

基于三维欧氏距离的瓷砖颜色匹配问题分析

冉智文

重庆交通大学机电与车辆工程学院 重庆 400074

摘要: 本文主要讨论了有关马赛克瓷砖选色等相关问题, 通过数据拟合分析将RGB值近似看成三维坐标系的xyz轴, 并通过赋予权重, 考虑瓷砖成本和表现力为厂家实际生产与测试提供理论参考。

关键词: 数据拟合分析; 瓷砖颜色; R、G、B值; 权重; 欧式距离; 贪心算法思想; 拟合程度比率; 回归曲线

Analysis of tile color matching problem based on three-dimensional Euclidean distance

Zhiwen Ran

Mechanotronics and Vehicle Engineering of Chongqing Jiaotong University, Chongqing, 400074

Abstract: This paper mainly discusses the color selection of mosaic tiles and other related problems. Through data fitting analysis, the RGB value is approximately regarded as the XYZ axis of the three-dimensional coordinate system, and by giving weight and considering the cost and expressiveness of tiles. It provides a theoretical reference for manufacturers' actual production and testing.

Keywords: Data fitting analysis; Tile color; RGB value; Weight; Euclidean distance; Greedy algorithm; Fitting degree rate; Regression curve

1 问题重述

1.1 问题背景

马赛克是已知最古老的装饰艺术之一。它是用小瓷砖或小陶片制作的图案。在现代, 马赛克更多的是一种瓷砖。它的体积很小。它通常由几十块小砖组成, 形成一块较大的砖^[1]。它具有小巧玲珑, 色彩鲜明的特点, 广泛应用于室内小面积的地面, 墙面和室外大小的墙面和地面。它以丰富多彩的形式成为装饰材料的宠儿, 受到时尚前卫家庭的喜爱。

1.2 问题提出

由于技术和成本的限制, 瓷砖的颜色只能是有限的。拼接图案时, 员工应先根据原图颜色选择颜色相近的瓷砖再拼接。如何减少手工选色的工作量, 让客户选择满意的瓷砖已成为商家关注的问题^[2]。

问题一 找到每种颜色最接近的瓷砖颜色并得出空间分布。

问题二 如果厂家技术升级, 准备研发新颜色的马赛克瓷砖。在忽略研发难度, 只考虑拼接后的综合表现力, 应该优先增加哪些颜色的瓷砖? 增加的颜色由1逐

渐到10请分别给出对应颜色的RGB编码值。

问题三 如果研发任意一种新颜色马赛克瓷砖的成本完全一致, 颜色对其无影响, 在成本和表现力的共同约束下, 新增哪几种颜色更好, 给出对应的RGB编码值。

2 问题假设

假设原始图像为24位真彩色格式, 即R、G、B三个颜色分量均为8位, 共有28三次方16777216种颜色。

1. 假设两个空间坐标点之间的距离近似为欧氏距离, 其换算公式为 $\Delta = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$, 其中, x_1, y_1, z_1 分别代表第一点的经纬度, x_2, y_2, z_2 代表第二点的经纬度。

2. 假设两种颜色之间的欧式距离大小与颜色表现力存在线性关系。

3. 忽略瓷砖颜色的表现力对R、G、B值的微分误差。

4. 假设所找数据真实、可靠、有代表性。

3 问题分析

通过对问题的研究, 我们发现新增瓷砖的表现力与原有瓷砖的RGB值有关, 我们建立一个空间直角坐标

系,将颜色RGB值分别设置为空间坐标系的xyz值,通过计算两种颜色之间的欧氏距离,用欧式距离的数值来体现新增颜色的表现力,欧氏距离越小,说明该新增颜色的表现力越强。

问题一 假设RGB看成空间坐标,可以计算瓷砖颜色与已有22种颜色的距离,寻找距离最短的记为最相似。最后可视化出自动匹配的效果

问题二 独立计算22种颜色RGB值的相邻距离,用贪心算法,优先填补最大间距,通过填补前后的均匀程度来判断效果。

问题三 通过表现力的增长曲线来确定所要增加颜色个数。

4 问题处理方法

问题一 使用 $\text{dist}(a, b')$ 函数, dist 函数就是欧式距离加权函数 $\text{dist}(W, P)$ 中: W —— $S \times R$ 的权值矩阵; P —— $R \times Q$ 的矩阵,表示 Q 个输入(列)向量。假设把RGB看成空间的坐标,我们定义了不同的距离来测量需求数据的每个点和22种颜色的距离。最后用 min 函数寻找最小值,这个最小值就是最短距离。最后可视化出自动匹配的效果,根据RGB信息自动匹配最相近的颜色。

问题二 利用贪心算法思想,在间距最大位置处优先插入第1个所需颜色,根据其对应的R、G、B值的得到第1个增加的瓷砖颜色,通过类比,采用自上而下的迭代计算方法,依次进行贪心计算,直到优先级增加到10个瓷砖颜色,并且备选颜色尽可能均匀地分布在空间中,以最大化颜色覆盖,达到最佳性能效果。

问题三 将第二问得出的结果放入直角坐标系,绘制成折线图,运用标准多目标优化模型可求出相应的最优解。

5 问题的解决

问题一

将颜色放入空间直角坐标系,RGB的值转换为xyz轴上的数值,由此颜色在空间中的位置被确定,通过计算两点之间的距离来确定两种颜色的相似程度。基于欧氏距离计算函数, $\text{dist}(a, b')$ 将所给出的416种颜色取出,分别计算与已有的22种颜色的距离,求出22个距离数值,再使用 $[\text{min_D}, \text{index}] = \text{min}(D, [], 2)$ 函数求出,该距离组中的最小值,运用该算法可以依次求出附件中416种颜色分别对应的最相似瓷砖颜色。

图像1 216种颜色空间分布

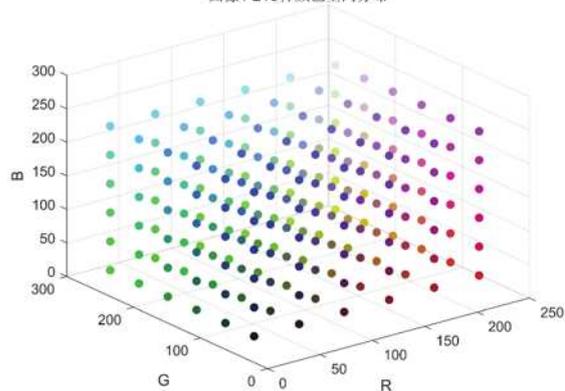


图1 216种颜色空间分布

图像2 200种颜色空间分布

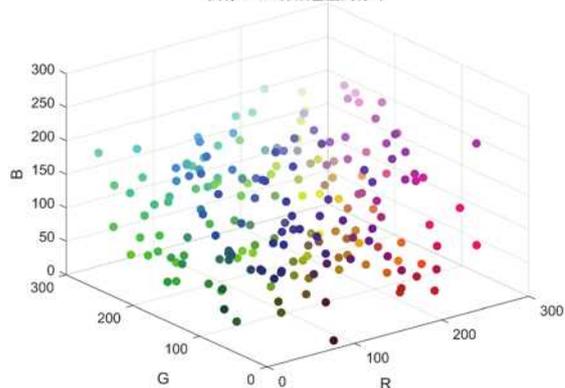


图2 200种颜色空间分布

问题二

因为生活中呈现的颜色比我们的瓷砖多,所以我们要开发新的颜色。这更方便我们寻找相似的颜色,增加图案的表现力。为了使现有的颜色更加均匀和富有表现力,我们在供应方面观察22种现有颜色的分布是否均匀。

在问题1中,我们将标题指定的数据平面化,但我们没有看到颜色的空间分布。只有理解了空间分布,我们才能以这种思维方式添加新的色彩。从空间分布图中我们发现,我们可以看出图1中216种颜色是有规律的等间隔采集数据。

对于这类数据,我们可以假设,只有使可选颜色尽可能均匀地分布在空间中,颜色覆盖才能达到最大,性能效果才会更好。如果要数据添加到原始数据中,则需要对原始数据的分布进行统计研究,以使添加的数据更加均匀。

我们发现已有的二十二种颜色RGB值是不均匀分布的,为了使该颜色更加均匀,我们使用贪心算法,优先在R, G, B图中填补最大的三个空缺的数值,从而确定所要增加的颜色RGB值。所谓贪心算法是指在解决问题

时，它总是做出当前的最佳选择。该算法简单易操作，能快速找到问题的最优解。它适用于解决问题中的颜色填充顺序问题^[3]。

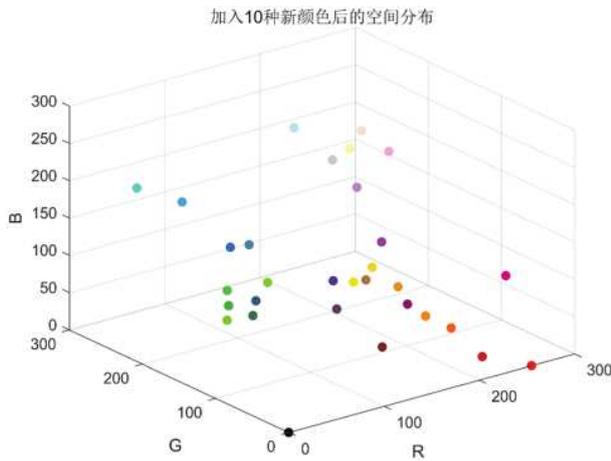


图3 插入10种颜色后空间分布

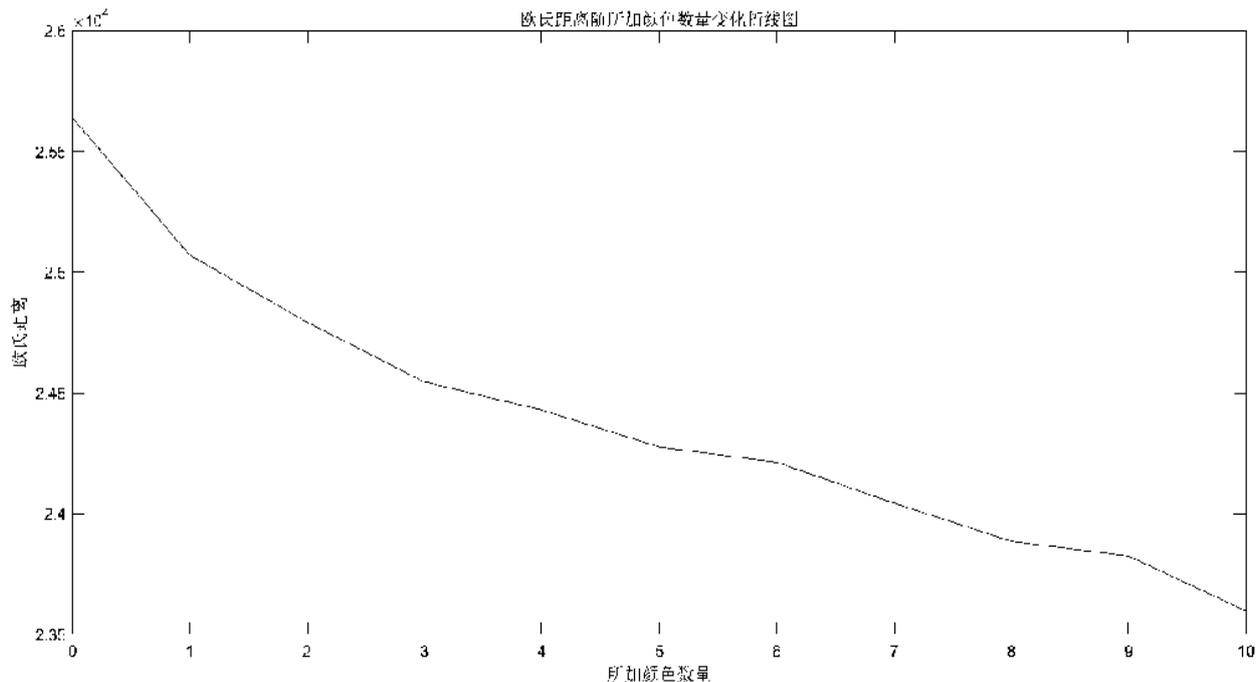


图4 欧式距离随所加颜色数量变化

以该折线图为基础，可以计算出回归方程，当 $x=6$ 时，拟合优度 R^2 最接近1，可以认为增加六种颜色，第三问的最优解，六种颜色的RGB值分别为[147, 29, 107], [162, 84, 160], [132, 44, 48], [210, 14, 20], [40, 97, 118], [105, 71, 96]

6 模型的优缺点

6.1 优点

1.在问题1中，将“寻找图像中每种颜色最接近的已有瓷砖颜色”问题转化为根据颜色的RGB编码值计算

比起原有的图像，该图中颜色分布更趋于均匀。说明我们的思路是正确的，接下来我们依次输出第二问所要求的十种颜色的RGB值[147, 129, 107][162, 84, 160][132, 44, 48][210, 14, 20][40, 97, 118][105, 71, 96][82, 191, 74][62, 135, 168][236, 158, 40][139, 210, 56]

问题三

假设单件瓷砖研发成本为 c 元，随着瓷砖件数的增加，可以假设成本函数为 $y=cx$ ($x=1, 2, \dots, 10$)，当 $x=10$ 的时候，成本最高。

假设每研发一种颜色，已有颜色都将增加一种，运用欧式距离算法公式，可以计算每增加一种新颜色瓷砖时，图1和图2与已有瓷砖颜色的最近距离，我们假设该距离为颜色间的拟合度。可以用回归方程 $\hat{y} = \hat{b}x + \hat{a}$ ($x=1, 2, \dots, 10$)来表示。

其欧式距离最小值问题，将颜色相近转化为距离最短，抽象化为具体。

2.在问题2中，利用贪心算法的思想，将10种新瓷砖颜色对图像整体表现力的综合影响分为几个子问题，有效地减少了10种新颜色之间的交互，使模型更加稳定。

3.本文的模型使用了matlab专用数学分析处理软件，代码简单易懂，过程清楚明白。

6.2 缺点

在问题1中,未能考虑相邻RGB编码值差值的方差而直接采用欧式距离公式。因为方差为0说明为同种颜色,在此前提下,欧式距离越大明暗差别越大。然而单一采用欧式距离公式会使得解决问题片面化,信服度降低。

在问题2中,虽然颜色越多,效果越好,但与添加新颜色相对应的效果改善的边际量通常会发生变化,标题假设成本呈线性增加。因此,同时考虑两个因素,瓷砖颜色的数量不应一直增加^[4]。

6.3改进

1.问题一可增加RGB编码值差的方差和欧式距离公式的权重比例,可设为0.8:0.2,并运用到问题一的代码中。

2.由于题目没有具体说明影响成本的因素,因此可以引入合理的假设.对“表现效果”的处理应与问题二一致或相关。同时,新增颜色的数量没有限制,可以不增加,也可以多于10种.新方案也可以在问题二基础上增

加成本因素来获得^[5]。

3.问题1、2的颜色RGB编码值矩阵数列表、空间分布图、单一编码值间距火柴杆图的绘制可作为解决问题前的准备工作,而不应将其拆开带入各问题中,降低问题间的连续性与逻辑性。

参考文献:

- [1]刘东.建筑材料[J].中国计量出版社,2010.06
- [2]Yvette, Red Dot Award Product Design.玻璃马赛克瓷砖 EbruSeries[J].设计,2016(20):19.
- [3]郭卫斌,杨建国.计算机导论[J].华东理工大学出版社,2012.08
- [4]胡新启,毛紫阳,丁义明,田代平.“马赛克瓷砖选色问题”评阅综述[J].数学建模及其应用,2021.03.09.
- [5]吴妍,吴靖轩,晏丽,向彩容.基于K-means对马赛克瓷砖选色问题的研究[J].中阿科技论坛(中英文),2021(11):111-113.