Hemispherical confocal imaging using turtleback reflector

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

| As Published | http://dx.doi.org/10.1007/978-3-642-19315-6_26 |
| Publisher | Springer |
| Version | Author’s final manuscript |
| Citable link | http://hdl.handle.net/1721.1/66111 |
| Terms of Use | Creative Commons Attribution-Noncommercial-Share Alike 3.0 |
| Detailed Terms | http://creativecommons.org/licenses/by-nc-sa/3.0/ |
Hemispherical Confocal Imaging using Turtleback Reflector

Yasuhiro Mukaigawa\textsuperscript{1,2}, Seiichi Tagawa\textsuperscript{1}, Jaewon Kim\textsuperscript{2}, Ramesh Raskar\textsuperscript{2}, Yasuyuki Matsushita\textsuperscript{3}, and Yasushi Yagi\textsuperscript{1}

\textsuperscript{1}Osaka University
\textsuperscript{2}MIT Media Lab
\textsuperscript{3}Microsoft Research Asia

Abstract. We propose a new imaging method called \textit{hemispherical confocal imaging} to clearly visualize a particular depth in a 3-D scene. The key optical component is a \textit{turtleback reflector} which is a specially designed polyhedral mirror. By combining the turtleback reflector with a coaxial pair of a camera and a projector, many virtual cameras and projectors are produced on a hemisphere with uniform density to synthesize a hemispherical aperture. In such an optical device, high frequency illumination can be focused at a particular depth in the scene to visualize only the depth with descattering. Then, the observed views are factorized into masking, attenuation, and texture terms to enhance visualization when obstacles are present. Experiments using a prototype system show that only the particular depth is effectively illuminated and hazes by scattering and attenuation can be recovered even when obstacles exist.

1 Introduction

Significant effort has been made to obtain cross-sectional views of a 3-D scene. Real scenes often include obstacles such as scattering materials or opaque occluders. To clearly visualize cross-sectional views as if the scene is cut at a plane, only a particular depth has to be illuminated and haze due to scattering and attenuation should be recovered.

The simplest way to observe a particular depth is to use a large aperture lens. The large aperture makes the DOF (depth of field) shallow, and the region outside the DOF is blurred. The synthetic aperture method \cite{24} mimics a large virtual aperture by combining many small apertures. However, obstacles are still bright and visible, while they are blurred. The confocal imaging \cite{15} simultaneously scans two confocal pinholes over a particular depth. Since both illumination and observation are focused, clear cross-sectional views are obtained. While still visible, obstacles are darkened and blurred. Moreover, scanning requires long measuring time.

Recently, Levoy et al. \cite{11} proposed a new imaging technique which combines synthetic aperture with the confocal imaging. Since this technique is based on light field analysis, only a particular depth can be illuminated without scanning. However, the synthesized aperture size is relatively small because rectangular
mirrors are aligned as a 2D array. Moreover, unwanted effects such as scattering and attenuation still remain.

In this paper, we propose a novel imaging method called hemispherical confocal imaging. To improve the imaging performance of the synthetic aperture confocal imaging [11], we designed turtle reflector which is a polyhedral mirror to approximate a hemispherical aperture with 180 degree of FOV. We introduce focused high frequency illumination using the turtleback reflector with a projector. This method can eliminate scattering on the focused depth, and make unfocused depth almost invisible. Moreover, we introduce factorization of observed views to eliminate attenuation.

Contribution

- By utilizing the new optical device, unfocused depth becomes almost invisible and scattering is eliminated. Moreover, the measurement is very fast, since no scanning is required.
- We have designed the turtleback reflector which is a polyhedral mirror circumscribed in an ellipsoid. The combination of the turtleback reflector, a projector, and a camera can synthesize a hemispherical wide aperture. The optical device can also be used for measuring the complete 8-D reflectance field on the hemisphere.
- A new imaging technique of the focused high frequency illumination is introduced. This technique enables us to separate direct and global components not on the surface but in the 3-D space because any projection pattern can be focused at the particular depth in the 3-D scene.

2 Related work

2.1 Reflectance field measurement

The optical device proposed in this paper can be regarded as an 8-D reflectance field measuring device. A 4-D slice of the 8-D reflectance field under a static illumination can be recorded by scanning a camera [12] or installing multiple cameras [24]. Alternatively, a high-resolution camera is combined with a micro lens array [1], a micro mirror array [22], or masks [25]. To vary illumination,Debevec et al. [4] rotated a light source, and Sen et al. [20] used a projector as a light source. Masselus et al. [13] rotated a projector, and Matusik et al. [14] rotated both a light source and a camera to measure 6-D reflectance field. Müller et al.[17] used 151 cameras with flashes.

In principle, complete 8-D reflectance field can be measured by densely installing many projectors and cameras on a hemisphere. However, it is difficult to realize such a system due to the cost and physical interference between devices. While rotating projector and camera solves these problems, capture process is impractically long. Recently, Garg et al. [6] and Levoy et al. [11] used multiple planar mirrors and Cossairt et al. [2] used a lens array to measure a part of 8-D
reflectance field, but the direction of the illumination and observation is limited to a narrow angle.

On the other hand, only our system can measure complete 8-D LF on the hemisphere covering the scene. Since our system utilizes the geometric property of an ellipsoid, many virtual cameras and projectors can be produced on a hemisphere with uniform density.

2.2 BRDF measurement using mirrors

For measuring bidirectional reflectance distribution function (BRDF), mirrors are often used to replace mechanical motion. Reflected lights can be effectively measured from all directions using a hemispherical mirror [26] or a paraboloidal mirror [3]. Recently, a wide variety of mirrors such as a cylindrical mirror [10], several plane mirrors [9], an ellipsoidal mirror [16], and a combination of a paraboloidal mirror and a specially-designed dome mirror [7] have been used in conjunction with a projector and a camera.

Our turtleback reflector design was inspired by the BRDF measurement using an ellipsoidal mirror [16]. They utilized a geometric property of a rotationally symmetric ellipsoid that all rays from one focal point reflect off the ellipsoidal mirror and reach the other focal point. On the other hand, we utilized a different geometric property of an ellipsoid that the total length from one focal point to the other focal point through any surface points is constant. By utilizing this characteristic, we can produce virtual cameras and projectors at a constant distance from the target just as they are on a hemisphere. Since our purpose is not BRDF measurement but visualization of cross-sectional views, we designed a polyhedral mirror circumscribed in an ellipsoid.

2.3 Descattering

Incident lights to murky liquid or translucent media scatter, and the appearance becomes blurred. To obtain clear views, descattering methods have been developed. Treibitz and Schechner [21] used polarizer under water. Assuming that only single scattering is observed in optically thin media, Narasimhan et al. [18] estimated 3-D shape with descattering and Gu et al. [8] estimated 3-D distribution of inhomogeneous scattering media.

Recently, Fuchs et al. [5] combined confocal imaging with descattering which utilize the fact that scattering components have low frequency. The principle of our hemispherical confocal imaging is similar to this approach. However, we combine two ideas of the focused illumination proposed by Levoy et al. [11] and the high frequency illumination proposed by Nayar et al. [19]. Our approach both shortens the measuring time and eliminates scattering.

3 Hemispherical confocal imaging

Let us assume that a 3-D scene is illuminated by a light source and observed by a camera as shown in Fig.1. Even if the camera is focused on a particular depth
focused depth
obstacles
scattering
occlusion
absorption
unfocused
depth
attenuation

Fig. 1. Illumination and reflection in a 3-D scene. It is difficult to observe a particular depth due to scattering and attenuation.

Table 1. Comparison of several imaging methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>unfocused depth</th>
<th>scanning</th>
<th>scattering</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synthetic aperture</td>
<td>bright</td>
<td>unnecessary</td>
<td>remaining</td>
</tr>
<tr>
<td>Confocal imaging</td>
<td>darken</td>
<td>necessary</td>
<td>remaining</td>
</tr>
<tr>
<td>Synthetic aperture confocal imaging[11]</td>
<td>unilluminated</td>
<td>unnecessary</td>
<td>remaining</td>
</tr>
<tr>
<td>Confocal imaging with descattering [5]</td>
<td>darken</td>
<td>necessary</td>
<td>reduced</td>
</tr>
<tr>
<td>Our hemispherical confocal imaging</td>
<td>unilluminated</td>
<td>unnecessary</td>
<td>reduced</td>
</tr>
</tbody>
</table>

in the scene, the captured image includes reflections from the entire scene. To observe the particular depth, only the depth should be illuminated. This means that both the illumination and observation should have a shallow DOF.

Even if we succeed in illuminating only the particular depth, clear views cannot be observed. The major reasons are scattering and attenuation. The scattering is caused by multi-bounce reflections in translucent media. By the scattering, the views become blurred. On the other hand, the attenuation is caused by occlusion due to obstacles or absorption due to low transparent media. By the attenuation, the illumination becomes nonuniform and the reflections are partially darkened. The following four functions are required to obtain clear views of a particular depth in a 3-D scene.

(a) The DOF should be as shallow as possible.
(b) Only the particular depth should be illuminated.
(c) Scattering should be eliminated.
(d) Attenuation should be eliminated.

To satisfy these requirements, we propose the hemispherical confocal imaging consisting of (1) specially designed turtleback reflector, and (2) focused high frequency illumination.

The turtleback reflector with coaxial camera and projector synthesizes a hemispherical aperture for both illumination and observation to solve (a). The focused high frequency illumination eliminates reflections from unfocused depth and global reflection to solve (b) and (c). Then, we factorized the observed views into masking, attenuation, and texture terms to solve (d). The merits and de-
Hemispherical Confocal Imaging using Turtleback Reflector

Fig. 2. The number of projectors and cameras of several imaging methods which use projector(s) and camera(s) for reflection analysis or reflectance field measurement.

(a) design using single ellipsoid  (b) design using dual ellipsoids

Fig. 3. Principle of the proposed optical device. Virtual projector and cameras are distributed on a hemisphere with uniform density by using the turtleback reflector.

merits and the number of projector and cameras of several imaging methods are summarized in Table 1 and Fig.2, respectively.

4 Optical design of turtleback reflector

4.1 Projectors and cameras on a hemisphere

In our system, projectors are used as light sources to illuminate a particular depth. If the aperture size of the projector is large enough, the projected pattern is focused only within the shallow DOF and blurred elsewhere. Although it is difficult to realize a large aperture using a single lens, the combination of a number of small apertures can easily synthesize a large aperture by the synthetic aperture technique[24].

To mimic an extremely large aperture, many projectors should be placed in every direction at a constant distance and uniform density. That is, ideal
locations of the projectors and cameras are on a hemisphere. If both the projectors and cameras are densely distributed on the hemisphere covering the target scene, the projectors and cameras can synthesize a hemispherical aperture with 180 degree of FOV for both illumination and observation.

4.2 Turtleback reflector for virtual projectors and cameras

It is difficult to place many projectors and cameras due to the cost and physical conflict. Therefore, we distribute many virtual projectors and cameras using mirrors. For this purpose we utilize the geometric property of an ellipsoid that the total length from one focal point to the other focal point through any surface points on a rotationally symmetric ellipsoid is a constant. A target object is placed at the focal point, and a projector and a camera are placed at the other focal point using beam splitter as illustrated in Fig.3(a). Planer mirrors are placed in the tangent planes to the ellipsoid. This is equivalent to hemispherical distribution of many virtual projectors and cameras with uniform density.

Actually, the design using single ellipsoid is difficult to construct because it requires real projector and camera with 180 degree of FOV. Hence, the hemisphere is evenly divided into two parts and two slanted ellipsoids are placed so that they share same focal point as shown in Fig.3(b). In this design using dual ellipsoids, a projector and a camera with a normal FOV can be used.

4.3 Design of optical device

We designed a new turtleback reflector which is a polyhedral mirror circumscribed in an ellipsoid as shown in Fig.4. The virtual projectors and cameras are placed at the nodes of a geodesic dome which is generated by dividing an icosahedron two times. While the optical device is composed of two symmetric parts, we constructed one half as a prototype to confirm the ability.

Figure 5(a) shows the frame to fix the mirror patches made by stereolithography. Fifty mirror patches are attached to the frame. The turtleback reflector is combined with a high-resolution camera (PointGrey, Grass-50S5C, 2448 × 2048) and a projector (KAIREN Projector X Pro920, 640 × 480).
5 Focused high frequency illumination

5.1 Illumination and reflection in a 3-D scene

To analyze the reflectance field in a 3-D scene, we need to know how lights illuminate points in a scene, and how the reflections are observed. We divide the 3-D scene into a set of small voxels. Let $L_k$ be a set of rays which illuminate the $k$-th voxel, and $R(L_k)$ be a set of reflected rays of $L_k$ at the voxel. Since the observed view of the entire scene by a camera is expressed by a sum of the reflected rays from all voxels, the view is presented by $\sum_k R(L_k)$.

Illuminations and reflections can be regarded as a sum of direct and global components [19]. As shown in Fig.6 (a), the illumination of the $k$-th voxel can be decomposed into the direct illumination $L^D_k$ and the global one $L^G_k$. Similarly, the reflection can also be decomposed into the direct reflection $R^D(L_k)$ and global one $R^G(L_k)$. That is,

$$R(L_k) = R^D(L_k) + R^G(L_k) \quad \text{and} \quad L_k = L^D_k + L^G_k.$$  \hspace{1cm} (1)

Hence, the observed view can be modeled as a sum of four components by

$$\sum_k R(L_k) = \sum_k (R^D(L^D_k) + R^G(L^D_k) + R^D(L^G_k) + R^G(L^G_k)).$$  \hspace{1cm} (2)

5.2 Focused illumination by multiple projectors

To obtain clear views of a particular depth in a 3-D scene, only the depth should be illuminated. Moreover, any global illuminations and global reflections should be eliminated to reduce scattering in the media. That is, the first term $R^D(L^D_k)$ in Eq.(2) should be measured separately.
Fig. 6. Focused high frequency illumination. The high frequency patterns are focused only on the particular depth. The projection is blurred out of the DOF.

By using our optical device, such special illumination and measurement can be realized. Since the virtual projectors surrounding the target scene can synthesize a large aperture, the DOF becomes very shallow. We combine the focused illumination technique proposed by Levoy et al. [11] and high frequency illumination technique proposed by Nayar et al. [19], and call the new technique focused high frequency illumination (FHFI for short). It is noted that our optical system enables the efficient combination of [11] and [19]. The former can project arbitrary focused patterns at any plane in the scene. Out of the focused plane, projected patterns become blurred. The latter can separate direct and global components on the 2-D surface. Our FHFI can separate direct and global components in the 3-D volume, because our optical device can synthesize a hemispherical aperture with 180 degree of FOV.

For the FHFI, high frequency checker board patterns are projected from each projector. The position of the white and black pixels are aligned at the depth as shown in Fig. 6 (b). This means that the high frequency illumination is focused only at a particular depth. The voxels in the scene are classified into unfocused voxels $U$, and focused and illuminated voxels $F1$, and focused but unilluminated voxels $F2$ as shown in Fig. 6 (c).

Compared to a white pattern, the average intensity of the high frequency illumination is darken because the half of pixels are black. Table 2 shows the relative intensities of the four reflection components for each voxel type. The global illumination to every voxel decreases by half. The direct illumination to $U$ also decreases by half because the projected patterns are blurred. The $F1$ receives full direct illumination, while the $F2$ receives no direct illumination. By combining these differences, \( \sum_{k \in F1 \cup F2} R^D(L^D_k) \) which presents only direct components from voxels at the focused depth can be separated.

Let $I_P$ be a captured image when voxels of $F1$ are illuminated but voxels of $F2$ are not illuminated. Let $I_N$ be a captured image when the inverse pattern is projected. Then, these images can expressed as

\[
I_P = \sum_{i \in F1} R \left( \frac{L_i^D}{2} + \frac{L_i^G}{2} \right) + \sum_{i \in F2} R \left( \frac{L_i^G}{2} \right) + \sum_{i \in U} R \left( \frac{L_i^D + L_i^G}{2} \right),
\]

\[
I_N = \sum_{i \in F1} R \left( \frac{L_i^G}{2} \right) + \sum_{i \in F2} R \left( \frac{L_i^D}{2} + \frac{L_i^G}{2} \right) + \sum_{i \in U} R \left( \frac{L_i^D + L_i^G}{2} \right).
\]
Table 2. Relative intensities of four reflection components for each voxel type.

<table>
<thead>
<tr>
<th></th>
<th>(R^p(L_i^L))</th>
<th>(R^p(L_i^U))</th>
<th>(R^G(L_i^D))</th>
<th>(R^G(L_i^U))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(U) (unfocused)</td>
<td>1/2</td>
<td>1/2</td>
<td>1/2</td>
<td>1/2</td>
</tr>
<tr>
<td>(F_1) (focused and illuminated)</td>
<td>1</td>
<td>1/2</td>
<td>1</td>
<td>1/2</td>
</tr>
<tr>
<td>(F_2) (focused and unilluminated)</td>
<td>0</td>
<td>1/2</td>
<td>0</td>
<td>1/2</td>
</tr>
</tbody>
</table>

By comparing two intensities at the same position in \(I_p\) and \(I_N\), we can make an image \(I_{\text{max}}\) which has larger intensities and an image \(I_{\text{min}}\) which has smaller intensities. Since global component has only low frequency \([5] [19]\),

\[
\sum_{i \in F_1} R^G(L_i^D) \simeq \sum_{i \in F_2} R^G(L_i^D). \tag{5}
\]

Therefore,

\[
I_{\text{max}} - I_{\text{min}} = \sum_{i \in F_1} R^D(L_i^D) + \sum_{i \in F_2} R^D(L_i^D) \pm (\sum_{i \in F_1} R^G(L_i^D) - \sum_{i \in F_2} R^G(L_i^D))
\]

\[
= \sum_{i \in F_1} R^D(L_i^D) + \sum_{i \in F_2} R^D(L_i^D)
\]

\[
= \sum_{i \in F_1 \cup F_2} R^D(L_i^D). \tag{6}
\]

This means that only the particular depth \((F_1 \cup F_2)\) can be directly illuminated without global illuminations, and only the direct reflections can be measured without global reflections. As shown in Table 1, our method does not illuminate unfocused depth. Since no scanning is necessary, the measurement is fast. Moreover, scattering which is a major global component in translucent media is eliminated.

6 Factorization of the observed views

Although only the focused depth is illuminated and scattering is eliminated by the FHFI, the view of the focused depth is still unclear due to attenuation of the incident and reflective lights. This is due to occlusion and absorption as shown in Fig.1. Occlusion casts sharp shadows because some lights are directly interrupted by obstacles, while attenuation usually makes smooth change because lighting powers are decreased by spatially distributed low transparent media.

Fortunately, the scene is observed by several virtual cameras. Even if some lights are attenuated, other cameras may observe the scene without attenuation. Hence, we try to estimate the texture which is not affected by attenuation based on the observation from multiple cameras. We assume that there are \(K\) virtual cameras and each camera has \(N\) pixels. Let \(O_{ij}\) be the intensity of the \(i\)-th pixel in the \(j\)-th camera. We model that the observed intensities are factorized as

\[
O_{ij} = M_{ij} A_{ij} T_i. \tag{7}
\]
Here, $M_{ij}$ is the *masking term* which has a value of 0 or 1. If the light is occluded by obstacles, the value becomes 0, otherwise it becomes 1. $A_{ij}$ is the *attenuation term* which expresses light attenuation due to absorption. $T_i$ is the *texture term* which expresses the reflectance of the particular depth. It is noted that only the texture term is independent to the viewing direction assuming Lambertian reflection. Figure 7 illustrates this relationship.

The flow of the factorization process is as follows

**STEP-1:** First, the masking term is decided. Since unfocused depths are not illuminated by the FHFI, obstacles can be easily distinguished using a simple threshold. After decision of the masking term, the following processes are done for pixels satisfying $M_{ij} = 1$.

**STEP-2:** The initial attenuation term is decided as $A_{ij} = 1$.

**STEP-3:** The texture term is calculated. Ideally, a unique reflectance should be estimated despite different camera $j$, but the observed intensities vary. This kind of problem is often seen in stereoscopy[23], so we used a median filter in a similar fashion by $T_i = \text{Median}(O_{ij}/A_{ij})$.

**STEP-4:** Update the attenuation term by $A_{ij} = O_{ij}/T_i$ to satisfy Eq.(7).

**STEP-5:** Smooth the attenuation term using a Gaussian function, because attenuation smoothly varies over the 3-D scene. Then go back to **STEP-3** until the texture term does not change.

By this factorization process, the observed views are decomposed to three terms and we can get clear texture of the particular depth without attenuation.

## 7 Experiments

### 7.1 Synthetic aperture using virtual cameras

First, we evaluated the ability of the synthetic aperture using the prototype system. A textured paper is covered by obstacle of yellow dense mesh as shown in Fig.8(a). A white uniform pattern was projected onto the scene. Figure 8(b) shows the captured image by the real camera. This image includes fifty views corresponding to fifty virtual cameras. Since all views are affected by the obstacle, it is difficult to see the texture of the sheet. Figure 8(c) shows the change of...
the appearance when the number of the virtual cameras increases to synthesize a large aperture. Since our optical device can synthesize a half of the hemispherical aperture, the obstacle is completely blurred and the texture becomes clear with increasing the number of virtual cameras.

7.2 Descattering by focused high frequency illumination

Next, we confirmed that the FHFI is effective for descattering in a 3-D volume. We covered some textured papers with a white plastic sheet. The left images of Fig.9 (a)(b)(c) show views when a white uniform pattern was projected. The appearances are blurred due to scattering. Checkered patterns in which white and black are replaced every three pixels are projected from the virtual projectors so that these patterns are aligned on the paper. Totally, eighteen images were captured by shifting the projecting pattern. The right images of Fig.9 (a)(b)(c) show the direct component when the high frequency patterns were projected. We can see that scattering in the 3-D scene is effectively reduced and the appearances become clear. While the descattering effect is not perfect, this is attributed to the low resolution of the virtual projectors in our current prototype system.

7.3 Factorization of the observed views

We confirmed the ability to visualize a particular depth in a 3-D scene by combining the FHFI and the factorization. Figure 10(a) shows the scene that an
orange mesh covers a textured paper and (f) shows all views from the virtual cameras under normal illumination. By simply averaging these views, a synthetic aperture image can be generated as shown in (b). Although the obstacle is blurred, the orange color of the mesh affects the paper.

The mesh becomes dark by the FHFI because it is not illuminated, while the paper is bright as shown in (g). By averaging these views, the dark mesh is blurred and the orange color correctly disappears as shown in (c). However, there are uneven dark regions due to attenuation. The factorization decomposes the observed views (g) into the masking term (h), the attenuation term (i), and the texture term (d). We can see that the attenuation can be reduced especially around the letter of the black ‘A’ and the red ‘D’, since the occlusion due to the mesh is regarded as masks.

---

1 Although there are fifty mirror patches, only forty eight patches were used because two patches were misaligned.
8 Limitations

- The resolution of the virtual projectors and cameras is low because the imaging areas of real projectors and cameras are divided for virtual ones.
- The observable area is narrow because all projectors must illuminate and all cameras must observe a common area. To enlarge the area, a large turtle reflector is necessary and it may be difficult to construct.
- The factorization is basically an ill-posed problem. For example, we can not distinguish two different scenes in which red texture is covered with colorless sheet and white texture is covered with red sheet. Some empirical constraints such as the smoothness of the attenuation are necessary.

9 Conclusion

We propose a new method of the hemispherical confocal imaging. This new imaging technique enables us to observe clear views of a particular depth in a 3-D scene. The originally designed turtleback reflector can divide the imaging areas so that a projector and a camera mimic a number of virtual projectors and cameras surrounding the scene. The combination of the focused high frequency illumination and the factorization can illuminate only the particular depth and eliminate scattering and attenuation. We have constructed a prototype system and confirmed the principles of the hemispherical aperture, descattering, and factorization.

One of our future works is to rebuild a more accurate optical device using a high resolution projector and evaluate the total performance, because the current prototype system only showed the principles separately. To develop some applications which can visualize any cross-sectional views of a translucent object is important. Another future work is to visualize the inside of the human body using infrared light.

Acknowledgement

This work has been supported by a Microsoft Research CORE5 project and Grants-in-Aid for Scientific Researches (21680017 and 21650038).

References