

## 基于能流信息场景检测的码率控制策略

张忠伟 刘贵忠 李宏亮 李永利

(西安交通大学电子与信息学院信息与通讯工程系 西安 710049)

**摘要:** 该文首先介绍了用能流信息进行运动估计的方法,然后把该运动信息运用到场景检测中,最后分析了场景切换下 MPEG-4 Q2 码率控制算法失效的原因,提出了基于场景切换补偿的码率控制策略,并在场景切换时对 MPEG-4 Q2 算法得到的目标比特数进行修正。实验结果表明,该方法可以在保证编码缓冲器不发生上溢的前提下改善切换点及后继帧图像编码的质量。

**关键词:** 运动估计, 场景切换检测, 码率控制

中图分类号: TN919.8

文献标识码: A

文章编号: 1009-5896(2005)12-1964-05

## Rate Control Scheme Based on Scene Change Detection Using Energy Flow Information

Zhang Zhong-wei Liu Gui-zhong Li Hong-liang Li Yong-li

(School of Electronic & Information, Xi'an Jiaotong University, Xi'an 710049, China)

**Abstract** Firstly, a method is introduced to estimate motion vectors between two pictures with energy flow information. Secondly, these motion vectors are applied to detect the scene change. Finally, from analyzing MPEG-4 Q2 rate-control algorithm with the situation of scene changes, an improved rate-control algorithm is proposed where the scene change is taken into account. In the improved algorithm, the target bit budget from MPEG-4 Q2 is modified when scene change occurs. Experimental results show that the proposed scheme can improve the picture quality at or after scene change without buffer overflow.

**Key words** Motion estimation, Scene change detection, Rate control

### 1 引言

MPEG-4 Q2 是一个可扩展的码率控制算法,根据不同的环境可为视频编码提供不同尺度、不同复杂度的码率控制方案<sup>[1]</sup>。MPEG-4 Q2 已被国际运动图像专家组(MPEG)采纳,作为最新一代视频编码标准MPEG-4<sup>[2]</sup>的标准码率控制算法之一,算法的核心是一个模拟信息率失真特性的二次函数模型。MPEG-4 Q2 码率控制分为4个步骤:初始化模型参数、估计目标比特数、计算量化参数和编码后更新模型参数。该算法在大多数情况下都能在保证图像质量的前提下有效控制恒定码率的视频流输出。但是,在场景切换处,MPEG-4 Q2 算法往往会失去有效性,造成在切换帧附近图像编码质量的

下降,所以有必要对其进行改进。

降低切换帧附近图像编码质量的下降,第一步必须先要检测场景切换位置,也就是首先要预测出场景切换的帧,然后采用相应的策略处理切换帧,主要策略分为两类:一类是编码方式转换,将切换点处的P帧强制转换为I帧,以保证切换点处的图像编码质量<sup>[3]</sup>;一类是根据率失真特性为切换帧分配相应的比特数,来减小因场景切换带来的编码质量下降<sup>[4]</sup>。

本文的安排如下,第2节介绍图像间运动矢量的能流估计方法,第3节是利用该运动矢量信息进行场景检测,第4节详细说明场景切换下的码率控制策略,第5节为实验验证及分析,第6节是本文的结束语。

2004-05-17 收到, 2005-05-08 改回

国家自然科学基金(60272072)和教育部高等学校博士学科点专项科研基金(2000069828)资助课题

## 2 运动矢量的能流估计方法

计算图像间运动矢量的传统方法是光流估计算法, 但光流只是三维物体的实际运动在图像平面上的投影, 所以它包含了许多虚假的运动信息, 使基于光流分析的视频分割方法造成镜头检测的漏检和误检。在我们的工作文献中<sup>[5]</sup>, 我们设计了利用能流信息估计运动矢量的技术, 该技术与传统的光流估计算法相比, 提高了运动矢量的估计精度, 大大减少了虚假的运动信息。

### 2.1 气体分子运动理论和麦克斯韦-玻尔兹曼分布定律

麦克斯韦是最早研究气体分子的速率分布规律的物理学家, 并于 1859 年从理论上推导出了理想气体分子在平衡状态下的速率分布函数<sup>[6]</sup>, 即麦克斯韦分布律:

$$f(v) = 4\pi \left( \frac{m}{2\pi kT} \right)^{3/2} v^2 e^{-mv^2/(2kT)} \quad (1)$$

式中  $m$  是气体分子的质量,  $T$  是平衡状态的温度,  $k$  是玻尔兹曼常数,  $v$  是气体分子的运动速率。由式(1)做出的理论曲线与实验曲线是极为接近的, 说明麦克斯韦分布律真实地描述了气体分子速率分布的客观规律。玻尔兹曼在麦克斯韦研究的基础上, 进一步研究了不同离散能量状态下微观粒子数的分布规律, 即麦克斯韦-玻尔兹曼分布律, 该定律的描述是: 在一个理想的气体平衡系统中, 如果微观粒子只能具有一系列不连续的能量, 其值由小到大排列为  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_i, \dots$ , 则在能量为  $\varepsilon_i$  的状态上分布的粒子数可表示为

$$N_i = C' e^{-\varepsilon_i/(kT)} \quad (2)$$

式中,  $C'$  是由系统状态所决定的玻尔兹曼常数; 在理想平衡状态下, 气体分子的能量  $\varepsilon_i$  主要是由其动能来决定的, 所以能量  $\varepsilon_i$  可写为:  $\varepsilon_i = mv_i^2/2$ 。由式(1)和式(2)可知: 平衡状态下, 气体分子的速率分布函数和分子数都是关于速率  $v$  的“钟”型函数。

麦克斯韦-玻尔兹曼分布律是气体分子动力学的一个基本定律, 它定量地描述了气体分子做随机运动时分子数的分布规律。该定律不单在气体分子动力学研究中起了巨大的推动作用, 而且在固体物理、激光等近代物理学中, 在研究微观粒子的随机运动规律时, 也得到了广泛的应用。

### 2.2 图像的温度场

视频信号也是随着场景的变化而随机变化的, 所以本文

把图像序列中的像素看作是随机运动的气体分子, 提出了一个“图像温度场”的概念, 并用麦克斯韦-玻尔兹曼分布律来描述图像序列中像素的变化规律。

假设视频序列中每帧图像内任何位置都具有一定的“温度”, 与气体温度的定义相一致, 并且这样的图像“温度”是由该位置处像素的“运动速度”和一个视频单元内具有该“运动速度”的像素总数所共同决定。把像素的亮度信息看作像素的“运动速度”  $v_i$ , 并统计一个 GOP 中具有该“运动速度”  $v_i$  的像素总数为  $N_i$ , 由式(2)可得

$$N_i = C' e^{mv_i^2/(2kT)} \quad (3)$$

对式(3)两端取自然对数, 整理后得

$$T = mv_i^2 / [2k(\ln C' - \ln N_i)] \quad (4)$$

式中  $m$  是像素的“质量”, 本文假设所有的像素都有相同的“质量”值 1;  $k$  和  $C'$  都是归一化常数。由式(4)可知, 图像任何位置处的温度  $T$  都是由该位置处像素的“运动速度”  $v_i$  与具有该“运动速度”  $v_i$  的像素总数  $N_i$  共同决定的。

### 2.3 能流余数方程和运动估计

根据推导光流方程的方法, 当定义了图像的温度之后, 假设沿着运动轨迹上图像的“温度场”保持不变, 就可以推导出一个全新的描述图像运动信息的偏微分方程, 即

$$\frac{\partial T(x, y, t)}{\partial x} u + \frac{\partial T(x, y, t)}{\partial y} v + \frac{\partial T(x, y, t)}{\partial t} = 0 \quad (5)$$

式中  $\partial T(x, y, t)/\partial x$  和  $\partial T(x, y, t)/\partial y$  是图像温度场分别在空间的水平和垂直方向上的梯度,  $\partial T(x, y, t)/\partial t$  是图像温度场在时间上的梯度,  $u = dx/dt$  和  $v = dy/dt$  表示水平和垂直方向上图像温度场的运动速度。

根据气体分子运动理论, 忽略分子之间的势能, 气体的温度是气体分子平均动能的度量。所以, 视频中像素的温度也是像素能量的度量, 沿着运动轨迹上温度场不变, 也就是能量保持不变, 所以该假设就称为能流约束条件, 式(5)即为本文的能流方程。根据式(5)我们可以利用回归的方法<sup>[7]</sup>估计出视频序列中连续两帧之间的运动矢量, 即能流信息。

图 1(a)和 1(b)给出了 News 测试序列的第 488 帧和 489 帧图像; 图 1(c)是两幅图像之间差分的绝对值图像, 表示事实上的运动信息。图 1(d), 1(e)分别是光流分析方法和能流分析方法估算出的运动场。

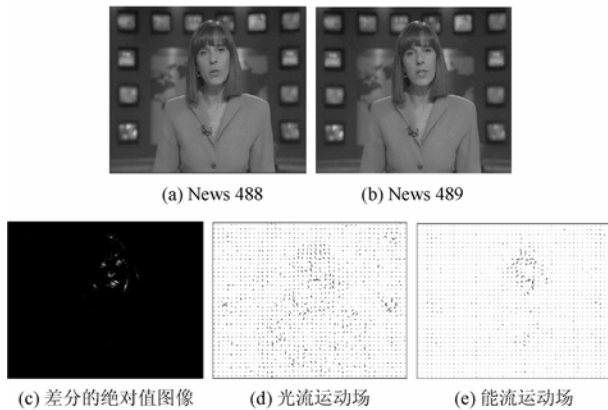


图1 光流方程与能流方程估算出的运动场之间的对比

从图1的比较可以看出,能流方程估算的运动场(能流信息)更加精确地捕捉到实际运动信息,而光流方程估算的运动场产生了大量的伪运动信息。

### 3 用能流信息进行镜头检测

由第2节的分析可知,图像温度场估算出的能流场比光流场更加接近图像实际的运动场,所以把能流场应用于镜头检测应该能够得到比用光流信息进行镜头检测更好的结果。把能流的平滑度与图像温度场空间梯度的比值  $D$  作为镜头检测的判据。

我们用 MPEG-7 提供的 Eyexam 作为测试序列。为了实验的方便我们对 Eyexam 做了进一步的编辑,编辑后的 700 帧图像中包含 7 个镜头的突变(cut)和 13 个镜头的渐变,其中镜头渐变包括 6 个镜头的溶解(dissolve)、4 个镜头的淡入(fade in)和 3 个镜头的淡出(fade out)。综合考虑检测结果的“检中率”和检验的“正确率”,  $D$  的阈值分别选为 7.5% 和 10%。如图 2 给出了用光流信息和能流信息进行镜头检测的结果,图中虚线和点线分别为  $D$  的阈值分别取 10% 和 7.5% 时的阈值线。

由图 2 可以看出,光流信息在镜头边界处也发生了扩散现象,造成无法确定镜头发生变化的准确位置。利用能流信息不但检测出了所有的镜头变化,而且可以精确定位镜头的变化位置。

### 4 场景切换下的码率控制策略

在场景切换点, MPEG-4 Q2 码率控制算法失去有效性,是因为切换点的图像的内容发生了很大的变化,如果再利用前面图像编码信息来预测该切换点的图像的编码性质,导致出现较大的预测误差,大大降低了重建视频的质量。

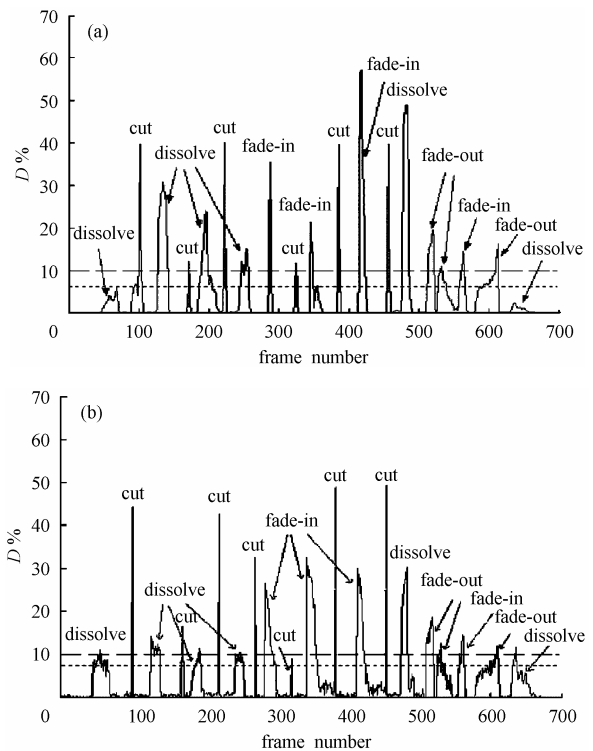


图2 利用光流信息 (a) 和能流信息 (b) 进行镜头检测的结果

为了消除场景切换对重建视频质量的影响, Farin 等人提出了变 GOP 结构的时间预测编码方法<sup>[8]</sup>。将场景中的第一个 P 帧作为 I 帧处理,也即以此为新 GOP 的开始。这里没有讨论 B 帧,因为 B 帧不作为任何帧的参考,图像质量下降可以不作考虑;另外 B 帧使用双向预测,至少和两个参考帧中的一帧具有相同的场景。图 3 给出了动态调整 GOP 结构的示意图。

频繁地调整 GOP 的结构,会使更多的 P 帧重新调整为 I 帧编码,尤其是在运动变化剧烈而场景没有发生变化的场景中,这样会使后继帧的目标比特预算大大低于目标码率,导致连续图像的质量发生剧烈波动,大大降低解码器重建视频的视觉效果。为此,我们采用视频编码前预处理的方法,首先用第 3 节引入的场景检测算法,预测出场景变化的帧,并把该位置存储于文件中,在视频编码过程中,如果遇到场景切换帧,编码器就把该帧作为 I 帧处理。预先进行场景检测适合于已存储的视频序列,如果在实时视频通信中,可以把该场景预测算法整合到视频编码器中,在编码前估计当前编码帧与前一帧的能流信息,判断当前帧是否为场景切换帧,如果是场景切换帧就把当前帧作为一个新 GOP 的第一帧,即 I 帧。

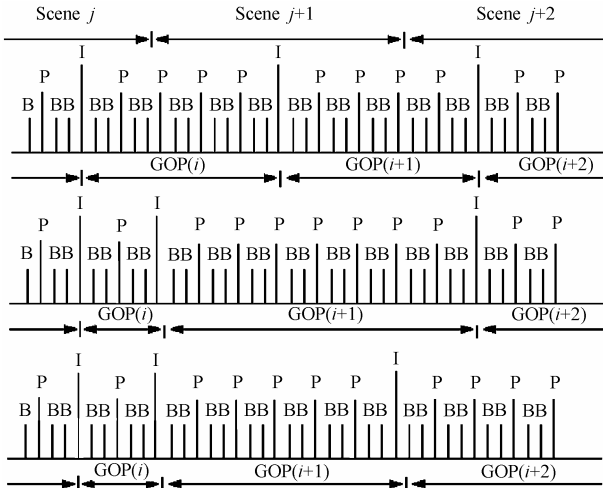


图 3 场景切换时 GOP 重组示意图

MPEG-4 Q2 码率控制算法的有效性很大程度上依赖于图像编码在时域上的相关性,

并不能随当前图像内容及预测效果而变化,因此不必要对其进行改进。如前所述,我们先用能流信息检测出场景切换帧,然后把它作为 I 帧进行处理,但该 I 帧的目标比特数如果不能精确预测,那么 MPEG-4 Q2 依然无法有效地处理场景切换问题。为此,在检测到场景切换后,我们用下面一种相对较简单的方法来估计该切换 I 帧的目标比特数。

根据编码理论, I 帧的编码比特数与信源方差之间的关系可用一个对数形式的函数来表示,由于 MPEG 编码以宏块为单位,所以定义宏块的编码比例系数:

$$r_{mb} = \ln(\sigma^2/16) \tag{6}$$

其中  $\sigma^2$  为宏块灰度值的方差; 16 为与图像信源分布参数有关的经验值。定义图像的码率比例系数为该帧图像中所有宏块的编码比例系数之和,则可按下式对由 MPEG-4 Q2 算法得到的目标比特数  $T_0$  进行修正:

$$T = T_0 \times \frac{r_{cur}}{r_{pre}} \tag{7}$$

其中  $r_{cur}$  和  $r_{pre}$  分别为当前 I 帧和前一个 I 帧的编码比例系数。

### 5 实验验证及分析

我们用能流信息检测场景切换帧,并用第 4 节中的算法修正切换帧的目标比特数,设计了一个改进的 MPEG-4 Q2 码率控制方案,通过与 MPEG-4 Q2 的码率效果相比较来说明改进 MPEG-4 Q2 方案的有效性。实验条件为:在 MPEG-4 编码器 MomuSys 上编码标准测试序列 Test1 和 Test2,这两

个测试序列都是由几个不同的标准视频编码序列的一部分片段组合而成,即我们把 Test1 和 Test2 都编辑成由 7 个场景促成的测试序列,长度都是 100 帧;编码的帧率为 10 帧/s,都采用实时视频通信中推荐的 IPPP... 时间预测结构;端端的最大时间延迟预置为 250ms。

当采用 MPEG-4 Q2 算法时, GOP 的固定大小为 10,即在一个 GOP 中第一个为 I 帧编码,其余 9 帧进行 P 帧编码。当采用本章改进的 MPEG-4 Q2 码率控制算法时,先用第 3 节的能流信息的场景切换检测算法检测出 Test1 和 Test2 中场景切换帧的位置,并预先存储到一个文件中,在编码时编码器先根据上述文件检测当前的待编码帧是不是场景切换帧,如果是场景切换帧就把它作为 I 帧编码,即作为一个新 GOP 的第一帧,再用第 4 节中的算法对该帧的目标比特数进一步修正,显然这是一种变 GOP 的编码方案。图 4 给出了采用本文改进码率控制算法和原始 MPEG-4 Q2 码率控制算法时,两个测试序列编码图像质量的变化曲线,其中编码质量的测度采用了图像灰度的峰值信噪比(PSNR,单位为分贝)。

从图4可以看出,用本文改进后的MPEG-4 Q2码率控制方案,不但大大提高了视频编码后图像质量变化的平稳性,而且使Test1和Test2的平均峰值信噪比提高了0.13dB和0.20dB。播放重建的视频也显示出,采用本文改进的码率控制方案取得了更好的视觉效果。

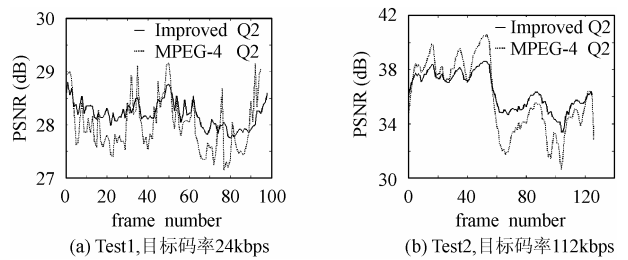


图 4 采用不同码率控制算法编码不同测试序列时,编码图像峰值信噪比(PSNR)的变化

### 6 结束语

本文利用能流分析方法准确地捕捉到了视频中连续两帧图像之间的运动信息;能流信息精确地检测出了场景切换的位置;用图像的编码比例系数修正场景切换帧的目标比特数。利用上述算法改进 MPEG-4 Q2 码率控制方案,改进后的方案大大改善了码率控制的效果,不但生成了更加平稳性的视频编码图像质量;而且还提高了编码视频的平均峰值信噪比。

## 参 考 文 献

- [1] Lee H J, Chiang T, Zhang Y Q. Scalable rate control for MPEG-4 video [J]. *IEEE Trans. on Circuits Syst. Video Technol.*, 2000, 10(9): 878 – 894.
- [2] MPEG-4 Video Verification Model v18.0, Coding of Moving Picture and Associated Audio N3908, ISO.IEC JTC1/ SC29/WG11, Jan, 2001.
- [3] Grecos C, Jiang J. On-line improvement of the rate-distortion performances in MPEG-2 rate control [J]. *IEEE Trans.on Circuits Syst. Video Technol.*, 2003, 13(6): 519 – 528.
- [4] Sun H, Kwok W. MPEG coding performance improvement by joint optimizing coding mode decisions and rate control [J]. *IEEE Trans.on Circuits Syst. Video Technol.*, 1997, 7(3): 449 – 458.
- [5] Zhang Z W, Liu G Z. Motion estimation based on energy flow equation [J]. *Processing in Natural Science*, 2003, 13(3): 228 – 232.
- [6] Jones E R. *Contemporary College Physics* [B], 3<sup>rd</sup> ed., Boston, McGraw Hill: 2001: 203 – 132.
- [7] Horn B K, Schunck B G. Determining optical flow [J]. *Artificial Intelligence*, 1981, 17(2): 185–204.
- [8] Farin D, Mache N, P H N De With. A software-based high-quality MPEG-2 encoder employing scene change detection and adaptive quantization [J]. *IEEE Trans. on Consumer Electronics*, 2002, 48(4): 2174 – 2193.
- 张忠伟: 男, 1975 年生, 博士生, 研究方向为多媒体通信、码率控制、图象处理、数据压缩、非平稳信号处理、小波分析.
- 刘贵忠: 男, 1962 年生, 教授, 博士生导师, 研究方向为通信、多媒体、非平稳信号处理.
- 李宏亮: 男, 1970 年生, 博士生, 研究方向为多媒体通信、小波分析、非平稳信号处理.
- 李永利: 男, 1975 年生, 博士, 研究方向为多媒体技术、实时传输控制、QoS、无线局域网.