

天津大学

硕士学位论文

颜色直方图识别新技术研究

姓名： 娄强

申请学位级别： 硕士

专业： 生物医学工程

指导教师： 肖松山

20070101

# 中文摘要

随着多媒体技术、计算机技术、通信技术及 Internet 网络的迅速发展,人们越来越多地接触到各种各样的图像信息,图像中的物体特性有很多,其中颜色是最为重要的特征之一。常用的颜色特征提取技术是颜色直方图,直方图是在研究颜色特征分析时最为重要的方法之一,直方图处理除了计算简单的特性外,同时还具有尺度不变性和旋转不变性的优点,但是直方图无法得到颜色组成之间的空间关系信息而且对光照变化敏感。本文主要对如下几个方面进行了分析:

1. 分析和研究了基于颜色直方图识别领域的一些关键技术,其内容包括颜色空间的选取、颜色直方图的描述方法比较、颜色直方图特征的相似性度量准则。
2. 对于颜色直方图特征的提取,本文主要采取的方式是根据 Shafer 二色反射模型,通过把物体表面反射的过程分为两种,即表面反射和本体反射。表面反射分量一般与入射光有相同的光谱分布,本体反射分量的光谱则是有光照的光谱功率分布和物体表面反射特性共同决定的。本体反射率代表了物体的颜色反射特性,可以表征物体颜色的本征特性。对基于颜色比率的颜色不变量模型进行了改进,通过新构建的颜色模型,去除掉表面反射分量,得到本体反射的分量,从而可以得到无关于光照和几何特性无关的本体反射率。即仅与物体颜色反射特性相关的不变量信息  $q_1$ 、 $q_2$ 、 $q_3$ ,通过与其他颜色模型(如  $r$ 、 $g$ 、 $b$  模型 HIS 模型等)的对比实验结果的分析,可以得知本文所提出的色彩模型在不同的光源照射下以及不同光照强度下的稳定性有一定的提高。
3. 由于颜色直方图反映的是颜色的全局统计信息,丢失了颜色的空间分布信息。为了在基于颜色直方图的方法中增加空间信息,本文提出应用图像分块以及配置权重并结合模糊颜色直方图进行处理,符合人体的视觉规律,能够在一定程度上模拟人的视觉过程,通过实验得到较好的结果。

本文针对直方图技术的缺点,提出并实施了基于二色反射模型和模糊颜色直方图的新思路,提高了光照稳定性并增加了空间信息,更接近模拟人眼视觉图像处理过程,对彩色图像处理技术发展提供了新途径。

**关键词:** 颜色直方图 光照不变量 颜色模型 空间直方图

# ABSTRACT

Along with the rapid development of multimedia technology, computer technology, communication and Internet network, the people make more and more contact with various image informations, Color in the image information is the best characteristic for display the object characteristic, the color histogram is one of most important methods for color characteristic analysis, besides the simple computation characteristic, the method of color histogram processing also has the advantages of scale invariance and rotate invariance, but the histogram is unable to obtain spatial relations information of the color composition. This article mainly has carried on the analysis for the following several aspects:

1. Some key technologies are analyzed and studied based on color histogram recognition domain, including the selection of color space, the comparison of color histogram description method, the measuring criterion for the similar characteristic of color histogram.
2. Regarding the extraction of color histogram characteristic, the Shafer reflection model is adopted; that is dividing the surface reflection process from object into two kinds, namely the superficial reflection and the body reflection. The superficial reflection component generally has the same spectrum distribution with the incident light, the body reflects the component the spectrum to have the illumination spectrum power distribution and the object surface echoing characteristic decides together. The spectral component of the body reflection is decided by the power distribution of incident light and reflection characteristic of the object surface together. The body reflection has represented the object color echoing characteristic, may attribute the object color the Eigen characteristic. The model improvement based on the color ratio color invariant is made through the new construction color model, the superficial reflection component is canceled, the component of the body reflection only obtained, thus body index of reflection irrelevant to illumination and the geometry characteristic may obtained. Namely invariant information  $q_1$ ,  $q_2$ ,  $q_3$  correlated only with object color echoing characteristic, through the experiment result analysis compared with other color model (for example  $r$ ,  $g$ ,  $b$  model, HIS model and so on), the color model in this paper proposed has certain stability enhancement under the

different light source as well as the different illumination strength.

3. The color histogram reflection is the color statistic information overall, and the color spatial distribution information is lost. In order to increase the spatial information based on the color histogram method, this article proposed to process the fuzzy color histogram using the image piecemeal as well as the disposition weight, conforms to the human body visual rule, it can simulate the human's visual process to a certain extent.

In view of the color histogram technology shortcoming, this paper proposed and implemented a new color-recognition method based on the Shafer reflection model and the fuzzy color histogram. Based on this color histogram method, the illumination stability is enhanced and the spatial information is increased. It can be used to simulate human's visual process, and has provided a new way to the color image processing technological development.

**KEY WORDS:** Color Histogram, Illumination invariant, Color model, Spatial Histogram ;

## 独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作和取得的研究成果，除了文中特别加以标注和致谢之处外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得 天津大学 或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

学位论文作者签名： 娄强 签字日期： 2007 年 2 月 6 日

## 学位论文版权使用授权书

本学位论文作者完全了解 天津大学 有关保留、使用学位论文的规定。特授权 天津大学 可以将学位论文的全部或部分内容编入有关数据库进行检索，并采用影印、缩印或扫描等复制手段保存、汇编以供查阅和借阅。同意学校向国家有关部门或机构送交论文的复印件和磁盘。

（保密的学位论文在解密后适用本授权说明）

学位论文作者签名：

娄强

导师签名：

肖松心

签字日期：2007 年 2 月 6 日

签字日期：2007 年 2 月 6 日

## 第一章 绪论

### 1.1 研究的背景和意义

自然世界是由不同的物体组成的,组成自然界的物体特性有很多,其中颜色是最为重要的特征之一,是构成现实世界的一个基本要素。人类的知觉系统对外部世界的感知是极大的依赖物体的颜色特性,对颜色的感知是人类知觉系统来认识世界的一个重要的因素。这一点已经成为认知科学和计算机科学领域的科研工作者的普遍共识。

随着计算机认知科学和人工智能的发展,人们要求计算机能够模拟人类的知觉系统对外部世界的感知能力,也就是要求计算机能够“认识”世界。计算机和图像处理等分支也就由此发展起来。颜色分析则是计算机视觉和数字图像处理中的一个重要的研究课题。

从计算机视觉的产生开始,颜色分析就引起了科研工作者的注意,颜色特征是应用最为广泛的视觉特征,主要原因在于颜色往往和图像中所包含物体或场景十分相关。此外,与其他的视觉特征相比,颜色特征对图像本身的尺寸、方向、视角依赖性较小,从而具有较高的鲁棒性。颜色特征的表达涉及到若干问题。首先,我们需要选择合适的颜色空间来描述颜色特征;其次,我们要采用一定的量化方法将颜色特征表达为向量的形式;最后,还要定义一相似度标准用来衡量图像之间在颜色上的相似性。直方图是在研究颜色特征分析时最为重要的方法之一,直方图颜色分析具有尺度、旋转不变性,但是直方图会丢失图像的空间信息,而且对光照变化也很敏感。

近年来,随着多媒体技术和计算机网络的飞速发展,全世界的数字图像的容量正以惊人的速度增长。然而,由于这些图像是无序地分布在世界各地,图像中包含的信息无法被有效地访问和利用。这就要求有一种能够快速而且准确地查找访问图像的技术,也就是所谓的图像检索技术。自从20世纪70年代以来,在数据库系统和计算机视觉两大研究领域的共同推动下,图像检索技术已逐渐成为一个非常活跃的研究领域。90年代初期,随着大规模数字图像库的出现,上述的问题变得越来越尖锐。为克服这些问题,基于内容的图像检索技术(content-based image retrieval)应运而生。区别于原有系统对图像进行人工标注的做法,基于内容的检索技术自动提取每幅图像的视觉内容特征作为其索引,如色彩、纹理、形状等。此后几年中,这个研究领域中的许多技术发展起来,一大批研究性的或

商用的图像检索系统被建立起来。这个领域的发展主要来归功于计算机视觉技术的进步。而颜色特征是在图像检索中应用最为广泛的视觉特征，主要原因在于颜色往往和图像中所包含的物体或场景十分相关。迄今为止，对颜色的研究已经有了近 30 年的历史，在这期间，提出了很多颜色分析方法以及很多应用于不同领域的颜色空间。可以应用在目标识别、遥感图像分析、医学图像处理以及图像检索等很多方面。其中颜色直方图是最为重要而且应用最为广泛的一种方法，但是在颜色直方图分析过程中还是有很多亟待要解决的问题，如颜色模型的选取，彩色图像的识别分割，以及环境光线的干扰等。

## 1.2 研究现状和方法

在颜色视觉的研究中，颜色恒常性具有相当的重要地位。所谓颜色恒常性，即当光照条件变化时人感知的物体颜色在一定程度上看来却保持不变。在颜色物体识别的应用中，实现具有颜色恒常特性的算法是十分重要的，否则当光照条件发生变化时算法的性能将大大下降，从而不能保证系统的实用性。

80 年代以来，随着物理模型的应用，颜色视觉研究得到了充分的发展。由于有限维线性模型的应用，颜色视觉研究得到了充分的发展。由于有限维线性模型的应用，使得光照光谱分布和物体表面发射率函数的表达式和计算得大大简化，提高了颜色计算的效率。而随着双色反射模型的提出，使得从双极性物体高光中提取光照信息成为可能，为恢复物体的本体颜色铺平了道路。但这些颜色恒常性算法大多都是通过对高光图像的分析来获得颜色恒常性，尽管具有较强的理论研究意义，在实际应用中难以有效应用。

1991 年，Novak<sup>[1]</sup>在对以往的算法分析的基础上，提出了监督颜色恒常性算法。其基本思想是：在视场中放置已知光谱反射特性的颜色片，通过颜色成像系统对这些颜色片的响应来恢复光照信息，进而获取颜色恒常性。这种方法，相对于其他颜色恒常性算法来说，具有简单、方便有效的特性，因而在颜色物体识别等研究中得到了广泛的应用，但其所求得颜色成像系统通道响应特性在许多情况下难于获取。

常用的颜色特征提取技术是颜色直方图，直方图是在研究颜色特征分析时最为重要的方法之一，它所描述的是不同颜色在整幅图像中所占的比例，而并不关心每种颜色所处的空间位置，即无法描述图像中的对象或者物体。直方图处理除了计算简单的特性外，同时还具有对平移和旋转不变性的优点，但是直方图无法得到颜色组成之间的空间关系信息，同时，全局颜色直方图只记录了颜色的全局统计信息，丢失了颜色的空间分布信息。为此，许多新的颜色特征提取方法被

提了出来。Stncker等和刘忠伟<sup>[2]</sup>等提出采用累加直方图和局部累加直方图来进一步增强直方图特征的提取效果；金韬<sup>[3]</sup>等利用一种颜色常量方法来减少光照度变化对颜色的影响；刘建峰<sup>[4]</sup>等对彩色空间进行K-L变换，从而将图像空间变换为正交空间，使得颜色分量在统计上具有不相关性；李向阳等对颜色的聚类进行了研究；王伟<sup>[5]</sup>等提出了构造模糊直方图的方法，等等。

基本颜色直方图虽然简单易用，但也存在一些缺点。它是对像素颜色在不同颜色区段分布的全局统计，对图像内容的反映能力一定程度上依赖具体的颜色量化方法以及颜色种类多寡。大多数自然拍摄的图像直方图结果都比较稀疏，不能很好体现统计特性；基于基本直方图的特征对量化边界很敏感，无论如何分割，都存在跨越边界附近的近似颜色像素被划入不同颜色段的可能，由此造成的映射扭曲影响并限制了直方图的应用。针对不足人们提出了很多改进，如颜色对方法、基于区域划分配不同视觉比重的直方图计算，基于交集计算的直方图距离测度等。这些方法对基本颜色直方图效果有一定改善，但并未有效解决基本颜色直方图中刚性区间划分和离散分类和人类对相似颜色的连续性视觉和心理认知不匹配造成的问题。人类对图像的认知客观上取决于画面中各种显著颜色比率及在图像中的分布。同时人对颜色差别的认知基于连续变化，对近似颜色不敏感。针对这些问题，本文提出一种基于模糊分类的改良直方图用于近似颜色匹配。实验表明，该方法与前述各种改进方法结合，一定程度上克服了上述缺点，能更好地匹配人类视觉认知特性，从而有效解决问题

### 1.3 本文的主要贡献和组织结构

第二章给出了颜色特征识别的概述，颜色空间的定义及分类，以及颜色直方图的基本定义。

第三章对于颜色直方图特征的提取，本文根据Shafer二色反射模型<sup>[6]</sup>，把物体表面反射的过程分表面反射和本体反射。本体反射率代表了物体的颜色反射特性，可以表征物体颜色的本征特性。通过颜色特征的直方图分析得到无关于光照和几何特性的本体反射率。通过对试验结果的分析，依据本文所提出的颜色模型得到的颜色直方图在不同光照角度和强度下的稳定性有一定的提高。

第四章对颜色直方图的不足进行了改进，直方图是在研究颜色特征分析时最为重要的方法之一，直方图颜色分析具有尺度、旋转不变性，但是直方图会丢失图像的空间信息，通过对空间颜色直方图不同构造方法的分析，应用模糊直方图和图像分块加入图像的空间信息，模拟人体的视觉机理，从而提高了物体识别率。

第五章总结了本文的主要工作和贡献，并对以后的工作提出了建议。



## 第二章 颜色直方图特征识别原理

### 2.1 颜色的视觉机制

颜色<sup>[7]</sup>是光作用于人眼视觉而产生,除了目标外观(形状、大小、位置、质地)外的视觉特性,包括区分相同结构、大小、形状的两个视场角之间差异的视觉现象或辐射特性、形成上述感受的光刺激特性以及能产生光刺激的物体特性等三个重要内容。颜色理论是20世纪初叶以物理光学、视觉生理学、生理学、心理学和工业测量技术为基础发展起来的边缘性、交叉性理论,颜色理论的发展,极大促进了颜色科学与技术的应用。

### 2.2 颜色及其视觉特性

#### 2.2.1 颜色的含义

颜色作为一种光对视觉系统的物理——心理反应,反映了光与视觉相互作用的规律。目前,客观世界颜色的呈色模式有光源色、反射色和透射色等。在实际应用中,可以分为:孔色、表面色、透膜色、透体色、镜面色、光泽、光源色等类型<sup>[8]</sup>。其中:

(1)孔色(Aperture color)是指眼睛焦点在小孔处,通过小孔看到的屏后边的颜色。观察孔色时,只能够看到颜色,而没有孔后物体的大小、形状、位置和质地的知觉,是最单纯的颜色,也叫心理物理色或感觉色。

(2)表面色(Surface color)是指观察不透明物体表面的颜色,是最常见的颜色表面色受照明光照条件和物体表面反射特性的共同影响。

(3)透膜色(Transparent film color)是指透过一层彩色薄膜看到的颜色。透膜色受照明光照条件和膜片透射特性的共同影响。

(4)透体色(Transparent volume color)是指透过一块稍有吸收和散射的透明物体(如冰、薄雾)看到的颜色。

(5)镜面色(Mirrored color)是指在平面反射镜中看到的物体颜色。

(6)光泽(Luster)是指物体表面极小部分强烈反射光而使其细微结构看不清的状态。

(7)光源色(Light source color)是指自发光体的颜色。

### 2.2.2 颜色的视觉特性

颜色是由彩色如红、橙、黄、绿、青、蓝、紫与中性色(非彩色)如黑、白、灰共同组成。其中彩色具有色相、饱和度和明度(亮度)三个基本属性,而中性色只有明度(亮度)属性。在一定照明和观察条件下,能够基于加色法和减色法原理来组合与匹配颜色,而且颜色相同的光,不论其光谱分布是否相同,在加色法混合中都能产生相同的颜色效果。

颜色视觉作为一种复杂的物理、生理与心理现象,是不同波长的光作用于视网膜后在人脑引起的主观感觉。有关资料证明<sup>[9]</sup>,人眼不仅能够识别对应于可见光光谱区一定波长的红、橙、黄、绿、青、蓝、紫等七种主要颜色,而且还可以识别颜色3-5nm波长增减的变化,最多可以分辨150余种不同的色相。显然,视网膜上存在上百种对不同波长光波起反应的视锥细胞或感光色素是不可能的。因此,人们一直在探索颜色视觉的特性。牛顿早期从物理学分析了颜色视觉,获得了某种颜色不仅可以由某种固定波长引起,而且可以由两种或更多种其它波长的混合作用引起。如著名的牛顿色环盘<sup>[10]</sup>的旋转,使光谱上七色光形成人眼视觉的白色感觉。采用红、绿、蓝三色光(非色料)通过合适的混合,就能够产生光谱上任何颜色的感觉。这种色光混合现象,表达了颜色视觉产生原理的基本规律,并已经广泛应用于彩色印刷、彩色摄影,彩色电视机,显示器等工业领域。

### 2.2.3 颜色视觉的机制

目前,颜色视觉机制理论的研究成果主要有T.Young-Helmhotz(1809年)的三原色理论,E.Hering(1876年)<sup>[11]</sup>的四色理论和P.L.walraven(1962年)<sup>[12]</sup>的阶段理论。其中在彩色复制工业领域中,以三色理论的广泛应用为代表。在颜色心理学领域中,以四色理论为代表。而阶段理论在日益增多的试验证实的条件下,融合了三色理论和四色理论的各自优势,已经成为人们关注的理论热点。

#### (1) T.Young-Helmhotz的三原色理论<sup>[13]</sup>

1809年T.Young根据牛顿的粒子打击振动论、配色实验及色散试验,首次提出了三原色理论。他认为人眼的视网膜上只有红(R)绿(G)蓝(B)三种基本锥体细胞(光接收器),每种锥体细胞对各种不同频率的光都有相应的响应。但只对某确定波长的光最为敏感。当三种锥体细胞在同时同等程度地被刺激时,则产生白色感觉。通过实验证明,采用红、绿、蓝三种光谱色的色光混合,能够产生客观世界中各种颜色感受。T.Young的三种接收器理论虽然能很好解释了色光混合规律,但却无法解释后象效应和颜色对比效应。

19世纪中叶,Helmhotz在肯定T.Young三色原理的基础上,认为人眼视网膜

上具有三种视神经细胞,每种神经细胞的兴奋都产生一种原色,并补充了两个假设,其一是在除去色刺激后的一定时间内,视网膜锥体细胞仍维持其激励状态,并将一定信息传入大脑;其二是在同一刺激持续一定时间后,被激励的神经细胞会疲劳,因而对同种刺激会降低其感受灵敏度,从而形成了T.Young-Helmhotz三原色理论(三分量理论)。这种改进的三原色理论可归纳为“物理刺激在神经末梢分解为三种基本形式,即三种独立的生理成分,这三种“成分”由中枢神经综合而产生色觉。

T.Young-Helmhotz理论很好地解释了正后象效应、负后象效应、对比效应和单色色盲。其最大优点是对色光配色规律的解释及其在彩色复制工业中的应用,并在视网膜感受器一级上,获得了心理物理学和生理学的实验验证,其中有代表的实验有:眼底反射分光密度实验、显微分光度实验和锥体细胞感受器电位光谱敏感度实验等。

但三原色理论存在三个感受曲线不唯一,不能解释色盲成对出现和全色盲等问题。如红色盲与绿色盲成对出现,全色盲无颜色感觉,但有明暗感觉。

#### (2) E.Hering的四色理论

1876年E.Hering从研究心理颜色视觉出发,提出了色觉的对立颜色理论。他认为视网膜上有三种光化学物质—视素,每种视素都能发生同化和异化两种生物化学变化,在同化过程中,视素产生合成,异化过程中则产生分解。同化异化的发生完全是由于不同光谱组成的色光刺激的结果。而同化异化的结果使人产生相应的对立颜色感觉。即R-G, Y-B, K-W等六种不同色觉,其对应关系的归纳如表2-1所示。

表 2-1 色觉对立关系

光化学物质	视网膜上的反应	产生的色觉
W (白) —K (黑)	异化 (分解)	W (白)
	同化 (合成)	K (黑)
R (红) —G (绿)	异化 (分解)	R (红)
	同化 (合成)	G (绿)
Y (黄) —B (蓝)	异化 (分解)	Y (黄)
	同化 (合成)	B (蓝)

在这个理论中,由于W-K、R-G与Y-B是互补色,即对立,因此也称为对立颜色理论。它与三原色理论的不同之处是除红、绿、蓝外,增加了黄色、黑色、白色,由于E.Hering理论不计算黑色与白色,因而简称为四色理论。

四色理论是针对三原色理论的不足而提出,解决了三原色理论无法解释色盲成对出现、颜色对比现象等问题,其对成对色盲出现的解释是视网膜上缺少某种对立的光化学物质,如红—绿色盲是视网膜缺少红—绿光化学物质,而全色盲是缺少红—绿,黄—蓝两种光化学物质,而只有黑—白光化学物质。对于颜色对比则是由于对立光化学物质的异化与同化,使得互补色之间形成强烈的对比现象,如将绿色放在红色背景下,会使得绿色更绿,红色更红。

但四色理论存在的问题是无法解释红、绿、蓝三原色能够混合匹配所有颜色的规律。在解剖学实验上,始终未能证实其在视网膜感受水平一级存在三种对立光化学物质。

### 2.2.4 颜色的心理描述

颜色是外界光刺激作用于人的视觉器官而产生的主观感觉。所以颜色特性既可以从客观刺激方面来衡量,也可以从观察者的主观感觉方面来描述。描述客观刺激的概念是心理物理学概念;描述观察者主观感觉的概念是心理学概念。确定光的心理物理量与心理量的关系是感觉心理学研究的重要任务。颜色视觉有三种特性,描述颜色的心理物理量是亮度、主波长和纯度,相应的心理量是明度、色调和饱和度。

颜色分两大类,非彩色和彩色。非彩色是指黑色、白色和介于这两者之间深浅不同的灰色。它们可以排成一个系列,由白色逐渐到浅灰、中灰、深灰直到黑色。这叫白黑系列或无色系列白黑系列。由白到黑的变化可以用一条直线代表,一端是纯白,另一端是纯黑。中间有着各种不同等级的灰色过渡。所谓灰色是相对的,比周围明亮的称为浅灰,比周围暗的称为深灰,灰色是最不饱和色之一。所谓纯白和纯黑也是相对而言的,并无绝对的标准,白雪接近纯白,黑绒接近纯黑,由白和黑按不同比例混合可得出各种灰色。白色和各种灰色是物体表面没有选择性的反射。白黑系列的非彩色的反射率代表物体的明度。反射率越高时接近白色,反射率低时接近黑色。一张洁白的纸的反射率可达85%以上,用来测量颜色,定标用的标准白板的反射率可达90%以上。一张黑纸的反射率可低至5%以下,黑色天鹅绒的反射率甚至可低于0.05%。

表示光的强度的心理物理学概念是亮度(Luminosity)。所有的光,不论是什么颜色都可以用亮度来测量。非彩色的白黑变化相应于白光的亮度变化。当白光的亮度非常高时,人眼就感觉到是白色的;当光的亮度很低时就感觉到发暗或发灰;无光时是黑色的。与亮度相应的心理学概念是明度(brightness)。明度是人眼对物体的明亮感觉,受视觉感受性和过去经验的影响。通常明度的变化相应于亮度的变化。物体表面或光源的亮度越高,人感觉到的明度就越高。但二者的关

系并不固定。若亮度的微小增加或减少达不到人眼的分辨阈限,眼睛就觉察不出明度的变化。这时亮度虽有变化而明度却不变。在暗环境中观察一张高反射率的书页与在亮的环境中观察一块低反射率的黑墨相比,虽然可能后者的亮度大于前者,但由于观察者已经知道它们是书页和黑墨,所以书页仍感觉为白色,而黑墨仍为黑色,有较低的明度。这是因为观察者有对书页和黑墨的记忆和经验,有周围其它物体的相对明度作参考,以及对不同照明条件的认识影响了明度感觉。

彩色系列或有色系列是指除了白黑系列以外的各种颜色。我们通常所说的颜色。即指颜色的第一个特性<sup>[4]</sup>是用心理物理量亮度和心理量明度来表示的。所有的光,不论是什么颜色都可以用光的亮度来定量。与非彩色相似,彩色光的亮度越高,人眼就感觉明亮,或者说有较高的明度。彩色物体表面的反射率越高,它的明度就越高。

表示彩色的第二个特性的心理物理学概念是主波长(dominant wavelength)。与主波长相应的心理学概念是色调(hue)。光谱是由不同波长的光组成的,用三棱镜可以把日光分解成光谱上不同波长的光,不同波长所引起的不同感觉就是色调。例如,700纳米波长光的色调是红色,579纳米波长光的色调是黄色,500纳米波长光的色调是绿色。等若将几种主波长不同的光按照适当的比例加以混合,则能产生不具有任何色调的感觉,也就是白色。事实上,只选择两种主波长不同的光以适当的比例加以混合也能产生白色。这样的一对主波长的光叫做互补波长。例如,600纳米的橙色和492纳米的蓝绿色是一对互补波长;575.5纳米的黄色和474.5纳米的蓝色也是互补波长;一对互补波长的色调叫互补色。光源的色调决定于人眼对辐射光的光谱组成产生的感觉。物体的色调决定于光源的光谱组成和物体表面反射(透射)的各波长的比例对人眼产生的感觉。例如,在日光下,一个物体反射480-560纳米波段的辐射,而相对吸收其它波长的辐射,那么该物体表面为绿色。

表示彩色第三个特性的心理物理学概念是颜色纯度(purity)。其相应的心理学概念是饱和度(saturation)。纯色是指没有混入白色的窄带单色光。在视觉上就是高饱和度的颜色。可见光谱的各种单色光是最饱和的彩色。当光谱色掺入白光成分越多时,就越不饱和。例如主波长为650纳米的光是非常纯的红光。如果把一定数量的白光加到这个红光上,混合的结果便产生粉红色。加入的白光越多,混合后的颜色光就越不纯,看起来就越不饱和。

光刺激的心理物理特性可以按亮度,主波长和纯度来确定。这些特性又分别同明度、色调、饱和度的主观感觉相联系。颜色可分为彩色和非彩色。光刺激如果没有主波长,这个光就是非彩色的白光,它没有纯度。然而所有视觉刺激都有亮度特性。亮度是彩色刺激和非彩色刺激的共同特性,而主波长和纯度表示刺激

是彩色的。

### 2.3 颜色的交互作用和颜色恒常性

我们在某一物体表面所看到的颜色不仅取决于这个表面本身产生的物理刺激，而且还取决于同时呈现在它周围的颜色。物体本身的颜色和它周围颜色的交互作用能影响被看表面的色调和明度。当被看的颜色向它周围颜色的对立方向转化，即向周围颜色的补色方向变化时，叫做颜色的同时交互作用，或颜色对比。例如，在红色背景中放一小块白纸或灰纸，用眼睛注视白纸几分钟白纸会表现出绿色；如果背景是黄色，白纸会出现蓝色；红和绿是互补色，黄和蓝也是互补色。

当在一个颜色（包括灰色）的周围呈现高亮度或低亮度刺激时，这个颜色就向其周围明度的对立方向转化，这叫做明度对比。白背景上的灰方块呈浅黑色，而黑背景上的灰方块则呈白色等。对比效应在视觉中有重要作用，明度对比更是这样：它与视觉中的颜色恒常性<sup>[15]</sup>相联系。一块煤在阳光下单位面积反射光比一张白纸在黑暗处时高一千倍。但我们仍然把煤看成是黑的，而把纸看成是白的灰的或黑的。这是由这个物体与周围物体的相对明度关系决定的。白纸不管在什么样的照明条件下都是白的，而煤仍然是黑的。这就是说尽管外界的条件发生变化，人们仍然根据物体的固有颜色来感知它们，这是颜色恒常性的表现。在一天的过程中，我们周围物体所受的照度会有很大的变化。中午时的照度要比日出时和日落时高几百倍。但在日常生活中，当照明条件变化时我们的视觉仍能保持对物体颜色的恒常性，这才使我们对周围物体有正确的认识。

颜色恒常性还表现在当光源的光谱成分发生变化时被观察物体的颜色在一定程度上看起来仍然保持不变。例如，室内不管由白炽灯的黄光或荧光灯的蓝光照明，书页纸看起来都是白色的。但如果让被试验者通过一个圆筒，只看到被照射物体的一小块面积，同时又不让他知道是用哪一种色光照射时，一张白纸在用红光照射时就会被看成是红色的。

一个物体的明度和色调不仅取决于当前的刺激，而且也与先前刺激的后效有关。这种后效叫相继交互作用，或相继对比效应。例如我们看一个红色方块一段时间后再注意看一个均匀的灰色表面，就会看到一个很快消失的绿色方块的映象，映象的颜色是诱导颜色的补色，这样产生的映象叫做负后象。对黑色和白色也是同样的情况黑色的负后象是白色，而白色的负后象是黑色。后象也可以是与原来刺激相同的颜色，这种后象叫做正后象。

## 2.4 颜色特征提取方法

颜色特征是在物体识别中应用最为广泛的视觉特征，主要原因在于颜色往往和图像中所包含的物体或场景十分相关。此外，与其他的视觉特征相比，颜色特征对图像本身的尺寸、方向、视角的依赖性较小，从而具有较高的鲁棒性。

颜色特征的表达涉及到若干问题。首先，我们需要选择合适的颜色空间来描述颜色特征；其次，我们要采用一定的量化方法将颜色特征表达为向量的形式；最后，还要定义一种相似度（距离）标准用来衡量图像之间在颜色上的相似性。在本节中，我们将主要讨论前两个问题，并介绍颜色直方图、颜色矩、颜色集、颜色聚合向量以及颜色相关图等颜色特征的表示方法。

### 2.4.1 颜色直方图

灰度直方图<sup>[16]</sup>是反映灰度图像的亮度概率分布的有效工具，得到广泛的应用和深入的研究。颜色直方图同样反映彩色图像的颜色信息在颜色空间的分布情况，多种系统的应用同样表明彩色直方图具有很好的表现能力。

颜色在传统的图像理解中未能受到普遍重视，然而相对于几何特征而言，颜色具有与生俱来的尺度不变性<sup>[17]</sup>（scale invaruance）和旋转不变性<sup>[18]</sup>（rotate invariance）。

颜色直方图是刻画颜色信息的一种有效方法，通过对每种颜色在图像内的像素数统计而构造，归一化的颜色直方图具有尺度不变性（scale invaruance）和旋转不变性（rotate invariance），特征的鲁棒性好，在对彩色图像的直方图特征提取过程中，颜色空间的选择是极为重要的一环。

#### 2.4.1.1 颜色空间

图像的每一个像素可以看作是3维颜色空间中的一个点，颜色空间有RGB，Munsell，CIEL\*a\*b，CIEL\*u\*v\*，HSV(orHSL，HSB)等，采用何种颜色空间没有一个标准，一般选择视觉一致性颜色空间。所谓视觉一致性颜色空间是指在颜色空间中定义一个颜色度量函数，如果在这个颜色度量下，人的视觉感觉差异比较大的颜色之间距离也大，视觉感觉差异比较小的颜色之间距离也小，具有这种性质的颜色空间即为具有视觉一致性的颜色空间。

##### (1)RGB颜色空间

RGB颜色空间在图像显示中是一种常用的颜色空间。它是三维空间包括R、G、B三个坐标轴。每一个坐标轴的值从0到255，共256个水平。但是RGB空间的

一个主要缺点是不具备视觉统一性。比如红色R(Red[255,0,0])、粉红色M(Magenta[255,0,255])、蓝色B(Blue[0,0,255])，按照 $L_2$ 距离，显然 $d_2(M, R) = d_2(M, B)$ ，但是对人的视觉感觉来说，粉红色应该更接近于红色。

## (2) CIEL\*a\*b\*颜色空间

国际照明委员会(CIE)<sup>[10]</sup>于1931年提出了CIE L\*a\*b\*颜色空间，从RGB颜色空间到CIEL\*a\*b\*，颜色空间的转换公式为：

RGB → CIE XYZ

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.607 & 0.174 & 0.200 \\ 0.299 & 0.587 & 0.114 \\ 0.000 & 0.066 & 1.116 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2-1)$$

CIE XYZ → CIE L\*a\*b\*

$$L = \begin{cases} 116 \times \left(\frac{Y}{Y_N}\right)^{1/3} & \text{当 } \frac{Y}{Y_N} > 0.008856 \\ 903.3 \times \frac{Y}{Y_N} & \text{否则} \end{cases} \quad \begin{cases} a = 500 \times \left(f\left(\frac{X}{X_N}\right) - f\left(\frac{Y}{Y_N}\right)\right) \\ b = 200 \times \left(f\left(\frac{Y}{Y_N}\right) - f\left(\frac{Z}{Z_N}\right)\right) \end{cases} \quad (2-2)$$

$$\text{这里 } f(t) = \begin{cases} t^{1/3} & \text{当 } t > 0.008856 \\ 7.787 \times t + 16/116 & \text{否则} \end{cases} \quad (2-3)$$

其中(X, Y)=(0.312713, 0.329016)为白色点的固定坐标。

## (3) CIEL\*u\*v\*颜色空间

从RGB颜色空间转换到CIE XYZ颜色空间，然后再转换到CIEL\*u\*v\*颜色空间，计算公式为：

$$L^* = \begin{cases} 116 \left(\frac{Y}{Y_n}\right)^{1/3} - 16 & \text{当 } \frac{Y}{Y_n} > 0.008856 \\ 903.3 \times \frac{Y}{Y_n} & \text{否则} \end{cases} \quad (2-4)$$

$$u^* = 13L^*(u' - u'_n) \quad (2-5)$$

$$v^* = 13L^*(v' - v'_n) \quad (2-6)$$

这里 $u', u'_n$ 和 $v', v'_n$ 分别为：

$$u' = \frac{4X}{X+15Y+3Z}, \quad v' = \frac{9Y}{X+15Y+3Z} \quad (2-7)$$

$$u'_n = \frac{4X_n}{X_n+15Y_n+3Z_n}, \quad v'_n = \frac{9Y_n}{X_n+15Y_n+3Z_n} \quad (2-8)$$

$X_n, Y_n, Z_n$ 为亮度值。



#### (4)HSV颜色空间

HSV颜色空间具有视觉一致性，Hue表示颜色的色调；Saturation表示颜色的纯度，即表示一种颜色中加入了多少白光；Value表示颜色值的大小。另外国际标准颜色组织提出采用H、S、V三维空间表示的Munsell颜色空间。RGB空间很容易变换到HSV颜色空间，从RGB颜色空间到HSV颜色空间的转换公式为：

$$h' = \begin{cases} \frac{(g-b)}{\delta} & \text{如果 } r = \max \\ \frac{2+(b-r)}{\delta} & \text{如果 } g = \max \\ \frac{4+(r-g)}{\delta} & \text{如果 } b = \max \end{cases} \quad (2-9)$$

$$v = \max, s = \frac{\max - \min}{\max}$$

$$h = h' \times 60$$

这里  $\max = \text{MAX}(r, g, b)$ ,  $\min = \text{MIN}(r, g, b)$ ,  $\delta = \max - \min$ 。(h, s, v)为在HSV颜色空间中对于RGB颜色空间中点(r, g, b)坐标，而且有：

$$(r, g, b) \in [0, 1.0]; h \in [0, 360]; (s, v) \in [1, 1.0]。$$

##### 2.4.1.2 颜色空间的量化

计算颜色直方图需要将颜色空间划分成若干个小的颜色区间，每个小区间成为直方图的一个bin。这个过程称为颜色量化（color quantization）。然后，通过计算颜色落在每个小区间内的像素数量可以得到颜色直方图。颜色量化有许多方法，例如向量量化、聚类方法或者神经网络方法。最为常用的做法是将颜色空间的各个分量（维度）均匀地进行划分。相比之下，聚类算法则会考虑到图像颜色特征在整个空间中的分布情况，从而避免出现某些bin中的像素数量非常稀疏的情况，使量化更为有效。另外，如果图像是RGB格式而直方图是HSV空间中的，我们可以预先建立从量化的RGB空间到量化的HSV空间之间的查找表（look-up table），从而加快直方图的计算过程。

上述的颜色量化方法会产生一定的问题。设想两幅图像的颜色直方图几乎相同，只是互相错开了一个bin，这时如果我们采用L1距离或者欧拉距离计算两者的相似度，会得到很小的相似度值。为了克服这个缺陷，需要考虑到相似但不同的颜色之间的相似度，一种方法是采用二次式距离；另一种方法是对颜色直方图事先进行平滑过滤，即每个bin中的像素对于相邻的几个bin也有贡献。这样，

相似但不相同颜色之间的相似度对直方图的相似度也有所贡献。

选择合适的颜色小区间（即直方图的bin）数目和颜色量化方法与具体应用的性能和效率要求有关。一般来说，颜色小区间的数目越多，直方图对颜色的分辨能力就越强。然而，bin的数目很大的颜色直方图会增加计算负担。而且对于某些应用来说，使用非常精细的颜色空间划分方法不一定能够提高效果，特别是对于不能容忍对相关图像错漏的那些应用。另一种有效减少直方图bin的数目的办法是只选用那些数值最大（即像素数目最多）的bin来构造图像特征，因为这些表示主要颜色的bin能够表达图像中大部分像素的颜色。实验证明这种方法并不会降低颜色直方图的检索效果。事实上，由于忽略了那些数值较小的bin，颜色直方图对噪声的敏感程度降低了，有时会使效果更好。

#### 2.4.2 颜色矩

另一种非常简单而有效的颜色特征使由Stricker和Orengo<sup>[20]</sup>所提出的颜色矩（color moments）。这种方法的数学基础在于图像中任何的颜色分布均可以用它的矩来表示。此外，由于颜色分布信息主要集中在低阶矩中，因此仅采用颜色的一阶矩（mean）、二阶矩（variance）和三阶矩（skewness）就足以表达图像的颜色分布。与颜色直方图相比，该方法的另一个好处在于无需对特征进行向量化。颜色的三个低次矩在数学上表达为：

$$\mu_i = \frac{1}{N} \sum_{j=1}^N p_{ij} \quad (2-10)$$

$$\sigma_i = \left( \frac{1}{N} \sum_{j=1}^N (p_{ij} - \mu_i)^2 \right)^{\frac{1}{2}} \quad (2-11)$$

$$s_i = \left( \frac{1}{N} \sum_{j=1}^N (p_{ij} - \mu_i)^3 \right)^{\frac{1}{3}} \quad (2-12)$$

其中  $p_{ij}$  是图像中第  $j$  个像素的第  $i$  个颜色分量。因此，图像的颜色矩一共只需要9个分量（3个颜色分量，每个分量上3个低阶矩），与其他的颜色特征相比是非常简洁的。在实际应用中为避免低次矩较弱的分辨能力，颜色矩常和其它特征结合使用，而且一般在使用其它特征前起到过滤缩小范围（narrow down）的作用。

### 2.4.3 颜色集

为支持大规模图像库中的快速查找, Smith和Chang<sup>[21]</sup>提出了用颜色集(color sets)作为对颜色直方图的一种近似。他们首先将RGB颜色空间转化成视觉均衡的颜色空间(如HSV空间), 并将颜色空间量化成若干个bin。然后, 他们用颜色自动分割技术将图像分为若干区域, 每个区域用量化颜色空间的某个颜色分量来索引, 从而将图像表达一个二进制的颜色索引集。在图像匹配中, 比较不同图像颜色集之间的距离和颜色区域的空间关系(包括区域的分离、包含、交等, 每种对应于不同得分)。

### 2.4.4 颜色聚合向量

针对颜色直方图和颜色矩无法表达图像颜色的空间位置的缺点, Pass提出了图像的颜色聚合向量<sup>[22]</sup>(color coherence vector)。它是颜色直方图的一种演变, 其核心思想是将属于直方图每一个bin的像素进行分为两部分: 如果该bin内的某些像素所占据的连续区域的面积大于给定的阈值, 则该区域内的像素作为聚合像素, 否则作为非聚合像素。假设 $\alpha_i$ 与 $\beta_i$ 分别代表直方图的第 $i$ 个bin中聚合像素和非聚合像素的数量, 图像的颜色聚合向量可以表达为 $\langle(\alpha_1, \beta_1), (\alpha_2, \beta_2), \dots, (\alpha_N, \beta_N)\rangle$ 。而 $\langle(\alpha_1 + \beta_1), (\alpha_2 + \beta_2), \dots, (\alpha_N + \beta_N)\rangle$ 就是该图像的颜色直方图。由于包含了颜色分布的空间信息, 颜色聚合向量相比颜色直方图可以达到更好的识别效果。

### 2.4.5 颜色相关图

颜色相关图(color correlogram)<sup>[23]</sup>是图像颜色分布的另一种表达方式。这种特征不但刻画了某一种颜色的像素数量占整个图像的比例, 还反映了不同颜色对之间的空间相关性。

假设 $I$ 表示整张图像的全部像素,  $I_{c(i)}$ 则表示颜色为 $c(i)$ 的所有像素。颜色相关图可以表达为:

$$\gamma_{i,j}^{(k)} = \frac{p_r}{p_i \in I_{c(i)}, p_j \in I_{c(j)}} [p_2 \in I_{c(j)} \mid |p_1 - p_2| = k] \quad (2-13)$$

其中 $i, j \in \{1, 2, \dots, N\}$ ,  $k \in \{1, 2, \dots, d\}$ ,  $|p_1 - p_2|$ 表示像素 $p_1$ 和 $p_2$ 之间的距离。

颜色相关图可以看作是一张用颜色对 $\langle i, j \rangle$ 索引的表, 其中 $\langle i, j \rangle$ 的第 $k$ 个分量表示颜色为 $c(i)$ 的像素和颜色为 $c(j)$ 的像素之间的距离小于 $k$ 的概率。如果考虑到任

何颜色之间的相关性，颜色相关图会变得非常复杂和庞大(空间复杂度为 $O(N^2d)$ )。一种简化的变种是颜色自动相关图 (color auto-correlogram)，它仅仅考察具有相同颜色的像素间的空间关系，因此空间复杂度降到 $O(Nd)$ 。

#### 2.4.6 局部颜色特征

以上各种颜色特征是基于全局颜色特征索引捕获了整幅图像颜色分布的信息，但是丢失了许多局部的颜色空间信息。Fuh等试图结合图像的颜色信息和颜色的分布空间信息检索图像，其主要思想是从图像中选择一些代表颜色，然后将图像划分成一定的矩形区域，每个区域中以一种主要的单一颜色作为代表，两个图像之间的相似性是两个图像之间具有相似颜色区域的重叠程度。Smith和Chang<sup>[24]</sup>采用颜色的自动分割方法，形成一个二进制颜色索引集，在图像匹配中，比较这些图像颜色集之间的距离。

### 2.5 小结

本章介绍了颜色产生的视觉机理，同时对产生颜色恒常性的原理作了阐述，列举了颜色特征提取的不同方法，着重介绍了基于直方图的颜色特征提取方法，并对基本的颜色空间作了介绍。

## 第三章 基于二色反射模型的颜色直方图识别

### 3.1 颜色识别

颜色能够在物体识别中提供重要的信息, Swain和Ballard<sup>[26]</sup>提出了的基于颜色直方图的物体识别方法, 是在物体识别中的一个重大的改进。但是, 基于颜色直方图的方法在环境光线不均匀时, 物体识别率会变得很差。Funt和Finlayson<sup>[26]</sup>通过retinex理论扩展了基于颜色直方图的方法, 通过相邻点的计算, 得到光照无关的表面特征描述, 但是前提是相邻点拥有相同的表面特征, 因此这种方法并不适用于表面变化剧烈的物体; Healey、Slater和Finlayson<sup>[27]</sup>通过颜色矩的方法提取光照不变量, 但是这种方法对于不规则的物体效果不好; 对于一个通过整体图像来识别物体的方法, 图像中不规则的部分通常会对物体识别造成影响, 为了减少影响, Slater和Healey通过把图像分割为若干小块代替对图像整体的检测。

在对彩色图像的直方图特征提取过程中, 颜色空间的选择是极为重要的一环。通过以上方法的描述, 可知在物体识别中, 图像通常受到光照变化, 物体摆放角度以及摄像机视角的影响, 为此, 本章希望找到一个简便, 分辨能力强并且对于干扰不敏感的颜色模型, 使得它能够独立于光照变化不受光源以及环境的影响。

### 3.2 二色反射模型

通过上一章的介绍, 我们可以知道应用颜色直方图来描述颜色特征, 首要的问题就是颜色空间的选择, RGB空间虽然有利于图像的显示, 但不是一个均匀的颜色空间, 在其上定义符合视觉特性的距离比较困难, 通过对人眼颜色视觉特性的研究和应用。我们比较上节所述的几个颜色空间光照稳定性, 通过运算找到能够独立于光照变化的颜色空间。

颜色信息最重要的优点之一是物体表面的颜色在几何形状和位置变化时比相应的图象辐射率稳定得多, 所以过去在颜色图象研究中把颜色看成是物体的不变特性, 把颜色的变化看作是噪声引起的。但是实际上颜色的变化在很大程度上与景物的光学反射特性有关。物体的反射特性造成了物体的颜色、高光 (highlight) 和阴影 (shading)。因此需要建立物体的颜色反射模型来说明由于高光和影调引起颜色变化的规律。这里我们介绍一种双色反射模型<sup>[28]</sup>

(dichromatic reflection model)。

当我们观察一个光亮的物体时，通常认为反射光由两种颜色组成，一种是高光区域的颜色，另一种是无光泽物体区域的颜色。双色反射模型可描述任意颜色的单光源照射不透明的绝缘材料（dielectric materials）时的反射现象。在合适的条件下这个模型也能很好地说明实际的颜色信息。

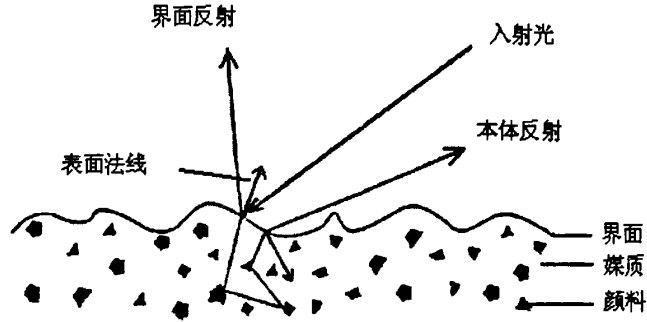


图3-1 光线在绝缘材料表面上的反射

图3-1所示为光线在不均匀的绝缘材料，如塑料和颜料，表面上的反射情况。根据Fresnel的反射定理<sup>[29]</sup>，一部分入射光线在物质的界面处被立即反射。一般来说，这部分反射光的颜色大致与光源相同。其余的入射光线将穿透进入物质内部。这部分光线被物质所散射，其中某些波长被吸收，然后再次射出物质表面。这部分光的颜色决定于光照的颜色和物质的反射特性。在界面处的反射过程通常被称为镜面反射、高光或光泽，而在物质内部的反射过程被称为漫反射或无光泽反射，把这两个过程分别称为界面反射和本体反射（body reflection）可能更为确切，因为这样能说明这两个反射过程在物理本质上的差别。

基于上述讨论，双色反射模型把景物中某一点的光谱辐射  $L(\lambda, i, e, g)$  表示为界面反射分量  $L_s(\lambda, i, e, g)$  和本体反射分量  $L_b(\lambda, i, e, g)$  之和，其中  $\lambda$  是波长， $i$  和  $e$  分别是光线的入射和出射角， $g$  是相位角。这个模型假设光照和物体的反射光谱特性与物体表面的方向无关，这是合理的近似，这样我们就可把这两个反射分量分解为描述颜色的光谱分量  $c(\lambda)$  和描述比例因子的幅度分量  $m(i, e, g) \in [0, 1]$ 。即：

$$\begin{aligned} L(\lambda, i, e, g) &= L_s(\lambda, i, e, g) + L_b(\lambda, i, e, g) \\ &= m_s(i, e, g)c_i(\lambda) + m_b(i, e, g)c_b(\lambda) \end{aligned} \quad (3-1)$$

当摄像机记录景物的图象时，通常采用R、G、B三个参数表示由物体表面反

射到摄像机的颜色光的光谱。根据颜色相加原理,在成象过程中摄像机把象素位置 $(x, y)$ 处的入射光线,通过三刺激值的积分,把颜色光谱变换成R、G、B三刺激值即 $C(x, y)=[R, G, B]$ 。在此过程中把每个波长上的颜色光经过颜色滤波器透射率和该波长上摄像机的响应特性加权后相加求和。因为这是线性变换,同时因为光度学的角度 $i$ ,  $e$ 和 $g$ 与位置 $x$ 和 $y$ 有关,所以根据双色反射模型我们可把颜色象素值 $C(x, y)$ 描述为景物中对应点处界面反射和本体反射的颜色向量,  $C_s$ 和 $C_b$ 的线性组合:

$$C(x, y) = m_s(i, e, g)C_s + m_b(i, e, g)C_b \quad (3-2)$$

在上述公式中颜色向量 $C_s$ 和 $C_b$ 对一种物质的表面来说是不变的,而比例因子 $m_s$ 和 $m_b$ 则随不同象素而变化。

为了表示在整个物体上颜色的变化情况,我们利用颜色空间的颜色直方图。建立这个直方图的方法是把物体上所有象素的颜色投影到颜色空间中去。双色反射模型假设景物中只有一个光源,没有背景光和物体之间的相互反射。在此假设下,物体上所有象素的颜色是界面反射颜色 $C_s$ 和本体反射颜色 $C_b$ 的线性组合。因此,在物体区域内颜色的变化取决于几何比例因子 $m_s$ 和 $m_b$ ,而 $C_s$ 和 $C_b$ 保持不变。与此相应, $C_s$ 和 $C_b$ 在颜色空间中形成一个双色平面。在此双色平面内颜色象素形成一个稠密的集群,它的形状与物体材料的几何特性、界面反射以及物体形状密切相关,我们可以利用这个关系来确定颜色象素集群的特征。为便于说明把颜色象素分类为无光泽象素、高光象素和剪裁象素。

无光泽象素是景物中在观察方向上只呈现出本体反射的象素的投影。这些象素的颜色只取决于本体反射的颜色,其幅度和物体的局部表面法线与光照方向之间的夹角成比例。所以在颜色空间中无光泽象素的颜色沿本体反射向量 $C_b$ 的方向形成一条无光泽线。

高光象素是景物中下述点的投影;这些点在观察者方向上同时呈现本体反射和界面反射。在高光区域中所有象素的颜色位于下述直线上:在此直线上本体反射为常数而只有界面反射量的变化。因此,这些象素颜色在颜色空间中形成的高光直线与界面反射向量 $C_s$ 平行,并且是从代表所有位置高光象素的本体反射分量的无光泽集群出发的。与相邻位置的高光象素(这些相邻象素的本体反射的幅度略有不同)组合在一起,所有的高光象素在颜色空间中形成一个斜的楔子形的高光集群。如果在景物中物体上不止一处有高光,那么每一处高光都在颜色空间形成一条光线。

剪裁颜色象素是一些反射光超过摄像机的动态范围的高光象素,取决于物体

的颜色,可能出现在一种原色的通道上超过动态范围,而在其它两个原色通道上没有超过的情况,这时高光集群在颜色立方体的边界面上发生弯曲,如果在两个原色通道饱和,那么就在颜色立方体边界的交边上出现高光集群的弯曲。在高光反射的中心点处可能在三个原色通道上都超过摄像机的动态范围,那么这时即使光照不是白色的,但这些点看起来也是白色的。

根据双色反射模型,利用对颜色集群的分析可以确定光照的颜色,从而可去除颜色光源的影响;此外还可检测和去除高光,分离后的无光泽图象对观察者的方向变化不敏感,从而可抽取可靠的特征用于立体视觉和运动视觉。同时高光图象本身也可提供关于光源和表面方向的信息。

### 3.3 不同颜色空间颜色不变性分析

彩色模型(也称作彩色空间或彩色系统)的用途是在某些标准下用通常可接受的方式简化彩色规范。本质上,彩色模型是坐标系统和子空间的规范。而不同的颜色模型受光照的影响程度是不同的,所以适当的选取或构造光照稳定性强的颜色模型,可以在光照条件变化时,仍能够保证颜色值的可靠性和算法的鲁棒性。

根据上一节描述的Shafer的二色反射理论,物体表面反射的过程分为两种,即表面反射和本体反射。表面反射分量一般与入射光有相同的光谱分布,本体反射分量的光谱则是有光照的光谱功率分布和物体表面反射特性共同决定的。本体反射率则代表了物体的颜色反射特性,可以表征物体颜色的本征特性。所以去除掉表面反射分量,再通过对本体反射的分量处理,从而可以得到无关于光照和几何特性无关的本体反射率。

#### 3.3.1 二色反射模型

Shafer认为对于一个非均质的非导电物体的一块无限小的表面区域,在光谱能量分布为 $e(\lambda)$ 的光源照射下,利用光谱敏感特性分别为 $f_R(\lambda)$ ,  $f_G(\lambda)$ 和 $f_B(\lambda)$ 的三个传感器获得的颜色值为:

$$C = m_b(\vec{n}, \vec{s}) \int_{\lambda} f_c(\lambda) e(\lambda) c_b(\lambda) d\lambda + m_s(\vec{n}, \vec{s}, \vec{v}) \int_{\lambda} f_c(\lambda) e(\lambda) c_s(\lambda) d\lambda \quad (3-3)$$

$f_R(\lambda)$ ,  $f_G(\lambda)$ 和 $f_B(\lambda)$ 分别表示红色、绿色和蓝色的传感器光谱响应,光源的光谱能量分布强度用 $e(\lambda)$ 来表示,  $C = \{R, G, B\}$ 表示第 $C$ 个传感器的响应值,



$c_b(\lambda)$  和  $c_s(\lambda)$  分别表示物体的体反射率和表面反射率。 $\lambda$  表示光源波长,  $\vec{n}$  为表面法线方向,  $\vec{s}$  表示光源方向,  $\vec{v}$  表示观察者的方向, 其中  $m_b$ 、 $m_s$  分别表示体反射和表面反射的几何变量。

通过第二章的讨论, 我们在此列出几种颜色空间, 分别讨论下它们在不同的光照情况下稳定性的对比。

颜色空间  $c_1$ 、 $c_2$  和  $c_3$ :

$$c_1 = \arctan\left(\frac{R}{\max\{G, B\}}\right) \quad (3-4)$$

$$c_2 = \arctan\left(\frac{G}{\max\{R, B\}}\right) \quad (3-5)$$

$$c_3 = \arctan\left(\frac{B}{\max\{R, G\}}\right) \quad (3-6)$$

颜色空间  $r$ ,  $g$ ,  $b$ :

$$r = \frac{R}{R+G+B}; g = \frac{G}{R+G+B}; b = \frac{B}{R+G+B} \quad (3-7)$$

颜色空间 HIS 的定义如下:

$$I = (R+G+B)/3 \quad (3-8)$$

$$H = \arctan\left(\frac{\sqrt{3}(G-B)}{(R-G)+(R-B)}\right) \quad (3-9)$$

$$S = 1 - \frac{\min(R, G, B)}{R+G+B} \quad (3-10)$$

### 3.3.2 不同物体特性及光照下模型的推导

以上是一些颜色空间定义, 下面我们就不同物体特性及光照情况进行讨论:

(1) 首先考虑中性界面反射 ( $c_s(\lambda) = c_s$ ) 和白光条件  $e(\lambda) = e$  因此公式(3-3)

变为

$$C = em_b(\vec{n}, \vec{s})k_c + em_s(\vec{n}, \vec{s}, \vec{v})c_s \int f_c(\lambda)d\lambda \quad (3-11)$$

其中我们设:

$$\int_{\lambda} f_R d\lambda = \int_{\lambda} f_G d\lambda = \int_{\lambda} f_B d\lambda = f \quad k_C = \int_{\lambda} f_C(\lambda) c_b(\lambda) d\lambda \quad (3-12)$$

将公式(3-12)代入公式(3-11):

$$C = em_b(\vec{n}, \vec{s}) k_c + em_s(\vec{n}, \vec{s}, \vec{v}) c_s f \quad (3-13)$$

对于表面粗糙的物体, 其反射主要有体反射决定, 即:

$$C_b = em_b(n, s) k_C \quad (3-14)$$

将式(3-14)体反射分量代入  $c_1$ 、 $c_2$  和  $c_3$  中得:

$$\begin{aligned} c_1(R_b, G_b, B_b) &= \arctan \left( \frac{em_b(n, s) k_R}{\max \{em_b(n, s) k_G, em_b(n, s) k_B\}} \right) \\ &= \arctan \left( \frac{k_R}{\max \{k_G, k_B\}} \right) \end{aligned} \quad (3-15)$$

$$\begin{aligned} c_2(R_b, G_b, B_b) &= \arctan \left( \frac{em_b(n, s) k_G}{\max \{em_b(n, s) k_R, em_b(n, s) k_B\}} \right) \\ &= \arctan \left( \frac{k_G}{\max \{k_R, k_B\}} \right) \end{aligned} \quad (3-16)$$

$$\begin{aligned} c_3(R_b, G_b, B_b) &= \arctan \left( \frac{em_b(n, s) k_B}{\max \{em_b(n, s) k_R, em_b(n, s) k_G\}} \right) \\ &= \arctan \left( \frac{k_B}{\max \{k_R, k_G\}} \right) \end{aligned} \quad (3-17)$$

将式(3-14)代入颜色空间  $r$ ,  $g$ ,  $b$  得:

$$r(R_b, G_b, B_b) = \frac{em_b(n, s) k_R}{em_b(n, s) (k_R + k_G + k_B)} = \frac{k_R}{k_R + k_G + k_B} \quad (3-18)$$

$$g(R_b, G_b, B_b) = \frac{em_b(n, s) k_G}{em_b(n, s) (k_R + k_G + k_B)} = \frac{k_G}{k_R + k_G + k_B} \quad (3-19)$$

$$b(R_b, G_b, B_b) = \frac{em_b(n, s) k_B}{em_b(n, s) (k_R + k_G + k_B)} = \frac{k_B}{k_R + k_G + k_B} \quad (3-20)$$

将式(3-14)代入颜色空间 HIS 中得:

$$\begin{aligned}
 H(R_b, G_b, B_b) &= \arctan\left(\frac{\sqrt{3}(G_b - B_b)}{(R_b - G_b) + (R_b - B_b)}\right) \\
 &= \arctan\left(\frac{\sqrt{3}em_b(n, s)(k_G - k_B)}{em_b(n, s)((k_R - k_G) - (k_R - k_B))}\right) \\
 &= \arctan\left(\frac{\sqrt{3}(k_G - k_B)}{(k_R - k_G) - (k_R - k_B)}\right)
 \end{aligned} \tag{3-21}$$

$$\begin{aligned}
 S(R_b, G_b, B_b) &= 1 - \frac{\min(em_b(n, s)k_R, em_b(n, s)k_G, em_b(n, s)k_B)}{em_b(n, s)(k_R + k_G + k_B)} \\
 &= 1 - \frac{\min(k_R, k_G, k_B)}{(k_R + k_G + k_B)}
 \end{aligned} \tag{3-22}$$

很显然, 最后推导出的结果仅与  $k$  有关, 即有效的去除了光照因子得到了只与物体表面特性有关的分量。所以以上的颜色模型在对表面粗糙的物体在不同光照下均有很好的稳定性。

以上讨论的是没有高光反射现象的情况下, 而当物体存在高光区域时, 即需要考虑:

$$C_s = em_s(n, s, v)c_s f \tag{3-23}$$

则 R、G、B 传感器响应为下式表示:

$$R_w = em_b(n, s)k_R + em_s(n, s, v)c_s f \tag{3-24}$$

$$G_w = em_b(n, s)k_G + em_s(n, s, v)c_s f \tag{3-25}$$

$$B_w = em_b(n, s)k_B + em_s(n, s, v)c_s f \tag{3-26}$$

将式(3-24)、(3-25)、(3-26)分别带入到上述颜色空间中, 避免重复说明, 仅保留有效的颜色空间变换。

$$\begin{aligned}
 H(R_w, G_w, B_w) &= \arctan\left(\frac{\sqrt{3}(G_w - B_w)}{(R_w - G_w) + (R_w - B_w)}\right) \\
 &= \arctan\left(\frac{\sqrt{3}em_b(n, s)(k_G - k_B)}{em_b(n, s)((k_R - k_G) - (k_R - k_B))}\right) \\
 &= \arctan\left(\frac{\sqrt{3}(k_G - k_B)}{(k_R - k_G) - (k_R - k_B)}\right)
 \end{aligned} \tag{3-27}$$

由代入颜色空间进行计算可知, 仅有 HIS 模型的 H 分量能够有效地去除光照因子, 得到仅与物体表面特性有关的分量。显然, 其他的光照模型均对高光区域敏感。

对于这种情况, 提出一种新的颜色模型  $l_1, l_2, l_3$ 。

$$l_1 = \frac{(R - G)^2}{(R - G)^2 + (R - B)^2 + (G - B)^2} \tag{3-28}$$

$$l_2 = \frac{(R - B)^2}{(R - G)^2 + (R - B)^2 + (G - B)^2} \tag{3-29}$$

$$l_3 = \frac{(G - B)^2}{(R - G)^2 + (R - B)^2 + (G - B)^2} \tag{3-30}$$

将式(3-24)、(3-25)、(3-26)带入到新建的颜色空间中:

$$\begin{aligned}
 l_1(R_w, G_w, B_w) &= \left( \frac{(R_w - G_w)^2}{(R_w - G_w)^2 + (R_w - B_w)^2 + (G_w - B_w)^2} \right) \\
 &= \frac{(em_b(n, s)(k_R - k_G))^2}{(em_b(n, s)(k_R - k_G))^2 + em_b(n, s)(k_R - k_B))^2 + em_b(n, s)(k_G - k_B))^2}
 \end{aligned} \tag{3-31}$$

$$\begin{aligned}
 &= \frac{(k_R - k_G)^2}{(k_R - k_G)^2 + (k_R - k_B)^2 + (k_G - k_B)^2} \\
 l_2(R_w, G_w, B_w) &= \frac{(k_R - k_B)^2}{(k_R - k_G)^2 + (k_R - k_B)^2 + (k_G - k_B)^2}
 \end{aligned} \tag{3-32}$$

$$l_3(R_w, G_w, B_w) = \frac{(k_G - k_B)^2}{(k_R - k_G)^2 + (k_R - k_B)^2 + (k_G - k_B)^2} \tag{3-33}$$

通过式(3-31)、(3-32)、(3-33)可得颜色模型  $l_1, l_2, l_3$  能够很好的解决光照变化及高光的影响。

(2)当不同颜色光源照射情况下时, 表面反射分量应为体反射:

$$C_c = m_b(n, s) \int_{\lambda} f_c(\lambda) e(\lambda) c_b(\lambda) d\lambda \quad (3-34)$$

由于摄像机彩色传感器的窄带效应, 故其光谱敏感特性可近似用  $\delta$  函数表示  $f_c(\lambda) = \delta(\lambda - \lambda_c)$ , 表面反射分量可以表示为:

$$C_c = m_b(n, s) e(\lambda_c) c_b(\lambda_c) \quad (3-35)$$

为了能够适应不同光照情况以及色光变化, 本文中我们采用基于比率的颜色模型可以更有利于消除光源及环境的影响。为此定义新的颜色模型:

$$Q_1(R_{\bar{x}_1}, R_{\bar{x}_2}, G_{\bar{x}_1}, G_{\bar{x}_2}) = \frac{(R_{\bar{x}_1} - B_{\bar{x}_1})(G_{\bar{x}_2} - B_{\bar{x}_2})}{(R_{\bar{x}_2} - B_{\bar{x}_2})(G_{\bar{x}_1} - B_{\bar{x}_1})} \quad (3-36)$$

$$Q_2(R_{\bar{x}_1}, R_{\bar{x}_2}, B_{\bar{x}_1}, B_{\bar{x}_2}) = \frac{(R_{\bar{x}_1} - G_{\bar{x}_1})(B_{\bar{x}_2} - G_{\bar{x}_2})}{(R_{\bar{x}_2} - G_{\bar{x}_2})(B_{\bar{x}_1} - G_{\bar{x}_1})} \quad (3-37)$$

$$Q_3(G_{\bar{x}_1}, G_{\bar{x}_2}, B_{\bar{x}_1}, B_{\bar{x}_2}) = \frac{(G_{\bar{x}_1} - R_{\bar{x}_1})(B_{\bar{x}_2} - R_{\bar{x}_2})}{(G_{\bar{x}_2} - R_{\bar{x}_2})(B_{\bar{x}_1} - R_{\bar{x}_1})} \quad (3-38)$$

我们假设光照在相邻像素是保持不变的即:  $e^x(\lambda_c) = e^{x+1}(\lambda_c)$

将式(3-33)带入(3-34)、(3-35)、(3-36)中得:

$$Q_1 = \frac{(k_{R_1} - k_{B_1}) * (k_{G_2} - k_{B_2})}{(k_{R_2} - k_{B_2}) * (k_{G_1} - k_{B_1})} \quad (3-39)$$

其中  $k_{R_1}$ 、 $k_{B_1}$ 、 $k_{G_1}$  分别表示相邻两点中一点的不同颜色的本体反射量, 其具体定义见公式(3-12)中的定义。

所以此时  $Q_1$  仅与体反射率有关。同理可得  $Q_2$ ,  $Q_3$  在不同色光照射下能够得到与无关于照明光线而仅与物体本身特性有关的常量。

### 3.4 实验

为了为了检验各种颜色模型的颜色特性, 采用AOIL图像数据库<sup>[30]</sup>, 它包含有一千个小物体的彩色图像, 图 3-1 是数据库中的部分物体图示。通过系统的变化视角, 光照角度, 可以得到一系列标准图像。图 3-3 所示图像库中的图像在不同角度和强度下的物体图像。



图 3-1 图像库中部分物体图像示例

图 3-2 表示了物体灯光照明的变化情况以及摄像机的角度变化，其中光源 1 到 5 即对应于图 3-3 来说分别对应于 11~15 的图像，16 将右边的两个灯打开，17 是将左边的两个灯打开，18 对应的将所有的灯打开的情况。c1~c3 对应的是图 3-2 中摄像机的不同位置。由此可以得到的图 3-3 中的 24 种不同光照情况下的图像。

图 3-4 表示了物体在不同照明光源色温变化下的图像，照明光色温变化从 2 1 7 5 到 3 0 7 5 k 变化，对应数据库中的代号如图 3-4 所示。

图 3-5 表示了物体分别向左和向右偏  $15^\circ$ ，在加上正中放置，成为一组 3 幅的图像。

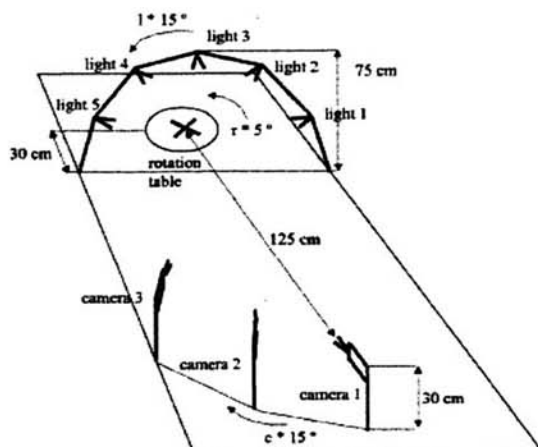


图 3-2 图像采集的试验设置



图 3-3 AOIL 图像数据库中的一组不同光照方向下的图像

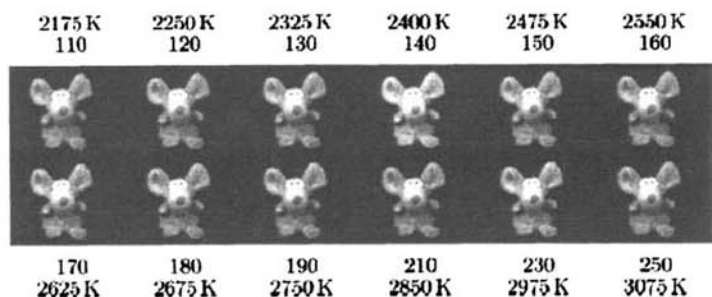


图 3-4 AOIL 图像数据库中的一组不同光源照射下的图像

### 3.5 试验结果分析及讨论

#### 3.5.1 不同照明情况以及摄像机的角度变化下的试验结果

通过第 3 节所提出的颜色模型对第 4 节中的图像库中的图像进行变换，得到的结果如下：

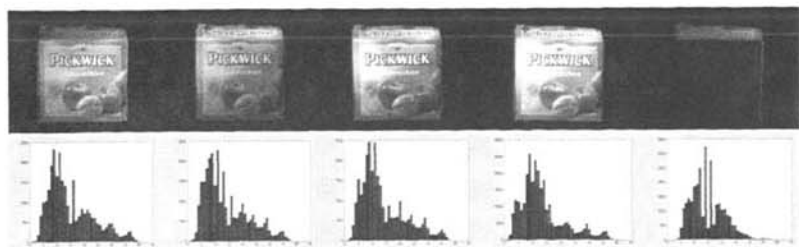


图 3-5 11c1, 14c1, 14c2, 14c3, 11c3 经过式(3-17)变换后的直方图

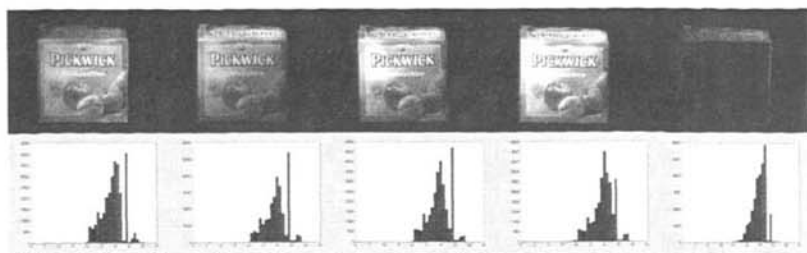


图 3-6  $l1c1$ ,  $l4c1$ ,  $l4c2$ ,  $l4c3$ ,  $l1c3$  经过式(3-20)变换后的直方图

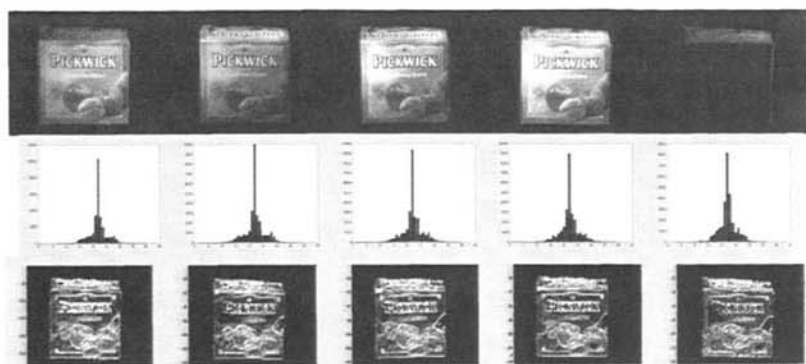


图 3-7  $l1c1$ ,  $l4c1$ ,  $l4c2$ ,  $l4c3$ ,  $l1c3$  经过式(3-36)变换后的直方图

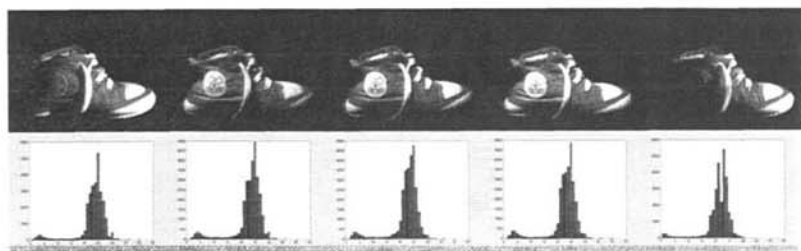


图 3-8  $l1c1$ ,  $l4c1$ ,  $l4c2$ ,  $l4c3$ ,  $l1c3$  经过式(3-17)变换后的直方图

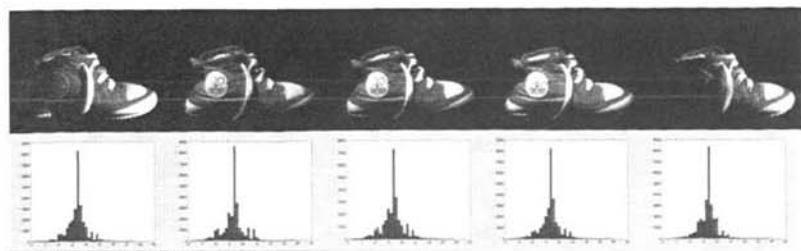


图 3-9  $l1c1$ ,  $l4c1$ ,  $l4c2$ ,  $l4c3$ ,  $l1c3$  经过式(3-36)变换后的直方图



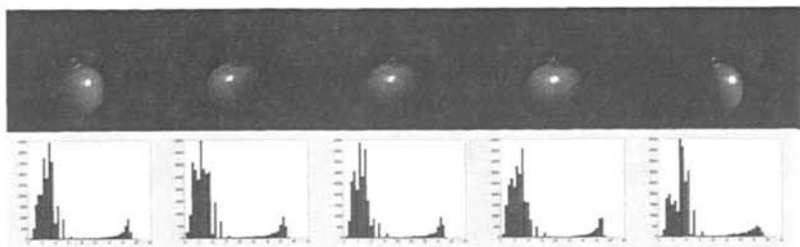
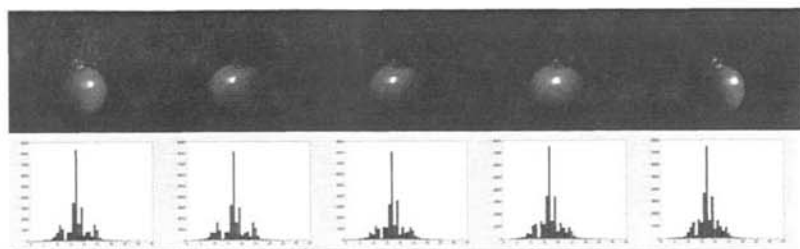
图 3-10  $11c1$ ,  $14c1$ ,  $14c2$ ,  $14c3$ ,  $11c3$  经过式(3-17)变换后的直方图图 3-11  $11c1$ ,  $14c1$ ,  $14c2$ ,  $14c3$ ,  $11c3$  经过式(3-36)变换后的直方图

图 3-5 是一组图像在经过颜色空间  $r$ 、 $g$ 、 $b$  变换后对其分量  $r$  作的直方图，从以上直方图特征来看不是很稳定，和我们之前推导出来的理论结果相对应， $r$ 、 $g$ 、 $b$  颜色空间可以在去除一定的光照影响，但是实际中实验的效果并不是很好。图 3-6 是一组图像在经过 HIS 空间变换后对  $H$  分量作的直方图，从直方图及其统计特性分析得知，HIS 空间也是光照稳定的，而且对于高光图片有较好的效果。图 3-7 是用本文所提出的颜色空间所作的直方图，可以看出直方图稳定性较好，对不同光照有较好的去除。

图 3-8 和图 3-9 是对一组球鞋作的直方图对比，这是一组具有多层次以及多种颜色的物体，可以看出本文提出的颜色模型能够适用于此种类型物体。

图 3-10 和图 3-11 是一组红色为基调的物体，通过直方图对比可以看出本文所提出的颜色模型可以适用于多种颜色的情况。

通过对不同图像应用同一颜色空间处理后的直方图的对比可知，不同物体的直方图有着较为明显的差异性可以用来作为物体识别的标准。

表 3-1 不同颜色空间直方图的统计特性

	序号	均值	方差	偏度	峰度	能量
Q1,Q2,Q3 模型 (Q1 分量)	L1c1	0.50572	0.083623	0.066487	1.7556	0.001432
	L1c2	0.49576	0.088394	0.000906	1.7911	0.001555
	L1c3	0.54561	0.084164	0.000841	1.7961	0.010342
	L2c1	0.48747	0.090945	-0.01239	1.8137	0.001464
	L2c2	0.49615	0.084959	0.092960	1.7565	0.001579
	L2c3	0.49590	0.088355	0.001238	1.8128	0.0017195
	L3c1	0.50815	0.081951	0.075871	1.7375	0.0016173
	L3c2	0.48425	0.093224	0.0021188	1.8074	0.0017287
	L3c3	0.50247	0.083615	0.078663	1.7673	0.0016671
	L4c1	0.48657	0.093034	-0.005284	1.7882	0.0016517
	L4c2	0.50504	0.08465	0.054337	1.7504	0.0017004
	L4c3	0.48767	0.092711	-0.023012	1.8137	0.0017456
r、g、b 模型 (r 分量)	L1c1	0.56571	0.037476	-0.74078	2.8892	0.00014456
	L1c2	0.38088	0.013077	0.29287	3.9238	0.00003455
	L1c3	0.29497	0.027332	0.93333	2.3661	0.00001655
	L2c1	0.61081	0.026894	-0.56816	3.0808	0.00098465
	L2c2	0.43900	0.027914	0.90088	4.2150	0.00046543
	L2c3	0.33969	0.050518	1.62450	4.4052	0.00068745
	L3c1	0.60934	0.031769	-0.91043	3.6756	0.00005856
	L3c2	0.60702	0.032877	0.00555	3.1707	0.00000654
	L3c3	0.62651	0.033144	-0.18627	3.0443	0.00043366
	L4c1	0.56097	0.035662	-0.10406	2.3973	0.00073018
	L4c2	0.65555	0.044001	-0.50484	2.5954	0.00060153
	L4c3	0.70732	0.031704	-1.71380	5.7679	0.00034605

表 3-1 中  $L_{ij}$  表示图 3-3 中的第  $i$  列  $j$  行的图像序号，均值、方差、偏度、峰度和能量如下定义，用来表示颜色模型直方图的统计特性<sup>[31]</sup>。表 3-1 所统计的数据是针对图 3-5 以及 3-7 所示的直方图统计结果。由表 3-1 中的统计数据可知在不同的光照情况下 Q1，Q2，Q3 模型中的 Q1 分量的直方图统计特性稳定，均值、方差变化较小，说明 Q1 分量的直方图分布在光照变化下基本保持不变，而通过与之对比的 r，g，b 模型的 r 分量的直方图统计特性的比较可知，r 分量均值差异较

大, 方差不稳定(如表 3-1 统计数据可示), 所以在光照变化强烈的情况下r分量的直方图无法作为物体识别的特征。通过以上讨论, 本文提出的颜色模型对光照变化有较好的鲁棒性。

$$\text{均值: } \mu = \sum_{b=0}^{L-1} b \frac{N(b)}{M} \quad (3-38)$$

$$\text{方差: } \sigma^2 = \sum_{b=0}^{L-1} (b - \mu)^2 \frac{N(b)}{M} \quad (3-39)$$

$$\text{偏度: } \eta_1 = \frac{1}{\sigma^3} \sum_{b=0}^{L-1} (b - \mu)^3 \frac{N(b)}{M} \quad (3-40)$$

$$\text{峰度: } \eta_2 = \frac{1}{\sigma^4} \sum_{b=0}^{L-1} (b - \mu)^4 \frac{N(b)}{M} \quad (3-41)$$

$$\text{能量: } \eta_3 = \sum_{b=0}^{L-1} \left( \frac{N(b)}{M} \right)^2 \quad (3-42)$$

其中 L 为量化层数, b 为量化层的值, N(b)为颜色值为 b 的像素数, M 为像素总数。均值用于描述直方图的整体取值范围, 方差用于描述直方图的相对平滑程度。偏度用于描述分布偏离与对称情况的度量, 峰度用于描述分布是在均值附近还是在两端。

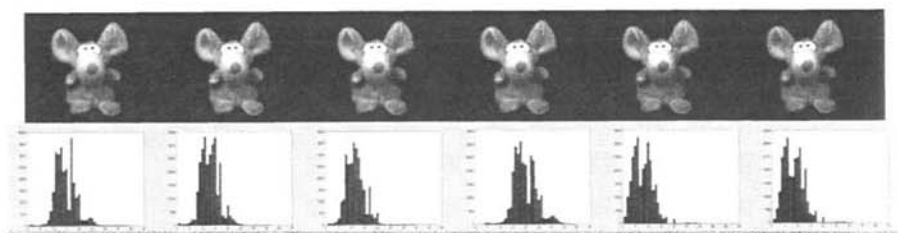


图 3-12 不同色温光源下照射的物体经过式(3-17)变换后的直方图

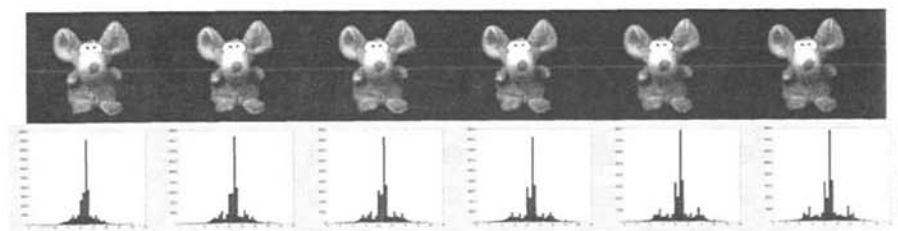


图 3-13 不同色温光源下照射的物体经过式(3-36)变换后的直方图

大, 方差不稳定(如表 3-1 统计数据可示), 所以在光照变化强烈的情况下r分量的直方图无法作为物体识别的特征。通过以上讨论, 本文提出的颜色模型对光照变化有较好的鲁棒性。

$$\text{均值: } \mu = \sum_{b=0}^{L-1} b \frac{N(b)}{M} \quad (3-38)$$

$$\text{方差: } \sigma^2 = \sum_{b=0}^{L-1} (b - \mu)^2 \frac{N(b)}{M} \quad (3-39)$$

$$\text{偏度: } \eta_1 = \frac{1}{\sigma^3} \sum_{b=0}^{L-1} (b - \mu)^3 \frac{N(b)}{M} \quad (3-40)$$

$$\text{峰度: } \eta_2 = \frac{1}{\sigma^4} \sum_{b=0}^{L-1} (b - \mu)^4 \frac{N(b)}{M} \quad (3-41)$$

$$\text{能量: } \eta_3 = \sum_{b=0}^{L-1} \left( \frac{N(b)}{M} \right)^2 \quad (3-42)$$

其中 L 为量化层数, b 为量化层的值, N(b)为颜色值为 b 的像素数, M 为像素总数。均值用于描述直方图的整体取值范围, 方差用于描述直方图的相对平滑程度。偏度用于描述分布偏离与对称情况的度量, 峰度用于描述分布是在均值附近还是在两端。

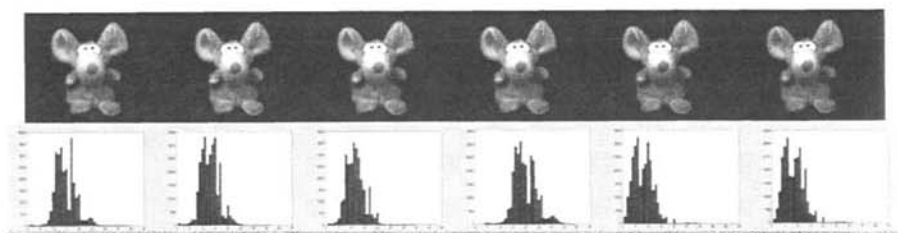


图 3-12 不同色温光源下照射的物体经过式(3-17)变换后的直方图

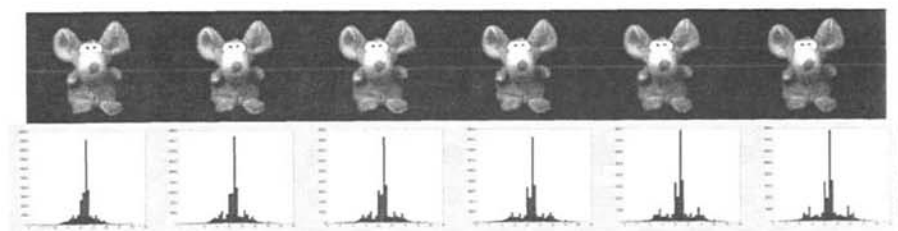


图 3-13 不同色温光源下照射的物体经过式(3-36)变换后的直方图

L4c3	66%	90%	89%	84%
L5c1	62%	82%	90%	88%
L5c2	58%	79%	75%	80%
L5c3	51%	73%	71%	76%

表 3-2 显示的是在不同的光照情况下（光照情况如图 3-2、3-3 所示）实验所得到的图像识别率，可以看出对于光照强度以及视角变化等，除了 r、g、b 颜色模型，在其他的颜色空间对这些变化均有较好的鲁棒性。本文所提出的颜色模型在不同光照下有较好的鲁棒性，但是在 l1c3 的光照情况下图像识别率较差。其原因是因为在图像光照强度低的情况下，图像的 R、G、B 分量趋于一致，当代入到颜色不变量公式 (3-36) 中，尤其是在  $(R=G=B)$  会导致颜色模型没有定义，造成数据失真，从而使得识别率下降。

表 3-3 不同光源变化下的颜色模型识别率

	r、g、b	H、I、S(H)	L1、l2、l3	Q1, q2, q3
I120	80%	85%	86%	95%
I130	72%	86%	85%	96%
I140	74%	80%	78%	93%
I150	69%	81%	76%	95%
I160	67%	74%	77%	91%
I170	64%	78%	77%	91%
I180	66%	77%	72%	94%
I190	60%	73%	75%	96%
I210	55%	71%	74%	90%
I230	55%	72%	72%	94%
I250	52%	70%	71%	90%

表 3-3 所示为物体在不同光源色温照射下不同颜色模型的识别率，通过所得数据看出本文所提出的颜色模型相比于其他的颜色空间对光源变化有较强的鲁棒性，可以有效的提取出颜色不变量信息。而其他几类颜色空间受光源变化影响较大。

表 3-4 不同颜色模型在各种光照变化下的稳定性

	r、g、b	H、I、S	L1、I2、I3	Q1, q2, q3
视角变化	√	√	√	√
高光	×	√	√	√
光照强度	√	√	√	√
光源变化	×	×	×	√

表 3-4 表示了通过实验验证后的不同颜色模型对各种光照变化下的稳定性，可以看出，本文所提出的颜色空间对各种不同的变化均有较好的稳定性。r、g、b 空间对高光以及光源变化后图像的识别均不理想，而提出的 HIS 空间和 I1、I2、I3 空间能够去除一定的高光影响，但对光源变化后的识别效果不佳，而本文提出的颜色模型能在一定程度上去除光照角度，高光以及光源变化的影响。

3.6 小结

本章根据 Shafer 二色反射模型，对基于颜色比率的颜色不变量模型进行了改进，通过新构建的颜色模型，去除掉表面反射分量，得到本体反射的分量，从而可以得到无关于光照和几何特性无关的本体反射率。即仅与物体颜色反射特性相关的不变量信息 q1、q2、q3，通过与其他颜色模型(如 r、g、b 模型 HIS 模型等)的对比实验结果的分析，可以得知本文所提出的色彩模型在不同的光源照射下以及不同光照强度下的稳定性有一定的提高。

## 第四章 空间颜色直方图

### 4.1 引言

基于颜色直方图的识别方法最早由Swain和Ballard提出，其计算简单，对图像的平移和旋转具有不变性。但是颜色直方图反映的是颜色的全局统计信息，丢失了颜色的空间分布信息。改进颜色直方图特征的方法包括：累加直方图方法，通过累加方式增加了直方图的鲁棒性，但仍丢失了空间分布信息；可视权重直方图方法，将颜色空间由RGB转换到huv空间，量化成少数可视相似的代表色，以与这些代表色距离的倒数为权重，得到权重直方图，但也没有颜色的空间分布信息。为了在基于颜色直方图的方法中增加空间信息，提出了很多方法，主要分三类：其一是直接反映空间信息，例如颜色相关向量CCV(color coherence vector)方法<sup>[32]</sup>和颜色相关图(color correlograms)<sup>[33]</sup>方法，取得较好的效果，但它们计算复杂，对图像的平移和旋转敏感；其二是采用分块方法<sup>[32]</sup>，通过将图像分成图像块来反映空间信息，但图像块与位置有关，对平移和旋转敏感；其三是采用分割方法<sup>[33]</sup>，按照某种特征将图像分割成不同的区域，再利用各个区域的颜色、形状等特征进行检索，这种方法用得较多，有不少取得较好结果。但由于分割本身是一个难题，分割的结果也很难准确。最近提出的方法有合并直方图MHM(merged histogram method)<sup>[34]</sup>方法，将小于给定可视距离范围内的颜色分别进行同一颜色通道内及不同颜色通道间的合并，以保留较少的主要颜色，再利用直方图相交方法进行检索。该方法数据存储量小，检索速度快，但也没有反映颜色的空间分布信息。

### 4.2 空间直方图方法概述

#### 4.2.1 局部颜色直方图和颜色对

全局颜色直方图虽然具有计算简单，对平移和旋转不敏感的优点，但无法捕捉颜色组成之间的空间关系，丢失了图像的空间信息。这无疑是不够准确的。两幅颜色直方图非常相似的图像其内容可能丝毫无相似之处。如图 4-1 所示，几幅图像的颜色直方图基本上是一致的，而从图像看他们并不是同一类物体。

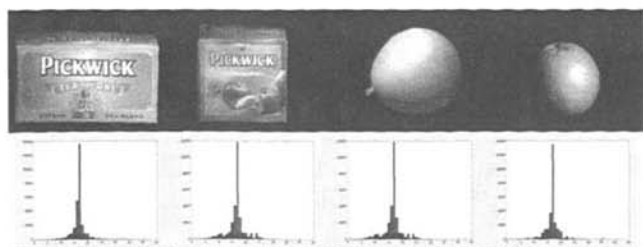


图 4-1 主色相近的图像颜色直方图

因此,在基于颜色的图像检索中引入空域的信息对于确保检索精度是十分重要的。事实上,颜色的不同空间分布极大地影响了人们对图像的相似性判断。鉴于此,解决方法之一就是基于图像空间的固定划分,即人为地将图像划分成适当的分块,然后为每个分块提取相应的局部颜色特征。我们知道分块太大则失去分块的意义,若分块太小则会增加检索过程的计算量。我们可以划分更多的子块,划分的子块越多,图像的距离矩阵携带的空间信息越完备,检索精度也越高,计算复杂度也越高。当然如果考虑图像中像素间的相关性,使得图像对小范围的旋转不敏感,可以采用重叠分块,分块间的重叠区域取为分块尺寸的 $1/4$ 。在分块的基础上,可以进一步根据子块直方图来进行相似度判定,如下是四种基本的检索算法:

- 1) 基于平均色。如果直接将色彩直方图建立索引<sup>[38]</sup>,索引的结构就会变成很高的维数(例如 64 或 256 维)。为了避免“维数的灾难”,我们可采用色彩直方图的三维压缩描述—平均色彩—作为一种较为简便的方式来摹拟全n维直方图的比较。也可用平均色彩比较过滤数据库,然后运用全色彩直方图对过滤下来的显著变小的图像集合进行比较。假设每一个存储色彩直方图的文件表示了一个三维色彩值,那么色彩直方图的平均色彩就可以定义为对应于标准色彩直方图的加权平均色彩。特别地,令  $C = [c_1, c_2, \dots, c_n]$  为一个三阶矩阵,该矩阵的第i列为颜色  $c_i = [\alpha_i, \beta_i, \gamma_i]^T$  分别表示三维色彩的各分量的大小。假设两个n维色彩直方图X和Y,它们对应的  $3 \times 1$  平均色彩矢量分别为:  $X_{ave} = CX$ ,  $Y_{ave} = CY$ 。定义平均色彩距离的平方为:

$$d_{ave}^2 = (X_{ave} - Y_{ave})^T (X_{ave} - Y_{ave}) = (X - Y)^T C^T C (X - Y) \quad (4-1)$$



整个图像的距离为： $\sum_{n=1}^N (d_{ave}^2(n))$ ，其中  $N$  是划分的子图的总数。虽然

平均色彩比较并不如全  $n$  维直方图准确，但是却具有较快的速度。

- 2) 基于主色匹配<sup>[36]</sup>。可以更进一步地简化索引的计算，仅仅找出直方图中占绝大多数的颜色，作为直方图的主色，定义距离的平方为：

$$d_{main}^2 = (X_{main} - Y_{main})^T (X_{main} - Y_{main})，整幅图像距离为 \sum_{n=1}^N (d_{main}^2(n))。从而$$

有效去除了不重要物体带来的颜色干扰，计算简单，节省特征存储空间。

- 3) 直接基于局部颜色直方图。当然也可以直接计算局部直方图全色彩的距离： $d^2 = (X - Y)^T (X - Y)$ ，同样计算所有的子图  $\sum_{n=1}^N (d^2(n))$ 。

- 4) 基于颜色变化比较。同时，还可以从颜色变化的角度考虑，任一子图与自身相邻的所有子图作平均色的比较，求得距离和作为本子图的特征值。若  $m \times m$  划分的图像，则结果保留在  $m \times m$  的二维数组中，作为特征向量，表征了图像子图之间的相似度。两幅图像的相似距离由计算特征向量的欧几里得距离得到。因此，如果一幅图像只有细微的颜色变化的话，它的特征向量数组的值将接近于 0，反之如果颜色变化较大的话，将接近于 1（经过归一化处理）。图 4-2 显示了  $4 \times 4$  分块的图像计算特征向量数组时考虑的三种情况。

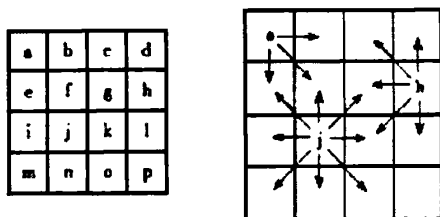


图 4-2 计算特征向量组

#### 4.2.2 加权颜色直方图

基于加权颜色直方图的图像检索算法，即将图像分割成  $(m \times n)$  个子块，这样就提供了一定程度的位置信息，而且对用户感兴趣的子块加大权重，提高检索的查准查全率。设  $G$  和  $S$  表示要进行比较的两幅图像， $N$  表示图片中出现的颜色级数， $g_k$  和  $S_k$  分别表示图像  $G$  和  $S$  中  $k$  级色出现的频数，则两幅图像的相似度可用如下公式表示：

$$Sim(G, S) = \frac{1}{N} \sum_{k=1}^N \left( 1 - \frac{|g_k - s_k|}{\max\{g_k, s_k\}} \right) \quad (4-2)$$

Sim 值越接近 1, 两幅图像越相似, 图 G 与 S 中的对应子图  $G_{ij}$  与  $S_{ij}$  间的相似性度量:

$$Sim_{ij}(G_{ij}, S_{ij}) = \frac{1}{N} \sum_{k=1}^N \left( 1 - \frac{|g_k - s_k|}{\max\{g_k, s_k\}} \right), (i=1, 2, \dots, m; j=1, 2, \dots, n) \quad (4-3)$$

其中 N 表示图中出现的颜色级数,  $g_k$  和  $s_k$  分别表示  $G_{ij}$  与  $S_{ij}$  中 k 级色出现的频数. 若引入子块权重  $W_{ij}$  则图 G 与 S 的相似度:

$$Sim(G, S) = \sum_{i=1}^m \sum_{j=1}^n W_{ij} \times Sim_{ij}(G_{ij}, S_{ij}) \quad (4-4)$$

其中  $W_{ij}$  的选取应根据图像的特点决定, 通常图像中间子块的权重较大, 或者用户指定区域的权重较大, 以反映图像的位置信息。

加权颜色直方图图像检索算法描述如下:

1) 将目标图像库中每一个图像文件分成  $(m \times n)$  子图像  $G_{ij}$  提取特征(局部颜色直方图), 形成特征库;

2) 计算检索图像 S 中各子图像  $S_{ij}$  的局部颜色直方图, 然后计算对应子图  $G_{ij}$

与  $S_{ij}$  间局部相似性  $Sim_{ij}(G_{ij}, S_{ij}) = \frac{1}{N} \sum_{k=1}^N \left( 1 - \frac{|g_k - s_k|}{\max\{g_k, s_k\}} \right)$  最后考虑子块权

重  $W_{ij}$  计算检索图片 S 和图片库中任一图片 G 的全图直方图相似值

$$Sim(G, S) = \sum_{i=1}^m \sum_{j=1}^n W_{ij} \times Sim_{ij}(G_{ij}, S_{ij});$$

3) 给定阈值  $\delta (0 < \delta < 1)$ , 将全图直方图相似值  $Sim(G, S) > \delta$  的图片 G 作为检索结果。

#### 4.2.3 相交颜色直方图

定义 1: 归一化直方图

设 H 为一个直方图,  $H(r)$  为 H 的一个分量, 如果满足  $\sum_r H(r) = 1$ , 称 H 为归一化直方图。如果一个直方图 H 不是归一化直方图, 可以用下式归一化:

$$H(r) = h(r) / \sum_r h(r) \quad (4-5)$$

定义 2: 归一化直方图的交

设  $H_1$  和  $H_2$  为两个归一化直方图, 则其交定义为:

$$Sim(H_1, H_2) = \sum_r \min(H_1(r), H_2(r)) \quad (4-6)$$

Swain和Ballad<sup>[37]</sup>的直方图相交算法就是根据  $Sim(H_1, H_2)$  值的大小来衡量两图像(其颜色直方图分别为  $H_1$  和  $H_2$ ) 的相似程度, 当其值为 1 时, 两图像最相似, 为 0 时, 两图像最不相似。

实际上, 直方图相交算法的基本思想是取两图像所有颜色的交集, 同时考虑了背景颜色。因此, 当两图像中占较大的背景颜色并不相同时, 即使占较小比例的前景目标完全相同, 其直方图的交集也不会很大, 导致检索结果不理想。由此看出, 如果是应用于背景颜色都相同的图像数据库检索、或应用于类似找出“红色占 30%, 绿色占 40% 的所有图像”的检索, 直方图相交算法非常有效。然而在实际当中, 往往要求检索含有某一个目标对象的一类图像, 如一朵花、一株草或者一条狗等, 这时候, 就必须考虑图像中背景颜色不同却含有相同目标对象的情况。因此, 有必要对直方图相交算法加以补充。

彩色图像中颜色的数目非常多, 如果颜色直方图匹配过程中对图像中所有的颜色都进行考虑, 其计算量必然巨大。经实验分析, 图像中的目标对象, 其占主要的颜色一般只有少数的几种或几十种, 为此, 我们利用颜色聚类<sup>[38]</sup>技术, 提出了一个类似Mehltre的参考颜色表<sup>[39]</sup>的基于示例图像的主颜色表算法。该算法为示例图像建立一个颜色表, 并将其作为主颜色表, 被检索图像的颜色直方图根据此主颜色表来建立, 不在主颜色表中出现的颜色被忽略。算法中, 主颜色表是属于示例图像的, 示例图像不同, 检索时采用的主颜色表也不同, 因而是一种动态的参考颜色表, 而Mehltre的参考颜色表是静态的。

我们知道, 如果示例图像和被检索图像中含有不相同的颜色, 那么这些不相同的颜色必定不是组成示例图像中要检索的目标对象的颜色, 或者说被检索图像中不含有要检索的目标。此外, 相同的目标对象在不同的图像中出现, 组成该目标对象的颜色应该是基本成比例的, 如果两图像归一化直方图中具有比例相差很大的颜色, 那么该颜色必定为杂色, 或者说其中一幅图像的这种颜色大部分属于杂色。依据这两个事实, 我们建立两图像的相交颜色直方图。

设示例图像为  $Q$ , 其主颜色表归一化直方图为  $H_Q$ 。被检索图像为  $J$ , 根据  $H_Q$

所建立的归一化颜色直方图为  $H_j$ ，则算法可描述为：

- (1) 过滤掉  $H_j$  中  $H_Q(r)$  与  $H_j(r)$  比例相差若干倍的颜色  $r$ ，本节算法中取 0.3，即：若  $\min(H_Q(r), H_j(r)) / \max(H_Q(r), H_j(r)) < t$ ，则  $H_j(r) = 0$   
其中  $H_Q(r) \in H_Q, H_j(r) \in H_j, t = 0.3$ ，处理后的  $H_j$  设为  $H'_j$ 。
- (2) 对  $H_Q$  过滤掉没有在  $H'_j$  中出现的颜色，得到  $H'_Q$ 。
- (3) 根据直方图相交算法计算  $H'_Q$  与  $H'_j$  的相交距离，设为  $Sim$ 。

### 4.3 基于模糊分类的图像颜色直方图研究

#### 4.3.1 模糊分类基本思想

颜色直方图用于表示图像中颜色分布。横轴表示颜色等级，纵轴表示在某个颜色等级中的像素在整幅图像中所占比例。基本构造思想可抽象为如下通用分类统计：

假定横轴被按序排列的  $n+1$  点  $\{x_0, x_1, \dots, x_i, \dots, x_n\}$  划分为  $C = \{C_0, C_1, \dots, C_i, \dots, C_n\}$   $n$  个抽象类，每个类  $C_i$  的取值区间为  $[x_{i-1}, x_i], 1 \leq i \leq n$ 。假定样本集  $\Omega$  中存在  $M$  个采样点，需要被划分到这个  $n$  类中。

对任意的点  $x \in \Omega$ ，如果  $x_{i-1} \leq x < x_i$ ，则有  $x \in C_i$ ， $x$  对应的隶属向量<sup>[40]</sup>为  $c^x = \{(c_1^x, c_2^x, \dots, c_k^x, \dots, c_n^x) \mid \begin{cases} c_k^x = 1 & k = i \\ c_k^x = 0 & \text{else} \end{cases}\}$ 。同时，样本集合  $\Omega$  的归一化统计直方图特征向量定义为：

$$H = \{H_1, H_2, \dots, H_i, \dots, H_n\} \quad (4-7)$$

$$\text{其中, } H_i = \frac{1}{M} \sum_{x \in \Omega}^{x_{i-1} \leq x < x_i} 1。$$

显然这是一种诸如“是”或“不是”的“0、1”分类统计方式。如前所述，人类对颜色的认识和分类和对很多其他客观事物一样，无法简单地以“0”或“1”方式给出。受模糊数学思想启发，我们认为使用模糊分类<sup>[41]</sup>更能恰当地描述人类对颜色的认知。图 4-3 表示某一样本的不同隶属分类情况。图 4-3(a)表示精确分类。对某个给定的样本  $x$ ，有且仅有一个  $i$ ，表征仅存在一个类别  $C_i$ ，使得  $x$  的隶属向量  $c^x$  对应  $c_i^x = 1, c_k^x = 0, 1 \leq k \leq n, \text{ if } k \neq i$ 。但如图 4-3(b)所示，按照模糊分类定义，却可能存在多个  $i$  的取值构成的子集  $\Delta$ ，使得多个类别  $C_i$  同时满足  $c_i^x \neq 0, 0 < c_i^x < 1, i \in \Delta$  以及  $\sum_{i \in \Delta} c_i^x = 1$ 。

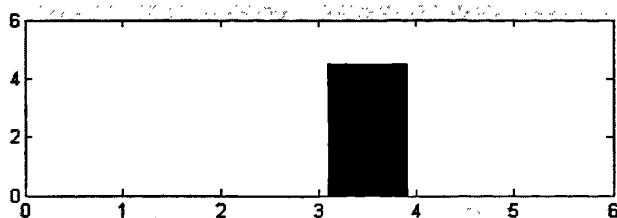


图 4-3(a)精确分类

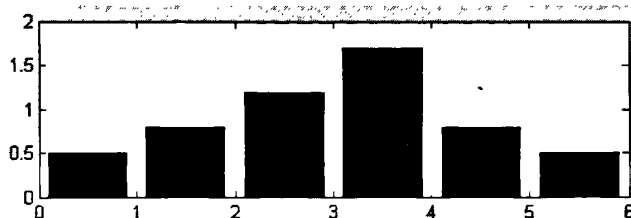


图 4-3(b)模糊分类

图 4-3 精确分类与模糊分类

对于模糊分类而言,  $c_i^x$  表明的实际上是样本  $x$  隶属于类别  $C_i$  的可能性或确信度, 这种模糊分类的表示方法直观上更符合上述人类视觉系统对颜色的认知规律。比如当我们无法精确指出某种程度的灰色到底是应该划为白色还是黑色时, 用一个二维特征向量  $f = (f_b, f_w)$  来描述灰色的程度应该是更好的方法。

### 4.3.2 模糊颜色直方图特征向量构造

假定彩色图像  $G$  宽  $W$  高  $H$ , 下面描述如何通过  $W \times H$  个像素点的模糊隶属向量  $c^x$  的统计, 最终得到表征颜色内容特征的模糊直方图特征向量  $f$ 。

首先, 将图像从RGB颜色空间转化到由色调、饱和度和亮度三元组  $(h, s, v)$  构成的颜色空间, 在颜色空间计算模糊颜色直方图。原因在于颜色空间的距离度量比颜色空间中的距离度量能更逼近地反映人类对画面内容差异的视觉感知。其一, 人类视觉可以独立地感知色调、饱和度, 亮度分量的变化; 其二, 空间中两点之间的欧氏距离<sup>[42]</sup>所代表的颜色之间的差异与人类主观感知的心理颜色差异具有近似线性的比例关系。

然后, 同上也讲图像  $G$  在纵横轴均匀分割为  $M \times N$  个子图像块  $G_\gamma$ ,  $1 \leq \gamma \leq M \times N$ , 每个子图像  $G_\gamma$  的宽高分别为  $W/M$  和  $H/N$ 。  $M, N$  的取值依

赖于  $W \times H$ 。每个子图块  $G_i$  考虑如下过程：

将 HSV 3 个不同颜色通道的取值范围分别量化为  $n^h, n^s, n^v$  个区间段。实验研究表明人类视觉对亮度、色调、饱和度各分量变化的敏感程度不一，因此针对不同的 HSV 颜色通道，我们采用非等间隔区间量化的方式，得到独立于任何具体图像的 3 个独立颜色通道上的区间段量化分类：

$$C^h = \{C_1^h, C_2^h, \dots, C_i^h, \dots, C_{n^h}^h\}, \quad i=1, 2, \dots, n^h \quad (4-8)$$

$$C^s = \{C_1^s, C_2^s, \dots, C_i^s, \dots, C_{n^s}^s\}, \quad i=1, 2, \dots, n^s \quad (4-9)$$

$$C^v = \{C_1^v, C_2^v, \dots, C_i^v, \dots, C_{n^v}^v\}, \quad i=1, 2, \dots, n^v \quad (4-10)$$

假定图像块  $G_i$  中任意像素点  $P$  在 HSV 颜色空间的 3 个不同颜色通道中取值为  $x=(x^h, x^s, x^v)$ 。以 H 通道中的  $x^h$  取值为例，讨论如何依据  $P$  的该分量将  $P$  模糊划分到  $C^h$  的  $n^h$  区间分类中：

为  $n^h$  个类别  $C_i^h$ ， $1 \leq i \leq n^h$  分别指定一个像素值  $x$  的  $x^h$  分量的值隶属置信度函数：

$$F_i^h(x), \quad 1 \leq i \leq n^h \quad (4-11)$$

任给一个  $x^h$ ，同时计算  $n^h$  个隶属函数  $F_i^h(x^h)$ ，得到像素点  $P$  依据  $x^h$  的取值隶属于  $C^h$  中所有  $n^h$  个类别  $C_i^h$ ， $1 \leq i \leq n^h$  的  $n^h$  维置信度或隶属向量值

$$C^{x^h} = \{C_1^{x^h}, C_2^{x^h}, \dots, C_i^{x^h}, \dots, C_{n^h}^{x^h}\} \quad (4-12)$$

针对  $n^h$  个类别  $C_i^h$  的隶属置信度函数  $F_i^h(x^h)$  可以有多种选择，常用的有三角形函数、梯形函数和高斯型函数。本文构造的三角型隶属函数  $F_i^h(x^h)$  为公式(4-13)所示，形状如 4-3 所示。其中  $\varphi$  为单位量化区间的宽度， $x_i$  是第  $i$  个区间的中间点的取值，值域范围在  $[0, 1]$  之间。

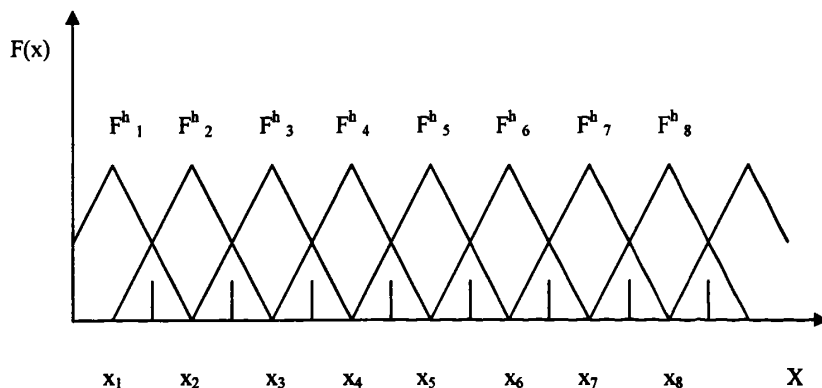


图 4-4 三角形隶属函数

$$F_i^h(x) = \begin{cases} 1 - \frac{|x - x_i|}{\varphi}, & |x - x_i| < \varphi, \\ 0, & |x - x_i| \geq \varphi \end{cases}, \quad i = 1, 2, \dots, n^h \quad (4-13)$$

选择简单的三角形隶属函数组  $F_i^h(x)$ ,  $1 \leq i \leq n^h$  进行计算, 对应图形如图 4-4 中所示。此刻  $\varphi$  表示三角形底边宽度的一半。

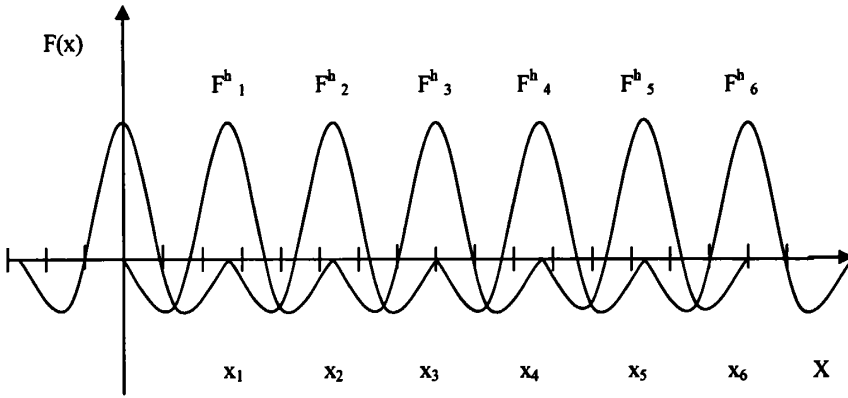


图 4-5 基于 marr 理论的隶属函数

$$F_i^h(x) = \begin{cases} \frac{-1}{\pi\sigma^4} \left(1 - \frac{(x - x_i)^2}{2\sigma^2}\right) e^{\frac{-(x-x_i)^2}{\sigma^2}}, & |x - x_i| < \varphi, \\ 0, & |x - x_i| \geq \varphi \end{cases}, \quad i = 1, 2, \dots, n^h \quad (4-14)$$

为了更好的描述人类视觉现象, 引入  $\text{marr}^{(48)}$  算子, 如图 4-5 所示, 相比于简单的三角形隶属函数, 基于  $\text{marr}$  的计算视觉理论的隶属函数能够更好的表示出人类的视觉现象, 更接近于实际的情况。函数解析式为公式 4-15, 其中  $\sigma$  是指高斯函数的标准差,  $\varphi$  是对应函数图形底的一半。

使用上述构造法, 可针对 HSV 颜色空间中每个独立通道, 为对应的独立类别  $C^h, C^s, C^v$  指定如下 3 组对应隶属函数:

$$F_i^h(x), \quad 1 \leq i \leq n^h \quad (4-15)$$

$$F_j^s(x), \quad 1 \leq j \leq n^s \quad (4-16)$$

$$F_k^v(x), \quad 1 \leq k \leq n^v \quad (4-17)$$

这三组函数区别仅在于定义域的细微差别, 因此不详细展开。

遍历  $G_\gamma$  中所有像素点  $P$ , 根据 3 组上述隶属函数, 从  $P$  点颜色值  $x$  在不同颜色通道中的颜色分量  $x^h, x^s, x^v$  得到分别隶属三组独立类别  $C^h, C^s, C^v$  中各个分类的 3 个置信度向量  $C^{x^h}, C^{x^s}, C^{x^v}$ 。

在此基础上按照下式生成子图像  $G_\gamma$  不同颜色通道上的模糊直方图统计向量  $H^h, H^s, H^v$ :

$$H^h = \{H_1^h, H_2^h, \dots, H_i^h, \dots, H_{n^h}^h\} \quad (4-18)$$

$$\text{其中, } H_i^h = \frac{W \bullet H}{M \bullet N} \sum_{x \in G_\gamma} c_i^{x^h} = \frac{W \times H}{M \times N} \sum_{x \in G_\gamma} F_i^h(x^h), \quad i=1,2,\dots,n^h$$

$$H^s = \{H_1^s, H_2^s, \dots, H_i^s, \dots, H_{n^s}^s\} \quad (4-19)$$

$$\text{其中, } H_j^s = \frac{W \bullet H}{M \bullet N} \sum_{x \in G_\gamma} c_j^{x^s} = \frac{W \times H}{M \times N} \sum_{x \in G_\gamma} F_j^s(x^s), \quad j=1,2,\dots,n^s$$

$$H^v = \{H_1^v, H_2^v, \dots, H_i^v, \dots, H_{n^v}^v\} \quad (4-20)$$

$$\text{其中, } H_k^v = \frac{W \bullet H}{M \bullet N} \sum_{x \in G_\gamma} c_k^{x^v} = \frac{W \times H}{M \times N} \sum_{x \in G_\gamma} F_k^v(x^v), \quad k=1,2,\dots,n^v$$

整个 HSV 空间则被 3 组独立的区间类  $C^h, C^s, C^v$ , 联合分割为  $n = n^h \cdot n^s \cdot n^v$  个量子空间, 记为  $C = \{C_1, C_2, \dots, C_l, \dots, C_n\}$ ,  $l=1,2,\dots,n$ 。则  $l$  与  $i, j, k$  可以定义如下对应关系:

$$l = n^i \cdot n^j \cdot (k-1) + n^i \cdot (j-1) + i \quad (4-21)$$

那么, 根据模糊数学乘法<sup>[4]</sup>, 在遍历像素  $p \in G_\gamma$  后, 得到图像块  $G_\gamma$  上定义的模糊直方图  $H_\gamma$  为:

$$H^\gamma = \{H_1^\gamma, H_2^\gamma, \dots, H_l^\gamma, \dots, H_{n^\gamma}^\gamma\}, \quad n^\gamma = n^h \cdot n^s \cdot n^v \quad (4-22)$$

其中,  $H_l^\gamma = H_l^h \cdot H_l^s \cdot H_l^v$ ,  $1 \leq l \leq n$ ,  $1 \leq i \leq n^h$ ,  $1 \leq j \leq n^s$ ,  $1 \leq k \leq n^v$ ,  $i, j, k$  满足公式(4-21)指出的关系。

得到所有子图像  $G_\gamma$  对应的模糊直方图  $H_\gamma, 1 \leq \gamma \leq M \times N$  后, 以类似方式得到



针对整个图像  $G$  的模糊直方图  $H = \{H_1, H_2, \dots, H_l, \dots, H_n\}$ , 与上述不同仅在于图像的大小差别。在此基础上最终通过加权平均得到如下所示的图像  $G$  的  $n$  维抽象颜色内容特征向量  $f$ :

$$f = f(f_1, f_2, \dots, f_l, \dots, f_n), \quad \text{其中 } f_l = \alpha H_l + \sum_{\forall G_i \in G} \beta_i H_l^i \quad (4-23)$$

其中  $\alpha, \beta_i$  为预先设定的可调整参数值, 反映全局模糊直方图和局部模糊直方图在特征向量构造中的比重。

#### 4.4 实验结果与性能分析

本节实验仍选取第三章第 4 节所述的数据库, 因为在上一章已详细论述数据库特征, 在此不赘述。

对于分块图像所构造的模糊直方图与整体图像的模糊直方图权值定义, 即公式(4-23), 我们设定为图像中心权值较大, 因为如数据库中的图像基本特征所示, 主要像素集中在图像的中心区域, 所以在此指定中间子块的权值为最大, 可以得到更多的空间信息, 另外图像子块大小的划分也需要符合人的视觉现象, 通过实验比较, 我们可以得到一个最佳的分割效果。图 4-6 是一些图像的划分后的分块效果。

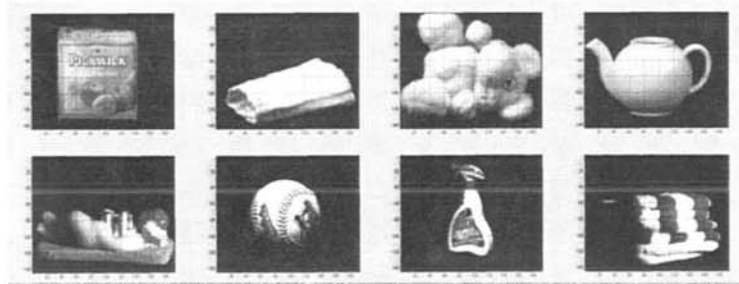


图 4-6 图像划分效果

我们对比这一章不同直方图的构造方法, 应用于数据库后所得到的识别率对比, 可以得出本章所采用的基于 marr 的计算视觉理论中所提到的墨西哥草帽函数作为隶属函数所构造的模糊直方图相比于其他的直方图构造方法有更好的识别效果, 其具体的理论依据在第 5 节中已经有详细阐述即证明。实验数据如表 4-1 所示。

实验中选择了 500 多幅真彩图作为测试图像，基本达到了预期效果。左起第一幅为待查询图。从检索结果中可以发现全局颜色直方图由于缺乏空间信息，检索出的结果不能保证图像基本的颜色分块信息。而基于分块的局部颜色直方图方法基本上能够解决一定的问题



图 4-7 应用整体直方图检测结果



图 4-8 应用局部颜色直方图的检测结果



图 4-9 应用颜色对检测的结果



图 4-10 应用模糊直方图(三角隶属函数)检测结果



图 4-11 应用模糊直方图(墨西哥草帽隶属函数)检测结果

通过以上不同检测方法结果的对比, 可以看出在应用全局直方图检测时, 由于自身的不足, 没有保留颜色的空间信息, 造成识别率较差, 而其他的几种不同的颜色直方图的改进中, 从示例的图像检索结果可以看出, 基本可以有效地去除空间形状有明显差异的图像结果; 在本例中, 由于待检索图像是一种类似于长方体的包装, 通过全局直方图, 很容易引入不同形状的而主色类似物体, 如在图 4-7 中的橙子和瓷器。而在如下其他的加入空间信息的图像检索结果中, 可以看出有效的屏蔽了这类错误。而通过模糊直方图的引入可以很好的分析图像的一些细节特征, 本文应用marr提出的视觉响应函数<sup>[46]</sup>作为模糊直方图的隶属函数, 从表 4-1 数据可以看出相对于三角形的隶属函数所构造的模糊直方图效果有一定的提升。

检索算法的优劣通常采用两个性能指标: 查准率(Precision)和查全率(Recall)<sup>[46]</sup>来评价检索算法。查准率指返回的结果集中有效图像的比率, 用于测量系统排除无关图像的能力。查全率指返回的结果中有效图像占数据库中所有相似图像的

数量的比率，用于测量系统检索相关图像的能力。查准率和查全率越高，说明该检索算法的效果越好。查准率  $P$  和查全率  $R$  可按公式  $P = \frac{R_A}{R_A + R_B}$  和  $R = \frac{R_A}{R_A + R_C}$  计算，其中  $R_A$  表示正确检索出的相关图像数<sup>[47]</sup>， $R_B$  表示检索出的无关图像数<sup>[48]</sup>， $R_C$  表示漏检的相关图像数。

表 4-1 不同直方图识别方法的平均识别率

	全局直 方图	局部颜色直 方图(加权)	颜色对	相交颜色 直方图	模糊直方图(三 角隶属函数)	模糊颜色直方图(墨 西哥草帽隶属函数)
查准率	53.4%	62.9%	67.6%	71.2%	75.6%	81.3%
查全率	59.8%	64.3%	68.2%	60.5%	70.2%	76.7%

从公式(4-23)可以看出来，在应用  $n$  维特征向量  $f$  进行图像分析比较时，距离测度简单地使用不同  $f$  间的欧氏距离，相似性测度使用欧氏距离的倒数<sup>[49]</sup>或向量余弦夹角公式<sup>[60]</sup>表征即可。人类视觉在图像感知时仅对图像中比率超过 2%~3% 以上的几十种显著颜色构成的连续子区域敏感。为进一步降低图像比较环节中次要颜色所在区间的差值影响图像间的相似测度，凸显主要颜色的作用，简单地对特征向量针对  $n$  个子空间的隶属度取值进行排序，前 50 个分量的取值不变，50 以后的对应分量值忽略为 0。由于采用模糊分类措施，次要颜色对特征向量的贡献仍然在邻近区间得以体现，并未全部抹杀。

#### 4.4.1 小结

为了解决基于颜色直方图识别方法无法提供颜色空间信息的先天不足，本章列举了现今主要的几种改进方法并作了分析，同时为解决像素统计稀疏<sup>[61]</sup>及颜色区间刚性划分<sup>[62]</sup>带来的缺陷，本章重点讨论了模糊直方图的构造。应用表明该方法在一定程度上改善了直方图特征向量在基于内容特征的图像分析应用中的效果。本文只是应用模拟人体视觉的函数来模拟人类视觉对颜色差异的感知，比“0、1”划分的效果要好，但是由于实际的颜色感知性质以及人体的视觉机理还不是很清楚，只能做近似的模拟。要想得到更好的效果，需要根据进一步实验得到更好的拟和颜色差异感知的经验曲线<sup>[63]</sup>。最后，可针对具体需求，与颜色直方图的其他优化进行组合，构成更好的颜色直方图特征向量方案。

## 第五章 总结与展望

### 5.1 总结

颜色特征的选择和利用是彩色图象识别中的关键问题之一。常用的颜色特征提取技术是颜色直方图，直方图是在研究颜色特征分析时最为重要的方法之一。而颜色特征由于同人眼的视觉认知生理心理、物理光学以及电子设备等的信息表达均有关系，所以在图象识别的研究中需要结合实际情况对它进行系统的研究和优化。图象自动识别的要求是快速地获得良好的识别结果，则选择多种有效的识别特征并对它们进行合理的优化对于提高识别结果的正确率、降低识别计算的复杂度、提高识别速度等十分重要。

本文的主要内容与贡献如下：

系统地介绍了基于颜色直方图研究的一些预备知识和若干关键技术，主要包括颜色空间、颜色直方图的构造方法，直方图与空间信息结合的主要方法，及检测算法的性能评价等。

对于颜色直方图特征的提取，本文根据 Shafer 二色反射模型，物体表面反射的过程分为两种，即表面反射和本体反射。而表面反射分量一般与入射光有相同的光谱分布，本体反射分量的光谱则是由光照的光谱功率分布和物体表面反射特性共同决定的。本体反射率代表了物体的颜色反射特性，可以表征物体颜色的本征特性。通过去除掉表面反射分量，并对本体反射的分量处理，从而可以得到无关于光照和几何特性无关的本体反射率。即仅与物体颜色反射特性相关的不变量信息，对二色反射模型的颜色模型进行了改进，通过对试验结果的分析，本文所提出的色彩模型在不同光照角度和强度下的稳定性有一定的提高。

为了解决基于颜色直方图识别方法无法提供颜色空间信息的先天不足，本文列举了现今主要的几种改进方法并作了分析，同时为解决像素统计稀疏及颜色区间刚性划分带来的缺陷，本章重点讨论了模糊直方图的构造。应用表明该方法在一定程度上改善了直方图特征向量在基于内容特征的图像分析应用中的效果。本文只是应用模拟人体视觉的函数来模拟人类视觉对颜色差异的感知，比“0、1”划分的效果要好，并且与颜色直方图的其他优化进行组合，构成更好的颜色直方图特征向量方案。

## 5.2 展望

在本论文的研究基础上，有待在下述几个方面进行进一步研究：

1. 对于颜色直方图特征的提取，本文在应用 Shafer 二色反射模型进行颜色不变量提取时，由于本文所提出的颜色空间的不足，对于灰度图象，以及彩色图像中的近灰度点进行变换时会造成较大误差，在以后的研究工作中需要进一步改进，以减少影响。
2. 本文提出了应用近似人类视觉特性的函数作为构造模糊直方图的隶属函数，但是由于实际的颜色感知性质以及人体的视觉机理还不是很清楚，只能做近似的模拟。要想得到更好的效果，需要根据进一步实验得到更好的拟和颜色差异感知的经验曲线。

## 参考文献

- [1] Novak, Carol L., Shafer Steven A., Supervised color constancy for machine vision Proc. SPIE Vol. 1453, 353-368, 1991
- [2] 刘忠伟, 利用局部累加直方图进行彩色图象检索, 中国图象图形学报, 3(7) 533-537, 1998
- [3] 金韬, 任秀丽, 图像检索中颜色特征的提取与匹配, 计算机辅助设计与图形学学报, Vol. 12 No. 6 459-462, 2000
- [4] 刘建峰, 李春茂, 基于小波变换的彩色自然图像数据库自动检索, 红外与毫米波学报, Vol. 19 No. 129~32, 2000
- [5] 王伟, 武德峰, 基于模糊分类的图像颜色直方图研究, 模糊系统与数学, Vol. 17 No. 4 94-98, 2003
- [6] S. A. Shafer, Using color to separate reflection components, COLOR Res. Appl. 10(4). 210-218. 1985
- [7] 荆其诚, 色度学, 科学出版社, 10, 1979
- [8] C. A. Rothwell, A. Zisserman, D. A. Forsyth, and J. L. Mundy, Planar object recognition using projective shape representation, Int. J. Comput. Vis., vol. 16, no. 1, 57-99, 1995
- [9] I. Weiss, Geometric invariants and object recognition, Int. J. Comput. Vis., vol. 10, no. 3, 207-231, 1993
- [10] T. Gevers and A. W. M. Smeulders, Image indexing using composite color and shape invariant features, in Proc. 6th Int. Conf. Computer Vision, Bombay, India. 576-581. 1998
- [11] B. V. Funt and G. D. Finlayson, Color constant color indexing, IEEE Trans. Pattern Anal. Mach. Intell., vol. 17, no. 5. 522-529. 1995.
- [12] M. J. Swain and .H. Ballard, Color indexing, Int. J. Comput. Vis., vol. 7, no. 1, 11-32, 1991.
- [13] L. van Gool, T. Moons, and D. Ungureanu, Affine photometric invariants for planar intensity patterns, in Proc. 4th Eur. Conf. Computer Vision, vol. 1, 642-651. 1996
- [14] F. Mindru, T. Moons, and L. van Gool, Recognizing color patterns irrespective of viewpoint and illumination, in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 368-373. 1999
- [15] B. W. Mel, Seemore: combining color, shape, and texture histogramming in a neurally inspired approach to visual object recognition, Neur. Comput. vol. 9, no. 4, 777-804, 1997

- [16] S.K.Nayar and R.M.Bolle, Reflectance based object recognition, Int. J. Comput. Vis. vol. 17, no. 3, 219~240, 1996
- [17] T.Gevers and A.W.M.Smeulders, Color-based object recognition, Pattern Recognit., vol. 32, no. 3, 453~464, 1999
- [18] J.R.Taylor, An Introduction to Error Analysis, the Study of Uncertainties in Physical Measurements. Sausalito, CA: Univ. Science Books, 1982
- [19] 章毓晋, 图像工程上册, 图像处理与分析, 北京:清华大学出版社, 1999
- [20] 汤海纓, 庄天戈, 计算机彩色模型在图像显示与分割中的应用, 计算机学报, 22(41) 375~382, 1999
- [21] 魏宝刚, 李向阳, 鲁东明, 潘云鹤, 彩色图像分割研究进展, 计算机科学, 26(4) 59~62, 1999
- [22] 徐旭, 朱森良, 梁倩卉, Waheed Sajjad, 一种用于 CBIR 系统的主色提取及表示方法, 计算机辅助设计与图形学学报, 11(5) 385~388, 1999
- [23] Kenneth. R. Castleman 数字图像处理, 电子工业出版社, 北京, 1998
- [24] 田玉敏, 林高生, 基于颜色特征的彩色图像检索方法, 西安电子科技大学学报, 29(1) 43~46, 2002
- [25] 杨静宇, 曹雨龙, 计算机图像处理及常用算法手册, 南京大学出版社, 1997
- [26] M. J. Swain, D.H. Ballard, Color indexing, Int. J. Computer. Vision 7(1). 11~32. 1991
- [27] B. V. Funt and G. D. Finlayson, "Color constant color indexing," IEEE Trans. Pattern Anal. Mach. Intell., vol. 17, no. 5, 522~529, 1995
- [28] E.H. Land, J.J. McCann, Lightness and retinex theory, J. Opt. Soc. Am. 61, 1~11. 1971
- [29] <http://staff.science.uva.nl/~aloi/>
- [30] Swain M. J., Ballard D. H., Color indexing. Intl. J. on Computer Vision, 7(1): 11~32, 1991
- [31] Stricker M., Orengo M., Similarity of color images. In: Proceeding of SPIE Storage and Retrieval for Image and Video Database, 381~392, 1995
- [32] M. Chantler, M. Schmidt, M. Petrou, and G. McGunnigle. The effect of illuminant rotation on texture filters: Lissajous' sellipses. In Proc. European Conference on Computer Vision, vol. 3, 289 - 303, 2002.
- [33] 王宇生, 陈纯, 一种新的基于色彩的图像检索算法, 计算机研究与发展, 39(1) 105~109, 2002



- [34] Dupont B. Bone marrow transplantation in sever combined immunodeficiency with an unrelated MLC compatible donor. In: White H J, Smith R, eds. Proceedings of the Third Annual Meeting of the International Society for Experimental Hematology. Houston International Society for Experimental Hematology, 44-46. 1997
- [35] Klinker, C. J., Shafer, S. A., and Kanade, T., Using a Color reflection model to separate highlights from Object Color, Proc. of the Int. Conf. On Computer Vision, New York, 145-150. 1987
- [36] Klinker, G. I., A Physical approach to Color image understanding, A. K. Peters, Welleslay, Massachusetts, 1993
- [37] Land, E. H., The retinal theory of Color Vision, Proceedings of the IEEE, Vol. 47, No. 1, 23-58. 1974
- [38] Land, E. H., The retinal theory of Color Vision, Scientific American, Vol. 237, No. 6, 108-128. 1977
- [39] Levine, M. D., Vision in man and machine, McGraw-Hill Book Company, 1985
- [40] OgleVE, StonebrakerM. Chabot: retrieval from a relational database of images. IEEE Computer, 28. 40-48. 1995
- [41] Flickner M, et al. Query by image and video content: the QBIC system. IEEE Computer, 28. 23-32. 1995
- [42] Smith J R, Chang S F. Visualseek: a fully automated content-based image query qsystem. Proc. ACM Multimedia Conf. 87-98. 1997
- [43] Swain M, Ballard D. Color indexing. International Journal of Computer Vision, 11-32. 1997
- [44] Niblack W, Berber R, et al. Extracting color features and dynamic matching for image data-base retrieval. IEEE Trans. on Circuits and Systems for Video Technology. 496-505. 1999
- [45] D. Slater, G. Healey, The illumination-invariant recognition of 3-D objects using local color invariants, IEEE Trans. PAMI 18(2)1996 206-211.
- [46] S. A. Shafer, Using color to separate reflection components, COLOR Res. Appl. 10 (4) 210-218. 1985
- [47] D. Forsyth, A novel algorithm for color constancy, Int. J. Comput. Vision 5-36. 1990. 5
- [48] B. V. Funt, M. S. Drew, Color constancy computation in near-mondrian scenes, 544-549. CVPR, IEEE Computer Society Press, Silver Spring, MD, 1988
- [49] T. Gevers, Color image invariant segmentation and retrieval, Ph. D. Thesis, ISBN 90-74795-51-X, University of Amsterdam, The Netherlands, 1996

- [50] G. D. Finlayson, M. S. Drew, B. V. Funt, Spectral sharpening: sensor transformations for improved color constancy, J. Opt. Soc. Am. 11 (5) 1553~1563. 1994
- [51] J. J. Koenderink, the Structure of Images, Biological Cybernetics. vol. 50, 363~370, 1984
- [52] J. J. Koenderink and A. J. van Doorn, Receptive Field Families, Biological Cybernetics, vol. 63, 291~297, 1990
- [53] T. Gevers, S. Ghebreab, and A. W. M. Smeulders, Color Invariant Snakes, Proc. Ninth British Machine Vision Conf., P. H. Lewis and M. S. Nixon, eds. 659~670, 1998

## 发表论文和科研情况说明

### 发表的论文：

肖松山 娄强 “基于颜色不变量的物体识别” 已投稿

### 参与的科研项目：

图像不变性识别研究

基于烟支空头检测的烟草机械改造

## 致 谢

本论文的工作是在我的导师肖松山副教授的悉心指导下完成的,肖松山副教授严谨的治学态度和科学的工作方法给了我极大的帮助和影响。在此衷心感谢三年来肖松山老师对我的关心和指导。

在实验室工作及撰写论文期间,吴永兴、兰子穆、宿曼、锁斌等同学对我论文中的研究工作给予了热情帮助,在此向他们表达我的感激之情。

另外也感谢家人,他们的理解和支持使我能够在学校专心完成我的学业。