

基于核聚类的 H.264 压缩域的运动对象分割方法

孔祥鹏,沙宁,邢薇

(哈尔滨工程大学 计算机科学与技术学院,黑龙江 哈尔滨 150001)

摘要:提出了一种基于核聚类的 H.264 压缩域的运动对象分割的方法。首先对运动矢量进行归一化处理,其次又引入了平均欧式距离的中值滤波的方法对运动矢量进行滤波去噪,再次利用减法聚类初始化聚类中心并且使用引入了核函数的模糊聚类进行聚类,最后将有效函数的判断结果作为分割条件,从而达到自适应分割出压缩域中的运动对象的目的。本方法对 Hall 视频序列进行了实验。实验证明,通过该方法可以较好地分割出视频序列中的运动对象。

关键词:核函数;模糊聚类;聚类有效性;减法聚类;运动矢量

中图分类号:TP301.4

文章标识码:A

文章编号:1673-629X(2013)08-0087-04

doi:10.3969/j.issn.1673-629X.2013.08.022

A Moving Object Extraction Method in H.264 Compressed Domain Based on Kernel Clustering

KONG Xiang-peng, SHA Ning, XING Wei

(College of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China)

Abstract: In this paper, a moving object extraction method in H.264 compressed domain based on kernel clustering is proposed. First, the motion vector is normalized. Second, the median filter using Average Euclidean Distance is used to process noise for motion vector. Third, the clustering center is initialized by the subtractive clustering and the use the fuzzy clustering to classify the data. Finally, the extracted condition is produced by the valid function and the moving object is got adaptively by the divided condition. The Hall experiment results show that the method can get the moving object very well.

Key words: kernel function; fuzzy clustering; cluster validity; subtractive clustering; motion vector

0 引言

当前数字科技迅猛发展,在现实生活中存储大量视频信息,如何有效地对其进行利用,是当下研究的热点。传统的提取视频中运动对象的分割方法是基于像素域的,但此种方法计算量大。基于压缩域的视频提取不需要对视频的数据流进行解码,计算量较少,更符合实时性的要求。H.264 视频压缩编码以其优良的压缩和编码效率得到了广泛的运用,基于 H.264 压缩域的运动矢量场的分割方法,作为一类分割方法已得到广泛的研究。

近几年提出的视频运动对象的分割方法有:Zeng^[1]等人提出了一种基于马尔可夫随机场的运动对象分割方法;Liu^[2]等人利用运动矢量的特性,使用统

计区域的方法合并相似的区域,分割出运动对象;Badu^[3]等人通过将运动矢量滤波和内插获取稠密的运动矢量场,然后使用仿射参数模型和 EM 算法获得运动对象的掩码,进而获取对象的轮廓。

文中提出了一种基于 H.264 压缩域的运动对象分割方法,使用核的 C-均值聚类算法,以及有效函数的判断指标,自适应分割出压缩域的视频运动对象,得到了较好的分割结果。

1 算法概述

文中从 H.264 压缩域中提取运动矢量场,首先利用阈值法去除运动矢量场中的野值点,其次进行中值滤波,平滑运动矢量场,最后通过使用核的 C-均值聚

收稿日期:2012-11-15

修回日期:2013-03-04

网络出版时间:2013-04-22

基金项目:黑龙江省自然科学基金项目(F201103)

作者简介:孔祥鹏(1985-),男,硕士,研究方向为数据库与知识工程;邢薇,教授,CCF 会员,研究方向为软件工程。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20130422.1729.068.html>

类算法和有效函数的指标判断相结合的方法产生自适应算法,通过自我学习,自适应地对运动矢量场进行分类,最后分割出压缩域中的运动对象。算法流程如图 1 所示。

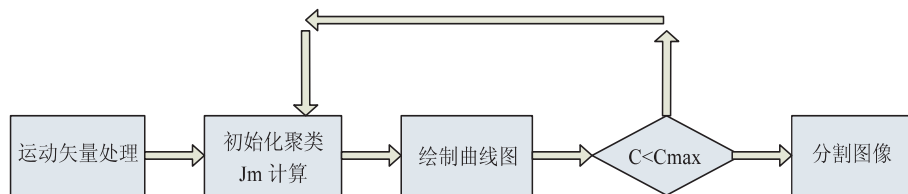


图 1 算法流程图

2 对运动矢量场的预处理

2.1 运动矢量场的归一化

H. 264 宏块(16×16)的划分方式分为 7 种分割方式,其中 4×4 为最小的宏块划分方式,因此,尺寸大于 4×4 宏块的宏块直接映射到自己所有的 4×4 的宏块,因此,得到了 4×4 宏块的均匀运动矢量宏块。

使用运动矢量之前需要对这些运动矢量进行归一化处理。

2.2 滤波去噪

由于随机噪声产生的野值点的运动矢量较小,因此可以通过设定一个阈值消除野值点。为了平滑噪声在运动对象的运动矢量场产生的噪声,采取中值滤波的方式平滑运动矢量场的运动矢量。

文献[4]提出了一种中值滤波的方法,文中在此使用的为平均欧式距离的方法。设窗口的大小为 N ,根据中值滤波的思想,此处 N 的选择为奇数,平均欧式距离公式为:

$$S_i = \left[\frac{1}{d} \sum_{i=1}^d (p_i - q_i)^2 \right]^{1/2} \quad (1)$$

其中 d 为数据的维数, i 为 1, ..., N 。计算出窗口内的所有 S_i , 排序所有 S_i , 得到最小 S_i 所对应的运动矢量的值,即为窗口的中值运动矢量值,这样就达到了中值滤波的目的。

2.3 运动矢量场的累积

宏块运动矢量归一化和滤波去噪之后,需对运动矢量进行前后帧的累积获取稠密的运动矢量场,运动矢量场的累积分为前向投影和后向投影,文中采取后向投影法。

设当前帧的运动矢量为 MV_n , 下帧的运动矢量为 MV_{n+1} , $n+1$ 帧往 n 帧投影,形成致密的投影矢量,这样运动矢量形成了累积。

3 聚类方法

3.1 减法聚类

减法聚类是利用高斯密度函数计算样本点的密度

函数,根据样本点的密度大小进行排序,选取聚类中心,且减法类与计算数据的数量呈线性关系,且与计算的数据的特征维数无直接关系^[5]。

设样本数据点 $X_i, i = 1, 2, \dots, n$, 且每个样本的维数为 s , 其样本密度函数为:

$$D_i = \sum_{j=1}^n \exp(-\alpha \|x_i - x_j\|^2) \quad (2)$$

$$\partial = \frac{4}{r_0^2} \quad (3)$$

减法聚类的计算过程:

(1) 使用公式(2)计算每个样本的点的密度大小,选取密度最高的样本为第一个聚类中心 x_{c1} , D_{c1} 为其密度,半径 r_0 为聚类中心的邻域。

(2) 设 x_{ci} 为第 i 的样本聚类中心,对应的密度大小为 D_{ci} , 每次选取新的聚类中心后,需要使用公式(4)修正每个样本的密度大小,最后使用公式(5)判断新的聚类中心是否成立。

$$D_i = D_i - D_{ci} \exp(-\beta \|x_i - x_k\|^2) \quad (4)$$

$$\beta = \frac{4}{r_\beta^2} \quad (5)$$

$$\frac{D_{ci+1}}{D_i} < \varepsilon \quad (6)$$

(3) 文献[5]中 r_0, r_β 都是常数,这种处理方式导致和数据相关性较小,因此文献[6]提出了一种确定 r_0, r_β 的方法:

$$r_a = \frac{1}{2} \min_k \{ \max \{ \|x_i - x_k\| \} \} \quad (7)$$

$$r_b = \frac{1}{2} \sqrt{\frac{1}{n(n-1)} \sum_{k=1}^n \sum_{i=1}^n \|x_i - x_k\|^2} \quad (8)$$

公式(7)、(8)可以替换 r_0, r_β 。且 r_a, r_b 值以外的样本点对聚类中心的密度影响较小,在样本数量一定的情况下, r_a, r_b 越大,聚类数目越少, r_a, r_b 越小,聚类数目越多。

3.2 核的模糊 C-均值聚类算法

模糊 C-均值聚类(FCM)算法由 Dunn 提出,经由 Bezdek 改进,发展起来的一种模糊聚类算法,现已成功运用于多个领域。FCM 由于对噪声和野值敏感,且算法倾向于发现致密的球类聚集结构,而对非凸的聚类结构效果并不理想,因此,针对以上问题,将核引入到聚类算法中,产生了核 C-均值聚类算法(KFCM)^[7]。

设样本聚类空间为 $X = \{x_1, x_2, \dots, x_n\}, x_j \in R^s, j = 1, 2, \dots, n$, 核的映射为 $\varphi: x \rightarrow \Phi(x), \Phi(\cdot)$ 为非线性函数,核 C-均值聚类目标函数^[4~7], 见公式(9):

$$J_m(U, V) = \sum_{i=1}^c \sum_{k=1}^N u_{ij}^m \| \Phi(x_j) - \Phi(v_i) \|^2 = \sum_{i=1}^c \sum_{k=1}^N u_{ij}^m (K(x_j, x_j) + K(v_i, v_i) - 2K(x_j, v_i)) \quad (9)$$

其中 $\Phi(x_j)$ 和 $\Phi(v_i)$ 分别表示样本空间和聚类中心向高维空间的映射,且

$$\Phi(v_i) = \sum_{k=1}^n u_{ik}^m \Phi(x_k) / \sum_{k=1}^n u_{ik}^m \quad (10)$$

由于不知道 Φ 的具体形式,需要利用核函数与特征空间的关系计算 $K(x_j, v_i)$ 和 $K(v_i, v_i)$, 因为 $K(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$, 所以可以得到公式(11), (12):

$$K(x_j, v_i) = \langle \Phi(x_j), \Phi(v_i) \rangle = \sum_{k=1}^n u_{ij}^m K(x_k, x_j) / \sum_{k=1}^n u_{ik}^m \quad (11)$$

$$K(v_i, v_i) = \langle \Phi(v_i), \Phi(v_i) \rangle = \sum_{k=1}^n \sum_{s=1}^n u_{ik}^m u_{is}^m K(x_k, x_s) / \left(\sum_{k=1}^n u_{ik}^m \right)^2 \quad (12)$$

引入核函数后的隶属度函数为公式(13):

$$u_{ij} = \frac{(1 / (\| \Phi(x_j) - \Phi(v_i) \|^2))^{1/(m-1)}}{\sum_{i=1}^c (1 / (\| \Phi(x_j) - \Phi(v_i) \|^2))^{1/(m-1)}} = \frac{(1 / (K(x_j, x_j) - 2K(x_j, v_i) + K(v_i, v_i)))^{1/(m-1)}}{\sum_{i=1}^c (1 / (K(x_j, x_j) - 2K(x_j, v_i) + K(v_i, v_i)))^{1/(m-1)}} \quad (13)$$

核函数发展到现在,已有多种用于实际当中,高斯核函数对野值点的敏感度低,因此文中选取的核函数为高斯核函数:

$$K(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right) \quad (14)$$

σ 为尺度参数。

3.3 有效指标的判断

KFCM 需要预先指定聚类数目 c , 但实际对于样本空间,最优的划分是无法预先确定的,因此,对应不同的划分,最后的样本的聚类划分效果是不同的。通过聚类有效性指标的判断,能够确定聚类划分的质量。同样,使用聚类有效性指标,能够对样本的最优聚类划分进行划分。当前比较有代表性的有 $V_{pe}, V_{pe}, V_{xb}, V_{fs}$ 等,其中由 Xie 和 Beni 提出的有效性函数^[8]:

$$V_{xb} = \frac{\sum_{i=1}^c \sum_{j=1}^n u_{ij}^m \|x_j - v_i\|^2}{n(\min_{j \neq k} \|v_j - v_k\|^2)} \quad (15)$$

在此,可将核函数引入有效性函数,且文献[9]证明是可行的,但是通过文献的实验结果可以看出引入核函数与原有效函数的效果一样,因此,文中使用公式(15)作为有效函数。

4 算法原理及步骤

文献[10]得出权重系数 m 的最佳取值空间为[1.5, 2.5], 一般的研究取值为 2, 因此文中的权重系数的取值为 2。文献[11]通过相关推导以及实验证明,对 c_{\max} 的选择是 $c_{\max} \leq \sqrt{n}$, 因此,文中的聚类数目 c 的搜索范围为 $1 \sim \sqrt{n}$ 。

计算聚类 c 的过程中,根据有效函数的值确定,文献[8, 12]提出了一种方法,画出有效性函数的聚类曲线图,选出第一个出现递减趋势的点做为 c , 通过实验数据证明,得到较好的验证。对聚类数目有规律性的选择,而不是随机性的选择,能够快速有效地确定聚类数目,减少计算的复杂性。

模糊矩阵进行分类的过程中,由于聚类数目是未知的,隶属度函数的初始化矩阵一般是随机初始化的,随机初始化的不确定性,影响隶属度矩阵的迭代计算及其他计算。因此,文中在初始化隶属度矩阵时,采用减法矩阵的密度函数的方法,确定聚类的中心,以及使用公式(13)未引入核函数的公式计算隶属度矩阵。使用这种方法初始化的隶属度矩阵,误差小,利于数据的准确性。

算法过程如下:

Step1 对运动矢量进行处理。

Step2 设置权重系数 m , 目标函数精度 ε , 初始化聚类数目 c 。

Step3 根据减法聚类的原理初始化聚类中心,并且依据公式(13)未引入核函数的公式初始化隶属度矩阵。

Step4 依据公式(9)计算目标函数 $J_m(n)$, 根据公式(13)更新隶属度矩阵。

Step5 若 $|J_m(n) - J_m(n-1)| < \varepsilon$, 则转向 Step6, 否则转向 Step4 继续计算。

Step6 利用公式(14)计算聚类划分的有效性,绘制曲线图。

Step7 若图像的 $c < \sqrt{n}$, 则 $c = c + 1$, 转向 Step3, 否则,根据画出的曲线图,获得第一个趋势减小的点即为聚类的划分点,并根据计算的隶属度函数获取聚类样本,进而分割出运动图像。

5 实验结果

通过采用 Hall 视频序列来验证文中提出的算法。实验平台的电脑配置为内存为 1G, Intel 双核 1.8G, H. 264 编解码软件版本为 JM7.5, Hall 视频测试序列为经过 JM7.5 编码,其中每帧的大小为 176×144, 视频序列的总帧数为 300, 处理每帧的速度为 35ms, 选取的视频序列帧为 154, 212, 如图 2 所示。

图 2 是 Hall 视频序列的分割结果,第二列是文献[13]的分割结果,第三列是文中的分割结果,通过实验结果可以看出,文中的算法,较好地分割实验数据中运动对象的轮廓。因此,通过上面的实验结果可以得出,文中提出的运动对象的处理方法是合理的。



图 2 Hall 序列的分割结果

6 结束语

文中基于 H. 264 压缩域提出了一种分割算法,首先对运动矢量进行处理,包括归一化、去噪、累积,然后将有效函数作为循环条件,通过使用核聚类对运动矢量进行分割,最后分割出运动对象。由于使用核聚类和自适应的分类方法,且初始化隶属度函数是使用当前数据进行初始化,而不是随机的初始化,因此能够准确地自学习地分割运动对象。实验结果表明能够有效分割出运动对象,但文中的视频分割序列是基于背景静止的条件,因此,对于摄像头运动的情况还需要进一步研究。

参考文献:

[1] Zeng W, Du J, Gao W. Robust moving object segmentation on H. 264/AVC compressed video using the block-based MRF

(上接第 86 页)

41(6):683-687.
[5] Badia L, Lindstrom M, Zander J, et al. An Economic Model for the Radio Resource Management in Multimedia Wireless Systems[J]. Computer Communications, 2004, 27(11): 1056-1064.
[6] 丛秋实,姜爱全,潘纪明. 无线/移动网络中基于价格的接纳控制方案[J]. 计算机工程与应用, 2006(33): 125-128.
[7] Senouci S M, Beylot A, Pujolle G. Call admission control in cellular networks: a reinforcement learning solution[J]. International Journal of Network Management, 2004, 14(2): 89-103.
[8] 洪 龙,肖奚安,朱梧楨. 中介真值程度的度量及其应用

model[J]. Real-time Imaging, 2005, 11(4): 290-299.

[2] Liu Z, Lu Y, Zhang Z. Real-time spatiotemporal segmentation of video objects in the H. 264 compressed domain[J]. Journal of Visual Communication and Image Representation, 2007, 18(3): 275-290.
[3] Babu R V, Ramakrishnan K R, Srinivasan S H. Video object segmentation: A compressed domain approach[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2004, 14(4): 462-474.
[4] Haykin S. Kalman filtering and neural networks[M]. New York: John Wiley and Sons, 2001.
[5] Chiu S L. Fuzzy model identification based on cluster estimation[J]. Journal of Intelligent and Fuzzy Systems, 1994, 2(3): 263-278.
[6] 裴继红,范九伦,谢维信. 聚类中心的初始化方法[J]. 电子科学学刊, 1999, 21(3): 320-325.
[7] 张 莉,周传达,焦李成. 核聚类算法[J]. 计算机学报, 2002, 25(6): 587-590.
[8] Xie X L, Beni G. A Validity Method for Fuzzy Clustering[J]. IEEE Trans on PAMI, 1991, 13(8): 841-847.
[9] 增运伟,金炜东,朱 明,等. 核模糊 C 均值算法的聚类有效性研究[J]. 计算机科学, 2007, 34(2): 207-210.
[10] Kim D W, Lee K Y, Lee D. Evaluation of the Performance of Clustering Algorithms in Kernel Induced Feature Space[J]. Pattern Recognition, 2005, 38(4): 607-611.
[11] Zahid N, Limouri M, Essaid A. A new cluster validity for fuzzy clustering[J]. Pattern Recognition, 1999, 32(7): 1089-1097.
[12] 于 剑,程乾生. 模糊聚类方法中的最佳聚类数的搜索范围[J]. 中国科学(E 辑), 2002, 32(2): 275-280.
[13] Venkatesh B R, Ramakrishnan K R, Srinivasan S H. Video Object Segmentation: A Compressed Domain Approach[J]. IEEE Transactions on Circuits Systems for Video Technology, 2004, 14(4): 462-474.
[1] [J]. 计算机学报, 2006, 29(12): 2186-2193.
[9] 洪 龙,肖奚安,朱梧楨. 中介真值程度的度量及其应用 (II) [J]. 计算机学报, 2007, 30(9): 1551-1558.
[10] Zhou Ningning, Zhao Zhengxu, Hong Long, et al. A New Image Edge Detection Algorithm Based on Measuring of Medium Truth Scale[C]//Proceeding of 5th IEEE International Conference on Networking, Sensing and Control. Sanya, China: [s. n.], 2008: 698-703.
[11] 成卫青,龚 俭,丁 伟. 基于流特性和真值程度的 VoIP 语音质量单端客观评价[J]. 通信学报, 2008, 29(4): 30-39.
[12] 赵焕臣,许树柏,和金生. 层次分析法[M]. 北京: 科学出版社, 1986: 23-32.

基于核聚类的H.264压缩域的运动对象分割方法

作者：[孔祥鹏](#)，[沙宁](#)，[邢薇](#)，[KONG Xiang-peng](#)，[SHA Ning](#)，[XING Wei](#)

作者单位：[哈尔滨工程大学 计算机科学与技术学院, 黑龙江 哈尔滨, 150001](#)

刊名：[计算机技术与发展](#)



英文刊名：[Computer Technology and Development](#)

年，卷(期)：2013(8)

本文链接：http://d.wanfangdata.com.cn/Periodical_wjfz201308022.aspx