Compressive Light Transport Sensing

Pieter Peers[*]   Dhruv K. Mahajan[†]   Bruce Lamond[*]   Abhijeet Ghosh[*]
Wojciech Matusik[§]   Ravi Ramamoorthi[‡]   Paul Debevec[*]

University of Southern California, Institute for Creative Technologies[*]
Columbia University[†]
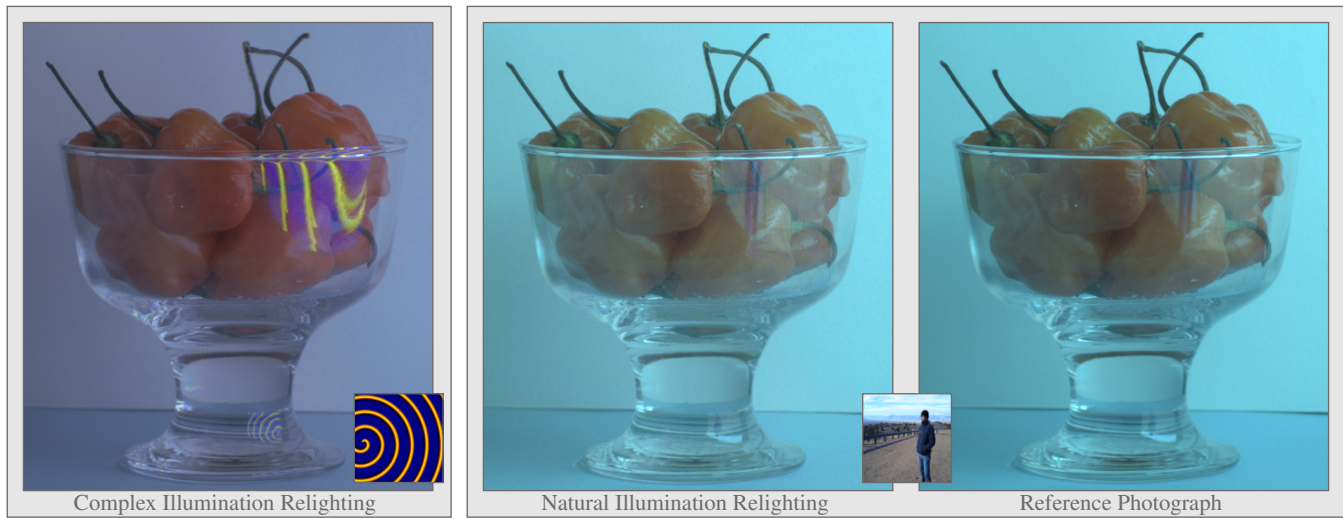University of California, Berkeley[‡]
Adobe Inc.[§]

Figure 1: An example of relit images of a scene generated from a reflectance field captured using just 1000 non-adaptive illumination patterns (emitted from the right onto the scene). The incident lighting resolution, and resolution of each reflectance function, is 128×128. Even though we only performed a small number of measurements, we are still able to capture and represent complex light transport paths. **Left:** the scene relit with a high frequency illumination condition (inset). **Middle:** the scene relit under a natural illumination condition (inset). **Right:** a ground truth reference photograph of the scene.

Abstract

In this paper we propose a new framework for capturing light transport data of a real scene, based on the recently developed theory of compressive sensing. Compressive sensing offers a solid mathematical framework to infer a sparse signal from a limited number of non-adaptive measurements. Besides introducing compressive sensing for fast acquisition of light transport to computer graphics, we develop several innovations that address specific challenges for image-based relighting, and which may have broader implications. We develop a novel hierarchical decoding algorithm that improves reconstruction quality by exploiting inter-pixel coherency relations. Additionally, we design new non-adaptive illumination patterns that minimize measurement noise and further improve reconstruction quality. We illustrate our framework by capturing detailed high-resolution reflectance fields for image-based relighting.

1 Introduction

The complexity in the appearance of real world scenes remains a principal challenge to simulate in computer graphics. Modeling and rendering such scenes under novel lighting with traditional computer graphics is an arduous task, which requires talent and experience. As a result, image-based representations have gained in popularity, wherein traditional modeling and rendering are replaced by image-based acquisition and relighting. Unfortunately, acquiring this image-based data is time and storage intensive. For instance, acquiring high-resolution datasets for scenes such as the bowl of peppers shown in Figure 1, can require tens of thousands of photographs and gigabytes of storage. In this paper we focus on acquiring light transport for the purpose of generating relit images of a scene using compressive sensing to greatly reduce acquisition time and storage.

Due to the acquisition complexity and storage requirements for such relightable datasets, many different methods have been developed to speed up the acquisition of light transport. Previous image-based techniques rely on sampling methods [Debevec et al. 2000; Wenger et al. 2005], adaptive methods [Matusik et al. 2004; Peers and Dutré 2003; Sen et al. 2005], or techniques using non-trivial measurement patterns [Peers and Dutré 2005] or specialized projector-camera setups [Garg et al. 2006].

A central concept in these image-based methods is a reflectance field, an 8D entity that abstracts the light transport through a scene in terms of incident and outgoing illumination on a bounding volume surrounding the scene. Both the incident and outgoing light field are 4D fields: for each position on the bounding volume (2D), all possible directions (2D) are considered. Capturing and handling these 8D fields is difficult. Therefore, most methods consider a reduced approximation. In this paper we will use the commonly used approximation proposed by Debevec *et al.* [2000], where the outgoing light field is reduced to a 2D field by fixing the viewpoint, and

where the incident light field is also reduced to a 2D field by assuming that incident illumination only varies positionally over the bounding volume.

The recently developed mathematical theory of compressive sensing [Donoho 2006; Candès 2006] offers a new theoretical background to improve image-based acquisition of light transport, and has potential for many other fields in computer graphics. Compressive sensing explores how to acquire a compressible signal from only a few *non-adaptive* measurements. Key to compressive sensing is the use of measurement patterns *incoherent* with the signal to be measured. In this paper, we investigate how compressive sensing can be adapted to acquire detailed reflectance fields for image-based relighting, reducing acquisition and storage costs by more than two orders of magnitude.

Conventional compressive sensing algorithms typically deal with a single function that needs to be measured and reconstructed. In the case of image-based relighting, many (reflectance) functions are sampled in parallel (i.e., one for each camera pixel), and each of these functions needs to be reconstructed individually using a compressive sensing algorithm. However, this is computationally expensive, and ignores spatial relationships between neighboring pixels' reflectance functions. Therefore, we have developed a novel hierarchical reconstruction algorithm that is able to extract the compressed signals using less computation, and with greater accuracy.

The "standard" measurement patterns used in compressive sensing assume an almost perfect acquisition system, where the only source of error is due to the observation of the measurements. In reality, other sources of error exist. For instance, in image-based relighting, measurements are performed by *emitting* measurement patterns from a controllable lighting device. This device can also introduce errors such as quantization. Although some research in compressive sensing has focused on designing optimal illumination patterns [Elad 2007; Weiss et al. 2007], these approaches are not suited for the scale required for our application. In this paper we propose a more practical approach, and give guidelines for designing illumination patterns that increase the signal to noise ratio of the measurements.

In summary, our contributions include:

- The introduction of compressive sensing for the acquisition of light transport, applied to image-based relighting.

- A novel hierarchical algorithm for decoding multiple compressive measurements that exploits spatial coherency to improve reconstruction quality and decrease computation costs.

- Practical guidelines for designing illumination patterns that improve signal to noise ratios of the measurements.

Notation. Throughout this paper we will use the following notational convention: a matrix is denoted by a bold capital, e.g., $\mathbf{M}$. A vector is denoted by a bold lower-case character, e.g., $\mathbf{v}$, and a scalar by a lower-case character, e.g., s. The j-th column of $\mathbf{M}$ is denoted as $\mathbf{m}_j$, the i-th row by $\mathbf{m}_{i,\cdot}$, and the element at the i-th row and j-th column as $m_{i,j}$. The i-th element of a vector $\mathbf{v}$ is denoted by v_i.

2 Related Work

In this section we will discuss related work, subdivided into two categories: measurement of light transport (Subsection 2.1), and compressive sensing (Subsection 2.2).

2.1 Measuring Light Transport

Measuring the light transport through a scene is a challenging problem, due to the dimensionality (8D) and the wide frequency behavior of real materials. In this section we will briefly discuss some of the milestones in this field relevant to our contributions.

Sampling Light Transport. Debevec *et al.* [2000] reduce the full 8D reflectance field to a 4D entity by fixing the viewpoint and assuming directional illumination only. The reduced reflectance field is then sampled using a *Light Stage*. A light source is moved to a finite number of positions (2D) around the subject, and a photograph is captured. Each photograph represents a 2D slice of the 4D reflectance field. Relighting can then be easily achieved by linearly recombining these images with appropriate weights. This method has been further extended to near instant acquisitions [Wenger et al. 2005], or to capture 6D reflectance fields, allowing relighting with more general 4D incident light fields [Masselus et al. 2003]. In general, the duration of the acquisition in these sampling based methods is proportional to the size of the *uncompressed* reflectance functions. In our method, however, the duration of the acquisition process is proportional to the size of the *compressed* reflectance functions. Since reflectance functions can be compressed very well without losing much accuracy [Masselus et al. 2004], this yields a significant speed-up in the acquisition process.

Adaptive Acquisition Methods. Peers Dutré [2003] propose to sample the reflectance field directly in the wavelet domain. To speed up the acquisition, an adaptive sampling scheme is used. A measurement oracle decides, based on previous samples from reflectance field, where to sample the wavelet domain next such that the potential for error reduction on the approximation of the full reflectance field is maximized. This method, however, has difficulty dealing with scenes containing many specular elements. Recently, Fuchs *et al.* [2007] presented an adaptive method to sample reflectance fields using point samples in the canonical domain. Other interesting recent developments are dual photography [Sen et al. 2005], and symmetric photography [Garg et al. 2006], where physical properties, reciprocity and symmetry respectively, of the transport matrix are exploited to speed up acquisition. Both use an adaptive acquisition scheme, which places most of the complexity on the acquisition system, as opposed to our method where the acquisition is very straightforward and most of the complexity is shifted towards post-processing where reflectance data is inferred from the observations after acquisition.

Beyond Conventional Sampling and Adaptive Methods. Sampling-based methods have a number of disadvantages, such as the practical limit on how many samples can be taken. Environment matting [Zongker et al. 1999; Chuang et al. 2000] proposes an alternative paradigm, but is unable to represent all types of light transport correctly. Matusik *et al.* [2004] and Peers Dutré [2005] presented methods that capture high-quality reflectance fields from a fixed set of non-adaptive illuminations patterns, natural illumination and wavelet noise respectively. Subsequently, an adaptive greedy algorithm is used to infer the reflectance functions. Both methods are very similar to compressive sensing in the sense that they use a fixed set of illumination conditions, and utilize an adaptive algorithm to decode the reflectance functions after acquisition. Our work differs from the above methods in that we are not restricted to a specific non-adaptive illumination pattern, but can choose from a whole family of patterns. Additionally, with the exception of [Matusik et al. 2004] none of the above methods leverage spatial coherence of the scenes' reflectance.

2.2 Compressive Sensing

In this subsection we give a brief introduction to compressive sensing, important terms, and relevant properties. Although compressive sensing [Donoho 2006; Candès 2006] is a relatively new research area, a large body of literature already exists. In order to keep this section concise, we will only discuss a select set of key points relevant to this paper. An extensive list of compressive sensing resources can be found at the compressive sensing resources web site [CSR].

Theoretical Background.   Consider a k-sparse discrete signal $\mathbf{x} \in \mathbb{R}^n$ that contains at most $k \le n$ non-zero elements (i.e., $||\mathbf{x}||_0 \le k$). Now define a measurement of this function by taking the dot product of the signal $\mathbf{x}$ and a measurement vector $\phi_j \in \mathbb{R}^n$: $y_j = \phi_j^T \mathbf{x}$. We can write m multiple measurements conveniently as a matrix-vector multiplication:

$$\mathbf{y} = \phi^T \mathbf{x}. \tag{1}$$

The $m \times n$ matrix ϕ^T is called a *measurement ensemble* in the context of compressive sensing. Conventionally (e.g., brute-force sampling), when the measurement ensemble has a rank n, the signal $\mathbf{x}$ can be faithfully reconstructed from the measurements, without any a priori knowledge of the signal. This implies that the number of measurements $m \ge n$. However, in compressive sensing the *a priori* knowledge that the signal is k-sparse is used to capture a signal in just $\mathscr{O}(k)$ *non-adaptive* measurements [Donoho 2006; Candès 2006].

Reconstruction. Before discussing the exact nature of these $\mathscr{O}(k)$ non-adaptive measurements, we need to consider the reconstruction $\bar{\mathbf{x}}$ of the original signal $\mathbf{x}$ from the measurements $\mathbf{y}$. This can be achieved by solving the following minimization problem:

$$\min_{\bar{\mathbf{x}}} ||\bar{\mathbf{x}}||_1, \quad subject\ to\ \ \mathbf{y} = \phi^T \bar{\mathbf{x}}. \tag{2}$$

This minimization searches for the approximation that explains the observed measurements under a specific measurement ensemble, and that is as sparse as possible. Sparseness is enforces by an ℓ_1-minimization, which is known to prefer sparse solutions, and which can be solved fairly efficiently using either the simplex algorithm [Dantzig 1963] or interior point methods [Wright 1997]. The optimization strategy that solves Equation (2) is also called *basis pursuit* [Chen et al. 2001; Donoho 2006]. Although basis pursuit is effective, the computational costs associated are still significant. As a result a number of alternative decoding strategies have been developed that solve minimization problems that approximate Equation (2).

Orthogonal Matching Pursuit (OMP), one of the first alternative decoding algorithms, and was studied by Tropp Gilbert [2005] in the context of compressive sensing. Although its accuracy is slightly below that of basis pursuit, it remains a very popular method due to its speed and ease of implementation. As opposed to basis pursuit, OMP is a greedy algorithm: it adds a single coefficient per iteration. Because coefficients are never removed, an erroneous addition of a coefficient can degrade the quality of the solution significantly. Recently, Needell Vershynin [2007b; 2007a] presented a promising variant of OMP, called Regularized Orthogonal Matching Pursuit (ROMP), that has better theoretical convergence properties, while maintaining the ease of implementation and speed of OMP. The decoding algorithm developed in this paper relies on the use of an existing reconstruction method for single functions in order to decode multiple measurements. Although any reconstruction method can be used, we opt for using ROMP.

A number of methods have considered using inter-measurement coherency relations to improve reconstruction results. Tropp *et al.* [2006] extend OMP to handle multiple measurements. However, unlike our method, it assumes that all measured signals have the same sparsity profile, i.e., they have the same non-zero coefficients. Cotter *et al.* [2005] make similar assumptions when decoding multiple observations. Ji *et al.* [2007a] extend their previous Bayesian compressive sensing method [Ji et al. 2007b] to decode measurements of multiple signals simultaneously, by using hierarchical Bayesian models. They show impressive improvements in decoding quality, but note that their method is sensitive to shifts in wavelet coefficients between different signals.

Relevant Properties. The above discussion assumed, unrealistically perhaps, that a signal is exactly k-sparse (i.e., at least $n - k$ terms are exactly zero). Many signals, such as reflectance functions, are not k-sparse but compressible [Liu et al. 2004; Masselus et al. 2004]. Candès Tao [2006] have shown that the same framework in compressive sensing also works very well for recovering the best k-term approximation of compressible signals.

The accuracy of the reconstruction of both sparse and compressible signals depends on two factors: measurement noise and approximation error. Candès [2006] showed that the error on the recovered signal is of order:

$$||\bar{\mathbf{x}} - \mathbf{x}||_2 \sim \mathscr{O}(\varepsilon_m) + \mathscr{O}(\varepsilon_a), \tag{3}$$

where ε_m represents errors due to measurement noise, and ε_a is the approximation error due to compression. Note, that the first term is optimal, i.e., the error is proportional to measurement noise and thus *not* amplified by compressive sensing. Similarly, the approximation error is also not amplified by the compressive measurements. In other words, given some target approximation error, compressive sensing is as good as the best adaptive methods, but without the overhead of taking adaptive measurements.

Finally, the number of measurements required to successfully reconstruct a k-term approximations of a signal of length n has been shown in the seminal work by Candès [Candès 2006] to be $k \log^c n$ measurements, with c some constant dependent on the measurement ensemble and reconstruction algorithm.

Measurement Ensemble. A remaining issue is what kind of measurement ensembles can be used to decode any measured sparse signal with a high probability of success. An important condition that determines whether a measurement ensemble is suitable for this goal, is the *restricted isometry condition (RIC)*. This condition, with parameters (k, ε), is defined as:

$$(1 - \varepsilon)||\mathbf{v}||_2 \le ||\phi^T \mathbf{v}||_2 \le (1 + \varepsilon)||\mathbf{v}||_2, \tag{4}$$

for every k-sparse vector $\mathbf{v}$. This condition states that the eigenvalues of any subset of at most k rows of the ensemble ϕ are at most ε removed from 1. This implies that any subset of at most k rows forms a basis for that particular subspace. This implies that an ensemble that adheres to the RIC form a basis for any k-sparse signal. Thus, enough information is available to reconstruct the original signal, on the condition that a sufficient number of measurements are recorded.

The concept of the RIC is straightforward, but it does not give a generative definition of practical measurement ensembles that meet this condition. It has been shown that assigning Gaussian or Bernoulli random values to each element of the ensemble ϕ, or randomly selected Fourier basis functions (i.e., randomly selected frequencies) fulfills these conditions [Baraniuk et al. 2008].

Randomly selected measurement ensembles have the advantage that they are general and easy to generate. However, they are not necessarily the most optimal for measuring any signal, nor are these ensembles optimized to the limitations of real world acquisition devices. Elad [2007] showed how to optimize the measurement ensemble, by minimizing mutual coherence between a dictionary of bases and the measurement ensemble. Although it improves reconstruction quality, it does not solve any practical issues with real measurement equipment. Weiss *et al.* [2007] developed a method called *Uncorrelated Component Analysis* (UCA) that attempts to learn the optimal measurement ensemble from a training dataset by maximizing the distance between the measured signals. Although their method is quite intriguing, it is not suitable for our particular problem for two reasons: First, the training dataset needs to be representative of the whole space of possible signals. Because we are dealing with signals defined on a large domain, a very large dataset is required, making the algorithm impractically slow. Second, UCA assumes that the signal is exactly sparse, and does not take into account approximation errors due to lossy compression.

3 Light Transport Sensing

In this section we explore how compressive sensing can be applied to acquiring light transport data, in particular to image-based relighting. As noted in numerous previous works [Ng et al. 2003; Peers and Dutré 2003], image-based relighting can be written compactly in matrix notation as:

$$\mathbf{c} = \mathbf{T}\mathbf{l}, \qquad (5)$$

where: $\mathbf{T}$ is a $p \times n$ matrix that defines the light transport between n light sources and p camera pixels; $\mathbf{c}$ represents these pixels in an observed camera image, stacked in a vector of length p; $\mathbf{l}$ represents the illumination conditions, stacked in a vector of length n. An illumination condition $\mathbf{l}$ can consist of any combination of point, directional, and area light sources. Each element in $\mathbf{l}$ indicates the emitted radiance of the corresponding light source. The relighting process consists of two stages: First there is the measurement stage that has as its goal determining the transport matrix $\mathbf{T}$, by observing the scene under a selection of m different illumination conditions $\mathbf{l}_j$, $j \leq m$; Second, there is the relighting stage, that has as its goal computing newly relit *virtual observations* given the measured transport matrix from the first stage, and a user-defined lighting condition. This second stage comes down to evaluating Equation (5) and is straightforward once the transport matrix is known.

For the acquisition stage, multiple illumination conditions $\mathbf{L} = [\mathbf{l}_0, ..., \mathbf{l}_m]$, and their corresponding observations $\mathbf{C} = [\mathbf{c}_0, ..., \mathbf{c}_m]$, can also be compactly denoted in matrix notation as:

$$\mathbf{C} = \mathbf{T}\mathbf{L}. \qquad (6)$$

Each row $\mathbf{t}_{i,\cdot}$ of the transport matrix $\mathbf{T}$ represents the reflectance function of the i-th pixel in the camera image. The observations of the i-th pixel are thus governed by an equation similar to Equation (1):

$$\mathbf{c}_{i,\cdot} = \mathbf{t}_{i,\cdot}\mathbf{L}, \qquad (7)$$

where $\mathbf{L}$ fulfills the role of the measurement ensemble ϕ, and the reflectance function $\mathbf{t}_{i,\cdot}$ corresponds to the discrete signal $\mathbf{x}$. Before compressive sensing can be applied to each reflectance function, we need to verify that these reflectance functions are either sparse or compressible. Previous research [Liu et al. 2004; Masselus et al. 2004] has shown that certain types of reflectance functions are indeed sparse or compressible in certain bases (e.g., spherical harmonics or wavelets). Suppose we have such a basis $\mathbf{B}$, and assume for simplicity that this basis is orthogonal (i.e., $\mathbf{B}^{-1} = \mathbf{B}^T$), then:

$$\mathbf{C} = \mathbf{T}\mathbf{L}, \qquad (8)$$
$$= \mathbf{T}(\mathbf{B}\mathbf{B}^T)\mathbf{L}, \qquad (9)$$
$$= \hat{\mathbf{T}}\mathbf{B}^T\mathbf{L}, \qquad (10)$$

where $\hat{\mathbf{T}} = \mathbf{T}\mathbf{B}$ is the transport matrix expressed in the basis $\mathbf{B}$. By choice of $\mathbf{B}$ we now know that $\hat{\mathbf{T}}$ is compressible, and thus suited for measurement by compressive sensing. Now define the illumination patterns as $\mathbf{L} = \mathbf{B}\phi$, where ϕ is one of the theoretical compressive sensing ensembles. Combining this with Equation (10) yields:

$$\mathbf{C} = \hat{\mathbf{T}}(\mathbf{B}^T\mathbf{B})\phi, \qquad (11)$$
$$= \hat{\mathbf{T}}\phi. \qquad (12)$$

We can now see that the observations $\mathbf{C}$ lit by these illumination conditions, directly map (after transposing both sides) to Equation (1) for each row $\hat{\mathbf{t}}_{i,\cdot}$, and that it still fits in an image-based relighting framework.

To summarize, the illumination patterns are the measurement (row) vectors ϕ_j from the ensemble, projected onto the inverse basis $\mathbf{B}^T$ (note that the transpose is because of the pre-multiplication in the definition of $\mathbf{L} = \mathbf{B}\phi$). To illustrate, suppose the measurement ensemble consists of independently and identically-distributed Gaussian random variables, and the basis is a Haar wavelet basis. By applying an *inverse* Haar wavelet transform on the Gaussian noise vectors, we obtain the illumination patterns. This effectively means that we are choosing Haar wavelet coefficients in a Gaussian way. Taking photographs of the scene illuminated by these illumination patterns yields the matrix of observations $\mathbf{C}$. Each row in this matrix (i.e., the observed values for a specific pixel location) yields a vector of measurements that can be used to reconstruct the original reflectance function for that pixel expressed in the basis $\mathbf{B}$.

Brute-force Compressive Light Transport Sensing. From the discussion above, we can now design a *brute-force compressive light transport sensing* algorithm. During acquisition, the illumination patterns as defined above, are emitted onto the scene, and for each a HDR photograph of the scene is recorded. Next, a reflectance function is inferred for each pixel separately by applying a compressive sensing decoding algorithm to the observation of only that pixel. As noted before, we use ROMP [Needell and Vershynin 2007b; Needell and Vershynin 2007a] in our implementation.

This brute force algorithm has the advantage that it is straightforward to implement, and all theoretical properties of compressive sensing are still valid. However, because each pixel is processed independently, there is no guarantee that spatial coherence is maintained. To illustrate, Figure 2 shows a synthetic scene consisting of
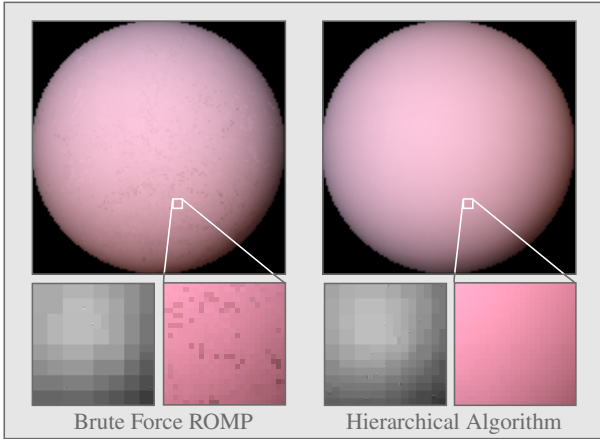
Figure 2: **Spatial coherency** of brute-force compressive light transport sensing (left) versus hierarchical compressive light transport sensing (right). The reflectance functions (128 × 128) are in both cases reconstructed from 1000 measurements using 100 Haar wavelet coefficients. For each algorithm, a detail is shown to better illustrate the differences in coherency. Additionally, a single reflectance function from this detail is shown on the bottom left of each example.

a diffuse ball. On the left a relighting is shown of the reflectance field computed using a brute-force approach, where each 128×128 reflectance function is reconstructed from 1000 Gaussian measurements (without adding measurement noise) using 100 Haar wavelet coefficients. While this shows that the brute-force approach is able to generate good results, some artifacts can be seen (i.e., spikes). On the right a relit result is shown of the same scene, reconstructed with the same number of coefficients per reflectance function, but with the algorithm presented in the next section that takes into account spatial coherency.

4 Hierarchical Compressive Sensing

In this section we introduce a novel hierarchical reconstruction algorithm that exploits spatial coherency to regularize the reconstructions from many simultaneously acquired compressive measurements. This is achieved by using a coarse-to-fine strategy to compute and refine reflectance functions. First we will provide some necessary mathematical insight into how a coarse-to-fine strategy can be used in the context of light transport sensing (Subsection 4.1). Next, we detail our algorithm in Subsection 4.2. We conclude this section with a brief discussion (Subsection 4.3).

4.1 Mathematical Basis

Studying the transport matrix $\hat{\mathbf{T}}$ gives us additional insight into its structure and how it can be used to develop a hierarchical algorithm. A key observation is that the projected transport matrix $\hat{\mathbf{T}} = \mathbf{TB}$ is only sparse/compressible over its rows, because the basis transform $\mathbf{B}$ only operates on the rows. Any coherency over the columns, and thus between different camera pixels, is not utilized.

In order to take advantage of these coherencies, we need to extend either the basis transform $\mathbf{B}$, or apply additional transformations to the transport matrix $\mathbf{T}$. Extending the compressive sensing algorithm to exploit inter-pixel coherencies during acquisition is not feasible, due to the natural parallelism of cameras (i.e., all pixels are captured at once), and thus the basis transform $\mathbf{B}$ cannot be adapted. However, we can still utilize inter-pixel relations during post-processing (decoding) by applying additional transformations on $\hat{\mathbf{T}}$.

To define these additional transformations, we note the similarity with image compression, where inter-pixel relations are used to minimize storage requirements. A common method to condense information is to apply a suitable basis transformation, such as a wavelet transform. After transforming the image, only the most important coefficients (i.e., with largest magnitudes) are stored. Again, this can be written compactly as a matrix multiplication similar to Equation (10). For example, an observed image $\mathbf{c}$ can be condensed into a few significant coefficients using a suitable basis $\mathbf{P}$ by the transformation: $\mathbf{P}^T \mathbf{c}$. Interestingly, when applying this to Equation (10) we obtain the following equation:

$$\mathbf{P}^T \mathbf{C} = (\mathbf{P}^T \hat{\mathbf{T}})(\mathbf{B}^T \mathbf{L}). \tag{13}$$

Observe that $\mathbf{P}^T \hat{\mathbf{T}} = \mathbf{P}^T \mathbf{TB}$, consists of *two* basis transformations on the transport matrix $\mathbf{T}$. The right transformation $\mathbf{B}$ operates on the rows (i.e., reflectance functions) of the transport matrix, while the left transformation $\mathbf{P}^T$ operates on the columns of the transport matrix (i.e., photographs). In other words, the right transform exploits coherencies within the reflectance functions, while the left transform condenses inter-pixel information. The resulting doubly transformed transport matrix is even more sparse, and thus is potentially inferable from fewer measurements at a higher accuracy.

4.2 Algorithm

We now develop our novel algorithm that exploits spatial relations. Equation (13) provides the necessary tool for utilizing inter-pixel relations during post-processing. In order to exploit inter-pixel relations at different scales, we employ a hierarchical basis to create a multi-resolution approximation of the acquired photographs, and to compute for each multi-resolution level, an approximation of the reflectance functions using compressive sensing. At each level, we initialize the compressive sensing reconstruction algorithm by the solution obtained in the previous level. The key assumption is that a coarse approximation at coarser levels, can be much more accurately estimated. During the compressive sensing decoding, coefficients are either updated, added or removed from this initial solution.

Instead of developing a novel compressive sensing reconstruction algorithm, we will use an existing compressive sensing decoding algorithm as a basis. To ensure maximum flexibility, and to anticipate future advances in the field of compressive sensing, we design our hierarchical algorithm such that *any* compressive sensing reconstruction algorithm can be used as a basis. As in section 3, we opt for using the Regularized Orthogonal Matching Pursuit (ROMP) [Needell and Vershynin 2007b; Needell and Vershynin 2007a] reconstruction algorithm in our implementation.

We will now discuss every step in detail of the presented hierarchical algorithm for which the pseudo-code is shown in Figure 3.

Multi-resolution initialization (Steps 1-2). Initialization of our algorithm starts by transforming the observations $\mathbf{C}$ using a suitable multi-resolution basis $\mathbf{P}$. This basis should exploit the coherencies within each column in the transport matrix $\hat{\mathbf{T}}$ as much as possible. This ensures that as much information as possible is condensed in coarser levels, and details are added in finer levels. A good choice for such a basis are bases suited for image compression, since each column corresponds to a recorded image of the scene. Probably the best known class of hierarchical basis functions that excel for compressing images are wavelets. In our implementation we use the Haar wavelet basis for $\mathbf{P}$.

Input: Observations $\mathbf{C}$ ($p \times m$ matrix) of the scene lit by the measurement ensemble ϕ ($n \times m$ matrix). k is a user-specified target sparsity. The wavelet transform $\mathbf{P}^T = (\prod_l \mathbf{P}_l)^T$, where $\mathbf{P}_l^T$ is the transform from wavelet level l to $l - 1$, $l = 0$ corresponds to the coarsest level, and $l = m_l$ to the finest level.

Output: For each pixel location i, a set S_i of coefficient indices, and their corresponding magnitudes $\hat{\mathbf{t}}_{i,\cdot}$, with $t'_{i,j} \neq 0$ only if $j \in S_i$.

Init: The initial set $S^{-1} = \emptyset$, and initial coefficients $\hat{\mathbf{T}}^{-1} = 0$. The coarsest level of the hierarchy: $\mathbf{C}^0 = \mathbf{P}^T \mathbf{C}$.

(1)   For every wavelet level: $l = 0 \rightarrow m_l$:
(2)   For every index i representing a scale coefficient in $\mathbf{C}^l$:
(3)   Copy initial the solution:
… $S_i^l = S_{parent(i)}^{l-1}$
… $\hat{\mathbf{t}}_{i,\cdot}^l = \hat{\mathbf{t}}_{parent(i),\cdot}^{l-1}$
(4)   Difference: $\mathbf{d}^T = \mathbf{c}_{i,\cdot}^l - \hat{\mathbf{t}}_{i,\cdot}^l \phi$
(5)   CS decode: $(R, \mathbf{r}) = decode(\mathbf{d}, \phi, k_{\mathbf{d}_l})$
(6)   Merge:
… $S_i^l \leftarrow S_i^l \cup R$
… $\forall j \in R : \hat{t}_{i,j}^l \leftarrow \hat{t}_{i,j}^l + r_j$
(7)   Limit size: $S_i^l \leftarrow$ indices of k_l largest coefficients in $\hat{\mathbf{t}}_{i,\cdot}^l$
(8)   Prune: $\forall j \in S_i^l$: if $|\hat{t}_{i,j}^l| < \delta$ then: $S_i^l \leftarrow S_i^l - j$
(9)   Update: $(\hat{\mathbf{t}}_{i,\cdot}^l)^T = \arg\min_{\mathbf{z}} ||\mathbf{c}_{i,\cdot}^l - \phi^T \mathbf{z}||_2$ s.t. $z_j \neq 0$ i.f.f. $j \in S_i^l$
(10) $\mathbf{C}^{l+1} \leftarrow \mathbf{P}_l \mathbf{C}^l$
(11) $S_i \leftarrow S_i^{m_l}$, and, $\hat{\mathbf{t}}_{i,\cdot} \leftarrow \hat{\mathbf{t}}_{i,\cdot}^{m_l}$

Figure 3: An overview of the hierarchical algorithm. At every level, and for every pixel, a compressive approximation is computed from the difference of the predicted measurement and the actual measurement. This predicted measurement is computed using the initial solution from the prior level.

Besides transforming the observations onto the wavelet basis $\mathbf{P}$, the initial set of active coefficients S^{-1} is set to empty, and the transport matrix $\hat{\mathbf{T}}^{-1}$ is set to zero. The set of active coefficients will contain the indices (of the approximation) of the reflectance field that are not zero. Only their corresponding magnitudes will be stored instead of the complete transport matrix to reduce storage costs.

Our multi-resolution algorithm iterates over all the multi-resolution levels (Step 1). For each level, an estimate for each reflectance function (at this resolution) is computed (Step 2). Note that the number of reflectance functions increases as the level increases. At the lowest level, only a single *global average* reflectance function is estimated. For each subsequent level, the number of estimated functions quadruples.

Difference Estimation (Steps 3-5).   These steps are the core of our algorithm. At each multi-resolution level, an estimate for each reflectance function is computed. As noted before, we would like to initialize the compressive sensing decoding algorithm from the corresponding reflectance function of the prior level's approximation. Unfortunately, many compressive sensing decoding algorithms do not provide a mechanism to refine an initial solution. Because we would like our method to work with any compressive sensing decoding algorithm, we cannot rely on the decoding algorithm to provide this mechanism for initialization.

To provide a mechanism to enforce initialization, even when the reconstruction algorithm does not support initialization, consider the following. A reasonable assumption is that the provided initialization is already a good match of the target reflectance function. Since it is a good match, the difference of the reflectance function and the initial approximation should be sparse, i.e., a few spikes and many (near) zero coefficients. Therefore, the difference is also a function suitable for compressive sensing. We can formalize this

by rewriting Equation (2) as:

$$\min_{\bar{\mathbf{d}}} ||\bar{\mathbf{x}}||_1, \quad subject\ to \quad \left\{ \begin{array}{l} \mathbf{y} = \phi^T \bar{\mathbf{x}}, \\ \bar{\mathbf{x}} = \bar{\mathbf{x}}_{init} + \bar{\mathbf{d}}, \end{array} \right. \quad (14)$$

where $\bar{\mathbf{x}}_{init}$ is the initial solution, and $\bar{\mathbf{d}}$ is the compressive approximation of the difference between the initial solution and the signal $\mathbf{x}$. Because $\bar{\mathbf{x}}_{init}$ remains constant during the ℓ_1 minimization, we can rewrite the minimization as:

$$\arg\min_{\bar{\mathbf{d}}} ||\bar{\mathbf{x}}||_1 = \arg\min_{\bar{\mathbf{d}}} ||\bar{\mathbf{d}}||_1. \quad (15)$$

Combining Equations (14) and (15), we obtain the following optimization rule:

$$\min_{\bar{\mathbf{d}}} ||\bar{\mathbf{d}}||_1, \quad subject\ to \quad \left(\mathbf{y} - \phi^T \bar{\mathbf{x}}_{init} \right) = \phi^T \bar{\mathbf{d}}. \quad (16)$$

Note that $\left(\mathbf{y} - \phi^T \bar{\mathbf{x}}_{init} \right)$ is a constant that can be precomputed before starting the minimization process. The minimization process itself is identical to Equation (2) (only $\mathbf{y}$ has been replaced by $\left(\mathbf{y} - \phi^T \bar{\mathbf{x}}_{init} \right)$, and $\bar{\mathbf{x}}$ by $\bar{\mathbf{d}}$), and thus the same algorithm can be used to compute this.

The above discussion gives a mechanism to enforce an initial approximation to any compressive sensing decoding algorithm, and covers steps (3) to (5) in the algorithm in Figure 3. First the solution from the previous level ($l - 1$) is copied into the set of non-zero coefficient-indices S_i^l, and into the corresponding magnitudes $\hat{\mathbf{t}}_{i,\cdot}^l$ in step (3). Next, the difference vector $\mathbf{d}$ is constructed similar to Equation (16), in step (4). Finally, in step (5) the minimization problem is solved using a compressive sensing reconstruction algorithm of choice. In our implementation we use ROMP as a decoding algorithm. The input parameters of the decoding algorithm are:

- The simulated measurements $\mathbf{d}$, computed in step (4).

- The measurement ensemble ϕ. As in the brute-force case, this ensemble defines the illumination patterns as: $\mathbf{L} = \mathbf{B} \phi$.

- The sparseness parameter $k_{\mathbf{d}_l}$ of the difference function. We set $k_{\mathbf{d}_l}$ to $k_{l-1} + 1$, where k_{l-1} is the maximum size of the initial solution. Basing the sparseness on the maximum size of the initial solution ensures that any incorrect coefficient in this initialization can be corrected. An in depth discussion regarding our choice of the sparseness and the implications of this choice can be found in Subsection 4.3.

The output is given by a set of computed coefficient-indices R and corresponding magnitudes $\mathbf{r}$. An overview of the hierarchical algorithm including the difference estimation is illustrated in figure 4.

Merging and Pruning (Steps 6-8).   Once an approximation for the difference function is found, an approximation for the complete reflectance function can be computed. This process consists of three steps: merging, size reduction, and pruning.

The first step, merging is trivial. The set of active coefficients S_i^l is updated by taking the union with the computed active coefficients R. The magnitudes $\mathbf{r}$ of the active coefficients R are summed to the corresponding magnitudes of S_i^l.
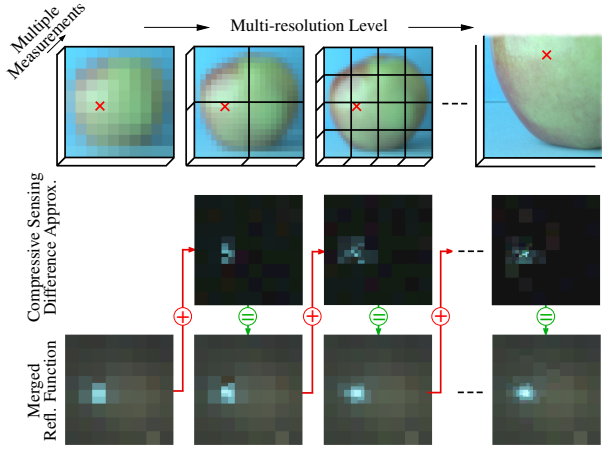
Figure 4: An overview of the hierarchical algorithm. For each multi-resolution level, compressive sensing is used to approximate the difference between the pixel's reflectance function and the approximation from the prior level. Merging the difference approximation and the previous level's approximation yields the desired function.

In addition, we need to ensure that the merged solution does not become too large, because otherwise the update step (Step 9) will become under-constrained. This also ensures that the target sparseness k of the final approximation of the reflectance field is met. Limiting the size of the approximation is obtained by keeping the k_l largest coefficients in the approximation. Smaller coefficients are set to zero. In our implementation the size k_l is defined as $2k_{l-1}+1$ coefficients, and where the size of the final solution (at level m_l) corresponds to the user-defined sparsity: $k_{m_l} = k$. In a normal situation the size of the merged solution should not exceed k_l, because $k_l = k_{l-1} + k_{\mathbf{d}_l}$. However, some compressive decoding algorithms, such as ROMP, can return more coefficients than requested. In Subsection 4.3, a detailed discussion is given on alternative functions for k_l and $k_{\mathbf{d}_l}$.

A third step takes care of removing coefficients that are close to zero. This can occur when the initial solution contains an erroneous coefficient. This will show up in the difference function as a large spike. The estimated magnitude for this spike will be of similar size (but with opposite sign) as the erroneous coefficient's magnitude, and thus cancel out by adding these magnitudes together. However, due to estimation errors these magnitudes will most likely not be exactly the same. Therefore, we remove any coefficient with a value below some threshold δ. This has the additional advantage that insignificant coefficients (mostly noise) will also be removed.

Update (Step 9). Finally, we update the coefficient magnitudes. This is achieved by computing a linear least squares solution for the non-zero coefficients such that the observations match the prediction, given the approximation and the ensemble, as well as possible.

In general, we found this step to be unnecessary, since the coefficient magnitudes obtained from the previous step are already very good. However, when many coefficient have been removed in the pruning step, some residual energy is unaccounted for. This unaccounted energy can potentially degrade the quality of the approximation, which can affect subsequent multi-resolution iterations. Some possible ways of minimizing the computational cost of this update step are:

- Only update when the total energy of the removed coefficients exceeds some pre-defined threshold. The difficulty in this case is setting the threshold. If it is too low, then too many updates are performed. If it is too high, potential errors can remain in the reconstruction.

- Similar to Peers Dutré [2005] a QR factorization can be maintained, and updated every step. This reduces the complexity of the least squares computation significantly.

In our implementation we opt for executing the update step fully at every level for robustness.

Multi-resolution Loop (Step 10-11). As noted before, a wavelet basis is used for $\mathbf{P}$. The wavelet basis transform is often described in terms of a cascading filter bank. The filter bank transforms a signal into two new signals: a high-pass filtered version, and a low-pass filter version. The exact frequency responses of the high-pass and low-pass filter depend on the specific wavelet used. To avoid duplication of information, both the high-pass and low-pass version are subsampled with half the number of samples. The filter banks are cascaded by repeatedly filtering the low-pass results. The cascade stops when the last low-pass version consists of only a single sample. Note, that the i-th low-pass version of a signal, can be reconstructed by adding the $i+1$-th high-pass and $i+1$-th low-pass version of the signal (both upsampled).

Now reconsider step (4). At this step we are actually computing the difference between two low-pass filtered versions of the same signal (at two different consecutive levels). $\mathbf{c}_{i,\cdot}^l$ represents the more detailed low-pass version, and $\hat{\mathbf{t}}_{i,\cdot}^l \phi$ the coarser low-pass version. It is tempting to equate this difference with the high-pass version of the signal (at the finer level), while in reality it is only an approximation, because the reflectance functions $\hat{\mathbf{t}}_{i,\cdot}^l$ at level l are only known approximately. It is for this reason that we explicitly compute the difference in step (4), and the low-pass version at level l in step (10). This last step is in fact just the reconstruction of the low-pass version at level l (i.e., to undo a single filter bank step of the wavelet transform).

Finally, the approximation of the last iteration is copied into the final solution (Step 11).

4.3 Discussion and Results

Parameters. The hierarchical algorithm takes two user-defined parameters as input, plus any additional parameters required by the basis decoding algorithm (Step 5): the threshold δ used in the pruning step (8), and the final sparseness k.

The threshold δ serves two goals. Its main goal is to determine when an erroneous coefficient needs to be removed, i.e., how similar the magnitudes need to be in the difference approximation and the initial solution. Second, it regulates denoising. Coefficients with a magnitude below this threshold are considered noise, and should therefore be removed from the approximation. Consequently, the value of this threshold should be set proportional to the magnitude of measurement noise. In the synthetic (noiseless) examples below, we set this parameter to some low value (e.g., 10^{-8}).

Choosing a specific sparseness k depends on a number of factors:

- The number of acquired (or planned) photographs is instrumental in setting an upper limit for the sparseness k. A lower bound on the number of measurements is given by $k \log^c n$ (Subsection 2.2), where n is the *uncompressed* size of the signal. In many cases, the budget for taking measurements is fixed beforehand, and thus the sparseness has to be chosen with respect to this budget.

- From the above it also follows that the resolution of the incident light field, and thus of the reflectance functions, plays a role in setting the sparseness. Note that it only depends on
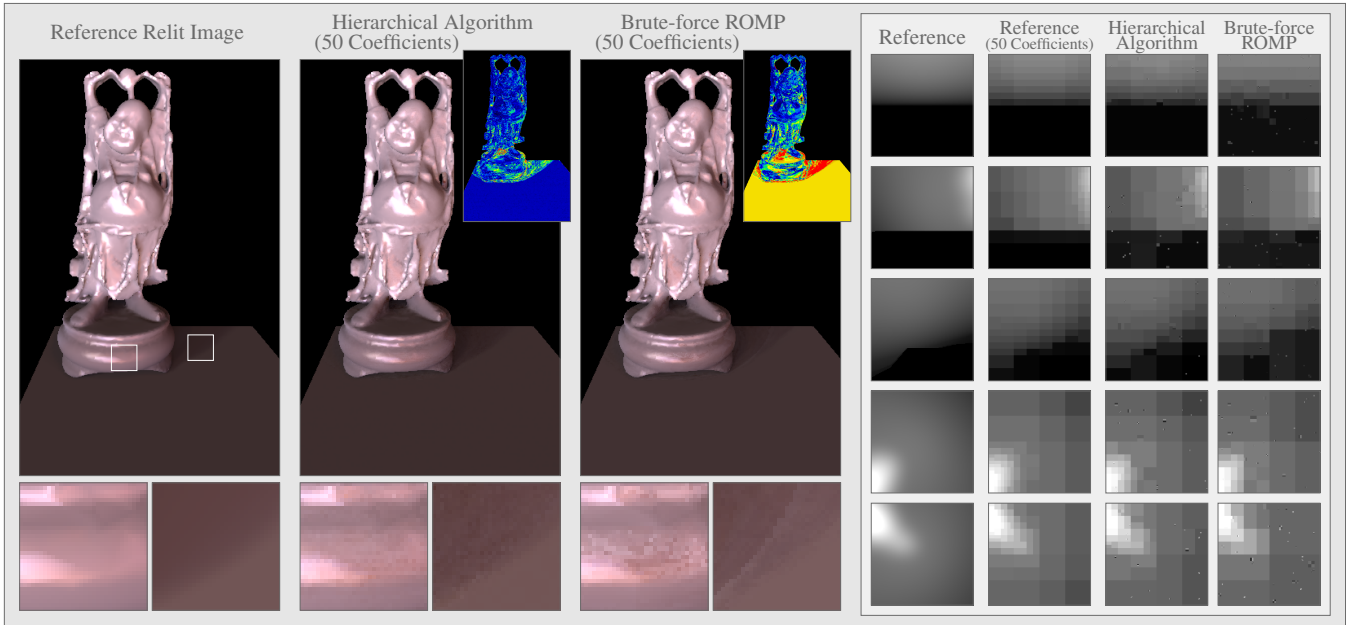
Figure 5: **Qualitative comparison. Left:** A relit image computed from a reference synthetic reflectance field. **Middle:** A relit image from a reflectance field (128 × 128 lighting resolution) inferred using the presented hierarchical algorithm from 512 (noiseless) measurements under a Gaussian ensemble. Each reflectance function is reconstructed using 50 Haar wavelet functions. **Right:** A relit image from a reflectance field computed using brute-force ROMP (512 Gaussian measurements, 50 coefficients per reflectance function). At the bottom of each of the relit images, two details are given. **Far right:** a comparison of five selected reflectance functions.

the logarithm of the resolution. This implies that compressive sensing enables greater efficiency when measuring high resolution reflectance functions.

- The desired accuracy of the approximation also plays an important role. A higher sparseness value implies that the approximation is coarser than when a low sparseness is used. Fortunately, reflectance functions can be well approximated using very sparse non-linear approximations. This is due to the limited number of high frequency features, such as specular peaks and shadowing, in these functions. Furthermore, higher resolution reflectance functions usually compress better than low resolution reflectance functions, augmenting the effectiveness of compressive sensing for measuring high resolution reflectance functions.

Qualitative Validation. To validate the quality of our method we have generated a full reflectance field of a synthetic scene. This allows us to generate ground truth relit images, and compare reflectance functions. Figure 5 shows such a scene containing the Buddha model with a glossy material on a diffuse underground. The leftmost image shows a reference relit image, and two details at the bottom. The middle image shows a relit image of a reflectance field (128 × 128 lighting resolution) inferred using our hierarchical algorithm from 512 (noiseless) measurements with a Gaussian ensemble[1] The sparseness k was set to 50. The inset shows the color-coded relative error image with the reference relit image. The right image shows a relit image of a reflectance field computed from the same measurements using a brute-force compressive light transport sensing algorithm (i.e., applying ROMP for every pixel independently). As can be seen, all three images are very similar. However, some differences can be seen in the detail images on the bottom. This is further corroborated by the difference images. Furthermore, the differences are more pronounced in the brute-force results.

---

[1]Adding measurement noise can mask the approximation errors made by the decoding algorithms. However, the conclusions of this validation would stay the same.
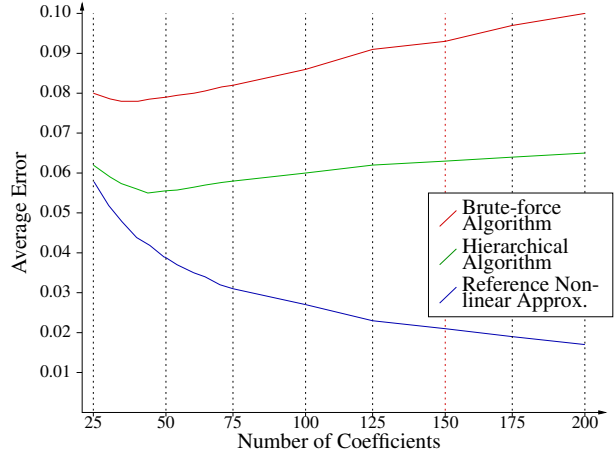

Figure 6: **Relative Error vs. Number of Coefficients.** The error in function of the number of coefficients for a fixed number of measurements (512 Gaussian measurements). In blue the error on a baseline non-linear Haar wavelet approximation is shown. In red the error on brute-force ROMP is plotted. Green shows the error of the presented hierarchical algorithm.

On the right of Figure 5, five selected reflectance functions from the Buddha scene are compared. The first column are the reference reflectance functions. The second column shows non-linear approximations using 50 Haar wavelet coefficients of the reference reflectance functions. These are the best possible approximations possible using 50 coefficients, and thus represent a baseline for comparison. The third and fourth columns are the computed compressive sensing approximations from the presented algorithm and the brute-force algorithm respectively. As can be seen, the reflectance functions computed using the hierarchical algorithm contain more details than the functions inferred using the brute-force technique. Both contain less detail than the baseline.

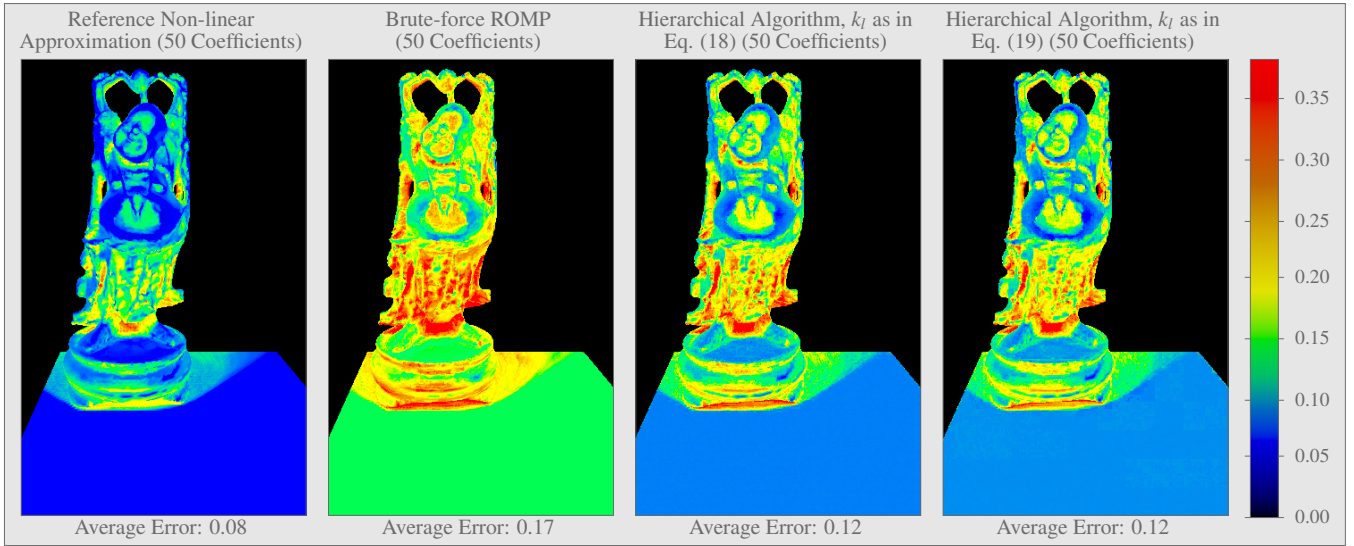**Quantitative Validation.** The qualitative comparison above

Reference Non-linear Approximation (50 Coefficients)	Brute-force ROMP (50 Coefficients)	Hierarchical Algorithm, k_l as in Eq. (18) (50 Coefficients)	Hierarchical Algorithm, k_l as in Eq. (19) (50 Coefficients)
Average Error: 0.08	Average Error: 0.17	Average Error: 0.12	Average Error: 0.12

Figure 7: **Error Comparison.** False color images of the relative error with respect to a reference reflectance field. From left to right: an error image showing a baseline 50 coefficient non-linear Haar wavelet approximation, the error on brute-force ROMP, the error on the presented hierarchical algorithm, and the error on the hierarchical algorithm using the alternative formulas (19) for the per-level sparseness and difference sparseness.

shows that the presented algorithm has the potential of generating better results. However, this qualitative comparison does not provide information on what the optimal number of coefficients are, or how much better the hierarchical algorithm is. A quantitative analysis can give an answer to these questions.

Figure 6 plots the error on the computed reflectance field (from 512 measurements) of a representative 32×32 pixel area of the scene shown in Figure 5 with respect to the reference reflectance field. The error is computed as:

$$\varepsilon = \frac{\sum_i |\mathbf{t}_{i,\cdot} - \bar{\mathbf{t}}_{i,\cdot}|_2}{p \sum_i |\mathbf{t}_{i,\cdot}|_2}, \tag{17}$$

where p is the number of reflectance functions, $\mathbf{t}_{i,\cdot}$ the i-th reference reflectance function, and $\bar{\mathbf{t}}_{i,\cdot}$ the corresponding computed approximation. The blue graph in Figure 6 represents the baseline error. This error is obtained by computing a non-linear Haar wavelet approximation of the reference reflectance field. The number of coefficients in this non-linear approximation is variable (horizontal axis). The green and red graph depicts the error on the solution obtained using the presented hierarchical algorithm, and the solution obtained from using brute-force ROMP respectively. In both cases we notice a decrease of error, reaching a minimum at about 50 coefficients[2], followed by a steady increase. The initial decrease can be explained by the fact that adding additional coefficients increases the representational power of the approximation. However, at a certain point, the number of measurements is no longer sufficient to uniquely select and compute coefficients. At this point, many possible *optimal* approximations exist that minimize the error between the real and simulated measurements. As a result, *noise* is introduced into the solution, and the approximation begins to deviate from the real solution. From this graph we can see that the presented hierarchical algorithm delivers a better approximation with a significantly lower error than the brute-force approach. Additionally, due to the spatial regularization, the error increases less rapidly after reaching the optimal number of coefficients (i.e., 50). This

suggests that over-estimating the optimal number of coefficients is less harmful to the reconstruction quality of our algorithm than for the brute-force approach.

Alternative Sparseness Settings. Two important variables in the presented hierarchical algorithm are the total sparseness per level k_l, and the sparseness of the difference $k_{\mathbf{d}_l}$. Previously we have set these variables to:

$$\begin{aligned} k_{\mathbf{d}_l} &= k_{l-1} + 1, \ and \\ k_l &= k_{l-1} + k_{\mathbf{d}_l}. \end{aligned} \tag{18}$$

Setting these variables to these formulas results in a doubling of the size of the reflectance functions with every consecutive level. Furthermore, it also means that at every level, two times the number of variables are added compared to the previous level. Each extra variable, requires additional computation time. Thus, more time is spent at the finer levels than at the coarser levels. This effect is further amplified by the fact that the number of reflectance functions also quadruples per level. Although, computation time is slightly below that of the brute-force approach, this still raises the question if other formulas for k_l and $k_{\mathbf{d}+l}$ can be found, that limit the number of coefficients that need to be computed at finer levels. One such alternative is:

$$\begin{aligned} k_{\mathbf{d}_l} &= \frac{k}{m_l}, \ and \\ k_l &= l \frac{k}{m_l}, \end{aligned} \tag{19}$$

where m_l is the total number of levels. In this case, the number of coefficients computed at each level stays constant, and thus yields a more efficient (in terms of computation time) algorithm.

Figure 7 shows a color coded error image comparing both variations, and the brute-force algorithm to a reference reflectance field. For each image the error is computed using Equation (17), and

---

[2]Increasing or decreasing the number of measurements, would similarly change the optimal number of measurements.
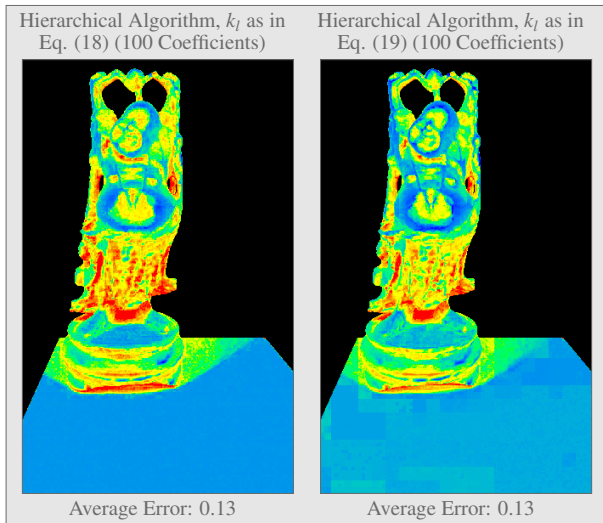
Figure 8: **Error Comparison on Non-optimal Sparseness Selections.** The hierarchical algorithm is used to compute a 100 term approximation from 512 Gaussian measurements. The left image uses the original per-level and difference sparseness formula (18). The right image uses formula (19) that reduces the computation costs, but at the cost of quality. The same false color scale as in Figure 7 is used for both images.

plotted per reflectance function. In all cases the reflectance functions were reconstructed using 50 coefficients computed from 512 Gaussian measurements. A baseline error image of a non-linear approximation using 50 Haar wavelet coefficients is also provided for reference. These error images confirm the conclusion from the quantitative analysis of before, and show that the hierarchical algorithm significantly outperforms the brute-force algorithm in quality. Both variations at the hierarchical algorithm yield a similar error. However, the alternative sparseness shows a bit more structure in the error image, as can for example be seen on the diffuse underground. This effect is further amplified when selecting a non-optimal sparseness k, as illustrated in Figure 8, where 100 coefficient approximations of both variations are compared. From this we can conclude that, although the originally proposed formula for the per-level sparseness and difference sparseness yield a less optimally performing algorithm, the quality gain (in spatial coherency) and robustness to non-optimal sparseness selection is better. Depending on the preferences of the user, either formula can be selected. In this paper we will use the improved quality variant, Equation 18).

5 Capturing High-resolution Reflectance Fields

In this section we discuss how to apply compressive sensing, and in particular our hierarchical algorithm, to the acquisition of real light transport data. In particular, we will focus on capturing high resolution reflectance fields for image-based relighting.

Although applicable to other kinds of acquisition setups, our discussion focuses on a setup, similar to that used in previous work [Zongker et al. 1999; Matusik et al. 2004; Peers and Dutré 2005], consisting of a single video camera and a CRT monitor as a controllable high resolution light field emitter. The CRT monitor is placed either to the side or behind the scene which is imaged from a fixed camera position. Additionally, in order to cover a larger portion of the sphere of incident lighting directions, we have also experimented by replacing the CRT by a hemispherical controllable light source [Peers et al. 2007]. Except for the coverage, this setup is identical to a CRT from a practical point of view.

5.1 Practical Issues

A straightforward application of compressive sensing could be the following: First, the emitter and the camera are radiometrically calibrated (e.g., inverting the gamma curve of the CRT), such that linear radiance values are measured or emitted. Next, each measurement vector ϕ_j from a measurement ensemble (for instance Gaussian random Haar wavelet coefficients) is de-linearized into an illumination pattern (2D image), and emitted from the CRT monitor. A high dynamic range photograph of the scene under this illumination condition is recorded. This recorded image is linearized and stored in a column of the observation matrix $\mathbf{c}_j$. This is repeated for all measurement vectors. Afterwards a decoding algorithm operates on the acquired data.

Unfortunately, this straightforward approach does not work well. The main reason for this is the interplay between measurement noise, quantization errors at the emitter and the normalization of wavelet bases.

- **Measurement noise** occurs for every measurement at each pixel on the camera sensor. Its magnitude is independent of the measurement pattern, or of the reflectance field. Therefore, its relative impact is more severe when a reflectance functions reflects little light towards the camera (i.e., *dark* reflectance functions)[3]

- **Quantization errors** occur at the camera and at the emitter. Quantization errors at the camera are usually modeled as part of measurement noise. Quantization errors at the emitter, however, are not. This error depends not only on the measurement patterns, but also on the reflectance functions, because every illumination element has a different quantization error. Given the vector of quantization errors $\mathbf{q}$ (i.e., the difference between each element in the theoretical measurement vector and the actually emitted measurement vector), the quantization error ε_{quant} on the measurement of a signal $\mathbf{x}$ can be expressed by:

$$\varepsilon_{quant} = \mathbf{q}^T \mathbf{x}. \qquad (20)$$

Note, that both $\mathbf{x}$ and $\mathbf{q}$ are expressed in the canonical domain, and not in the wavelet domain. If the measurement vector is random enough, then the elements in $\mathbf{q}$ are randomly distributed. In such a case, ε_{quant} can be seen as a weighted sum, where the signal $\mathbf{x}$ acts as a weighting vector, of random variables. If we assume that $\mathbf{x}$ contains k non-zero elements of approximately the same magnitude, then is the quantization error ε_{quant} normally distributed with a standard deviation proportional to $\frac{1}{\sqrt{k}}$ according to the central limit theorem. From this it follows that measurements of reflectance functions with a large support (e.g., diffuse reflectance functions) suffer less from quantizations errors, while measurement of reflectance functions with a compact support (e.g., specular reflectance functions) suffer more from quantization errors.

- **Wavelet normalization** ensures that the energy content of wavelets at different scales are the same. This implies that wavelets with a large spatial support (i.e., coarse wavelets) will be scaled more than wavelets with a small spatial support

---

[3]Real camera systems are also influenced by shotnoise in addition to additive noise. However, in this analysis we will omit shotnoise. Future research is necessary to investigate the exact effects of this factor on the design guidelines presented in this section.

(i.e., fine wavelets). For 2D wavelet bases, this scaling is proportional to the wavelet level l. In the specific case of assigning randomly distributed values to Haar wavelet coefficients, the scale of each wavelet level l is according $\frac{1}{4}\left(2^{-(m_l-l)}\right)$. The factor $\frac{1}{4}$ ensures that any generated pattern fits in the $[0..1]$ range.

The interplay of these three factors causes problems when estimating the coefficients of coarse wavelets. These coefficients are the response of the reflectance function with the randomly weighted coarse wavelet functions. However, due to the normalization, the amplitude of these coarse wavelet functions is very small. When the support of the reflectance function falls below some threshold, the response of these coarse wavelets can fall below the quantization error, and thus many measurement need to be performed to reliably infer these coefficients. This effect is is further amplified by measurement noise.

A potential solution would be to remove the quantization error from the system, by using measurement ensembles in which the effects of quantization are included. Unfortunately, the coefficients of coarse wavelet coefficients will be mapped in a large number of cases to zero or near-zero values, and thus no information about these coefficients is measured. Furthermore, the random distribution of the coefficients is significantly altered, and the RIC is potentially broken.

Peers Dutré [2005] circumvented the above problems by using a different normalization of the Haar wavelet basis. Instead of normalizing according to energy content, they use an alternate normalization such that all wavelet basis functions have a similar amplitude. The effect is that each wavelet level takes an equal portion of the emitter's dynamic range (i.e., $\frac{1}{4m_l}$). Although this alternative normalization would also work to some degree in compressive light transport sensing, it still suffers from some issues. First, this solution does not scale well with respect to sample resolution. For example, a 128×128 sampling results in only 5 intensity values per wavelet level (assuming 140 linear and regularly spaced intensities). Second, by using an alternative normalization of the wavelet basis, the non-linear approximation of the reflectance functions is also altered, and thus the obtained approximations are not optimal in an L_2 sense.

5.2 Illumination Ensemble Design

In this subsection we will design a novel measurement ensemble that does not suffer from the issues described above. First, general design guidelines are presented that improve the signal to noise ratio (SNR) of the ensemble. Next, we introduce a binary measurement ensemble that does not suffer from the issues of the *standard* compressive sensing ensembles and ensures good signal to noise ratios.

Design Guidelines. In order to design measurement patterns that are suited for compressive light transport sensing we define the following guidelines. First, to ensure that the measurement patterns fall within the theory of compressive sensing, we strictly adhere to the restricted isometry condition (RIC) defined in Subsection 2.2. Second, in order to optimize the signal to noise ratio of the measurement patterns we keep in mind the following:

- Increasing the dynamic range of the measurement of a coefficient by a factor c, increases SNR by a factor c^2.

- Increasing the number of measurements of a coefficient by a factor c, increases SNR by a factor c.

In order words, increasing dynamic range for the measurement of a coefficient is more effective than increasing the number of measurements for that coefficient.

Segregated Ensembles. In order to design new measurement patterns, we start from segregated measurement ensembles, and extend them using the design guidelines above.

In [Donoho and Tsaig 2006; Donoho 2006] segregated measurement ensembles are considered. In the context of a measurement ensemble defined on the wavelet domain, a segregated ensemble corresponds to the subdivision of the domain in disjunct partitions, and where each partition only covers wavelet coefficients from a specific wavelet level (or a subset of levels). A single measurement vector has random distributed values assigned only to a single partition. All coefficients in the other partitions are set to zero.

Now, consider the Haar wavelet basis, and a subdivision in m_l partitions $P_i, i \in [0, .., m_l - 1]$ (i.e., each partition only contains a single wavelet level). Because a measurement vector only contains wavelet coefficients from a single wavelet level, the number of overlapping (in terms of spatial support) wavelet basis functions is limited, unlike the *standard* measurement ensembles, were the overlap is maximal. For example, for the Haar wavelet, only three Haar functions can overlap (i.e., the three wavelet orientations for the same wavelet position), while in the standard case this overlap is $3m_l$ functions. Consequently, the dynamic range can be increased by a factor $2^{(m_l-i)}$ for the measurement vectors corresponding to each partition P_i compared to the *standard* measurement ensembles. However, segregation also comes at a cost. Each coefficient, will only occur as many times as there are measurement vectors associated with the coefficient's corresponding partition. In order to maintain SNR, the number of measurement vectors associated with each partition should be at least $m\left(\frac{1}{2^{(m_l-i)}}\right)^2$. If we assign the minimum number of measurement vectors to each partition, than we only need (in the limit) $\frac{m}{3}$ measurements. In other words, only one third of the measurement budget has been used. This implies that if we take triple the number of measurement vectors per partition, that the SNR triples.

A different perspective on segregated measurement ensembles is that a signal is split up into many unrelated sub-signals. Each of these sub-signals is sparse, and can be inferred independently from the other sub-signals. The measurement ensemble for inferring each sub-signal can thus be optimized to stretch the limitation of the measurement devices maximally. In the proposed segregation, the normalization of wavelet basis functions is exploited to increase the dynamic range during measurement. This different perspective also reveals an important point omitted in the above discussion. As noted before in Subsection 2.2, the number of compressive measurements required to infer a k-sparse signal is $k \log^c n$. Thus, this depends on the number of measurements, and the size of the signal n. This rule also holds for the segregated sub-signals. The optimal partitions size, that maintains the SNR, might not be realistic when taking this into account, especially for the coarse wavelet levels. In general, the sparsity of the sub-signals increases as the wavelet level increases. Therefore, more measurements than previously suggested (to maintain the SNR) are focused on the coarser partitions. The net effect is that the SNR increases more on the coarser wavelet coefficients, and less on the finer wavelet coefficients.

Binary Segregated Ensemble. From the previous discussion, it can be seen that an increase in SNR is attained by reducing the overlap of wavelet basis functions for each spatial location. At this point, there is still an overlap of three basis functions per spatial location. This raises the question whether this can be further reduced.
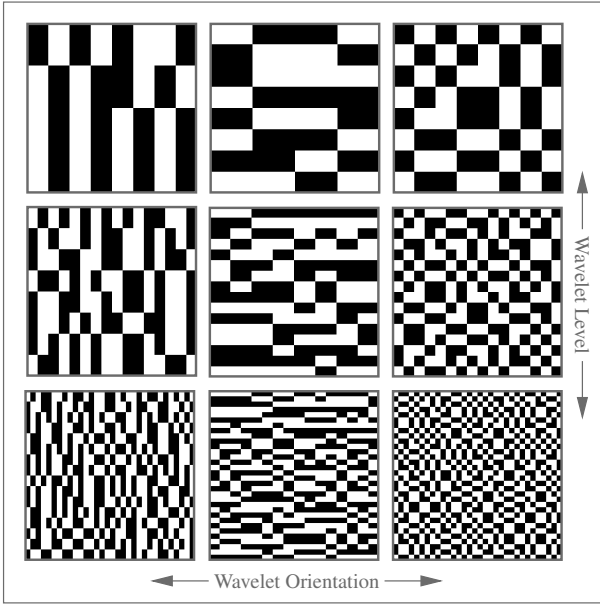
Figure 9: **Measurement Patterns.** Examples of binary segregated measurement patterns. Three different wavelet levels are shown, and for each level, the three different wavelet orientations are shown.

Consider the subdivision of each partition P_i into three new partitions $P_i^j, j \in \{0, 1, 2\}$, where each new partition only contains a single wavelet level as well as wavelet orientation[4]. This ensures that no overlap exists when using the Haar wavelet. Dynamic range increases by a factor 4, while the number of measurements per coefficient decreases by a factor of 3. The total gain in SNR from this segregation is thus $\frac{16}{3}$.

In order to further maximize SNR, we will assign Bernoulli random variables ($P(-1) = 0.5$ and $P(+1) = 0.5$) to the coefficients in each partition. The resulting measurement vectors will consist of only two intensity values. These binary patterns have the advantage that no quantization errors are made when emitting these patterns, and that no radiometric calibration of the emitter (i.e., gamma correction) is necessary. This further reduces the possibility for calibration and measurement errors. A selected set of nine random measurement patterns are depicted in Figure 9.

As demonstrated, the binary segregated patterns yield a better SNR than many of the standard measurement ensembles. Furthermore, they also fulfill the restricted isometry condition (Equation (4)), and thus are suited for compressive measurements. A disadvantage is that the size of the partitions needs to be selected before acquisition. The optimal size is related to the distribution of the coefficients in the to-be-measured functions, which is unknown. However, in most cases general properties are known, and a well educated choice of the partition size can be made.

5.3   Direct Measurement

As discussed in Subsection 2.2, at least $m = k \log^c n$ measurements are required to ensure a successful decoding of a k-sparse signal. From this it follows that compressive sensing has a small overhead

---

[4]A 2D wavelet basis is the result of the outer product of 1D scale and wavelet basis functions for the horizontal and vertical axes. In the non-standard wavelet basis, each spatial location has three orientations for each wavelet level: horizontal scale $\times$ vertical wavelet, horizontal wavelet $\times$ vertical scale, and horizontal wavelet $\times$ vertical wavelet.

of $\frac{m}{k}$ measurements. Segregating the ensemble into disjoint partitions P_i, effectively subdivides the signal in #P independent parts. As a result, each partition has its own sparsity k_{P_i}, and thus an overhead of $\frac{m_{P_i}}{k_{P_i}}$, where m_{P_i} corresponds to the number of measurement vectors assigned to the segregated set P_i. For the fine detail wavelet coefficients, this ratio is still acceptable because the probability of a coefficient being important is very small. For the coarse wavelet levels, however, this ratio can actually be larger than the probability that a coefficient is among the k most important ones. It would be suboptimal to use compressive sensing in such a case. In general, if the probability of a coefficient being amongst the k largest coefficients is less than $\frac{k}{m}$ then a compressive sensing approach is optimal. If, however, the probability is larger than this ratio then a direct measurement of the coefficients is better suited.

The binary patterns proposed in Subsection 5.2 easily allow splitting the measurement ensemble into direct and compressive sensing measurements. To maintain a good signal to noise ratio, we opt for using Hadamard patterns to perform the direct measurements. Practically, we found that the first three wavelet levels are more efficiently captured using direct measurements.

6   Results

All the results in this section have a lighting resolution of 128×128, and are captured using 991 patterns:

- 64 direct measurements, i.e., the 3 coarsest wavelet levels.

- 153 compressive measurements for the 4-th level, i.e., 51 compressive measurements per orientation for this level.

- 258 compressive measurements for each of the remaining three levels, i.e., 86 measurements per orientation for these levels.

As noted before, we use Hadamard patterns for the direct measurements, and the binary segregated patterns developed in Section 5 for the compressive measurements. Each reflectance function is reconstructed using 128 wavelet coefficients in total (of which 64 form the direct measurements), and the resolution of the incident light field is 128×128. This yields two orders of magnitude acquisition speed-up and data reduction.

Figure 10 shows a glass fish scene captured using the hemispherical emitter of Peers *et al.* [2007]. On the left, a relit image of the glass object is shown under a complex illumination condition (shown in the inset). In the middle a reference image is shown. This reference image is generated by recording a photograph of the scene while emitting the complex illumination condition. The color responses and gamma curves of the emitter were carefully measured to ensure that both scenes are compared under exactly the same illumination conditions. On the right, a false color image of the relative error between the relit and reference image is shown. As can be seen, both the relit and reference photograph are very similar, especially the reflected high frequency patterns are well preserved (e.g., on the dorsal fin). There are some differences due to the non-linear approximation of the reflectance functions, and measurement noise. For example, a faint blue caustic can be observed on the red background in the reference image. This caustic is less visible in the relit image. However, as can be seen in the false color error plot, the error on these caustics falls well below the average error. The error is biggest on areas with low albedo, and little incident illumination, in the reference photograph, and is dominated by measurement noise. Because this image represents a relative error plot, measurement noise (an absolute effect) is amplified as the pixel's albedo decreases. The total average relative error for this scene under this particular illumination condition is 7.3%.
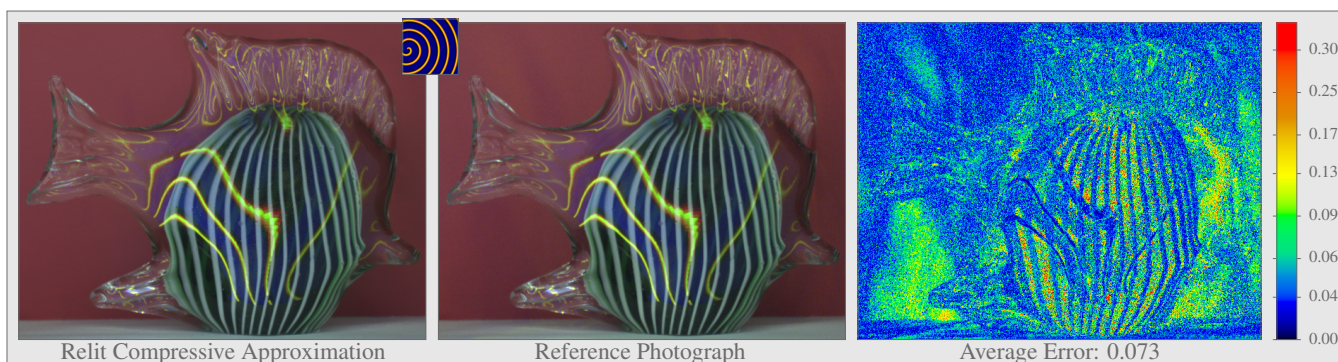
Figure 10: **Quantitative Comparison of an Acquired Reflectance Field. Left:** a relit image of a reflectance field (128×128 lighting resolution) inferred from 991 measurements (64 direct, and 927 compressive), relit using the complex illumination shown in the inset. **Middle:** a reference photograph of the scene lit by the same complex illumination conditions. **Right:** a false color relative error plot of the relit and reference image.

Figures 1 and 11 show three additional scenes acquired using the presented method and illumination patterns. The illumination patterns were emitted from a CRT monitor positioned to the right of the scene.

Figure 1 shows a glass container filled with Habanero peppers. The left subimage shows a relit image under complex illumination conditions (shown in the inset). As can be seen, the specular reflections (primary and secondary) are nicely reproduced. The right subfigure shows a comparison of a relit image and a reference photograph under a natural illumination condition, i.e., a photograph (shown in the inset). Both images are very similar.

The left column of Figure 11 shows a model motorbike containing specular (exhaust pipe), glossy (body of the bike), and diffuse materials (ground plate and backdrop). Figure 11.a shows the motorbike relit using the directly measured components only. The incident light field is shown in the inset on the bottom right of the photograph. Note that most glossy and specular reflections are missing. In Figure 11.b the same illumination is applied, but now with the compressive sensing coefficients included. As can be seen, the direct measurements encode most of the diffuse reflection effects, and a sampling resolution of 8×8 for a single side of the hemicube suffices. Note, that due to measurement noise, it is difficult to obtain high resolution reflectance fields using only direct measurements (with the same exposure times). Figure 11.c shows the motorbike illuminated by a more challenging illumination condition: three colored light sources. Some noise can be observed on the exhaust pipe. As noted in Subsection 2.2 measurement noise is not amplified due to the RIC.

The right column of Figure 11 shows a scene containing a textured ceramic object and an apple. The ceramic object contains a strong specular component on top of a textured diffuse layer. The apple also has a rich texture, and consists of a glossy reflective layer on top of a scattering medium. A comparison between a ground truth photograph (Figure 11.d) and a relit image (Figure 11.e) under a natural illumination condition (shown in the inset) shows that the presented method is capable of capturing scenes with visual accuracy. Figure 11.f shows the scene under three colored light sources with different intensities. Again, this shows that our method can deal with more extreme lighting conditions.

Figure 12 shows six selected reflectance functions from the scenes in Figure 11. For each reflectance function, two different exposures (not necessarily the same for each function) are shown to better show the full extent of the reconstructed functions. Note that our method is capable of capturing sharp specular peaks, glossy peaks, and diffuse lobes.

Figure 13 shows an interesting experiment. The reflectance field of the toy motorbike is captured using 927 compressive and 64 direct measurements using the hemispherical emitter. Next, a relit image is created by illuminating it by a single illumination element (out of 128×128). To better illustrate the effect, the image has been brightened 4069 times (left), and 16 times (right). As expected, especially after brightening the image 4096 times, some measurement noise is visible in the relit images. Unless a very long exposure photograph were taken, recording a reference photograph under similar lighting conditions would likely exhibit similar or worse measurement noise.

7 Discussion

We have shown that compressive sensing is a valid method to capture light transport in a scene. Compared to adaptive methods the acquisition setup is less complex, and thus more robust. However, the post-processing phase is computationally more expensive. In effect, the complexity of the acquisition phase has been shifted to the post-processing phase. This shift can be justified by observing that progress in computational power is much faster, than the rate of improvement in sensor sensitivity. Adaptive methods on the other hand have the advantage that they can adapt to the nature of the scene and avoid potential problem areas (e.g., oversaturation, external light contamination), and can even recapture parts of the scene if necessary. Compressive sensing, however, does not have that luxury, and contaminations of the measured data can affect the quality of a captured dataset adversely. In controlled acquisition environments, the need for such robustness is of lesser concern.

In the remainder of this section we will discuss the differences between the presented method and two of the most closely related previous works, i.e., [Matusik et al. 2004], and [Peers and Dutré 2005].

7.1 Comparison to Related Methods

Matusik *et al.* [2004] infers reflectance fields from natural illumination, i.e., photographs of natural scenes. Each reflectance function is represented by a sum of non-overlapping kernels, i.e., box functions. Each reflectance function is inferred using a greedy algorithm, where at each step a kernel is split horizontally or vertically. The splitting direction, and kernel to split are selected such that the error decreases maximally. A quadratic programming approach is used to update the magnitudes of each kernel function after splitting such that non-negativity of the reflectance function is maintained and error is minimized. Because the complexity of this algorithm increases quadratically with the number of box func-

Figure 11: Acquired Reflectance Fields. Left column, a relit toy motorbike is shown. **(a) and (b)** are lit by the illumination condition shown in the inset. Subfigure (a) only shows the effect of directly measured coefficients, while (b) also includes the coefficients obtained by compressive sensing. **(c)** shows the motorbike relit by three colored light sources. The right column shows a ceramic object and an apple. A reference photograph of the scene **(d)** lit by the illumination condition (shown in the inset) is compared to a computed result **(e)**. **(f)** shows the scene under three small colored light sources. Reflectance functions of the six marked pixels are shown in Figure 12.

Figure 12: **Reflectance Function Approximations.** Six selected reflectance functions from the scenes in Figure 11. Two different exposures of each reflectance function are shown to better illustrate the different details at different scales.
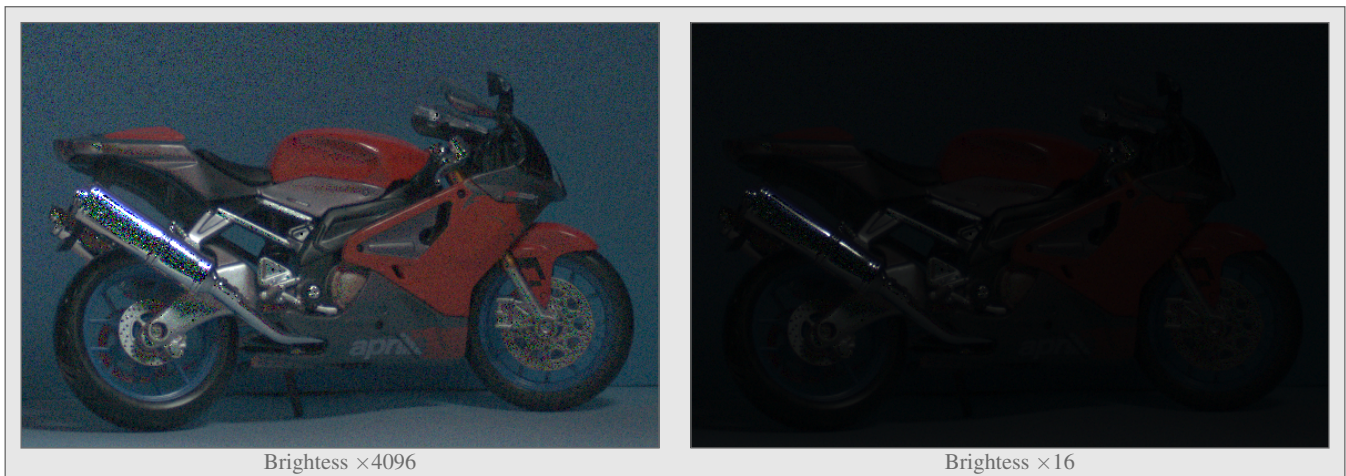

Figure 13: **Extreme Illumination Condition.** A scene containing a motor bike illuminated by a single illumination pixel (out of 128×128). The left image has been brightened 4096 times, and the right 16 times.

tions, an energy-based heuristic is used to preselect the kernel to be split. Finally, to ensure spatial coherency, the kernel configuration of neighboring pixels' reflectance functions are also verified as a potential kernel subdivision. If the quadratic programming solution is less optimal, then the neighbor's kernel configuration is used instead. Additionally, the kernel configuration is shifted around to check whether this gives a more optimal configuration. This spatial regularization is only performed at the end of the algorithm.

Peers Dutré [2005] also infer reflectance functions in a greedy fashion. Their input consists of wavelet noise patterns: normal distributed random Haar wavelet coefficients. The Haar wavelet basis used is non-standardly normalized, i.e., equal amplitude instead of the standard equal energy normalization. The reconstruction algorithm reconstructs a non-linear Haar (also non-standardly normalized) wavelet approximation of each reflectance function independently in a greedy fashion based on the heuristic that only coefficients need to be considered for which their parent is already included in the approximation. Of all potential candidate coefficients, the one that reduces the error most is included. This corresponds to the candidate wavelet coefficient with the largest estimated coefficient.

7.2 Comparison

We will break the comparison into four focus points: the illumination patterns, the representation and reconstruction of reflectance functions, spatial coherency, and flexibility of the method.

Illumination Patterns. One of the strengths, and at the same time weaknesses, of [Matusik et al. 2004] is the ability to infer reflectance fields from natural illumination. By not imposing hard constraints on the measurement ensemble, it is possible to compute reflectance fields for large outside scenes such as whole cities. However, it is also a weakness in the sense that it is unclear what the influence of the measurement ensemble is on the reconstruction, and what the limits of the system are. This makes it very hard to compare this method since no optimal illumination is known.

The method of Peers Dutré [2005], takes an opposite approach, and imposes strict rules on the generation of the wavelet noise patterns. This has the advantage that strict bounds are imposed on the performance of the algorithm, but at the cost of flexibility in selecting patterns that are optimized to the characteristics of the acquisition setup.

The presented method also puts strict conditions on which kind of illumination patterns can be used to ensure optimal inference of

reflectance fields. However, these restrictions are much less strict than those of [Peers and Dutré 2005], allowing one to optimize the patterns for the limitations of the acquisition setup. Neither of the previous methods addresses the design of illumination patterns.

Reflectance Function Representation and Reconstruction. Matusik *et al.* [2004] presented two methods for inferring the reflectance functions. However, both methods suffer from some shortcomings. First, it has been previously shown in [Peers and Dutré 2005] that the maximum energy subdivision heuristic is suboptimal. Splitting the kernel with the highest energy does not necessary yield the best error reduction. Second, without the splitting heuristic all possible subdivisions need to be tested. However, in such a case the method is only practical for 25 to 30 subdivisions, after that the method becomes computationally too expensive. The presented method, on the other hand, scales much better with respect to the desired approximation quality.

The method of Peers Dutré [2005] computes a non-linear wavelet approximation of each reflectance function that minimizes the L_2 error in the wavelet domain. However, due to the non-standard normalization, this does not correspond to an optimization of the L_2 error in the canonical domain, i.e., with respect to a reference reflectance function, but instead to a weighted L_2 norm, where the weighting places more importance on low frequency wavelets. Furthermore, only wavelet basis functions can be used to represent the reflectance functions. Both are inherent to their greedy algorithm and are necessary for correct execution. Our method, on the other hand, is able to infer reflectance functions for any basis, regardless of normalization.

Another key difference is that both previous methods only add kernels or basis functions, and are thus unable to undo previous suboptimal decisions. The presented method has the ability to remove coefficients at every level of the hierarchical algorithm, even if the basis compressive sensing reconstruction method does not support removal of coefficients.

Spatial Coherency. The spatial correction method of [Matusik et al. 2004] warrants further discussion since our method also utilizes spatial information. A key difference is that Matusik *et al.* [2004]'s method copies complete configurations, while our method is able to partially copy configurations, and build upon these. Because they copy configurations, no increase in information is attained, nor is bad information removed. Additionally, our method not only uses neighborhood information from direct neighbors, but considers multiple neighborhoods of different sizes by using a multi-resolution approach. This multi-resolution approach regularizes the system at multiple scales, yielding substantially better results. As noted in [Matusik et al. 2004], their spatial correction improves the quality when the number of photographs is small, thus when their system is ill-conditioned. In a well-conditioned case, little improvement is noted. In our system, we note improvements irrespective of the conditioning.

Flexibility. While it is obvious from the above discussion that the presented method will yield better reconstructions, it is not the only strength of the presented technique. An important advantage of our method is its flexibility. First, there is the flexibility in designing illumination patterns (as shown in Section 5). Previous methods do not offer any flexibility at all [Peers and Dutré 2005], or offer too much unconstrained flexibility [Matusik et al. 2004] such that it is hard to design optimal patterns. Second, the presented method has the flexibility of choosing any existing compressive sensing reconstruction algorithm as a basis. This allows to trade off speed versus accuracy, and allows for using future advanced reconstruction algorithms in the field of compressive sensing. Finally, due to its foundation in compressive sensing, our method has the advantage

of a well-defined and flexible theoretical framework. We believe that it is this flexibility that is the most important advantage over previous methods.

7.3 Limitations

Currently compressive light transport sensing is mostly limited by the quality of the measurements. Synthetic examples have shown that our method can accurately reconstruct sharp shadow boundaries, and smooth reflectance functions. The limited SNR of the acquisition devices, however, limits the accuracy in real scenes. Small coefficients are especially difficult to infer accurately. Fortunately, these small coefficients are the least important in a non-linear approximation. However, it does limit how many coefficients can be computed from a limited number of measurements. It should be noted that even in a brute-force acquisition (e.g., using Hadamard measurement patterns), or using current adaptive methods, these will encounter the same problems when using similar measurement devices.

An advantage of compressive light transport sensing is that it trades acquisition complexity against post-processing computation time. As a result, acquisition is straightforward. A disadvantage is that the required computation time can be quite significant. Our unoptimized code requires approximately 15 hours for the synthetic scene (512×512 resolution) shown in Figure 5. We anticipate that optimized code could reduce this significantly. Furthermore, the hierarchical algorithm can be easily parallelized to exploit the capabilities of new multi-core processors.

8 Conclusion and Future Work

We have presented a novel framework for capturing light transport based on the theory of compressive sensing which captures the light transport data of a real scene using a small set of non-adaptive illumination patterns. A new hierarchical algorithm was developed that exploits spatial coherencies to improve reconstruction quality. Additionally, practical guidelines to design illumination patterns were presented to ensure a good signal to noise ratio during acquisition of real scenes. We have applied the presented framework to an image-based relighting setup, and showed that we were able to achieve high quality acquisitions, with detailed reflectance functions, and relit results indistinguishable from reference photographs. Additionally, these results are obtained using less than two orders of magnitude of measurements compared to a brute force acquisition.

For future work, we would like to apply this framework to other applications in computer graphics. Additionally, we would also like to extend the presented technique to different setups for capturing light transport data, such as BRDFs, BTFs, and 8D reflectance fields. Further research into improving the signal to noise ratio for light transport acquisition is necessary. Additionally, we would like to refine the ensemble design guidelines to minimize the impact of camera shotnoise.

sition or the policy of the US Government, and no official endorsement should be inferred.

References

BARANIUK, R., DAVENPORT, M., DEVORE, R., AND WAKIN, M. 2008. A simple proof of the restricted isometry property for random matrices. *Constructive Approximation* (Jan.).

CANDÈS, E., AND TAO, T. 2006. Near optimal signal recovery from random projections: Universal encoding strategies? *IEEE Trans. on Information Theory 52*, 12 (Dec.), 5406–5425.

CANDÈS, E., ROMBERG, J., AND TAO, T. 2006. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics 59*, 8 (Aug.), 1207–1223.

CANDÈS, E. 2006. Compressive sampling. In *Int. Congress of Mathematics*, no. 3, 1433–1452.

CHEN, S. S., DONOHO, D. L., AND SAUNDERS, M. A. 2001. Atomic decomposition by basis pursuit. *SIAM Rev. 43*, 1, 129–159.

CHUANG, Y.-Y., ZONGKER, D. E., HINDORFF, J., CURLESS, B., SALESIN, D. H., AND SZELISKI, R. 2000. Environment matting extensions: Towards higher accuracy and real-time capture. In *Proceedings of ACM SIGGRAPH 2000*, Computer Graphics Proceedings, Annual Conference Series, 121–130.

COTTER, S. F., RAO, B. D., ENGAN, K., AND KREUTZ-DELGADO, K. 2005. Sparse solutions to linear inverse problems with multiple measurement vectors. *IEEE Transactions on Signal Processing 53*, 7 (July).

CSR. Compressive sensing resources. `http://www.dsp.ece.rice.edu/cs/`.

DANTZIG, G. 1963. *Linear Programming and Extensions*. Princeton University Press.

DEBEVEC, P., HAWKINS, T., TCHOU, C., DUIKER, H.-P., SAROKIN, W., AND SAGAR, M. 2000. Acquiring the reflectance field of a human face. In *Proceedings of ACM SIGGRAPH 2000*, Computer Graphics Proceedings, Annual Conference Series, 145–156.

DONOHO, D., AND TSAIG, Y. 2006. Extensions of compressed sensing. *Signal Processing 86*, 3 (Mar.), 533–548.

DONOHO, D. 2006. Compressed sensing. *IEEE Trans. on Information Theory 52*, 4 (Apr.), 1289–1306.

ELAD, M. 2007. Optimized projections for compressed sensing. *IEEE Transactions on Signal Processing 55*, 12 (Dec.), 5695–5702.

FUCHS, M., BLANZ, V., LENSCH, H. P., AND SEIDEL, H.-P. 2007. Adaptive sampling of reflectance fields. *ACM Transactions on Graphics 26*, 2 (June), 10:1–10:18.

GARG, G., TALVALA, E.-V., LEVOY, M., AND LENSCH, H. P. A. 2006. Symmetric photography: Exploiting data-sparseness in reflectance fields. In *Rendering Techniques 2006: 17th Eurographics Workshop on Rendering*, 251–262.

JI, S., DUNSON, D., AND CARIN, L. 2007. Multi-task compressive sensing. In *Preprint*.

JI, S., XUE, Y., AND CARIN, L. 2007. Bayesian compressive sensing. *IEEE Trans. on Signal Processing*.

LIU, X., SLOAN, P.-P., SHUM, H.-Y., AND SNYDER, J. 2004. All-frequency precomputed radiance transfer for glossy objects. In *Rendering Techniques 2004: 15th Eurographics Workshop on Rendering*, 337–344.

MASSELUS, V., PEERS, P., DUTRÉ, P., AND WILLEMS, Y. D. 2003. Relighting with 4d incident light fields. *ACM Transactions on Graphics 22*, 3 (July), 613–620.

MASSELUS, V., PEERS, P., DUTRÉ, P., AND WILLEMS, Y. D. 2004. Smooth reconstruction and compact representation of reflectance functions for image-based relighting. In *Rendering Techniques 2004: 15th Eurographics Workshop on Rendering*, 287–298.

MATUSIK, W., LOPER, M., AND PFISTER, H. 2004. Progressively-refined reflectance functions from natural illumination. In *Rendering Techniques 2004: 15th Eurographics Workshop on Rendering*, 299–308.

NEEDELL, D., AND VERSHYNIN, R. 2007. Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit. In *Preprint*.

NEEDELL, D., AND VERSHYNIN, R. 2007. Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit. In *Preprint*.

NG, R., RAMAMOORTHI, R., AND HANRAHAN, P. 2003. All-frequency shadows using non-linear wavelet lighting approximation. *ACM Transactions on Graphics 22*, 3 (July), 376–381.

PEERS, P., AND DUTRÉ, P. 2003. Wavelet environment matting. In *Eurographics Symposium on Rendering: 14th Eurographics Workshop on Rendering*, 157–166.

PEERS, P., AND DUTRÉ, P. 2005. Inferring reflectance functions from wavelet noise. In *Rendering Techniques 2005: 16th Eurographics Workshop on Rendering*, 173–182.

PEERS, P., HAWKINS, T., AND DEBEVEC, P. 2007. A reflective light stage. Tech. rep., ICT-USC, Dec. ICT-TR-04.2006.

SEN, P., CHEN, B., GARG, G., MARSCHNER, S. R., HOROWITZ, M., LEVOY, M., AND LENSCH, H. P. A. 2005. Dual photography. *ACM Transactions on Graphics 24*, 3 (Aug.), 745–755.

TROPP, J., AND GILBERT, A. 2005. Signal recovery from partial information via orthogonal matching pursuit. In *Preprint*.

TROPP, J. A., GILBERT, A. C., , AND STRAUSS, M. J. 2006. Algorithms for simultaneous sparse approximation. part i: Greedy pursuit. *Signal Processing, special issue "Sparse approximations in signal and image processing" 86* (Apr.), 572–588.

WEISS, Y., CHANG, H. S., AND FREEMAN, W. T. 2007. Learning compressed sensing. In *Allerton Conference on Communication, Control, and Computing*.

WENGER, A., GARDNER, A., TCHOU, C., UNGER, J., HAWKINS, T., AND DEBEVEC, P. 2005. Performance relighting and reflectance transformation with time-multiplexed illumination. *ACM Transactions on Graphics 24*, 3 (Aug.), 756–764.

WRIGHT, S. 1997. *Primal-Dual Interior Point Methods*. SIAM.

ZONGKER, D. E., WERNER, D. M., CURLESS, B., AND SALESIN, D. H. 1999. Environment matting and compositing. In *Proceedings of SIGGRAPH 99*, Computer Graphics Proceedings, Annual Conference Series, 205–214.