

A Bidirectional Light Field - Hologram Transform

Remo Ziegler¹, Simon Bucheli¹, Lukas Ahrenberg², Marcus Magnor³, Markus Gross^{1 †}

¹ETH Zurich, Switzerland

²Max-Planck-Institute für Informatik, Germany

³TU Braunschweig, Germany

Abstract

In this paper, we propose a novel framework to represent visual information. Extending the notion of conventional image-based rendering, our framework makes joint use of both light fields and holograms as complementary representations. We demonstrate how light fields can be transformed into holograms, and vice versa. By exploiting the advantages of either representation, our proposed dual representation and processing pipeline is able to overcome the limitations inherent to light fields and holograms alone. We show various examples from synthetic and real light fields to digital holograms demonstrating advantages of either representation, such as speckle-free images, ghosting-free images, aliasing-free recording, natural light recording, aperture-dependent effects and real-time rendering which can all be achieved using the same framework. Capturing holograms under white light illumination is one promising application for future work.

Categories and Subject Descriptors (according to ACM CCS): I.3.0 [Computer Graphics]: General

1. Introduction

Throughout computer graphics rendering, geometric (ray) optics is frequently being adopted as a physical model of the image formation process, for some very good reasons: geometric optics is a mathematically simple and yet surprisingly powerful model that is able to explain and also quantitatively describe most optical effects that we can perceive with our eyes. Given all necessary information about a scene, geometric optics is regularly sufficient to achieve fast as well as realistic rendering performance. Nevertheless, geometric optics also possesses a number of limitations. Most prominently, any scene to be rendered must be represented rather inelegantly in terms of 3D geometry, texture, and local reflectance characteristics. Obtaining these separate descriptions of real-world scenes proves tedious, time-consuming, and expensive.

To overcome this drawback, image-based rendering techniques, and specifically light field rendering [LH96] have been proposed. In light field rendering, a (large) set of photographs taken from various different positions all around



the scene are used to represent the visual appearance of the scene. Unfortunately, very large numbers of photos are needed to guarantee aliasing-free light field rendering results [CCST00, IMG00], which is why subsequent image-based rendering techniques again resort to additional (approximate) geometry information to interpolate views from much reduced numbers of photographs.

With holograms, an alternative representation of visual scene appearance is known. Based on wave optics, holography is mathematically considerably more demanding than geometric optics. The need for monochromatic, coherent illumination during acquisition and speckle patterns during display additionally seem to argue against considering holograms in the context of computer graphics rendering. On the other hand, holograms represent visual scene appearance in the most elegant way, containing any possible view from a continuous viewport region without aliasing. In many ways,

[†] {zieglerr.grossm}@inf.ethz.ch, tyberis@gmail.com,
ahrenberg@mpi-inf.mpg.de, magnor@cg.cs.tu-bs.de

holograms are complementary to light fields, see Tab.1. Geometric optics turns out to be simply the approximation of wave optics in the limit of infinitesimally small wavelength [BW59].

In this paper, we propose to use both the light field and the hologram representation of a scene’s visual appearance in tandem. Our goal is to perform processing steps on that respective representation for which the processing step can be done easier, faster, or more accurately. To switch between either representation, we describe how to map from the hologram to the light field representation, and vice versa. These mappings give us the freedom to exploit the advantages of either representation. The advantages and limitations of light fields and holograms are summarized in Table 1.

The core technical contributions of this paper are two functions to transform between holograms and light fields. A key ingredient of the forward transform includes a novel algorithm to reconstruct depth from arbitrary input light fields by exploiting 4D epipolar volume representations. Our mapping functions provide a solid theoretical basis to record full-parallax holograms by means of light field cameras and they enable us to convert any input light field into a hologram for output on future holographic displays. In addition, the wave optics representation of the hologram allows for a variety of sophisticated processing algorithms and the computed depth proxy effectively eliminates ghosting artifacts of the input light fields.

In Sect. 3, we discuss the properties inherent to the light field and the hologram representation. An overview of our framework is presented in Sect. 4, followed by the description of the forward transform from the light field to the hologram in Sect. 5, operations on the hologram in Sect. 6, and the inverse transformation elaborating the essential physical characteristics in Sect. 7. To demonstrate the advantages of our proposed dual light field-hologram representation, we present results for real and synthetic light fields as well as digitally recorded holograms (DRH) in Sect. 9 before we conclude with an outlook on future work.

2. Related Work

In a paper in 1936 [Ger36], Gershun introduced the concept of light fields for the first time. He described it as the amount of light traveling in every direction through every point in space using light vectors. In 1996 Levoy and Hanrahan in [LH96] and Gortler et al. in [GGSC96] presented two similar practical ways of capturing and representing scenes for computer graphics independently, based on Gershun’s theory. Many publications have drawn their work upon the light field as well as the lumigraph representation. Various publications focussing on sampling requirements [CCST00], rendering and filtering [IMG00, SYGM03], reducing ghosting artifacts and aliasing, as well as numerous capturing setups consisting of a camera array [WJV*05, YEBM02] or one shot capturing devices such as in [Ng05, LNA*06] keep exploiting the big potential of this field in various ways.

Table 1: Advantages and Disadvantages

Compare	Hologram	Light Field
Function dimension	2D	4D
Light representation	wave	ray
Single-shot acquisition	Yes	Yes
Refocusing	Yes	Yes
Natural light recording	No	Yes
Speckle free	No	Yes
Real time rendering	No	Yes
Aliasing free	Yes	No
Scene independent sampling	Yes	No
Phase information for depth encoding	Yes	No
Recording without optical elements	Yes	No
Compression	Yes	Yes
Combination with geometrical representations	Yes	No

A good overview of recent work in this field is presented in [Lev06].

Holograms can be computer generated from synthetic data and rendered either on a conventional display as in [ZKG07] or rendered onto holographic displays as presented in [LG95, Mat05]. A real scene can be captured on a holographic film or digitally captured by a CCD camera only if illuminated by monochromatic laser light. This is a severe restriction, since for many scenes the light cannot be controlled in such a meticulous way. DeBitetto presented a two-step model to record holographic stereograms under white light illumination in [DeB]. Halle studied the characterization of sampling-related image artifacts and presented different ways of reducing or eliminating aliasing artifacts of holographic stereograms in [Hal94]. The artifacts originate from using a plane as the depth approximation of the scene.

Numerous publications deal with the problem of depth reconstruction from multi-view input. Many algorithms are based on the Epipolar-Plane Image (EPI) representation or on the related Epipolar Volumes, which were first introduced by Bolles et al. in [BBH87]. Although most of the work assumes Lambertian surfaces, various approaches remove specular effects such as [CKS*05, BN95, LLL*02] while few publications [DYW05, STM06] reconstruct surfaces with almost arbitrary BRDF. However, these methods require additional information about the reflection properties, assume light transport constancy requiring multiple acquisitions under different illumination or are not using the full redundancy of a camera array used to capture a light field.

In our work we present a way of reconstructing depth from light fields with almost arbitrary BRDF. Based on the extracted depth information, a scene captured under white light can be transformed into a hologram featuring full parallax. Since the correct depth gets encoded into the hologram, the images created from the hologram do not show any ghosting artifacts and operations such as refocusing

and varying depth of field are still possible. The highest frequency of the BRDF reconstructed from the hologram will however, not be higher than the one captured by the light field.

3. Representation

Holograms and light fields have been parameterized in numerous ways. In Sect. 3.1 and Sect. 3.2 we describe the specific representations for the hologram and the light field, which are used throughout the paper.

3.1. Parametrization of Light Fields

There exist different parameterizations of light fields. In our paper we will either use the popular two-plane parametrization $LF(u, v, s, t)$ as presented by Levoy [LH96] (see Fig. 1b) or consider the light field as angular parametrization $LF(u, v, \theta, \phi)$ dependent of position on the uv -plane and direction dependent on θ and ϕ as in Fig. 1c.

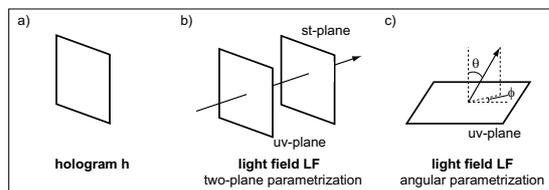


Figure 1: a) depicts the representation of a hologram. b) and c) show two different representations of a light field.

3.2. Parametrization of Holograms

In general, a hologram is a real valued function describing the intensity of the static interference pattern of a complex valued object wave with a complex valued reference wave. The original object wave can be reconstructed from the hologram. In the following we will use the term "hologram" in the spirit of the *wavefield*, which is a complex valued wave function $U(u, v)$, instead of a real valued intensity field. This simplification does not have an influence on the transformation from a hologram to a light field, since the complex valued wave function can be reconstructed from a real valued hologram.

4. Light Field Mapping Pipeline

In this paper we describe a pipeline based on a novel mapping \mathbf{M} and its inverse \mathbf{M}^{-1} giving the possibility to transform a light field into a holographic representation and vice versa. The holographic data representation is similar to a light field in that the hologram as well as the light field measure the light propagated through a plane in space into all directions. The input to our pipeline depicted in Fig. 2 is a pure light field without any depth information. \mathbf{M} (cf. Sect. 5) transforms the light field into a holographic representation. A core ingredient of \mathbf{M} and a core contribution of this paper is a method to extract depth from the input light field (cf. Sect. 5.1). If accurate depth information is available for the light field it can optionally be added to the input

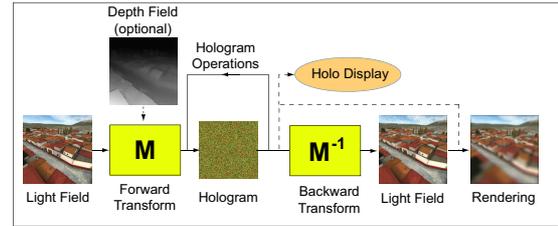


Figure 2: The input to the pipeline is a pure light field, with an option of providing an accurate depth map for every view. Using \mathbf{M} the light field is transformed to a holographic representation, where functions such as compression and progressive transmission can be applied directly on the hologram. A mapping function \mathbf{M}^{-1} allows an inverse transformation into a light field, from which images from different view points can be rendered in real time.

of \mathbf{M} , increasing the quality of the holographic representation as described in [Hal94, CCST00]. Different algorithms can be applied to the manipulation of the hologram, such as compression, progressive transmission, wavefront propagation simulating diffractive effects and others. In Sect. 6 we present a rendering technique, a compression algorithm and study effects of loss of data. Arbitrary parallax images can be rendered efficiently from the holographic representation as long as the COP of the virtual camera lies on the holographic plane. For arbitrary viewpoints we present an inverse mapping \mathbf{M}^{-1} (cf. Sect. 7), transforming the holographic representation back into a light field representation, from which it can be rendered to arbitrary viewpoints in real time.

5. Forward Mapping

The forward mapping \mathbf{M} takes a pure light field and maps it to a hologram. \mathbf{M} consists of two main steps, namely a depth reconstruction from light fields with almost arbitrary BRDF, and hologram evaluation based on the reconstructed depth proxy and the light field if available. An optional accurate depth field can be added to \mathbf{M} , making a depth reconstruction of the light field obsolete, speeding up the mapping and slightly enhancing the accuracy of the forward mapping in case of inaccurate automatic depth reconstruction. Intermediate steps of the forward mapping are shown in Fig. 3.

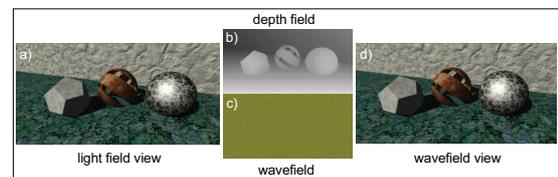


Figure 3: For every input view of the light field a) a depth map b) is reconstructed, which is used to evaluate the wavefield c). A rendering of the generated wavefield is depicted in d).

Since such a depth map is usually not at hand, we present

a novel depth reconstruction method from light field data based on a 4D Epipolar Volume representation.

5.1. Depth Reconstruction from Light Fields

Our method takes advantage of the full light field information and redundancy captured by cameras aligned to a matrix instead of a line. We call the resulting per view depth map the *depth field*. Throughout the paper, depicted depth maps are in fact always disparity maps.

Representation The key advantage of the Epipolar-Plane Image (EPI) representation $EPI(u, s)$ (cf. Fig. 4a) is the collocation of corresponding pixels from different views on one line l_c . In case of Lambertian scenes, such lines are consisting of a constant color in absence of occlusions. Furthermore, the inclination of l_c is dependent on the depth. Line l_c corresponds to a plane p_c in our 4D Epipolar Volume $\widetilde{EV}(u, v, s, t)$. In all our examples $\Delta v = \Delta u$ the inclination in s and t is the same. The plane p_c can also be interpreted as the set of all the samples of the plenoptic function of one point P sampled by the light field.

Discretization Assuming a continuous light field, every point in space leads to a continuous line in the EPI as long as occlusion is ignored. However, the rasterization of l_c (cf. Fig. 4b) at an inclination smaller than 45° will lead to dotted lines, which are hard to be recognized using any filter or edge detector. The same problem arises when trying to fit the inclination of planes p_s in $\widetilde{EV}(u, v, s, t)$. Therefore, we compute a sequence of sheared spaces by progressively changing the shear factor s_{xy} corresponding to an inclination, such that $\widetilde{EV}'(u', v', s', t') = \widetilde{EV}(u, v, s + s_{xy} \cdot u, t + s_{xy} \cdot v)$ and check only shear planes $p_s = \widetilde{EV}'_s(u', v', s', t')$ orthogonal to the s and t direction. The reconstructed depth precision can be improved by increasing the number of shears.

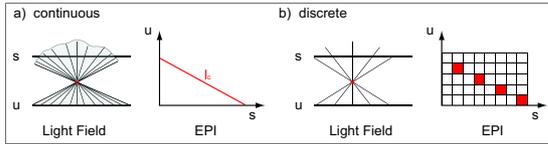


Figure 4: a) shows the continuous light field and its corresponding EPI. b) shows a discrete light field and its corresponding discrete EPI.

Frequency Minimization Criteria In the case of a Lambertian scene the color of the plane p_s is constant, as long as its pixels stemming from every light field image $LF(u, v, \cdot, \cdot)$ corresponds to the same point in the scene. This consistency criterion can be evaluated by minimizing the variance over p_s . In case of arbitrary BRDFs, the variance will fail (cf. Fig. 6) in most of the depth reconstructions of specular objects. By comparing Fourier transforms of different shear planes containing $\widetilde{EV}'(u, v, s, t)$ we observe predominantly low frequencies if the shear corresponds to the depth of the

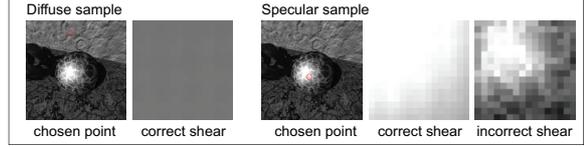


Figure 5: Shear planes for a diffuse and a specular point.

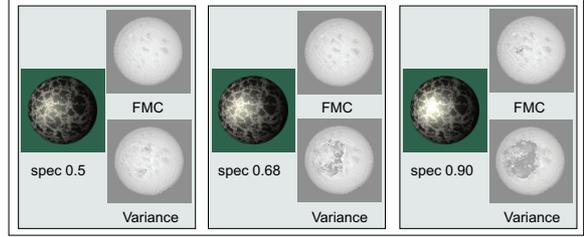


Figure 6: Depth reconstructions based on FMC and variance of the same sphere with varying specular coefficients are compared after the first pass of the 2-pass algorithm of Sect. 5.1. The variance leads to holes at specular reflections.

pixel at that position, albeit points showing specular highlights. This is based on the fact that the specular highlight becomes smaller for non-matching shear planes leading to higher frequencies in the spectrum. Additional texture magnifies the high frequencies for non-matching shears as well. Therefore, we introduce the following novel criteria based on the Fourier Transform \mathcal{F} , which we will refer to as *Frequency Minimization Criteria (FMC)*:

$$FMC(u, v, s, t, s_{xy}) = \sum_{P \in \mathcal{F}\{\cdot\}} w(P) \cdot |\mathcal{F}\{(p_s - \bar{p}_s) \cdot f_{apod}\}(P)|^2 \quad (1)$$

$$\bar{p}_s = \text{mean}(\widetilde{EV}'_s(u', v', s', t')) \quad (2)$$

$$z_R = \frac{\Delta u}{\frac{d_{px}}{N} \cdot \tan(\frac{\vartheta}{2})} \cdot 2 \quad (3)$$

with f_{apod} being an apodization function and $w(P)$ a frequency dependent weighting function. We use a weighted sum of the power spectrum, penalizing high frequencies and disregarding lower frequencies. This approach gives a finer control of the shear plane analysis than the variance. The shear s_{xy} corresponding to $\min_{s_{xy} \in S} (FMC(u, v, s, t, s_{xy}))$, with S being the set of all possible shears bounded by the closest and farthest object, leads to depth z_R (cf. Eq.(3)) of the ray $LF(u, v, s, t)$. d_{px} is the disparity in pixels, N the number of pixels in one view and ϑ the field of view (FoV).

Gray Channel Evaluation Since FMC can only be evaluated on monochrome images, we transform our color images into gray scale images using a technique presented in [GD05]. In order to transfer the gray levels to all the other views of the light field, we create either a big image containing all of the images for small light fields, or evaluate the first principal component of the first image and use it as an

axis to which all the colors of all the images are projected to. More elaborate but slower versions of color transformations such as presented in [GOTG05, KRJ05], did not seem to achieve better results according to our requirements.

Occlusion and Specular Highlights Occluded points always lie further away than the occluding points and hence lie on a shear plane with a larger shear factor s_{xy} . As soon as a foreground point is detected, all pixels corresponding to it are ignored for further FMC calculations. In order to avoid high frequencies due to missing pixels, a 3×3 Gauss filter is applied on the shear plane, leading to pixels usable to fill the holes. Furthermore, sharp features at occlusions as well as thin structures are preserved, since our algorithm ignores any kind of neighborhood for correspondence matching. Therefore, occlusions are modeled correctly and included in our reconstruction method.

Since the FMC finds the optimal shear despite specular highlights, we do not have to handle them in any particular way.

Algorithm We evaluate the FMC from the smallest shear, corresponding to the closest point, to the largest shear, corresponding to the farthest point, in order to find the global minimum per sample of the light field. We suggest two methods which focus on speed and accuracy respectively. In a one-pass algorithm, the global minimum is chosen once a local minimum has not been changed after the last w steps and the minimum frequency is below a certain threshold, where w is any number of shear steps (cf. Algo. 1). If the variance over p_s is smaller than a certain threshold, we do not evaluate the FMC, but choose the current s_{xy} as the optimal fit.

Input : Light Field

Output: Depth Field

% initialize FMC

$FMC(u, v, s, t, s_{xy}) = \infty$;

for $s_{xy} = \text{smallShear}$ **to** largeShear **do**

for all (s, t) **do**

 eval $FMC(\frac{u_{max}}{2}, \frac{v_{max}}{2}, s, t, s_{xy})$ for p_s ;

if $\text{Var}(\widetilde{EV}_s'(u', v', s', t')) < \text{Threshold}$ **then**
 choose shear;

 remove pixels from p_s ;

else if $FMC(u, v, s, t, s_{xy}) >$

$prev_w(FMC(u, v, s, t, s_{xy}))$ **then**

 assign prev. s_{xy} to pixels of p_s ;

 remove pixels from p_s ;

end

end

end

Algorithm 1: Depth reconstruction of the one-pass algorithm. $prev_w(\cdot)$ takes the minimum of the last w steps.

A more robust but slower two-pass algorithm removes the points for which a global minimum below a certain thresh-

old has been found after completing the first pass. In a second pass the remaining points are detected by evaluating the minimum FMC for the remaining samples. For a very complex scene the number of passes could be adapted.

Scenes showing reflections of surrounding objects will not be reconstructed properly, since altering colors caused by other objects will lead to high frequencies and are leading to an arbitrary FMC. In the case of big homogenous patches, no unique solution exists due to the lack of information in the light field. In this case we select the first minimal FMC to evaluate the depth.

5.2. Hologram Evaluation

The holography pipeline presented in [ZKG07] can be extended in order to handle the evaluation of a light field with a corresponding depth field. Instead of merging the depth maps from all the views to one sampled scene, we set one point source P_{uv} along each ray $LF(u, v, \theta, \phi)$ corresponding to a frequency component of the hologram at a depth corresponding to the depth field in order to minimize speckles in the reconstruction (cf. Sect. 7). Each P_{uv} is evaluated over the entire tile T_{uv} of size $\Delta u \times \Delta v$ as depicted in Fig. 7a. By evaluating one point source per ray, we implicitly include knowledge of view dependent occlusion and reflection properties of the scene captured by the light field. Efficient per point wavefield evaluation is enhanced by a hardware based implementation of point source evaluation. The contribution of every point source can be added up to obtain the wavefield.

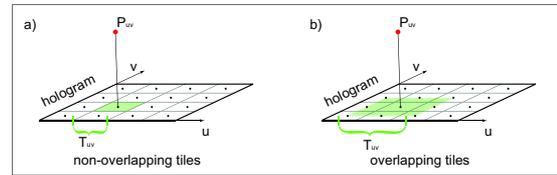


Figure 7: A point source P_{uv} lying on the ray $LF(u, v, s, t)$ will only be evaluated on the tile T_{uv} conserving knowledge of occlusion and reflection captured by the light field.

Overlapping Tiles Since the captured light field leads to a discrete sampling of the BRDF of points on the scene surface, discontinuities between two neighboring tiles can occur. The maximal discontinuity is dependent on the maximum frequency of the BRDF. In order to avoid ringing artifacts during the inverse mapping \mathbf{M}^{-1} we overlap and blend the tiles as shown in Fig. 7b. The overlap can be freely chosen between 0 to maximally Δu or Δv . The overlapping parts are linearly blended. Note that the blending does not have an influence on the interpolation of intensities between different BRDF samples for novel viewpoints, but solely avoids ringing artifacts when evaluating novel viewpoints.

Choice of Wavelength If the hologram is only used as an intermediate representation and not as an interference pattern for holographic screens, we are able to choose a wave

length with more flexibility. On the one hand, the wave length should be as big as possible in order to keep the required sampling fulfilling the Nyquist criteria as low as possible. On the other hand, the wavelength has to be short enough in order to guarantee the required resolution for the back transformation \mathbf{M}^{-1} into the light field leading to Sect. 7.

Speckle elimination Speckles occur if multiple point sources create the same frequency and the phases are distributed in a way canceling each other out. This is a known physical phenomenon inherent to coherent light modeling. Point sources create the same frequency if they lie in the same frequency band and therefore, in approximately the same direction from the center of the aperture (cf. Sect. 7). In order to reduce speckle noise in the final views, we only evaluate every n^{th} point source to create a hologram, leading to a number of n holograms.

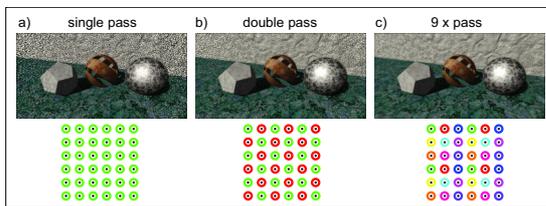


Figure 8: Speckle reduction through selective point source rendering. Each point source surrounded by the same colored ring will be evaluated on the same hologram. Mostly two holograms are sufficient to achieve speckle suppression.

The final image resulting from the holographic rendering or inverse mapping is a sum of the images of the n holograms. Using this technique we are able to improve the image quality from a straight forward evaluation Fig. 8a with $n = 1$ showing speckle to Fig. 8b and Fig. 8c being almost speckle free. To avert speckles created from corresponding points, we set identical phases for all of them. By increasing the aperture size over several tiles, speckles become visible at the straight forward transformation. However, since the bigger aperture size is leading to a higher resolution with high-frequency speckle, we low-pass filter and down sample the image to create an almost speckle free smaller image. The resized smaller image still has the same resolution as the corresponding light field view would have had.

6. Hologram Operations

A practical application of the transform \mathbf{M} is to create unique input for a holographic screen. Moreover, a holographic representation has various advantages, such as smooth parallax rendering without ghosting artifacts, robustness regarding data loss and diffraction simulation.

Smooth parallax rendering The hologram rendering is based on the holographic pipeline presented in [ZKG07]. Setting the aperture and viewpoint for a hologram rendering will handle interpolation of intensity information from the

light field implicitly. Since depth information of the scene is encoded in the phase of the hologram no ghosting artifacts are visible in novel viewpoints as shown in Fig. 9c and Fig. 9d. Although light fields do show ghosting as in Fig. 9b if no depth is known, they are not prone to speckles as holograms are. By choosing the optimal focal plane, the light field images can be improved considerably as depicted in Fig. 9a. Ghosting reduction has been studied intensively in [CCST00, IMG00, SYGM03].

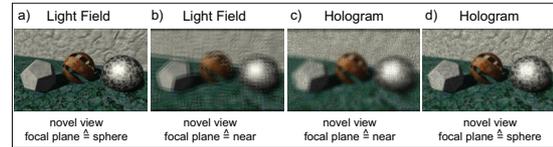


Figure 9: The same novel viewpoint is rendered for the light field and the hologram using various focal plane distances.

Effects of Loss of Data The hologram stores the information of the scene in the form of a wavefront. Therefore, every point of the scene has an influence on every pixel of the hologram as long as it is not occluded. This means that we can cut out parts of the hologram and still retain information about all the points in the scene as long as at least one part of the evaluation of all the points is still visible. If the aperture is chosen large enough in order to never fully lay over the cut out part, images for all viewpoints can still be reconstructed. Artifacts can occur if the cut out parts are not apodized, since high intensity differences can lead to ringing artifacts.

Compression Light field compression was already addressed in the pioneering work of Levoy and Hanrahan [LH96] and Gortler et al. [GGSC96]. Since then many compression strategies have been presented, which were most often inspired by standard image or video compression techniques. In contrast to this, hologram compression does not lend itself to standard image compression since the reconstruction quality can depend on the entire frequency range. Naughton et al. [NFM*01] showed the limited use of lossless compression for holograms. A better strategy for hologram compression is non linear quantization [SNJ06]. This preserves the spatial detail while still requiring relatively few bits per complex number.

7. Inverse Mapping

The inverse mapping \mathbf{M}^{-1} cannot be implemented as a straight forward inverse of \mathbf{M} , since the point sources are combined to one hologram. A wave based inverse propagation would lead to a lot of problems due to the limited aperture size. Furthermore, the complex valued spatially dependent point spread functions would have to be deconvolved in order to reconstruct each point source independent of one another.

Instead we render images at positions (u, v) on the holographic plane leading to directional components θ and ϕ (cf. Sect. 8), which can be interpreted as samples of an angular

light field $LF(u, v, \theta, \phi)$. By applying a perspective distortion as described in [ZKG07] the angular components can be transformed into a two-plane parametrization $LF(u, v, s, t)$. The transformation from the hologram to a light field can be done for any hologram, for which the original wavefront of the scene can be restored.

Desired light field resolution In the first step we have to determine the desired resolution $\Delta\theta$ and $\Delta\phi$ and the spacing Δu and Δv defining the light field. According to Sect. 8.2 the centers of the artificial apertures are set on every sample position (u, v) of the hologram. The size of the aperture a has to be chosen such that the minimal angular resolution $\Delta\alpha = \arcsin\left(\frac{\lambda}{a}\right)$ corresponds to $\min(\Delta\theta, \Delta\phi)$. For every aperture, a lens function and apodization function is multiplied with the wavefield before getting the directional components through the Short-Term Fourier Transform (STFT) described in Sect. 8.1. The resolution of u is limited by the number of samples on the hologram in this dimension.

Upper boundary of angular resolution The best possible resolution, which can theoretically be achieved for the transformation depends on the maximal depth extension of the visible part of the scene $\Delta\hat{z}_{visScene}$. The depth of field Δz (cf. App. B) has to be congruent with $\Delta\hat{z}_{visScene}$ allowing a maximal aperture size a leading to the highest resolution $\Delta\alpha$ for a given wavelength λ . Different techniques to elaborate $\Delta\hat{z}_{visScene}$ are found in numerous papers. We are using a technique proposed in [ZKG07] for depth reconstruction from the hologram.

8. Transformation

The angular spectrum $\mathcal{A}(\cdot)$ as presented in [Goo68] is a Fourier transform $\mathcal{F}(\cdot)$ decomposing the complex valued wavefront $U(u, v)$ into a collection of plane wave components propagating in different directions \mathbf{k} dependent on the spatial frequencies v_u and v_v (see Eq.(4)). The vector \mathbf{k} is defined as $\mathbf{k} = k \cdot (\alpha, \beta, \gamma)$ with $k = 2 \cdot \pi / \lambda$ being the wave number, λ the wavelength and (α, β, γ) being the unit vector pointing into the direction of the wave propagation. The components of the vector are called *directional cosines* and are related to the spatial frequencies by $\alpha = v_u \cdot \lambda$, $\beta = v_v \cdot \lambda$ and $\gamma = \sqrt{1 - \alpha^2 - \beta^2}$.

$$\begin{aligned} \mathcal{A}(v_u, v_v) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(u, v) e^{-2\pi i(v_u u + v_v v)} du dv \\ &= \mathcal{F}\{U(u, v)\} \end{aligned} \quad (4)$$

Every spatial frequency extends over the entire uv -domain and can, therefore, not be spatially localized. Nevertheless, [Goo68] shows that *local spatial frequencies* $v_l u$ and $v_l v$ can be obtained by a spatially limited Fourier transform as long as the phase $\varphi(u, v)$ does not change too rapidly (see Sect. 8.1).

8.1. Local Spatial Frequencies

We employ the Short-Term Fourier Transform (STFT) also known as the Sliding-Window Transform, where the wavefront $U(u, v)$ to be Fourier transformed is multiplied by a window function $h(u, v)$, which is nonzero for a limited area around the origin. The resulting spectrum $S(\cdot)$ is called the *local frequency spectrum* and is defined as follows:

$$S(v_u, v_v, x, y; h) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{U}(u, v) e^{-2\pi i(v_u u + v_v v)} du dv \quad (5)$$

$$\hat{U}(u, v) = U(u, v) h(u - x, v - y). \quad (6)$$

The multiplication by $h(\cdot)$ suppresses $U(\cdot)$ outside the window and causes a localization. However, since this transformation is governed by the Heisenberg-Gabor inequality, as shown in [Fla99] we cannot get a perfect localization in the spatial domain as well as in the frequency domain.

Considering the analysis of a wavefront, we can say that the better the localization of the directional components, the less directions can be specified. Nevertheless, applying a lens localizes the frequencies for points at specific depths. In the following section we use the principle of STFT, but improve the quality of localization for certain depths.

8.2. Aperture

The window $h(\cdot)$ can be regarded as an aperture S_A which blocks the incoming wavefront outside of it. By evaluating the wavefield $U(u, v)$ from a point source P on S_A and transforming it using the STFT we obtain the directional components of the planar waves describing $U(u, v)$. For a point P of finite distance (cf. Fig. 10a) $U(u, v)$ leads to several planar waves, and therefore no localization in the frequency domain.

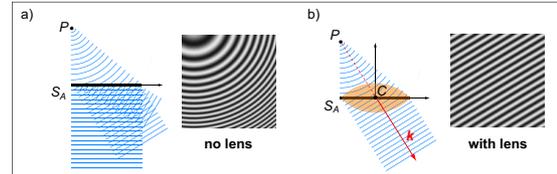


Figure 10: a) shows the frequency distribution over the whole aperture caused by a point source. b) shows the frequency distribution when using a lens.

By introducing a lens with focal length f as in Eq.(7), the incoming wavefront from P can be transformed into a single plane wave as shown in Fig. 10b. Introducing a lens does not only have the benefit of creating a single plane wave, but also gives information about the spatial location of the directional wave. Under the assumption of P being perfectly in focus, P lies on the line defined by the center of the aperture C and the directional vector \mathbf{k} .

$$\ell(u, v) = e^{-ikr}, \text{ with } r = \sqrt{u^2 + v^2 + f^2} \quad (7)$$

A lens has theoretically exactly one focal plane, which

lies at distance z from the lens and therefore, only planar surfaces parallel to the aperture could be transformed into rays. In practice however, every capturing system has a certain resolution, which determines a Circle of Confusion (CoC) (cf. App. A). Taking this CoC into account we can define a Depth of Field (DoF) (cf. App. B) in which all the point sources appear in focus and can be transformed into rays going through C . If the whole object is considered to lie in the depth of field of a lens, the frequency distribution of each point will not extend over more than one discrete frequency measure and the wavefield of the scene at the aperture can be transformed into a light field. To get the highest angular resolution (cf. App. C) and therefore the biggest aperture we have to achieve a tight fit of the DoF around the object including diffraction.

The lens function, aperture size, wavefront sampling and the wavelength define the resulting CoC and DoF yielding a light field with a specific maximal resolution for s and t as well as a maximal FoV for every position (u, v) . The dependence between these characteristics are elaborated in the following appendices App. A, App. B, App. C and App. D.

9. Results

All hologram renderings presented in this section are only computed to give an illustration of direct output of holographic content on future generation holographic displays. They do not compete with the light field renderings which are by far more efficient for conventional 2D framebuffer displays. We show the versatility and the power of \mathbf{M} and \mathbf{M}^{-1} by applying it to several examples, such as synthetic light fields, real light fields and digitally recorded holograms. The rendered images can be evaluated directly from the holographic representation or through light field rendering. We implemented a light field renderer using a spatial method capable of simulating different aperture sizes as well as focal length for viewpoints in the uv -plane. A more efficient implementation has been presented in [Ng05] and would have to be used if real-time performance was a requirement. Evaluations of a hologram from a light field and depth field as well as all the renderings from the holograms have been integrated into the pipeline presented in [ZKG07].

9.1. Forward Mapping

We compute three **synthetic** scenes shown in Fig. 11. The dataset in Fig. 11a is a POV-Ray rendered $384 \times 192 \times 16 \times 16$ light field for which our depth map reconstruction algorithm requires 60 minutes on a Pentium 4 with 3.2GHz. The result is a depth map with 163 possible depth values per light field sample, while handling various difficulties such as occlusions, areas of low texture and specular highlights correctly. Fig. 11b and Fig. 11c depict a light field rendered using RenderMan13 of the procedurally generated Pompeii scene presented in [MWH*06]. Both scenes contain a very big depth range, which is bigger than the depth of field of the camera used for the holographic rendering. Regions being slightly out of focus are therefore spread over multi-

ple frequencies leading to some speckle noise. According to Sect. 5.2 speckle diminishes for bigger apertures again. The renderings of the hologram are not primarily shown to demonstrate holographic rendering, but to show a possible view, which could be generated on a holographic screen. The human eye would transform the wavefront into an image and therefore determine aperture size and focal length.

Most importantly we transform a **real** light field (cf. Fig. 11d) into a hologram in order to show that our method can be applied to capture holograms under white light illumination. The depth map reconstruction shows some artifacts since we had no camera calibration and the images suffered from lens distortion. However, the depth map is still precise enough in order not to show any ghosting.

9.2. Inverse Mapping

The transformation from the hologram to a light field can be done for any hologram, for which the original wavefront of the scene can be restored. The third column of every sequence of Fig. 11 represents a reconstructed light field view by applying \mathbf{M}^{-1} . Direct comparisons show some distortions at off-axis rays for cameras with a big FoV.

Furthermore, we transformed a digitally recorded hologram into a light field in order to show the versatility of our framework and transform Fig. 12. Our proposed speckle reduction cannot be applied to digitally recorded holograms, so the final renderings are speckle prone.

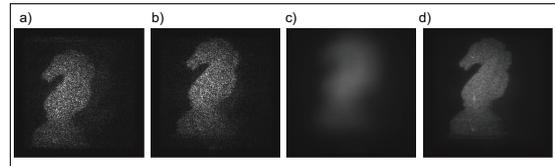


Figure 12: a) and b) are light fields generated from a digitally captured hologram rendered by a small aperture at two positions. c) shows a big aperture with short focal length and d) a big aperture with focal length on object.

Limitations A limitation of this framework consists in transforming arbitrary light fields featuring strong reflections and transparent objects, since those regions can fail during depth reconstruction. Furthermore, the lack of visible rays of a point at the border of the light field might not provide enough information for a robust depth estimation in all the views. In our examples at least 20% of the rays have to be visible to reconstruct the depth of a point. The resulting holes are filled through interpolation of surrounding depth values. Inaccurate depth could lead to ghosting for novel viewpoints.

For scenes with a big depth extent speckles can be noticed. Therefore, the bigger the depth extent of the scene the more holograms have to be evaluated for a perfect image. Furthermore, the applied lens model leads to aberrations for non-paraxial rays and can therefore lead to speckles.

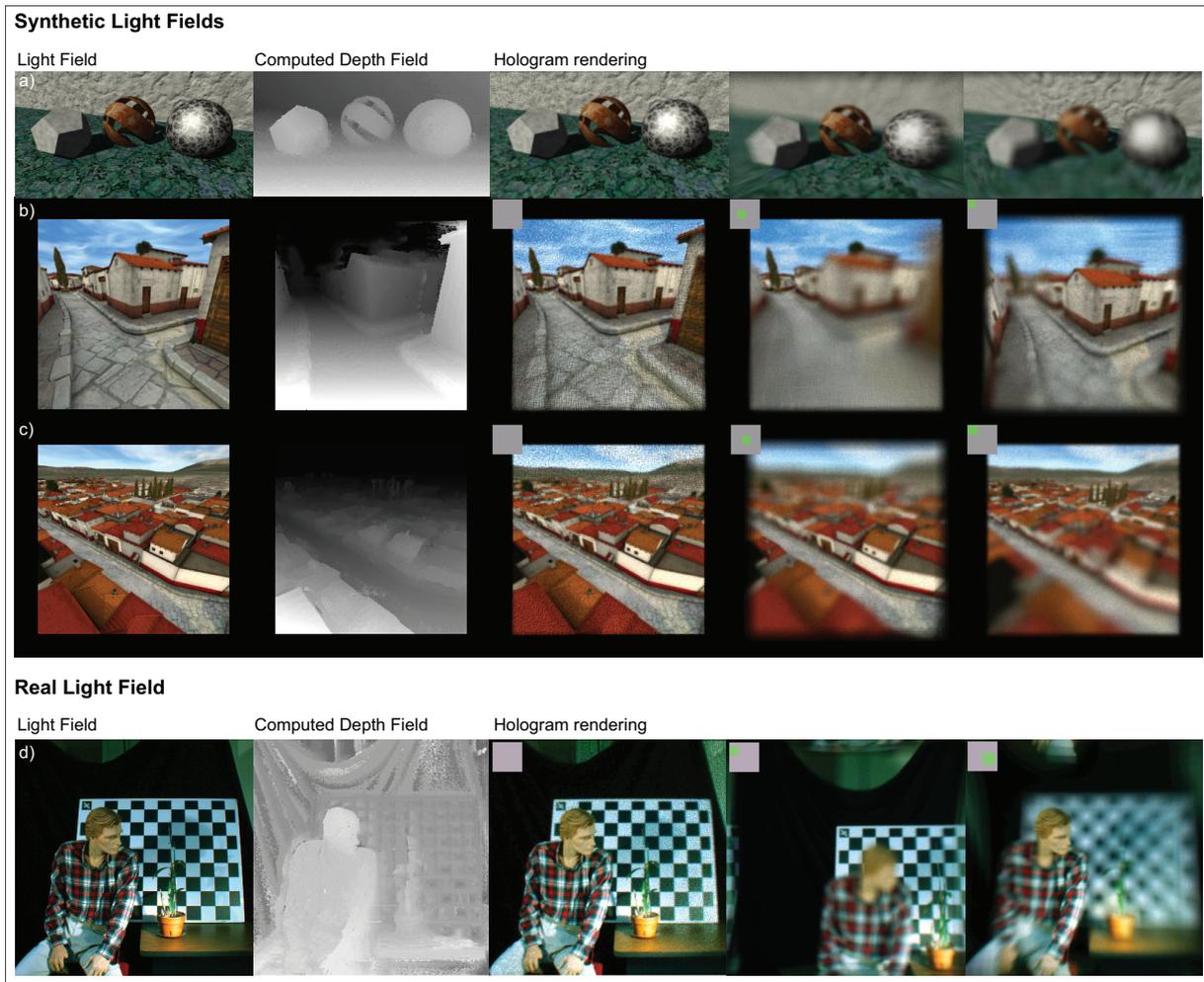


Figure 11: All the image sequences show an original view of the light field input, its corresponding depth field, the rendered transformation into a hologram and two arbitrary views of the hologram with varying aperture and focal length. The grey square symbolizes the hologram, while the green square shows the position and aperture size of the camera.

10. Conclusions

In this paper we presented a fundamental mapping from light fields to holograms and demonstrated the versatility on multiple examples. For the forward mapping we introduced a novel 3D reconstruction technique based on frequency spectrum analysis capable of evaluating depth despite of occlusions, specular highlights, and low texture information. The created depth field provides the base for a forward transform into a hologram. Most importantly this gives the possibility to capture full parallax holograms under natural illumination, which has not been possible so far. This creates a big potential for future work in this field. Furthermore, the inverse mapping operation allows for digitally captured holograms to be rendered in real-time using the light field representation.

11. Future Work

Based on the elaborated mapping operation, holograms can be captured using a light field camera as presented in [Ng05] and rendered on a holographic screen as presented by Qinetiq in [SCS05]. This technique can take advantage of the realism and detail preserving benefits of a real light field while giving the possibility of a 3D output on a holographic screen. Furthermore, the 3D reconstruction technique can be used for ghosting reduction in light field rendering without having to blur any part of the scene. Various lens effects can further be used to create realistic looking renderings for general graphics processing. Finally, digital holograms not requiring optical elements for acquisition, can be rendered in real-time after mapping them into light fields. Therefore, future work can benefit in numerous ways from the fundamental map-

ping by taking advantage of either representation, depending on the needs.

Acknowledgement

We want to thank Leonard McMillan for providing a light field dataset from a real scene as well as Pascal Mueller for rendering several light fields from the Pompeii scene. Furthermore, we want to thank Ervin Kolenovic and Jan Mueller from BIAS for providing the digital hologram, Tim Weyrich and Miguel A. Otaduy for helpful comments, as well as Marc Levoy for very valuable comments on the final version. The POV-Ray code for one of the objects used in the artificial scene was provided by Tor Olav Kristensen. Lukas Ahrenberg is supported by the EC within FP6 under Grant 511568 with the acronym 3DTV.

References

- [BBH87] BOLLES R. C., BAKER H. H., H. MARIMONT D.: Epipolar-Plane Image Analysis: An Approach to Determining Structure from Motion. In *International Journal of Computer Vision* (1987), pp. 7–55.
- [BN95] BHAT D., NAYAR S.: Stereo in the presence of specular reflection. *iccv 00* (1995), 1086.
- [BW59] BORN M., WOLF E.: *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Pergamon Press, 1959.
- [CCST00] CHAI J.-X., CHAN S.-C., SHUM H.-Y., TONG X.: Plenoptic sampling. In *SIGGRAPH '00: Proceedings of the 27th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 2000), ACM Press/Addison-Wesley Publishing Co., pp. 307–318.
- [CKS*05] CRIMINISI A., KANG S. B., SWAMINATHAN R., SZELISKI R., ANANDAN P.: Extracting layers and analyzing their specular properties using epipolar-plane-image analysis. *Comput. Vis. Image Underst.* 97, 1 (2005), 51–85.
- [DeB] DEBITETTO D.: Holographic panoramic stereograms synthesized from white light recordings. *Applied Optics*.
- [DYW05] DAVIS J. E., YANG R., WANG L.: Brdf invariant stereo using light transport constancy. In *ICCV '05: Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1* (Washington, DC, USA, 2005), IEEE Computer Society, pp. 436–443.
- [Fla99] FLANDRIN P.: *Time-frequency/time-scale analysis*, vol. 10 of *Wavelet Analysis and its Applications*. Academic Press Inc., San Diego, CA, 1999. With a preface by Yves Meyer, Translated from the French by Joachim Stöckler.
- [GD05] GRUNDLAND M., DODGSON N. A.: *The decolorize algorithm for contrast enhancing, color to grayscale conversion*. Tech. Rep. UCAM-CL-TR-649, University of Cambridge, Computer Laboratory, Oct. 2005.
- [Ger36] GERSHUN A.: The light field. *Translated by E. Moon and G. Timoshenko in Journal of Mathematics and Physics XVIII*, 1 (1936), 51–151.
- [GGSC96] GORTLER S. J., GRZESZCZUK R., SZELISKI R., COHEN M. F.: The lumigraph. *Computer Graphics 30*, Annual Conference Series (1996), 43–54.
- [Goo68] GOODMAN J. W.: *Introduction to Fourier Optics*. McGraw-Hill Book Company, San Francisco, 1968.
- [GOTG05] GOOCH A. A., OLSEN S. C., TUMBLIN J., GOOCH B.: Color2gray: salience-preserving color removal. *ACM Trans. Graph.* 24, 3 (2005), 634–639.
- [Hal94] HALLE M.: Holographic stereograms as discrete imaging systems, 1994.
- [Hop55] HOPKINS H. H.: The Frequency Response of a Defocused Optical System. *Royal Society of London Proceedings Series A 231* (July 1955), 91–103.
- [IMG00] ISAKSEN A., MCMILLAN L., GORTLER S. J.: Dynamically reparameterized light fields. In *SIGGRAPH '00: Proceedings of the 27th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 2000), ACM Press/Addison-Wesley Publishing Co., pp. 297–306.
- [Jan97] JANSSON P.: *Deconvolution of Images and Spectra*. Academic Press, New York, 1997.
- [KRJ05] KARL R., ROBERT G., JAMES W.: Re-coloring images for gamuts of lower dimension. *Computer Graphics Forum* 24, 3 (2005), 423–432.
- [Lev06] LEVOY M.: Light fields and computational imaging. *Computer* 39, 8 (2006), 46–55.
- [LG95] LUCENTE M., GALYEAN T. A.: Rendering interactive holographic images. In *SIGGRAPH '95: Proceedings of the 22nd annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 1995), ACM Press, pp. 387–394.
- [LH96] LEVOY M., HANRAHAN P.: Light field rendering. *Computer Graphics 30*, Annual Conference Series (1996), 31–42.
- [LLL*02] LI Y., LIN S., LU H., KANG S. B., SHUM H.-Y.: Multibaseline stereo in the presence of specular reflections. In *ICPR '02: Proceedings of the 16th International Conference on Pattern Recognition (ICPR'02) Volume 3* (Washington, DC, USA, 2002), IEEE Computer Society, p. 30573.
- [LNA*06] LEVOY M., NG R., ADAMS A., FOOTER M., HOROWITZ M.: Light field microscopy. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Papers* (New York, NY, USA, 2006), ACM Press, pp. 924–934.
- [Mat05] MATSUSHIMA K.: Computer-generated holograms for three-dimensional surface objects with shade and texture. *Applied Optics* 44 (August 2005), 4607–4614.

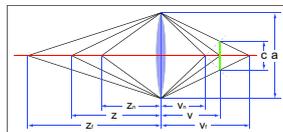
- [MWH*06] MÜLLER P., WONKA P., HAEGLER S., ULMER A., GOOL L. V.: Procedural modeling of buildings. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Papers* (New York, NY, USA, 2006), ACM Press, pp. 614–623.
- [NFM*01] NAUGHTON T. J., FRAUEL Y., MATOBA O., BERTAUX N., JAVIDI B., TAJAHUERCE E.: *Compression and numerical reconstruction of digital holograms*. Tech. Rep. NUIIM/SS/01/13, Signals and Systems Group, National University of Ireland, Maynooth, December 2001.
- [Ng05] NG R.: Fourier slice photography. *ACM Trans. Graph.* 24, 3 (2005), 735–744.
- [SCS05] SLINGER C., CAMERON C., STANLEY M.: Computer-generated holography as a generic display technology. *Computer* 38, 8 (2005), 46–53.
- [SNJ06] SHORTT A. E., NAUGHTON T. J., JAVIDI B.: A companding approach for nonuniform quantization of digital holograms of three-dimensional objects. *Opt. Express* 14 (2006), 5129–5134.
- [STM06] STICH T., TEVS A., MAGNOR M.: Global Depth from Epipolar Volumes - A General Framework for Reconstructing Non-Lambertian Surfaces. In *Third International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT)* (2006), pp. 1–8.
- [SYGM03] STEWART J., YU J., GORTLER S., MCMIL L.: A new reconstruction filter for undersampled light fields. pp. 150–156.
- [WJV*05] WILBURN B., JOSHI N., VAISH V., TALVALA E.-V., ANTUNEZ E., BARTH A., ADAMS A., HOROWITZ M., LEVOY M.: High performance imaging using large camera arrays. *ACM Trans. Graph.* 24, 3 (2005), 765–776.
- [YEPM02] YANG J., EVERETT M., BUEHLER C., MCMILLAN L.: A realtime distributed light field camera. pp. 77–86.
- [ZKG07] ZIEGLER R., KAUFMANN P., GROSS M.: A framework for holographic scene representation and image synthesis. *IEEE Transactions on Visualization and Computer Graphics* 13, 2 (2007), 403–415.

Appendix A: Circle of Confusion

The Circle of Confusion (CoC) c is defined as the size of the circle to which an idealized point will diverge when the lens is focused at different length. Assuming a ray representation of light, the CoC caused from defocus $c_{defocus}$ is

$$c_{defocus} = \frac{f^2(z_f - z_n)}{N(2z_n z_f - f(z_n + z_f))}, \quad (8)$$

This implies that the CoC at focal distance f is zero. Even with a perfect lens, a point will not lead to a point in the image but to the Airy disk governed by diffraction when simulating light as waves. The CoC limited by diffraction c_{diff} is given by $c_{diff} \approx 2.44\lambda N(1+m)$ with m being a magnification factor $m = \frac{f}{z-f}$.



Despite the physical limitation of the minimal CoC given by diffraction c_{diff} , the CoC can also be restricted by the resolution of the discretizing media (e.g. film grain, pixel size of a camera) and therefore, regarded to be in focus. In our case the CoC is limited by the angular resolution $\Delta\alpha$ given by the aperture size a and the wavelength λ for a hologram and by the pixel size $\max(\Delta s, \Delta t)$ of the CCD when capturing a light field.

Appendix B: Depth of Field

The Depth of Field (DoF) $\Delta z = z_f - z_n$ as shown in Fig. A is the distance between the closest point in focus and the farthest point in focus, where a point in focus is determined by the CoC. The DoF considering only geometrical optics is given as

$$\Delta z = \frac{2zNcf^2(z-f)}{f^4 - (Nc(z-f))^2}. \quad (9)$$

There is no simple numerical expression to combine the effects of defocus and diffraction besides using an empirical expression. Furthermore, the perception of sharpness is not solely dependent of the finest resolution but also on contrast. The ability to transfer contrast of an input pattern with a given frequency ν by a diffraction-limited lens with defocus can be described by the Optical Transfer Function $OTF(c_{defocus}, N, \nu)$ as shown in [Hop55] by

$$OTF(c_{defocus}, N, \nu) = \begin{cases} \frac{4}{\pi g} \int_0^{\sqrt{1-s^2}} \sin(\sqrt{1-y^2}-s) ds & \text{if } s \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

where $s = \lambda \nu N(1+m)$ and $g = \pi \nu c_{defocus}$. In [Jan97] and [BW59] the maximal resolution according to the Rayleigh criterion is given, as long as two points can be separated so that a 19% dip appears between the peaks. Applying this threshold to the OTF leads to a maximal spatial frequency ν , which can be resolved by the lens. Diffraction can be ignored if defocus is sufficiently big or N sufficiently small leading to a simplified $OTF(c, \nu) = \frac{2J_1(g)}{g}$ with $J_1(\cdot)$ being the first-order Bessel function of the first kind. Fig. 13a) shows three curves for the OTF corresponding to the plane in focus without defocus, the OTF dependent of defocus and diffraction at the DoF limits and the OTF ignoring diffraction.

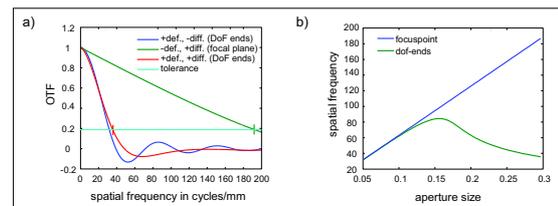


Figure 13: The graph in a) shows the OTF at focal distance in green and DoF limits in red, as well as the OTF dependent of defocus only in blue. In graph b), the curve for an $OTF=0.19$ at focal distance and at DoF limits is given in blue and green, respectively.

In Fig. 13b) the spatial frequency ν depending of the f-Number N with $OTF=0.19$ and constant DoF is depicted. This gives the possibility to choose a desired resolution represented as a spatial frequency and setting the depth of field including the complete scene in order to obtain the biggest possible aperture.

Appendix C: Resolution / Sampling

Different resolutions and samplings have to be considered for the recording and the transformation of a hologram and a light field. By *resolution* we mean the maximal number of samples in a certain dimension, whereas *sampling* refers to the size of one sample. We take into account the maximal number of pixels n_{Xmax} of a resulting image and the number of resulting camera positions n_{Umax} . For simplicity the resolution and sampling is always given for one dimension and can be handled analogously for the other dimension.

Holograms with resolution n_{holo}^2 and extent a_{holo} can have a minimal angular sampling of $\Delta\alpha = \arcsin\left(\frac{\lambda}{a_{holo}}\right)$. The angular resolution is defined by the number of samples n_{apt} of the aperture simulating the camera. The maximal resolution $n_{max} = \frac{2a_{holo}}{\lambda}$ is achieved if the FoV (see Sect. D) is 180° . Furthermore, the number of useful camera positions for the transformation equals n_{holo} in each dimension.

Angular parameterized Light Fields are most easily compared to holograms. The angular resolution of the image is given by the number of samples in θ -dimension with the maximal resolution $n_{max} = \frac{180^\circ}{\Delta\theta}$ with $\Delta\theta$ being the sampling distance. The number of possible camera positions equals the number of samples in u .

For **Two-plane parameterized Light Fields**, the maximum resolution of a rendered image is equal to the maximum number of samples in the s direction. The number of different views depends on the number of samples in the u dimension.

Comparison By neglecting compression we can see that the hologram is capable to store more different views with high resolution, than any of the light field representations. This efficiency is due to the sampling of the wavefield rather than the sampling of different viewpoints, where information like depth is being disregarded. However, holograms are prone to speckle noise because of the coherent light. Speckle size can be defined as $d_{sp} = \frac{\lambda b}{a}$ with b being the distance of the point source to the imaging system and a being the aperture size.

Appendix D: Field of View

The maximal FoV α_{max} of a hologram can be determined by the sample size Δu as $\alpha = \arcsin\left(\frac{\lambda}{2\Delta u}\right)$. This implies a phase difference between two samples of maximally π just reaching the Nyquist frequency. The FoV α of a freely chosen aperture depends on the number of samples n inside the aperture, since one sample corresponds to an angular sampling of $\Delta\alpha$ and the FoV to $\alpha = \Delta\alpha \cdot n$ not exceeding α_{max} .

The maximal FoV α_{max} for the light field depends on the parametrization. For $LF(u, v, \theta, \phi)$ the maximum is defined as $\alpha_{max} = \max(\theta, \phi)$ considering θ and ϕ of all rays. For a two-plane parametrization $LF(u, v, s, t)$ the FoV depends on the extent and the distance between the uv -plane and the st -plane. Furthermore the view frustum is sheared if only rays going through both planes are considered.