

基于图像分割和对象跟踪的新闻 视频镜头边界检测方法

徐新文 李国辉 朱 为

(国防科技大学信息系统与管理学院,长沙 410073)

摘要 镜头边界检测是许多多媒体应用的一个重要步骤,而现有的镜头边界检测方法大都是首先提取视频帧低层视觉特征,然后构造相异性测度函数,但由于这些方法对低层特征变化、对象运动、摄像机运动和视频质量较敏感,为克服此问题,提出了一种基于图像分割和对象跟踪的镜头边界检测方法。该方法首先采用分区直方图对镜头进行预检,然后利用基于小波分析的无监督图像分割和对象跟踪技术,通过构造相异性测度函数来对镜头边界进行确认。由于分区直方图方法作为第1过滤器,可有效地减少图像分割和对象跟踪的视频帧数目,从而提高了整个算法的效率,而基于小波变换的无监督图像分割和对象跟踪,则对以上问题具有较好的鲁棒性。在3个多小时的 CCTV 和 CNN 新闻视频实验中,获得了 97.2% 查准率和 96.4% 的查全率。

关键词 镜头边界检测 图像分割 对象跟踪 分区直方图

中图法分类号:TP391.41 TN941.1 文献标识码:A 文章编号:1006-8961(2009)08-1594-07

A Shot Boundary Detection Method for News Video Based on Image Segmentation and Object Tracking

XU Xin-wen, LI Guo-hui, ZHU Wei

(School of Information System and Management, National University of Defense Technology, Changsha 410073)

Abstract As a critical step in many multimedia applications, shot boundary detection attracts many research interests in recent years. Most present methods measure the similarity among video frames based on its low-level features. However, they are sensitive to the change in brightness, color, motion of object, camera motions and the quality of video. This paper proposes an innovative shot boundary detection method for news video based on image segmentation and object tracking. It combines three main techniques, namely, the partitioned histogram comparison method, the image segmentation based on wavelet analysis and the object tracking. The partitioned histogram comparison is used as the first filter to effectively reduce the number of video frames which need segmentation and object tracking. The unsupervised image segmentation based on wavelet analysis and object tracking is robust to those problems mentioned above. The efficacy of the proposed method is extensively tested with more than 3 hours of CCTV and CNN news programs, and that 96.4% recall with 97.2% precision has been achieved.

Keywords shot boundary detection, image segmentation, object tracking, partitioned histogram

1 引言

众所周知,镜头边界检测是许多多媒体应用的

一个重要步骤。它是基于内容的视频访问和视频挖掘工作的第1步,它的精度直接影响后续工作的准确度。一个镜头是指摄像机一次连续拍摄所记录的图像序列,而视频则是由一个个镜头采用各种方式

基金项目:国家自然科学基金项目(60243006);国家教育部博士点基金项目(20069998022)

收稿日期:2008-01-11;改回日期:2008-05-28

第一作者简介:徐新文(1974~),男,国防科技大学信息系统与管理学院系统工程专业博士研究生。主要研究方向为多媒体信息系统和多媒体数据挖掘。E-mail:xinwen_xu@126.com

串连而成。根据镜头连接方式不同,可将镜头按切换方式不同分为突变过渡(abrupt transition)和渐变过渡(gradual transition)两类。其中突变是将两个镜头直接连接在一起,中间没有任何视频编辑特效;渐变则是将两个镜头通过视频编辑特效连接在一起。根据视频编辑特效的不同,镜头渐变可细分成淡入淡出(fade)、溶解(dissolve)、划变(wipe)等。

镜头边界检测就是检测镜头间的连接方式,并将视频分成一个个物理镜头。现有的镜头边界检测方法^[1-4]有很多,按视频源的编码方式可分为基于像素(非压缩)域的方法和基于压缩变换域的方法。基于像素域的镜头检测方法主要有:

(1)像素差异法 该方法首先定义像素相异性测度,然后计算连续两帧图像的帧间差异,并用其与特定阈值进行比较,以检测镜头是否发生改变。对于渐变镜头,一般采用双阈值的方法^[5]进行检测,其中高阈值用于过滤突变镜头,低阈值用于确定渐变过渡。该方法虽简单直接,但对物体运动和摄像机运动很敏感;

(2)直方图法^[6] 它能有效地克服由帧内对象的局部运动产生的影响,但在统计直方图的过程中,易丢失图像像素点的位置信息,无法反映图像的整体内容,因此易发生漏检;

(3)块匹配法 与像素差异法不同,基于块匹配的方法是先将每一帧图像划分成 k 个块,然后通过比较对应的块对连续帧之间的相似性进行估计,该方法利用了图像局部特征来抑制噪声以及摄像机和物体运动的影响,同时保留了一定的位置信息,但对复杂场景变化和激烈运动的视频,其检测效果仍不理想;

(4)边界变化率法^[7] 其主要思想是通过计算边界的变化程度来确定镜头的边界,而计算边界的改变则是通过对全局运动信息进行估计来确定出摄像机的运动参数,并用该全局运动向量来对齐相邻两帧,以消除摄像机运动参数对边缘比较的影响。虽然该方法能减少摄像机运动的影响,但为了抵消运动带来的影响,还需要进行运动补偿,因此计算量较大。再者,当前镜头中如果存在多种运动时,则易产生误检;

(5)光流法和模型法 它们一般只适用于特定类型的镜头渐变的检测,且对噪声与运动很敏感。

基于压缩变换域的镜头检测方法主要有基于离散余弦变换(DCT)系数和基于小波变换的方法。

Yeo 和 Liu 等人提出了一种利用 MPEG 压缩视频的 DC 序列进行镜头边界检测的方法^[8]。由于该方法直接在压缩域中操作,因此虽然计算速度快,但在具有相似像素值,而具有不同密度分布的两帧之间则容易发生误检。基于小波变换的镜头检测方法^[9-10]是在子带域上对镜头边界进行检测的一种方法,其基本思想是首先对图像进行小波分解后,再分别对它的低频部分和高频部分进行分析和处理。在其低频部分利用彩色直方图(color histogram)的方法可以检测出突变;在其高频部分应用边界计数(edge count)、边界谱平均(edge spectrum average)以及双彩色差分(double chromatic difference)等算法即可分辨出渐变镜头边界。但该方法在小波基选择与相异性测度函数的构造上较为困难。另外,文献[11]、[12]通过模糊粗糙集理论和 Adaboost 方法结合多种特征以及文献[13]、[14]采用动态的自适应的阈值选择虽然改善了镜头边界检测的精度与鲁棒性,但仍有较大的提升空间。

本文针对非压缩视频数据,提出了一种基于图像分割和对象跟踪的镜头边界检测新方法。其包括3个主要技术,即分区直方图方法、基于小波变换的无监督图像分割算法和对象跟踪技术。该方法首先采用分区直方图方法作为第1过滤器来检测出任何可能的候选镜头边界;然后在小波分析的基础上,再利用无监督图像分割算法自动提取候选镜头边界附近帧的分割掩码图,并进行帧间差比较;最后通过对分割结果中的对象(或分块)进行跟踪分析,以进一步对候选镜头边界进行确认。真实新闻视频的实验结果证明,本文方法是有效的和鲁棒的。

2 分区直方图特征分析

在已往的研究和实际应用中,计算帧间直方图差值法已被证明是简单,且高效的一种算法。但是,由于全局直方图统计丢失了像素的位置信息,从而导致它无法检测颜色直方图的统计特征相同,但位置分布不同的镜头边界。因此,本文采用分区颜色直方图差值法作为镜头边界检测的第1过滤器,并采用严格的阈值来尽量不错过任何可能的镜头边界。

分区直方图法的主要思想是:首先将视频帧分成 N 个子块,然后计算每个子块各自的颜色直方图,最后将各个子块的颜色直方图之差相加来得到

最后的帧间差。比如,对一个 3 维颜色空间 S , 可将其 3 个颜色分量分别量化为 L_1, L_2, L_3 级。令 $L = L_1 \times L_2 \times L_3$, 则 3 维颜色空间就被量化为 L 个级, 而连续两帧 f_i 和 f_{i-1} 的差 $D_H(f_i, f_{i-1})$ 则为

$$D_H(f_i, f_{i-1}) = \sum_{n=1}^N d(H_i(n), H_{i-1}(n)) \quad (1)$$

其中, $H_i(n)$ 为第 i 帧的第 n 个子块的颜色直方图, $d(H_i(n), H_{i-1}(n))$ 为连续两帧间的对应于第 n 个子块的直方图差, 其计算公式为

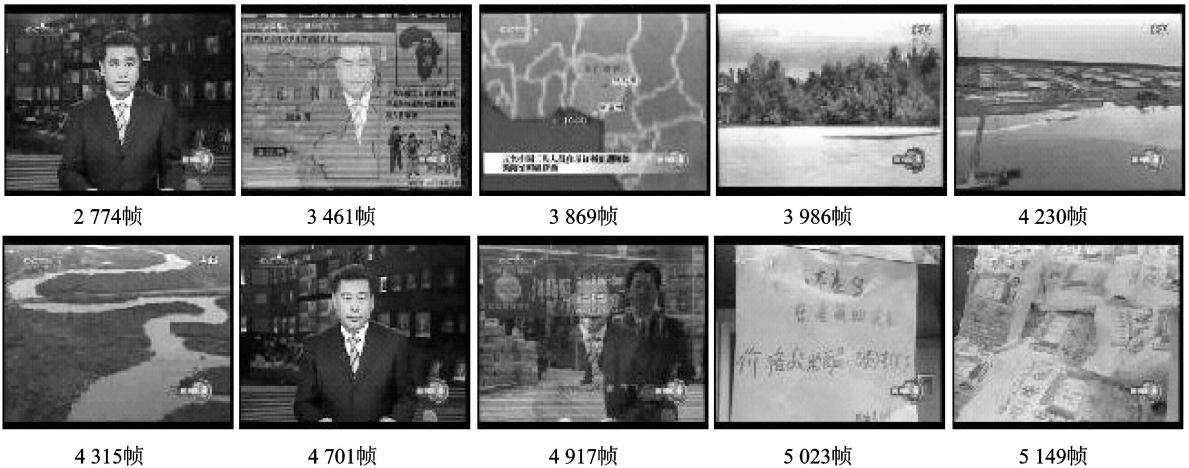
$$d(H_i(n), H_{i-1}(n)) = \sum_{l=1}^L |h_i^{(n)}(l) - h_{i-1}^{(n)}(l)| \quad (2)$$

其中, $h_i^{(n)}(l)$ 是第 i 帧中第 n 个子块的像素颜色分量值为 l 级的像素占有所有像素的比例。

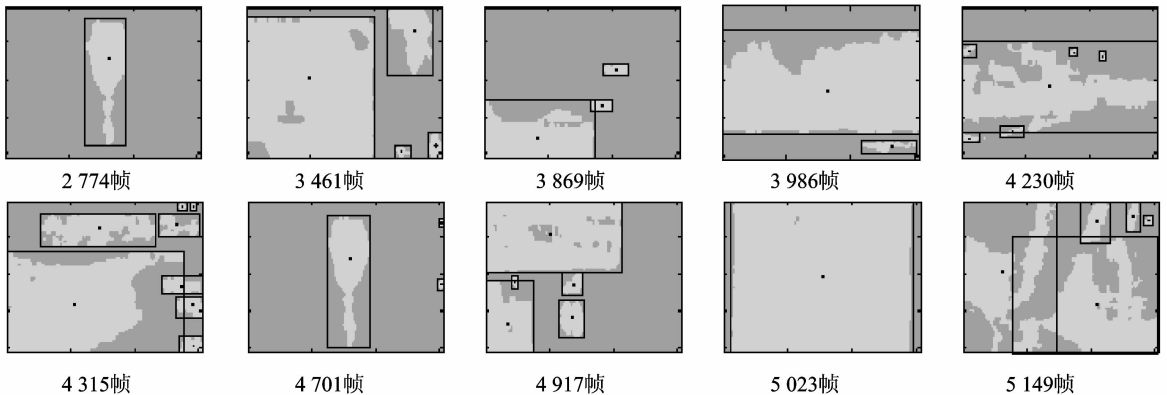
3 图像分割和对象跟踪

3.1 基于小波的无监督图像分割算法

结合小波分析, 采用分块和类参数同时评估 (SPCPE)^[15] 的无监督图像分割算法来进行视频帧分割, 并首先阐明类和分块这两个概念, 其中类是特征的统计描述, 它包括图像中相同特征的所有区域; 分块是类的一个实例。图 1 为类与分块的例子, 其中, 图 1(a) 为视频帧序列图像, 图 1(b) 为图 1(a) 相应帧的分割掩码图, 图中的亮区与暗区分别表示两个不同的类, 且每个对象有一个边界框和一个重心, 它们就是图像分割时要提取的结果。



(a) CCTV 新闻视频序列



(b) 图 1(a) 相应视频帧的分割掩码图

图 1 镜头边界与对象掩码图

Fig. 1 Shot boundary and segmentation mask map

SPCPE 是一个视频帧的无监督分割方法,它可用来对分块和类参数同时进行估计。实际上,由于类描述及其参数是不容易得到的,且图像分割是无用户干预的,因此不知道像素应该属于哪个类。SPCPE 算法将分块和类视为随机的变量。

假设分别有类 ω_1 , 类 ω_2 两类,其分块变量集合 $C = \{c_1, c_2\}$, 则类就由参数 $\theta = \{\theta_1, \theta_2\}$ 确定。假定所有像素值为 $y_{i,j}$, 且像素都属于类 $\omega_k (k=1, 2)$, 则形成一个向量 Y_k , 矩阵 Φ 中的每一行为 $(1, i, j, i \cdot j)$, $\alpha_k = (\alpha_{k,0}, \dots, \alpha_{k,3})^T$ 为参数向量。

$$y_{i,j} = \alpha_{k,0} + \alpha_{k,1}i + \alpha_{k,2}j + \alpha_{k,3}i \cdot j \quad (3)$$

$$\forall (i,j) y_{i,j} \in c_k$$

$$Y_k = \Phi \alpha_k \quad (4)$$

$$\hat{\alpha}_k = (\Phi^T \Phi)^{-1} \Phi^T Y_k \quad (5)$$

对于图像数据 Y , 其最佳的分块被估计为分块变量的最大后验概率。类参数 $\theta = \{\theta_1, \theta_2\}$ 和分块变量集合 $C = \{c_1, c_2\}$ 被估计为

$$\begin{aligned} (\hat{C}, \hat{\theta}) &= \text{Arg max}_{(C, \theta)} P(C, \theta | Y) \\ &= \text{Arg max}_{(C, \theta)} P(Y | C, \theta) P(C, \theta) \end{aligned} \quad (6)$$

取函数 $J(C, \theta)$ 的极小值, 根据上面的假设, 则联合估计可以简化为以下形式:

$$(\hat{C}, \hat{\theta}) = \text{Arg min}_{(C, \theta)} J(c_1, c_2, \theta_1, \theta_2) \quad (7)$$

$$J(c_1, c_2, \theta_1, \theta_2) = \sum_{y_{i,j} \in c_1} -\ln P_1(y_{i,j}; \theta_1) + \sum_{y_{i,j} \in c_2} -\ln P_2(y_{i,j}; \theta_2) \quad (8)$$

这样对于每个类来说, 分割问题就变为了同时估计分块和类参数的问题。关于参数的估计, 对每个指定的类可以采用式(5)来直接计算其参数。关于类分块估计, 若给像素 $y_{i,j}$ 指定一类, 且该类对应一个最低的 $-\ln P_k(y_{i,j}, \theta_k)$, 则其判定规则如下:

$$\begin{cases} y_{i,j} \in \hat{c}_1 & -\ln p_1(y_{i,j}) \leq -\ln p_2(y_{i,j}) \\ y_{i,j} \in \hat{c}_2 & \text{其他} \end{cases} \quad (9)$$

该算法先从第1个候选镜头边界处的视频帧开始进行分割, 并计算相应的类参数; 然后用这些数据和类参数估算出一个新的分块, 并将类参数和分块通过反复迭代进行优化, 直到没有进一步的变化。函数 J 的极小值可以产生一个局部最小值, 从而保证了迭代算法的收敛性。

原始的 SPCPE 算法是从随机产生的初始分块开始的, 但不同的初始分块产生的局部最小值不同。通常, 随机产生的初始分块经常导致不理想的分割

结果。可见, 初始分割会影响分割计算的效率和最终分割的效果。目前有以下两种方法进行初始分割: 一种是随机直线分割法, 即用随机直线将原始图像分割成不同的区域, 以表示不同的类; 另一种方法是通过选择理想的预定义模板来生成初始分块。尽管通过结合以上两种方法的分割结果得到改善, 但仍不能获取图像的实际轮廓和对象的分布。以往研究表明, 小波分析有助于图像分割, 所以本文提出一种基于小波变换的无监督图像分割方法。该方法先通过选用适当的小波进行分析来去除高频部分的大量子带, 这样图像中可能的区域将更加清晰。基于算法的有效性和效率考虑, 本文采用 Daubechies 小波对帧图像进行预处理, 且分解层为1。分割时, 首先通过合并各通道中的突出点来获得一个粗略的初始分割, 并将其作为 SPCPE 的输入。实验表明, 通过多分辨率小波分析检测出的突出点, 大多数为原图像的突出点, 从而有效地改善了最终的分割结果。

3.2 对象跟踪技术

对象跟踪的基本思想是首先从视频帧的分割结果中提取出离散的对象(或分块), 并计算分块的重心、位置和大小, 然后通过计算相邻帧中对象间的差异进行对象跟踪。

(1) 对象提取 本文采用线合并算法来提取离散对象, 它是一种既快速又简单的分块提取方法。其基本思想是逐行或逐列对分割视频帧进行扫描, 如果行数少, 则进行逐行扫描, 否则进行逐列扫描, 以发现新的分块, 并将相邻分块加以合并。这样一次扫描后就能得到所有的分块的重心、位置和大小。位置定义为能包含对象的最小矩形框 $B(P_L, P_R)$, P_L 和 P_R 分别为矩形框的左上角和右下角的点, 其坐标分别表示为 (x_L, y_L) 和 (x_R, y_R) 。大小为对象包含的像素数 N 。重心为 $p_c = (x_c, y_c)$, 重心坐标计算公式如下:

$$\begin{aligned} x_c &= \left(\sum_{n=1}^N x_n \right) / N, y_c \\ &= \left(\sum_{n=1}^N y_n \right) / N \end{aligned} \quad (10)$$

其中, N 为对象包含的像素数, (x_n, y_n) 为对象中第 n 个像素的坐标。

(2) 对象跟踪 相邻帧间对象跟踪技术的主要思想就是计算相邻两帧间相关对象的重心、位置和大小间的距离, 当它小于某一阈值时, 则认为它们是同一对象。跟踪时, 首先计算两视频帧 f_i 和 f_{i-1} 中

对象间的重心距离,即

$$D_C = \sqrt{(x_i^C - x_{i-1}^C)^2 + (y_i^C - y_{i-1}^C)^2} \quad (11)$$

对于小于某一阈值的两对象,则再计算对象边界框位置间的距离

$$D_B = \frac{1}{2} \left(\sqrt{(x_i^L - x_{i-1}^L)^2 + (y_i^L - y_{i-1}^L)^2} + \sqrt{(x_i^R - x_{i-1}^R)^2 + (y_i^R - y_{i-1}^R)^2} \right) \quad (12)$$

也就是对象边界框的对应顶点的距离;然后计算它们的面积差,只有以上 3 个数值均小于各自的阈值,才判定两个对象是匹配的。

4 镜头边界检测

本文提出的基于图像分割和对象跟踪的视频镜头边界检测方法步骤如下:

(1) 比较连续两帧 f_i, f_{i-1} 的直方图,用式(1)计算其直方图的差 $D_H(f_i, f_{i-1})$ 。如果 $D_H(f_i, f_{i-1})$ 小于直方图阈值 T_H ,则此处不是镜头边界,继续处理下一帧,否则进行下一步。

(2) 分割当前帧和前一帧,并获取它们的分割掩码图 G_i 和 G_{i-1} 。

(3) 计算连续两视频帧间的分割掩码图的差: $D_M(f_i, f_{i-1}) = |G_i - G_{i-1}|$ 。由于分割掩码图为二值数据,直接对掩码图进行异或运算即可得出两帧图像间的差异。用 N_{diff} 表示为 $D_M(f_i, f_{i-1})$ 中非零值的个数, N_{all} 表示总像素数,则两视频帧间掩码图中像素改变百分比为 $R(f_i, f_{i-1}) = N_{diff}/N_{all}$ 。

(4) 检查 $R(f_i, f_{i-1})$

if $R(f_i, f_{i-1}) < T_1$

此处不是镜头边界,返回步骤(1)继续处理下一帧。

else

if $R(f_i, f_{i-1}) < T_2$

对连续两帧进行对象跟踪,计算当前帧与前一帧间的对象的不匹配数目 $N_o(f_i, f_{i-1})$,两帧中不匹配对象的个数为两帧中除匹配对象外的所有对象数之和。

if $N_o(f_i, f_{i-1}) < T_0$

此处不是镜头边界,返回步骤(1)继续处理下一帧。

else

此处是镜头边界,返回步骤(1)继续处理下一帧。

endif;

else

此处是镜头边界,返回步骤(1)处理下一帧。

endif;

endif;

其中, T_H 为分区直方图差的阈值, T_1 和 T_2 为两分割掩码图的变化百分比的双阈值, T_0 为对象不匹配数目的阈值,它们均来自实验中的经验数据。

5 实验结果与分析

为验证本文算法的视频镜头边界的检测效果,用中国中央电视台(CCTV)和美国有线电视新闻网(CNN)的共 7 段新闻视频对算法进行了测试实验,实验时,首先去除台标和字幕区域,以减少其对图像分割和对象跟踪的影响,视频的详细信息如表 1 所示。为了衡量检测结果的好坏,采用查准率 R_{recall} 和查全率 $R_{precision}$ 作为评价指标,其定义如下:

$$\begin{cases} R_{recall} = n_c / (n_c + n_m) \\ R_{precision} = n_c / (n_c + n_f) \end{cases} \quad (13)$$

其中, n_c, n_m, n_f 分别表示正确、漏检和误检的镜头数。

表 1 新闻视频的详细信息

Tab. 1 Information of news program

新闻视频	CCTV 新闻					CNN 新闻		总共
	1	2	3	4	5	6	7	
长度	30'00"	30'00"	30'00"	30'00"	30'00"	30'00"	28'30"	208'30"
镜头边界数	464	386	465	445	465	331	321	2 877
突变镜头数	455	380	455	433	461	33	56	2 270
渐变镜头数	9	6	10	12	4	298	265	607

实验时,首先用人工方法对视频进行切分,并将其分割结果作为正确的镜头边界。考虑到镜头渐变

时,镜头边界位置存在差异,一般认为误差在 10 帧以内都是允许的。为了证明本文方法的鲁棒性,采

用经典的 Twin Comparison 方法^[5]对相同的视频片段进行了检测比较,比较结果列于表 2。表 2 中,7 段新闻视频共有 2 877 个镜头边界,本文方法检测到 2 857 个镜头边界,其中包括 91 个漏检和 71 个误检,查全率和查准率分别达到 96.4% 和 97.2%。漏检的原因主要是由于镜头渐变边界处帧间亮度和颜色变化小,致使分区直方图方法产生漏检而直接影响最终的检测结果;误检主要是由于场景变化复杂,加之不规则快速摄像机运动所造成的。对于分区直方图差阈值 T_H 、分割掩码图变化百分比的双阈值 T_1 和 T_2 以及对象匹配数目阈值 T_0 ,均采用滑动窗口法和高斯模型相结合的方法,在局部范围内利用高斯模型求取。局部均值用以计算阈值;局部方差评估阈值的可靠性,当方差大时,均值不可靠,则采用

前一均值计算阈值;反之可靠,可直接用均值计算阈值。 T_H 和 T_1 应尽量小些,而 T_2 应尽量大些,以避免过滤阶段发生漏检影而响最后的性能,但由于其增加了图像分割和对象跟踪处理的帧数,从而降低了整个检测方法的速度。对于对象匹配数目阈值,可采用交叉验证法进行确定。而且它们对最后的查准率和查全率影响不明显。总之,基于图像分割的方法,由于提取了对象级的特征,从而有效地克服了基于图像低层特征的检测因亮度和颜色的变化而造成的对检测的影响。另外,对象跟踪分析对摄像机的一些运动如平推、摇摆运动以及对象运动具有较强的鲁棒性。图 1 为一新闻视频的镜头边界变换与分割掩码图的例子。

表 2 本文方法与 Twin Comparison 方法的检测结果比较

Tab. 2 Compare between the proposed method and Twin Comparison method

测试视频	本文方法					Twin Comparison 方法				
	检出数	漏检数	误检数	查全率(%)	查准率(%)	检出数	漏检数	误检数	查全率(%)	查准率(%)
1	463	7	6	98.5	98.7	471	32	39	93.1	91.7
2	381	9	4	97.7	99.0	403	15	32	96.1	92.1
3	468	7	10	98.5	97.9	474	36	45	92.3	90.5
4	443	8	6	98.2	98.6	464	30	49	93.3	89.4
5	466	2	3	99.6	99.4	473	28	36	94.0	92.4
6	322	32	23	90.3	92.9	344	39	52	88.2	84.9
7	314	26	19	91.9	93.9	338	29	46	91.0	86.4
总共	2 857	91	71	96.4	97.2	2 967	209	299	92.6	89.6

6 结 论

本文针对非压缩视频数据,提出了一种基于图像分割和对象跟踪的镜头边界检测新方法,其包括 3 个主要技术,即分区直方图检测法、基于小波变换的无监督图像分割算法和对象跟踪技术。其中分区直方图方法是通过严格的阈值设定来尽可能检测出候选镜头边界;基于小波变换的无监督图像分割算法是通过自动提取候选镜头边界附近帧的分割掩码图来进行帧间相异度检测,即先确认大部分镜头边界,再通过对分割结果中的对象进行跟踪分析来对候选镜头边界进行进一步确认。分区直方图方法作为第 1 过滤器,可有效地减少图像分割和对象跟踪的视频帧数目,从而提高了整个方法的效率;基于小

波变换的无监督图像分割算法和对象跟踪技术能较好地克服低层特征变化、对象运动、摄像机运动和视频质量好对镜头边界检测的影响。在 3 个多小时 CCTV 和 CNN 新闻视频的检测实验中获得 97.2% 查准率和 96.4% 的查全率。

在进一步的工作中,可将本文的方法与其他信息(如音频、字幕等)相结合,以期取得更好的检测效果。同时,可利用本文算法中提取的对象级信息结合其他方面的信息来进行关键帧提取、新闻视频故事分割、新闻事件的跟踪与挖掘。

参考文献 (References)

- 1 Fu Chang-jian, Li Guo-hui, Hu Jun-tao. Video hierarchical structure mining[J]. Computer Engineering and Science, 2006, 26(6): 159-162. [付畅俭, 李国辉, 胡军涛. 视频层次结构挖掘[J]. 计算机工程与科学, 2006, 26(6): 159-162.]

- 2 Boreczky J S, Rowe L A. Comparison of video shot boundary detection techniques[A]. In: Proceedings of the SPIE Conference on Storage and Retrieval for Still Images and Video Databases IV [C], San Jose, California, USA, 1996:170-179.
- 3 Ford Ralph M, Robson Craig, Temple Daniel, *et al.* Metrics for shot boundary detection in digital video sequences [J]. *Multimedia Systems*, 2000, **8**(1):37-46.
- 4 Zhu Xi, Lin Xing-gang. Survey on video temporal segmentation[J]. *Chinese Journal of Computer*, 2004, **27**(8):1027-1035. [朱曦, 林行刚. 视频镜头时域分割方法的研究[J]. *计算机学报*, 2004, **27**(8):1027-1035.]
- 5 Zhang H J, Kankanhallil Atreyi, Smoliar Stephen W. Automatic partitioning of full motion video[J]. *Multimedia Systems*, 1993, **1**(1):10-28.
- 6 Nagasaka A, Tanaka Y. Automatic video indexing and full-video search for object appearances[A]. In: Proceedings of IFIP Second Workshop Conference on Visual Database System II[C], Budapest, Hungary, 1992:113-127.
- 7 Zabin R, Miller J, Mai K. A feature-based algorithm for detecting and classification production effects[J]. *Multimedia Systems*, 1999, **7**(2):119-128
- 8 Yeo B L, Liu B. Rapid scene analysis on compressed videos[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 1995, **5**(6):533-544.
- 9 Yu H H, Wolf W. A hierarchical multiresolution video shot transition detection scheme [J]. *Computer Vision and Image Understanding*, 1999, **75**(1-2):196-213.
- 10 Han Bing, Gao Xin-bo, Ji Hong-bing. A hierarchical and multiresolution detection method of shot transition for video shots segmentation[J]. *Computer Science*, 2006, **33**(6):255-231. [韩冰, 高新波, 姬红兵. 一种分层的和多分辨率的镜头边界检测方法[J]. *计算机科学*, 2006, **33**(6):255-231.]
- 11 Han Bing, Gao Xin-bo, Ji Hong-bing. A shot boundary detection method for news video based on rough fuzzy sets[J]. *Acta Electronica Sinica*, 2006, **34**(6):1085-1089. [韩冰, 高新波, 姬红兵. 基于模糊粗糙集的新闻视频镜头边界检测方法[J]. *电子学报*, 2006, **34**(6):1085-1089.]
- 12 Zhao Zhi-cheng, Cai An-ni. Shot boundary detection algorithm in compressed domain based on adaboost and fuzzy theory [A]. In: Proceedings of the Second International Conference on Advances in Natural Computation (ICNC 2006) [C], Xi'an, China, 2006:617-626.
- 13 Zhu Xing-quan, Xue Xiang-yang, Wu Li-de. An autom threshold detection method in video shot segmentation[J]. *Journal Computer Research and Development*, 2000, **37**(1):80-85. [朱兴全, 薛向阳, 吴立德. 一种自动门限选取的视频 Shot 分割方法[J]. *计算机研究与发展*, 2000, **37**(1):80-85.]
- 14 Truong B T, Dorai C, Venkatesh S. New enhancements to cut, fade, and dissolve detection processes in video Segmentation [A]. In: Proceedings of the 8th ACM International Conference on Multimedia [C], Marina del Rey, California, USA, 2000:219-227.
- 15 Sista S, Kashyap R L. Unsupervised video segmentation object tracking[J]. *Computers in Industry*, 2000, **42**(2):127-146.