基于细节层次与最小生成树的三维地形识别与检索*

肖 俊+, 庄越挺, 吴 飞

(浙江大学 人工智能研究所,浙江 杭州 310027)

Recognition and Retrieval of 3D Terrain Based on Level of Detail and Minimum Spanning Tree

XIAO Jun+, ZHUANG Yue-Ting, WU Fei

(Institute of Artificial Intelligence, Zhejiang University, Hangzhou 310027, China)

+ Corresponding author: Phn: 86-571-87951853, Fax: 86-571-87951947, E-mail: xiaojunboy@yeah.net

http://www.cs.zju.edu.cn

Received 2002-07-27; Accepted 2003-01-20

Xiao J, Zhuang YT, Wu F. Recognition and retrieval of 3D terrain based on level of detail and minimum spanning tree. *Journal of Software*, 2003,14(11):1955~1963.

http://www.jos.org.cn/1000-9825/14/1955.htm

Abstract: Image, video, audio and graphics are information media in multimedia. In order to use and manage them effectively, the contents implied by them are needed to analyze. Currently, content-based image (video) and audio retrieval make some progress. However, there is no a very efficient method to perform a similar graphics retrieval, especially a similar 3D graphics retrieval. In this paper, an algorithm is presented to implement the similar 3D object recognition and retrieval. In this algorithm, 3D features are first obtained after the meshes of a 3D object are reduced through level of detail. Since the dimension of the extracted 3D features is very huge, minimum spanning tree (MST) are used to reduce the features, then the recognition of similar 3D objects are realized by Support Vector Machine (SVM). The proposed algorithm works well when it is used to recognize and retrieve a 3D terrain.

Key words: level of detail; minimum spanning tree; support vector machine; 3D retrieval and recognition; feature reduction

摘 要: 图像、视频、音频和图形等均是多媒体数据流中的信息载体,对上述数据所蕴涵的内容进行分析,可以极大地方便人们对它们的使用与管理.基于内容的图像(视频)和音频检索已经取得了不少进展,但是对于图形,特别是3D图形进行识别与检索的有效方法还很少见.提出了对相似3D物体识别与检索的算法,在这个算法中,首先使用细节层次模型对 3D 物体进行三角面片约减,然后提取 3D 物体的特征.由于所提取的特征维数很大,最小生成树(minimum spanning tree,简称MST)被用来对每一个3D物体的特征进行约减,基于约减后的特征,实现了基于支持向

* Supported by the National Natural Science Foundation of China under Grant No.60272031 (国家自然科学基金); the National Research Foundation for the Doctoral Program of Higher Education of China No.20010335049 (国家教育部博士点基金); Zhejiang Provincial Natural Science Foundation of China under Grant No.ZD0212 (浙江省自然科学基金)

第一作者简介: 肖俊(1979－),男,贵州铜仁人,博士生,主要研究领域为机器学习,计算机视觉.

量机的 3D 物体识别与检索方法.这个算法被使用到 3D 丘陵与山地的地形识别中,取得了良好效果.

关键词: 细节层次;最小生成树;支持向量机;3D 检索与识别;特征约减

中图法分类号: TP391 **文献标识码**: A

多媒体数据流中的每种媒质,如图像、视频、音频和图形等,均是信息载体,对它们所蕴涵的语义内容进行分析,可以极大地方便人们对多媒体数据的有效利用与管理.

但是,上面这些媒质不像文本信息那样,可以通过关键字等文字信息对其表达的内容进行标注,于是,基于内容的多媒体分析与检索技术应运而生.

在多媒体分析与检索中,基于内容的视频(图像)检索技术,一般是通过提取颜色、纹理、形状和运动等视觉特征来表征视频(图像)内容所蕴涵的语义,对视频流信息进行结构化,达到检索目的[1,2].作为多媒体重要媒质之一的音频蕴涵了丰富语义,从 20 世纪 90 年代中期开始的基于内容的音频检索就研究如何提取音频信息流中的语义信息,以方便对音频数据进行检索[3,4].

应该讲,无论是基于内容的图像(视频)检索,还是基于内容的音频检索,均是基于视觉或听觉内容相似的检索,而非语义内容检索.因此,在这类数据的检索中,定义一个与人感知一致的相似度衡量方法是非常重要的[5].

尽管如此,随着计算机图形学研究的进展,目前在互联网上,3D 虚拟物体越来越多,如何实现对这些 3D 物体的相似检索,还缺乏一种有效手段.与图像(视频)和音频检索不同,3D 物体检索面临如下挑战:(1) 一般而言,3D 物体均是由三角面片构成的,一个普通 3D 物体所包含的三角面片一般超过万个,复杂的 3D 物体所包含的三角面片可以上亿[6].因此,在检索相似 3D 物体之前,需要对这些三角面片按照细节层次进行约减,以减少计算量;(2) 如何定义 3D 物体的特征,使用这些特征来表征 3D 物体,是第 2 个难题;(3) 即使 3D 物体进行了三角面片约减,其提取的特征仍然很大,特征之间存在冗余,如果不对这些特征通过映射而降维,就会损坏识别效果;(4) 对相似 3D 物体需要构造一个鲁棒的分类识别学习机,使其有良好的学习推广(generalization)能力,即这个分类器不仅要对测试数据表现出良好的识别能力,而且也要对实际未知数据的识别能力表现良好.

基于 3D 相似物体识别与检索所面临的上述挑战,本文提出先使用细节层次和最小扩展树法实现 3D 物体特征约减,然后支持向量机被用来识别的 3D 检索方法,并且将其应用到 3D 地形识别中去.

本文第 1 节介绍了约减三角面片的细节层次原理.第 2 节介绍了提取的 3D 特征.第 3 节介绍使用最小生成树(minimum spanning tree,简称 MST)对所提取的 3D 特征进行降维.第 4 节介绍使用支持向量机(support vector machine,简称 SVM)完成相似 3D 物体的识别与检索.本文最后两节给出实验比较和结论.

1 细节层次原理

在生成 3D 物体的过程中,为了加速图形生成以保证图形的实时绘制,一般使用 3D 物体细节层次模型 LoD(level of detail)[7,8].细节层次模型的主要原理是根据物体对于观察者的重要性选择该物体绘制的细节.LoD 可以在不同层次和不同视觉的条件下,采用不同精细程度的模型来表示同一物体,提高场景的显示速度,实现实时交互.

在 3D 物体显示中,最主要的细节就是构成 3D 物体的三角面片数目.与 3D 图形绘制追求真实感与实时性不同,在3D物体识别与检索中,只要得到最小数目的三角面片,并且这些三角面片完全表征了3D物体,就可以从最小的三角面片中提取 3D 物体特征,以减少不必要的计算量.图 1 给出了用 LoD 进行三角面片约减的示例.由图 1 可以看到,通过三角面片的删除、边压缩和面片收缩等方法,将一个 3D 物体的三角面片减少,可以达到细节层次目的.

在众多文献所提到的 LoD 算法中,一种比较常用的算法就是基于四叉树(quadtrees)的 LoD 算法,这种算法的基本思想极为简单,即利用一个距离的阈值来控制四叉树递归运算的深度,当这个阈值比较大时,得到较少的三角面片数量,反之则得到较多的三角面片.本文的实验中就是采用了这种方法进行 LoD 的计算.

图 2~图 4 分别给出了同一地形在不同个数三角面片下的显示效果.可以看出,三角面片数量越多,3D 地形

的细节越丰富(如图 2 所示);当三角面片约减到一定程度时,3D 地形的细节基本丢失(如图 4 所示);但是,当三角面片数量在一定范围内时,3D 地形的细节基本被保留下来(如图 3 所示).

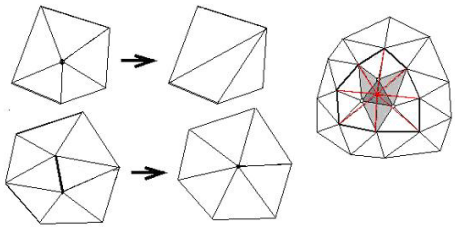

Fig.1 LoD principle

图 1 LoD 工作原理

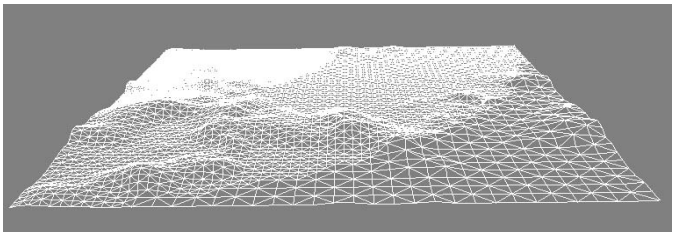

Fig.2 3D Terrain after LoD (About 10 000 triangles)

图 2 LoD 后的 3D 地形(约 10 000 个三角面片)

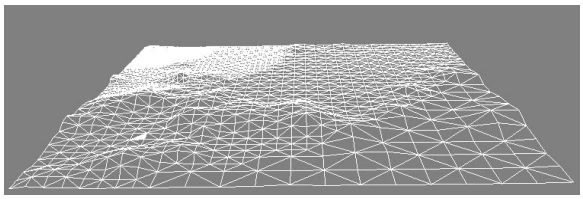

Fig.3 3D Terrain after LoD (About 4 000 triangles)

图 3 LoD 后的 3D 地形(约 4 000 个三角面片)

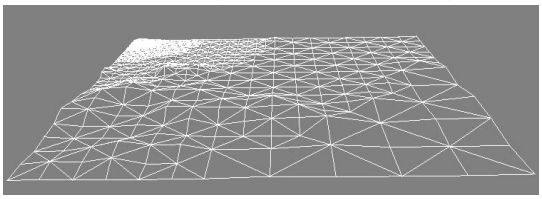

Fig.4 3D Terrain after LoD (About 1 000 triangles)

图 4 LoD 后的 3D 地形(约 1 000 个三角面片)

2 三维地形特征提取

得到经过 LoD 模型处理的 3D 物体之后,就需要从中提取特征,进行 3D 物体识别与检索.

由于本文是介绍相似地形识别与检索的,所以下面对 3D 地形特征提取作详细解释.与图像识别中特征提取不同,三维地形的特征选取是一个相当复杂的工作,因为所抽取的特征要表征一个 3D 地形在空间中的最基本和最显著的特点(如形状、大小和位置等).

在 3D 地形的识别中,比如对于丘陵地区和山地这两种不同地形的识别与检索,本文提取了如下特征:

(1) 地表中每个三角面片的法向量与空间 z 轴的夹角.这个特征可以反映出 3D 地形起伏不平的程度:对于丘陵地表而言,所有三角面片的法向量分布会集中在几个区间之内;而对于起伏剧烈的山地而言,所有三角面片的法向量分布要分散很多.

(2) 每个三角面片与坐标原点构成的四面体体积.在总体高度相对较低,地表起伏相对不大的丘陵地形中,这个体积值就是一个相对较小的数值,而且分布区间比较小;而在高度相对较高,地表起伏剧烈的山地中,这个值相对较大,而且分布的区间也要大得多.对地形样本中的每一个三角面片,采用如下公式计算它与坐标原点构成的四面体体积:

$$V=\left|\frac{1}{6}(-x_3y_2z_1+x_2y_3z_1+x_3y_1z_2-x_1y_3z_2-x_2y_1z_3+x_1y_2z_3)\right|,$$

其中 $x_i,y_i,z_i(i=1,2,3)$ 分别表示三角面片 3 个顶点的坐标.

(3) 每个三角面片的矩值.可以提取每个三角面片的零阶及高阶矩作为特征进行 3D 识别与检索.以下是计算三角面片的矩时所使用的公式:

$$M_{000}=\frac{1}{6}(-x_3y_2z_1+x_2y_3z_1+x_3y_1z_2-x_1y_3z_2-x_2y_1z_3+x_1y_2z_3),$$

$$M_{100}=\frac{1}{4}(x_1+x_2+x_3)M_{000},$$

$$M_{010}=\frac{1}{4}(y_1+y_2+y_3)M_{000},$$

$$M_{001}=\frac{1}{4}(z_1+z_2+z_3)M_{000},$$

$$M_{200}=\frac{1}{10}(x_1^2+x_2^2+x_3^2+x_1x_2+x_2x_3+x_1x_3)M_{000},$$

$$M_{020}=\frac{1}{10}(y_1^2+y_2^2+y_3^2+y_1y_2+y_2y_3+y_1y_3)M_{000},$$

$$M_{002}=\frac{1}{10}(z_1^2+z_2^2+z_3^2+z_1z_2+z_2z_3+z_1z_3)M_{000},$$

$$M_{300}=\frac{1}{20}\left[x_1^3+x_2^3+x_3^3+x_1^2(x_2+x_3)+x_2^2(x_1+x_3)+x_3^2(x_1+x_2)+x_1x_2x_3\right]M_{000},$$

$$M_{300}=\frac{1}{20}\left[y_1^3+y_2^3+y_3^3+y_1^2(y_2+y_3)+y_2^2(y_1+y_3)+y_3^2(y_1+y_2)+y_1y_2y_3\right]M_{000},$$

$$M_{300}=\frac{1}{20}\left[z_1^3+z_2^3+z_3^3+z_1^2(z_2+z_3)+z_2^2(z_1+z_3)+z_3^2(z_1+z_2)+z_1z_2z_3\right]M_{000}.$$

其中, $x_i,y_i,z_i(i=1,2,3)$ 分别表示三角面片 3 个定点的坐标.

对于三维物体的识别而言,如何在数学层面上归纳出一个空间三维物体的几何特征是一个重要但又很困难的问题.在现有的研究工作中,通常提取空间物体的矩(moment)来描述该物体在三维空间中的几何特性[9],因为矩可以表征诸如物体的对称轴等几何特点.空间物体的不同形式的矩可以通过基本矩理论[10]来求得.

这样,从每个三角面片可以提取 12 个 3D 特征.也就是说,如果一个经过约减后的 3D 地形包含 4 000 个三角面片,从这个 3D 地形中所提取的抽取特征就是 4000×12.如果直接将 4000×12 用于后续计算,将使计算变得异常复杂.

3 基于最小生成树的 3D 特征约减

从上面可以看到,即使 3D 地形经过了三角面片约减,其所提取的 3D 特征也非常多,如果直接使用这些特征进行识别,就无法保证识别及检索过程的实时性,这也是不必要的,因为这些特征中存在冗余.

为了减少 3D 特征之间的冗余,在进行识别与检索之前,可以对上面得到的原始特征进行约减.本文对从每个 3D 样本所提取的特征集都建立一个最小生成树,然后在这个最小生成树的基础上采用聚类算法来消除特征的冗余性,达到特征约减的目的.

3.1 三维地形样本的最小生成树表示

当使用最小生成树来表示一个 3D 地形样本时,设 $D=\{d_i\}$ 是从一个 3D 地形样本得到的特征数据集, $d_i=[f_i^1,f_i^2,...,f_i^j]$,它代表了样本中第 i 个三角面片的特征向量,其中分量 f_i^j 表示第 i 个三角面片的第 j 个特征值.

在 D 的基础上,如下定义一个有权重的无向图 $G(D)=(V,E)$,其中 $G(D)$的顶点集为 $V=\{d_i|d_i\in D\}$,边的集合为 $E=\{(d_i,d_j)|d_i,d_j\in D,i\neq j\}$.可以看出,$G(D)$是一个全连通的无向图,图中的每条边$(u,v)\in E$ 都具有一个权重值,它代表了两个顶点(三角面片)u 和 v 之间的距离$\rho(u,v)$.$\rho(u,v)$可以采用很多方法去定义,本文使用的是两个三角面片特征向量之间的欧拉距离.

每个无向图均表示了一个 3D 物体中所有三角面片的特征向量之间的空间拓扑关系,为了达到消除 3D 特征冗余的目的,需要消除无向图中的回路,也就是要得到无向图的最小生成树.无向图 $G(D)$的最小生成树 T 定义如下:(1) T 包括了图 $G(D)$中的所有顶点;(2) T 中没有回路;(3) T 中所有边的距离之和最小.图 5 给出了一个二维平面上数据集的最小生成树示例,图 5 右边是 12 个点所构成的最小生成树,左边是根据最小生成树将这 12 点归属到 3 个自然的群集,然后使用每个属类数据的质心表示这个属类中的所有数据点,达到数据约减的目的.从另外角度来看,得到最小生成树的过程,也是对原数据点进行聚类的过程.

构造最小生成树的方法很多,在对任意 3D 物体进行细节层次处理得到了三角面片之后,本文采用 Prim 算法计算这些三角面片的最小生成树,基于最小生成树进行特征约减.

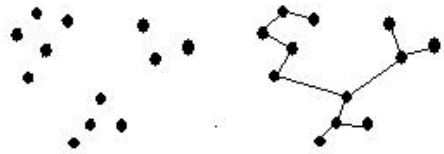

Fig.5 An MST representation of a set of data points

图 5 数据集的 MST 表示

3.2 基于最小生成树的聚类算法

为了取得较好的聚类效果,不同的聚类问题需要不同的目标函数.本节介绍一种基于最小生成树的聚类算法和其相应的目标函数,用来对每一个 3D 样本的三角面片进行聚类分析,达到特征约减目的.

假设 D 是从一个 3D 样本中得到的特征数据集,ρ表示两两数据之间的距离.基于最小生成树的聚类算法目标是在得到 D 的最小生成树后,依据最小生成树将 D 中数据聚类成 K 个不同子类 $C_i(1\le i\le K)$.

C_i 成为子类的条件可以如下定义[11]:对于 C_i 的任意一个划分 $C_i=C_{i1}\cup C_{i2}$,如果 d 是满足条件 $d\in D-C_{i1}$ 且距离 C_{i1} 最近的点,那么必然有 $d\in C_{i2}$.也可以用如下的公式来表示:

$$\arg\min_{d\in D-C_1}\{\min\{\rho(d,c)\mid c\in C_1\}\}\in C_2.$$

也就是说,类中数据之间的距离总会比类之间数据的距离要小.

如果要把 D 中数据聚类成如上定义的 K 个子类,基于最小生成树进行聚类的目标函数就是要把得到的最小生成树划分为 K 个子树,使得 $K(K>0)$个子树的所有边之和为最小.不难证明,只要去掉最小生成树中最长的 $K-1$ 条边,就得到了 K 个子树,并且满足所提出的目标函数.这样,就把 D 中特征数据聚类成了 K 个子集,每一个子集在图中表现为一棵子树.

3.3 特征约减

在对每一个 3D 样本中的三角面片进行基于最小扩展树的聚类分析之后,接下来就可以进行特征的约减.由于在构造图 $G(D)$时使用的是每个三角面片的特征向量间的欧拉距离,这就说明在聚类分析中处于同一个子树中的三角面片具有相近的特征向量值.所以只要计算出每一个子树所具有的特征向量的平均值(质心),就可以用质心来代表这一类三角面片的特征.假设经过细节层次处理后的3D样本中有N个三角面片,在经过特征提取后得到一个大小为 $N\times l$(l 为提取出的特征个数)的特征矩阵,由于 N 非常大,所以这是一个非常大的矩阵.然而在使用上述的聚类分析之后,这个特征矩阵的大小变为 $K\times l$(K 为聚类时的子树个数),由于一般 $K<<N$,所以特征得到了很大的约减.

本文之所以采取基于最小生成树的聚类算法,主要原因在于:(1) 最小生成树使得一个多维数据的聚类问题转化为图论中子树的划分问题,而这个过程不会丢失关键信息;(2) 树这种数据结构使得聚类算法的实现很简单,对类的形式化定义很清楚;(3) 传统的聚类方法对群集的边界要求很高,即群集之间的边界不能太复杂,但是基于最小生成树的聚类方法对这一点并不敏感.

4 基于支持向量机的 3D 识别

在进行相似 3D 物体识别时,需要为每类 3D 物体构造一个分类模板,然后使用这个模板实现相似 3D 物体识别.

由于支持向量机具有小样本识别正确率高和良好的学习推广能力,这里使用支持向量机实现相似 3D 物体的识别与分类.支持向量机起源于统计学习理论,它研究如何构造学习机,实现模式分类问题[12].支持向量机使用结构风险最小化(structural risk minimization,简称 SRM 准则)原理构造决策超平面,使每一类数据之间的分类间隔(margin)最大.SRM 准则认为:学习机对未知数据分类所产生的实际风险是由两部分组成的,如果 $0\le\eta\le1$,则

满足如下关系:

$$R \le R_{emp} + \sqrt{\frac{h(\log(2n/h)+1)-\log(\eta/4)}{n}},$$

其中,R 是实际风险,不等式的右边叫做风险边界,R_{emp} 称为经验风险,$\sqrt{\frac{h(\log(2n/h)+1)-\log(\eta/4)}{n}}$ 叫做“VC 置信值”,n 是训练样本个数,h 是学习机的 VC 维(h 反映了学习机的复杂程度).

支持向量机的思想就是在样本数目适宜的前提下,选取比较好的 VC 维 h,使经验风险 R_{emp} 和置信值达到一个折衷,最终使实际风险 R 变小.

对于两类问题分类,存在线性可分和线性不可分两种支持向量机,但是在实际中,为了将两类模式尽可能分类开来,一般要构造非线性可分的支持向量机.Cover 定理指出[13]:一个复杂的模式识别分类问题,在高维空间比低维空间更容易线性可分.支持向量机就是首先通过核函数把训练样本中的低维数据映射到高维特征空间,然后在高维特征空间构造一个最佳分类平面.由于构造的核函数满足 Mercer 条件,所以在训练中只需考虑核函数 K,而不必明确知道低维向高维的映射函数 Φ[12].

SVM 中研究最多的核函数主要有 3 类:多项式、径向基函数(RBF)和多层 Sigmoidal 神经网络.本文使用的是 RBF 核函数,形式为 $K(x,y)=\exp\left\{-\frac{|x-y|^2}{\sigma^2}\right\}$.

在训练支持向量机模板时,收集某类 3D 物体的正样本和负样本,通过细节层次约减三角面片后,提取特征,进行基于最小生成树的聚类,将聚类质心进行训练,得到识别这类 3D 物体的分类器模板.当用户提交一个要识别的 3D 物体时,同样经过细节处理和特征抽取,由训练好的支持向量机判断用户所提交的 3D 样本属于哪一类,从而完成 3D 样本的识别.

由于本文处理的是两类地形识别,所以在 SVM 训练时,将一类地形样本标注为+1,另外一类地形样本则标注为–1.对于每个测试样本,可以使用训练好的 SVM,通过判断其输出是正还是负,来达到 3D 地形识别的目的.

5 实验数据与对比

5.1 实验结果

本文对丘陵和山地两类不同的 3D 地形进行了识别与检索.其中,丘陵的地形相对平缓(如图 6 所示),而山地的地形相对起伏较大(如图 7 所示).

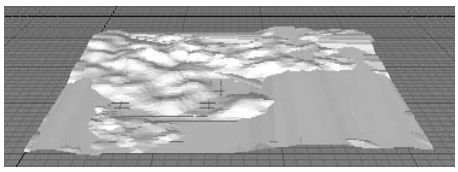

Fig.6 3D Terrain (Hill)

图 6 3D 地形(丘陵)

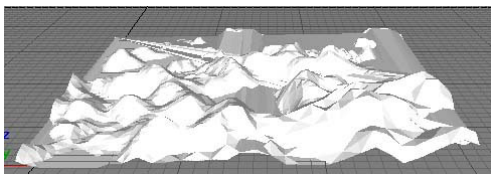

Fig.7 3D Terrain (Mountain)

图 7 3D 地形(山地)

这个实验中总共生成了 60 个 3D 地形样本,丘陵和山地各 30 个(所生成的样本均包含 10 000 个三角面片).从生成的样本中随机抽取 50 个样本(丘陵和山地各 25 个)作为训练样本,对其进行细节层次处理、特征提取和特征约减,将聚类约减后的特征作为 SVM 的输入进行训练.剩下的 10 个同样处理,作为测试样本.

本文的实验是在大约 4 000 个三角面片的细节层次下进行的,得到的识别正确率为 70%.同时,为了考察不同细节层次和不同特征约减策略对 3D 地形识别所产生的效果不同,我们还给出了在不同细节层次下,使用原始特征和主成分分析(principle component analysis,简称 PCA)特征约减方法[14]的实验结果.

在实验中,对每一个地形样本采用统一的 LoD 算法,使得在识别时具有可比性.在提取特征时,坐标原点可以选取在样本的任意一角,对识别和检索结果没有影响.在 VC 和 Matlab 环境下进行了实验,计算效率见表 1.

Table 1 Efficiency of algorithm

表 1 算法效率

Step	Time
LoD	<1s (VC)
Feature extraction	<3.8s (Matlab)
MST reduction	5min (Matlab)
SVM training	<0.2s (Matlab)
SVM recognition	<0.1s (Matlab)

可以看到,除了用 MST 对特征进行约减以外,其他各个步骤所需时间都很少.在特征约减这一步上所花的时间较多,有以下两个原因:(1) 样本本身的数据量就很大(超过 4 000 个三角面片);(2) 本文采用的是较为简单的 MST 生成算法,如果对 MST 算法进行优化,可以提高计算性能.

表 2 给出了使用训练好的支持向量机,在不同的细节层次下使用不同特征约减方法所取得的识别正确率(括号中的数字表示进行 LoD 处理后 3D 地形样本所包含的三角面片数.

Table 2 Precison of recognition under different LoD and feature reduction

表 2 不同细节层次和特征约减下识别正确率

	First level LoD (10 000) (%)	Second level LoD (4 000) (%)	Third level LoD (1 000) (%)
Original features	43.33	53.33	60
PCA	62.33	66.67	61
MST	64.33	70	63.33

从表 2 的实验结果可以看出,在对 3D 地形进行识别时,采取细节层次处理是一个较好的策略.3D 地形样本经过 LoD 处理后,非但没有对识别结果产生很坏的影响,甚至还提高了识别正确率.这一点在表 1 中得到了很好的体现,二级 LoD 后的识别正确率明显高于进行一级 LoD 处理后的正确率.

此外,采用一种好的特征约减策略对识别结果影响也很大.比如,使用最小生成树方法和 PCA 方法对原始特征进行约减后,就使之比原始特征取得了较好的识别正确率.并且相对于 PCA 方法而言,MST 特征约减策略取得了更好的识别效果.

最后也是最关键的一点,在如此小样本的实验中,最终能取得较高识别正确率,主要归功于一个强大的分类器的构建,这也说明支持向量机能够在样本不是很多的情况下,实现较高识别率.

实验中出现误判的原因有两个方面,首先是因为样本的数量不是很多(只用了 50 个训练样本);其二,样本的产生和训练样本的选取都是随机的.

在基于内容的多媒体检索中,一般采用基于例子的多媒体检索方式[4,5],即用户需要提交一个检索例子(如图像、视频例子和音频例子),然后从被检索数据库找到与用户提交例子相似的多媒体数据.

本文在相似 3D 检索中也采用这种方式,用户提交一个 3D 地形例子的检索请求后,提取其 3D 约减特征,接着计算这个 3D 地形例子与被检索 3D 地形数据库中所有样本之间的欧拉距离,返回前 10 个与查询 3D 地形例子距离最近的 3D 地形,作为检索结果.

按照查全率和查准率标准,表 3 和表 4 中分别给出了使用 PCA 特征约减方法和 MST 方法的检索结果.从表 3 可以看出,在第 2 级 LoD 层次上检索取得了相对较好的结果,说明先对 3D 地形样本进行 LoD 处理后再提取相关特征是可行的方法.从表 3 和表 4 的比较来看,使用最小扩展树方法也取得了优于 PCA 方法的检索正确率.但是也可以看到,60%的检索正确率并不是很理想,这是因为采用 MST 聚类的特征约减算法在一定程度上使得样本的特征表达能力受到了削弱(对于某一个样本的特征向量来说,它实际来自于 MST 的每一棵子树中所有节点的特征向量的平均值).

Table 3 Result of retrieval under PCA

表 3 PCA 约减下的 3D 检索结果

	First level LoD (10 000) (%)	Second level LoD (4 000) (%)	Third level LoD (1 000) (%)
Precision	57	58	55.50
Recall	40.14	41.57	39.50

Table 4 Result of retrieval under MST

表 4 MST 方法下的 3D 检索结果

	Second level LoD (4 000) (%)
Precision	60
Recall	42.27

5.2 与其他方法的比较

在 Cha Zhang 和 Tsuhan Chen 的一篇文章[15]中,曾经提出了一种基于手工标注与主动学习机制的三维物体识别与检索算法.在这个算法中,首先系统初始化时通过手工标注大量的训练样本,然后让系统在接下来的运行中根据用户的反馈来对所有的样本重新进行自动标注.可以看到,这种识别与检索机制在对三维飞机模型的实验中取得了较好的识别正确率.但是这种较高的识别正确率是建立在对大量训练样本(上百个)的手工标注基础上的,而且在系统根据用户的反馈做调整时所进行的计算量非常之大.

本文则采用了一种完全不同的识别与检索机制(三维样本→LoD 处理→特征提取→使用 MST 特征约减→使用 SVM 识别),与前面所讲的算法相比,本文实验中所采用的算法有以下特点:首先,对训练样本的数量要求不高(正、负样本各 25 个);其次,计算主要集中在 SVM 的训练上,在使用训练好的 SVM 进行地形样本识别时用得很少.

6 结论和今后工作

本文对 3D 地形使用了细节层次处理和基于最小生成树的特征约减技术,通过支持向量机进行识别与检索,取得了较好的效果.原始的 3D 地形样本所包含的三角面片数量较多,从而提取出的特征之间也存在较多冗余,损害了识别与检索结果;而进行多次 LoD 后的 3D 地形虽然包含三角面片数量较少,但是已经丢失了较多的信息,同样不利于识别与检索.所以,对原始的 3D 地形进行合适层次的 LoD 处理是获得较好结果的关键.

今后的工作包括如下两个方面:(1) 继续研究能够反映 3D 地形性质的良好特征;(2) 本文的实验结果是在小测试样本下取得的,今后应该将这个算法使用到更多的 3D 数据中,以增强算法的适应性.

致谢 在此,我们向对本文的工作给予支持和建议的同行,尤其是浙江大学新利智能多媒体实验室、浙大-微软视觉感知联合实验室的老师和同学表示感谢!

References:

[1] Zhan GH, Zhuang YT, Wu Y. Indexing video using global and local features. Journal of Computer Aided Design & Computer Graphics, 2000,12(12):911～916 (in Chinese with English abstract).

[2] He QF, Li GJ. An approach to combining dominant color of partition with relevance feedback in image retrieval. Journal of Computer Aided Design & Computer Graphics, 2001,13(10):912～917 (in Chinese with English abstract).

[3] Foote J. An overview of audio information retrieval. Multimedia Systems, 1999,7(1):2~11.

[4] Wu F, Zhuang YT, Zhang Y, Pan YH. Hidden markovia model based audio semantic retrieval. Journal of Intelligence & Pattern Recognition, 2001,14(1):104～108 (in Chinese with English abstract).

[5] Wu Y, Zhuang YT, Pan YH. Video similarity measurement. Journal of Computer Aided Design & Computer Graphics, 2001, 13(3):384～388 (in Chinese with English abstract).

[6] Levoy M. The digital michelangelo project: creating a 3D archive of his sculptures using laser scanning. In: Proceedings of the 2nd International Conference on 3-D Digital Imaging and Modeling. IEEE/IEE Electronic Library, 1999. 2~13.

[7] Seo H, Thalmann AF. LOD management on animating face models. In: Thalmann D, ed. Proceedings of the IEEE Virtual Reality 2000. New Brunswick: IEEE Computer Society, 2000. 161~168.

[8] Lindstrom P, Koller D, Ribarsky W, Hodges LF, Faust N, Turner G. Real-Time, continuous level of detail rendering of height fields. In: Proceedings of the SIGGRAPH. New Orleans: Addison Wesley Professional, 1996. 109~118. http://www.gvu.gatech.edu/people/peter.lindstrom/papers/siggraph96/siggraph96.pdf.

[9] Reeves AP, Prokop RJ, Andrews SE, Kuhl FP. Three dimensional shape analysis using moments and Fourier descriptors. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1988,10(6):937~943.

[10] Reiss TH. The revised fundamental theorem of moment in-variants. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1991,13(8):830～834.

[11] Xu Y, Olman V, Xu D. Minimum spanning trees for gene expression data clustering. Genome Informatics, 2001,11:24~33.

[12] Vapnik VN. The Nature of Statistical Learning Theory. Springer, 1995.

[13] Cover TM. Geometrical and statistical properties of systems of linear inequalities with applications to pattern recognition. IEEE Transactions on Electronic Computers, 1965,EC-14(3):326~334.

[14] Yambor WS. Analysis of PCA-based and fisher discriminant-based image recognition algorithms. Technical Report, CS-00-103, Computer Science, Colorado State University, 2001.

[15] Zhang C, Chen T. Active learning for information retrieval: Using 3D models as an example. Technical Report, AMP01-04, Pittsburgh: Advanced Multimedia Processing Laboratory, Carnegie Mellon University, 2001.

附中文参考文献:

[1] 詹国华,庄越挺,吴翌.基于全局与局部特征的视频索引模型.计算机辅助设计与图形学学报,2000,12(12):911~916.

[2] 何清法,李国杰.综合分块主色和相关反馈技术的图像检索方法.计算机辅助设计与图形学学报,2001,13(10):912~917.

[4] 吴飞,庄越挺,张引,潘云鹤.基于隐马尔可夫链的音频语义检索.模式识别与人工智能,2001,14(1):104~108.

[5] 吴翌,庄越挺,潘云鹤.视频相似度的衡量.计算机辅助设计与图形学学报,2001,13(3):384~388.

中国计算机学会 2003 年部分活动计划

11 月 17~19 日 西安

The IEEE Twelfth Asian Test Symposium (ATS'03)

主办：IEEE Computer Society, Test Technology Technical Council

承办：容错计算专委

联系人： 李晓维 电话：010-62654541 Email: ats03@ict.ac.cn Http://ats03.ict.ac.cn

11 月 20~21 日 西安

4th IEEE Workshop on RTL and High Level Testing (WRTLT'03)

主办：IEEE Computer Society Test Technology Technical Council

承办：容错计算专委

联系人：李华伟 电话：010-62565533-5673 Email: wrtlt03@ict.ac.cn Http://wrtlt03.ict.ac.cn

11 月 21 日 北京

学术报告会：语义 Web

主办：中国计算机学会

联系：学会办公室 010-62527486 E-mail: yocsef@ict.ac.cn

12 月 20 日 北京

专题论坛：电子政务误区何在

主办：中国计算机学会

联系：学会办公室 010-62527486 E-mail: yocsef@ict.ac.cn