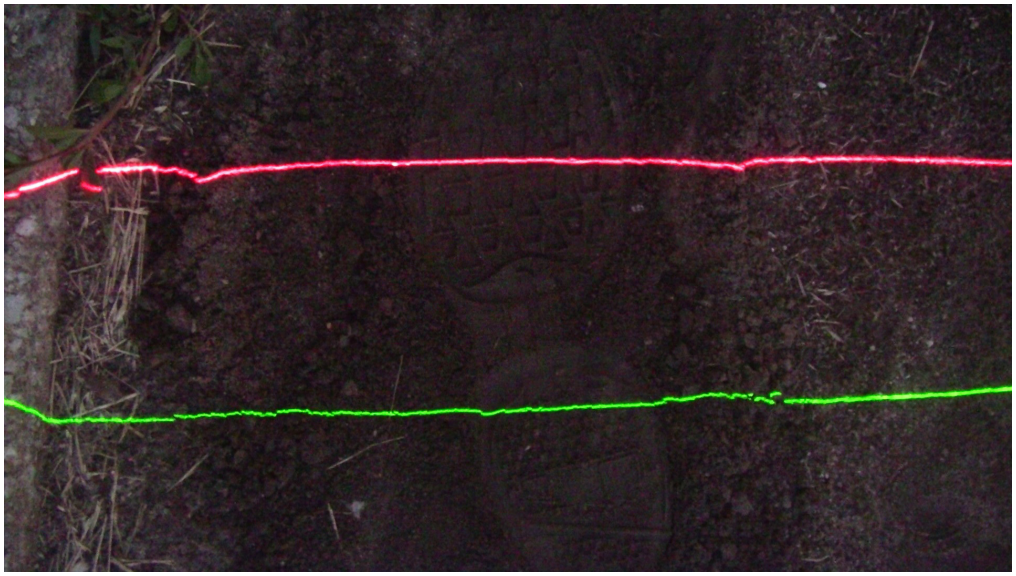


Figure 38: A sample frame from the scan of Figure 37

motion). Some of these values will depend on the speed with which the impression is scanned and the imaging geometry such as the distance of the camera assembly from the evidence, focal length, etc. This project demonstrated a proof of concept that the desired 3D scanning accuracies can be achieved.

In addition to building the hardware of the device, we have also developed the software and a calibration method that is integrated with the scanning process making the task of the field technician much easier. The calibration method we used is integrated in the scanning process and eliminates the requirement of pre-calibrating the system which can become stale in the field due to the pre-calibrated configuration being changed during transportation and setup. We also used two different colored line laser modules to eliminate difficulties due to occlusions.

The prototype device costs about as much as an accessorized digital camera with an associated computer. The total cost for building our prototype device as reported here, including the computers and tablet devices for user interaction with the scanner was around $6500 (see Table 1). We realize that, in order to successfully commercialize this device, the costs will be higher due to engineering, development, and other factors such as personnel, sales, and marketing. We believe, however, that the final cost of such a product will still be competitive with the currently available commercial stand-alone laser range finders which can cost $35,000 on the market. Moreover, unlike our prototype, these currently available range sensors are not necessarily designed to satisfy the accuracy and resolution requirements of forensic imaging applications.

Additionally, when such a system is commercialized, the engineering of the product will address some of the shortcomings of our current prototype. For example, we used the off-the-shelf HD camcorder for the imaging device. However, in a final commercial product, both the camera and the assembly that includes the line lasers can be much more carefully designed to be rugged and with better optical qualities. Other parts of the system such as the legs would probably have to be redesigned by professional engineers to the proper specs to be used in the field by practitioners. The line laser lights may be better selected for quality and intensity for bright light situations. In addition, the system may require a redesign for two pass operation: One using narrow band filters in order to detect the laser lights in bright light situations, and a second pass to capture the color texture image. Even though the mechanical actuators movements are very precisely

Figure 39: z range: 271.9mm, 235.2mm. Brighter points are closer to camera.

Figure 40: A sample frame from the scan of Figure 39

controlled, this may create alignment issues between the two images captured, which would have to be solved. However, we believe this misalignment would be sufficiently small that the software solution would not be very difficult. Finally, a good engineering design would have a user interface that would give more real-time feedback at the scene in case the impression may need to be scanned again. This feedback does can be a lower resolution image to indicate to the practitioner whether the image captured looks reasonable. Once he/she is convinced that the scan is good, the high quality images can be obtained by processing at the lab.

One final issue is how these images may be used for comparison. With the 3D depth data, investigators can further have the depth of the impression on soft ground, in addition to the traditional two-dimensional photographic images. The information upon which comparisons are based becomes richer with proper metric measurements. Future research could try to automate some of this comparison and compute a goodness-of-fit measure to aid the examiner.

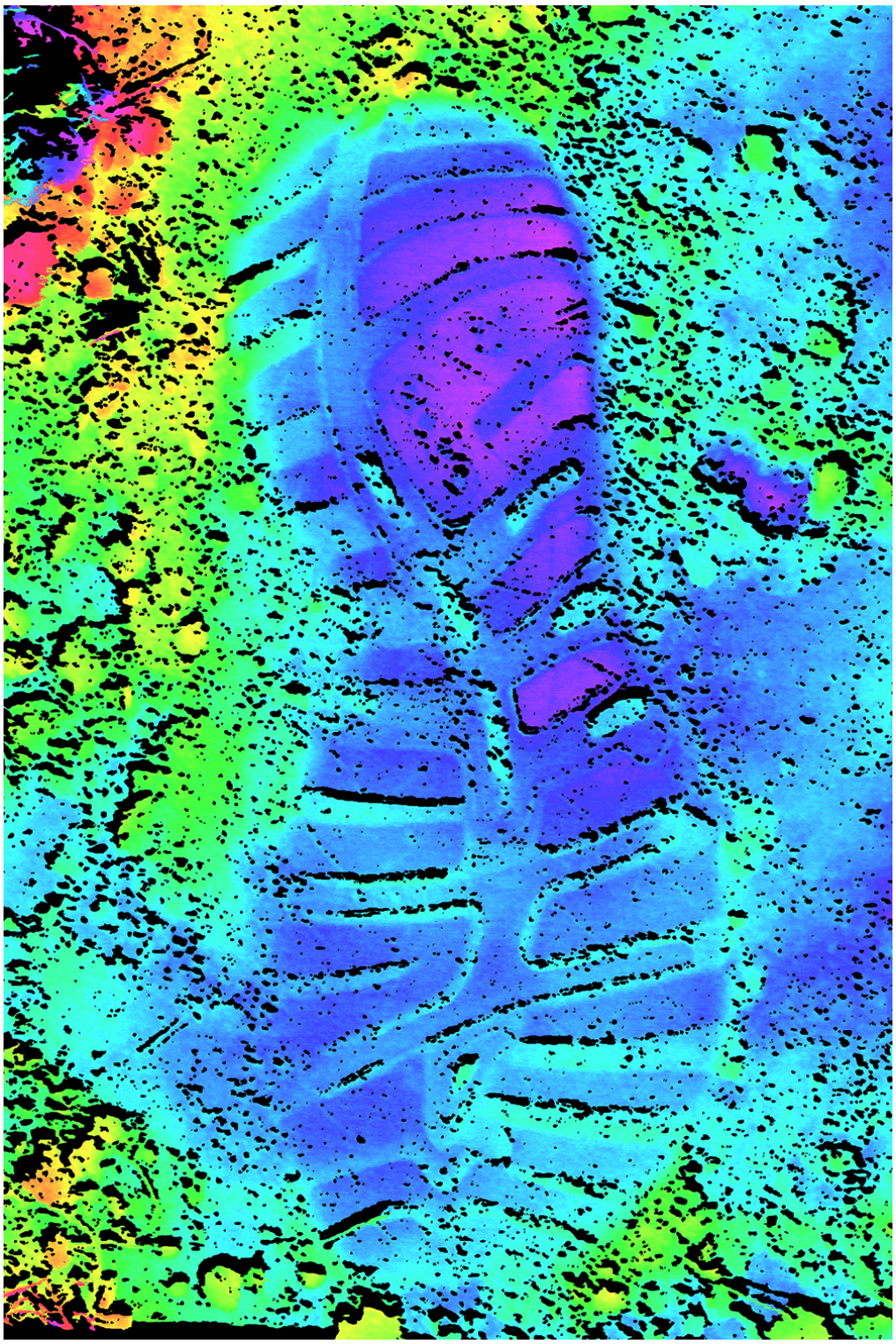


Figure 41: Psuedo-color encoded heightmap of scan of Figure 39. This is another way to visualize the depth values that enables us to see the details in the depth data.

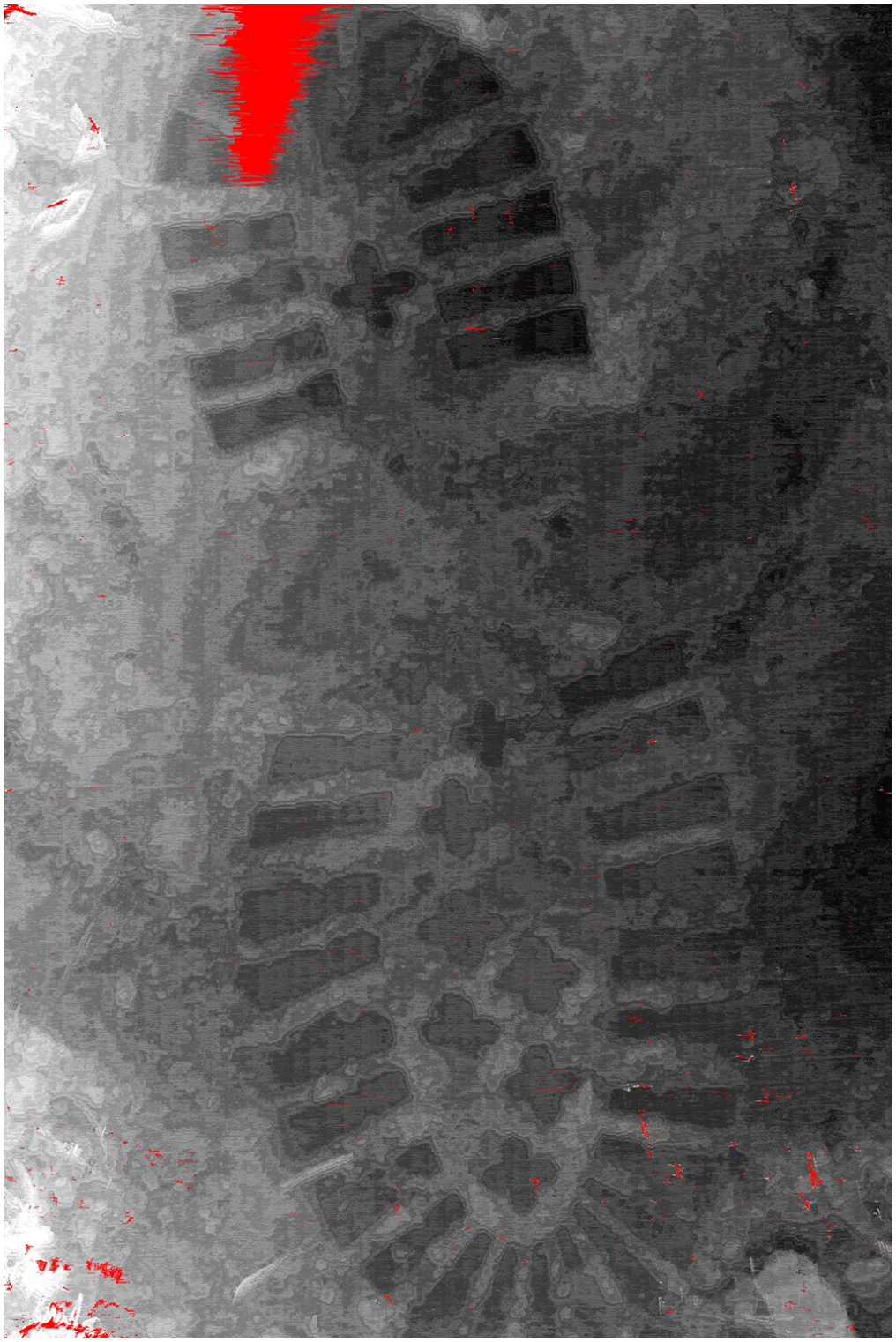


Figure 42: z range: 275.1mm, 241.6mm. Brighter points are closer to camera.

Figure 43: A sample frame from the scan of Figure 42

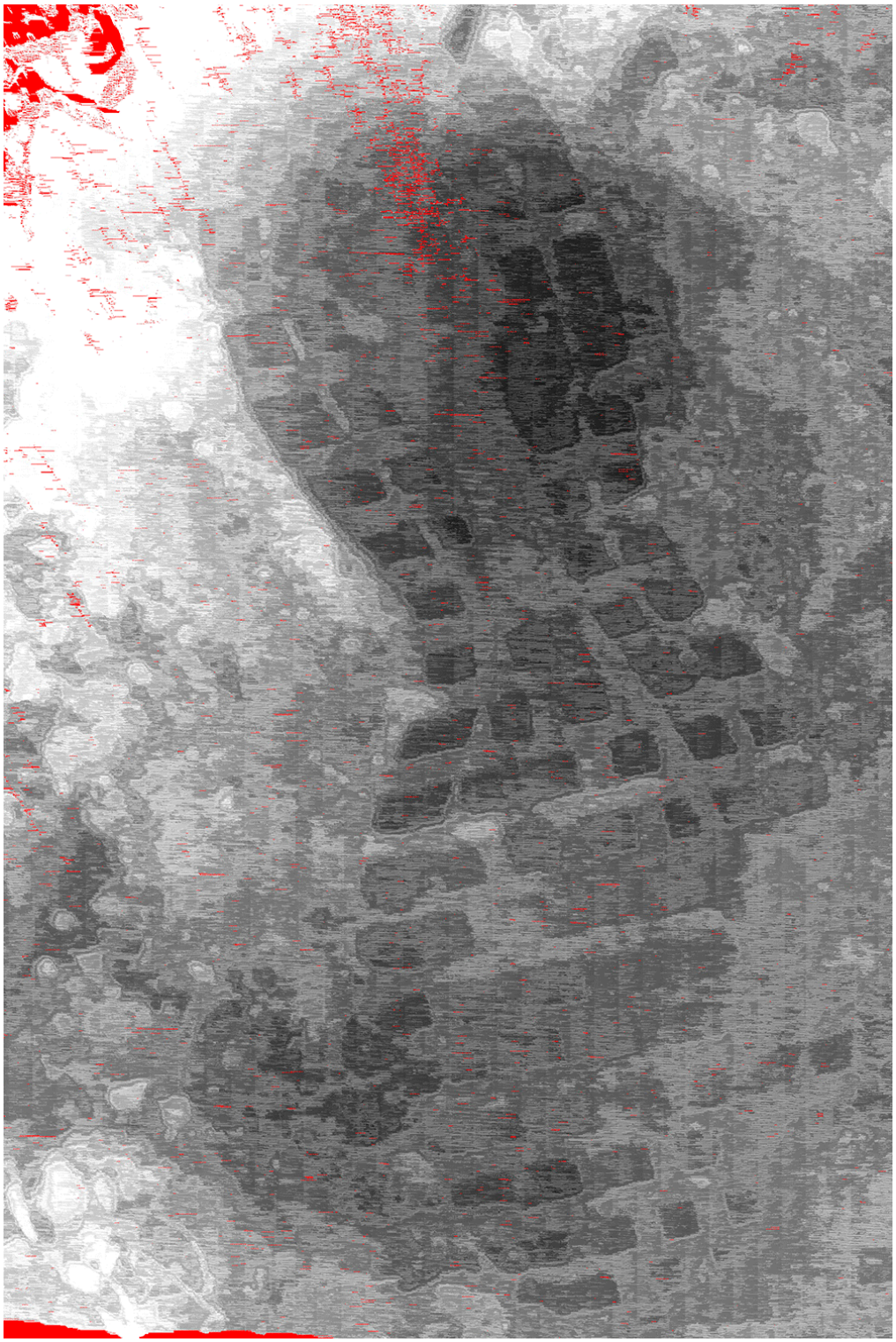


Figure 44: z range: 217.5mm, 192.0mm. Brighter points are closer to camera.

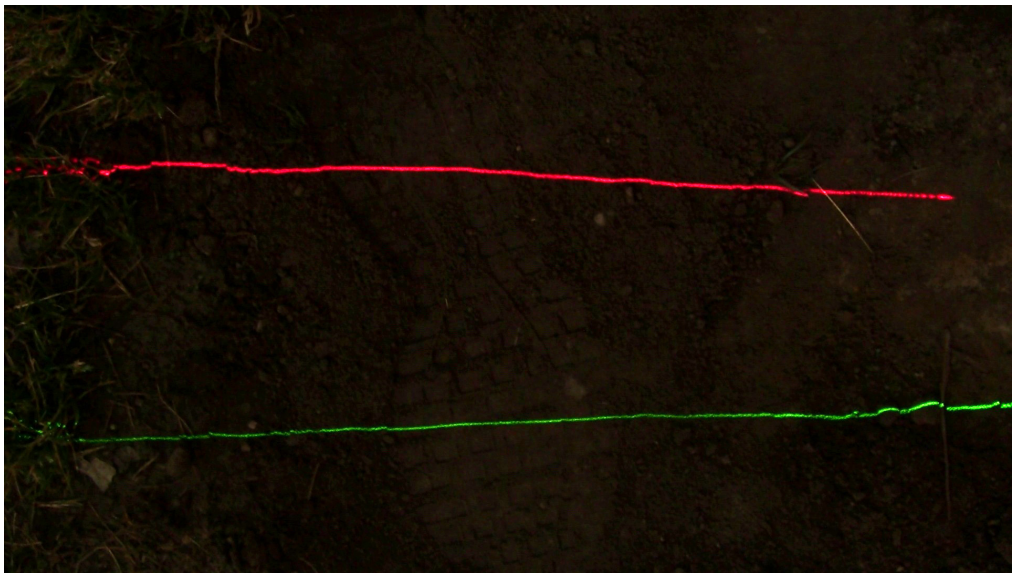


Figure 45: A sample frame from the scan of Figure 44

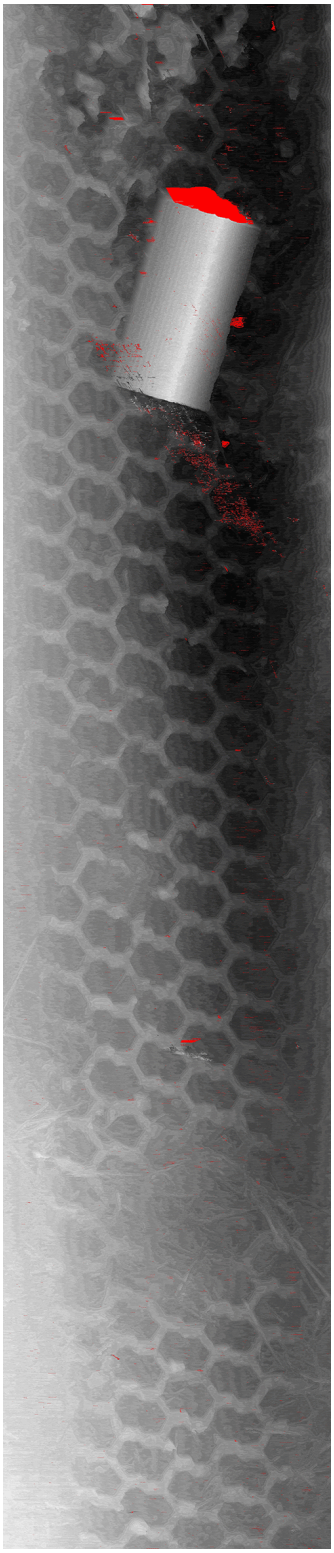


Figure 46: A long span tire impression scanned with our device. z range: 307.1mm, 259.2mm. Brighter points are closer to camera. x span: 169.2mm and y span (which is the length of the impression scanned): 800mm (80cm which is about half of the rail length)

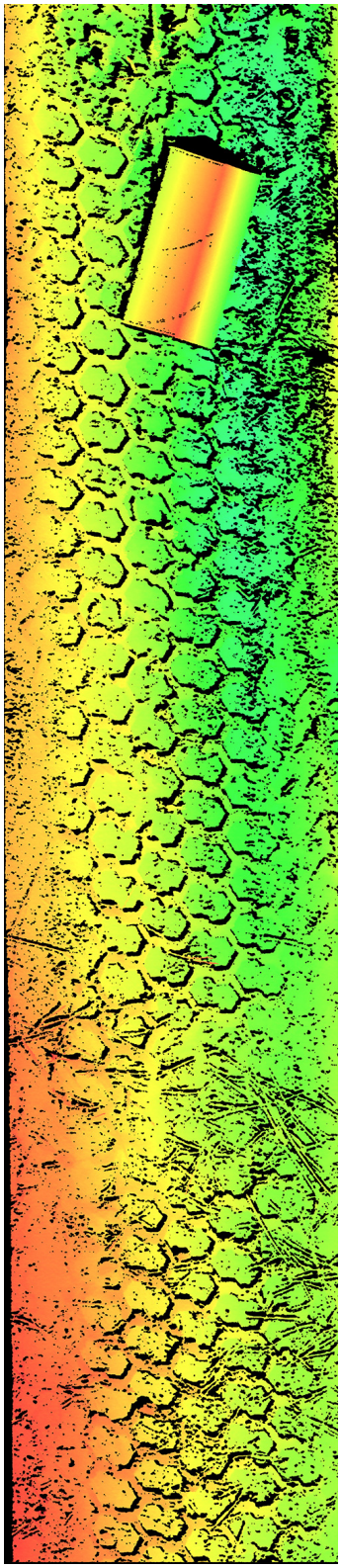


Figure 47: Psuedo-colored heightmap for the scan of Figure 46

Figure 48: A sample frame from the scan of Figure 46

Figure 49: Color texture images captured during the scans along with 3D depth maps.

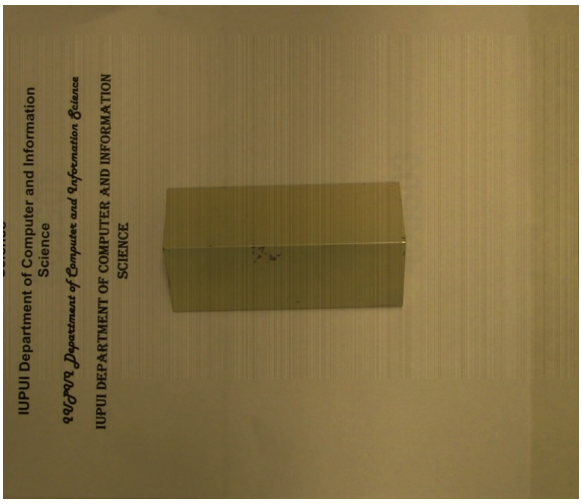


Figure 50: Low resolution (800×1088 pixels) input image

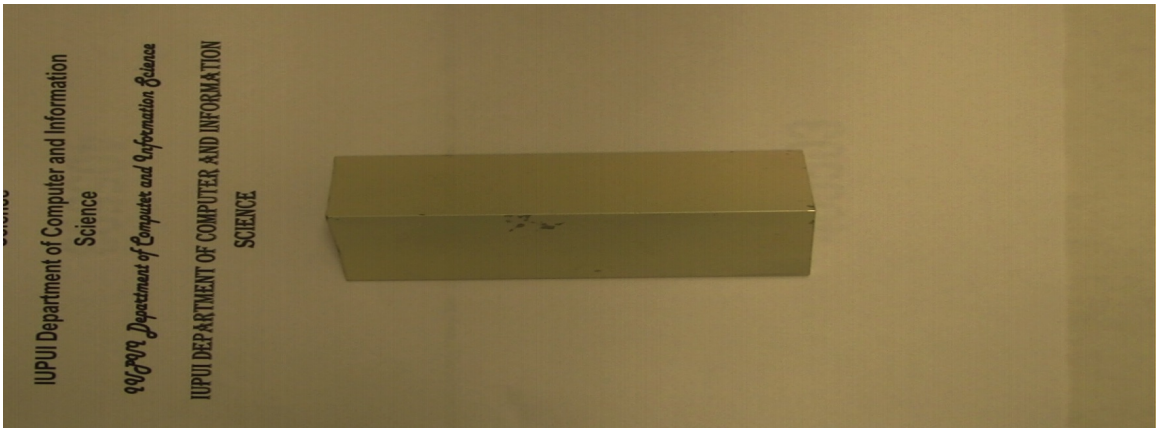


Figure 51: Computed high Resolution Image 800×2176 pixels) image

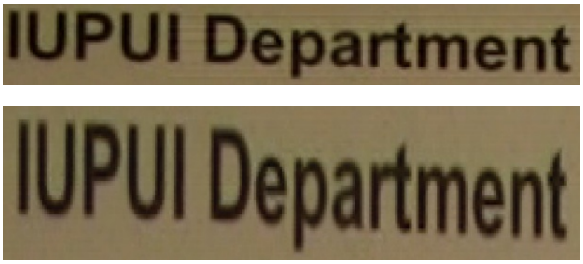


Figure 52: Close up of low resolution (Top) and High Resolution Image (Bottom) Height from camera $= 18''$, Speed $= 1$ pulse per second.

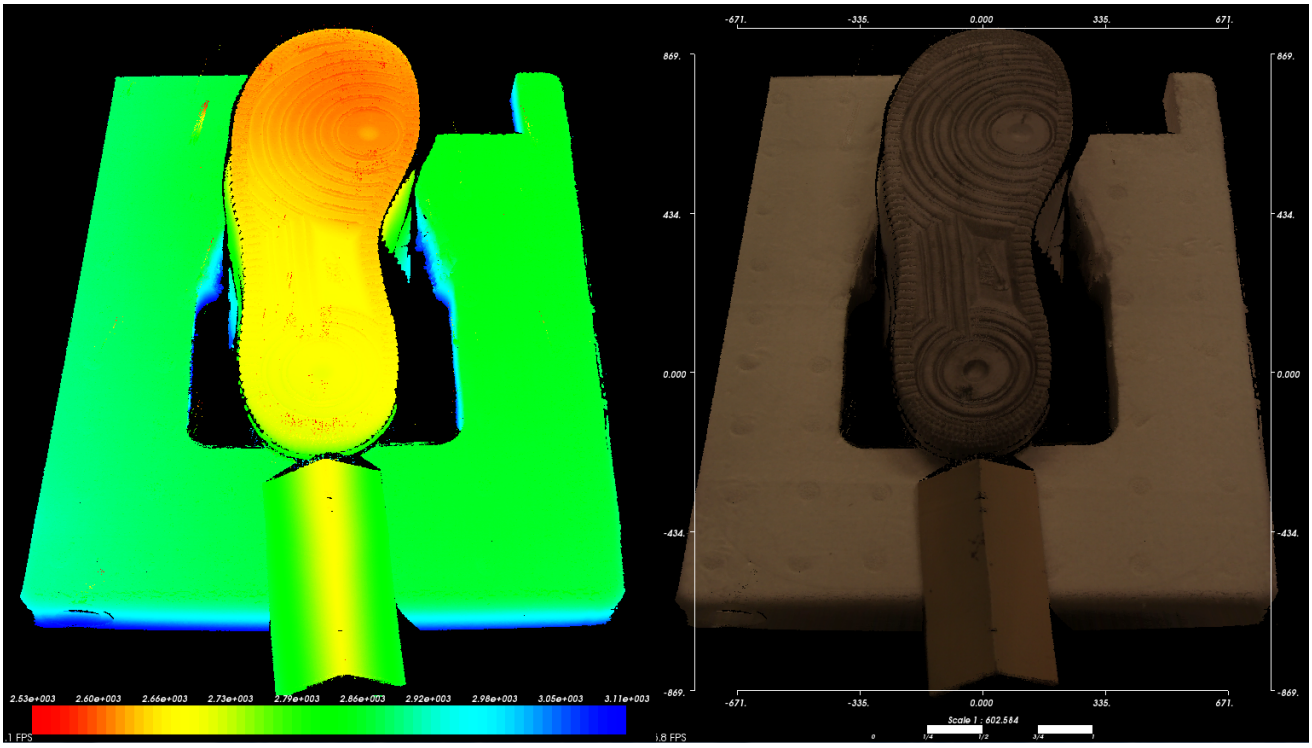


Figure 53: 3D point cloud result of the shoe scan. left: color encoded along z axis (colors toward red are closer to the camera and colors toward blue are farther from the camera). right: texture added.

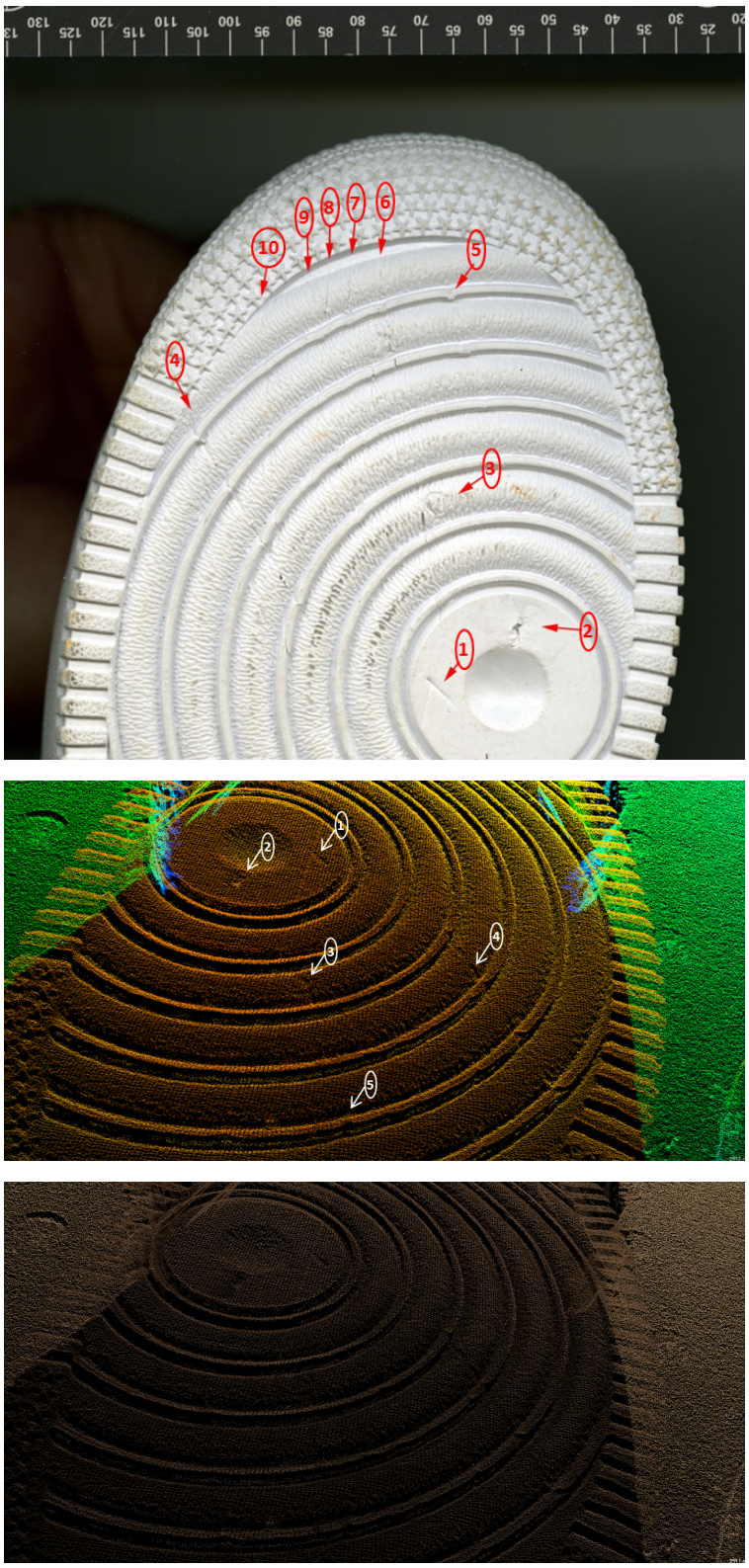


Figure 54: Scan results of a shoe with known characteristics. top: points of interest provided by expert. middle: 3D point cloud, height along z axis encoded with pseudocolor. bottom: 3D point cloud with captured color texture applied.

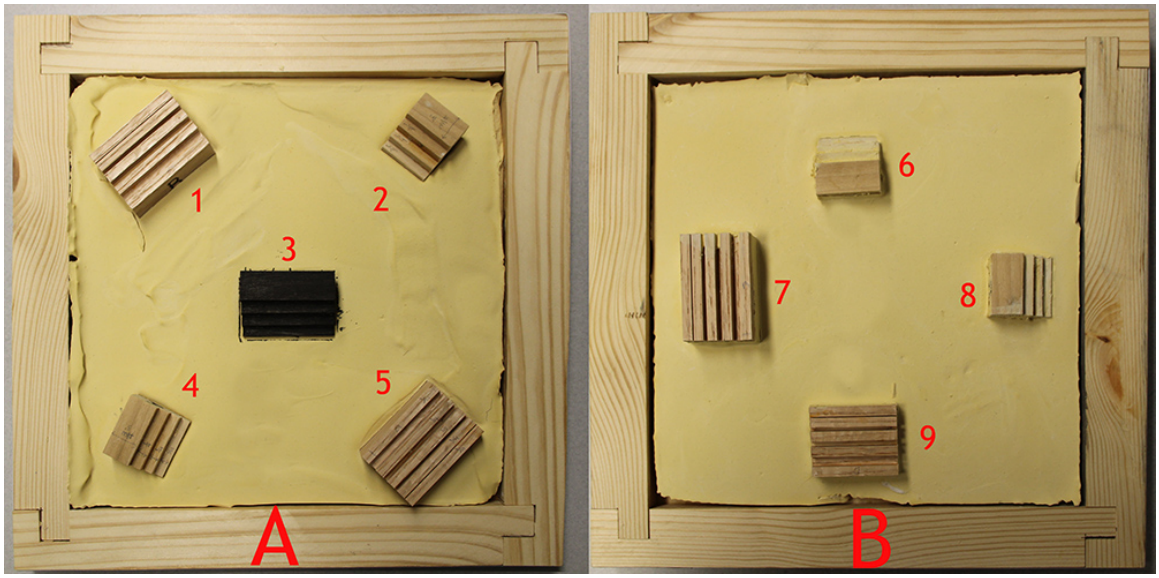


Figure 55: Test setup. Wooden blocks embedded in dental stone plates.

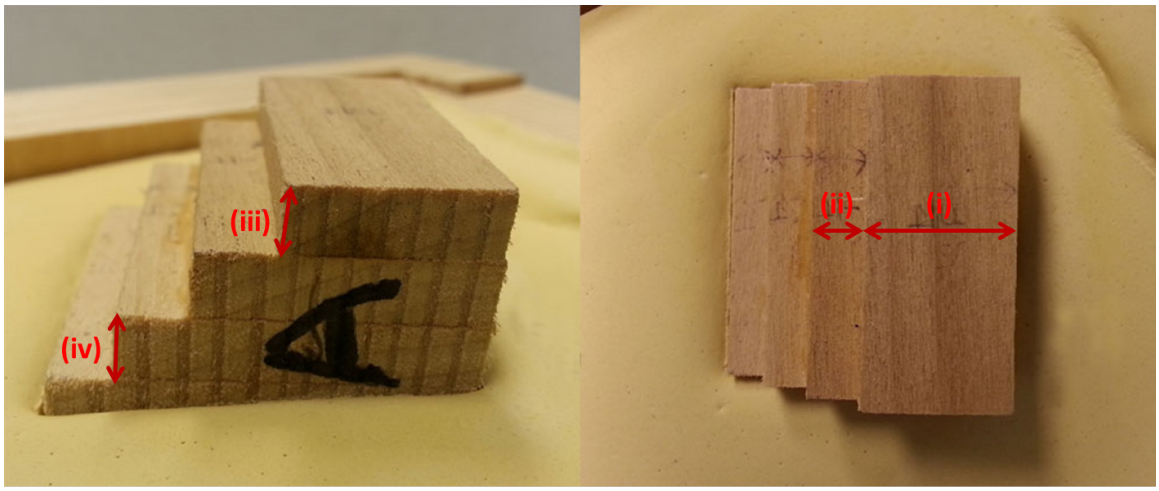


Figure 56: Wooden block with steps. Measurements (i) and (ii) are along xy plane; the width indicated by (ii) is $1/4$ inch. Measurements (iii) and (iv) are along z axis; the heights (iii) and (iv) are $1/4$ inches.

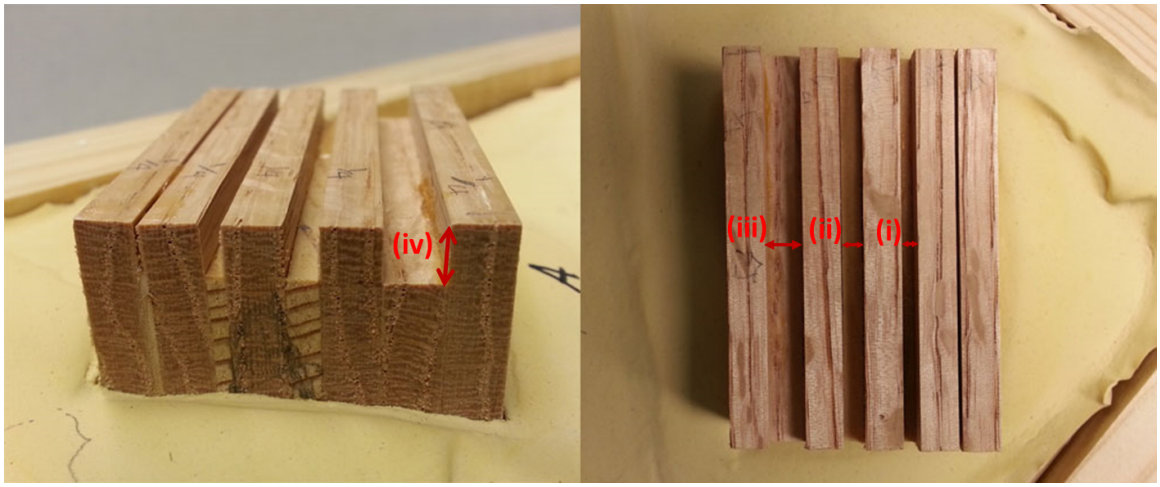


Figure 57: Wooden block with slots. Measurements (i), (ii) and (iii) are along xy plane. Measurement (iv) is along the z axis. The depth of the slots (represented by (iv)) is $1/4$ inch. The widths indicated by (iii), (ii), (i) are $1/4$, $1/8$, and $1/16$ inch, respectively. The narrowest slot on the right not labeled is $1/32$ inch wide.

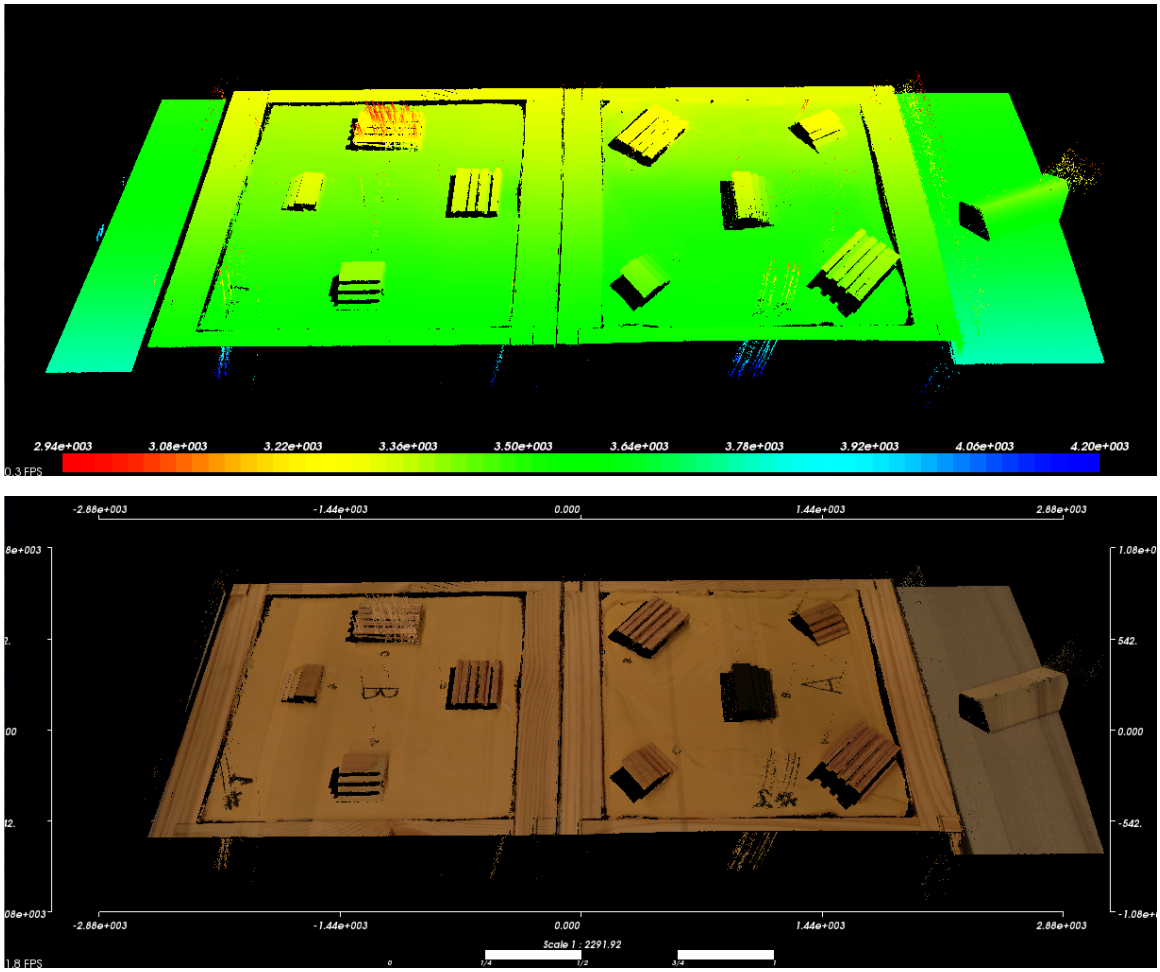


Figure 58: 3D Point cloud results of the accuracy test. Plates are positioned in the sequence of A followed by B. Top: color encoded heights. Lower values (red) are closer to the camera and larger values (blue) are farther from the camera. Bottom: textured point cloud in which the registered color values are displayed mapped to the height values.

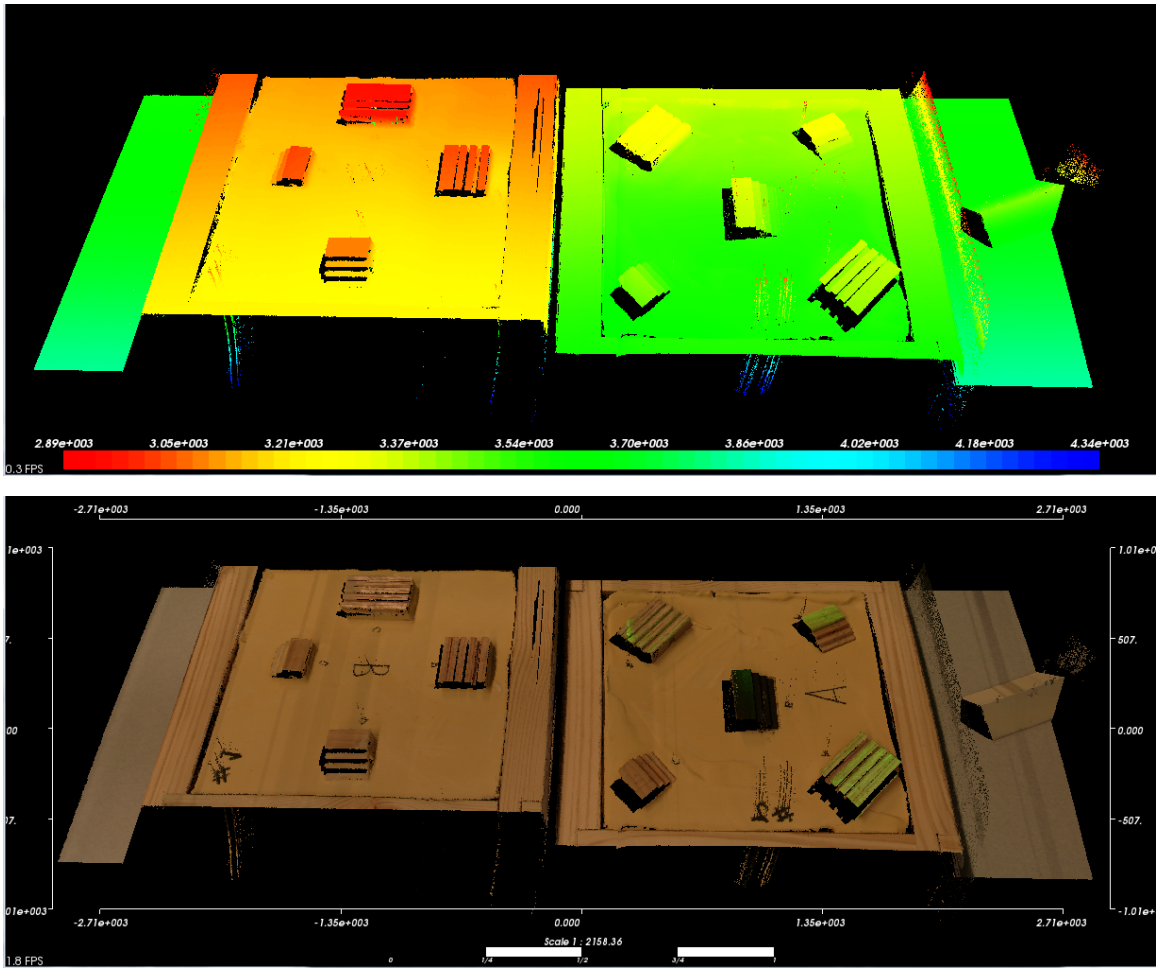


Figure 59: 3D Point cloud results of the accuracy test. Plates are positioned in the sequence of A followed by B, but B is lifted 40mm above the surface of A. Top: color encoded heights. Bottom: textured point cloud.

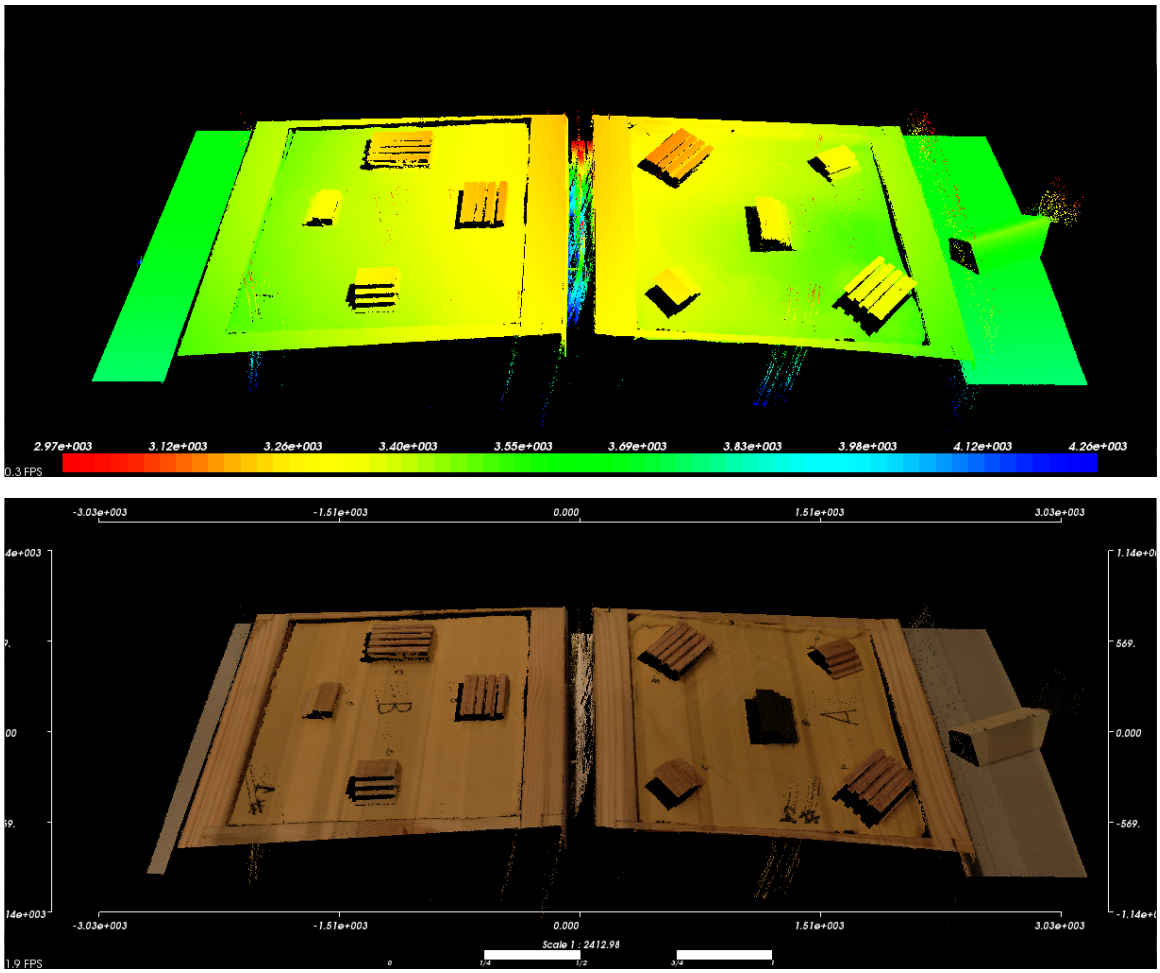


Figure 60: 3D Point cloud results of the accuracy test. Plates are positioned in the sequence of A followed by B, but both plated tilted by 5 degrees. Top: color encoded heights. Bottom: textured point cloud.

6 REFERENCES

[1] SWGTREAD Guidelines. http://www.swgtread.org/. http://www.swgtread.org/.

[2] W.J. Bodziak. *Footwear impression evidence: detection, recovery, and examination*. CRC Press, 1999.

[3] S. Borman and R. L. Stevenson. Super-resolution from image sequences-a review. In *Proc. Midwest Symp. Circuits and Systems*, pages 374–378, 1998.

[4] T. Botterill, S. Mills, and R. Green. Design and calibration of a hybrid computer vision and structured light 3D imaging system. In *Proc. 5th Int Automation, Robotics and Applications (ICARA) Conf*, pages 441–446, 2011.

[5] C. Chen and A. Kak. Modeling and calibration of a structured light scanner for 3-D robot vision. In *Proc. IEEE Int. Conf. Robotics and Automation*, volume 4, pages 807–815, 1987.

[6] Fei Zheng and Bin Kong. Calibration of linear structured light system by planar checkerboard. In *Proc. Int Information Acquisition Conf*, pages 344–346, 2004.

[7] R. B. Fisher and D. K. Naidu. A Comparison of Algorithms for Subpixel Peak Detection. In *in Image Technology, Advances in Image Processing, Multimedia and Machine Vision*, pages 385–404. Springer-Verlag, 1996.

[8] M.M. Houck and J.A. Siegel. *Fundamentals of forensic science*. Academic Press, 2nd edition, 2010.

[9] Jiang Yu Zheng. A flexible laser range sensor based on spatial-temporal analysis. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, volume 4, pages 740 –743 vol.4, 2000.

[10] P. McDonald. *Tire imprint evidence*. CRC Press, 1992.

[11] Michael B. Smith. The Forensic Analysis of Footwear Impression Evidence. *Forensic Science Communications*, 11(3), July 2009.

[12] Min Zhang and Dehua Li. An on-site calibration technique for line structured light 3D scanner. In *Proc. Asia-Pacific Conf. Computational Intelligence and Industrial Applications PACIIA 2009*, volume 2, pages 30–33, 2009.

[13] M. Pardonge, B. Walravens, M. Demeyere, K. T. Duy, R. Olszewski, H. Reychler, B. Raucent, and C. Eugene. 3D Reconstitution of Dentition by Laser Triangulation Under Structured Lighting. In *Proc. IEEE Int. Workshop Medical Measurement and Applications MEMEA '07*, pages 1–4, 2007.

[14] Gopal Ranjan and ASR Rao. *Basic and Applied Soil Mechanics*. New Age International, 2007.

[15] K.H. Strobl, W. Sepp, E. Wahl, T. Bodenmuller, M. Suppa, J.F. Seara, and G. Hirzinger. The DLR multisensory Hand-Guided Device: the Laser Stripe Profiler. In *Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on*, volume 2, pages 1927 – 1932 Vol.2, 26-may 1, 2004.

[16] Ursula Buck, Nicola Albertini, Silvio Naether, and Michael J. Thali. 3D documentation of footwear impressions and tyre tracks in snow with high resolution optical surface scanning. *Forensic Science International*, 171(2-3):157 – 164, 2007.

[17] Y. F. Wang and A. Pandey. Interpretation of 3D structure and motion using structured lighting. In *Proc. Workshop Interpretation of 3D Scenes*, pages 84–90, 1989.

[18] J.Y. Zheng and Zhong Li Zhang. Virtual recovery of excavated relics. *Computer Graphics and Applications, IEEE*, 19(3):6 –11, may/jun 1999.

[19] Assaf Zomet and Shmuel Peleg. Super-Resolution from Multiple Images Having Arbitrary Mutual Motion. In Subhasis Chaudhuri, editor, *Super-Resolution Imaging*, volume 632 of *The Kluwer International Series in Engineering and Computer Science*, pages 195–209. Springer US, 2002.

7 DISSEMINATION OF RESULTS

The results of this project were disseminated in the following venues up to now:

1. A poster was presented at the NIJ conference in Washington, DC in June 2011.

2. A paper was accepted to the International Workshop on Depth Image Analysis (WDIA 2012) for presentation and publication which was presented in November 2012. The full paper will be published in Lecture Notes on Computer Science during 2013.

3. The device and the research results were presented to the Scientific Working Group on Imaging Technology (SWGIT) in the January 2013 meeting.

4. A paper was accepted for presentation and publication to the Workshop on Applications of Computer Vision (WACV 2013) and was presented in January 2013. The paper was published in the conference proceedings. The exact citation is:

   - Ruwan Egoda Gamage, Abhishek Joshi, Jiang Yu Zheng, Mihran Tuceryan. "A High Resolution 3D Tire and Footprint Impression Acquisition for Forensics Applications," in Proc. of Workshop on Applications of Computer Vision (WACV), Clearwater, Florida, pp. 317-322, January, 2013. DOI link: 10.1109/WACV.2013.6475035, and (PDF link).

5. One abstract was submitted to the AAFS 2013 conference. It was presented in the criminalistics session on Friday, Feb 22, 2013. The abstract was published in the Proceedings of the 65th Meeting of the American Academy of Forensic Sciences (AAFS 2013), Washington, DC, Feb 18-23, 2013. (PDF Link to abstract)

6. A Technical Report has been published in the IUPUI Department of Computer & Information Science. Citation: Ruwan Egoda Gamage, Abhishek Joshi, Jiang Yu Zheng, Mihran Tuceryan, "Designing a High Resolution 3D Imaging Device for Footprint and Tire Track Impressions at Crime Scenes," Technical Report #TR-CIS-0416-12, Department of Computer & Information Science, Indiana University Purdue University Indianapolis (IUPUI), April 16, 2012. URL: `http://www.cs.iupui.edu/~tuceryan/tech-reports/TR-CIS-0416-12.pdf`