

本文节选自《基于MPEG-7与内容的图像检索技术的研究》。

MPEG-7 标准中视觉描述工具包括基本结构和描述符。本文主要介绍各描述符。

(1) 颜色描述符

MPEG-7 主要定义了七种颜色描述符：颜色空间、主颜色、颜色的量化、颜色直

方图、颜色布局、GOF/GOP 颜色、颜色结构、对颜色的特征和结构进行描述。他们之间的关系如图 2.2 所示。

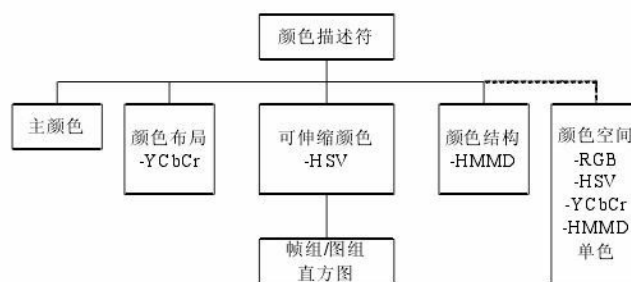


图 2.2 MPEG-7 颜色描述符

①颜色空间,该描述符主要用于其它基于颜色的描述。当前描述所支持的颜色空间有:RGB、YCbCr、HSV、HMMD、关于 RGB 的线性变换矩阵、单色。

②颜色量化,该描述符定义了颜色空间的均匀量化。量化产生的维(bin)的数目是可配置的,这样使得各种应用具有更大的灵活性。要使这个描述符在 MPEG-7 背景下有应用意义,例如表示主颜色值的含义,必须结合其它颜色描述符。

③主颜色,该描述符最适用于表示局部(对象或图像区域)特征,几种颜色就足以表达我们感兴趣区域的颜色信息。当然,它也可以用于整个图像,例如旗帜图像或彩色商标图像。颜色量化用于提取每个区域/图像的少数代表颜色,并相应的计算出区域中的每种量化颜色所占的百分比。同时还定义了整个描述符的空间相关性,用于相似性检索。

④颜色布局,该描述符以一种紧凑的形式,有效的表达了颜色的空间分布。这种紧凑性以很小的计算代价,带来高速的浏览和检索。它提供图像与图像的匹配和超高速的片断与片断的匹配,这些匹配要求大量相似性计算的重复。由于该描述符表达了颜色特征的布局信息,因此它可以提供相当友好的用户接口,例如使用其它颜色描述符中均不支持的手绘草图查询。

⑤可伸缩颜色,该描述符是 HSV 颜色空间的颜色直方图(用 Haar 变换编码)。根据维的数目和比特表示的精度,它的二进制表示在一定数据速率范围内是可伸缩的。这个描述符主要用于图像与图像的匹配和基于颜色特征的检索,检索的精度随着描述中使用的比特数目的增加而增加。

⑥颜色结构,该描述符是一个颜色特征描述符,它既包括颜色内容信息(类似于颜色直方图),又包括内容的结构信息。它的主要功能是图像与图像的匹配,主要用于静态图像检索,在这里一幅图像可能由一个单一矩形或者任意形状、可能是非连通的区域组成。提取的方法是:通过考虑一个 8×8 像素的结构化元素中的所有颜色,将颜色结构信息加入该描述符中,而不是单独考虑每个像素。

⑦帧组/图组颜色,该描述符将用于静态图像的可伸缩颜色描述符扩展到对视频片段或静态图像集合的颜色描述。在 Haar 变换之前,用附加的两个比特定义如何计算颜色直方图,是均值、中值还是相交。

(2) 纹理描述符

MPEG-7 推荐了三种纹理描述符,同质纹理描述符(Homogenous Texture Descriptors)、纹理浏览描述符(Texture Browsing Descriptors)和边缘直方图描述符(Edge Histogram Descriptors)。

①同质纹理描述符

同质纹理作为一个重要的视觉基本特征,主要用于大量相似图案的搜索和浏览。一幅图像可看作由同质纹理以马赛克形式拼接而成的,所以与这些区域关联的纹理特征可以作为索引来检索图像。例如,用户浏览一个航空图像数据库,可能想识别图像集合中停车场。当从远处观察时,汽车规则(以相等间隔)停放的停车场就是一个极好的同质纹理图案的例子。同样的,从空中或是卫星拍摄的农田和植被也是同质纹理的示例。同质纹理描述符使用 62 个数字(每一个都量化为 8bits)提供了纹理信息的量化表示。这 62 个数字的计算方法是:将图像用一组带有方向和尺度参数的 Gabor 滤波器进行滤波,一共是六个方向和五个尺度。一共得到 30 个滤波结果,每一个滤波结果频率域的一阶矩和二阶矩就是两个描述数字。

②纹理浏览描述符

描述纹理的感知特性,例如规则性、方向性和粗糙性。计算方法和同质纹理描述符类似,首先使用一组带有方向和尺度参数的 Gabor 滤波器进行滤波,然后通过分析滤波结果,找到纹理主要的方向。接着分析滤波后的图像沿着这两个(第二个主方向是可选的)主方向投影,来确定纹理的规则性和稀疏性。最多只需要 12 比特。同质纹理描述符和浏览描述符提供了表示相似纹理区域(Homogeneous Texture Regions)的多尺度方法。

③边缘直方图描述符

描述了 5 种边缘(4 种方向边缘和一种无方向边缘)的空间分布统计。因为对于图像感知,边缘扮演了一个重要角色,所以它可以用于相似语义的图像检索。因此它的主要目标在于图像与图像的匹配(通过示例或草图),特别是边缘分布不规则的自然图像。如果边缘直方图描述符与其它描述符如颜色直方图相结合,将可以显著提高图像检索的性能。

(3) 形状描述符

MPEG-7 定义了三种形状描述符:基于区域的形状(RegionShape)、基于轮廓的形状(Contour Shape)和三维形状(Shape 3D)。

①基于区域的形状

一个对象的形状可能只包括一个单一区域,也有可能包括多个区域,同时在对象中还可能有一些孔(见图 2.3)。基于区域的形状描述符充分利用了一个帧内组成形状的所有像素,它可以描述任何形状,不但可以描述有一个单一连通区域的简单形状(如图(a)和图(b)),而且可以描述有对象中一些孔或几个不向连区域组成的复杂形状(如图(c)、图(d)和图(e))。基于区域的形状描述符不仅能够有效描述不同目标形状,而且对于沿对象边界的较小变形具有一定的健壮性。

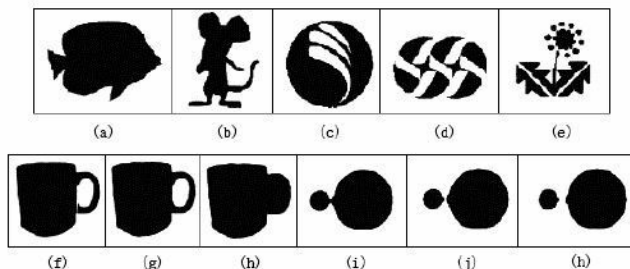


图 2.3 各种形状的例子

②基于轮廓的形状

基于轮廓的形状描述符提取了目标和区域基于轮廓获得的形状,一般采用曲率空间表示形状信息,这种形状描述符反映了人的感知特性,又称为曲率尺度空间(Curvature Scale-space, CSS)表示,它由轮廓的曲率尺度空间表达式构成,该表达式具有如下重要特征(如图 2.4 所示):

较好的提取了形状特有特征,并能用于相似性检索;

反映了人类视觉系统的感知性,具有良好的概括性;

对于不是很明显的运动、形状的部风重叠、投影变换均具有健壮性;

表达式简洁。

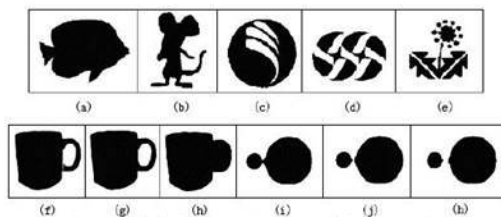


图 2.3 各种形状的例子

③三维形状

考虑到多媒体技术、虚拟世界和增强现实技术的持续发展,三维内容也成为当今多媒体信息系统的普遍特征。大多数情况下,三维信息是用多边形网格来表示的。MPEG-4 的 SNHC 组研究这个问题并开发了有效的三维网格模型编码技术。在 MPEG-7 标准的框架中,要求对三维信息实现智能的基于内容的提取,用以查找、检索和浏览三维模型库。三维形状特征描述符对三维网格模型进行本质的形状描述,它善于挖掘三维表面的局部特征。

原文地址: <http://www.cnki.net/KCMS/detail/detail.aspx?QueryID=4&CurRec=2&recid=&filename=2008145927.nh&dbname=CMFD0911&dbcode=CMFD&pr=&urlid=&yx=&uid=WEEvREcwSIJHSlidSdnQ1V11YzRFR1o1VWJjc3EwQmJhdExWR3NMbFdRSW9oczZOUmNyMkwxczBMdDJxc2JZPQ==&v=MTlwNTJGeXJsVnlzS1YxMjdGcks4RzlkT3FKRWJQSUVl4ZVgqTHV4WVM3RGgxVDNxVHJXTTFGckNVUkxtZVp1UnU=>

文章标签: [mpeg-7](#) [颜色直方图](#) [颜色](#) [纹理](#) [形状](#)

个人分类: [MPEG7/图像检索](#)

