A High Resolution 3D Tire and Footprint Impression Acquisition for Forensics Applications

Ruwan Egoda Gamage, Abhishek Joshi, Jiang Yu Zheng, and Mihran Tuceryan

Indiana University Purdue University Indianapolis (IUPUI), Indianapolis IN 46202, USA,
{rjegodag,abhjoshi,jzheng,tuceryan}@cs.iupui.edu

Abstract. This paper presents a method with which 3D images of tire track and footprint impressions at crime scenes can be captured with high fidelity, while capturing high resolution 2D color texture images simultaneously. The resulting device is portable, easy to use, is non-destructive of the evidence, and saves time at crime scenes. The same technique can also be used in the laboratory to create 3D depth images of suspect tires or shoe soles. Computer-based pattern matching technology can be used to assist in matching and comparison tasks. The device produces better quality data at a close range obtained in a larger field compared to existing devices. It avoids problems related to occlusions by using two lasers and can digitize long spans of impressions in one scan. The method includes a calibration method which is integrated into the scanning process on site, thus avoiding problems with pre-calibrated configurations becoming stale during transportation and setup.

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 [1, 3, 4]. 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 paper, 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. The method can provide a depth resolution of around 0.5mm and high resolution color image.

Related Work: 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 [1, 3, 4]. The requirements for imaging

such evidence are specified by the Scientific Working Group for Shoe print and Tire Tread Evidence (SWGTREAD) guidelines [5]. 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 [7, 6]. 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 [2]. 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.

2 Design of the 3D Imaging System

2.1 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.

The two different colored lasers are used to eliminate the blind spots due to occluding surfaces. The actuator moves the entire assembly of camcorder and laser lights with fixed geometry along the rail in a linear motion. Each video frame provides one line of 3D depth values and all the lines are stacked to form the 3D depth image. The images obtained from the red and green lasers are processed separately and fused at the end to obtain a single, consistent image.

2.2 Image Processing Algorithms

Laser stripe detection: 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.

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. The assumption is that there will not be large surface depth variation in the tire treads and shoe soles. 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 $[\text{hue}_{min}, \text{hue}_{max}]$, which is determined by the color of the particular laser light (green or red). A dilate morphological operation is applied to the candidate pixels which form a 5-10 pixel wide band within which the peak points constitute the location of the laser

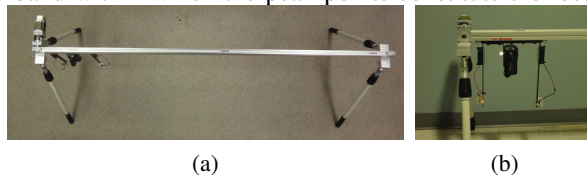

(a)   (b)

Fig. 1: The device prototype.

stripe. These peaks are the pixels with the maximum Value (brightness) for the pixel in the HSV representation within the band being searched.

Color texture image extraction: Simultaneously with the laser stripe detection, we also extract a high resolution color image of the impression evidence. This is done by extracting the pixels along the $y = 0$ line of each video frame having the origin at the center of the image. The color image captured by this process has the following properties: (i) The spatial resolution of the image along the scan direction is dependent on the speed with which the linear actuator moves. The slower the actuator moves, the higher the image resolution in the scan direction because the video is being captured at a fixed 30fps—the distance between the scan lines of successive video frames will be smaller as the actuator moves more slowly; (ii) 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 the highest resolution video frame is 1920×1080 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.

3   System Calibration

We have integrated the calibration of the geometry into the scanning and data collection process. This eliminates the risk that a pre-calibrated configuration can be invalidated during the transportation and set-up of the device. The only requirement in the field is that the criminalist places the calibration object in the scene for at least one scan.

We use an L-shaped calibration object (shown in Figure 2) 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 scan.
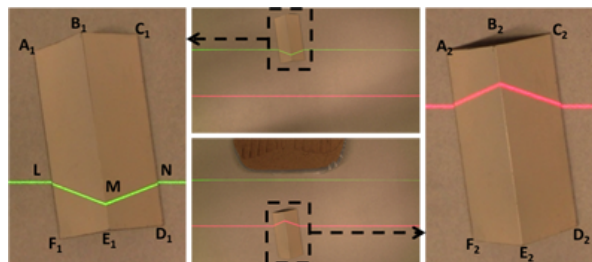

Fig. 2: Frames are captured when each laser scan over the calibration object at time t_1 and t_2.

Currently, we operate the system in its slowest speed of 1.3138 mm/s in order to obtain highest scan resolution along the y axis. We perform a calibration of the video camera in order to correct for radial lens distortion at widest zoom settings, and we computed a single focal length f for this setting to be 2110 in pixel units along both x and y directions. We correct for this radial distortion in every captured frame 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 and the rail motion direction. Then we use the calibration object (Figure 2) to calculate the orientations of the red and green laser planes. 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.

Coordinate System Definitions: We use three main coordinate systems

1. *o-xyz*: original camera coordinate system. Coordinates are measured in pixel units.
2. *O-XYZ*: ideal image coordinate system is still camera centered, but corrected for the roll and tilt of the camera with respect to the rail coordinates. The Y axis is parallel to the rail motion and the Z axis points downwards where the camera points. Coordinates are in pixel units.
3. $O\text{-}X'Y'Z'$: rail coordinate system. The coordinate system is aligned with the ideal image coordinate system orientation and related to it with a translation. The coordinates are in metric units.

The calibration procedure: Ideally, the camera should be connected to the rail, 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. 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 2, 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 $(A_i, B_i, C_i, D_i, E_i, F_i,$ 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.

First, we calculate the vanishing point from the corresponding points in two frames as described above. Let this vanishing point be (x_v, y_v).

Second, we compute the pose of the camera with respect to the rail from this vanishing point (for *O-XYZ* transformation). The camera roll, θ, (around its optical axis) is given by $\theta = \tan^{-1}(x_v/y_v)$. The camera tilt, α, between the optical axis and the Z axis in the ideal coordinate system *O-XYZ* is given by $(x_v/|y_v|) \tan^{-1}\left(f/\sqrt{x_v^2 + y_v^2}\right)$.

Next, the calculated roll and the tilt of the camera is used to obtain the transformation from the image coordinates (*o-xyz*) to the ideal coordinates (*O-XYZ*). This transformation

is given by $\mathbf{T} = \mathbf{R}_x(\alpha)\mathbf{R}_z(\theta)$, where the $\mathbf{R}_x(\alpha)$ and $\mathbf{R}_z(\theta)$ are the rotation transformations around the x and z axes, respectively. Note that because the linear motion is along the y axis, we do not need to calculate the third rotation angle, pan.

Computing rail coordinates of a point on the calibration object: 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 $Z_{a_1}/f = Y_{a_1}/Y$ and $Z_{a_1}/f = X_{a_1}/X$. Therefore, the projection of A_1 onto the $Z = f$ plane is obtained by $A_{1f} = (X_{a_1}f/Z_{a_1}, Y_{a_1}f/Z_{a_1}, f)$. Similarly, the projection of A_2 onto the $Z = f$ plane is given by $A_{2f} = (X_{a_2}f/Z_{a_2}, Y_{a_2}f/Z_{a_2}, f)$.

Finding the Z value in rail coordinate system: Assume the distance traveled between the first frame and the second frame is d (in metric units). Then by similar triangles, we obtain

$$Z' = \frac{d}{(Y_{a_2}/Z_{a_2} - Y_{a_1}/Z_{a_1})} \tag{1}$$

Finding the point A_1 in rail coordinate system at time $t = t_1$: Considering the edges of a triangle, we obtain $Z_{a_1}/Z' = X_{a_1}/X'$ and $Z_{a_1}/Z' = Y_{a_1}/Y'$. Therefore the ideal coordinates of the point A_1 at time $t = t_1$ is given by $A_1' = (X_{a_1}Z'/Z_{a_1}, Y_{a_1}Z'/Z_{a_1}, Z')$. Points corresponding to B_1, C_1, D_1, E_1, and F_1 are similarly computed in the rail coordinate system.

Finding the laser plane: Let's consider the green laser plane. First, we transform the pixel locations of L, M, and N (in Figure 2) to the ideal image coordinate system. 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:

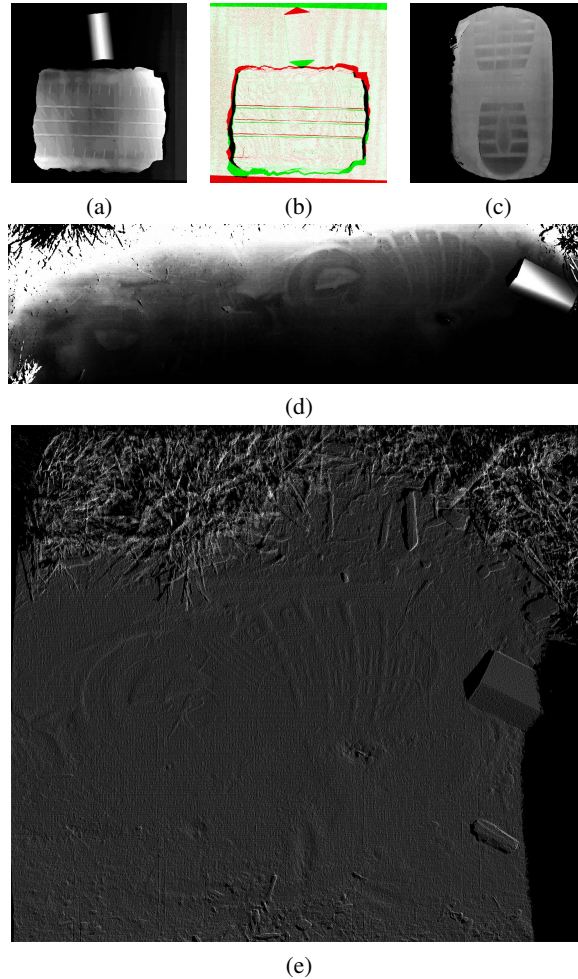$$\mathbf{N} \cdot (\mathbf{P} - \mathbf{P}_a) = 0$$

Fig. 3: Computed impression image results: (a) Fused height map for a tire impression; (b) Contributions of each laser for this fusion. (c) A computed shoe print impression image. (d) A long scan of two shoe prints. (e) Height map of the shoe print region in the scan of (d).

4   Computing the 3D Impression Image

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 Eq 1, 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 number 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.

5 Experimental Results

To test the system we have scanned tire and shoe impressions. Figure 3a shows the computed impression image for one such scan. The scan video file consisted 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.

Figure 3a shows the fused result of the two independently processed images with the red and green laser. Figure 3b shows the contributions of each laser to the final result color coded. Over most of the image the two computed height maps from the two lasers agree. Where one image has data and the other does not due to occlusions, the fusion

process fills in these pixels with the appropriate value. Figure 3c shows the fused depth image for a shoe print. And Figure 3d shows a longer scan of shoe print impressions.

We also captured some impressions in different materials such as snow and sand. The laser stripes in snow became very faint.

Accuracy and Resolution Measurements: We have determined the following accuracy measurements for the system: (i) Rail speed = 1.3138 mm/s, fps = 30, $f = 2110$; (ii) Resolution along Y' axis = 1.3138 / 30 = 0.0438mm; (iii) Resolution along X' axis (at a Z' distance of 500mm) = Z' / f = 500 / 2110 = 0.2369 mm; and (iv) Empirically, we have observed that we can resolve 0.5mm in Z' values as reflected in detectable differences of at least 1 pixel in the disparity image computed. (v) The device takes approximately 20 minutes to scan a 1.75m long surface. Note that even though 20 minutes seems like a long time, this is a great improvement for the current practices of obtaining impression evidence in the field. Currently, if the impression evidence is captured by making physical casts, not only can the process take longer, but it could also destroy the evidence in the process.

6 Conclusion

In summary, we have developed an inexpensive high resolution 3D impression device for digitizing shoe and tire impressions in crime scenes. The device can also be used for many other objects such as ancient calligraphy on stones. We used two laser modules to eliminate occlusions and improved performance. 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. Compared to current practices in forensics, the device can greatly improve and speed up the process of collecting impression evidence in the field. Moreover, currently, in order to scan long tire tracks, multiple photographs need to be taken along the track and stitched together. Our device can capture a 3D impression of such long tire tracks in one single scan.

References

1. Bodziak, W.: Footwear impression evidence: detection, recovery, and examination. CRC Press (1999)
2. Buck, U., Albertini, N., Naether, S., Thali, M.J.: 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), http://www.sciencedirect.com/science/article/B6T6W-4MJBTMV-2/2/037b6c5ec7b1b9745f53cbbc2f2ce5d4
3. Houck, M., Siegel, J.: Fundamentals of forensic science. Academic Press, 2nd edn. (2010)
4. McDonald, P.: Tire imprint evidence. CRC Press (1992)
5. Scientific Working Group for Shoeprint and Tire Tread Evidence: SWGTREAD Guidelines. http://www.swgtread.org/, http://www.swgtread.org/
6. Zheng, J.Y.: A flexible laser range sensor based on spatial-temporal analysis. In: Pattern Recognition, 2000. Proceedings. 15th International Conference on. vol. 4, pp. 740 –743 vol.4 (2000)
7. Zheng, J., Zhang, Z.L.: Virtual recovery of excavated relics. Computer Graphics and Applications, IEEE 19(3), 6 –11 (may/jun 1999)