

基于样本空间学习算法的 彩色匹配方法研究¹⁾

苏晓红 王亚东 郭茂祖

(哈尔滨工业大学计算机科学与工程系 哈尔滨 150001)

(E-mail: ydwang@mlg.hit.edu.cn)

马培军

(日本 Canotec 公司)

摘要 提出基于样本空间学习算法的 RGB(Red, Green, Blue)→CMYK(Cyan, Magenta, Yellow, black)颜色空间变换方法,仅需测试400组纯色色块,即可在不同的打印(抖动)模式、墨水和纸张条件下,对RGB→CMYK转换规律进行有效的学习。该算法学习能力强、打印速度快,通过在模型中增加高次项,提高了彩色匹配的精度,进一步解决了彩色匹配的自动化和通用性问题。同时,借助于CIELab均匀颜色空间对图像进行动态范围调整,还起到了减小色偏、突出图像细节、增强图像对比度的效果。

关键词 彩色匹配, 样本空间, 机器学习, 多项式回归, 色差。

RESEARCH OF COLOR MATCHING METHOD BASED ON LEARNING ALGORITHM IN VERSION SPACE

SU Xiao-Hong WANG Ya-Dong GUO Mao-Zu

(Department of Computer Science and Engineering, Harbin Institute of Technology, Harbin 150001)

(E-mail: ydwang@mlg.hit.edu.cn)

MA Pei-Jun

(Canotec Corporation Incorporated, Japan)

Abstract A color space transformation method from RGB to CMYK based on knowledge discovery is presented in this paper. The RGB→CMYK for different color ink-jet printers and paper can be found out after only four hundred pure color blocks are tested. This method has a good learning ability and fast printing speed. By adding high-order item to the model, the precision of color matching is improved, and at the same time the problem of automatic color matching and generalization is solved. With the aid of CIELAB uniform color space, the RGB→CMYK laws are well adjusted to stress the image details and enhance image

1)国家自然科学基金(69975005)及国际合作项目彩色匹配系统资金资助项目。

contrast.

Key words Color matching, version space, machine learning, multiple polynomial regression, color difference.

1 引言

彩色匹配技术是当前彩色印刷领域的一项重要的世界性技术难题,其实质就是要在不同设备间保持色彩的一致性^[1],这在办公自动化系统(尤其是在网络办公系统)中有着极其重要的应用。

传统的彩色匹配技术有 Neugebauer 方程^[2]、一维非线性函数、线性矩阵转换^[3]、多维表转换^[4]等方法,这些方法在实际应用中均需根据色彩专家的经验知识针对不同的输出设备、墨水和纸张进行调整,确定新的匹配参数,工作量大、不能自动化、缺乏通用性、难以实用和推广。文献[5]提出基于机器学习的彩色匹配方法,给出一种解决彩色匹配自动化和通用性问题的有效途径;但在实际应用于 RGB 到 CMYK 颜色空间转换时,由于选取的转换模型简单,且存在训练数据多、工作量大、偏色、图像细节分辨率低等不足。本文基于样本空间学习算法,较文献[5]做了如下重要改进。

1) 在 RGB→CMYK 规律的学习中,转换模型的选取及模型中系数的学习算法是问题的关键,它直接影响着打印图像的质量。本文采用多项式回归技术计算模型中的系数,这种方法考虑了模型的非线性,可以通过在模型中增加高次项提高模型的精度,而系数的计算方法不变,具有一定的通用性,同时弥补了文献[5]中提出的线性模型精度差的缺点。

2) 在从 RGB→CMYK 颜色空间转换时,文献[5]中提出的方法需印制 1 000 个标准常用色块(包括纯色色块和混色色块),然后用测光仪 ColorTron 分别测出其反射光(R,G,B)值形成训练数据集,工作量很大。本文提出的改进方法,仅需测试 400 个标准纯色色块就可对 RGB→CMYK 规律进行有效的学习,降低了彩色匹配算法的空间和时间复杂性,而且取得了更好的彩色匹配效果。

3) 借助于 CIELab 均匀颜色空间,通过对纯色数据的学习得到人眼对 C,M,Y,K 纯色色块分辨力的饱和值,然后将图像的色彩和亮度变化的动态范围线性压缩到学习出的饱和值范围内,从而达到减小色偏、突出图像的色彩和亮度细节、增强图像对比度的效果。

2 RGB→CMYK 颜色空间转换的难点问题

从 RGB 到 CMYK 的颜色空间转换,其实质就是要解决一种复杂的多维对多维的数据空间对应问题,即 $(C, M, Y, K) = F(R, G, B)$ 。这里 RGB 颜色空间代表计算机屏幕等发光体的颜色空间,是加色空间,当分别用 8bit 二进制数表示 R,G,B 值时,其数据组合有 $2^8 \times 2^8 \times 2^8 = 2^{24}$ 种;而 CMYK 颜色空间代表打印机输出墨水形成的反射光的颜色空间,是减色空间,CMYK 的值一般也用 8bit 二进制数表示,参照 PhotoShop 软件设定为 0~100,则其数据组合有 $100 \times 100 \times 100 \times 100 = 10^8$ 种。显然,RGB 和 CMYK 两种颜色空间的大小是不一样的,其表示如图 1 所示。

本文所要解决的问题就是将 RGB 空间中的某一点 (R_m, G_m, B_m) 对应到 CMYK 空间的某一点 (C_n, M_n, Y_n, K_n) 上,使得计算机屏幕上的发光点 (R_m, G_m, B_m) 所表达的颜色与打印机在白纸上打印墨水 (C_n, M_n, Y_n, K_n) 后形成的反光点 (R'_m, G'_m, B'_m) 所表达的颜色是一致的。显然,这种色彩匹配是相当困难的,主要表现在以下几个方面。

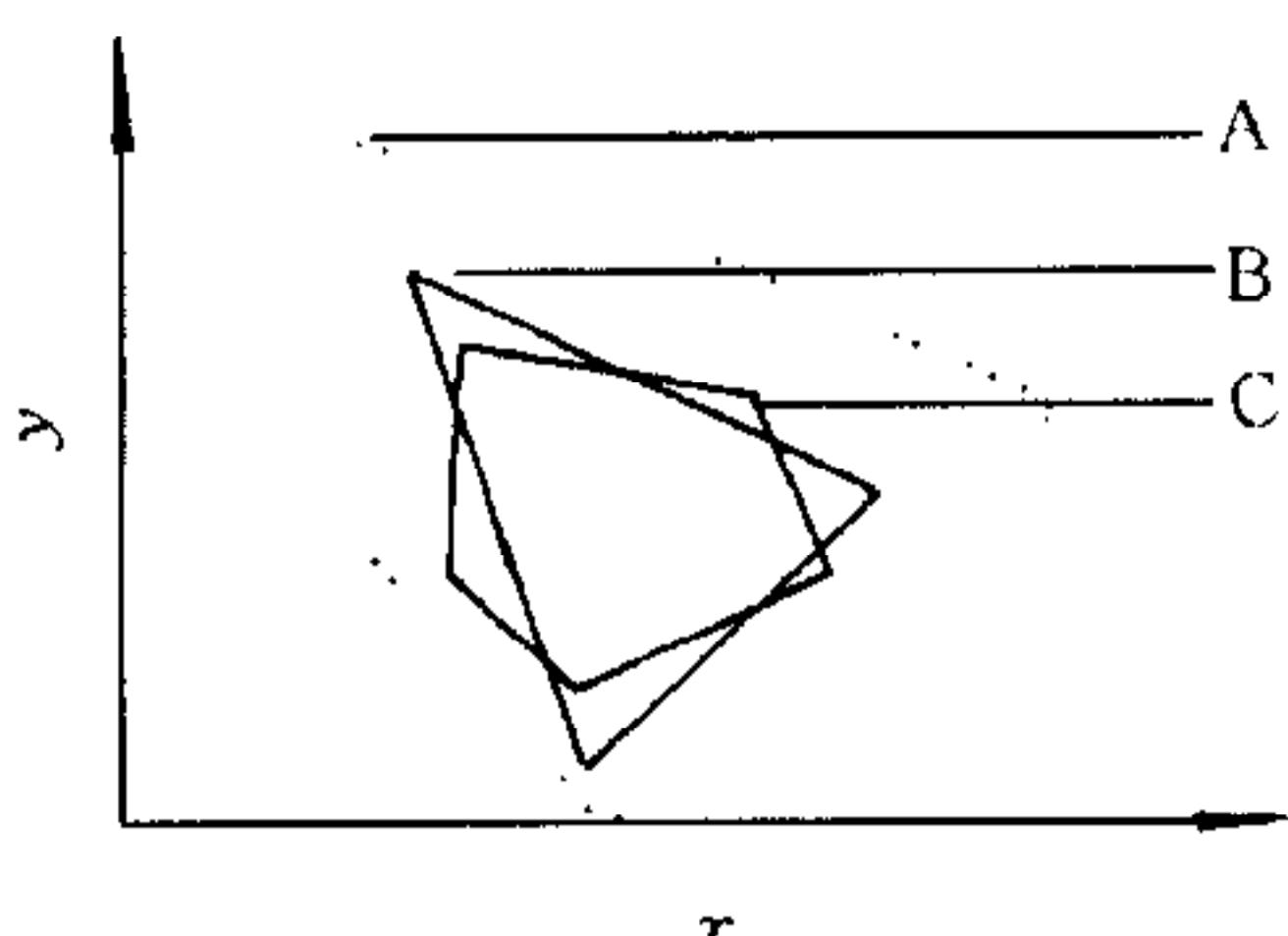


图1 Lab 颜色模式的颜色光谱

A—Lab 模式 B—RGB 模式 C—CMYK 模式
选取困难,且实时处理时大量的查表和插值运算又会导致打印速度过慢,从而无法实用。

2)RGB 和 CMYK 空间都具有非线性特点,在 RGB 或 CMYK 空间中距离相近的两点在实际的颜色表达上可能有较大差别,而距离较远的两点,却可能具有相近的颜色,因此,RGB 和 CMYK 空间应该存在一种非线性关系,而实际中是很难找到一个合适的非线性函数的,通常要根据色彩专家的经验反复试验获得,实现比较困难,也不具通用性。

3)对于加色法混色的色匹配可通过简单的数学式演算求得。而对于减色法混色的色匹配则较为复杂。因表达式中含有比朗伯比尔法则^[6]还要复杂的高次非线性项,故不能简单地求出其反函数,需采用牛顿迭代法逐次逼近求解,不仅误差大,而且实现速度也很慢。

3 RGB→CMY 转换的学习算法

3.1 学习数据的获取

机器学习首先要提供学习用的数据。先考虑如何分辨颜色,当从计算机屏幕上接收颜色信号时,实际上是发光体 (R, G, B) 的值直接进入人眼,而当把墨水打印在纸张上形成颜色时,是纸张上的墨水 (C, M, Y, K) 对一束照射在纸上的白光 $R=G=B=255$ 进行不同程度上的吸收之后,反射到人眼中形成一组新的 (R', G', B') 值,那么对应于屏幕上的一点 (R, G, B) ,如果能使 (R, G, B) 与 (R', G', B') 基本一致,则彩色匹配的目的也就达到了,这里要满足的条件是:计算 (R, G, B) 和 (R', G', B') 两种颜色在 CIELab 均匀颜色空间中的色差 ΔE_{ab} , ΔE_{ab} 应满足小于一个色彩分辨率误差 e 。

基于上述思想,提出一种分治策略来解决彩色匹配问题,通过对样本空间数据的学习,分别确定 C, M, Y, K 与 R, G, B 之间的单函数关系: $C=f_c(R, G, B)$, $M=f_m(R, G, B)$, $Y=f_y(R, G, B)$, $K=f_k(R, G, B)$, 然后在此基础上综合形成一种多维对多维的数据关系。

现给定一种彩色打印机,用选定的墨水和纸张打印一组 C, M, Y, K 纯色色块,其中 $C: 1 \sim 100 (M=Y=K=0)$; $M: 1 \sim 100 (C=Y=K=0)$; $Y: 1 \sim 100 (C=M=K=0)$; $K: 1 \sim 100 (C=M=Y=0)$, 然后再用 ColorTron 分别测定其各自的反射光 R, G, B 值,形成 400 组独立的测试数据(样本),即 $(C_i: R_i, G_i, B_i)$, $(M_i: R_i, G_i, B_i)$, $(Y_i: R_i, G_i, B_i)$, $(K_i: R_i, G_i,$

1)RGB 和 CMYK 两个颜色空间的大小不一致,而且不是完全包含的关系,是一种交叉关系,因而 RGB 空间有一些颜色肯定是无法在 CMYK 空间中正确表达的,只能给出近似的对应颜色。同时因两个空间的数据组合数非常庞大,所以要建立由 $RGB \rightarrow CMYK$ 的转换表必须印制和测量 2^{24} 种不同的色块,这显然是不现实的。若将色块数限制在一个实用的范围内,就必须用复

杂的插值法来计算表中没有的数据,不仅变换表的数据选取困难,且实时处理时大量的查表和插值运算又会导致打印速度过慢,从而无法实用。

2)RGB 和 CMYK 空间都具有非线性特点,在 RGB 或 CMYK 空间中距离相近的两点在实际的颜色表达上可能有较大差别,而距离较远的两点,却可能具有相近的颜色,因此,RGB 和 CMYK 空间应该存在一种非线性关系,而实际中是很难找到一个合适的非线性函数的,通常要根据色彩专家的经验反复试验获得,实现比较困难,也不具通用性。

3)对于加色法混色的色匹配可通过简单的数学式演算求得。而对于减色法混色的色匹配则较为复杂。因表达式中含有比朗伯比尔法则^[6]还要复杂的高次非线性项,故不能简单地求出其反函数,需采用牛顿迭代法逐次逼近求解,不仅误差大,而且实现速度也很慢。

B_i), 其中 $i=1,2,\dots,100$. 至此, 便得到了机器学习所需的样本数据, 改变打印模式、墨水和纸张等条件时, 可重新获取上述数据, 将这些数据存入硬盘, 供学习程序使用. 通过学习程序对特定的打印模式、墨水和纸张条件下的测试数据进行学习, 发现 $\text{RGB} \rightarrow \text{CMYK}$ 转换规律, 形成知识库, 供打印驱动程序执行 $\text{RGB} \rightarrow \text{CMYK}$ 转换, 并实施彩色图像打印时使用. 这种方法的最突出的优点就是实现了彩色匹配的自动化, 学习程序和打印程序具有通用性.

3.2 $\text{RGB} \rightarrow \text{CMY}$ 转换规律的研究

RGB 与 CMY 之间在理论上存在如下关系:

$$C = W-R, M = W-G, Y = W-B.$$

从色彩吸收的角度来理解, 即青色墨水 C 从白光 W 中吸收红光 R , 品红墨水 M 从白光中吸收绿光 G , 黄色墨水 Y 从白光中吸收蓝光 B , 然而由于墨水的纯度、散射性、纸张的质量以及抖动方式的不同等因素的影响, 使得这种纯理论上的对应关系在实际应用时存在很大的误差, 不易调整. 实际上, 输出的 C 值虽然主要依赖于输入的 R 值, 但与输入的 G, B 值也有关系, 同样 M 值和 Y 值也与输入的 R, G, B 都有关系. 根据这一思想, 对前文测得的 C, M, Y 300 组测试数据, 建立如下颜色空间转换模型:

$$\begin{cases} C = \alpha_{11}(R_w - R) + \alpha_{12}(G_w - G) + \alpha_{13}(B_w - B) + \epsilon_1, \\ M = \alpha_{21}(R_w - R) + \alpha_{22}(G_w - G) + \alpha_{23}(B_w - B) + \epsilon_2, \\ Y = \alpha_{31}(R_w - R) + \alpha_{32}(G_w - G) + \alpha_{33}(B_w - B) + \epsilon_3. \end{cases} \quad (1)$$

其中 (R_w, G_w, B_w) 是用 ColorTron 测得的一组白纸对白光的反射光值的平均值, 对于白纸显然有 $C=M=Y=0$.

对 C, M, Y 300 组数据, 有

$$\begin{cases} C_i = \alpha_{11}(R_w - R_i) + \alpha_{12}(G_w - G_i) + \alpha_{13}(B_w - B_i) + \epsilon_{1i}, \\ M_i = \alpha_{21}(R_w - R_i) + \alpha_{22}(G_w - G_i) + \alpha_{23}(B_w - B_i) + \epsilon_{2i}, \\ Y_i = \alpha_{31}(R_w - R_i) + \alpha_{32}(G_w - G_i) + \alpha_{33}(B_w - B_i) + \epsilon_{3i}, \end{cases} \quad (2)$$

$$i = 1, 2, \dots, 100.$$

其中 ϵ_{ki} ($k=1, 2, 3, i=1, 2, \dots, 100$) 独立同分布, $E(\epsilon_{ki})=0, V(\epsilon_{ki})=\sigma$. 记

$$\boldsymbol{\epsilon}_1 = [\epsilon_{11} \epsilon_{12} \dots \epsilon_{1,100}]^T, \boldsymbol{\epsilon}_2 = [\epsilon_{21} \epsilon_{22} \dots \epsilon_{2,100}]^T, \boldsymbol{\epsilon}_3 = [\epsilon_{31} \epsilon_{32} \dots \epsilon_{3,100}]^T,$$

$$\mathbf{c} = [C_1 C_2 \dots C_{100}]^T, \mathbf{m} = [M_1 M_2 \dots M_{100}]^T, \mathbf{y} = [Y_1 Y_2 \dots Y_{100}]^T,$$

$$X_k = \begin{bmatrix} R_w - R_1 & R_w - R_2 & \dots & R_w - R_{100} \\ G_w - G_1 & G_w - G_2 & \dots & G_w - G_{100} \\ B_w - B_1 & B_w - B_2 & \dots & B_w - B_{100} \end{bmatrix},$$

其中 $k=1, 2, 3$, 分别对应 C, M, Y 的 100 组数据.

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\alpha}_1^T \\ \boldsymbol{\alpha}_2^T \\ \boldsymbol{\alpha}_3^T \end{bmatrix},$$

这里 $\boldsymbol{\alpha}_1^T = [\alpha_{11} \alpha_{12} \alpha_{13}]$, $\boldsymbol{\alpha}_2^T = [\alpha_{21} \alpha_{22} \alpha_{23}]$, $\boldsymbol{\alpha}_3^T = [\alpha_{31} \alpha_{32} \alpha_{33}]$, 于是式(2)可写成

$$\mathbf{c} = \boldsymbol{\alpha}_1^T X_1 + \boldsymbol{\epsilon}_1, \mathbf{m} = \boldsymbol{\alpha}_2^T X_2 + \boldsymbol{\epsilon}_2, \mathbf{y} = \boldsymbol{\alpha}_3^T X_3 + \boldsymbol{\epsilon}_3. \quad (3)$$

上式中 $\mathbf{c}, \mathbf{m}, \mathbf{y}$ 是已知的, 称为观测向量; X_k 也是已知, 称为结构矩阵, 一般假定 X_k 是满秩矩阵, 从而 $\text{rank}(X_k X_k^T) = \text{rank}(X_k)$, 故 $(X_k X_k^T)^{-1}$ 存在, 且为 3×3 对称矩阵.

于是, RGB→CMY 颜色空间转换问题就化为了多项式回归分析问题, 即化为对系数矩阵 A 作估计的问题. 用多项式回归分析方法^[7], 对系数矩阵 A 中第 k 行系数 α_k 作估计, 得到

$$\bar{\mathbf{a}}_1 = (X_1 X_1^T)^{-1} X_1 \mathbf{c}, \quad \bar{\mathbf{a}}_2 = (X_2 X_2^T)^{-1} X_2 \mathbf{m}, \quad \bar{\mathbf{a}}_3 = (X_3 X_3^T)^{-1} X_3 \mathbf{y}. \quad (4)$$

易知, $\bar{\alpha}_k$ 是 α_k 的使误差平方和达到最小的估计, 且是 α_k 的最小方差线性无偏估计.

进一步地, 为提高模型精度, 在式(1)假设的模型中增加高次项, 可得如下几个模型:

$$C = \alpha_{11}(R_w - R) + \alpha_{12}(G_w - G) + \alpha_{13}(B_w - B) + \alpha_{14}(R_w - R)(G_w - G) +$$

$$\alpha_{15}(R_w - R)(B_w - B) + \alpha_{16}(G_w - G)(B_w - B),$$

$$M = \alpha_{21}(R_w - R) + \alpha_{22}(G_w - G) + \alpha_{23}(B_w - B) + \alpha_{24}(R_w - R)(G_w - G) +$$

$$\alpha_{25}(R_w - R)(B_w - B) + \alpha_{26}(G_w - G)(B_w - B),$$

$$Y = \alpha_{31}(R_w - R) + \alpha_{32}(G_w - G) + \alpha_{33}(B_w - B) + \alpha_{34}(R_w - R)(G_w - G) +$$

$$\alpha_{35}(R_w - R)(B_w - B) + \alpha_{36}(G_w - G)(B_w - B).$$

此时, $X_k X_k^T$ 为 6×6 对称矩阵, A 为 3×6 系数矩阵.

$$C = \alpha_{11}(R_w - R) + \alpha_{12}(G_w - G) + \alpha_{13}(B_w - B) + \alpha_{14}(R_w - R)(G_w - G) +$$

$$\alpha_{15}(R_w - R)(B_w - B) + \alpha_{16}(G_w - G)(B_w - B) + \alpha_{17}(R_w - R)^2 +$$

$$\alpha_{18}(G_w - G)^2 + \alpha_{19}(B_w - B)^2 + \alpha_{110}(R_w - R)(G_w - G)(B_w - B),$$

$$M = \alpha_{21}(R_w - R) + \alpha_{22}(G_w - G) + \alpha_{23}(B_w - B) + \alpha_{24}(R_w - R)(G_w - G) +$$

$$\alpha_{25}(R_w - R)(B_w - B) + \alpha_{26}(G_w - G)(B_w - B) + \alpha_{27}(R_w - R)^2 +$$

$$\alpha_{28}(G_w - G)^2 + \alpha_{29}(B_w - B)^2 + \alpha_{210}(R_w - R)(G_w - G)(B_w - B),$$

$$Y = \alpha_{31}(R_w - R) + \alpha_{32}(G_w - G) + \alpha_{33}(B_w - B) + \alpha_{34}(R_w - R)(G_w - G) +$$

$$\alpha_{35}(R_w - R)(B_w - B) + \alpha_{36}(G_w - G)(B_w - B) + \alpha_{37}(R_w - R)^2 +$$

$$\alpha_{38}(G_w - G)^2 + \alpha_{39}(B_w - B)^2 + \alpha_{310}(R_w - R)(G_w - G)(B_w - B).$$

此时, $X_k X_k^T$ 为 10×10 对称矩阵, A 为 3×10 系数矩阵.

在利用式(4)对上述模型中的 α_k 作估计时, 随着模型精度的提高, 增加的计算量主要集中在对 $X_k X_k^T$ 求逆的运算上, 学习程序的计算量将随 $X_k X_k^T$ 阶数的增加而增加, 因此, 在选择模型时应在精度和计算量之间取折衷.

得到系数矩阵 A 后, 自然也就得到了 RGB→CMY 转换算法. 若使用三色打印, 则可直接用打印驱动程序调用这些数据实现 RGB→CMY 转换并进行彩色图象打印; 若使用四色打印, 还需寻找 RGB→K 的规律, 对黑色 K 进行单独处理.

3.3 黑色 K 的处理

四色打印中, K 的引入是为了提高黑色纯度, 虽然理论上 C, M, Y 墨水等量混合在一起能产生黑色, 但由于所有打印墨水都不是真正的纯色, 所以它们实际产生的是一种浑浊的褐色. 因此, 必须混合进一些黑色墨水来产生真正的黑色, 同时要在 C, M, Y 中减去这部分 K , 即用 K 代替等量的 C, M, Y 来产生黑色.

显示的象素越暗, 则打印时使用的 K 色墨水应越多, 所以 RGB 与 K 之间存在着一种近似的反比关系. 如果用 $K = \alpha/\max(R, G, B) + \beta$ (α, β 为待发现的参数)^[5] 来描述 RGB

$\rightarrow K$ 规律,即用 $\max(R, G, B)$ 与 K 的函数关系近似代替 (R, G, B) 与 K 的函数关系,一方面,它没有充分利用训练数据提供的信息,学习的知识不全面;另一方面,它没有考虑人眼的视觉特性,发现的规律又不甚准确,因此,本文重新考虑 $RGB \rightarrow K$ 规律的知识发现方法.

由于人眼的视觉心理特性,实际观察图像时得到的亮度感觉并不完全与图像亮度信号一致.由韦伯-费赫涅尔定理^[8]可知,主观亮度与图像亮度信号的对数成线性关系,从而导致感觉上相同的亮度差别,对应于暗区的亮度值变化小,而对应于亮区的亮度值变化大,这与视觉实验可察觉的增量韦伯比是一致的.因此,用 $\log(1/Y)$ 与 K 的函数关系可以较好地反映 (R, G, B) 与 K 的函数关系,运用科学发现方法,通过对 K 的 100 组测试数据的学习,发现 K 与 $\log(1/Y)$ 的函数关系 $K = f_k(\log(1/Y))$,它是一个三次多项式,这里 Y 是由 (R, G, B) 计算出的亮度值.

3.4 基于色差的 C, M, Y, K 值调整

以前为了防止墨水过多导致纸被打透,只是简单地将 C, M, Y 压缩到一个经验参数内.为了增加学习能力,本文将这一经验参数用学习的办法来得到.考虑到人眼对彩色的分辨率低于对亮度的分辨率,对于两种相同色调不同饱和度的颜色,只有当计算出二者的色差 ΔE_{ab} 大于某一误差值 e 时,人眼才能区分出这两种色彩,而仅从数据 $\Delta C, \Delta M, \Delta Y$ 来判别其颜色差异是不准确的,因此,利用学习算法从纯色训练数据中先分别学习出人眼对 C, M, Y, K 纯色色块分辨力的饱和值 $C_{max}, M_{max}, Y_{max}, K_{max}$,然后将图像的色彩和亮度变化的动态范围压缩到学习出的饱和值范围内.

首先,将训练数据中的 (R, G, B) 变换到 CIEXYZ 颜色空间中,用式

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 2.7689 & 1.7517 & 1.1302 \\ 1.0002 & 4.5907 & 0.0600 \\ 0.0000 & 0.0565 & 5.5943 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

求出其对应的三刺激值 X, Y, Z .

由于在 CIE1931XYZ 颜色空间的 xy 色度图中各等色差域(麦克亚当椭圆)对应的坐标变化量很不均匀,这种不均匀性给衡量颜色差别带来不便,因此必须将 (X, Y, Z) 转换到 CIELab 均匀颜色空间^[7]中,求出其对应的色度值

$$\begin{aligned} L &= 116(Y/Y_n)^{1/3} - 16, \quad a = 500\{(X/X_n)^{1/3} - (Y/Y_n)^{1/3}\}, \\ b &= 200\{(Y/Y_n)^{1/3} - (Z/Z_n)^{1/3}\}, \end{aligned}$$

其中 X_n, Y_n, Z_n 为 2° 视场下 CIE 标准照明体 D_{65} 的三刺激值, $X_n = 95.0170, Y_n = 100.0, Z_n = 108.8130$, 颜色 (L_1, a_1, b_1) 与颜色 (L_2, a_2, b_2) 的色差 ΔE_{ab} 为

$$\Delta E_{ab} = \{(\Delta L)^2 + (\Delta a)^2 + (\Delta b)^2\}^{1/2},$$

式中 $\Delta L = L_1 - L_2, \Delta a = a_1 - a_2, \Delta b = b_1 - b_2$.

当 $\Delta E_{ab} < e$ (e 为误差) 时,人眼将不能区分出被检验的两种颜色,如果仍按两种不同的色彩进行打印,则此范围内的颜色细节势必会丢失.

根据上述思想,通过对训练数据的学习,在找到人眼对 CMYK 颜色分辨力的饱和值后,将图像的色彩和亮度变化的动态范围压缩到学习出的饱和值范围内,以达到增强图像对比度、突出图像的色彩和亮度细节的目的.

4 实验结果比较

在 Canon BJC-465彩色喷墨式打印机、误差扩散抖动方式、BCI-22e 墨水、HR101纸的条件下,在 Power Macintosh 7200/90计算机、PowerPC 8.0操作系统下,通过对样本数据的学习,打印了一些具有代表性的国际印刷标准图像,如图2~4所示。



图2 Photoshop 打印的结果



图3 先前研制的 HITCMM 软件打印的结果

可以看出用 Photoshop 打印的图2颜色较为正常,但细节过渡较为生硬、不自然;用先前研制的 HITCMM 软件打印的图3细节过渡较为自然,但存在一定的色偏;而用改进的 HITCMM 软件打印的图4细节过渡自然,细节分辨率也有所提高,同时色偏也有所减小。

5 结束语

彩色匹配是当前彩色印刷领域公认的一项世界性技术难题,本文通过对样本空间数据的学习得到 RGB→CMYK 色彩空间的转换方法,其优点主要有

- 1) 具有学习功能,易于进一步改进;
- 2) 具有通用性,针对不同的抖动模式、墨水和纸张均适用;
- 3) 不需进行大量的查表计算,打印速度快;



图4 改进的 HITCMM 软件打印的结果

- 4) 需要测试的实验数据少,工作量小;
- 5) 打印的图像细节过渡较为自然,且细节分辨率有所提高,对比度也增强了.

存在的问题是,通过变换后的空间有一定的减小,对于 RGB 空间中那些无法在 CMYK 空间中正确表达的颜色如何在 CMYK 空间中更好地近似表达等方面的问题还有待进一步地研究. 另外,考虑到墨水的打印顺序的影响,如果在算法中加入对混色数据的学习将会进一步提高打印图像的质量,目前这项工作我们正在研究之中.

参 考 文 献

- 1 Fairchild MD, Berns R S. Image color-appearance specification through extension of CIELAB. *Color Research and Application*, 1993, **18**(3):178~190
- 2 Bezerra C de M, Hawkyand C J, Kulube H M, Reyner S. Color matching for ink-jet prints on paper. *Color Research and Application*. 1998, **23**(1):18~26
- 3 Okatsu, Murai, Kita. High accurate color transformation based on UCR. In: Japan Hardcopy Papers Collection'94, 1994. 177~180
- 4 Hung Po Chieh. Colorimetric calibration in electronic imaging devices using look-up-table model and interpolations. *J. Electronic Imaging*, 1993, **25**(2):53~61
- 5 钱国良,陈彬等. 基于机器学习的彩色匹配技术. *软件学报*, 1998, **9**(11):845~850
- 6 [日]大田 登著,刘中本译. 色彩工学. 西安:西安交通大学出版社, 1997. 164~166
- 7 王学仁,温忠磷. 应用回归分析. 重庆:重庆大学出版社, 1989. 35~48
- 8 余松煜,周源华,吴时光. 数字图像处理. 北京:电子工业出版社, 1989. 26~32

苏晓红 女,1966年生,硕士,副教授. 主要研究领域为计算机图形学、数字图像处理、彩色匹配技术.

王亚东 男,1964年生,副教授. 主要研究领域为专家系统、机器学习、知识工程.

郭茂祖 男,1966年生,博士,副教授. 主要研究领域为机器学习、彩色匹配技术、非数值并行计算.

马培军 男,1963年生,教授. 主要研究领域为人工智能与专家系统、机器学习、彩色匹配技术.