

# 低码率视频编码的优化方法\*

贺玉文, 王琪, 袁昱, 钟玉琢, 杨士强

(清华大学 计算机科学与技术系, 北京 100084)

E-mail: heyw@media.cs.tsinghua.edu.cn

http://imlab.cs.tsinghua.edu.cn

**摘要:** 采用的低码率视频编码方法是以全局运动补偿和局部运动补偿为核心的混合编码方法. 主要针对 MPEG-4 中原有算法计算量大、复杂度高的特点进行了编码系统优化, 将这一编码方法实用化. 该优化方法是针对占编码系统 70% 计算量的全局运动估计进行的, 采用了基于运动特征的运动估计方法和鲁棒的目标函数, 并在优化方法中引入了三级金字塔的分层计算. 从对比实验结果来看, 优化方法对不同运动类别视频的编码都是有效的, 系统编码速度提高了 3 倍以上. 目前, 这项优化方法已经被 MPEG-4 视频编码优化小组采用.

**关键词:** 低码率编码; 视频编码; 全局运动; 运动补偿; 运动估计

中图法分类号: TP391 文献标识码: A

本文对低码率下视频压缩编码方法进行了研究, 所采用的低码率视频编码方法是以全局运动补偿和局部运动补偿为核心的混合编码方法. 本文主要是针对 MPEG-4 中原有算法计算量大、复杂度高的特点进行了编码系统优化, 将这一编码方法实用化. 文中优化方法是针对占编码系统 70% 计算量的全局运动估计进行的, 采用了基于运动特征的运动估计方法和鲁棒的目标函数, 并在优化方法中引入了三级金字塔的分层计算. 从对比实验结果来看, 优化方法对不同运动类别视频的编码都是有效的, 系统编码速度提高了 3 倍以上. 目前, 这项优化方法已经被 MPEG-4 视频编码优化小组采用.

随着计算机和通信技术的迅猛发展, 图像压缩编码技术在数字媒体存储、通信等领域发挥了及其重要的作用. 为适应多媒体存储型应用发展的需要, ISO(国际标准化组织)于 1988 年前后成立了运动图像专家组 MPEG(Moving Picture Experts Group), 开展运动图像压缩编码标准制定工作, 并在 1993 年和 1994 年分别制定了 MPEG-1(IS-11172)和 MPEG-2(IS-13818), 这两个编码标准都获得了成功, 并在存储型应用领域中被广泛使用, 其中 MPEG-1 标准被 VCD 行业所采用, 而 MPEG-2 标准则被用于高清晰度数字电视(HDTV)和电影节目存储. 另一方面, ITU-T(国际通信联盟)的高级编码专家工作组也一直致力于运动图像编码技术研究, 并制定相关编码标准, 如 H.120(1984), H.261(1991), H.263(1996)等, 这些国际标准在通信领域得到了广泛使用.

目前由于网络发展很快, 多媒体应用逐渐由存储应用型转为网络传输型(例如视频点播), 实现较高质量和稳定的视频传输是这些网络多媒体应用的一个关键技术. 通过宽带网络点播多媒体节目已经成为现实, 而在无线网络和低带宽网络环境下进行视频传输是当前网络多媒体应用的一个技术难点, 如果能够提高在低带宽下的视频压缩性能, 那么这将使未来通过 Internet 网络点播视频成为可能, 这也将把多媒体应用带入到世界每个角落. 当前如果要在低带宽网络(传输率为 10kbps~100kbps)条件下传输高质量压缩的视频数据, 还存在许多技术困难. 目前 ISO/IEC JTC1 SC29 WG11 MPEG 运动图像专家组正制定多媒体编码标准 MPEG-4(IS-14496)<sup>[1,2]</sup>, 以

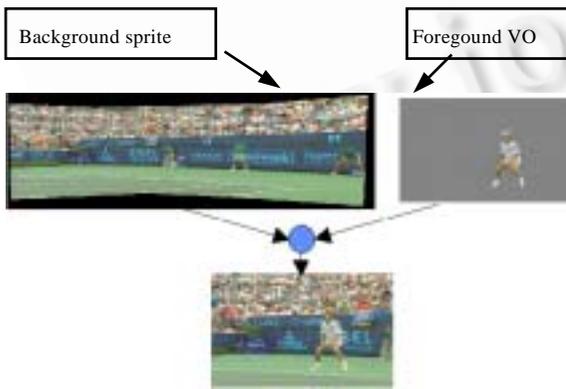
\* 收稿日期: 2000-12-18; 修改日期: 2001-03-21

基金项目: 国家 863 高科技发展计划资助项目(863-306-ZT03-09-1)

作者简介: 贺玉文(1974-), 男, 江苏丹阳人, 博士生, 主要研究领域为低码率视频压缩和传输; 王琪(1973-), 男, 北京人, 博士, 主要研究领域为数据压缩; 袁昱(1975-), 男, 四川重庆人, 博士生, 主要研究领域为数据压缩; 钟玉琢(1938-), 男, 辽宁沈阳人, 教授, 博士生导师, 主要研究领域为多媒体计算机技术; 杨士强(1952-), 男, 山东高唐人, 教授, 主要研究领域为多媒体计算机技术与系统, 多媒体信息处理.

促进下一代多媒体通信应用,其中 MPEG-4 标准中引入了一些低码率下新的编码技术,如全局运动补偿编码和 2D 三角网格等,同时 ITU-T 的高级视频编码专家组也正在制定下一代视频编码标准 H.26L<sup>[3,4]</sup>,这一标准的目标是要进一步提高编码效率,尤其是提高低码率下的编码效率(提高压缩比).为了攻克低码率下的技术难点,MPEG 运动图像专家工作组和 ITU-T 高级编码专家组进行了积极的合作,MPEG-4 低码率下的一些编码技术参考了 H.263 中的技术,而 H.263++中的一些低码率下的容错技术则借鉴了 MPEG-4 中的技术(如可变长度 slice 结构).MPEG-4 和 H.26L 中低码率编码技术都是针对无线网络特性的,如带宽低、误码率高等,其中许多编码技术都将是下一代通信标准的主流技术<sup>[5]</sup>.

MPEG-4 中的 sprite 编码方法是专门针对视频中背景编码特点而提出的一种低码率编码方法(图 1 是 sprite 编码示意图).它的基本思想是,首先通过全局运动估计得到全局运动参数,然后根据运动模型参数将这一视频段中的背景图像通过拼接构成有关背景的全景图,这个背景图像就称为 sprite 图,视频编码时对这个 sprite 图进行编码,然后对前景进行单独编码,而在解码时首先对背景 sprite 图进行解码,然后再对每帧的前景进行解码,并根据每帧运动参数得到当前帧背景在 sprite 图像中的位置,再通过叠加方法将背景和前景合成出完整的视频帧.与传统的运动补偿编码方法<sup>[6]</sup>(MPEG-1/MPEG-2/H.263)相比,这种编码方法的好处是,最大限度地去除了背景的冗余度.目前由于 sprite 图像一般都有很大的分辨率,因此这种编码方法有较大的延迟性和复杂度(生成 sprite 图和对 sprite 图像编码),因此在实际使用时较多采用由它变形简化得到的全局运动补偿编码方法,在编码时不必对一个视频段生成 sprite 图像,只是利用前一帧作为已有的“sprite”图像,对当前帧进行全局运动补偿和局部运动补偿两种编码方法相结合的混合编码,其中全局运动补偿是指,使用复杂模型(仿射或者透视模型)对图像进行整体的运动补偿;局部运动补偿是指,对局部小区域(宏块)利用平移模型进行运动补偿.这是因为与局部运动补偿编码相比,全局运动补偿更适合背景编码(背景运动具有一致性),图 2 显示了全局运动补偿编码和局部运动补偿编码在对背景编码时的性能比较,从图中可以看出,在相同码率下全局运动补偿编码要比局部运动补偿编码在图像质量上提高 1.5~3dB 左右,这说明全局运动补偿编码更加适合对背景的编码.但是全局运动补偿编码的计算量和复杂度太大,Smolic<sup>[7]</sup>,Kim<sup>[8]</sup>和 Konrad<sup>[9]</sup>等人作了一些有意义的改进,其中 Konrad 提出了三级金字塔计算和基于运动补偿残差直方图去除噪声的方法,提高了 MPEG-4 中已有全局运动补偿编码的效率,但其编码速度还比较慢.由于 MPEG-4 的复杂度妨碍了这一标准的推广和使用,因此 MPEG 组织为了推广 MPEG-4 的使用,专门成立了一个 MPEG-4 优化小组对其进行全面优化<sup>[10,11]</sup>,本文对 MPEG-4 中全局运动补偿这种低码率编码方法进行了研究<sup>[12,13]</sup>,提出了比较有效的优化方法,提高了系统性能(编码速度提高了 3 倍以上),这一基于特征的全局运动估计优化方法目前已被 MPEG-4 所采用<sup>[10]</sup>.



背景 sprite 图, 一个前景 VO.

Fig.1 Sprite coding

图 1 Sprite 编码

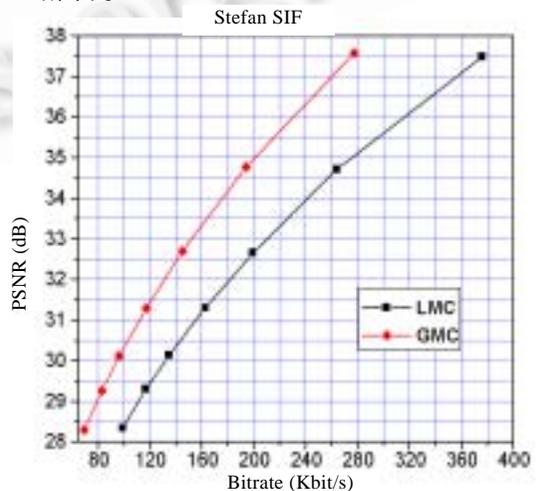


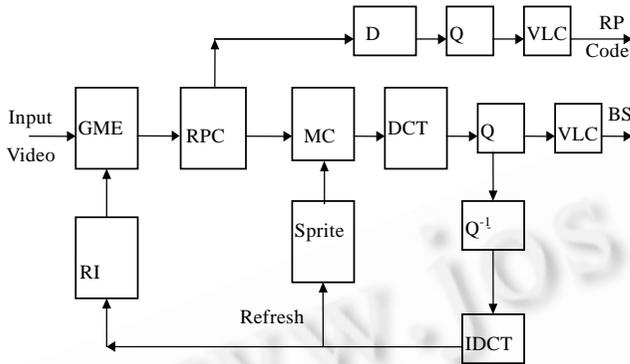
Fig.2 Comparison between GMC and LMC for background object's coding

图 2 全局运动补偿(GMC)和局部运动补偿(LMC)编码对背景编码时的性能比较

本文第 1 节详细讲述低码率视频编码方法和本文提出的优化方法以验证优化方法的有效性和优越性能. 第 2 节将讲述对比实验结果,包括静态 sprite 编码和全局运动补偿编码的比较结果,这些实验对比条件是 MPEG-4 视频组统一规定,并通过了第三方的验证<sup>[14,15]</sup>.第 3 节将对本文所提出的优化方法进行一个小结,并讨论以后研究的一些问题.

### 1 低码率视频编码方法及其优化技术

从目前研究来看,实用的低码率视频编码方法是全局运动补偿编码和局部运动补偿编码相结合的混合编



视频输入, 全局运动估计, 重构图像, 参考点计算, 差分, 运动补偿, 更新, 参考点码流, 码流.

Fig.3 Flowchart of Sprite coding

图 3 Sprite 编码的流程图

码方法,因为全局运动补偿编码适合背景区域编码,而局部运动补偿适合局部复杂运动区域编码,两者的有机结合可以提高系统整体的编码性能,这种编码方法属于简化的 sprite 编码.图 3 是 sprite 编码的流程图,基本步骤是对输入视频图像帧首先作全局运动估计,得到全局运动模型参数,然后计算出参考点位置,对参考点的运动矢量进行差分编码,得到参考点编码码流;而在得到全局运动参数之后,根据当前已生成的 sprite 图像进行全局运动补偿,对补偿后的残差进行 DCT 变换,然后再进行量化和可变长度编码,得到残差编码码流;在编码器端还要对已编码的当前

帧进行解码(逆量化,逆 DCT 变换),并将解码后的重构图像保存到缓冲区中,然后根据已得到的全局运动估计参数将当前解码图像进行变换,并添加到当前 sprite 图像上,从而更新 sprite 图像,以用于下一帧的全局运动补偿.

在 sprite 编码和全局运动补偿编码中的核心技术就是全局运动估计,它直接影响到全局运动补偿的效率,且其复杂度较高,需要较大的计算量,统计表明,全局运动估计时间占整个全局运动补偿编码时间的 70%,本文针对这一点对 sprite 编码中的全局运动估计进行了优化研究.第 1.1 节将讲述全局运动估计原理,第 1.2 节讲述全局运动估计的优化方法.

#### 1.1 全局运动估计原理

用  $[x_t, y_t]^T$  表示某像素点在当前图像中的位置,用  $[x_{t-1}, y_{t-1}]^T$  表示该像素点在前一参考图像中的对应位置.它们的对应关系用六参数仿射参数模型表示:

$$\begin{cases} x_{t-1} = ax_t + by_t + c \\ y_{t-1} = dx_t + ey_t + f \end{cases} \quad (1)$$

用  $I(x,y)$ 和  $I'(x',y')$ 分别表示当前图像和前一帧图像.用  $\theta = (a, b, c, d, e, f)^T$  表示参数矢量,则目标能量函数为:  $R(\theta) = \sum_{x_t, y_t} w[I'(x'_{t-1}, y'_{t-1}) - I(x_t, y_t)]^2$ ,全局运动估计就是求解待定参数  $\theta$ ,使得  $R(\theta)$  取得最小值,从而得到参数的最佳估计.假设已有第  $k$  步的计算参数  $\theta_k$ ,通过目标函数  $R(\theta)$  在该点作泰勒展开以近似  $R(\theta)$ :

$$R(\theta) \approx R(\theta_k) + g_k^T (\Delta\theta_k) + \frac{1}{2} (\Delta\theta_k)^T H_k (\Delta\theta_k), \quad (2)$$

其中  $g_k$  和  $H_k$  分别是  $R(\theta)$  在  $\theta_k$  处的梯度向量和 Hessian 矩阵:

$$g_k = J_k^T W \gamma_k, \quad H_k = J_k^T W J_k + \sum_i \gamma_i w_i H_{ik},$$

$\gamma_k = [\gamma_1 \gamma_2 \dots \gamma_N]^T$  表示  $\theta_k$  处的残差,

$$J_k = \partial \gamma_k / \partial \theta = \begin{bmatrix} \partial r_i / \partial \theta_1 & \dots & \partial r_i / \partial \theta_6 \\ \vdots & \vdots & \vdots \\ \partial r_N / \partial \theta_1 & \dots & \partial r_N / \partial \theta_6 \end{bmatrix},$$

$$\partial r_i / \partial \theta_1 = \partial r_i / \partial a = (\partial I' / \partial x_{t-1})|_{(x'_{t-1}, y'_{t-1})} \cdot x_t, \quad \partial r_i / \partial \theta_2 = \partial r_i / \partial b = (\partial I' / \partial x_{t-1})|_{(x'_{t-1}, y'_{t-1})} \cdot y_t,$$

$$\partial r_i / \partial \theta_3 = \partial r_i / \partial c = (\partial I' / \partial x_{t-1})|_{(x'_{t-1}, y'_{t-1})}, \quad \partial r_i / \partial \theta_4 = \partial r_i / \partial d = (\partial I' / \partial y_{t-1})|_{(x'_{t-1}, y'_{t-1})} \cdot x_t,$$

$$\partial r_i / \partial \theta_5 = \partial r_i / \partial e = (\partial I' / \partial y_{t-1})|_{(x'_{t-1}, y'_{t-1})} \cdot y_t, \quad \partial r_i / \partial \theta_6 = \partial r_i / \partial f = (\partial I' / \partial y_{t-1})|_{(x'_{t-1}, y'_{t-1})},$$

$W$  是对角方阵且  $W_{ii} = w_i$ ,  $H_{ik}$  是  $\gamma_i$  的 Hessian 矩阵,如果残差  $\gamma$  比较小,那么可以作如下近似:  $H_k \approx J_k^T W J_k$ , 由方程  $\partial R(\theta) / \partial \theta = 0$  和等式(2)可以得到:

$$J_k^T W J_k (\Delta \theta_k) = -J_k^T W \gamma_k. \tag{3}$$

根据方程式(3)就可以得到  $\theta_k$  处的增量,这样就可以求出下一步参数:  $\theta_{k+1} = \theta_k + \Delta \theta_k$ . 这样迭代,最后可以逐步减小目标函数,达到优化目的.详细算法可以参考文献[2,16,17].

### 1.2 全局运动估计优化方法

在 MPEG-4<sup>[2]</sup> sprite 编码中采用 Konrad<sup>[9]</sup>提出的方法,该方法是基于三级金字塔多分辨率运动估计,在每层中又采用了基于残差直方图的方法去除残差大的像素点,然后用基于梯度下降的方法进行优化参数计算,得到每一层的运动模型参数.这种方法的优点是分辨率计算可以加快估计速度,同时还具有一定的鲁棒性.但是由于在每层中都采用了全部像素进行计算,计算量非常大,而且准确性也难以得到保证,这样的复杂度使得 sprite 编码系统性能受到很大限制,本文针对这一点进行了算法上的优化,在计算过程中将利用运动特征点代替全体像素计算,并在金字塔的每层计算中都采用收敛性较好的 Levenberg-Marquadet 方法进行参数估计计算,而不是基于梯度的方法.从时间和空间变化特性出发,本文将采用公式(4)所定义的像素集合作为运动特征像素集合,进行有效的运动估计计算.因为空间边缘处像素亮度值对运动比较敏感,能够比较准确地反映运动变化情况.

$$\{ (x, y) \mid (DThreshold_{gradient} < (|I_x| + |I_y|) < UThreshold_{gradient}) \text{ AND } (|I_t| > Threshold_t) \}. \tag{4}$$

其中  $I_x, I_y$  分别表示像素亮度空间梯度的 X, Y 方向分量,  $I_t$  表示像素亮度的时间变化率.在优化算法中采用的阈值是经过实验比较得出的(主要考虑能够取得较好的系统编码速度和 PSNR),如下所示:

$$DThreshold_{gradient} = 1.25 \text{Mean}(|I_x| + |I_y|),$$

$$UThreshold_{gradient} = 2.0 \text{Mean}(|I_x| + |I_y|),$$

$$Threshold_t = 1.0 \text{Mean}(|I_t|).$$

优化后的全局运动估计方法由以下 5 步组成:首先利用高斯滤波器生成图像帧的三级金字塔低分辨率图像,然后进行以下 5 步分层计算:

第 1 步.计算金字塔最上层的运动参数:采用三步搜索方法估计最初平移参数(式(1)中的  $c, f$  分量),这对大平移运动非常有效,可以增加以后计算的收敛性;用残差直方图排除残差大的 10% 像素点,然后利用剩余像素点用 Gauss-Newton 方法估计该层全局运动参数.

第 2 步.将最上层估计出的参数映射到中间层(由于分辨率不同,所以要修改平移分量  $c, f$ ).

第 3 步.计算中间层全局运动参数:根据从上一层得到的初始参数值进行运动补偿,去除残差直方图中大的 10% 像素;根据式(4)选取运动特征点;在运动特征点基础上,用 Levenberg-Marquadet 方法估计中间层运动参数.

第 4 步.将中间层估计出的参数映射到最底层.

第 5 步.计算最底层的运动参数:计算  $16 \times 16$  大小块的残差和,然后去除残差和较大的 30% 的块区域;根据式(4)从剩余像素中选择运动特征点;在运动特征点基础上,用 Levenberg-Marquadet 方法估计金字塔最底层的运动参数.

## 2 实验比较

为了验证本文提出的优化方法性能,我们根据 MPEG-4 视频优化小组制定的实验条件<sup>[11]</sup>进行了与 MPEG-4 中原有编码方法的对比实验,测试了本文优化方法在 MPEG-4 静态 sprite 编码和全局运动补偿编码中的效果.表 1 列出了实验中所用的视频序列,它们都是长度为 300 帧的 MPEG-4 运动编码的标准测试序列.这些序列中包含了不同的运动情况,运动规律相对一般视频都要复杂些,基本上包括了平移、旋转、放缩等,而且序列中既有背景运动(摄像机运动引起),有前景运动,而前景运动包括多个运动物体运动.另外,这些运动基本都是变速运动.对 sprite 编码而言,这些序列具有很好的代表性,通过这些序列的实验比较可以看出本文提出的优化方法的性能.

**Table 1** Testing sequences

表 1 测试序列

Sequence (300frames)	Format	Motion
1.CoastGuard (Rect)	QCIF (30fps)	Large motion for background and foreground
2.Foreman (Rect)	QCIF (30fps)	Small motion for background, large motion for foreground
3.Stefan (Rect)	QCIF (30fps)	Large motion for background and foreground
4.Bus (Rect)	SIF (30fps)	Large motion for background and foreground
5.TableTennis (Rect)	SIF (30fps)	Small motion for background, large local motion for foreground
6.Stefan (Rect)	SIF (30fps)	Large motion for background and foreground
7.Stefan (VO0)	SIF (30fps)	Large motion for background
8.Basketball (Rect)	CIF (30fps)	Variable motion for background, large motion for foreground
9.CoastGuard (Rect.)	CIF (30fps)	Large motion for background and foreground
10.Foreman (Rect)	CIF (30fps)	Small motion for background, large motion for foreground
11.FlowerGarden (Rect)	SIF (30fps)	Large motion for background

背景前景都有大运动, 背景小运动,前景大运动, 背景小运动,前景有局部大运动,  
背景大运动, 背景有变速运动,前景大运动.

将本文提出的优化的全局运动估计方法和 MPEG-4 校验模型(VM)已有方法在静态 sprite 编码中进行了比较,比较结果见表 2.

**Table 2** Comparison results of static sprite coding experiment

表 2 静态 sprite 编码实验结果比较

Seq.	VM Sprite GME Rate [kb/s] SNRY [dB]	Proposed sprite new GME Rate [kb/s] SNRY [dB]	VM sprite GME	Proposed sprite new GME
			VM GME time: Proposed GME time	
Coast G. (Rect.)	31.78	44.54		
	12.44	12.45	7.92:1	
Coast G. (VO3)	23.54	24.35		
	13.05	13.08	5.75:1	
Stefan (Rect.)	340.97	286.47		
	15.57	15.20	7.93:1	
Stefan (VO0)	94.00	99.72		
	20.41	29.42	7.80:1	

从表中可以看出,使用本文提出的优化全局运动算法的静态 sprite 编码效果比 MPEG-4 原有算法要有所改进,如 stefan(VO0)的峰值信噪比都有明显提高,而全局运动估计速度提高了 7 倍多.

表 3 是全局运动补偿编码的对比实验结果,所用的参考程序是 MPEG-4 FPDAM1 Microsoft<sup>[18]</sup>平台,其中 GME Time-Ratio 是指对 MPEG-4 中全局运动估计优化前和优化后运动时间之比,GMC Time-Ratio 是系统优化前后编码时间之比.从表中可以看到,在各种不同码率下(24kbps~1024kbps),全局运动补偿编码系统中全局运动估计速度提高了 7.5 倍左右(6.3~10.2),系统编码速度都有不同程度的提高(2.5~5.3),平均提高了 3.5 倍左右.而系统编码后峰值信噪比只下降了 0.06dB,这在解码的图像质量上很难分辨,在 2000 年 10 月第 54 次 MPEG 工作会议上进行了主观测试,整体评价是优化前后图像质量没有差别.从这个实验可以看出,在具有不同运动的视频序列中,本文提出的优化方法对低码率下全局运动补偿编码方法的性能提高都是极其有效的.

**Table 3** Comparison results of global motion compensation coding experiment

**表 3** 全局运动补偿(GMC)对比实验结果

Bitrate (kbps)	MPEG-4 GMC PSNR-Y[dB]	Proposed GMC PSNR-Y[dB]	GME time-ratio	GMC time-ratio
24	27.35	27.29	9.62	5.02
48	29.98	29.92	10.19	5.3
75	25.47	25.34	6.36	3.81
112	28.93	28.86	7.01	2.71
256	22.68	22.65	6.25	2.64
512	30.99	30.97	7.17	2.51
1 024	32.2	32.18	6.55	2.67
Avg	28.23	28.17	7.59	3.5

注:GME Time-Ratio 是指优化前全局运动估计时间和优化后时间之比;GMC Time-Ratio 是指优化前全局运动补偿编码时间和优化后时间之比。

图 4 和图 5 分别是 Coastguard (QCIF)序列在 48kbps 码率下全运动补偿编码时每帧图像峰值信噪比(PSNR)和全局运动估计(GME)时间的比较。从图 5 可以看出,采用优化算法后全局运动时间明显降低,优化前全局运动估计每帧平均时间是 401.13ms,而优化后则只有 43.0ms,系统速度提高非常明显。从图 4 看出,优化后编码的亮度信号峰值信噪比没有下降很多,优化前的平均信噪比为 29.05dB,而优化后为 29.04dB。图 6 和图 7 分别是 Bus 序列在 112kbps 码率下全运动补偿编码时每帧峰值信噪比(PSNR)和全局运动估计(GME)时间的比较,从图 7 可以看出,采用优化算法后全局运动速度明显提高,优化前全局运动估计每帧平均时间是 852.94ms,而优化后则只有 145.60ms。从图 6 看出,优化后编码的亮度信号峰值信噪比还有所提高,优化前的平均信噪比为 26.91dB,而优化后为 26.93dB。图 8 和图 9 分别是 Basketball 序列在 256kbps 码率下全运动补偿编码时每帧峰值信噪比(PSNR)和全局运动估计(GME)时间的比较,优化前全局运动估计每帧平均时间是 1184.95ms,而优化后则只有 189.57ms,系统速度提高非常明显。从图 8 看出,优化后编码的亮度信号峰值信噪比略有下降,优化前的平均信噪比为 22.68dB,而优化后为 22.65dB。

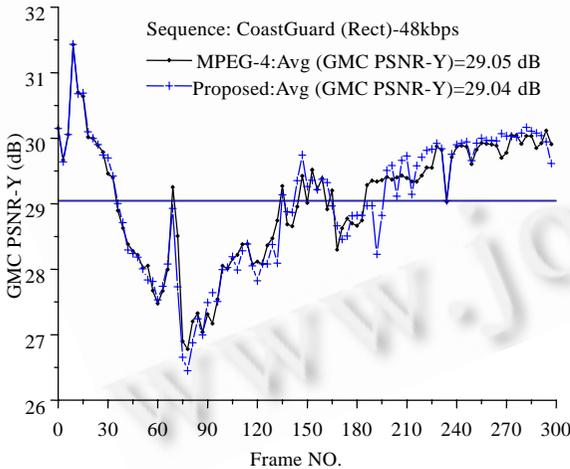


Fig.4 Y-PSNR comparison of Coastguard (QCIF) sequence at 48kbps

图 4 Coastguard(QCIF)序列在 48kbps 时编码的亮度信号峰值信噪比(PSNR)比较

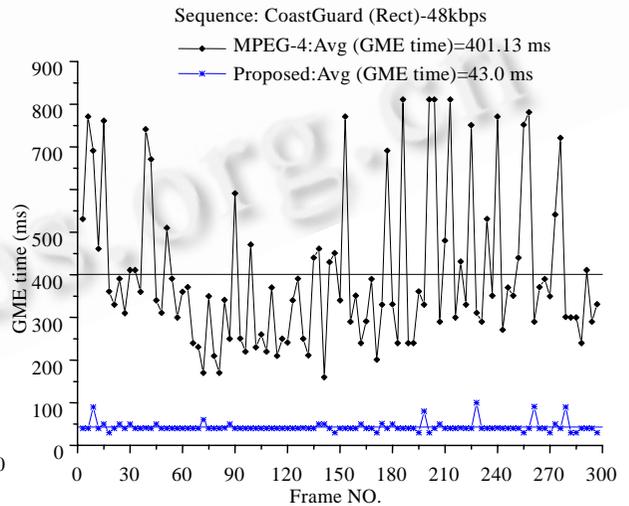


Fig.5 Global motion estimation time comparison of Coastguard (QCIF) sequence at 48kbps

图 5 Coastguard(QCIF)序列在 48kbps 编码时全局运动估计时间比较

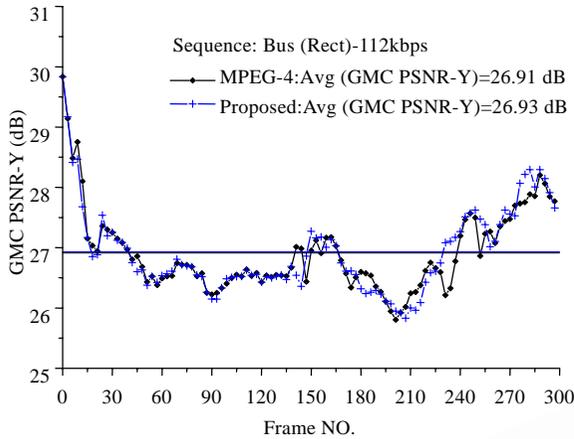


Fig.6 Y-PSNR comparison of Bus (SIF) sequence at 112kbps

图6 Bus(SIF)序列在 112kbps 时编码的亮度信号峰值信噪比(PSNR)比较

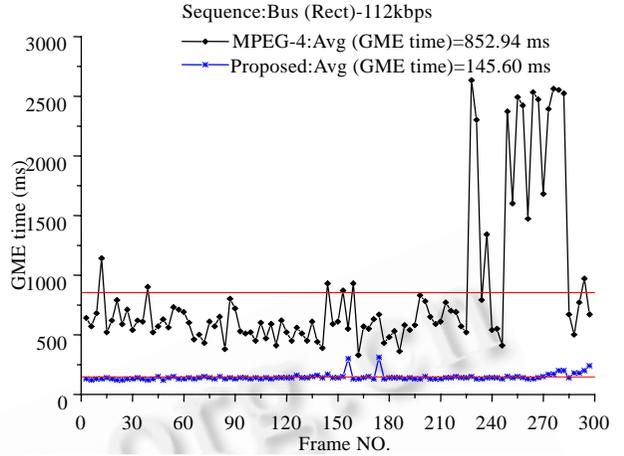


Fig.7 Global motion estimation time comparison of Bus (SIF) sequence at 112kbps

图7 Bus(SIF)序列在 112kbps 编码时全局运动估计时间比较

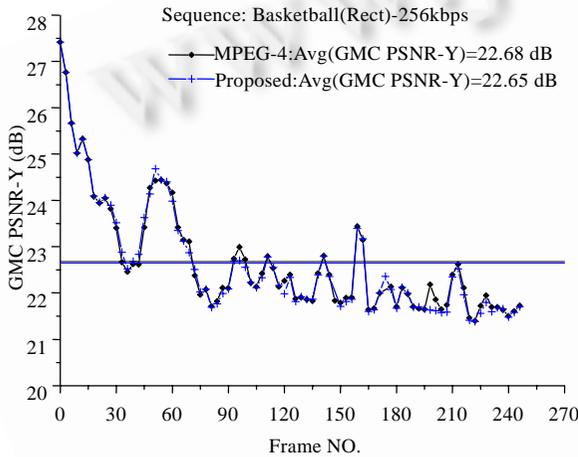


Fig.8 Y-PSNR comparison of basketball (CIF) sequence at 256kbps

图8 Basketball (CIF)序列在 256kbps 时编码的亮度信号峰值信噪比(PSNR)比较

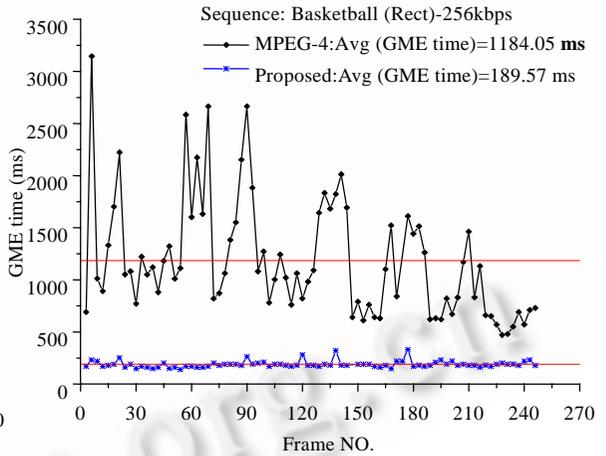


Fig.9 Global motion estimation time comparison of Basket (CIF) sequence at 256kbps

图9 Basketball (CIF)序列在 256kbps 编码时全局运动估计时间比较

### 3 小结

本文提出了对 MPEG-4 低码率下 sprite 编码和全局运动补偿编码方法的优化方法,这主要是通过对这些编码方法中的核心技术——全局运动估计进行优化.在低码率编码方法中,为了提高压缩性能和图像质量,编码系统都具有相当的复杂度,包括许多复杂的视频分析方法,这些都使得编码性能优越的编码方法难以应用到实际编码系统中.本文对这些编码方法进行了优化,其中重点优化了全局运动估计方法,主要是采用运动特征点代替全部像素,并进行鲁棒的目标函数进行计算,提高了这些方法中关键部分全局运动估计的速度,从而在保证质量的前提下提高系统的编码速度,极大地促进了 sprite 编码和全局运动补偿编码的实用性.基于运动特征的运动估计方法非常适合运动估计,这些运动特征能够很好地反映运动所导致的亮度变化,与一般运动估计方法相比,减少了计算量,而且能够保证运动估计的可靠性.本文提出的优化方法可以在更广泛的领域内得到使用.目前本文提出的基于运动特征的运动估计方法还可以进一步提高性能.因为在目前方法中并没有考虑初始值的预测,而相邻帧之间的运动具有较强的相关性,我们下一步将对运动估计的初值预测进行研究.

致谢 这项研究受到微软亚洲研究院与清华大学联合实验室的资助,在此表示感谢.

### References:

- [1] ISO/IEC JTC1/SC29/WG11. Overview of the MPEG-4 standard. No.N3536, Beijing, 2000. <http://www.cselt.it/mpeg>.
- [2] ISO/IEC JTC1/SC29/WG11 MPEG Video Group. MPEG-4 video verification model version 15.0. No. MPEG98/N3093, Hawaii, 1999.
- [3] ITU-Telecommunication Standardization Sector, Study Group 16. Temporary Document 83-E(P) Rev.1. WP 3/16 Meeting Plenary Report, Geneva, 2000.
- [4] ITU-Telecommunications Standardization Sector, STUDY GROUP 16, Video Coding Experts Group (Question 15). H.26L test model long term number 5 (TML-5) draft0. Eleventh Meeting: Portland, Oregon, USA, 2000.
- [5] Sikora, T. MPEG digital video-coding standards. IEEE Signal Processing Magazine, 1997,14(5):82~100.
- [6] Mitchell, J.L. MPEG Video Compression Standard. New York: Chapman & Hall, 1997.
- [7] Smolic, A., Sikora, T., Ohm, J-R. Long-Term global motion estimation and its application for sprite coding, content description, and segmentation. IEEE Transactions on Circuits and Systems for Video Technology, 1999,9(8):1227~1242.
- [8] Kim, E.T., Kim, H.M. Fast and robust parameter estimation method for global motion compensation in the video coder. IEEE Transactions on Consumer Electronics, 1999,45(1):76~83.
- [9] Konrad, J., Dufaux, F. Improved global motion estimation for N3. No. MPEG97/M3096. San Jose, 1998.
- [10] ISO/IEC JTC1/SC29/WG11MPEG Video Group. Optimization model. ISO/IEC JTC1/SC29/WG11 N3675. LaBaule, France, 2000.
- [11] ISO/IEC WG11 MPEG Video Group. Encoder optimization core experiment descriptions. ISO/IEC/JTC1/SC29/WG11, No.MPEG00/N3523, Beijing, 2000.
- [12] He, Yu-wen, Qi, Wei, Yang, Shi-qiang, *et al.* Feature-Based fast and robust global motion estimation technique for sprite coding. In: Meeting of ISO/IEC JTC1/SC29/WG11, MPEG00/M6226. Beijing, 2000.
- [13] He, Yu-wen, Qi Wei, Yang, Shi-qiang, *et al.* Experiment results for feature-based fast and robust global motion estimation technique for sprite coding (FFRGMET). In: ISO/IEC JTC1/SC29/WG11, MPEG00/M6548. LaBaule, 2000.
- [14] Tourapis, A.M., Au, O.C. Verification results of the feature-based fast and robust global motion estimation technique (FFRGMET) for spriteCoding. In: ISO/IEC JTC1/SC29/WG11, MPEG00/M6552. LaBaule, 2000.
- [15] Wu, Feng. Cross-Verification results of feature-based fast and robust global motion estimation technique (FFRGMET) for sprite coding. In: ISO/IEC JTC1/SC29/WG11, MPEG00/M6553. LaBaule, 2000.
- [16] Meyer, F.G., Bouthemy, P. Region-Based tracking using Affine motion models in long image sequences. CVGIP: Image Understanding, 1994,60(2):119~140.
- [17] Striller, C., Konrad, J. Estimating motion in image sequences, a tutorial on modeling and computation of 2D motion. IEEE Signal Processing Magazine, 1999,16(4):70~91.
- [18] Microsoft-FPDAM1-1.0-000703.<ftp://ftp.tnt.uni-hannover.de>.

## An Optimization Approach on Low Bite Rate Video Coding\*

HE Yu-wen, WANG Qi, YUAN Yu, ZHONG Yu-zhuo, YANG Shi-qiang

(Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

E-mail: [heyw@media.cs.tsinghua.edu.cn](mailto:heyw@media.cs.tsinghua.edu.cn)

<http://imlab.cs.tsinghua.edu.cn>

**Abstract:** The coding method in low bit rate discussed in this paper is a hybrid encoding architecture, which combines global motion compensation and local motion compensation coding. In order to solve the complexity and large quantity of calculations of global motion compensation in this kind of hybrid encoder, the feature based global motion estimation and robust object function and three-level pyramid calculation are proposed to improve the performance of encoder system, especially in speed. This optimal global motion estimation technique enhances the

performance of coding system substantially. Some comparison experiments are made in order to validate the effect of the proposed optimization method with different resolutions and sequences in different bit rates. The encoder system speeds up more than 3 times after optimization through comparisons. The proposed method proposed is an attractive technique, which has been accepted by MPEG-4 video encoder optimization group.

**Key words:** low bit rate coding; video coding; global motion; motion compensation; motion estimation

\* Received December 18, 2000; accepted March 21, 2001

Supported by the National High Technology Development 863 Program of China under Grant No.863-306-ZT03-09-1

XX

### 第 9 届 Rough 集、模糊集、数据挖掘与粒度计算国际学术会议(RSFDGrC 2003)

#### 征文通知

RSFDGrC 2003 是 RSFDGrC 系列国际会议的第九届会议，该系列会议每两年举行一次。RSFDGrC 2003 是该系列会议第一次在中国举办。RSFDGrC 2003 将由国家自然科学基金会、重庆邮电学院等有关单位共同主办，于 2003 年 5 月 26 日~29 日在重庆召开。会议将针对 Rough 集、模糊集、数据挖掘、粒度计算等计算智能的理论与应用进行学术研究与讨论。

#### 一、 征文范围

Rough 集理论及应用	计算智能	机器学习	文字计算
Fuzzy 集理论及应用	粒度计算	软计算及其应用	进化计算
Petri 网	软计算的逻辑基础	非经典逻辑	神经网络
软计算复杂性	空间推理	统计推理	智能 Agent
多标准决策分析	决策支持系统	知识发现与数据挖掘	多 Agent 技术
网络智能	集成智能系统	近似推理与不确定推理	数据仓库
模式识别与图像处理	其他有关领域		

#### 二、 征文要求

- (1) 论文未被其他会议、期刊录用或发表；
- (2) 大会工作语言为英语；
- (3) 来稿一式 4 份；
- (4) 为联系方便，请务必提供作者的姓名、单位、通信地址、电话、传真及 EMAIL 地址；
- (5) 论文版面格式及投稿信息请参看会议主页。

#### 三、 重要日期

特邀小组会议申请：2002 年 11 月 10 日（收到日期）  
 征文截止日期：2002 年 12 月 10 日（收到日期）  
 录用通知日期：2003 年 1 月 10 日（发出日期）  
 提交正式论文截止日期：2003 年 2 月 10 日（收到日期）

#### 四、 联系方式

联系地址：400065 重庆市南岸区黄桷垭镇堡上园 1 号      重庆邮电学院计算机科学与技术研究所  
 联系人：RSFDGrC 2003 程序委员会      王国胤教授  
 联系电话：023-62460066      传真：023-62461882  
 E-mail：wanggy@cqupt.edu.cn      Http://www.cqupt.edu.cn/rsfdgrc