



(12)发明专利申请

(10)申请公布号 CN 111800629 A

(43)申请公布日 2020.10.20

(21)申请号 201910279326.5 *H04N 19/42*(2014.01)

(22)申请日 2019.04.09 *H04N 19/44*(2014.01)

(71)申请人 华为技术有限公司 *H04N 19/61*(2014.01)

地址 518129 广东省深圳市龙岗区坂田华为总部办公楼 *H04N 19/625*(2014.01)

申请人 韩国科学技术院 *H04N 19/70*(2014.01)

*H04N 19/91*(2014.01)

*H04N 19/96*(2014.01)

(72)发明人 周川 金慕淳 *G06N 3/04*(2006.01)

(74)专利代理机构 北京龙双利达知识产权代理有限公司 11329

代理人 张振 张欣

(51)Int.Cl.

*H04N 19/103*(2014.01)

*H04N 19/159*(2014.01)

*H04N 19/176*(2014.01)

*H04N 19/186*(2014.01)

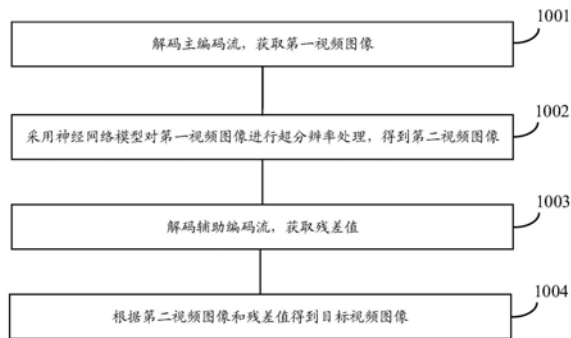
权利要求书3页 说明书35页 附图15页

(54)发明名称

视频解码方法、编码方法以及视频解码器和编码器

(57)摘要

本申请提供了视频解码方法、编码方法以及视频解码器和视频编码器。其中，视频解码方法包括：分别解析主编码流和辅助编码流，以获得分辨率较低的第一视频图像和残差值，然后利用预设的卷积神经网络对该第一视频图像进行超分辨率提升，以得到分辨率较高的第二视频图像，最后再根据第二视频图像和残差值得到最终的目标视频图像。本申请中，通过预设的神经网络模型能够更好地对分辨率较低的第一视频图像进行超分辨率处理，进而得到分辨率较高的第二视频图像，进而能够根据第二视频图像与残差值的一次叠加就能得到最终的目标视频图像，与传统方案中采用层层叠加的方式才能得到最终的视频图像的方式相比，能够简化解码流程。



1. 一种视频解码方法,其特征在于,包括:  
解码主编码流,获取第一视频图像;  
采用神经网络模型对所述第一视频图像进行超分辨率处理,得到第二视频图像;  
解码辅助编码流,获取残差值;  
将所述第二视频图像的像素值与所述残差值叠加,得到目标视频图像。
2. 如权利要求1所述的方法,其特征在于,所述神经网络模型是根据训练样本图像和训练目标图像训练得到的,所述训练目标图像是与所述目标视频图像的分辨率相同的视频图像,所述训练样本图像是对所述训练目标图像进行下采样、编码和解码处理得到的视频图像。
3. 如权利要求1或2所述的方法,其特征在于,所述采用神经网络模型对所述第一视频图像进行超分辨率处理,得到第二视频图像,包括:  
采用所述神经网络模型对所述第一视频图像的像素点亮度分量值进行超分辨率处理,得到所述第二视频图像的像素点亮度分量值;  
所述将所述第二视频图像的像素值与所述残差值叠加,得到目标视频图像,包括:  
将所述第二视频图像的像素点亮度分量值与所述残差值叠加,得到所述目标视频图像的像素点亮度分量值。
4. 如权利要求3所述的方法,其特征在于,所述方法还包括:  
对所述第一视频图像的像素点色度分量值进行插值处理,得到所述目标视频图像的像素点色度分量值。
5. 如权利要求1-4中任一项所述的方法,其特征在于,所述第一视频图像为高动态范围HDR图像或者标准动态范围SDR图像。
6. 一种视频编码方法,其特征在于,包括:  
对初始视频图像进行下采样和编码处理,得到主编码流;  
对所述主编码流进行解码,得到第一视频图像;  
采用神经网络模型对所述第一视频图像进行超分辨率处理,得到与所述初始视频图像的分辨率相同的第二视频图像;  
确定所述初始视频图像相对于所述第二视频图像的残差值;  
对所述残差值进行编码,得到辅助编码流。
7. 如权利要求6所述的方法,其特征在于,所述神经网络模型是根据训练样本图像和训练目标图像进行训练得到的,所述训练目标图像是与所述初始视频图像分辨率相同的视频图像,所述训练样本图像是对所述训练目标图像进行下采样、编码和解码处理后得到的视频图像。
8. 如权利要求6或7所述的方法,其特征在于,所述采用神经网络模型对所述第一视频图像进行超分辨率处理,得到与所述初始视频图像的分辨率相同的第二视频图像,包括:  
采用所述神经网络模型对所述第一视频图像的像素点亮度分量值进行超分辨率处理,得到所述第二视频图像的像素点亮度分量值;  
所述确定所述初始视频图像相对于所述第二视频图像的残差值,包括:  
将所述初始视频图像的像素点亮度分量值相对于所述第二视频图像的像素点亮度分量值的差值,确定为所述残差值。

9. 如权利要求6-8中任一项所述的方法,其特征在于,所述初始视频图像为高动态范围HDR图像或者标准动态范围SDR图像。

10. 一种视频解码器,其特征在于,包括:

存储器,用于存储程序;

处理器,用于执行所述存储器存储的程序,当所述处理器执行所述存储器存储的程序时,所述处理器用于:

解码主编码流,获取第一视频图像;

采用神经网络模型对所述第一视频图像进行超分辨率处理,得到第二视频图像;

解码辅助编码流,获取残差值;

将所述第二视频图像的像素值与所述残差值叠加,得到目标视频图像;

11. 如权利要求10所述的视频解码器,其特征在于,所述神经网络模型是根据训练样本图像和训练目标图像训练得到的,所述训练目标图像是与所述目标视频图像的分辨率相同的视频图像,所述训练样本图像是对所述训练目标图像进行下采样、编码和解码处理得到的视频图像。

12. 如权利要求10或11所述的视频解码器,其特征在于,所述处理器用于:

采用所述神经网络模型对所述第一视频图像的像素点亮度分量值进行超分辨率处理,得到所述第二视频图像的像素点亮度分量值;

将所述第二视频图像的像素点亮度分量值与所述残差值叠加,得到所述目标视频图像的像素点亮度分量值。

13. 如权利要求12所述的视频解码器,其特征在于,所述处理器还用于:

对所述第一视频图像的像素点色度分量值进行插值处理,得到所述目标视频图像的像素点色度分量值。

14. 如权利要求10-13中任一项所述的视频解码器,其特征在于,所述第一视频图像为高动态范围HDR图像或者标准动态范围SDR图像。

15. 一种视频编码器,其特征在于,包括:

存储器,用于存储程序;

处理器,用于执行所述存储器存储的程序,当所述处理器执行所述存储器存储的程序时,所述处理器用于:

对初始视频图像进行下采样和编码处理,得到主编码流;

对所述主编码流进行解码,得到第一视频图像;

采用神经网络模型对所述第一视频图像进行超分辨率处理,得到与所述初始视频图像的分辨率相同的第二视频图像;

确定所述初始视频图像相对于所述第二视频图像的残差值;

对所述残差值进行编码,得到辅助编码流。

16. 如权利要求15所述的视频编码器,其特征在于,所述神经网络模型是根据训练样本图像和训练目标图像进行训练得到的,所述训练目标图像是与所述初始视频图像分辨率相同的视频图像,所述训练样本图像是对所述训练目标图像进行下采样、编码和解码处理后得到的视频图像。

17. 如权利要求15或16所述的视频编码器,其特征在于,所述处理器用于:

采用所述神经网络模型对所述第一视频图像的像素点亮度分量值进行超分辨率处理，得到所述第二视频图像的像素点亮度分量值；

将所述初始视频图像的像素点亮度分量值相对于所述第二视频图像的像素点亮度分量值的差值，确定为所述残差值。

18. 如权利要求15-17中任一项所述的视频编码器，其特征在于，所述初始视频图像为高动态范围HDR图像或者标准动态范围SDR图像。

19. 一种解码视频数据的设备，其特征在于，包括：

接收器以及如权利要求10-14中任一项所述的视频解码器，所述接收器用于接收视频数据的主编码流和所述视频数据的辅助编码流，并将所述主编码流和所述辅助编码流输入到所述视频解码器进行解码。

20. 一种编码视频数据的设备，其特征在于，包括：

接收器以及如权利要求15-18中任一项所述的视频编码器，所述接收器用于接收初始视频图像，并将所述初始视频图像输入到所述视频编码器进行编码。

## 视频解码方法、编码方法以及视频解码器和编码器

### 技术领域

[0001] 本申请涉及视频编解码技术领域,并且更具体的涉及一种视频解码方法、编码方法以及视频解码器和编码器。

### 背景技术

[0002] 重构视频编码(reconstructive video coding,RVC)是一种视频内容压缩编码的技术。运用RVC技术对视频图像进行编码可以生成主编码流和辅助编码流这两路编码流。其中,主编码流为分辨率较低的视频码流,播放器可以直接解码并播放;辅助编码流的作用是协助播放器对主编码流进行分辨率的提升,以获得分辨率较高的视频图像(例如,通过解析主编码流可以获得分辨率为1080p的视频图像,接下来,再通过辅助编码流,可以得到分辨率为4K的视频图像)。

[0003] 在传统的RVC技术中,编码端在编码时一般会根据原始视频图像确定多个不同层级分辨率的视频图像,然后再确定不同层级分辨率的视频图像相对于原始视频图像的下降残差,接下来,编码端就可以对不同层级分辨率的视频图像相对于原始图像的下降残差和最低层级分辨率的视频图像(多个不同层级分辨率的视频图像中分辨率最低的图像为最低层级分辨率的视频图像)进行编码,生成码流。解码端在解码时可以通过解码最低层级分辨率的视频图像以及不同层级分辨率的视频图像相对于原始频图像的下降残差来逐渐获得最终的视频图像(最终的视频图像的分辨率与原视频图像的分辨率相同)。

[0004] 在利用上述传统RVC技术对视频图像进行编码和解码时,编码端需要获得不同层级的分辨率的视频图像,并需要对不同层级分辨率的视频图像相对于原始视频图像的下降残差进行编码,涉及的不同层级的视频图像较多。另外,解码端在解码时需要先获取最低层级分辨率的视频图像,然后根据该最低层级分辨率的视频图像和不同层级分辨率的视频图像相对于原始视频图像的下降残差进行层层叠加才能得到与原始视频图像分辨率相同的视频图像,编码和解码的过程比较复杂。

### 发明内容

[0005] 本申请提供一种视频编码方法、解码方法以及视频编码器和视频解码器,以降低编解码复杂度,提高编解码效率。

[0006] 第一方面,提供了一种视频解码方法,该方法包括:解码主编码流,获取第一视频图像;采用神经网络模型对第一视频图像进行超分辨率处理,得到第二视频图像;解码辅助编码流,获取残差值;根据第二视频图像和残差值得到目标视频图像。

[0007] 上述神经网络模型可以是预先训练好的一个神经网络模型。

[0008] 可选地,上述神经网络模型是根据训练样本图像和训练目标图像训练得到的,该训练目标图像是与目标视频图像的分辨率相同的视频图像,该训练样本图像是对训练目标图像进行下采样、编码和解码处理得到的视频图像。

[0009] 应理解,上述训练样本图像是神经网络模型的输入,上述训练目标图像是神经网络

络模型训练时对照的样本(神经网络模型在训练时输出的视频图像与训练目标图像的差异要尽可能的小)。

[0010] 在对上述神经网络模型进行训练时,可以将训练样本图像输入到神经网络模型中,然后通过对比神经网络模型的输出图像与训练目标的差异来确定神经网络模型的模型参数(主要包括神经网络模型中的各个激活函数的多个输入参数的权重值)。

[0011] 具体地,可以采用大量的训练样本图像对神经网络模型进行训练,计算神经网络模型的输出图像与目标训练图像的差异,并将神经网络模型的输出图像与训练目标图像的差异满足预设要求时的神经网络模型的模型参数的参数值确定为神经网络模型的模型参数的最终取值,这样就完成了神经网络模型的训练。完成训练后的神经网络模型就可以用于对第一视频图像进行超分辨率处理。

[0012] 可选地,上述根据第二视频图像和残差值得到目标视频图像,包括:将第二视频图像的像素值与上述残差值叠加,得到目标视频图像。

[0013] 应理解,上述残差值包括第二视频图像中的每个像素点的像素值对应残差值。上述对第二视频图像的像素值与残差值叠加是指将第二视频图像的每个像素点的像素值与其对应的残差值相加,进而得到目标视频图像中相应像素点的像素值。

[0014] 本申请中,通过神经网络模型能够更好地对分辨率较低的第一视频图像进行超分辨率处理,进而得到分辨率较高的第二视频图像,接下来,就能够通过第二视频图像以及解码得到的残差值进行一次叠加得到最终的目标视频图像,与传统方案中采用低分辨率的视频图像与多个层级的残差值进行层层叠加得到最终的视频图像的方式相比,能够简化解码流程。

[0015] 另外,本申请还能够灵活地根据需要来对神经网络模型进行优化调整,本申请方案的灵活性更高。

[0016] 结合第一方面,在第一方面的某些实现方式中,当上述残差值仅包含像素点亮度分量值对应的残差时,上述采用神经网络模型对第一视频图像进行超分辨率处理,得到第二视频图像,包括:采用神经网络模型对第一视频图像的像素点亮度分量值进行超分辨率处理,得到第二视频图像的像素点亮度分量值;上述将第二视频图像的像素值与残差值叠加,得到目标视频图像,包括:将第二视频图像的像素点亮度分量值与残差值叠加,得到目标视频图像的像素点亮度分量值。

[0017] 本申请中,神经网络模型仅对第一视频图像的像素点亮度分量值进行处理,能够减少神经网络模型进行超分辨处理时的计算量,提高视频解码的效率。

[0018] 具体地,相对于图像的色度,人眼对图像的亮度更加敏感。因此,本申请中可以采用神经网络模型仅对第一视频图像的像素点亮度分量值进行超分辨率处理,而第二视频图像的像素点的色度分量值可以通过传统的插值算法计算得到的,使得本申请能够在保证视觉体验的同时,减少视频解码的计算复杂度。

[0019] 结合第一方面,在第一方面的某些实现方式中,上述方法还包括:对第一视频图像的像素点色度分量值进行插值处理,得到目标视频图像的像素点色度分量值。

[0020] 可选地,当上述残差值包含像素点亮度分量值和像素点色度分量值对应的残差时,上述采用神经网络模型对第一视频图像进行超分辨率处理,得到第二视频图像,包括:采用神经网络模型对第一视频图像的像素点亮度分量值和像素点色度分量值进行超分辨

率处理,得到第二视频图像的像素点亮度分量值和像素点色度分量值。

[0021] 应理解,在本申请中,既可以采用神经网络模型来计算得到第二视频图像的像素点亮度分量值和像素点色度分量值,也可以只采用神经模型来计算得到第二视频图像的像素点亮度分量值,而采用其他的方法(如插值方法)得到第二视频图像的像素点色度分量值。

[0022] 上述像素点亮度分量值可以称为像素点的亮度分量值,上述像素点色度分量值可以称为像素点的色度分量值。为了描述方便,本申请统一称为像素点亮度分量值和像素点色度分量值。

[0023] 图像除了可以采用RGB格式表示之外,还可以采用YUV格式来表示,其中,Y表示明亮度(Luminance或Luma),也就是灰度值,U表示色度(Chrominance),V表示浓度(Chroma)。

[0024] 上述像素点亮度分量值可以表示为Y,上述像素点色度分量值可以包括U和V。

[0025] 结合第一方面,在第一方面的某些实现方式中,第一视频图像为高动态范围HDR图像或者标准动态范围SDR图像。

[0026] 本申请的解码方法不仅适用于标准动态范围(standard dynamic range,SDR)图像,也适用于高动态范围(high dynamic range image,HDR)图像。

[0027] 第二方面,提供了一种视频编码方法,该方法包括:对初始视频图像进行下采样和编码处理,得到主编码流;对主编码流进行解码,得到第一视频图像;采用神经网络模型对第一视频图像进行超分辨率处理,得到与初始视频图像的分辨率相同的第二视频图像;确定初始视频图像相对于第二视频图像的残差值;对残差值进行编码,得到辅助编码流。

[0028] 可选地,上述神经网络模型是根据训练样本图像和训练目标图像进行训练得到的,训练目标图像是与初始视频图像分辨率相同的视频图像,训练样本图像是对训练目标图像进行下采样、编码和解码处理后得到的视频图像。

[0029] 应理解,第二方面中的视频编码方法可以是与第一方面中的视频解码方法相对应的方法,上述第一方面中的视频解码方法解码的主编码流和辅助编码流可以是上述第二方面中的视频编码方法生成的。

[0030] 上述第一方面中的目标视频图像可以与上述第二方面中的初始视频图像的分辨率相同。

[0031] 应理解,上述第二方面中的目标视频图像的图像内容与初始视频图像的图像内容基本一致(由于在编码和解码过程中有一些损失,因此,目标视频图像相对于初始视频图像可能有一定的偏差)。

[0032] 上述第二方面中的神经网络模型的模型参数可以与第一方面中的神经网络模型的模型参数一致。上述第一方面中对神经网络模型的限定和解释同样适用于第二方面中的神经网络模型。

[0033] 可选地,上述确定初始视频图像相对于第二视频图像的残差值,包括:确定初始视频图像的像素值相对于第二视频图像的像素值的残差值。

[0034] 具体地,上述残差值可以是初始视频图像中的像素点的像素值与第二视频图像中对应的像素点的像素值的差值。

[0035] 本申请中,在生成主编码流之后,能够通过神经网络模型对主编码流解码得到的视频图像进行超分辨率处理,从而得到与初始视频图像分辨率相同的视频图像,接下来,就

可以直接生成初始视频图像相对于超分辨率处理得到的视频图像的残差值,并根据残差值生成辅助码流,与传统编码方案中需要根据初始视频图像得到多层分辨率等级不同的视频图像以及对应等级的残差值的方式相比,能够简化编码的流程。此外,相对于现有技术中需要向解码端传输多层分辨率等级不同的视频图像相比,

[0036] 结合第二方面,在第二方面的某些实现方式中,采用神经网络模型对第一视频图像进行超分辨率处理,得到与初始视频图像的分辨率相同的第二视频图像,包括:采用神经网络模型对第一视频图像的像素点亮度分量值进行超分辨率处理,得到第二视频图像的像素点亮度分量值;确定初始视频图像相对于第二视频图像的残差值,包括:将初始视频图像的像素点亮度分量值相对于第二视频图像的像素点亮度分量值的差值,确定为残差值。

[0037] 本申请中,神经网络模型仅对第一视频图像的像素点亮度分量值进行处理,能够减少神经网络模型进行超分辨率处理时的计算量,提高视频解码的效率。

[0038] 应理解,相对于色度,人眼对图像的亮度更加敏感。因此,本申请中可以采用神经网络模型仅对第一视频图像的像素点亮度分量值进行超分辨率处理,而第二视频图像的像素点的色度分量值可以通过传统的插值算法计算得到的,使得本申请能够在保证视觉体验的同时,减少视频解码的计算复杂度。

[0039] 可选地,上述采用神经网络模型对第一视频图像进行超分辨率处理,得到第二视频图像,包括:采用神经网络模型对第一视频图像的像素点亮度分量值和像素点色度分量值进行超分辨率处理,得到第二视频图像的像素点亮度分量值和像素点色度分量值。

[0040] 可选地,确定初始视频图像相对于第二视频图像的残差值,包括:将初始视频图像的像素点亮度分量值和像素点色度分量值分别相对于第二视频图像的像素点亮度分量值和像素点色度分量值的差值,确定为上述残差值。

[0041] 在本申请中,既可以采用神经网络模型来计算得到第二视频图像的像素点亮度分量值和像素点色度分量值,也可以只采用神经模型来计算得到第二视频图像的像素点亮度分量值,而采用其他的方法(如插值方法)得到第二视频图像的像素点色度分量值。

[0042] 结合第二方面,在第二方面的某些实现方式中,初始视频图像为高动态范围HDR图像或者标准动态范围SDR图像。

[0043] 第三方面,提供了一种视频解码方法,该方法包括:解码主编码流,获取第一视频图像,第一视频图像为标准动态范围SDR视频图像;采用神经网络模型对第一视频图像进行处理,得到第二视频图像,第二视频图像为高动态范围HDR视频图像,且第二视频图像的分辨率大于第一视频图像的分辨率;解码辅助编码流,获取残差值;将第二视频图像的像素值与残差值叠加,得到目标视频图像。

[0044] 其中,上述神经网络模型是根据训练样本图像和训练目标图像训练得到的,训练目标图像是与目标视频图像的分辨率相同的视频图像,训练样本图像是对训练目标图像进行下采样、编码和解码处理得到的视频图像。

[0045] 第三方面中的神经网络模型与第一方面中的神经网络模型类似,第一方面中对神经网络相关的限定和解释也适用于第三方面中的神经网络模型。

[0046] 本申请中,当解码得到的是SDR视频图像时,通过神经网络模型进行处理,能够得到分辨率更高的HDR视频图像,接下来,就能够通过第二视频图像以及解码得到的残差值进行一次叠加得到最终的目标视频图像,本申请能够适用于仅支持SDR视频图像编码和解码



的设备。

[0047] 另外,本申请与传统方案中采用低分辨率的视频图像与多个层级的残差值进行层层叠加得到最终的视频图像的方式相比,能够简化解码流程。本申请还能够灵活地根据需要来对神经网络模型进行优化调整,本申请方案的灵活性更高。

[0048] 结合第三方面,在第三方面的某些实现方式中,采用神经网络模型对第一视频图像进行处理,得到第二视频图像,包括:采用神经网络模型对第一视频图像进行超分辨率处理和反向色调映射处理,得到第二视频图像。

[0049] 其中,超分辨率处理用于提升第一视频图像的分辨率,用于得到分辨率较高的视频图像,反向色调映射处理用于提升分辨率较高的视频图像的像素精度,以得到第二视频图像。

[0050] 结合第三方面,在第三方面的某些实现方式中,当上述残差值仅包含像素点亮度分量值对应的残差时,采用神经网络模型对第一视频图像进行处理,得到第二视频图像,包括:采用神经网络模型对第一视频图像的像素点亮度分量值进行处理,得到第二视频图像的像素点亮度分量值;将第二视频图像的像素值与残差值叠加,得到目标视频图像,包括:将第二视频图像的像素点亮度分量值与残差值叠加,得到目标视频图像的像素点亮度分量值。

[0051] 本申请中,神经网络模型仅对第一视频图像的像素点亮度分量值进行处理,能够减少神经网络模型进行超分辨处理时的计算量,提高视频解码的效率。

[0052] 应理解,相对于色度,人眼对图像的亮度更加敏感。因此,本申请中可以采用神经网络模型仅对第一视频图像的像素点亮度分量值进行超分辨率处理,而第二视频图像的像素点的色度分量值可以通过传统的插值算法计算得到的,使得本申请能够在保证视觉体验的同时,减少视频解码的计算复杂度。

[0053] 结合第三方面,在第三方面的某些实现方式中,上述方法还包括:对第一视频图像的像素点色度分量值进行插值处理,得到目标视频图像的像素点色度分量值。

[0054] 可选地,当上述残差值包含像素点亮度分量值和像素点色度分量值对应的残差时,上述采用神经网络模型对第一视频图像进行超分辨率处理,得到第二视频图像,包括:采用神经网络模型对第一视频图像的像素点亮度分量值和像素点色度分量值进行超分辨率处理,得到第二视频图像的像素点亮度分量值和像素点色度分量值。

[0055] 在本申请中,既可以采用神经网络模型来计算得到第二视频图像的像素点亮度分量值和像素点色度分量值,也可以只采用神经模型来计算得到第二视频图像的像素点亮度分量值,而采用其他的方法(如插值方法)得到第二视频图像的像素点色度分量值。

[0056] 第四方面,提供了一种视频编码方法,该方法包括:对初始视频图像进行处理,得到处理后的视频图像,其中,初始视频图像为高动态范围HDR视频图像,处理后的视频图像为标准动态范围SDR视频图像;对处理后的视频图像进行编码,得到主编码流;对主编码流进行解码,得到第一视频图像;采用神经网络模型对第一视频图像进行处理,得到第二视频图像,第二视频图像为HDR视频图像,且第二视频图像的分辨率与初始视频图像的分辨率相同;确定初始视频图像相对于第二视频图像的残差值;对残差值进行编码,得到辅助编码流。

[0057] 其中,上述神经网络模型是根据训练样本图像和训练目标图像进行训练得到的,

训练目标图像是与初始视频图像分辨率相同的视频图像,训练样本图像是对训练目标图像进行下采样、编码和解码处理后得到的视频图像。

[0058] 应理解,第四方面中的视频编码方法可以是与第二方面中的视频解码方法相对应的方法,上述第三方面中的视频解码方法解码的主编码流和辅助编码流可以是上述第四方面中的视频编码方法生成的。

[0059] 上述第三方面中的目标视频图像可以与上述第四方面中的初始视频图像的分辨率相同。

[0060] 上述第四方面中的神经网络模型的模型参数可以与第一方面中的神经网络模型的模型参数一致。上述第一方面中对神经网络模型的限定和解释同样适用于第四方面中的神经网络模型。

[0061] 可选地,上述确定初始视频图像相对于第二视频图像的残差值,包括:确定初始视频图像的像素值相对于第二视频图像的像素值的残差值。

[0062] 具体地,上述残差值可以是初始视频图像中的像素点的像素值与第二视频图像中对应的像素点的像素值的差值。

[0063] 本申请中,对于HDR图像,可以先将该HDR图像转化为SDR图像,然后再进行后续的编码处理,能够适用于仅支持对SDR图像进行编码的设备。

[0064] 进一步的,在生成主编码流之后,能够通过神经网络模型对主编码流解码得到的视频图像进行超分辨率处理,从而得到与初始视频图像分辨率相同的视频图像,接下来,就可以直接生成初始视频图像相对于超分辨率处理得到的视频图像的残差值,并根据残差值生成辅助码流,与传统编码方案中需要根据初始视频图像得到多层分辨率等级不同的视频图像以及对应等级的残差值的方式相比,能够简化编码的流程。

[0065] 结合第四方面,在第四方面的某些实现方式中,采用神经网络模型对初始视频图像进行处理,得到处理后的视频图像,包括:采用神经网络模型对初始视频图像进行下采样和色调映射处理,得到处理后的视频图像。

[0066] 其中,下采样用于降低初始视频图像的分辨率,得到分辨率较低的视频图像,色调映射处理用于对分辨率较低的视频图像进行处理,以提升该分辨率较低的视频图像的像素精度,得到处理后的视频图像。

[0067] 结合第四方面,在第四方面的某些实现方式中,采用神经网络模型对第一视频图像进行处理,得到第二视频图像,包括:采用神经网络模型对第一视频图像进行超分辨率处理和反向色调映射处理,得到第二视频图像。

[0068] 其中,超分辨率处理用于提升第一视频图像的分辨率,用于得到分辨率较高的视频图像,反向色调映射处理用于提升分辨率较高的视频图像的像素精度,以得到第二视频图像。

[0069] 结合第四方面,在第四方面的某些实现方式中,采用神经网络模型对第一视频图像进行处理,得到第二视频图像,包括:采用神经网络模型对第一视频图像的像素点亮度分量值进行处理,得到第二视频图像的像素点亮度分量值;确定初始视频图像相对于第二视频图像的残差值,包括:将初始视频图像的像素点亮度分量值相对于第二视频图像的像素点亮度分量值的差值,确定为上述残差值。

[0070] 可选地,上述采用神经网络模型对第一视频图像进行处理,得到第二视频图像,包

括:采用神经网络模型对第一视频图像的像素点亮度分量值和像素点色度分量值进行处理,得到第二视频图像的像素点亮度分量值和像素点色度分量值。

[0071] 上述采用神经网络模型对第一视频图像的像素点亮度分量值和像素点色度分量值进行处理,具体包括:采用神经网络模型对第一视频图像的像素点亮度分量值和像素点色度分量值进行超分辨率处理和反向色调映射处理。

[0072] 可选地,确定初始视频图像相对于第二视频图像的残差值,包括:将初始视频图像的像素点亮度分量值和像素点色度分量值分别相对于第二视频图像的像素点亮度分量值和像素点色度分量值的差值,确定为上述残差值。

[0073] 在本申请中,既可以采用神经网络模型来计算得到第二视频图像的像素点亮度分量值和像素点色度分量值,也可以只采用神经模型来计算得到第二视频图像的像素点亮度分量值,而采用其他的方法(如插值方法)得到第二视频图像的像素点色度分量值。

[0074] 第五方面,提供了一种解码视频数据的设备,该设备包括:接收器,用于接收视频数据的主编码流和辅助编码流输入,并将所述主编码流和辅助编码流输入到视频解码器进行解码;视频解码器,用于实施第一方面或第三方面的任意一种方法的部分或全部步骤。

[0075] 第六方面,提供了一种编码视频数据的设备,该设备包括:接收器,用于接收初始视频图像,并将所述初始视频图像输入到视频编码器进行编码;视频编码器,用于实施第二方面或第四方面的任意一种方法的部分或全部步骤。

[0076] 第七方面,本申请实施例提供一种解码视频数据的设备,包括:存储器和处理器,所述处理器调用存储在所述存储器中的程序代码以执行第一方面或第三方面的任意一种方法的部分或全部步骤。

[0077] 第八方面,本申请实施例提供一种编码视频数据的设备,包括:存储器和处理器,所述处理器调用存储在所述存储器中的程序代码以执行第二方面或第四方面的任意一种方法的部分或全部步骤。

[0078] 可选地,上述存储器为非易失性存储器。

[0079] 可选地,上述存储器与处理器互相耦合在一起。

[0080] 第九方面,本申请实施例提供一种计算机可读存储介质,所述计算机可读存储介质存储了程序代码,其中,所述程序代码包括用于执行第一方面、第二方面、第三方面以及第四方面中的任意一个方面中的方法的部分或全部步骤的指令。

[0081] 第十方面,本申请实施例提供一种计算机程序产品,当所述计算机程序产品在计算机上运行时,使得所述计算机执行第一方面、第二方面、第三方面以及第四方面中的任意一个方面中的方法的部分或全部步骤。

## 附图说明

[0082] 图1是用于实现本申请实施例的视频编码解码系统实例的示意性框图;

[0083] 图2是用于实现本申请实施例的视频编码器实例的示意性结构框图;

[0084] 图3是用于实现本申请实施例的视频解码器实例的示意性结构框图;

[0085] 图4是用于实现本申请实施例的视频处理系统实例的示意性结构框图;

[0086] 图5是用于实现本申请实施例的视频处理设备实例的示意性结构框图;

[0087] 图6是用于实现本申请实施例的编码装置或者解码装置实例的示意性框图;

- [0088] 图7是本申请实施例的视频编码和视频解码的流程的示意图；
- [0089] 图8是本申请实施例的视频解码方法的示意性流程图；
- [0090] 图9是本申请实施例的视频解码方法的流程图；
- [0091] 图10是本申请实施例训练神经网络模型的示意性流程图；
- [0092] 图11是本申请实施例的训练神经网络模型的训练系统的示意性框图；
- [0093] 图12是本申请实施例的采用神经网络模型对第一视频图像进行处理的示意图；
- [0094] 图13是本申请实施例的卷积神经网络的结构示意图；
- [0095] 图14是本申请实施例的视频编码方法的示意性流程图；
- [0096] 图15是本申请实施例的视频编码方法的示意性流程图；
- [0097] 图16是本申请实施例的视频编码和视频解码的流程的示意图；
- [0098] 图17是本申请实施例的视频解码方法的示意性流程图；
- [0099] 图18是本申请实施例的视频编码方法的示意性流程图；
- [0100] 图19是本申请实施例的视频解码方法的流程图；
- [0101] 图20是本申请实施例的视频编码方法的流程图；
- [0102] 图21是本申请实施例的视频编码和视频解码的流程的示意图；
- [0103] 图22是本申请实施例的视频解码器的示意性框图；
- [0104] 图23是本申请实施例的视频编码器的示意性框图；
- [0105] 图24是本申请实施例的视频解码器的示意性框图；
- [0106] 图25是本申请实施例的视频编码器的示意性框图。

### 具体实施方式

[0107] 下面将结合附图,对本申请中的技术方案进行描述。

[0108] 本申请描述的实施例不仅可以在本文中描述的情况中应用,还可以在其它方面中使用,并且可以包括附图中未描绘的结构或逻辑变化。本申请中的详细描述不应以限制性的意义来理解,本申请的范围应由所附权利要求书界定。

[0109] 本申请所描述的方法揭示的内容可以同样适用于执行所述方法的对应设备或系统,反之亦然。具体地,在本申请中,如果描述了某个方法的相应的步骤,则对应的设备可以包含执行该方法对应步骤的单元或者模块,即使附图中未明确描述或说明对应的设备包括相应的单元或者模块。

[0110] 本申请实施例所涉及的视频编码和视频解码不仅可以应用于现有的视频编码标准中(例如,H.264、高性能视频编码(high efficiency video coding,HEVC)等标准),还可以应用于未来的视频编码标准中(如H.266标准)。本申请的实施方式部分使用的术语仅用于对本申请的具体实施例进行解释,而非用于限定本申请。

[0111] 为了便于后续更好地理解本申请实施例的方案,下面先对视频编码和视频解码过程中可能涉及的一些基本概念进行简单介绍。

[0112] 视频编码通常是指处理形成视频或视频序列的图片序列。在视频编码领域,术语“图片(picture)”、“帧(frame)”或“图像(image)”可以用作同义词。视频编码在源侧执行,通常包括处理(例如,通过压缩)原始视频图片以减少表示该视频图片所需的数据量,从而更高效地存储和/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

重构视频图片。在本申请实施例中,编码部分和解码部分的组合也称为编解码(编码和解码)。

[0113] 视频序列包括一系列图像 (picture), 图像可以被进一步划分为切片 (slice), 切片还可以再被划分为块 (block)。视频编码时, 一般以块为单位进行编码处理, 在一些新的视频编码标准中, 块的概念被进一步扩展。比如, 在H.264标准中有宏块 (macroblock, MB), 宏块可进一步划分成多个可用于预测编码的预测块 (partition)。在高性能视频编码 (high efficiency video coding, HEVC) 标准中, 编码时采用的是编码单元 (coding unit, CU), 预测单元 (prediction unit, PU) 和变换单元 (transform unit, TU) 等基本概念, 这些基本单元还可以基于树结构划分方式进行进一步的划分。例如, CU可以按照二叉树划分为更小的CU, 而更小的CU还可以继续划分, 从而形成一种二叉树结构, CU是对编码图像进行划分和编码的基本单元。对于PU和TU也有类似的树结构, PU可以对应预测块, 是预测编码的基本单元。对CU按照划分模式进一步划分成多个PU。TU可以对应变换块, 是对预测残差进行变换的基本单元。然而, 无论CU, PU还是TU, 本质上都属于块 (或称图像块) 的概念。

[0114] 为了更好地理解本申请实施例的视频编码方法和视频解码方法的应用场景, 下面先结合图1对本申请实施例的可能应用的系统架构进行描述。

[0115] 图1是用于实现本申请实施例的视频编码解码系统实例的示意性框图。

[0116] 如图1所示, 视频编码及解码系统10包括源设备12和目的地设备14, 源设备12可以输出编码后的视频数据, 源设备12可以被称为视频编码装置。目的地设备14可对由源设备12输出的编码后的视频数据进行解码, 目的地设备14可被称为视频解码装置。

[0117] 源设备12、目的地设备14可以包含一个或多个处理器以及耦合到所述一个或多个处理器的存储器。所述存储器可包含但不限于只读存储器 (read-only memory, ROM)、随机存取存储器 (random access memory, RAM)、可擦写可编程只读存储器 (erasable programmable read-only memory, EPROM)、快闪存储器或可用于以可由计算机存取的指令或数据结构的形式存储所要的程序代码的任何其它媒体, 如本文所描述。

[0118] 源设备12和目的地设备14可以包括各种装置, 包含桌上型计算机、移动计算装置、笔记型 (例如, 膝上型) 计算机、平板计算机、机顶盒、例如所谓的“智能”电话等电话手持机、电视机、相机、显示装置、数字媒体播放器、视频游戏控制台、车载计算机、无线通信设备或其类似者。

[0119] 如图1所示, 源设备12和目的地设备14既可以是彼此独立的设备 (此时视频编码及解码系统10由两个独立的设备源设备12和目的地设备14构成), 或者, 源设备12和目的地设备14仅仅是视频编码及解码系统10这个设备的组成部分 (此时, 视频编码及解码系统10为单独的设备, 源设备12和目的地设备14构成该设备的不同模块)。

[0120] 上述源设备12和目的地设备14之间可通过链路13进行通信连接, 目的地设备14可经由链路13从源设备12接收经编码视频数据。链路13可包括能够将经编码视频数据从源设备12移动到目的地设备14的一个或多个媒体或装置。在一个实例中, 链路13可包括使得源设备12能够实时将编码视频数据直接发送到目的地设备14的一个或多个通信媒体。在此实例中, 源设备12可根据通信标准 (例如无线通信协议) 来调制经编码视频数据, 且可将经调制的视频数据发射到目的地设备14。

[0121] 上述一个或多个通信媒体可包含无线和/或有线通信媒体, 例如射频 (radio

frequency, RF) 频谱或一个或多个物理传输线。上述一个或多个通信媒体可形成基于分组的网络的一部分, 基于分组的网络例如为局域网、广域网或全球网络(例如, 因特网)。上述一个或多个通信媒体可包含路由器、交换机、基站或促进从源设备12到目的地设备14的通信的其它设备。

[0122] 源设备12可以包括编码器20。可选地, 源设备12还可以包括图片源16、图片预处理器18、以及通信接口22。在具体实现时, 编码器20、图片源16、图片预处理器18、以及通信接口22可以是源设备12中的硬件部件, 也可以是源设备12中的软件程序。下面对源设备12中的各个模块或者单元进行描述。

[0123] 图片源16, 可以包括任何类别的图片捕获设备(或者可以是任何类别的图片捕获设备)。这里的图片捕获设备可以是用于捕获现实世界图片和/或任何类别的图片或评论(对于屏幕内容编码, 屏幕上的一些文字也认为是待编码的图片或图像的一部分)的生成设备, 例如, 用于生成计算机动画图片的计算机图形处理器, 或用于获取和/或提供现实世界图片、计算机动画图片(例如, 屏幕内容、虚拟现实(virtual reality, VR)图片)的任何类别设备, 和/或其任何组合(例如, 实景(augmented reality, AR)图片)。

[0124] 图片预处理器18, 用于接收原始图片数据17并对原始图片数据17执行预处理, 以获取经预处理的图片19或经预处理的图片数据19。例如, 图片预处理器18执行的预处理可以包括整修、色彩格式转换(例如, 从RGB格式转换为YUV格式)、调色或去噪。

[0125] 编码器20(或称为视频编码器20), 用于接收经预处理的图片数据19, 采用相关预测模式(例如, 帧间预测模式, 帧内预测模式)对经预处理的图片数据19进行编码处理, 从而得到编码后的图片数据21。

[0126] 在一些实施例中, 编码器20可以用于执行后文所描述的编码方法的各个步骤, 以实现本申请所描述的视频编码方法在编码侧的应用。

[0127] 通信接口22, 可用于接收编码后的图片数据21, 并可通过链路13将编码后的图片数据21传输至目的地设备14或任何其它设备(如存储器), 所述其它设备可为任何用于解码或存储的设备。通信接口22可例如用于将编码后的图片数据21封装成合适的格式, 例如数据包, 以在链路13上传输。

[0128] 目的地设备14包括解码器30。可选地, 目的地设备14还可以包括通信接口28、图片后处理器32和显示设备34。下面对目的地设备14中的各个模块或者单元进行描述。

[0129] 通信接口28, 可用于从源设备12或任何其它源接收编码后的图片数据21, 所述任何其它源例如为存储设备, 存储设备例如为经编码图片数据存储设备。通信接口28可以用于藉由源设备12和目的地设备14之间的链路13或藉由任何类别的网络传输或接收编码后的图片数据21, 链路13例如为直接有线或无线连接, 任何类别的网络例如为有线或无线网络或其任何组合, 或任何类别的私网和公网, 或其任何组合。通信接口28可以例如用于解封通信接口22所传输的数据包以获取编码后的图片数据21。

[0130] 通信接口28和通信接口22都可以配置为单向通信接口或者双向通信接口, 以及可以用于例如发送和接收消息来建立连接、确认和交换任何其它与通信链路和/或例如经编码图片数据传输的数据传输有关的信息。

[0131] 解码器30(或称为解码器30), 用于接收编码后的图片数据21并提供经解码图片数据31或经解码图片31(下文将进一步基于图3或图4或图5描述解码器30的结构细节)。在一

些实施例中,解码器30可以用于执行后文所描述的各个实施例,以实现本申请所描述的色度块预测方法在解码侧的应用。

[0132] 图片后处理器32,用于对经解码图片数据31(经过解码得到的图片数据31,也称为经重构图片数据)执行后处理,以得到图片数据33(图片数据33是图片后处理器32对图片数据31进行后处理之后得到的图片数据)。图片后处理器32执行的后处理可以包括:色彩格式转换(例如,从YUV格式转换为RGB格式)、调色、整修或重采样,或任何其它处理,还可用于图片数据33传输至显示设备34。

[0133] 显示设备34,用于接收图片数据33以向例如用户或观看者显示图片。显示设备34可以为或可以包括任何类别的用于呈现经重构图片的显示器,例如,集成的或外部的显示器或监视器。另外,显示器可以包括液晶显示器(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LED显示器、硅基液晶(liquid crystal on silicon,LCoS)、数字光处理器(digital light processor,DLP)或任何类别的其它显示器。

[0134] 本领域技术人员基于描述明显可知,不同单元的功能性或图1所示的源设备12和/或目的地设备14的功能性的存在和(准确)划分可能根据实际设备和应用有所不同。源设备12和目的地设备14可以包括各种设备中的任一个,包含任何类别的手持或静止设备,例如,笔记本或膝上型计算机、移动电话、智能手机、平板或平板计算机、台式计算机、机顶盒、电视机、相机、车载设备、显示设备、数字媒体播放器、视频游戏控制台、视频流式传输设备(例如内容服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,并可以不使用或使用任何类别的操作系统。

[0135] 编码器20和解码器30都可以实施为各种合适电路中的任一个,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、硬件或其任何组合。如果部分地以软件实施所述技术,则设备可将软件的指令存储于合适的非暂时性计算机可读存储介质中,且可使用一个或多个处理器以硬件执行指令从而执行本申请的技术。前述内容(包含硬件、软件、硬件与软件的组合等)中的任一者可视为一个或多个处理器。

[0136] 在一些情况下,图1中所示视频编码及解码系统10仅为示例,本申请的技术可以适用于不必包含编码和解码设备之间的任何数据通信的视频编码设置(例如,视频编码或视频解码)。在其它实例中,数据可从本地存储器检索、在网络上流式传输等。视频编码设备可以对数据进行编码并且将数据存储到存储器,和/或视频解码设备可以从存储器检索数据并且对数据进行解码。在一些实例中,由并不彼此通信而是仅编码数据到存储器和/或从存储器检索数据且解码数据的设备执行编码和解码。

[0137] 图2是用于实现本申请实施例的视频编码器实例的示意性结构框图。

[0138] 图2所示的编码器20可以称为混合型视频编码器或根据混合型视频编解码器的视频编码器。在图2的实例中,编码器20包括残差计算单元204、变换处理单元206、量化单元208、逆量化单元210、逆变换处理单元212、重构单元214、缓冲器216、环路滤波器单元220、经解码图片缓冲器(decoded picture buffer,DPB) 230、预测处理单元260和熵编码单元270。

[0139] 其中,预测处理单元260可以包含帧间预测单元244、帧内预测单元254和模式选择单元262。帧间预测单元244可以包含运动估计单元和运动补偿单元(图2中未示出)。

[0140] 编码器20通过例如输入202,接收图片201或图片201的图像块203,例如,形成视频或视频序列的图片序列中的图片。图像块203也可以称为当前图片块或待编码图片块,图片201可以称为当前图片或待编码图片(尤其是在视频编码中将当前图片与其它图片区分开时)。

[0141] 编码器20的实施例可以包括分割单元(图2中未示出),用于将图片201分割成多个例如图像块203的块,通常分割成多个不重叠的块。分割单元可以用于对视频序列中所有图片使用相同的块大小以及定义块大小的对应栅格,或用于在图片或子集或图片群组之间更改块大小,并将每个图片分割成对应的块。

[0142] 在一个实例中,编码器20的预测处理单元260可以用于执行上述分割技术的任何组合。

[0143] 残差计算单元204用于基于图片图像块203和预测块265(计算残差块205,例如,通过逐样本(逐像素)将图片图像块203的样本值减去预测块265的样本值,以在样本域中获取残差块205)。

[0144] 变换处理单元206用于在残差块205的样本值上应用例如离散余弦变换(discrete cosine transform,DCT)或离散正弦变换(discrete sine transform,DST)的变换,以在变换域中获取变换系数207。变换系数207也可以称为变换残差系数,并在变换域中表示残差块205。

[0145] 变换处理单元206可以用于应用DCT/DST的整数近似值,例如为HEVC/H.265指定的变换。与正交DCT变换相比,这种整数近似值通常由某一因子按比例缩放。为了维持经正变换和逆变换处理的残差块的范数,应用额外比例缩放因子作为变换过程的一部分。比例缩放因子通常是基于某些约束条件选择的,例如,比例缩放因子是用于移位运算的2的幂、变换系数的位深度、准确性和实施成本之间的权衡等。例如,在解码器30侧通过例如逆变换处理单元212为逆变换(以及在编码器20侧通过例如逆变换处理单元212为对应逆变换)指定具体比例缩放因子,以及相应地,可以在编码器20侧通过变换处理单元206为正变换指定对应比例缩放因子。

[0146] 量化单元208用于例如通过应用标量量化或向量量化来量化变换系数207,以获取经量化变换系数209。经量化变换系数209也可以称为经量化残差系数209。量化过程可以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,可以应用不同的标度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。可以通过量化参数(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。

[0147] 逆量化单元210用于在经量化系数上应用量化单元208的逆量化,以获取经反量化系数211,例如,基于或使用与量化单元208相同的量化步长,应用量化单元208应用的量化方案的逆量化方案。经反量化系数211也可以称为经反量化残差系数211,对应于变换系数207,虽然由于量化造成的损耗通常与变换系数不相同。



[0148] 逆变换处理单元212用于应用变换处理单元206应用的变换的逆变换,例如,逆离散余弦变换(discrete cosine transform,DCT)或逆离散正弦变换(discrete sine transform,DST),以在样本域中获取逆变换块213。逆变换块213也可以称为逆变换经反量化块213或逆变换残差块213。

[0149] 重构单元214(例如,求和器214)用于将逆变换块213(即经重构残差块213)添加至预测块265,以在样本域中获取经重构块215,例如,将经重构残差块213的样本值与预测块265的样本值相加。

[0150] 环路滤波器单元220(或简称“环路滤波器”220)用于对经重构块215进行滤波以获取经滤波块221,从而顺利进行像素转变或提高视频质量。环路滤波器单元220旨在表示一个或多个环路滤波器,例如去块滤波器、样本自适应偏移(sample-adaptive offset,SAO)滤波器或其它滤波器,例如双边滤波器、自适应环路滤波器(adaptive loop filter,ALF),或锐化或平滑滤波器,或协同滤波器。尽管环路滤波器单元220在图2中示出为环内滤波器,但在其它配置中,环路滤波器单元220可实施为环后滤波器。经滤波块221也可以称为经滤波的经重构块221。经解码图片缓冲器230可以在环路滤波器单元220对经重构编码块执行滤波操作之后存储经重构编码块。

[0151] 编码器20(对应地,环路滤波器单元220)的实施例可以用于输出环路滤波器参数(例如,样本自适应偏移信息),例如,直接输出或由熵编码单元270或任何其它熵编码单元熵编码后输出,例如使得解码器30可以接收并应用相同的环路滤波器参数用于解码。

[0152] 解码图片缓冲器(decoded picture buffer,DPB)230可以为存储参考图片数据供编码器20编码视频数据之用的参考图片存储器。DPB 230可由多种存储器设备中的任一个形成,例如动态随机存储器(dynamic random access memory,DRAM)(包含同步DRAM(synchronous DRAM,SDRAM)、磁阻式RAM(magnetoresistive RAM,MRAM)、电阻式RAM(resistive RAM,RRAM)或其它类型的存储器设备。可以由同一存储器设备或单独的存储器设备提供DPB 230和缓冲器216。在某一实例中,经解码图片缓冲器(decoded picture buffer,DPB)230用于存储经滤波块221。经解码图片缓冲器230可以进一步用于存储同一当前图片或例如先前经重构图片的不同图片的其它先前的经滤波块,例如先前经重构和经滤波块221,以及可以提供完整的先前经重构亦即经解码图片(和对应参考块和样本)和/或部分经重构当前图片(和对应参考块和样本),例如用于帧间预测。在某一实例中,如果经重构块215无需环内滤波而得以重构,则经解码图片缓冲器(decoded picture buffer,DPB)230用于存储经重构块215。

[0153] 预测处理单元260,也称为块预测处理单元260,用于接收或获取图像块203(当前图片201的当前图像块203)和经重构图片数据,例如来自缓冲器216的同一(当前)图片的参考样本和/或来自经解码图片缓冲器230的一个或多个先前经解码图片的参考图片数据231,以及用于处理这类数据进行预测,即提供可以为经帧间预测块245或经帧内预测块255的预测块265。

[0154] 模式选择单元262可以用于选择预测模式(例如帧内或帧间预测模式)和/或对应的用作预测块265的预测块245或255,以计算残差块205和重构经重构块215。

[0155] 模式选择单元262的实施例可以用于选择预测模式(例如,从预测处理单元260所支持的那些预测模式中选择),所述预测模式提供最佳匹配或者说最小残差(最小残差意味

着传输或存储中更好的压缩),或提供最小信令开销(最小信令开销意味着传输或存储中更好的压缩),或同时考虑或平衡以上两者。模式选择单元262可以用于基于码率失真优化(rate distortion optimization,RDO)确定预测模式,即选择提供最小码率失真优化的预测模式,或选择相关码率失真至少满足预测模式选择标准的预测模式。

[0156] 图3是用于实现本申请实施例的视频解码器实例的示意性结构框图。

[0157] 如图3所示,视频解码器30用于接收例如由编码器20编码的经编码图片数据21(例如,经编码比特流),以获取经解码图片231。在解码过程期间,视频解码器30从视频编码器20接收视频数据,例如表示经编码视频条带的图片块的经编码视频比特流及相关联的语法元素。

[0158] 在图3的实例中,解码器30包括熵解码单元304、逆量化单元310、逆变换处理单元312、重构单元314(例如求和器314)、缓冲器316、环路滤波器320、经解码图片缓冲器330以及预测处理单元360。预测处理单元360可以包含帧间预测单元344、帧内预测单元354和模式选择单元362。在一些实例中,视频解码器30可执行大体上与参照图2的视频编码器20描述的编码遍次互逆的解码遍次。

[0159] 熵解码单元304用于对编码后的图片数据21执行熵解码,以获取例如经量化系数309和/或经解码的编码参数(图3中未示出),例如,帧间预测、帧内预测参数、环路滤波器参数和/或其它语法元素中(经解码)的任意一个或全部。熵解码单元304进一步用于将帧间预测参数、帧内预测参数和/或其它语法元素转发至预测处理单元360。视频解码器30可接收视频条带层级和/或视频块层级的语法元素。

[0160] 逆量化单元310功能上可与逆量化单元110相同,逆变换处理单元312功能上可与逆变换处理单元212相同,重构单元314功能上可与重构单元214相同,缓冲器316功能上可与缓冲器216相同,环路滤波器320功能上可与环路滤波器220相同,经解码图片缓冲器330功能上可与经解码图片缓冲器230相同。

[0161] 预测处理单元360可以包括帧间预测单元344和帧内预测单元354,其中帧间预测单元344功能上可以类似于帧间预测单元244,帧内预测单元354功能上可以类似于帧内预测单元254。预测处理单元360通常用于执行块预测和/或从经编码数据21获取预测块365,以及从例如熵解码单元304(显式地或隐式地)接收或获取预测相关参数和/或关于所选择的预测模式的信息。

[0162] 逆量化单元310可用于逆量化(即,反量化)在比特流中提供且由熵解码单元304解码的经量化变换系数。逆量化过程可包含使用由视频编码器20针对视频条带中的每一视频块所计算的量化参数来确定应该应用的量化程度并同样确定应该应用的逆量化程度。

[0163] 逆变换处理单元312用于将逆变换(例如,逆DCT、逆整数变换或概念上类似的逆变换过程)应用于变换系数,以便在像素域中产生残差块。

[0164] 重构单元314(例如,求和器314)用于将逆变换块313(即经重构残差块313)添加到预测块365,以在样本域中获取经重构块315,例如通过将经重构残差块313的样本值与预测块365的样本值相加。

[0165] 环路滤波器单元320(在编码循环期间或在编码循环之后)用于对经重构块315进行滤波以获取经滤波块321,从而顺利进行像素转变或提高视频质量。在一个实例中,环路滤波器单元320可以用于执行下文描述的滤波技术的任意组合。环路滤波器单元320旨在表

示一个或多个环路滤波器,例如去块滤波器、样本自适应偏移(sample-adaptive offset, SAO)滤波器或其它滤波器,例如双边滤波器、自适应环路滤波器(adaptive loop filter, ALF),或锐化或平滑滤波器,或协同滤波器。尽管环路滤波器单元320在图3中示出为环内滤波器,但在其它配置中,环路滤波器单元320可实施为环后滤波器。

[0166] 随后将给定帧或图片中的经解码视频块321存储在存储用于后续运动补偿的参考图片的经解码图片缓冲器330中。

[0167] 解码器30用于例如,藉由输出332输出经解码图片31,以向用户呈现或供用户查看。

[0168] 视频解码器30的其它变型可用于对压缩的比特流进行解码。例如,解码器30可以在没有环路滤波器单元320的情况下生成输出视频流。例如,基于非变换的解码器30可以在没有针对某些块或帧的逆变换处理单元312的情况下直接逆量化残差信号。在另一实施方式中,视频解码器30可以具有组合成单个单元的逆量化单元310和逆变换处理单元312。

[0169] 具体的,在本申请实施例中,解码器30用于实现后文实施例中描述的视频解码法。

[0170] 应当理解的是,本申请中的视频编码器可以只包括视频编码器30中的部分模块,例如,本申请中的视频编码器可以包括划分单元和图像编码单元。其中,图像编码单元可以由预测单元、变换单元、量化单元和熵编码单元中的一种或者多种单元组成。

[0171] 图4是用于实现本申请实施例的视频处理系统实例的示意性结构框图。

[0172] 如图4所示,视频处理系统40可以是包含图2的编码器20和/或图3的解码器30的系统。视频处理系统40可以包含成像设备41、视频处理设备46(包括编码器20或解码器30)、天线42和显示设备45。其中,编码器20和解码器30可以采用逻辑电路来实现,逻辑电路可以包含专用集成电路(ASIC)逻辑、图形处理器、通用处理器等。

[0173] 如图4所示,成像设备41、天线42、视频处理设备46和显示设备45之间能够互相通信。

[0174] 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。成像设备用于采集原始的视频图像,具体可以为摄像机。

[0175] 在一些实例中,编码器20和解码器30可以包含图像缓冲器和图形处理单元。图形处理单元可以通信耦合至图像缓冲器。

[0176] 图5是用于实现本申请实施例的视频处理设备实例的示意性结构框图。

[0177] 图5中的视频处理设备400适于实施本文所描述的实施例。在一个实施例中,视频处理设备400可以包括视频解码器(例如图3的解码器30)或视频编码器(例如图2的编码器20)。此时,视频处理设备可以称为解码视频数据的设备或编码视频数据的设备。

[0178] 在另一个实施例中,视频处理设备400可以是上述图3的解码器30或图2的编码器20中的一个或多个组件。

[0179] 视频处理设备400包括:用于接收数据的入口端口410和接收器420,用于进行视频编码的编码器或进行视频解码的解码器,用于传输数据的发射器440和出口端口450,以及,用于存储数据的存储器460。视频处理设备400还可以包括与入口端口410、接收器单元420、发射器单元440和出口端口450耦合的光电转换组件和电光(EO)组件,用于光信号或电信号的出口或入口。

[0180] 当视频处理设备为解码视频数据的设备时,接收器420可以用于接收视频数据的主编码流和辅助编码流,并将该主编码流和辅助编码流输入到编码器。当视频处理设备为编码视频数据的设备时,接收器420还可以接收初始视频图像,并将初始视频图像输入到解码器进行处理。

[0181] 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,用于在选择性地执行这些程序时存储程序,并存储在程序执行过程中读取的指令和数据。

[0182] 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(ROM)、随机存取存储器(RAM)、随机存取存储器(ternary content-addressable memory,TCAM)和/或静态随机存取存储器(SRAM)。

[0183] 此外,上述的视频处理设备可以包括处理器,其用于处理入口端口输入的指令,来启动上面的视频编码器执行视频编码,或启动上面的视频解码器执行视频解码。

[0184] 图6是用于实现本申请实施例的编码装置或者解码装置实例的示意性框图。

[0185] 图6是根据一示例性实施例的可用作图1中的源设备12和目的地设备14中的一个或两个的装置500的简化框图。装置500可以实现本申请的技术。换言之,图6为本申请实施例的编码设备或解码设备(简称为处理设备500)的一种实现方式的示意性框图。其中,处理设备500可以包括处理器510、存储器530和总线系统550。其中,处理器和存储器通过总线系统相连,该存储器用于存储指令,该处理器用于执行该存储器存储的指令。处理设备的存储器存储程序代码,且处理器可以调用存储器中存储的程序代码执行本申请描述的各种视频编码或解码方法,尤其是各种新的图像块的划分方法。为避免重复,这里不再详细描述。

[0186] 在本申请实施例中,该处理器510可以是中央处理单元(central processing unit,CPU),该处理器510还可以是其他通用处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现成可编程门阵列(FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

[0187] 该存储器530可以包括只读存储器(ROM)设备或者随机存取存储器(RAM)设备。任何其他适宜类型的存储设备也可以用作存储器530。存储器530可以包括由处理器510使用总线550访问的代码和数据531。存储器530可以进一步包括操作系统533和应用程序535,该应用程序535包括允许处理器510执行本申请描述的视频编码或解码方法的至少一个程序。例如,应用程序535可以包括应用1至N,其进一步包括执行在本申请描述的视频编码或解码方法的视频编码或解码应用(简称视频处理应用)。

[0188] 该总线系统550除包括数据总线之外,还可以包括电源总线、控制总线和状态信号总线等。但是为了清楚说明起见,在图中将各种总线都标为总线系统550。

[0189] 可选的,处理设备500还可以包括一个或多个输出设备,诸如显示器570。在一个示例中,显示器570可以是触感显示器,其将显示器与可操作地感测触摸输入的触感单元合并。显示器570可以经由总线550连接到处理器510。

[0190] 另外,在本申请中,可以采用神经网络模型对低分辨率的视频图像进行超分辨率处理,以得到分辨率较高的视频图像。这里的神经网络可以是卷积神经网络(convolutional neuron network,CNN)、深度神经网络(deep neural network,DNN)等等,

下面对这些神经网络的基本概念以及这些神经网络相关的其他内容进行介绍。

### [0191] (1) 神经网络

[0192] 神经网络可以是由神经元组成的,神经元可以是指以 $x_s$ 和截距1为输入的运算单元,该运算单元的输出可以为:

$$[0193] \quad h_{w,b}(x) = f(W^T x) = f\left(\sum_{s=1}^n W_s x_s + b\right) \quad (1-1)$$

[0194] 其中, $s=1,2,\dots,n$ , $n$ 为大于1的自然数, $W_s$ 为 $x_s$ 的权重, $b$ 为神经单元的偏置。 $f$ 为神经单元的激活函数(activation functions),用于将非线性特性引入神经网络中,来将神经元中的输入信号转换为输出信号。该激活函数的输出信号可以作为下一层卷积层的输入。激活函数可以是sigmoid函数。神经网络是将许多个上述单一的神经元联结在一起形成的网络,即一个神经元的输出可以是另一个神经元的输入。每个神经元的输入可以与前一层的局部接受域相连,来提取局部接受域的特征,局部接受域可以是由若干个神经元组成的区域。

### [0195] (2) 深度神经网络

[0196] 深度神经网络(deep neural network,DNN),也称多层神经网络,可以理解为具有很多层隐含层的神经网络,这里的“很多”并没有特别的度量标准。从DNN按不同层的位置划分,DNN内部的神经网络可以分为三类:输入层,隐含层,输出层。一般来说第一层是输入层,最后一层是输出层,中间的层数都是隐含层。层与层之间是全连接的,也就是说,第 $i$ 层的任意一个神经元一定与第 $i+1$ 层的任意一个神经元相连。

[0197] 虽然DNN看起来很复杂,但是就每一层的工作来说,其实并不复杂,简单来说就是如下线性关系表达式: $\vec{y} = \alpha(W \cdot \vec{x} + \vec{b})$ ,其中, $\vec{x}$ 是输入向量, $\vec{y}$ 是输出向量, $\vec{b}$ 是偏移向量, $W$ 是权重矩阵(也称系数), $\alpha(\cdot)$ 是激活函数。每一层仅仅是对输入向量 $\vec{x}$ 经过的操作得到输出向量 $\vec{y}$ 。由于DNN层数多,则系数 $W$ 和偏移向量 $\vec{b}$ 的数量也就很多了。

[0198] 这些参数在DNN中的定义如下所述:以系数 $W$ 为例:假设在一个三层的DNN中,第二层的第4个神经元到第三层的第2个神经元的线性系数定义为 $W_{24}^3$ 。上标3代表系数 $W$ 所在的层数,而下标对应的是输出的第三层索引2和输入的第二层索引4。

[0199] 综上,第 $L-1$ 层的第 $k$ 个神经元到第 $L$ 层的第 $j$ 个神经元的系数定义为 $W_{jk}^L$ 。

[0200] 需要注意的是,输入层是没有 $W$ 参数的。在深度神经网络中,更多的隐含层让网络更能够刻画现实世界中的复杂情形。理论上而言,参数越多的模型复杂度越高,“容量”也就越大,也就意味着它能完成更复杂的学习任务。训练深度神经网络也就是学习权重矩阵的过程,其最终目的是得到训练好的深度神经网络的所有层的权重矩阵(由很多层的向量 $W$ 形成的权重矩阵)。

### [0201] (3) 卷积神经网络

[0202] 卷积神经网络(convolutional neuron network,CNN)是一种带有卷积结构的深度神经网络。卷积神经网络包含了一个由卷积层和子采样层构成的特征抽取器。该特征抽取器可以看作是滤波器,卷积过程可以看作是使用一个可训练的滤波器与一个输入的图像或者卷积特征平面(feature map)做卷积。卷积层是指卷积神经网络中对输入信号进行卷积处理的神经元层。在卷积神经网络的卷积层中,一个神经元可以只与部分邻层神经元连

接。一个卷积层中,通常包含若干个特征平面,每个特征平面可以由一些矩形排列的神经单元组成。同一特征平面的神经单元共享权重,这里共享的权重就是卷积核。共享权重可以理解为提取图像信息的方式与位置无关。这其中隐含的原理是:图像的某一部分的统计信息与其他部分是一样的。即意味着在某一部分学习的图像信息也能用在另一部分上。所以对于图像上的所有位置,都能使用同样的学习得到的图像信息。在同一卷积层中,可以使用多个卷积核来提取不同的图像信息,一般地,卷积核数量越多,卷积操作反映的图像信息越丰富。

[0203] 卷积核可以以随机大小的矩阵的形式初始化,在卷积神经网络的训练过程中卷积核可以通过学习得到合理的权重。另外,共享权重带来的直接好处是减少卷积神经网络各层之间的连接,同时又降低了过拟合的风险。

[0204] (4) 损失函数

[0205] 在训练深度神经网络的过程中,因为希望深度神经网络的输出尽可能的接近真正想要预测的值,所以可以通过比较当前网络的预测值和真正想要的目标值,再根据两者之间的差异情况来更新每一层神经网络的权重向量(当然,在第一次更新之前通常会有初始化的过程,即为深度神经网络中的各层预先配置参数),比如,如果网络的预测值高了,就调整权重向量让它预测低一些,不断的调整,直到深度神经网络能够预测出真正想要的目标值或与真正想要的目标值非常接近的值。因此,就需要预先定义“如何比较预测值和目标值之间的差异”,这便是损失函数(loss function)或目标函数(objective function),它们是用于衡量预测值和目标值的差异的重要方程。其中,以损失函数举例,损失函数的输出值(loss)越高表示差异越大,那么深度神经网络的训练就变成了尽可能缩小这个loss的过程。

[0206] 为了更好地理解本申请实施例的视频编码方法和视频解码方法的整个过程,下面先结合图7对本申请实施例的视频编码和视频解码的整个过程进行大致的介绍。

[0207] 图7是本申请实施例的视频编码和视频解码过程的示意性流程图。

[0208] 如图7所示,待编码图像经过编码器编码后生成主编码流和辅助编码流,解码器在获取到主编码流和辅助编码流之后,对主编码流和辅助编码流进行解码,得到待显示的视频图像,接下来,解码器可以将待显示视频图像送入到显示设备中进行显示。图7中的编码器具体可以是上文中的编码器20,图7中的解码器具体可以是上文中的解码器30。

[0209] 本申请实施例的视频编码方法可以由图7中所示的编码器来执行,本申请实施例的视频解码方法可以由图7中所示的解码器来执行。

[0210] 图8是本申请实施例的视频解码方法的示意性流程图。图8所示的方法可以由视频解码器(例如,图1或图2中所示的解码器30)来执行,图8所示的方法包括步骤1001至1004,下面对这些步骤进行详细的介绍。

[0211] 1001、解码主编码流,获取第一视频图像。

[0212] 上述主编码流可以是编码端先对分辨率较高的原始图像进行降低分辨率处理,然后再对降低分辨率处理后得到的较低分辨率的视频图像进行编码后生成的码流。

[0213] 例如,初始视频图像(也可以称为原始视频图像)的分辨率为4K,编码端通过对初始视频图像进行下采样处理,得到2K的视频图像,接下来,对该2K的视频图像进行编码,得到主编码流。

[0214] 应理解,在步骤1001中,可以按照H.264/H.265/H.266等视频标准中规定的解码方式对主编码流进行解码,以得到第一视频图像。

[0215] 上述第一视频图像既可以是一帧图像,也可以是多帧(可以是连续的多帧)图像。

[0216] 1002、采用神经网络模型对第一视频图像进行超分辨率处理,得到第二视频图像。

[0217] 一般地,超分辨率处理是指通过对低分辨率的视频图像进行变换(或称为超分辨率重建),以得到高分辨率的视频图像的过程。

[0218] 例如,对一个图像分辨率为2K(2K resolution)的视频图像进行超分辨率处理,得到一个图像分辨率为4K(4K resolution)的视频图像。

[0219] 上述神经网络模型可以是根据训练样本图像和训练目标图像训练得到的,该训练目标图像是与目标视频图像的分辨率相同的视频图像,该训练样本图像是对训练目标图像进行下采样、编码和解码处理得到的视频图像。

[0220] 上述训练样本图像是神经网络模型的输入,上述训练目标图像是神经网络模型训练时的目标,神经网络模型在训练时输出的视频图像与训练目标图像的差异要尽可能的小。

[0221] 1003、解码辅助编码流,获取残差值。

[0222] 应理解,在步骤1003中,可以按照H.264/H.265/H.266等视频解码标准中规定的解码方式对辅助编码流进行解码,以得到残差值。

[0223] 1004、根据第二视频图像和残差值得到目标视频图像。

[0224] 可选地,上述根据第二视频图像和残差值得到目标视频图像,具体包括:将第二视频图像的像素值与上述残差值叠加,得到目标视频图像。

[0225] 应理解,上述残差值包括第二视频图像中的每个像素点的像素值对应残差值。上述对第二视频图像的像素值与残差值叠加是指将第二视频图像的每个像素点的像素值与其对应的残差值相加,进而得到目标视频图像中相应像素点的像素值。

[0226] 本申请中,通过神经网络模型能够更好地对分辨率较低的第一视频图像进行超分辨率处理,进而得到分辨率较高的第二视频图像,接下来,就能够通过第二视频图像以及解码得到的残差值进行一次叠加得到最终的目标视频图像,与传统方案中采用低分辨率的视频图像与多个层级的残差值进行层层叠加得到最终的视频图像的方式相比,能够简化解码流程,提高解码效率,提升解码器的性能。

[0227] 另外,本申请还能够灵活地根据需要来对神经网络模型进行优化调整,本申请方案的灵活性更高。

[0228] 具体地,本申请中可以根据超分辨率处理时分辨率提升程度的不同来灵活设置神经网络模型的模型参数。

[0229] 例如,当神经网络模型原来进行超分辨处理时是将视频图像的分辨率从1080p提升到2K,而现在需要神经网络模型在进行超分辨率处理时,是将分辨率从1080p提升到4K。那么,可以再采用一些训练数据(这些数据可以包括一些分辨率为4K的视频图像,以及对这些分辨率为4K的视频图像进行下采样、编码和解码处理得到的分辨率为1080p视频图像)对原来的神经网络模型重新进行训练,从而对原来的神经网络模型的模型参数进行调整,得到新的神经网络模型。

[0230] 可选地,第一视频图像为高动态范围HDR图像或者标准动态范围SDR图像。

[0231] 本申请的解码方法不仅适用于SDR图像,也适用于HDR图像。

[0232] 为了更好地理解图8所示的本申请实施例的视频解码方法的解码过程,下面结合图9以具体的实例对本申请实施例的视频解码方法的流程进行描述。

[0233] 图9是本申请实施例的视频解码方法的流程图。图9所示的视频解码方法包括以下步骤:

[0234] 2001、解码辅助编码流,得到残差值;

[0235] 2002、解码主编码流得到视频图像3 (2K);

[0236] 2003、对视频图像3 (2K) 进行超分辨率处理,得到视频图像4 (4K);

[0237] 2004、将视频图像4 (4K) 与残差值进行叠加,得到最终的目标视频图像。

[0238] 应理解,步骤2001至步骤2004的具体实现细节可以参见图8所示的方法中的相关描述。

[0239] 在图9所示的解码流程中,通过对主编码流进行解码,得到的是分辨率为2K的视频图像,接下来,再通过超分辨率处理,能够将图像的分辨率从2K提升到4K,最终根据视频图像4 (4K) 和残差值叠加就能得到最终的目标视频图像,其中,目标视频图像的分辨率也为4K,并且目标视频图像既可以是SDR视频图像,也可以是HDR视频图像。

[0240] 本申请中,在根据神经网络模型对第一视频图像进行超分辨率处理之前,可以采用大量的训练样本图像以及对应的训练目标图像进行训练。

[0241] 在训练上述神经网络模型时,可以计算神经网络模型的输出图像与目标训练图像的差异,并将神经网络模型的输出图像与训练目标图像的差异满足预设要求(例如,神经网络模型的输出图像与训练目标图像的差异值小于预设数值)时的神经网络模型的模型参数的参数值确定为神经网络模型的模型参数的最终取值,这样就完成了神经网络模型的训练。完成训练后的神经网络模型就可以用于对第一视频图像进行超分辨率处理。

[0242] 下面结合图10至图12对本申请中神经网络模型的训练的相关内容进行描述。

[0243] 图10是本申请实施例训练神经网络模型的示意性流程图。

[0244] 如图10所示,神经网络模型的训练过程包括步骤3001至3003,下面对这些步骤进行介绍。

[0245] 3001、对原始视频图像进行下采样和编码,得到主编码流。

[0246] 对原始视频图像进行下采样后,能够降低原始视频图像的分辨率,得到分辨率较低的视频图像,接下来,可以对该分辨率较低的视频图像进行编码,得到码流。

[0247] 例如,原始视频图像的分辨率为4K,通过下采样后得到的视频图像的分辨率为2K,那么,对该分辨率为2K的视频图像进行编码,可以得到主编码流。

[0248] 3002、对主编码流进行解码,得到低分辨率视频图像。

[0249] 例如,假设步骤3001中下采样后得到的视频图像的分辨率为2K,那么,步骤3002解码主编码流得到的视频图像的分辨率也为2K。

[0250] 3003、根据训练样本图像和训练目标图像对神经网络模型进行训练。

[0251] 其中,训练样本图像就是步骤3002中解码得到的低分辨率视频图像,该训练样本图像作为神经网络模型训练时输入的视频图像,而训练目标图像就是原始视频图像,该训练目标图像作为神经网络模型训练时的目标,通过训练,使得神经网络模型输出的视频图像与训练目标图像尽可能的接近,并将神经网络模型输出的视频图像与训练目标图像的差异



异满足预设要求时的对应的模型参数作为神经网络模型的最终的模型参数。

[0252] 图11是本申请实施例的训练神经网络模型的训练系统的示意性框图。

[0253] 如图11所示,训练系统100包括执行设备110、训练设备120、数据库130和图像采集设备140。下面对这些设备进行简单的介绍。

[0254] 图像采集设备140用于采集神经网络训练时用到的视频图像,图像采集设备140采集到的视频图像可以包括上述训练样本图像和训练目标图像,图像采集设备140可以将采集到的视频图像存储到数据库130中。

[0255] 训练设备120用于从数据库中获取训练用的视频图像(具体可以是训练样本图像和训练目标图像),并根据这些视频图像对目标模型/规则101进行训练,这里的目标模型/规则101可以是本申请实施例中的神经网络模型。

[0256] 执行设备110是包含计算模块111、I/O接口112和预处理模块113的设备,其中,计算模块111包括目标模型/规则101(相当于神经网络模型在计算模块中实现)。在训练设备完成了对目标模型/规则101的训练之后,执行设备110就可以对输入视频图像进行处理,得到输出视频图像了。

[0257] 其中,执行设备110的计算模块111在对输入视频图像进行处理之前,还可以先采用输入视频图像进行一些预处理,然后再由计算模块111对输入视频图像进行处理(也可以不对输入视频图像进行预处理,而由计算模块111直接对输入视频图像进行处理)。

[0258] 上述执行设备110既可以是视频编码器,也可以是视频解码器,当执行设备110是视频解码器时,执行设备110可以对解码得到的分辨率较低的视频图像进行超分辨率处理,以得到分辨率较高的视频图像。当执行设备110是视频编码器时,执行设备110可以对原始视频图像下采样后得到的视频图像进行超分辨率处理,以得到分辨率较高的视频图像。

[0259] 上述步骤1003解码得到的残差值可以只包含像素点亮度分量值对应的残差,也可以既包含像素点亮度分量值对应的残差又包含像素点色度分量值对应的残差。这两种情况下超分辨率处理以及最终得到目标视频图像的具体实现方式有所不同,下面分别进行详细介绍。

[0260] 情况一:步骤1003解码得到的残差值既包含像素点亮度分量值对应的残差又包含像素点色度分量值对应的残差。

[0261] 在情况一中,上述步骤1002中采用神经网络模型对第一视频图像进行超分辨率处理,得到第二视频图像,具体包括:采用神经网络模型对第一视频图像的像素点亮度分量值和像素点色度分量值进行超分辨率处理,得到第二视频图像的像素点亮度分量值和像素点色度分量值;上述步骤1004中将第二视频图像的像素值与残差值叠加,得到目标视频图像,具体包括:将第二视频图像的像素点亮度分量值和像素点色度分量值与残差值叠加,得到目标视频图像。

[0262] 应理解,在情况一中,残差值中既包含与像素点亮度分量值对应的残差又包含与像素点色度分量值对应的残差。

[0263] 上述将第二视频图像的像素点亮度分量值和像素点色度分量值与残差值叠加,得到目标视频图像,具体是将第二视频图像的像素点亮度分量值与像素点亮度分量值对应的残差叠加,得到目标视频图像的像素点亮度分量值,以及将第二视频图像的像素点色度分量值与像素点色度分量值对应的残差叠加,得到目标视频图像的像素点色度分量值,从而

得到目标视频图像的像素点亮度分量值和像素点色度分量值(得到了目标视频图像的像素点亮度分量值和像素点色度分量值,也就得到了目标视频图像的像素值了)。

[0264] 情况二:步骤1003解码得到的残差值仅包含像素点亮度分量值对应的残差。

[0265] 在情况二中,上述步骤1002中采用神经网络模型对第一视频图像进行超分辨率处理,得到第二视频图像,具体包括:采用神经网络模型对第一视频图像的像素点亮度分量值进行超分辨率处理,得到第二视频图像的像素点亮度分量值;上述步骤1004中将第二视频图像的像素值与残差值叠加,得到目标视频图像,包括:将第二视频图像的像素点亮度分量值与残差值叠加,得到目标视频图像的像素点亮度分量值。

[0266] 本申请中,神经网络模型仅对第一视频图像的像素点亮度分量值进行处理,能够减少神经网络模型进行超分辨处理时的计算量,提高视频解码的效率。

[0267] 具体地,相对于色度,人眼对图像的亮度更加敏感。因此,本申请中可以采用神经网络模型仅对第一视频图像的像素点亮度分量值进行超分辨率处理,而第二视频图像的像素点的色度分量值可以通过传统的插值算法计算得到的,使得本申请能够在保证视觉体验的同时,减少视频解码的计算复杂度。

[0268] 可选地,图8所示的方法还包括:对第一视频图像的像素点色度分量值进行插值处理,得到目标视频图像的像素点色度分量值。

[0269] 应理解,在本申请中,既可以采用神经网络模型来计算得到第二视频图像的像素点亮度分量值和像素点色度分量值,也可以只采用神经网络模型来计算得到第二视频图像的像素点亮度分量值,而采用其他的方法(如插值方法)得到第二视频图像的像素点色度分量值。

[0270] 上述像素点亮度分量值可以称为像素点的亮度分量值,上述像素点色度分量值可以称为像素点的色度分量值。为了描述方便,本申请统一称为像素点亮度分量值和像素点色度分量值。

[0271] 本申请中的视频图像既可以采用RGB格式表示,也可以是采用YUV格式来表示,其中,Y表示明亮度(Luminance或Luma),也就是灰度值,U表示色度(Chrominance),V表示浓度(Chroma)。

[0272] 上述像素点亮度分量值可以表示为Y,上述像素点色度分量值可以包括U和V。

[0273] 图12是本申请实施例的采用卷积神经网络对第一视频图像进行处理的示意图。

[0274] 如图12所示,可以采用卷积神经网络对第一视频图像的像素点亮度分量值进行超分辨率处理,得到第二视频图像的像素点的亮度分量值。另外,为了提高最终得到的第二视频图像的像素点的亮度分量值的准确性,还可以综合采用卷积神经网络和插值算法(例如,可以采用双三次插值(Bicubic插值))来获取第二视频图像的像素点的亮度分量值。

[0275] 如图12所示,可以采用卷积神经网络对第一视频图像的像素点亮度分量值进行处理,得到第二视频图像的像素点亮度分量值,采用插值处理算法对第一视频图像的像素点亮度分量值进行处理,得到第二视频图像的像素点的亮度分量值,接下来,再对两种方式得到的像素点亮度分量值进行求和就能够得到第二视频图像的像素点亮度分量值了。

[0276] 应理解,在图12中,第一视频图像可以是对原始视频图像进行下采样之后得到的较低分辨率的视频图像,第二视频图像可以是与第一视频图像的分辨率相同的视频图像。

[0277] 在对第一视频图像的像素点亮度分量值进行处理时,为了便于神经网络进行处

理,可以先将第一视频图像的像素点亮度分量值的取值范围由 $[0, 255]$ 线性化为 $[0, 1]$ ,然后再减去0.5,使得第一视频图像的像素点亮度分量值的取值范围变为 $[-0.5, 0.5]$ ,然后再采用卷积神经网络对第一视频图像的像素点亮度分量值进行超分辨率处理。

[0278] 在图12中,卷积神经网络只需要对第一视频图像的像素点亮度分量值进行超分辨率处理即可,能够降低卷积神经网络处理的复杂度。

[0279] 应理解,图12所示的流程计算出了第二视频图像的像素点的亮度分量值,而第二视频图像的像素点色度分量值可以通过对第一视频图像的像素点色度分量值进行插值运算得到,这样就得到了第二视频图像的像素点色度分量值和像素点亮度分量值,也就得到了第二视频图像。

[0280] 另外,在图12中,当对卷积神经网络进行训练时,还可以再计算第二视频图像相对于原始视频图像的损失,以修正卷积神经网络的参数,以减小第二视频图像相对于原始视频图像的损失。此外,这里可以通过计算损失函数的方式确定第二视频图像相对于原始视频图像的损失。这里的损失函数可以是均方误差(mean squared error, MSE)等等。

[0281] 图13是本申请实施例的卷积神经网络的结构示意图。

[0282] 应理解,图12中的卷积神经网络的具体结构可以如图13所示。在图13中,卷积神经网络中包括多个卷积层、多个激活函数层以及像素重组层,卷积神经网络中间层的部分结果会跳跃地向后传输。

[0283] 卷积层是卷积神经网络的基本构造单元,能够对图像的特征进行提取,输出特征矩阵。在图13中,最后一个卷积层的通道数为4,其余的卷积层的通道数为64,每个通道对应一个卷积核,用于输出一个特征矩阵。对于最后一个卷积层来说,输出的是4个特征矩阵,接下来,像素重组层可以将这4个特征矩阵叠加成一个高分辨率的特征矩阵,像素点的个数变为原来的4倍。

[0284] 另外,图13中的激活函数层的系数可以是0.1。

[0285] 应理解,图13中卷积神经网络的结构以及卷积神经网络中的卷积层的通道数以及激活函数层的系数仅为示例,本申请中的卷积神经网络的还可以是其它结构,本申请中的卷积神经网络中的卷积层的通道数以及激活函数层的系数还可以是其它数值。

[0286] 上面结合图8从解码端的角度对本申请实施例的视频解码方法进行了详细的介绍,下面结合图14从编码端的角度对本申请实施例的视频编码方法进行介绍,应理解,图14所示的编码方法与图8所示的解码方法是对应的,图8所示的解码方法能够对图14所示的编码方法产生的主编码流和辅助编码流进行解码进而得到最终的目标视频图像,为了简洁,下面在介绍图14所示的编码方法时适当省略重复的描述。

[0287] 图14是本申请实施例的视频编码方法的示意性流程图。图14所示的方法可以由视频编码器(例如,上文中图1中所示的编码器20或者图2中所示的编码器20)来执行,应理解,图14所示的方法可以按照H.264/H.265/H.266等视频标准中规定方式进行编码和/或解码。图14所示的方法包括步骤4001至4005,下面对这些步骤进行介绍。

[0288] 4001、对初始视频图像进行下采样和编码处理,得到主编码流。

[0289] 在步骤4001中,通过对初始视频图像进行下采样操作,可以得到分辨率较低的视频图像(该分辨率较低的视频图像的分辨率低于初始视频图像的分辨率),接下来,再对该分辨率较低的视频图像进行编码处理得到主编码流。

[0290] 例如,初始视频图像的分辨率为4K,编码端通过对初始视频图像进行下采样处理,得到2K的视频图像,接下来,对该2K的视频图像进行编码,得到主编码流。上述初始视频图像可以来自于成像设备,例如摄像机。

[0291] 另外,上述步骤4001中的编码处理可以具体包括预测、变换、量化和熵编码等处理过程,通过这些处理过程能够最终得到主编码流。

[0292] 可选地,上述初始视频图像为HDR图像或者SDR图像。

[0293] 图14所示的本申请实施例的视频编码方法即可以适用于SDR图像,也可以适用于HDR图像。

[0294] 4002、对主编码流进行解码,得到第一视频图像。

[0295] 步骤4002实现的具体过程与步骤1001相同,步骤4002实现的具体过程可以参见步骤1001的相关描述。

[0296] 4003、采用神经网络模型对第一视频图像进行超分辨率处理,得到与初始视频图像的分辨率相同的第二视频图像。

[0297] 其中,上述神经网络模型是根据训练样本图像和训练目标图像进行训练得到的,训练目标图像是与初始视频图像分辨率相同的视频图像,训练样本图像是对训练目标图像进行下采样、编码和解码处理后得到的视频图像。

[0298] 步骤4003实现的具体过程与步骤1002相同,步骤4003实现的具体过程可以参见步骤1002的相关描述。

[0299] 4004、确定初始视频图像相对于第二视频图像的残差值。

[0300] 本申请中,先通过步骤4002对主编码流进行解码,然后再通过步骤4003对解码得到的第一视频图像进行编码,能够使得步骤4004最终确定的残差值能够反映出编码损失,从而使得解码端能够根据解码得到的残差值来获得初始视频图像。

[0301] 4005、对残差值进行编码,得到辅助编码流。

[0302] 步骤4005中,可以按照RVC技术中对残差值进行编码、进而生成辅助编码流的方式进行编码。

[0303] 应理解,图14所示的方法(步骤4003)中的神经网络模型的模型参数可以与图8所示的方法(步骤1002)中的神经网络模型的模型参数一致,图8所示的方法中对神经网络模型的限定和解释同样适用于图14所示的方法中的神经网络模型。

[0304] 本申请中,在生成主编码流之后,能够通过神经网络模型对主编码流解码得到的视频图像进行超分辨率处理,从而得到与初始视频图像分辨率相同的视频图像,接下来,就可以直接生成初始视频图像相对于超分辨率处理得到的视频图像的残差值,并根据残差值生成辅助码流,与传统编码方案中需要根据初始视频图像得到多层分辨率等级不同的视频图像以及对应等级的残差值的方式相比,本实施例能够简化编码的流程。

[0305] 可选地,作为一个实施例,上述步骤4003中采用神经网络模型对第一视频图像进行超分辨率处理,得到与初始视频图像的分辨率相同的第二视频图像,包括:采用神经网络模型对第一视频图像的像素点亮度分量值进行超分辨率处理,得到第二视频图像的像素点亮度分量值;上述4004中确定初始视频图像相对于第二视频图像的残差值,包括:将初始视频图像的像素点亮度分量值相对于第二视频图像的像素点亮度分量值的差值,确定为残差值。

[0306] 应理解,当图14所示的方法编码得到的残差值为初始视频图像的像素点亮度分量值相对于第二视频图像的像素点亮度分量值的差值时,解码端在解码辅助编码流得到残差之后,也是将该残差值与获得的第二视频图像的像素点亮度分量值进行叠加,进而得到最终的目标视频图像的像素点亮度分量值,而目标视频图像的像素点色度分量值可以通过其他插值的方法得到。

[0307] 可选地,作为一个实施例,上述步骤4003中采用神经网络模型对第一视频图像进行超分辨率处理,得到与初始视频图像的分辨率相同的第二视频图像,包括:采用神经网络模型对第一视频图像的像素点亮度分量值和像素点色度分量值进行超分辨率处理,得到第二视频图像的像素点亮度分量值和像素点色度分量值;上述4004中确定初始视频图像相对于第二视频图像的残差值,包括:将初始视频图像的像素点亮度分量值和像素点色度分量值分别相对于第二视频图像的像素点亮度分量值和像素点色度分量值的差值,确定为残差值。

[0308] 应理解,当图14所示的方法编码得到的残差值为初始视频图像的像素点亮度分量值和像素点色度分量值分别相对于第二视频图像的像素点亮度分量值和像素点色度分量值的差值时,解码端在解码辅助编码流得到残差之后,也是将该残差值与获得的第二视频图像的像素点亮度分量值和像素点色度分量值进行叠加,进而得到最终的目标视频图像的像素点亮度分量值和像素点色度分量值,从而直接得到最终的目标视频图像。(这里的目标视频图像相当于图14所示的方法中的初始视频图像,但是相对于初始视频图像可能有少许失真)。

[0309] 本申请中,神经网络模型仅对第一视频图像的像素点亮度分量值进行处理,能够减少神经网络模型进行超分辨处理时的计算量,提高视频解码的效率。

[0310] 相对于色度,人眼对图像的亮度更加敏感。因此,本申请中可以采用神经网络模型仅对第一视频图像的像素点亮度分量值进行超分辨率处理,而第二视频图像的像素点的色度分量值可以通过传统的插值算法计算得到的,使得本申请能够在保证视觉体验的同时,减少视频解码的计算复杂度。

[0311] 为了更好地理解图14所示的本申请实施例的视频编码方法的编码过程,下面结合图15以具体的实例对本申请实施例的视频编码方法的流程进行描述。

[0312] 图15是本申请实施例的视频编码方法的流程图。图15所示的视频编码方法包括以下步骤:

[0313] 5001、对视频图像1进行下采样操作,得到视频图像2。

[0314] 其中,视频图像1为分辨率为4K的视频图像,视频图像2为分辨率为2K的视频图像,通过步骤5001中的下采样操作,使得输入的视频图像的分辨率降低了一倍(由4K降低到了2K)。

[0315] 5002、对视频图像2进行编码,得到主编码流。

[0316] 5003、对主编码流进行解码,得到视频图像3。

[0317] 其中,视频图像3是分辨率为2K的视频图像,视频图像3的分辨率与视频图像2的分辨率相同,视频图像3与视频图像2的图像内容基本相同(经过编码和解码,可能导致一部分失真,使得视频图像3的图像内容与视频图像2的图像内容有所差异)。

[0318] 通过步骤5002和步骤5003对视频图像2的编码和解码,又得到了分辨率为2K的视

频图像3。

[0319] 5004、对视频图像3进行超分辨率处理,得到视频图像4。

[0320] 其中,视频图像4的分辨率为4K,通过对视频图像3进行超分辨率处理,能够将视频图像的分辨率提升一倍。在步骤5004中,可以采用神经网络模型对视频图像3进行超分辨率处理,得到视频图像4,具体过程与上述步骤4003类似,这里不再详细描述。

[0321] 5005、根据视频图像1和视频图像4确定残差值。

[0322] 视频图像1和视频图像4均是分辨率为4K的视频图像,通过视频图像1的像素值与视频图像4的像素值进行做差,能够得到残差值(视频图像1的像素值与视频图像4的像素值的差值)。

[0323] 此外,这里的残差值可以只是视频图像1的像素点亮度分量值与视频图像4的像素点的亮度分量值的差,或者,这里的残差值也可以是视频图像1的像素点亮度分量值和像素点色度分量值分别与视频图像4的像素点亮度分量值和像素点色度分量值的差。

[0324] 5006、对残差值进行编码,得到辅助编码流。

[0325] 经过上述步骤5001至步骤5006完成了对视频图像1的编码过程,得到了主编码流和辅助编码流。

[0326] 应理解,图15所示的编码过程可以是与图9所示的解码过程相对应的,图9所示的解码过程能够对图15所示的编码过程编码得到的主编码流和辅助编码流进行解码。

[0327] 上文结合图8和图14分别从解码端和编码端的角度对本申请实施例的视频解码方法和视频编码方法进行了详细的介绍,为了更好地理解本申请实施例的视频编码和视频解码的整个过程,下面结合图16对本申请实施例的视频编码过程和视频解码的过程进行介绍。

[0328] 图16是本申请实施例的视频解码方法和视频编码过程的示意性流程图。图16所示的方法可以由视频解码器(例如,图1或图2中所示的解码器30)和视频编码器(例如,上文中图1中所示的编码器20或者图2中所示的编码器20)来共同执行,

[0329] 图16所示的视频编码和视频解码过程包括步骤6001至步骤6010,下面对这些步骤进行介绍。

[0330] 6001、对初始视频图像进行下采样,获得下采样后得到的视频图像。

[0331] 6002、对下采样后得到的视频图像进行编码,得到主编码流。

[0332] 6003、对主编码流进行解码,得到第一视频图像。

[0333] 6004、对第一视频图像进行超分辨率处理,得到第二视频图像。

[0334] 具体地,可以采用神经网络模型对第一视频图像进行超分辨率处理,得到第二视频图像,具体内容可以参见上文中对步骤4003的相关描述。

[0335] 6005、根据初始视频图像和第二视频图像确定残差值。

[0336] 6006、对残差值进行编码,得到辅助编码流。

[0337] 上述步骤6001至步骤6006可以由视频编码器来执行,上述步骤6001至步骤6006相当于图14所示的方法中的步骤4001至步骤4005(都是对输入的初始视频图像进行处理,得到主编码流和辅助编码流)。其中,步骤6001和步骤6002相当于步骤4001,用于对初始视频图像进行下采样处理和编码处理,以得到主编码流。步骤6003相当于步骤4002,步骤6004相当于步骤4003,步骤6005相当于步骤4004,步骤6006相当于步骤4005,步骤6001至步骤6006

的描述可以参见步骤4001至步骤4005的相关内容。具体的执行过程参考上面实施例的描述。

[0338] 主编码流和辅助编码流输出到解码端,解码端执行如下的解码过程。

[0339] 6007、对辅助编码流进行解码,得到残差值。

[0340] 6008、对主编码流进行解码,得到第一视频图像。

[0341] 6009、对第一视频图像进行超分辨率处理,得到第二视频图像。

[0342] 6010、将第二视频图像与残差值叠加,得到目标视频图像。

[0343] 上述步骤6007至步骤6010可以由视频编码器来执行,上述步骤6007至步骤6010相当于图8所示的方法中的步骤1001至步骤1004(通过对主编码流和辅助编码流进行解码,以得到最终的目标视频图像)。其中,步骤6008相当于图8所示的方法中的步骤1003,步骤6008相当于图8所示的方法中的步骤1001,步骤6009中具体可以采用神经网络模型对第一视频图像进行超分辨率处理,步骤6009相当于图8所示的方法中的步骤1002,步骤6010相当于图8所示的方法中的步骤1004。具体的执行过程参考上面实施例的描述。

[0344] 对于某些视频编码设备或者视频解码设备来说,可能只支持SDR视频图像的编码或者解码,而无法直接对HDR视频图像进行编码或者解码,在这种情况下,可以先将HDR视频图像转化为SDR视频图像然后再进行处理,下面结合图17和图18对这种情况下的视频编码和视频解码进行详细的描述。

[0345] 图17是本申请实施例的视频解码方法的示意性流程图。图17所示的方法可以由视频解码器(例如,图1或图2中所示的解码器30)来执行,图17所示的方法包括步骤7001至7006,下面对这些步骤进行详细的介绍。

[0346] 7001、解码主编码流,获取第一视频图像,所述第一视频图像为SDR视频图像;

[0347] 7002、采用神经网络模型对所述第一视频图像进行处理,得到第二视频图像;

[0348] 其中,第二视频图像为高动态范围HDR视频图像,且第二视频图像的分辨率大于第一视频图像的分辨率;

[0349] 7003、解码辅助编码流,获取残差值;

[0350] 7004、将所述第二视频图像的像素值与所述残差值叠加,得到目标视频图像。

[0351] 其中,所述神经网络模型是根据训练样本图像和训练目标图像训练得到的,所述训练目标图像是与所述目标视频图像的分辨率相同的视频图像,所述训练样本图像是对所述训练目标图像进行下采样、编码和解码处理得到的视频图像。

[0352] 图17所示的方法中的神经网络模型与图8所示的方法的神经网络模型类似,图8所示的方法中对神经网络相关的限定和解释也适用于图17所示的方法中的神经网络模型。

[0353] 本申请中,当解码得到的是SDR视频图像时,通过神经网络模型进行处理,能够得到分辨率更高的HDR视频图像,接下来,就能够通过第二视频图像以及解码得到的残差值进行一次叠加得到最终的目标视频图像,本申请能够适用于支持SDR视频图像编码和解码的设备。

[0354] 另外,本申请与传统方案中采用低分辨率的视频图像与多个层级的残差值进行层层叠加得到最终的视频图像的方式相比,能够简化解码流程。本申请还能够灵活地根据需要来对神经网络模型进行优化调整,本申请方案的灵活性更高。

[0355] 可选地,作为一个实施例,上述步骤7002中采用神经网络模型对第一视频图像进

行处理,得到第二视频图像,包括:采用神经网络模型对第一视频图像进行超分辨率处理和反向色调映射处理,得到第二视频图像。

[0356] 其中,超分辨率处理用于提升第一视频图像的分辨率,用于得到分辨率较高的视频图像,反向色调映射处理用于提升分辨率较高的视频图像的像素精度,以得到第二视频图像。

[0357] 可选地,作为一个实施例,当上述残差值仅包含像素点亮度分量值对应的残差时,上述步骤7002中采用神经网络模型对第一视频图像进行处理,得到第二视频图像,包括:采用神经网络模型对第一视频图像的像素点亮度分量值进行处理,得到第二视频图像的像素点亮度分量值;将第二视频图像的像素值与残差值叠加,得到目标视频图像,包括:将第二视频图像的像素点亮度分量值与残差值叠加,得到目标视频图像的像素点亮度分量值。

[0358] 可选地,作为一个实施例,图17所示的方法还包括:对第一视频图像的像素点色度分量值进行插值处理,得到目标视频图像的像素点色度分量值。

[0359] 本申请中,神经网络模型仅对第一视频图像的像素点亮度分量值进行处理,能够减少神经网络模型进行超分辨处理时的计算量,提高视频解码的效率。

[0360] 相对于色度,人眼对图像的亮度更加敏感。因此,本申请中可以采用神经网络模型仅对第一视频图像的像素点亮度分量值进行超分辨率处理,而第二视频图像的像素点的色度分量值可以通过传统的插值算法计算得到的,使得本申请能够在保证视觉体验的同时,减少视频解码的计算复杂度。

[0361] 可选地,当上述残差值包含像素点亮度分量值和像素点色度分量值对应的残差时,上述采用神经网络模型对第一视频图像进行超分辨率处理,得到第二视频图像,包括:采用神经网络模型对第一视频图像的像素点亮度分量值和像素点色度分量值进行超分辨率处理,得到第二视频图像的像素点亮度分量值和像素点色度分量值。

[0362] 在本申请中,既可以采用神经网络模型来计算得到第二视频图像的像素点亮度分量值和像素点色度分量值,也可以只采用神经模型来计算得到第二视频图像的像素点亮度分量值,而采用其他的方法(如插值方法)得到第二视频图像的像素点色度分量值。

[0363] 图18是本申请实施例的视频编码方法的示意性流程图。图18所示的方法可以由视频编码器(例如,上文中图1中所示的编码器20或者图2中所示的编码器20)来执行,应理解,图18所示的方法可以按照H.264/H.265/H.266等视频标准中规定方式进行编码和/或解码。图18所示的方法包括步骤8001至8006,下面对这些步骤进行介绍。

[0364] 8001、对初始视频图像进行处理,得到处理后的视频图像。

[0365] 其中,初始视频图像为高动态范围HDR视频图像,处理后的视频图像为标准动态范围SDR视频图像。

[0366] 8002、对处理后的视频图像进行编码,得到主编码流。

[0367] 8003、对主编码流进行解码,得到第一视频图像。

[0368] 8004、采用神经网络模型对第一视频图像进行处理,得到第二视频图像,第二视频图像为HDR视频图像,且第二视频图像的分辨率与初始视频图像的分辨率相同。

[0369] 其中,上述神经网络模型是根据训练样本图像和训练目标图像进行训练得到的,训练目标图像是与初始视频图像分辨率相同的视频图像,训练样本图像是对训练目标图像进行下采样、编码和解码处理后得到的视频图像。



[0370] 8005、确定初始视频图像相对于第二视频图像的残差值。

[0371] 8006、对残差值进行编码,得到辅助编码流。

[0372] 应理解,图18所示的编码过程可以是与图17所示的解码过程相对应的,图18所示的编码过程能够对图17所示的编码过程编码得到的主编码流和辅助编码流进行解码。

[0373] 图18所示的方法中的神经网络模型与图14所示的方法的神经网络模型类似,图18所示的方法中对神经网络相关的限定和解释也适用于图14所示的方法中的神经网络模型。

[0374] 可选地,作为一个实施例,上述步骤8005中确定初始视频图像相对于第二视频图像的残差值,包括:确定初始视频图像的像素值相对于第二视频图像的像素值的残差值。

[0375] 具体地,上述残差值可以是初始视频图像中的像素点的像素值与第二视频图像中对应的像素点的像素值的差值。

[0376] 本申请中,对于HDR图像,可以先将该HDR图像转化为SDR图像,然后再进行后续的编码处理,能够适用于仅支持对SDR图像进行编码的设备。

[0377] 进一步的,在生成主编码流之后,能够通过神经网络模型对主编码流解码得到的视频图像进行超分辨率处理,从而得到与初始视频图像分辨率相同的视频图像,接下来,就可以直接生成初始视频图像相对于超分辨率处理得到的视频图像的残差值,并根据残差值生成辅助码流,与传统编码方案中需要根据初始视频图像得到多层分辨率等级不同的视频图像以及对应等级的残差值的方式相比,能够简化编码的流程。

[0378] 可选地,作为一个实施例,上述步骤8004中采用神经网络模型对初始视频图像进行处理,得到处理后的视频图像,包括:采用神经网络模型对初始视频图像进行下采样和色调映射处理,得到处理后的视频图像。

[0379] 其中,下采样用于降低初始视频图像的分辨率,得到分辨率较低的视频图像,色调映射处理用于对分辨率较低的视频图像进行处理,以提升该分辨率较低的视频图像的像素精度,得到处理后的视频图像。

[0380] 可选地,作为一个实施例,上述步骤8004中采用神经网络模型对第一视频图像进行处理,得到第二视频图像,包括:采用神经网络模型对第一视频图像进行超分辨率处理和反向色调映射处理,得到第二视频图像。

[0381] 其中,超分辨率处理用于提升第一视频图像的分辨率,用于得到分辨率较高的视频图像,反向色调映射处理用于提升分辨率较高的视频图像的像素精度,以得到第二视频图像。

[0382] 可选地,作为一个实施例,上述步骤8004中采用神经网络模型对第一视频图像进行处理,得到第二视频图像,包括:采用神经网络模型对第一视频图像的像素点亮度分量值进行处理,得到第二视频图像的像素点亮度分量值;上述步骤8004中确定初始视频图像相对于第二视频图像的残差值,包括:将初始视频图像的像素点亮度分量值相对于第二视频图像的像素点亮度分量值的差值,确定为上述残差值。

[0383] 可选地,上述步骤8004中采用神经网络模型对第一视频图像进行处理(可以是超分辨率处理和反向色调映射处理),得到第二视频图像,包括:采用神经网络模型对第一视频图像的像素点亮度分量值和像素点色度分量值进行超分辨率处理,得到第二视频图像的像素点亮度分量值和像素点色度分量值;上述步骤8005中确定初始视频图像相对于第二视频图像的残差值,包括:将初始视频图像的像素点亮度分量值和像素点色度分量值分别相

对于第二视频图像的像素点亮度分量值和像素点色度分量值的差值,确定为上述残差值。

[0384] 在本申请中,既可以采用神经网络模型来计算得到第二视频图像的像素点亮度分量值和像素点色度分量值,也可以只采用神经模型来计算得到第二视频图像的像素点亮度分量值,而采用其他的方法(如插值方法)得到第二视频图像的像素点色度分量值。

[0385] 为了更好地理解图17所示的本申请实施例的视频解码方法的解码过程以及图18所示的本申请实施例的视频编码方法编码过程,下面结合图19和图20以具体的实例对本申请实施例的视频解码方法的流程和视频编码方法的流程进行描述。

[0386] 图19是本申请实施例的视频解码方法的流程图。图19所示的视频解码方法包括以下步骤:

[0387] 9001、解码辅助编码流,得到残差值;

[0388] 9002、解码主编码流得到视频图像4(2K、SDR);

[0389] 9003、对视频图像4(2K、SDR)进行超分辨率处理和反向色调映射,得到视频图像5(4K、HDR);

[0390] 9004、将视频图像5(4K、HDR)与残差值进行叠加,得到最终的目标视频图像。

[0391] 其中,视频图像4为SDR图像,视频图像4的分辨率为2K,在对视频图像4进行超分辨率处理之后,能够将分辨率从2K提升到4K,并且通过反向色调映射,能够将SDR图像变换为HDR图像,使得得到得视频图像5是分辨率为4K的HDR图像。

[0392] 图20是本申请实施例的视频编码方法的流程图。图20所示的视频编码方法包括以下步骤:

[0393] 10001、对视频图像1进行下采样操作,得到视频图像2。

[0394] 其中,视频图像1是分辨率为4K的HDR图像,经过下采样得到的视频图像2是分辨率为2K的HDR图像,通过步骤5001中的下采样操作,使得输入的视频图像的分辨率降低了一倍(由4K降低到了2K)。

[0395] 10002、对视频图像2进行色调映射,得到视频图像3。

[0396] 本申请中,通过色调映射处理,能够将视频图像由HDR图像转化为SDR图像,从而使原本不能直接对HDR视频图像进行编码/解码的设备也能够处理HDR图像,能够提高编解码设备的兼容性,使得编解码设备既能够处理HDR图像,也能够处理SDR图像。

[0397] 例如,视频图像2为10bit的像素精度的HDR图像,通过色调映射,能够得到像素精度为8bit的SDR图像,接下来,编码设备就可以对SDR图像进行编码了。

[0398] 10003、对视频图像3进行编码,得到主编码流。

[0399] 10004、对主编码流进行解码,得到视频图像4。

[0400] 其中,视频图像4是分辨率为2K的HDR视频图像,视频图像4的分辨率与视频图像3的分辨率相同,视频图像4与视频图像3的图像内容基本相同(经过编码和解码,尽可能导致一部分失真,使得视频图像4的图像内容与视频图像3的图像内容有所差异)。

[0401] 通过步骤10003和步骤10004对视频图像3的编码和解码,又得到了分辨率为2K的视频图像4。

[0402] 10005、对视频图像3进行超分辨率处理和反向色调映射处理,得到视频图像5。

[0403] 其中,视频图像5是分辨率为4K的HDR图像,通过对视频图像4进行超分辨率处理,能够将视频图像的分辨率提升一倍。在步骤10005中,可以采用神经网络模型对视频图像4

进行超分辨率处理和反向色调映射,得到视频图像5,具体过程与图18所示的方法中的步骤8004类似,这里不再详细描述。

[0404] 10006、根据视频图像1和视频图像5确定残差值。

[0405] 视频图像1和视频图像4均是分辨率为4K的HDR图像,通过视频图像1的像素值与视频图像5的像素值进行做差,能够得到残差值(视频图像1的像素值与视频图像5的像素值的差值)。

[0406] 当然,这里的残差值可以只是视频图像1的像素点亮度分量值与视频图像5的像素点的亮度分量值的差,或者,这里的残差值也可以是视频图像1的像素点亮度分量值和像素点色度分量值分别与视频图像5的像素点亮度分量值和像素点色度分量值的差。

[0407] 10007、对残差值进行编码,得到辅助编码流。

[0408] 经过上述步骤10001至步骤10007完成了对视频图像1的编码过程,得到了主编码流和辅助编码流。

[0409] 应理解,图20所示的编码过程可以是与图19所示的解码过程相对应的,图19所示的解码过程能够对图20所示的编码过程编码得到的主编码流和辅助编码流进行解码。

[0410] 上文结合图17和图18分别从解码端和编码端的角度对本申请实施例的视频解码方法和视频编码方法进行了详细的介绍,为了更好地理解本申请实施例的视频编码和视频解码的整个过程,下面结合图21对本申请实施例的视频编码过程和视频解码的过程进行介绍。

[0411] 图21是本申请实施例的视频解码方法和视频编码过程的示意性流程图。图21所示的方法可以由视频解码器(例如,图1或图2中所示的解码器30)和视频编码器(例如,上文中图1中所示的编码器20或者图2中所示的编码器20)来共同执行,

[0412] 图21所示的视频编码和视频解码过程包括步骤20001至步骤20011,下面对这些步骤进行介绍。

[0413] 20001、对视频图像1进行下采样处理,获得下采样后得到的视频图像。

[0414] 20002、对上述下采样后得到的视频图像进行色调映射处理,获得色调映射后得到的视频图像。

[0415] 20003、对上述色调映射后得到的视频图像进行编码,得到主编码流。

[0416] 20004、对主编码流进行解码,得到第一视频图像。

[0417] 20005、采用神经网络模型对第一视频图像进行超分辨率处理和反向色调映射,得到第二视频图像。

[0418] 上述第二视频图像为HDR视频图像,且第二视频图像的分辨率与初始视频图像的分辨率相同。

[0419] 其中,上述神经网络模型是根据训练样本图像和训练目标图像进行训练得到的,该训练目标图像是与初始视频图像分辨率相同的视频图像,该训练样本图像是对训练目标图像进行下采样、编码和解码处理后得到的视频图像。

[0420] 20006、确定初始视频图像相对于第二视频图像的残差值。

[0421] 20007、对残差值进行编码,得到辅助编码流。

[0422] 上述步骤20001至步骤20007可以由视频编码器来执行,上述步骤20001至步骤20007相当于图18所示的方法中的步骤8001至步骤8005(都是对输入的初始视频图像进行

处理,得到主编码流和辅助编码流)。

[0423] 20008、解码辅助编码流,获取残差值。

[0424] 20009、解码主编码流,获取第一视频图像,该第一视频图像为SDR视频图像;

[0425] 20010、采用神经网络模型对第一视频图像进行超分辨率处理和反向色调映射处理,得到第二视频图像。

[0426] 上述第二视频图像为高动态范围HDR视频图像,且第二视频图像的分辨率大于第一视频图像的分辨率;

[0427] 其中,上述神经网络模型是根据训练样本图像和训练目标图像训练得到的,该训练目标图像是与目标视频图像的分辨率相同的视频图像,该训练样本图像是对训练目标图像进行下采样、编码和解码处理得到的视频图像。

[0428] 20011、将第二视频图像的像素值与残差值叠加,得到目标视频图像。

[0429] 上述步骤20008至步骤20011可以由视频解码器来执行,上述步骤20008至步骤20011相当于图17所示的方法中的步骤7001至步骤7004(都是通过对主编码流和辅助编码流进行解码,以得到最终的目标视频图像)。

[0430] 上文结合附图对本申请实施例的视频解码方法和视频编码方法进行了详细的介绍,下面结合图22至图25分别对本申请实施例的视频解码器和视频编码器进行介绍,应理解,图22至图25所示的视频解码器和视频编码器能够分别执行本申请实施例的视频解码方法和视频编码方法中的各个步骤。为了避免不必要的重复,下面在介绍本申请实施例的视频解码器和视频编码器时适当省略重复的描述。

[0431] 图22是本申请实施例的视频解码器的示意性框图。图22所示的视频解码器30000包括:解码模块30010和处理模块30020。

[0432] 视频解码器30000中的解码模块30010和处理模块30020可以执行本申请实施例的视频解码方法的各个步骤,例如,解码模块30010和处理模块30020既能够执行图8所示的方法中的步骤1001至步骤1004,又能够执行图17所示的方法中的步骤7001至步骤7004。

[0433] 当解码模块30010和处理模块30020执行图8所示的方法中的步骤1001至步骤1004时,解码模块30010和处理模块30020的具体作用如下:

[0434] 解码模块30010,用于解码主编码流,获取第一视频图像;

[0435] 处理模块30020,用于采用神经网络模型对所述第一视频图像进行超分辨率处理,得到第二视频图像;

[0436] 所述解码模块30010还用于解码辅助编码流,获取残差值;

[0437] 所述处理模块30020还用于将所述第二视频图像的像素值与所述残差值叠加,得到目标视频图像;

[0438] 可选地,上述神经网络模型是根据训练样本图像和训练目标图像训练得到的,该训练目标图像是与所述目标视频图像的分辨率相同的视频图像,该训练样本图像是对所述训练目标图像进行下采样、编码和解码处理得到的视频图像。

[0439] 当解码模块30010和处理模块30020执行图17所示的方法中的步骤7001至步骤7004时,解码模块30010和处理模块30020的具体作用如下:

[0440] 解码模块30010,用于解码主编码流,获取第一视频图像,第一视频图像为标准动态范围SDR视频图像;

[0441] 处理模块30020,用于采用神经网络模型对第一视频图像进行处理,得到第二视频图像,第二视频图像为高动态范围HDR视频图像,且第二视频图像的分辨率大于第一视频图像的分辨率;

[0442] 解码模块30010,用于解码辅助编码流,获取残差值;

[0443] 所述处理模块30020还用于将第二视频图像的像素值与残差值叠加,得到目标视频图像。

[0444] 可选地,上述神经网络模型是根据训练样本图像和训练目标图像训练得到的,该训练目标图像是与所述目标视频图像的分辨率相同的视频图像,该训练样本图像是对所述训练目标图像进行下采样、编码和解码处理得到的视频图像。

[0445] 图23是本申请实施例的视频编码器的示意性框图。图23所示的视频编码器40000包括:编码模块40010、解码模块40020和处理模块40030。

[0446] 视频编码器40000中的解码模块30010和处理模块30020可以执行本申请实施例的视频编码方法的各个步骤,例如,解码模块30010和处理模块30020既能够执行图14所示的方法中的步骤4001至步骤4005,又能够执行图18所示的方法中的步骤8001至步骤8005。

[0447] 当视频编码器40000执行图14所示的方法中的步骤4001至步骤4005时,视频编码器40000中的各个模块具体作用如下。

[0448] 编码模块40010,用于对初始视频图像进行下采样和编码处理,得到主编码流;

[0449] 解码模块40020,用于对所述主编码流进行解码,得到第一视频图像;

[0450] 处理模块40030,用于采用神经网络模型对所述第一视频图像进行超分辨率处理,得到与所述初始视频图像的分辨率相同的第二视频图像;

[0451] 所述处理模块40030还用于确定所述初始视频图像相对于所述第二视频图像的残差值;

[0452] 所述编码模块40010还用于对所述残差值进行编码,得到辅助编码流。

[0453] 可选地,上述神经网络模型是根据训练样本图像和训练目标图像进行训练得到的,该训练目标图像是与上述初始视频图像分辨率相同的视频图像,该训练样本图像是对训练目标图像进行下采样、编码和解码处理后得到的视频图像。

[0454] 当视频编码器40000执行图18所示的方法中的步骤8001至步骤8005时,视频编码器40000中的各个模块具体作用如下。

[0455] 处理模块40030,用于对初始视频图像进行处理,得到处理后的视频图像,其中,初始视频图像为高动态范围HDR视频图像,处理后的视频图像为标准动态范围SDR视频图像;

[0456] 编码模块40010,用于对处理后的视频图像进行编码,得到主编码流;

[0457] 解码模块40020,用于对主编码流进行解码,得到第一视频图像;采用神经网络模型对第一视频图像进行处理,得到第二视频图像,第二视频图像为HDR视频图像,且第二视频图像的分辨率与初始视频图像的分辨率相同;

[0458] 所述处理模块40030还用于确定初始视频图像相对于第二视频图像的残差值;

[0459] 所述处理模块40030还用于对残差值进行编码,得到辅助编码流。

[0460] 可选地,上述神经网络模型是根据训练样本图像和训练目标图像进行训练得到的,该训练目标图像是与上述初始视频图像分辨率相同的视频图像,该训练样本图像是对训练目标图像进行下采样、编码和解码处理后得到的视频图像。

[0461] 图24是本申请实施例的视频解码器的示意性框图。图24所示的视频解码器50000包括：

[0462] 存储器50010,用于存储程序；

[0463] 处理器50020,用于执行存储器50010存储的程序,当存储器50010存储的程序被执行时,所述处理器50020用于执行本申请实施例的视频解码方法的各个步骤。

[0464] 图25是本申请实施例的视频编码器的示意性框图。图25所示的视频编码器60000包括：

[0465] 存储器60010,用于存储程序；

[0466] 处理器60020,用于执行存储器60010存储的程序,当存储器60010存储的程序被执行时,所述处理器60020用于执行本申请实施例的视频编码方法的各个步骤。

[0467] 图22中的视频解码器30000和图24中的视频解码器50000的具体结构可以如图2中的编码器20所示,图23中的视频编码器40000和图25中的视频编码器60000的具体结构可以如图2中的编码器20所示。

[0468] 本领域普通技术人员可以意识到,结合本文中所公开的实施例描述的各示例的单元及算法步骤,能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出本申请的范围。

[0469] 所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述描述的系统、装置和单元的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0470] 在本申请所提供的几个实施例中,应该理解到,所揭露的系统、装置和方法,可以通过其它的方式实现。例如,以上所描述的装置实施例仅仅是示意性的,例如,所述单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口,装置或单元的间接耦合或通信连接,可以是电性,机械或其它的形式。

[0471] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0472] 另外,在本申请各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。

[0473] 所述功能如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读取存储介质中。基于这样的理解,本申请的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备)执行本申请各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(read-only memory,ROM)、随机存取存储器(random access memory,RAM)、磁碟或者光盘等各种可以存储程序代码的介质。

[0474] 以上所述,仅为本申请的具体实施方式,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以所述权利要求的保护范围为准。

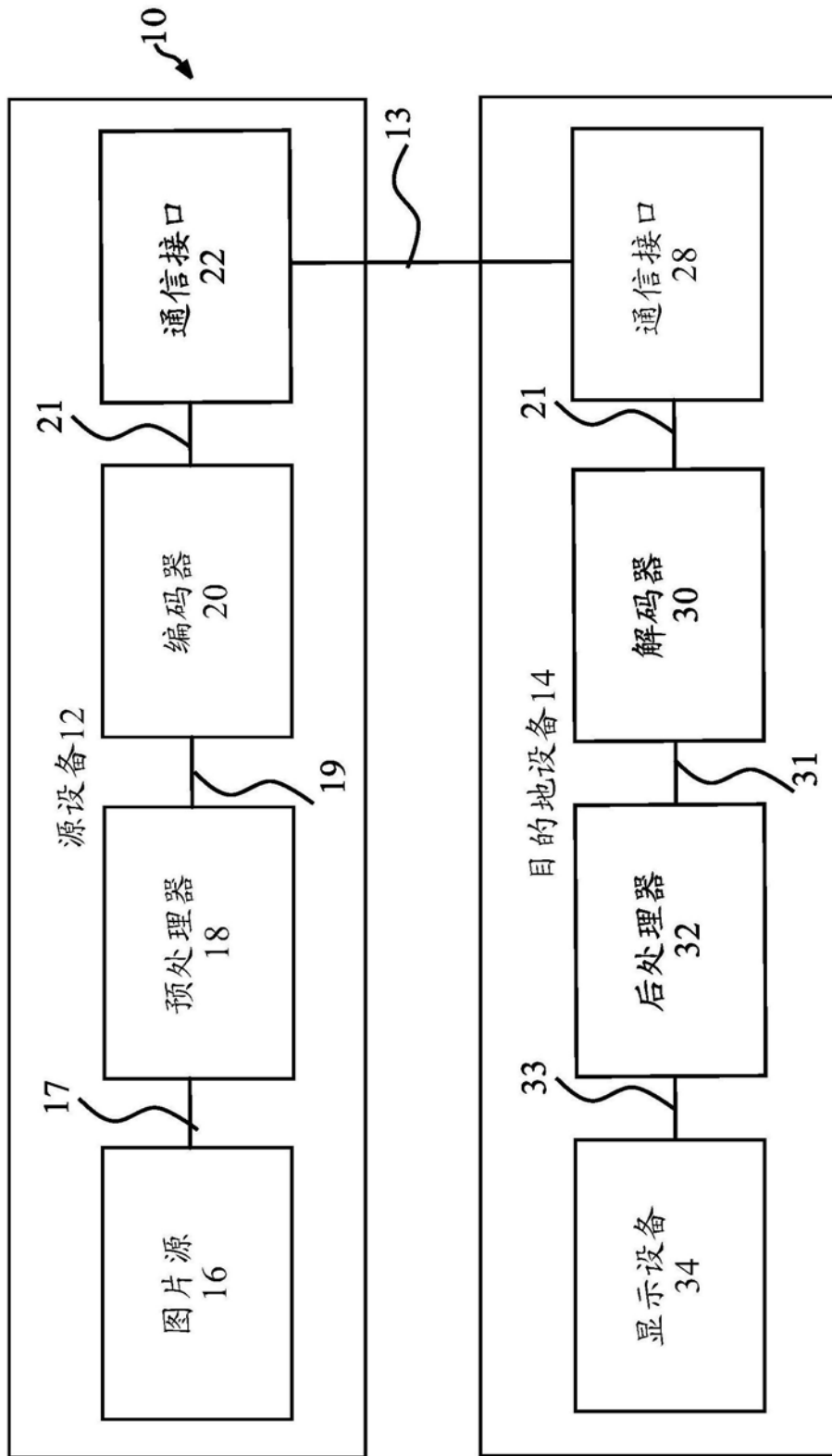


图1



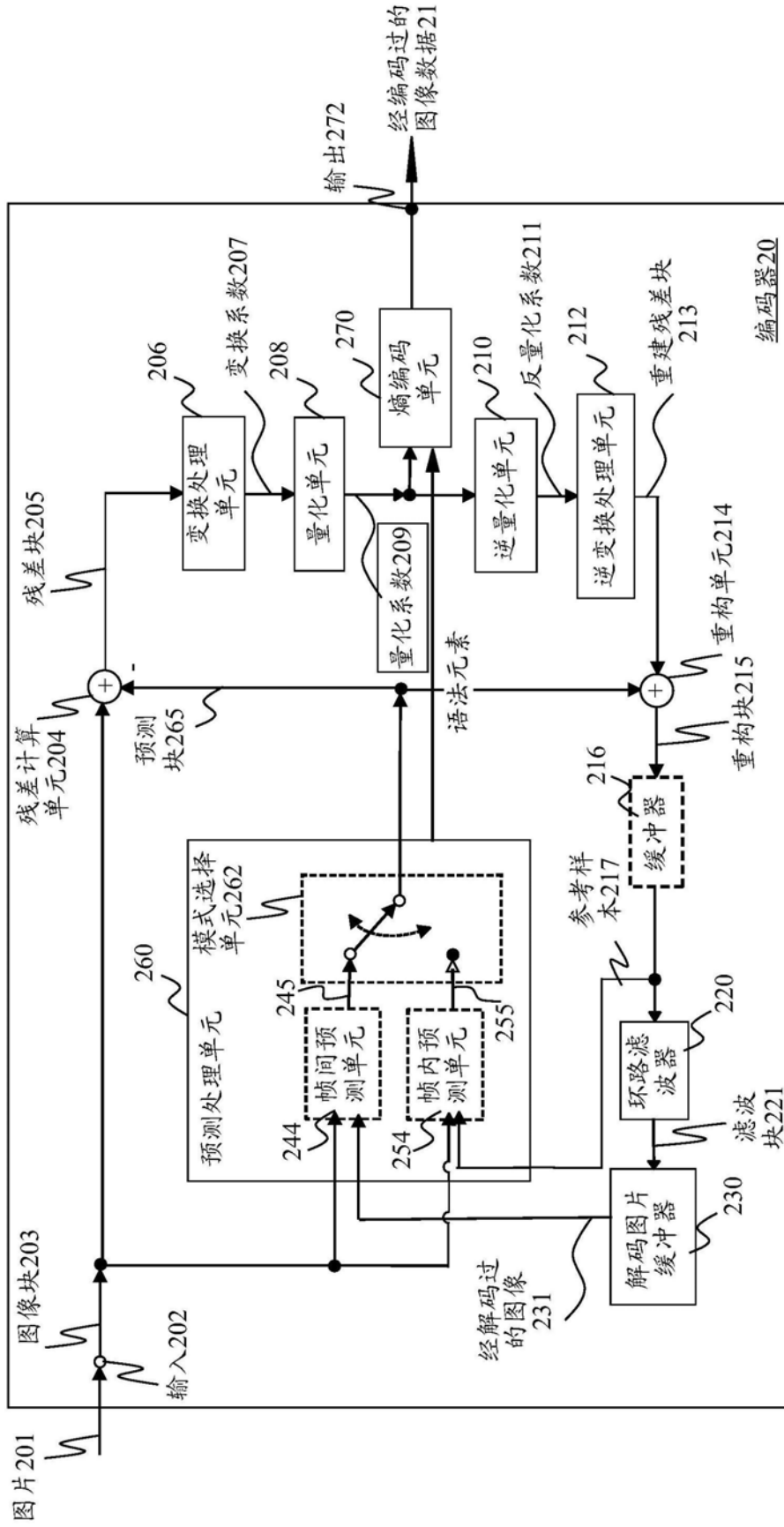


图2

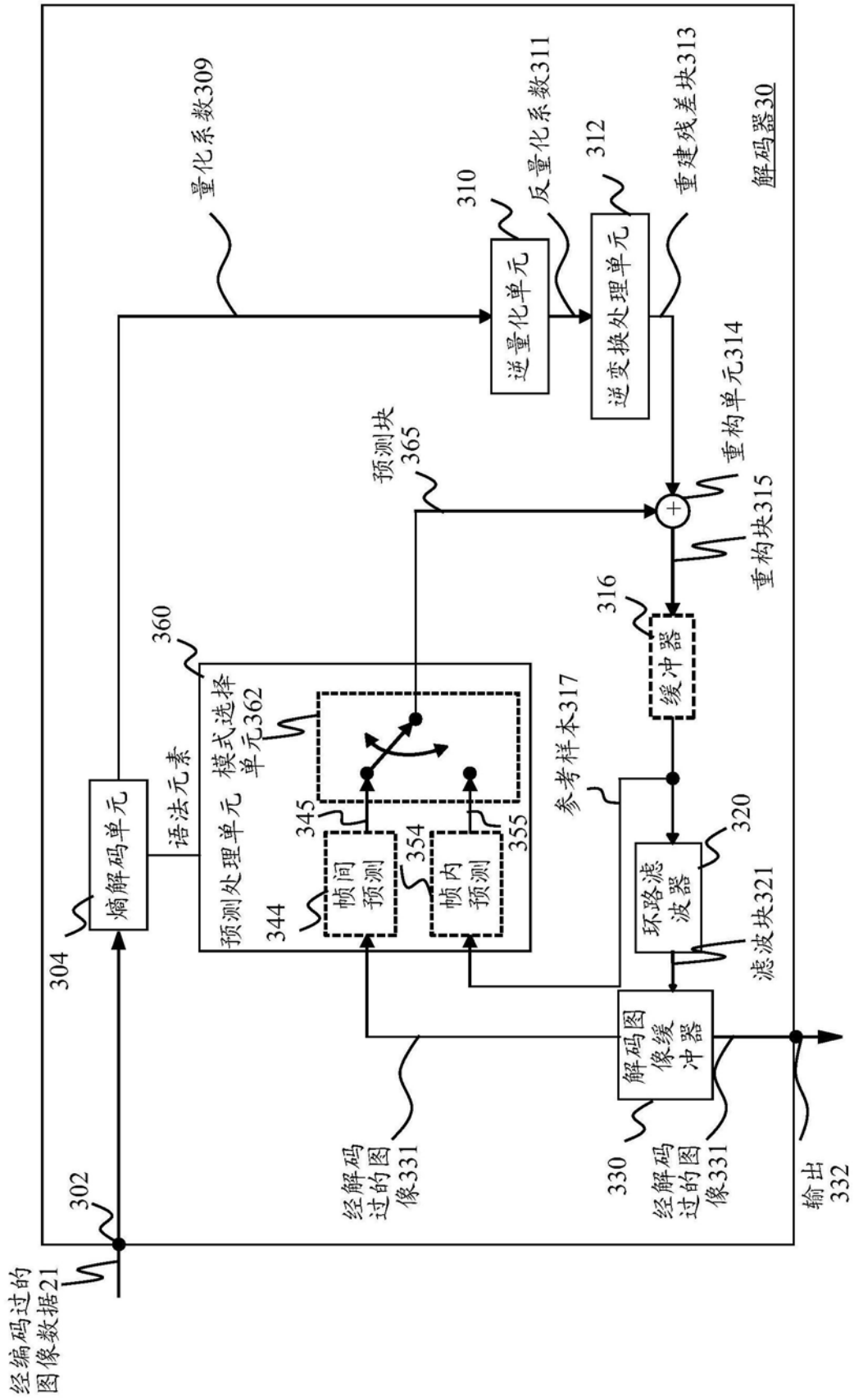


图3

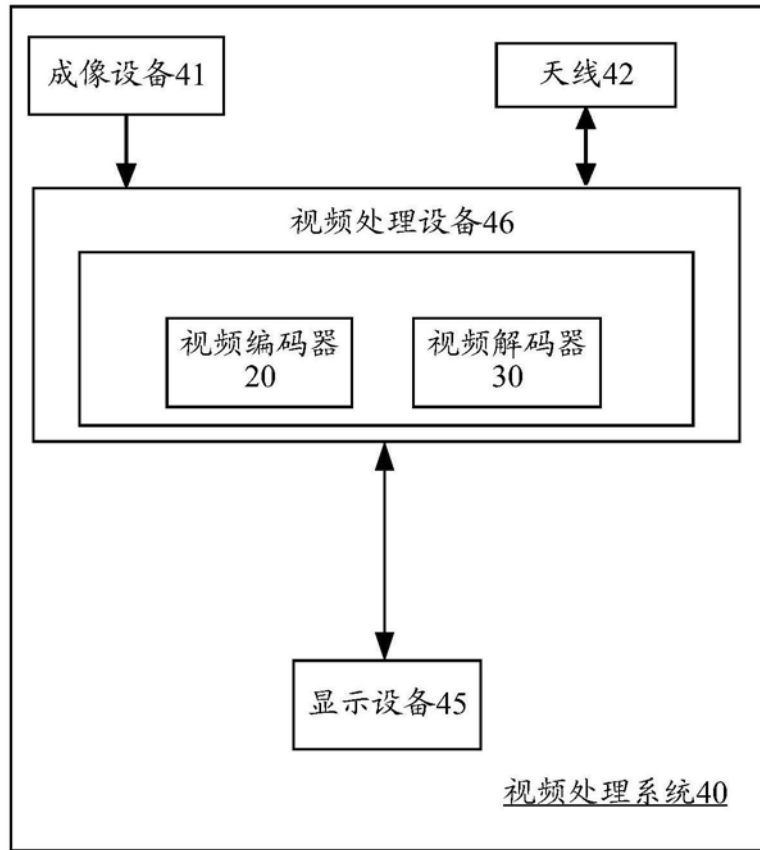


图4

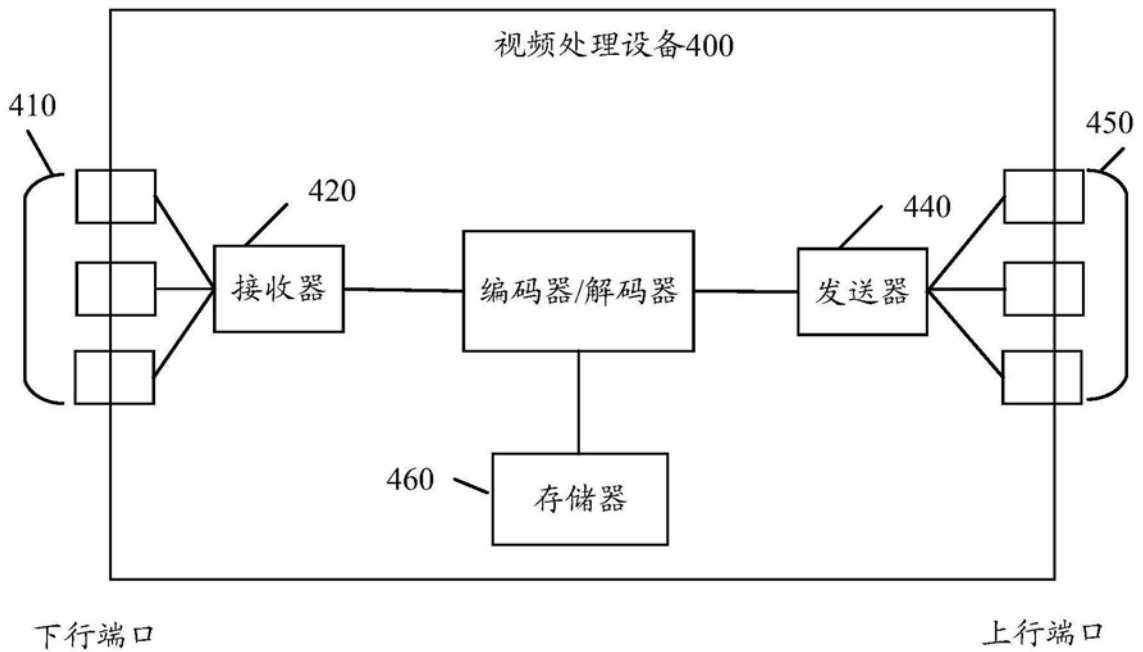


图5

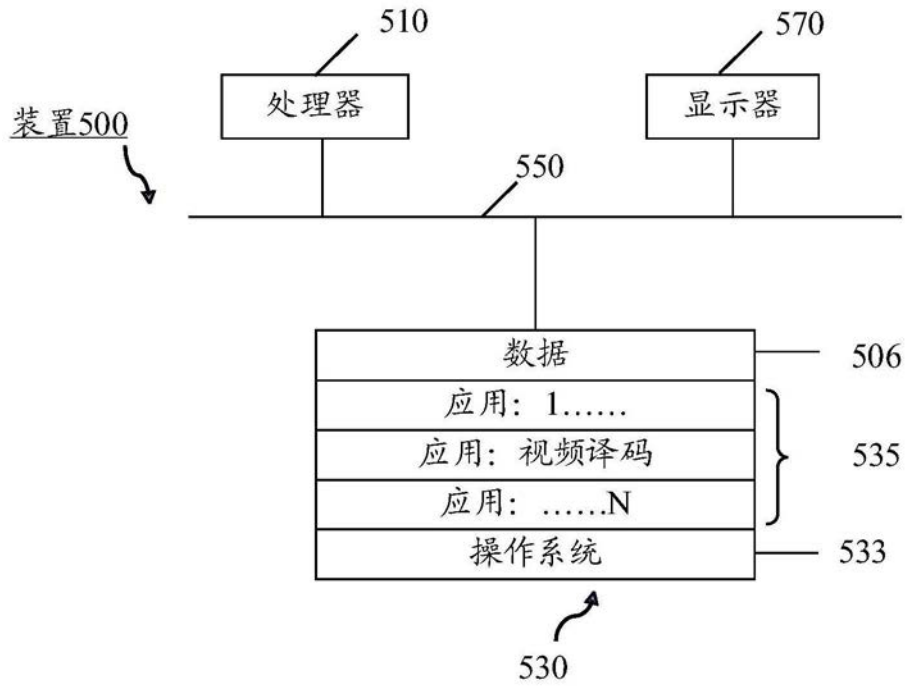


图6

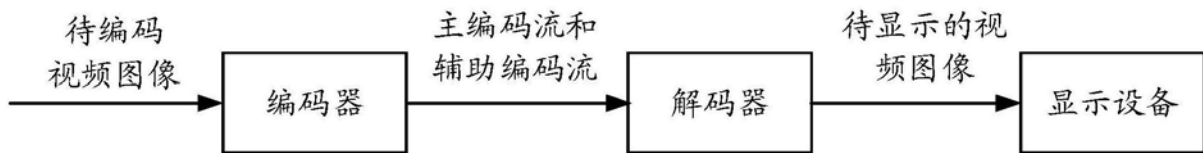


图7

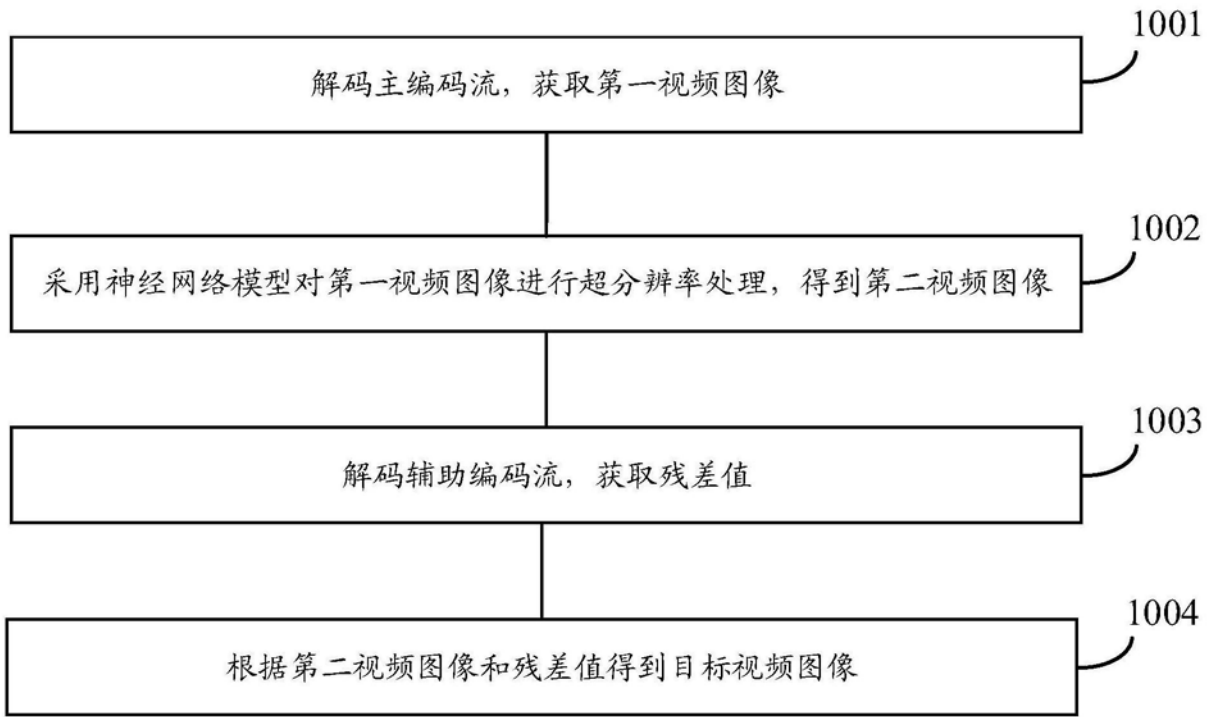


图8

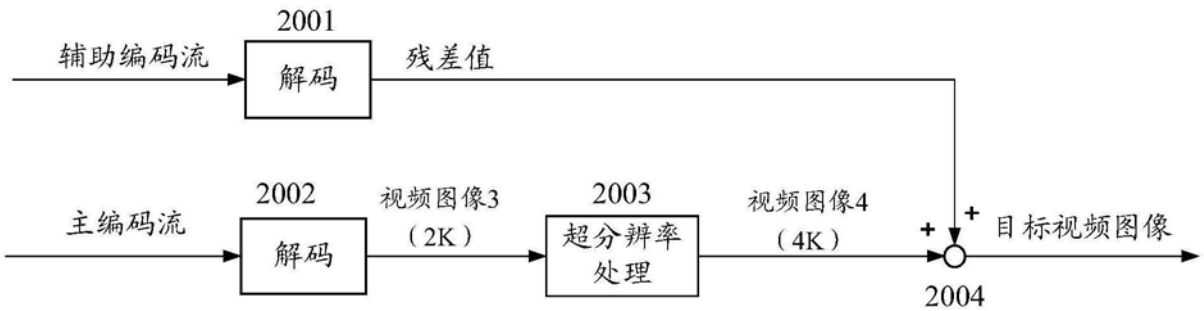


图9

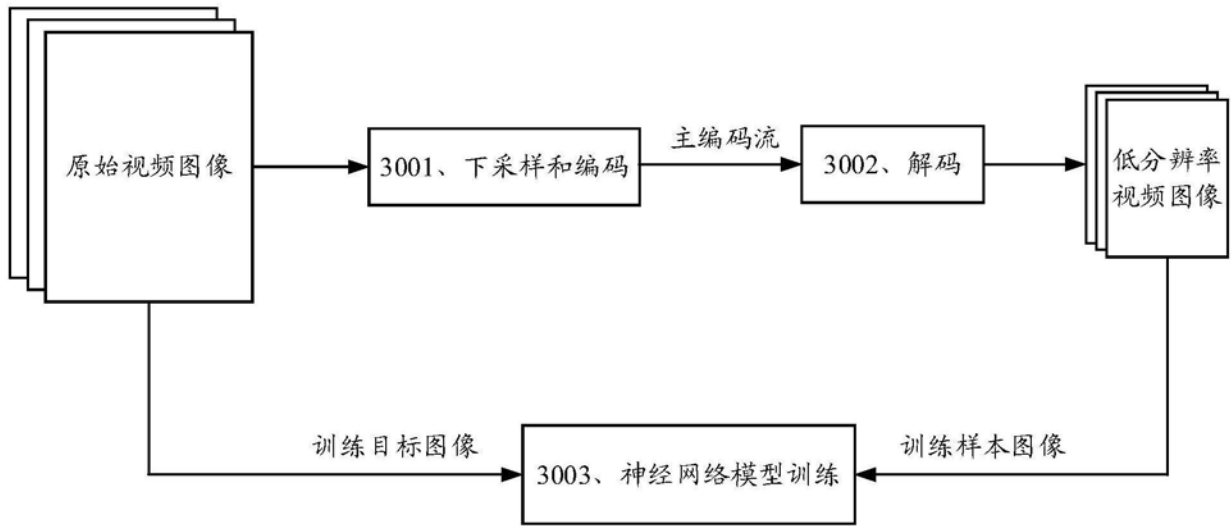


图10

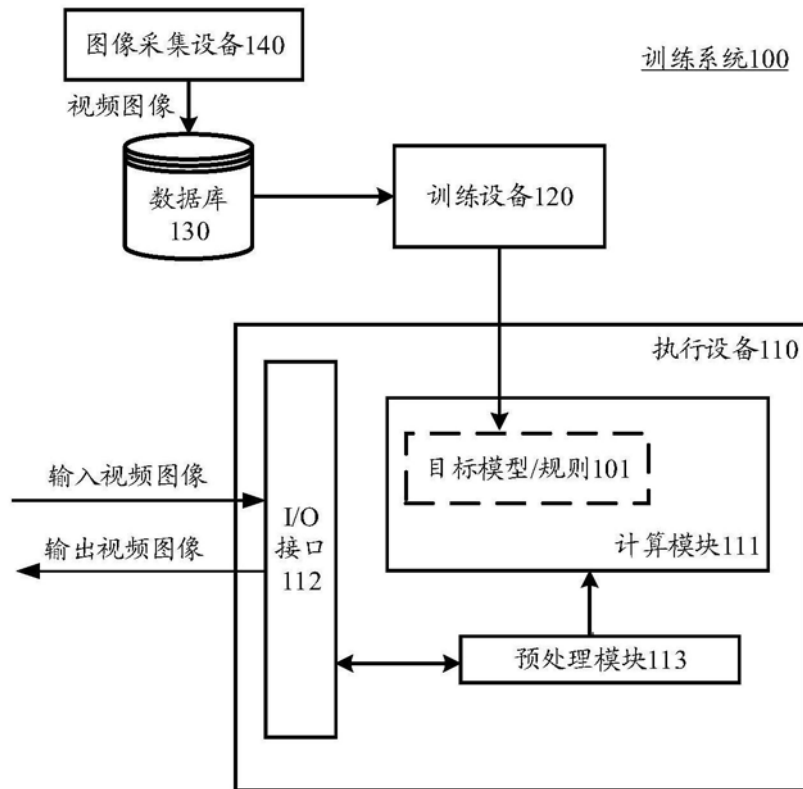


图11

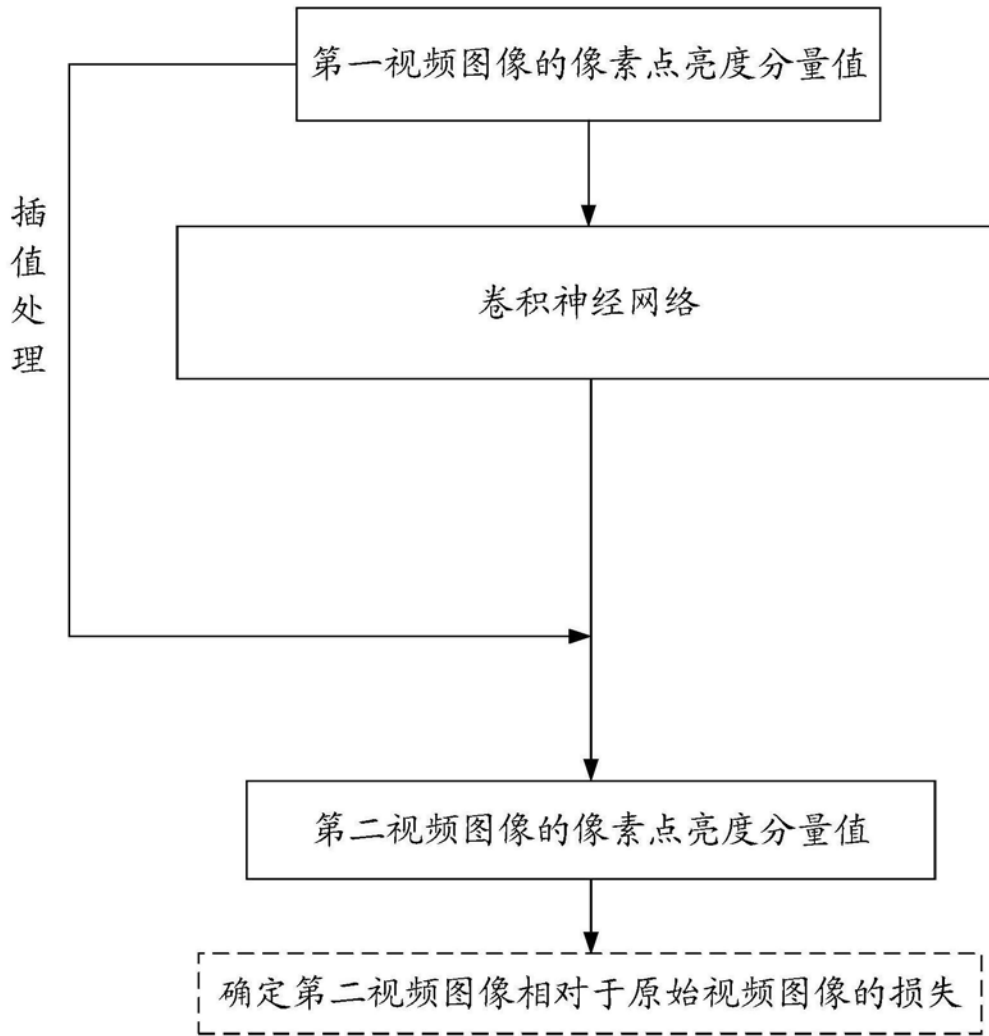


图12

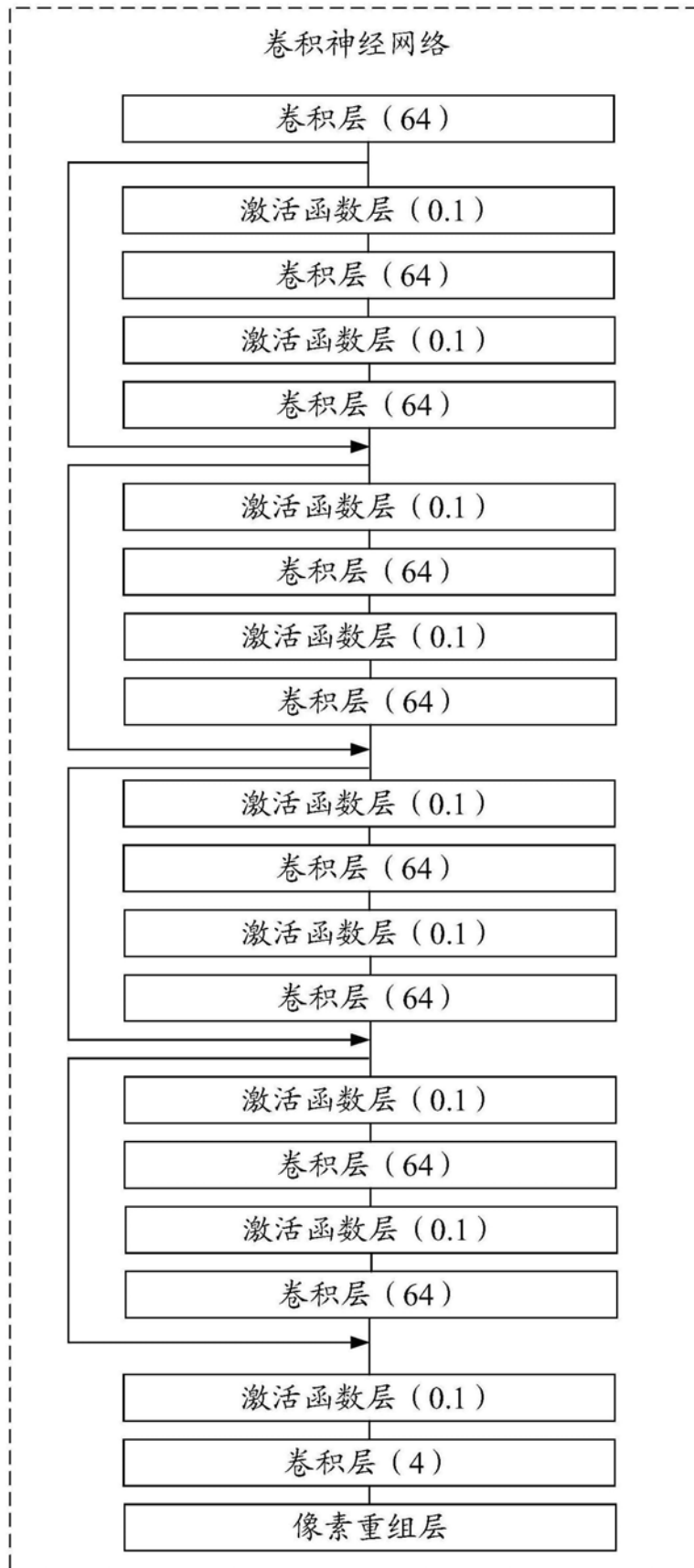


图13



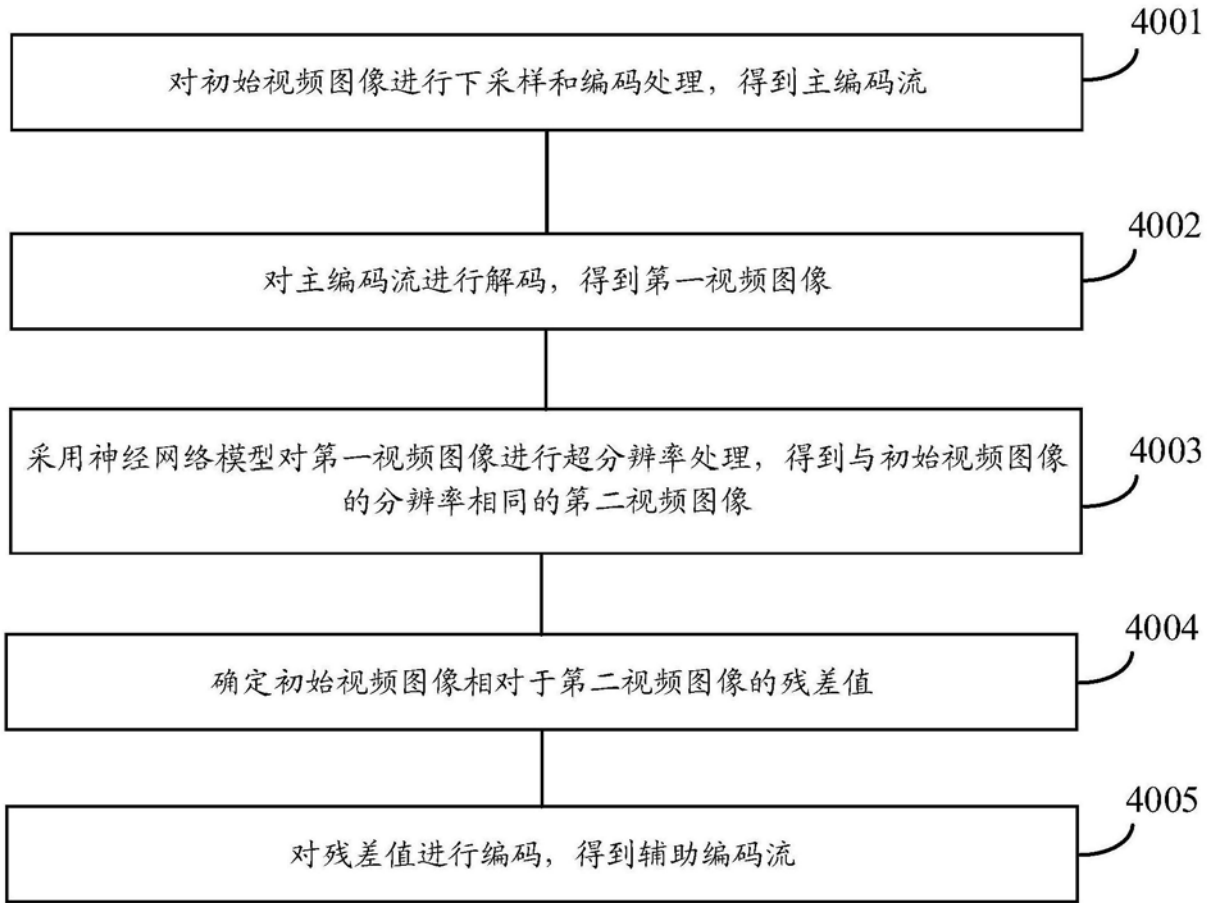


图14

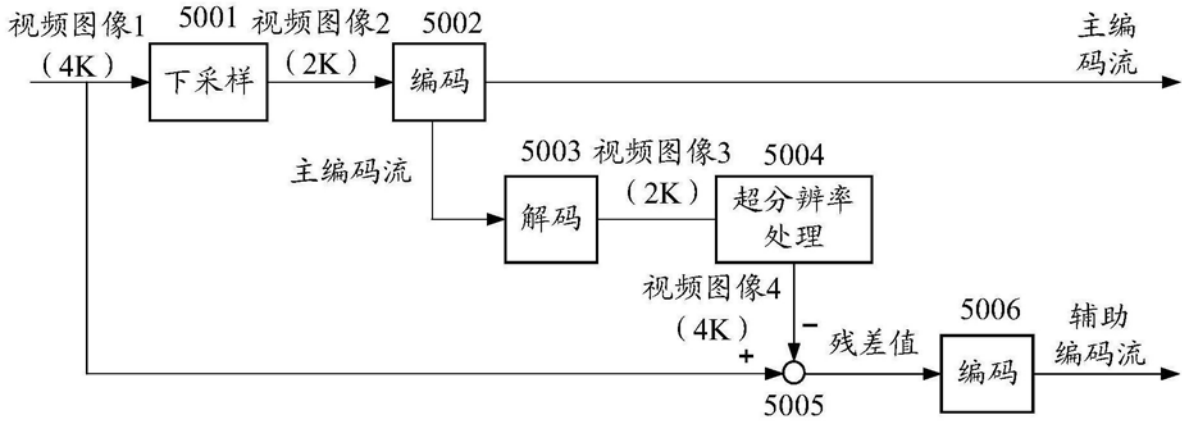


图15

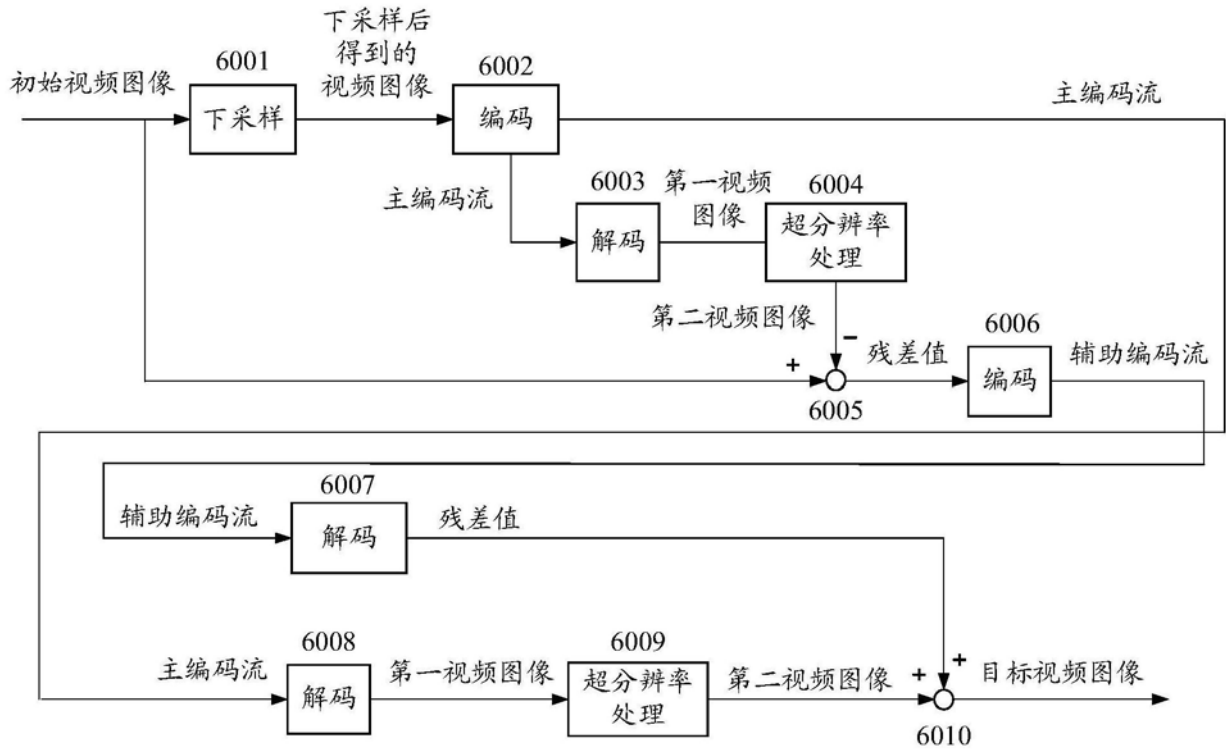


图16

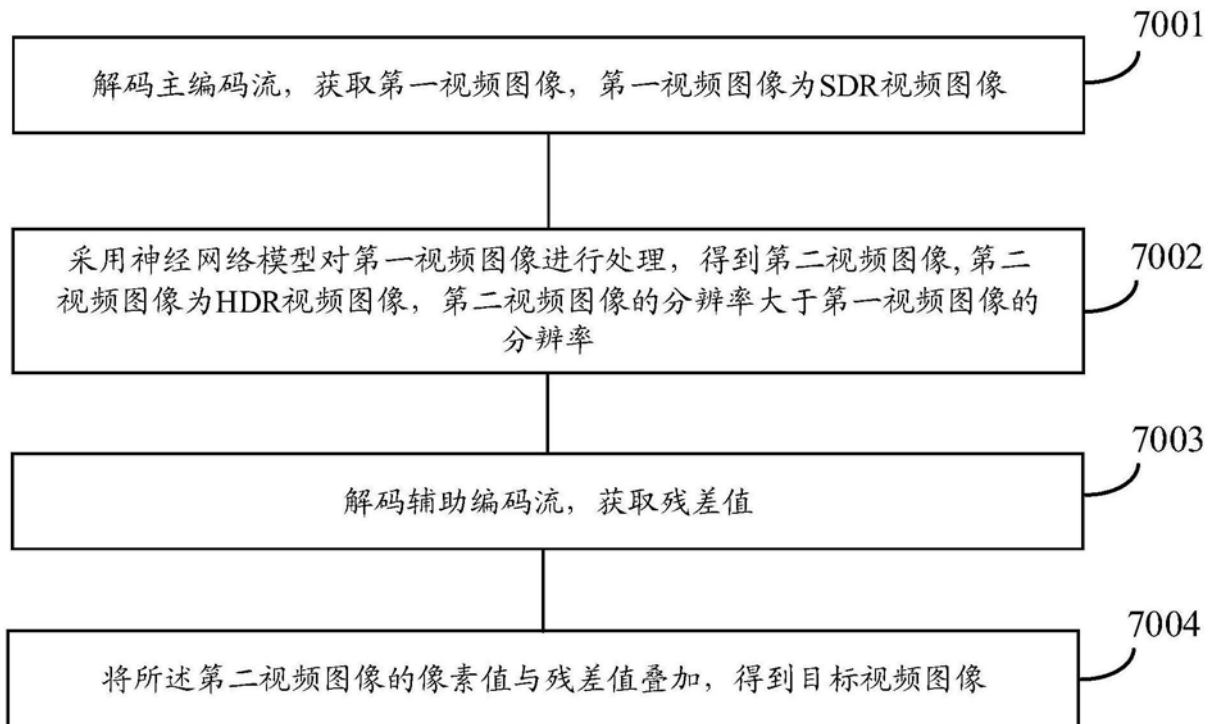


图17



图18

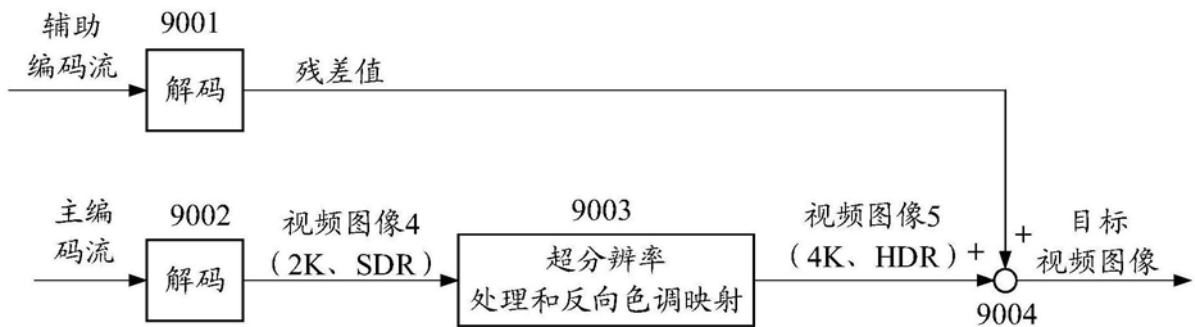


图19

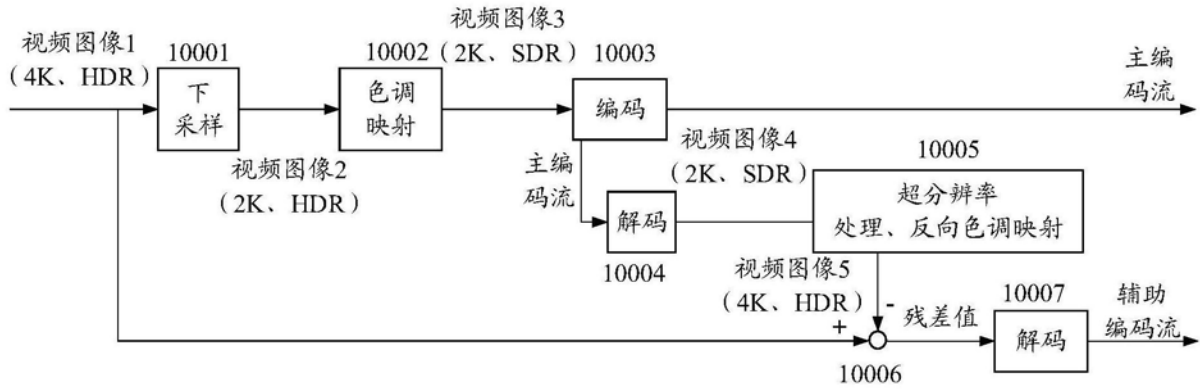


图20

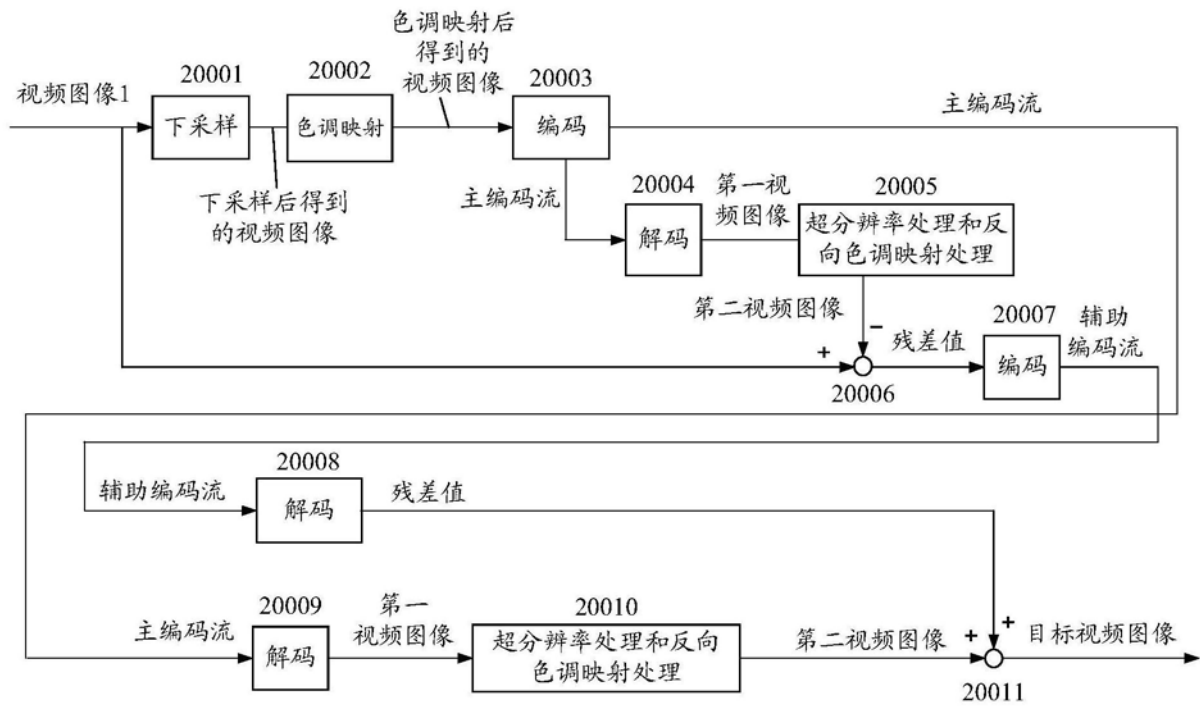


图21

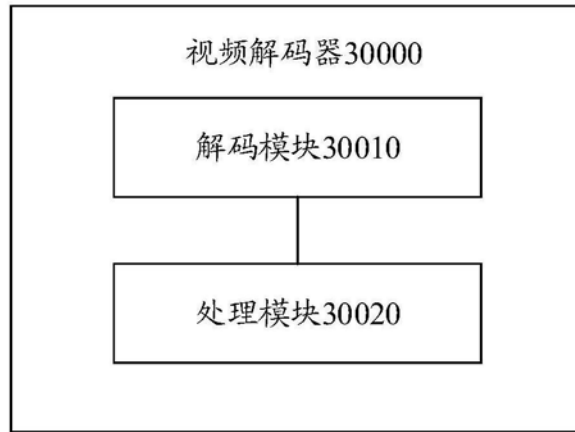


图22

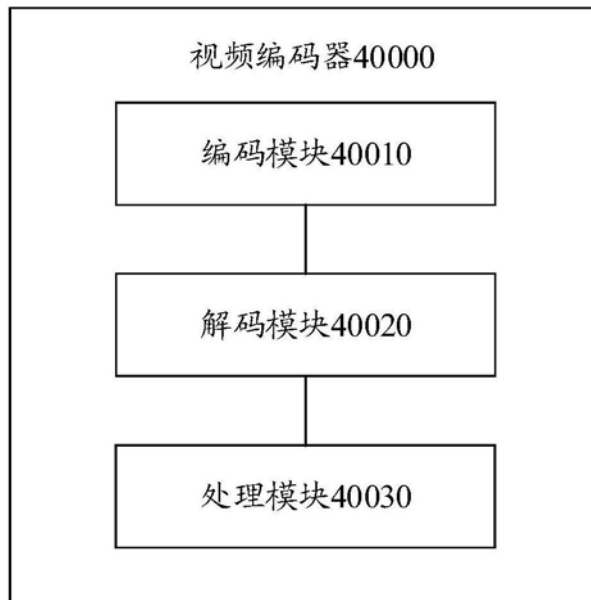


图23

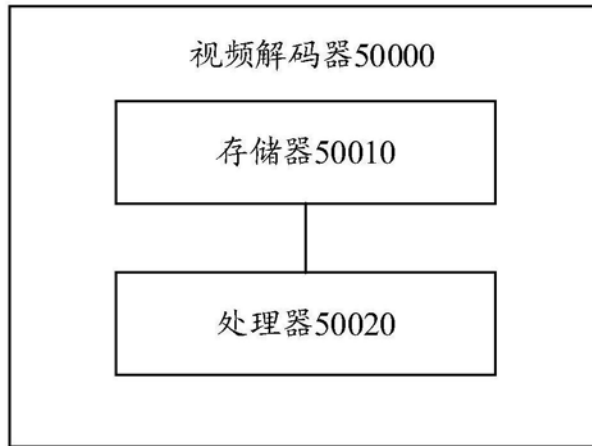


图24

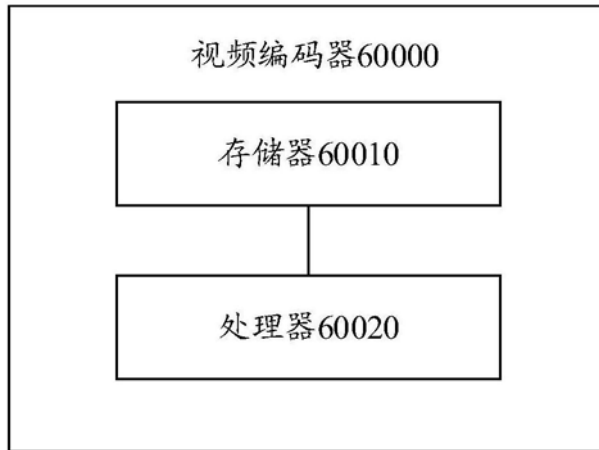


图25