

非均匀三维查找表的快速查找技术^{*})

丁二锐 曾平 罗雪梅 王义峰

(西安电子科技大学外部设备研究所 西安 710071)

摘要 利用四面体插值算法进行色彩校正时,通过样本集直接建立起来的三维查找表往往是不均匀的,为了生成符合 ICC 规范的 Profile,必须通过插值得到均匀的 Profile 数据。在均匀化过程中,耗时最严重的是如何准确定位插值颜色点所在的四面体。从数据的邻域相关性,除了基于历史的局部搜索法,同时从粗定位和细定位角度提出了基于辅助表的快速定位算法。实验结果证明任何一种加速算法都有效缩短了 Profile 数据生成的时间,同时结合两种加速算法的效果更佳。

关键词 色彩管理,色彩校正,四面体算法,ICC Profile,快速查找

Technologies on Fast Search for Non-uniform 3D-LUT

DING Er-Rui ZENG Ping LUO Xue-Mei WANG Yi-Feng

(Research Institute of Peripherals, Xidian University, Xi'an 710071)

Abstract In the process of color calibration using tetrahedral interpolation, 3D-LUT created directly by color patch sample is usually not uniform, which contradicts the ICC profile specification and makes it necessary to obtain the uniform profile data by interpolation. However, the time-consuming location of proper tetra for interpolation for uniform data is a serious problem for creating the profile data. Two fast searching algorithms, namely, local searching based on history and fast location using auxiliary table, are presented from the perspective of neighborhood correlation and location at the rough and detailed level. Experimental results indicate either of the fast searching algorithms can effectively shorten the time of creating profile data and the combination of two fast searching algorithms can perform better than either of them.

Keywords Color management, Color calibration, Tetrahedral interpolation, ICC profile, Fast searching

1 引言

为了解决各种成像设备如显示器、打印机等等之间成像色彩的一致性,色彩管理技术通过引入与设备无关颜色空间,建立设备的彩色特性描述,并按用户要求进行颜色空间转换,即色彩校正。彩色校正的主要任务是补偿因设个体的颜色空间的基的偏差,其实现方法可以划分为多重回归、三维插值、神经网络和模糊逻辑等^[1]。

作为三维插值的一种算法,四面体插值算法^[1,2]由于其在实现代价和计算速度上的优势已广泛应用于实际的设备校正中。色彩校正前,通常需要一样本集来描述设备特性。然而通过样本集直接建立起来的三维查找表是不均匀的,即是非均匀的三维查找表。ICC Profile 规范^[3]要求最后的数据的源色彩空间必须是均匀的。这样以来,直接建立起的三维查找表数据不能直接利用,必须经过四面体插值算法得到均匀的数据。插值的过程就意味着必须定位非均匀三维查找表中的四面体。但是如何快速精确定位到插值所需的四面体确实是一个具有挑战性的问题。本文从两种不同角度,即,数据的邻域相关性和粗定位以及细定位角度,提出了两种不同的加速算法:基于历史的局部搜索算法和基于辅助表的快速定位

算法。实验证明两种算法均可以达到加速生成 Profile 数据的效果,同时结合两种算法进一步提高了生成速度。

2 ICC 规范约束下的 Profile 生成

ICC Profile 规范^[3]要求查找表数据的域(源色彩空间)必须包含包括 0.0 到 65535.0 之间的所有实数。第一个数据应该位于 0.0,最后一个数据位于 65535.0,中间的值依 65535.0/(M-1)间隔均匀分布,其中 M 为三维查找表每一维上的格点数。

在生成 A2B 数据的时候,由于 CMYK 的数据在每一维上切分的,所以查找并插值的过程比较简单,如图 1 所示。

样本集生成的时候不知道设备特性,所以均匀分色样本集在设备无关色彩空间中是不均匀的,也就是说,即使样本集的 CMYK 数据是有规律的,其对应的 Lab 数据在 Lab 空间也是不均匀的。所以 CMYK 的空间中的立方体在 Lab 空间中就变成了一个变形的六面体^[4]。因此,在生成 Profile 的 B2A 数据的时候,从 Lab 到 CMYK 的插值过程中,不能一次定位到需要插值点所在的四面体位置。为此,需要找到快速定位(搜索)四面体的算法来均匀化数据,即反向均匀化^[5]。以下讨论的两种加速定位(搜索)方法虽假设在生成 B2A 数据时提出,但具有普适性,亦可用于 A2B 数据生成过程。

^{*}基金项目:国防预研基金资助项目(413160501);西安电子科技大学研究生创新基金资助课题(05008)。丁二锐 博士研究生,主要研究方向为色彩校正,图形图像处理。曾平 博士生导师,主要研究方向为图形图像处理,电子成像技术,色彩管理。罗雪梅 博士研究生,主要研究方向为色外观技术。王义峰 博士研究生,讲师,主要研究方向为自适应色域匹配。



图1 A2B插值过程

3 两种加速算法的原理

3.1 原始方法

在没有分析需要查找数据之间的关系的时候,很显然,最为直接的方法就是在非均匀三维查找表中从头至尾查找每个四面体。这种方法虽可确保查找到插值四面体,但其速度显然逊色许多。

3.2 基于历史的局部搜索法

该算法的依据主要来源与数字图像处理中像素邻域相关性^[6]。对于图像的任一个像素,通常它周围的像素的颜色值和它的颜色值是比较接近的,即,这些像素都在一个邻域中,因此也就可以通过周围邻域像素的颜色值来估计中心像素的颜色值。

在生成 B2A 数据的时候,需要均匀采样 Lab 空间进行反向均匀化。对 Lab 空间采样的数据(Profile B2A 数据)是按照 Lab 高维到低维的顺序存放的,相邻两个 Lab 数据至多在一个维上变化,也就是说,这两个值是相近的。基于上述原因,如果已经查到一个 Lab 所在四面体,那么下一个 Lab 就应该在这个四面体的一个邻域中可以找到。

由此得到基于历史的局部搜索法的算法的形式化描述如下:

```

Algorithm LocalSearching
{
  //NLab为 Lab 需要插值的数目,即,Profile 数据的数目;NTetra为四面体的数目;
  for(i=0;i<NLab;i++)
  {
    if (不是第一查找的 Lab)
    {
      for (j=StartI; abs(j-StartI)<εTetra; j++)//在邻域 εTetra 中查找
      {
        if Search_TetraList_Interpolation()
        {
          StartI=j; //若在邻域 εTetra 中,记录查到四面体的位置
          else//若未在邻域 εTetra 中查找到,则查找所有四面体
          break;
        }
      }
    }
    else//记录第一次查找位置
    {
      for(j=0;j<NTetra;j++)
      {
        if Search_TetraList_Interpolation()
        {
          StartI=j; //记录第一次查到四面体的位置
        }
      }
    }
    for (j=0;j<NTetra;j++)//在邻域 εTetra 中未查找到,查找所有四面体
    {
      if Search_TetraList_Interpolation()
      {
        StartI=j; //记录查到四面体的位置
      }
    }
  }
}
    
```

3.3 利用辅助表快速定位法

在四面体插值算法中,判断一个点是否在某个四面体中的时候,主要是判断该点与这个四面体的四个面的位置关系。如果存在从某个面上的点指向该点的向量和该面指向四面体内部的法向量的数量积为负数,那么就可以说明这个点在四面体序列之外。这种方法由于在各个面上都要这样计算,所以对于查找很多数据来是非常费时的。如果可以知道包络这个四面体的外接长方体,那么可以首先粗略地判断是否在这个长方体中。判断是否在长方体中比较简单,只需要比较边界值即可。如果在长方体中,则进一步再用上述数量积的判断方法精确判断出是否在该四面体中。将这些长方体的边界存成一张辅助表,那么上述的过程就是通过这张表来快速定

位的。

这种算法的形式化描述如下所示:

```

Algorithm FastLocation
{
  for(i=0;i<NLab;i++)
  {
    if (Value.L<Tetra.MaxL && Value.L>Tetra.MinL)
    {
      if (Value.a<Tetra.Maxa && Value.a>Tetra.Mina)
      {
        if (Value.b<Tetra.Maxb && Value.b>Tetra.Minb)//粗定位,判断是否在外接长方体内
        {
          if IsInTetra //细定位,判断是否在四面体内。若在,插值;否则,查找下个满足粗定位的四面体。
          {
            Interpolation();
          }
          else
          {
            continue;
          }
        }
      }
    }
  }
}
    
```

4 理论及实验结果分析

以下讨论对原始方法以及上述两种加速算法的时间效率分析。为讨论方便,记原始方法为 M1,基于历史的局部搜索法为 M2,利用辅助表实现快速查找算法为 M3。

理论分析:设非均匀三位查找表共有四面体 N 个,而且要搜索重新划分源色彩空间的 L 个值(Profile 数据)。(一般说来, $L \approx d * N$ 。当源色彩空间的划分和最后生成 Profile 数据的级数确定之后, d 为一个常数)。

首先对于 M1 来说,由于查找 L 个值的时候需要从头到尾查找 N 个四面体,所以时间复杂度显然就是 $d * N * N$,即就是 $O(N^2)$ 。

对于 M2 来说,设查找第 i 值时所确定的邻域为在上次查找四面体的前后 ω_i ($1 \leq i \leq L$) 个四面体中查找。设 $\omega = \text{Max}(\omega_i)$,那么采用该方法的时间复杂度为 $2 * \omega * d * N$,即就是 $O(N)$ 。

对于 M3 来说,由于粗略判断是否在长方体中的时间非常小,但是由于要搜索整个四面体序列 N ,综合考虑后认为对于每个需要查找的值来说时间为 D (D 为一实常数),那么要查找 L 个值的时间复杂度为 $D * L$ ($D * d * N$),即就是 $O(N)$ 。

综上,各种算法的理论分析如表 1 所示。

表 1 理论分析

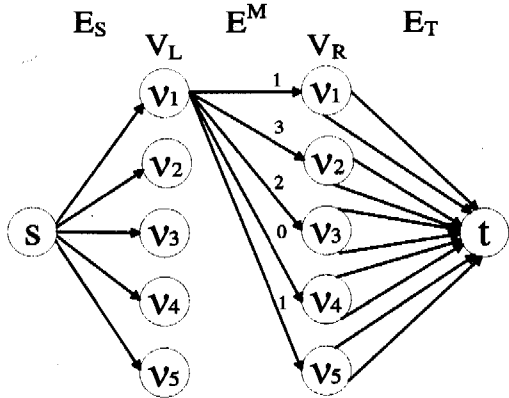
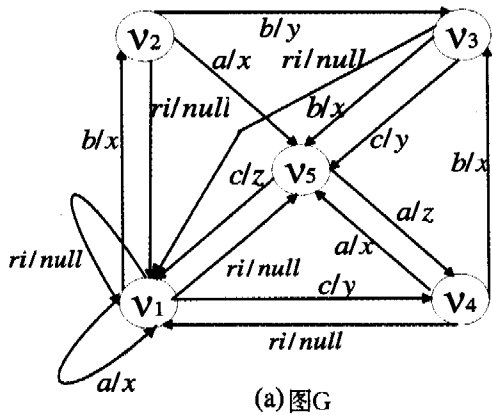
算法	M1	M2	M3
时间复杂度	$O(N^2)$	$O(N)$	$O(N)$

试验分析:试验环境为:CPU 为 PIV 1.8GHz,内存为 256M。样本集的数目为 2912。Lab 空间 17 级均分。测试过程为建立生成 9 级 A2B 数据和 17 级 B2A 数据。其时间测试如表 2 所示。

表 2 实验分析

算法	M1	M2	M3
时间/秒(A2B)	17.375	1.75	4.688
时间/秒(A2B & B2A)	54.953	37.86	21.125

(下转第 276 页)



(a) 图G

(b) 图GM

图 1

表 1 CFSM 各顶点的多 UIO 序列

状态	UIO 序列	UIO 尾状态	MUIO
v_1	a/x a/x	v_1	UIO_1^1
	a/x b/x	v_2	UIO_2^1
	c/y b/x	v_3	UIO_3^1
	a/x c/y	v_4	UIO_4^1
	b/x a/x	v_5	UIO_5^1

5 算法仿真实验

算法仿真是在 Simu-BlueTSGT 实验环境下进行的。

(上接第 259 页)

显然,分别使基于历史的局部搜索法和利用辅助表实现的快速查找算法比从原始查找方法快了好多。在生成 A2B 的数据时候,基于历史的局部搜索法显然优于利用辅助表实现的快速查找算法。但是当还需生成 B2A 数据的时候,前者却比后者慢得多。究其原因,是因为在使用前者的时候,一旦发现所要的数据没有在邻域当中时,算法必须查找整个非均匀三维查找表的四面体。鉴于此种情况,结合两种算法,即在生成 A2B 数据的时候,采用基于历史的局部搜索法;在生成 B2A 数据的时候,采用利用辅助表实现的快速定位算法。最后整个的生成过程所需的时间为 18.532s。

结论 考虑到生成 Profile 数据之间的相关性,本文提出了基于历史的局部搜索算法,该算法利用历史数据的信息来加速生成新数据,从而有效提高了 Profile 的生成速度。利用辅助表的快速定位方法,避开一部分判断四面体的开销,通过粗定位和细定位进一步提高了生成 Profile 数据的效率。但

Simu-BlueTSGT 是 BlueTSGT 系统的一个仿真系统,直接调用 C 语言书写 CFSM 对应的有向图的中间形式。为简化仿真,在仿真实验中,直接输入本文算法解释实例的有向图的中间形式,仿真结果如下:

启动多 UIO 优选算法生成的测试序列为:

```
ri/null@a/x@a/x@a/x@b/x@b/x@b/x@c/z@ri/null
@a/x@b/x@ri/null@a/x@b/x@a/x@c/z@c/y@b/x@c/y
@a/z@b/x@c/y@ri/null@b/x@a/x@c/z@c/y@b/x@ri/
null@a/x@b/x@b/y@b/x@a/z@a/x@a/z@b/x@b/x@a/
z@ri/null@c/y@b/x@c/y@c/z
```

不启动多 UIO 优选算法生成的测试序列为:

```
ri/null@b/x@b/y@c/y@c/z@a/x@b/x@a/x@a/z@
b/x@c/y@c/z@b/x@a/x@ri/null@b/x@a/x@a/z@a/x@
c/y@c/z@b/x@c/y@a/z@b/x@b/x@c/z@b/x@b/y@b/x
@c/z@b/x@a/x@c/z@b/x@ri/null@b/x@a/x@a/z@b/x
@ri/null@b/x@a/x@a/z@b/x@b/x@c/z@ri/null@b/x@
a/x@a/z@ri/null@b/x@a/x@c/z
```

长度分别为 44 和 55。

可见,简化后的算法确实优化了测试序列的费用。

结论 本论文参考了基于乡间中国邮递员遍历算法利用状态的 UIO 序列获得测试序列的方法,考虑到使用此方法生成测试序列时, G^{test} 的度: $\Delta(G^{test}) = \sum_{i=1}^n |\xi(v_i)|$ 的减少导致了相应的测试序列总时间费用的减少,以及一个状态的 UIO 序列不止一个的事实,提出了一种给每个状态选择一个适当的 UIO 序列的优选算法,使得 $\Delta(G^{test})$ 最小,从而使得最终生成的测试序列总时间费用减少,达到优化测试序列费用的目的。

参考文献

- 1 Aho A V, Dabhura A T, Lee D, Uyar M U. An optimization technique for protocol conformance test generation based on UIO sequences and rural Chinese postman tours. In: Protocol Specification, Testing and Verification, VIII, S. Aggarwal and K. K. Sabnani, eds. New York: Elsevier North-Holland, 1998. 75~86
- 2 Bondy J A, Murty U S R. Graph Theory With Applications. New York: Elsevier North Holland, 1976
- 3 Aho A V, Dabhura A T, Lee D, Uyar M U. An optimization technique for protocol conformance test generation based on UIO sequences and rural Chinese postman tours. In: Proc. IEEE INFOCOM, 1997
- 4 Miller R E, Paul S. On the Generation of Minimal-Length Conformance Tests for Communication Protocols. In: Proc. of IEEE Int'l Communication Conference. Atlanta: IEEE Computer Society, 1993
- 5 周炯磐著. 通信网理论基础. 北京邮电大学出版社, 1991. 124~134

需要注意的是在基于历史的局部搜索法中邻域确定得不合适可导致搜索时间的增加,因此下一步的工作重点可放在如何确定适当的邻域来最大化该算法的效率。

参考文献

- 1 Kang H R. Printer-Related Color Processing Techniques. In: Proceedings of SPIE, 1995, 2413, 410~419
- 2 Kang H R. Comparisons of Three-Dimensional Interpolation Techniques by Simulations. In: Proceedings of SPIE, 1995, 2414, 104~114
- 3 International Color Consortium. ICC Profile Specification. USA: ICC, 2001
- 4 史瑞芝. 用查表插值法实现 CIE Lab 与 CMYK 颜色空间的转换. 测绘学院学报, 2002, 19(3): 224~226
- 5 乔新亮. 面向彩色打印的 3D-LUT 优化技术. [学位论文]. 西安: 西安电子科技大学, 2002
- 6 Sonka M, Hlavac V, Boyle R. Image Processing, Analysis, and Machine Vision. (2nd Version). Beijing: PTPH & Thomson Learning, 2002