

一种在 MPEG 压缩域中检测镜头扫换的方法

刘 阳, 吴志美

(中国科学院软件研究所, 北京 100080)

摘 要: 当前对镜头切变已能够提供有效的检测方法,但渐变的检测一直是一个难题,也是当前视频分割新的研究热点. 本文提出一种在 MPEG 压缩域中检测扫换(Wipe)的方法,它通过分析 B 帧和 P 帧中宏块编码类型的时空分布特点来进行检测,利用哈夫变换判断变化区域连续性,最后使用 DC 图像的直方图分析进一步验证. 它不仅能定位扫换的起止帧,而且能得到扫换方向等特性. 本方法需要比较小的运算量,同时对运动和其他镜头切换有比较好的区分性.

关键词: 视频分割; 扫换检测; MPEG; 宏块编码类型; 哈夫变换

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2002) 05-0741-04

A Wipe Transition Detection Method In MPEG Compressed Domain

LIU Yang, WU Zhi-mei

(Institute of Software, Chinese Academy of Sciences, Beijing 100080, China)

Abstract: Now many video segmentation algorithms can detect cuts efficiently, but the detection of gradual transitions is left as a challenge, which is becoming a new active research region. In this paper, a new wipe transition detection method in MPEG compressed domain is presented based on the spatio-temporal distribution of the macroblock coding types in B frames and P frames. Hough transform is utilized to analyze the continuity of change regions. At last, the histogram analysis of DC-image is used to reconfirm the detection. This method can get not only the beginning and end frame number of the wipe transition, but also the transition patterns, such as the wipe direction. It requires less calculations and can distinguish wipes from other transition types and motions in some cases.

Key words: video segmentation; wipe detection; MPEG; macroblock coding type; Hough transform

1 引言

视频分割是视频分析中最基本也最重要的一项内容,它的基本任务是根据相邻帧的相关性将视频序列分割成相对独立的基本单元——镜头(Shot). 根据镜头连接方式不同可将镜头切换分为切变(Cut)和渐变(Gradual Transition)两类. 常见的渐变方式主要有三种:淡入淡出、叠化和扫换(Wipe,有时译为划变或刮划). 也有学者认为淡入淡出是叠化的一种特殊形式. 文[1]给出的扫换定义为:一般指在镜头切换中有一条线扫过屏幕,使得当前镜头慢慢被下一个镜头覆盖. 图1给出了一个简单的扫换例子,随着现代视频编辑技术的发展,出现了一些更加复杂的扫换方式.

早期的视频分割方法主要针对切变,其中有一些可检测渐变,但只能判定是否存在渐变,而不能区分渐变类型和准确定位起止点. 检测渐变比检测切变困难,当前研究热点日趋集中到渐变检测上来. 文[2]对淡入淡出和叠化的检测方法作了综述,但是对扫换的检测讨论地相对比较少. 当前提出的扫换检测方法主要有文[3~6],它们均基于像素分析,需要比较大的解码运算量,而且大都对运动敏感. 最近文[7]提出在

MPEG压缩域中检测扫换,该法不需完全解码 MPEG,运算量大大降低,但很难将扫换和运动区分开,也不能得到扫换模式和准确定位扫换起止点,同时忽略了 P 帧中的有用信息. 本文提出一种在 MPEG 压缩域中进行扫换检测的方法,它利用 B 帧和 P 帧中宏块编码类型的时空分布特点进行检测,最后使用 DC 图像的直方图分析进一步验证检测结果.

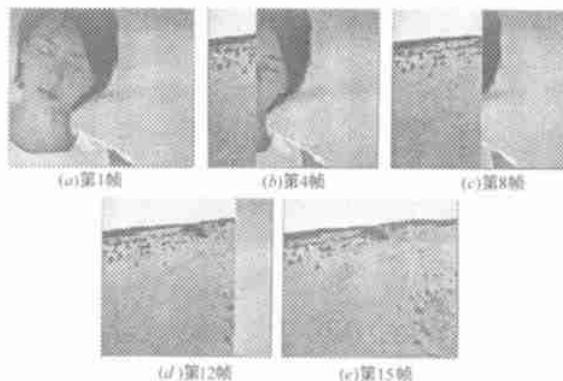


图1 一个简单的扫换例子(从(a)到(e))

收稿日期:2001-08-23;修回日期:2001-11-28

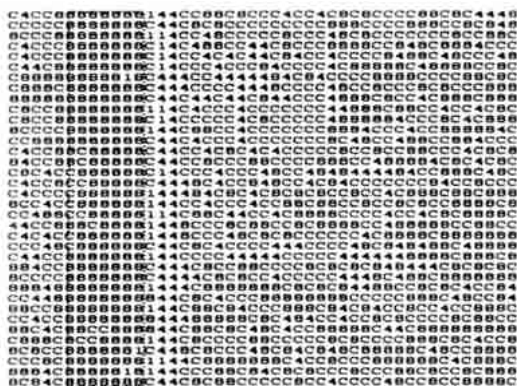
基金项目:国家重点基础研究发展规划(973 计划)(No. G1998030407);北京科委资助课题(H011710010123)

2 扫描中宏块编码类型的时空分布特性分析

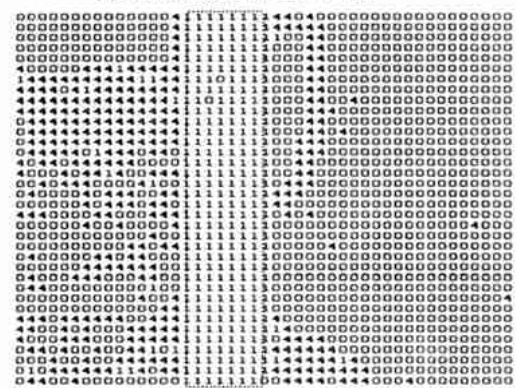
2.1 扫描中宏块编码类型的时空分布特性

MPEG流中存在的运动编码信息在某种程度上反映了帧之间的内容相关性. 扫描过程中当前帧的某个特定区域被下一个镜头替换, 这必然会影响到该区域内的宏块编码方式. 对 P 帧变化区域宏块, 在前面参考帧中搜索不到匹配块, 所以采用内部编码方式; 对 B 帧变化区域宏块, 在后面参考帧中能够搜索到匹配块, 所以采用向后预测编码. 为叙述方便, 下文将 P 帧中帧内编码的宏块和 B 帧中向后预测编码的宏块称为备检宏块. 下面以典型扫描例如水平或对角线方式的直线滑动 (slide) 为例论述, 在本文最后将讨论复杂扫描模式检测.

在空间分布特性上, 相邻帧变化区域是连续而不是间断的, 而且该区域一般以帧图像边缘为边界, 相应的备检宏块的空间分布也会有这样的特性. 图 2 给出了图 1 的第 4B 帧和第 8P 帧的对应宏块编码类型的空间分布图, 可以看到 8P 帧中采用内部编码的宏块构成一个连续的近似矩形区域, 而且该区域以图像的边缘为上下边界.



(a) 第 4B 帧的宏块编码空间分布图



(b) 第 8P 帧的宏块编码空间分布图

(其中 1 表示采用帧内编码, 8 表示采用向后预测编码)

图 2 B 帧和 P 帧的宏块编码空间分布

在时间分布特性上, 在扫描中连续帧的变化区域也应该是连续的, 前一帧的变化区域和后一帧的变化区域的像素坐标应保持一定连续性而不是间断的, 而且将变化区域累加应覆盖整个帧图像.

2.2 与其他类型切换方式及运动的区分性

其它形式的镜头切换不存在以上的备检宏块的时空分布特性, 切变不存在时间分布连续性, 淡入淡出和叠化的变化区域一般存在于整个帧内, 而且将连续时间上的变化区域累加是没有意义的.

各种视频分割算法必须面临的一个公共难题是如何最大限度的将对象运动和摄像机运动区分. 一般情况下, 对象运动是在帧图像内部进行, 而且连续帧的变化区域的累加一般不会正好覆盖整个帧图像. 对于大部分的摄像机运动, 相邻帧的变化会涉及整个帧图像, 而扫描中相邻帧的变化区域只能是帧图像的一个子集.

3 扫描检测算法

在具体实现中, 应首先分析扫描的空间分布特性, 如符合则继续分析其时间分布特性.

3.1 相邻帧变化区域的空间分布特性分析

根据以上讨论, 具有扫描特性的相邻帧变化区域应同时满足以下条件:

由备检宏块组成

构成一个具有某种规则几何特性的连续区域

连续区域是以帧图像的边缘为边界

发生变化区域的比例应该满足以下不等式:

$$T_{sl} \leq F_c / F_t \leq T_{sh} \quad (1)$$

其中 F_c 表示当前帧的变化区域尺寸, F_t 表示整个帧图像尺寸, T_{sl} 表示变化区域的最小比例 (例如对图 1 的扫描可限制变化区域至少包括两条连续垂线代表的宏块), T_{sh} 表示变化区域的最大比例, 根据一般扫描所经历的帧长度并结合试验经验, 可以限制 T_{sl} 和 T_{sh} 分别为 0.02 和 0.20.

3.2 哈夫变换与变化区域连续性检测

为得到变化区域的空间特性还需检测其连续性. 将宏块作为分析的基本单位, 可用一个点代表一个宏块. 这样需首先检测某些特定点是否共线, 然后判定相临曲线是否构成连续区域. 例如对图 1 所示的水平扫描, 欲分析备检宏块是否可构成与帧图像等高的矩形区域, 可简单统计备检宏块在对应垂线上的比例来判断是否可形成一条直线, 然后将相邻直线代表的宏块组成一个矩形区域. 但扫描模式是多样的, 需检测任意的规则几何图形, 根据哈夫 (Hough) 变换的基本原理设计了一种检测任意规则曲线形式变化区域的方法. 它可检测任何形式的曲线, 下面以检测任意方向的直线为例说明.

哈夫变换将图像空间中的直线转换到参数空间进行描述, 对所有可能落在直线边界上的点进行统计, 根据统计数据确定该点属于该直线的程度. 因为存在类似于图 1 的垂线, 宜采用极坐标方式. 设直线方程为: $x \cos \theta + y \sin \theta = \rho$, 该直线的任一点 (x, y) 可在参数空间 (θ, ρ) 中描述为一条正弦曲线, 而直线上的所有点都将通过同一点 (θ, ρ) . 图 3 所示直角坐标系上直线的三个点, 在参数空间的描述为三条相交的正弦曲线. 通过检测 (θ, ρ) 的累积分布求得具有最大值的 (θ, ρ) 点, 进而求得参数方程, 得到该直线的参数描述. 文献 [8] 给出了哈夫变换的具体实现算法.

据此可求出该条直线所经过的宏块个数 N_c , 同时能统计

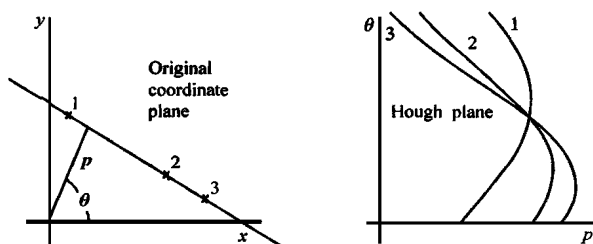


图 3 哈夫变换的基本原理图

出该直线上实际备检宏块个数 N_d , 可使用以下不等式判定备检宏块能否组成连续直线:

$$N_d / N_c \geq T_{hl} \quad (2)$$

其中 T_{hl} 限制直线上备检宏块应占的最小比例阈值, 一般情况下其取值应大于 0.65. 在实际应用时, 为减小计算量, 需要对上面的和作区间取值并取整^[8], 计算 N_c 时, 需要对宏块的坐标作适当的近似取整. 最后将相邻直线构成一个连续区域, 并记录每帧变化区域的坐标.

3.3 连续帧变化区域时间分布特性分析

在空间特性满足扫换特性前提下, 需要检测连续帧的变化区域是否有以下的时间特性: 一方面, 连续帧的变化区域坐标应保持一定连续性而不间断, 相对单调递增或递减; 另一方面, 连续帧的变化区域的叠加应覆盖整个帧图像, 可使用以下不等式判断:

$$T_{il} \leq \sum_{i=b}^e R_i \leq T_{ih} \quad (3)$$

其中 R_i 表示第 i 帧的变化区域比例, 要求第 i 帧的坐标与第 $i-1$ 帧的坐标保持连续, b 和 e 分别表示检测得到的扫换开始和结束帧号(可根据变化区域起始和结束点得到), T_{il} 和 T_{ih} 分别是预先设定的最小和最大阈值, 它们的取值会对检测结果有比较大的影响. 根据样本分析和试验经验, T_{il} 应该大于 0.75, T_{ih} 应该小于 1.25. 这样可克服其他检测方法采用的根据扫换经历帧长度进行检测的缺陷.

在以上检测中可分析得到扫换特性, 例如可根据哈夫变换求得曲线参数和随时间变化的轨迹来确定扫换模式, 可根据每帧变化率和经历的帧长度得到扫换速度.

3.4 利用 DC 图像的直方图分析进一步验证检测

扫换的开始和结束帧属于不同的镜头, 可通过使用检测切变的方法来分析开始帧和结束帧的差异性进一步验证检测结果. 采用压缩域的 DC 图像直方图分析这一简单有效的方法验证检测结果. 因为 I 帧能快速形成 DC 图像, 可将验证的开始帧和结束帧扩展到 $[b, e]$, b 是离 b 帧最近的 I 帧号, e 是离 e 帧最近的 I 帧号.

3.5 复杂扫换模式的检测

扫换模式有多样性, 仅 Adobe Premiere 5.0 可编辑的就有 50 多种, 目前难以定义统一的检测模型, 因而扫换被认为是最难检测的镜头切换方式^[5], 当前的检测方法大都检测其中一类. 基于以上的基本原理, 调整相应空间特性分析模型即可检测大部分的复杂扫换模式. 例如在图 4(a) 和 (b) 所示的 iris-around 扫换中, 边界为一个以图像中心为圆心的椭圆, 且

椭圆区域随扫换过程逐渐变大.

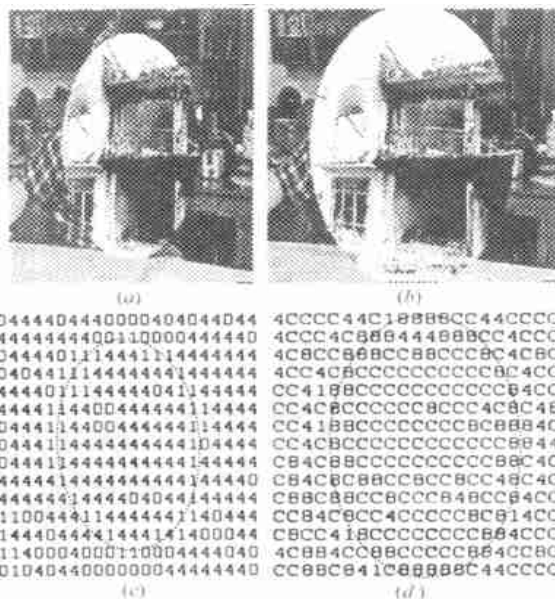


图 4 iris-around 方式扫换与对应的宏块编码空间分布图. (a)、(b) 表示扫换中的典型帧, (c)、(d) 分别表示 P 帧和 B 帧的宏块编码类型的空间分布图. (其中 1 表示采用帧内编码, 8 表示采用向后预测)

如图 4(c) 和 (d) 所示, 备检宏块在空间上组成了与扫换特性一致的椭圆形闭合区域, 利用上面的哈夫变换可分析备检宏块是否相应地构成一个椭圆, 时间分布特性分析与上文论述原理类似. 对更加复杂的扫换模式, 备检宏块的空间特性也会变得复杂, 我们可将备检宏块作为目标像素, 利用图形学中的边界检测方法^[8]检测备检宏块组成的闭合区间.

4 试验结果

我们收集了两类视频片段. 一类取自 MTV 节目, 另一类取自电影《星球大战》(Star Wars), 该影片作为具有代表性的视频分割的测试集被广泛应用, 它们包含了大部分典型的扫换方式. 使用 DVMPEG V5.0 ($TM5, M=3, N=15$) 编码. 试验结果见表 1.

表 1 算法检测结果

视频集名	实际数	正确检测数	误检数	检出率(%)	精度(%)
MTV	9	7	1	77.8	87.5
星球大战	12	9	1	75	90
合计	21	16	2	76.2	88.9

大部分漏检的扫换是本方法不适合的复杂的扫换方式, 另一个漏检是因为扫换边界模糊引起的, 同时将一个包含快速飞行的飞机误检为水平扫换. 在试验中各个阈值按前文所述的参考值统一取值, 可使用自适应阈值的方法提高检测性能^[9]. 同其它在压缩域中分析的算法一样, 试验中我们发现算法的性能会受到 MPEG 编码方案的影响, 文献[7]对此进行详细的讨论. 另外试验表明 P 帧的信息比 B 帧的信息更有效, 在某些情况下仅可用 P 帧的内部编码宏块快速检测.

5 结论

本文提出一种在 MPEG 压缩域进行扫换检测的方法,它不仅能确定扫换的开始点和结束点,而且能检测扫换模式。本方法对运动和其他镜头切换有比较好的区分性,同时因为在 MPEG 压缩域中进行检测,不需要完全解码,要求比较少的运算量。当前面临的主要问题是:如何对所有的扫换模式定义统一的扫换模型。

参考文献:

- [1] D Bordwell, K Thompson. Film Art: An Introduction (2nd Edition) [M]. New York: Random House, 1986.
- [2] R Lienhart. Comparison of automatic shot boundary detection algorithms [A]. SPIE Conf. On Storage and Retrieval for Image and Video Databases [C]. San Jose, CA: SPIE3656, 1999. 290 - 301.
- [3] H Yu, W Wolf. A multi-resolution video segmentation scheme for wipe transition identification [A]. Proc Of ICASSP 1998 [C]. Washington: 1998. 2965 - 2968.
- [4] A M Alattar. Wipe scene change detector for use with video compression algorithms and MPEG7 [J]. IEEE Trans on Consumer Electronics, 1998: 44(1): 43 - 51.
- [5] J Nam, A H Tewfik. Wipe transition detection using polynomial interpolation [A]. SPIE Conf On Storage and Retrieval for Media Databases [C]. San Jose, CA: SPIE 4315, 2001. 231 - 241.
- [6] 金红, 周源华. 一种基于模型的扫换检测方法 [J]. 软件学报, 2001, 12(3): 468 - 474.
- [7] S C Pei, Y Z Chou. Efficient and effective wipe detection in MPEG compressed video based on the macroblock information [A]. Inter. Conf On Image Processing 2000 [C]. Vancouver, BC, 2000. 953 - 956.
- [8] 章毓晋. 图象分割 [M]. 北京: 科学出版社, 2001.
- [9] Y Yusoff, W Christmas, J Kittler. Video shot cut detection using adaptive thresholding [A]. British Machine Vision Conference 2000 [C]. Bristol: 2000. 340 - 349.

作者简介:



刘 阳 男, 1975 年 6 月出生于山东省, 中国科学院软件研究所博士生, 主要研究方向是视频分析和多媒体通信等。



吴志美 男, 1942 年 12 月出生于江苏省, 中国科学院软件研究所多媒体通信和网络工程研究中心主任研究员, 博士生导师; 主要研究领域有多媒体通信, 网络融合与互通, 协议测试与验证。

(上接第 740 页)

3 异动信号报警器的特点

(1) 利用网络功能设计异动信号报警器, 可以提高报警的综合性能, 系统简单, 成本低, 适用范围广, 精度高, 可在网上报警。

(2) 有利于把复杂、分散的工业自动控制报警向小型化、简易化、多功能报警方向扩展, 也可进行中断控制, 智能控制, 过程控制和远程控制。

(3) 异动信号报警器突出了实用的特点, 有广泛的应用前景。

4 小结

结合利用通讯网络功能设计的异动信号报警器有综合功能强, 有明显的实用性, 成本低等特点, 但也会由于网络中存在病毒问题影响到报警, 会使报警信号失真, 因此主动防范网络病毒问题是今后解决的一个问题。