(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 111182254 B
(45)授权公告日 2022.06.24

(21)申请号 202010005855.9
(22)申请日 2020.01.03
(65)同一申请的已公布的文献号
申请公布号 CN 111182254 A
(43)申请公布日 2020.05.19
(73)专利权人 北京百度网讯科技有限公司
地址 100085 北京市海淀区上地十街10号百度大厦2层
(72)发明人 李超 文石磊 丁二锐
(74)专利代理机构 北京品源专利代理有限公司 11332
专利代理师 孟金喆
(51)Int.Cl.
H04N 7/01(2006.01)
G06T 3/40(2006.01)
G06T 5/50(2006.01)

(56)对比文件
CN 109272450 A,2019.01.25
CN 104778659 A,2015.07.15
CN 107578377 A,2018.01.12
CN 106709875 A,2017.05.24
CN 109961396 A,2019.07.02
CN 108550115 A,2018.09.18
CN 110473144 A,2019.11.19
US 2019370608 A1,2019.12.05

审查员 李维维

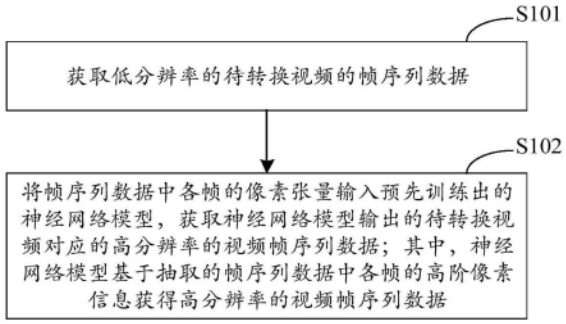
权利要求书3页 说明书12页 附图3页

(54)发明名称

一种视频处理方法、装置、设备及存储介质

(57)摘要

本申请公开了一种视频处理方法、装置、设备及存储介质，涉及计算机视觉技术领域。具体实现方案为：获取低分辨率的待转换视频的帧序列数据；将帧序列数据中各帧的像素张量输入预先训练出的神经网络模型，获取神经网络模型输出的待转换视频对应的高分辨率的视频帧序列数据；其中，神经网络模型基于抽取的帧序列数据中各帧的高阶像素信息获得高分辨率的视频帧序列数据。本申请实施例的技术方案在进行视频分辨率转换时，借助视频中的高阶像素信息，使得在神经网络进行数据处理的过程中，能够有效保留图像边缘及纹理等信息，从而实现了边缘、纹理信息的保真，避免了图片过度平滑的情况，使得画面更加自然，进而提升了超分辨率的视觉效果。



CN 111182254 B

1.一种视频处理方法,其特征在于,包括:

获取低分辨率的待转换视频的帧序列数据;

将所述帧序列数据中各帧的像素张量输入预先训练出的神经网络模型,获取所述神经网络模型输出的所述待转换视频对应的高分辨率的视频帧序列数据;所述神经网络模型在训练过程中输出的帧序列数据与真实高分辨率帧序列数据之间的距离偏差趋于稳定时训练完成;

其中,所述神经网络模型基于抽取的所述帧序列数据中各帧的高阶像素信息获得所述高分辨率的视频帧序列数据;

所述高阶像素信息包括帧序列数据中的帧图片的图像边缘信息以及纹理信息;

初始建立的神经网络模型包括:顺序连接的高阶像素特征学习模块、时序特征建模模块、多帧上采样模块和高阶像素损失建模模块;其中,

所述高阶像素特征学习模块,用于基于样本视频的帧序列数据中各帧的像素张量计算各帧的高阶像素信息;并根据所述高阶像素信息和所述像素张量,得到所述样本视频的帧序列数据中各帧的特征张量;

所述时序特征建模模块,用于对所述样本视频的帧序列数据中各帧的特征张量,在时间维度上进行特征变换,以使变换后各帧的时序特征参考相邻帧;

所述多帧上采样模块,用于基于所述变换后的所述样本视频的帧序列数据中各帧的特征张量,对所述样本视频的帧序列数据中的各帧进行放大,得到所述样本视频的超分视频帧序列;

所述高阶像素损失建模模块,用于分别获取所述超分视频帧序列中各帧的第一高阶像素信息,和真实高分辨率帧序列数据中各帧的第二高阶像素信息,并根据所述第一高阶像素信息与所述第二高阶像素信息之间的第一误差、以及所述超分视频帧序列中各帧的像素张量与所述真实高分辨率帧序列数据中各帧的像素张量之间的第二误差,调整神经网络模型中待学习的网络参数。

2.根据权利要求1所述的方法,其特征在于,在获取低分辨率的待转换视频的帧序列数据之前,所述方法还包括:

获取低分辨率的样本视频的帧序列数据,以及所述样本视频的真实高分辨率帧序列数据;

将获取的所述样本视频的帧序列数据和真实高分辨率帧序列数据,作为样本数据对初始建立的神经网络模型进行训练。

3.根据权利要求1所述的方法,其特征在于,所述高阶像素特征学习模块包括:

高阶像素信息抽取层,用于基于所述样本视频的帧序列数据中各帧的像素张量,计算各帧的高阶像素信息;

高阶特征信息学习层,用于对由所述高阶像素信息构成的高阶像素张量和所述像素张量进行非线性变换,得到所述样本视频的帧序列数据中各帧的特征张量。

4.根据权利要求1所述的方法,其特征在于,所述神经网络模型中包括多组高阶像素特征学习模块和时序特征建模模块顺序连接的组合模块,每组所述组合模块顺序连接,最后一组组合模块的时序特征建模模块与所述多帧上采样模块相连。

5.根据权利要求3所述的方法,其特征在于,所述高阶像素信息抽取层包括至少一个卷

积权重确定的二维卷积层；

所述高阶特征信息学习层包括至少一个卷积权重需通过训练确定的三维卷积层；

所述时序特征建模模块包括至少一个三维卷积层；

所述多帧上采样模块包括至少一个卷积权重需通过训练确定的二维卷积层和像素重排列层；

所述高阶像素损失建模模块包括高阶像素信息抽取层和计算模块。

6.根据权利要求1-5中任一项所述的方法，其特征在于，所述预先训练出的神经网络模型包括：顺序连接的高阶像素特征学习模块、时序特征建模模块和多帧上采样模块；

所述高阶像素特征学习模块，用于基于所述待转换视频的帧序列数据中各帧的像素张量计算各帧的高阶像素信息，并根据该高阶像素信息和该像素张量，得到所述待转换视频的帧序列数据中各帧的特征张量；

所述时序特征建模块，用于对所述待转换视频的帧序列数据中各帧的特征张量，在时间维度上进行特征变换，以使变换后各帧的时序特征参考相邻帧；

所述多帧上采样模块，用于基于所述变换后的所述待转换视频的帧序列数据中各帧的特征张量，对所述待转换视频的帧序列数据中的各帧进行放大，得到所述待转换视频的高分辨率的视频帧序列数据。

7.根据权利要求1-5任一项所述的方法，其特征在于，所述高阶像素信息包括各像素点与周围像素点的二阶导数信息。

8.一种视频处理装置，其特征在于，包括：

帧序列数据获取单元，用于获取低分辨率的待转换视频的帧序列数据；

帧序列数据转换单元，用于将所述帧序列数据中各帧的像素张量输入预先训练出的神经网络模型，获取所述神经网络模型输出的所述待转换视频对应的高分辨率的视频帧序列数据；所述神经网络模型在训练过程中输出的帧序列数据与真实高分辨率帧序列数据之间的距离偏差趋于稳定时训练完成；

其中，所述神经网络模型基于抽取的所述帧序列数据中各帧的高阶像素信息获得所述高分辨率的视频帧序列数据；

所述高阶像素信息包括帧序列数据中的帧图片的图像边缘信息以及纹理信息；

初始建立的神经网络模型包括：顺序连接的高阶像素特征学习模块、时序特征建模模块、多帧上采样模块和高阶像素损失建模模块；其中，

所述高阶像素特征学习模块，用于基于样本视频的帧序列数据中各帧的像素张量计算各帧的高阶像素信息；并根据所述高阶像素信息和所述像素张量，得到所述样本视频的帧序列数据中各帧的特征张量；

所述时序特征建模模块，用于对所述样本视频的帧序列数据中各帧的特征张量，在时间维度上进行特征变换，以使变换后各帧的时序特征参考相邻帧；

所述多帧上采样模块，用于基于所述变换后的所述样本视频的帧序列数据中各帧的特征张量，对所述样本视频的帧序列数据中的各帧进行放大，得到所述样本视频的超分视频帧序列；

所述高阶像素损失建模模块，用于分别获取所述超分视频帧序列中各帧的第一高阶像素信息，和真实高分辨率帧序列数据中各帧的第二高阶像素信息，并根据所述第一高阶像

素信息与所述第二高阶像素信息之间的第一误差、以及所述超分视频帧序列中各帧的像素张量与所述真实高分辨率帧序列数据中各帧的像素张量之间的第二误差，调整神经网络模型中待学习的网络参数。

9.一种电子设备，其特征在于，包括：

至少一个处理器；以及

与所述至少一个处理器通信连接的存储器；其中，

所述存储器存储有可被所述至少一个处理器执行的指令，所述指令被所述至少一个处理器执行，以使所述至少一个处理器能够执行权利要求1-7中任一项所述的一种视频处理方法。

10.一种存储有计算机指令的非瞬时计算机可读存储介质，其特征在于，所述计算机指令用于使所述计算机执行权利要求1-7中任一项所述的一种视频处理方法。

一种视频处理方法、装置、设备及存储介质

技术领域

[0001] 本申请涉及数据处理技术，尤其涉及计算机视觉技术领域。

背景技术

[0002] 视频超分辨技术可以将低分辨率的视频转换成高分辨率的视频。

[0003] 现有技术中，通常采用神经网络模型将低分辨率的视频转化成高分辨率的视频。而神经网络模型单纯的以重构像素损失函数作为监督信息，会产生画面过度平滑等现象，致使画面视觉效果较差。

发明内容

[0004] 本申请实施例提供了一种视频处理方法、装置、涉及及存储介质，以避免转换后的视频画面过度平滑，使画面更加自然，提升超分辨率的视觉效果。

[0005] 第一方面，本申请实施例提供了一种视频处理方法，包括：

[0006] 获取低分辨率的待转换视频的帧序列数据；

[0007] 将所述帧序列数据中各帧的像素张量输入预先训练出的神经网络模型，获取所述神经网络模型输出的所述待转换视频对应的高分辨率的视频帧序列数据；

[0008] 其中，所述神经网络模型基于抽取的所述帧序列数据中各帧的高阶像素信息获得所述高分辨率的视频帧序列数据。

[0009] 本申请实施例通过获取低分辨率的待转换视频的帧序列数据；将帧序列数据中各帧的像素张量输入预先训练出的神经网络模型，并由神经网络模型基于抽取的帧序列数据中各帧的高阶像素信息，获得高分辨率的视频帧序列数据并输出，从而将低分辨率的待转换视频的帧序列数据转换为高分辨率的视频帧序列数据。上述技术方案在进行视频分辨率转换时，借助视频中的高阶像素信息，使得在神经网络进行数据处理的过程中，能够有效保留图像边缘及纹理等信息，从而实现了边缘、纹理信息的保真，避免了图片过度平滑的情况，使得画面更加自然，进而提升了超分辨率的视觉效果。

[0010] 可选的，在获取低分辨率的待转换视频的帧序列数据之前，所述方法还包括：

[0011] 获取低分辨率的样本视频的帧序列数据，以及所述样本视频的真实高分辨率帧序列数据；

[0012] 将获取的所述样本视频的帧序列数据和真实高分辨率帧序列数据，作为样本数据对初始建立的神经网络模型进行训练。

[0013] 上述申请中的一个可选实施方式，通过在获取低分辨率的待转换视频的帧序列数据之前，追加神经网络模型的训练操作，并具体通过获取低分辨率的样本视频的帧序列数据以及相应的真实高分辨率帧序列数据，作为样本数据对初始建立的神经网络模型进行训练，从而完善了神经网络模型的训练机制，为低分辨率的视频向高分辨率视频的转换提供保障。

[0014] 可选的，所述初始建立的神经网络模型包括：顺序连接的高阶像素特征学习模块、

时序特征建模模块、多帧上采样模块和高阶像素损失建模模块;其中,

[0015] 所述高阶像素特征学习模块,用于基于所述样本视频的帧序列数据中各帧的像素张量计算各帧的高阶像素信息;并根据所述高阶像素信息和所述像素张量,得到所述样本视频的帧序列数据中各帧的特征张量;

[0016] 所述时序特征建模模块,用于对所述样本视频的帧序列数据中各帧的特征张量,在时间维度上进行特征变换,以使变换后各帧的时序特征参考相邻帧;

[0017] 所述多帧上采样模块,用于基于所述变换后的所述样本视频的帧序列数据中各帧的特征张量,对所述样本视频的帧序列数据中的各帧进行放大,得到所述样本视频的超分视频帧序列;

[0018] 所述高阶像素损失建模模块,用于分别获取所述超分视频帧序列中各帧的第一高阶像素信息,和所述真实高分辨率帧序列数据中各帧的第二高阶像素信息,并根据所述第一高阶像素信息与所述第二高阶像素信息之间的第一误差、以及所述超分视频帧序列中各帧的像素张量与所述真实高分辨率帧序列数据中各帧的像素张量之间的第二误差,调整神经网络模型中待学习的网络参数。

[0019] 上述申请中的一个可选实施方式,通过细化神经网络模型的结构组成,以及各部分结构的功能,完善了神经网络模型中的数据处理机制,从而保证基于顺序连接的高阶像素特征学习模块、时序特征及建模模块、多帧上采样模块和高阶像素损失建模模块,能够达到将低分辨率视频向高分辨率视频转换的效果,同时保证转换结果的边缘、纹理保真,画面自然。

[0020] 可选的,所述高阶像素特征学习模块包括:

[0021] 高阶像素信息抽取层,用于基于所述样本视频的帧序列数据中各帧的像素张量,计算各帧的高阶像素信息;

[0022] 高阶特征信息学习层,用于对由所述高阶像素信息构成的高阶像素张量和所述像素张量进行非线性变换,得到所述样本视频的帧序列数据中各帧的特征张量。

[0023] 上述申请中的一个可选实施方式,通过将高阶像素特征学习模块划分为进行高阶像素信息的高阶像素信息抽取层,以及进行非线性变换的高阶特征信息学习层,从而实现两部分功能的有效隔离,同时使得所抽取的特征在包含边缘、纹理等高阶信息的同时,能够引入非线性细节,实现了对特征信息的进一步挖掘,使得所提取的特征信息更具代表性,为高分辨率视频的边缘、纹理等信息的高度保真奠定基础。

[0024] 可选的,所述神经网络模型中包括多组高阶像素特征学习模块和时序特征建模模块顺序连接的组合模块,每组所述组合模块顺序连接,最后一组组合模块的时序特征建模模块与所述多帧上采样模块相连。

[0025] 上述申请中的一个可选实施方式,通过在时序特征建模模块与多帧上采样模块之间,顺序连接多组高阶像素特征学习模块和时序特征建模模块,从而提高了所提取的特征信息的准确度,为高分辨率视频的边缘、纹理等信息的高度保真奠定基础。

[0026] 可选的,所述高阶像素信息抽取层包括至少一个卷积权重确定的二维卷积层;

[0027] 所述高阶特征信息学习层包括至少一个卷积权重需通过训练确定的三维卷积层;

[0028] 所述时序特征建模模块包括至少一个三维卷积层;

[0029] 所述多帧上采样模块包括至少一个卷积权重需通过训练确定的二维卷积层和像

素重排列层；

[0030] 所述高阶像素损失建模模块包括高阶像素信息抽取层和计算模块。

[0031] 上述申请中的一个可选实施方式，通过细化神经网络模型中各模块组成，完善了不同模块的运行机制，为各模块功能的实现提供技术支撑。

[0032] 可选的，所述预先训练出的神经网络模型包括：顺序连接的高阶像素特征学习模块、时序特征建模模块和多帧上采样模块；

[0033] 所述高阶像素特征学习模块，用于基于所述待转换视频的帧序列数据中各帧的像素张量计算各帧的高阶像素信息，并根据该高阶像素信息和该像素张量，得到所述待转换视频的帧序列数据中各帧的特征张量；

[0034] 所述时序特征建模块，用于对所述待转换视频的帧序列数据中各帧的特征张量，在时间维度上进行特征变换，以使变换后各帧的时序特征参考相邻帧；

[0035] 所述多帧上采样模块，用于基于所述变换后的所述待转换视频的帧序列数据中各帧的特征张量，对所述待转换视频的帧序列数据中的各帧进行放大，得到所述待转换视频的高分辨率的视频帧序列数据。

[0036] 上述申请中的一个可选实施方式，通过明确训练好的神经网络模型内部各模块间的连接关系，并且完善不同模块对待转换视频对应信息的处理机制，为低分辨率的待转换视频向高分辨率视频的转换提供了保障。

[0037] 可选的，所述高阶像素信息包括各像素点与周围像素点的二阶导数信息。

[0038] 上述申请中的一个可选实施方式，完善了高阶像素信息的构成方案，同时通过对视频帧序列数据的各帧中的各像素点与周围像素点的二阶导数信息的确定，有效提取帧图像中的边缘、纹理信息等，为高阶像素信息的构成提供数据支撑。

[0039] 第二方面，本申请实施例还提供了一种视频处理装置，包括：

[0040] 帧序列数据获取单元，用于获取低分辨率的待转换视频的帧序列数据；

[0041] 帧序列数据转换单元，用于将所述帧序列数据中各帧的像素张量输入预先训练出的神经网络模型，获取所述神经网络模型输出的所述待转换视频对应的高分辨率的视频帧序列数据；

[0042] 其中，所述神经网络模型基于抽取的所述帧序列数据中各帧的高阶像素信息获得所述高分辨率的视频帧序列数据。

[0043] 第三方面，本申请实施例还提供了一种电子设备，包括：

[0044] 至少一个处理器；以及

[0045] 与所述至少一个处理器通信连接的存储器；其中，

[0046] 所述存储器存储有可被所述至少一个处理器执行的指令，所述指令被所述至少一个处理器执行，以使所述至少一个处理器能够执行如第一方面实施例所提供的一种视频处理方法。

[0047] 第四方面，本申请实施例还提供了一种存储有计算机指令的非瞬时计算机可读存储介质，其特征在于，所述计算机指令用于使所述计算机执行如第一方面实施例所提供的一种视频处理方法。

[0048] 上述可选方式所具有的其他效果将在下文中结合具体实施例加以说明。

附图说明

[0049] 附图用于更好地理解本方案,不构成对本申请的限定。其中:

[0050] 图1是本申请实施例一中的一种视频处理方法的流程图;

[0051] 图2A是本申请实施例二中的一种视频处理方法的流程图;

[0052] 图2B是本申请实施例二中的一种初始构建的神经网络模型的结构图;

[0053] 图2C是本申请实施例二中的一种初始构建的改进后的神经网络模型的结构图;

[0054] 图3是本申请实施例三中的一种视频处理装置的结构图;

[0055] 图4是用来实现本申请实施例的视频处理方法的电子设备的框图。

具体实施方式

[0056] 以下结合附图对本申请的示范性实施例做出说明,其中包括本申请实施例的各种细节以助于理解,应当将它们认为仅仅是示范性的。因此,本领域普通技术人员应当认识到,可以对这里描述的实施例做出各种改变和修改,而不会背离本申请的范围和精神。同样,为了清楚和简明,以下的描述中省略了对公知功能和结构的描述。

[0057] 实施例一

[0058] 图1是本申请实施例一中的一种视频处理方法的流程图,本申请实施例适用于将低分辨率的视频转化为高分辨率视频的情况,该方法由视频处理装置执行,该装置由软件和/或硬件实现,并具体配置于具备一定数据运算能力的电子设备中。

[0059] 如图1所示的一种视频处理方法,包括:

[0060] S101、获取低分辨率的待转换视频的帧序列数据。

[0061] 其中,帧序列数据可以理解为待转换视频中按照视频播放顺序所排列的帧图片的集合。

[0062] 示例性地,可以在电子设备本地、与电子设备关联的其他存储设备或云端中预先存储待转换视频的帧序列数据,并在需要时直接从相应的存储位置进行待转换视频的帧序列数据的获取。

[0063] S102、将所述帧序列数据中各帧的像素张量输入预先训练出的神经网络模型,获取所述神经网络模型输出的所述待转换视频对应的高分辨率的视频帧序列数据;其中,所述神经网络模型基于抽取的所述帧序列数据中各帧的高阶像素信息获得所述高分辨率的视频帧序列数据。

[0064] 将帧序列数据中各帧的像素张量,作为输入数据输入至预先训练出的神经网络模型中;神经网络模型抽取帧序列数据中各帧的高阶像素信息,并基于所抽取的高阶像素信息将低分辨率的待转换视频,转化为高分辨率的视频帧序列数据,并加以输出。

[0065] 其中,神经网络模型可以预先通过大量的低分辨率的样本视频的帧序列数据进行训练得到。在所训练的神经网络模型中,会预先对样本视频的帧序列数据进行高阶像素信息的提取,并基于提取的高阶像素信息对神经网络模型中的模型参数进行训练,并通过不断调整模型参数,使得模型的输出结果与样本视频对应的真实高分辨率帧序列数据之间的距离偏差逐渐减小并趋于稳定。

[0066] 在本申请实施例的一个可选实施方式中,高阶像素信息包括帧序列数据中的帧图片的图像边缘以及纹理等信息。可选的,图片边缘以及纹理等信息可以通过对帧图片中各

像素点与周围像素点的二阶导数确定，并将得到的二阶导数信息作为高阶像素信息。

[0067] 示例性地，二阶导数信息可以通过二维卷积核对像素张量在宽、高维度上进行卷积操作得到。其中，二维卷积核可以由技术人员根据需要或经验值进行设定，还可以通过大量试验反复确定。

[0068] 需要说明的是，由于本申请通过对待转换视频的帧序列数据中各帧的高阶像素信息进行处理，实现了低分辨率的视频向高分辨率视频的转换，无需利用光流或动作补偿等技术，减少了视频处理过程的数据计算量，同时避免了额外误差的引入所导致的超分效果退化现象的发生。

[0069] 本申请实施例通过获取低分辨率的待转换视频的帧序列数据；将帧序列数据中各帧的像素张量输入预先训练出的神经网络模型，并由神经网络模型基于抽取的帧序列数据中各帧的高阶像素信息，获得高分辨率的视频帧序列数据并输出，从而将低分辨率的待转换视频的帧序列数据转换为高分辨率的视频帧序列数据。上述技术方案在进行视频分辨率转换时，借助视频中的高阶像素信息，使得在神经网络进行数据处理的过程中，能够有效保留图像边缘及纹理等信息，从而实现了边缘、纹理信息的保真，避免了图片过度平滑的情况，使得画面更加自然，进而提升了超分辨率的视觉效果。

[0070] 实施例二

[0071] 图2A是本申请实施例二中的一种视频处理方法的流程图，本申请实施例在上述各实施例的技术方案的基础上，进行了优化改进。

[0072] 进一步地，在操作“获取低分辨率的待转换视频的帧序列数据”之前，追加“获取低分辨率的样本视频的帧序列数据，以及所述样本视频的真实高分辨率帧序列数据；将获取的所述样本视频的帧序列数据和真实高分辨率帧序列数据，作为样本数据对初始建立的神经网络模型进行训练”，以完善神经网络模型的训练机制。

[0073] 如图2A所示的一种视频处理方法，包括：

[0074] S201、获取低分辨率的样本视频的帧序列数据，以及所述样本视频的真实高分辨率帧序列数据。

[0075] S202、将获取的所述样本视频的帧序列数据和真实高分辨率帧序列数据，作为样本数据对初始建立的神经网络模型进行训练。

[0076] 在本申请实施例的一种可选实施方式中，通过低分辨率的样本视频的帧序列数据对神经网络模型的模型参数进行训练，通过不断调整模型参数的参数值，使得模型输出的帧序列数据与真实高分辨率帧序列数据之间的距离偏差逐渐减小，直至距离偏差趋于稳定，得到训练好的神经网络模型。

[0077] 参见图2B所示的初始构建的神经网络模型的模型结构图。其中，所述初始建立的神经网络模型包括：顺序连接的高阶像素特征学习模块10、时序特征建模模块20、多帧上采样模块30和高阶像素损失建模模块40。

[0078] 示例性地，高阶像素特征学习模块10，用于基于所述样本视频的帧序列数据中各帧的像素张量计算各帧的高阶像素信息；并根据所述高阶像素信息和所述像素张量，得到所述样本视频的帧序列数据中各帧的特征张量。

[0079] 为了实现高阶像素学习模块10中不能功能的有效隔离，还可以通过设置数据层的方式，对高阶像素学习模块10进行划分。具体的，高阶像素特征学习模块10包括高阶像素信

息抽取层11和高阶特征信息学习层12。其中，高阶像素信息抽取层11，用于基于所述样本视频的帧序列数据中各帧的像素张量，计算各帧的高阶像素信息，以实现对样本视频的帧序列数据中高阶特征的提取。其中，高阶特征信息学习层12，用于对由所述高阶像素信息构成的高阶像素张量和所述像素张量进行非线性变换，得到所述样本视频的帧序列数据中各帧的特征张量，以在提取的高阶特征中引入非线性信息，从而实现了对提取的高阶特征进行深度挖掘，凸显了所提取的特征信息的代表性和准确性，为转换后的视频的像素质量的提高提供准备。

[0080] 可选的，在高阶像素信息抽取层11中设置有至少一个卷积权重确定的二维卷积层，用于进行高阶像素信息的计算操作，以进行高阶特征的提取，从而降低特征维度。典型的，在计算各帧的高阶像素信息时，通过二阶卷积核对样本视频的帧序列数据中各帧的像素张量在宽、高维度上进行卷积操作得到。其中，二阶卷积核可以由技术人员根据需要或经验值进行设定，还可以通过大量试验反复确定。

[0081] 典型的，在计算高阶像素信息时，所采用的二阶卷积核为：$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$。

[0082] 可选的，在高阶特征信息学习层12中设置有至少一个卷积权重需要训练确定的三维卷积层，以将所提取的高阶像素信息映射至高维空间。

[0083] 示例性地，所述时序特征建模模块20，用于对所述样本视频的帧序列数据中各帧的特征张量，在时间维度上进行特征变换，以使变换后各帧的时序特征参考相邻帧。

[0084] 可选的，时序特征建模模块20中设置有至少一个三维卷积层，对样本视频的帧序列数据中各帧的特征张量，进行卷积处理，从而达到建模时序信息的目的。

[0085] 示例性地，所述多帧上采样模块30，用于基于所述变换后的所述样本视频的帧序列数据中各帧的特征张量，对所述样本视频的帧序列数据中的各帧进行放大，得到所述样本视频的超分视频帧序列。

[0086] 可选的，多帧上采样模块30设置有至少一个卷积权重需要通过训练确定的二维卷积层31和像素重排列层32。其中，二维卷积层31用于对输入的在时间维度上进行特征变换后的样本视频的帧序列数据中各帧的特征张量进行放大处理，使得输出的各帧的宽、高为样本视频的帧序列数据中尺寸的r(r>1)倍。其中，像素重排列层32用于对二维卷积层31输出的放大后的帧序列数据进行像素重排，将像素点从特征维度转换至空间维度，得到样本视频尺寸r倍的超分视频帧序列。其中，超分视频帧序列的分辨率高于样本视频的分辨率。其中，r可以由技术人员根据需要或经验值进行设定。

[0087] 示例性地，所述高阶像素损失建模模块40，用于分别获取所述超分视频帧序列中各帧的第一高阶像素信息，和所述真实高分辨率帧序列数据中各帧的第二高阶像素信息，并根据所述第一高阶像素信息与所述第二高阶像素信息之间的第一误差(比如第一均方误差)、以及所述超分视频帧序列中各帧的像素张量与所述真实高分辨率帧序列数据中各帧的像素张量之间的第二误差(比如第二均方误差)，调整神经网络模型中待学习的网络参数，具体的，以第一误差和第二误差小于设定阈值为学习目的。

[0088] 可选的，高阶像素损失建模模块40设置有高阶像素信息抽取层41和计算模块42。其中，高阶像素信息抽取层41，用于确定超分视频帧序列中各帧的第一高阶像素信息和真

实高分辨率帧序列数据中各帧的第二高阶像素信息。其中，计算模块42，用于计算第一高阶像素信息与第二高阶像素信息之间的第一误差，以及超分视频帧序列中各帧的像素张量与真实高分辨率帧序列数据中各帧的像素张量之间的第二误差；根据第一误差和第二误差，调整神经网络模型中待学习的网络参数。

[0089] 可以理解的是，为了避免神经网络模块中的代码冗余，高阶像素损失建模模块40中的高阶像素信息抽取层41，可以通过复用高阶像素特征学习模块10中的高阶像素信息抽取层11，并直接进行输出的高阶像素信息的获取即可。

[0090] 需要说明的是，计算模块32在进行神经网络模型中待学习的网络参数的调整时，可以借助梯度下降等算法加以实现，以使第一误差逐渐减小且趋于稳定，以及第二误差逐渐减小且趋于稳定。

[0091] 可以理解的是，在对神经网络模型训练过程中，还可以采用计算模块32中的第一误差和第二误差，对所训练的神经网络模型的模型性能进行评价。

[0092] 在本申请实施例的一种可选实施方式中，为了进一步提高神经网络模型进行视频分辨率转化的转化精度，还可以对图2B所示的神经网络模型进行优化改进。

[0093] 参加图2C所示的初始构建的改进后的神经网络模型，其中，神经网络模型中包括多组高阶像素特征学习模块10和时序特征建模模块20顺序连接的组合模块100，每组所述组合模块顺序100连接，最后一组组合模块100的时序特征建模模块20与所述多帧上采样模块30相连。

[0094] 其中，神经网络模型中的第一组组合模块100中的高阶像素特征学习模块10和时序特征建模模块20用于将输入的视频的帧序列数据中各帧的像素张量转化成特征张量，以实现对所输入视频的帧序列数据的特征提取和转换；神经网络模型中的第一组组合模块100之后顺次连接的其他组合模块100，用于对所转化的特征张量进一步进行学习，以提高所学习特征张量的准确度，从而提升神经网络模型的模型精度。

[0095] S203、获取低分辨率的待转换视频的帧序列数据。

[0096] S204、将所述帧序列数据中各帧的像素张量输入预先训练出的神经网络模型，获取所述神经网络模型输出的所述待转换视频对应的高分辨率的视频帧序列数据；其中，所述神经网络模型基于抽取的所述帧序列数据中各帧的高阶像素信息获得所述高分辨率的视频帧序列数据。

[0097] 在对采用上述任一实施例的技术方案所训练好的神经网络模型进行使用时，所使用的神经网络模型与训练阶段所构建的神经网络模型相似。由于在模型使用阶段无需对神经网络模型中的模型参数进行调整，因此训练出的神经网络模型中，仅包括顺序连接的高阶像素特征学习模块、时序特征建模模块和多帧上采样模块，不包括高阶像素损失建模模块。

[0098] 在本申请实施例的一种可选实施方式中，在使用训练好的神经网络模型对低分辨率的待转换视频的帧序列数据进行处理时，不同模块对帧序列数据所执行的处理操作如下：

[0099] 所述高阶像素特征学习模块，用于基于所述待转换视频的帧序列数据中各帧的像素张量计算各帧的高阶像素信息，并根据该高阶像素信息和该像素张量，得到所述待转换视频的帧序列数据中各帧的特征张量。

[0100] 示例性地，高阶像素特征学习模块包括高阶像素信息抽取层和高阶特征信息学习层。其中，高阶像素信息抽取层，用于基于待转换视频的帧序列数据中各帧的像素张量，计算各帧的高阶像素信息，以实现待转换视频的帧序列数据中高阶特征的提取。其中，高阶特征信息学习层，用于对由高阶像素信息构成的高阶像素张量和像素张量进行非线性变换，得到待转换视频的帧序列数据中各帧的特征张量，以在提取的高阶特征中引入非线性信息，从而实现了对提取的高阶特征进行深度挖掘，凸显了所提取的特征信息的代表性和准确性，为转换后的视频的像素质量的提高提供准备。

[0101] 所述时序特征建模块，用于对所述待转换视频的帧序列数据中各帧的特征张量，在时间维度上进行特征变换，以使变换后各帧的时序特征参考相邻帧。

[0102] 所述多帧上采样模块，用于基于所述变换后的所述待转换视频的帧序列数据中各帧的特征张量，对所述待转换视频的帧序列数据中的各帧进行放大，得到所述待转换视频的高分辨率的视频帧序列数据。

[0103] 示例性地，多帧上采样模块设置有至少一个卷积权重需要通过训练确定的二维卷积层和像素重排列层。其中，二维卷积层用于对输入的在时间维度上进行特征变换后的待转换视频的帧序列数据中各帧的特征张量进行放大处理，使得输出的各帧的宽、高为样本视频的帧序列数据中尺寸的r(r>1)倍。其中，像素重排列层用于对二维卷积层输出的放大后的帧序列数据进行像素重排，将像素点从特征维度转换至空间维度，得到待转换视频尺寸r倍的超分视频帧序列作为高分辨率的视频帧序列数据。

[0104] 本申请实施例通过在获取低分辨率的待转换视频的帧序列数据之前，追加神经网络模型的训练操作，并具体通过获取低分辨率的样本视频的帧序列数据，以及样本视频的真实高分辨率帧序列数据作为样本数据，对初始建立的神经网络模型进行训练，以得到训练好的神经网络模型，为待转换视频的高分辨率转换提供保障。

[0105] 实施例三

[0106] 图3是本申请实施例三中的一种视频处理装置的结构图，本申请实施例适用于将低分辨率的视频转化为高分辨率视频的情况，该装置由软件和/或硬件实现，并具体配置于具备一定数据运算能力的电子设备中。

[0107] 如图3所示的一种视频处理装置300，包括：帧序列数据获取单元301和帧序列数据转换单元302。其中，

[0108] 帧序列数据获取单元301，用于获取低分辨率的待转换视频的帧序列数据；

[0109] 帧序列数据转换单元302，用于将所述帧序列数据中各帧的像素张量输入预先训练出的神经网络模型，获取所述神经网络模型输出的所述待转换视频对应的高分辨率的视频帧序列数据；

[0110] 其中，所述神经网络模型基于抽取的所述帧序列数据中各帧的高阶像素信息获得所述高分辨率的视频帧序列数据。

[0111] 本申请实施例通过帧序列数据获取单元获取低分辨率的待转换视频的帧序列数据；通过帧序列数据转换单元将帧序列数据中各帧的像素张量输入预先训练出的神经网络模型，并由神经网络模型基于抽取的帧序列数据中各帧的高阶像素信息，获得高分辨率的视频帧序列数据并输出，从而将低分辨率的待转换视频的帧序列数据转换为高分辨率的视频帧序列数据。上述技术方案在进行视频分辨率转换时，借助视频中的高阶像素信息，使得

在神经网络进行数据处理的过程中，能够有效保留图像边缘及纹理等信息，从而实现了边缘、纹理信息的保真，避免了图片过度平滑的情况，使得画面更加自然，进而提升了超分辨率的视觉效果。

[0112] 进一步地，该装置还包括，神经网络模型训练单元，具体用于：

[0113] 在获取低分辨率的待转换视频的帧序列数据之前，获取低分辨率的样本视频的帧序列数据，以及所述样本视频的真实高分辨率帧序列数据；

[0114] 将获取的所述样本视频的帧序列数据和真实高分辨率帧序列数据，作为样本数据对初始建立的神经网络模型进行训练。

[0115] 进一步地，所述初始建立的神经网络模型包括：顺序连接的高阶像素特征学习模块、时序特征建模模块、多帧上采样模块和高阶像素损失建模模块；其中，

[0116] 所述高阶像素特征学习模块，用于基于所述样本视频的帧序列数据中各帧的像素张量计算各帧的高阶像素信息；并根据所述高阶像素信息和所述像素张量，得到所述样本视频的帧序列数据中各帧的特征张量；

[0117] 所述时序特征建模模块，用于对所述样本视频的帧序列数据中各帧的特征张量，在时间维度上进行特征变换，以使变换后各帧的时序特征参考相邻帧；

[0118] 所述多帧上采样模块，用于基于所述变换后的所述样本视频的帧序列数据中各帧的特征张量，对所述样本视频的帧序列数据中的各帧进行放大，得到所述样本视频的超分视频帧序列；

[0119] 所述高阶像素损失建模模块，用于分别获取所述超分视频帧序列中各帧的第一高阶像素信息，和所述真实高分辨率帧序列数据中各帧的第二高阶像素信息，并根据所述第一高阶像素信息与所述第二高阶像素信息之间的第一误差、以及所述超分视频帧序列中各帧的像素张量与所述真实高分辨率帧序列数据中各帧的像素张量之间的第二误差，调整神经网络模型中待学习的网络参数。

[0120] 进一步地，所述高阶像素特征学习模块包括：

[0121] 高阶像素信息抽取层，用于基于所述样本视频的帧序列数据中各帧的像素张量，计算各帧的高阶像素信息；

[0122] 高阶特征信息学习层，用于对由所述高阶像素信息构成的高阶像素张量和所述像素张量进行非线性变换，得到所述样本视频的帧序列数据中各帧的特征张量。

[0123] 进一步地，所述神经网络模型中包括多组高阶像素特征学习模块和时序特征建模模块顺序连接的组合模块，每组所述组合模块顺序连接，最后一组组合模块的时序特征建模模块与所述多帧上采样模块相连。

[0124] 进一步地，所述高阶像素信息抽取层包括至少一个卷积权重确定的二维卷积层；

[0125] 所述高阶特征信息学习层包括至少一个卷积权重需通过训练确定的三维卷积层；

[0126] 所述时序特征建模模块包括至少一个三维卷积层；

[0127] 所述多帧上采样模块包括至少一个卷积权重需通过训练确定的二维卷积层和像素重排列层；

[0128] 所述高阶像素损失建模模块包括高阶像素信息抽取层和计算模块。

[0129] 进一步地，所述预先训练出的神经网络模型包括：顺序连接的高阶像素特征学习模块、时序特征建模模块和多帧上采样模块；

[0130] 所述高阶像素特征学习模块，用于基于所述待转换视频的帧序列数据中各帧的像素张量计算各帧的高阶像素信息，并根据该高阶像素信息和该像素张量，得到所述待转换视频的帧序列数据中各帧的特征张量；

[0131] 所述时序特征建模块，用于对所述待转换视频的帧序列数据中各帧的特征张量，在时间维度上进行特征变换，以使变换后各帧的时序特征参考相邻帧；

[0132] 所述多帧上采样模块，用于基于所述变换后的所述待转换视频的帧序列数据中各帧的特征张量，对所述待转换视频的帧序列数据中的各帧进行放大，得到所述待转换视频的高分辨率的视频帧序列数据。

[0133] 进一步地，所述高阶像素信息包括各像素点与周围像素点的二阶导数信息。

[0134] 上述视频处理装置可执行本申请任意实施例所提供的视频处理方法，具备执行视频处理方法相应的功能模块和有益效果。

[0135] 实施例四

[0136] 根据本申请的实施例，本申请还提供了一种电子设备和一种可读存储介质。

[0137] 如图4所示，是实现本申请实施例的视频处理方法的电子设备的框图。电子设备旨在表示各种形式的数字计算机，诸如，膝上型计算机、台式计算机、工作台、个人数字助理、服务器、刀片式服务器、大型计算机、和其它适合的计算机。电子设备还可以表示各种形式的移动装置，诸如，个人数字处理、蜂窝电话、智能电话、可穿戴设备和其它类似的计算装置。本文所示的部件、它们的连接和关系、以及它们的功能仅仅作为示例，并且不意在限制本文中描述的和/或者要求的本申请的实现。

[0138] 如图4所示，该电子设备包括：一个或多个处理器401、存储器402，以及用于连接各部件的接口，包括高速接口和低速接口。各个部件利用不同的总线互相连接，并且可以被安装在公共主板上或者根据需要以其它方式安装。处理器可以对在电子设备内执行的指令进行处理，包括存储在存储器中或者存储器上以在外部输入/输出装置(诸如，耦合至接口的显示设备)上显示GUI的图形信息的指令。在其它实施方式中，若需要，可以将多个处理器和/或多条总线与多个存储器和多个存储器一起使用。同样，可以连接多个电子设备，各个设备提供部分必要的操作(例如，作为服务器阵列、一组刀片式服务器、或者多处理器系统)。图4中以一个处理器401为例。

[0139] 存储器402即为本申请所提供的非瞬时计算机可读存储介质。其中，所述存储器存储有可由至少一个处理器执行的指令，以使所述至少一个处理器执行本申请所提供的视频处理方法。本申请的非瞬时计算机可读存储介质存储计算机指令，该计算机指令用于使计算机执行本申请所提供的视频处理方法。

[0140] 存储器402作为一种非瞬时计算机可读存储介质，可用于存储非瞬时软件程序、非瞬时计算机可执行程序以及模块，如本申请实施例中的视频处理方法对应的程序指令/模块(例如，附图3所示的帧序列数据获取单元301和帧序列数据转换单元302)。处理器401通过运行存储在存储器402中的非瞬时软件程序、指令以及模块，从而执行服务器的各种功能应用以及数据处理，即实现上述方法实施例中的视频处理方法。

[0141] 存储器402可以包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、至少一个功能所需要的应用程序；存储数据区可存储执行视频处理方法的电子设备的使用所创建的数据等。此外，存储器402可以包括高速随机存取存储器，还可以包括非瞬时

存储器，例如至少一个磁盘存储器件、闪存器件、或其他非瞬时固态存储器件。在一些实施例中，存储器402可选包括相对于处理器401远程设置的存储器，这些远程存储器可以通过网络连接至执行视频处理方法的电子设备。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0142] 执行视频处理方法的电子设备还可以包括：输入装置403和输出装置404。处理器401、存储器402、输入装置403和输出装置404可以通过总线或者其他方式连接，图4中以通过总线连接为例。

[0143] 输入装置403可接收输入的数字或字符信息，以及产生与执行视频处理方法的电子设备的用户设置以及功能控制有关的键信号输入，例如触摸屏、小键盘、鼠标、轨迹板、触摸板、指示杆、一个或者多个鼠标按钮、轨迹球、操纵杆等输入装置。输出装置404可以包括显示设备、辅助照明装置(例如，LED)和触觉反馈装置(例如，振动电机)等。该显示设备可以包括但不限于，液晶显示器(LCD)、发光二极管(LED)显示器和等离子体显示器。在一些实施方式中，显示设备可以是触摸屏。

[0144] 此处描述的系统和技术的各种实施方式可以在数字电子电路系统、集成电路系统、专用ASIC(专用集成电路)、计算机硬件、固件、软件、和/或它们的组合中实现。这些各种实施方式可以包括：实施在一个或者多个计算机程序中，该一个或者多个计算机程序可在包括至少一个可编程处理器的可编程系统上执行和/或解释，该可编程处理器可以是专用或者通用可编程处理器，可以从存储系统、至少一个输入装置、和至少一个输出装置接收数据和指令，并且将数据和指令传输至该存储系统、该至少一个输入装置、和该至少一个输出装置。

[0145] 这些计算程序(也称作程序、软件、软件应用、或者代码)包括可编程处理器的机器指令，并且可以利用高级过程和/或面向对象的编程语言、和/或汇编/机器语言来实施这些计算程序。如本文使用的，术语“机器可读介质”和“计算机可读介质”指的是用于将机器指令和/或数据提供给可编程处理器的任何计算机程序产品、设备、和/或装置(例如，磁盘、光盘、存储器、可编程逻辑装置(PLD))，包括，接收作为机器可读信号的机器指令的机器可读介质。术语“机器可读信号”指的是用于将机器指令和/或数据提供给可编程处理器的任何信号。

[0146] 为了提供与用户的交互，可以在计算机上实施此处描述的系统和技术，该计算机具有：用于向用户显示信息的显示装置(例如，CRT(阴极射线管)或者LCD(液晶显示器)监视器)；以及键盘和指向装置(例如，鼠标或者轨迹球)，用户可以通过该键盘和该指向装置来将输入提供给计算机。其它种类的装置还可以用于提供与用户的交互；例如，提供给用户的反馈可以是任何形式的传感反馈(例如，视觉反馈、听觉反馈、或者触觉反馈)；并且可以用任何形式(包括声输入、语音输入或者、触觉输入)来接收来自用户的输入。

[0147] 可以将此处描述的系统和技术实施在包括后台部件的计算系统(例如，作为数据服务器)、或者包括中间件部件的计算系统(例如，应用服务器)、或者包括前端部件的计算系统(例如，具有图形用户界面或者网络浏览器的用户计算机，用户可以通过该图形用户界面或者该网络浏览器来与此处描述的系统和技术的实施方式交互)、或者包括这种后台部件、中间件部件、或者前端部件的任何组合的计算系统中。可以通过任何形式或者介质的数字数据通信(例如，通信网络)来将系统的部件相互连接。通信网络的示例包括：局域网

(LAN)、广域网(WAN)和互联网。

[0148] 计算机系统可以包括客户端和服务器。客户端和服务器一般远离彼此并且通常通过通信网络进行交互。通过在相应的计算机上运行并且彼此具有客户端-服务器关系的计算机程序来产生客户端和服务器的关系。

[0149] 根据本申请实施例的技术方案,通过获取低分辨率的待转换视频的帧序列数据;将帧序列数据中各帧的像素张量输入预先训练出的神经网络模型,并由神经网络模型基于抽取的帧序列数据中各帧的高阶像素信息,获得高分辨率的视频帧序列数据并输出,从而将低分辨率的待转换视频的帧序列数据转换为高分辨率的视频帧序列数据。上述技术方案在进行视频分辨率转换时,借助视频中的高阶像素信息,使得在神经网络进行数据处理的过程中,能够有效保留图像边缘及纹理等信息,从而实现了边缘、纹理信息的保真,避免了图片过度平滑的情况,使得画面更加自然,进而提升了超分辨率的视觉效果。

[0150] 应该理解,可以使用上面所示的各种形式的流程,重新排序、增加或删除步骤。例如,本发申请中记载的各步骤可以并行地执行也可以顺序地执行也可以不同的次序执行,只要能够实现本申请公开的技术方案所期望的结果,本文在此不进行限制。

[0151] 上述具体实施方式,并不构成对本申请保护范围的限制。本领域技术人员应该明白的是,根据设计要求和其他因素,可以进行各种修改、组合、子组合和替代。任何在本申请的精神和原则之内所作的修改、等同替换和改进等,均应包含在本申请保护范围之内。

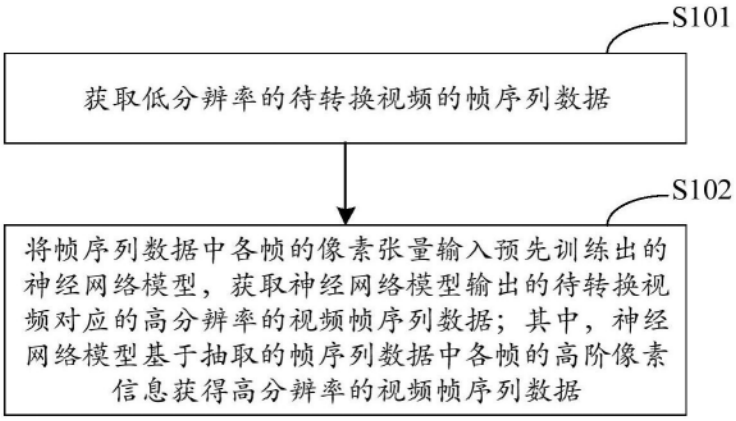

S101
获取低分辨率的待转换视频的帧序列数据
S102
将帧序列数据中各帧的像素张量输入预先训练出的神经网络模型，获取神经网络模型输出的待转换视频对应的高分辨率的视频帧序列数据；其中，神经网络模型基于抽取的帧序列数据中各帧的高阶像素信息获得高分辨率的视频帧序列数据

图1

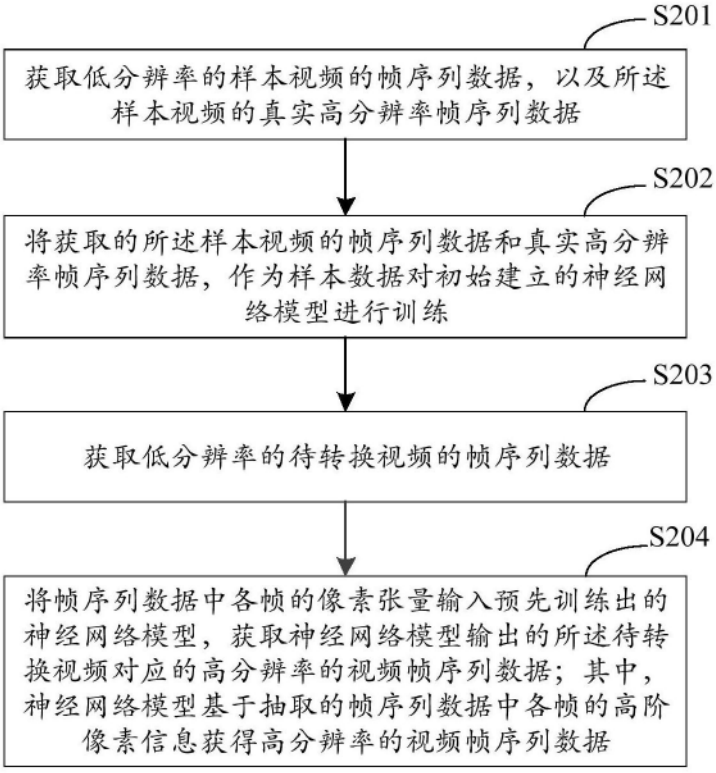

S201
获取低分辨率的样本视频的帧序列数据，以及所述样本视频的真实高分辨率帧序列数据
S202
将获取的所述样本视频的帧序列数据和真实高分辨率帧序列数据，作为样本数据对初始建立的神经网络模型进行训练
S203
获取低分辨率的待转换视频的帧序列数据
S204
将帧序列数据中各帧的像素张量输入预先训练出的神经网络模型，获取神经网络模型输出的所述待转换视频对应的高分辨率的视频帧序列数据；其中，神经网络模型基于抽取的帧序列数据中各帧的高阶像素信息获得高分辨率的视频帧序列数据

图2A

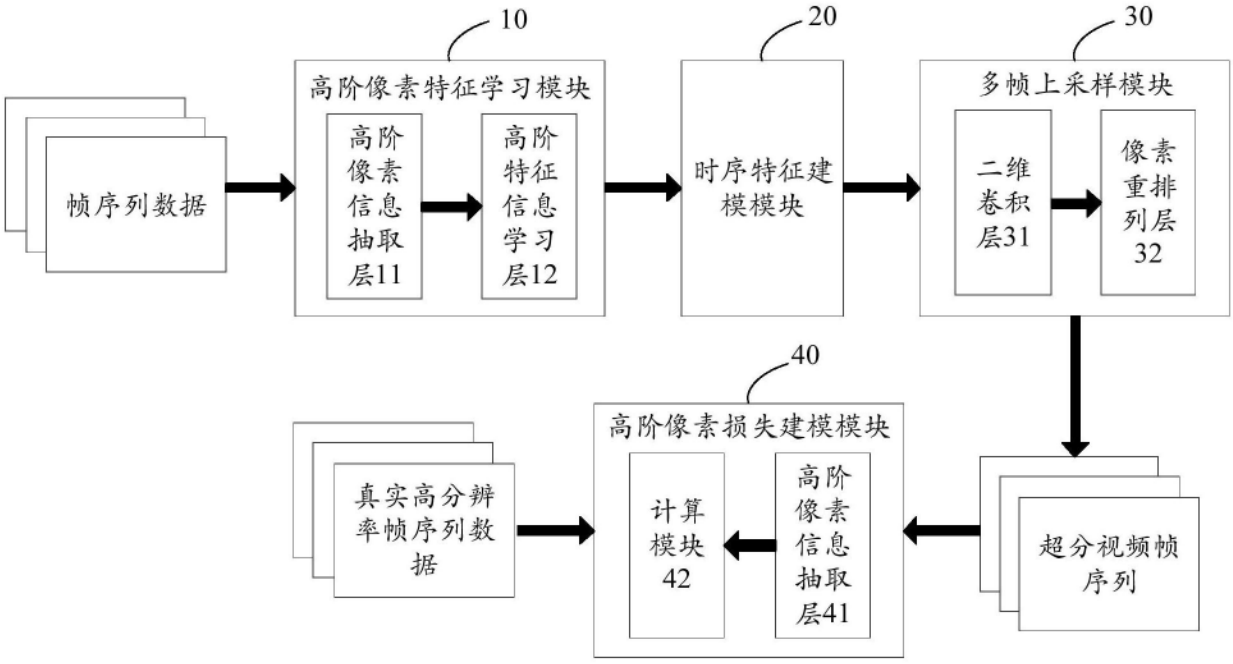

10
高阶像素特征学习模块
高阶像素信息抽取层11
高阶特征信息学习层12
20
时序特征建模模块
30
多帧上采样模块
二维卷积层31
像素重排列层32
帧序列数据
40
高阶像素损失建模模块
计算模块42
高阶像素信息抽取层41
真实高分辨率帧序列数据
超分视频帧序列

图2B

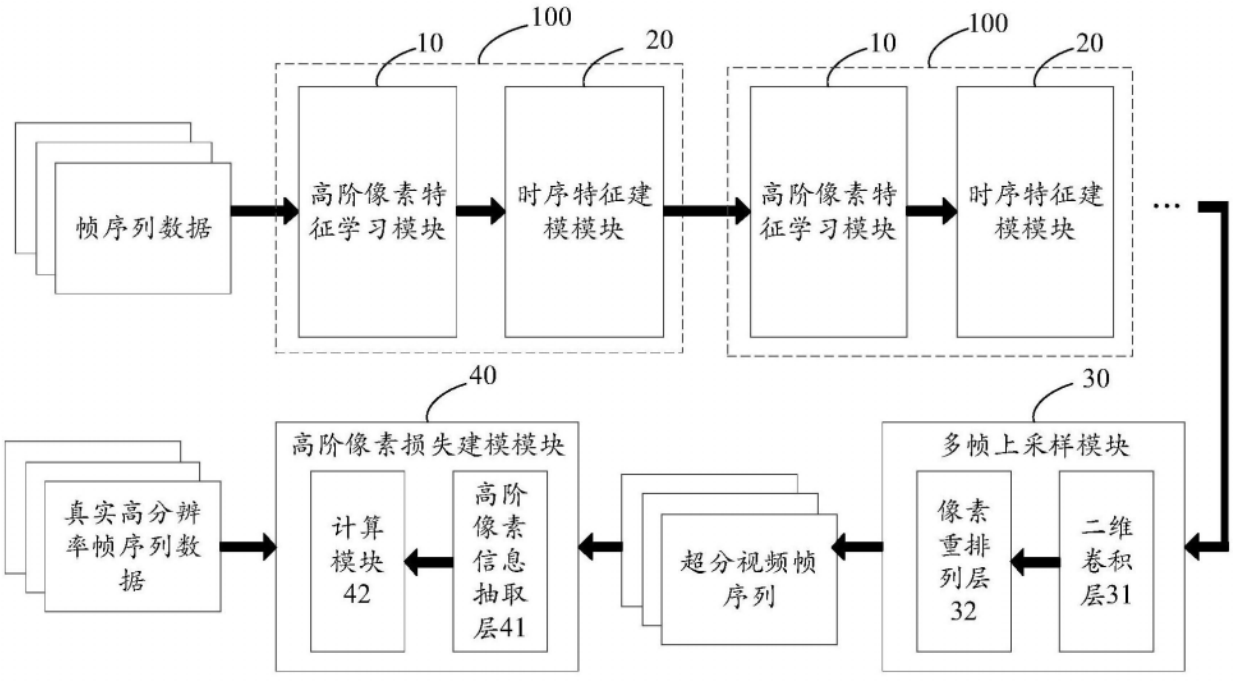

100
10
20
帧序列数据
高阶像素特征学习模块
时序特征建模模块
...
30
多帧上采样模块
像素重排列层32
二维卷积层31
40
高阶像素损失建模模块
计算模块42
高阶像素信息抽取层41
真实高分辨率帧序列数据
超分视频帧序列

图2C

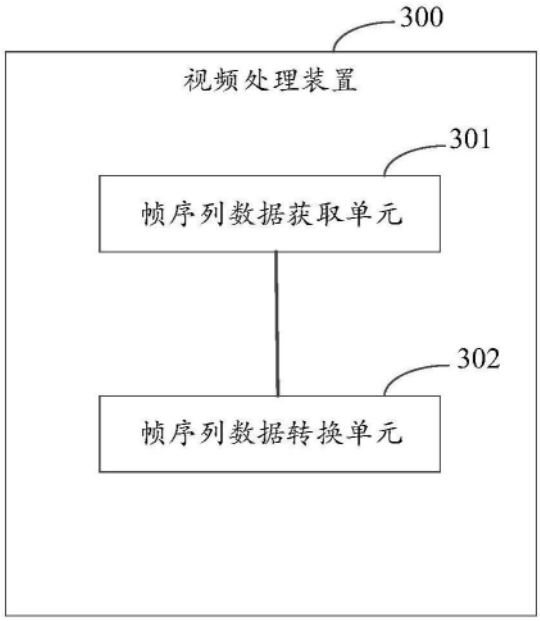

300
视频处理装置
301
帧序列数据获取单元
302
帧序列数据转换单元

图3

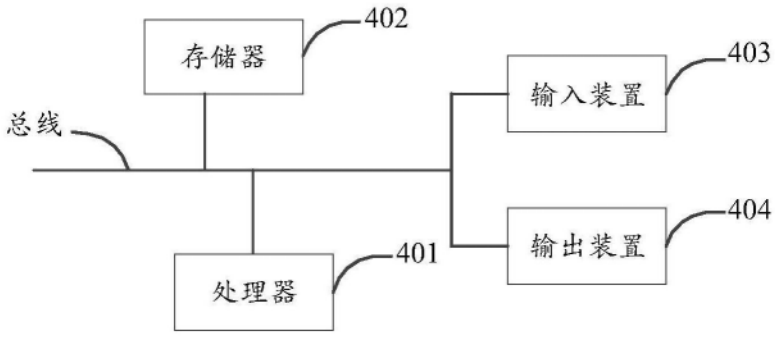

402
存储器
403
输入装置
总线
404
输出装置
401
处理器

图4