

Estimation of reflectance from camera responses by the regularized local linear model

Wei-Feng Zhang,^{1,2,*} Gongguo Tang,² Dao-Qing Dai,³ and Arye Nehorai²

¹Department of Applied Mathematics, South China Agricultural University, Guangzhou 510642, China

²Department of Electrical and Systems Engineering, Washington University in St. Louis, St. Louis, Missouri 63130, USA

³Center for Computer Vision and Department of Mathematics, Sun Yat-Sen (Zhongshan) University, Guangzhou 510275, China

*Corresponding author: zhangwf@scau.edu.cn

Received July 27, 2011; revised August 17, 2011; accepted August 29, 2011;

posted August 30, 2011 (Doc. ID 151837); published September 30, 2011

Because of the limited approximation capability of using fixed basis functions, the performance of reflectance estimation obtained by traditional linear models will not be optimal. We propose an approach based on the regularized local linear model. Our approach performs efficiently and knowledge of the spectral power distribution of the illuminant and the spectral sensitivities of the camera is not needed. Experimental results show that the proposed method performs better than some well-known methods in terms of both reflectance error and colorimetric error. © 2011 Optical Society of America

OCIS codes: 330.1710, 330.1730.

Many applications require accurate color reproduction, such as art painting archiving and online product exhibition. The most informative way to describe a color is to give its spectral reflectance. If we know the spectral reflectance of an object, it is possible to simulate color values under arbitrary imaging conditions. Since digital cameras are the most widely used devices for color acquisition, developing accurate reflectance estimation methods from three-channel or multichannel camera responses has received increasing attention [1–6].

Suppose the spectral characteristics are sampled uniformly in the visible wavelengths from 400–700 nm, and the number of samples is n . The number of imaging channels is m . The camera response $\mathbf{x} \in \mathbb{R}^m$ can be modeled as

$$\mathbf{x} = \mathbf{S}\mathbf{L}\mathbf{r} + \mathbf{e}, \quad (1)$$

where $\mathbf{r} \in \mathbb{R}^n$ is the spectral reflectance of the object, \mathbf{S} is an $m \times n$ matrix of spectral sensitivities of the camera sensors, \mathbf{L} is an $n \times n$ diagonal matrix having the illuminant's spectral power distribution in the diagonal, and \mathbf{e} is an $m \times 1$ column vector of system noise. For abbreviation, Eq. (1) is usually expressed as

$$\mathbf{x} = \mathbf{M}\mathbf{r} + \mathbf{e}, \quad (2)$$

where $\mathbf{M} = \mathbf{S}\mathbf{L}$ denotes the spectral responsivity matrix.

Given a camera response \mathbf{x} , the reflectance estimation problem attempts to find a pattern $\hat{\mathbf{r}} \in \mathbb{R}^n$ that has the minimized error compared with the actual reflectance \mathbf{r} . Many methods have been developed to tackle this problem [1–6]. These methods can be roughly divided into three categories: the Wiener estimation, finite-dimensional modeling, and regression methods. Wiener estimation has the form

$$\hat{\mathbf{r}} = \mathbf{K}_r \mathbf{M}^T (\mathbf{M} \mathbf{K}_r \mathbf{M}^T + \mathbf{K}_e)^{-1} \mathbf{x}, \quad (3)$$

where \mathbf{K}_r is the covariance matrix of the reflectance and \mathbf{K}_e is the covariance matrix of noise. Performance of Wiener estimation will be poor if \mathbf{K}_r and \mathbf{K}_e are not

calculated accurately [1,2]. Furthermore, Wiener estimation needs to know the spectral responsivity matrix \mathbf{M} . Since it is difficult to measure the spectral responsivity of the imaging system, which could change as a function of environmental parameters such as temperature or humidity, methods that do not require this prior knowledge could be useful. Finite-dimensional modeling exploits the inherent smoothness of reflectance with the use of low-dimensional linear models [3]. The reflectance \mathbf{r} is approximated by a weighted sum of l (usually 5–8) basis functions as

$$\hat{\mathbf{r}} = \mathbf{B}\mathbf{w}, \quad (4)$$

where \mathbf{B} is an $n \times s$ matrix whose column i is the i th basis function B_i , and \mathbf{w} is an $s \times 1$ column vector denoting the basis coefficients. If the spectral responsivity matrix \mathbf{M} is known, the basis coefficients \mathbf{w} can be computed either by pseudoinverse or by nonlinear optimization. The regression methods learn a set of functions

$$\hat{r}_i = f_i(\mathbf{x}), \quad i = 1, \dots, n, \quad (5)$$

from training samples [4,5]. Then the reflectance can be estimated as $\hat{\mathbf{r}} = (\hat{r}_1, \dots, \hat{r}_n)^T$. It is not necessary to know the spectral responsivity matrix \mathbf{M} , but only the spectral reflectance and camera responses of the training samples are used, thus making the method more applicable. However, since the regression functions are built separately for each sampled wavelength, the global smoothness of reflectance is not considered.

Reflectance \mathbf{r} can be seen as a point that is distributed in the n -dimensional vector space \mathbb{R}^n . Using the traditional finite-dimensional modeling in Eq. (4) can then be seen to approximate \mathbf{r} in a low-dimensional linear subspace determined by the vectors $\{B_i, i = 1, \dots, l\}$. Since there is no evidence that the reflectance vectors of natural objects are completely located in a low-dimensional linear space, using fixed basis functions will have limited approximation capability. Lancel *et al.* proved that dictionary based sparse recovery method can typically outperform traditional linear models with fixed basis

functions [6]. However, their method is computationally intensive and needs to know the spectral responsivity matrix \mathbf{M} .

In this Letter, we propose a novel reflectance estimation approach based on the regularized local linear model (RLLM). The reflectance \mathbf{r} for a camera response \mathbf{x} is approximated adaptively by the linear combination of k reflectances from the training set that have similar camera responses to \mathbf{x} . The idea is inspired by the regularized locality preserving method for the preimage problem in kernel principal component analysis [7]. Our method is efficient and we do not need to know the spectral responsivity matrix \mathbf{M} .

Local linear model. Although, in a global view, the reflectance vectors are not located in a low-dimensional linear subspace, we expect each reflectance and its neighbors to lie on or close to a locally linear low-dimensional subspace. Since similar reflectances usually have similar camera responses, we can build the local linear model as

$$\mathbf{r} = \sum_{i=1}^k w_i \mathbf{r}_i, \quad (6)$$

where $\{\mathbf{r}_i, i = 1, \dots, k\}$ are reflectances from the training set that have similar camera responses to \mathbf{x} . The reconstruction weight $\mathbf{w} = (w_1, \dots, w_k)^T$ in Eq. (6) is computed by regularized linear regression as follows.

Regularized linear regression. For a given camera response \mathbf{x} , let $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ be its k nearest neighbors by Euclidean distance from the training set, and $\{\mathbf{r}_1, \dots, \mathbf{r}_k\}$ be the corresponding reflectances of $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$. By substituting Eq. (6) into Eq. (2), we have the relationship

$$\mathbf{x} \approx \sum_{i=1}^k w_i \mathbf{x}_i. \quad (7)$$

Table 1. Summary of Datasets

Dataset	Name	Size
Training set	Munsell Matte	100, 300, 600, 900
Testing set	Agfa IT8.7/2	288
Validation set	Paper	426

Thus we can learn the local linear reconstruction weight $\mathbf{w} = (w_1, \dots, w_k)^T$ in Eq. (6) by solving the regularized linear regression problem

$$\arg \min_{\mathbf{w}} \left\| \mathbf{x} - \sum_{i=1}^k w_i \mathbf{x}_i \right\|^2 + \lambda \|\mathbf{w}\|^2, \quad (8)$$

where the regularization parameter λ is used to prevent overfitting. Let \mathbf{X} be an $m \times k$ matrix in which column i is the i th nearest camera responses \mathbf{x}_i . Then the solution of Eq. (8) can be given by

$$\mathbf{w} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^T \mathbf{x}, \quad (9)$$

where \mathbf{I} is a $k \times k$ identity matrix.

Experimental results. We compare the performance of our method with five other methods: linear least squares regression (LLS), finite-dimensional modeling using the first five or 10 principal component analysis (PCA) basis vectors, whose basis coefficients are computed by linear least squares regression, and kernel regression with the Gaussian kernel and the Doehon spline kernel [4]. Note that our experiments are under the assumption that the spectral responsivity matrix \mathbf{M} is unknown, so Wiener estimation and other methods that need to know \mathbf{M} are not compared here.

Table 2. Estimation Errors of Different Methods on the Validation Set^a

Error Term	Training Size		LLS	PCA-5	PCA-10	G-Kernel	D-Kernel	RLLM
RMSE	100	Mean	0.0513	0.0513	0.0513	0.0505	0.0512	0.0475
		Max	0.0552	0.0552	0.0552	0.0550	0.0544	0.0521
	300	Mean	0.0511	0.0511	0.0511	0.0489	0.0501	0.0457
		Max	0.0533	0.0533	0.0533	0.0519	0.0529	0.0492
	600	Mean	0.0505	0.0505	0.0505	0.0481	0.0497	0.0444
		Max	0.0513	0.0513	0.0513	0.0498	0.0513	0.0454
	900	Mean	0.0507	0.0507	0.0507	0.0479	0.0499	0.0444
		Max	0.0513	0.0513	0.0513	0.0488	0.0510	0.0452
	100	Mean	1.8683	1.8694	1.8682	1.9194	2.0592	1.7004
		Max	2.3566	2.3510	2.3566	2.7644	3.9811	2.0473
	300	Mean	1.8356	1.8377	1.8356	1.6736	1.9102	1.5622
		Max	1.9508	1.9522	1.9508	1.9601	2.3052	1.7912
$\Delta E_{00}(D65)$	600	Mean	1.8051	1.8078	1.8051	1.5900	1.7891	1.4928
		Max	1.8437	1.8465	1.8437	1.8010	1.8605	1.5546
	900	Mean	1.8108	1.8134	1.8108	1.5678	1.7851	1.5122
		Max	1.8514	1.8553	1.8514	1.6248	1.8340	1.5660
	100	Mean	3.3735	3.3788	3.3735	2.9204	3.4429	2.5485
		Max	3.9730	3.9574	3.9730	3.3906	4.1571	3.0835
	300	Mean	3.3208	3.3263	3.3208	2.5791	3.2446	2.2283
		Max	3.5680	3.5707	3.5680	3.0125	3.6729	2.4980
	600	Mean	3.2760	3.2809	3.2760	2.4743	3.1428	2.1220
		Max	3.3769	3.3825	3.3769	2.6460	3.4343	2.2323
	900	Mean	3.2832	3.2883	3.2832	2.4296	3.1249	2.1290
		Max	3.3515	3.3587	3.3515	2.5211	3.3743	2.2062

^aFor each training data size, the results are summarized over 50 trials. In each row the best result is rendered in bold type.

Three datasets of spectral reflectance are used for evaluation and are shown in Table 1. They are all obtained from [8]. Camera responses are simulated based on the spectral sensitivity curves of a real three-channel camera measured by Barnard *et al.* [9]. The illumination light is assumed to be the Commission Internationale de l'Eclairage illuminant D65. The spectra of datasets and the sensor sensitivities are sampled from 400 to 700 nm in 10 nm increments.

Different-sized training sets are randomly selected from the 1269 samples of the Munsell dataset. The training set is used to build the regression functions for the regression methods, compute the PCA basis set for the finite-dimensional modeling, and compute the k nearest neighbors for our method. The testing set is used to select the optimal parameters of the kernel regression methods and our method. The best parameters are chosen empirically from predefined sets. The idea is to find parameters that minimize the reflectance error on the testing set. Finally, the generalization performance of the six methods is assessed on the validation set.

Table 2 shows the reflectance errors and colorimetric errors of different methods on the validation set. For each training data size, the results are summarized over 50 trials. As can be seen, our method provides the lowest reflectance errors and colorimetric errors in every training set size. Kernel regression methods seem to perform better than the linear least squares regression method. In the case of small training set size (100 samples), kernel regression provides slightly lower reflectance errors than the linear least squares regression but higher colorimetric errors. However, when increasing the number of training samples, it is evident that kernel regression performs much better than linear least squares regression. This verifies that the nonlinear regression methods are more suitable for large training set sizes. The performance of finite-dimensional modeling with PCA basis sets is close to that of the linear least squares regression method. The more the principal components are used, the closer the performance is. This proves, in a sense, that the performance of reflectance approximation in the global space cannot be improved by using a fixed low-dimensional subspace in it.

The performance of our method depends on the choices of the regularization parameter λ and the number of nearest neighbors k . Through evaluation, we find our method is not sensitive to these two parameters. The results of our method in Table 2 are given by setting $k = 20$ and $\lambda = 0.0001$. For the different-sized randomly selected training sets, setting $k = 20$ and $\lambda = 0.0001$ can always give satisfactory results. The results given by $k = 20, 30, 40$, and 50 are very similar. The camera responses of the experimental data in Table 2 are simulated with no noise added. We have also investigated the spectral estimation accuracy under two different levels of simulated independent Gaussian noise. We find our method still performs the best.

In summary, our local linear model focuses on the individual reflectance estimation performance rather than on building a global model that fits all the training samples. It performs adaptively and efficiently, and needs mainly the low-order matrix inverse operation. Furthermore, our method does not require prior knowledge of the spectral power distribution of the illuminant and the spectral sensitivities of the camera.

This work was supported in part by the National Natural Science Foundation of China (NSFC, grants 60903094, 11171354, and 90920007).

References

1. N. Shimano, K. Terai, and M. Hironaga, *J. Opt. Soc. Am. A* **24**, 3211 (2007).
2. H. L. Shen and J. H. Xin, *Opt. Lett.* **32**, 96 (2007).
3. X. Zhang and H. Xu, *J. Opt. Soc. Am. A* **25**, 371 (2008).
4. V. Heikkinen, T. Jetsu, J. Parkkinen, M. Hauta-Kasari, T. Jaaskelainen, and S. D. Lee, *J. Opt. Soc. Am. A* **24**, 2673 (2007).
5. W.-F. Zhang and D.-Q. Dai, *J. Opt. Soc. Am. A* **25**, 2286 (2008).
6. S. Lansel, M. Parmar, and B. A. Wandell, *Proc. SPIE* **7246**, 72460D (2009).
7. W.-S. Zheng and J.-H. Lai, in *Proceedings of the 18th International Conference on Pattern Recognition* (2006), Vol. 2, pp. 456–459.
8. Spectral database, University of Joensuu Color Group, <http://spectral.joensuu.fi/>.
9. K. Barnard, L. Martin, B. Funt, and A. Coath, *Color Res. Appl.* **27**, 147 (2002).