[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
H04N 7/34 (2006.01)

[12] 发明专利申请公开说明书

[21] 申请号 200510030401.2

[43] 公开日 2006 年 11 月 8 日

[11] 公开号 CN 1859576A

[22] 申请日 2005.10.11

[21] 申请号 200510030401.2

[71] 申请人 华为技术有限公司

地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

[72] 发明人 熊联欢 谭毅华 王 樟 刘阿军

[74] 专利代理机构 上海明成云知识产权代理有限公司

代理人 竺 云

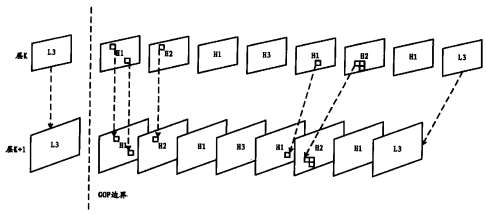
权利要求书 2 页 说明书 8 页 附图 4 页

[54] 发明名称

对空间分层编码视频图象的上采样方法及其系统

[57] 摘要

本发明涉及视频图象压缩技术，公开了一种对空间分层编码视频图象的上采样方法及其系统，使得在上采样过程中，可以减少计算复杂度，而编码性能基本不变。 本发明中，利用人眼对色度分量的敏感度远远低于亮度分量的原理，在 I_BL 层间预测或残差图象层间预测的上采样过程中，对色度分量采用较亮量分量更简单的滤波器，从而有效减少了计算复杂度，而编码性能基本不变。 在残差图象上采样过程中，至少在亮度分量上使用较为复杂的 [－155 －1] /8 滤波器，从而提高编码性能。



1. 一种对空间分层编码视频图象的上采样方法，其中，使用第一滤波器对图层的亮度分量进行上采样，使用第二滤波器对图层的色度分量进行上采样，其特征在于，第二滤波器比第一滤波器简单。

2. 根据权利要求 1 所述的对空间分层编码视频图象的上采样方法，其特征在于，所述上采样应用于“I_BL”层间预测或残差图象层间预测。

3. 根据权利要求 2 所述的对空间分层编码视频图象的上采样方法，其特征在于，“I_BL”层间预测上采样时，所述第一滤波器是[1 -5 20 20 -5 1]/32，所述第二滤波器是[-1 5 5 -1]/8。

4. 根据权利要求 2 所述的对空间分层编码视频图象的上采样方法，其特征在于，残差图象层间预测上采样时，所述第一滤波器是[-1 5 5 -1]/8，所述第二滤波器是[1 1]/2。

5. 一种对空间分层编码视频图象的上采样系统，其中至少包含用于对图层的亮度分量进行上采样的第一滤波器，和对图层的色度分量进行上采样的第二滤波器，其特征在于，第二滤波器比第一滤波器简单。

6. 根据权利要求 5 所述的对空间分层编码视频图象的上采样系统，其特征在于，所述上采样应用于“I_BL”层间预测或残差图象层间预测。

7. 根据权利要求 6 所述的对空间分层编码视频图象的上采样系统，其特征在于，应用于“I_BL”层间预测上采样时，所述第一滤波器是[1 -5 20 20 -5 1]/32，所述第二滤波器是[-1 5 5 -1]/8。

8. 根据权利要求 6 所述的对空间分层编码视频图象的上采样系统，其特征在于，应用于残差图象层间预测上采样时，所述第一滤波器是[-1 5 5 -1]/8，所述第二滤波器是[1 1]/2。

9. 一种对空间分层编码视频图象的上采样方法，其特征在于，在空间

分层编码中，对于残差图象层间预测，在残差图象上采样过程中，亮度分量和色度分量上采样都采用滤波器 [-1 5 5 -1]/8。

对空间分层编码视频图象的上采样方法及其系统

技术领域

本发明涉及视频图象压缩技术，特别涉及空间分层编码视频图象的上采样技术。

背景技术

H.264 视频压缩编码标准由国际电信联盟电信标准部（International Telecommunication Union Telecommunication Standardization Sector，简称“ITU-T”）联合国际标准组织（International Organization for Standardization，简称“ISO”）和国际电工委员会（International Electrotechnical Commission，简称“IEC”）的运动图像专家组（Moving Picture Expert Group，简称“MPEG”）制定。

目前，H.264 已经逐渐成为多媒体通信中的主流标准，大量采用 H.264 标准的多媒体实时通信产品，比如会议电视、可视电话、第三代（3rd Generation，简称“3G”）移动通信终端，以及网络流媒体产品先后问世。可以说，是否支持 H.264 已经成为这个市场领域中决定产品竞争力的关键因素。尤其是随着 3G 移动通信系统的出现和网际协议（Internet Protocol，简称“IP”）网络的迅速发展，视频网络通信正逐步成为通信的主要业务之一。

从发展来看，ITU-T 继制定了 H.261、H.263、H.263+等视频压缩标准后，于 2003 年正式发布了 H.264 标准，它同时也是 MPEG-4 第 10 部分的主要内容。H.264 标准的制定更加有效地提高了视频编码效率和网络适配性，随着 H.264 的推广和使用，IP 网络和移动无线网络的多媒体通信进入了一个飞跃发展的新阶段。

下面简单介绍 H.264 标准的消息构成及发送机制：H.264 标准采用分层模式，定义了视频编码层(Video Coding Layer，简称“VCL”)和网络抽象层（Network Abstraction Layer，简称“NAL”），后者专为网络传输设计，能适应不同网络中的视频传输，进一步提高网络的“亲和性”。H.264 引入了面向 IP 包的编码机制，有利于网络中的分组传输，支持网络中视频的流媒体传输；具有较强的抗误码特性，特别适应丢包率高、干扰严重的无线视频传输的要求。H.264 的所有待传送数据，包括图像数据及其他消息均封装为统一格式的包传送，即网络抽象层单元（NAL Unit，简称“NALU”）。每个 NALU 是一个一定语法元素的可变长字节字符串，包括包含一个字节的头信息，可用来表示数据类型，以及若干整数字节的负荷数据。一个 NAL 单元可以携带一个编码片、各自类型数据分割或一个序列或图像参数集。为了加强数据可靠性，每帧图像都被分为若干个条带（Slice），每个 Slice 由一个 NALU 承载，Slice 又是由若干个更小的宏块组成，即为最小的处理单元。一般的说，前后帧对应位置的 Slice 相互关联，不同位置的 Slice 相互独立，这样可以避免 Slice 之间发生误码相互扩散。

H.264 数据包含非参考帧的纹理数据、序列参数、图像参数、补充增强消息（Supplemental Enhancement Information，简称“SEI”）、参考帧纹理数据等。其中，SEI 消息是在 H.264 视频的解码、显示及其它方面起辅助作用的消息的统称。

图 1 示出 H.264 的压缩处理框架。H.264 的基本处理单位为 16×16 宏块，采用了多帧参考、帧内预测、多宏块模式、4×4 整数变换和量化、环滤波、1/4 象素运动预测、基于上下文的自适应可变长编码（Context-based Adaptive Variable Length Coding，简称“CAVLC”）和上下文自适应的二进制算术编码（Context-Adaptive Binary Arithmetic Coding，简称“CABAC”）熵编码等先进技术，因此 H.264 的压缩效率比 MPEG-2、H.263、MPEG-4 ASP 提高一倍多。

从2005年1月的第MPEG71次会议到2006年6月，联合视频组（Joint Video Team，简称“JVT”）将进行H.264分层编码标准制定工作，基本层兼容H.264 main profile，采用了MCTF的算法框架，由此可以较好地实现空间（spatial）分层、时间（temporal）分层、质量（quality或SNR）分层、复杂度（complexity）分层等功能。其中，JVT SVC（Joint Video Team Scalable video coding，暂译为“联合视频组视频分层编码”）的最新参考模型为JSVM3（Joint Scalable Video Model 3，暂译为“联合视频分层编码模型3”），图2示出上述SVC算法框图。需要指出的是，时间分解过程可以采用如图3所示的，基于B帧分解的方法，或如图4所示的，基于MCTF分解的方法。

具体地说，在帧内预测（INTRA prediction）方面，JSVM3基本层采用H.264帧内预测方法，增强层在H.264的预测模式基础上增加了一种预测模式：I_BL。I_BL就是将当前层的宏块从低层对应宏块逐象素进行预测，如图5所示。

此外，对于增强层宏块残差图象，即预测求差后的差值图象，也可以从基本层或较低层相应宏块的残差图象进行预测，方法与I_BL类似。

对于空间分层编码来说，在进行增强层I_BL预测或者残差图象预测时，基本层或较低层相应宏块必须经过上采样（upsampling）过程。上采样是重采样（Resampling）的一种，重采样用于采样信号的放大或缩小。假设原采样点位于整数坐标(0，1，2，…)，重采样后，新采样点之间的间距为a，则当a>1时，为下采样（downsampling）；当a<1时为上采样（upsampling）。

在JSVM3中，I_BL预测时的上采样滤波器为较为复杂的[1 -5 20 20 -5 1]/32，残差图象预测时的上采样滤波器为[1 1]/2，亮度图象和色度图象都采用同一个滤波器。

在实际应用中，上述方案存在以下问题：对色度分量采用与亮度分量相同的较为复杂的6阶滤波器，使色度分量上采样过程的计算复杂度过高。

在残差图象上采样过程中，采用的滤波器[1 1]/2 过于简单，从而影响编码性能。

发明内容

有鉴于此，本发明的主要目的在于提供一种对空间分层编码视频图象的上采样方法及其系统，使得在上采样过程中，可以减少计算复杂度，而编码性能基本不变。

本发明另一个发明目的是在残差图象上采样过程中提高编码性能。

为实现上述目的，本发明提供了一种对空间分层编码视频图象的上采样方法，其中，使用第一滤波器对图层的亮度分量进行上采样，使用第二滤波器对图层的色度分量进行上采样，第二滤波器比第一滤波器简单。

其中，所述上采样应用于“I_BL”层间预测或残差图象层间预测。

此外在所述方法中，“I_BL”层间预测上采样时，所述第一滤波器是[1 -5 20 20 -5 1]/32，所述第二滤波器是[-1 5 5 -1]/8。

此外在所述方法中，残差图象层间预测上采样时，所述第一滤波器是[-1 5 5 -1]/8，所述第二滤波器是[1 1]/2。

本发明还提供了一种对空间分层编码视频图象的上采样系统，其中至少包含用于对图层的亮度分量进行上采样的第一滤波器，和对图层的色度分量进行上采样的第二滤波器，第二滤波器比第一滤波器简单。

其中，所述上采样应用于“I_BL”层间预测或残差图象层间预测。

此外在所述系统中，应用于“I_BL”层间预测上采样时，所述第一滤波器是[1 -5 20 20 -5 1]/32，所述第二滤波器是[-1 5 5 -1]/8。

此外在所述系统中，应用于残差图象层间预测上采样时，所述第一滤波器是[-1 5 5 -1]/8，所述第二滤波器是[1 1]/2。

本发明还提供了一种对空间分层编码视频图象的上采样方法，在空间分层编码中，对于残差图象层间预测，在残差图象上采样过程中，亮度分量和色度分量上采样都采用滤波器 [-1 5 5 -1]/8。

通过比较可以发现，本发明的技术方案与现有技术的主要区别在于，利用人眼对色度分量的敏感度远远低于亮度分量的原理，在 I_BL 层间预测或残差图象层间预测的上采样过程中，对色度分量采用较亮量分量更简单的滤波器，从而有效减少了计算复杂度，而编码性能基本不变。

在残差图象上采样过程中，本发明提出至少在亮度分量上使用较为复杂的[-1 5 5 -1]/8 滤波器，从而提高编码性能。

附图说明

图 1 是现有技术中 H.264 压缩处理框架的示意图；

图 2 是现有技术中 SVC 算法框图；

图 3 是现有技术中基于 B 帧的分解框图；

图 4 是现有技术中基于 MCTF 的分解框图；

图 5 是现有技术中增强层的 I_BL 预测模式示意图。

具体实施方式

为使本发明的目的、技术方案和优点更加清楚，下面将结合附图对本发明作进一步地详细描述。

本发明的要点在于，基于人眼对色度分量的敏感度远远低于亮度分量的自然规律，在上采样过程中，对亮度分量和色度分量的上采样采用不同的滤

波器，其中色度分量所用的滤波器比亮度分量所用的滤波器更为简单。

这会带来两个后果，一个后果是色度分量的编码性能不可避免地有所下降，另一个后果是色度分量的计算复杂度也相应下降。前一个后果是不利的，但由于人眼对色度分量的敏感度远远低于亮度分量的自然规律，色度分量上性能的下降在人眼看来很不明显，换句话说普通人几乎察觉不出图像在颜色上已经发生了变化。后一个效果有有利的，实实在在地减少了运算器件的工作量，使同样的设备可以承载更多的业务，使处理时延更小，画面更为流畅。

下面说明一下本发明的实施方式的应用环境。在分层编码中，增强层编码过程需要充分利用基本层或者较低层的信息，以便去除冗余性，提高编码效率，例如，SVC 分别采用了 I_BL 层间预测、残差图象层间预测、运动向量层间预测等技术。对于空间分层编码来说，在进行增强层 I_BL 预测或者残差图象预测时，基本层或较低层相应宏块必须经过上采样（upsampling）过程，以方便进行求差预测。在 JSVM3 中，I_BL 预测时的上采样滤波器为[1 -5 20 20 -5 1]/32，残差图象预测时的上采样滤波器为[1 1]/2，亮度图象和色度图象都采用同一个滤波器。

在本发明的第一实施方式中，在 I_BL 预测时，亮度分量和色度分量上采样采用不同的滤波器，亮度分量上采样滤波器采用[1 -5 20 20 -5 1]/32，色度分量上采样滤波器采用较为简单的[-1 5 5 -1]/8。也就是说将 JSVM3 中的色度分量采样滤波器从[1 -5 20 20 -5 1]/32 改为[-1 5 5 -1]/8。使用[1 -5 20 20 -5 1]/32 滤波器时，每一个采样点的计算需要作 6 次乘法运算、5 次加法运算和一次除法运算，而使用[-1 5 5 -1]/8 滤波器时，每一个采样点的计算需要作 4 次乘法运算、3 次加法运算和一次除法运算，大致上减少了三分之一的运算量。而由于人眼对色度分量不太敏感，所以在实际视觉上感受不到太大的差异。

在本发明的第二实施方式中，在残差图象上采样过程中，亮度分量和色

度分量上采样也采用不同的滤波器，亮度分量上采样采用滤波器[-1 5 5 -1]/8，而色度分量上采样采用滤波器[1 1]/2。如果色度分量上也采用[-1 5 5 -1]/8 滤波器，则每一个采样点的计算需要作 4 次乘法运算、3 次加法运算和一次除法运算，现在使用[1 1]/2 只要做一次加法和一次除法，如果考虑到除 2 的整数次方可以用较快速的移位运算代替的话，运算量可能比[-1 5 5 -1]/8 滤波器减少了一个数量级。本实施方式中，亮度分量采用了比 JSVM3 更高的[-1 5 5 -1]/8 滤波器，这可以提高编码性能，而计算复杂度的增加则在可接受的范围内。

在本发明的第三实施方式中，为了提高编码性能，在残差图象上采样过程中，亮度分量和色度分量上采样都采用滤波器[-1 5 5 -1]/8。为了证明该滤波器的有效性，本发明对一些典型序列进行了仿真测试。下表为对 Harbour 序列 300 帧图象、QCIF 和 CIF 两层编码的测试结果性能对比，其中本发明技术结果为 I_BL 预测中色度分量上采样滤波器采用滤波器[-1 5 5 -1]/8、残差图象预测中亮度分量和色度分量上采样都采用滤波器[-1 5 5 -1]/8 的测试结果。

	原技术结果				本发明技术结果			
QP	码流量	Y	Cb	Cr	码流量	Y	Cb	Cr
8	59496.902	50.7719	52.0737	52.632	59095.658	50.7719	52.0702	52.6277
12	43774.38	47.5296	49.3939	50.1502	43404.455	47.527	49.3866	50.14
16	30060.991	44.475	47.3343	48.4443	29752.754	44.474	47.3262	48.4323
20	19368.697	41.399	45.8185	47.3728	19133.755	41.3966	45.8097	47.3585
24	11262.018	38.4222	44.5876	46.5447	11124.914	38.4233	44.5841	46.5474
28	5870.7624	35.6978	43.3869	45.5161	5799.3328	35.6983	43.378	45.5082
32	2873.7904	33.2981	42.2992	44.3998	2847.3504	33.3022	42.3175	44.4073
36	1412.7112	31.1973	41.2843	43.2547	1403.5464	31.1975	41.2769	43.247

测试结果显示，采用本发明技术后码流量有所下降，而亮度分量信噪比 PSNR 值有所增加、色度分量信噪比 PSNR 值基本一致，从而编码性能有所提高。

需要指出的是，以上的实施方式只是本发明技术方案的一些典型应用，除了上述滤波器[-1 5 5 -1]/8 外，在 I_BL 上采样过程中，色度分量上采样还

可以采用其它滤波器；在残差图象上采样过程中，亮度分量和色度分量上采样也可以采用其它滤波器。

虽然通过参照本发明的某些优选实施方式，已经对本发明进行了图示和描述，但本领域的普通技术人员应该明白，可以在形式上和细节上对其作各种改变，而不偏离本发明的精神和范围。

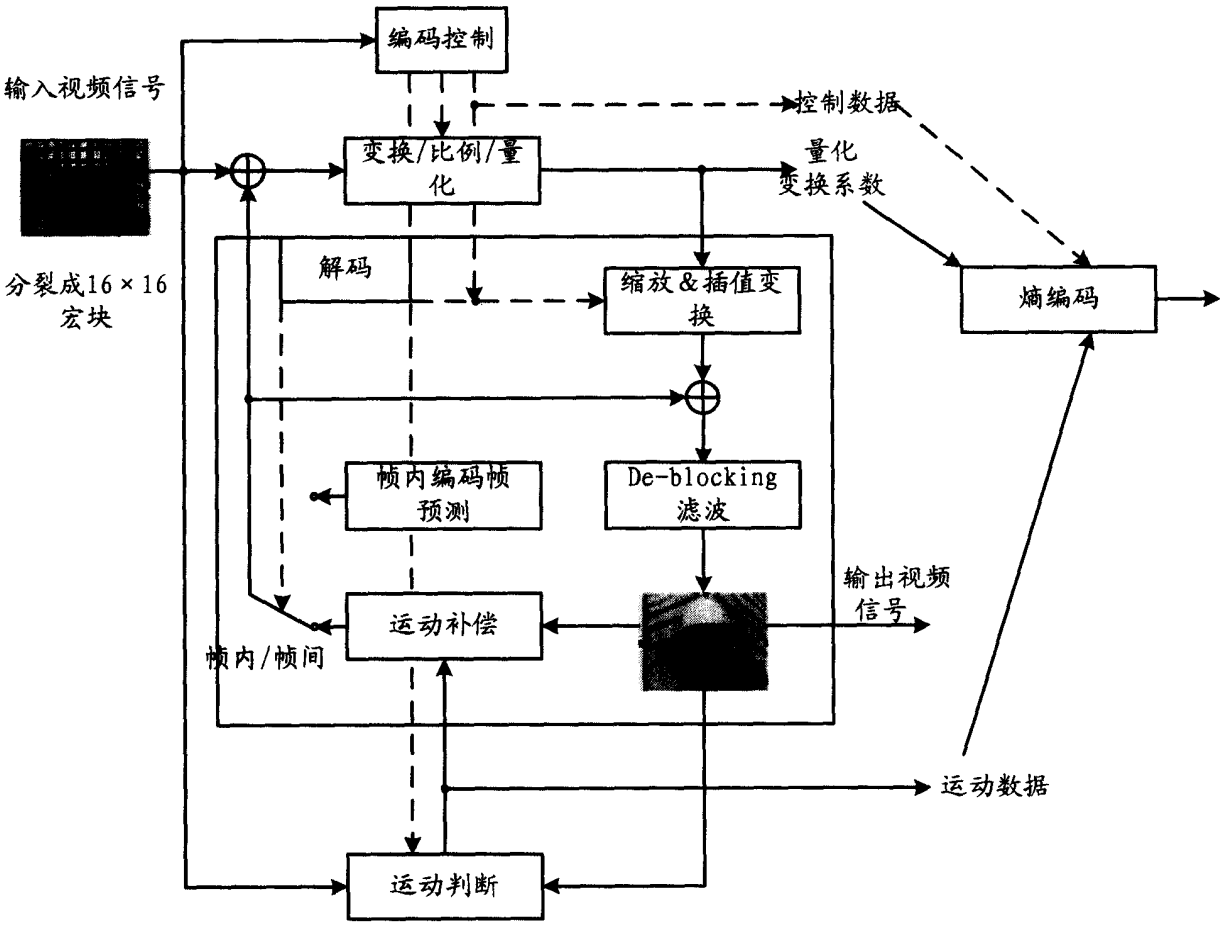

编码控制
输入视频信号
控制数据
变换/比例/量化
量化
变换系数
分裂成16×16
宏块
解码
缩放&插值变换
熵编码
帧内编码帧预测
De-blocking滤波
输出视频信号
运动补偿
帧内/帧间
运动数据
运动判断

图 1

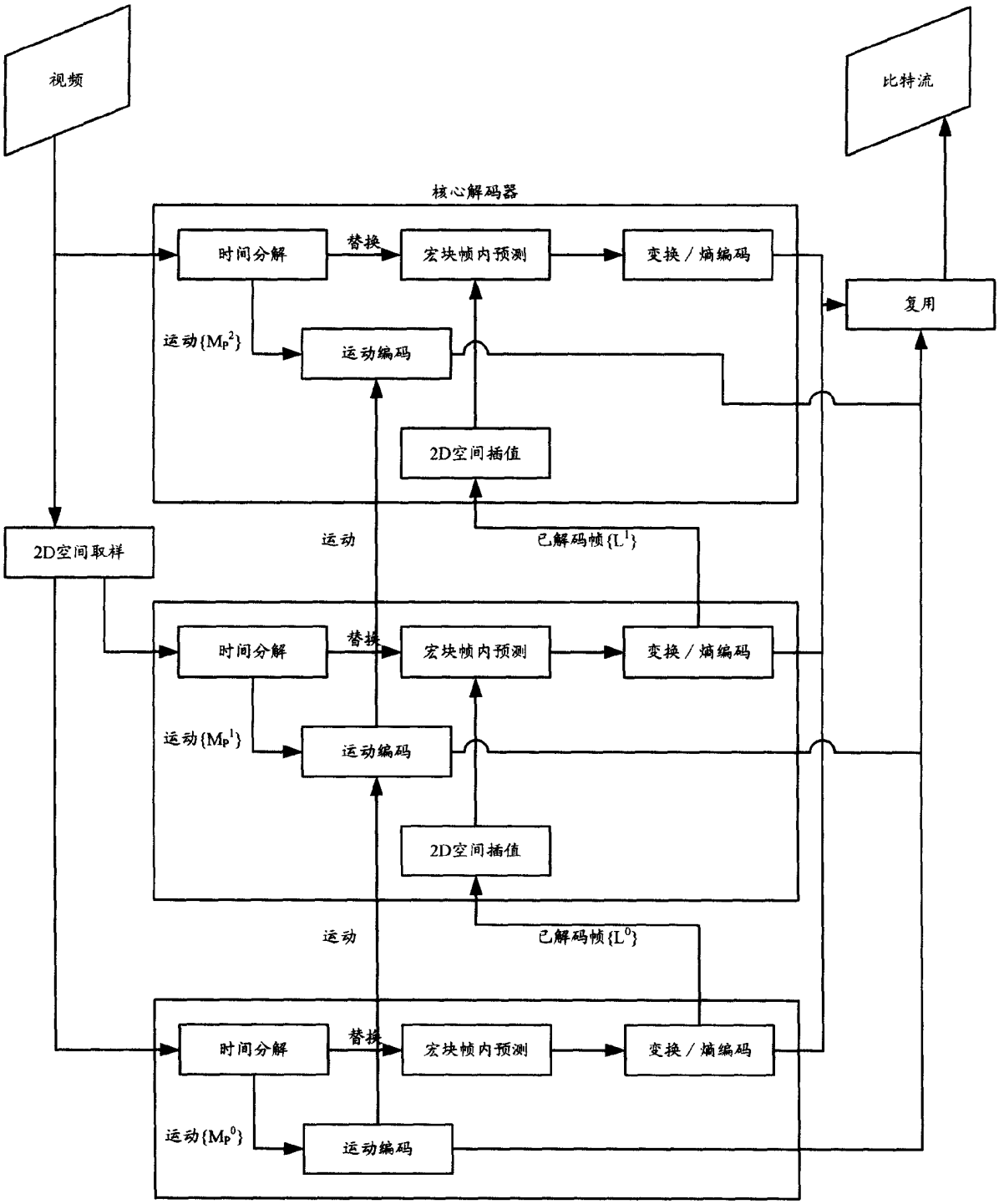

视频
比特流
核心解码器
时间分解
替换
宏块帧内预测
变换／熵编码
复用
运动{MP2}
运动编码
2D空间插值
2D空间取样
运动
已解码帧{L1}
时间分解
替换
宏块帧内预测
变换／熵编码
运动{MP1}
运动编码
2D空间插值
运动
已解码帧{L0}
时间分解
替换
宏块帧内预测
变换／熵编码
运动{MP0}
运动编码

图 2

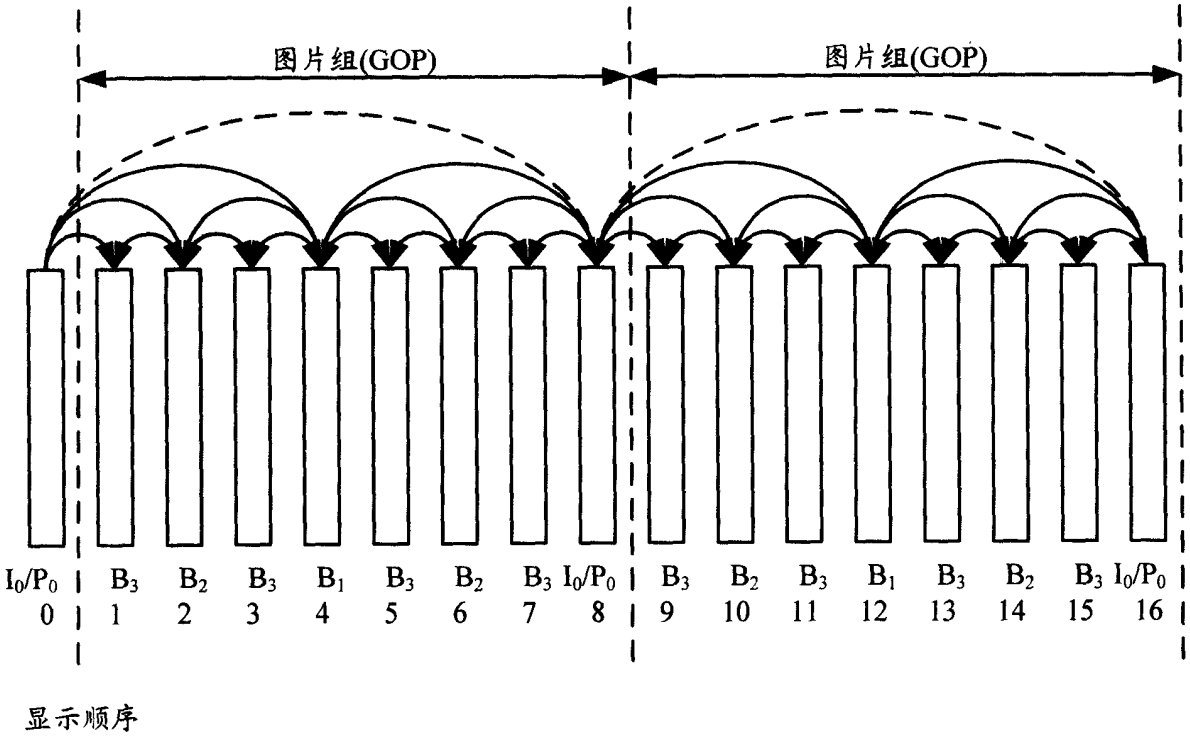

图片组(GOP)
图片组(GOP)
I0/P0 B3 B2 B3 B1 B3 B2 B3 I0/P0 B3 B2 B3 B1 B3 B2 B3 I0/P0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
显示顺序

图 3

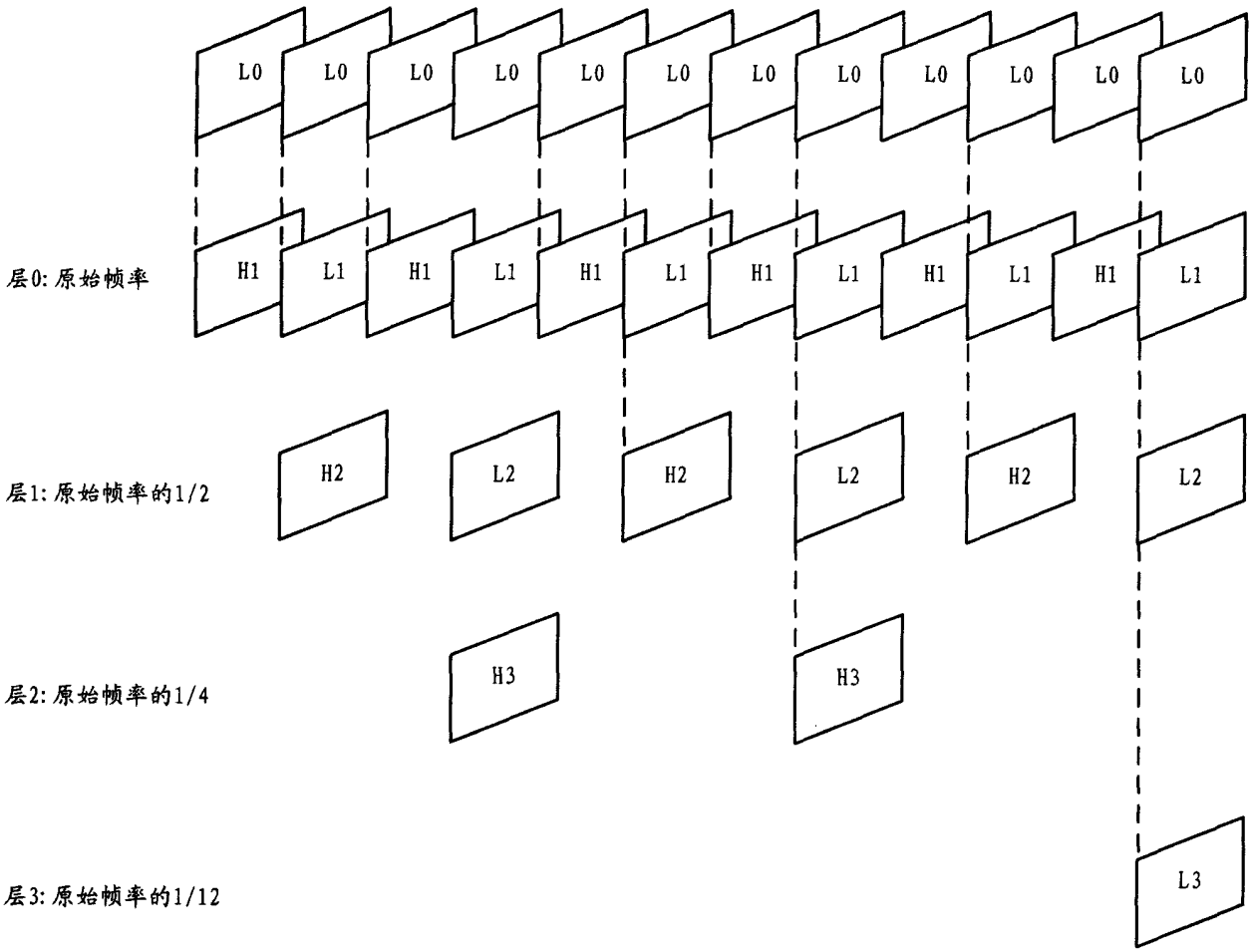

层0: 原始帧率
层1: 原始帧率的1/2
层2: 原始帧率的1/4
层3: 原始帧率的1/12
L0
H1
L1
H2
L2
H3
L3

图 4

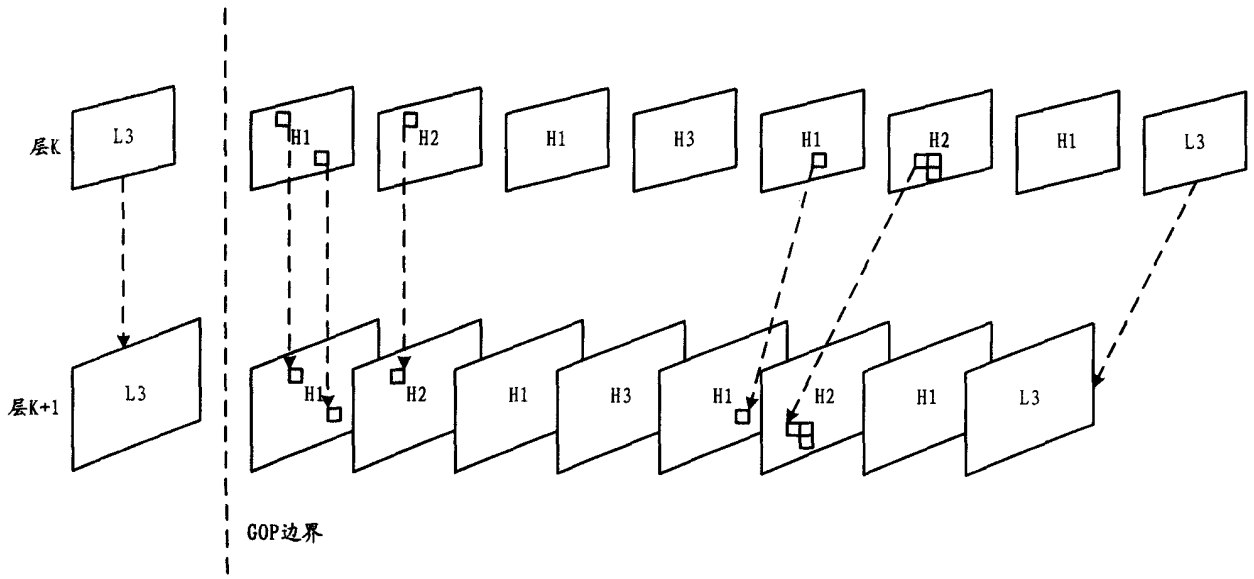

层K
层K+1
L3
H1
H2
H3
GOP边界

图 5