

扫描仪色彩特征文件的生成及分析

Analysis of the Profile Generation for Scanners

吴桂芹 (西安工程大学 管理学院 陕西 西安 710048)

贺菲菲 (中兴通讯西安分公司 陕西 西安 710065)

摘要: 针对需要经常校正扫描仪并重新建立符合设备当时颜色特性文件这一实际情况, 本文设计并实现了扫描仪颜色特性文件生成程序。参照国际颜色委员会制定的 ICC Profile 标准格式, 依据矩阵模型和查找表模型分别实现了以线性插值法和多项式回归法为核心的两种扫描仪颜色特性文件生成算法, 并给出了两种模型在相同测试环境下生成的颜色特征文件的测试结果和误差分析。通过多项误差值的比较和分析得出查找表模型比矩阵模型的转换精度高。

关键词: 颜色管理 ICC 标准 设备特征描述文件 矩阵模型 查找表模型

1 引言

随着信息化和数字化时代的到来, 图像信息作为主要的信息载体, 在我们的生活中无所不在, 人们对精确传递图像信息的要求日益增高。扫描仪作为专业计算机图像输入设备, 对图像数据的准确复制起到了不可低估的作用。

国际彩色联盟 (ICC) 色彩标准已经得到了广泛认可, 大多数硬件设备在出厂时已带有各自的颜色特性文件 (Profile), 但是随着设备使用时间的增加以及环境的变化, 硬件参数会发生变化, 所以我们需要经常校正设备并重新建立符合设备当时特性的 ICC Profile 文件, 这样才能充分发挥设备的性能, 这也是用软件来补偿硬件性能的一种方法。ICC 利用一个与设备无关的色彩空间 (Profile Connection Space, 缩写 PCS) 作为色彩空间转换的中间媒介空间^[1], 通过量化彩色设备色彩空间 (RGB 或 CMYK) 和 PCS 空间之间的联系, 生成该设备的颜色特征描述文件 (ICC Profile 文件), 建立设备色彩空间和 PCS 空间的色彩转换模式, 实现色彩在各个设备之间传递的一致性。

2 工作原理及流程

随着设备的老化扫描仪的扫描特性和表色特性也会发生变化, 必须通过重新确定扫描仪的工作基准来控制设备并修正设备偏差以使得扫描仪能够实现正确

的色彩表达, 忠实于原稿^[2]。为了生成相对准确的扫描仪颜色特征文件, 首先需要校准扫描仪, 即保证对扫描仪进行校准后, 对于同一份原稿, 不论什么时候扫描, 都应当能获得相同的图像数据。然后扫描一张标准色标, 目前常用的 IT8 色标, 覆盖了整个 CIE L * a * b * 色彩区域的采样, 底部带有一个中性灰度梯尺。色标上的色块由已校准的分光光度仪测量其色度值 L * a * b *, 从而生成色标的 L * a * b * 参数表。扫描仪颜色空间是 RGB, 要建立某个扫描仪的特征文件时, 用该扫描仪扫描色标并获得色标上每一块色块的 RGB 值, 这样, 就可以建立一张 RGB 和 L * a * b * 之间的转换速查表, 它可用来将扫描仪上生成的 RGB 文件的某一点映射到 L * a * b * 空间上, 根据这个映射关系可以生成扫描仪特征文件。

ICC Profile 的文件结构, 主要为三部分: 文件头, 标签表和标签元素数据^[3]。根据设备类型的不同, ICC Profile 中所含的必要标签不同, 而且对于同一种设备, 由于生成 ICC Profile 的算法模型不同也会产生不同的标签组合。色彩管理系统在转换色彩时会从标签表里根据所需要的标签的名称找到该标签位置, 然后读出这个标签的数据来进行色彩转换的数学计算。

对于彩色扫描仪, ICC Profile 规范提出了两种模型算法, 分别是“阶调/矩阵模型”和“查找表模型”, 它们的原理都是通过对标准色卡的扫描值 (RGB 色彩空

间)和标准值(Lab 或 XYZ 色彩空间)进行比较,做出定量分析,然后按照规定的格式写到相应的标签里。但是它们的计算精度不一样,而且最后生成的文件大小也不一样。

2.1 矩阵模型

矩阵模型如图 1 所示。其中,三条阶调再生曲线(TRC)是三个一维查找表(1D Lut),三个色彩分量 R、G、B 分别通过相应的 TRC 转换为线性值,再通过 3×3 矩阵转换为设备无关空间 CIEXYZ。该算法的优点是简单,生成的设备颜色特性文件小,处理速度快;缺点是精度不高,比较适合于显示器和线性设备。但是目前大多数扫描仪设备都不是线性的,所以不常用矩阵模型。

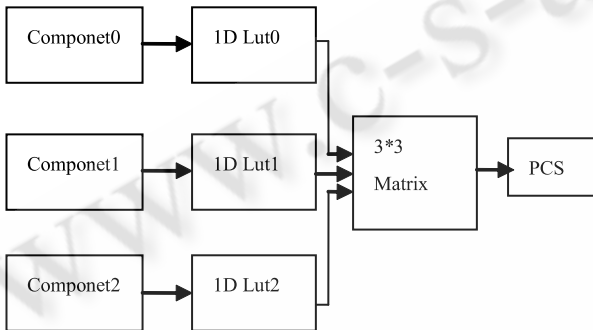


图 1 矩阵模型处理流程^[3]

图 1 中, Component0、Component1、Component2 表示 R、G、B 三个颜色通道,把图像里提取出来的 R、G、B 颜色值先代入 TRC 曲线,经过了 TRC 后,将图像的 R、G、B 值线性化了,并且归一到了 0 到 1 之间。然后将这些值通过 3×3 矩阵,最后转到 PCS 空间,这里的 PCS 空间一般为 XYZ。

在矩阵模型的必要标签里面 redTRCTag、greenTRCTag、blueTRCTag、redColorantTag、greenColorantTag、blueColorantTag 这六个标签记录了用矩阵模型进行颜色转换所需要的 TRC 曲线和矩阵。

2.2 查找表模型

查找表模型如图 2 所示。

其中 3×3 矩阵是专门针对输入数据是 XYZ 值的,而扫描仪的色彩空间是 RGB,所以这个 3×3 矩阵被定义为单位阵;每个一维输入表是一个一维数组,将输入的 R、G、B 值从 0 ~ 255 规化到 0 ~ 65535 之间,分别存储到三个一维数组里;一维输入表的输出是多维表的

输入,通过多维表完成从扫描仪 RGB 色彩空间向 PCS 色彩空间的转换,最后的结果也需要规化到 0 ~ 65535 之间;每个输出表也是一维数组,数据经输出表变换后即可得到从一个色彩空间向另一个色彩空间的转换结果。

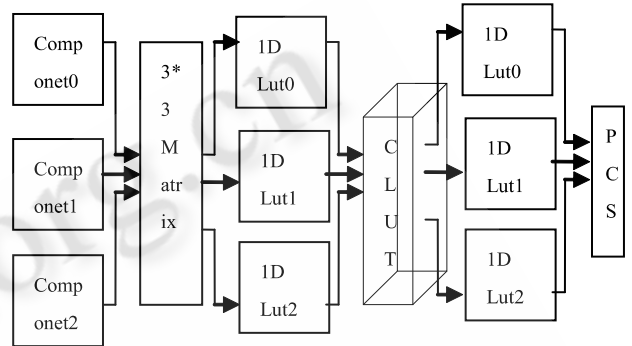


图 2 查找表模型处理流程^[3]

这种模型的优点是精度高,灵活性高,适合任何设备,但是数据转换复杂,插值时可能会出现数据的不确定性,生成的特征描述文件庞大,处理速度较慢,是现在最常用的算法模型。

在查找表模型的必要标签里,所有的颜色转换值都存储在 AtoB0 标签里的。多维表 CLUT 是整个 AtoB0 标签的重中之重,为了提高 CLUT 的精度,要对其进行栅格化。栅格点的多少直接影响到颜色数据转换的精度。为了兼顾速度,要在保证颜色转换精度的前提下尽量减少栅格点的数目。

3 算法具体实现流程

算法流程如下^[4]:

(1) 准备数据。用分光测色仪测量得到的 IT8.7 标准色卡的标准 CIEXYZ 或 CIELAB 数据。扫描仪在校正后的稳定状态下对色卡进行扫描的实际图像,并根据图像文件数据读出每个色块的 RGB 值,这两组数据都以 *.txt 文件格式存储;

(2) 查找表模型直接跳转到步骤 5。建立 TRC 曲线。根据 24 级灰度尺的实际 RGB 值和标准亮度 Y 值,用线性插值的方法,分别建立 R 与 Y、G 与 Y、B 与 Y 的关系,由于标准 Y 值达不到 100,RGB 值也达不到 255,所以在插值之前都要分别除以它们的最大值,以保证插值以后的数值达到 1,算法针对每个 TRC 曲线共计算

出了 4096 个点,在写入 TRCTag 标签之前,还要将这些归一化的值乘以 65535;

(3) 线性化 624 个扫描色块的实际 RGB 值。根据步骤 2 得出的 TRC 曲线将 624 个色块的 RGB 值都进行线性化;

(4) 计算 3 × 3 转换矩阵。根据线性化的 RGB 值和标准 CIEXYZ 的值建立转换关系,如式 (1) 所示:

$$\begin{cases} X=r_1R'+g_1G'+b_1B' \\ Y=r_2R'+g_2G'+b_2B' \\ Z=r_3R'+g_3G'+b_3B' \end{cases} \quad (1)$$

R' G' B' 为已知的 624 个扫描色块的实际 RGB 线性化值,XYZ 是归一化的标准 CIEXYZ 值,将它们代入矩阵方程,求出转换矩阵。求出来的系数矩阵被写入 redColorantTag、greenColorantTag、blueColorantTag 中;

(5) 对于查找表模型,核心是 AtoB0 标签中多维表的生成,通过 3 × 11 多项式求出实际 RGB 值和标准 CIEXYZ 值之间的关系,具体如式 (2) 所示:

$$\begin{cases} X=A_0+A_1R+A_2G+A_3B+A_4RG+A_5RB+A_6GB+A_7R^2+A_8G^2+A_9B^2+A_{10}RGB \\ Y=B_0+B_1R+B_2G+B_3B+B_4RG+B_5RB+B_6GB+B_7R^2+B_8G^2+B_9B^2+B_{10}RGB \\ Z=C_0+C_1R+C_2G+C_3B+C_4RG+C_5RB+C_6GB+C_7R^2+C_8G^2+C_9B^2+C_{10}RGB \end{cases} \quad (2)$$

(6) ICC Profile 文件中所有需要写入的数据都已经准备好了,按照 ICC 标准建立该扫描仪的 ICC Profile 文件。

4 实验结果及分析

软件测试步骤如下:

(1) 用分光测色仪测量色标卡的标准 CIEXYZ 值或 CIELAB 值;

(2) 用扫描仪扫描色标卡 KODAK Q-60 IT8.7/2,分辨率为 300dpi,其他参数设为默认值,扫描图像存为 bmp 格式;

(3) 运行软件,生成该扫描仪的 ICC Profile 文件;

(4) 使用 ICC Profile Inspector 查看生成的 ICC Profile 里面的标签值是否正确;

(5) 计算实际扫描色卡经过生成的 ICC Profile 文件中的模型处理以后的值,求出色差 ΔE,色差公式如式 (3):

$$\Delta E = \sqrt{(L - L_0)^2 + (a - a_0)^2 + (b - b_0)^2} \quad (3)$$

下面给出在相同测试环境下生成的颜色特征文件

的测试结果和误差分析。

4.1 矩阵模型结果

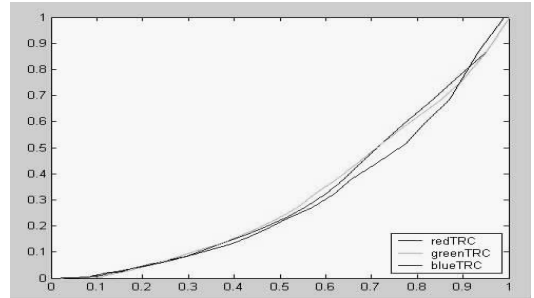


图 3 TRC 曲线比较

图 3 表现了生成的 ICC Profile 文件中的 redTRC、greenTRC、blueTRC 三条 TRC 曲线的比较结果。它是通过扫描后的色标卡图像中的 24 级灰度尺的实际 RGB 色彩分量和标准 CIEXYZ 中的 Y 分量建立的线性关系,RGB 分量值和 Y 分量值都是归一化以后进行线性插值的。所有需要通过矩阵模型进行颜色管理的 RGB 图像数据都要先经过这三条曲线进行线性化。从图中我们可以看出,这三条 TRC 曲线的形状基本一致,说明扫描仪的灰平衡性能比较好。由于使用的是线性插值法,要求曲线保证通过数据点,所以曲线不是很平滑。

当然,从图 3 中我们只能看到这三条 TRC 曲线的形状,下面以 redTRC 为例,使用 ICC Profile Inspector 看一下 redTRCTag 的情况,如图 4 所示,横坐标表示 R 分量的值,它是将 0 ~ 255 之间的数据规范化到 0 ~ 4096 里,线性化以后的 Y 值也需要规划到 0 ~ 65535 之间。从图上可以看出来,这条曲线是采用了 4096 个采样点

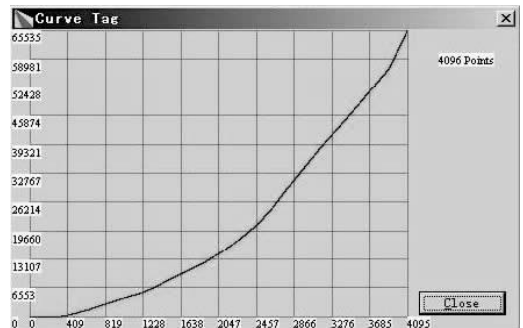


图 4 redTRCTag 曲线图

进行描述的。

将实际扫描的色卡中 264 个色块的 RGB 值通过模型转换到 PCS 空间, 计算每一个色块的色差值 ΔE , 然

后对结果进行统计和误差分析, 如表 1 所示:

表 1 矩阵模型色差统计

色差范围	$0 < \Delta E < 1$	$1 < \Delta E < 2$	$2 < \Delta E < 3$	$3 < \Delta E < 4$	$4 < \Delta E < 5$	$5 < \Delta E < 6$	$\Delta E > 6$
色块统计	32	86	47	25	15	15	44
平均色差	3.68	最大色差	26.53	方差	15.18	标准差	3.90

从表 1 中看出, 色块的色差值有 63% 小于 3, 21% 在 3 到 6 之间, 另外 16% 大于 6。其中平均色差 $\Delta E = 3.68$, 最大色差 $\max \Delta E = 26.53$ 。

4.2 查找表模型结果

在实现查找表模型是采用了 3×11 多项式回归算法, 通过色卡扫描后的实际 RGB 分量和标准色卡 XYZ 值进行矩阵运算, 得到 3×11 多项式的系数矩阵, 如式 (4) 所示:

$$\begin{cases} X=0.0061-0.0304R-0.0075G-0.0135B+0.3244RG+0.0444RB+0.0543BG-0.391R^2+0.1478G^2+0.1388B^2+0.0318RGB \\ Y=0.0062-0.0179R-0.0276G-0.0103B+0.3152RG+0.0182RB+0.1448BG+0.1697R^2+0.3795G^2+0.0234B^2-0.0363RGB \\ Z=-0.0070+0.0731R+0.0635G-0.0388B-0.1723RG-0.0803RB-0.1255BG-0.0033R^2+0.0696G^2+0.718B^2+0.4848RGB \end{cases} \quad (4)$$

表 2 是对查找表模型的色差统计数据, 统计方法和矩阵模型一致。从表中的数据可以得出: 色块色差值在 3 以内的有 63%, 在 3 到 6 之间的占 26%, 11% 的大于 6。平均色差值 $\Delta E = 3.14$, 最大色差 $\max \Delta E = 16.55$ 。

表 2 查找表模型色差统计

色差范围	$0 < \Delta E < 1$	$1 < \Delta E < 2$	$2 < \Delta E < 3$	$3 < \Delta E < 4$	$4 < \Delta E < 5$	$5 < \Delta E < 6$	$\Delta E > 6$
色块统计	35	63	67	38	18	12	31
平均色差	3.14	最大色差	16.55	方差	6.18	标准差	2.49

4.3 误差分析

通过对矩阵模型和查找表模型的试验结果进行分析, 发现查找表模型比矩阵模型的平均色差小 0.54, 最大色差小了 10, 具体的色差分布比较如图 5 所示:

致, 所以要提高这些色块的转换精度, 必须根据它们的特点进行专门的校正, 必要的时候可以采用分区域校正。由于采用了 CIEXYZ 和实际 RGB 值来建立转换关系, 在将 CIEXYZ 转换到 CIELAB 空间的时候, 也会带来较大的误差。

参考文献

- 1 赵小梅. 浅谈印前系统中的色彩管理. 机电信息, 2005, 10: 13 - 14.
- 2 杨尚昆. 基于 ICC 标准的输入设备色彩管理的实现. 今日印刷, 2007, 11: 7 - 8.
- 3 International Color Consortium. ICC Profile Format Specification. Version 4.0.0, 2001, 12: 20 - 87.
- 4 贺菲菲. 专业图像处理软件——“扫描仪颜色特征文件生成插件”的研究[硕士论文]. 西安: 西安电子科技大学, 2005.

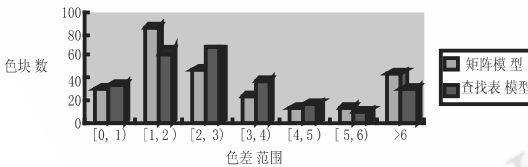


图 5 矩阵模型与查找表模型色差范围比较

从图中看出, 两个模型色差范围分布形态基本一致, 但是查找表模型中色块的色差的色差值较平均, 而且大于 6 的色块数少于矩阵模型, 由于查找表的最大误差是 16.55, 小于矩阵模型将近 10, 所以从整体上比较得出: 查找表模型的精度高于矩阵模型。另外, 从实验数据发现, 色差值大于 6 的色块基本分布在饱和度高或亮度的色块中, 由于亮度低, 色相难以分辨, 因此色差比较大。由于所有的色块的转换精度并不一