

Robust Luminance and Chromaticity for Matte Regression in Polynomial Texture Mapping

Mingjing Zhang and Mark S. Drew

School of Computing Science,
Simon Fraser University
{mza44,mark}@cs.sfu.ca

Abstract. Polynomial Texture Mapping (PTM) is a technique employed in a variety of settings, from museums to in-the-field image capture to multi-illuminant microscopy. It consists of illuminating the surface in question with lights from a collection of light directions, each light in turn. To date, the most accurate interpolation employed in PTM consists of two stages: a matte regression stage followed by a further specular/shadow interpolation. For the first stage, recovering an underlying matte model so as to acquire surface albedo, normals and chromaticity, PTM employs polynomial regression at each pixel, mapping light-direction to luminance. A more accurate model excludes outlier values deriving from specularities and shadows by employing a robust regression from 6-D polynomials to 1-D luminance. Robust methods are guaranteed to automatically find the best representation of the underlying matte content. Here, we retain the idea of using robust methods but instead investigate using a much simpler robust 1-D mode-finder, acting on luminance and on chromaticity components. We then go on to increase accuracy by carrying out 3-D to 1-D regression: this strikes a balance between the best method and the fastest method, with greatly diminished complexity and another large speedup. We show that little accuracy is lost using this much simpler method, and demonstrate the effectiveness of the new method on several image datasets.

1 Introduction

Polynomial Texture Maps (PTM) [1] use 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. 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, in its original Least Squares (LS) formulation at least, 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 LS-based PTM method already in use at major museums in the U.S. 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 [3].

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 3-vectors \mathbf{a} . Then in the original embodiment, a 6-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 recovery of surface properties – surface normal, colour, and albedo – plus the ability to *re-light* pixels using the regression parameters obtained. For re-lighting, the idea is simply that if the regression from in-sample light-directions \mathbf{a} to 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 directions.

In [4], we extend PTM in two ways: (1) First, the LS regression for the underlying matte model is replaced by a robust regression, the Least Median of Squares (LMS) method [5]. 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 knowledge of correctly matte pixels in hand, surface normals, albedos, and pixel chromaticity are more accurately recovered; (2) Secondly, [4] further adds 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 takes care of features such as specularities and shadows that 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 [6,7]. 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 [4].

The main contribution of this paper is to investigate replacing the robust regression from polynomials in \mathbf{a} onto L by two far simpler but still robust methods: first we apply a robust mode-finder to either the components of chromaticity – the pixel colour independent of brightness – or the luminance. The result is a fast identification of outliers, as opposed to a very slow such identification in 6-D to 1-D robust regression. If we are willing to tolerate more time complexity, but with more accuracy, then we can use a 3-D to 1-D robust regression. This reduces PTM to PST but with an essential difference: we use a fully robust method. We found that this simplification introduced little reduction in accuracy compared to a full-polynomial robust version but with a large speedup. Exact solutions for LMS estimation require $O(n^d \log n)$ time. Here we have $n \approx 50$ and either $d = 6$ or $d = 3$ or, in the case of mode-finding, $d = 1$;

the lower bound on algorithmic implementations of LMS have complexity $O(n^d)$ (cf. [5], p. 206), meaning a reduction of 5 (for $d = 3$) and then a further 3 (for $d = 1$) orders of magnitude.

With robust outlier identification in hand, we can then go on to apply a full 6-D to 1-D regression as usual in PTM, but using only (all) the light direction information for inlier directions, resulting in a much simpler technique overall.

For the fastest, robust-mode based method the insight here is that, inside specular or shadow areas, the pixel colour or brightness changes abruptly from some approximate form of a matte colour to instead a chromaticity/luminance associated with shadows or specularities, which will be near the unsaturated region of colour space.

By examining pixel reconstruction error compared to the best, robust 6-D method, we found in fact that for a mode-based method, purely using luminance without chrominance was best. Using mode-finding for 1-D data, in essence we are inverting the process of the robust version of PTM: in the method [4], the robust regression randomly suggests sets of 6 lighting directions and solves for regression coefficients. The optimum set is that regression delivering the minimum robust error, with residuals identified as too large (according to another robust measure) flagged as outliers. Here, in contrast, we do not first seek potential regression parameters but, instead, simply flag outliers as light-directions giving chromaticities/luminances not near the mode value for that pixel. Then we go on to find a 6-D to 1-D regression as usual, but using a simple trimmed LS method and all the inlier directions at once. The 3-D to 1-D linear robust regression stands between the two extremes of best- and simplest-regression.

2 Matte Modelling Using PTM

2.1 Luminance

PTM models smooth dependence of images on lighting direction via polynomial regression. Here we briefly recapitulate PTM as amended by [4]. 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 $\boldsymbol{\rho}^i$, and we also make use of luminance images, $L^i = \sum_{k=1}^3 \rho_k^i$. Colour is re-inserted later, below.

Robust 6-D Regression. With [4] we use a 6-D vector \mathbf{p} for each light direction 3-vector \mathbf{a} as follows:

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

Then at each pixel (x, y) separately, we can seek a polynomial regression 6-vector of coefficients $\mathbf{c}(x, y)$ in a simple model

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

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

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

The original PTM [1] solves a similar equation, using LS; the recent version [4] solves eq. (3) using a robust LMS regression [5] from matrix \mathbf{P} to vector \mathbf{L} .

Robust Modes. In a mode-finder based approach, we adopt (3), but solve for \mathbf{c} using only inlier directions \mathbf{a} , with the decision of whether a direction is an inlier based not on the relation of luminance to light direction as in PTM but instead on whether chrominance or luminance is in the cluster around the mode colour over n lights, for this pixel, or is instead excluded from that cluster. This is a far simpler problem. For reference, we call this method Method:MODE.

The simple method outlined in [8] is similar in a sense. That method simply drops the top 10-percentile of luminance values associated with lights (at the current pixel) and also drops the bottom 50% of luminances. Then, coefficient values sought are found using a triangular function to weight lighting directions in the resulting range. Along with [4] we refer to this simple method as Method:QUANTILE, and denote the original PTM method as Method:LS. The robust method [4] based on 6-D LMS is denoted Method:LMS. The difference from [8] in this paper is that, in keeping with the most accurate method, Method:LMS, here we seek a *robust* way to identify outliers, rather than simply using a heuristic. That is, we let the data itself dictate what are in- and outliers.

Robust Linear-PTM. Instead of eq. (3) we could use only the *linear* part of \mathbf{p} , i.e., the lighting directions \mathbf{a} themselves; however here we apply a *robust* regression, and solve for new, 3-D coefficients \mathbf{c} in

$$\mathbf{A} \mathbf{c}(x, y) = \mathbf{L}(x, y) \quad (4)$$

where the entire collection of light directions \mathbf{a} is given by the $n \times 3$ matrix \mathbf{A} . This 3-D approach turns out to be remarkably successful and accurate, notwithstanding a large reduction in complexity compared to 6-D. Let us call this Method:LINEAR.

Re-lighting. For the above methods, we obtain regression coefficients \mathbf{c} directly from robust LMS, from Method:LMS or Method:LINEAR. For Method:MODE, we use a trimmed version of (3), wherein we drop outliers identified by the mode-finder, to obtain 6-D \mathbf{c} . For all methods, matte interpolation proceeds by generating a new luminance value L for new light direction \mathbf{a}' via interpolation $L(x, y) = \max[\mathbf{p}(\mathbf{a}') \mathbf{c}(x, y), 0]$ or the dimension-3 version for Method:LINEAR, with negative values set to zero. Colour is re-introduced by multiplying by the chromaticity times the albedo as follows.

2.2 Colour, Normals, 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 be Lambertian shading,

meaning surface normal dotted into light direction), times albedo α : i.e., $L = s\alpha$. The chromaticity χ is *defined* as RGB colour ρ independent of intensity:

$$\rho = L\chi, \quad L = s\alpha, \quad \chi \equiv \{R, G, B\}/(R + G + B) \quad (5)$$

Suppose our robust regression below delivers binary weights ω , with $\omega = 0$ for outliers. With [4], we recover a robust estimate of chromaticity χ as the median of inlier values, for $k = 1..3$:

$$\chi_k = \operatorname{median}_{i \in (\omega \equiv 1)} (\rho_k^i / L^i) \quad (6)$$

And 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

$$\tilde{\mathbf{n}} = (\mathbf{A}(\omega^0))^+ \mathbf{L}(\omega^0); \quad \alpha = \|\tilde{\mathbf{n}}\|, \quad \hat{\mathbf{n}} = \tilde{\mathbf{n}} / \alpha \quad (7)$$

where \mathbf{A}^+ is the Moore-Penrose pseudoinverse. Method:QUANTILE instead uses matrices weighted by the triangular functions.

With chromaticity χ in hand, eq. (5) gives RGB pixel values ρ for the interpolated luminance L . And (7) above also gives us the properties albedo α and surface normal $\hat{\mathbf{n}}$ intrinsic to the surface.

3 Robust Chromaticity/Luminance Modes

Firstly, let us start by using Method:LMS, regressing from 6-D $\mathbf{p}(\mathbf{a})$ to luminance L [4]. That is, we are solving eq. (3) for \mathbf{c} , by robustly finding the best slope. Fig. 1(a) displays one image from a dataset of images taken from $n = 50$ angles.¹ Here, a yellow \mathbf{x} shows a single pixel, one that takes on matte, shadowed, and specular values depending on the lighting direction. Fig. 1(b) shows luminance values L at this pixel, in sorted- L order, for $n = 50$. LMS regression automatically identifies outliers, here shadows and specularities, that do not adhere closely enough to a matte model (3). However the objective of LMS is mainly to trim outlier values so as to obtain the best possible values of slopes $\mathbf{c}(x, y)$. LMS is guaranteed to perform optimally provided at least $\lfloor n/2 \rfloor + 1$ values are indeed inliers. Here we also flag as outliers any unphysical negative values \hat{L} .

Thus in our situation we are using LMS to perform a regression from 6-D values \mathbf{P} onto 1-D luminance values L . The band shown in Fig. 1(a) is automatically chosen by LMS such that within the band, residuals $\hat{L} - L$ that are too large in absolute value are flagged as outliers. The figure shows measured luminance values as black dots, regression estimates \hat{L} as blue circles, and outliers with a red box.

¹ Data courteously supplied by Tom Malzbender, HP.

The complexity of LMS is high, but has strong mathematical guarantees. The advantage of using a robust mechanism for finding slopes is that we will then also find demonstrably best estimates of surface normals and albedo [4]. How LMS proceeds is repeatedly sub-sampling to get best slope parameters \mathbf{c} , establishing an initial set of inliers, and then re-calculating the inlier band using only the previous set of survivors.

Alternatives could be: ignoring the problem of outliers (e.g., [1]); or perhaps applying a heuristic, as in [8], and making a simple guess for potential outliers.

3.1 Luminance Modes

In this paper we wish to strike a balance between a simple but fast method, on the one hand, and a best-accuracy but high-complexity regression on the other. Since regression (3) is aimed at luminance, let us consider what we can learn from luminance alone, without regard for an underlying matte model. Since luminance is scalar, what we require in this simpler approach is a *robust* method of flagging luminance values that are potential shadows or specular highlights. Therefore we make use of the much simpler robust LMS regression denoted “location finding” [5], meaning determining the location of the *mode* of scalar values. Fig. 1(c) again shows sorted-luminance values, but now the inlier band and outlier determination is determined simply using the luminance values themselves, without regard to the 6-D to 1-D matte model. Outliers are initially determined simply by whether L values are sufficiently close to the mode value, shown as a horizontal line. The inlier band is indicated by another two horizontal lines bracketing this mode value.

Again, negative approximations \hat{L} are deemed outliers. In order to obtain luminance approximations we go back to the full regression model (3) but restrict the calculation of slopes \mathbf{c} to a trimmed LS calculation as follows:

$$\mathbf{c} = (\mathbf{P}(\boldsymbol{\omega}))^+ \mathbf{L}(\boldsymbol{\omega}), \quad \hat{L} = \mathbf{P} \mathbf{c} \quad (8)$$

where now $\boldsymbol{\omega}$ are binary weights from the much simpler, 1-D LMS location finder.

3.2 Chromaticity Modes

Another approach might be to use the observation that in fact in a shadow or specularity, the colour of the pixel changes substantially. Consider again the chromaticity $\boldsymbol{\chi}$ in eq. (5). Suppose we carry out our 1-D LMS mode-finder idea simply on the scalar green-chromaticity value χ_2 . Plotting red-green 2-D chromaticity values in Fig. 1(d), we show outlier values circled in red. Indeed, this idea does determine outlier chromaticities.

So in this case we would carry out the indicated 1-D regression, followed by generation of inlier slopes and approximations of matte luminance \hat{L} just as in (8) above, but now for inliers determined by chromaticity, not luminance. And, of course, we can also combine Luminance and Chromaticity to use both kinds of mode-finder to help identify outliers.

We found that in order of accuracy for such 1-D estimation approaches, for generating \hat{L} compared to the best method, Method:LMS, we have an order Lum > (Green & Red & Lum) > Green > (Green & Red) > Method:QUANTILE where “>” means better accuracy. I.e., using Luminance alone is always best, (R & G) seems to be slightly worse than Green, and (G & R & Lum) is between Green and Lum; Method:QUANTILE is worst.

Therefore, Method:MODE comes down to using the considerably faster 1-D robust mode-finder as a new mechanism for determining matte image output as per eq. (8). To encompass both matte plus specular/shadow information, we then go on to include RBF networks for modelling this non-smooth information, as in [4].

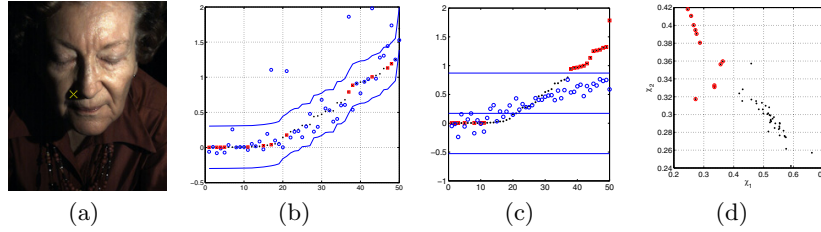


Fig. 1. (a): One original image, with light with spherical angles $\{\theta, \phi\} = \{39.8^\circ, -0.4^\circ\}$. (b): 6-D to 1-D LMS regression: black dots are measured data, blue circles are regression estimates \hat{L} , and outliers are indicated with a red box. (c): Using mode of Luminance: mode value is center horizontal line, bracketed by horizontal inlier band; inliers include negative- \hat{L} lights. (d): Red vs. Green chromaticity, with outliers for Green mode in red.

4 Specularities and Shadows

Following [4] we adopt an RBF network approach for the remaining luminance not explained by the matte model (3). The “excursion” $\boldsymbol{\eta}$ is defined as the set of $(N \times 3 \times n)$ non-matte colour values not explained by the \mathbf{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

$$\mathbf{R}_{matte} = \mathbf{P} \mathbf{C} \boldsymbol{\chi}, \quad \text{or} \quad \mathbf{R}_{matte} = \mathbf{A} \mathbf{C} \boldsymbol{\chi} \quad \text{in Method:LINEAR} \quad (9)$$

where \mathbf{C} is the collection of all luminance-regression slopes. Including colour, all RBF quantities become functions of the colour channel as well.

Then an $(N \times 3 \times n)$ set of non-matte excursion colour values $\boldsymbol{\eta}$ is defined for our input set of colour images, via $\boldsymbol{\eta} = \mathbf{R} - \mathbf{R}_{matte}$ where \mathbf{R} is the $(N \times 3 \times n)$ set of input images. We follow [4] in carrying out RBF interpolation for interpolant light directions. But here we use the much faster luminance-mode approach Method:MODE or Method:LINEAR 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 $\hat{\eta} = RBF(\mathbf{a}')$. Thus one arrives at an overall interpolant

$$\hat{R} = \hat{R}_{matte}(\mathbf{a}') + \hat{\eta}(\mathbf{a}') \quad (10)$$

5 Accuracy

We can determine the best balance to strike between simplicity and accuracy by challenging each method using Gaussian noise added to synthetic data, thus with ground truth known. Let us use the synthetic-Mozart image dataset ($n = 50$ images), comprised of (inherently noisy) range-finder data with Lambertian shading plus specularity.

Fig. 2 shows results for four methods comparing both accuracy of recovery of surface normals as well as albedo, for this dataset. We see that the fastest method, Method:MODE, does not compare well, whereas the more costly but still relatively fast Method:LINEAR compares well with ostensibly the most accurate Method:LMS and a good deal better than the simple Method:QUANTILE. What we learn here is that a robust approach does pay, but we can use a simpler (albeit still robust) method, at least for such basically Lambertian plus shadows plus specularity data.

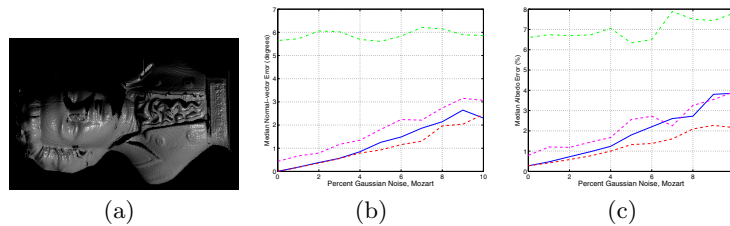


Fig. 2. Gaussian-noise sensitivity for synthetic Mozart image set: (b): One example out of 50 input images. (b): Median angular error for recovered surface normal – blue is Method:LMS, green is Method:MODE, magenta is Method:QUANTILE, and red is Method:Linear. (c): Median percent albedo error.

6 Results and Conclusions

Interpolation results, along with matte image and normal vectors recovered, are shown in Fig. 3.² Given the success and speed of Method:Linear in Fig. 2, in this paper we adopt that method for generating the matte part of interpolants.

² Row 2: Data courteously supplied by Roberto Scopigno, Visual Computing Laboratory, ISTI-CNR, Pisa, Italy; Rows 3-7: Data courteously supplied by Mark Mudge of Cultural Heritage Imaging.

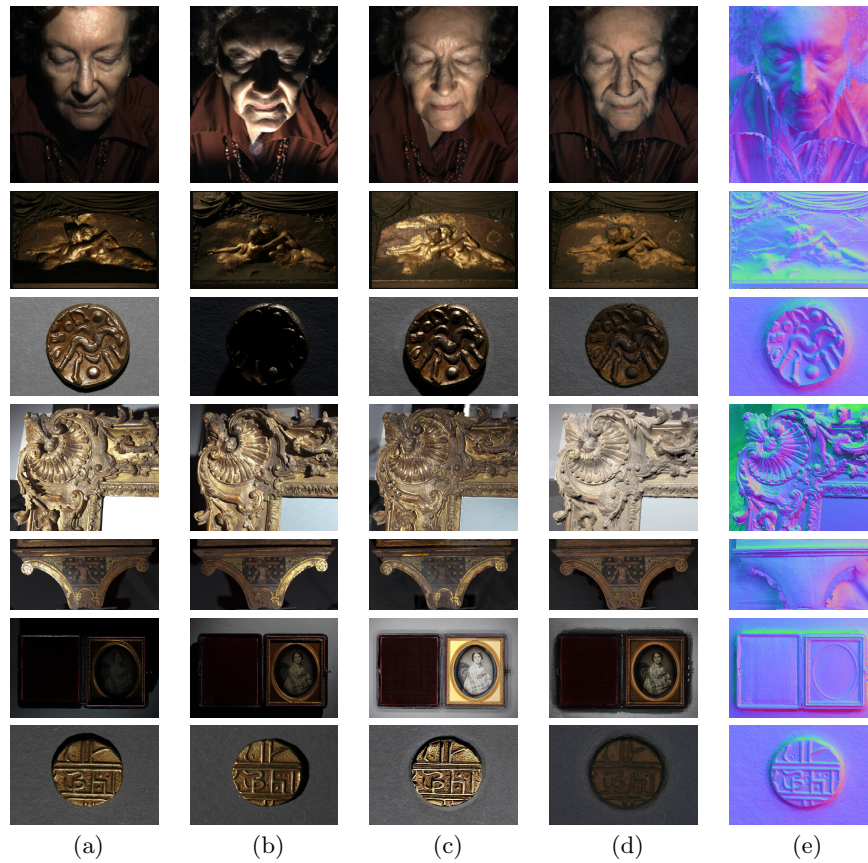


Fig. 3. (a,b): Input images with lighting directions to be interpolated; (c): interpolants for direction given by the mean of directions for (a,b); (d): robust matte image for interpolated angle; (e): normals. Directions for (a,b,c) are row 1: $\{\theta, \phi\} = \{24.5^\circ, 58.7^\circ\}, \{90.0^\circ, -90.0^\circ\} \rightarrow \{36.8^\circ, -71.6^\circ\}$; row 2: $\{52.2^\circ, -63.4^\circ\}, \{83.0^\circ, 134.5^\circ\} \rightarrow \{25.0^\circ, 179.9^\circ\}$; row 3: $\{23.3^\circ, 85.1^\circ\}, \{75.3^\circ, 80.8^\circ\} \rightarrow \{49.3^\circ, 82.1^\circ\}$; row 4: $\{39.2^\circ, -177.1^\circ\}, \{23.5^\circ, 32.4^\circ\} \rightarrow \{11.6^\circ, 148.3^\circ\}$; row 5: $\{34.7^\circ, -142.9^\circ\}, \{42.4^\circ, -31.4^\circ\} \rightarrow \{24.3^\circ, -80.1^\circ\}$; row 6: $\{75.3^\circ, -99.2^\circ\}, \{62.2^\circ, 100.1^\circ\} \rightarrow \{24.0^\circ, -164.8^\circ\}$; row 7: $\{36.6^\circ, -30.5^\circ\}, \{36.7^\circ, -120.5^\circ\} \rightarrow \{27.7^\circ, -75.48^\circ\}$;

The importance of achieving a good model for the matte surface cannot be overstressed, in that the underlying matte foundation gives us the surface properties desired – surface normals, albedo, chromaticity. On top of that fundamental structure, we also wish to derive convincing, accurate, and informative interpolants. To date, the RBF approach is the most straightforward, in that it involves no (spatial) Reflectance Sharing [6], which needs geometric knowledge.

From the results, one can see that a combination of using a fast matte model, as proposed here for the first part of the interpolation process, along with accurate RBF modelling, does generate interpolants with useful specular and shadow information. Further work will entail approaches to speed up the second part of the interpolation as well.

References

1. Malzbender, T., Gelb, D., Wolters, H.: Polynomial texture maps. In: Computer Graphics, SIGGRAPH 2001 Proceedings, pp. 519–528 (2001)
2. Woodham, R.J.: Photometric method for determining surface orientation from multiple images. *Optical Engineering* 19, 139–144 (1980)
3. Mudge, M.: Cultural Heritage Imaging.: Personal Communication (2010)
4. Drew, M., Hel-Or, Y., Malzbender, T., Hajari, N.: Robust estimation of surface properties and interpolation of shadow/specularity components. *Image and Vision Computing* 30, 317–331 (2012)
5. Rousseeuw, P.J., Leroy, A.M.: *Robust Regression and Outlier Detection*. Wiley (1987)
6. Zickler, T., Enrique, S., Ramamoorthi, R., Belhumeur, P.: Reflectance sharing: image-based rendering from a sparse set of images. In: *Eurographics Symposium on Rendering*, pp. 253–265 (2005)
7. Zickler, T., Ramamoorthi, R., Enrique, S., Belhumeur, P.: Reflectance sharing: Predicting appearance from a sparse set of images of a known shape. *IEEE Trans. on Patt. Anal. and Mach. Intell.* 28, 1287–1302 (2006)
8. Wenger, A., Gardner, A., Tchou, C., Unger, J., Hawkins, T., Debevec, P.: Performance relighting and reflectance transformation with time-multiplexed illumination. *ACM Trans. Graph.* 24, 756–764 (2005)