A FREQUENCY-BASED APPROACH FOR EFFICIENT PLENOPTIC SAMPLING

Philippe Lambert, Jean-Daniel Deschênes and Patrick Hébert

Computer Vision and Systems Laboratory
Laval University
{plambert,desche07,hebert}@gel.ulaval.ca

ABSTRACT

In image-based light field rendering, many efforts have been made to improve the efficiency of the modeling. This efficiency pertains to the photorealism and compression of a given model. We propose a new way to improve these two aspects. We show that it is better to sample the plenoptic function on a surface that minimizes the frequency content of all its lumispheres. We also demonstrate an additional constraint based on the visual hull that further guides the sampling of the plenoptic function around an object. We propose a corresponding algorithm that relies on images alone and only suppose that the modeled object is opaque. This algorithm is then validated on both real and synthetic data sets.

1. INTRODUCTION

The plenoptic function [1] describes all photons that pass through each point in every direction inside a volume of interest (their intensity, polarization and wavelength at any time). The purpose of modeling the plenoptic function is to allow the photorealistic rendering of a scene from any viewpoint. It distinguishes itself from the use of a 3D model, as it is independent of the scene's geometry. It also enables the reproduction of a complex reflectance or transmittance function. Since we will not be concerned by polarization nor by time variation, we will refer to the plenoptic function as defining only the intensity and wavelength of photons. Hence, the plenoptic function can be measured at any given point by aiming cameras at this point and recording the intensity of colors from several directions.

Assuming air is a transparent medium, the knowledge of the plenoptic function on any closed surface enclosing an object allows the photorealistic rendering of any view of this object from outside the enclosed volume [2]. Thus, modeling the plenoptic function on a surface rather than in a volume considerably reduces the amount of data required to render the object. In theory, one could choose any such surface to model the plenoptic function. For instance, both the Light field Rendering [2] and Lumigraph [3] approaches choose to use a cube since it is straightforward to parameterize. Nevertheless, using such an arbitrary surface requires a greater amount of photographs to achieve the same rendering quality than when using geometric information about the object. The difference in the number of photographs can reach an order of magnitude [4, 5]. This result suggests exploiting range data in conjunction with light field. However, since the light field does not intrinsically require the measurement of the scene geometry, we prefer keeping this advantage and rely only on photographs as input data. This suggests using photographs to deduce depth information. Unfortunately, such a process is often ambiguous (as shown in [6]) and suffers from the difficulties inherent to solving for the well-known correspondence problem. To overcome these difficulties, a least-commitment approach leads to the photo hull [7], but we must hypothesize one or several reflectance models that are usually not available.

Ramanathan *et al.* [8] showed that it is possible to refine a geometry model to improve light field compression efficiency by using images alone. They start from the visual hull and then refine its geometry, an idea which we build on. However, they still aim at finding the object's surface. As our first contribution, we instead justify seeking a surface that minimizes the high frequency component of all its plenoptic function samples, namely lumispheres. This surface is not necessarily the object's surface. Moreover, the search for this optimal surface assumes neither a specific geometric model nor any reflectance function. When sampling the plenoptic function, it is required to account for the visibility of each point. We show that we can approximate this information by using the visual hull. We further contribute to the plenoptic function sampling study. Instead of determining the number or position of photographs to achieve un-aliased rendering such as in [9, 10], we seek the most efficient sampling surface for a given set of viewpoints.

In section 2, we begin by justifying the criterion to seek the optimal sampling surface and address the visibility problem. In section 3, theses ideas are developed to suit a more general object. Finally, experimental results are presented in section 4.

2. WHICH SURFACE TO SAMPLE?

Let $Pv = P(\vec{x},\theta,\varphi)$ stand for the plenoptic function P, defined in a volume V, for any point $\vec{x}$ = (x,y,z) of the domain V, and any direction $\theta \in [-\pi,\pi]$, $\varphi \in [0,\pi/2]$. Let (u,v) be the image coordinates of the projection of $\vec{x}$ in the image plane I of a camera viewing $\vec{x}$ along the direction given by (θ,φ). The intensities of the three primary colors I(u,v) = [R,G,B] recorded by this camera are called a sample value of $P(\vec{x},\theta,\varphi)$. As a graphical representation of Pv, we use a unit sphere $S_{\vec{x}}$ located at any point $\vec{x}$ of V. This sphere parameterizes the directions (θ,φ) of the rays traversing its center and its surface displays the colors viewed along those directions. We refer to this sphere as a lumisphere, a term introduced by Wood et al. [5]. Using this terminology, the knowledge of the plenoptic function on a specific surface amounts to the knowledge of an infinite set of lumispheres located on this surface. We will refer to this surface as the sampling surface. In practical cases however, the number of known rays on each lumisphere and the number of lumispheres are finite since both the number of camera views and their resolution are limited.

2.1 Minimum frequency surface

When rendering an image from a specific viewpoint using the sampled plenoptic function, we must compute the color of the rays that pass through each pixel of the rendered image. We might therefore need to obtain the color of a ray that does not originate from a sampled lumisphere, as shown at the top of figure 1. Since the distance *d* between the sampling surface and the actual object's surface is unknown, it is not possible to determine the actual rays from which to interpolate the rendered ray. One way to cope with this problem is to assume that the closest available lumispheres to the intersection between the rendered ray and the sampling surface, approximate the missing one. Under this assumption, the color of the rendered ray can be obtained by interpolating the colors of the rays parallel to the rendered ray originating from these closest lumispheres. These interpolating rays might again be unknown. We can therefore compute these rays by interpolating their colors from those of neighboring known rays in each lumisphere as shown at the bottom of figure 1.

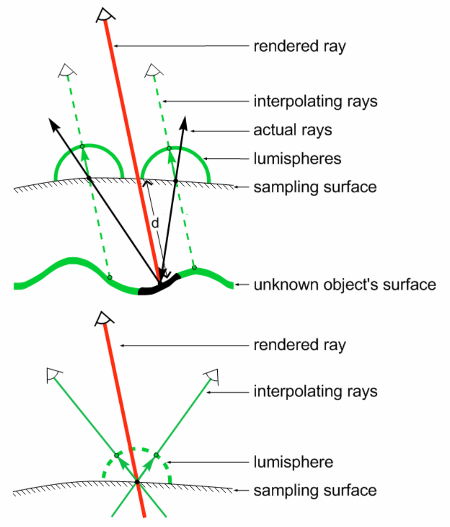
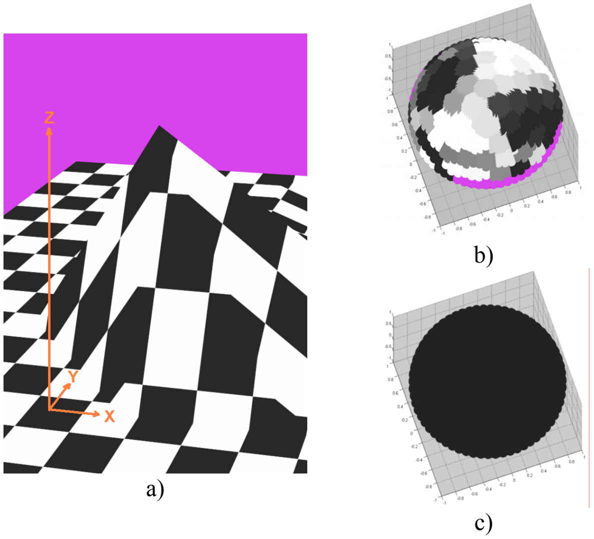


Figure 1: Top - When a rendered ray does not fall on a known lumisphere, its value is interpolated from its closest neighbors in the closest lumispheres available. Bottom - When a ray is rendered from a known lumisphere, we must interpolate its value from its closest available neighbors.

These two interpolation levels illustrate the cause of rendering errors, which thus essentially arise from the reconstruction of the lumispheres' signals.

For a given set of images, the reconstruction of each lumisphere's signal will be aliased if it contains higher frequencies than the limit imposed by the sampling achieved. Figure 2 illustrates that the location of a lumisphere affects its frequency content. For a textured object, a lumisphere far from its surface encompasses higher frequency components since it describes photons originating from a larger sustained solid angle on the object. Thus, since we are free to select any closed surface bounding the object as the sampling surface, we should select one such that its lumispheres present the lower frequency content. We will call this surface, the *minimum frequency surface*. By using the *minimum frequency surface*, the lumispheres' signals will be less aliased and consequently so will be the rendered images.



a)

b)

c)

Figure 2: A lumisphere far from a textured object exhibits higher frequency content. a) Image displaying the object and its reference frame. b) A lumisphere located above the object's surface on the Z-axis. c) A lumisphere located at the level of the object's surface on the Z-axis.

The frequency content $freq(\vec{x})$ of a lumisphere located at $\vec{x}$ can be measured using the following expression:

$$freq(\vec{x}) = \sum_{s}\sum_{t}\left(F(\vec{x},s,t)^2\right) - F(\vec{x},0,0)^2, \qquad (1)$$

where $F(\vec{x},s,t) = \underset{(\theta,\varphi)}{TFd}(P(\vec{x},\theta,\varphi))$ and TFd denotes the Fourier transform. We remove the contribution of the zero frequency term to ensure that any uniform lumisphere will exhibit the same frequency content. Since $P(\vec{x},\theta,\varphi)$ is not uniformly sampled in (θ,φ), it was interpolated in order to be normalized to KxK (with K=64 typically) discretized values of (θ,φ). The selected discretization was fixed empirically. In the case of a graph surface as shown in figure 2, the *minimum* *frequency surface* can be obtained by using the following procedure. For all points (x_i,y_i) on a regular 2-D grid, we search along the Z-axis for the lumisphere located at (x_i,y_i,z_i) that encompasses less frequency content. The heights found will serve as anchor points for the sampling surface sought. It then suffices to build a mesh from the acquired points and feed it to a Light Field Mapping algorithm [4].

The *minimum frequency surface* is not necessarily the object's surface. Actually, in the case of a uniformly colored cavity as shown in figure 3 for instance, the lumispheres along the surface shown will exhibit the same frequency content than that on the actual surface.

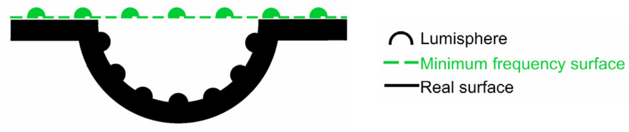


Figure 3: A case where the minimum frequency surface is not the real surface.

2.2 Visual hull constraint and visibility

When the model of a specific object is required, it is desirable to exclude information arising from the background to improve memory efficiency. Conversely, missing any ray coming from the object would create an incomplete model. This observation leads us to another constraint on the surface sought, that we translate into the following theorem.

Theorem 1: Visual hull constraint theorem
A closed surface emits all photons coming from an object and does not emit any photon from the background if and only if it shares the same visual hull as that of the object.

Proof:
First, suppose that the closed surface is missing one photon coming from the object. Then, it is possible to find a viewpoint (for example orthogonally to the photon's direction) such that the apparent silhouette will miss a part of the object's silhouette. Thus, the visual hull of the surface will itself miss a part of the object's visual hull. The surface therefore would not share the same visual hull as the object.
Secondly, suppose that the closed surface emits a photon coming from the background. Then, it is possible to find a view (again, orthogonally to the photon's direction) such that the apparent silhouette will be different from that of the object. The surface therefore would not share the same visual hull as the object.
The converse statement follows from similar arguments. □

Due to self-occlusion, it is also necessary to account for the visibility of a point when sampling the plenoptic function at its location. This exact information could be obtained by measuring the object's geometry. However, since we do not want to rely on this information, we choose instead to use the visual hull as an initial approximation of the real visibility. Thus, we are now interested in searching for the *minimum frequency surface* that follows the *visual hull constraint* and visibility. We will now describe an algorithm that performs this task for a general object.

3. MINIMUM FREQUENCY SURFACE SEARCH

As the *visual hull constraint* theorem and visibility suggest, we start with the visual hull of the object itself. We assume that the number of viewpoints is enough to estimate the object's visual hull. In order to do so, we exploit a regular volumetric grid of lumispheres and remove any lumisphere that exhibits any photon coming from the background. These photons are easily identifiable by using color keying. Once all the points inside the visual hull have been obtained, we use a marching cube algorithm to find a mesh of this visual hull. The distance field D($\vec{x}$) needed to perform this operation is obtained by the following procedure. All voxels inside the visual hull are set to a unique positive distance from the surface while all the voxels outside are set at the same distance but negative. The zero level isosurface is then the desired hull. This surface is used to obtain the visibility clues that are necessary to generate coherent lumispheres. The surface is then refined based on the frequency criterion. Since we must preserve the *visual hull constraint*, we only modify the surface's points that are inside a locally planar region. This is then equivalent to the case of a graph surface described previously. The displacement is made in the direction of the inward normal computed from the isosurface, and is such that for each point it stops at the location where the lumisphere presents the lowest frequency content. We also fixed a minimum gain required to perform the displacement in order to avoid useless modification. If $\vec{q}$ is the modified version of $\vec{q}_0$, a point of a planar region of the isosurface, it is obtained by:

$$\vec{q} = \vec{q}_0 + t \cdot \nabla D \mid freq(\vec{q}) = \min_t(freq(\vec{q}_o + t \cdot \nabla D)), \qquad (2)$$

where $t \in [0, t_{\max}]$, D is the distance field such that ∇D is a vector pointing inside the visual hull. Since any modification to a surface creates new visible points or occludes some of them, it is difficult to maintain a coherent visibility while having a low order computation complexity. We found that keeping the initial visibility computed from an opaque visual hull was a sufficient approximation to the real visibility. More precisely, we assume that a modified point is visible by the same set of cameras that sees its initial position on the visual hull. Once we obtain the modified points, we use the same mesh topology as that of the visual hull mesh and obtain the desired minimum frequency surface.

4. RESULTS

Both synthetic and real image sets were used in the experiments. In both cases, the acquisition consisted in taking several calibrated pictures of an object against a known and fixed background color. We rely on approximately one hundred pictures in both cases. The real set was acquired using a 6.3 megapixel Canon EOS Rebel Digital. The voxel grid was limited to 150^3 voxels. Finally, we use the Light Field Mapping algorithm [4] to generate the light field map from our data sets on the visual hull and on the *minimum frequency surface*.

A synthetic object was selected to demonstrate that the *minimum frequency surface* converges to the real surface when the object is lambertian and textured. It also intends to show a case where traditional passive stereo techniques would have difficulty in solving for the correspondence problem. Renderings are presented in figure 3 and 4. Figure 5 compares the quality and compression of the *minimum frequency surface* to that of the visual hull. The memory compression measure we use is the number of SVD decomposition terms in the light field map.

The volumetric grid discretization is the first apparent distorting factor of our results. Its effects can be seen both on the mesh of figure 4 c) and on the aliased contours of figure 4 e). Moreover, our assumption that the visibility of the *minimum frequency surface* is well approximated by the visibility of the visual hull might not always be valid. In order to cope with this last limitation, we would need to allow the algorithm to iterate by using the previous computed surface as the visibility for the next iteration.

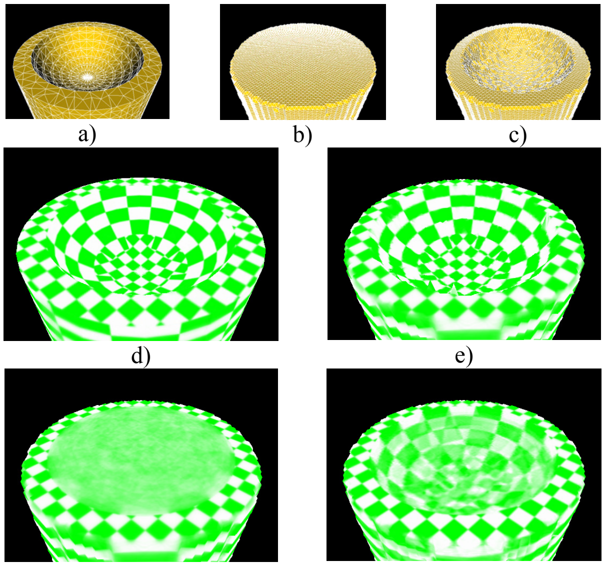


Figure 3: Synthetic data set a) Theoretical mesh b) Visual hull's mesh, c) *Minimum frequency surface's* mesh d) Reference image e) Light field map rendering using the first SVD decomposition term on the *minimum frequency surface*, f) Light field map rendering using the first SVD decomposition term on the visual hull, g) Light field map rendering using the first four SVD decomposition terms on the visual hull.

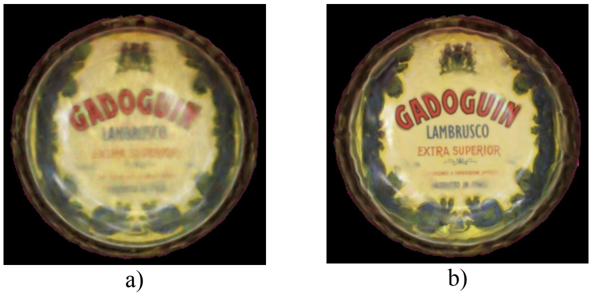


Figure 4: Light field map rendering (3 SVD terms) of real data set a) using the visual hull, b) using the *minimum frequency surface.*

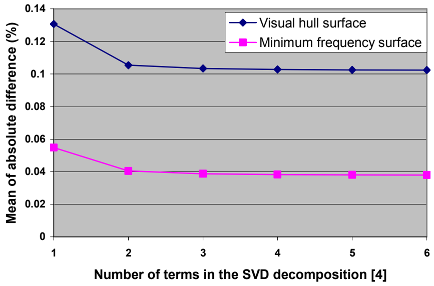


Figure 5: Mean of the absolute difference between pixel values of an actual image and renderings of the same viewpoint.

5. CONCLUSION

The proposed algorithm increases the light field mapping efficiency of visual hull based technique. It also provides new insights into the field of plenoptic sampling. However, it assumes the *minimum frequency surface* shares the same topology and visibility as those of the visual hull. This is not true in some cases. We leave to further work the goal of providing an algorithm that releases this constraint.

6. REFERENCES

[1] E. H. Adelson and J. R. Bergen, "The Plenoptic Function and the Elements of Early Vision", In Computational Models of Visual Processing, Eds. Movshon and Landy, MIT Press, 1991, pp. 3-20,

[2] M. Levoy and P. Hanrahan, "Light Field Rendering," in Computer Graphics (Proc. SIGGRAPH96), Aug. 1996, pp. 31-42.

[3] S. J. Gortler, R. Grzeszczuk, R. Szeliski and M. F. Cohen, "The lumigraph," in Computer Graphics (Proc. SIGGRAPH96), Aug. 1996, pp. 43-54.

[4] W.-C. Chen, J.-Y. Bouguet, M. H. Chu, R. Grzeszczuk, "Light Field Mapping: Efficient Representation and Hardware Rendering of Surface Light Fields." ACM Transactions on Graphics. 21 (3), 2002, pp. 447-456.

[5] D. Wood, D. Azuma, K. Aldinger, B. Curless, T. Duchamp, D. Salesin,and W. Stuetzle, "Surface light fields for 3D photography," in Computer Graphics (Proc. SIGGRAPH00), New Orleans, Aug. 2000, pp. 287-296.

[6] S. Baker, T. Sim, and T. Kanade, "When is the Shape of a Scene Unique Given its Light-Field: A Fundamental Theorem of 3D Vision?," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 25, No. 1, Jan. 2003, pp. 100-109.

[7] K. Kutulakos and S. Seitz, "A theory of shape by space carving," in Proc. Int. Conf. Computer Vision (ICCV99), vol. 1, Sept. 1999, pp. 307-314.

[8] P. Ramanathan, E. Steinbach, P. Eisert,B. Girod, "Geometry refinement for light field compression," Proc. Int. Conf. on Image Processing (ICIP02), Sept. 2002, pp. 225-228

[9] Z. Lin and H.-Y. Shum, "A Geometric Analysis of Light Field Rendering," Int. J. of Computer Vision, Vol. 58, No. 2, Jul. 2004, pp. 121-138.

[10] J.-X. Chai, X. Tong, S.-C. Chan, and H.-Y. Shum, "Plenoptic sampling," in Computer Graphics (Proc. SIGGRAPH00), Aug. 2000, pp. 307-318.