PLENOPTIC SPHERICAL SAMPLING

Luigi Bagnato, Pascal Frossard, Pierre Vandergheynst

Ecole Polytechnique Fédérale de Lausanne (EPFL)
Signal Processing Laboratory (LTS2/LTS4), Lausanne, Switzerland

ABSTRACT

We present a novel plenoptic sampling scheme that permits an efficient representation of the full light ray field in a space limited by a convex closed surface. We show that a convenient way to sample the light ray field around an observer consists in using a discrete set of perspective imagers with overlapping field-of-views that are distributed on a closed convex surface and looking along the normal to the surface. Taking inspiration from the vision system of flying insects, we choose to constrain the cameras on a sphere of finite radius. Building on spectral analysis we propose a sampling scheme that permits to reconstruct the spherical light field without aliasing. We validate our framework through experiments in a synthetic environment and we show that our constructive sampling scheme permits to effectively reconstruct the light field without artifacts.

Index Terms— Plenoptic sampling, Spherical Imaging, omnidirectional imaging, Image based rendering

1. INTRODUCTION

Massive changes in the way we acquire and consume video content has lead in the last decade to the raise of a new paradigm for the representation of 3D scenes, called Image Based Rendering (IBR). In applications such as 3DTV or free-viewpoint TV the user can choose the viewpoint and the orientation of the camera. While reconstructing the full 3D model of a scene is still a technological challenge under realtime constraints, image based solutions appear promising as long as the light field is properly sampled to guarantee a given image quality. The corresponding sampling problem then consists in finding a good sampling pattern of the light field around the observer.

The concept of light fields for the elaboration of scenes information was first devised by Leonardo da Vinci [1]. Further studies conducted by Adelson and Bergen [2] resulted in the definition of the *plenoptic* function as the most complete representation of scene over time. In the last decade several parametrizations have been proposed for the plenoptic function: the light field [3] and the lumigraph [4] are among the most popular ones. In these works a 4D parametrization of the plenoptic function is proposed under the assumption that the scene is confined into a cubic bounding box, or, in other words, each light ray is represented by the intersection with two parallel planes. One of the biggest advantages of such parametrization lies in its simple analytic description and efficient implementation. Furthermore in [5] the authors perform a spectral analysis of the light field signal and show that the support of the signal is bounded by the minimum and maximum depth of the scene. However, as pointed out in [6], the use of the representation based on two planes has a major drawback: the limited visual angle of the representation. Several couples of planes could be used to cover the convex hull of an object, but then artifacts are visible in the reconstruction at the boundaries. The authors in [6] refers to this problem as the *disparity problem*. To solve the problem they propose to use a spherical surface for the parametrization of the light field and choose a sampling scheme based on a polyedra approximation of the spherical surface. While their approach offers a way to sample uniformly the light field, their scheme is difficult to apply in practical situations and no clear indication is provided on the number of samples necessary for a good rendering. Furthermore a good knowledge of the scene geometry is required. A rather different approach, called concentric mosaics [7], consists in capturing the scene with a camera mounted at the end of a rotating level beam. As the beam rotates regular images are acquired. This is a 3D representation of the light field, because the virtual observer can only move inside the circle described by the rotating beam. The major drawback of this representation is the vertical distortion caused by the lack of vertical samples.

Inspired by the efficient visual system of flying insects we propose a new 4D sampling scheme of the light field. The common fly has two faceted eyes composed of several thousands simple sensors called ommatidias [8] that provide plenty of planar overlapping views of the world. Mimicking the faceted eye concept, a spherical light field camera can be realized by layering perspective CMOS image sensors over the surface of a spherical structure. One example of such a camera has been recently introduced in [9], where the authors show that the implementation of such a system is feasible with current technologies. Since the cameras are placed on a spherical surface and point outwards, they can capture the light rays coming from the convex hull of the scene around, similarly to concentric mosaic or panoramic video systems.

The sphere is a closed convex surface with constant curvature, so it permits a complete angular coverage while keeping a simple topology. One direct consequence is that it can be efficiently parametrized in spherical coordinates, using only the zenithal and azimuthal angle. Assuming that the air is transparent and there is no radial fall-off of the light intensity, we propose a 4D representation of the light field based on a two-sphere parametrization of the plenoptic function. We also develop a spectral analysis of the captured light field assuming that the scene is at constant depth from the center of the sphere and propose a constructive sampling strategy that guarantees the absence of aliasing during the reconstruction of the visual scene. We finally validate our framework with experiments in a synthetic environment.

The paper is organized as follows: In Section 2 we describe the proposed light field parametrization and the scene modeling. In Section 3 we find an optimal sampling scheme based on the spectral analysis of the plenoptic function. In Section 4 we test the proposed scheme in a synthetic environment.

This work has been partially supported by the Swiss National Science Foundation under grant 200021-125651.

2. LIGHT FIELD PARAMETRIZATION

The plenoptic function $\mathcal{L}(\mathbf{x}, \omega, t, \lambda)$, is the function that represents the intensity (or radiance) that a perfect observer inside a given 3-D scene would record at any position $\mathbf{x} \in \mathbb{R}^3$, direction $\omega \in S^2$, assuming no attenuation of light intensity in free space, time t and wavelength λ. In the rest of the paper, for simplicity of notation and without loss of generality, we will drop the temporal variable t and the variable λ, assuming that we acquire a grayscale image. Let us consider a convex surface layered with a discrete set of pinhole cameras. Assuming an ideal behavior of the cameras, i.e., they sample correctly the light information passing through their focal point, it is not difficult to show that if the convex surface contains a sufficient number of pinhole cameras, the system is an ideal plenoptic sampler, meaning that we can reconstruct the plenoptic function anywhere in the space delimitated by the surface. In the rest of the paper we assume that the convex surface is a sphere. We assume to have N pinhole cameras distributed around the surface of a sphere of radius R in positions $\mathbf{x}_i$. Each camera i can capture light in all the directions ω up to a maximum angle of α_i with the surface normal in $\mathbf{x}_i$. The angle α_i is a parameter of the problem and must be in the range $\alpha_i \in [0, \pi/2]$. Such a choice for the field-of-view (FOV) of the cameras, together with the assumption of the convexity of the imaging surface, exclude the possibility of self-visual-occlusions. The full set of directions that each pinhole camera is able to acquire are thus contained in the positive half space defined by the tangent plane of the surface in the point $\mathbf{x}_i$. Since we are working on the sphere a natural choice is to work in spherical coordinates. A point on the spherical camera surface is then parametrized as $\mathbf{x} = (\phi_\mathbf{x}, \theta_\mathbf{x}, R)$, while a direction by $\omega = (\phi_\omega, \theta_\omega)$ as shown in Fig. 1. In the following we will assume that the coordinates are with respect to a common reference system.

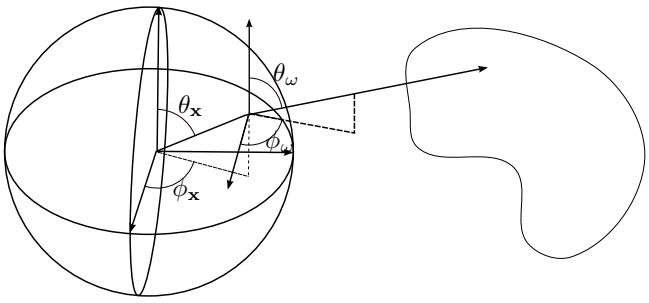

Fig. 1. Schematic representation of the chosen parametrization of the plenoptic function.

3. SPECTRAL ANALYSIS

We propose now a spectral analysis of the plenoptic function parametrized as described in Section 2. We start from [5] but we extend its results to a more general sampling scheme. In [5] the authors assume that the cameras are placed on a plane, so the sampling of plenoptic function is only extended to half of the full 3D space. We generalize their results by using a spherical camera surface. The reader can easily see that a plane can be seen as a degenerated sphere with infinite curvature.

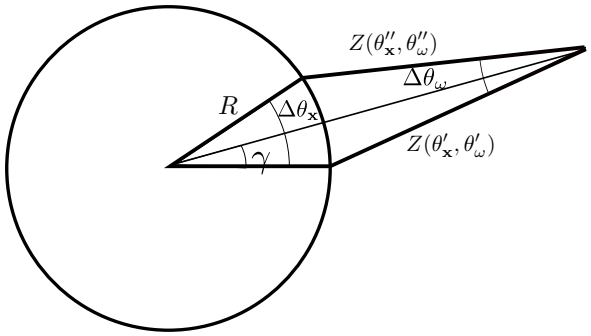

Fig. 2. Geometrical relationship between angular disparities.

The spectrum of the captured light field is given by the spherical harmonics expansion with coefficients L given by:

$$L(m_\mathbf{x}, l_\mathbf{x}, m_\omega, l_\omega) = \int_{S^2} d\Omega(\phi_\mathbf{x}, \theta_\mathbf{x}) Y_{l_\mathbf{x} m_\mathbf{x}}(\phi_\mathbf{x}, \theta_\mathbf{x}) \int_{S^2} d\Omega(\phi_\omega, \theta_\omega) \mathcal{L}(\phi_\mathbf{x}, \theta_\mathbf{x}, \phi_\omega, \theta_\omega) Y_{l_\omega m_\omega}(\phi_\omega, \theta_\omega) \tag{1}$$

The spectrum on the sphere is discrete because it is defined on a closed domain. If we discretize the sphere using an equiangular grid, we can use the theorem from Driscoll and Healy [10]. It tells us that, if the function is bandlimited at $(L_\omega, L_\mathbf{x})$ then we can calculate exactly the spherical harmonic coefficients using the available light field samples as long as their number is sufficiently high. The expression in Eq. 1 is very complex to use directly for deriving good sampling strategies. We will therefore first try to derive some geometrical relationships to estimate some bounds depending on the depth of the scene.

With reference to Fig. 2 we derive the following geometric relationship on a $\phi = constant$ plane:

$$\Delta\theta'_\omega = \sin^{-1}\left(\frac{R}{Z(\theta'_\mathbf{x}, \theta'_\omega)} \sin\gamma\right) \tag{2}$$

$$\Delta\theta''_\omega = \sin^{-1}\left(\frac{R}{Z(\theta''_\mathbf{x}, \theta''_\omega)} \sin\left(\Delta\theta_\mathbf{x} - \gamma\right)\right) \tag{3}$$

$$\Delta\theta_\omega = \Delta\theta'_\omega + \Delta\theta''_\omega = \sin^{-1}\left(\frac{R}{Z(\theta'_\mathbf{x}, \theta'_\omega)} \sin\gamma\right) + \sin^{-1}\left(\frac{R}{Z(\theta''_\mathbf{x}, \theta''_\omega)} \sin\left(\Delta\theta_\mathbf{x} - \gamma\right)\right). \tag{4}$$

This relates the depth of the scene $Z(\theta_\mathbf{x}, \theta_\omega)$ to the angular disparity $\Delta\theta_\omega$ and $\Delta\theta_\mathbf{x}$. In the equation the auxiliary variable γ is the direction of the light ray coming straight from the scene to the center of the sphere. Similar relationships are more complicated in the $\theta = constant$ plane since we have to consider the dependency on $\theta_\mathbf{x}$ and θ_ω:

$$\Delta\phi_\omega = \Delta\phi'_\omega + \Delta\phi''_\omega = \sin^{-1}\left(\frac{R\sin\theta_\mathbf{x}}{Z(\phi'_\mathbf{x}, \phi'_\omega)\sin\theta_\omega} \sin\gamma\right) + \sin^{-1}\left(\frac{R\sin\theta_\mathbf{x}}{Z(\phi''_\mathbf{x}, \phi''_\omega)\sin\theta_\omega} \sin\left(\Delta\phi_\mathbf{x} - \gamma\right)\right). \tag{5}$$

This equation relates the depth of the scene $Z(\phi_\mathbf{x}, \phi_\omega)$ to the angular disparity $\Delta\phi_\omega$ and $\Delta\phi_\mathbf{x}$.

The equations (2) to (5) show a non linear relationship between the disparities, the radius of the spherical camera surface and the

distance to the scene. The use of these geometrical relationships in the analysis of the optimal plenoptic sampling is not straightforward, since they do not simplify the expression of the integral in Eq. (1).

To linearize equation Eq. (5) we assume that $Z(\phi''_{\mathbf{x}}, \phi''_{\omega}) \simeq Z(\gamma,\gamma)$ and $Z(\phi'_{\mathbf{x}}, \phi'_{\omega}) \simeq Z(\gamma,\gamma)$ which is an approximation with an error of 2% if $Z > 3R$ or, equivalently, if the field of view of each camera is small. The expression then simplify into:

$$\Delta\phi_{\omega} = \frac{R\sin\theta_{\mathbf{x}}}{Z(\gamma,\gamma)\sin\theta_{\omega}}\Delta\phi_{\mathbf{x}} \tag{6}$$

Similarly, for Eq. (2) we will have:

$$\Delta\theta_{\omega} = \frac{R}{Z(\gamma,\gamma)}\Delta\theta_{\mathbf{x}} \tag{7}$$

We now derive the spectrum support for a simple scene at constant depth from the camera surface, i.e. we take $Z(\gamma,\gamma) = z_0$.

On a spherical domain the best representation of the light field would be obtained using spherical harmonics with the coefficient expressed in Eq. (1). Since we already made the assumption of small FOV we can, in fact, use the classic Fourier analysis where the complex exponentials will be a fairly good basis for the light field. In other words, since we do not integrate over all the sphere, due to the FOV assumption, we can locally approximate the sphere with its tangent plane. We can further notice that, if we perform the following bijective mapping:

$$s_{\omega}(\phi_{\omega},\theta_{\omega}) = \phi_{\omega}\sin\theta_{\omega} \tag{8}$$

$$s_{\mathbf{x}}(\phi_{\mathbf{x}},\theta_{\mathbf{x}}) = \phi_{\mathbf{x}}\sin\theta_{\mathbf{x}} \tag{9}$$

which describe the arc lengths s_{ω} and s_{ω} on the sphere, then Eq. (6) simplify as $\Delta s_{\omega} = \frac{R}{Z(\gamma,\gamma)}\Delta s_{\phi}$. In practice the previous mapping introduces the elements of the metric tensor on the sphere, $\sin\theta_{\omega}$ and $\sin\theta_{\mathbf{x}}$, inside the parametrization such that for small variations of θ and ϕ, Δs and $\Delta\theta$ represent the coordinates of an orthonormal basis on the tangent plane.

Let us assume that the scene is lambertian, i.e., it reflects light isotropically. If we take as reference camera the one in ($\phi_{\mathbf{x}} = 0, \theta_{\mathbf{x}} = \bar{\theta}_{\mathbf{x}}$) and we use the lambertian property of the scene we can write $\mathcal{L}_s(s_{\mathbf{x}},\theta_{\mathbf{x}},s_{\omega},\theta_{\omega}) = \mathcal{L}_s(0,\bar{\theta}_{\mathbf{x}},s_{\omega}+\frac{R}{z_0}s_{\mathbf{x}},\theta_{\omega}+\frac{R}{z_0}(\bar{\theta}_{\mathbf{x}}-\theta_{\mathbf{x}}))$, where $\mathcal{L}_s(s_{\mathbf{x}},\theta_{\mathbf{x}},s_{\omega},\theta_{\omega}) = \mathcal{L}(\phi_{\mathbf{x}},\theta_{\mathbf{x}},\phi_{\omega},\theta_{\omega})$. The Fourier transform of the light field generated by a scene at constant depth is:

$$\begin{aligned}
L(\Omega_{s_{\mathbf{x}}},\Omega_{\theta_{\mathbf{x}}},\Omega_{s_{\omega}},\Omega_{\theta_{\omega}}) = \int d\theta_{\omega} e^{-j\Omega_{\theta_{\omega}}\theta_{\omega}} \int ds_{\omega} e^{-j\Omega_{s_{\omega}}s_{\omega}} \\
\int ds_{\mathbf{x}} \int d\theta_{\mathbf{x}} \mathcal{L}_s(s_{\mathbf{x}},\theta_{\mathbf{x}},s_{\omega},\theta_{\omega}) e^{-j\Omega_{s_{\mathbf{x}}}s_{\mathbf{x}}} e^{-j\Omega_{\theta_{\mathbf{x}}}\theta_{\mathbf{x}}} = \\
\int d\theta_{\omega} e^{-j\Omega_{\theta_{\omega}}\theta_{\omega}} \int ds_{\omega} e^{-j\Omega_{s_{\omega}}s_{\omega}} L_s(0,\bar{\theta}_{\mathbf{x}},s_{\omega},\theta_{\omega}) e^{-j\frac{R}{z_0}\bar{\theta}_{\mathbf{x}}} \\
\int ds_{\mathbf{x}} e^{-j(\Omega_{s_{\mathbf{x}}}+\frac{R}{z_0}\Omega_{s_{\omega}})s_{\mathbf{x}}} \int d\theta_{\mathbf{x}} e^{-j(\Omega_{\theta_{\mathbf{x}}}+\frac{R}{z_0}\Omega_{\theta_{\omega}})\theta_{\mathbf{x}}} = \\
= F\{\mathcal{L}_s(0,\bar{\theta}_{\mathbf{x}},\theta_{\omega},\phi_{\omega})\} e^{-j\frac{R}{z_0}\bar{\theta}_{\mathbf{x}}} \delta(\Omega_{\theta_{\mathbf{x}}} + \frac{R}{z_0}\Omega_{\theta_{\omega}}) \\
\delta(\Omega_{s_{\mathbf{x}}} + \frac{R}{z_0}\Omega_{s_{\omega}})
\end{aligned} \tag{10}$$

The analysis in Eq.(10) neglects the windowing effect due to a non-infinite domain of integration under the assumption (common in the literature [5, 11]) that is negligible. What emerges is that, if we consider a slice of the spectrum on the plane $(\Omega_{s_{\mathbf{x}}}, \Omega_{s_{\omega}})$, we observe that the spectrum is constrained on a family of lines given by $\Omega_{s_{\mathbf{x}}} + \frac{R}{z_0}\Omega_{s_{\omega}} = 0$ (and similarly $\Omega_{\theta_{\mathbf{x}}} + \frac{R}{z_0}\Omega_{\theta_{\omega}} = 0$), whose slope depends on the ratio R/z_0. A similar results was obtained with a different procedure in [11] but only applied to concentric mosaics.

Using the methodology defined in [5], assuming a regular sampling of the light field and that the scene is confined between two spheres at distances z_{min} and z_{max}, we find that to avoid aliasing of the spectral replicas the following sampling condition must be respected for $\theta_{\mathbf{x}}$:

$$\Delta\theta_{\mathbf{x}} \leq \frac{2\Delta\theta_{\omega}}{R(\frac{1}{z_{min}} - \frac{1}{z_{max}})} \tag{11}$$

a similarly for $s_{\mathbf{x}}$:

$$\Delta s_{\mathbf{x}} \leq \frac{2\Delta s_{\omega}}{R(\frac{1}{z_{min}} - \frac{1}{z_{max}})} \tag{12}$$

For the details of the derivations we refer to [5].

4. EXPERIMENTAL RESULTS

In this section we present some experimental results to test the sampling conditions given in Eq. (11) and Eq. (12). We simulate an hemispherical arrangement of cameras, looking outwards along the normal to the surface, inside the 3D modeling environment Blender. Each camera has a planar sensor with a resolution of 360x270 pixels and an horizontal field of view of $\pi/3$. We use a regular spatial sampling $\Delta s_{\mathbf{x}} = \Delta\theta_{\mathbf{x}} = \pi/12$. To distribute the cameras over the hemisphere we start by subdividing the hemisphere into 6 layers identified by the positions $\theta^k_{\mathbf{x}} = \pi/12 * k$ with $k = 1, 2, ..., 6$. In each layer we position an equally distributed number of cameras given by $N^k = \lfloor \frac{2\pi}{\pi/12}\sin\theta^k_{\mathbf{x}} \rfloor$ for a total number of 100 cameras. We place the hemispherical arrangement approximatively in the middle of a living room of size 20x20x3.

To show the aliasing effect due to an insufficient sampling we reconstruct the view of a virtual observer in the center of the hemispherical surface containing the cameras as shown in Fig. 3. For simplicity of visualization the output image is represented on an equiangular grid of size 2048x563, for an equivalent angular resolution of $\pi/1024$ which is similar to the one of the planar cameras. In fact, the image has been calculated on the "irregular" grid, similar to the one we use to position the cameras, and the missing points on the equiangular grid are obtained by interpolation for display purpose only. In other words, the proposed scheme permits to save roughly 30% of computational time by reducing the number of samples contained into an equiangular grid.

It is out of the scope of the paper to address the reconstruction algorithm used to interpolate the light field, and we plan to address this topic in future works. For the purpose of this analysis we can say that the images are reconstructed in a two step process: first each camera image is focused at a given distance and remapped on the output irregular grid. Once the images are remapped on a common reference system and easy to access the output image is obtained by filtering. In this paper we used a large aperture filter to render the images.

In Fig. 3 we show two different renderings obtained by varying the radius R of the hemisphere. In both cases the system is focused on the close chair which is visible on the bottom in the right lower third, which is at a distance of about 5. In the top image we used a radius $R = 0.2$ which respect the sampling conditions defined in Eq. (11) and Eq. (12), while in the bottom image we used a bigger radius $R = 0.5$ and the conditions are not respected anymore. We can

Fig. 3. Reconstruction results with focus plane on the closest object, the chair on the bottom right of the image. Top: the sampling conditions are respected. Bottom: the sampling conditions are not respected and artifacts in the form of double lines are visible in the image.

clearly see that when we use $R = 0.5$ visually disturbing aliasing effects appear in the image in the form of double lines.

5. CONCLUSIONS

We propose a novel spherical 4D light field sampling scheme. We choose a two sphere parametrization of the plenoptic function where each light ray is described by the intersection with two concentric spheres. Using spherical coordinates and small FOV approximations we perform a spectral analysis of the light field assuming that the scene is at a given distance from the center of the sphere. The proposed sampling scheme is tested in a synthetic environment and the results show that the reconstructions are free from aliasing if the proposed sampling conditions are respected.

6. REFERENCES

[1] Leonardo and I. Richte, *The Notebooks of Leonardo Da Vinci*, Oxford University Press, 1980.

[2] Edward H. Adelson and James R. Bergen, "The plenoptic function and the elements of early vision," in *Computational Models of Visual Processing*. 1991, pp. 3–20, MIT Press.

[3] M. Levoy and P. Hanrahan, "Light Field Rendering," in *SIGGRAPH '96, Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*. ACM, 1996, pp. 31–42.

[4] S.J. Gortler, R. Grzeszczuk, R. Szeliski, and M.F. Cohen, "The Lumigraph," in *SIGGRAPH '96 , Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*. ACM, 1996, pp. 43–54.

[5] Jin X. Chai, Xin Tong, Shing C. Chan, and Heung Y. Shum, "Plenoptic sampling," in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, New York, NY, USA, 2000, SIGGRAPH '00, pp. 307–318, ACM Press/Addison-Wesley Publishing Co.

[6] Emilio Camahort, Apostolos Lerios, and Don Fussell, "Uniformly sampled light fields," Tech. Rep., Austin, TX, USA, 1998.

[7] Heung Y. Shum and Li W. He, "Rendering with concentric mosaics," in *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, New York, NY, USA, 1999, SIGGRAPH '99, pp. 299–306, ACM Press/Addison-Wesley Publishing Co.

[8] R. Zbikowski, "Fly like a fly [micro-air vehicle]," *Spectrum, IEEE*, vol. 42, no. 11, pp. 46 – 51, 2005.

[9] Hossein Afshari, Laurent Jacques, Luigi Bagnato, Alexandre Schmid, P. Vandergheynst, and Y. Leblebici, "Hardware implementation of an omnidirectional camerawith real-time 3D imaging capability," in *3DTV Conference, 2011*, May 2011, pp. 1–4.

[10] J. R. Driscoll and D. M. Healy, "Computing fourier transforms and convolutions on the 2-Sphere," *Advances in Applied Mathematics*, vol. 15, no. 2, pp. 202–250, June 1994.

[11] Cha Zhang and Tsuhan Chen, "Spectral analysis for sampling image-based rendering data," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 13, no. 11, pp. 1038–1050, Nov. 2003.