Transfer Efficiency and Depth Invariance in Computational Cameras

Jongmin Baek
Stanford University
jbaek@cs.stanford.edu

Abstract

Recent advances in computational cameras achieve extension of depth of field by modulating the aperture of an imaging system, either spatially or temporally. They are, however, accompanied by loss of image detail, the chief cause of which is low and/or depth-varying frequency response of such systems. In this paper, we examine the tradeoff between achieving depth invariance and maintaining high transfer efficiency by providing a mathematical framework for analyzing the transfer function of these computational cameras. Using this framework, we prove mathematical bounds on the efficacy of the tradeoff. These bounds lead to observations on the fundamental limitations of computational cameras. In particular, we show that some existing designs are already near-optimal in our metrics.

1. Introduction

The depth of field (DOF) of an imaging system refers to the range of depths over which objects appear sharp. Photographers traditionally obtained larger DOF by stopping down and sacrificing light. More recently, however, computational approaches have enabled larger DOF with little or no loss of light. In these approaches, the image acquisition process is modified by coding the aperture spatially or temporally, and this coding is computationally inverted afterwards. The fidelity of the output is directly related to the optical transfer function (OTF) of the system, which describes the per-frequency attenuation. As such, an OTF that is high at every spatial frequency is desirable.

The aim of this paper is to characterize how having a high OTF relates to the depth invariance of a computational camera, both of which are common goals in the literature. We show that the OTF exhibits a very specific structure that leads to a tighter bound on its magnitude than previous work, and more importantly, we establish a mathematical tradeoff between its magnitude and depth invariance.

By analyzing existing computational cameras in the context of these new bounds, we demonstrate the surprising optimality of some existing designs. Specifically, we show (1) focus sweep[16] is near-optimally efficient in achieving depth-invariant OTF, and (2) for any given spatial frequency, there exists a (tuned) cubic phase plate[5] that is near-optimally efficient in achieving depth-invariant MTF (the magnitude of the OTF) at that frequency.

The rest of the paper is organized as follows. In Section 2, we discuss assumptions and prior work, and briefly review the definition of the OTF. In Section 3, we formalize metrics for measuring transfer efficiency and depth invariance, and discuss limitations. In Section 4, we prove a structural property of the OTF and use it to bound the scores of a computational camera in these metrics. In Section 5, we apply these results to existing designs in order to account for their specific performances. Conclusions and future work are discussed in Section 6.

2. Background

We make the wave-optics assumptions that are common in the computational camera and optics literature about image formation: a diffraction-limited, incoherent, single-capture imaging system observes a Lambertian scene with locally constant depth. Under these conditions, the observed image equals the scene attenuated frequency-wise by the optical transfer function (OTF). The scene is then reconstructed by inverting this attenuation process[9].

The OTF is determined by the aperture of the imaging system[7, 10, 19], and this fact has given rise to novel aperture designs that in turn optimize the OTF in some regard. In fact, many computational cameras rely on spatiotemporally modulating the aperture in order to extend the depth of field—such as phase coding[5, 6, 12, 21], amplitude coding[11, 15, 18, 22], and temporal coding[16]—though other approaches exist, notably stereo[20], the plenoptic camera[17], and multiple capture[8].

2.1. The optical transfer function

Here we review the definition of the OTF. Consider an aperture function $P : \mathbb{R}^2 \to \mathbb{C}$. We stipulate that coordinates and transmission are normalized such that the aperture is a unit square, and the maximum power transmitted is 1:

$$\forall |x|, |y| \leq \tfrac{1}{2},\ |P(x,y)| \leq 1, \quad \text{(0 otherwise.)} \qquad (2.1)$$

Let (f_x, f_y) be the spatial frequency of interest, normalized such that the diffraction limit is at ± 1, and ψ be the defocus parameter[5, 7], which measures the deviation from the well-known Lensmaker's equation, hence depth:

$$\psi := \frac{\pi}{4} \frac{\text{[area of aperture]}}{\text{[wavelength of light]}} \left(\frac{1}{f} - \frac{1}{d_i} - \frac{1}{d_o} \right). \quad (2.2)$$

In this scale, the traditional Hopkins criterion for misfocus is equivalent to $\psi \simeq \pm 4$[3], and wavefront coding in [5] extends depth of field to $\psi \simeq \pm 120$. The OTF at frequency (f_x, f_y) and depth ψ is then the following[7]:

$$OTF_\psi(f_x, f_y) = \iint P\left(t_1 + \tfrac{f_x}{2}, t_2 + \tfrac{f_y}{2}\right) P^*\left(t_1 - \tfrac{f_x}{2}, t_2 - \tfrac{f_y}{2}\right) \times \exp\left\{2i\psi(t_1 f_x + t_2 f_y)\right\} dt_1 dt_2. \quad (2.3)$$

The magnitude of the OTF is called the *modulation transfer function* (MTF). Note that the notion of OTF can easily be extended onto a temporally modulated aperture by simply taking the average of the OTF over time.

2.2. Prior work on analysis of the OTF

The OTF of a standard lens is well-understood[7, 10]. Each computational imaging system in literature is accompanied by analysis of its OTF, and there have also been efforts to compare them[8, 12, 25]. However, despite their abundance, there has not been commensurate effort to study the OTF of *general* computational cameras with spatiotemporally modulated apertures. Only recently, it has been remarked that the total amount of energy in the OTF at each nonzero frequency (f_x, f_y) is bounded:[2, 12]

$$\int_{-\infty}^{\infty} |OTF_\psi(f_x, f_y)|^2 d\psi \leq \frac{\pi}{f_x^2 f_y^2} K_x^2 (K_y - K_x/3), \quad (2.4)$$

where $K_x = f_x - |f_x|^2$, $K_y = f_y - |f_y|^2$ for brevity and $K_x \leq K_y$ without loss of generality. However, this bound is not useful by itself: the maximum is achieved by a standard lens, which exhibits a high OTF only when in focus.

3. Metrics for computational cameras

Discussing the transfer efficiency and depth invariance of computational cameras requires appropriate metrics to quantify them. We focus on the mathematical limits imposed by the OTF per frequency, as opposed to external effects (such as perception, noise level, and the reconstruction algorithm), and define them solely as functions of the OTF.

3.1. Measuring transfer efficiency

The transfer efficiency at the spatial frequency (f_x, f_y) may be measured by the average power of the OTF at that frequency[2, 12]. Letting $\psi \in [-S, S]$ be the depth range of interest, it can naturally be defined as,

$$\frac{1}{2S} \int_{-S}^{S} |OTF_\psi(f_x, f_y)|^2 d\psi. \quad (3.1)$$

The power of the OTF above directly determines the signal-to-noise ratio of the observed image[2]. Note that Eq.(2.4) immediately puts a bound on this expression[12].

3.2. Measuring depth invariance

As observed in Section 2.2, transfer efficiency alone is not useful in evaluating OTFs. This necessitates a separate metric that measures the degree to which the OTF (or MTF) stays consistently high across depths. One such measure is the variance with respect to ψ:

$$\text{Var}_\psi(OTF_\psi(f_x, f_y)) = \frac{1}{2S} \int_{-S}^{S} |OTF_\psi(f_x, f_y)|^2 d\psi - \left| \frac{1}{2S} \int_{-S}^{S} OTF_\psi(f_x, f_y) d\psi \right|^2. \quad (3.2)$$

$$\text{Var}_\psi(MTF_\psi(f_x, f_y)) = \frac{1}{2S} \int_{-S}^{S} |MTF_\psi(f_x, f_y)|^2 d\psi - \left| \frac{1}{2S} \int_{-S}^{S} MTF_\psi(f_x, f_y) d\psi \right|^2. \quad (3.3)$$

These formulae follow from the definition of variance as $E[X^2] - E[X]^2$. They are consistent with the spirit of existing computational cameras that explicitly aim for depth-invariant transfer functions, as to avoid reconstruction errors that arise from misjudging the depth. In fact, the expected per-frequency reconstruction error, when the average MTF is used for reconstruction, is proportional to the variance:

$$\frac{1}{2S} \int_{-S}^{S} \left| \frac{c \cdot MTF_\psi}{E_\theta[MTF_\theta]} - c \right|^2 d\psi = |c|^2 \frac{\text{Var}_\psi(MTF_\psi)}{(E_\psi[MTF_\psi])^2}. \quad (3.4)$$

Here the arguments f_x, f_y are implicit, and c is the signal strength. Other metrics for quantifying DOF exist, such as the Hilbert-space angle for varying depths[21] or the worst-case MTF[12]. The former is analytically intractable. The latter will be briefly discussed in Section 5.

3.3. Limitations

The analysis in this paper stops short of making explicit statements on the quality of the reconstructed image. While the suggested metrics are motivated to maintain the OTF high and therefore invertible, calculating the reconstruction error in manners of [8, 25] requires decisions on the noise level, the choice of deconvolution algorithm, the fidelity of depth or kernel estimation, design-specific analyses, et cetera. We avoid making these choices, which are required for assessing an existing imaging system under a fully specified condition. Instead, our analysis complements such evaluations by exploring the mathematical structures of *all* OTFs, and making general assertions about them.

4. Mathematical limits on the tradeoff

Ideally, for each frequency (f_x, f_y), we would like to maximize the transfer efficiency of a camera, while min-

imizing its depth variance. Mathematically, we have the following objectives arising from (3.1) and (3.2):

$$\text{Maximize } \frac{1}{2S}\int_{-S}^{S} |OTF(\psi)|^2 d\psi,$$

$$\text{Minimize } \frac{1}{2S}\int_{-S}^{S} |OTF(\psi)|^2 d\psi - \left|\frac{1}{2S}\int_{-S}^{S} OTF(\psi)d\psi\right|^2.$$

(The arguments f_x, f_y are implicit. Substitute in MTF in case (3.3) is used.) Solving each in isolation leads to trivial results: transfer efficiency is maximized by a standard lens, as will be shown, and variance is minimized by a total occluder. Either solution is abysmal for the other metric. Hence, we must consider the two metrics in conjunction.

Unfortunately, a cursory glance reveals that they cannot be optimized simultaneously, since they have a common expression, namely (3.1). This represents a fundamental tradeoff in transfer efficiency and depth invariance. The degree to which we can independently optimize these two objectives depends on the difference:

$$\left|\frac{1}{2S}\int_{-S}^{S} OTF(\psi)d\psi\right|^2, \tag{4.1}$$

$$\left|\frac{1}{2S}\int_{-S}^{S} MTF(\psi)d\psi\right|^2. \tag{4.2}$$

Our goal in the remainder of this section is to demonstrate a fairly tight upper bound to (3.1), (4.1) and (4.2). Later in Section 5.1, these bounds will provide context to the performance of existing computational cameras.

4.1. Structure of the OTF in Fourier domain

We wish to maximize some objective functions involving the OTF of a spatiotemporally modulated aperture, at a particular frequency (f_x, f_y). The OTF, however, cannot take an arbitrary form. For instance, (2.4) tells us that the total power contained in the OTF at this frequency is bounded. Since our optimization is over the space of all possible OTFs, such constraints must be taken into account.

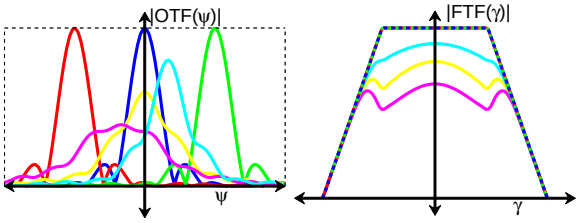
The traditional definition of the OTF as supplied in (2.3) does little to reveal any other structural constraints that may exist. Figure 1(a) shows the OTF at $(f_x, f_y) = (0.1, 0.5)$ as a single-variable function of ψ, for six different spatially coded apertures. No conspicuous structure is perceivable.

Instead, let us examine the Fourier duals of these OTFs with respect to ψ. Elect the following change of variables in (2.3): $t := t_1 f_x$ and $\gamma := (t_1 f_x + t_2 f_y)/\pi$. This yields,

$$OTF(\psi) = \frac{\pi}{f_x f_y}\iint P\left(\frac{t}{f_x} + \frac{f_x}{2}, \frac{\gamma\pi - t}{f_y} + \frac{f_y}{2}\right) \times P^*\left(\frac{t}{f_x} - \frac{f_x}{2}, \frac{\gamma\pi - t}{f_y} - \frac{f_y}{2}\right)\exp(2\pi i\psi\gamma)dtd\gamma. \tag{4.3}$$

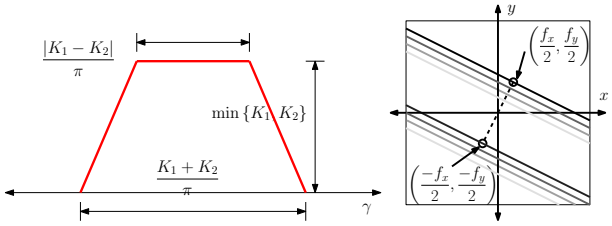
We immediately deduce the Fourier dual with respect to ψ:

$$\mathcal{F}\{OTF\}(\gamma) = \frac{\pi}{f_x f_y}\int P\left(\frac{t}{f_x} + \frac{f_x}{2}, \frac{\gamma\pi - t}{f_y} + \frac{f_y}{2}\right)$$



(a) The magnitude of OTFs (b) The magnitude of FTFs

Figure 1. **Left**: the OTFs at frequency $(f_x, f_y) = (0.1, 0.5)$ for several apertures as a function of ψ. **Right**: the magnitude of the corresponding FTFs, which are the Fourier duals of the OTFs.



(a) $R(\gamma)$ (b) Calculating *FTF*

Figure 2. **Left**: the measure of the support of the integrand in (4.4), namely $R(\gamma)$. Its dimensions depend on K_x and K_y. **Right**: illustration of how $FTF(\gamma)$ is computed. At each γ, $FTF(\gamma)$ equals the correlation of two lines offset by (f_x, f_y) and normal to the dotted line. As γ varies, the two lines sweep through the aperture.

$$\times P^*\left(\frac{t}{f_x} - \frac{f_x}{2}, \frac{\gamma\pi - t}{f_y} - \frac{f_y}{2}\right)dt. \tag{4.4}$$

Let us use the shorthand $FTF(\gamma) := \mathcal{F}\{OTF\}(\gamma)$ hereafter. The geometric interpretation of (4.4) is the correlation of the aperture with itself at an offset (f_x, f_y), where the coordinates $\left(\frac{t}{f_x} \pm \frac{f_x}{2}, \frac{\gamma\pi - t}{f_y} \pm \frac{f_y}{2}\right)$ as a parametric function of t trace out a pair of lines of slope $-f_x/f_y$, as seen in Figure 2(b). The domain of integration in (4.4) for which the integrand is nonzero satisfies the following:

$$\left|\frac{t}{f_x} \pm \frac{f_x}{2}\right| \le \frac{1}{2}, \qquad \left|\frac{\gamma\pi - t}{f_y} \pm \frac{f_y}{2}\right| \le \frac{1}{2}.$$

The two inequalities combine into

$$\max\left\{\frac{-K_x}{2}, \frac{-K_y}{2} + \gamma\pi\right\} \le t \le \min\left\{\frac{K_x}{2}, \frac{K_y}{2} + \gamma\pi\right\}. \tag{4.5}$$

Then, the range of valid values for t as a function of γ forms a trapezoid, as seen in Figure 2(a). Denote this by $R(\gamma)$; in other words, $R(\gamma)$ is the measure of the support of the integrand in (4.4). Most importantly, $R(\gamma)$ is extremely germane to the shape of FTF:

Proposition 1. *For any spatiotemporally modulated aperture and at any frequency* (f_x, f_y),

$$\forall\gamma,\ |FTF(\gamma)| \le \frac{\pi}{f_x f_y}R(\gamma). \tag{4.6}$$

Proof. Since the magnitude of P is at most 1, the integrand in (4.4) is bounded by 1 in magnitude. Hence, the integral is bounded by the area of its support, and (4.6) follows. In case of temporal modulation, the bound applies at each time slice, so the resulting FTF is bounded likewise. □

This bound is tight for a standard lens, since all terms add up constructively. Figure 1(b) shows the FTF of the apertures previously shown in Figure 1(a). As predicted, the magnitude of the FTFs is capped by a trapezoid.

Surprisingly, the converse of Proposition 1 is also true:

Proposition 2. *Fix f_x, f_y. Let $g(\gamma)$ be bounded in magnitude by $\pi/(f_x f_y)R(\gamma)$. Then there exists a spatially modulated aperture whose FTF at frequency (f_x, f_y) equals g.*

Proof. Each value of $FTF(\gamma)$ is computed from the correlation of two parallel line segments over the aperture. It is possible to code the phase of the two lines to obtain any desired value of $FTF(\gamma)$ below the upper bound. □

Proposition 2 tells us that there is no further structural constraint on the FTF beyond Proposition 1, and in turn, the OTF. Thus, we have *fully* characterized the OTF of an arbitrary spatiotemporally modulated aperture at a given frequency. Also, because the FTF has a finite support, it is a more suitable representation for optimization than the OTF.

A simple calculation yields that the power contained in the trapezoid (scaled by $\pi/(f_x f_y)$) is at most:

$$\int_{-\infty}^{\infty} \left| \frac{\pi}{f_x f_y} R(\gamma) \right|^2 d\gamma = \frac{\pi}{f_x^2 f_y^2} K_x^2 (K_y - K_x/3). \quad (4.7)$$

By Parseval's Theorem[4], this also bounds the power of the OTF, and in fact, agrees with (2.4). We finally remark that the FTF is also a slice of the "generalized OTF" in [14].

4.2. Applying the structure of the OTF

Now that we have completely characterized the per-frequency structure of an OTF, let us apply it to our tasks. First, consider the transfer efficiency as quantified by (3.1). Rewriting the expression in terms of the FTF, we obtain

$$\begin{aligned}
\frac{1}{2S}\int_{-S}^{S} |OTF(\psi)|^2 d\psi &= \frac{1}{2S}\int_{-S}^{S} \left| \int FTF(\gamma) e^{2\pi i \gamma\psi} d\gamma \right|^2 d\psi \\
&= \iint FTF(\gamma_1) \left[\int_{-S}^{S} e^{2\pi i(\gamma_1-\gamma_2)\psi} \frac{d\psi}{2S} \right] FTF^*(\gamma_2) d\gamma_1 d\gamma_2 \\
&= \iint FTF(\gamma_1) \text{sinc}\left(2S(\gamma_1-\gamma_2)\right) FTF^*(\gamma_2) d\gamma_1 d\gamma_2 \\
&\le \max_g \iint g(\gamma_1) \text{sinc}\left(2S(\gamma_1-\gamma_2)\right) g^*(\gamma_2) d\gamma_1 d\gamma_2, \quad (4.8)
\end{aligned}$$

where $\forall\psi,\ |g(\psi)| \le \pi/(f_x f_y)R(\psi)$ via Prop. 1. Because $\text{sinc}\left(2S(\gamma_1-\gamma_2)\right)$ is Hermitian, the spectral theorem for compact self-adjoint operators[23] tells us that it can be decomposed into orthonormal eigenfunctions with real eigenvalues. In fact, (4.8) is maximized when g is the eigenfunction with the largest eigenvalue λ, and the maximum achieved is λ multiplied by the power of g.

Remark: This argument is best understood as a continuous analogue of the spectral theorem for matrices: if A is a Hermitian matrix, the expression $v^* A v$ is maximized by letting v be the eigenvector of A with the largest eigenvalue.

Since the power of f is bounded by (4.7), we obtain:

Proposition 3. *For any spatiotemporally modulated aperture and at any frequency (f_x, f_y),*

$$\frac{1}{2S}\int_{-S}^{S} |OTF(\psi)|^2 d\psi \le \frac{\pi}{f_x^2 f_y^2} K_x^2 (K_y - K_x/3)\lambda. \quad (4.9)$$

The largest eigenvalue λ can be found numerically by approximating $\text{sinc}\left(2S(\gamma_1-\gamma_2)\right)$ as a matrix. Recall that this operator has a finite support, namely $|\gamma_1|, |\gamma_2| \le \frac{K_1+K_2}{2\pi}$, so representing it as a matrix is possible.

The left half of Figure 3 shows the bounds obtained for $S = 20$ and compares them against [2, 12]. As shown in the figure, we improve upon the existing bounds. In addition, we can conclude that the standard lens is near-optimally transfer-efficient at most frequencies.

Next, we bound the expression in (4.1), the difference between transfer efficiency and depth variance. Equivalently, we bound its square root, which is the magnitude of the average OTF:

$$\begin{aligned}
\left|\frac{1}{2S}\int_{-S}^{S} OTF(\psi)d\psi\right| &= \left|\frac{1}{2S}\int_{-S}^{S}\int FTF(\gamma)e^{2\pi i\gamma\psi}d\gamma d\psi\right| \\
&= \left|\int FTF(\gamma)\text{sinc}\left(2S\gamma\right)d\gamma\right| \le \int |FTF(\gamma)\text{sinc}\left(2S\gamma\right)|\,d\gamma \\
&\le \int \frac{\pi}{f_x f_y} R(\gamma)|\text{sinc}\left(2S\gamma\right)|d\gamma. \quad (4.10)
\end{aligned}$$

The inequality in (4.10) is tight when

$$FTF(\gamma) = \frac{\pi}{f_x f_y} R(\gamma)\text{sign}\left(\text{sinc}\left(2S\gamma\right)\right). \quad (4.11)$$

However, if S is large, the magnitude of the average OTF can be nicely approximated:

Proposition 4. *For any spatiotemporally modulated aperture and at any frequency (f_x, f_y) and S large enough,*

$$\left|\frac{1}{2S}\int_{-S}^{S} OTF(\psi)d\psi\right| \lessapprox \frac{\pi}{2Sf_x f_y}\min\{K_x, K_y\}. \quad (4.12)$$

Proof. If S is large enough,

$$\begin{aligned}
\left|\frac{1}{2S}\int_{-S}^{S} OTF(\psi)d\psi\right| &\simeq \left|\frac{1}{2S}\int_{-\infty}^{\infty} OTF(\psi)d\psi\right| \\
&= \left|\frac{1}{2S}FTF(0)\right| \le \frac{\pi}{2Sf_x f_y}\min\{K_x, K_y\}.
\end{aligned}$$

The last line follows from Proposition 1. □

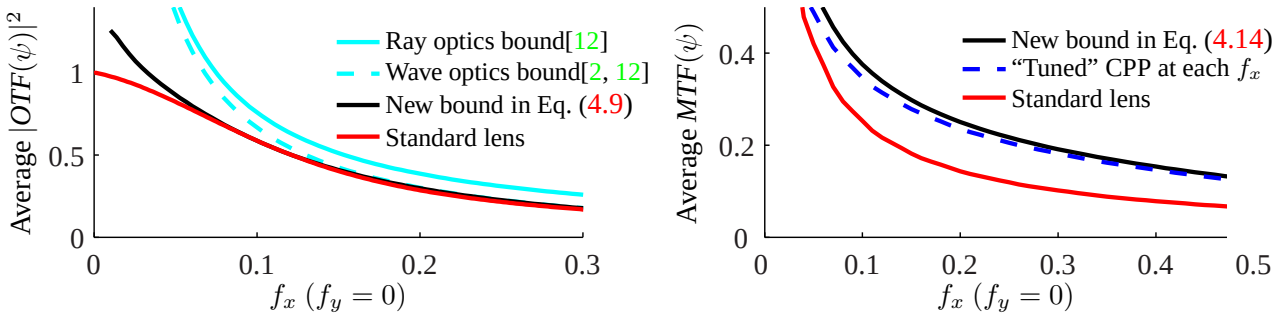


Figure 3. **Left**: comparison of upper bounds on the average power of OTF, for $\psi \in [-20, 20]$. We visualize the bounds along a 1D slice of the frequency domain. Note that our bound in (4.9) is tighter than the previous bounds significantly. **Right**: upper bound on the average MTF for $\psi \in [-100, 100]$. Once again, we visualize the bound along a 1D slice of the frequency domain. At each frequency, we also plot the average MTF of the standard lens, and that of the best-performing cubic phase plate at that frequency, from (5.5). For both figures, the diffraction limit is at ± 1. The bounds along other 1D slices of the frequency domain look sufficiently the same and are omitted.

Lastly, we bound the square root of (4.2), the difference between transfer efficiency and depth invariance in case we define them in terms of the MTF.

$$\begin{aligned}
&\left|\frac{1}{2S}\int_{-S}^{S} MTF(\psi)d\psi\right| = \frac{1}{2S}\int_{-S}^{S}\left|\int FTF(\gamma)e^{2\pi i\gamma\psi}d\gamma\right|d\psi \\
&= \frac{1}{2S}\int_{-S}^{S} e^{-ig(\psi)}\int FTF(\gamma)e^{2\pi i\gamma\psi}d\gamma d\psi, \\
&\quad \text{where } e^{ig(\psi)} \text{ is the phase of } OTF(\psi), \\
&= \max_h \frac{1}{2S}\left|\int_{-S}^{S}\int FTF(\gamma)e^{2\pi i\gamma\psi}e^{-ih(\psi)}d\gamma d\psi\right| \quad (4.13) \\
&\le \sigma\left(\int |FTF(\gamma)|^2 d\gamma\right)\left(\frac{1}{2S}\int_{-S}^{S}\left|e^{-ih(\psi)}\right|^2 d\psi\right), \\
&\quad \text{where } \sigma \text{ is the largest singular value of } e^{2\pi i\gamma\psi} \\
&\quad \text{with support } |\psi| \le S,\ |\gamma| \le (K_x + K_y)/(2\pi), \\
&= \sigma\frac{\pi}{f_x^2 f_y^2}K_x^2(K_y - K_x/3). \quad (4.14)
\end{aligned}$$

The expression in (4.13) is equivalent to a rank-1 approximation: we wish to find a separable (rank-1) function $FTF(\gamma)e^{-ih(\psi)}$ that best correlates to $e^{2\pi i\gamma\psi}$, which is non-separable. The rank-1 approximation of $e^{2\pi i\gamma\psi}$ can be found via its singular value decomposition (SVD)[13]. The right half of Figure 3 visualizes the bound, along with the average MTF for the standard lens and a "tuned" cubic phase plate at each given spatial frequency (optimized for that specific frequency only.)

5. Evaluations of existing designs

In Section 4, we bounded the tradeoff between transfer efficiency and depth invariance for spatiotemporally modulated apertures. Now we test how existing computational cameras fare under these metrics. We consider the standard lens, cubic phase plate (CPP)[5], focus sweep[16], and lattice focal lens[12]. For the definition of these designs and derivation of their OTFs, see the supplementary report[1].

All these designs represent methods of trading off transfer efficiency for depth of field, with tunable parameters: the f-stop for a standard lens, the thickness of CPP, the sweep distance for focus sweep, and the choice of lenslet size and focal power for lattice focal lens. Figures 4 and 5 visualize the tradeoff for these designs at particular frequencies and in aggregate, respectively, over their parameter spaces. Note that when aggregating the per-frequency values, the f-number of the system determines which frequencies to aggregate, since the frequencies are normalized by the diffraction limit. Increasing the f-number lowers the diffraction limit, so the cut-off frequency, if one exists, should scale up by the same factor. The per-frequency analysis remains independent of the f-number, and so does the optimality argument that follows in Section 5.1.

The figures are scatterplots showing the performance of several computational camera: the x-axis measures the transfer efficiency (the average power of OTF over depth), and the y-axis measures the normalized standard deviation of the MTF (or the OTF, depending on the plot.) This quantity is given by dividing the variance of the MTF (or the OTF) by transfer efficiency, and taking its square root. (Formally, we plot $\sqrt{(3.3)/(3.1)}$ or $\sqrt{(3.2)/(3.1)}$ against (3.1).) As we claimed, all families trade off between the two axes.

All these designs outperform the standard lens in terms of the efficacy of the tradeoff. In particular, focus sweep excels at obtaining depth-invariant OTF at all frequencies (the right column of the figures). The cubic phase plate excels at obtaining depth-invariant MTF for axial frequencies (the first and third rows, left column of Figure 4). The lattice focal lens performs well in non-axis-aligned low frequencies.

5.1. Optimality of existing designs

We can formalize the observations made on Figures 4, 5 and give an argument as to why such results are produced. First, consider the focus sweep, which translates the sensor along the optical axis during exposure in order to focus on all planes in some depth range $[-R, R]$. Typically R is slightly larger than S. Its OTF is the time-average of the OTF for the defocused standard lens, and is depth-

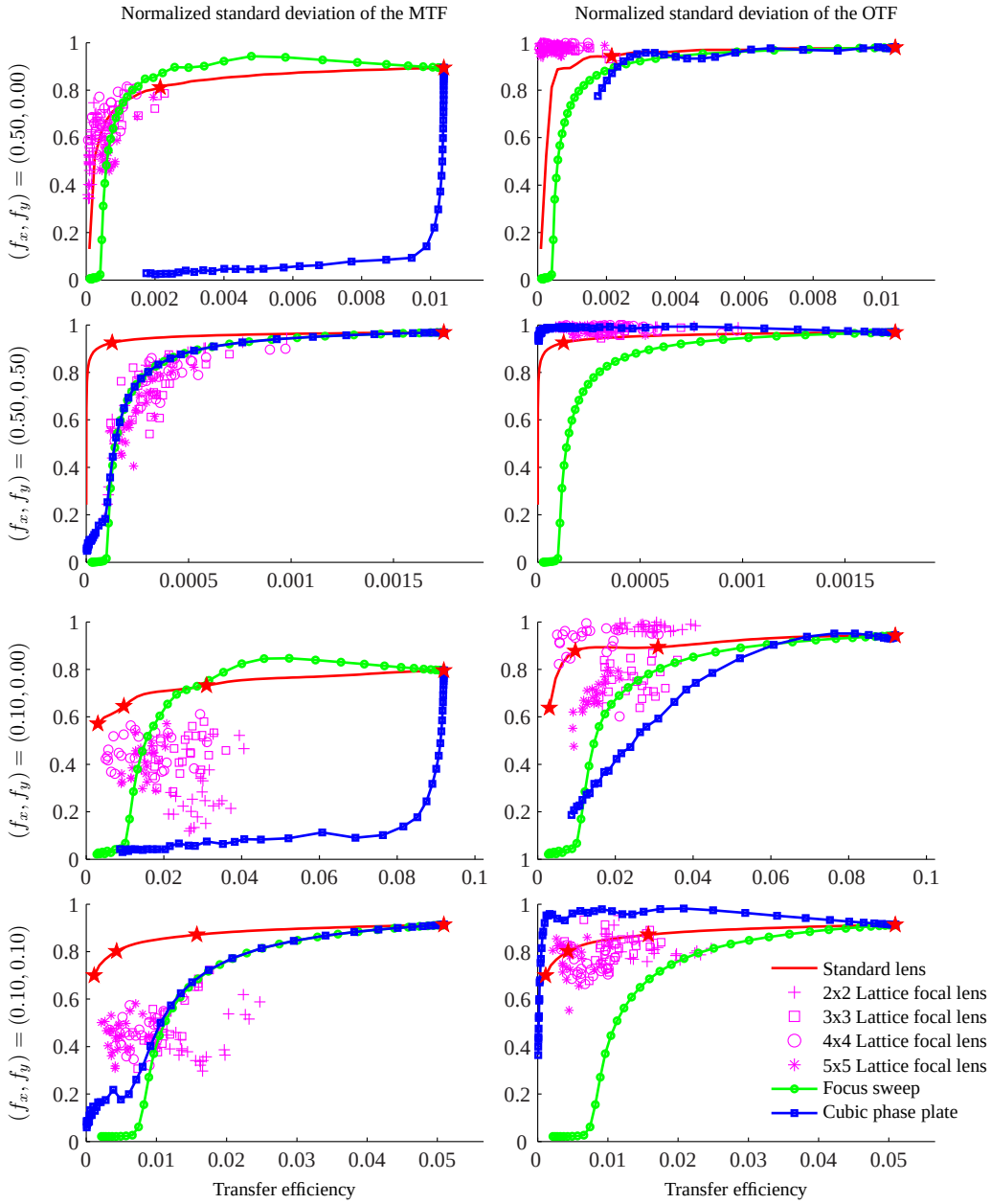


Figure 4. Comparison of various computational cameras in trading off transfer efficiency for depth invariance. The depth range in interest is $\psi \in [-150, 150]$. Each row corresponds to a particular frequency (f_x, f_y). The left and the right columns measure the *normalized standard deviation* of the MTF or the OTF, respectively, against the transfer efficiency. For standard lens, each star denotes a discrete f-stop. The sweet spot is bottom-right, where transfer efficiency is maximized and depth variance is zero.

independent. Hence, assuming S is large enough,

$$\left|\frac{1}{2S}\int_{-S}^{S} OTF(\psi)d\psi\right| \simeq |OTF(0)| \text{ due to depth invariance,}$$

$$= \left|\frac{1}{2R}\int_{-R}^{R} OTF \text{ standard lens}(\psi)d\psi\right|$$

$$\simeq \left|\frac{1}{2S}\int_{-\infty}^{\infty} OTF \text{ standard lens}(\psi)d\psi\right|$$

$$= \left|\frac{1}{2S} FTF \text{ standard lens}(0)\right|$$

$$= \frac{\pi}{2Sf_xf_y}\min\{K_x, K_y\}. \tag{5.1}$$

Note that the expression in (5.1) equals the upper bound from Proposition 4. Therefore, we may conclude:

Proposition 5. *Focus sweep is near-optimally transfer-efficient for obtaining a depth-invariant OTF.*

Next, we study the cubic phase plate, in which the phase of the aperture varies according to a cubic monomial:

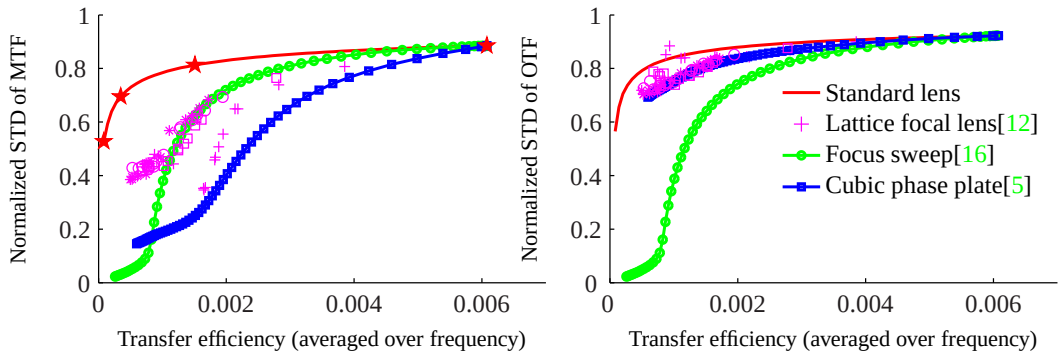


Figure 5. Comparison of various computational cameras in trading off transfer efficiency for depth invariance. The plots are the same as in Figure 4, but the results are aggregated over frequency. The normalized standard variance is computed by averaging the per-frequency variance, normalizing by the average OTF power, and taking the square root.

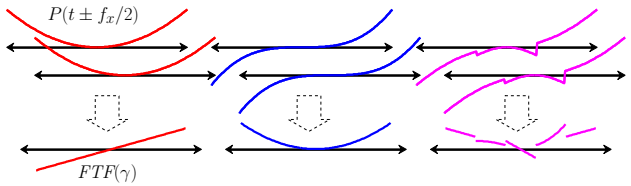


Figure 6. The phase of the FTF for various 1D apertures. **Left**: a standard lens, **Center**: a CPP, **Right**: lattice focal lens. A standard lens has quadratic phase, so shifting it and correlating it with itself yields linear phase. Similarly, the CPP yields quadratic phase. The lattice focal lens, however, yields both in a piecewise fashion.

$P(x,y) = \exp\{i\alpha(x^3+y^3)\}$. Consider the problem of selecting the cubic coefficient α. For simplicity, let the frequency be axis-aligned: $f_x \simeq 0$. Then the FTF converges to a rectangle of width K_y/π and height π/f_y, with quadratic phase whose leading coefficient is $3\alpha f_y$. In comparison, a standard lens would yield a linear phase. See Figure 6.

The average MTF can be computed from the FTF as before. Solely for the purpose of estimating α, we drop the linear phase term in the integrand, since integrating over a quadratic is roughly invariant to shears[5, 24]:

$$\frac{1}{2S}\int_{-S}^{S}\left|\int_{-K_y/(2\pi)}^{K_y/(2\pi)}\frac{\pi}{f_y}\exp\{i3\alpha f_y\gamma^2\}d\gamma\right|d\psi. \tag{5.2}$$

We elect a change of variable so that the domain of integration is independent of K_y:

$$\frac{1}{2S}\int_{-S}^{S}(1-|f_y|)\left|\int_{-1/2}^{1/2}\exp\left\{i\left(\frac{3\alpha f_y K_y^2}{\pi^2}\right)\gamma^2\right\}d\gamma\right|d\psi. \tag{5.3}$$

Because the domain of integration is fixed, there is a single optimal value $\hat{\beta}$ as the leading coefficient of the quadratic:

$$\hat{\beta} := \text{argmax}_{\hat{\beta}}\int_{-S}^{S}\left|\int_{-1/2}^{1/2}\exp\left\{i\hat{\beta}\gamma^2\right\}d\gamma\right|d\psi. \tag{5.4}$$

Consequently, (5.3) is maximized if $3\alpha f_y K_y^2/\pi^2 = \hat{\beta}$, or,

$$\alpha = (\hat{\beta}\pi^2)/(3K_y^2 f_y). \tag{5.5}$$

However, because α is a fixed constant (thus independent of f_y), (5.5) can hold only at a single frequency. Conversely, for each frequency, there exists a value of α that is optimal.

Furthermore, the cubic phase plate is fully efficient at axis-aligned frequencies[12]: its FTF achieves the full trapezoid. Since it yields depth-invariant MTF, it must be optimal under our criteria. In case (f_x, f_y) is not axis-aligned, rotating the phase function yields analogous results. We observed previously in Figure 3 that such "tuned" cubic phase plate nearly achieves the upper bound.

Proposition 6. *For each frequency (f_x, f_y), there exists a "tuned" cubic phase plate that is near-optimally transfer-efficient in obtaining depth-invariant MTF at this frequency.*

We remark that using the worst-case MTF[2, 12] instead to measure performance would not alter our conclusion:

$$\begin{aligned}\text{minimum MTF of a lens} &\leq \text{average MTF of the lens}\\ &\leq \text{average MTF of cubic phase plate}\\ &\simeq \text{minimum MTF of cubic phase plate.}\end{aligned}$$

Lastly, we consider the lattice focal lens, a collection of small lenslets focused at various depths. The phase of its aperture is a piecewise quadratic, so the integrand of the FTF becomes a complex exponential of piecewise linear/quadratic phase (Figure 6). Thus we expect its behavior to be between that of a standard lens and the CPP, as shown in Figure 7: it is more transfer-efficient than the CPP at non-axis-aligned low frequencies, but much less depth-invariant.

6. Conclusion

We have quantified the tradeoff between transfer efficiency and depth invariance for a general family of computational cameras. We first characterized the structure of the OTF of an arbitrary spatiotemporally modulated aperture at any given frequency (f_x, f_y) using its Fourier dual with respect to misfocus, namely the FTF. Using this newfound structure of the OTF, we advanced mathematical claims about the tradeoff, including, to our knowledge, the tightest bounds yet derived on the average OTF and MTF. We were also able to explain the tradeoff observed in various existing

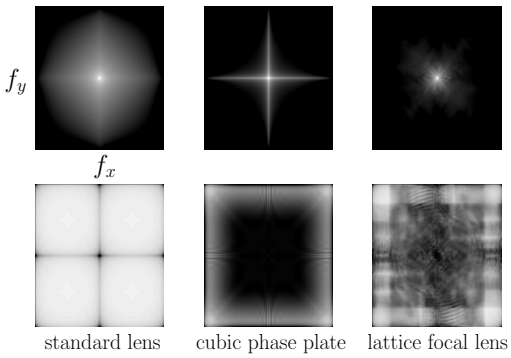


Figure 7. Comparison of the cubic phase plate and lattice focal lens. Here $\psi \in [-150, 150]$. The CPP uses $\alpha = 200$, and lattice focal lens has 4-by-4 lenslets with equally spaced focal power in the desired range. **Top**: the average power as a function of frequency, plotted on a log scale. **Bottom**: the average normalized variance as a function of frequency. The standard lens exhibits maximal power but no depth invariance; the CPP shows depth invariance; the lattice focal lens falls in between.

designs, and show that some are already near-optimal under our metric.

While we saw that there exists a near-optimal CPP for each frequency (Fig. 3 and Prop. 6), it remains unclear whether there is a single lens optimal for producing depth-invariant MTF at every frequency. We conjecture that such design is impossible under the current assumptions.

Future work may consider the same theoretical problem of analyzing OTFs and depth invariance under a weaker set of assumptions. For instance, we may relax the assumptions on image formation to permit multiple captures from the same viewpoint, in the manner of [8]. This may lead to looser theoretical bounds than imposed on a single-capture system. Another direction of future work calls for incorporating a specific deconvolution method and camera noise model, and seeing if interesting mathematical claims can still be advanced on the relationship among the OTF, depth invariance and the resulting image fidelity. This will render the work more numerical in nature, but will reveal more practical insights on computational cameras.

Acknowledgment I would like to thank Zhengyun Zhang, Andrew Adams, Marc Levoy and the anonymous reviewers for their helpful comments and discussions.

References

[1] J. Baek. Fast computation of the OTFs for various computational cameras. Technical report, Stanford University, 2010. Supplied as additional material `supplement.pdf`.

[2] S. Bagheri, P. E. X. Silveira, and G. Barbastathis. Signal-to-noise-ratio limit to the depth-of-field extension for imaging systems with an arbitrary pupil function. *J. Opt. Soc. Am. A*, 26(4), 2009.

[3] H. Bartelt, J. Ojeda-Castañeda, and E. S. Enrique. Misfocus tolerance seen by simple intersection of the ambiguity function. *Applied Optics*, 23(16), 1984.

[4] R. Bracewell. *The Fourier Transform & Its Applications*. McGraw-Hill, 1999.

[5] E. R. Dowski and W. T. Cathey. Extended depth of field through wave-front coding. *Applied Optics*, 34(11), 1995.

[6] N. George and W. Chi. Extended depth of field using a logarithmic asphere. *J. Opt. A: Pure Appl. Optics*, 5, 2003.

[7] J. W. Goodman. *Introduction to Fourier Optics*. Roberts & Company, 2005.

[8] S. W. Hasinoff, K. N. Kutulakos, F. Durand, and W. T. Freeman. Time-constrained photography. In *IEEE ICCV*, 2009.

[9] G. Hausler. A method to increase the depth of focus by two step image processing. *Opt. Commun.*, 6(1), 1972.

[10] H. H. Hopkins. The frequency response of a defocused optical system. *Proceeding of the Royal Society of London. Series A. Mathematical and Physical*, 231(1184), 1955.

[11] A. Levin, R. Fergus, F. Durand, and W. T. Freeman. Image and depth from a conventional camera with a coded aperture. In *ACM SIGGRAPH*, 2007.

[12] A. Levin, S. W. Hasinoff, P. Green, F. Durand, and W. T. Freeman. 4D frequency analysis of computational cameras for depth of field extension. In *ACM SIGGRAPH*, 2009.

[13] C. D. Manning, P. Raghavan, and H. Schütze. *Introduction to Information Retrieval*. Cambridge University Press, 2008.

[14] L. Mertz. *Transformations in Optics*. John Wiley & Sons, Inc., 1965.

[15] M. Mino and Y. Okano. Improvement in the optical transfer function of a defocused optical system through the use of shaded apertures. *Applied Optics*, 10(10), 1971.

[16] H. Nagahara, S. Kuthirummal, C. Zhou, and S. Nayar. Flexible depth of field photography. In *ECCV*, 2008.

[17] R. Ng, M. Levoy, M. Bréif, G. Duval, M. Horowitz, and P. Hanrahan. Light field photography with a hand-held plenoptic camera. Technical report, Stanford University, 2005.

[18] J. Ojeda-Castañeda, E. Tepichin, and A. Diaz. Arbitrary high focal depth with a quasioptimum real and positive transmittance apodizer. *Applied Optics*, 28(13), 1989.

[19] A. Papoulis. Ambiguity function in fourier optics. *J. Opt. Soc. Am.*, 64, 1974.

[20] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *IJCV*, 47(1-3), 2002.

[21] S. S. Sherif, W. T. Cathey, and E. R. Dowski. Phase plate to extend the depth of field of incoherent hybrid imaging systems. *Applied Optics*, 43(13), 2004.

[22] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin. Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing. In *ACM SIGGRAPH*, 2007.

[23] N. Young. *An introduction to Hilbert Space*. Cambridge University Press, 1988.

[24] Z. Zhang and M. Levoy. Wigner distributions and how they relate to the light field. In *IEEE ICCP*, 2009.

[25] C. Zhou and S. Nayar. What are good apertures for defocus deblurring? In *IEEE ICCP*, 2009.