

一种 MPEG-4 视频流的运动目标检测算法

王 泉, 董宝鸳, 田玉敏

(西安电子科技大学 计算机外部设备研究所, 陕西 西安 710071)

摘要: 为了同时提高 MPEG-4 视频流中运动目标检测的精度和速度, 利用 MPEG 中已有的运动信息——离散余弦系数和运动矢量, 提出了一种用于视频中运动目标检测的算法. 该算法首先从 MPEG-4 压缩流中获取运动矢量, 由运动矢量标识出运动区域, 在标识出的运动区域上取离散余弦系数并求互帧差的四阶矩得到运动宏块, 最后通过区域生长得到运动目标. 该算法充分利用了 MPEG-4 码流中蕴含的运动信息, 避免了仅由运动矢量检测运动目标时误差太大的缺点, 也避免了计算全部离散余弦系数时的大量加乘法运算, 从而可以实时地对视频流中的运动目标进行跟踪.

关键词: 目标跟踪; 离散余弦系数; 运动矢量; MPEG-4

中图分类号: TN919.81 **文献标识码:** A **文章编号:** 1001-2400(2007)06-0869-04

A motion object detection algorithm for MPEG-4 video

WANG Quan, DONG Bao-yuan, TIAN Yu-min

(Research Inst. of Peripherals, Xidian Univ., Xi'an 710071, China)

Abstract: Focusing on the problem of moving object detection and tracking on MPEG-4 compressed streams by a stationary camera, an algorithm for detecting the moving object using DC coefficients and motion vectors of the motion information in MPEG is presented. The algorithm is first to mark the moving regions based on the motion vectors, then get the moving macro blocks by obtaining the DC coefficient of each marked block in MPEG data and calculating the fourth-order detector of inter-frame difference, and finally obtain the moving object by region growth. The computational complexity of only computing DC coefficients and the detection error only through motion vectors are reduced by the method. The algorithm presented makes the real detecting and tracking of the moving object possible.

Key Words: object tracking; DC coefficients; motion vectors; MPEG-4

运动目标检测与跟踪技术是随着数字视频技术的发展而产生的一个新的研究课题. 视频中的运动目标检测是指把视频帧图像按一定的标准分割成区域, 从视频序列中分割和提取运动对象的一种技术. 目前视频中运动目标检测技术主要分两类: 像素域算法和压缩域算法. 基于像素域的检测算法已经比较成熟, 而压缩域的检测算法尚处于发展阶段. 由于大量的多媒体数据是以压缩形式存储和传输的, 所以研究压缩域中的视频运动目标检测算法有很大的实用价值.

压缩域中的目标检测算法大致分为两类: 一类是利用视频压缩域提供的运动矢量进行检测^[1,2], 一类是利用离散余弦变换(DCT)系数^[3,4]进行检测. 基于运动矢量的检测方法充分利用了压缩视频中运动矢量表示出了当前帧图像与参考帧在对应位置上宏块间的位置偏移这一特点, 而基于 DCT 系数的检测方法则充分利用了运动目标所在的区域即是图像中亮度变化最明显的区域这一特点. 笔者将 DC 系数与运动矢量检测方法相结合, 首先从 MPEG-4 压缩流中获取运动矢量, 由运动矢量标识出可能的运动区域, 在这些区域上通过求取 DC 系数并求互帧差的四阶矩得到真正的运动目标. 该方法的特点在于充分利用了 MPEG-4 码流中蕴含的信息, 避免了仅由运动矢量检测运动目标的误差太大的问题, 也避免了计算全部 DC 系数时的大运算量, 从而可以实时地对视频流中的运动目标进行跟踪.

收稿日期: 2007-05-18

基金项目: 陕西省自然科学基金资助(2006F48)

作者简介: 王 泉(1970-), 男, 副教授.

1 结合运动矢量和 DC 系数的运动目标检测算法

1.1 基于运动矢量的运动区域初步提取

运动目标检测预处理是笔者提出的算法中的重要一步. 它通过对运动矢量的判断, 要标示出视频帧中可能的运动区域. 通过从 MPEG-4 视频流中直接获得运动矢量, 能够快速地标示出 P, B 帧中的运动宏块^[5]. 对每一个帧为宏块(INTRA-MB), 如果该宏块为不编码宏块(SKIP-MB)或宏块的运动矢量小于阈值 T_1 , 那么该宏块属于背景的宏块 S_{bac} , 否则属于运动目标区域的宏块集合 S_{obj} .

然而仅利用运动矢量的提取算法误差较大, 会有噪声的干扰, 影响分割结果, 因此笔者进一步利用 DC 系数差检测运动目标.

1.2 基于 DC 系数的运动区域精确定位

参考文献[3]给出了将运动区域与背景分离的具体过程. 它是在互帧差图像上逐点移动一个小的窗口区域, 然后将每个窗口区域内 4 次互帧差和的平均值与视频图像中背景区域估计的高斯噪声的均方差的平均值进行比较, 以确定前景区域和背景区域. 但是此算法中完成每一 P 帧宏块的帧互差计算需要 15 次乘除法和 23 次加减法(B 帧更多), 一个 352×288 格式的视频序列包含 396 个宏块, 那么就需要 5940 次乘除法和 9108 次加减法. 为了减少计算复杂度, 在由第一步运动矢量模值判定所产生的运动区域的基础上, 仅求取已标识出的可能运动目标的宏块互帧差, 其他宏块的互帧差可认为是零. 假设一个视频中运动目标占整个图像大小的 $1/5$, 那么完成 DC 系数及 4 次矩计算的乘除法为 1188 次, 加减法 1822 次, 从而提高了运算的效率. 具体运动目标的检测过程如下.

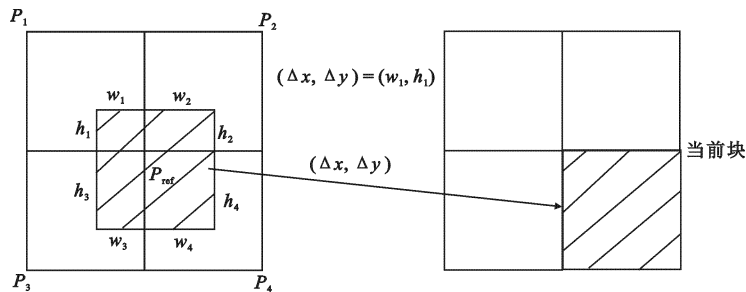


图 1 参考块、当前块和原始块示意图

1.2.1 获取 DC 系数

I 帧直接从数据流中提取 DC 系数^[3]; P 和 B 帧必须利用运动信息来求 DC 系数. 如图 1 所示, P_{ref} 为当前参考块, P_1, \dots, P_4 为与 P_{ref} 重合的 4 个原始的相邻图像块, h_i, w_i ($i = 1, 2, 3, 4$) 分别是 P_{ref} 与 P_i 重合区域的高与宽, C_i 为第 P_i 个块的 DC 系数. P_1, \dots, P_4 的阴影区域部分移动的位移矢量为 $(\Delta x, \Delta y)$. 当前宏块的 DC 系数等于 C_{ref} , 是由参考块在参考帧中占 4 个原始块的比例求得的.

$$P \text{ 帧: } \quad C_{ref} = \sum_{i=1}^4 \frac{h_i w_i}{64} C_i(f) \quad ; \quad (1)$$

$$B \text{ 帧: } \quad C_{ref} = \frac{1}{2} \left(\sum_{i=1}^4 \frac{h_i w_i}{64} C_i(f) + \sum_{i=1}^4 \frac{h_i w_i}{64} C_i(b) \right) \quad . \quad (2)$$

1.2.2 目标提取

设 $d(x, y)$ 为帧间宏块 (x, y) 的亮度差, $m_4(x_0, y_0)$ 为宏块 (x_0, y_0) 的 4 次矩, 则

$$d(x, y) = I_t(x, y) - I_{t-1}(x, y) = C_{diff} \quad , \quad (3)$$

$$m_4(x_0, y_0) = \frac{1}{N} \sum_{(x, y) \in w(x_0, y_0)} (d(x, y) - \bar{d}(x_0, y_0))^4 \quad , \quad (4)$$

其中 C_{diff} 是相邻帧对应宏块的 DC 系数差, $\bar{d}(x_0, y_0)$ 是移动窗口内的抽样平均, 即 $\bar{d}(x_0, y_0) =$

$$\frac{1}{N} \sum_{(x, y) \in w(x_0, y_0)} d(x, y), \text{ 移动窗口的大小为 } N = 9.$$

算法在第一步已标识出的 S_{obj} 区域中逐块计算 $m_4(x_0, y_0)$, 并与一个阈值比较(这个阈值是背景信号的方差), 按照文献[3]给出的方法检测出运动宏块:

$$b(x_0, y_0, t) = \begin{cases} S_{obj} & , \quad m_4(x_0, y_0) > c\sigma_i^2 \\ S_{bac} & , \quad \text{其他,} \end{cases} \quad (5)$$

其中 c 为常量, 最佳值 $c = 81$, σ_i 为背景区域信号的均方差.

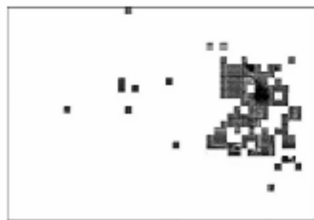
1.3 区域生长

按照上述方法获得的运动目标可能出现中空的现象(图 2(b)). 为了分割出较为完整的运动目标, 笔者使用八邻域的增长法, 遍历图中的每一个宏块, 进行区域生长. (x, y) 为帧中宏块的编号, $\delta(x, y)$ 为 (x, y) 的八邻域, $C_{average}$ 为 (x, y) 八邻域的 DC 系数差绝对值的平均, 生长规则如下:

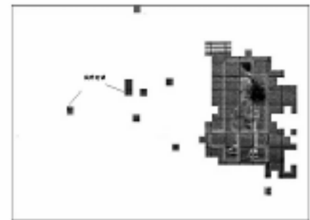
$$\begin{aligned} & \text{if } ((x, y) \in S_{bac} \ \& \ C_{average} \geq T_2) \quad , \\ & \text{then } (x, y) \in S_{obj} \quad , \\ & \text{else } (x, y) \in S_{bac} \quad . \end{aligned} \quad (6)$$



(a) 视频原图



(b) 运动目标区域



(c) 区域生长后的运动区域图

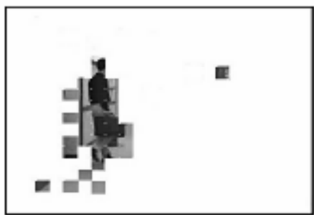
图 2 实验结果($T_1 = 5.0, T_2 = 10.0$)

1.4 局外宏块的剔除

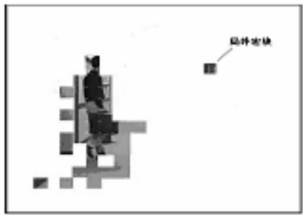
经过上述算法获得的运动目标区域通常会包含有一些本不属于运动目标的宏块, 称之为局外宏块(如图 2(c)和图 3(c)所示). 局外宏块会干扰运动目标质心的计算, 参与运动的估计从而影响估计精度, 因此应该剔除局外宏块. 局外宏块的剔除方法是: 先让所有的宏块参与运动目标区域的估计, 然后由基于距离的孤立点检测算法^[6]剔除这些宏块, 让其余宏块参与运动轨迹的描述.



(a) hall-monitor 中第 11 帧



(b) 运动目标区域



(c) 区域生长后的运动区域图

图 3 实验结果($T_1 = 5.0, T_2 = 5.0$)

1.5 运动轨迹

在获得所有帧中的运动目标区域后, 只需简单地根据运动宏块的位置求得运动区域的质心坐标, 即可勾勒出视频中某一段时间内运动目标在场景中的运动轨迹.

2 实验结果

为了检测上述算法的效果, 采用 Microsoft 的 MPEG-4 参考编解码软件对 test 和 hall-monitor 序列进行编码, test 为 720×480 像素/帧, hall-monitor 序列格式为 CIF 352×288 像素/帧, 视频流帧结构序列为 "IPBBPBBPBBPBBP...". 实验在同一 Pentium4 CPU 2.80 GHz, 512RAM 的 PC 机上对同样的视频流进行处理. 图 2 的阈值为 $T_1 = 5.0, T_2 = 10.0$; 图 3 为 $T_1 = 5.0, T_2 = 5.0$.

时间分析见表 1.

表 1 几种算法的比较

算法	全解码	仅采用互帧差四阶矩的算法	笔者提出的算法
Demo	2.031	0.974	0.625
Hall-monitor	1.265	0.638	0.421

实验结果表明,笔者提出的算法与仅采用互帧差四阶矩的算法相比,效率要高.区域生长后的运动区域要比仅通过 DC 系数差检测算法得到的运动区域更加完整.整个算法只需要对 MPEG-4 压缩流部分解码,时间仅是全部解码的 1/3,同时比传统的对整帧的 DC 系数运算时间节省了 25%~30%.此算法的效率主要决定于运动目标在帧中的比例,比例越小,效率越高.

3 结束语

笔者将运动矢量和 DC 系数相结合,采用区域生长和局外宏块剔除的方法进一步提高了运动目标检测的精度.由于所用的目标检测算法是基于 MPEG-4 压缩域的算法,对 I 帧、P 帧和 B 帧的处理单位都是 16×16 的宏块.因此,目标跟踪的矩形框往往不能紧凑地套在运动目标上,考虑到这是一个侧重于取得目标运行轨迹的非基于先验知识的目标跟踪算法,这些误差对目标运动轨迹的取得影响不大.

参考文献:

- [1] Yoneyama A, Nakajima Y, Yanagihara H, et al. Moving Object Detection and Identification from MPEG Coded Data [C]//International Conference on Image Processing; Vol 2. Japan: IEEE, 1999: 934-938.
- [2] Liang Zhonghua, Wang Ping, Tan Zheng. Moving Object Detection from MPEG Bit Stream[C]//Proceedings of 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing; Vol 3. Hong Kong: IEEE, 2003: 701-704.
- [3] Zeng Wei, Gao Wen, Zhao Debin. Automatic Moving Object Extraction in MPEG Video[C]//Proceedings of 2003 IEEE International Symposium on Circuits and Systems; Vol 2. Thailand: IEEE, 2003: 524-527.
- [4] Chen Tianhang, Liu Shaohui, Yao Hongxun, et al. Robust Video Watermarking Based on DC Coefficients of Selected Blocks[C]//IEEE the Fourth International Conference on Machine Learning and Cybernetics; Vol 9. Guangzhou: IEEE, 2005: 5 273-5 278.
- [5] 陈杉,于鸿洋.基于 MPEG 压缩域的运动对象检测方法[J].信号处理,2004,20(6):628-631.
- [6] 陆声链,林士敏.基于距离的孤立点检测及其应用[J].计算机与数字工程,2004,32(5):94-97.

(编辑:郭 华)

简 讯

2007 年 9 月 3 日~6 日,由我校承办的 2007 全国大学生电子设计竞赛陕西地区比赛在我校举行,来自陕西地区 36 所高校的 427 支参赛队参加了比赛,我校选派了 35 个代表队参加.

摘自《西电情况》2007.9.18