Compressive Sensing for Light Fields

Ingomar Wesp

Institute of Computer Graphics

Johannes Kepler University Linz

16th March 2011

Contents

1 Introduction

Knowing the light transport of a scene in terms of the linear contribution of incident illumination to radient light opens up a number of exciting applications in multiple fields. Once the relation between a number of emitters and a set of receptors has been determined for a given setup, synthetic images of the scene under arbitrary illumination can be generated, provided that the artificial illumination would have been expressible in terms of a configuration of the original emitters.

The potential use of measured light transport ranges from traditional areas such as image-based relighting (e.g. for the purpose of compositing foreign objects into captured scenes) to applications in microscopy, where a post-acquisition modification of the illumination might reveal details in a specimen that would otherwise have remained hidden – even if the original specimen is not available anymore. Due to the reciprocity principle, captured light transport can even be used to match incident illumination to a desired outcome, for example to compensate for undesirable conditions in a projector-observer setup. When measuring the reflectance field along angular as well as spatial dimensions, the captured light transport effectively reflects the BSSRDF of a surface, and can therefor be used to determine material properties by matching the recorded reflectance field against known material models.

Unfortunately, capturing the light transport for a given scene is known to be a time-consuming process. While there are adaptive approaches that can speed up acquisition considerably under favorable circumstances, they require real-time computation at the time of acquisition and thereby complicate the capturing setup.

The fairly young field of compressive sensing paves the way for a new and entirely non-adaptive approach that recovers an approximation of the light transport from a small set of fixed measurements. The entire recovery is performed independently of the actual acquisition in a post-processing step, effectively trading acquisition time for processing time.

This work tries to outline the fundamentals of compressed sensing as well as two slightly different ways of applying the theory to the task of capturing the light transport of a scene. In addition to comparing these two known approaches in terms of their advantages and disadvantages in real-world setups, it also attempts to show how easily they can be extended from the common 4D case to a full 8D reflectance field. It closes by conceptually comparing the compressed sensing approach to more traditional adaptive capturing schemes.

A Short Remark on Notation

Unfortunately, the notation found in the literature referenced by this work is highly inconsistent. Existing sources follow different conventions with regards to the symbols and operators they use, even when they aim to express the very same things. This work tries to follow the notation used in [25] with a few exceptions – most notably the definition of Φ, where using the convention from [25] would have caused some inconvenience.

Even though I tried to carefully introduce the symbols used in this work at the time they first

appear, I know that an overview of recurring symbols reduces the reader's burden of having to scan the text for these introductorial passages. Therefor, such an overview can be found at the end of this document.

2 Related Work

This section, which comprises the majority of this work, introduces the concept of light transport, outlines the acquisition problem and attempts to show how the theory of compressive sensing can be used in order to alleviate it.

2.1 Light Transport

2.1.1 Plenoptic Function, Light Field, Reflectance Field

In geometrical optics, given an arbitrary scene under fixed lighting conditions, the radiance leaving each point in space (x, y, z) in the direction given by two angles (tilt θ, azimuth ϕ) can be expressed as a five-dimensional function in (x, y, z, θ, ϕ) called the *plenoptic function* [1]. Since the plenoptic function covers all light moving through the scene as a result of some predetermined lighting, it provides all the data necessary to be able to reconstruct images recorded from any arbitrary viewpoint. [1]

Fixing the observer at a position outside of a closed, convex surface that encompasses all objects in the scene and assuming a totally transmissive medium allows a simplified representation as a four-dimensional function $P'(u, v, \theta, \phi)$, where the point (x, y, z) is replaced by coordinate pair (u, v) that refers to a point on the enclosing surface [13]. This function is commonly referred to as a 4D *light field*. The reduction in dimensionality stems from the knowledge that radiance traveling along a ray remains constant once it is outside of the enclosing surface, as it cannot be influenced either by occluders (which, by definition, are all encapsulated by the surface) or by the medium itself (since it is assumed to be neutral). Because light outside the surface remains unaffected, the plenoptic function P for all points outside the surface can be derived from P' by following the ray from the point (x, y, z) until it intersects with the surface at (u, v).

In computer vision, the light direction is often expressed not in terms of the angles (θ, ϕ), but by the intersection of the ray with two planes at some coordinates (u, v) in the first and (s, t) in the second plane, leading to the alternate parametrization $P''(u, v, s, t)$. When a light field is represented in this way, it is often called a "light slab".

Figure 1 attempts to illustrate the different ways to parameterize the light field and the plenoptic function.

[1] Wavelength and time are ignored in this model on purpose, as this work deals only with static scenes and trichromatic imaging. If time and wavelength are of concern, the parameterization of the plenoptic function can be easily extended to one that includes additional parameters t and λ.

In order to take altering lighting conditions into account, one can extend the function of the radiant light field into one that is parameterized for all possible incident light fields. This function, which describes the emitted light under all incident light fields, is named *reflectance field*. It is expressed in terms of an incident light field $R_i(u_i, v_i, \theta_i, \phi_i)$ and a radient light field $R_r(u_r, v_r, \theta_r, \phi_r)$, resulting in the eight-dimensional parameterization of the reflectance field R as $R(u_i, v_i, \theta_i, \phi_i, u_r, v_r, \theta_r, \phi_r)$. As noted in [13], this definition coincides with the definition of the BSSRDF [22] for the bounding surface, except that the definition of the reflectance field does not necessarily require that the surface exists as a part of the scene. This equivalence implies that the reflectance field of any arbitrarily complex scene can be just as well represented through a BSSRDF of some convex surface bounding that same scene (see figure 2).

Since many applications do not require the reflectance field in all eight dimensions, it is common to assume a finite set of fixed positions of light sources and light receptors with a

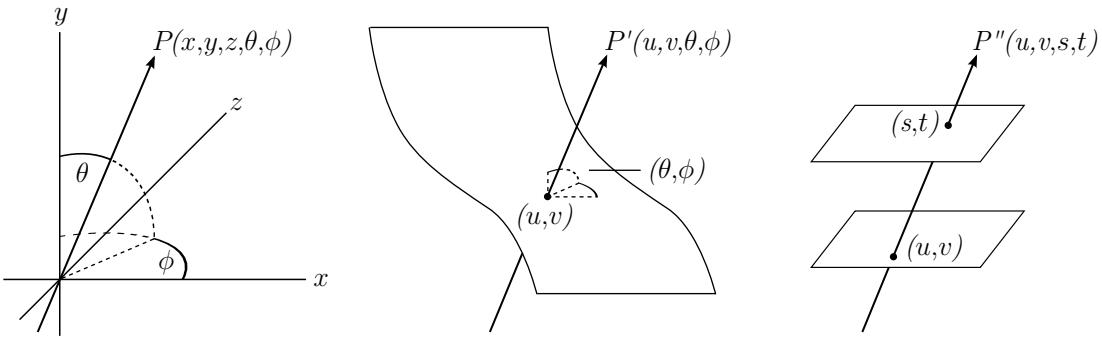

Figure 1: Left: Five-dimensional parameterization of the plenoptic function as radiance passing through the point (x, y, z) with azimuth ϕ and tilt θ. Center: Light field composed of rays intersecting surface at coordinates (u, v) and exiting at angle (θ, ϕ) into unoccluded space. Right: "Light slab" parameterization by intersecting a ray with two planes at coordinates (u, v) and (s, t) respectively.

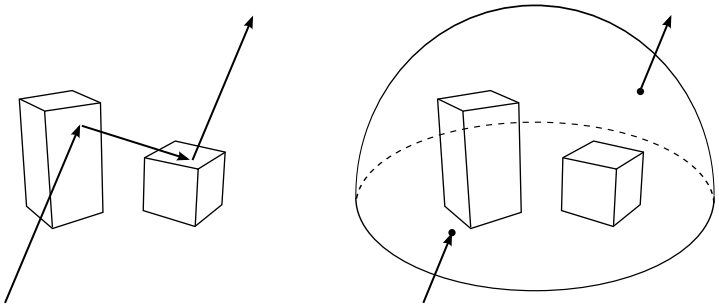

Figure 2: For an observer positioned outside of the convex surface, the reflectance field of the scene (left) is indistinguishable from a suitably chosen BSSRDF applied to the surface (right).

constrained field of view, thereby taking a discrete slices of the full reflectance field. Even under these constraints, depending on the problem at hand, a variety of different emitter-receptor setups are imaginable.

In [13], a so-called light stage is used to illuminate a subject with a spotlight that moves along a spherical hull. Camera images of the scene lit by different light stage configurations are taken and treated as if each image would correspond to a different light source. More conventionally, as in [28], a single projector is used as an array of light sources. In such a projector-camera setup, the pixels of the projector are regarded as individual light sources that are used to illuminate the scene recorded by a camera that acts as a sampling device.

2.1.2 The Light Transport Matrix

Regardless of dimensionality, any set of finite, scalar measurements can be stacked into a one-dimensional vector representation. Since light transport - the flow of light from a set of emitters to a set of receptors - is linear, the correspondence between incident and measured radiance is expressible in vector-matrix form as

$$c = Tl \tag{1}$$

where each element in the $p \times 1$ vector c corresponds to measured radiance at a light receptor and each element in the $n \times 1$ vector l represents the radiance emitted by a light source [23]. The $p \times n$ *light transport matrix* T expresses the linear mapping between the two, describing how much light flows from every emitter through the scene to every receptor. Every row T_i in T weights the contributions of each individual emitter towards the result in c_i. Since the rows of T linearly combine sampled incident radiance into samples of measured radiance, they are – in allusion to the definition of the reflectance field – also referred to as *reflectance functions*.

The relation above can be easily extended to express a set of different lighting conditions and their corresponding responses by replacing l with an $n \times m$ matrix L, in which each column vector L_j represents a single pattern projected into the scene (see figure 3):

$$C = TL \tag{2}$$

The $p \times m$ matrix C resulting from multiplying T by L contains m column vectors C_j, each of which records the response for the corresponding pattern L_j.

In other words, knowing the light transport matrix T for one particular setup of light sources and emitters allows one to reconstruct receptor responses for any lighting condition that can be produced by the light sources that correspond to the individual row indices of L. Depending on how the scene was set up, T can therefor be used to represent a discrete sampling of the scene's reflectance field or slices thereof.

In order to allow for a less general treatment of T, the next few sections of this work will focus on the 4D projector-camera setup, as it is both easy to illustrate and simple to extend into a sampling of a higher dimensional function when necessary. In this setup, C_j can be understood to be a stacked 2D camera image created by illuminating the scene with the projected 2D pattern stacked in L_j.

2.1.3 Acquisition of Light Transport Data

The problem of acquiring the reflectance function T_i for the ith receptor can be reduced to a general sampling problem, in which we project a number of patterns into the scene and record the corresponding responses. By considering T_i to be some unknown signal, we can recover it by taking a set of patterns L and recording the response vector C_i, which contains the responses of receptor i to each pattern L_j:

$$C_i = T_i L$$

Transposing both sides yields the equation system in $Ax = b$ form:

$$(C_i)^T = (T_i L)^T = L^T (T_i)^T \tag{3}$$

Finding a solution for T_i, however, requires that L^T be of rank n or, in other words, that the number of measurements taken with linearly independent patterns equals at least the number of different light sources. Even though receptor responses can usually be captured in parallel (e.g. in a single camera image), acquiring the light transport for a today's projector resolutions would require more than 10^6 separate images to be taken – a situation that is undesirable if acquisition time is of concern [2].

Based on the observation that in many scenes, some emitters do not conflict in terms of the receptors they influence, adaptive capturing schemes have been developed that attempt to

[2] In this case, storage would most likely also be an issue, but one that could be alleviated at the cost of exactness by performing on-the-fly image compression.

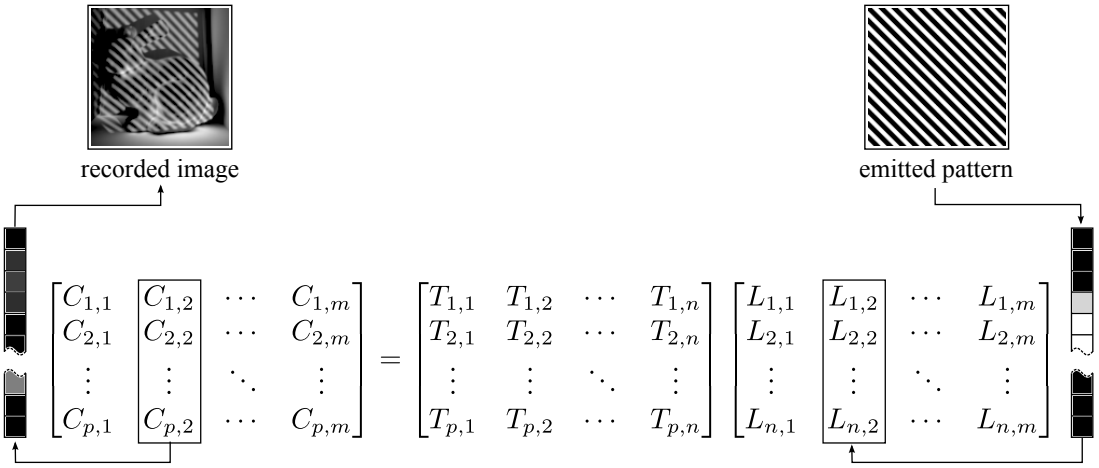

Figure 3: Illustration of the role of the light transport matrix in a 4D projector-camera setup. 2D patterns are stacked into individual column vectors of L. After multiplying with the light transport matrix, each column of the resulting matrix C contains a stacked image of the scene relit with the pattern from the corresponding column in L.

capture multiple reflectance function coefficients at once [23, 27].

Methods like the one presented in [27] work very well for scenes that are not dominated by global illumination effects. Conditions with a lot of low frequency contribution affecting all receptors, however, impede their ability to separate the images into independent partitions, which in the worst case degrades their performance to the level of a brute-force approach, where each emitter's contribution is sampled independently. An additional drawback is the necessity that the captured images need to be processed on the fly in order to produce future patterns, which increases setup effort and can slow down acquisition depending on the available computing power and the complexity of the method.

The theory of compressive sensing, which will be introduced in section 2.2, alleviates this issue by allowing for the reconstruction of approximations of reflectance functions from a number of *non-adaptive* measurements that is still much lower than the number of individual light sources. It thus combines the benefits of adaptive approaches (fewer measurements) with the advantage of not having to perform complex computations while the acquisition phase is in progress.

2.1.4 Dual Images

Due to the reciprocity principle originally formulated for specular reflection by Helmholtz [17], extended to diffuse reflection by Rayleigh [26] and generalized to physically valid BSDFs in [32], we know that the paths taken by light through the scene remain the same when reversing light direction. This assertion allows for an interesting observation with regards to the individual reflectance functions: Due to the symmetry of light flow, every reflectance function T_i not only describes the influence of all light sources on receptor i, but it also allows us to reverse the roles of receptor and emitters; T_i can be alternatively seen as weighing the contribution of emitted light from an hypothetical emitter (that has the same positional and directional properties as the receptor i) towards hypothetical receptors (again, with the same properties as the previous emitters). In the 4D case, T_i is effectively an image of the scene taken from the position of the projector and illuminated only by a single camera pixel i. Such a computed image is called a *dual image* as opposed to the *primal image*s resulting from equation 2.

This insight implies that the transposed light transport matrix enables us to express all possible dual images resulting from "illuminating" the scene with the primal receptors, an idea extensively explored in [27]. The dual light transport from emitters to receptors can thus be formulated as

$$L' = T^T C'$$

Note, however, that in general T is not orthogonal (TT^T is not necessarily the identity matrix). In fact, when sampling real world scenes or physically accurate simulations, it is mandatory that $TT^T \neq I$, because some of the emitted light is "lost" by being absorbed or scattered onto a path that never reaches a receptor.

In the following section, the concept of dual images will help us not only to provide an intuitively graspable visualization of reflectance functions and their reconstructions, but it will also serve as an underpinning for finding a suitable basis for compressing reflectance functions.

2.2 Compressive Sensing

The theory of compressive sensing – as applicable to this work – deals with the reconstruction of sparse signals from a fraction of the number of linear measurements that would be necessary to reconstruct arbitrary signals of the same length. In order to understand how this can be accomplished, it is necessary to introduce some underlying terminology first.

2.2.1 Sparsity and Compressibility

An arbitrary discrete signal in n dimensions, whose scalar values are stored in the vector $x \in \mathbb{R}^n$ is said to be k-sparse if and only if the number of non-zero elements of x is smaller or equal to k, that is if and only if the cardinality of its support ($|supp(x)| = \|x\|_0) \leq k$.

Signals that are not necessarily k-sparse can still be k-compressible if the error produced by reducing them to their k components of highest absolute value is acceptably small. More formally, one can say that x is compressible if its k-term approximation z, created from x by preserving only the k components of highest absolute magnitude, yields an approximation error $\|x - z\|_X$ (in some norm $\|\cdot\|_X$) that decreases rapidly with increasing k. This is the case for signals which contain only a few big, but many small components, such as signals, whose component magnitudes decay by a power law. As the magnitude of their elements – ordered by decreasing absolute value – decreases rapidly, they can be recovered from just a few significant components with little error.

Since many signals are compressible only in certain bases, it is practical to define a signal x as "k-compressible in some basis B" if only the result $\hat{x}$ of transforming x into the the transform domain B ($\hat{x} = Bx$) is k-compressible.

Finding such bases for different kinds of signals has been the subject of significant research in the past decades, as transforming a signal into a basis where it can be faithfully approximated by a sparse representation is fundamental for signal processing tasks such as lossy compression. Aside from sinusoids, different types of wavelet bases are probably among the most commonly used decompositions used to this end.

While the actual compression, where the signal is transformed into a sparse representation in the domain B, is traditionally performed as a post processing step applied to a full, uncompressed measurement, compressive sensing aims at shifting this transformation into the measurement process itself.

2.2.2 Reconstruction of Sparse Signals through Basis Pursuit

A fundamental question addressed by the research in compressed sensing reads as follows: Assume $x \in \mathbb{R}^n$ to be a k-sparse signal, with $k \ll n$, and define a set of *measurements* y of the signal as the matrix product between some known $m \times n$ *measurement ensemble* Φ and x:

$$y = \Phi x \tag{4}$$

How many measurements m (rows of Φ) do we need in order to faithfully reconstruct x? Conventionally, we would assume that the number of measurements would need to be at least n ($m = n$ suffices to uniquely solve equation 4 for x if Φ is nonsingular). Excitingly, it is shown in [6], that if x is indeed k-sparse and Φ fulfills certain conditions, x can be recovered without error from just $\mathcal{O}(k)$ measurements by solving the ℓ_0 minimization

$$min\|\tilde{x}\|_0 \text{ , such that } y = \Phi\tilde{x}$$

Unfortunately, even if we know how to construct a suitable measurement ensemble, performing the ℓ_0 minimization poses a combinatorial problem of exponential complexity, so computing this solution is not tractable except in the most trivial cases [3]. While this result itself is impractical, is shows that sufficient information about x is retained in y to make such a reconstruction possible. Maybe there are other solution strategies that might recover x at lower computational cost, if only under additional preconditions.

It was first shown for discrete Fourier coefficients [6] and later refined as well as generalized to coefficients in other transforms [7, 8], that indeed if certain properties of Φ hold, the ℓ_1 minimization

$$min\|\tilde{x}\|_1 = min\sum_{i=1}^{n}|\tilde{x}_i| \text{ , such that } y = \Phi\tilde{x}$$

will also recover the sparsest $\tilde{x}$ that solves $y = \Phi\tilde{x}$ from just $\mathcal{O}(k \cdot log(n))$ measurements [12]. In other words, for certain Φ and a sufficient amount $m \ll n$ of measurements, it is possible to replace the ℓ_0 minimization by an ℓ_1 minimization - also called *basis pursuit* - without affecting the outcome. Unlike ℓ_0 minimization, for which no algorithms of polynomial time complexity are known [4], basis pursuit can be mapped to linear programming, and thus can feasibly solved by algorithms such as the interior point method or the simplex algorithm [11].

Up until now, outlining the exact conditions imposed on the nature of Φ has been carefully avoided, just as we didn't bother to give an exact lower bound for m in terms of sparsity and signal length.

Finding necessary and sufficient conditions for the suitability of a measurement ensemble for the task of sparse reconstruction is a subject of ongoing research. As summarized in [5], the idea of sparse recovery hinges on the fact that the columns of the measurement ensemble Φ need to be incoherent with the basis in which the signal is sparse (the canonical basis, if the signal is sparse without transform), as this essentially ensures that every non-zero signal component has

[3] Although the exponential nature of ℓ_0 minimization might be intuitively obvious, [7] contains a mapping of ℓ_0 minimization to the subset-sum problem and thereby proofs its NP-completeness.

[4] In fact, if such an algorithm were to be found, it would be a solution to the P = NP? problem in favor of P = NP.

an effect on many individual measurements – or in other words, that for any sparse signal x, the measurement vector y will *not* be sparse.

For the past few years, a significant body of research was based upon measurement ensembles that fulfill a condition called the *restricted isometry property* (RIP). Introduced in [7], in not only enforces incoherence (if the number of rows m is small), but ensembles obeying the RIP can also be shown to be robust against measurement errors and to allow recovery of good k-term approximations for signals that are not exactly k-sparse, but compressible [10].

The restricted isometry property with parameters (s, ϵ) holds for a measurement ensemble Φ if and only if the following holds for every s-sparse vector v and some ϵ smaller than 1:

$$(1 - \epsilon)\|v\|_2 \leq \|\Phi v\|_2 \leq (1 + \epsilon)\|v\|_2 \tag{5}$$

This property requires that any set of columns from Φ with size $\leq s$ forms an approximate orthonormal set. It can be used to find deterministic bounds on the maximum size of ϵ for exact recovery to occur. Particularly, in [3], where the implications of the RIP are discussed in detail, it is shown that exact recovery of every k-sparse vector is guaranteed to occur under ℓ_1 minimization if the RIP holds for sparsity $s = 2k$ and $\epsilon \leq \sqrt{2} - 1$.

As a sidenote, the RIP has been shown to be a sufficient condition for recovery, but it also suffers from a number of problems: Not only is checking whether a given matrix fulfills the RIP an NP-hard problem, but the RIP is also known to be too strict (that is, it is a sufficient condition, but not a necessary one), which is why effort is spent on defining properties that are more pleasing in this regard. Among multiple proposals by other authors, there is one by Candes and Plan [9] that builds a probabilistic theory of compressed sensing which is not based upon the RIP.

Regardless of the ongoing work to find better properties, some random matrices – e.g. ones drawn from a Gaussian or Bernoulli distribution – have been shown to obey the RIP with high probability and are therefor suited for compressed sensing (see [8] for proofs).

2.2.3 Reconstruction via Greedy Approximation

While basis pursuit indeed offers a tractable solution the reconstruction problem, the computational cost is still significant. In 2007, Gilbert and Tropp [31] suggested using an already known sparse approximation algorithm called *Orthogonal Matching Pursuit* (OMP) in order to recover the signal more efficiently. OMP is a greedy algorithm that recovers a single component of x with each iteration resulting in an overall complexity of $\mathcal{O}(k \cdot m \cdot n)$, which means it scales considerably better than any known ℓ_1 minimization.

In order to find a sparse solution for $y = \Phi x$, OMP takes the measurement ensemble Φ, the measurement vector y, and a target sparsity $\tilde{k}$ as input. If x is indeed $\tilde{k}$-sparse (i.e. $k = \tilde{k}$), then y is just a linear combination of $\tilde{k}$ columns from Φ, so OMP attempts to select columns from Φ which appear to be participating in y. It does this by iteratively looking for the column Φ_j that correlates most strongly with the measurement residual r (initialized with $r \leftarrow y$ for the first iteration). This is the column at index j, for which the inner product with the measurement

residual is maximal:

$$j = \arg\max_{\hat{j}} |\Phi_{\hat{j}} r^T|$$

By concatenating all previously selected columns into $\tilde{\Phi}$, an approximation $\tilde{x}$ for the signal vector in terms of $\tilde{\Phi}$ is found by solving the least squares problem

$$\tilde{x} = \arg\min_{\hat{x}} \|y - \tilde{\Phi}\hat{x}\|_2$$

From this approximation, a new residual r can be computed by subtracting the portion of the measurements explained by $\tilde{x}$ from the actual measurements y. For the next iteration, the steps above are repeated with the newly acquired residual r. After $\tilde{k}$ iterations, OMP terminates and returns an approximation for x constructed from $\tilde{x}$ by populating an empty vector with the $\tilde{k}$ components of $\tilde{x}$ at the indices found for j in each iteration.

Unfortunately, while it can be shown that OMP can solve the recovery problem for some random ensembles, the general guarantees for successful recovery are significantly weaker than the ones provided by basis pursuit. In particular, while OMP can recover measurements taken with subgaussian ensembles, it is not yet known whether it is suited for all ensembles that fulfill the RIP. It is also known that OMP requires somewhat more measurements than ℓ_1 minimization in order to allow exact recovery.

Algorithm 1: Orthogonal Matching Pursuit (OMP)

Input: Measurements $y \in \mathbb{R}^m$, measurement ensemble $\Phi \in \mathbb{R}^{m \times n}$, target sparsity $\tilde{k}$
Output: Approximation a with sparsity $\tilde{k}$
begin
  initialize residual $r \leftarrow y$;
  initialize index list $I \leftarrow ()$;
  initialize chosen columns $\tilde{\Phi} \leftarrow []$;
  for $t \leftarrow 1$ *to* $\tilde{k}$ **do**
    find biggest correlation $j \leftarrow \arg\max_{\hat{j}} |\Phi_{\hat{j}} r^T|$;
    add j to index list $I \leftarrow I$ concat (j) ;
    set matrix of chosen columns $\tilde{\Phi} \leftarrow \Phi_I$;
    find approximation $\tilde{x} \leftarrow \arg\min_{\hat{x}} \|y - \tilde{\Phi}\hat{x}\|_2$;
    compute new residual $r \leftarrow y - \tilde{\Phi}\tilde{x}$;
  end
  initialize extension from $\tilde{x}$ to n dimensions as $a \leftarrow 0^n$;
  $t \leftarrow 1$;
  for $i \in I$ **do**
    set $a_i \leftarrow \tilde{x}_t$;
    increment $t \leftarrow t + 1$;
  end
  return a
end

An extension of OMP called *Regularized Orthogonal Matching Pursuit* (ROMP) addresses some of these issues [21]. Instead of restoring a single signal component per iteration, ROMP recovers an entire set, determined by looking at the observation vector $u = \Phi^T y = \Phi^T \Phi x$. The idea is that since every $\tilde{k}$ columns of Φ are required by the RIP to form an approximate orthonormal set, u should be a reasonable local approximation of x. Thus, in every iteration, ROMP first finds the largest uniform subset of the biggest components in u and chooses the indices of this subset as indices of x to be recovered. Then, as in OMP a least squares problem is solved to get an approximation for x and a new residual is determined, from which in turn a new observation vector u for the next iteration is computed.

Algorithm 2: Regularized Orthogonal Matching Pursuit (ROMP)

Input: Measurements $y \in \mathbb{R}^m$, measurement ensemble $\Phi \in \mathbb{R}^{m \times n}$, target sparsity $\tilde{k}$
Output: Approximation a
begin
  initialize residual $r \leftarrow y$;
  initialize index list $I \leftarrow ()$;
  initialize chosen columns $\tilde{\Phi} \leftarrow []$;
  while $\|r\|_2 > 0$ **do**
    set observation vector $u \leftarrow \Phi^T y$;
    if $\|u\|_0 > \tilde{k}$ **then**
      let J be the set of the $\tilde{k}$ indices in u with the highest absolute value ;
    end
    else
      $J \leftarrow \{i | u_i \neq 0\}$;
    end
    from all subsets of $J_k \subset J$, chose the $J_0 \subset J$ with the maximum $\|u_{J_0}\|_2$, that satisfies $|u_i| < 2 \cdot |u_j|$ for all $i, j \in J_0$;
    add indices in J_0 to index list $I \leftarrow I$ concat J_0 ;
    set matrix of chosen columns $\tilde{\Phi} \leftarrow \Phi_I$;
    find approximation $\tilde{x} \leftarrow \arg\min_{\hat{x}} \|y - \tilde{\Phi}\hat{x}\|_2$;
    compute new residual $r \leftarrow y - \tilde{\Phi}\tilde{x}$;
  end
  initialize extension from $\tilde{x}$ to n dimensions as $a \leftarrow 0^n$;
  $t \leftarrow 1$;
  for $i \in I$ **do**
    set $a_i \leftarrow \tilde{x}_t$;
    increment $t \leftarrow t + 1$;
  end
  return a
end

For measurement ensembles obeying the RIP, ROMP effectively provides similar guarantees as basis pursuit, albeit at only a computational cost comparable to the one of OMP.

2.3 Applying Compressive Sensing to Light Transport Acquisition

Regardless of the reconstruction approach, in order to apply compressed sensing to the sampling of light transport, we have yet to map the acquisition problem from section 2.1.3 to the compressed sensing framework just introduced. Remembering equation 3, we know that a reflectance function T_i (the ith row of the light transport matrix T) can be inferred by solving the linear system

$$(C_i)^T = L^T (T_i)^T$$

It is quite obvious that this can be mapped to a compressed sensing problem

$$y = \Phi x$$

with $y = (C_i)^T$, $x = (T_i)^T$ and $\Phi = L^T$ [28, 25]. So in order to recover the light transport matrix T, one simply needs to subject the scene to illumination conditions that obey the RIP, stack them into individual rows of Φ and record the response vector C_i for every receptor i.

The approach outlined above, however, only works if the reflectance functions are sparse or compressible in the canonical basis, which would require that each receptor is only influenced by just a few light sources. In this case, the dual image would be almost completely black with just a few high frequency spikes. While conditions like these occur in simulated scenes that consider direct illumination only, most real-life scenes that involve some amount of diffuse materials will show global illumination effects such as a low-frequency contribution from the majority of light sources due to diffuse interreflection.

2.3.1 Exploiting Compressibility in a Transform Domain

Fortunately, the reconstruction process can be adopted so that instead of recovering the reflectance functions in the canonical basis, it finds a sparse approximation in a chosen transform domain. The assumption that the reflectance functions are compressible seems reasonable at least in the 4D case, where – as we know from section 2.1.4 – they represent the dual images lit by virtual light sources corresponding to the receptors used for capturing the scene. Since the reflectance functions are nothing but images of a real-world scene, any compression basis that is suitable for image compression should allow for sparse approximation. See section 4 for an extension of this assumption to 8D reflectance fields.

That still leaves the question of how to incorporate the basis transform into the compressed sensing framework. As we will see, previous works on this topic have found different approaches for tackling this problem, each with their own benefits and disadvantages.

The approach taken in [25] is based on the observation that one can recover the reflectance functions expressed in the orthonormal basis B ($B^T = B^{-1}$) by recording images created from projections of the measurement ensemble into the inverse transform. Since BB^T is the identity, we can expand on equation 2 as follows:

$$C = TL = TBB^T L$$

By incorporating the measurement ensemble as $L = B\Phi^T$ and defining the matrix of compressed reflectance functions $\hat{T}$ as $\hat{T} = TB$ we get

$$\begin{aligned} C &= TBB^T L \\ &= \hat{T}B^T L \\ &= \hat{T}B^T B\Phi^T \\ &= \hat{T}\Phi^T \end{aligned}$$

Transposing both sides yields

$$C^T = (\hat{T}\Phi^T)^T = \Phi\hat{T}^T$$

which for each row i in C and $\hat{T}$ means

$$(C_i)^T = \Phi(\hat{T}_i)^T \tag{6}$$

This maps exactly to the outline of a set of compressive sensing measurements as shown in equation 4. Choosing the original patterns L in a way such that they are suitable for compressive sensing would thus allow us to reconstruct the compressed reflectance functions $\hat{T}$ from the measurement ensemble $\Phi = L^T B$, which is the inverse transform of the individual patterns in L with regards to the basis B.

But how do we know that our newly gained Φ is even suitable for compressed sensing? Fortunately, as outlined in [30], certain random ensembles, such as ones drawn from a Gaussian or Bernoulli distribution have a very high probability of being incoherent with any fixed basis and should therefor be suitable for sparse recovery regardless of the domain they are applied to. Since there is no known deterministic way of creating an ensemble that obeys the RIP anyways, restricting ourselves to this kind of random matrices does not impose an actual limitation. In practical terms, we would create a set of random patterns, perform an inverse transform and then subject our scene to them. By applying an inverse transform on the images gathered this way, we can reconstruct the canonical reflectance functions in T.

An alternative way to pose the problem was shown in [28]. There, the measurement ensemble Φ is defined as $\Phi = L^T B^T$, which is the forward transform of patterns into the basis B. Unlike in [25], however, the scene is only illuminated by the patterns in L, but recovered against Φ. By supplying the reconstruction algorithm with the transformed patterns combined with the results of the subjecting the scene to the canonical ones, this approach basically shifts the basis transform into the scene. Since the recovery algorithm will attempt to explain the recorded responses in terms of the sparsest solution to equation 6, the result should still be the same as with the previous approach: The reconstruction algorithm will again recover $(\hat{T}_i)^T$, the reflectance function represented in the basis B.

The latter approach has considerable practical advantages, because it does not rely on emitting transformed coefficients. For some measurement ensembles – like the Bernoulli ensemble – this removes the necessity for calibrating the emitter equipment, as only two levels need to emitted.

2.3.2 Exploiting Interpixel Coherence in the Primal Images

Both methods outlined above can be used to reconstruct a complete light transport matrix by simply applying the reconstruction algorithm to every single reflectance function separately.

As noted in [25], it is reasonable to assume that not only the reflectance functions themselves, but also the primal images are compressible. Peers et al. demonstrated a hierarchical extension of their reconstruction approach that makes use of this additional compressibility between individual columns of the transport matrix. Essentially, instead of recovering every reflectance function on its own, their algorithm first reconstructs a global reflectance function and then refines it iteratively by recovering only the differences to the next resolution level until the resolution equals the number of pixels in the primal image. This approach is inherently similar to Haar compression: It describes the final reflectance function in terms of a hierarchical set of multiresolution averages that are constructed by applying differences to their respective parent levels.

Unfortunately, the method in the form it was outlined in [25] requires special segregated measurement patterns that restrict the compression basis for both, the rows and the columns, to a Haar basis. This is in contrast to the method from [28], which is applicable to more efficient compression schemes, as long as they can be expressed by a basis transform.

3 Method

In order to test the viability of applying compressed sensing to light transport acquisition, a set of test scenes was modeled in the 3D graphics suite *Blender* [15]. Since the internal renderer provided by Blender uses a local illumination model only, the scenes were exported and rendered in *LuxRender*, a physics based, unbiased renderer [19].

3.1 Acquisition Setup

With the scenes ready to be rendered, a 4D reference light transport matrix for fixed projector and camera positions was acquired by recording linear-response HDR images of the scenes consecutively illuminated by every single projector pixel at maximum intensity. Please note that in real-life scenarios, this approach is suboptimal, since the dimly lit images would suffer from a considerable amount of amplifier noise. Instead, as in [25], the reflectance functions can be measured in full using Hadamard patterns, which should allow for a better signal-to-noise ratio (SNR). Since images generated by LuxRender, however, do not suffer from thermal sensor noise, this way of capturing light transport was chosen for its advantage of providing partial light

transport matrices while the rendering process was still in progress.

At the time of writing, LuxRender employs a path-based model for light transport and provides different samplers and surface integrators to that end. In the configuration used for rendering the reference images, a sampler implementing *Metropolis Light Transport* (MLT) [33] was used together with a surface integrator that implements bidirectional path tracing. Due to restrictions both in memory and processing power, all renderings were performed at (or reduced to) a resolution $p \leq 256 \times 256$ at a reflectance function resolution of $n \leq 128 \times 128$. When storing the coefficients as single precision floating point values, setting p to 256^2 and n to 128^2 causes the light transport matrix T to occupy 2^{32} bytes ($= 4$ GiB) of memory. Since light transport was acquired for the three color channels in the RGB model, the total amount of gathered data even for such a modest resolution requires 12 GiB in storage. Note that, while the sparse reconstructions can be stored as a sparse matrix (and thus require considerably less memory both in RAM and on disk), the acquired reference images cannot be represented sparsely without data loss.

3.2 Reconstructing a Single, Global Reflectance Function

In order to get an estimate of the achievable quality of reconstruction, a single global reflectance function for a single receptor was recovered using the second method described in section 2.3.1. This approach is similar in principle to the "single pixel imaging" experiments described in [14] in that it recovers the dual image from a set of measurements recorded with only a single receptor. In the setup used here, the receptor response, however, was not acquired directly, but computed by downscaling the images that were already captured at higher resolutions to a single pixel.

The reconstruction algorithm used was a slightly modified version of the ROMP reference implementation provided at [20]. Besides adopting Needell's MATLAB code so that it runs also in GNU Octave, the number of coefficients returned by ROMP was also limited to the same value that is used as target sparsity.

Since emitter calibration is not an issue in simulated scenes, there is no reason not to chose measurement ensembles that contain intermediate values. In order to show the effect of different random ensembles, the scene was subjected to both, Gaussian and Bernoulli ensembles. Results for different measurement counts and target sparsities are shown side-by-side with the reference dual images at comparable compression levels in figure 4.

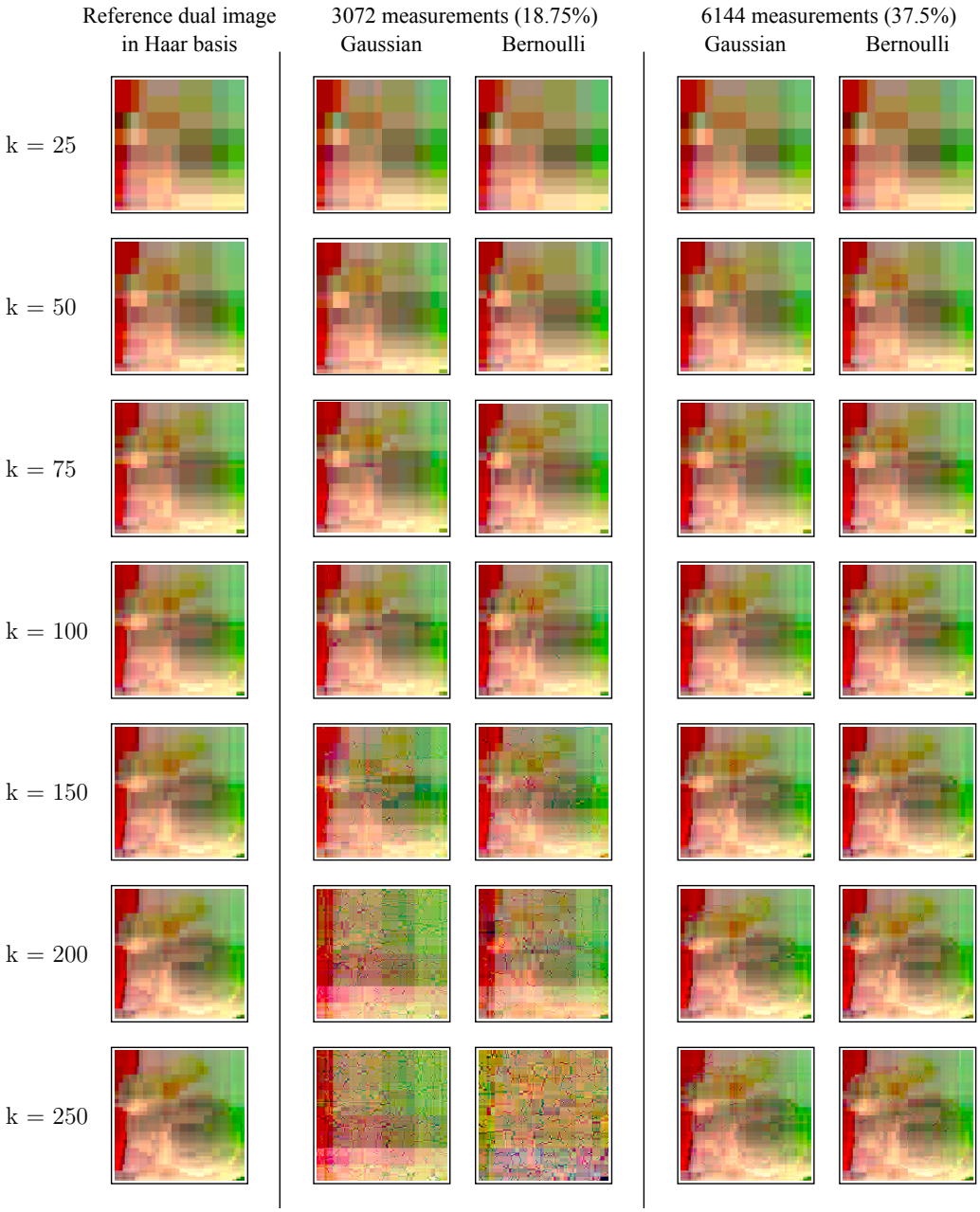

Figure 4: Reconstructions of sparse approximations vs. the reference dual image compressed at the same sparsity level. As can be clearly seen in the middle columns, setting k to a value that is too high for the available number of measurements will cause ROMP to pick up incorrect coefficients that show up as noise in the reconstruction result.

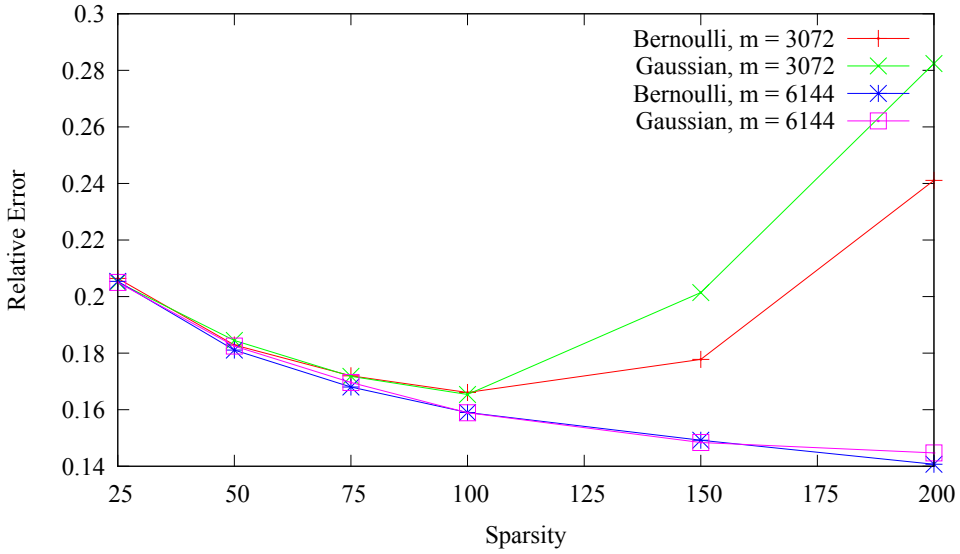

Figure 5: Average relative luminance error of the dual image reconstructions shown in figure 4 against the uncompressed reference reflectance function in the canonical domain. Errors are computed as $\epsilon = \|\tilde{x} - x\|_2 \cdot \|x\|_2^{-1}$. Again, the effect of capturing not enough images for a given sparsity is clearly visible as the error increases rapidly for $m = 3072$ and $k > 100$.

3.3 Reconstructing 4D Light Transport

For a full reconstruction of a set of 128×128 reflectance functions, each of which are again of the resolution 128×128, the approach from [28] was chosen together with a Haar compression basis. The main reason for this choice is not just the relative ease of implementation, but also the fact that the gathered measurements can be easily re-used for reconstruction in different transform domains. Unlike with the methods presented in [25], if different compression schemes are to be investigated, it is not necessary to acquire a new set of images, but reconstruction in a different compression basis can be perfomed with the images that have already been taken. This flexibility has considerable advantages for comparative research, since it avoids changes in noise levels or other capturing artifacts that can not easily be reproduced.

Furthermore, the Bernoulli pattern illumination scheme can be applied as-is to higher dimensional reflectance functions, such as the ones required to express incident light fields, whereas the approach from [25] would require novel pattern designs.

As in the single pixel camera experiments, the entire set of reflectance functions was reconstructed separately for the three linear-coded RGB channels.

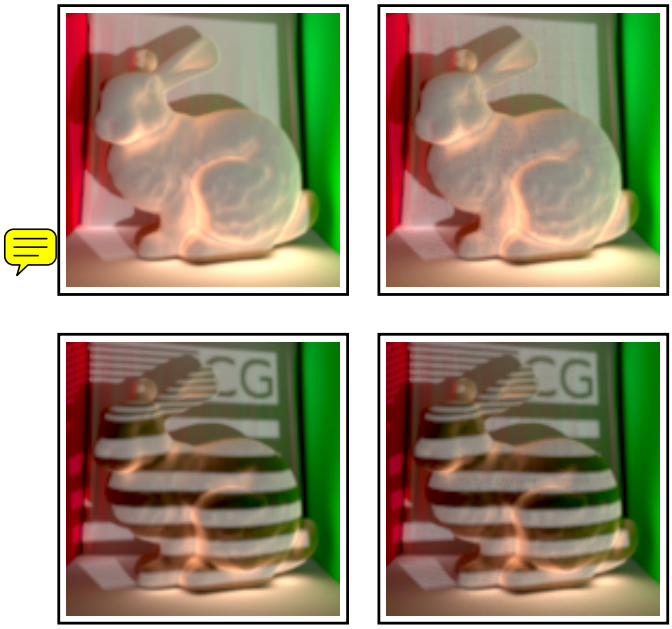

Figure 6: Left: Reference images of the scene relit at full intensity (top) and with a custom pattern (bottom). Right: The result of applying the same illumination to a light transport matrix that has been reconstructed from 3072 Bernoulli pattern measurements ($3/16$ of the number of measurements that would be required for a complete sampling of the reflectance functions) and a sparsity of 100. At close inspection, some high-frequency noise can be observed, but the overall result is quite faithful.

4 Discussion

As can be seen from the example reconstructions performed within the scope of this work, compressive sensing can be used to faithfully approximate light transport from just fraction of the measurements that would be required with conventional brute-force sampling. Since the number of required measurements scales linearly with the sparsity and just logarithmically with the number of emitters, this ratio should improve considerably with higher resolution reflectance functions.

Furthermore, it is worth pointing out that the approach used here (and even the same acquired data) can be applied as-is to other transform domains that are better suited for image compression, which should provide for even better results at comparable sparsity levels.

4.1 Practical Limitations

In contrast to the computer generated images used in this work, real-life capturing scenarios impose a number of problems that do not occur when dealing with simulated scenes.

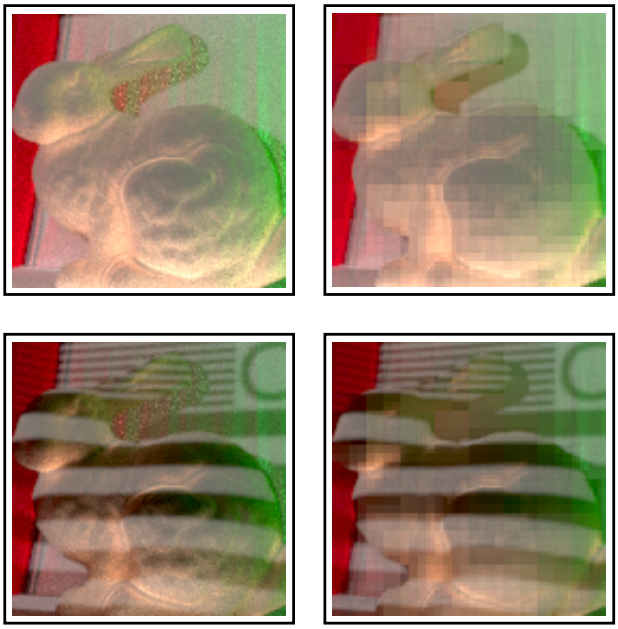

Figure 7: Under the same pretext as in figure 6, this figure shows the reference dual image and its reconstruction. Unlike in the primal images, which are not affected as strongly by the errors introduced by compressing the reflectance functions used for relighting them, compression artifacts are clearly visible in the dual image approximations (please note that the visible noise in the reference images to the left is caused by the renderer and should thus be considered part of the ground truth).

4.1.1 Limited Dynamic Range

Unfortunately, the dynamic range of common camera sensors is usually too low to capture both directly and indirectly lit portions of the scene with sufficient detail. The common approach to compensate for that is to repeat every single measurement multiple times with different exposure times (exposure bracketing). The individual images for a single illumination condition can then be combined into a linear-response HDR image. If the receptor parameters are known, this does not pose a problem functionally, but the need for multiple shots per measurement obviously increases acquisition time.

4.1.2 Emitter Calibration and Quantization Errors

If the measurement ensemble contains intermediate values that fall in the range between the minimum and maximum emitter output, the emitters need to be calibrated for linearity. But even when the output is strictly linear, it is usually limited to a finite number of different intensity steps. This limitation causes errors due to quantization (mapping intermediate values to the most suitable output intensity).

While Peers and Dutre [24] have suggested an alternative wavelet normalization scheme that maps all levels in the emitted wavelet patterns to an equal proportion of the emmiters' dynamic range, quantization errors and the need for emitter calibration can be avoided altogether when using measurement ensembles that draw only from two values, which can easily be mapped to the minimum and maximum output. Bernoulli patterns or the segregated measurement ensembles from [25] are two examples for such measurement schemes.

4.1.3 Projecting Negative Values

Random measurement ensembles such as the ones drawn from a Gaussian or a Bernoulli distribution contain negative values, which can obviously not be used directly in a typical acquisition setup. One common approach to simulate the emission of "negative light" is to take two separate images, one illuminated by the positive range of the pattern and one for the (sign-inverted) negative one. These two images can then be combined into a single image just by subtracting the latter from the former.

Similarly to the need for exposure bracketing, this effectively increases the number of images that need to be taken by a constant factor.

4.1.4 Measurement Noise

Since real-life measurements always show some degree of measurement noise due to external influences such as thermal effects or photon counting variations [2], it is important to consider how prone the reconstruction process is to these kinds of errors.

Fortunately, it has been shown in [4], that the reconstruction remains in the same order of magnitude as the measurement error plus the error introduced due to approximation. Measurement noise is thus not artificially amplified by the reconstruction process.

4.2 Extension to 8D Light Transport

Because the application of compressive sensing to light transport acquisition makes no assumptions about the nature (or dimensionality) of the unknown signal, it can be easily extended to higher dimensional reflectance functions as long as the compressibility hypothesis holds. Since light fields are expressible as four-dimensional functions with two dimensions each in the spatial and angular domain, it would stand to reason that good sparse approximations can be found in a 4D wavelet decomposition. This approach seems particularly obvious, as it comes as a natural extension of the 2D wavelet transform used in the usual projector-camera setup.

One example of how to construct a refinement scheme that generates a suitable 4D basis for light fields from 1D wavelet transforms is outlined in [18]: Based on the observation that coherence *within* the spatial and angular dimensions is generally stronger than the coherence between these two domains, a decomposition scheme is proposed that constructs the 4D wavelet basis by separately applying a nonstandard refinement to both dimensions and then combining them

through a standard refinement[5].

Due to the general nature of the reconstruction approach from [28], this only changes the basis transform used in the reconstruction step; the Bernoulli pattern measurement ensemble can be used without modification for light field illumination.

4.3 Comparison with Adaptive Approaches

When comparing the results achievable through compressed sensing with adaptive approaches like the ones in [27] or [16], it is important to note that they are somewhat complementary, because typically used compression bases perform well in situations where these approaches struggle (and vice versa). In particular, the strategy of avoiding conflicts between emitters cannot efficiently capture global illumination effects (such as low-frequency contribution from many emitters) from a small number of measurements. On the other hand, wavelet compression bases tend to lose detail when the reflectance functions contain high-frequency contributions at low magnitudes.

Since neither conflicting regions, nor the compressibility of the reflectance functions is known in advance, the required capture time for a desired result in both, adaptive and non-adaptive approaches can only be predicted by an educated guess. When using compressed sensing, choosing the number of measurements is similar to setting a desired target quality for a lossy compression scheme: The acquisition time is fixed, but the outcome is only known after the compression has been perfomed. Adaptive approaches, on the other hand, will keep capturing as long as necessary, which might lead to excessive acquisition times if the scene does not lend itself to the parallelization scheme. In this case, the need for illuminating with only a single emitter can also cause SNR problems due to the low illumination level.

By using Daubechies wavelets, which are far better suited for image compression than the Haar transform used in the examples here, Sen and Darabi achieved results in [28] that are comparable to the outcomes of the adaptive dual-imaging approach from [27] with roughly the same number of captured images, but without the need for real-time processing while the acquisition is in progress.

Another advantage of using compressed sensing is that fixed-pattern acquisition (e.g. with Bernoulli patterns) lends itself to hardware implementations. Since the capturing process does not rely on any computations outside of random number generation and image capture, the entire acquisition pipeline is easy to implement in low-cost consumer grade components. The actual reconstruction can then be performed offline on a cluster or well-equipped workstation. Although ROMP is fairly efficient, the computations required for reconstructing high-resolution light transport matrices still take in the order of days on consumer-grade PCs, but if the reflectance functions are to be reconstructed independently, the reconstruction step is very natural

[5] A broad overview that outlines the basic properties of nonstandard and standard wavelet construction is provided in [29]

to parallelize – both, in terms of data (responses for individual pixels) and computation.

Notation Overview

n	number of emitters		
p	number of receptors		
m	number of measurements		
$L \in \mathbb{R}^{n \times m}$	emitter pattern matrix	$\Phi \in \mathbb{R}^{m \times n}$	measurement ensemble
$C \in \mathbb{R}^{p \times m}$	receptor response matrix		
$C_i \in \mathbb{R}^{1 \times m}$	response vector for receptor i	$y \in \mathbb{R}^{m \times 1}$	measurement vector
$T \in \mathbb{R}^{p \times n}$	light transport matrix		
$T_i \in \mathbb{R}^{1 \times n}$	ith reflectance function (row of T)	$x \in \mathbb{R}^{n \times 1}$	unknown signal
$\hat{T}_i \in \mathbb{R}^{1 \times n}$	T_i in transform domain	$\hat{x} \in \mathbb{R}^{n \times 1}$	x in transform domain

Figure 8: An overview showing the the most important symbols used throughout this work. When the application of compressed sensing to light transport acquisition maps an entity from the compressed sensing theory to a symbol used to describe light transport, this is indicated through positioning on the same line.

References

[1] Edward H. Adelson and James R. Bergen. The Plenoptic Function and the Elements of Early Vision. In *Computational Models of Visual Processing*, pages 3–20. MIT Press, Cambridge, MA, 1991.

[2] A.C. Bovik. *Handbook of Image and Video Processing*. Elsevier, 2005.

[3] E. Candes. The Restricted Isometry Property and its Implications for Compressed Sensing. *Comptes Rendus Mathematique*, 346(9-10):589–592, May 2008.

[4] E. J. Candes. Compressive Sampling. In *Proceedings of the International Congress of Mathematicians*, Madrid, Spain, 2006.

[5] E. J. Candes and J. Romberg. Sparsity and Incoherence in Compressive Sampling. *Inverse Problems*, 23(3):969–985, June 2007.

[6] E. J. Candes, J. Romberg, and T. Tao. Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information. *IEEE Transactions on Information Theory*, 52(2):489–509, February 2006.

[7] E. J. Candes and T. Tao. Decoding by Linear Programming. *IEEE Transactions on Information Theory*, 51(12):4203–4215, December 2005.

[8] E.J. Candes and T. Tao. Near-Optimal Signal Recovery from Random Projections: Universal Encoding Strategies? *Information Theory, IEEE Transactions on*, 52(12):5406–5425, 2006.

[9] Emmanuel J. Candes and Yaniv Plan. A Probabilistic and RIPless Theory of Compressed Sensing. 2010.

[10] Emmanuel J. Candes, Justin K. Romberg, and Terence Tao. Stable Signal Recovery from Incomplete and Inaccurate Measurements. *Communications on Pure and Applied Mathematics*, 59(8):1207–1223, 2006.

[11] Scott S. Chen, David L. Donoho, and Michael A. Saunders. Atomic Decomposition by Basis Pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, 1999.

[12] Albert Cohen, Wolfgang Dahmen, and Ronald Devore. Compressed Sensing and Best k-term Approximation. Technical report, 2006.

[13] Paul Debevec, Tim Hawkins, Chris Tchou, Haarm-Pieter Duiker, Westley Sarokin, and Mark Sagar. Acquiring the Reflectance Field of a Human Face. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '00, pages 145–156, New York, NY, 2000. ACM Press/Addison-Wesley Publishing Co.

[14] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, Ting Sun, K. F. Kelly, and R. G. Baraniuk. Single-Pixel Imaging via Compressive Sampling. *IEEE Signal Processing Magazine*, 25(2):83–91, March 2008.

[15] The Blender Foundation. The Blender Website, February 2011. http://www.blender.org/.

[16] Gaurav Garg, Eino V. Talvala, Marc Levoy, and Hendrik Lensch. Symmetric Photography: Exploiting Data-sparseness in Reflectance Fields. In *Eurographics Symposium on Rendering (EGSR)*, 2006.

[17] Hermann von Helmholtz. *Handbuch der physiologischen Optik*. Leopold Voss, Leipzig, 3rd edition, 1867.

[18] Janne Kontkanen, Emmanuel Turquin, Nicolas Holzschuch, and François Sillion. Wavelet Radiance Transport for Interactive Indirect Lighting. In Wolfgang Heidrich and Thomas Akenine-Möller, editors, *Rendering Techniques 2006 (Eurographics Symposium on Rendering)*. Eurographics, 2006.

[19] LuxRender. The LuxRender Website, February 2011. http://www.luxrender.net/.

[20] D. Needell. ROMP Reference Implementation for MATLAB. http://www-stat.stanford.edu/~dneedell/romp.m.

[21] D. Needell and R. Vershynin. Uniform Uncertainty Principle and Signal Recovery via Regularized Orthogonal Matching Pursuit. *Foundations of Computational Mathematics*, 9(3):317–334, 2009.

[22] Fred E. Nicodemus, J. C. Richmond, J. J. Hsia, I. W. Ginsberg, and T. Limperis. *Geometrical Considerations and Nomenclature for Reflectance*. National Bureau of Standards (US), Washington DC, 1977.

[23] Pieter Peers and Philip Dutré. Wavelet Environment Matting. In *Proceedings of the 14th Eurographics workshop on Rendering*, EGRW '03, pages 157–166, Aire-la-Ville, Switzerland, 2003. Eurographics Association.

[24] Pieter Peers and Philip Dutré. Inferring Reflectance Functions from Wavelet Noise. In *Rendering Techniques*, pages 173–182, 2005.

[25] Pieter Peers, Dhruv K. Mahajan, Bruce Lamond, Abhijeet Ghosh, Wojciech Matusik, Ravi Ramamoorthi, and Paul Debevec. Compressive Light Transport Sensing. *ACM Transactions on Graphics*, 28(1):1–18, 2009.

[26] John William Strutt Lord Rayleigh. On the Law of Reciprocity in Diffuse Reflexion. *Philosophical Magazine Series 5*, 49:324–325, 1900.

[27] Pradeep Sen, Billy Chen, Gaurav Garg, Stephen R. Marschner, Mark Horowitz, Marc Levoy, and Hendrik P. A. Lensch. Dual Photography. *ACM Transactions on Graphics*, 24:745–755, 2005.

[28] Pradeep Sen and Soheil Darabi. Compressive Dual Photography. *Computer Graphics Forum (Proceedings of Eurographics 2009)*, 28(2), 2009.

[29] Eric J. Stollnitz, Tony D. Derose, and David H. Salesin. Wavelets for Computer Graphics: A Primer, Part 1. *IEEE Computer Graphics and Applications*, 15(3):76–84, 1995.

[30] Dharmpal Takhar, Jason N. Laska, Michael B. Wakin, Marco F. Duarte, Dror Baron, Shriram Sarvotham, Kevin F. Kelly, and Richard G. Baraniuk. A New Compressive Imaging Camera Architecture using Optical-Domain Compression. In *in Proc. of Computational Imaging IV at SPIE Electronic Imaging*, pages 43–52, 2006.

[31] J. A. Tropp and A. C. Gilbert. Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit. *Information Theory, IEEE Transactions on*, 53(12):4655–4666, 2007.

[32] Eric Veach. *Robust Monte Carlo Methods for Light Transport Simulation*. PhD thesis, Stanford, CA, 1998.

[33] Eric Veach and Leonidas J. Guibas. Metropolis Light Transport. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '97, pages 65–76, New York, NY, USA, 1997. ACM Press/Addison-Wesley Publishing Co.

Acknowledgements

The bunny model depicted in some of the figures is a simplified version of the *Stanford Bunny* model generously provided by the *Stanford University Computer Graphics Lab* through their 3D graphics repository at <`http://graphics.stanford.edu/data/3Dscanrep/`>.