

《数据可视化》第一次作业

姓名：刘佰川 学号：16307130214

1. 为学生信息设计一个距离度量

两个学生的数据分别表示为 $\langle V_{11}, V_{12}, V_{13}, V_{14} \rangle$, $\langle V_{21}, V_{22}, V_{23}, V_{24} \rangle$ 。

观察数据的形式， V_1 、 V_2 表示学生的身高和体重，是比值型的距离。直接采用曼哈顿距离，距离越大代表相似度越低。 $D1_{ij} = |V_{i1} - V_{j1}|$, $D2_{ij} = |V_{i2} - V_{j2}|$ 。

并对其进行归一化，将范围压缩至 $[0,1]$ 。 $D1_{ij} = D1_{ij} / \max(D1_{ij})$, $D2_{ij} = D2_{ij} / \max(D2_{ij})$ 。

出生地是类别属性，我将出生在同一城市的距离设置为 0，出生在不同城市的距离设置为 1。 $D3_{ij} = (V_{i3} \neq V_{j3})$

第四个变量是年级，从实际情况来看，相近年级的学生的差异应该更小，即距离函数最小。一到五年级分别由数字 1-5 进行表示，差异最大的情况就是一年级和五年级的差异，可以用数字 $5-1=4$ 来进行表示。类似于身高体重的曼哈顿的距离， $D4_{ij} = |V_{i4} - V_{j4}| / 4$ 。

因为四个特征的数据都在 $[0,1]$ 之间，所以直接选取平均值作为两个学生之间的距离度量。 $D = (D1_{ij} + D2_{ij} + D3_{ij} + D4_{ij}) / 4$ 。

2. 选择实现 RGB 颜色模式和 HSB 颜色模式的相互转化。

自然界中所有的颜色都可以用红、绿、蓝(RGB)这三种颜色波长的不同强度组合而得。因此，这三种光常被人们称为三基色或三原色。对 RGB 三基色各进行 8 位编码就构成了大约 16.7 万种颜色。

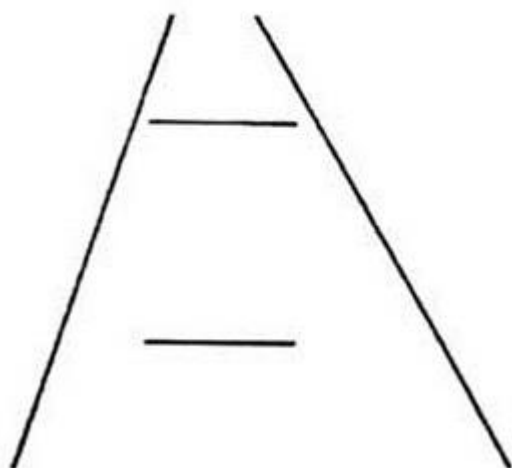
HSB(HSV) 通过色相/饱和度/亮度三要素来表达颜色。

H(Hue):表示颜色的类型(例如红色,绿色或者黄色).取值范围为 0—360.其中每一个值代表一种颜色。

S(Saturation):颜色的饱和度.从 0 到 1.有时候也称为纯度.(0 表示灰度图,1 表示纯的颜色)

B(Brightness or Value):颜色的明亮程度.从 0 到 1.(0 表示黑色,1 表示特定饱和度的颜色)

转化代码见附件。



3.

视错觉原理：两条斜线会让大脑认为横线有远近，远处的跟近处的看起来一样长的话，远处的更长



托兰斯肯弯曲幻觉：这三个圆弧看起来弯曲程度差别很大，但实际它们是完全一样，只是下面两个比上面那个短一些。视觉神经末梢最开始只是按照短线段解释世界。当线段的相关位置在一个更大的空间范围延伸概括后，弯曲才被感知到。所以如果给定的是一条曲线的一小部分，视觉系统往往不能察觉它是曲线。

4. 见代码附件

在快排上利用 n 降低复杂度。(递归进行运算时超过 n 值的部分则不再计算)

测试数据

[3,1,2,4,2,5,4]

N 值	输出结果
1	1
2	1,2

3	1,2,2
4	1,2,2,3
5	1,2,2,3,4
6	1,2,2,3,4,4
7	1,2,2,3,4,4,5

[7,6,4,5,1,2,3,8]

N 值	输出结果
1	1
2	1,2
3	1,2,3
4	1,2,3,4
5	1,2,3,4,5
6	1,2,3,4,5,6
7	1,2,3,4,5,6,7