Analyzing Depth from Coded Aperture Sets

Anat Levin

Dep. of Computer Science and Applied Math, The Weizmann Institute of Science

Abstract. Computational depth estimation is a central task in computer vision and graphics. A large variety of strategies have been introduced in the past relying on viewpoint variations, defocus changes and general aperture codes. However, the tradeoffs between such designs are not well understood. Depth estimation from computational camera measurements is a highly non-linear process and therefore most research attempts to evaluate depth estimation strategies rely on numerical simulations. Previous attempts to design computational cameras with good depth discrimination optimized highly non-linear and non-convex scores, and hence it is not clear if the constructed designs are optimal. In this paper we address the problem of depth discrimination from J images captured using J arbitrary codes placed within one fixed lens aperture. We analyze the desired properties of discriminative codes under a geometric optics model and propose an *upper bound* on the best possible discrimination. We show that under a multiplicative noise model, the half ring codes discovered by Zhou et al. [1] are near-optimal. When a large number of images are allowed, a multi-aperture camera [2] dividing the aperture into multiple annular rings provides near-optimal discrimination. In contrast, the plenoptic camera of [5] which divides the aperture into compact support circles can achieve at most 50% of the optimal discrimination bound.

1 Introduction

Estimating scene depth from image measurements is a central goal of computer vision research. Historical depth estimation strategies utilize viewpoint or defocus cues. This includes stereo [3] and plenoptic cameras [4, 5], depth from focus and depth from defocus techniques [6, 2]. Recent computational cameras combine and extend these strategies using coded apertures [7, 8, 1] and phase masks [9, 10].

The large variety of computational cameras calls for a systematic way to compare their performance and understand limitations. However, despite the large amount of research on the subject, the problem is far from being understood. Historical comparisons between stereo and DFD approaches have attracted a lot of research [11–13], leading to different conclusions based on different experimental setups. An important analytic analysis is proposed by Schechner and Kiryati [11], who point that to compare the two approaches the same physical dimensions should be used, and the stereo baseline should be equal to the lens aperture. Despite the important contribution of their analysis, many open questions remain. In particular, they model noise sensitivity on a per-frequency basis and do not analyze how the discrimination information is combined over multiple frequencies.

The problem of computational camera analysis is gaining increased attention [11, 14–17, 1, 10]. Recently, success has been achieved in understanding the related problem of removing defocus blur given depth [17, 10, 16]. However, while the accuracy of depth estimation has an important effect on the quality of defocus deblurring, depth discrimination accuracy is often omitted from the analysis [17, 10], or evaluated numerically only [16, 14]. Some authors [1, 8] address the depth discrimination problem directly and propose explicit scores for discrimination accuracy. However, depth discrimination is a non linear process and the proposed discrimination scores are highly non linear as well. Specifically, Zhou et al. [1] searched for discriminative codes using a genetic algorithm. While the discovered codes are interesting, it is not clear if they are optimal since there is no way to test whether the global optimum of the discrimination score was reached by the optimization algorithm. This optimization also did not offer concrete understanding of the characteristics of good codes.

In this manuscript we address depth discrimination in a setting similar to [1]. One is allowed to capture $J \geq 2$ images of a scene using J different aperture code masks. All images, however, are taken by a fixed static camera, using a standard lens at a fixed focus setting. The aperture codes are allowed to have any general shape within the maximal lens aperture width. This problem formulation is general enough to cover most existing depth estimation strategies. For example, depth from defocus can be expressed using disc aperture masks with different widths, and stereo can be expressed using code masks with holes allowing light at opposite ends of the aperture. We note that in this setting all designs have the same physical dimensions since they are bounded within the same maximal aperture which provides an upper bound on their discrimination performance. We restrict the discussion to the geometric optics model and our results are valid only up to the extent at which this model is valid.

We ask what is the quality of depth discrimination that a given codes set can provide, and what are the best results one can hope to achieve. We build on the discrimination score derived by [1] but notice that it can be analyzed analytically using the derivatives of the code spectra, and larger variations improve discrimination. We use Parseval's theorem and the fact that the primal support of the codes is bounded to show that the maximal derivative power is bounded. This analysis allows us to derive an analytic *upper bound* on the maximal possible depth discrimination. It also provides an understanding of the desired properties of good aperture codes. We use this to analyze existing depth discrimination strategies. For the case of $J = 2$ images and multiplicative noise, we show that the half ring codes discovered by [1] are indeed near optimal. When a large number J of images is allowed, near optimal discrimination can be obtained when the aperture is divided to multiple annular rings, along the lines of the multi-aperture camera of [2]. In contrast, dividing the aperture into small compact squares or circles as done by the plenoptic camera of [5] can achieve no more than 50% of the discrimination bound.

2 Depth Discrimination from Coded Apertures

Problem formulation: We consider a camera with a standard lens focused at depth d_0. J images are captured via J codes which block light at different regions of the aperture. The codes can be expressed as J functions $a^1, \ldots, a^J$ bounded within a maximal aperture radius R. That is: $\forall x, y\ 0 \leq a^j_{(x,y)} \leq 1,\ a^j_{(x,y)} = 0$ if $|x|^2 + |y|^2 > R^2$. We also restrict the J codes to disjoint parts of the aperture so that the J images can be captured simultaneously (e.g. [5, 2]). That is, for each x, y there is a single j for which $a^j_{(x,y)} > 0$. However, many of our results hold without the disjointness requirement.

Let I denote an ideal sharp version of a scene, and assume we observe J images $B^1, \ldots, B^J$. If the depth is locally constant the observed images of an object at depth d can be described as a local convolution of the sharp version I with Point Spread Functions (PSFs) ϕ^d. In the frequency domain the imaging is expressed as multiplication with an Optical Transfer Function (OTF) $\hat{\phi}^d$: [1]

$$\hat{B}^j_{\omega_{x,y}} = \hat{\phi}^{d,j}_{\omega_{x,y}} \hat{I}_{\omega_{x,y}} + n^j_{\omega_{x,y}}, \tag{1}$$

where $\omega_{x,y} = (\omega_x, \omega_y)$ denote spatial frequency coordinates and n^j is an imaging noise. Through this paper we index depth using the corresponding light field slope $s = (d - d_0)/d$, since the PSF and OTF vary as a linear function of s and not of d. Using a geometric optics model, it was shown [7, 8] that the PSFs and OTFs are scaled versions of the aperture codes:

$$\hat{\phi}^{s,j}_{\omega_{x,y}} = \hat{a}^j_{(s \cdot \omega_{x,y})} \tag{2}$$

The scaled PSF model, however, does not take into account wave optics effects. The geometric optics model is a reasonable approximation to the true optics when the holes in the code are not too small. The optimality arguments in this paper are only valid to the extant at which the geometric model is valid.

Noise model: We follow the affine noise model (e.g. [16]). For simplicity this model assumes the noise is a zero mean Gaussian whose variance η^2 is constant over the image, and it is a combination of a constant additive term, the read noise, and a multiplicative term, the photon noise:

$$(\eta^j)^2 = \alpha^j \eta^2_{mult} + \eta^2_{add}, \tag{3}$$

where α^j is the amount of light gathered by the j'th aperture $\alpha^j = \iint a^j d\omega_x d\omega_y$. For modern sensors under good illumination conditions the noise is dominated mostly by the multiplicative term and $(\eta^j)^2 \approx \alpha^j \eta^2_{mult}$. We often assume that all J apertures have equal area and omit the j index from α, η.

Given a set of observed images $B^1, \ldots, B^J$, estimating the scene depth is equivalent to estimating at every local image window a slope index s such that locally B^j was blurred with $\phi^{s,j}$. Our goal in this paper is to analyze the quality of depth discrimination that a set of aperture codes can provide. We note that all codes are bounded within the same maximal aperture radius R and the maximal radius is the parameter upper-bounding the performance of all codes. We start with a brief review of the depth discrimination score proposed in [1]. In Section 3, we derive bounds on this discrimination score and study the desired properties of optimal discrimination codes.

[1] Through this manuscript we use $\hat{}$ to denote the Fourier transform of the corresponding signal.

2.1 Descrimination Score

One often assumes a zero mean Gaussian prior on the sharp signal $\hat{I}$. For simplicity the classical $1/f^2$ law is used and the variance in frequency $\omega_{x,y}$ is set to $\sigma^2_{\omega_{x,y}} = 1/|\omega_{x,y}|^2$. We denote with bold fonts the J dimensional vectors $\hat{\boldsymbol{B}}_{\omega_{x,y}} = [\hat{B}^1_{\omega_{x,y}}, \ldots, \hat{B}^J_{\omega_{x,y}}]$, $\hat{\boldsymbol{\phi}}^s_{\omega_{x,y}} = [\hat{\phi}^{s,1}_{\omega_{x,y}}, \ldots, \hat{\phi}^{s,J}_{\omega_{x,y}}]$, $\hat{\boldsymbol{a}}_{\omega_{x,y}} = [\hat{a}^1_{\omega_{x,y}}, \ldots, \hat{a}^J_{\omega_{x,y}}]$. The probability of the observed images factorizes as an independent product over individual frequencies $P(\hat{B}^1, \ldots, \hat{B}^J) = \prod_{\omega_{x,y}} P(\hat{\boldsymbol{B}}_{\omega_{x,y}})$. Since $\hat{\boldsymbol{B}}_{\omega_{x,y}}$ is obtained from $\hat{I}$ as a linear transformation plus Gaussian noise, $P(\hat{\boldsymbol{B}}_{\omega_{x,y}})$ follows a Gaussian distribution with covariance $\Psi^s_{\omega_{x,y}} = \hat{\boldsymbol{\phi}}^s_{\omega_{x,y}} \sigma^2_{\omega_{x,y}} \hat{\boldsymbol{\phi}}^{s*}_{\omega_{x,y}} + \eta^2 \mathbb{I}$, where $\mathbb{I}$ denotes a $J \times J$ identity matrix and * denotes the conjugate transpose. Let $U^s_{\omega_{x,y}}$ be a $J \times (J-1)$ matrix completing $\hat{\boldsymbol{\phi}}^s_{\omega_{x,y}}$ to an orthogonal basis, that is, the $J \times J$ matrix $\tilde{U}^s_{\omega_{x,y}} = [\frac{\hat{\boldsymbol{\phi}}^s_{\omega_{x,y}}}{|\hat{\boldsymbol{\phi}}^s_{\omega_{x,y}}|}, U^s_{\omega_{x,y}}]$ is orthogonal. We can then express:

$$\Psi^s_{\omega_{x,y}} = \tilde{U}^s_{\omega_{x,y}} D\left(|\hat{\boldsymbol{\phi}}^s_{\omega_{x,y}}|^2 \sigma^2_{\omega_{x,y}} + \eta^2, \eta^2, \ldots, \eta^2\right) \tilde{U}^*_{\omega_{x,y}}, \tag{4}$$

$${\Psi^s_{\omega_{x,y}}}^{-1} = \tilde{U}^s_{\omega_{x,y}} D\left(\frac{1}{|\hat{\boldsymbol{\phi}}^s_{\omega_{x,y}}|^2 \sigma^2_{\omega_{x,y}} + \eta^2}, \frac{1}{\eta^2}, \ldots, \frac{1}{\eta^2}\right) \tilde{U}^*_{\omega_{x,y}}, \tag{5}$$

where $D(\cdots)$ denotes a diagonal matrix. We assume the signal variance is sufficiently above the noise level: $|\hat{\boldsymbol{\phi}}^s_{\omega_{x,y}}|\sigma^2_{\omega_{x,y}} > \eta^2$ (other frequencies provide little discrimination and their contribution can be ignored). Ψ allows high variance along the OTFs direction $\hat{\boldsymbol{\phi}}_{\omega_{x,y}}$ and low variance at all orthogonal directions. The expected negative log likelihood of an observation whose correct depth is s_0 under possible explanation s is:

$$E_{P(B|s_0)}[-2\log P(B|s)] = E_{P(B|s_0)}[B^* \Psi^s B] + \log|\Psi^s| \tag{6}$$

$$= \sum\nolimits_{\omega_{x,y}} \left[\frac{|\hat{\boldsymbol{\phi}}^{s_0}_{\omega_{x,y}}|^2 \sigma^2_{\omega_{x,y}} + \eta^2}{|\hat{\boldsymbol{\phi}}^{s_0}_{\omega_{x,y}}|^2 \eta^2} \left| U^{s*}_{\omega_{x,y}} \hat{\boldsymbol{\phi}}^{s_0}_{\omega_{x,y}} \right|^2 + \frac{\eta^2}{|\hat{\boldsymbol{\phi}}^s_{\omega_{x,y}}|^2 (|\hat{\boldsymbol{\phi}}^s_{\omega_{x,y}}|^2 \sigma^2 + \eta^2)} \left| \hat{\boldsymbol{\phi}}^{s*}_{\omega_{x,y}} U^{s_0}_{\omega_{x,y}} \right|^2 \right.$$

$$\left. + \frac{|\hat{\boldsymbol{\phi}}^{s_0}_{\omega_{x,y}}|^2 \sigma^2_{\omega_{x,y}} + \eta^2}{|\hat{\boldsymbol{\phi}}^s_{\omega_{x,y}}|^2 |\hat{\boldsymbol{\phi}}^{s_0}_{\omega_{x,y}}|^2 (|\hat{\boldsymbol{\phi}}^s_{\omega_{x,y}}|^2 \sigma^2_{\omega_{x,y}} + \eta^2)} \left| \hat{\boldsymbol{\phi}}^{s*}_{\omega_{x,y}} \hat{\boldsymbol{\phi}}^{s_0}_{\omega_{x,y}} \right|^2 + \left| U^{s*}_{\omega_{x,y}} U^{s_0}_{\omega_{x,y}} \right|^2 + \log|\Psi^s_{\omega_{x,y}}| \right] \tag{7}$$

$$\approx \sum\nolimits_{\omega_{x,y}} \frac{\sigma^2_{\omega_{x,y}}}{\eta^2} \left| U^{s*}_{\omega_{x,y}} \hat{\boldsymbol{\phi}}^{s_0}_{\omega_{x,y}} \right|^2 \tag{8}$$

where the approximation of Eq. (8) follows from the fact that Eq. (7) is dominated by the first term when the noise is small relative to the signal, and the $\log|\Psi^s_{\omega_{x,y}}|$ term is relatively constant. That means that a good discrimination is obtained when $\hat{\boldsymbol{\phi}}^s$ and $\hat{\boldsymbol{\phi}}^{s_0}$ are as orthogonal as possible.

The discrimination score in Eq. (8) is a simple extension of the one derived by [1] from the two image case to the J image case. In [1] the discrimination score was evaluated discretely over a sample of s values. Since discrimination is usually most challenging at the neighborhood of the true solution[2] we propose to replace the discrete sample with

[2] This model does not penalize the symmetry of the PSF in front and behind the focus depth

an analytic derivative of the OTF as a function of depth: $\nabla_s \hat{\phi}^s_{\omega_{x,y}} = \frac{\partial \hat{\boldsymbol{\phi}}^s_{\omega_{x,y}}}{\partial s}$ ($\nabla_s \hat{\phi}^s_{\omega_{x,y}}$ is a J-dimensional vector). We note that since $U^{s*}_{\omega_{x,y}} \hat{\boldsymbol{\phi}}^s_{\omega_{x,y}} = 0$, $U^{s*}_{\omega_{x,y}} \hat{\boldsymbol{\phi}}^{s_0}_{\omega_{x,y}} = U^{s*}_{\omega_{x,y}} (\hat{\boldsymbol{\phi}}^{s_0}_{\omega_{x,y}} - \hat{\boldsymbol{\phi}}^s_{\omega_{x,y}}) \approx U^{s*}_{\omega_{x,y}} \nabla_s \hat{\phi}^s_{\omega_{x,y}}$. We denote by $\mathfrak{D}^s_{\omega_{x,y}}(\hat{\boldsymbol{\phi}})$ the local discrimination score at frequency $\omega_{x,y}$:

$$\mathfrak{D}^s_{\omega_{x,y}}(\hat{\boldsymbol{\phi}}) = \frac{\sigma^2_{\omega_{x,y}}}{\eta^2} \left| U^{s*}_{\omega_{x,y}} \nabla_s \hat{\phi}^s_{\omega_{x,y}} \right|^2 = \frac{\sigma^2_{\omega_{x,y}}}{\eta^2} \left(\left| \nabla_s \hat{\phi}^s_{\omega_{x,y}} \right|^2 - \frac{1}{|\hat{\boldsymbol{\phi}}^s_{\omega_{x,y}}|^2} \left| \hat{\boldsymbol{\phi}}^{s*}_{\omega_{x,y}} \nabla_s \hat{\phi}^s_{\omega_{x,y}} \right|^2 \right), \tag{9}$$

which implies that discrimination is maximized when there is a large variation of the OTFs as a function of s, in the direction orthogonal to $\hat{\boldsymbol{\phi}}^s$. We wish to find OTFs maximizing discrimination integrated over all frequencies (up to spatial resolution Ω):

$$\mathfrak{D}^s(\hat{\boldsymbol{\phi}}) = \int_{-\Omega}^{\Omega} \int_{-\Omega}^{\Omega} \mathfrak{D}^s_{\omega_{x,y}}(\hat{\boldsymbol{\phi}}) d\omega_x d\omega_y. \tag{10}$$

3 Discrimination Budget

To understand how to maximize the discrimination score we study some of the physical constraints on the OTF. Let a^j , $\hat{a}^j$ denote the aperture code of the j'th view and its Fourier transform. The PSF and OTF are known to be scaled versions of the aperture code [7, 8]:$\hat{\phi}^{s,j}_{\omega_{x,y}} = \hat{a}^j_{(s \cdot \omega_{x,y})}$. This implies that

$$\nabla_s \hat{\phi}^{s,j}_{\omega_{x,y}} = |\omega_{x,y}| \nabla_{\omega_{x,y}} \hat{a}^j_{(s\omega_{x,y})} = |\omega_{x,y}| \left(\frac{\omega_x}{|\omega_{x,y}|} \frac{\partial \hat{a}^j}{\partial \omega_x} + \frac{\omega_y}{|\omega_{x,y}|} \frac{\partial \hat{a}^j}{\partial \omega_y} \right). \tag{11}$$

The $|\omega_{x,y}|$ factor multiplying the derivative in Eq. (11) is canceled by $\sigma^2_{\omega_{x,y}} = 1/|\omega_{x,y}|^2$ in Eq. (9), and the discrimination score of Eq. (9) can be expressed as a function of the derivatives of the aperture spectra $\hat{a}^j$ weighting all entries equally:

Definition 1. Consider a set of aperture codes $\hat{\boldsymbol{a}}$. The *depth discrimination score at spatial frequency* $\omega_{x,y}$ is defined as

$$\mathfrak{D}_{\omega_{x,y}}(\hat{\boldsymbol{a}}) = \frac{1}{\eta^2} \left(|\nabla_{\omega_{x,y}} \hat{\boldsymbol{a}}_{\omega_{x,y}}|^2 - \frac{\left| \hat{\boldsymbol{a}}^*_{\omega_{x,y}} \nabla_{\omega_{x,y}} \hat{\boldsymbol{a}}_{\omega_{x,y}} \right|^2}{|\hat{\boldsymbol{a}}_{\omega_{x,y}}|^2} \right) \tag{12}$$

and the *total discrimination score* at depth s as

$$\mathfrak{D}^s(\hat{\boldsymbol{a}}) = \frac{1}{s^2} \int_{-s\Omega}^{s\Omega} \int_{-s\Omega}^{s\Omega} \mathfrak{D}_{\omega_{x,y}}(\hat{\boldsymbol{a}}) d\omega_x d\omega_y. \tag{13}$$

We note that the discrimination around depth s is integrated up to a cut-off frequency $s\Omega$ but multiplied by the density $1/s^2$. We often omit the s index from Eq. (13) and consider the case $s = 1$ which is usually the more challenging case. (the $1/s^2$ factor

improves the score of low s values and in addition, there is usually more energy at the low frequencies).

The discrimination score cannot be made arbitrarily large. Our goal is to show that the best possible discrimination score is bounded and then understand the desired properties of aperture codes with optimal discrimination. We define two useful quantities:

Definition 2. The *center-oriented derivative power at frequency* $\omega_{x,y}$ is

$$\mathfrak{C}_{\omega_{x,y}}(\hat{\boldsymbol{a}}) = \frac{1}{\eta^2}\sum_j \left|\nabla_{\omega_{x,y}}\hat{a}^j_{\omega_{x,y}}\right|^2 = \frac{1}{\eta^2}\sum_j \left|\frac{\omega_x}{|\omega_{x,y}|}\frac{\partial \hat{a}^j}{\partial \omega_x} + \frac{\omega_y}{|\omega_{x,y}|}\frac{\partial \hat{a}^j}{\partial \omega_y}\right|^2, \qquad (14)$$

the *total center-oriented derivative power* is $\mathfrak{C}^s(\hat{\boldsymbol{a}}) = \frac{1}{s^2}\int_{-s\Omega}^{s\Omega}\int_{-s\Omega}^{s\Omega}\mathfrak{C}_{\omega_{x,y}}(\hat{\boldsymbol{a}})d\omega_x d\omega_y$.

Definition 3. The *gradient power at frequency* $\omega_{x,y}$ is

$$\mathfrak{G}_{\omega_{x,y}}(\hat{\boldsymbol{a}}) = \frac{1}{\eta^2}\sum_j \left|\nabla \hat{a}^j_{\omega_{x,y}}\right|^2 = \frac{1}{\eta^2}\sum_j \left[\left|\frac{\partial \hat{a}^j}{\partial \omega_x}\right|^2 + \left|\frac{\partial \hat{a}^j}{\partial \omega_y}\right|^2\right] \qquad (15)$$

and the *total gradient power* is $\mathfrak{G}^s(\hat{\boldsymbol{a}}) = \frac{1}{s^2}\int_{-s\Omega}^{s\Omega}\int_{-s\Omega}^{s\Omega}\mathfrak{G}_{\omega_{x,y}}(\hat{\boldsymbol{a}})d\omega_x d\omega_y$.

Note that while we use the terms derivative and gradient, the quantities in Eqs. (14) and (15) are actually normalized by the noise variance η^2. The definition of the discrimination score in Eq. (12) implies that it is bounded by the center-oriented derivative power $\mathfrak{D}_{\omega_{x,y}}(\hat{\boldsymbol{a}}) \leq \mathfrak{C}_{\omega_{x,y}}(\hat{\boldsymbol{a}})$. Also the oriented derivative power is bounded by the gradient power $\mathfrak{C}_{\omega_{x,y}}(\hat{\boldsymbol{a}}) \leq \mathfrak{G}_{\omega_{x,y}}(\hat{\boldsymbol{a}})$. We observe that the total gradient power cannot be made arbitrarily high. We show that there is a fixed energy budget which is determined by the primal support of a^j, and this budget is preserved in the frequency domain as a consequence of Parseval's theorem. Bounding the gradient power provides a bound on the best possible discrimination score.

Claim 1 *Let $a^1, \ldots a^J$ be a set of disjoint codes inside an aperture of radius R. Their total gradient power is bounded and satisfies*

$$\sum_{j=1}^{J}\frac{1}{\eta^2}\iint\left|\nabla\hat{a}^j_{\omega_{x,y}}\right|^2 d\omega_x d\omega_y = \sum_{j=1}^{J}\frac{1}{\eta^2}\iint\left|a^j_{(x,y)}\right|^2(x^2+y^2)dxdy \;\leq\; \mathfrak{B} \qquad (16)$$

for

$$\mathfrak{B} = \max_{0<\lambda<R}\frac{\frac{\pi}{2}\left(R^4-(R-\lambda)^4\right)}{\frac{\pi}{J}(R^2-(R-\lambda)^2)\eta^2_{mult}+\eta^2_{add}} \qquad (17)$$

Proof. Differentiating $\hat{a}^j$ can be expressed as a convolution of $\hat{a}^j$ with derivative filters f_{ω_x}, such that $\nabla_{\omega_x}\hat{a}^j = f_{\omega_x}\otimes\hat{a}^j$. In the primal domain this convolution translates to multiplication $\hat{f}_{(x,y)}\cdot a^j_{(x,y)}$. The Fourier transform of an ideal derivative filter is $|\hat{f}_{(x,y)}| = |x|$. Parseval's theorem implies that the frequency derivatives power is preserved in the primal domain, and thus:

$$\iint\left|f_{\omega_x}\otimes\hat{a}^j\right|^2 d\omega_x d\omega_y = \iint|a^j_{(x,y)}|^2x^2dxdy. \qquad (18)$$

A similar property applies for $\nabla_{\omega_y}\hat{a}^j$. Therefore

$$\frac{1}{\eta^2}\iint\left|\nabla\hat{a}^j_{\omega_{x,y}}\right|^2 d\omega_x d\omega_y = \frac{1}{\eta^2}\iint\left|a^j_{(x,y)}\right|^2 (x^2+y^2)dxdy. \tag{19}$$

We now ask what is the maximal value that the RHS of Eq. (19) can obtain. If a^j is open over a large area the integral value is increased, but if $\eta^2_{mult} > 0$, a larger aperture area also increases the noise (see Eq. (3)). Thus, the optimal aperture area should trade discrimination v.s. noise, and depends on the ratio between η^2_{mult} to η^2_{add}. However, for every fix aperture area, we can ask what is the maximal value of the integral in the RHS of Eq. (19), among all codes with the same fixed area. The highest value is obtained by a ring attached to the aperture boundaries, since the x^2+y^2 values averaged in Eq. (19) are large when they are adjacent to the aperture boundary. Let r^λ denote a code open at an outer ring of width λ, $r^\lambda_{(x,y)} = 1$ iff $(R-\lambda)^2 \le x^2+y^2 \le R^2$. The gradient power of the code r^λ is the highest among all codes with the same area. Standard calculus implies that the area and gradient power of a ring r^λ are

$$\iint r^\lambda_{(x,y)}dxdy = \pi\left(R^2-(R-\lambda)^2\right), \quad \iint r^\lambda_{(x,y)}(x^2+y^2)dxdy = \frac{\pi}{2}\left(R^4-(R-\lambda)^4\right). \tag{20}$$

Therefore the gradient power of a code is bounded by:

$$\frac{1}{\eta^2}\iint\left|\nabla\hat{a}^j_{\omega_{x,y}}\right|^2 d\omega_x d\omega_y \le \max_{0<\lambda<R}\frac{\frac{\pi}{2}\left(R^4-(R-\lambda)^4\right)}{\pi(R^2-(R-\lambda)^2)\eta^2_{mult}+\eta^2_{add}}. \tag{21}$$

For J disjoint codes, the best gradient power is obtained when their union forms an outer ring and Eq. (16) follows. □

Corollary 1. *The total depth discrimination score, oriented derivative power and gradient power are bounded in the following order*

$$\mathfrak{D}(\hat{\boldsymbol{a}}) \le \mathfrak{C}(\hat{\boldsymbol{a}}) \le \mathfrak{G}(\hat{\boldsymbol{a}}) \le \mathfrak{B}. \tag{22}$$

Optimal discrimination codes should therefore satisfy the following three properties:

1. *The total gradient power should approach the bound* $\mathfrak{G}(\hat{\boldsymbol{a}}) \to \mathfrak{B}$.
2. *For every* ω_x, ω_y*, the oriented derivative power should approach the gradient power* $\mathfrak{C}_{\omega_{x,y}}(\hat{\boldsymbol{a}}) \to \mathfrak{G}_{\omega_{x,y}}(\hat{\boldsymbol{a}})$
3. *For every* ω_x, ω_y*, the discrimination score should approach the oriented derivative power* $\mathfrak{D}_{\omega_{x,y}}(\hat{\boldsymbol{a}}) \to \mathfrak{C}_{\omega_{x,y}}(\hat{\boldsymbol{a}})$

To understand the first property note that the proof of Claim 1 implies that the gradient power is maximized when the codes let in light at a ring at the periphery of the aperture. The exact ring width is a function of the ratio between the multiplicative and additive noise components. If the noise is fully additive ($\eta_{mult} = 0$) it is best to collect light over the entire aperture area. When the noise is mostly multiplicative ($\eta^2 \approx \alpha\eta^2_{mult}$), the best is to have a very narrow ring at the periphery of the aperture.

The oriented derivative power is equal to the gradient power if and only if for all js, the gradient direction at spatial frequency $\omega_{x,y}$ equals $\omega_{x,y}$. Having all gradients oriented toward the center implies that $\hat{a}^j$ should be radially symmetric.

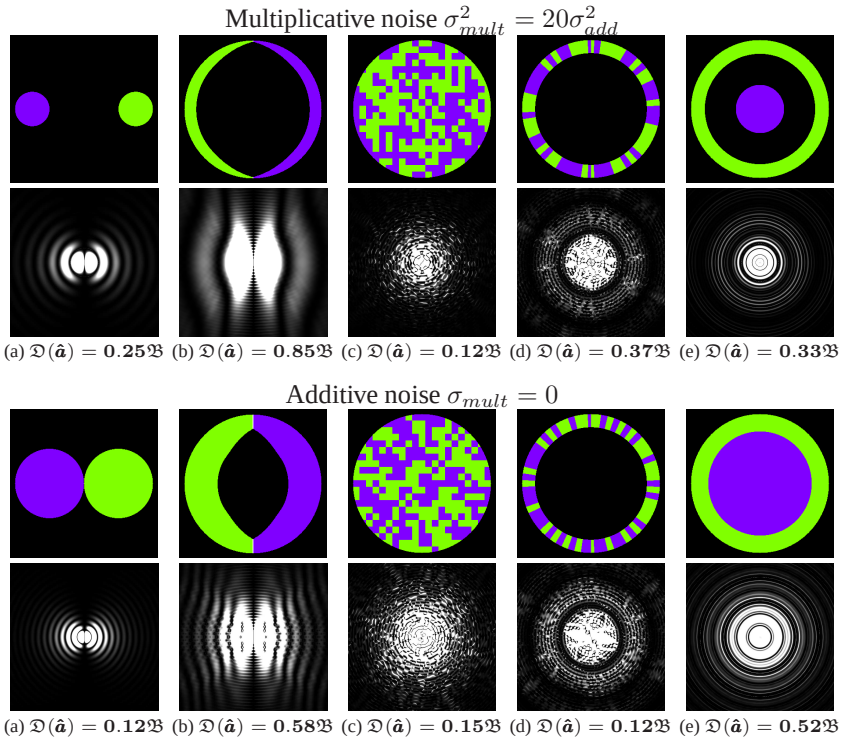


Fig. 1: *Aperture codes and their discrimination score. Top group: multiplicative noise ($\sigma^2_{mult} = 20\sigma^2_{add}$), Lower group: additive noise ($\sigma^2_{mult} = 0$). Each group visualizes in the upper row aperture codes (primal domain) and in the second row their discrimination score $\mathfrak{D}_{\omega_{x,y}}(\hat{\boldsymbol{a}})$ provided in each spatial frequency (frequency domain). The portion of the upper bound utilized by $\mathfrak{D}(\hat{\boldsymbol{a}})$ is reported at the bottom.*

The last property of Corollary 1 requires that the OTFs vector is orthogonal to its derivative. This property is the hardest to analyze, but in the next section we consider examples of codes which approach this requirement.

4 Analyzing Aperture Code Designs

In this section we consider a few aperture code designs and analyze their optimality. We start with designs capturing two images and then move to multiple images.

4.1 Two image designs

We consider the aperture code pairs visualized in Figure 1. The first one is a standard stereo setting with two disc holes. Another design is two halves of a ring (bananas),

inspired by the optimized codes of [1]. We then consider pseudo random codes in each view. We consider random codes in the entire aperture area (Fig 1(c)) and only in the periphery ring (Fig 1(d)). Finally we consider a depth from defocus like pair (Fig 1(e))- an outside ring plus an inner disc.

We consider two noise situations, well illuminated scenes at which $\eta_{mult}^2 = 20\eta_{add}^2$, and the purely additive noise case $\eta_{mult} = 0$. Apart for the random designs, we searched numerically for good parameters in each pair family (e.g. the width of the rings and discs, or the exact shape of the banana parameterized as a spline with 5 keypoints). The parameters optimization is done independently for each noise situation. For multiplicative noise, narrower holes or rings at the periphery are favored and at the additive noise case wide code areas are selected.

For each pair, the second row of Fig 1 visualizes the map of discrimination score at each spatial frequency. At the bottom we report the portion of the bound achieved by the total discrimination scores. The best discrimination is obtained by the two halved rings, which at the multiplicative case utilize 85% of the gradient power upper bound. All other designs utilize at most 50% of the bound. Below we discuss some of the interesting properties of each design.

Note that in designs (a-d) the codes are flipped versions of each other $a^1_{(x,y)} = a^2_{(-x,-y)}$, and hence the spectra are conjugate. To analyze the discrimination scores of code pairs of this type we derive the following lemmas. The first lemma shows that the spectrum derivative consists of the magnitude derivative plus the phase derivative, and for conjugate spectra the portion of the derivative lost in the non orthogonal direction is the derivative of the magnitude. Using this observation we will later aim to show that for discriminative code pairs the magnitude derivative is small relative to the phase.

Lemma 1. *Let $\hat{a}^1, \hat{a}^2$ be a conjugate pair of aperture spectra,*

$$\hat{a}^1_{\omega_{x,y}} = m_{\omega_{x,y}} e^{i\zeta_{\omega_{x,y}}}, \quad \hat{a}^2_{\omega_{x,y}} = m_{\omega_{x,y}} e^{-i\zeta_{\omega_{x,y}}}. \tag{23}$$

where $m_{\omega_{x,y}}$ is the magnitude and $\zeta_{\omega_{x,y}}$ the phase. The discrimination score equals the magnitude times the phase derivative power which is also the derivative power minus the magnitude derivative power.

$$\mathfrak{D}_{\omega_{x,y}}(\hat{\boldsymbol{a}}) = 2\left|m_{\omega_{x,y}}\right|^2 \left|\nabla_{\omega_{x,y}} \zeta_{\omega_{x,y}}\right|^2 = \left|\nabla_{\omega_{x,y}} \hat{\boldsymbol{a}}\right|^2 - 2\left|\nabla_{\omega_{x,y}} |\hat{a}|\right|^2. \tag{24}$$

Proof. The derivatives are the sum of the magnitude derivative and the phase derivative:

$$\nabla_{\omega_{x,y}} \hat{a}^1_{\omega_{x,y}} = e^{i\zeta_{\omega x,y}} \nabla_{\omega_{x,y}} m_{\omega_{x,y}} + m_{\omega_{x,y}} e^{i\zeta_{\omega_{x,y}}} i \nabla_{\omega_{x,y}} \zeta_{\omega_{x,y}} \tag{25}$$

$$\nabla_{\omega_{x,y}} \hat{a}^2_{\omega_{x,y}} = e^{-i\zeta_{\omega x,y}} \nabla_{\omega_{x,y}} m_{\omega_{x,y}} - m_{\omega_{x,y}} e^{-i\zeta_{\omega_{x,y}}} i \nabla_{\omega_{x,y}} \zeta_{\omega_{x,y}}. \tag{26}$$

Since the two phase derivatives have opposite signs they cancel each other when we take the inner product with $\hat{\boldsymbol{a}}$ and we are left with the magnitude derivative only:

$$\frac{\left|\hat{\boldsymbol{a}}_{\omega_{x,y}}{}^* \nabla_{\omega_{x,y}} \hat{\boldsymbol{a}}_{\omega_{x,y}}\right|^2}{\left|\hat{\boldsymbol{a}}_{\omega_{x,y}}\right|^2} = 2\left|\nabla_{\omega_{x,y}} m_{\omega_{x,y}}\right|^2 \tag{27}$$

and Eq. (24) follows from the definition in Eq. (12). □

Next, we note that to analyze the spectra $\hat{\boldsymbol{a}}$ along direction θ we can apply the slicing theorem [18] and look at the projection of the primal codes $\boldsymbol{a}$ in the orthogonal direction. Let $\rho^{j,\theta}$ denote the projection of a^j onto direction θ

$$\rho^{j,\theta}_{(x)} = \int a^j_{(x\overrightarrow{\theta}+y\overrightarrow{\theta}^+)} dy, \tag{28}$$

for $\overrightarrow{\theta} = (cos(\theta), sin(\theta))$, $\overrightarrow{\theta}^+ = (-sin(\theta), cos(\theta))$. Let $|\rho^{j,\theta}|^2 = \int |\rho^{j,\theta}_{(x)}|^2 dx$ denote the total power and $\tau^{j,\theta}$ the power center of mass:

$$\tau^{j,\theta} = \frac{1}{|\rho^{j,\theta}|^2} \int |\rho^{j,\theta}_{(x)}|^2 x \, dx. \tag{29}$$

If a^1, a^2 are a flipped pair, their powers are equal and their mass centers satisfies $\tau^{1,\theta} = -\tau^{2,\theta}$. Therefore we can think of $2|\tau^{j,\theta}|$ as the average disparity along direction θ.

The following lemma shows that if the projection has a relatively wide disparity, there will be a large discrimination in that direction. Figure 2 visualizes projections of the codes in Figure 1(a-d). As we will explain below, the more discriminative codes have a wider disparity in most orientations.

Lemma 2. *Let a^1, a^2 be a pair of flipped aperture codes. The frequency domain discrimination score along direction θ is at least the square of the averaged disparity times the total power*

$$\int \mathfrak{D}_{\omega\overrightarrow{\theta}}(\hat{\boldsymbol{a}}) d\omega \geq \frac{2}{\eta^2} |\tau^{j,\theta}|^2 |\rho^{j,\theta}|^2 \tag{30}$$

Proof. The slicing theorem [18] implies that a slice of $\hat{a}^j$ in direction $\overrightarrow{\theta}$ is the 1D Fourier transform of $\rho^{j,\theta}$. Therefore, we can apply a 1D variant of Claim 1 to compute the total derivative magnitude of $\hat{a}^j$ along direction $\overrightarrow{\theta}$.

$$\int \left| \nabla_{\overrightarrow{\theta}} \hat{a}^j(\omega\overrightarrow{\theta}) \right|^2 d\omega = \int (x\rho^{j,\theta}_{(x)})^2 dx. \tag{31}$$

We note that

$$\int (x\rho^{j,\theta}_{(x)})^2 dx = \int (x-\tau^{j,\theta})^2 (\rho^{j,\theta}_{(x)})^2 dx \; + \; |\tau^{j,\theta}|^2 |\rho^{j,\theta}|^2. \tag{32}$$

Our goal is to show that

$$\int \left| \nabla_{\overrightarrow{\theta}} m_{\omega\overrightarrow{\theta}} \right|^2 d\omega \leq \int (x-\tau^{j,\theta})^2 |\rho^{j,\theta}_{(x)}|^2 dx, \tag{33}$$

and then Eq. (30) will follow directly from Lemma 1 (the factor 2 is because we sum over 2 code spectra). For that note that for any phase $\chi_{\omega\overrightarrow{\theta}}$,

$$\left| \nabla_{\overrightarrow{\theta}} \left(m_{\omega\overrightarrow{\theta}} e^{i\chi_{\omega\overrightarrow{\theta}}} \right) \right|^2 = \left| \nabla_{\overrightarrow{\theta}} m_{\omega\overrightarrow{\theta}} \right|^2 + \left| m_{\omega\overrightarrow{\theta}} \right|^2 \left| \nabla_{\overrightarrow{\theta}} \chi_{\omega\overrightarrow{\theta}} \right|^2 \geq \left| \nabla_{\overrightarrow{\theta}} m_{\omega\overrightarrow{\theta}} \right|^2. \tag{34}$$

In particular, we can choose a phase χ such that $m_{\omega\overrightarrow{\theta}} e^{i\chi_{\omega\overrightarrow{\theta}}}$ will be the Fourier transform of $|\rho^{j,\theta}_{(x-\tau^{j,\theta})}|^2$ (that is, a centered version of ρ). Applying Claim 1 again, the total derivative power in that centered version is

$$\int (x\rho^{j,\theta}_{(x-\tau^{j,\theta})})^2 dx = \int (x-\tau^{j,\theta})^2 (\rho^{j,\theta}_{(x)})^2 dx \tag{35}$$

and Eq. (33) follows. □

Another intuitive argument which follows from the above proof is that when most of the mass of $\rho^{j,\theta}$ is located in a narrow region, but the average disparity is large, the magnitude derivative power is small relative to the total derivative power. This is because the total derivative power is $\int (x\rho^{j,\theta}_{(x)})^2 dx$, and the magnitude derivative power is bounded by $\int (x\rho^{j,\theta}_{(x-\tau^{j,\theta})})^2 dx$. The x^2 values to which $(\rho^{j,\theta}_{(x-\tau^{j,\theta})})^2$ assigns non-zero weight are small compared to the non-zero values of $(\rho^{j,\theta}_{(x)})^2$. Since Lemma 1 shows that the portion of the derivative power lost in the non-orthogonal direction is the magnitude derivative power, if the magnitude derivative power is low, most of the derivative power $\mathfrak{C}(\hat{\boldsymbol{a}})$ contributes to the discrimination score $\mathfrak{D}(\hat{\boldsymbol{a}})$.

We now use Lemmas 1,2 to analyze the code pairs in Fig 1.

Stereo pair: (Fig 1(a)). The main problem with this design is that it provides disparity only in the horizontal direction. The discrimination map in Fig 1(a) is high around the horizontal spatial frequency axis and low around the vertical one. Similarly, the projections in Fig 2(a) show that the average disparity is large in the vertical direction, but reduces as the projection angle changes, with zero disparity horizontally. Stereo violates the second optimal discrimination property of Corollary 1. The spectra are not radially symmetric, but the gradient is mostly in the horizontal direction at all spatial frequencies. To see this, let $m_{\omega_{x,y}}$ denote the spectrum of a disc centered at $(0,0)$. The spectrum of a disc centered at a point $\tau = [\tau_x, \tau_y]$ is a phase shifted version of m:

$$\hat{a}^1_{\omega_{x,y}} = m_{\omega_{x,y}} e^{i(\tau_x\omega_x + \tau_y\omega_y)}, \quad \hat{a}^2_{\omega_{x,y}} = m_{\omega_{x,y}} e^{-i(\tau_x\omega_x + \tau_y\omega_y).} \tag{36}$$

According to Lemma 1, the magnitude derivative does not contribute to the discrimination, and the phase gradient has a constant direction τ. In fact, this implies that the total discrimination score of a stereo pair cannot pass 50% of the total gradient power. To see this, note that the discrimination score at frequency $\omega_{x,y}$ is a function of the inner produce of τ and the direction $\omega_{x,y}/|\omega_{x,y}|$: $\mathfrak{D}_{\omega_{x,y}}(\hat{a}) = \frac{2|m_{\omega_{x,y}}|^2}{|\omega_{x,y}|^2} \left|\omega^*_{x,y}\tau\right|^2$ (the factor 2 results from summing $J = 2$ spectra). Since m is radially symmetric, averaged over all directions, this inner product utilizes only half of the power of τ. Hence:

$$\mathfrak{D}_{\omega_{x,y}}(\hat{a}) = \iint \frac{2|m_{\omega_{x,y}}|^2}{|\omega_{x,y}|^2} \left|\omega^*_{x,y}\tau\right|^2 d\omega_x d\omega_y = \iint |m_{\omega_{x,y}}|^2 |\tau|^2 d\omega_x d\omega_y \leq \frac{1}{2}\mathfrak{G}_{\omega_{x,y}}(\hat{a}). \tag{37}$$

In Fig 1(a) we see that for multiplicative noise two smaller holes at the far ends of the aperture are preferred, while for additive noise it is better to have wide holes at the expense of reducing the disparity between them.

Halved rings: (Fig 1(b)). Our numerical calculation shows that for multiplicative noise this design utilizes 85% of the bound. There is no simple closed-form formula for these code spectra. However, we provide below a few intuitive arguments to justify this success. The important property of halved rings is a large disparity along most orientations. Fig 2(b) visualizes the 1D oriented code projections. Apart from the horizontal projection, the mean disparity at most other orientations is close to the aperture boundary, and hence Lemma 2 implies high discrimination at most orientations.

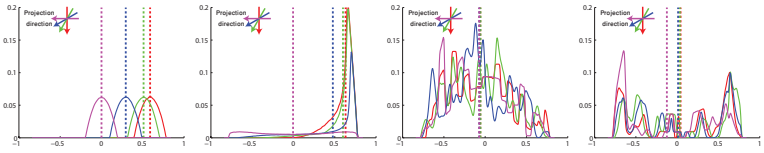


Fig. 2: *1D projections ρ_θ for the first 4 codes of Figure 1, multiplicative noise. We plot projections at 4 sampled orientations 0^o (red), 30^o (green), 60^o (blue) and 90^o (magenta). For each projection direction the center of mass (mean disparity) is marked by a dashed line.*

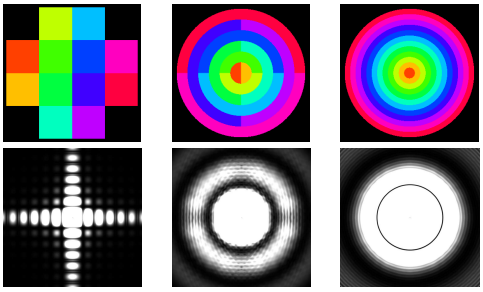
For this near-optimal pair we can verify empirically that all three properties of Corollary 1 hold. Property 1: the open code area is next to the periphery of the aperture disc. Property 2: for most orientations, the projections in Fig 2(b) are similar, and have the shape of a narrow peak at the right plus a tail (the exact tail shape varies between different projection directions). Since most projections are similar, most oriented slices from the spectra are similar, and the spectra are relatively radially symmetric except of a narrow angle range around the vertical direction. Property 3: according to the discussion after Lemma 2 the fact that the 1D projections have most mass around a narrow peak implies that the magnitude derivative power is small, which implies by Lemma 1 that most of the derivative power is discrimination power.

When the noise is additive the codes area is wider and the optimal shape resemble the one discovered by Zhou et al. [1] (who assume an additive noise). The banana's shape of the ends reduces the degeneracy around the vertical direction.

Pseudo random codes: (Fig 1(c,d)) obtain the worst discrimination scores. In terms of Lemma 2, the fact that light is collected from all around the aperture means that there is no real disparity, and the mean projections in Fig 2(c,d) are close to zero. Also, these codes violate properties 2 and 3 of Corollary 1. They are not radially symmetric and the derivatives are not orthogonal to the OTFs. In the multiplicative noise case the full aperture code (Fig 1(c)) has another major problem since it lets in light over the entire aperture area and not only around the boundaries. This is solved by a code in the outer ring (Fig 1(d)), but properties 2,3 are still problematic.

Ring and disc pair: (Fig 1(e)). This design should simulate the concept of depth from defocus (DFD). Intuitively, its drawback is enabling only half of the possible disparity-from the aperture boundary to the center, instead of from end to end. While this design is perfectly radially symmetric and satisfies property 2 of Corollary 1, for multiplicative noise it violates property 1, as it lets in light at the center of the aperture and holes at the center of the aperture do not contribute to spectra gradients. This design also violates property 3 since the OTFs vector is not orthogonal to its derivative.

Stereo v.s. depth from defocus: Our analysis predicts that DFD overcomes stereo. The disadvantages of stereo are having disparity only along one axis, and in the additive noise case it also suffers because it blocks light. This is consistent with previously reported evaluations of stereo v.s. DFD under similar physical dimensions [11–13].



(a) $\mathfrak{D}(\hat{\mathbf{a}}) = \mathbf{0.48}\mathfrak{B}$ (b) $\mathfrak{D}(\hat{\mathbf{a}}) = \mathbf{0.88}\mathfrak{B}$ (c) $\mathfrak{D}(\hat{\mathbf{a}}) = \mathbf{0.85}\mathfrak{B}$

Fig. 3: *Top row: Aperture codes which divide the aperture into multiple images. Second row: The discrimination $\mathfrak{D}_{\omega_{x,y}}(\hat{\mathbf{a}})$ they provide in each spatial frequency. At the bottom we provide the portion of the upper bound achieved by $\mathfrak{D}(\hat{\mathbf{a}})$ for the additive noise case*

4.2 Multiple image designs

We now consider some designs which capture a large number of images. We note that while an image captured from an inner aperture area is less discriminative than an image from an outer ring, adding additional images always improves the discrimination score. Therefore if we are allowed to capture a large number of images we do want to utilize the entire aperture area. If the images number is unlimited, better results can be obtained if we restrict the area of each code such that $\alpha \to 0$ and the noise reduces $\eta^2 \to \eta_{add}^2$.

The first example divides the aperture into multiple subsquares (Fig 3(a)). It is suboptimal and must lose at least 50% of the gradient power budget, since as in the stereo case the gradients have a constant direction. This design is similar to the plenoptic camera implementation of [5] which divides the aperture area and captures multiple views ([5] divides the sensor area as well).

In contrast, Fig 3(b,c) show two near-optimal designs utilizing over 85% of the bound. The first one generalizes Fig 1(b) with a set of half rings. To remove the degeneracy in one direction, it alternates vertical and horizontal pairs. The second near-optimal design is a set of annular rings, like the multi-aperture camera of [2]. Despite the sub-optimality of the ring and disc in Fig 1(e), multiple rings can approach the bound.

5 Discussion

In this paper we have analyzed the depth discrimination accuracy provided by a general set of aperture codes. We propose an analytic upper bound on the best achievable discrimination, and study the desired characteristics of an optimal solution. We show that under multiplicative noise, the two half-ring codes of [1] provide near-optimal discrimination. When a large number of images are allowed, a multi-aperture camera [2] dividing the aperture into annular rings provides near-optimal discrimination. In contrast, a plenoptic camera can achieve at most 50% of the bound.

Our analysis and bounds can be extended to more general families of computational cameras such as cameras including phase plates (optical elements with non standard curvature) and not only amplitude masks. This, however, requires an analysis of the lens kernels in the 4D light field spectrum space, as proposed in [10].

References

1. Zhou, C., Lin, S., Nayar, S.K.: Coded Aperture Pairs for Depth from Defocus. In: ICCV. (2009)
2. Green, P., Sun, W., Matusik, W., Durand, F.: Multi-aperture photography. SIGGRAPH (2007)
3. Scharstein, D., Szeliski, R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. IJCV (2002)
4. Adelson, E., Wang, J.: Single lens stereo with a plenoptic camera. IEEE PAMI (1992)
5. Georgeiv, T., Zheng, K., Curless, B., Salesin, D., Nayar, S., Intwala, C.: Spatio-angular resolution tradeoffs in integral photography. In: EGSR. (2006)
6. Chaudhuri, S., Rajagopalan, A.: Depth from defocus: A real aperture imaging approach. Springer-Verlag, New York (1999)
7. Veeraraghavan, A., Raskar, R., Agrawal, A., Mohan, A., Tumblin, J.: Dappled photography: Mask-enhanced cameras for heterodyned light fields and coded aperture refocusing. SIGGRAPH (2007)
8. Levin, A., Fergus, R., Durand, F., Freeman, W.: Image and depth from a conventional camera with a coded aperture. SIGGRAPH (2007)
9. Dowski, E., Cathey, W.: Single-lens single-image incoherent passive-ranging systems. App Opt (1994)
10. Levin, A., Hasinoff, S., Green, P., Durand, F., Freeman, W.: 4D frequency analysis of computational cameras for depth of field extension. SIGGRAPH (2009)
11. Schechner, Y., Kiryati, N.: Depth from defocus vs. stereo: How different really are they. IJCV (2000)
12. Farid, H., Simoncelli, E.: Range estimation by optical differentiation. JOSA (1998)
13. Vaish, V., Levoy, M., Szeliski, R., Zitnick, C., Kang, S.: Reconstructing occluded surfaces using synthetic apertures: Stereo, focus and robust measures. In: CVPR. (2006)
14. Levin, A., Freeman, W., Durand, F.: Understanding camera trade-offs through a Bayesian analysis of light field projections. In: ECCV. (2008)
15. Hasinoff, S., Kutulakos, K.: Light-efficient photography. In: ECCV. (2008)
16. Hasinoff, S., Kutulakos, K., Durand, F., Freeman, W.: Time-constrained photography. In: ICCV. (2009)
17. Zhou, C., Nayar, S.K.: What are Good Apertures for Defocus Deblurring? In: IEEE International Conference on Computational Photography. (2009)
18. Ng, R.: Fourier slice photography. SIGGRAPH (2005)