

# 一种基于视频相关性的溶解镜头检测方法<sup>①</sup>

## A Dissolve Detection Method Based on Video Correlation

邱建雄<sup>1</sup> 黄少年<sup>2</sup>

(1.长沙学院 计算机系 湖南 长沙 410073; 2.湖南商学院 计算机与电子工程学院 湖南 长沙 410000)

**摘要:** 镜头边界检测在基于内容的视频检索中起着关键的作用,而如何识别渐变镜头是镜头检测中的难点。根据理想视频模型和基于镜头相关性的视频模型间的误差提出了一种溶解镜头检测方法,该方法首先根据溶解模型方差曲线的一次导数和二次导数特征来提取溶解变化的候选区间,然后根据溶解镜头模型误差来验证候选区间以确定最终的溶解变化区间。试验表明:该方法可以有效、准确地检测到溶解镜头。

**关键词:** 溶解 镜头 场景检测 视频 检索

### 1 引言

随着多媒体压缩编码技术、网络技术和个人计算机技术的飞速发展,使得各种各样的多媒体信息在人们的日常生活中广泛传播。视频数据的飞速增长也使如何有效地对其进行管理和检索成为一个新的研究课题。其中,镜头分割是视频检索系统的基础,其检测结果将直接影响到系统性能。

一般来说,视频镜头可分为突变镜头和渐变镜头两种。渐变又可分为淡入、淡出、溶解、划变等几种。对于突变镜头的检测目前已经有很多方法<sup>[1]</sup>,但对很多视频序列,尤其是电影,由于其含有多种渐变类型,目前还没有一种统一的简单有效的检测各种渐变的方法。“双阈值法”<sup>[2]</sup>是一种较典型的渐变检测方法,但该方法的阈值难于确定。而基于边缘特征变化率<sup>[3]</sup>、灰度直方图的扩张与收缩<sup>[4]</sup>、图像亮度的均值与方差<sup>[5]</sup>关系等都是基于某一种渐变类型的检测方法。其中基于边缘特征的方法需要对每一帧进行特征提取,计算量大,对画面的变化较敏感;使用方差和均值关系的方法通过判断方差变化的顶点来检测溶解镜头,但在有较强的对象运动和镜头操作时,容易造成漏检。

在以上研究的基础上,本文分析了溶解镜头模型误差,提出了一种新的溶解镜头的检测方法。该方法

首先通过计算溶解镜头方差曲线的一次导数和二次导数来提取溶解变化的候选区域,然后通过溶解镜头模型的误差来验证溶解变化区域。

### 2 溶解镜头检测方法

#### 2.1 相关定义

在视频制作的后期通常将两个视频信号进行简单的叠加,形成新的视频序列。若两个镜头分别为  $A_1$ ,  $A_2$ , 则在渐变过程中,  $A_1$  的信号强度从 100% 逐渐变为 0%;  $A_2$  的信号则从 0% 逐渐变为 100%。典型的溶解镜头如图 1 所示。淡入/淡出则可看作是溶解镜头的特例。

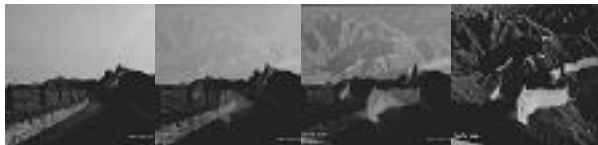


图 1 溶解变换

若  $p(x,y,t)$  表示前一个镜头,  $q(x,y,t)$  为后一镜头,  $f(x,y,t)$  表示由  $p(x,y,t)$  和  $q(x,y,t)$  组成的中间镜头。若  $p(x,y,t)$ 、 $q(x,y,t)$  均为各态经历的随机过程, 则  $p(x,y,t)$ 、 $q(x,y,t)$  可分别表示为  $p(x,y)$ 、 $q(x,y)$ 。则溶

<sup>①</sup> 基金项目:国家自然科学基金(60873016);国家高技术研究发展计划(863)(2009AA01Z102)

收稿时间:2009-02-07

解镜头模型可表示为如下公式：

$$f(x,y,t) = \begin{cases} p(x,y) & t < t_1 \\ \alpha(t)p(x,y) + \beta(t)q(x,y) & t_1 \leq t \leq t_2 \\ q(x,y) & t > t_2 \end{cases} \quad (1)$$

其中  $t_1$ 、 $t_2$  为溶解变化发生的区间， $x$ 、 $y$ 、 $t$  分别为表示水平轴、垂直轴和时间轴上的连续变量。参数  $\alpha(t)$  设置为  $(t_2-t)/(t_2-t_1)$ ，在整个溶解变化的过程中， $\alpha(t)$  是一个递减函数。而  $\beta(t) = 1 - \alpha(t)$ ，它在溶解变换中为一个递增函数。在溶解变换期间， $f(x,y,t)$  的方差变换可表示为：

$$\begin{aligned} \sigma_f(t) &= E[(f - \bar{f})^2] \\ &= E[(\alpha(t)p(x,y) + \beta(t)q(x,y) - \alpha(t)\overline{p(x,y)} - \beta(t)\overline{q(x,y)})^2] \\ &= \alpha^2(t)\sigma_p^2 + \beta^2(t)\sigma_q^2 + 2\alpha(t)\beta(t)E[(p(x,y) - \overline{p(x,y)})(q(x,y) - \overline{q(x,y)})] \end{aligned} \quad (2)$$

其中： $\bar{f}$ 、 $\overline{p(x,y)}$ 、 $\overline{q(x,y)}$  分别表示  $f$ 、 $p(x,y)$ 、 $q(x,y)$  的均值。

如果  $p(x,y)$  和  $q(x,y)$  是独立的，则  $p(x,y)$  与  $q(x,y)$  协方差为零，则公式(2)可以近似的表示成为：

$$\begin{aligned} \sigma_f(t) &\approx \alpha^2(t)\sigma_p^2 + \beta^2(t)\sigma_q^2 \\ &= (\sigma_p^2 + \sigma_q^2)\alpha^2(t) - 2\sigma_p\sigma_q\alpha(t) + \sigma_q^2 \end{aligned} \quad (3)$$

从公式(3)可以看出，溶解镜头的方差曲线是一个开口向下的抛物线。在溶解变化的中心处，即  $\alpha(t) = 1/2$  时，理想抛物线模型的方差  $\hat{\sigma}_{center}$  可定义如下：

$$\hat{\sigma}_{center} = \frac{\sigma_p^2 + \sigma_q^2}{4} \quad (4)$$

在实际的是视频序列中，理想的溶解变化模型方差曲线  $\sigma_f(t)$  会受到其变换前后镜头的内容以及其方差误差的影响，即公式(2)中的第三项。因此，为准确的检测到溶解镜头，本文考虑了方差误差的影响。

### 2.2 镜头相关性的选择

在早期使用抛物线特征进行渐变镜头检测的方法中，紧紧考虑了理想抛物线模型的特征，而没有考虑到邻近镜头对渐变镜头的影响，而这些邻近镜头之间的相互关系实际上又影响了溶解镜头的检测，因此为了精确的检测溶解镜头，必须考虑他们之间的相互关系。

当一个理性融入镜头发生在区间  $[p,q]$  时，实际的方差曲线  $\sigma_{real}(t)$  和理想模型方差曲线  $\sigma_{ideal}$  间的差异我

们称之为溶解模型误差。则该误差可表示为：

$$\begin{aligned} D_{p,q}(t) &= 2\alpha(t)\beta(t)E[(p - \bar{p})(q - \bar{q})] \\ &= 2\alpha(t)\beta(t)\mu_{pq} \end{aligned} \quad (5)$$

其中  $\mu_{pq}$  是镜头间  $p,q$  的协方差。

本文将  $\mu_{pq}$  使用的  $p,q$  标准偏移量进行归一化如下：

$$\rho_{p,q} = \frac{\mu_{p,q}}{\sigma_p\sigma_q} \quad (6)$$

若用等式(6)代替等式(5)，则溶解模型的误差表示如下：

$$D_{p,q}(t) = 2\alpha(t)\beta(t)\sigma_p\sigma_q\mu_{p,q} \quad (7)$$

在溶解变化的中心处，即  $\alpha(t) = 1/2$  时，其最大的误差可表示为：

$$D_{max} = \frac{\sigma_p\sigma_q\mu_{pq}}{2} \quad (8)$$

当区域  $[p,q]$  间的相关性是  $C$  时，最大的误差  $D_{max,C}$  变为  $\frac{\sigma_p\sigma_q C}{2}$ ，若在任意的其它区域，其最大的误差  $D_{max}$  均小于  $D_{max,C}$ ，则该区域可被认为是相关性小于  $C$  的溶解模型。因此具有相关性  $C$  的最大溶解误差  $D_{max,C}$  成为检测所有相关性小于  $C$  的溶解模型的阈值。对于任意的区域  $[p,q]$ ，其阈值定义如下：

$$D_{max,C} = \frac{\sigma_p\sigma_q C}{2} \quad (9)$$

其中，参数  $C$  为经验参数。

### 2.3 溶解镜头检测算法

本文的溶解镜头检测算法分为两步，首先根据方差曲线的一次导数和二次导数提取候选溶解变化区域，然后根据 3.2 节的阈值验证候选区域。

#### ① 候选区域的提取

从公式(3)可知，溶解镜头的方差曲线是一个开口向下的抛物线，而抛物线的一次导数为直线，二次导数则是一个固定的常数。利用这些特征，本文使用方差曲线的一次导数和二次导数曲线来确定候选区域。下图 2 显示了一个溶解镜头的亮度信息方差变化曲线。图 3、图 4 分别显示了其一次导数变化、二次导

数曲线变化的情况。

为了在其一次导数曲线中确定候选区域, 首先将其导数值从负数变为正数的零交叉点作为候选区域的中心, 然后将离中心点最近的左边第一个零导数值点作为区域的起始点; 同理, 将离中心点最近的右边第一个零导数值点作为区域的结束点。以中心点为边界, 左边的区域成为左区间, 右边的区域成为右区间(如图2所示)。考虑到在实际的视频序列中, 渐变镜头一般持续1s(30帧)左右。因此, 对用上述方法检测到的候选区间, 若其长度小于5, 则舍弃该区域。对于选定的候选区域, 在其二次导数曲线内找其左区域的最小值和右区域内的最小值来形成最终的溶解镜头候选的最小值和右区域内的最小值来形成最终的溶解镜头候选区域(如图3所示)。

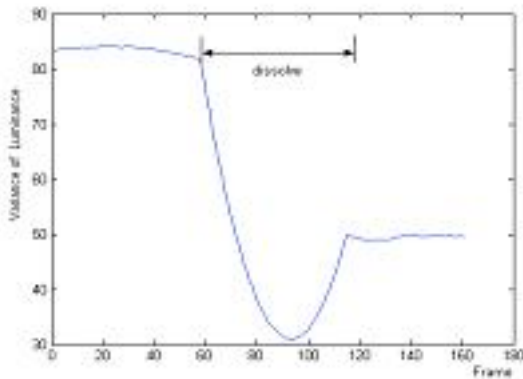


图2 溶解镜头亮度信息方差曲线

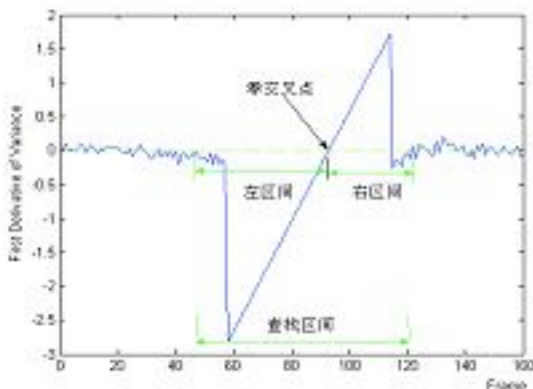


图3 一次导数曲线

为了在其一次导数曲线中确定候选区域, 首先将其导数值从负数变为正数的零交叉点作为候选区域的中心, 然后将离中心点最近的左边第一个零导数值点作为区域的起始点; 同理, 将离中心点最近的右边第

一个零导数值点作为区域的结束点。以中心点为边界, 左边的区域成为左区间, 右边的区域成为右区间(如图2所示)。考虑到在实际的视频序列中, 渐变镜头一般持续1s(30帧)左右。因此, 对用上述方法检测到的候选区间, 若其长度小于5, 则舍弃该区域。对于选定的候选区域, 在其二次导数曲线内找其左区域。

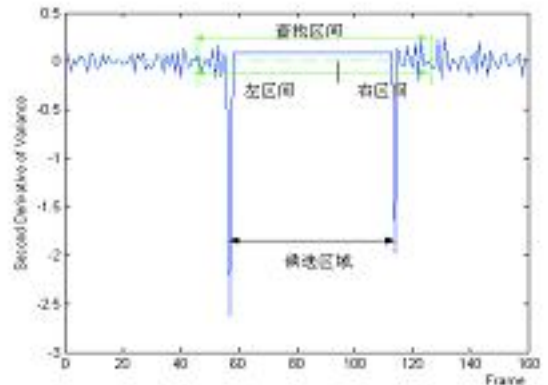


图4 二次导数曲线

## ② 候选区域的验证

上述的候选区域仅仅是根据其一次导数、二次导数的特征来计算的, 但由于镜头内的对象运动和相机操作, 也可能造成误检; 因此需要将真正的溶解镜头的抛物线特征和其它原因引起的特征进行区分, 本文使用3.1节的溶解模型误差进行候选区域的验证。其最大模型误差用方差曲线的中心值  $\sigma_f(\text{center})$  与理想模型的中心值  $\hat{\sigma}_{\text{center}}$  的差来表示。具体的验证算法如下:

Step1: 输入候选溶解变化区域  $r$ ;

Step2: 对查找区域  $r$  用公式(2)计算  $\sigma_f(\text{center})$

Step3: 计算二次导数曲线确定变化的起点和终点, 根据给定的参数计算  $D_{\text{max},c}$ , 并按公式(4)计算其理想溶解变化情况下的中心方差  $\hat{\sigma}_{\text{center}}$ 。

Step4: 若  $\sigma_f(\text{center}) - \hat{\sigma}_{\text{center}} < D_{\text{max},c}$ , 则认为该区域为溶解变化区间。

Step5: 输出溶解变化区间的两 endpoints。

## 3 试验结果分析

实验数据包括我们用 premiere6.0 编辑的一些溶解镜头视频序列以及从网络上下载的 Trec2002 video 测试集, 帧率为 30 帧/秒。试验环境是 P5-1.8GHzCPU+256M 内存的 PC 机和 Windows XP 中文版操作系统。

为了定量的评判试验的结果,我们需要一个统一的对视频边界的认同。在试验中,通过课题研究小组的人员共同商量来手工的标记镜头边界。对于溶解镜头的起始帧和结束帧,若试验检测结果与人工标记结果的误差在2帧之内时,都认为是正确的检测。镜头检测的查全率(Recall)与查准率(Precision)的定义如下:

查全率=正确检测数/(正确检测数+漏检数)

查准率=正确检测数/(正确检测数+误检数)

试验中,我们首先对用 premiere6.0 编辑的6个溶解镜头使用不同的参数进行检测,分别命名为 Test1、Test2、Test3、Test4、Test5、Test6,其检测的结果如下表1所示。表中溶解区域表示溶解镜头发生的区间, $\rho$ 为溶解镜头发生前后视频序列间的相关性。

表1 参数的溶解镜头检测关系

视频名	溶解区域	$\rho$	参数C取值		
			0.2	0.3	0.4
Test1	55~92	-0.1	T	T	T
Test2	37~120	0.32	F	F	T
Test3	83~112	-0.03	T	T	T
Test4	74~135	-0.02	T	T	T
Test5	38~80	0.25	F	T	T
Test6	60~75	0.48	F	F	F

从对上表的分析可知,当 $\rho$ 的取值小于参数C时,该溶解镜头可以被正确的检测(标志T表示正确检测,F表示错误检测)。而当前后镜头的相关性 $\rho$ 大于C时,则不能正确检测。如对Test2的检测, $\rho=0.32$ ,参数C取值0.2,0.3时,均没有正确的检测到溶解镜头。当参数C的值增大到接近1时,其检测的溶解镜头区域不仅包含真正的溶解镜头区域,还包含由于镜头操作或对象运动引起的错误检测区域。因此,对参数的设置需要恰当的考虑视频序列的相关性。

在试验中,还将本文的算法和传统的文献[6]中的方法进行了比较,其中本文参数C的取值为0.5。其检测的结果如下表2所示。可以看到本文算法比文献[5]中方法在查全率和查准率上有一定的提高。

## 4 结论

本文考虑了视频序列相关性对溶解镜头检测的影响并根据视频序列的相关性和其方差曲线建立溶解镜头误差模型。在此基础上,提出了一种基于模型误差的溶解镜头检测方法。试验表明在考虑视频相关性的前提下,该方法可以有效的检测到溶解镜头。

表2 溶解镜头检测结果比较

	本文算法	文献[6]算法
镜头总数	58	58
检测到的镜头数	60	65
正确检测镜头数	55	44
查全率	94.8%	75.8%
查准率	91.7%	67.7%

## 参考文献

- Lienhart R. Comparison of automatic shot boundary detection algorithm. Proc. of SPIE Storage and Retrieval for Image and Video Databases VII, San Jose, CA, U.S.A. 2000(2):290-301.
- Browne P, Smeaton AF, Murphy N, O'Connor N, Marlow, Berrut C. Evaluation and combining digital video shot boundary detection algorithms. Proc. of the Fourth Irish Machine Vision and Information Processing Conference, Queens University Belfast, 2000:136-148.
- Zabin R, Miller J, Mai K. Feature-Based Algorithm for Detecting and Classifying Scene Breaks. ACM Conference of multimedia Systems, 1995.
- Fernando WAC, Canagarajah CN, Bull DR. Fade in and Fade-out Detection in Video Sequence Using Histograms. ISCAS, 2000(4):748-752.
- Truong Bt, Dorai C, Venkatesh S. Improved Fade and Dissolve Detection for Reliable Video. Image Processing, 2000:961-964.
- Lienhart R, Reliablr Dissolve Detection. Computer Vision and Image Understanding, 1998,71(1):94-109. Segmentation.