

一种基于实例的图像自动语义标注方法

纪 颖

(哈尔滨理工大学 计算机科学与技术学院, 黑龙江 哈尔滨 150080)

摘 要: 在基于内容的图像检索中, 图像的标注字能够缩小图像的高级语义和低级视觉内容之间的差距, 并方便检索. 但手工标注费时费力且结果具有主观不一致性, 而图像的自动语义标注能够将图像的视觉特征转化为图像的标注字信息, 为用户的使用带来了极大的方便. 本文提出了一种基于实例的图像自动语义标注方法. 该方法的优点是可以根据以往的标注经验自动确定图像标注信息, 克服了手工标注的缺点, 也可以方便地扩展为半自动标注, 为标注者提供了一个简单方便的标注接口.

关键词: 图像检索; 视觉特征; 语义; 自动标注

中图分类号: TP317.4 **文献标识码:** A **文章编号:** 1007-2683(2009)01-0038-05

An Instance-based Method for Automatic Image Semantics Annotation

Ji Ying

(School of Computer Science and Technology, Harbin University of Science and Technology, Harbin 150080, China)

Abstract: In content-based image retrieval, annotations of image can not only reduce the gap between high-grade semantics and low-grade visual content, but is also convenient for retrieval. As we all know, manual-annotation is time-consuming, strength-consuming, and the annotation results may be subjectively different, while the automatic image annotation can transform the visual features into annotations. This benefits users a lot. In this paper, we propose an instance-based method for automatic image annotation. The advantage of this method is that it can determine the image annotation according to previous annotation experiences, which overcomes the shortcomings of manual-annotation. Moreover, the method can be extended to semi-automatic annotation, which provides users with a simple and convenient interface for annotation.

Key words: image retrieval; visual feature; semantics; automatic annotation

1 引 言

近年来,随着多媒体技术和计算机网络的飞速发展,全世界数字图像的容量正以惊人的速度增长,这些数字图像中包含了大量有用的信息.然而,由于这些图像是无序地分布在世界各地,需要人们用图像检索技术来访问和利用图像中包含的有效信息.在基于内容的图像检索中,图像的标注字能够较好

地表达图像的语义内容,能够缩小图像的高级语义和低级视觉内容之间的差距,因此图像标注字也正逐渐引起人们的重视.为了构造一个支持语义检索的图像检索系统,从而获取有效的图像语义是基础.传统的手工标注不仅需要耗费巨大的人力物力,效率低下,而且由于不同的人对图像的理解存在一定的差异,手工标注的信息将会产生不一致性,直接影响检索的结果,并且很难针对大型的图像库实施标注^[1].而图像的自动语义标注能够将图像的视觉特

征转化为图像的标注字信息,这为用户的使用带来了极大的方便,也克服了手工标注费时费力的缺点,因此,图像的自动语义标注成为了支持图像语义检索的一个重要技术,也成为了图像检索领域一个十分活跃的研究分支。

2 图像自动语义标注中的关键技术

图像自动语义标注的流程为:首先提取图像的视觉特征作为图像的语义描述,然后根据某种特征匹配模型,用示例查询的方法进行基于内容的图像检索,最后根据检索结果对示例图像进行自动标注。算法描述如图 1 所示。

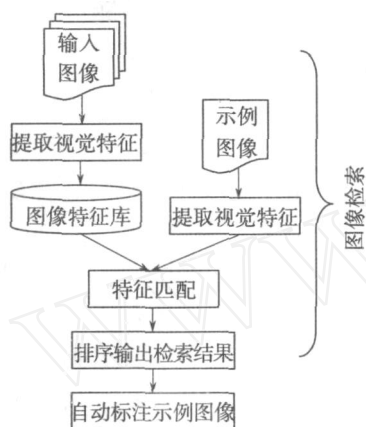


图 1 图像自动语义标注流程

2.1 图像视觉特征提取

图像视觉特征的提取和表达是基于内容图像检索技术的基础。图像的视觉特征主要包括颜色、纹理、形状和空间关系等。图像特征从各个方面描述了图像的内在语义,从而可以作为图像的抽象表示。

颜色特征是被最广泛使用的视觉特征,主要原因在于颜色特征的定义比较明确,特征提取方法比较简单,并且颜色特征能较好地体现出图像中所包含的物体或场景。此外,与其他的视觉特征相比,颜色特征对图像本身的尺寸和方向的依赖性较小,从而具有更高的鲁棒性。颜色特征反映了彩色图像的整体特性。根据颜色与空间属性的关系,颜色特征的表示有颜色直方图、颜色矩、颜色集、颜色聚合向量以及颜色相关图等。

纹理是图像中一个重要而又难以描述的特征,纹理特征是一种不依赖于颜色或亮度的反映图像中同质现象的视觉特征。此特征包含了物体表面结构组织排列的重要信息以及它们与周围环境的联系。

纹理特征的表示方法有共生矩阵、Tamura 纹理特征、小波变换以及 Gabor 滤波器组的纹理特征提取方法。其中,共生矩阵和 Tamura 属于统计方法,小波变换和 Gabor 滤波器组则属于频谱分析方法。

形状特征包括轮廓特征和区域特征两类。前者关心的是物体的外边界,而后者则关心整个形状区域。这两类形状特征的最典型方法分别是傅立叶描述符和形状无关矩。

2.2 图像特征匹配模型的选择

由于视觉特征大都表示成向量的形式,因此常用的相似度衡量模型都采用向量空间模型,在计算的时候这些模型将视觉特征看作是向量空间中的点,并计算两个点(两幅图像)之间的距离来衡量图像特征间的相似度。在图像检索中常用的距离公式包括欧拉距离、直方图相交、二次距离和马氏距离。

2.3 图像的自动语义标注

图像自动语义标注的核心是通过对训练数据集的学习,构建图像与语义关键词之间的标注模型。

3 基于实例的图像自动语义标注技术原理

在对图像自动标注前首先要进行检索出示例图像的相似图像,而视觉特征的提取和图像特征的匹配决定着图像检索结果的好坏,因此二者也是图像自动标注的关键技术。

3.1 颜色特征提取

对彩色图像的研究,必须在彩色图像的特定颜色空间中进行。

3.1.1 颜色空间模型的选取

实际应用中常用的颜色空间有 RGB、YIQ、Lab、HSV 等。一般认为,RGB 颜色空间与人眼的感知差别很大,而 HSV 颜色空间直接对应于人眼视觉特性三要素:色调 H (Hue)、饱和度 S (Saturation) 和亮度 V (Value)。在这 3 个分量中,色调尤其影响着人类的视觉判断。由于 HSV 空间各轴在视觉上彼此无关,空间距离又符合人眼视觉特征,并且从 RGB 到 HSV 的转换是一个简单且快速的非线性变换,所以,本文使用 HSV 颜色空间作为彩色图像的色彩空间。

3.1.2 颜色特征提取

为了使 HSV 颜色模型能更好地符合人眼的视觉特性,首先将图像中每一像素的红、绿、蓝分量值: R, G, B 转化为亮度、饱和度、色调分量值 H, S, V ($[0, 360], [0, 1], [0, 1]$)^[2], 然后根据 HSV 颜色空间中颜色特征进行非等间隔量化^[3]:

$$\begin{aligned}
 H &= \begin{cases} 0 & \text{if } H \in (315, 20] \\ 1 & \text{if } H \in (20, 40] \\ 2 & \text{if } H \in (40, 75] \\ 3 & \text{if } H \in (75, 155] \\ 4 & \text{if } H \in (155, 190] \\ 5 & \text{if } H \in (190, 270] \\ 6 & \text{if } H \in (270, 295] \\ 7 & \text{if } H \in (295, 315] \end{cases} \\
 S &= \begin{cases} 1 & \text{if } H \in [0, 0.2) \\ 1 & \text{if } H \in [0.2, 0.7) \\ 1 & \text{if } H \in [0.7, 1) \end{cases} \\
 V &= \begin{cases} 1 & \text{if } H \in [0, 0.2) \\ 1 & \text{if } H \in [0.2, 0.7) \\ 1 & \text{if } H \in [0.7, 1) \end{cases} \quad (1)
 \end{aligned}$$

在运算中对 H , S , V 三维特征矢量取不同的权值而组合成一维特征矢量, 以便于计算. 在这三个矢量中, 人眼对于颜色的划分主要是依据色调 H , 其次是饱和度 S , 最后才是亮度 V . 同时根据 H , S , V 的量化级数和其频带宽度可取组合后的一维矢量 L .

$$G = HQ_sQ_v + SQ_v + V \quad (2)$$

其中: Q_s 和 Q_v 分别是 S 和 V 的量化级数, 取 $Q_s = 3$, $Q_v = 3$, 上式可表示为

$$G = 9H + 3S + V \quad (3)$$

G 的取值范围为 $[0, 71]$, 是一个 72 柄的一维直方图, 这样 H , S , V 三个分量就在一维矢量上分布开来, 减少了时空的复杂度, 再对特征矢量 G 构造累加直方图作为图像的颜色特征.

3.2 纹理特征提取

本文针对纹理特征的复杂程度, 采用两种纹理特征提取方法.

3.2.1 灰度共生矩阵

Haralick^[4]于 1973 年提出的灰度共生矩阵方法被认为是一个重要而有效的提取纹理特征的方法. 灰度共生矩阵是描述在 θ 方向上, 相隔 d 像元距离的一对像元, 分别具有灰度值 i 和 j 的出现概率, 其元素可记为 $P(i, j|d, \theta)$, 当 θ 和 d 选定时, 也可简记为 $P_{i,j}$. 显然灰度共生矩阵是一个对称矩阵, 其阶数由图像中的灰度级个数决定. 灰度共生矩阵的各元素值由下式求得, 即

$$P(i, j|d, \theta) = \frac{P(i, j|d, \theta)}{\sum_i \sum_j P(i, j|d, \theta)} \quad (4)$$

灰度共生矩阵方法是考虑处于几何位置的一对对像素的灰度相关性并以这一对对像素出现某种灰

度的条件概率来表征纹理, 可以定量地描述图像纹理特征. 本文计算和选用能量、惯性矩、熵、相关、局部平稳这五种特征.

1) 能量, 即

$$E(d, \theta) = \sum_{i,j} \{P(i, j|d, \theta)\}^2$$

他是对图像灰度分布均匀性的度量. 当共生矩阵中的元素 $P(i, j|d, \theta)$ 较集中地分布在主对角线附近时, 从图像的整体来观察纹理较粗, 相应的能量值较大; 反之能量值较小.

2) 惯性矩, 即

$$I(d, \theta) = \sum_{i,j} (i - j)^2 P(i, j|d, \theta)$$

可以理解为图像纹理清晰程度. 在图像中, 纹理的沟纹越深, 其数值就越大, 图像的视觉效果越清晰.

3) 熵, 即

$$H(d, \theta) = - \sum_{i,j} \{P(i, j|d, \theta)\} \cdot \log\{P(i, j|d, \theta)\}$$

熵是对图像的信息量的度量, 纹理信息也属于图像信息. 若图像没有任何纹理, 则灰度共生矩阵几乎为零矩阵, 熵值接近为零. 若图像充满细纹理, 当灰度共生矩阵中各 $P(i, j|d, \theta)$ 数值相差不大且较分散时, 熵值较大, 反之, 若 $P(i, j|d, \theta)$ 的数值较集中分布, 且数值差别较大时, 熵值较小.

4) 相关, 如下式:

$$C(d, \theta) = \frac{\sum_{i,j} (i - \mu_x)(j - \mu_y) P(i, j|d, \theta)}{\sigma_x \sigma_y}$$

$$\begin{aligned}
 \text{其中: } \mu_x &= \sum_i i \sum_j P(i, j|d, \theta); \mu_y = \sum_j j \sum_i P(i, j|d, \theta); \sigma_x = \\
 &\sum_i (i - \mu_x)^2 \sum_j P(i, j|d, \theta); \sigma_y = \\
 &\sum_j (j - \mu_y)^2 \sum_i P(i, j|d, \theta).
 \end{aligned}$$

相关是用来描述矩阵中行或列元素之间相似程度的, 它是灰度线性关系的度量. 例如, 某图像具有水平方向的纹理, 则图像在水平方向上的灰度共生矩阵的相关值往往大于其他方向灰度共生矩阵的相关值.

5) 局部平稳, 如下式:

$$L(d, \theta) = \sum_{i,j} \frac{1}{1 + (i - j)^2} P(i, j|d, \theta)$$

基于共生矩阵提取图像纹理特征的方法目前较为成熟、实用, 也是人们公认的一种纹理分析的重要方法^[5]. 本文利用此方法提取简单纹理图像的纹理

特征.

3.2.2 小波变换提取特征

近年来,建立在多尺度分析和空频分析基础上的纹理分析方法得到了广泛关注,而小波变换作为一种多分辨率分析工具,为不同尺度上信号的分析 and 表征提供了精确而统一的框架.小波变换在纹理分析中的应用是 Mallat首先提出的,随后研究者提出了许多基于小波变换的纹理分类方法.

本文首先通过 Daubechies4小波对图像进行分解,然后提取一层分解的四个频带小波系数的均值和方差作为图像的纹理特征^[6].

一幅图像的四频带分布如图 2所示.

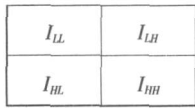


图 2 图像的四频带分布

进行一层变换后, (i, j) 点的小波系数为 c_{ij} ,则任一频带的均值和方差计算公式为

$$\mu = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N c_{ij} \quad (5)$$

$$D = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (c_{ij} - \mu)^2} \quad (6)$$

纹理特征向量为

$$T = [\mu_{LL}, DLL, \mu_{HL}, DHL, \mu_{LH}, DLH, \mu_{HH}, DHH]$$

小波分解提取纹理特征的方法适用于纹理特征比较复杂的图像.

3.3 加权欧氏距离度量

欧氏距离是最常用的度量方式,通常用来计算图像特征的各分量之间正交或无关的图像间的距离.

加权的欧氏距离度量公式如下:

$$D = (A - B)^t C (A - B) \quad (7)$$

式中: A 和 B 表示两幅图像对应的特征; C 是权值矩阵; $C_{n \times n} = [c_{ij}]$, 权 c_{ij} 表示 A 中第 i 个元素和 B 中第 j 个元素的相似程度.

本文对颜色特征向量设置权重为 0.7,对纹理特征向量设置权重为 0.3,这是因为实验发现,在检索时颜色特征对图像的影响比较大.

3.4 基于实例的图像自动语义标注

根据图像检索的结果,本文采用基于实例的机器翻译方法 (Example - Based Machine Translation, EBMP)^[7]对图像进行标注.基于实例的机器翻译的基本思想是在已经收集的双语实例库中检索出与待

翻译部分最相似的翻译实例 (通常是句子),再对实例的译文通过替换等一系列变形操作实现翻译.将基于实例的机器翻译思想应用于自动图像标注中,把带有标注的训练样本集当作一种标注经验,在给出了一幅图像的视觉描述后,从经验库中检索出与之视觉相似的图像,并且通过模仿这些例子图像的标注,对图像实施标注^[8].

4 实验结果分析

图像库中的图像千差万别,有时图像灰度起伏变化剧烈,呈现出很强的纹理特性,有时图像由若干个平稳但颜色有差异的区域构成.颜色和纹理都只是描述了图像的一部分特征,单纯的只用一种特征来表征图像的信息并不能全面地描述图像所包含的内容.为了提高自动标注的准确性,首先要提高图像检索的精确度,本文对简单纹理图像采用颜色和共生矩阵纹理综合特征,示例如图 3所示;对复杂纹理图像采用颜色和小波分解纹理综合特征检索标注图像,示例见图 4.

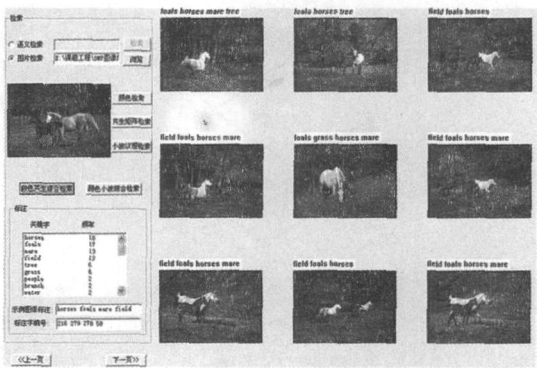


图 3 简单纹理图像检索与标注

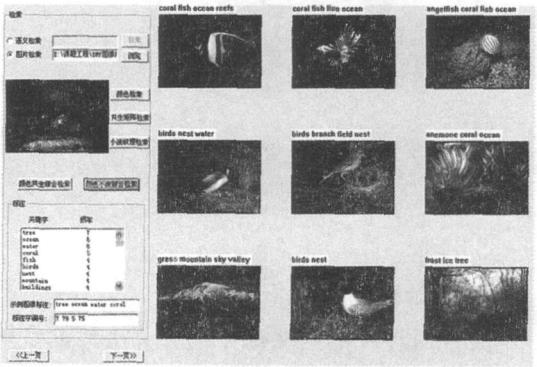


图 4 复杂纹理图像检索与标注

表 1显示了此语义自动标注系统的标注性能.

表 1 标注性能

评价准则	平均精确度 / (%)
查全率	36.1
查准率	40.2

5 结 语

图像的自动语义标注能够将图像的视觉特征转化为图像的标注字信息,这为用户的使用带来了极大的方便.本文提出了一种基于实例的自动语义标注方法,该方法首先针对不同纹理图像进行图像检索,检索出与示例图像最相似的图像,然后根据标注实例确定示例图像的标注信息.

本文给出的这种标注方法容易受到噪声的污染,也不能明显提高自动标注的性能,但是它可以方便地扩展为半自动标注方式,为标注者提供了一个简单方便的标注接口.如何提高检索的准确度以及标注的精确度仍然是今后要深入研究的问题.

(编辑:付长纓)

(上接第 37 页)

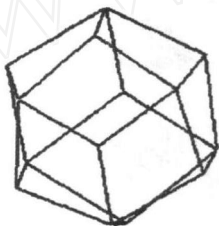


图 7 本文算法细分 1 次

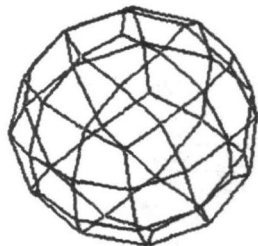


图 8 本文算法细分 2 次

5 结 语

本文针对传统的网格二分细分算法数据量增长较快的问题,给出了一种基于四边形网格的细分算法,该算法不仅具有 4 - 8 细分算法所拥有的数据量

参 考 文 献:

- [1] 吴越聪.多媒体交叉参照检索和语义自动标注[D].杭州:浙江大学,2005.
- [2] HNANG J. Color-spatial Image Indexing and Applications[D]. New York, USA: Cornell University, 1998: 20 - 32.
- [3] 刘鹏宇.基于内容的图像特征提取算法的研究[D].长春:吉林大学,2004.
- [4] HARALICK R M, SHANGMUGAM K. Texture Feature for Image Classification [J]. IEEE Transactions on Systems, 1973, SMC - 3 (6): 768 - 780.
- [5] 龚 坚.彩色图像分割和纹理分隔[D].南京:东南大学,1998,38 - 40.
- [6] WANG JAMES Z. SMPLcity: Semantics-Sensitive Integrated Matching for Picture Libraries[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23 (9): 947 - 963.
- [7] SOMERS H. Review Article: Example - based Machine Translation[J]. Machine Translation, 1999, 14 (2): 113 - 157.
- [8] 于林森.图像检索中缩小语义差距的几个关键技术研究[D].哈尔滨:哈尔滨工业大学,2006.

增长相对缓慢的优点,还具有 $\sqrt{2}$ 细分算法几何操作简单有效的特点.由于本文细分算法对初始网格的拓扑变更,因此第一次细分会产生扭曲现象,但后面的细分会逐步光滑.今后研究工作的重点是保持初始网格的几何特性,实现简单可行的几何操作,使得细分数据量增长缓慢.

参 考 文 献:

- [1] ZORND, SCHRODER P. Subdivision for Modeling and Animation[C]//SIGGRAPH 2000 Course Notes, Course 23. New Orleans: ACM SIGGRAPH, 2000: 13 - 16.
- [2] LOOP C. Smooth Subdivision Surfaces Based on Triangles [D]. Utah: University of Utah, 2005.
- [3] WANG G J, WANG G Z, ZHENG J M. Computer Aided Geometric Design[M]. Beijing: China Higher Education Press, 2006.
- [4] 李桂清,马维银,鲍虎军.带尖锐特征的 Loop 细分曲面拟合系统[J].计算机辅助设计与图形学学报,2005,17 (6): 1179 - 1185.
- [5] WANG Jianwei, ZHANG Zeyin, HUANG Daren. Edge-flipping Subdivision for Quad Meshes[J]. Journal of Computer-Aided Design & Computer Graphics, 2004 (1): 23 - 28.

(编辑:付长纓)