

# 数字视频特技镜头转换检测算法的分析\*

白雪生, 张子银, 徐光祐, 杨士强, 史元春

(清华大学 计算机科学与技术系 人机交互与媒体集成研究所,北京 100084)

E-mail: xgy-dcs@mail.tsinghua.edu.cn

http://media.cs.tsinghua.edu.cn

**摘要:** 随着数字设备在视频制作中的广泛应用,数字特技(DVE)也越来越多地被用于镜头转换,这就要求视频检索能有效地对 DVE 镜头转换进行检测.通过对 DVE 镜头转换的模型和特性的分析,指出了目前基于二维时空图处理算法的局限性,进而提出了一种基于转换曲面的 DVE 检测算法以解决这个问题.并对在实际 DVE 视频上的镜头检测结果进行了分析,取得了很好的结果.

**关键词:** 镜头检测;镜头切变;数字视频特技;时空图分析

**中图分类号:** TP391 **文献标识码:** A

随着数字化技术的推广,数字设备在视频制作中越来越广泛的应用,数字视频特技(digital video effect,简称 DVE)也被越来越多地用于镜头间的转换.而在基于内容视频检索中,镜头的分割和检测是视频数据结构化的基础<sup>[1,2]</sup>,这就提出了对 DVE 镜头转换的检测问题.

与镜头切变(cut)和渐变(dissolve)不同,数字特技(DVE)转换是通过后一镜头在图像域的不断扩张来完成的,故也称 wipe,不同的扩张模式对应于不同的 wipe 模板(SMPTE 定义了一系列与特定号码相对应的转换模板).目前,针对一些比较常见的 wipe(如水平、垂直、对角等)已提出了一些检测算法,其绝大部分都是建立在对视频通过特定采样方案得到的二维时空图进行处理的基础上<sup>[3-5]</sup>.但随着数字编辑硬件和软件功能越来越强,许多产品支持的(DVE)转换更加丰富、多样,上述基于二维时空图处理的算法已不能满足 DVE 镜头转换检测的需要.

在本文中,我们将通过对 DVE 镜头转换的模型和特性的分析,指出二维时空图处理算法的局限性——理论上对任一种采样方案(无论是单个的连续采样方案还是连续采样方案的组合)都可以设计出一种 wipe 模板,使得基于此采样方案的镜头检测不能将该 wipe 与镜头切变有效分开.在此基础上,我们提出了一种基于转换曲面的 DVE 检测算法,并对该算法的检测性能进行了分析.

本文第 1 节首先阐述 DVE 转换的模型和特点.第 2 节对二维时空图的定义和构成进行介绍,并对 DVE 检测算法的局限性进行分析.第 3 节给出了基于转换曲面的 DVE 检测算法,并在第 4 节给出了实验结果和分析.

## 1 数字视频特技转换的模型和特点

对 DVE 镜头转换分析可知,尽管不同的 DVE 对应于不同的模板,但它们都具有一个共同点,即转换过程中后一视频信号在帧图像中占据的区域不断扩张,面积也从 0 增加到整幅图像.根据这个特点,我们对 DVE 提出如

\* 收稿日期: 2000-11-13; 修改日期: 2001-02-26

基金项目: 国家 863 高科技发展计划资助项目(863-306-ZT04-02-01;863-306-QN99-01);国家教育部博士点基金资助项目(99000307)

作者简介: 白雪生(1972 - ),男,辽宁辽阳人,博士,讲师,主要研究领域为基于内容检索,计算机视觉;张子银(1970 - ),男,黑龙江牡丹江人,硕士生,主要研究领域为基于内容检索;徐光祐(1940 - ),男,上海人,教授,博士生导师,主要研究领域为计算机视觉,多媒体技术;杨士强(1952 - ),男,山东高唐人,教授,博士生导师,主要研究领域为多媒体技术;史元春(1967 - ),女,河南汤阴人,博士,副教授,主要研究领域为计算机支持的协同工作,多媒体技术.

下的区域扩张模型:

记  $W_N$  为一个长度为  $N$  的 wipe 过程,它由一个长度为  $N$  的模板序列  $W_N = \{M_i, i = 0, 1, \dots, N-1\}$  组成.其中  $M_i$  为定义在二维图像平面上的点集,满足  $M_0 = S, M_{N-1} = \emptyset$ , 即  $M_0$  为全集,  $M_{N-1}$  为空集.相邻  $M_i$  间满足关系

$$M_{i+1} \subseteq M_i. \quad (1)$$

根据此模型,可知 DVE 转换具有以下特点:

DVE 转换对后一视频来说是渐进扩张性的,即图像空间上的某一像素一旦在其内容改变(为后一视频内容)后就一直播放后一视频的内容.

相对于 DVE 转换的区域扩张速度而言,转换前后镜头内的运动较小.

## 2 对二维时空图算法的分析

### 2.1 二维时空图的定义与获取

目前,多数 DVE 检测算法都是针对简单、常见的几种 wipe,通过对视频流采样得到的二维时空图进行处理来实现的.在此,我们首先对二维时空图形式化描述如下:

令  $f_v(u, v, t)$  为视频流  $V$  在时刻  $t$  像素  $(u, v)$  处的值,则视频流  $V$  可表示为

$$V = \{f_v(u, v, t)\}, \quad u, v, t \in \{0, 1, 2, \dots\}, \quad (2)$$

记  $f_v(u, v, t)$  所在帧对应 DC(direct current, 直流)图像像素  $(u, v)$  处的值为  $f_{\text{Thumbnail}}(u, v, t)$ , 则相应 DC 图像序列

$$V_{\text{Thumbnail}} = \{f_{\text{Thumbnail}}(u, v, t)\}, \quad u, v, t \in \{0, 1, 2, \dots\}, \quad (3)$$

则其时空图  $VR$  可表示为

$$VR = \{f_{VR}(z, t)\} = \{f_{\text{Thumbnail}}(u(z), v(z), t)\}, \quad (4)$$

其中  $u(z)$  和  $v(z)$  分别为  $z$  的一维函数.

从上述定义可以看出,时空图构成了一幅从 DC 图像序列中采样形成的二维图像.对同一 DC 图像,根据函数  $u(z)$  和  $v(z)$  的不同,可得到不同的时空图.图 1 给出了一些常用的简单采样方案.图 1 中(a)~(d)分别对应水平、垂直、对角和双对角的采样方案.选择以左上角为原点、 $u$  轴向右、 $v$  轴向下的图像坐标系,令  $h$  和  $w$  分别为图像的高和宽,则在图像沿选定的水平线采样的图 1(a)中,其函数可用  $u(z)=z, v(z)=y_0$  来表示;图 1(b)则沿选定的垂直线采样,其函数  $u(z)=x_0, v(z)=z$ ;在图 1(c)中,图像沿单对角线采样,其函数  $u(z)=z, v(z)=h^*z/w$ ;图 1(d)中,图像沿双对角线采样,其函数可分段表示为  $u(z)=z, v(z)=h^*z/w, z < w; u(z)=2^*w-z, v(z)=h^*(z-w)/w, z > w$ .

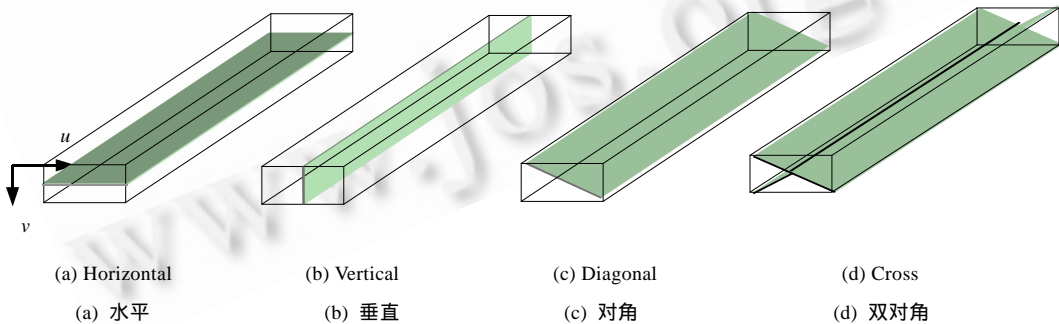


Fig.1 Sampling patterns often used

图 1 常用的时空图采样方案

不同镜头转换在时空图上表现形式不同.令  $t$  轴为时空图横轴, $z$  轴为纵轴,则同列像素采自同帧图像.在镜头内部,相邻两列间像素相差不大;当发生镜头切变时,切变前后两列分别取自不同镜头,因而会出现明显的垂直边缘.而 wipe 发生期间,同帧图像上的像素分别取自前后两个镜头,一般会有明显的差别.若采样方案能保证 wipe 变化在  $z$  轴上连续,则此差别在时空图上体现为边缘,其形状取决于 wipe 类型和时空图采样方案.表 1 给出了 5 种常见 wipe(水平、垂直、扩张、吸收、对角)在上述 4 种采样方案下的时空图边缘形状.

Table 1 Shape of spatial-temporal edge under common sampling pattern and wipe mode

表 1 常见 wipe 和采样方案下的时空图边缘形状

Shape of spatial-temporal edge		Wipe Mode				
		Horizontal	Vertical	Expand	Absorb	Diagonal
Sampling pattern	Horizontal					
	Vertical					
	Diagonal					
	Cross					

时空图边缘类型, 采样方案, 水平, 垂直, 对角, 双对角, 扩张, 吸收.

2.2 基于二维时空图边缘检测的算法及其局限性分析

由上述分析可知,镜头切变对应的边缘在各种采样方案时空图上都表现为垂直的边缘;而对较为常见的规则 wipe,在连续采样方案——即采样函数 $(u(z),v(z))$ 构成一条对  $z$  连续的曲线——如前 3 种(水平、垂直、对角)采样方案下,这些边缘点构成了一条将时空图分成两部分的连续非垂直边缘.据此可通过判断边缘垂直性来对镜头切变和上述规则 wipe 进行检测和区分.

但亦可看出,对这 3 种采样方案都存在一种 wipe,其时空图上的边缘为垂直边缘,不易与镜头切变区分开.一种可有效区分镜头切变与上述 wipe 的方法是采用不同连续采样方案的时空图联合判断,即只有在所有采样方案时空图上均表现为垂直边缘的镜头转换才为切变,在任一采样方案时空图上表现出非垂直边缘的镜头转换都可视为 wipe.表 1 中的双对角采样方案就是这样一种时空图组合,可以看出该方案可较好地地区分镜头切变与常见的 wipe.

但此时时空图联合判断方法对更普遍和灵活的 DVE 转换很难应用.从理论上讲,采用二维时空图方法,对任一采样方案(无论是单个连续采样方案还是连续采样方案的组合)都可设计一种 wipe,使其在时空图上边缘为垂直边缘.实际上,对采样方案组合 $\{S_i=(u_{iz},v_{iz}),z=1,2,\dots,n_i\},i=1,2,\dots,m\}$ ,定义点集 $S_p = \bigcup_i S_i$ .只要 $S_p$ 不覆盖整幅图像,考察长度  $N$  的 wipe 过程 $W_N = \{M_i, i=0,1,\dots,N-1\}$ , $M_i$ 为定义在二维图像平面上的点集,满足 $M_0 = S$ , $M_{N-1} = \emptyset$ , $M_{i+1} \subseteq M_i$ ,且对某一 $1 < k < N$ ,有

$$M_k \cap S_p = \emptyset, \tag{5}$$

$$S_p \subseteq M_{k-1}, \tag{6}$$

即在  $k$  处所有采样点同时从前一视频信号变为后一视频信号.显然,在上述采样方案构成的时空图上,此 wipe 过程中的  $k$  处将出现垂直边缘.

3 基于边界曲面的 DVE 检测算法

上述问题是由时空图在二维图像上进行一维采样的本质所决定的.为了解决这个问题,我们提出了如下的基于 DC 图像序列上边界曲面检测的通用 DVE 转换检测思想:

计算 DC 图像序列  $f_{Thumbnail}(u, v, t)$  的亮度时间差分图像:

$$d(u, v, t) = |f_{\text{Thumbnail}}(u, v, t) - f_{\text{Thumbnail}}(u, v, t-1)|, \quad (7)$$

再计算以当前时刻  $t$  为中心、长度为  $L$  的时间窗内的中值  $\text{median}_L(u, v, t)$ , 并根据如下准则来判断一个像素是否是边缘点:

$$b(u, v, t) = \begin{cases} 1 & d(u, v, t) > K * \text{median}_L(u, v, t) \\ 0 & \text{其他} \end{cases}. \quad (8)$$

$K$  为比值阈值. 据此得到的  $b(u, v, t)$  构成二值图像序列. 在其对应的  $(u, v, t)$  空间中, 镜头切变及 wipe 导致的值为 1 的点构成二维曲面, 曲面延伸至 4 个边界平面  $u=0, v=0, u=w$  和  $v=h$ , 是前后镜头转换的边界. 镜头切变对应边界是一个与时间轴  $t$  垂直的平面, 其他平面与曲面则对应于各式各样的 wipe. 时空图上的边缘实际上是根据采样方案对此二维曲面的采样, 如图 2 所示.

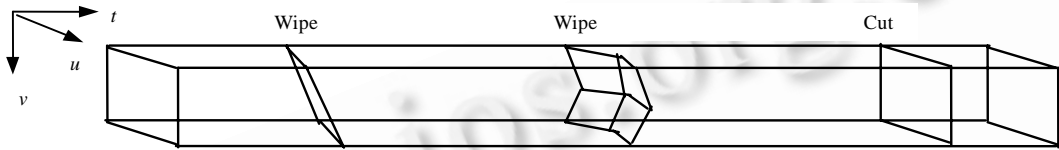


Fig.2 2-D surface corresponding to cut and wipe in  $(u, v, t)$  space of DC image sequence

图 2 Wipe 与镜头切变在 DC 图像序列  $(u, v, t)$  空间中对应的二维曲面

对  $b(u, v, t)$ , 分别计算如下度量

$$m_{\text{cut}}(t) = \sum_{u, v} b(u, v, t), \quad (9)$$

$$m_{\text{wipe}}(t) = \sum_{k=0}^{k < W} \sum_{u, v} b(u, v, t+k), \quad (10)$$

其中  $W$  为时间窗长度, 一般选择为 wipe 的最大转换时间. 显然,  $m_{\text{cut}}(t)$  和  $m_{\text{wipe}}(t)$  分别为当前帧图像中和从当前开始时间窗中的边缘点数. 切变发生时,  $m_{\text{cut}}(t)$  上出现单点峰值,  $m_{\text{wipe}}(t)$  上则出现等于时间窗长度的峰值序列; 而当 wipe 发生时,  $m_{\text{wipe}}(t)$  上将出现峰值高台. 针对此特点, 我们采用了计算如下定义的次大比值序列

$$R_{\text{cut}}(t) = \frac{m_{\text{cut}}(t)}{2\text{nd max}_{t' \in [t-5, t+5]} \{m_{\text{cut}}(t')\}}, \quad (11)$$

并对其进行阈值检测的方法来检测镜头切变的出现. 在无镜头切变的情形下, 再对  $m_{\text{wipe}}(t)$  计算边缘点占整幅 DC 图像的面积比值序列

$$R_{\text{wipe}}(t) = \frac{m_{\text{wipe}}(t)}{h * w}, \quad (12)$$

并对其进行阈值检测来检测 DVE 镜头转换的出现. 这样就可完成对镜头切变和 DVE 镜头转换的检测和区分.

#### 4 实验结果与分析

利用本文中提出的算法, 我们设计并实现了对镜头切变和 DVE 镜头转换进行检测和区分的系统, 并对由数字视频编辑软件 (Adobe Premiere)、硬件 (V-Star 卡) 得到的含有视频切变及 DVE 镜头转换的视频节目进行了测试. 实现中, 边缘点检测时间窗长度设置为 10 帧, 比值  $K$  根据稳健分析理论设置为  $3 * 1.4826 = 4.4478$  (无镜头转换情况下同一像素的各帧灰度值可用高斯分布来描述, 其方差可用  $\sigma = 1.4826 * \text{median}_L(u, v, t)$  来近似估计, 我们选择  $3\sigma$  作为边缘点检测阈值);  $m_{\text{wipe}}(t)$  计算时间窗长度  $W$  设为 50 帧. 次大比值序列  $R_{\text{cut}}(t)$  的阈值选择为 3, 而面积比值序列  $R_{\text{wipe}}(t)$  的阈值则选择为 50%. 对几种 DVE 镜头转换的检测结果如图 3 所示.

图 3 给出了对两个镜头切变和 4 个 DVE 镜头转换的处理结果. 每个处理结果包括 5 幅图像, 分别对应于检出时的当前帧图像、帧间差图像 (由于帧间差一般较小, 肉眼不易观察到)、边缘点图像以及前后两镜头的代表帧. 最上面一行显示的是对两个镜头切变的检测结果, 下面两行从左到右依次为角形对角 wipe 以及块形、扇形

和星形 3 种不规则形状 DVE 转换的处理结果.从图中可以看出,当前后两镜头某些对应像素值较为接近时,此时即使发生转换,对应像素点也不作为边缘点检出.为此,我们将  $R_{\text{wipe}}(t)$  阈值选择为较低的 50%,较好地避免了这些因素的影响.

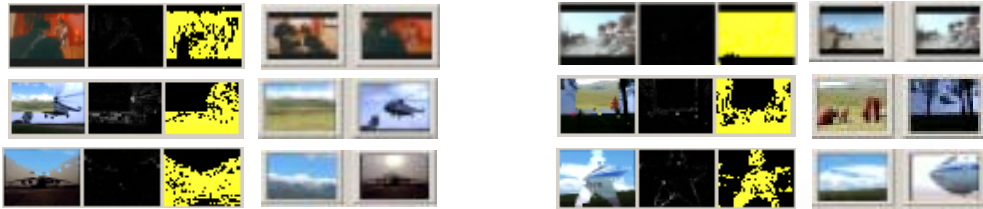


Fig.3 Detection results on cut and DVE shot transitions

图 3 对镜头切变和 DVE 镜头转换的检测结果

我们采用上述方法对“大话西游”、“逃学威龙”等影视和通过视频特技编辑软件生成的视频进行了检测.实验结果表明,对镜头切变,该方法在 3 149 个切变中漏检 15 个(主要发生在前后镜头背景接近且存在运动的情况下),误检 8 个(发生在打斗、剧烈运动与爆炸等情况下);而对 DVE 镜头转换,该方法检出了全部 31 个转换,发生了一处误检.分析其原因,是在视频中存在较大运动时,算法检出了较多的边缘点,超过了  $R_{\text{wipe}}(t)$  50% 的阈值所致.但考虑到前后镜头对应像素值可能比较接近的因素,单纯提高阈值在避免误检的同时将会导致漏检,因此并不是一个好办法.我们设计并提供了一个人机交互的检测结果编辑界面来解决此问题.从整体来说,该算法对镜头切变检出率为 99.5%,误检率为 0.25%,对 DVE 镜头转换检出率达到 100%,误检率为 3.3%,且未发生镜头切变与 DVE 镜头转换的错认,很好地完成了对镜头切变和 DVE 转换的检测和区分.

#### References:

- [1] Zhang, H.J., Kankanhalli, A., Smoliar, S.W. Automatic partitioning of animate video. Technical Report, Institute of Systems Science, National University of Singapore, 1992.
- [2] Yeo, B.L., Liu, B. Rapid scene analysis on compressed video. IEEE Transactions on Circuits and Systems for Video Technology, 1995,5(6):533~544.
- [3] Lienhart, R. Comparasion of automatic shot boundary detection algorithms. In: Yeung, M.M., Yeo, B., Bouman, C.A., ed. Proceedings of the SPIE, Storage and Retrieval for Image and Video databases VII, Vol 3656. Los Alamitos: IEEE Computer Society Press, 1998. 290~301.
- [4] Kobla, V., Dementhon, D., Doermann, D. Special effect edit detection using video trails: a comparison with existing techniques. In: Yeung, M.M., Yeo, B., Bouman, C.A., ed. Proceedings of the SPIE, Storage and Retrieval for Image and Video databases VII, Vol 3656. Los Alamitos: IEEE Computer Society Press, 1998. 302~313.
- [5] Kim, H., Park, S.J., Lee, J. *et al.* Processing of partial video data for detection of wipes. In: Yeung, M.M., Yeo, B., Bouman, C.A., ed. Proceedings of the SPIE Vol.3656, Storage and Retrieval for Image and Video databases VII, Vol 3656. Los Alamitos: IEEE Computer Society Press, 1998. 280~289.

## Analysis of Digital Video Effect Shot Transition Detection Algorithm\*

BAI Xue-sheng, ZHANG Zi-yin, XU Guang-you, YANG Shi-qiang, SHI Yuan-chun

(Institute of Human-Computer Interaction and Media Integration, Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

E-mail: xgy-dcs@mail.tsinghua.edu.cn

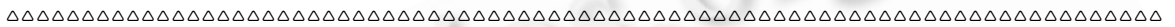
<http://media.cs.tsinghua.edu.cn>

**Abstract:** With the popularity of digital equipment in video producing, digital video effect (DVE) has been more and more frequently used in shot transitions, so that the demand of DVE transition detection is proposed for the video indexing and retrieval system. In this paper, the limitation of the exiting 2-D spatial-temporal image analysis methods is pointed out through the analysis of DVE transition model, and a transition-surface-based DVE detection method is proposed. Basic thoughts and the implementation of this are detailed, as well as the experimental result and analysis.

**Key words:** shot detection; cut; digital video effect; spatial-temporal-image analysis

\* Received November 13, 2000; accepted February 26, 2001

Supported by the National High Technology Development 863 Program of China under Grant Nos.863-306-ZT04-02-01, 863-306-QN99-01; the National Research Foundation for the Doctoral Program of Higher Education of China under Grant No.99000307



### 第 12 届全国信息存储技术会议 征文通知

随着信息技术在当今社会各方面应用的不断深入,信息存储解决方案在信息系统中占有越来越重要的地位,与信息存储相关的产品将拥有更加广阔的市场。为了促进和加强存储技术和产品的交流,中国计算机学会信息存储技术专业委员会决定于 2002 年 11 月 1 日~3 日在上海召开第 12 届全国信息存储技术学术会议。本次会议由中国计算机学会信息存储技术专业委员会主办,复旦大学计算机科学与工程系承办。会议将通过学术报告、专题讨论、产品展示等多种形式,就信息存储的最新进展和发展趋势开展深入、广泛的学术交流,并特邀著名专家学者作专题报告。

#### 一、征文范围

欢迎从事信息技术研究、开发、应用的各界人士,就下列领域(但不限于)中所涉及的信息存储技术方面的内容踊跃来稿:国内外存储技术的发展现状及趋势;国内外存储产品的市场状况及应用前景探讨;存储技术及产品的标准;存储领域中的核心技术及实现研究;存储系统解决方案;SAN&NAS 技术的应用特点及发展趋势;存储相关芯片的设计与应用;海量信息存储的发展;智能存储技术研究;光盘库系统和近线存储技术;管理信息系统中的信息存储技术;金融、电信等大型应用领域中的信息存储技术及产品;高性能计算和多媒体系统中的信息存储技术及产品;信息存储系统的安全性和可靠性;高性能计算机系统

#### 二、征文要求

应征学术论文应是未正式发表过的研究成果,字数(含中英文摘要、关键字与参考文献)不超过 8000 字。请在论文首页上注明联系作者的通信地址、邮政编码、联系电话和 E-mail 地址。

作者可书面或电子投稿:书面投稿一式 3 份,无论录用与否均不退稿;电子投稿的稿件请采用 PostScript,PDF,HTML,RTF,Word 文档或带分行符的纯文本文件等格式。

征文截止: 2002 年 9 月 15 日 录用通知: 2002 年 9 月 30 日

书面投稿请寄: 200433 复旦大学计算机科学与工程系, 杨伟敏 薛向阳(收)

电子投稿的 E-mail 地址: cs211@fudan.edu.cn; xyxue@fudan.edu.cn