Low Cost Rapid Acquisition of Bidirectional Texture Functions for Fabrics

Banafsheh Azari
CogVis/MMC,
Faculty of Media,
Bauhaus-University Weimar
Bauhausstrasse 11
99423 Weimar, Germany
banafsheh.azari@uni-weimar.de

Sven Bertel
Usability,
Faculty of Media,
Bauhaus-University Weimar
Bauhausstrasse 11
99423 Weimar, Germany
sven.bertel@uni-weimar.de

Charles A. Wüthrich
CogVis/MMC,
Faculty of Media,
Bauhaus-University Weimar
Bauhausstrasse 11
99423 Weimar, Germany
charles.wuethrich@uni-weimar.de

ABSTRACT

Creating photo realistic images from real world complex materials is a great challenge: the reflectance function of materials, especially fabrics, has glossy or specular highlights, reflectance anisotropy and retroreflections. This increases greatly the complexity of rendering. Bidirectional Texture Functions (BTFs), i.e. 2D textures acquired under varying illumination and viewing directions, have been used to render complex materials. However, the acquisition of textures for BTF requires up to now expensive setups and the measurement process is very time-consuming as the directional dependent parameters (lighting and viewing directions) have to be controlled accurately. This paper will present in detail a new low cost programmable device for the rapid acquisition of BTF datasets. The device allows to acquire BTF databases at a fraction of the cost of available setups, and allows to experiment when a texture resolution and sample density increase in the parameter space is not perceivable by an observer of the renderings. The paper proves that using smaller resolution textures and decreasing the samples in parameter space does not lead to a loss of picture quality.

Keywords
Bidirectional Texture Function Acquisition, Photorealistic Rendering, Image Quality, Hardware Devices.

1 INTRODUCTION

Rendering real world fabrics has been one of the big challenges in Computer Graphics. Fabrics exhibit a complex reflection behaviour, which depends, among other factors on the meso- and micro-structures of the thread, on the type of weaving, which influences the position of the thread in the fabric, on the interreflections between the components of the fabric, and on surface and subsurface scattering of light. For highly realistic material rendering the reflectance of the fabric must be simulated accurately. The amount of light reflected by a material is not only dependent on the light’s incident direction, but also on the angle from which a viewer looks at the fabric. Most woven materials have small three-dimensional details so that some part of surface can be occluded to the viewer depending on his position, a phenomenon which is called self occlusion. For materials exhibiting small-scale bumps, depending on the illuminant position, self shadowing can occur: surface irregularities cast shadows onto other parts of the surface. Moreover, incoming light gets reflected, refracted and scattered by the fibers in the thread, leading to an even more complex reflection behaviour.

Therefore, the surface reflectance of a material is often modeled by a four dimensional function which is dependent on the viewing and lighting vector. Two additional dimensions are required for an exact rendering function due to thread color patterns and visible geometry variances due to the weaving or knitting pattern. A material function capturing all these effects has six dimensions and is therefore neither easy to design nor to evaluate efficiently [1]. In theory, all these components need to be modeled individually to simulate the reflection behaviour of the fabric.

Bidirectional Texture Functions (BTFs), introduced by Dana et al. [2], represent an alternative solution to exact rendering: instead of doing the complex modeling, pictures of the fabric taken at different modeling and viewing directions are used as textures for rendering, implicitly integrating into the rendering step the reflectance properties of the surface. A BTF contains all information on reflectance of a set of points of a surface.
under a particular lighting and viewing condition, and is described as follows:

$$S_{\text{BTF}} = \int_{p \in P} (\theta_i, \phi_i, x_p, y_p, \theta_o, \phi_o) \delta p,$$

where $p$ is a discrete point of the sampled surface $P$, the variables $\theta_i$, $\phi_i$, $\theta_o$, and $\phi_o$ are the angles describing the direction of the incoming and the reflected light ray ($L_i$, $L_o$) and $N$ is the surface normal. Furthermore the parameter $x_p$ and $y_p$ describe the point of incidence and reflectance of the light ray. Given a BTF with varying $\theta_i$, $\phi_i$, $\theta_o$, and $\phi_o$, one can accurately describe the reflection behavior of a sampled surface area.

In practice, BTFs use large collections of digitally acquired pictures of a material taken at discretely varying illumination and viewing angles. When a simulation of the material needs to be computed for rendering, the viewing and illumination vectors are used to pick matching textures from the collection of scanned textures, and, if the angles do not match with angles of the corresponding textures, neighbouring textures are interpolated at the point to be rendered.

A big disadvantage of BTFs is that state of the art measurement devices require expensive robotics setups and that the measurement process is very time consuming since direction dependent parameters (light- and view-direction) have to be controlled accurately. Otherwise the resulting data will be poor. Moreover, the size of BTF data can range from hundreds of megabytes to several gigabytes, since in the ideal case a large number of high resolution pictures have to be used. For real time rendering this is a big disadvantage, since either the entire collection of pictures needs to be kept in the computer memory, or computationally expensive methods have to be used to intelligently load/unload the textures. In this context, several authors focused on efficient compression methods for BTF data (including reflectance models based on linear factorization and pixel-wise bidirectional reflection distribution functions, in short BRDFs, which are the general reflection model from which BTFs are derived [3]).

The focus was rarely set on the perceived quality of the results of compression or loading/unloading mipped-mapped textures [4–12], and while the existing approaches are technically well motivated, we believe that before starting to choose how and how strongly BTF data should be compressed, it makes sense to first take a step back and see how many measured samples at which resolution are needed to have the same perceived quality when rendered instead of using a complete database at the highest possible resolution, or automatically degraded texture downscalings not taking into account the final user. If the texture database is perceptually sound, then it can be reduced in its number of texture samples.

In this paper we present two basic improvements to the use of BTFs for rendering: firstly we want to address the cost of BTF acquisition by introducing a flexible low cost step motor setup for BTF acquisition allowing to generate a high quality BTF database taken at user defined arbitrary angles. Secondly, we want to adapt the number of acquired textures to the perceptual quality of the renderings, so that the database size is not overblotted and can fit better in memory when rendered. In order to do this, we will use Daly’s Visual Difference Predictor (VDP; [13, 14]) to prove that the reduced dataset acquired through our device does not lead to perceivable differences for the rendered images for a viewer.

In the next section, we will introduce how we plan to reduce the BTF database. Next the BTF measurement setup will be described in detail. Then an experimental evaluation of the BTF acquisition setup results will be presented, proving that no perceivable differences in the renderings are made by reducing the BTF database angle steps. Finally we will present some conclusions, and an outlook.

## 2 REDUCING THE SAMPLE DENSITY

In [10–12] the authors propose to reduce the BTF dataset size by down-sampling resolution and view/illumination angles and proved that perceived quality did not decrease. The outcomes could help prevent capturing redundant images with high resolution from a sample and this will reduce the acquisition time significantly. We propose a preprocessing step before starting to acquire the complete database to determine the down-sampling threshold. As this threshold depends on the surface characteristics of a material, each sample should be tested individually.

The first step in the proposed preprocessing method is to generate solely samples required to texturing a section of a sphere, as shown in Table 1. The produced database therefore covers a fourth of the BTF database (22*22=484 samples) as opposed to a complete BTF database (81*81=6561 samples), which we will refer to as ‘BTF’. In the next step this database is down-sampled using reduced densities and resolution
Four down-sampling schemes are adopted. In the first scheme we reduced the resolution of the each texture from 265 x 256 to 128 x 128; it will be referred in this paper as ‘BTF-R’. In order to obtain considerable reduction of BTF dataset size we adopted two different BTF sampling schemes denoted as A, B from [11]. While scheme ‘A’ preserves the original sampling of elevation angle \( \theta \) but reduces the number of azimuthal samples along angle \( \phi \), scheme ‘B’ reduce sampling for both angles (see Table 2).

It should be noticed that BTFs require directional sampling of both illumination (\( \theta_i, \phi_i \)) and view directions (\( \theta_o, \phi_o \)) and in these two directions different sampling schemes can be adopted without limiting practical usage of the data. We choose three down-sampled BTF datasets. The first two are straightforward and down-sampled both illumination and viewing directions in the same way, using a combination of the schemes AxA and BxB, which we will refer to as ‘BTF-A’ and ‘BTF-B’. The third one used scheme B on just view directions (‘BTF-C’). Consequently, four down-sampled datasets are generated (see Table 3).

The samples in each databases are then used to render the sphere in order to check the influence of down-sampling on the perceived quality. This can be done either by using a Subjective Quality Metrics or an Objective Quality Metrics [15].

Since human beings are the users in most image-processing applications, the most reliable way of assessing the quality of an image is by using Subjective Quality Metrics. Indeed, the mean opinion score (MOS), a subjective quality measure requiring human observers, has been long regarded as the best method of image quality measurement. However, the MOS method is expensive, and it is usually too slow to be useful in real-world applications.

To solve the problem Objective Quality Metrics have been proposed. The goal of these metrics is to design

<table>
<thead>
<tr>
<th>Database</th>
<th>Number of Samples</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>BTF</td>
<td>(22 x 22)</td>
<td>256 x 256</td>
</tr>
<tr>
<td>BTF-R</td>
<td>(22 x 22)</td>
<td>128 x 128</td>
</tr>
<tr>
<td>BTF-A</td>
<td>(A x A)</td>
<td>256 x 256</td>
</tr>
<tr>
<td>BTF-B</td>
<td>(B x B)</td>
<td>256 x 256</td>
</tr>
</tbody>
</table>

Table 3: The Proposed BTF Database (BTF) compared with the down-sampled BTF Database using reduced resolution (BTF-R) and densities (BTF-A, BTF-B and BTF-C).
mathematical models that are able to predict the quality of an image accurately and automatically. An ideal method should be able to mimic the quality predictions of an average human observer.

One of the most popular and widely used objective quality metric based on models of the human vision system is Daly’s Visual Difference Predictor (VDP; [13, 14]). We used VDP to assess visual differences between the rendered spheres by different down-sampling schemes and to find a compression threshold. Based on this threshold a compressed BTF database could be acquired without capturing redundant images which reduce strongly the acquiring time. To test the method introduced above a measurement setup for the acquisition of BTF data has been built, which will be explained in detail in the next section.

3 ACQUISITION SETUP

The acquisition of 2D textures is a very simple process which can be performed using a standard 2D scanner or an off-the-shelf digital camera and image-processing software. On the contrary, the acquisition of BTFs requires a complex and controlled measurement environment. Since BTF acquisition can be seen as physical measurement of real-world reflection, special attention has to be paid to device calibration and image registration. Otherwise the measurements will contain inaccuracies which may generate visible rendering artifacts.

3.1 Prior Works

Dana et al. [2] built the first BTF measurement device. A robot arm is used in this device to orient the texture sample at arbitrary orientations and the camera and light orbit around the sample. 205 combinations of light and view directions are sampled for each material, and more than 60 materials have been measured and published1. Due to the sparse sampling, it is not practical to use the measured data for rendering directly.

More recently, researchers have built similar setups and provided measurements at higher angular resolutions [16–18]. Significantly to the gonioreflectometer case where one value is measured at a time for each lighting and viewing directions, however, unlike the in gonioreflectometer case where one value is measured per setting, in these newer methods each sample is a texture.

For a fast high quality acquisition of BTFs Müller et al. [19] propose an array of 151 digital still cameras mounted on a hemispherical gantry. The on-camera flashes serve as light source. By synchronizing the cameras, 151*151=22801 images can be captured in 151 time steps and the authors report a measurement time of about 40 minutes. In this setup no moving parts are needed. Hence, the region of interest is known for every camera and consequently, there is no need for a time-consuming detection of the region of interest. While this is a big improvement in terms of measurement time, the setup is large and expensive.

Han and Perlin [20] introduced a measurement setup based on a kaleidoscope which allows viewing a sample from multiple angles at the same time through multiple reflections. Illumination is provided by a projector pointing into the kaleidoscope. By selectively illuminating a small group of pixels, the light direction can be controlled. Since there is no moving part in this setup, measurement is very fast. However, the equipment is difficult to build and calibrate. In addition, due to multiple reflections in the optical path, the resulting quality tends to be rather low.

Dana and Wang [21] proposed a setup based on a parabolic mirror. While their setup can provide higher quality measurements than the kaleidoscope setup, they can only capture a single spatial location at a time. As a result, it does not offer any acceleration compared to the gonioreflectometer-like approaches.

The standard gonioreflectometer-like approaches allow to capture high–quality BTFs reliably. Their drawback is the speed - several hours are needed and this makes measured BTFs an expensive resource. Using mirrors may be a promising approach in the future, but the quality of the measurements of current systems remains dubious.

Using a camera array sensor greatly reduces measurement times at the expense of the costs for a large number of cameras. In the next section the proposed measurement device will be introduced in detail.

3.2 The Proposed Measurement Device

During the measurement the light source and the sensor are positioned at various angles covering the entire hemisphere above a flat sample of a homogeneous material. In other words, the system allows acquiring images from all possible angles of illumination and of camera perspective. The proposed device (Figure 2) is the result of our attempts to find a setup covering the 4 degree of freedom available in a BTF.

We set our reference coordinate as shown in Figure 2. The origin is placed on the center of the sample. The sample can rotate about the x, y and z axes. While the light can rotate about z-axis using a step motor, the camera is fixed. The camera and light are directed to the center of the sample. The system has 4 degree of freedom and is appropriate for anisotropic material. To rotate the sample we decided to use a combination of three step motors [1-3], (see Figure 2).

With these motions three degree of freedom are achieved ( φ, θ, ϕ). To reach the additional degree

1 http://www1.cs.columbia.edu/CAVE/software/curet/
of freedom $\theta_i$, the light should rotate in the altitude direction. For this reason, we mounted our light source on an axes and rotate it with a wheel which can move with the help of the step motor number 4 (see Figure 2). The length of the light radius is adjustable.

The system is composed of different parts. The main component is an Arduino Mega 2560 which is equipped with a RAMPS 1.4 board and the Marlin operating system. The Arduino takes commands from a host PC and controls the motors and the remote control of the camera. This is done by using a serial connection between the Arduino and the host PC. The commands are transmitted as Gcodes$^2$.

**Hardware:** According to the hardware producer, the relationship between the voltage on the potentiometer and the motor current is given by,

$$A = \frac{V_{ref}}{8 \cdot R_s}$$

where $A$ is the motor current and $V_{ref}$ is the voltage on the potentiometer: in our case the drivers have a resistance $R_s$ of 0.1Ω. This formula is driver specific: if another driver is used, the values have to be updated or another formula could be necessary. The system uses four step motors to move the sample in all directions.

We decided to choose SM42051 and E7126-0140 step motors$^3$. The SM42051 has 0.196 Nm torque with max rated current of 0.6 A and is used to rotate the sample around the x and y axes, which we will refer to as S1 and S2. Because of the higher friction force by rotating the light and sample about z axis two E7126-0140 step motors with 1.6 Nm torque has been chosen (S3, S4). The motors are connected to the according S1, S2, S3 and S4 axes of the RAMPS 1.4 board. S1, S2 and S3 move the sample while S4 moves the light source.

Each step motor has a 1.8° step resolution. Therefore in order to rotate the S1, S2 and S3 axis by 9°, 5 steps are needed. To achieve adequate leveraging, the S4 axis is equipped with a gear. The gear ratio is 9 : 120. Each tooth of the small gear wheel corresponds to a 3° movement of the light, which is the smallest possible movement of Z axis. To move the small gear wheel one tooth further, 22.22 motor steps are necessary. To have a reference, the S1, S2 and S4 axes have end-stops which are triggered whenever the corresponding axis reaches its maximum or minimum rotation.

**Software:** The system uses the 3D printer software Marlin as operating system. The Marlin firmware is customized for this purpose. Therefore, the file configuration header is changed at specific points, so that the

---

$^2$ Gcode is a control language for CNC (or Reprap) machines

$^3$ [http://www.emisgmbh.de/schrittmotoren.html](http://www.emisgmbh.de/schrittmotoren.html)
Figure 4: Objects rendered by the proposed BTF and down-sampled BTF.

The system is connected to the host PC using a serial connection and receives commands from that PC. As host software for the PC Pronterface is used. After connection to the Arduino, various commands can be send to the Arduino.

The camera is the center piece of hardware in our measurement setup. Therefore special attention was paid to choose the camera. We selected a Nikon D750 DSLR camera, a high-end and full format digital camera intended for professional photography. The camera captures the material sample’s appearance at different positions in raw format at a resolution of 6016 x 3375 pixels. A fixed length SIGMA camera lens (105 mm F/2.8) is mounted on the camera. Via an IR Remote Control the camera’s shutter is released.

The other important piece of hardware in the measurement setup is the light source therefore it should be selected carefully as well. The decision for a specific light source was based on the emitter geometry and the lamp’s photometric properties. An OSTAR-Lighting LED Light Source [Osram GmbH]. Besides a controlled environment and suitable equipment, we applied a number of standard algorithms to further increase the quality of the images that are used as input to our measurements. These algorithms are Geometric and Colorimetric Camera Calibration.

Geometric calibration involves the recovery of a camera’s extrinsic and intrinsic parameters. While the intrinsic parameters relate the camera’s coordinate system to the idealized coordinate system, the extrinsic parameters relate the camera’s coordinate system to a fixed world coordinate system and specify its position and orientation in space. The actual transformation of the camera’s lens system is described by its intrinsic parameters. To extract the feature points from the calibration images an implementation of the Harris detector [22] included in Bouguet’s camera calibration toolbox was used.

To achieve the best possible color reproduction, the camera has to be color calibrated as well. In order to relate the recorded color to well defined standards, color management systems have become a standard tool. Hereby, an image of a test target with well known properties (the Macbeth ColorChecker) was taken and processed in the same way as all later images.

After the measurement the raw image data has been converted into a set of rectified registered images. Registration is done by projecting all sample images onto the plane which is defined by the frontal view. In a final step, the textures are cut out of the raw reprojected images and resized appropriately. Finally the technique presented by Goesele et al. [23] was used to reduce the fixed pattern noise.

4 EXPERIMENT AND RESULTS

To generate high quality real world input data for appearance measurements a special purpose digital photo studio has been built. Special attention was paid to carefully control the illumination and image capturing conditions in order to be able to acquire exact data about the surface properties of samples using readily available digital camera technology. We set the distance of the light and camera to the sample to one meter and \( \Delta \theta \) to 15°. We choose two planar samples with the size of 10 x 10 cm², Cord-Brown and Cord-Red (shown in Figure 3). 484 raw images with the resolution of 6061 x 3375 were captured for each sample. After the measurement the raw image data are projected onto the plane which is defined by the frontal view (\( \phi = 0^\circ \), \( \theta = 0^\circ \)). To be able to conduct an automatic registration

\[ \text{http://www.vision.caltech.edu/bouguetj} \]
we have attached point markers visible at the corners of the sample holder in Figure 3. Consequently four down-sampled databases are generated out of each of these two databases: BTF-A, BTF-B, BTF-C and BTF-R (as explained in section 2). For each of the five texture data sets, a three dimensional textured model of a section of a sphere was rendered through the standard BTF rendering method [3] at a screen resolution of 1920x1084 pixels.

An objective quality metric introduced by Daly (VDP, [13, 14]) has been used to assess the perceived quality differences between objects rendered by complete and down-sampled databases. VDP simulates low level human perception for known viewing conditions (in our case: a resolution of 1920 x1080 pixels at an observer’s distance of 0.7m). Figure 5 shows the visually perceivable differences per image pair as predicted by VDP. Each pair consisted of a rendering using BTF dataset and one of four downsamped BTF datasets.

It can be seen that there are not a significant perceivable quality differences between BTF and BTF_A in both of the samples while the Cord-Brown react more sensitive to the down-sampling by BTF_B and BTF_C than BTF-Red. The last row of the Figure 5 shows that the resolution reduction from 256*256 to 128*128 is not perceivable in BTF-Red. According to this information for both of the databases it is possible to reduce the number of generated samples as scheme A without losing quality: this decrease the acquisition time to 26% of the acquisition of complete database. Because less pictures need to be taken, the acquisition time becomes 4.5 hours instead of 6. In Cord-Red the captured images could have the half of the resolution, which reduce the database size 50%.

5 CONCLUSION

In this paper we presented a new low cost programmable BTF database acquisition device based on standard off the shelf components, step motors, a semiprofessional camera and a standard LED illumination source capable of capturing high quality databases. The device cuts the cost of existing database acquisition setup by a factor of hundreds. Since the positions of the illumination source and the orientation of the sample to be acquired can be chosen at will and therefore cover all four degrees of freedom of the parameter space, the device allows to investigate if smaller databases obtained through undersampling the parameter space allow perceptually sound renderings which show no perceptual difference with respect to a higher sampling of the parameters space.

Daly’s VDP results show that in our case both the texture resolution as well as a reduction of the samples to 26% of the number of samples used in widespread databases do not deteriorate significantly the perceived quality. Furthermore, also the time spent in the acquisition of the database is also reduced to little more than one fourth.

The new device appears therefore to be an excellent compromise, cutting significantly the costs in the acquisition process (to approximately €400 plus the Camera and the lens). Moreover, its programmability allows to conceive new experiments aimed at understanding the limits at which increasing the number of samples in the database, as well as the resolution of the acquired textures makes little sense since the observer of the rendered objects does not perceive any differences.

In future work, we plan to use the device as a basis for new experiments aimed at shedding a light in the rela-
tionship between high quality rendering and the perception of observers of rendered images. This should be relevant for the Computer Graphics, Image Processing and Image Compression communities alike.

6 REFERENCES


