

Supplementary Material

Paper ID 1526

The organization of the material is as follows. In Section 1, some additional results of occluder removal are shown. Some example reconstructions for Internet photo collections are shown in Section 2. Experiments on BRDF databases are discussed in Section 3. The approach we took for processing color images is described in Section 4.

1 Occluder Removal Results

Figures 1 and 2 show some additional results.

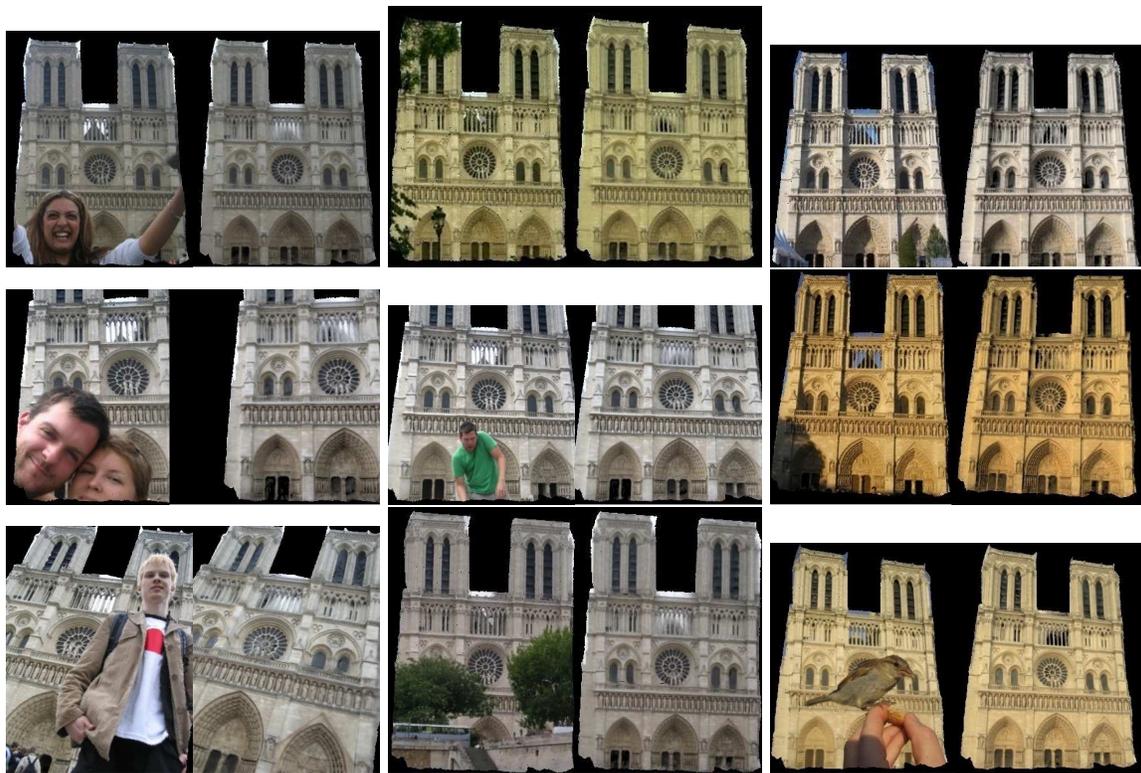


Figure 1: Occluder Removal Results

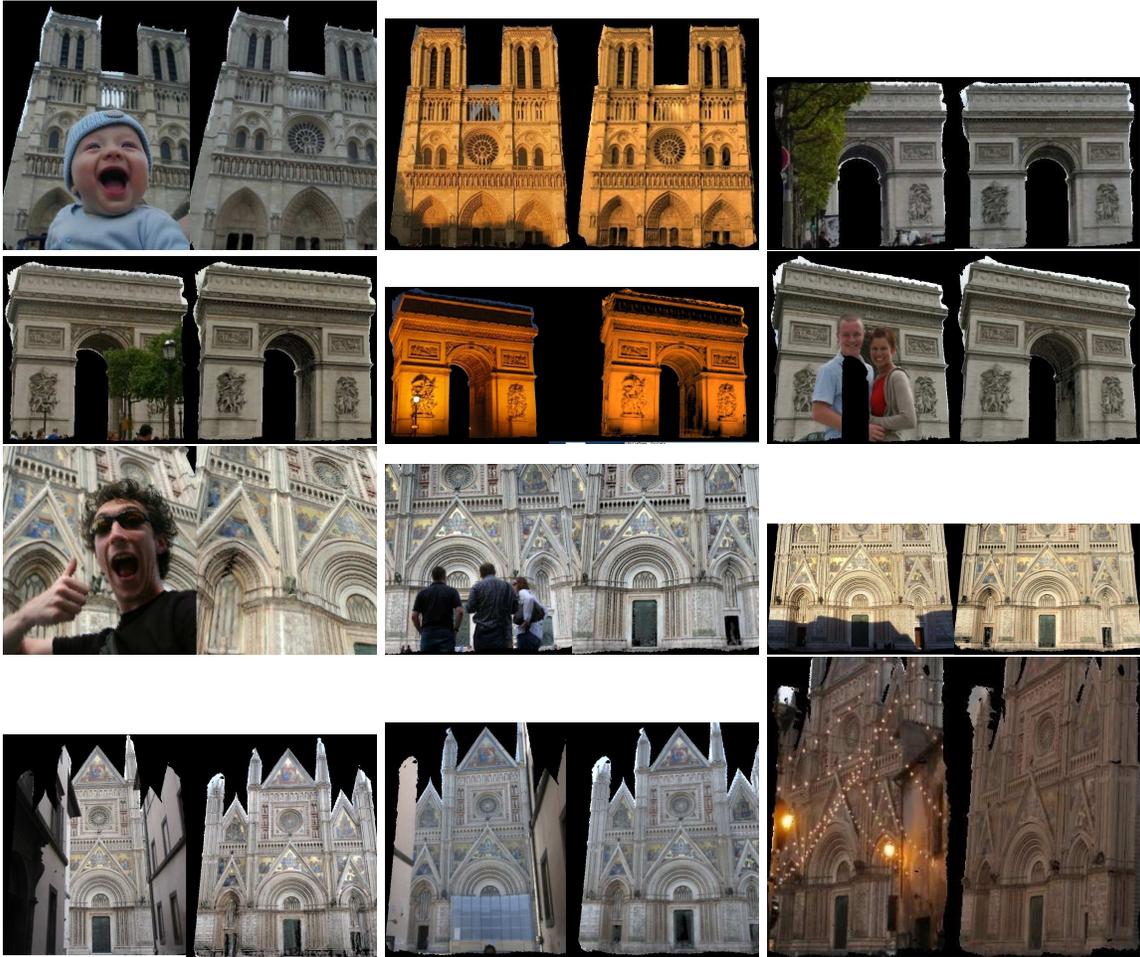


Figure 2: Occluder Removal Results

2 Reconstruction of Internet Photo Collections

Some example reconstructions with varying number of basis images are shown in Figure 3.

Figure 4 shows the fall of relative RMS error vs the number of basis images (for the green channel of images). The relative RMS error is measured as the RMS error divided by the RMS value of all images in a dataset.

As was mentioned in the paper, there exists sufficient parallax in the images of Moon for Structure from motion to be able to register the photos. Figure 5 shows a video of the reconstruction.

3 Experiments on CURET BRDF Database

We performed experiments on BRDF databases to ascertain the range of materials present in real world images. We looked at two different databases of BRDFs – MERL BRDF database [2] which has 100 different materials, and CURET database [1], which has 61 different materials. CURET is more representative of real world materials (e.g., paper, grass, cloth, etc.) whereas MERL is

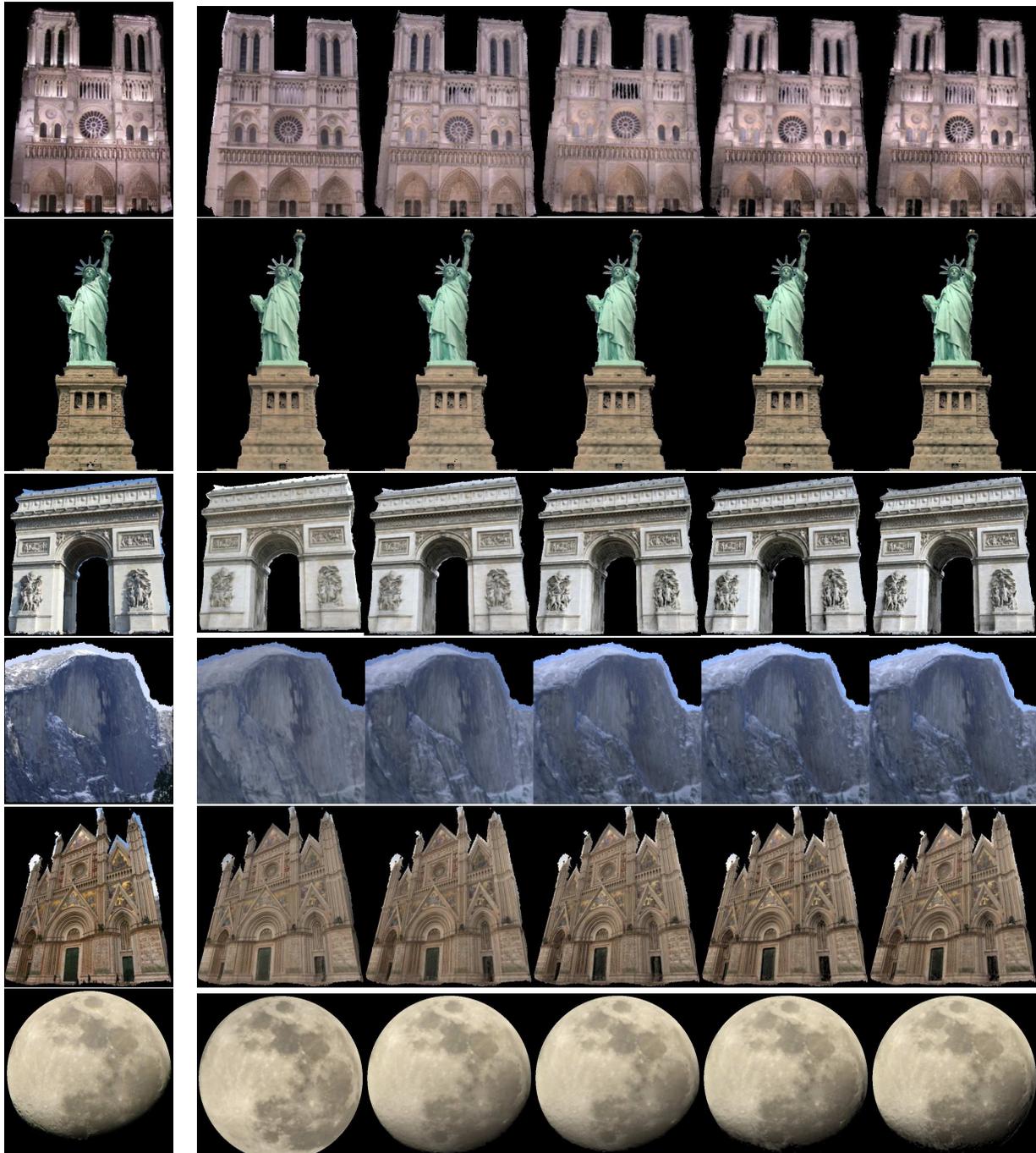


Figure 3: Left: An image from the dataset. Right: Corresponding reconstructions using 1, 3, 5, 10 and 20 basis images

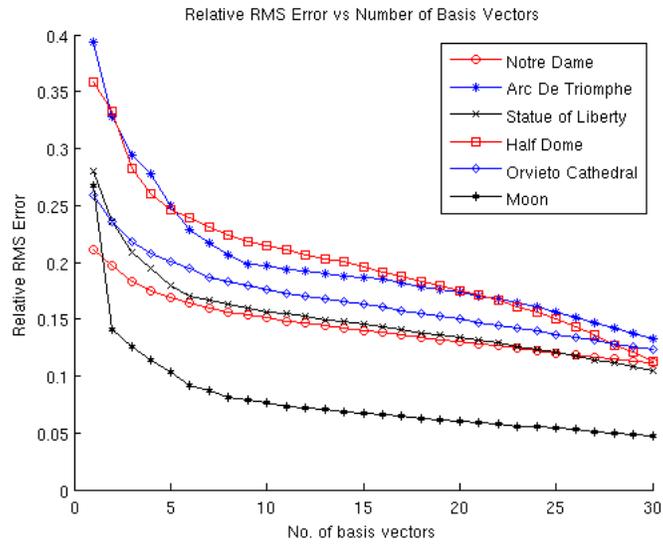


Figure 4: Relative RMS Error vs number of basis images. Even though the reconstruction error is around 12% (due to noise, occluders, etc.) for most datasets even with 30 basis images, reconstructions with only 10 basis images look visually similar.

(Loading Movie..)

Figure 5: A video showing point cloud corresponding to the reconstruction from the moon images. Click to play. The video file is also provided separately in the directory if you have trouble playing.

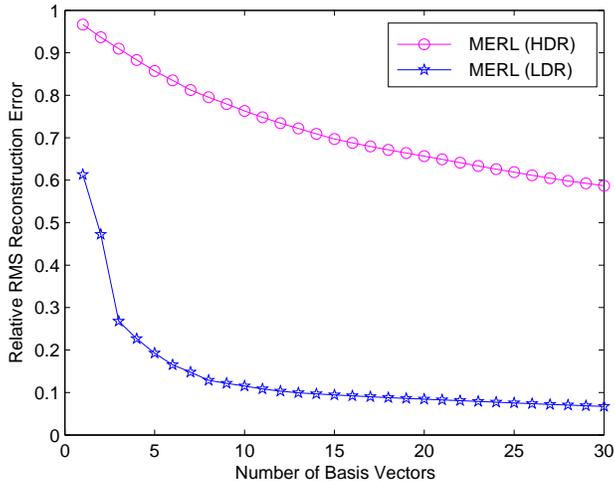


Figure 6: Reconstruction Error vs Number of basis images for HDR and LDR images rendered using materials in the MERL database. A hundred 50×50 images of a sphere of each material were rendered under different illumination conditions. Basis images were computed using this collection of all 100×100 images. The relative RMS error was measured as the ratio of the RMS error to the RMS of rendered images. The LDR image curve is seen to be much lower, indicating that much of the error in the HDR curve is due to sharp specularities.

restricted to machined and painted spheres. Since our goal is to understand the dimensionality of real-world scenes, we chose to focus on CURET.

Second, the prevalence of specularities in MERL in combination with high dynamic range capture, makes approximating through linear models much more difficult, as the highlights alone capture vast majority of image energy. We found, however, that converting the images to a standard 8-bit-per-channel dynamic range (and clamping highlights to 255) yields a reasonable fit (See Figure 6). Here, to convert HDR images to LDR images, we quantized the range of each image so that 90% of the pixels fall in the range $[0, 255]$ and clamped the pixels outside this range to 255.

CURET database is used for the results in the following sections.

3.1 Appearance Space of each Material

We first analyze how many basis images are required to span the appearance space of each material. For every material, a hundred 50×50 images of a sphere were rendered. Each image was lit by a distant directional white light whose direction was randomly chosen (uniformly distributed over the front facing hemisphere). The rendered images were reduced to grayscale and SVD was used to compute the basis images corresponding to each material. The green (middle) curve in Figure 7 shows the fall of *relative RMS error* as the number of basis images used to model the appearance are increased. The curve is averaged over all 61 materials for CURET. From the graph, we can see that the error falls to less than 10% after only 6 basis images.

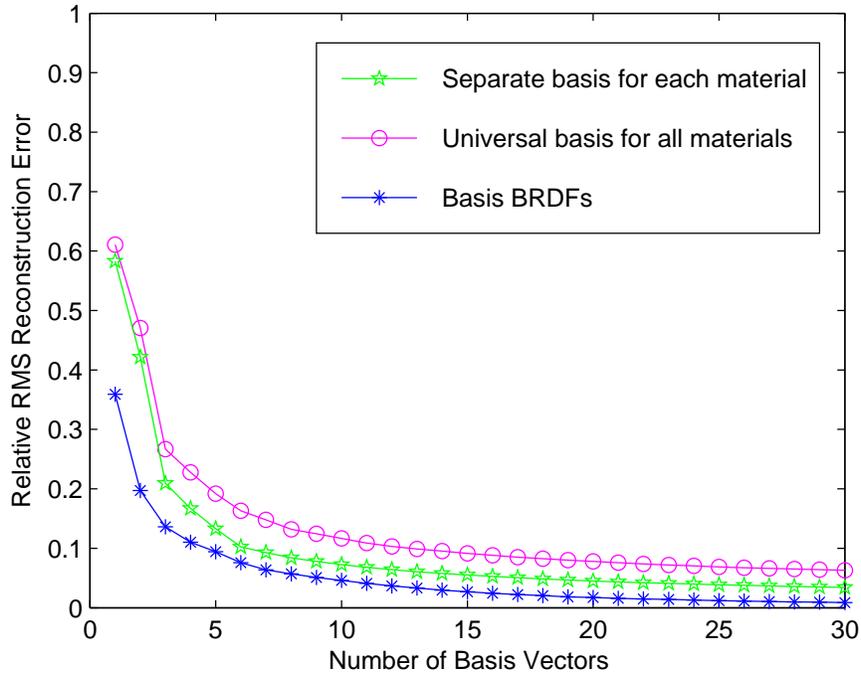


Figure 7: The fall in relative RMS error vs number of basis vectors. Three curves are shown. First, where a separate basis was computed for each material (The curve shown is averaged over all 61 materials) . Note that reconstruction error is less than 10% with only 6 basis images. Second, the case where a *universal basis* was computed for all images of all materials. Note that this curve is only slightly above the first curve indicating that the space of all materials can be spanned by a small basis. The last curve corresponds to the case where every BRDF was treated as a vector and basis BRDFs were computed. It is lower than the curve corresponding to universal basis as all the specularities are at the same location for all BRDF vectors while the location of the specularities in the rendered images depend on the scene geometry and position of the light sources.

3.2 Appearance Space of all Materials

It was shown in Section 2.2.1 of the paper that the upper bound on the dimensionality is reduced in case the BRDFs in a scene are contained in a low dimensional linear subspace. To gauge the range of materials present in the CURET database, we computed *basis BRDFs* (treating each BRDF as a vector). The BRDFs corresponding to the three color channels were concatenated to form a single large vector. The blue (bottom) curve in Figure 7 shows the reconstruction error vs number of basis BRDFs.

We also run SVD over *all sphere images* rendered in the previous section to calculate *universal basis images* and measure the reconstruction error versus the number of basis images (The pink (top) curve in Figure 7). The curve is marginally above the curve obtained by calculating a separate basis for each material indicating that the same basis can be shared across a large number of materials.

Reconstruction accuracy of each material is shown in figure 8. The first six *universal basis* images are shown in Figure 9, five of which are very similar to the basis images used by Ramamoorthi [3] to span the appearance of a Lambertian sphere.

Some example reconstructions are shown in Figure 10.

3.3 BRDF across Color Channels

We also test the assumption made in Section 2.3.2 of the paper i.e BRDF of a material does not depend on wavelength. For each material, we consider the $3 \times N$ matrix obtained by stacking the BRDFs of the 3 channels, where N is the number of samples in each BRDF for each channel. The mean ratio of first singular value to the second singular value was found to be 49.61 with a minimum of 4.55, indicating that for almost every material, the matrix is close to rank 1 and hence the BRDFs can be approximated as being the same across color channels save for a scaling factor.

4 Processing Color Images for Internet Photo Collections

A naive method to process color images is to treat every channel independently. However, this does not perform well for photo collections as even the handpicked set of clean images contains outliers (e.g. cast shadows) that can produce artifacts (Figure 11).

First, assume that the assumptions made in Section 2.3.2 and 2.3.3 of the paper hold and hence the profile of a pixel x corresponding to, say the red channel, $\mathbf{P}_x^{\text{red}}$ can be written as $\mathbf{P}_x^{\text{red}} = f^{\text{red}}(x)\mathbf{P}_x^{\text{green}}$ where $f^{\text{red}}(x)$ depends on $k_{\text{red}}(x)$ and $k_{\text{green}}(x)$ defined by Equation 18 in the paper. However, Internet photos are not captured by identical cameras and the spectrum of light is also different for different photos (Spectrum in the evening is very different from the spectrum during the day). Intuitively, the combined effect of these two factors can be thought of as individual scaling applied to the channels of a particular image, i.e., $\mathbf{P}_x^{\text{red}}(i) = g^{\text{red}}(i)f^{\text{red}}(x)\mathbf{P}_x^{\text{green}}(i)$ where $g^{\text{red}}(i)$ is the scaling applied to the red channel of the i^{th} image. $f^{\text{red}}(x)$ can be thought of as a measure of the red color of the pixel (The measurements are relative to the green channel of the image).

Now, if we can recover g^{red} for all images and f^{red} for all pixels, then by calculating the basis for just the green channel, we can also recover the red channel. We found that the following simple method works well in practice. We assume that there exists a dominant value in $g^{\text{red}}(i)$, say $g_{\text{dom}}^{\text{red}}$. Note that $g^{\text{red}}(i)$ depends on the interaction of the camera sensor and the illumination. So, this

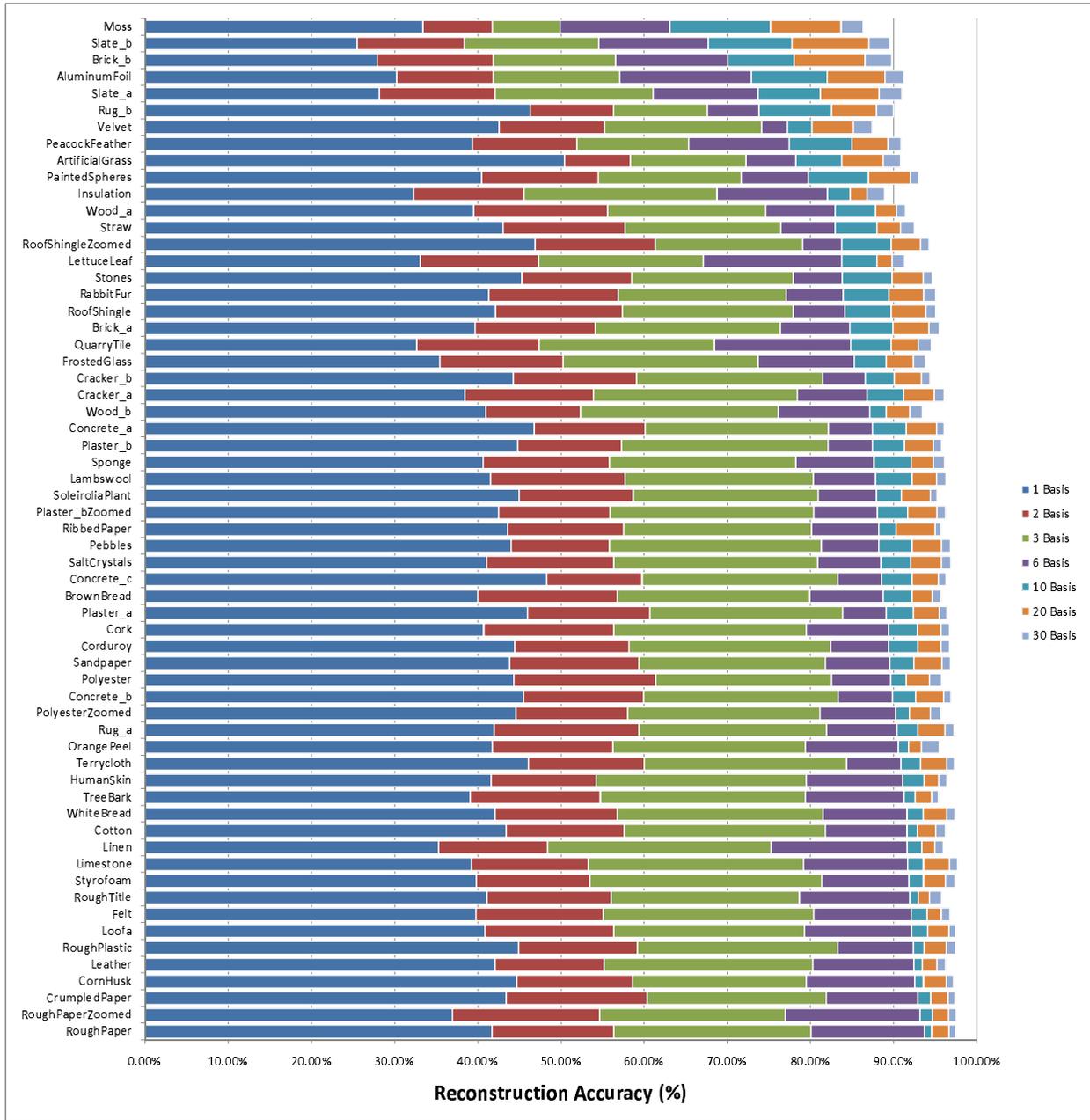


Figure 8: Reconstruction accuracy achieved for each material in the CURET database using 1, 2, 3, 6, 10, 20 and 30 basis images respectively. The basis was computed using the set of all images of all materials (universal basis). Best seen in color.

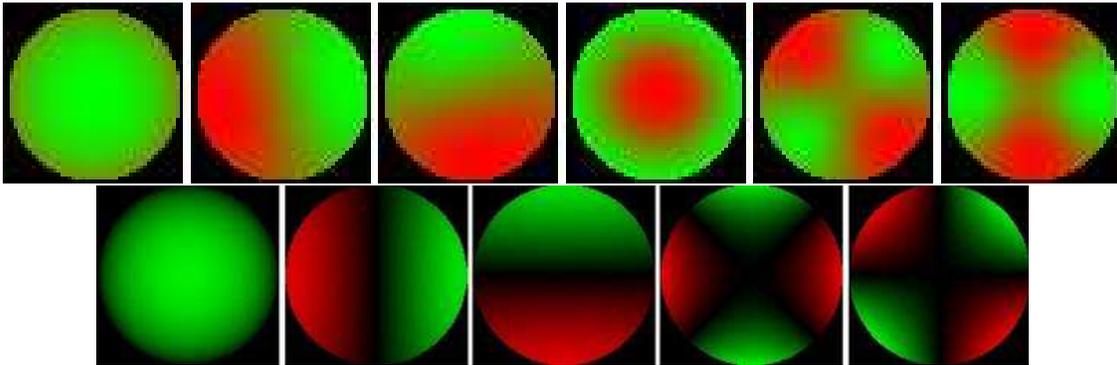


Figure 9: The top row shows the first six *universal basis* images computed by performing a SVD over all sphere images rendered using materials in the CURET database. Five of the six are remarkably similar to the 5 basis images used by Ramamoorthi [3] to span the appearance space of a Lambertian sphere (bottom row), i.e., the Lambertian basis augmented with one additional basis is sufficient to give a reconstruction accuracy of about 85% on average for the materials in the CURET database

assumption translates to saying that a large number of photos are taken by similar cameras under lighting conditions with similar spectra (This is plausible as the spectrum of the light remains the same for the larger part of the day). Hence, we can recover $f^{red}(x)$ by

$$f^{red}(x)g_{dom}^{red} = \text{median}_i \left(\frac{\mathbf{P}_x^{red}(i)}{\mathbf{P}_x^{green}(i)} \right) \quad (1)$$

where the median is being taken over i , i.e, along the profile. Once, we have recovered $f^{red}(x)$, one can also recover $g^{red}(i)$ by

$$\frac{g^{red}(i)}{g_{dom}^{red}} = \text{median}_x \left(\frac{\mathbf{I}_i^{red}(x)}{f^{red}(x)g_{dom}^{red} \mathbf{I}_i^{green}(x)} \right) \quad (2)$$

where now the median is taken over x i.e. an image. The blue channel can also be processed similarly.

Figure 4 shows the relative RMS error in three color channels for the Notre Dame and Arc De Triomphe datasets. The curves corresponding to the red and blue channels are seen to be only marginally above the green channel curve.

Let us now formally see under what physical conditions is the above intuition correct. We still make the assumption that in a *particular* image, the spectrum of light is same for all light sources which allows us to write $L_{f_i}(\omega', \lambda)$ as $K^i(\lambda)L'_{f_i}(\omega')$. Even though illumination in outdoor scenes is often modeled using two distinct light sources – sunlight and skylight which have different spectra, one can assume that one of them dominates, i.e. the intensity of one is much stronger than the other. For e.g., sunlight will dominate on a clear sunny day while skylight will dominate on an overcast day. Repeating the analysis of Section 2.3.2 (where we now have $K^i(\lambda)$ instead of $K(\lambda)$), it can be seen that $\mathbf{P}_x^{c_1}(\mathbf{i})/k_{c_1}(x, i) = \mathbf{P}_x^{c_2}/k_{c_2}(x, i)$. where

$$k_{c_l}(x, i) = \int s_{c_l}(\lambda)K^i(\lambda)\alpha^x(\lambda)d\lambda \quad (3)$$

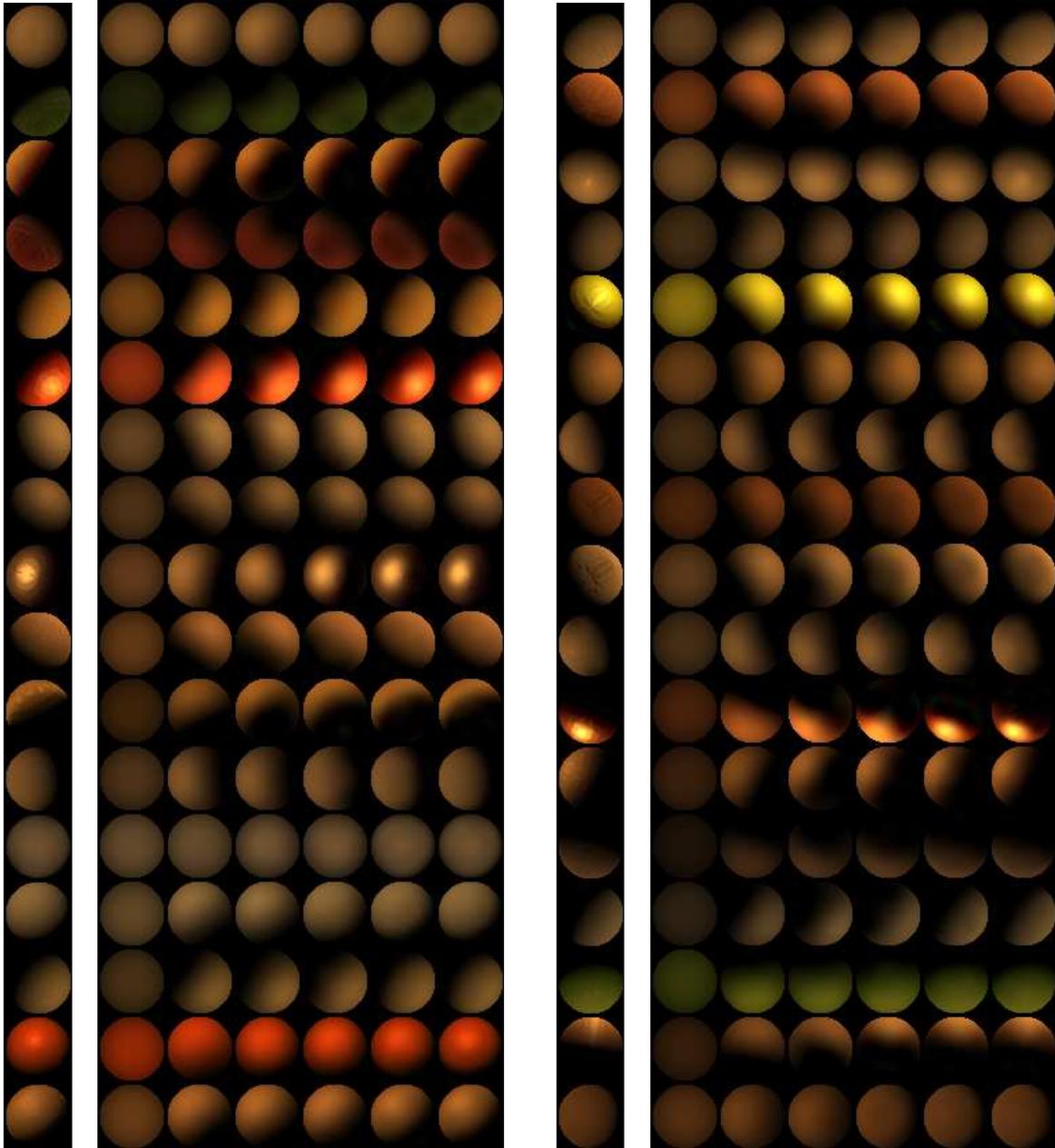


Figure 10: Some samples from the CURET BRDF database and corresponding reconstructions using 1, 3, 5, 10, 20 and 30 basis images. A single basis was learnt from images of all the materials. The reconstructions look visually similar (except for smoothing of highlight in some cases) indicating that the same basis may be shared across multiple materials.

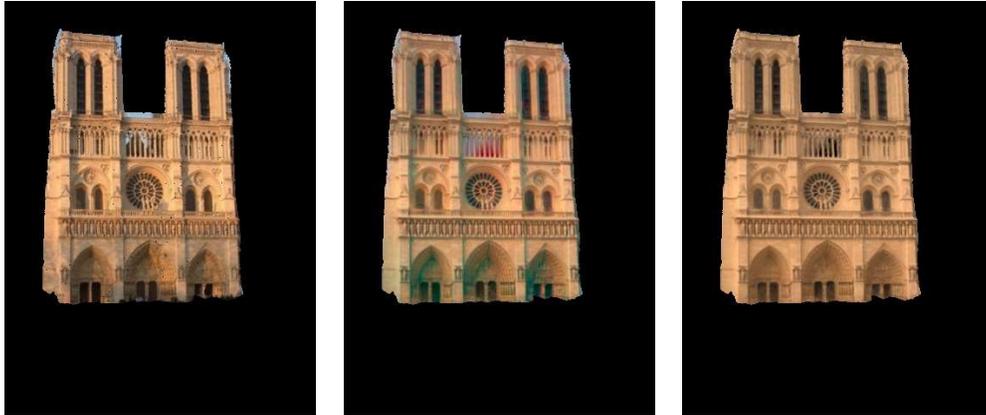


Figure 11: The image on the left is the original image. The middle image shows the reconstruction obtained using 5 basis images when processing the color channels independently. Notice the *false color* in regions of shadow or incorrect geometry. The figure on the right is the reconstruction obtained by the described approach.

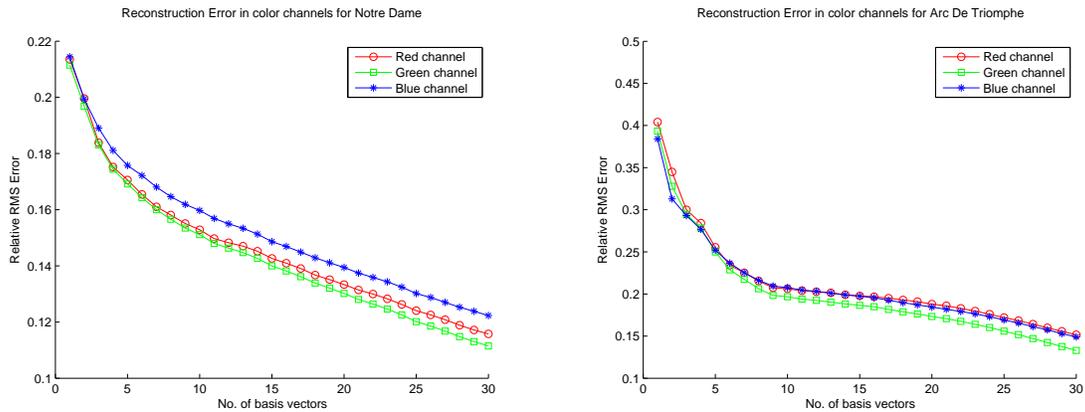


Figure 12: Relative RMS error in red, blue and green channels for the Notre Dame and Arc De Triomphe datasets. The blue and red curves are only marginally above the green curve indicating that the accuracy achieved in the reconstruction of the red and blue channel by the described approach is similar to the accuracy in the green channel (via SVD).



Figure 13: The first figure shows the original image. The second shows the result of robust projection. It does not work well as the number of outlier pixels is large. The third figure shows the precomputed outliers (in red). The last figure shows the result of robust projection with precomputed outliers.

i.e., k_{c_i} depends on i as well unlike Equation 18. We want to be able to write the above integral in the form $f^{c_i}(x)g^{c_i}(i)$. One assumption under which this holds true is when the albedos remain constant over the range of λ over which the support of spectral responses varies and hence can be taken out of the integral. This essentially translates to saying that the support of the spectral response of, say the red sensor, stays in a small neighborhood of a particular wavelength across different cameras where the albedo can be expected to remain constant.

4.1 Precomputing Outliers

The above approach can also be used to precompute outliers when projecting new images onto a basis. Assuming we have $f^{red}(x)$ computed for the scene (from the set of images used to compute the basis), given a new image, we can compute $g^{red}(i)$ using Equation 2. With this information, one can mark pixels where $|\mathbf{I}_i^{red}(x) - f^{red}(x)g^{red}(i)\mathbf{I}_i^{green}(x)|$ (and a similar expression for the blue channel) is beyond a certain threshold as outliers. Figure 13 shows an example.

References

- [1] K. Dana, B. Van-Ginneken, S. Nayar, and J. Koenderink. Reflectance and Texture of Real World Surfaces. *ACM Trans. on Graphics*, 18(1):1–34, Jan 1999.
- [2] W. Matusik, H. Pfister, M. Br, and L. Mcmillan. A data-driven reflectance model. *ACM Trans. on Graphics*, 22:759–769, 2003.
- [3] R. Ramamoorthi. Analytic PCA construction for theoretical analysis of lighting variability in images of a lambertian object. *PAMI*, 24(10):1322–1333, October 2002.