Effects of Approximate Filtering on the Appearance of Bidirectional Texture Functions

Adrian Jarabo, Hongzhi Wu, Julie Dorsey, Holly Rushmeier, and Diego Gutierrez

Abstract—The BTF data structure was a breakthrough for appearance modeling in computer graphics. More research is needed though to make BTFs practical in rendering applications. We present the first systematic study of the effects of approximate filtering on the appearance of BTFs, by exploring the spatial, angular and temporal domains over a varied set of stimuli. We perform our initial experiments on simple geometry and lighting, and verify our observations on more complex settings. We consider multi-dimensional filtering versus conventional mipmapping, and find that multi-dimensional filtering produces superior results. We examine the trade off between under- and oversampling, and find that different filtering strategies can be applied in each domain, while maintaining visual equivalence with respect to a ground truth. For example, we find that preserving contrast is more important in static than dynamic images, indicating greater levels of spatial filtering are possible for animations. We find that filtering can be performed more aggressively in the angular domain than in the spatial. Additionally, we find that high-level visual descriptors of the BTF are linked to the perceptual performance of pre-filtered approximations. In turn, some of these high-level descriptors correlate with low level statistics of the BTF. We show six different practical applications of applying our findings to improving filtering, rendering and compression strategies.

Index Terms—BTF, perception, filtering

◆

1 INTRODUCTION

MANY computer graphics applications require accurate depiction of visually rich material appearance. Bidirectional Texture Functions (BTFs) represent complex spatially and angularly-varying appearance, including effects such as self-shadowing, inter-reflections or subsurface scattering. BTFs are usually captured by taking photographs of a material sample under different combinations of light and view directions.

Filtering is required to render BTFs without aliasing. Mathematically exact filtering is infeasible with finite processing power and/or memory. Approximate filtering techniques are needed, with the acceptability of approximations measured by their perceptual impact. In this paper we present the first systematic study of the effects of multi-dimensional filtering on the appearance of BTFs. Our goal is to understand how approximate filtering of BTFs along the spatial, angular and temporal dimensions (both for a moving light source and a moving camera) affects the perceived visual quality of the results (see Figure 1).

Previous studies have focused on developing efficient *compression* algorithms (see [1], [2]). Recently, Filip et al. showed that different materials actually require different compression strategies [3]. Another related line of work deals with *filtering* strategies, where

---

- *A. Jarabo and D. Gutierrez are with Universidad de Zaragoza.*
- *H. Wu is with Yale University and State Key Lab of CAD & CG, Zhejiang University.*
- *J. Dorsey and H. Rushmeier are with Yale University.*

pre-computations are used for efficient rendering, by computing multiple pre-filtered representation of appearance at different viewing distances. Given the sophisticated, non-linear illumination effects stored in a BTF, relatively complex filters are employed [4], although visual appearance is usually not taken into account.

Our work complements these two lines of research: We study how different filtering strategies affect the perceived appearance of a varied set of BTFs, by means of systematic psychophysical experiments. We use a varied subset of the Bonn and UCSD BTF databases, and analyze them both globally and according to high-level descriptors of their visual properties.

Contributions Our main findings are:

- Approximate pre-filtered representations of BTFs can be used without affecting visual equivalence with a multi-sampled reference solution
- For static scenes, there is a great tolerance for aliasing in the spatial and angular domains, which is preferred to over-blurred BTFs. For dynamic scenes, this trend is reversed, and over-blurred animations are preferred over temporal aliasing artifacts. These findings correlate well with known mechanisms of human perception
- The angular domain can be more aggressively filtered, while filtering in the spatial domain rapidly affects visual equivalence
- High-level descriptors of the BTFs (such as e.g. *glossy* or *structured*) correlate well with perceptually equivalent levels of filtering. This sug-
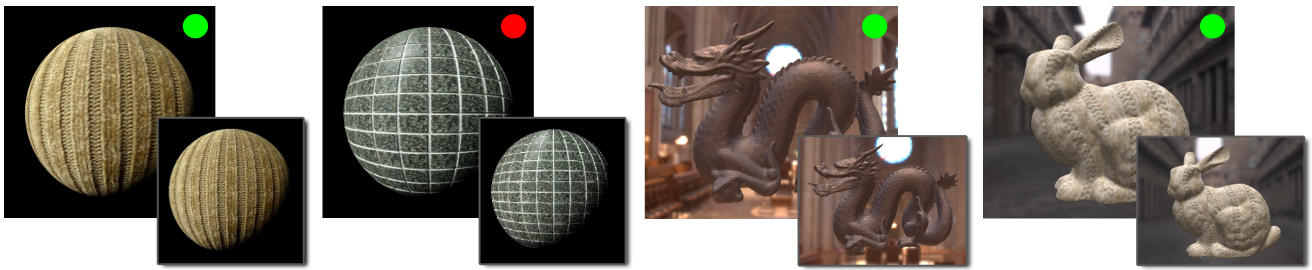
Fig. 1: We perform psychophysical experiments on simple geometries to examine filtering strategies for BTF rendering. From left to right: (1) Our experiments reveal that the filtered, aggressively undersampled representation (smaller image) of the BTF is perceived as visually equivalent to its reference, multi sampled version (larger image). (2) However, the same experiments show that the same strategy on the second sphere is not considered visually equivalent. We apply our observations to produce filtered examples of more complex scenes (3 and 4) which experiments show maintain visual equivalence.

gests the necessity and usefulness of standardized high-level descriptors in material and BTF databases. In turn, these high-level descriptors are also correlated with low-level statistics of the BTF

- We show that our findings generalize to different geometries and illumination conditions, and propose different practical applications in filtering, rendering and compression strategies

This is the first work to systematically analyze the perceptual effect on material appearance of different approximate BTF filtering strategies. While we did not consider all possible combinations of all parameters, we believe that our findings should provide enough traction to motivate future work, for which we plan to make our stimuli and data publicly available. We hope that this work can also inspire future compression and filtering strategies for BTFs, which in turn may lead to more sophisticated rendering and editing algorithms.

2 PREVIOUS WORK

BTF Compression Dana et al. [5] first introduced BTF as an image-based representation for material appearance. Mueller et al. [1] presented a comprehensive survey on BTF acquisition, synthesis and rendering. Several BTF compression methods were compared, but none of them took into account the perception of BTF at different viewing distances or under different light directions. Another excellent survey including subsequent work in the field can be found in [2].

More recently, Ruiters et al. [6] achieved high compression rates for single-level BTFs, by fitting a small set of basis functions based on tensor decomposition. Tsai et al. [7] further pushes the idea to a k-clustered tensor approximation, which is suitable for efficient real-time rendering of the compressed BTFs. Havran et al. [8] compressed the BTF by adopting a multidimensional conditional probability density function in conjunction with vector quantization. Mipmapping was handled by directly applying the same algorithm

to averaged BTF data. In contrast to our work, no perceptual factors were considered.

Hierarchical Representation of Appearance Various algorithms have been proposed to pre-filter material appearance, focusing on texture maps [9], normal maps [10], [11], or general geometry and BRDFs [12]. A survey for reflectance filtering algorithms was recently presented by Bruneton and Neyret [4]. Ma et al. [13] pre-filtered BTFs using Principal Component Analysis (PCA) and parametric fitting of Phong models after a Laplacian transformation. Impressive real-time LOD rendering results were achieved. However, their method is purely numerical, and no evaluation of how the BTF appearance is perceived for different levels of detail is provided.

Perception The study of visual perception in the context of computer graphics has produced many useful results [14]. Some example applications include selective rendering [15], global illumination [16], [17] and motion blur [18]. Rogowitz and Rushmeier [19] found that simplified geometries were perceived differently for static and animated cases. Ramanarayanan et al. [20] and Vangorp et al. [21] studied how object geometry, material, and illumination interact to provide information about appearance. Krivanek et al. [22] investigated the relationship between parameters of the virtual point lights algorithm with the perception of different materials.

Meseth et al. [23] evaluated the rendering quality obtained by using BTFs instead of simple 2D textures. A perception-based metric for single-level BTFs was derived in [24], in order to achieve higher compression rates. Filip et al. [3], [25] ran psychophysical tests to discard perceptually unimportant data in the BTF. Although closer to our approach, our work differs in several ways: First, we do not aim at reducing the input dataset of 2D texture images that define the BTF; instead, we focus on efficient strategies to sample such multidimensional data, and on studying the effect of pre-filtering BTFs in their perceived appearance. Second, Filip et al. use a statistical description of
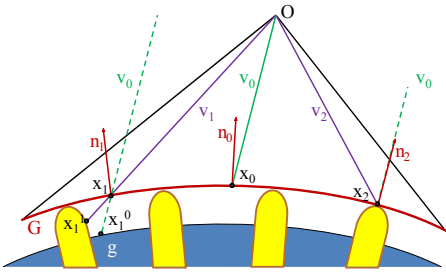
Fig. 2: This diagram shows the geometry of filtering the reflectance of surface G (red, representing the area covered in one pixel). Appearance is modeled with a BTF describing the underlying meso-geometry g (blue-yellow surface). O denotes the viewpoint (please note that we have exaggerated its proximity to the pixel for illustration purposes), while v_0 (green) is the viewing direction (vector $O - x_0$), assumed to be constant. Sampling points x_1 and x_2 from the constant or from the correct view directions will yield different results. Note that even using an orthogonal projection (i.e. constant viewing direction v_0), the angle between the viewing direction and the normal at the differential points in the surface varies according to the surface's curvature, which creates an effective non-zero solid angle.

the BTFs, whereas we rely on more intuitive, high-level material descriptors. Third, they compare results based on per-pixel visible differences between the compressed and the original data, which is not related to higher-level visual properties; instead, our work relies on the concept of *visual equivalence*, where visibly different images are considered equivalent if they convey the same impression of appearance. Our work offers applications beyond data compression, ranging from level-of-detail or filtering techniques for BTFs, to optimized rendering. Our findings could be applied in conjunction with compression techniques, resulting in good filtering of reduced data.

3 PROBLEM STATEMENT

Filtering a BTF Rendering a richly textured surface at different distances without aliasing artifacts requires high sampling rates at considerable computational cost. To address this issue, one common approach is to pre-filter the appearance of the surface [4]. However, to obtain the exact solution, we would need to compute the filtered appearance for all light and view conditions. This is impractical, due to huge storage and computation requirements. Instead, the filtered appearance is typically approximated, by pre-filtering only a subset of viewing distances, and then building a hierarchy of level-of-detail (LOD) representations.

Texture *mipmapping* is one such technique to reduce aliasing for surfaces decorated with 2D plain textures. This technique creates a hierarchy of different versions of the texture, where each level is a down-sampled

version of the previous level. When the surface is rendered, the level is chosen so that the ratio $r = t{:}p$ (texel to pixel) is preserved at 1:1. However, it is difficult to extend texture mipmapping to BTFs, since it makes assumptions about the underlying surface geometry, which may not hold for complex surface representations as BTFs. For example, one assumption is that the normal n of the surface and the light l and view v directions are uniform inside the fraction of the surface contained in each pixel. This is not valid for BTFs, as four additional directional dimensions defining l and v need to be sampled, to take into account the effect of the underlying meso-geometry.

Figure 2 shows this. The red line represents the portion of geometry G covered by one pixel, which has an underlying meso-structure g (the blue and yellow shape) modeled with a BTF. The resulting outgoing radiance of the pixel is the integral of the reflected light. However, this constant view vector will introduce errors, due to parallax and to the directional dependence of the reflectance. For point x_1 in G, the sampled point in g would be $g(x_1, v_0) = x_1^0$, yielding a wrong blue sample. Using the correct view direction v_1 we obtain $g(x_1, v_1) = x_1^1$, a correct yellow sample (parallax problem). For x_2, the sampled reflected light is likely to be different for both view directions; this is particularly problematic for glossy surfaces.

Similar problems occur when the incoming light direction l is assumed to be constant. This means that, in practice, a correct BTF filtering should not only preserve a correct ratio $t{:}p$ in the spatial domain, but also filter the view v and light l directions correctly.

Perceptual consequences of filtering BTF While necessary to avoid the artifacts described above, approximately pre-filtering the BTF might produce new, different artifacts: if the filter size is too small, aliasing artifacts might appear on the surface. On the other hand, if the filter is too large, the appearance of the surface will get blurred, decreasing overall contrast and detail.

Our main goal is to evaluate under which conditions an approximately pre-filtered BTF is considered *visually equivalent* [20], [22] to the ground-truth image. We want to explore how different kernel sizes applied on the spatial, directional and temporal domains of the BTF affect the appearance of the surface at different viewing and lighting conditions, both for static images and animation. Particularly, we are interested in the following questions:

- Is it possible to approximately pre-filter BTFs while maintaining visual equivalence with a multi-sampled reference?
- Is this perceived equivalence correlated with high-level visual properties of the surface?
- What kind of artifacts (e.g. aliasing, blur) are more easily accepted by the human visual system? Under what conditions?

- Can different sampling strategies be applied to the different domains of the BTF? What is the interplay between domains?
- How do different filtering kernels affect visual equivalence?
- Does the distance to the camera (i.e. projected area in the pixel) affect visual equivalence?
- Does the motion of camera and light sources affect visual equivalence?

4 OVERVIEW OF THE EXPERIMENTS

By applying different filtering kernels on the domains of the BTF (spatial and angular), we want to analyze how each domain affects the appearance of the filtered BTFs, as well as the interplay between them. We have designed three different experiments: The first one evaluates the perception of filtered BTFs on static images, seen at different distances and under different illumination directions. The second one analyzes the effect of a varying illumination vector l, while the third analyzes the effect for variations in the view vector v.

For all the experiments, we use a simple scene (rendered at 512×512 pixels) consisting of a sphere, with material appearance modeled using a BTF, illuminated by a single directional light. We choose this setup following the work by Filip et al. [3], whose results suggest that simple geometry with directional illumination is less forgiving than more complex geometry and illumination. Vangorp et al. [21] found that simple objects like an sphere are actually not appropriate for depicting material. However, the authors only explore reflectance and explicitly leave spatially varying effects (e.g. textures or BTFs) as a future line of research. On the contrary, the work by Filip et al. focuses on the particular case of BTF's, so their findings are better suited for our research. Generalization to different object shapes and illuminations is later evaluated in Section 8.

We describe here common aspects of all the experiments. We also show how our multidimensional filtering strategy has advantages over classic (spatial) mipmapping by means of a pilot study. Last, we give specific details of each experiment in subsequent sections.

Input BTFs We use the Bonn [26] and UCSD [27] databases. These BTFs present a low to moderate angular frequency, restricted by the capabilities of the acquisition devices. This in turn limits the types of materials that can be represented using BTFs. Therefore, we limit our study to the materials that BTFs excel at representing. We use captured BTFs since they are the most common and used in practice. We visually analyzed high-resolution multisampled renderings of several BTFs in order to detect pixel-wise registration inaccuracies, and discarded unsuitable ones. In the end, we keep sixteen different BTFs (see Figure 3).
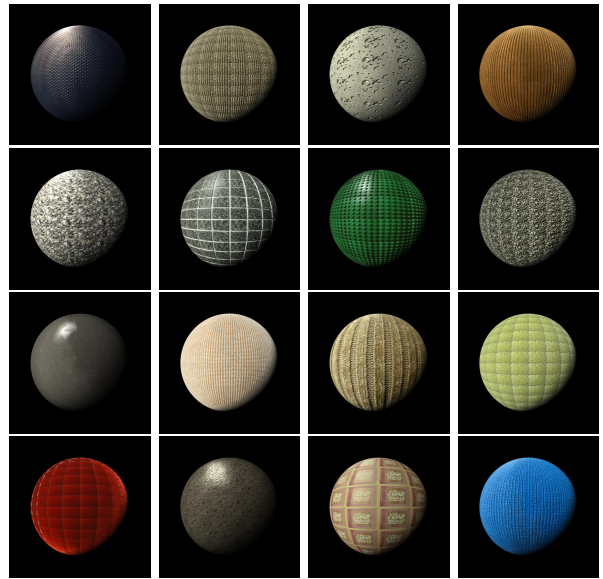

Fig. 3: Reference renderings of the stimulus BTFs. From top to bottom, and from left to right: *Cambrils*, *Carpet*, *Ceiling*, *Corduroy*; *Floortile*, *Impala*, *Lego*, *Lichen*; *Pinktile*, *Proposte*, *Pulli*, *Sponge*; *Velvet*, *Walkway*, *Wallpaper* and *Wool*.
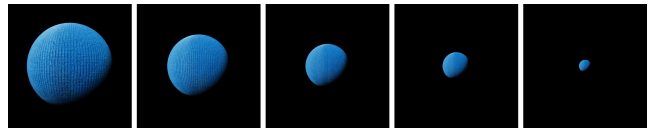

Fig. 4: Sphere viewed under the different distances used in Experiment 1, from left to right d_0, d_1, d_2, d_3, and d_4. The sphere is rendered using BTF *Wool* iluminated with light direction l_1.

Each one is made up of up to 151×151 images with a resolution up to 256^2 pixels. They have been chosen to represent a wide range of surfaces, showing different levels of complexity, both in reflectance and mesogeometry. We use uncompressed versions of the BTF, to avoid artifacts. Since these BTFs represent very different materials, we classify them according to a set of high level properties describing their appearance, by means of a pilot study performed by ten participants. These properties are based on previous work on texture [28] and BRDF [29] classification. The purpose of this descriptor-based classification is to analyze the results both globally, and at descriptor level. Table 1 shows the result of the pilot study. More details and the individual description of the BTFs are included in the supplementary material.

Filtering kernels We test three different filters in our experiments: *box*, *Gaussian* and *Lanczos*. The box filter is widely used, since it is the simplest and most efficient; additionally, it is implemented by default in graphics hardware, although it has known poor performance with high frequencies [4]. The Gaussian filter usually behaves better, and offers a good com-

BTF	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Cambrils	○	○	●	○	●	○	○	●	○	●	●	○	○	○	○
Carpet	○	○	●	●	●	○	●	○	○	○	○	●	○	○	○
Ceiling	○	○	○	●	○	●	○	●	●	○	○	○	●	○	●
Corduroy	○	○	●	○	○	●	○	●	●	○	○	●	○	○	○
Floortile	○	○	○	○	●	●	○	○	○	●	○	●	○	○	○
Impala	●	○	○	●	●	○	○	●	●	○	○	○	○	○	●
Lego	○	○	●	○	●	○	○	●	○	●	○	●	○	○	●
Lichen	●	○	○	●	●	○	●	●	●	○	○	○	○	○	●
Pinktile	●	○	●	○	○	○	●	○	○	○	○	○	○	○	○
Proposte	○	●	○	●	○	○	○	●	○	○	○	●	●	○	○
Pulli	○	○	●	○	●	●	○	●	○	○	○	●	●	●	○
Sponge	○	●	○	○	●	○	○	●	○	○	○	●	●	●	○
Velvet	○	○	○	●	○	○	●	○	○	○	○	●	●	○	○
Walkway	○	●	○	●	●	○	○	●	○	●	○	●	○	○	○
Wallpaper	○	●	●	○	●	○	○	○	○	○	○	●	●	○	○
Wool	○	○	●	○	●	○	○	●	○	●	○	●	○	●	○

TABLE 1: BTFs with their tagged descriptors, according to our pilot study. Green ● and red ○ circles indicate whether the property has been associated or not to the BTF. The properties tagged are: (1) *high-contrast*, (2) *granular*, (3) *structured*, (4) *rough*, (5) *feature-dense*, (6) *complex-structure*, (7) *flat*, (8) *relief*, (9) *sharp-relief*, (10) *smooth-relief*, (11) *glossy*, (12) *color*, (13) *light*, (14) *soft*, and (15) *hard*.

Fig. 5: Left: Incoming light direction ω_i at each point in a texel varies from ω_c in the middle of the texel. Middle: To solve this, each level of the spatially filtered mipmapping hierarchy is extended to contain several filtered versions computed for different solid angles Ω. Right: Radiance in texel t for solid angle Ω_0 is obtained by sampling the immediate lower level texels in the hierarchy t_{i-1}^1 and t_{i-1}^2 at directions ω_c^1 and ω_c^2 respectively.

promise between final result and cost. Finally, the Lanczos filter is the finite filter that best models the ideal *sinc* kernel, but is the most computationally expensive. The implementation and parameters of the kernels follow PBRT [30]. We use isotropic filtering for the three cases. Anisotropic filtering would improve the results, so using isoptropic filtering is a *worst case* scenario. Additionally, it keeps the number of explored dimensions tractable.

Pre-filtering the BTFs If we opt for pre-filtering only in the spatial domain of the BTF, each texture image used to represent the BTF is mipmapped from a spatial resolution of 256×256 down to 1×1, thus yielding a hierarchy of nine levels. Each scene is rendered from five different distances d_i. The closest distance d_0 is chosen so that $r = 1$ (the ratio texel to pixel) in the center of the rendered image of the sphere. Successive distances are set so that r at distance d_i is 2^2 times the ratio at d_{i-1}. Thus, for d_1 we have $r = 4$ and for the final d_4 we have $r = 256$ (Figure 4). Distances d_i vary directly with view vector v; in the following we use d_i as the distance where the objects are viewed, and v as the view vector.

Filtering only in the spatial domain in BTFs might lead to incorrect appearance, as described in Section 3 and shown in Figure 5 (left). Let Ω_0 be the effective solid angle subtended by a pixel and the light source, which accounts for the variation of the angle between the light direction and the normal in the points of the surface (see Figure 2). Assuming that all incoming light on a pixel is from the same direction ω_c (sampled at the center) is wrong, since different points receive light from different directions ω_i. Since these directions depend on Ω_0, we create
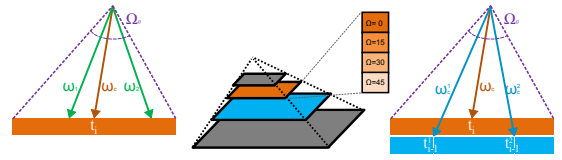
several filtered versions for each level of the mipmap hierarchy, varying it from $0°$ to $45°$ in 15-degree increments (Figure 5 (middle)). We found that four levels are enough for our BTFs. Then, to accurately compute incoming light at a given filtered texel t_i, we take into account the immediate lower level in the hierarchy t_{i-1}, and sample its center point instead (Figure 5 (right)). The same approach works for the viewing direction v.

At rendering time, the pre-filtered hierarchy is accessed based on the ratio r for the spatial domain, which is computed using the derivatives of the texture coordinates at the pixel. This gives us a texel t at hierarchy level k. Then, the light and view solid angles are used to choose the pre-filtered version of t, as described above. Thus our multidimensional mipmapping hierarchy is accessed by specifying three input parameters: r, Ω_l (light) Ω_v (view). These last two are computed in run-time using the derivatives in the light and view directions with respect to the normal of the surface, which allows accounting for surfaces with a varying normal within a pixel. Note that in our experiments the normal varies smoothly within the pixel, so this variation can be captured using the pixel derivatives. For geometries with more complex geometry, level-of-detail techniques can be applied [31]. Linear interpolation between hierarchy levels is used.

Similar to the common bias term used in mipmapping, we introduce a scale factor s to guide the selection of the level in the pre-filtered hierarchy. This factor is used to scale the input parameters used to traverse the pre-filtered hierarchy r, Ω_l, and Ω_v. Intuitively, $s < 1$ means under-sampling the BTF, while $s > 1$ means blurring it. We note that two key characteristics that make BTF filtering interesting are the different nature of the spatial and angular domains, and the interplay between them. We thus define three different scales: s_x for the spatial domain, and s_l and s_v for the light and the view directions in the angular domain, respectively. In our experiments, we set $s_x \in S_1 = \{.25, .5, 1, 2, 4\}$. For the angular do-
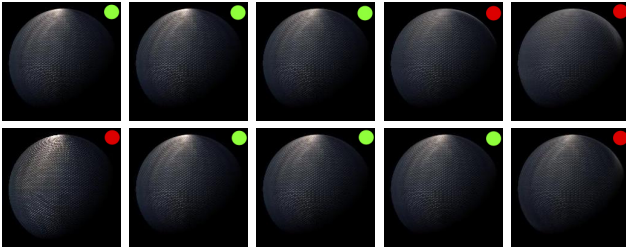
Fig. 6: Example of filtering in the angular domain. The BTF *Cambrils* is filtered varying s_v (top) and s_l (bottom) with values (from left to right) .25, 1, 2, 8, and 32, at distance d_2 and illuminated from direction l_1. The inset dots means whether the image has been considered visually equivalent to a reference (green) or not (red), based on a pilot study.

Fig. 7: Sphere illuminated from the four light vectors l_i used in Experiment 1, from left to right l_1, l_2, l_3 and l_4. This example shows the sphere rendered using the BTF *Pulli* with scale $s = 1$ at distance d_2.
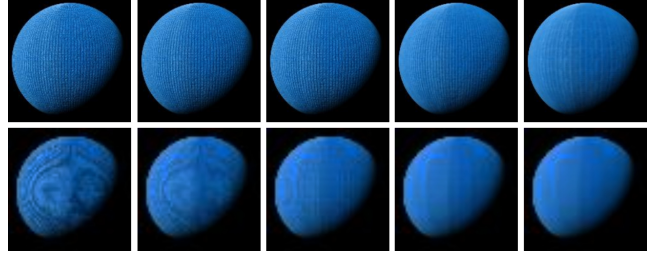

Fig. 8: Example of filtering all domains with same scale s. The BTF *Wool* is shown at distances d_2 (top) and d_4 (bottom), filtered varying s with values (from left to right) .25, .5, 1, 2, and 4, illuminated from direction l_1.

main, we set s_l and s_v to $S_2 = \{.25, .5, 1, 2, 4, 8, 16, 32\}$, based on an initial exploratory pilot study. Figure 6 shows an example of the *Cambrils* BTF under varying parameters of s_v and s_l. To be able to explicitly explore angular filtering while keeping the experiment tractable, we only vary one of the scale factors each time, while fixing the other two at 1 (no filter size biasing). Additionally, we also explore the interplay between domains by analyzing all five cases where $s_x = s_i = s_o = s$. This yields a total of 23 scale combinations.

Comparison with mipmapping A first concern is whether the proposed multidimensional filtering has advantages over classic, simple mipmapping in the spatial domain, which assumes that the normal, light and view directions remain constant within the pixel. We evaluate this by means of a pilot study, run on sixteen participants in our laboratory. The study followed a randomized two-alternative-choice (2AFC) design, similar to the used subsequently in Experiment 1 (Section 5). The user is shown test pairs consisting of an image filtered with either the multidimensional filtering or classic mipmapping, and a multi-sampled ground truth, and he has to select which of them represents a reference material more accurately. The results show that our multidimensional filtering performs significantly better ($F = 28.8$, $p < 0.01$) with a significance level of 99%, so it is indeed a better strategy for BTFs than mipmapping. Details about this pilot study can be found in the supplementary material.

Mechanical Turk Similar to many previous studies in computer graphics [32]–[35], we use Amazon Mechanical Turk (MTurk) as a source of participants for the experiments. A reasonable concern when using MTurk as a source of participants in user studies is the possible effect that uncontrolled viewing conditions (such as display resolution, brightness or environment light) may have in the data. On the other hand, using MTurk allows to have a significantly larger number of participants, much larger than what can be achieved

in controlled lab sessions, which reduces variance. Heer and Bostok [36] showed that MTurk can actually be used in visual psychophysical experiments obtaining valid results; they replicated classic perception experiments in MTurk and compared the results with those from controlled experiments, finding a good match between the two sources of data. Nevertheless, we wanted to check the reliability of our MTurk data. We thus repeated the main experiments described in this paper under controlled conditions in our lab, and found that the results are indeed consistent. Details and analysis can be found in the supplementary material.

Data Analysis We seek to establish a threshold where the visual equivalence between the reference and a given stimuli is lost, either due to aliasing or excessive blurring. Following previous work [20], [22], we cut through the data, using a conventional 75% 2AFC threshold value, where 50% is pure guessing. Additionally, to study the effect of each variable and to find significant trends in the data, we use N-ways Analysis of Variance (ANOVA), focusing on main effects and interactions between variables. Significant effects are further analyzed by using a Tukey-Kramer post-hoc analysis, to detect potential false-positives [37].

To obtain more meaningful analysis, we also analyze the effect of the high-level appearance descriptors explained in Section 4. In the following sections we describe each experiment, and summarize our main findings derived from them; the complete analysis can be found in the supplementary material.

5　EXPERIMENT 1: STATIC SCENE

We first take into account both the spatial and angular domains on static images.

5.1　Description

Stimuli The stimuli used are the result of combining our sixteen BTFs rendered at four different distances d using all 23 possible combination of scales $s = < s_x, s_i, s_o >$, for each of the three filtering kernels. To account for the light vector, we include four different incoming directions $l_i = \{\phi_i, \theta_i\}$, chosen to light the sphere from significantly different angles and offering a wide range of depictions. In particular, we take: $l_1 = \{-45°, 45°\}$, $l_2 = \{0°, 22°\}$, $l_3 = \{30°, 0°\}$ and $l_4 = \{100°, -28°\}$ (Figure 7). We define $\{\phi = 0, \theta = 0\}$ as the direction from the camera to the center of the sphere.

This makes a total of 17664 stimuli images. To handle such large number, we split the experiment into four smaller parts. In Part 1 we use the same scale for all domains ($s = s_x = s_i = s_o$), only one light direction (l_1), and the box filter, making a total of 320 test images. An example for BTF *Wool* and distances d_2 and d_4 is shown in Figure 8; the full set of stimuli can be found in the supplementary material. In Parts 2 and 3 we explore the light direction l and the filter kernel respectively, so we reduce the number of images in other dimensions. First, we reduce the number of BTF by clustering the original sixteen according to the results of the Part 1 (Section 5.2), and taking representatives from each cluster. This yields eight BTFs: *Cambrils*, *Corduroy*, *Impala*, *Proposte*, *Pulli*, *Velvet*, *Wallpaper* and *Wool* (please refer to the supplementary material for the full description of the clustering process). Additionally, we only use distances d_2 and d_4. Although no significant effect on d was found in Part 1 of the experiment, we opt for a somewhat conservative approach and analyze close/medium and long views. This makes a pool of 320 and 240 images respectively.

Finally, in Part 4 we explore the effect of scale s in the spatial and angular domains *independently*; as argued before, this is of particular interest for BTF filtering. We vary the scale factor in one domain at a time, while keeping the other two set to 1 (no scaling) (see Section 4). We again use the reduced stimuli from Parts 2 and 3 (eight BTFs, two distances d), using only the box filter. Since no significant effect with light direction l was found in Part 2, we only use two light directions l_1 and l_3. This gives us a total of 608 stimuli.

Experimental procedure A total of 350 subjects took part in the first part of the experiment (236 male, 78 female; some participants did not report gender), 650 in the second (336 M, 233 F), 250 in the third (90 M, 140 F), and 625 in the fourth (300 M, 234 F), with ages between 18 and 64. All of them reported normal acuity and color vision; some of them had a computer
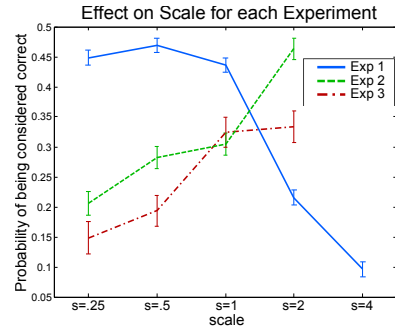

Fig. 10: Main effects found for scale s in the results of Experiment 1 (Section 5), Experiment 2 (Section 6), and Experiment 3 (Section 7). In static scenes, visual equivalence drops significantly for over-blurred BTFs ($s > 1$). However, when introducing dynamics in the stimulus (i.e. moving light or camera), blur results into higher visual equivalence, in contrast with visual equivalence for under-sampled BTFs ($s < 1$). In the x-axis are the filter scales s, while the y-axis represents the visual equivalence for each scale.

graphics and/or artistic background. The minimum required screen resolution was 1024×768 pixels. A written explanation was given at the beginning of each experiment, although none of the participants were familiar with the final goal.

The experiment followed a fully randomized two-alternative-forced-choice (2AFC) design. Each participant was shown an average of around 20 tests, with a time limit of 20 seconds to answer each one. Three images are shown simultaneously, depicting a sphere with the same BTF. On the top, a multi-sampled close view of the sphere is shown as reference (rendered with super-sampling and jittered anti-aliasing to avoid artifacts, using a Gaussian filter for reconstruction); on the bottom, two test images from the stimulus set are shown. One shows the ground-truth, multi-sampled render of the BTF, while the other shows a pre-filtered representation, rendered with one sample-per-pixel. The position (left or right) is randomized. The two images are rendered from a more distant point of view, and under a different illumination than the reference, to avoid matching tasks in image space. We apply an antialiasing mask to the edges of the spheres. The subject is asked: *"Which image represents the reference material on top more accurately?"*. We randomly introduced a few control tests with clear artifacts in one of the images, and discarded the results from participants who failed to provide the right answer to *all* of them, keeping in the end about 90 % of the total.

5.2　Results

Figure 9 (a) summarizes the results for the BTFs categorized with each property for the same scale for all dimensions; we further discuss each domain
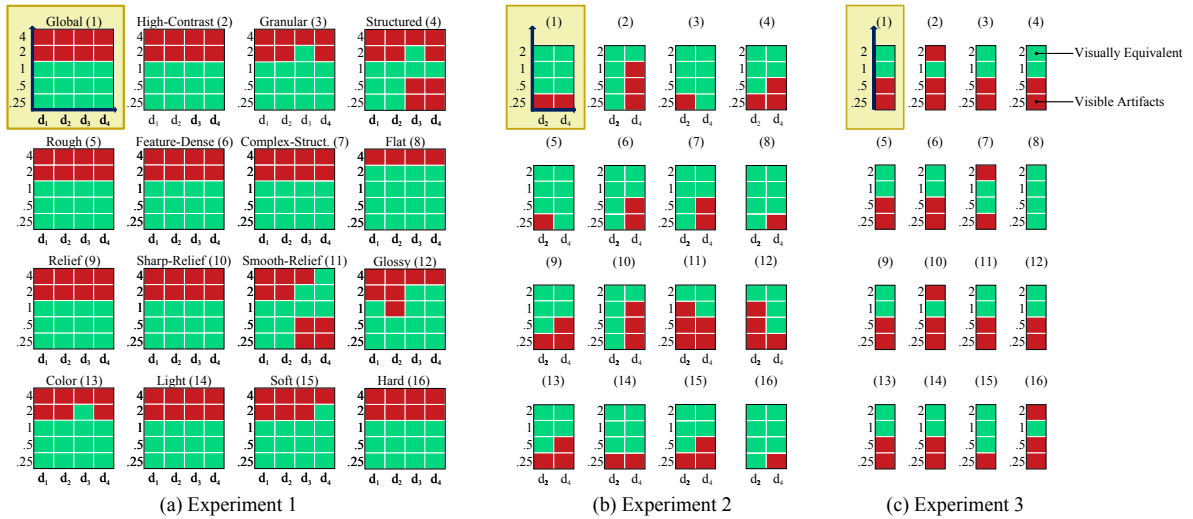
Fig. 9: Results of (a) Experiment 1 (Section 5), (b) Experiment 2 (Section 6), and (c) Experiment 3 (Section 7), globally and for each property studied. Green means that visual equivalence is found, while red means that visible differences in form of artifacts are detected. For Experiment 1 and Experiment 2, each graph shows the effect on the visual equivalence for the combination of scales s (y-axis) and distance d (x-axis). For Experiment 3, visual equivalence depends only on the scales s (y-axis).

separately later in the section. Green indicates visual equivalence with the reference for a given distance d and scale s (75% 2AFC threshold), while red indicates visual differences in form of artifacts, globally and for each visual property used to categorize the BTFs. For the second and third parts of this first experiment, no significant differences were found between light directions l ($F = 1.34$, $p > 0.25$), nor between different kernels ($F = 1.32$, $p > 0.26$). A significant interaction effect between l and d was found ($F = 2.03$, $p < 0.05$), but the post-hoc test showed no significant differences between pairs. Finding no effect on the kernels used is surprising: one might expected that both the *Lanczos* and *Gaussian* kernels would performed significantly better than *box* filter. However, we found in this experiment that filter type did not affect the visual equivalence of BTFs. This is interesting when designing efficient rendering strategies.

A significant trend is found in scale s ($F = 198.05$, $p < 0.05$): under-sampled ($s < 1$) BTFs are overall considered visually equivalent to the reference, while over-blurred ($s > 1$) BTFs are, in most cases, found non-equivalent ($s = 4$ in particular is almost never considered visually equivalent). This indicates that aliasing on the surface appearance is preferred for static images. Blurring reduces contrast, and it is well-known that the human visual system is especially sensitive to contrast, which may explain this result. A two-way interaction between filter scale s and distance d ($F = 5.22$, $p < 0.05$) shows that this trend is consistent across all distances tested (Figure 11).

Globally, no effect was found between distance d and visual equivalence ($F = 0.43$, $p > 0.73$). This is surprising, since it was expected that larger distances
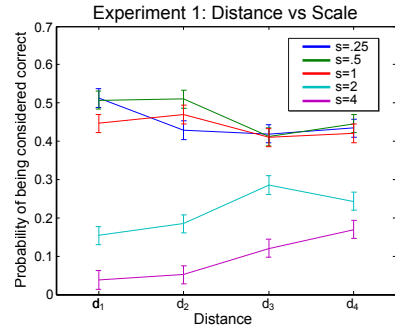

Fig. 11: Interaction effect between d and s in Experiment 1. The trends shown that the even when d increases, under-sampled BTFs ($s < 1$) still perform better than over-blurring ($s > 1$) for our range of distances tested.

would be more forgiving. However, if we focus on individual properties, it can be seen that in *structured*, *repeatable* or *regular* BTFs when distance is increased, over-blurred BTFs are preferred over under-sampled, probably because artifacts in repeated structured patterns are easier to spot, in contrast with BTFs with *complex structure*, where aliasing is preferred.

BTFs with *relief* are significantly harder to filter than *flat* BTFs ($F = 34.56$, $p < 0.05$). This was intuitively expected, but our tests confirm the statistical significance of the difference. A closer look reveals another interesting finding: for *sharp* relief, under-sampling ($s < 1$) is always preferred, while blur ($s > 1$) is considered non-equivalent at any distance d; in contrast, for long distances, blur is preferred in BTFs with *smooth* relief.
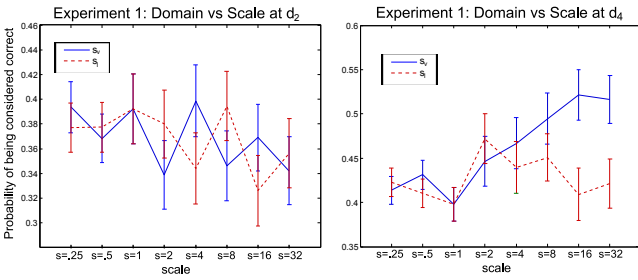
We found a significant difference on sampling each

Fig. 12: Effect of varying the scale on each of the two angular domains of the BTF (s_v, s_l) for distances d_2 (left) and d_4 (right) in static scenarios. Each domain is scaled independently, while fixing the other two domains to $s = 1$ (i.e. $s_i = 1$ is common for all domains).
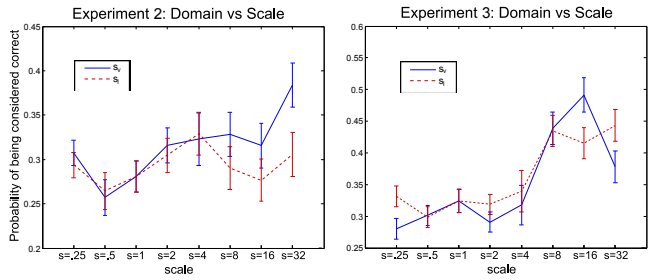

Fig. 13: Effect of varying the scale on each of the two angular domains of the BTF (s_v, s_l) for moving the light source (left) and the camera (right). Only the scale of one domain is fixed at a time, while fixing the other two domains to $s = 1$ (i.e. $s_i = 1$ is common for all domains).

domain separately: varying only the scale in the spatial domain s_x rates significantly lower than varying the scale of the angular dimensions ($F = 77.71$, $p < 0.05$), but do not present significant differences with scaling all domains at the same time (i.e. $s_x = s_v = s_l$). This means that artifacts in the spatial domain due to under-sampling or blur are more noticeable than in the angular domain, which can be subsampled more aggressively. Focusing on the angular dimensions, the results shows that errors in the light domain are easier to spot than in the view domain. A post-hoc test shows that this effect is dependent on the distance d to the object: for close-medium distance d_2 no significant trend is observed when varying s_v and s_l (Figure 12 (left)), whereas for longer views the view dimension can be further subsampled (Figure 12 (right)).

6 EXPERIMENT 2: TIME-VARYING ILLUMINATION

Experiment 2 analyzes the effect of dynamic illumination conditions, with the light vector l orbiting around the sphere.

6.1 Description

Stimuli Adding a new parameter again makes the complete set of stimulus too large. Additionally, each test is now longer, since it involves animations. We thus perform additional simplifications in our multidimensional parameter space, based on the results of Experiment 1: we start from the reduced stimuli from Part 2 and Part 3 of Experiment 1 i.e. eight BTFs times two distances d. Since no significant effect on the filtering kernel was found (see Section 5.2), we only use the box filter, because it represents the theoretical worst-case scenario. We analyze the 23 different combinations of $< s_x, s_l, s_v >$, making a total of 368 animations.

Experimental procedure A total of 517 (320 M, 160 F) different participants took part, with ages between 18

and 64. The experiment was carried out under similar conditions as Experiment 1. Training tests were presented prior to the experiment. Each test shows two animations (ground-truth and filtered) of the light vector l rotating around the sphere. No reference animation was shown, since three videos at the same time would be too confusing. The participants are given 30 seconds to answer each test. The video is played in loop mode until the user answers or time-out is reached. We ask: *"Which scene better represents the material without artifacts?"*. A brief explanation of what we mean by *artifacts* is given at the beginning. Since we rely on a high-quality video reproduction, we add an additional confidence test: if a participant reaches time-out more than 3 times, we conservatively assume poor playback quality and reject the participant's data. In the end we kept the responses from about 80% of the participants.

6.2 Results

As before, we first focus on scaling all dimensions at the same time, and then we discuss each domain separately. Our results show that visual equivalence can be achieved in at least one pre-filtered representation for all BTFs (Figure 9 (b)). A significant effect was found for scale s ($F = 34.39$, $p < 0.05$, see Figure 10): Larger scales obtain better rates, which means that artifacts produced by undersampling ($s < 1$) the light domain become too noticeable. Even at close distances, over-blurred BTFs are considered visually equivalent. This is the contrary to the static case analyzed in Experiment 1, where contrast was the main feature to be preserved.

However, analyzing the individual properties of the BTFs, some other interesting properties can be inferred from our results: On most high-frequency BTFs where oversampling may blur out recognizable patterns (i.e. properties *high-contrast*, *complex-structure*, *sharp-relief* and *feature-dense*), some spatial aliasing is preferred at close distances.

Analyzing filtering each BTF domain separately, we find that, again, the angular domains can be filtered with values of s_v and s_l higher than the spatial scale s_x, confirming that artifacts in the spatial domain are more easily spotted by a human observer than in the angular domain. We found no significant effect between the different angular dimensions ($F = 2.66$, $p > 0.1$), nor in the interaction effect between the dimensions and the scale ($F = 0.17$, $p > 0.77$). However, an interesting trend can be observed for s_v (Figure 13 (left)), where over blurring in the view domain $s_v = 32$ rates significantly higher than the baseline $s_v = 1$.

7 EXPERIMENT 3: TIME-VARYING VIEW

Last, we study how the perception of pre-filtered BTFs is affected by changes in the view vector v.

7.1 Description

Stimuli We use the eight BTFs selected previously, rendered using the scales s_i from Experiment 2 and the *box* kernel for filtering, and render an animation of a receding sphere, from d_0 to d_4. Since in Experiment 1 no significant effect was found on light direction l, we reduce the directions used to two, l_1 and l_3. This makes again a total of 368 animations.

Experimental procedure A total of 545 (346 M, 168 F) participants took part of the experiment, with ages between 18 and 67. The videos are shown in loop mode, with the camera moving back and forth perpendicular to the screen, so that the projected area of the sphere decreases linearly. Again, no ground-truth animation is shown. The rest of the experiment (procedure and question asked) is similar to Experiment 2. We kept 75 % of the participants in this experiment.

7.2 Results

Figure 9 (c) summarizes our results for $s = s_x = s_v = s_l$. A clear pattern emerges: undersampled pre-filtered representations ($s < 1$) are not considered visually equivalent to the reference. This indicates that temporal aliasing artifacts due to varying the view vector v are the most disturbing. This holds true except for surfaces with very high-frequency appearance (*high-contrast*, *complex-structure* or *sharp-relief*), which confirms our findings in our previous experiment: The detailed appearance of these BTFs is over-smoothed, and the exact filter size should be used.

Comparing these results with Experiment 1, the visual equivalence of pre-filtered BTFs drops significantly (Figure 10). Under-sampled representations, which were considered equivalent in the static scenario, are now viewed as non-equivalent for most tested BTFs ($F = 34.29$, $p < 0.05$).

Again, we observe no significant differences between scaling all parameters and scaling only in the spatial domain s_x ($F = 1.33$, $p > 0.26$). However, there is an interesting trend as s_v and s_l increase, where the visual equivalence of the pre-filtered BTF also tends to increase (Figure 13 (right)). This again suggests that over-blurring the angular domain of the BTF is preferred by the human observer.

8 GENERALIZATION

We now show how our results generalize to more complex geometry and illumination configurations. Additionally, we discuss an automatic method to assign properties to the BTFs, based on statistical analyses.

Geometry and illumination We have conducted an additional study with new geometries and illumination conditions. We follow the work of Ramanarayanan et al. [20], and add two new geometries of increasing complexity: the Stanford Bunny and Dragon. For illumination, we introduce natural illumination modeled with the *Uffizi* and *Grace* environment maps from [38]. We choose these two maps as representatives of low- and high-frequency illumination, respectively. The generalization study was performed using the same procedure as Experiment 1, with two distances (d_2 and d_4) and the same scale s for the three domains. We compare the ground truth with multiple samples per pixel against a prefiltered version, with just one sample per pixel. From the intersection point of that sample 2048 rays are thrown over the hemisphere to compute the illumination from the environment map. We found no significant differences from using different geometry ($F = 0.97, p > 0.37$) or different illumination ($F = 1.54, p > 0.21$), so we conclude that our findings generalize well. Details and example stimuli for the BTF *Cambrils* can be found in the supplementary material. Figure 14 illustrates this generalization: the BTFs shown are rendered using filtering scales $s \neq 1$ for distances d_2 and d_4, which were found to be visually equivalent in Experiment 1 (see Section 5). Figure 15, produced with the proposed rendering algorithm explained later in this section, shows additional evidence.

Statistical analysis In this paper we have used high-level descriptors of the BTFs. Adding such tags would be a simple, non-taxing step in future acquisition protocols. However, for existing untagged BTF databases, automatic classification is also interesting. As a first step towards such automatic model, we have explored how a set of low-level statistics correlate with our categorization and our results. These statistics include common first-order image statistics [39] and textural features [40] on both the spatial and the angular domains, plus the BTF statistics proposed by Filip et al. [3]. We have measured the correlation between

Fig. 14: Examples of equivalent pre-filtered representations of BTFs for different geometries under natural illumination. The larger image shows the multi-sampled reference image at close distance, while the smaller images show the visually equivalent, pre-filtered representations for distances d_2 and d_4, and scale $s = .25$. From left to right: *Wallpaper*, *Cambrils* and *Pulli*.
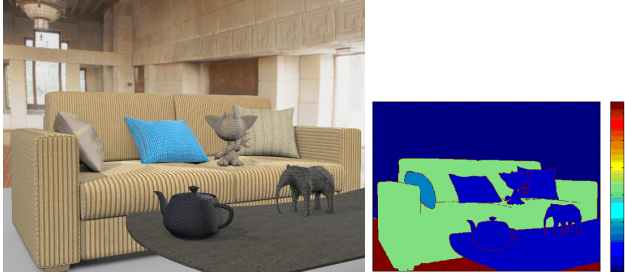

Fig. 15: Example of the application on adaptive rendering using our findings to guide sampling. For each ratio $texel : pixel$ defined by the projection of the pixel in the BTF, the visual properties of the BTFs are used to guide the number of samples (left) on the surface, based on the results summarized in Figure 9. The base samples-per-pixel spp_0 is 128. Note that the sampling is guided only by the visual properties of the surface; surfaces with no tagged descriptors (e.g. the floor) receive spp_0 samples-per-pixel (spp). Using this approach we use an average of 51.2 spp, as opposed to the 128 spp needed without using any adaptive scheme.

these statistics and the results from the pilot categorization study using Pearson correlation [37], which evaluates the linear association r between the two variables tested, with $r \in [-1, 1]$ ($r = 1$ being perfect linear correlation, $r = -1$ perfect inverse correlation, and $r = 0$ no correlation). We only keep correlations with a significance of 95% ($p < 0.05$). We have found several strong correlations ($|r| > .8$) for some properties. For instance, between *Granularity* and the directionality in the angular domain (kurtosis and skewness of the histogram of directions), between *Glossiness* and the kurtosis and skewness of luminances, or between *Structureness* and the directionality statistics in the spatial domain. Please refer to the supplementary material for the complete description of the statistics analyzed, and the full list of correlations found. While finding a complete and robust set statistics that match higher-level perceptual properties of BTFs is outside the scope of this paper, we believe these preliminary findings show great promise, while opening a new avenue of interesting future work.

9 PRACTICAL APPLICABILITY

We propose a set of practical applications to improve the performance of rendering and compression of BTFs without compromising their final appearance.

Filtering kernels The first immediate application of our study has to do with the choice of filtering kernel: No significant effect was found from using different kernels in static scenes. This means that the cheapest filter kernel (box filter) can be used in several scenarios. In our implementation, this is more than 8 times faster than Gaussian and Lanczos filters.

Rendering For all properties considered in all different scenarios tested (static, moving light and moving camera), there is at least one pre-filtered representation of the BTF (rendered with just one sample) that is visually equivalent to a multi-sampled reference (see Figure 9). This has direct applications in rendering, since it means that the isotropic pre-filtering proposed in Section 4 can be used with just one sample per pixel, achieving visual equivalence to a multi-sampled reference. Figure 16 shows side-by-side comparisons between our pre-filtered versions and the multi-sampled ground truth, rendered using ray-tracing. While not identical pixel-wise, our pre-filtered versions are considered visually equivalent, and have been rendered with a speed-up factor of up to $61.1\times$ in the most extreme case, and between $6\times$ and $9\times$ on average (see Table 2). Additional results and comparisons, including timings for each of them, can be found in the supplementary material.

Another rendering application lies in the context of adaptive rendering schemes, based on the visual properties of BTFs. It works as follows: once the ray hits a surface shaded with a BTF, we can access the visual properties defining the underlying BTF (manually tagged, or explored through statistical analysis as explained above). With these properties, and the value d obtained using the texel footprint in the pixel (computed with the derivatives of the ray in the pixel), we obtain the minimum scale s that achieves visual equivalence (Figure 9). This scale s determines the number of samples per pixel spp_p needed to get good results at pixel p ($spp_p = s * spp_0$, with spp_0 a predefined, baseline samples-per-pixel value). Following
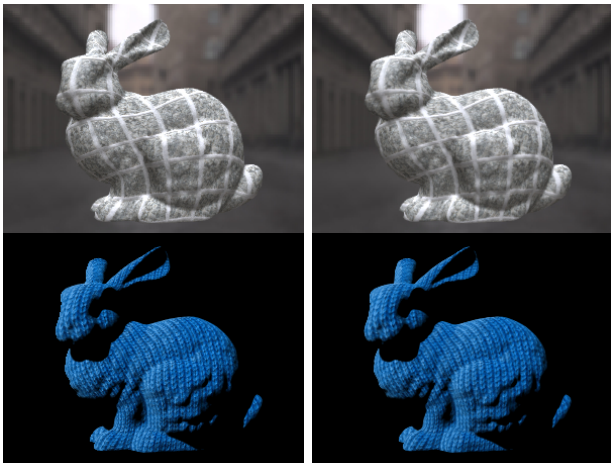
Fig. 16: Comparison between a pre-filtered BTF (left) and its multi-sampled ground truth (right) at distance d_2. Top: *Impala*. Bottom: *Wool*.

Image	PF Time	MS Time	Speed-Up
Wallpaper	6'47''	56'32''	$\times 8.33$
Cambrils	7'26''	1h05'01''	$\times 8.74$
Pulli	2'24''	16'15''	$\times 6.77$
Impala	6'47''	56'57'	$\times 8.39$
Wool	0'1.8''	1'10''	$\times 61.1$

TABLE 2: Rendering times for the images in Figure 14 and Figure 16, both using the pre-filtered representation (PF) and the multi-sampled ground truth (MS). Image *Wool* is illuminated with one light sample per pixel; this makes its speed up increase significantly.

the results from Experiment 1 (Section 5), we can reduce the number of samples up to a factor of 16 in static images. Figure 15 shows an example of a scene rendered using this adaptive approach. Note that this scheme only adapts the sampling to the surface being rendered, taking into account the surface properties. Other adaptive sampling schemes, such as sampling complex illumination patterns, are not handled by this scheme.

Our third rendering application involves BTF level-of-detail techniques such as the work by Ma et al. [13]. The authors build a Laplacian pyramid to model the pre-filtered LoD hierarchy. This is traversed for reconstruction during rendering time, starting from the lowest frequency (most pre-filtered) and adding higher-frequencies until reaching the appropriate level of detail. Based on the results summarized in Figure 9, the traversal of the pyramid can be stopped at lower-frequency levels, reducing reconstruction costs while maintaining visual equivalence for $s > 1$.

Compression Over-blurred BTFs are in general considered visually equivalent for dynamic scenes. This has two immediate consequences in compression schemes: first, it allows us to discard high-resolution levels in the level-of-detail hierarchy, which effectively reduces memory requirements. Second, since
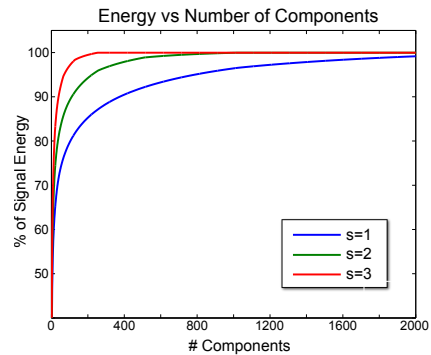

Fig. 17: Percentage of the signal energy stored by the N principal components (x-axis) of the compressed BTF, averaged for all BTFs. Each line represents the results of compressing the original BTF ($s = 1$, blue), and its overblurred versions with $s = 2$ (green) and $s = 4$ (red). We refer to the supplementary material for the individual graphs of each BTF.

low-frequencies generally contain most of the signal energy while being easier to approximate without noticeable artifacts, more aggressive compressions can be employed. As a proof-of-concept, we have implemented a common PCA-based BTF compression algorithm [26], and applied it to all our BTFs using $s = 2$ and $s = 4$. On average, using the same reconstruction error threshold, we keep 99% of the original signal energy using only 28.5% and 8% of the coefficients needed for the original version, respectively (Figure 17).

Finally, the fact that the angular domain may be aggressively filtered without decreasing visual equivalence can also be applied to compression. Blurring in the angular domain makes it possible to represent angular effects with a small number of spherical harmonic basis functions. Representation with a small number of spherical harmonic terms is also advantageous for rendering with Precomputed Radiance Transfer techniques [41].

10 DISCUSSION AND FUTURE WORK

Important conclusions can be drawn from our experiments. For static scenes, there is a clear preference for under-sampled images, which might be due to the fact that humans are very sensitive to contrast, which decreases with blurring. The trend is reversed when considering motion (both light and camera). This may be due not only to the lack of flickering artifacts, but also to motion de-blurring mechanisms in our visual system, which make moving images appear sharper than they actually are [42]. This preference seems to be related to the motion blur effect [18]. Future work needs to be conducted to examine the relationship between pre-filtering and motion blur.

We have shown that filtering in the spatial domain rapidly affects visual equivalence, while the angular

domain can be more aggressively filtered. In fact, the trend observed in the angular domain is that over-filtering is preferred over under-sampling; this supports the results from our pilot study, which shown that a multidimensional filtering produces more visually equivalent results than just filtering in the spatial domain (i.e. $s_v = s_l = 0$). In particular, at long distances the domain of the viewing vector can be filtered the most before losing visual equivalence, specially at long distances.

Another important finding is that high-level properties of the BTFs have a significant effect on how filtering affects appearance perception. For instance, under-sampling BTFs with clear symmetric structure quickly introduces artifacts that affect visual equivalence. This generalizes a conclusion of a recent study on media retargeting [33], which also identified structure as one of the most important image features that should be preserved. On the other hand, BTFs with complex high-frequency structure are in general more forgiving with under-sampling, specially for static scenes. This is in accordance with previous work [43], possibly due to the effect of visual masking (high-frequency features mask high-frequency artifacts). For BTFs with sharp relief, under-sampling $s < 1$ is preferred, even at the cost of visible aliasing. This is probably because relief is a very dominant and salient feature with strong parallax effects, which give the BTF a very three-dimensional, appealing appearance. In static BTFs with a high specular component, under-sampling is in general preferred. This makes sense since the specular highlight is usually a salient high-frequency feature. On the other hand, on animated scenarios, this behaviour is again reversed: aliasing becomes too distracting, making over-blur preferred by the observers. These findings suggest that high-level descriptors of BTFs could be used when devising an optimal sampling strategy, as we have shown in Section 9. We have found some correlation between these high-level BTF descriptors and low-level statistics, which we hope spur future work in this direction. Last, we have proposed several applications in BTF filtering, compression and rendering, directly derived from our results.

Although we have shown that our results generalize well, our work is limited by our choice of stimuli and filtering parameters. More BTFs could be added to future experiments, along with different lighting schemes, parameters, BTF descriptors and the combinations of these. Additionally, our conclusions are restricted to materials that can be accurately represented using BTFs, which are in general diffuse or moderately glossy. Although our work focuses on BTFs, we believe that some of the insights and methodology could be extrapolated to study the perception of other reflectance representations, such as SV-BRDFs, provided that the filtered surface's reflectance has moderate angular frequency. We hope that new studies and potential future applications can leverage our stimuli set, which can be accessed at http://giga.cps.unizar.es/~ajarabo/pubs/btfTVCG14/.

ACKNOWLEDGEMENTS

REFERENCES

[1] G. Müller, J. Meseth, M. Sattler, R. Sarlette, and R. Klein, "Acquisition, synthesis and rendering of bidirectional texture functions," in *Eurographics 2004, State of the Art Reports*, 2004.

[2] J. Filip and M. Haindl, "Bidirectional texture function modeling: A state of the art survey," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 31, 2009.

[3] J. Filip, M. J. Chantler, P. R. Green, and M. Haindl, "A psychophysically validated metric for bidirectional texture data reduction," *ACM Trans. Graph.*, vol. 27, no. 5, 2008.

[4] E. Bruneton and F. Neyret, "A survey of nonlinear prefiltering methods for efficient and accurate surface shading," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, 2012.

[5] K. J. Dana, B. van Ginneken, S. K. Nayar, and J. J. Koenderink, "Reflectance and texture of real-world surfaces," *ACM Trans. Graph.*, vol. 18, no. 1, 1999.

[6] R. Ruiters and R. Klein, "BTF compression via sparse tensor decomposition," *Computer Graphics Forum*, vol. 28, no. 4, 2009.

[7] Y.-T. Tsai and Z.-C. Shih, "K-clustered tensor approximation: A sparse multilinear model for real-time rendering," *ACM Trans. Graph.*, vol. 31, no. 3, 2012.

[8] V. Havran, J. Filip, and K. Myszkowski, "Bidirectional texture function compression based on the multilevel vector quantization," *Computer Graphics Forum*, vol. 27, no. 1, 2010.

[9] L. Williams, "Pyramidal parametrics," in *Proceedings of SIGGRAPH '83*, 1983.

[10] P. Tan, S. Lin, L. Quan, B. Guo, and H.-Y. Shum, "Multiresolution reflectance filtering," in *Proceedings of Eurographics Symposium on Rendering 2005*, 2005.

[11] C. Han, B. Sun, R. Ramamoorthi, and E. Grinspun, "Frequency domain normal map filtering," *ACM Trans. Graph.*, vol. 26, no. 3, 2007.

[12] H. Wu, J. Dorsey, and H. Rushmeier, "Characteristic point maps," *Computer Graphics Forum*, vol. 28, no. 4, 2009.

[13] W.-C. Ma, S.-H. Chao, Y.-T. Tseng, Y.-Y. Chuang, C.-F. Chang, B.-Y. Chen, and M. Ouhyoung, "Level-of-detail representation of bidirectional texture functions for real-time rendering," in *Proceedings of I3D '05*, 2005.

[14] A. McNamara, K. Mania, and D. Gutierrez, "Perception in graphics, visualization, virtual environments and animation," in *SIGGRAPH Asia 2011 Courses*, 2011.

[15] W. A. Stokes, J. A. Ferwerda, B. Walter, and D. P. Greenberg, "Perceptual illumination components: a new approach to efficient, high quality global illumination rendering," *ACM Trans. Graph.*, vol. 23, no. 3, 2004.

[16] K. Myszkowski, T. Tawara, H. Akamine, and H.-P. Seidel, "Perception-guided global illumination solution for animation rendering," in *Proceedings of SIGGRAPH '01*. New York, NY, USA: ACM, 2001.

[17] A. Jarabo, T. Van Eyck, V. Sundstedt, K. Bala, D. Gutierrez, and C. O'Sullivan, "Crowd light: Evaluating the perceived fidelity of illuminated dynamic scenes," *Computer Graphics Forum*, vol. 31, no. 2, 2012.

[18] F. Navarro, S. Castillo, F. J. Serón, and D. Gutierrez, "Perceptual considerations for motion blur rendering," *ACM Trans. Appl. Percept.*, vol. 8, no. 3, 2011.

[19] B. Rogowitz and H. Rushmeier, "Are image quality metrics adequate to evaluate the quality of geometric objects?" in *Proceedings of SPIE, Vol 4299. Human Vision and Electronic Imaging VI*, 2001.

[20] G. Ramanarayanan, J. Ferwerda, B. Walter, and K. Bala, "Visual equivalence: towards a new standard for image fidelity," *ACM Trans. Graph.*, vol. 26, no. 3, 2007.

[21] P. Vangorp, J. Laurijssen, and P. Dutré, "The influence of shape on the perception of material reflectance," *ACM Trans. Graph.*, vol. 26, no. 3, 2007.

[22] J. Křivánek, J. A. Ferwerda, and K. Bala, "Effects of global illumination approximations on material appearance," *ACM Trans. Graph.*, vol. 29, no. 4, 2010.

[23] J. Meseth, G. Müller, R. Klein, F. Röder, and M. Arnold, "Verification of rendering quality from measured btfs," in *Proceedings of APGV '06*, 2006.

[24] M. Guthe, G. Müller, M. Schneider, and R. Klein, "BTF-CIELab: A perceptual difference measure for quality assessment and compression of BTFs," *Computer Graphics Forum*, vol. 28, no. 1, 2009.

[25] J. Filip, M. J. Chantler, and M. Haindl, "On uniform resampling and gaze analysis of bidirectional texture functions," *ACM Trans. Appl. Percept.*, vol. 6, no. 3, 2009.

[26] M. Sattler, R. Sarlette, and R. Klein, "Efficient and realistic visualization of cloth," in *Proceedings of Eurographics Symposium on Rendering 2003*, 2003.

[27] M. L. Koudelka, S. Magda, P. N. Belhumeur, and D. J. Kriegman, "Acquisition, compression, and synthesis of bidirectional texture functions," in *In ICCV 03 Workshop on Texture Analysis and Synthesis*, 2003.

[28] A. R. Rao and G. L. Lohse, "Towards a texture naming system: Identifying relevant dimensions of texture," *Vision Research*, vol. 36, no. 11, 1996.

[29] W. Matusik, H. Pfister, M. Brand, and L. McMillan, "A data-driven reflectance model," *ACM Trans. Graph.*, vol. 22, no. 3, 2003.

[30] M. Pharr and G. Humphreys, *Physically based rendering: From theory to implementation*. Morgan Kaufmann, 2010.

[31] D. Luebke, B. Watson, J. D. Cohen, M. Reddy, and A. Varshney, *Level of Detail for 3D Graphics*. New York, NY, USA: Elsevier Science Inc., 2002.

[32] F. Cole, K. Sanik, D. DeCarlo, A. Finkelstein, T. Funkhouser, S. Rusinkiewicz, and M. Singh, "How well do line drawings depict shape?" *ACM Trans. Graph.*, vol. 28, no. 3, 2009.

[33] M. Rubinstein, D. Gutierrez, O. Sorkine, and A. Shamir, "A comparative study of image retargeting," *ACM Trans. Graph.*, vol. 29, no. 5, 2010.

[34] P. O'Donovan, A. Agarwala, and A. Hertzmann, "Color compatibility from large datasets," *ACM Trans. Graph.*, vol. 30, no. 4, 2011.

[35] A. Bousseau, J. P. O'Shea, F. Durand, R. Ramamoorthi, and M. Agrawala, "Gloss perception in painterly and cartoon rendering," *ACM Trans. Graph.*, vol. 32, no. 2, 2013.

[36] J. Heer and M. Bostock, "Crowdsourcing graphical perception: using mechanical turk to assess visualization design," in *Proceedings of the CHI '10*, 2010.

[37] D. Cunningham and C. Wallraven, *Experimental Design: From User Studies to Psychophysics*. A K Peters/CRC Press, 2011.

[38] P. Debevec, "Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography," in *Proceedings of SIGGRAPH '98*, 1998.

[39] T. Pouli, D. W. Cunningham, and E. Reinhard, "A survey of image statistics relevant to computer graphics," *Computer Graphics Forum*, vol. 30, no. 6, 2011.

[40] H. Tamura, S. Mori, and T. Yamawaki, "Textural features corresponding to visual perception," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 8, no. 6, 1978.

[41] P.-P. Sloan, J. Kautz, and J. Snyder, "Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments," *ACM Trans. Graph.*, vol. 21, no. 3, 2002.

[42] D. C. Burr and J. Ross, "Contrast sensitivity at high velocities," *Vision Research*, vol. 22, no. 4, 1982.

[43] J. A. Ferwerda, P. Shirley, S. N. Pattanaik, and D. P. Greenberg, "A model of visual masking for computer graphics," in *Proceedings of SIGGRAPH '97*, 1997.

Adrian Jarabo Adrian Jarabo is a Ph.D. student at Universidad de Zaragoza (Spain), in the field of computer science. He obtained his B.Sc. and M.Sc. in computer engineering in 2010 and 2012, respectively. His research interests include photo-realistic rendering, applied perception in rendering, and computational photography.

Hongzhi Wu Hongzhi Wu received B.Sc. in computer science from Fudan University in 2006, and Ph.D. in computer science from Yale University in 2012. He is currently an assistant professor in State Key Lab of CAD & CG, Zhejiang University. His research interests include appearance modeling, design and rendering. He has served on the program committees of EGSR and SCCG.

Julie Dorsey Julie Dorsey is a Professor of Computer Science at Yale University, where she teaches computer graphics. She came to Yale in 2002 from MIT, where she held tenured appointments in both the Department of Electrical Engineering and Computer Science (EECS) and the School of Architecture. She received undergraduate degrees in architecture and graduate degrees in computer science from Cornell University.

Holly Rushmeier Holly Rushmeier is a professor and chair of the Yale Department of Computer Science. Her research interests include shape and appearance capture, applications of perception in computer graphics, modeling material appearance and developing computational tools for cultural heritage.

Diego Gutierrez Diego Gutierrez is a tenured Associate Professor at the Universidad de Zaragoza, in Spain, where he leads the Graphics and Imaging Lab. His research interests include global illumination, computational photography and applied perception. He's served as Papers Chair for several conferences, like EGSR 2012 or APGV 2011. Hes also served on many Program Committees, including SIGGRAPH, SIGGRAPH Asia and Eurographics, and is an Associate Editor of four journals (MIT Press Presence, IEEE Computer Graphics & Applications, ACM Transactions on Applied Perception, and Computers & Graphics).