

基于双层动态筛选训练样本的光谱重建算法

刘士伟^{1,2}, 刘真¹, 田全慧¹, 张建青¹

(1.上海理工大学, 上海 200093; 2.河南牧业经济学院, 郑州 450006)

摘要: **目的** 研究光谱重建过程中训练样本筛选方法对光谱重建精度的影响。**方法** 利用违逆的方法对测试样本 Munsell 样本和 ColorChecker SG 样本进行光谱重构, 训练样本分别选择未经筛选的 Munsell 样本集、经过动态聚类筛选的和经过文中提出的双重动态筛选的 Munsell 样本集, 然后比较 3 种样本筛选方法得到的光谱重构精度。**结果** 实验结果表明, 经过双层动态筛选的训练样本重构精度无论是均方根误差(RMSE)、拟合优度(GFC)还是不同光源下(A, D50 和 F2)的色差, 明显高于动态聚类分析的样本和未经筛选的样本。**结论** 提出了一种新的样本筛选方法, 该筛选方法效果良好, 具有一定的先进性。

关键词: 双层动态样本筛选; 光谱重建; 样本筛选; 重建精度

中图分类号: TS802.3; TS801.3 文献标识码: A 文章编号: 1001-3563(2017)03-0160-05

Spectral Reconstruction Algorithm Based on Dual Dynamic Training Samples Selection Method

LIU Shi-wei^{1,2}, LIU Zhen¹, TIAN Quan-hui¹, ZHANG Jian-qing¹

(1.University of Shanghai for Science and Technology, Shanghai 200093, China;

2.Henan University of Animal Husbandry and Economy, Zhengzhou 450006, China)

ABSTRACT: The work aims to study the influence of the training sample selection method in the process of spectral reconstruction on the spectral reconstruction accuracy. Munsell and ColorChecker SG (test samples) were reconstructed by using the method of pseudo inverse. Training samples were selected from unscreened Munsell sets and the Munsell sets selected through dynamic clustering and dual dynamic selection proposed in the paper. Then the spectral reconstruction accuracy was obtained by comparing three sample selection methods. The experimental results showed that the reconstruction accuracy of training samples subject to double dynamic selection was apparently higher than that of the samples analyzed by dynamic clustering and the unscreened samples, whether it was root-mean-square error (RMSE), goodness of fit (GFC) or color chromatic error under different light sources (A, D50, and F2). A new sample selection method is proposed. The selection method brings good effects and it is advanced to some extent.

KEY WORDS: dual dynamic training samples selection; spectral reconstruction; sample selection; reconstruction accuracy

在光谱重建算法中, 基于训练样本的光谱重建是目前光谱重建的主要研究方向, 主要算法有主成分分析法、维纳估计、R-矩阵法、非负矩阵分解、伪逆法等^[1-3]。基于训练样本的光谱重建方法一般需要通过训练样本建立光谱反射率与设备驱动值之间的转换矩阵, 然后再使用重建样本的设备驱动值与转换矩阵来重建光谱, 不需要成像系统各器件的光谱性能, 方便快捷^[4]。如何选择训练样本对光谱重建至关重要, 其样本选择情况直接影响光谱重建精度。选择训

练样本需要考虑两方面的问题: 首先是训练样本的代表性; 其次, 要考虑训练样本与重建样本的相关性。Munsell(孟塞尔)样本集包含了的色块范围较广, 因此, Munsell 样本集常常被用来作为训练样本。

目前的训练样本选择方法, 可以分为 2 种。第 1 种是固定子分区法, 其主要思想是把光谱空间按照一定的颜色特性进行划分。如 Garcia-Beltran 等^[5]通过色相对光谱空间进行酒类分区; Lee 等^[6]基于颜色特征划分颜色空间; Ayala 等^[7]和 Xiandou Zhang 等^[8]

收稿日期: 2016-08-13

基金项目: 上海理工大学科技发展项目 (16KJFZ017)

作者简介: 刘士伟 (1975—), 女, 上海理工大学博士生, 主要研究方向为光谱图像复制, 色彩管理等。

通讯作者: 刘真 (1953—), 女, 上海理工大学教授、博导, 主要研究方向为印刷光学工程、色彩再现理论与应用等。

根据 Munsell 色相对光谱空间进行划分。以上理论根据预先定义好的色相角进行空间划分,在临近分区线周围的颜色点特征较难界定,导致子分区划分的不够准确。另一种样本筛选方法是根据颜色点之间的距离或色差值的大小进行选择。Cheung 等提出的训练样本选择方法^[9-10],通过样本之间的欧几里德距离来实现,较多的考虑了样本之间的无关性而忽略了训练样本相关性的影响;如 Ansari 等^[11]使用色差对训练样本进行筛选;Agahian 等^[12]和 Babaei 等^[13]选择色差的大小做为加权函数,然后进行光谱重构。基于色差的训练样本筛选方法只能选择临近的样本集,这样会影响在高色差区域内训练样本的选择。常规的光谱重构算法在重建光谱反射率时,如果使用样本集全部的数据进行重建,仅得到一个转换矩阵,转换矩阵中用到训练样本重构样本的相关性较低,导致重构光谱精度

较差;如果光谱重建过程中,根据每个测试样本的光谱特征动态生成一个子样本集,使用与重构样本相似度的样本进行光谱重建,光谱误差将会显著降低,因此,提出了一种双层动态筛选训练样本的方法。

1 训练样本双重动态筛选

1) 理论依据。根据色相环原理可知^[14],类似色是在色相环上任意 60° 以内的颜色,通过绘制 Munsell 样本光谱反射率图发现,如果色相角小于 60° ,光谱反射率的特征形状相似度非常高。随机选择的蓝绿样本光谱见图 1b,图 1a, c 分别为与所选样本色相角相差 30° 的光谱,因此,如果选用色相角对训练样本进行筛选,将能获得与重构样本相关度较高的训练样本子集。

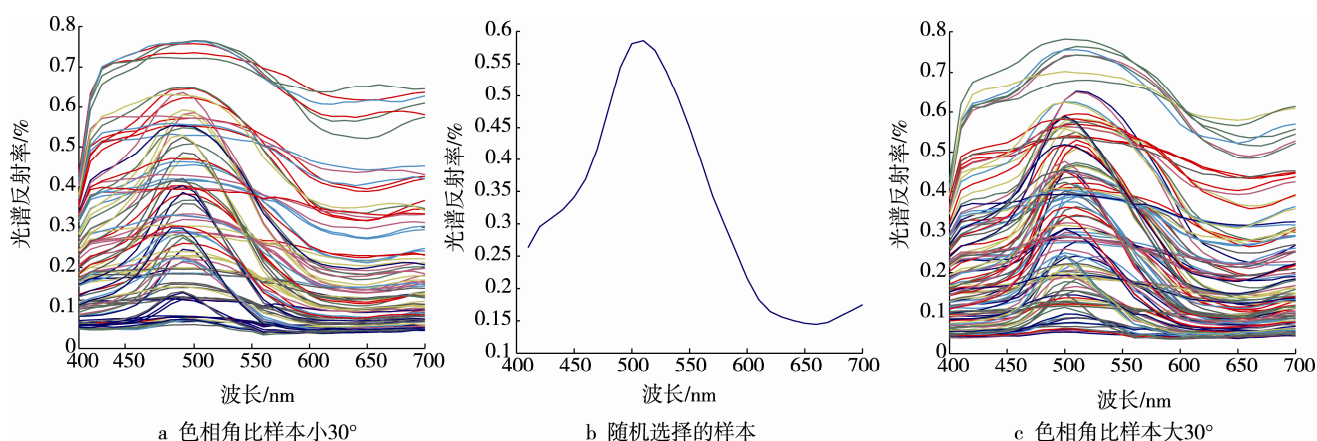


图 1 随机挑选的样本光谱及在 Munsell 样本集中筛选的样本
Fig.1 The randomly selected sample spectrum and the samples chosen from the Munsell set

通过色相环原理可知,以色相角进行空间划分得到的子空间,会包括该角度区域内所有的彩色色块和非彩色色块。如果能够通过再次动态筛选,根据重构样本将彩色色块和非彩色色块分开,必然会提高训练样本与重构样本的相似度,提高重构精度。光谱的相关系数是评

价 2 条光谱相似程度的性能指标,比利用光谱角评价相似程度精确度更高^[15],因此选择光谱的相关系数作为再次进行样本筛选的依据。例如重构样本是随机选择的蓝绿色块,将非彩色色块从动态子空间剔除,最终筛选出来的样本与重构样本的光谱相似度更高,见图 2。

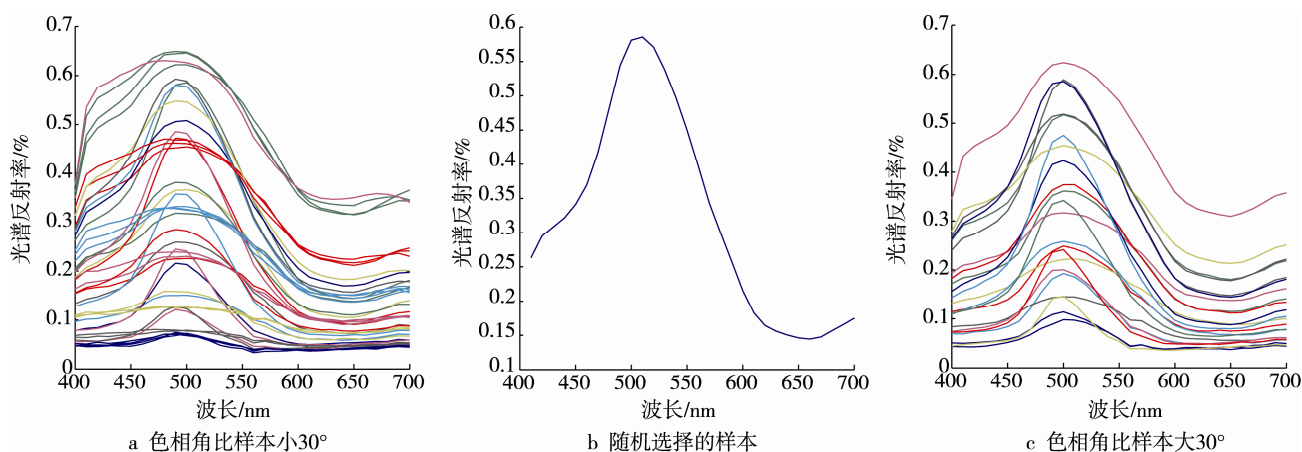


图 2 随机挑选的样本光谱及在 Munsell 样本集中经过 2 次筛选的样本
Fig.2 The randomly selected sample spectrum and the samples twice chosen from the Munsell set

2) 算法实现。选择 Munsell 样本作为训练样本集, 使用 Munsell 样本和 ColorChecker SG 作为重构样本, 以 10 nm 为间隔在 400 ~ 700 nm 范围内对所有样本进行均匀采样。

首先, 通过式(1)计算训练样本和测试样本中每个样本的 CIE XYZ 值(色块三刺激值), 然后根据 CIE XYZ 值再计算 CIELab, 再计算得到 CIE $L^*C^*H^*$ (颜色空间, L 表示明度, C 表示饱和度, H 表示色调角)。

$$\begin{aligned} X &= k \int r(\lambda) I(\lambda) \bar{x}(\lambda) d\lambda \\ Y &= k \int r(\lambda) I(\lambda) \bar{y}(\lambda) d\lambda \\ Z &= k \int r(\lambda) I(\lambda) \bar{z}(\lambda) d\lambda \end{aligned} \quad (1)$$

式中: k 为修正系数, $k=100/\int I(\lambda) \bar{y}(\lambda) d\lambda$; $r(\lambda)$ 代表光谱反射率; $I(\lambda)$ 为光源光谱相对功率分布, 在计算色度的时候统一选用 CIE D65 光源; $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$ CIE 颜色匹配函数。式(1)也可以写成矩阵形式:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = A^T r \quad (2)$$

其次, 以重构样本的色相角为中心, 在训练样本的 CIE $L^*C^*H^*$ 色空间, 以角度 θ 为间隔, 获取动态子空间。利用动态子中间的训练样本计算出转换矩阵 $(A^T)^+$, 然后通过公式(3)计算得到重构样本的初始光谱反射率。

$$R = (A^T)^+ \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (3)$$

最后, 将重构样本的初始光谱反射率与初次筛选得到的动态子分区的训练样本进行一一对比, 以光谱相关系数作为筛选指标, 在动态子分区中再次筛选出与重构样本光谱特征相似的训练样本。相关系数计算见式(4)。

$$P(x, y) = \frac{(x - \bar{x})^T (y - \bar{y})}{\left[(x - \bar{x})^T (x - \bar{x}) \quad (y - \bar{y})^T (y - \bar{y}) \right]^{\frac{1}{2}}} \quad (4)$$

这里 x, y 分别表示训练样本与重构样本光谱向量, $\bar{x}, \bar{y}$ 为 x, y 的均值。

2 实验与结果

为了验证双重动态筛选算法的精度及准确性, 选择伪逆法^[3]进行光谱重构; 训练样本分别选择没有筛选的 Munsell 样本集、动态聚类筛选的 Munsell 样本集、文中提出的双重动态筛选的 Munsell 样本集。为了定量分析不同算法的结果, 采用均方根误差(RMSE)、拟合优度(GFC)和不同光源下(A, D50 和 F2)的色差来评价光谱重构精度。色差是衡量光谱重构的一个重要指标, 计算色差需要在一定的光源和确定的标准观察者条件下计算。

表 1 基于伪逆算法的重构结果
Tab.1 The value of spectral reconstruction using the method of pseudo inverse

		Munsell 样本			ColorChecker SG		
		没有筛选	动态聚类	双层动态	没有筛选	动态聚类	双层动态
光谱 误差	平均RMSE	0.0233	0.0163	0.0075	0.0474	0.0370	0.0260
	最大RMSE	0.1759	0.1369	0.0637	0.1754	0.1382	0.1340
	最小GFC	0.7655	0.8305	0.9571	0.9384	0.9504	0.9032
	平均GFC	0.9936	0.9967	0.9992	0.9846	0.9893	0.9945
	最大GFC	0.9999	1.0000	1.0000	0.9994	0.9997	0.9999
在光源 A 下的色差	平均 ΔE_{ab}	1.5475	0.8532	0.6611	2.3503	1.6633	1.1520
	Var.	3.0017	1.0473	0.6216	9.6388	5.1304	2.5957
	最大 ΔE_{ab}	11.5247	7.0838	6.0788	13.3779	9.7395	7.8319
	$\Delta E_{ab} > 3$	12.9236	5.3586	2.9157	26.4386	20.0000	11.4286
在光源D50下的色差	平均 ΔE_{ab}	0.4707	0.2601	0.2077	0.7531	0.5305	0.3642
	Var.	0.2448	0.0911	0.0610	0.9508	0.4704	0.2362
	最大 ΔE_{ab}	3.1911	1.9429	1.9249	4.6767	2.8278	2.5186
	$\Delta E_{ab} > 3$	0.1576	0.0000	0.0000	5.0000	0.0000	0.0000
在光源F12下的色差	平均 ΔE_{ab}	1.2280	0.7102	0.5419	1.8756	1.4400	0.9348
	Var.	1.7630	0.5697	0.3105	7.6920	4.2713	1.4163
	最大 ΔE_{ab}	11.9383	7.0017	4.9132	14.3177	10.8345	5.9046
	$\Delta E_{ab} > 3$	7.8802	2.0489	0.7092	20.0000	17.8571	9.2857

注: Var. 为 CIELAB 色差的方差; $\Delta E_{ab} > 3$ 表示测试样本色差超过 3 的百分比。

3 结果分析

从表 1 中可看出经过双层动态样本筛选的重建精度最高,明显高于单层动态聚类 and 没有经过样本筛选的精度。为了清楚地表现双层动态样本筛选的筛选原理,随机挑选了一个重构样本并利用 2 种筛选方法从 Munsell 测试样本中筛选出 20 条光谱,见图 3。双层动态样本筛选出来的光谱样本与测试样本光谱形状相似,而单层聚类动态样本筛选出来的样本与重构样本光谱接近。

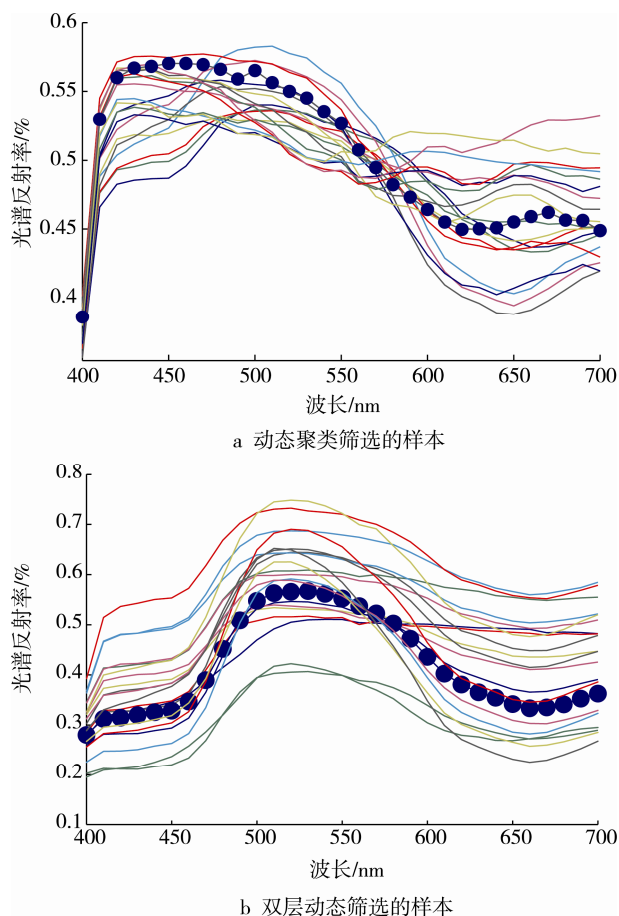


图 3 测试样本及筛选的 20 条训练样本光谱
Fig.3 Test samples and 20 selected training spectrum

在使用 Munsell 样本集作为重构样本时,比使用 ColorChecker SG 样本作为重构样本重建精度要高。这主要是因为不同样本集中的色块的特性不同,适合于不同的应用场景。ColorChecker SG 作为重构样本在 Munsell 样本集能够筛选出来的样本集光谱的相似程度不够高。在光源 A 和 F12 观测条件下的 CIELAB 色差值明显高于在光源 D50 观测条件下的色差值。分析其原因,发现 D50 光源与文中在计算三刺激值时选择的参考光源 D65 光源光谱曲线相似度较高,而 A 光源和 F12 光源与 D65 光源光谱曲线有较大差别,见图 4,因此,与参考光源光谱曲线相似度越高,得到的色差值越小。实验结果跟 Ayala 等和

Zhang Xiandou 等得到的结论一致;使用与参考光源相似度较低的观测光源如光源 A 和 F12,则能够更加准确地反映色度重建精度。

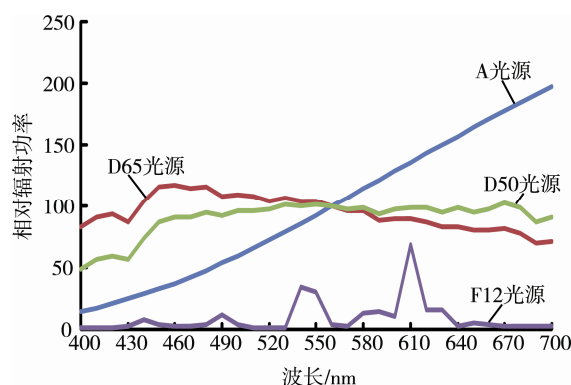


图 4 光源的光谱功率分布曲线
Fig.4 The spectral distribution curve of the light source

4 结语

文中通过双层动态样本筛选方法通过伪逆法进行光谱重构,在光谱重建过程中,没有仅仅使用一个转换矩阵,而是根据每一个测试样本的光谱特征动态筛选训练样本,生成自适应转换矩阵,并且最后重构的光谱误差和色差重建精度与没有经过筛选的和经过动态聚类筛选的样本重建进行了比较。结果表明,双层动态样本筛选方法是一种较为合理的样本筛选方法。

参考文献:

- [1] 刘振, 万晓霞, 黄新国, 等. 基于宽带多通道的光谱反射率重建方法研究[J]. 光谱学与光谱分析, 2013, 33(4): 1076—1081.
LIU Zhen, WAN Xiao-xia, HUANG Xin-guo, et al. The Study on Spectral Reflectance Reconstruction Based on Wideband Multi-spectral Acquisition System[J]. Spectroscopy and Spectral Analysis, 2013, 33(4): 1076—1081.
- [2] 于海琦, 刘真, 张雷洪, 等. 样本特征对光谱图像重构影响的研究[J]. 包装工程, 2014, 35(13): 144—149.
YU Hai-qi, LIU Zhen, ZHANG Lei-hong, et al. Effects of Sample Characteristics on Spectral Image Reconstruction[J]. Packaging Engineering, 2014, 35(13): 144—149.
- [3] 汪军, 强俊. 面向应用的数字图像处理课程教学模式的研究[J]. 重庆工商大学学报(自然科学版), 2015, 32(5): 94—98.
WANG Jun, QIANG Jun. Research on Application oriented Teaching Model of Digital Image Processing Course[J]. Journal of Chongqing Technology and Business University(Natural Science Edition), 2015, 32(5): 94—98.

- [4] 王晋. 基于数字成像设备的光谱测量精度评价[J]. 包装工程, 2014, 35(13): 133—138.
WANG Jin. Accuracy Evaluation of Spectral Measurement Based on Digital Imaging Equipment[J]. Packaging Engineering, 2014, 35(13): 133—138.
- [5] GARCÍA-BELTRÁN A, NIEVES J L, HERNÁNDEZ-ANDRÉS J, et al. Linear Bases for Spectral Reflectance Functions of Acrylic Paints[J]. Color Research and Application, 1998, 23(1): 39—45.
- [6] LEE C H, PARK K H, HA Y H, et al. Surface Reflectance Estimation Using the Principal Components of Similar Colors[J]. Journal of Imaging Science and Technology, 2007, 51(2): 166—174.
- [7] AYALA F, ECHÁVARRI J F, RENET P, et al. Use of Three Tristimulus Values from Surface Reflectance Spectra to Calculate the Principal Components for Reconstructing These Spectra by Using Only Three Eigenvectors[J]. JOSA A, 2006, 23(8): 2020—2026.
- [8] ZHANG X, XU H. Reconstructing Spectral Reflectance by Dividing Spectral Space and Extending the Principal Components in Principal Component Analysis[J]. JOSA A, 2008, 25(2): 371—378.
- [9] IBRAHIM A, TOMINAGA S, HORIUCHI T. Invariant Representation for Spectral Reflectance Images and Its Application[J]. EURASIP Journal on Image and Video Processing, 2011, 2011(1): 1—12.
- [10] CHEUNG V, WESTLAND S. Methods for Optimal Color Selection[J]. Journal of Imaging Science & Technology, 2006, 50(5): 481—488.
- [11] ANSARI K, AMIRSHAHI S H, MORADIAN S. Recovery of Reflectance Spectra from CIE Tristimulus Values Using a Progressive Database Selection Technique[J]. Coloration Technology, 2006, 122(3): 128—134.
- [12] AGAHIAN F, AMIRSHAHI S A, AMIRSHAHI S H. Reconstruction of Reflectance Spectra Using Weighted Principal Component Analysis[J]. Color Research and Application, 2010, 33(5): 360—371.
- [13] BABAEI V, AMIRSHAHI S H, AGAHIAN F. Using Weighted Pseudo-inverse Method for Reconstruction of Reflectance Spectra and Analyzing the Dataset in Terms of Normality[J]. Color Research and Application, 2011, 36(4): 295—305.
- [14] 胡威捷, 汤顺青, 朱正芳. 现代颜色技术原理及应用[M]. 北京: 北京理工大学出版社, 2007.
HU Wei-jie, TANG Shun-qing, ZHU Zheng-fang. Modern Color Technology Principle and Application [M]. Beijing: Beijing Science and Technology University Press, 2007.
- [15] 王立国, 赵春晖. 高光谱图像处理技术[M]. 北京: 国防工业出版社, 2013.
WANG Li-guo, Zhao Chun-hui. Processing Techniques of Hyperspectral Imagery[M]. Beijing: National Defence Industry Press, 2013.