Illumination Estimation via Thin-Plate Spline Interpolation

Weihua Xiong*[1], Lilong Shi**, Brian Funt**, Sung-Su Kim***, Byoung-Ho Kang***, Sung-Duk Lee*** and Chang-Yeong Kim***
*Omnivision Corporation, CA, USA[1]
**Simon Fraser University, Burnaby, Canada
***Samsung Advanced Institute of Technology, Korean

Abstract

Thin-plate spline interpolation is used to interpolate the color of the incident scene illumination from an image of the scene. The algorithm can be used to provide color constancy under changing illumination conditions, and automatic white balancing for digital cameras. Thin-plate splines interpolate over a non-uniformly sampled input space, which in this case is a set of training images and associated illumination chromaticities. Tests of the thin-plate spline method on a large set of real images demonstrate that the method estimates the color of the incident illumination quite accurately.

Keywords

colour constancy, illumination estimation, automatic white balancing, thin-plate spline interpolation

Introduction

A new approach to illumination estimation for color constancy and automatic white balancing is developed based on the technique of thin-plate spline interpolation. We describe the illumination in terms of its chromaticity components *r*, *g* which can be viewed as a functions of the image *I* ; namely, $r = f_r(I)$ and $g = f_g(I)$. The problem of illumination estimation becomes that of estimating these two functions. In this paper, we treat their estimation as a problem of interpolation over a set of training images.

Interpolation is a common problem and there are many well-established interpolation methods[1]. The majority of these methods, such as bilinear or bi-cubic interpolation, are based on interpolation over training data sampled on a uniform grid. However, we can not uniformly sample the space of images, so interpolation over a non-uniformly sampled space is required. Thin-plate spline interpolation is an effective interpolation method under these conditions, and has been widely used in the context of deforming one image into registration with another. In the case of illumination estimation, TPS maps image information to the r-chromaticity and g-chromaticity values of the illumination.

Thin Plate Spline Method

As is typical of interpolation methods in general, thin-plate spline (TPS) interpolation constructs a function that matches a given set of data values y_i, corresponding to a given set of data vectors $\overline{X_i} = [X_{i,1}, X_{i,2}, \ldots X_{i,D}]$, in the sense that $y_i = f(\overline{X_i})$.

TPS interpolation was originally designed for 2-dimensional image registration[2-5]. In the color context, it has been extended TPS to 3 dimensions, and successfully applied to the problem of camera and color display calibration [6]. Compared with other methods, TPS has been found to be quite stable and accurate in terms of finding a unique solution without having to tune a lot of parameters. Here, we extend TPS to n-dimensions and apply it to the problem of estimating the chromaticity of a scene's overall incident illumination from an image of that scene. Many previous methods [7-9] have used a color histogram as the input data; however, for TPS we use image thumbnails as input. The thumbnails are 8x8 images created by averaging the underlying pixels in the original input image. These thumbnails in chromaticity coordinates become input vectors of size 8 x 8 x 2 = 128.

TPS for illumination estimation requires a "training" set of *N* images along with their corresponding illumination chromaticity values $\{(I_{i,1}, I_{i,2}, \ldots I_{i,128}), (r_i, g_i)\}$. TPS determines parameters w_i and a_j controlling the two mapping functions f_r, f_g, such that

$$(r_i, g_i) = (f_r(I_{i,1}, I_{i,2}, \ldots I_{i,128}),\ f_g(I_{i,1}, I_{i,2}, \ldots I_{i,128})).$$

The mapping function f_r, is defined as

$$f_r(I_1', I_2', \ldots I_{128}') = \sum_{i=1}^{N} w_i U(\| (I_1', I_2', \ldots I_{128}') - (I_{i,1}, I_{i,}, \ldots I_{i,128}) \|) + a_0 + \sum_{j=1}^{128} a_j I'_j$$
$$\text{where} \quad U(x) = x^2 \log x \qquad (1)$$

The function f_g is defined similarly. The weights w_i control a non-linear term, and the a_j, control an additional linear term.

Each training set pair (an image plus its illumination chromaticity) provides 2 equations. For the i^{th} training image we have

$$\begin{cases} r_i = f_r(I_{i,1}, I_{i,2}, \ldots I_{i,128}) \\ g_i = f_g(I_{i,1}, I_{i,2}, \ldots I_{i,128}) \end{cases} \qquad (2)$$

In addition, a smoothness constraint is imposed by minimizing the bending energy. In the original TPS formulation [1], the bending energy function was defined in 2D. Here we generalize it to higher dimensions defined as α_i *(i= 1..128)*:

$$J(f_r) = \sum_{\alpha_1 + \ldots + \alpha_{128} = 128} \frac{128!}{\alpha_1! \ldots \alpha_{128}!} \int \frac{\partial^{128} f_X}{\partial I_1^{\alpha_1} \ldots \partial I_{128}^{\alpha_{128}}} dI_1 \ldots dI_{128} \qquad (3)$$

where $J(f_r)$ is the total bending energy described in terms of the curvature of f_r. Following [10-12], the energy will be minimized when

$$\sum w_i = 0$$
$$\sum I_{i,1} w_i = 0$$
$$\sum I_{i,2} w_i = 0$$
$$\ldots$$
$$\sum I_{i,128} w_i = 0 \qquad (4)$$

[1] This work was done while the author was a PH.D. student at Simon Fraser University.

Xiong, W., Shi, L., Funt, B., Kim, S., Kang, B., and Lee, S.D.,"Illumination Estimation via Thin-Plate Spline Interpolation," Proc. 15th IS&T Color Imaging Conference, Albuquerque, Nov. 2007.

For each of f_r and f_g, we have *(N+129)* equations in *(N+129)* linear can be uniquely determined using matrix operations. Define *L, W, K, Q* and $\bar{U}$ as follows:

$$L = \begin{bmatrix} \bar{U} & Q \\ Q^T & O \end{bmatrix}$$

$W = [w_1, w_2, \ldots, w_N, a_0, a_1, a_2, \ldots, a_{128}]^T$
$K = [X_1, X_2, \ldots X_N, 0, 0, 0, \ldots. 0]^T$

$$Q = \begin{bmatrix} 1\ I_{1,1}\ I_{1,2} \ldots I_{1,128} \\ 1\ I_{2,1}\ I_{2,2} \ldots I_{2,128} \\ \ldots \\ 1\ I_{N,128}\ I_{N,128}\ \ldots I_{N,128} \end{bmatrix}$$

$$\bar{U} = \begin{bmatrix} 0 & U_{1,2} \ldots\ldots\ldots U_{1,N} \\ U_{2,1} & 0 \ldots\ldots\ldots\ldots U_{2,N} \\ .. & \\ U_{N,1} \ldots\ldots\ldots\ldots\ldots\ldots\ldots 0 & \end{bmatrix}$$

where $U_{ij} = U(\|(I_{i1} \ldots I_{i128}) - (I_{j1} \ldots I_{j128})\|)$ and 0 is the *129x129* matrix of zeroes.

The *(N+129)* equations can then be written $K=LW$, and the solution can be obtained as $W = L^{-1}K$.

Experiments

We implemented the TPS illumination-estimation method in Matlab and conducted experiments to compare its performance to that of other illumination-estimations methods.

Several different error measures are used to evaluate performance. For each image, the distance between the measured actual illumination chromaticity (r_a,g_a) and that estimated by an algorithm (r_e,g_e) is calculated as:

$$E_{i-dist} = \sqrt{(r_a - r_e)^2 + (g_a - g_e)^2} \qquad (5)$$

For a test set of N images we report the root mean square (RMS), mean, and median distance [13] . The RMS is defined in the standard way as:

$$RMS_{dist} = \frac{1}{N}\sqrt{\sum_{i=1}^{N} E_{i-dist}^2} \qquad (6)$$

Given illumination chromaticity r and g, the other component can be obtained as $b = 1 - r - g$. This allows us to compute the angular error between two 3-dimensonal chromaticity vectors as:

$$E_{i-angular} = COS^{-1}\left[\frac{(r_a, g_a, b_a) \circ (r_e, g_e, b_e)}{\sqrt{r_a^2 + g_a^2 + b_a^2} \times \sqrt{r_e^2 + g_e^2 + b_e^2}}\right] \times \frac{2\pi}{360} \quad (7)$$

As with the distance measure, we also compute the RMS, mean, and median angular error over the test set of images.

To evaluate whether there is a significant difference in the performance of two competing methods, the Wilcoxon signed-rank is applied [13]. The threshold for accepting or rejecting the null hypothesis is set to 0.01.

The first experiment is from Barnard's calibrated 321 SONY images [14]. We evaluate the illumination error using the leave-one-out cross-validation procedure [14]. In the leave-one-out procedure, one image is selected for testing, and the remaining 320 images are used for training in order to determine the required parameters. This is repeated 321 times, with a different image left out each time.

The results with corresponding results for the Shades of Gray (SoG)[15], Support Vector Regression (SVR)[7], Max RGB (MAX)[16], and Grayworld (GW)[17] methods are listed in the Table 1 and 2.

We next consider Cardei's [20] set of 900 uncalibrated images taken using a variety of different digital cameras manufactured by Kodak, Olympus, HP, Fuji Polaroid, PDC, Canon, Ricoh and Toshiba. The illumination RGB values for these images were measured from a gray card placed in each scene. Leave-one-out experiments are used once again. The results are shown in tables 3 and 4.

The final test is based on the 7,661 real images extracted from over 2 hours of digital video acquired with a SONY VX-2000. Ciurea et. al. [18] set up a special camera with a matte gray ball attached to it. This was made to appear at a fixed location near the right-bottom corner of each video frame. The average chromaticity value of the pixels in the brightest region is assumed to reflect the RGB of the true scene illumination. To ensure that the grayball has no effect on the results, all images were cropped on the right to remove the grayball. The remaining images are 240 by 240 pixels.

The whole image dataset includes a wide variety of indoor and outdoor scenes, including many with people in them. Since neighboring images in the database tend to be related, we partitioned the database into two independent sets based on geographical location. Subset A includes 3,581 images, and subset B includes 4080. Subset A contains images from the Apache Trail, Burnaby Mountain, Camelback Mountain, CIC 2002 and Deer Lake. Subset B contains images from completely different locations including False Creek, Granville Island Market, Marine Drive, Metrotown shopping center, Scottsdale, Simon Fraser University and Whiteclyff Park. We then used Subset A for training and B for testing and vice versa. The results are listed in Table 5. Tables 6 and 7 give the corresponding Wilcoxon sign test results. The combined errors and corresponding Wilcoxon sign test result from both tests are shown in Table 8 and Table 9.

Conclusion

The problem of estimating the chromaticity of the overall scene illumination is formulated in terms of interpolation over a non-uniformly sampled data set. The chromaticity is viewed as a function of the image and the set of training images is non-uniformly spaced. Thin-plate spline interpolation is an excellent interpolation technique for these conditions and has been shown to work well for illumination estimation in particular. TPS calculates its result based on a weighted combination of the entire set of training data. Hence, for efficiency it is important to keep that set as small as possible, and how best to prune the training set is a direction for future research. Overall, the experiments on real images show the accuracy of TPS illumination estimation to be very good.

Acknowledgement

This research was funded by the Samsung Advanced Institute of Technology.

References

[1] G.D. Knott, *Interpolating Cubic Splines*, Birkhauser Inc, 2000

[2] F. L. Bookstein. “Principal warps: thin-plate splines and decomposition of deformations.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, Issue 6, pp. 567–585 (1989)

[3] M. H. Davis, A. Khotanzad, D. Flamig, and S. Harms. “A physics-based coordinate transformation for 3-d image matching,” *IEEE Tranactions on. Medical Imaging*, Vol. 16, Issue 3, pp. 317–328, (1997)

[4] N. Arad and D. Reisfeld, “Image warping using few anchor points and radial functions,” *Computer Graphics forum,* vol. 14, pp. 35–46 (1995)

[5] Asker M. Bazen and Sabih H. Gerez, “Elastic minutiae matching by means of thin-plate spline models,” *International Conference on Pattern Recognition*, (2002).

[6] W. Xiong and B. Funt, “Nonlinear RGB-to-XYZ Mapping for Device Calibration,” Proceedings of CIC’13, pp. 200-204, (2005)

[7] B.V. Funt, and W.H. Xiong. “Estimating Illumination Chromaticity via Support Vector Regression,” *Proceedings of CIC’12*, pp. 47-52, (2004)

[8] V. Cardei, B. Funt. and K. Barnard, “Estimating the Scene Illumination Chromaticity Using a Neural Network,” *Journal of the Optical Society of America A,* Vol. 19, No. 12, pp. 2374-2386 (2002)

[9] G. D. Finlayson, S. D. Hordley and P. M. Hubel, “Color by Correlation: A Simple, Unifying Framework for Color Constancy,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol.23, No. 11, pp. 1209-1221, (2001)

[10] A. Zandifar, S. Lim, R. Duraiswami, N. Gumerov, and L. S. Davis, "Multi-level Fast Multipole Method for Thin Plate Spline Evaluation,” *IEEE International Conference on Image Processing*, pp. 1683-1686, (2004)

[11] F.L. Bookstein, *Morphometric Tools for Landmark Data Geometry and Biology*, Chapter 2, pp. 23-41, (1991)

[12] J. Meinguet, “Multivariate Interpolation at Arbitrary Points Made Simple,” *Journal of Applied Mathematics and Physics*, Vol.30, pp. 292-304, (1979)

[13] S.D. Hordley, and G.D. Finlayson, “Re-evaluating colour constancy algorithms,” *Proc. 17th ICPR*, Vol. 1, pp. 76-79, (2004)

[14] K. Barnard, L. Martin, A. Coath, and B. Funt, “A Comparison of Computational Colour Constancy Algorithms. Part Two: Experiments on Image Data,” *IEEE Transactions on Image Processing,* No. 11, 2002, pp. 985-996 , (2002)

[15] G. D. Finlayson and E. Trezzi, “Shades of Gray and Colour Constancy,” *Proc. IS&T/SID Twelfth Color Imaging Conference: Color Science, Systems and Applications*, *Society for Imaging Science and Technology*, Scottsdale, pp. 37-41, (2004)

[16] G.D. Finlayson. “Retinex viewed as a gamut mapping theory of color constancy,” *Proc. AIC International Color Association Color 97* Vol. 2, pp. 527-530, (1997)

[17] B. Buchsbaum. “A spatial processor model for object color perception,” *Journal of the Franklin Institute*, Vol. 31, pp. 1-26, (1980)

[18] F. Ciurea and B. Funt, “A Large Image Database for Color Constancy Research,” *Proc. IS&T.SID Eleventh Color Imaging Conference, Society for Imaging Science and Technology*, pp. 160-163, (2003)

[19] K. Barnard, L. Martin, and B. Funt, “Colour by correlation in a three dimensional colour space,” *6th European Conference on Computer Vision,* Springer, pp. 375-389, (2000)

[20] V. Cardei, personal communication, (2005)

Method	SVR Dimension/ Norm Power	Median Angle	Max Angle	RMS Angle	Median Dist($\times 10^2$)	RMS Dist ($\times 10^2$)	Max Dist ($\times 10^2$)
TPS		0.64	14.43	2.10	0.53	1.55	10.42
SVR	2D	4.65	22.99	10.06	3.41	7.5	16.41
	3D	2.17	24.66	8.069	3.07	6.3	16.03
SoG	6	3.97	28.70	9.027	2.83	6.21	19.77
Max RGB		6.44	36.24	12.28	4.46	8.25	25.01
GW		7.04	37.31	13.58	5.68	11.12	35.38

Table 1 Comparison of TPS to 2D and 3D SVR performance, SoG, Max RGB, Grayworld performance. The results involve real-data training and testing on the 321 SONY images. Errors are based on leave-one-out cross-validation, and are reported in terms of both the RMS angular chromaticity and distance error measures.

	TPS	2D SVR	3D SVR	SoG (norm power = 6)	Max RGB	GW
TPS		+	+	+	+	+
2D SVR	-		-	=	+	+
3D SVR	-	+		+	+	+
SoG (norm power = 6)	-	=	-		+	+
Max RGB	-	-	-	-		-
GW	-	-	-	-	+	

Table 2 Comparison of the different algorithms via the Wilcoxon signed-rank test. A ‘+’ means the algorithm listed in the corresponding row is better than the one in corresponding column. A ‘-‘ indicates the opposite, and an ‘=’ indicates that the performance of the respective algorithms is statistically equivalent.

Method	Dimension/Norm Power	Median Angle	RMS Angle	Max Angle	Median Dist ($\times 10^2$)	Mean Dist ($\times 10^2$)	RMS Dist ($\times 10^2$)	Max Dist ($\times 10^2$)
TPS(rg)	2	2.26	3.86	22.23	1.72	2.22	2.92	18.29
SVR (no resampling)	2D	2.40	4.47	20.43	1.74	2.40	3.27	18.40
	3D	2.02	3.94	17.46	1.40	2.09	2.94	15.42
SVR(with resampling)	3D	2.07	3.91	10.57	1.55	2.03	2.72	6.42
C-by-C	2D	-	-	-	-	2.92	3.89	-
NN	2D	-	-	-	-	2.26	2.76	-
SoG	6	3.02	4.99	19.71	2.19	2.96	3.80	15.96
Max RGB		2.96	6.39	27.16	2.17	3.36	4.75	22.79
GW		4.34	6.65	31.44	3.17	4.12	5.26	29.99

Table 3 Comparison of Composition Solution and TPS to that of SVR, Color by Correlation, the Neural Network, SoG, Max RGB, Grayworld. The tests are based on leave-one-out cross validation on a database of 900 uncalibrated images. The entries for C-by-C and NN are from [8] (Table 7 page 2385).

	TPS	2D SVR	3D SVR	3D SVR (with resampling)	SoG (norm power = 6)	MAX RGB	GW
TPS		+	-	-	+	+	+
2D SVR	-		-	-	+	+	+
3D SVR	+	+		=	+	+	+
3D SVR (with resampling)	+	+	=		+	+	+
SoG (norm power = 6)	-	-	-	-		=	+
MAX RGB	-	-	-	-	=		+
GW	-	-	-	-	-	-	

Table 4 Evaluation of the performance results from Table 3 using the Wilcoxon signed-rank test. Labeling '+', <'-', <'=' as in the caption for Table 2.

Method	Training and Test Sets	Angular Degrees			Distance($\times 10^2$)		
		Median	RMS	Max	Median	RMS	Max
TPS	Train: Subset A Test: Subset B	4.52	7.02	34.81	3.37	5.19	25.78
3D SVR		4.53	6.76	24.55	4.11	5.03	18.62
SoG (norm = 6)		6.71	8.93	37.01	4.83	6.59	27.99
MAX RGB		10.33	12.81	27.42	6.99	9.14	21.72
GW		6.83	9.66	43.84	5.25	7.82	45.09
TPS	Train: Subset B Test: Subset A	4.58	6.83	27.62	3.31	4.99	29.37
3D SVR		5.33	7.32	24.80	3.91	5.29	16.68
SoG (norm = 6)		6.71	8.92	37.01	4.83	6.59	27.99
MAX RGB		9.23	11.32	26.76	6.76	8.39	21.55
GW		7.83	10.66	43.84	6.25	8.81	45.09

Table 5 Comparison of TPS error to 3D SVR, SoG, Max RGB, and Grayworld. Training is based on all the images in the given subset.

Method	TPS	3D SVR	SoG (norm power = 6)	MAX	GW
TPS		=	+	+	+
3D SVR	=		+	+	+
SoG (norm power = 6)	=	-		-	-
MAX	-	-	-		-
GW	-	-	-	+	

Table 6 Comparison of the algorithms based on the Wilcoxon signed-rank test on angular error. Training set is Subset A. Test set for all methods is Subset B. Labeling '+', '-', '=' as explained in the caption for Table 2.

Method	TPS	3D SVR	SoG (norm power = 6)	MAX	GW
TPS		=	+	+	+
3D SVR	=		+	+	+
SoG (norm power = 6)	-	-		+	+
MAX	-	-	-		=
GW	-	-	-	=	

Table 7 Comparison of the algorithms based on the Wilcoxon signed-rank test on angular error. Training set is Subset B. Test set for all methods is Subset A. Labeling '+', '-', '=' as explained in the caption for Table 2.

Method	Angular Degrees			Distance($\times 10^2$)		
	Median	RMS	Max	Median	RMS	Max
TPS	4.56	6.93	34.18	3.35	5.09	25.78
3D SVR	4.91	7.03	24.80	3.62	5.16	18.62
SoG	6.71	8.93	37.01	4.83	6.59	27.99
MAX RGB	9.65	12.13	27.42	6.86	8.80	21.72
GW	6.82	9.66	43.84	5.25	7.82	45.09

Table 8 Comparison of TPS error to 3D SVR, SoG, Max RGB, and Grayworld. The results involve real-data training and testing on disjoint sets of 7,661 images from the Ciurea data set.

Method	TPS	3D SVR	SoG (norm power = 6)	MAX	GW
TPS		=	+	+	+
3D SVR	=		+	+	+
SoG (norm power = 6)	-	-		-	=
MAX	-	-	+		+
GW	-	-	=	-	

Table 9 Comparison of the performance based on the Wilcoxon signed-rank test. Labeling '+', '-', '=' as explained in the caption for Table 2.