

基于多元非线性回归模型的彩色扫描仪特征化

李娟, 李斌, 张逸新

(江南大学, 无锡 214122)

摘要: 研究了基于多元非线性回归模型的彩色扫描仪特征化方法, 分析了多元回归的阶数以及输出色空间的选择对非线性多项式转换模型是否存在影响。研究表明: 训练组样本特征优化随着阶数增加而增加, 测试组样本在高阶数非线性回归下优化精度下降, 多项式方程呈现病态, 泛化能力下降; 回归模型采用 RGB 到 $L^*a^*b^*$ 的转换比采用 RGB 到 XYZ 的转换, 优化精度略高。结果说明, 利用多元非线性回归模型可满足扫描仪特征化的精度要求。

关键词: 扫描仪; 扫描仪特征化; 多元非线性回归模型

中图分类号: TS801.3; TS803 **文献标识码:** A **文章编号:** 1001-3563(2011)15-0110-03

Characterization of Color Scanners with Nonlinear Polynomial Regression Model

LI Juan, LI Bin, ZHANG Yi-xin

(Jiangnan University, Wuxi 214122, China)

Abstract: The influence of order number and selection of color spaces on polynomial regression model was analyzed. Color scanner characterization method based on nonlinear polynomial regression model was studied. Experimental results showed that as the order number features increase, the characterizing precision on training results of the nonlinear polynomial regression increases and the precision of experimental results decline in the high order. Polynomial equation present ill, and its generalization ability drops; characterizing precision of the polynomial regression of RGB to CIE transformation is higher than RGB to XYZ. The results showed that polynomial regression model can satisfy the accuracy requirement of scanner characterization.

Key words: scanner; characterization of scanner; polynomial regression

随着科技生活的进步, 彩色扫描仪已成为彩色复制领域中常用设备, 其颜色特性化的精度直接影响色彩再现的质量。色度特征化为彩色扫描仪特征化的主要方法^[1], 即将扫描仪获取色域的 RGB (红绿蓝) 值转换为与设备无关的 CIE 色度值^[2]。查找表法、神经网络方法以及多元回归为彩色扫描仪色度特征化主要方法。多元回归法转换效率高^[3], 实现过程相对查找表、神经网络方法较为容易, 但在多项式阶数、回归色空间的选择^[4]等相关问题上还需进一步探讨。

训练样本对扫描仪的优化精度随着回归模型的阶数增加而提高, 而当回归模型阶数的增加到一定值时, 对测试样本优化却相应下降, 也就是说导致回归方程成高危病态方程。能否在保证训练样本优化精度前提下, 使得未参与非线性回归的测试样本仍能保

持优化度。对扫描仪色彩特征优化问题, 也就是寻找适当的多项式阶数以保证扫描仪色域的整体优化度。H. L. Shen, Simone Bianco 等人对 RGB 值线性化, 着力于寻找创新的回归模型, 以达到更好的优化结果^[5-6]。王勇、徐海松等人研究用不同项数、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, 结果表明多项式项数高于 11 项, 回归模型泛化能力下降, 且优化精度不再有明显程度提高。对于扫描仪输出颜色空间问题的讨论, 相对于 CIEXYZ 色彩空间的不均匀性, CIE 为颜色均匀的色彩空间, 是否会提高扫描仪的特征化精度, 不同的理论模型研究结果也不尽一致。鉴于目前扫描仪特征化的研究现状, 笔者研究彩色扫描仪 RGB 空间与 CIEXYZ 和 CIE 空间之间的多元回归模型以及多项式阶数, 对彩色扫描

收稿日期: 2011-04-29

作者简介: 李娟(1987-), 女, 江苏宿迁人, 江南大学硕士生, 主攻印刷色彩管理。

仪多元非线性回归模型精度的影响。

1 彩色扫描仪回归模型

彩色扫描仪的 RGB 到 CIE 色彩空间转换的多项式回归法^[7-8]是基于源色空间色度值 RGB 和目标色空间色度值 CIE 之间能被一组同时成立的等式所联系的假设。在源色空间和目标色空间中选择颜色样本(色块),根据选定的多项式建模^[9],通过求解得到的多项式系数矩阵,直接把源色空间的色度值转换为目标色空间色度值。

设彩色扫描仪输入色彩空间为 RGB,色彩输出空间为 L^*, a^*, b^* 。设定色样的 L^*, a^*, b^* 色度值分别是 R, G, B 值的非线性多项式函数,则:

$$\begin{cases} L^* = \sum b_{xy} R^p G^q B^r \\ a^* = \sum b_{yj} R^p G^q B^r \\ b^* = \sum b_{zj} R^p G^q B^r \end{cases} \quad (1)$$

式中 p, q, r 取值 $0, 1, \dots, n; j=1, 2, 3, \dots$ 。矩阵形式表达为:

$$[L^* \quad a^* \quad b^*]^T = \mathbf{B}_{3 \times L} \times \boldsymbol{\rho}_i \quad (2)$$

$\mathbf{B}_{3 \times L}$ 为系数矩阵; $\boldsymbol{\rho}_i (i=1, 2, \dots, L)$ 为多项式矩阵; L 为多项式列项数。

$n=1, L=4$ 时:

$$\boldsymbol{\rho}_4 = [R \quad G \quad B \quad 1]^T \quad (3)$$

$n=2, L=10$ 时:

$$\boldsymbol{\rho}_{10} = [RB \quad RG \quad GB \quad R^2 \quad G^2 \quad B^2 \quad \rho_4]^T \quad (4)$$

$n=3, L=20$ 时:

$$\boldsymbol{\rho}_{20} = [RGB \quad R^3 \quad G^3 \quad B^3 \quad R^2G \quad G^2B \quad B^2R \quad G^2R \quad B^2G \quad R^2B \quad \rho_{10}]^T \quad (5)$$

$n=4, L=35$ 时:

$$\boldsymbol{\rho}_{35} = [R^4 \quad G^4 \quad B^4 \quad R^2GB \quad RG^2B \quad RGB^2 \quad R^3G \quad R^3B \quad G^3R \quad G^3B \quad B^3R \quad B^3G \quad R^2G^2 \quad R^2B^2 \quad G^2B^2 \quad \rho_{20}]^T \quad (6)$$

基于最小二乘法,系数矩阵 $\mathbf{B}_{3 \times L}$ 可按式(7)求解:

$$\mathbf{B}_{3 \times L} = [L^* \quad a^* \quad b^*]^T [(\boldsymbol{\rho}_i)^T \boldsymbol{\rho}_i]^{-1} (\boldsymbol{\rho}_i)^T \quad (7)$$

式中右上角的“ -1 ”表示矩阵求逆。选定 n, L 值后,将对应的 $\boldsymbol{\rho}_i$ 和彩色扫描仪输出变量 L^*, a^*, b^* 与输入变量 R, G, B 代入(2)式,由(7)求得系数矩阵 $\mathbf{B}_{3 \times L}$ 。同理,CIEXYZ 色彩空间的转换也是模拟上述过程进行拟合,并得到优化结果。

2 彩色扫描仪特性化多元回归法实验

2.1 设备和材料

实验选用扫描仪 Microtek ScanMaker 3750i,扫描分辨率设置为 300 dpi(每英寸点数)。Kodak 专业相纸印刷的 IT8.7/2 标准色卡作为标准反射色卡,选用 X-Rite eye-one 测得标准照明体 D50 下的 CIE 和 CIEXYZ 值。

2.2 实验结果分析

将 IT8.7/2 标准色卡中的 288 个色块按其色卡上的排列顺序均匀分为 2 组,其中 144 个色块作为训练样本,另外 144 个色块作为测试样本。通过训练样本得到系数建立回归模型后,通过 CIE1976 和 CIEXYZ 颜色空间的色差 ΔE_{ab}^* 来评价测试组色度值与优化值之间的色差大小。进而对模型的预测精度进行评价,并通过色差最大值、最小值、平均值来评价不同色彩空间和不同阶数的影响。

用训练样本对扫描仪的优化情况见表 1,也就是

表 1 训练组样本的色差

Tab.1 ΔE_{ab}^* of training group specimen

色差	阶数			
	2 阶	3 阶	4 阶	
RGB 与 $L^* a^* b^*$ 的回归	最大色差	8.6331	7.0167	16.6781
	最小色差	0.1430	0.1082	0.0779
	平均色差	2.2428	1.3306	1.7810
RGB 与 XYZ 的回归	最大色差	18.7785	10.5578	19.0933
	最小色差	0.8597	0.2128	0.2331
	平均色差	6.1006	4.4300	7.4521

色卡中选取的训练样本实际测试色度值与非线性回归模型计算所得色度值之间的最大、最小和平均色度值色差。

由表 1 可以看出,随着多元回归阶数的增大,色差均依次减小,4 阶回归模型预测效果最好。阶数越高,多元回归模型预测的效果越好。同时,可以看出 RGB 与 $L^* a^* b^*$ 之间多元回归模型优化精度总体上要比 RGB 与 XYZ 之间多元回归模型好。

为了证明回归模型适用性,选用测试样本来进一步验证。色卡中选取的测试样本实际测试色度值与回归模型预测色度值之间的最大、最小和平均色差见表 2。

表 2 测试组样本的色差

Tab.2 ΔE_{ab}^* of tested group specimen

色差		阶数		
		2 阶	3 阶	4 阶
RGB 与 $L^* a^* b^*$ 的回归	最大色差	11.222 3	5.990 1	4.119 8
	最小色差	0.206 7	0.456 7	0.077 8
	平均色差	2.052 4	1.192 0	0.838 2
RGB 与 XYZ 的回归	最大色差	27.891 6	14.408 0	8.013 8
	最小色差	1.125 1	0.370 8	0.157 3
	平均色差	4.078 6	2.231 0	1.256 9

从表 2 中可以看出,测试组数据与训练组数据不同的是 4 阶的回归模型最大色差、平均色差均比 3 阶回归模型大,表明 4 阶回归模型已经导致泛化能力下降。阶数高不能保证测试组优化精度,也就是说回归模型已呈病态方程,在实际生产将会产生严重的弊端。由理论研究的数据得出,扫描仪特征化应选择 3 阶多元非线性回归模型。

3 结论

利用多元非线性回归方法对扫描仪进行了特征化的研究,分析讨论了多元回归精度的阶数、色空间选择等影响因素,得出如下结果:训练组样本特征优化随着阶数增加而增加,测试组样本在高阶数下优化精度下降,多项式方程呈现病态,泛化能力下降,研究表明扫描仪特征化应选择 3 阶多元回归模型;CIE 的色彩空间均匀性使得 RGB 与 $L^* a^* b^*$ 之间多元回归

模型总体上要比 RGB 与 XYZ 之间多元回归模型好,优化精度更高。

参考文献:

- [1] 王勇,徐海松.基于多项式回归模型的扫描仪色度特征化[J].光学学报,2007,27(6):1135-1138.
- [2] 刘伟军.扫描仪的颜色校正实现[D].中山大学(软件工程),2007.
- [3] 谢政,李建平,汤泽滢.非线性最优化[M].北京:国防科技大学出版社,2003.
- [4] 胡成发.印刷色彩与色度学[M].北京:印刷工业出版社,1993.
- [5] SHEN H L, XIN J H. Colorimetric and Spectral Characterization of a Color Scanner Using Local Statistics[J]. Imaging Sci Technol, 2004, 48(4): 342-346.
- [6] BIANCO S, GASPARINI F, SCHETTINI R, et al. Polynomial Modeling and Optimization for Colorimetric Characterization of Scanners[J]. Journal of Electronic Imaging, 2008, 17(4): 1-13.
- [7] 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.
- [8] 李斌,张扬,张逸新.基于 SVM 的彩色扫描仪特征化[J].包装工程,2011,32(3):81-83.
- [9] 韩喜君,石俊生,黄小乔.彩色打印机特征化多元回归模型的研究[J].光学技术,2011,37(1):25-31.
- [10] BERNIS R S, SHYU M J. Colorimetric Characterization of a Desktop Drum Scanner Using a Spectral Model[J]. Electron Imaging, 1995, 4(4): 360-372.

(上接第 95 页)

4 结语

将 SIWAREX U 应用于瓦线单面机的压力控制,荷重信号的稳定性有了明显的提高,满足了纸板生产过程中对压力的要求,实现了生产过程中压力辊压力的自动控制,且压力设定值可以通过网络从瓦线中央控制室或车间办公室给定,实现生产管理智能化。

参考文献:

- [1] 杨瑞奉.瓦楞纸箱生产实用技术[M].北京:化学工业出

版社,2006.

- [2] 邢力.瓦楞纸箱粘合强度的解释[J].包装工程,1997,18(6):82-83.
- [3] 郭世刚. PLC 的人机接口与编程[J].微计算机信息,2006,22(19):42-45.
- [4] Simatic S7-300 可编程序控制器[M].西门子公司,2005.
- [5] 李转芳,康存锋. PLC 在自动称重和封装设备中的应用[J].包装工程,2010,31(15):85-88.
- [6] 西门子 SIWAREX U U 使用手册[K].西门子公司,2005.
- [7] STEP7 V5.0 User Manual[M]. SIEMENS AG, 1999.