

# Continuous Observation and Identification of Weather Phenomena via KNN

Sizhao Liu, Guanlei Xu

Dalian Navy Academy, Dalian Liaoning  
Email: 278574378@qq.com

Received: Jan. 3<sup>rd</sup>, 2018; accepted: Jan. 17<sup>th</sup>, 2018; published: Jan. 24<sup>th</sup>, 2018

---

## Abstract

Multi-level weather classification is the basis and important technology of many potential applications such as video surveillance and intelligent transportation. Most of the existing weather classification methods only consider weather conditions that are sunny or sunny cloudy and mainly concentrate in a fixed scene, such as mass tourism and traffic scenarios. Based on the HSV image color model, this paper mainly analyzes and studies the brightness of each weather image, comparing the brightness, and the average and variance of its brightness value. Using the nearest neighbor method, the selected images are classified and identified, so as to automatically identify various types of weather. The experimental results confirm the feasibility of this method.

## Keywords

Weather Classification, HSV Image Color Model, The Nearest Neighbor Method, Image Recognition

---

# 一种基于KNN的天气现象连续观测识别方法

刘泗照, 徐冠雷

大连舰艇学院, 辽宁 大连  
Email: 278574378@qq.com

收稿日期: 2018年1月3日; 录用日期: 2018年1月17日; 发布日期: 2018年1月24日

---

## 摘 要

多级天气分类是视频监控和智能交通等诸多潜在应用的基础和重要技术。大多数现有的天气分类方法只考虑天气晴朗或晴天多云的天气条件, 并且主要集中在一个固定的场景, 如大众旅游和交通情景。本文基于HSV图像色彩模型, 主要对各天气现象图片的亮度进行分析和研究, 进行亮度间的对比, 比较其亮

度值的平均值与方差, 采用最近邻方法, 对所选图像进行分类与识别, 达到自动识别各类天气的目的。实验结果证实了此研究方法的可行性。

## 关键词

天气分类, HSV图像色彩模型, 最近邻方法, 图像识别

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着互联网技术与数字化的飞速发展, 以及电子数码产品的普及, 人们能够获取的数字图像数据已经越来越多[1]。图像作为一种表现直观、内容丰富的多媒体信息, 在各行各业中也得到越来越广泛的应用, 如数字多媒体图书馆、医学图像应用管理、卫星遥感图像等[2]。目前, 图像处理方法的重要性源于两个主要应用领域: 改善图示信息以便人们解释; 为存储、传输和表示而对图像数据进行处理, 以便于机器自动理解。

图像特征提取是自动识别图像的基础。获得高效的图像不仅可以降低计算复杂度, 还可以准确地表述图像的原始信息[3]。目前, 常用的图像底层特征包括颜色、纹理和形状等。此外, 新颖合适的图像特征提取方法对别的领域如人工智能、模式识别等也会有一定的促进作用。因此, 如何有效的提取低层视觉特征成为基于内容的图像识别中很关键的一个步骤, 也成为图像识别领域研究的热点内容。

图像理解与智能视频技术的不断发展, 对室外环境下各种天气现象, 尤其是恶劣天气的现象的检测与识别提出了极为迫切的需求, 同时在气象领域, 对于天气现象的观测仍然依靠人工观测, 天气现象信息的采集耗费了大量的人力物力, 近几年来基于室外图像与视频数据的天气现象自动识别研究受到了较为广泛的关注[4]。

目前, 计算机视觉领域中的大多数现有方法都是基于户外图像或视频中的天气明确的假设[5]。然而, 不同的天气条件如雨, 雪或烟雾会降低图像或视频的质量。这种影响可能会显著降低依赖于图像/视频特征提取或视觉注意建模的户外视觉系统的性能。因此, 天气分类的应用很多, 如天气条件的检测和观察, 图像/视频分析, 视频监控系统的可靠性改进。

本文基于光学图像处理技术, 对晴阴雨雪气象的天气特征进行了提取并研究, 为天气现象的自动观测提供光学技术、方法和手段。

## 2. HSV 空间模型

HSV 模型在 1978 年由埃尔维·雷·史密斯创立。

HSV 色彩模型从 CIE 三维颜色空间演变而来, 它采用的是用户直观的色彩描述方法, 它跟孟塞尔显色系统的 HVC 球型色立体较接近[6]。只不过 HSV 色彩模型是一个倒立的六菱锥, 只相当于孟塞尔球型色立体的一半(南半球), 所以不含黑色的纯净颜色都处于六菱锥顶面的一个色平面上[7]。在 HSV 六菱锥色彩模型(如图 1)中, 色相(H)处于平行于六菱锥顶面的色平面上, 它们围绕中心轴 V 旋转和变化, 红、黄、绿、青、蓝、品红六个标准色分别相隔 60 度。色彩明度(B)沿六菱锥中心轴 V 从上至下变化, 中心轴顶端呈白色(V = 1), 底端呈黑色(V = 0), 它们表示无彩色系的灰度颜色。色彩饱和度(S)沿水平方向变

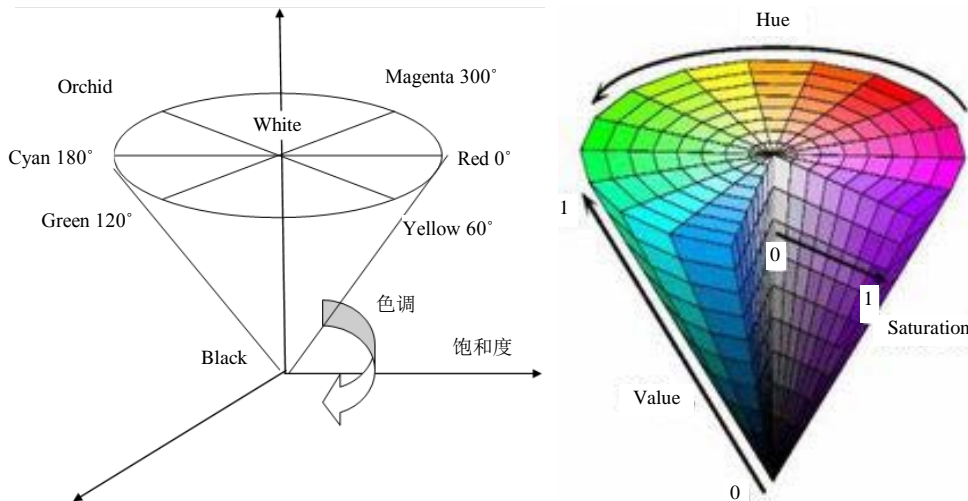


Figure 1. Hexagonal cone color model diagram  
图 1. 六菱锥色彩模型图

化, 越接近六菱锥中心轴的色彩, 其饱和度越低, 六边形正中心的色彩饱和度为零( $S = 0$ ), 与最高明度的  $V = 1$  相重合, 最高饱和度的颜色则处于六边形外框的边缘线上( $S = 1$ ) [8]。

HSV 模型通常用于计算机图形应用中。在用户必须选择一个颜色应用于特定图形元素各种应用环境中时, 经常使用 HSV 色轮(如图 2)。在其中, 色相表示为圆环; 可以使用一个独立的三角形来表示饱和度和明度。典型的, 这个三角形的垂直轴指示饱和度, 而水平轴表示明度。在这种方式下, 选择颜色可以首先在圆环中选择色相, 在从三角形中选择想要的饱和度和明度。

HSV 模型的另一种可视方法是圆锥体(如图 3)。在这种表示中, 色相被表示为绕圆锥中心轴的角度, 饱和度被表示为从圆锥的横截面的圆心到这个点的距离, 明度被表示为从圆锥的横截面的圆心到顶点的距离。某些表示使用了六棱锥体。这种方法更适合在一个单一物体中展示这个 HSV 色彩空间; 但是由于它的三维本质, 它不适合在二维计算机界面中选择颜色[9]。

HSV 色彩空间还可以表示为类似于上述圆锥体的圆柱体, 色相沿着圆柱体的外圆周变化, 饱和度沿着从横截面的圆心的距离变化, 明度沿着横截面到底面和顶面的距离而变化[10]。这种表示可能被认为是 HSV 色彩空间的更精确的数学模型; 但是在实际中可区分出的饱和度和色相的级别数目随着明度接近黑色而减少。此外计算机典型的用有限精度范围来存储 RGB 值; 这约束了精度, 再加上人类颜色感知的限制, 使圆锥体表示在多数情况下更实用[11]。

### 3. K-邻近分类算法

K-邻近分类算法可以解释为有那么一堆已经知道分类的数据, 然后当一个新数据进入的时候, 就开始跟训练数据里的每个点求距离, 然后挑选这个训练数据最近的 K 个点看看这几个点属于什么类型, 然后用少数服从多数的原则, 给新数据归类(如图 4) [12]。

最近邻分类的解释是: 令  $D^n = \{x_1, x_2, \dots, x_n\}$ , 其中每一个样本  $x_i$  所属的类别均已标记。对于参数样本点  $x$ , 在集合  $D^n$  中距离它最近的点记为  $x'$  那么最近邻分类规则就是把点  $x$  分为  $x'$  所属的类别[13]。

## 4. 基于 HSV 空间模型和 K-邻近分类算法实验

### 4.1. 图像 HSV 模型的建立

在图像分类的过程中, 我们之所以选择提取图像的主要特征, 是因为直接处理原始图像得到的向量

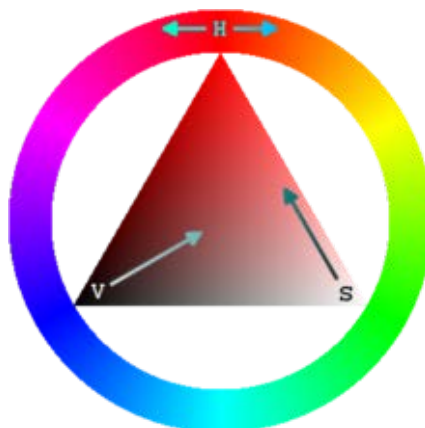


Figure 2. Color wheel diagram of HSV  
图 2. HSV 色轮

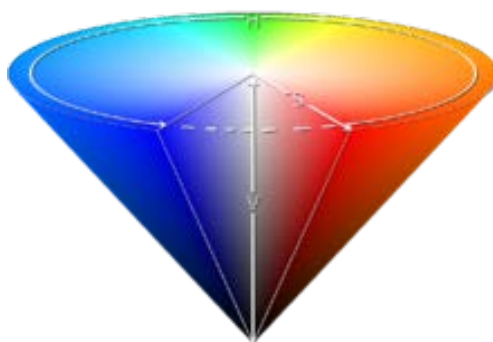


Figure 3. Cone of HSV  
图 3. HSV 圆锥体

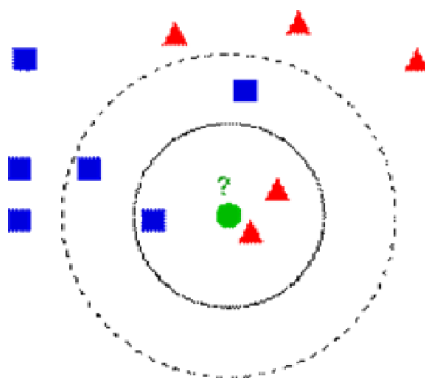
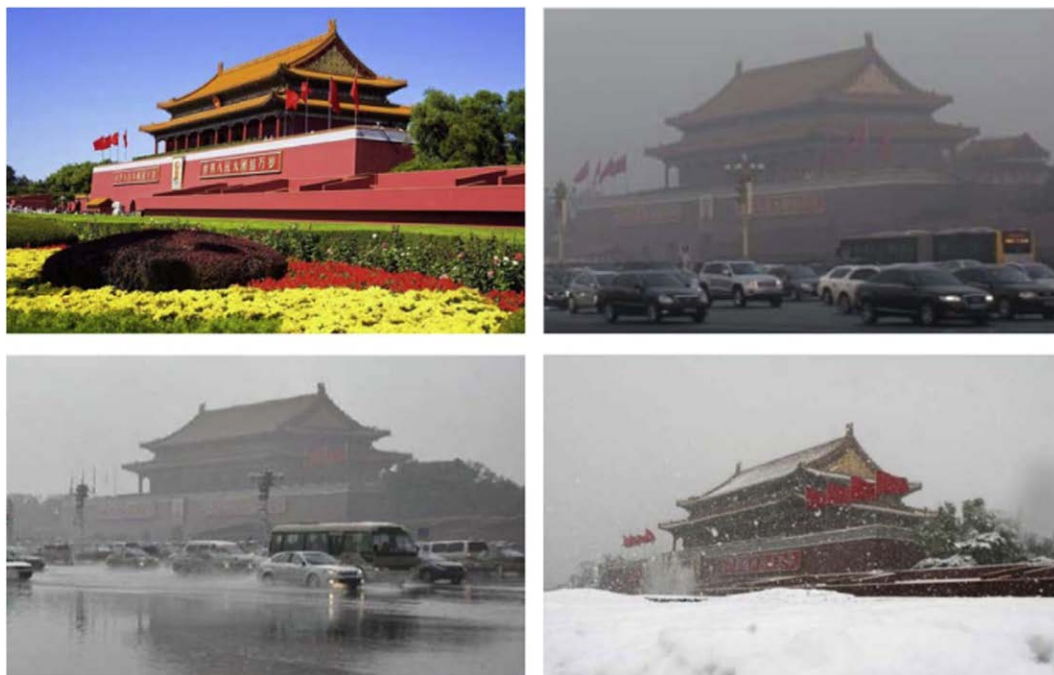


Figure 4. K-nearest neighbor classification algorithm diagram  
图 4. K-邻近分类算法示意图

维度不易分类。因此，提取图像的主要特征就成为了图像分类的关键。实际来讲，HSV 分类是从两类问题的分类所提出的，但是现在我们所要做的工作是对多类图像进行分类，这就需要将两类的问题扩展到多类[14]。本文要通过三种特征将图 5 中四种天气图像分类。

表 1 反应了四种天气 H、S、V 三个通道的特征数据情况，其中  $H_{avg}$  表示色调均值； $H_{var}$  表示色调标准差； $S_{avg}$  表示饱和度均值； $S_{var}$  表示饱和度标准差； $V_{avg}$  表示亮度均值； $V_{var}$  表示亮度标



**Figure 5.** Four kinds of weather including sunny, cloudy, rainy and snowy  
**图 5.** 晴、阴、雨、雪四种天气图

**Table 1.** Classification results of HSV  
**表 1.** HSV 分类结果

序号	类别	H_avg	H_var	S_avg	S_var	V_avg	V_var
1	晴	0.5330	0.1042	0.4828	0.2188	0.7977	0.1319
2	阴	0.4242	0.2128	0.1711	0.1348	0.5245	0.2126
3	雨	0.3863	0.1891	0.3672	0.1634	0.4011	0.2082
4	雪	0.4258	0.1866	0.1082	0.0970	0.7401	0.1786

准差。本文共采用了 800 张天气图像, 其中晴、阴、雨、雪四种天气中每类天气图像各 200 张用来提取图像特征, 建立分类器, 即四种天气已知样本库, 分类结果如表 1。

#### 4.2. $k = 1$ 与 $k = 3$ 时邻近法的比较

在 4.1 建立四种已知天气样本库的基础上, 选取晴、雨、阴、雪四种天气图像各 200 张作为待识别图像样本库, 经图像识别,  $k = 1$  时分类结果如表 2,  $k = 3$  时分类结果如表 3。

从实验结果来看, 当  $k = 1$  时晴天的正确识别率最高, 到达 85%, 雨天的正确识别率最低, 为 65%。 $k = 3$  时晴天的正确识别率最高, 到达 80%, 雨天的正确识别率最低, 为 62%。可见  $k = 1$  比  $k = 3$  时的天气正确识别率要高。但是, 从人们视觉的直观感受来看, 阴天图像与雪天图像整体上都呈现白色, 两种天气的亮度和饱和度大致相同, 因此, 无论是  $k = 1$  的情形还是  $k = 3$  的情形都存在将阴天天气错判为雪天天气的情况。

## 5. 结论

基于光学图像的天气识别技术是当前研究比较热门的领域, 涉及了计算机视觉、模式识别、图像处

**Table 2.** Weather recognition rate when k is 1  
**表 2.** k = 1 时的天气识别率

天气类别	样本数	识别准确数	准确率
晴	200	170	85%
阴	200	146	73%
雨	200	130	65%
雪	200	142	71%

**Table 3.** Weather recognition rate when k is 3  
**表 3.** k = 3 时的天气识别率

天气类别	样本数	识别准确数	准确率
晴	200	160	80%
阴	200	140	70%
雨	200	124	62%
雪	200	136	68%

理、分类聚类分析等众多方面的技术[15]。基于 HSV 空间模型和 K-邻近分类算法识别是本文的主要研究内容:

- 1) 利用 HSV 色彩分析法并结合近邻方法对图像进行比较归类, 较为准确的对图像进行了自动识别, 达到了预期的目的;
- 2) 在特征提取后, 我利用 K-NN 对待识别检测图像进行了识别检测, 检测正确率普遍为 70%左右, 准确度有待进一步提高;
- 3) 图像识别技术作为当前比较热门的专业, 应该有很多的研究方法, 因此可进一步拓宽研究渠道, 进而增加图像识别的准确率。

该研究成果可以应用于天气条件的检测和观察, 图像/视频的分析, 视频监控系统的可靠性改进等方面。

## 参考文献 (References)

- [1] 布洛基. 智能车辆[M]. 北京: 人民交通出版社, 2002.
- [2] 郑南宁. 计算机视觉与模式识别[M]. 北京: 国防工业出版社, 1998.
- [3] 尼克松, 阿瓜多, 李实英, 杨高波. 特征提取与图像处理[M]. 北京: 电子工业出版社, 2010.
- [4] 穆尔. MATLAB 实用教程[M]. 北京: 电子工业出版社, 2010.
- [5] Gonzalez, R.C., Woods, R.E., 阮秋琦, 阮宇智. 数字图像处理[M]. 第 3 版. 北京: 电子工业出版社, 2001.
- [6] 张彩华. 基于颜色和纹理特征的图像检索[D]: [硕士学位论文]. 哈尔滨: 哈尔滨理工大学, 2010.
- [7] 顾宁. 基于颜色和纹理特征的图像检索算法研究[D]: [硕士学位论文]. 南京: 南京邮电大学, 2011.
- [8] 姚华静. 基于颜色和纹理特征的图像检索技术研究[D]: [硕士学位论文]. 南京: 河海大学, 2007.
- [9] 王惠明, 史萍. 图像纹理特征的提取方法[J]. 中国传媒大学学报(自然科学版), 2006, 13(1): 49-52.
- [10] 谢菲. 图像纹理特征的提取和图像分类系统研究及实现[D]: [硕士学位论文]. 成都: 电子科技大学, 2009.
- [11] 程涛. 基于颜色和形状特征的图像检索[D]: [硕士学位论文]. 西安: 西北大学, 2010.
- [12] 肖靛. 基于支持向量机的图像分类研究[D]: [硕士学位论文]. 上海: 同济大学, 2006.

- [13] 刘晓勇, 付辉. 一种快速 AP 聚类算法[J]. 山东大学学报(工学版), 2011, 41(4): 20-23.
- [14] 赵莹, 高隽, 汪荣贵, 胡静. 一种新的广义最近邻方法研究[J]. 电子学报, 2004, 32(s1): 196-198.
- [15] 钱丽丽, 施鹏飞. 近邻传播算法在非监督图像聚类中的应用[J]. 微型电脑应用, 2011, 27(2): 5+40-42.

**知网检索的两种方式:**

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>  
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2325-6753, 即可查询
2. 打开知网首页 <http://cnki.net/>  
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>  
期刊邮箱: [jisp@hanspub.org](mailto:jisp@hanspub.org)