

基于生物视觉的颜色直方图特征提取^①张悦^② 霍宏方^③ 涛^③

(上海交通大学自动化系系统控制与信息处理教育部重点实验室 上海 200240)

摘 要 根据人类视觉通路对颜色的感知过程,建立了颜色感知的层次性描述模型。用该模型模拟了自然光经眼睛进入视网膜(RETINA),到侧膝体(LGN),再到 V1 阶段的处理过程,最终在 V1 阶段的 double-opponent 神经元响应下形成具有方向信息的颜色特征。针对传统的颜色直方图包含信息较少,特征存在维数高、精度低的缺点,基于此模型,提出了描述具有对比度敏感、具有方向选择性和颜色恒常性的层次性颜色特征的方法。利用该特征对遥感图像进行了分类实验,实验表明了该特征比广泛采用的颜色直方图特征有更低的特征维数和更好的分类效果。对不同光照的图像进行了分类实验,验证了该特征具有颜色恒常性和更好的稳健性。

关键词 腹侧通路,层次性,double-opponent,颜色直方图,图像分类

0 引 言

颜色是一种非常重要的特征,广泛应用于计算机视觉领域。彩色图像相对于灰度图像来说,包含了更多的信息,更有助于我们理解图像。颜色特征中最常用的就是颜色直方图特征,Swain^[1]等最早提出了用颜色直方图进行特征提取的方法。颜色直方图是从统计意义上分析颜色分量所占的比例,具有复杂度低,对旋转、缩放等几何变换具有不变性的优点,因而被广泛应用于图像检索、分类识别等领域^[2,3]。除了最基本的红绿蓝(RGB)直方图外,还有很多经过空间变换颜色直方图,如色调-饱和度-亮度(HSI)直方图和 rg 直方图等^[4]。然而目前的颜色直方图特征是从计算机视觉角度出发提出的,忽略了图像的空间信息,而且对光照变化也很敏感^[4],难以完整地表达出所需要的信息。自然场景中目标的颜色是生物视觉感知的最重要的低层次特征之一,根据人与哺乳动物视觉系统具有层次性处理的特点,颜色的处理沿着视网膜(RETINA)-侧膝体(lateral geniculate nucleus, LGN)-V1-V2-PIT-IT 腹侧通路完成的^[5]。随着近年来随着神经科学研究

的不断深入及层次性神经编码机制的提出^[6],很多计算机视觉研究学者从生物视觉对颜色感知的角度出发,研究更符合生物视觉颜色处理机理的特征提取方法。Zhang 等^[7]提出的基于生物视觉的颜色特征,结合形状特征,在目标识别、场景分类中取得了较好效果。但是由于结合了分层最大化(Hierarchical MAX, HMAX)模型^[6]和尺度不变特征转换(scale-invariant feature transform, SIFT)等特征,使得计算相对复杂。本文从视觉的腹侧通路出发,重点研究了沿着 RETINA-LGN-V1 通路的颜色加工过程,建立了颜色感知的层次性描述模型,提出了一种相对简单同时包含方向信息的颜色特征描述方法。该方法模拟每个阶段的生物学机理,通过数学建模得到 V1 阶段的颜色输出,然后利用这些信息构造直方图特征。最后利用支持向量机(SVM)进行遥感图像和颜色恒常性图像分类,并和现有的方法进行了对比。

1 视觉通路颜色感知的机理

1.1 RETINA

颜色的形成离不开光照,光线进入人眼后首先

① 973 计划(2012CB719903),国家自然科学基金委创新研究群体(X198144),国家自然科学基金青年科学基金(41101386)和国家自然科学基金(41071256)资助项目。

② 男,1986 年生,硕士;研究方向:计算机视觉,数字图像处理。

③ 通讯作者, E-mail: tfang@sjtu.edu.cn

(收稿日期:2013-09-10)

要经过视网膜处理。在视网膜上,遍布锥细胞和杆细胞,其中杆细胞负责明暗光线的接收,而负责色彩的则是分别对长波 L、中波 M 和短波 S 敏感的三种锥细胞。为了简便起见,这里直接用图像的三个 RGB 通道来近似模拟 LMS 锥细胞的响应。

1.2 LGN

锥细胞将接收到的信号通过轴突传递到 LGN。LGN 区神经元的感受野由 RETINA 区的锥细胞组成,它们的中心和外周区域是兴奋和抑制的对抗关系,称为锥对抗 (cone-opponent) 细胞。颜色经过 cone-opponent 型的神经元在 LGN 区分别形成 R (红) - G (绿) 通路 (L - M)、B (蓝) - Y (黄) 通路 (S - (L + M)) 与亮度 (Luminance) 通路 (L + M)^[8] 三种不同的通路。其中 R - G 和 B - Y 通路又分别各包含两种, R - G 按照中心区对红色刺激兴奋和抑制分为 R + G - 和 R - G +, B - Y 也是如此, Luminance 通路由于不包含颜色对抗信息,所以只有一个通路。LGN 形成的通路如图 1 所示,包含了 4 个对抗通路和一个非对抗通路。

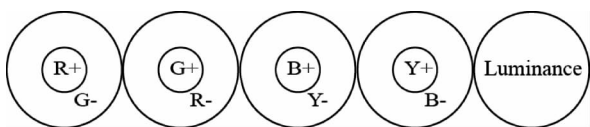


图 1 LGN 形成的五种 cone-opponent 颜色通路

1.3 V1

LGN 形成的三种颜色通路进入 V1 区,在双重对抗 (double-opponent) 型神经元的作用下又形成了三种新的颜色通路,与 LGN 类似,可以用 RG、BY 和 LUM (Luminance) 表示^[9]。这里的双重 (double) 是指色度信息和空间信息,只有二者同时满足一定条件才能获得这种神经元的最大响应。比如中心区受红光刺激兴奋的 double-opponent 神经元,需要红色区域在绿色的背景下才能获得,在其他颜色背景下或者红色区域在绿色区域旁边都不能获得。

但是,与 LGN 不同,在 V1 区, double-opponent 型细胞的感受野是不对称的^[10],因此这三种颜色通路表现出了对方向的选择性。这些 double-opponent 细胞是颜色对比度和颜色恒常性的基础,也是理解颜色在皮层中表达的关键所在。V1 形成的通路如图 2 所示,同样这里的 RG、BY 通路也是各包含两种通路,分别记为 RGGR 和 GRRG、BYYB 和 YBBY。

颜色信息通过从 RETINA 到 LGN 再到 V1 颜色

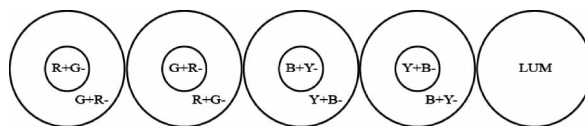


图 2 V1 形成的五种 double-opponent 通路

的神经编码过程,在 V1 阶段输出的颜色编码特征将同时具有对不同波长的颜色处理能力,具有对比度敏感、方向选择性和颜色恒常性等特点。

2 颜色感知的层次性描述模型

基于上述腹侧通路颜色感知的机理,本文建立了颜色感知的层次性描述模型,基于此模型,提出了具有对比度敏感、方向选择性的层次性颜色特征描述方法。

2.1 模型具体描述

图 3 所示为 RETINA-LGN-V1 阶段的层次性颜色特征描述模型及直方图特征提取的过程,图中虚线表示抑制或者相减的关系。

为了研究方便,Shapley^[11]等人构造出感受野的数学模型,即感受野的空间 - 色度敏感函数为

$$R(x, y, \lambda) = a_L L(\lambda) r_L(x, y) + a_M M(\lambda) r_M(x, y) + a_S S(\lambda) r_S(x, y) \quad (1)$$

在图像处理中, (x, y) 表示图像的坐标; $L(\lambda)$ 、 $M(\lambda)$ 、 $S(\lambda)$ 分别是三种锥细胞的光谱响应函数,可以简单地对应图像的 R、G、B 三通道; r_L 、 r_M 、 r_S 是每个锥细胞输入的空间敏感度分布,也就是其感受野形状,在 LGN 中可以近似用高斯差分 (Difference of Gaussians, DOG) 函数来代替, V1 阶段普遍用 Gabor 函数来表示; a_L 、 a_S 、 a_M 为系数。

DOG 函数形式如下:

$$DOG(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) - \frac{1}{2\pi K^2 \sigma^2} \exp\left(-\frac{x^2 + y^2}{2K^2 \sigma^2}\right) \quad (2)$$

其中 (x, y) 表示坐标, σ 表示标准差, K 表示系数。

具体实施时,比如 R + G - 的对抗,即中心区是红色,外周是绿色的对抗,先用 DOG 滤波器的正分量对图像的 R 通道进行卷积,负分量对 G 通道进行卷积,然后再利用式 (1) 求出 R + G - 的通道响应。

图 4 表示 3×3 的像素点经过 DOG 滤波后的图像,可以看出被绿色包围的红色部分经过 R + G - 的 DOG 滤波处理后,红色部分得到了显著增强 (白

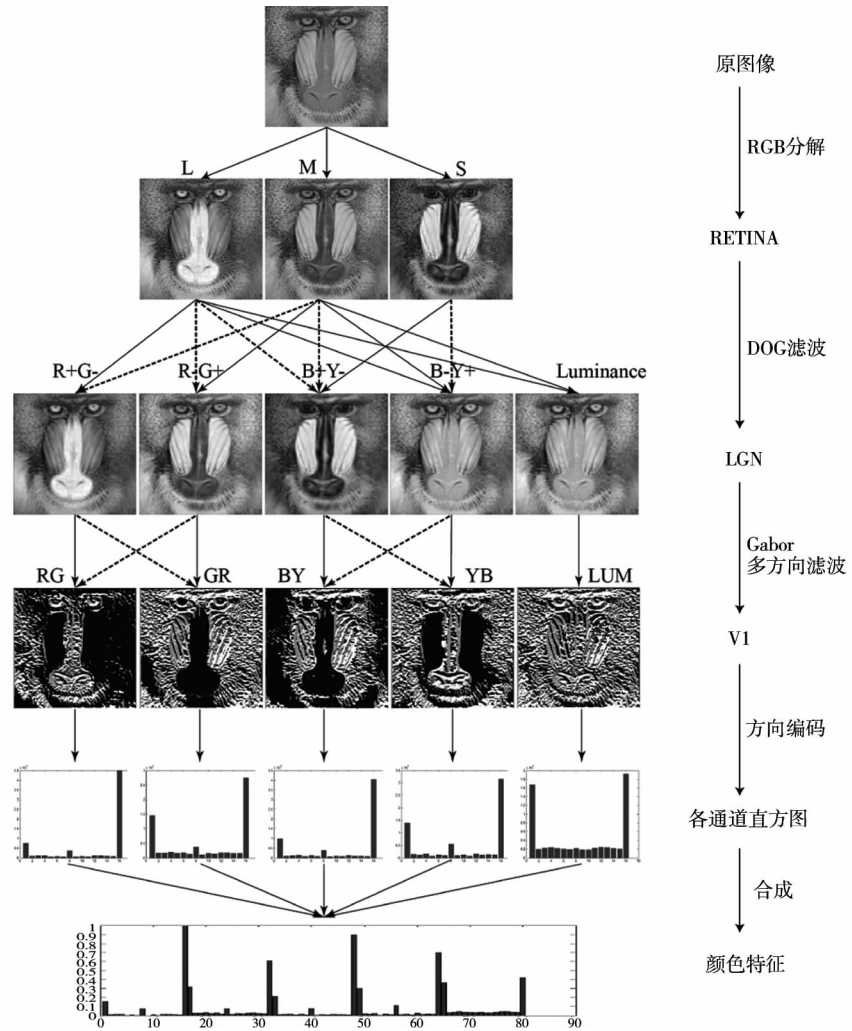
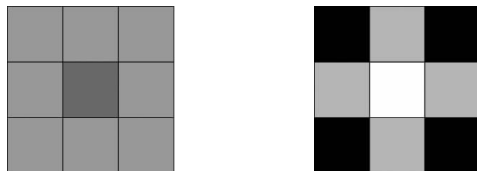


图3 颜色感知的层次性描述模型及特征提取

色表示最显著,黑色表示最不显著,灰色介于两者之间),而邻域和外周部分则相应得到抑制,这种结构体现了颜色对比度敏感性,可以用它检测出特定的颜色空间关系。



(a) 原图 (b) 经过对抗神经元响应后的图像

图4 DOG 滤波器模拟 cone-opponent 神经元的响应

同理,可以得出 $R+G-$ 、 $R-G+$ 、 $B+Y-$ 、 $B-Y+$ 和 Luminance 共 5 个通道,对于 Luminance 亮度通路,目前大都采取 RGB 三个通道的线性组合形成,由于亮度信息不包含颜色对抗性,直接用 DOG 滤波器对 RGB 的线性组合进行滤波获得。

在 V1 阶段,用 Gabor 滤波器表示 double-opponent 细胞的感受野。因为考虑到不同的颜色对抗,所以采用奇滤波,即 Gabor 滤波器的虚部:

$$Gabor(x, y) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \sin\left(2\pi \frac{x'}{\lambda}\right) \quad (3)$$

其中, $x' = x\cos\theta + y\sin\theta$, $y' = -x\sin\theta + y\cos\theta$, (x, y) 表示坐标, σ 表示标准差, λ 表示正弦函数的波长, θ 表示方向。

实际中,图像经过 LGN 区的 cone-opponent 细胞后,形成了 5 个颜色通道(Luminance 虽然不包含对抗信息,但是它对颜色的形成也具有重要作用)。用 Gabor 滤波器的正负分量(图 5 (b) 所示的红色(A)和蓝色(B)区域)对 LGN 的通道分别进行滤波,比如用正分量对 $R+G-$ 通道响应进行卷积,负分量对 $R-G+$ 卷积,两者相加得到其中一个 double-opponent 通道响应 RGGR,其他通道响应也用类

似方法获得。通过 Gabor 滤波器的 θ 可以调节方向,从而每个通道都能得到几种不同的方向响应。

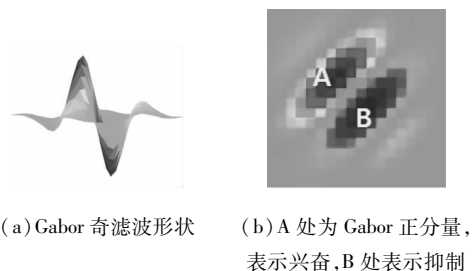


图 5 Gabor 奇滤波的形状及特性

选择这些方向中最大的响应作为最终的该通道响应,并通过对这些最大响应对应的方向进行编码,组成 5 个颜色通道的方向图。最后再利用这些方向图编码信息提取直方图特征,将 5 个通道的直方图合成为一个直方图向量,进行归一化之后作为一幅图像的特征。

2.2 对比度敏感性和方向选择性可视化表达

经过 V1 区的处理,形成具有对比度敏感性和方向选择性的颜色特征,可以通过图 6 直观展示出来。

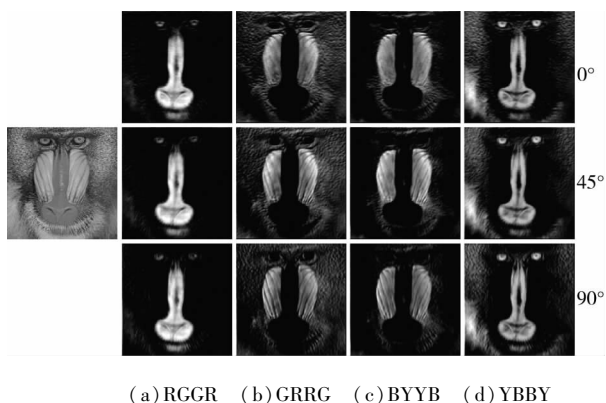


图 6 V1 区三个方向上的颜色对立输出

图 6 给出了不同颜色通道在不同方向下的响应, RGGR、GRRG、BYYB 和 YBBY 分别表示 4 种双对抗的颜色通道,亮度通路不包含颜色对抗信息,没有展示出来。图 6(a) 表示 RGGR 通道在 0° 、 45° 和 90° 方向上的响应。对同一列而言,图中上中下三幅图像的差别可以通过狒狒脸上的毛的走向来判断,比如 RGGR 在 0° 方向上的响应,表明红绿对抗明显的地方的朝向会是水平的,这里包含了两种信息,一种是对比度敏感,即红色部分的显著增强和绿色部分的抑制,另一种是方向选择性,即将水平方向的信息

息选择性地凸显出来。其他通道类似,分别如图 6(b) GRRG、图 6(c) BYYB 与图 6(d) YBBY 所示。

可见,通过 V1 区 double-opponent 神经元的响应,颜色特征表现出对比度敏感和方向选择性。

3 实验与分析

根据本文模型,一幅图像的颜色特征提取具体过程如下:

(1) 提取出彩色图像 RGB 三个通道的响应。

(2) 用 DOG 滤波器对 RGB 三个通道响应进行滤波,滤波方式见模型具体描述,其中式(1)中系数的选择如下:对于 R-G 神经元, $a_L = -a_M$, 为简便起见,取 $a_L = -a_M = 1$, $a_s = 0$; 对于 B-Y 型的神经元,取 $a_s = -(a_L + a_M) = 1$; 对于 Luminance 神经元, $a_L = a_M = 1$, $a_s = 0$ 。式(2)中 DOG 滤波器, σ 取 1, K 取 5, 窗口大小 5×5 。由此形成 R+G-、R-G+、B+Y-、B-Y+ 和 Luminance 共 5 个通道响应。

(3) Gabor 对上层的 5 个通道进行滤波,滤波方式见模型具体描述,式(3)中, γ 取 2, σ 取 3, λ 取 5, 窗口大小 5×5 , 若 $(0 \sim 180^\circ)$ 等分成 n 份(n 的具体数值在后文中有讨论), 则 θ 取值为 $(0, \frac{\pi}{n}, \frac{2\pi}{n}, \dots, \frac{(n-1)\pi}{n})$, 经过滤波形成 RGGR、GRRG、BYYB、YBBY 和 LUM 5 个通道。

(4) 对 5 个通道 n 个方向中最大响应对应的方向进行编码,每个方向编码为一个数值,这样一个通道编码范围为 $(0 \sim n-1)$ 。利用这些方向图编码信息提取直方图特征,直方图 bin 个数取 n , 这样每个通道特征维数为 n , 再将 5 个通道的直方图合成为一个向量,进行归一化之后作为特征向量,维数为 $5n$ 。

归一化之后的特征向量即本文提出的方法提取出图像的颜色特征,用该特征进行下面的实验。

实验所用计算机配置环境为:

CPU: Intel Xeon X5690 3.47GHz, 内存: 32.0GB, 操作系统: 64 位 Windows Server 2008, 开发平台: MATLAB R2013a。

3.1 遥感图像分类实验

实验采用 SVM 分类器对 21 类遥感数据集^[12]进行分类,数据集共有 21 类,每类 100 幅图像,每幅大小为 256×256 像素。这 21 类包含了较为复杂和

丰富的地物结构,但类与类之间有些存在颜色上的相似性,有些存在形状上的相似性,且大多数图像的空间结构关系也较为复杂,使得该数据集分类难度较大。实验采用农田、飞机、建筑、稠密居民区、森

林、港口、移动住房、河流、稀疏居民区和储油罐这 10 类进行,图 7 列出了这 10 类的样图。经过本文的方法提取特征后,用 LIBSVM 工具箱^[13]对 SVM 分类器进行训练。

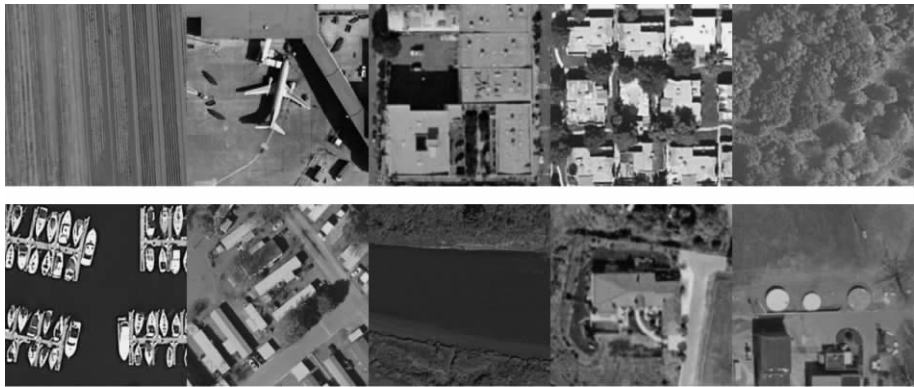


图 7 待分类的遥感数据集

为了研究特征维数对分类精度的影响,对($0^{\circ} \sim 180^{\circ}$)方向分别进行了 8、16、24、32、40、48、56、64 和 72 等分,最终提取的特征分别为 40、80、120、160、200、240、280、320 和 360 维,训练样本和测试样本各取 50,随机取样,重复 10 次,得到 10 个不同的训练集和测试集。利用 libsvm 自带的 5 折交叉验证进行训练,5 折的平均值作为分类精度,用上面得到的 10 个样本集进行 10 次训练,再求其均值(即 10 次 5 折的平均值)作为总体分类精度,分类结果如图 8 所示。从图中可以看出,随着特征维数的逐渐增加,总体分类精度有所提升,当维数增加到 240 维时,分类精度达到最大值,随后精度反而有所下降,而且高的特征维数会增加计算时间,因此没有必要再增加维数进行讨论,本文最终采用 240 维向量作为特征。

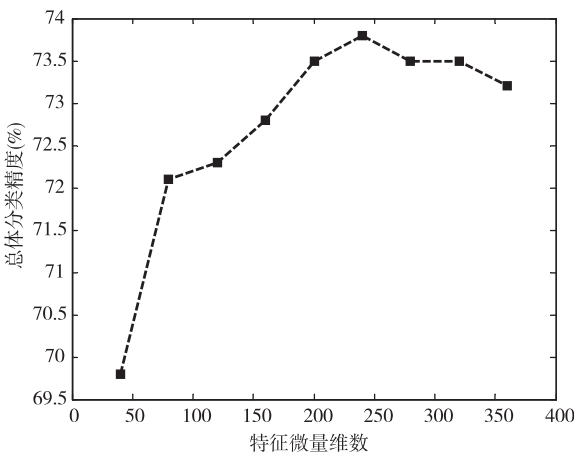


图 8 特征向量维数对分类精度的影响

本文算法与普遍采用的 RGB 直方图、rg 直方图和 HSI 直方图进行了对比。RGB 直方图每个通道分别进行了 50、100 的量化,分别生成了 150 和 300 维的向量;rg 直方图每通道分别采取 50、100、150 量化;而 HSI 三通道分别采用 16、4、4 的均匀量化,最后用 $L = 16H + 4S + I$ 表示特征,特征维数 256 维。

表 1 所示为不同颜色直方图的总体分类精度,由表中可以看出,RGB 直方图特征在特征维数为 150 和 300 维时,分类精度提升不多;rg 直方图在维数为 300 时效果最好,但是精度仍然太低;HSI 直方图特征为 256 维,但精度不高。这三种直方图特征即使取最高的特征维数,最好的分类精度依然比本文方法的低 5% 以上。

表 1 不同直方图分类精度

直方图特征	特征维数	精度(%)
RGB	150	68.1
	300	68.9
rg	100	50.8
	200	54.3
	300	55.5
HSI	256	58.6
本文方法	240	73.8

为了分析遥感图像各类别的分类情况,表 2 给出了本文方法的分类结果混淆矩阵,表中每行表示

50 个测试集中每幅图像的分类情况,对角线上的表示正确分类的图像数,非对角项表示错分的数目。由表 2 可以看出,大部分类别都能有超过半数的测试图像被正确分类,从而验证了本文方法对遥感图像分类的有效性和正确性。

表 2 图像分类实验的混淆矩阵

类别	分类情况									
农田	48	0	0	0	0	0	1	1	0	0
飞机	0	38	3	1	0	0	3	2	1	2
建筑	1	1	24	9	0	0	5	1	2	7
稠密居民区	0	0	5	32	0	0	6	1	2	4
森林	0	0	0	0	50	0	0	0	0	0
港口	0	1	0	0	0	49	0	0	0	0
移动住房	4	0	5	0	1	0	40	0	0	0
河流	1	2	0	3	3	6	2	27	5	1
稀疏居民区	0	4	0	1	1	0	1	7	36	0
储油罐	0	5	11	3	1	0	2	0	2	26

3.2 颜色恒常性图像分类实验

颜色恒常性是视觉系统重要特性之一,本实验将验证本文提出的特征对不同光照的影响。实验图像采用 Barnard 等^[14]所用的颜色恒常性图像集,展示的是同一目标在不同光照下的图像,选取了其中的 10 类,每类 10 幅图像进行分类实验,这 10 类分别为 ball、blocks1、clothes4、hdr_mondrian、macbeth、monkey、munsell1、munsell5、paper 和 sml_mondrian。其中部分样本如图 9 所示。

该实验主要研究不同特征对光照区分能力,因此实验时固定各方法的特征维数分别为:RGB 300 维,rg 300 维,HSI 256 维,本文方法 80 维。训练集和测试集各取 5,每种特征都进行了 10 次 5 折交叉验证实验。不同特征的分类实验结果如图 10 所示。从图 10 中可以看出,本文方法在 10 次实验中都能维持在较高精度(95% 以上),而其他直方图特征总体分类精度不高,并且在每次实验中精度都有很大变化,rg 和 HSI 直方图尤为明显。这体现了本文方法提取的特征对光照的依赖较小,而且较为稳定。

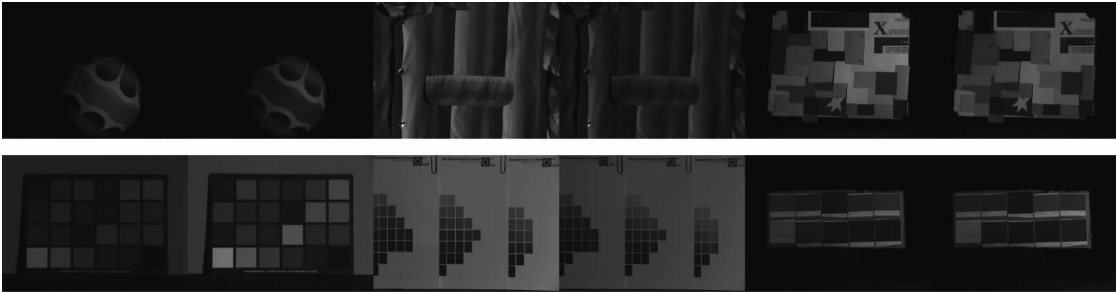


图 9 部分颜色恒常性图片

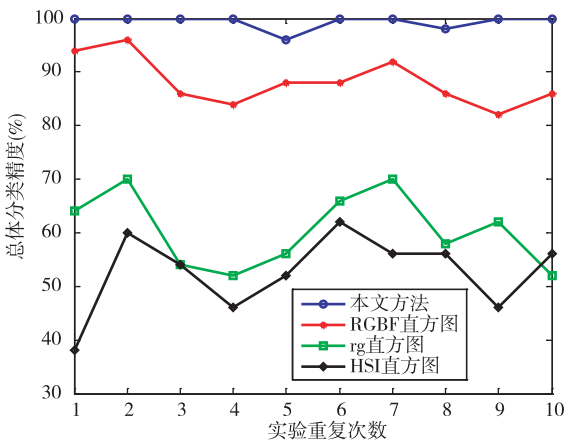


图 10 不同特征的精度及实验次数的影响

4 结论

本文从生物视觉的角度出发,通过模拟人类视觉通路对颜色的感知过程,建立了颜色感知的层次性描述模型,该模型重点研究了腹侧通路的 RETINA-LGN-V1 阶段,其中对 cone-opponent 和 double-opponent 两种神经元的模拟,使得模型能够对颜色对比度敏感、对方向具有选择性。在此基础上,针对现有的直方图方法的包含信息少、维数高并且效果不是很好的缺点,提出了含有方向信息的颜色直方图特征。最后进行了分类实验验证,结果表明该方法计算过程简单,特征维数低,能够稳健表达不同光照下的颜色特征,分类精度较高。

目前的研究还存在很多问题,比如本文只研究了整个颜色视觉通路的一小部分,后面更高层的V2、V4和IT区对颜色处理更加精细,对颜色特征的表达很有帮助^[5]。但是本文对颜色感知的层次性描述模型研究将为以后继续在视觉皮层的高层模拟颜色的感知过程奠定了基础。而且虽然本文在颜色特征中用到了方向信息,丰富了颜色的表达能力,但是这只是简单的结合,要想取得更好的效果,颜色必须和形状信息结合起来^[7],这些都是以后研究的重点。

参考文献

- [1] Swain M J, Ballard D H. Color indexing. *International journal of computer vision*, 1991, 7(1): 11-32
- [2] 窦建军,文俊,刘重庆. 基于颜色直方图的图像检索技术. *红外与激光工程*, 2005, 34(1): 84-88
- [3] 张鑫,刘秉权,张德园等. 空间金字塔颜色直方图在图像分类中的应用. *计算机工程与应用*, 2010, 46(018): 152-155
- [4] Van De Sande K E, Gevers T, Snoek C G. Evaluating color descriptors for object and scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(9): 1582-1596
- [5] Conway B R. Color vision, cones, and color-coding in the cortex. *The Neuroscientist*, 2009, 15(3): 274-290
- [6] Serre T, Wolf L, Bileschi S, et al. Robust object recognition with cortex-like mechanisms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, 29(3): 411-426
- [7] Zhang J, Barhomi Y, Serre T. A new biologically inspired color image descriptor. In: *Proceedings of the 2012 European Conference on Computer Vision, Heidelberg, Germany: Springer Berlin*, 2012. 312-324
- [8] Boynton G M. Color Vision: How the Cortex Represents Color. *Current Biology*, 2001, 12(24): 838-840
- [9] Chatterjee S, Callaway E M. Parallel colour-opponent pathways to primary visual cortex. *Nature*, 2003, 426(6967): 668-671
- [10] Conway B R, Livingstone M S. Spatial and temporal properties of cone signals in alert macaque primary visual cortex. *The Journal of neuroscience*, 2006, 26(42): 10826-10846
- [11] Shapley R, Hawken M J. Color in the cortex: single- and double-opponent cells. *Vision research*, 2011, 51(7): 701-717
- [12] Yang Y, Newsam S. Bag-of-visual-words and spatial extensions for land-use classification. In: *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM GIS)*, 2010, 270-279
- [13] Chang C C, Lin C J. LIBSVM--A Library for Support Vector Machines, <http://www.csie.ntu.edu.tw/~cjlin/libsvm/index.html>
- [14] Barnard K, Martin L, Funt B, et al. A data set for color research. *Color Research & Application*, 2002, 27(3): 147-151

Color histogram feature extraction based on biological vision

Zhang Yue, Huo hong, Fang Tao

(Key Laboratory of System Control and Information Processing, Ministry of Education,
Department of Automation, Shanghai Jiao Tong University, Shanghai 200240)

Abstract

According to the color perception of human visual pathways, a hierarchical model for description of color perception was established. By using the model, the neural processing for the natural light entering through the eye into the retina, the lateral geniculate nucleus (LGN), then to the sytage of V1, was simulated, and eventually, a color feature with orientation information formed in the response of double-opponent neurons. As traditional color histogram features have the shortcomings of less information, high dimension and low accuracy, a method for description of a hierarchical color feature with the contrast sensitiveness, direction selectivity and color constancy, was proposed based on this model. The experiment on remote sensing image classification was conducted by using the fenture, and the results show that the proposed method could achieve the better accuracy and the lower dimension than traditional ways. The classification of the images under different illuminations verifies that the proposed feature has the color constancy and better stability.

Key words: ventral pathway, hierarchy, double-opponent, color histogram, image classification