

分层编解码的方法及装置

技术领域

本发明实施例涉及基于人工智能（AI）的视频或图像压缩技术领域，尤其涉及一种分层编解码的方法及装置。

背景技术

视频压缩编解码技术在多媒体服务、广播、视频通信和存储等领域都有广泛的应用，例如广播数字电视、互联网和移动网络上的视频传输、视频聊天和视频会议等实时会话应用、DVD 和蓝光光盘、视频内容采集和编辑系统以及可携式摄像机的安全应用。

即使在影片较短的情况下也需要对大量的视频数据进行描述，当数据要在带宽容量受限的网络中发送或以其它方式传输时，这样可能会造成困难。因此，视频数据通常要先压缩然后在现代电信网络中传输。由于内存资源可能有限，当在存储设备上存储视频时，视频的大小也可能成为问题。视频压缩设备通常在信源侧使用软件和/或硬件，以在传输或存储之前对视频数据进行编码，从而减少用来表示数字视频图像所需的数据量。然后，压缩的数据在目的地侧由视频解压缩设备接收。在有限的网络资源以及对更高视频质量的需求不断增长的情况下，需要改进压缩和解压缩技术，这些改进的技术能够提高压缩率而几乎不影响图像质量。

近年来，将深度学习应用于在端到端的图像编解码技术领域逐渐成为一种趋势。在采用混合架构的视频编码器和视频解码器中，对特征图进行熵编码时，假定特征值满足零均值的高斯分布，超先验结构估计高斯分布的方差，得到特征值的概率分布模型，算术编码模块基于估计的概率分布对特征图进行熵编码。为了使解码端同样能准确地估计特征图的概率分布，超先验结构中模块提取估计概率分布的隐变量，隐变量经由量化、算术编码作为边信息传递到解码端。在这样的机制下，输入图像为 YUV 格式时，Y、U、V 分量的码率占比固定。但是由于图像内容色彩特性不同，YUV 分量码率固定，会导致编码的图像失真较大。

发明内容

本申请提供一种分层编解码的方法及装置，能够适配不同色彩特性的图像内容。

在本申请中，分层编解码是指把视频信号分为第一信号分量和第二信号分量；或者

把视频信号分为第一信号分量，第二信号分量和第三信号分量。所述第一信号分量为 Y 分量，所述第二信号分量为 UV 分量，U 分量或 V 分量。当所述第二信号分量为 U 分量时，所述第三信号分量为 V 分量，或当所述第二信号分量为 V 分量，所述第三信号分量为 U 分量。

第一方面，本申请提供一种编码方法。所述编码方法包括：把所述视频信号的第一信号分量的控制信号作用于所述第一信号分量的第一特征图，获得所述第一信号分量的第二特征图，其中所述第一信号分量的控制信号通过学习获得；把所述视频信号的第二信号分量的控制信号作用于所述第二信号分量的第一特征图，获得所述第二信号分量的第二特征图，其中所述第二信号分量的控制信号通过学习获得；以及根据所述第一信号分量的第二特征图和所述第二信号分量的第二特征图，获得所述视频信号的码流。

在一种可能的实现方式中，根据所述第一信号分量的第二特征图和所述第二信号分量的第二特征图，获得所述视频信号的码流包括：对所述第一信号分量的第二特征图和所述第二信号分量的第二特征图，或者，对所述第一信号分量的第二特征图和经神经网络处理的所述第二信号分量的第二特征图，或者，对经神经网络处理的所述第一信号分量的第二特征图和所述第二信号分量的第二特征图，或者，对经神经网络处理的所述第一信号分量的第二特征图和经神经网络处理的所述第二信号分量的第二特征图，进行熵编码，以获得所述视频信号的码流。

在一种可能的实现方式中，根据所述第一信号分量的第二特征图和所述第二信号分量的第二特征图，获得所述视频信号的码流包括：对所述第一信号分量的第二特征图和所述第二信号分量的第二特征图，进行联合处理，获得联合的特征图，对所述联合的特征图进行熵编码，以获得所述视频信号的码流；或者，对所述第一信号分量的第二特征图和经神经网络处理的所述第二信号分量的第二特征图，进行联合处理，获得联合的特征图，对所述联合的特征图进行熵编码，以获得所述视频信号的码流；或者，对经神经网络处理的所述第一信号分量的第二特征图和所述第二信号分量的第二特征图，进行联合处理，获得联合的特征图，对所述联合的特征图进行熵编码，以获得所述视频信号的码流；或者，对经神经网络处理的所述第一信号分量的第二特征图和经神经网络处理的所述第二信号分量的第二特征图，进行联合处理，获得联合的特征图，对所述联合的特征图进行熵编码，以获得所述视频信号的码流。

在一种可能的实现方式中，根据所述第一信号分量的质量因子从 N 个候选第一控制信号中，获得所述第一信号分量的控制信号，其中 N 为大于 1 的整数；以及根据所述第

二信号分量的质量因子从M个候选第二控制信号中,获得所述第二信号分量的控制信号,其中M为大于1的整数。N和M可以相等或者不等,本申请对此不做限定。

在一种可能的实现方式中,当所述第一信号分量为Y分量,所述第二信号分量为UV分量时,通过学习生成所述Y分量的控制信号矩阵 $\{q_{y1}, q_{y2}, \dots, q_{yi} \dots q_{yN}\}$ 和所述UV分量的控制信号矩阵 $\{q_{uv1}, q_{uv2}, \dots, q_{uvj}, \dots, q_{uvM}\}$,其中N和M为大于1的整数;根据所述Y分量的质量因子的索引i得到所述第一信号分量的控制信号 q_{yi} ,根据所述UV分量的质量因子的索引j得到所述第二信号分量的控制信号 q_{uvj} 。此时所述视频信号的码流中包括所述Y分量的质量因子的索引i和所述UV分量的质量因子的索引j。或者

在另一种可能的实现方式中,当所述第一信号分量为Y分量,所述第二信号分量为UV分量时,通过学习生成所述视频信号的控制信号矩阵 $\{q_{c1}, q_{c2}, \dots, q_{ci} \dots q_{cN}\}$,其中c为2代表Y分量和UV分量,N为大于1的整数;根据所述视频信号的质量因子的索引i得到包含所述第一信号分量和所述第二信号分量的控制信号 q_{ci} 。此时所述视频信号的码流中包括所述视频信号的质量因子的索引i。或者

在再一种可能的实现方式中,当所述第一信号分量为Y分量,所述第二信号分量为UV分量时,通过全连接网络实现,则将所述Y分量的质量因子作为全连接网络的输入,输出所述Y分量的控制信号;将所述UV分量的质量因子作为全连接网络的输入,输出所述UV分量的控制信号。此时所述视频信号的码流中包括所述Y分量的质量因子和所述UV分量的质量因子。

在一种可能的实现方式中,当所述第二信号分量为U分量或者V分量时,所述方法还包括:把所述视频信号的第三信号分量的控制信号作用于所述第三信号分量的第一特征图,获得所述第三信号分量的第二特征图,其中所述第三信号分量的控制信号通过学习获得,其中当所述第二信号分量为U分量时,所述第三信号分量为V分量,或当所述第二信号分量为V分量,所述第三信号分量为U分量。

在一种可能的实现方式中,当所述第一信号分量为Y分量,所述第二信号分量为U分量时,所述第三信号分量为V分量,通过学习生成所述Y分量的控制信号矩阵 $\{q_{y1}, q_{y2}, \dots, q_{yi} \dots q_{yN}\}$ 、所述U分量的控制信号矩阵 $\{q_{u1}, q_{u2}, \dots, q_{uj}, \dots, q_{uM}\}$ 、所述V分量的控制信号矩阵 $\{q_{v1}, q_{v2}, \dots, q_{vk}, \dots, q_{vL}\}$,其中N,M和L为大于1的整数;根据所述Y分量质量因子的索引i得到所述第一信号分量的控制信号 q_{yi} ;根据所述U分量质量因子的索引j得到所述第二信号分量的控制信号 q_{uj} ;以及根据所述V分量质量因子的索引k得到所述第三信号分量的控制信号 q_{vk} 。此时所述视频信号的码流中包括所述Y分量的质量因

子的索引 i 、所述 U 分量的质量因子的索引 j 、和所述 V 分量质量因子的索引 k 。或者

在另一种可能的实现方式中，当所述第一信号分量为 Y 分量，所述第二信号分量为 U 分量时，所述第三信号分量为 V 分量，通过学习生成所述视频信号的控制信号矩阵 $\{q_{c1}, q_{c2}, \dots, q_{ci} \dots q_{cN}\}$ ，其中 c 为 3 代表 Y 分量、 U 分量和 V 分量， N 为大于 1 的整数；根据所述视频信号的质量因子的索引 i 得到包含所述第一信号分量、所述第二信号分量和所述第三信号分量的控制信号 q_{ci} 。此时所述视频信号的码流中包括所述视频信号的质量因子的索引 i 。或者

在再一种可能的实现方式中，当所述第一信号分量为 Y 分量，所述第二信号分量为 U 分量时，所述第三信号分量为 V 分量，通过全连接网络实现，则将所述 Y 分量的质量因子作为全连接网络的输入，输出 Y 分量的控制信号；将所述 U 分量的质量因子作为全连接网络的输入，输出 U 分量的控制信号；以及将所述 V 分量的质量因子作为全连接网络的输入，输出 V 分量的控制信号。此时所述视频信号的码流中包括所述 Y 分量的质量因子、所述 U 分量的质量因子和所述 V 分量的质量因子。

第二方面，本申请提供一种解码方法。所述解码方法包括：获得所述视频信号的码流；对所述码流进行熵解码以获得所述视频信号的第一信号分量的特征图和所述视频信号的第二信号分量的特征图；根据所述第一信号分量的响应信号和所述第一信号分量的特征图，获得所述第一信号分量的重建图，其中所述第一信号分量的响应信号通过学习获得；根据所述第二信号分量的响应信号和所述第二信号分量的特征图，获得所述第二信号分量的重建图，其中所述第二信号分量的响应信号通过学习获得；以及根据所述第一信号分量的重建图和所述第二信号分量的重建图，重建所述视频信号。

需要说明的是，本申请中解码端的响应信号与编码端的控制信号类似，为了区分解码端称为响应信号，编码端称为控制信号。解码端的响应信号包括响应向量，或者包括响应向量和偏移向量。

在一种可能的实现方式中，所述码流中还包括所述第一信号分量的质量因子信息和所述第二信号分量的质量因子信息，其中所述第一信号分量的质量因子信息为所述第一信号分量的质量因子或者所述第一信号分量的质量因子的索引，所述第二信号分量的质量因子信息为所述第二信号分量的质量因子或者所述第二信号分量的质量因子的索引；通过所述第一信号分量的质量因子信息，获得所述第一信号分量的响应信号；通过所述第二信号分量的质量因子信息，获得所述第二信号分量的响应信号。

其中，当所述第一信号分量的质量因子信息为所述第一信号分量的质量因子时，所

述第一信号分量的质量因子取值为 N 中的一个；当所述第一信号分量的质量因子信息为所述第一信号分量的质量因子的索引时，所述第一信号分量的质量因子的索引的取值范围为 0 至 $N-1$ 或者 1 至 N ，其中 N 为大于 1 的整数；

当所述第二信号分量的质量因子信息为所述第二信号分量的质量因子时，所述第二信号分量的质量因子取值为 M 中的一个；当所述第二信号分量的质量因子信息为所述第二信号分量的质量因子的索引时，所述第二信号分量的质量因子的索引的取值范围为 0 至 $M-1$ 或者 1 至 M ，其中 M 为大于 1 的整数。

在一种可能的实现方式中，当所述第一信号分量为 Y 分量，所述第二信号分量为 UV 分量时，若所述码流中包括所述 Y 分量的质量因子的索引 i 和所述 UV 分量的质量因子的索引 j ，通过学习生成所述第一信号分量的响应信号矩阵 $\{g_{y1}, g_{y2}, \dots, g_{yi} \dots g_{yN}\}$ 、和所述第二信号分量的响应信号矩阵 $\{g_{uv1}, g_{uv2}, \dots, g_{uvj}, \dots, g_{uvM}\}$ ，其中 N 和 M 为大于 1 的整数；根据所述 Y 分量的质量因子的索引 i 得到所述第一信号分量的响应信号 g_{yi} ；根据所述 UV 分量的质量因子的索引 j 得到所述第二信号分量的响应信号 g_{uvj} 。或者

在另一种可能的实现方式中，当所述第一信号分量为 Y 分量，所述第二信号分量为 UV 分量时，若所述码流中包括所述视频信号的质量因子的索引 i ，通过学习生成所述视频信号的响应信号矩阵 $\{g_{c1}, g_{c2}, \dots, g_{ci} \dots g_{cN}\}$ ，其中 c 为 2 代表 Y 分量和 UV 分量， N 为大于 1 的整数；根据所述视频信号的质量因子的索引 i 得到包含所述第一信号分量和所述第二信号分量的响应信号 g_{ci} 。

在再一种可能的实现方式中，当所述第一信号分量为 Y 分量，所述第二信号分量为 UV 分量时，若所述码流中包括所述第一信号分量的质量因子和第二信号分量的质量因子，通过全连接网络实现，则将所述 Y 分量的质量因子作为全连接网络的输入，输出所述 Y 分量的响应信号；将所述 UV 分量的质量因子作为全连接网络的输入，输出所述 UV 分量的响应信号。

在一种可能的实现方式中，当所述第二信号分量为 U 分量或者 V 分量时，所述方法还包括：对所述码流进行熵解码以获得所述视频信号的第三信号分量的特征图；根据所述第三信号分量的响应信号和所述第三信号分量的特征图，获得所述第三信号分量的重建图，其中所述第三信号分量的响应信号通过学习获得，其中当所述第二信号分量为 U 分量时，所述第三信号分量为 V 分量，或当所述第二信号分量为 V 分量，所述第三信号分量为 U 分量。则所述重建所述视频信号包括：根据所述第一信号分量的重建图、所述第二信号分量的重建图和和所述第三信号分量的重建图，重建所述视频信号。

在一种可能的实现方式中,当所述第一信号分量为 Y 分量,所述第二信号分量为 U 分量时,所述第三信号分量为 V 分量,若所述码流中包括所述 Y 分量的质量因子的索引 i、所述 U 分量的质量因子的索引 j,和所述 V 分量的质量因子的索引 k,通过学习生成所述第一信号分量的响应信号矩阵 $\{g_{y1}, g_{y2}, \dots, g_{yi} \dots g_{yN}\}$ 、所述第二信号分量的响应信号矩阵 $\{g_{u1}, g_{u2}, \dots, g_{uj} \dots, g_{uM}\}$ 、所述第三信号分量的响应信号矩阵 $\{g_{v1}, g_{v2}, \dots, g_{vk} \dots g_{vL}\}$, 其中 N, M 和 L 为大于 1 的整数;根据所述 Y 分量质量因子的索引 i 得到所述第一信号分量的响应信号 g_{yi} ; 根据所述 U 分量质量因子的索引 j 得到所述第二信号分量的响应信号 g_{uj} ; 以及根据所述 V 分量质量因子的索引 k 得到所述第三信号分量的响应信号 g_{vk} 。或者

在另一种可能的实现方式中,当所述第一信号分量为 Y 分量,所述第二信号分量为 U 分量时,所述第三信号分量为 V 分量,若所述码流中包括所述视频信号的质量因子的索引 i,通过学习生成所述视频信号的响应信号矩阵 $\{g_{c1}, g_{c2}, \dots, g_{ci} \dots g_{cN}\}$, 其中 c 为 3 代表 Y 分量、U 分量和 V 分量, N 为大于 1 的整数;根据所述视频信号的质量因子的索引 i 得到包含所述第一信号分量、所述第二信号分量和所述第三信号分量的响应信号 g_{ci} 。或者

在再一种可能的实现方式中,当所述第一信号分量为 Y 分量,所述第二信号分量为 U 分量时,所述第三信号分量为 V 分量,若所述码流中包括所述第一信号分量的质量因子、所述第二信号分量的质量因子和所述第三信号分量的质量因子,通过全连接网络实现,则将所述 Y 分量的质量因子作为全连接网络的输入,输出 Y 分量的响应信号;将所述 U 分量的质量因子作为全连接网络的输入,输出 U 分量的响应信号;以及将所述 V 分量的质量因子作为全连接网络的输入,输出 V 分量的响应信号。

第三方面,本申请提供一种编码器,包括处理电路,用于执行根据上述第一方面及第一方面任一项所述的方法。

第四方面,本申请提供一种解码器,包括处理电路,用于执行上述第二方面及第二方面任一项所述的方法。

第五方面,本申请提供一种计算机程序产品,包括程序代码,当其在计算机或处理器上执行时,用于执行上述第一方面及第一方面任一项、上述第二方面及第二方面任一项所述的方法。

第六方面,本申请提供一种编码器,包括:一个或多个处理器;非瞬时性计算机可读存储介质,耦合到所述处理器并存储由所述处理器执行的程序,其中所述程序在由所述处理器执行时,使得所述解码器执行上述第一方面及第一方面任一项所述的方法。

第七方面,本申请提供一种解码器,包括:一个或多个处理器;非瞬时性计算机可

读存储介质，耦合到所述处理器并存储由所述处理器执行的程序，其中所述程序在由所述处理器执行时，使得所述编码器执行上述第二方面及第二方面任一项所述的方法所述的方法。

第八方面，本申请提供一种非瞬时性计算机可读存储介质，包括程序代码，当其由计算机设备执行时，用于执行上述第一方面及第一方面任一项、上述第二方面及第二方面任一项所述的方法。

第九方面，本发明涉及编码装置，具有实现上述第一方面或第一方面任一项的方法实施例中的行为的功能。所述功能可以通过硬件实现，也可以通过硬件执行相应的软件实现。所述硬件或软件包括一个或多个与上述功能相对应的模块。在一个可能的设计中，所述编码装置包括：第一控制模块，用于把所述视频信号的第一信号分量的控制信号作用于所述第一信号分量的第一特征图，获得所述第一信号分量的第二特征图，其中所述第一信号分量的控制信号通过学习获得；第二控制模块，用于把所述视频信号的第二信号分量的控制信号作用于所述第二信号分量的第一特征图，获得所述第二信号分量的第二特征图，其中所述第二信号分量的控制信号通过学习获得；以及编码模块，用于根据所述第一信号分量的第二特征图和所述第二信号分量的第二特征图，获得所述视频信号的码流。这些模块可以执行上述第一方面或第一方面任一项方法示例中的相应功能，具体参见方法示例中的详细描述，此处不做赘述。

第十方面，本发明涉及解码装置，具有实现上述第二方面或第二方面任一项的方法实施例中的行为的功能。所述功能可以通过硬件实现，也可以通过硬件执行相应的软件实现。所述硬件或软件包括一个或多个与上述功能相对应的模块。在一个可能的设计中，所述解码装置包括：解码模块，用于获得所述视频信号的码流，对所述码流进行熵解码以获得所述视频信号的第一信号分量的特征图和所述视频信号的第二信号分量的特征图；第一控制模块，用于根据所述第一信号分量的响应信号和所述第一信号分量的特征图，获得所述第一信号分量的重建图，其中所述第一信号分量的响应信号通过学习获得；第二控制模块，用于根据所述第二信号分量的响应信号和所述第二信号分量的特征图，获得所述第二信号分量的重建图，其中所述第二信号分量的响应信号通过学习获得；以及重建模块，用于根据所述第一信号分量的重建图和所述第二信号分量的重建图，重建所述视频信号。这些模块可以执行上述第二方面或第二方面任一项方法示例中的相应功能，具体参见方法示例中的详细描述，此处不做赘述。

在现有端到端图像编码中，对于某一特定网络学习优化时，按固定 Y、U、V 分量权

重值进行优化，因此，YUV 分量的码率固定。由于不同图像色彩特性不同，固定码率分配会导致部分视频图像内容上编码性能较差的现象。本申请上述各方面通过各信号分量的控制信号对相应的信号分量的特征图分别进行控制，从而可以支持 YUV 分量间的码率分配，达到适配不同色彩特性的图像内容。

附图及以下说明中将详细描述一个或多个实施例。其它特征、目的和优点在说明、附图以及权利要求中是显而易见的。

附图说明

下面对本申请实施例用到的附图进行介绍。

图 1A 为用于实现本发明实施例的视频译码系统示例的框图，其中该系统基于深度学习来编码或解码视频图像；

图 1B 为用于实现本发明实施例的视频译码系统示例另一示例的框图，其中该系统基于深度学习来编码或解码视频图像；

图 1C 为用于实现本发明实施例的视频译码系统再一示例的框图，其中该视频编码器和/或视频解码器基于深度学习来编码或解码视频图像；

图 2 为用于实现本发明实施例的视频编码器实例示例的框图，其中该视频编码器 20 基于深度学习来编码视频图像；

图 3 为用于实现本发明实施例的视频解码器实例示例的框图，其中该视频解码器 30 基于深度学习来解码视频图像；

图 4 为用于实现本发明实施例的视频译码装置的示意性框图；

图 5 为用于实现本发明实施例的视频译码装置的示意性框图；

图 6 为 YUV 格式示意图；

图 7A 是本申请实施例提供的分层编解码结构的示意图；

图 7B 是基于图 7A 的编码方法的一种实施例；

图 7C 是基于图 7A 的解码方法的一种实施例；

图 7D 是本申请实施例提供的分层编解码结构的另一示意图；

图 8A 是本申请实施例提供的分层编解码结构的示意图；

图 8B 是图 8A 的一种实施例；

图 9A 是本申请实施例提供的分层编解码结构的示意图；

图 9B 是图 9A 的一种实施例；

图 10 是示出根据本申请一种实施例的编码装置 1000 的结构示意图；

图 11 是示出根据本申请一种实施例的解码装置 1100 的结构示意图。

具体实施方式

本申请实施例提供一种基于 AI 的视频图像压缩技术，具体提供一种分层编解码中的方法及装置，以改进传统的基于端到端的混合视频编解码系统。

视频编码通常是指处理形成视频或视频序列的图像序列。在视频编码领域，术语“图像 (picture)”、“帧 (frame)”或“图片 (image)”可以用作同义词。视频编码（或通常称为编码）包括视频编码和视频解码两部分。视频编码在源侧执行，通常包括处理（例如，压缩）原始视频图像以减少表示该视频图像所需的数据量（从而更高效存储和/或传输）。视频解码在目的地侧执行，通常包括相对于编码器作逆处理，以重建视频图像。实施例涉及的视频图像（或通常称为图像）的“编码”应理解为视频图像或视频序列的“编码”或“解码”。编码部分和解码部分也合称为编解码（编码和解码，CODEC）。

在无损视频编码情况下，可以重建原始视频图像，即重建的视频图像与原始视频图像具有相同的质量（假设存储或传输期间没有传输损耗或其它数据丢失）。在有损视频编码情况下，通过量化等执行进一步压缩，来减少表示视频图像所需的数据量，而解码器侧无法完全重建视频图像，即重建的视频图像的质量比原始视频图像的质量较低或较差。

几个视频编码标准属于“有损混合型视频编解码”（即，将像素域中的空间和时间预测与变换域中用于应用量化的 2D 变换编码结合）。视频序列中的每个图像通常分割成不重叠的块集合，通常在块级上进行编码。换句话说，编码器通常在块（视频块）级处理及编码视频，例如，通过空间（帧内）预测和时间（帧间）预测来产生预测块；从当前块（当前处理/待处理的块）中减去预测块，得到残差块；在变换域中变换残差块并量化残差块，以减少待传输（压缩）的数据量，而解码器侧将相对于编码器的逆处理部分应用于编码或压缩的块，以重建用于表示的当前块。另外，编码器需要重复解码器的处理步骤，使得编码器和解码器生成相同的预测（例如，帧内预测和帧间预测）和/或重建像素，用于处理，即编码后续块。

在以下译码系统 10 的实施例中，编码器 20 和解码器 30 根据图 1B 至图 3 进行描述。

图 1A 为示例性译码系统的示意性框图，如图 1A 所示，使用视频采集设备将视频采集以后，经过一系列的前处理，再对处理以后的视频进行压缩编码，得到编码码流。使

用发送模块将码流经传输网络发送到接收模块，经解码器进行解码以后，便可渲染显示。除此之外，视频编码以后的码流也可以直接进行存储。

图 1B 为示例性译码系统 10 的示意性框图，例如可以利用本申请技术的视频译码系统 10（或简称为译码系统 10）。视频译码系统 10 中的视频编码器 20（或简称为编码器 20）和视频解码器 30（或简称为解码器 30）代表可用于根据本申请中描述的各种示例执行各技术的设备等。

如图 1B 所示，译码系统 10 包括源设备 12，源设备 12 用于将编码图像等编码图像数据 21 提供给用于对编码图像数据 21 进行解码的目的设备 14。

源设备 12 包括编码器 20，另外即可选地，可包括图像源 16、图像预处理器等预处理器（或预处理单元）18、通信接口（或通信单元）22。

图像源 16 可包括或可以为任意类型的用于捕获现实世界图像等的图像捕获设备，和/或任意类型的图像生成设备，例如用于生成计算机动画图像的计算机图形处理器或任意类型的用于获取和/或提供现实世界图像、计算机生成图像（例如，屏幕内容、虚拟现实（virtual reality, VR）图像和/或其任意组合（例如增强现实（augmented reality, AR）图像）的设备。所述图像源可以为存储上述图像中的任意图像的任意类型的内存或存储器。

为了区分预处理器（或预处理单元）18 执行的处理，图像（或图像数据）17 也可称为原始图像（或原始图像数据）17。

预处理器 18 用于接收（原始）图像数据 17，并对图像数据 17 进行预处理，得到预处理图像（预处理图像数据）19。例如，预处理器 18 执行的预处理可包括修剪、颜色格式转换（例如从 RGB 转换为 YCbCr）、调色或去噪。可以理解的是，预处理单元 18 可以为可选组件。

视频编码器（或编码器）20 用于接收预处理图像数据 19 并提供编码图像数据 21（下面将根据图 2 等进一步描述）。

源设备 12 中的通信接口 22 可用于：接收编码图像数据 21 并通过通信信道 13 向目的设备 14 等另一设备或任何其它设备发送编码图像数据 21（或其它任意处理后的版本），以便存储或直接重建。

目的设备 14 包括解码器 30，另外即可选地，可包括通信接口（或通信单元）28、后处理器（或后处理单元）32 和显示设备 34。

目的设备 14 中的通信接口 28 用于直接从源设备 12 或从存储设备等任意其它源设备

接收编码图像数据 21（或其它任意处理后的版本），例如，存储设备为编码图像数据存储设备，并将编码图像数据 21 提供给解码器 30。

通信接口 22 和通信接口 28 可用于通过源设备 12 与目的设备 14 之间的直连通信链路，例如直接有线或无线连接等，或者通过任意类型的网络，例如有线网络、无线网络或其任意组合、任意类型的私网和公网或其任意类型的组合，发送或接收编码图像数据（或编码数据）21。

例如，通信接口 22 可用于将编码图像数据 21 封装为报文等合适的格式，和/或使用任意类型的传输编码或处理来处理所述编码后的图像数据，以便在通信链路或通信网络上进行传输。

通信接口 28 与通信接口 22 对应，例如，可用于接收传输数据，并使用任意类型的对应传输解码或处理和/或解封装对传输数据进行处理，得到编码图像数据 21。

通信接口 22 和通信接口 28 均可配置为如图 1B 中从源设备 12 指向目的设备 14 的对应通信信道 13 的箭头所指示的单向通信接口，或双向通信接口，并且可用于发送和接收消息等，以建立连接，确认并交换与通信链路和/或例如编码后的图像数据传输等数据传输相关的任何其它信息，等等。

视频解码器（或解码器）30 用于接收编码图像数据 21 并提供解码图像（或解码图像数据）31（下面将根据图 3 等进一步描述）。

后处理器 32 用于对解码后的图像等解码图像数据 31（也称为重建后的图像数据）进行后处理，得到后处理后的图像等后处理图像数据 33。后处理单元 32 执行的后处理可以包括例如颜色格式转换（例如从 YCbCr 转换为 RGB）、调色、修剪或重采样，或者用于产生供显示设备 34 等显示的解码图像数据 31 等任何其它处理。

显示设备 34 用于接收后处理图像数据 33，以向用户或观看者等显示图像。显示设备 34 可以为或包括任意类型的用于表示重建后图像的显示器，例如，集成或外部显示屏或显示器。例如，显示屏可包括液晶显示器（liquid crystal display, LCD）、有机发光二极管（organic light emitting diode, OLED）显示器、等离子显示器、投影仪、微型 LED 显示器、硅基液晶显示器（liquid crystal on silicon, LCoS）、数字光处理器（digital light processor, DLP）或任意类型的其它显示屏。

译码系统 10 还包括训练引擎 25，训练引擎 25 用于训练编码器 20 或解码器 30 以对重构图像进行分层编码处理。

本申请实施例中训练数据包括：训练矩阵集合，该训练矩阵集合包括图像块的滤波

前亮度矩阵、量化步长矩阵和滤波后亮度矩阵，其中滤波前亮度矩阵中的对应位置的像素点对应于对应图像块中的对应位置的像素的滤波前的亮度值，量化步长矩阵中的对应位置的像素点对应于对应图像块中的对应位置的像素的亮度值对应的量化步长值，滤波后亮度矩阵中的对应位置的像素点对应于对应图像块中的对应位置的像素的滤波后的亮度值。

训练矩阵集合中的多个矩阵例如可以以图 6a 至 6c 所示的方式输入训练引擎 25。如图 6a 所示，将训练矩阵集合中的多个矩阵直接输入训练引擎 25，该多个矩阵均是二维矩阵。如图 6b 所示，选取训练矩阵集合中的多个矩阵的部分或全部做合并处理得到多维矩阵，再将该多维矩阵输入训练引擎 25。如图 6c 所示，选取训练矩阵集合中的多个矩阵的部分或全部做相加（或相乘）处理得到二维矩阵，再将该二维矩阵输入训练引擎 25。

上述训练数据可以存入数据库（未示意）中，训练引擎 25 基于训练数据训练得到目标模型（例如：可以是用于分层编解码的神经网络等）。需要说明的是，本申请实施例对于训练数据的来源不做限定，例如可以是云端或其他地方获取训练数据进行模型训练。

训练引擎 25 训练目标模型的过程使得滤波前像素逼近原始像素值。每个训练过程可以使用 64 个图像的小批量大小和 $1e-4$ 的初始学习率，遵循步长大小为 10。在训练数据可以通过编码器在不同 QP 量化参数设置下生成的数据。目标模型能够用于实现本申请实施例提供的分层编解码方法，即，将重构得到的图像或图像块通过相关预处理后输入该目标模型，可以得到滤波后的图像或图像块。本申请实施例中的目标模型具体可以为滤波网络，下文将结合图 7A-7D 详细说明目标模型。

训练引擎 25 训练得到的目标模型可以应用于译码系统 10 中，例如，应用于图 1B 所示的源设备 12（例如编码器 20）或目的设备 14（例如解码器 30）。训练引擎 25 可以在云端训练得到目标模型，然后译码系统 10 从云端下载并使用该目标模型；或者，训练引擎 25 可以在云端训练得到目标模型并使用该目标模型，译码系统 10 从云端直接获取处理结果。

尽管图 1B 示出了源设备 12 和目的设备 14 作为独立的设备，但设备实施例也可以同时包括源设备 12 和目的设备 14 或同时包括源设备 12 和目的设备 14 的功能，即同时包括源设备 12 或对应功能和目的设备 14 或对应功能。在这些实施例中，源设备 12 或对应功能和目的设备 14 或对应功能可以使用相同硬件和/或软件或通过单独的硬件和/或软件或其任意组合来实现。

根据描述，图 1B 所示的源设备 12 和/或目的设备 14 中的不同单元或功能的存在和（准确）划分可能根据实际设备和应用而有所不同。

编码器 20（例如视频编码器 20）或解码器 30（例如视频解码器 30）或两者都可通过如图 1C 所示的处理电路实现，例如一个或多个微处理器、数字信号处理器（digital signal processor, DSP）、专用集成电路（application-specific integrated circuit, ASIC）、现场可编程门阵列（field-programmable gate array, FPGA）、离散逻辑、硬件、视频编码专用处理器或其任意组合。编码器 20 可以通过处理电路 46 实现，以包含参照图 2 编码器 20 论述的各种模块和/或本文描述的任何其它编码器系统或子系统。解码器 30 可以通过处理电路 46 实现，以包含参照图 3 解码器 30 论述的各种模块和/或本文描述的任何其它解码器系统或子系统。所述处理电路 46 可用于执行下文论述的各种操作。如图 5 所示，如果部分技术在软件中实施，则设备可以将软件的指令存储在合适的非瞬时性计算机可读存储介质中，并且使用一个或多个处理器在硬件中执行指令，从而执行本发明技术。视频编码器 20 和视频解码器 30 中的其中一个可作为组合编解码器（encoder/decoder, CODEC）的一部分集成在单个设备中，如图 1C 所示。

源设备 12 和目的设备 14 可包括各种设备中的任一种，包括任意类型的手持设备或固定设备，例如，笔记本电脑或膝上型电脑、手机、智能手机、平板或平板电脑、相机、台式计算机、机顶盒、电视机、显示设备、数字媒体播放器、视频游戏控制台、视频流设备（例如，内容业务服务器或内容分发服务器）、广播接收设备、广播发射设备，等等，并可以不使用或使用任意类型的操作系统。在一些情况下，源设备 12 和目的设备 14 可配备用于无线通信的组件。因此，源设备 12 和目的设备 14 可以是无线通信设备。

在一些情况下，图 1B 所示的视频译码系统 10 仅仅是示例性的，本申请提供的技术可适用于视频编码设置（例如，视频编码或视频解码），这些设置不一定包括编码设备与解码设备之间的任何数据通信。在其它示例中，数据从本地存储器中检索，通过网络发送，等等。视频编码设备可以对数据进行编码并将数据存储到存储器中，和/或视频解码设备可以从存储器中检索数据并对数据进行解码。在一些示例中，编码和解码由相互不通信而只是编码数据到存储器和/或从存储器中检索并解码数据的设备来执行。

图 1C 是根据一示例性实施例的包含图 2 的视频编码器 20 和/或图 3 的视频解码器 30 的视频译码系统 40 的实例的说明图。视频译码系统 40 可以包含成像设备 41、视频编码器 20、视频解码器 30（和/或藉由处理电路 46 实施的视频编/解码器）、天线 42、一个或多个处理器 43、一个或多个内存存储器 44 和/或显示设备 45。

如图 1C 所示, 成像设备 41、天线 42、处理电路 46、视频编码器 20、视频解码器 30、处理器 43、内存存储器 44 和/或显示设备 45 能够互相通信。在不同实例中, 视频译码系统 40 可以只包含视频编码器 20 或只包含视频解码器 30。

在一些实例中, 天线 42 可以用于传输或接收视频数据的经编码比特流。另外, 在一些实例中, 显示设备 45 可以用于呈现视频数据。处理电路 46 可以包含专用集成电路(application-specific integrated circuit, ASIC) 逻辑、图形处理器、通用处理器等。视频译码系统 40 也可以包含可选的处理器 43, 该可选处理器 43 类似地可以包含专用集成电路(application-specific integrated circuit, ASIC) 逻辑、图形处理器、通用处理器等。另外, 内存存储器 44 可以是任何类型的存储器, 例如易失性存储器(例如, 静态随机存取存储器(static random access memory, SRAM)、动态随机存储器(dynamic random access memory, DRAM) 等) 或非易失性存储器(例如, 闪存等) 等。在非限制性实例中, 内存存储器 44 可以由超速缓存内存实施。在其它实例中, 处理电路 46 可以包含存储器(例如, 缓存等) 用于实施图像缓冲器等。

在一些实例中, 通过逻辑电路实施的视频编码器 20 可以包含(例如, 通过处理电路 46 或内存存储器 44 实施的) 图像缓冲器和(例如, 通过处理电路 46 实施的) 图形处理单元。图形处理单元可以通信耦合至图像缓冲器。图形处理单元可以包含通过处理电路 46 实施的视频编码器 20, 以实施参照图 2 和/或本文中所描述的任何其它编码器系统或子系统所论述的各种模块。逻辑电路可以用于执行本文所论述的各种操作。

在一些实例中, 视频解码器 30 可以以类似方式通过处理电路 46 实施, 以实施参照图 3 的视频解码器 30 和/或本文中所描述的任何其它解码器系统或子系统所论述的各种模块。在一些实例中, 逻辑电路实施的视频解码器 30 可以包含(通过处理电路 46 或内存存储器 44 实施的) 图像缓冲器和(例如, 通过处理电路 46 实施的) 图形处理单元。图形处理单元可以通信耦合至图像缓冲器。图形处理单元可以包含通过处理电路 46 实施的视频解码器 30, 以实施参照图 3 和/或本文中所描述的任何其它解码器系统或子系统所论述的各种模块。

在一些实例中, 天线 42 可以用于接收视频数据的经编码比特流。如所论述, 经编码比特流可以包含本文所论述的与编码视频帧相关的数据、指示符、索引值、模式选择数据等, 例如与编码分割相关的数据(例如, 变换系数或经量化变换系数, (如所论述的) 可选指示符, 和/或定义编码分割的数据)。视频译码系统 40 还可包含耦合至天线 42 并用于解码经编码比特流的视频解码器 30。显示设备 45 用于呈现视频帧。

应理解，本申请实施例中对于参考视频编码器 20 所描述的实例，视频解码器 30 可以用于执行相反过程。关于信令语法元素，视频解码器 30 可以用于接收并解析这种语法元素，相应地解码相关视频数据。在一些例子中，视频编码器 20 可以将语法元素熵编码成经编码视频比特流。在此类实例中，视频解码器 30 可以解析这种语法元素，并相应地解码相关视频数据。

为便于描述，参考通用视频编码(Versatile video coding, VVC)参考软件或由 ITU-T 视频编码专家组(Video Coding Experts Group, VCEG)和 ISO/IEC 运动图像专家组(Motion Picture Experts Group, MPEG)的视频编码联合工作组(Joint Collaboration Team on Video Coding, JCT-VC)开发的高性能视频编码(High-Efficiency Video Coding, HEVC)描述本发明实施例。本领域普通技术人员理解本发明实施例不限于 HEVC 或 VVC。

编码器和编码方法

图 2 为用于实现本申请技术的视频编码器 20 的示例的示意性框图。在图 2 的示例中，视频编码器 20 包括输入端(或输入接口) 201、残差计算单元 204、变换处理单元 206、量化单元 208、反量化单元 210、逆变换处理单元 212、重建单元 214、环路滤波器 220、解码图像缓冲器(decoded picture buffer, DPB) 230、模式选择单元 260、熵编码单元 270 和输出端(或输出接口) 272。模式选择单元 260 可包括帧间预测单元 244、帧内预测单元 254 和分割单元 262。帧间预测单元 244 可包括运动估计单元和运动补偿单元(未示出)。图 2 所示的视频编码器 20 也可称为混合型视频编码器或基于混合型视频编解码器的视频编码器。

残差计算单元 204、变换处理单元 206、量化单元 208 和模式选择单元 260 组成编码器 20 的前向信号路径，而反量化单元 210、逆变换处理单元 212、重建单元 214、缓冲器 216、环路滤波器 220、解码图像缓冲器(decoded picture buffer, DPB) 230、帧间预测单元 244 和帧内预测单元 254 组成编码器的后向信号路径，其中编码器 20 的后向信号路径对应于解码器的信号路径(参见图 3 中的解码器 30)。反量化单元 210、逆变换处理单元 212、重建单元 214、环路滤波器 220、解码图像缓冲器 230、帧间预测单元 244 和帧内预测单元 254 还组成视频编码器 20 的“内置解码器”。

量化

量化单元 208 用于通过例如标量量化或矢量量化对变换系数 207 进行量化，得到量化变换系数 209。量化变换系数 209 也可称为量化残差系数 209。

量化过程可减少与部分或全部变换系数 207 有关的位深度。例如，可在量化期间将 n

位变换系数向下舍入到 m 位变换系数, 其中 n 大于 m 。可通过调整量化参数(quantization parameter, QP) 修改量化程度。例如, 对于标量量化, 可以应用不同程度的比例来实现较细或较粗的量化。较小量化步长对应较细量化, 而较大量化步长对应较粗量化。可通过量化参数 (quantization parameter, QP) 指示合适的量化步长。例如, 量化参数可以为合适的量化步长的预定义集合的索引。例如, 较小的量化参数可对应精细量化 (较小量化步长), 较大的量化参数可对应粗糙量化 (较大量化步长), 反之亦然。量化可包括除以量化步长, 而反量化单元 210 等执行的对应或逆解量化可包括乘以量化步长。根据例如 HEVC 一些标准的实施例可用于使用量化参数来确定量化步长。一般而言, 可以根据量化参数使用包含除法的等式的定点近似来计算量化步长。可以引入其它比例缩放因子来进行量化和解量化, 以恢复可能由于在用于量化步长和量化参数的等式的定点近似中使用的比例而修改的残差块的范数。在一种示例性实现方式中, 可以合并逆变换和解量化的比例。或者, 可以使用自定义量化表并在比特流中等将其从编码器向解码器指示。量化是有损操作, 其中量化步长越大, 损耗越大。

在一个实施例中, 视频编码器 20 (对应地, 量化单元 208) 可用于输出量化参数 (quantization parameter, QP), 例如, 直接输出或由熵编码单元 270 进行编码或压缩后输出, 例如使得视频解码器 30 可接收并使用量化参数进行解码。

反量化

反量化单元 210 用于对量化系数执行量化单元 208 的反量化, 得到解量化系数 211, 例如, 根据或使用与量化单元 208 相同的量化步长执行与量化单元 208 所执行的量化方案的反量化方案。解量化系数 211 也可称为解量化残差系数 211, 对应于变换系数 207, 但是由于量化造成损耗, 反量化系数 211 通常与变换系数不完全相同。

重建

重建单元 214 (例如, 求和器 214) 用于将变换块 213 (即重建残差块 213) 添加到预测块 265, 以在像素域中得到重建块 215, 例如, 将重建残差块 213 的像素点值和预测块 265 的像素点值相加。

滤波

环路滤波器单元 220 (或简称“环路滤波器” 220) 用于对重建块 215 进行滤波, 得到滤波块 221, 或通常用于对重建像素点进行滤波以得到滤波像素点值。例如, 环路滤波器单元用于顺利进行像素转变或提高视频质量。环路滤波器单元 220 可包括一个或多个环路滤波器, 例如去块滤波器、像素点自适应偏移 (sample-adaptive offset, SAO) 滤

波器或一个或多个其它滤波器，例如自适应环路滤波器 (adaptive loop filter, ALF)、噪声抑制滤波器 (noise suppression filter, NSF) 或任意组合。例如，环路滤波器单元 220 可以包括去块滤波器、SAO 滤波器和 ALF 滤波器。滤波过程的顺序可以是去块滤波器、SAO 滤波器和 ALF 滤波器。再例如，增加一个称为具有色度缩放的亮度映射 (luma mapping with chroma scaling, LMCS) (即自适应环内整形器) 的过程。该过程在去块之前执行。再例如，去块滤波过程也可以应用于内部子块边缘，例如仿射子块边缘、ATMVP 子块边缘、子块变换 (sub-block transform, SBT) 边缘和内子部分 (intra sub-partition, ISP) 边缘。尽管环路滤波器单元 220 在图 2 中示为环路滤波器，但在其它配置中，环路滤波器单元 220 可以实现为环后滤波器。滤波块 221 也可称为滤波重建块 221。

在一个实施例中，视频编码器 20 (对应地，环路滤波器单元 220) 可用于输出环路滤波器参数 (例如 SAO 滤波参数、ALF 滤波参数或 LMCS 参数)，例如，直接输出或由熵编码单元 270 进行熵编码后输出，例如使得解码器 30 可接收并使用相同或不同的环路滤波器参数进行解码。

解码器和解码方法

图 3 示出了用于实现本申请技术的示例性视频解码器 30。视频解码器 30 用于接收例如由编码器 20 编码的编码图像数据 21 (例如编码比特流 21)，得到解码图像 331。编码图像数据或比特流包括用于解码所述编码图像数据的信息，例如表示编码视频片 (和/或编码区块组或编码区块) 的图像块的数据和相关的语法元素。

在图 3 的示例中，解码器 30 包括熵解码单元 304、反量化单元 310、逆变换处理单元 312、重建单元 314 (例如求和器 314)、环路滤波器 320、解码图像缓冲器 (DBP) 330、模式应用单元 360、帧间预测单元 344 和帧内预测单元 354。帧间预测单元 344 可以为或包括运动补偿单元。在一些示例中，视频解码器 30 可执行大体上与参照图 2 的视频编码器 100 描述的编码过程相反的解码过程。

如编码器 20 所述，反量化单元 210、逆变换处理单元 212、重建单元 214、环路滤波器 220、解码图像缓冲器 DPB230、帧间预测单元 344 和帧内预测单元 354 还组成视频编码器 20 的“内置解码器”。相应地，反量化单元 310 在功能上可与反量化单元 110 相同，逆变换处理单元 312 在功能上可与逆变换处理单元 122 相同，重建单元 314 在功能上可与重建单元 214 相同，环路滤波器 320 在功能上可与环路滤波器 220 相同，解码图像缓冲器 330 在功能上可与解码图像缓冲器 230 相同。因此，视频编码器 20 的相应单元和功能的解释相应地适用于视频解码器 30 的相应单元和功能。

反量化

反量化单元 310 可用于从编码图像数据 21 (例如通过熵解码单元 304 解析和/或解码) 接收量化参数 (quantization parameter, QP) (或一般为与反量化相关的信息) 和量化系数, 并基于所述量化参数对所述解码的量化系数 309 进行反量化以获得反量化系数 311, 所述反量化系数 311 也可以称为变换系数 311。反量化过程可包括使用视频编码器 20 为视频片中的每个视频块计算的量化参数来确定量化程度, 同样也确定需要执行的反量化的程度。

重建

重建单元 314 (例如, 求和器 314) 用于将重建残差块 313 添加到预测块 365, 以在像素域中得到重建块 315, 例如, 将重建残差块 313 的像素点值和预测块 365 的像素点值相加。

滤波

环路滤波器单元 320 (在编码环路中或之后) 用于对重建块 315 进行滤波, 得到滤波块 321, 从而顺利进行像素转变或提高视频质量等。环路滤波器单元 320 可包括一个或多个环路滤波器, 例如去块滤波器、像素点自适应偏移 (sample-adaptive offset, SAO) 滤波器或一个或多个其它滤波器, 例如自适应环路滤波器 (adaptive loop filter, ALF)、噪声抑制滤波器 (noise suppression filter, NSF) 或任意组合。例如, 环路滤波器单元 220 可以包括去块滤波器、SAO 滤波器和 ALF 滤波器。滤波过程的顺序可以是去块滤波器、SAO 滤波器和 ALF 滤波器。再例如, 增加一个称为具有色度缩放的亮度映射 (luma mapping with chroma scaling, LMCS) (即自适应环内整形器) 的过程。该过程在去块之前执行。再例如, 去块滤波过程也可以应用于内部子块边缘, 例如仿射子块边缘、ATMVP 子块边缘、子块变换 (sub-block transform, SBT) 边缘和内子部分 (intra sub-partition, ISP) 边缘。尽管环路滤波器单元 320 在图 3 中示为环路滤波器, 但在其它配置中, 环路滤波器单元 320 可以实现为环后滤波器。

解码器 30 用于通过输出端 312 等输出解码图像 311, 向用户显示或供用户查看。

尽管上述实施例主要描述了视频编解码, 但应注意的是, 译码系统 10、编码器 20 和解码器 30 的实施例以及本文描述的其它实施例也可以用于静止图像处理或编解码, 即视频编解码中独立于任何先前或连续图像的单个图像的处理或编解码。一般情况下, 如果图像处理仅限于单个图像 17, 帧间预测单元 244 (编码器) 和帧间预测单元 344 (解码器) 可能不可用。视频编码器 20 和视频解码器 30 的所有其它功能 (也称为工具或技术)

同样可用于静态图像处理,例如残差计算 204/304、变换 206、量化 208、反量化 210/310、(逆)变换 212/312、分割 262/362、帧内预测 254/354 和/或环路滤波 220/320、熵编码 270 和熵解码 304。

图 4 为本发明实施例提供的视频译码设备 400 的示意图。视频译码设备 400 适用于实现本文描述的公开实施例。在一个实施例中,视频译码设备 400 可以是解码器,例如图 1B 中的视频解码器 30,也可以是编码器,例如图 1B 中的视频编码器 20。

视频译码设备 400 包括:用于接收数据的入端口 410 (或输入端口 410) 和接收单元 (receiver unit, Rx) 420; 用于处理数据的处理器、逻辑单元或中央处理器 (central processing unit, CPU) 430; 例如,这里的处理器 430 可以是神经网络处理器 430; 用于传输数据的发送单元 (transmitter unit, Tx) 440 和出端口 450 (或输出端口 450); 用于存储数据的存储器 460。视频译码设备 400 还可包括耦合到入端口 410、接收单元 420、发送单元 440 和出端口 450 的光电 (optical-to-electrical, OE) 组件和光电 (electrical-to-optical, EO) 组件,用于光信号或电信号的出口或入口。

处理器 430 通过硬件和软件实现。处理器 430 可实现为一个或多个处理器芯片、核 (例如,多核处理器)、FPGA、ASIC 和 DSP。处理器 430 与入端口 410、接收单元 420、发送单元 440、出端口 450 和存储器 460 通信。处理器 430 包括译码模块 470 (例如,基于神经网络 (neural networks, NN) 的译码模块 470)。译码模块 470 实施上文所公开的实施例。例如,译码模块 470 执行、处理、准备或提供各种编码操作。因此,通过译码模块 470 为视频译码设备 400 的功能提供了实质性的改进,并且影响了视频译码设备 400 到不同状态的切换。或者,以存储在存储器 460 中并由处理器 430 执行的指令来实现译码模块 470。

存储器 460 包括一个或多个磁盘、磁带机和固态硬盘,可以用作溢出数据存储设备,用于在选择执行程序时存储此类程序,并且存储在程序执行过程中读取的指令和数据。存储器 460 可以是易失性和/或非易失性的,可以是只读存储器 (read-only memory, ROM)、随机存取存储器 (random access memory, RAM)、三态内容寻址存储器 (ternary content-addressable memory, TCAM) 和/或静态随机存取存储器 (static random-access memory, SRAM)。

图 5 为示例性实施例提供的装置 500 的简化框图,装置 500 可用作图 1B 中的源设备 12 和目的设备 14 中的任一个或两个。

装置 500 中的处理器 502 可以是中央处理器。或者,处理器 502 可以是现有的或今

后将研发出的能够操控或处理信息的任何其它类型设备或多个设备。虽然可以使用如图所示的处理器 502 等单个处理器来实施已公开的实现方式，但使用一个以上的处理器速度更快和效率更高。

在一种实现方式中，装置 500 中的存储器 504 可以是只读存储器 (ROM) 设备或随机存取存储器 (RAM) 设备。任何其它合适类型的存储设备都可以用作存储器 504。存储器 504 可以包括处理器 502 通过总线 512 访问的代码和数据 506。存储器 504 还可包括操作系统 508 和应用程序 510，应用程序 510 包括允许处理器 502 执行本文所述方法的至少一个程序。例如，应用程序 510 可以包括应用 1 至 N，还包括执行本文所述方法的视频译码应用。

装置 500 还可以包括一个或多个输出设备，例如显示器 518。在一个示例中，显示器 518 可以是将显示器与可用于感测触摸输入的触敏元件组合的触敏显示器。显示器 518 可以通过总线 512 耦合到处理器 502。

虽然装置 500 中的总线 512 在本文中描述为单个总线，但是总线 512 可以包括多个总线。此外，辅助存储器可以直接耦合到装置 500 的其它组件或通过网络访问，并且可以包括存储卡等单个集成单元或多个存储卡等多个单元。因此，装置 500 可以具有各种各样的配置。

本发明实施例涉及一种面向色彩分量码率分配的 AI 图像有损编码方案，可应用于图 1A 至图 5 所描述的视频译码系统、编码器和解码器。

需要说明的是，本申请提供的方法主要用于 YUV 分量码率分配过程，此过程主要由编码端控制，为使得解码端更适配，也可在解码器中增加相应控制单元。

视频图像信号通常包括一个亮度分量与两个色度分量。亮度分量通常使用符号 Y 表示，色度分量通常使用符号 U、V 表示。如图 7 (a) 至 (c) 所示，常用的 YUV 格式包括如下格式，图 7 中又表示亮度分量采样点，圈表示每色度分量采样点：

4:4:4 格式：表示色度分量没有下采样；

4:2:2 格式：表示色度分量相对于亮度分量进行 2:1 的水平下采样，没有竖直下采样。对于每两个 U 采样点或 V 采样点，每个扫描行都包含四个 Y 采样点；

4:2:0 格式：表示色度分量相对于亮度分量进行 2:1 的水平下采样与 2:1 的竖直下采样。

在视频图像采用 YUV 4:2:0 格式的情况下，若图像块的亮度分量为 $2N \times 2N$ 大小的图像块，则图像块的色度分量为 $N \times N$ 大小的图像块。本发明实施例将以 4:2:0 格式为例解释本发明

方案技术。但可以理解的，本发明技术方案除了用于YUV4:2:0格式外，还可以用于YUV其它格式，或者其它视频图像格式中不同分量间的相互预测，如RGB格式等。另一方面，当前块可以是方形块，也可以是非方形的矩形块或其它形状区域，本发明实施例提供的技术方案同样适用。

为方便叙述，本发明实施例采用第一信号分量和第二信号分量的表述。若图像信号包括亮度信号分量与色度信号分量，则第一信号分量可以是色度分量，第二信号分量可以是亮度分量；若图像信号包括R、G、B三个信号分量，则第一信号分量可以是R、G、B三个信号分量中任意一个信号分量，第二信号分量可以是与第一信号分量不同的R、G、B三个信号分量中的一个信号分量；若按照其它方式将图像信号分解为多个信号分量，则可采用类似的方法规定第一信号分量与第二信号分量。

如下面具体描述，本申请实施例可以输入质量因子到码率控制模块（或称为码率分配控制模块），该模块生成各分量特征图的控制信号；各分量的控制信号的控制向量与对应的特征图相乘，得到量化后即待编码的特征值。

如图7A所示，为本发明实施例提供的面向YUV码率分配的AI图像译码系统700，可用于视频图像编码器和解码器。如图7A所示，图像译码系统700包括编码端的第一信号分量处理模块（如Y分量处理模块）、第二信号分量处理模块（如UV分量处理模块）、码率分配控制模块、熵编码模块；以及解码端的第一信号分量处理模块（Y分量处理模块2）、第二信号分量处理模块（如UV分量处理模块2）、熵解码模块。图像译码系统700可选地包含联合处理模块、联合处理模块2、以及质量响应模块（也可称为码率分配控制模块2、或者码率控制模块2）。图像译码系统700中Y分量质量因子、UV分量质量因子输入码率分配控制模块，该模块输出控制信号分别作用于Y分量处理模块的Y分量特征图、UV分量处理模块的UV分量特征图（可以称为第一特征图），输出各信号分量的第二特征图，从而实现Y、UV的码率分配。随后，根据各信号分量的第二特征图，获得视频信号的码流。比如，Y分量处理模块、UV分量处理模块输出的特征图直接级联在一起，或Y分量处理模块、UV分量处理模块输出的特征图直接相加，形成编码器最终输出的特征图，对该特征图进行熵编码。可选的，Y分量处理模块、UV分量处理模块输出的特征图输入联合处理模块，得到编码器最终输出的特征图，对该特征图进行熵编码。

在图7A所示的架构中，图7B为一种编码方法的实施例。步骤701，根据第一信号分量的质量因子，获得所述第一信号分量的控制信号。步骤702，根据第二信号分量的质量因子，获得所述第二信号分量的控制信号。比如，图7B所示的实施例可以根据所述第一

信号分量的质量因子从 N 个候选第一控制信号中, 获得所述第一信号分量的控制信号, 其中 N 为大于 1 的整数; 以及根据所述第二信号分量的质量因子从 M 个候选第二控制信号中, 获得所述第二信号分量的控制信号, 其中 M 为大于 1 的整数。 N 和 M 可以相等或者不等, 本申请对此不作限定。

步骤 703, 把所述第一信号分量的控制信号作用于所述第一信号分量的第一特征图, 获得所述第一信号分量的第二特征图。步骤 704, 把所述第二信号分量的控制信号作用于所述第二信号分量的第一特征图, 获得所述第二信号分量的第二特征图。

比如, 在一种实施例中所述控制信号由网络学习产生, 作用于 Y 分量处理模块、 UV 分量处理模块中各模块中至少一层网络的特征图 (可以称为第一特征图), 然后输出第二特征图。例如, 作用于最后一层网络的输出。

步骤 705, 根据所述第一信号分量的第二特征图和所述第二信号分量的第二特征图, 获得所述视频信号的码流。

由于控制信号可以作用于 Y 分量处理模块、 UV 分量处理模块中各模块中任何一层网络的第一特征图, 因此输出第二特征图之后, 可以对第二特征图继续进行神经网络的处理。则相应地, 根据 Y 分量的第二特征图和 UV 分量的第二特征图, 获得所述视频信号的码流包括:

对所述第一信号分量的第二特征图和所述第二信号分量的第二特征图, 进行熵编码, 以获得所述视频信号的码流; 或者,

对所述第一信号分量的第二特征图和经神经网络处理的所述第二信号分量的第二特征图, 进行熵编码, 以获得所述视频信号的码流; 或者,

对经神经网络处理的所述第一信号分量的第二特征图和所述第二信号分量的第二特征图, 进行熵编码, 以获得所述视频信号的码流; 或者,

对经神经网络处理的所述第一信号分量的第二特征图和经神经网络处理的所述第二信号分量的第二特征图, 进行熵编码, 以获得所述视频信号的码流。

当图像译码系统 700 包含联合处理模块, 根据所述第一信号分量的第二特征图和所述第二信号分量的第二特征图, 获得所述视频信号的码流包括:

对所述第一信号分量的第二特征图和所述第二信号分量的第二特征图, 进行联合处理, 获得联合的特征图, 对所述联合的特征图进行熵编码, 以获得所述视频信号的码流; 或者,

对所述第一信号分量的第二特征图和经神经网络处理的所述第二信号分量的第二特

征图，进行联合处理，获得联合的特征图，对所述联合的特征图进行熵编码，以获得所述视频信号的码流；或者，

对神经网络处理的所述第一信号分量的第二特征图和所述第二信号分量的第二特征图，进行联合处理，获得联合的特征图，对所述联合的特征图进行熵编码，以获得所述视频信号的码流；或者，

对神经网络处理的所述第一信号分量的第二特征图和神经网络处理的所述第二信号分量的第二特征图，进行联合处理，获得联合的特征图，对所述联合的特征图进行熵编码，以获得所述视频信号的码流。

具体说明如下：

当控制信号包括 Y、UV 特征图的控制向量：码率分配控制模块学习生成所述第一信号分量的 N 个候选第一控制信号（如控制向量矩阵 $\{q_{y1}, q_{y2}, \dots, q_{yi} \dots q_{yN}\}$ ）、所述第二信号分量的 M 个候选第二控制信号（如控制向量矩阵 $\{q_{uv1}, q_{uv2}, \dots, q_{uvj}, \dots, q_{uvM}\}$ ），使用时根据所述 Y 分量的质量因子的索引 i 得到所述第一信号分量的控制信号 q_{yi} ，根据所述 UV 分量的质量因子的索引 j 得到所述第二信号分量的控制信号 q_{uvj} ，其中 N 和 M 为大于 1 的整数。然后，控制向量通过与对应的特征图逐通道相乘实现不同分量特征图的控制。此时发送给解码端的视频信号的码流中包括所述 Y 分量的质量因子的索引 i 和所述 UV 分量的质量因子的索引 j。

当控制信号包括 Y、UV 特征图的控制向量及偏移向量：如前述方法，根据所述 Y 分量的质量因子的索引 i 得到所述第一信号分量的控制向量 q_{yi} 、偏移向量 b_{yi} ；根据所述 UV 分量的质量因子的索引 j 得到所述第二信号分量的控制向量 q_{uvj} 、偏移向量 b_{uvj} 。然后，控制向量通过与对应的特征图逐通道相乘，然后加上对应的偏移向量来实现不同分量特征图的控制。

在另一种实施例中，Y 分量的控制信号和 UV 分量的控制信号作为二元组，码率分配控制模块学习生成视频信号的 N 个候选控制信号（如控制向量矩阵 $\{q_{c1}, q_{c2}, \dots, q_{ci} \dots q_{cN}\}$ ），此时，c 为 2，每个控制向量 q_{ci} 中既包括第一信号分量的控制向量，也包括第二信号分量的控制分量。然后通过视频信号的质量因子的索引 i 得到包含所述第一信号分量和所述第二信号分量的控制信号 q_{ci} 。偏移向量类似实现，即视频信号的每个偏移向量中既包括第一信号分量的偏移分量，也包括第二信号分量的偏移分量。此时发送给解码端的视频信号的码流中包括所述视频信号的质量因子的索引 i。

在再一种实施例中，Y 分量质量因子、UV 分量质量因子作为全连接网络的输入，输

出控制向量 $\mathbf{q}_{yi}$ 、控制向量 $\mathbf{q}_{uvj}$ ，通过控制向量与对应的特征图逐通道相乘，实现不同分量特征图的控制。Y分量质量因子、UV分量质量因子作为全连接网络的输入，还可以输出偏移向量 $\mathbf{b}_{yi}$ 、偏移向量 $\mathbf{b}_{uvj}$ ，然后通过控制向量与对应的特征图逐通道相乘，偏移向量与对应的特征图逐通道相加，实现不同分量特征图的控制。此时发送给解码端的视频信号的码流中包括所述Y分量的质量因子和所述UV分量的质量因子。

解码端对接收到的码流进行熵解码得到特征图，特征图分解为Y分量特征图、UV分量特征图。可选的，熵解码得到特征图先输入联合处理子模块2，得到Y分量特征图、UV分量特征图。

Y分量特征图、UV分量特征图分别输入Y分量处理模块2、UV分量处理模块2输出Y分量重建图、UV分量重建图。可选的，Y分量质量因子、UV分量质量因子输入质量响应模块，该模块输出响应信号分别作用于Y分量处理模块的Y分量特征图、UV分量处理模块的UV分量特征图，从而实现Y、UV分量的自适应质量响应。质量响应也可以称为质量控制，只是为了与编码端的质量控制进行区分，在解码端称为质量响应。

具体来说，以图7C为例，步骤711，解码端从编码端获得视频信号的码流，对所述码流进行熵解码以获得所述视频信号的第一信号分量（如Y分量）的特征图和所述视频信号的第二信号分量（如UV分量）的特征图。

解码端还会从码流中获得所述第一信号分量的质量因子信息和所述第二信号分量的质量因子信息，其中所述第一信号分量的质量因子信息为所述第一信号分量的质量因子或者所述第一信号分量的质量因子的索引，所述第二信号分量的质量因子信息为所述第二信号分量的质量因子或者所述第二信号分量的质量因子的索引。然后通过所述第一信号分量的质量因子信息，获得所述第一信号分量的响应信号；通过所述第二信号分量的质量因子信息，获得所述第二信号分量的响应信号。当所述第一信号分量的质量因子信息为所述第一信号分量的质量因子时，所述第一信号分量的质量因子取值为N个中的一个；当所述第一信号分量的质量因子信息为所述第一信号分量的质量因子的索引时，所述第一信号分量的质量因子的索引的取值范围为0至N-1或者1至N，其中N为大于1的整数。类似的，当所述第二信号分量的质量因子信息为所述第二信号分量的质量因子时，所述第二信号分量的质量因子取值为M个中的一个；当所述第二信号分量的质量因子信息为所述第二信号分量的质量因子的索引时，所述第二信号分量的质量因子的索引的取值范围为0至M-1或者1至M，其中M为大于1的整数。

当编码端传递过来的为联合的特征图时，解码端还需要对所述联合的特征图进行熵

解码,并经神经网络处理获得所述第一信号分量的特征图和所述第二信号分量的特征图。

步骤 712,通过所述第一信号分量的质量因子信息,获得所述第一信号分量的响应信号。步骤 713,通过所述第二信号分量的质量因子信息,获得所述第二信号分量的响应信号。

在一种实施例中,若所述码流中包括所述 Y 分量的质量因子的索引 i 和所述 UV 分量的质量因子的索引 j,则解码端需要通过学习生成所述第一信号分量的响应信号矩阵 $\{g_{y1}, g_{y2}, \dots, g_{yi} \dots g_{yN}\}$ 、和所述第二信号分量的响应信号矩阵 $\{g_{uv1}, g_{uv2}, \dots, g_{uvj}, \dots, g_{uvM}\}$,其中 N 和 M 为大于 1 的整数;可选地,通过对编码端第一信号分量的控制信号矩阵 $\{q_{y1}, q_{y2}, \dots, q_{yi} \dots q_{yN}\}$ 取倒数得到所述第一信号分量的响应信号矩阵 $\{g_{y1}, g_{y2}, \dots, g_{yi} \dots g_{yN}\}$ 、编码端第一信号分量的控制信号矩阵 $\{q_{uv1}, q_{uv2}, \dots, q_{uvi} \dots q_{uvN}\}$ 取倒数得到和所述第二信号分量的响应信号矩阵 $\{g_{uv1}, g_{uv2}, \dots, g_{uvj}, \dots, g_{uvM}\}$,其中 N 和 M 为大于 1 的整数;根据所述 Y 分量的质量因子的索引 i 得到所述第一信号分量的响应信号 g_{yi} ;根据所述 UV 分量的质量因子的索引 j 得到所述第二信号分量的响应信号 g_{uvj} 。

在另一种实施例中,若所述码流中包括所述视频信号的质量因子的索引 i,则解码端需要通过学习生成所述视频信号的响应信号矩阵 $\{g_{c1}, g_{c2}, \dots, g_{ci} \dots g_{cN}\}$,其中 c 为 2 代表 Y 分量和 UV 分量, N 为大于 1 的整数;可选地,通过对编码端视频信号的控制信号矩阵 $\{q_{c1}, q_{c2}, \dots, q_{ci} \dots q_{cN}\}$ 取倒数,得到所述视频信号的响应信号矩阵 $\{g_{c1}, g_{c2}, \dots, g_{ci} \dots g_{cN}\}$,其中 c 为 2 代表 Y 分量和 UV 分量, N 为大于 1 的整数;根据所述视频信号的质量因子的索引 i 得到包含所述第一信号分量和所述第二信号分量的响应信号 g_{ci} 。

在再一种实施例中,若所述码流中包括所述第一信号分量的质量因子和第二信号分量的质量因子,则所述解码端将所述 Y 分量的质量因子作为全连接网络的输入,输出所述 Y 分量的响应信号;将所述 UV 分量的质量因子作为全连接网络的输入,输出所述 UV 分量的响应信号。

步骤 714,根据所述第一信号分量的响应信号和所述第一信号分量的特征图,获得所述第一信号分量的重建图。步骤 715,根据所述第二信号分量的响应信号和所述第二信号分量的特征图,获得所述第二信号分量的重建图。

当所述响应信号包括响应向量,则根据所述第一信号分量的响应信号和所述第一信号分量的特征图,获得所述第一信号分量的重建图包括:

将所述第一信号分量的响应向量和所述第一信号分量的特征图相乘,获得所述第一信号分量的重建图;或将所述第一信号分量的响应向量和所述第一信号分量的特征图相

乘后并且再经神经网络处理，获得所述第一信号分量的重建图。

则根据所述第二信号分量的响应信号和所述第二信号分量的特征图，获得所述第二信号分量的重建图包括：

将所述第二信号分量的响应向量和所述第二信号分量的特征图相乘，获得所述第二信号分量的重建图；和将所述第二信号分量的响应向量和所述第二信号分量的特征图相乘后并且再经神经网络处理，获得所述第二信号分量的重建图。

当所述响应信号包括响应向量和偏移向量，则根据所述第一信号分量的响应信号和所述第一信号分量的特征图，获得所述第一信号分量的重建图包括：

将所述第一信号分量的响应向量与所述第一信号分量的特征图相乘，再与所述第一信号分量的偏移向量相加，获得所述第一信号分量的重建图；或将所述第一信号分量的响应向量与所述第一信号分量的特征图相乘，再与所述第一信号分量的偏移向量相加后并且再经神经网络处理，获得所述第一信号分量的重建图。

则根据所述第二信号分量的响应信号和所述第二信号分量的特征图，获得所述第二信号分量的重建图包括：

将所述第二信号分量的响应向量与所述第二信号分量的特征图相乘，再与所述第二信号分量的偏移向量相加，获得所述第二信号分量的重建图；或将所述第二信号分量的响应向量与所述第二信号分量的特征图相乘，再与所述第二信号分量的偏移向量相加，并且再经神经网络处理，获得所述第二信号分量的重建图。

步骤 716 根据所述第一信号分量的重建图和所述第二信号分量的重建图，重建所述视频信号。

如图7D所示，为本发明实施例提供的面向YUV码率分配的AI图像译码系统710，可用于视频图像编码器和解码器。如图7D所示，图像译码系统710包括编码端的第一信号分量处理模块（如Y分量处理模块）、第二信号分量处理模块（如U分量处理模块）、第三信号分量处理模块（如V分量处理模块）、码率分配控制模块、熵编码模块；以及解码端的第一信号分量处理模块（Y分量处理模块2）、第二信号分量处理模块（如U分量处理模块2）、第三信号分量处理模块（如V分量处理模块2）、熵解码模块。图像译码系统710可选地包含联合处理模块、联合处理模块2、以及质量响应模块（也可称为码率分配控制模块2、或者码率控制模块2）。图像译码系统710中Y分量质量因子、U分量质量因子以及V分量质量因子输入码率分配控制模块，该模块输出控制信号分别作用于Y分量处理模块的Y分量特征图、U分量处理模块的U分量特征图、以及V分量处理模块的V分量特征图（可以称为

第一特征图)，输出各信号分量的第二特征图，从而实现Y、U、V的码率分配。随后，根据各信号分量的第二特征图，获得视频信号的码流。比如，Y分量处理模块、U分量处理模块、V分量处理模块输出的特征图直接级联在一起，或Y分量处理模块、U分量处理模块、V分量处理模块输出的特征图直接相加，形成编码器最终输出的特征图，对该特征图进行熵编码。可选的，Y分量处理模块、U分量处理模块和V分量处理模块输出的特征图输入联合处理模块，得到编码器最终输出的特征图，对该特征图进行熵编码。

图 7D 所示实施例可以根据所述第一信号分量(Y 分量)的质量因子从 N 个候选第一控制信号中，获得所述第一信号分量的控制信号；根据所述第二信号分量(U 分量)的质量因子从 M 个候选第二控制信号中，获得所述第二信号分量的控制信号；根据所述第三信号分量(V 分量)的质量因子从 L 个候选第三控制信号中，获得所述第三信号分量的控制信号。N、M 和 L 为大于 1 的整数，可以相等或者不等，本申请对此不作限定。

在图 7D 所示的架构中，编解码方法与图 7B 和 7C 类似，说明如下：

比如，在一种实施例中所述控制信号由网络学习产生，作用于 Y 分量处理模块、U 分量处理模块、V 分量处理模块中各模块中至少一层网络的特征图(可以称为第一特征图)，然后输出第二特征图。例如，作用于最后一层网络的输出。由于控制信号可以作用于 Y 分量处理模块、U 分量处理模块中、V 分量处理模块中各模块中任何一层网络的第一特征图，因此输出第二特征图之后，可以对第二特征图继续进行神经网络的处理。则相应地，根据 Y 分量的第二特征图、U 分量的第二特征图和 V 分量的第二特征图，获得所述视频信号的码流包括：

对所述第一信号分量的第二特征图、所述第二信号分量的第二特征图和所述第三信号分量的第二特征图，进行熵编码，以获得所述视频信号的码流；或者，

对所述第一信号分量的第二特征图、经神经网络处理的所述第二信号分量的第二特征图，和所述第三信号分量的第二特征图，进行熵编码，以获得所述视频信号的码流；或者，

对所述第一信号分量的第二特征图、经神经网络处理的所述第二信号分量的第二特征图，和经神经网络处理的所述第三信号分量的第二特征图，进行熵编码，以获得所述视频信号的码流；或者，

对所述第一信号分量的第二特征图、所述第二信号分量的第二特征图，和经神经网络处理的所述第三信号分量的第二特征图，进行熵编码，以获得所述视频信号的码流；或者，

对神经网络处理的所述第一信号分量的第二特征图、所述第二信号分量的第二特征图和所述第三信号分量的第二特征图，进行熵编码，以获得所述视频信号的码流；或者，

对神经网络处理的所述第一信号分量的第二特征图、经神经网络处理的所述第二信号分量的第二特征图，和所述第三信号分量的第二特征图，进行熵编码，以获得所述视频信号的码流；或者

对神经网络处理的所述第一信号分量的第二特征图、经神经网络处理的所述第二信号分量的第二特征图，和经神经网络处理的所述第三信号分量的第二特征图，进行熵编码，以获得所述视频信号的码流；或者，

对神经网络处理的所述第一信号分量的第二特征图、所述第二信号分量的第二特征图，和经神经网络处理的所述第三信号分量的第二特征图，进行熵编码，以获得所述视频信号的码流。

当图像译码系统 710 包含联合处理模块时，还可以对前述获得的第二特征图或者经过处理过的特征图以及组合，进行联合处理，获得联合的特征图，对所述联合的特征图进行熵编码，以获得所述视频信号的码流。

具体说明如下：

当控制信号包括 Y、U、V 特征图的控制向量：码率分配控制模块学习生成所述第一信号分量的 N 个候选第一控制信号（如控制向量矩阵 $\{q_{y1}, q_{y2}, \dots, q_{yi} \dots q_{yN}\}$ ）、所述第二信号分量的 M 个候选第二控制信号（如控制向量矩阵 $\{q_{u1}, q_{u2}, \dots, q_{uj} \dots, q_{uM}\}$ ）和所述第三信号分量的 L 个候选第三控制信号（如控制向量矩阵 $\{q_{v1}, q_{v2}, \dots, q_{vk} \dots q_{vL}\}$ ），使用时根据所述 Y 分量质量因子的索引 i 得到所述第一信号分量的控制信号 q_{yi} ；根据所述 U 分量质量因子的索引 j 得到所述第二信号分量的控制信号 q_{uj} ；以及根据所述 V 分量质量因子的索引 k 得到所述第三信号分量的控制信号 q_{vk} 。然后，控制向量通过与对应的特征图逐通道相乘实现不同分量特征图的控制。此时发送给解码端的视频信号的码流中包括所述 Y 分量的质量因子的索引 i、所述 U 分量的质量因子的索引 j、和所述 V 分量质量因子的索引 k。

当控制信号包括 Y、U、V 特征图的控制向量及偏移向量：如前述方法，使用时使用时根据所述 Y 分量质量因子的索引 i 得到所述第一信号分量的控制向量 q_{yi} 、偏移向量 b_{yi} ；根据所述 U 分量质量因子的索引 j 得到所述第二信号分量的控制向量 q_{uj} 、偏移向量 b_{uj} ；以及根据所述 V 分量质量因子的索引 k 得到所述第三信号分量的控制向量 q_{vk} 、偏移向量

b_{vi} 。然后，控制向量通过与对应的特征图逐通道相乘，然后加上对应的偏移向量来实现不同分量特征图的控制。

在另一种实施例中，Y分量的控制信号、U分量、V分量的控制信号作为三元组，码率分配控制模块学习生成视频信号的N个候选控制信号（如控制向量矩阵 $\{q_{c1}, q_{c2}, \dots, q_{ci} \dots q_{cN}\}$ ），此时，c为3，每个控制向量 q_{ci} 中既包括第一信号分量的控制分量，也包括第二信号分量的控制向量，还包括第三信号分量的控制向量。然后通过视频信号的质量因子的索引i得到包含所述第一信号分量、所述第二信号分量和第三信号分量的控制信号 q_{ci} 。偏移向量类似实现，即视频信号的每个偏移向量中既包括第一信号分量的偏移分量，也包括第二信号分量的偏移分量，还包括第三信号分量的偏移分量。此时发送给解码端的视频信号的码流中包括所述视频信号的质量因子的索引i

在再一种实施例中，Y分量质量因子、U分量质量因子、V分量质量因子作为全连接网络的输入，输出控制向量 q_{yi} 、控制向量 q_{ui} 、和控制向量 q_{vk} ，通过控制向量与对应的特征图逐通道相乘，实现不同分量特征图的控制。Y分量质量因子、U分量质量因子、V分量质量因子作为全连接网络的输入，还可以输出偏移向量 b_{yi} 、偏移向量 b_{ui} 、和偏移向量 b_{vk} ，然后通过控制向量与对应的特征图逐通道相乘，偏移向量与对应的特征图逐通道相加，实现不同分量特征图的控制。此时发送给解码端的视频信号的码流中包括所述视频信号的码流中包括所述Y分量的质量因子、所述U分量的质量因子和所述V分量的质量因子。

解码端对接收到的码流进行熵解码得到特征图，特征图分解为Y分量特征图、U分量特征图、和V分量特征图。可选的，熵解码得到特征图先输入联合处理子模块2，得到Y分量特征图、U分量特征图、V分量特征图。

Y分量特征图、U分量特征图、V分量特征图分别输入Y分量处理模块2、U分量特征图2、V分量处理模块2输出Y分量重建图、U分量重建图、V分量重建图。可选的，Y分量质量因子、U分量质量因子、V分量质量因子输入质量响应模块，该模块输出控制信号分别作用于Y分量处理模块的Y分量特征图、U分量处理模块的U分量特征图、V分量处理模块的V分量特征图，从而实现Y、U、V分量的自适应质量响应。

所述响应信号产生方式与控制信号类似，只是便于区分，编码端称为控制信号，解码端称为响应信号。

具体来说，解码端从编码端获得视频信号的码流；对所述码流进行熵解码以获得所述视频信号的第一信号分量（如Y分量）的特征图、所述视频信号的第二信号分量（如U

分量)的特征图、和所述视频信号的第二信号分量(如V分量)的特征图;根据所述第一信号分量的响应信号和所述第一信号分量的特征图,获得所述第一信号分量的重建图,其中所述第一信号分量的响应信号通过学习获得;根据所述第二信号分量的响应信号和所述第二信号分量的特征图,获得所述第二信号分量的重建图,其中所述第二信号分量的响应信号通过学习获得;根据所述第三信号分量的响应信号和所述第三信号分量的特征图,获得所述第三信号分量的重建图,其中所述第三信号分量的响应信号通过学习获得,以及根据所述第一信号分量的重建图、所述第二信号分量的重建图、和所述第三信号分量的重建图,重建所述视频信号。

解码端还会从码流中获得所述第一信号分量的质量因子信息、所述第二信号分量的质量因子信息和所述第三信号分量的质量因子信息。其中所述第一信号分量和第二信号分量的质量因子信息与图7A实施例类似。类似的,所述第三信号分量的质量因子信息也可以为所述第三信号分量的质量因子或者所述第三信号分量的质量因子的索引。然后通过所述第三信号分量的质量因子信息,获得所述第三信号分量的响应信号。当所述第三信号分量的质量因子信息为所述第三信号分量的质量因子时,所述第三信号分量的质量因子取值为L个中的一个;当所述第三信号分量的质量因子信息为所述第三信号分量的质量因子的索引时,所述第一信号分量的质量因子的索引的取值范围为0至L-1或者1至L,其中L为大于1的整数。L,M,N可以相等或者不等,本申请对此不作限定。

当编码端传递过来的为联合的特征图时,解码端还需要对所述联合的特征图进行解码,并经神经网络处理获得所述第一信号分量的特征图、所述第二信号分量的特征图、和所述第三信号分量的特征图。

在一种实施例中,若所述码流中包括所述Y分量的质量因子的索引i所述U分量的质量因子的索引j,和所述V分量的质量因子的索引k,则解码端需要通过学习生成所述第一信号分量的响应信号矩阵 $\{g_{y1}, g_{y2}, \dots, g_{yi} \dots g_{yN}\}$ 、所述第二信号分量的响应信号矩阵 $\{g_{u1}, g_{u2}, \dots, g_{uj} \dots, g_{uM}\}$ 、所述第三信号分量的响应信号矩阵 $\{g_{v1}, g_{v2}, \dots, g_{vk} \dots g_{vL}\}$,其中N,M和L为大于1的整数;可选地,通过对编码端第一信号分量的控制信号矩阵 $\{q_{y1}, q_{y2}, \dots, q_{yi} \dots q_{yN}\}$ 取倒数得到所述第一信号分量的响应信号矩阵 $\{g_{y1}, g_{y2}, \dots, g_{yi} \dots g_{yN}\}$,编码端第二信号分量的控制信号矩阵 $\{q_{u1}, q_{u2}, \dots, q_{ui} \dots q_{uM}\}$ 取倒数得到所述第二信号分量的响应信号矩阵 $\{g_{u1}, g_{u2}, \dots, g_{uj}, \dots, g_{uM}\}$,和编码端第三信号分量的控制信号矩阵 $\{q_{v1}, q_{v2}, \dots, q_{vi} \dots q_{vL}\}$ 取倒数得到所述第三信号分量的响应信号矩阵 $\{g_{v1}, g_{v2}, \dots, g_{vk} \dots g_{vL}\}$,其中N,M和L为大于1的整数;根据所述Y分量质量因子的索引

i 得到所述第一信号分量的响应信号 g_{yi} ；根据所述 U 分量质量因子的索引 j 得到所述第二信号分量的响应信号 g_{uj} ；以及根据所述 V 分量质量因子的索引 k 得到所述第三信号分量的响应信号 g_{vk} 。

在另一种实施例中，若所述码流中包括所述视频信号的质量因子的索引 i ，则解码端需要通过学习生成所述视频信号的响应信号矩阵 $\{g_{c1}, g_{c2}, \dots, g_{ci} \dots g_{cN}\}$ ，其中 c 为 3 代表 Y 分量、 U 分量和 V 分量， N 为大于 1 的整数；可选地，通过对编码端视频信号的控制信号矩阵 $\{q_{c1}, q_{c2}, \dots, q_{ci} \dots q_{cN}\}$ 取倒数，得到所述视频信号的响应信号矩阵 $\{g_{c1}, g_{c2}, \dots, g_{ci} \dots g_{cN}\}$ ，其中 c 为 3 代表 Y 分量、 U 分量和 V 分量， N 为大于 1 的整数；根据所述视频信号的质量因子的索引 i 得到包含所述第一信号分量、所述第二信号分量和所述第三信号分量的响应信号 g_{ci} 。

在再一种实施例中，若所述码流中包括所述第一信号分量的质量因子、所述第二信号分量的质量因子和所述第三信号分量的质量因子，则所述解码端将所述 Y 分量的质量因子作为全连接网络的输入，输出所述 Y 分量的响应信号；将所述 U 分量的质量因子作为全连接网络的输入，输出所述 U 分量的响应信号；将所述 V 分量的质量因子作为全连接网络的输入，输出所述 V 分量的响应信号。

获得所述第一信号分量和第二信号分量的重建图如图 7A 类似，此处不再赘述。

在一种实施例中，当所述响应信号包括响应向量，则根据所述第三信号分量的响应信号和所述第三信号分量的特征图，获得所述第三信号分量的重建图包括：

将所述第三信号分量的响应向量和所述第三信号分量的特征图相乘，获得所述第三信号分量的重建图；或将所述第三信号分量的响应向量和所述第三信号分量的特征图相乘后并且再经神经网络处理，获得所述第三信号分量的重建图。

在再一种实施例中，当所述响应信号包括响应向量和偏移向量，则根据所述第三信号分量的响应信号和所述第三信号分量的特征图，获得所述第三信号分量的重建图包括：

将所述第三信号分量的响应向量与所述第三信号分量的特征图相乘，再与所述第三信号分量的偏移向量相加，获得所述第三信号分量的重建图；或将所述第三信号分量的响应向量与所述第三信号分量的特征图相乘，再与所述第三信号分量的偏移向量相加后并且再经神经网络处理，获得所述第三信号分量的重建图。

实施例 7A 对 UV 分量组合合并处理，7D 对 Y 、 U 、 V 三个分量单独进行处理，也可以进行组合合并处理，例如组合成 Y 、 UV 两种分量，或 Y 、 U 、 V 的其他组合。

图 8A 和 8B 给出一种具体的实施例。本实施例的技术方案整体框图如图 8A 所示，

其中, Y 分量质量因子、UV 分量质量因子输入码率分配控制模块, 该模块输出控制向量 q_{yi} 、控制向量 q_{uvi} 分别作用于 Y 分量处理模块输出的 Y 分量特征图、UV 分量处理模块输出的 UV 分量特征图, 从而实现 YUV 分量的码率分配。

解码端通过 Y 分量质量因子、UV 分量质量因子输入质量响应模块, 该模块输出控制向量 g_{yi} 、控制向量 g_{uvi} 分别作用于 Y 分量特征图、UV 分量特征图, 从而实现 YUV 分量的各自质量增益相应。

本实施例不对 Y 分量处理模块、UV 分量处理模块、联合处理模块、概率估计模块、Y 分量处理模块 2、UV 分量处理模块 2、联合处理模块 2 的具体网络结构做约束, 为便于理解图 8B 给出一种具体示例。

第一步, 获取 Y、UV 分量的特征图:

将 Y、UV 分量分别输入 Y 分量处理模块、UV 分量处理模块, 网络输出 Y、UV 分量的特征图。以图 8B 为例, Y 分量处理模块包含两层卷积层和两层非线性层, 两个卷积层中水平、垂直方向下采样因子均为 2, Y 分量处理模块输出 Y 分量特征图。UV 分量处理模块包含两层卷积层和两层非线性层, 第一层卷积层水平和垂直方向下采样因子为 1, 即不进行无下采样操作。UV 分量处理模块第二层卷积层水平和垂直方向下采样因子为 2。UV 分量处理模块输出 UV 分量特征图。经过上述网络处理后, 对于 YUV420 的数据格式, Y 分量特征图与 UV 分量特征图宽、高相同。

类似 YUV420 数据格式的处理方式, 对于 YUV444、YUV422 等数据格式, 通过控制卷积层数目、水平和垂直方向的下采样因子使得 Y 分量特征图与 UV 分量特征图宽、高相同。

第二步, Y 分量质量因子、UV 分量质量因子输入码率分配模块获取控制向量 q_{yi} 、控制向量 q_{uvi} , 控制向量 q_{yi} 、控制向量 q_{uvi} 与 Y、UV 分量的特征图逐通道相乘得到处理后的 Y、UV 分量的特征图, 处理后的 Y、UV 分量的特征图相加或级联在一起, 输入联合处理模块, 输出待编码的特征图。

码率分配模块由控制矩阵 Q_y 和 Q_{uv} 组成, 将 Y 分量质量因子、UV 分量质量因子作为控制矩阵 Q_y 和 Q_{uv} 的索引值, 从 Q_y 和 Q_{uv} 索引得到控制向量 q_{yi} 、控制向量 q_{uvi} 。其中, 控制矩阵 Q_y 和 Q_{uv} 经网络学习得到。其中, Y 分量质量因子、UV 分量质量因子为设定的任意值。

以图 8B 为例, 控制矩阵 Q_y 为 $K \times N$ 大小的二维矩阵, 控制矩阵 Q_{uv} 为 $L \times M$ 大小的二维矩阵, 两个矩阵中的每个元素为网络可学习的参数。其中 K 表示 Y 分量特征图数目, L 表示 UV 分量特征图数目, N 表示 N 组 Y 分量质量因子候选值, M 表示 M 组 UV 分量质量

因子候选值。以 N 为 4、 M 为 4 为例， Y 分量质量因子可选值为 $\{0.5, 0.7, 0.8, 1.0\}$ ， UV 分量质量因子可选值为 $\{0.15, 0.2, 0.25, 0.3\}$ 。

第三步，待编码的特征图输入熵编码模块，输出码流。以图 8B 为例，待编码的特征图输入编码的特征图 Hyper Entropy 模块，输出待编码符号的概率分布。基于所述待编码符号的概率分布进行算术编码，输出码流。同时， Y 分量质量因子、 UV 分量质量因子信息写入码流。

Y 分量质量因子、 UV 分量质量因子信息可采用如下三种方式进行表达写入码流：

方案一：预定义 Y 分量和 UV 分量质量因子候选值数目及候选值，将 Y 分量质量因子、 UV 分量质量因子在各自候选列表中的索引号传递至解码端。以 N 为 4、 M 为 3 为例， Y 分量质量因子可选值为 $\{0.5, 0.7, 0.8, 1.0\}$ ， UV 分量质量因子可选值为 $\{0.15, 0.2, 0.25\}$ ，将 Y 分量索引号 i ， UV 分量索引号 j 写入码流， i, j 取值为 0, 1, 2, 3，当 i 为 1 时，表示 Y 分量质量因子为 0.7，当 j 为 0 时，表示 UV 分量质量因子为 0.15。

方案二：预定义 Y 分量和 UV 分量质量因子组合后的候选值数目及候选值，例如， Y 、 UV 分量质量因子组合值的候选数目为 6，候选列表为 $\{(0.5, 0.25), (0.7, 0.15), (0.7, 0.25), (0.8, 0.1), (0.8, 0.2), (1.0, 0.2)\}$ ，将索引号 i 写入码流， i 取值为 0, 1, 2, 3, 4, 5，当 i 为 1 时，表示 Y 、 U 、 V 分量质量因子为 $(0.7, 0.15)$ 。

方案三：直接将 Y 分量质量因子和 UV 分量质量因子写入码流传递至解码端，例如，将 $(1.0, 0.2)$ 写入码流。

第四步，码流输入熵解码模块，进行算术解码得到特征图及 Y 分量质量因子、 UV 分量质量因子。以图 8B 为例，基于 Hyper Entropy 模块估计的概率分布进行算术解码。

第五步，解码得到特征图输入联合处理模块 2，输出通道数为 M 的特征图，将通道数为 M 的特征图拆分为通道数为 K 的 Y 分量特征图、通道数为 L 的 UV 分量特征图。拆分方案保证 $K \leq M$ ， $L \leq M$ 即可。其中，当 $K=L=M$ 时，表示将 Y 分类特征图与 UV 分量特征图相同，均为通道数为 M 的特征图。以图 8B 为例，联合处理模块 2 包含二层卷积层，一层非线性层。

第六步， Y 分量质量因子、 UV 分量质量因子输入质量响应模块获取响应向量 g_{yi} 、响应向量 g_{uvi} ，响应向量 g_{yi} 、响应向量 g_{uvi} 与 Y 、 UV 分量的特征图逐通道相乘得到质量增益后的 Y 、 UV 分量的特征图。所述质量增益后的 Y 、 UV 分量的特征图分别输入 Y 分量处理模块 2、 UV 分量处理模块 2，输出 Y 分量重建图、 UV 分量重建图。

质量响应模块由响应矩阵 G_y 和 G_{uv} 组成，将解码得到的 Y 分量质量因子、 UV 分量

质量因子作为响应矩阵 G_y 和 G_{uv} 的索引值,从 G_y 和 G_{uv} 索引得到响应向量 g_{yi} 、响应向量 g_{uvi} 。

其中,响应矩阵 G_y 和 G_{uv} 经网络学习得到。以图 8B 为例,响应矩阵 G_y 为 $K \times N$ 大小的二维矩阵,响应矩阵 G_{uv} 为 $L \times M$ 大小的二维矩阵,两个矩阵中的每个元素为网络可学习的参数。其中 K 表示 Y 分量特征图数目, L 表示 UV 分量特征图数目, N 表示 N 组 Y 分量质量因子候选值, M 表示 M 组 UV 分量质量因子候选值。

可选地,对控制矩阵 Q_y 和 Q_{uv} 分别取倒数得到响应矩阵 G_y 和 G_{uv} 。

本实施例对上述步骤 1 至步骤 7 涉及的网络模块及控制矩阵参数进行训练学习。具体的,本申请使用自适应矩估计 (Adaptive Moment Estimation, Adam) 优化算法对神经网络进行优化训练, ImgeNet 数据集为训练数据集。由于本网络结构面向图像编码,训练优化目标为最小化码率-失真联合损失函数,其函数表达式为:

$$R + \lambda(w_y D_y + w_u D_u + w_v D_v) \\ = E[-\log_2 p(y)] + \lambda w_y E\|x_y - x_y'\|_2 + \lambda w_u E\|x_u - x_u'\|_2 + \lambda w_v E\|x_v - x_v'\|_2$$

其中, $p(y)$ 代表概率估计器估计的概率分布, x_y 为 Y 分量原始值, x_y' 为 Y 分量重建图, x_u 为 U 分量原始值, x_u' 为 U 分量重建图, x_v 为 V 分量原始值, x_v' 为 V 分量重建图, w_y 为 Y 分量的质量因子, w_u 为 U 分量的质量因子, w_v 为 V 分量的质量因子。 λ 为常数, 匹配目标码率。

以 N 为 4 为例, Y 、 U 、 V 分量质量因子 (w_y, w_u, w_v) 可选值为 $\{(0.5, 0.25, 0.25), (0.7, 0.4, 0.4), (0.8, 0.1, 0.1), (1.0, 0.2, 0.2)\}$ 。网络训练时, 权重组索引号 i 从 $\{0, 1, 2, 3\}$ 随机选取, 根据 i 确定权重值组 (w_{yi}, w_{ui}, w_{vi}) 及待学习的控制向量 q_{yi} 、控制向量 q_{uvi} , 根据优化目标对网络模块及控制矩阵参数进行训练学习。

本实施例中提及编解码器中去除联合处理模块或/和联合处理模块 2 或/和质量响应模块时, 本申请其他实施例中依然适用。

本实施例给出 Y 、 U 、 V 分量组合为 Y 、 UV 两种分量的情况下的技术方案, 对 Y 、 U 、 V 的其他组合例如 $\{YU, V\}$ 、 $\{YV, U\}$, 本申请的解决思路依然适用。

同样, 将 UV 分量进一步拆分为 U 分量、 V 分量分别处理, 本申请的解决思路依然适用。

在现有端到端图像编码中, 对于某一特定网络学习优化时, 按固定 Y 、 U 、 V 分量权重值进行优化, 因此, YUV 分量的码率固定。由于不同图像色彩特性不同, 固定码率分配会导致部分视频图像内容上编码性能较差的现象。可简单通过按照多组不同 Y 、 U 、 V 分量权重值训练多个模型, 以实现 YUV 分量不同码率分配, 但这样会增加模型数目, 且训

练多个模型花费大量计算资源及时间。相比于现有技术，本申请基于Y、U、V分量权重值导出网络学习得到的控制向量，根据控制向量对Y、UV分量的特征图进行不同程度的失真控制，实现Y、UV分量的码率分配。从而本申请具有以下优势：

- 1) 支持YUV分量间的码率分配，达到适配不同色彩特性的图像内容。
- 2) 减少训练多个模型的时间花费，同时也减少模型新增的网络参数量。

图9A和9B给出一种具体的实施例。在图9A和9B实施例的基础上，本实施例采用U分量处理模块、V分量处理模块分别处理U分量、V分量数据。本实施例，将Y分量质量因子、U分量质量因子、V分量质量因子作为码率分配控制模块的输入，输出控制信号对Y分量处理模块、U分量处理模块、V分量处理模块中的任意层特征图进行处理，实现Y、U、V分量间的码率分配。解码端，将Y分量质量因子、U分量质量因子、V分量质量因子作为质量响应模块的输入，输出控制信号对Y分量处理模块、U分量处理模块、V分量处理模块中的任意层特征图进行质量增益响应。本申请不对码率分配控制模块、质量响应模块、Y分量处理模块、U分量处理模块、联合处理模块、Y分量处理模块2、U分量处理模块2、V分量处理模块2、联合处理模块2、熵编码模块、熵解码模块的具体网络结构做约束，为便于理解图9A给出一种具体示例。

以图9B为例，第一步，Y分量质量因子、U分量质量因子、V分量质量因子输入码率分配控制模块，该模块由全连接网络组成，模块输出控制信号：控制向量、偏移向量。

第二步，待编码Y、U、V信号分别输入Y分量处理模块、U分量处理模块、V分量处理模块。以Y分量处理模块为例，该模块中每个卷积层输出的特征图与其相对应的控制向量逐通道相乘，然后与其相对应的偏移向量逐通道相加。模块中非线性层输出的特征图与其相对应的控制向量逐通道相乘。Y分量处理模块中每层网络的输出均经码率控制模块的控制信号进行处理。U、V分量的处理方式与Y分量类似。

图9B给出一种具体网络结构的示意图，Y分量处理模块、U分量处理模块、V分量处理模块的网络结构除第一层卷积层外，其他网络层均相同。对于YUV422格式，Y分量处理模块中第一层卷积层水平和垂直方向下采样因子为2。U分量处理模块、V分量处理模块中第一层卷积层水平下采样因子为1即不进行下采样操作，垂直方向下采样因子为2。对于YUV420格式，Y分量处理模块中第一层卷积层水平和垂直方向下采样因子为2。U分量处理模块、V分量处理模块中第一层卷积层水平和垂直方向下采样因子为1，即均不进行下采样操作。

第三步，将Y分量特征图、U分量特征图、V分量特征图级联拼接在一起形成待编

码的特征图，输入熵编码模块，输出码流。以图 9B 为例，待编码的特征图输入编码的特征图 Hyper Entropy 模块，输出待编码符号的概率分布。基于所述待编码符号的概率分布进行算术编码，输出码流。同时，Y 分量质量因子、U 分量质量因子、U 分量质量因子信息写入码流。

第四步，码流输入熵解码模块，进行算术解码得到特征图及 Y 分量质量因子、U 分量质量因子、V 分量质量因子信息。以图 9B 为例，基于 Hyper Entropy 模块估计的概率分布进行算术解码。

第六步，将解码得到的特征图输入联合处理模块 2，输出的特征图。

第七步，Y 分量质量因子、U 分量质量因子、V 分量质量因子输入质量响应模块获取响应向量 g_{yi} 、响应向量 g_{ui} 、响应向量 g_{vi} 。响应向量 g_{yi} 与 Y 分量处理模块 2 的中第二层卷积层输出的特征图逐通道相乘得到质量增益后的特征图，U、V 分量处理过程类似。Y 分量处理模块 2、U 分量处理模块 2、V 分量处理模块 2 输出 Y、U、V 分量的重建图。

Y 分量质量因子、U 分量质量因子、V 分量质量因子输入质量响应模块，该模块由全连接网络组成，模块输出响应向量 g_{yi} 、响应向量 g_{ui} 、响应向量 g_{vi} 。

可选的，类似实施例一响应向量获取方式，质量响应模块由响应矩阵 G_y 、 G_u 和 G_v 组成，将解码得到的 Y 分量质量因子、U 分量质量因子、V 分量质量因子作为响应矩阵 G_y 、 G_u 和 G_v 的索引值，从 G_y 、 G_u 和 G_v 索引得到响应向量 g_{yi} 、响应向量 g_{ui} 、响应向量 g_{vi} 。其中，响应矩阵 G_y 、 G_u 和 G_v 经网络学习得到。

网络训练过程与图 8A 和 8B 实施例类似，不再赘述。

本实施例中控制信号作用于 Y 分量处理模块、U 分量处理模块、V 分量处理模块中每层网络的输出，可选的，控制信号只作用于 Y 分量处理模块、U 分量处理模块、V 分量处理模块中部分网络的输出。

本实施例中响应信号只作用于 Y 分量处理模块、U 分量处理模块、V 分量处理模块中间一层网络的输出，可选的，控制信号作用于 Y 分量处理模块、U 分量处理模块、V 分量处理模块中任意一层或多层网络的输出。

本实施例中提及编解码器去除联合处理模块 2 或/和质量响应模块时，本申请技术依然适用。

本实施例中提及编解码器中增加联合处理模块时，本申请技术依然适用。

本实施例给出 Y、U、V 作为三种分量分别处理的情况下的技术方案，对 Y、U、V 的其他组合例如 {YU、V}、{YV、U}、{Y、UV}，本申请技术的解决思路依然适用。

根据前面实施例的描述,本申请中本发明中 YUV 不同分量的质量因子输入码率分配控制模块,该模块输出控制信号分别作用于不同分量的特征图,从而实现不同分量的码率分配。所述不同分量可指 Y、U、V 三种分量, Y、UV 两种分量,或 Y、U、V 的其他组合。

可选的,所述控制信号指控制向量 q_i ,根据不同分量的质量因子产生:网络学习得到同分量的权重矩阵 $\{q_{c1}, q_{c2}, \dots, q_{cN}\}$,其中 c 为 2 或 3 代表不同分量数目, N 为质量因子候选值数目。使用时根据不同分量的质量因子索引得到不同分量相应的控制向量 q_{ci} 。

可选的,所述控制信号指控制向量 q 、偏移向量 b ,不同分量的质量因子作为全连接网络的输入,输出不同分量相应的控制向量 q 、偏移向量 b 。

因此,本申请提供的实施例能够:

- 1) 适配不同色彩特性的图像内容, 通过控制向量支持 YUV 分量间的码率分配。
- 2) 减少训练多个模型的时间花费,同时也减少模型新增的网络参数量。

图 10 是示出根据本申请一种实施例的编码装置 1000 的结构示意图。该编码装置 1000 可以对应于视频编码器 20。该编码装置 1000 包括第一控制模块 1001、第二控制模块 1002 和编码模块 1003。其中,第一控制模块 1001,用于把所述视频信号的第一信号分量的控制信号作用于所述第一信号分量的第一特征图,获得所述第一信号分量的第二特征图,其中所述第一信号分量的控制信号通过学习获得;第二控制模块 1002,用于把所述视频信号的第二信号分量的控制信号作用于所述第二信号分量的第一特征图,获得所述第二信号分量的第二特征图,其中所述第二信号分量的控制信号通过学习获得;以及编码模块 1003,用于根据所述第一信号分量的第二特征图和所述第二信号分量的第二特征图,获得所述视频信号的码流。该编码装置 1000 还可以包括前面实施例介绍的码率分配控制模块。编码装置 1000 用于实现前述实施例介绍的编码方法,详细的功能参见前面实施例的描述,此处不再重复描述。

图 11 是示出根据本申请一种实施例的解码装置 1100 的结构示意图。该解码装置 1100 可以对应于视频解码器 30。该解码装置 1100 包括解码模块 1101、第一控制模块 1102、第二控制模块 1103 和重建模块 1104。其中,解码模块 1101,用于获得所述视频信号的码流,对所述码流进行熵解码以获得所述视频信号的第一信号分量的特征图和所述视频信号的第二信号分量的特征图;第一控制模块 1102,用于根据所述第一信号分量的响应信号和所述第一信号分量的特征图,获得所述第一信号分量的重建图,其中所述第一信号分量的响应信号通过学习获得;第二控制模块 1103,用于根据所述第二信号分量的响应信号和所述第二信号分量的特征图,获得所述第二信号分量的重建图,其中所述第二

信号分量的响应信号通过学习获得；以及重建模块 1104，用于根据所述第一信号分量的重建图和所述第二信号分量的重建图，重建所述视频信号。该解码装置 1100 还可以包括前面实施例介绍的质量响应模块。解码装置 1100 用于实现前述实施例介绍的解码方法，详细的功能参见前面实施例的描述，此处不再重复描述。

本领域技术人员能够领会，结合本文公开描述的各种说明性逻辑框、模块和算法步骤所描述的功能可以硬件、软件、固件或其任何组合来实施。如果以软件来实施，那么各种说明性逻辑框、模块、和步骤描述的功能可作为一或多个指令或代码在计算机可读媒体上存储或传输，且由基于硬件的处理单元执行。计算机可读媒体可包含计算机可读存储媒体，其对应于有形媒体，例如数据存储媒体，或包括任何促进将计算机程序从一处传送到另一处的媒体（例如，根据通信协议）的通信媒体。以此方式，计算机可读媒体大体上可对应于（1）非暂时性的有形计算机可读存储媒体，或（2）通信媒体，例如信号或载波。数据存储媒体可为可由一或多个计算机或一或多个处理器存取以检索用于实施本申请中描述的技术的指令、代码和/或数据结构的任何可用媒体。计算机程序产品可包含计算机可读媒体。

作为实例而非限制，此类计算机可读存储媒体可包括 RAM、ROM、EEPROM、CD-ROM 或其它光盘存储装置、磁盘存储装置或其它磁性存储装置、快闪存储器或可用来存储指令或数据结构形式的所要程序代码并且可由计算机存取的任何其它媒体。并且，任何连接被恰当地称作计算机可读媒体。举例来说，如果使用同轴缆线、光纤缆线、双绞线、数字订户线(DSL)或例如红外线、无线电和微波等无线技术从网站、服务器或其它远程源传输指令，那么同轴缆线、光纤缆线、双绞线、DSL 或例如红外线、无线电和微波等无线技术包含在媒体的定义中。但是，应理解，所述计算机可读存储媒体和数据存储媒体并不包括连接、载波、信号或其它暂时媒体，而是实际上针对于非暂时性有形存储媒体。如本文中所使用，磁盘和光盘包含压缩光盘(CD)、激光光盘、光学光盘、数字多功能光盘(DVD)和蓝光光盘，其中磁盘通常以磁性方式再现数据，而光盘利用激光以光学方式再现数据。以上各项的组合也应包含在计算机可读媒体的范围内。

可通过例如一或多个数字信号处理器(DSP)、通用微处理器、专用集成电路(ASIC)、现场可编程逻辑阵列(FPGA)或其它等效集成或离散逻辑电路等一或多个处理器来执行指令。因此，如本文中所使用的术语“处理器”可指前述结构或适合于实施本文中所描述的技术的任一其它结构中的任一者。另外，在一些方面中，本文中所描述的各种说明性逻辑框、模块、和步骤所描述的功能可以提供于经配置以用于编码和解码的专用硬件和/

或软件模块内，或者并入在组合编解码器中。而且，所述技术可完全实施于一或多个电路或逻辑元件中。

本申请的技术可在各种各样的装置或设备中实施，包含无线手持机、集成电路(IC)或一组 IC(例如，芯片组)。本申请中描述各种组件、模块或单元是为了强调用于执行所揭示的技术的装置的功能方面，但未必需要由不同硬件单元实现。实际上，如上文所描述，各种单元可结合合适的软件和/或固件组合在编解码器硬件单元中，或者通过互操作硬件单元（包含如上文所描述的一或多个处理器）来提供。

以上所述，仅为本申请示例性的具体实施方式，但本申请的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本申请揭露的技术范围内，可轻易想到的变化或替换，都应涵盖在本申请的保护范围之内。因此，本申请的保护范围应该以权利要求的保护范围为准。

权 利 要 求 书

1、一种视频信号的编码方法，其特征在于，包括：

把所述视频信号的第一信号分量的控制信号作用于所述第一信号分量的第一特征图，获得所述第一信号分量的第二特征图，其中所述第一信号分量的控制信号通过学习获得；

把所述视频信号的第二信号分量的控制信号作用于所述第二信号分量的第一特征图，获得所述第二信号分量的第二特征图，其中所述第二信号分量的控制信号通过学习获得；

以及

根据所述第一信号分量的第二特征图和所述第二信号分量的第二特征图，获得所述视频信号的码流。

2、如权利要求1所述的方法，其特征在于，所述方法还包括：

根据所述第一信号分量的质量因子从N个候选第一控制信号中，获得所述第一信号分量的控制信号，其中N为大于1的整数；以及

根据所述第二信号分量的质量因子从M个候选第二控制信号中，获得所述第二信号分量的控制信号，其中M为大于1的整数。

3、如权利要求1或2所述的方法，其特征在于，根据所述第一信号分量的第二特征图和所述第二信号分量的第二特征图，获得所述视频信号的码流包括：

对所述第一信号分量的第二特征图和所述第二信号分量的第二特征图，进行熵编码，以获得所述视频信号的码流；或者，

对所述第一信号分量的第二特征图和经神经网络处理的所述第二信号分量的第二特征图，进行熵编码，以获得所述视频信号的码流；或者，

对经神经网络处理的所述第一信号分量的第二特征图和所述第二信号分量的第二特征图，进行熵编码，以获得所述视频信号的码流；或者，

对经神经网络处理的所述第一信号分量的第二特征图和经神经网络处理的所述第二信号分量的第二特征图，进行熵编码，以获得所述视频信号的码流。

4、如权利要求1或2所述的方法，其特征在于，根据所述第一信号分量的第二特征图和所述第二信号分量的第二特征图，获得所述视频信号的码流包括：

对所述第一信号分量的第二特征图和所述第二信号分量的第二特征图，进行联合处理，获得联合的特征图，对所述联合的特征图进行熵编码，以获得所述视频信号的码流；或者，

对所述第一信号分量的第二特征图和经神经网络处理的所述第二信号分量的第二特

征图, 进行联合处理, 获得联合的特征图, 对所述联合的特征图进行熵编码, 以获得所述视频信号的码流; 或者,

对神经网络处理的所述第一信号分量的第二特征图和所述第二信号分量的第二特征图, 进行联合处理, 获得联合的特征图, 对所述联合的特征图进行熵编码, 以获得所述视频信号的码流; 或者,

对神经网络处理的所述第一信号分量的第二特征图和神经网络处理的所述第二信号分量的第二特征图, 进行联合处理, 获得联合的特征图, 对所述联合的特征图进行熵编码, 以获得所述视频信号的码流。

5、如权利要求 1-4 任一所述的方法, 其特征在于, 所述第一信号分量为 Y 分量, 所述第二信号分量为 UV 分量, U 分量或 V 分量。

6、如权利要求 5 所述的方法, 其特征在于, 当所述第二信号分量为 UV 分量时, 所述方法包括:

通过学习生成所述 Y 分量的控制信号矩阵 $\{q_{y1}, q_{y2}, \dots, q_{yi} \dots q_{yN}\}$ 和所述 UV 分量的控制信号矩阵 $\{q_{uv1}, q_{uv2}, \dots, q_{uvj}, \dots, q_{uvM}\}$, 其中 N 和 M 为大于 1 的整数;

根据所述 Y 分量的质量因子的索引 i 得到所述第一信号分量的控制信号 q_{yi} ;

根据所述 UV 分量的质量因子的索引 j 得到所述第二信号分量的控制信号 q_{uvj} 。

7、如权利要求 5 所述的方法, 其特征在于, 当所述第二信号分量为 UV 分量时, 所述方法包括:

通过学习生成所述视频信号的控制信号矩阵 $\{q_{c1}, q_{c2}, \dots, q_{ci} \dots q_{cN}\}$, 其中 c 为 2 代表 Y 分量和 UV 分量, N 为大于 1 的整数;

根据所述视频信号的质量因子的索引 i 得到包含所述第一信号分量和所述第二信号分量的控制信号 q_{ci} 。

8、如权利要求 6 或 7 所述的方法, 其特征在于, 所述视频信号的码流中包括所述 Y 分量的质量因子的索引 i 和所述 UV 分量的质量因子的索引 j; 或包括所述视频信号的质量因子的索引 i。

9、如权利要求 5 所述的方法, 其特征在于, 当所述第二信号分量为 UV 分量时, 所述方法包括:

将所述 Y 分量的质量因子作为全连接网络的输入, 输出所述 Y 分量的控制信号;

将所述 UV 分量的质量因子作为全连接网络的输入, 输出所述 UV 分量的控制信号。

10、如权利要求 9 所述的方法, 其特征在于, 所述视频信号的码流中包括所述 Y 分

量的质量因子和所述UV分量的质量因子。

11、如权利要求5所述的方法，其特征在于，当所述第二信号分量为U分量或者V分量时，所述方法还包括：

把所述视频信号的第三信号分量的控制信号作用于所述第三信号分量的第一特征图，获得所述第三信号分量的第二特征图，其中所述第三信号分量的控制信号通过学习获得，其中当所述第二信号分量为U分量时，所述第三信号分量为V分量，或当所述第二信号分量为V分量，所述第三信号分量为U分量。

12、如权利要求11所述的方法，其特征在于，所述方法包括：

通过学习生成所述Y分量的控制信号矩阵 $\{q_{y1}, q_{y2}, \dots, q_{yi} \dots q_{yN}\}$ 、所述U分量的控制信号矩阵 $\{q_{u1}, q_{u2}, \dots, q_{uj} \dots, q_{uM}\}$ 、所述V分量的控制信号矩阵 $\{q_{v1}, q_{v2}, \dots, q_{vk} \dots q_{vL}\}$ ，其中N、M和L为大于1的整数；

根据所述Y分量质量因子的索引i得到所述第一信号分量的控制信号 q_{yi} ；

根据所述U分量质量因子的索引j得到所述第二信号分量的控制信号 q_{uj} ；以及

根据所述V分量质量因子的索引k得到所述第三信号分量的控制信号 q_{vk} 。

13、如权利要求11所述的方法，其特征在于，所述方法包括：

通过学习生成所述视频信号的控制信号矩阵 $\{q_{c1}, q_{c2}, \dots, q_{ci} \dots q_{cN}\}$ ，其中c为3代表Y分量、U分量和V分量，N为大于1的整数；

根据所述视频信号的质量因子的索引i得到包含所述第一信号分量、所述第二信号分量和所述第三信号分量的控制信号 q_{ci} 。

14、如权利要求12或13所述的方法，其特征在于，所述视频信号的码流中包括所述Y分量的质量因子的索引i、所述U分量的质量因子的索引j、和所述V分量质量因子的索引k；或包括所述视频信号的质量因子的索引i。

15、如权利要求11所述的方法，其特征在于，所述方法包括：

将所述Y分量的质量因子作为全连接网络的输入，输出Y分量的控制信号；

将所述U分量的质量因子作为全连接网络的输入，输出U分量的控制信号；以及

将所述V分量的质量因子作为全连接网络的输入，输出V分量的控制信号。

16、如权利要求15所述的方法，其特征在于，所述视频信号的码流中包括所述Y分量的质量因子、所述U分量的质量因子和所述V分量的质量因子。

17、如权利要求1-16任一所述的方法，其特征在于，当所述控制信号包括控制向量，所述方法包括：

所述第一信号分量的控制向量和所述第一信号分量的第一特征图相乘，获得所述第一信号分量的第二特征图；

所述第二信号分量的控制向量和所述第一信号分量的第二特征图相乘，获得所述第二信号分量的第二特征图。

18、如权利要求 1-16 任一所述的方法，其特征在于，当所述控制信号包括控制向量和偏移向量，所述方法包括：

所述第一信号分量的控制向量和所述第一信号分量的第一特征图相乘，然后加上所述第一信号分量的偏移向量，获得所述第一信号分量的第二特征图；

所述第二信号分量的控制向量和所述第一信号分量的第二特征图相乘，然后加上所述第二信号分量的偏移向量，获得所述第二信号分量的第二特征图。

19、如权利要求 1-18 任一所述的方法，其特征在于，所述第一信号分量的第一特征图为所述第一信号分量通过至少一层卷积和/或至少一层非线性处理获得；

所述第二信号分量的第一特征图为所述第二信号分量通过至少一层卷积和/或至少一层非线性处理获得。

20、如权利要求 19 所述的方法，其特征在于，所述第一信号分量的第一特征图为所述第一信号分量通过两层卷积层，下采样因子均为 2，两层非线性层获得；

所述第二信号分量的第一特征图为所述第二信号分量通过一层卷积层，无下采样操作，一层卷积层，下采样因子为 2，两层非线性层获得。

21、一种视频信号的解码方法，其特征在于，包括：

获得所述视频信号的码流；

对所述码流进行熵解码以获得所述第一信号分量的特征图和所述视频信号的第二信号分量的特征图；

根据所述第一信号分量的响应信号和所述第一信号分量的特征图，获得所述第一信号分量的重建图，其中所述第一信号分量的响应信号通过学习获得；

根据所述第二信号分量的响应信号和所述第二信号分量的特征图，获得所述第二信号分量的重建图，其中所述第二信号分量的响应信号通过学习获得；以及

根据所述第一信号分量的重建图和所述第二信号分量的重建图，重建所述视频信号。

22、如权利要求 21 所述的方法，其特征在于，所述码流中还包括所述第一信号分量的质量因子信息和所述第二信号分量的质量因子信息，其中所述第一信号分量的质量因子信息为所述第一信号分量的质量因子或者所述第一信号分量的质量因子的索引，所述

第二信号分量的质量因子信息为所述第二信号分量的质量因子或者所述第二信号分量的质量因子的索引；

通过所述第一信号分量的质量因子信息，获得所述第一信号分量的响应信号；

通过所述第二信号分量的质量因子信息，获得所述第二信号分量的响应信号。

23、如权利要求 22 所述的方法，其特征在于，当所述第一信号分量的质量因子信息为所述第一信号分量的质量因子时，所述第一信号分量的质量因子取值为 N 个中的一个；当所述第一信号分量的质量因子信息为所述第一信号分量的质量因子的索引时，所述第一信号分量的质量因子的索引的取值范围为 0 至 $N-1$ 或者 1 至 N ，其中 N 为大于 1 的整数；

当所述第二信号分量的质量因子信息为所述第二信号分量的质量因子时，所述第二信号分量的质量因子取值为 M 个中的一个；当所述第二信号分量的质量因子信息为所述第二信号分量的质量因子的索引时，所述第二信号分量的质量因子的索引的取值范围为 0 至 $M-1$ 或者 1 至 M ，其中 M 为大于 1 的整数。

24、如权利要求 21-23 所述的方法，其特征在于，所述码流中包括联合的特征图，

对所述联合的特征图进行熵解码，并经神经网络处理获得所述第一信号分量的特征图和所述第二信号分量的特征图。

25、如权利要求 21-24 任一所述的方法，其特征在于，所述第一信号分量为 Y 分量，所述第二信号分量为 UV 分量， U 分量或 V 分量。

26、如权利要求 25 所述的方法，其特征在于，当所述第二信号分量为 UV 分量时，若所述码流中包括所述 Y 分量的质量因子的索引 i 和所述 UV 分量的质量因子的索引 j ，所述方法包括：

通过学习生成所述第一信号分量的响应信号矩阵 $\{g_{y1}, g_{y2}, \dots, g_{yi} \dots g_{yN}\}$ 、和所述第二信号分量的响应信号矩阵 $\{g_{uv1}, g_{uv2}, \dots, g_{uvj}, \dots, g_{uvM}\}$ ，其中 N 和 M 为大于 1 的整数；

根据所述 Y 分量的质量因子的索引 i 得到所述第一信号分量的响应信号 g_{yi} ；

根据所述 UV 分量的质量因子的索引 j 得到所述第二信号分量的响应信号 g_{uvj} 。

27、如权利要求 25 所述的方法，其特征在于，当所述第二信号分量为 UV 分量时，若所述码流中包括所述视频信号的质量因子的索引 i ，所述方法包括：

通过学习生成所述视频信号的响应信号矩阵 $\{g_{c1}, g_{c2}, \dots, g_{ci} \dots g_{cN}\}$ ，其中 c 为 2 代表 Y 分量和 UV 分量， N 为大于 1 的整数；

根据所述视频信号的质量因子的索引 i 得到包含所述第一信号分量和所述第二信号分量的响应信号 g_{ci} 。

28、如权利要求 25 所述的方法，其特征在于，当所述第二信号分量为 UV 分量时，若所述码流中包括所述第一信号分量的质量因子和第二信号分量的质量因子，所述方法包括：

将所述 Y 分量的质量因子作为全连接网络的输入，输出所述 Y 分量的响应信号；

将所述 UV 分量的质量因子作为全连接网络的输入，输出所述 UV 分量的响应信号。

29、如权利要求 25 所述的方法，其特征在于，当所述第二信号分量为 U 分量或者 V 分量时，所述方法还包括：

对所述码流进行熵解码以获得所述视频信号的第三信号分量的特征图；

根据所述第三信号分量的响应信号和所述第三信号分量的特征图，获得所述第三信号分量的重建图，其中所述第三信号分量的响应信号通过学习获得，其中当所述第二信号分量为 U 分量时，所述第三信号分量为 V 分量，或当所述第二信号分量为 V 分量，所述第三信号分量为 U 分量；

则所述重建所述视频信号包括：

根据所述第一信号分量的重建图、所述第二信号分量的重建图和和所述第三信号分量的重建图，重建所述视频信号。

30、如权利要求 29 所述的方法，其特征在于，所述码流中还包括所述第三信号分量的质量因子信息，其中所述第三信号分量的质量因子信息为所述第三信号分量的质量因子或者所述第三信号分量的质量因子的索引，所述第三信号分量的质量因子取值为 L 中的一个，所述第三信号分量的质量因子的索引的取值范围为 0 至 L-1 或者 1 至 L，其中 L 为大于 1 的整数；

通过所述第三信号分量的质量因子信息，获得所述第三信号分量的响应信号。

31、如权利要求 30 所述的方法，其特征在于，若所述码流中包括所述 Y 分量的质量因子的索引 i、所述 U 分量的质量因子的索引 j，和所述 V 分量的质量因子的索引 k，所述方法包括：

通过学习生成所述第一信号分量的响应信号矩阵 $\{g_{y1}, g_{y2}, \dots, g_{yi} \dots g_{yN}\}$ 、所述第二信号分量的响应信号矩阵 $\{g_{u1}, g_{u2}, \dots, g_{uj} \dots, g_{uM}\}$ 、所述第三信号分量的响应信号矩阵 $\{g_{v1}, g_{v2}, \dots, g_{vk} \dots g_{vL}\}$ ，其中 N, M 和 L 为大于 1 的整数；

根据所述 Y 分量质量因子的索引 i 得到所述第一信号分量的响应信号 g_{yi} ；

根据所述 U 分量质量因子的索引 j 得到所述第二信号分量的响应信号 g_{uj} ；以及

根据所述 V 分量质量因子的索引 k 得到所述第三信号分量的响应信号 g_{vk} 。

32、如权利要求 29 所述的方法，其特征在于，若所述码流中包括所述视频信号的质量因子的索引 i ，所述方法包括：

通过学习生成所述视频信号的响应信号矩阵 $\{g_{c1}, g_{c2}, \dots, g_{ci} \dots g_{cN}\}$ ，其中 c 为 3 代表 Y 分量、U 分量和 V 分量，N 为大于 1 的整数；

根据所述视频信号的质量因子的索引 i 得到包含所述第一信号分量、所述第二信号分量和所述第三信号分量的响应信号 g_{ci} 。

33、如权利要求 29 所述的方法，其特征在于，若所述码流中包括所述第一信号分量的质量因子、所述第二信号分量的质量因子和所述第三信号分量的质量因子，所述方法包括：

将所述 Y 分量的质量因子作为全连接网络的输入，输出所述 Y 分量的响应信号；

将所述 U 分量的质量因子作为全连接网络的输入，输出所述 U 分量的响应信号；

将所述 V 分量的质量因子作为全连接网络的输入，输出所述 V 分量的响应信号。

34、如权利要求 21-33 任一所述的方法，其特征在于，当所述响应信号包括响应向量，则根据所述第一信号分量的响应信号和所述第一信号分量的特征图，获得所述第一信号分量的重建图包括：

将所述第一信号分量的响应向量和所述第一信号分量的特征图相乘，获得所述第一信号分量的重建图；或将所述第一信号分量的响应向量和所述第一信号分量的特征图相乘后并且再经神经网络处理，获得所述第一信号分量的重建图；

则根据所述第二信号分量的响应信号和所述第二信号分量的特征图，获得所述第二信号分量的重建图包括：

将所述第二信号分量的响应向量和所述第二信号分量的特征图相乘，获得所述第二信号分量的重建图；或将所述第二信号分量的响应向量和所述第二信号分量的特征图相乘后并且再经神经网络处理，获得所述第二信号分量的重建图。

35、如权利要求 21-33 任一所述的方法，其特征在于，当所述响应信号包括响应向量和偏移向量，则根据所述第一信号分量的响应信号和所述第一信号分量的特征图，获得所述第一信号分量的重建图包括：

将所述第一信号分量的响应向量与所述第一信号分量的特征图相乘，再与所述第一信号分量的偏移向量相加，获得所述第一信号分量的重建图；或将所述第一信号分量的响应向量与所述第一信号分量的特征图相乘，再与所述第一信号分量的偏移向量相加后并且再经神经网络处理，获得所述第一信号分量的重建图；

则根据所述第二信号分量的响应信号和所述第二信号分量的特征图，获得所述第二信号分量的重建图包括：

将所述第二信号分量的响应向量与所述第二信号分量的特征图相乘，再与所述第二信号分量的偏移向量相加，获得所述第二信号分量的重建图；或将所述第二信号分量的响应向量与所述第二信号分量的特征图相乘，再与所述第二信号分量的偏移向量相加，并且再经神经网络处理，获得所述第二信号分量的重建图。

36、一种编码器，其特征在于，包括处理电路，用于执行权利要求 1-20 任一项所述的方法。

37、一种解码器，其特征在于，包括处理电路，用于执行权利要求 21-35 任一项所述的方法。

38、一种计算机程序产品，其特征在于，包括程序代码，当其在计算机或处理器上执行时，用于执行权利要求 1-35 任一项所述的方法。

39、一种编码器，其特征在于，包括：

一个或多个处理器；

非瞬时性计算机可读存储介质，耦合到所述处理器并存储由所述处理器执行的程序，其中所述程序在由所述处理器执行时，使得所述编码器执行权利要求 1-20 任一项所述的方法。

40、一种解码器，其特征在于，包括：

一个或多个处理器；

非瞬时性计算机可读存储介质，耦合到所述处理器并存储由所述处理器执行的程序，其中所述程序在由所述处理器执行时，使得所述解码器执行权利要求 21-35 任一项所述的方法。

41、一种非瞬时性计算机可读存储介质，其特征在于，包括程序代码，当其由计算机设备执行时，用于执行权利要求 1-35 任一项所述的方法。