

基于模糊聚类的 H.264 压缩域视频对象分割

轩腾蛟, 康志伟

(湖南大学信息科学与工程学院, 长沙 410082)

摘要: 鉴于压缩域视频运动分割方法在分割速度上的优越性, 提出一种基于 H.264 的压缩域视频运动对象分割方法, 对初始的运动矢量场进行去噪、中值滤波、校正和累积处理, 得到更可靠的运动矢量场, 用改进的模糊 C-均值聚类算法分割出视频序列中的运动对象。实验结果表明, 该方法可以快速准确地提取出视频序列中的运动对象。

关键词: H.264 标准; 压缩域; 对象分割; 中值滤波; 运动矢量; 模糊 C-均值聚类

Video Object Segmentation in H.264 Compressed Domain Based on Fuzzy Clustering

XUAN Teng-jiao, KANG Zhi-wei

(College of Information Science and Engineering, Hunan University, Changsha 410082, China)

【Abstract】 According to the advantage of segmentation speed in the compressed domain, this paper proposes a method of video object segmentation based on H.264. It eliminates the noise, median filter, correction and accumulation to processing initial motion vectors and uses fuzzy C-means clustering algorithm to process motion vectors and segments the moving objects. Experimental results show that the method can get the moving object fast and accuracy.

【Key words】 H.264 standard; compressed domain; object segmentation; median filtering; motion vector; fuzzy C-means clustering

DOI: 10.3969/j.issn.1000-3428.2012.03.088

1 概述

随着多媒体技术的飞速发展和第 2 代视频编码技术的出现, 视频对象分割技术展现出广阔的应用前景, 可用于视频编码、视频检索、视频监控、模式识别等。从数字视频信号中提取出有意义的运动对象信息, 是准确进行视频运动对象分割的关键技术。传统的视频对象分割技术是在像素域中进行的, 目前大部分视频码流是以压缩形式存储的, 在像素域进行视频运动对象提取, 需要对压缩码流进行完全解码, 这样增加了分割耗时, 而且像素域分割算法运算复杂度大, 需要大量的计算时间和存储空间, 达不到高效处理压缩域视频流的要求, 所以, 从压缩视频流中直接提取运动对象已经成为必然。新一代视频压缩标准 H.264 具有良好的网络亲和性和优异的压缩性能, 利用 H.264 标准进行压缩域视频运动对象分割意义重大。

目前压缩域视频运动对象的分割方法大致可分为 3 类:

- (1) 基于变化检测的分割方法;
- (2) 基于时空联合的分割方法;
- (3) 基于运动矢量场的分割方法。本文采用方法(3)。

近年来出现了很多基于 H.264 的视频运动对象分割方法: 文献[1]提出了一种基于块的 MRF 模型运动对象分割方法, 利用 MRF 能量最小化把相似的 MV 合并成运动对象, 模型将 H.264 中块的不同尺寸作为一个线索加以研究, 但是仅限于背景是静止的视频序列。文献[2]提出了一种对运动背景和静止背景都适用的基于匹配矩阵的视频运动对象分割方法, 首先归一化提取的运动矢量场, 然后利用连续多帧的后向投影得到可靠的运动矢量场, 对其进行全局的运动补偿来获得各块的残差, 最后利用区域生长算法得到运动相似的区域, 进行基于块的运动对象分割, 但该方法的运算量较大。文献[3]通过时域累积和空域内插的方法估计出一个稠密的

运动矢量场, 利用 EM 聚类算法聚类, 达到分割出运动对象的目的。本文提出一种基于 H.264 压缩标准的压缩域视频运动对象分割方法, 利用改进的模糊 C-均值算法聚类, 不但分割结果更加准确, 而且节省了运算时间, 满足视频的实时性要求。

2 本文方法流程

本文方法直接从压缩视频流中提取出运动矢量场, 无须对压缩视频流进行解压缩处理。提取的初始运动矢量场经过去噪声处理, 得到平滑的运动矢量场, 然后对运动矢量场进行校验, 通过后向迭代的方法累积运动矢量场, 得到可靠、稠密的运动矢量场, 最后提取出运动对象, 其流程如图 1 所示。

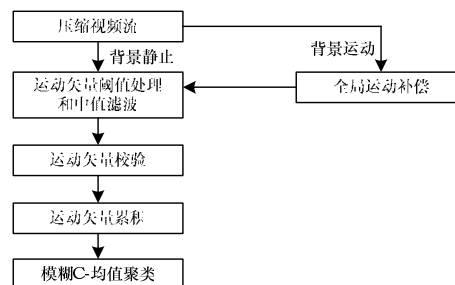


图 1 本文方法流程

3 运动矢量场处理

初始的运动矢量场是从压缩视频流中直接提取出来的,

基金项目: 国家自然科学基金资助项目(60702065)

作者简介: 轩腾蛟(1988—), 男, 硕士, 主研方向: 信号与图像处理; 康志伟, 教授

收稿日期: 2011-07-14 E-mail: xuantengjiao2@163.com

是稀疏的、有噪声的、不可靠的，需要对其进行以下处理，使其变得稠密、可靠。

3.1 运动矢量场噪声处理

实验用的是背景静止的视频序列，无须进行全局运动补偿。如果是背景运动的视频序列，需要先进行全局运动补偿，然后对运动矢量场进行处理。

图像静止区域往往会产生数值较小的噪声，可以设置一个合适的阈值 T 来去除这类噪声(一般设置 $T=2$)。阈值的设置十分重要，阈值过大则有可能将一些运动块滤掉，阈值过小则有可能将有背景噪声的块当成运动块，实验采用的方法是将运动矢量 2 个方向分量绝对值的和作为一个参数，如式(1)所示：

$$MV = \begin{cases} MV & |MV_x| + |MV_y| \geq T \\ 0 & |MV_x| + |MV_y| < T \end{cases} \quad (1)$$

经过上式的处理，图像静止区域零值附近的噪声被滤掉了。

中值滤波是一种典型的低通滤波，对象边缘运动矢量和噪声运动矢量是不同的，中值滤波器可以在保护运动对象边缘的同时滤除运动矢量场中一些孤立的点。中值滤波的原理是把以点 (x,y) 为中心的窗口内所有运动矢量的数值从大到小排序，取其中间值作为当前点的值。实验采用 3×3 中值滤波器。

3.2 运动矢量场的校验

经过上面处理的运动矢量场中零值附近的噪声和一些孤立点噪声被去除，但是运动矢量场可能还含有其他噪声，有些块的运动矢量并不准确，需要做进一步处理，本文对运动矢量的幅值进行处理，对那些幅值突然变化的运动矢量场，利用式(2)增加运动矢量的可靠性：

$$DMV_{(i,j)} = \frac{\left| \frac{MV_{(i,j)} - \sum_{(i,j) \in \theta} \frac{MV_{(i,j)}}{N}}{\sum_{(i,j) \in \theta} \frac{MV_{(i,j)}}{N}} \right|}{\sum_{(i,j) \in \theta} \frac{MV_{(i,j)}}{N}} \quad (2)$$

其中， $DMV_{(i,j)}$ 为点 (i,j) 处的运动矢量幅值 $MV_{(i,j)}$ 背离相邻非零点运动矢量的程度； θ 表示与 $MV_{(i,j)}$ 相邻的 8 个点； N 为邻域中非零点的个数； $MV_{(i,j)}$ 为邻域非零点的运动矢量。可以为 $DMV_{(i,j)}$ 设定一个阈值 T_1 ，如果点 (i,j) 处的 $DMV_{(i,j)}$ 大于 T_1 ，就用该点邻域的平均值代替该点的运动矢量值，否则，保留该点原来的运动矢量值。

3.3 运动矢量场累积

为了得到更稠密、更可靠的运动矢量场，需要对运动矢量场进行累积，一般选取连续的 2 帧~3 帧进行累积。本文采用连续多帧的后向迭代进行累积。设 $f(x,y,t)$ 、 $f(x,y,t+1)$ 为视频序列中连续两帧，运动矢量分别是 MV_t 、 MV_{t+1} 。把运动矢量 MV_{t+1} 向 t 帧投影，投影得到的运动矢量为 TMV_{t+1} ，这样 t 帧的运动矢量是 MV_t 和 TMV_{t+1} 的合成，就被致密了。运动矢量场累积能很好地去除提取运动对象出现空洞现象。

3.4 模糊 C-均值聚类

聚类是按照某种准则将物体或者抽象的对象进行分类的过程。国内外学者对其进行了广泛的研究，提出了很多分类方法。文献[4]提出基于 FCM 的自适应增量式聚类算法 AIFCM，该算法结合密度和集合的思想，给出一种自动确定

聚类初始中心的方法，能在聚类过程中动态改变聚类结果，改善聚类的质量，减少人为的主观因素，获得比较符合用户需求的聚类结果，并能在原有聚类结果的基础上简单地处理更新数据，过滤噪声数据，较好地避免大量重复计算。文献[5]提出的模糊紧性分离性(Fuzzy Compactness and Separation, FCS)算法规定有固定的核边界，而一个核中的所有对象具有相同的隶属度值而不能被辨识，造成分类准确率下降；文献[6]对初始聚类中心的选择进行了研究，采用改进的基于相似性阈值和最小距离原则的聚类方法对样本粗聚类，得到各个类的中心，得到了很好的效果，但是增加了聚类的计算量，不适合类数目已经确定的样本。

本文提出一种改进的模糊 C-均值聚类算法，该方法优化了初始聚类中心的选择，使初始聚类中心更准确。

设 $X=(X_1, X_2, \dots, X_n)$ 为被分类的样本集合，改进的 FCM 算法的步骤为：

Step1 设定类间的最小距离 $a(a>0)$ 、模糊因子 m (一般 $m=2$) 和精度值 ζ ，统计样本 X 中出现最多的个体 X_i ，然后以 R 为半径：

$$R = \frac{1}{2} \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^n \sum_{l=1}^n \|X_i - X_l\|^2} \quad (3)$$

在样本 X 中找出符合 $|X_j - X_i| \leq R$ 的样本，求这些样本的平均值作为第 1 个聚类中心。

Step2 同理，在剩下的样本中找出与第 1 个聚类中心的距离大于 a 且出现次数最多的一个样本，然后以 R 为半径按照 Step1 求出第 2 个聚类中心。

Step3 重复 Step2，直到找到 c 类聚类中心 $W_i(k)$ ， $i=1, 2, \dots, c$ 为止。

Step4 设置迭代次数 $k=1$ ，以 Step3 中的结果作为初始聚类中心 $W_i(k)$ ， $i=1, 2, \dots, c$ 。

Step5 根据 $W_i(k)$ 按照式(4)求出隶属度矩阵 $\mu_{i,j}(k)$ ， $i=1, 2, \dots, c, j=1, 2, \dots, n$ ：

$$\mu_{i,j} = \frac{1}{\sum_{k=1}^c \left[\frac{d_{ij}^2(X_j - W_i)}{d_{ij}^2(X_j - W_k)} \right]^{\frac{1}{m-1}}} \quad (4)$$

Step6 根据隶属度矩阵 $\mu_{i,j}(k)$ ，按照式(5)修正所有的聚类中心 $W_i(k+1)$ ， $i=1, 2, \dots, c$ ：

$$W_i = \frac{\sum_{l=1}^n (\mu_{il}^m X_l)}{\sum_{l=1}^n \mu_{il}^m} \quad (5)$$

Step7 计算误差：

$$\varepsilon = \sum_{i=1}^c \|W_i(k+1) - W_i(k)\|^2 \quad (6)$$

如果 $\varepsilon < \zeta$ ，输出聚类中心，否则， $k=k+1$ ，转 Step5。

Step8 如果 $d^2_{i,j}(X_j - W_i) < d^2_{i,j}(X_j - W_k)$ ，其中， $k=1, 2, \dots, c$ ； $i=2, \dots, c$ ； $k \neq 1$ ，则将 X_j 归入第 i 类。

4 实验结果与分析

实验采用典型的 Claire 和 table tennis 序列，2 个视频序列均为 qcif 格式。利用 JM86 在编码端编码。视频序列的帧格式为 IPPP，帧率为 28 帧/s。实验中几个阈值的设定十分重要，合理的阈值才能得到较好的实验结果。实验在聚类算法中引入搜索半径，明确了第 1 个类别的中心，减少了迭代次

数, 从而减少了运算时间。实验结果如图 2 和图 3 所示。

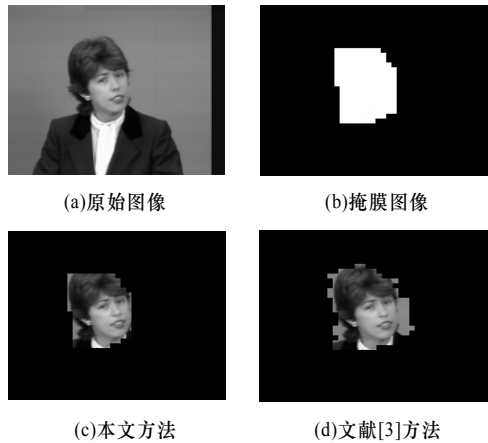


图 2 Claire 序列实验结果

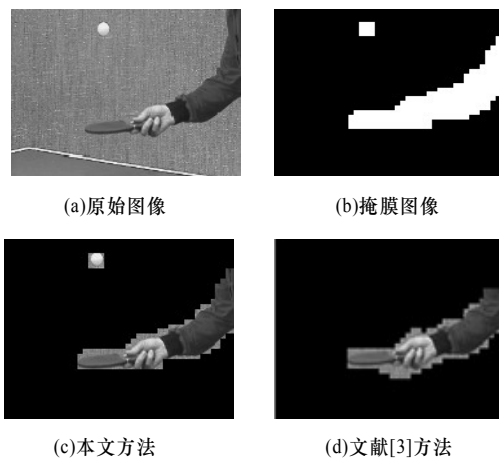


图 3 tennis 序列实验结果

从图 2 可以看出, 本文方法提取的运动对象相对于文献[3]的方法轮廓更小, 且消除了一些锯齿状背景区域; 从

(上接第 266 页)

在对网络故障分析中发现, 蓄意攻击对整个网络的连通性会产生较大的影响。所以, 在故障分析后, 对二部图进行了一些优化: 在二部图的“下集”中搜索与每一个 HUB 节点直接相连的“上集”节点(在此二部图中, 度为 24 的节点为 HUB 节点), 找到与“上集”节点直接相连的度最小的“下集”节点, 建立集合 $N = (v_1, v_2, \dots, v_n)$ 。在“上集”中添加新的节点, 并将集合中的节点与新节点相连。以优化的二部图为基础再次进行蓄意攻击后发现, MCR 值维持在 0.83 左右, 为 0 的 Laplacian 特征根在 2 个左右, 但平均路径长度仍然较高。

通过实验可以得出以下 2 点结论: (1)兰州市因为其狭长的地理特点, 使得公共交通网络的发展受到限制。(2)兰州相对于北京等地的节点较少, 使得度分布相对不明显。

6 结束语

本文通过对二部图进行优化, 使网络的连通和性能有了一定的改善, 但是平均路径长度在实际网络中仍然比较重要, 如何使得网络在受到一定的蓄意攻击下依然有较小的平均路径长度, 同时保证网络的连通和性能, 是今后所要研究的方向之一。此外, 当一个网络并不能很清晰地给出其连接关系的时候, 图中的一些边就会缺失, 那么如何通过二部图预测缺失的边也是今后所要研究的一个方向。

图 3 可以看出, 本文方法可以准确地分割出手臂和乒乓球, 而文献[3]只分割出了手臂。因此, 本文方法可以较好地分割出运动对象。

5 结束语

本文提出一种基于 H.264 压缩域视频运动对象的分割方法, 从压缩码流中直接提取运动矢量场, 经过滤波去噪、校验、运动矢量场累积, 最后利用改进的模糊 C-均值聚类算法进行分类, 准确地提取出运动对象, 实验结果表明该方法可以较好地分割出运动对象, 满足视频的实时性要求。由于实验采用的是迭代算法, 因此提取的运动对象存在轮廓偏大等不足之处, 有待于下一步改进。

参考文献

- [1] Zeng Wei, Du Jun, Gao Wen, et al. Robust Moving Object Segmentation on H.264/AVC Compressed Video Using the Block-based MRF Model[J]. Real-time Imaging, 2005, 11(4): 290-299.
- [2] Liu Zhi, Lu Yu, Zhang Zhaoyang. 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] 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.
- [4] 张忠平, 陈丽萍, 王爱杰, 等. 基于自适应模糊 C-均值的增量式聚类算法[J]. 计算机工程, 2009, 35(6): 60-62, 65.
- [5] Wu Kuolung, Yu Jian, Yang Minshen. A Novel Fuzzy Clustering Algorithm Based on a Fuzzy Scatter Matrix with Optimality Tests[J]. Pattern Recognition Letters, 2005, 26(5): 639-652.
- [6] 张新波. 两阶段模糊 C-均值聚类算法[J]. 电路与系统学报, 2005, 10(2): 117-121.

编辑 张正兴

参考文献

- [1] Duncan J, Steven H. Collective Dynamics of Small-world Networks[J]. Nature, 1998, 393(6684): 440-442.
- [2] Barabási A, Albert R. Emergence of Scaling in Random Networks[J]. Science, 1999, 286(5439): 509-511.
- [3] Guillaume J L, Latapy M. Bipartite Graphs as Models of Complex Networks[J]. Physica A, 2006, 371(2): 795-813.
- [4] Molloy M, Reed B. The Size of the Giant Component of a Random Graph with a Given Degree Sequence[J]. Combinatorics, Probability and Computing, 1998, 7(3): 295-305.
- [5] Dorogovtsev S N, Mendes J F F. Evolution of Networks[J]. Advances in Physics, 2002, 51(4): 1079-1187.
- [6] Sienkiewicz J, Janusz A. Statistical Analysis of 22 Public Transport Networks in Poland[J]. Physical Review E, 2005, 72(4): 46-50.
- [7] 汪小帆, 李翔, 陈关荣. 复杂网络理论及其应用[M]. 北京: 清华大学出版社, 2006.
- [8] Yang Xuhua, Wang Bo, Wang Wanliang. Research on Some Bus Transport Networks with Random Overlapping Clique Structure[J]. Communications in Theoretical Physics, 2008, 50(5): 1249-1254.
- [9] 叶东海, 蒋国平, 宋玉蓉. 多局域世界复杂网络中的病毒传播研究[J]. 计算机工程, 2010, 36(23): 130-132.

编辑 任吉慧

