https://doi.org/10.2352/ISSN.2470-1173.2021.18.3DIA-069

3D Shape Estimation for Smooth Surfaces Using Double Mirror System*

Yin Wang[1]***, Davi He***[2]***, Zillion Lin***[2]***, George Chiu***[1]***, Jan Allebach***[1]***,***
[1] ***Purdue University, West Lafayette, Indiana 47906, U.S.A***
[2] ***Sunvalleytek International Inc., Shenzhen, CHINA***

Abstract

In this paper, a low cost, single camera, double mirror system that can be built in a desktop nail printer will be described. The usage of this system is to capture an image of a fingernail and to generate the 3D shape of the nail. The nail's depth map will be estimated from this rendered 3D nail shape. The paper will describe the camera calibration process and explain the calibration theory for this proposed system. Then a 3D reconstruction method will be introduced, as well. Experimental results will be shown in the paper, which illustrate the accuracy of the system to handle the rendering task.

1. Introduction

With the development of inkjet printing technology, inkjet printers have been widely used in the office and in the household. At the same time, three-dimensional measurement is a popular research topic in computer vision [1][2]. We want to combine inkjet printing technology with a 3D measurement method to produce a small, desktop size printer, which can be used to measure the depth map of the fingernail, and print any desired pattern on it. The reason why we want to do the 3D measurement is that without considering the depth map, the curved surface of the fingernail will cause image deformation during the printing process. In order to solve this problem, last year, we proposed to use structured light [3][4] to estimate the depth map of a fingernail, with a single-camera, single-projector system [5]. This year, we propose a simpler system that contains a single camera, and double mirror. The calibration process can be simplified compared to the structured light system; and the reconstruction accuracy will be evaluated as well.

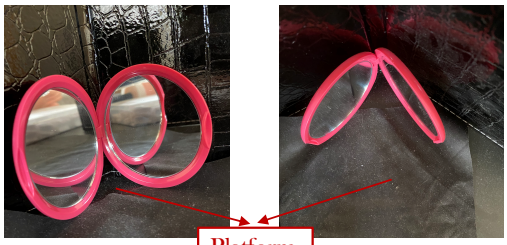
This article will have the following structure. First, Sec. 2 will give an overview for the nail printer product, and introduce the system layout. Sec. 3 will include the camera calibration method and the calibrated result with the proposed system. In Sec. 4, two different reconstruction methods that can be used to reconstruct the 3D map of the fingernail will be introduced, and the reconstruction result will be shown in this section, as well. Last but not least, the conclusion will be stated in Sec. 5

2. Overview

The product we are trying to build is a small size portable printer. The nail printer has the size of 210mm $\times$ 210mm $\times$ 210mm, which is a cube and has a screen built on top of it, which can be used as the interface for the customer to choose the desired pattern and see the position of their nail inside the printer. Fig. 1 shows an image of the nail printer.

The printer in Fig. 1 does not have mirrors built inside. In order to verify that our strategy would work for solving the problem, a single camera, double mirror system is built outside the printer. As shown in Fig. 2, the pink round shaped mirrors are used, and the two mirrors are standing on the same platform, which is highlighted in the Fig. 2. The second essential tool is an uncalibrated camera. In this experiment, we are using the same camera as we have used in the structured light system [5], which is a Logitech C920S Webcam, shown in Fig. 3 [6]. The camera's resolution is 960×720. That is sufficient to capture an image with the desired quality; and the camera's parameters (focal length, white balance, etc) can be manually adjusted, accordingly. A sample image with a fingernail in the proposed experimental setup is shown in Fig. 4. We can see from the figure that the nail must be standing between the two mirrors on the same platform so that both mirrors will be reflecting the image of the fingernail.

Figure 1: Nail printer overview.



(a) Front view (b) Top view

Figure 2: Double mirrors standing on a platform.

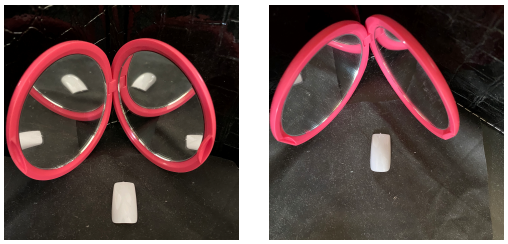
When taking an image using the proposed experimental setup, there are several requirements that must be met: The first

*Research supported by Sunvalleytek International Inc, Shenzhen, CHINA.

Figure 3: Logitech C920S Webcam.

one is from the camera point of view. The camera's focal length f and principal point $\mathbf{p_0}$, should stay the same in each image. Two images will be taken using the same focal length f and the same principal point $\mathbf{p_0}$. The only difference between the two images is the camera's physical position. In other words, only the camera's external parameters will be changed between the two shots. Another requirement is that the angle and the physical position of the two mirrors cannot be changed between taking the two images. Also, the object's position cannot be changed, as well.

In the camera image plane, we have to make sure that 5 views of the object should be visible in each camera image plane. These are: a view directly on the object, O_R; two direct reflections from the two mirrors: O_{v1} and O_{v2}; two reflection of reflections: O_{v12} and O_{v21}. Suppose the left mirror is mirror 1 and the right mirror is mirror 2. The view O_{v1} is the reflection of the real object O_R in mirror 1, and O_{v2} is the real object's reflection in mirror 2. Then, the reflection of O_{v1} in mirror 2 is called O_{v12}, and the reflection of O_{v2} in mirror 1 is O_{v21}. A sample image with the five views can be seen in Fig. 5. We need to pay attention that the orders of the 5 views will always stay the same in each image. This is very critical in the further calculations.



(a) Front view (b) Top view

Figure 4: Double mirror system with the nail standing in the middle.

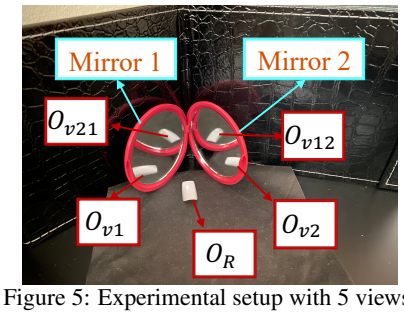


Figure 5: Experimental setup with 5 views.

Another thing to which we need to pay attention is that, in each captured image, all 5 views of the object must be fully visible. Partially visible views will not be accepted; and will cause some reconstruction errors in the final results. For example, Fig. 6 is a bad example of the 5 views, where the nail is partially invisible in view O_{v2}. Fig. 7 is a good example of the 5 views in the camera image plane, from two different angles.

Figure 6: A bad example of the 5 views.

Figure 7: A good example of the 5 views in the camera image plane taken from two different angles.

3. Camera Calibration

We have introduced how to set up the experiment environment with a single camera, double mirror system in the previous section. Now we are going to introduce how to calibrate the camera under the given experimental setup. All the calculations and geometries are based on the pin-hole camera model [7]. The camera system can be regarded as a 3D coordinates system, and the camera center is located at the origin, as shown in Fig. 8. Also, the camera image plane is always perpendicular to one of the axes, the z axis. The intersection between the z axis and the camera image plane is the principal point ($\mathbf{p_0}$). The distance from the camera center to the principal point is the focal length (f) of our camera.

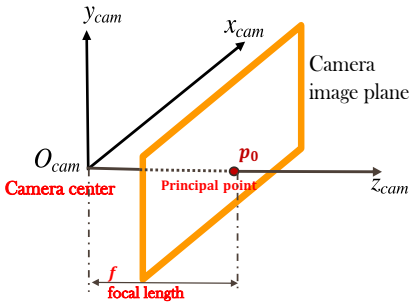


Figure 8: Illustration of basic camera model.

3.1 Epipolar Geometry

An epipolar geometry has been commonly used to describe the cameras intrinsic to the computer vision area when we have a pair of cameras. The basic epipolar geometry is shown in Fig. 9. Here, $\vec{C}$ and $\vec{C}'$ are the two camera centers, $\vec{X}$ is a point in the world coordinate system, its reflections on the two camera image planes are $\vec{x}$ and $\vec{x}'$, respectively. And the epipoles $\vec{e}$ and $\vec{e}'$ are the intersection points between the two camera image planes and

the line connecting the two camera centers. From Fig. 9, we can see that any pixel $\vec{x}$ in the left camera image plane must have a corresponding pixel $\vec{x}'$ in the right camera image plane. And the image $\vec{x}'$ of any world point $\vec{X}$ must lie on the right epipolar line l'. Because of the definition of the epipolar line, all epipolar lines must pass through right epipole $\vec{e}'$, and the position of the epipolar line l' must depend on the position of $\vec{x}$ in the left camera image plane. So, the position of epipole $\vec{e}$ is a critical parameter to determine the camera's position and direction.

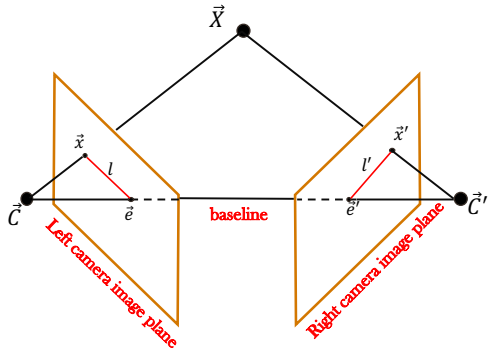


Figure 9: Epipolar geometry.

Now, let us consider a situation where we only have one camera, but there is a mirror stand in front of the camera, as shown in Fig. 10. The camera center is C, and there are two points in the 3D world coordinate system, point A and point B. If we connect point A and the camera center C, the intersection between that line and the camera center plane will be a 2D point a. And we will have point b in the camera image plane in the same way. Because of the mirror, the reflections of point A and point B in the mirror will be point A' and point B', respectively. These two points will also have their projections on the camera image plane as a' and b', respectively. For the camera center C, there will also be a reflection on the other side of the mirror, we call this the virtual camera center C'. The intersection between the line CC' and the camera image plane is the epipole $\vec{e}$. And the line aa' and the line bb' on the camera image plane will also intersect at point $\vec{e}$, as well. So, any pair of lines on the camera image plane must pass through point $\vec{e}$. And point $\vec{e}$ can be used to represent the mirror's normal direction, because the line CC' must be perpendicular to the mirror plane.

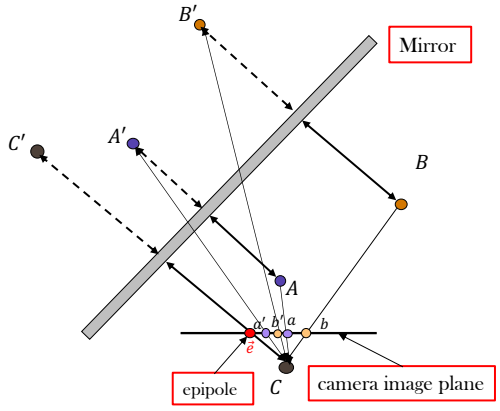


Figure 10: Single camera, single mirror scheme illustration.

In Fig. 11, we present an example image in the single camera and single mirror scheme. The nail in the bottom part of the image is the real object. The other one is its reflection in the mirror. Suppose point a, a' and point b, b' are the two corresponding points between the real object and its reflection. The intersection between aa' and bb' must be the epipole point $\vec{e}$ that we are looking for.

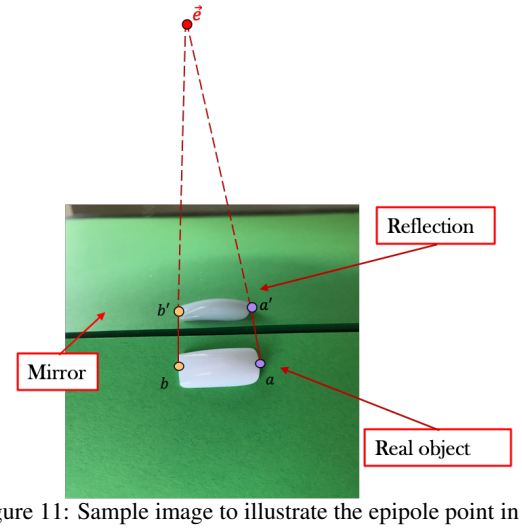


Figure 11: Sample image to illustrate the epipole point in the single camera, single mirror scheme.

3.2 Double Mirror Scheme

Having introduced the single camera single mirror scenario, let us switch to the system that we proposed at the beginning of the paper: the single camera, double mirror system, where we will have one more mirror. Because of the reflection properties of the mirror, we will have more cameras that need to be considered in this case: the real camera center C_R, the direct reflection C_{V1} of C_R in mirror 1, the direct reflection C_{V2} of C_R in mirror 2, the reflection C_{V12} of C_{V1} in mirror 2, and the reflection C_{V21} of C_{V2} in mirror 1. There are two more camera centers that need to be considered, the reflection C_{V121} of C_{V21} in mirror 2, and the reflection C_{V212} of C_{V12} in mirror 1. So, we will have 1 real camera center and 6 virtual camera centers. Since finding all the camera centers is based on the reflection of the two mirrors that stand on a platform, the 7 camera centers must be on the same plane. We call it the camera center plane. This geometry is first proposed by [8], and an intuitive geometry illustration can be found in Fig. 12.

In this camera image plane, we have total of 7 camera centers, and 4 epipoles. Fig. 13 shows how the four epipoles are found. The intersection of the two corresponding points of O_R and O_{v1} will be the epipole e_{v1}. A pair of corresponding points O_{v2} and O_{v21} will intersect at epipole e_{v1}, as well. The same thing will happen at epipole e_{v2}. If we find a pair of corresponding points between the object O_{v21} and the object O_{v1}, the intersection of the lines will be the epipole of camera C_{V121}. We call this epipole point e_{v121}. At the same time, from a pair of corresponding points of O_{v2} and O_{v12}, we can define the epipole e_{v212}, which is the epipole of the virtual camera C_{V212}.

After having all the camera centers and epipoles, we can calculate the pricinple point $\mathbf{p_0}$ and the focal length f of the real camera in the camera center plane. First, we will introduce two

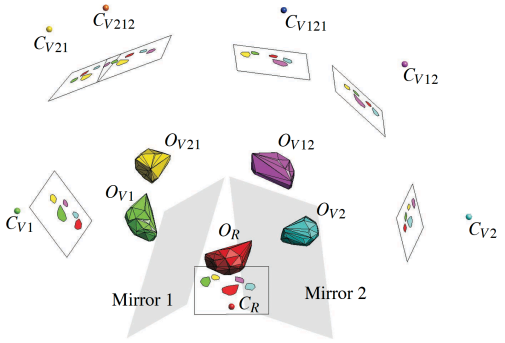


Figure 12: One real camera center C_R and six virtual cameras C_{V1}, C_{V2}, C_{V12}, C_{V21}, C_{V121}, and C_{V212} on the camera center plane. (Image is from [8]).

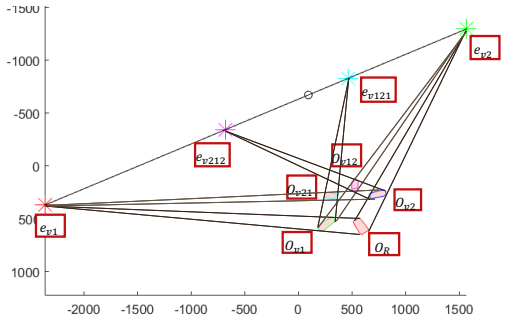


Figure 13: Five silhouettes and four epipoles in the real camera image plane.

dummy mirrors: dummy mirror 1 and dummy mirror 2. Fig. 14 shows the imaging geometry with two real mirrors and all 7 camera centers in the camera center plane. Next, we introduce a new mirror, dummy mirror 1, which has the same direction as the mirror 1, but is positioned in the middle of the real camera center C_R and epipole e_{V1}. Point E is the reflection of e_{V1} in the mirror 2, and point F is the reflection of point E in the dummy mirror 1. If we draw a line to connect point F with the camera center C_R, the line must intersect with image plane at point e_{V121}. This means that C_{V121}, e_{V121}, point F, and point C_R must be co-linear. The intersection of these lines to the dummy mirror 1 is called point G. Then, we only keep the blue lines in the Fig. 14, and introduce another dummy mirror, which is dummy mirror 2 in Fig. 15. The direction of dummy mirror 2 is the same as mirror 2, but is positioned in the middle of C_R and e_{V2}. Similar to the way in which we introduced points E, F, and G, using the dummy mirror 2, we introduce point M, the reflection of e_{V2} in dummy mirror 1, and point N, the reflection of M in dummy mirror 2. Then, the intersection of line NC_R and the image plane must be the epipole point e_{V121}, and points N, J, K, C_R must be co-linear, as well. We define a, b, and c as the distance between e_{V1} and e_{V212}, the distance between e_{V212} and e_{V121}, and the distance between e_{V121} and e_{V2}, respectively.

In Fig. 15, there are several similar triangles. Because $\triangle e_{V1}e_{V121}G \cong \triangle e_{V2}e_{V121}C_R$, we will have:

$$\frac{d(e_{V1},G)}{d(C_R,e_{V2})} = \frac{(a+b)}{c} \tag{1}$$

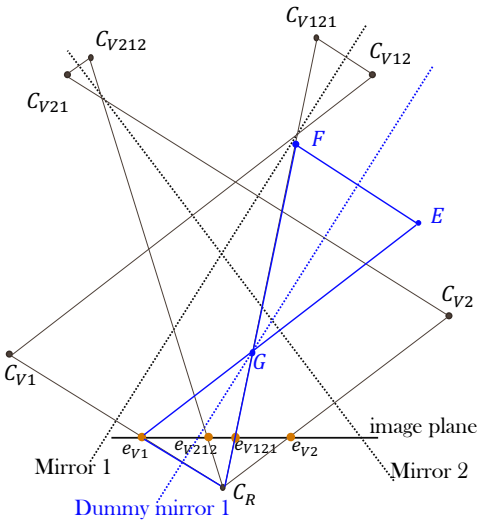


Figure 14: Imaging geometry with two real mirrors and one dummy mirror. (Image is from [8]).

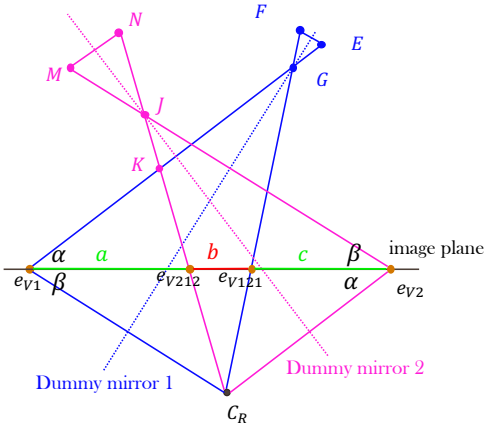


Figure 15: Imaging geometry with two dummy mirrors. (Image is from [8]).

And since $\triangle e_{V2}Je_{V212} \cong \triangle e_{V1}C_Re_{V212}$, we have the equation:

$$\frac{d(e_{V2},J)}{d(e_{V1},C_R)} = \frac{(b+c)}{a} \tag{2}$$

Next, we can observe that since $\triangle e_{V1}GC_R \cong \triangle e_{V2}JC_R$, we can express the similar relationship as:

$$\frac{d(C_R,e_{V1})}{d(C_R,e_{V2})} = \frac{d(e_{V1},G)}{d(e_{V2},J)} \tag{3}$$

Combining Eqns. (1), (2), and (3), we can express $d(C_R,e_{V1})$ in terms of $d(C_R,e_{V2})$ as:

$$d(C_R,e_{V1}) = \frac{\sqrt{a(a+b)}}{\sqrt{c(b+c)}} d(C_R,e_{V2}) \tag{4}$$

Until now, all three sides in $\triangle e_{V1}GC_R$ can be expressed using $d(C_R,e_{V2})$. In that case, using the cosine rule, we can express $\angle Ge_{V1}C_R$ using a, b, and c as:

$$\cos(\angle Ge_{V1}C_R) = \cos(\alpha+\beta) = \frac{1}{2}\frac{\sqrt{c(c+b)a(a+b)}}{(c+b)(a+b)} \tag{5}$$

Then, we can determine the lengths of all the sides in $\triangle e_{V1}C_R e_{V2}$, because $\angle e_{V1}C_R e_{V2} = 180° - \alpha - \beta$. Now, let f_Π represent the distance from the camera center C_R to the epipole line, which is also the height in $\triangle e_{V1}C_R e_{V2}$. We can express f_Π as:

$$f_\Pi = \frac{1}{2}\frac{\sqrt{3ac+4ab+4cb+4b^2}(a+b+c)\sqrt{ac}}{a^2+ab+c^2+cb+ac} \tag{6}$$

Let the point $\mathbf{p}_\Pi$ represent the point on the epipole line that is closest to the camera center C_R. Then, the point $\mathbf{p}_\Pi$ can be expressed using a, b, and c as:

$$\mathbf{p}_\Pi = \mathbf{e_{V1}} + \frac{1}{2}\frac{(2a+2b+c)a(a+b+c)}{a^2+ab+c^2+bc+ac}\frac{\mathbf{e_{V1}}-\mathbf{e_{V2}}}{|\mathbf{e_{V1}}-\mathbf{e_{V2}}|} \tag{7}$$

Here $\mathbf{e_{V1}}$ and $\mathbf{e_{V2}}$ represent the 2D image coordinates of points e_{V1} and e_{V2} on the camera image plane. A visualization of f_Π and $\mathbf{p}_\Pi$ is shown in Fig. 16.

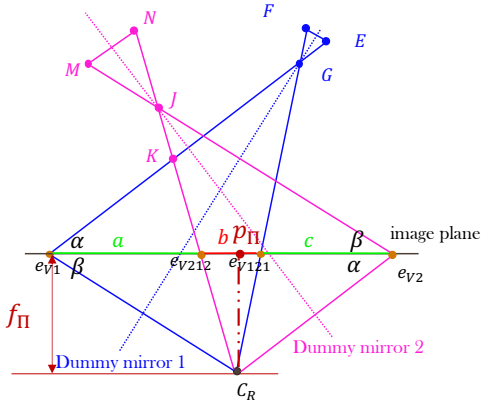


Figure 16: Illustration of distance f_Π and point $\mathbf{p}_\Pi$.

Recall that the principal point $\mathbf{p_0}$ must be on the camera image plane. So the line that passes through $\mathbf{p}_\Pi$ and is perpendicular to the epipole line must contain the point $\mathbf{p_0}$. In that case, if we have two images from two different viewing angles, under the same camera parameters, we can calculate the position $\mathbf{p_0}$ from two such images. Fig. 17 and Fig. 18 illustrate how we calculate $\mathbf{p_0}$ from the two viewing angles. In Fig. 17, we can determine that point $\mathbf{p_0}$ must be in the direction of $\vec{l}_1$, which is perpendicular to the epipole line. And in Fig. 18, $\mathbf{p_0}$ must also lie in the direction of $\vec{l}_2$. So the intersection of line $\vec{l}_1$ and line $\vec{l}_2$ is the position of the principal point $\mathbf{p_0}$.

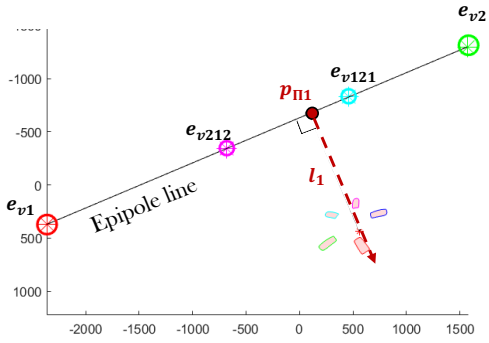


Figure 17: Camera image plane view 1.

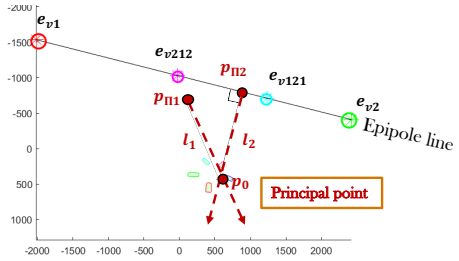


Figure 18: Camera image plane view 2, with the principal point $\mathbf{p_0}$.

At this moment, we can express f_Π, C_R, $\mathbf{p}_\Pi$, and $\mathbf{p_0}$ in terms of a, b, c, and $\mathbf{e_{V1}}$ and $\mathbf{e_{V2}}$. The relationship between f_Π, f, C_R, $\mathbf{p}_\Pi$, and $\mathbf{p_0}$ can be viewed in Fig. 19. So we can express the focal length of the camera f using the known parameters $\mathbf{p}_\Pi$, $\mathbf{p_0}$, and f_Π as:

$$f = \sqrt{f_\Pi^2 - ||\mathbf{p_0} - \mathbf{p}_\Pi||^2} \tag{8}$$

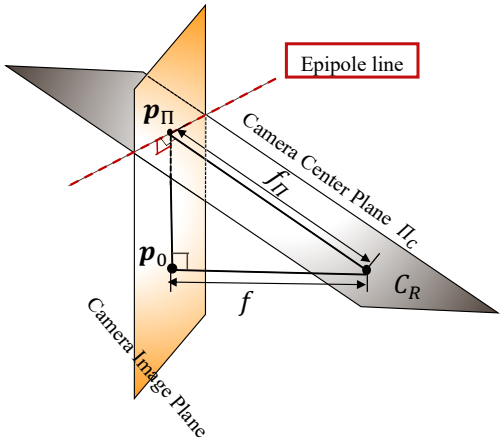


Figure 19: 3D visualization of the relationship between the camera center plane and the camera image plane, and visualization of the relationship between f_Π, f, C_R, $\mathbf{p}_\Pi$, and $\mathbf{p_0}$.

Then the normal directions of the two mirrors $\mathbf{m_1}$ and $\mathbf{m_2}$ can be represented as:

$$\mathbf{m_1} = -\begin{bmatrix}\mathbf{e_{V1}}-\mathbf{p_0}\\ f\end{bmatrix}, \mathbf{m_2} = -\begin{bmatrix}\mathbf{e_{V2}}-\mathbf{p_0}\\ f\end{bmatrix} \tag{9}$$

The minus signs indicate that we want the normal directions to point from the mirrors toward the camera center C_R. Also, note that $\mathbf{m_1}$ and $\mathbf{m_2}$ are 3D directions. Here $\mathbf{e_{V1}}-\mathbf{p_0}$ and $\mathbf{e_{V2}}-\mathbf{p_0}$ represent the 2D directions on the camera image plane, and f is the length from the camera image plane to the camera center, and its direction is perpendicular to the camera image plane. So combining $\mathbf{e_{V1}}-\mathbf{p_0}$, $\mathbf{e_{V2}}-\mathbf{p_0}$ with f, we can express the two mirrors' normal directions in 3D coordinates.

Then, given any point, its reflection by the mirror can be represented as: (Suppose the mirror has unit normal direction $\mathbf{m} = [m_x \quad m_y \quad m_z]^T$)

$$R = \begin{bmatrix} -m_x^2+m_y^2+m_z^2 & -2m_xm_y & -2m_xm_z \\ -2m_xm_y & m_x^2-m_y^2+m_z^2 & -2m_ym_z \\ -2m_xm_z & -2m_ym_z & m_x^2+m_y^2-m_z^2 \end{bmatrix} \tag{10}$$

Table 1: Difference between the camera calibration results based on the double mirror system and the checkerboard method.

	$\mathbf{p_0}$ (pixels)	f (pixels)
Double mirror method	$(575.96, 426.69)$	1017
Bouguet's camera calibration Toolbox [9]	$(568.65, 430.44)$	1064
Difference	$(\mathbf{1.29\%}, \mathbf{0.87\%})$	$\mathbf{4.42\%}$

3.3 Camera Calibration Result

Fig. 20 and Fig. 21 show the camera calibration result from two different viewing angles. The image labeled (a) is the original image in the camera image plane, and (b) is the binary image that segments the nail from the background. And the third column (c) shows the epipole lines in the two images, where all four epipoles lie on the epipole line. And the last column (d) shows the directions that are perpendicular to the epipole line. The intersection of two such lines can help us to find the position of the camera's principal point $\mathbf{p_0}$.

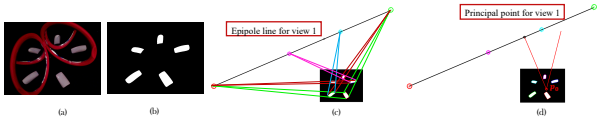


Figure 20: Camera calibration View 1. The two red lines in (d) indicate the directions that are perpendicular to the epipole line in the two different views, respectively.

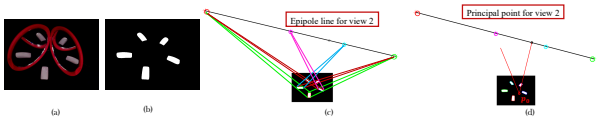


Figure 21: Camera calibration View 2. The two red lines in (d) indicate the directions that are perpendicular to the epipole line in the two different views, respectively.

From our calculation, the camera's principal point's position is $(575.96, 426.69)$ in the camera image plane. And the camera's focal length is $f = 1017$ pixels. In order to estimate the camera calibration accuracy, we compare the calibration results with the results from the checkerboard method [7][9], where about 10 images will be taken by positioning the camera in front of a checkerboard. The comparison result can be viewed in Table 1. We can see that the difference between the principal points' positions is quite small, which is no more than 1.3% of the whole field of view (960 pixels $\times$ 720 pixels). For the focal length, the difference is about 4.4%, which is a 47 pixel difference.

One thing that needs to be noted is that we only compared these two parameters individually. The comparisons are just a reference. They are not necessarily a criterion for a good estimation. There is no direct relationship between a high calibration quality and the small difference between two calibration results. A high quality camera calibration result must provide a good mapping between the 3D world coordinates point to the 2D camera image coordinates point throughout the field of view.

4. 3D Reconstruction Under the Double Mirror System

In this section, we will consider the following topic: Under the single camera, double mirror system, how do we do the mapping from 3D world coordinates to the 2D image coordinates and also back-project the 2D image points to the world 3D points, using the camera calibration results from Sec. 3. A number of articles have introduced methods that combine the visual hull and the shape of the silhouettes to do the 3D reconstruction [10][11][12][13]. We will be focusing on our specific object – a fake nail, to discuss the problem in this section.

4.1 Visual Hull 3D Reconstruction

The visual hull is a concept of 3D reconstruction using the Shape-From-Silhouette (SFS) technique, which is first introduced by Baumgart [14]. The basic idea of SFS is to create a 3D representation of the object by mapping from 2D points on the silhouettes in the camera image, under several different viewpoints. Fig. 22 gives a good visualization of the visual hull principles [15]. In Fig. 22(a), we take images from two different viewpoints, so two silhouettes are formed on the two camera image planes. In Fig. 22(b), when connecting the camera center with the silhouette's edge points, a certain volume will be bounded. The bounded volume is called the visual cone. Each camera center from each viewpoint will form its visual cone with a corresponding silhouette. From the intersection of several visual cones, we can estimate the general 3D shape of the original object. As shown in Fig. 22(c), we can estimate the shape of the object from the intersection of the two visual cones.

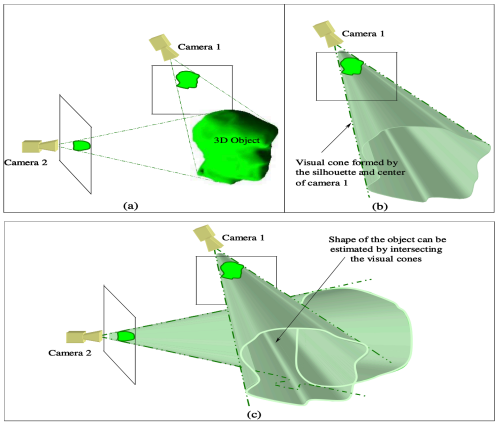


Figure 22: Visual hull illustration. (Image is from [15]).

So, the basic working process for 3D reconstruction using the visual hull is:

1. Take several images from different viewpoints.
2. Binarize each image into an object mask to get a silhouette of the object.
3. Combining the camera calibrated intrinsic and extrinsic parameters, back-project the silhouettes' points to the 3D coordinates.
4. Find the intersection of multiple viewing cones, the intersection volume will be called the visual hull (The concept is proposed in [16], and further discussed in [17] and [18].), and will yield the estimation of the original 3D object.

4.2 Two Ways for the 3D Reconstruction from the Visual Hull

There are two ways that we can reconstruct the 3D object from the visual hull. The first is two-dimensional surface based reconstruction. The other one is three-dimensional volume based reconstruction. Fig. 23 shows how to do the 3D reconstruction using two dimensional surface based method. In this figure, C^1, C^2, C^3, and C^4 are the camera centers from different viewpoints. S_1^1, S_1^2, S_1^3, and S_1^4 are the silhouettes corresponding to the four camera centers, respectively. Basically, we want to find the direct intersections of the 2D visual wedges.

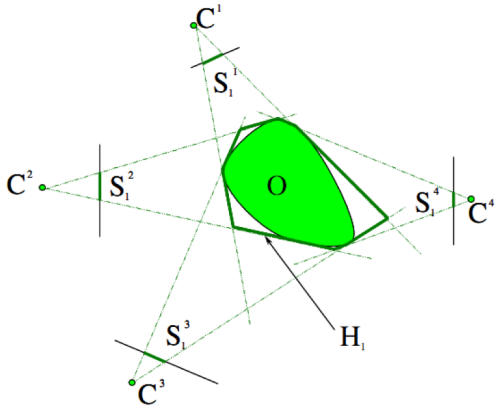


Figure 23: Visual hull reconstruction method 1 – Two dimensional surface based 3D representation.(Image is from [15]).

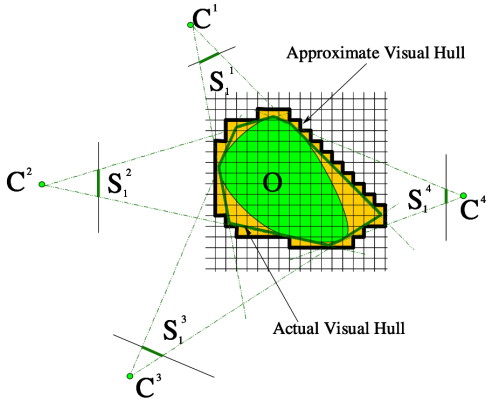


Figure 24: Visual hull reconstruction method 1 – Three dimensional surface based 3D representation.(Image is from [15].)

There are many different ways to do the 3D reconstruction using the three-dimensional volume based method. Several approaches have been proposed [19][20][21][22] using the volume based 3D reconstruction. The one that is called 'voxel based' is the most commonly used. Fig. 24 illustrates the 'voxel based' method. Basically, we divid the volume of interest into multiple discrete voxels. For each voxel, we back-project it to the 2D camera image plane. If the back-projected area lies completely outside the silhouette area, then we classify this voxel as an outside voxel, which means it does not belong to the 3D object region. The pseudo code for this method is summarized in Fig. 25 (Suppose we have K silhouettes.)

- We divide the volume of interest into $N \times N \times N$ discrete voxels v_n, $n = 1 \cdots N^3$.
- Initialize all the N^3 voxels as inside voxels.
- For $n = 1$ to N^3 {
  - For $k = 1$ to K {
    - Back-project v_n into the k^{th} image plane.
    - If the back-projected area lies completely outside the k^{th} silhouette, then classify v_n as an outside voxel.

  }
}
- The Visual hull can be regarded as the union of inside voxels v_n

Figure 25: Pseudo code for 'voxel based' 3D reconstruction method.

We can observe from Fig. 23 and Fig. 24 that using the surface based visual hull reconstruction method, if the object consists of a lot of irregular surfaces and textures, like a stone, the computational complexity will be extremely high and will require a lot more views to do the reconstruction. But if the object's surface is relatively flat and there is not much texture on the object, the surface based method can be a good choice. Buehler et al. [23] proposed a fast algorithm for computing visual cone intersections for a 3D polyhedral object, which will be helpful in our scenario, because the nail can be regarded as the intersection of several polyhedral shapes and does not have irregular textures on top of it. For the volume based 3D reconstruction method, the processing efficiency is much higher and it has a better representation for the objects with irregular, textured surfaces, like stones. But the drawback is also very obvious: The volume of the reconstruction result will be significantly bigger than the true volume. It can be seen in Fig. 24 that the shaded color yellow is the three dimensional reconstruction result and the bounded dark green is the two dimensional reconstruction result, the volume of yellow is always bigger than the volume of green. $VH(yellow) \supseteq VH(green)$. After all these discussions, considering the smooth surface of the fingernail and the computational efficiency using the method in [23], we will use a two dimensional surface based representation to reconstruct the nail's shape in the following experiment.

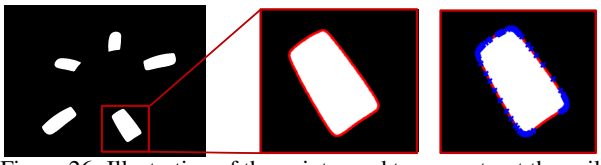


Figure 26: Illustration of the points used to reconstruct the nail surface. Red points: edge points. Blue points: points that are on the mask's convex hull.

4.3 Nail Reconstruction Result

Because we want to approximate general 3D objects by polyhedral shapes when intersecting visual cones [23], the edge points in each silhouette will be critical in the calculation. In order to decrease the computational complexity, the points that lie inside the silhouette will not be calculated in this case. Fig. 26 shows a sample silhouette image from one viewpoint in the double mirror single camera system. If we only look at the silhouette that is bounded by the red box, the red points are the edge points of this silhouette. But in order to increase the computational efficiency and because of the idea of polyhedral intersection, among these edge points, we can only keep the points that belong to the convex hull of the mask to do the reconstruction. The rightmost image in Fig. 26 shows the points on the edge that belong to the

convex hull. These points are highlighted with blue crosses.

If we take a closer look at part of the image in one silhouette, as shown in Fig. 27, only a small fraction of the edge points are used to back project from the camera image plane to the visual hull. That will increase the efficiency and decrease the computational complexity dramatically.

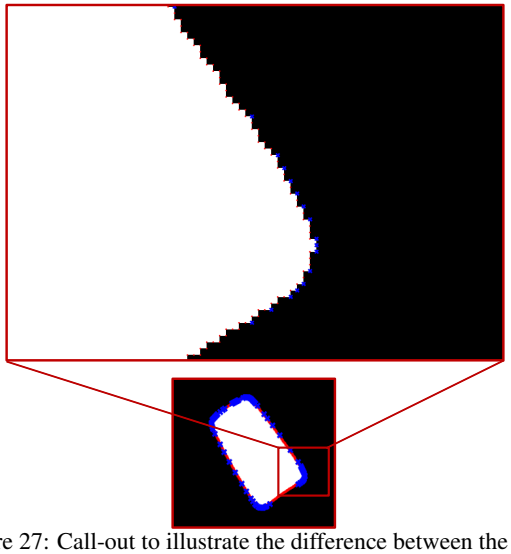


Figure 27: Call-out to illustrate the difference between the edge points and the convex hull points in one silhouette.

Based on the information above, two sets of experiments are conducted to reconstruct the nails' shape from the silhouette images.

Comparing Relative Depth

The goal of this experiment is to determine whether the reconstructed ratio (width/depth, length/depth) is correct or not. We are using four different nail sizes to do the experiment as shown in Fig. 28. The length (x), width (y), and height (z) for each nail is measured using a digital caliper as shown in Fig. 29. We compare the ratio of x/z and y/z between the ground truth and the reconstruction results to estimate the reconstruction accuracy of our system. (Note that the depth is the maximum height of the nail). The experimental results are listed in Fig. 30. The

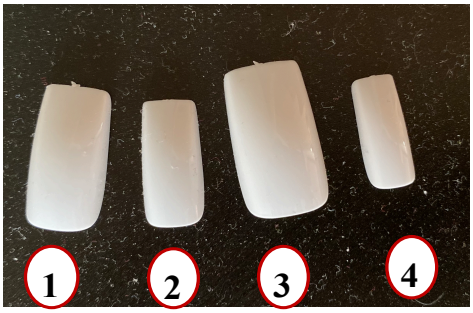


Figure 28: Four different nail sizes.

columns with red bounding boxes are the ground truth, double mirror results and the accuracy of the ratio of length/depth (x/z). We can see that the accuracy is above 93% for all four nails. The columns with blue bounding boxes are the ground truth, double

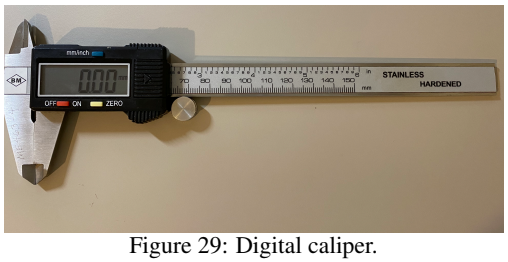


Figure 29: Digital caliper.

mirror results and the accuracy of the ratio of width/depth (y/z). The accuracy is above 92% for all four nails. In this case, we can come to the conclusion that based on the relative depth, the reconstruction accuracy is above 92%.

	x len (mm)	y len (mm)	z len (mm)	x/z (gt)	y/z (gt)	x/z (double mirror)	y/z (double mirror)	Accuracy (x/z)	Accuracy (y/z)
Nail 1	22.33	10.93	5.23	4.27	2.09	4.20	2.01	98%	96%
Nail 2	21.30	9.90	3.96	5.38	2.50	5.15	2.30	96%	92%
Nail 3	24.65	12.93	5.72	4.31	2.26	4.00	2.18	93%	96%
Nail 4	17.20	7.44	3.32	5.18	2.24	5.09	2.31	98%	97%

Figure 30: Relative nail dimensions comparison results.

Reference Height

The goal of this experiment is to measure the absolute height of the nails, with the knowledge of a known reference height. We will also need four nails, but one nail's maximum height is known in advance. Fig. 31 shows the four nails with Nail 4's height known as 3.32 mm. As we did before, we will take two images for each nail from two different viewpoints, using the same camera setting. But this time, the two viewpoints will always stay the same for all four nails. In other words, we will have two fixed viewpoints; and we will take images for all the nails from the two fixed viewpoints. Also, the mirror's position and the angle between the two mirrors cannot be changed when we take pictures of all four nails. We use the ratios between the other three unknown nails' maximum heights to the one known nail's maximum height in the reconstruction results. And the ratio will be used to multiply the known nail's height, which is measured using a digital caliper. The multiplication results will be the nails' estimated absolute maximum heights. And we will compare the estimated results to the real maximum heights.

Fig. 32 shows the estimation results. We can observe that the absolute differences between the real height and the estimated height are around $0.1mm$ to $0.3mm$, which means the accuracy is above 93% in the experiment. So if we can locate a reference height in the camera image plane, the nails' maximum height can be estimated and the depth map can be calculated in this case. In our nail printer product, it will be reasonable to put a small ball inside of the printer, and make sure this small ball is visible in the camera image plane, and the dimension of the small ball will be known in advance. In this case, we can estimate each nail's absolute maximum height using the proposed double mirror and single camera system within a small size printer. And we are guaranteed high accuracy.

4.4 one shot Reconstruction

In this part, we will simplify the reconstruction process from two shots images to just one shot and compare the performance between these two methods. First we will introduce the differ-

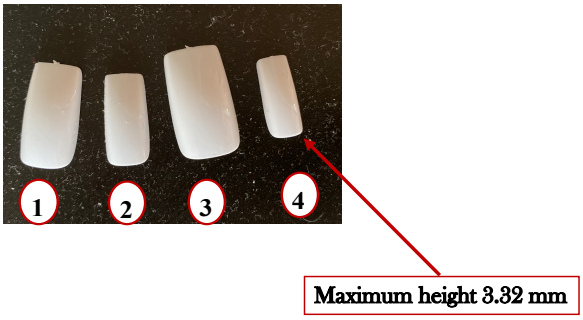


Figure 31: Four different nails where one nail's maximum height is known.

	z len (gt) (mm)	z len (scaled)	Ratio	Height (mm) (estimated)	Difference (mm)	Accuracy
Nail 1	5.23	0.293	1.48	5.12	0.11	98%
Nail 2	3.96	0.210	1.06	3.70	0.26	93%
Nail 3	5.72	0.343	1.73	5.97	-0.25	96%
Nail 4	3.32	0.198	N/A	N/A	N/A	N/A

Figure 32: Absolute maximum heights comparison.

ence in the camera calibration process when we are using one shot method.

From Sec. 3 we know that in order to locate the position of the principal point $\mathbf{p_0}$ in the camera image plane, we need to find the intersection point of the lines that are perpendicular to the epipole line in two different viewpoints, as shown in Fig. 33. But if just one image is taken for each object, there is no way to find the intersection point. In that case, we simply estimate the principal point's location to be at the center of the image plane.

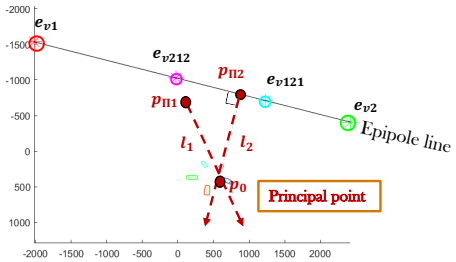


Figure 33: Finding the principal point under the two shots condition.

Fig. 34 shows the difference of the principal point's position between the two shots condition and the one shot condition. It is very obvious that the location of principle point shifts about 100 pixels in both x and y directions. But as we stated at the end of Sec. 3, a single parameter cannot guarantee a good or bad calibration result. The way to justify the calibration result is to see whether it can provide a good mapping between the 2D image plane point to the 3D world coordinates point. So we will take a look at the 3D reconstruction results next.

The reconstruction results for the four nails in Fig. 28 under the two shots condition and the one shot condition are shown in Fig. 35 and Fig. 36, respectively. We can see that the nails' shapes do not change much from two shots to one shot. This means that even though we may have a worse estimation of the principal point in the one shot condition, the mapping between the 2D image plane and the 3D object is still reasonable and acceptable.

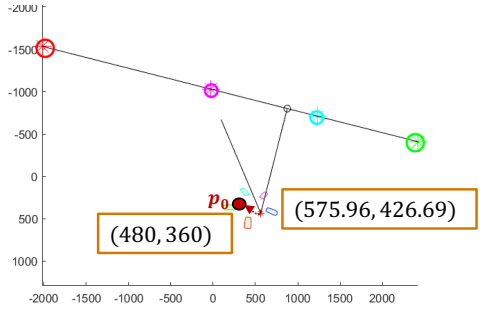


Figure 34: Finding the principal point under the one shot condition.

Figure 35: two shots reconstruction results.

Figure 36: one shot reconstruction results.

5. Conclusion

In this paper, we introduced how to build a low-cost single camera, double mirror system that can be fit into a portable desktop size nail printer. With this system, we can use the reflection of the mirrors to calibrate the camera with only two images. By applying the visual hull 3D reconstruction method, the depth map of the fingernail can be acquired. A further simplified one shot method has been introduced, which can be used to further improve the computational efficiency. Four different nail sizes have been tested for both two shots and one shot method, and the rendering results are very convincing. Also, based on the comparison of the relative depth and the comparison of the absolute depth, the accuracy is over 92%. Considering the actual size of the nails, the accuracy of the approach is acceptable for the final product.

References

[1] J. Dong, K. R. Bengtson, B. Robinson, and J. P. Allebach, "Low-cost structured-light based 3D capture system design." In Three-Dimensional Image Processing, Measurement (3DIPM), and Applications 2014, vol. 9013, p. 90130B. International Society for Optics and Photonics, 2014.

[2] L. Yang, K. R. Bengtson, L. Li, and J. P. Allebach, "Design and decoding of an M-array pattern for low-cost structured light 3D reconstruction systems." In 2013 IEEE International Conference on Image Processing, pp. 2168-2172. IEEE, 2013.

[3] D. Lanman , and G. Taubin, "Build your own 3D scanner: 3D photography for beginners." In ACM SIGGRAPH 2009 Courses, pp. 1-94. 2009.

[4] C. Albitar, P. Graebling, and C. Doignon, "Design of a monochromatic pattern for a robust structured light coding." In 2007 IEEE International Conference on Image Processing, vol. 6, pp. VI-529. IEEE, 2007.

[5] Y. Wang, D. He, Z. Lin, B.Robinson, G. Chiu, and J. Allebach, "3D shape estimation for smooth surfaces using structured light patters." 3D Measurement and Data Processing, vol. 2020, no. 17, pp. 35–1, 2020.

[6] Logitech C920S Webcam [Online].Available:https://www.logitech.com/en-us/product/hd-pro-webcam-c920s?crid=34

[7] Z. Zhang, "A flexible new technique for camera calibration." IEEE Transactions on Pattern Analysis and Machine Intelligence 22, no. 11 (2000): 1330-1334.

[8] K. Forbes, F. Nicolls, G. de Jager, and A. Voigt, "Shape-from-silhouette with two mirrors and an uncalibrated camera." In European Conference on Computer Vision, Springer, 2006, pp. 165-178, Berlin, Heidelberg.

[9] J.Bouguet, "Camera calibration toolbox for Matlab." (2010). http://www.vision.caltech.edu/bouguetj/calib-doc.

[10] P. Lai, A. Yilmaz, "Projective reconstruction of building shape from silhouette images acquired from uncalibrated cameras," In ISPRS Congress Beijing 2008, Proceedings of Commission III.

[11] K. Forbes, A. Voigt, and N. Bodika. " Visual hulls from single uncalibrated snapshots using two planar mirrors," IAPR, 2004.

[12] RN. PP, and S. Jabbar, "Efficient 3D visual hull reconstruction based on marching cube algorithm." International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), pp. 1-6, IEEE, 2015.

[13] C. Esteban, and F. Schmitt, "Silhouette and stereo fusion for 3D object modeling." Computer Vision and Image Understanding, vol. 96, no. 3, pp. 367-392, 2004.

[14] P. Besl, "Geometric modeling and computer vision." Proceedings of the IEEE, vol. 76, no. 8, pp. 936-958, 1988.

[15] K. Cheung, "Visual hull construction, alignment and refinement for human kinematic modeling, motion tracking and rendering." PhD thesis, Tech. Report, CMU-RI-TR-03-44, Robotics Institute, Carnegie Mellon University, 2003.

[16] A. Laurentini, "The visual hull: A new tool for contour-based image understanding." Proc. 7th Scandinavian Conf. Image Analysis, vol. 993, pp. 1002, 1991.

[17] A. Laurentini, "The visual hull concept for silhouette-based image understanding." IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 16, no. 2, pp. 150-162, 1994.

[18] A. Laurentini, "How far 3D shapes can be understood from 2D silhouettes." IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 17, no. 2, pp. 188-195, 1995.

[19] M. Potmesil, "Generating octree models of 3D objects from their silhouettes in a sequence of images." Computer Vision, Graphics, and Image Processing, vol. 40, no. 1, pp. 1-29, 1987.

[20] H. Noborio, S. Fukuda, and S. Arimoto, "Construction of the octree approximating a three-dimensional object by using multiple views." IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 10, no. 6, pp. 769–782, 1988.

[21] N. Ahuja, and J. Veenstra, "Generating octrees from object silhouettes in orthographic views." IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 11, no. 2, pp. 137-149, 1989.

[22] R. Szeliski, "Rapid octree construction from image sequences." CVGIP: Image understanding, vol. 58, no. 1, pp. 23-32, 1993.

[23] W. Matusik, C. Buehler, and L. McMillan, "Polyhedral visual hulls for real-time rendering." Eurographics Workshop on Rendering Techniques, pp. 115–125, 2001.

Author Biography

Yin Wang *received her B.S.(2016) in Electrical Engineering from Purdue University and received her Ph.D. from Purdue University in 2020. Her research focuses on color image processing, computer vision, 3D reconstruction, and deep learning applications.*

Peng(Davi) He *(Note: 'He' is his last name) is a software manager in Sunvalleytek International Inc, Shenzhen, Guangdong, China. He received his B.S in Communication Engineering from Hunan University of Arts and Science, Changde, Hunan, China in 2012. His current research interest including artificial intelligence and robot.*

Zillion(Zhenqing) Lin *is a technical director in Sunvalleytek International Inc, Shenzhen, Guangdong, China. He received his BS in Microelectronics Engineering from University of Electronic Science and Technology of China, Chengdu, Sichuan, China in 2008 and MS in Biomedical Engineering from Shenzhen University, Shenzhen, Guangdong, China. His current research interest including artificial intelligence and robot.*

George T. Chiu *is a Professor in the School of Mechanical Engineering with courtesy appointments in the School of Electrical and Computer Engineering and the Department of Psychological Sciences at Purdue University. He also serves as the Assistant Dean for Global Engineering Programs and Partnership for the College of Engineering. Dr. Chiu received the B.S. degree in Mechanical Engineering from the National Taiwan University in 1985 and the M.S. and Ph.D. degrees in Mechanical Engineering from the University of California at Berkeley, in 1990 and 1994, respectively. From September 2011 to June 2014, he served as the Program Director for the Control Systems Program at the National Science Foundation. His current research interests are mechatronics and dynamic systems and control with applications to digital printing and imaging systems, digital fabrications and functional printing, human motor control, motion and vibration perception and control. He received the 2012 NSF Director's Collaboration Award and the 2010 IEEE Transactions on Control System Technology Outstanding Paper Award. He served as the Editor-in-Chief for the IEEE/ASME Transactions on Mechatronics from 2017-19 and as the Editor for the Journal of Imaging Science and Technology from 2012-14. Dr. Chiu served on the Executive Committee of the ASME Dynamic Systems and Control Division (DSCD) from 2007 to 2014 and as the Chair of the Division from 2012-13. He is a Fellow of ASME and a Fellow of the Society for Imaging Science and Technology (IS&T)*

Jan P. Allebach *is Hewlett-Packard Distinguished Professor of Electrical and Computer Engineering at Purdue University. Allebach is a Fellow of IEEE, IS&T, and SPIE. He was named Electronic Imaging Scientist of the Year by IS&T and SPIE, and was named Honorary Member of IS&T, the highest award that IS&T bestows. He has received the IEEE Daniel E. Noble Award, the IS&T/OSA Edwin Land Medal, the IS&T Gutenberg Prize, is a Fellow of the National Academy of Inventors, and is a member of the National Academy of Engineering.*

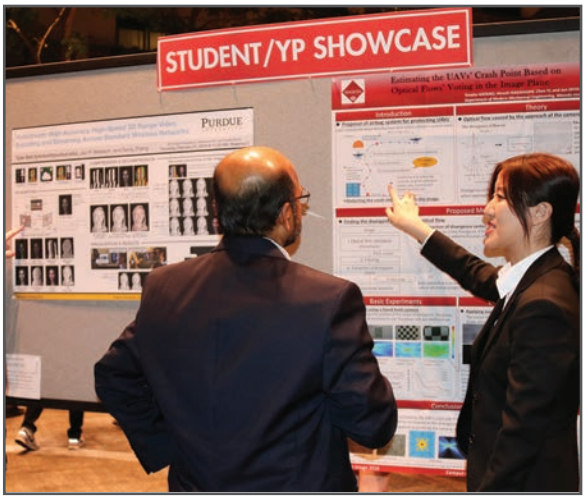
JOIN US AT THE NEXT EI!

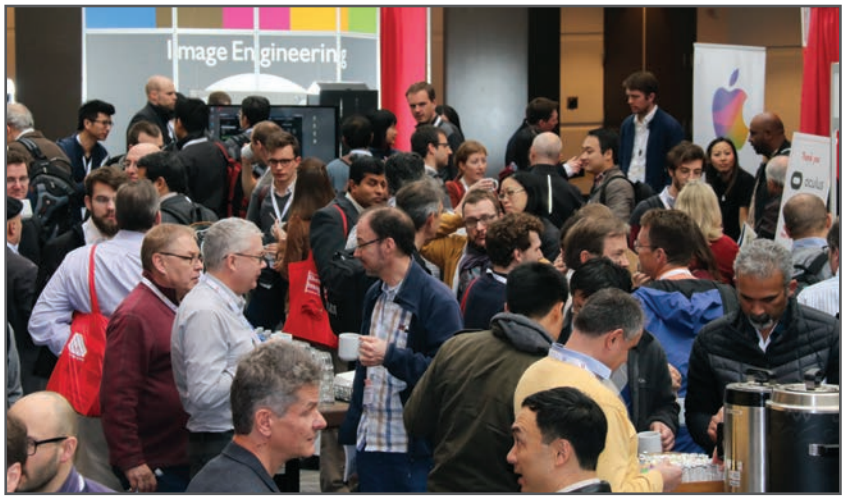
IS&T International Symposium on

Electronic Imaging

SCIENCE AND TECHNOLOGY

Imaging across applications . . . Where industry and academia meet!





• SHORT COURSES • EXHIBITS • DEMONSTRATION SESSION • PLENARY TALKS •
• INTERACTIVE PAPER SESSION • SPECIAL EVENTS • TECHNICAL SESSIONS •

www.electronicimaging.org

IS&T
imaging.org