



(21) 申请号 202311782836.7

(22) 申请日 2023.12.22

(71) 申请人 浙江省北大信息技术高等研究院

地址 311200 浙江省杭州市萧山区宁围镇

市心北路857号288-1室

申请人 北京大学

(72) 发明人 马思伟 蒋云 滕波 常建慧

高文

(74) 专利代理机构 北京辰权知识产权代理有限公司

公司 11619

专利代理师 付婧

(51) Int. Cl.

H04N 19/124 (2014.01)

H04N 19/13 (2014.01)

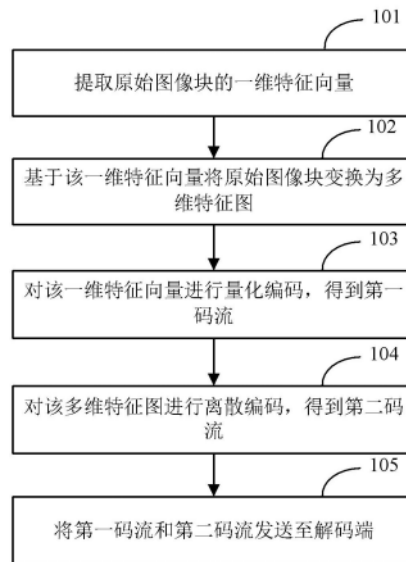
权利要求书3页 说明书12页 附图5页

(54) 发明名称

一种图像处理方法、装置、系统、电子设备及存储介质

(57) 摘要

本申请公开了一种图像处理方法、装置、系统、电子设备及存储介质,应用于编码端,方法包括:提取原始图像块的一维特征向量;基于所述一维特征向量将所述原始图像块变换为多维特征图;对所述一维特征向量进行量化编码,得到第一码流;对所述多维特征图进行离散编码,得到第二码流,从而实现空域无关向量和多维特征图的高效压缩;将所述第一码流和所述第二码流发送至解码端。由于编码码流由两层码流构成,包含图像的不同层信息,这样即使在低码率下也不会损失很多信息,因此由两层码流重建获得的图像,可以改善视觉效果,提升视觉体验。



1. 一种图像处理方法,其特征在于,应用于编码端,所述方法包括:
提取原始图像块的一维特征向量;
基于所述一维特征向量将所述原始图像块变换为多维特征图;
对所述一维特征向量进行量化编码,得到第一码流;
对所述多维特征图进行离散编码,得到第二码流;
将所述第一码流和所述第二码流发送至解码端。
2. 根据权利要求1所述的方法,其特征在于,所述基于所述一维特征向量将所述原始图像块变换为多维特征图,包括:
通过预设的扩散模型利用所述一维特征向量对所述原始图像块进行多次迭代加噪,得到多维特征图。
3. 根据权利要求2所述的方法,其特征在于,所述扩散模型包括第一噪声预测网络和加噪层;所述通过预设的扩散模型利用所述一维特征向量对所述原始图像块进行多次迭代加噪,包括:
针对每次加噪过程,通过所述第一噪声预测网络根据上一次加噪结果和所述一维特征向量预测当前次的噪声;
通过所述加噪层将所述第一噪声预测网络输出的噪声加入到上一次加噪结果;
其中,对于首次加噪过程,上一次加噪结果为所述原始图像块。
4. 根据权利要求3所述的方法,其特征在于,通过所述加噪层将所述第一噪声预测网络输出的噪声加入到上一次加噪结果,包括:
获取上一次使用的噪声水平系数和当前次的噪声水平系数;
根据上一次使用的噪声水平系数、当前次的噪声水平系数、所述第一噪声预测网络输出的噪声、所述上一次加噪结果获得当前次的加噪结果。
5. 根据权利要求1所述的方法,其特征在于,所述对所述一维特征向量进行量化编码,得到第一码流,包括:
通过预设的量化编码模型对所述一维特征向量进行量化编码,得到第一码流。
6. 根据权利要求5所述的方法,其特征在于,所述量化编码模型包括特征编码网络、量化层、熵估计网络;
所述通过预设的量化编码模型对所述一维特征向量进行量化编码,得到第一码流,包括:
通过所述特征编码网络对所述一维特征向量进行下采样处理,得到下采样特征;
通过所述量化层将所述下采样特征的特征值转换为整型数值,得到量化特征;
通过所述熵估计网络将所述量化特征编码为第一码流。
7. 根据权利要求1所述的方法,其特征在于,所述对所述多维特征图进行离散编码,得到第二码流,包括:
通过预设的离散编码模型对所述多维特征图进行离散编码,得到第二码流。
8. 根据权利要求7所述的方法,其特征在于,所述离散编码模型包括特征编码网络、离散表示层、码流转换层;
所述通过预设的离散编码模型对所述多维特征图进行离散编码,得到第二码流,包括:
通过所述特征编码网络对所述多维特征图进行下采样处理,得到下采样特征;

通过所述离散表示层对所述下采样特征进行离散表示,得到索引矩阵;
通过所述码流转换层将所述索引矩阵转换为第二码流。

9. 根据权利要求8所述的方法,其特征在于,所述通过所述离散表示层对所述下采样特征进行离散表示,得到索引矩阵,包括:

在预设特征字典中,为所述下采样特征的每维特征向量匹配最相似的向量;
利用与每维特征向量匹配的向量的索引获得索引矩阵。

10. 一种图像处理方法,其特征在于,应用于解码端,所述方法包括:

接收来自编码端的第一码流和第二码流;
对所述第一码流进行反量化解码,得到一维特征向量;
对所述第二码流进行反离散解码,得到多维特征图;
基于所述一维特征向量对所述多维特征图进行图像重建,得到重建图像块。

11. 根据权利要求10所述的方法,其特征在于,所述对所述第一码流进行反量化解码,得到一维特征向量,包括:

通过预设的反量化解码模型对所述第一码流进行反量化解码,得到一维特征向量;
所述反量化解码模型包括熵解码网络、反量化层和特征解码网络;所述通过预设的反量化解码模型对所述第一码流进行反量化解码,得到一维特征向量,包括:

通过所述熵解码网络将所述第一码流解码为量化特征;
通过所述反量化层将所述量化特征转换为浮点数值,得到反量化特征;
通过所述特征解码网络对所述反量化特征进行上采样处理,得到一维特征向量。

12. 根据权利要求10所述的方法,其特征在于,所述对所述第二码流进行反离散解码,得到多维特征图,包括:

通过预设的离散解码模型对所述第二码流进行反离散解码,得到多维特征图;
所述离散解码模型包括码流解码层、反离散表示层和特征解码网络;
所述通过预设的离散解码模型对所述第二码流进行反离散解码,得到多维特征图,包括:

通过所述码流解码层将所述第二码流解码为索引矩阵;
通过所述反离散表示层基于预设特征字典将所述索引矩阵转换为连续特征;
通过所述特征解码网络对所述连续特征进行上采样处理,得到多维特征图。

13. 根据权利要求10所述的方法,其特征在于,所述基于所述一维特征向量对所述多维特征图进行图像重建,得到重建图像块,包括:

通过预设的扩散模型利用所述一维特征向量对所述多维特征图进行多次迭代去噪,得到重建图像块。

14. 根据权利要求13所述的方法,其特征在于,所述扩散模型包括第二噪声预测网络和去噪层;所述通过预设的扩散模型利用所述一维特征向量对所述多维特征图进行多次迭代去噪,得到重建图像块,包括:

针对每次去噪过程,通过所述第二噪声预测网络根据上一次去噪结果和所述一维特征向量预测当前次的噪声;

通过所述去噪层利用所述第二噪声预测网络输出的噪声对上一次去噪结果进行去噪;
其中,对于首次去噪过程,上一次去噪结果为所述多维特征图。

15. 根据权利要求14所述的方法, 其特征在于, 通过所述去噪层利用所述第二噪声预测网络输出的噪声对上一次去噪结果进行去噪, 包括:

获取上一次使用的噪声水平系数和当前次的噪声水平系数;

根据上一次使用的噪声水平系数、当前次的噪声水平系数、所述第二噪声预测网络输出的噪声、所述上一次去噪结果获得当前次的去噪结果。

16. 一种图像处理装置, 其特征在于, 应用于编码端, 所述装置包括:

第一提取模块, 用于提取原始图像块的一维特征向量;

第二提取模块, 用于基于所述一维特征向量将所述原始图像块变换为多维特征图;

第一编码模块, 用于对所述一维特征向量进行量化编码, 得到第一码流;

第二编码模块, 用于对所述多维特征图进行离散编码, 得到第二码流;

发送模块, 用于将所述第一码流和所述第二码流发送至解码端。

17. 一种图像处理装置, 其特征在于, 应用于解码端, 所述装置包括:

接收模块, 接收来自编码端的第一码流和第二码流;

第一解码模块, 用于对所述第一码流进行反量化解码, 得到一维特征向量;

第二解码模块, 用于对所述第二码流进行反离散解码, 得到多维特征图;

重建模块, 用于基于所述一维特征向量对所述多维特征图进行图像重建, 得到重建图像块。

18. 一种图像处理系统, 其特征在于, 所述系统包括:

编码端, 用于执行权利要求1-9任一项所述的图像处理方法;

解码端, 用于执行权利要求10-15任一项所述的图像处理方法。

19. 一种电子设备, 包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序, 其特征在于, 所述处理器执行所述程序以实现如权利要求1-15任一项所述的方法。

20. 一种计算机可读存储介质, 其上存储有计算机程序, 其特征在于, 所述程序被处理器执行以实现如权利要求1-15任一项所述的方法。

一种图像处理方法、装置、系统、电子设备及存储介质

技术领域

[0001] 本申请涉及图像处理技术领域,具体涉及一种图像处理方法、装置、电子设备及存储介质。

背景技术

[0002] 随着人工智能的发展,传统图像编码标准(如JPEG(Joint Photographic Experts Group,联合图像专家组格式)、MPEG(Moving Picture Expert Group,运动图像专家组格式)、AVS(Audio Video coding Standard,信源编码标准)、BPG(Better Portable Graphics,更好的可移植图形)等)是面向人类视觉的像素级优化,无法很好支持各式各样的机器视觉要求,人们尝试将神经网络加入到图像编码中,实现更加智能、高效的图像编码,因此基于深度学习的神经网络编码成为研究重点。但无论是传统图像编码标准还是神经网络编码技术的编码码流都是由单一码流构成,在低码率下信息损失大,图像重建效果差,视觉体验效果差。

发明内容

[0003] 本申请的目的是针对上述现有技术的不足提出的一种图像处理方法、装置、电子设备及存储介质,该目的是通过以下技术方案实现的。

[0004] 本申请的第一方面提出了一种图像处理方法,应用于编码端,所述方法包括:

提取原始图像块的一维特征向量;

基于所述一维特征向量将所述原始图像块变换为多维特征图;

对所述一维特征向量进行量化编码,得到第一码流;

对所述多维特征图进行离散编码,得到第二码流;

将所述第一码流和所述第二码流发送至解码端。

[0005] 基于上述第一方面所述的图像处理方法,本申请至少具有如下有益效果或优点:

通过提取原始图像块的一维特征向量,也即一个低维的空域无关向量,该向量包含了图像的纹理结构信息,然后基于这个一维特征向量再将原始图像块变换为多维特征图,该多维特征图不仅包含图像的空域相关信息,还会包含图像的高频信息,考虑到空域无关向量和多维特征图分别包含了图像的不同层面信息,通过对一维特征向量量化编码为第一码流,而对多维特征图离散编码为第二码流,以实现空域无关向量和多维特征图的高效压缩。由于编码码流由两层码流构成,包含图像的不同层信息,这样即使在低码率下也不会损失很多信息,因此由两层码流重建获得的图像,可以改善视觉效果,提升视觉体验。

[0006] 可选的,所述基于所述一维特征向量将所述原始图像块变换为多维特征图,包括:通过预设的扩散模型利用所述一维特征向量对所述原始图像块进行多次迭代加噪,得到多维特征图。

[0007] 可选的,所述扩散模型包括第一噪声预测网络和加噪层;所述通过预设的扩散模型利用所述一维特征向量对所述原始图像块进行多次迭代加噪,包括:针对每次加噪过程,

通过所述第一噪声预测网络根据上一次加噪结果和所述一维特征向量预测当前次的噪声；通过所述加噪层将所述第一噪声预测网络输出的噪声加入到上一次加噪结果；其中，对于首次加噪过程，上一次加噪结果为所述原始图像块。

[0008] 可选的，通过所述加噪层将所述第一噪声预测网络输出的噪声加入到上一次加噪结果，包括：获取上一次使用的噪声水平系数和当前次的噪声水平系数；根据上一次使用的噪声水平系数、当前次的噪声水平系数、所述第一噪声预测网络输出的噪声、所述上一次加噪结果获得当前次的加噪结果。

[0009] 可选的，所述对所述一维特征向量进行量化编码，得到第一码流，包括：通过预设的量化编码模型对所述一维特征向量进行量化编码，得到第一码流。

[0010] 可选的，所述量化编码模型包括特征编码网络、量化层、熵估计网络；所述通过预设的量化编码模型对所述一维特征向量进行量化编码，得到第一码流，包括：通过所述特征编码网络对所述一维特征向量进行下采样处理，得到下采样特征；通过所述量化层将所述下采样特征的特征值转换为整型数值，得到量化特征；通过所述熵估计网络将所述量化特征编码为第一码流。

[0011] 可选的，所述对所述多维特征图进行离散编码，得到第二码流，包括：通过预设的离散编码模型对所述多维特征图进行离散编码，得到第二码流。

[0012] 可选的，所述离散编码模型包括特征编码网络、离散表示层、码流转换层；所述通过预设的离散编码模型对所述多维特征图进行离散编码，得到第二码流，包括：通过所述特征编码网络对所述多维特征图进行下采样处理，得到下采样特征；通过所述离散表示层对所述下采样特征进行离散表示，得到索引矩阵；通过所述码流转换层将所述索引矩阵转换为第二码流。

[0013] 可选的，所述通过所述离散表示层对所述下采样特征进行离散表示，得到索引矩阵，包括：在预设特征字典中，为所述下采样特征的每维特征向量匹配最相似的向量；利用与每维特征向量匹配的向量的索引获得索引矩阵。

[0014] 本申请的第二方面提出了一种图像处理方法，应用于解码端，所述方法包括：

接收来自编码端的第一码流和第二码流；

对所述第一码流进行反量化解码，得到一维特征向量；

对所述第二码流进行反离散解码，得到多维特征图；

基于所述一维特征向量对所述多维特征图进行图像重建，得到重建图像块。

[0015] 基于上述第二方面所述的图像处理方法，本申请至少具有如下有益效果或优点：

由于接收的第一码流和第二码流包含了图像的不同层面信息，第一码流包含有图像的纹理风格信息，第二码流包含图像的高频信息和结构信息，因此由第一码流和第二码流解码获得的重建图像，可以提升图像视觉效果，并且即使在带宽限制条件下也不会损失很多信息。

[0016] 可选的，所述对所述第一码流进行反量化解码，得到一维特征向量，包括：通过预设的反量化解码模型对所述第一码流进行反量化解码，得到一维特征向量；

所述反量化解码模型包括熵解码网络、反量化层和特征解码网络；所述通过预设的反量化解码模型对所述第一码流进行反量化解码，得到一维特征向量，包括：通过所述熵解码网络将所述第一码流解码为量化特征；通过所述反量化层将所述量化特征转换为浮点

数值,得到反量化特征;通过所述特征解码网络对所述反量化特征进行上采样处理,得到一维特征向量。

[0017] 可选的,所述对所述第二码流进行反离散解码,得到多维特征图,包括:通过预设的离散解码模型对所述第二码流进行反离散解码,得到多维特征图;

所述离散解码模型包括码流解码层、反离散表示层和特征解码网络;所述通过预设的离散解码模型对所述第二码流进行反离散解码,得到多维特征图,包括:通过所述码流解码层将所述第二码流解码为索引矩阵;通过所述反离散表示层基于预设特征字典将所述索引矩阵转换为连续特征;通过所述特征解码网络对所述连续特征进行上采样处理,得到多维特征图。

[0018] 可选的,所述基于所述一维特征向量对所述多维特征图进行图像重建,得到重建图像块,包括:通过预设的扩散模型利用所述一维特征向量对所述多维特征图进行多次迭代去噪,得到重建图像块。

[0019] 可选的,所述扩散模型包括第二噪声预测网络和去噪层;所述通过预设的扩散模型利用所述一维特征向量对所述多维特征图进行多次迭代去噪,得到重建图像块,包括:针对每次去噪过程,通过所述第二噪声预测网络根据上一次去噪结果和所述一维特征向量预测当前次的噪声;通过所述去噪层利用所述第二噪声预测网络输出的噪声对上一次去噪结果进行去噪;其中,对于首次去噪过程,上一次去噪结果为所述多维特征图。

[0020] 可选的,通过所述去噪层利用所述第二噪声预测网络输出的噪声对上一次去噪结果进行去噪,包括:获取上一次使用的噪声水平系数和当前次的噪声水平系数;根据上一次使用的噪声水平系数、当前次的噪声水平系数、所述第二噪声预测网络输出的噪声、所述上一次去噪结果获得当前次的去噪结果。

[0021] 本申请的第三方面提出了一种图像处理装置,应用于编码端,所述装置包括:
第一提取模块,用于提取原始图像块的一维特征向量;
第二提取模块,用于基于所述一维特征向量将所述原始图像块变换为多维特征图;

第一编码模块,用于对所述一维特征向量进行量化编码,得到第一码流;

第二编码模块,用于对所述多维特征图进行离散编码,得到第二码流;

发送模块,用于将所述第一码流和所述第二码流发送至解码端。

[0022] 本申请的第四方面提出了一种图像处理装置,应用于解码端,所述装置包括:
接收模块,接收来自编码端的第一码流和第二码流;
第一解码模块,用于对所述第一码流进行反量化解码,得到一维特征向量;
第二解码模块,用于对所述第二码流进行反离散解码,得到多维特征图;
重建模块,用于基于所述一维特征向量对所述多维特征图进行图像重建,得到重建图像块。

[0023] 本申请的第五方面提出了一种图像处理系统,所述系统包括:

编码端,用于执行第一方面所述的图像处理方法;

解码端,用于执行第二方面所述的图像处理方法。

[0024] 本申请的第六方面提出了一种电子设备,包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,所述处理器执行所述程序以实现如上述第一方面或第

二方面所述的方法。

[0025] 本申请的第七方面提出了一种计算机可读存储介质,其上存储有计算机程序,所述程序被处理器执行以实现如上述第一方面或第二方面所述的方法。

[0026] 上述说明仅是本申请技术方案的概述,为了能够更清楚了解本申请的技术手段,可依照说明书的内容予以实施,并且为了让本申请的上述和其他目的、特征和优点能够更明显易懂,以下特举本申请的具体实施方式。

附图说明

[0027] 此处所说明的附图用来提供对本申请的进一步理解,构成本申请的一部分,本申请的示意性实施例及其说明用于解释本申请,并不构成对本申请的不当限定。在附图中:

图1为根据一示例性实施例示出的一种图像处理方法的实施例流程图;

图2为根据一示例性实施例示出的一种扩散模型的结构示意图;

图3为根据一示例性实施例示出的一种量化编码模型的结构示意图;

图4为根据一示例性实施例示出的一种离散编码模型的结构示意图;

图5为根据一示例性实施例示出的另一种图像处理方法的实施例流程图;

图6为根据一示例性实施例示出的一种量化解码模型的结构示意图;

图7为根据一示例性实施例示出的一种离散解码模型的结构示意图;

图8-图11为根据一示例性实施例示出的实现图像处理的示意图。

具体实施方式

[0028] 这里将详细地对示例性实施例进行说明,其示例表示在附图中。下面的描述涉及附图时,除非另有表示,不同附图中的相同数字表示相同或相似的要素。以下示例性实施例中所描述的实施方式并不代表与本申请相一致的所有实施方式。相反,它们仅是与如所附权利要求书中所详述的、本申请的一些方面相一致的装置和方法的例子。

[0029] 在本申请使用的术语是仅仅出于描述特定实施例的目的,而非旨在限制本申请。在本申请和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式,除非上下文清楚地表示其他含义。还应当理解,本文中使用的术语“和/或”是指并包含一个或多个相关联的列出项目的任何或所有可能组合。

[0030] 应当理解,尽管在本申请可能采用术语第一、第二、第三等来描述各种信息,但这些信息不应限于这些术语。这些术语仅用来将同一类型的信息彼此区分开。例如,在不脱离本申请范围的情况下,第一信息也可以被称为第二信息,类似地,第二信息也可以被称为第一信息。取决于语境,如在此所使用的词语“如果”可以被解释成为“在……时”或“当……时”或“响应于确定”。

[0031] 实施例一:

图1为根据一示例性实施例示出的一种图像处理方法的实施例流程图,该图像处理方法应用于编码端,编码端用于将原始图像压缩编码为码流。如图1所示,该图像处理方法包括如下步骤:

步骤101:提取原始图像块的一维特征向量。

[0032] 在该步骤中,原始图像块可以是按照预设块尺寸对待编码图像进行划分之后获得

的其中一块图像,其尺寸大小为预设块尺寸,例如32*32。该一维特征向量的视觉内容表示包含了图像纹理结构,由于特征的维度比较低,因此其表征的图像空域相关性比较弱,空域上是无关的。

[0033] 在一可选实施例中,可以通过使用一个预设的卷积编码器对原始图像块进行特征提取,得到一维特征向量。具体地,特征提取公式如下:

$$z_t = \text{Enc}(x_0)$$

[0034] 其中,Enc()表示卷积编码器, x_0 表示输入的原始图像块。

[0035] 步骤102:基于该一维特征向量将原始图像块变换为多维特征图。

[0036] 在该步骤中,多维特征图包含了原始图像块中的图像高频信息和图像结构信息的视觉内容表示。值得注意的是,该多维特征图的尺寸与原始图像块的尺寸一致。

[0037] 在一可选实施例中,可以通过预设的扩散模型利用该一维特征向量对原始图像块进行多次迭代加噪,得到多维特征图。

[0038] 其中,扩散模型包括前向扩散过程和后向过程,在前向扩散过程对原始图像块进行迭代加噪,得到的多维特征图即为一张噪声图,由于扩散模型具有保留数据语义结构的能力,并且使用一维特征向量调制噪声的特征分布,而不是每次加入随机采样高斯噪声,因此对原始图像进行多次迭代加噪后得到的近似正态分布的噪声图,包含有图像高频信息和图像结构信息,同时还保持着图像的空域分辨率,特征之间具有空间相关性。

[0039] 在一具体实施方式中,如图2所示的扩散模型,在前向扩散过程包括第一噪声预测网络和加噪层,在对原始图像块进行多次迭代加噪过程中,对于每次加噪过程,都通过第一噪声预测网络根据上一次加噪结果和一维特征向量预测当前次的噪声,然后通过加噪层将第一噪声预测网络输出的噪声加入到上一次加噪结果上。

[0040] 其中,对于首次加噪过程,上一次加噪结果即为原始图像块。通过在前向扩散过程使用噪声预测网络估计的确定噪声代替随机采样高斯噪声,将传统的随机扩散过程转化为确定性的编码过程,由于噪声预测引入一维特征向量和上一次加噪结果的条件用于调制噪声预测网络的特征分布,因此预测噪声能够保留空域相关信息和丰富的结构信息,从而经过多次迭代加噪后得到的噪声图具有空域相关性,同时还包含高频信息和图像结构信息。

[0041] 具体地,针对加噪层的具体加噪实现过程包括:获取上一次使用的噪声水平系数和当前次的噪声水平系数,然后根据上一次使用的噪声水平系数、当前次的噪声水平系数、第一噪声预测网络输出的噪声、上一次加噪结果获得当前次的加噪结果。

[0042] 加噪公式可选为:

$$x_t = \frac{\sqrt{\alpha_t}(x_{t-1} - \sqrt{1 - \alpha_{t-1}}\varepsilon_\theta(x_{t-1}, t-1, z_t))}{\sqrt{\alpha_{t-1}}} + \sqrt{1 - \alpha_t}\varepsilon_\theta(x_{t-1}, t-1, z_t)$$

[0043] 其中, x_t 表示当前次的加噪结果, x_{t-1} 表示上一次加噪结果, $\varepsilon_\theta(x_{t-1}, t-1, z_t)$ 表示第一噪声预测网络输出的噪声, z_t 表示所述一维特征向量, α_t 表示当前次的噪声水平系数, α_{t-1} 表示上一次的噪声水平系数。值得注意的是,噪声水平系数表示噪声程度,取值范围0-1,并且使用的噪声水平系数逐步增加,即当前次的噪声水平系数大于上一次的噪声水平系数。

[0044] 上述加噪公式仅为一种示例性说明,对于其他可实现加噪结果的公式也在本申请

保护范围内,因此本申请不限定加载公式的具体形式。

[0045] 需要说明的是,由图2可以看出,扩散模型的反向过程(也即去噪过程)也会使用噪声预测网络进行噪声预测,以用于去噪,该反向过程在描述图像重建时详细说明,为了更好的实现图像重建,保持加噪、去噪的一致性,第一噪声预测网络可以采用反向过程中使用的噪声预测网络,也就是说,第一噪声预测网络与第二噪声预测网络相同。当然,第一噪声预测网络也可以采用与第二噪声预测网络不同的网络结构。

[0046] 基于上述描述,通过使用扩散模型驱动图像视觉内容的高效表示,由于扩散模型在生成视觉内容表示过程中,不需要考虑图像的场景内容,因此可以实现对任意图像场景内容的高效表示,相比现有的生成式图像编码技术对复杂数据分布的学习能力有限,模型参数需要面向图像特定场景优化,无法应用于复杂多变的内容场景,本申请实施例使用扩散模型进行图像视觉内容高效表示的适应性更强。

[0047] 此外,由于使用扩散模型驱动图像视觉内容表示,无需考虑图像场景内容,可以将图像划分成多个图像块分别进行编码,因此即使是高分辨率内容的图像,也可以实现高效编码,相比现有的生成式图像编码技术对于高分辨率内容的训练受限于资源、算法设计等问题,无法应用至高分辨率图像内容编码,本申请实施例使用扩散模型实现编码,可以实现对任意分辨率内容的高效编码。

[0048] 步骤103:对该一维特征向量进行量化编码,得到第一码流。

[0049] 在该步骤中,由于一维特征向量的空域相关性很弱,而局部相关性比较强,因此可以使用基于局部相关性的编码方法,将特征向量在特征维度上进行量化编码,例如基于端到端的图像编码方法、基于块的视频帧内编码方法等。

[0050] 在一可选实施例中,可以通过预设的量化编码模型对所述一维特征向量进行量化编码,得到第一码流,从而实现端到端的图像编码。

[0051] 在一种实施方式中,如图3所示的量化编码模型,包括特征编码网络、量化层以及熵估计网络。编码过程包括:通过特征编码网络对一维特征向量进行下采样处理,得到下采样特征,然后通过量化层将下采样特征的特征值转换为整型数值,从而得到量化特征,最后通过熵估计网络量化特征编码为第一码流。

[0052] 其中,量化特征中整型特征值的取值多样化,因此需要使用熵估计网络的概率表将量化特征中整型特征值编码为二进制的第二码流。

[0053] 如前所述,一维特征向量包含有图像纹理信息,因此将一维特征向量量化编码获得的第一码流会携带有图像纹理信息,便于解码端重建图像纹理。

[0054] 步骤104:对该多维特征图进行离散编码,得到第二码流。

[0055] 在该步骤中,由于多维特征图的局部相关性比较弱,而空域相关性很强,因此不适用基于局部相关性的编码方法,而适用于基于离散表示的编码方法。

[0056] 在一可选实施例中,可以通过预设的离散编码模型对多维特征图进行离散编码,得到第二码流,从而实现离散表示的图像编码。

[0057] 在一种实施方式中,如图4所示的离散编码模型,包括特征编码网络、离散表示层以及码流转换层。编码过程包括:通过特征编码网络对多维特征图进行下采样处理,得到下采样特征,然后通过离散表示层对下采样特征进行离散表示,从而得到索引矩阵,最后通过码流转换层将索引矩阵转换为第二码流。

[0058] 其中,索引矩阵中索引值来源于离散表示层中使用的特征字典包含的向量的索引,因此索引值的取值范围固定,可以在码流转换层设计一个转换函数将索引矩阵中索引值转换为二进制的第二码流。

[0059] 如前所述,多维特征图包含有图像高频信息和结构信息,因此将多维特征图离散编码获得的第二码流会携带有图像高频信息和结构信息,便于解码端重建图像。

[0060] 具体地,对于离散表示层,在训练过程中,通过对多维特征图进行聚类,学习一个特征字典,该特征字典定义了一个离散隐变量空间,又称为嵌入空间,其大小为 $K \times D$, K 表示嵌入向量的个数, D 表示每个嵌入向量的长度。

[0061] 基于此,离散表示过程包括:在预设特征字典中,为下采样特征的每维特征向量匹配最相似的向量,从而利用与每维特征向量匹配的向量的索引获得索引矩阵。其中,特征向量在特征字典中相似向量的匹配方式可以选择欧式距离最近的向量。

[0062] 步骤105:将第一码流和第二码流发送至解码端。

[0063] 至此,完成上述图1所示的图像处理流程,通过提取原始图像块的一维特征向量,也即一个低维的空域无关向量,该向量包含了图像的纹理结构信息,然后基于这个一维特征向量再将原始图像块变换为多维特征图,该多维特征图不仅包含图像的空域相关信息,还会包含图像的高频信息,考虑到空域无关向量和多维特征图分别包含了图像的不同层面信息,通过对一维特征向量量化编码为第一码流,而对多维特征图离散编码为第二码流,以实现空域无关向量和多维特征图的高效压缩。由于编码码流由两层码流构成,包含图像的不同层信息,这样即使在低码率下也不会损失很多信息,因此由两层码流重建获得的图像,可以改善视觉效果,提升视觉体验。

[0064] 实施例二:

图5为根据一示例性实施例示出的另一种图像处理方法的实施例流程图,基于上述图1所示实施例给出的是编码端的图像处理方案,在此基础上,本实施例给出的图像处理方法应用于解码端,如图5所示,该图像处理方法包括如下步骤:

步骤501:接收来自编码端的第一码流和第二码流。

[0065] 步骤502:对该第一码流进行反量化解码,得到一维特征向量。

[0066] 在该步骤中,如前所述,第一码流是使用量化编码模型编码获得,相应的,第一码流的解码过程,可以通过预设的反量化解码模型对第一码流进行反量化解码,得到一维特征向量,从而实现第一码流的解码。

[0067] 在一可行实施方式中,如图6所示的反量化解码模型,包括熵解码网络、反量化层和特征解码网络,解码过程包括:通过熵解码网络将第一码流解码为量化特征,然后通过反量化层将量化特征转换为浮点数值,得到反量化特征,最后通过特征解码网络对反量化特征进行上采样处理,得到一维特征向量。

[0068] 步骤503:对该第二码流进行反离散解码,得到多维特征图。

[0069] 在该步骤中,如前所述,第二码流是使用离散编码模型编码获得,相应的,第二码流的解码过程,可以通过预设离散解码模型对第二码流进行反离散解码,得到多维特征图;从而实现第二码流的解码。

[0070] 在一可行实施方式中,如图7所示的离散解码模型,包括码流解码层、反离散表示层和特征解码网络,解码过程包括:通过码流解码层将第二码流解码为索引矩阵,然后通过

反离散表示层基于预设特征字典将索引矩阵转换为连续特征,最后通过特征解码网络对连续特征进行上采样处理,得到多维特征图。

[0071] 其中,反离散表示层使用的特征字典与离散表示层使用的是相同的特征字典,以便于解码索引矩阵。基于上述所述的离散表示过程的基础上,索引矩阵转换为连续特征的过程可以是将索引矩阵中的每个索引利用特征字典中与索引对应的嵌入向量表示,从而得到连续特征。

[0072] 步骤504:基于一维特征向量对多维特征图进行图像重建,得到重建图像块。

[0073] 如前所述,由于多维特征图是由扩散模型的前向扩散过程获得,因此由多维特征图进行图像生成的实现由扩散模型的后向过程获得,也即由后向过程进行去噪解码。

[0074] 基于此,通过预设的扩散模型利用一维特征向量对多维特征图进行多次迭代去噪,以得到重建图像块。

[0075] 在一可行实施方式中,如图2所示的扩散模型,扩散模型的后向过程包括第二噪声预测网络和去噪层,在对多维特征图进行多次迭代去噪过程中,对于每次去噪过程,都通过第二噪声预测网络根据上一次去噪结果和一维特征向量预测当前次的噪声,然后通过去噪层利用第二噪声预测网络输出的噪声对上一次去噪结果进行去噪;

其中,对于首次去噪过程,上一次去噪结果为多维特征图。具体地,针对去噪层的具体去噪实现过程包括:获取上一次使用的噪声水平系数和当前次的噪声水平系数,然后根据上一次使用的噪声水平系数、当前次的噪声水平系数、第二噪声预测网络输出的噪声、上一次去噪结果获得当前次的去噪结果。

[0076] 基于上述所述的加噪公式,对应的去噪公式可选为:

$$x_{t-1} = \sqrt{\alpha_{t-1}} \frac{(x_t - \sqrt{1 - \alpha_t} \varepsilon_\theta(x_t, t, z_t))}{\sqrt{\alpha_t}} + \sqrt{1 - \alpha_{t-1}} \varepsilon_\theta(x_t, t, z_t)$$

[0077] 其中, x_t 表示上一次的去噪结果, x_{t-1} 表示当前次的去噪结果, $\varepsilon_\theta(x_t, t, z_t)$ 表示第二噪声预测网络输出的噪声, z_t 表示所述一维特征向量, α_t 表示上一轮的噪声水平系数, α_{t-1} 表示当前轮的噪声水平系数。

[0078] 至此,完成上述图5所示的图像处理流程,由于接收的第一码流和第二码流包含了图像的不同层面信息,第一码流包含有图像的纹理风格信息,第二码流包含图像的高频信息和结构信息,因此由第一码流和第二码流解码获得的重建图像,可以提升图像视觉效果,并且即使在带宽限制条件下也不会损失很多信息。

[0079] 本申请实施例的执行主体可以是应用程序、服务、实例、软件形态的功能模块、虚拟机(Virtual Machine, VM)、容器或云服务器等,或者具有数据处理功能的硬件设备(如服务器或终端设备)或硬件芯片(如CPU、GPU、FPGA、NPU、AI加速卡或DPU)等。实现图像处理的装置可以部署在提供相应服务的应用方的计算设备或提供算力、存储和网络资源的云计算平台上,云计算平台对外提供服务的模式可以是IaaS(Infrastructure as a Service,基础设施即服务)、PaaS(Platform as a Service,平台即服务)、SaaS(Software as a Service,软件即服务)或DaaS(Data as a Service,数据即服务)。以平台提供SaaS软件即服务(Software as a Service)为例,云计算平台可以利用自身的计算资源提供图像处理模型的训练或图像处理模块的功能执行,具体的应用架构可以根据服务需求进行搭建。例

如,平台可以向使用平台资源的应用方或个人提供基于上述模型的构建服务,进一步基于相关客户端或服务器等设备提交的图像处理请求调用上述模型和实现在线或离线图像处理的功能。

[0080] 基于上述图1和图5所示实施例基础上,本申请还提出一种图像处理系统,该系统包括编码端和解码端,其中编码端用于执行上述图1所示实施例的图像处理方法,解码端用于执行上述图5所示实施例的图像处理方法。

[0081] 下面将编码端与解码端结合在一起,对本申请提出的图像处理方法进行详细阐述。

[0082] 在编码端,首先,如图8所示,通过卷积编码器提取原始图像块 x_0 的空域无关特征向量 Z_l (也即一维特征向量),并基于扩散模型的扩散过程,利用空域无关特征向量 Z_l 对原始图像块 x_0 进行逐次加噪,输出噪声图 X_T (也即多维特征图);然后,如图9所示,通过量化编码模型对空域无关特征向量 Z_l 依次进行特征编码、量化处理以及熵估计,得到第一码流;如图10所示,通过离散编码模型对噪声图 X_T 依次进行特征编码、离散表示,得到索引矩阵,并将索引矩阵转换为第二码流;

[0083] 在解码端,首先,如图9所示,通过反量化解码模型对第一码流进行反量化、特征解码,得到解码后的空域无关特征向量 Z'_l ,如图10所示,通过离散解码模型对第二码流依次进行反离散表示、特征解码,得到重建后的噪声图 $\hat{X}_T$;然后,如图8所示,基于扩散模型的后向过程,利用重建后的空域无关特征向量 Z'_l 对重建后的噪声图 $\hat{X}_T$ 进行逐次去噪,输出重建后的图像块。

[0084] 基于上述描述,如图11所示,使用同一输入图像,通过HiFiC方案对原始图进行编码和重建后,得到图a,通过JPEG方案对原始图进行编码和重建后得到图b,使用本申请方案对原始图进行编码和重建后得到图c,在原始图、图a、图b、图c中相同区域截图与原始图的区域截图对比可以发现,使用HiFiC方案和JPEG方案得到的重建后效果均有伪影,而使用本申请方案得到的重建后效果没有伪影,与原始图效果几乎一致。

[0085] 与前述图像处理方法的实施例相对应,本申请还提供了一种图像处理装置的实施例,该装置用于执行上述图1所示实施例提供的图像处理方法,应用于编码端,该图像处理装置包括:

第一提取模块,用于提取原始图像块的一维特征向量;

第二提取模块,用于基于所述一维特征向量将所述原始图像块变换为多维特征图;

第一编码模块,用于对所述一维特征向量进行量化编码,得到第一码流;

第二编码模块,用于对所述多维特征图进行离散编码,得到第二码流;

发送模块,用于将所述第一码流和所述第二码流发送至解码端。

[0086] 所述基于所述一维特征向量将所述原始图像块变换为多维特征图,包括:

在一可选实现方式中,所述第二提取模块,具体用于通过预设的扩散模型利用所述一维特征向量对所述原始图像块进行多次迭代加噪,得到多维特征图。

[0087] 在一可选实现方式中,所述扩散模型包括第一噪声预测网络和加噪层;所述第二

提取模块,具体用于针对每次加噪过程,通过所述第一噪声预测网络根据上一次加噪结果和所述一维特征向量预测当前次的噪声;通过所述加噪层将所述第一噪声预测网络输出的噪声加入到上一次加噪结果上;其中,对于首次加噪过程,上一次加噪结果为所述原始图像块。

[0088] 在一可选实现方式中,所述第二提取模块,具体用于针对通过所述加噪层将所述第一噪声预测网络输出的噪声加入到上一次加噪结果上的过程,获取上一次使用的噪声水平系数和当前次的噪声水平系数;根据上一次使用的噪声水平系数、当前次的噪声水平系数、所述第一噪声预测网络输出的噪声、所述上一次加噪结果获得当前次的加噪结果。

[0089] 在一可选实现方式中,所述第一编码模块,具体用于通过预设的量化编码模型对所述一维特征向量进行量化编码,得到第一码流。

[0090] 在一可选实现方式中,所述量化编码模型包括特征编码网络、量化层、熵估计网络;所述第一编码模块,具体用于通过所述特征编码网络对所述一维特征向量进行下采样处理,得到下采样特征;通过所述量化层将所述下采样特征的特征值转换为整型数值,得到量化特征;通过所述熵估计网络将所述量化特征编码为第一码流。

[0091] 在一可选实现方式中,所述第二编码模块,具体用于通过预设的离散编码模型对所述多维特征图进行离散编码,得到第二码流。

[0092] 在一可选实现方式中,所述离散编码模型包括特征编码网络、离散表示层、码流转换层;所述第二编码模块,具体用于通过所述特征编码网络对所述多维特征图进行下采样处理,得到下采样特征;通过所述离散表示层对所述下采样特征进行离散表示,得到索引矩阵;通过所述码流转换层将所述索引矩阵编码为第二码流。

[0093] 在一可选实现方式中,所述第二编码模块,具体用于在通过所述离散表示层对所述下采样特征进行离散表示,得到索引矩阵过程中,在预设特征字典中,为所述下采样特征的每维特征向量匹配最相似的向量;利用与每维特征向量匹配的向量的索引获得索引矩阵。

[0094] 与前述图像处理方法的实施例相对应,本申请还提供了另一种图像处理装置的实施例,该装置用于执行上述图5所示实施例提供的图像处理方法,应用于解码端,该图像处理装置包括:

接收模块,接收来自编码端的第一码流和第二码流;

第一解码模块,用于对所述第一码流进行反量化解码,得到一维特征向量;

第二解码模块,用于对所述第二码流进行反离散解码,得到多维特征图;

重建模块,用于基于所述一维特征向量对所述多维特征图进行图像重建,得到重建图像块。

[0095] 在一可选实现方式中,所述第一解码模块,具体用于通过预设的反量化解码模型对所述第一码流进行反量化解码,得到一维特征向量;

所述反量化解码模型包括熵解码网络、反量化层和特征解码网络;所述第一解码模块,具体用于通过所述熵解码网络将所述第一码流解码为量化特征;通过所述反量化层将所述量化特征转换为浮点数值,得到反量化特征;通过所述特征解码网络对所述反量化特征进行上采样处理,得到一维特征向量。

[0096] 在一可选实现方式中,所述第二解码模块,具体用于通过预设的离散解码模型对

所述第二码流进行反离散解码,得到多维特征图;

所述离散解码模型包括码流解码层、反离散表示层和特征解码网络;所述第二解码模块,具体用于通过所述码流解码层将所述第二码流解码为索引矩阵;通过所述反离散表示层基于预设特征字典将所述索引矩阵转换为连续特征;通过所述特征解码网络对所述连续特征进行上采样处理,得到多维特征图。

[0097] 在一可选实现方式中,所述重建模块,具体用于通过预设的扩散模型利用所述一维特征向量对所述多维特征图进行多次迭代去噪,得到重建图像块。

[0098] 在一可选实现方式中,所述扩散模型包括第二噪声预测网络和去噪层;所述重建模块,具体用于针对每次去噪过程,通过所述第二噪声预测网络根据上一次去噪结果和所述一维特征向量预测当前次的噪声;通过所述去噪层利用所述第二噪声预测网络输出的噪声对上一次去噪结果进行去噪;其中,对于首次去噪过程,上一次去噪结果为所述多维特征图。

[0099] 在一可选实现方式中,所述重建模块,具体用于针对通过所述去噪层利用所述第二噪声预测网络输出的噪声对上一次去噪结果进行去噪的过程,获取上一次使用的噪声水平系数和当前次的噪声水平系数;根据上一次使用的噪声水平系数、当前次的噪声水平系数、所述第二噪声预测网络输出的噪声、所述上一次去噪结果获得当前次的去噪结果。

[0100] 上述装置中各个单元的功能和作用的实现过程具体详见上述方法中对应步骤的实现过程,在此不再赘述。

[0101] 对于装置实施例而言,由于其基本对应于方法实施例,所以相关之处参见方法实施例的部分说明即可。以上所描述的装置实施例仅仅是示意性的,其中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模块来实现本申请方案的目的。本领域普通技术人员在不付出创造性劳动的情况下,即可以理解并实施。

[0102] 本申请实施方式还提供一种与前述实施方式所提供的图像处理方法对应的电子设备,以执行上述图像处理方法。该电子设备包括:通信接口、处理器、存储器和总线;其中,通信接口、处理器和存储器通过总线完成相互间的通信。处理器通过读取并执行存储器中与图像处理方法的控制逻辑对应的机器可执行指令,可执行上文描述的图像处理方法,该方法的具体内容参见上述实施例,此处不再累述。

[0103] 本申请中提到的存储器可以是任何电子、磁性、光学或其它物理存储装置,可以包含存储信息,如可执行指令、数据等等。具体地,存储器可以是RAM(Random Access Memory, 随机存取存储器)、闪存、存储驱动器(如硬盘驱动器)、任何类型的存储盘(如光盘、DVD等),或者类似的存储介质,或者它们的组合。通过至少一个通信接口(可以是有线或者无线)实现该系统网元与至少一个其他网元之间的通信连接,可以使用互联网、广域网、本地网、城域网等。

[0104] 总线可以是ISA总线、PCI总线或EISA总线等。所述总线可以分为地址总线、数据总线、控制总线等。其中,存储器用于存储程序,所述处理器在接收到执行指令后,执行所述程序。

[0105] 处理器可能是一种集成电路芯片,具有信号的处理能力。在实现过程中,上述方法

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。上述的处理器可以是通用处理器,包括网络处理器(Network Processor,简称NP)、数字信号处理器(DSP)、专用集成电路(ASIC)、现成可编程门阵列(FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。可以实现或者执行本申请实施例中的公开的各方法、步骤及逻辑框图。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。结合本申请实施例所公开的方法的步骤可以直接体现为硬件译码处理器执行完成,或者用译码处理器中的硬件及软件模块组合执行完成。

[0106] 本申请实施例提供的电子设备与本申请实施例提供的图像处理方法出于相同的发明构思,具有与其采用、运行或实现的方法相同的有益效果。

[0107] 本领域技术人员在考虑说明书及实践这里公开的发明后,将容易想到本申请的其它实施方案。本申请旨在涵盖本申请的任何变型、用途或者适应性变化,这些变型、用途或者适应性变化遵循本申请的一般性原理并包括本申请未公开的本技术领域中的公知常识或惯用技术手段。说明书和实施例仅被视为示例性的,本申请的真正范围和精神由下面的权利要求指出。

[0108] 还需要说明的是,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、商品或者设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、商品或者设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括所述要素的过程、方法、商品或者设备中还存在另外的相同要素。

[0109] 以上所述仅为本申请的较佳实施例而已,并不用以限制本申请,凡在本申请的精神和原则之内,所做的任何修改、等同替换、改进等,均应包含在本申请保护的范围之内。

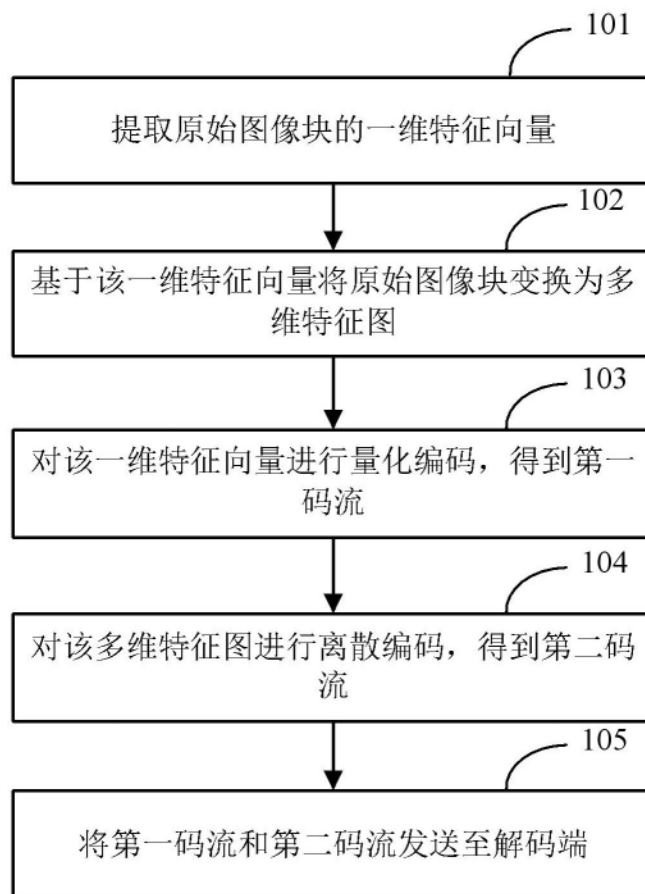


图1

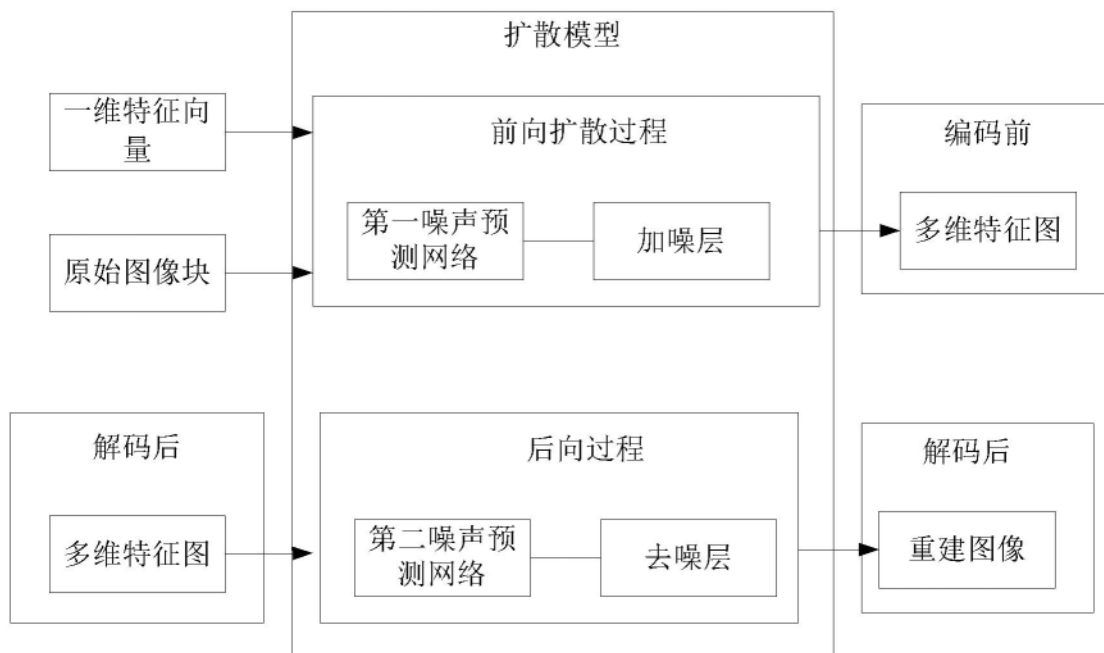


图2

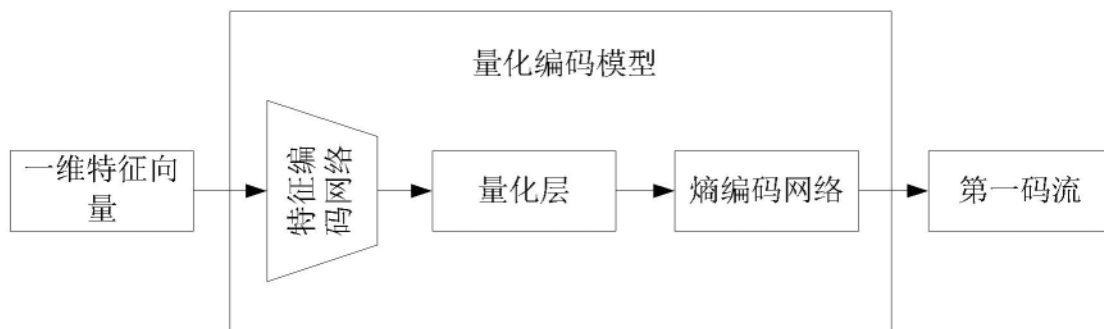


图3

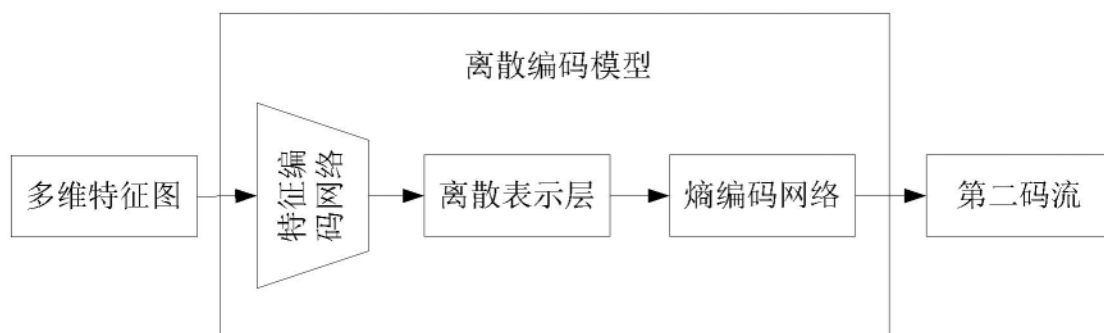


图4

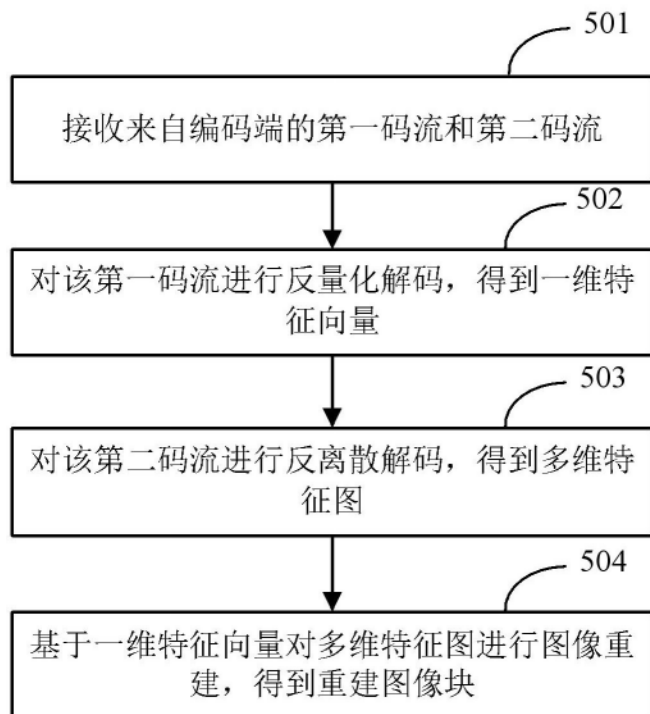


图5

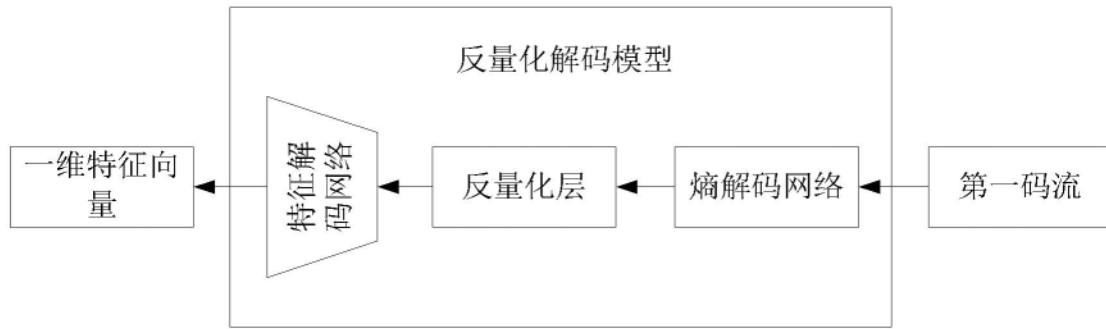


图6

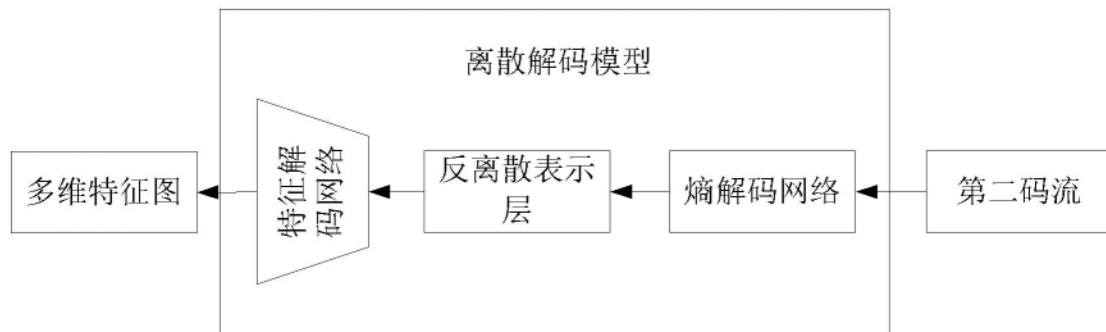


图7

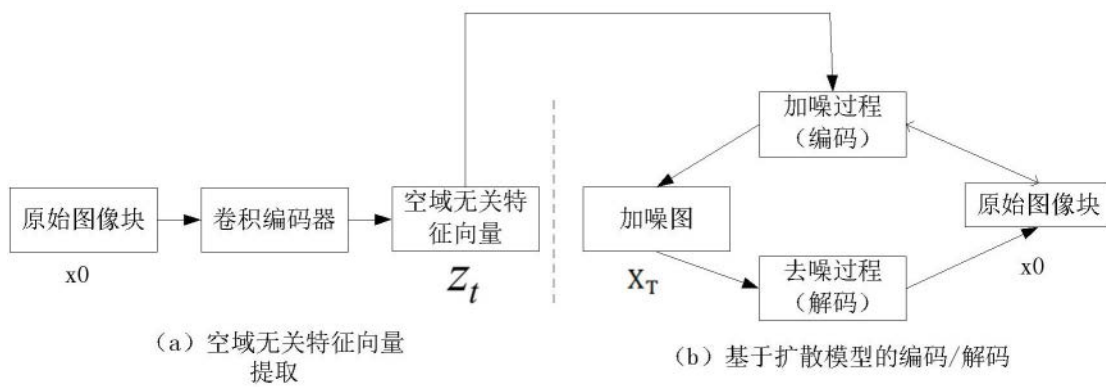


图8

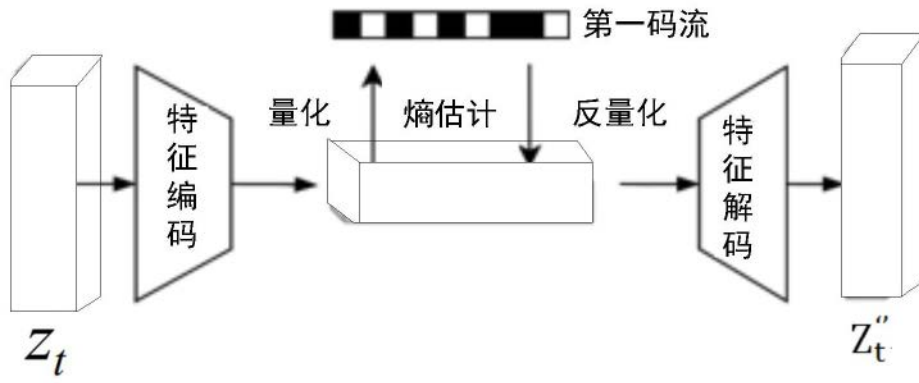


图9

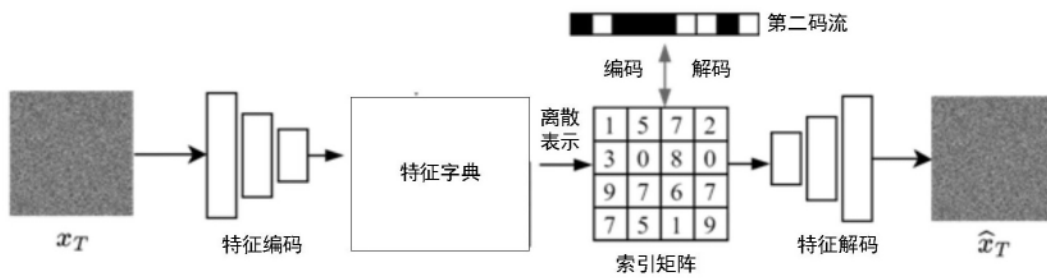


图10

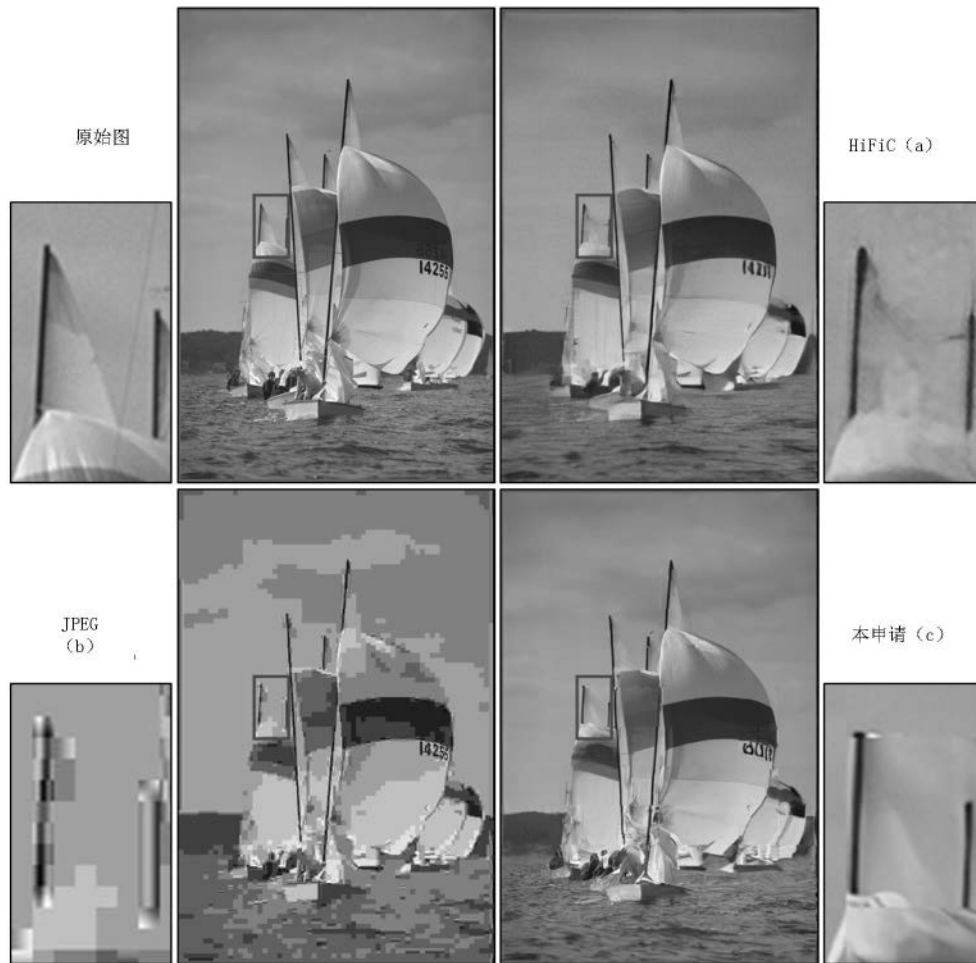


图11