

## 基于感兴趣区的 MPEG-4 FGS 增强层码率分配算法

陈 强 卓 力 沈兰荪

(北京工业大学信号与信息处理研究室 北京 100022)

**摘 要:** 该文提出了一种基于感兴趣区的 MPEG-4 FGS 增强层码率分配算法。该算法在建立一种针对感兴趣区的率失真模型的基础上, 在等质量的约束条件下, 可以实现选择性增强后 FGS 增强层码率的优化分配。实验结果表明, 与平均码率分配算法相比, 该算法可以显著降低视频图像中人眼感兴趣区域重建质量的波动。同时, 该算法的计算复杂度低, 完全能满足视频服务器实时码率分配的要求。

**关键词:** MPEG-4 FGS, 选择性增强, 感兴趣区, 率失真模型, 码率分配

**中图分类号:** TN919.8      **文献标识码:** A      **文章编号:** 1009-5896(2005)03-0402-05

## ROI-Based Rate Allocation Algorithm for MPEG-4 FGS

Chen Qiang    Zhuo Li    Shen Lan-Sun

(Signal and Information Processing Lab, Beijing University of Technology, Beijing 100022, China)

**Abstract** A ROI-based rate allocation algorithm for the MPEG-4 FGS enhancement layer is proposed in this paper. On the basis of the ROI-oriented rate-distortion model, the proposed algorithm can perform optimal rate allocation for the enhancement layer after selective enhancement under equal image quality constraint. The experimental results show that, compared with the conventional average rate allocation algorithm, the proposed algorithm can significantly eliminate the quality fluctuation of ROI. Simultaneously, the computational complexity of the proposed method is low enough to meet the requirements of the real time rate allocation at the video server.

**Key words** MPEG-4 FGS, Selective enhancement, Region Of Interest (ROI), Rate-distortion model, Rate allocation

### 1 引言

随着 Internet 和无线通信技术的迅速发展, 视频流已成为重要的网络多媒体传输方式。但由于网络的异构性以及缺乏必要的 QoS 保障, 网络带宽会在很大的范围内产生波动。因此适应网络传输带宽变化, 充分利用有限带宽成为目前视频压缩和传输中的关键问题。为此, 在 MPEG-4 标准中定义了一种细粒度可扩展 (Fine Granularity Scalability, FGS) 视频编码方法<sup>[1]</sup>。该方法将视频编码分为基本层和增强层两个码流, 其中基本层码流用以满足最低的网络带宽要求, 而增强层采用了比特平面 (bit plane) 编码方法, 码流可在任意位置进行截断, 这样码率就可以随着网络带宽的波动进行连续调整。在发送视频流的过程中, 视频服务器可以实时地根据网络带宽的变化, 为各帧增强层码流分配适当的码率并进行截断, 从而可以充分利用网络带宽。

FGS 编码技术的另一个突出特点是可对视频图像中人眼感兴趣区进行选择增强 (Selective enhancement)<sup>[1]</sup>。研究发现, 人们在观察与理解图像时常常会不自觉地对某些区

域产生兴趣, 而整幅图像的视觉质量往往取决于感兴趣区 (Region Of Interest, ROI) 的图像质量<sup>[2]</sup>。例如, 在头肩视频序列中, 人眼最关注的是序列中的人脸区域, 而对背景区域通常并不注意。在网络带宽有限的条件下, 如果能对人眼感兴趣区域优先编码并传输, 则人眼感兴趣区就能获得很高的重建质量, 从而可以大大提高图像的主观感受水平。

人眼视觉系统对相邻帧之间的图像质量波动非常敏感, 因此在有限的网络带宽条件下, 如何合理地各帧增强层分配码率是一个非常重要的问题。传统的平均码率分配算法是将当前可用的网络带宽平均分配给各帧, 这种算法非常简单, 但它没有考虑各帧图像率失真特性之间的差异, 从而造成相邻帧之间图像重建质量的严重波动。目前, 人们已经提出了几种基于图像率失真模型的码率分配算法来降低相邻帧间的质量波动。文献[3]采用指数形式的率失真模型, 文献[4,5]则分别采用了分段线性的率失真模型。但是上述码率分配算法都只适用于一般的 FGS 增强层码流。当进行选择增强后, 整幅图像的率失真特性将会产生明显的变化, 所以上

述各种算法将不能再有效地应用于进行选择性增强后的 FGS 增强层码流。

本文首先针对 FGS 选择性增强后的率失真特性,建立了感兴趣区的分段二次率失真模型。基于该模型,提出了一种 FGS 选择性增强码率优化分配算法。实验结果表明,该算法可以显著降低相邻帧感兴趣区内的质量波动,提高重建视频的主观质量。

## 2 选择性增强的率失真模型

### 2.1 FGS 选择性增强编码原理

MPEG-4 FGS 采用比特平面编码方法对增强层进行编码,该方法按照比特平面的重要性程度依次编码及放置码流。因此,如果有选择地将人眼感兴趣区的系数所在的比特平面进行上移(Up-shift),就可保证这部分系数能够优先编码并传输。这样在带宽有限的条件下对增强层码流进行截断时,感兴趣区部分的重要信息就能保留下来,感兴趣区的解码质量也因此能够增强。图 1 所示的是选择性增强(提升 2 个比特平面)前、后感兴趣区域的重建质量对比。可以看出,在相同的码率下,进行选择增强后人眼感兴趣区的重建质量有很大提高。

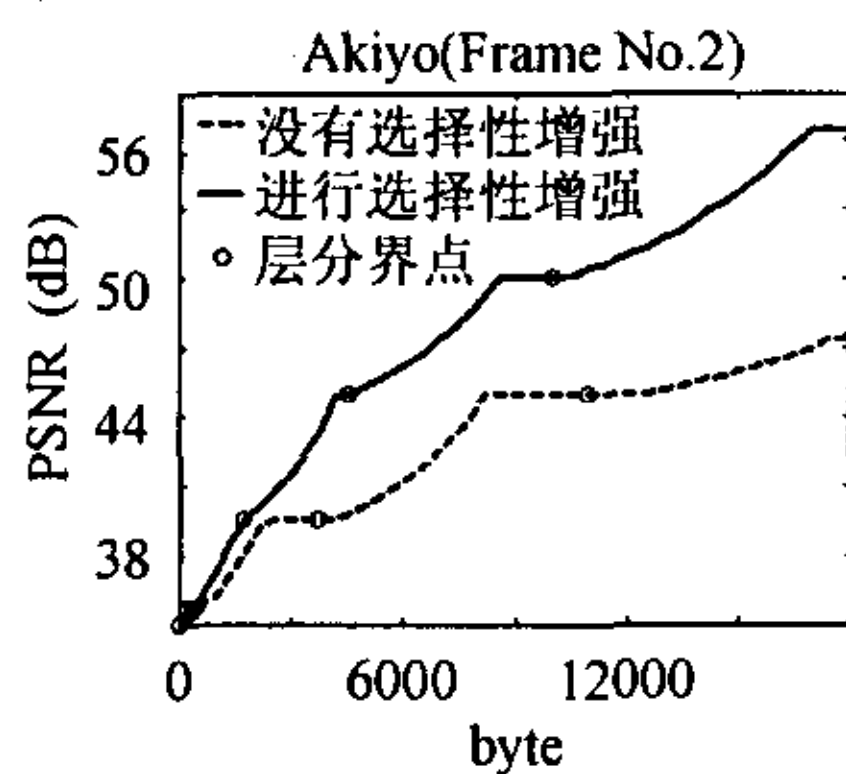


图 1 进行选择增强前、后感兴趣区的质量对比

### 2.2 FGS 选择性增强的率失真模型

率失真模型的准确性直接影响到码率分配算法的有效性。对于一般的未进行选择增强的 FGS 增强层码流来说,同一比特平面中每个比特的重要性程度是相同的,所以率失真特性呈现出分段特性,即一个比特平面对应着一个分段区域。通过大量实验我们发现,如果对每个分段分别采用二次模型就能很准确地描述实际的率失真曲线,如图 2 中的虚线所示。然而,当进行选择增强后,在同一个比特平面对应于感兴趣区和非感兴趣区系数的各比特的重要性程度已不再相同,从而使增强层码流的率失真曲线发生了显著的变化,如图 2 中的实线所示。可以看出,这时各比特平面的曲线很不规则,很难再用统一的函数形式进行描述。

但是考虑到人眼的视觉特性,人眼对于感兴趣区的质量波动会非常敏感,而背景区域即使有一定的波动,人眼一般也不会注意。因此本文建立了针对感兴趣区的率失真模型。

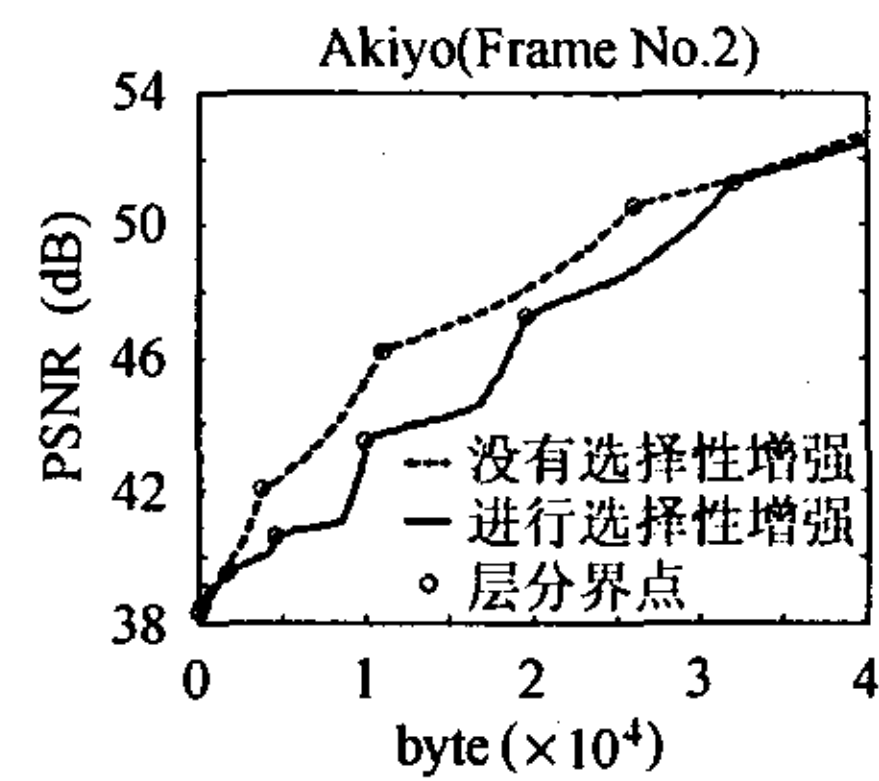


图 2 选择性增强前、后增强层码流的率失真特性

需要指出,由于视频服务器在截断码流时,是对增强层的整体码流进行截断。因此,本文将感兴趣区的率失真关系描述为感兴趣区质量和总码率(包括感兴趣区和非感兴趣区)之间的关系。

图 1 描述的是感兴趣区的率失真曲线,与图 2 中的虚线相比,存在两个明显的差异:第一,每个分段的开始和结尾部分的曲线与 X 轴(码率)是平行的。第二,在每个分段中间一段与 X 轴非平行的曲线上呈现出细小的锯齿状。原因如下:在编码图 3 中的背景区域 A, B 时,产生的编码信息不会对感兴趣区的质量有任何贡献,因而造成了率失真曲线各段之间的不连续性。而在各比特平面中间一段码流中,由于包含了感兴趣区的码流,码率和质量之间因此呈现单调上升的现象。但对背景区域 C, D 产生的编码信息同样不会改善感兴趣区的质量,因此在率失真曲线上就形成了一些微小的锯齿状线段(实际也是很小的水平线段)。

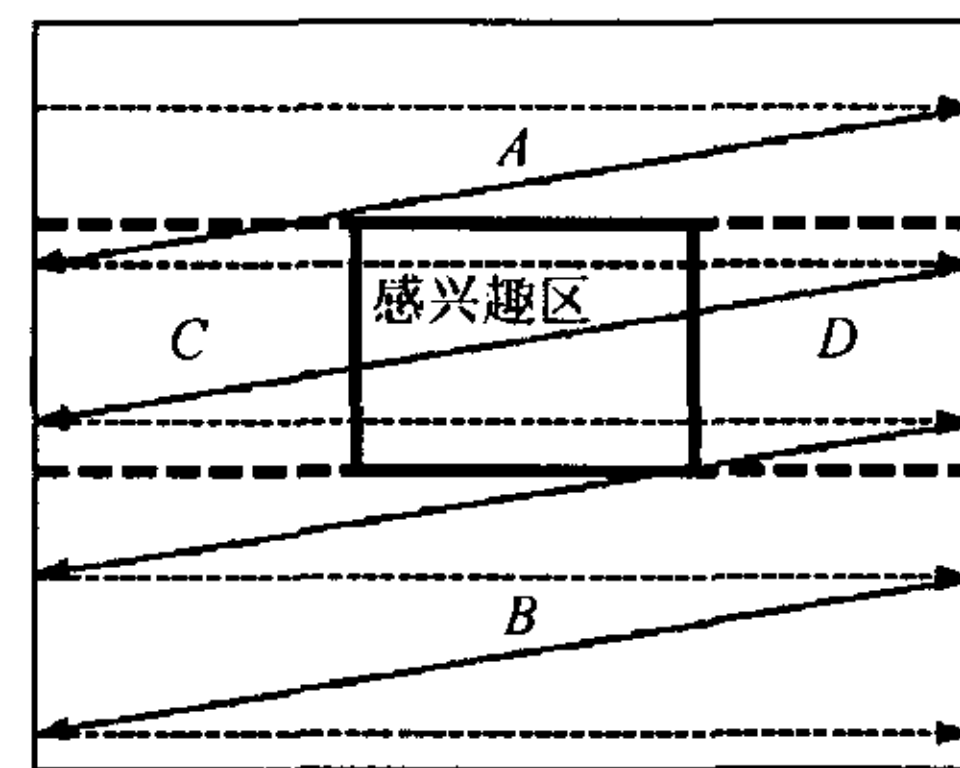


图 3 比特平面编码顺序示意图

根据上述分析,本文建立了感兴趣区质量和增强层整体码率之间的分段二次模型,即对各帧图像的每一比特平面中的感兴趣区分别建立一个二次模型,模型的基本形式为

$$R = A(i, j)Q^2 + B(i, j)Q + C(i, j) \quad (1)$$

其中,  $Q$  表示感兴趣区的图像质量(用 PSNR 度量),  $R$  表示整体码率,  $A(i, j)$ ,  $B(i, j)$  和  $C(i, j)$  为模型的参数,  $i$  代表帧号,  $j$  代表比特平面号。在求解每一层的模型参数时所需的 3 个采样点分别选择为各比特平面中感兴趣区曲线的开始点  $P_0(i, j)$ , 结束点  $P_2(i, j)$  以及  $P_0$  和  $P_2$  的中点  $P_1(i, j)$ 。根据这 3 个点的坐标就可计算出二次模型的参数  $A(i, j)$ ,  $B(i, j)$  和  $C(i, j)$ 。图 4 所示的就是针对感兴趣区的分段二次率失真



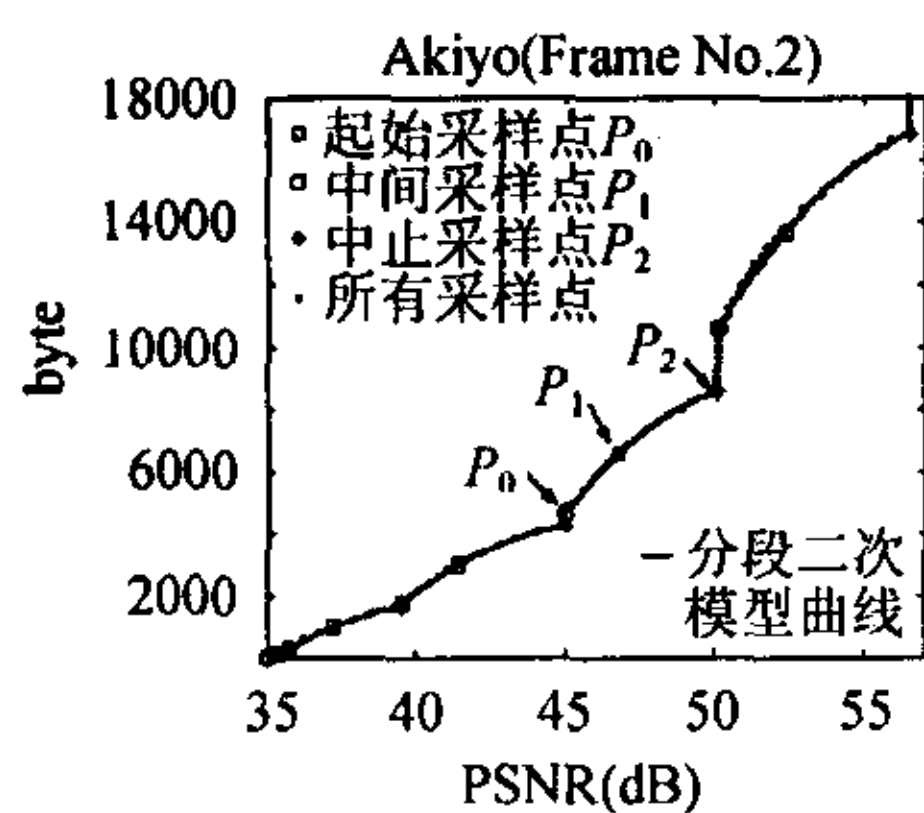


图4 基于感兴趣区的分段二次率失真模型

模型。可以看出,这种分段二次模型可以非常精确地描述实际的感兴趣区质量与整体码率之间的率失真关系。

由于 DCT 变换是一种归一 (Unitary) 变换,像素域的失真可以通过 DCT 系数来进行度量,即  $d(I, \hat{I}) = d(C, \hat{C})$ 。其中  $I$  表示原始图像,  $\hat{I}$  表示重建图像,  $C$  表示 DCT 系数,  $\hat{C}$  表示重建后的 DCT 系数。因此对于 FGS 来说,率失真特性可以在编码过程中同时得到。为了码率分配算法的需要,这些采样点的信息以及计算出的各模型系数都存储在服务器中,用于码率分配。与视频压缩码流的大小相比,这些额外存储的信息占用空间很小,不会为服务器造成存储负担。

### 3 基于感兴趣区的码率分配算法

#### 3.1 算法描述

在本文中,将基于感兴趣区的码率分配算法的目标定为,在满足可用带宽的约束条件下,使相邻帧人眼感兴趣区的重建质量波动达到最小。设以  $N$  帧为一组进行码率分配,  $R_{\text{sum}}$  表示这  $N$  帧可用码率的总比特数,  $R_i$  表示分配给第  $i$  帧增强层的比特数,  $Q_i$  表示在  $R_i$  下第  $i$  帧中感兴趣区的解码质量,这样码率分配问题可以表示为

$$\min \left( \sum_{i=0}^{N-2} |Q_i(R_i) - Q_{i+1}(R_{i+1})| \right), \quad \text{s.t.} \quad \sum_{i=0}^{N-1} R_i \leq R_{\text{sum}} \quad (2)$$

对式(2)的求解可以根据这  $N$  帧的率失真曲线,采用全局搜索的方法,直接寻找到一组最优解  $(R_0^O, R_1^O, \dots, R_{N-1}^O)$ , 使得  $\sum_{i=0}^{N-2} |Q_i(R_i^O) - Q_{i+1}(R_{i+1}^O)|$  为最小值,这种方法最为直接,但是计算量会随着帧数  $N$  的增加迅速上升,无法实现码率的实时分配。

考虑到如下两个事实:(1)码率分配的理想结果是使相邻帧感兴趣区的质量不产生任何波动,各帧感兴趣区的质量完全相等,即  $Q_i = Q_{\text{opt}}$ ,  $0 \leq i \leq N-1$ ;(2)码率  $R_i$  和质量  $Q_i$  二者之间是单调上升的关系。为了简化求解过程,本文用  $R_i(Q_i)$  替代  $Q_i(R_i)$  表示各帧的率失真函数,即将码率  $R_i$  描述为质量  $Q_i$  的函数。这样问题可以转化为在满足  $\sum_{i=0}^{N-1} R_i(Q_{\text{opt}}) \leq R_{\text{sum}}$  的条件下,寻找最优的相同质量  $Q_{\text{opt}}$ , 然后通过  $Q_{\text{opt}}$  与各帧的率失真关系  $R_i(Q_i)$  求出各帧增强层应

该分配的码率  $R_i^O = R_i(Q_{\text{opt}})$ ,  $0 \leq i \leq N-1$ 。

#### 3.2 算法实现

这种码率分配算法的有效性主要取决于两点:(1)精确的率失真模型,(2)需要快速准确计算出最优相同质量  $Q_{\text{opt}}$ 。如前所述,本文提出的分段二次模型可很好地满足第(1)点要求。对于第(2)点,首先根据各帧的率失真函数  $R_i(Q)$ , 定义  $N$  帧的联合率失真函数:

$$R(Q) = \sum_{i=0}^{N-1} R_i(Q) = \sum_{i=0}^{N-1} (A(i, j)Q^2 + B(i, j)Q + C(i, j)) \quad (3)$$

然后设:

$$f(Q) = \sum_{i=0}^{N-1} R_i(Q) - R_{\text{sum}} \quad (4)$$

显然  $f(Q)$  是单调递增的。当  $Q = Q_{\text{opt}}$  时,有

$$f(Q) = 0 \quad (5)$$

根据式(3)和式(5)可以得到

$$\sum_{i=0}^{N-1} (A(i, j)Q_{\text{opt}}^2 + B(i, j)Q_{\text{opt}} + C(i, j)) - R_{\text{sum}} = 0 \quad (6)$$

将式(6)进一步改写成如下形式:

$$Q_{\text{opt}} \left( Q_{\text{opt}} \sum_{i=0}^{N-1} A(i, j) + \sum_{i=0}^{N-1} B(i, j) \right) + \left( \sum_{i=0}^{N-1} C(i, j) - R_{\text{sum}} \right) = 0 \quad (7)$$

其中  $R_{\text{sum}}$  为已知,  $A(i, j)$ ,  $B(i, j)$  和  $C(i, j)$  已经在编码过程中得到,可以根据服务器端存储的信息来确定。求解式(7)就可以得到  $Q_{\text{opt}}$ , 然后根据式(1)和  $Q_{\text{opt}}$  就可得到最终的码率分配结果  $R_i^O$ ,  $i = 0, \dots, N-1$ 。

#### 3.3 算法复杂度分析

FGS 增强层的码率分配算法不能太复杂,否则将会影响视频服务器的性能。下面我们对本文算法的复杂度进行分析。从前面的分析可以看出:由于模型参数  $A(i, j)$ ,  $B(i, j)$  和  $C(i, j)$  可以在编码过程中得到,因此视频服务器端的计算量主要集中在根据式(7)和式(1)求最优  $Q_{\text{opt}}$  和  $R_i^O$ 。考虑到  $f(Q)$  是单调的,本文采用“二分法”来求解  $Q_{\text{opt}}$ , 方法如下:

- Step 1 根据  $R_{\text{sum}}$  选择初始区间  $[Q_1, Q_2]$ , 使  $f(Q_1)f(Q_2) < 0$ .
- Step 2  $Q = (Q_1 + Q_2)/2$ , Calculate  $f(Q)$ .
- Step 3 If  $(f(Q) \cdot f(Q_1)) > 0$  Then  $Q_1 = Q$ ; Else  $Q_2 = Q$ .
- Step 4 If  $(|f(Q)| < \varepsilon)$  Then Go to End; Else Go to Step 2.
- End  $Q_{\text{opt}} = Q$ .

其中  $\varepsilon$  是预定的阈值。在二分法中,计算量集中在求  $f(Q)$  部分,根据式(7),计算一次  $f(Q)$  仅需要 2 次乘法和  $3N+3$  次加法,而计算  $R_i$ ,  $i = 0, \dots, N-1$  共需要  $2N$  次乘法和  $2N$  次加法,因此本文算法的计算量和  $N$  成正比,完全能满足视频服务器实时码率分配的要求。

#### 4 实验结果及分析

标准测试序列和实验条件如表1所示。在实验中,基本层编码格式为每10帧含一个I帧,其余9帧为P帧,并将感兴趣区定义为包含人脸区域的矩形(位置由矩形左上角和右下角坐标表示)。为了衡量质量波动,本文采用了平均质量波动和最大质量波动两个指标:

$$\text{平均质量波动: } \varepsilon_{\text{avg}} = (1/N-1) \sum_{i=0}^{N-2} |Q_i - Q_{i+1}|$$

$$\text{最大质量波动: } \varepsilon_{\text{max}} = \max_{0 \leq i \leq N-1} (Q_i) - \min_{0 \leq i \leq N-1} (Q_i)$$

表1 测试序列和测试条件

视频序列	图像格式	帧数	基本层码率(kbps)	增强层码率(kbps)	帧率(fps)	感兴趣区位置
Foreman	CIF	100	128	480	10	(96,64) (272,272)
Akiyo	CIF	100	128	480	10	(128,32) (224,192)
Silent	CIF	100	128	480	10	(128,16) (208,128)
Carphone	QCIF	100	32	120	10	(48,16) (128,112)
Claire	QCIF	100	32	120	10	(64,16) (112,96)
Miss A	QCIF	100	32	120	10	(48,16) (128,112)
Salesman	QCIF	100	32	120	10	(80,16) (128,80)

表2 视频图像质量的平均波动和最大波动对比(dB)

视频序列	平均码率分配算法		本文提出的基于感兴趣区码率分配算法			
			以20帧为一组进行分配		以100帧为一组进行分配	
	$\varepsilon_{\text{avg}}$ (dB)	$\varepsilon_{\text{max}}$ (dB)	$\varepsilon_{\text{avg}}$ (dB)	$\varepsilon_{\text{max}}$ (dB)	$\varepsilon_{\text{avg}}$ (dB)	$\varepsilon_{\text{max}}$ (dB)
Foreman	0.52	5.37	0.06	1.63	0.04	0.99
Akiyo	0.38	3.97	0.06	1.31	0.05	0.19
Silent	0.17	1.37	0.08	0.75	0.06	0.32
Carphone	0.16	1.77	0.04	0.91	0.03	0.17
Claire	0.30	3.91	0.06	1.00	0.06	0.28
Miss A	0.33	3.58	0.11	2.58	0.09	0.47
Salesman	0.28	5.37	0.18	4.46	0.17	1.04

需要说明的是,码率分配组中包含的帧数将会影响到最后的分配结果。从图5可以看出,同一码率分配组中各帧的重建质量波动很小,但不同帧组内各帧的重建质量则仍然存在着一定波动。在具体的应用中,可以根据具体的情况决定一个帧组内包括的帧数。当网络带宽变化比较平缓时,可以将更多的帧作为一组进行码率分配,这样可进一步消除质量波动。如表2所示,以100帧为一组进行码率分配时,基本可使最大波动降低到1dB以下。

表2是对各序列分别采用传统的码率平均分配算法和本文所提出的基于感兴趣区的码率分配算法进行码率分配后,感兴趣区解码质量的波动对比结果。可以看出,传统的平均码率分配算法会造成感兴趣区重建质量的明显波动,而本文提出的基于感兴趣区的码率分配算法则可以明显降低这种波动。图5表示了分别以20帧、50帧和100帧为一组进行码率分配后得到的结果。同一码率分配组中的各帧可以获得几乎相等的重建质量。

#### 5 结论

本文在详细分析了选择性增强后感兴趣区率失真曲线的特点后,建立了一种分段二次模型来表示感兴趣区的重建质量和增强层整体码率之间的关系。基于此模型,提出了一种基于感兴趣区的码率分配算法。该算法在满足网络带宽的约束条件下,可以最优分配各帧增强层的码率,从而降低相邻帧之间感兴趣区重建质量的波动,提高视频图像的主观感

受水平。实验结果表明,与传统的平均码率分配算法相比,本文方法可以明显降低感兴趣区的质量波动,特别是当网络带宽变化平缓时,可以将更多的帧作为一组进行码率分配,这样可进一步降低质量波动。同时,本文方法的计算复杂度低,完全能满足视频服务器实时码率分配的要求。

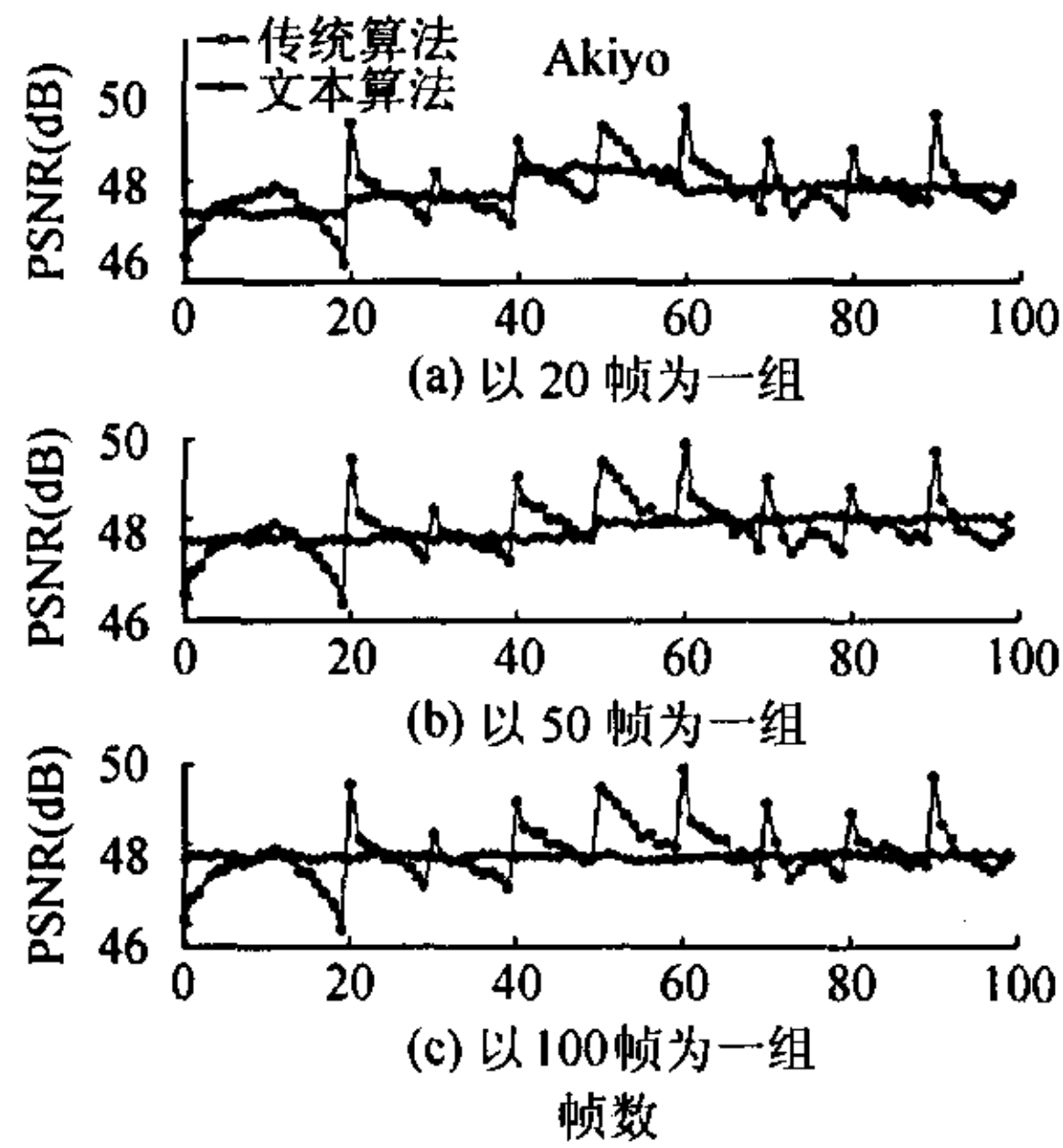


图5 码率分配中的帧组大小对减小质量波动的影响

### 参考文献

[1] Li Wei-ping. Overview of fine granularity scalability in MPEG-4

video standard. *IEEE Trans. on Circuits and Systems for Video Technology*, 2001, 11(3): 301 – 317.

- [2] 沈兰荪, 卓力, 田栋, 汪孔桥. 视频编码与低速率传输. 北京: 电子工业出版社, 2001, 12: 91 – 92.
- [3] 王琪, 赵黎, 吴枫, 李世鹏, 钟玉琢. 可伸缩性视频编码的码率分配算法. *电子学报*, 2002, 30(2): 205 – 209.
- [4] Zhang Xi-Min, Vetro A, Shi Yun Q, Sun Huifang. Constant quality constrained rate allocation for FGS coded video. *IEEE Trans. on Circuits and Systems for Video Technology*, 2003, 13(2): 121 – 130.
- [5] Zhao Xue-jun, He Yuwen, Yang Shiqiang, Zhong Yuzhuo. Rate allocation of equal image quality for MPEG-4 FGS video streaming. *12<sup>th</sup> International Packet Video Workgroup (PV'2002)*, Pittsburgh, April 2002: 24 – 26.

陈强: 男, 1979年生, 硕士生, 研究方向为视频编码与传输技术.

卓力: 女, 1971年生, 副教授, 研究方向为图像视频编码与传输、智能化信息处理.

沈兰荪: 男, 1938年生, 教授, 博士生导师, 研究方向为图像与视频信号的压缩编码、处理与传输等.