

文章编号:1673-9469(2014)01-0087-03

doi:10.3969/j.issn.1673-9469.2014.01.023

三维蚁堆算法在视频关键帧提取中的应用

吴开兴, 沈志佳

(河北工程大学 信息与电气工程学院, 河北 邯郸 056038)

摘要:为了有效、且自适应的提取出视频中的关键帧,提出了一种改进的三维蚁堆新算法。该算法首先提取每一帧中 $H-S-V$ 颜色空间的三维特征向量,并将其表示为 $H-S-V$ 三维欧式空间中的点,之后通过改进的三维蚁堆算法,自适应的聚类,从而提取出视频中的关键帧。通过MATLAB仿真并与传统算法对比,结果表明:相对于传统算法,该算法的查全率和查准率都有了一定程度的提高。

关键词:视频检索; 关键帧提取; 蚁堆算法

中图分类号:TP301.6

文献标识码:A

The algorithm of key frame extraction based on improved three - dimensional ant heap

WU Kai - xing, SHEN Zhi - jia

(College of Information and Electrical Engineering, Hebei University of Engineering, Hebei Handan 056038, China)

Abstract: The key frame extraction is one of the key technologies in Content - based video retrieval (CBVR). In order to extract self - adaptively and effectively the key frames in the video, this text proposes an algorithm of key frame extraction basing on improved three - dimensional ant heap. First, this algorithm extracts three - dimensional feature vectors from color space of $H-S-V$ for every frame and shows them as the points of three - dimensional European Space. Then, we divide all the frames into clusters through the algorithm of improved three - dimensional ant heap to extract the key frames in video. Through simulating the new arithmetic through the MATLAB and comparing it with the traditional arithmetic, results show that the recall ratio and precision ratio are improved in some extent.

Key words: video retrieval; key frame extraction; algorithm of ant heap

随着计算机的普及和网络技术的不断发展,越来越多的视频数据不断涌现。如何对这些数据进行高效的存储、浏览和检索便显得十分重要^[1-2]。在基于内容的视频检索中,关键帧提取是最关键的技术之一。目前关键帧提取技术方法很多^[3-10],但多数存在提取效果不理想或算法复杂等问题。本文主要是在传统蚁堆算法的基础上,提出了一种改进的三维蚁堆算法,将传统算法的蚂蚁运动路径扩展到三维平面。则三维平面与三维视频帧特征向量的帧数一致,从而提高了聚类的查全率和查准率。

1 传统的蚁堆算法

蚂蚁是一种社会性昆虫,具有自组织、自适应、分布式、相互通信、相互合作等特点。模拟蚂蚁的这些特性,我们可以解决许多很复杂的问题。Deneubourg等人从蚂蚁堆积尸体中受到启发,提出了一种基本模型BM(Basic Model)。Lumer和Faieta在BM的基础上进行了改进,增加了数据对象的相似性表达式,提出了用于数据聚类的LF算法。LF算法的基本思想如下:将 N 个数据对象随机的投入到 $M \times M$ 的网格中,将 A 只蚂蚁也随机的投入到该网格中,蚂蚁沿着网格随机移动,蚂蚁

收稿日期:2013-09-23

作者简介:吴开兴(1962-),男,汉族,陕西渭南人,硕士,教授,从事多媒体技术、管理信息系统方面的研究。

遇到数据对象后,根据该数据对象与蚂蚁可视范围内数据对象的相似度,以一定的概率负载数据对象,负载数据对象的蚂蚁沿着网格随机移动,当遇到空网格后,根据负载与可视范围内数据对象的相似度,以一定的概率卸载负载,然后未负载的蚂蚁继续随机移动。

设蚂蚁的可视范围为 r , 则相似度计算函数为如下公式^[8]:

$$f(o_i) = \sum_{o_j \in \text{neigh}(r)} [1 - d(o_i, o_j)/a] \quad (1)$$

$d(o_i, o_j)$ 为数据对象 o_i 与 o_j 的距离,通常用欧式距离, a 为相似性参数, $\text{neigh}(r)$ 表示以 r 为半径的圆形区域。

P_p 和 P_d 分别为蚂蚁拾起对象和放下对象的概率,计算如公式(2)(3)^[8]所示。

$$P_p = \begin{cases} 1 & f(o_i) \leq 0 \\ 1 - k \times f(o_i) & 0 < f(o_i) \leq 1/k \\ 0 & f(o_i) > 1/k \end{cases} \quad (2)$$

$$P_d = \begin{cases} 1 & f(o_i) \geq 1/k \\ k \times f(o_i) & 0 < f(o_i) < 1/k \\ 0 & f(o_i) \leq 0 \end{cases} \quad (3)$$

2 改进的蚁堆算法在关键帧提取中的应用

传统的蚁堆算法,是基于二维平面的,数据对象是以随机的方式投入到二维网格上。本文在传统蚁堆算法的基础上,将其扩展到三维平面,数据对象以其在 $H-S-V$ 空间中的三维特征向量为坐标,分布在三维立体网格上。这样,数据对象之间的欧式距离,便是其在三维空间中的直线距离,距离近的数据对象聚集在一起,因此,数据对象已在一定程度完成了初始聚类。

首先进行视频帧特征向量的提取。视频帧的每个像素均可以在 $H-S-V$ 空间中表现出来,其中 H 的取值范围为 $[0, 360]$, S 和 V 的取值范围均为 $[0, 1]$,为了使视频帧的距离计算更加有效,将 S 和 V 的取值乘以 360,使 S 和 V 的取值范围也扩展到 $[0, 360]$,这样就得到了关于每个象素的三维特征向量。计算第 n 帧整幅图像的平均特征向量 X_n , 两帧之间的欧式距离 $d(o_i, o_j)$ 便可由公式(4)表示:

$$d(o_i, o_j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + (x_{i3} - x_{j3})^2} \quad (4)$$

其中 x_{nk} 为特征向量 X_n 的 k 维分量。

将每一个视频帧映射到三维的欧式空间,这样,两帧之间的欧式距离便可由三维空间中两点

的直线距离表示。

一段镜头便可由三维空间中一系列的散点表示,如下图1。

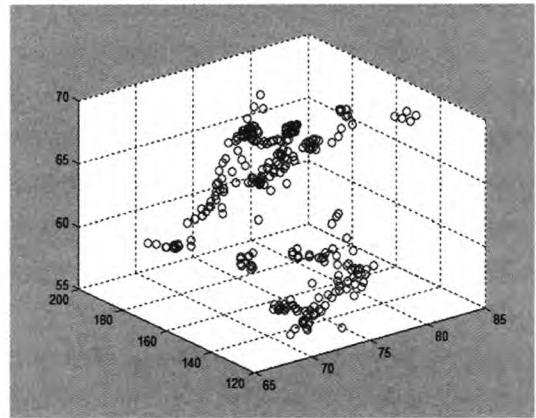


图1 视频帧向量三维空间散点图

Fig.1 Scatter diagram of frames in three-dimensional space

之后,我们通过对传统蚁堆算法进行改进,使其适应于 $H-S-V$ 三维空间。具体的改进方面如下,

(1) 将数据对象按照其在 $H-S-V$ 中的坐标放在三维数据空间中,由于数据对象在三维空间中的实际距离,即是其欧式距离,从某种程度说,数据对象,已经完成了初始聚类。

(2) 由于数据对象在三维空间中已经有一定的聚类性质,初始放置蚂蚁时,将蚂蚁均匀的分布在三维空间中,并且使每一只蚂蚁,在一个已有数据对象的网格内,避免蚂蚁对于负载的饥饿。

(3) 数据对象周围的数据对象有很大几率属于同一聚类,因此,蚂蚁放下负载后,自动寻找与此数据对象最近的负载点,减少蚂蚁的无意义运动。

(4) 蚂蚁拾起负载后,有很大概率,在其负载周围再次放下负载,因此,蚂蚁以负载初始位置为圆心,在半径 r_1 为 1 的球面内运动,如果在此球面内不能放下负载,则半径 $r_1 + 1$,直到蚂蚁放下负载为止。此改进点的目的是使蚂蚁在较短的步数内,在正确位置放下负载,这也是因为数据对象之间的欧式距离就是其实际距离,欧式距离近的数据对象已经聚集在一起的缘故。

(5) 如果蚂蚁欲放下负载的位置已有其它的数据对象,则在此位置的邻域内求其放下对象的概率。如果所有的领域都不能放下,则蚂蚁回到原先的运动路径,继续在以 r_1 为半径的球面内移动。

(6) 定义数据对象的优先级,如果,蚂蚁经过

较少的步骤,即放下数据对象,说明数据对象已经在其正确位置,则此数据对象需要再次移动的概率很低,我们就赋予此数据对象较低的优先级,否则赋予较高的优先级,当条件一样时,即步骤 3 中存在多个最近负载点时,蚂蚁优先负载优先级较高的负载点,如果优先级一致则随机选取。

(7) 参数 a 的确定非常依赖使用者的经验,使算法缺少普适性。算法中 a 初始值设为 0.1,如果迭代时,蚂蚁放下对象失败次数与迭代次数的比值大于 0.99,则 a 增加 0.01,否则减少 0.01。

(8) 当蚁堆算法迭代到一定次数 t 后,数据对象已完成了进一步的聚类,按照文献[11]所述方法,再一次提取每一帧的 $H-S-V$ 颜色直方图。之后以每一帧中 $H-S-V$ 颜色直方图的 72 维特征向量为标准,以蚁堆算法提取的聚类中心和聚类个数作为参考,通过 K -均值聚类算法完成最终的聚类。则最终聚类中,距离每一个最终聚类中心距离最小的维特征向量所代表的视频帧,即为我们所需要的关键帧。最终聚类选取 72 维颜色直方图作为特征向量是因为,采用 72 维特征向量,比前述三维特征向量在计算欧式距离时更精确,最终聚类的结果也会更准确。

3 实验结果与讨论

利用三维蚁堆算法提取出的基于前图 1 所代表镜头的关键帧如下图 2 所示。

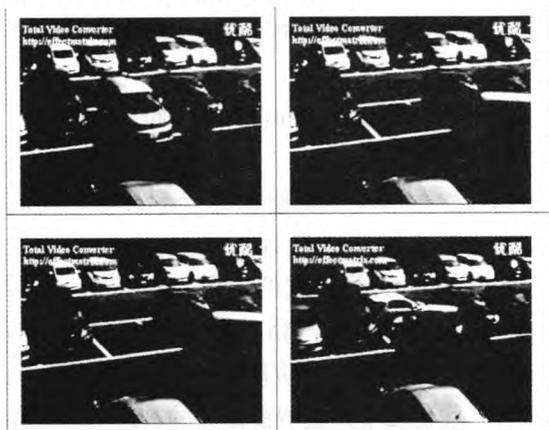


图2 视频关键帧序列

Fig.2 Sequence of key frames

可见提取出的关键帧分别代表了停车场某一个车位,汽车的动态变化,提取出的关键帧具有代表性。

利用本算法提取出的某个演讲镜头的关键帧如图 3 所示。



图3视频关键帧

Fig.3 Key frame of video

由于该演讲镜头内视频内容变化很少,所以仅提取出了唯一的帧,由此可见,本算法具有自适应性,能够根据镜头内容的复杂程度,自适应的提取出不同数量的关键帧。

为了验证算法的有效性,本文采用了 120 个不同的镜头来对算法进行分析。具体的实验结果如下图 4、图 5 所示。

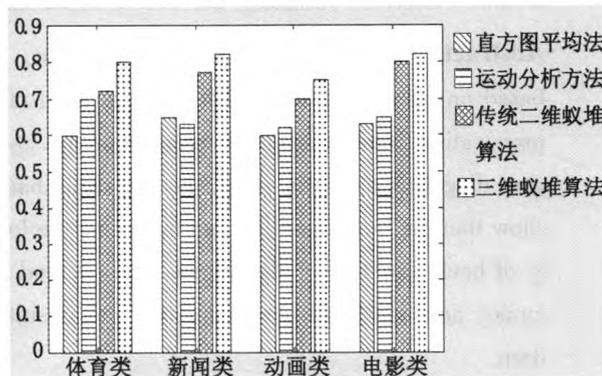


图4 各算法查准率对比图

Fig.4 Comparison of precision ratio

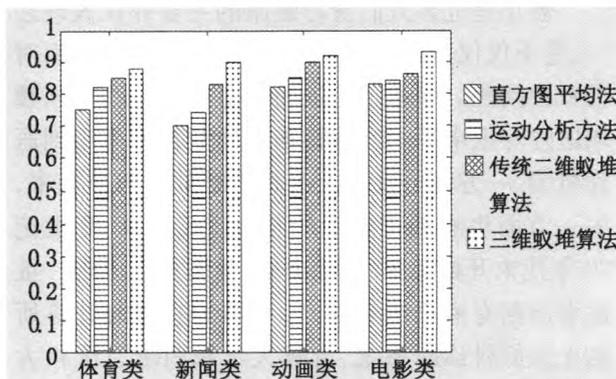


图5 各算法查全率对比图

Fig.5 Comparison of recall ratio

可见,无论是在体育类,新闻类,动画类,电影类的视频中,本文算法的查全率和查准率,都优于其它的传统算法。(下转第 101 页)

由表1可以得出,优化后的内门架在质量减少12%的情况下,最大等效应力降低了7%,最大等效位移降低了6.89%。在满足设计要求的前提下实现了优化目标,优化效果良好。

4 结论

(1)内门架的强度满足设计要求,但在滚轮轴独柱连接处以及腰板外侧与翼缘板外侧连接处存在应力集中现象,影响其疲劳寿命。在改进设计中,可通过材料力学性能的提示加以解决。

(2)内门架自下而上其位移变形量逐渐增大,在顶端达到最大值。虽然满足设计要求,但为了使其工作时更稳定、安全,可通过将横梁的布置高度提高或增加横梁数目等措施来增强内门架的刚度。

(3)优化后的内门架在质量减少12%的情况下,最大等效应力降低7%,最大等效位移降低了6.89%。

参考文献:

[1] 卞学良, 罗明军, 穆希辉, 等. 全向蓄电池侧面叉车外门架的有限元分析及优化[J]. 工程机械, 2006(7): 35-38.

(上接第89页)

4 结论

(1)本算法能够根据镜头内容的复杂程度,自适应地提取出不同数量,且具有代表性的关键帧。

(2)在体育类、新闻类、动画类、电影类的视频中,本文算法的查全率和查准率都优于传统的直方图法,运动分析法和二维蚁堆算法。

参考文献:

[1] TANIGUCHI Y. An intuitive and efficient access interface to real-time incoming video based on automatic indexing [C]// Proc of ACM Multimedia. San Francisco, 1995: 25-33.
 [2] 汪勤. 基于视频图像处理的无人值守变电站在线检测的研究[J]. 四川理工学院学报:自然科学版, 2013, 26(3): 91-94.
 [3] 瞿中, 高腾飞, 张庆庆. 一种改进的视频关键帧提取算法研究[J]. 计算机科学, 2012(8): 300-303.
 [4] 柴饶军, 马彩文. 图像序列中目标关键帧快速搜索算法[J]. 光子学报, 2004(10): 1233-1235.
 [5] WOLF W. Key frame selection by motion analysis[C]//

[2] 周京京, 穆希辉, 马振书, 等. 全向侧面叉车门架有限元分析及运动仿真[J]. 起重运输机械, 2009(2): 65-68.
 [3] 汪宇, 王东方. 基于 ANSYS Workbench 的立式加工中心床身有限元分析和优化设计[J]. 制造业自动化, 2009, 31(9): 129-131.
 [4] 慕灿. 复杂结构 UG NX 模型导入 ANSYS Workbench 的方法研究[J]. 四川理工学院学报:自然科学版, 2013, 26(4): 33-36.
 [5] 浦广益. ANSYS Workbench 12 基础教程与实例详解[M]. 北京:中国水利水电出版社, 2010.
 [6] 周政亮, 王贵飞, 丛明. 基于 ANSYS Workbench 的主轴箱有限元分析及优化设计[J]. 组合机床与自动化加工技术, 2012(3): 17-20.
 [7] 李兵, 何正嘉, 陈雪峰. ANSYS Workbench 设计、仿真与优化[M]. 北京:清华大学出版社, 2008.
 [8] 马静敏, 范云霄. 基于 ANSYS Workbench 的太阳能热水器支架优化设计[J]. 轻工机械, 2011, 29(5): 97-101.
 [9] 刘学文, 刘康, 刘光磊, 等. 龙门加工中心横梁部件静态特性分析[J]. 四川理工学院学报:自然科学版, 2013, 26(5): 44-47.
 [10] 凌桂龙, 丁金滨, 温正. ANSYS WorkBench 13.0 从入门到精通[M]. 北京:清华大学出版社, 2012.

(责任编辑 王利君)

Proc of IEEE International Conference on Acoustics Speech and Signal Processing. Atlanta, 1996: 1228-1231.

[6] 顾家玉, 覃团发, 陈慧婷. 一种基于 MPEG-7 颜色特征和块运动信息的关键帧提取方法[J]. 广西大学学报:自然科学版, 2010(2): 310-314.
 [7] YANG SHUPING, LIN XINGGANG. Key frame extraction using unsupervised clustering based on a statistical model[J]. Tsinghua Science and Technology, 2005(2): 169-173.
 [8] HUA MAN, JIANG PENG. A feature weighed clustering based key-frames Extraction method[C]. // Proceedings of the 2009 International Forum on Information Technology and Applications. Piscataway, 2009: 69-72.
 [9] 张建国, 刘海燕, 孙淑敏. 改进的蚁群算法与凝聚相结合的关键帧提取[J]. 计算机工程与应用, 2013(3): 222-226.
 [10] 贾存锋, 朱加雷, 焦向东, 等. GMAW 熔滴过渡高速摄像系统与熔滴边缘提取[J]. 河北科技大学学报, 2013, 34(4): 316-319.
 [11] 王娟, 孔兵, 贾巧丽, 等. 基于颜色特征的图像检索技术[J]. 计算机系统应用, 2011(7): 160-164.

(责任编辑 王利君)