Volume 0 (1981), Number 0 pp. 1–12 COMPUTER GRAPHICS *forum*

Robust Fitting on Poorly Sampled Data for Surface Light Field Rendering and Image Relighting

K. Vanhoey, B. Sauvage, O. Génevaux, F. Larue and J.-M. Dischler

ICube, Université de Strasbourg, CNRS

Abstract

2D parametric color functions are widely used in Image-Based Rendering and Image Relighting. They make it possible to express the color of a point depending on a continuous directional parameter: the viewing or the incident light direction. Producing such functions from acquired data is promising but difficult. Indeed, an intensive acquisition process resulting in dense and uniform sampling is not always possible. Conversely, a simpler acquisition process results in sparse, scattered and noisy data on which parametric functions can hardly be fitted without introducing artifacts.

Within this context, we present two contributions. The first one is a robust least-squares based method for fitting 2D parametric color functions on sparse and scattered data. Our method works for any amount and distribution of acquired data, as well as for any function expressed as a linear combination of basis functions. We tested our fitting for both image-based rendering (surface light fields) and image relighting using polynomials and spherical harmonics. The second one is a statistical analysis to measure the robustness of any fitting method. This measure assesses a trade-off between precision of the fitting and stability w.r.t. input sampling conditions. This analysis along with visual results confirm that our fitting method is robust and reduces reconstruction artifacts for poorly sampled data while preserving the precision for a dense and uniform sampling.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation—Digitizing and scanning I.4.1 [Computer Graphics]: Digitization and Image Capture—Reflectance

1. Introduction

Acquiring real-world data using geometric scans and/or photographs has received a lot of attention over the last two decades. Setups and algorithms have made it possible to recreate a subset of the plenoptic function [MB95] based on geometric information and/or images [SKC03, ZC04, Deb05]. Sophisticated acquisition techniques [DvGNK99, DHT*00, MGW01] based on domes of light sources and/or cameras allow the acquisition of both geometry and photometry with unprecedented detail and accuracy. Modeling acquired data by efficient representations (surface light fields [MRP98, WAA*00, CBCG02], Polynomial Texture Maps [MGW01], etc.) enables straightforward visualization and compression.

When an intensive acquisition campaign can be performed in controlled conditions, good results are obtained. It generally requires to move the object or the devices freely, which is not always feasible, especially for large scale objects. In many cases, a light and versatile acquisition process may be the only solution. Moreover, the image-based modeling and rendering community recently aims at simplifying acquisition. Such a process, namely one that is based on *portable devices* and that considers *non-controlled* environments, results in scattered input samples that tend to be sparse and noisy.

Our purpose is twofold. We want to enable artifact-less visualization on highly detailed models with free walk-through of the scene, even when considering poorly sampled data. Moreover, we want to contribute to the simplification of acquisition processes looked for by the community. Whereas previous methods assume dense and/or uniformly distributed data, we conversely consider the case of sparse and non-uniform sampling. In this context, we investigate the fitting of 2D parametric functions.

Our first contribution is a simple robust constrained least-squares (CLS) fitting method able to fit functions on extremely sparse data (section 4). These functions are expressed as a linear combination of basis functions and can

in particular represent surface light fields (SLF). To do so, we add a stabilization energy to the unconstrained least-squares (ULS) fitting process that prevents visual artifacts and makes the fitting robust to input sampling conditions. We demonstrate that better fitting quality is achieved for poorly sampled data obtained from simple and lightweight acquisition devices.
Stability problems mainly arise in poor sampling conditions. They are due to input noise which is a consequence of accuracy limits of acquisition devices and of reconstruction algorithms. The most frequently used visual examination is not sufficient to assess these problems. Therefore our second contribution is a statistical bias-variance analysis that is carried out by a bootstrap method (section 5). This tool allows for (i) measuring the stability with respect to input sampling conditions, (ii) balancing between precision and stability, and (iii) comparing fitting methods and function bases. It is easy to implement and provides a quantitative analysis for any fitting method.

The paper is organized as follows. Section 2 presents related work. In section 3 we state the formulation of the problem. Our new fitting method is then described in section 4. Section 5 describes a statistical analysis for assessing the robustness of any fitting method. We discuss results in section 6 and conclude in section 7.

2. Related work

2.1. Image-based modeling and rendering

The field of image-based modeling and rendering aims at reconstructing a continuous approximation of a scene based on a set of photographs. That is, it seeks to produce a virtual scene allowing free walk-through. A photograph is a set of pixels defining an instantaneous representation of the wavelength (color) (1D), observed from a specific location (3D) with a specific viewing direction (2D) at a given time (1D). It is therefore a sampling of the 7D plenoptic function [MB95]. As shown in [ZC04], successive simplifications enable to reduce the plenoptic approximation problem to the reconstruction of a 4D signal. This signal is called a light field and depends on two dimensions related to location and two others to the viewing direction.

Light field rendering [LH96] and lumigraph [GGSC96] were the first efficient representations of this 4D space. Visualization is real-time but the acquisition process, solely based on photographs, remains tedious and dense sampling is required for good results. Light fields taking advantage of the geometrical information (meshes) were designed to enable image re-projection on it, thus improving accuracy and reducing memory consumption while exploiting modern graphics hardware in a better way. View-dependent texture mapping [DTM96, DYB98] updates textures based on photographs to project them onto very coarse geometry, for instance for approximating city-like scenes. It assumes sparse data, but is not designed to handle detailed geometry. When fine geometrical detail is favored, surface light fields (SLF) offer a better solution. There are two approaches for representing them: global factorization methods [MRP98, NSI01, CHLG05] that process the light field as a combination of eigen-textures, and local methods [WAA*00, CBCG02, CL06] that express the color by defining a 2D function on the visible hemisphere independently at each surface point (usually vertices or texels).

Image relighting techniques handle a scene viewed from a fixed point of view, the resulting image being defined as a function of the lighting conditions. Conversely to SLF, the geometry is ignored in this case, discarding sampling problems induced by self-occlusion. However, to completely avoid reconstruction artifacts, highly controllable acquisition conditions are required. For image relighting too, both global interpolation/factorization and local representations coexist. Global methods are useful for defining an image as a function of general lighting conditions (an environment map for instance). Light transfer matrices form a recent and powerful tool but are exploitable and precise only when the lighting environment can be modified at will during acquisition [WDT*09, PML*09, OK10]. Local representations on the other hand represent each pixel color as a function of the direction of a unique point light source, parametrized on a hemisphere. Here again, it is required to use video footage for dense acquisition [MDA02] or rigid devices [DHT*00, HED05, FLBS07, FBLS07] able to control light sources uniformly arranged around the object.

During an on-site acquisition campaign, lighting conditions and/or viewpoint sampling may not be controllable. In this paper we therefore address the problem of robustness in fitting 2D parametric functions on input data for SLF rendering and image relighting.

2.2. Fitting 2D parametric functions

A light field can be defined as a color for every surface point (2D) and for every viewing/lighting direction (2D) and can therefore be expressed by defining a 2D hemispherical function per point. The fitting process consists in estimating the parameters of such a function in order to fit the input samples. Input samples are a set of colors located on the visible hemisphere. This location is deduced from projection of photographs onto the surface (SLF) or from the light source position (for image relighting).

Non-linear functions (for example [LFTG97]) are precise but require dense sampling [WDR11] and are rather used for higher-dimensioned (4D) hemispherical functions. In 2D, linear combinations of basis functions are often used owing to their simplicity. They enable to express the fitting problem as a set of linear equations. Therefore, the fitting and evaluation are simple and efficient. The most commonly used are spherical harmonics [Mac48], polynomials [MGW01], spherical wavelets [SS95] and lumispheres [WAA*00].

We solve the problem of fitting hemispherical functions to sample data by constrained least-squares. Woods et al. [WAA*00] used this approach to solve the under-constriction problem resulting from occlusions. Lam et al. [LLW06] constrain a fitting such that the post-processing compression noise is minimized. By contrast we study the impact of the input sampling on the fitting process.

2.3. Statistical analysis

In the field of statistical learning, fitting is known as *regression*. Our representation for the fitted functions corresponds to a *linear basis expansion*. The stabilization is called a *regularization method*. It is however a sensitive method that requires tuning. In this context, bias-variance analysis is a powerful assessment tool [HTF01]. For our purpose it enables both to compare fitting methods and to balance between precision and stability.

This analysis mainly evaluates the *expected prediction error* which measures the method's capability in successfully fitting on new data sets. It may be implemented in many ways. Since the SLFs can be arbitrarily complicated and since several perturbations are mixed during data acquisition and processing, we are not able to define a precise statistical model of the data. Thus implementations requiring much *a priori* knowledge are difficult to apply: general methods (*e.g.* cross-validation or bootstrap) better suit our purpose. Another obstacle is the scarcity of the data. The analysis is made easier if a large data set is available for each hemisphere. However, dense acquisition of real world data is tedious, and we aim at validating the fitting in sparse sampling conditions. Therefore we implemented this analysis via bootstrap [HTF01] which is more stable in this setting.

3. Definitions

For the sake of clarity we will consider in the following only SLF. However, the proposed methods and results hold for image relighting techniques, as illustrated in section 6. SLF is also the most difficult case: since it is defined on the geometry, auto-occlusions occur and a complex reconstruction pipeline worsens data sampling quality. Our purpose is then to represent the color of a 3D model for any viewing direction. So we consider the fitting of scalar-valued functions over the visible hemisphere (one per color channel) for each spatial location on a surface. We note that all results still hold for functions defined over the entire sphere.

3.1. Least-Squares Fitting

A function $f(\omega)$ defined over the hemisphere Ω is represented as a linear combination of K basis functions by

$$f(\omega) = \sum_{i=1}^{K} c_i \varphi_i(\omega) = C^T \Phi(\omega) \tag{1}$$

where $C = [c_1, \ldots, c_K]^T$ is the vector of coefficients, and $\Phi(\omega) = [\varphi_1(\omega), \ldots, \varphi_K(\omega)]^T$ is the vector of basis functions.

As an input, we are given a set of N samples $\{(\omega_1, v_1), \ldots, (\omega_N, v_N)\}$ defined by a direction $\omega_n \in \Omega$ and a color coordinate v_n.

The problem of fitting f onto the samples is commonly solved by least-squares, *i.e.* by minimizing the mean square error (MSE):

$$E_{MSE} = \frac{1}{N} \left\| \begin{bmatrix} \Phi(\omega_1)^T \\ \vdots \\ \Phi(\omega_N)^T \end{bmatrix} C - \begin{bmatrix} v_1 \\ \vdots \\ v_N \end{bmatrix} \right\|^2 \tag{2}$$

For dense sampling ($N \gg K$) the problem is in general well-constrained and the minimization is stable. Sparse sampling may result in under-constriction and instability. Both drawbacks are often overcome by minimizing a weighted combination of E_{MSE} and a stabilization energy E_{stab}:

$$\underset{C}{\operatorname{argmin}} (1-\lambda) E_{MSE} + \lambda E_{stab} \tag{3}$$

3.2. Scattered data classification

We consider two different properties of the sampling distribution that impact on the stability of the fitting step:

Sparsity. Due to complex geometry and the inability to take large amounts of photographs, many surface points are visible from only a few images. Therefore, handling a few samples (N) compared to the number of basis functions (K) is a key problem for fitting. We call the data *very sparse* when $N \ll K$. *Sparse* data refers to $N \approx K$. *Dense* data refers to $N \gg K$.

Uniformity. Large parts of the hemisphere may contain no samples. This can be due to unreachable viewing directions (*e.g.* top views of large objects) or to self-occlusion. We call a distribution uniform when samples are spread over the entire hemisphere. A distribution is non-uniform when there are no samples in significant parts of it. More than half of the hemisphere may be concerned.

The more sparse and non-uniform the sampling, the more difficult the stabilization.

4. Robust Fitting Method

Stability issues are mainly caused by sparsity and are magnified in non-covered areas. For example with a low amount of color-noise in figure 1a (input samples contain only white and magenta), Unconstrained Least Squares (ULS) fitting introduces high frequencies that make unwanted colors appear (fig. 1b).

In the setting of eq. (3), the challenge consists in choosing E_{stab} and λ. Though many choices would numerically overcome the under-constriction, analyzing the causes of the

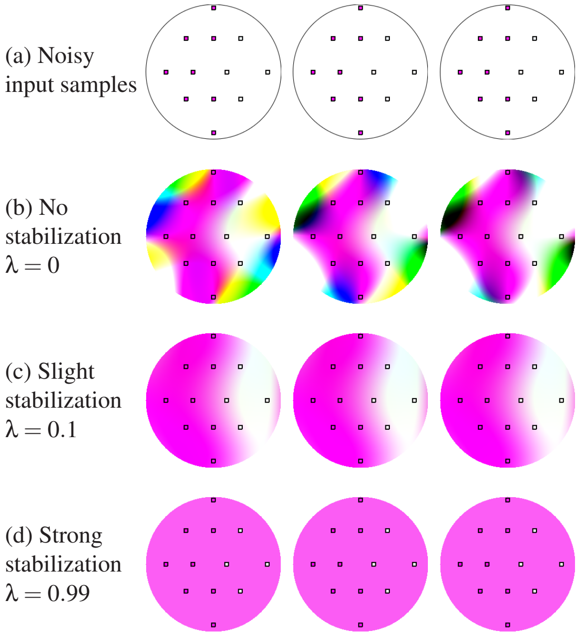


Figure 1: Hemispherical functions are fitted to input samples (1a) containing 1% color noise. Without any stabilization (1b) the fitted functions show huge color variations. A slight stabilization (1c) prevents variations. With a strong stabilization (1d) the fitted functions are over-smoothed.

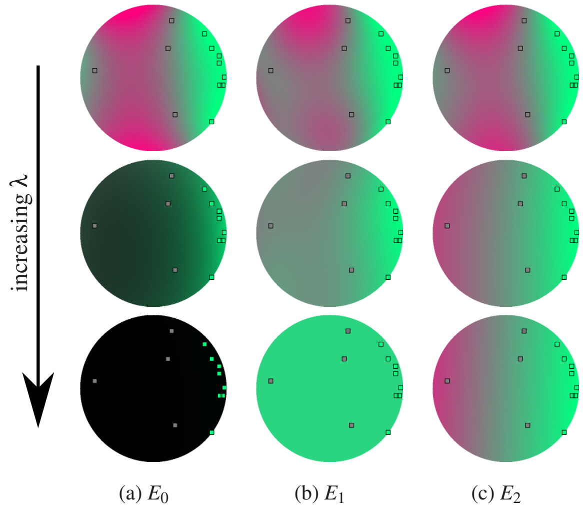


Figure 2: Comparison of the energies E_{stab}. All functions fit the same input green and gray samples. E_0, E_1 and E_2 behave differently as λ increases.

instability and the results we expect is essential for making a wise choice.

We tested two different energies that have been used for Constrained Least Squares (CLS) fitting in related works. By analyzing their drawbacks, we propose a new energy that aims at improving the robustness of the fitting. We label these energies according to the derivation order of the fitted function f.

The 0-order energy

$$E_{stab} = E_0 = \frac{1}{Area(\Omega)} \int_{\Omega} \|f\|^2 \tag{4}$$

has been used for the fitting process of Spherical Harmonics in the CLS technique presented in [LLW06] to limit post-processing compression noise. An important drawback in our context is that increasing the stabilization (i.e. λ) pulls the fitted function towards zero (i.e. black color) in regions where no or few samples are located, as shown in figure 2a. This is compensated in [LLW06] by fixing a target energy, which is intricate in our case due to poor sampling conditions.

The 2^{nd}-order (bending or thin-plate) energy is defined in an (x, y) parametric domain by

$$E_{stab} = E_2 = \iint \frac{\partial^2 f}{\partial^2 x} + 2\frac{\partial^2 f}{\partial x \partial y} + \frac{\partial^2 f}{\partial^2 y} \tag{5}$$

On a domain without border (e.g. an entire sphere), E_2 is equal to $\frac{1}{Area(\Omega)} \int_{\Omega} \|\Delta f\|^2$, where Δ is the Laplace-Beltrami operator [Wah81]. It was used in [WAA*00] for fitting lumispheres (locally supported functions parametrized on the sphere) in order to cope with the occlusion problem which under-determines the fitting. E_2 penalizes high frequencies. However, it sometimes fails to prevent unexpected extrapolated values because linear functions have $E_2 = 0$. This can be observed in figure 2c: though the input samples contain only green and gray, magenta still appears even after strong stabilization (high λ).

We argue that a better energy should prevent strong color variations (i.e. high frequencies) in areas of the hemisphere not covered by samples. It should also prevent the functions to exceed the range of the color space in these areas: such out-of-bound color values would create artifacts when clamped during rendering. In other words the energy should flatten the function. Therefore we introduce the 1^{st}-order energy

$$E_{stab} = E_1 = \frac{1}{Area(\Omega)} \int_{\Omega} \|\nabla f\|^2 \tag{6}$$

The relevance of E_1 lies in its goal: when increasing λ or when the number of samples N equals 1, it pulls f towards a constant value, which represents a diffuse color. This behavior is observed when increasing λ in figure 2b. E_1 can be computed in any function basis as a convex quadratic form w.r.t. the coefficients. Depending on the basis, integration can be analytical or numerical.

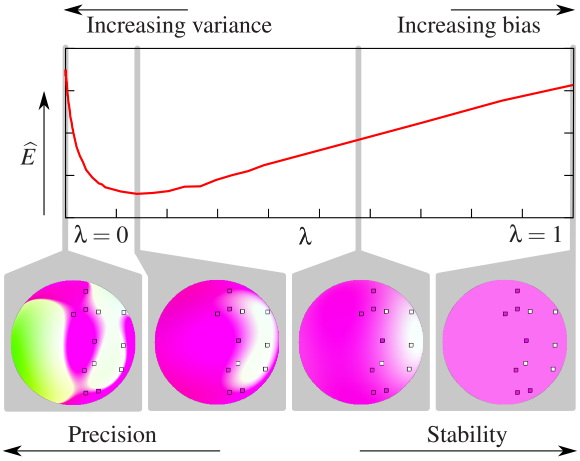


Figure 3: Interpretation of expected prediction error graphs. $\lambda = 0$ corresponds to unconstrained least-squares. Small λ values imply high precision but low stability: the error $\widehat{E}$ is penalized by high variance. Large λ values imply low precision but high stability: the error $\widehat{E}$ is penalized by high bias.

5. Statistical Analysis

The first goal of any fitting method is to precisely fit to input data. However, other issues (*e.g.* stability, smoothness) may also be important. We investigate the stability, defined as the resistance of the method to input sampling conditions. Precision and stability change conversely when intensifying the stabilization (*i.e.* increasing λ in eq. 3): the stability increases while the precision decreases. A tool for assessing a trade-off is useful.

5.1. Bias versus variance

Figure 1 provides an intuition of this trade-off. Three input sample sets (fig. 1a) have imperceptible differences. Without any stabilization (fig. 1b) the precision is high but the result suffers strong variance (the fitted functions differ from each other). The variance measures how the actual fittings deviate from their mean. A slight stabilization (fig. 1c) overcomes the problem (the three functions are alike) while still closely fitting to the input samples. A strong stabilization (fig. 1d) flattens the resulting function: the precision drops (poor approximation of the input samples) because bias is introduced. The bias measures how the actual fittings deviate from "true values".

The *expected prediction error* is a statistical measure which includes both variance and bias. It measures the capability of a method in successfully fitting on new data sets. As shown in figure 3 it can be plotted against λ. As λ grows, this error first decreases because stabilization reduces the variance. Then it increases because stabilization raises the bias.

This curve allows for selecting λ: the minimum error is

(a) Scanner setup (b) An input photo (c) All viewpoints

Figure 4: Acquisition procedure for the double dragon model. 4a: the scanner setup used for geometry measurement. 4b: one of the photographs taken during photometry acquisition. 4c: placements of the 200 photographs acquired using a hand-held camera.

an optimal trade-off between precision and stability. In the purpose of comparing different methods (choice of E_{stab}), several curves can be compared on the same graph. In section 6 we also compare curves in order to evaluate the impact of a change in the basis, in the color space, or in the sampling density.

5.2. Bootstrap

In order to compute the trade-off curves the expected prediction error has to be estimated. As discussed in section 2.3 our implementation has to cope with both scarce input data and little knowledge about data distribution. Indeed, when resulting from an acquisition of real objects, the input data are perturbed by color and direction noise resulting from registration problems and the lack of accuracy of the acquisition devices. Besides, registration failures or self-occlusions may result in missing viewing directions. It is difficult to derive not only a statistical model of such perturbations but also the corresponding expected prediction error. As a consequence, we estimate the error using a bootstrap process, which does not require any statistical data model.

The process is given a set $\mathcal{T}$ of N samples as input. The main idea is to repeatedly divide $\mathcal{T}$ into a training set (called bootstrap) and a validation set. The training sets are used to fit the functions while the validation sets are used to estimate the error. The bootstrap process consists of 3 stages [HTF01]:

- B bootstraps $\{\mathcal{T}_1, \ldots, \mathcal{T}_B\}$ are formed. Each $\mathcal{T}_b$ is a set of N samples randomly drawn with replacement from $\mathcal{T}$. Note that redundancy will appear within $\mathcal{T}_b$.
- A function f_b is fitted to each bootstrap $\mathcal{T}_b$.
- For each sample $n \in \mathcal{T}$ the squared error is averaged over the bootstraps that do not contain n. The error is eventually estimated by the average over the samples:

$$\widehat{E} = \frac{1}{N} \sum_{n=1}^{N} \frac{1}{\#\{b/n \notin \mathcal{T}_b\}} \sum_{b/n \notin \mathcal{T}_b} (y_n - f_b(x_n))^2 \tag{7}$$

This estimation is computed for every λ in order to plot the curve.

(a) Input pictures

(b) No stabilization (ULS): $PSNR = 21.51dB$

(c) Slight stabilization (CLS): $PSNR = 19.59dB$

(d) Strong stabilization (CLS): $PSNR = 14.51dB$

Figure 5: Reconstruction of the double dragon model (polynomial basis of degree 4, stabilization with the E_1 energy). Images on the two left-most columns match one of the input viewpoints (fig. 5a). Without stabilization, artifacts appear when the virtual camera moves away from the acquired viewpoints (fig. 5b). A slightly stabilized fitting removes artifacts while still preserving view-dependent features (fig. 5c). A too strong stabilization removes both artifacts and view-dependent features (fig. 5d).

Bootstrap is efficient for sparse data (N small) because one can increase B independently of N. This is especially useful for testing real-world data whose density can hardly be controlled. By using $B = 10^4$ bootstraps our experiments achieve about 2 significant digits for $\widehat{E}$.

5.3. Measuring precision

$\widehat{E}$ may be quite large because it estimates the expected error on new data sets. It is useful for selecting λ and for comparing methods but it does not measure the actual fitting precision.

In the context of very sparse data, which we know to be too sparse to capture complex material behavior, it is vain to try to reproduce an underlying ground truth. Indeed, this would require strong assumptions or external knowledge such as lighting conditions or ad-hoc material models. The versatile acquisition process we consider hinders such assumptions. Therefore, we will not use a reference fitting to compare any result to. The input samples are considered as the only truth. Thus, the precision of a fitted function is directly given by E_{MSE} (defined in eq. (2)). Since the range of values is very large we use the peak signal-to-noise ratio [GG92]:

$$PSNR = 10\log_{10}\left(\frac{E_{max}}{E_{MSE}}\right) \tag{8}$$

such that an absolute difference beneficially represents a relative ratio: $10dB$ represents a factor of ten.

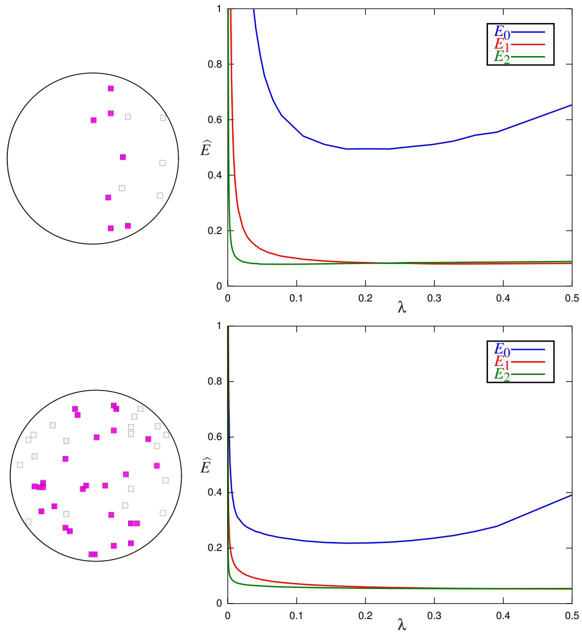


Figure 6: Error graphs for different input data sets using spherical harmonics ($l = 4$). Top: sparse and non uniform. Bottom: dense and uniform.

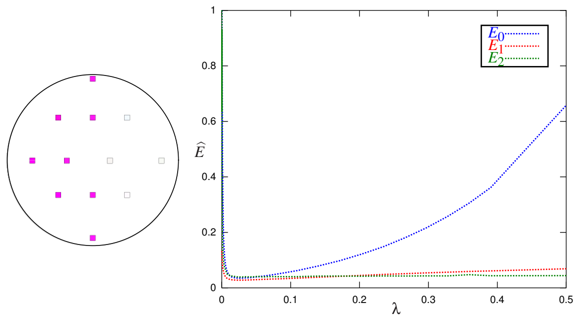


Figure 7: Numerical comparison of energies. Polynomial basis ($d = 4$) fitting on a sparse and uniform sampling.

6. Results

The need for stabilization depends on many parameters:

- The noise induced by the acquisition and reconstruction pipeline.
- The sampling distribution. According to our classification of section 3.2 we tested various sampling qualities.
- The basis functions. We experimented with two widely-used bases: Spherical Harmonics (SH) and Polynomial Basis (PB). Polynomials of degree 2 are common for representing Polynomial Texture Maps [MGW01]: we extended it to higher degrees and applied it also to SLFs.
- The size of the basis. We tested SH bases with $l = 2$ up to 20 levels (containing l^2 functions). We tested PBs of degree $d = 2$ up to 4 (containing $(d+1)(d+2)/2$ functions).
- The color space.

In the present section we discuss all these parameters using numerical results from our statistical analysis. Since the user may also be interested in a subjective quality of the fitting (he may expect "pleasant" results), numerical and visual results are also compared.

6.1. Experimental data

Experiments were performed on four different types of data revealing different properties of the fitting process.

Standalone light fields show point-wise fitting behavior. Besides, they are used as input to our statistical analysis in order to generate quantitative results with controlled sampling conditions. We therefore built different samplings corresponding to the properties listed in section 3.2. We computed the graphs of these fittings for different stabilization energies and bases.

Real-world SLF data were acquired as illustrated in figure 4. Geometry is acquired with a 3D-scanner operating with structured light. A unique point cloud is obtained by registering and merging different scans. A polygonal mesh is then reconstructed using the Poisson Surface Reconstruction algorithm [KBH06]. Photographs are taken with a hand-held high-resolution camera. Pattern-based camera calibration and picture registration is used for sample projection.
Such real data emphasize the need for stabilization: even if hundreds of photographs are taken, many hemispheres are poorly sampled because some directions are obstructed by the object or by the ground.
The *double dragon* model (fig. 5) is interesting for its very detailed features both in geometry (subtle details and compound curves) and appearance (different colors and reflectance properties). Photometry was acquired using 200 photographs (fig. 4c). SLFs were computed per vertex (1.8M vertices). Hemispheres have in average 81 samples but 5.5% have less than 20 samples.
The *Mask* model (figs. 13 and 12) illustrates the case of an intensive measurement of a real object: 672 photographs were taken and registered to the geometry. The mesh has been parametrized to a texture image and SLFs were computed per texel (2.8M texels).

Real-world image relighting data illustrate our method when no occlusions arise. Indeed, for image relighting, geometry is not accounted for and all color samples are always defined for every pixel. Thus, it mainly shows the results when the light direction moves away from the acquired samples. For the *Capitello* (fig. 11), 36 photographs were acquired with a resolution of 1000×750 pixels.

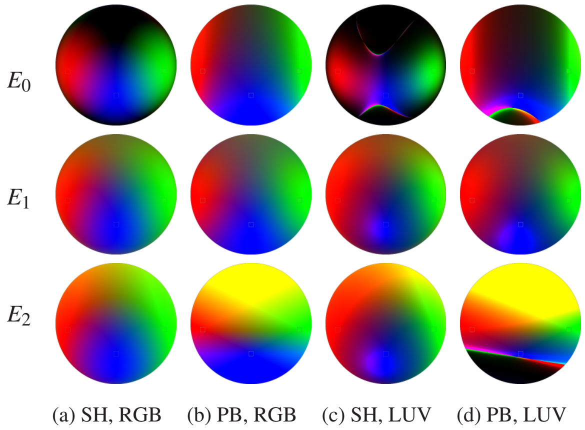


Figure 8: Evaluation of the consistency w.r.t. color space and function basis of E_0 ($\lambda = 0.01$), E_1 ($\lambda = 0.01$) and E_2 ($\lambda = 0.001$). CLS fitting is applied to only three (respectively red, blue and green) samples in two bases (SH, $l = 4$, and PB, $d = 4$) for two color spaces (RGB and CIELUV).

Synthetic data were created by making virtual photographs based on Monte-Carlo ray-tracing. Since the lighting environment is controlled and since noise induced by the acquisition pipeline is avoided, such data has the advantage of isolating the sampling condition issues.
The *Stanford Bunny* (fig. 9) was lit with 3 light spots. 51 virtual photographs are not uniformly distributed around the object. An average of 11 input samples per hemisphere is obtained. 6.2% of the hemispheres have less than 5 samples.

6.2. Robustness issues

The need for stabilization. Despite intensive photographic acquisitions, real data nearly always suffer from poorly sampled hemispheres. The double dragon model (fig. 5) contains almost every situation listed in section 3.2:

- hemispheres oriented upwards are densely and uniformly sampled;
- hemispheres oriented downwards are very sparsely sampled;
- in-between hemispheres may be densely but non-uniformly sampled due to occlusions.

A naive unconstrained fitting (ULS) is not stable enough to prevent artifacts when the viewing direction moves away from the acquired viewpoints (see fig. 5b on the right). A stabilization may eliminate most of them (fig. 5c). Yet it must be slight enough (*i.e.* λ small) to preserve the viewing direction dependent features: in our example of figure 5, a slight stabilization ($\lambda = 0.01$) only reduces the average PSNR from $21.51dB$ to $19.59dB$.

Numerical results (fig. 6) confirm these visual observations. For sparse and non-uniform sampling (top), the prediction error of the ULS fitting (any curve at $\lambda = 0$) is high. A slight stabilization ($\lambda = 0.05$) improves the stability. The same stabilization is also efficient for dense sampling (bottom) while preserving the precision: the PSNR barely decreases from $10dB$ (ULS) to $9dB$ (with E_1). This suggests that a slight stabilization is effective when needed while limiting precision loss when the sampling is good.

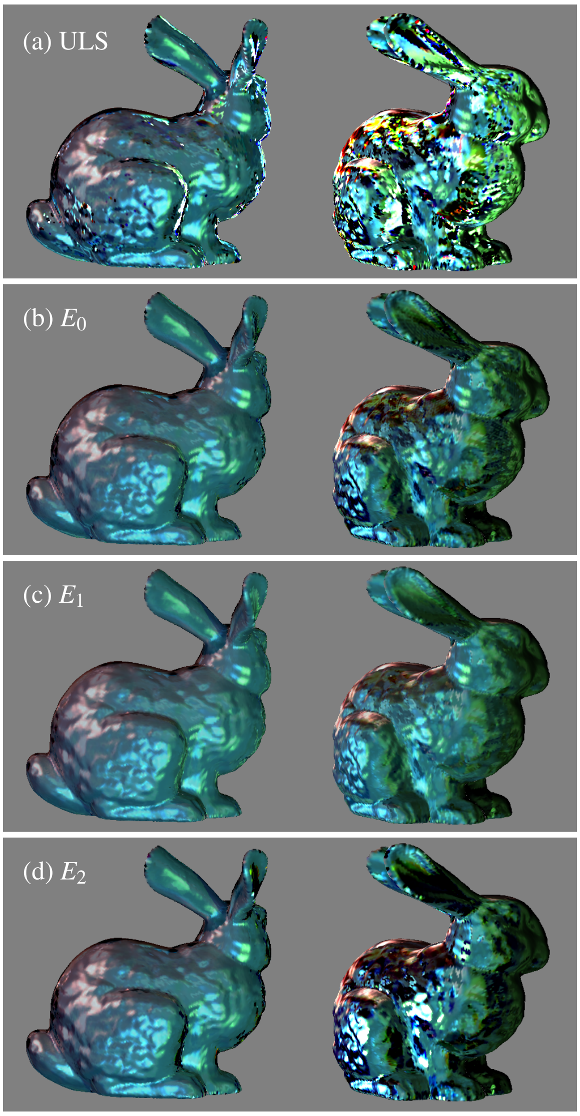
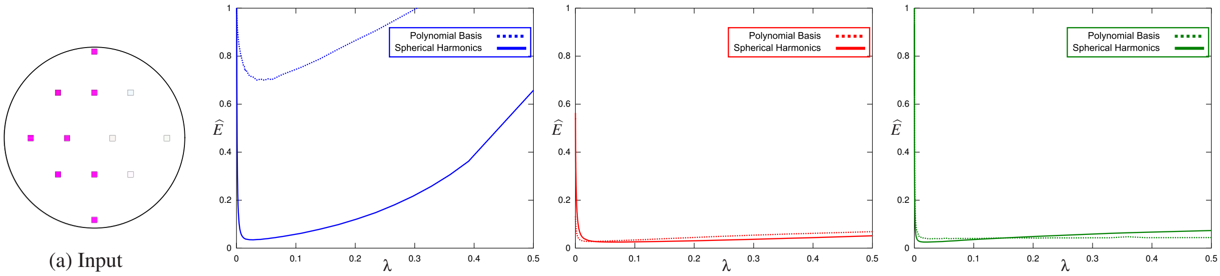


Figure 9: Virtual light field using $SH(l = 4)$. Visual comparison of energies: our energy E_1 (in row (c)) provides the most consistent results.

Comparing the stabilization energies is a key point of the present study. When the sampling is dense or uniform the three energies reach similar trade-offs (see fig. 7). Artifacts arise for poor samplings.



(a) Input

(b) Energy E_0

(c) Energy E_1

(d) Energy E_2

Figure 10: Comparison between polynomial basis ($d = 4$, $K = 15$, dashed lines) and spherical harmonics ($l = 4$, $K = 16$, solid lines): fitting on a sparse and uniform sample. E_0 fails to stabilize in SH bases because of the sparsity while E_1 and E_2 provide similar results in PB and SH.

As mentioned in section 4, E_0 tends to produce black color in non-sampled areas of the hemisphere. This can be observed in figure 8 top, and causes artifacts on the bunny (fig. 9b). The prediction error contains much bias, even for small λ values. Therefore, the error curves are unable to reach low values (figure 6 top).

On the other side E_2 may fail to prevent extrapolated values (yellow regions in figure 8 bottom). This is illustrated in figure 9d by excessive intensity variations.

In contrast our energy E_1 fairly interpolates between the 3 input colors only (fig. 8 middle). The result on the bunny (fig. 9c) is better balanced between preserving specular highlights and preventing artifacts.

Robustness w.r.t. the basis. It may be interesting for the user to choose the function basis independently of the fitting method. Therefore we analyze the behavior of each E_{stab} w.r.t. the basis. By comparing figures 8a vs 8b and 8c vs 8d, we see that E_0 and E_2 do not behave the same when the basis changes whereas our energy E_1 provides very similar and predictable results. This emphasizes that the artifacts discussed in the previous paragraph appear rather with spherical harmonics for E_0 and rather in polynomial bases for E_2. Figure 10 compares the PB and SH bases with a similar number of basis functions on a sparse and uniform sample: E_0 exhibits much higher errors in SH bases.

Robustness w.r.t. the size of the basis. Complex lighting can only be captured when a rich enough function basis is used. Indeed, a too poor basis will induce an incompressible model bias due to its inability to represent complex inputs, resulting in over-smoothing. This can be observed in figure 13 where highly specular effects are smoothed when decreasing the number of basis functions.

To represent complex reflection behaviors, the user may be interested in safely increasing the number of basis functions. When doing so, consistent results are expected. Figure 11 compares image relighting obtained from a conventional bi-quadratic PTM [MGW01] and a bi-quartic PTM. Compared to ULS, E_1 provides consistent fittings when changing the size of the basis.

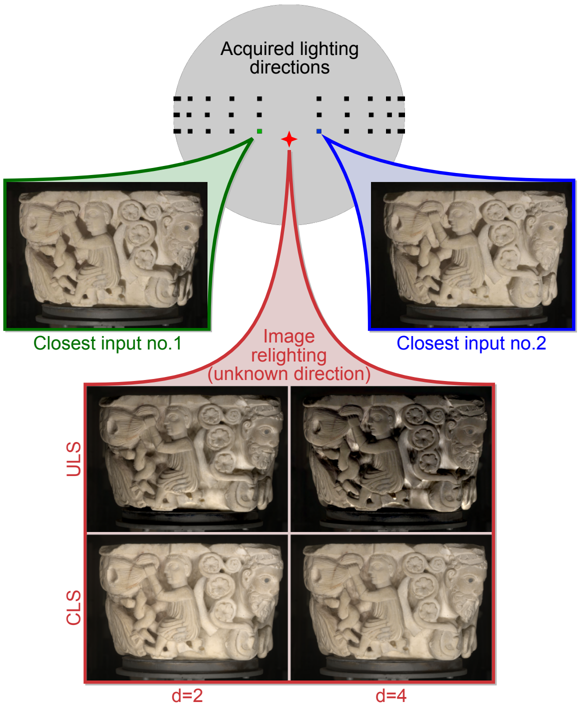


Figure 11: Robustness w.r.t. the size of the basis: CLS fitting is consistent when changing the degree. For an image relighting application, photographs were taken corresponding to 36 point light source directions. The fittings are performed with PB of degree 2 and 4 (using ULS or CLS with E_1 and $\lambda = 0.01$). The disk represents the lighting directions on the hemisphere: squares are input directions (real pictures) while the red star is a relighting direction (virtual picture).

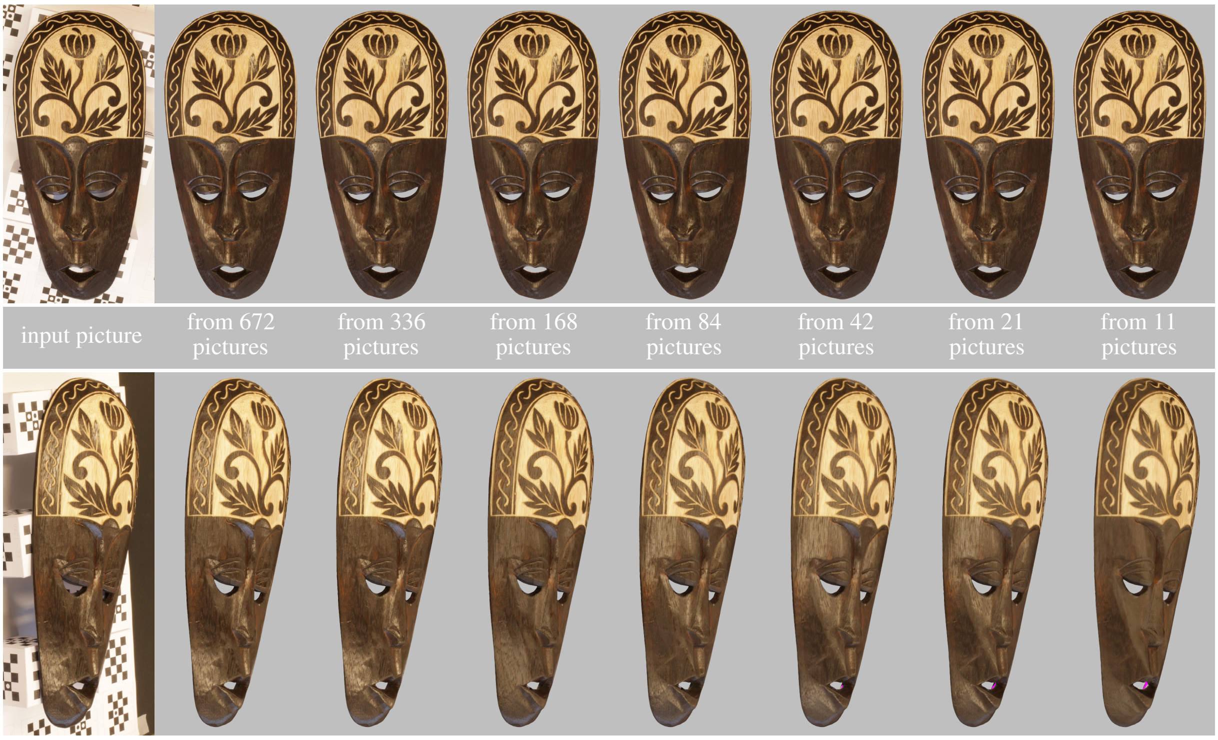


Figure 12: Robustness w.r.t. sparsity. The SLF for the mask model has been reconstructed from different view samplings roughly uniformly distributed. Fitting was performed in SH ($l = 20$) with E_1 and $\lambda = 0.01$. The input picture of the first line is included in all samplings (down to 11 pictures): few differences are visible. The input picture of the second line is discarded from the samplings below 336 pictures: the result looks smoothed but remains consistent.

Robustness w.r.t. sparsity. Figure 12 shows the reconstruction of the mask for various numbers of input pictures. By using a very large amount of basis functions, model bias is almost discarded such that the fitting method can be finely observed on this picture. Though degradation is unavoidable, our method provides consistent results when the number of samples decreases.

Robustness w.r.t. color space. We mainly used RGB but other color spaces might be appropriate. We experimented with the CIELAB and the CIELUV space. The same robustness problems occur and the same ranges of λ are effective for reducing artifacts.

It may be interesting for the user to choose the color space independently of the fitting method. By comparing the figures 8a vs 8c and 8b vs 8d, we see that, in contrast to E_0 and E_2, our energy E_1 is much more consistent when changing the color space. Note also that the discontinuities with E_0 and E_2 in CIELUV caused by out-of-bound color values are prevented with E_1, as it was designed for (see section 4).

6.3. Balancing precision and stability

The weighting factor λ in equation (3) balances between E_{MSE} and E_{stab}, i.e. between precision and stability of the fitting. A good weighting value λ can be defined as high enough to discard visual artifacts (high variance) and low enough to avoid over-smoothing (high bias). Real data exhibit various sampling conditions over the surface. Computing the optimal λ for each surface point would be computationally intensive. Instead we suggest that a constant value can be chosen. When applied on standalone light fields our statistical analysis is useful for evaluating the robustness and comparing the energies. In order to determine λ we have to estimate $\widehat{E}$ on real-world data. Therefore we applied the following pragmatic protocol.

We computed the error curves for 10^4 surface points of the double dragon. From these 10^4 curves we plotted the point-wise $95th$, $98th$ and $99th$ percentile curves (figure 14) which roughly represent the behavior of the 5%, 2% and 1% worst sampled points. The very low $95th$ percentile curve proves that most of the surface points are well sampled and barely require stabilization: λ must be as close as possible to the optimum for these points in order to avoid global over-smoothing. The $98th$ and $99th$ percentile curves represent

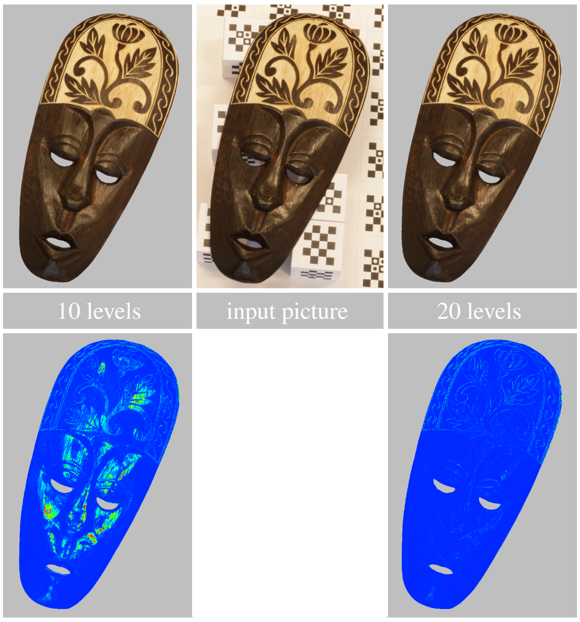


Figure 13: The larger the basis, the lower the bias: a poor basis is unable to represent complex reflection behaviors. The right-hand side image shows a SLF reconstruction using 20 SH levels (E_1 stabilization, $\lambda = 0.01$). Below, the difference w.r.t. the reference image (center) suggests that the bias is very low. Conversely, the left-hand side SLF reconstruction was fitted on 10 SH levels. Under equivalent stabilization parameters, the loss in quality suggests that the degradation is indeed due to the basis size.

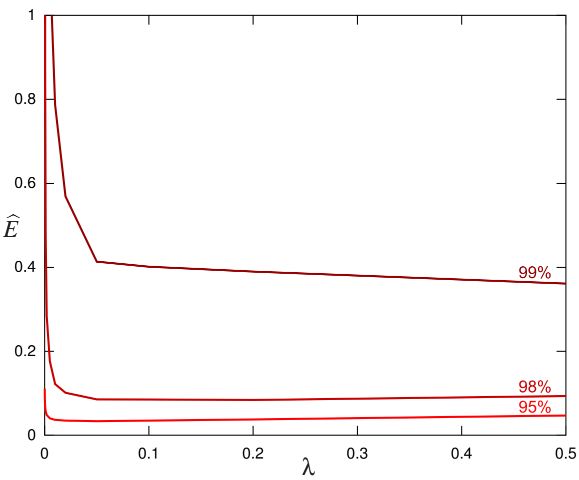


Figure 14: Point-wise 95*th*, 98*th* and 99*th* percentile error curves. Fittings are performed in SH basis ($l = 4$) stabilized with E_1 on 10000 surface points of the double dragon.

the minority part of poorly sampled surface points: λ must be just high enough for these points in order to avoid local artifacts. We deduce that E_1 performs well for $\lambda \in [0.01, 0.05]$. Similar tests show that $\lambda \in [0.01, 0.05]$ is suitable for E_0 and that $\lambda \in [0.001, 0.005]$ is suitable for E_2.

These actual values of λ are derived from one object but our tests suggest they remain valid for others. We conjecture that they rather depend on the noise magnitude, which is not controlled in the statistical analysis. Since the noise closely depends on the acquisition devices and on the reconstruction algorithms, it would be worth applying the same protocol when changing the acquisition process.

7. Conclusion

In this paper we investigated the reconstruction of color functions for surface light field rendering and image relighting by fitting parametric functions. We laid out the need for stabilization in case of sparse and non-uniform sampling in order to prevent fitting artifacts while preserving high frequencies present in the input data. We defined a new stabilization energy that makes the fitting robust and compared it with two other known energies. Our experiments were held on synthetic standalone light fields, synthetic models with simulated photographic acquisition and real-world acquired data. Both visual and numerical results show that our energy improves the stability with respect to the sparsity of the sampling, to the non-uniformity of the sampling, to the function basis (polynomials and spherical harmonics were tested), to the size of the basis, and to the color space.

To complement usual visual results, we provided a statistical analysis for measuring the robustness of any fitting method. We also derived a pragmatic protocol for balancing between precision and robustness, such that no additional adjustment is needed to handle poorly and well sampled data in the same process. It is easy to implement and can be used to tune the fitting on its own acquired data.

In this study we processed the data independently at each point/texel. Within this setting, stabilization proved to be effective in removing artifacts. However the fitted functions will always exhibit few variations when the data are too sparse to locally capture complex material behavior. To further improve the results, we could take the spatial coherence into account by propagating directional information between neighboring texels. Besides, additional knowledge about the material and/or the environment could be introduced.

Acknowledgements

We would like to thank Sophie Lèbre for her precious advice concerning statistical tools and the VCG group of the CNR of Pisa for giving us the *Capitello* acquired data.

References

[CBCG02] CHEN W.-C., BOUGUET J.-Y., CHU M. H., GRZESZCZUK R.: Light field mapping : Efficient representation and rendering of surface light fields. *ACM Transactions on Graphics (TOG) 21*, 3 (2002), 447–456. 1, 2

[CHLG05] COOMBE G., HANTAK C., LASTRA A., GRZESZCZUK R.: Online construction of surface light fields. In *Rendering Techniques* (2005), pp. 83–90. 2

[CL06] COOMBE G., LASTRA A.: An incremental weighted least squares approach to surface lights fields. In *GRAPP* (2006), pp. 84–91. 2

[Deb05] DEBEVEC P.: Image-based lighting. In *ACM SIGGRAPH 2005 Courses* (New York, USA, 2005), SIGGRAPH '05, ACM. 1

[DHT*00] DEBEVEC P., HAWKINS T., TCHOU C., DUIKER H.-P., SAROKIN W., SAGAR M.: Acquiring the reflectance field of a human face. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques* (New York, USA, 2000), SIGGRAPH '00, ACM Press/Addison-Wesley Publishing Co., pp. 145–156. 1, 2

[DTM96] DEBEVEC P. E., TAYLOR C. J., MALIK J.: Modeling and rendering architecture from photographs: a hybrid geometry- and image-based approach. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques* (New York, USA, 1996), SIGGRAPH '96, ACM, pp. 11–20. 2

[DvGNK99] DANA K. J., VAN GINNEKEN B., NAYAR S. K., KOENDERINK J. J.: Reflectance and texture of real-world surfaces. *ACM Trans. Graph. 18*, 1 (1999), 1–34. 1

[DYB98] DEBEVEC P., YU Y., BOSHOKOV G.: *Efficient View-Dependent Image-Based Rendering with Projective Texture-Mapping*. Tech. Rep. UCB/CSD-98-1003, EECS Department, University of California, Berkeley, 1998. 2

[FBLS07] FUCHS M., BLANZ V., LENSCH H. P., SEIDEL H.-P.: Adaptive sampling of reflectance fields. *ACM Trans. Graph. 26*, 2 (June 2007). 2

[FLBS07] FUCHS M., LENSCH H. P. A., BLANZ V., SEIDEL H.-P.: Superresolution reflectance fields: Synthesizing images for intermediate light directions. *Comput. Graph. Forum 26*, 3 (2007), 447–456. 2

[GG92] GERSHO A., GRAY R. M.: *Vector quantization and signal compression*. Kluwer Academic Publishers, 1992. 6

[GGSC96] GORTLER S. J., GRZESZCZUK R., SZELISKI R., COHEN M. F.: The lumigraph. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques* (New York, USA, 1996), SIGGRAPH '96, ACM, pp. 43–54. 2

[HED05] HAWKINS T., EINARSSON P., DEBEVEC P. E.: A dual light stage. In *Rendering Techniques* (2005), pp. 91–98. 2

[HTF01] HASTIE T., TIBSHIRANI R., FRIEDMAN J.: *The Elements of Statistical Learning*. Springer Series in Statistics. Springer New York Inc., New York, USA, 2001. 3, 5

[KBH06] KAZHDAN M., BOLITHO M., HOPPE H.: Poisson surface reconstruction. In *Proceedings of the fourth Eurographics symposium on Geometry processing* (Aire-la-Ville, Switzerland, Switzerland, 2006), SGP '06, Eurographics Association, pp. 61–70. 7

[LFTG97] LAFORTUNE E. P. F., FOO S.-C., TORRANCE K. E., GREENBERG D. P.: Non-linear approximation of reflectance functions. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques* (New York, USA, 1997), SIGGRAPH '97, ACM Press/Addison-Wesley Publishing Co., pp. 117–126. 2

[LH96] LEVOY M., HANRAHAN P.: Light field rendering. In *SIGGRAPH* (1996), pp. 31–42. 2

[LLW06] LAM P.-M., LEUNG C.-S., WONG T.-T.: Noise-resistant fitting for spherical harmonics. *IEEE Transactions on Visualization and Computer Graphics 12* (March 2006), 254–265. 3, 4

[Mac48] MACROBERT T. M.: *Spherical harmonics: an elementary treatise on harmonic functions, with applications*. Dover Publications, 1948. 2

[MB95] MCMILLAN L., BISHOP G.: Plenoptic modeling: an image-based rendering system. In *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques* (New York, USA, 1995), SIGGRAPH '95, ACM, pp. 39–46. 1, 2

[MDA02] MASSELUS V., DUTRÉ P., ANRYS F.: The free-form light stage. In *Proceedings of the 13th Eurographics workshop on Rendering* (Aire-la-Ville, Switzerland, Switzerland, 2002), EGRW '02, Eurographics Association, pp. 247–256. 2

[MGW01] MALZBENDER T., GELB D., WOLTERS H.: Polynomial texture maps. In *Computer Graphics, SIGGRAPH 2001 Proceedings* (2001), pp. 519–528. 1, 2, 7, 9

[MRP98] MILLER G. S. P., RUBIN S. M., PONCELEON D. B.: Lazy decomposition of surface light fields for precomputed global illumination. In *Rendering Techniques* (1998), pp. 281–292. 1, 2

[NSI01] NISHINO K., SATO Y., IKEUCHI K.: Eigen-texture method: Appearance compression and synthesis based on a 3d model. *IEEE Trans. Pattern Anal. Mach. Intell. 23*, 11 (Nov. 2001), 1257–1265. 2

[OK10] O'TOOLE M., KUTULAKOS K. N.: Optical computing for fast light transport analysis. *ACM Trans. Graph. 29*, 6 (Dec. 2010), 164:1–164:12. 2

[PML*09] PEERS P., MAHAJAN D. K., LAMOND B., GHOSH A., MATUSIK W., RAMAMOORTHI R., DEBEVEC P.: Compressive light transport sensing. *ACM Trans. Graph. 28*, 1 (Feb. 2009), 3:1–3:18. 2

[SKC03] SHUM H.-Y., KANG S. B., CHAN S.-C.: Survey of image-based representations and compression techniques. *IEEE Trans. Circuits Syst. Video Techn. 13*, 11 (2003), 1020–1037. 1

[SS95] SCHRÖDER P., SWELDENS W.: Spherical wavelets: efficiently representing functions on the sphere. In *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques* (New York, USA, 1995), SIGGRAPH '95, ACM, pp. 161–172. 2

[WAA*00] WOOD D. N., AZUMA D. I., ALDINGER K., CURLESS B., DUCHAMP T., SALESIN D. H., STUETZLE W.: Surface light fields for 3d photography. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques* (New York, USA, 2000), SIGGRAPH '00, ACM Press/Addison-Wesley Publishing Co., pp. 287–296. 1, 2, 3, 4

[Wah81] WAHBA G.: Spline Interpolation and Smoothing on the Sphere. *Siam Journal on Scientific and Statistical Computing 2* (1981). 4

[WDR11] WU H., DORSEY J., RUSHMEIER H.: A sparse parametric mixture model for btf compression, editing and rendering. *Computer Graphics Forum 30*, 2 (2011), 465–473. 2

[WDT*09] WANG J., DONG Y., TONG X., LIN Z., GUO B.: Kernel nystrom method for light transport. *ACM Trans. Graph. 28*, 3 (July 2009), 29:1–29:10. 2

[ZC04] ZHANG C., CHEN T.: A survey on image-based rendering - representation, sampling and compression. *Sig. Proc.: Image Comm. 19*, 1 (2004), 1–28. 1, 2