

文章编号: 0253-2239(2007)06-1135-4

# 基于多项式回归模型的扫描仪色度特征化\*

王 勇 徐海松

(浙江大学现代光学仪器国家重点实验室, 杭州 310027)

**摘要:** 为了实现扫描仪精确的色度特征化,详细研究了多种不同项数、不同变换类型多项式回归模型,多项式的项数分别为 3,5,9,11,18 和 20,变换类型包括扫描仪 RGB 空间到 CIEXYZ 或 CIELAB 空间的多种变换。以 ANSI IT8.7/2 标准色卡为色靶,对同一台扫描仪进行了色度特征化。实验结果表明,多项式项数小于 11 时,不同类型模型间的精度差异大,且线性化 RGB 能提高模型精度;项数大于 11 时,模型间精度差异缩小,而线性化 RGB 反而会降低模型精度。实验中,多项式项数为 20 时的 RGB 到 CIEXYZ 变换模型特征化精度最好,平均色差和最大色差分别为 0.832 和 2.988 CIELAB 色差单位。

**关键词:** 色度学;色度特征化;多项式回归;扫描仪

中图分类号: O432.3 文献标识码: A

## Colorimetric Characterization for Scanner Based on Polynomial Regression Models

Wang Yong Xu Haisong

(State Key Laboratory of Modern Optical Instrumentation, Zhejiang University, Hangzhou 310027)

**Abstract:** In order to realize accurate colorimetric characterization for a scanner, several polynomial regression models with different number of items and different conversion types were evaluated. The items of polynomial included 3, 5, 9, 11, 18 and 20, and the conversion types consisted of manifold transformation from scanner RGB space to standard CIEXYZ or CIELAB colorimetric space. An identical scanner was colorimetrically characterized with these polynomial models utilizing ANSI IT8.7/2 as color targets. Experimental results showed that the performance of these models were quite different and the linearization of scanner RGB signals improved the model accuracy when the items of polynomial were less than 11, while the differences disappeared and the linearization decreased the model accuracy along with the items more than 11. In this experiment, the polynomial conversion from raw scanner RGB to CIEXYZ with 20 items gave the best performance with the mean/maxmium color defferences of 0.832/2.988 CIELAB units.

**Key words:** colorimetry; colorimetric characterization; polynomial regression; scanner

## 1 引 言

扫描仪作为常用的图像获取设备,其颜色特征化的精度对于高质量的颜色再现至关重要。扫描仪的颜色特征化方法主要分为两类。第一类方法是色度特征化<sup>[1~7]</sup>,直接将 RGB(红绿蓝)值转换为 CIE 色度值。这种方法比较直接,就是寻求扫描仪 RGB 值与 CIE 色度值之间的映射关系,如多项式回归模

型和神经网络模型<sup>[8]</sup>等。第二类方法是光谱特征化<sup>[9~12]</sup>,由 RGB 值重构图像的光谱数据。多项式回归模型是扫描仪色度特征化的主要方法,但仍然存在以下问题:1)确定合适的多项式项数。一般,多项式项数增加,模型对于回归样本精度会提高,而对于未参与回归样本精度会下降。最优的项数应能使模型对于未参与回归样本仍保持足够的精度。2)是

\* 国家自然科学基金(60578011)资助课题。

作者简介:王 勇(1976—),男,安徽肥东人,博士研究生,主要从事数字图像设备颜色管理方面的研究。

E-mail: wangyong\_zju@163.com

导师简介:徐海松(1966—),男,江苏吴江人,教授,博士生导师,主要从事颜色科学、视觉与图像工程、光电技术和光学工程等方面的研究。E-mail: chsxu@zju.edu.cn

收稿日期:2006-08-08;收到修改稿日期:2006-11-16

否需要扫描仪信号线性化。Kang 和 Hardeberg<sup>[1,4]</sup>的研究中对扫描仪信号进行了线性化,而 Hong 和徐艳芳等<sup>[2,9]</sup>的研究中没有类似的处理。3) 选择合适的输出颜色空间。有研究<sup>[2,4]</sup>认为  $RGB \rightarrow L^* a^* b^*$  变换多项式模型在均匀颜色空间上进行优化,可以得到更高的特征化精度。但是,目前仍没有这方面的详细研究报道。本文研究和讨论了不同项数、不同变换类型的多项式回归模型来实现扫描仪的色度特征化。

## 2 扫描仪色度特征化理论

在进行扫描仪色度特征化时,首先扫描包含一系列色块的色靶(Color target)图像,如 ANSI IT8.7/2 标准色卡,获得各个色块的扫描仪 RGB 值;用测色仪器对色块进行测量得到每个色块的 CIE 色

度值。然后,建立色块的 RGB 值(或线性化 RGB 值)与 CIE 色度值的变换模型。最后,通过模型计算值与测试值间的色差<sup>[13,14]</sup>评价模型精度。

### 2.1 RGB 的线性化

扫描仪信号的线性化,也称作灰平衡校正,使处理后的扫描仪信号是亮度的线性函数。记灰平衡校正后扫描仪信号为  $(R_1, G_1, B_1)$ , 则

$$R_1 = g_R(R), G_1 = g_G(G), B_1 = g_B(B), \quad (1)$$

式中  $g_R, g_G, g_B$  分别是  $R, G, B$  通道的线性化函数。

### 2.2 多项式回归

记输入矩阵  $R$ , 转换矩阵  $M$  和输出矩阵  $H$ , 则多项式模型可表示为

$$H = RM, \quad (2)$$

式中输入矩阵  $R$  为  $N \times m$  矩阵,  $N$  为色块数,  $m$  对应多项式项数。如(3)式,  $\rho_m$  为输入矩阵  $R$  的行向量:

$$\begin{aligned} \rho_3 &= [R \ G \ B], \\ \rho_5 &= [1 \ RGB \ \rho_3], \\ \rho_9 &= [\rho_3 \ RG \ RB \ GB \ R^2 \ G^2 \ B^2], \\ \rho_{11} &= [1 \ RGB \ \rho_9], \\ \rho_{18} &= [\rho_9 \ RG^2 \ RB^2 \ GR^2 \ GB^2 \ BR^2 \ BG^2 \ R^3 \ G^3 \ B^3], \\ \rho_{20} &= [1 \ RGB \ \rho_{18}], \end{aligned} \quad (3)$$

$R, G, B$  可以是线性化前或后的值。输出矩阵  $H$  为  $N \times 3$  矩阵, 行向量可以由色块的 XYZ 或  $L^* a^* b^*$  值组成。若  $\rho = \rho_3$ , 则(2)式表示简单线性变换。以多项式扩展  $\rho$  能得到更精确的色度变换精度, 本文选择3项、5项、9项、11项、18项和20项等多项式模型进行实验和讨论, 其中5项、11项和20项分别是在3项、9项和18项基础上增加常数项和项扩展。对应于不同的  $\rho$ , 转换矩阵  $M$  分别是  $3 \times 3, 5 \times 3, 9 \times 3, 11 \times 3, 18 \times 3$  和  $20 \times 3$  矩阵。基于最小二乘法, 变换矩阵  $M$  按(4)式求解:

$$M = (R^T R)^{-1} R^T H, \quad (4)$$

式中  $R^T$  表示  $R$  的转置, 右上角的“ $-1$ ”表示矩阵求逆。

### 2.3 多项式回归特征化模型

根据(2)式中输入矩阵和输出矩阵行向量不同, 具体的多项式模型可以分为5种, 分别记作 TA、TB、TC、TD 和 TE, 如图1所示。TA 和 TB 为徐艳芳等采用的模型, TC、TD 和 TE 为 Hardeberg 采用的模型。  $RGB^{1/3}$  表示对线性化后的 RGB 求  $1/3$  次方。

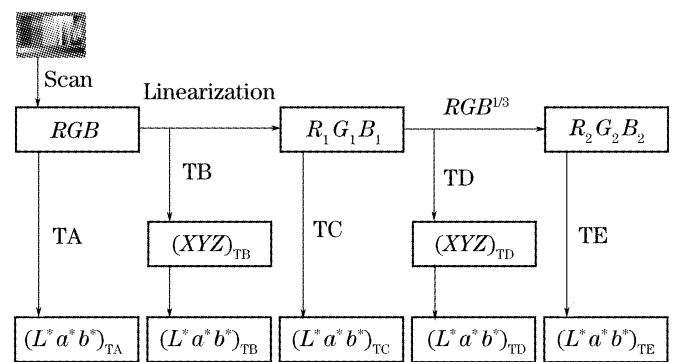


图1 多项式回归模型变换流程

Fig.1 Conversional workflow for polynomial regression model

## 3 实 验

### 3.1 实验设备和材料

扫描仪为 Epson Perfection1270, 扫描分辨率设置为 300 dpi(每英寸点数)。色靶为 Kodak 专业相纸制作的 IT8.7/2 标准色卡。所有色块的光谱反射比用 X-Rite SP64 测得, 并采用标准照明体 D65 和 CIE1931 标准色度观察者计算 XYZ 和  $L^* a^* b^*$  值。

### 3.2 实验结果分析

若线性化函数  $g_R, g_G, g_B$  为 2 阶多项式函数,  $a_1, a_2, a_3$  为多项式系数。则

$$g_C(C) = a_1C^2 + a_2C + a_3, \quad C = R, G, B \quad (5)$$

将 24 个灰色块的 RGB 值归一化到  $[0, 1]$  范围内, 与色块亮度值  $Y$  数据回归得到线性化函数系数以及拟合方均根差 (Root-mean-square error, RMSE) 列于表 1。由表中数据可见, 2 阶多项式能很好地描述亮度与 RGB 之间的非线性关系。

将 IT8.7/2 标准色卡中的色块按其在色卡上的排列次序均匀分成两部分, 其中 144 个色块作为

回归样本, 另外 144 个色块作为检验样本。扫描仪特征化精度由模型计算值与测试值之间的 CIELAB 色差来评价, 全部色块的平均/最大 CIELAB 色差 (Mean/ Max  $\Delta E_{ab}^*$ ) 列于表 2。

表 1 RGB 线性化 2 阶多项式系数和拟合方均根差

Table 1 Coefficients of two-order polynomial for RGB linearization and root-mean-square error

Channel	$a_1$	$a_2$	$a_3$	RMSE
R	0.6268	0.2532	0.02498	0.005306
G	0.6344	0.2830	0.02844	0.004574
B	0.5511	0.3184	0.02337	0.007422

表 2 多项式回归模型实验结果 (Mean/Max  $\Delta E_{ab}^*$ )

Table 2 Experimental results of polynomial regression models (Mean/Max  $\Delta E_{ab}^*$ )

Transformation matrix $M$	TA	TB	TC	TD	TE
$3 \times 3$	11.44/23.99	8.00/57.08	19.27/37.51	2.88/16.48	5.00/16.49
$5 \times 3$	3.78/18.41	9.11/72.60	8.80/30.75	2.90/16.73	2.80/15.70
$9 \times 3$	5.21/18.25	2.46/12.95	8.23/17.90	2.01/9.34	2.22/7.39
$11 \times 3$	1.64/8.13	1.81/8.83	3.56/16.42	1.98/8.49	1.70/5.74
$18 \times 3$	2.77/15.34	1.51/10.81	3.69/13.93	1.74/9.73	1.14/8.05
$20 \times 3$	1.25/7.58	1.04/4.61	2.00/15.18	1.64/6.61	1.09/6.52

采用  $3 \times 3$  线性变换矩阵, TA、TB、TC 和 TE 模型变换精度都很差, TD 模型平均色差为 2.88, 明显优于其他模型, 但最大色差为 16.48, 远远超过了可接受色差范围 ( $\Delta E_{ab}^* \leq 6$ )<sup>[4]</sup>。当变换矩阵为  $5 \times 3$  时, TA、TC 和 TE 模型的精度有明显的提高, 其中 TE 模型平均色差在所有模型中最小, 但最大色差仍不能接受。当变换矩阵为  $9 \times 3$  时, TA 模型的平均色差增加, 其它模型的平均色差持续减小, 但仍存在较大色差。变换矩阵扩展到  $11 \times 3$  时, 除 TC 模型外, 其他模型间差异已经不大, 其中 TE 模型的平均色差和最大色差最小, 分别为 1.70 和 5.74。继续增加变换矩阵的项数至  $18 \times 3$  时, TA 和 TC 模型的平均色差增加, TB、TD 和 TE 的平均色差仍继续减小; 除了 TC 模型的最大色差略有减小外, 其他模型的最大色差都有所增加。最后, 变换矩阵项数增加到  $20 \times 3$  时, 各模型的平均色差都降到 2 以下, 其中 TB 和 TE 的平均色差接近 1; TA 和 TC 模型的最大色差仍然偏大, TD、TE 模型的最大色差基本可以接受, 而 TB 的最大色差仅为 4.61 CIELAB 色差单位。

TC 模型在所有模型中精度最差, 原因可能是扫描仪的 RGB 经线性化后, 与 XYZ 之间存在近似的线性变换关系, 而  $XYZ \rightarrow L^* a^* b^*$  变换中存在近似指数形式的非线性关系, 这种非线性关系难以用

多项式回归模型进行描述。另外, 当多项式的项数增加到 11 以上后, 无论是变换到 CIELAB 还是 CIEXYZ 空间, RGB 线性化都只能使模型的精度下降。因此, 只有输出为 CIELAB 颜色空间, 并且多项式项数较少时, 才需要对扫描仪 RGB 信号线性化以提高模型的精度。

多项式中增加常数项和 RGB 项 (5 项、11 项和 20 项) 可以提高模型的精度, 尤其对于输出为 CIELAB 空间的 TA、TC 和 TE 模型, 效果明显。

一般来说, 多项式项数过高, 多项式回归模型的泛化能力会下降。表 3 列出了 TB 模型的部分实验结果。当多项式项数为 20 时, 检验样本的精度与回归样本比较, 平均色差增加 0.52, 最大色差增加 2.36, 但没有出现异常的局部极大值, 因此, 20 项 TB 模型能够满足泛化要求。

表 3 TB 多项式回归模型实验结果

Table 3 Experimental results of TB polynomial regression model

Transformation matrix $M$	Regression sample		Check sample	
	Mean	Max	Mean	Max
	$\Delta E_{ab}^*$	$\Delta E_{ab}^*$	$\Delta E_{ab}^*$	$\Delta E_{ab}^*$
$11 \times 3$	1.54	7.16	2.08	8.83
$18 \times 3$	1.11	6.79	1.91	10.81
$20 \times 3$	0.78	2.25	1.30	4.61

为了与其他研究者的研究结果作比较, 将

IT8.7/2标准色卡所有色块用于模型的回归和检验, 20项TB模型平均色差/最大色差为0.832/2.988 CIELAB色差单位,与Hardeberg认为极好(excellent)的实验结果[1.006/5.515(FUJI IT8.7/2色卡, AGFA Arcus II扫描仪)和0.918/4.666(AGFA IT8.7/2色卡, AGFA Arcus II扫描仪)]相比略有提高,也优于徐艳芳等用多项式回归和神经网络模型对多台扫描仪颜色特征化的最好结果[1.12/5.00和1.40/5.80(AGFA IT8.7/2色卡, Microtek ScanMaker V600扫描仪)],Shi等的多项式回归模型实验结果[1.71/10.24(Kodak IT8.7/2部分色卡, MUSTEK Paragon扫描仪)]以及Schettini等的神经网络模型实验结果[2.02/5.29(AGFA IT8.7/2色卡, Sharp JP600扫描仪)]。

## 4 结 论

通过对扫描仪色度特征化多项式回归模型的实验测试和详细分析比较可知,当多项式项数小于11时,变换方式对模型的精度影响较大,扫描仪信号线性化能够明显提高模型的精度;但随着项数继续增加,不同变换方式模型间的差别减小,扫描仪信号线性化反而会降低模型的精度。多项式中增加常量项和RGB项可以提高模型的精度,尤其对于输出为CIELAB空间的TA、TC和TE模型,效果明显。大部分情况下,输出为CIEXYZ空间时的多项式模型精度优于CIELAB空间。Hardeberg提出的 $R_2G_2B_2 \rightarrow L^* a^* b^*$ (TE模型)明显优于 $R_1G_1B_1 \rightarrow L^* a^* b^*$ (TC模型)或(TA模型),其缺点是使变换过程过于复杂。当多项式项数大于或等于9时,综合考虑精度和运算效率,TB模型是最佳选择。

## 参 考 文 献

1 Henry R. Kang. Color scanner calibration[J]. *J. Imaging*

- Science and Technology*, 1992, **36**(2):162~170
- 2 Xu Yanfang, Liu Wenyao, Zuo Kunlong *et al.*. Characterization of color scanners[J]. *Optics and Precision Engineering*, 2003, **12**(1): 15~20 (in Chinese)  
徐艳芳,刘文耀,左坤隆等. 彩色扫描仪的特征化[J]. *光学精密工程*, 2003, **12**(1): 15~20
- 3 R. Schettini, B. Barolo, E. Boldrin. Colorimetric calibration of color scanners by back-propagation[J]. *Pattern Recognition Letters*, 1995, **16**(10): 1051~1056
- 4 Jon Yngve Hardeberg. Acquisition and reproduction of colour images; colorimetric and multispectral approaches[D]. Paris: Ecole Nationale Supérieure des Télécommunications, 1999. 36~51
- 5 Miaohong Shi, Glenn Healey. Using reflectance models for color scanner calibration[J]. *J. Opt. Soc. Am. A*, 2002, **19**(4): 645~656
- 6 Guowei Hong, M. Ronnier Luo, Peter A. Rhodes. A study of digital camera colorimetric characterization based on polynomial modeling[J]. *Color Research and Application*, 2001, **26**(1): 76~84
- 7 Wang Yong, Xu Hai-song. Colorimetric characterization of liquid crystal display using an improved two-stage model[J]. *Chin. Opt. Lett.*, 2006, **4**(7): 432~434
- 8 Mu Baozhong, Yu Feihong. Color space converseon of digital photofinishing by neural network[J]. *Chin. Opt. Lett.*, 2005, **3**(9): 556~558
- 9 Roy S. Berns, M. J. Shyu. Colorimetric characterization of a desktop drum scanner using a spectral model[J]. *J. Electronic Imaging*, 1995, **4**(4): 360~372
- 10 Hui-liang Shen, John H. Xin. Colorimetric and spectral characterization of a color scanner using local statistics[J]. *J. Imaging Science and Technology*, 2004, **48**(4): 342~346
- 11 Wang Yong, Xu Hai-song. Spectral characterization of scanner based on PCA and BP ANN[J]. *Chin. Opt. Lett.*, 2005, **3**(12): 725~728
- 12 Wang Yong, Xu Hai-song. Spectral characterization of a flat panel color scanner using PCA method[C]. *Proc. SPIE*, 2006, **6033**: 144~148
- 13 Wang Zhehong, Xu Haisong. Lightness threshold evaluateon in visual color matching[J]. *Acta Optica Sinica*, 2006, **26**(8): 1274~1278 (in Chinese)  
汪哲弘,徐海松. 颜色视觉匹配中明度阈值的评价[J]. *光学学报*, 2006, **26**(8): 1274~1278
- 14 Cui Guihua, Li Wei, Fan Qiumei *et al.*. Study of the parametric effects in colour-difference evaluation using CRT display[J]. *Acta Optica Sinica*, 2001, **21**(4): 426~432 (in Chinese)  
崔桂华,李为,范秋梅等. 用阴极射线管显示器研究影响色差知觉的因素[J]. *光学学报*, 2001, **21**(4): 426~432