

一种基于 H. 264/AVC 的压缩域运动对象分割方法

冯 镔 肖 非 朱光喜 袁 姣

(华中科技大学电子与信息工程系, 武汉 430074)

摘 要 目前大部分压缩域视频对象的分割方法主要面向 MPEG 系列视频标准,且算法建模复杂。为了解决这一问题,现提出了一种新的基于 H. 264/AVC 的压缩域时空联合运动对象分割(TSMOS)算法。该方法主要利用压缩码流中的 DCT 系数和运动矢量信息进行对象分割,并首先利用相邻帧 DCT 系数之差提取运动对象轮廓,同时通过对轮廓进行形态学和抗噪声处理来得到粗糙的运动对象帧差掩码;然后采用运动向量归一化、噪声向量滤除、权值扩展向量中值(WEVM)滤波及前帧分割结果后向投影技术来得到对象的运动掩码;最后通过引入有效机制合并帧差掩码和运动掩码来分割运动对象。实验证明,该算法可取得较好的分割效果。

关键词 压缩域 DCT 系数 运动向量 权值扩展向量中值滤波 对象分割

中图法分类号: TN919.81 TN 911-73 文献标识码: A 文章编号: 1006-8961(2009)07-1327-07

Moving Object Segmentation in Compressed Domain for H. 264/AVC

FENG Bin, XIAO Fei, ZHU Guang-xi, YUAN Jiao

(Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074)

Abstract Most of the current moving object segmentation algorithms in compressed domain were focused on the MPEGx video standards which were complexity intensive due to the use of complex mathematical models. In order to resolve these problems, a new temporal-spatio method for moving object segmentation (TSMOS) in H. 264/AVC compressed domain was proposed. DCT coefficients and motion vectors (MV) in bits stream were mainly used in this algorithm. Firstly, the differential mask of moving object was abstracted using the DCT coefficients difference (QDCD) between adjacent frames and followed by morphological processing. Then the methods of normalizing MV, weighted extended vector median (WEVM) filter on the MV fields and the existent segmentation results projection were utilized to obtain the moving mask of object. Finally, accurate moving objects segmentation results were achieved by combining the differential mask and the moving mask. The experiments on H. 264/AVC test sequences demonstrated the validity of the proposed methods.

Keywords compressed domain, DCT coefficients, motion vectors, weighted extended vector median filter, object segmentation

1 引 言

视频对象分割是智能视频处理与分析中的一项重要技术,由于在此基础上可以进行对象跟踪和基于对象的压缩编码等技术的研究,其已广泛应用于视频监控、视频检索和基于视频的人机交互等领域,

因此受到了研究人员广泛的关注。由于传统的视频对象分割技术是在像素域中进行的,其不仅要完全解码全部压缩码流,而且分割算法复杂,需要大量的计算时间和存储空间,不能满足高效地处理大量压缩视频的需求,所以研究直接从压缩视频中提取对象的方法尤为重要^[1]。

压缩域运动对象分割算法主要是利用压缩码流

基金项目:国家自然科学基金项目(60502023);湖北省自然科学基金项目(2007ABA090)

收稿日期:2007-09-24;改回日期:2008-03-27

第一作者简介:冯 镔(1978 ~),男。2006 年于华中科技大学获工学博士学位,现为华中科技大学电子与信息工程系讲师。主要研究方向为多媒体通信与模式识别。E-mail: frobby@163.com

中的运动向量(MV)和离散余弦变换(DCT)系数来进行运动对象分割。运动向量可看作是压缩域中的光流场,其能反映运动对象的大致位置。DCT系数是图像灰度值在频域中的表示,其能反映图像的能量和纹理信息。纵观这些算法,从所采用的压缩域信息和判定准则的角度看,压缩域运动对象的分割方法基本可分为以下3类:基于运动向量的分割方法^[2-4]、基于变化检测的分割方法^[5-6]和联合时空域的分割方法^[7-8]。

基于运动向量的分割方法是利用压缩码流中的MV合并聚类来获得运动一致性区域。Nakajima等人通过分析P、B帧中的运动矢量与相邻宏块和参考帧中对应宏块的运动向量的相关性来检测运动对象^[2]。对于帧内块是通过判定其是否作为后续帧中某运动对象的参考块来进行检测。Badu等人首先通过对运动矢量进行滤波和内插来得到密集的运动向量场,再先后采用仿射参数模型和EM算法来估计运动对象的个数和提取对象的掩码;最后结合像素域方法得到精确的对象轮廓^[3]。刘志等人提出了一种在H.264/AVC上分割对象的方法^[4],即首先对累计运动向量场进行全局运动补偿,然后利用运动相似性,采用统计区域生长算法分割多个区域,最后结合向量方向角直方图来提取运动对象。基于变化检测的分割方法是利用变换域DCT系数的特征来检测由于图像中的对象运动而带来的变化区域。文献[5]是通过度量相邻块和相邻帧中对应块的直流系数(DC)和交流系数(AC)的相似性来合并区域,如果某一区域的平均时域变化大于整帧图像的平均时域变化,则此区域为前景,否则为背景。文献[6]是通过压缩码流进行部分解码,先建立原始图像的DC图像,再利用前景和背景信号的不同分布特征,结合高阶矩的方法来实现对DC图像的对象进行分割。联合时空域的分割方法是结合时域MV和空域DCT系数来分割运动对象。文献[7]提出先利用运动向量的空域一致性和相邻帧投影技术来定位出运动对象的大致位置,再对运动区域的宏块DC系数,采用最大熵失真聚类准则来分割运动对象。文献[8]则先建立基于块的马尔可夫随机场(MRF)的标记模型,再利用运动向量和DC系数来设置不同的能量函数,最后结合高置信度优先(HCF)的迭代算法,通过求解能量最小的标记场来实现运动对象的分割。

基于MV的对象分割方法能有效利用运动的时

空域一致性,适用于有明显运动的刚体对象分割,但易造成欠分割。对于受噪声影响较大的压缩码流,因为此时不能从码流中得到可靠的运动向量,而且由于存在较多的帧内编码块,所以基于变化检测的对象分割方法更为适用,但其分割结果往往大于实际运动对象区域,即造成过分割。相比之下,联合时空域对象分割算法虽可得到较好的分割结果,但需要复杂的建模和迭代运算,计算量大。此外,目前的压缩域运动对象提取多是针对MPEG系列的视频编码标准,而针对最新压缩标准H.264/AVC的对象分割算法则研究不多。从视频编码的角度看,H.264/AVC引入了许多不同于MPEG的编码技术,如帧内预测。此时,由于帧内预测块的DCT系数和帧间预测块的DCT系数一样,是由预测残差经变换而得到的,因而不能利用它来提取DC图像。可见,很多现有的基于MPEG的压缩域对象分割算法不能应用于最新的视频标准中。

针对以上的问题,本文提出了一种新的基于H.264/AVC压缩域的时空域运动对象分割(TSMOS)算法。该算法先分别从DCT系数和运动向量的角度,结合形态滤波、权值扩展向量中值滤波

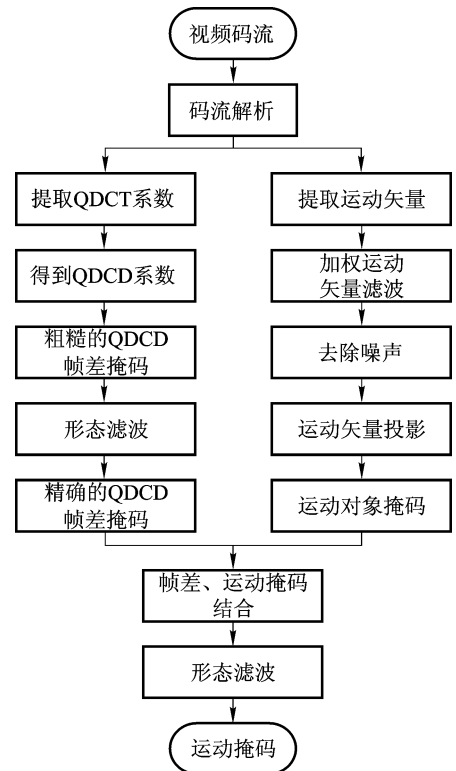


图1 TSMOS算法流程图

Fig. 1 The flowchart of TSMOS algorithm

(WEVM)等技术得到不同的分割掩码,再采用有效的机制合并这些掩码来分割运动对象(如图1所示)。实验表明,该算法具有良好的分割效果,且可以推广应用到 MPEG 压缩域中。

2 时空域运动对象分割算法

本文提出的 TSMOS 算法的基本流程如图1所示,主要包括以下3个部分:(1)在时域中提取运动对象的帧差掩码;(2)在空域中提取对象的运动掩码;(3)合并由时域和空域中分别得到的掩码,并进行后处理。该算法是利用从压缩域中获得的信息,在时域上,采用求取相邻帧帧差 DCT 系数和形态滤波、噪声处理等技术来分割帧差掩码;在空域上,则采用 WEVM 滤波方法,并利用相邻块运动向量的一致性来检测运动块,以及结合运动投影方法来得到运动掩码;最后通过引入一套准则有机地结合时空域的分割结果来获得准确的运动对象。

2.1 时域运动对象分割

压缩码流中的量化 DCT(QDCT)系数是由运动估计的残差经过变换和量化后得到的,其在一定程度上可以反映运动估计的准确性。一个块的 QDCT 系数越大,则表明运动估计越不准确,反之则越准确。H.264/AVC 采用的是基于块的混合预测编码模式,每个块内所有像素都用同一个运动矢量来表征其运动,虽然这对于大片平坦背景或运动一致的区域而言是准确的,且可以得到较小的 QDCT 值,但对于内部包含多种运动的对象,如平移、旋转、缩放等来说,如果用统一的运动矢量来表示一个块内所有像素的运动,则显得粒度太粗,并将造成 QDCT 值较大。根据这个特点,可以通过 QDCD (quantized DCT difference) 系数来分割静止背景下运动对象的大致区域。对于 H.264/AVC 中的每个像素,QDCD 系数 $D(i)$ 的定义如式(1)所示, $Q_{cur}(i)$ 表示当前帧内某个分块内索引为第 i 个像素量化后的 DCT 系数, $Q_{ref}(i)$ 则表示与当前帧相邻的前一帧或前几帧的对应位置像素的量化 DCT 系数。如果一个分块包含 N 个像素,且其 $D(i)$ 满足式(2),则表明该分块是运动对象的候选块。其中 $W(i)$ 表示权值,且权值大小与像素在分块内的 Zig-Zag 扫描位置有关。一般来说,低频位置系数的权值要比高频位置系数的权值要大。 T 为实验阈值, N 为分块中包含的像素点的个数,对于一个 4×4 的分块, N 值为 16。

$$D(i) = |Q_{ref}(i) - Q_{cur}(i)| \quad (1)$$

$$\text{如果} ((\max(D(i)) > 1) \parallel (\sum_{i=1}^N W(i)D(i) > T)) \quad (2)$$

候选块

否则

非候选块

对于某些视频序列来说,运动对象在场景推进的过程中,会出现局部停滞或运动缓慢的现象,此时需要通过累积当前帧前面 l 帧的 QDCT 值来获得较完整的对象轮廓。 l 的选择是由对象的运动速度和运动剧烈程度来决定的。 l 越大,则提取的轮廓越完整,同时引入的噪声越多; l 过小,则会丢失部分对象边界,进而影响后续的分割结果。可以通过式(3)来确定 l 的值,其中 Ω 是当前帧中具有运动向量的分块的集合, m 表示 Ω 中分块的数目, i 为当前帧中具有运动矢量的分块的索引值, K_i 和 V_i 分别为与索引值为 i 的分块对应的 DC 系数值和运动矢量, T_1^{MV} 和 T^{DC} 为对应于运动矢量和 DC 系数计算时的门限。

$$l = \begin{cases} 1 & \frac{1}{m} \sum_{i \in \Omega} |V_i| > T_1^{MV}, \frac{1}{m} \sum_{i \in \Omega} |K_i|^2 > T^{DC} \\ 3 & \text{其他} \end{cases} \quad (3)$$

经过 QDCD 判决后所得到的候选运动对象轮廓是以块为单位的粗糙帧差掩码,对此帧差掩码进行形态滤波,如膨胀、收缩、开闭等运算,即可以得到进一步相对精确的帧差掩码。但此时可能存在许多离散的噪声点,考虑到对象的连通性,根据 MRF 运动场原理,如果候选运动块的邻域内的大多数块表现为运动块状态时,则当前块才确定为运动块,否则认为此块为背景块。这样可有效消除通过多帧 QDCD 值累计来获取运动对象帧差掩码所带来的噪声。

图2是利用 QDCD 系数分割运动对象的结果示意图,由该图可以看出,QDCD 方法能够较准确地分割出运动对象的大致区域,甚至连较小的运动对象也能检测出来,如图2(d)中右边一个人。

一般来说,运动对象的边界块很难在参考帧中找到精确的匹配块,这将导致运动估计不准确,因为预测残差相对较大,所以采用 QDCD 方法获得的帧差掩码能较准确地分割出运动对象的轮廓。然而对于大片平坦的运动对象区域及对象局部停滞的区域来说,即使运动估计较准确,帧差掩码仍可能检测不出来,而下一小节介绍的空域运动对象分割方法正好可以弥补这方面的不足。

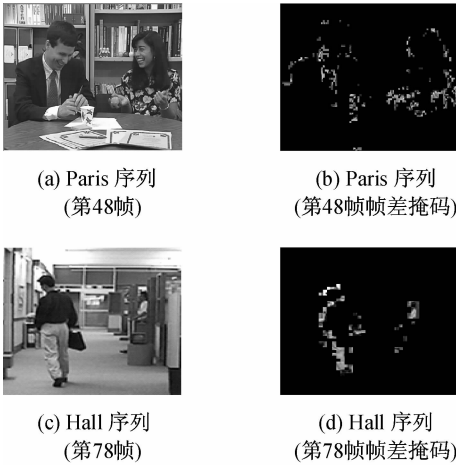


图 2 QDCD 系数分割结果

Fig. 2 The segment results using QDCD coefficients

2.2 空域运动对象分割

空域运动对象分割主要是利用相邻块运动的相关性特征来分割运动对象。由于运动相关性是指物体内部的运动大体上是一致的,因此组成运动对象的分块的运动矢量的大小和方向具有相关性。压缩码流中包含的运动矢量是在编码端通过运动估计得到的,但由于 MPEGx 和 H. 26x 等压缩标准采用的是块匹配的运动估计技术,其主要考虑的是码流的压缩效率,而非运动估计的准确性,因此运动矢量并不能代表对象真实的运动情况。此外,由于外界环境变化、摄像头移动,以及噪声的引入都可能在码流中产生相应的运动矢量,因此在运动分割前必须尽可能将这些不利于运动分割的运动矢量滤除。

在滤波之前需要将运动矢量进行时域和空域的归一化,具体方法可参考文献[4]。此外,由于随机噪声带来的运动矢量较小,因此可简单通过设定一个门限的方法来滤除噪声。超过该门限的即认为是候选运动块,接下来的滤波处理只在候选运动块中进行,这就可在一定程度上降低运算量。本文采用的滤波算法是 WEVM 算法,因为其能有效地去除随机噪声,同时保留运动对象的细节信息,如对对象边界等。WEVM 算法的定义如式(4)^[9]所示,其中权值向量中值滤波结果 V_{WVM} 和加权滤波结果 V_{WAVE} 的定义分别如式(5)和式(6)所示,

$$\text{如果 } \sum_{i=1}^N W_i \|\omega \otimes (V_i - V_{WAVE})\|_p < \sum_{i=1}^N W_i \|\omega \otimes (V_i - V_{WVM})\|_p \quad (4)$$

则 $V_{WEVM} = V_{WAVE}$

$$\text{否则 } V_{WEVM} = V_{WVM}$$

$$\text{其中, } V_{WVM} \in V_i \quad i = 1, 2, \dots, N$$

$$\sum_{i=1}^N W_i \|\omega \otimes (V_i - V_{WVM})\|_p \leq \sum_{i=1}^N W_i \|\omega \otimes (V_i - V_j)\|_p \quad j = 1, 2, \dots, N \quad (5)$$

$$V_{WAVE} = \frac{\sum_{i=1}^N W_i V_i}{\sum_{i=1}^N W_i} \quad (6)$$

式中, V_i 表示第 i 个候选块的运动矢量, $\|\cdot\|$ 表示范数运算, W_i 表示参与滤波的各个运动矢量的权重, $\omega = \{\omega_1, \omega_2\}$ 表示运动矢量的权重向量, \otimes 表示矢量内积。

在本文实验中,取 $p=2$,滤波窗口大小为 3×3 ,位于窗口中心位置的运动矢量的权值为 2,其他位置的权值为 1,这样就可以有效地保证较小运动对象的运动矢量不被滤掉;运动矢量的 x 分量和 y 分量的权值均为 1,即滤波强度相同。对于滤波后的每个块的运动矢量,若满足 $\|V_i\| \geq T_2^{MV}$,则检测为运动掩码 R^{MV} 。这里 $R^{MV} = \{r_1, r_2, r_3, \dots\}$, r_i 表示连通区域。但此时得到的运动掩码中仍可能存在噪声区域,例如光强发生变化的背景区域,而采用比重法则可有效去除这种情况(如式(7)所示),其中 Q_i 为帧差掩码投射到运动掩码上与 r_i 交叠的区域, S 为该区域所包含的面积。

$$r_i = \begin{cases} \text{噪声} & \frac{S(Q_i)}{S(r_i)} < 0.12, S(r_i) < 64 \\ \text{对象} & \text{其他} \end{cases} \quad (7)$$

一般来说,物体的运动在时间上具有连续性和一致性,且由于连续两帧之间的时间间隔较短,故可利用运动投影技术来完善对象的分割结果。运动投影分为前向投影和后向投影两种,本文进行空域对象分割是采用后向投影,而时空域对象分割则采用前向投影,即先将已经得到的上一帧的分割结果后向投影到当前帧,再结合对象的边界轮廓信息得到当前帧完整的运动掩码。设当前帧为 F_n ,前一帧为 F_{n-1} , F_{n-1} 中的任一帧间预测块为 $B_{n-1}(i, j)$, i, j 表示此块在帧中的水平和垂直坐标。若块 $B_{n-1}(i, j)$ 的运动矢量为 $V_{n-1}(i, j) = (v_{n-1}^x, v_{n-1}^y)$,则此块投影到当前帧 F_n 中的块为 $B_n(p, q)$,这里 $p = i - v_{n-1}^x/s, q = j - v_{n-1}^y/s$,这里 $s = w/g$,其中 w 为分块像素宽度, g 为运动矢量精度。对于 H. 264/AVC 而

言,由于分块模式的最小宽度为4,运动矢量精度为 $1/4$,所以对应的 s 为16。

2.3 时空域运动对象分割

由前两节的论述可知,通过时域运动对象分割可得到比较精确的运动对象轮廓,但对于运动一致的区域,由于运动估计较准确,如果量化后的DCT系数较小,则不能通过帧差掩码反映这部分运动区域,如Hall序列中左边一个人的后背,Paris序列中右边一个人的头发(参见图1(a)和图1(c))。而空域分割法的思路则是基于对象运动的相关性,即运动一致的区域属于同一个运动对象。对于刚体运动对象来说,一方面对象内部的运动应该具有一致性,表现为对象内部相邻块的运动矢量的方向和幅度应相差不大;另一方面由于相邻帧间的时间间隔较短,运动对象近似于匀速运动,所以时域上对应块的运动矢量的方向和幅度也具有相似性。然而由于运动对象的边界块往往与其内部块运动不一致,故基于运动矢量的分割方法不能准确划分出对象的轮廓,而且,对于H.264/AVC帧间预测中的skip(不编码模式)块和帧内预测块而言,空域分割方法也无法进行运动检测,因此,需要将时域分割出的帧差掩码与空域分割出的运动掩码进行有机的结合,才可以得到完整准确的运动对象掩码。

第 n 帧的帧差掩码 R_n^{OD} 和运动矢量掩码 R_n^{MV} 的结合准则如下:设 $N(B_n(i, j))$ 表示第 n 帧中位置为 (i, j) 处分块的邻域,且 $N_1(B_n(i, j)) < N_2(B_n(i, j))$, $B_{ref}(r, s)$ 表示与块 $B_n(i, j)$ 对应的参考块, O 为对象。

如果 $(B_n(i, j) \in R_n^{OD}) \quad B_n(i, j) \in O$
 否则,如果 $(B_{ref}(r, s) \in R_{n-1}^{OD}) \quad B_n(i, j) \in O$
 否则,如果 $((B_n(i, j) \in R_n^{MV}) \&\& (B_n(p, q) \in O) \&\& (B_n(p, q) \in N_1(B_n(i, j)))) \quad B_n(i, j) \in O$
 否则,如果 $((B_n(i, j) \in R_n^{MV}) \&\& (B_n(r, s) \in O) \&\& (B_n(r, s) \in R_n^{MV}) \&\& (B_n(r, s) \in N_2(B_n(i, j)))) \&\& (\|V_n(i, j) - V_n(r, s)\| < T_3^{MV})$

$B_n(i, j) \in O$

否则 $B_n(i, j) \notin O$

时空域结合分割算法得到的运动对象掩码还可能包括噪声带来的白色掩码,以及由运动性质决定的出现在对象内部的黑色掩码。一般来说,因噪声区域的面积相对于运动对象而言较小,故可以通过计算各个连通区域的面积,把面积小于一定阈值的白色掩码视为噪声去除。对于掩码中由于运动对象

局部运动的不连续性而产生的空洞,则可以采用二次扫描的方法进行有效的填充。在进行运动对象提取时,首先对分割的掩码从左到右进行搜索,将最左端模型点和最右端模型点之间的像素点划为候选对象区域;再对每一列从上到下进行搜索,将最上端模型点和最下端模型点间的像素点划为候选对象区域;然后将两种方式扫描获得的候选对象区域相与即可有效地提取封闭的视频对象;最后采用形态开闭运算对分割的掩码进行后处理,以有效平滑边界毛刺来得到完整准确的运动对象掩码。这里用到的虽然只是较为常见的膨胀和腐蚀方法,但是却明显地提高了提取运动对象的准确性。

3 实验结果

本文选取Hall和Paris两个序列作为实验场景序列,序列格式为CIF(352×288 pixels),序列长度均为150帧。序列中的运动对象具有多个方向的运动,Hall序列中包含水平运动(与镜头方向垂直)、垂直运动(与镜头方向平行)、静止以及对象的局部运动等,其符合视频监控的特点;而Paris序列则属于头肩序列,其符合可视电话、视频会议的特征。

压缩域对象分割算法对视频数据较敏感,若采用不同的编码器设置,则可能会导致不同的分割结果。本文采用H.264/AVC的参考编解码软件JM86对序列进行编码,帧结构为IPPP,帧率为25fps,参考帧数为1帧,量化步长 $QP=36$ 。算法中的各参数设置为: $R(i)=1, T=2, T_{DC}=2, T_1^{MV}=8, T_2^{MV}=3, T_3^{MV}=4, N_1$ 为 3×3 大小邻域, N_2 为 5×5 大小邻域。本文还将TSMOS算法得到的分割结果与文献[7]算法的分割结果进行了比较,图3是TSMOS算法和文献[7]算法分割结果的主观对比。

此外,本文采用分割精度和分割率^[10]两个指标来衡量算法的客观性能。分割精度是指检测到的真实前景运动对象占检测到的所有前景对象的百分比;分割率是指检测到的真实前景运动对象占实际前景运动对象的百分比。算法分割精度越高,则误检率就越低;分割率越高,则表示有越多的运动对象被分割出来。为了统计方便,本文以 4×4 块为单位来计算算法的分割精度和分割率。图4和图5给出了时空域运动对象分割算法和文献[7]算法分别针对Hall和Paris两个序列的运动对象进行分割的分割率、分割精度的对比实验结果,其中横坐标为视频

序列的帧数,两种算法的平均分割率和分割精度见表 1。从图 4 和图 5 中可以看到,Hall 和 Paris 两个序列在分割率和分割精度上呈现不同的趋势,这是由对象的运动波动所造成的。如 Hall 序列中的运动对象在 110 帧至 130 帧间突然静止,而 Paris 序列则在整个过程中两个对象都有较大的运动幅度,这都使得对应的分割率和分割精度出现波动。

从以上对比实验结果来看,TSMOS 算法的分割效果较文献[7]算法的分割结果要好。在复杂度方面,文献[7]算法在聚类中是采用迭代和线性回归

的方法,由于其迭代次数事先不能确定,同时计算中使用了较多的求指数和对数运算,而且在对边界块进行区分时,还需要统计所有对应块的概率,因此复杂度很大。而 TSMOS 算法则不需要复杂的数学建模、聚类和迭代运算,所以从理论上分析,计算量要比文献[7]算法小。鉴于代码优化等问题,本文未给出算法具体的运行时间比较。此外,还可看出算法的分割精度和分割率是相互制约的,若分割率高,分割精度则相对较低,反之亦然。

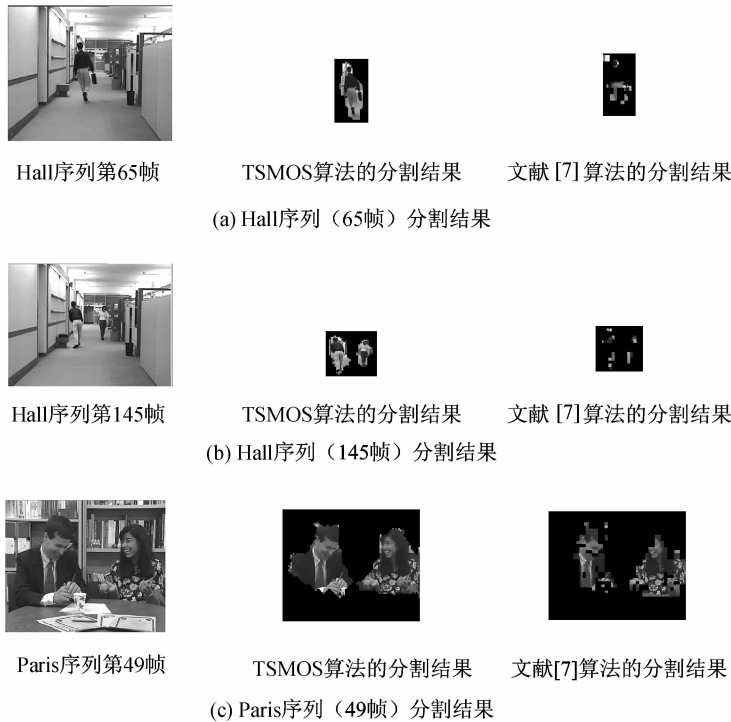


图 3 主观分割结果对比示意图

Fig. 3 The subjective segment results between TSMOS and [7] algorithm

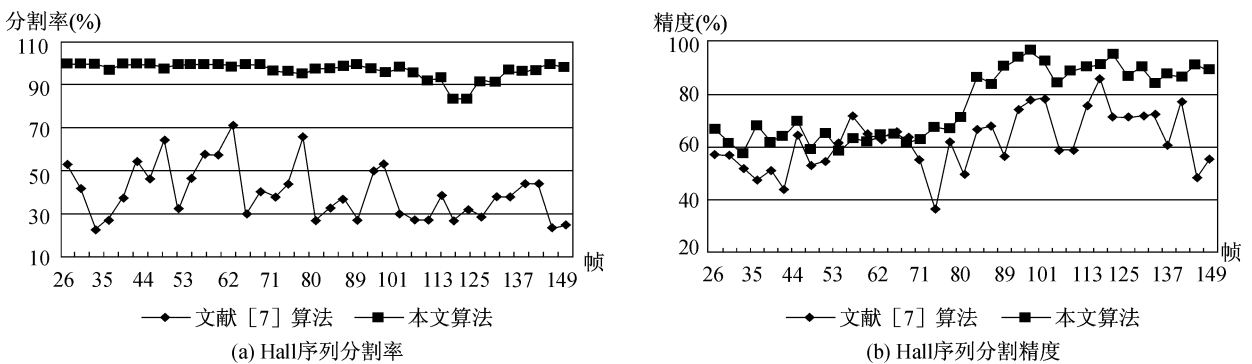


图 4 Hall 序列分割率、分割精度对比结果

Fig. 4 The objective segment results between TSMOS and [7] algorithm on sequence Hall

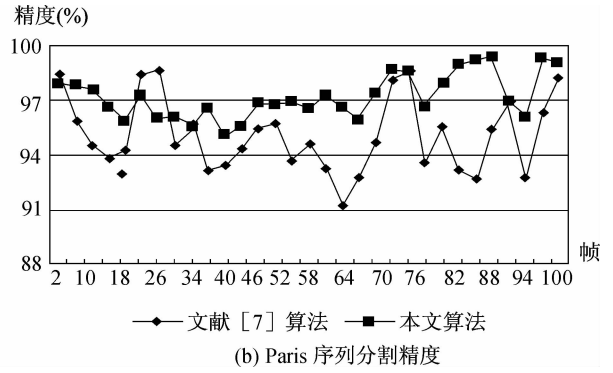
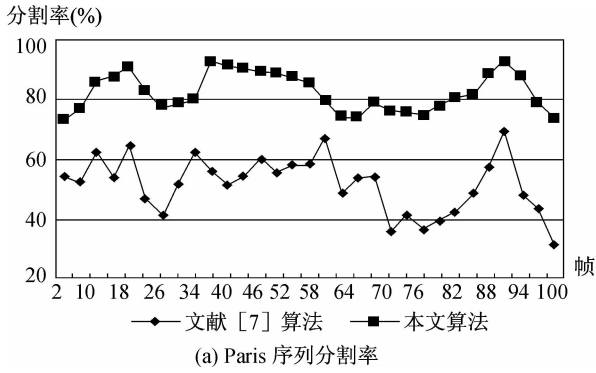


图 5 Paris 序列分割率、分割精度对比结果

Fig. 5 The objective segment results between TSMOS and [7] algorithm on sequence Paris

表 1 Hall, Paris 序列平均分割率、分割精度

Tab.1 The average objective segment results between TSMOS and [7] algorithm

序列	TSMOS 算法		文献[7]算法	
	分割率 (%)	精度 (%)	分割率 (%)	精度 (%)
Hall	96	76	40	62
Paris	82	97	52	95

4 结 论

本文提出了一种新的时空域相结合的运动对象分割方法,该新方法充分利用了压缩码流中的 DCT 系数,并首次引入 QDCD 的概念,结合形态滤波方法在压缩域中提取对象轮廓;然后本文还对压缩域中的运动向量,采用 WEVM 滤波技术来得到平滑的向量场,并利用相邻块的运动相关性来检测运动块;最后提出一套简单有效的准则,用来合并由 DCT 系数和运动向量的分割结果。实验表明,本文算法具有较高、较稳定的分割率和分割精度。下一步的研究将结合像素域的分割技术改进本文算法,以得到更理想的运动对象分割结果。

参考文献 (References)

- Mezaris V, Kompatsiaris I, Boulgouris N V, et al. Real-time compressed-domain spatiotemporal segmentation and ontologies for video indexing and retrieval[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2004, 14(5):606-621.
- Nakajima Y, Yoneyama A, Yanagihara H, et al. Moving object

- detection from MPEG coded data [A]. In: Proceedings of SPIE Conference on Visual Communications and Image Processing [C], San Jose, CA, USA, 1998, 3309: 988-996.
- Babu R V, Ramakrishnan K R, Srinivasan S H. Video object segmentation;a compressed domain approach[J]. IEEE Transactions on Circuits Systems for Video Technology, 2004, 14(4):462-474.
- Liu Zhi, Xu Jian-feng, Lu Yu, et al. A real-time H.264 compressed domain based moving object segmentation algorithm[J]. Journal of Image and Graphics, 2006, 11(11): 1614-1618. [刘志, 徐剑峰, 陆宇等. 基于 H.264 压缩域的实时运动对象分割算法[J]. 中国图象图形学报, 2006, 11(11): 1614-1618.]
- Sukmarg O, Rao K R. Fast object detection and segmentation in MPEG compressed domain[A]. In: Proceedings of TENCON 2000 [C], Kuala Lumpur, Malaysia, 2000: 364-368.
- Zeng W, Gao W, Zhao D. Automatic moving object extraction in MPEG video [A]. In: Proceedings of the IEEE International Symposium on Circuits and Systems 2003 [C], Bangkok, Thailand, 2003, 11:524-527.
- Eng H L, MA K K. Spatiotemporal segmentation of moving video objects over MEPG compressed domainA[A]. In: Proceedings of the IEEE International Conference on Multimedia and Expo 2000 [C], New York City, NY, USA, 2000, 3: 1531-1534.
- Zeng W, Du J, Gao W, et al. Robust moving object segmentation on H.264/AVC compressed video using the block-based MRF model [J]. Real-Time Imaging, 2005, 11(4): 290-299.
- Kassam S A, Aburdene M. Multivariate median filters and their extensions[A]. In:Proceedings of the IEEE International Symposium on Circuits and Systems [C], Singapore, 1991: 85-88.
- Ashraf M A A, Bashar M A A, Lee S Y. Fast and robust object detection framework in compressed domain[A]. In: Proceedings of the IEEE Sixth International Symposium on Multimedia Software Engineering 2004 [C], Miami, FL, USA, 2004: 210-217.