Light-Efficient Photography

Samuel W. Hasinoff and Kiriakos N. Kutulakos*

Dept. of Computer Science, University of Toronto
{hasinoff,kyros}@cs.toronto.edu

Abstract. We consider the problem of imaging a scene with a given depth of field at a given exposure level in the shortest amount of time possible. We show that by (1) collecting a sequence of photos and (2) controlling the aperture, focus and exposure time of each photo individually, we can span the given depth of field in less total time than it takes to expose a single narrower-aperture photo. Using this as a starting point, we obtain two key results. First, for lenses with continuously-variable apertures, we derive a closed-form solution for the *globally optimal* capture sequence, *i.e.*, that collects light from the specified depth of field in the most efficient way possible. Second, for lenses with discrete apertures, we derive an integer programming problem whose solution is the optimal sequence. Our results are applicable to off-the-shelf cameras and typical photography conditions, and advocate the use of dense, wide-aperture photo sequences as a light-efficient alternative to single-shot, narrow-aperture photography.

1 Introduction

Two of the most important choices when taking a photo are the photo's exposure level and its depth of field. Ideally, these choices will result in a photo whose subject is free of noise or pixel saturation [1, 2], and appears in-focus. These choices, however, come with a severe time constraint: in order to take a photo that has both a specific exposure level and a specific depth of field, we must expose the camera's sensor for a length of time dictated by the optics of the lens. Moreover, the larger the depth of field, the longer we must wait for the sensor to reach the chosen exposure level. In practice, this makes it impossible to efficiently take sharp and well-exposed photos of a poorly-illuminated subject that spans a wide range of distances from the camera. To get a good exposure level, we must compromise something – accepting either a smaller depth of field (incurring defocus blur [3–6]) or a longer exposure (incurring motion blur [7–9]).

In this paper we seek to overcome the time constraint imposed by lens optics, by capturing a sequence of photos rather than just one. We show that if the aperture, exposure time, and focus setting of each photo is selected appropriately, we can span a given depth of field with a given exposure level *in less total time than it takes to expose a single photo* (Fig. 1). This novel observation is based on a simple fact: even though wide apertures have a narrow depth of field (DOF), they are much more efficient than narrow apertures in gathering light from within their depth of field. Hence, even though

* This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under the RGPIN program and by an Ontario Premier's Research Excellence Award.

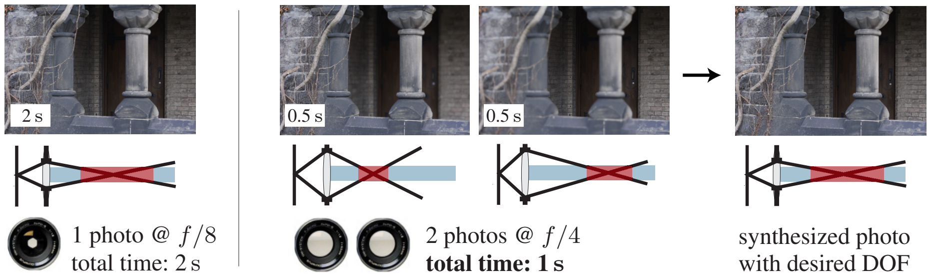

Fig. 1. *Left:* Traditional single-shot photography. The desired depth of field is shaded (red). *Right:* Light-efficient photography. Two wide-aperture photos span the same DOF as a single-shot narrow-aperture photo. Each wide-aperture photo requires $1/4$ the time to reach the exposure level of the single-shot photo, resulting in a $2\times$ net speedup for the total exposure time.

it is not possible to span a wide DOF with a single wide-aperture photo, it is possible to span it with several of them, and to do so very efficiently.

Using this observation as a starting point, we develop a general theory of *light-efficient photography* that addresses four questions: (1) under what conditions is capturing photo sequences with "synthetic" DOFs more efficient than single-shot photography? (2) How can we characterize the set of sequences that are *globally optimal* for a given DOF and exposure level, *i.e.* whose total exposure time is the shortest possible? (3) How can we compute such sequences automatically for a specific camera, depth of field, and exposure level? (4) Finally, how do we convert the captured sequence into a single photo with the specified depth of field and exposure level?

Little is known about how to gather light efficiently from a specified DOF. Research on computational photography has not investigated the light-gathering ability of existing methods, and has not considered the problem of optimizing exposure time for a desired DOF and exposure level. For example, even though there has been great interest in manipulating a camera's DOF through optical [10–13] or computational [2, 5, 14–18] means, current approaches do so without regard to exposure time – they simply assume that the shutter remains open as long as necessary to reach the desired exposure level. This assumption is also used for high-dynamic range photography [2, 19], where the shutter must remain open for long periods in order to capture low-radiance regions in a scene. In contrast, here we capture photos with camera settings that are carefully chosen to minimize total exposure time for the desired DOF and exposure level.

Since shorter total exposure times reduce motion blur, our work can be thought of as complementary to recent *synthetic shutter* approaches whose goal is to reduce such blur. Instead of controlling aperture and focus, these techniques divide a given exposure interval into several shorter ones, with the same total exposure (*e.g.*, n photos, each with $1/n$ the exposure time [9]; two photos, one with long and one with short exposure [8]; or one photo where the shutter opens and closes intermittently during the exposure [7]). These techniques do not increase light-efficiency but can be readily combined with our work, to confer the advantages of both methods.

Moreover, our approach can be thought of as complementary to work on light field cameras [13, 17, 18], which are based on an orthogonal tradeoff between resolution and directional sampling. Compared to regular wide-aperture photography, these designs do not have the ability to extend the DOF when their reduced resolution is taken into account. Along similar lines, wavefront coding [11] exploits special optics to extend the DOF with no change in exposure time by using another orthogonal tradeoff – accepting lower signal-to-noise ratio for higher frequencies.

The final step in light-efficient photography involves merging the captured photos to create a new one (Fig. 1). As such, our work is related to the well-known technique of extended-depth-of-field imaging, which has found wide use in microscopy [18] and macro photography [17, 20].

Our work offers four contributions over the state of the art. First, we develop a theory that leads to provably-efficient light-gathering strategies, and applies both to off-the-shelf cameras and to advanced camera designs [7, 9] under typical photography conditions. Second, from a practical standpoint, our analysis shows that the optimal (or near-optimal) strategies are very simple: for example, in the continuous case, a strategy using the widest-possible aperture for all photos is either globally optimal or it is very close to it (in a quantifiable sense). Third, our experiments with real scenes suggest that it is possible to compute good-quality synthesized photos using readily-available algorithms. Fourth, we show that despite requiring less total exposure time than a single narrow-aperture shot, light-efficient photography provides more information about the scene (*i.e.*, depth) and allows post-capture control of aperture and focus.

2 The Exposure Time *vs*. Depth of Field Tradeoff

The *exposure level* of a photo is the total radiant energy integrated by the camera's entire sensor while the shutter is open. The exposure level can influence significantly the quality of a captured photo because when there is no saturation or thermal noise, a pixel's signal-to-noise ratio (SNR) always increases with higher exposure levels [1]. For this reason, most modern cameras can automate the task of choosing an exposure level that provides high SNR for most pixels and causes little or no saturation.

Lens-based camera systems provide only two ways to control exposure level – the diameter of their aperture and the exposure time. We assume that all light passing through the aperture will reach the sensor plane, and that the average irradiance measured over this aperture is independent of the aperture's diameter. In this case, the exposure level L is equal to

$$L \;=\; \tau\, D^2 \;\; , \tag{1}$$

where τ is exposure time, D is the effective aperture diameter, and the units of L are chosen appropriately.

Now suppose that we have chosen a desired exposure level L^*. How can we capture a photo at this exposure level? Equation (1) suggests that there are only two general strategies for doing this – either choose a long exposure time and a small aperture diameter, or choose a large aperture diameter and a short exposure time. Unfortunately, both strategies have important side-effects: increasing exposure time can introduce motion blur when we photograph moving scenes [8, 9]; opening the lens aperture, on the

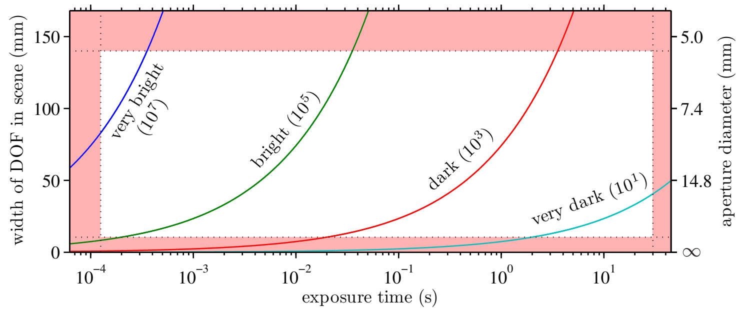

Fig. 2. Each curve represents all pairs (τ, D) for which $\tau D^2 = L^*$ in a specific scene. Shaded zones correspond to pairs outside the camera limits (valid settings were $\tau \in [1/8000\,\mathrm{s},\ 30\,\mathrm{s}]$ and $D \in [f/16, f/1.2]$ with $f = 85\,\mathrm{mm}$). Also shown is the DOF corresponding to each diameter D. The maximum acceptable blur was set to $c = 25\,\mu\mathrm{m}$, or about 3 pixels in our camera. Different curves represent scenes with different average radiance (relative units shown in brackets).

other hand, affects the photo's *depth of field (DOF)*, *i.e.*, the range of distances where scene points do not appear out of focus. These side-effects lead to an important tradeoff between a photo's exposure time and its depth of field (Fig. 2):

> **Exposure Time *vs*. Depth of field Tradeoff:** *We can either achieve a desired exposure level L^* with short exposure times and a narrow DOF, or with long exposure times and a wide DOF.*

In practice, the exposure time *vs*. DOF tradeoff limits the range of scenes that can be photographed at a given exposure level (Fig. 2). This range depends on scene radiance, the physical limits of the camera (*i.e.*, range of possible apertures and shutter speeds), as well as subjective factors (*i.e.*, acceptable levels of motion blur and defocus blur).

Our goal is to "break" this tradeoff by seeking novel photo acquisition strategies that capture a given depth of field at the desired exposure level L^* much faster than traditional optics would predict. We briefly describe below the basic geometry and relations governing a photo's depth of field, as they are particularly important for our analysis.

2.1 Depth of Field Geometry

We assume that focus and defocus obey the standard thin lens model [3, 21]. This model relates three positive quantities (Eq. (A) in Table 1): the focus setting v, defined as the distance from the sensor plane to the lens; the distance d from the lens to the in-focus scene plane; and the focal length f, representing the "focusing power" of the lens.

Apart from the idealized pinhole, all apertures induce spatially-varying amounts of defocus for points in the scene (Fig. 3a). If the lens focus setting is v, all points at distance d from the lens will be in-focus. A scene point at distance $d' \neq d$, however, will be defocused: its image will be a circle on the sensor plane whose diameter b is called the *blur diameter*. For any given distance d, the thin-lens model tells us exactly what focus setting we should use to bring the plane at distance d into focus, and what the blur diameter will be for points away from this plane (Eqs. (B) and (C), respectively).

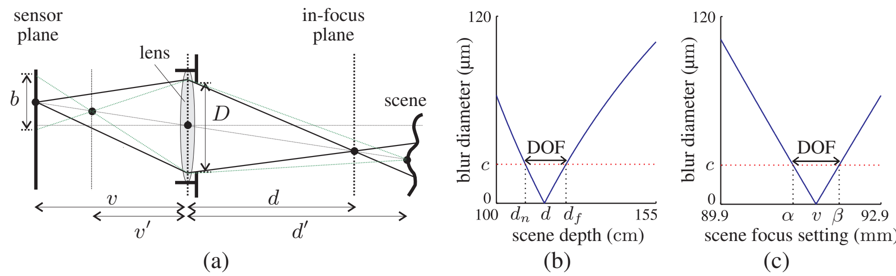

Fig. 3. (a) Blur geometry for a thin lens. (b) Blur diameter as a function of distance to a scene point. The plot is for a lens with $f = 85\,\mathrm{mm}$, focused at $117\,\mathrm{cm}$ with an aperture diameter of $5.31\,\mathrm{mm}$ (*i.e.*, an $f/16$ aperture in photography terminology). (c) Blur diameter and DOF represented in the space of focus settings.

(A) Thin lens law	*(B) Focus for distance d*	*(C) Blur diameter for distance d'*	*(D) Aper. diam. for DOF $[\alpha, \beta]$*	*(E) Focus for DOF $[\alpha, \beta]$*	*(F) DOF for aper. diam. D, focus v*
$\frac{1}{v} + \frac{1}{d} = \frac{1}{f}$	$v = \frac{fd}{d-f}$	$b = D\frac{f\lvert d'-d\rvert}{d'(d-f)}$	$D = c\,\frac{\beta+\alpha}{\beta-\alpha}$	$v = \frac{2\,\alpha\,\beta}{\alpha+\beta}$	$\alpha, \beta = \frac{Dv}{D \pm c}$

Table 1. Eqs. (A)–(F): Basic equations governing focus and DOFs for the thin-lens model.

For a given aperture and focus setting, the *depth of field* is the interval of distances in the scene whose blur diameter is below a maximum acceptable size c (Fig. 3b).

Since every distance in the scene corresponds to a unique focus setting (Eq. (B)), every DOF can also be expressed as an interval $[\alpha, \beta]$ in the space of focus settings. This alternate DOF representation gives us especially simple relations for the aperture and focus setting that produce a given DOF (Eqs. (D) and (E)) and, conversely, for the DOF produced by a given aperture and focus setting (Eq. (F)). We adopt this DOF representation for the rest of the paper (Fig. 3c).

A key property of the depth of field is that it shrinks when the aperture diameter increases: from Eq. (C) it follows that for a given out-of-focus distance, larger apertures always produce larger blur diameters. This equation is the root cause of the exposure time *vs*. depth of field tradeoff.

3 The Synthetic DOF Advantage

Suppose that we want to capture a single photo with a specific exposure level L^* *and* a specific depth of field $[\alpha, \beta]$. How quickly can we capture this photo? The basic DOF geometry of Sect. 2.1 tells us we have no choice: there is only one aperture diameter that can span the given depth of field (Eq. (D)), and only one exposure time that can achieve a given exposure level with that diameter (Eq. (1)). This exposure time is[1]

$$\tau^{one} \;=\; L^* \cdot \left(\frac{\beta-\alpha}{c\,(\beta+\alpha)}\right)^2 \;. \tag{2}$$

[1] The apertures and exposure times of real cameras span finite intervals and, in many cases, take discrete values. Hence, in practice, Eq. (2) holds only approximately.

The key idea of our approach is that while lens optics do not allow us to reduce this time without compromising the DOF or the exposure level, we *can* reduce it by taking more photos. This is based on a simple observation that takes advantage of the different rates at which exposure time and DOF change: if we increase the aperture diameter and adjust exposure time to maintain a constant exposure level, its DOF shrinks (at a rate of about $1/D$), but the exposure time shrinks much faster (at a rate of $1/D^2$). This opens the possibility of "breaking" the exposure time *vs*. DOF tradeoff by capturing a sequence of photos that jointly span the DOF in less total time than τ^{one} (Fig. 1).

Our goal is to study this idea in its full generality, by finding capture strategies that are provably time-optimal. We therefore start from first principles, by formally defining the notion of a *capture sequence* and of its *synthetic depth of field*:

Definition 1 (Photo Tuple). *A tuple* $\langle\, D,\ \tau,\ v \,\rangle$ *that specifies a photo's aperture diameter, exposure time, and focus setting, respectively.*

Definition 2 (Capture Sequence). *A finite ordered sequence of photo tuples.*

Definition 3 (Synthetic Depth of Field). *The union of DOFs of all photo tuples in a capture sequence.*

We will use two efficiency measures: the *total exposure time* of a sequence is the sum of the exposure times of all its photos; the *total capture time*, on the other hand, is the actual time it takes to capture the photos with a specific camera. This time is equal to the total exposure time, plus any overhead caused by camera internals (computational and mechanical). We now consider the following general problem:

> **Light-Efficient Photography**: *Given a set* $\mathcal{D}$ *of available aperture diameters, construct a capture sequence such that: (1) its synthetic DOF is equal to* $[\alpha, \beta]$*; (2) all its photos have exposure level* L^**; (3) the total exposure time (or capture time) is smaller than* τ^{one}*; and (4) this time is a global minimum over all finite capture sequences.*

Intuitively, whenever such a capture sequence exists, it can be thought of as being optimally more efficient than single-shot photography in gathering light. Below we analyze three instances of the light-efficient photography problem. In all cases, we assume that the exposure level L^*, depth of field $[\alpha, \beta]$, and aperture set $\mathcal{D}$ are known and fixed.

Noise Properties. All photos we consider have similar noise, because most noise sources (photon, sensor, and quantization noise) depend only on exposure level, which we hold constant. The only exception is thermal noise, which increases with exposure time [1], and so will be lower for light-efficient sequences with shorter exposures.

4 Theory of Light-Efficient Photography

4.1 Continuously-Variable Aperture Diameters

Many manual-focus SLR lenses allow their aperture diameter to vary continuously within some interval $\mathcal{D} = [D_{min},\ D_{max}]$. In this case, we prove that the optimal capture sequence has an especially simple form – it is unique, it uses the same aperture diameter for all tuples, and this diameter is either the maximum possible or a diameter close to that maximum.

More specifically, consider the following special class of capture sequences:

Definition 4 (Sequences with Sequential DOFs). *A capture sequence has sequential DOFs if for every pair of adjacent photo tuples, the right endpoint of the first tuple's DOF is the left endpoint of the second.*

The following theorem states that the solution to the light-efficient photography problem is a specific sequence from this class:

Theorem 1 (Optimal Capture Sequence for Continuous Apertures). *(1) If the DOF endpoints satisfy $\beta < (7+4\sqrt{3})\alpha$, the sequence that globally minimizes total exposure time is a sequence with sequential DOFs whose tuples all have the same aperture. (2) Define $D(k)$ and n as follows:*

$$D(k) \;=\; c\,\frac{\sqrt[k]{\beta}+\sqrt[k]{\alpha}}{\sqrt[k]{\beta}-\sqrt[k]{\alpha}}\;, \qquad n \;=\; \left\lfloor \frac{\log\frac{\alpha}{\beta}}{\log\left(\frac{D_{max}-c}{D_{max}+c}\right)} \right\rfloor \;. \tag{3}$$

The aperture diameter D^* and length n^* of the optimal sequence is given by

$$D^* \;=\; \begin{cases} D(n) & \text{if } \frac{D(n)}{D_{max}} > \sqrt{\frac{n}{n+1}} \\ D_{max} & \text{otherwise.} \end{cases} \qquad n^* \;=\; \begin{cases} n & \text{if } \frac{D(n)}{D_{max}} > \sqrt{\frac{n}{n+1}} \\ n+1 & \text{otherwise.} \end{cases} \;. \tag{4}$$

Theorem 1 specifies the optimal sequence indirectly, via a "recipe" for calculating the optimal length and the optimal aperture diameter (Eqs. (3) and (4)). Informally, this calculation involves three steps. The first step defines the quantity $D(k)$; in our proof of Theorem 1 (see Appendix A), we show that this quantity represents the only aperture diameter that can be used to "tile" the interval $[\alpha,\beta]$ with exactly k photo tuples of the same aperture. The second step defines the quantity n; in our proof, we show that this represents the largest number of photos we can use to tile the interval $[\alpha,\beta]$ with photo tuples of the same aperture. The third step involves choosing between two "candidates" for the optimal solution – one with n tuples and one with $n+1$.

Theorem 1 makes explicit the somewhat counter-intuitive fact that the most light-efficient way to span a given DOF $[\alpha,\beta]$ is to use images whose DOFs are very narrow. This fact applies broadly, because Theorem 1's inequality condition for α and β is satisfied for all lenses for consumer photography that we are aware of (*e.g.*, see [22]).[2] See Fig. 4 for an application of this theorem to a practical example.

Note that Theorem 1 specifies the number of tuples in the optimal sequence and their aperture diameter, but does not specify their exposure times or focus settings. The following lemma shows that specifying those quantities is not necessary because they are determined uniquely. Importantly, Lemma 1 gives us a recursive formula for computing the exposure time and focus setting of each tuple in the sequence:

[2] To violate the condition, the minimum focusing distance must be under $1.077f$, measured from the lens center.

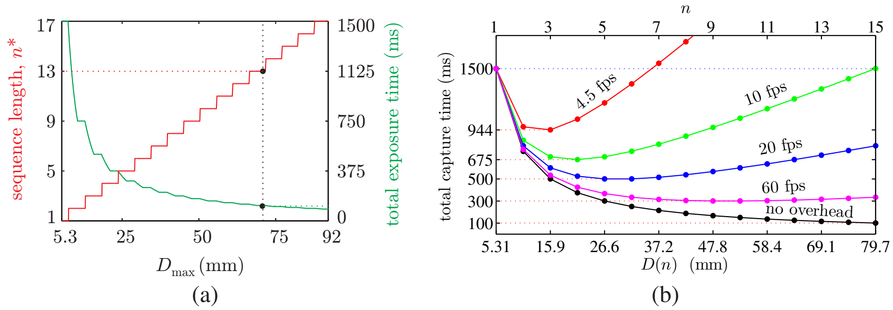

Fig. 4. (a) Optimal light-efficient photography of a "dark" subject spanning a DOF of $[110\,\text{cm}, 124\,\text{cm}]$, using an $f = 85\,\text{mm}$ lens with a continuously-variable aperture. In this example, we can use a $f/16$ aperture ($5.3\,\text{mm}$ diameter) to cover the DOF with a single photo, which requires a $1.5\,\text{s}$ exposure to obtain the desired exposure level. The plot illustrates the optimal sequences when the aperture diameter is restricted to a range $[f/16, D_{\max}]$: for each value of the maximum aperture, $D_{\max}$, Theorem 1 gives a unique optimal sequence. The graph shows the number of images n^* (red) and total exposure time (green) of this sequence. As $D_{\max}$ increases, the total exposure time of the optimal sequence falls dramatically: for lenses with an $f/1.2$ maximum aperture ($71\,\text{mm}$), synthetic DOFs confer a $13\times$ speedup over single-shot photography for the same exposure level. (b) The effect of camera overhead for various frame-per-second (fps) rates. Each point represents the total capture time of a sequence that spans the DOF and whose photos all use the diameter $D(n)$ indicated. Even though overhead reduces the efficiency of long sequences, synthetic DOFs are faster than single-shot photography even for low fps rates.

Lemma 1 (Construction of Sequences with Sequential DOFs). *Given a left DOF endpoint α, every ordered sequence $D_1, \ldots, D_n$ of aperture diameters defines a unique capture sequence with sequential DOFs whose n tuples are*

$$\langle\, D_i\,,\ \frac{L^*}{{D_i}^2}\,,\ \frac{D_i + c}{D_i}\,\alpha_i\,\rangle\ ,\quad i\ =\ 1,\ldots,n\ , \tag{5}$$

with α_i given by the following recursive relation:

$$\alpha_i\ =\ \begin{cases} \alpha & \textit{if } i = 1\ , \\ \frac{D_i + c}{D_i - c}\,\alpha_{i-1} & \textit{otherwise.} \end{cases} \tag{6}$$

4.2 Discrete Aperture Diameters

Modern auto-focus lenses often restrict the aperture diameter to a discrete set of choices, $\mathcal{D} = \{D_1, \ldots, D_m\}$. These diameters form a geometric progression, spaced so that the aperture area doubles every two or three steps. Unlike the continuous case, the optimal capture sequence is not unique and may contain several distinct aperture diameters. To find an optimal sequence, we reduce the problem to integer linear programming [23]:

Theorem 2 (Optimal Capture Sequence for Discrete Apertures). *There exists an optimal capture sequence with sequential DOFs whose tuples have a non-decreasing*

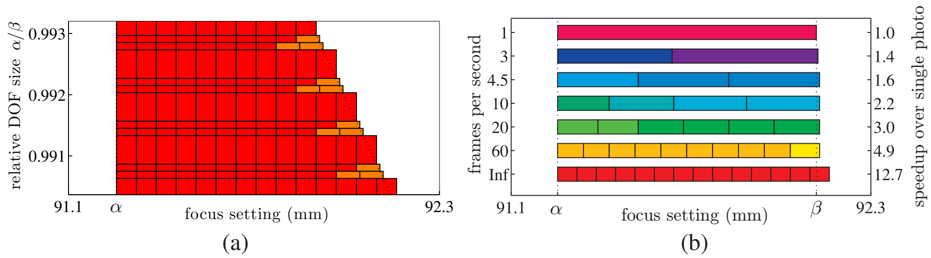

Fig. 5. Optimal light-efficient photography with discrete apertures, shown for a Canon EF85mm 1.2L lens (23 apertures, illustrated in different colors). (a) For a depth of field whose left endpoint is α, we show optimal capture sequences for a range of relative DOF sizes $\frac{\alpha}{\beta}$. These sequences can be read horizontally, with subintervals corresponding to the apertures determined by Theorem 2. Note that when the DOF is large, the optimal sequence approximates the continuous case. The diagonal dotted line indicates the DOF to be spanned. (b) Visualizing the optimal capture sequence as a function of the camera overhead for the DOF $[\alpha, \beta]$. Note that with higher overhead, the optimal sequence involves fewer photos with larger DOFs (*i.e.*, smaller apertures).

sequence of aperture diameters. Moreover, if n_i is the number of times diameter D_i appears in the sequence, the multiplicities $n_1, \ldots, n_m$ satisfy the integer program

$$\textit{minimize} \quad \sum_{i=1}^{m} n_i \, \frac{L^*}{{D_i}^2} \tag{7}$$

$$\textit{subject to} \quad \sum_{i=1}^{m} n_i \log \frac{D_i - c}{D_i + c} \;\leq\; \log \frac{\alpha}{\beta} \tag{8}$$

$$n_i \geq 0 \;\; \textit{and integer} \; . \tag{9}$$

See [24] for a proof. As with Theorem 1, Theorem 2 does not specify the focus settings in the optimal capture sequence. We use Lemma 1 for this purpose, which explicitly constructs it from the apertures and their multiplicities.

While it is not possible to obtain a closed-form expression for the optimal sequence, solving the integer program for any desired DOF is straightforward. We use a simple branch-and-bound method based on successive relaxations to linear programming [23]. Moreover, since the optimal sequence depends only on the relative DOF size $\frac{\alpha}{\beta}$, we pre-compute it for all possible DOFs and store the results in a lookup table (Fig. 5a).

4.3 Discrete Aperture Diameters Plus Overhead

Our treatment of discrete apertures generalizes easily to account for camera overhead. We model overhead as a per-shot constant, τ^{over}, that expresses the minimum delay between the time that the shutter closes and the time it is ready to open again for the next photo. To find the optimal sequence, we modify the objective function of Theorem 2 so that it measures total capture time rather than total exposure time:

$$\textit{minimize} \quad \sum_{i=1}^{m} n_i \left[\, \tau^{over} \; + \; \frac{L^*}{{D_i}^2} \,\right] \; . \tag{10}$$

Clearly, a non-negligible overhead penalizes long capture sequences and reduces the synthetic DOF advantage. Despite this, Fig. 5b shows that synthetic DOFs offer significant speedups even for current off-the-shelf cameras. These speedups will be amplified further as camera manufacturers continue to improve frame-per-second rates.

5 Depth of Field Compositing and Resynthesis

DOF Compositing. To reproduce the desired DOF, we use a variant of the Photomontage method [20], based on maximizing a simple "focus measure" that evaluates local contrast according to the difference-of-Gaussians filter. In this method, each pixel in the composite has a label that indicates the input photo for which the pixel is in-focus. These labels are optimized with a Markov random field network that is biased toward piecewise smoothness. The resulting composite is a blend of the input photos, performed in the gradient domain to reduce artifacts at label boundaries.

3D Reconstruction. The DOF compositing operation produces a coarse depth map as an intermediate step. This is because labels correspond to input photos, and each input photo defines an in-focus depth according to the focus setting with which it was captured. We found this depth map to be sufficient for good-quality resynthesis, although a more sophisticated depth-from-defocus analysis is also possible [6].

Synthesizing Photos for Novel Focus Settings and Aperture Diameters. To synthesize novel photos, we generalize DOF compositing and take advantage of the different levels of defocus throughout the capture sequence. We proceed in four basic steps. First, given a specific focus and aperture setting, we use Eq. (C) and the coarse depth map to assign a blur diameter to each pixel in the final composite. Second, we use Eq. (C) again to determine, for each pixel in the composite, the input photo whose blur diameter that corresponds to the pixel's depth matches most closely.[3] Third, for each depth layer, we synthesize a photo under the assumption that the entire scene is at that depth, and is observed with the novel focus and aperture setting. To do this, we use the blur diameter for this depth to define an interpolation between two of the input photos. We currently interpolate using simple linear cross-fading, which we found to be adequate when the DOF is sampled densely enough (*i.e.*, with 5 or more images). Fourth, we generate the final composite by merging all these synthesized images into one photo using the same gradient domain blending as in DOF compositing, with the same depth labels.

6 Experimental Results

Figure 6 shows results and timings for two experiments, performed with two different cameras – a high-end digital SLR and a compact digital camera (see [24] for more results and videos). All photos were captured at the same exposure level for each experiment. In each case, we captured (1) a narrow-aperture photo and (2) the optimal capture sequence for the equivalent DOF and the particular camera. To compensate for the distortions that occur with changes in focus setting, we align the photos according to a one-time calibration method that fits a radial magnification model to focus setting [25]. To determine the maximum acceptable blur diameter c for each camera, we evaluated focus using a resolution chart. The values we found, $5\,\mu\mathrm{m}$ (1.4 pixels) and $25\,\mu\mathrm{m}$ (3.5 pixels) respectively, agree with standard values [21].

[3] Note each blur diameter is consistent with two depths (Fig. 3b). We resolve the ambiguity by choosing the matching input photo whose focus setting is closest to the synthetic focus setting.

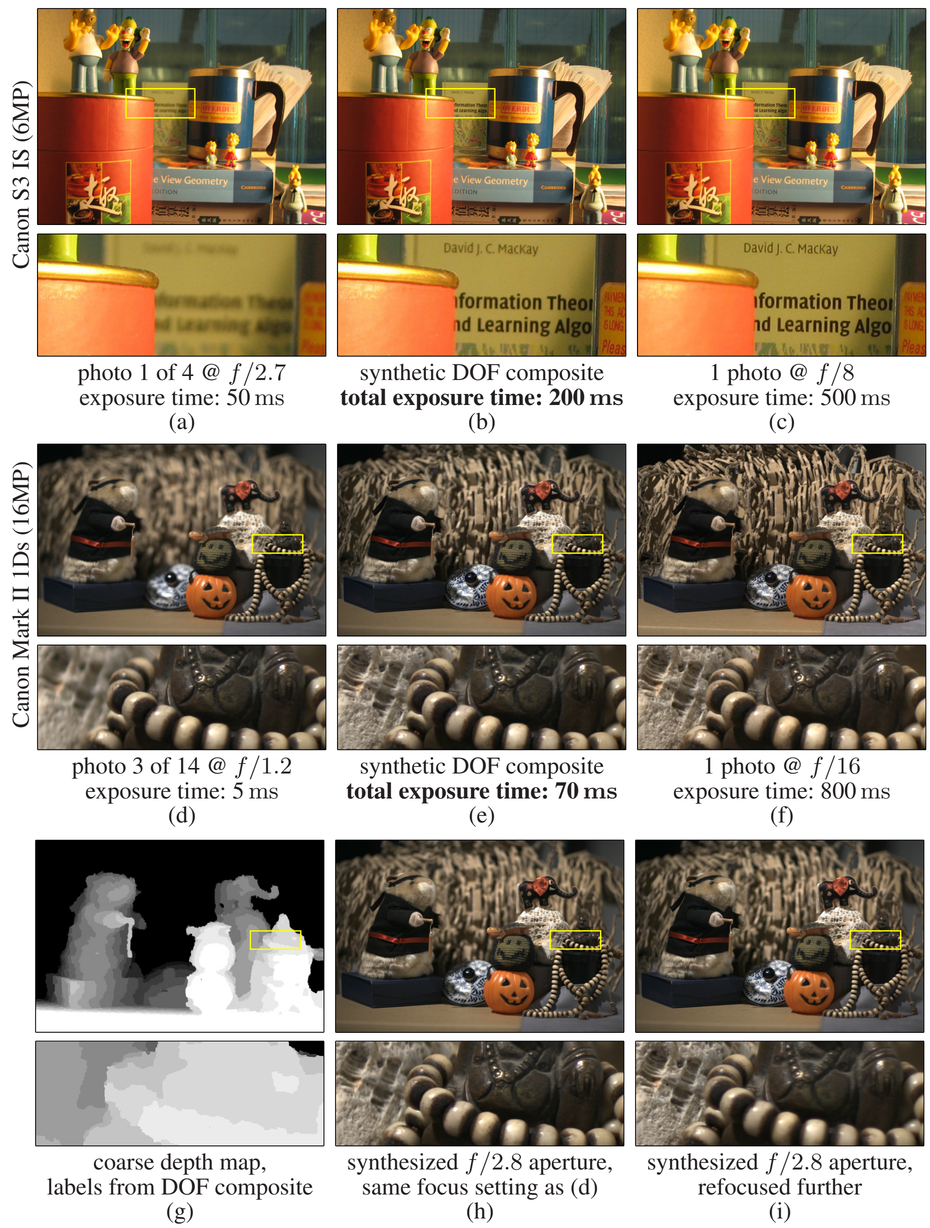

Fig. 6. Light-efficient photography timings and synthesis, for several real scenes, captured using a compact digital camera and a digital SLR. (a,d) Sample wide-aperture photo from the synthetic DOF sequence. (b,e) DOF composites synthesized from this sequence. (c,f) Narrow-aperture photos spanning an equivalent DOF, but with much longer exposure time. (g) Coarse depth map, computed from the labeling we used to compute (e). (h) Synthetically changing aperture size, focused at the same setting as (d). (i) Synthetically changing the focus setting as well.

Timing Comparisons and Optimal Capture Sequences. To determine the optimal capture sequences, we assumed zero camera overhead and applied Theorem 2 for the chosen DOF and exposure level, according to the specifications of each camera and lens. The optimal sequences involved spanning the DOF using the largest aperture in both cases. As Fig. 6 shows, these sequences led to significant speedups in exposure time – $2.5\times$ and $11.9\times$ for the compact digital camera and digital SLR, respectively.

DOF Compositing. Figures 6b and 6e show that despite the availability of just a coarse depth map, our compositing scheme is able to reproduce high-frequency detail over the whole DOF without noticeable artifacts, even in the vicinity of depth discontinuities. Note that while the synthesized photos satisfy our goal of spanning a specific DOF, objects outside that DOF will appear more defocused than in the corresponding narrow-aperture photo (*e.g.*, see the background in Figs. 6e–f). While increased background defocus may be desirable (*e.g.*, for portrait or macro photography), it is also possible to capture sequences of photos to reproduce arbitrary levels of defocus outside the DOF.

Depth Maps and DOF Compositing. Despite being more efficient to capture, sequences with synthetic DOFs provide 3D shape information at no extra acquisition cost (Fig. 6g). Figures 6h–i show results of using this depth map to compute novel images whose aperture and focus setting was changed synthetically according to Sect. 5.

Implementation Details. Neither of our cameras provide the ability to control focus remotely. For our compact camera we used modified firmware that enables scripting [26], while for our SLR we used a computer-controlled motor to drive the focusing ring mechanically. Both methods incur high overhead and limit us to about 1 fps.

While light-efficient photography is not practical in this context, it will become increasingly so, as newer cameras begin to provide focus control and to increase frame-per-second rates. For example, the Canon EOS-1Ds Mark III provides remote focus control for all Canon EF lenses, and the Casio EX-F1 can capture 60 fps at 6MP.

7 Concluding Remarks

In this paper we studied the use of dense, wide-aperture photo sequences as a light-efficient alternative to single-shot, narrow-aperture photography. While our emphasis has been on the underlying theory, we believe our method has great practical potential.

We are currently investigating several extensions to the basic approach. These include designing light-efficient strategies (1) for spanning arbitrary defocus profiles, rather than just the DOF; (2) improving efficiency by taking advantage of the camera's auto-focus sensor; and (3) operating under a highly-restricted time-budget, for which it becomes important to weigh the tradeoff between noise and defocus.

References

1. Healey, G.E., Kondepudy, R.: Radiometric CCD camera calibration and noise estimation. TPAMI **16**(3) (1994) 267–276

2. Hasinoff, S.W., Kutulakos, K.N.: A layer-based restoration framework for variable-aperture photography. In: ICCV. (2007)
3. Pentland, A.P.: A new sense for depth of field. TPAMI **9**(4) (1987) 523–531
4. Krotkov, E.: Focusing. IJCV **1**(3) (1987) 223–237
5. Hiura, S., Matsuyama, T.: Depth measurement by the multi-focus camera. In: CVPR. (1998) 953–959
6. Watanabe, M., Nayar, S.K.: Rational filters for passive depth from defocus. IJCV **27**(3) (1998) 203–225
7. Raskar, R., Agrawal, A., Tumblin, J.: Coded exposure photography: motion deblurring using fluttered shutter. In: SIGGRAPH. (2006) 795–804
8. Yuan, L., Sun, J., Quan, L., Shum, H.Y.: Image deblurring with blurred/noisy image pairs. In: SIGGRAPH. (2007)
9. Telleen, J., Sullivan, A., Yee, J., Gunawardane, P., Wang, O., Collins, I., Davis, J.: Synthetic shutter speed imaging. In: Proc. Eurographics. (2007) 591–598
10. Farid, H., Simoncelli, E.P.: Range estimation by optical differentiation. JOSA A **15**(7) (1998) 1777–1786
11. Cathey, W.T., Dowski, E.R.: New paradigm for imaging systems. Applied Optics **41**(29) (2002) 6080–6092
12. Levin, A., Fergus, R., Durand, F., Freeman, W.T.: Image and depth from a conventional camera with a coded aperture. In: SIGGRAPH. (2007)
13. Veeraraghavan, A., Raskar, R., Agrawal, A., Mohan, A., Tumblin, J.: Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing. In: SIGGRAPH. (2007)
14. Aizawa, K., Kodama, K., Kubota, A.: Producing object-based special effects by fusing multiple differently focused images. TCSVT **10**(2) (2000)
15. Chaudhuri, S.: Defocus morphing in real aperture images. JOSA A **22**(11) (2005) 2357–2365
16. Hasinoff, S.W., Kutulakos, K.N.: Confocal stereo. In: ECCV. Volume 1. (2006) 620–634
17. Ng, R.: Fourier slice photography. In: SIGGRAPH. (2005) 735–744
18. Levoy, M., Ng, R., Adams, A., Footer, M., Horowitz, M.: Light field microscopy. In: SIGGRAPH. (2006) 924–934
19. Debevec, P., Malik, J.: Recovering high dynamic range radiance maps from photographs. In: SIGGRAPH. (1997) 369–378
20. Agarwala, A., *et al*.: Interactive digital photomontage. In: SIGGRAPH. (2004) 294–302
21. Smith, W.J.: Modern Optical Engineering. 3rd edn. McGraw-Hill, New York (2000)
22. Canon lens chart, `http://www.usa.canon.com/app/pdf/lens/`.
23. Nocedal, J., Wright, S.J.: Numerical Optimization. Springer (1999)
24. `http://www.cs.toronto.edu/~kyros/research/lightefficient/`.
25. Willson, R., Shafer, S.: What is the center of the image? JOSA A **11**(11) (1994) 2946–2955
26. CHDK, `http://chdk.wikia.com/`.

A Proof of Theorem 1

Theorem 1 follows as a consequence of Lemma 1 and four additional lemmas. We first state Lemmas 2–5 below and then prove a subset of them, along with a proof sketch the theorem. All missing proofs can be found in [24].

Lemma 2 (Efficiency of Sequential DOFs). *For every sequence $\mathcal{S}$, there is a sequence $\mathcal{S}'$ with sequential DOFs that spans the same synthetic DOF and whose total exposure time is no larger.*

Lemma 3 (Permutation of Sequential DOFs). *Given the left endpoint, α, every permutation of $D_1, \ldots, D_n$ defines a capture sequence with sequential DOFs that has the same synthetic DOF and the same total exposure time.*

Lemma 4 (Optimality of Maximizing the Number of Photos). *Among all sequences with up to n tuples whose synthetic DOF is $[\alpha, \beta]$, the sequence that minimizes total exposure time has exactly n of them.*

Lemma 5 (Optimality of Equal-Aperture Sequences). *If $\beta < (7+4\sqrt{3})\alpha$, then among all capture sequences with n tuples whose synthetic DOF is $[\alpha, \beta]$, the sequence that minimizes total exposure time uses the same aperture for all tuples. Furthermore, this aperture is equal to*

$$D(n) \;=\; c\,\frac{\sqrt[n]{\beta}+\sqrt[n]{\alpha}}{\sqrt[n]{\beta}-\sqrt[n]{\alpha}} \;\; . \tag{11}$$

Proof of Lemma 1. We proceed inductively, by defining photo tuples whose DOFs "tile" the interval $[\alpha, \beta]$ from left to right. For the base case, the left endpoint of the first tuple's DOF must be $\alpha_1 = \alpha$. Now consider the i-th tuple. Equation (D) implies that the left endpoint α_i and the aperture diameter D_i determine the DOF's right endpoint uniquely:

$$\beta_i \;=\; \frac{D_i+c}{D_i-c}\,\alpha_i \;\; . \tag{12}$$

The tuple's focus setting in Eq. (5) now follows by applying Eq. (E) to the interval $[\alpha_i, \beta_i]$. Finally, since the DOFs of tuple i and $i+1$ are sequential, we have $\alpha_{i+1} = \beta_i$. ☐

Proof of Lemma 4. From Lemma 2 it follows that among all sequences up to length n whose DOF is $[\alpha, \beta]$, there is a sequence $\mathcal{S}^*$ with minimum total exposure time whose tuples have sequential DOFs. Furthermore, Lemmas 1 and 3 imply that this capture sequence is fully determined by a sequence of n' aperture settings, $D_1 \le D_2 \le \cdots \le D_{n'}$, for some $n' \le n$. These settings partition the interval $[\alpha, \beta]$ into n' sub-intervals, whose endpoints are given by Eq. (6):

$$\alpha = \;\alpha_1 \;<\; \overbrace{\alpha_2 \;<\; \cdots \;<\; \alpha_{n'}}^{\text{determined by } \mathcal{S}^*} \;<\; \beta_{n'} \;= \beta \;\; . \tag{13}$$

It therefore suffices to show that placing $n'-1$ points in $[\alpha, \beta]$ is most efficient when $n' = n$. To do this, we show that splitting a sub-interval always produces a more efficient capture sequence.

Consider the case $n = 2$, where the sub-interval to be split is actually equal to $[\alpha, \beta]$. Let $x \in [\alpha, \beta]$ be a splitting point. The exposure time for the sub-intervals $[\alpha, x]$ and $[x, \beta]$ can be obtained by combining Eqs. (D) and (1):

$$\tau(x) \;=\; \frac{L}{c^2}\left(\frac{x-\alpha}{x+\alpha}\right)^2 + \frac{L}{c^2}\left(\frac{\beta-x}{\beta+x}\right)^2 \;\; , \tag{14}$$

Differentiating Eq. (14) and evaluating it for $x = \alpha$ we obtain

$$\left.\frac{d\tau}{dx}\right|_{x=\alpha} \;=\; -\frac{4L}{c^2}\,\frac{(\beta-\alpha)\,\beta}{(\beta+\alpha)^3} \;<\; 0 \;\; . \tag{15}$$

Similarly, it is possible to show that $\frac{d\tau}{dx}$ is positive for $x = \beta$. Since $\tau(x)$ is continuous in $[\alpha, \beta]$, it follows that the minimum of $\tau(x)$ occurs strictly inside the interval. Hence, splitting the interval always reduces total exposure time. The general case for n intervals follows by induction. ☐

Proof Sketch of Theorem 1. We proceed in four steps. First, we consider sequences whose synthetic DOF is equal to $[\alpha, \beta]$. From Lemmas 4 and 5 it follows that the most efficient sequence, $\mathcal{S}'$, among this set has diameter and length given by Eq. (3). Second, we show that sequences with a larger synthetic DOF that are potentially more efficient can have at most one more tuple. Third, we show that the most efficient of these sequences, $\mathcal{S}''$, uses a single diameter equal to $D_{\max}$. Finally, the decision rule in Eq. (4) follows by comparing the total exposure times of $\mathcal{S}'$ and $\mathcal{S}''$.