

Linear Volumetric Focus for Light Field Cameras

DONALD G. DANSEREAU, OSCAR PIZARRO, and STEFAN B. WILLIAMS
University of Sydney

We demonstrate that the redundant information in light field imagery allows volumetric focus, an improvement of signal quality that maintains focus over a controllable range of depths. To do this, we derive the frequency-domain region of support of the light field, finding it to be the 4D hyperfan at the intersection of a dual fan and a hypercone, and design a filter with correspondingly shaped passband. Drawing examples from the Stanford Light Field Archive and images captured using a commercially available lenslet-based plenoptic camera, we demonstrate that the hyperfan outperforms competing methods including planar focus, fan-shaped antialiasing, and nonlinear image and video denoising techniques. We show the hyperfan preserves depth of field, making it a single-step all-in-focus denoising filter suitable for general-purpose light field rendering. We include results for different noise types and levels, through murky water and particulate matter, in real-world scenarios, and evaluated using a variety of metrics. We show that the hyperfan's performance scales with aperture count, and demonstrate the inclusion of aliased components for high-quality rendering.

Categories and Subject Descriptors: I.4.1 [**Image Processing and Computer Vision**]: Digitization and Image Capture—*Imaging geometry*; I.4.3 [**Image Processing and Computer Vision**]: Enhancement—*Filtering*

General Terms: Algorithms, Theory, Design

Additional Key Words and Phrases: Light field filtering, digital refocusing, plenoptic cameras, computational photography, Fourier analysis

ACM Reference Format:

Donald G. Dansereau, Oscar Pizarro, and Stefan B. Williams. 2015. Linear volumetric focus for light field cameras. *ACM Trans. Graph.* 34, 2, Article 15 (February 2015), 20 pages.

DOI: <http://dx.doi.org/10.1145/2665074>

1. INTRODUCTION

Focus has existed almost as long as photography, and is employed in all modern cameras. Focus is used to selectively emphasize elements

This work is supported in part by the Australian Centre for Field Robotics, the New South Wales State Government, The University of Sydney, and the Australian Government's International Postgraduate Research Scholarship and Endeavour Research Fellowship.

Authors' addresses: D. G. Dansereau (corresponding author), O. Pizarro, S. B. Williams, Australian Centre for Field Robotics, School of Aerospace, Mechanical and Mechatronic Engineering, University of Sydney, NSW, Australia; email: donald.dansereau@gmail.com.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permission from Permission@acm.org.

© 2015 ACM 0730-0301/2015/02-ART15 \$15.00
DOI: <http://dx.doi.org/10.1145/2665074>

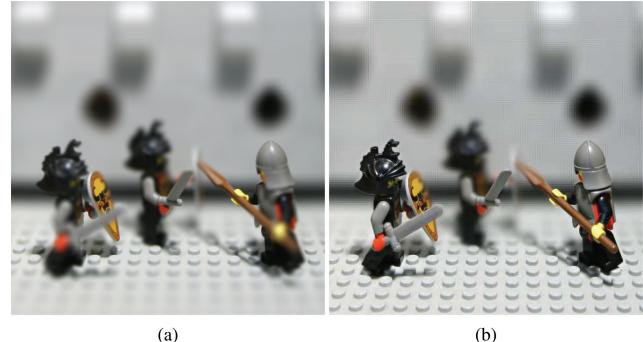


Fig. 1. (a) Conventional focus improves SNR but maintains sharpness on a single plane; (b) in this work we demonstrate volumetric focus dramatically improving SNR while maintaining focus over a controllable range of depths. Code and example light fields are available online at <http://marine.acfr.usyd.edu.au/permlinks/Plenoptic>. Original light field courtesy the Stanford Computer Graphics Laboratory.

of a scene, controlling the level and shape of blur—the “bokeh”—to yield an aesthetically pleasing result. It is easy to forget that a key motivation for focus, and probably the original reason it came about, is not to blur background elements but to gather more light, shortening exposure times and increasing signal-to-noise ratio (SNR). The side-effect of this enhanced light gathering is a narrowed depth of field, and applications that benefit from both a large depth of field and light gathering must strike a balance between the two. Indeed, this trade-off can impact or even prevent imaging in low contrast—at night, through murky water, smoke, cloud, fog, or dust—or where exposure times are limited due to motion.

In combating low contrast, the obvious approach of increasing illumination is not always applicable: Large scenes cannot always be effectively lit, illumination power budgets are typically limited, and, in the presence of scattering media, backscatter can negate any advantage gained by increasing illumination. The alternative approach of increasing exposure duration also finds limited success where dynamic scenes or platforms lead to motion blur.

Interfering scene elements such as snow, rain, underwater particulate matter, and other heterogeneous occluders can further complicate imaging. These are distinct from low-contrast scenarios in that the fundamental limitation is not a lack of signal, with imaging ultimately limited by sensor noise, but rather interference present within the signal itself. Increasing illumination or sensitivity will not help remove partial occluders; they will simply be imaged with higher fidelity. Widening a camera's aperture does help remove occluders, not by virtue of gathering more light, but rather by increasing depth selectivity to better isolate desired scene content.

The ability to better control the trade-offs associated with focus is clearly desirable, with potential applications including mobile robotics, autonomous driving, consumer photography, and surveillance. Recent developments in computational photography have demonstrated the potential of such an approach, with methods in focal sweep, flutter shutter, and multiple-exposure-duration video all allowing a camera to gather more light than would normally be

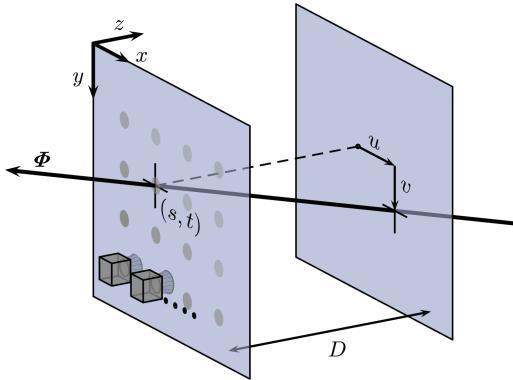


Fig. 2. Two-plane parameterizations of light rays. Shown is the *relative* two-plane parameterization. The points of intersection of a ray with two parallel planes completely describe its position and orientation in space. By convention, the s, t plane is closer to the camera, and the u, v plane is closer to the scene.

possible for a given depth of field and exposure time [Nagahara et al. 2008; Raskar et al. 2006; Agrawal et al. 2009]. Quantitative analyses confirm these techniques offer significant benefit in low-SNR applications [Cossairt et al. 2012; Mitra et al. 2013].

This article employs light field imaging to break the trade-offs of conventional imaging. We present *volumetric focus*, a method that improves SNR while maintaining sharp focus over a user-selected range of depths, rather than at a single depth as in conventional focus; examples of conventional and volumetric focus are depicted in Figure 1. In contrast to previous all-in-focus or extended depth-of-field techniques, the proposed method employs a simple, linear single-step filter to combine information from across the light field.

The remainder of this article is organized as follows: We provide background on light field imaging and related work in Section 2 and develop the light field characteristics central to the article in Section 4. These characteristics are exploited in Section 5 to derive volumetric focus filters. Sections 6 and 7 show results for camera array and lenslet-based light fields, giving quantitative and qualitative analyses of the volumetric filter’s performance. The work concludes with discussion and directions for future work in Section 8.

2. BACKGROUND

Whereas a conventional camera measures variations in light as a function of direction for rays passing through a single position, a light field (also “plenoptic”) camera encodes variations in light as a function of both direction and position. This can be achieved by introducing coded masks [Veeraraghavan et al. 2007] or lens arrays [Ng et al. 2005] into the optical path of a camera, or by constructing a grid of conventional cameras [Wilburn et al. 2005] depicted as boxes in Figure 2. The resulting images enable a range of new capabilities, including featureless or linear methods for traditionally nonlinear, iterative, or nondeterministic tasks including depth-selective filtering, distractor isolation, and visual odometry [Ng et al. 2005; Dansereau and Bruton 2007; Dansereau and Williams 2011; Yang et al. 2007; Dansereau et al. 2011].

Light field imaging offers important benefits in challenging imaging conditions, most notably in breaking the conventional trade-off between depth of field and SNR. Both arrays of cameras and lenslet-based cameras gather significantly more light for a given depth of field than conventional cameras [Ng et al. 2005]. Specifically, in both an array of $N \times N$ cameras and a lenslet-based plenoptic

camera with $N \times N$ pixels per lenslet, the increase in light gathering for a given depth of field is N^2 —a huge improvement. In both cases, the effective baseline also increases, increasing depth selectivity and the ability to reject occluders.

However, the redundant light that plenoptic cameras capture must be combined computationally in order to yield the maximum benefit. Combining light field information to improve SNR is the main focus of this article and it is not without precedent. It is well established that a light field contains sufficient information to allow post-capture focus through appropriate filtering [Isaksen et al. 2000; Ng 2005]. This virtual focus demonstrates similar properties to conventional focus: It combines light coming from different directions to increase SNR and simultaneously offers depth selectivity, blurring scene elements that fall outside a plane of focus. Because plenoptic focus can be tuned after the imagery has been captured, there is no need to decide ahead of time on a single focal setting.

In this work we generalize planar to volumetric focus. As in planar focus, volumetric focus combines light coming from different directions to increase SNR. Unlike conventional focus, volumetric focus keeps a *range* of depths in focus, blurring scene elements outside the focal *volume* as in Figure 1. The filter we present is useful where planar focus is useful: in ameliorating low contrast due to lack of illumination, murky water or other attenuating media, and in seeing around heterogeneous occluders. Volumetric focus can simplify system design by offering different trade-offs in depth of field and SNR than are possible with planar focus. This is particularly important where large baselines are present: An array of cameras sharply focused at a single depth will display a high SNR, but over a very narrow depth of field. Many applications deal with nonplanar scenes, and so the ability to put a volume in focus becomes highly desirable. Volumetric focus also simplifies applications in which a variable focal plane, adjusted to match the scene content, can be replaced with a fixed focal volume designed to encompass all typical scene depths.

Throughout this article we employ the *relative* two-plane parameterization depicted in Figure 2, in which light rays are described by their points of intersection with two parallel planes: an s, t plane, by convention closest to the camera, and a u, v plane at distance D , by convention closer to the scene. The continuous-domain light field signal $L(s, t, u, v)$ describes all light rays passing through the s, t and u, v planes. In the case of a light field camera array, it is often most convenient to place the camera apertures within and aligned with the s, t plane, as depicted in Figure 2. One of the advantages of the relative two-plane parameterization is that if one selects D to equal the focal length of the cameras in the array, u and v then coincide with physical coordinates on the image sensor. One can then think of the s, t plane as selecting a camera, and u, v as selecting a pixel.

3. RELATED WORK

Denoising of conventional imagery is a rich and active area of research, and a good review is provided by Buades et al. [2005]. See also Guleryuz [2007] for modern overcomplete dictionary developments, and Aharon et al. [2006] and Elad and Aharon [2006] for a singular value decomposition generalization of K-means for learning dictionaries directly from noisy imagery. Because we are dealing with high-dimensional imagery, video denoising is also relevant, including recent advances in block matching and filtering [Dabov et al. 2007].

Alternative approaches to low-light and contrast-limited imaging have appeared in the realm of computational photography. Levoy et al. [2004] demonstrate an active illumination generalization of

confocal imaging, allowing effective imaging through turbid media. A converse of this structured light approach, in which the position of the camera is varied rather than that of the illumination source, yields the light-field-based method explored in this article. O’Toole et al. [2012] augment the structured light method by including a variable camera mask, allowing a range of light transport phenomena to be investigated through completely optical processes. Relevant capabilities of this system are depth selectivity and the improvement of contrast through turbid media.

Iterative variational Bayesian frameworks have been explored for combining light measured across many apertures [Bishop and Favaro 2012; Goldluecke and Wanner 2013]. Our work differs significantly in its complexity; we present a single, noniterative linear filter as a means of combining images from across the light field, offering a simpler and potentially more robust solution. Yu et al. [2013] tackle denoising of light fields measured using reflective spheres, employing a robust image registration technique. Again, our work differs in its level of complexity by offering a linear, noniterative solution.

Several techniques for enhancing a camera’s depth of field or light gathering ability have emerged from the domain of computational photography. These include focal sweep, flutter shutter, and motion blur mitigation from multiple-exposure-time video [Nagahara et al. 2008; Raskar et al. 2006; Agrawal et al. 2009]. Because of the fundamental differences in approach, these offer significantly different performance trade-offs to the method we present.

The key principle underlying much of this article is ultimately parallax motion and its consequences in the 4D frequency domain. Parallax motion is a common thread throughout light field research and indeed much of computer vision, including stereo and multiple-camera geometry and structure from motion. As early as 1987 the manifestation of parallax in 2D light field slices was being explored [Bolles et al. 1987]. That work examines the characteristic straight lines arising in “epipolar images”, 2D slices of the light field in spatial and angular dimensions. These straight lines were the basis for depth estimation from lenslet-based plenoptic cameras in Adelson and Wang [2002], and similar ideas were later elaborated in general 4D light fields [Dansereau and Bruton 2004].

Similar developments often arise in disparate fields, and it is interesting that evolution itself may have stumbled upon depth estimation from parallax motion in lenticular arrays, in the form of insect compound eyes [Bitsakos and Fermüller 2006]. A year before that work was published, Neumann et al. [2005] proposed an artificial compound eye sensor for egomotion estimation, based on a spatio-temporal generalization of parallax motion. Spatial-domain light field manifolds are also discussed in more detail in Berent and Dragotti [2007] and Gu et al. [1997].

Exploiting parallax motion in light fields is not limited to depth estimation, and indeed one of its first applications was in filtering. Levoy and Hanrahan [1996] included a discussion of spatial-domain antialiasing filters, employing the properties of the light field to improve rendering quality. In this article we show that parallax motion has consequences in the frequency domain, namely that the frequency-domain region of support (ROS) of a light field is a fan-like shape which we call a hyperfan. The frequency content of light fields has been the subject of extensive research [Chai et al. 2000; Chan and Shum 2000; Durand et al. 2005; Freeman et al. 2009], with the frequency plane being a commonly identified feature. To the authors’ knowledge, the first frequency-planar light field filter was proposed by Isaksen et al. [2000], and the same idea has since reappeared with minor variations, including efficient recursive and frequency-slicing approaches for carrying out light field focus [Dansereau and Bruton 2003; Ng 2005].

Volumetric focus is a generalization of planar focus, and an example is discussed in Dansereau and Bruton [2007]. That work proposes the dual fan as the frequency-domain ROS of a light field, and employs multiple-branch filter banks to approximate the dual-fan shape. Earlier work had proposed a two-branch filter bank to approximate a fan shape, under different terminology [Stewart et al. 2003]. In the present work it is shown that the dual fan is a projection of the much more selective frequency hyperfan underlying light fields.

Levin et al. [2009] and Levin and Durand [2010] discuss the light field’s frequency-domain ROS in terms of a dimensionality gap, the idea that light field images lie on a 3D *focal manifold* in 4D frequency space. In Levin et al. [2009] the focal manifold is used to analyze a novel, physical lens design that displays extended depth of field by virtue of collecting light over many discrete focal depths. Levin and Durand [2010] employ the focal manifold in derivations of 2D deconvolution kernels for rendering from focal stacks and sparse collections of viewpoints. That same work discusses aliasing in terms of the focal manifold, and concludes by rendering wide depth-of-field images from a stack of more narrowly focused antialiased images produced using methods from Lumsdaine and Georgiev [2009].

Our work differs in specifically identifying the frequency-domain ROS of the light field as the 4D *hyperfan* shape at the intersection of a *hypercone* and a dual fan. We effect tunable, postcapture volumetric focus by surrounding the frequency hyperfan with a novel, linear, single-step and irreducibly 4D hyperfan filter. We demonstrate the frequency hyperfan to show important theoretical and practical performance gains over previously described filters in low-contrast, wide depth-of-field scenarios.

This article builds on Dansereau et al. [2013a], introducing spatial-domain and hybrid implementations, addressing aliased passband components, discussing the scaling of selectivity with sample count, and providing more extensive results. A more detailed treatment can be found in Dansereau [2014].

4. THE MANY FACES OF PARALLAX

In this section we explore the spatial- and frequency- domain behaviors of light fields, starting with parallax motion and concluding with a set of rules that, under a few assumptions, all light fields follow. In subsequent sections we design linear filters which exploit these rules to carry out volumetric focus.

4.1 Parallax in 2D

We begin by investigating the case of a single point $\mathbf{P} = [P_x, P_y, P_z]$ in an arbitrary scene, in 2D. The rays emanating from \mathbf{P} can be described using a simple set of rules. As depicted in Figure 3(a), if one begins with a ray that intersects \mathbf{P} (highlighted), then translates that ray’s point of intersection along s , its point of intersection along u must follow at a proportional rate in order for the ray to maintain its intersection with \mathbf{P} . In other words, the rays emanating from \mathbf{P} follow a linear relationship in s and u . This is the light field manifestation of parallax motion [Bolles et al. 1987; Dansereau and Bruton 2003].

We can write the linear relationship relating s and u , and its generalization in the vertical dimensions t and v , as

$$\begin{bmatrix} u \\ v \end{bmatrix} = \left(\frac{D}{P_z} \right) \begin{bmatrix} P_x - s \\ P_y - t \end{bmatrix}, \quad (1)$$

where D is the plane separation in the two-plane parameterization as depicted in Figure 2. We can visualize this relationship as shown in Figure 3(b). We label the line supporting \mathbf{P} ’s rays λ . Recall that we

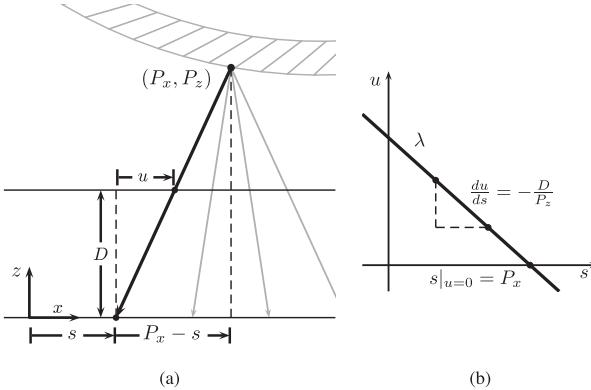


Fig. 3. Parallax in the light field: the point-plane correspondence. (a) For all rays originating at a point \mathbf{P} in space, u varies linearly with s , and by extension v with t ; (b) this describes a line λ in the 2D s, u plane and, by extension, in the t, v plane.

are operating under the *relative* two-plane parameterization; under the *absolute* two-plane parameterization a similar linear relationship will hold, but with different slopes and offsets. Notice how the slope of the line relating s and u is determined entirely by the depth of \mathbf{P} in the scene. An immediate consequence of this is that a scene containing many points at the same depth will yield parallel lines in s, u , and in t, v .

Thus far we have discussed only the support of \mathbf{P} 's rays, and said nothing of their values. In a totally unconstrained scene we can say very little. \mathbf{P} may lie on a mirrored surface, and there can be arbitrarily many occlusions within the scene, in which case the values along λ can be almost anything. Thankfully, much of the light measured in natural scenes is diffusely reflected. Trees, grass, dirt, rocks, kelp, coral, sand... just about everything occurring naturally is *primarily* diffuse except for water, as confirmed in studies measuring the bidirectional reflectance distribution functions (BRDFs) of natural materials [Dana et al. 1999]. Similarly, the energy in occlusions will generally be minimal given the limited baseline of our cameras and following arguments stemming from scene statistics [Geisler 2008; Ruderman 1997].

As such, we adopt the assumptions of a diffuse, Lambertian scene [Lambert 1760] with no occlusion, allowing us to say that the line λ corresponding to each point \mathbf{P} in the scene is constant valued. Considering the case of multiple points, we can see that the light field slices must consist of multiple, constant-valued lines. Because the orientation of a line depends only on the depth of its corresponding point, a scene consisting of surface elements at a single depth will yield light field slices of parallel, constant-valued lines.

We now consider the implications of these observations in the frequency domain. The 2D Fourier transform of a set of parallel, constant-valued lines is an orthogonal line that passes through the origin. This fact can be derived mathematically [Dansereau 2003], or understood intuitively by realizing that a function which is constant valued in a certain direction will exist as a frequency-domain delta function along that direction.

More formally, the frequency-domain ROS of the Lambertian surface at depth P_z can be described as

$$\Omega_s/\Omega_u = \Omega_t/\Omega_v = D/P_z, \quad (2)$$

where Ω is the continuous-domain light field frequency space.

Generalizing for a scene containing a range of depths is possible through superposition: A scene comprising surface elements at

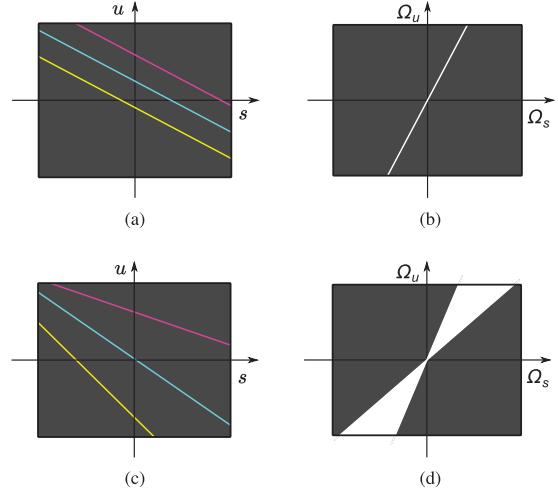


Fig. 4. The relationship between Lambertian scenes and their frequency-domain regions of support: (a) Points at a single depth, shown in s and u , correspond to (b) a 2D frequency-domain line; (c) points over a range of depths correspond to (d) a 2D frequency-domain fan.

many depths will exist as a superposition of lines in the 2D light field. This can be seen by allowing P_z in (2) to sweep through a range of depths corresponding to the scene extents,

$$Z_{MIN} < P_z < Z_{MAX}. \quad (3)$$

The resulting shape is a 2D fan [Chai et al. 2000]. The relationships between Lambertian scenes and their frequency-domain regions of support are depicted in 2D in Figure 4.

Recall that we have ignored the effects of occlusion and specular reflection. The curious reader is referred to Durand et al. [2005] for a discussion of specularly reflective surfaces and occlusions in the context of the light field, Maeno et al. [2013] for scenes with refractive objects, Ji et al. [2013] for an excellent treatment of the more complex case of refractive gas flows, and Raskar et al. [2008] for situations where the camera itself contributes complex lens flare effects.

4.2 Generalizing to 4D

We now generalize the observations made in 2D in the previous section to the 4D light field. We begin with the relationship depicted in Figure 3, that is expressed as a system of two linear equations (1). In 4D, each of these linear equations describes a *hyperplane* [Dansereau 2003], because it imposes a single linear constraint on the four dimensions. The two hyperplanes described by (1) are depicted as 2D slices of 2D images in Figures 5(a) and (b). This visualization of the 4D light field as an array of slices is akin to tiling the images captured by the cameras of an array. In this case the light field is sliced as an array of t, v slices arranged according to their s, u positions. Notice the compact convention we follow in labeling these axes.

Applying both Eqs. (1) simultaneously results in an *intersection* of the two hyperplanes. The situation is closely analogous to the intersection, in 3D, of two planes: Each plane is described by a single linear equation, and the combination of the two equations is the line where the two planes intersect. In the same way, our two linear equations describe two hyperplanes that intersect to form a plane in 4D space, as depicted in Figure 5(c). The consequence of

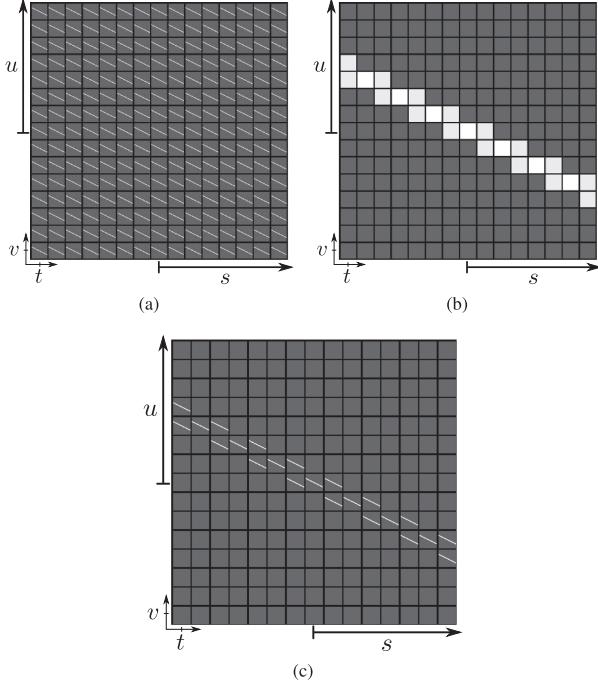


Fig. 5. Two 4D hyperplanes (a) and (b) intersect to form a plane (c).

these observations is that a point in space, P , corresponds to a plane in the 4D light field.

In 2D, we saw that a Lambertian surface at a single depth has a linear 2D frequency-domain ROS. Generalizing this to 4D follows exactly the same procedure as before. Each 2D linear ROS corresponds to a 4D *frequency hyperplane*, and the simultaneous application of the two hyperplanes intersects to form a 4D plane. The result is that a Lambertian surface at a single depth has a 4D *frequency-planar* ROS. Figure 6(a) depicts three points at a single depth in a scene, and (b) depicts the corresponding 4D frequency-domain ROS.

Generalizing to multiple depths must be performed in 4D. Simply applying the fan-shaped ROS depicted in Figure 4(d) in both \$s\$, \$u\$ and \$t\$, \$v\$ dimensions yields a dual-fan *volume* [Dansereau and Bruton 2007], while the true shape of the light field's ROS, we shall see, is a 3D *manifold* embedded in 4D space. This has strong parallels to the 3D example of attempting to describe the surface of a cone as the intersection of a circle and two triangles, yielding a family of shapes that are not generally cones, and most of which are volumes, not surfaces.

We begin instead from spatial points at different depths as visualized in 4D. Figure 6(c) depicts three points at different depths, and their corresponding ROS is depicted in (d). The latter is the superposition of planes like the one in (b) at different orientations. We denote this new manifold the *hyperfan* because it is constructed by sweeping a plane through a range of angles, akin to sweeping a line through 2D space to form a fan.

A more mathematically driven approach considers (2) and (3) together, resulting in *three* constraints describing the frequency-domain ROS of the light field:

$$m_{MIN} < \Omega_s / \Omega_u < m_{MAX}, \quad (4)$$

$$m_{MIN} < \Omega_t / \Omega_v < m_{MAX}, \quad (5)$$

$$\Omega_s / \Omega_u = \Omega_t / \Omega_v. \quad (6)$$

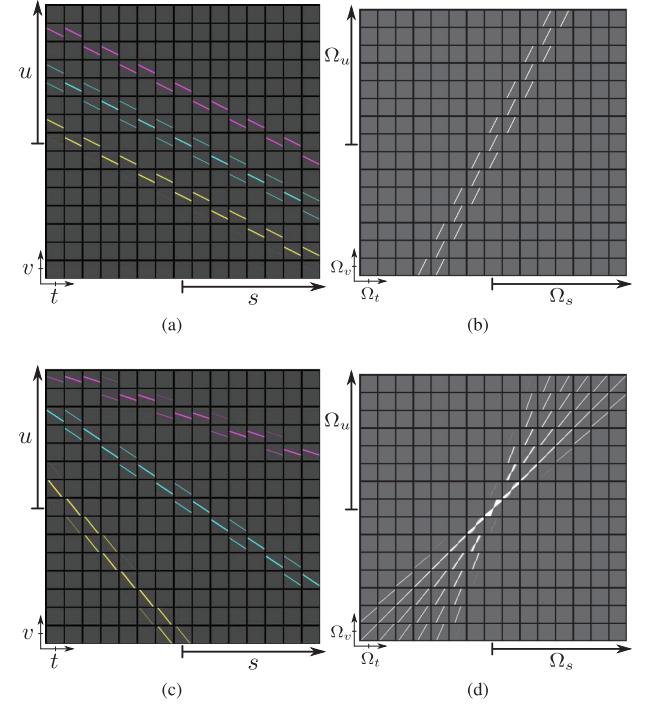


Fig. 6. Deriving the frequency-domain ROS of the light field in 4D: Points at a single depth (a) have a frequency-planar ROS (b), while points over a range of depths (c) have an ROS that is a superposition of planes at different orientations (d). We denote this sweep of planes a *hyperfan*.

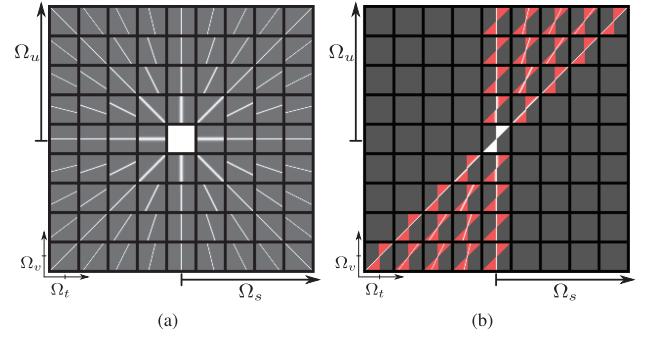


Fig. 7. Decomposing the hyperfan into (a) the 4D frequency hypercone (6), which constrains slopes in two pairs of dimensions, and (b) the dual fan (4), (5), shown in red. The shape at their intersection, shown in white in (b), is the hyperfan.

The first two constraints, (4) and (5), describe the dual fan [Dansereau and Bruton 2007]. We shall demonstrate in the following section that the third constraint (6) describes a *hypercone*. The hypercone is depicted on its own in Figure 7(a), and in 7(b) the dual fan is depicted in red while the intersection of the two, the hyperfan, is shown in white.

The hypercone restricts two pairs of slopes to be equal in the frequency domain. The physical interpretation of this constraint is that an object's apparent motion in the horizontal light field dimensions \$s\$ and \$u\$ should equal its apparent motion in the vertical directions \$t\$ and \$v\$. Recall that the slope of the line \$\lambda\$ supporting a point depends on the depth of the point in the scene, \$P_z\$. It makes

sense that, regardless of the value of this slope, it should be equal in horizontal and vertical directions. Noise will not in general follow this rule, and so the hypercone shape gives us a high degree of selectivity against noise.

The dual fan imposes depth limits on the scene by constraining the range of valid slopes. In the following sections we will construct a volumetric focus filter by combining the depth selectivity of the dual fan and the noise rejection of the hypercone.

4.3 Hyperfans and Hypercones

To see why (6) describes a hypercone, we begin with the standard form

$$R_s^2 + R_u^2 - R_t^2 - R_v^2 = 0, \quad (7)$$

which describes a 4D saddle or hyperbolic cone—this differs from the 4D spherical cone in the sign of the third term. To show equivalence with (6), we transform the coordinate axes by applying rotations of $-\pi/4$ in the Ω_s , Ω_v and Ω_t , Ω_u planes, yielding

$$\begin{bmatrix} R_s \\ R_t \\ R_u \\ R_v \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \Omega_s + \Omega_v \\ \Omega_t + \Omega_u \\ \Omega_t - \Omega_u \\ \Omega_s - \Omega_v \end{bmatrix}. \quad (8)$$

Substituting the rotated coordinates into (7) and simplifying yields the form shown in (6), thus the two forms are rotated views of the same shape. The rotated form of the hypercone (7) is depicted in Figure 8(a), alongside some other rotations of the same shape.

We have made much in this article of the distinction between the dual fan and the hyperfan. As we shall see, the difference made by treating the hyperfan as an inseparable 4D shape is significant, especially in regard to improving SNR in low-contrast applications.

5. THE 4D HYPERFAN FILTER

Having described a frequency-domain ROS for the light field, we proceed to design a linear filter that selectively passes it. We begin by implementing the filter in the frequency domain, computing the input's discrete Fourier transform (DFT), multiplying by the filter's magnitude response in the frequency domain, and then computing the inverse DFT. We explore spatial-domain implementation in the following section. Note that we describe the filter in terms of the continuous-domain frequency space Ω , and that practical implementation requires appropriate adjustment of filter parameters to reflect the sample rate of the discrete light field [Dansereau 2003].

Because the frequency hyperfan lies at the intersection of a dual fan and a hypercone as depicted in Figure 7, one way forward is to describe each of these passbands and take their product. As we proceed we will evaluate the theoretical selectivity of each passband as the fractional 4D Nyquist volume that it passes, with smaller fractions corresponding to higher selectivity.

Starting with the dual-fan passband, we note that this is itself the product of two 2D fan filters [Dansereau and Bruton 2007]

$$H_{DF}(\Omega) = H_{FAN}^{2D}(\Omega_s, \Omega_u, \theta_1, \theta_2) H_{FAN}^{2D}(\Omega_t, \Omega_v, \theta_1, \theta_2), \quad (9)$$

where each 2D fan is implemented by passing all points within the prescribed angular range θ_1 to θ_2 . The 2D fan filter and the process of selecting θ values for a desired depth range are described in Ansari [1987] and Dansereau and Bruton [2007].

The fractional 2D area passed by each 2D fan has a lower bound α_{DF} determined by the range $[\theta_1, \theta_2]$. We apply Gaussian smoothing to reduce ringing artifacts, surrounding the fan by a tunable bandwidth and increasing the passband area by β_{DF} . Because the same

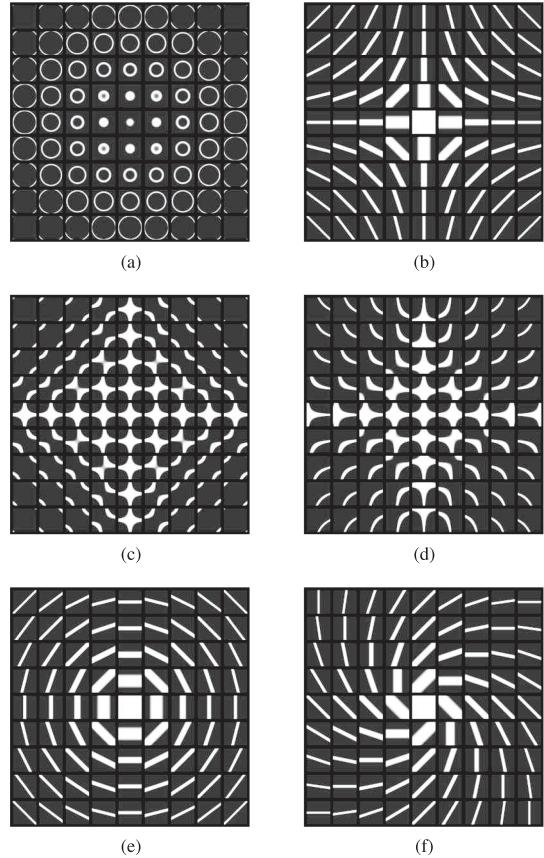


Fig. 8. Visualizing the 4D hypercone does not come naturally but, by inspecting tilings under a variety of rotations, we can construct an intuition for its nature. (a) When rotated as in (6) circles are revealed which grow with distance from the center, highlighting the shape's cone-like nature; (b) this rotation elicits the contour lines of a saddle shape; (c)–(f) further reveal the complex beauty of this shape, at turns eliciting circles, spirals, saddles, and crosses.

selectivity is applied in Ω_s , Ω_u and in Ω_t , Ω_v , the fractional volume passed by the 4D dual fan is given by the square

$$V_{DF} = (\alpha_{DF} + \beta_{DF})^2. \quad (10)$$

The ideal hypercone (6) is a 3D manifold, not a 4D volume, and so practical implementation requires surrounding the hypercone by a bandwidth β_{HC} . We propose the filter with magnitude response

$$H_{HC}(\Omega) = \exp \left(- \left[\frac{(\Omega_s \Omega_v - \Omega_t \Omega_u)}{\beta_{HC}^2 / \sqrt{2 \ln 2}} \right]^2 \right), \quad (11)$$

where β_{HC} is the 3-dB bandwidth measured as the radius of the hypercone at the origin; this is the radius of the cone in the rotated R_s , R_u and R_t , R_v planes. The magnitude of the numerator of the exponential increases with distance from the ideal hypercone shape, and so the filter rolls off in a Gaussian-like manner from the ideal passband. Note that the filter offers no selectivity near the origin, but this is consistent given that the underlying constraint (6) provides no information to do so.

For analysis we begin by ignoring the Gaussian rolloff, approximating the hypercone filter as having constant thickness related to the 3-dB bandwidth β_{HC} through a constant factor κ . Examining

Figure 7(a), this implies every Ω_t, Ω_v slice, with the exception of the origin, will pass a constant fraction of its area. Including the effect of the Gaussian rolloff increases the total admitted volume by another constant factor which we absorb into κ , for a fractional volume passed by the hypercone, given by

$$V_{HC} = \kappa \beta_{HC}. \quad (12)$$

The hyperfan filter is simply the product of the hypercone and dual fan

$$H_{HF} = H_{HC} H_{DF}. \quad (13)$$

Referring to Figure 7(b), we notice that every nonzero Ω_t, Ω_v slice of the hyperfan will pass a mean area of $\kappa \beta_{HC}$, and from the dual fan $\alpha_{DF} + \beta_{DF}$ describes the ratio of nonzero slices. The fractional volume passed by the hyperfan filter is therefore the product

$$V_{HF} = \kappa \beta_{HC} (\alpha_{DF} + \beta_{DF}). \quad (14)$$

Notice the minimum volume passed by the dual fan is α_{DF}^2 , while the minimum for the hyperfan is zero, that is, the hyperfan offers direct control, via β_{HC} , of the total signal energy passed, and therefore presents significantly greater selectivity than the equivalent dual-fan filter. Note also that both the dual-fan and hyperfan filters degenerate gracefully to frequency-planar filters as their depth ranges approach zero.

5.1 Scaling Selectivity

Here we examine the change in filter selectivity as the number of light field samples increases. The trivial example of a planar filter focused at a single depth $P_z = D$ has a maximally selective passband given by the plane $\omega_s = \omega_t = 0$ [Ng 2005]. In a light field having $N = [N_s, N_t, N_u, N_v]$ samples, this yields $N_u \times N_v$ nonzero frequency-domain passband entries, and so the selectivity of this filter can be written as $N_u N_v / (N_s N_t N_u N_v) = 1 / (N_s N_t)$. Increasing the number of samples in both s and t by a factor of M results in an increase in selectivity of M^2 [Ng et al. 2005].

Extending this approach to the hypercone, we note that its passband grows with all dimensions. Because each 2D slice of the light field contains a 1D line, the selectivity can be written as $\sqrt{N_s N_t} N_u N_v / (N_s N_t N_u N_v) = 1 / \sqrt{N_s N_t}$. The square root $\sqrt{N_s N_t}$ reflects the mean area covered within each 2D slice. Increasing the number of samples in both s and t by a factor of M now results in an increase in selectivity of M , rather than M^2 , for the planar filter.

The hyperfan filter narrows the hypercone to a tuneable range of depths. At one extreme, it presents very little depth selectivity, resulting in a filter that behaves much like the hypercone, and at the other extreme it focuses on a very narrow range of depths, yielding a behavior similar to the planar filter. In most cases, the hyperfan occupies the space in between, offering selectivity between $1 / \sqrt{N_s N_t}$ and $1 / (N_s N_t)$.

The preceding discussion assumes an increase in sample count by adding additional spatial samples at the same sample rate, for example, when growing a camera array without changing the camera spacing. If one instead increases the sample rate by packing cameras more tightly into the same space, or by increasing the number of pixels per lenslet in a lenslet-based camera, an additional effect must be considered: Under these conditions, the angle subtended by the hyperfan decreases for a given depth of field. This is a consequence of frequency scaling in ω_s and ω_t , and allows a higher selectivity because the narrower fan more closely resembles the planar case. Conversely, if the sample rate in u and v is increased, the hyperfan angles must be widened for a given depth of field, reducing selectivity.

As a concrete example, for a 4×4 camera array a planar filter has a maximum selectivity of $1 / (N_s N_t) = 1 / 16$, a hypercone filter $1 / \sqrt{N_s N_t} = 1 / 4$, and a hyperfan filter varies between these extents as a function of the selected depth of field and the u, v resolution of the cameras.

5.2 Including Aliased Components

Arrays of cameras have discontinuous sampling patterns in the s, t plane due to gaps between apertures. As in the case of resampling a 2D image at discrete points, this yields aliasing. However, unlike the example of resampling an image, we do not have the possibility of applying an antialiasing filter prior to the sampling operation, and camera array light fields consequently feature aliasing in the s and t dimensions. This aliasing varies with depth and is most pronounced in those elements forming the greatest slopes in s, u and t, v .

Aliased components are sometimes desirable; indeed, u, v slices of camera array light fields generally contain important edge detail that is aliased in s, t . Because they occur in a predictable manner, it is possible to extend our definition for the light field's ROS, and our filters' passbands, to include these aliased components. In the Results section we demonstrate this approach passes desired edge detail that would otherwise be destroyed. An evident drawback of including aliased components in this manner is a decrease in selectivity.

A simple way to incorporate aliased components into a filter is to construct a magnitude response H over an extended domain that incorporates an integer multiple of the desired domain size in ω_s, ω_t . Such an extended magnitude response is depicted in 2D in Figure 9(a), for which the domain in ω_s is extended to include two additional bands. The magnitude response is then collapsed to the desired size by taking the maximum value over the bands, as depicted in Figure 9(b). Extension of this method to 4D is straightforward.

5.3 Memory and Complexity

If we implement the hyperfan filter in the frequency domain, the filtering process is one of applying a discrete Fourier transform, its inverse, and a per-sample complex multiplication. Computation time for an N -sample light field is therefore constant and of complexity $O(N \log N)$ when using the fast Fourier transform (FFT).

We operate on the three color channels separately, and so the memory requirement is for a single color channel at a time. Two buffers are required beyond the input light field buffer: the filter magnitude buffer, and a complex buffer to contain the DFT. The input light field comprises 8-bit integers, but for simplicity our implementation operates on single- or double-precision floats. For a color light field of N samples total, our total additional memory requirement, for double precision, is

$$M = (8 + 16)N/3 = 8N. \quad (15)$$

In practical terms, the $128 \times 128 \times 17 \times 17$ light fields shown in the Results section occupy approximately $N = 14$ MBytes. Filtering required an $8N/3 = 38$ MByte double-precision buffer to hold the filter's magnitude response, and a $16N/3 = 76$ MByte complex double-precision buffer to hold the DFT of the input, for a total of $8N = 114$ MBytes. The single-precision implementation requires half the memory.

For the full-resolution Stanford Archive light fields, for example, the $1024 \times 1024 \times 17 \times 17$ -sample Tarot light fields, the input buffer itself occupies 909 MBytes, and the additional memory requirements associated with a double-precision filter are 7272 MBytes. Most modern computers have sufficient memory to support such an

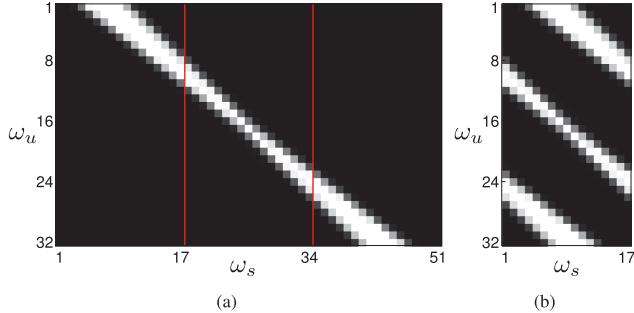


Fig. 9. Including aliased components in the filter passband: (a) The magnitude response is extended over two additional bands as delineated by the red lines and (b) reduced to the desired size by taking the maximum value over the bands.

operation, but in lightweight mobile applications a more memory-efficient spatial-domain implementation may be desirable.

5.4 Spatial-Domain Implementation

For very large light fields, for example, the full-resolution versions of the Stanford Archive light fields, directly computing the full 4D DFT may be prohibitively memory intensive on smaller systems. For this reason, a spatial-domain filter implementation may be desirable. By constructing a spatial-domain finite impulse response (FIR) filter with impulse response $h(i, j, k, l)$, we can compute the output light field a single pixel at a time. The key advantage of this is lower memory utilization: The output buffer need not be the full light field size if only a subset of the output is needed. This would be the case, for example, when only a 2D subset of the output light field is required. Furthermore, the filter buffer—in this case the impulse response h —will not in general be as large a structure as the full light field L . The total memory utilization of a spatial implementation will therefore be much lower than for a frequency-domain implementation.

As a concrete example, for the $1024 \times 1024 \times 17 \times 17$ -sample 3-channel Tarot light fields, rendering a single 2D output image requires only a $1024 \times 1024 \times 3$ -sample output buffer, plus a buffer to store the impulse response h , which, as we shall show, can be quite modest, between 1 and 16MBytes. As such, the total memory requirement for the spatial implementation is 20MBytes or lower, a significant improvement over the 7272MBytes required by the DFT-based implementation.

Where spatial implementation suffers is in processing time¹. Convolution over millions of samples is much more complex than Fourier-based multiplicative filtering. If, however, only a 2D output slice is required, the spatial convolution method can outpace the frequency-domain implementation because the latter treats the entire signal during the DFT, while the former can focus on those parts of the light field required for the 2D output. The filter appropriate to a given application will therefore depend on the nature of the desired output, the size of the input, and memory availability.

¹This observation applies mostly to general-purpose computing. Though the total operation count may be higher, the highly parallel nature of spatial implementations can make them better suited to parallel architectures, leading to significantly faster runtimes on specialized hardware such as graphics processing units (GPUs), field programmable gate arrays (FPGAs) and application-specific integrated circuits (ASICs).

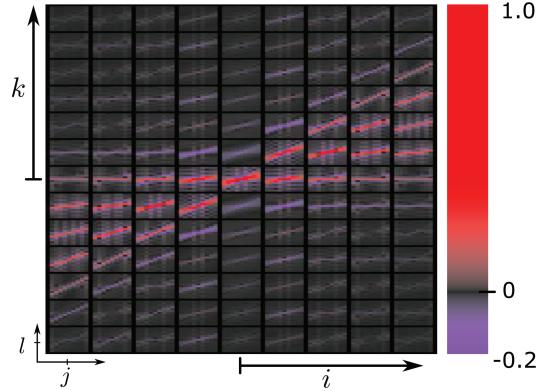


Fig. 10. A typical hyperfan filter impulse response. This example is for a $9 \times 9 \times 13 \times 13$ filter passing a range of slopes between 0 and 1. The overall shape resembles a superposition of planar filters, but with the inclusion of orthogonal highpass components that appear as ringing.

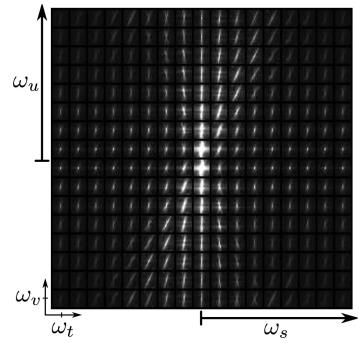


Fig. 11. The maximum magnitude per frequency component over the first six Stanford light fields, showing a characteristic hyperfan shape. Compare with Figure 7(b).

A key factor allowing us to constrain the size of the impulse response h is the range of parallax motion typical of real-world light fields. Apparent motion is usually restricted to a small fraction of the total u, v plane, for the simple reason that it is impractical to design a camera otherwise. Even arrays of cameras with relatively large baselines are seldom designed to display more apparent motion than a fraction of the u, v plane, as so doing would yield excessive aliasing.

The size of the impulse response required for a given volumetric focus task is directly related to the slopes that it must support. If the desired depth range projects at most to an apparent motion of 10 pixels, then the resulting impulse response will not need to be more than 10 pixels wide in u and v . In general we assume that the whole s, t range is to be covered, as so doing maximizes selectivity, and we select the impulse response's size in u and v to conservatively include the maximum apparent motion we might want to include in the passband.

Having chosen a size for the impulse response, we proceed to build the appropriately sized hyperfan in the frequency domain as in the frequency-domain implementation, then take its inverse DFT. To avoid windowing artifacts, we pad the frequency-domain shape to a larger size; for the Stanford light fields, we pad to a hypercube of size 32 or 64 samples in each dimension.

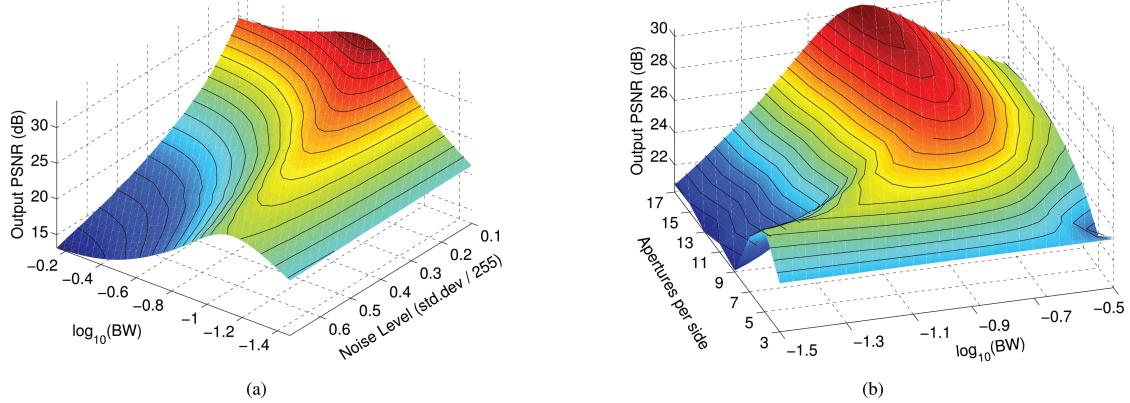


Fig. 12. Optimal bandwidth shifting with (a) noise level and (b) aperture count.

A typical impulse response h is shown in Figure 10. Hyperfan impulse responses typically have many samples with very low magnitudes. As such, a simple optimization discards low-magnitude samples, effectively speeding convolution. The number of samples to retain in the impulse response can be exposed as a parameter of the filter, and we will show in the Results section that less than 5% of the samples are typically required for high-quality results.

5.5 Hybrid Implementations

We have seen that the frequency-domain implementation is memory intensive but fast, while the spatial-domain method employs less memory at the cost of slower performance. Overlap-add and overlap-save methods are a well-established means of balancing memory utilization and speed [Rabiner and Gold 1975]. As an example, for the $1024 \times 1024 \times 17 \times 17$ -sample three-channel Tarot light fields, the FFT can be applied to blocks of size $32 \times 32 \times 128 \times 128$, with results saved into a strip of size $17 \times 17 \times 128 \times 1056 \times 3$ that cycles back onto the input buffer once a region of the input is no longer needed. The additional memory requirement over the input buffer, including the $32 \times 32 \times 128 \times 128$ magnitude response, is 603MBbytes for single-precision buffers. This is again a significant savings over the DFT-based implementation, but less so than for the purely spatial-domain implementation. The speed of this method lies between the frequency- and spatial-domain methods.

In the experiments that follow, we demonstrate spatial-, hybrid- and frequency-domain implementations. For smaller light fields such as those captured by the Lytro, we employ the frequency-domain implementation as it is practical and efficient. Downsampled versions of the Stanford light fields are also processed in this manner. For larger light fields such as full-sized Stanford light fields, the memory requirements of the frequency-domain implementation made it less desirable, and so the more memory-efficient spatial-domain and hybrid implementations are employed; note that these two methods produce practically identical output.

6. EXPERIMENTS: STANFORD LIGHT FIELDS

The Stanford Light Field Archive² is a publicly accessible database suitable for evaluating light field filtering techniques. The 12 light fields we utilize, listed in the legend of Figure 17, all contain 17×17

²<http://lightfield.stanford.edu/>.

aperture positions in s, t . Aperture positions are close enough to an ideal grid that ignoring the deviation results in negligible degradation to output quality. Each image in s, t is rectified in u, v , and the light fields are in the two-plane parameterization. Light field geometry varies across the dataset: Grid spacing is not identical, plane separation varies, and image aspect and resolution vary, meaning fan extents θ need to be tuned on a per-light-field basis. An alternative would have been to convert the light fields to a uniform relative two-plane parameterization and use generic fan extents.

In noise rejection experiments, the Stanford light fields were generally downsampled to a maximum u, v size of 128×128 pixels to reduce memory requirements, though full-resolution spatial-domain and hybrid results were also evaluated. Numeric results are for monochrome versions of the light fields. When an experiment calls for fewer than 17×17 apertures, we discard apertures at the edge of the light field, retaining the central portion. For consistency across experiments for which aperture counts can vary, metrics report on the central image in s, t .

Displayed results are produced by taking a 2D slice of the 4D filtered light field at the center of s, t . In this sense the filter is acting as a rendering algorithm, though it only renders at the native u, v resolution of the light field. As future work we consider extending the filter to also perform interpolation to greater resolutions. This work in no way precludes the application of existing 4D-to-2D plenoptic rendering methods [Lumsdaine and Georgiev 2008; Bishop and Favaro 2012; Wanner and Goldluecke 2013].

In the following section, further validation of the hyperfan filter is carried out on imagery collected using a commercially available Lytro lenslet-based light field camera. This imagery includes low-light and turbid media examples. The raw lenslet images are decoded to a 9×9 array of images, each having 380×380 pixels, following Dansereau et al. [2013b]. Compared with the 17×17 images of the Stanford light fields, we expect significantly less selectivity. However, there is still a potential 81-fold redundancy in the imagery (practically slightly less due to lenslet vignetting), allowing significant noise rejection to be demonstrated.

As empirical evidence of the frequency-hyperfan ROS of light fields, we computed the DFT of the first six of the 12 Stanford light fields, scaled to a common size, and selected the maximum magnitude at each frequency. The result, shown in Figure 11, establishes the bounds of the light fields in frequency space; the hyperfan shape is clearly evident. Note that this is true despite the varying light field geometries and the presence of occlusions, non-Lambertian surfaces, and aliasing.

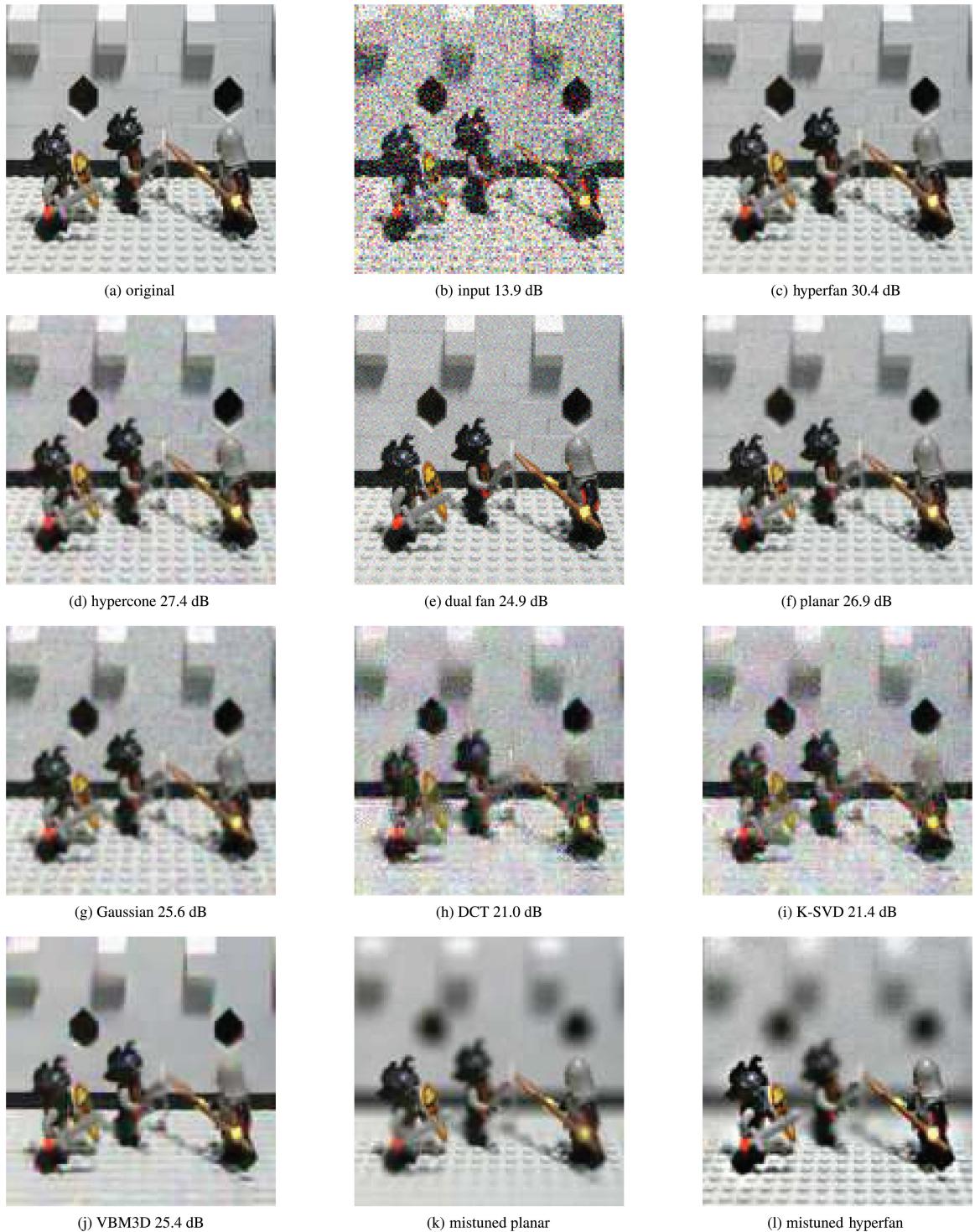


Fig. 13. Filtering results for the Stanford “Lego Knights” light field: (a) The original scene; (b) with additive white Gaussian noise; (c)–(i) show filter outputs; the depth-tunable results (c), (e), and (f) are at the PSNR-optimal balance between noise rejection and reduction in depth of field; the effects of mistuning are exaggerated in (k),(l); the hyperfan output is visually superior, with the nonlinear methods providing the most jarring artifacts, the Gaussian and planar reducing edge content, and the dual fan and hypercone being less selective to noise. Original light field courtesy of the Stanford Computer Graphics Laboratory.



Fig. 14. Filtering the “Tarot Coarse” light field for synthetic noise based on a camera model including quantization, Poisson, Gaussian, and salt-and-pepper noise. (a) The original light field; (b) the low-light image prior to salt-and-pepper noise and gain control; (c) the gain-adjusted input including salt-and-pepper noise, and (d)–(k) the filter outputs; light refracting through the crystal ball violates the depth constraints, leading to attenuation of that content and a lower PSNR for depth-selective filters (d), (f), and (g); the hyperfan nevertheless arguably provides the most visually appealing result. Original light field courtesy of the Stanford Computer Graphics Laboratory.

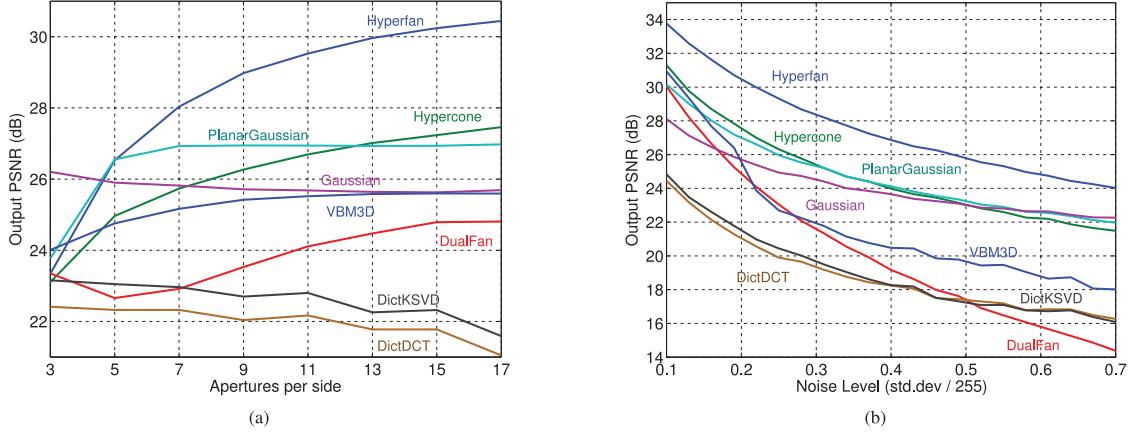


Fig. 15. Performance of the evaluated methods versus (a) aperture count; (b) noise level. The hyperfan generally shows the best performance.

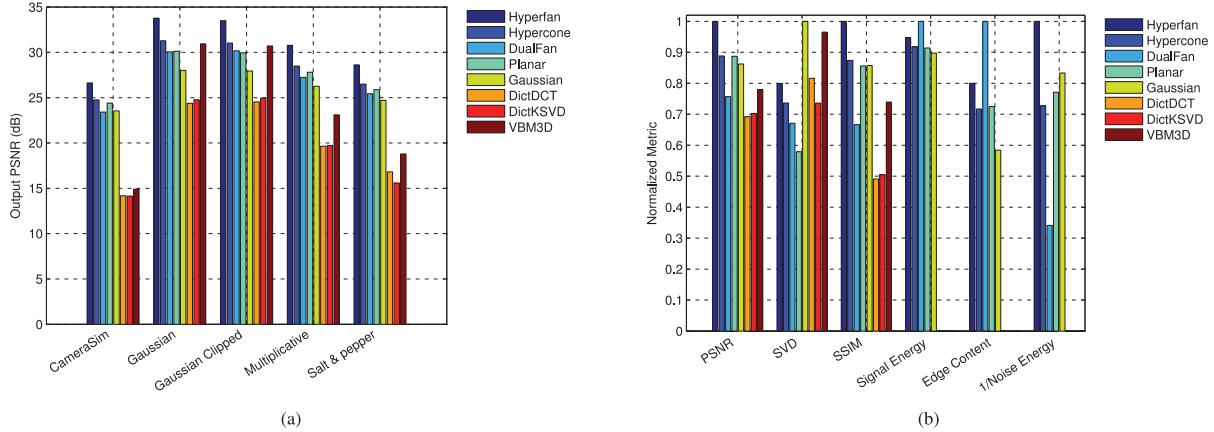


Fig. 16. Performance of the evaluated methods versus (a) noise type; (b) metric. The hyperfan generally shows the best performance.

6.1 The Methods

We test a range of linear filters on the Stanford light fields, including the three described in this article: the hyperfan (13), the hypercone (11), and the dual fan (9). If our earlier assertions are correct, the hyperfan will be the most selective of these, although how the hypercone alone behaves will also prove interesting.

We further test a 4D Gaussian filter as well as a 4D planar Gaussian filter that is the basis for synthetic refocusing of light fields [Dansereau and Bruton 2003; Ng et al. 2005; Ng 2005]. Dictionary-based image denoising approaches do not exploit the structure of the light field; nevertheless, by collapsing the light field into a tiling of images, we test the overcomplete discrete cosine transform (DCT) [Guleryuz 2007] and K-SVD methods [Elad and Aharon 2006; Aharon et al. 2006]. Finally, we test the block-matching and filtering approach V-BM3D [Dabov et al. 2007] by applying it over sequences of frames constructed along the s, t dimensions.

6.2 Tuning

The hyperfan has four tunable parameters: the two depth limits and filter rolloff associated with the dual-fan filter, and the bandwidth associated with the hypercone. The optimal values for these de-

pend on the range of depths occupied by the scene, the number of apertures in the light field, the noise level, and the light field parameterization.

If no prior knowledge of scene depth is available, a great deal of selectivity is nevertheless possible, as the valid range of plane angles present in any light field is limited [Levin et al. 2009]. In the relative two-plane parameterization, for example, all planes must lie within the first and third quadrants in Ω_s, Ω_u and Ω_t, Ω_v , that is, the plane angles are restricted to a 90° range. This observation allows the fan limits to be pretuned for generic scenes, leaving only the hypercone bandwidth to be tuned. Of course, knowledge of a more selective depth range allows for more aggressive filtering.

For fixed fan angles and selectivity, Figure 12 demonstrates the dependence of the optimal hypercone bandwidth on input noise level and aperture count. We leave derivation of closed-form expressions for these optima as future work; the following results are for filters tuned to their PSNR-optimal bandwidths and depth limits through exhaustive search.

6.3 Evaluation

Figures 13 and 14 are typical of the output from each filter—numerical results are the peak signal-to-noise ratio (PSNR), assuming the uncorrupted input to be ideal. Figure 13 introduces

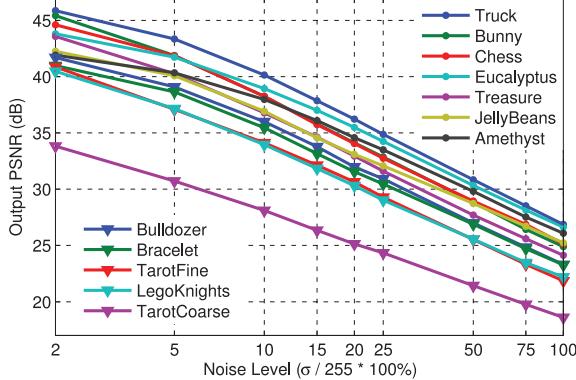


Fig. 17. Hyperfan filter output PSNR (dB) over a range of noise levels for the Stanford Archive.

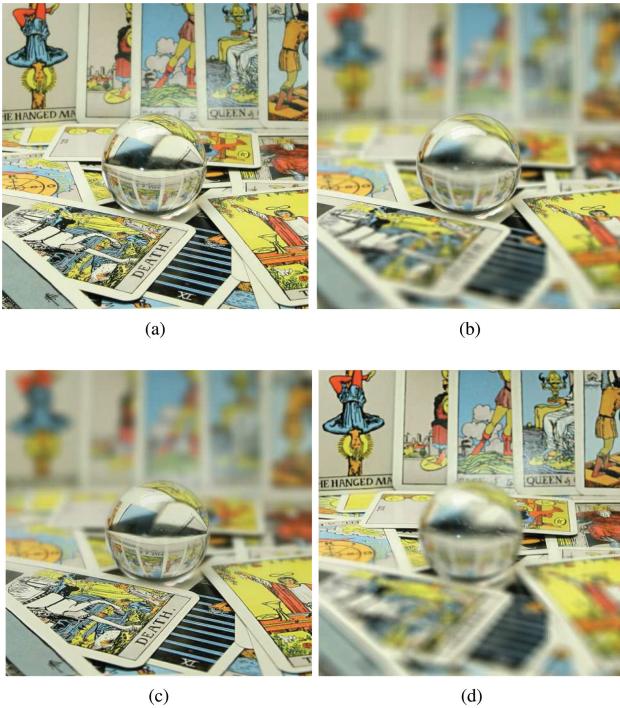


Fig. 18. Examples of volumetric focus applied using a spatial-domain filter implementation. Only the pixels shown in these 2D slices of the 4D output light field were computed, saving significant processing time and memory. (a) A slice of the input light field; (b) filtered with a narrow depth of field centered on the crystal ball; (c) filtered with a wide depth of field containing elements near the camera including the ball; (d) filtered with a wide depth of field containing elements farther from the camera and excluding the ball. Notice that the image within the crystal ball behaves similarly to foreground scene elements and, as such, passes most clearly in (c). Original light field courtesy of the Stanford Computer Graphics Laboratory.

additive Gaussian noise to the light field, while Figure 14 introduces a model of low-light camera noise, including quantization to 32 levels, intensity-dependent Poisson noise, additive Gaussian noise ($\sigma = 5\%$ maximum pixel value), and salt and pepper noise (5% density).



Fig. 19. Example of a multiple-passband filter constructed as the superposition of two hyperfans. Here only a volume surrounding the crystal ball is left out of the focal volume. Notice how the crystal ball content is nevertheless left clear, as it behaves similarly to objects closer to the camera and, as such, does not conform to the parallax motion of the stop-band signal. Original light field courtesy of the Stanford Computer Graphics Laboratory.

Visually, the hyperfan outperforms the other filters in all cases, though this will not always be true: Scene elements that violate the underlying assumptions of Lambertian and nonoccluding scenes will not generally conform to the hyperfan passband, and so the filter will attenuate these elements. If a scene were dominated by such elements, the filter could perform poorly. Note, for example, the severely attenuated crystal ball content in Figure 14(d) which has resulted in a decreased PSNR. Because the content being refracted through the ball takes on apparent motion matching scene elements close to the camera and outside the passband range, it has been attenuated. These limitations are not always so jarring; the specular highlights on the Lego knights' helmets are mostly retained, for example, while the noise is mostly rejected. Furthermore, some applications can actually benefit from removal of non-Lambertian and occluding energy, for example, geometric reconstruction and visual odometry.

Figures 15 and 16 show each method's performance for the "Lego Knights" light field over a range of aperture counts, for a variety of noise types, over a range of input noise levels, and evaluated with a range of metrics. Note that the hyperfan outperforms the others for aperture counts of five or more and continues to improve significantly with aperture count (note the logarithmic vertical scale), confirming the scalability of the approach.

The metrics depicted in Figure 16(d) are normalized to a maximum value of one. These represent the mean result over 21 levels of additive Gaussian noise with $\sigma = 10\%$ to 70% maximum pixel value. The first three metrics are, in order: PSNR, an SVD-based similarity measure [Shnayderman et al. 2006], and a structural similarity measure SSIM [Wang et al. 2004]. The remaining three metrics apply only to linear methods and linear noise, as they rely on separating the filter's treatment of noise and signal. By filtering the original image and the noise alone, the attenuation to each can be evaluated separately. Shown, in order, are the energy remaining when filtering the original image, the edge content of that filtered image measured as the mean magnitude of the first derivative of the image, and the inverse of the energy remaining in the filtered noise signal. Because of normalization, the best performance for all metrics is one.

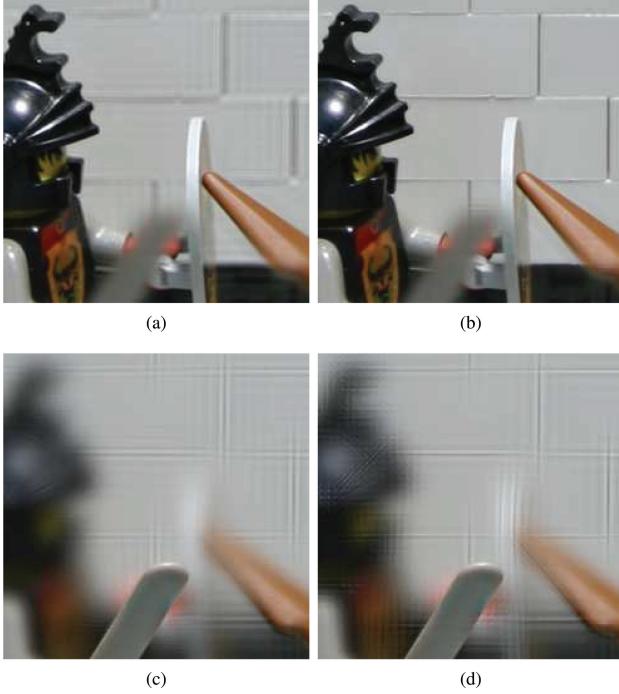


Fig. 20. Hyperfan filter outputs without (left) and with (right) inclusion of aliased components, for filters passing distant (top) and nearby (bottom) volumes. Note a significant sharpening of distant passband elements in the top-right figure due to inclusion of the aliased components. These same components masquerade as passband elements in the bottom figures, appearing as high-frequency lines. Original light field courtesy of the Stanford Computer Graphics Laboratory.

Inspecting the metric results, the humble Gaussian filter takes on a prominent position in the first three metrics, even taking the lead for the SVD metric. Note, however, that the Gaussian also attenuates the most edge content. All linear methods are similar in passing signal energy, and the dual fan outperforms the hyperfan in edge content, although it also does a poor job of attenuating noise energy, thus its weak PSNR. The nonlinear methods do well according to the SVD but a visual analysis shows that the artifacts they introduce are jarring to the human visual system. On the whole, the hyperfan attenuates the most noise energy while passing the second-to-best edge content, surpassed in this respect only by the poorly selective dual fan. The hyperfan also dominates in structural similarity and PSNR, outperformed by its nonlinear counterparts only in the SVD metric.

Drawing on the variety of light fields available in Stanford's archive, Figure 17 shows the hyperfan's performance over a range of inputs. Notice the proportional falloff in output PSNR as input noise increases. The output quality throughout these results is high and consistent, despite the varying presence of occlusions, specular reflections, and refractions in the light fields—all phenomena that break the assumptions behind the filter. The weakest performance is for “Tarot Coarse”, which we attribute to refraction in the scene as seen in Figure 14(d).

6.4 Spatial-Domain Implementation

We employed the spatial-domain implementation described in Section 5.4 to demonstrate volumetric focus on the full-resolution

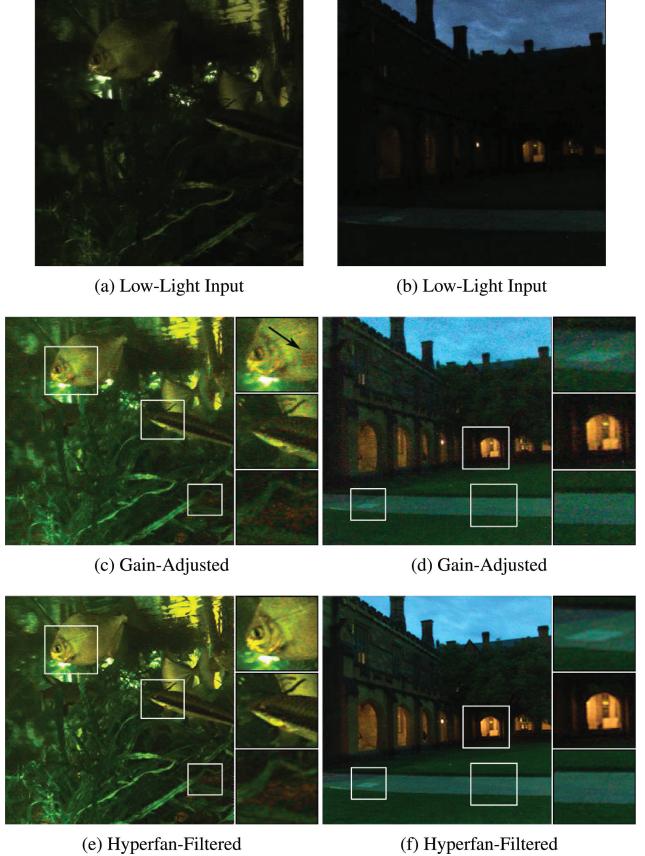


Fig. 21. Filtering low-light imagery from a Lytro consumer-grade light field camera: (a,b) Low-contrast input, (c,d) gain-adjusted images, and (e,f) filter output, showing a visible improvement in SNR. The filtered results demonstrate both noise rejection and depth selectivity, with specks of dirt on the side of the aquarium being attenuated based on depth – one such dirt speck is indicated by the black arrow.

Stanford Archive light fields. Examples are shown in Figure 18. We found that the number of nonzero impulse response samples required to obtain high-quality results varies with the depth of field of the passband signal. The narrow-passband filter employed to generate Figure 18(b) was well approximated with 2,000 impulse response entries, while the wider depth-of-field examples (c) and (d) required 40,000 samples.

It is possible to synthesize interesting filters by combining multiple hyperfans. This can be accomplished by taking the maximum magnitude response across two or more filters, for example, yielding a single-step linear filter with a complex passband. Figure 19 shows the result of including most of the tarot scene in the passband, with the exception of a narrow volume surrounding the crystal ball. Such disconnected focal regions cannot be obtained using conventional cameras, and might be useful in removing occluding interference from scenes comprising objects of interest at multiple depths, for example.

6.5 Aliased Components

Some of the Stanford light fields contain substantial aliasing in s and t . Figure 20 demonstrates the impact of including these aliased

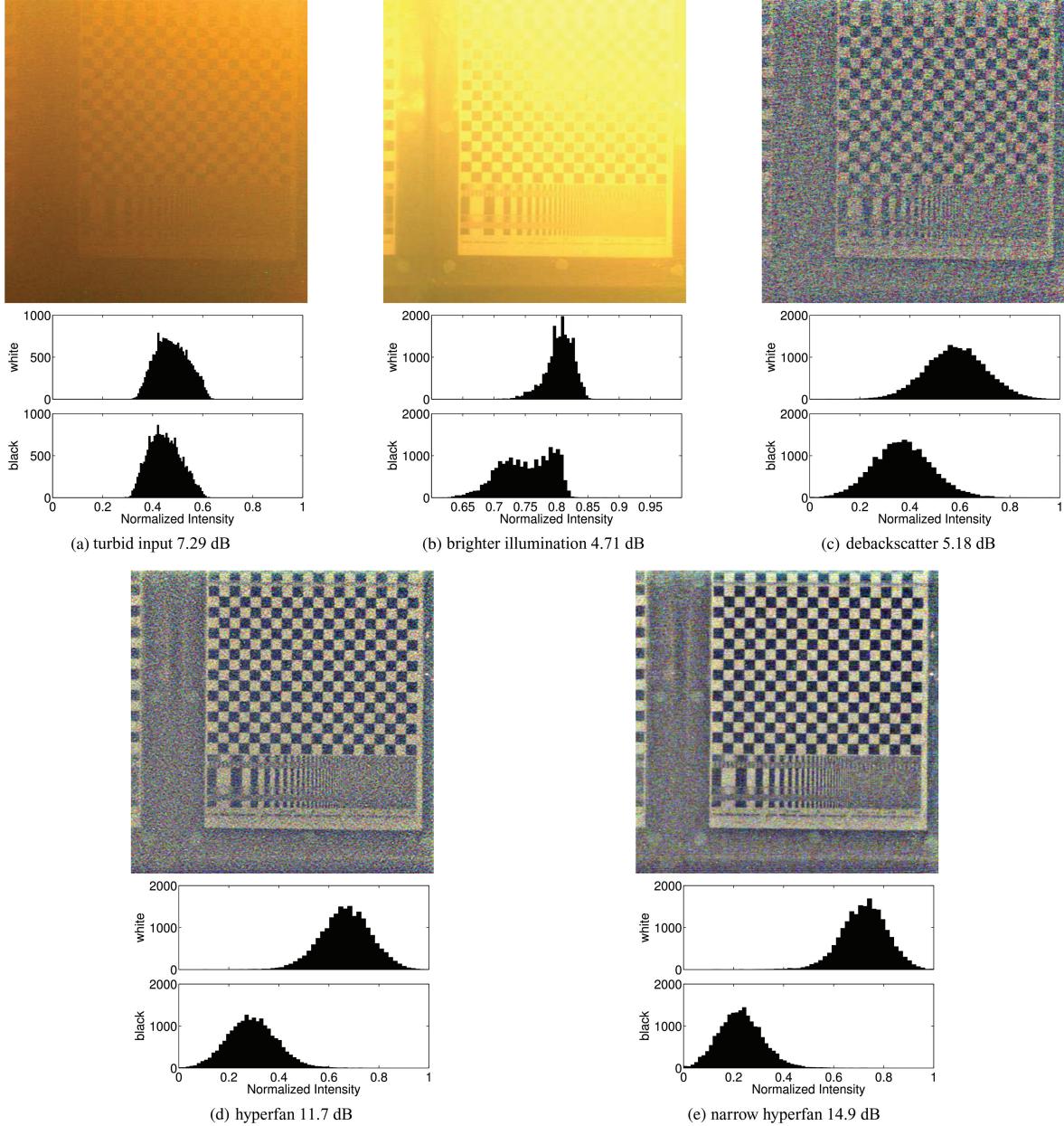


Fig. 22. A demonstration of imaging in a turbid medium. The histograms beneath each image indicate the distribution of pixel intensities in white and black checkerboard squares, and numeric values are CNR for the same. (a) The low-contrast input is not ameliorated by (b) adding light, due to backscatter and saturation (note the change in scale on the histograms); (c) backscatter compensation increases contrast but is noise limited, while (d) hyperfan filtering significantly reduces noise, yielding higher-CNR results; (e) further improvement is possible by trading off depth of field in this planar scene.

components in the passband of a hyperfan filter as described in Section 5.2. Note the dramatic improvement in passband performance on the back wall of the Lego scene; this scene content appears at a large slope in s, u , showing substantial aliasing in s and t . Note also that the presence of aliased components in the stop-band signal, shown in the bottom two images, is probably impossible to remove by purely linear means, as the undesired and desired signals overlap in the frequency domain.

7. EXPERIMENTS: LENSLET-BASED CAMERA

Validation was carried out on imagery collected using a Lytro consumer-grade lenslet-based hand-held light field camera; typical low-contrast results are depicted in Figure 21. The left column depicts a low-light aquarium scene, and the right depicts a low-light outdoor scene. Inspection of the unfiltered and filtered images shows that the hyperfan filter has significantly attenuated the noise.

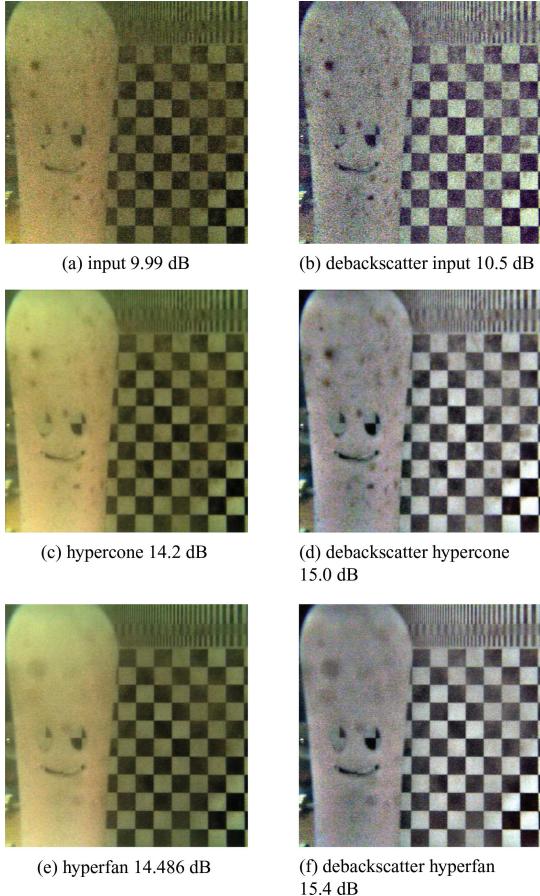


Fig. 23. A scene with suspended particulate matter and relatively clear water. Numerical results are CNR over the checkerboard region of the image, and images in the right column have been backscatter compensated. Relative to the input (top) the hypercone filter reduces noise (center), but does not attenuate particulate occluders. The hyperfan filter reduces noise and attenuates the occluders while maintaining focus over the scene's volume (bottom).

Note also that the specks of dirt on the side of the aquarium in the top row have been rejected by the depth selectivity of the filter; one of these is indicated by a black arrow in the inset depicting a Silver Dollar fish.

7.1 Murky Water and Particulate Matter

Figure 22 depicts a checkerboard as imaged through turbid water. The histograms beneath each image show the distribution of pixel intensities corresponding to white (top) and black (bottom) checkerboard squares, where intensity is taken as the mean of the three color channels. Numeric values are Contrast-to-noise-ratio CNR rather than PSNR, because this is more reflective of the quality of images in the presence of a scattering medium. PSNR neglects the biasing effect of backscatter, which effectively limits the range and contrast of a signal. Contrast was taken as the difference between the means of pixels belonging to white and black checkerboard squares, and noise level as the standard deviation of pixels from their respective distribution means.

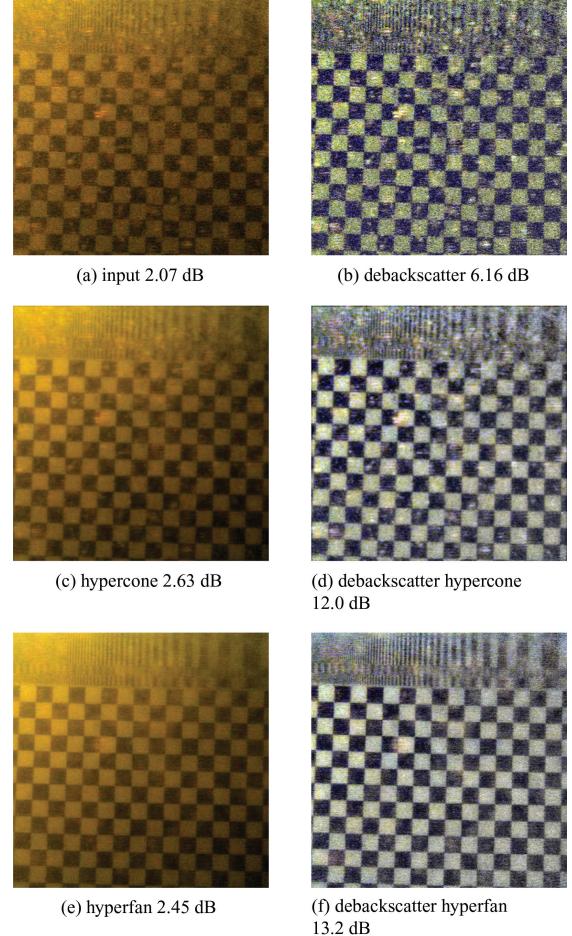


Fig. 24. Similar to Figure 23, except here the water is considerably more turbid, showing a lower CNR in the input image (top) and a greater advantage in applying the hypercone filter without depth selectivity (center). Again the hyperfan filter attenuates occluders (bottom).

In Figure 22 illumination and camera were co-located, resulting in significant backscatter as seen in (a). The result of increasing illumination is depicted in (b), where saturation and backscatter have limited the efficacy of this approach both visually and in terms of CNR. The result of gain-adjusting the input is shown in (c), including removal of a low-frequency biasing term caused by backscatter. The biasing term was estimated by low-pass filtering in the u and v dimensions. Notable is the similarity of this adjusted image to a gain-adjusted low-light image, where noise has limited the extent to which contrast can be enhanced. The final two images show the output of the hyperfan filter tuned to two different depth ranges: the first is for a wide depth range including content between the camera and the checkerboard, while the final image is for a narrow filter more closely matching the geometry of this constant-depth scene. In all cases, the noise reduction effected by the hyperfan filter has been significant visually and in terms of CNR.

In applications involving heterogeneous occluders, such as snow, rain, or particulate suspended in water, the depth selectivity of the hypercone filter becomes an asset in reducing the influence of the interfering elements. Figures 23 and 24 show scenes imaged through

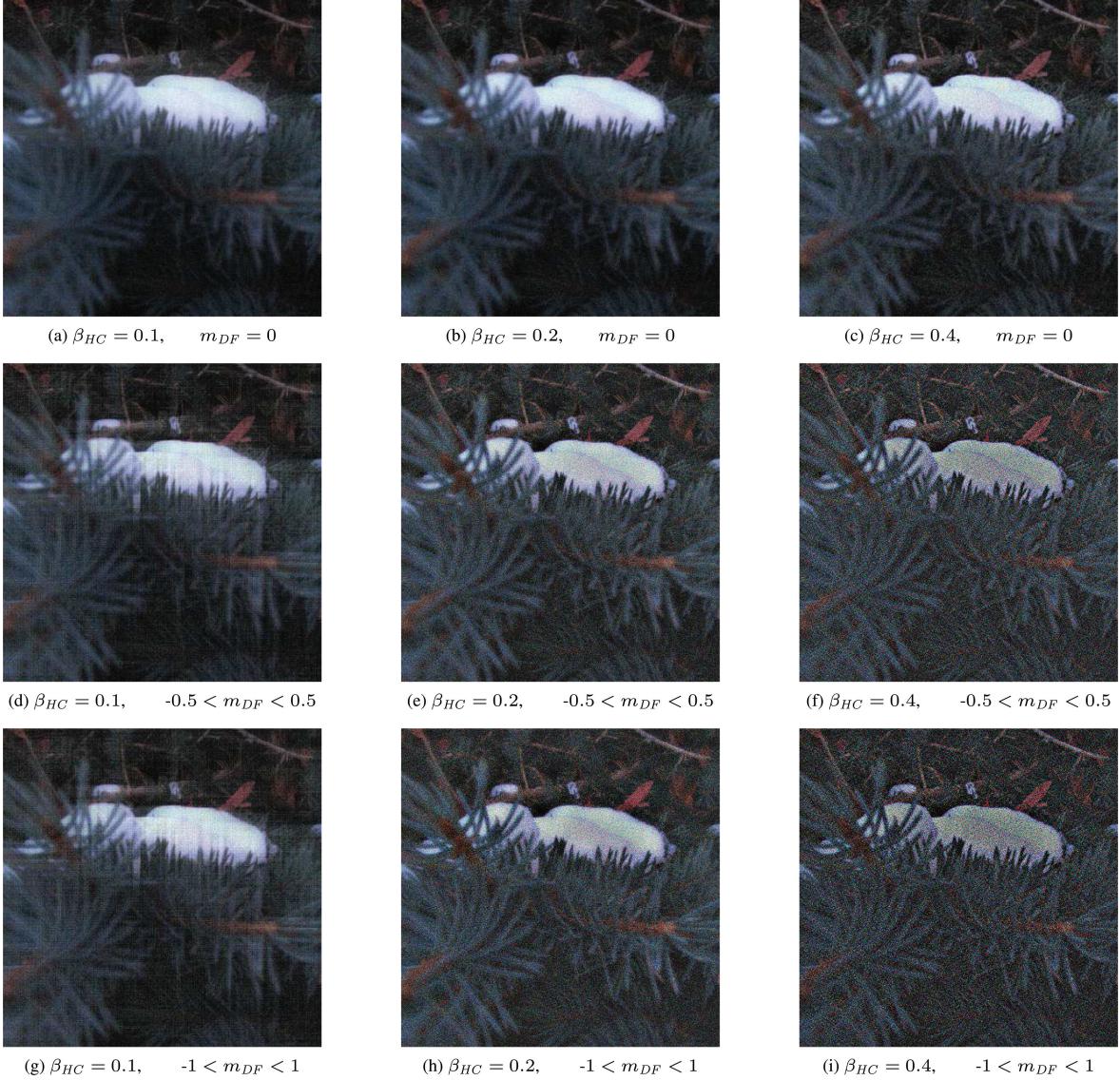


Fig. 25. Visual impact of the fan angle and hypercone bandwidth of a hyperfan filter. Columns from left to right correspond to hypercone bandwidths of 0.1, 0.2, and 0.4, and rows from top to bottom correspond to dual-fans subtending slopes m_{DF} of 0, ± 0.5 and ± 1 . Note that wider fans admit a greater range of depths, broadening the depth of field, but also admitting more noise. Narrower hypercones admit less noise, but attenuate non-Lambertian effects including occlusions (this appears as streaking in the left column of images).

fine, suspended particulate matter. The hypercone filter increases the CNR of the images, but has little effect on the particulate matter, while the hyperfan both reduces noise and attenuates the occluding particles. We attribute the decrease in CNR between the hypercone and hyperfan output in Figures 24(c) and (e) to the nonstationary mean across the image caused by backscatter, which is not accounted for in the CNR metric. Note that the CNR for the corresponding backscatter-compensated images reflects the qualitative improvement in these images.

Figure 23 features clearer water than Figure 24 and there is therefore less advantage in applying the hypercone. That scene also includes a foreground element positioned approximately halfway between the checkerboard and the camera, requiring that a volu-

metric focal region be utilized to keep all scene elements in focus. This figure underlines that particle attenuation is not achieved by the same mechanism as noise reduction. There is adequate illumination in this scene, and the noise level is low. All the scene elements, including the particulate matter, conform to the rules of parallax motion and will therefore fall within the frequency hypercone in the light field. It is the depth selectivity of the hyperfan that allows us to single out the desired scene elements.

Note that CNR is a useful but inaccurately named measure in this context, as the “noise” value includes interference from the particles. A more accurate term would be the contrast-to-noise-and-interference ratio, similar to the carrier-to-noise-and-interference ratio employed in telecommunications [Proakis and Salehi 2007].

7.2 Visual Impact of Bandwidth

To investigate the visual impact of the dual-fan angle and hypercone bandwidth, hyperfan filters over a range of settings were applied to light fields taken in low light, or with noise artificially added. Figure 25 shows a typical result: wider fans admit wider depths of field, but attenuate noise less effectively. At the same time, narrower hypercone bandwidths attenuate noise more effectively, but also attenuate potentially desirable non-Lambertian effects, including occlusion.

8. DISCUSSION AND FUTURE DIRECTIONS

We have established that the frequency-domain ROS of a light field image is a hyperfan at the intersection of a dual fan and a hypercone. We have designed, implemented, and tested a novel volumetric focus filter that selectively passes this ROS. This approach to light field denoising is linear and featureless, operating efficiently as a single-step, constant-runtime filter.

We have demonstrated the filter outperforming a range of linear and nonlinear alternatives over a range of conditions, including noise type, noise level, aperture count, and scene content. Test scenes included examples of occlusion, non-Lambertian surfaces, attenuating media, and interference.

Numeric results were shown for 12 light fields from the Stanford Light Field Archive, including representative images and quantitative results over a range of metrics. The filter was shown effective at removing noise in all cases, generally outperforming the other methods we evaluated including planar, dual-fan, overcomplete DCT, K-SVD, and video-based VBR3D methods. We also showed that the hyperfan filter’s performance scales with aperture count.

Further results demonstrated the filter on imagery collected with the Lytro consumer-grade light field camera, including scenes with low light, turbid (murky) water, and suspended underwater particulate matter. We showed how increased illumination can lead to saturation in the presence of backscatter, effectively limiting how much light can be employed to mitigate contrast limits in underwater imaging. The hyperfan filter was shown to significantly improve CNR and visibly improve image quality. Finally, we demonstrated that, where aliased components are present, an inclusion of these components in the filter’s passband can significantly improve output performance.

There are several immediate avenues for future work. Automated means of selecting filter parameters would be desirable, and we believe the hyperfan filter could be useful for a range of interesting tasks, including compression and interpolation. The inverse-DFT FIR-based approach we presented was only one of many possible approaches to spatial-domain implementation. The FIR filter design might benefit from iterative refinement similar to that presented in Cetin et al. [1997], for example, and recursive infinite impulse response (IIR) filters may be more appropriate in some applications, particularly where hardware implementation would be of benefit.

The 2012 paper “When Does Computational Imaging Improve Performance?” and follow-on work [Cossairt et al. 2012; Mitra et al. 2013] provide theoretical bounds on image improvement and relate it to absolute light levels. It would be interesting to evaluate the hyperfan filter in this context, and against other computational photography techniques such as focal sweep and flutter shutter [Nagahara et al. 2008; Raskar et al. 2006].

Finally, the hyperfan filter can attenuate desired occluding edges. In the case of the wide depth-of-field Lego Knights scene, for

example, some ghosting is visible in desired, occluding foreground elements. A means of better dealing with these occlusions would be desirable, perhaps through detection and refinement of small subsets of the light field using a more complex method, like the variational Bayesian framework proposed in Goldluecke and Wanner [2013], or by employing a form of median filtering like that proposed in Vaish et al. [2006].

REFERENCES

- E. H. Adelson and J. Y. A. WANG. 2002. Single lens stereo with a plenoptic camera. *IEEE Trans. Pattern Anal. Mach. Intell.* 14, 2, 99–106.
- A. Agrawal, Y. Xu, and R. Raskar. 2009. Invertible motion blur in video. *ACM Trans. Graph.* 28, 3.
- M. Aharon, M. Elad, and A. Bruckstein. 2006. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. Signal Process.* 54, 11, 4311–4322.
- R. Ansari. 1987. Efficient IIR and FIR fan filters. *IEEE Trans. Circ. Syst.* 34, 8, 941–945.
- J. Berent and P. L. Dragotti. 2007. Plenoptic manifolds. *Signal Process. Mag.* 24, 6, 34–44.
- T. E. Bishop and P. Favaro. 2012. The light field camera: Extended depth of field, aliasing, and superresolution. *IEEE Trans. Pattern Anal. Mach. Intell.* 34, 5, 972–986.
- K. Bitsakos and C. Fermuller. 2006. Depth estimation using the compound eye of dipteran flies. *Biol. Cybernet.* 95, 5, 487–501.
- R. Bolles, H. Baker, and D. Marimont. 1987. Epipolar-plane image analysis: An approach to determining structure from motion. *Intl. J. Comput. Vis.* 1, 1, 7–55.
- A. Buades, B. Coll, and J.-M. Morel. 2005. A review of image denoising algorithms, with a new one. *SIAM J. Multiscale Model. Simul.* 4, 2, 490–530.
- A. Cetin, O. Gerek, and Y. Yardimci. 1997. Equiripple FIR filter design by the FFT algorithm. *IEEE Signal Process. Mag.* 14, 2, 60–64.
- J. Chai, X. Tong, S. Chan, and H. Shum. 2000. Plenoptic sampling. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH’00)*. 307–318.
- S.-C. Chan and H.-Y. Shum. 2000. A spectral analysis for light field rendering. In *Proceedings of the International Conference on Image Processing (ICIP’00)*. Vol. 2, 25–28.
- O. Cossairt, M. Gupta, and S. Nayar. 2012. When does computational imaging improve performance? *IEEE Trans. Image Process.* 22, 2, 447–458.
- K. Dabov, A. Foi, and K. Egiazarian. 2007. Video denoising by sparse 3D transform-domain collaborative filtering. In *Proceedings of the 15th European Signal Processing Conference (EURIPCO’07)*. 7.
- K. J. Dana, B. Van Ginneken, S. K. Nayar, and J. J. Koenderink. 1999. Reflectance and texture of real-world surfaces. *ACM Trans. Graph.* 18, 1, 1–34.
- D. G. Dansereau. 2003. 4D light field processing and its application to computer vision. M.S. thesis, Department of Electrical and Computer Engineering, University of Calgary. <http://www-personal.acfr.usyd.edu.au/ddan1654/dansereau2003MscThesis.pdf>.
- D. G. Dansereau. 2014. Plenoptic signal processing for robust vision in field robotics. Ph.D. thesis, Australian Centre for Field Robotics, School of Aerospace, Mechanical and Mechatronic Engineering, The University of Sydney.
- D. G. Dansereau, D. L. Bongiorno, O. Pizarro, and S. B. Williams. 2013a. Light field image denoising using a linear 4D frequency-hyperfan all-in-focus filter. In *Proceedings of the SPIE Conference on Computational Imaging (SPIE’13)*. Vol. 8657.

- D. G. Dansereau, O. Pizarro, and S. B. Williams. 2013b. Decoding, calibration and rectification for lenslet-based plenoptic cameras. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'13)*. 1027–1034.
- D. G. Dansereau and L. T. Bruton. 2003. A 4D frequency-planar IIR filter and its application to light field processing. In *Proceedings of the International Symposium on Circuits and Systems (ISCAS'03)*. Vol. 4. 476–479.
- D. G. Dansereau and L. T. Bruton. 2004. Gradient-based depth estimation from 4D light fields. In *Proceedings of the International Symposium on Circuits and Systems (ISCAS'04)*. Vol. 3. 549–552.
- D. G. Dansereau and L. T. Bruton. 2007. A 4-D dual-fan filter bank for depth filtering in light fields. *IEEE Trans. Signal Process.* 55, 2, 542–549.
- D. G. Dansereau, I. Mahon, O. Pizarro, and S. B. Williams. 2011. Plenoptic flow: Closed-form visual odometry for light field cameras. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'11)*. 4455–4462.
- D. G. Dansereau and S. B. Williams. 2011. Seabed modeling and distractor extraction for mobile AUVS using light field filtering. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'11)*. 1634–1639.
- F. Durand, N. Holzschuch, C. Soler, E. Chan, and F. Sillion. 2005. A frequency analysis of light transport. In *Proceedings of the 32nd International Conference on Computer Graphics and Interactive Techniques: Papers (SIGGRAPH'05)*. 1115–1126.
- M. Elad and M. Aharon. 2006. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans. Image Process.* 15, 12, 3736–3745.
- W. Freeman, A. Levin, S. Hasinoff, P. Green, and F. Durand. 2009. 4D frequency analysis of computational cameras for depth of field extension. Tech. rep. MIT-CSAIL-TR-2009-019. <http://dspace.mit.edu/handle/1721.1/45513>.
- W. S. Geisler. 2008. Visual perception and the statistical properties of natural scenes. *Ann. Rev. Psychol.* 59, 167–192.
- B. Goldluecke and S. Wanner. 2013. The variational structure of disparity and regularization of 4D light fields. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'13)*. 1003–1010.
- X. Gu, S. J. Gortler, and M. F. Cohen. 1997. Polyhedral geometry and the two-plane parameterization. In *Proceedings of the Eurographics Workshop on Rendering Techniques*. Springer, 1–12.
- O. G. Guleryuz. 2007. Weighted averaging for denoising with overcomplete dictionaries. *IEEE Trans. Image Process.* 16, 12, 3020–3034.
- A. Isaksen, L. McMillan, and S. Gortler. 2000. Dynamically reparameterized light fields. In *Proceedings of the International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH'00)*. 297–306.
- Y. Ji, J. Ye, and J. Yu. 2013. Reconstructing gas flows using light-path approximation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'13)*. 2507–2514.
- J.-H. Lambert. 1760. *Photometria, sive de Mensura et gradibus luminis, colorum et umbrae*. Eberhard Klett.
- A. Levin and F. Durand. 2010. Linear view synthesis using a dimensionality gap light field prior. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'10)*. 1831–1838.
- A. Levin, S. Hasinoff, P. Green, F. Durand, and W. Freeman. 2009. 4D frequency analysis of computational cameras for depth of field extension. *ACM Trans. Graph.* 28, 3.
- M. Levoy, B. Chen, V. Vaish, M. Horowitz, I. McDowall, and M. Bolas. 2004. Synthetic aperture confocal imaging. *ACM Trans. Graph.* 23, 3, 825–834.
- M. Levoy and P. Hanrahan. 1996. Light field rendering. In *Proceedings of the International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH'96)*. 31–42.
- A. Lumsdaine and T. Georgiev. 2008. Full resolution lightfield rendering. Tech. rep., Adobe Systems. <http://tgeorgiev.net/FullResolution.pdf>.
- A. Lumsdaine and T. Georgiev. 2009. The focused plenoptic camera. In *Proceedings of the IEEE International Conference on Computational Photography (ICCP'09)*. 1–8.
- K. Maeno, H. Nagahara, A. Shimada, and R.-I. Taniguchi. 2013. Light field distortion feature for transparent object recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'13)*. 2786–2793.
- K. Mitra, O. Cossairt, and A. Veeraraghavan. 2013. A frame-work for the analysis of computational imaging systems with practical applications. <http://arxiv.org/pdf/1308.1981.pdf>.
- H. Nagahara, S. Kuthirummal, C. Zhou, and S. K. Nayar. 2008. Flexible depth of field photography. In *Proceedings of the European Conference on Computer Vision (ECCV'08)*. Springer, 60–73.
- J. Neumann, C. Fermuller, Y. Aloimonos, and V. Brajovic. 2005. Compound eye sensor for 3D ego motion estimation. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'05)*. Vol. 4. 3712–3717.
- R. Ng. 2005. Fourier slice photography. *ACM Trans. Graph.* 24, 3, 735–744.
- R. Ng, M. Levoy, M. Bredif, G. Duval, M. Horowitz, and P. Hanrahan. 2005. Light field photography with a hand-held plenoptic camera. Tech. rep. CSTR 2, Department of Computer Science, Stanford University. <https://graphics.stanford.edu/papers/lfcamera/lfcamera-150dpi.pdf>.
- M. O'TOOLE, R. Raskar, and K. N. Kutulakos. 2012. Primal-dual coding to probe light transport. In *Proceedings of the International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH'12)*. Vol. 31. 39.
- J. G. Proakis and M. Salehi. 2007. *Digital Communications* 5th Ed. McGraw-Hill.
- L. R. Rabiner and B. Gold. 1975. *Theory and Application of Digital Signal Processing*, Volume 1. Prentice-Hall.
- R. Raskar, A. Agrawal, and J. Tumblin. 2006. Coded exposure photography: Motion deblurring using fluttered shutter. *ACM Trans. Graph.* 25, 3, 795–804.
- R. Raskar, A. Agrawal, C. A. Wilson, and A. Veeraraghavan. 2008. Glare aware photography: 4D ray sampling for reducing glare effects of camera lenses. *ACM Trans. Graph.* 27, 3.
- D. L. Ruderman. 1997. Origins of scaling in natural images. *Vis. Res.* 37, 23, 3385–3398.
- A. Shnayderman, A. Gusev, and A. Eskicioglu. 2006. An SVD-based grayscale image quality measure for local and global assessment. *IEEE Trans. Image Process.* 15, 2, 422–429.
- J. Stewart, J. Yu, S. J. Gortler, and L. McMillan. 2003. A new reconstruction filter for undersampled light fields. In *Proceedings of the 14th Eurographics Workshop on Rendering (EGRW'03)*. Eurographics Association, 150–156.
- V. Vaish, M. Levoy, R. Szeliski, C. Zitnick, and S. Kang. 2006. Reconstructing occluded surfaces using synthetic apertures: Stereo, focus and robust measures. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'06)*. Vol. 2. 2331–2338.
- A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin. 2007. Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing. *ACM Trans. Graph.* 26, 3.
- Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. 2004. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* 13, 4, 600–612.

- S. Wanner and B. Goldluecke. 2013. Variational light field analysis for disparity estimation and super-resolution. *IEEE Trans. Pattern Anal. Mach. Intell.* 36, 3, 606–619.
- B. Wilburn, N. Joshi, V. Vaish, E. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, and M. Levoy. 2005. High performance imaging using large camera arrays. *ACM Trans. Graph.* 24, 3, 765–776.
- H. Yang, M. Pollefeys, G. Welch, J. Frahm, and A. Ilie. 2007. Differential camera tracking through linearizing the local appearance manifold. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'07)*. 1–8.
- Z. Yu, X. Guo, X. Chen, and Y. Yu. 2013. Catadioptric array photography for low light imaging. In *Proceedings of the 2nd IEEE International Workshop on Computational Cameras and Displays (CCD'13)*.

Received March 2014; revised June 2014; accepted August 2014