



(12) 发明专利申请

(10) 申请公布号 CN 116939218 A

(43) 申请公布日 2023. 10. 24

(21) 申请号 202210365196.9

(22) 申请日 2022.04.08

(71) 申请人 华为技术有限公司

地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

(72) 发明人 毛珏 赵寅 杨海涛 张恋

(74) 专利代理机构 北京润泽恒知识产权代理有限公司 11319

专利代理师 王洪

(51) Int. Cl.

H04N 19/182 (2014.01)

H04N 19/176 (2014.01)

H04N 19/167 (2014.01)

H04N 19/149 (2014.01)

H04N 19/91 (2014.01)

权利要求书6页 说明书35页 附图11页

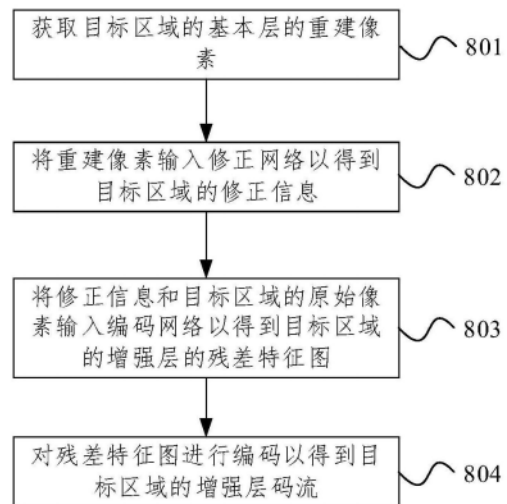
(54) 发明名称

区域增强层的编解码方法和装置

(57) 摘要

本申请提供一种区域增强层的编解码方法和装置。本申请区域增强层的编码方法,包括:获取目标区域的基本层的重建像素;将所述重建像素输入修正网络以得到所述目标区域的修正信息;将所述修正信息和所述目标区域的原始像素输入编码网络以得到所述目标区域的增强层的残差特征图;对所述残差特征图进行编码以得到所述目标区域的增强层码流。本申请可以提高增强层的编码效率和编码精度。

800



1. 一种区域增强层的编码方法,其特征在于,包括:
 - 获取目标区域的基本层的重建像素;
 - 将所述重建像素输入修正网络以得到所述目标区域的修正信息;
 - 将所述修正信息和所述目标区域的原始像素输入编码网络以得到所述目标区域的增强层的残差特征图;
 - 对所述残差特征图进行编码以得到所述目标区域的增强层码流。
2. 根据权利要求1所述的方法,其特征在于,所述将所述重建像素输入修正网络以得到所述目标区域的修正信息,包括:
 - 将所述重建像素输入所述修正网络以得到所述目标区域的多个像素值和多个特征值的至少之一,所述修正信息为所述多个像素值或者所述多个特征值。
3. 根据权利要求1或2所述的方法,其特征在于,所述对所述残差特征图进行编码以得到所述目标区域的增强层码流,包括:
 - 根据所述修正信息获取多个概率分布,所述多个概率分布与所述残差特征图包含的多个特征值对应;
 - 根据所述多个概率分布分别对所述残差特征图中的对应特征值进行熵编码以得到所述增强层码流。
4. 根据权利要求3所述的方法,其特征在于,所述根据所述修正信息获取多个概率分布,包括:
 - 将所述修正信息输入概率估计网络以得到所述多个概率分布。
5. 根据权利要求3所述的方法,其特征在于,所述根据所述修正信息获取多个概率分布,包括:
 - 根据所述修正信息和所述残差特征图的重建边信息获取所述多个概率分布。
6. 根据权利要求5所述的方法,其特征在于,所述根据所述修正信息和所述残差特征图的重建边信息获取所述多个概率分布,包括:
 - 当所述修正信息为多个特征值时,将所述重建边信息输入边信息处理网络以得到第一特征图;
 - 将所述多个特征值和所述第一特征图输入概率估计网络以得到所述多个概率分布。
7. 根据权利要求5所述的方法,其特征在于,所述根据所述修正信息和所述残差特征图的重建边信息获取所述多个概率分布,包括:
 - 当所述修正信息为多个像素值时,将所述多个像素值输入特征估计网络以得到第二特征图;
 - 将所述重建边信息输入边信息处理网络以得到第一特征图;
 - 将所述第一特征图和所述第二特征图输入概率估计网络以得到所述多个概率分布。
8. 根据权利要求1或2所述的方法,其特征在于,所述对所述残差特征图进行编码以得到所述目标区域的增强层码流,包括:
 - 根据所述残差特征图的重建边信息获取多个概率分布,所述多个概率分布与所述残差特征图包含的多个特征值对应;
 - 根据所述多个概率分布分别对所述残差特征图中的对应特征值进行熵编码以得到所述增强层码流。

9. 根据权利要求8所述的方法,其特征在于,所述根据所述残差特征图的重建边信息获取多个概率分布,包括:

将所述重建边信息输入概率估计网络以得到所述多个概率分布。

10. 根据权利要求8所述的方法,其特征在于,所述根据所述残差特征图的重建边信息获取多个概率分布,包括:

根据所述重建边信息和所述重建像素获取所述多个概率分布。

11. 根据权利要求10所述的方法,其特征在于,所述根据所述重建边信息和所述重建像素获取所述多个概率分布,包括:

将所述重建像素输入特征估计网络以得到第三特征图;

将所述重建边信息输入边信息处理网络以得到第一特征图;

将所述第一特征图和所述第三特征图输入概率估计网络以得到所述多个概率分布。

12. 根据权利要求5-11中任一项所述的方法,其特征在于,所述方法还包括:

将所述残差特征图输入边信息提取网络以得到所述残差特征图的边信息;

对所述边信息或经量化的所述边信息进行熵编码,并写入码流。

13. 根据权利要求1-12中任一项所述的方法,其特征在于,所述编码网络包括第一编码网络;所述将所述修正信息和所述目标区域的原始像素输入编码网络以得到所述目标区域的增强层的残差特征图,包括:

当所述修正信息为多个像素值时,对所述原始像素和所述修正信息中对应的像素值求差;

将求差后的结果输入所述第一编码网络以得到所述残差特征图。

14. 根据权利要求1-12中任一项所述的方法,其特征在于,所述编码网络包括第二编码网络;所述将所述修正信息和所述目标区域的原始像素输入编码网络以得到所述目标区域的增强层的残差特征图,包括:

将所述原始像素输入所述第二编码网络;

当所述修正信息为多个特征值时,对所述第二编码网络中的任意一个卷积层的输出和所述修正信息中对应的特征值求差;

将求差后的结果输入所述第二编码网络中的任意一个卷积层之后的网络层以得到所述残差特征图。

15. 根据权利要求1-14中任一项所述的方法,其特征在于,所述方法还包括:

对所述目标区域所属的图像进行编码以得到所述图像的基本层码流;

对所述基本层码流进行解码以得到所述图像的基本层的重建图;

根据所述重建图确定需要增强的至少一个区域,则所述目标区域为所述至少一个区域的其中之一。

16. 根据权利要求15所述的方法,其特征在于,所述根据所述重建图确定需要增强的至少一个区域,包括:

对所述重建图进行划分得到多个区域;

将所述多个区域中方差大于第一阈值的区域确定为所述至少一个区域;或者,

确定所述多个区域中梯度大于第二阈值的像素占总像素的比例,将所述比例大于第三阈值的区域确定为所述至少一个区域。

17. 根据权利要求5-12中任一项所述的方法,其特征在于:

将所述残差特征图的边信息作为所述残差特征图的重建边信息。

18. 一种区域增强层的解码方法,其特征在于,包括:

获取目标区域的基本层的重建像素;

将所述重建像素输入修正网络以得到所述目标区域的修正信息;

获取所述目标区域的增强层码流;

对所述增强层码流进行解码以得到所述目标区域的增强层的残差特征图;

将所述残差特征图和所述修正信息输入解码网络以得到所述目标区域的增强层的重建像素。

19. 根据权利要求18所述的方法,其特征在于,所述将所述重建像素输入修正网络以得到所述目标区域的修正信息,包括:

将所述重建像素输入所述修正网络以得到所述目标区域的多个像素值和多个特征值的至少之一,所述修正信息为所述多个像素值或者所述多个特征值。

20. 根据权利要求18或19所述的方法,其特征在于,所述对所述增强层码流进行解码以得到所述目标区域的增强层的残差特征图,包括:

根据所述修正信息获取所述多个概率分布,所述多个概率分布与所述增强层码流包含的多个特征值码流对应;

根据所述多个概率分布分别对所述增强层码流中的对应特征值码流进行熵解码以得到所述残差特征图。

21. 根据权利要求20所述的方法,其特征在于,所述根据所述修正信息获取所述多个概率分布,包括:

将所述修正信息输入概率估计网络以得到所述多个概率分布。

22. 根据权利要求20所述的方法,其特征在于,所述根据所述修正信息获取所述多个概率分布,包括:

根据所述修正信息和所述残差特征图的重建边信息获取所述多个概率分布。

23. 根据权利要求22所述的方法,其特征在于,所述根据所述修正信息和所述残差特征图的重建边信息获取所述多个概率分布,包括:

当所述修正信息为多个特征值时,将所述重建边信息输入边信息处理网络以得到第一特征图;

将所述多个特征值和所述第一特征图输入概率估计网络以得到所述多个概率分布。

24. 根据权利要求22所述的方法,其特征在于,所述根据所述修正信息和所述残差特征图的重建边信息获取所述多个概率分布,包括:

当所述修正信息为多个像素值时,将所述多个像素值输入特征估计网络以得到第二特征图;

将所述重建边信息输入边信息处理网络以得到第一特征图;

将所述第一特征图和所述第二特征图输入概率估计网络以得到所述多个概率分布。

25. 根据权利要求18或19所述的方法,其特征在于,所述对所述增强层码流进行解码以得到所述目标区域的增强层的残差特征图,包括:

根据所述残差特征图的重建边信息获取所述多个概率分布,所述多个概率分布与所述

增强层码流包含的多个特征值码流对应；

根据所述多个概率分布分别对所述增强层码流中的对应特征值码流进行熵解码以得到所述残差特征图。

26. 根据权利要求25所述的方法,其特征在于,所述根据所述残差特征图的重建边信息获取所述多个概率分布,包括:

将所述重建边信息输入概率估计网络以得到所述多个概率分布。

27. 根据权利要求25所述的方法,其特征在于,所述根据所述残差特征图的重建边信息获取所述多个概率分布,包括:

根据所述重建边信息和所述重建像素获取所述多个概率分布。

28. 根据权利要求27所述的方法,其特征在于,所述根据所述重建边信息和所述重建像素获取所述多个概率分布,包括:

将所述重建像素输入特征估计网络以得到第三特征图;

将所述重建边信息输入边信息处理网络以得到第一特征图;

将所述第一特征图和所述第三特征图输入概率估计网络以得到所述多个概率分布。

29. 根据权利要求22-28中任一项所述的方法,其特征在于,所述方法还包括:

将所述残差特征图输入边信息提取网络以得到所述残差特征图的边信息;

将所述边信息作为所述残差特征图的重建边信息。

30. 根据权利要求22-28中任一项所述的方法,其特征在于,所述方法还包括:

获取所述目标区域的边信息码流;

解析所述边信息码流以得到所述重建边信息。

31. 根据权利要求18-30中任一项所述的方法,其特征在于,所述解码网络包括第一解码网络;所述将所述残差特征图和所述修正信息输入解码网络以得到所述目标区域的增强层的重建像素,包括:

将所述残差特征图输入所述第一解码网络以得到所述目标区域的增强层的重建残差像素;

当所述修正信息为多个像素值时,对所述重建残差像素和所述修正信息中对应的像素值求和以得到所述重建像素。

32. 根据权利要求18-30中任一项所述的方法,其特征在于,所述解码网络包括第二解码网络;所述将所述残差特征图和所述修正信息输入解码网络以得到所述目标区域的增强层的重建像素,包括:

将所述残差特征图输入所述第二解码网络;

当所述修正信息为多个特征值时,对所述第二解码网络中的任意一个卷积层的输出和所述修正信息中对应的特征值求和;

将求和后的结果输入所述第二解码网络中的任意一个卷积层之后的网络层以得到所述重建像素。

33. 根据权利要求18-32中任一项所述的方法,其特征在于,所述方法还包括:

获取所述目标区域所属的图像的基本层码流;

解析所述基本层码流以得到所述图像的基本层的重建图;

根据所述重建图确定需要增强的至少一个区域,则所述目标区域为所述至少一个区域

的其中之一。

34. 根据权利要求33所述的方法,其特征在于,所述根据所述重建图确定需要增强的至少一个区域,包括:

对所述重建图进行划分得到多个区域;

将所述多个区域中方差大于第一阈值的区域确定为所述至少一个区域;或者,

确定所述多个区域中梯度大于第二阈值的像素占总像素的比例,将所述比例大于第三阈值的区域确定为所述至少一个区域。

35. 一种编码装置,其特征在于,包括:

获取模块,用于获取目标区域的基本层的重建像素;

处理模块,用于将所述重建像素输入修正网络以得到所述目标区域的修正信息;将所述修正信息和所述目标区域的原始像素输入编码网络以得到所述目标区域的增强层的残差特征图;

编码模块,用于对所述残差特征图进行编码以得到所述目标区域的增强层码流。

36. 根据权利要求35所述的装置,其特征在于,所述处理模块,具体用于将所述重建像素输入所述修正网络以得到所述目标区域的多个像素值和多个特征值的至少之一,所述修正信息为所述多个像素值或者所述多个特征值。

37. 根据权利要求35或36所述的装置,其特征在于,所述编码模块,具体用于执行权利要求3-17中任一项所述的方法。

38. 一种解码装置,其特征在于,包括:

获取模块,用于获取目标区域的基本层的重建像素;

处理模块,用于将所述重建像素输入修正网络以得到所述目标区域的修正信息;

所述获取模块,还用于获取所述目标区域的增强层码流;

解码模块,用于对所述增强层码流进行解码以得到所述目标区域的增强层的残差特征图;

所述处理模块,还用于将所述残差特征图和所述修正信息输入解码网络以得到所述目标区域的增强层的重建像素。

39. 根据权利要求38所述的装置,其特征在于,所述处理模块,具体用于将所述重建像素输入所述修正网络以得到所述目标区域的多个像素值和多个特征值的至少之一,所述修正信息为所述多个像素值或者所述多个特征值。

40. 根据权利要求38或39所述的装置,其特征在于,所述解码模块,具体用于执行权利要求20-34中任一项所述的方法。

41. 一种编码器,其特征在于,包括:

一个或多个处理器;

存储器,用于存储一个或多个程序;

当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现如权利要求1-17中任一项所述的方法。

42. 一种解码器,其特征在于,包括:

一个或多个处理器;

存储器,用于存储一个或多个程序;

当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现如权利要求18-34中任一项所述的方法。

43.一种计算机可读存储介质,其特征在于,包括计算机程序,所述计算机程序在计算机上被执行时,使得所述计算机执行权利要求1-34中任一项所述的方法。

44.一种计算机程序产品,所述计算机程序产品中包含指令,其特征在于,当所述指令在计算机或处理器上运行时,使得所述计算机或所述处理器实现上述权利要求1-34中任一项所述的方法。

区域增强层的编解码方法和装置

技术领域

[0001] 本申请涉及视频编解码技术,尤其涉及一种区域增强层的编解码方法和装置。

背景技术

[0002] 在视频传输过程中,尤其是多用户场景下,由于多用户间的网络带宽、设备处理能力存在一定的差异,因而提出了面向不同用户进行自适应的码率调控需求。可分级视频译码,又称可伸缩视频译码,提出时域、空域、质量可分层概念,在基本层的基础上增加增强层的信息,可以获得比基本层更高帧率、更高分辨率、更高质量的视频内容。不同用户可以选择是否需要增强层的码流以匹配各自的网络带宽和设备处理能力。

[0003] 但是,可分级视频译码仍然存在复杂度较高、编解码效率较低的问题。

发明内容

[0004] 本申请提供一种区域增强层的编解码方法和装置,以提高增强层的编码效率和编码精准度。

[0005] 第一方面,本申请提供一种区域增强层的编码方法,包括:获取目标区域的基本层的重建像素;将所述重建像素输入修正网络以得到所述目标区域的修正信息;将所述修正信息和所述目标区域的原始像素输入编码网络以得到所述目标区域的增强层的残差特征图;对所述残差特征图进行编码以得到所述目标区域的增强层码流。

[0006] 本申请实施例,通过修正网络对基本层的重建像素去除对AI增强层编码无益的噪声信号,以得到修正信息,再根据该修正信息对目标区域的增强层的残差特征图进行编码,一方面只对有需要的区域(目标区域)进行增强层编码,可以降低增强层编码的复杂度,提高增强层编码效率,另一方面基于修正信息进行编码,可以提高编码的精准度。

[0007] 目标区域意在表示本申请实施例的方案在一次编码过程中所关注并处理的图像块的位置,而目标区域的形状可以是规则的矩形或正方形,也可以是不规则图形,对此不做具体限定。

[0008] 在编码过程中,初始得到的图像块可以称作原始块,其所包含的像素可以称作原始像素;重建的到的图像块可以称作重建块,其所包含的像素可以称作重建像素。而在可分级视频译码中,虽然不同层的分辨率、帧率或质量不同,但编码过程大致相似,尤其是各层均包括初始图像块和重建图像块。

[0009] 因此在本申请实施例中,针对目标区域,可以有以下概念:

[0010] 在基本层,初始得到的区域所包含的像素称作该区域的基本层的原始像素,重建得到的区域所包含的像素称作该区域的基本层的重建像素。

[0011] 在增强层,初始得到的区域所包含的像素称作该区域的增强层的原始像素,重建得到的区域所包含的像素称作该区域的增强层的重建像素。

[0012] 需要说明的是,上述在描述原始像素或重建像素时均未提到像素数量的概念,但应理解,区域中所包含的像素通常是多个像素,本申请实施例对此不做具体限定。

[0013] 在一种可能的实现方式中,将所述重建像素输入所述修正网络以得到多个像素值和多个特征值的至少之一,所述修正信息为所述多个像素值或者所述多个特征值。

[0014] 本申请实施例可以采用神经网络来实现修正网络。示例性的,使用四个卷积层/反卷积层与三个激活层穿插级联构成的神经网络来构建修正网络,修正网络的输入为目标区域的基本层的重建像素,输出为目标区域对应的修正信息,其作用是去除对AI增强层编码无益的噪声信号。其中,卷积层的卷积核尺寸可设置为 3×3 ,输出特征图的通道数设置为M,且每层卷积层的宽高均进行2倍下采样。应理解,前述示例并不构成具体限定。其中,卷积核大小、特征图通道数、下采样倍数、下采样次数、卷积层数、激活层数均可调整,本申请实施例对此不作具体限定。

[0015] 本申请实施例中,可以获取多个概率分布,该多个概率分布与残差特征图包含的多个特征值对应,再根据多个概率分布分别对残差特征图中的对应特征值进行熵编码以得到增强层码流。

[0016] 目标区域的增强层的残差特征图包含多个特征值,为了对该多个特征值进行熵编码,需要得到该多个特征值各自的概率分布。以下是多个概率分布的几种获取方法:

[0017] 1、根据修正信息获取多个概率分布。

[0018] 1.1、将修正信息输入概率估计网络以得到多个概率分布。

[0019] 概率估计网络也可以包括卷积层和GDN,至于概率估计网络是否其它激活函数不做限定,而本申请实施例对卷积层的卷积层数不限定,对卷积核的大小也不做限定。本申请实施例中,首先使用概率分布模型进行建模,然后将修正信息输入概率估计网络得到模型参数,将模型参数代入概率分布模型中,得到概率分布。其中,概率分布模型可以是单高斯模型(gaussian single model,GSM)、非对称高斯模型、混合高斯模型(gaussian mixture model,GMM)或者拉普拉斯分布模型(laplace distribution)。当概率分布为高斯分布时,模型参数为高斯分布的均值参数 μ 和方差 σ 的值。当概率分布为拉普拉斯分布时,模型参数为拉普拉斯分布的位置参数 μ 、尺度参数 b 的值。应理解,除上述概率分布模型之外,还可以采用其它模型,对此不做具体限定。

[0020] 1.2、根据修正信息和残差特征图的重建边信息获取多个概率分布。

[0021] 本申请实施例中,可以将目标区域的增强层的残差特征图输入边信息提取网络以得到残差特征图的边信息,对该边信息进行熵编码,并写入码流,与此同时将前述残差特征图的边信息作为残差特征图的重建边信息。

[0022] a、当修正信息为多个特征值时,将重建边信息输入边信息处理网络以得到第一特征图,将多个特征值和第一特征图输入概率估计网络以得到多个概率分布。

[0023] b、当修正信息为多个像素值时,将多个像素值输入特征估计网络(estimate)以得到第二特征图,将重建边信息输入边信息处理网络以得到第一特征图,将第一特征图和第二特征图输入概率估计网络以得到多个概率分布。

[0024] 2、根据残差特征图的重建边信息获取多个概率分布。

[0025] 2.1、将重建边信息输入概率估计网络以得到多个概率分布。

[0026] 2.2、根据重建边信息和重建像素获取多个概率分布。

[0027] 本申请实施例中,可以将残差特征图的重建像素输入特征估计网络以得到第三特征图,将重建边信息输入边信息处理网络以得到第一特征图,将第一特征图和第三特征图

输入概率估计网络以得到多个概率分布。

[0028] 在一种可能的实现方式中,所述方法还包括:将所述残差特征图输入边信息提取网络以得到所述残差特征图的边信息;对所述边信息或经量化的所述边信息进行熵编码,并写入码流。

[0029] 在一种可能的实现方式中,所述编码网络包括第一编码网络;所述将所述修正信息和所述目标区域的原始像素输入编码网络以得到所述目标区域的增强层的残差特征图,包括:当所述修正信息为多个像素值时,对所述原始像素和所述修正信息中对应的像素值求差;将求差后的结果输入所述第一编码网络以得到所述残差特征图。例如,对应的可以理解为位置对应,即将所述原始像素和所述修正信息中对应位置的像素值求差。

[0030] 在一种可能的实现方式中,所述编码网络包括第二编码网络;所述将所述修正信息和所述目标区域的原始像素输入编码网络以得到所述目标区域的增强层的残差特征图,包括:将所述原始像素输入所述第二编码网络;当所述修正信息为多个特征值时,对所述第二编码网络中的任意一个卷积层的输出和所述修正信息中对应的特征值求差;将求差后的结果输入所述第二编码网络中的任意一个卷积层之后的网络层以得到所述残差特征图。例如,对应的可以理解为位置对应,即将所述第二编码网络中的任意一个卷积层的输出和所述修正信息中对应位置的特征值求差。

[0031] 如上所述,修正信息可以有两种情况,一种是多个像素值,另一种是多个特征值。相应的,编码网络也可以采用两种结构。本申请实施例中,编码端的编码网络(Encoder),其输入为修正信息和目标区域的原始像素,输出为目标区域的增强层的残差特征图。需要说明的是,除了上述两种编码网络的示例外,本申请实施例还可以用其他结构的编码网络,对此不做具体限定。

[0032] 在一种可能的实现方式中,获取目标区域的基本层的重建像素可以包括:对所述目标区域所属的图像进行编码以得到所述图像的基本层码流;对所述基本层码流进行解码以得到所述图像的基本层的重建图;根据所述重建图确定需要增强的至少一个区域,则所述目标区域为所述至少一个区域的其中之一。

[0033] 其中,根据所述重建图确定需要增强的至少一个区域,包括:对所述重建图进行划分得到多个区域;将所述多个区域中方差大于第一阈值的区域确定为所述至少一个区域;或者,确定所述多个区域中梯度大于第二阈值的像素占总像素的比例,将所述比例大于第三阈值的区域确定为所述至少一个区域。例如,区域的方差大于阈值 t_1 , $t_1 > 0$,可以认为该区域的纹理比较复杂,因此需要做增强处理,以提高图像质量;或者,区域中梯度大于第二阈值的像素占总像素的比例大于第三阈值。例如,区域中梯度大于阈值 a 的像素占总像素的比例大于阈值 t_2 , $a > 0$, $0 < t_2 < 1$,也可以认为该区域的纹理比较复杂,因此需要做增强处理,以提高图像质量。

[0034] 针对基本层,编码端对原始图进行编码得到基本层码流,再对基本层码流进行解码得到基本层的重建图。例如,VVC编码器对原始图进行编码得到基本层码流,VVC解码器对基本层码流进行解码得到基本层的重建图。应理解,针对基本层还可以采用其它编码器,HEVC编解码器、AVC编解码器,本申请实施例对此不做具体限定。

[0035] 在上述多个区域中,根据图像的属性或特征,可能不是所有的区域都需要做增强处理,某一个或多个区域只需要做基本层的编解码即可满足画质需求,因此本申请实施例

中的目标区域是需要做增强处理的区域,而在后续的增强层编码中可以只对该目标区域进行增强层的编解码,这样不需要对整帧图像进行增强处理,可以提高图像的编解码效率。

[0036] 在一种可能的实现方式中,所述方法还包括:将所述残差特征图的边信息作为所述残差特征图的重建边信息。

[0037] 第二方面,本申请提供一种区域增强层的解码方法,包括:获取目标区域的基本层的重建像素;将所述重建像素输入修正网络以得到所述目标区域的修正信息;获取所述目标区域的增强层码流;对所述增强层码流进行解码以得到所述目标区域的增强层的残差特征图;将所述残差特征图和所述修正信息输入解码网络以得到所述目标区域的增强层的重建像素。

[0038] 本申请实施例,通过修正网络对基本层的重建像素去除对AI增强层编码无益的噪声信号,以得到修正信息,再根据该修正信息对增强层码流进行解码,一方面只对有需要的区域(目标区域)进行增强层解码,可以降低增强层解码的复杂度,提高增强层解码效率,另一方面基于修正信息进行解码,可以提高解码的精准度。

[0039] 在一种可能的实现方式中,所述将所述重建像素输入修正网络以得到所述目标区域的修正信息,包括:将所述重建像素输入所述修正网络以得到所述目标区域的多个像素值和多个特征值的至少之一,所述修正信息为所述多个像素值或者所述多个特征值。

[0040] 在一种可能的实现方式中,所述对所述增强层码流进行解码以得到所述目标区域的增强层的残差特征图,包括:根据所述修正信息获取所述多个概率分布,所述多个概率分布与所述增强层码流包含的多个特征值码流对应;根据所述多个概率分布分别对所述增强层码流中的对应特征值码流进行熵解码以得到所述残差特征图。

[0041] 在一种可能的实现方式中,所述根据所述修正信息获取所述多个概率分布,包括:将所述修正信息输入概率估计网络以得到所述多个概率分布。

[0042] 在一种可能的实现方式中,所述根据所述修正信息获取所述多个概率分布,包括:根据所述修正信息和所述残差特征图的重建边信息获取所述多个概率分布。

[0043] 在一种可能的实现方式中,所述根据所述修正信息和所述残差特征图的重建边信息获取所述多个概率分布,包括:当所述修正信息为多个特征值时,将所述重建边信息输入边信息处理网络以得到第一特征图;将所述多个特征值和所述第一特征图输入概率估计网络以得到所述多个概率分布。

[0044] 在一种可能的实现方式中,所述根据所述修正信息和所述残差特征图的重建边信息获取所述多个概率分布,包括:当所述修正信息为多个像素值时,将所述多个像素值输入特征估计网络以得到第二特征图;将所述重建边信息输入边信息处理网络以得到第一特征图;将所述第一特征图和所述第二特征图输入概率估计网络以得到所述多个概率分布。

[0045] 在一种可能的实现方式中,所述对所述增强层码流进行解码以得到所述目标区域的增强层的残差特征图,包括:根据所述残差特征图的重建边信息获取所述多个概率分布,所述多个概率分布与所述增强层码流包含的多个特征值码流对应;根据所述多个概率分布分别对所述增强层码流中的对应特征值码流进行熵解码以得到所述残差特征图。

[0046] 在一种可能的实现方式中,所述根据所述残差特征图的重建边信息获取所述多个概率分布,包括:将所述重建边信息输入概率估计网络以得到所述多个概率分布。

[0047] 在一种可能的实现方式中,所述根据所述残差特征图的重建边信息获取所述多个

概率分布,包括:根据所述重建边信息和所述重建像素获取所述多个概率分布。

[0048] 在一种可能的实现方式中,所述根据所述重建边信息和所述重建像素获取所述多个概率分布,包括:将所述重建像素输入特征估计网络以得到第三特征图;将所述重建边信息输入边信息处理网络以得到第一特征图;将所述第一特征图和所述第三特征图输入概率估计网络以得到所述多个概率分布。

[0049] 在一种可能的实现方式中,所述方法还包括:将所述残差特征图输入边信息提取网络以得到所述残差特征图的边信息;将所述边信息作为所述残差特征图的重建边信息。

[0050] 在一种可能的实现方式中,所述方法还包括:获取所述目标区域的边信息码流;解析所述边信息码流以得到所述重建边信息。

[0051] 在一种可能的实现方式中,所述解码网络包括第一解码网络;所述将所述残差特征图和所述修正信息输入解码网络以得到所述目标区域的增强层的重建像素,包括:将所述残差特征图输入所述第一解码网络以得到所述目标区域的增强层的重建残差像素;当所述修正信息为多个像素值时,对所述重建残差像素和所述修正信息中对应的像素值求和以得到所述重建像素。

[0052] 在一种可能的实现方式中,所述解码网络包括第二解码网络;所述将所述残差特征图和所述修正信息输入解码网络以得到所述目标区域的增强层的重建像素,包括:将所述残差特征图输入所述第二解码网络;当所述修正信息为多个特征值时,对所述第二解码网络中的任意一个卷积层的输出和所述修正信息中对应的特征值求和;将求和后的结果输入所述第二解码网络中的任意一个卷积层之后的网络层以得到所述重建像素。

[0053] 在一种可能的实现方式中,所述方法还包括:获取所述目标区域所属的图像的基本层码流;解析所述基本层码流以得到所述图像的基本层的重建图;根据所述重建图确定需要增强的至少一个区域,则所述目标区域为所述至少一个区域的其中之一。

[0054] 在一种可能的实现方式中,所述根据所述重建图确定需要增强的至少一个区域,包括:对所述重建图进行划分得到多个区域;将所述多个区域中方差大于第一阈值的区域确定为所述至少一个区域;或者,确定所述多个区域中梯度大于第二阈值的像素占总像素的比例,将所述比例大于第三阈值的区域确定为所述至少一个区域。

[0055] 第三方面,本申请提供一种编码装置,包括:获取模块,用于获取目标区域的基本层的重建像素;处理模块,用于将所述重建像素输入修正网络以得到所述目标区域的修正信息;将所述修正信息和所述目标区域的原始像素输入编码网络以得到所述目标区域的增强层的残差特征图;编码模块,用于对所述残差特征图进行编码以得到所述目标区域的增强层码流。

[0056] 在一种可能的实现方式中,所述处理模块,具体用于将所述重建像素输入所述修正网络以得到所述目标区域的多个像素值和多个特征值的至少之一,所述修正信息为所述多个像素值或者所述多个特征值。

[0057] 在一种可能的实现方式中,所述编码模块,具体用于根据所述修正信息获取多个概率分布,所述多个概率分布与所述残差特征图包含的多个特征值对应;根据所述多个概率分布分别对所述残差特征图中的对应特征值进行熵编码以得到所述增强层码流。

[0058] 在一种可能的实现方式中,所述编码模块,具体用于将所述修正信息输入概率估计网络以得到所述多个概率分布。

[0059] 在一种可能的实现方式中,所述编码模块,具体用于根据所述修正信息和所述残差特征图的重建边信息获取所述多个概率分布。

[0060] 在一种可能的实现方式中,所述编码模块,具体用于当所述修正信息为多个特征值时,将所述重建边信息输入边信息处理网络以得到第一特征图;将所述多个特征值和所述第一特征图输入概率估计网络以得到所述多个概率分布。

[0061] 在一种可能的实现方式中,所述编码模块,具体用于当所述修正信息为多个像素值时,将所述多个像素值输入特征估计网络以得到第二特征图;将所述重建边信息输入边信息处理网络以得到第一特征图;将所述第一特征图和所述第二特征图输入概率估计网络以得到所述多个概率分布。

[0062] 在一种可能的实现方式中,所述编码模块,具体用于根据所述残差特征图的重建边信息获取多个概率分布,所述多个概率分布与所述残差特征图包含的多个特征值对应;根据所述多个概率分布分别对所述残差特征图中的对应特征值进行熵编码以得到所述增强层码流。

[0063] 在一种可能的实现方式中,所述编码模块,具体用于将所述重建边信息输入概率估计网络以得到所述多个概率分布。

[0064] 在一种可能的实现方式中,所述编码模块,具体用于根据所述重建边信息和所述重建像素获取所述多个概率分布。

[0065] 在一种可能的实现方式中,所述编码模块,具体用于将所述重建像素输入特征估计网络以得到第三特征图;将所述重建边信息输入边信息处理网络以得到第一特征图;将所述第一特征图和所述第三特征图输入概率估计网络以得到所述多个概率分布。

[0066] 在一种可能的实现方式中,所述编码模块,还用于将所述残差特征图输入边信息提取网络以得到所述残差特征图的边信息;对所述边信息或经量化的所述边信息进行熵编码,并写入码流。

[0067] 在一种可能的实现方式中,所述编码网络包括第一编码网络;所述编码模块,具体用于当所述修正信息为多个像素值时,对所述原始像素和所述修正信息中对应的像素值求差;将求差后的结果输入所述第一编码网络以得到所述残差特征图。

[0068] 在一种可能的实现方式中,所述编码网络包括第二编码网络;所述编码模块,具体用于将所述原始像素输入所述第二编码网络;当所述修正信息为多个特征值时,对所述第二编码网络中的任意一个卷积层的输出和所述修正信息中对应的特征值求差;将求差后的结果输入所述第二编码网络中的任意一个卷积层之后的网络层以得到所述残差特征图。

[0069] 在一种可能的实现方式中,所述编码模块,还用于对所述目标区域所属的图像进行编码以得到所述图像的基本层码流;对所述基本层码流进行解码以得到所述图像的基本层的重建图;根据所述重建图确定需要增强的至少一个区域,则所述目标区域为所述至少一个区域的其中之一。

[0070] 在一种可能的实现方式中,所述编码模块,具体用于对所述重建图进行划分得到多个区域;将所述多个区域中方差大于第一阈值的区域确定为所述至少一个区域;或者,确定所述多个区域中梯度大于第二阈值的像素占总像素的比例,将所述比例大于第三阈值的区域确定为所述至少一个区域。

[0071] 在一种可能的实现方式中,所述编码模块,还用于将所述残差特征图的边信息作

为所述残差特征图的重建边信息。

[0072] 第四方面,本申请提供一种解码装置,包括:获取模块,用于获取目标区域的基本层的重建像素;处理模块,用于将所述重建像素输入修正网络以得到所述目标区域的修正信息;所述获取模块,还用于获取所述目标区域的增强层码流;解码模块,用于对所述增强层码流进行解码以得到所述目标区域的增强层的残差特征图;所述处理模块,还用于将所述残差特征图和所述修正信息输入解码网络以得到所述目标区域的增强层的重建像素。

[0073] 在一种可能的实现方式中,所述处理模块,具体用于将所述重建像素输入所述修正网络以得到所述目标区域的多个像素值和多个特征值的至少之一,所述修正信息为所述多个像素值或者所述多个特征值。

[0074] 在一种可能的实现方式中,所述解码模块,具体用于根据所述修正信息获取所述多个概率分布,所述多个概率分布与所述增强层码流包含的多个特征值码流对应;根据所述多个概率分布分别对所述增强层码流中的对应特征值码流进行熵解码以得到所述残差特征图。

[0075] 在一种可能的实现方式中,所述解码模块,具体用于将所述修正信息输入概率估计网络以得到所述多个概率分布。

[0076] 在一种可能的实现方式中,所述解码模块,具体用于根据所述修正信息和所述残差特征图的重建边信息获取所述多个概率分布。

[0077] 在一种可能的实现方式中,所述解码模块,具体用于当所述修正信息为多个特征值时,将所述重建边信息输入边信息处理网络以得到第一特征图;将所述多个特征值和所述第一特征图输入概率估计网络以得到所述多个概率分布。

[0078] 在一种可能的实现方式中,所述解码模块,具体用于当所述修正信息为多个像素值时,将所述多个像素值输入特征估计网络以得到第二特征图;将所述重建边信息输入边信息处理网络以得到第一特征图;将所述第一特征图和所述第二特征图输入概率估计网络以得到所述多个概率分布。

[0079] 在一种可能的实现方式中,所述解码模块,具体用于根据所述残差特征图的重建边信息获取所述多个概率分布,所述多个概率分布与所述增强层码流包含的多个特征值码流对应;根据所述多个概率分布分别对所述增强层码流中的对应特征值码流进行熵解码以得到所述残差特征图。

[0080] 在一种可能的实现方式中,所述解码模块,具体用于将所述重建边信息输入概率估计网络以得到所述多个概率分布。

[0081] 在一种可能的实现方式中,所述解码模块,具体用于根据所述重建边信息和所述重建像素获取所述多个概率分布。

[0082] 在一种可能的实现方式中,所述解码模块,具体用于将所述重建像素输入特征估计网络以得到第三特征图;将所述重建边信息输入边信息处理网络以得到第一特征图;将所述第一特征图和所述第三特征图输入概率估计网络以得到所述多个概率分布。

[0083] 在一种可能的实现方式中,所述解码模块,还用于将所述残差特征图输入边信息提取网络以得到所述残差特征图的边信息;将所述边信息作为所述残差特征图的重建边信息。

[0084] 在一种可能的实现方式中,所述解码模块,还用于获取所述目标区域的边信息码

流;解析所述边信息码流以得到所述重建边信息。

[0085] 在一种可能的实现方式中,所述解码网络包括第一解码网络;所述解码模块,具体用于将所述残差特征图输入所述第一解码网络以得到所述目标区域的增强层的重建残差像素;当所述修正信息为多个像素值时,对所述重建残差像素和所述修正信息中对应的像素值求和以得到所述重建像素。

[0086] 在一种可能的实现方式中,所述解码网络包括第二解码网络;所述解码模块,具体用于将所述残差特征图输入所述第二解码网络;当所述修正信息为多个特征值时,对所述第二解码网络中的任意一个卷积层的输出和所述修正信息中对应的特征值求和;将求和后的结果输入所述第二解码网络中的任意一个卷积层之后的网络层以得到所述重建像素。

[0087] 在一种可能的实现方式中,所述解码模块,还用于获取所述目标区域所属的图像的基本层码流;解析所述基本层码流以得到所述图像的基本层的重建图;根据所述重建图确定需要增强的至少一个区域,则所述目标区域为所述至少一个区域的其中之一。

[0088] 在一种可能的实现方式中,所述解码模块,具体用于对所述重建图进行划分得到多个区域;将所述多个区域中方差大于第一阈值的区域确定为所述至少一个区域;或者,确定所述多个区域中梯度大于第二阈值的像素占总像素的比例,将所述比例大于第三阈值的区域确定为所述至少一个区域。

[0089] 第五方面,本申请提供一种编码器,包括:一个或多个处理器;存储器,用于存储一个或多个程序;当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现如上述第一方面中任一项所述的方法。

[0090] 第六方面,本申请提供一种解码器,包括:一个或多个处理器;存储器,用于存储一个或多个程序;当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现如上述第二方面中任一项所述的方法。

[0091] 第七方面,本申请提供一种计算机可读存储介质,包括计算机程序,所述计算机程序在计算机上被执行时,使得所述计算机执行上述第一至二方面中任一项所述的方法。

[0092] 第八方面,本申请提供一种计算机程序产品,所述计算机程序产品中包含指令,当所述指令在计算机或处理器上运行时,使得所述计算机或所述处理器实现上述第一至二方面中任一项所述的方法。

附图说明

[0093] 图1为本申请可分级视频译码的一个示例性的层级示意图;

[0094] 图2A为示例性的译码系统10的示意性框图;

[0095] 图2B是视频译码系统40的实例的说明图;

[0096] 图3为本发明实施例提供的视频译码设备400的示意图;

[0097] 图4为示例性实施例提供的装置500的简化框图;

[0098] 图5为本申请实施例的应用场景的示例图;

[0099] 图6为本申请实施例的应用场景的示例图;

[0100] 图7为本申请实施例的应用场景的示例图;

[0101] 图8为本申请实施例的区域增强层的编码方法的过程800的流程图;

[0102] 图9a为修正网络的一个示例性的示意图;

- [0103] 图9b为修正网络的一个示例性的示意图；
- [0104] 图9c为修正网络的一个示例性的示意图；
- [0105] 图9d为修正网络的一个示例性的示意图；
- [0106] 图10a为编码网络的一个示例性的示意图；
- [0107] 图10b为编码网络的一个示例性的示意图；
- [0108] 图11a为概率估计网络的一个示例性的示意图；
- [0109] 图11b为概率估计网络的一个示例性的示意图；
- [0110] 图12为本申请实施例的区域增强层的解码方法的过程1200的流程图；
- [0111] 图13a为解码网络的一个示例性的示意图；
- [0112] 图13b为解码网络的一个示例性的示意图；
- [0113] 图14为编解码流程的一个示例性的示意图；
- [0114] 图15为编解码流程的一个示例性的示意图；
- [0115] 图16为编解码流程的一个示例性的示意图；
- [0116] 图17为编解码流程的一个示例性的示意图；
- [0117] 图18为本申请实施例编码装置1800的一个示例性的结构示意图；
- [0118] 图19为本申请实施例解码装置1900的一个示例性的结构示意图。

具体实施方式

[0119] 为使本申请的目的、技术方案和优点更加清楚，下面将结合本申请中的附图，对本申请中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

[0120] 本申请的说明书实施例和权利要求书及附图中的术语“第一”、“第二”等仅用于区分描述的目的，而不能理解为指示或暗示相对重要性，也不能理解为指示或暗示顺序。此外，术语“包括”和“具有”以及他们的任何变形，意图在于覆盖不排他的包含，例如，包含了一系列步骤或单元。方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元，而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0121] 应当理解，在本申请中，“至少一个(项)”是指一个或者多个，“多个”是指两个或两个以上。“和/或”，用于描述关联对象的关联关系，表示可以存在三种关系，例如，“A和/或B”可以表示：只存在A，只存在B以及同时存在A和B三种情况，其中A，B可以是单数或者复数。字符“/”一般表示前后关联对象是一种“或”的关系。“以下至少一项(个)”或其类似表达，是指这些项中的任意组合，包括单项(个)或复数项(个)的任意组合。例如，a，b或c中的至少一项(个)，可以表示：a，b，c，“a和b”，“a和c”，“b和c”，或“a和b和c”，其中a，b，c可以是单个，也可以是多个。

[0122] 本申请实施例涉及神经网络的应用，为了便于理解，下面先对相关名词或术语进行解释说明：

[0123] 1、神经网络

[0124] 神经网络(neural network, NN)是机器学习模型，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：

$$[0125] \quad h_{w,b}(x) = f(W^T x) = f(\sum_{s=1}^n W_s x_s + b) \quad (1-1)$$

[0126] 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, 来将神经单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ReLU等非线性函数。神经网络是将许多个上述单一的神经单元联结在一起形成的网络, 即一个神经单元的输出可以是另一个神经单元的输入。每个神经单元的输入可以与前一层的局部感受野 (local receptive field) 相连, 来提取局部感受野的特征, 局部感受野可以由若干个神经单元组成的区域。

[0127] 2、多层感知器 (multi-layer perception, MLP)

[0128] MLP是一种简单的深度神经网络 (deep neural network, DNN) (不同层之间是全连接的), 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。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: $\vec{y} = \alpha(W\vec{x} + \vec{b})$, 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} 经过如此简单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。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。总结就是: 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 。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。训练深度神经网络的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)。

[0129] 3、卷积神经网络

[0130] 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, 是一种深度学习 (deep learning) 架构, 深度学习架构是指通过机器学习的算法, 在不同的抽象层级上进行多个层次的学习。作为一种深度学习架构, CNN是一种前馈 (feed-forward) 人工神经网络, 该前馈人工神经网络中的各个神经元可以对输入其中的图像作出响应。卷积神经网络包含了一个由卷积层和池化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。

[0131] 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。卷积层可以包括很多个卷积算子, 卷积算子也称为核, 其在图像处理中的作用相当于一个从输入图像矩阵中提取特定信息的过滤器, 卷积算子本质上可以是一个权重矩阵, 这个权重矩阵通常被

预先定义,在对图像进行卷积操作的过程中,权重矩阵通常在输入图像上沿着水平方向一个像素接着一个像素(或两个像素接着两个像素……这取决于步长stride的取值)的进行处理,从而完成从图像中提取特定特征的工作。该权重矩阵的大小应该与图像的大小相关,需要注意的是,权重矩阵的纵深维度(depth dimension)和输入图像的纵深维度是相同的,在进行卷积运算的过程中,权重矩阵会延伸到输入图像的整个深度。因此,和一个单一的权重矩阵进行卷积会产生一个单一纵深维度的卷积化输出,但是大多数情况下不使用单一权重矩阵,而是应用多个尺寸(行×列)相同的权重矩阵,即多个同型矩阵。每个权重矩阵的输出被堆叠起来形成卷积图像的纵深维度,这里的维度可以理解为由上面所述的“多个”来决定。不同的权重矩阵可以用来提取图像中不同的特征,例如一个权重矩阵用来提取图像边缘信息,另一个权重矩阵用来提取图像的特定颜色,又一个权重矩阵用来对图像中不需要的噪点进行模糊化等。该多个权重矩阵尺寸(行×列)相同,经过该多个尺寸相同的权重矩阵提取后的特征图的尺寸也相同,再将提取到的多个尺寸相同的特征图合并形成卷积运算的输出。这些权重矩阵中的权重值在实际应用中需要经过大量的训练得到,通过训练得到的权重值形成的各个权重矩阵可以用来从输入图像中提取信息,从而使得卷积神经网络进行正确的预测。当卷积神经网络有多个卷积层的时候,初始的卷积层往往提取较多的一般特征,该一般特征也可以称之为低级别的特征;随着卷积神经网络深度的加深,越往后的卷积层提取到的特征越来越复杂,比如高级别的语义之类的特征,语义越高的特征越适用于待解决的问题。

[0132] 由于常常需要减少训练参数的数量,因此卷积层之后常常需要周期性的引入池化层,可以是一层卷积层后面跟一层池化层,也可以是多层卷积层后面接一层或多层池化层。在图像处理过程中,池化层的唯一目的就是减少图像的空间大小。池化层可以包括平均池化算子和/或最大池化算子,以用于对输入图像进行采样得到较小尺寸的图像。平均池化算子可以在特定范围内对图像中的像素值进行计算产生平均值作为平均池化的结果。最大池化算子可以在特定范围内取该范围内值最大的像素作为最大池化的结果。另外,就像卷积层中用权重矩阵的大小应该与图像尺寸相关一样,池化层中的运算符也应该与图像的大小相关。通过池化层处理后输出的图像尺寸可以小于输入池化层的图像的尺寸,池化层输出的图像中每个像素点表示输入池化层的图像的对应子区域的平均值或最大值。

[0133] 在经过卷积层/池化层的处理后,卷积神经网络还不足以输出所需要的输出信息。因为如前所述,卷积层/池化层只会提取特征,并减少输入图像带来的参数。然而为了生成最终的输出信息(所需要的类信息或其他相关信息),卷积神经网络需要利用神经网络层来生成一个或者一组所需要的类的数量的输出。因此,在神经网络层中可以包括多层隐含层,该多层隐含层中所包含的参数可以根据具体的任务类型的相关训练数据进行预先训练得到,例如该任务类型可以包括图像识别,图像分类,图像超分辨率重建等等。

[0134] 可选的,在神经网络层中的多层隐含层之后,还包括整个卷积神经网络的输出层,该输出层具有类似分类交叉熵的损失函数,具体用于计算预测误差,一旦整个卷积神经网络的前向传播完成,反向传播就会开始更新前面提到的各层的权重值以及偏差,以减少卷积神经网络的损失,及卷积神经网络通过输出层输出的结果和理想结果之间的误差。

[0135] 4、循环神经网络

[0136] 循环神经网络(recurrent neural networks,RNN)是用来处理序列数据的。在传

统的神经网络模型中,是从输入层到隐含层再到输出层,层与层之间是全连接的,而对于每一层层内之间的各个节点是无连接的。这种普通的神经网络虽然解决了很多难题,但是却仍然对很多问题却无能无力。例如,你要预测句子的下一个单词是什么,一般需要用到前面的单词,因为一个句子中前后单词并不是独立的。RNN之所以称为循环神经网络,即一个序列当前的输出与前面的输出也有关。具体的表现形式为网络会对前面的信息进行记忆并应用于当前输出的计算中,即隐含层本层之间的节点不再无连接而是有连接的,并且隐含层的输入不仅包括输入层的输出还包括上一时刻隐含层的输出。理论上,RNN能够对任何长度的序列数据进行处理。对于RNN的训练和对传统的CNN或DNN的训练一样。同样使用误差反向传播算法,不过有一点区别:即,如果将RNN进行网络展开,那么其中的参数,如 W ,是共享的;而如上举例上述的传统神经网络却不是这样。并且在使用梯度下降算法中,每一步的输出不仅依赖当前步的网络,还依赖前面若干步网络的状态。该学习算法称为基于时间的反向传播算法(Back propagation Through Time,BPTT)。

[0137] 既然已经有了卷积神经网络,为什么还要循环神经网络?原因很简单,在卷积神经网络中,有一个前提假设是:元素之间是相互独立的,输入与输出也是独立的,比如猫和狗。但现实世界中,很多元素都是相互连接的,比如股票随时间的变化,再比如一个人说了:我喜欢旅游,其中最喜欢的地方是云南,以后有机会一定要去。这里填空,人类应该都知道是填“云南”。因为人类会根据上下文的内容进行推断,但如何让机器做到这一步?RNN就应运而生了。RNN旨在让机器像人一样拥有记忆的能力。因此,RNN的输出就需要依赖当前的输入信息和历史的记忆信息。

[0138] 5、损失函数

[0139] 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。

[0140] 6、反向传播算法

[0141] 卷积神经网络可以采用误差反向传播(back propagation,BP)算法在训练过程中修正初始的超分辨率模型中参数的大小,使得超分辨率模型的重建误差损失越来越小。具体地,前向传递输入信号直至输出会产生误差损失,通过反向传播误差损失信息来更新初始的超分辨率模型中参数,从而使误差损失收敛。反向传播算法是以误差损失为主导的反向传播运动,旨在得到最优的超分辨率模型的参数,例如权重矩阵。

[0142] 7、生成式对抗网络

[0143] 生成式对抗网络(generative adversarial networks,GAN)是一种深度学习模型。该模型中至少包括两个模块:一个模块是生成模型(Generative Model),另一个模块是

判别模型(Discriminative Model),通过这两个模块互相博弈学习,从而产生更好的输出。生成模型和判别模型都可以是神经网络,具体可以是深度神经网络,或者卷积神经网络。GAN的基本原理如下:以生成图片的GAN为例,假设有两个网络,G(Generator)和D(Discriminator),其中G是一个生成图片的网络,它接收一个随机的噪声 z ,通过这个噪声生成图片,记做 $G(z)$;D是一个判别网络,用于判别一张图片是不是“真实的”。它的输入参数是 x , x 代表一张图片,输出 $D(x)$ 代表 x 为真实图片的概率,如果为1,就代表100%是真实的图片,如果为0,就代表不可能是真实的图片。在对该生成式对抗网络进行训练的过程中,生成网络G的目标就是尽可能生成真实的图片去欺骗判别网络D,而判别网络D的目标就是尽量把G生成的图片和真实的图片区分开来。这样,G和D就构成了一个动态的“博弈”过程,也即“生成式对抗网络”中的“对抗”。最后博弈的结果,在理想的状态下,G可以生成足以“以假乱真”的图片 $G(z)$,而D难以判定G生成的图片究竟是不是真实的,即 $D(G(z))=0.5$ 。这样就得到了一个优异的生成模型G,它可以用来生成图片。

[0144] 本申请实施例涉及可分级视频译码,为了便于理解,下面先对相关名词或术语进行解释说明:

[0145] 可分级视频译码,又称可伸缩视频编解码,是当前视频译码标准的扩展编码标准(一般为高级视频译码(advanced video coding,AVC)(H.264)的扩展标准可伸缩视频译码(scalable video coding,SVC),或高效率视频译码(high efficiency video coding,HEVC)(H.265)的扩展标准可伸缩高效视频译码(scalable high efficiency video coding,SHVC))。可分级视频译码的出现主要是为了解决实时视频传输中出现的由于网络带宽实时变化带来的丢包和时延抖动问题。

[0146] 可分级视频译码中的基本结构可称作层级,可分级视频译码技术通过对原始的图像块进行空域分级(分辨率分级),可以得到不同分辨率的层级的码流。分辨率可以是指图像块的以像素为单位的尺寸大小,低层级的分辨率较低,而高层级的分辨率不低于低层级的分辨率;或者,通过对原始的图像块进行时域分级(帧率分级),可以得到不同帧率的层级的码流。帧率可以是指单位时间内视频包含的图像帧数,低层级的帧率较低,而高层级的帧率不低于低层级的帧率;或者,通过对原始的图像块进行质量域分级,可以得到不同编码质量的层级的码流。编码质量可以是指视频的品质,低层级的图像失真程度较大,而高层级的图像失真程度不高于低层级的图像失真程度。

[0147] 通常,被称作基本层的层级是可分级视频译码中的最底层。在空域分级中,基本层图像块使用最低分辨率进行编码;在时域分级中,基本层图像块使用最低帧率进行编码;在质量域分级中,基本层图像块使用最高QP或是最低码率进行编码。即基本层是可分级视频译码中品质最低的一层。被称作增强层的层级是可分级视频译码中在基本层之上的层级,由低到高可以分为多个增强层。最低层增强层依据基本层获得的编码信息,编码得到的合并码流,其编码分辨率比基本层高,或是帧率比基本层高,或是码率比基本层大。较高层增强层可以依据较低层增强层的编码信息,来编码更高品质的图像块。

[0148] 例如,图1为本申请可分级视频译码的一个示例性的层级示意图,如图1所示,原始图像块送入可分级编码器后,根据不同的编码配置可分层为基本层图像块B和增强层图像块($E_1 \sim E_n, n \geq 1$),再分别编码得到包含基本层码流和增强层码流的码流。基本层码流一般是对图像块采用最低分辨率、最低帧率或者最低编码质量参数得到的码流。增强层码流是

以基本层作为基础,叠加采用高分辨率、高帧率或者高编码质量参数对图像块进行编码得到的码流。随着增强层层数增加,编码的空域层级、时域层级或者质量层级也会越来越高。编码器向解码器传输码流时,优先保证基本层码流的传输,当网络有余量时,逐步传输越来越高层级的码流。解码器先接收基本层码流并解码,然后根据收到的增强层码流,按照从低层级到高层级的顺序,逐层解码空域、时域或者质量的层级越来越高的码流,然后将较高层级的解码信息叠加在较低层级的重建块上,获得较高分辨率、较高帧率或者较高质量的重建块。

[0149] 如上,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(图像块)级处理即编码视频,例如,通过空间(帧内)预测和时间(帧间)预测来产生预测块;从图像块(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,可以减少待传输(压缩)的数据量。编码器还需要经过反量化和反变换获得重建残差块,然后将重建残差块的像素点值和预测块的像素点值相加以获得重建块。基本层的重建块是指针对原始的图像块分层得到的基本层图像块执行上述操作所得到的重建块。

[0150] 下文中的区域可以是指整帧图像中的最大编码单元(largest coding unit, LCU),或者整帧图像划分后得到的图像块,或者整帧图像中的感兴趣区域(region of interest, ROI)(即在图像中进行指定的某个需要处理的图像区域),或者一帧图像中的分片图像(Slice)。

[0151] 随着深度学习在各领域表现优异,研究人员提出基于深度学习的端到端图像编码方案。结合该方案,相关技术提供了一种图像分层编码框架,其中,基本层使用基于多功能视频译码(versatile video coding, VVC)的编解码器,增强层使用基于人工智能(artificial intelligence, AI)的编解码器。该方案可参照如下描述:

[0152] 编码端:

[0153] 基本层, VVC编码器对输入图像 x 进行编码得到码流, VVC解码器对码流进行解码得到基本层的重建图像 x_c 。

[0154] 增强层,将 x 和 x_c 输入编码网络(Encoder),得到增强层的残差特征图(y)。一个分支上,将 y 输入边信息提取网络,得到 y 的边信息(z),对 z 进行量化得到经量化的边信息(\hat{z}),对 \hat{z} 进行熵编码并写入码流;再对前述码流进行熵解码得到经解码的边信息(\tilde{z}),对 \tilde{z} 进行反量化得到重建边信息(\tilde{z})。将 \tilde{z} 输入概率估计网络,得到 \hat{y} 的概率分布。另一个分支上,对 y 进行量化得到经量化的增强层的残差特征图(\hat{y}),根据上述概率分布对 \hat{y} 进行熵编码并写入码流;再根据上述概率分布对前述码流进行熵解码得到经解码的残差特征图 \hat{y} ,对 \hat{y} 进行反量化得到重建残差特征图(\tilde{y})。将 x_c 和 \tilde{y} 输入解码网络(Decoder),得到增强层的重建图 y_c 。

[0155] 解码端:

[0156] 基本层, VVC解码器解析码流得到基本层的重建图像 x_c 。

[0157] 增强层,解析码流得到经解码的边信息(\tilde{z}),对 \tilde{z} 进行反量化得到重建边信息(\tilde{z})。将 \tilde{z} 输入概率估计网络,得到 \hat{y} 的概率分布。根据上述概率分布对前述码流进行熵解码得到经解码的残差特征图 \hat{y} ,对 \hat{y} 进行反量化得到重建残差特征图(\tilde{y})。将 x_c 和 \tilde{y} 输入解码网络

(Decoder),得到增强层的重建图yc。

[0158] 上述方案中,在增强层使用AI图像编解码的方法,在MS-结构相似性(structural similarity,SSIM)、峰值信噪比(peak signal to noise ratio,PSNR)等客观质量度量上优于传统图像编码方案。但是上述方案中采用的AI图像编解码的网络复杂度较高,且有较高的算力需求,导致增强层的编解码效率偏低。

[0159] 为此,本申请实施例提供了一种区域增强层的编解码方法,以提高增强层的编解码效率。以下对本申请实施例的方案所适用的系统和/或场景进行说明。

[0160] 图2A为示例性的译码系统10的示意性框图。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)可用于执行本申请实施例中描述的各种示例的方案。

[0161] 如图2A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。

[0162] 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。

[0163] 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。

[0164] 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。

[0165] 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(或预处理图像数据)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。

[0166] 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3等进一步描述)。

[0167] 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。

[0168] 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。

[0169] 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。

[0170] 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。

[0171] 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。

[0172] 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。

[0173] 通信接口22和通信接口28均可配置为如图2A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。

[0174] 视频解码器(或解码器)30用于接收编码图像数据21并提供解码图像数据(或解码图像数据)31(下面将根据图4等进一步描述)。

[0175] 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。

[0176] 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。

[0177] 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,尤其是编码器20或解码器30中用到的神经网络(下文将详细描述)。

[0178] 本申请实施例中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神经网络。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,例如可以从云端或其他地方获取训练数据进行模型训练。

[0179] 尽管图2A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。

[0180] 根据描述,图2A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,这对技术人员来说是显而易见的。

[0181] 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2B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和解码器30分别可以通过处理电路46实现。所述处理电路46可用于执行下文论述的各种操作。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,从而执行本申请技术。编码器20和解码器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2B所示。

[0182] 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、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),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。

[0183] 在一些情况下,图2A所示的译码系统10仅仅是示例性的,本申请提供的技术可适用于视频译码设备(例如,视频编码或视频解码),这些设备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。

[0184] 图2B是视频译码系统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。

[0185] 如图2B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。

[0186] 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。

[0187] 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。逻辑电路可以用于执行本文所论述的各种操作。

[0188] 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4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。

[0189] 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。

[0190] 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。

[0191] 为便于描述,参考通用视频译码(Versatile video coding,VVC)参考软件或由ITU-T视频译码专家组(Video Coding Experts Group,VCEG)和ISO/IEC运动图像专家组(Motion Picture Experts Group,MPEG)的视频译码联合工作组(Joint Collaboration Team on Video Coding,JCT-VC)开发的高性能视频译码(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。

[0192] 图3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2A中的视频解码器30,也可以是编码器,例如图2A中的视频编码器20。

[0193] 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)组件和电光(electrical-to-optical,E0)组件,用于光信号或电信号的出口或入口。

[0194] 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。

[0195] 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)。

[0196] 图4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2A中的源设备12和目的设备14中的任一个或两个。

[0197] 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。

[0198] 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。

[0199] 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。

[0200] 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。

[0201] 图5为本申请实施例的应用场景的示例图,如图5所示,该应用场景可以是终端、云服务器、视频监控中涉及图像/视频采集、存储或传输的业务,例如,终端拍照/录像、相册、云相册、视频监控等。

[0202] 编码端:摄像头 (Camera) 采集图像/视频。人工智能 (artificial intelligence, AI) 图像/视频译码网络对图像/视频进行从特征提取得到冗余度较低的图像特征,进而基于图像特征进行压缩得到码流/图像文件。

[0203] 解码端:当需要输出图像/视频时,AI图像/视频解码网络对码流/图像文件进行解压缩得到图像特征,再对图像特征进行反特征提取得到重建的图像/视频。

[0204] 存储/传输模块将压缩得到的码流/图像文件针对不同业务进行存储(例如,终端拍照、视频监控、云服务器等)或传输(例如,云服务,直播技术等)。

[0205] 图6为本申请实施例的应用场景的示例图,如图6所示,该应用场景可以是终端、视频监控中涉及图像/视频采集、存储或传输的业务,例如,终端相册、视频监控、直播等。

[0206] 编码端:编码网络将图像/视频变换成冗余度更低的图像特征,其通常包含非线性变换单元,具有非线性特性。熵估计网络负责计算图像特征中各个数据的编码概率。熵编码网络根据各个数据对应的概率,对图像特征进行无损编码得到码流/图像文件,进一步降低图像压缩过程中的数据传输量。

[0207] 解码端:熵解码网络根据各个数据对应的概率,对码流/图像文件进行无损解码得到重建的图像特征。解码网络对熵解码输出的图像特征进行反变换,将其解析为图像/视频。和编码网络对应,其通常包含非线性变换单元,具有非线性特性。保存模块将码流/图像文件保存至终端的对应存储位置。加载模块从终端的对应存储位置加载码流/图像文件,并输入到熵解码网络。

[0208] 图7为本申请实施例的应用场景的示例图,如图7所示,该应用场景可以是云、视频监控中涉及图像/视频采集、存储或传输的业务,例如,云相册、视频监控、直播等。

[0209] 编码端:本地获取图像/视频,对其进行图像(JPEG)编码得到压缩图像/视频,之后向云端发送压缩图像/视频。云端对压缩图像/视频进行JPEG解码得到图像/视频,之后对图像/视频进行压缩得到码流/图像文件并存储。

[0210] 解码端:当本地需要从云端获取图像/视频时,云端对码流/图像文件进行解压缩得到图像/视频,之后对图像/视频进行JPEG编码,得到压缩图像/视频,向本地发送压缩图像/视频。本地对压缩图像/视频进行JPEG解码得到图像/视频。云端的结构以及各个模块的用途可以参考图7的结构以及各个模块的用途,本申请实施例在此不做赘述。

[0211] 基于上述编/解码网络和应用场景,本申请实施例提供了一种图像编/解码方法,以实现高效的非线性变换处理,提升图像/视频压缩算法中的率失真性能。

[0212] 图8为本申请实施例的区域增强层的编码方法的过程800的流程图。过程800可由视频编码器20执行。过程800描述为一系列的步骤或操作,应当理解的是,过程800可以以各种顺序执行和/或同时发生,不限于图8所示的执行顺序。假设具有多个图像帧的视频数据流正在使用视频编码器20,执行包括如下步骤的过程800来对区域增强层进行编码。过程800可以包括:

[0213] 步骤801、获取目标区域的基本层的重建像素。

[0214] 通常,视频序列中的每帧图像可以分割成不重叠的图像块集合,然后在图像块级上进行编码。换句话说,编码器通常在块(图像块)级处理即编码视频,例如,通过空间(帧内)预测和时间(帧间)预测来产生预测块;从图像块(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,可以减少待传输(压缩)的数据量。编码器还需要经过反量化和反变换获得重建残差块,然后将重建残差块的像素点值和预测块的像素点值相加以获得重建块。本申请实施例中,图像帧包含的区域可以是指整帧图像中的最大编码单元(largest coding unit,LCU),或者整帧图像划分后得到的图像块,或者整帧图像中的感兴趣区域(region of interest,ROI)(即在图像中进行指定的某个需要处理的图像区域),等等。应理解,除前述几种情况外,区域还可以是以其他方式描述的局部图像,对此不做具体限定。

[0215] 基于此,目标区域意在表示本申请实施例的方案在一次编码过程中所关注并处理的图像块的位置,而目标区域的形状可以是规则的矩形或正方形,也可以是不规则图形,对此不做具体限定。

[0216] 如上所述,在编码过程中,初始得到的图像块可以称作原始块,其所包含的像素可以称作原始像素;重建的到的图像块可以称作重建块,其所包含的像素可以称作重建像素。而在可分级视频译码中,虽然不同层的分辨率、帧率或质量不同,但编码过程大致相似,尤其是各层均包括初始图像块和重建图像块。

[0217] 因此在本申请实施例中,针对目标区域,可以有以下概念:

[0218] 在基本层,初始得到的区域所包含的像素称作该区域的基本层的原始像素,重建得到的区域所包含的像素称作该区域的基本层的重建像素。

[0219] 在增强层,初始得到的区域所包含的像素称作该区域的增强层的原始像素,重建得到的区域所包含的像素称作该区域的增强层的重建像素。

[0220] 需要说明的是,上述在描述原始像素或重建像素时均未提到像素数量的概念,但应理解,区域中所包含的像素通常是多个像素,本申请实施例对此不做具体限定。

[0221] 本申请实施例中,获取目标区域的基本层的重建像素可以包括:对目标区域所属的图像进行编码以得到图像的基本层码流,再对基本层码流进行解码以得到图像的基本层的重建图,然后根据重建图确定需要增强的至少一个区域,而目标区域为至少一个区域的其中之一。

[0222] 针对基本层,编码端对原始图进行编码得到基本层码流,再对基本层码流进行解码得到基本层的重建图。例如,VVC编码器对原始图进行编码得到基本层码流,VVC解码器对基本层码流进行解码得到基本层的重建图。应理解,针对基本层还可以采用其它编码器,HEVC编解码器、AVC编解码器,本申请实施例对此不做具体限定。

[0223] 在得到基本层的重建图后,可以对该重建图进行划分得到多个区域,该多个区域可以参照上述区域的相关描述,例如,多个LCU、多个图像块、多个ROI等。需要说明的是,区域可以是以多种方式描述的局部图像,相应的,获得多个区域的划分方式也可以采用多种方式,本申请实施例对此不做具体限定。

[0224] 在上述多个区域中,根据图像的属性或特征,可能不是所有的区域都需要做增强处理,某一个或多个区域只需要做基本层的编解码即可满足画质需求,因此本申请实施例中的目标区域是需要做增强处理的区域,而在后续的增强层编码中可以只对该目标区域进行增强层的编解码,这样不需要对整帧图像进行增强处理,可以提高图像的编解码效率。

[0225] 在多个区域中确定可以作为目标区域的可以包括至少一个区域,该至少一个区域满足条件:区域的方差大于第一阈值。例如,区域的方差大于阈值 t_1 , $t_1 > 0$,可以认为该区域的纹理比较复杂,因此需要做增强处理,以提高图像质量;或者,区域中梯度大于第二阈值的像素占总像素的比例大于第三阈值。例如,区域中梯度大于阈值 a 的像素占总像素的比例大于阈值 t_2 , $a > 0$, $0 < t_2 < 1$,也可以认为该区域的纹理比较复杂,因此需要做增强处理,以提高图像质量。

[0226] 经过上述过程,编码端可以将上述至少一个区域中的任意一个区域作为当前的目标区域,从基本层的重建图中提取与目标区域对应的位置的像素即为目标区域的基本层的重建像素。

[0227] 步骤802、将重建像素输入修正网络以得到目标区域的修正信息。

[0228] 修正网络的输入为目标区域的基本层的重建像素,输出为目标区域对应的修正信息,其作用是去除对AI增强层编码无益的噪声信号。修正网络可以由卷积层(conv)和激活层(ReLU)构成,至于修正网络是否包括激活层(generalized divisive normalization, GDN)不做限定,是否有其它激活函数也不做限定,而本申请实施例对卷积层的卷积层数不限定,对卷积核的大小也不做限定,例如,卷积核可以采用 3×3 、 5×5 或 7×7 。

[0229] 在一种可能的实现方式中,将目标区域的基本层的重建像素输入修正网络以得到目标区域的多个像素值和多个特征值的至少之一,修正信息可以是多个像素值或者多个特征值。

[0230] 本申请实施例可以采用神经网络来实现修正网络。示例性的,使用四个卷积层/反卷积层与三个激活层穿插级联构成的神经网络来构建修正网络。其中,每层卷积层的卷积核尺寸可设置为 3×3 ,输出特征图的通道数设置为 M ,且每层卷积层的宽高均进行2倍下采

样。应理解,前述示例并不构成具体限定。其中,卷积核大小、特征图通道数、下采样倍数、下采样次数、卷积层数、激活层数均可调整,本申请实施例对此不作具体限定。

[0231] 经过上述修正网络的处理,可以输出多个像素值和/或多个特征值,应理解,修正网络的输入是重建像素,该重建像素是目标区域范围内的,因此输出的多个像素值和/或多个特征值,即使与目标区域内的像素不是一一对应,也可以认为该多个像素值和/或多个特征值仍属于目标区域范围内,亦称作多个像素值和/或多个特征值和目标区域对应。例如,修正信息可以是目标区域的基本层的重建像素的上采样值。目标区域的基本层的重建像素的分辨率与目标区域的增强层的重建像素的分辨率相同或不同,本申请实施例对此不做具体限定。

[0232] 图9a为修正网络的一个示例性的示意图,如图9a所示,该修正网络包括6个卷积层,5个激活层ReLU,卷积核的尺寸为 3×3 。将目标区域的基本层的重建像素输入该修正网络,输出多个像素值。该修正网络可以对重建像素进行去噪去除对AI增强层编码无益的噪声信号,得到多个像素值。

[0233] 图9b为修正网络的一个示例性的示意图,如图9b所示,该修正网络包括6个卷积层,5个激活层ReLU,卷积核的尺寸为 3×3 。将目标区域的基本层的重建像素输入该修正网络,输出多个特征值。

[0234] 图9c为修正网络的一个示例性的示意图,如图9c所示,该修正网络包括8个卷积层,6个激活层ReLU,卷积核的尺寸为 3×3 。将目标区域的基本层的重建像素输入该修正网络,从修正网络的不同层输出多个像素值和多个特征值。

[0235] 图9d为修正网络的一个示例性的示意图,如图9d所示,该修正网络包括5个卷积层和一个反卷积层,5个激活层ReLU,卷积核的尺寸为 3×3 。修正网络的反卷积层的卷积核大小可设置为 3×3 ,输出特征图的通道数设置为48(还可以是其他值,在此不做限定),且反卷积层的宽高进行2倍上采样,最后一层卷积的输出通道为M,当M为3时,输出多个像素值;当M为48时,输出多个特征值。此时,修正信息的分辨率为基本层重建像素的分辨率的两倍。

[0236] 需要说明的是,除了上述四种修正网络的示例外,本申请实施例还可以用其他结构的修正网络,对此不做具体限定。

[0237] 步骤803、将修正信息和目标区域的原始像素输入编码网络以得到目标区域的增强层的残差特征图。

[0238] 编码端可以从原始图像中提取与目标区域对应的位置的像素即为目标区域的原始像素。

[0239] 如上所述,修正信息可以有两种情况,一种是多个像素值,另一种是多个特征值。相应的,编码网络也可以采用两种结构。本申请实施例中,编码端的编码网络(Encoder),其输入为修正信息和目标区域的原始像素,输出为目标区域的增强层的残差特征图。

[0240] 图10a为编码网络的一个示例性的示意图,如图10a所示,该编码网络可以是第一编码网络,包括4个卷积层和3个GDN。先对原始像素和修正信息(多个像素值)中对应的像素值求差,再将求差后的结果输入第一编码网络以得到残差特征图。本实施例中,由于修正信息(多个像素值)对应于像素域,因此可以与目标区域的原始像素直接求差。

[0241] 图10b为编码网络的一个示例性的示意图,如图10b所示,该编码网络可以是第二编码网络,包括4个卷积层和3个GDN。先将原始像素输入第二编码网络,对第二编码网络中

的任意一个卷积层(例如第2个卷积层)的输出和修正信息(多个特征值)中对应的特征值求差,再将求差后的结果输入第二编码网络中的任意一个卷积层(例如第2个卷积层)之后的网络层以得到残差特征图。本实施例中,由于修正信息(多个特征值)对应于特征域,因此要先将原始像素输入第二编码网络转换到特征域,然后再与多个特征值求差。

[0242] 需要说明的是,除了上述两种编码网络的示例外,本申请实施例还可以用其他结构的编码网络,对此不做具体限定。

[0243] 步骤804、对残差特征图进行编码以得到目标区域的增强层码流。

[0244] 本申请实施例中,可以获取多个概率分布,该多个概率分布与残差特征图包含的多个特征值对应,再根据多个概率分布分别对残差特征图中的对应特征值进行熵编码以得到增强层码流。

[0245] 目标区域的增强层的残差特征图包含多个特征值,为了对该多个特征值进行熵编码,需要得到该多个特征值各自的概率分布。以下是多个概率分布的几种获取方法:

[0246] 1、根据修正信息获取多个概率分布。

[0247] 1.1、将修正信息输入概率估计网络以得到多个概率分布。

[0248] 概率估计网络也可以包括卷积层和GDN,至于概率估计网络是否其它激活函数不做限定,而本申请实施例对卷积层的卷积层数不限定,对卷积核的大小也不做限定。本申请实施例中,首先使用概率分布模型进行建模,然后将修正信息输入概率估计网络得到模型参数,将模型参数代入概率分布模型中,得到概率分布。其中,概率分布模型可以是单高斯模型(gaussian single model,GSM)、非对称高斯模型、混合高斯模型(gaussian mixture model,GMM)或者拉普拉斯分布模型(laplace distribution)。当概率分布为高斯分布时,模型参数为高斯分布的均值参数 μ 和方差 σ 的值。当概率分布为拉普拉斯分布时,模型参数为拉普拉斯分布的位置参数 μ 、尺度参数 b 的值。应理解,除上述概率分布模型之外,还可以采用其它模型,对此不做具体限定。

[0249] 图11a为概率估计网络的一个示例性的示意图,如图11a所示,该概率估计网络包括4个卷积层和3个GDN,其输入修正信息为多个像素值。

[0250] 图11b为概率估计网络的一个示例性的示意图,如图11b所示,该概率估计网络包括2个卷积层和2个GDN,其输入修正信息为多个特征值。

[0251] 需要说明的是,除了上述两种概率估计网络的示例外,本申请实施例还可以用其他结构的概率估计网络,对此不做具体限定。

[0252] 1.2、根据修正信息和残差特征图的重建边信息获取多个概率分布。

[0253] a、当修正信息为多个特征值时,将重建边信息输入边信息处理网络以得到第一特征图,将多个特征值和第一特征图输入概率估计网络以得到多个概率分布。

[0254] 本申请实施例中,可以将目标区域的增强层的残差特征图输入边信息提取网络以得到残差特征图的边信息,对该边信息进行熵编码,并写入码流,与此同时将前述残差特征图的边信息作为残差特征图的重建边信息。边信息可以认为是对目标区域的增强层的残差特征图进一步进行特征提取得到的与前述残差特征图维度相同的特征图,因此边信息提取网络用于对目标区域的增强层的残差特征图进一步进行特征提取得到的与前述残差特征图维度相同的特征图。

[0255] 边信息处理网络可以对边信息进行特征提取以输出与残差特征图分辨率相同的

第一特征图。例如,边信息处理网络中使用三个反卷积层与两个激活层穿插级联构成的神经网络来实现前述功能。

[0256] 概率估计网络可以参照上文描述,此处不再赘述。

[0257] b、当修正信息为多个像素值时,将多个像素值输入特征估计网络(estimate)以得到第二特征图,将重建边信息输入边信息处理网络以得到第一特征图,将第一特征图和第二特征图输入概率估计网络以得到多个概率分布。

[0258] 特征估计网络可以将像素域表示的像素值(多个像素值)转为特征域表示的特征值(第二特征图)。该特征估计网络可以采用图11b所示的概率估计网络的结构,与概率估计网络的区别在输入、输出以及训练过程,而正是由于输入、输出和训练过程的不同,即使网络结构相同(即包含的层结构相同),仍然可以视作是不同的网络,并完成不同的功能。

[0259] 边信息处理网络和概率估计网络可以参照上文描述,此处不再赘述。

[0260] 2、根据残差特征图的重建边信息获取多个概率分布。

[0261] 2.1、将重建边信息输入概率估计网络以得到多个概率分布。

[0262] 概率网络可以参照上文描述,此处不再赘述。

[0263] 2.2、根据重建边信息和重建像素获取多个概率分布。

[0264] 本申请实施例中,可以将残差特征图的重建像素输入特征估计网络以得到第三特征图,将重建边信息输入边信息处理网络以得到第一特征图,将第一特征图和第三特征图输入概率估计网络以得到多个概率分布。

[0265] 特征估计网络、边信息处理网络和概率估计网络可以参照上文描述,此处不再赘述。

[0266] 本申请实施例,对于选定的目标区域的增强层采用AI编码,在此基础上,通过修正网络对基本层的重建像素去除对AI增强层编码无益的噪声信号,以得到修正信息,再根据该修正信息对目标区域的增强层的残差特征图进行编码,一方面只对有需要的区域(目标区域)进行增强层编码,可以降低增强层编码的复杂度,提高增强层编码效率,另一方面基于修正信息进行编码,可以提高编码的精准度。

[0267] 图12为本申请实施例的区域增强层的解码方法的过程1200的流程图。过程1200可由视频解码器30执行。过程1200描述为一系列的步骤或操作,应当理解的是,过程1200可以以各种顺序执行和/或同时发生,不限于图12所示的执行顺序。假设码流正在使用视频解码器30,执行包括如下步骤的过程1200来对码流进行解码以得到区域增强层重建像素。过程1200可以包括:

[0268] 步骤1201、获取目标区域的基本层的重建像素。

[0269] 与编码端相对应,解码端可以接收来自编码端的码流,该码流包含了编码端对图像的原始像素编码后得到的基本层码流,因此解码端通过对基本层码流解码得到图像的基本层的重建图。

[0270] 本申请实施例中,解码端可以采用与编码端相同的方式获取目标区域的基本层的重建像素,可以参照步骤801的描述,此处不再赘述。

[0271] 步骤1202、将重建像素输入修正网络以得到目标区域的修正信息。

[0272] 步骤1202可以参照步骤802的描述,此处不再赘述。

[0273] 步骤1203、获取目标区域的增强层码流。

[0274] 步骤804中,编码端对残差特征图进行编码以得到目标区域的增强层码流,因此,相应的,解码端接收的码流中还包括目标区域的增强层码流。

[0275] 步骤1204、对增强层码流进行解码以得到目标区域的增强层的残差特征图。

[0276] 本申请实施例中,可以获取多个概率分布,该多个概率分布与增强层码流包含的多个特征值码流对应,再根据多个概率分布分别对增强层码流中的对应特征值码流进行熵解码以得到目标区域的残差特征图。

[0277] 增强层码流包含的多个特征值码流,为了对该多个特征值码流进行熵解码,需要得到该多个特征值码流各自的概率分布。多个概率分布的几种获取方法可以参照步骤804的描述,此处不再赘述。

[0278] 步骤1205、将残差特征图和修正信息输入解码网络以得到目标区域的增强层的重建像素。

[0279] 与步骤803相对应,编码端通过编码网络,根据输入的修正信息和目标区域的原始像素得到目标区域的增强层的残差特征图,解码端则是通过解码网络,反向根据输入的残差特征图和修正信息得到目标区域的增强层的重建像素。

[0280] 修正信息可以有两种情况,一种是多个像素值,另一种是多个特征值。相应的,解码网络也可以采用两种结构。本申请实施例中,解码端的解码网络(Decoder),其输入为修正信息和目标区域的增强层的残差特征图,输出为目标区域的增强层的重建像素。

[0281] 图13a为解码网络的一个示例性的示意图,如图13a所示,该解码网络可以是第一解码网络,包括4个卷积层和3个GDN。先将残差特征图输入第一解码网络以得到目标区域的增强层的重建残差像素,再对重建残差像素和修正信息中对应的像素值求和以得到重建像素,修正信息为多个像素值。本实施例中,由于修正信息(多个像素值)对应于像素域,因此可以与目标区域的增强层的重建残差像素直接求和。

[0282] 图13b为解码网络的一个示例性的示意图,如图13b所示,该解码网络可以是第二解码网络,包括4个卷积层和3个GDN。先将残差特征图输入第二解码网络,对第二解码网络中的任意一个卷积层(例如第2个卷积层)的输出和修正信息(多个特征值)中对应的特征值求和,再将求和后的结果输入第二解码网络中的任意一个卷积层(例如第2个卷积层)之后的网络层以得到重建像素。本实施例中,由于修正信息(多个特征值)对应于特征域,因此要先将残差特征图输入第二解码网络转换到特征域,然后再与多个特征值求和。

[0283] 需要说明的是,除了上述两种解码网络的示例外,本申请实施例还可以用其他结构的解码网络,对此不做具体限定。

[0284] 本申请实施例,对于增强层码流采用AI解码,在此基础上,通过修正网络对基本层的重建像素去除对AI增强层编码无益的噪声信号,以得到修正信息,再根据该修正信息对增强层码流进行解码,一方面只对有需要的区域(目标区域)进行增强层解码,可以降低增强层解码的复杂度,提高增强层解码效率,另一方面基于修正信息进行解码,可以提高解码的精准度。

[0285] 以下通过几个具体的实施例,对本申请实施例的方案进行说明。

[0286] 实施例一

[0287] 图14为编解码流程的一个示例性的示意图,如图14所示,本实施例的流程如下文所述。

[0288] 编码端:

[0289] 1、基本层的编码器 (Encoder1) 对原始图像 x 进行编码得到基本层码流 (Bitstream1), 基本层的解码器 (Decoder1) 对 Bitstream1 进行解码得到重建图像 x_c 。

[0290] 其中, Encoder1 和 Decoder1 可以使用传统的视频译码标准进行基本层的编解码, 例如, H.264/AVC, H.265/HEVC 或 H.266/VVC 标准, 还可以使用现有的 JPEG 图像编码标准进行基本层的编解码, 对此不做具体限定。

[0291] 2、在 x_c 中确定需要增强的至少一个区域, 其中之一为目标区域 x_{c1} 。

[0292] 确定方法可以包括选择纹理比较复杂的区域, 选择人眼感兴趣区域, 或者, 随机选取一个或者多个区域。该步骤可以采用相关技术, 对此不做具体限定。

[0293] 例如, 当区域的方差大于阈值 t_1 时, 则该区域为纹理比较复杂的区域, 可作为目标区域; 或者, 当区域内梯度大于阈值 a 的像素比例大于 t_2 时, 则判定该区域为纹理比较复杂的区域, 可作为目标区域。其中, t_1 和 a 为大于 0 的数, t_2 为 0 到 1 的数。

[0294] 3、将 x_{c1} 的基本层的重建像素输入修正网络, 输出修正信息 (p)。

[0295] 其中, 修正网络的作用是去除对 AI 增强层编码无益的噪声信号, 修正网络可以输出多个像素域和/或多个特征域, 修正信息包含前述多个像素域或者多个特征域。示例性地, 可以使用以下三种方法:

[0296] 方法一: 将 x_{c1} 的基本层的重建像素输入如图 9a 所示的修正网络, 输出多个像素值, 作为 p 。

[0297] 方法二: 将 x_{c1} 的基本层的重建像素输入如图 9b 所示的修正网络, 输出多个特征值, 作为 p 。

[0298] 方法三: 将 x_{c1} 的基本层的重建像素输入如图 9c 所示的修正网络, 输出多个像素值和多个特征值, 将其中之一作为 p 。

[0299] 4、将 x 和 p 输入增强层的编码器 (Encoder2), 输出目标区域的增强层的残差特征图 (y), y 经过量化 (Q) 得到量化后的残差特征图 (\hat{y})。

[0300] 修正信息有两种: 像素值或特征值。均可以用于输入到 Encoder2, 只是 Encoder2 的结构上有区别, 如果是像素值, Encoder2 采用图 10a 所示结构; 如果是特征值, Encoder2 采用图 10b 所示结构。

[0301] 需要说明的是, 本申请实施例对具体的量化方法不做限定, 还可以是, 将 y 中的浮点数进行截断得到整数, 或者还可以是根据预设的量化步长进行量化得到量化后的特征值。

[0302] 5、将 p 输入概率估计网络, 输出 \hat{y} 的每个特征值的概率分布, 根据概率分布对 \hat{y} 进行熵编码 (AE), 得到增强层码流 (Bitstream2)。

[0303] 修正信息有两种: 像素值或特征值。均可以输入概率估计网络, 只是概率估计网络的结构上有区别, 如果是像素值, 概率估计网络采用图 11a 所示结构; 如果是特征值, 概率估计网络采用图 11b 所示结构。

[0304] 解码端:

[0305] 1、基本层解码器 (Decoder1) 解析基本层码流 (Bitstream1) 得到重建图像 x_c 。

[0306] 2、在 x_c 中确定需要增强的至少一个区域, 其中之一为目标区域 x_{c1} 。

[0307] 3、将 x_{c1} 的基本层的重建像素输入修正网络, 输出修正信息 (p)。

[0308] 4、将 p 输入概率估计网络,输出 \hat{y} 的每个特征值的概率分布,根据概率分布对增强层码流(Bitstream2)进行熵解码(AD)得到 x_{c1} 的增强层的残差特征图 \hat{y} , \hat{y} 经过反量化(IQ)得到反量化的残差特征图(\tilde{y})。

[0309] 5、将 p 和 \tilde{y} 输入到增强层解码器(Decoder2),输出 x_{c1} 的增强层的重建像素(x_d)。

[0310] 6、将 x_c 中对应于 x_{c1} 位置的像素替换为 x_d 的像素,得到增强后的重建图像。

[0311] 实施例二

[0312] 图15为编解码流程的一个示例性的示意图,如图15所示,本实施例与实施例一的区别在于:概率分布的获取方法。

[0313] 编码端:

[0314] 1、基本层的编码器(Encoder1)对原始图像 x 进行编码得到基本层码流(Bitstream1),基本层的解码器(Decoder1)对Bitstream1进行解码得到重建图像 x_c 。

[0315] 其中,Encoder1和Decoder1可以使用传统的视频译码标准进行基本层的编解码,例如,H.264/AVC,H.265/HEVC或H.266/VVC标准,还可以使用现有的JPEG图像编码标准进行基本层的编解码,对此不做具体限定。

[0316] 2、在 x_c 中确定需要增强的至少一个区域,其中之一为目标区域 x_{c1} 。

[0317] 确定方法可以包括选择纹理比较复杂的区域,选择人眼感兴趣区域,或者,随机选取一个或者多个区域。该步骤可以采用相关技术,对此不做具体限定。

[0318] 例如,当区域的方差大于阈值 t_1 时,则该区域为纹理比较复杂的区域,可作为目标区域;或者,当区域内梯度大于阈值 a 的像素比例大于 t_2 时,则判定该区域为纹理比较复杂的区域,可作为目标区域。其中, t_1 和 a 为大于0的数, t_2 为0到1的数。

[0319] 3、将 x_{c1} 的基本层重建像素输入修正网络,输出修正信息(p)。

[0320] 其中,修正网络的作用是去除对AI增强层编码无益的噪声信号,修正网络可以输出多个像素域和/或多个特征域,修正信息为多个像素域或多个特征域。示例性地,可以使用以下三种方法:

[0321] 方法一:将 x_{c1} 的基本层重建像素输入如图9a所示的修正网络,输出多个像素值,作为 p 。

[0322] 方法二:将 x_{c1} 的基本层重建像素输入如图9b所示的修正网络,输出多个特征值,作为 p 。

[0323] 方法三:将 x_{c1} 的基本层重建像素输入如图9c所示的修正网络,输出多个像素值和多个特征值,将其中之一作为 p 。

[0324] 4、将 x 和 p 输入增强层的编码器(Encoder2),输出目标区域的增强层的残差特征图(y), y 经过量化(Q)得到量化后的残差特征图(\hat{y})。

[0325] 修正信息有两种:像素值或特征值。均可以用于输入到Encoder2,只是Encoder2的结构上有区别,如果是像素值,Encoder2采用图10a所示结构;如果是特征值,Encoder2采用图10b所示结构。

[0326] 需要说明的是,本申请实施例对具体的量化方法不做限定,还可以是,将 y 中的浮点数进行截断得到整数,或者还可以是根据预设的量化步长进行量化得到量化后的特征值。

[0327] 5、将 y 输入边信息提取网络得到 y 的边信息 z ,对 z 量化(Q)得到量化后的边信息(\hat{z}),对 \hat{z} 进行熵编码(AE)得到码流(Bitstream3),又对Bitstream3进行熵解码(AD)得到重建边信息(\hat{z})。此外,编码端可以将Bitstream3传输给解码端。

[0328] 6、将 \hat{z} 输入概率估计网络,输出 \hat{y} 的每个特征值的概率分布,根据概率分布对 \hat{y} 进行熵编码(AE),得到增强层码流(Bitstream2)。

[0329] 解码端:

[0330] 1、基本层解码器(Decoder1)解析基本层码流(Bitstream1)得到重建图像 x_c 。

[0331] 2、在 x_c 中确定需要增强的至少一个区域,其中之一为目标区域 x_{c1} 。

[0332] 3、将 x_{c1} 的基本层重建像素输入修正网络,输出修正信息(p)。

[0333] 4、解析码流(Bitstream3)得到重建边信息(\hat{z})。

[0334] 5、将 \hat{z} 输入概率估计网络,输出增强层码流(Bitstream2)的每个特征码流的概率分布,根据概率分布对Bitstream2进行熵编码(AE),得到目标区域的增强层的残差特征图(\hat{y})。

[0335] 6、将 p 和 \hat{y} 输入到增强层解码器(Decoder2),输出 x_{c1} 的增强层重建像素(x_d)。

[0336] 7、将 x_c 中对应于 x_{c1} 位置的像素替换为 x_d 的像素,得到增强后的重建图像。

[0337] 实施例三

[0338] 图16为编解码流程的一个示例性的示意图,如图16所示,本实施例与实施例二的区别在于:概率分布的获取方法。

[0339] 编码端:

[0340] 1、基本层的编码器(Encoder1)对原始图像 x 进行编码得到基本层码流(Bitstream1),基本层的解码器(Decoder1)对Bitstream1进行解码得到重建图像 x_c 。

[0341] 其中,Encoder1和Decoder1可以使用传统的视频译码标准进行基本层的编解码,例如,H.264/AVC,H.265/HEVC或H.266/VVC标准,还可以使用现有的JPEG图像编码标准进行基本层的编解码,对此不做具体限定。

[0342] 2、在 x_c 中确定需要增强的至少一个区域,其中之一为目标区域 x_{c1} 。

[0343] 确定方法可以包括选择纹理比较复杂的区域,选择人眼感兴趣区域,或者,随机选取一个或者多个区域。该步骤可以采用相关技术,对此不做具体限定。

[0344] 例如,当区域的方差大于阈值 t_1 时,则该区域为纹理比较复杂的区域,可作为目标区域;或者,当区域内梯度大于阈值 a 的像素比例大于 t_2 时,则判定该区域为纹理比较复杂的区域,可作为目标区域。其中, t_1 和 a 为大于0的数, t_2 为0到1的数。

[0345] 3、将 x_{c1} 的基本层重建像素输入修正网络,输出修正信息(p)。

[0346] 其中,修正网络的作用是去除对AI增强层编码无益的噪声信号,修正网络可以输出多个像素域和/或多个特征域,修正信息为多个像素域或多个特征域。示例性地,可以使用以下三种方法:

[0347] 方法一:将 x_{c1} 的基本层重建像素输入如图9a所示的修正网络,输出多个像素值,作为 p 。

[0348] 方法二:将 x_{c1} 的基本层重建像素输入如图9b所示的修正网络,输出多个特征值,作为 p 。

[0349] 方法三:将 x_{c1} 的基本层重建像素输入如图9c所示的修正网络,输出多个像素值

和多个特征值,将其中之一作为 p 。

[0350] 4、将 x 和 p 输入增强层的编码器(Encoder2),输出目标区域的增强层的残差特征图(y), y 经过量化(Q)得到量化后的残差特征图(\hat{y})。

[0351] 修正信息有两种:像素值或特征值。均可以用于输入到Encoder2,只是Encoder2的结构上有区别,如果是像素值,Encoder2采用图10a所示结构;如果是特征值,Encoder2采用图10b所示结构。

[0352] 需要说明的是,本申请实施例对具体的量化方法不做限定,还可以是,将 y 中的浮点数进行截断得到整数,或者还可以是根据预设的量化步长进行量化得到量化后的特征值。

[0353] 5、将 y 输入边信息提取网络得到 y 的边信息 z ,对 z 量化(Q)得到量化后的边信息(\hat{z}),对 \hat{z} 进行熵编码(AE)得到码流(Bitstream3),又对Bitstream3进行熵解码(AD)得到重建边信息(\hat{z})。此外,编码端可以将Bitstream3传输给解码端。

[0354] 6、将 \hat{z} 输入边信息处理网络得到特征域的重建边信息。

[0355] 7、如果 p 是多个像素值,则需要将多个像素值输入特征估计网络进行特征提取,如果 p 是多个特征值,则不需要输入特征估计网络。

[0356] 8、将特征域的重建边信息和特征域的 p 输入概率估计网络,输出 \hat{y} 的每个特征值的概率分布,根据概率分布对 \hat{y} 进行熵编码(AE),得到增强层码流(Bitstream2)。

[0357] 解码端:

[0358] 1、基本层解码器(Decoder1)解析基本层码流(Bitstream1)得到重建图像 x_c 。

[0359] 2、在 x_c 中确定需要增强的至少一个区域,其中之一为目标区域 x_{c1} 。

[0360] 3、将 x_{c1} 的基本层重建像素输入修正网络,输出修正信息(p)。

[0361] 4、解析码流(Bitstream3)得到重建边信息(\hat{z})。

[0362] 5、将 \hat{z} 输入边信息处理网络得到特征域的重建边信息。

[0363] 6、如果 p 是多个像素值,则需要将多个像素值输入特征估计网络进行特征提取,如果 p 是多个特征值,则不需要输入特征估计网络。

[0364] 7、将特征域的重建边信息和特征域的 p 输入概率估计网络,输出增强层码流(Bitstream2)的每个特征码流的概率分布,根据概率分布对Bitstream2进行熵编码(AE),得到目标区域的增强层的残差特征图(\hat{y})。

[0365] 8、将 p 和 \hat{y} 输入到增强层解码器(Decoder2),输出 x_{c1} 的增强层重建像素(x_d)。

[0366] 9、将 x_c 中对应于 x_{c1} 位置的像素替换为 x_d 的像素,得到增强后的重建图像。

[0367] 实施例四

[0368] 图17为编解码流程的一个示例性的示意图,如图17所示,本实施例与实施例三的区别在于:概率分布的获取方法。

[0369] 编码端:

[0370] 1、基本层的编码器(Encoder1)对原始图像 x 进行编码得到基本层码流(Bitstream1),基本层的解码器(Decoder1)对Bitstream1进行解码得到重建图像 x_c 。

[0371] 其中,Encoder1和Decoder1可以使用传统的视频译码标准进行基本层的编解码,例如,H.264/AVC,H.265/HEVC或H.266/VVC标准,还可以使用现有的JPEG图像编码标准进行基本层的编解码,对此不做具体限定。

[0372] 2、在xc中确定需要增强的至少一个区域,其中之一为目标区域xc1。

[0373] 确定方法可以包括选择纹理比较复杂的区域,选择人眼感兴趣区域,或者,随机选取一个或者多个区域。该步骤可以采用相关技术,对此不做具体限定。

[0374] 例如,当区域的方差大于阈值t1时,则该区域为纹理比较复杂的区域,可作为目标区域;或者,当区域内梯度大于阈值a的像素比例大于t2时,则判定该区域为纹理比较复杂的区域,可作为目标区域。其中,t1和a为大于0的数,t2为0到1的数。

[0375] 3、将xc1的基本层的重建像素输入修正网络,输出修正信息(p)。

[0376] 其中,修正网络的作用是去除对AI增强层编码无益的噪声信号,修正网络可以输出多个像素域和/或多个特征域,修正信息为多个像素域或多个特征域。示例性地,可以使用以下三种方法:

[0377] 方法一:将xc1的基本层的重建像素输入如图9a所示的修正网络,输出多个像素值,作为p。

[0378] 方法二:将xc1的基本层的重建像素输入如图9b所示的修正网络,输出多个特征值,作为p。

[0379] 方法三:将xc1的基本层的重建像素输入如图9c所示的修正网络,输出多个像素值和多个特征值,将其中之一作为p。

[0380] 4、将x和p输入增强层的编码器(Encoder2),输出目标区域的增强层的残差特征图(y),y经过量化(Q)得到量化后的残差特征图(\hat{y})。

[0381] 修正信息有两种:像素值或特征值。均可以用于输入到Encoder2,只是Encoder2的结构上有区别,如果是像素值,Encoder2采用图10a所示结构;如果是特征值,Encoder2采用图10b所示结构。

[0382] 需要说明的是,本申请实施例对具体的量化方法不做限定,还可以是,将y中的浮点数进行截断得到整数,或者还可以是根据预设的量化步长进行量化得到量化后的特征值。

[0383] 5、将y输入边信息提取网络得到y的边信息z,对z量化(Q)得到量化后的边信息(\hat{z}),对 \hat{z} 进行熵编码(AE)得到码流(Bitstream3),又对Bitstream3进行熵解码(AD)得到重建边信息(\hat{z})。此外,编码端可以将Bitstream3传输给解码端。

[0384] 6、将 \hat{z} 输入边信息处理网络得到特征域的重建边信息。

[0385] 7、将xc1输入特征估计网络进行特征提取。

[0386] 8、将同属于特征域的重建边信息和xc1输入概率估计网络,输出 \hat{y} 的每个特征值的概率分布,根据概率分布对 \hat{y} 进行熵编码(AE),得到增强层码流(Bitstream2)。

[0387] 解码端:

[0388] 1、基本层解码器(Decoder1)解析基本层码流(Bitstream1)得到重建图像xc。

[0389] 2、在xc中确定需要增强的至少一个区域,其中之一为目标区域xc1。

[0390] 3、将xc1的基本层的重建像素输入修正网络,输出修正信息(p)。

[0391] 4、解析码流(Bitstream3)得到重建边信息(\hat{z})。

[0392] 5、将 \hat{z} 输入边信息处理网络得到特征域的重建边信息。

[0393] 6、将xc1输入特征估计网络进行特征提取。

[0394] 7、将同属于特征域的重建边信息和xc1输入概率估计网络,输出增强层码流(Bitstream2)的每个特征码流的概率分布,根据概率分布对Bitstream2进行熵编码(AE),得到目标区域的增强层的残差特征图(\hat{y})。

[0395] 8、将p和 \hat{y} 输入到增强层解码器(Decoder2),输出xc1的增强层的重建像素(xd)。

[0396] 9、将xc中对应于xc1位置的像素替换为xd的像素,得到增强后的重建图像。

[0397] 图18为本申请实施例编码装置1800的一个示例性的结构示意图,如图18所示,本实施例的编码装置1800可以应用于编码端20。该编码装置1800可以包括:获取模块1801、处理模块1802和编码模块1803。其中,

[0398] 获取模块1801,用于获取目标区域的基本层的重建像素;处理模块1802,用于将所述重建像素输入修正网络以得到所述目标区域的修正信息;将所述修正信息和所述目标区域的原始像素输入编码网络以得到所述目标区域的增强层的残差特征图;编码模块1803,用于对所述残差特征图进行编码以得到所述目标区域的增强层码流。

[0399] 在一种可能的实现方式中,所述处理模块1802,具体用于将所述重建像素输入所述修正网络以得到所述目标区域的多个像素值和多个特征值的至少之一,所述修正信息为所述多个像素值或者所述多个特征值。

[0400] 在一种可能的实现方式中,所述编码模块1803,具体用于根据所述修正信息获取多个概率分布,所述多个概率分布与所述残差特征图包含的多个特征值对应;根据所述多个概率分布分别对所述残差特征图中的对应特征值进行熵编码以得到所述增强层码流。

[0401] 在一种可能的实现方式中,所述编码模块1803,具体用于将所述修正信息输入概率估计网络以得到所述多个概率分布。

[0402] 在一种可能的实现方式中,所述编码模块1803,具体用于根据所述修正信息和所述残差特征图的重建边信息获取所述多个概率分布。

[0403] 在一种可能的实现方式中,所述编码模块1803,具体用于当所述修正信息为多个特征值时,将所述重建边信息输入边信息处理网络以得到第一特征图;将所述多个特征值和所述第一特征图输入概率估计网络以得到所述多个概率分布。

[0404] 在一种可能的实现方式中,所述编码模块1803,具体用于当所述修正信息为多个像素值时,将所述多个像素值输入特征估计网络以得到第二特征图;将所述重建边信息输入边信息处理网络以得到第一特征图;将所述第一特征图和所述第二特征图输入概率估计网络以得到所述多个概率分布。

[0405] 在一种可能的实现方式中,所述编码模块1803,具体用于根据所述残差特征图的重建边信息获取多个概率分布,所述多个概率分布与所述残差特征图包含的多个特征值对应;根据所述多个概率分布分别对所述残差特征图中的对应特征值进行熵编码以得到所述增强层码流。

[0406] 在一种可能的实现方式中,所述编码模块1803,具体用于将所述重建边信息输入概率估计网络以得到所述多个概率分布。

[0407] 在一种可能的实现方式中,所述编码模块1803,具体用于根据所述重建边信息和所述重建像素获取所述多个概率分布。

[0408] 在一种可能的实现方式中,所述编码模块1803,具体用于将所述重建像素输入特征估计网络以得到第三特征图;将所述重建边信息输入边信息处理网络以得到第一特征

图;将所述第一特征图和所述第三特征图输入概率估计网络以得到所述多个概率分布。

[0409] 在一种可能的实现方式中,所述编码模块1803,还用于将所述残差特征图输入边信息提取网络以得到所述残差特征图的边信息;对所述边信息或经量化的所述边信息进行熵编码,并写入码流。

[0410] 在一种可能的实现方式中,所述编码网络包括第一编码网络;所述编码模块1803,具体用于当所述修正信息为多个像素值时,对所述原始像素和所述修正信息中对应的像素值求差;将求差后的结果输入所述第一编码网络以得到所述残差特征图。

[0411] 在一种可能的实现方式中,所述编码网络包括第二编码网络;所述编码模块1803,具体用于将所述原始像素输入所述第二编码网络;当所述修正信息为多个特征值时,对所述第二编码网络中的任意一个卷积层的输出和所述修正信息中对应的特征值求差;将求差后的结果输入所述第二编码网络中的任意一个卷积层之后的网络层以得到所述残差特征图。

[0412] 在一种可能的实现方式中,所述编码模块1803,还用于对所述目标区域所属的图像进行编码以得到所述图像的基本层码流;对所述基本层码流进行解码以得到所述图像的基本层的重建图;根据所述重建图确定需要增强的至少一个区域,则所述目标区域为所述至少一个区域的其中之一。

[0413] 在一种可能的实现方式中,所述编码模块1803,具体用于对所述重建图进行划分得到多个区域;将所述多个区域中方差大于第一阈值的区域确定为所述至少一个区域;或者,确定所述多个区域中梯度大于第二阈值的像素占总像素的比例,将所述比例大于第三阈值的区域确定为所述至少一个区域。

[0414] 在一种可能的实现方式中,所述编码模块1803,还用于将所述残差特征图的边信息作为所述残差特征图的重建边信息。

[0415] 本实施例的装置,可以用于执行图8所示方法实施例的技术方案,其实现原理和技术效果类似,此处不再赘述。

[0416] 图19为本申请实施例解码装置1900的一个示例性的结构示意图,如图19所示,本实施例的解码装置1900可以应用于解码端30。该解码装置1900可以包括:获取模块1901、处理模块1902和解码模块1903。其中,

[0417] 获取模块1901,用于获取目标区域的基本层的重建像素;处理模块1902,用于将所述重建像素输入修正网络以得到所述目标区域的修正信息;所述获取模块1901,还用于获取所述目标区域的增强层码流;解码模块1903,用于对所述增强层码流进行解码以得到所述目标区域的增强层的残差特征图;所述处理模块1902,还用于将所述残差特征图和所述修正信息输入解码网络以得到所述目标区域的增强层的重建像素。

[0418] 在一种可能的实现方式中,所述处理模块1902,具体用于将所述重建像素输入所述修正网络以得到所述目标区域的多个像素值和多个特征值的至少之一,所述修正信息为所述多个像素值或者所述多个特征值。

[0419] 在一种可能的实现方式中,所述解码模块1903,具体用于根据所述修正信息获取所述多个概率分布,所述多个概率分布与所述增强层码流包含的多个特征值码流对应;根据所述多个概率分布分别对所述增强层码流中的对应特征值码流进行熵解码以得到所述残差特征图。

[0420] 在一种可能的实现方式中,所述解码模块1903,具体用于将所述修正信息输入概率估计网络以得到所述多个概率分布。

[0421] 在一种可能的实现方式中,所述解码模块1903,具体用于根据所述修正信息和所述残差特征图的重建边信息获取所述多个概率分布。

[0422] 在一种可能的实现方式中,所述解码模块1903,具体用于当所述修正信息为多个特征值时,将所述重建边信息输入边信息处理网络以得到第一特征图;将所述多个特征值和所述第一特征图输入概率估计网络以得到所述多个概率分布。

[0423] 在一种可能的实现方式中,所述解码模块1903,具体用于当所述修正信息为多个像素值时,将所述多个像素值输入特征估计网络以得到第二特征图;将所述重建边信息输入边信息处理网络以得到第一特征图;将所述第一特征图和所述第二特征图输入概率估计网络以得到所述多个概率分布。

[0424] 在一种可能的实现方式中,所述解码模块1903,具体用于根据所述残差特征图的重建边信息获取所述多个概率分布,所述多个概率分布与所述增强层码流包含的多个特征值码流对应;根据所述多个概率分布分别对所述增强层码流中的对应特征值码流进行熵解码以得到所述残差特征图。

[0425] 在一种可能的实现方式中,所述解码模块1903,具体用于将所述重建边信息输入概率估计网络以得到所述多个概率分布。

[0426] 在一种可能的实现方式中,所述解码模块1903,具体用于根据所述重建边信息和所述重建像素获取所述多个概率分布。

[0427] 在一种可能的实现方式中,所述解码模块1903,具体用于将所述重建像素输入特征估计网络以得到第三特征图;将所述重建边信息输入边信息处理网络以得到第一特征图;将所述第一特征图和所述第三特征图输入概率估计网络以得到所述多个概率分布。

[0428] 在一种可能的实现方式中,所述解码模块1903,还用于将所述残差特征图输入边信息提取网络以得到所述残差特征图的边信息;将所述边信息作为所述残差特征图的重建边信息。

[0429] 在一种可能的实现方式中,所述解码模块1903,还用于获取所述目标区域的边信息码流;解析所述边信息码流以得到所述重建边信息。

[0430] 在一种可能的实现方式中,所述解码网络包括第一解码网络;所述解码模块1903,具体用于将所述残差特征图输入所述第一解码网络以得到所述目标区域的增强层的重建残差像素;当所述修正信息为多个像素值时,对所述重建残差像素和所述修正信息中对应的像素值求和以得到所述重建像素。

[0431] 在一种可能的实现方式中,所述解码网络包括第二解码网络;所述解码模块1903,具体用于将所述残差特征图输入所述第二解码网络;当所述修正信息为多个特征值时,对所述第二解码网络中的任意一个卷积层的输出和所述修正信息中对应的特征值求和;将求和后的结果输入所述第二解码网络中的任意一个卷积层之后的网络层以得到所述重建像素。

[0432] 在一种可能的实现方式中,所述解码模块1903,还用于获取所述目标区域所属的图像的基本层码流;解析所述基本层码流以得到所述图像的基本层的重建图;根据所述重建图确定需要增强的至少一个区域,则所述目标区域为所述至少一个区域的其中之一。

[0433] 在一种可能的实现方式中,所述解码模块1903,具体用于对所述重建图进行划分得到多个区域;将所述多个区域中方差大于第一阈值的区域确定为所述至少一个区域;或者,确定所述多个区域中梯度大于第二阈值的像素占总像素的比例,将所述比例大于第三阈值的区域确定为所述至少一个区域。

[0434] 本实施例的装置,可以用于执行图12所示方法实施例的技术方案,其实现原理和技术效果类似,此处不再赘述。

[0435] 在实现过程中,上述方法实施例的各步骤可以通过处理器中的硬件的集成逻辑电路或者软件形式的指令完成。处理器可以是通用处理器、数字信号处理器(digital signal processor, DSP)、特定应用集成电路(application-specific integrated circuit, ASIC)、现场可编程门阵列(field programmable gate array, FPGA)或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。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。本申请实施例公开的方法的步骤可以直接体现为硬件编码处理器执行完成,或者用编码处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器,闪存、只读存储器,可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器,处理器读取存储器中的信息,结合其硬件完成上述方法的步骤。

[0436] 上述各实施例中提及的存储器可以是易失性存储器或非易失性存储器,或可包括易失性和非易失性存储器两者。其中,非易失性存储器可以是只读存储器(read-only memory, ROM)、可编程只读存储器(programmable ROM, PROM)、可擦除可编程只读存储器(erasable PROM, EPROM)、电可擦除可编程只读存储器(electrically EPROM, EEPROM)或闪存。易失性存储器可以是随机存取存储器(random access memory, RAM),其用作外部高速缓存。通过示例性但不是限制性说明,许多形式的RAM可用,例如静态随机存取存储器(static RAM, SRAM)、动态随机存取存储器(dynamic RAM, DRAM)、同步动态随机存取存储器(synchronous DRAM, SDRAM)、双倍数据速率同步动态随机存取存储器(double data rate SDRAM, DDR SDRAM)、增强型同步动态随机存取存储器(enhanced SDRAM, ESDRAM)、同步连接动态随机存取存储器(synchlink DRAM, SLDRAM)和直接内存总线随机存取存储器(direct rambus RAM, DR RAM)。应注意,本文描述的系统和方法的存储器旨在包括但不限于这些和任意其它适合类型的存储器。

[0437] 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。

[0438] 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。

[0439] 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。

[0440] 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。

[0441] 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。

[0442] 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(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。

[0443] 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。

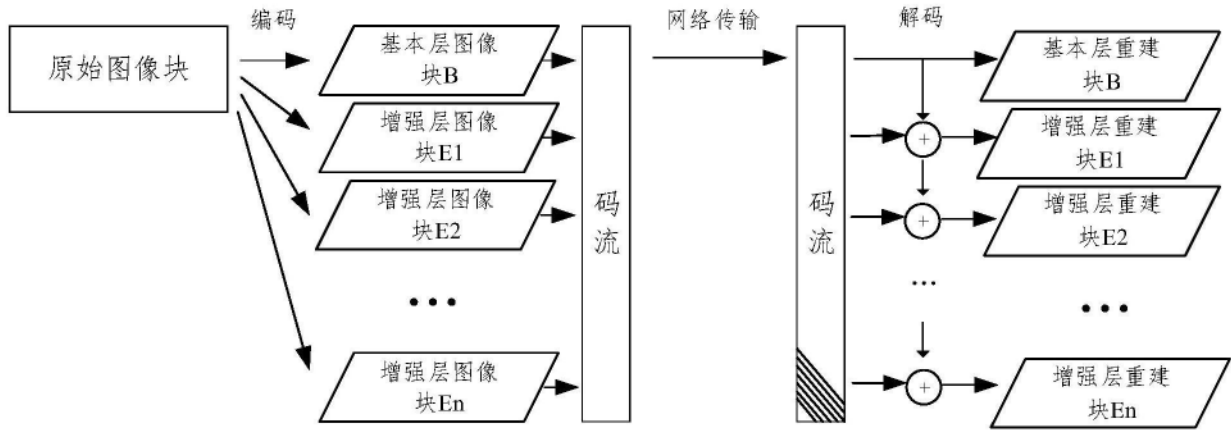


图1

10

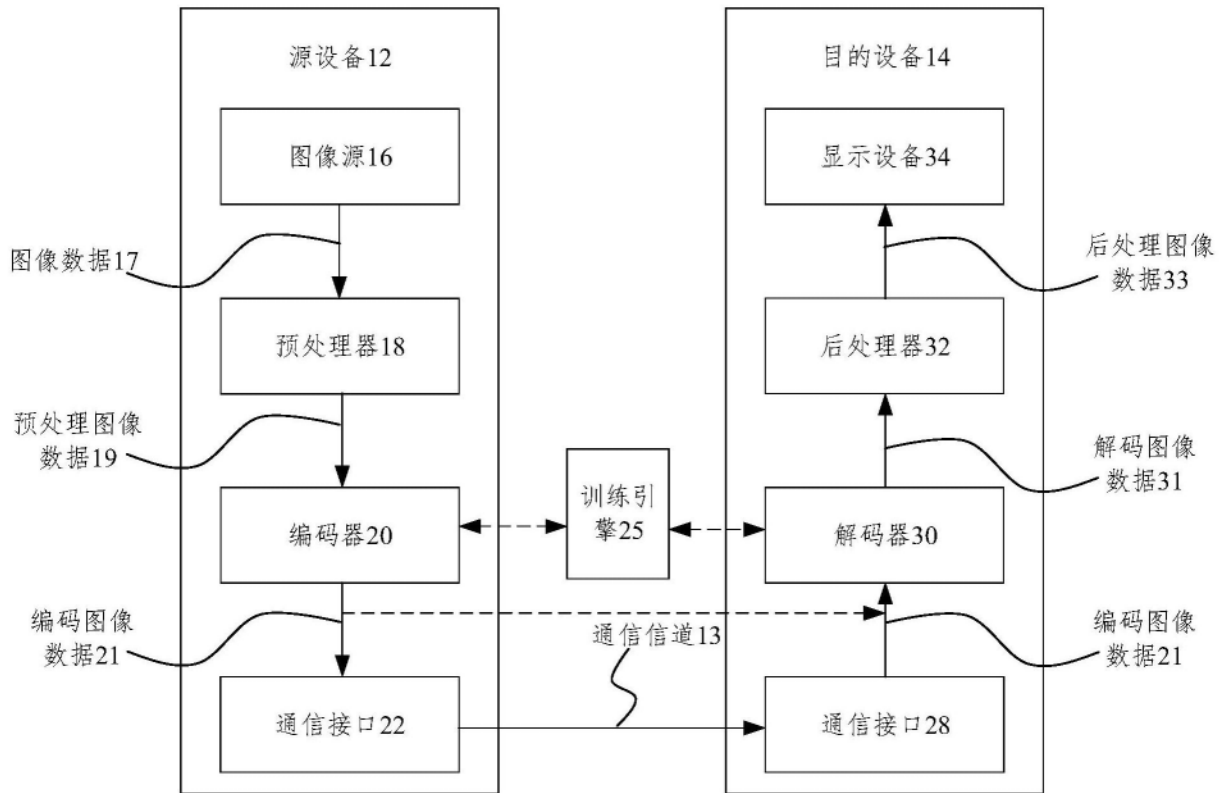


图2A

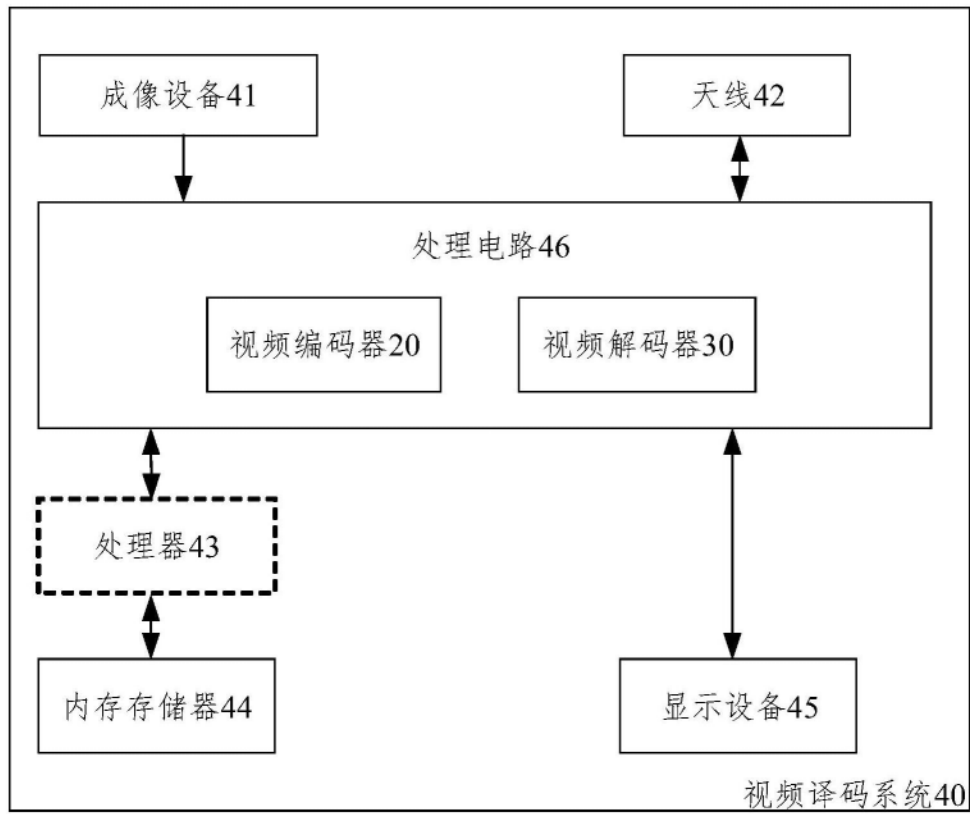


图2B

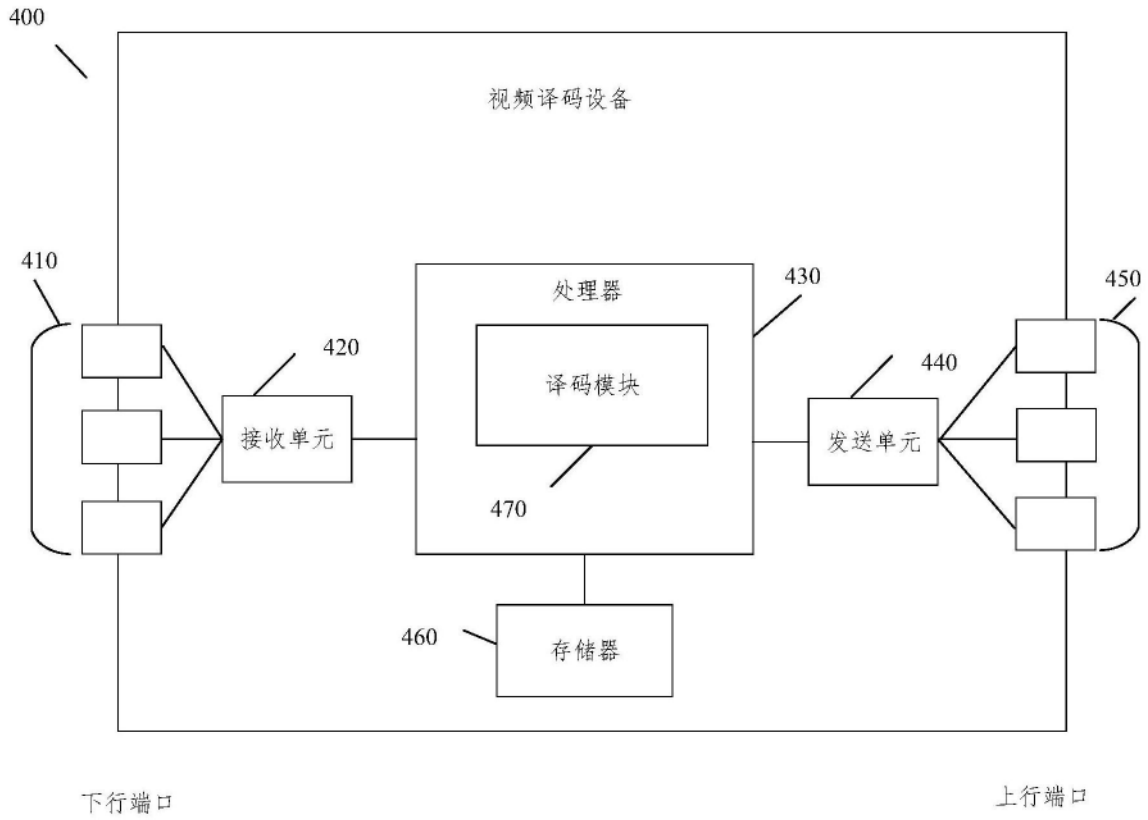


图3

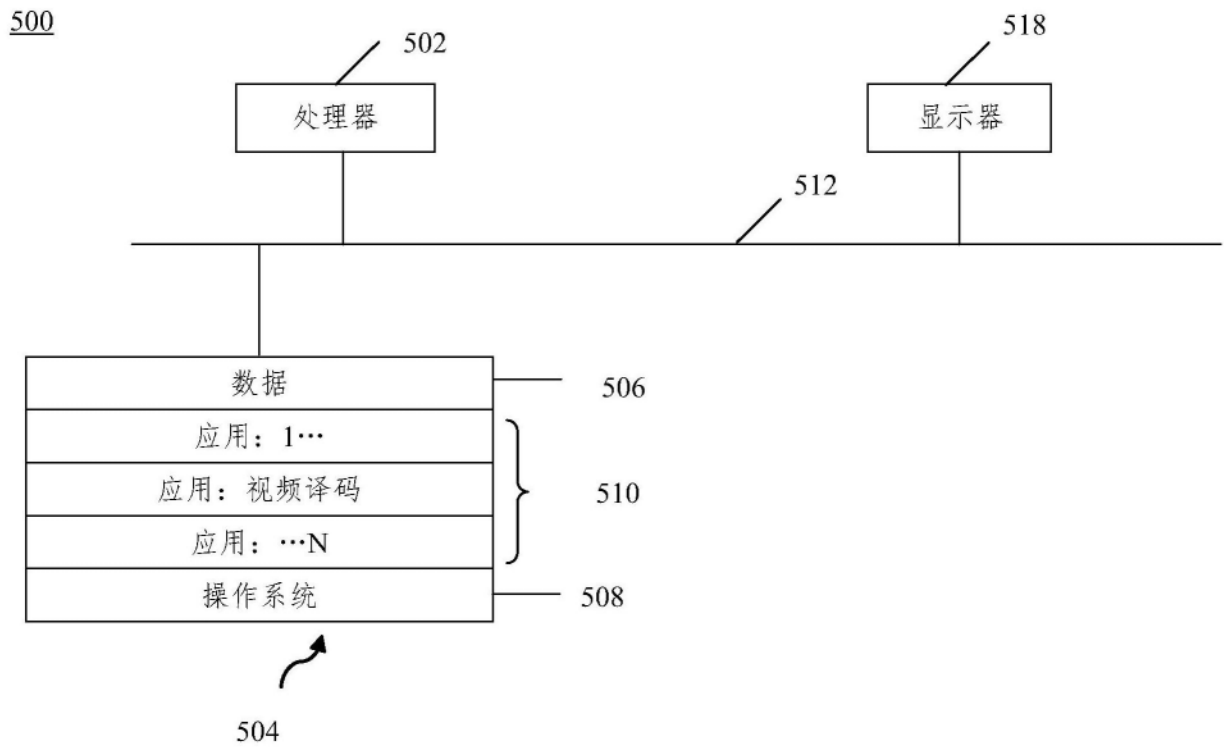


图4

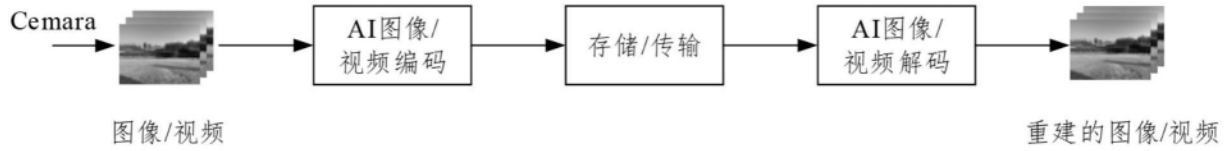


图5

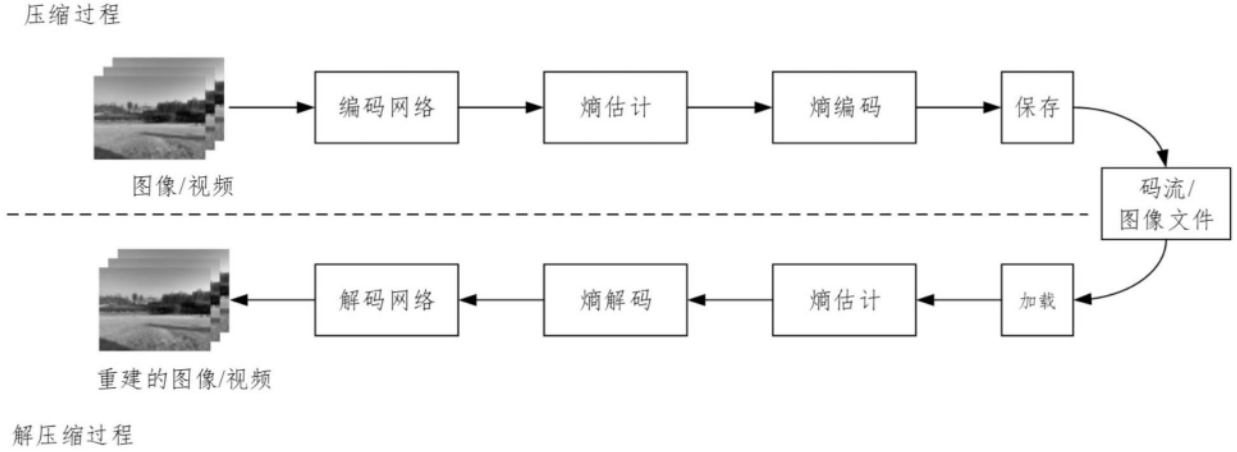


图6

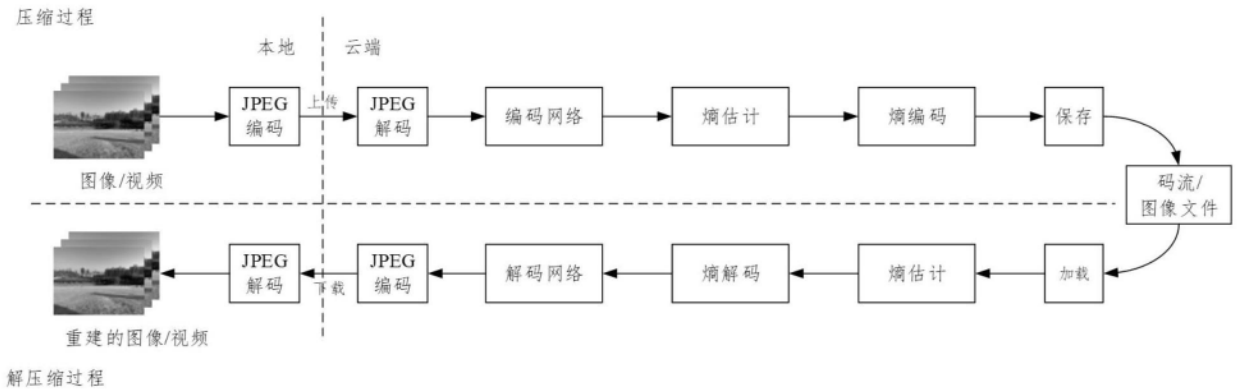


图7

800

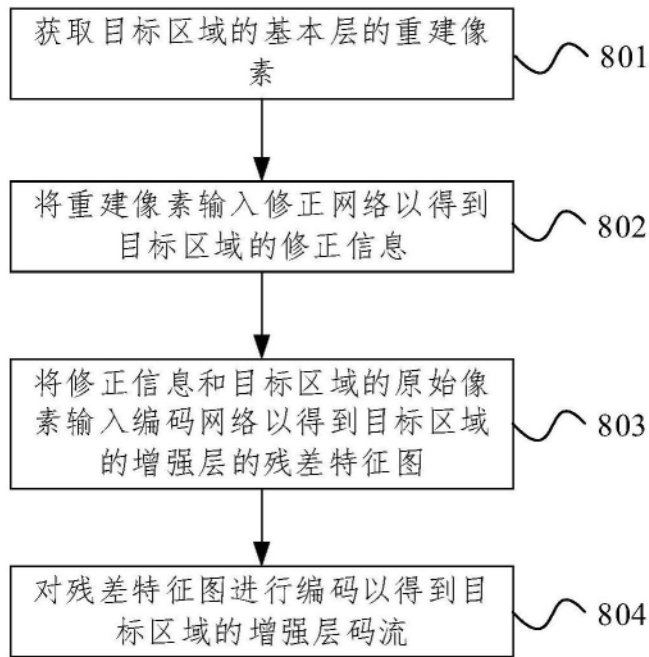


图8

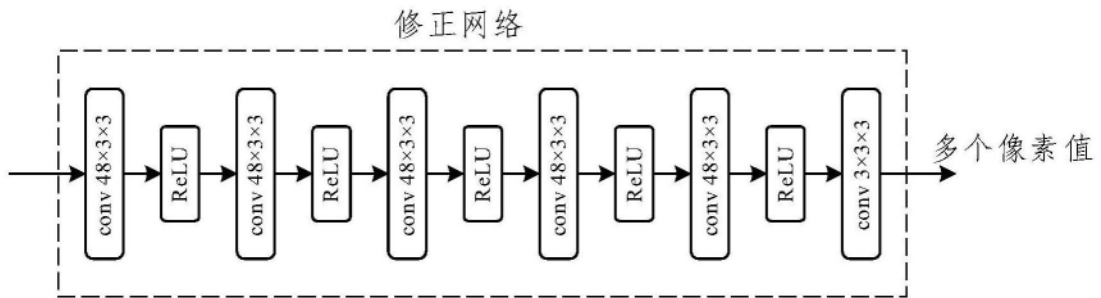


图9a

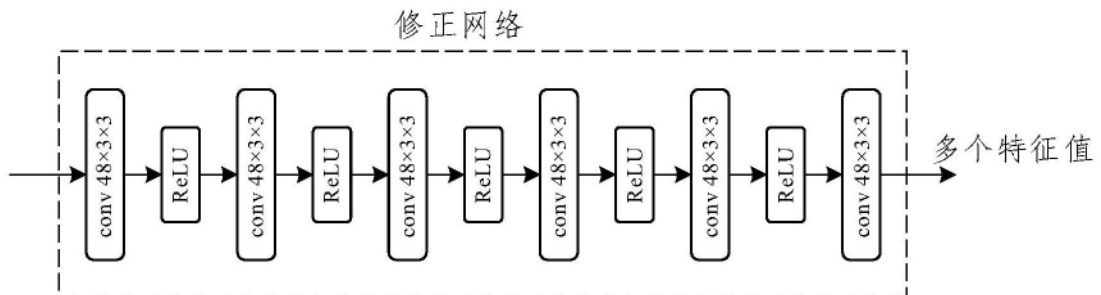


图9b

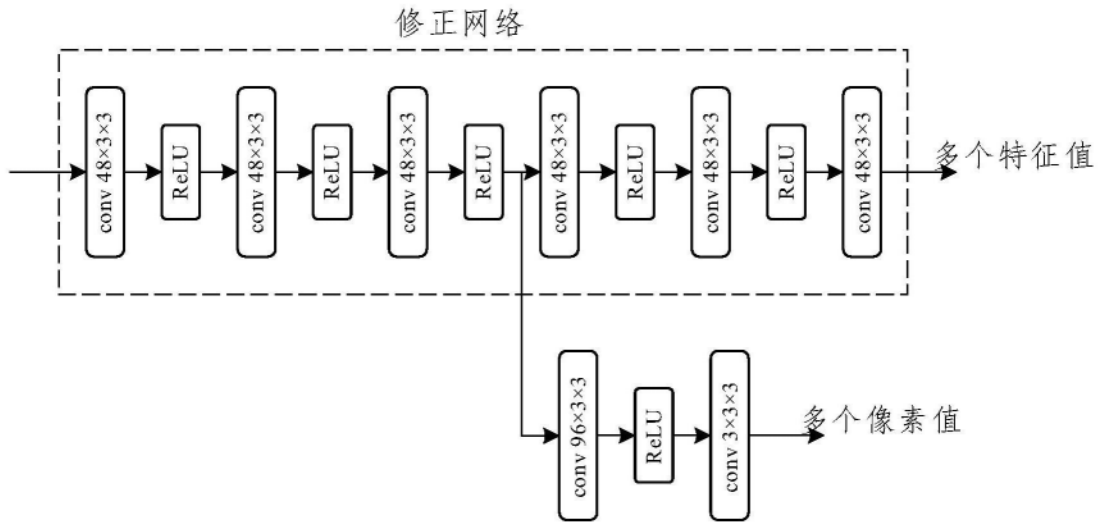


图9c

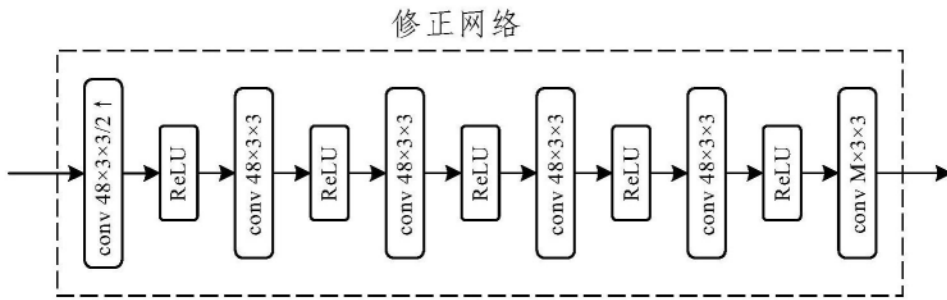


图9d

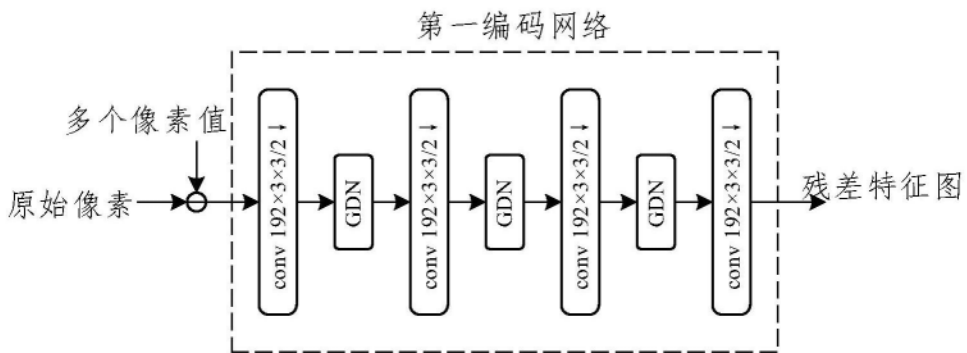


图10a

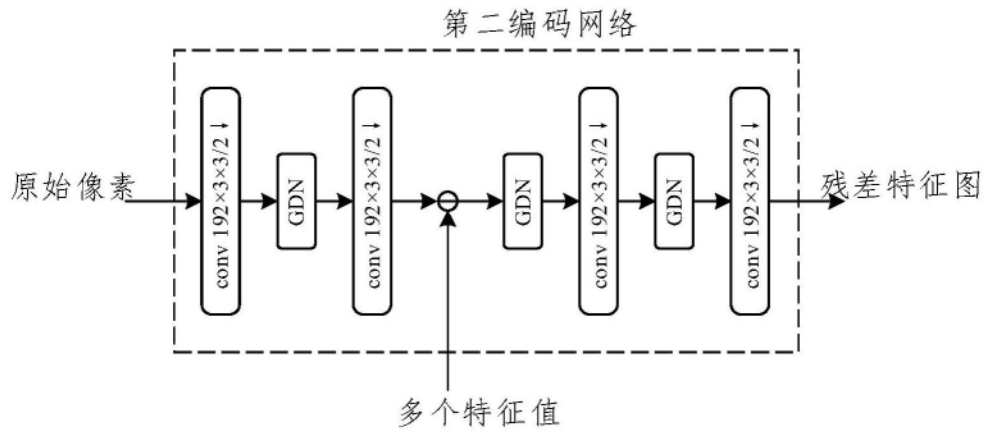


图10b

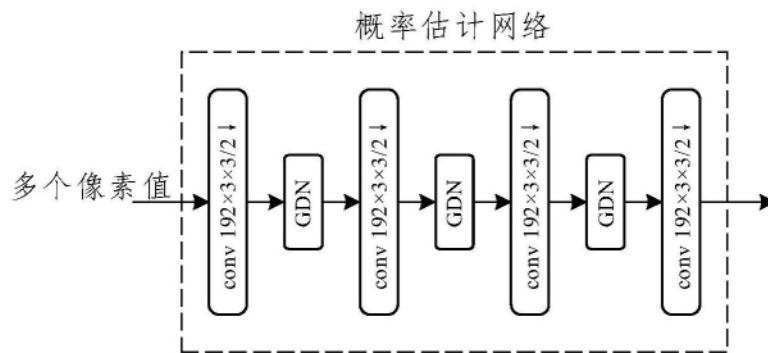


图11a

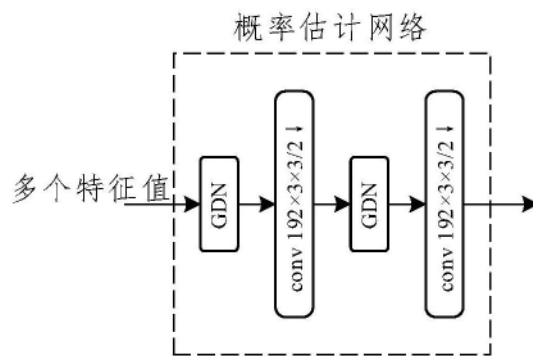


图11b

1200

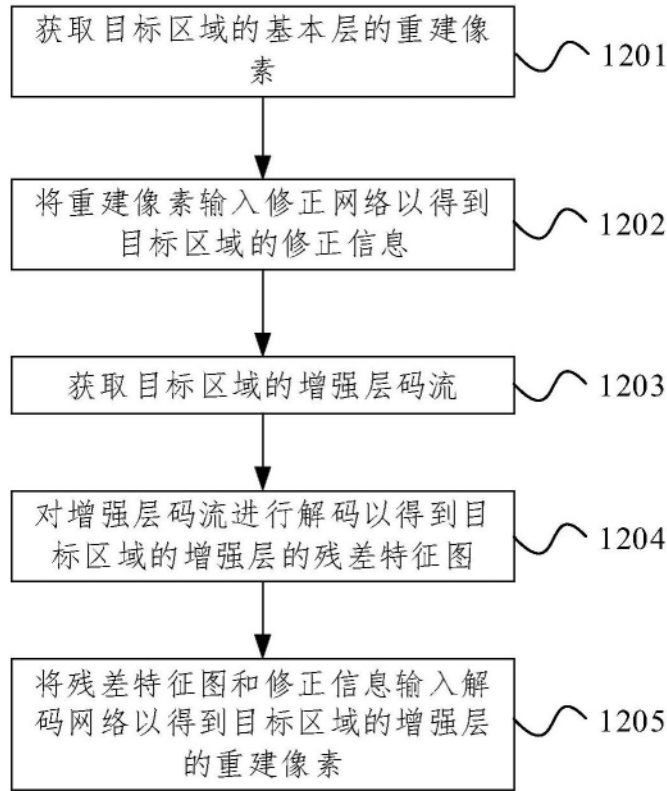


图12

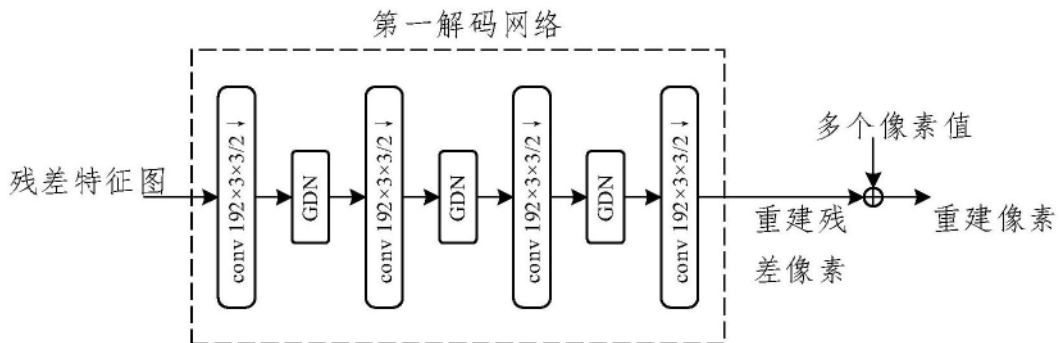


图13a

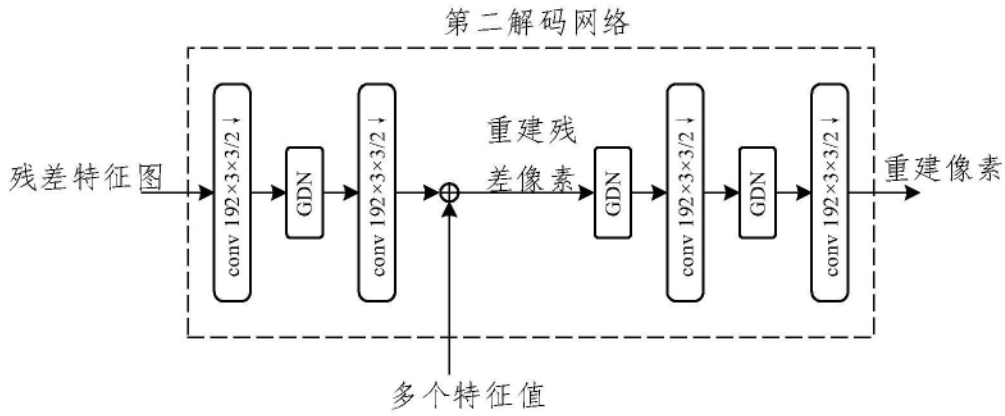


图13b

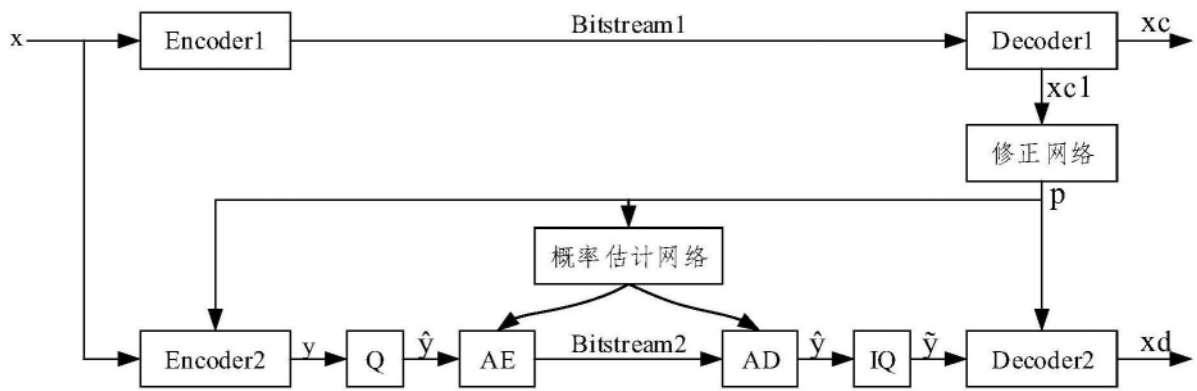


图14

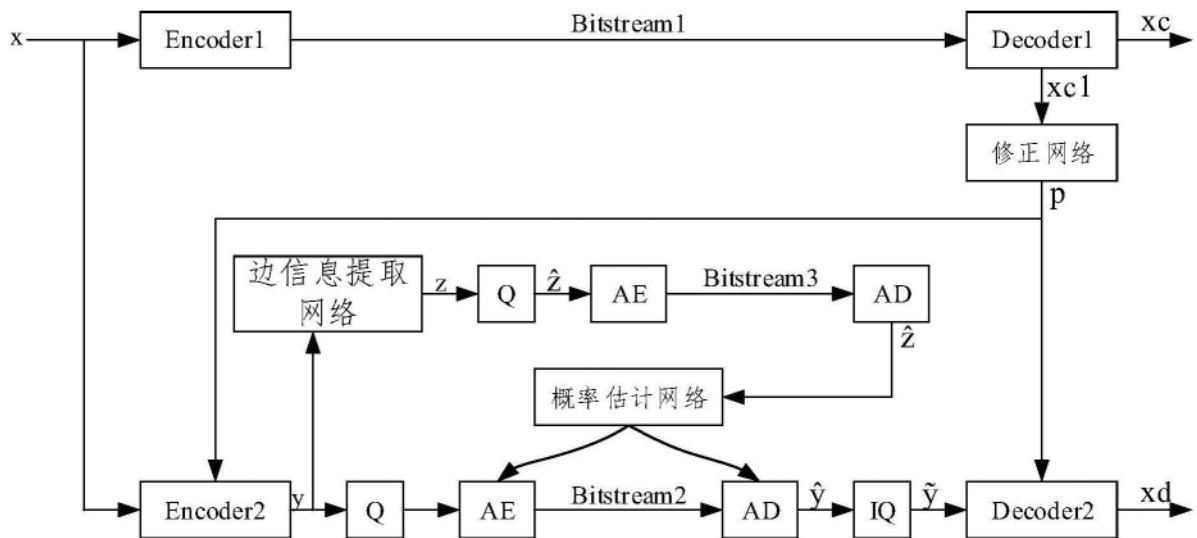


图15

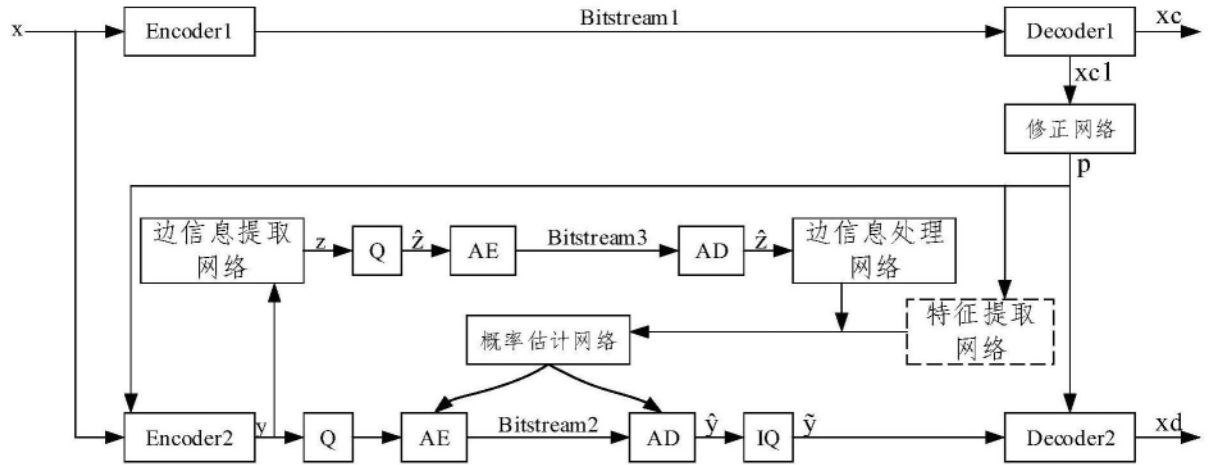


图16

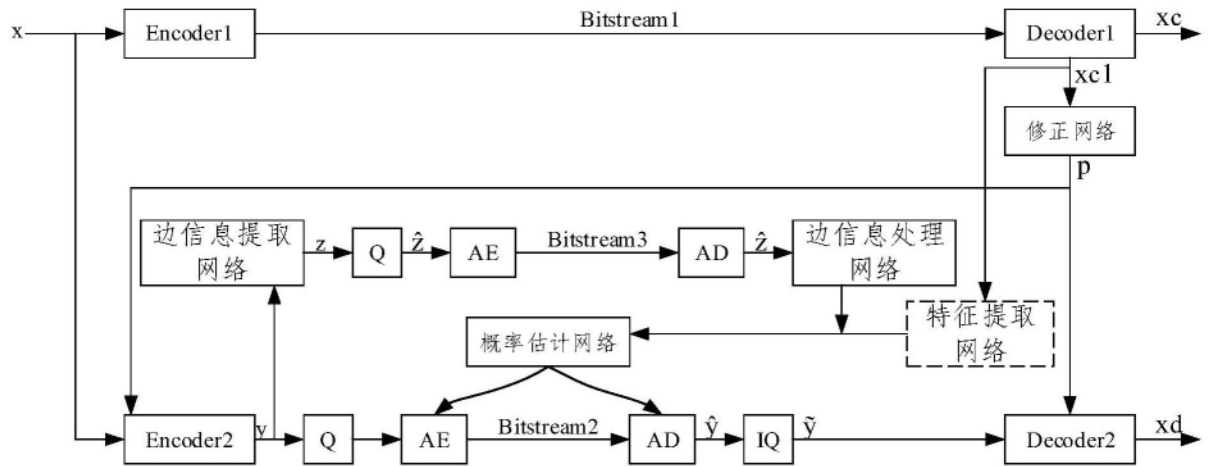


图17

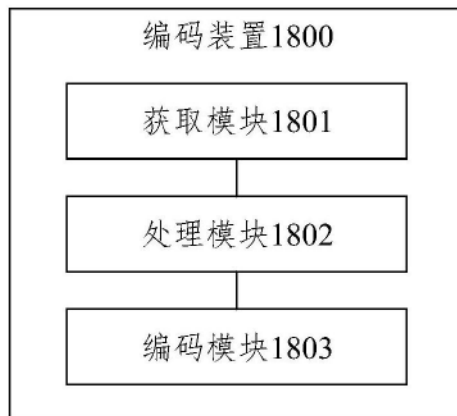


图18

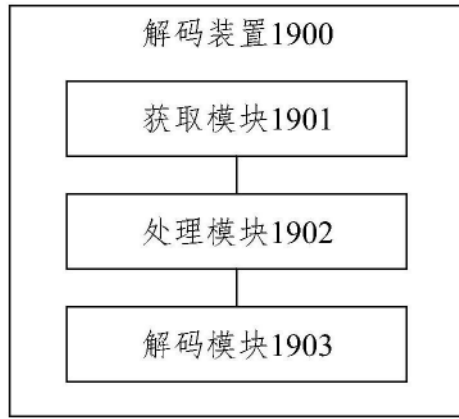


图19