

基于感知的右视点质量可分级编码算法

吴爱红, 郁梅, 邵枫, 彭宗举, 蒋刚毅

(宁波大学 信息科学与工程学院, 浙江 宁波 315211)

摘要: 为了减少立体视频传输如此庞大的数据量, 根据人眼的立体掩蔽效应, 提出了基于感知的右视点质量可分级编码算法. 该算法将宏块分为 3 类: 平坦块、纹理块和边缘块, 通过不同的量化步长使得右视点图像块的质量可分级. 实验证明根据左视点的图像质量, 右视点在人眼感知范围内码率减少 25%~37%. 该算法能在人眼主观察觉范围内很好地提高码率, 进一步消除人眼立体视觉冗余.

关键词: 立体视频; 掩蔽效应; 可分级; 感知

中图分类号: TN919.81

文献标识码: A

文章编号: 1001-5132 (2011) 01-0121-07

随着计算机、通信及网络技术的发展, 增强视觉现实感和逼真感的三维视频技术成为当前的一个热点, 已经吸引了越来越多的国内外学者, 并在各个领域得到了广泛应用, 如三维电视(3DTV)、远程教育、远程工业控制、三维视频会议系统、虚拟现实系统、远程医疗、遥控机器人、自动导航、消费电子等诸多方面^[1-3].

随着第 3 代移动通信的发展, 手机等移动终端的立体显示成为一种可能, 也成为新的立体视频编码研究热点. 手机将从单纯的通讯工具逐步变成具备视频电话、移动电视、视频游戏和视频播放等功能的个人娱乐终端. 但庞大的立体视频数据量对于有限的无线带宽、有限的存储能力和实时编解码都成为一种挑战, 于是一种有效的立体视频数据压缩算法是解决这一挑战的关键.

在解决立体视频压缩效率问题上, 很多学者都有各自的改进. 立体视频编码结构右最初的 simulcast 编码、基于视差估计的编码、联合编码、到 IBP 结构编码和 HBP 结构编码, 这些编码方法在压缩效率上都有很大的提高. 早期的心理学研究表明, 立体视觉中存在掩蔽效应, 既构成立体图像的 2 个视点, 质量好的视点图像质量对立体图像的整体质量贡献较大^[4]. 根据这一视觉特性,

Fehn C 等提出了非对称编码^[5], 即将一个视点按原分辨率编码, 另一分辨率采用下采样进行低分辨率编码. 基于这一方案, 文献[6]提出了基于残差下采样的立体视频编码算法, 即对右视点的预测残差进行下采样, 减少了解码右通道解码时需要将参考帧下采样的复杂度, 同时基于块的下采样能得到更高的压缩比. 文献[7]提出了视频加深度的编码方案, 编码端只传输左视点的视频序列和右视点相对右视点的深度图序列, 数据量较小的深度图序列为在带宽有限的编码传输中减少了压力. 由于非对称编码中左右通道的数据量是不一样的, 为了充分利用带宽, 文献[8]通过实验证明混合分辨率编码中的低分辨率视点最佳分配大约为总码率的 30%到 35%.

混合分辨率编码主要是利用视点间的空间相关性. 视频加深度编码方法利用了立体视频的深度信息, 该方法虽能有效地降低码率, 但在绘制端绘制出右视点提高了解码的复杂度, 这对手机等移动终端来说又是另一个瓶颈. 人眼的视觉特性不仅在视点间有一定的掩蔽效应, 在时间和空间及不同序列内容都有一定的掩蔽效应. 杨小康等^[9]在研究最小可察觉变化步长(Just Noticeable Difference, JND)建模时, 提出 JND 不仅与时空对比敏感度有

关,还与图像序列的适应亮度及图像的内容有关。

基于人眼感知特性,笔者提出了基于感知的右视点质量可分级编码算法,该算法将右视点图像进行块分类,分别为平坦块、纹理块和边缘块,对不同类的块进行不同量化步长量化,实现右视点块的质量可分级。

1 基于移动终端的立体视频编码结构

视频编码计算量非常大,常常需要借助于专用设备或硬件完成,尤其对于交互式视频通信或实时视频应用,视频的编码和解码必须在给定的时间间隔内完成。这就要求视频编码和解码算法的计算复杂度不能够超出处理器的处理能力,有效降低视频编码计算量是这类应用首要解决的问题^[10]。H.264 编码性能的提高是以编码复杂度的提高为代价的。和 MPEG-4 的编码器相比, H.264 的编码复杂度要高出 10 倍左右^[11]。而立体视频需要对左右 2 个通道进行编码,其复杂度至少是单通道编码复杂度的 2 倍。目前,国际上一些著名手机和相机数码厂商已经开始搭载立体拍摄功能,通常移动终端的计算能力有限,对于立体视频的实时应用提出了很大的挑战。这就要求在视频编码和解码算法的计算复杂度不能够超出处理器的处理能力前提下,尽可能地降低立体视频编码时间是非常有必要的。

H.264 是一种高效的单通道视频编码标准,但对于立体视频编码来说,并不是对左右通道分别采用高效的单通道视频编码,就可以达到很好的压缩效果。因为立体视频编码除了要考虑每个通道帧内图像的空间冗余度和帧间图像之间的时间冗余度外,还要考虑左右通道图像之间的空间冗余度。对于后者,通常可以利用视差估计补偿的方法来去除其冗余度。同时,立体视频的应用还不够广泛,传统的单通道视频仍然是市场的主流。单通道视频向立体视频过渡,其间必然要经历两者共存的局面^[12]。为了使传统的单通道视频显示设备能正常地播放立体视频节目,必须考虑视点可分级的问题,即立体视频流既可完全解码 2 个通道的数据,也可只解码 1 个通道的数据。根据不同客户的实际需要,编码需考虑立体编码同时向下兼容单视点解码。

目前作为 mobile-3DTV 领军人物的 Tempere 大学对立体视频编码主要采用图 1 的结构,每个编码帧有 3 个参考帧,其中右视点中一参考帧参考左视点编码帧。之所以使用 P 帧而没有使用率失真性能更好的 B 帧,是考虑到手机平台相对硬件编解码性能较弱,使用 P 帧可以降低其解码复杂度。

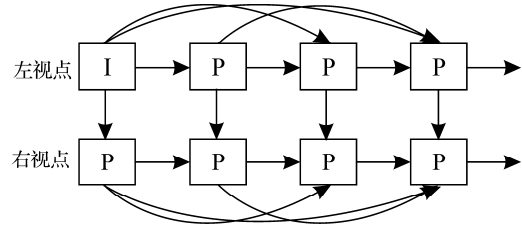


图 1 立体视频编码框架

考虑到移动终端的内存和计算能力有限,从实时的角度考虑,参考左右视点一共参考 6 个参考帧对快速编码是一严峻挑战。笔者对其方案进行了改进,它基本结构为:右通道的每个像素块需要进行 2 种方式的预测,一种是基于右通道先前帧图像的 MCP 方式;另一种是基于左通道图像的 DCP 方式,然后从中选择预测误差较小的一种,如图 2 所示^[13]。

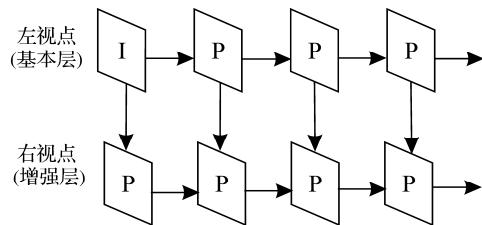


图 2 基于移动终端的立体编码框架

该编码方案采用 IPPP 的基本档次编码,同时考虑了时间和视点间的相关性,考虑到移动终端的存储能力,所以设置的参考帧个数较少,对于左视点只有 1 帧时间参考帧,而右视点分别有 1 帧时间参考帧和 1 帧空间参考帧。

2 基于感知的右视点质量可分级编码算法

不同的图像压缩会产生不同的失真,不同的失真对图像的影响也是不一样的。结合各失真类型对图像细节的影响,模糊失真会使图像中的边缘强度变弱,平坦区域增多; JPEG 压缩失真所造成的块效应会造成伪轮廓;高斯白噪声则会使得纹理区域增多。Ran 在对图像属性的研究结果中将图像分为 3 个区域:边缘、平坦、纹理区域,指

出针对不同的区域在评价中应给出不同的权重^[14]. 结合这些主观评价依据, 笔者提出基于感知的右视点质量可分级编码算法, 在人眼的 JND 范围内, 对不同复杂度的图像给予不同的量化步长, 实现右视点质量可分级编码.

2.1 右视点图像块的分类

对图像的分类, 不同文献方法不一, 有的在空域分, 有的在频域分, 主要有 2 种, 一种是根据图像的梯度来分, 一种是根据频域的频率分. 笔者根据块亮度的复杂度将块分为平坦块、纹理块和边缘块.

(1) 块类型的粗判断. 根据块的亮度平均绝对差值来判断块的类别, 先根据这个差值粗略地判断出 2 个阈值, 将判断结果与人眼感知结合, 最终给出最佳判决条件.

27	7	8	12	11	12	10	12	9	7	13	13	15	15	15	6	5	4	29	21
13	8	6	9	13	14	7	6	7	5	6	12	24	12	10	6	5	3	23	16
5	9	5	7	7	15	8	6	4	5	7	11	23	13	8	7	5	4	18	15
7	16	5	7	22	14	8	8	5	5	9	12	20	9	7	7	6	6	16	17
8	5	5	8	26	12	4	10	2	4	10	12	18	7	6	7	6	6	14	14
7	6	5	9	29	12	4	11	3	3	9	15	14	12	2	3	5	4	10	11
5	7	5	9	23	12	6	11	3	12	17	9	11	2	2	5	5	3	6	10
3	3	5	8	17	13	6	7	7	0	0	2	51	4	4	4	5	3	5	11
2	5	4	10	14	12	5	9	5	60	23	27	24	50	5	3	5	1	4	10
1	5	6	8	14	11	6	6	40	10	27	15	7	27	3	3	6	4	4	11
1	3	6	10	12	7	4	6	65	0	22	3	19	33	5	4	4	2	4	13
1	1	7	11	13	6	4	8	44	78	10	16	9	32	2	5	2	1	3	15
1	6	8	12	13	9	5	6	20	16	26	23	53	30	5	4	6	1	3	11
1	4	11	13	9	6	2	8	19	16	11	16	59	28	7	6	10	8	11	29
1	6	12	13	8	3	3	11	28	15	29	18	2	10	18	5	4	9	8	

图 3 部分 akko 块平均绝对差值



图 4 对应绝对差值的部分图像

从图 4 对应的数据(图 3)可以看出, 在不平坦或者说深度不一致的地方平均绝对差值比较大, 根据这个特点可粗略的取出 2 个阈值, 根据块亮度内容的复杂度进行块分类. 初步将阈值定义为 20 和 10, 若块的平均绝对差值大于 20 判断为纹理块, 在 10 和 20 之间判为边缘块, 小于 10 判为平坦块. 实验结果见图 5 和图 6, 其中, 白色框标记的表示纹理块, 黑色框标记的表示边缘块, 没有标记的表示平坦区域. 在 akko 标记结果中可看出, 有过多的边缘块, 在背景上, 出现大量的边缘块, 这样的标记结果对编码带来一定的浪费或者使得原本简

单的编码变得更加复杂. 在 puppy 标记结果中可看出, 图像中判为边缘块的部分太多, 甚至一些人眼认为是纹理的地方, 也标记成了边缘块, 明显存在一些不合理性.



图 5 akko 初步标记结果

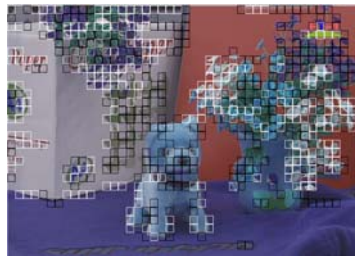


图 6 puppy 初步标记结果

(2) 块类别精细判断. 在文献[9]中通过实验证明背景平均亮度大于 127 时, JND 值与块亮度成线性关系. 根据这一人眼视觉特性, 将块的亮度分为 2 段, 且 2 段的阈值不同, 对于块的平均亮度大于 127 的块, 阈值分别为 20 和 10; 对块的平均亮度小于 127 的块, 阈值分别为 15 和 10. 标记结果见图 7 和图 8.



图 7 akko 精细标记结果

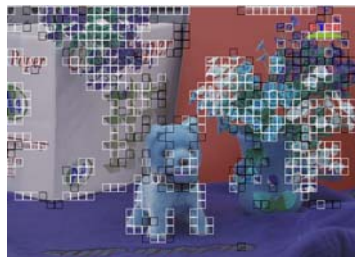


图 8 puppy 精细标记结果

与初步标记结果相比, 有花的地方边缘块减

少了很多, 相应的转换成了平坦块和纹理块. 但在 akko 图像中原本是边缘块被误判为平坦块, 如图像中屏风的纹理. 在 puppy 图像中, 标记结果还是比初步标记准确, 在图像中有花的区域基本上都标记成了纹理块, 在地毯上基本是平坦块, 但地毯上的文字误判为平坦块, 但对编码来说, 这些误判不会影响编码中的质量.

2.2 基于感知的右视点质量可分级编码算法

在块分类基础上, 对右视点进行质量可分级编码, 其算法流程如图 9 所示.

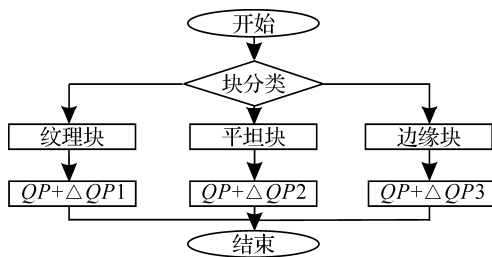


图 9 右视点质量可分级编码流程图

2.2.1 ΔQPX 的选取

根据人眼视觉系统的掩蔽效应, 人眼对纹理区域的掩蔽比较多, 对平坦区域的失真较容易感知, 但人眼对边缘区域的失真更容易察觉, 根据这一特性, 粗略的 ΔQP1 取 2, ΔQP2 取 1, ΔQP3 取 0. 实验条件分别为: 实验平台选用的是 JM8.6 校验模型修改的立体编码平台, 该平台是 JM8.6 的视点扩展, 同时又加入了视点间的参考; 实验编码结构选用的是较简单的 IPPP 结构, 主要是考虑编码结构的复杂度; 编码帧数采用的是 60 帧; 考虑到解码复杂度和存储能力, GOP 长度选用 12 帧; 参考帧数设定为 2 帧; 测试 QP 选用的是 22、27、32 和 37; 测试序列选用的是 akko 和 puppy 2 组.

2.2.2 实验结果及分析

表 1 和表 2 分别是左右通道相同 QP 下的编码结果, 从中可以看出, akko 序列右视点的质量要比左视点的质量好, 而 puppy 的右视点质量比左视点的质量要差. 主要是因为 akko 序列的拍摄相机间距 5 cm, puppy 相机间距为 6.5 cm, 故 akko 序列右视点视点间的相关性比较强, 而 puppy 的时间相关性较强, 故右视点采用时间和视点间参考时, akko 的右视点质量要好于左视点, 而 puppy 左视点的质量要好于右视点的质量.

从主观结果看, 原始 QP 编码和质量分级编码

效果基本一致, 如图 10 所示. 质量可分级主要思想是对人眼不敏感的纹理区域采用大的量化(在原始 QP 上加 2), 对于人眼感兴趣的边缘区域保护起来(原 QP 不变), 在人眼比较敏感的平坦区域采用中等量化(在原始 QP 上加 1), 使得编码结果主观质量保持基本不变的基础上减少码率.

表 1 akko 原始 QP 编码结果

原始 akko	V _l		V _r		ΔPSNR
	QP	PSNR 码率	PSNR 码率		
22	43.20	2916.25	43.30	2152.51	-0.10
27	40.20	1391.53	40.30	1010.39	-0.10
32	36.97	764.85	37.02	526.55	-0.05
37	33.94	443.62	33.88	281.06	0.06

表 2 puppy 原始 QP 编码结果

原始 puppy	V _l		V _r		ΔPSNR
	QP	PSNR 码率	PSNR 码率		
22	41.36	2230.42	40.72	2531.51	0.56
27	38.58	930.51	37.55	705.94	1.03
32	35.81	542.61	34.61	306.77	1.20
37	32.98	330.85	31.70	149.49	1.28

从表 3 中看出, PSNR 下降在 0.16 dB 到 0.38 dB, 加上原始 QP 右视点质量比左视点质量要好, 即右视点可分级编码后右视点质量的下降完全在人眼不可察觉的范围内. 虽然 puppy 下降的质量要多(表 4), 但基本上也在人眼立体的 JND (相差 2 dB) 内. 由于在判断块的类别时纹理区域和边缘区域



akko 原始 QP(32)



akko 原始 QP(32)+ΔQP



图 10 初步右视点初步质量可分级编码主观图像

较少, 平滑区域较多, 且 ΔQPX 的取值较小, 故码率不能在原始的基础上有很好的减少. akko 减少的码率在 25%到 29%, puppy 减少的码率在 12%到 20%. 对于 puppy 来说由于平坦块较多, 标记为边缘的块较少, 被保护的块也很少, 故导致质量下降较严重, 根据人眼视觉系统特性, 质量好的图像 JND 值较大, 如果用固定的 ΔQPX 对不同的块进行质量可分级不能很好地利用这个主观感知来提高压缩比.

表 3 akko 初步选取 ΔQPX 的编码结果

akko	原始 V_r		V_r		$\Delta PSNR$ /dB	减少码 率/%
	QP	PSNR	码率	PSNR		
22	43.30	2152.51	42.92	1565.19	0.38	27.29
27	40.30	1010.39	40.02	714.79	0.28	29.56
32	37.02	526.55	36.86	382.05	0.16	27.44
37	33.88	281.06	33.73	210.95	0.16	24.95

表 4 puppy 初步选取 ΔQPX 的编码结果

puppy	原始 V_r		V_r		$\Delta PSNR$ /dB	减少码 率/%
	QP	PSNR	码率	PSNR		
22	40.72	2531.51	40.10	2003.84	0.62	20.84
27	37.55	705.94	37.03	614.86	0.52	12.90
32	34.61	306.77	34.05	264.74	0.56	13.70
37	31.70	149.49	31.23	126.87	0.47	15.13

2.2.3 ΔQPX 的精细

在 H.264 中, 量化步长共 52 个值, 详见表 5, 其中, QP 是量化参数, 是量化步长的序号. 当 QP 取最小值 0 时, 代表最精细的量化; 当 QP 取最大

值 51 时, 代表最粗燥的量化. QP 每增加 6, Q_{step} 增加 1 倍^[15].

表 5 H.264 编码器量化表

QP	Q_{step}	QP	Q_{step}	QP	Q_{step}	QP	Q_{step}
0	0.625	1	0.6875	2	0.8125	3	0.875
4	1	5	1.125	6	1.25	7	1.375
8	1.625	9	1.75	10	2	11	2.25
12	2.5	13	2.75	14	3.25	15	3.5
16	4	17	4.5	18	5	19	5.5
20	6.5	21	7	22	8	23	9
24	10	25	11	26	13	27	14
28	16	29	18	30	20	31	22
32	26	33	28	34	32	35	35
36	40	37	44	38	52	39	56
40	64	41	72	42	80	43	88
44	104	45	112	46	128	47	144
48	160	49	176	50	208	51	224

根据人眼的视觉特性, 在质量好的立体图像中, 人眼的 JND 值比质量差的图像要大, 提出 ΔQPX 自适应选取, 即根据左视点的基本 QP 值来选取不同的 ΔQPX 量化步长有 52 个值, 且 QP 每增加 6, Q_{step} 增加 1 倍. 根据这一特点得到自适应的 ΔQPX , 具体算法如下:

- (1) 将左视点的基本 QP 除以 6 向下取整得到一个整数 QP_w ;
- (2) 若是纹理块, 且 $\text{floor}(QP_{max}/6) - QP_w$ 不为 0, 则 $\Delta QP1 = \text{floor}(QP_{max}/6) - QP_w$;
- (3) 若是平坦块, 且 $\text{floor}(QP_{max}/6) - QP_w - 2$ 不为 0, 则 $\Delta QP2 = \text{floor}(QP_{max}/6) - QP_w - 2$;
- (4) 若是边缘块, 且 $\text{floor}(QP_{max}/6) - QP_w - 4$ 不为 0, 则 $\Delta QP3 = \text{floor}(QP_{max}/6) - QP_w - 4$.

该算法目的是在 JND 值大的情况下用大的 ΔQPX 进行质量分级, JND 小的情况下, 由于左图像的质量本身就很差, 如果通过右视点的下降质量, 可能导致整体质量下降严重, 甚至导致人眼视觉的烦感, 故根据左视点的质量来确定 ΔQPX 的大概范围.

3 实验结果及分析

图 11 是精细 ΔQPX 后编码重建的主观图像, 从图像看, 增加 ΔQPX 和未增加的主观质量基本

一致, 没有因为右视点的质量下降而导致整体主观质量的下降. 相比初步 ΔQPX 的编码结果, 精细的 ΔQPX 编码结果在小 QP 的情况下右视点的客观质量下降更多, 码率也相应减少更多. 在大 QP 下, 为了保护整体的图像质量, ΔQPX 会自适应的下调, 从而导致客观质量比初步 ΔQPX 编码结果质量下降的更少, 码率相应的减少要少. 精细 ΔQPX 后的编码结果见表 6 和表 7. 在小 QP 时, 初步 ΔQPX 的编码结果中码率只提高 20.84% 和 12.90%, 而在精细 ΔQPX 的编码中, 码率能够减少 36.70% 和 20.03%.



akko 原始 $QP(32)$



akko 原始 $QP(32)+\Delta QP$



puppy 原始 $QP(37)$



puppy 原始 $QP(37)+\Delta QP$

图 11 精细右视点质量可分级编码主观图像

由于增大 QP 能使原始数据丢失更多, 且由于 H.264/AVC 标准采用的块编码, 各宏块采用不同的 QP 会出现块效应, 所以导致编码结果中 PSNR 下降很大. 又因为立体视频中, 只对右视点利用 QP 使得质量可分级, 左视点编码 QP 不变, 所以立体的主观质量基本保持不变.

表 6 akko 精细选取 ΔQPX 的编码结果

akko	原始 V_r		V_r		$\Delta PSNR/$ dB	减少码 率/%
	QP	PSNR	PSNR	码率		
22	43.30	2 152.51	42.56	1 374.27	0.74	36.16
27	40.30	1 010.39	39.94	709.08	0.36	29.82
32	37.02	526.55	36.82	381.74	0.20	27.50
37	33.88	281.06	33.77	216.93	0.11	22.82

表 7 puppy 精细选取 ΔQPX 的编码结果

puppy	原始 V_r		V_r		$\Delta PSNR/$ dB	减少码 率/%
	QP	PSNR	PSNR	码率		
22	40.72	2531.51	39.11	1602.56	1.61	36.70
27	37.55	705.94	36.58	564.54	0.97	20.03
32	34.61	306.77	33.87	253.18	0.74	17.47
37	31.70	149.49	31.23	129.16	0.47	13.60

4 结论

根据图像块的纹理复杂程度对序列进行块的分类型, 通过不同的阈值将图像块分为平坦块、边缘块和纹理块. 通过标记结果可以看出, 图像块的分类型基本与人眼的注意力相符合, 但由于阈值的固定不变性, 使得部分区域的标记存在不合理性. 同时, 根据人眼对立体视频的掩蔽效应, 即左右视点质量相差在一定范围内人眼感觉不出左右视点的质量差异, 提出了右视点质量可分级编码. 先后采用了 2 种不同增加量化步长的方式使得右视点实现质量可分级编码. 实验结果表明, 2 种方法都能在主观质量不变的情况下相应地减少码率, 但精细自适应的量化步长的选取更有利于利用人眼的 JND, 使得在主观质量不变的基础上压缩比更高. 实验中压缩比并不是很高, 一部分是因为前部分块分类准确度的影响, 另一部分是没有充分利用到人眼的 JND 值, 使得客观质量下降码率却不能很好地减少.

由于只根据图像纹理的复杂度自定义阈值, 没有过多的统计分析实验得出块分类的阈值存在

一定的局限性, 使得块分类部分存在不合理性. 同时人眼对量化失真比较敏感, 故通过改变 QP 来提高码率存在一定的局限性. 下一步工作可以根据定量的统计分析得出一个自适应的阈值, 使得块在空域上的分类算法得到改进. 立体视频中人眼 JND 值的大小与图像本身的内容以及初始的质量有关, 可进一步根据人眼的 JND 模型来确定具体的 JND 值, 充分利用这个 JND 值更进一步在主观质量基本不变的情况下提高压缩比. 同时根据人眼不同失真的敏感程度, 改变失真类型(如采用下采样而产生的人眼不太敏感的模糊失真)来对右视点进行可分级编码.

参考文献:

- [1] Xu L Q, Loeffler A, Sheppard P J, et al. True-view video-conferencing system through 3-D impression of telepresence [J]. BT Technology Journal, 1999, 17(1):59-68.
- [2] Izquierdo E. Stereo matching for enhanced telepresence in three-dimensional video communications[J]. IEEE Trans on CSVT, 1997, 7(4):629-643.
- [3] Wang Rushang, Wang Yao. Multiview video sequence analysis, compression and virtual viewpoint synthesis[J]. IEEE Trans on CSVT, 2000, 10(3):397-410.
- [4] Julesz B. Foundations of cyclopean perception[M]. Chicago: The University of Chicago Press, 1971:54-56.
- [5] Fehn C, Kauff P, Cho S, et al. Asymmetric coding of stereoscopic video for transmission over T-DMB[C]//Proceedings of 3DTV Conference, Kos, 2007:1-4.
- [6] 王知嘉, 何文革. 基于残差下采样的立体视频编码算法[J]. 电力系统通信, 2009, 30(4):68-71.
- [7] Merkle Y, Wang K, Müller A, et al. Video plus depth compression for mobile 3D services[C]//Proceeding of the 2009 IEEE 3DTV Conference, Potsdam, 2009:1-4.
- [8] Brust H, Smolic A, Müller K, et al. Mixed resolution coding of stereoscopic video for mobile devices[C]//Proceeding of the 2009 IEEE 3DTV Conference, Potsdam, 2009:1-4.
- [9] Yang X K, Ling W S, Lu Z K, et al. Just noticeable distortion model and its applications in video coding[J]. Signal Processing: Image Communication, 2005, 20(7):662-680.
- [10] 叶斌. 基于 H.264/AVC 的先进立体视频编码研究[D]. 宁波: 宁波大学, 2008.
- [11] Sullivan G, Wiegand T. Rate-distortion optimization for video compression[J]. IEEE Signal Processing Magazine, 1998(11):74-90.
- [12] Cho S, Hur N, Kim J, et al. Carriage of 3D Audio-visual services by T-DMB[C]//Proceeding of IEEE International Conference on Multimedia and Expo, Toronto, Ont, 2006:2165-2168.
- [13] 杨海龙. 基于移动终端的立体视频编解码研究[D]. 宁波: 宁波大学, 2010.
- [14] Ran X. A perceptually motivated three-component image model-part I: Description of the model[J]. IEEE Transaction on Image Processing. 1995, 4(4):401-415.
- [15] 毕厚杰. 新一代视频压缩编码标准——H.264/AVC[M]. 北京: 人民邮电出版社, 2005:115-117.

A Perception Based Algorithm of Right View Quality Scalable Video Coding

WU Ai-hong, YU Mei, SHAO Feng, PENG Zong-ju, JIANG Gang-yi

(Faculty of Information Science and Technology, Ningbo University, Ningbo 315211, China)

Abstract: Based on the masking effect of human eyes, a quality scalable video coding algorithm is proposed for right view aiming at reducing large amount of data in stereo video transmission. The algorithm is derived from the perception of human visual system, and the macro blocks are used to classify flat blocks, edge blocks and texture. These blocks of different types have the different addition QP which can be inferred from that of the left view. As a result, the right view features in scalable quality that can be identified from one area to another. The result shows that according to the quality of left view images, the bit rate for right view can be reduced by about 25%-37%. The proposed algorithm can well increase bit rate within the human subjective perception range, and further eliminate eye stereo vision redundancy.

Key words: stereo video; masking effect; scalable; perception

(责任编辑 史小丽)