To appear in the ACM SIGGRAPH Asia 2008 conference proceedings

A Psychophysically Validated Metric for Bidirectional Texture Data Reduction

Jiří Filip[1]* Michael J. Chantler[1] Patrick R. Green[1] Michal Haindl[2]
Heriot-Watt University[1] Institute of Information Theory and Automation of the ASCR[2]

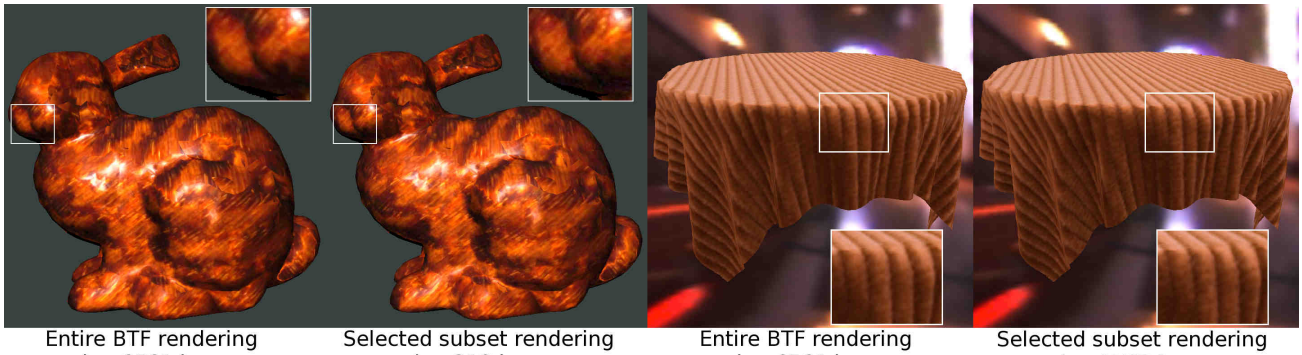


Figure 1: *Examples of BTF visual equivalence for two samples, different objects and illuminations.*

Abstract

Bidirectional Texture Functions (BTF) are commonly thought to provide the most realistic perceptual experience of materials from rendered images. The key to providing efficient compression of BTFs is the decision as to how much of the data should be preserved. We use psychophysical experiments to show that this decision depends critically upon the material concerned. Furthermore, we develop a BTF derived metric that enables us to automatically set a material's compression parameters in such a way as to provide users with a predefined perceptual quality. We investigate the correlation of three different BTF metrics with psychophysically derived data. Eight materials were presented to eleven naive observers who were asked to judge the perceived quality of BTF renderings as the amount of preserved data was varied. The metric showing the highest correlation with the thresholds set by the observers was the mean variance of individual BTF images. This metric was then used to automatically determine the material-specific compression parameters used in a vector quantisation scheme. The results were successfully validated in an experiment with six additional materials and eighteen observers. We show that using the psychophysically reduced BTF data significantly improves performance of a PCA-based compression method. On average, we were able to increase the compression ratios, and decrease processing times, by a factor of four without any differences being perceived.

CR Categories: I.3.7 [Three-Dimensional Graphics and Realism]: Color, shading, shadowing, and texture— [J.4]: Social and Behavioral Sciences—Psychology

Keywords: Surface texture, BTF, texture perception, phychophysical experiment, texture compression, perceptual metric

1 Introduction

In many industrial sectors, demand is currently increasing for accurate virtual representations of real-world materials. Important application areas include safety simulations and computer-aided design. In the first area, the main concern is choosing the right material to fulfill given safety limits of reflectance, while in the second the aim is to avoid costly and time consuming design cycles of material selection, solid model production and visual evaluation. These tasks, among others, require accurate photo-realistic representation of real material samples dependent on different illumination and viewing conditions.

The first real illumination and view dependent surface texture representation was the Bidirectional Texture Function (BTF), introduced by Dana et al. [1999]. A BTF is a six-dimensional function representing the appearance of a material sample surface for variable illumination $\omega_i(\theta_i, \varphi_i)$ and view $\omega_v(\theta_v, \varphi_v)$ directions, resulting in a six-dimensional monospectral function $BTF(x, y, \theta_i, \varphi_i, \theta_v, \varphi_v)$, where θ are elevation and φ azimuthal angles. Compared to the four-dimensional BRDF, a BTF depends on two additional parameters, a planar position (x, y) over a material surface. The BTF preserves such effects as masking, shadowing, inter-reflections and sub-surface scattering. In recent years, various BTF measurement systems have appeared as well as methods of interactive BTF editing [Kautz et al. 2007]. Although material visualization using BTFs provides superb visual quality, even a relatively small sample (e.g., 256x256) occupies several gigabytes of data in its raw form and is not suitable for real-time rendering. There has been much research aimed at developing efficient compression techniques that also allow computationally cheap reconstruction and visualization of BTFs (see [Müller et al. 2005] for a recent summary). However, all of these approaches have fixed or predefined compression parameters regardless of the sample's properties, and do not exploit the fact that each sample requires different amounts of data to provide the same level of perceptual fidelity.

Our contribution. This paper makes two main contributions. First we show explicitly, using a psychophysical experiment, that the amount of data required to represent a material to a given degree of perceived fidelity is dependent upon that material's characteristics. Second, we exploit this fact to develop a BTF metric that allows us to automatically set the appropriate compression parameters. Our technique makes it possible to significantly increase the compression ratio of many standard BTF compression methods while maintaining the same visual quality as the full dataset (see Figure 1).

*e-mail: filipj@macs.hw.ac.uk

Paper organisation. In the next section we survey related work (Section 2) while Section 3 provides an overview of our approach. Section 4 describes the BTF datasets that were used, their statistical analysis, and more importantly proposes the three variance-based metrics to be investigated. Section 5 proposes the core BTF vector quantization model. Section 6 describes the psychophysical experiment that was used to investigate the relationship between quantisation parameters and perceived rendering quality of different materials. Section 7 analyses the correlation between the three metrics and the material-dependent quantisation parameters required for given perceptual fidelity. The method's performance is validated and its limitations are discussed in Section 8, while its applications for BTF compression and optimal sampling are described in Section 9. Section 10 concludes the paper and proposes directions for further research.

2 Prior Work

The method proposed in this paper has several distinct and important aspects. These are dimensionality surface estimation, compression, and psychophysical aspects of bidirectional texture functions.

Dimensionality estimation. The dimensionality surface of materials represented by means of BTF was first analysed by Suen and Healey [2000], where a correlation analysis showed that BTF dimensionality increases almost linearly with elevation angles of illumination and view. However, this study did not investigate any correlation between BTF dimensionality and human perception. In this paper, we focus on description of BTF dimensionality and variability by means of a perceptually related metric.

Compression. A number of works have been published on methods for compression of BTF samples, in order to overcome the disadvantages of their massive size. A recent overview of this field can be found in [Müller et al. 2005]. The methods are based mainly on linear basis decomposition (PCA, wavelets) or on fitting BRDF models to pixel-wise BTF data. Another branch of efficient methods is based on probabilistic modelling [Haindl and Filip 2007]. Unfortunately, none of the compression approaches attempt to perform any perceptual analysis, and they use fixed or predefined parameters regardless of compressed BTF samples. Instead of compressing all BTF images, as is common in this field [Müller et al. 2005], we propose to compress just the perceptually relevant subset of them. Our work is most closely related to clustering-based techniques such as [Leung and Malik 2001] that decompose BTFs into individual textons that are further compressed. Unlike this top-down clustering method performed in a pixel-wise domain, we suggest a bottom-up vector quantization approach in the domain of illumination and view direction angles. This allows us not only to reduce the size of datasets, but also to obtain information about the sampling that is most appropriate for specific materials.

Psychophysical aspects. Applications of psychophysical methods have so far been restricted to investigations of how surface properties and the shape of real-world materials are perceived. Padilla et al. [2008] developed a model of perceived roughness in fractal surfaces. Ho et al. [2008] found that roughness perception is correlated with texture contrast. Lawson et al. [2003] showed that human performance in matching 3D shapes is lower for varying view directions. Ostrovsky [2005] pointed out that illumination inconsistency is hard to detect in geometrically irregular scenes. Ramanarayanan et al. [2007] developed metrics that predict the visual equivalence of rendered objects under warping and blurring of illumination and warping of object surfaces. Daly [1993] presented a metric that predicts the perceived similarity of images on the basis of known low-level processes in human vision. A psychophysically-based model of light reflection with two perceptually meaningful uniform dimensions was developed by Pellacini et al. [2000]. Fleming et al. [2003] pointed out the importance of real-world illumination for correct matching of surface roughness and specularity.

Limited work has been carried out on the perceptual aspects of real BRDF/BTF measurements. te Pas and Pont [2005a] showed that changes in surface BRDF and illumination are often confounded, but that adding complex illumination or 3D texture improves the matching. Dependency of the perception of a light source direction on surface BRDF and its 3D shape was shown in [te Pas and Pont 2005b; Khang et al. 2006]. Matusik et al. [2003] psychophysically evaluated large sets of BRDF samples, and showed that there are consistent transitions in perceived properties between different samples. Vangorp et al. [2007] found that object shape considerably influences perception of BRDF samples. Meseth et al. [2006] showed that real-world scenes represented using BTFs were perceived as almost identical while those using flat textures modulated by BRDFs scored much lower. Filip et al. [2008] psychophysically analysed several BTF samples and suggested several general rules for their appropriate uniform resampling.

Although this research yielded interesting results, most of the methods discussed above focus on the perception or compression of materials represented by BRDF or BTF depending on the position or type of illumination. However, we are not aware of any work that applied psychophysical methods to systematically assess the sample dependent perceptual effects of BTF compression. Unlike other research dealing with the perceptual classification of images rendered using BRDF or BTF, our aim was to find the relationship between the perceptual features of a BTF sample and the appropriate settings of its compression parameters. These parameters can then be used for reducing the number of BTF images, without any effects of the reduction being perceptible by an average human observer.

3 Overview of the Method

A general overview of the proposed method is given in Figure 2. First, a set of statistical features of each sample is computed. Then a vector quantization model is proposed that removes BTF images selectively, depending on a predefined set of quantization thresholds. The rendered images are used as stimuli in the psychophysical experiment, which establishes the largest degree of degradation at which rendered images are not distinguishable from those generated from the original full BTF set. This process is repeated for a number of different materials, object shapes, and illuminations. The correlation of three precomputed statistical features with the psychophysically measured data is analysed and the one with the highest correlation is chosen and psychophysically scaled. This scaling can, for any novel BTF sample, automatically establish sample variance-based quantization thresholds based on the sample variability, yielding the perceptually important subset of BTF images.

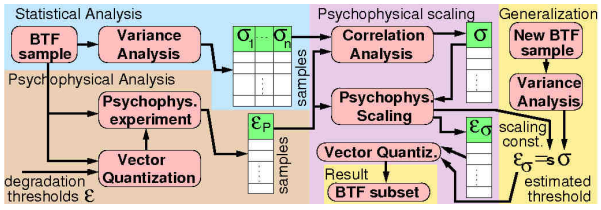


Figure 2: *Scheme of the proposed method of BTF analysis.*

4 BTF Samples Statistical Analysis

We have used measurements from the BTF Database Bonn [1] as data samples. These data have planar resolution $N \times M = 256 \times 256$

[1] http://btf.cs.uni-bonn.de/

pixels and illumination and viewing directions ($n_i \times n_v = 81 \times 81$) producing uniform sampling of a hemisphere above a material sample (see Figure 3). The following BTF samples were used in our experiments (see Figure 10): aluminum profile (*alu*), corduroy fabric (*corduroy*), dark cushion fabric (*fabric d.*), artificial dark leather (*leather d.*), artificial light leather (*leather l.*), glazed tile with white pointing (*impalla*), dark lacquered wood (*wood d.*), and knitted wool (*wool*). To avoid high resource requirements during processing of complete BTF samples, and visible seams when mapping these data onto test 3D objects, we applied the image tiling approach [Somol and Haindl 2005]. We cut one to five seamless BTF tiles for each tested sample. The size and number of tiles (second column of Tab 1) was chosen depending on the regularity and spatial frequency of the surface pattern of the sample, and on the memory resources available.

First of all we analysed the importance of colour information in the tested samples. We did so by calculating the variance of the samples in individual channels of the perceptually uniform CIE LAB color-space. The overall variance in individual channels for all tested samples is shown in columns 3-5 of Table 1. From these values we concluded that data from the luminance channel vary the most. This difference between channels is least clear for the *wood* sample, followed by the *wool* and *corduroy* samples, which exhibit strong subsurface scattering and / or translucency effects. To reduce the processing times of our method we decided to use only the luminance values, however, some BTF samples having large color variations may require the use of full spectral information.

BTF	# of	BTF σ_1			tested BTF variances		
sample	tiles	L	a	b	σ_1	σ_2	σ_3
alu	5	**26.1**	1.7	2.6	26.1	26.4	8.6
corduroy	5	**22.7**	4.7	13.0	22.7	18.4	11.5
fabric d.	5	**16.1**	2.1	2.4	16.1	14.0	6.8
leather d.	2	**18.5**	2.0	1.9	18.5	18.6	3.6
leather l.	2	**25.6**	2.8	7.9	25.6	15.1	4.2
impalla	1	**29.5**	2.3	3.3	29.5	19.4	17.5
wood d.	1	**13.9**	12.8	15.2	13.9	13.8	4.5
wool	5	**22.2**	3.7	17.9	22.2	14.8	7.9

Table 1: *Variance values of tested samples in individual channels of CIE LAB color-space, and values of the proposed metrics.*

Different materials, and therefore their representation by means of BTF samples, exhibit different degrees of variability in their appearance. This variability depends on illumination and view direction, and consequently on optical properties of the underlying micro-structure of the sample, on its roughness or on its tendency to complex inter-reflection and sub-surface scattering effects. We believe that understanding the variability of the sample is the key for development of optimal compression algorithms of BTFs representing the real materials. For this reason we tested several measures of BTF sample variance with the goal of finding the one that correlates best with human perception; in other words, to find a computationally efficient and perceptually plausible measure of the sensitivity of a sample to perceptual degradation explained in the following section. For this reason we tested the following three metrics:

$$\begin{aligned} \sigma_1 &= Var_{x,y,i,v}(BTF_L(x,y,i,v)) \ , \\ \sigma_2 &= \frac{1}{MN}\sum_{\forall x,y} Var_{i,v}(BTF_L(x,y,i,v)) \ , \\ \sigma_3 &= \frac{1}{n_i n_v}\sum_{\forall i,v} Var_{x,y}(BTF_L(x,y,i,v)) \ , \end{aligned} \qquad (1)$$

where σ_1 computes the total variance of a BTF luminance, σ_2 computes the mean variance over individual pixel-wise BRDF images of size $n_i \times n_v$, and σ_3 gives the mean variance over BTF images of size $M \times N$. All of the variances were computed in the CIE LAB luminance channel only. Values of all these variances for all tested samples are shown in the least three columns of Table 1.

Later in the paper, we will compare how each of these variances correlate with a perceptually estimated degradation threshold of the vector quantization model proposed in the following section.

5 Vector Quantization of BTF Images

Choosing a subset of perceptually important BTF images from each BTF sample is the core of the proposed method. As the first step in such a BTF subset selection we need to evaluate similarity of all pairs of images in the sample. As there are 6561 images representing each sample, one can imagine that such a similarity measure has to be relatively simple to evaluate. We did so by computation of the symmetric similarity matrix $A(n_i n_v \times n_i n_v)$ where each value represents the average pixel-wise luminance difference for a corresponding pair of BTF images. The next step is to choose the perceptually important BTF subset. This is achieved by substitution of redundant / similar images by an index of the most similar image. This information is stored in a positive index of array $I(n_i \times n_v)$. The substitution algorithm, outlined in Table 2, has as inputs the precomputed similarity matrix A, predefined luminance difference threshold ε, and maximum allowed view direction euclidean difference over hemisphere $\mathsf{hemiDist}(.)$ (we use a default threshold $\eta = 0.8$ that corresponds to maximum view angles distortion $\sim 45^o$). The algorithm starts with initialization of the substitution index I (lines 1,2). A corresponding BTF image is substituted when the difference between both images $m(i_1, v_1)$ and $n(i_2, v_2)$ (lines 4,6) is less than ε and their view direction difference is less than η (line 7). Parameters $i \in 0 \ldots n_i - 1$ and $v \in 0 \ldots n_v - 1$ are indices of individual measurements points circling the hemisphere above the sample starting at its pole. If the image was substituted previously the substitution is overridden if differences in the possible substitution are lower than those of the previous one (lines 8-10). The output of the algorithm is the index array $I(i, v)$, which has either a negative value at positions where the original image was preserved, or positive indices in range $(0 \ldots n_i n_v - 1)$, specifying the index of one of the preserved images, at positions where the image was substituted. The reconstruction of an arbitrary BTF

1	**for** $k(0 \ldots n_i - 1), l(0 \ldots n_v - 1)$
2	$I(k,l) = -1$
3	**for** $i_1(0 \ldots n_i - 1), v_1(0 \ldots n_v - 1)$
4	$m = v_1 n_i + i_1$, if(m=n) break
5	**for** $i_2(0 \ldots n_i - 1), v_2(0 \ldots n_v - 1)$
6	$n = v_2 n_i + i_2$
7	**if** $A(m,n) < \varepsilon$ **and** $\mathsf{hemiDist}(v_1, v_2) < \eta$
8	**if** $I(i_2, v_2) = -1$ **or** $A(m,n) < A(I(i_2, v_2), n)$
9	**if** $m \neq n$ **and** $I(i_1, v_1) = -1$
10	$I(i_2, v_2) = m$

Table 2: *Algorithm of BTF subset selection.*

image from its subset, for a given combination of the measured illumination $i_a \in (0 \ldots n_i)$ and view $v_a \in (0 \ldots n_v)$ indices, can be written as follows

$$BTF_{i,v} \quad \begin{cases} i = i_a, \ v = v_a & \text{if } I(i_a, v_a) < 0 \\ i = \frac{I(i_a, v_a)}{n_i}, \ v = I(i_a, v_a) - i & \text{otherwise} \end{cases} . \qquad (2)$$

Figure 4 shows the relationship between the luminance difference threshold ε and the number of preserved BTF images. From this figure we can see that, for all tested samples, the number of images decreases dramatically as ε increases from 1.0 to 9.0, and then only slightly as ε increases beyond 12.0.

An example of a sample degradation introduced by the proposed BTF images vector quantization scheme is shown in Figure 5. Here,

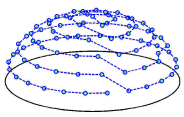


Figure 3: *81 sampling directions.*

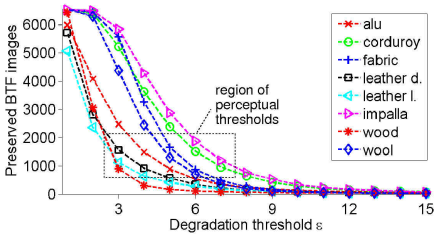


Figure 4: *Number of preserved BTF images dependent on the luminance threshold ε.*

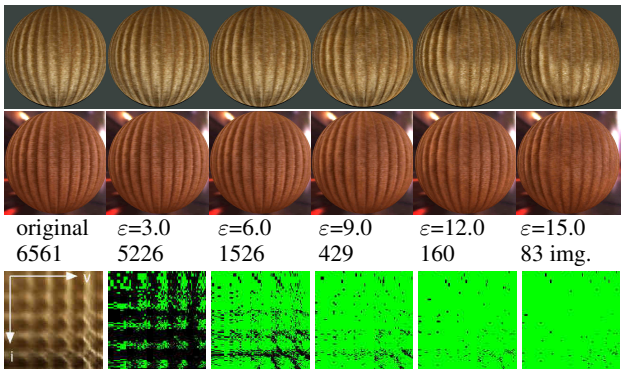


Figure 5: *Renderings of a sphere covered with the corduroy BTF sample, under point-light and environment illumination (first and second rows), with increasing values of the luminance difference threshold ε (third row). Fourth row – number of preserved images. Fifth row – substitution maps with removed samples as green (rows – illumination direction, columns – viewing direction, top left corner – pole of the sampled hemisphere) – see Figure 3).*

a sphere is covered by a *corduroy* BTF sample illuminated by either point-light or environmental illumination. The sample is gradually degraded from the original set by increasing the luminance difference threshold ε to 15.0. The figure also shows substitution maps, which indicate which of the 6561 ($81 \times 81 = n_i \times n_v$) BTF images have been substituted at five values of ε from 3.0 to 15.0. In these maps, green indicates substituted and black indicates preserved BTF images. The reference BRDF image (spatially averaged original BTF) is also shown on the left.

From these results, we can see that the proposed vector quantization model allows efficient control of the rendered BTF quality by means of the luminance difference threshold ε. The greater the threshold is, the fewer BTF images are preserved and used at the rendering stage, and the greater is the perceptual error between original and degraded BTF samples. Note that the model is based on pixel-wise differences in CIE LAB color space, so any imperfections in the perceptual uniformity of this space, or in spatial registration of individual BTF images, can affect the results that it yields. Additionally, different underlying materials have completely different structural properties, and these differences lead to variable view occlusion errors. For these reasons, we expect that each sample will require a unique degradation threshold for optimal perceptual results. In the following section we will exploit a way to use psychophysical data to set these thresholds.

6 Psychophysical Experiment

The goal of the experiment was to identify thresholds for the proposed vector quantization model for different combinations of tested material BTF, illumination environment and 3D object. At these thresholds, we aim to achieve the minimum size of the BTF image set consistent with there being no perceptible difference from a rendering using the entire image set.

Experimental Stimuli. As experimental stimuli we have used pairs of static images of size 800×800, representing a material BTF rendered on a 3D object. Each pair consisted of a rendering using the full original dataset and one using a degraded set, with ε values 3, 6, 9, 12, 15 or 18. Pairs of images were displayed simultaneously, side-by-side. Three different 3D objects were used; *sphere*, *tablecloth* and *bunny* [2], and the *sphere* was rendered with three different illuminations; point-light, and two environments, *grace* [3] and *grassplain* (see Figure 6). The point-light was positioned slightly above

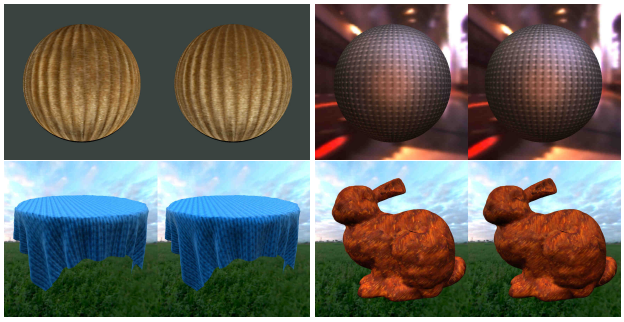


Figure 6: *Four examples of experimental stimuli showing the three test objects, and illumination environments for $\varepsilon = 9.0$.*

the viewing direction corresponding to illumination of the room in which testing took place, and avoiding extensive hard shadows on object surfaces. We chose to test the *sphere* under different types of illumination because its geometry provides a wide range of illumination and viewing combinations without introducing unwanted effects of higher curvatures. The environment maps were approximated by a set of 144 discrete point-lights [Havran et al. 2005]. The remaining two objects were illuminated by the *grassplain* environment only, as we considered that this natural environment would be most familiar to the participants. The background of the point-light illuminated stimuli, and the remaining space on the screen, was set to dark gray. Given eight material BTFs, six degradation levels and five object/illumination combinations, the total number of stimuli was 240.

Participants. Eleven paid observers (six males, five females) participated in the experiments. All were students or University employees working in different fields, were less than 35 years of age, and had normal or corrected to normal vision. All were naive with respect to the purpose and design of the experiment.

Experimental Procedure. The participants were shown the 240 stimuli in a random order and asked a yes-no question: *'Can you spot any differences in the material covering the two objects?'* Note that this question is a strict test of the ability to detect a difference and allows participants to make a 'yes' response on the basis of only a minor local difference. Participants were given as much time as they wanted to make their decision, up to a maximum of 10 seconds; after this time, the response was always recorded as 'no'. This procedure was followed to prevent participants from adopting a strategy of making an exhaustive pixel-wise comparison of the images, however, it turned out that more than 90% of the responses were made in less than 10 seconds. There was a pause of two seconds between stimului presentations, and participants took on average less than 40 minutes to perform the whole experiment, which was split into two sessions. All stimuli were presented on a calibrated 20.1" NEC2090UXi LCD display (60 Hz, resolution

[2] http://graphics.stanford.edu/data/3Dscanrep/
[3] http://www.debevec.org

1600×1200, brightness $120cd/m^2$, 6500K, gamma 2.2). The experiment was performed under dim room lighting. Participants viewed the screen at a distance of 0.9m, so that each object in a pair subtended approximately 9^o of visual angle.

Analysis and Discussion of Experimental Results. When participants reported a difference between the rendered images their response was assigned a value of 1, and otherwise 0. By averaging the responses of all participants, we obtained psychometric data relating average response to the luminance difference threshold ε of the vector quantization model (Section 5). There are five such datasets for each BTF sample; three for the *sphere* under different illuminations, and two additional for the different objects illuminated by the *grassplain* environment. We have computed standard variances of the psychometric data for all tested samples with average value $\sigma^2 = 0.14$ and have not found any significant trends except increased values for the *alu* sample as we will explain later. The psychophysical data obtained can be represented by the psychometric function $\psi(x)$ [Wichmann and Hill 2001], which specifies the relationship between the underlying probability ψ of positive response and the stimulus intensity x

$$\psi(x;\alpha,\beta,\gamma,\lambda) = \gamma + (1-\gamma-\lambda)F(x;\alpha,\beta) \ , \qquad (3)$$

where F is data fitting function with parameters α and β, γ specifies guess rate (i.e., response to zero stimulus), and λ miss rate (i.e., incorrect response for large stimulus). The functions were fitted to the measured data using the *psignifit* [4] tool. We have tested several distributions to fit our data (Weibull, Gumbel, sigmoid) and finally have obtained the best fitting using the Weibull cumulative distribution, which is most commonly used in life data analysis due to its flexibility. It is described as

$$F(x,\alpha,\beta) = 1 - \exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right] \qquad (4)$$

for $x \geq 0$, where $\beta > 0$ is the shape parameter and $\alpha > 0$ is the scale parameter of the distribution. The resulting psychometric functions averaged for all eight tested BTF samples are shown in Figure 7. The left graph shows functions corresponding to the same object (*sphere*) with different illuminations while the right graph shows functions for different objects lit by the same environment (*grassplain*). Goodness of the fit was evaluated and shown as confidence intervals for individual functions at response levels (0.25, 0.5, and 0.75). If the data distribution were Gaussian, these intervals would correspond to a standard deviation from the mean values in the range ± 1.

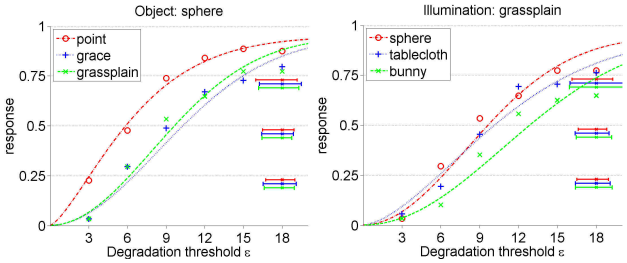


Figure 7: *Averaged psychometric functions for all tested BTF samples (left) for three different BTF illuminations (point-light and environments* grace*,* grassplain*), (right) for three different objects (*sphere*,* tablecloth*, and* bunny*).*

From these graphs we can conclude the following:

- **Varying illumination:** The degradation is easier to spot for point-light than for environment illumination, where its effects are hidden in sample structure due to interpolation between discrete sets of point lights. The performance for both environments was almost identical.
- **Varying object:** The more complex the object, the more difficult it is to spot the degradation. The effects of degradation are harder to detect in the *bunny* than in the *sphere* and the *tablecloth*. This is consistent with the fact that the latter objects contain large surfaces with low curvature (compare objects in Figure 6), where the degradation can be detected more easily without distortion by surface curvature. We therefore believe that the conclusion given in [Vangorp et al. 2007] stating that 'the sphere is not very-well suited object for material (BRDF) discrimination tasks' cannot be generalized to BTF data.

These results suggest that, the less uniform the illumination environment or the curvature of the scene object is, the more sensitive the participants were to degradation of the BTF data.

Individual psychometric functions for all BTF samples (mapped on a sphere) for varying illumination are shown in Figure 8, together with confidence intervals for their fitting at the response level 0.25. Observing the solid black psychometric curves fitted to data averaged for all environments, we can compare sensitivity of participants' responses to the same levels of degradation for different materials. The results suggest that the samples most sensitive to degradation are the leathers (*leather d.*, *leather l.*) while those least sensitive are the *alu* and *impalla* samples. The appearance of *alu* and *impalla* samples seems to be preserved regardless of degradation level apart from subtle differences in holes or pointing in the material structure. The samples most sensitive to a change of the illumination environment (see Figure 8) were *alu*, *fabric d.*, and *leather l.*. For changes in the object the most sensitive were *leather d.*, *leather l.*, and *wood d.*.

The following section outlines a way of using these experimental data to determine which of the statistical measures introduced in Section 4 is optimal for setting a compression threshold, and how it can be scaled perceptually.

7 Psychophysical Scaling

To exploit the psychophysical results, presented in the previous section, we need to determine the greatest level of degradation of individual BTF samples that is not detectable by the average human observer. To do this, we use the values of the threshold ε at which a difference between rendered images is detected by only 25% of participants. Note that due to the applied context of our work we use a more stringent value than 50%, which is standard in psychophysical research. These thresholds are computed from the fitted psychophysical functions [Wichmann and Hill 2001] for different illuminations (see Figure 8) and objects using equation

$$\hat{\varepsilon}^{p=0.25} = \alpha \sqrt[\beta]{\ln\left(\frac{1-\gamma-\lambda}{1-0.25-\lambda}\right)} \ , \qquad (5)$$

where α, β are estimated parameters of the Weibull distribution and γ and λ are guess and miss rates estimated during the fitting of (4). The perceptual thresholds $\hat{\varepsilon}$ estimated using equation (5) for all psychometric functions are shown in Table 3. The perceptual thresholds for different illuminations of object *sphere* are shown in the columns 2-4 of this table followed by their average in the fifth column. Thresholds for different objects illuminated by the *grassplain* environment, shown in the last three columns of Table 3, are generally higher. This is caused by the type of illumination and by finer BTF resolution and higher shape variability of the *tablecloth* and *bunny* objects (see Figure 6). Examples of renderings comparing the original data with the perceptually reduced subsets for

[4] http://www.bootstrap-software.org/psignifit/

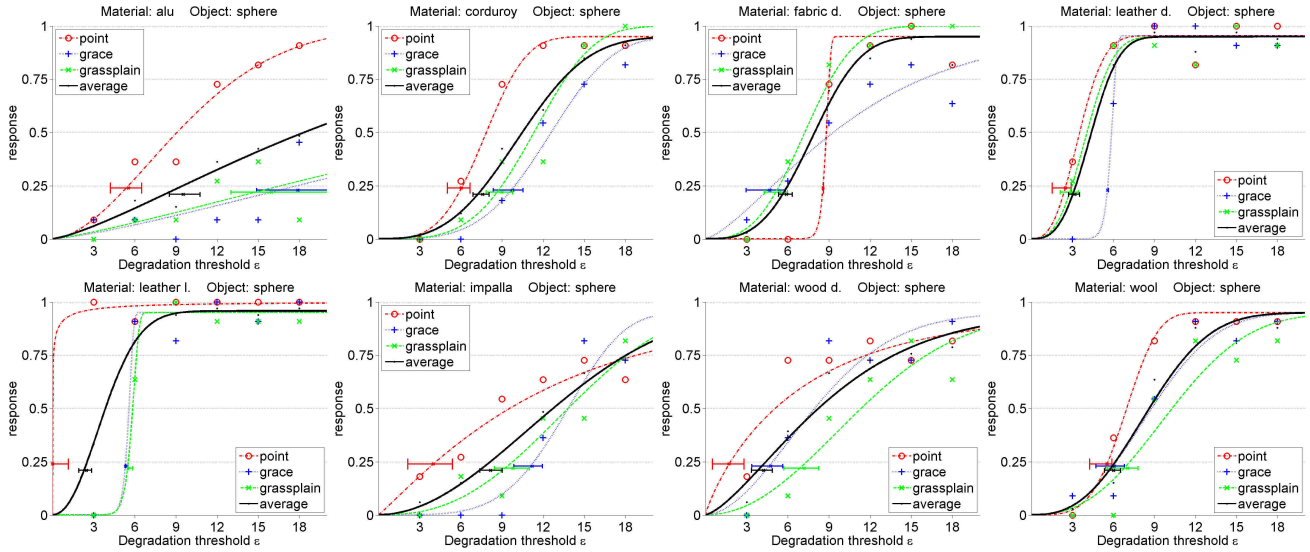


Figure 8: *Psychometric functions for all tested BTF samples (alu, corduroy, fabric d., leather l., leather d., impalla, wood d., wool) for three different BTF illuminations (point-light and environments* grace, grassplain*) and for their averaged values. All the BTF samples were enlarged to* 1024×1024 *pixels and shown mapped on a sphere.*

BTF sample	$\hat{\varepsilon}$ – different illuminations				$\hat{\varepsilon}$ – different objects		
	point	grace	grass	AVG	sphere	table	bunny
alu	5.5	17.9	16.0	**13.1**	16.0	10.6	25.4
corduroy	6.0	9.8	9.0	**7.6**	9.0	8.6	12.7
fabric d.	8.6	4.7	5.4	**5.9**	5.4	6.3	6.0
leather d.	2.5	5.6	2.9	**3.2**	2.9	3.3	8.4
leather l.	0.0	5.3	5.5	**2.5**	5.5	4.4	3.4
impalla	4.0	11.2	9.8	**8.2**	9.8	11.5	14.2
wood d.	1.7	4.8	7.2	**4.2**	7.2	6.4	7.9
wool	5.5	6.1	7.1	**6.0**	7.1	7.9	9.0

Table 3: *Summary of the perceptual thresholds* $\hat{\varepsilon}$ *estimated by the experiment for all tested samples, illuminations, and objects.*

wood ($\hat{\varepsilon} = 4.0$) and for *corduroy* ($\hat{\varepsilon} = 6.0$) samples are shown in Figure 1.

The psychophysically estimated degradation thresholds provide us with information about the sensitivity of the perceptual appearance of individual BTF samples to introduced degradation. In the next step we attempt to match the estimated psychophysical thresholds $\hat{\varepsilon}$ with the most similar statistical feature, i.e., one of variances σ_1, σ_2, and σ_3 precomputed in the last three columns of Table 1. To do so we used average measured $\hat{\varepsilon}$ values for object *sphere* from the fourth column of Table 3 and compared their correlation with values of the individual variances. Using the averaged values is justified by their overall best confidence intervals (see Figure 8). The correlation coefficient $R_{X,Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y}$ was used for comparison of vectors X, Y, where μ and σ are their means and variances. The computed coefficients for the tested variances are $R_{\sigma_1} = 0.48$, $R_{\sigma_2} = 0.67$, and $R_{\sigma_3} = 0.76$. The mean variance computed over BTF images, i.e., σ_3, exhibits the highest correlation with the averaged psychophysical thresholds $\hat{\varepsilon}$. The correlation with σ_3 is even greater ($R = 0.90$) when the thresholds for sample *alu* are not taken into account. This can be justified on the basis of the wide confidence intervals in fitting the psychometric functions for this sample (see Figure 8 top-left).

We have found that variance σ_3 is the best predictor of the extent to which a sample can be degraded before an observer can detect the effects on the rendered image. However, values of σ_3 are only proportional to psychophysically estimated thresholds $\hat{\varepsilon}$, and have to be scaled to be useful as an absolute estimate of threshold values. Due to the fact that psychophysical thresholds vary considerably, depending on the illumination scheme, the scaling of σ_3 was done for individual lighting models separately. For each illumination a different scaling coefficient s was computed, by means of solving a set of linear equations $\Sigma_3 s = \hat{E}$, where Σ_3 and $\hat{E}$ are vectors of values σ_3 and $\hat{\varepsilon}$. Finally, the threshold ε_σ for any new BTF sample is estimated by means of simple scaling $\varepsilon_\sigma = s\sigma_3$. In Figure 9 it is shown that the estimated and scaled values ε_σ (even group of bars) reliably follow the psychophysically measured ones $\hat{\varepsilon}$ (odd group of bars) for most of the tested samples and illumination conditions, represented by the individual bars. The exception is the sample *alu*, where the difference in the two thresholds is probably caused by inconsistent psychometric data (see the wide confidence intervals for fitting the function in Figure 8 top-left). The errorbars on $\hat{\varepsilon}$ values represent confidence intervals of psychometric functions fitting. The estimated ε_σ values are conservative, i.e., lower than measured thresholds $\hat{\varepsilon}$, in all cases but the sample *impalla* where the computed variance σ_3, and ε_σ are increased due to the distinct reflectance properties of the tiles and the pointing material. From Figure 9 it is also apparent that thresholds for both environment illuminations are very similar and often considerably higher than for point light. Values averaged for all lighting models are shown in the yellow bars. These average estimated ε_σ values generally correspond to the lower elbow of samples' degradation graphs in Figure 4 as depicted by the rectangular area. To assure reliable threshold estimation for samples with high variance σ_3 and to preserve correct shadowing effects in their rough structure we limit ε_σ values to 6.0. Using the higher ε_σ thresholds would require deeper analysis of the sample's visual properties to identify the particular reasons for the increased variance. This will be a subject of our future research.

Additionally, the bottom of Figure 9 shows original σ_3 for individual samples and numbers of preserved BTF images k_A corresponding to the average threshold (yellow bar of ε_σ). Since the original samples comprise 6561 images, we conclude that only $10-35\%$ of the original BTF images are needed to maintain the same visual appearance as the original samples. Figure 10 compares the original renderings of all the tested samples under point-light and *grace* illumination with rendering of their perceptual subsets obtained using the estimated thresholds ε_σ shown in Figure 9.

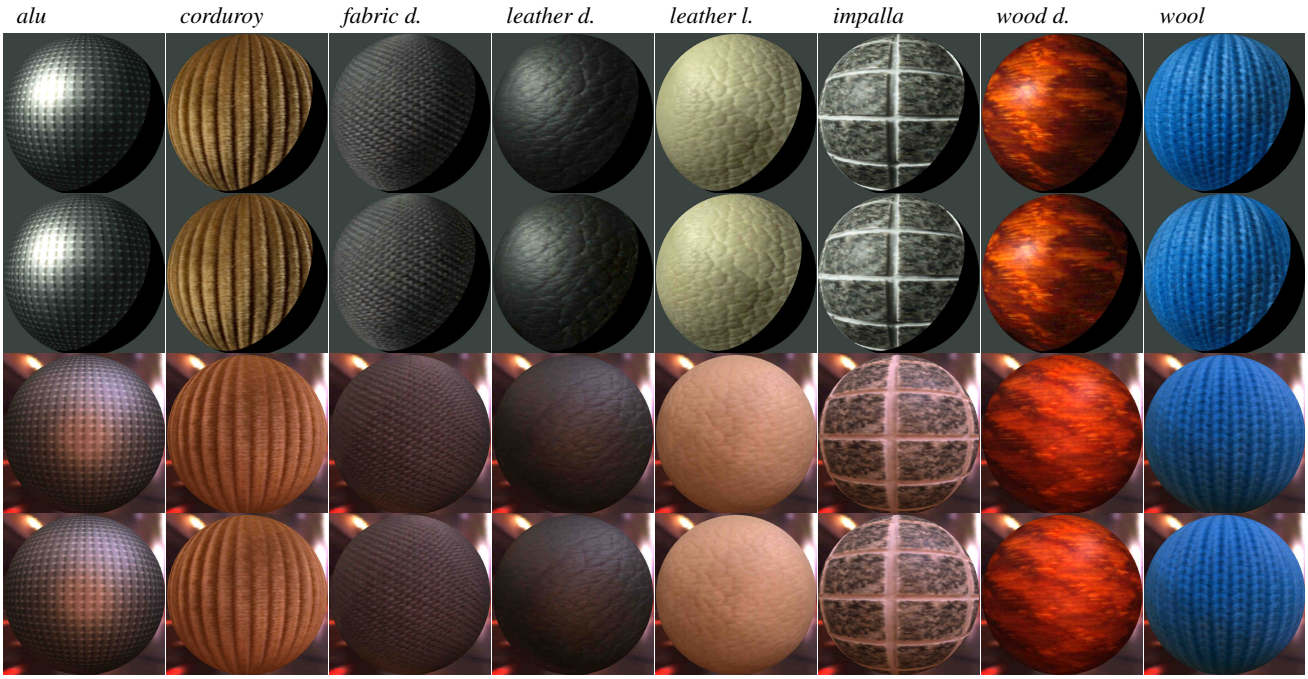
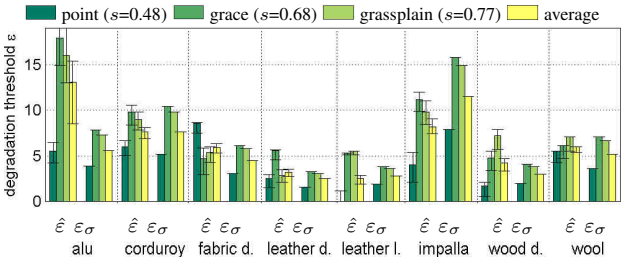
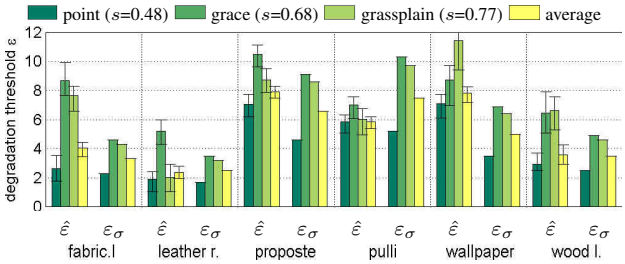


Figure 10: *Original samples renderings compared with their perceptually equivalent renderings from the subset of the sample, for point-light from left (the first two rows) and* grace *environment (the last two rows). The estimated thresholds ε_σ are shown in Figure 9.*



	alu	corduroy	fabric d.	leather d.	leather l.	impalla	wood d.	wool
k_A:	625	1526	2322	2005	1244	1887	852	1137
σ_3:	8.6	11.5	6.8	3.6	4.2	17.5	4.5	7.9

Figure 9: *Comparison of the measured perceptual thresholds $\hat{\varepsilon}$ with the ε_σ values estimated by scaling of sample variance for all illumination types and tested samples; average number of the preserved BTF images k_A and original values of σ_3 variance.*



	fabric.l	leather r.	proposte	pulli	wallpaper	wood l.
k_A:	2221	1514	1564	2095	971	293
σ_3:	5.1	3.8	10.1	11.4	7.6	4.5

Figure 11: *Comparison of the measured perceptual thresholds $\hat{\varepsilon}$ with the ε_σ values estimated by scaling of the sample variance for all illumination types and validation samples; average number of the preserved BTF images k_A and original values of σ_3 variance.*

8 Validation and Discussion of the Method

In this section, we validate the method for automatically setting thresholds that was described in the previous section. We used six new BTF samples (see Figure 12): light fabric (*fabric l.*), real light leather (*leather r.*), cushion fabric (*proposte*), knitted wool (*pulli*), *wallpaper*, and light lacquered wood (*wood l.*).

We performed an additional experiment to validate the thresholds ε_σ obtained by scaling of the variance σ_3. The experimental procedure and setup were the same as in the main experiment. Only the *sphere* object was used for the six tested samples, and so, with the three illumination types used previously, we obtained 108 stimuli. Eighteen participants took part, only three of whom had been participants in the main experiment. We fitted psychometric functions to the data in the same way as in the main experiment and determined perceptual degradation thresholds at the 25% level. Figure 11 shows the perceptually measured thresholds $\hat{\varepsilon}$ compared with estimated and scaled thresholds ε_σ, for different samples and illumination types. Original σ_3 values and numbers of preserved BTF images k_A corresponding to the average threshold (yellow bar of ε_σ) are shown at the bottom of the figure. This time, the thresholds ε_σ were obtained by means of scaling the σ_3 values using the scaling constants s for each illumination type (top of Figure 11) estimated in the previous section. We can see that the estimated values ε_σ follow the perceptual ones $\hat{\varepsilon}$ for most of the samples and illumination types, and that the estimated values ε_σ are conservative, i.e., lower, with respect to the psychophysically measured values $\hat{\varepsilon}$. The only exception is the sample *pulli*, with a very rough translucent structure, where the estimated thresholds for environment illumination suggest higher degradation than the values obtained by experiment. On the other hand, for sample *wallpaper*, the estimated values are quite low when compared to the measured ones. This is probably caused by the colorful pattern on the material sample increasing the variance σ_3, without an increase in the structural complexity of the sample. The performance of the method on the validation samples for point-light illumination is shown in Figure 12. The first row shows renderings of original samples while the second row shows renderings obtained by their subsets using

estimated threshold ε_σ for point-light. There are no obvious differences, as would be expected in most cases due to the conservative nature of the ε_σ values.

We conclude from the results that the proposed variance-based threshold estimation method is more reliable for samples of low color variation, and that it provides conservative, but perhaps not optimal thresholds for the majority of BTF samples. To sum up, we believe that reliable degradation thresholds can be estimated automatically and scaled by experimental data obtained for completely different samples as shown in the validation experiment. Thus the psychophysical experiment only needs to be run once with a representative set of samples, and then we can use the results to scale any new measured BTF sample. However, to achieve the highest scaling accuracy such a pilot experiment should comprise representative types of the samples being scaled. Our results also show, that the estimated thresholds for *grace* and *grassplain* environment illuminations are very similar, so we believe that many different environments can be represented by the same threshold without significant loss of accuracy. Note, that the texture resolution on the sphere during the experiments was 1024×1024 pixels and thus some of the faint degradation artifacts might become even less apparent for higher resolutions. Additionally, the *sphere* is a very simple object, and the perceptual thresholds for shapes with higher curvature variability are considerably higher (see the last three columns in Table 3). Due to these facts, we believe that in real applications the thresholds may be higher than those estimated in this paper, resulting in even greater reduction in sample size.

9 Applications of the Method

This section illustrates possible applications of the proposed psychophysically driven reduction of BTF samples for their optimal compression and sampling. In this section, for degradation of samples, we have used the average measured thresholds $\hat{\varepsilon}$ that were rounded and limited as shown in Table 4. Note, that the estimated and scaled values ε_σ can be used in the same way.

9.1 BTF Compression

The compression attained by the proposed BTF sample reduction method, can be significantly improved by application of published image-based BTF compression algorithms summarized in [Müller et al. 2005].

The Compression Algorithm. We have chosen the algorithm from [Müller et al. 2003] due to its robustness, relatively high compression ratio, reconstruction quality, and feasible hardware implementation. The method is based on iterative clustering of BTF data with their further refinement in each cluster by a dedicated local PCA model. The squared PCA reconstruction error is used as a distance measure in the clustering process. The described BTF factorization can be stated as

$$\mathrm{BTF}(x,y,i,v) \approx \sum_{j=1}^{c} \alpha_l(j,[x,y]) E_l(I(i,v),j) + \mu_l(I(i,v)), \tag{6}$$

where i, v are illumination and view direction indices, l is a cluster index from a look-up table of size $(M \times N)$ for required planar coordinates (x, y), c is the number of preserved principal components for each cluster, α are PCA weights, E are either eigen-images or eigen-BRDFs, and μ is the mean vector of the cluster. We have applied the method in eigen-BRDF representation as suggested by Müller et al. However, instead of PCA compression of full apparent BRDF sets (LPCA) (i.e., $\mathrm{BTF}(x, y; \forall\omega_i; \forall\omega_v)$ of size $n_i \times n_v$, in each cluster), we compressed only the k psychophysically important combinations of illumination and viewing directions (LPCA+) (i.e., sparse BRDF sets only). By row-wise scanning we rearranged these sparse eigen-images into eigen-vectors $E(k \times c)$. During rendering we restore preserved k BRDF values (6) and the missing values are filled in using the information stored in the substitution index $I(i, v)$ (Section 5, (2)).

Results. The parameters of the algorithm were set to 5 clusters r and 10 PCA eigen-images c representing the data in each cluster. With this configuration, the LPCA method achieves compression of the full BTF set of 1:48 (parameter size $MN + 3rk(1 + 4c)$, where $M \times N = 100 \times 100$ pixels and $k = 6561$). When it is applied not to the full set, but to the perceptually important subset obtained by the proposed method (LPCA+), the results shown in the fourth column of Table 4 are obtained (typically $k \leq 2000$). Although the compression ratio is variable, according to the estimated threshold and therefore the number of preserved BTF images for each sample is variable, it is always higher than for the LPCA method. Pixel-wise comparisons, computed in CIE LAB color space, of the results of using the full BTF data set with the results obtained from sets compressed by each of the two methods are shown in the fifth column of Table 4. These indicate that the LPCA+ method yields only a slightly greater error than does the LPCA method. Additionally, the BTF sample analysis using the proposed modification was on average $4\times$ faster and required $3\times$ lower memory resources during processing. Our CPU rendering implementation achieves $\sim$3 frames/second for both variants and a BTF resolution of 512×512 pixels. We believe that a GPU implementation can benefit from considerably less values k to be interpolated.

BTF sample name	est. $\hat{\varepsilon}$	images. k	LPCA+ C.R.	ΔE in CIE LAB LPCA	LPCA+
alu	6	**544**	**1:261**	3.2	6.8
corduroy	6	**1562**	**1:144**	29.8	30.2
fabric d.	5	**1669**	**1:136**	4.2	7.3
leather d.	3	**1565**	**1:142**	19.6	20.0
leather l.	3	**1129**	**1:177**	22.6	22.9
impalla	6	**1887**	**1:125**	6.9	8.2
wood d.	4	**319**	**1:319**	30.1	32.9
wool	6	**707**	**1:230**	2.6	5.2

Table 4: *Comparison of the tested BTF samples in terms of estimated thresholds $\hat{\varepsilon}$ and corresponding number of preserved images. Comparison of BTF compression ratio using the entire BTF data (LPCA) and using only their perceptually important BTF subset (LPCA+), in terms of compression ratios and pixel-wise error.*

Psychophysical Test of the Effects of Compression. We carried out a further psychophysical experiment to determine whether any difference in quality of rendered images could be detected when the proposed compression method (LPCA+) was used, compared to the use of LPCA alone. The stimuli were pairs of rendered spheres, using the same eight materials as in the main experiment, and three illuminations (point-light and *grace*, used before, and *st.peters*). One of the spheres was rendered after applying the LPCA algorithm to the whole BTF sample, and the other after applying it to the subset yielded by the proposed degradation method (LPCA+). Otherwise, the experimental procedure and setup were as in the main experiment. Eleven paid participants took part, only three of whom had been participants in the main experiment. Average responses with twice the standard error are shown in Figure 13. Using the 25% criterion for detecting differences between images, which was used in obtaining thresholds $\hat{\varepsilon}$ from psychometric functions (Section 6), the results for environment illuminations (*grace*, *st.peters*) indicate that differences between the images could not be detected. This is also the case for point-light illumination of five of the samples, the exceptions being *alu*, *leather l.*, and *impalla*. Therefore, only in the case of these three samples, rendered under point-light illumination, is there evidence that the greater degree of compression achieved by the proposed method causes any reduction of image

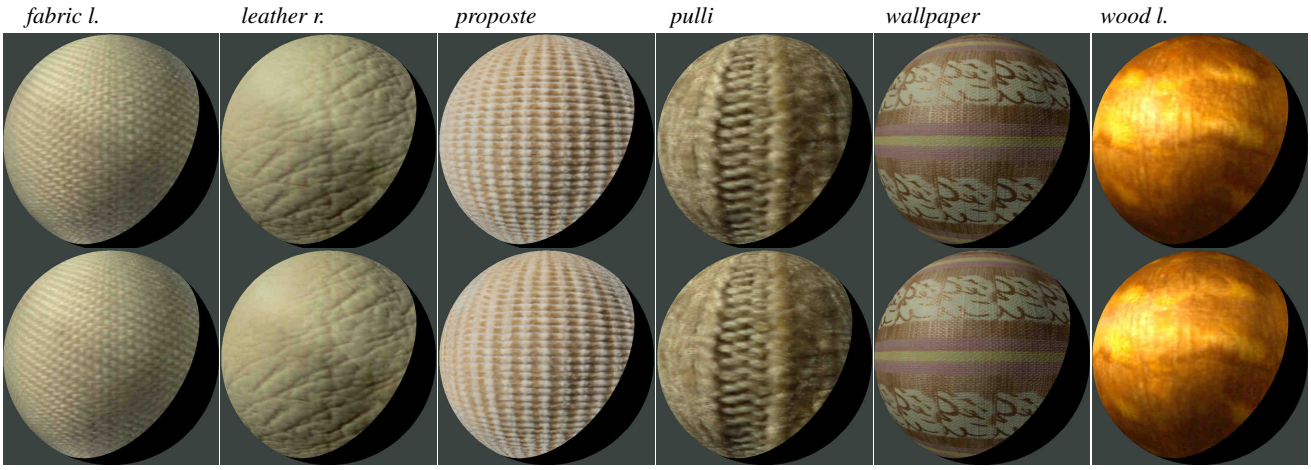


Figure 12: *Original renderings of validation samples (the top row) compared with their perceptually equivalent renderings from the subset of the sample (the bottom row), for point-light from left. The estimated thresholds ε_σ are shown in Figure 11.*

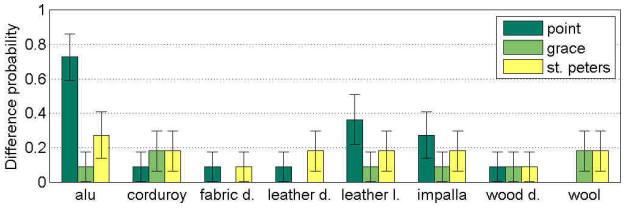


Figure 13: *Results of BTF compression verification study - average perceived difference for individual materials for point-light,* grace*, and* st.peters *illumination environments.*

quality. These results are illustrated in Figure 14. Here, renderings of these three materials over the *tablecloth* object, under point light illumination, are displayed. For each one, the results of using the full original BTF sample are compared with the results from the two levels of compression achieved with the LPCA and LPCA+ methods. The images indicate that, even in these worst cases, the effects of LPCA+ compression on image quality are slight.

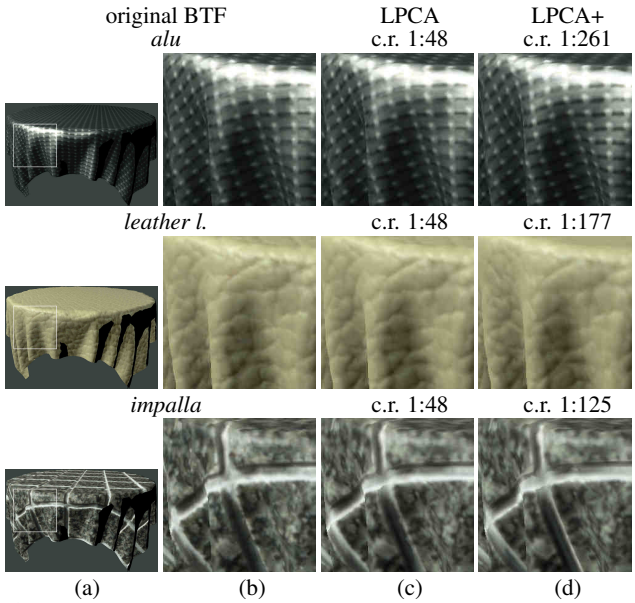


Figure 14: *Comparison of original BTF rendering (a),(b) with compression of full BTF sample (LPCA) (c) and its perceptually important subset only (LPCA+) (d).*

9.2 BTF Sampling

By examining the preserved combinations of illumination and viewing directions in the perceptually reduced BTF subsets, we can obtain interesting information about the perceptually optimal sampling of different types of materials when BTF data are acquired. Figure 15 shows the distribution of illumination and view directions of the preserved BTF subset as black dots in substitution maps for all tested samples. It is clear from the figure that, to obtain optimal results, different BTF samples require different sampling strategies, depending on the structural and reflectance properties of the underlying material. For most of the materials we can see a characteristic increase of sampling density in illumination directions opposite to the viewing direction (diagonal stripes). This suggests that the specular highlight is the main visual feature of such samples, particularly for *leather d.* and *leather l.*. In contrast, the sampling for *wood d.*, *corduroy*, and *impalla* is more uniform. In the case of *wood d.*, the material has a smooth surface and exhibits significant anisotropic reflectance, producing strong subsurface scattering effects. The other two samples (*corduroy* and *impalla*) have rough surface structure, (i.e, very variable surface height), and more uniform sampling is therefore required to preserve occlusions and shadows reliably. This would be the case for any samples having distinct wrinkles, holes or pointing in the surface. Some materials (*fabric d.*, *wool*) exhibit combinations of both uniform and specular sampling, where the latter prevails for higher elevations of illumination and viewing angles.

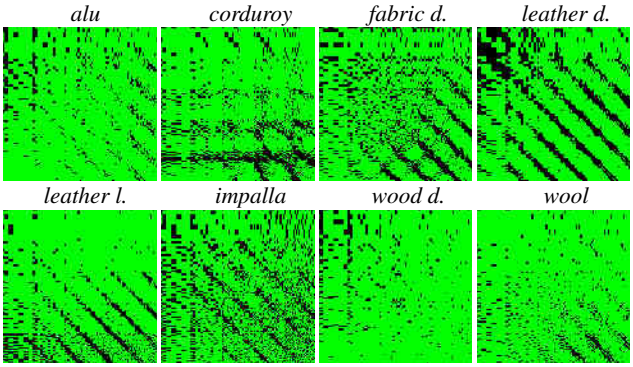


Figure 15: *Substitution maps for all tested samples, showing preserved combinations of illumination and viewing directions (in black) of the estimated perceptually important subset. Please refer to Figure 5 for map description.*

10 Conclusions and Future Work

In this paper, we have presented a novel application of psychophysical methods to the problems of analysis and optimal reduction of view and illumination dependent texture data (BTF). An analysis of eight different BTF samples shows that the average variance of BTF sample images is strongly correlated with the way that human observers judge the sample's degradation introduced by the proposed vector quantization of BTF images. Furthermore, we have proposed a general method for optimal, sample dependent setting of quantization thresholds. Although general application of the method requires an initial psychophysical experiment to estimate scaling values, these can then be used for automatically setting the thresholds of any new BTF sample. The method provides us with reliable and mostly conservative thresholds, as was validated using eighteen observers and six new samples. We believe that this is the first time that rigorous psychophysical investigation has shown that uniform sampling introduces significant perceptual redundancy into BTF data, and furthermore that this redundancy can be detected and efficiently removed using a simple statistical analysis of a BTF sample. Additionally, we have shown that by using reduced BTF data that provides a predefined level of perceived quality, a significant improvement in the performance of a PCA-based compression method can be obtained. We have increased the compression more than $4\times$ while reducing processing time also approximately $4\times$ compared to that required to compress the entire BTF dataset, without a difference in visual appearance being perceptible.

In future we intend to investigate the underlying physical properties of BTF samples to determine both the causes of their specific variances, and how our perception interacts with these properties.

Acknowledgments We would like to thank Bonn University for providing the BTF samples, V. Havran for help with environmental lighting, the subjects for participation in the experiments and the reviewers for their helpful comments. J. Filip was supported by the EC Marie Curie Intra-European Fellowship No.41358. This research was also supported by EPSRC grants GR/S12395, EP/F02553X/1, GAČR grant 102/08/0593 and 1ET400750407.

References

DALY, S. 1993. The visible differences predictor: an algorithm for the assesment of image fidelity. *Digital Images and Human Vision*, 179–206.

DANA, K. J., VAN GINNEKEN, B., NAYAR, S. K., AND KOENDERINK, J. J. 1999. Reflectance and texture of real-world surfaces. *ACM Transactions on Graphics 18*, 1, 1–34.

FILIP, J., CHANTLER, M., AND HAINDL, M. 2008. On optimal resampling of view and illumination dependent textures. In *5th Symposium on Applied Perception in Graphics and Visualization*, 131–134.

FLEMING, R. W., DROR, R. O., AND ADELSON, E. H. 2003. Real-world illumination and perception of surface reflectance properties. In *Journal of Vision*, vol. 3, 347–368.

HAINDL, M., AND FILIP, J. 2007. Extreme compression and modelling of bidirectional texture function. *IEEE Trans. on Pattern Analysis and Machine Intelligence, 29*, 10, 1859–1865.

HAVRAN, V., SMYK, M., KRAWCZYK, G., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2005. Interactive system for dynamic scene lighting using captured video environment maps. In *Eurographics Symposium on Rendering*, 31–42,311.

HO, Y. X., LANDY, M. S., AND MALONEY, L. T. 2008. Conjoint measurement of gloss and surface texture. *Psychological Science 19*, 2, 196–204.

KAUTZ, J., BOULOS, S., AND DURAND, F. 2007. Interactive editing and modelling of bidirectional texture functions. *ACM Transactions on Graphics 26*, 3, 53.

KHANG, B.-G., KOENDERINK, J. J., AND KAPPERS, A. M. L. 2006. Perception of illumination direction in images of 3-D convex objects: Influence of surface materials and light fields. *Perception 35*, 5, 625–645.

LAWSON, R., BÜLTHOFF, H. H., AND DUMBELL, S. 2003. Interactions between view changes and shape changes in picture - picture matching. *Perception 34*, 12, 1465–1498.

LEUNG, T., AND MALIK, J. 2001. Representing and recognizing the visual appearance of materials using three-dimensional textons. *International Journal of Computer Vision 43*, 1, 29–44.

MATUSIK, W., PFISTER, H. P. BRAND, M., AND MCMILLAN, L. 2003. A data-driven reflectance model. *ACM Transactions on Graphics 22*, 3, 759–769.

MESETH, J., MÜLLER, G., KLEIN, R., RÖDER, F., AND ARNOLD, M. 2006. Verification of rendering quality from measured BTFs. In *3rd Symposium on Applied perception in Graphics and Visualization*, 127–134.

MÜLLER, G., MESETH, J., AND KLEIN, R. 2003. Compression and real-time rendering of measured BTFs using local PCA. In *Vision, Modeling and Visualisation 2003*, 271–280.

MÜLLER, G., MESETH, J., SATTLER, M., SARLETTE, R., AND KLEIN, R. 2005. Acquisition, synthesis and rendering of bidirectional texture functions. *Computer Graphics Forum 24*, 1, 83–110.

OSTROVSKY, Y., CAVANAGH, P., AND SINHA, P. 2005. Perceiving illumination inconsistences in scenes. *Perception 34*, 1301–1314.

PADILLA, S., DRBOHLAV, O., GREEN, P., SPENCE, A., AND CHANTLER, M. 2008. Perceived roughness in $\frac{1}{f^\beta}$ noise surfaces. *Vision Research 48*, 1791–1797.

PELLACINI, F., FERWERDA, J. A., AND GREENBERG, D. P. 2000. Toward a psychophysically-based light reflection model for image synthesis. In *27th International Conference on computer Graphics and Interactive Techniques*, 55–64.

RAMANARAYANAN, G., FERWERDA, J., WALTER, B., AND BALA, K. 2007. Visual equivalence: towards a new standard for image fidelity. *ACM Trans. on Graphics 26*, 3, 76:1–76:10.

SOMOL, P., AND HAINDL, M. 2005. Novel path search algorithm for image stitching and advanced texture tiling. In *Computer Graph., Visual. and Computer Vision, WSCG05*, 155–162.

SUEN, P., AND HEALEY, G. 2000. The analysis and recognition of real-world textures in three dimensions. *IEEE Transactions on Pattern Analysis and Machine Intelligence 22*, 5, 491–503.

TE PAS, S. F., AND PONT, S. C. 2005. A comparison of material and illumination discrimination performance for real rough, real smooth and computer generated smooth spheres. In *2nd Symp. on Applied Perception in Graphics and Visualization*, 57–58.

TE PAS, S. F., AND PONT, S. C. 2005. Estimations of light-source direction depend critically on material BRDFs. *Perception, ECVP Abstract Supplement 34*, 212.

VANGORP, P., LAURIJSSEN, J., AND DUTRE, P. 2007. The influence of shape on the perception of material reflectance. *ACM Transactions on Graphics 26*, 3, 77:1–77:10.

WICHMANN, F. A., AND HILL, N. J. 2001. The psychometric function: I. fitting, sampling, and goodness of fit. *Perception & Psychophysics 63*, 8, 1293–1313.