

RESEARCH

Open Access

Efficient robust image interpolation and surface properties using polynomial texture mapping

Mingjing Zhang and Mark S Drew*

Abstract

Polynomial texture mapping (PTM) uses simple polynomial regression to interpolate and re-light image sets taken from a fixed camera but under different illumination directions. PTM is an extension of the classical photometric stereo (PST), replacing the simple Lambertian model employed by the latter with a polynomial one. The advantage and hence wide use of PTM is that it provides some effectiveness in interpolating appearance including more complex phenomena such as interreflections, specularities and shadowing. In addition, PTM provides estimates of surface properties, i.e., chromaticity, albedo and surface normals. The most accurate model to date utilizes multivariate Least Median of Squares (LMS) robust regression to generate a basic matte model, followed by radial basis function (RBF) interpolation to give accurate interpolants of appearance. However, robust multivariate modelling is slow. Here we show that the robust regression can find acceptably accurate inlier sets using a much less burdensome 1D LMS robust regression (or 'mode-finder'). We also show that one can produce good quality appearance interpolants, plus accurate surface properties using PTM before the additional RBF stage, provided one increases the dimensionality beyond 6D and still uses robust regression. Moreover, we model luminance and chromaticity separately, with dimensions 16 and 9 respectively. It is this separation of colour channels that allows us to maintain a relatively low dimensionality for the modelling. Another observation we show here is that in contrast to current thinking, using the original idea of polynomial terms in the lighting direction outperforms the use of hemispherical harmonics (HSH) for matte appearance modelling. For the RBF stage, we use Tikhonov regularization, which makes a substantial difference in performance. The radial functions used here are Gaussians; however, to date the Gaussian dispersion width and the value of the Tikhonov parameter have been fixed. Here we show that one can extend a theorem from graphics that generates a very fast error measure for an otherwise difficult leave-one-out error analysis. Using our extension of the theorem, we can optimize on both the Gaussian width and the Tikhonov parameter.

Keywords: Polynomial texture mapping; Photometric stereo; Radial basis functions; Hemispherical harmonics; Robust regression

1 Introduction

Polynomial texture mapping (PTM) [1] uses a single fixed digital camera at constant exposure, with a set of n images captured using lighting from different directions. A typical rig would consist of a hemisphere of xenon flash lamps imaging an object, where directions to each light is known (Figure 1a). The basic idea in PTM is to improve on a simple Lambertian model for matte

content, whereby the three components of the light direction are mapped to luminance, by extending the model to include a low-order polynomial of lighting-direction components. The strength of PTM, in comparison to a simple Lambertian photometric stereo (PST) [2] is that PTM can better model real radiance and to some extent grasp intricate dependencies due to self-shadowing and interreflections. Usually, some 40 to 80 images are captured. The better capture of details is the driving force behind the interest in this technique evinced by many museum professionals, with the original least squares

*Correspondence: mark@cs.sfu.ca
School of Computing Science, Simon Fraser University, Vancouver,
British Columbia V5A 1S6, Canada

(LS)-based PTM method already in use at major museums in the USA, including the Smithsonian, the Museum of Modern Art and the Fine Arts Museums of San Francisco, and is planned for the Metropolitan and the Louvre (M. Mudge, personal communication, Cultural Heritage Imaging). As well, some work has involved applying PTM *in situ* for such applications as imaging palaeolithic rock art [3]. In such situations, one has to *recover* lighting directions from the specular patch on a reflective sphere [4]; such a ‘highlight’ method [5] can also be applied to museum capture of small objects or to microscopic image capture.

PTM generates a matte model for the surface, where luminance (or RGB) is modelled at each pixel via a polynomial regression from light-direction components to luminance. Say, e.g. there are $n = 50$ images, with n known normalized light-direction three-vectors \mathbf{a} . Then in the original embodiment, a six-term polynomial model is fitted at each pixel separately, regressing onto that pixel’s n luminance values using LS regression. The main objectives of PTM are the ability to *re-light* pixels using the regression parameters obtained, as well as the recovery of surface properties: surface normal, colour and albedo. For re-lighting, the idea is simply that if the regression from the n in-sample light-directions \mathbf{a} to n luminance values, L is known then substituting a new \mathbf{a} will generate a new L , thus yielding a simple interpolation scheme for new, out-of-sample, light directions \mathbf{a} .

In [6], we extend PTM in three ways: First, the six-term polynomial is changed so as to allow purely linear terms to model purely linear luminance exactly. Secondly, the LS regression for the underlying matte model is replaced by a robust regression, the least median of squares (LMS) method [7]. This means that only a majority of the n pixel values obtained at each pixel need be actually matte, with specularities and shadows automatically identified as outliers. With correctly identified matte pixels in hand, surface normals, albedos and pixel chromaticity are more accurately recovered. Thirdly, authors in [6] further add an additional interpolation level by modelling the part of in-sample pixel values that is not completely explained by the matte PTM model via a radial basis function (RBF) interpolation. The RBF model does a much better job of modelling features such as specularities that depend strongly on the lighting direction. As well, the RBF approach can make use of any shadow information to help model interpolated shadows, which change abruptly with lighting direction. The interpolation is still local to each pixel and, thus, does not attempt to bridge non-specular locales as in reflectance sharing, for example [8,9]. In reflectance sharing, a *known* surface geometry is assumed, as opposed to the present paper. Here, we rely on the idea that there is at least a small

contribution to specularity at any pixel, e.g. the sheen on skin or paintwork, so that we need not share across neighbouring pixels and can employ the RBF approach from [6]. For cast shadows, a more difficult feature to model, at each pixel the RBF model will utilize whatever shadow content is actually present across the whole set of n images from n lights.

The current study is aimed at further refining and improving the PTM + RBF pipeline as was employed in [6], as well as exploring different combinations of basis functions. The main contributions of this work are threefold:

1. We introduce a more efficient, ‘mode-finder’ regression method to replace the computationally intensive multivariate LMS regression in the matte modelling stage. Compared to the 6D LMS regression, the mode-finder effectively reduces the number of unknowns from 6 to 1 and thus greatly reduces the processing time from $O(n^6 \log n)$ to $O(n \log n)$ ([7], p. 206). We found that this simplification introduces little reduction of accuracy. Although technically the mode-finder regression approach can be applied to the mode of either luminance or any colour components, we show that the mode of luminance provides the highest accuracy. How a robust mode-finder works is simple: from the n luminance values at the current pixel, select one randomly; continue and adopt as the best estimate of the ‘mode’ that luminance which delivers the least median of squared residuals. What makes the LMS method powerful is that it provides strong mathematical guarantees on the performance given by choosing a much smaller subset than a simple exhaustive search and it also delivers an inlier band, automatically, thus classifying luminance values as usable or not. The multivariate version of LMS is similar: for 6D LMS, e.g., we randomly select six luminances and find residuals for a polynomial regression. Again, the number of selections is tremendously smaller than an exhaustive search, but nonetheless is very slow compared to a 1D search.
2. We explore different combinations of basis functions for PTM. Firstly, we extend the classical polynomial models from 6D to 16D. Moreover, due to another observation we made that luminance reconstruction has a far greater impact on the re-lighted image quality than the reconstruction of chromaticity, we can reduce the dimension for chromaticity modelling with little loss of accuracy. Reducing the number of float regression coefficients makes a difference, when we multiply by millions of pixels. We found that 16D for luminance + 9D for chromaticity is a good balance between dimensionality and accuracy.

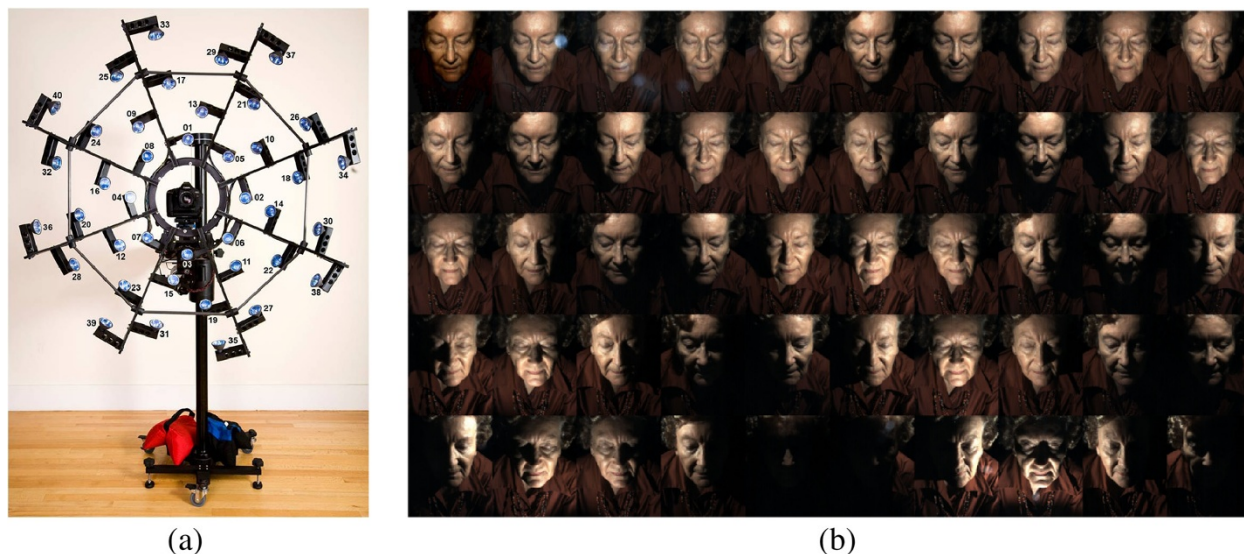


Figure 1 A typical PTM rig and an example dataset. **(a)** A 40-light rig for capturing PTM datasets (courteously supplied by Cultural Heritage Imaging). **(b)** A PTM dataset of 50 images (courteously supplied by Tom Malzbender, Hewlett-Packard).

Secondly, we compared the performance of the hemispherical harmonics (HSH) basis against the polynomial basis of the same order and found that, surprisingly, the polynomial model outperforms HSH in terms of the quality of appearance reconstruction, especially at large incident angles.

3. We adopt a method to mathematically determine the optimal parameters used for the RBF interpolation stage. Previously in [6], we made use of an RBF network consisting of Gaussian radial functions to model the non-Lambertian contribution. The parameters in this model, including the Gaussian dispersion σ and Tikhonov regularization coefficient τ , were taken heuristically and remained constant across all pixels. In this work, we start off from a theorem that minimizes error in a leave-one-out analysis by optimizing the Gaussian dispersion parameter. Such a theorem is not new, but here we extend its use to whole images and three-channel colour. More importantly, however, we also extend the theorem to optimize over the Tikhonov regularization. For the fairly large size matrices being inverted, these optimizations matter and make a substantial difference to results obtained.

Note that contributions 1 and 3 are direct improvements over the methodology of the PTM + RBF pipeline: Contribution 1 is aimed at increasing the efficiency of the first stage - matte modelling; contribution 3 is devoted to the optimization of the second stage - RBF interpolation. On the other hand, the goal of contribution 2 is to find

an optimal set of basis functions. The discoveries made in contribution 2 can be applied to the matte modelling stage of PTM + RBF, as well as regular PTM with no RBF interpolation.

This paper is organized as follows: In Section 2 we review previous work in this area, and in Section 3 we provide a brief recapitulation of the PTM method. In Section 4 we introduce the notion of our contribution 1 - using a robust mode-finder instead of a full multivariate robust regression and explicate how we use the mode-finder and trimmed LS to realize outlier detection and recover surface properties. In Section 5, focusing on contribution 2, we test the appearance reconstruction with PTM separately applied to luminance and chromaticity and compare the reconstructed matte appearance for PTM and for HSH. In Section 6 we describe our contribution 3, i.e. how to use an optimized version of the RBF framework to interpolate specularities and shadows on reconstructed images. Finally, Section 7 presents concluding remarks.

2 Related work

Many methods for detecting outlier pixels in photometric methods have been proposed. Early examples include a four-light PST approach in which the values yielding significantly differing albedos are excluded [10-12]. In a similar five-light PST method [13], the highest and the lowest values, presumably corresponding to highlights and shadows, are simply discarded. Another four-light method [14] explicitly includes ambient illumination and surface integrability and adopts an iterative strategy using

current surface estimates to accept or reject each additional light based on a threshold indicating a shadowed value. The problem with these methods is that they rely on throwing away only a small number of outlier pixel values, whereas our robust methods in the current and previous studies allow up to 50% of the pixel values discarded as outliers.

More recently, Willems et al. [15] used an iterative method to estimate normals. Initially, the pixel values within a certain range (10 to 240 out of 255) were used to estimate an initial normal map. In each of the following iterations, error residuals in normals for all lighting directions are computed and the normals are updated based only on those directions with small residuals. Sun et al. [16] showed that at least six light sources are needed to guarantee that every location on the surface is illuminated by at least three lights. They proposed a decision algorithm to discard only doubtful pixels, rather than throwing away all pixel values that lie outside a certain range. However, the validity of their method is based on the assumption that out of the six values for each pixel, there is at most one highlight pixels and two shadowed pixels. Julia et al. [17] utilized a factorization technique to decompose the luminance matrix into surface and light source matrices. The shadow and highlight pixels are considered as missing data, with the objective of reducing their influence on the result. Wu et al. [18] formulated the problem of surface normal recovery as a rank minimization problem, which can be solved via convex optimization. Their method is able to handle specularities and shadows as well as other non-Lambertian deviations. Compared to these methods, the algorithm proposed here is a good deal simpler, while producing excellent results.

A small number of recent studies utilize probability models as a mechanism to try to incorporate handling shadows and highlights into the PST formulation. Tang et al. [19] model normal orientations and discontinuities with two coupled Markov random fields (MRFs). They proposed a tensorial belief propagation method to solve the *maximum a posteriori* problem in the Markov network. Chandraker et al. [20] formulate PST as a shadow labelling problem where the labels of each pixel's neighbours are taken into consideration, enforcing the smoothness of the shadowed region, and approximate the solution via a fast iterative graph-cut method. Another study [21] employs a maximum-likelihood (ML) imaging model for PST. In their method, an inlier map modelled via MRF is included in the ML model. However, the initial values of the inlier map would directly influence the final result, whereas our methods do not depend on the choice of any prior.

Yang et al. [22] include a dichromatic reflection model into PST and associated method for both estimating

surface normals as well as separating the diffuse and specular components, based on a surface chromaticity invariant. Their method is able to reduce the specular effect even when the specular-free observability assumption (that is, each pixel is diffuse in at least one input image) is violated. However, this method does not address shadows and fails on surfaces that mix their own colours into the reflected highlights, such as metallic materials. Moreover, their method also requires knowledge of the lighting chromaticity - they suggest a simple white-patch estimator - whereas in our method, we have no such requirement. Kherada et al. [23] proposed a component-based mapping method. They decompose the captured images into direct and global components - single bounce of light from a surface, as opposed to illumination onto a point that is interreflected from all other points of the scene. They then model matte, shadow and specularly separately within each component. Their method is stated to provide a better appearance reconstruction than the original PTM [1], although at the cost of a much heavier computational load, but depends on a training phase and requires accurate disambiguation of direct and global contributions.

Aside from the polynomial basis, it is possible to use other types of basis function in PTM, as long as they provide a good approximation of the light-reflectance interaction. Spherical harmonics (SH), the angular portion of a set of solutions to the Laplace's equations defined on a sphere, appear to be a good candidate for this purpose. Due to their appealing mathematical properties, they have been extensively applied in a great variety of topics in computer graphics, such as the modelling of BRDFs [24], early work on image-based rendering and re-lighting [25,26], BRDF shading [27], irradiance environment maps [28], precomputed radiance transfer [29,30], distant lighting [31,32] and lighting-invariant object recognition [33]. However, in the context of PTM, we note that the incoming and outgoing lights are defined only on the upper hemisphere. Therefore, representation of such a hemispherical function using basis functions defined over the full spherical domain introduces discontinuities at the boundary of the hemisphere and requires a large number of coefficients [34]. Thus, it is more natural to map these functions to a basis set defined only over the upper hemisphere. In [34], a HSH basis derived from SH using shifted associated Legendre polynomials was proposed. This basis has been applied in surface modelling under distant illumination [35] and in shape description and reconstruction of surfaces [36]. Recent progress on HSH includes a HSH-based Helmholtz bidirectional reflectance basis [37] and noise-resistant Eigen hemispherical harmonics. In this study, we incorporate the HSH basis as proposed in [34] into the framework of PTM and compare its performance with the polynomial basis.

PTM and other similar reflectance transformation imaging (RTI) methods have found extensive applications in cultural heritage imaging and art conservation. Earl et al. use PTM to capture and visually examine a great variety of ancient artefacts, including bronze busts, coins, paintings, ceramics and cuneiform inscriptions [38-41]. Duffy [42] employed a highlighted RTI method to record the prehistoric rock inscriptions and carvings at the Roughting Linn rock site, UK. Padfield et al. [43] adopted PTM to digitally capture paintings in order to monitor their physical changes during conservation. These applications demonstrate the ability of PTM to visually enhance the captured images via different display modes, most notably specular enhancement and diffuse gain, allowing for inspection of features such as fingerprints and erasure marks that are otherwise much less visually prominent in regular images.

3 Matte modelling using PTM

3.1 Luminance

PTM models smooth dependence of images on lighting direction via polynomial regression. Here we briefly recapitulate PTM as amended by [6]: Suppose n images of a scene are taken with a fixed-position camera and lighting from $i = 1..n$ different lighting directions $\mathbf{a}^i = (u^i, v^i, w^i)^T$. Let each RGB image acquired be denoted ρ^i , and we also make use of luminance images, $L^i = \sum_{k=1}^3 \rho_k^i$. Colour is re-inserted later, as is described in Section 3.4. It is also possible to ‘multiplex’ illumination by combining several lights at once in order to decrease noise [44], but here we simply use one light at a time.

In [6] we use a 6D vector polynomial \mathbf{p} for each normalized light direction three-vector \mathbf{a} as follows:

$$\mathbf{p}(\mathbf{a}) = (u, v, w, u^2, uv, 1), \quad \text{where } w = \sqrt{1 - u^2 - v^2} \quad (1)$$

This differs from the original PTM formulation [1] in that originally the polynomial used had been $(u, v, u^2, v^2, uv, 1)$, which unfortunately does not model a true Lambertian (linear) surface well since it must warp a non-linear model to suit linear data.

Then at each pixel (x, y) separately, we can seek a polynomial regression six-vector of coefficients $\mathbf{c}(x, y)$ in a simple model, regressing lighting directions onto luminance:

$$\begin{bmatrix} \mathbf{p}(\mathbf{a}^1) \\ \mathbf{p}(\mathbf{a}^2) \\ \dots \\ \mathbf{p}(\mathbf{a}^n) \end{bmatrix} \mathbf{c}(x, y) = \begin{bmatrix} L^1(x, y) \\ L^2(x, y) \\ \dots \\ L^n(x, y) \end{bmatrix} \quad (2)$$

E.g. if $n = 50$, then we could write this as

$$\begin{matrix} \mathbf{P} & \mathbf{c}(x, y) & = & \mathbf{L}(x, y) \\ 50 \times 6 & 6 \times 1 & & 50 \times 1 \end{matrix} \quad (3)$$

An example dataset (code named *Barb*) for PTM is displayed in Figure 1b, which was captured with a 50-light dome (i.e. $n = 50$) similar to the one shown in Figure 1a. The dataset *Barb* has large specular and shadowed regions, which cannot be well addressed by the classical PTM model, and such datasets have typically been avoided. Thus, we find *Barb* an ideal representative dataset to test the accuracy and/or robustness of a re-lighting method. On other such difficult datasets we have tried, very similar results were found (see [6] for depictions of shiny and shadowed datasets).

3.2 Robust 6D regression

In our recent version of PTM [6], we solve Equation 3 using a robust LMS regression [7]. The purpose of robust regression is to (1) isolate the matte and specular/shadow components and allow the latter to be more cleanly modelled with an additional RBF interpolation stage and (2) identify the non-matte outliers so that more accurate surface normals as well as other reflectance properties can be obtained with LS. The LMS algorithm as applied in [6] is summarized as follows [7]:

While the 6D LMS regression is slow, it is guaranteed to omit distracting features such as specularities and shadows. Due to the 50% breakdown point of LMS, it requires that at least half plus 1 of the luminance observations belong to a base matte reflectance that can be sufficiently addressed by a polynomial model. Fortunately, this requirement is satisfied for most pixels in real-world datasets. This regression method will be referred to Method:LMS in the following text.

3.3 Re-lighting

The re-lighting of images for PTM is fairly straightforward. Given a new light direction \mathbf{a}' and estimated polynomial coefficients $\mathbf{c}(x, y)$, the approximated luminance can be expressed as:

$$L'(x, y) = \max[\mathbf{p}(\mathbf{a}') \mathbf{c}(x, y), 0] \quad (10)$$

Note that with Method:LMS, $\mathbf{c}(x, y)$ was obtained from a trimmed LS where only the matte observations are used. Therefore, the resulting $L'(x, y)$ is expected to show matte-only contents as well, and non-matte components can be later addressed by other methods (such as the RBF interpolation we will describe in Section 6). This contrasts the robust methods with Method:LS, which uses only PTM to capture both the matte and non-matte components (to some degree) at the same time. Also note that in Equation 10 only luminance is recovered. Colour would

Algorithm: LMS

1. Initialize the iteration counter $q = 1$.
2. Sample a random subset of d indices:
 $J_q \subset \{1, 2, \dots, n\}$ ($|J_q| = d$). In our 6D PTM model, $d = 6$. Get the subset of the polynomial of lighting directions and their corresponding luminance observations indexed by J_q , denoted by \mathbf{P}_{J_q} and \mathbf{L}_{J_q} respectively.
3. Use LS on this subset of observations/lighting directions to estimate the polynomial coefficients \mathbf{c}_{J_q} :

$$\mathbf{c}_{J_q} = \mathbf{P}_{J_q}^\dagger \mathbf{L}_{J_q} \quad (4)$$

4. Use the current estimation \mathbf{c}_{J_q} to approximate the real observation \mathbf{L} and calculate the squared residuals $r_1^2, r_2^2, \dots, r_n^2$ for the n observations in \mathbf{L} , collectively stored as $\mathbf{R}^2 = (r_1^2, r_2^2, \dots, r_n^2)$:

$$\mathbf{R} = \mathbf{L} - \mathbf{P}\mathbf{c}_{J_q} \quad (5)$$

$$\mathbf{R}^2 = \mathbf{R}^T \mathbf{R} \quad (6)$$

5. Get the median M_{J_q} of the n residuals r_1^2, \dots, r_n^2 in \mathbf{R}^2 :

$$M_{J_q} = \text{median}_{i=1, \dots, n} r_i^2 \quad (7)$$

6. Let $q \leftarrow q + 1$ and go back to Step 2 until $q > m$. Here we choose $m = 1,500$ for our typical datasets where $n = 50$ and $d = 6$.
7. Find the smallest median of squared residuals M_{\min} such that:

$$M_{\min} = \min_{q=1, \dots, m} M_{J_q} \quad (8)$$

and keep the estimation of coefficient set \mathbf{c}_{\min} that yields M_{\min} .

8. Calculate the robust standard deviation σ :

$$\sigma = 1.4826 \left(1 + \frac{5}{n-d} \right) \sqrt{M_{\min}} \quad (9)$$

9. Obtain the squared residuals r_i^2 ($i = 1, \dots, n$) with respect to \mathbf{c}_{\min} for each of the n observations. Maintain a binary weight vector $\boldsymbol{\omega}$ of n elements, where $\omega_i = 1$ if $r_i^2 \leq (2.5\sigma)$ and $\omega_i = 0$ otherwise.
 10. Perform trimmed LS using only the observations L_i and polynomial of lighting vector $\mathbf{p}(\mathbf{a}^i)$ with $\omega_i = 1$.
-

albedo α : i.e. $L = s\alpha$. The chromaticity $\boldsymbol{\chi}$ is defined as RGB colour $\boldsymbol{\rho}$, made independent of intensity by dividing by the L_1 norm:

$$\boldsymbol{\rho} = L\boldsymbol{\chi}, \quad L = s\alpha, \quad \boldsymbol{\chi} \equiv \{R, G, B\}/(R + G + B) \quad (11)$$

Suppose our robust regression below delivers binary weights ω , with $\omega = 0$ for outliers. As in [6], once inliers are identified we recover a robust estimate of chromaticity $\boldsymbol{\chi}$ as the median of inlier values, for $k = 1..3$:

$$\chi_k = \text{median}_{i \in (\omega \equiv 1)} (\rho_k^i / L^i) \quad (12)$$

In addition, an estimate of surface normal \mathbf{n} is given by a trimmed PST: with the collection of directions \mathbf{a} stored in the $n \times 3$ matrix \mathbf{A} , suppose ω^0 is an index variable giving the inlier subset of light directions: $\omega^0 = (\omega \equiv 1)$. Using just the inlier subset, a trimmed version of PST gives an estimate of normalized surface normal $\hat{\mathbf{n}}$ and albedo α via

$$\tilde{\mathbf{n}} = (\mathbf{A}(\omega^0))^\dagger \mathbf{L}(\omega^0); \quad \alpha = \|\tilde{\mathbf{n}}\|, \quad \hat{\mathbf{n}} = \tilde{\mathbf{n}}/\alpha \quad (13)$$

where \mathbf{A}^\dagger is the Moore-Penrose pseudoinverse. Other weighting functions are also possible, such as the triangular function used by Method:QUANTILE which we will briefly describe in Section 4.1.

With chromaticity $\boldsymbol{\chi}$ in hand, Equation 11 gives RGB pixel values $\boldsymbol{\rho}$ for the interpolated luminance L , and (13) above also gives us the properties albedo α and surface normal $\hat{\mathbf{n}}$ intrinsic to the surface.

Institutional users of the PTM approach are indeed interested in appearance modelling for re-lighting, but they are also separately interested in surface properties, especially accurate surface normals, which carry much of the shape information.

4 Robust chromaticity/luminance modes

In this section, we present our first main contribution. As we mentioned in Section 3.2, despite its high robustness LMS can be very slow. Therefore, it is necessary to find a less computationally expensive robust method. Here, we suggest a simplified form of LMS - the mode-finder approach.

4.1 Robust mode-finder algorithm

The basic idea of a mode-finder is first to identify a central value of either luminance or chromaticity, termed ‘mode’ across all the observations at every pixel then perform trimmed LS only using the observations that are with a certain range around the mode. This is a far simpler problem than LMS. For reference, we call this new method Method:MODE, which can be achieved with the following algorithm [7]:

be re-introduced by multiplying the chromaticity and the albedo as in Equation 11 as discussed next.

3.4 Colour, normals and albedo

The luminance L consists of the sum of colour components: $L = R + G + B$. Luminance is given by the shading s (e.g. this could in the simplest case be Lambertian shading, meaning surface normal dotted into light direction) times

Algorithm: MODE

1. Initialize the iteration counter $q = 1$.
2. Compute the n squared residuals r_i^2 ($i = 1, \dots, n$) with respect to the scalar observation indexed by q :

$$r_i^2 = (L_i - L_q)^2 \quad (14)$$

3. Find the median M_q of r_i^2 :

$$M_q = \text{median}_{i=1, \dots, n} r_i^2 \quad (15)$$

4. Let $q \leftarrow q + 1$ and go back to step 2 until $q > n$ ($n = 50$ for our typical datasets).
5. Find the smallest median of squared residuals M_{\min} such that

$$M_{\min} = \min_{q=1, \dots, n} M_q \quad (16)$$

and keep the observation of L_{\min} that yields M_{\min} .

6. Same as steps 8 to 10 in Algorithm: LMS, except that in mode-finder, the dimensionality $d = 1$ and the residuals are calculated with respect to L_{\min} .
-

The rationale of Method:MODE is that non-matte outlier observations usually take extreme values in luminance (for instance, shadowed and specular pixels may have an intensity close to 0 and 1, respectively), or their chromaticity may deviate from other matte observations (for instance, specular observations are usually more desaturated whereas shadowed regions appear darker).

Method:MODE may seem to be merely another example of previous thresholding methods. In a typical method of this type [45], the top 10% and the bottom 50% of luminance observations are simply discarded. Then, coefficient values sought are found using a triangular function to weight lighting directions in the resulting range. As in [6], we refer to this simple method as Method:QUANTILE and denote the original PTM method as Method:LS. However, Method:MODE is different from Method:QUANTILE in that the inlier range is calculated based on the distribution of the observation values rather than the empirical values and heuristic triangle functions previously employed. Simply put, Method:MODE lets the data itself dictate what values are in- and outliers.

4.2 Mode-finder versus LMS

In essence, both Method:LMS and Method:MODE attempt to fit a mathematical model to as many data points as possible by minimizing the median of residuals and then identify an inlier range around the fitted model. All observations that fall outside of this range are deemed outliers. The only difference between the two methods is

the mathematical model used: Method:LMS fits the data with a 6D polynomial model, whereas Method:MODE approximates the observations with one single scalar constant, i.e. a 1D mathematical model.

To see how the outlier identification works in the two methods, we study a particular pixel in the *Barb* dataset (marked by a yellow cross in Figure 2a). In Figure 2b,c, the actual luminance observations at this pixel location from 50 lighting conditions are represented as either black solid dots (if they are identified as inliers) or red crosses (for outliers) and are sorted in ascending order. For comparison, the approximated luminance values are shown as blue circles. An observation is classified as outlier if (1) its value is outside the inlier band, marked with green shade enclosed by blue dashed lines or (2) its *approximated* value (blue circle) is negative. Note that the major difference between Method:LMS (Figure 2b) and Method:Mode (Figure 2c) is that the 6D polynomial model in LMS generates an inlier band that closely approximates the actual data curve, whereas the 1D constant model in Method:MODE creates a wider, horizontal band. Despite this seemingly crucial difference, Method:MODE as a matter of fact correctly captures most of the outliers identified by Method:LMS. Although Method:MODE may throw away more data points than necessary, it would not negatively affect the accuracy of estimated polynomial coefficients since these unnecessarily excluded data points are matte anyway and a robust method is not affected by the sum of squared residuals as in LS.

Figure 3 shows a more detailed comparison on outlier estimation and surface property recovery using LMS and mode-finder. Since there is no ground truth data for these properties available, we simply adopt the results obtained with the full 6D LMS method as our 'gold standard' [6] and compare the relative performance of mode-finder against it. Figure 3a displays accuracy of outlier detection in terms of precision, recall and f-statistic, and shows that as long as we use modes for luminance we can achieve a very accurate set of outliers. Results using luminance are shown using white bars. The black bars represent the results obtained by the chromaticity mode, which will be covered in Section 4.3. Figure 3b shows the results for recovered surface normal vectors using outlier detection based on the simpler mode-finder, compared to Method:LMS: the median angular error is 3.03° , which is quite small. Figure 3c shows error in three-vector chromaticity, again measured in terms of angle: the median error is 5.93° , which is quite acceptable. Figure 3d shows errors in albedo - the median is only 0.0037 (where the maximum correct albedo is 1.5855). Such small differences are quite reasonable as a tradeoff with having a much less complex algorithm.

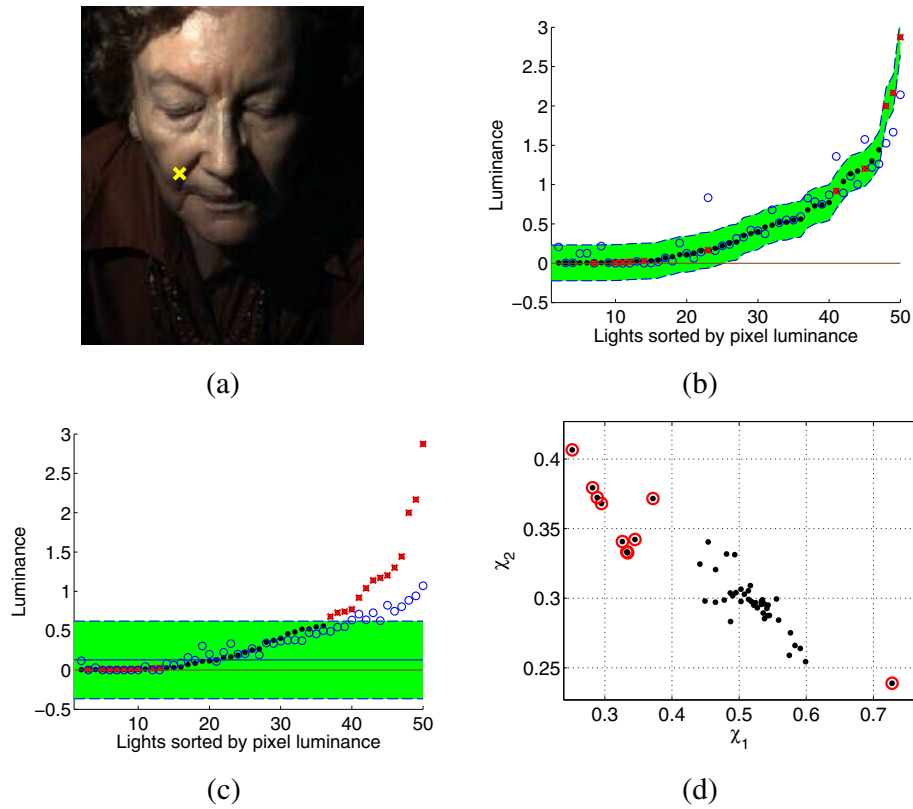


Figure 2 Comparison of outlier detection with LMS and with mode-finding approach. (a) One original image; consider pixel at yellow ‘x’. **(b)** Outlier detection with 6D to 1D LMS regression. Here, the pixels are displayed in ascending order sorted by luminance. The approximated values are shown as blue circles. Inlier pixel values (at this 1 pixel, over the set of 50 lights) with estimated \hat{L} values that fall within the green inlier band are displayed as black dots and outlier measured luminances, including values with negative luminance estimates that fall below the horizontal line at $L = 0$, as black dots with red crosses. The blue dashed lines indicate the boundaries of the inlier band automatically identified by LMS. **(c)** Outlier detection with luminance-mode finder. Here the blue solid line shows the location of the (scalar) mode, bracketed by a horizontal inlier band; inliers also exclude negative- \hat{L} lights. **(d)** Red vs. green chromaticity, with outliers for green mode showing red circles (see Section 4.3).

4.3 Luminance versus chromaticity modes

As mentioned earlier in Subsection 4.1, the mode-finder can be applied on luminance but as well could be applied to colour components, since non-matte observations tend to have an altered chromaticity. For example, in Figure 2c, we have shown the outliers identified by Method:MODE on luminance. In Figure 2d, we apply mode-finder on green chromaticity only and find that the observations with outlying green components (red circles) tend to have outlying red chromaticities as well. In addition, the chromaticity outliers are also expected to largely overlap with the luminance outliers.

It is also possible to combine outliers obtained from different chromaticities or even mix luminance/chromaticity outliers in the hope of getting a more accurate outlier estimation. For example, we can estimate outliers using green chromaticity (this subset of outlier indices are denoted c_{green}) and red chromaticity (c_{red}) at the same time, and then take the outliers c that appear in both c_{green} and c_{red} ,

i.e. $c = c_{\text{green}} \cap c_{\text{red}}$. We refer to such a combined method as ‘green & red’.

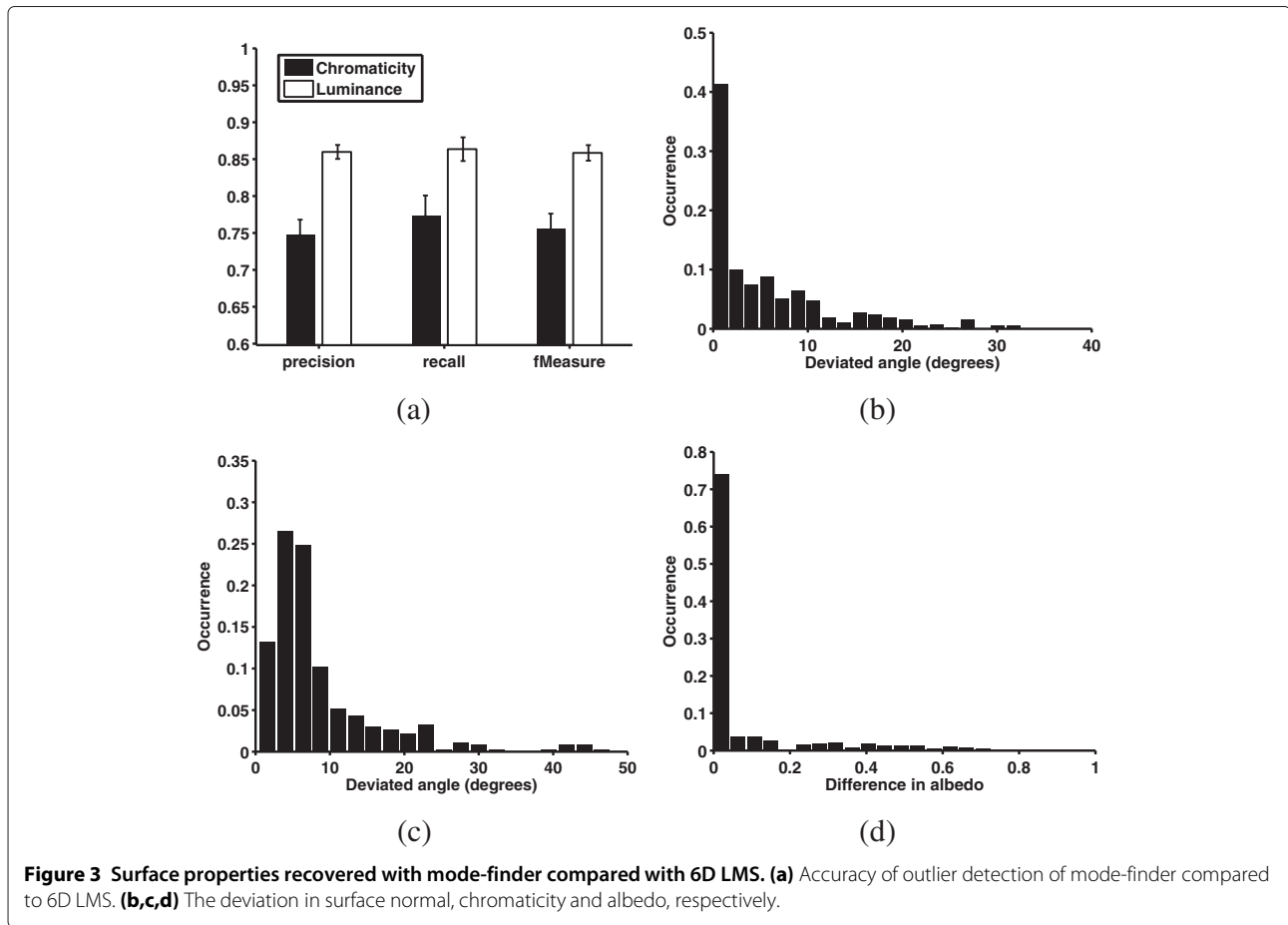
Now the question is: which combination of modalities gives the best approximated appearance? We found [46] that in terms of peak signal-to-noise ratio (PSNR) accuracy of the reconstructed appearance for Method:MODE, we have an ordering:

$$\text{Lum} > (\text{green \& red \& lum}) > \text{green} > (\text{green \& red}) > \text{lum (Method:QUANTILE)}$$

where ‘>’ means better accuracy; using luminance alone is always best, (green & red) seems to be slightly worse than green only, and (green & red & lum) is between green and luminance. In comparison, using luminance with Method:QUANTILE has the worst performance.

5 Higher-dimensional LS-based PTM and hemispherical harmonics

In this section, we present our second contribution. First, we investigate what can be gained by increasing the



dimensionality of the classical PTM model above 6D without including robust regression. In addition, we apply PTM with different dimensions to model luminance and chromaticity separately. The objective of this part of the investigation is to show that one can, in fact, go quite a long way towards accuracy of appearance modelling using only high-dimensional smooth regression, without the final step of RBF modelling, provided we separate modelling of luminance and chrominance.

Secondly, aside from polynomials, other sets of basis functions can be used to model lighting-surface interaction. One notable example is HSH [34] - it has also been suggested that one could replace a PTM polynomial basis by HSH instead [47]. HSH is mathematically very similar to SH which have already been extensively employed in computer graphics. The key difference between HSH and SH is that HSH is only defined for light directions that live on an upper hemisphere, making it more appropriate for our experimental setup.

The conclusions we reach are that (1) a higher dimension does indeed substantially improve the quality of the reconstructed appearance; (2) if we *split* the problem into modelling luminance and chrominance separately,

rather than applying PTM to each component of colour, then we can reduce the dimensionality for chrominance, compared to that for luminance - we find that 16D for luminance and 9D for chrominance work well; and (3) surprisingly, PTM works better than HSH. Note that every dataset we tried behaved this same way.

5.1 Separation of luminance and chromaticity using LS-based PTM

Our first observation is that the quality of the reconstructed images has a positive correlation with the dimensionality of PTM. Suppose we model luminance only, using an LS-based simple PTM. Figure 4a shows accuracy of the approximated input image set, in terms of PSNR, for different dimensionalities d . In order to calculate the overall PSNR between the original and the approximated set of images, we make the individual images into collages, as the one shown in Figure 1b, and compute the similarity between the original and approximated collages. Here we traverse d values 1, 4, 6, 9 and 16. We see that the reconstructed image quality improves steadily as dimensionality increases for both PTM and HSH (which will be covered in Section 5.2), and in fact PTM produces an

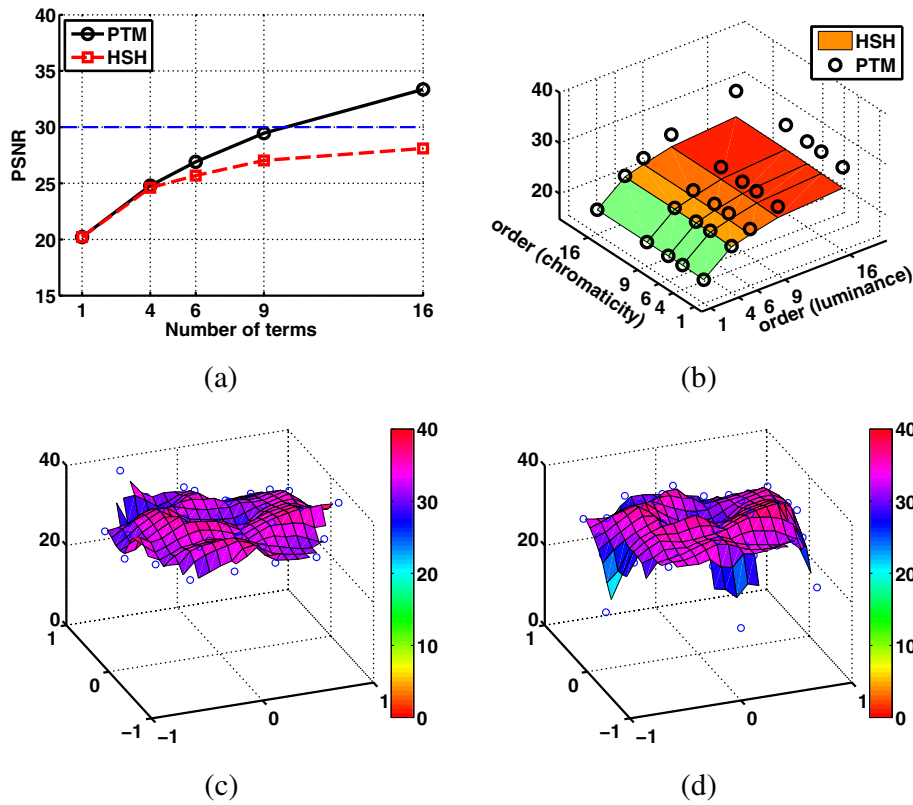


Figure 4 Quality of entire image set reconstructed with non-robust, LS-based PTM and HSH over range of dimensionalities. **(a)** PSNRs for PTM (black curve) and HSH (red, dashed curve), for luminance images, over values of the basis set dimension; the horizontal blue dashed line indicates PSNR = 30. Here the PSNR value displayed is for the entire set of input images compared to the approximated set. **(b)** PSNRs for PTM (scattered circles) and HSH (surface) in RGB images versus the dimensions for luminance and chromaticity. **(c,d)** PSNRs for approximated images for each of the lighting direction images, for PTM and HSH, respectively. PSNR is plotted as against the x and y components of the lighting direction. Blue circles indicate PSNRs for individual reconstructed image in the dataset, and the coloured surface shows an interpolation surface. Here, 16D for luminance and 9D for chromaticity are used.

acceptable (chosen to be $\text{PSNR} \geq 30$ dB) reconstruction at $d = 16$.

Second, we also investigate modelling the luminance and chromaticity separately, using different dimensionalities for each. (Note that only two of the components of χ need be modelled, since $\sum_{k=1}^3 \chi_k \equiv 1$). Figure 4b shows results for dimension of luminance versus chrominance, for HSH (coloured surface) and PTM (black circles). We see that while a higher dimension for luminance is important (as in Figure 4a), the accuracy of approximation of chrominance is only mildly dependent upon dimension. The actual PSNR values plotted in Figure 4b are shown in Table 1.

Due to the two observations made above, we conclude that the quality of the reconstructed images is mainly determined by the luminance, rather than the chromaticities. Hence, in order to achieve a high PSNR with a given dimensionality, it is reasonable to assign a higher dimensionality for luminance and a relatively lower dimensionality for chromaticities. Here we adopt $d = 16$ for

luminance and $d = 9$ for chromaticities, making the total number of dimensions $16 + 9 \times 2 = 34$.

5.2 Comparison of higher-dimension PTM and HSH

Using the LS-based approach, we use either a polynomial matrix P or an HSH equivalent, which we denote as S . When we solve Equation 3, we also prudently include some Tikhonov regularization [48] in solving for c . The solution of Equation 3 is thus

$$c = P^\dagger L \quad \text{or} \quad c = S^\dagger L \quad (17)$$

where \dagger indicates forming a pseudoinverse using a small amount of regularization, with Tikhonov parameter (denoted τ) of, say, $\tau = 10^{-3}$.

We relegate the definition of HSH to Appendix 1. There we list explicitly the definition of the first 16 HSH basis functions, along with the first 16 PTM polynomials.

Recall that in Figure 4a, HSH is consistently outperformed by PTM of the same dimension. Even at a high

Table 1 Comparison of PTM and HSH over various dimensionalities

Chrom terms		Lum terms				
		1	4	6	9	16
PTM basis	1	19.66	24.19	26.15	28.40	31.15
	4	19.84	24.50	26.53	28.74	31.55
	6	19.82	24.48	26.55	28.79	31.73
	9	19.77	24.45	26.50	28.99	32.27
	16	19.75	24.37	26.45	28.96	32.59
HSH basis	1	19.66	24.02	25.03	26.27	27.16
	4	19.81	24.20	25.19	26.39	27.28
	6	19.79	24.19	25.24	26.45	27.38
	9	19.76	24.17	25.22	26.54	27.50
	16	19.75	24.14	25.21	26.54	27.57

PSNR values for least-squares matte regression over ranges of dimensionalities for luminance and chromaticity. The italicized value is obtained with the combination of dimensionalities 16D for luminance and 9D for chromaticity.

dimension $d = 16$, HSH still cannot produce an acceptable result. Similar results are shown in Figure 4b and Table 1.

We further compare the PSNR for each individual image in the dataset. Figure 4c,d shows PSNR for approximation of each image in the colour image set, using PTM and HSH, respectively. Here, as described in Section 5.1, $d = 16$ for luminance and $d = 9$ for chrominance are used. We see that as well as producing higher PSNR values, PTM also does not lose too much accuracy for lighting directions with large incident angles (lights low to the object), whereas HSH does very poorly at these boundary points.

In Table 2 we summarize statistics for PSNRs in Figure 4c,d and as well include results for applying PTM or HSH to each component of RGB separately: to be comparable with dimensionality of 16 for luminance and 9 for chromaticity (for each of two components), making a total of 34 dimensions, here we model R,G,B with 11D each.

For comparison, we also include results for the RBF modelling in Section 6 below: the PSNR values are not (machine-) infinite because Tikhonov regularization moves the approximation slightly away from exactly reproducing input images.

6 Specularities and shadows: RBF modelling

Following [6] we adopt an RBF network approach for the remaining luminance not explained by the matte model Equation 3. For N -pixel images, the ‘excursion’ H is defined as the set of $(N \times 3 \times n)$ non-matte colour values not explained by the R_{matte} given by the basic PTM matte Equation 3, now extended to functions of the colour channel as well: the approximated colour matte image is given by

$$R_{\text{matte}} = PC\chi, \quad (18)$$

where C is the collection of all luminance-regression slopes. Since we include colour, all RBF quantities become functions of the colour channel as well. Throughout, we use the mode-finder efficient robust outlier finder to determine coefficients C .

Then a set of non-matte excursion colour values H is defined for our input set of colour images, via $H = R - R_{\text{matte}}$ where R is the $(N \times 3 \times n)$ set of input images. We follow [6] in carrying out RBF interpolation for interpolant light directions. But here we use the much

Table 2 Comparison of modelling luminance + chromaticity and RGB

		Mean	Median	Mean bottom quartile	Mean top quartile
Lum + Chrom (16D + 9D×2)	PTM	32.60	32.80	28.58	36.37
	HSH	31.80	32.86	24.67	36.51
RGB (11D×3)	PTM	30.35	30.49	25.71	34.20
	HSH	30.33	31.28	23.68	34.95
RBF	47.94	46.47	36.60	61.63	

PSNR statistics for PTM and HSH, using LS + regularization, for dimensionalities 16 and 9 for luminance and chromaticity and similarly for modelling R,G,B separately with 11D each. For completeness, we also show values for RBF modelling.

faster luminance-mode approach Method:MODE for generating matte images and also for recovering the surface chromaticity, surface normal and albedo.

For a particular input dataset, the RBF network models the interpolated excursion solely based on the direction to a new light \mathbf{a}' : an estimate is given by $\hat{\eta} = \text{RBF}(\mathbf{a}')$. Thus, one arrives at an overall interpolant

$$\hat{R} = \hat{R}_{\text{matte}}(\mathbf{a}') + \hat{\eta}(\mathbf{a}') \quad (19)$$

Since in general we do not possess ground-truth data for acquired image sets, we can characterize the accuracy of appearance-interpolation methods by a leave-one-out analysis. In this approach, we carry out the entire image modelling task but omit, in turn, each of the input set images, thus yielding a modelling dimensionality decreased by 1. Since we know the left-out image's appearance, we can generate an error characteristic by comparing the interpolated image with the actual one.

We will summarize how to use RBF interpolation and appearance reconstruction in Sections 6.1 and 6.2, respectively. Then in Section 6.3, we present a method to optimize the parameters of the radial Gaussian function, which serves as the third contribution in this work.

6.1 RBF

A brief recapitulation of the RBF calculation is in order, so as to explain the mechanism of developing a leave-one-out error measurement below.

As in [6], we first generate a matte interpolation structure from in-sample input images and then use RBF to model the excursion \mathbf{H} , for the part of the input image which cannot be explained by a matte model. So first we model the luminance \mathbf{L} , using either PTM or HSH. E.g. if we decide to use a 16D polynomial $p(\mathbf{A})$, then luminance for in-sample images is modelled by $\mathbf{L}_{\text{matte}} = \mathbf{C} (\mathbf{p}(\mathbf{A}))^\dagger$, where \mathbf{C} is the set of polynomial coefficients. If there are N pixels and n lights, then $\mathbf{L}_{\text{matte}}$ is $N \times n$ and \mathbf{C} is $N \times 16$, and the polynomial term above is $16 \times n$.

We obtain an $N \times 3$ set of chromaticities as in Equation 12 from which we can generate a matte colour image model for in-sample images $\mathbf{R}_{\text{matte}}$, for each if the $i = 1..n$ lighting directions, via

$$\mathbf{R}_{\text{matte}}^i = \text{diag}(\mathbf{L}_{\text{matte}}^i) \boldsymbol{\chi}, i = 1..n \quad (20)$$

The dimensionality of $\mathbf{R}_{\text{matte}}$ is $N \times 3 \times n$. The set of excursions for all the input images \mathbf{H} has this same dimensionality, and $\mathbf{H} = \mathbf{R} - \mathbf{R}_{\text{matte}}$. Because the RBF modelling adopted in [6] includes a so-called polynomial term (actually, linear here), we have to extend \mathbf{H} with a set of $N \times 3 \times 4$ zeros. Call this extended excursion \mathbf{H}' .

For interpolation, we need a set of RBF coefficients $\boldsymbol{\Psi}'$, with dimensionality $N \times 3 \times (n + 4)$. We adopt Gaussian RBF basis functions $\phi(\|\mathbf{a} - \mathbf{a}^i\|)$, $i = 1..n$ (although

of course other functions might be tried, such as multi-quadratic or inverse-multi-quadratic). We call the set $\phi(\|\mathbf{a}^i - \mathbf{a}'\|)$ matrix Φ . Then Φ is extended into an $(n+4) \times (n+4)$ matrix Φ' as in [6].

Then we calculate and store the RBF coefficients $\boldsymbol{\Psi}'$ over all the input lights as follows:

$$\boldsymbol{\Psi}' = \mathbf{H}' (\Phi')^\dagger \quad (21)$$

where the \dagger means the Moore-Penrose pseudoinverse, guarding against reduced rank.

However, here we also extend the pseudoinverse to include some Tikhonov regularization:

$$(\Phi')^\dagger = \left(\Phi'^T \Phi' + \tau \mathbf{I}_{(n+4)} \right) \Phi'^T \quad (22)$$

with Tikhonov parameter τ . Below, we mean to optimize this parameter using a clever mathematical theorem borrowed for this work.

6.2 Appearance reconstruction

Given a novel lighting direction \mathbf{a} , appearance reconstruction from PTM coefficients \mathbf{C} and RBF coefficients $\boldsymbol{\Psi}'$ is quite straightforward: we generate a matte image by multiplying PTM coefficient matrix \mathbf{C} by its corresponding combination of polynomial $\mathbf{p}(\mathbf{a})$ and then use recovered chromaticity $\boldsymbol{\chi}$ to form a colour matte image. Then we form a new Gaussian function ϕ from new lighting direction \mathbf{a} and simply multiply ϕ times the prestored RBF excursion coefficient set $\boldsymbol{\Psi}'$ to generate a single-image excursion value η . The Gaussian radial basis function has the explicit form $\phi(\mathbf{a}_i, \mathbf{a}_j, \sigma) = \exp(-r^2/\sigma^2)$, with radius r for light-direction vectors \mathbf{a}_i and \mathbf{a}_j given by $r = \|\mathbf{a}_i - \mathbf{a}_j\|$.

6.3 Optimization of dispersion σ and of Tikhonov parameter τ

In this subsection, we describe our third contribution, i.e. finding the best values for the Gaussian dispersion σ and the Tikhonov coefficient τ so as to optimize the reconstructed appearance. Since we have no ground truth for real input image sets, we test the accuracy of appearance modelling by simply leaving out one of the n input images at a time and attempting to reconstruct the left-out image.

To this end, here we borrow the work in [49] in determining a best value of the Gaussian dispersion parameter σ to minimize the leave-one-out error. However, here we mean to apply the method given in [49] to a whole image at once and include colour, extend RBF modelling to include the additional polynomial term and, finally and importantly, extend [49] to include Tikhonov regularization and its optimization.

The work [49] defines the optimum σ as that yielding the smallest error in reconstructing a leave-one-out image, using only the information from the other images.

E.g. if the input set consists of 50 images, then we follow through matte and then RBF modelling using only 49 images and attempt to reconstruct the 50th image, and then repeat for each of the 50 light directions.

Modelling on the theorem given in [49] in Appendix 2, we generalize the theorem, which is aimed at optimizing RBF over the dispersion parameter σ , to also optimize over Tikhonov parameter τ . The resulting calculation from this theorem is so fast that it is simple to run any unconstrained non-linear optimizer such as the subspace trust-region method [50].

We find that an approximate colour image reconstruction, for the k th leave-one-out image, is simply as follows:

$$E = \Psi' / \nu \tag{23}$$

$$\hat{R} = R - E$$

where the *error image* E is simply formed from the RBF coefficients Ψ' , and a vector ν generated as the solution to the following simple equation in terms of the $(n + 4) \times (n + 4)$ identity matrix I :

$$\Phi' \nu = I \tag{24}$$

This theorem means that one can very rapidly assess the error generated in a leave-one-out analysis of RBF modelling. Figure 5a shows the PSNR between the actual input image set and the result of matte plus RBF modelling, for an optimal choice of σ and τ . Unsurprisingly, we see that RBF interpolation does best in the center of the cluster of lighting directions and worse when there is less supporting information, near the boundary of the cluster of light directions. We take as the optimum dispersion σ and Tikhonov parameter value τ as those which deliver the highest leave-one-out median PSNR over the set overall.

Table 3 shows PSNR statistics for this leave-out-out RBF test. In comparison, we show in Figure 5b and also in the second line of Table 3 the results of a leave-one-out test using PTM matte modelling alone for dimensions 16 and 9 for luminance and chrominance, with no RBF stage. We notice that in a challenging leave-one-out test for interpolation, PTM does reasonably well. To put these plots in perspective, in Figure 5c, we also show the results for PTM + RBF in a leave-all-in setting: of course, the PSNR for PTM + RBF for leave-all-in is by far the best accuracy. In Figure 5d we show the in-sample correct image closest to the mean value of PSNR values for all leave-one-out RBF

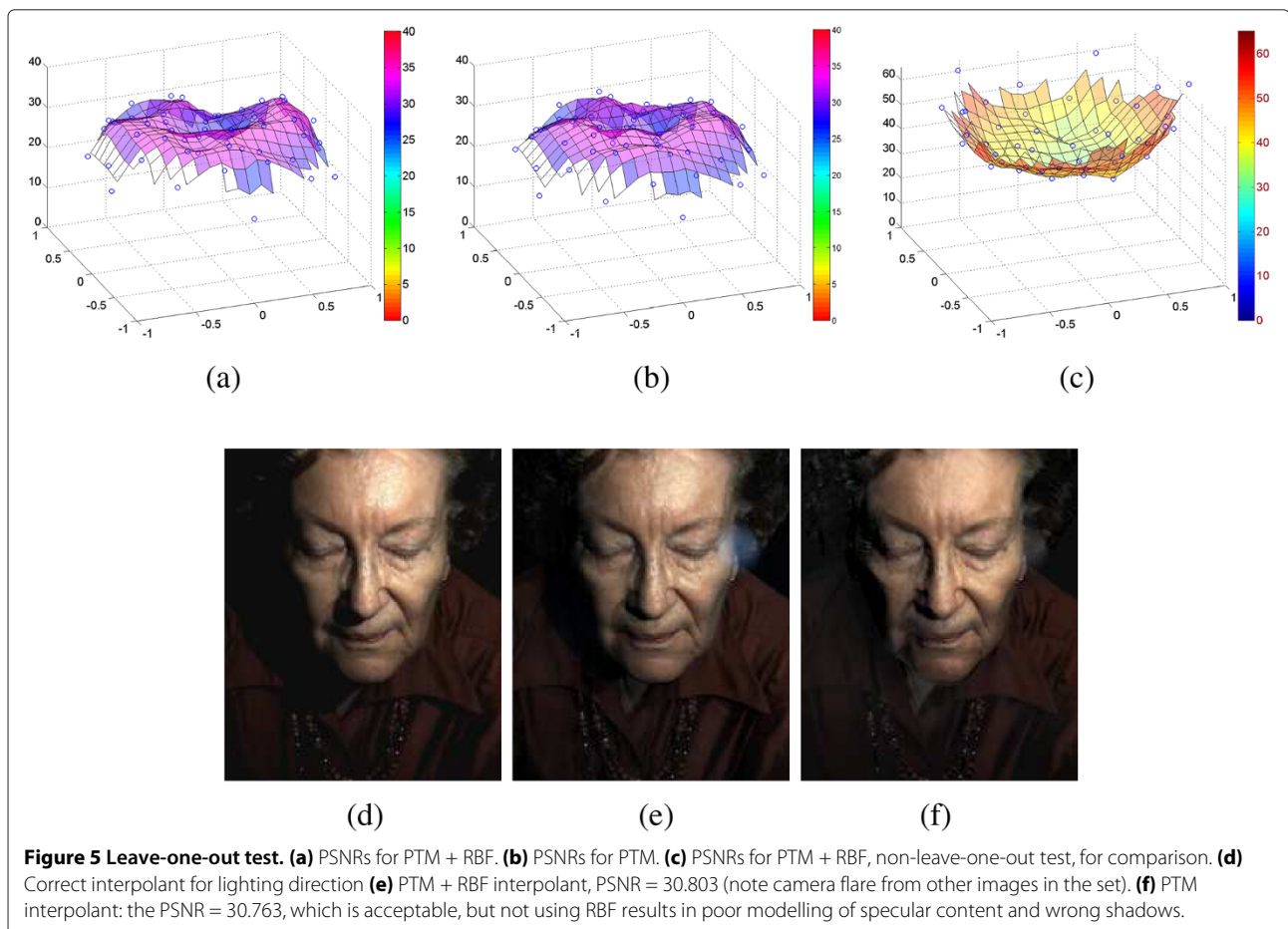


Table 3 PSNR statistics for leave-one-out test, using PTM + RBF, and using only PTM

	Mean	Median	Mean bottom quartile	Mean top quartile
PTM + RBF	30.18	31.24	22.77	35.68
PTM	29.15	29.54	22.23	34.16

modelling, and in Figure 5e,f, we show the interpolants from using PTM + RBF and from using just PTM, respectively. Clearly, RBF provides a substantial boost in visual appearance, although PTM itself (with no RBF stage), with the higher dimensions we have specified, does produce a reasonable image. Nevertheless, qualitatively, using RBF does much better in that, without RBF, specularities are not well modelled and the shadows are wrong.

7 Conclusions

In this paper, we have set out tests and conclusions that improve PTM modelling for appearance interpolation and surface property recovery. We found that increasing PTM dimensionality has a substantial effect on accuracy, more for the luminance channel than for colour. We found that a dimension of 16 for luminance and 9 for chromaticity, modelling luminance and chromaticity separately, delivered good performance. We found that for determining outliers, we could have almost as good accuracy using a much less burdensome robust 1D 'location finder' as in a more accurate but slower robust multivariate processing.

A second stage of modelling using RBF interpolation provides a large boost in accuracy of appearance modelling. Here we showed that Tikhonov regularization in calculating RBF coefficients was important, since we are inverting large matrices; and moreover we incorporated optimizing the Tikhonov parameter into an optimization theorem that had been initially aimed at only generating a best choice of Gaussian dispersion parameter for radial basis function networks.

Future work will include developing a real-time viewer including the new insights gained here.

Appendix 1: hemispherical harmonics

HSH are derived from spherical harmonics (SH) as an alternative set of basis functions on the unit sphere that are particularly aimed at non-negative function values. The familiar SH are defined as [51]

$$Y_l^m(\theta, \phi) = K_l^m e^{im\phi} P_l^{|m|}(\cos\theta), l \in N, -l \leq m \leq l \quad (25)$$

where $\theta \in [0, \pi]$ is the altitude angle, and $\phi \in [0, 2\pi]$ the azimuth angle. P_l^m are the associated Legendre

polynomials, orthogonal polynomial basis functions over $[-1, +1]$, and K_l^m are the normalization factors for these.

$$P_l^m(x) = \frac{(-1)^m}{2^l l!} \sqrt{(1-x^2)^m} \frac{d^{l+m}}{dx^{l+m}}(x^2-1)^l \quad (26)$$

$$K_l^m = \sqrt{\frac{(2l+1)(l-|m|)!}{4\pi(l+|m|)!}}$$

In the context of computer graphics, real-valued functions as follows are often preferred:

$$Y_l^m = \begin{cases} \sqrt{2} K_l^m \cos(m\phi) P_l^m(\cos\theta), & m > 0 \\ \sqrt{2} K_l^m \sin(-m\phi) P_l^{-m}(\cos\theta), & m < 0 \\ K_l^0 P_l^0(\cos\theta), & m = 0 \end{cases} \quad (27)$$

However, since in graphics the incident and reflected lights are all distributed on an upper hemisphere, it requires a large number of coefficients to handle the discontinuities at the boundary of the hemisphere when the mapping is represented with basis defined on a full sphere [34]. Thus, it is more natural to use an HSH basis instead. In this study, we used the HSH model proposed in [34]:

$$H_l^m = \begin{cases} \sqrt{2} \tilde{K}_l^m \cos(m\phi) \tilde{P}_l^m(\cos\theta) & m > 0 \\ \sqrt{2} \tilde{K}_l^m \sin(-m\phi) \tilde{P}_l^{-m}(\cos\theta) & m < 0 \\ \tilde{K}_l^0 \tilde{P}_l^0(\cos\theta) & m = 0 \end{cases} \quad (28)$$

where \tilde{P}_l^m and \tilde{K}_l^m are the 'shifted' associated Legendre polynomials and the hemispherical normalization factors, respectively, defined as follows:

$$\tilde{P}_l^m(x) = P_l^m(2x-1) \quad (29)$$

$$K_l^m = \sqrt{\frac{(2l+1)(l-|m|)!}{2\pi(l+|m|)!}}$$

Now the hemispherical functions are defined only over the upper hemisphere, $\theta \in [0, \pi/2]$, $\phi \in [0, 2\pi]$.

Figure 6 shows the first three 'bands' of the HSH, i.e. $l = 0..2$, and the first 16 functions are stated explicitly in Equation 31.

Similarly, we can also consider the polynomial basis in Equation 1 as a set of functions defined on the hemisphere by representing the lighting direction (u, v, w) with spherical polar coordinates: $u = \sin\theta \cos\phi$, $v = \sin\theta \sin\phi$, $w = \cos\theta$, so e.g. the PTM basis functions in Equation 1 are given by

$$(\sin\theta \cos\phi, \sin\theta \sin\phi, \cos\theta, \sin^2\theta \cos^2\phi, \sin^2\theta \cos\phi \sin\phi, 1) \quad (30)$$

For comparison, a selection of nine polynomial terms are visualized as surface plots in Figure 7, and the first 16 polynomial terms are listed in Equation 32

$$H_i = H_i^m; i = ((l + 1)l - m) + 1; \text{ Order} = (l + 1) :$$

Order 1:

$$H_1(\theta, \phi) = 1/\sqrt{(2\pi)}$$

Order 2:

$$H_2(\theta, \phi) = \sqrt{(6/\pi)}(\cos(\phi)\sqrt{(\cos(\theta) - \cos(\theta)^2)})$$

$$H_3(\theta, \phi) = \sqrt{(3/(2\pi))}(-1 + 2 \cos(\theta))$$

$$H_4(\theta, \phi) = \sqrt{(6/\pi)}(\sin(\phi)\sqrt{(\cos(\theta) - \cos(\theta)^2)})$$

Order 3:

$$H_5(\theta, \phi) = \sqrt{(30/\pi)}(\cos(2\phi)(-\cos(\theta) + \cos(\theta)^2))$$

$$H_6(\theta, \phi) = \sqrt{(30/\pi)}(\cos(\phi)(-1 + 2 \cos(\theta))\sqrt{(\cos(\theta) - \cos(\theta)^2)})$$

$$H_7(\theta, \phi) = \sqrt{(5/(2\pi))}(1 - 6 \cos(\theta) + 6 \cos(\theta)^2)$$

$$H_8(\theta, \phi) = \sqrt{(30/\pi)}(\sin(\phi)(-1 + 2 \cos(\theta))\sqrt{(\cos(\theta) - \cos(\theta)^2)})$$

$$H_9(\theta, \phi) = \sqrt{(30/\pi)}((-\cos(\theta) + \cos(\theta)^2) \sin(2\phi))$$

(31)

Order 4:

$$H_{10}(\theta, \phi) = 2\sqrt{(35/\pi)} \cos(3\phi)(\cos(\theta) - \cos(\theta)^2)^{3/2}$$

$$H_{11}(\theta, \phi) = \sqrt{(210/\pi)} \cos(2\phi)(-1 + 2 \cos(\theta))(-\cos(\theta) + \cos(\theta)^2)$$

$$H_{12}(\theta, \phi) = 2\sqrt{(21/\pi)} \cos(\phi)\sqrt{(\cos(\theta) - \cos(\theta)^2)}(1 - 5 \cos(\theta) + 5 \cos(\theta)^2)$$

$$H_{13}(\theta, \phi) = \sqrt{(7/(2\pi))}(-1 + 12 \cos(\theta) - 30 \cos(\theta)^2 + 20 \cos(\theta)^3)$$

$$H_{14}(\theta, \phi) = 2\sqrt{(21/\pi)} \sin(\phi)\sqrt{(\cos(\theta) - \cos(\theta)^2)}(1 - 5 \cos(\theta) + 5 \cos(\theta)^2)$$

$$H_{15}(\theta, \phi) = \sqrt{(210/\pi)}(-1 + 2 \cos(\theta))(-\cos(\theta) + \cos(\theta)^2) \sin(2\phi)$$

$$H_{16}(\theta, \phi) = 2\sqrt{(35/\pi)} \sin(3\phi)(\cos(\theta) - \cos(\theta)^2)^{3/2}$$

Constant term:

$$P_1 = 1$$

Linear terms:

$$P_2 = u = \sin(\theta) \cos(\phi)$$

$$P_3 = v = \sin(\theta) \sin(\phi)$$

$$P_4 = w = \cos(\phi)$$

(32)

Quadratic terms:

$$P_5 = u^2 = \sin^2(\theta) \cos^2(\phi)$$

$$P_6 = uw = \sin(\theta) \cos^2(\phi)$$

$$P_7 = uv = \sin^2(\theta) \cos(\phi) \sin(\phi)$$

$$P_8 = vw = \sin(\theta) \cos(\phi) \sin(\phi)$$

$$P_9 = v^2 = \cos^2(\phi)$$

Cubic terms:

$$P_{10} = u^3 = \sin^3(\theta) \cos^3(\phi)$$

$$P_{11} = u^2v = \sin^3(\theta) \cos^2(\phi) \sin(\phi)$$

$$P_{12} = u^2w = \sin^2(\theta) \cos^3(\phi)$$

$$P_{13} = uvw = \sin^2(\theta) \cos^2(\phi) \sin(\phi)$$

$$P_{14} = v^2u = \sin^3(\theta) \sin^2(\phi) \cos(\phi)$$

$$P_{15} = v^2w = \sin^2(\theta) \sin^2(\phi) \cos(\phi)$$

$$P_{16} = v^3 = \sin^3(\theta) \sin^3(\phi)$$

Appendix 2: leave-one-out optimization in RBF

It is useful to state explicitly how the optimization theorem in [49] goes over to the situation when Tikhonov regularization comes into play.

Firstly, we utilize three-band colour image data, rather than scalar data, and process whole images at once using vectorized programming in Matlab. However for clarity, below we state matters as they pertain to a single pixel and in one colour band.

Suppose there are n lights and n input values at a pixel, e.g. for our exemplar dataset $n = 50$. Then we make $(n + 4) \times (n + 4)$ matrix $\Phi(\sigma)$, where here we are explicitly including dependence on a variable dispersion value σ . For the $(n + 4)$ vector of excursion values \mathbf{H} (extended by four zeros to include the 'polynomial' RBF part), we begin by solving for the $(n + 4)$ vector set of RBF coefficients ψ , which is the vector solution for the modelling equation

$$\mathbf{H} = \Phi \psi$$

However, instead of simply using a matrix inverse in order to guard against numerical instability, we make use of the Tikhonov regularized inverse from Equation 22:

$$\psi = \Phi(\sigma, \tau)^\dagger \mathbf{H}$$

so that in fact we generate only approximate, not exact, approximations $\hat{\mathbf{H}}$ for in-sample lighting directions:

$$\hat{\mathbf{H}} = \Phi \psi = \Phi \Phi^\dagger \mathbf{H}$$

Then the main task is interpolation to any new light \mathbf{a}' via

$$\eta' = \sum_{j=1}^{(n+4)} \psi_j^T \phi(\mathbf{a}_j - \mathbf{a}')$$

where η' is the scalar value of interpolated excursion (for this pixel and colour channel).

Now we mean to consider the leave-one-out problem, meaning that all the matrices and vectors have extent $(n + 3)$ because the k th input-image case has been omitted. Suppose we denote this case using superscript (k) . That is, we aim for a solution $\psi^{(k)}$ of

$$\mathbf{H}^{(k)} = \Phi^{(k)} \psi^{(k)} \quad (33)$$

Firstly, consider the following *Lemma*: if vector \mathbf{v} has $\mathbf{v}_k = 0$, then

$$\text{if } \mathbf{A}\mathbf{v} = \mathbf{b}, \text{ then } \mathbf{A}^{(k)}\mathbf{v}^{(k)} = \mathbf{b}^{(k)}$$

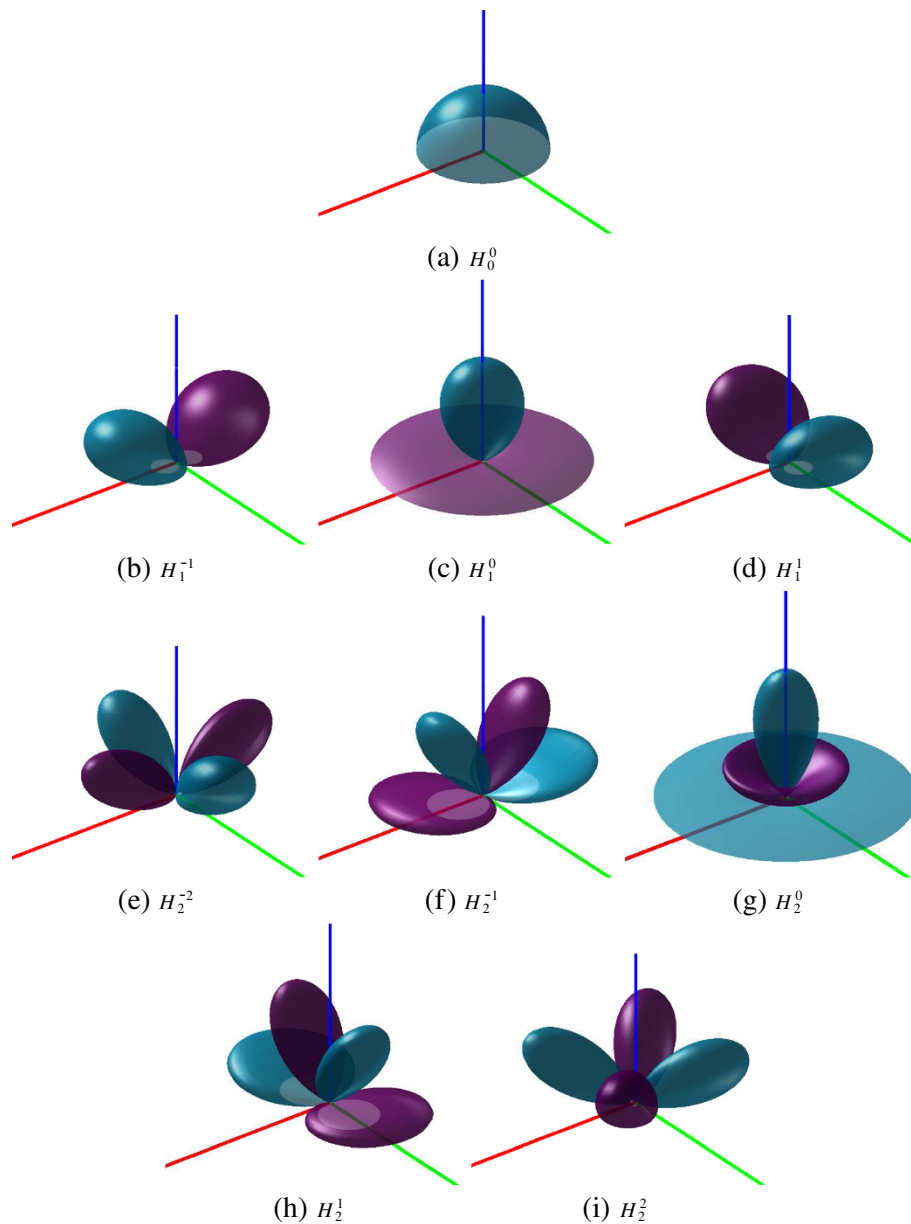


Figure 6 Visualization of the first three bands of hemispherical harmonics. (a) H_0^0 . (b) H_1^{-1} . (c) H_1^0 . (d) H_1^1 . (e) H_2^{-2} . (f) H_2^{-1} . (g) H_2^0 . (h) H_2^1 . (i) H_2^2 . The distance r from the origin to any point (θ, ϕ, r) on the plot surface is proportional to the value of H_l^m at direction (θ, ϕ) , with cyan indicating positive values and purple negative. Red, green and blue indicate x, y and z axes, respectively.

That is, if we know the not-reduced-dimension equation holds, then for the special situation in which $v_k = 0$, we can simply omit whatever value b_k may take on, for the reduced-dimension problem indicated by (k) .

Now consider an auxiliary full-dimension vector \mathbf{v} defined such that

$$\mathbf{v} = \Phi^+ \mathbf{e}_k$$

where \mathbf{e}_k is the k th column of the unit matrix.

Now define a new vector

$$\boldsymbol{\beta} = \boldsymbol{\psi} - (\psi_k/v_k)\mathbf{v}$$

Notice that the k th component of $\boldsymbol{\beta}$ is zero.

Now evaluate $\Phi\boldsymbol{\beta}$:

$$\Phi\boldsymbol{\beta} = \Phi\boldsymbol{\psi} - (\psi_k/v_k)\Phi\mathbf{v} = \hat{\boldsymbol{\eta}} - (\psi_k/v_k)\mathbf{e}_k$$

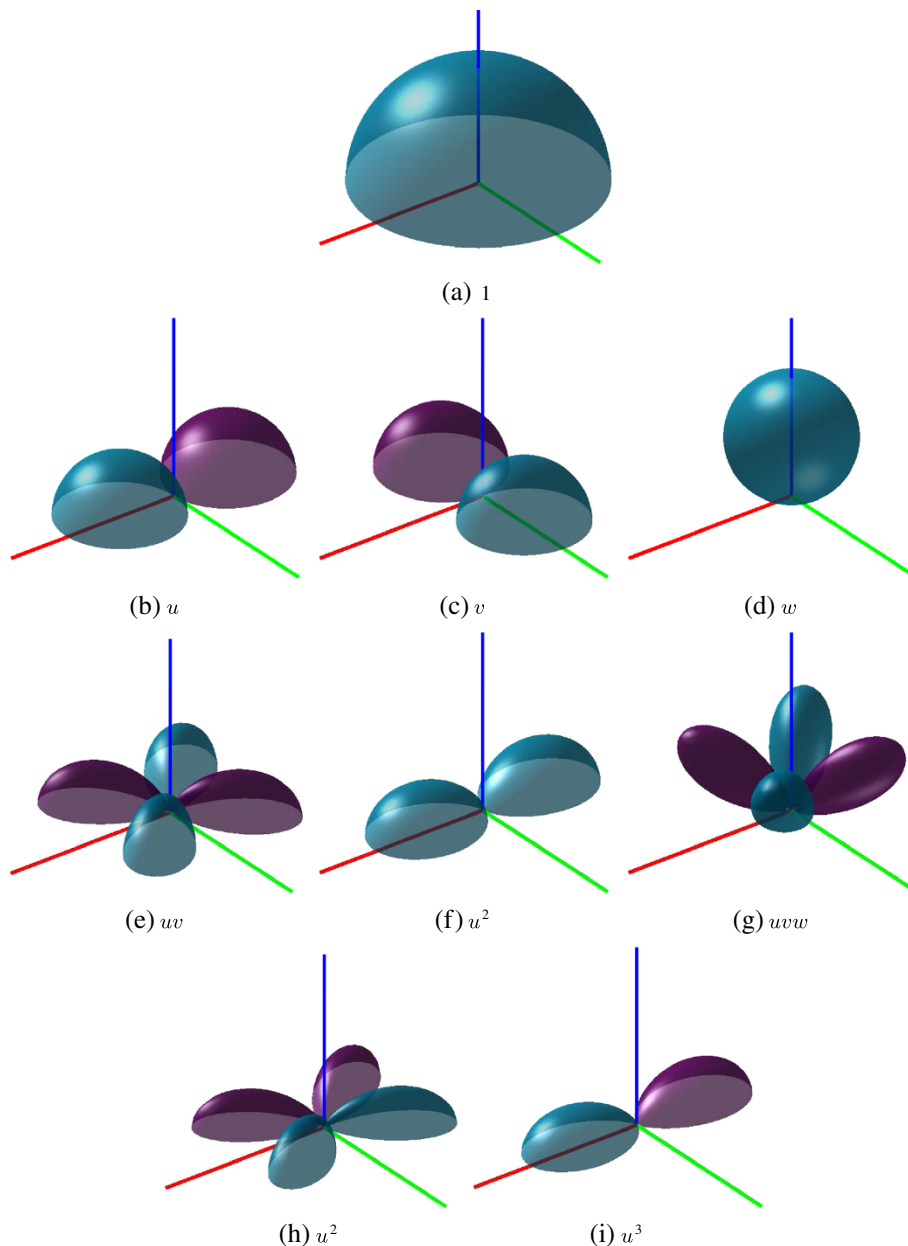


Figure 7 Visualization of selected polynomial basis functions. (a) 1, (b) u , (c) v , (d) w , (e) uv , (f) u^2 , (g) uvw , (h) u^2 , (i) u^3 . The distance r from the origin to any point (θ, ϕ, r) on the plot surface is proportional to the value of the polynomial term P at direction (θ, ϕ) , with cyan indicating positive values and purple negative. Red, green, blue indicate x, y, z axes, respectively.

Hence, by our lemma, β is the sought solution for the leave-one-out set of coefficients $\psi^{(k)}$; however, this statement is approximate and not exact because $\hat{\eta}$ is only approximately (but very close to being equal to) η .

So in order to optimize on σ and τ , we need only to generate the error estimate E_k for the k th case,

$$E_k = -(\psi_k/v_k)$$

for each of the $k = 1..n$ left-out lights, and apply some appropriate error measure such as median (E_k) for choosing the least-error solution:

$$\min_{\{\sigma, \tau\}} \text{median}_{k=1}^n E_k(\sigma, \tau) \quad (34)$$

In practise, we found that utilizing this leave-one-out calculation is very fast and generates smaller interpolation errors when the resulting solution pair $\{\sigma, \tau\}$ is used for

general interpolation for the dataset being optimized for by this leave-one-out procedure.

Competing interests

The authors declare that they have no competing interests.

Received: 10 August 2013 Accepted: 16 April 2014

Published: 1 May 2014

References

1. T Malzbender, D Gelb, H Wolters, Polynomial texture maps, in *Proceedings of Computer Graphics, SIGGRAPH*, vol. 2001 (ACM Los Angeles, California, 2001), pp. 519–528
2. RJ Woodham, Photometric method for determining surface orientation from multiple images. *Opt. Eng.* **19**, 139–144 (1980)
3. M Mudge, T Malzbender, C Schroer, M Lum, New reflection transformation imaging methods for rock art and multiple-viewpoint display, in *7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage* (Eurographics, Nicosia, Cyprus, 2006)
4. K Sunkavalli, T Zickler, H Pfister, Visibility subspaces: uncalibrated photometric stereo with shadows, in *11th European Conference on Computer Vision—ECCV2010* (Springer Heraklion, 2010), pp. 251–264
5. G Earl, K Martinez, T Malzbender, Archaeological applications of polynomial texture mapping: analysis, conservation and representation. *J. Archaeological Sci.* **37**, 1–11 (2010)
6. M Drew, Y Hel-Or, T Malzbender, N Hajari, Robust estimation of surface properties and interpolation of shadow/specularity components. *Image Vis. Comput.* **30**(4-5), 317–331 (2012)
7. PJ Rousseeuw, AM Leroy, *Robust Regression and Outlier Detection*. (Wiley, New York, 1987)
8. T Zickler, S Enrique, R Ramamoorthi, P Belhumeur, Reflectance sharing: image-based rendering from a sparse set of images, in *Eurographics Symposium on Rendering Techniques* (Konstanz, Germany, 2005), pp. 253–265
9. T Zickler, R Ramamoorthi, S Enrique, P Belhumeur, Reflectance sharing: predicting appearance from a sparse set of images of a known shape. *IEEE Trans. Patt. Anal. Mach. Intell.* **28**, 1287–1302 (2006)
10. EN Coleman Jr, R Jain, Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry. *Comput. Graph. Image Process.* **18**, 309–328 (1982)
11. F Solomon, K Ikeuchi, Extracting the shape and roughness of specular lobe objects using four light photometric stereo. *IEEE Trans. Patt. Anal. Mach. Intell.* **18**, 449–454 (1996)
12. S Barsky, M Petrou, The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *IEEE Trans. Patt. Anal. Mach. Intell.* **25**(10), 1239–1252 (2003)
13. H Rushmeier, G Taubin, A Guéziec, Applying shape from lighting variation to bump map capture, in *Eurographics Rendering Techniques 97* (Springer Vienna, 1997), pp. 35–44
14. A Yuille, D Snow, Shape and albedo from multiple images using integrability, in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 1997* (San Juan, Puerto Rico, 1997), pp. 158–164
15. G Willems, F Verbiest, W Moreau, H Hameeuw, K Van Lerberghe, L Van Gool, Easy and cost-effective coneiform digitizing, in *Proceedings of 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (Short and Project Papers)* (Pisa, Italy, 2005), pp. 73–80
16. J Sun, M Smith, L Smith, S Midha, J Bamber, Object surface recovery using a multi-light photometric stereo technique for non-Lambertian surfaces subject to shadows and specularities. *Image Vis. Comput.* **25**, 1050–1057 (2007)
17. Julià C, F Lumberras, AD Sappa, A factorization-based approach to photometric stereo. *Int. J. Imag. Syst. Tech.* **21**, 115–119 (2011)
18. L Wu, A Ganesh, B Shi, Y Matsushita, Y Wang, Y Ma, Robust photometric stereo via low-rank matrix completion and recovery, in *Computer Vision—ACCV2010*, ed. by Kimmel R, Klette R, Sugimoto A, and Lecture notes in computer science, no. 6494. (Springer Berlin Heidelberg, 2011), pp. 703–717
19. KL Tang, CK Tang, TT Wong, Dense photometric stereo using tensorial belief propagation, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2005*, vol. 1 (San Diego, California, 2005), pp. 132–139
20. M Chandraker, S Agarwal, D Kriegman, ShadowCuts: photometric stereo with shadows, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2007* (Minneapolis, Minnesota, 2007), pp. 1–8
21. F Verbiest, L Van Gool, Photometric stereo with coherent outlier handling and confidence estimation, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2008* (Anchorage, Alaska, 2008), pp. 1–8
22. Q Yang, N Ahuja, Surface reflectance and normal estimation from photometric stereo. *Comput. Vis. Image Underst.* **116**(7), 793–802 (2012)
23. S Kherada, P Pandey, A Nambodiri, Improving realism of 3D texture using component based modeling, in *2012 IEEE Workshop on Applications of Computer Vision (WACV)* (Breckenridge, Colorado, 2012), pp. 41–47
24. SH Westin, JR Arvo, KE Torrance, Predicting reflectance functions from complex surfaces, in *Proceedings of the 19th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '92* (New York, 1992), pp. 255–264
25. TT Wong, PA Heng, SH Or, WY Ng, Image-based rendering with controllable illumination, in *Proceedings of the Eurographics Workshop on Rendering Techniques 1997* (Springer Vienna, 1997), pp. 13–22
26. JS Nimeroff, E Simoncelli, J Dorsey, Efficient re-rendering of naturally illuminated environments, in *Photorealistic Rendering Techniques, Focus on Computer Graphics* (Springer Berlin Heidelberg, 1995), pp. 373–388
27. J Kautz, PP Sloan, J Snyder, Fast, arbitrary BRDF shading for low-frequency lighting using spherical harmonics, in *Proceedings of the 13th Eurographics Workshop on Rendering Techniques* (Pisa, Italy, 2002), pp. 291–296
28. R Ramamoorthi, P Hanrahan, An efficient representation for irradiance environment maps, in *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques* (Los Angeles, California, 2001), pp. 497–500
29. PP Sloan, J Kautz, J Snyder, Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments, in *ACM Transactions on Graphics (TOG)*, vol. 21 (ACM, New York, 2002), pp. 527–536
30. PP Sloan, J Sloan, J Hart, J Snyder, Clustered principal components for precomputed radiance transfer. *ACM Trans. Graph.* **22**(3), 382–391 (2003)
31. R Basri, DW Jacobs, Lambertian reflectance and linear subspaces. *IEEE Trans. Patt. Anal. Mach. Intell.* **25**, 218–233 (2003)
32. R Basri, D Jacobs, I Kemelmacher, Photometric stereo with general, unknown lighting. *Int. J. Comput. Vis.* **72**, 239–257 (2007)
33. L Zhang, D Samaras, Pose invariant face recognition under arbitrary unknown lighting using spherical harmonics, in *Biometric Authentication. Lecture notes in computer science, vol 3087* (Springer Heidelberg, 2004), pp. 10–23
34. P Gautron, J Křivánek, SN Pattanaik, K Bouatouch, A novel hemispherical basis for accurate and efficient rendering, in *Eurographics Symposium on Rendering Techniques 2004* (Eurographics Association, Norköping, Sweden, 2004), pp. 321–330
35. S Elhabian, H Rara, A Farag, 2011 Canadian Conference on Computer and Robot Vision (CRV) (St. Johns, Newfoundland, 2011), pp. 293–300
36. H Huang, L Zhang, D Samaras, L Shen, R Zhang, F Makedon, J Pearlman, Hemispherical harmonic surface description and applications to medical image analysis, in *Third International Symposium on 3D Data Processing, Visualization, and Transmission* (Chapel Hill, North Carolina, 2006), pp. 381–388
37. S Elhabian, H Rara, A Farag, Towards accurate and efficient representation of image irradiance of convex-Lambertian objects under unknown near lighting, in *2011 IEEE International Conference on Computer Vision (ICCV)* (Barcelona, Spain, 2011), pp. 1732–1737
38. G Earl, K Martinez, T Malzbender, Archaeological applications of polynomial texture mapping: analysis, conservation and representation. *J. Archaeol. Sci.* **37**(8), 2040–2050 (2010)
39. G Earl, G Beale, K Martinez, H Pagi, Polynomial texture mapping and related imaging technologies for the recording, analysis and presentation of archaeological materials, in *ISPRS Commission V Midterm Symposium* (Newcastle, 21–24 June 2010), pp. 218–223
40. G Earl, PJ Basford, AS Bischoff, A Bowman, C Crowther, J Dahl, M Hodgson, K Martinez, L Isaksen, H Pagi, KE Piquette, E Kotoula, Reflectance transformation imaging systems for ancient documentary artefacts, in *EVA London 2011: Electronic Visualisation and the Arts* (London, 2011)

41. R Bridgman, G Earl, Experiencing lustre: polynomial texture mapping of medieval pottery at the Fitzwilliam Museum, in *Proceedings of the 7th International Congress of the Archaeology of the Ancient Near East (7th ICAANE)*, ed. by R Matthews, J Curtis, M Seymour, A Fletcher, A Gascoigne, C Glatz, SJ Simpson, H Taylor, J Tubb, R Chapman, and Ancient & Modern Issues in Cultural Heritage. Colour & Light in Architecture, Art & Material Culture. Islamic Archeology. (Harrasowitz, London, 2012), pp. 497–512
42. S Duffly, Polynomial texture mapping at Roughting Linn rock art site, in *Proceedings of the ISPRS Commission V Mid-Term Symposium: Close Range Image Measurement Techniques* (Newcastle, 21–24 June 2010), pp. 213–217
43. J Padfield, D Saunders, T Malzbender, Polynomial texture mapping: a new tool for examining the surface of paintings. *ICOM Comm. Conserv.* **1**, 504–510 (2005)
44. Y Schechner, S Nayar, P Belhumeur, Multiplexing for optimal lighting. *IEEE Trans. Patt. Anal. Mach. Intell.* **29**(8), 1339–1354 (2007)
45. A Wenger, A Gardner, C Tchou, J Unger, T Hawkins, P Debevec, Performance relighting and reflectance transformation with time-multiplexed illumination. *ACM Trans. Graph.* **24**(3), 756–764 (2005)
46. M Zhang, MS Drew, Robust luminance and chromaticity for matte regression in polynomial texture mapping, in *Workshops and Demonstrations in Computer Vision–ECCV 2012* Springer Firenze, Italy, 2012), pp. 360–369
47. M Mudge, J Davis, R Scopigno, M Doerr, A Chalmers, O Wang, P Gunawardane, T Malzbender, Image-based empirical information acquisition, scientific reliability, and long-term digital preservation for the natural sciences and cultural heritage, in *Eurographics Tutorials* (Crete, 14–18 April 2008
48. A Tikhonov, V Arsenin, *Solutions of Ill-Posed Problems*. Wiley, New York, 1977)
49. S Rippa, An algorithm for selecting a good value for the parameter c in radial basis function interpolation. *Adv. Comput. Math.* **11**(2–3), 193–210 (1999)
50. T Coleman, Y Li, An interior, trust region approach for nonlinear minimization subject to bounds. *SIAM J. Optimiz.* **6**, 418–445 (1996)
51. M Abramowitz, I Stegun, *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables* (Dover, New York, 1965)

doi:10.1186/1687-5281-2014-25

Cite this article as: Zhang and Drew: Efficient robust image interpolation and surface properties using polynomial texture mapping. *EURASIP Journal on Image and Video Processing* 2014 **2014**:25.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com