

一种利用视觉掩蔽效应的感兴趣区域 视频编码方法

刘琼¹⁾ 胡瑞敏¹⁾ 韩镇²⁾

¹⁾(武汉大学计算机学院, 武汉 430072) ²⁾(武汉大学国家多媒体软件工程技术研究中心, 武汉 430072)

摘要 传统感兴趣区域视频编码方法通过整体降低 ROI 区域的量化参数来实现 ROI 质量优化,会造成整体率失真性能的下降,尤其在低码率时下降更加严重。针对这一问题,结合人眼掩蔽特性的精细 Grid 量化方法在文中提出,可有效提高 ROI 质量并降低整体质量的损失。而且该量化方法能够根据 ROI 尺寸大小对初始感知权重进行自适应调节,适用于 ROI 特征不同的视频序列。实验结果表明,对于不同特征的标准测试序列,在中低码率条件下和 H.264 相比,本文算法能平均提高 ROI 的 PSNR 值 1.905 dB,主观感受有明显提升。和同类 ROI 算法相比,ROI 质量平均提升 0.47 dB,整体质量提高 0.625 dB。

关键词 视觉掩蔽效应 视频编码 感兴趣区域 量化

中图分类号: TN919.81 文献标志码: A 文章编号: 1006-8961(2010)06-861-06

Region-of-Interest Video Coding Algorithm Based on Masking Effect

LIU Qiong¹⁾, HU Ruimin¹⁾, HAN Zhen²⁾

¹⁾(School of Computer, Wuhan University, Wuhan 430072)

²⁾(National Engineering Research Center for Multimedia Software, Wuhan 430072)

Abstract Usually, the Region of Interest(ROI) video coding algorithm is implemented by using a smaller quantization parameter(QP) or more bit-plane inside ROI. It caused the degradation of quality by the artifacts of boundary and additional bit-rate burden in ROI. This paper proposes the grid quantization algorithm to avoid these disadvantages. The macroblocks inside of ROI are divided into three kinds: Grid A, Grid B and Gradient. Different partition has different quantization parameter based on visual masking effect. The test results show that the proposed algorithm can promote the ROI objective quality about 1.9 dB on average, and the subjective quality is also improved obviously. Compared to other ROI quantization algorithm, the ROI quality is improved by 0.47 dB, while the whole picture quality is improved by 0.625 dB. The proposed algorithm can reduce the bit-rate burden of ROI and maintain subjective quality.

Keywords masking effects, video coding, region-of-interest, quantization

0 引言

视频压缩是视频图像处理领域的重点研究内容,它可以有效节省视频图像信息存储空间、减少传输信道的容量。压缩比和重建质量之间的矛盾一直

是压缩领域的重要问题。为此,人们提出了感兴趣区域(ROI)编码技术。其基本思想是对图像中感兴趣的区域,采用低压缩比的压缩,以获得较高的图像质量;而对于背景图像,则采用高压比压缩。

目前 ROI 视频编码方法主要分为两类:位平面提升方法和量化系数(QP)调节方法。其中,位平面

基金项目:国家自然科学基金项目(60772106);国家科技支撑计划项目(2008BAH30812,2008BAH30809);国家科技重大专项项目(2009ZX03003-009,2008ZX03003-04,2009ZX03002-009)

收稿日期:2008-09-04;改回日期:2009-02-03

第一作者简介:刘琼(1981—),女,武汉大学计算机应用专业博士研究生。主要研究方向为视频编码与通信。E-mail: liuq_who@163.com

提升方法以位平面编码技术为基础,通过对一帧中重要区域的系数进行位平面提升来实现 ROI 的质量优化。MPEG-4 FGS(精细可分级)框架中的选择性增强(SE)机制^[1-2]就是在质量增强层对 ROI 区域的系数进行选择性的位平面提升,以实现精细的 ROI 质量优化。文献[3]以此为基础,结合人脸检测算法,有效地实现了脸部的精细质量增强。但是由于位平面提升方法增加了位平面的总数目,不可避免地增加了码流开销。文献[4]指出宏块级别的选择性增强会导致超过 2 dB 的整体客观质量下降。

量化系数调节方法的基本思想是对 ROI 设置比其他区域更小的 QP 值,以提高 ROI 的视频质量。QP 值的确定主要有两种方法:一是基于区域(宏块)的码率分配模型^[5],对 ROI 区域设置比其他区域更多的目标码率,然后根据码率模型计算出各个宏块的 QP 值;二是预先设定 ROI 区域和其他区域的 QP 值比例,然后编码器初始化一个 QP 值,其他 QP 值则依据比例计算得到^[6-7]。相对于位平面提升算法来说,量化系数调节方法具有复杂度低、容易实现的特点。但是该类方法会带来率失真性能的下落。这是因为该类方法基于的前提假设是:码率的提升一定会带来质量的提升。但是人眼视觉特性的研究表明,人眼对于图像质量变化存在一定的察觉阈值,在某些情况下,码率的提升不一定会带来主观质量的提升。目前的研究较少考虑到人眼的主观视觉特性,所以通常会导致 ROI 占用过多的码率,引起整体率失真性能下降。特别是在 ROI 面积比较大的情况下,这个问题更加明显。

因此,本文从人眼视觉特性出发,提出基于视觉掩蔽特性的 ROI 视频编码方法,在同等码率条件下提高 ROI 重建质量。

1 视觉掩蔽效应

图像质量的变化可以看做是对视觉的一种刺激,能够引起察觉的最小刺激量,叫做差别感知阈值^[8]。也就是说,当质量变化在一定阈值内时,人眼是无法识别出来的^[9]。而感知阈值和视觉掩蔽效应(Masking)密切相关。掩蔽效应一般解释为一个信号的感知被另一种信号所阻挡,也就是由于其他信号的影响(干扰),该信号差别感知阈值被提升。比如,一个失真信号容易在图像平滑区域被察觉,而在高频成分丰富的区域可以被掩盖。该现象

被称为纹理掩蔽效应(texture masking),即人类视觉系统对于图像中某块区域内容的感知程度随着周围相邻频率的增大而呈现出下降趋势。

在视频压缩系统中引入视觉掩蔽效应,在不影响主观视觉的前提下减小码率代价,可以有效地提高编码率失真性能,改善低比特率下重建图像的视觉效果。

根据视觉掩蔽效应和 ROI 的一般特征,作如下考虑:ROI 区域一般包括运动物体、人脸头像等比较复杂的纹理区。所以,可以利用纹理掩蔽效应对感兴趣区域内部进行编码优化,解决传统分区量化方法中纹理复杂区域往往占用过多的码率,而质量又没有得到应有提升的缺陷。

2 基于视觉掩蔽效应的 ROI 编码方法

ROI 的确定可以通过以下两种方式完成:一种是通过用户对其感兴趣的区域进行指定,称为人为检测的感兴趣区域(human detected region of interest);另一种是根据被处理图像的特征利用算法获得感兴趣区域(algorithmically detected region of interest)^[10-11]。在采用检测算法确定 ROI 位置的前提下,主要研究如何通过优化 ROI 区域的量化模型,提高 ROI 算法的编码性能。

2.1 精细 Grid 量化模型

当 ROI 位置确定后,一帧图像中所有的宏块可以被分为 ROI 宏块和非 ROI 宏块。同时,可在 ROI 边界上增加过渡带^[7],提供质量上的过渡,避免 ROI 到非 ROI 的质量突变。通过分别对非 ROI 宏块、过渡区宏块和 ROI 宏块采用依次递减的量化参数来实现 ROI 的质量优化。进一步,对 ROI 内部区域的宏块再进行一次分类,利用纹理掩蔽效应,对 ROI 内部区域的 QP 值进行加权修正。

首先,将 ROI 内部所有的宏块样本划分为两类:GridA 和 GridB,具体如图 1 所示。对其中的一类样本(比如 GridB 宏块)进行纹理掩蔽权重调整,以降低码率。而对另一类宏块(如 GridA 宏块)则采用相对更小的量化参数,用来提升 ROI 的质量。根据视觉掩蔽效应,由于其他信号的影响,信号本身的差别感知阈值可以被提升。当采用如图 1 所示的格状掩模时,GridA 和 GridB 两类样本间隔出现,GridB 宏块周围环绕质量较好的 GridA 宏块,可以最大程度地提升 GridB 宏块的误差感知阈值。只要将

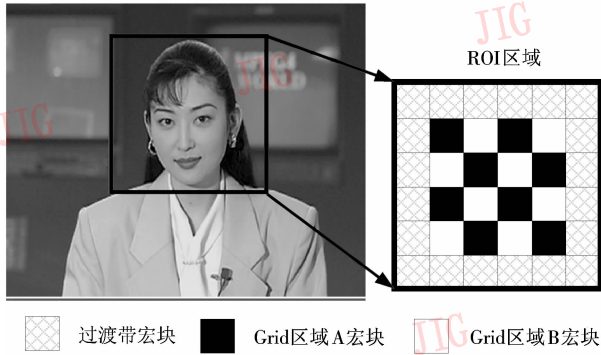


图1 Grid 量化模型区域划分示意图

Fig.1 Grid quantization algorithm

两类样本的质量差别控制在一定阈值之内,人眼将很难察觉到质量的变化。但是,如果对 GridB 宏块的质量调整过大,会使得 GridA 宏块和 GridB 宏块的质量相差过大,导致感兴趣区域产生明显的区域误差,影响重建质量;如果对 GridB 宏块的质量调节不够,则无法对码率进行有效降低。所以,GridB 宏块的质量调节需要在充分考虑 GridA 宏块和 GridB 宏块之间质量过渡性的前提下降低码率。本文算法通过将 GridB 宏块 QP 值的修正权重和过渡带中最接近 ROI 的宏块样本的权重设置为一样来解决 GridA 宏块和 GridB 宏块的质量差别控制问题。这是因为,文献[7]中过渡带的思想本身就考虑了视觉上的过渡性,能够较好地隐藏质量的变化;同时 ROI 区域中有一半以上的宏块采用了更小的 QP 值,可有效降低码率,提高率失真性能。而且,过渡带的 QP 值本身就需要计算出来,所以无需额外的计算量。

综上所述,对于每一帧图像来说,所有的宏块可以分为非 ROI 宏块、过渡带宏块、Grid 区域 A 宏块和 Grid 区域 B 宏块这 4 个类型,其中后面 3 个类型的宏块都属于 ROI 区域。下面分别对这 4 个类型宏块的 QP 值进行计算:

设 QP_{init} 是每一帧的初始 QP 值,一般在编码器初始化时给出。 P 为 ROI 的感知权重, QP_{temp} 为计算各个区域 QP 值的临时变量, QP_{NROI} 为计算后的非 ROI 宏块的量化参数,则有:

$$QP_{NROI} = QP_{init} \quad (1)$$

$$QP_{temp} = QP_{NROI}/P \quad (2)$$

设 W 代表过渡带的宽度, i 代表当前过渡带宏块离 ROI 边界的距离,过渡带的量化参数为 QP_i , ($i=0, 1, \dots, W-1$), 则有:

$$QP_i = (QP_{NROI} - QP_{temp}) \times (W - i)/(W + 1) + QP_{temp} \quad (3)$$

设 Grid 区域 A 的量化参数为 QP_{gridA} , Grid 区域 B 的量化参数为 QP_{gridB} , 则有:

$$QP_{gridA} = QP_{temp} \quad (4)$$

$$QP_{gridB} = (QP_{ROI} + \min(QP_i) + 1)/2 \quad (5)$$

将式(1)(2)带入式(3)~(5), 分别得到式(6)~(8), 即得到过渡带宏块、Grid A 宏块和 Grid B 宏块的 QP 值计算公式如下:

$$QP_i = (QP_{init} - QP_{init}/P) \times (W - i)/(W + 1) + QP_{init}/P \quad (6)$$

$$QP_{gridA} = QP_{init}/P \quad (7)$$

$$QP_{gridB} = (QP_{init}/P + \min(QP_i) + 1)/2 \quad (8)$$

通过以上公式的计算,在给定初始量化参数值 QP_{init} 和 ROI 感知权重 P 的情况下,可以得到各个不同视觉区域的 QP 值,进行编码。

2.2 根据 ROI 特征的自适应调整

当 ROI 面积增大时,由于图像中具有较小量化参数的宏块数目增多,所以对整体率失真性能的影响会加大,需要对量化模型进行进一步的调整。

根据人类视觉心理特征,对于同样内容的感兴趣度随着 ROI 面积的增大而降低。所以,感知权重 P 应当随 ROI 面积的变化而调整。当 ROI 的面积逐渐增大时,相对的兴趣程度就会降低, P 应该减小,以降低 ROI 占用的码率;而当 ROI 面积减小时, P 应该增大。基于这个思想,对 Grid 量化模型的初始感知权重 P 进行基于面积的自适应调整。设 $S_{picture}$ 是整幅图像的面积, S_{ROI} 是 ROI 的面积,则

$$P = \frac{\alpha S_{picture}}{k S_{ROI} + S_{picture}} \quad (9)$$

式中, α 为用户选定的常量,代表了 ROI 的重要程度。 k 为 ROI 面积对兴趣度的影响参数,大量实验结果表明, $k=1.2$ 时符合大多数序列情况。

2.3 编码框架

在传统编码框架的基础上,主要对量化模块进行改进。其主要思想是在一般编码框架的量化模块之前,通过 Grid 量化模型对不同视觉区域的宏块计算量化参数。而不同区域的感知权重受 ROI 位置信息、ROI 特征的约束。其中,ROI 位置信息决定了当前宏块是否属于 ROI 区域。ROI 的主要特征在这里主要表现为 ROI 的面积,这是因为在 ROI 面积较大时,更容易造成整体率失真性能的大幅下降,需要

进一步对权重进行调整。整体的编码框架示意图如图 2 所示。

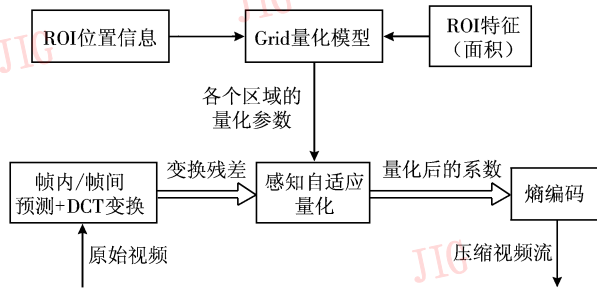


图 2 基于视觉感知的 ROI 量化算法编码框架

Fig. 2 ROI encoder framework of perceptual quantization algorithm

3 实验结果与分析

本文算法在 H. 264 的参考软件平台 JM10. 2 上实现。测试条件如表 1 所示。

表 1 测试条件
Tab. 1 Test conditions

测试帧数	30 帧
运动搜索范围	16 像素
RDO(率失真优化)	0
参考帧数目	1 帧
序列结构	IPPP
帧率	30 帧/s
测试平台	JM10. 2

在国际标准测试序列中,选择 ROI 面积大小不同和运动程度不同的 3 个典型测试序列进行实验,如表 2 所示。

设 $\alpha = 2$, 则利用式(9)计算得到 News, Foreman 和 Coastguard 的 ROI 感知权重分别为 1. 2, 1. 3 和 1. 5。设过渡带宽度为 16 个像素,即一个宏块。采用峰值信噪比(PSNR)和编码输出码流的比特率(bitrate)作为客观性能参数。本实验在常用带宽 256k 和 128k 下进行测试,以 JM10. 2 的算法为基准算法,以文献[7]算法为同类算法进行性能比较。对每个算法的重建图像的整体 PSNR 值(P_{average})和 ROI 的 PSNR 值(P_{ROI})进行统计。 P_{average} 的计算公式如下:

表 2 测试序列
Tab. 2 Test sequences

序列(cif)	ROI 选择	序列特征描述
News		ROI 面积约占 50%, ROI 内部纹理丰富,且局部运动较多。背景基本静止。
Foreman		ROI 面积约占 40%, ROI 纹理比较平滑。序列整体运动程度中等。
Coastguard		ROI 面积约占 10%, ROI 纹理复杂但基本无运动。背景纹理复杂存在全局运动。

$$P_{\text{average}} = \frac{10 \lg(2^n - 1)^2}{\frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N [f'_i(x, y) - f_i(x, y)]^2} \quad (10)$$

式中, n 为 8, M 和 N 分别表示图像的宽度和高度, $f'_i(x, y)$ 和 $f_i(x, y)$ 分别表示像素 $f(x, y)$ 的重建值和原始值。

P_{ROI} 的计算公式如下:

$$P_{\text{ROI}} = \frac{10 \lg(2^n - 1)^2}{\frac{1}{M' \times N'} \sum_{i=1}^{M'} \sum_{j=1}^{N'} [f'_{\text{ROI}}(x, y) - f_{\text{ROI}}(x, y)]^2} \quad (11)$$

式中, n 为 8, M' 和 N' 分别表示 ROI 的宽度和高度, $f'_{\text{ROI}}(x, y)$ 和 $f_{\text{ROI}}(x, y)$ 分别表示像素 $f(x, y)$ 的重建值和原始值,且像素 $f(x, y)$ 属于 ROI 区域。

以 JM10. 2 的测试结果为基准,分别求文献[7]算法和本文算法与基准之间的差值。测试结果如表 3 所示,其中负数表示 PSNR 值有所降低;正数表示 PSNR 值有所升高。由于带宽条件是一定的,所以 PSNR 值的高低可以代表率失真性能的高低。原则上,ROI 算法的目标是能够更好地提高 ROI 的率失真性能,同时尽可能减少整体率失真性能的降低。

对比表 3 的结果可以看出,采用 ROI 算法后(文献[7]和本文算法),ROI 的客观质量分别得到 1. 435 dB 和 1. 905 dB 的增益。但是文献[7]算法的整体 PSNR 值平均降低 2. 145 dB,这会造成明显可察觉的主观质量损失,并且整体率失真性能有大幅降低。而本文算法将整体 PSNR 值的损失降低为 1. 52 dB,能够有效缓解这个问题。

表3 测试结果
Tab.3 Test results

序列	带宽/k	PSNR/dB		
		文献[7]算法	本文算法	
News	256	$\Delta P_{\text{average}}$	-2.14	-1.48
		ΔP_{ROI}	-0.34	0.06
	128	$\Delta P_{\text{average}}$	-2.24	-1.65
		ΔP_{ROI}	-0.15	0.21
Foreman	256	$\Delta P_{\text{average}}$	-1.83	-1.24
		ΔP_{ROI}	2.33	2.48
	128	$\Delta P_{\text{average}}$	-3.77	-2.77
		ΔP_{ROI}	0.51	1.33
Coastguard	256	$\Delta P_{\text{average}}$	-1.2	-0.72
		ΔP_{ROI}	3.78	4.14
	128	$\Delta P_{\text{average}}$	-1.69	-1.28
		ΔP_{ROI}	2.48	2.94
平均值	$\Delta P_{\text{average}}$	-2.145	-1.52	
	ΔP_{ROI}	1.435	1.905	

分析各个测试序列在带宽分别为256 k和128 k下的测试结果,可以看出在256 k时,ROI的质量提高更加明显,同时整体率失真性能的损失更小。这说明在中码率下 ROI算法可以在有效提升 ROI质量的同时也保证整体图像有良好的视觉质量。

对比表2中所列出的各个序列的 ROI面积,可以发现传统的量化方法在 ROI面积比较大时(比如 News序列,ROI占50%以上的面积),采用 ROI算法后反而使 ROI的质量有略微的降低(分别为-0.34 dB和-0.15 dB)。产生这个现象的原因是,当 ROI面积比较大时,由于 ROI部分都采用了更小的量化参数,那么导致 ROI部分的码率增加过多;而背景部分面积较小,难以均衡 ROI带来的额外码率负担,因而引起率失真性能的严重下降,所以在相同带宽下,ROI质量不仅不能够提升,反而有略微的降低。但是本文算法在这种情况下,ROI的质量仍然能够得到一定的提升(0.06 dB和0.21 dB),并且整体 PSNR值的降低也更小。

在计算复杂度方面:与通常的固定比例分区量化方法相比,本文算法和文献[7]算法虽然需要额外计算过渡带宏块的量化系数(见式(3)),但是计算方法简单,复杂度略有增加;与文献[7]算法相比,虽然本文算法增加了对内部宏块的进一步细分,但是 GridA宏块的量化系数和文献[7]中的 ROI内

部宏块计算方法一致,GridB宏块的量化系数采用过渡带边界的量化系数,没有计算量的增加。

ROI算法的主观质量比较如图3所示,其中人的脸部设为 ROI区域,图3(a)是使用 ROI算法的效果,图3(b)是没有使用 ROI算法的效果。可以看到,在同样带宽条件下,使用 ROI算法后,人脸部的细节信息,比如眼睛、鼻子等部分的轮廓以及皮肤的平滑程度,都能得到更好的呈现,主观感受可以得到明显提升。



(a) 128kbps, $P_{\text{ROI}} = 31.49\text{dB}$ (b) 128kbps, $P_{\text{ROI}} = 30.16\text{dB}$

图3 Grid量化算法的主观效果对比(Foreman)

Fig.3 Subjective result of Grid quantization algorithm

4 结论

本文提出基于视觉掩蔽特性的精细 Grid模型的量化模型,该算法增加对 ROI内部的子区域划分,改善传统 ROI量化方法带来的边界效应和整体率失真性能下降。并且针对视频序列中 ROI面积特征,对量化模型的视觉感知进行自适应调节。实验结果表明,本文算法在同等带宽条件下,对于 ROI面积、纹理特征和运动特征不同的各类序列,平均提高 ROI客观质量达1.905 dB,比同类算法高出约0.5 dB。并且 ROI区域的质量有明显的主观效果提升。

参考文献(References)

- [1] Peng W H, Chai T. Adding selective enhancement in scalable video coding for region-of-interest functionality[C]//Proceeding of IEEE International Symposium on Circuits and Systems. Island of Kos, Greece: IEEE Press, 2006: 3089-3093.
- [2] Li Weiping. Overview of fine granularity scalability in MPEG-4 video standard[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2001, 11(3): 301-318.
- [3] Schaar M. Content-based selective enhancement for streaming video[C]//Proceeding of International Conference on Image

Processing. Thessaloniki, Greece: IEEE Press, 2001: 977-980.

- [4] He Yuwen. Block-based fine granularity scalable video coding for content-aware streaming [C]//Proceeding of International Conference on Image Processing. Rochester, New York, USA: IEEE Press, 2002: 22-25.
- [5] Li Ziyin, Zhu Shanan, Liu Lifang. ROI-codec-supported adaptive rate control algorithm [J]. Opto-Electronic Engineering, 2006, 33(4): 105-110. [李子印, 朱善安, 刘丽芳. 支持 ROI 优先编码策略的自适应码率控制算法[J]. 光电工程, 2006, 33(4): 105-110.]
- [6] Chen Gang, Dai Lirong. An improved ROI scalable coding scheme based on FMO [J]. Computer Simulation, 2007, 24(1): 101-106. [陈刚, 戴礼荣. 一种利用 FMO 特性实现 ROI 可伸缩编码的方法[J]. 2007, 24(1): 101-106.]
- [7] Lu Zhongkang, Li Zhengguo. JVT-0056: Perceptual Region-of-interest (ROI) based Scalable Video Coding [R]. Busan, Korea: Joint Video Team, 2005.
- [8] Jayant N. Signal compression: Technology targets and research directions [J]. IEEE Journal of Select Areas Communication, 1992, 10: 314-323.
- [9] Gershon B. An analytical derivation of visual nonlinearity [J]. IEEE Transactions on Biomedical Engineering, 1980, 27(5): 237-242.
- [10] Privitera C M, Stark L W. Algorithm for defining visual Region-of-Interest: comparison with eye fixations [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(9): 970-982.
- [11] Stark L W, Choi Y. Experiment Metaphysics: The scanpath as an epistemological mechanism [C]//Advance in Psychology: Visual Attention and Congition. Amsterdam, Netherland: Elsevier Science, 1996: 3-69.