



(12) 发明专利

(10) 授权公告号 CN 112887739 B

(45) 授权公告日 2024. 11. 26

(21) 申请号 202010273164.7  
 (22) 申请日 2020.04.09  
 (65) 同一申请的已公布的文献号  
 申请公布号 CN 112887739 A  
 (43) 申请公布日 2021.06.01  
 (30) 优先权数据  
 10-2019-0157761 2019.11.29 KR  
 (73) 专利权人 三星电子株式会社  
 地址 韩国京畿道水原市  
 (72) 发明人 金敏镐 姜羽锡 罗尚权  
 (74) 专利代理机构 北京铭硕知识产权代理有限公司 11286  
 专利代理师 朱志玲 曾世骁  
 (51) Int. Cl.  
 H04N 19/85 (2014.01)

H04N 19/154 (2014.01)  
 G06N 3/044 (2023.01)  
 G06N 3/0464 (2023.01)  
 G06N 3/09 (2023.01)

(56) 对比文件

US 2017085872 A1, 2017.03.23  
 CN 110121749 A, 2019.08.13  
 KR 20090080830 A, 2009.07.27

Eirina Bourtsoulatzte et al. Deep Video Precoding. IEEE Trans. Circ. Syst. for Video Technology, Arxiv.org, Cornell University Library, 201 Olin Library Cornell University Ithaca, NY 14853. 2019, 文章第一部分, 图1.

审查员 王瑞

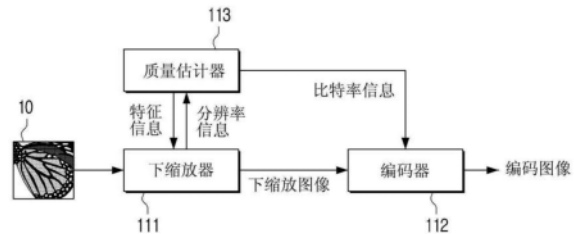
权利要求书2页 说明书39页 附图25页

(54) 发明名称

电子设备、系统及其控制方法

(57) 摘要

公开了一种电子设备、系统及其控制方法。所述电子设备包括: 处理器, 被配置为使用训练的第一人工智能 (AI) 模型对图像进行下缩放并且对下缩放图像进行编码, 并且该处理器被配置为基于使用训练的第二AI模型获得的所述图像的质量信息控制对所述图像进行下缩放, 并且第二AI模型可使用从第一AI模型获得的所述图像的特征信息被训练。



1. 一种电子设备,包括:  
处理器,被配置为:  
使用训练的第二人工智能AI模型获得图像的质量信息;  
基于所述图像的质量信息确定所述图像的下缩放比率和下缩放图像的比特率;  
使用训练的第一AI模型基于确定的下缩放比率对图像进行下缩放;并且  
基于确定的比特率对下缩放图像进行编码,  
其中,第二AI模型使用从第一AI模型获得的所述图像的特征信息被训练,  
其中,处理器还被配置为基于根据所述图像的当前帧获得的质量信息控制对当前帧进行下缩放,或者基于根据所述图像的至少一个先前帧预测的质量信息控制对当前帧进行下缩放。
2. 如权利要求1所述的电子设备,其中,所述图像的质量信息包括和与所述图像相关的分辨率信息和比特率信息不同组合相应的多个质量标识符,  
其中,处理器被配置为:  
基于与所述多个质量标识符中的任意一个质量标识符相应的分辨率信息和比特率信息,确定所述图像的下缩放比率和下缩放图像的比特率,并且  
基于确定的下缩放比率控制对所述图像进行下缩放,并且基于确定的比特率控制对下缩放图像进行编码。
3. 如权利要求1所述的电子设备,还包括:  
通信接口,包括通信电路,  
其中,处理器被配置为获得通信接口连接到的网络的状态信息,并且基于网络的状态信息和所述图像的质量信息控制对所述图像进行下缩放。
4. 如权利要求3所述的电子设备,其中,所述图像的质量信息包括和与所述图像相关的关于多个分辨率的信息和关于多个比特率的信息的不同组合相应的多个质量标识符,  
其中,处理器被配置为:  
基于网络的状态信息,确定所述多个质量标识符中的至少一个质量标识符;基于目标质量标识符,确定所述至少一个质量标识符中的一个质量标识符;  
基于与确定的所述一个质量标识符相应的分辨率信息控制对所述图像进行下缩放,并且基于与确定的所述一个质量标识符相应的关于比特率的信息控制对下缩放图像进行编码。
5. 如权利要求1所述的电子设备,其中,第二AI模型包括至少一个神经网络层,并且  
其中,所述至少一个神经网络层的参数与第一AI模型中包括的至少一些层的参数被共同训练。
6. 如权利要求1所述的电子设备,其中,第二AI模型包括被训练为基于与至少一个先前帧相应的特征信息来预测当前帧的质量信息的至少一个循环神经网络RNN层。
7. 如权利要求1所述的电子设备,其中,所述图像的质量信息包括和与所述图像相关的关于多个分辨率的信息和关于多个比特率的信息的不同组合相应的多个质量标识符,并且  
其中,第二AI模型利用和与训练图像相关的关于多个分辨率的信息和关于多个比特率的信息的不同组合相应的多个质量标识符与通过将所述训练图像输入到第二AI模型而输出的多个质量标识符之间的差作为损失函数,基于监督训练而被训练。

8. 如权利要求1所述的电子设备,其中,处理器被配置为:  
基于编码后的图像产生数据包并发送产生的数据包,并且  
其中,分辨率信息和比特率信息被包括在所述数据包的报头中。

9. 一种包括发送设备和接收设备的系统,所述系统包括:

发送设备,被配置为使用训练的第二人工智能AI模型获得图像的质量信息,基于所述图像的质量信息确定所述图像的下缩放比率和下缩放图像的比特率,使用训练的第一AI模型基于确定的下缩放比率对图像进行下缩放,并且基于确定的比特率对下缩放图像进行编码;以及

接收设备,被配置为对从发送设备接收到的编码图像进行解码,并且通过对解码图像进行上缩放来获得输出图像,

其中,第二AI模型被配置为使用从第一AI模型获得的所述图像的特征信息被训练,

其中,发送设备被配置为基于根据所述图像的当前帧获得的质量信息控制对当前帧进行下缩放,或者基于根据所述图像的至少一个先前帧预测的质量信息控制对当前帧进行下缩放。

10. 如权利要求9所述的系统,其中,接收到的编码图像包括比特率信息或分辨率信息中的至少一个,并且

其中,接收设备被配置为基于接收到的编码图像中包括的比特率信息对编码图像进行解码,并且基于接收到的编码图像中包括的分辨率信息对解码图像进行上缩放。

11. 一种用于控制电子设备的方法,所述方法包括:

使用训练的第二人工智能AI模型获得图像的质量信息;

基于所述图像的质量信息确定所述图像的下缩放比率和下缩放图像的比特率;

使用训练的第一AI模型基于确定的下缩放比率对图像进行下缩放;并且

基于确定的比特率对下缩放图像进行编码,

其中,第二AI模型使用从第一AI模型获得的所述图像的特征信息被训练,

其中,所述方法还包括:基于根据所述图像的当前帧获得的质量信息控制对当前帧进行下缩放,或者基于根据所述图像的至少一个先前帧预测的质量信息控制对当前帧进行下缩放。

## 电子设备、系统及其控制方法

### 技术领域

[0001] 本公开涉及一种电子设备及其控制方法,并且例如,涉及一种使用人工智能(AI)模型执行图像处理的电子设备、系统及其控制方法。

### 背景技术

[0002] 随着电子技术的发展,已经开发并发布了各种类型的电子设备。近年来,已经越来越多地开发了在各种场所(诸如,住宅、办公室、公共场所等)中使用的显示设备。

[0003] 随着削减传统付费广播(诸如,地面/有线/IPTV/卫星广播)的剪线趋势、通过公共互联网网络的云上(OTT)服务(诸如,Netflix/Amazon Prime)日益激增,以及单一媒体广播平台(诸如,YouTube/Twitter/Facebook等)的激增,消费者对实时现场流传输应用的需求正在增加。

[0004] OTT服务提供商为实时性重要的内容(诸如,实时广播或体育赛事)提供现场流传输服务,并且在将来,正在准备对需要与观看者交互的服务(诸如,实时投注)的扩展。单人媒体广播平台正在提供一种使能够通过聊天窗口与观看者交互的服务,并且正在演变为能够进行多方视频通话/视频会议的平台。

[0005] 在该实时现场流传输应用中,影响观看者的感觉质量的非常重要的因素是图像的质量和延迟。例如,由于使用公共互联网网络提供服务(诸如,供应链(现场事件>捕捉>编码>第一英里传递>CDN传播>最后一英里传递>用户装置)),因此应利用给定网络环境中的最优图像质量执行流传输,并且由于服务重视实时性,因此消费者感觉到的延迟应被最小化或减少。例如,诸如体育赛事的服务应通过使延迟小于或等于一秒来使剧透效果最小化或减少剧透效果,并且在诸如视频通话的交互式服务中,必须利用至少小于或等于200毫秒的延迟提供该服务,使得消费者不会感到延迟。

[0006] 由于经由互联网和物联网(IoT)/全IP趋势的服务的激增,所以不仅为移动装置(诸如,移动电话或平板电脑),而且还为笔记本电脑、智能TV等提供通信模块(诸如,能够进行互联网协议(IP)通信的Wi-Fi)。在住宅中,正在开发用于装置之间直接流传输的应用场景。

[0007] 屏幕镜像是装置间直接流传输应用场景的代表性示例,其中,已经在相对较小的屏幕(诸如,移动电话和平板电脑)上再现的内容被发送到大屏幕(诸如,用于观看的TV)。在TV行业中,大体积的部件(诸如,电源部件、缆线连接器等)被划分到外部盒中,以保持薄的显示器形状尺寸。这里,可使用无线通信模块利用直接数据流传输来替换外部盒与显示器之间的复杂的缆线连接。

[0008] 当将装置间直接流传输应用中的数据路径和通过互联网的实时现场流传输应用的供应链进行比较时,端系统的技术相似性仅在不同的网络环境下非常高,并且观看者对感觉质量的要求相对较高。例如,无线TV场景的图像质量应接近视觉无损,并且延迟应在10毫秒内,以防止或减少用户的感觉得质量下降。

## 发明内容

[0009] 本公开提供了一种能够预测被实时流传输的图像的质量并以最优和/或提高的分辨率和比特率处理该图像的电子设备、系统及其控制方法。

[0010] 本公开的实施例可解决上述问题,并且可提供一种电子设备,包括:处理器,被配置为使用训练的第一人工智能(AI)模型对图像进行下缩放并且对下缩放图像进行编码,并且该处理器被配置为基于使用训练的第二AI模型获得的图像的质量信息来控制对图像进行下缩放,并且第二AI模型可使用从第一AI模型获得的图像的特征信息被训练。

[0011] 根据示例实施例,一种电子设备包括:处理器,被配置为使用训练的第一人工智能(AI)模型对图像进行下缩放并且对下缩放图像进行编码。这里,处理器可基于使用训练的第二AI模型获得的图像的质量信息来控制对图像进行下缩放,并且第二AI模型可使用从第一AI模型获得的图像的特征信息被训练。

[0012] 根据示例实施例,处理器可基于使用第二AI模型获得的图像的质量信息来控制对下缩放图像进行编码。

[0013] 根据示例实施例,图像的质量信息包括和与图像相关的分辨率信息和比特率信息的不同组合相应的多个质量标识符。处理器可基于与所述多个质量标识符中的任何一个质量标识符相应的分辨率信息和比特率信息来确定图像的下缩放比率 and 下缩放图像的比特率,并且基于确定的下缩放比率控制对图像进行下缩放,并且基于确定的比特率控制对下缩放图像进行编码。

[0014] 根据示例实施例,电子设备可包括:通信接口,包括通信电路。处理器可获得通信接口连接到的网络的状态信息,并且基于网络的状态信息和图像的质量信息来控制对图像进行下缩放。

[0015] 根据示例实施例,图像的质量信息包括和与图像相关的关于多个分辨率的信息和关于多个比特率的信息的不同组合相应的多个质量标识符。处理器可基于网络的状态信息确定所述多个质量标识符中的至少一个质量标识符,基于目标质量标识符确定所述至少一个质量标识符中的一个质量标识符,并且基于与确定的一个质量标识符相应的分辨率信息来控制对图像进行下缩放,并基于与确定的一个质量标识符相应的关于比特率的信息来控制对下缩放图像进行编码。

[0016] 根据示例实施例,第二AI模型包括至少一个神经网络层,并且所述至少一个神经网络层的参数可与第一AI模型中包括的至少一些层的参数被共同训练。

[0017] 根据示例实施例,其中,处理器可基于根据图像的当前帧获得的质量信息来控制对当前帧进行下缩放,或者基于根据图像的至少一个先前帧预测的质量信息来控制对当前帧进行下缩放。

[0018] 根据示例实施例,第二AI模型包括被训练为基于与至少一个先前帧相应的特征信息来预测当前帧的质量信息的至少一个循环神经网络(RNN)层。

[0019] 根据示例实施例,图像的质量信息包括和与图像相关的关于多个分辨率的信息和关于多个比特率的信息的不同组合相应的多个质量标识符,并且第二AI模型可利用和与训练图像相关的关于多个分辨率的信息和关于多个比特率的信息的不同组合相应的多个质量标识符与通过将训练图像输入到第二AI模型而输出的多个质量标识符之间的差作为损失函数,基于监督训练而被训练。

[0020] 根据示例实施例,处理器可基于图像的质量信息来确定图像的下缩放比率和下缩放图像的比特率,基于确定的下缩放比率对图像进行下缩放,并基于确定的比特率对下缩放图像进行编码,并且基于编码图像产生数据包并发送产生的数据包,并且分辨率信息和比特率信息被包括在数据包的报头中。

[0021] 根据示例实施例,一种用于控制电子设备的方法,所述方法包括:使用训练的第一人工智能(AI)模型对图像进行下缩放;并且对下缩放图像进行编码,其中,对图像进行下缩放的步骤包括基于使用训练的第二AI模型获得的图像的质量信息来控制对图像进行下缩放,并且其中,第二AI模型使用从第一AI模型获得的图像的特征信息被训练。

[0022] 根据示例实施例,对下缩放图像进行编码的步骤包括基于使用第二AI模型获得的图像的质量信息来控制对下缩放图像进行编码。

[0023] 根据示例实施例,图像的质量信息包括和与图像相关的分辨率信息和比特率信息不同组合相应的多个质量标识符,对图像进行下缩放的步骤包括基于所述多个质量标识符中的任意一个质量标识符确定图像的下缩放比率并基于确定的下缩放比率来控制对图像进行下缩放,并且对下缩放图像进行编码的步骤包括基于质量标识符中的任意一个质量标识符确定下缩放图像的比特率并基于确定的比特率来控制对下缩放图像进行编码。

[0024] 根据示例实施例,所述方法还包括:获得网络的状态信息,其中,对图像进行下缩放的步骤包括基于网络的状态信息和图像的质量信息来控制对图像进行下缩放,并且其中,对下缩放图像进行编码的步骤包括基于网络的状态信息和图像的质量信息来控制对下缩放图像进行编码。

[0025] 根据示例实施例,图像的质量信息包括和与图像相关的关于多个分辨率的信息和关于多个比特率的信息的不同组合相应的多个质量标识符,并且所述方法还包括:基于网络的状态信息确定所述多个质量标识符中的至少一个质量标识符,并且基于目标质量标识符确定所述至少一个质量标识符中的任意一个质量标识符,其中,对图像进行下缩放的步骤包括基于与确定的所述任意一个质量标识符相应的分辨率信息来控制对图像进行下缩放,并且其中,对下缩放图像进行编码的步骤包括基于与确定的所述任意一个质量标识符相应的关于比特率的信息来控制对下缩放图像进行编码。

[0026] 根据示例实施例,第二AI模型包括至少一个神经网络层,并且其中,所述至少一个神经网络层的参数与在第一AI模型中包括的至少一些层的参数被共同训练。

[0027] 根据示例实施例,对图像进行下缩放的步骤包括:基于根据图像的当前帧获得的质量信息来控制对当前帧进行下缩放;或者基于根据图像的至少一个先前帧预测的质量信息来控制对当前帧进行下缩放。

[0028] 根据示例实施例,一种包括发送设备和接收设备的系统包括:发送设备,包括被配置为使用训练的第一人工智能(AI)模型对图像进行下缩放并对下缩放图像进行编码的电路;以及接收设备,包括被配置为对从发送设备接收到的图像进行解码并通过对解码图像进行上缩放来获得输出图像的电路,其中,发送设备被配置为基于使用训练的第二AI模型获得的图像的质量信息来控制对图像进行下缩放,并且第二AI模型被配置为使用从第一AI模型获得的图像的特征信息被训练。

[0029] 根据示例实施例,发送设备被配置为基于使用第二AI模型获得的图像的质量信息来控制对下缩放图像进行编码。

[0030] 根据示例实施例,接收到的图像包括比特率信息或分辨率信息中的至少一个,接收设备被配置为:基于接收到的图像中包括的比特率信息对编码图像进行解码,并且基于接收到的图像中包括的分辨率信息对解码图像进行上缩放。

## 附图说明

[0031] 从以下结合附图进行的详细描述中,本公开的特定实施例的以上和其它方面和优点将更加明显,其中:

[0032] 图1A是示出根据实施例的基于人工智能神经网络的示例图像处理方法的示意图;

[0033] 图1B是示出根据实施例的基于人工智能神经网络的示例图像处理方法的示意图;

[0034] 图1C是示出根据实施例的基于人工智能神经网络的示例图像处理方法的示意图;

[0035] 图2A是示出根据实施例的电子设备的示例配置的框图;

[0036] 图2B是示出根据实施例的电子设备的示例配置的框图;

[0037] 图3是示出根据实施例的处理器示例操作的框图;

[0038] 图4是示出根据实施例的第一人工智能模型与第二人工智能模型之间的示例关系的示意图;

[0039] 图5是示出根据实施例的第二人工智能模型的示例学习方法的示意图;

[0040] 图6是示出根据实施例的第二人工智能模型的示例学习方法的示意图;

[0041] 图7是示出根据实施例的第二人工智能模型的示例学习方法的示意图;

[0042] 图8A是示出根据各种实施例的使用当前帧或先前帧的质量信息的示例的示意图;

[0043] 图8B是示出根据各种实施例的使用当前帧或先前帧的质量信息的示例的示意图;

[0044] 图9是示出根据实施例的电子设备的示例配置的框图;

[0045] 图10A是示出根据实施例的使用人工智能模型的示例上缩放方法的示意图;

[0046] 图10B是示出根据实施例的使用人工智能模型的示例上缩放方法的示意图;

[0047] 图10C是示出根据实施例的使用人工智能模型的示例上缩放方法的示意图;

[0048] 图11A是示出根据实施例的示例电子设备和示例接收设备的示意图;

[0049] 图11B是示出根据实施例的示例电子设备和示例接收设备的框图;

[0050] 图12是示出根据实施例的发送和接收系统的示例操作的示意图;

[0051] 图13是示出根据实施例的示例质量信息的示意图;

[0052] 图14是示出根据实施例的用于控制电子设备的示例方法的流程图;

[0053] 图15是示出根据实施例的示例AI编码处理和示例AI解码处理的示意图;

[0054] 图16是示出根据实施例的示例AI解码设备的示例配置的框图;

[0055] 图17是示出根据实施例的用于AI上缩放的示例第二深度神经网络(DNN)的示意图;

[0056] 图18是示出根据实施例的通过卷积层的示例卷积运算的示意图;

[0057] 图19是示出根据实施例的各种与图像相关的信息与DNN配置信息之间的示例映射关系的示意图;

[0058] 图20是示出根据实施例的包括多个帧的示例第二图像的示意图;

[0059] 图21是示出根据实施例的AI编码设备的示例配置的框图;

[0060] 图22是示出根据实施例的用于对图像进行AI下缩放的第一DNN的示意图;

[0061] 图23是示出根据实施例的用于训练第一DNN和第二DNN的示例方法的示意图;

[0062] 图24是示出根据实施例的由训练装置对第一DNN和第二DNN的示例训练处理的信号流程图;以及

[0063] 图25是示出根据实施例的用于AI下缩放的示例设备和用于AI上缩放的示例设备的框图。

[0064] 在整个附图中,相同的参考标号可用于表示相同(或相似)的元件。

### 具体实施方式

[0065] 本公开可包括各种修改和实施例,并且本公开的具体实施例已经在附图中通过示例的方式被示出并且在本文中被更详细地描述。然而,应理解,本公开不旨在限于本公开的实施例,而是旨在涵盖落入各种实施例和本公开的精神和范围内的所有修改、等同物和替换物。

[0066] 在以下描述中,当确定本公开的主题可能使本公开的要点模糊时,可省略对现有技术的详细描述。此外,在本文的描述中使用的数字(例如,第一、第二等)仅是用于将一个组件与另一组件区分开的标识符。

[0067] 此外,当元件在本文中被称为“正被连接到”另一元件或“被连接到”另一元件时,应理解该元件可被直接连接到或被连接到另一元件,但是除非另有说明,否则该元件可被连接或经由它们之间的另一元件被连接。

[0068] 在本公开中,由“一部分(单元)”、“模块”等表示的元件可被划分为两个或更多个组件,或者一个组件可被划分为针对每个更具体的功能的两个以上的组件。此外,以下描述的组件中的每个组件除了可执行这些组件负责的主要功能之外,还可另外执行其他组件负责的功能的一些功能或所有功能的功能,并且理解到组件中的每个组件负责的功能中的一些功能可被其他组件执行。

[0069] 在本公开中,“图像”或“画面”可表示静止图像、包括多个连续的静止图像(或帧)的运动画面或视频。

[0070] 此外,术语“深度神经网络(DNN)”是模拟脑神经的人工神经网络模型的非限制说明性示例,并且不限于使用特定算法的人工神经网络模型。

[0071] 在本公开中,“参数”可指例如在形成神经网络的每层的运算处理中使用的值,并且可包括例如在将输入值应用于预定等式时所使用的权重值。此外,可以以矩阵形式表示所述参数。所述参数可被设置为训练的结果,并且可根据需要经由单独的训练数据被更新。

[0072] 在本公开中,“第一深度神经网络(DNN)”可指例如用于对图像进行AI下缩放的DNN,“第二DNN”可指例如用于对图像进行AI上缩放的DNN。

[0073] 在本公开中,“DNN配置信息”可指例如与形成DNN的组件相关的信息,并且可包括上述参数。使用DNN配置信息,可设置第一DNN或第二DNN。

[0074] 在本公开中,“图像”可指例如经过AI编码的图像,并且“第一图像”可指例如作为AI编码处理中进行AI下缩放的结果而获取的图像。此外,“第二图像”可指例如在AI解码处理期间通过第一解码处理获取的图像,并且“第三图像”可指例如在AI解码处理中通过对第二图像进行AI上缩放而获取的图像。

[0075] 在本公开中,“AI下缩放”可指例如基于例如AI降低图像的分辨率的处理,并且“第一编码”可指例如通过基于频率转换的图像编码方法的编码处理。“第一解码”可指例如通

过基于频率转换的图像重建方法的解码处理,并且“AI上缩放”可指例如基于例如AI增加图像的分辨率的处理。

[0076] 图1A、图1B和图1C是示出根据实施例的基于人工智能(AI)神经网络的示例图像处理方法的示图。

[0077] 为了通过网络对高清晰度/高分辨率图像(诸如,4K和8K)进行流传输,可减少网络请求带宽的图像编码技术和上缩放/下缩放技术很重要。对于图像编码技术,广泛使用标准编解码器(诸如,H.264/265AVC、VP8/9和AV1算法),对于OTT服务提供商,例如,针对H.265,可将4K压缩到大约15Mbps用于服务。为了为每个用户根据不同的网络环境提供服务,用于以图像分辨率和比特率的各种组合进行压缩的技术是上缩放/下缩放技术。例如,当将以大约15Mbps的水平发送8K图像时,发送端10可执行将图像下缩放到4K,并且可将该图像发送到编解码编码器,并且接收端20可执行上缩放,以便将编解码编码器的4K输出恢复为8K图像。尽管在上缩放/下缩放中使用了简单的插值(诸如,双线性或双三次),但是最近的研究表示,通过使用例如卷积神经网络(CNN)进行上缩放/下缩放可进一步提高消费者的感觉质量。例如,该方法容易与任何压缩编解码器兼容,因此可通过应用于当前广泛使用的H.265/VP9标准编解码器被容易地缩放。

[0078] 通常,压缩效率根据图像分辨率和图像类型而不同。例如,如果具有不同特征的两个图像以相同的分辨率和比特率被压缩,则例如如图1B中所示,消费者的感觉质量下降更可能对看起来相对复杂的下方图像的情况更敏感。另一方面,在看起来相对不太复杂的上方图像的示例中,压缩效率良好,因此,发送具有较低比特率图像可减少不必要地浪费的网络带宽。

[0079] 如图1C的上面曲线图中所示,由于随着分辨率越高,增加比特率有助于提高用户的感觉得质量,因此当根据图像的类型优化比特率时,也考虑分辨率。如图1C的下面曲线图中所示,内容提供商找到用于各种分辨率和比特率组合的最优点以进行编码,以为每个用户的给定网络环境提供具有最优质量的服务。

[0080] 在视频点播(VoD)服务的示例中,图像已被预先编码并被存储在用于进行服务的服务器中,并且图像质量可通过投入足够的时间和计算能力被离线最优化,但是在实时现场流传输应用中,由于延迟和计算资源的限制,所以可能无法使用该方法。

[0081] 在下文中,将更详细地描述在内容的实时流传输中提供具有最优化和/或改进的质量的内容的各种实施例。

[0082] 图2A和图2B是示出根据实施例的电子设备的示例配置的框图。

[0083] 参照图2A,电子设备100包括处理器(例如,包括处理电路)110,并且参照图2B,电子设备100'包括处理器(例如,包括处理电路)110和存储器120。

[0084] 根据实施例的关于AI模型的信息可被存储在处理器110的内部存储器中或外部存储器(例如,存储器120)中,因此,分别示出了图2A和图2B。在下文中,将参照图2B描述实施例。

[0085] 电子设备100可被实现为可提供内容的各种类型的设备,诸如,服务器(例如,但不限于内容提供服务器、个人计算机(PC)等)。电子设备100可以是建立云计算环境的系统本身,但不限于此。电子设备100可被实现为处理图像并将图像发送到外部电子设备的各种类型的图像处理设备,诸如,例如,但不限于机顶盒、数字通用盘(DVD)播放器、游戏机(例

如, Xbox™、PlayStation™)、虚拟现实 (VR) 装置、平板电脑、移动电话、电子书、台式计算机、膝上型计算机、上网本计算机、个人数字助理 (PDA)、便携式多媒体播放器 (PMP)、MP3 播放器、移动医疗装置、家用电器 (诸如, 冰箱)、相机、摄像机、电子框架、可穿戴装置 (例如, 头戴式装置 (HMD)、智能手表、电子服装、电子手镯、电子项链、眼镜等)、近眼显示器 (NED) 等。

[0086] 根据示例, 电子设备100可将图像转换为具有各种分辨率的图像和各种压缩图像, 并且发送所述图像。通过电子设备100处理的图像可以是例如由内容提供商准备的图像, 但不限于此, 并且可以是在该图像上另外处理的图像。在下文中, 为了便于描述, 假设通过电子设备100处理图像。

[0087] 例如, 电子设备100可将图像 (或输入图像) 转换为至少一个图像, 不仅是720×480的标准清晰度 (SD)、1280×720的高清晰度 (HD)、1920×1080的全高清晰度 (FHD)、2560×1440的四倍高清晰度 (QHD)、3840×2160的4K超高清晰度 (UHD)、7680×4320的8K UHD的图像, 而且是更高分辨率 (例如, 16K、32K) 的图像, 并且发送该图像。电子设备100可以以压缩形式 (诸如, 运动画面专家组 (MPEG) (例如, MP2、MP4、MP7等)、联合图像专家组 (JPEG)、高级视频编码 (AVC)、H.264、H.265、高效图像编解码器 (HEVC)、VC-1、VP8、VP9、AOMedian图像1 (AV1) 等, 但本公开不限于此) 压缩图像。

[0088] 存储器120可存储本公开的各种实施例所需的数据。根据数据使用目的, 存储器120可被实现为嵌入在电子设备100中的存储器, 或者可被实现为电子设备100中的可拆卸的存储器。例如, 用于驱动电子设备100的数据可被存储在嵌入在电子设备100中的存储器中, 并且用于电子设备100的附加功能的数据可被存储在对于电子设备100可拆卸的存储器中。嵌入在电子设备100中的存储器可以是易失性存储器 (诸如, 动态随机存取存储器 (DRAM)、静态随机存取存储器 (SRAM)、同步动态随机存取存储器 (SDRAM)) 或非易失性存储器 (例如, 一次性可编程ROM (OTPROM)、可编程ROM (PROM)、可擦除可编程ROM (EPROM)、电可擦除可编程ROM (EEPROM)、掩膜ROM、闪速ROM、闪存 (例如, NAND闪存或NOR闪存)、硬盘驱动器或固态驱动器 (SSD) 等), 并且在存储器可拆卸地安装到电子设备100的情况下, 存储器可被实现为存储卡 (例如, 紧凑式闪存 (CF)、安全数字 (SD)、微型安全数字 (micro-SD)、迷你安全数字 (mini-SD)、极限数字 (xD)、多媒体卡 (MMC) 等)、可连接到USB端口的的外部存储器 (例如, USB存储器) 等, 但本公开不限于此。

[0089] 根据实施例, 存储器120可存储包括用于控制电子设备100的至少一个指令或多个指令的计算机程序。

[0090] 根据另一示例, 存储器120可存储关于包括多个层的AI模型的信息。存储关于AI模型的信息可指例如存储与AI模型的运算相关的各种信息, 例如, 关于AI模型中包括的多个层的信息、关于在所述多个层中的每个层中使用的参数 (例如, 滤波器系数、偏差等) 的信息。例如, 存储器120可存储关于根据示例实施例的被训练为对图像进行下缩放的第一AI模型的信息。存储器120可存储关于根据示例实施例的被训练为获得图像的质量信息的第二AI模型的信息。然而, 应理解, 关于AI模型的信息可根据以下描述的处理器110的实现方式被存储在处理器110的内部存储器中。例如, 如果以专用硬件实现处理器110, 则可将关于AI模型的信息存储在处理器110的内部存储器中。

[0091] 根据另一实施例, 存储器120可存储从外部装置、外部存储介质 (例如, 通用串行总线 (USB))、外部服务器 (例如, 网络盘 (web hard)) 接收到的图像。在此, 图像可以是数字运

动图像,但不限于此。

[0092] 根据另一示例,存储器120可例如但不限于存储下缩放所需的信息、压缩所需的信息、用于执行例如降噪、细节增强、色调映射、对比度增强、颜色增强、帧速率转换等中的至少一个的图像质量处理所需的各种信息。存储器120可存储通过图像处理产生的最终输出图像。

[0093] 根据示例实施例,存储器120可被实现为存储在根据本公开的各种操作中产生的数据的单个存储器。然而,根据另一实施例,存储器120可被实现为包括多个存储器,其中,每个存储器存储不同类型的数据,或者分别存储在不同步骤中产生的数据。

[0094] 在示例中已经描述了各种数据被存储在处理器110的外部存储器120中,但是根据至少一个示例,以上数据中的至少一个数据可被存储在处理器110的内部存储器中。

[0095] 处理器110可包括各种处理电路,并可被电连接到存储器120以控制电子设备100的全部操作。处理器110可被配置有一个或多个处理器。例如,处理器110可通过执行存储在存储器120中的至少一个指令来执行根据本公开的各种实施例的电子设备100的操作。

[0096] 根据实施例的处理器110可被实现有例如但不限于:用于数字图像信号的图像处理的数字信号处理器(DSP)、微处理器、图形处理器(GPU)、AI处理器、神经处理器(NPU)、时间控制器(TCON)等,但不限于此。处理器110可包括例如但不限于中央处理器(CPU)、微控制器单元(MCU)、微处理器(MPU)、控制器、应用处理器(AP)、通信处理器(CP)、高级精简指令集计算(RISC)机器(ARM)处理器、专用处理器中的一个或多个,或者可被定义为相应的术语。可以以嵌入了处理算法的片上系统(SoC)类型或大规模集成(LSI)类型、专用集成电路(ASIC)、或者以现场可编程门阵列(FPGA)类型实现处理器110。

[0097] 此外,根据示例实施例的用于执行AI模型的处理器110可以是通用处理器(诸如,中央处理器(CPU)、应用处理器(AP)、数字信号处理器(DSP)、专用处理器等)、纯图形处理器(诸如,图形处理器(GPU)、视觉处理器(VPU))、纯AI处理器(诸如,神经网络处理器(NPU)等),但不限于此。处理器110可根据存储在存储器120中的预定义的运算规则或AI模型来执行控制以处理输入数据。如果处理器110是纯AI处理器,则可被设计有专用于特定AI模型的处理的硬件结构。例如,专用于特定AI模型的处理的硬件可被设计到硬件芯片(诸如,专用集成电路(ASIC)、现场可编程门阵列(FPGA)等)中。

[0098] 处理器110处理输入数据以获得输出数据。输入数据可包括例如但不限于文本、图像、用户语音等中的至少一个。例如,可经由用户输入(诸如,能够与外部装置进行通信的通信器(例如,包括通信电路)、键盘或触摸板、相机、麦克风等)输入输入数据。输出数据可根据AI模型的类型采用各种形式。例如,输出数据可以是具有提高的分辨率的图像、图像中包括的对象相关信息、与语音相应的文本等。

[0099] 根据示例,处理器110可对图像进行图像处理以获得输出图像。图像可包括静止图像、多个连续的静止图像(或帧)。图像可以是原始图像,但不限于此。图像处理可以是包括图像增强、图像重建、图像变换、图像分析、图像理解或图像压缩等中的至少一个的数字图像处理,但不限于此。

[0100] 根据示例实施例,处理器110可使用AI模型对图像(例如,原始图像)执行图像处理。例如,处理器110可加载和使用存储在存储器120(例如,诸如动态随机存取存储器(DRAM)的外部存储器)中的AI模型相关信息以使用AI模型。

[0101] 根据示例实施例,处理器110可对图像进行下缩放并且对下缩放图像进行编码(或压缩)。此外,处理器110可基于图像的质量信息来控制对图像进行下缩放或对下缩放图像进行编码中的至少一个。可使用第一AI模型执行对图像进行下缩放。可通过第二AI模型获得图像的质量信息。在这种情况下,可基于从第一AI模型获得的图像的特征信息学习第二AI模型。

[0102] 图3是示出根据实施例的处理器110的示例操作的示图。

[0103] 根据示例实施例,处理器110可包括下缩放器(例如,包括处理电路和/或可执行程序元件)111、编码器(例如,包括处理电路和/或可执行程序元件)112以及质量估计器(例如,包括处理电路和/或可执行程序元件)113。为了便于描述,下缩放器111、编码器112和质量估计器113对处理器110的功能进行区分,并且可在处理器110内用至少一个软件、至少一个硬件或它们的组合来实现下缩放器111、编码器112和质量估计器113。

[0104] 根据示例实施例,下缩放器111可包括各种处理电路和/或可执行的程序元件,并且使用第一AI模型对图像进行下缩放。第一AI模型的输出可以是下缩放图像。此外,第一AI模型可联合用于对图像进行上缩放的AI模型被共同训练,其中,用于对图像进行上缩放的AI模型包括在用于从电子设备100接收处理后的图像的接收装置(图1A和图20)中。

[0105] 被训练的AI模型可指例如通过学习算法使用多个训练数据被训练的基本AI模型(例如,包括任意随机参数的AI模型),从而产生被设置为执行预期属性(或目的)的预定动作规则或AI模型。可通过单独的服务器和/或系统实现学习,但不限于此,并且可在电子设备中实现学习。学习算法的示例包括但不限于监督学习、无监督学习、半监督学习或强化学习。

[0106] 第一AI模型可被实现为例如但不限于卷积神经网络(CNN)、循环神经网络(RNN)、受限玻尔兹曼机(RBM)、深度置信网络(DBN)、双向循环深度神经网络(BRDNN)、深度Q网络等,但不限于此。

[0107] 关于下缩放,例如,可使用至少一个插值,诸如,例如,但不限于双线性插值、最近邻插值、双三次插值、反卷积插值、子像素卷积插值、多相插值、三线性插值、线性插值等。

[0108] 根据实施例,编码器112可包括各种处理电路和/或可执行的程序元件,并且通过下缩放器111对下缩放图像进行编码(或压缩)。例如,处理器110可使用各种压缩方法(诸如,例如,但不限于运动画面专家组(MPEG)(例如,MP2、MP4、MP7等)、联合图像专家组(JPEG)、高级视频编码(AVC)、H.264、H.265、高效图像编解码器(HEVC)、VC-1、VP8、VP9、AOMedian图像1(AV1)等)中的一种压缩方法对下缩放图像进行编码。

[0109] 质量估计器113可包括各种处理电路和/或可执行的程序元件,并且估计图像的质量并基于估计的质量控制下缩放器111或编码器112中的至少一个。例如,质量估计器113可获得图像的质量信息,并且基于获得的质量信息控制下缩放器111的下缩放处理或编码器112的编码处理中的至少一个。质量信息可包括与分辨率信息或比特率(或压缩强度)信息中的至少一个相应的质量标识符。然而,可与可控制下缩放器111或编码器112中的至少一个的其他类型的信息一起实现质量信息,但不限于此。例如,质量信息可包括与下缩放信息(例如,下缩放比率)或比特率(或压缩强度)信息中的至少一个相应的质量标识符。

[0110] 当在图像中获得与对于图像最优的质量信息相应的分辨率信息和比特率信息时,质量估计器113可将获得的分辨率信息提供给下缩放器111,并且将获得的比特率信息提供

给编码器112。也就是说,处理器110可基于图像质量实时地控制图像的最优分辨率和比特率。

[0111] 根据实施例,质量估计器113可使用被训练为获得图像的质量信息的第二AI模型来获得图像的质量信息,并且可基于获得的质量信息控制对图像进行下缩放或对下缩放图像进行编码中的至少一个。可训练第二AI模型以使用从第一AI模型获得的图像的特征信息来获得图像的质量信息。作为示例,第二AI模型可估计并输出针对预定义分辨率和比特率组合的矢量形式的图像质量标识符。矢量形式可指例如针对预定义分辨率和比特率组合而获得的各种质量标识符。然而,在另一示例中,第二AI模型可估计并输出针对预定义下缩放比率和比特率组合的矢量形式的图像质量标识符。

[0112] 图4是示出根据实施例的第一人工智能模型与第二人工智能模型之间的示例关系的示图。

[0113] 如图4中所示,第一AI模型410可包括多个神经网络层,并且所述多个神经网络层中的每个神经网络层可包括多个参数。第一AI模型410可通过先前层的运算结果和多个参数之间的运算来执行神经网络运算。

[0114] 根据示例,在将卷积滤波器应用于任意层之后,可通过激活功能(例如,整流线性单元(ReLU)运算)输出运算数据。在这种情况下,从该任意层输出的运算数据是多通道数据,例如,可输出64个特征图(或激活图)数据并将其提供给下一层。根据示例,可将特征图数据存储在存储器(内部缓冲器或外部存储器)中,然后将其提供给下一层。然而,在图4中未示出该配置。参数可指例如滤波器的权重(或系数)。在这种情况下,第一AI模型410可使用各种类型的激活函数(诸如,例如,但不限于标识函数、对数S形函数、双曲正切(tanh)函数、ReLU函数、带泄漏ReLU函数等)执行运算。

[0115] 如图4中所示,第二AI模型420可包括多个神经网络层,并且所述多个神经网络层中的每个神经网络层可包括多个参数。可结合在第一AI模型410中包括的至少一个层411来训练在第二AI模型420中包括的至少一个层421。

[0116] 作为示例,第二AI模型420可被实现为与第一AI模型420共享至少一个层411。作为另一示例,第二AI模型420可被实现为仅包括使用从第一AI模型410输出的特征信息的一个层421。然而,为了便于描述,假设第二AI模型420被实现为与第一AI模型420共享至少一个层411。

[0117] 例如,第二AI模型420可包括至少一个层421,至少一个层421用于接收从第一AI模型410中包括的多个层中的特定层412输出的特征图数据。因此,可以以包括第一AI模型410中包括的层411中的至少一些层和添加的至少一个层421的形式来实现第二AI模型420。特定层412可以是属于原始分辨率域的层。例如,由于从第二AI模型输出的质量信息的分辨率信息被反馈到第一AI模型410以控制第一AI模型410的下缩放,因此第一AI模型410可通过插值运算(例如,双三次)来执行下缩放,其中,从所述域的层输出的特征图数据可被提供给第二AI模型420。因此,应将执行下缩放之前的特征图数据提供给第二AI模型420。

[0118] 根据示例实施例,第二AI模型420可被实现为使与第一AI模型410共享的至少一个层411的参数固定,并且训练不与第一AI模型410共享的其余层421的参数。例如,可将迁移学习用于训练第二AI模型420。迁移学习可指例如重新使用和学习在另一应用中针对特定应用已经训练好的模型的方法。例如,假设存在如图5的上部中所示的训练好的用于对图像

中的动物进行分类的模型。当产生如图5的下部中所示的具有更多类型的动物的模型时,可重新使用现有的训练的模型而无需从开始进行重新训练,并且可仅额外训练分类器部分以产生新的模型。过渡学习方法具有利用少量数据快速学习数据的优点,并且已知在性能方面是极好的。当已经存在训练的模型并且试图解决与该模型相似的问题时,过度学习方法是有效的方法。这样,可将第二AI模型420实现为使用训练好的用于下缩放的第一AI模型410的至少一些层411,并且仅进一步训练其余的附加层421。

[0119] 然而,如上所述,即使第二AI模型420被实现为仅包括用于接收从第一AI模型410中包括的至少一个层输出的特征图数据的一个层421,第二AI模型420中包括的至少一个层421的参数也可与在第一AI模型410中包括的至少一些层411的参数被共同训练。

[0120] 通过实现第二AI模型420共享从第一AI模型410获得的图像的特征信息,可有效地训练第二AI模型420并且使H/W实现复杂度最小化并且/或者降低H/W实现复杂度。

[0121] 根据示例实施例,第二AI模型420可包括用于提取输入图像的特征信息的卷积神经网络(CNN)和用于将二维图像转换为一维质量信息的全连接网络(FCN)。在一些情况下,第二AI模型420还可包括用于预测图像的每一帧的时间变化的循环神经网络(RNN)层。这样,第二AI模型420可被实现为包括适合于预测图像质量的各种类型的AI模型。

[0122] 根据实施例,从第二AI模型420输出的图像的质量信息可包括和与图像相关联的多个比特率信息和多个分辨率信息的不同组合相应的多个质量标识符。所述多个分辨率信息和所述多个比特率信息的组合可被预定义用于学习第二AI模型420。

[0123] 根据实施例,图像的质量信息可包括与分辨率信息和比特率信息相应的质量索引。例如,分辨率信息可包括例如但不限于标准清晰度(SD)、高清晰度(HD)、全高清晰度(FHD)、四倍高清晰度(QHD)、4K超高清晰度(4KUHD)、8K超高清晰度(8K UHD)或其它,但不限于此。此外,比特率信息可以是图像的比特率信息(例如,比特率\_0、比特率\_1、比特率\_2、...、比特率\_N)的形式。此外,质量标识符可以是VQ度量\_0、VQ度量\_1、VQ度量\_2、...、VQ度量\_N的形式。VQ度量可被实现为可测量图像质量的各种质量度量,诸如,例如,但不限于峰值信噪比(PSNR)、图像多元法评估融合(VMAF)、结构相似度(SSIM)、自然度图像质量评估器(NIQE)、MSAD、均方误差(MSE)等。

[0124] 根据示例,从第二AI模型420输出的图像的质量信息可以是包括分辨率、比特率和质量标识符的一对数据类型。例如,可将第二AI模型420训练为输出用于预定义分辨率和比特率组合的质量标识符。

[0125] 例如,从第二AI模型420输出的图像的质量信息可包括与不同分辨率信息和不同比特率信息相应的多个质量标识符,诸如,例如,但不限于(FHD,比特率\_0,VQ度量\_0)、(FHD,比特率\_1,VQ度量\_1)、(FHD,比特率\_2,VQ度量\_2)、(FHD,比特率\_3,VQ度量\_3)、(2K,比特率\_4,VQ度量\_4)、(2K,比特率\_5,VQ度量\_5)、(2K,比特率\_6,VQ度量\_6)、(2K,比特率\_7,VQ度量\_7)、(4K,比特率\_8,VQ度量\_8)、(4K,比特率\_9,VQ度量\_9)和(4K,比特率\_10,VQ度量\_10)、(4K,比特率\_11,VQ度量\_11)。1至N仅是用于标识比特率和VQ度量的数值,并且比特率和VQ度量可由各种值来表示。

[0126] 根据实施例,可基于多个第一样本质量标识符与多个第二质量标识符之间的差来训练第二AI模型420,其中,所述多个第一样本质量标识符相应于与样本图像相关联的多个比特率信息和多个分辨率信息的不同组合,所述多个第二质量标识符通过将样本图像输入

到第二AI模型420被输出。第一样本质量标识符可以是由接收设备200根据电子设备100处理的图像重建的重建图像的实际质量标识符。

[0127] 根据实施例,如图6中所示,可通过监督学习使用损失函数由从第二AI模型420输出的质量信息与实际图像的质量信息之间的差来训练第二AI模型420。可通过以下操作来获得实际图像的质量信息:根据示例对电子设备100中的原始训练图像进行下缩放和编码、通过网络将图像发送到接收设备200、然后对从接收设备200接收到的图像进行解码和上缩放以对重建图像与原始训练图像进行比较。如图6中所示,第二AI模型420可包括用于提取输入图像的特征信息的卷积神经网络(CNN)和用于将二维图像转换为一维质量信息的全连接网络(FCN)。在一些情况下,第二AI模型420还可包括用于估计图像的每个帧的时间变化的循环神经网络(RNN)层。

[0128] 例如,当将实际图像质量标识符实现为PSNR时,可基于如下所示的等式1和2计算指示实际图像质量标识符的PSNR。

[0129] 【等式1】

$$[0130] \quad PSNR=10*\log_{10}\frac{255^2}{MSE}$$

[0131] 【等式2】

$$[0132] \quad MSE=\frac{1}{X*Y}\sum_{i=1}^Q\sum_{j=1}^Q|C(i,j)-R(i,j)|^2$$

[0133] 这里,C(i,j)是图像,并且R(i,j)是重建图像。

[0134] 返回图2A和图2B,处理器110可基于与从第二AI模型420输出的多个质量标识符中的任意一个质量标识符相应的分辨率信息和比特率信息,确定图像的下缩放比率 and 下缩放图像的比特率。

[0135] 根据示例,处理器110可基于多个质量标识符中的最佳质量标识符确定图像的下缩放比率和下缩放图像的比特率,例如,与第一质量标识符相应的分辨率信息和比特率信息。例如,处理器110可基于与第一质量标识符相应的分辨率信息控制第一AI模型410的下缩放运算,并且可基于与第一质量标识符相应的比特率信息控制从第一AI模型410输出的下缩放图像的压缩强度。然而,最高质量标识符是示例,并且处理器110可利用多个质量标识符中的平均质量标识符,或者可使用与预定义的阈值质量标识符相应的质量标识符(例如,与比最高质量小10%的质量相应的质量标识符)。

[0136] 根据另一实施例,处理器110可获得网络状态信息,并且基于网络的状态信息和图像的质量信息控制对图像进行下缩放或对下缩放图像进行编码中的至少一个。网络状态信息可包括例如但不限于网络的可用带宽、数据丢包率、数据包的往返时间、数据包的延迟度量梯度、接收信号强度标识符(RSSI)信息、通信信道信息、链路速度信息、信道干扰信息、重试率信息等中至少一个。网络状态信息可以是受网络的拥塞、距离和传输速度影响的各种网络相关信息。数据包的往返时间(或往返延迟)可指例如数据包从网络到接收端往返所花费的时间。数据包的延迟梯度可以是在发送端发送数据包的发送时间间隔与在接收端接收数据包的接收时间间隔之间的差。通信信道信息可包括关于信道和频率中的至少一个的信息。链路速度信息可包括关于从发送端发送到接收端的数据包的速率的信息。可基于使用

各种信道RF参数(诸如,信道活动时间、信道繁忙时间和信道传输时间)测量的每个信道的干扰因子来获得信道干扰信息。重试率信息可指示以秒为单位重试的传输数据包(TX数据包)的百分比。

[0137] 根据示例,可从经由网络与电子设备100和100'连接的内容接收装置(或内容输出装置)(例如,图1A和图20)接收网络状态信息。在该示例中,内容接收装置(或内容输出装置)(例如,图1A和图20)可监视网络状态以收集网络状态信息,并且如果存在预定时段或电子设备100和100'的请求,则内容接收装置可将收集的网络状态信息发送到电子设备100和100'。例如,如果利用基于互联网的传输控制协议(TCP)/互联网协议(IP)通信协议实现网络,则TCP/IP可被实现为链路层、互联网层、传输层和应用层。在该示例中,可通过在应用层中实时监视网络状态来收集网络状态信息,但不限于此。然而,在这种情况下,电子设备100、100'可通过直接监视网络状态来收集网络状态信息。

[0138] 根据示例,处理器110可基于网络的状态信息和图像的质量信息来控制对图像进行下缩放或对下缩放图像进行编码中的至少一个。例如,处理器110可基于网络的状态信息确定多个质量标识符中的至少一个质量标识符,并且可基于目标质量标识符确定所述至少一个质量标识符中的质量标识符。在这种情况下,处理器110可基于在网络状态下当前可用的比特率来确定一个质量标识符。当存在可在当前网络状态下传输的多个比特率时(即,与相应比特率相应的多个质量标识符),处理器110可选择最低比特率的质量标识符。因此,可尽可能地减少网络带宽。

[0139] 处理器110可基于与确定的图像质量标识符相应的分辨率信息和比特率信息来确定图像的下缩放比率和下缩放图像的比特率。

[0140] 例如,如果假设图像的质量标识符矢量值被称为图像质量度量(VQM),网络状态矢量值被称为网络状态(NS),并且目标VQM值(最大质量标准(例如,PSNR 45dB或更高))被设置,则最优分辨率和比特率可例如通过以下方法被确定:i)估计从第二AI模型输出的多个VQM中的可在当前网络状态下被传输的VQM集合(VQM'),ii)获得VQM'中的满足目标VQM值的集合,并且iii)如果不存在满足目标VQM值的集合(例如,网络条件差),则将目标VQM降低10%并重复步骤i)和ii)。如果存在满足目标VQM值的集合,则选择具有最低比特率的比特率和分辨率组合。

[0141] 根据另一示例,从第二AI模型420输出的图像的质量信息可另外包括以包括分辨率、比特率和质量标识符的形式的成对数据以及该数据的概率信息。

[0142] 例如,从第二AI模型420输出的质量信息可包括与不同的比特率信息和不同的分辨率信息相应的多个质量标识符,诸如,例如,但不限于(FHD,比特率\_0,VQ度量\_0,概率\_0)、(FHD,比特率\_1,VQ标识符\_1,概率\_1)、(FHD,比特率\_2,VQ标识符\_2,概率\_2)、(FHD,比特率\_3,VQ标识符\_3,概率\_3)、(2K,比特率\_4,VQ标识符\_4,概率\_4)、(2K,比特率\_5,VQ度量\_5,概率\_5)、(2K,比特率\_6,VQ度量\_6,概率\_6)、(2K,比特率\_7,VQ度量\_7,概率\_07)、(4K,比特率\_8,VQ度量\_8,概率\_8)、(4K,比特率\_9,VQ度量\_9,概率\_9)、(4K,比特率\_10,VQ度量\_10,概率\_10)、和(4K,比特率\_11,VQ度量\_11,概率\_11)以及与每个质量标识符相应的概率值。在此,1至N是用于标识比特率和VQ度量的数字,并且比特率、VQ度量和概率可用各种数字来表示。

[0143] 概率值可指例如可提供阈值范围的的质量的最小比特率的概率。例如,参照图7,如

果将电子设备100所需的阈值范围的质量设置为范围710,则特定比特率720可以是满足分辨率阈值范围的质量的最小比特率。然而,为了便于描述,假设分辨率为1。这样,概率值可指示特定数据对与最小比特率相应的概率。如果即使压缩到比特率A和比特率B( $A < B$ ),电子设备100也提供相似的质量和临界范围质量,则就压缩效率而言,压缩到小比特率A而不是大比特率B是有利的。因此,概率值可很大。

[0144] 在这种情况下,第二AI模型可被训练为不仅另外输出包括图像质量信息(诸如,分辨率、比特率和质量标识符)的格式的配对数据,而且还另外输出配对数据的概率信息。

[0145] 根据以上实施例,在给定的网络环境中,可有效地使用网络带宽,并且可使用户的感觉质量最大化。

[0146] 返回图2A和图2B,根据实施例,处理器110可基于从图像的当前帧预测的(或估计的)质量信息来控制对当前帧进行下缩放或编码中的至少一个。根据另一实施例,处理器110可基于从图像的至少一个先前帧预测的质量信息来控制对当前帧进行下缩放或编码中的至少一个。

[0147] 图8A和图8B是示出根据各种实施例的使用当前帧或先前帧的质量信息的示例的示图。

[0148] 图8A是根据实施例的由处理器110基于与当前帧相应的质量信息来控制对当前帧进行下缩放或编码中的至少一个的示例的示图。

[0149] 参照图8A,当将第一图像帧(当前图像帧)输入到第一AI模型810时,处理器110可将从第一AI模型810中包括的特定层输出的图像特征信息(例如,与第一图像帧(当前图像帧)相应的特征图数据)输入到第二AI模型820。特定层可以是属于原始分辨率域的层,例如,与执行插值处理之前的分辨率相应的层。处理器110可将从特定层输出的第一图像帧(当前图像帧)的特征信息(例如,特征图数据)存储在帧缓冲器80中,并且基于从第二AI模型820输出的质量信息对存储在帧缓冲器80中的特征图数据进行下缩放,以获得下缩放的帧。因此,下缩放的帧可以是一帧延迟(例如,基于60FPS大约16.7ms),并且被输入到图像编码器。然而,可提高预测准确性并且降低复杂性。

[0150] 图8B是示出根据另一实施例的处理器110基于与图像的先前帧相应的质量信息来控制对当前帧进行下缩放或编码中的至少一个的示例的示图。例如,处理器110可基于第N1、第N-2、...第N-k帧来预测第N帧的质量信息,并且基于此来控制对第N帧进行下缩放或编码中的至少一个。

[0151] 参照图8B,当第二图像帧(先前图像帧)被输入到第一AI模型810时,处理器110可将从第一AI模型810中包括的特定层输出的图像特征信息(例如,与第二图像帧(先前图像帧)相应的特征图数据)输入到第二AI模型820。特定层可以是在执行实质性下缩放(例如,双三次处理)之前属于图像的分辨率域的层。处理器110可将从特定层输出的第二图像帧(先前图像帧)的特征信息(即,特征图数据)存储在帧缓冲器81(例如,先进先出(FIFO))中,并且基于存储在帧缓冲器81中的特征图数据获得质量信息。当将第三图像帧(当前帧)输入到第一AI模型810时,处理器110可基于根据至少一个第二图像帧(先前图像帧)预测的质量信息对第三图像帧(当前帧)执行下缩放(例如,双三次)。因此,可将下缩放后的当前帧在无延迟的情况下输入到图像编码器。

[0152] 为此,如图8B中所示,第二AI模型830可包括至少一个循环神经网络(RNN)层,其

中,循环神经网络层被训练为基于与至少一个先前帧相应的特征信息预测当前帧的质量信息。这是因为,在图8B的实施例中,由于RNN是适合于对利用时间差和顺序地输入的数据进行分析并基于该分析进行未来预测的网络,因此应基于先前图像帧的质量信息预测当前图像帧的质量信息。

[0153] 在图8A和图8B中,在第二AI模型820的最后层中,将全连接网络(FCN)划分为三个FCN以划分并表示不同的分辨率(例如,图6的4K、2K、FHD),但是可将该FCN实现为一个FCN。

[0154] 在以上实施例中,已经说明了在最后的神经网络层的后端执行第一AI模型810中的双三次处理,但是可在中间神经网络层的后端执行双三次处理,在这种情况下,可在双三次处理之前从神经网络层输出的特征图数据提供给第二AI模型820、830。

[0155] 在以上实施例中,第一AI模型810、第二AI模型820和830中包括的插值处理(例如,双三次)和神经网络层可被实现为软件、硬件或它们的组合(例如,处理电路和/或可执行程序元件)。

[0156] 诸如上述方法,处理器110可基于编码图像产生数据包(或帧),并且发送产生的数据包。根据一个示例,处理器110可将包括编码图像的数据包(或帧)发送到接收设备20(图1A)。在该示例中,分辨率信息和比特率信息可被包括在数据包的报头区域中,并且编码图像可被包括在数据包的有效载荷区域中。根据实施例,可以以至少一个帧单位改变分辨率和比特率,并且可在数据包的报头中包括与至少一个帧单位相应的分辨率和比特率。

[0157] 图9是示出根据实施例的电子设备的示例配置的框图。

[0158] 图9中所示的电子设备可被实现为用于从图2A至图2B中所示的电子设备100和100'接收内容并输出内容的设备,并且可被称为例如与电子设备100和100'区别开的接收设备。

[0159] 参照图9,接收设备200包括存储器210、通信接口(例如,包括通信电路)220和处理器(例如,包括处理电路)230。

[0160] 接收设备200可被实现为例如电视(TV)或机顶盒,但不限于此,并且可被实现为具有显示功能的各种类型的装置(诸如,例如,但不限于智能电话、平板个人计算机(PC)、移动电话、电子书、台式计算机、膝上型计算机、上网本计算机、个人数字助理(PDA)、便携式多媒体播放器(PMP)、运动画面专家组阶段1或阶段2(MPEG-1或MPEG-2)音频层3(MP3)播放器、移动医疗装置、相机、摄像机、电子相框、可穿戴装置(例如,头戴式装置(HMD)、智能手表、电子服装、电子手镯、电子项链等)、近眼显示器(NED),大画幅显示器(LFD)、数字标牌、数字信息显示器(DID)/图像墙、投影仪显示器等),但不限于此,并且可以是使用AI模型处理数据的设备。

[0161] 根据示例,各种分辨率的图像和各种压缩图像可被接收。例如,接收设备200可接收SD、HD、FHD、4K UHD、8K UHD或更多中的至少一种分辨率的图像。电子设备100可接收具有诸如MPEG、JPEG、AVC、H.264、H.265、HEVC、VC-1、VP8、VP9、AV1(AOMedian图像1)等的压缩形式的图像。接收到的图像可以是来自电子设备100和100'被编码的图像。

[0162] 存储器210可存储各种实施例所需的数据。存储器210的实现格式与图2B的存储器120的格式相似,并且将不被进一步描述。

[0163] 根据示例,存储器210可存储至少一个指令或包括用于控制接收设备200的指令的计算机程序。

[0164] 根据另一示例,存储器210可存储关于包括多个层的AI模型的信息。存储关于AI模型的信息可指例如存储与AI模型的运算相关的各种信息,例如关于AI模型中包括的多个层的信息、关于在所述多个层中的每个层中使用的参数(例如,滤波器系数、偏差等)的信息等。例如,存储器210可存储关于根据示例实施例被训练为对图像进行上缩放的AI模型的信息。可根据以下描述的处理器230的实现方式将关于AI模型的信息存储在处理器230的内部存储器中。例如,如果处理器230被实现为专用硬件,则关于AI模型的信息可被存储在处理器230的内部存储器中。

[0165] 根据另一示例,存储器210可存储从外部装置、外部存储介质(例如,USB)、外部服务器(例如,网络盘)接收到的图像。图像可以是数字运动图像,但不限于此。

[0166] 根据另一示例,存储器210可存储上缩放所需的信息、算法、图像质量参数等、压缩所需的信息、用于执行例如降噪、细节增强、色调映射、对比度增强、色彩增强或帧速率转换中的至少一个的图像质量处理所需的各种信息。存储器210可存储通过图像处理产生的最终输出图像。

[0167] 在以上实施例中,尽管已经描述了各种数据被存储在处理器230的外部存储器210中,但是前述数据的至少一部分可根据接收设备200或处理器230的至少一个示例被存储在处理器230的内部存储器中。

[0168] 通信接口220可包括各种处理电路并且接收各种类型的内容。例如,通信接口220可通过经由通信方法(诸如,基于接入点(AP)的Wi-Fi(无线LAN网络)、蓝牙、Zigbee、有线/无线局域网(LAN)、广域网(WAN)、以太网、IEEE 1394、高清晰度多媒体接口(HDMI)、通用串行总线(USB)、移动高清链接(MHL)、高级加密标准(AES)/欧洲广播联盟(EBU)、光纤、同轴等)从外部装置(例如,源装置)、外部存储介质(例如,通用串行总线(USB)装置)、外部服务器(例如,网络服务器等)进行流传输或下载来接收图像信号。根据实施例,通信接口220可通过互联网通信网络从电子设备100、100'接收实时图像。

[0169] 处理器230可包括各种处理电路并被电连接到存储器210,并且可控制接收设备200的全部操作。处理器230可包括一个或多个处理器。处理器230可通过执行存储在存储器210中的至少一个指令来执行根据各种实施例的接收设备200的操作。处理器230的实现格式与图2A和图2B的处理器110的格式相同或相似,并且将不被进一步描述。

[0170] 处理器230可通过处理输入数据来获得输出数据。输入数据可包括文本、图像或用户语音中的至少一个。例如,可通过可与外部装置通信的通信器、用户输入器(诸如,键盘或触摸板、相机、麦克风等)输入输入数据。根据AI模型的类型,输出数据可具有各种格式。例如,输出数据可以是具有提高的分辨率的图像、与图像中包括的对象相关的信息、与语音相应的文本等。

[0171] 根据实施例,处理器230可对输入图像进行图像处理以获得输出图像。图像处理可以是包括图像增强、图像恢复、图像变换、图像分析、图像理解或图像压缩中的至少一个的数字图像处理。在一个示例中,当输入图像是压缩图像时,处理器230可对压缩图像进行解码并执行图像处理。

[0172] 根据实施例,当接收到输入图像(例如,在被电子设备100、100'下缩放之后被编码的图像)时,处理器230可通过对接收到的图像进行解码(或释放压缩)并对解码图像进行上缩放来获得输出图像。

[0173] 根据示例,处理器230可基于接收到的图像的压缩方法对接收到的图像进行解码。处理器230可使用AI模型对解码图像进行上缩放。

[0174] 根据实施例,由于可以以至少一个帧单位改变图像分辨率和图像比特率,因此处理器230需要知道发送的图像的分辨率和比特率。例如,当处理器230接收到数据包类型的输入图像时,处理器230可对数据包进行解包以获得报头区域中包括的分辨率信息和比特率信息,并且获得有效载荷区域中包括的编码图像。报头区域中包括的分辨率信息和比特率信息可以是与在电子设备100和100'中执行的下缩放相关联的分辨率信息以及与图像压缩相关联的比特率信息。

[0175] 处理器230可基于获得的比特率信息对被编码的图像执行解码,并且基于获得的分辨率信息对解码图像执行上缩放。

[0176] 图10A、图10B和图10C是更详细地示出根据实施例的使用人工智能模型的示例上缩放方法的示意图。

[0177] 根据实施例的用于上缩放的AI模型1010可被实现为用于超分辨率处理的学习网络模型。超分辨率可指例如通过一系列媒体处理将低分辨率图像转换成高分辨率图像的处理。

[0178] 根据示例,如图10A中所示,处理器230可使用包括多个神经网络层的AI模型1010对解码图像11进行上缩放。所述多个神经网络层中的每个神经网络层包括多个参数(或多个权重值),并且可通过先前层的结果与多个参数之间的运算来执行神经网络运算。可通过AI模型的学习结果优化所述多个神经网络层中包括的参数。例如,参数可被更新,使得由AI模型获得的损失值或代价值在学习处理期间被减小或最小化。人工神经网络可包括深度神经网络(DNN)(例如,但不限于卷积神经网络(CNN)、循环神经网络(RNN)、通用对抗网络(GAN)、受限玻尔兹曼机(RBM)、深度置信网络(DBN)、双向循环深度神经网络(BRDNN)、深度Q网络等),但不限于上述示例。

[0179] 根据另一示例,如图10B中所示,处理器230可对解码图像11执行插值处理1020,并且将插值图像12输入到AI模型1010以获得残差图像13。例如,AI模型1010可被实现为残差神经网络。残差神经网络的多个层中的每个层都可使用包括不同参数的滤波器产生针对插值图像的残差图像。然而,AI模型1010不必仅产生残差图像,并且可根据AI模型1010的实施例以各种方式处理输入图像,并且输出处理后的图像。在这种情况下,处理器230可将插值处理后的图像12与残差图像13组合以获得输出图像14,例如,高分辨率图像。插值处理可包括例如将具有低分辨率的图像缩放为高分辨率的处理,并且至少一种插值技术(例如,但不限于双线性插值、最近邻插值、双三次插值、反卷积插值、子像素卷积插值、多相插值、三线性插值、线性插值等)可被使用。残差图像可指仅包括残差信息的图像。残差信息可包括但不限于边缘方向、边缘强度、噪声信息或纹理信息中的至少一个作为根据输入图像与参考图像之间的差的信息。根据另一示例,残差信息可包括灰度信息、亮度信息或伽马信息中的至少一个。

[0180] 根据另一示例,如图10C中所示,处理器230可将解码图像11输入到AI模型1010以获得残差图像13',并且对残差图像13'进行插值1030以获得插值残差图像13"。处理器230还可对解码图像11执行插值处理1020以获得插值图像12。然后处理器230可将插值图像12与插值残差图像13"组合以获得输出图像15,例如,高分辨率图像。根据图10C所示的实施例

例,与图10B所示的实施例不同,可将解码图像11输入到AI模型1010以获得残差图像13'。

[0181] 然而,根据另一实施例,还可包括除AI模型1010之外的AI模型。在该示例中,AI模型1010和另一AI模型可顺序地进行运算或者并行地进行运算。作为示例,处理器230可将解码图像11输入到AI模型1010,将AI模型1010的输出输入到另一AI模型,然后基于从另一AI模型输出的图像来获得输出图像。作为另一示例,处理器230可将解码图像11输入到AI模型1010和另一AI模型中的每一个,并且基于从AI模型1010和另一AI模型并行输出的多个图像来获得输出图像。例如,AI模型1010可以是产生第一残差图像的模式,并且另一AI模型可包括产生第二残差图像的模式。AI模型1010可指例如用于对分辨率进行上缩放的模型,并且另一AI模型可以是用于上述各种图像处理(例如,降噪)中的一个图像处理的模型。AI模型1010可以是用于对象区域处理的模型,并且另一AI模型可以是用于背景区域处理的模型。

[0182] 图11A和图11B是示出根据实施例的电子设备和接收设备的示例的框图。

[0183] 图11A示出根据实施例的电子设备的示例。根据图11A,电子设备100"包括处理器(例如,包括处理电路)110、存储器120和通信接口(例如,包括通信电路)130。处理器110和存储器120与图2A和2B的配置重复,并且将不被进一步描述。

[0184] 通信接口130可包括各种通信电路并且与外部装置通信。例如,通信接口130可通过经由通信方法(诸如,例如,但不限于基于接入点(AP)的Wi-Fi(无线LAN网络)、蓝牙、Zigbee、有线/无线局域网(LAN)、广域网(WAN)、以太网、IEEE 1394、高清晰度多媒体接口(HDMI)、通用串行总线(USB)、移动高清链接(MHL)、高级加密标准(AES)/欧洲广播联盟(EBU)、光学、同轴等)从外部装置(例如,源装置)、外部存储介质(例如,通用串行总线(USB)装置)、外部服务器(例如,网络服务器等)进行流传输或下载来接收图像信号。根据实施例,通信接口130可通过互联网通信网络向图9中所示的接收设备200发送实时图像。

[0185] 图11B示出根据实施例的接收设备的示例。参照图11B,接收设备200'包括存储器210、通信接口(例如,包括通信电路)220、处理器(例如,包括处理电路)230、显示器240、输出器(例如,包括输出电路)250和用户接口260。存储器210、通信接口220和处理器230与图9的配置重复,因此将不被进一步描述。

[0186] 显示器240可被实现为包括自发光元件的显示器或包括非自发光元件和背光的显示器。例如,显示器240可被实现为各种类型的显示器,诸如,例如,但不限于液晶显示器(LCD)、有机发光二极管(OLED)显示器、发光二极管(LED)、微型LED、迷你LED、等离子显示面板(PDP)、量子点(QD)显示器、量子点发光二极管(QLED)等。在显示器240中,还可包括可被实现为a-si TFT、低温多晶硅(LTPS)TFT、有机TFT(OTFT)的背光单元、驱动电路等。显示器140可被实现为被耦合到触摸传感器的触摸屏、柔性显示器、可弯曲显示器、三维(3D)显示器、物理地连接多个显示模块的显示器等。处理器230可控制显示器240输出根据各种实施例获得的输出图像。这里,输出图像可以是4K、8K或更高的高分辨率图像。

[0187] 输出器250可包括各种输出电路并且输出声音信号。例如,输出器250可将由处理器230处理的数字声音信号转换成模拟声音信号,放大并输出该模拟声音信号。例如,输出器250可包括各种输出电路,诸如,例如,但不限于能够输出至少一个声道的至少一个扬声器单元、D/A转换器、音频放大器等。根据示例,输出器250可被实现为输出各种多声道声音信号。处理器230可根据输入图像的增强处理来控制输出器250处理输入的声音信号。例如,处理器230可将输入的两声道声音信号转换为虚拟多声道(例如,5.1声道)声音信号,识别

接收设备200'所处的位置以将该信号处理为对空间优化的立体声音信号,或者根据输入图像的类型(例如,内容类型)提供优化的声音信号。

[0188] 用户接口260可包括各种用户接口电路,并且可被实现为装置,诸如,例如,但不限于按钮、触摸板、鼠标和键盘、触摸屏、能够执行上述显示功能和操作输入功能的远程控制收发器等。远程控制收发器可经由至少一种通信方法(诸如,红外线通信、蓝牙通信或Wi-Fi通信)从外部远程控制器接收远程控制信号,或者发送远程控制信号。

[0189] 根据实施例,接收设备200'还可包括调谐器和解调器。调谐器(未示出)可通过调谐经由天线接收的射频(RF)广播信号中的用户选择的信道或所有预先存储的信道来接收RF广播信号。解调器(未示出)可接收和解调由调谐器转换的数字中频(IF)信号和IF(DIF)信号,并且执行信道解码等。根据示例实施例的经由调谐器接收到的输入图像可经由解调器(未示出)被处理,然后被提供给处理器230以进行根据示例实施例的图像处理。

[0190] 图12是示出根据实施例的示例发送和接收系统的示例操作的示图。

[0191] 在图12中,虚线的左上侧可表示发送设备(例如,电子设备100、100'、100'')的操作,并且右侧可表示接收设备(例如,接收设备200和200')的操作。为了便于描述,假设左侧的操作由电子设备100执行并且右侧的操作由接收设备200执行。

[0192] 电子设备100可包括下缩放器(例如,包括处理电路和/或可执行程序元件)1210、视频编码器(例如,包括处理电路和/或可执行程序元件)1220、发送器1230、视频质量估计器(例如,包括处理电路和/或可执行程序元件)1240和视频质量控制器(例如,包括处理电路和/或可执行程序元件)1250,并且接收设备200可包括接收器1260、视频解码器(例如,包括处理电路和/或可执行程序元件)1270和上缩放器(例如,包括处理电路和/或可执行程序元件)1280。每个结构1210至1280可用IP实现以用于各种处理。IP可指例如可重用的功能块,并且可被实现为至少一个硬件或软件功能块或者至少一个软件和至少一个硬件的组合。电子设备100中的每个组件1210-1250可被实现在至少一个处理芯片内,并且接收设备200中的每个组件1260-1280也可被实现在至少一个处理芯片内。

[0193] 下缩放器1210可包括各种处理电路和/或可执行程序元件,并且使用第一AI模型410、810对图像帧进行下缩放。图像(例如,视频)编码器1220可包括各种处理电路和/或可执行程序元件,并且对下缩放的图像帧进行编码。发送器1230可包括各种发送电路,并且将编码的图像帧发送到接收设备200。

[0194] 在该示例中,图像质量估计器1240可包括各种处理电路和/或可执行程序元件,并且基于在下缩放器1210的下缩放处理中获得的图像帧的特征信息来估计图像帧的质量。图像帧的特征信息可以是例如从第一AI模型410或810中包括的特定层输出的特征图信息。

[0195] 例如,可利用第二AI模型420、820来实现图像质量估计器1240。第二AI模型420、820可包括第一AI模型410、810包括的至少一些层、至少一些层1241、以及用于将图像帧的特征信息转换为质量信息的附加层1242。在一些情况下,如图8B中所示,还可包括用于预测图像的时间变化的RNN层1243。从第二AI模型420、820输出的当前图像帧(或先前图像帧)的质量信息可包括与多个分辨率和比特率集合相应的多个质量标识符(VQ\_度量)。例如,如图13中所示,质量信息可以是与每个帧索引相应的质量信息,例如,分辨率和比特率。

[0196] 图像(例如,视频)质量控制器1250可包括各种处理电路和/或可执行程序元件,并且考虑从第二AI模型420、820输出的多个质量标识符中的电子设备100的目标质量索引、网

络状态等来选择一个质量标识符,并且将与选择的质量标识符相应的分辨率信息和比特率信息分别提供给下缩放器1210和图像编码器1220。可基于当前图像帧获得的提供的分辨率信息和比特率信息,或者提供的分辨率信息和比特率信息可以是基于先前图像帧预测的信息。

[0197] 下缩放器1210和图像编码器1220可分别基于接收到的分辨率信息和比特率信息执行下缩放和编码。

[0198] 发送器1230可包括各种电路,并且对从图像编码器1220输出的编码图像进行打包,并将图像发送到接收设备200。

[0199] 当从电子设备100接收到数据包时,接收设备200中包括的接收器1260可将接收到的数据包拆包为压缩图像格式。图像(例如,视频)解码器1270可包括各种处理电路和/或可执行程序元件,并且对从接收器1260输出的图像数据进行解码(或解压缩)。上缩放器1280可包括各种处理电路和/或可执行程序元件,并且将解码图像数据上缩放到输出分辨率以获得输出图像。可通过显示器显示获得的输出图像。接收到的数据包的报头可包括与下缩放器1210的下缩放相关联的分辨率信息和与图像编码器1220的编码相关联的比特率信息。图像解码器1270可基于报头中包括的比特率信息执行解码,并且上缩放器1280可基于报头中包括的分辨率信息执行上缩放。

[0200] 在图12所示的实施例中,可使用从第二AI模型420和820输出的质量信息与实际图像的质量信息之间的差作为损失函数,利用监督学习训练第二AI模型420、820的学习。根据示例,可通过将原始训练图像与经由下缩放器1210、图像编码器1220、发送器1230、接收器1260、图像解码器1270和上缩放器1280重建的图像进行比较来获得实际图像的质量信息。

[0201] 图14是示出根据实施例的用于控制电子设备的示例方法的流程图。

[0202] 根据图14的电子设备的控制方法,在操作S1410,电子设备100、100'和100"使用训练的第一AI模型对图像进行下缩放。在操作S1420,电子设备100、100'和100"对下缩放图像进行编码。

[0203] 在操作S1430,可基于使用训练的第二AI模型获得的图像的质量信息来控制使用第一AI模型对图像进行下缩放。可使用从第一AI模型获得的图像的特征信息训练第二AI模型。

[0204] 在操作S1420,可控制基于使用第二AI模型获得的图像的质量信息对下缩放图像进行编码。

[0205] 图像的质量信息可包括和与图像相关联的分辨率信息和比特率信息不同组合相应的多个质量标识符。在操作S1410,可基于多个质量标识符中的任意一个的质量标识符来确定图像的下缩放比率,并且可基于确定的下缩放比率控制对图像进行下缩放。在操作S1420,可基于多个质量标识符中的任意一个的质量标识符来确定下缩放图像的比特率,并且可基于确定的比特率控制对下缩放图像进行编码。

[0206] 此外,控制电子设备100、100'、100"的方法还可包括获得网络的状态信息。在这种情况下,在操作S1410,可基于网络的状态信息和图像的质量信息控制对图像进行下缩放。在操作S1420,可基于网络的状态信息和图像的质量信息控制对下缩放图像进行编码。

[0207] 控制电子设备100、100'、100"的方法还可包括:基于网络的状态信息确定多个质量标识符中的至少一个质量标识符;并且基于目标质量标识符确定所述至少一个质量标识

符中的质量标识符。在操作S1410,可基于与确定的质量标识符相应的分辨率信息控制对图像进行下缩放。此外,在操作S1420,可基于与确定的质量标识符相应的比特率信息控制对下缩放图像进行编码。

[0208] 第二AI模型可包括例如至少一个神经网络层,并且至少一个神经网络层的参数可与第一AI模型中包括的至少一些层的参数相关联地被训练。

[0209] 在操作S1410,可基于根据图像的当前帧获得的质量信息来控制对当前帧进行下缩放,并且可基于根据图像的至少一个先前帧预测的质量信息来控制对当前帧进行下缩放。

[0210] 第二AI模型可包括例如至少一个RNN层,其中,该至少一个RNN层被训练为基于与至少一个先前帧相应的特征信息预测当前帧的质量信息。

[0211] 可利用通过将和与训练图像相关的多个分辨率信息和多个比特率信息不同组合相应的多个实际质量标识符输入到第二AI模型而输出的多个质量标识符之间的差作为损失函数,利用监督学习来训练第二AI模型。在图6中已经描述了实际的质量标识符,并且将不重复详细描述。

[0212] 控制电子设备100、100'和100"的方法还可包括基于编码图像产生数据包并发送产生的数据包。图像的分辨率信息和比特率信息可被包括在数据包的报头区域中。

[0213] 根据上述各种实施例,可基于实时预测的图像质量和网络状态来控制最优分辨率和比特率。因此,可有效地利用网络带宽,并且可使给定网络环境中的消费者的感觉质量最大化。特别地,通过以最优方式控制具有延迟限制的实时现场流传输中的图像的分辨率和比特率,可有效地使用网络带宽。

[0214] 通过共享在下缩放处理期间获得的图像的特征信息来预测图像质量,可对AI模型进行有效训练,并且最小化和/或降低H/W实现复杂度。

[0215] 将参照附图描述根据实施例的AI编码处理和AI解码处理。

[0216] 图15是示出根据实施例的示例AI编码处理和示例AI解码处理的示图。

[0217] 如上所述,当图像的分辨率增加时,用于编码/解码的信息处理量增加,并且需要用于增强图像的编码效率和解码效率的方法。

[0218] 如图15所示,根据实施例,可通过对具有高分辨率的图像1505进行AI下缩放1510来获得第一图像1515。通过对具有相对较小分辨率的第一图像1515执行第一编码1520和第一解码1530,与对图像1505执行第一编码1520和第一解码1530相比,可显著降低将被处理的比特率。

[0219] 参照图15,在实施例中,在AI编码处理中,通过对图像1505进行AI下缩放1510来获得第一图像1515,并且对第一图像1515进行第一编码1520。在AI解码处理中,接收包括作为AI编码的结果而获得的AI数据和包括图像数据的AI编码数据,通过第一解码1530获得第二图像1535,并通过第二图像1535进行AI上缩放1540获得第三图像1545。

[0220] 更详细地参照AI编码处理,在接收到图像1505时,对图像1505进行AI下缩放1510,以获得预定分辨率的第一图像1515或预定图像。基于AI执行AI下缩放1510,其中,由于如果用于AI下缩放1510的AI和用于AI上缩放1540的AI被单独训练,则作为AI编码的目标的图像1505与通过AI解码重建的第三图像1545之间的差异增大,所以用于AI下缩放1510的AI应与用于第二图像1535的AI上缩放1540的AI被共同训练。

[0221] 在实施例中,为了在AI编码处理和AI解码处理中保持共同关系,可使用AI数据。因此,通过AI编码处理获得的AI数据需要包括指示上缩放目标的信息,并且在AI解码处理中,需要根据基于AI数据识别出的上缩放目标对第二图像1535进行AI上缩放1540。

[0222] 用于AI下缩放1510的AI和用于AI上缩放1540的AI可被实现为例如但不限于DNN。如以下参照图23所述,由于通过在预定目标下共享损失信息来共同训练第一DNN和第二DNN,因此AI编码设备将当共同训练第一DNN和第二DNN时使用的目标信息提供给AI解码设备,并且AI解码设备可基于接收到的目标信息利用目标分辨率对第二图像1535进行AI上缩放1540。

[0223] 当更详细地描述第一编码1520和第一解码1530(例如,图1A中所示的发送端10和接收端20)时,从图像1505AI下缩放的第一图像1515可经由第一编码1520而具有减少的信息量。第一编码1520可包括:通过对第一图像1515进行预测产生预测数据,产生与第一图像1515和预测数据之间的差相应的残差数据,将作为空间域分量的残差数据变换为频域分量,对变换为频域分量的残差数据进行量化,对量化后的残差数据进行熵编码等。可通过使用频率变换的图像压缩方法(诸如,例如,但不限于MPEG-2、H.264AVC、MPEG-4、HEVC、VC-1、VP8、VP9和AV1等)中的一个图像压缩方法来实现第一编码处理1520。

[0224] 可通过对图像数据进行第一解码1530来重建与第一图像1515相应的第二图像1535。第一解码1530可包括:对图像数据进行熵解码以产生量化的残差数据;对量化的残差数据进行反量化;将频域分量的残差数据变换为空间域分量;产生预测数据;并且使用预测数据和残差数据重建第二图像1535等。可通过与在第一编码1520中使用的使用频率变换的图像压缩方法(诸如,例如,但不限于MPEG-2、H.264、MPEG-4、HEVC、VC-1、VP8、VP9、AV1等)中的一个图像压缩方法相应的图像重建方法来实现第一解码1530。

[0225] 通过AI编码处理获得的AI编码数据可包括作为对第一图像1515进行第一编码1520的结果而获得的图像数据以及与对图像1505进行AI下缩放1515相关的AI数据。

[0226] 可以以比特流的形式发送图像数据。图像数据可包括基于第一图像1515中的像素值获得的数据,例如,作为第一图像1515与第一图像1515的预测数据之间的差的残差数据。此外,图像数据包括在对第一图像1515进行第一编码1520中使用的信息。例如,图像数据可包括用于对第一图像1515进行第一编码1520的预测模式信息和运动信息、第一编码1520中使用的量化参数相关信息等。可根据第一编码1520中使用的图像压缩方法的规则(例如,使用频率变换的图像压缩方法(诸如,例如,但不限于MPEG-2、AVC、H.264AVC、MPEG-4、HEVC、VC-1、VP8、VP9、AV1等)的语法)来产生图像数据。

[0227] 可基于第二DNN将AI数据用于AI上缩放1540。如上所述,第一DNN和第二DNN被共同训练,并且AI数据包括使通过第二DNN对第二图像1535进行AI上缩放1540能够被执行的信息。在AI解码处理中,可基于AI数据将第二图像1535AI上缩放1540为目标分辨率和/或质量。

[0228] AI数据可与图像数据一起以比特流格式被发送。可选地,根据实施例,可以以与图像数据分离的帧或数据包发送AI数据。可通过相同的网络或不同的网络发送作为AI编码的结果而获得的图像数据和AI数据。

[0229] 图16是示出根据实施例的示例AI解码设备1600的示例配置的框图。

[0230] 参照图16,根据实施例的AI解码设备1600可包括接收器1610和AI解码器1630。接

收器1610可包括通信器(例如,包括通信电路)1612、解析器(例如,包括处理电路和/或可执行程序元件)1614和输出器(例如,包括输出电路)1616。AI解码器1630可包括第一解码器(例如,包括处理电路和/或可执行程序元件)1632和AI上缩放器(例如,包括处理电路和/或可执行程序元件)1634。

[0231] 接收器1610可包括各种接收电路,并且接收并解析作为AI编码的结果而获得的AI编码数据,划分图像数据和AI数据,并将数据输出到AI解码器1630。

[0232] 例如,通信器1612可包括各种通信电路,并且通过网络接收作为AI编码的结果而获得的AI编码数据。作为AI编码的结果而获得的AI编码数据包括图像数据和AI数据。可通过相同的网络或不同的网络接收图像数据和AI数据。

[0233] 解析器1614可包括各种处理电路和/或可执行程序元件,并且对由通信器1612接收的AI编码数据进行解析,并将该数据划分为图像数据和AI数据。例如,通过读取从通信器1612获得的数据的报头来划分数据是图像数据还是AI数据。在一个示例中,解析器1614通过经由通信器1612接收到的数据的报头划分图像数据和AI数据,将划分的图像数据和AI数据传递到输出器1616,并且输出器1616(可包括各种输出电路)将各个划分的数据传递到第一解码器1632和AI上缩放器1634。可确定AI编码数据中包括的图像数据是通过预定编解码器(例如,MPEG-2、H.264、MPEG-4、HEVC、VC-1、VP8、VP9或AV1)获得的图像数据。在该示例中,可通过输出器1616将相应的信息发送到第一解码器1632,从而可利用识别出的编解码器处理图像数据。

[0234] 在实施例中,可从包括磁性介质(诸如,硬盘、软光盘、磁带)、光学可记录介质(诸如,CD-ROM和DVD)、磁光介质(诸如,光磁软盘)等的数据存储介质获得通过解析器1614解析的AI编码数据。

[0235] 第一解码器1632可包括各种处理电路和/或可执行程序元件,并且基于图像数据重建与第一图像1515相应的第二图像1535。将通过第一解码器1632获得的第二图像1535提供给AI上缩放器1634。根据实施例,还可将图像数据中包括的第一解码相关信息(诸如,预测模式信息、运动信息、量化参数信息等)提供给AI上缩放器1634。

[0236] 接收AI数据的AI上缩放器1634可包括各种处理电路和/或可执行程序元件,并且基于AI数据对第二图像1535进行AI上缩放。根据实施例,可通过进一步使用与第一编码相关的信息(诸如,图像数据中包括的预测模式信息和量化参数信息等)执行AI上缩放。

[0237] 根据实施例的接收器1610和AI解码器1630被描述为单独的设备,但是可通过单个处理器被实现。在这种情况下,可将接收器1610和AI解码器1630实现为专用处理器,并且可通过通用处理器(诸如,例如,但不限于应用处理器(AP)、中央处理器(CPU)、图形处理器(GPU)和S/W)的组合实现接收器1610和AI解码器1630。此外,在专用处理器的情况下,可通过包括用于实现实施例的存储器或用于使用外部存储器的存储器处理器来实现专用处理器。

[0238] 接收器1610和AI解码器1630可包括多个处理器,可被实现为专用处理器的组合,或者可通过多个通用处理器(诸如,AP、CPU、GPU和S/W)之间的组合被实现。类似地,可将AI上缩放器1634和第一解码器1632实现为不同的处理器。

[0239] 提供给AI上缩放器1634的AI数据包括用于对第二图像1535进行AI上缩放的信息。上缩放目标应与第一DNN的下缩放相应。因此,AI数据需要包括用于识别第一DNN的下缩放

目标的信息。

[0240] 作为AI数据中包括的信息的示例,存在关于图像1505的分辨率与第一图像1515的分辨率之间的差的信息以及与第一图像1515相关的信息。

[0241] 差信息可被表示为例如关于第一图像1515相对于图像1505的分辨率变换的程度的信息(例如,分辨率变换率信息)。由于第一图像1515的分辨率通过重建的第二图像1535的分辨率而可知,并且可通过已知的分辨率确认分辨率变换的程度,因此可仅由图像1505的分辨率信息表示差信息。分辨率信息可被表示为水平/垂直方向上的屏幕尺寸,并且可被表示为比率(16:9、4:3等)和一个轴上的尺寸。此外,可以以索引或标志的形式表示预定分辨率信息。

[0242] 与第一图像1515相关的信息可包括关于作为对第一图像1515进行第一编码的结果而获得的图像数据的比特率和用于对第一图像1515进行第一编码的编解码器类型中的至少一个的信息。

[0243] AI上缩放器1634可基于AI数据中包括的差信息和与第一图像1515相关的信息中的至少一个来确定第二图像1535的上缩放目标。上缩放目标可指示例如是否将第二图像1535上缩放到哪个分辨率。一旦已经确定了上缩放目标,AI上缩放器1634就经由第二DNN对第二图像1535进行AI上缩放以获得与上缩放目标相应的第三图像1545。

[0244] 在描述通过AI上缩放器1634对第二图像1535进行AI上缩放以与上缩放目标相应的方法之前,将参照图3和图4描述通过第二DNN进行AI上缩放的处理。

[0245] 图17是示出用于对第二图像1535进行AI上缩放的第二DNN 300的示图,并且图18示出通过图17中示出的第一卷积层310进行的示例卷积运算。

[0246] 参照图17,第二图像1535被输入到第一卷积层310。图17中所示的第一卷积层310中所示的 $3 \times 3 \times 4$ 示出了使用尺寸为 $3 \times 3$ 的四个滤波器核对一个输入图像执行卷积处理。作为卷积处理的结果,通过四个滤波器核产生四个特征图。每个特征图表示第二图像1535的唯一特征。例如,每个特征图可指示第二图像1535的垂直方向特征、水平方向特征或边缘特征。

[0247] 参照图18,将更详细地描述第一卷积层310中的卷积操作。

[0248] 可通过在第一卷积层310中使用的尺寸为 $3 \times 3$ 的滤波器核430的参数和与其相应的第二图像1535中的像素值之间的乘法运算和加法运算来产生一个特征图450。由于在第一卷积层310中使用了四个滤波器核,因此可使用四个滤波器核通过卷积运算处理来产生四个特征图。

[0249] 在图18中,在第二图像1535中显示的I1至I49表示第二图像1535的像素,并且在滤波器核430中显示的F1至F9表示滤波器核430的参数。在特征图450中显示的M1至M9表示特征图450的样本。

[0250] 在图18中,第二图像1535包括49个像素,但这仅是说明性的非限制性示例,并且当第二图像1535的分辨率为4K时,例如,可包括 $3840 \times 2160$ 个像素。

[0251] 在卷积运算中,可分别执行第二图像1535的I1、I2、I3、I8、I9、I10、I15、I16、I17的像素值中的每个像素值与滤波器核430的F1、F2、F3、F4、F5、F6、F7、F8和F9中的每一个之间的乘法运算,并且可将乘法运算的结果值进行组合(例如,加法运算)的值分配为特征图450的M1的值。如果卷积运算的步幅为2,则可执行第二图像1535的I3、I4、I5、I10、I11、I12、

I17、I18和I19的像素值中的每个像素值与滤波器核430的F1、F2、F3、F4、F5、F6、F7、F8和F9中的每一个的乘法运算,并且可将对乘法的结果值进行组合的值分配给特征图450的M2的值。

[0252] 当滤波器核430沿着步幅移动直到到达第二图像1535的最后一个像素时,可在第二图像1535的像素值与滤波器核430的参数之间执行卷积运算,并且可获得具有预定尺寸的特征图450。

[0253] 根据实施例,通过对第一DNN和第二DNN进行共同训练,可优化第二DNN的参数,例如,第二DNN的卷积层中使用的滤波器核的参数(例如,滤波器核430的F1、F2、F3、F4、F5、F6、F7、F8和F9的值)。如上所述,AI上缩放器1634可基于AI数据确定与第一DNN的下缩放目标相应的上缩放目标,并且将与确定的上缩放目标相应的参数确定为在第二DNN的卷积层中使用的滤波器核的参数。

[0254] 第一DNN和第二DNN中包括的卷积层可根据参照图18描述的卷积运算处理来执行处理,但是图18的卷积运算处理仅是示例,并且不限于此。

[0255] 参照图17,从第一卷积层310输出的特征图被输入到第一激活层320。第一激活层320可对每个特征图分配非线性特征。第一激活层320可包括但不限于S形函数、Tanh数学函数、ReLU函数等,但不限于此。

[0256] 在第一激活层320中分配非线性特征表示改变特征图的一些样本值并输出改变后的一些样本值。通过应用非线性特征执行所述改变。

[0257] 第一激活层320确定是否将从第一卷积层310输出的特征图的样本值传递到第二卷积层330。例如,特征图的样本值中的一些样本值被第一激活层320激活并被传递到第二卷积层330,并且一些样本值被第一激活层320停用且不被传递到第二卷积层330。通过第一激活层320突出由特征图表示的第二图像1535的固有特征。

[0258] 从第一激活层320输出的特征图325被输入到第二卷积层330。图17中示出的特征图325中的一个特征图是通过相对于图18在第一激活层320中处理特征图450而获得的结果。

[0259] 第二卷积层330中指示的 $3 \times 3 \times 4$ 示出了使用尺寸为 $3 \times 3$ 的四个滤波器核对输入特征图325进行卷积处理。第二卷积层330的输出被输入到第二激活层340。第二激活层340可将非线性特征发放给输入数据。

[0260] 从第二激活层340输出的特征图345被输入到第三卷积层350。图17中所示的第三卷积层350中所示的 $3 \times 3 \times 1$ 示出了使用尺寸为 $3 \times 3$ 的一个滤波器核执行卷积处理以产生一个输出图像。第三卷积层350并作为用于输出最终图像的层使用一个滤波器核产生一个输出。根据示例,第三卷积层350可经由卷积运算结果输出第三图像1545。

[0261] 如下所述,指示第二DNN 300的第一卷积层310、第二卷积层330和第三卷积层350中的滤波器核的数量、滤波器核的参数等的DNN配置信息可以是多个,但是多个DNN配置信息应与第一DNN的多个DNN配置信息相关联。可通过共同训练第一DNN和第二DNN来实现第二DNN的多个DNN配置信息与第一DNN的多个DNN配置信息之间的关联。

[0262] 尽管图17示出第二DNN 300包括三个卷积层310、330、350和两个激活层320、340,但这仅是一个示例,并且根据实施例,卷积层和激活层的数量可以以广泛的方式变化。此外,根据实施例,可经由RNN实现第二DNN 300。在该示例中,根据实施例,将第二DNN 300的

CNN结构改变为RNN结构。

[0263] 在示例实施例中, AI上缩放器1634可包括用于上述卷积运算和激活层的运算的至少一个算术逻辑单元 (ALU)。可利用处理器实现ALU算法。对于卷积运算, ALU可包括乘法器和加法器, 其中, 乘法器在滤波器核的样本值与从第二图像1535或先前层输出的特征图的样本值之间执行乘法运算, 加法器将乘法的结果值相加。对于激活层的运算, ALU可包括乘法器和比较器, 其中, 乘法器用于将输入样本值乘以在预定的S形函数、Tanh函数或ReLU函数中使用的权重, 比较器用于将乘法的结果与预定值进行比较以确定是否将输入样本值传递到下一层。

[0264] 在下文中, 将描述通过AI上缩放器1634对第二图像1535进行AI上缩放以与上缩放目标相应的方法。

[0265] 在实施例中, AI上缩放器1634可存储可在第二DNN中设置的多个DNN配置信息。

[0266] DNN配置信息可包括关于第二DNN中包括的卷积层的数量、针对每个卷积层的滤波器核的数量以及每个滤波器核的参数中的至少一个的信息。多个DNN配置信息可分别与各种上缩放目标相应, 并且第二DNN可基于与特定上缩放目标相应的DNN配置信息进行运算。根据DNN配置信息, 第二DNN可具有不同的结构。例如, 第二DNN可根据任一DNN配置信息包括三个卷积层, 并且第二DNN可根据其它DNN配置信息包括四个卷积层。

[0267] 在实施例中, DNN配置信息可仅包括用于第二DNN的滤波器核的参数。第二DNN的结构可不改变, 而是仅内部滤波器核的参数可根据DNN配置信息改变。

[0268] AI上缩放器1634可获得多个DNN配置信息中的用于对第二图像1535进行AI上缩放的DNN配置信息。多个DNN配置信息中的每个DNN配置信息用于获得预定分辨率和/或预定质量的第三图像1545, 并且与第一DNN被共同训练。

[0269] 例如, 多个DNN配置信息中的任意一个DNN配置信息可包括用于获得分辨率为2K (2048\*1080) 的第二图像1535的两倍的4K (4096\*2160) 的第三图像1545的信息, 并且另一DNN配置信息可包括用于获得分辨率为2K (2048\*1080) 的第二图像1535的四倍的8K (8192\*4320) 的第三图像1545的信息。

[0270] 多个DNN配置信息中的每个DNN配置信息与AI编码设备600的第一DNN的DNN配置信息相关联地被产生, 并且AI上缩放器1634根据与第一DNN的DNN配置信息的缩小比率相应的放大比率获得在多个DNN配置信息中的一个DNN配置信息。为此, AI上缩放器1634应检查第一DNN的信息。为了AI上缩放器1634检查第一DNN的信息, 根据实施例的AI解码设备1600从AI编码设备600接收包括第一DNN的信息的AI数据。

[0271] 换句话说, AI上缩放器1634可使用从AI编码设备600接收到的信息识别由用于获得第一图像1515的第一DNN的DNN配置信息所标靶的信息, 并且获得被共同训练的第二DNN的DNN配置信息。

[0272] 当在多个DNN配置信息中获得DNN配置信息或对第二图像1535进行AI上缩放时, 可基于根据获得的DNN配置信息进行运算的第二DNN处理输入数据。

[0273] 例如, 如果获得一个DNN配置信息, 则对于图17中所示的第二DNN 300的第一卷积层310、第二卷积层330和第三卷积层350中的每一个, 可将每个层中包括的滤波器核的数量和滤波器核的参数设置为获得的DNN配置信息中包括的值。

[0274] 例如, 如果将图17中所示的第二DNN的卷积层中的任意一个卷积层中使用的 $3 \times 3$

滤波器核的参数设置为{1,1,1,1,1,1,1,1,1},然后改变DNN配置信息,则可用改变后的DNN配置信息中包括的{2,2,2,2,2,2,2,2,2}替换所述参数。

[0275] AI上缩放器1634可基于AI数据中包括的信息获得多个DNN配置信息中的对第二图像1535进行上缩放的DNN配置信息。将更详细地描述用于获得DNN配置信息的AI数据。

[0276] 在实施例中,AI上缩放器1634可基于AI数据中的包括差信息获得多个DNN配置信息的用于对第二图像1535进行上缩放的DNN配置信息。例如,基于差信息,如果识别出图像1505的分辨率(例如,4K(4096\*2160))是第一图像1515的分辨率(例如,2K(2048\*1080))的两倍,则AI上缩放器1634可获取可将第二图像1535的分辨率提高到两倍的DNN配置信息。

[0277] 在另一实施例中,AI上缩放器1634可基于与AI数据中包括的与第一图像1515相关的信息,获得多个DNN配置信息中的用于对第二图像1535进行AI上缩放的DNN配置信息。AI上缩放器1634可预先确定图像相关信息与DNN配置信息之间的映射关系,并且获得被映射到与第一图像1515相关的信息的DNN配置信息。

[0278] 图19是示出各种图像相关信息与DNN配置信息之间的示例映射关系的示图。

[0279] 参照图19的示例,可看出,AI编码/AI解码处理不仅仅考虑分辨率的改变。如图19中所示,可通过分别考虑分辨率(诸如,SD、HD、全HD)、比特率(诸如,10Mbps、15Mbps、20Mbps)和编解码器信息(诸如,AV1、H.264、HEVC)来选择DNN配置信息。为此,必须与编码和解码处理共同执行考虑AI训练处理中的元件中的每个元件的训练(见图23)。

[0280] 根据训练的内容,当如图19中所示基于包括编解码器类型、图像的分辨率等的图像相关信息提供多个DNN配置信息时,可基于在AI解码处理中接收到的关于第一图像1515的信息获得用于对第二图像1535进行AI上缩放的DNN配置信息。

[0281] AI上缩放器1634将图19的左侧中示出的图像相关信息和右侧的DNN配置信息进行匹配,并且可使用根据图像相关信息的DNN配置信息。

[0282] 如图19中所示,如果识别出第一图像1515的分辨率是SD,作为对第一图像1515进行第一编码而获得的图像数据的比特率是10Mbps,并且通过AV1编解码器对第一图像1515进行第一编码,则AI上缩放器1634可使用多个DNN配置信息中的A DNN配置信息。

[0283] 如果从与第一图像1515相关的信息识别出第一图像1515的分辨率是HD,作为第一编码结果获得的图像数据的比特率是15Mbps,并且通过H.264编解码器对第一图像1515进行第一编码,则AI上缩放器1634可使用多个DNN配置信息中的B DNN配置信息。

[0284] 此外,如果从与第一图像1515相关的信息识别出第一图像1515的分辨率是全HD,作为对第一图像1515进行第一编码的结果而获得的图像数据的比特率是20Mbps,并且通过HEVC编解码器对第一图像1515进行第一编码,则AI上缩放器1634可使用多个DNN配置信息中的C DNN配置信息,并且如果识别出第一图像1515的分辨率为全HD,作为对第一图像1515进行第一编码的结果而获得的比特率是15Mbps,并且通过HEVC编解码器对第一图像1515进行第一编码,则AI上缩放器1634可使用多个DNN配置信息中的D DNN配置信息。根据作为对第一图像1515进行第一编码的结果而获得的图像数据的比特率是20Mbps还是15Mbps来选择C DNN配置信息和D DNN配置信息中的一个。当通过相同的编解码器对具有相同分辨率的第一图像1515进行第一编码时,图像数据的比特率彼此不同可指例如恢复的图像的图像质量彼此不同。因此,可基于预定的图像质量共同训练第一DNN和第二DNN,因此,AI上缩放器1634可根据指示第二图像1535的图像质量的图像数据的比特率来获得DNN配置信息。

[0285] 在另一实施例中, AI上缩放器1634可考虑到从第一解码器1632提供的信息(预测模式信息、运动信息、量化参数信息等)和AI数据中包括的第一图像1515相关信息, 获得多个DNN配置信息中的用于对第二图像1535进行AI上缩放的DNN配置信息。例如, AI上缩放器1634可从第一解码器1632接收在第一图像1515的第一编码处理中使用的量化参数信息, 从AI数据识别作为对第一图像1515进行编码的结果而获得的图像数据的比特率, 并且获得与量化参数和比特率相应的DNN配置信息。根据图像的复杂度, 相同的比特率可使重建图像具有不同的图像质量, 并且在第一图像1515中, 每个帧的比特率可不同, 作为表示整个第一图像1515的值, 其中, 在整个第一图像1515中, 第一编码图像1515被编码。因此, 考虑到可从第一解码器1632获得的针对每个帧的预测模式信息、运动信息和/或量化参数, 与仅使用AI数据相比, 可获得适合于第二图像1535的DNN配置信息。

[0286] 此外, 根据实施例, AI数据可包括相互约定的DNN配置信息的标识符。DNN配置信息的标识符是用于识别第一DNN与第二DNN之间的一对共同训练的DNN配置信息的信息, 使得第二图像1535被AI上缩放到与第一DNN的下缩放目标相应的上缩放目标。在获得AI数据中包括的DNN配置信息的标识符之后, AI上缩放器1634可使用与DNN配置信息的标识符相应的DNN配置信息对第二图像1535进行AI上缩放。例如, 可预先确定指示在第一DNN中可配置的多个DNN配置信息中的每个DNN配置信息的标识符以及指示在第二DNN中可配置的多个DNN配置信息中的每个DNN配置信息的标识符。在这种情况下, 可为针对第一DNN和第二DNN中的每一个可配置的每对DNN配置信息指定相同的标识符。AI数据可包括在第一DNN中设置的用于对图像1505进行AI下缩放的DNN配置信息的标识符。接收AI数据的AI上缩放器1634可使用多个DNN配置信息的由AI数据中包括的标识符指示的DNN配置信息对第二图像1535进行AI上缩放。

[0287] 根据实施例, AI数据可包括DNN配置信息。AI上缩放器1634可获得AI数据中包括的DNN配置信息, 然后使用相应的DNN配置信息对第二图像1535进行AI上缩放。

[0288] 根据实施方式, 如果以查找表的形式存储DNN配置信息(例如, 卷积层的数量、用于每个卷积层的滤波器核的数量、每个滤波器核的参数等)的信息, 则AI上缩放器1634可基于AI数据中包括的信息对查找表值的选择的部分进行组合以获得DNN配置信息, 并且使用获得的DNN配置信息对第二图像1535进行AI上缩放。

[0289] 根据实施例, 当确定了与上缩放目标相应的DNN结构时, AI上缩放器1634可获得与确定的DNN结构相应的DNN配置信息, 例如, 滤波器核的参数。

[0290] AI上缩放器1634通过包括与第一DNN相关的信息的AI数据获得第二DNN的DNN配置信息, 并且通过利用获得的DNN配置信息设置的第二DNN对第二图像1535进行AI上缩放, 其中, 与通过直接分析第二图像1535的特征进行上缩放相比, 上述操作可减少存储器使用量以及计算量。

[0291] 在实施例中, 当第二图像1535包括多个帧时, AI上缩放器1634可按照预定数量的帧独立地获得DNN配置信息, 或者获得用于整个帧的公共DNN配置信息。

[0292] 图20是示出包括多个帧的第二图像1535的示意图。

[0293] 如图20中所示, 第二图像1535可包括与 $t_0$ 至 $t_n$ 相应的帧。

[0294] 在示例中, AI上缩放器1634可经由AI数据获得第二DNN的DNN配置信息, 并且基于获得的DNN配置信息对与 $t_0$ 至 $t_n$ 相应的帧进行AI上缩放。也就是说, 可基于公共DNN配置信

息对与t0至tn相应的帧进行AI上缩放。

[0295] 在另一示例中,针对包括预定数量的帧的每个组,AI上缩放器1634可利用从AI数据获得的“**A**”DNN配置信息对与t0至tn相应的帧中的一些帧(例如,与t0至ta相应的帧)进行AI上缩放,并且利用从AI数据获得的“**B**”DNN配置对与ta+1至tb相应的帧进行AI上缩放。AI上缩放器1634可利用“**C**”DNN配置信息对与tb+1至tn相应的帧进行AI上缩放。换句话说,AI上缩放器1634可独立地获得用于包括多个帧中的预定数量的帧的每个组的DNN配置信息,并且可按照独立地获得的DNN配置信息对每个组中包括的帧进行AI上缩放。

[0296] 在另一实施例中,AI上缩放器1634可独立地获得用于第二图像1535的每个帧的DNN配置信息。也就是说,当第二图像1535包括三个帧时,AI上缩放器1634可利用针对第一帧获取的DNN配置信息对第一帧进行AI上缩放,利用针对第二帧获取的DNN配置信息对第二帧进行AI上缩放,并且利用针对第三帧获取的DNN配置信息对第三帧进行AI上缩放。由于可针对第二图像1535的每个帧独立地确定模式信息、量化参数信息等,因此,可根据获得DNN配置信息的方式,基于从第一解码器1632提供的信息(预测模式信息、运动信息、量化参数信息等)和AI数据中包括的第一图像1515的相关信息按照第二图像1535的每个帧独立地获得DNN配置信息。

[0297] 在另一示例中,AI数据可包括指示直到哪一帧为止基于AI数据获得的DNN配置信息有效的信息。例如,如果在AI数据中包括在ta帧之前DNN配置信息有效的信息,则AI上缩放器1634可利用基于AI数据获得的DNN配置信息对t0至ta帧进行AI上缩放。如果在另一AI数据中包括直到tn帧为止DNN配置信息有效的信息,则AI上缩放器1634可基于根据另一AI数据获得的DNN配置信息对ta+1至tn帧进行AI上缩放。

[0298] 参照图21,将描述用于对图像1505进行AI编码的AI编码设备600。

[0299] 图21是示出根据实施例的AI编码设备600的示例配置的框图。

[0300] 参照图21,AI编码设备600可包括AI编码器(例如,包括处理电路和/或可执行程序元件)610和发送器(例如,包括发送电路)630。AI编码器610可包括AI下缩放器(例如,包括处理电路和/或可执行程序元件)612和第一编码器(例如,包括处理电路和/或可执行程序元件)614。发送器630可包括数据处理器(例如,包括处理电路和/或可执行程序元件)632和通信器(例如,包括通信电路)634。

[0301] 尽管图21将AI编码器610和发送器630示出为单独的装置,但是AI编码器610和发送器630可通过单个处理器被实现。在这种情况下,专用处理器可实现,或者通用处理器(诸如,AP、CPU或GPU和S/W)的组合可实现。此外,在专用处理器的情况下,包括用于实现实施例的存储器或用于使用外部存储器的存储器处理器可实现专用处理器。

[0302] AI编码器610和发送器630可由多个处理器组成。在这种情况下,专用处理器的组合可实现,或者多个通用处理器(诸如,AP、CPU或GPU和S/W)的组合可实现。还可用不同的处理器来实现AI下缩放器612和第一编码器614。

[0303] AI编码器610可包括各种处理电路和/或可执行程序元件,并且执行对图像1505的AI下缩放以及对第一图像1515的第一编码,并将AI数据和图像数据发送到发送器630。

[0304] 图像数据包括作为对第一图像1515进行第一编码的结果而获得的数据。图像数据可包括基于第一图像1515中的像素值获得的数据,例如,作为第一图像1515与第一图像1515的预测数据之间的差的残差数据。图像数据包括在第一图像1515的第一编码处理中使

用的信息。例如,图像数据可包括用于对第一图像1515进行第一编码的预测模式信息、运动信息、以及用于对第一图像1515进行第一编码的量化参数相关信息。

[0305] AI数据包括使AI上缩放器1634能够对第二图像1535进行AI上缩放到与第一DNN的下缩放目标相应的上缩放目标的信息。在一个示例中,AI数据可包括图像1505与第一图像1515之间的差信息。AI数据还可包括与第一图像1515相关的信息。与第一图像1515相关的信息可包括关于第一图像1515的分辨率、作为对第一图像1515进行第一编码的结果而获得的图像数据的比特率、或在第一图像1515进行第一编码中使用的编解码器类型中的至少一个的信息。

[0306] 在实施例中,AI数据可包括相互约定的DNN配置信息的标识符,使得第二图像1535被AI上缩放到与第一DNN的下缩放相应的上缩放目标。

[0307] 在实施例中,AI数据可包括可被设置为第二DNN的DNN配置信息。

[0308] AI下缩放器612可包括各种处理电路和/或可执行程序元件,并且获得通过第一DNN从图像1505进行AI下缩放的第一图像1515。AI下缩放器612可基于预定标准确定图像1505的下缩放目标。

[0309] 为了获取与下缩放目标相应的第一图像1515,AI下缩放器612可存储可设置为第一DNN的多个DNN配置信息。AI下缩放器612可获得多个DNN配置信息中的与下缩放目标相应的DNN配置信息,并且通过利用获得的DNN配置信息设置的第一DNN对图像1505进行AI下缩放。

[0310] 可训练多个DNN配置信息中的每个DNN配置信息以获得第一图像1515的预定分辨率和/或预定图像。例如,多个DNN配置信息中的任意一个DNN配置信息可包括用于获取分辨率为图像1505的1/2倍的第一图像1515的信息,例如,4K(4096\*2160)的图像1505的1/2倍的2K(2048\*1080)的第一图像1515,并且其他DNN配置信息可包括关于分辨率为图像1505的1/4倍的第一图像1515的信息,例如,8K(8192\*4320)的图像1505的1/4倍的2K(2048\*1080)的第一图像1515。

[0311] 根据实施例,如果以查找表的形式存储DNN配置信息中包括的信息(例如,卷积层的数量、用于每个卷积层的滤波器核的数量、每个滤波器核的参数等),AI下缩放器612可根据下缩放目标将查找表值的选择的部分进行组合以获得DNN配置信息,并且可使用获得的DNN配置信息对图像1505进行AI下缩放。

[0312] 根据实施例,AI下缩放器612可确定与下缩放目标相应的DNN的结构,并且获得与确定的DNN的结构相应的DNN配置信息,例如,滤波器核的参数。

[0313] 通过共同训练第一DNN和第二DNN,用于对图像1505进行AI下缩放的多个DNN配置信息可具有最优值。每个DNN配置信息包括第一DNN中包括的卷积层的数量、按照卷积层的滤波器核的数量或每个滤波器核的参数中的至少一个。

[0314] AI下缩放器612可利用用于对图像1505进行AI下缩放而确定的DNN配置信息来设置第一DNN,以经由第一DNN获得第一图像1515的预定分辨率和/或预定图像质量。当获得多个DNN配置信息中的用于对图像1505进行AI下缩放的DNN配置信息时,第一DNN中的每个层可基于DNN配置信息中包括的信息处理输入数据。

[0315] 下面将描述通过AI下缩放器612确定下缩放目标的方法。下缩放目标可指示需要获得具有从图像1505减少多少分辨率的第一图像1515。

[0316] 在实施例中, AI下缩放器612可基于压缩率(例如, 图像1505与第一图像1515之间的分辨率的差、目标比特率)、压缩质量(例如, 比特率类型)、压缩历史信息或图像1505的类型中的至少一个确定下缩放目标。

[0317] 在示例中, AI下缩放器612可基于从用户输入的预设压缩率或压缩质量确定下缩放目标。

[0318] 作为另一示例, AI下缩放器612可使用存储在AI编码设备600中的压缩历史信息确定下缩放目标。例如, 根据AI编码设备600可用的压缩历史信息, 可确定用户的优选编码质量或压缩率, 并且可根据基于压缩历史信息确定的编码质量确定下缩放目标。例如, 可根据基于压缩历史信息最常使用的编码质量确定第一图像1515的分辨率、图像质量等。

[0319] 作为另一示例, AI下缩放器612可基于根据压缩历史信息已经被使用了超过预定阈值的编码质量(例如, 已经被使用了超过预定阈值的编码质量的平均质量)确定下缩放目标。

[0320] 作为另一示例, AI下缩放器612可基于图像1505的分辨率、类型(例如, 文件格式)等确定下缩放目标。

[0321] 在实施例中, 当图像1505包括多个帧时, AI下缩放器612可按照预定数量的帧独立地确定下缩放目标, 或者确定针对整个帧的公共下缩放目标。

[0322] 在示例中, AI下缩放器612可将图像1505的帧划分为预定数量的组, 并且独立地确定针对每个组的下缩放目标。可为每个组确定相同或不同的下缩放目标。针对每个组, 组中包括的帧的数量可相同或不同。

[0323] 在另一示例中, AI下缩放器612可按照图像1505的帧独立地确定下缩放目标。针对每个帧, 可确定相同或不同的下缩放目标。

[0324] 在下文中, 将描述作为AI下缩放的示例基础的第一DNN 700的示例结构。

[0325] 图22是示出根据实施例的用于对图像进行AI下缩放的第一示例DNN 700的示意图。

[0326] 如图22中所示, 图像1505被输入到第一卷积层710。第一卷积层710使用尺寸为 $5 \times 5$ 的32个滤波器核对图像1505执行卷积处理。将作为卷积处理的结果而产生的32个特征图输入到第一激活层720。第一激活层720可将非线性特征分配给32个特征图。

[0327] 第一激活层720确定是否将从第一卷积层710输出的特征图的样本值传递到第二卷积层730。例如, 特征图的一些样本值被第一激活层720激活并被传递到第二卷积层730, 并且一些样本值被第一激活层720停用且不被传递到第二卷积层730。通过第一激活层720突出由从第一卷积层710输出的特征图表示的信息。

[0328] 第一激活层720的输出725被输入到第二卷积层730。第二卷积层730使用尺寸为 $5 \times 5$ 的32个滤波器核对输入数据执行卷积处理。将作为卷积处理的结果而产生的32个特征图输入到第二激活层740, 并且第二激活层740可将非线性特征分配给32个特征图。

[0329] 第二激活层740的输出745被输入到第三卷积层750。第三卷积层750使用尺寸为 $5 \times 5$ 的一个滤波器核对输入数据执行卷积处理。作为卷积处理的结果, 可从第三卷积层750输出一个图像。第三卷积层750是用于输出最终图像的层并使用一个滤波器核获得一个输出。根据示例, 第三卷积层750可经由卷积运算结果输出第一图像1515。

[0330] 表示第一DNN 700的第一卷积层710、第二卷积层730和第三卷积层750中的滤波器核的数量、滤波器核的参数等的DNN配置信息可以是多个, 其中, 所述多个DNN配置信息应与

第二DNN的多个DNN配置信息相关联。可通过共同训练第一DNN和第二DNN来实现第一DNN的多个DNN配置信息与第二DNN的多个DNN配置信息之间的关联。

[0331] 图22示出第一DNN 700包括三个卷积层710、730、750和两个激活层720、740,但这仅是一个示例,并且根据实施例,卷积层和激活层的数量可广泛地变化。此外,根据实施例,可经由RNN实现第一DNN 700。这种情况可指例如根据示例将第一DNN 700的CNN结构改变为RNN。

[0332] 在示例实施例中,AI下缩放器612可包括用于卷积运算和激活层的运算的至少一个ALU。可用处理器实现ALU。对于卷积运算,ALU可包括乘法器和加法器,其中,乘法器用于在滤波器核的样本值与从图像1505或先前层输出的特征图的样本值之间执行乘法运算,加法器将乘法的结果值相加。此外,对于激活层的运算,ALU可包括乘法器和比较器,其中,乘法器用于将输入样本值乘以在预定的S形函数、Tanh函数或ReLU函数中使用的权重,比较器用于将乘法的结果与预定值进行比较以确定是否将输入样本值传递到下一层。

[0333] 参照回图22,从AI下缩放器612接收第一图像1515的第一编码器614可对第一图像1515进行第一编码以减少第一图像1515具有的信息量。作为通过第一编码器614进行第一编码的结果,可获得与第一图像1515相应的图像数据。

[0334] 数据处理器632可包括各种处理电路和/或可执行程序元件,并且执行处理使得AI数据或图像数据中的至少一个可以以预定格式被发送。例如,如果将以比特流的形式发送AI数据和图像数据,则数据处理器632处理AI数据,使得AI数据以比特流的形式被表示,并且通过通信器634以一个比特流的形式发送AI数据和图像数据。作为另一示例,数据处理器632处理AI数据,使得AI数据以比特流的形式被表示,并且通过通信器634发送与AI数据相应的比特流以及与图像数据相应的比特流中的每一个。作为另一示例,数据处理器632处理AI数据,使得AI数据被表示为帧或数据包,并且通过通信器634以比特流的形式发送图像数据并且以帧或数据包的形式发送AI数据。

[0335] 发送器630可包括各种电路,并且通过网络发送作为AI编码的结果而获得的AI编码数据。作为AI编码的结果而获得的AI编码数据包括图像数据和AI数据。可通过相同的网络或不同的网络发送图像数据和AI数据。

[0336] 在实施例中,作为由数据处理器632处理的结果而获得的AI编码数据可被存储在包括磁性介质(诸如,硬盘、软光盘、磁带)、光学可记录介质(诸如,CD-ROM和DVD)、磁光介质(诸如,光磁软盘)等的数据存储介质中。

[0337] 在下文中,将参照图23描述共同训练第一DNN 700和第二DNN 300的方法。

[0338] 图23是示出根据实施例的用于训练第一DNN 700和第二DNN 300的示例方法的示意图。

[0339] 在示例实施例中,通过AI解码处理将AI编码的图像1505重建为第三图像1545,并且为了保持作为AI解码的结果而获得的第三图像1545与图像1505之间的相似性,需要AI编码处理和AI解码处理的相关性。也就是说,必须在AI解码处理中重建在AI编码处理中损失的信息,这需要共同训练第一DNN 700和第二DNN 300。

[0340] 为了准确的AI解码,需要减少与如图23中所示的第三训练图像804与原始训练图像801之间的比较结果相应的质量损失信息2330。质量损失信息2330用于训练第一DNN 700和第二DNN 300两者。

[0341] 将描述图23中示出的训练处理。

[0342] 在图23中,原始训练图像801是未经过AI下缩放的图像,并且第一训练图像802是从原始训练图像801被AI下缩放后的图像。第三训练图像804是从第一训练图像802被AI上缩放后的图像。

[0343] 原始训练图像801包括静止图像或包括多个帧的运动图像。在示例实施例中,原始训练图像801可包括从静止图像或包括多个帧的运动图像提取的亮度图像。在示例实施例中,原始训练图像801可包括从静止图像或由多个帧组成的运动图像提取的补丁图像。当原始训练图像801包括多个帧时,第一训练图像802、第二训练图像和第三训练图像804也包括多个帧。当将原始训练图像801的多个帧被顺序地输入到第一DNN 700时,可通过第一DNN 700和第二DNN 300顺序地获取第一训练图像802、第二训练图像和第三训练图像804的多个帧。

[0344] 为了共同训练第一DNN 700和第二DNN 300,原始训练图像801被输入到第一DNN 700。输入到第一DNN 700的原始训练图像801被AI下缩放并被输出为第一训练图像802,并且第一训练图像802被输入到第二DNN 300。作为针对第一训练图像802进行AI上缩放的结果,输出第三训练图像804。

[0345] 参照回图23,根据实施例,第一训练图像802被输入到第二DNN 300,并且通过对第一训练图像802进行第一编码和第一解码处理获得的第二训练图像可被输入到第二DNN 300。MPEG-2、H.264、MPEG-4、HEVC、VC-1、VP8、VP9和AV1中的任意一个都可用于将第二训练图像输入到第二DNN中。例如,在对第一训练图像802进行第一编码和对与第一训练图像802相应的图像数据进行第一解码中,可使用MPEG-2、H.264、MPEG-4、HEVC、VC-1、VP8、VP9和AV1中的任意一个。

[0346] 参照回图23,除了经由第一DNN 700输出第一训练图像802之外,还从原始训练图像801获得传统下缩放的减小的训练图像803。传统下缩放可包括双线性缩放、双三次缩放、lanczos缩放或阶梯缩放中的至少一个。

[0347] 为了防止和/或减少第一图像1515的结构特征明显偏离图像1505的结构特征,获得保留原始训练图像801的结构特征的减小的训练图像803。

[0348] 在训练之前,可利用预定的DNN配置信息设置第一DNN 700和第二DNN 300。随着训练进行,可确定结构损失信息2310、复杂度损失信息2320和质量损失信息2330。

[0349] 可基于减小的训练图像803和第一训练图像802的比较结果确定结构损失信息2310。在一个示例中,结构损失信息2310可与减小的训练图像803的结构信息和第一训练图像802的结构信息之间的差相应。结构信息可包括可从图像中提取的图像的各种特征,诸如,亮度、对比度、直方图等。结构损失信息2310指示原始训练图像801的结构信息在第一训练图像802中被保留的程度。结构损失信息2310越小,第一训练图像802的结构信息与原始训练图像801的结构信息越相似。

[0350] 可基于第一训练图像802的空间复杂度确定复杂度损失信息2320。在一个示例中,可使用第一训练图像802的总方差值作为空间复杂度。复杂度损失信息2320与通过对第一训练图像802进行第一编码而获得的图像数据的比特率相关。定义了复杂度损失信息2320越小,图像的比特率越小。

[0351] 可基于原始训练图像801和第三训练图像804的比较结果确定质量损失信息2330。

质量损失信息2330可包括以下值中的至少一个值:针对原始训练图像801与第三训练图像804之间的差的L1-范数值、L2-范数值、结构相似度(SSIM)值、峰值信噪比-人类视觉系统(PSNR-HVS)值、多尺度SSIM(MS-SSIM)值、方差插值因子(VIF)值、视频质量多方法评估融合(VMAF)值。质量损失信息2330指示第三训练图像804与原始训练图像801相似的程度。质量损失信息2330越小,第三训练图像804与原始训练图像801越相似。

[0352] 参照图23,结构损失信息2310、复杂度损失信息2320和质量损失信息2330用于训练第一DNN 700,并且质量损失信息2330用于训练第二DNN 300。也就是说,质量损失信息2330用于训练第一DNN 700和第二DNN 300两者。

[0353] 第一DNN 700可更新参数,使得基于结构损失信息2310、复杂度损失信息2320和质量损失信息2330确定的最终损失信息被减少或最小化。第二DNN 300可更新参数,使得质量损失信息2330被减少或最小化。

[0354] 用于训练第一DNN 700和第二DNN 300的最终损失信息可被确定为下面的等式3。

[0355] **【等式3】**

[0356]  $LossDS = a * \text{结构损失信息} + b * \text{复杂度损失信息} + c * \text{质量损失信息}$

[0357]  $LossUS = d * \text{质量损失信息}$

[0358] 在等式3中,LossDS表示对于第一DNN 700的训练应被减少或最小化的最终损失信息,并且LossUS表示对于第二DNN 300的训练应被减少或最小化的最终损失信息。此外,a、b、c和d可与预定权重相应。

[0359] 第一DNN 700在等式3的LossDS被减少的方向上更新参数,并且第二DNN 300在LossUS被减少的方向上更新参数。当根据从训练处理推导出的LossDS更新第一DNN 700的参数时,基于更新的参数获得的第一训练图像802与先前训练处理中的第一训练图像802不同,使得第三训练图像804也与先前训练处理中的第三训练图像804不同。当第三训练图像804与先前训练处理的第三训练图像804不同时,质量损失信息2330也被新确定,因此,第二DNN 300更新参数。当质量损失信息2330被新确定时,LossDS也被新确定,因此,第一DNN 700根据新确定的LossDS更新参数。也就是说,第一DNN 700的参数更新导致第二DNN 300的参数更新,并且第二DNN 300的参数更新导致第一DNN 700的参数更新。换句话说,由于通过共享质量损失信息2330来共同训练第一DNN 700和第二DNN 300,因此可彼此相关地优化第一DNN 700的参数和第二DNN 300的参数。

[0360] 参照等式3,已知根据质量损失信息2330确定LossUS,但这仅是说明性示例,并且可基于结构损失信息2310或复杂度损失信息2320中的至少一个和质量损失信息2330确定LossUS。

[0361] 已经描述了AI解码设备1600的AI上缩放器1634和AI编码设备600的AI下缩放器612存储多个DNN配置信息。将描述对存储在AI上缩放器1634和AI下缩放器612中的多个DNN配置信息中的每个DNN配置信息进行训练的方法。

[0362] 如结合等式3所述,在第一DNN 700的情况下,考虑以下项来更新参数:第一训练图像802的结构信息与原始训练图像801的结构信息之间的相似程度(结构损失信息2310)、通过第一训练图像802的第一编码结果获得的图像数据的比特率(复杂度损失信息2320)、以及第三训练图像804与原始训练图像801之间的差(质量损失信息2330)。

[0363] 可更新第一DNN 700的参数,使得可获得与原始训练图像801的结构信息相似的第

一训练图像802,且从第一编码获得的图像的比特率与原始训练图像801相似,并且使得用于对第一训练图像802进行AI上缩放的第二DNN 300获得与原始训练图像801相似的第三训练图像804。

[0364] 通过调节等式3的a、b和c的权重,第一DNN 700的参数被优化的方向不同。例如,如果将b的权重确定为高,则第一DNN 700的参数可被更新,使得重要性被赋予比第三训练图像804的质量低的比特率。此外,当将c的权重确定为高时,第一DNN 700的参数可被更新,使得与更高的比特率相比,重要性被赋予提高第三训练图像804的质量或者保持原始训练图像801的结构信息。

[0365] 此外,第一DNN 700的参数可根据用于对第一训练图像802进行第一编码的编解码器的类型而不同,由于取决于编解码器的类型,因此将被输入到第二DNN 300的第二训练图像可不同。

[0366] 可基于权重a、权重b、权重c以及用于对第一训练图像802进行第一编码的编解码器的类型共同更新第一DNN 700的参数和第二DNN 300的参数。因此,可通过将权重a、权重b和权重c确定为预定值、将编解码器的类型确定为预定类型、然后训练第一DNN 700和第二DNN 300,来确定共同优化的第一DNN 700的参数和第二DNN 300的参数。

[0367] 在改变权重a、权重b、权重c和编解码器的类型之后,可通过训练第一DNN 700和第二DNN300确定彼此相关地被优化的第一DNN 700的参数和第二DNN 300的参数。换句话说,当在分别改变权重a、权重b、权重c的值和编解码器类型的情况下训练第一DNN 700和第二DNN 300时,可由第一DNN 700和第二DNN 300确定被共同训练的多个DNN配置信息。

[0368] 如以上参照图19所述,可将第一DNN 700和第二DNN 300的多个DNN配置信息映射到第一图像相关信息。为了建立映射关系,可根据特定的比特率利用特定的编解码器对从第一DNN 700输出的第一训练图像802进行第一编码,并且可将通过对通过第一编码获得的比特流进行第一解码而获得的第二训练图像输入到第二DNN 300。也就是说,在设置环境使得通过特定编解码器以特定比特率对特定分辨率的第一训练图像802进行第一编码之后,通过训练第一DNN 700和第二DNN 300,可确定第一训练图像802的分辨率、用于对第一训练图像802进行第一编码的编解码器的类型、以及被映射到作为对第一训练图像802进行第一编码的结果而获得的比特流的比特率的DNN配置信息对。通过改变第一训练图像802的分辨率,可确定用于对第一训练图像802进行第一编码的编解码器的类型、根据对第一训练图像802进行第一编码获得的比特流的比特率以及第一DNN 700和第二DNN 300的第一图像相关信息与多个DNN配置信息之间的映射关系。

[0369] 图24是示出根据实施例的由训练装置1000对第一DNN 700和第二DNN 300的示例训练处理的信号流程图。

[0370] 参照图24,可由训练装置1000执行第一DNN 700和第二DNN 300的训练。训练装置1000可包括第一DNN 700和第二DNN300。训练装置1000可以是例如AI编码设备600或单独的服务器。通过训练获得的第二DNN 300的DNN配置信息被存储在AI解码设备1600中。

[0371] 参照图24,在操作S840和S845,训练装置1000初始设置第一DNN 700和第二DNN 300的DNN配置信息。因此,第一DNN 700和第二DNN 300可根据预定的DNN配置信息进行操作。DNN配置信息可包括关于第一DNN 700和第二DNN 300中包括的卷积层的数量、用于每个卷积层的滤波器核的数量、用于每个卷积层的滤波器核的尺寸、以及每个滤波器核的参数

中的至少一个的信息。

[0372] 在操作S850,训练装置1000将原始训练图像801输入到第一DNN 700。原始训练图像801可包括静止图像或运动图像中的至少一帧。

[0373] 在操作S855,第一DNN 700根据初始设置的DNN配置信息处理原始训练图像801,并且输出从原始训练图像801被AI下缩放的第一训练图像802。尽管图23示出从第一DNN 700输出的第一训练图像802被直接输入到第二DNN 300,但是可通过训练装置1000将从第一DNN 700输出的第一训练图像802输入到第二DNN 300。此外,训练装置1000可利用预定的编解码器对第一训练图像802进行第一编码和第一解码,然后将第二训练图像输入到第二DNN 300中。

[0374] 在操作S860,第二DNN 300根据初始设置的DNN设置信息处理第一训练图像802或第二训练图像,并且输出从第一训练图像802或第二训练图像被AI上缩放的第三训练图像804。

[0375] 在操作S865,训练装置1000基于第一训练图像计算复杂度损失信息2320。

[0376] 在操作S870,训练装置1000通过将减小的训练图像803和第一训练图像802进行比较来计算结构损失信息2310。

[0377] 在操作S875,训练装置1000通过将原始训练图像801和第三训练图像804进行比较来计算质量损失信息2330。

[0378] 在操作S880,第一DNN 700基于最终损失信息通过反向传播处理更新初始设置的DNN配置信息。训练装置1000可基于复杂度损失信息2320、结构损失信息2310和质量损失信息2330计算用于训练第一DNN 700的最终损失信息。

[0379] 在操作S885,第二DNN 300基于质量损失信息或最终损失信息通过反向传播处理更新初始设置的DNN配置信息。训练装置1000可基于质量损失信息2330计算用于训练第二DNN 300的最终损失信息。

[0380] 然后,训练装置1000、第一DNN 700和第二DNN 300通过重复操作S850至S885直到最终损失信息被最小化和/或减少来更新DNN配置信息。在每个重复处理期间,第一DNN 700和第二DNN 300根据在先前处理中更新的DNN配置信息进行操作。

[0381] 在下面的表1中,根据实施例示出了对图像1505进行AI编码和AI解码的情况以及利用HEVC对图像1505进行编码和解码的情况的效果。

[0382] 表1

内容	分辨率	帧的数量	信息量 (比特率) (Mbps)		主观图像质量评分 (VMAF)	
			HEVC	AI 编码/ AI 解码	HEVC	AI 编码/ AI 解码
[0383] 内容.01	8K (7680 x 4320)	300 帧	46.3	21.4	94.80	93.54
内容.02			46.3	21.6	98.05	98.98
内容.03			46.3	22.7	96.08	96.00
内容.04			46.1	22.1	86.26	92.00
内容.05			45.4	22.7	93.42	92.98
内容.06			46.3	23.0	95.99	95.61
		平均	46.11	22.25	94.10	94.85

[0384] 如表1中所示,可看出,即使对8K分辨率的包括300帧的内容进行AI编码和AI解码

的主观图像质量高于通过HEVC进行编码和解码的主观图像质量,比特率也降低了50%或更多。

[0385] 图25是示出根据实施例的用于对图像1505进行AI下缩放的设备10'和用于对第二图像1535进行AI上缩放的设备20'的示意图。

[0386] 设备10'接收图像1505,并且使用AI下缩放器2524和基于变换的编码器2526向设备20'提供图像数据22和AI数据23。在示例实施例中,图像数据22与图15的图像数据相应,并且AI数据23与图15的AI数据相应。此外,在示例实施例中,基于变换的编码器2526与图21的第一编码器614相应,并且AI下缩放器2524与图21的AI下缩放器612相应。

[0387] 设备20'接收AI数据23和图像数据22,并且使用基于变换的解码器2546和AI上缩放器2544获得第三图像1545。在示例实施例中,基于变换的解码器2546与图16的第一解码器1632相应,并且AI上缩放器2544与图16的AI上缩放器1634相应。

[0388] 在示例实施例中,设备10'包括计算机程序CPU和存储器,其中,计算机程序包括指令。该计算机程序被存储在存储器中。在示例实施例中,在通过CPU执行计算机程序时,设备10'执行结合图25描述的功能。在示例实施例中,通过专用硬件芯片和/或CPU执行结合图25描述的功能。

[0389] 在示例实施例中,设备20'包括计算机程序CPU和存储器,其中,计算机程序包括指令。该计算机程序被存储在存储器中。在示例实施例中,在通过CPU执行计算机程序时,设备20'执行结合图25描述的功能。在示例实施例中,通过专用硬件芯片和/或CPU执行结合图25描述的功能。

[0390] 在图25中,配置控制器2522接收一个或更多个输入值21。在示例实施例中,一个或更多个输入值21可包括用于AI下缩放器2524和AI上缩放器2544的目标分辨率差、图像数据22的比特率、图像数据22的比特率类型(例如,可变比特率类型、恒定比特率类型或平均比特率类型等)或对于基于变换的编码器2526的编解码器类型中的至少一个。一个或更多个输入值21可被预先存储在设备10'中,或者可包括从用户接收到的值。

[0391] 配置控制器2522基于接收到的输入值21控制AI下缩放器2524和基于变换的编码器2526的操作。在示例实施例中,配置控制器2522根据接收到的输入值21获得用于AI下缩放器2524的DNN配置信息,并且利用获得的DNN配置信息设置AI下缩放器2524。在示例实施例中,配置控制器2522可将接收到的输入值21传递给AI下缩放器2524,并且AI下缩放器2524可基于接收到的输入值21获得用于对图像1505进行AI下缩放的DNN配置信息。在示例实施例中,配置控制器2522可将附加信息(例如,应用了AI下缩放的格式(亮度分量、色度分量、红色分量、绿色分量或蓝色分量等)信息、高动态范围(HDR)的色调映射信息等)与输入值21一起提供给AI下缩放器2524,并且AI下缩放器2524可考虑输入值21和附加信息获得DNN配置信息。在示例实施例中,配置控制器2522将接收到的输入值21的至少一部分传递给基于变换的编码器2526,以使基于变换的编码器2526利用特定的比特率值、特定类型的比特率和特定编解码器对第一图像1515进行第一编码。

[0392] AI下缩放器2524接收图像1505并且执行结合图15、21、22、23或24中的至少一个描述的操作以获得第一图像1515。

[0393] 在示例实施例中,AI数据23被提供给设备20'。AI数据23可包括图像1505与第一图像1515之间的分辨率差信息或与第一图像1515相关的信息中的至少一个。可基于输入值21

的目标分辨率差确定分辨率差信息,并且可基于目标比特率、比特率类型和编解码器类型中的至少一个确定与第一图像1515相关的信息。在示例实施例中,AI数据23可包括在AI上缩放处理中使用的参数。可从AI下缩放器2524将AI度量数据提供给设备20'。

[0394] 通过基于变换的编码器2526处理第一图像1505以获得图像数据22,并且将图像数据22发送到设备20'。基于变换的编码器2526可根据MPEG-2、H.264AVC、MPEG-4、HEVC、VC-1、VP8,VP9或AV1处理第一图像2515。

[0395] 配置控制器2542基于AI数据23控制AI上缩放器2544的操作。在示例实施例中,配置控制器2542根据接收到的AI数据23获得用于AI上缩放器2544的DNN配置信息,并且利用获得的DNN配置信息设置AI上缩放器2544。在示例实施例中,配置控制器2542将接收到的AI数据23传递给AI上缩放器2544,并且AI上缩放器2544可基于AI数据23获得用于对第二图像1535进行AI上缩放的DNN配置信息。在示例实施例中,配置控制器2542可将附加信息(例如,应用了AI上缩放的YCbCr颜色格式(亮度分量、色度分量、红色分量、绿色分量或蓝色分量等)信息、HDR的色调映射信息等)与AI数据23一起提供给AI上缩放器2544,并且AI上缩放器2544可考虑AI数据23和附加信息获得DNN配置信息。在示例实施例中,AI上缩放器2544可从配置控制器2542接收AI数据23,从基于变换的解码器2546接收预测模式信息、运动信息和量化参数信息中的至少一个,并且基于预测模式信息、运动信息和量化参数信息中的至少一个以及AI数据23获得DNN配置信息。

[0396] 基于变换的解码器2546处理图像数据22以重建第二图像1535。基于变换的解码器2546可根据MPEG-2、H.264AVC、MPEG-4、HEVC、VC-1、VP8、VP9或AV1处理图像数据。

[0397] AI上缩放器2544可通过基于设置的DNN设置信息对从基于变换的解码器2546提供的第二图像1535进行AI上缩放来获得第三图像1545。

[0398] AI下缩放器2524包括第一DNN,并且AI上缩放器2544可包括第二DNN,其中,根据结合图23和图24描述的训练方法训练用于第一DNN和第二DNN的DNN配置信息。

[0399] 各种实施例除了可被应用于图像处理装置(诸如,机顶盒)之外,还可被应用于所有能够进行图像处理的电子设备,诸如,服务器。此外,可通过在电子设备中设置的嵌入式服务器或图像处理装置的外部服务器执行上述各种实施例。

[0400] 根据各种实施例的方法可被实现为可安装到现有电子设备的软件或应用的格式。

[0401] 可通过现有电子设备的软件升级或仅硬件升级来实现根据各种实施例的方法。

[0402] 同时,可以以软件来实现本公开的各种实施例,其中,所述软件包括存储在可由机器(例如,计算机)读取的机器可读存储介质上的指令。包括根据公开的实施例的图像处理设备(例如,图像处理设备A)的设备可从存储介质调用指令,并且执行调用的指令。当由处理器执行所述指令时,处理器可在处理器的控制下直接或使用其它组件来执行与所述指令相应的功能。所述指令可包括由编译器产生的代码或可由解释器执行的代码。可以以非暂时性存储介质的形式提供机器可读存储介质。在此,“非暂时性”存储介质可不包括信号,而是有形的,并且不对数据被半永久地存储在存储介质中的情况和数据被临时存储在存储介质中的情况进行区分。

[0403] 根据实施例,根据上述实施例的方法可被包括在计算机程序产品中。该计算机程序产品可作为产品在卖方与消费者之间被交易。可以以机器可读存储介质(例如,光盘只读存储器(CD-ROM))的形式或通过应用商店(例如,Play Store™)在线发布该计算机程序产

品,或者直接在线发布该计算机程序产品。在在线发布的情况下,计算机程序产品的至少一部分可被至少临时存储或临时产生在制造商的服务器、应用商店的服务器或机器可读存储介质(诸如,中继服务器的存储器)中。

[0404] 根据各种示例实施例,上述元件中的各个元件(例如,模块或程序)可包括单个实体或多个实体。根据实施例,可省略上述相应的元件中的至少一个元件或操作,或者可添加至少一个其他元件或操作。可选或另外地,可对多个组件(例如,模块或程序)进行组合以形成单个实体。在这种情况下,集成的实体可以以与由集成之前的多个元件中的相应的元件执行的功能相同或相似的方式,执行所述多个元件中的每个元件的元件的至少一个功能的功能。可连续地、并行地、重复地或启发性地执行根据各种实施例的模块、程序模块或由其它元件执行的操作,或者可根据不同的顺序执行至少一些操作,可省略至少一些操作,或者可将其它操作添加于此。

[0405] 尽管已经参照各种示例实施例示出和描述了各种实施例,但是本公开不限于特定实施例或附图,并且本领域普通技术人员将理解,在不脱离包括权利要求及其等同物的本公开的精神和范围的情况下,可对其进行形式和细节上的各种改变。

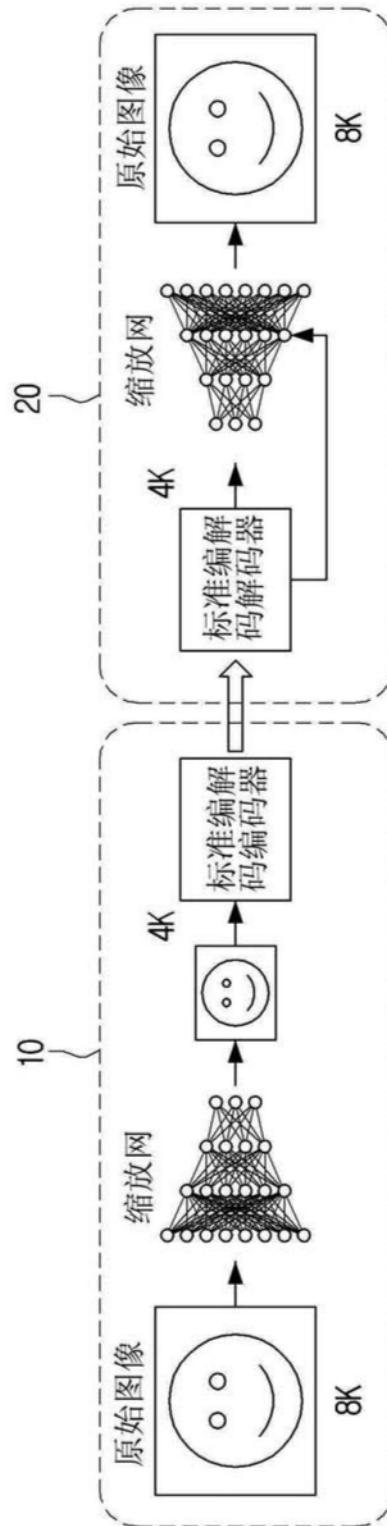


图1A

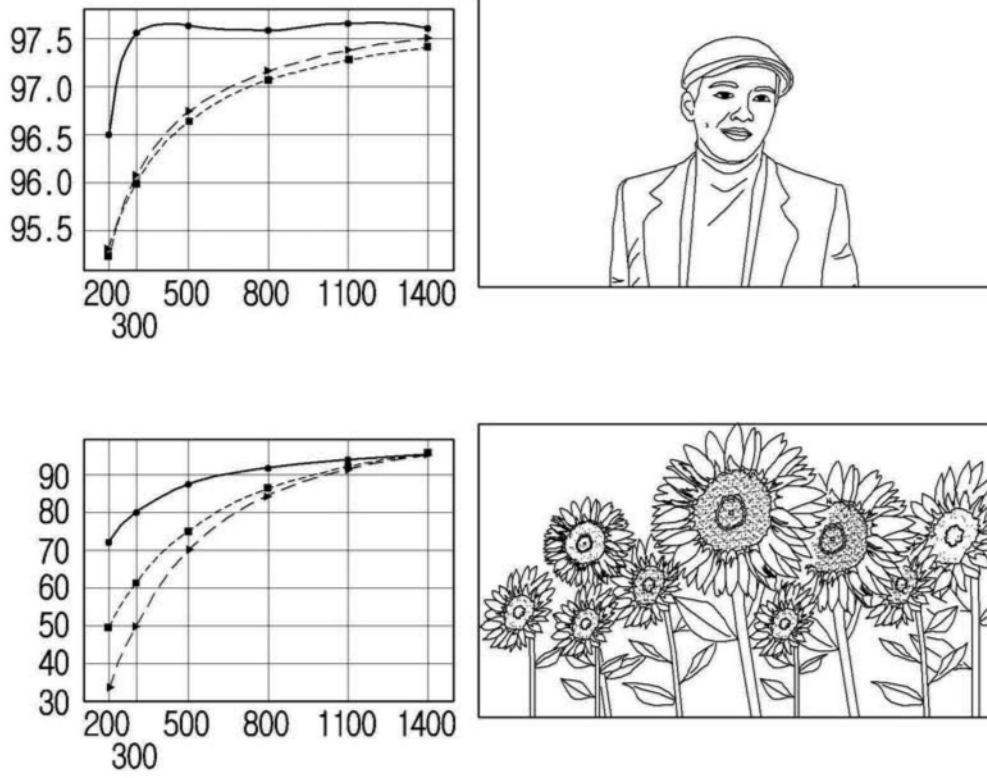


图1B

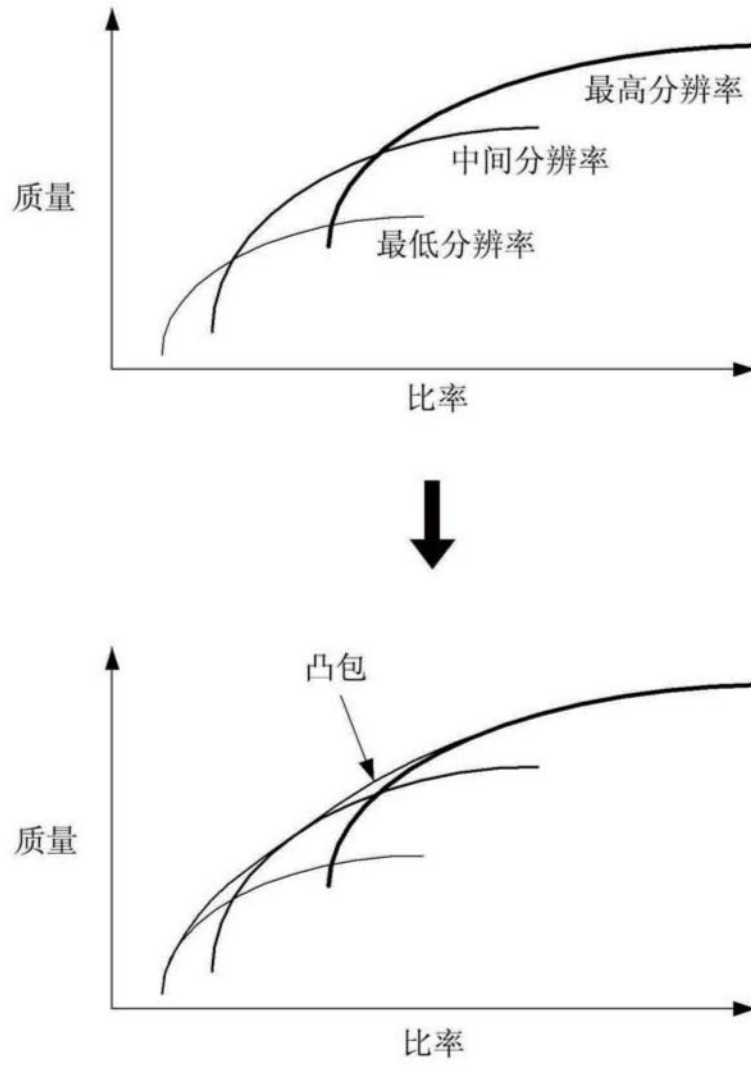


图1C

100



图2A

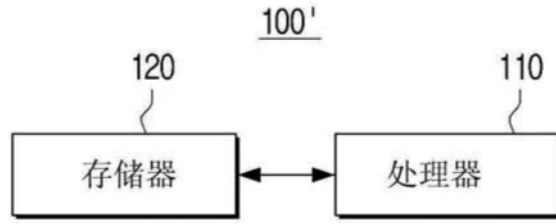


图2B

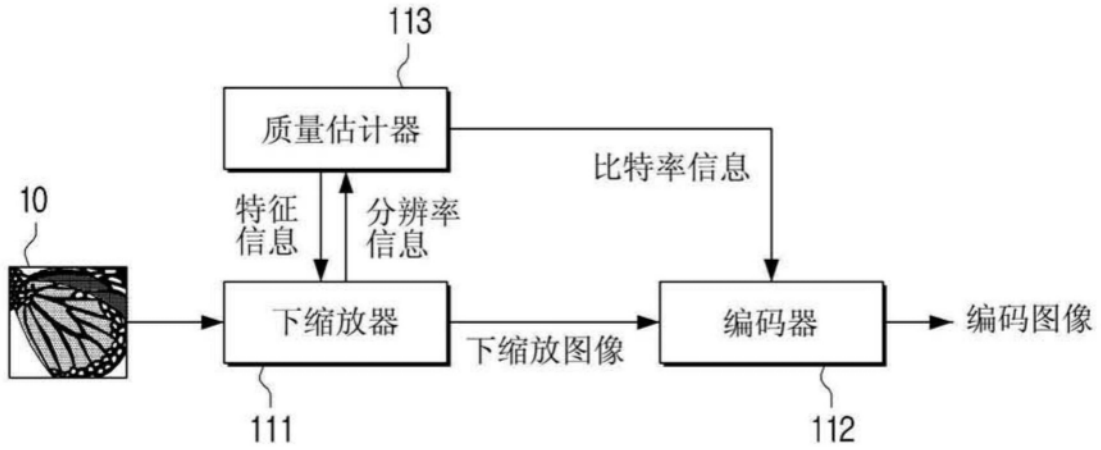


图3

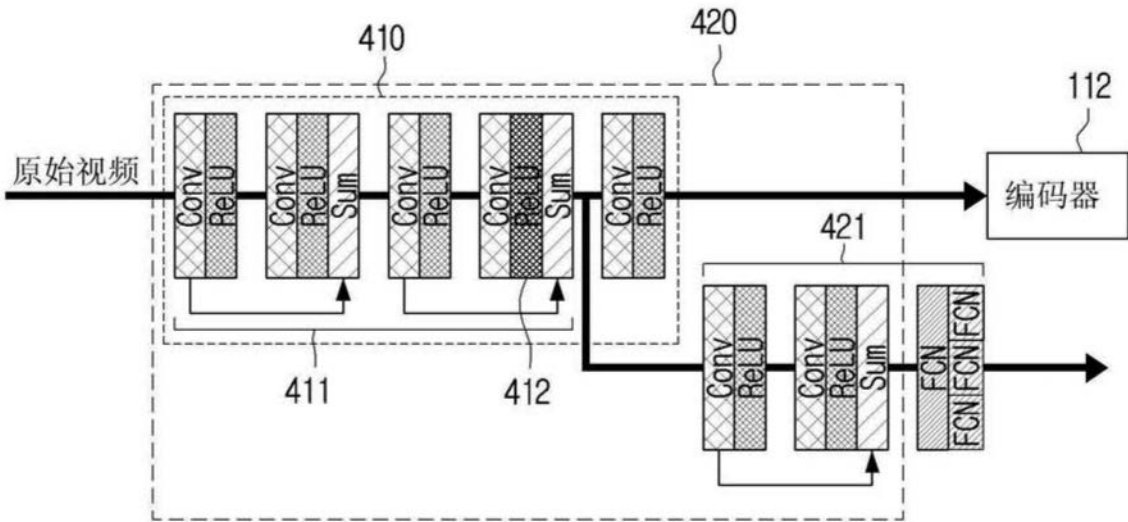


图4

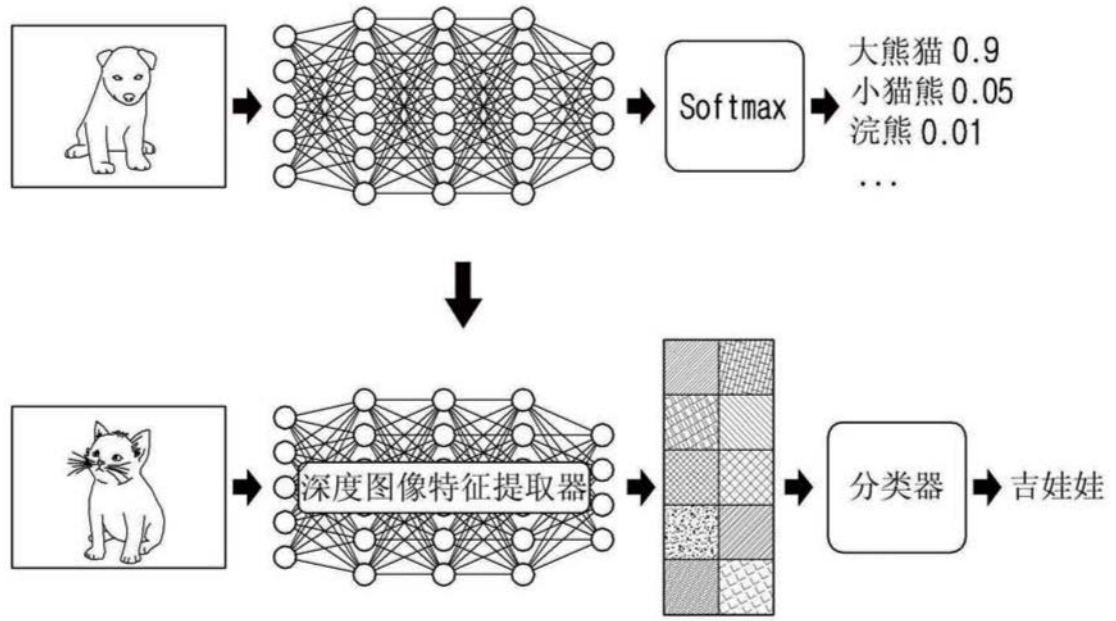


图5

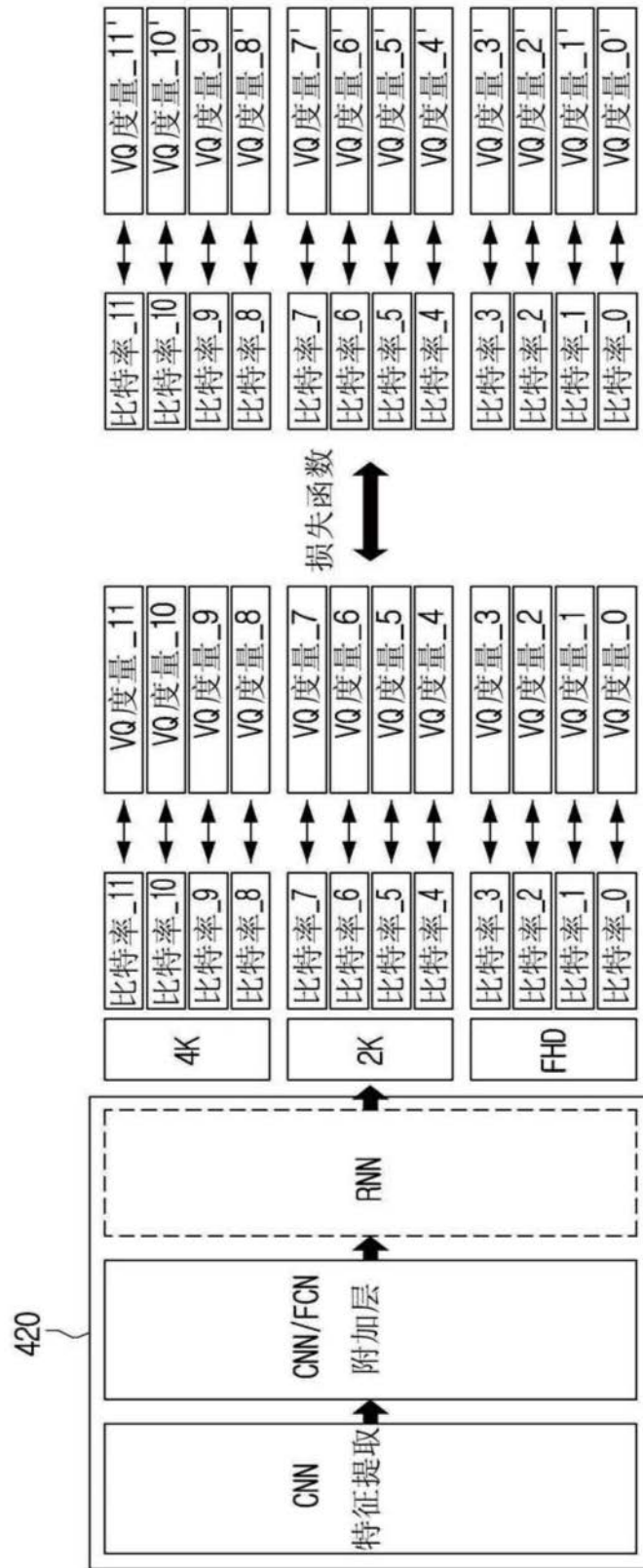


图6

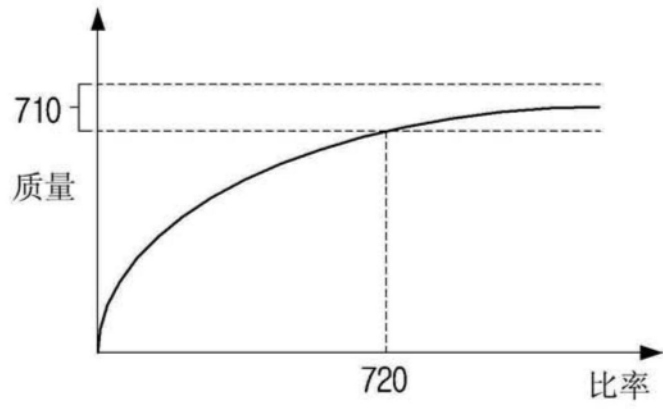


图7

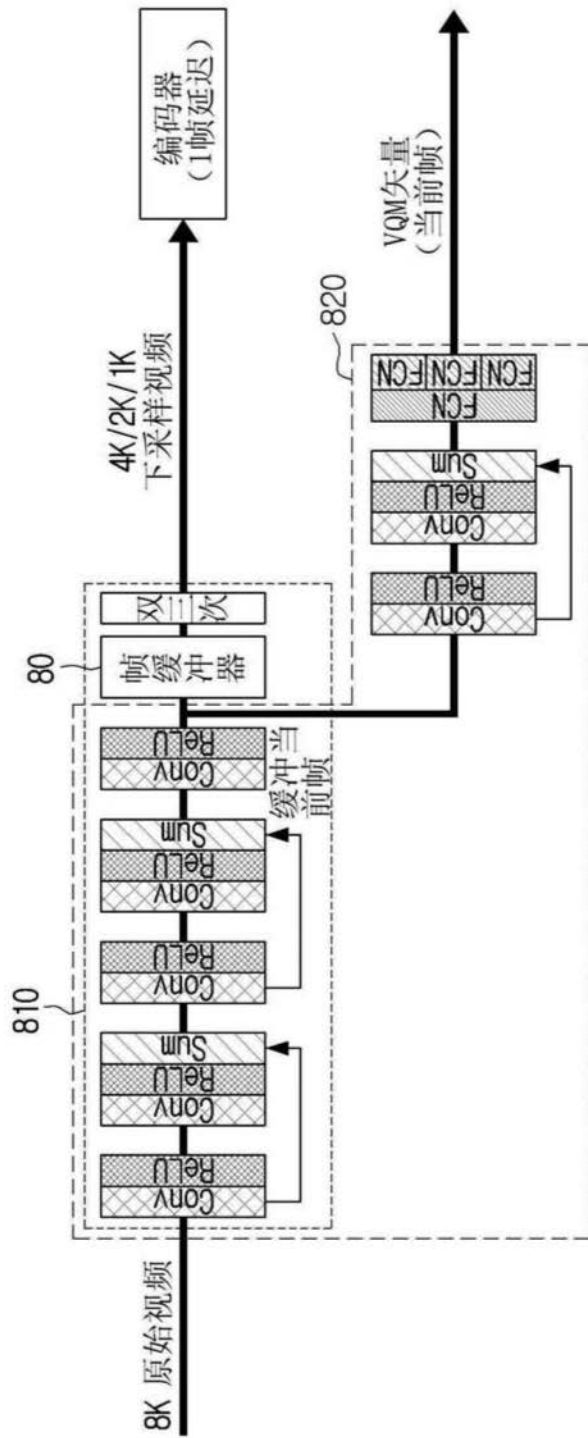


图8A

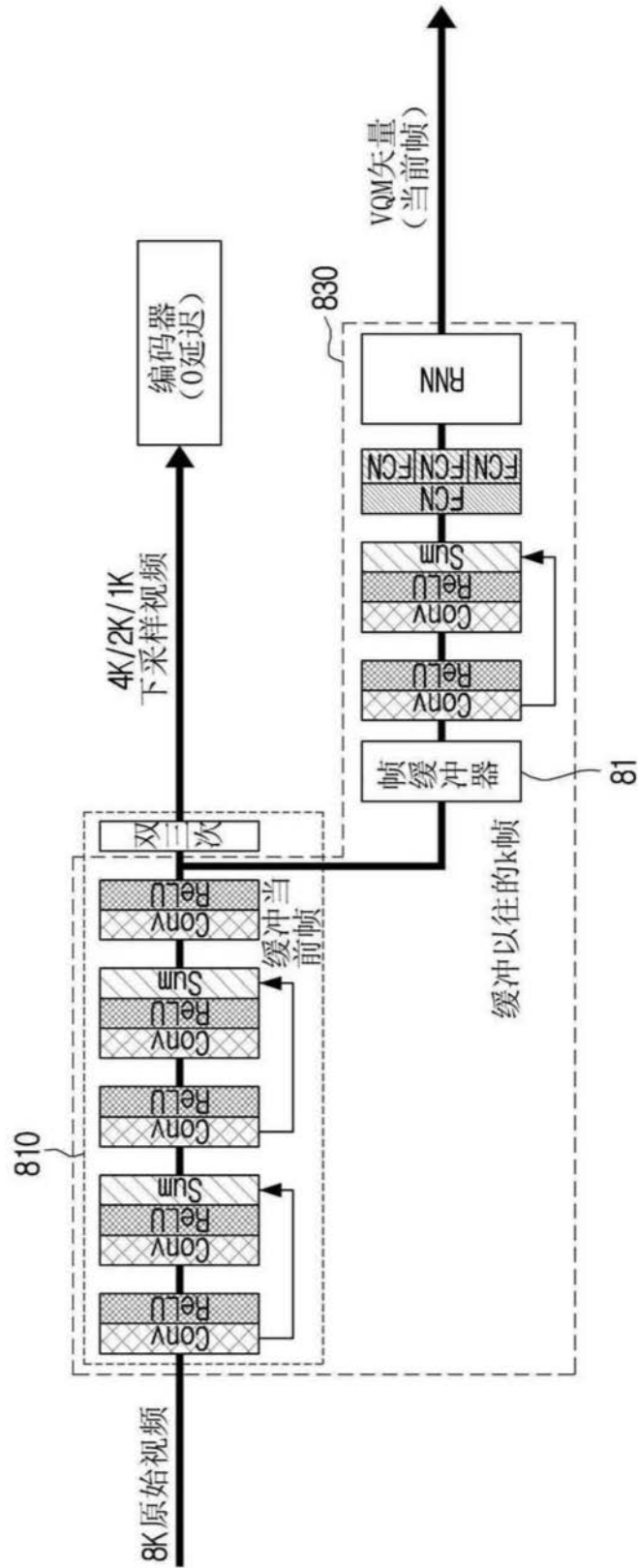


图8B

200



图9

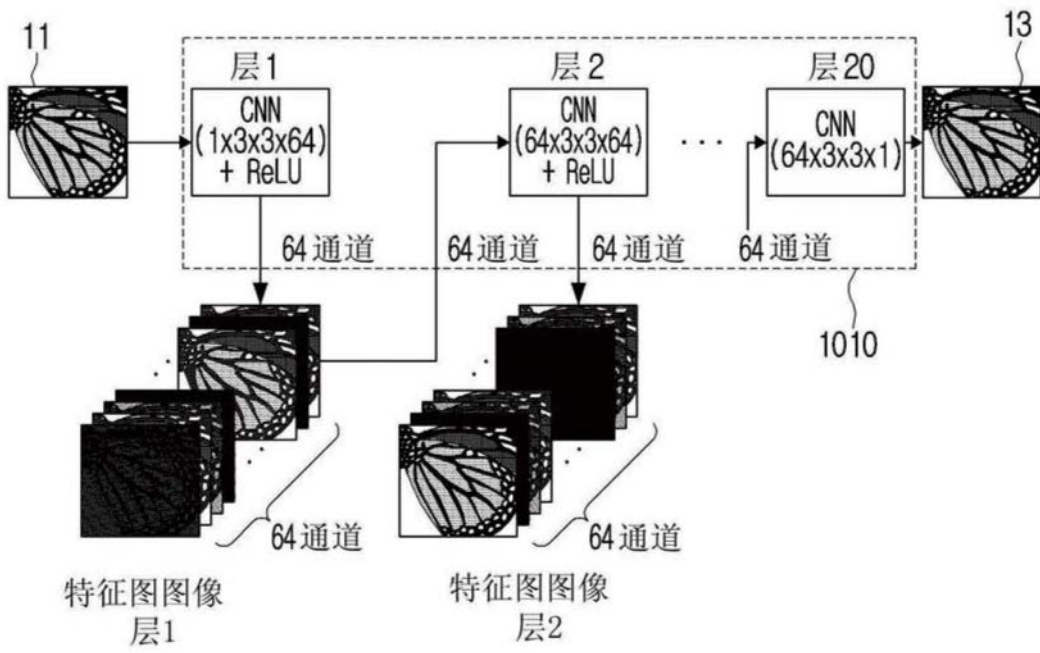


图10A

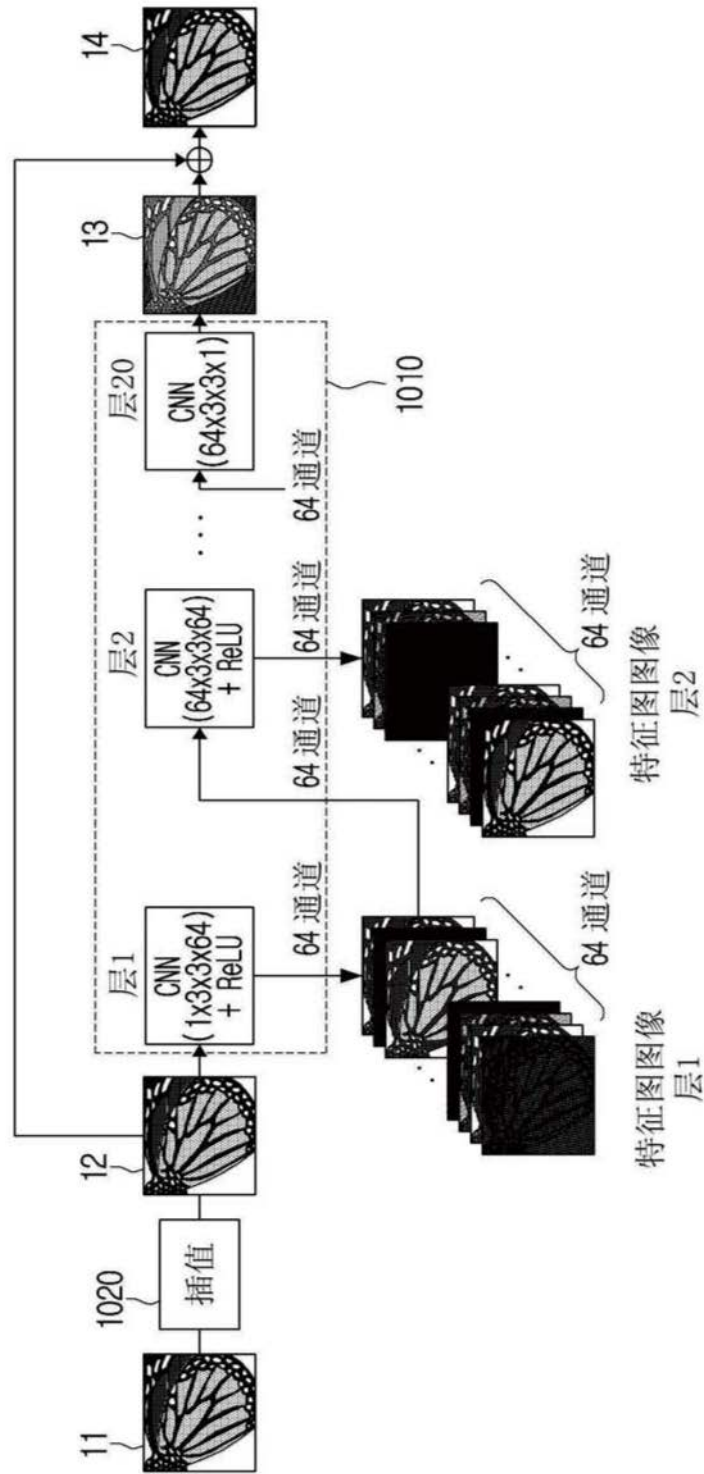


图10B

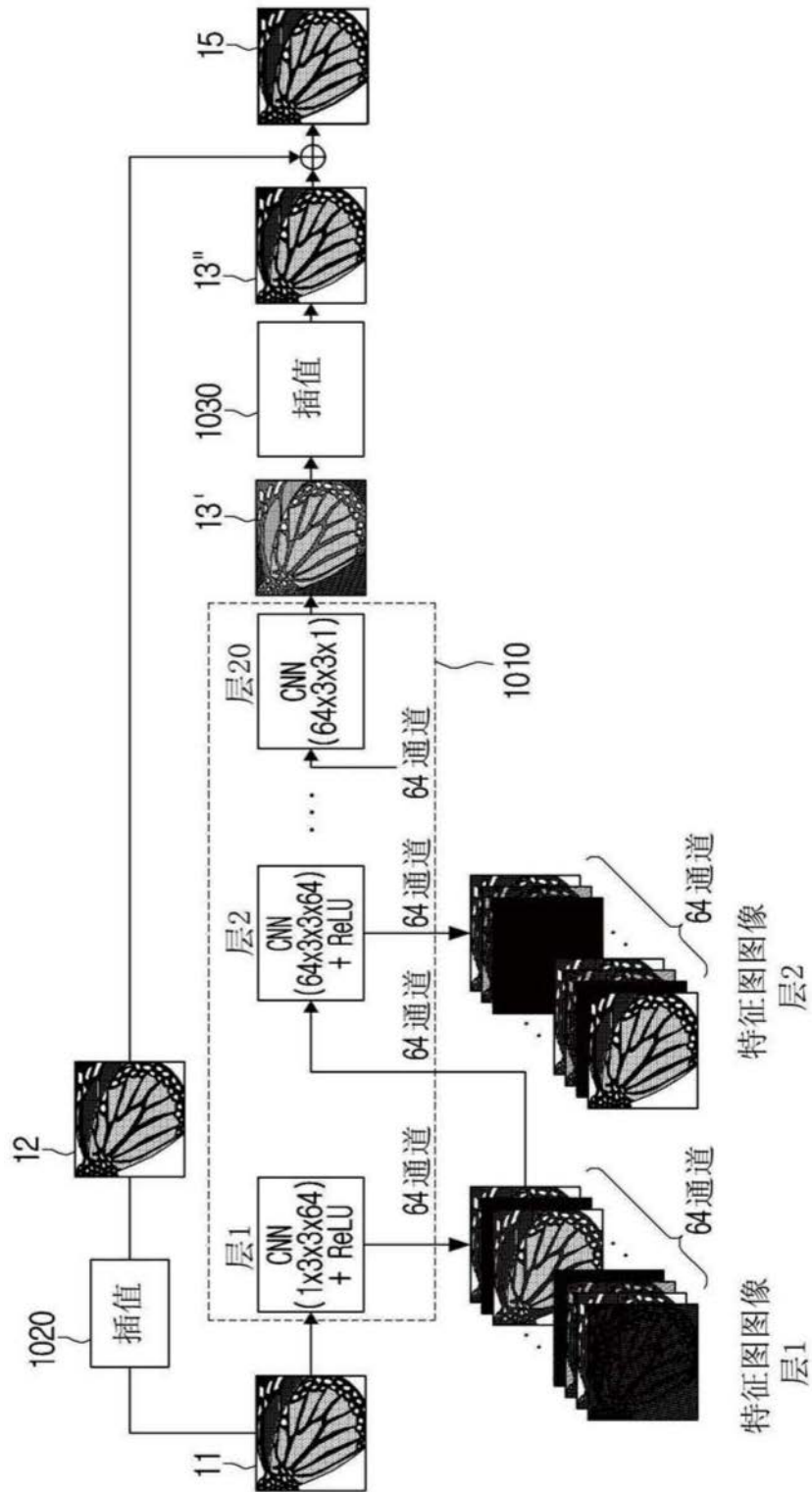


图10C

100''



图11A

200'

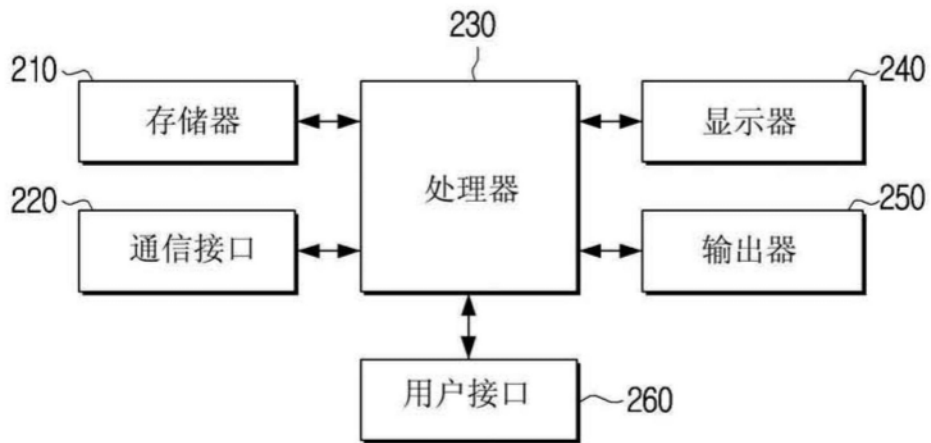


图11B

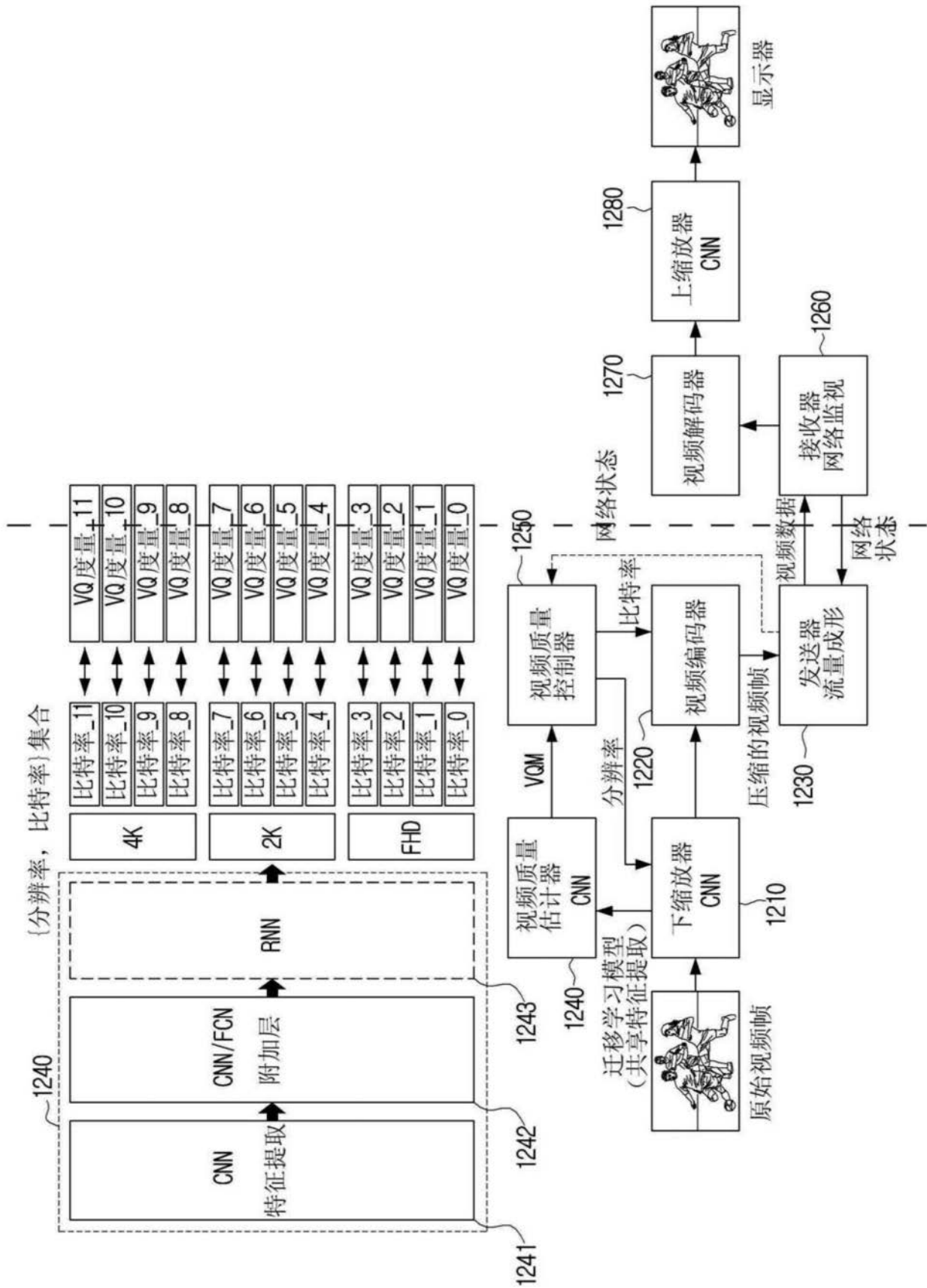


图12

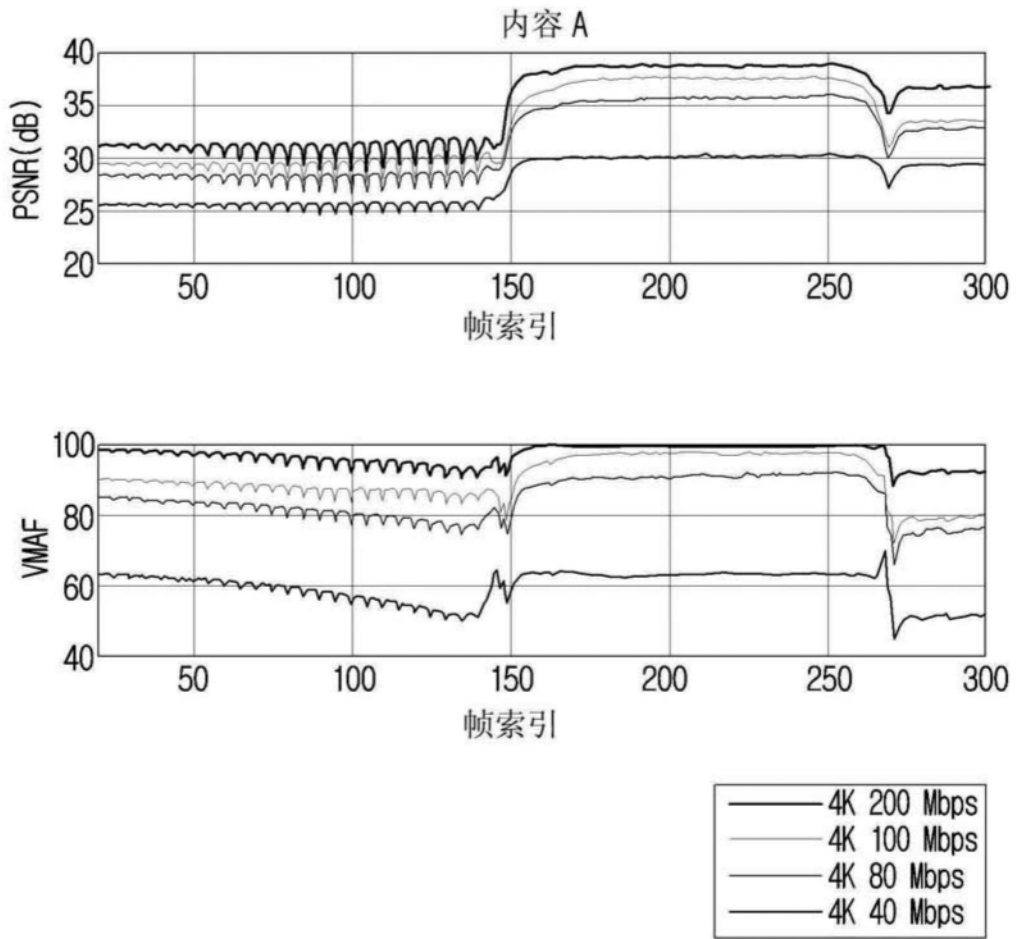


图13

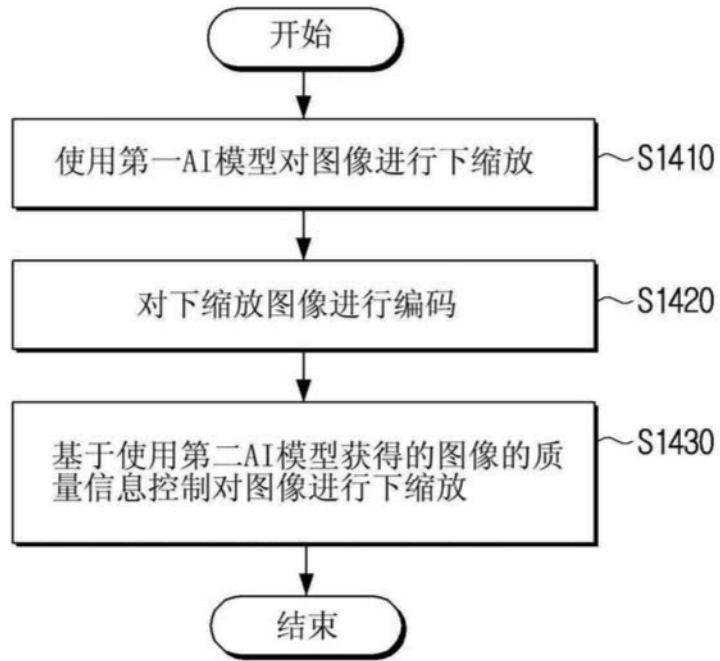


图14

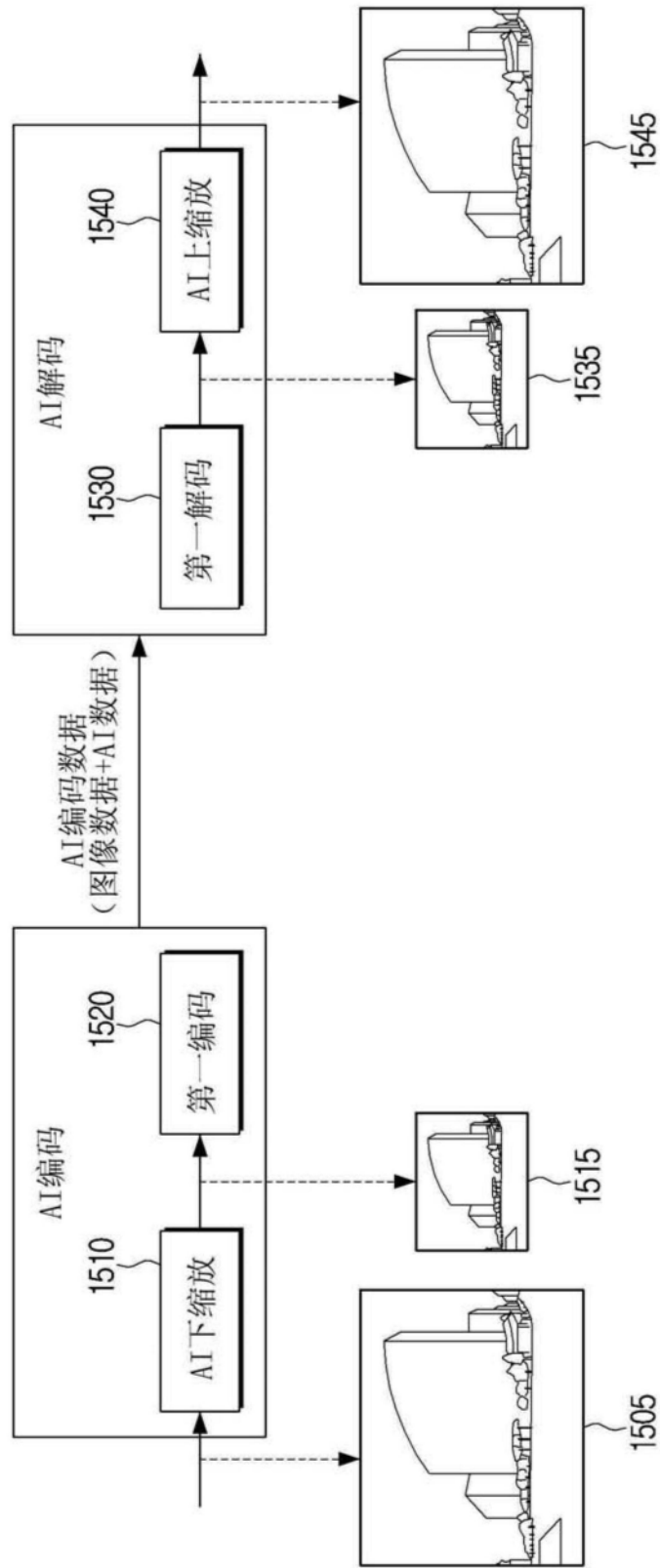


图15

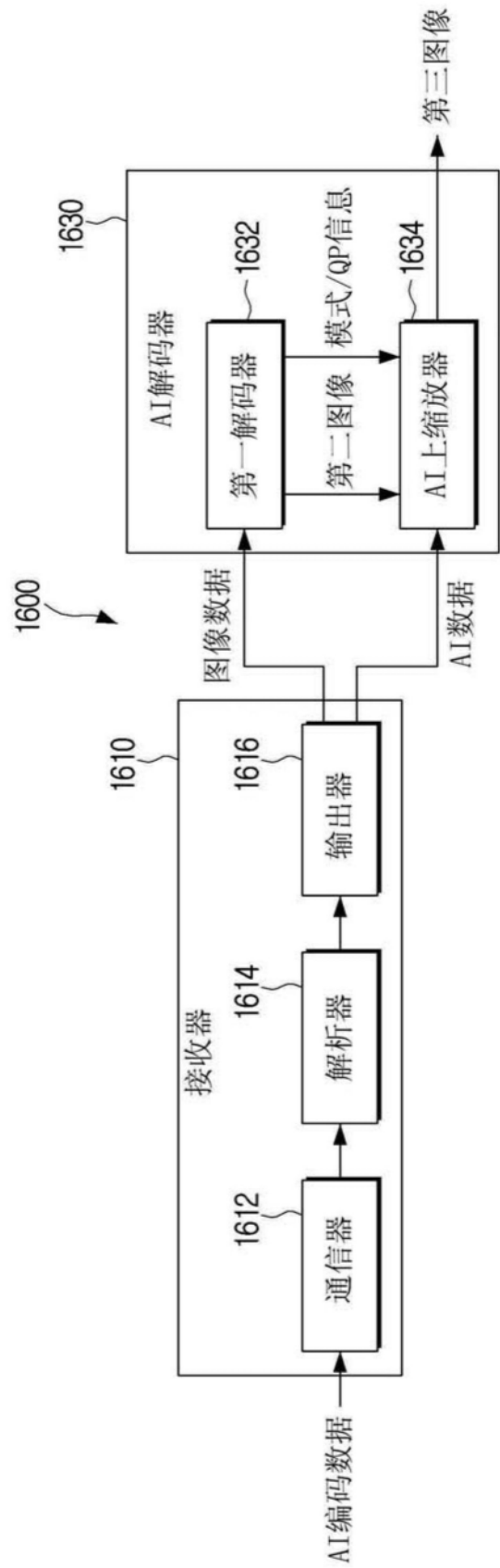


图16

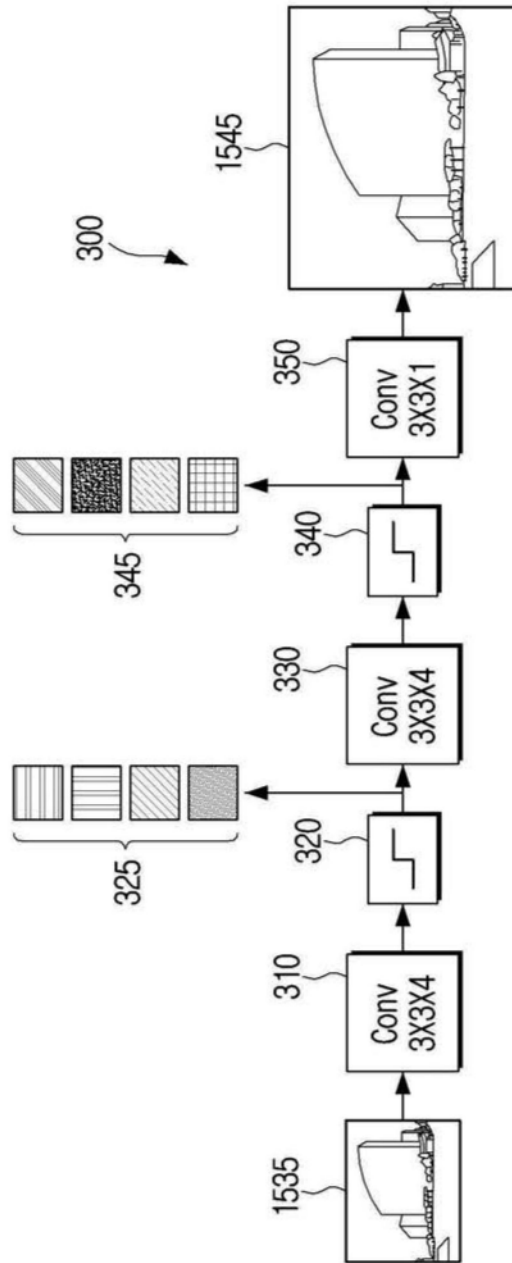


图17

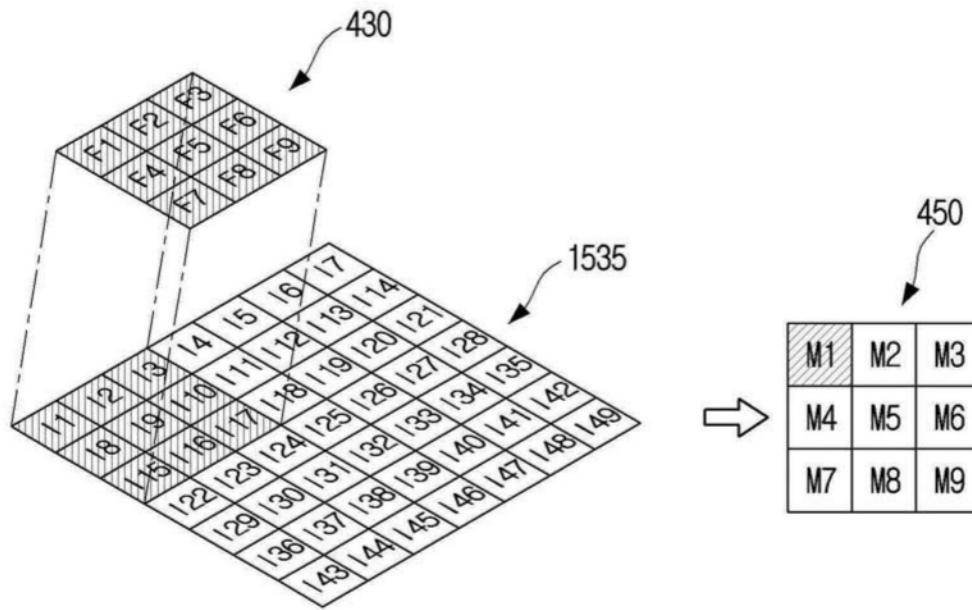


图18

图像相关信息	DNN配置信息
SD, 10Mbps, AV1	A DNN配置信息
HD, 15Mbps, H.264	B DNN配置信息
Full HD, 20Mbps, HEVC	C DNN配置信息
Full HD, 15Mbps, HEVC	D DNN配置信息

图19

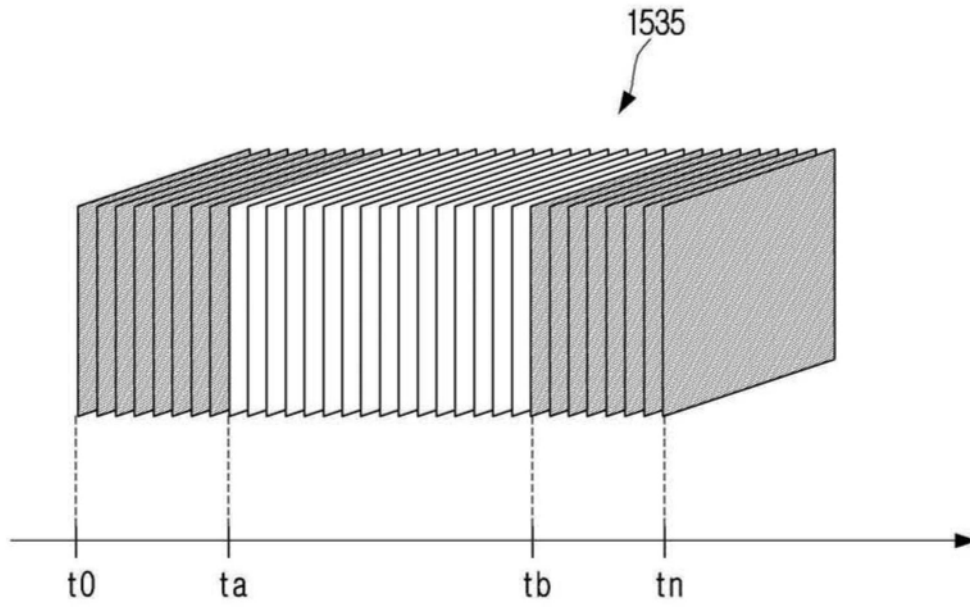


图20

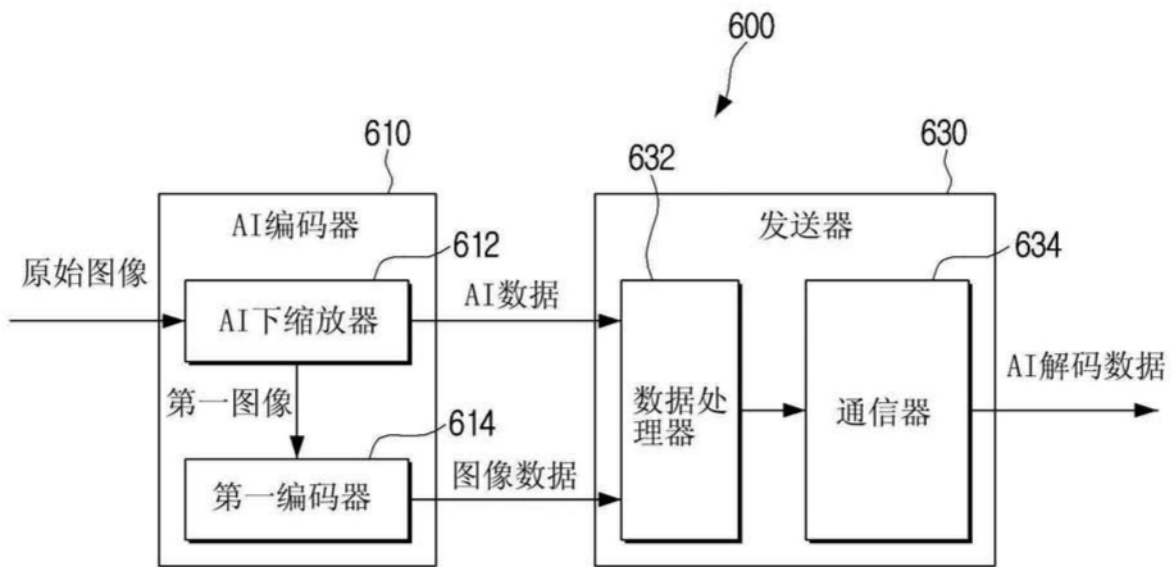


图21

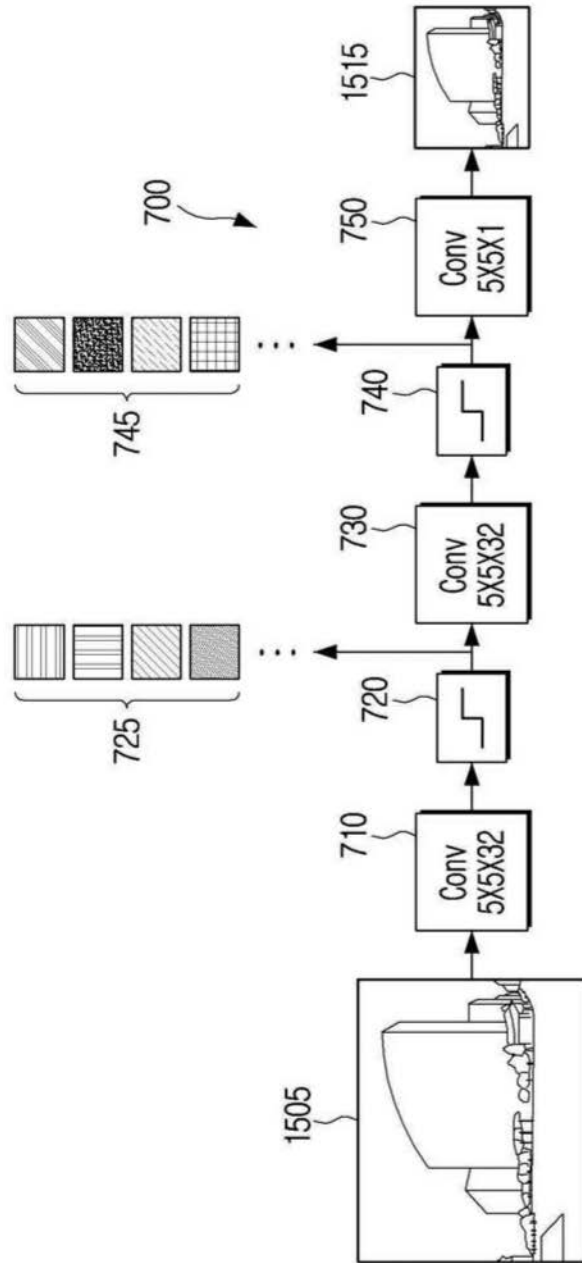


图22



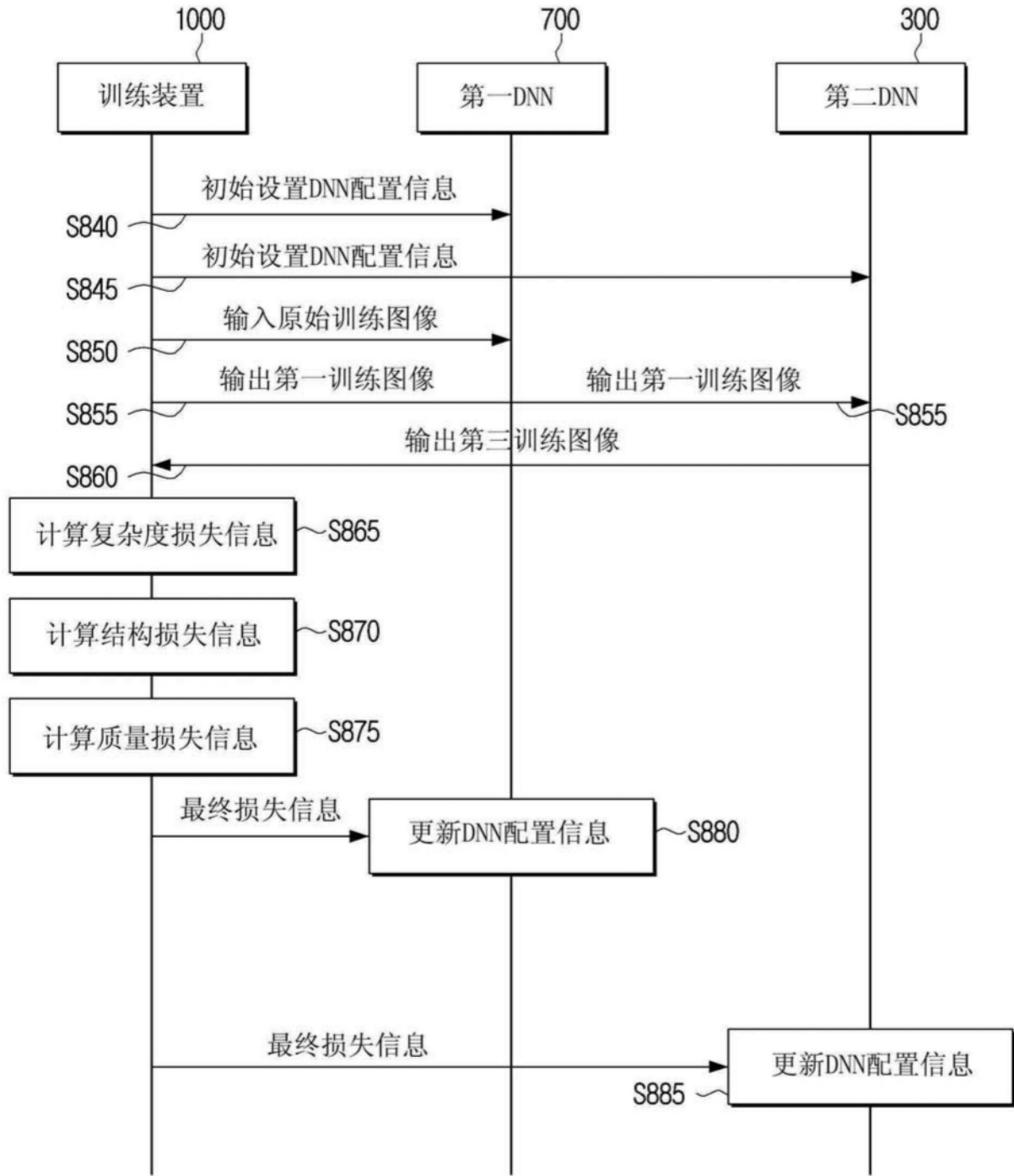


图24

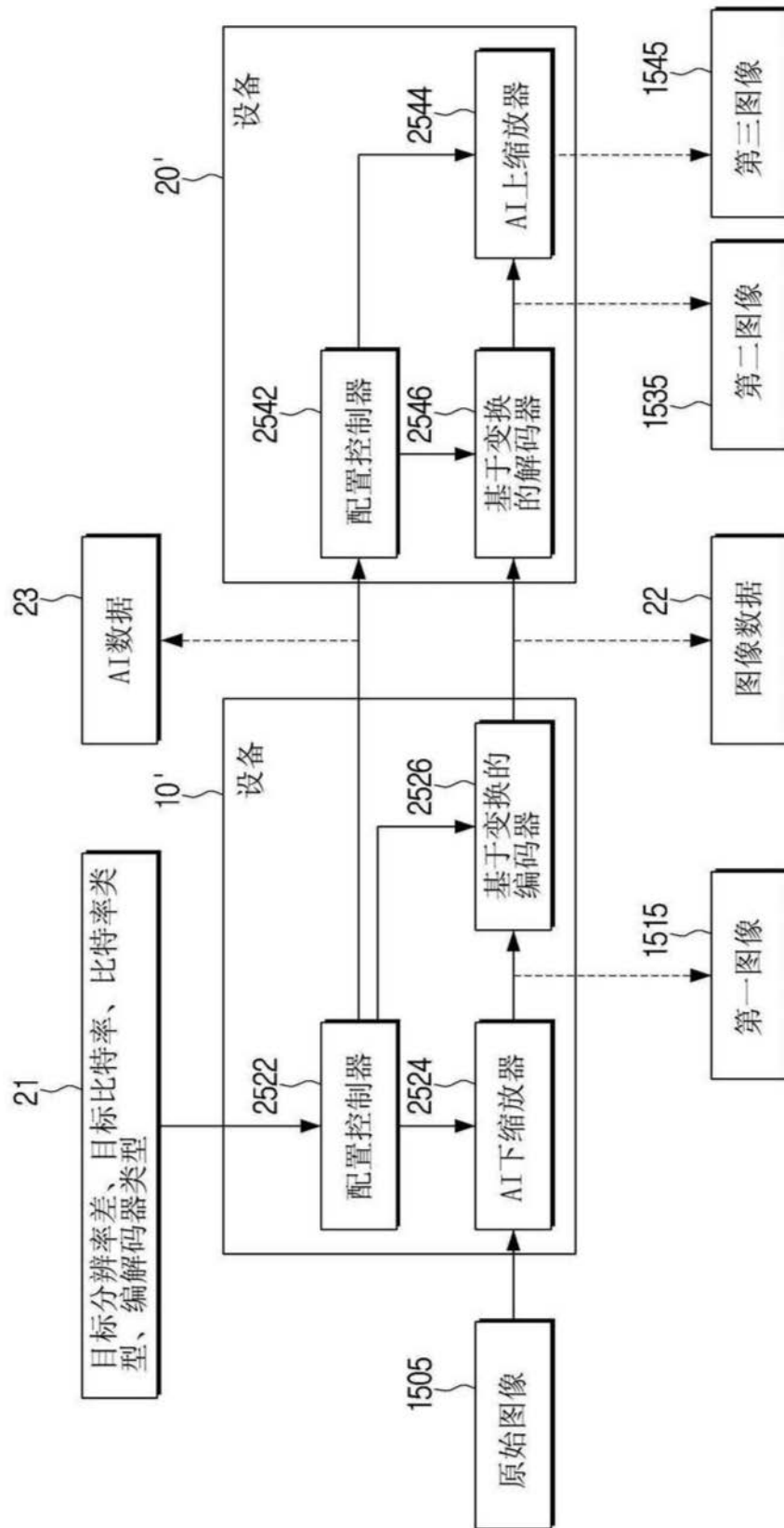


图25