

(12) 发明专利

(10) 授权公告号 CN 102073748 B

(45) 授权公告日 2012. 07. 25

(21) 申请号 201110054624. 8

CN 101315663 A, 2008. 12. 03,

(22) 申请日 2011. 03. 08

CN 101241504 A, 2008. 08. 13,

(73) 专利权人 武汉大学

审查员 张文

地址 430072 湖北省武汉市武昌区珞珈山武汉大学

(72) 发明人 邵振峰 朱先强 刘军

(74) 专利代理机构 武汉科皓知识产权代理事务所(特殊普通合伙) 42222

代理人 张火春

(51) Int. Cl.

G06F 17/30(2006. 01)

G06K 9/62(2006. 01)

(56) 对比文件

CN 101692224 A, 2010. 04. 07,

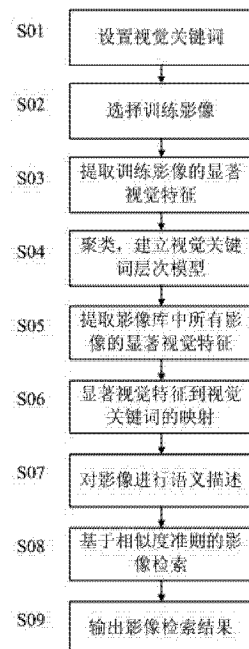
权利要求书 1 页 说明书 9 页 附图 2 页

(54) 发明名称

一种基于视觉关键词的遥感影像语义检索方法

(57) 摘要

本发明涉及一种基于视觉关键词的遥感影像语义检索方法,包括以下步骤:先设定描述影像库中影像内容的视觉关键词,从影像库中选取训练影像,提取每一幅训练影像的显著视觉特征,包括显著点、主色调和纹理,然后通过聚类算法的聚类中心得到关键模式,采用高斯混合模型建立视觉关键词层次模型,再提取影像库中所有影像的显著视觉特征,设定权重参数,构建描述该影像语义的视觉关键词特征向量,最后按照相似度准则计算待检索影像与所有影像的相似度,并按照相似度高低顺序输出检索结果。本发明通过视觉关键词建立低层显著视觉特征与高层语义信息之间的关联,能有效提高影像检索的查全率和查准率,同时本发明提供的技术方案具有良好的扩展性。



1. 一种基于视觉关键词的遥感影像语义检索方法,其特征在于包括以下步骤:

步骤一,设置能够描述影像库中影像类型的视觉关键词,并从影像库中分别选出能反映各视觉关键词的若干幅影像,作为训练影像;

步骤二,提取所有训练影像的各类显著视觉特征,提取的显著视觉特征包括显著点、对象驱动的主色调和纹理;

步骤三,对所有训练影像,将得到的各类显著视觉特征分别采用聚类算法进行聚类,得到与视觉关键词个数相等的聚类中心,将每个聚类中心映射为一个关键模式;采用高斯混合模型拟合任一显著视觉特征属于每类视觉关键词的概率密度函数,高斯混合模型参数估计来自训练影像,拟合方法采用期望最大化估计方法,从而建立起视觉关键词层次模型;

步骤四,采用与步骤二一致的方式分别提取影像库中所有影像的各类显著视觉特征;

步骤五,针对影像库中的每一幅影像,由步骤三所得概率密度函数计算显著视觉特征属于每类视觉关键词的概率,若属于某类视觉关键词的概率最大,则认为显著视觉特征属于该类视觉关键词,从而实现显著视觉特征到视觉关键词的映射;

步骤六,针对影像库中的每一幅影像,根据预设的设定权重参数,统计每类视觉关键词在该影像中出现的频率,进而构建描述该影像语义的视觉关键词特征向量;进行设定权重参数的预设时,显著点特征赋以均值权重,主色调特征和纹理特征以自身对象区域的面积为权重;

步骤七,采用预设的相似性度量准则,通过视觉关键词特征向量计算待检索影像与影像库中所有影像的相似性,将检索结果按照相似性从高到低排序并输出。

2. 根据权利要求1所述基于视觉关键词的遥感影像语义检索方法,其特征在于:提取显著视觉特征的实现方式如下,

(1) 利用 SIFT 图像局部特征描述算子提取所有训练影像的显著点,从而获取影像的显著点特征,每个显著点用一个 128 维特征向量来表示;

(2) 对所有训练影像进行基于 Quick Shift 算法的过分割,对过分割结果进行区域合并,然后对一致性对象区域采用 HSV 模型,根据其色调通道的量化结果提取出各区域的主色调,从而获取影像的主色调特征,每一个对象区域的主色调特征用一个特征向量来表示;

(3) 对所有训练影像进行基于 Quick Shift 算法的过分割,对过分割结果进行区域合并,然后对一致性对象区域采用小波变换,获取各尺度高频分量的均值和方差作为纹理描述子,从而获取影像的纹理特征,每一个对象区域的纹理特征用一个特征向量来表示。

3. 根据权利要求1或2所述基于视觉关键词的遥感影像语义检索方法,其特征在于:步骤三中,所采用的聚类算法为 K 均值或 ISODATA 算法。

4. 根据权利要求1或2所述基于视觉关键词的遥感影像语义检索方法,其特征在于:步骤七中,预设的相似性度量准则为 KL 散度的一阶近似距离。

一种基于视觉关键词的遥感影像语义检索方法

[0001] 技术领域

[0002] 本发明涉及图像处理技术领域,更具体地说,涉及一种基于视觉关键词的遥感影像语义检索方法。

[0003] 背景技术

[0004] 遥感影像数据应用正面临着“数据又多又少”的矛盾。一方面,随着航空航天及各类传感器技术、计算机网络技术、数据库技术等飞速发展,可获取的各种遥感影像数据产品,特别是高空间分辨率遥感影像数据每天都在以惊人的速度增长;另一方面,在如此浩瀚的遥感影像数据仓库中,人们却普遍感觉到要想快速查找到感兴趣的目标并非易事。这是由于遥感影像数据本身具有空间性、多样性、复杂性和海量性等特点,使得目前缺乏对海量遥感影像数据的有效检索方法,已经阻碍到遥感影像数据的应用。遥感影像的高效检索是解决海量遥感数据和人们对遥感数据应用日益增长的需求之间的矛盾的关键,是目前遥感应用领域亟待解决的难题,也是学科研究的前沿。

[0005] 在遥感影像检索研究所涉及的各项关键技术中,目前的研究重点主要集中在遥感影像的可视化特征(包括光谱特征、纹理特征、形状特征及组合特征)提取及其相似性匹配算法上,其中对纹理特征的研究和应用最为广泛和深入,目标形状特征的描述和提取相对光谱特征、纹理特征而言是一个非常复杂的问题,至今还没有“形状”的确切数学定义,目前基于内容的影像检索中,目标的形状通常采用边缘和区域特征来描述,但是对于目标的边缘和区域特征描述算子及其形状相似性匹配的研究仍不尽完善;基于目标形状特征描述和提取的困难,尽管人们越来越认识到其在遥感影像检索中的重要意义,但是研究成果却十分有限。在基于组合特征的遥感影像检索方面,主要有基于色调和纹理组合特征的检索。数据预处理(自动分块组织或预处理)和可视化特征相结合的算法也是以纹理特征为主。

[0006] 由于低层可视化特征不能直观地反映影像的语义信息,在没有专家库或者领域知识库的辅助下,通常都会产生“所求非所得”的检索结果。要解决这一问题,提高检索效率和检索准确率,在检索方法上必须突破对可视化特征的依赖。遥感影像高层语义特征包含了人们对影像内容的理解,基于语义的检索方法不仅顾及颜色、纹理、形状等视觉特征,而且注重对影像内容的语义描述,因此语义检索比基于视觉特征的内容检索更加充分、准确,智能性更高。然而,目前基于语义的遥感影像检索仍停留在探索阶段。

发明内容

[0007] 本发明的目的在于针对现有技术的缺点和不足,提供一种基于视觉关键词的遥感影像语义检索方法,通过符合人类视觉感知特性的影像分析方法,将复杂的遥感影像特征抽象为具有语义信息的视觉关键词,通过视觉关键词建立底层特征、中层对象及高层语义信息之间的关联,提供的方法能适用于各种不同类型遥感影像检索领域。

[0008] 本发明所采用的技术方案是一种基于视觉关键词的遥感影像检索语义方法,包括以下步骤:

[0009] 步骤一,设置能够描述影像库中影像类型的视觉关键词,并从影像库中分别选出能反映各视觉关键词的若干幅影像,作为训练影像;

[0010] 步骤二,提取所有训练影像的各类显著视觉特征;

[0011] 步骤三,对所有训练影像,将得到的各类显著视觉特征分别采用聚类算法进行聚类,得到与视觉关键词个数相等的聚类中心,将每个聚类中心映射为一个关键模式;采用高斯混合模型拟合任一显著视觉特征属于每类视觉关键词的概率密度函数,高斯混合模型参数估计来自训练影像,拟合方法采用期望最大化估计方法,从而建立起视觉关键词层次模型;

[0012] 步骤四,采用与步骤二一致的方式分别提取影像库中所有影像的各类显著视觉特征;

[0013] 步骤五,针对影像库中的每一幅影像,由步骤三所得概率密度函数计算显著视觉特征属于每类视觉关键词的概率,若属于某类视觉关键词的概率最大,则认为显著视觉特征属于该类视觉关键词,从而实现显著视觉特征到视觉关键词的映射;

[0014] 步骤六,针对影像库中的每一幅影像,根据预设的设定权重参数,统计每类视觉关键词在该影像中出现的频率,进而构建描述该影像语义的视觉关键词特征向量;

[0015] 步骤七,采用预设的相似性度量准则,通过视觉关键词特征向量计算待检索影像与影像库中所有影像的相似性,将检索结果按照相似性从高到低排序并输出。

[0016] 而且,步骤二和步骤四中,提取的显著视觉特征包括显著点、对象驱动的主色调和纹理。

[0017] 而且,提取显著视觉特征的实现方式如下,

[0018] (1) 利用 SIFT 图像局部特征描述算子提取所有训练影像的显著点,从而获取影像的显著点特征,每个显著点用一个 128 维特征向量来表示;

[0019] (2) 对所有训练影像进行基于 Quick Shift 算法的过分割,对过分割结果进行区域合并,然后对一致性对象区域采用 HSV 模型,根据其色调通道的量化结果提取出各区域的主色调,从而获取影像的主色调特征,每一个对象区域的主色调特征用一个特征向量来表示;

[0020] (3) 对所有训练影像进行基于 Quick Shift 算法的过分割,对过分割结果进行区域合并,然后对一致性对象区域采用小波变换,获取各尺度高频分量的均值和方差作为纹理描述子,从而获取影像的纹理特征,每一个对象区域的纹理特征用一个特征向量来表示。

[0021] 而且,步骤六中,设置权重参数时,显著点特征赋以均值权重,主色调特征和纹理特征以自身对象区域的面积为权重。

[0022] 而且,步骤三中,所采用的聚类算法为 K 均值或 ISODATA 算法。

[0023] 而且,步骤七中,预设的相似性度量准则为 KL 散度的一阶近似距离。

[0024] 本发明提供的技术方案的有益效果为,通过视觉关键词的层次模型建立低层显著视觉特征与高层语义信息之间的关联,缩小了低层显著视觉特征与高层语义之间的“语义鸿沟”,为从海量遥感影像库中快速定位和查找感兴趣目标提供一个新的解决途径,能有效提高影像检索的查全率和查准率。同时本发明提供的技术方案具有良好的扩展性,所采用的显著视觉特征包括但不限于本发明中所使用的显著点、主色调和纹理,只要是符合人类视觉特性的特征,都能顺利地纳入本发明提供的技术方案中。

附图说明

[0025] 图 1 为本发明实施例的流程图。

[0026] 图 2 为本发明实施例的效果示意图。

具体实施方式

[0027] 本发明提出的基于视觉关键词的遥感影像语义检索方法先设置反映影像库内容的视觉关键词,选择训练影像,提取显著视觉特征,然后建立视觉关键词层次模型,实现低层视觉特征与高层语义之间的关联,对遥感影像进行语义建模和描述,最后采用相似度准则对影像库中的影像进行检索。其中主要包括训练影像显著视觉特征提取、建立视觉关键词层次模型、遥感影像语义建模和基于相似度准则的影像检索四个过程。

[0028] 为详细说明具体实施方式,参见图 1,实施例流程如下:

[0029] 步骤 S01,设定用于描述影像库中影像内容的视觉关键词。

[0030] 实施例所采用的数据来自于 2009-12-27 采集的郑州城区 WorldView 影像,影像空间分辨率为 0.5 米,大小为 8740*11644,将影像按照 Tiles 分块方式分割为 320*320 大小的子块,构成 1036 幅子影像的检索影像库。由于遥感影像覆盖面积大,地物种类复杂,根据地物在影像上所呈现出的特征,可将地表的覆盖类型分为以下八类:农田、裸露地、道路、密集居住区、稀疏居住区、广场、立交桥、绿地。因此,实施例设定了八类视觉关键词:农田、裸露地、道路、密集居住区、稀疏居住区、广场、立交桥、绿地。

[0031] 步骤 S02,根据设定的视觉关键词,从影像库中找出能反映这些关键词的地物内容单一的影像,用做训练影像。

[0032] 实施例分别选择与农田、裸露地、道路、密集居住区、稀疏居住区、广场、立交桥、绿地相应类型纯净的影像块作为训练样本。

[0033] 步骤 S03,采用特征提取算法,对每一幅训练影像,提取显著视觉特征。

[0034] 遥感影像覆盖地物种类繁多,单一的特征空间很难形成对地物的有效区分,本发明实施例选择代表局部特征的显著点、对象驱动的主色调及纹理特征三大类显著视觉特征。具体实施时,还可以根据需要进行选择形状等特征。

[0035] 为便于实施参考起见,下面对这三类显著视觉特征的提取分别描述:

[0036] (1) 显著点特征提取:对于遥感影像来说,角点是表示和分析影像的重要特征,从显著点的邻域中提取的图像特征可以有效地反映图像的局部信息,并且当人们关注一幅影像时,往往容易被图像中显著的部分吸引,其中部分视觉焦点即影像中的角点。本发明采用 SIFT 图像局部特征描述算子提取显著点特征。SIFT 特征向量对旋转、尺度缩放、亮度变化保持不变,能尽可能减少不同空间分辨率、不同光照条件对显著点特征提取的影响。

[0037] (2) 主色调特征提取:采用 Quick Shift 分割算法,利用空间一致性和颜色一致性对影像进行过分割,然后对过分割所得到的对象区域进行合并,得到合并后的对象区域,

[0038] HSV 模型直接对应于人眼色彩视觉特征的三要素,三个颜色通道各自独立,根据其色调通道的量化结果可提取出图像的主色调。本发明首先将色调通道量化为 M 个子区,将上述过分割合并后提取的每一个对象区域分别用量化后的主色调直方图表示, A 为影像分割后的某一区域,则对象 O 的主色调特征矢量可表示如下式所示,其中 H_i 为

第 i 类色调出现的频率, i 取 1 至 n 。

[0039]

$$C_{(R,O)} = (h_1, h_2, \dots, h_n)$$

[0040] (3) 纹理特征提取: 采用 Quick Shift 分割算法对影像进行过分割, 然后采用主色调特征提取方法中所述的对象区域合并方法进行对象区域的合并, 最后采用多尺度多方向纹理特征描述方法进行纹理特征提取。本发明采用小波变换后各尺度高频分量的均值和方差作为纹理描述子, 这种描述方法所得的特征向量维数低、效率高并且具有一定的代表性, 同时本发明对小波变换系数进行了归一化处理, R 为影像分割后的某一区域, 对象 O 的纹理特征矢量表示如下式所示, $\bar{\mu}_i, \sigma_i$ 为第 i 个分量的归一化均值和方差, i 取 1 至 n , n 为分量总数, 等于尺度数量的三倍。

[0041]

$$T_{(R,O)} = (\bar{\mu}_1, \sigma_1, \bar{\mu}_2, \sigma_2, \dots, \bar{\mu}_n, \sigma_n)$$

[0042] 具体实施时, 可以在实现 (2) 和 (3) 时, 为过分割过程设定限定条件, 以提高分割质量, 例如限定这些合并后的区域满足以下三个条件:

[0043] a. 对象区域内部差别应尽可能小;

[0044] b. 对象与其周围邻接对象之间差别应较大;

[0045] c. 对象区域面积应大于某一阈值。

[0046] 条件 a 的目的是限定对象为纯净端元, 提高语义赋值的准确性; 条件 b 控制对象区域合并的程度; 条件 c 的主要目的是剔除干扰视觉判定的琐碎区域, 突出显著的主要特征, 提高算法的效率。假设 R 为影像分割后的某一区域, 区域 R 的内部差别 $D_{(R)}$ 定义为:

[0047]

$$D_{(R)} = w_1 \cdot \sigma_R + w_2 \cdot \sigma_R$$

[0048] 其中 w_1, w_2 为特征权重, 满足 $w_1 + w_2 = 1$; σ_R 为区域内部颜色的标准差, σ_R 为区域形状指数, 定义如下所示:

[0049]

$$\sigma_R = \sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 / n - 1} \quad X_i \in R_C, \bar{X} = \sum_{i=1}^n X_i / n$$

[0050]

$$\sigma_R = A_R \frac{l_R}{b_R}$$

[0051] R_C 为区域 R 内部颜色灰度值的集合, A_R 为区域 R 的面积, l_R 为区域周长, b_R 为区域最小外接矩形周长。区域间的差别定义公式如下所示:

[0052]

$$D_{(R,R')} = D_{R \cup R'} \cdot A_{R \cup R'} - (D_{(R)} \cdot A_R + D_{(R')} \cdot A_{R'})$$

[0053] R' 为区域 R 的某一邻接区域, $D_{R \cup R'}, A_{R \cup R'}$ 分别为合并后区域的内部差别和面积。

进行区域合并时,首先判定区域面积,若满足条件 c 则遍历该区域的邻接区域,当 $D_{(u,v)}$ 小于一定的阈值时则合并区域,否则不对对象进行操作;若不满足条件 c 则遍历邻接区域选择 $D_{(u,v)}$ 值最小的区域进行合并。

[0054] 步骤 S04,对所有训练影像,将得到的各类显著视觉特征分别采用聚类算法进行聚类,得到与视觉关键词个数相等的聚类中心,将每个聚类中心映射为一个关键模式;采用高斯混合模型拟合任一显著视觉特征属于每类视觉关键词的概率密度函数,高斯混合模型参数估计来自训练影像,拟合方法采用期望最大化估计方法,从而建立起视觉关键词层次模型。

[0055] 实施例建立视觉关键词层次模型的具体实现过程:由一幅训练影像可以提取出的三类显著视觉特征,包括显著点、主色调和纹理。对于每一类显著视觉特征的所有特征向量,采用 K 均值或者 ISODATA 聚类方法进行聚类,得到与视觉关键词个数一致的聚类中心,将每个聚类中心映射为一个关键模式。采用高斯混合模型拟合任一特征向量属于每类视觉关键词的概率密度函数,模型参数估计来自训练影像,方法采用期望最大化估计方法,从而建立起视觉关键词层次模型。因为实施例中,每个类别 GMM 模型中包含的广义高斯分布分量(GGD)数量为 8。通过训练样本影像特征空间的子空间的中心也即关键模式,得到各关键模式的高斯分布,多个关键模式的独立组合也即高斯分布的合并构成一个含有语义的关键词,整幅影像表示为影像中各类语义关键词的分布直方图,至此即可完成无语义的视觉特征到含语义的关键词标签的建模过程。

[0056] 为便于实施参考起见,本步骤提供相关说明如下:

[0057] 遥感影像可表示为从像素到局部显著特征或基元、目标对象和场景的层次模型,在模型各层次上都包含一系列描述视觉信息的视觉词汇,从而形成影像表达场景中语义标签与图像特征的连接。设某一影像视觉词汇定义为集合 \mathcal{O} ,其中 i 为词汇类型标识, α_i 为任一视觉词汇元素, N 为词汇类型总数, R 为整个可能的词汇空间集合。

[0058]

$$\mathcal{O} = \{\alpha_i | \alpha_i \in R, i = 1, 2, \dots, N\}$$

[0059] 视觉词汇的聚合可以产生任意尺度的影像,其中一些聚合属于全局聚合,这些词汇组合可还原出影像中大部分信息,文中将这些词汇组合模式称之为关键模式,以上式为基础,某一关键模式 KP 定义为:

[0060]

$$KP = \{\cup \alpha_i | \alpha_i \in R, i = 1, 2, \dots, N\}$$

[0061] 关键模式集合需满足近似完备条件即 $\bigcup_{i=1}^M KP_i \cong f_{(n)}$,其中 M 为关键模式总数, $f_{(n)}$

为图像特征空间。因此,视觉关键词模型即为对关键模式的建模。自动聚类算法在特征空间自动寻找聚类中心方面得到广泛的应用,本发明即通过常用的 K 均值或 ISODATA 自动聚类算法从庞杂的视觉词汇表中寻找关键模式;高斯混合模型能够用参数化的方法描述样本空间中的数据分布,将高斯混合模型的参数作为图像的特征具有简洁高效的优点,假设特征空间中视觉词汇分布服从高斯分布,则关键模式分布服从高斯混合分布 GMMs,每个类别

的语义关键词即由 K 个词汇空间的聚类中心组成, 以特征分布 $x \in R^d$ 为例, 其属于关键词 m 的概率密度函数可表示为:

[0062]

$$P(x|\theta_m) = \sum_{j=1}^K a_{mj} \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma_{mj}|^{\frac{1}{2}}} \exp[-\frac{1}{2}(x-\mu_{mj})^T \Sigma_{mj}^{-1} (x-\mu_{mj})]$$

[0063] 其中 R^d 为特征分布空间, K 为 GMM 模型的维数, θ_m 为视觉关键词 m 的 GMM 模型参数, $\theta_m = \{a_{mj}, \mu_{mj}, \Sigma_{mj}\}_{j=1}^K$, a_{mj} 为模型中第 j 个高斯变量的混合系数, μ_{mj} 为第 j 个高斯变量的均值, T 表示矩阵的转置, Σ_{mj} 为相应的协方差矩阵, d 为样本维数。模型参数估计以所选取的各类别影像为训练数据, 方法采用期望最大化估计方法, 每个类别的视觉关键词对应特征空间唯一的 GMM 分布。

[0064] 以下以显著点特征为例, 详细描述采用高斯混合模型拟合任一特征向量属于每类视觉关键词的概率密度函数的过程: 首先采用聚类算法将 128 维的 SIFT 特征向量聚类为 K 个子空间, 每个子空间的中心代表一个关键模式, 假设每个 SIFT 特征关键模式服从高斯分布 $N(x, \mu, \Sigma)$, 则可利用 GMM 模型拟合 K 个 SIFT 特征关键模式的组合表达一个 SIFT 类型的语义关键词。

[0065] 假设影像语义关键词数目为 M , 则训练后影像 I 的 SIFT 类型视觉关键词层次模型 $VKM_{(I, SIFT)}$ 表示为下式所示:

[0066]

$$VKM_{(I, SIFT)} = \{UP_j | j=1, 2, \dots, M\}, \text{ where } P(j|x) = \sum_{i=1}^K a_i N_j(x, \mu_i, \Sigma_i)$$

[0067] 其中 a_i 为第 i 个关键模式对应的高斯混合模型系数, N_j 为 x 属于第 j 个关键词的概率分布, 是一个高斯分布, μ_i 为第 i 个关键模式对应的特征矢量, Σ_i 第 i 个子空间的协方差矩阵。计算所提取的显著点描述子 x 属于第 j 个关键词的后验概率 $P(j|x)$, 采用最大后验概率分类器 MAP (maximum a posteriori) 将 SIFT 描述矢量 x 标记为 j^* 关键词, 其中:

[0068]

$$j^* = \arg \max [P(j|x)]$$

[0069] 如果将每幅影像看成由视觉关键词库中的若干关键词组成的“文本”, 则在文本检索技术中的经典统计方法 TF-IDF (term frequent-Inverse document frequency) 即词条频率 - 倒排文本频率, 可用以评估字词对于一个文件集或一个语料库中某份文件的重要程度。字词的重要性随它在文件中出现的次数成正比增加, 但同时会随着它在语料库中出现的频率成反比下降。本发明以单个点为统计单元, 影像 I 基于视觉关键词的 SIFT 特征矢量可表示如下式所示:

$$[0070] \quad S_{(I, SIFT)} = (P_{(w1)}, P_{(w2)}, \dots, P_{(wM)}), \quad P_{(wi)} = \frac{n_i}{n} \log \frac{F}{n_i}$$

[0071] $P_{(m)}$ 即为视觉关键词 i 的词条频率 - 倒排文本频率, i 取 1 至 M , 其中: I 是输入影像, n_i 是第 i 个关键词在影像 I 中出现的次数, n_i 为影像 I 中关键词的总数, n_i 为第 i 个关键词在整个影像库中出现的次数, F 为整个影像库中的影像数目; $TF_i = n_i / n_i$ 即为词条频率, $IDF_i = \log F/n_i$ 为倒排文本频率。

[0072] 主色调和纹理的拟合过程与显著点的过程是类似的, 本发明不予赘述。经过拟合, 即建立起了视觉关键词层次模型, 实现了低层视觉特征到高层语义关键词之间的联系。

[0073] 步骤 S05, 采用特征提取算法, 对影像库中的每一幅影像, 都提取显著视觉特征。实施例提取的显著视觉特征包括显著点、主色调和纹理, 得到一系列描述影像特征的特征向量, 即相应的显著点特征向量、主色调特征向量和纹理特征向量。具体实施过程与步骤 S03 一致即可。

[0074] 步骤 S06, 根据步骤 S04 中建立的视觉关键词层次模型, 将影像库中每一幅影像的三类视觉特征的特征矢量代入视觉关键词层次模型中, 计算每类视觉特征中的所有特征向量属于视觉关键词的概率, 按 $J^* = \arg \max [P(J | x)]$ 的原则将特征向量映射为相应的视觉关键词, 从而建立所有视觉特征的特征向量与视觉关键词的对应关系。

[0075] 步骤 S07, 针对影像库中的每一幅影像, 根据预设的设定权重参数, 统计每类视觉关键词在该影像中出现的频率, 进而构建描述该影像语义的视觉关键词特征向量, 从而实现对影像基于视觉关键词的语义建模和描述。具体实施时, 采用不同的权重参数, 得到的检索结果是不一样的, 可以由本领域技术人员事先根据经验设定。

[0076] 总结实施例中遥感影像语义建模的具体实现过程: 在步骤 S05, 对影像库中的所有影像均按照上面的步骤 S03, 先提取显著视觉特征, 包括显著点、主色调和纹理; 然后在步骤 S06, 在视觉关键词层次模型的支持下, 按照 $J^* = \arg \max [P(J | x)]$ 的原则, 每幅影像提取的单个特征矢量均可映射为有限个数的关键词, $I = \{KW_i\}, i = 1, 2, \dots, n$, n 为影像中特征点或对象的个数。但是每个关键词标识影像的贡献率并非完全一样的, 以区域对象为例, 一般来说影像中心区域或较大面积区域对影像解译的贡献要大于边角面积较小的区域, 本发明实施例在步骤 S07 中, 对特征点采用均匀权重, 即每个点特征关键模式的权重均相同, 对区域关键词(主色调、纹理)采用面积因子作为权重参数, 则可以统计出每个关键词在影像中出现的频率, 以单类特征为例, 影像 I 的关键词建模后的特征矢量 \bar{I}_n 可表示为

[0077]

$$\bar{I}_n = \{w_1, w_2, \dots, w_k\}$$

[0078] K 为关键词数目, w_i 为关键词 i 出现的频率, i 取 1 至 K , 对特征点而言, $w_i = n_i / n$, n_i 为第 i 个点特征关键模型出现的次数, n 为所有点特征关键模型出现的总次数; 对主色调或纹理特征而言, $w_i = \sum A_i / A$, A 为影像总面积, A_i 为关键词 i 在影像中的面积大小。由此即可得到影像的归一化视觉关键词特征矢量。

[0079] 步骤 S08, 设置三类视觉关键词特征向量的权重, 以 KL 散度的一阶近似距离为相似性度量方式, 通过视觉关键词特征向量计算待检索影像与影像库中所有影像的相似性。

具体事实时,也可以采用其他现有技术作为相似性度量准则,例如余弦距离、KL 距离, KL 散度的一阶近似距离,欧氏距离,马氏距离。

[0080] 实施例中基于相似度准则的影像检索的具体实现过程:采用特征提取算法,可获取影像中三种类型的关键词分别描述影像的显著点、主色调和纹理特征,关键词的描述借鉴文本表示方法中的 TF-IDF 描述方法,因此,每幅图像都可以描述为某一类或几类视觉关键词的特征矢量 \vec{I} 。

[0081] 从信息论的角度来分析,提取出的特征矢量所表达的含义是一幅影像中视觉关键词出现的概率分布关系,假设各视觉关键词之间相互独立,且服从概率密度函数 $p(X; \theta)$ 的分布,则两幅影像之间的距离也可表示为 Kullback-Leibler 散度,如下式所示:

[0082]

$$D(p(X; \theta) \| p(X; \theta_0)) = \int p(X; \theta) \log \frac{p(X; \theta)}{p(X; \theta_0)} dx$$

[0083] 其中下标 i 和 j 区别表示两幅影像。

[0084] 视觉关键词特征矢量之间的 KL 散度计算公式为:

[0085]

$$KL(KW_{I_1}, KW_{I_2}) = \sum_{n=1}^K (KW_{I_1}^{(n)} - KW_{I_2}^{(n)}) \log \frac{KW_{I_1}^{(n)}}{KW_{I_2}^{(n)}}$$

[0086] 其中下标 I_1 和 I_2 区别表示两幅影像, K 为关键词总数, n 为关键词序号,上式包含对数运算,计算效率低,选择其一阶近似距离 χ^2 可有效降低复杂度,如下:

[0087]

$$\chi^2(KW_{I_1}, KW_{I_2}) = \sum_{n=1}^K \frac{(KW_{I_1}^{(n)} - KW_{I_2}^{(n)})^2}{KW_{I_1}^{(n)} + KW_{I_2}^{(n)}} = \sum_{n=1}^K \chi^2(KW_{I_1}^{(n)}, KW_{I_2}^{(n)})$$

[0088] 一幅影像可以采用显著点、主色调、纹理三类特征矢量进行表达,影像语义分布的相似性大小采用三类特征的加权距离进行计算,如下:

[0089]

$$S_{(I_i, I_j)} = \alpha_1 \cdot dis(S, I_i, I_j) + \alpha_2 \cdot dis(C, I_i, I_j) + \alpha_3 \cdot dis(T, I_i, I_j)$$

[0090] 其中 $0 < \alpha_1 < 1, 0 < \alpha_2 < 1, 0 < \alpha_3 < 1, \alpha_1 + \alpha_2 + \alpha_3 = 1$, S, C, T 分别代表显著点、主色调、纹理, I_i, I_j 区别标识两幅影像。本发明中关键词特征矢量表达方式类似于影像直方图,只是其输入的是不同特征类型的语义关键词,而不是影像的灰度值,这种表达方式与影像大小无关,只需要设定相同的语义维数,不同大小的影像都可采用上式进行相似性度量。

[0091] 步骤 S09,将相似性按照从高到低的顺序进行排序,输出检索结果。

[0092] 在以上步骤 S08 计算检索影像的语义关键词特征矢量与影像库中所有影像的语义关键词特征矢量的相似性后,按照相似性从高到低的顺序输出结果影像,即为检索结果。

[0093] 为了便于说明本发明技术方案效果,采用不同权重的 8 类影像的平均查准率,对比单一纹理特征、单一主色调特征、单一显著点特征与本发明实施例所用综合特征的情况分别进行检索,然后对所得结果进行对比评价,如图 2 所示。定量化评价采取平均查准率,

即返回的前 16 幅影像中含有相似影像的比例。相似性度量准则采用了 KL 散度的一阶近似距离,综合特征检索中显著点、主色调、主纹理三类视觉关键词的权重分别设定为 0.5, 0.25, 0.25。可见,本发明技术方案的准确率更高。

[0094] 以上内容是结合最佳实施方案对本发明说做的进一步详细说明,不能认定本发明的具体实施只限于这些说明。本领域的技术人员应该理解,在不脱离由所附权利要求书限定的情况下,可以在细节上进行各种修改,都应当视为属于本发明的保护范围。

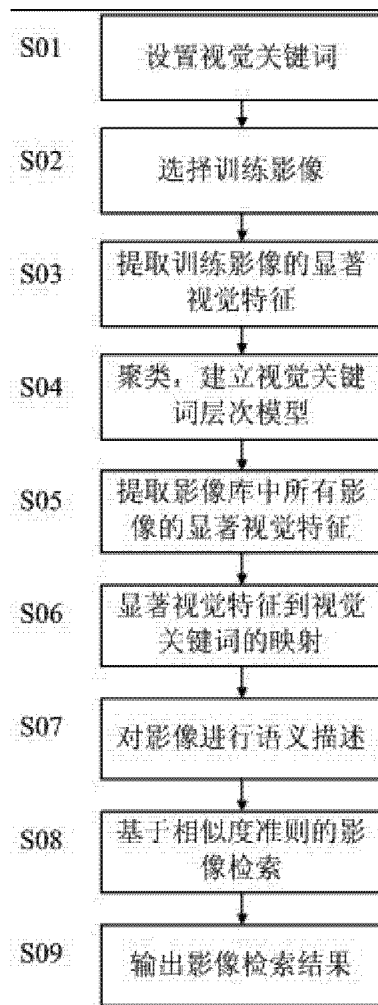


图 1

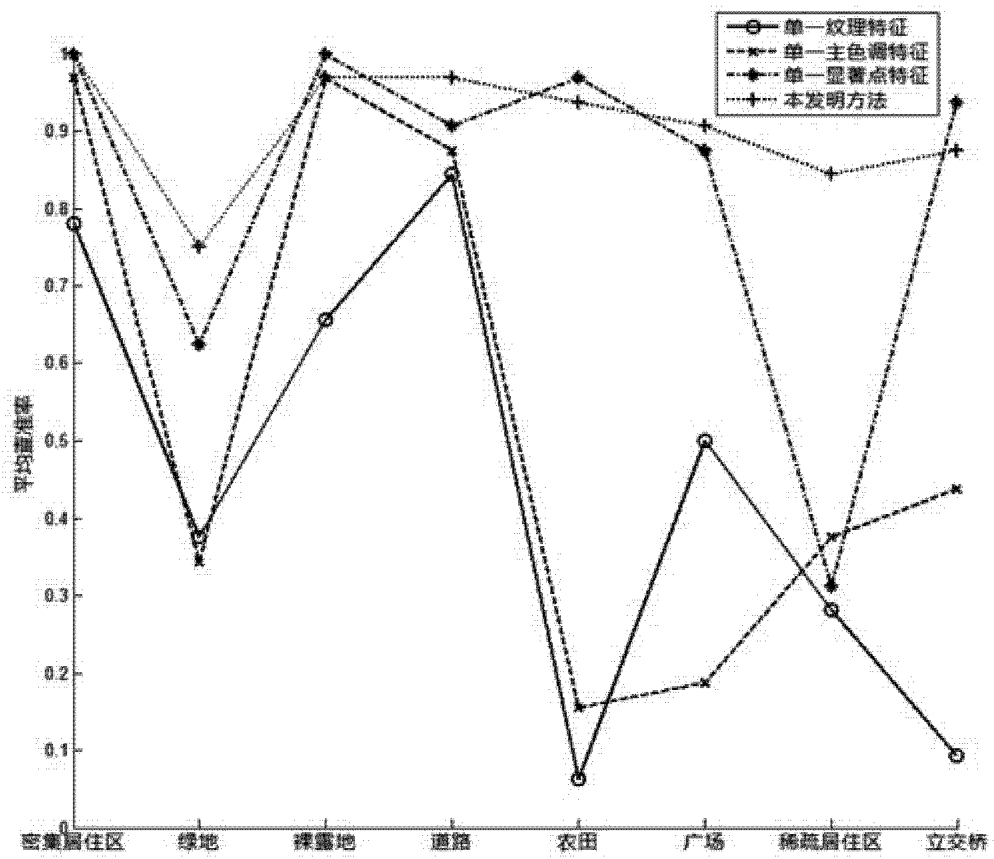


图 2