Appearance Bending: A Perceptual Editing Paradigm for Data-Driven Material Models

M. Mylo\textsuperscript{1, 2}, M. Giesel\textsuperscript{3}, Q. Zaidi\textsuperscript{4}, M. Hullin\textsuperscript{1} and R. Klein\textsuperscript{1}

\textsuperscript{1}Institute of Computer Science II, Computer Graphics, Universität Bonn, Germany
\textsuperscript{2}Fraunhofer FKIE, Germany
\textsuperscript{3}School of Psychology and Neuroscience, University of St Andrews, UK
\textsuperscript{4}Graduate Center for Vision Research, State University of New York College of Optometry, USA

Abstract

Data-driven representations of material appearance play an important role in a wide range of applications. Unlike with analytical models, however, the intuitive and efficient editing of tabulated reflectance data is still an open problem. In this work, we introduce appearance bending, a set of image-based manipulation operators, such as thicken, inflate, and roughen, that implement recent insights from perceptual studies. In particular, we exploit a link between certain perceived visual properties of a material, and specific bands in its spectrum of spatial frequencies or octaves of a wavelet decomposition. The result is an editing interface that produces plausible results at interactive rates, even for drastic manipulations. We present the effectiveness of our method on a database of bidirectional texture functions (BTFs) for a variety of material samples.

CCS Concepts

- Computing methodologies → Image processing; Image-based rendering;

1. Introduction

Accurate models of material appearance are one of the key ingredients to high-quality rendering. Due to the complexity of real-world materials, deriving such models from first principles is often unfeasible or impossible. Instead, more and more application fields are resorting to data-driven material representations that are based on reflectometric measurements and correspond to ray-space light transport tensors. Despite their practical benefits, such models are time-consuming to acquire, data-intensive to store, and nearly impossible to edit. Only very recently, researchers started developing approaches to manipulate, re-synthesize and interpolate between material models in a plausible way [RSK13].

With this work, we contribute to this emerging direction of research by introducing appearance bending, a family of image-based editing operators that work on tabulated reflectance data. Our technique builds upon recent findings from perception research [GZ13] that relate bands of spatial frequencies to certain perceived visual qualities of a material. We generalize this idea to higher-dimensional reflectance distributions and show that many aspects of material appearance, including mesostructure, local and nonlocal shading, are gracefully handled by our signal-processing edits.
Finally, we demonstrate edits of a variety of materials represented as bidirectional texture function (BTF) [DvGNK97]. The key insight of our work is that by modifying appearance data, rather than 2-dimensional imagery, much more drastic edits can be performed and still yield plausible results. Appearance bending is orthogonal to existing appearance editing approaches and can therefore be combined with texture synthesis and material interpolation.

2. Problem Definition and Related Work

Before we discuss the works that served as inspiration and foundation for appearance bending, we will introduce the problem to which we are aiming to contribute. Our focus lies on advanced representations of spatially varying material appearance, and in particular the bidirectional texture function (BTF), which is capable of capturing a wide range of optical phenomena including nonlocal shading and volumetric effects. The BTF $B$ is defined as [DvGNK97]

$$B : A \times \Theta_i \times \Theta_\omega \rightarrow \mathbb{R} \quad (1)$$

where $A$ is the spatial domain, i.e., the extent of a material, and $\Theta_i$ and $\Theta_\omega$ are the space of all (incident) lighting and (outgoing) viewing directions, respectively. $B(\vec{x}, \vec{\theta}_i, \vec{\theta}_\omega)$ is the amount of light scattered at point $\vec{x}$ from direction $\vec{\theta}_i$ into direction $\vec{\theta}_\omega$. Like other data-driven models for material appearance, a BTF is a high-dimensional collection of sampled data and therefore not easily factorized into meaningful components like color, reflectance, texture and mesostructure. Consequently, the development of decompositions and user interfaces for the editing of such material models is still an active field of research.

Our goal is to provide an interface for the user to perform certain directions of manipulation that produce plausible and predictable results. By *plausible*, we mean that moderate amounts of bending should produce an outcome that is again recognized as a material. The *predictability* of an edit should be defined not in physical terms, but rather by psycho-visual perception of certain material qualities. It turns out that the key to meeting both requirements is in the combination of an advanced appearance representation (the BTF) with the findings from recent perceptual studies, namely a connection between perceived material affordances and signal processing made by Giesel and Zaidi [GZ11, GZ13]. In the following, we will summarize existing approaches to editing BTFs and SVBRDFs, and then outline the key points of Giesel et al.'s work.

2.1. Editing Data-Driven Appearance Models

Since the BTF describes material reflectance by its spatial and angular variation alike, it can be approached either as a collection of spatially varying angular reflectance distributions, or as a collection of textures that vary by angle.

The majority of prior work approximates material appearance in terms of analytical models. Those analytical models are mostly based on spatially varying BRDFs (SVBRDFs) that are associated with the geometry of the object surface. Following the pioneering work by Lensch et al. [LK01] to recover SVBRDFs in a practical way from a small number of input images, other researchers went on fitting similar models to BTF data [RK09, WDR11]. Since SVBRDFs readily separate reflectance from geometry, they lend themselves to a number of editing techniques. Among the approaches proposed are the transfer of reflectance functions from one measured material to another [ATDP11], the use of one-dimensional nonlinear appearance manifolds to simulate aging processes [WTL*06], as well as techniques using graph-based [PL07] and low-rank representations [AP08] or deep neural networks [EIKM16] to propagate edits over similar regions on a surface. These are the works that we consider most representative for a larger body of prior art in SVBRDF fitting and editing. Due to the underlying model assumptions, all of them share similar problems with materials with a complex surface structure that may not adequately be represented as SVBRDFs. Haindl and Havlíček [HH17] suggest a stochastic process based BTF-model.

A second family of techniques deals with BTFs as with textures. For instance, Xu et al. [XWT*09] exploited the fact that An and Pellacini’s Appearance Propagation [AP08] is not bound to a special representation by transferring it to textures and BTFs. Kautz et al. [KJ07] showed that operators from picture editing to the spatial or the angular domain of a BTF may bring reasonable results. Müller et al. [MSK07] presented a texture synthesis approach for BTFs that guides the placement of local features using a given mesostructure constraint. More recent publications [HLLC17] concentrate on the usability of the editing scheme. An algorithm that is particularly suited for repetitive textures was introduced by Haindl and Hatka [HH05] with their BTF Roller. Closely related to our frequency based material bending scheme is the recently published band setting scheme by Boyadzhiev et al. [BBPA15].

For the compression of BTFs, which is used in our algorithm, we refer the reader to the exhaustive state-of-the-art survey by Filip and Haindl [FH09].

An up-to-date state of the art report has been published by Schmidt et al. [SPN*15].

2.2. Visual Perception of Material Affordances

With appearance bending, we introduce an approach to manipulating BTFs that is orthogonal to existing texture synthesis and editing techniques. It is based on simple signal processing operations that are rooted in perception research. We will spend the remainder of this section on an overview of Giesel and Zaidi’s work [GZ11,GZ13] that serves as the foundation for our technique.

The term *affordance*, as used in psychology [Gib86], refers to an idea of use that is communicated by an object. Affordances are not only transported by the shape of the object, but also by the material it consists of [Adc01] – the distinction between “hard” and “soft” materials or “water-absorbent” and “water-repellent” materials can make all the difference for the use case that they suggest to us. According to Giesel and Zaidi, there is a strong connection between the frequency content of a material and some of its affordance channels. The authors conducted extensive psychophysical experiments on textures of 150×150 pixels, and concluded that specific frequency bands in monocular (2D) textures correspond to unique material properties, and that simple amplification or attenuation of those bands leads to high-level appearance alteration:
Table 1: Three frequency bands, and the opposing material properties that they correspond to.

<table>
<thead>
<tr>
<th>Frequency [cpd] (cycles per degree)</th>
<th>Associated opposing properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.57 – 2.29</td>
<td>straight – undulated</td>
</tr>
<tr>
<td>2.29 – 4.28</td>
<td>flat – thick</td>
</tr>
<tr>
<td>6.57 – 15.14</td>
<td>soft – rough</td>
</tr>
</tbody>
</table>

3. Appearance Bending: Scale-Space Manipulation of Materials

Mathematically speaking, our goal is to transform a given BTF $B$ into an edited, or “bent”, BTF $B^E$, a process we abstract as a bending operator $E$:

$$B^E = EB$$  \hspace{1cm} (2)

$E$, in turn, is composed as the chain

$$E := R^{-1} \circ M^{-1} \circ D \circ M \circ R$$  \hspace{1cm} (3)

where $R$ is a range transform and $B$ is an editing operator, which consists of a multiplicative kernel $D$ acting pointwise on the coefficients in the basis $M$. In the following, we will motivate and discuss the individual components of this chain, and how we address the differences and technical challenges that arise when marrying Giesel and Zaidi’s frequency-based operations to data-driven reflectance models.

3.1. Choice of Basis and Scale

When it comes to selecting the basis in which the editing shall be performed, the most straightforward choice is a Fourier transform ($M = F$), which was also used by Giesel and Zaidi. This means that certain frequency components will be amplified or attenuated globally. In Section 3.5, we will see that Fourier editing may fail for certain multi-component materials. Thanks to the generality of appearance bending, however, it is possible just as well to operate in a wavelet basis, also enabling local manipulations of inhomogeneous materials. For the time being, let us constrain our discussion on Fourier editing, i.e., $M^{-1} \circ D \circ M$ amplifies certain Fourier coefficients linearly.

The kernel $D$ performs a pointwise multiplication in frequency domain, with a gain factor $k$ that determines the amount of effect desired:

$$D^k(v) = \begin{cases} k, & \text{if } v_{\min} < |v| < v_{\max} \\ 1, & \text{otherwise} \end{cases}$$  \hspace{1cm} (4)

The kernel’s ring-shaped support in Fourier domain (Figure 2) needs to be derived from Table 1. For design purposes it may be advised to use a less abrupt transition between edited and non edited regions by applying an appropriate filter. While Giesel and Zaidi used textures of 150 px × 150 px subtending 3.5° of visual angle, we want to make use of the full native resolution of the BTF database (346 px × 346 px) for Fourier editing, and cropped versions (256 px × 256 px) for “power-of-2 convenience” in wavelet editing. At a resolution of 6.92 px/mm (160 dpi), this corresponds to sample sizes of 50 mm and 37 mm, respectively, and roughly the same visual angle when viewing the real-world samples at 70 cm distance. This leads to a conversion factor of 12.211 mm/° to transform the original frequency bands to our working resolution:

<table>
<thead>
<tr>
<th>Frequency [cpm] (cycles per mm)</th>
<th>Associated opposing properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0467 – 0.1874</td>
<td>straight – undulated</td>
</tr>
<tr>
<td>0.1874 – 0.3503</td>
<td>flat – thick</td>
</tr>
<tr>
<td>0.5378 – 1.2392</td>
<td>soft – rough</td>
</tr>
</tbody>
</table>

We call the corresponding kernels $D_h$ (straighten–undulate), $D_i$ (flatten–thicken), and $D_r$ (soften–roughen).

Although the Fourier spectrum of a texture is complex-valued, this manner of editing will not introduce an imaginary component to the signal, as long as the kernel is symmetric about the origin ($v = 0$).

3.2. Dimensionality of Data

The black-box character of a BTF (essentially, a 6-dimensional table of reflectance samples) makes it particularly hard to edit. In accordance to Giesel and Zaidi’s 2-dimensional procedure, we treat the BTF as a collection of 2D textures which are processed independently of each other. The consistency of this approach is guaranteed by the structure-preserving nature of the editing operator (see Section 3.3) and the linearity of the rendering step. In Section 4, we show that, despite its simplicity, this approach is by far superior to prior work on manipulating single 2D textures. By editing a relightable representation of the material, much more realistic results can be attained.

3.3. Color Spaces and Dynamic Range

As in Giesel and Zaidi’s perceptual studies, we perform our manipulations on the intensity channel in YUV color space. Contrary to their example, however, our data is of much higher dynamic range, and so the same arithmetic produces significant artifacts (Figure 3). Therefore, we use the range transform operator $R$ to transform the BTF data to a logarithmic scale before the editing step, and transform the result back by its inverse, an exponential function. This is in accordance with the observation from vision research that the human visual system performs a logarithmic range compression of the incoming optical signal.

© 2017 The Author(s)

With this, we now have all the required ingredients to express the full bending operator as per Equation 3 as

$$E_{\{u,t,r\}} = \text{Exp} \circ F^{-1} \circ D_{\{u,t,r\}} \circ F \circ \text{Ln}$$ (5)

### 3.4. Compression

Appearance bending seamlessly integrates into systems operating on compressed data, provided that the decompression operator is linear. Suitable compression schemes include full matrix factorization [KMBK03] or Müller’s decorrelated full matrix factorization [Müller09]. The tabulated data are unrolled into a two-dimensional matrix where textures form the columns and the local reflectance distributions (apparent BRDF, short: aBRDF), form the rows. To compress the BTF, the PCA is calculated and the resulting eigenbase is truncated by setting a certain threshold.

Due to the linearity of the reconstruction step, the bending operator can be applied directly to the eigen-textures. For the lower frequencies, this yields the same results as manipulating the whole base. Manipulation of higher frequencies may suffer from the loss of such frequencies (fine-scale details) in compression.

### 3.5. Wavelet-Based Editing

Thus far, our discussion of appearance bending has been focused on Fourier-domain edits. As it turns out, for many textures with a certain degree of randomness, the Fourier domain is a good choice. For others, however, we have become aware that the global support of the basis functions can cause objectionable artifacts in the form of ringing – textures consisting of multiple basis materials are particularly susceptible. A wavelet basis is better suited to these situations.

Contrary to the tight spacing of frequencies in Fourier domain, wavelets are organized in octaves. In order to emulate the effect of the main frequency bands as given by Giesel and Zaidi [GZ13], we round the given cutoff frequencies upwards to the nearest octave, avoiding that the resulting bands overlap. For a 256px × 256px image corresponding to an edge length of 3.7 cm, we come up with the following octave bands:

- 0.5 0.66 1.5 2
- 0.65 0.75 1.4 1.8
- 1.0 – 3
- 0.0467 – 0.1874
- 0.7 0.85 1.4 1.7
- 0.5378 – 1.2392
- 4 - 5
- 0.85 0.9 1.2 1.3
- 0.1874 – 0.3503
- 5

We compared different wavelets and obtained the best results with wavelets from the Daubechies-series (D). To obtain a scaling that most closely resembles that of Fourier-domain edits, we employed Mantiuk’s HDR-VDP-2 predictor [MKRH11] to formulate a minimization problem:

$$k_{\text{opt.}} = \arg \min_k ||W E_{\{u,t,r\}} k - F E_{\{u,t,r\}}||_{\text{HDR-VDP-2}}$$ (6)

We sampled the range of $k$ for Fourier editing at four values (0.5, 0.66, 1.5 and 2.0), and found the following scalings to be appropriate:

<table>
<thead>
<tr>
<th>Operator</th>
<th>Freq band [cpmn]</th>
<th>Octave</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undulation</td>
<td>0.0467 – 0.1874</td>
<td>1 - 3</td>
</tr>
<tr>
<td>Thicken</td>
<td>0.1874 – 0.3503</td>
<td>4 - 5</td>
</tr>
<tr>
<td>Roughen</td>
<td>0.5378 – 1.2392</td>
<td>5</td>
</tr>
</tbody>
</table>

Linear regression reveals the following approximate relations between the Fourier and the wavelet gain factors:

\[
\begin{align*}
W_k u & \approx 0.66 (F_k u - 1) \\
W_k t & \approx 0.77 (F_k t - 1) \\
W_k r & \approx 0.31 (F_k r - 1)
\end{align*}
\]

Figure 4 shows plots of the HDR-VDP-2 error as a function of gain factor and wavelet type. Our experiments suggest to use wavelets with a number of vanishing moments greater or equal to 6 for the “roughen” operation; for “thicken” and “undulate”, vanishing moments of 3 and 2, respectively are sufficient. We note that higher vanishing moments increase the spatial support, so we use the lowest-order wavelet that produces no visible artifacts.

As expected, wavelets can act as a good replacement for the Fourier basis (Figure 5), and their use pays off particularly for materials with strong spatial variation (Figure 6).

### 4. Results

In this section, we present a variety of results that were obtained using appearance bending. Although we process the textures purely in a signal processing framework, the edits often appear to correspond...
Figure 5: A comparison of Fourier- and wavelet-edited materials. Note that the strength of the editing operator, $k$, has been adapted to obtain the best possible match according to HDR-VDP-2 [MKRH11].

Figure 6: Comparison of Fourier (top) and wavelet (bottom) editing for the “roughen” operator. From left to right: $k = 0.5$, $k = 1$ (identity), $k = 2$. Note the absence of regularly patterned artifacts in the wavelet results, even for a gain factor of $k = 2$.

Figure 7: $E_R$ applied to a wool material to a semantic meaning. Figure 7, which shows a wool material processed with the “roughen” operator at various gain levels. The edit amplifies the structure of the knitware on the finest scale, conveying an impression of fiber-level detail. Figures 8 and 10 show various operators applied to a selection of material classes ranging from grainy leather over textiles to stone and wallpaper. It becomes evident that not every material responds to each operator in the exact same way – however, within a class of materials, the results are consistent (Figure 9(a)). We note that uncompressed materials contain more high-frequency details and are therefore better bending candidates - at an increased computational cost.

4.1. Extreme edits

In Figure 8, we see the editing result of a material before and after editing in two different presentations: as a flat texture and on a 3D material with environment lighting. The 3D version is more forgiving, producing a plausible impression of a material even for extreme edits. We attribute this to the generally increased realism of the scene, but in particular to the low-pass nature of the environment illumination.

4.2. Comparison with image based editing

In this section, we want to compare the result of editing materials to directly editing on the rendered image. Of course the bending operators might have been applied directly to the rendered images. But then the result of the operations depended on the distance of the material surface from the camera and on the angular between the material surface and the image plane. This effect shall be demonstrated for the undulation operator.

In the left column of figure 9, we rendered the material probe first and applied the undulation operation afterwards. In the right column we applied the operator to the material base vectors. Shifting of the frequency window has been done appropriately as you can see in the yellow framed image sections: the results are nearly indistinguishable. But the yellow framed image sections show that pure manipulations of the image spectrum may not account for perspective distortion.
4.3. Performance

On compressed BTF data with 100 (Y), 50 (U) and 50 (V) principal components, data processing took no more than 0.5 s (Fourier), 1.2 s (D1 wavelet) and 4.8 s (D6 wavelet). For uncompressed data, the time scales up linearly with the number of textures, resulting in processing times of about 5 minutes for applying the a Fourier-based operator and 12 to 45 minutes for wavelet-based edits. All experiments have been made on an Intel Dual Core 6600.

5. Discussion and Future Work

The edits performed by appearance bending do, of course, not change the surface geometry. When moving an object that is textured with a strongly edited BTF, the parallax effect (or lack thereof) may result in a lack of realism.

In the current version, appearance bending can only strengthen or weaken such qualities that are present in the original material. The transfer of frequencies from one material to the other would be an interesting prospective and will be the subject of further investigation.

It would also be worthwhile to establish a stronger relation between the frequency band manipulations performed to the material, and the filtering occurring in the visual cortex [Dau79]. It remains to be seen to which extent a deeper understanding of human perception can be used to further increase the meaningfulness and realism of editing interfaces for appearance.

6. Conclusions

Appearance bending, or the editing of materials by scaling bands of spatial frequencies, has proven to be a versatile tool. The transfer from images to complex material representations such as the BTF yields results that match the original data in realism. We are hopeful that further developments in psycho-visual perception will further contribute to advance the state of the art in semantically meaningful and predictable editing processes for appearance.

References


The undulation operator applied to different materials.

(a) Different edits on the same material (blue wallpaper). From top to bottom: undulate, thicken, roughen.

(b) Effect of the “roughen” operator on three different materials. From top to bottom: Wool, Corduroy, Carpet.

**Figure 10:** Result renderings. For a discussion, see Section 4.