Principal Observation Ray Calibration
for Tiled-Lens-Array Integral Imaging Display

Weiming Li[1]   Haitao Wang[1]   Mingcai Zhou[1]   Shandong Wang[1]   Shaohui Jiao[1]

Xing Mei[1,3]   Tao Hong[1]   Hoyoung Lee[2]   Jiyeun Kim[1,2]

[1] Samsung Advanced Institute of Technology China (SAIT China), Beijing, China

[2] Samsung Advanced Institute of Technology (SAIT), Yongin, Korea

[3] National Laboratory of Pattern Recognition (NLPR), CASIA, Beijing, China

{weiming.li,ht.wang,mingcai.zhou,sd.wang,sh.jiao,tao31.hong,

hyounglee,jiyeun.kim}@samsung.com, xmei@nlpr.ia.ac.cn

Abstract

Integral imaging display (IID) is a promising technology to provide realistic 3D image without glasses. To achieve a large screen IID with a reasonable fabrication cost, a potential solution is a tiled-lens-array IID (TLA-IID). However, TLA-IIDs are subject to 3D image artifacts when there are even slight misalignments between the lens arrays. This work aims at compensating these artifacts by calibrating the lens array poses with a camera and including them in a ray model used for rendering the 3D image. Since the lens arrays are transparent, this task is challenging for traditional calibration methods. In this paper, we propose a novel calibration method based on defining a set of principle observation rays that pass lens centers of the TLA and the camera's optical center. The method is able to determine the lens array poses with only one camera at an arbitrary unknown position without using any additional markers. The principle observation rays are automatically extracted using a structured light based method from a dense correspondence map between the displayed and captured pixels. Experiments show that lens array misalignments can be estimated with a standard deviation smaller than 0.4 pixels. Based on this, 3D image artifacts are shown to be effectively removed in a test TLA-IID with challenging misalignments.

1. Introduction

Integral imaging display (IID) is a promising 3D display technology that allows users to see natural color 3D images with continuous horizontal and vertical parallax without wearing glasses. Originated by Lippmann [8], IID has regained interests in recent decades with the development of electronic technologies that make IID hopeful to be com-
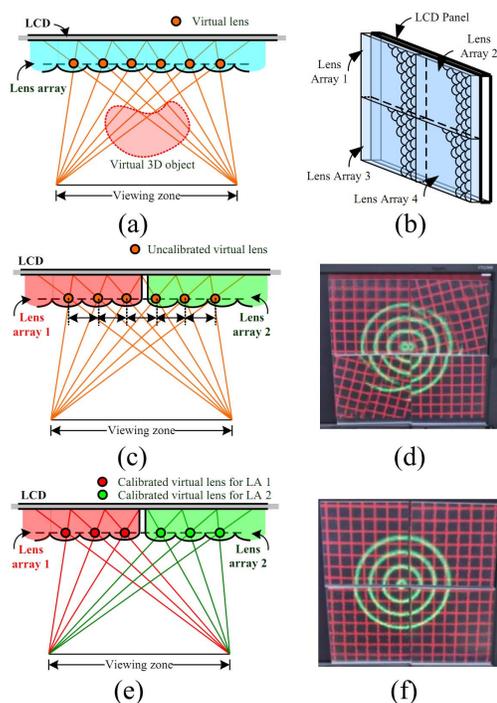

Figure 1. The proposed calibration method aims at eliminating 3D image artifacts for a tiled-lens array (TLA) integral imaging display (IID). (a) shows the principle of an IID. (b) is an example of TLA-IID. Uncalibrated virtual lenses mismatch the actual ones as shown in (c), which leads to 3D image artifacts as seen in (d). Calibrated virtual lenses are consistent with the actual ones as in (e), which creates correct 3D image as seen in (f).

mercialized. A survey of IID technology can be found in the literature [7, 9, 5]. Basically, a typical IID consists of a high resolution LCD panel and a lens array as shown in figure 1(a). A 2D image named as integral image is displayed on the LCD and it is refracted by the lens array to different

directions to form images in 3D space. The integral image is created by simulating 3D objects and light rays through a virtual lens array (a computational model) from a viewing zone. When the virtual lens array is consistent with the physical lens array, correct 3D image can be seen within the viewing zone.

To provide a more realistic and immersive experience, developing a large screen IID is desired. Though the size and resolution of nowadays commercialized LCDs have been ever increasing, building a large lens array with high precision at reasonable cost is still a challenging task for today's optics fabrication industry. To work around this, one alternative is to build a tiled lens array by mosaic of smaller lens arrays as shown in figure 1(b), which we refer to as a tiled-lens-array IID (TLA-IID). In a TLA-IID, each lens array is small and thus is easier to fabricate with precision at low cost. However, difficulties arise to align these lens arrays due to the precision limit of mounting tools and errors induced by mechanics and temperature variations during usage. With the presence of misalignments, virtual lenses without calibration are not consistent with the physical ones (see figure 1(c)), which would result in 3D image artifacts across different lens arrays as shown in figure 1(d). In order to compensate these errors and provide a consistent 3D image across all the lens arrays from all the viewpoints in the viewing zone, calibration is necessary. Here the calibration aims at obtaining the actual lens array poses with precision to create a correct (calibrated) virtual lens array. Then by using these calibrated virtual lenses (see figure 1(e)) for rendering 3D image, the 3D image artifacts can be removed as shown in figure 1(f).

There were some previous works on calibrating tiled 3D displays. Dependent on the working principals to create 3D images, the calibration methods are different. A large tiled 3D display named Varrier was reported in [4, 12], which consists of 35 autostereoscopic displays. The system was calibrated with a stereo camera pair, where each camera deals with one of the stereo views. As for our system, an extension of this method would require an expensive camera array to handle the tens of views in both horizontal and vertical directions. In [6], the authors presented a multi-viewer tiled automultiscopic display which consists of an array of lenticular based multi-view displays. The calibration is performed by adjusting parameters manually until color stripe patterns projected on a white card from the displays are visually aligned. When a large number of displays are tiled and need alignment by both rotation and translation, manual efforts for such adjustment would increase drastically. A camera based method was proposed in [13] for an IID with a lens array. The authors corrected errors in 3D images by shifting pixels of the LCD image with interpolation. Instead of using interpolation, our work aims at obtaining a correct geometry model for the lenses and rendering the 3D scene accordingly. To the best of our knowledge, our effort is among the first camera based methods to calibrate an IID with tiled lens arrays.

In addition, we notice that there are some previous calibration works on multi-projector displays [3, 10, 14, 1].

These works perform well in removing 2D image artifacts projected on a large-format surface screen. Different from these, we mainly focuses on correcting 3D image conveyed through lens arrays with IID methods. The 3D image here presents different image contents for different viewing directions, which all need to be aligned simultaneously.

In general, this work follows a computer vision approach to calibrate a TLA-IID based on camera captured images and 3D constraints. Note that the method is designed to work with the TLA-IID as-built after assembly. Compared to traditional calibration problems, some special challenges are: (1) the lens arrays are transparent and they are not directly visible to camera; (2) after assembly, light rays from the LCD are refracted by lenses which bend their propagation directions; (3) high precision is required to estimate the lens array poses in order to ensure correct 3D images at various views simultaneously. It is known that to avoid artifacts of 3D images such as cross talk, estimated displacement error should be smaller than half the size of an LCD pixel. For the above issues, contributions of this work are:

(1) The geometry of light rays captured by a camera through the lens array refraction is formulized. Particularly, we define a set of principle observations rays (PORs) that pass the lens centers and the camera's optical center. Based on the PORs, the lens array pose with respect to the LCD can be explicitly computed with images taken at only one camera pose without any additional calibration markers, which leads to an efficient calibration procedure.

(2) An automatic method is proposed to extract the PORs. The problem is that PORs are defined by lenses that are transparent to camera. For this, we explore structured lights to establish a dense correspondence map between the LCD pixels and camera pixels. With this map, the lenses are allowed to be extracted with precision based on detecting geometric variations, which is robust to illumination-induced image intensity noises.

(3) With the calibrated virtual lens arrays, an efficient method is presented to create a ray model that relates a correct light ray with each integral image pixel. The ray model allows creating integral images which show correct 3D images despite of the TLA-IID misalignment errors.

The calibration method does not need any markers to be attached to the TLA-IID and it can be performed under natural light with only one camera. This makes the method easy to be performed within a short time at the site of usage. Experimental results show that our method performs well with a precision of less than 50 um (0.4 pixels in regard to the used LCD). After calibration, 3D images can be viewed correctly from all the viewing directions.

The proposed method serves as a software approach to compensate optical hardware errors and provides a cost-effective solution for building 3D naked-eye display with large screens. This is useful for a large variety of applications such as 3D content presentation for public advertisement, entertainment, education, and exhibition where large screens are desired with reasonable cost.

2. The TLA integral imaging display system

An experimental TLA-IID was setup, for which the proposed calibration method was applied and tested. The TLA-IID consists of an LCD panel and four lens arrays (figure 2). Here the LCD monitor is a ViewSonic 2290b, which has a pixel pitch of 0.1245 mm and a resolution of 3840 pixels by 2400 pixels. Each of the four lens arrays is a hexagon lens array, where the lens pitch in horizontal direction is 2.59 mm and the lens pitch in vertical direction is 1.92 mm. The focal length of the lens array is 7 mm and the thickness is the same. The width of a lens array plate is 147.74 mm (57 lenses) and the height is 143.86 mm (74 lenses).

In practical TLA-IIDs, there are usually some misalignments between the lens arrays. For providing a challenging test in our experiments, two screws are intentionally inserted as spacers to enlarge rotational and translational misalignments among the lens arrays as shown in figure 2(b) and 2(d). Due to this, a gap is shown near the center of the TLA-IID's region of interest (ROI) as shown in figure 2(c).

3. Calibration of the TLA-IID

3.1. Definitions and assumptions

The lens arrays in the TLA are computationally modeled by a set of virtual lens arrays (VLAs). With these, calibration aims at estimating poses of the VLAs with respect to the LCD. Then the calibrated VLAs are used to create a ray model for rendering correct 3D image (section 4).

Specifically, this work assumes that: (1) Each virtual lens in the VLAs is a thin lens with a unique optical center. (2) All the virtual lens centers are on a plane parallel to the LCD plane at a distance that equals to the lens focal length. (3) The parameters of each VLA are known as provided by its design, which include lens shape, lens arrangement, lens pitch, focal length, and lens array size. Finally, the camera used is considered as a pinhole camera.

Following these, the VLA poses are constrained in a 2D plane as shown in figure 3. Then the purpose of calibration is to obtain the rotation and translation of each VLA with

Figure 3. Definition of the VLA poses to the LCD.

respect to the LCD.

3.2. Principle observation rays

Basically, this work follows a computer vision approach by exploring camera images and 3D geometric constraints. The light ray geometry when a camera observes the LCD panel through a lens array is shown in figure 4. Since the camera is set at a distance to the lens array and each lens is small, the light rays captured in the camera's image related to a single lens span a tiny angle and can be considered to be parallel. Note that this can be valid even the camera is not an orthogonal camera. Considering the assumption that the LCD panel is set at the focal distance of the lens, it is well known by optics that these parallel light rays would focus on an LCD pixel.

It can be seen that among all the light rays captured by the camera from a specific lens, only the light ray through the lens's optical center does not change its propagation direction. We refer to such a light ray that passes both the
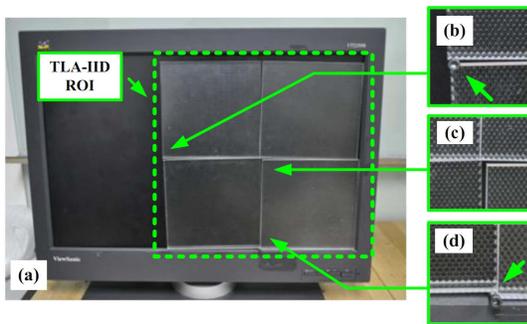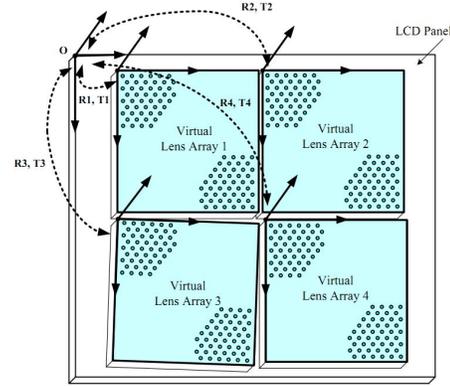

Figure 2. Setup of an experimental TLA-IID. The four lens arrays constitute a region of interest (ROI) as shown in (a). (b) and (d) show the two screws used as spacers to create misalignments for test. (c) shows the detail of a local region near the ROI center.
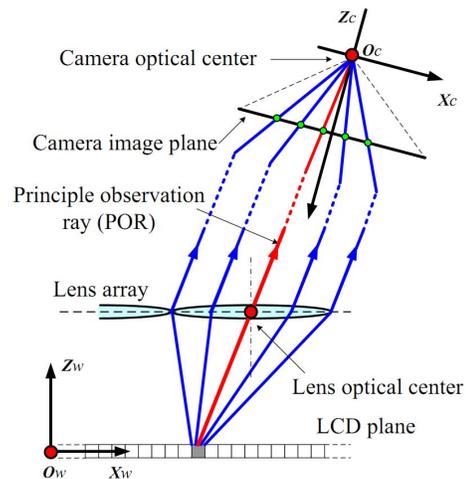

Figure 4. Illustration of the imaging geometry when the LCD is seen by a camera through the lens array. (Not to scale.)
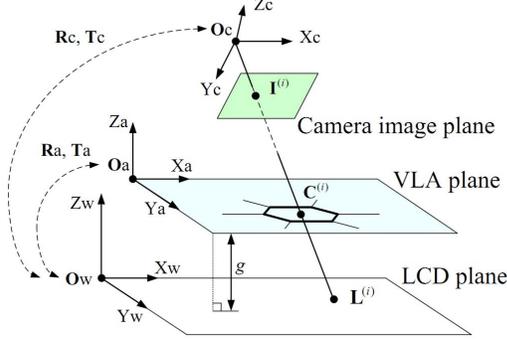
Figure 5. Illustration of the principal observation ray (POR) model to estimate virtual lens array (VLA) pose.

camera's optical center and a lens's optical center as a principle observation ray (POR). Each lens in the lens array provides a related POR. The set of PORs are consistent with the geometry when the pinhole camera is used to observe the LCD plane without the lens array. Thus, by exploring the 3D geometric constraints between the LCD, the PORs, and the camera image, we find that the VLA's pose with respect to the LCD can be estimated explicitly, for which the details will be introduced in section 3.3.

As shown in figure 4, the pixel related to a POR in the camera's image can be extracted as the center of a lens in the image. A challenge here is that the lenses are transparent and their boundaries are not directly visible. To solve this problem, we explore structured light technology. Structured light image patterns are displayed on the LCD and they are captured by a camera. By decoding the captured structured light images, a dense correspondence map between the LCD pixels and camera pixels can be obtained. Then the lens boundaries are extracted by detecting the drastic variations in the correspondence map. The LCD pixel locations related with the PORs are also known from the LCD-camera map once the POR image pixels are located. The structure light method in this work is specified in section 5.1.

3.3. VLA pose estimation with PORs

To estimate VLA pose, we define coordinate systems as shown in figure 5. A world coordinate system $O_w X_w Y_w Z_w$ is defined on the LCD plane. A VLA coordinate system $O_a X_a Y_a Z_a$ is defined on the VLA plane. The camera's coordinate system is defined as $O_c X_c Y_c Z_c$, where O_c is the camera's optical center.

Consider the center of the ith lens in the kth lens array. Let it be denoted by $C^{(i)}=[C_x^{(i)}, C_y^{(i)}, C_z^{(i)}]^T$ in $O_w X_w Y_w Z_w$ and by $\hat{C}^{(i)}=[\hat{C}_x^{(i)}, \hat{C}_y^{(i)}, \hat{C}_z^{(i)}]^T$ in $O_a X_a Y_a Z_a$. Following the assumptions in section 3.1, the motion between $O_w X_w Y_w Z_w$ and $O_a X_a Y_a Z_a$ is simplified to a 2D rotation and translation and we have:

$$\begin{bmatrix} C_x^{(i)} \\ C_y^{(i)} \end{bmatrix} = \begin{bmatrix} cos(a^{(k)}) & -sin(a^{(k)}) \\ sin(a^{(k)}) & cos(a^{(k)}) \end{bmatrix} \begin{bmatrix} \hat{C}_x^{(i)} \\ \hat{C}_y^{(i)} \end{bmatrix} + \mathbf{t}^{(k)} \tag{1}$$

where $a^{(k)}$ is the rotational angle of the kth lens array in the VLA plane and $\mathbf{t}^{(k)}=[t_x^{(k)}, t_y^{(k)}]$ is translation of the kth lens array. Here $C_z^{(i)} = g$, where g is the distance between the lens center to the LCD. As assumed in section 3.1, g equals to the lens's focal length that is known.

Figure 5 shows the example of a POR that passes a lens optical center $\mathbf{C}^{(i)}$. The POR originates from an LCD pixel $\mathbf{L}^{(i)}$ and it is captured by a camera pixel $\mathbf{I}^{(i)}$. Based on structured light decoding result, the LCD-to-image pixel correspondence pair $\{\mathbf{L}^{(i)}, \mathbf{I}^{(i)}\}$ is known.

Given a set of $\{\mathbf{L}^{(i)}, \mathbf{I}^{(i)}\}$ for a number of PORs, the camera's rotation and translation $\mathbf{R}_c$ and $\mathbf{T}_c$ with respect to $O_w X_w Y_w Z_w$ can be estimated with a well-known planar homography based method [15], where the LCD is used as an active planar calibration object. Then the camera's optical center $\mathbf{O}_c=[O_x, O_y, O_z]^T$ is known. Now, consider the three points $\mathbf{O}_c$, $\mathbf{C}^{(i)}$, and $\mathbf{L}^{(i)}$, which are located on the same POR. Based on the co-linear property, it can be derived that:

$$\begin{cases} (L_x^{(i)} - O_x)(C_z^{(i)} - O_z) + (C_x^{(i)} - O_x)O_z = 0 \\ (L_y^{(i)} - O_y)(C_z^{(i)} - O_z) + (C_y^{(i)} - O_y)O_z = 0 \end{cases} \tag{2}$$

This can be written as:

$$\begin{cases} a_1^{(i)} C_x^{(i)} + b_1^{(i)} C_y^{(i)} + c_1^{(i)} C_z^{(i)} = d_1^{(i)} \\ a_2^{(i)} C_x^{(i)} + b_2^{(i)} C_y^{(i)} + c_2^{(i)} C_z^{(i)} = d_2^{(i)} \end{cases} \tag{3}$$

where

$$a_1^{(i)} = O_z, b_1^{(i)} = 0, c_1^{(i)} = L_x^{(i)} - O_x, d_1^{(i)} = L_x^{(i)} O_z$$
$$a_2^{(i)} = 0, b_2^{(i)} = O_z, c_2^{(i)} = L_y^{(i)} - O_y, d_2^{(i)} = L_y^{(i)} O_z$$

Here since $\mathbf{O}_c$ and $\mathbf{L}^{(i)}$ are known, the parameters in the equation (3) are known.

Then by substituting equation (1) to equation (3) and let the unknown be $\mathbf{G} =[cos(a^{(k)}), sin(a^{(k)}), t_x^{(k)}, t_y^{(k)}]^T$, we have

$$\begin{bmatrix} a_1^{(i)} \hat{C}_x^{(i)} & -a_1^{(i)} \hat{C}_y^{(i)} & a_1^{(i)} & 0 \\ b_2^{(i)} \hat{C}_y^{(i)} & -b_1^{(i)} \hat{C}_x^{(i)} & 0 & b_2^{(i)} \end{bmatrix} \mathbf{G} = \begin{bmatrix} cd_1^{(i)} - c_1^{(i)} g \\ d_2^{(i)} - c_2^{(i)} g \end{bmatrix} \tag{4}$$

Let equation (4) be rewritten as: $\mathbf{A}^{(i)} \mathbf{G} = \mathbf{D}^{(i)}$, then by combining a number of M PORs from the same VLA:

$$\mathbf{AG} = \mathbf{D} \tag{5}$$

where

$$\mathbf{A} = [(\mathbf{A}^{(1)})^T, (\mathbf{A}^{(2)})^T, \ldots, (\mathbf{A}^{(M)})^T]^T,$$
$$\mathbf{D} = [(\mathbf{D}^{(1)})^T, (\mathbf{D}^{(2)})^T, \ldots, (\mathbf{D}^{(M)})^T]^T.$$

Then, the unknown can be obtained with a least mean

square estimation as:

$$\mathbf{G} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{D} \tag{6}$$

The above computations are performed for each VLA to obtain its pose in $\mathbf{O}_w\mathbf{X}_w\mathbf{Y}_w\mathbf{Z}_w$.

3.4. The calibration method

The calibration method needs no additional devices and requires only one camera to be placed at an unknown position towards the TLA-IID. The calibration method can be summarized as the following steps:

(**a**) Display a set of structured light images on the TLA-IID and capture each image with the camera at a fixed pose.

(**b**) Decode the structured light images to obtain a dense correspondence map between the LCD pixels and the camera pixels (an LCD-camera map).

(**c**) Extract lens boundaries from the LCD-camera map. Based on these, detect the number of lens arrays and extract the set of POR pixels for each lens array with image processing.

(**d**) Estimate the camera pose in $\mathbf{O}_w\mathbf{X}_w\mathbf{Y}_w\mathbf{Z}_w$ using the POR pixels with a planar homography method [15].

(**e**) For each of the VLA, estimate its pose in $\mathbf{O}_w\mathbf{X}_w\mathbf{Y}_w\mathbf{Z}_w$ with the method in section 3.3.

For a number of N_A lens arrays, the calibration result is a set of parameters $a^{(k)}$, $t_x^{(k)}$, $t_y^{(k)}$, where $k = 1,2,\ldots, N_A$ and the parameters are defined as in equation (1).

4. Create ray model with a calibrated TLA-IID

With integral imaging technology, 3D image of a scene model is formed by displaying an integral image on the LCD. The integral image is a 2D image created by simulating light ray intersections with the scene surfaces according to a ray model. Specifically, the ray model relates each pixel in the integral image to a light ray in 3D space. To create correct 3D image, the ray model must be consistent with the actual IID geometry. For a TLA-IID in this work, this condition is ensured by creating the ray model based on the calibrated VLAs as given by the method in section 3.

The key issue for creating the ray model is to assign each pixel in the integral image to a proper virtual lens in the VLAs according to the viewing geometry. When this is determined, the light ray related to the pixel can be obtained simply as a straight line passing the pixel and the optical center of that virtual lens.

Let an integral image pixel P be represented by its image coordinate $[m,n]^T$. Given the LCD pixel pitch, the physical location of P in the world coordinate $\mathbf{O}_w\mathbf{X}_w\mathbf{Y}_w\mathbf{Z}_w$ (see figure 6 and figure 5) is known and it is denoted as $[U(m,n), V(m,n), 0]^T$. Based on the calibrated VLA poses, since the lens array structure and shape parameters are known, the positions of all the virtual lens centers in $\mathbf{O}_w\mathbf{X}_w\mathbf{Y}_w\mathbf{Z}_w$ can be computed explicitly. Let them be represented as $\mathbf{H} = \{\mathbf{H}^{(j)}\} = \{[s^{(j)}, t^{(j)}, g]^T\}$, where $j = 1,2,\ldots,N_L$, N_L is the number of all virtual lenses in VLAs, and g equals to the lens's focal length. Assume that

P is related with the jth virtual lens, then the virtual lens center can be represented as $[S(m,n), T(m,n), g]^T$, where $S(m,n) = s^{(j)}$, $T(m,n) = t^{(j)}$. Then the light ray function in 3D space for the integral image pixel P can be determined by joining the two points $[U(m,n), V(m,n), 0]^T$ and $[S(m,n), T(m,n), g]^T$.

Figure 6 shows viewing geometry of the TLA-IID, according to which the integral image pixels are mapped to the virtual lens centers. A method to perform this mapping is proposed and summarized as the following steps:

(**1**) Based on the calibration result, compute virtual lens centers in all the VLAs and store them in the set $\mathbf{H}$.

(**2**) Assign a view distance D for the TLA-IID. Then a viewing center $\mathbf{V}_c$ can be determined, which is aligned with the center of the LCD's region of interests at distance D.

(**3**) Let an integral image pixel be represented by its coordinate $[m,n]^T$. Initialize the two pixel-to-lens matrixes by setting $S(m,n) = 0$ and $T(m,n) = 0$ for all $[m,n]^T$, where $S(m,n)$ and $T(m,n)$ are the horizontal and vertical coordinates of the virtual lens center related with the pixel $[m,n]^T$ respectively.

(**4**) For each virtual lens center $\mathbf{H}^{(j)}$ in $\mathbf{H}$, do the following. Project $\mathbf{V}_c$ though $\mathbf{H}^{(j)}$ to a point on the LCD. Centered at the projection point, define a local image window in the integral image and set the image size to be $1.2p(D + g)/(Ds)$, where p is the lens pitch size, s is the LCD pixel pitch size, and 1.2 is an empirical factor that assures overlapping of adjacent image windows. For each pixel $[q, w]^T$ in the image window:

(4.a) If $S(q,w) = 0$, let $S(q,w) = s(j)$, $T(q,w) = t^{(j)}$.

(4.b) If $S(q,w) \neq 0$, project the current pixel through the point $[S(q,w), T(q,w), g]^T$ to a plane parallel with the VLA plane at distance D to get a projection point $\mathbf{P}_1$. Also, project the current pixel through $\mathbf{H}^{(j)}$ to the same plane to get a projection point $\mathbf{P}_2$. If $|\mathbf{V}_c\mathbf{P}_1| \geq |\mathbf{V}_c\mathbf{P}_2|$, let
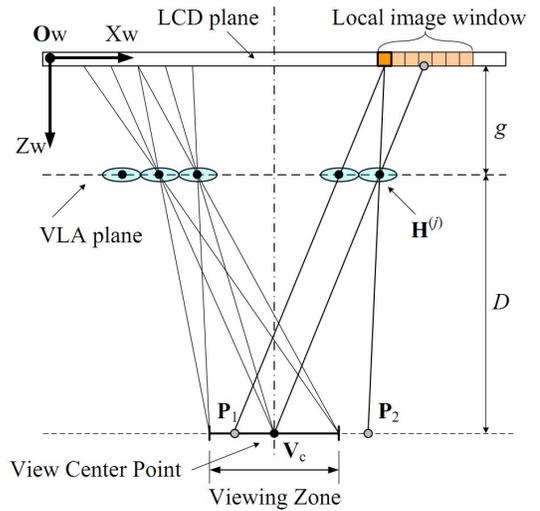

Figure 6. Illustration of the method for creating ray model based on the calibrated virtual lens arrays.

$S(q,w) = s^{(j)}$ and $T(q,w) = t^{(j)}$. Otherwise, let $S(q,w)$ and $T(q,w)$ remain their current values.

(**5**) Output the ray model that describes the light rays for all the integral image pixels with a two plane representation $\{U(m,n), V(m,n), S(m,n), T(m,n)\}$, where one plane is the LCD plane and the other plane is the VLA plane that is parallel to the LCD plane at a distance of g.

This work uses the above ray model to render integral images by intersecting light rays with virtual 3D objects and assign the surface colors at the intersection points to the related integral image pixels.

5. Experiments

5.1. Calibration procedure of the TLA-IID

For image capture, a Nikon D5000 DSLR camera is used with an 18 mm lens. The camera is connected to PC through USB and it is controlled by a program to capture the structured light image sequence automatically. The camera is calibrated with the toolbox [2].

For structured light, we used sinusoidal phase shift patterns and multi-frequency heterodyne phase unwrapping [11]. Specifically, we found a three-frequency five-step phase shift implementation gives satisfactory results. To encode LCD pixels on both horizontal (X) and vertical (Y) directions, 30 images are used. An example of one of the structured light images on the TLA-IID is shown in figure 7(a). Figure 7(b) shows an obtained LCD-camera map for the LCD's X-direction, where the color coded value related with a certain camera pixel is the X coordinate of the LCD pixel seen through the lens array. Figure 7(d) shows a local region where the lens structure can be seen. The structure appears because all the rays seen through a lens come from a same LCD pixel as explained in section 3.2.

Based on the LCD-camera map, lens boundary extraction is performed by detecting pixels where the LCD-camera map varies drastically. A DoG filter is applied and
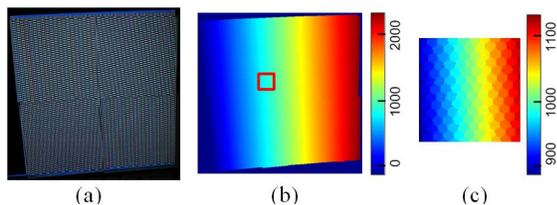

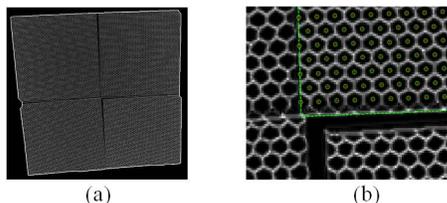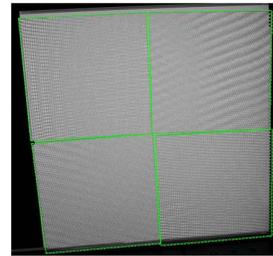Figure 9. Re-projection of the VLA boundaries back onto a camera captured image for the TLA-IID. (See the supplementary material for magnified views.)

the result is shown in figure 8(a). Figure 8(b) shows a local region near the screen center, where the lens array structure and misalignments can be well seen. Since the VLA geometry is known, a homography transform of a VLA can be found by image feature fitting, which provides the POR pixels as the lens image centers. The cut lenses at the boundary are excluded in the computation since their image features are not stable. The POR pixels extracted from one of the VLAs is superimposed with figure 8(b).

Then based on the method in section 3.2, the pose of each VLA is computed. Figure 9 shows the re-projection of the four VLA boundaries on an image captured at the same pose with the structured light capture. Here an all white image is displayed on the effective LCD area for the TLA-IID. Environment lights are turned off to make the transparent lens arrays as visible as possible. From the dim Morie patterns (due to the sub-pixel and black matrix structure of the LCD), the extracted VLA boundaries can be seen to be well aligned with the real ones. It is noticed that some outer VLA boundaries are out of the LCD's effective area. This is actually also consistent with the reality since the VLA plane is above the LCD with a distance (the lens's focal length) and the TLA-IID is viewed from an oblique angle.

5.2. Precision evaluation of the calibration result

Precision of the calibration method is evaluated with a cross verification approach. The calibration procedure is

Figure 7. Structured light pattern decoding results.

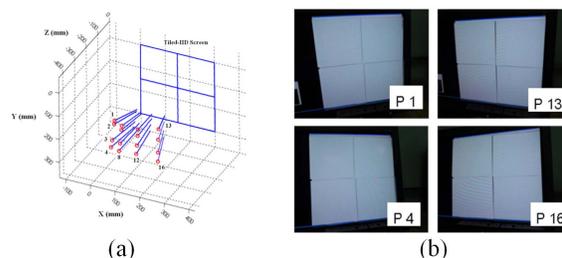Figure 8. Results of lens boundary and POR pixel extraction.

Figure 10. The camera is placed to 16 poses for precision evaluation. In (a), the red dots show the camera positions. The blue lines the camera view directions. (b) shows camera images for the TLA-IID at four of the 16 poses.

Items	Parameters	Mean	Std.
VLA 1	$a^{(1)}$	-0.0074 deg.	0.0157 deg.
	$t_x^{(1)}$	0.2202 mm	0.0386 mm
	$t_y^{(1)}$	5.8516 mm	0.0334 mm
VLA 2	$a^{(2)}$	1.1575 deg.	0.0094 deg.
	$t_x^{(2)}$	151.1391 mm	0.0275 mm
	$t_y^{(2)}$	5.8555 mm	0.0427 mm
VLA 3	$a^{(3)}$	-1.1377 deg.	0.0121 deg.
	$t_x^{(3)}$	1.1734 mm	0.0443 mm
	$t_y^{(3)}$	152.6087 mm	0.0487 mm
VLA 4	$a^{(4)}$	0.0328 deg.	0.0183 deg.
	$t_x^{(4)}$	151.7821 mm	0.0383 mm
	$t_y^{(4)}$	152.5969 mm	0.0479 mm

Table 1. Calibration results at different camera poses.

tested repeatedly for a number of times, for each of which the camera is placed to a different pose within the viewing zone. Ideally, the calibration results in these tests are supposed to be the same since the TLA-IID hardware configuration does not change. Therefore, the deviation of the calibration results in different tests reflects the calibration precision.

In our test, the camera is placed to 16 different poses as shown in figure 10(a), at each of which the calibration procedure is performed. In figure 10(a), the red dots show the camera positions and the blue lines attached to them indicate their principle view directions. In general, the camera is pointed towards the center of the TLA-IID. Here, we set the camera image resolution to 2144 pixels by 1424 pixels.

Table 1 lists the calibration results for the four VLAs at the 16 camera poses. Among the VLAs, the maximum standard deviation of angle estimation is 0.018 degrees. The maximum standard deviation of translation estimation is 48.7 um, which equals to 0.4 pixels when measured against the pixel pitch of the used display (0.1245 mm).

Here we also test the calibration by adding pixel position offsets to the POR pixels along a selected direction. The offsets increase from zero to the radius of a lens. Again, the standard deviation is used to evaluate the estimation precision. The test results are shown in figure 11. Figure 11(a)

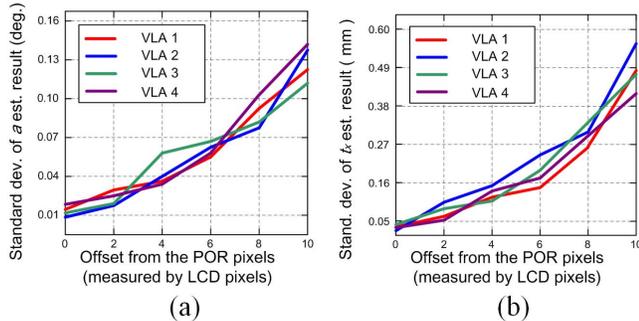(a)   (b)

Figure 11. Calibration precision test result when the image features deviate from the POR pixels.

shows the standard deviation of the angle estimation result with respect to an offset (in pixels of the LCD) from the POR pixels. It can be seen that, the precision decreases as the offsets increase in all the four VLAs. Figure 11(b) shows the test for translation estimation for t_x. Also, the estimation precision decreases with the features' deviation from the POR pixels. We also tested the results for estimating t_y, whose result is similar. This test shows that using POR pixels to estimate the VLA geometry plays an important role to the precision of calibration results.

5.3. Effects to remove 3D image artifacts

Finally, the calibration precision is examined by its effects for removing 3D artifacts.

First, we created a 3D scene with two well-defined planar geometry patterns. As shown in figure 12(a), the scene contains a front depth layer, which is 15 cm before the screen. On the front depth layer, a set of concentric circles are defined. Behind the screen about 10 cm is a rear depth plane, on which horizontal and vertical lines form a grid equally spaced. The two planes span a 3D volume with significant geometric texture patterns, whose distortions are clear to observe for testing the calibration effects. Figure 12(b) shows a 3D image without calibration, for which the integral image is created without considering the lens misalignment. It can be seen that the 3D image is severely distorted.

Figure 13(a) shows a 3D image after calibration. The subfigures 13(b) and 13(c) show local image regions. It can be seen that the misalignments between the lens arrays still exist. However, the 3D image distortions are corrected and the depth patterns are shown correctly. In figure 14, a Coca-Cola bottle was placed as a reference object in front of the screen with its cap aligned with the bottom rim of the green

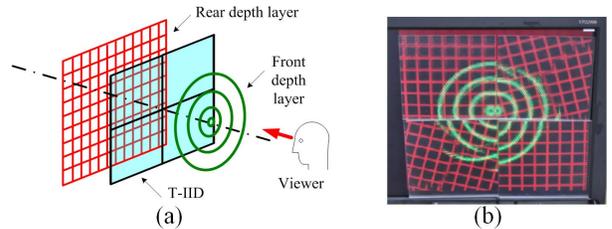
(a)   (b)

Figure 12. (a) A 3D test scene created to evaluation calibration effect. (b) 3D image of scene (a) without calibration. (Video in supplementary material.)

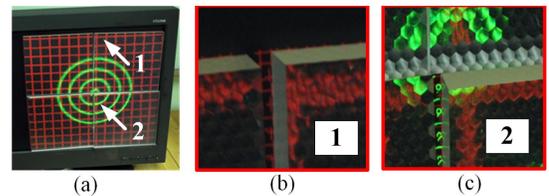(a)   (b)   (c)

Figure 13. After calibration, the TLA-IID shows a correct 3D image (a) though the hardware error in (b) and (c) still exist.
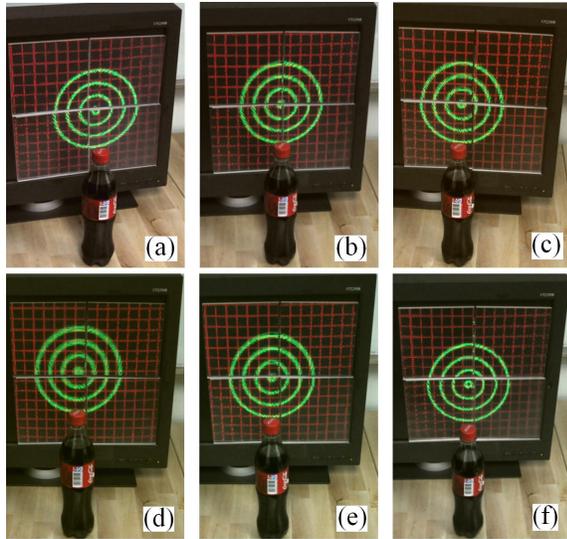
Figure 14. 3D images viewed from different angles after calibration. Subfigures (a), (b), and (c) show horizontal view variations. Subfigures (d), (e), and (f) show vertical view variations. (See video in supplementary materials.)
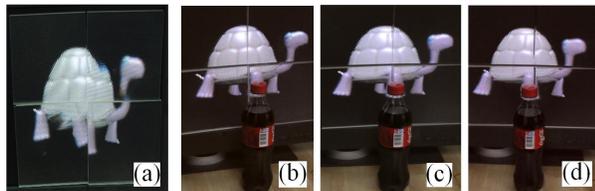

Figure 15. Experiment results with a CG tortoise model. Subfigure (a) shows a 3D image before calibration. Subfigures (b),(c), and (d) show 3D images corrected after calibration. (See video in supplementary materials.)

rings in 3D space. Note that under different view angles, the image of the rings move with the cap of the Cola bottle as if a real 3D object were placed on top of it. In each of the view angles, the rings and line grid shows correct geometry without distortions.

Also we show the experimental results with a CG generated tortoise model in figure 15. Here the bottle cap is aligned in 3D to the right front foot of the tortoise. This shows that the images are successfully corrected in these different views simultaneously.

In practice, there are some cases where warps appear in lens arrays when their fabrication quality is limited. These make the lens arrays not perfectly parallel to the LCD and introduce errors. Addressing such issues would require additional theory and test, which we leave for a future work.

6. Conclusions

This work shows that 3D image artifacts in a TLA-IID due to lens array misalignments can be effectively removed by calibration. By properly formulating the POR geometry and using structured light based methods, the proposed calibration can be performed with only one camera without any additional markers in a highly automatic manner. This provides an efficient software approach to compensate the hardware errors and leads to a cost-effective solution for building a large screen 3D integral imaging display.

7. Acknowledgements

References

[1] D. G. Aliaga, Y. H. Yeung, A. Law, B. Sajadi, and A. Majumder. Fast high-resolution appearance editing using superimposed projections. *ACM Transactions on Graphics*, 31(2):1–12, 2012.

[2] J. Bouguet. Camera calibration toolbox for matlab. http://www.vision.caltech.edu/bouguetj.

[3] H. Chen, R. Sukthankar, G. Wallace, and K. Li. Scalable alignment of large-format multi-projector displays using camera homography trees. *In proceedings of IEEE Visualization*, 2002.

[4] J. Ge, D. Sandin, T. Peterka, T. Margolis, and T. DeFanti. Camera based automatic calibration for the Varrier system. *In proceedings of IEEE CVPR*, 2005.

[5] Y. Kim, K. Hong, and B. Lee. Recent researches based on integral imaging display method. *3D Research*, 1(1):17–27, 2010.

[6] R. Kooima, A. Prudhomme, J. Schulze, D. Sandin, and T. DeFanti. A multi-viewer tiled autostereoscopic virtual reality display. *In proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology*, 2010.

[7] B. Lee, J.-H. Park, and S.-W. Min. Three-dimensional display and information processing based on integral imaging, 2006. T.-C. Poon, ed.(Springer, New York, USA).

[8] G. Lippmann. La photographie integrale. *Academy of Sciences, Paris, Comptes-Rendus*, 146(1):446–451, 1908.

[9] J.-H. Park, K. Hong, and B. Lee. Recent progress in three-dimensional information processing based on integral imaging. *Applied Optics*, 48(1):77–94, 2009.

[10] B. Sajadi and A. Majumder. Auto-calibration of multi-projector CAVE-like immersive environments. *IEEE Transactions on Visualization and Computer Graphics*, 8(3):381–393, 2012.

[11] J. Salvi, S. Fernandez, T. Pribanic, and X. Llado. A state of the art in structured light patterns for surface profilometry. *Pattern Recognition*, 43(8):2666–2680, 2010.

[12] D. J. Sandin, T. Margolis, J. G. J. Ge, T. Peterka, and T. A. DeFanti. The Varrier autostereoscopic virtual reality display. *In proceedings of ACM SIGGRAPH*, 2005.

[13] T. Teraguchi, H. Yamashita, K. Masamune, T. Dohi, and H. Liao. Three-dimensional auto-stereoscopic animated image with a long viewing distance using high-precision image correction. *In proceedings of ACM SIGGRAPH*, 2009.

[14] R. Yang, A. Majumder, and M. Brown. Camera-based calibration techniques for seamless multiprojector displays. *IEEE Transactions on Visualization and Computer Graphics*, 11(2):193–206, 2005.

[15] Z. Zhang. A flexible new technique for camera calibration. *IEEE Trans. PAMI*, 22(11):1330–1334, 2000.