

# 空间可扩展视频编码中增强层上的 帧内预测算法研究

王 樟<sup>1),2)</sup> 柳 健<sup>2)</sup> 田金文<sup>2)</sup>

<sup>1)</sup>(华中科技大学电子与信息工程系, 武汉 430074) <sup>2)</sup>(华中科技大学图像所, 武汉 430074)

**摘 要** 针对空间可扩展视频编码,提出了一种新的增强层上的帧内预测算法。该算法充分考虑了相邻层间所固有的空域相关性,在无法正确获取本层邻块预测值的情况下,利用基本层上插值放大后的解码,通过重构宏块来对增强层上的编码宏块进行帧内预测。此外,还在预测过程中,通过引入与图像纹理特征相关联的自适应加权因子对 $(W_v, W_h)$ 来进一步提高新算法的编码性能。实验结果表明,与H.264标准中的帧内预测算法相比,新算法在保持原有码率和不增加额外的计算复杂度的前提下,可一定程度地提高亮度分量的PSNR值,因此可以作为空间可扩展视频编码中增强层上的帧内编码方案。

**关键词** H.264 帧内预测 可扩展视频编码 空间增强层

中图法分类号:TP919.81 文献标识码:A 文章编号:1006-8961(2008)04-0612-06

## Spatially Scalable Video Coding with Intra Prediction in Enhancement Layer

WANG Zhang<sup>1),2)</sup>, LIU Jian<sup>2)</sup>, TIAN Jin-wen<sup>2)</sup>

<sup>1)</sup>(Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074)

<sup>2)</sup>(Institute for Pattern Recognition and Artificial Intelligence, Huazhong University of Science and Technology, Wuhan 430074)

**Abstract** In this paper, a new intra prediction algorithm for enhancement layer in spatially scalable video coding is proposed. It takes the spatial correlation between adjacent layers into consideration, and makes use of the magnified decoded and reconstructed base layer block for intra prediction in the lack of neighboring predicting samples. Furthermore, two adaptive weight factors  $(W_v, W_h)$  related to the texture characteristic within an image are introduced to improve the coding efficiency of proposed algorithm. Experimental results show that the PSNR value of luminance component has been increased and both coding bit-rate and computation complexity are preserved. As a result, it can be used as an efficient intra coding scheme for enhancement layer in spatially scalable video coding.

**Keywords** H.264, intra prediction, scalable video coding, spatial enhancement layer

## 1 引 言

近年来多媒体通信技术得到了广泛的应用与发展,特别是流式视频应用在 Internet 和无线网络上正变得越来越流行,如视频会议、可视电话、3G 视频业务<sup>[1]</sup>等等。这就对现在或将来的视频编码标准提出了一个新的要求,即编码后的视频流必须能应付传输

信道带宽的不断变化。为了达到这一目标,有关学者曾经在 H.264 标准的基础上做过一些尝试,比如 Kurceren 等人提出了在多个单层码流之间的自动切换技术<sup>[2]</sup>。尽管如此,上述方法却无法提供一种基于 H.264 标准的更高效的可扩展编码框架结构。

目前,由 MPEG 和 ITU-T 联合组成的 JVT 组织正在进行一项新的可扩展视频编码(scalable video coding, SVC)标准的制定工作<sup>[3]</sup>。SVC 以 H.264 标

基金项目:国家自然科学基金项目(60572048)

收稿日期:2006-06-26;改回日期:2006-11-06

第一作者简介:王 樟(1983 ~ ),男,华中科技大学博士研究生。研究方向为视频编码与网络传输,多媒体网络通信。E-mail: jiaozu-hust@sina.com.cn

准为基础,要求在单一编码的视频流中同时提供时域、空域以及质量域3个方面的可扩展性。其中,在空域可扩展性方面,空间基本层编码方案将完全与H.264标准兼容,而在空间增强层上,则通过采用多种层间预测技术,诸如纹理预测、运动预测以及残差预测等来极大地提高其编码效率。对于层间纹理预测,SVC还引入了一种新的增强层宏块编码模式Intra<sub>BL</sub><sup>[4]</sup>,其原理是利用基本层上的已解码重构宏块来逐个像素点地预测本层宏块,并且只编码两者之间的残差块数据。尽管如此,Intra<sub>BL</sub>模式在应用过程中却有一个严格限制条件,即用来预测的基本层宏块必须是Intra模式编码的。

帧内编码技术作为提高Intra帧编码效率的有效方法,已成为视频或图像压缩领域里最重要的技术手段之一。例如,最早在H.263+和MPEG-4标准中,就采用了基于离散余弦变换(DCT)域的帧内预测算法。而在H.264标准中,帧内预测是基于空域相关性<sup>[5]</sup>来进行的,其预测精度较以往的频域算法有显著提高。但是,所有这些帧内预测算法仅仅使用了本层中的预测数据,这就使得在这些预测数据无法获取的情况下导致帧内预测失效。

鉴于此,本文在空间可扩展视频编码方面,提出了一种新的用于增强层上的帧内预测算法。该算法首先在空间基本层仍采用与H.264标准兼容的编码方式,而对于空间增强层,则充分利用相邻层间所固有的空域相关性和基本层中的对应块数据进行帧内预测。另外,还引入了一对与图像纹理特征相关联的自适应加权因子( $W_v$ ,  $W_h$ ),用来提高该算法的预测精度,以便改善整体编码性能。

## 2 H.264 中的帧内编码方法

### 2.1 帧内编码框架

H.264标准中的帧内编码框架<sup>[6]</sup>如图1所示。该编码框架是当视频序列信号输入到编码器中进行编码时,首先根据帧存储器中的已编码块对每个输入的帧宏块进行帧内预测,预测值是从当前块的左方和上方的已编码邻块里获得;然后,只对宏块原始值和预测值之间的残差进行变换和量化,同时将一路码流经过反量化和反变换之后存入帧存储器,以作为后续宏块帧内预测的参考块;最后,将量化后的残差数据以及预测模式等其他语法元素经过熵编码后输出到码流。

### 2.2 帧内预测方向模式

众所周知,在自然场景图像中,前景和背景通常具有一定的纹理特性,按其方向性可以分为水平、垂

直和倾斜纹理。H.264标准中的帧内预测算法,正是基于像素点在空域中的纹理方向特性以及子块间的相关性进行预测,从而去除了空间冗余。现以亮度分量来进行说明,标准中包含2种预测类型<sup>[7]</sup>,即基于 $16 \times 16$ 宏块的Intra<sub>16 \times 16</sub>类型和基于 $4 \times 4$ 子块的Intra<sub>4 \times 4</sub>类型。

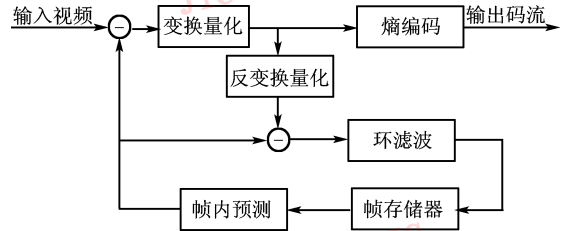


图1 H.264标准中帧内编码框架

Fig.1 Intra coding framework in H.264 standard

其中,Intra<sub>4 \times 4</sub>类型是将一个 $16 \times 16$ 宏块细分成16个 $4 \times 4$ 子块,并分别对每个子块进行预测。为了提高预测精度,共定义了9种子模式(见图2)。其中,除均值预测子模式以外,其余子模式都与图像中可能出现的纹理方向相对应。该Intra<sub>4 \times 4</sub>类型主要适用于包含丰富细节信息的图像区域。

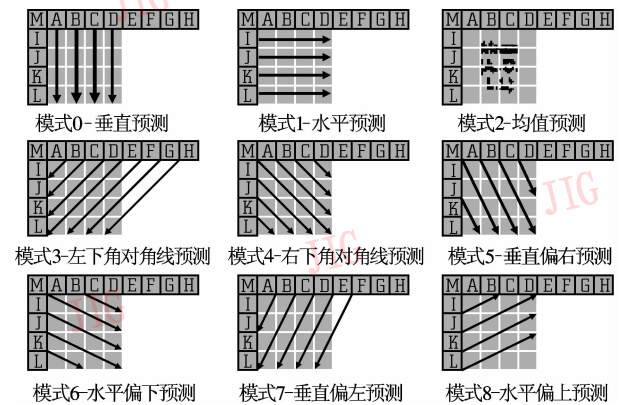


图2 H.264标准中Intra<sub>4 \times 4</sub>预测方向模式

Fig.2 Nine Intra<sub>4 \times 4</sub> prediction modes in H.264 standard

Intra<sub>16 \times 16</sub>类型是直接对整个宏块进行帧内预测,无需分块。由于大尺寸图像块的空间相关性一般要小于小尺寸的情况,所以此类型仅包含垂直、水平、均值和平面4种子模式。该Intra<sub>16 \times 16</sub>类型主要适用于内部较为平坦的图像区域。

## 3 新的层间帧内预测算法

### 3.1 空间可扩展性的编码框架

一个空间可扩展视频编码的结构框图<sup>[8]</sup>如图3

所示。该结构考虑以下 2 个空间层:一个基本层和一个增强层。增强层中的输入视频序列,经过一个 2 维下采样滤波之后就形成了基本层的输入信号。这时可以分别对 2 个空间层单独进行编码。在图 3 中,基本层编码方案是与 H. 264 兼容的基于分块的混合视频编码;而增强层编码方案则进一步采用多种层间预测技术来提高编码效率。这样将产生 2 个码流,其分别对应视频图像的不同空间分辨率。

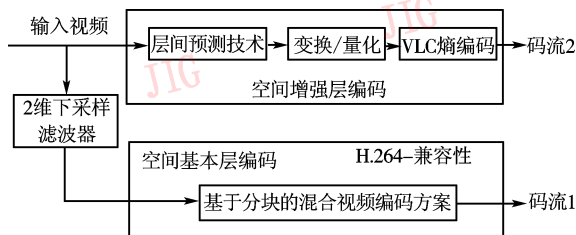


图 3 空间可扩展性编码框架

Fig. 3 Video coding framework with spatial scalability

### 3.2 层间帧内预测原理

由 3.1 节描述可知,基本层输入信号一般是由增强层中的输入视频序列经过 2 维下采样滤波得来的<sup>[9]</sup>。更进一步地,由于在一个 GOP (group of picture) 编码范围以内,当对增强层中的图像帧宏块进行编码的时候,对应的基本层图像帧已经解码重构过了,因此,如果把基本层重构帧插值放大到与增强层同样大小的空间分辨率(见图 4),那么就可以利用它来预测编码增强层上的图像帧。

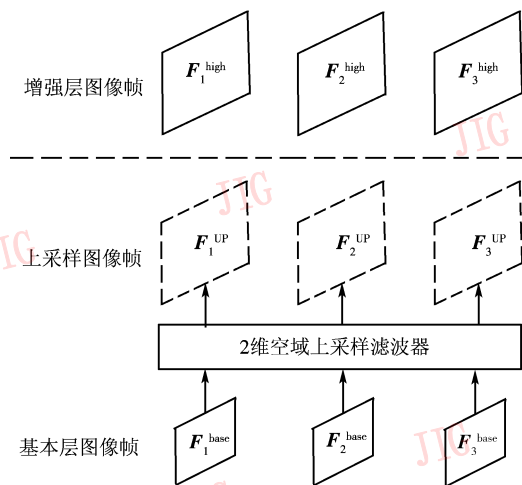


图 4 图像帧的层间空域关系

Fig. 4 Inter layer frame relationship

如图 4 所示,首先在一个 GOP 内,解码重构基本层的第 1 帧图像  $F_1^{\text{base}}$ ,然后通过将一个 2 维的上采样

滤波器作用于  $F_1^{\text{base}}$ ,就可以得到插值放大后的图像  $F_1^{\text{UP}}$ ,并且  $F_1^{\text{UP}}$  和增强层第 1 帧高分辨率图像  $F_1^{\text{high}}$  具有相同的空间分辨率。这样一来,图像  $F_1^{\text{UP}}$  中的宏块像素值就可以用来预测编码图像  $F_1^{\text{high}}$  中的对应块。

由于上采样滤波器的性能优劣直接决定了预测精度,因此选择合适的滤波器就成为本文算法的一个关键前提。此外,由于人眼视觉性质对亮度分量较之色度分量要敏感得多,所以可通过滤波器性能的对比实验,并通过权衡计算复杂度和预测精度两个因素来对图像不同分量的上采样滤波器进行选择:对于亮度分量,可使用 6-tap 滤波器;而对于色度分量,则使用 4-tap 滤波器或双线性滤波器。这 3 种滤波器分列如下:

Y 分量的 6-tap 滤波:  $[1, -5, 20, 20, -5, 1]/32$

Cb/Cr 分量的 4-tap 滤波:  $[-1, 5, 5, -1]/8$

Cb/Cr 分量的双线性滤波:  $[1, 1]/2$

具体来说,本文在对比实验中是以最复杂的 6-tap 滤波器为基准,由于当简化亮度分量时,PSNR 值下降幅度很大,高达 0.2 ~ 0.5 dB,因此亮度分量必须采用 6-tap 滤波器;当简化色度分量时,PSNR 值变化很小,只有 0.000 1dB 数量级大小,这说明色度分量可以采用简化的 4-tap 或双线性滤波器。

基于以上观点,本文提出了一种新的层间帧内预测算法。该算法重点考虑在本层邻块数据无法正确获得的情况下(如首行或首列宏块),基本层上的重构放大宏块可以作为参考的预测数据。为便于说明,首先介绍如下定义:  $MB_b$  (下角 b 代表 base,下同),是经过上采样放大后的基本层重构宏块,而  $MB_e$  (下角 e 代表 enhance,下同)则是增强层图像帧中的当前待编码宏块,两者对应同一图像区域。  $U_e$  和  $L_e$  表示  $MB_e$  的上方和左方邻块中的预测行/列像素值,而  $U_b$ 、 $D_b$ 、 $L_b$  和  $R_b$  则是  $MB_b$  的四周(上、下、左、右)边界,分别对应着首行、末行、首列和末列像素值。  $\hat{V}$  表示垂直方向上的预测数据,而  $\hat{H}$  则表示水平方向上的预测数据。本文引入 2 个与图像纹理特征相关联的加权因子  $W_v$  和  $W_h$  ( $0 < W_v, W_h < 1$ ),分别作为垂直和水平方向上的预测权系数。本算法的核心流程见图 5。

从图 5 可以看出,增强层上的层间帧内预测算法主要针对以下 4 种情况作出相应处理:

(1)  $U_e$  和  $L_e$  都能正确获取——此时,算法同基

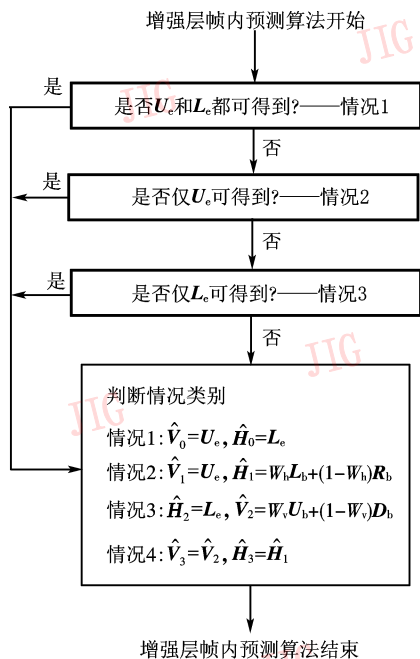


图5 层间帧内预测算法流程图

Fig. 5 Inter layer intra prediction algorithm flow chart

本层一致,即满足  $\hat{V} = U_e, \hat{H} = L_e$ 。

(2) 仅有  $U_e$  能正确获取——此时,预测数据在垂直方向上采用  $U_e$ ,而在水平方向上则采用  $L_b$  和  $R_b$  的加权和,即  $\hat{V} = U_e, \hat{H} = W_h L_b + (1 - W_h) R_b$ 。

(3) 仅有  $L_e$  能正确获取——此时,预测数据在水平方向上采用  $L_e$ ,而在垂直方向上则采用  $U_b$  和  $D_b$  的加权和,即  $\hat{H} = L_e, \hat{V} = W_v U_b + (1 - W_v) D_b$ 。

(4)  $U_e$  和  $L_e$  都无法正确获取——此时,预测数据全部采用基本层上的预测像素值,在水平方向上采用  $L_b$  和  $R_b$  的加权和,而在垂直方向上则采用  $U_b$  和  $D_b$  的加权和,即  $\hat{V} = W_v U_b + (1 - W_v) D_b, \hat{H} = W_h L_b + (1 - W_h) R_b$ 。

### 3.3 自适应权系数的确定

由于  $W_v$  和  $W_h$  分别是垂直和水平方向上的预测权系数,其能够反映图像的纹理特性,所以要求它们的取值必须能随着当前图像纹理的变化而自适应地动态改变。但是反过来看,如果对视频序列中的每帧图像都编码这对权系数,那么无疑会因额外增加相当数量的码率而影响整体的编码性能。因此,本文采用一种称之为基于上下文预测模式判断(context-adaptive mode estimation, CAME)的策略来动态确定  $W_v$  和  $W_h$  的取值。

首先,按照方向性,可划分出水平、垂直和倾斜

这3种主要的图像纹理。对于不同的纹理类型,权系数取值如下:

水平纹理:  $W_h = 0.75$ ;

垂直纹理:  $W_v = 0.75$ ;

倾斜纹理:  $(W_v, W_h) = (0.5, 0.5)$ 。

然后,通过 CAME 策略来预判当前编码图像块的纹理类型,并据此来选择当前帧内预测过程中所要使用的权系数值。CAME 策略的关键步骤描述如下:

(1) 如果上方或左方邻块的帧内编码是垂直模式,则判定当前块是垂直纹理。

(2) 如果上方或左方邻块的帧内编码是水平模式,则判定当前块是水平纹理。

(3) 如果是其他模式,则判定当前块是倾斜纹理。

最后需要说明的是,一般不需要对权系数值进行编码,并传输到解码端,这是因为 CAME 策略关于权系数的确定过程在编码端和解码端是一致的。由此可见,新的层间帧内预测算法只是在编解码前增加了一个纹理预判过程,而并没有额外增加预测运算的计算复杂度,因为它只是扩展了获取预测数据的途径,而并没有改变 H. 264 中现有的方向预测模式。

## 4 实验结果与分析

所有的编码实验都是在当前最新版本的 SVC 标准测试模型<sup>[10]</sup>——JSVM-4.0 软件平台上进行的。为验证不同编码方案的性能,本文将测试 SLA 和 MLA 这2种编码方案。其中,对于 SLA 编码方案,基本层和增强层都采用 H. 264 标准中的原始帧内预测算法;而对于 MLA 编码方案,基本层上的算法同 SLA 编码方案,但是增强层是采用本文提出的层间帧内预测算法。

实验是以 SVC 标准中的 Football 和 Bus 作为测试序列,编码的帧率设为 15 fps。GOP 长度和 Intra 帧的编码周期都设为 1,这样所有的视频帧图像都编码为 Intra 帧。为了更全面地测试在不同码率条件下的编码性能,一共选取了 8 个 QP 值,其范围是 (8, 36) 并以 4 为增量单位。同时,为了加快编码执行速度,取消 MCTF 时域分解过程<sup>[11]</sup>中的更新步骤。由于加权因子 ( $W_v, W_h$ ) 与图像纹理性质相关联,因此对 Football 和 Bus 序列取不同的初始权系数值,分别为 (0.5, 0.5) 和 (0.5, 0.75)。实验平台的输入参数配置如表 1、表 2 所示。

表 1 中 FGS, CABAC 分别为细粒度可扩展性 (fine granular scalability, FGS) 及基于上下文自适应的二进制算法编码 (content-adaptive binary arithmetic coding, CABAC)

表 1 公共参数配置  
Tab. 1 Common configuration

扩展位	FGS 层数	帧率 (fps)	QP 取值
1	0	15	8 ~ 36
自适应 QP	解码环数	帧内模式	熵编码
1	1	1	CABAC

表 2 差异参数配置  
Tab. 2 Different configuration

	基本层	增强层
序列帧格式	QCIF	CIF
H. 264 兼容性	是	否
层间预测机制	无	自适应
闭环级别	L 级	L + H 级

由于 SLA 和 MLA 两种编码方案的基本层所使用的帧内预测算法是相同的,故基本层的编码结果完全一样。据此,可只统计增强层上的实验结果,并给出编码性能的对比。实验结果统计了编码码率和亮度分量的峰值信噪比 (PSNR-Y) 值 (见表 3)。

表 3 不同编码方案码率和 PSNR-Y 值实验数据  
Tab. 3 Bit-rate and PSNR-Y values

测试序列	QP 值	SLA 编码方案		MLA 编码方案	
		码率 (kbps)	PSNR-Y 值 (dB)	码率 (kbps)	PSNR-Y 值 (dB)
Bus	8	9 600.59	52.947	9 604.47	53.109
	12	7 545.04	49.513	7 548.95	49.662
	16	5 859.33	46.055	5 863.42	46.183
	20	4 339.10	42.379	4 342.76	42.515
	24	3 166.23	39.096	3 169.73	39.211
	28	2 228.83	35.890	2 231.18	36.000
	32	1 468.82	32.662	1 470.65	32.754
	36	932.74	29.848	933.07	29.925
Football	8	7 487.51	52.764	7 492.45	52.913
	12	5 629.84	49.672	5 634.13	49.824
	16	4 168.02	46.631	4 172.92	46.776
	20	2 944.19	43.350	2 947.08	43.501
	24	2 076.95	40.388	2 078.37	40.520
	28	1 409.89	37.449	1 411.43	37.565
	32	890.06	34.524	893.64	34.615
	36	544.14	31.982	544.56	32.076

从表 3 的统计结果中可以看出,对于测试序列 Bus 和 Football,其编码输出的平均码率波动范围都

很小,几乎可以忽略不计。同时,PSNR-Y 值在每个 QP 取值点上都有提高,并且满足以下规律:随着 QP 取值的增加,PSNR-Y 值提高的幅度也不断增大。由此可见,MLA 编码方案的总体性能要优于 SLA 编码方案,这也说明了本文提出的层间帧内预测算法能提高增强层上的编码效率,并可改进编码性能。关于码率波动范围和 PSNR-Y 值的升高幅度,具体的统计分析数据如表 4 所示。

表 4 Bus 和 Football 序列码率和 PSNR-Y 值统计分析值  
Tab. 4 Analysis of the bit-rate and PSNR-Y

测试序列	码率波动 (kbps)	PSNR-Y 值	PSNR-Y 值	PSNR-Y 值
		最小增量 (dB)	最大增量 (dB)	平均增量 (dB)
Bus	[0, 4]	0.077	0.162	0.12
Football	[0, 5]	0.091	0.152	0.13

图 6 和图 7 分别给出了 Football 和 Bus 序列的码率/PSNR-Y 值相关曲线图。

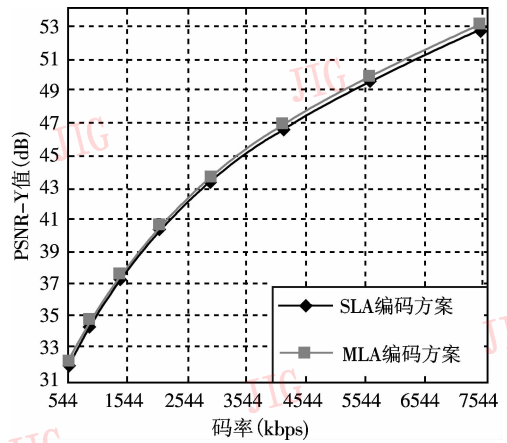


图 6 Football 序列的码率与 PSNR-Y 值相关曲线图

Fig. 6 Bit-rate and PSNR-Y curve for Football

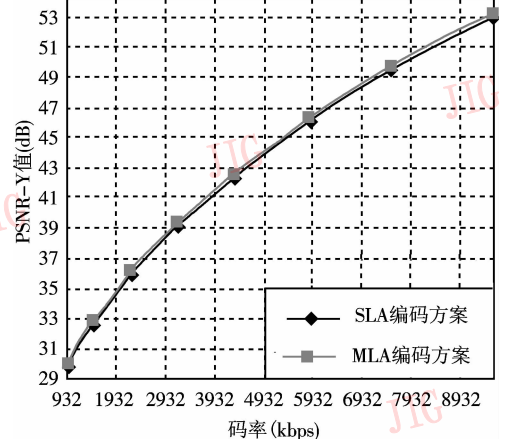


图 7 Bus 序列的码率与 PSNR-Y 值相关曲线图

Fig. 7 Bit-rate and PSNR-Y curve for Bus

## 5 结 论

本文针对空间可扩展视频编码,提出了增强层上的一种新的帧内预测算法。该新算法对于空间基本层,仍采用与 H. 264 兼容的编码方式,而对空间增强层则充分利用相邻层间的空域相关信息,通过引入与图像纹理特征相关联的加权因子( $W_v$ ,  $W_h$ )来提高其编码性能。实验结果表明,与 H. 264 中的原始帧内预测算法相比,新算法在保持原有码率和计算复杂度的前提下,还能够一定程度地提高亮度分量的 PSNR 值,从而改进了空间增强层上的编码性能。例如,对于测试序列 Football 和 Bus 编码后的图像,其 PSNR-Y 值的平均增益分别达到了 0.13dB 和 0.12dB。因此,新算法可以应用到空间可扩展视频编码的增强层帧内编码中去。

另一方面,如果自适应权系数  $W_v$  和  $W_h$  的取值与图像纹理特性越匹配,那么帧内预测精度也会越高。本文算法由于只考虑了 3 种主要纹理类型,并且分别赋予不同的权系数值,因此其只能在一定程度上提高预测精度。如何使权系数能更准确地匹配视频序列中不断变化着的图像纹理,将成为下一步研究工作的重点。诸如细分纹理类型、增强 CAME 策略精度、更有效地确定权系数初值等等,都是值得深入研究的问题。

### 参考文献 (References)

1 Greenbaum G. Remarks on the H. 26L Project: Steaming video

- requirements for Next Generation Video Compression Standards [A]. In: JVT-Q15-G-11 of ITU-T SG16/Q. 15 Meeting [C], Paris, France 1999.
- 2 Kurceren R, karczewicz M. Improved SP-frame encoding [A]. In: JVT-VCEG-M73 of ITU-T SG16/Q. 15 Meeting [C], Austin, TX, USA, 2001.
- 3 Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG. Scalable video coding-joint draft 4 [A]. In: JVT-Q201 of 17th Meeting [C], Nice, France, 2005.
- 4 Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG. Intra-BL prediction considering phase shift [A]. In: JVT-0023 of 15th Meeting [C], Busan, Korea, 2005.
- 5 Halbach T. Performance comparison: H.26L intra coding vs. JPEG2000 [A]. In: JVT 4th Meeting [C], Klagenfurt, Austria, 2002.
- 6 Wiegand T, Sullivan G, Luthra A. Overview of the H.264/AVC video coding standard [J]. IEEE Transactions on Circuits, System and Video for Video Technology, 2003, 13(7): 560 ~ 576.
- 7 H. 264/MPEG-4 Part 10 White Paper: Prediction of Intra Macroblocks [EB/OL]. <http://www.vcodex.com/h264.html>.
- 8 Ohm J. Advances in scalable video coding [J]. Proceedings of the IEEE, 2005, 93(2): 42 ~ 56.
- 9 Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG. Adaptive up-sampling for spatially scalable coding [A]. In: JVT-0010 of 15th Meeting [C], Busan, Korea, 2005.
- 10 Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG. Joint scalable video model - JSVM-4 [A]. In: JVT-Q202 of 17th Meeting [C], Nice, France, 2005.
- 11 Golwelkar A, Woods J. Scalable video compression using longer motion compensated temporal filters [A]. In: Proceedings of SPIE Visual Communication Image Processing (VCIP03) [C], Lugano, Switzerland, 2003: 1406 ~ 1417.