

# 一种 MPEG 压缩域的运动对象快速提取算法

周启亚<sup>1</sup>, 何春梅<sup>2</sup>, 杨高波<sup>3</sup>

(1. 湖南铁路科技职业技术学院 信息技术系, 湖南 株洲 412000;

2. 南京理工大学 计算机科学与技术学院, 南京 210094;

3. 湖南大学 计算机与通信学院, 长沙 410082)

(zhoukia@126.com)

**摘要:**针对目前压缩域直接得到的运动矢量过于稀疏且不可靠, 导致运动对象分割精度不高等问题, 提出一种适合于 MPEG 视频流的压缩域运动对象提取算法。它先对 MPEG 码流进行熵解码得到的运动矢量场进行可靠性度量校正、置密和滤波处理, 并采用基于块的区域增长算法得到每个对象的大致运动区域; 然后解码属于运动区域的块, 并把被还原的每个位置的像素值分布视为高斯分布, 用统计的方法建立其高斯模型, 通过阈值判断可靠地提取出最终具有像素精度运动对象。仿真实验表明, 该算法可以有效地去除运动矢量噪声的影响, 分割精度显著提高, 并且具有较高的处理速度。对于 QCIF 格式的视频码流, 每秒可以处理约 50 帧。

**关键词:**运动对象提取; 区域增长; 高斯模型; MPEG

**中图分类号:** TP391.41 **文献标志码:** A

## New algorithm for moving object extraction in MPEG compressed domain

ZHOU Qi-ya<sup>1</sup>, HE Chun-mei<sup>2</sup>, YANG Gao-bo<sup>3</sup>

(1. Department of Information Technology, Hunan Railway College of Science and Technology, Zhuzhou Hunan 412000, China;

2. School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing Jiangsu 210094, China;

3. School of Computer and Communication, Hunan University, Changsha Hunan 410082, China)

**Abstract:** The motion vector field directly obtained from the video stream is sparse and unreliable. It will lead to low accuracy of the extracted motion objects. This paper proposed a novel compressed-domain motion object extraction algorithm for MPEG video stream. The motion vector field directly obtained from the video stream was firstly processed by reliability test, densification and filtering. Block-based region merging was used to obtain the rough region of every motion object. These motion regions were decoded to get the original pixel values, which were modeled with Gaussian mixture model by statistical method. The final accurate motion objects were extracted by threshold decision. Simulation results demonstrate that the proposed approach can get rid of the noisy motion vectors, then the accuracy of segmentation is greatly improved, and a high processing speed is achieved. For video stream in QCIF format, 50 frames per second can be achieved.

**Key words:** moving object extraction; region-growing; Gaussian model; MPEG

## 0 引言

视频中运动目标的提取是视频压缩、视频内容分析等应用的基础, 具有重要的应用价值。最初的研究工作是从 MPEG-4 引入的基于内容的视频编码开始的, 主要在像素域进行, 要求在编码前将视频分割为具有语义意见的实体。由于视频数据量大, 通常是以压缩的形式进行存储和传输, 近年来运动对象分割提取的研究热点逐步由像素域过渡到压缩域<sup>[1]</sup>。

直接在压缩域进行运动目标提取可以显著降低运算量, 有助于实现实时分割。一方面, 从 MPEG 视频流中易于得到用于运动补偿的运动矢量, 可以省略像素域分割算法复杂耗时的运动估计过程; 另一方面, 采用部分 DCT 系数或仅采用直流 DCT 系数, 节省完全解码所需的时间, 并使算法在相对较小的尺寸 (DC 图像) 上进行。文献[2]提出一种基于变化检测的 MPEG 压缩视频对象分割方法。该方法利用 DC 图重建技术获取压缩视频帧的图像信息, 并利用高阶统计技术检

测运动的宏块。为了有效确定背景噪声, 还提出了一种自适应阈值计算方法来自动检测背景区域, 并由此估算当前帧的背景噪声。当压缩视频码率较低, 或出现大部分运动估计失效的宏块时, 该算法能稳定地检测出运动对象。针对 MPEG 视频流中的运动矢量与每个宏块的真实运动不尽一致的缺陷, 文献[3]提出了一种基于 Kernel 的多线索对象分割算法 (KMC)。它主要利用运动信息, 检测和标定摄像机的运动, 然后找到运动对象的核。分割从核 (核被定义为具有可信运动矢量的纹理区域) 开始进行检测和合并, 并利用色彩和纹理信息以进一步改善分割结果。文献[4]提出了一种基于贝叶斯 (Bayesian) 统计模型的分割算法, 从 MPEG 视频流中分割与跟踪运动区域。它针对从已编码的视频流分析有语义意义的运动实例和提取时, 空视频结构信息的困难, 算法通过 Gibbs-Markov 随机场理论与贝叶斯估计框架扩充 stochastic 的运动一致模型, 以保证分割结果符合预设的可能性、局部/区域一致性和轮廓平滑程度。我们此前提出在空域上提取 I 帧 DC + 2AC 低分辨率灰度图对象的轮廓特征, 再融合时域上 P

收稿日期: 2008-09-10; 修回日期: 2008-11-20。 **基金项目:** 国家自然科学基金资助项目 (60472061)。

**作者简介:** 周启亚 (1980-), 男, 湖南湘乡人, 硕士, 主要研究方向: 视频编码、视频分割; 何春梅 (1981-), 女, 湖南邵阳人, 博士研究生, 主要研究方向: 模糊智能系统; 杨高波 (1974-), 男, 湖南岳阳人, 教授, 博士, 主要研究方向: 图像处理、视频信号处理。

帧运动矢量的投影,结合空域和时域的信息,快速准确地实现了直接从压缩域中提取运动对象<sup>[5]</sup>。

显然在大多数已有的算法中,运动矢量是压缩域用于提取运动对象的主要线索,然而,视频码流中包含的运动向量用于运动补偿,通常采用块匹配的方法获得,与真实的运动向量相比存在很多的噪声。而且,相对于像素域的光流场,压缩域得到的运动矢量场要稀疏得多。直接使用这些运动矢量进行分割会造成运动目标的错分割和漏分割。因此,压缩域运动对象提取的关键在于运动矢量的致密化和可靠性判决。

本文针对以上问题,提出一种适合于 MPEG 视频流的压缩域运动对象提取算法。有效的预处理操作使得用于运动对象提取的运动矢量场更加可靠,大部分操作基于压缩域处理使得算法具有实时性。鲁棒性与实时性是该算法的主要优势。

## 1 算法概述

MPEG 主要采用帧间基于块的匹配技术来消除时域冗余而获得较高的编码效率。帧间编码又分为两种方式:前向预测编码和双向预测编码。MPEG-2 中对于 P 帧,有前向预测、帧内和跳过未编码三种类型的宏块<sup>[6]</sup>。除了后两种情况,进行前向预测编码的 P 帧宏块都有一个运动矢量,所有宏块的运动矢量视为运动矢量场。由于一个宏块(16×16 大小的像素块)仅有 1 个运动矢量,如图 2(a)是 Table 序列第 118 帧提取的原始运动矢量场(编码为 P 帧),每个小蓝点表示有 1 个运动矢量,可以看出,提取的运动矢量场非常稀疏。

同时,从压缩流中提取的运动矢量是在编码端通过最小化 SAD 而计算求得的,目的是为了获取尽可能高的编码效率,因此,这些运动矢量并不一定能准确反映物体的真实运动,只能视为有噪声的运动矢量。具体地,运动矢量的噪声通常表现为两种形式:1)伪运动矢量,即物体没有运动但在运动矢量中却有表示运动的运动矢量部分;2)漏运动矢量,即物体本身具有真实运动,但在运动矢量中却没有表示运动的运动矢量部分。这样,直接根据初始运动矢量场提取的运动区域与实际的运动对象会不一致,会导致错分割和漏分割。如图 2(c)和图 2(d)所示,它们分别表示初始运动区域范围和直接对这些初始运动区域部分进行解码得到的原始像素值,显然图中运动员左手的部分区域被漏掉,这就是由于未经处理的运动矢量场受噪声的影响,初始运动区域与实际的运动对象不大一致。

针对上述压缩域运动矢量场过于稀疏且因噪声影响不可靠的问题,本算法的研究重点是如何先对其进行可靠性度量和致密化,再用于提取运动对象,并尽可能提高提取对象的精度。算法的流程如图 1 所示,对于由 MPEG1/2 压缩的视频码流,先进行压缩域基于图像组(GOP)的处理。通过部分解码提取每个 GOP 中第一个 P 帧的运动矢量,得到原始的运动矢量场,然后进行可靠性度量与校正、置密和滤波处理,得到相对准确的运动矢量场。然后,采用基于块的区域生长算法得到每个视频对象对应的大致运动区域。在接下来的像素域处理时,进行有效的折中:通过部分解码视频流,即对每帧属于前面得到的大致运动区域内的块解码得到它们的像素值,然后用高斯统计模型进行建模,提取出感兴趣的运动对象。

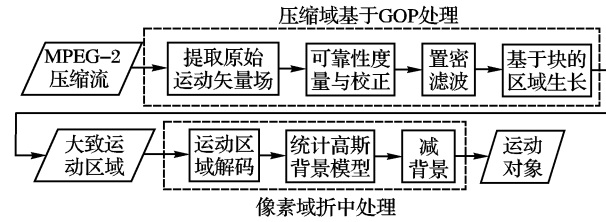


图1 算法流程

## 2 算法实现

### 2.1 运动矢量的可靠性度量与处理

在提取原始运动矢量场时,对于 P 帧,为了得到一个完整的运动矢量场,对跳过未编码的宏块,赋值为前一宏块相同的值,对帧内编码的宏块,取值为邻域 8 个宏块的均值。对于 B 帧,除了要进行前向预测,还要进行后向预测,然后插值得到双向预测的结果,编码比 P 帧复杂,算法暂未考虑。本文首先对运动矢量进行可靠性度量,求得那些表示物体真实运动的运动矢量,对那些是真实运动部分又没运动矢量的块进行校正。然后在空域上对运动矢量场进行置密,把宏块的运动矢量赋给其组成的四个块,使每个块(8×8 像素)具有一个运动矢量;在时域上进行运动矢量的迭代累积投影,以增强显著的运动信息。再对运动矢量场进行空、时域滤波,该滤波方法比单独进行空域中值滤波与时域滤波的效果好。在此需说明的是,运动矢量的投影方法不同,得到的最终运动矢量场会有所差异,在实验中我们采用按面积比例投影的方法<sup>[7]</sup>。

### 2.2 可靠性度量规则

#### 2.2.1 残差度量

对 P 帧前向预测编码的宏块,其编码数据为当前宏块进行运动补偿后的残差数据及相应的运动矢量,残差数据经 DCT 变换再进行熵编码。残差数据可以用来度量当前块与参考块的匹配程度,如果残差的平方和越大,当前块与参考块的匹配度越小,该块的运动矢量越不可靠。该规则定义如下:

$$weight\_one = \exp\left(-\sum err^2/K\right) = \exp\left(-\sum err_{DCT}^2/K\right) \quad (1)$$

其中  $K$  为残差数据的方差。第二个等式根据 Parseval 关系式得到,即信号经 DCT 后整体能量不会发生变化,  $err_{DCT}$  可直接从码流中得到,  $\sum err_{DCT}^2$  易于计算。

#### 2.2.2 空域连续性度量

由光流场理论可知,同一对象区域的运动矢量在空域上具有连续性,即光流场具有平滑约束性,这一结论在下节中也会用到。而对象边缘运动矢量和噪声运动矢量在强度或方向上通常会发生突变,因此,我们利用一个块的运动矢量与其相邻块运动矢量的连续性对运动矢量的可靠性进行度量,若块与其邻域块的运动矢量满足平滑性,该块的运动矢量越可靠。其规则定义如下:

$$weight\_two = \exp(-|mv - med\{neighbor\ of\ mv\}|) \quad (2)$$

其中  $med\{\cdot\}$  为中值滤波器。

图 2 中,(a)为第 118 帧的原始运动矢量场,编码为 P 帧;(b)为对(a)置密后的运动矢量场;(c)为初始运动区域示意图,黑色表示运动区域;(d)为初始运动区域部分解码,由于未经处理的运动矢量场受噪声的影响,初始运动区域与实际

的运动对象不大一致;(e)为处理后的运动区域示意图;(f)为图(e)部分解码后的结果。

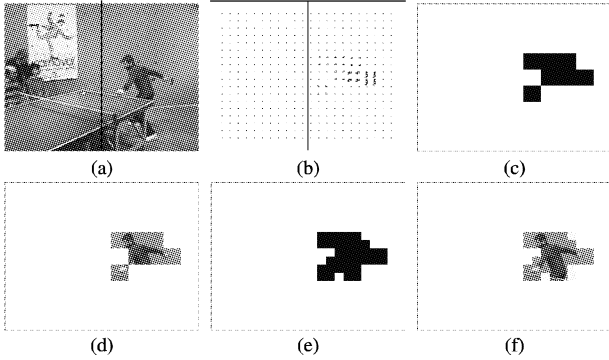


图2 Table\_Tennis 序列运动矢量场处理过程

### 2.2.3 质地特征度量

对于大区域具有相同质地又纹理丰富的对象,这部分的运动矢量一般呈现随机性,与物体的真实运动不一致,但这些运动矢量又可靠地表达了对象的运动信息,因此我们考虑用对象的质地特征来度量其运动矢量的可靠性。DCT 后块的 AC 系数可以表征块的质地特征,但对于大区域具有相同质地又纹理丰富的情况不能用 AC 系数来表征。文献[8]中用 MAD (Mean Absolute Difference) 规则来衡量块间的相似度,求出当前块与周围一定范围内最相似的块,即 MAD 取最小值的块,此时相应的 MAD 值记为 minMAD。若 minMAD 越大,说明当前块与周围块差异越大,那么运动搜索时错误匹配的可能性越小,相应的运动矢量越可靠。规则描述如下:

$$weight\_three = \alpha \cdot \min MAD \quad (3)$$

其中,  $\alpha$  为正态化系数。

### 2.3 空时域滤波

用  $V^t(i, j)$  表示第  $t$  帧位置为  $(i, j)$  处宏块的运动矢量,  $N(i, j)$  为与宏块  $(i, j)$  空域相邻一定范围内的所有宏块,定义  $SN = \{V^t(i, j) | (i, j) \in N(i, j)\}$ ,  $SN$  即为与宏块  $(i, j)$  空域相邻一定范围内所有宏块运动矢量的集合。对于  $SN$  中的任意运动矢量  $v$ ,映射到时域相邻帧中的某些块,把这些块的所有运动矢量集合记为  $TN(v)$ ,显然  $TN(v)$  为  $v$  的函数。空时域滤波方法描述为<sup>[9]</sup>:

$$F^t(i, j) = \underset{v}{\operatorname{argmin}} \sum_{y \in SN} (v - y)^2 + \sum_{z \in TN(v)} (v - z)^2 \quad (4)$$

该滤波方法由于同时考虑了当前矢量与空域和时域相邻矢量间的关系,能取得较好的滤波效果,比单独用空域中值滤波或时域滤波好,空时域滤波是对基本的向量中值滤波的扩展应用。

### 2.4 运动矢量基于块的区域增长

#### 2.4.1 运动矢量一致性模型

理想情况下位于同一刚性物体上的块具有相同的运动矢量<sup>[10]</sup>。令  $B(p, q)$  表示当前帧中位置为第  $p$  列,第  $q$  行处的块,块大小为  $W \times W$ ,其运动矢量记为  $MV(p, q) = [MV_x(p, q), MV_y(p, q)]^T$ 。则对于属于同一对象的任意两个块  $B(p, q)$  和  $B(p', q')$ ,下式成立:

$$(p'W - pW)(MV_x(p', q') - MV_x(p, q)) + (q'W - qW)(MV_y(p', q') - MV_y(p, q)) = 0 \quad (5)$$

简记为:

$$[p' - p, q' - q][MV(p', q') - MV(p, q)] = 0 \quad (6)$$

也就是说,若两块的运动矢量满足式(6),则认为这两个块位于同一对象。然而,上面的运动一致性模型在理想情况下才成立,考虑实际中由于运动矢量与物体的真实运动间的偏差及量化误差的影响,在我们的算法中块的一致性衡量标准修改如下:

$$d(B(p, q), B(p', q')) = | [p' - p, q' - q][MV(p', q') - MV(p, q)] | < \varepsilon \quad (7)$$

其中  $\varepsilon$  为阈值,不同的序列取值不同,在实验中我们取值为 0.1~0.8。

#### 2.4.2 区域增长

对前面直接从压缩流中提取并经处理的运动矢量场,运用运动矢量一致性模型(式(7))就可实现运动分割,即把视频帧根据其各块运动矢量一致性的特点分割成不同的对象或区域。这种根据块运动一致性的分组方法来实现运动分割,其本质是一种对运动矢量基于块的区域增长算法。

实现区域增长,关键是种子点的选择与增长规则的确定。与一般的区域增长算法不同,本算法增长过程操作的基本元素是块的运动矢量,实验中我们选择与其直接相邻的四个块的运动矢量严格满足运动一致性模型的那个块作为种子点,即种子点块的运动矢量  $B(p_s, q_s)$  满足:

$$d(B(p_s, q_s), B(p_s + i, q_s + j)) < \varepsilon; \quad \forall i, j \in \{-1, 0, 1\}, |i| + |j| = 1 \quad (8)$$

在确定了种子点后,种子点与其相邻的四个块首先进行区域增长形成新的区域  $R = \{B(p_s, q_s), B(p_s - 1, q_s), B(p_s, q_s - 1), B(p_s + 1, q_s), B(p_s, q_s + 1)\}$ ,接下来对与  $R$  邻接且未作处理的块  $B(p, q)$  进行增长,增长规则为:

$$\frac{\sum_{B(p_r, q_r) \in R} d(B(p, q), B(p_r, q_r))}{\|R\|} < \varepsilon \quad (9)$$

其中  $\|R\|$  表示已增长的区域中块的数目,  $B(p_r, q_r)$  表示已增长的区域中的块。若上式成立,把块  $B(p, q)$  加入该区域,形成新的区域  $R$ 。重复以上过程直至区域  $R$  不再增长,接着有两种可能的处理情况:1) 若还存在未处理的块,再选择新的种子点用相同的规则进行增长;2) 若还存在未处理的块且对当前增长规则区域不再增长,则调节阈值  $\varepsilon$  再进行增长,直至所有块都分组到各自相应的区域,区域增长完毕。

显然在上述确定种子点的过程中,种子点周围需存在四个直接相邻的块满足运动一致性模型,否则不能找到种子点。当运动物体的跨度不超过 5 个块,如对物体开始进入场景或将离开场景的情况,用上面的方法不能正确处理。对这种情况种子点的选取操作简化如下,若一个块与竖直相邻的两块满足运动一致性模型,我们选择该块作为种子点,即种子点块  $B(p_s, q_s)$  的运动矢量满足:

$$d(B(p_s, q_s), B(p_s, q_s + i)) < \varepsilon; i = \pm 1 \quad (10)$$

再用上面的方法进行区域增长,得到运动一致性区域。这样对于对象进入场景或离开场景的情况也能作出正确提取。

#### 2.5 运动区域部分解码

经上述运动矢量基于块的区域增长过程,我们得到运动对象的大致区域,由于压缩域的信息只具有块的精度,提取的区域还包含较多的背景部分信息。为了获取像素级精度的对

象边缘,在此我们先对属于运动区域的块进行离散余弦逆变换(Inverse Discrete Cosine Transform, IDCT),把运动区域还原成像素域。IDCT 是一个非常耗时的过程,值得注意的是我们只对部分块进行操作,其基本思想是以这部分时间代价,再借用像素域运动对象提取的方法来换取像素级的对象精度,也就是在时间与精度间做的折中。

## 2.6 基于高斯模型的运动对象提取

对运动区域进行部分解码还原成像素域后,我们借用像素域中的方法来提取运动对象。像素域提取运动对象的方法有很多种,主要有基于帧间差分的方法和基于减背景的方法,帧间差分的方法利用对连续两帧或多帧进行时域相减来分离前景和背景对象,减背景的方法利用当前帧减背景模型而分离运动对象。由于受光照变化、背景复杂度、噪声等多种因素的影响,上述两类方法具有各自的特点,针对不同应用场合,其提取效果会有所差异,执行时间也不尽相同<sup>[11]</sup>。

这里我们对每像素建立高斯分布模型,通过统计方法求得像素的概率密度函数(常满足高斯分布), $B_i(x, y) \sim N(\mu, \sigma)$ :

$$|r(x, y) - \mu_r(x, y)| < \tau \cdot \sigma_r(x, y) \quad (11)$$

$$|g(x, y) - \mu_g(x, y)| < \tau \cdot \sigma_g(x, y) \quad (12)$$

$$|b(x, y) - \mu_b(x, y)| < \tau \cdot \sigma_b(x, y) \quad (13)$$

以上三式同时满足,则认为该像素为背景像素。 $r, g, b$  分别表示坐标为  $(x, y)$  像素的三个颜色分量值,  $\mu_r, \mu_g, \mu_b$  为每个分量的均值,  $\sigma_r(x, y), \sigma_g(x, y), \sigma_b(x, y)$  为每个分量的方差,  $\tau$  为阈值,通常取值为 20 ~ 60。

## 3 实验结果及分析

在配置 Intel Celeron 2.4 GHz CPU, 256 MB DDR 内存的 PC 上,用 VC++6.0 平台对多个标准的视频序列进行测试。图 3~5 分别给出了对 Table\_tennis 序列、Hall\_monitor 序列和 Bream 序列(均为 qcif 格式)测试的部分结果。

首先,算法具有实时性。算法只用到了压缩流中的运动矢量来求得对象的大致运动区域,由于一个块(8 像素×8 像素)才一个运动矢量,基于块的操作使处理的数据量大大减少;在处理运动矢量场时,我们还做了以下简化处理,对属于同一视频场景或 GOP 中的帧,只处理前面几帧得到大致的运动区域,而对于后续帧使用与前面得到的相同区域进行解码,从而避免对每帧的运动矢量都进行重复的处理,这样可以大大节约处理时间。不过这样处理的前提条件是部分解码后实际的运动对象应包含在运动区域内。从后面的实验结果可以发现,有些序列部分解码的区域包含了除对象的较大背景部分,其目的就是多帧使用了相同的运动区域,以保证每帧中的运动对象包含在区域之内。对于 I 帧,由于没有运动矢量,我们按与 P 帧所求得的相同运动区域进行部分解码。对运动对象的部分解码时间取决于运动物体的大小,物体越大,进行 IDCT 的块数越多,消耗的时间越多。我们是以这部分的时间代价来提高了对象的精度,但一般仍能满足实时性,兼顾了实时与精度要求,表 1 列出了 Tennis 序列第 123 帧、Monitor 序列第 54 帧、Bream 序列第 124 帧提取过程中各主要步骤的处理时间。

算法还具有以下优势。1)能正确处理提取过程中的对

象进入场景、离开场景,同时多个对象等较复杂的情况,如 Hall\_monitor 序列,对人刚从房间出来进入走廊,从走廊进入房间及两人同时在走廊的情况都能进行正确提取。2)能准确提取运动区域,例如对 Table\_tennis 序列第 123 帧,我们准确提取了人的上半身,在短时间内,打乒乓球时脚是站立不动的,只有上半身在运动,这与自然状态一致。3)能适合不同应用场合,当只需作粗略的运动分析时,用基于块的区域增长算法基本上就能满足,见实验结果中第二列。当要得到更精确的对象轮廓形状信息时,我们以第二列中的结果作为预分割,对预提取的只有块精度的运动对象区域,用基于像素的方法来进一步提高所提取对象的精度。4)由于在压缩域处理时只用到了运动矢量信息,该算法同样适合于从 H. 264 压缩域中提取运动对象。

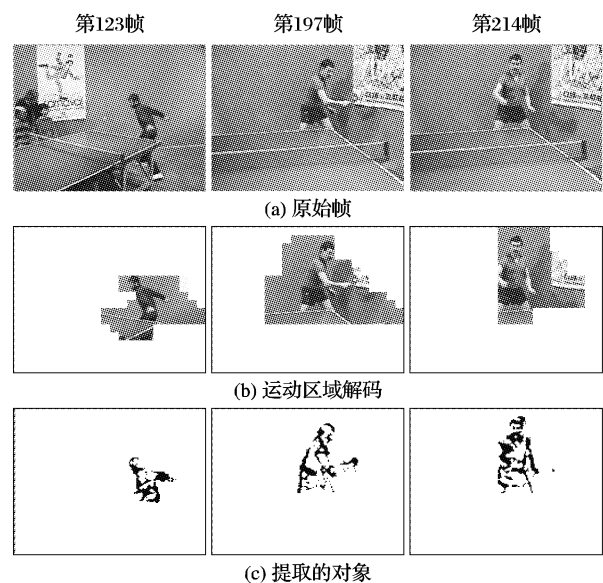


图3 Table\_Tennis 序列实验结果

表1 算法执行时的时间性能参考

帧	压缩域 处理时 间/ms	像素域处理时间/ms		
		IDCT 块数 (总块数: 396)	统计高斯 背景模型	减背景 (提取运动对象)
Tennis 序列 (第 123 帧)	0.182	61	0.856	0.521
Monitor 序列 (第 54 帧)	0.254	75	1.007	0.66
Bream 序列 (第 124 帧)	0.252	124	1.842	0.732

不足之处:1)算法对小于一个块大小的物体不能准确提取,如 Table\_tennis 序列第 197 帧中的乒乓球没有提取出来;2)在处理运动矢量场时,实现鲁棒的全局运动补偿算法是我们下一步工作,这样会使得算法既适用于静态背景序列,也能对动态背景序列进行全局运动补偿后提取出运动对象。

## 4 结语

本文提出了一种有效的 MPEG 压缩域运动对象提取算法,直接从 MPEG 压缩流提取原始运动矢量场后,对其进行处理得到可靠的运动矢量场,然后用基于块的区域增长算法求得运动一致性区域,再对运动区域解码还原成像素域,用高斯模型提取像素级精度的运动对象。这是在时间与精度间须知

的折中,以少许的时间代价来提高对象的精度。

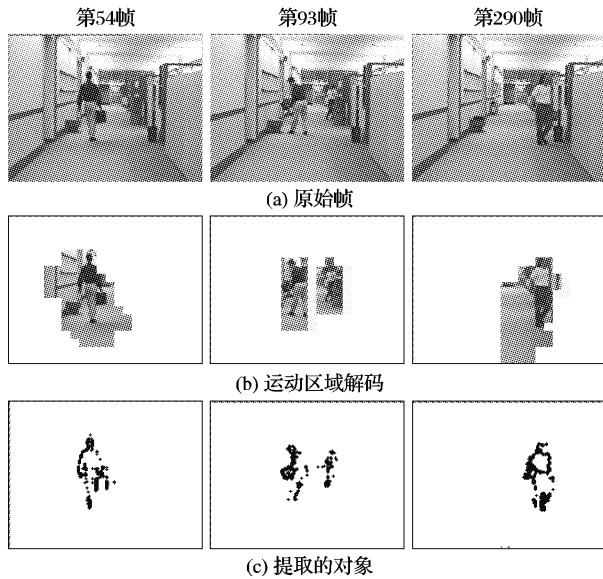


图4 Hall\_Monitor 序列实验结果

#### 参考文献:

- [1] 杨高波, 张兆扬. MPEG-4 框架下的视频对象分割及其关键技术分析[J]. 通信学报, 2007, 28(10): 109-122.
- [2] WEI ZENG, WEN GAO, DEBIN ZHAO. Automatic moving object extraction in MPEG video[C]//Proceedings of the 2003 International Symposium on Circuits and Systems. Bangkok, Thailand: IEEE, 2003, 2: 524-527.
- [3] WANG JIAN, LI ZE-NAN. Kernel-based multiple cue algorithm for object segmentation[C]//Proceedings of Image and Video Communications and Processing 2000, Proceedings of SPIE 4067. [S. l.]: SPIE, 2000: 462-473.
- [4] SIRIPONG T, UWE R, HEUER J, *et al.* Bayesian method for motion segmentation and tracking in compressed videos [C]// LNCS 3663. Berlin: Springer, 2005: 277-284.
- [5] ZHOU QIYA, YANG GAOBO, CHEN WEWIEI, *et al.* A fast and accurate moving object extraction scheme in the MPEG compressed domain[C]// Proceedings of the Fourth International Conference on Image and Graphics. Washington, DC: IEEE Computer Society, 2007: 592-597.
- [6] ISO/IEC JTC1/SC29/CD13818-2, MPEG-2, Coding of Moving Pictures and Associated Audio[S]. ISO, 1994.
- [7] PORIKLI F. Real-time video object segmentation for MPEG encoded video sequences[C]// Proceedings of SPIE Conference on Real-Time Image VIII. San Jose: SPIE, 2004, 5297: 195-203.
- [8] DONG LAN, ZOGHLAMI I, SCHWARTZ S C. Object tracking in compressed video with confidence measures[C]// Proceedings of IEEE International Conference on Multimedia and Expo. Toronto: IEEE, 2006: 753-756.
- [9] IBRAHIM M M, RAO S. Motion analysis in compressed video - An hybrid approach[C]// Proceedings of IEEE Workshop on Motion and Video Computing. Washington, DC: IEEE Computer Society, 2007: 17.
- [10] CHUNG R H Y, CHIN F Y L, WONG K-Y K. Efficient block-based motion segmentation method using motion vector consistency [C]// Proceedings of Machine Vision Application 2005. Japan: [s. n.], 2005: 550-553.
- [11] BENEZETH Y, EMILE B, ROSENBERGER C. Comparative study on foreground detection algorithm for human detection[C]// Proceedings of the Fourth International Conference on Image and Graphics. Washington, DC: IEEE Computer Society, 2007: 661-666.

(上接第 728 页)

## 4 结语

本文提出了一个基于轮廓方向变化率多尺度积的角点检测新算法。通过使用不同尺度的 B-样条函数对轮廓进行演化,再计算演化后轮廓方向变化率,把轮廓方向变化率累计相乘,从而得到轮廓方向变化率多尺度积,计算多尺度积的局部极值,最后针对此局部极值设置一个全局阈值来获得角点。定位性能分析和实验结果可以表明 MSM 算法具备如下优点: 1)能够检测出尽量多的真实角点,同时只有少量伪角点产生,即具备较高的检测率;2)具有好的定位性能,其定位误差不大于一个像素;3)对噪声的鲁棒性较强。

#### 参考文献:

- [1] LINDBERG T. Scale space theory: A basic tool for analyzing structures at different scales [J]. Journal Applied Statistics, 1994, 21(2): 223-261.
- [2] HARRIS C, STEPHENS M. A combined corner and edge detector [C]// Proceedings of the Fourth Alvey Vision Conference. Manchester: [s. n.], 1988: 147-151.
- [3] 陈白帆, 蔡自兴. 基于尺度空间理论的 Harris 角点检测[J]. 中南大学学报, 2005, 36(5): 751-754.
- [4] MOKHTARIAN F, SUOMELA R. Robust image corner detection through curvature scale space [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(12): 1376-1381.
- [5] RATTARANGSI A, CHIN R T. Scale-based detection of corners of planar curves [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992, 14(4): 430-449.
- [6] MOKHTARIAN F, SUOMELA R. Enhancing the curvature scale space corner detector[C]// Proceedings of Scandinavian Conference on Image Analysis. Bergen, Norway: [s. n.], 2001: 145-152.
- [7] HE X C, YUNG N H C. Curvature scale space corner detector with adaptive threshold and dynamic region of support [C]// IEEE Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04). Washington, DC: IEEE, 2004, 2: 791-794.
- [8] LEE J S, SUN Y N, CHEN C H. Multi-scale corner detection by using wavelet transform [J]. IEEE Transactions on Image Processing, 1995, 4(1): 100-104.
- [9] QUDDUS A, GABBOUJ M. Wavelet-based corner detection technique using optimal scale [J]. Pattern Recognition Letters, 2002, 23(1/3): 215-220.
- [10] WANG Y P, LEE S L. Scale-space derived from B-splines [J]. IEEE Transactions on Pattern Analysis Machine Intelligence, 1998, 20(10): 1040-1055.
- [11] SADLER B M, SWAMI A. On multiscale wavelet analysis for step estimation [C]// International Conference on Acoustics, Speech, and Signal Processing (ICASSP'98). Seattle: IEEE, 1998, 3: 1517-1520.
- [12] BAO P, ZHANG L. Canny edge detection enhancement by scale multiplication [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(9): 1485-1490.
- [13] YE H C - H. Wavelet - based corner detection using eigenvectors of covariance matrices [J]. Pattern Recognition Letters, 2003, 24(15): 2797-2806.