

目标基视频编码中的运动目标提取与跟踪新算法

朱仲杰^{1,2}, 蒋刚毅^{2,3}, 郁梅^{2,3}, 王让定², 吴训威^{1,2}

(1. 浙江大学信息与电子工程系, 浙江杭州 310027; 2. 宁波大学电路与系统研究所, 浙江宁波 315211;
3. 西安电子科技大学 ISN 国家重点实验室, 陕西西安 710071)

摘要: 自动、快速的视频目标提取与跟踪是目标基视频编码中的一项关键技术. 本文提出一种运动目标提取与跟踪新算法. 首先, 根据多帧运动信息和高阶统计检测方法得到二值运动掩模图像, 然后提出一种改进分水岭算法对运动区域及其周围部分进行分割. 将二者所得结果进行投影运算, 得到最终运动目标. 最后提出一种运动目标跟踪新算法, 能对目标进行有效的跟踪. 实验结果说明了本文算法的有效性.

关键词: 视频目标分割; 有限区域分割; 改进分水岭算法; 运动跟踪

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2003) 09-1426-03

New Algorithm for Extracting and Tracking Moving Object in Object2Based Video Coding

ZHU Zhongjie^{1,2}, JIANG Gangyi^{2,3}, YU Mei^{2,3}, WANG Rangding², WU Xunwei^{1,2}

(1. Dept. of Information and Electronics Engineering, Zhejiang University, Hangzhou, Jiangsu 310027, China;
2. Institute of Circuits and Systems, Ningbo University, Ningbo, Jiangsu 315211, China;
3. National Key Laboratory of ISN, Xidian University, Xidan, Shaanxi 710071, China)

Abstract: Fast and automatic segmentation and tracking of video object is a key technology in object2based video coding. A algorithm for extracting and tracking moving object is proposed. First, a binary motion image is achieved based on high order statistics detection and motion information of multiple frames. Then, an improved watershed algorithm is used to segment motion region and its surrounding areas. Thus the moving object with fine edges can be extracted with the spatial and temporal segmented results. Finally, a new algorithm is proposed to efficiently track moving object. Experimental results show that the proposed algorithm is efficient.

Key words: video object segmentation; segmentation of limited region; improved watershed algorithm; motion track

1 引言

第二代视频编码方案MPEG2引入视频目标的概念来实现基于内容的表示^[1]. 视频场景被分割成若干个视频目标, 不同的视频目标根据其结构特点采用不同的编码策略. 这种方案具有高压缩比、良好的视觉效果、基于内容交互和基于内容分级扩展等诸多优点^[2]. 视频运动目标分割与跟踪是目标基视频编码中的一项关键技术. 任意场景下视频目标的自动提取是个难点^[3,4]. 本文首先基于高阶统计模型和多帧运动信息提取完整连续的运动区域. 其次, 提出一种改进分水岭算法对运动区域及其附近区域进行空域分割. 然后, 将空域分割结果和时域提取结果进行投影运算, 得到具有精确边缘的运动目标. 最后提出一种简单、快速的跟踪算法对后续图像中的运动目标进行跟踪. 通过对测试序列的模拟试验, 表明本文算法是有效的.

2 高阶统计模型与运动检测

在运动目标的提取中, 运动是最重要的特征. 根据光流法通过计算绝对位移帧差可进行运动检测. 定义连续两帧的帧差图像为

$$d_{k, k+1}(x, y) = m_{k+1}(x, y) - m_k(x, y) + n_{k+1}(x, y) - n_k(x, y) \quad (1)$$

式中, $m(x, y)$ 表示运动区域(包括物体运动、暴露和遮挡等区域), $n(x, y)$ 为图像噪声. 理论上, $d_{k, k+1}(x, y)$ 为零的点代表静止区域, 但由于存在观察噪声, 使得实际上静止区域的 $d_{k, k+1}(x, y)$ 并不等于零. 为了抑止噪声, 精确提取属于运动区域的象素点, 采用高阶统计的检测方法. 运动区域检测过程可看作是从高斯信号中提取随机非高斯信号的过程^[5]. 为了判断象素 (x, y) 是否属运动区域, 先在帧差图像上计算以 (x, y) 为中心的矩形窗口 $w(x, y)$ (窗口尺寸: $N_w @ N_w$) 的四阶矩

$$m_{k, k+1}^{(4)}(x, y) = \frac{1}{N_w^2} \sum_{(p, q) \in w(x, y)} (d_{k, k+1}(p, q) + d_{k, k+1}(x, y))^4 \quad (2)$$

其中, $d_{k, k+1}(x, y) = \frac{1}{N_w} \sum_{(p, q) \in w(x, y)} d_{k, k+1}(p, q)$ 是窗口 $w(x, y)$ 内帧差信号均值. 给定阈值 $T = u(R_{od}^2)^2$, 可得

$$m_{k, k+1}(x, y) = \begin{cases} 1, & m_{k, k+1}^{(4)}(x, y) \setminus T \\ 0, & \text{其它} \end{cases} \quad (3)$$

收稿日期: 20020625; 修回日期: 20021228

基金项目: 浙江省青年人才基金(No. RC01057); 教育部留学回国人员科研基金(No. 2001331); 西安电子科技大学 ISN 国家重点实验室开放基金(No. H02025); 浙江省自然科学基金(No. 601017); 宁波市重点博士基金(No. 01J20300205, 2003A61006)

$m_{k, k+1}(x, y) = 1$ 表示像素点 (x, y) 属于运动区域, 反之则属于静止区域. 式中, u 是由实验确定的常系数, R_{ad}^2 是噪声方差, 由下式给出

$$R_{ad}^2 = \frac{1}{N_s^2} \sum_{(p, q) \in S} (d_{k, k+1}(p, q) - d_{k, k+1})^2 \quad (4)$$

式中, S 属于图像静止背景区域, N_s 为所选取区域尺寸大小.

以上得到的 $\{m_{k, k+1}(x, y)\}$ 仅仅是利用两帧的帧差信息, 提取的运动区域可能不够完整. 为了提出完整的运动区域, 采用连续多帧的帧差信息提取二值运动掩模图像 $\{m(x, y)\}$, 其中 $m(x, y) = \max\{m_{k, k+1}(x, y) | k = 0 \sim M-1\}$ (5) 式中, M 表示参与运动区域提取的图像总共有 $M+1$ 帧. 最后, 对 $m(x, y)$ 进行形态开运算和闭运算以消除凸凹区, 得到最终的二值运动掩模.

3 基于改进分水岭算法的空域分割

分水岭算法是一种经典的图像分割工具, 传统的分水岭算法容易产生过分割现象. 本文提出一种改进分水岭算法, 以减少过分割现象, 降低算法复杂度. 算法主要步骤简述如下:

(1) 首先计算输入图像的形态梯度图像, 根据形态梯度图像中像素灰度最大与最小值, 确定一个灰度尺度. 从最小灰度的像素开始, 像素差值小于等于的所有像素划归同一个等级. 整帧图像划分为一定数量的等级;

(2) 取出第一个等级的所有像素以分配标号: 相邻像素点分配相同标号, 孤立像素点分配一个独立标号;

(3) 取出下一个等级中的所有像素点, 按步骤(4)-(5)给每个像素分配标号;

(4) 在像素邻域内, 水线标号除外, 若存在已分配了标号的像素点, 且标号相同, 则赋予该像素相同的标号;

(5) 在像素邻域内, 水线标号除外, 若存在已分配了标号的像素点, 但标号不同, 则赋予该像素水线标号;

(6) 在像素邻域内, 水线标号除外, 不存在已分配了标号的像素点, 则赋予该像素一个新标号, 表示新区域;

(7) 重复(3)-(5), 直到所有等级中的所有像素都被分配了标号;

(8) 分配相同标号的所有像素属于同一区域. 检查图像分割后的每个区域的大小, 如果某个区域的像素个数少于 N , 则撤销该区域内所有像素已分配的标号, 重复(1)-(7), 重新给每个像素分配标号. 其中, N 表示区域的临界尺寸, 它由图像的尺寸大小来确定;

(9) 重复(1)-(8)直到图像中的所有像素都分配标号, 且不存在尺寸小于 N 的区域.

4 运动跟踪新算法

本文提出一种基于目标边界轮廓的跟踪新算法, 具有简单、计算量小等特点. 简述如下:

(1) 首先提取运动目标的边界轮廓. 提取的边界轮廓要求满足如下原则: 边界线上的每个边界点, 首尾两个边界点除外, 在其八邻域内必须存在且只存在两个相邻的边界点. 其中, 首尾两个边界点在其八邻域内必须存在且只存在一个边界点. 如果不满足上述原则, 则要进行细化处理. 例如, 图 1

(a) 中黑点表示的边界就不满足上述原则, 细化后的边界如图 1(b) 或图 1(c) 所示;

(2) 搜索某一边界点在下一帧图像中的匹配边界点时, 根据其前面的已知两个相邻边界点的位置来确定搜索区域. 设 A、B 表示待搜索的边界点 C 的前面两个已知相邻边界点, 则 A、B 的位置有四种情况, 如图 2(a)-(d) 所示. 根据 A、B 点的位置特

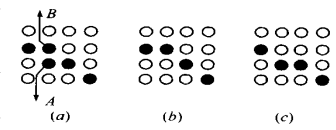


图 1 边界示意图

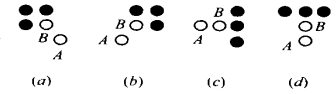


图 2 前后两个已知相邻边界点的四种位置

点, 分别在其前方的三个候选像素点中搜索 C 点的匹配点;

(3) 初始边界点(第一、第二个边界点)不能按照(2)中的方法确定搜索范围. 对于第一个边界点, 给定一个搜索区域, 在此区域内根据一定的匹配准则搜索最佳匹配点. 搜索区域的大小可根据图像序列的运动速度给定, 文中搜索区域大小为 $8@8$;

(4) 如果边界轮廓不止一条曲线, 要对每条边界曲线分别进行跟踪匹配;

(5) 匹配准则按照最小 MSE 准则, 匹配窗口采用 $16@16$ 的图像块.

5 目标提取与跟踪算法步骤

运动目标提取与跟踪的基本步骤如下:

(1) 利用高阶统计检测, 根据多帧的帧差信息提取完整连续的运动区域, 得到最终的二值运动掩模图像 $m(x, y)$;

(2) 由 $m(x, y)$ 确定空域分割的图像区域;

(3) 对待分割的图像区域计算形态梯度图像并化简, 利用改进分水岭算法进行空域分割;

(4) 对分割后的图像与 $m(x, y)$ 进行投影运算, 得到最终的运动目标二值掩模图像 $m_o(x, y)$. 设 $A_i (i = 0, +, N-1)$ 为空域分割后得到的各个区域, N_{A_i} 为 A_i 的大小. 给定阈值 T , 由如下投影运算得到最终的 $m_o(x, y)$

$$m_{oi}(x, y) = \begin{cases} 1, & (x, y) \in A_i \text{ 且 } \frac{1}{N_{A_i}} \sum_{(x,y) \in A_i} m(x, y) \geq T, \\ & i = 0, +, N-1 \\ 0, & \text{其它} \end{cases} \quad (6)$$

$$m_o(x, y) = Y\{m_{oi}(x, y)\}, \quad i = 0, +, N-1 \quad (7)$$

(5) 由 $m_o(x, y)$ 和原图像提取出运动目标, 并进行形态滤波后处理, 得到最终运动目标结果;

(6) 由运动跟踪算法对后续图像进行跟踪, 提取后续图像中的运动目标.

6 实验结果

为了检验新算法的有效性, 进行了计算机仿真测试. 图 3 是 susie 图像序列的试验结果. 首先, 根据多帧帧差信息, 运用高阶统计检测方法, 得到如图 3(b) 所示的二值运动掩模图像. 然后, 根据二值运动掩模图像确定空域分割的图像区域,

本文中分割的图像区域是运动区域及其周围五十个像素点. 图 3(c) 是化简后的待分割形态梯度图像, 图 3(d) 是形态处理后的最终提取的运动目标, 图 3(e) 是运动跟踪的结果, 图 3(f) 和图 3(g) 分别是传统分水岭算法和改进分水岭算法的实验结果. 图 4 是 missa 图像序列的实验结果. 从图 3 和图 4 的结果可以看出, 应用本文算法能有效地提取运动目标, 运动跟踪算法则能对目标进行有效的跟踪.

7 结论

利用多帧的帧差信息和经典的高阶统计检测方法, 能有

效从噪声信号中提取运动信息. 分水岭算法是一种有力的图像分割工具, 但传统的分水岭算法容易产生过分割现象, 本文提出的改进分水岭算法能有效降低过分割现象. 而采用有限区域分割能降低空域分割的计算量. 本文提出的运动跟踪算法计算量很小, 边界跟踪是单向、连续地进行的, 能保证跟踪搜索到的后续轮廓也是单向、连续的. 实验结果表明了本文算法的有效性.



图 3 susie 图像序列实验结果



图 4 missa 图像序列实验结果

参考文献:

- [1] Mech R, Wollbom M. A noise robust method for 2 D shape estimation of moving objects in video sequences considering a moving camera[J]. Signal Processing, 1998, 66(2): 203- 217.
- [2] Kim M, Choi J, Kim D, Lee H et al. A VOP generation tool: Automatic segmentation of moving objects in image sequences based on spatial-temporal information[J]. IEEE Trans on Circuits and Systems for Video Technology, 1999, 9(8): 1216- 1226.
- [3] Castagno R, Ebrahimi T, Kunt M. Video segmentation based on multiple features for interactive multimedia applications[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1998, 8(5): 562- 571.
- [4] Gu C, Lee M C. Semiautomatic segmentation and tracking of semantic video objects[J]. IEEE Trans on Circuits and Systems for Video Technology, 1998, 8(5): 572- 584.
- [5] Neri A, Colonnese S, Russo G. Automatic moving objects and background segmentation by means of higher order statistics[A]. Franklin T Luk. SPIE3024[C]. San Jose, California, USA, 1997. 8- 14.

作者简介:



朱仲杰 男, 1976 年 10 月出生于安徽安庆, 现为浙江大学信电系博士研究生, 主要研究领域为多视域视频图像编码、无线视频编码与通信、视频水印与信息安全等.



蒋刚毅 男, 1964 年 8 月出生于浙江绍兴, 博士, 教授, 主要研究领域为图像处理与视频信号编码、多媒体信息传输与安全技术研究、基于视觉的智能控制.

郁梅 女, 1968 年 12 月出生于四川成都, 博士, 副教授, 主要研究领域为网络视频编码与通信、立体视频图像编码、计算机视觉.