A Portable Setup for Fast Material Appearance Acquisition
A Portable Setup for Fast Material Appearance Acquisition

Jiří Filip  Radomír Vávra  Mikuláš Krupička

Abstract—A photo-realistic representation of material appearance can be achieved by means of bidirectional texture function (BTF) capturing a material’s appearance for varying illumination, viewing directions, and spatial pixel coordinates. BTF captures many non-local effects in material structure such as inter-reflections, occlusions, shadowing, or scattering. The acquisition of BTF data is usually time and resource-intensive due to the high dimensionality of BTF data. This results in expensive, complex measurement setups and/or excessively long measurement times. We propose an approximate BTF acquisition setup based on a simple, affordable mechanical gantry containing a consumer camera and two LED lights. It captures a very limited subset of material surface images by shooting several video sequences. A psychophysical study comparing captured and reconstructed data with the reference BTFs of seven tested materials revealed that results of our method show a promising visual quality. As it allows for fast, inexpensive, acquisition of approximate BTFs, this method can be beneficial to visualization applications demanding less accuracy, where BTF utilization has previously been limited.

Keywords—measurement setup, material appearance, BTF, ABRDF, visual psychophysics.

I. INTRODUCTION

Reproduction of the appearance of real-world materials in virtual environments has been one of the ultimate challenges of computer graphics. Therefore, methods of material appearance representation, acquisition, and rendering have already received a lot of attention. The required material representations depend on the complexity of the material’s appearance. They start with a bidirectional reflectance distribution function (BRDF) describing distribution of energy reflected in the viewing direction when illuminated from a specific direction. As the BRDF cannot capture a material’s spatial structure, it has been extended to spatially-varying BRDF (SVBRDF) describing the material’s surface appearance by means of a collection of independent BRDFs. This representation allows an already quite efficient approximation of material appearance, mostly based on a wide range of analytical BRDF models. However, the BRDF’s constraints – mainly light and view direction reciprocity – limits the applicability of SVBRDFs to nearly flat opaque surfaces. On the contrary, the bidirectional texture function (BTF) [1] does not share these restrictions due to simultaneous measurement of non-local effects in rough material structures, such as occlusions, masking, sub-surface scattering, or inter-reflections. A monospectral BTF is a six-dimensional function $BTF(x, y, \theta, \varphi, \theta_i, \varphi_i)$ representing the appearance of a material sample as a surface point with coordinates $(x, y)$ for variable illumination $I(\theta_i, \varphi_i)$ and view $V(\theta, \varphi)$ directions, where $\theta$ and $\varphi$ are elevation and azimuthal angles, respectively, as shown in Fig 1-a.

![Fig. 1. (a) A BTF parameterization over material surface, (b) a BTF pixel's BRFD unwrap into a 2D image.](image)

As the BTF data achieves photo-realistic representation of material appearances without the need for lengthy fitting or tweaking of parameters, it has high application potential mainly in areas requiring physically correct visualizations ranging from computer-aided interior design, visual safety simulations and medical visualizations in dermatology, to digitization of cultural heritage objects. The measurement of BTF is, due to its high dimensionality, very time- and storage-space-demanding. While the storage space issue cannot be easily resolved at the measurement stage and the data are always subject to compression and modeling, the duration of the measurement largely depends on the measurement setup design as well as on the types of the sensors and illumination used. To the best of our knowledge, a majority of the current BTF acquisition setups (except [2]) are based on either expensive hardware or specialized equipment demanding laboratory assembly and calibration. As such setups are usually composed of research and custom build devices, the resulting measured data will reflect their high development and purchase costs. This consequently limits the number of publicly available BTF samples as well as their usage in real applications.

Contribution of the paper: The main contribution of this paper is a practically verified and technically simple setup for acquisition and reconstruction of approximate BTF from a planar material sample. It allows rapid acquisition of a BTF subset in six minutes and the fully automatic reconstruction of the entire BTF dataset in under one hour using an inexpensive measurement setup based on an affordable mechanical gantry, a consumer camera and LED lighting. Our setup does not impose any additional restrictions on the type of material or its properties (e.g., isotropy, reciprocity, opacity), when compared to other BTF setups. We have psychophysically compared

All authors are with the Institute of Information Theory and Automation of the ASCR, Prague, Czech Republic, Radomír Vávra and Mikuláš Krupička are also with Faculty of Information Technology, Czech Technical University in Prague, E-mail: filipj@utia.cas.cz
our results on seven materials with their ground-true BTF measurements.

The paper is structured as follows. Section II sets the work in the context of previous work. Section III explains the principle of data acquisition and reconstruction. Section IV describes measurement and data processing procedures, while Section V shows results of the proposed acquisition setup.

II. PRIOR WORK

The proposed work relates to methods of SVBRDF and BTF acquisition and their reconstruction from sparse measurements.

As the SVBRDF is restricted by its definition to opaque and almost flat surfaces, its acquisition techniques make use of BRDF reciprocity. A severe limitation of a majority of these approaches is that they capture only isotropic SVBRDF, which is hardly the case for most spatially non-homogeneous real-world materials. On the other hand, BTF is a more general material appearance representation. Such data were initially captured by setups based on gonioreflectometers realizing the required four mechanical degrees of freedom (DOF) of camera/light/sample movement, e.g., [3]. Because the measurement times were too long, certain setups were used to reduce the required number of DOF using parabolic mirrors [4], or a kaleidoscope [5]. They allowed the capture of many viewing directions simultaneously, but at the cost of a limited range of surface height or elevation angles. The measurement time can also be reduced to approximately two hours by simultaneously using multiple lights and sensors [6], at a high financial cost of such a setup. A light stage originally designed for human face capturing was used by Gu et al. [7] for fast measurement of time-varying processes appearance. The measurement takes only 30 seconds, however, only 16 views are recorded.

Although the BTF measurement is a very demanding task, not many approximate measurement approaches have been proposed so far. An existing statistical acquisition approach [2] allows for fast and inexpensive measurement of a BTF; however, it requires a large sample of the regular material having uniform statistical properties, which is cut and positioned in different orientations with respect to the camera to achieve several viewing directions. The requirement of several sample specimens with the same statistical properties limits practical applicability of the approach to only spatially regular samples. Moreover, the need for extraction of the sample from its original environment prevents many portable measurement scenarios.

In contrast, the setup presented in this paper allows approximate BTF measurement and reconstruction of a wide range of materials, without restriction on their properties, required sample preparation, or their extraction from original environments. For a pixel-wise BTF reconstruction from our sparse measurements we applied the BRDF acquisition and reconstruction method introduced in [8]. Its extension to adaptive measurement of multiple slices and more general reconstruction has been proposed in [9]. Although [8] also provides results of BTF reconstruction from the proposed sparse representation, a practical BTF acquisition method has been missing. Therefore, this paper’s main contributions over [8] consist of the extension of setup’s measurement ability to approximate BTF instead of BRDF without introducing additional measurement constraints, which required us to solve a number of practical challenges.

III. SPARSE DATA ACQUISITION AND INTERPOLATION

Each BTF can be viewed as a collection of pixel-wise apparent BRDFs (ABRDFs), which do not follow the BRDF properties due to non-local effects in a material structure like shadowing, masking, etc. If we process individual color channels separately, each pixel’s ABRDF can be represented by a four-dimensional function \(\text{ABRDF}(\theta, \phi, \theta_v, \varphi_v)\). ABRDF is the most general data representation of material reflectance dependent on local illumination \(I(\theta, \varphi)\) and view \(V(\theta_v, \varphi_v)\) directions. Its typical parameterization by elevation \(\theta\) and azimuthal \(\varphi\) angles is shown in Fig. 1-a. ABRDF can represent dependence of view and illumination directions of a single BTF pixel. A projection of the 4D ABRDF by means of a 2D image is shown in Fig. 1-b. Note that individual rectangles (an example is shown in red) represent 2D subspaces of a 4D ABRDF at constant elevations \((\theta_v/\theta)\). These subspaces are toroidal. That is, data of the highest \(\varphi \approx 2\pi\) are followed by data of the lowest \(\varphi \approx 0\).

In this paper we use ABRDF reconstruction from sparse data proposed in [8]. This method is based on two perpendicular slices measured across azimuthal angles for fixed light and camera elevations as shown in Fig. 2-A (red and blue). As the slices are perpendicular to anisotropic (the red one) and specular highlights (the blue one) they bear enough information to approximate azimuthally-dependent behavior of the ABRDF. A principle of the ABRDF reconstruction method is sketched in Fig. 2-B. First, the sparse samples are measured at predefined locations (azimuthal angles) for defined elevations. Then four sampled ABRDF toroidal subspaces are reconstructed from the values of the slices, and finally the remaining values are interpolated. All the BTF pixels, i.e., ABRDFs are processed separately in this way.

\[\text{ABRDF}(\theta, \phi, \theta_v, \varphi_v)\]

\[\text{reference data} \quad \text{8 slices} \quad \text{reconstruction} \quad \text{interpolation}\]

Fig. 2. (A) A schema of ABRDF subspace representation using two slices and (B) principle of entire ABRDF reconstruction from four recorded subspaces [8]: (a) the reference, (b) sparse-sampling of eight slices, (c) reconstructions of elevations where the slices were measured, (d) missing data interpolation.
IV. APPROXIMATE BTF MEASUREMENT

This Section describes a fast, practical measurement setup of capturing the slices in BTF space using a consumer camera and a LED point-light source, followed by a complete BTF dataset reconstruction from such measurements.

A. Acquisition Setup

The setup realizing our measurement of the BTF slices is shown in Fig. 3. It consists of a mechanical gantry holding two arms rotating synchronously in either the same or opposite direction. The gantry was built using a Merkur toy construction set\(^1\). Contrary to its initial version presented in [8], requiring manual movement of the arms, we use a more reliable solution with a single DC motor (4.5V) run at a constant speed. Using additional gears guarantees accurate arms synchronization and allows us to switch the mutual rotation directions of the arms (see Fig. 3-top-right). One of the arms holds two LED Cree XLamp XM-L light sources with 20\(^\circ\) frosted optics (see Fig. 3-bottom-right). Contrary to the original design we improved arms weight balancing and added a rotation slip-ring to avoid clumsy LED power supply (0.7A/3V) wires. The second arm has two positions for attachment of a Panasonic Lumix DMC-FT3 camera. One advantage of this camera is that it does not have a protruding lens, which could possibly block the arm bearing the lights. Elevations of the LEDs and camera in both positions are fixed at 30\(^\circ\) and 65\(^\circ\). The setup can be constructed in 10 hours, for less than $350 and can hold almost any compact camera. The setup’s dimensions are 0.6×0.6×0.4 m, with a weight of 6 kg. The frame holding the rotating arms can be folded down on the support platform, so the setup’s size can even be reduced to allow easy transportation and the use for field measurements of samples that cannot be removed from their environments.

A material sample is placed below the setup under the arms’ rotation axis (see Fig. 3). The axial slice data are measured using rotation of the mutually fixed light and sensor around the sample, while the diagonal slice data are obtained by mutually opposite movements of the light and sensor with respect to the sample. In both cases, the camera and light travel full circle around the sample and return to their initial positions.

\(^1\)http://www.merkur.co.uk/

B. Sparse BTF Data Capturing

The camera records the material sample’s appearance at different arm positions as a video sequence at a resolution of 1280×720 pixels. Both \(s_A\) and \(s_D\) slices are recorded for two different elevations of the camera (C1, C2) and light (L1, L2); therefore, eight slices are measured at approximate elevations \(\theta_i/\theta_u = [30^\circ/30^\circ, 30^\circ/65^\circ, 65^\circ/30^\circ, 65^\circ/65^\circ]\) as shown in Fig. 2-B-b. Recording of a single slice takes 30 seconds and the entire set of eight slices is captured in six minutes – during which time the LEDs are switched four times, gears twice, and camera position once. Note that the fully automatic device assembled with two cameras would need only two minutes. Camera zoom and white balance are fixed during the recording to allow the calibrations described in Section IV-C. The camera’s image stabilizer was switched on and the global exposure level was set to \(-\frac{3}{2}\) EV to avoid overexposure. Note that, although the exposure level of individual measured images is unknown, the global exposure level adjustment by an uniform change of exposure time/aperture can be used to shift dynamic range of the whole recorded image sequence.

Due to a short distance of the camera and LED lights from the sample, we have limited the size of the measured sample to 30×30 mm, which is sufficient for a wide range of materials. However, shadows cast by the registration frame (see Fig. 4) limits the effective size to \(\approx 80\%\). Although a larger sample size can be used, we limited it due to a large span of illumination and view angles across sample’s plane, possibly causing spatial reflectance non-uniformity in the captured images. Finally, we cut the smallest possible repetitive tiles found near the sample’s image center. A white border with a white dot was attached around the sample (see Fig. 4) for detection of the camera orientation with respect to the sample, and for frame registration.

C. Data Processing and Calibrations

Basic processing steps applied to the measured data are outlined in Fig. 4 and explained in more detail in this Section.

Frame Filtering and Registration – Eight slices are recorded as videos at a frame rate of 30 fps, i.e., 900 frames per slice. As the camera provides M-JPEG non-interleaved variable bit-rate format storing all recorded frames (i.e., not only the key-frames), the error introduced by video coding is negligible compared to the ABRDF reconstruction error. From each of the eight video sequences, 24 frames are extracted corresponding to the sampling of azimuthal angles \(\varphi_i/\varphi_u\) every 15\(^\circ\). However, as not all of the frames are

---

Fig. 3. The proposed BTF measurement device.

Fig. 4. The pipeline of each slice’s frame processing.
sharp due to motion blur we search in the neighborhood of ± 4 frames for the sharpest image minimizing the edge-based blur metric. As this approach might miss intensity at specular reflections, we additionally scan the entire sequence for the two brightest frames, which are always sampled. To capture even very sharp specular highlights, four samples 1º apart from specular reflections are recorded as well. This leaves us with a set of 30 frames per slice. A subset of the recorded frames is used for camera calibration² and further for geometric distortion compensation of all frames. The frame registration itself is performed in two steps. First, all images are registered based on the sample’s border line detection using the Hough transform and computing homography between their intersections and desired corner coordinates. Second, as the measured material sample’s plane is usually at least 0.1 mm below the registration plane defined by the white border, we detect and compensate for this height and angular misalignment using the iterative fitting method [10]. This method uses PCA-based image compression as the alignment quality measure. Finally, the registered images are cropped to a size of 300×300 pixels, yielding a resolution of 340 DPI. If the sequences were recorded in a Full HD resolution (1920×1080 pixels) the BTF resolution would approach 500 DPI.

Exposure and Light Non-Uniformity Compensation – Unfortunately, most compact cameras adapt their exposure depending on the amount of light coming from the scene, which is also true for the camera we used. On a positive note, this enabled us to capture as much information as possible, even using a limited dynamic range of the camera’s sensor (8bits/channel). However, the information about exposure throughout the video sequence could not be retrieved from an EXIF header as is possible for still photos. Another related problem is the spatial non-uniformity of illumination, which is caused by a limited distance of LEDs from the sample and the fact that we use point-light instead of directional illumination. Therefore, we compensate for the exposure fluctuations as well as for the spatial non-uniformity of illumination of the original images I using the intensity of black uniform material with known BRDF at the locations beyond the white border surrounding the measured sample (Fig. 4). A compensation image C is computed for each frame by linear interpolation of the black material intensity. First, the originally measured frames in slices I are resampled to the azimuthally uniform grid IC, which step is necessary for ABRDF subspace reconstruction. Pixels of the monospectral correction image K are obtained from the center of the image divided by value interpolated from measured intensities behind the white frame, i.e., this value includes the known reference BRDF of the black material B divided by the compensation image C for the corresponding angles

\[ K_i(x, y) = \frac{1}{3} \sum_{j=1}^{3} B(\xi(i), j) C_i(x, y, j), \]  

where \( i = 1 \ldots m \) is the number of the frame in the slice of size \( m \), \( j \) is a color channel and \( \xi \) is a known mapping function \( \xi(i) \rightarrow (\theta_i, \phi_i, \theta_v, \phi_v) \).

The compensated image \( I_C \) is obtained as

\[ I_{C,i}(x, y, j) = I_{G,i}(x, y, j)K_i(x, y). \]  

Reference BRDFs of the black target B were obtained from the UTIA BTF database³.

Obtaining View and Illumination Directions – When all frames have been registered and compensated, their corresponding illumination and viewing angles must be identified. The camera viewing angles \( \theta_i/\phi_i \) are obtained from the camera’s extrinsic parameters, given the known camera calibration and corner points of the sample’s borders. Coordinates of these points are obtained from the image registration based on the camera calibration. As the viewing angles are known and the lighting support arm is mechanically coupled with the camera support arm (Fig. 3), the illumination azimuth angle can be computed as: \( \varphi_i = \varphi_{vi} + \alpha \) for the axial slice \( s_A \), and \( \varphi_i = 2\pi + \alpha - \varphi_{vi} \) for the diagonal slice \( s_D \). The elevation angles \( \theta_i \) are estimated from the fixed vertical positions of the LEDs in the setup (Fig. 3).

Colorimetric Calibration – The compensated images are further colorimetrically compensated using a transformation matrix relating measured and known color values of the ColorGauge Micro Target (35×41 mm) in the least-square sense.

Entire BTF Space Reconstruction – Since direct use of material images for BTF rendering would exhibit distractive seams on textured surfaces, we employ an image-tiling approach to find seamless tiles. Finally, the compensated and tiled images of the material’s appearance for the known illumination and viewing directions (measured in slices) are used for reconstructing the remaining directions. Due to a lower computation complexity we used step-wise linear interpolation of the entire illumination and view space instead of the more accurate but substantially slower global interpolation approach as shown in [8].

Timings – A typical timeframe for data processing using a desktop PC with Intel i7-3610QM 2.3 GHz is as follows: video sequences decoding, 1 minute; obtaining sharp frames, 15 minutes; image registration and registration plane alignment, 15 minutes; image-tiling, 5 minutes – in total 36 minutes. As the reconstruction of a single BTF pixel from the slices takes 0.2 second, the reconstruction of a BTF tile of the size 128² takes under one hour using a single core or 17 minutes using four cores. To summarize, our method allows measurement and reconstruction of BTF in under one hour.

V. RESULTS

We used seven BTF samples from the UTIA BTF database³ for evaluation of our method. These measurements cover a hemisphere of illuminations/views in 81 directions [3]. One advantage of this database is that it provides physical specimens of some of the measured samples for research purposes. Therefore, we can directly compare measured reference BTF data with their approximation as captured by our acquisition setup.

²http://www.vision.caltech.edu/bouguetj/calib_doc/
³http://btf.utia.cas.cz
We selected different types of materials for measurement as illustrated in Fig. 5: non-woven fabric (fabric03), upholstery with a height profile (fabric04), woven fabric (fabric38), and corduroy-like upholstery (fabric78), artificial leather (leather01), raw wood (wood01), and rough sandpaper (sandpaper01), most of them highly anisotropic as is clear from their spatially averaged ABRDFs shown in the second row.

As the reference measurements have a higher resolution (353 or 1071 DPI) than the data resolution captured by our camera (340 DPI), we downsampled the reference data to match our lower resolution. We also attempted to cut similar tiles from both the reference data and the captured BTF datasets to achieve a fair comparison of our measurements with the reference BTF data.

A comparison of the reference BTF data with our sparsely measured and reconstructed BTF datasets is shown in Fig. 6. For the purpose of distinguishing between differences introduced by the reconstruction process and those resulting from the proposed acquisition technique, the reference measurements (the left column) are compared with two types of results. The first one (the middle column) reconstructs BTF from the subset of reference measurements (240 images) while the second one (the right column) reconstructs BTF from the same subset of images, however, measured by the proposed acquisition setup.

These images show a close resemblance of both results to the reference data. The main difference between them is in color hue caused by differences in the acquisition processes as discussed in Section VI. Smoother appearance of the material wood01 results from variations in the amount of sanding that has been performed on the raw wood surface. Note that, although we use the same materials as the reference measurement system, we cannot achieve pixel-wise alignment between the reference and the proposed measurements. Therefore, the accuracy of the proposed method cannot be assessed by any of the standard pixel-wise quality evaluation metrics. This type of comparison is possible only between data in the first two columns. The difference values for PSNR[db], SSIM, and VDP2 are shown on the right sides of the images.

To objectively evaluate our measurement method we ran a psychophysical study with eight naive subjects. They were asked: Which of the two renderings looks more realistic? When we compared the renderings of complete and sparse reference data (the first vs. second column in Fig. 6) on average 37% of our subjects preferred renderings from sparse reference data, and when we compared renderings of entire reference data with our sparse measurements (the first vs. third columns in Fig. 6) 38% of our subjects preferred renderings from our data. These results are encouraging and show that most of the error is resulting from the reconstruction method and not from our acquisition technique, and that over one third of our subjects considered our data more realistic than the reference.

VI. DISCUSSION

A. Advantages

A notable advantage of our setup is its ability to quickly (≈ 5 minutes) measure almost any flat and slightly rough materials without needing to extract them from their environment. Compared to the SVBRDF measurement and representation approaches, the proposed method is not limited to the restrictions imposed by the BRDF properties. Therefore, it can be found useful for fast and inexpensive approximate BTF acquisition of many materials. Any height difference between the material surface and the registration plane is compensated for and the entire BTF is reconstructed in under one hour.

B. Limitations

To achieve fast and portable appearance measurements, our method restricts the number of measured images to 240 and this fact is reflected in certain limitations of its visual accuracy. Although there are not any restrictions imposed on the measured materials when compared with other BTF capture setups, the results have shown the following general limitations:

- **Lower sharpness of structure details** – results from geometrical deformation of the structure’s features, which is caused partly by mechanical vibrations during the measurement and partly by very sparse sampling of the azimuthal space, as well as by interpolation of the data at missing elevations. The interpolation causes blurring due to improper highlight extrapolation for low elevation angles. Another reason for the lower contrast is a low dynamic range of the camera sensor, where certain details are lost after the exposure compensation (e.g., white dots in fabric04); therefore, a sensor with a dynamic range over 8bits/channel would help.

- **Color hue differences** – are due to different dynamic range and spectral response of the RGB sensors used for the reference and our data acquisition, and due to different calibration targets used. Another source of these differences can be slight color variations across the specimen plane (e.g., sandpaper01).

- **Visible repeatable seams** – are caused by tiling with the aid of only a single BTF tile and by less than ideal illumination non-uniformity compensation, and are apparent for samples represented by a very large tile.

- **Limited sample size** – A larger sample size is not a severe limitation of our setup, as larger areas can be, due to a high speed of the measurement, scanned sequentially; and this approach does not compromise the setup portability. Alternatively, the setup can be built in a larger size for only minor additional costs.
Highly specular samples – could be inappropriately represented using the proposed fixed-sampling approach and would require adaptive sampling based either on initial material scan or on a step-wise adaptive approach [9].

Even though the reconstruction from sparse samples is not physically correct mainly in terms of properly shading structural elements, the performed perceptual study has confirmed that our method captures the look-and-feel of the material’s appearance.

VII. CONCLUSIONS

We present a fast and inexpensive setup for material appearance acquisition in the form of an approximate bidirectional texture function (BTF). The proposed acquisition setup is based purely on consumer hardware and is easy to build. The data acquisition and subsequent fully automatic reconstruction of the entire BTF dataset is fast and computationally non-intensive. The measurement process records a material’s appearance using eight video sequences, from which only 240 frames are taken to approximate the entire BTF of the measured sample. The promising performance of this method has been thoroughly psychophysically compared with reference BTF measurements.

Although the presented setup has certain limitations and does not capture exact detailed appearance for some materials, we believe that its speed, simplicity, and portability will make these approximate BTF measurements accessible even to such applications for which the standard BTF acquisition methods are too expensive.

In our future work we plan to employ material-dependent sampling along the measured slices and render material appearance on GPU directly from the sparsely measured dataset without the need for reconstructing a complete BTF dataset.

Acknowledgments We thank all volunteers for participation in the psychophysical experiment. This research has been supported by the Czech Science Foundation grants 103/11/0335, 14-02652S and EC Marie Curie ERG 239294.

REFERENCES

Fig. 6. A comparison of renderings: (left) using the entire reference BTF datasets (6561 images), (middle) using BTF reconstruction from a sparse subset of reference BTF (240 images), (right) using BTF reconstruction of the proposed measurements (240 images). At the end of each row are PSNR[dB]/SSIM/VDP2 values between the first two images.