

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2022 年 12 月 8 日 (08.12.2022)



(10) 国际公布号
WO 2022/253249 A1

- (51) 国际专利分类号:
H04N 19/13 (2014.01) G10L 25/30 (2013.01)
- (21) 国际申请号: PCT/CN2022/096510
- (22) 国际申请日: 2022 年 6 月 1 日 (01.06.2022)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202110616029.2 2021年6月2日 (02.06.2021) CN
202110674299.9 2021年6月17日 (17.06.2021) CN
202111091143.4 2021年9月17日 (17.09.2021) CN
- (71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东

省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN).

(72) 发明人: 毛珏(MAO, Jue); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 赵寅(ZHAO, Yin); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 闫宁(YAN, Ning); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 杨海涛(YANG, Haitao); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 张恋(ZHANG, Lian); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 王晶(WANG, Jing); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 师一博(SHI, Yi)

(54) Title: FEATURE DATA ENCODING METHOD AND APPARATUS AND FEATURE DATA DECODING METHOD AND APPARATUS

(54) 发明名称: 特征数据编解码方法和装置

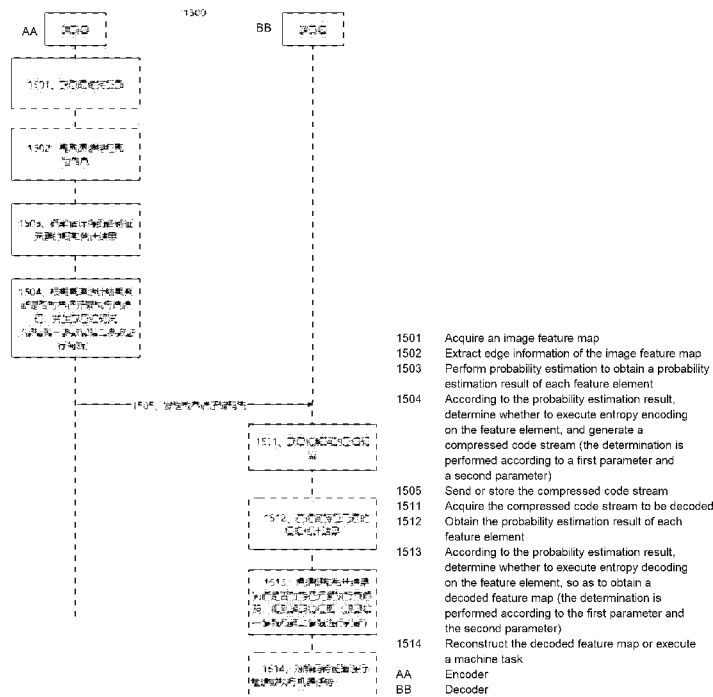


图 11A

(57) Abstract: The present application relates to the technical field of encoding and decoding an image or audio on the basis of artificial intelligence (AI), and specifically relates to the technical field of encoding and decoding an image feature map or audio feature variable on the basis of a neural network. Provided are an image or audio encoding method and apparatus and an image or audio decoding method and apparatus. The encoding method comprises: acquiring a target to be encoded, wherein said target comprises a plurality of feature elements, and the plurality of feature elements comprise a first feature element. The method further comprises: acquiring

WO 2022/253249 A1

Yibo); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

- (81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。
- (84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

a probability estimation result of the first feature element, and determining, according to the probability estimation result of the first feature element, whether to execute entropy encoding on the first feature element; and only when it is determined that it is necessary to execute entropy encoding on the first feature element, executing entropy encoding on the first feature element. By means of the present application, whether to encode a feature element is determined according to a probability estimation result. In this way, encoding and decoding complexity can be reduced without affecting the encoding and decoding performance. A target to be encoded comprises an image feature map or an audio feature variable.

(57) 摘要: 本申请提供了图像或音频编解码方法和装置, 涉及基于人工智能(AI)的图像或音频的编解码技术领域, 具体涉及基于神经网络的图像特征图或音频特征变量的编解码技术领域。其中编码方法包括: 获取待编码目标, 所述待编码目标包括多个特征元素, 所述多个特征元素包括第一特征元素。所述方法还包括获取所述第一特征元素的概率估计结果, 根据所述第一特征元素的概率估计结果, 判断是否对所述第一特征元素执行熵编码; 仅当判断出需要对所述第一特征元素执行熵编码时, 对所述第一特征元素执行熵编码。本申请根据概率估计结果来判断是否对特征元素编码。这样, 能够在不影响编解码性能情况下降低编解码复杂度, 其中待编码目标包括图像特征图或音频特征变量。

特征数据编解码方法和装置

技术领域

5 本发明实施例涉及基于人工智能 (AI) 的图像或音频压缩技术领域, 尤其涉及一种特征数据编解码方法及装置。

背景技术

10 图像或音频编码和解码 (简称为编解码) 广泛用于数字图像或音频应用, 例如广播数字电视、互联网和移动网络上的图像或音频传输、视频或语音聊天、和视频或语音会议等实时会话应用、DVD 和蓝光光盘、图像或音频内容采集和编辑系统以及可携式摄像机的安全应用。视频由多帧图像组成, 因此本申请中的图像可以是单独的图像, 也可以为视频中的图像。

15 即使在影片较短的情况下也需要对大量的视频数据进行描述, 当数据要在带宽容量受限的网络中发送或以其它方式传输时, 这样可能会造成困难。因此, 图像 (或音频) 数据通常要先压缩然后在现代电信网络中传输。由于内存资源可能有限, 当在存储设备上存储视频时, 图像 (或音频) 的大小也可能成为问题。图像 (或音频) 压缩设备通常在信源侧使用软件和/或硬件, 以在传输或存储之前对图像 (或音频) 数据进行编码, 从而减少用来表示数字图像 (或音频) 所需的数据量。然后, 压缩的数据在目的地侧由图像 (或音频) 解压缩设备接收。在有限的网络资源以及对更高图像 (或音频) 质量的需求不断增长的情
20 况下, 需要改进压缩和解压缩技术, 这些改进的技术能够提高压缩率而几乎不影响图像 (或音频) 质量。

近年来, 将深度学习应用于在图像 (或音频) 编解码领域逐渐成为一种趋势。如谷歌已连续几年在 CVPR (IEEE Conference on Computer Vision and Pattern Recognition) 会议上组织 CLIC (Challenge on Learned Image Compression) 专题竞赛, CLIC 专注使用深度神经网络来提升图像的压缩效率, 在 2020 年 CLIC 中还加入了图像挑战类别。基于竞赛方案的性能评估, 当前基于深度学习技术的图像编解码方案的综合压缩效率已经与最新一代视频
25 图像编解码标准 VVC (Versatile Video Coding) 相当, 而且在提升用户感知质量方面有独特优势。

30 VVC 的视频标准制定工作已于 2020 年 6 月完成, 标准收纳几乎所有能够带来显著压缩效率提升的技术算法。因此沿传统信号处理路径继续研究新型的压缩编码算法在短时间内难以获得大的技术突破。区别于传统图像算法通过人工设计来对图像压缩的各模块进行优化, 端到端 AI 的图像压缩是作为一个整体共同进行优化, 因此 AI 图像压缩方案的压缩

效果更好。变分自编码器 (Variational Autoencoder, AE) 方法是当前 AI 图像有损压缩技术的主流技术方案。在目前的主流技术方案是待编码图像通过编码网络获得图像特征图, 并进一步对图像特征图执行熵编码, 但是熵编码过程存在着复杂度过高的问题。

5 发明内容

本申请提供一种特征数据的编解码方法和装置, 能够在不影响编解码性能情况下降低编解码复杂度。

第一方面, 提供了一种特征数据的编码方法, 包括:

获取待编码特征数据, 所述待编码特征数据包括多个特征元素, 所述多个特征元素包

10 括第一特征元素;

获取所述第一特征元素的概率估计结果;

根据所述第一特征元素的概率估计结果, 判断是否对所述第一特征元素执行熵编码;

仅当判断出需要对所述第一特征元素执行熵编码时, 对所述第一特征元素执行熵编码。

15 其中, 所述特征数据包括图像特征图, 或音频特征变量, 或图像特征图和音频特征变量。可以为编码网络所输出的一维、二维或多维数据, 其中每个数据均为特征元素。需要说明的是, 本申请中特征点和特征元素的含义相同。

具体的, 所述第一特征元素为待编码特征数据中的任意一待编码特征元素。

20 一种可能性中, 获取所述第一特征元素的概率估计结果的概率估计过程可以通过概率估计网络实现; 在另一种可能性中, 概率估计过程可以采用传统非网络的概率估计方法对特征数据进行概率估计。

需要说明的是, 当只有边信息作为概率估计的输入时, 可以并行输出各特征元素的概率估计结果; 当概率估计的输入包括有上下文信息时, 需要串行输出各特征元素概率估计结果。其中所述边信息为特征数据输入神经网络进一步提取得到的特征信息, 所述边信息包含的特征元素的个数比特征数据的特征元素少。可选地, 可以将特征数据的边信息编入

25 码流。
一种可能性中, 当所述特征数据的第一特征元素不满足预设条件下, 不需要对所述特征数据的第一特征元素执行熵编码。

具体的, 假如当前的第一特征元素为特征数据的第 P 个特征元素, 则完成第 P 个特征元素的判断和根据判断结果执行或不执行熵编码后, 开始特征数据的第 P+1 个特征元素的判断和根据判断结果执行或不执行熵编码过程, 其中 P 为正整数且 P 小于 M, 其中 M 为整个的特征数据中特征元素的数量。比如对第二特征元素, 当判断不需要对所述第二特征元

素执行熵编码时，则对所述第二特征元素跳过执行熵编码。

上述技术方案中，通过对每个待编码的特征元素进行判定是否需要执行熵编码，从而跳过某些特征元素的熵编码过程，可以显著减少需执行熵编码的元素个数。这样，可以降低熵编码复杂度。

5 在一种可能的实现方式中，判断是否对所述第一特征元素执行熵编码包括：当所述第一特征元素的概率估计结果满足预设条件时，判断需要对所述第一特征元素执行熵编码；或当所述第一特征元素的概率估计结果不满足预设条件时，判断不需要对所述第一特征元素执行熵编码。

10 在一种可能的实现方式中，所述第一特征元素的概率估计结果为所述第一特征元素取值为 k 的概率值，所述预设条件为第一特征元素取值为 k 的概率值小于或等于第一阈值，其中 k 为整数。

k 为上述第一特征元素取值所可能的取值范围中的某一取值。比如，第一特征元素可以取值的范围为【-255, 255】。 k 可以设置为 0，则对概率值小于或者等于 0.5 的第一特征元素，执行熵编码。对概率值大于 0.5 的第一特征元素，不执行熵编码。

15 在一种可能的实现方式中，所述第一特征元素取值为 k 的概率值为所述第一特征元素的所有可能的取值的概率值中的最大概率值。

其中，编码码流在低码率情况所选定的第一阈值小于编码码流在高码率情况所选定的第一阈值。具体码率高低与图像的分辨率以及图像内容相关，以公开的 Kodak 数据集为例，低于 0.5bpp 为低码率，反之为高码率。

20 在某一码率情况下，所述第一阈值可以根据实际需要进行配置，此处不做限定。

上述技术方案中，通过灵活的第一阈值设定方式使得产生可以根据要求灵活的降低熵编码复杂度。

在一种可能的实现方式中，所述第一特征元素的概率估计结果包括所述第一特征元素概率分布的第一参数和第二参数。

25 则当所述概率分布为高斯分布，所述第一特征元素概率分布的第一参数为所述第一特征元素高斯分布的均值，所述第一特征元素概率分布的第二参数为所述第一特征元素高斯分布的方差；或当所述概率分布为拉普拉斯分布，所述第一特征元素概率分布的第一参数为所述第一特征元素拉普拉斯分布的位置参数，所述第一特征元素概率分布的第二参数为所述第一特征元素拉普拉斯分布的尺度参数。则所述预设条件可以为以下任意一种：

30 所述第一特征元素的概率分布的第一参数与所述第一特征元素取值 k 的差的绝对值

大于或等于第二阈值；或

所述第一特征元素的概率分布的第二参数大于或等于第三阈值；或

所述第一特征元素的概率分布的第一参数与所述第一特征元素取值 k 的差的绝对值与所述第一特征元素的概率分布的第二参数的和大于或等于第四阈值。

5 则当所述概率分布为混合高斯分布时，所述第一特征元素概率分布的第一参数为所述第一特征元素混合高斯分布的均值，所述第一特征元素概率分布的第二参数为所述第一特征元素混合高斯分布的方差，则所述预设条件可以为以下任意一种：

所述第一特征元素的混合高斯分布的所有均值与所述第一特征元素的取值为 k 的差的绝对值之和与所述第一特征元素的混合高斯分布的任一方差的和大于或等于第五阈值；

10 或

所述第一特征元素的混合高斯分布的任一均值与所述第一特征元素的取值为 k 的差大于或等于第六阈值；或

所述第一特征元素的混合高斯分布的任一方差大于或等于第七阈值。

15 则当所述概率分布为非对称高斯分布时，所述第一特征元素概率分布的第一参数为所述第一特征元素非对称高斯分布的均值，所述第一特征元素概率分布的第二参数为所述第一特征元素非对称高斯分布的第一方差和第二方差，则所述预设条件可以为以下任意一种：

所述第一特征元素的非对称高斯分布的均值与所述第一特征元素的取值为 k 的差的绝对值大于或等于第八阈值；或

所述第一特征元素的非对称高斯分布的第一方差大于或等于第九阈值；

20 所述第一特征元素的非对称高斯分布的第二方差大于或等于第十阈值。

所述第一特征元素的概率分布为混合高斯分布情况，确定所述第一特征元素的判断取值范围，当所述第一特征元素的概率分布的多个均值均不在所述第一特征元素的判断取值范围。

25 所述第一特征元素的概率分布为高斯分布情况，确定所述第一特征元素的判断取值范围，当所述第一特征元素的概率分布的均值不在所述第一特征元素的判断取值范围。

所述第一特征元素的概率分布为高斯分布情况，确定所述第一特征元素的判断取值范围，判断取值范围中包括了多个所述第一特征元素可能的取值，当所述第一特征元素的高斯分布的均值参数与所述第一特征元素的判断取值范围中的每个取值的差的绝对值大于或等于第十一阈值，或所述第一特征元素的概率分布的方差大于或等于第十二阈值。

30 所述第一特征元素的取值不在所述第一特征元素的判断取值范围。

所述第一特征元素的取值对应的概率值小于或等于第十三阈值。

在一种可能的实现方式中，所述方法还包括：构建第一阈值的阈值候选列表，将所述第一阈值放入所述第一阈值的阈值候选列表中且对应有所述第一阈值的索引号，将所述第一阈值的索引号写入编码码流，其中所述第一阈值的阈值候选列表的长度可以设置为 T；
5 T 为大于或等于 1 的整数。可以理解的，所述其他的阈值可以采用如第一阈值的阈值候选列表构建方式，且有对应阈值的索引号和写入编码码流中。

具体地，将所述索引号写入码流，可将其保存在序列头（sequence header）、图像头（picture header）、Slice/条带（slice header）或 SEI（supplemental enhancement information）中传送到解码端，还可以使用其他方法，在此不做限定。构建候选列表的
10 方式不做限定。

另一种可能性中，根据概率估计结果输入生成网络得到所述决策信息。所述生成网络可以为卷积网络，可以包括多个网络层，任意一网络层可以为卷积层、归一化层、非线性激活层等。

在一种可能的实现方式中，将所述特征数据的概率估计结果输入生成网络以得到所述
15 第一特征元素的决策信息，所述决策信息用于指示是否对所述第一特征元素执行熵编码。

在一种可能的实现方式中，所述特征数据的决策信息为决策图，决策图也可以称为决策图 map。决策图优选的为二元图，二元图也可以称为二元图 map。二元图中特征元素的决策信息取值通常为 0 或 1。因此当所述决策图中对应所述第一特征元素所在位置的值为预设值时，需要对所述第一特征元素执行熵编码；当所述决策图中对应所述第一特征元素
20 所在位置的值为不为预设值时，不需要对所述第一特征元素执行熵编码。

在一种可能的实现方式中，所述特征数据中的特征元素的决策信息为预设值。所述决策信息的预设值取值通常为 1，因此当所述决策信息为预设值时，需要对所述第一特征元素执行熵编码；当所述决策信息不为预设值时，不需要对所述第一特征元素执行熵编码。决策信息可以为标识或者标识的值。判断是否对所述第一特征元素执行熵编码取决于所述
25 标识或者标识的值是否为预设值，为预设值时，需要对所述第一特征元素执行熵编码；不为预设值时，不需要对所述第一特征元素执行熵编码。所述特征数据中各特征元素的决策信息的集合也可以为浮点数，也就是说取值可以为除 0 和 1 外的其他值。这时候，可以通过设置预设值，当所述第一特征元素的决策信息的值等于或者大于预设值时，判断需要对所述第一特征元素执行熵编码；或当所述第一特征元素的决策信息的值小于预设值时，
30 判断不需要对所述第一特征元素执行熵编码。

在一种可能的实现方式中，所述方法还包括：待编码图像经过编码网络获取所述特征数据；待编码图像经过编码网络后经过取整获取所述特征数据；或待编码图像经过编码网络后经过量化和取整获取所述特征数据。

其中编码网络可以采用自编码器结构。编码网络可以为卷积神经网络。编码网络可以包括多个子网络，每个子网络包含一个或多个卷积层。子网络间的网络结构可以互为相同或不同。

其中待编码图像可以是原始图像，也可以是残差图像。

应理解，待编码图像可以为 RGB 格式或 YUV、RAW 等表示格式，待编码图像在输入编码网络前可以进行预处理操作，预处理操作可以包括转换、块划分、滤波、剪枝等操作。

10 应理解，允许在同一时间戳内或同一时刻将多个待编码图像或多个待编码图像块输入编解码网络进行处理以得到特征数据。

第二方面，提供了一种特征数据的解码方法，包括：

获取待解码特征数据的码流；

所述待解码特征数据包括多个特征元素，所述多个特征元素包括第一特征元素；

15 获取所述第一特征元素的概率估计结果；

根据所述第一特征元素的概率估计结果，判断是否对所述第一特征元素执行熵解码；仅当判断出需要对所述第一特征元素执行熵解码时，对所述第一特征元素执行熵解码。

可以理解，所述第一特征元素为待解码特征数据中的任意特征元素，当待解码特征数据中所有特征元素完成了所述判断并根据判断结果执行或不执行熵解码后得到了解码特征数据。

20 其中，所述待解码特征数据可以为二维或多维数据，其中每个数据均为特征元素。需要说明的是，本申请中特征点和特征元素的含义相同。

具体的，所述第一特征元素为待解码特征数据中的任意一待解码特征元素。

一种可能性中，获取所述第一特征元素的概率估计结果的概率估计过程可以通过概率估计网络实现；在另一种可能性中，概率估计过程可以采用传统非网络的概率估计方法对特征数据进行概率估计。

需要说明的是，当只有边信息作为概率估计的输入时，可以并行输出各特征元素的概率估计结果；当概率估计的输入包括有上下文信息时，需要串行输出各特征元素概率估计结果。其中，所述边信息包含的特征元素的个数比特征数据的特征元素少。

30 一种可能性中，码流中包含了边信息，解码码流过程需要对边信息进行解码。

具体的，特征数据中的每个特征元素的判断过程包括了条件判断以及根据条件判断结果决定是否执行熵解码。

一种可能性中，熵解码可以通过神经网络方式实现。

在另一种可能性中，熵解码可以通过传统熵解码方式实现。

5 具体的，假如当前的第一特征元素为特征数据的第 P 个特征元素，则完成第 P 个特征元素的判断和根据判断结果执行或不执行熵解码后，开始特征数据的第 P+1 个特征元素的判断和根据判断结果执行或不执行熵解码过程，其中 P 为正整数且 P 小于 M，其中 M 为整个的特征数据中特征元素的数量。比如对第二特征元素，当判断不需要对所述第二特征元素执行熵解码时，则对所述第二特征元素跳过执行熵解码。

10 上述技术方案中，通过对每个待解码的特征元素进行判定是否需要执行熵解码，从而跳过某些特征元素的熵解码过程，可以显著减少需执行熵解码的元素个数。这样，可以降低熵解码复杂度。

在一种可能的实现方式中，所述判断是否对所述特征数据的第一特征元素执行熵解码包括：当所述特征数据的第一特征元素的概率估计结果满足预设条件，判断需要对所述第
15 一特征元素执行熵解码；或当所述第一特征元素的概率估计结果不满足预设条件时，判断不需要对所述第一特征元素熵解码，将所述第一特征元素的特征值设置为 k；其中 k 为整数。

在一种可能的实现方式中，所述第一特征元素的概率估计结果为所述第一特征元素取
20 值为 k 的概率值，所述预设条件为第一特征元素取值为 k 的概率值小于或等于第一阈值，其中 k 为整数。

一种可能性中，所述第一特征元素在不满足所述预设条件下设置为 k。比如，第一特征元素可以取值的范围为【-255，255】。k 可以设置为 0，则对概率值小于或者等于 0.5 的第一特征元素，执行熵编码。对概率值大于 0.5 的第一特征元素，不执行熵编码。

在另一种可能性中，所述第一特征元素在不满足所述预设条件下通过列表确定取值。

25 在另一种可能性中，所述第一特征元素在不满足所述预设条件下设置为固定整数值。k 为上述第一特征元素取值所可能的取值范围中的某一取值。

在一种可能性中，k 为上述第一特征元素中所有可能的取值范围中最大概率所对应的值。

其中，解码码流在低码率情况所选定的第一阈值小于解码码流在高码率情况所选定的
30 第一阈值。具体码率高低与图像的分辨率以及图像内容相关，以公开的 Kodak 数据集为例，

低于 0.5bpp 为低码率，反之为高码率。

在某一码率情况下，所述第一阈值可以根据实际需要进行配置，此处不做限定。

上述技术方案中，通过灵活的第一阈值设定方式使得产生可以根据要求灵活的降低熵解码复杂度。

5 在一种可能的实现方式中，所述第一特征元素的概率估计结果包括所述第一特征元素概率分布的第一参数和第二参数。

10 则当所述概率分布为高斯分布，所述第一特征元素概率分布的第一参数为所述第一特征元素高斯分布的均值，所述第一特征元素概率分布的第二参数为所述第一特征元素高斯分布的方差；或当所述概率分布为拉普拉斯分布，所述第一特征元素概率分布的第一参数为所述第一特征元素拉普拉斯分布的位置参数，所述第一特征元素概率分布的第二参数为所述第一特征元素拉普拉斯分布的尺度参数。则所述预设条件可以为以下任意一种：

所述第一特征元素的概率分布的第一参数与所述第一特征元素取值 k 的差的绝对值大于或等于第二阈值；或

所述第一特征元素的第二参数大于或等于第三阈值；或

15 所述第一特征元素的概率分布的第一参数与所述第一特征元素取值 k 的差的绝对值与所述第一特征元素的概率分布的第二参数的和大于或等于第四阈值。

则当所述概率分布为混合高斯分布时，所述第一特征元素概率分布的第一参数为所述第一特征元素混合高斯分布的均值，所述第一特征元素概率分布的第二参数为所述第一特征元素混合高斯分布的方差，则所述预设条件可以为以下任意一种：

20 所述第一特征元素的混合高斯分布的所有均值与所述第一特征元素的取值为 k 的差的绝对值之和与所述第一特征元素的混合高斯分布的任一方差的和大于或等于第五阈值；或

所述第一特征元素的混合高斯分布的任一均值与所述第一特征元素的取值为 k 的差大于第六阈值；或

25 所述第一特征元素的混合高斯分布的任一方差大于或等于第七阈值。

则当所述概率分布为非对称高斯分布时，所述第一特征元素概率分布的第一参数为所述第一特征元素非对称高斯分布的均值，所述第一特征元素概率分布的第二参数为所述第一特征元素非对称高斯分布的第一方差和第二方差，则所述预设条件可以为以下任意一种：

30 所述第一特征元素的非对称高斯分布的均值参数与所述第一特征元素的取值为 k 的差的绝对值大于第八阈值；或

所述第一特征元素的非对称高斯分布的第一方差大于或等于第九阈值；

所述第一特征元素的非对称高斯分布的第二方差大于或等于第十阈值。

所述第一特征元素的概率分布为混合高斯分布情况，确定所述第一特征元素的判断取值范围，当所述第一特征元素的概率分布的多个均值均不在所述第一特征元素的判断取值范围。

所述第一特征元素的概率分布为高斯分布情况，确定所述第一特征元素的判断取值范围，当所述第一特征元素的概率分布的均值不在所述第一特征元素的判断取值范围。

所述第一特征元素的概率分布为高斯分布情况，确定所述第一特征元素的判断取值范围，判断取值范围中包括了多个所述第一特征元素可能的取值，当所述第一特征元素的高斯分布的均值参数与所述第一特征元素的判断取值范围中的每个取值的差的绝对值大于或等于第十一阈值，或所述第一特征元素的概率分布的方差大于或等于第十二阈值。

所述第一特征元素取值为 k 不在所述第一特征元素的判断取值范围。

所述第一特征元素取值为 k 对应的概率值小于或等于第十三阈值。

在一种可能的实现方式中，构建第一阈值的阈值候选列表，通过对所述码流进行解码以得到所述第一阈值的阈值候选列表的索引号，将所述第一阈值的索引号所对应所述第一阈值的阈值候选列表位置的值作为所述第一阈值的值，其中所述第一阈值的阈值候选列表的长度可以设置为 T ； T 为大于或等于 1 的整数。可以理解的，所述其他任意的阈值可以采用如第一阈值的阈值候选列表构建方式，且可以解码对应阈值的索引号，并根据所述索引号选取构建列表里的值作为阈值。

另一种可能性中，根据概率估计结果输入生成网络得到所述决策信息。所述生成网络可以为卷积网络，可以包括多个网络层，任意一网络层可以为卷积层、归一化层、非线性激活层等。

在一种可能的实现方式中，将所述特征数据的概率估计结果输入生成网络以得到所述第一特征元素的决策信息，所述决策信息用于指示是否对所述第一特征元素执行熵解码。

在一种可能的实现方式中，所述特征数据中各特征元素的决策信息为决策图，决策图也可以称为决策图 **map**。决策图优选的为二元图，二元图也可以称为二元图 **map**。二元图中特征元素的决策信息取值通常为 0 或 1。因此当所述决策图中对应所述第一特征元素所在位置的值为预设值时，需要对所述第一特征元素执行熵解码；当所述决策图中对应所述第一特征元素所在位置的值为不为预设值时，不需要对所述第一特征元素执行熵解码。

所述特征数据中各特征元素的决策信息的集合也可以为浮点数，也就是说取值可以为

除 0 和 1 外的其他值。这时候，可以通过设置预设值，当所述第一特征元素的决策信息的值等于或者大于预设值时，判断需要对所述第一特征元素执行熵解码；或当所述第一特征元素的决策信息的值小于预设值时，判断不需要对所述第一特征元素执行熵解码。

在一种可能的实现方式中，所述特征数据经过解码网络以得到重建图像。

5 在另一种可能的实现方式中，所述特征数据经过解码网络以得到面向机器任务数据，具体的，所述特征数据经过面向机器任务模块以得到面向机器任务数据，所述面向机器模块包括目标识别网络，分类网络或者语义分割网络。

第三方面，提供了一种特征数据编码装置，包括：

10 获得模块，用于获取待编码特征数据，所述待编码特征数据包括多个特征元素，所述多个特征元素包括第一特征元素，以及用于获取所述中第一特征元素的概率估计结果；

编码模块，用于根据所述第一特征元素的概率估计结果，判断是否对所述第一特征元素执行熵编码；仅当判断出需要对所述第一特征元素执行熵编码时，对所述第一特征元素执行熵编码。

15 上述获得模块、编码模块的进一步实现功能可以参考第一方面或者第一方面的任意一种实现方式，此处不再赘述。

第四方面，提供了一种特征数据解码装置，包括：

获得模块，用于获取待解码特征数据的码流，所述待解码特征数据包括多个特征元素，所述多个特征元素包括第一特征元素；获取所述第一特征元素的概率估计结果；

20 解码模块，用于根据所述第一特征元素的概率估计结果，判断是否对所述第一特征元素执行熵解码；仅当判断出需要对所述第一特征元素执行熵解码时，对所述第一特征元素执行熵解码。

上述获得模块、解码模块的进一步实现功能可以参考第二方面或者第二方面的任意一种实现方式，此处不再赘述。

25 第五方面，本申请提供一种编码器，包括处理电路，用于判断根据上述第一方面及第一方面任一项所述的方法。

第六方面，本申请提供一种解码器，包括处理电路，用于判断上述第二方面及第二方面任一项所述的方法。

30 第七方面，本申请提供一种计算机程序产品，包括程序代码，当其在计算机或处理器上判断时，用于判断上述第一方面及第一方面任一项、上述第二方面及第二方面任一项所述的方法。

第八方面，本申请提供一种编码器，包括：一个或多个处理器；非瞬时性计算机可读存储介质，耦合到所述处理器并存储由所述处理器判断的程序，其中所述程序在由所述处理器判断时，使得所述解码器判断上述第一方面及第一方面任一项所述的方法。

第九方面，本申请提供一种解码器，包括：一个或多个处理器；非瞬时性计算机可读存储介质，耦合到所述处理器并存储由所述处理器判断的程序，其中所述程序在由所述处理器判断时，使得所述编码器判断上述第二方面及第二方面任一项所述的方法所述的方法。

第十方面，本申请提供一种非瞬时性计算机可读存储介质，包括程序代码，当其由计算机设备判断时，用于判断上述第一方面及第一方面任一项、上述第二方面及第二方面任一项所述的方法。

第十一方面，本发明涉及编码装置，具有实现上述第一方面或第一方面任一项的方法实施例中行为的功能。所述功能可以通过硬件实现，也可以通过硬件判断相应的软件实现。所述硬件或软件包括一个或多个与上述功能相对应的模块。在一个可能的设计中，所述编码装置包括：获得模块，用于将原始图像或残差图像经过编码网络变换到特征空间，提取特征数据用来压缩。另外对特征数据进行概率估计获取特征数据各特征元素的概率估计结果；编码模块，利用特征数据各特征元素的概率估计结果，用于通过一定条件来判断特征数据中各特征元素是否执行熵编码并完成所述特征数据中所有特征元素的编码过程以得到特征数据的编码码流。这些模块可以判断上述第一方面或第一方面任一项方法示例中的相应功能，具体参见方法示例中的详细描述，此处不做赘述。

第十二方面，本发明涉及解码装置，具有实现上述第二方面或第二方面任一项的方法实施例中行为的功能。所述功能可以通过硬件实现，也可以通过硬件判断相应的软件实现。所述硬件或软件包括一个或多个与上述功能相对应的模块。在一个可能的设计中，所述解码装置包括：获得模块，用于获取待解码特征数据的码流，并根据待解码特征数据的码流，进行概率估计以得到特征数据各特征元素的概率估计结果；解码模块，利用特征数据各特征元素的概率估计结果，通过一定条件来判断特征数据中各特征元素是否执行熵解码并完成所述特征数据中所有特征元素的解码过程以得到所述特征数据，并对所述特征数据进行解码以得到重建图像或面向机器任务数据。这些模块可以判断上述第二方面或第二方面任一项方法示例中的相应功能，具体参见方法示例中的详细描述，此处不做赘述。

第十三方面，提供了一种特征数据的编码方法，包括：

获取待编码特征数据，所述特征数据包括多个特征元素，所述多个特征元素包括第一特征元素；

获取所述特征数据的边信息，对所述特征数据的边信息输入联合网络以得到所述第一特征元素的决策信息；

根据所述第一特征元素的决策信息，判断是否对所述第一特征元素执行熵编码；

仅当判断出需要对所述第一特征元素执行熵编码时，对所述第一特征元素执行熵编码。

5 所述特征数据为编码网络所输出的一维、二维或多维数据，其中每个数据均为特征元素。

一种可能性中，将特征数据的边信息编入码流。所述边信息为特征数据输入神经网络进一步提取得到的特征信息，所述边信息包含的特征元素的个数比特征数据的特征元素少。

所述第一特征元素为所述特征数据中的任意特征元素。

10 一种可能性中，所述特征数据各特征元素的决策信息的集合可以以决策图等方式来进行表示。其中决策图为一维、二维或多维图像数据且与所述特征数据的尺寸一致。

一种可能性中，联合网络还输出所述第一特征元素的概率估计结果，所述第一特征元素的概率估计结果包括所述第一特征元素的概率值，和/或所述概率分布的第一参数和所述概率分布的第二参数。

15 上述技术方案中，通过对每个待编码的特征元素进行判定是否需要执行熵编码，从而跳过某些特征元素的熵编码过程，可以显著减少需执行熵编码的元素个数。这样，可以降低熵编码复杂度。

在一种可能性中，当所述决策图中对应第一特征元素位置的值为预设值时，需要对所述第一特征元素执行熵编码；当所述决策图中对应第一特征元素位置的值为不为预设值时，不需要对所述第一特征元素执行熵编码。

20 第十四方面，提供了一种特征数据的解码方法，包括：

获取待解码特征数据的码流和所述待解码特征数据的边信息；

所述待解码特征数据包括多个特征元素，所述多个特征元素包括第一特征元素；

对所述待解码特征数据的边信息输入联合网络以得到所述第一特征元素的决策信息；

25 根据所述第一特征元素的决策信息，判断是否对所述第一特征元素执行熵解码；

仅当判断出需要对所述第一特征元素执行熵解码时，对所述第一特征元素执行熵解码。

一种可能性中，解码待解码特征数据的码流以得到边信息。所述边信息包含的特征元素的个数比特征数据的特征元素少。

所述第一特征元素为所述特征数据中的任意特征元素。

30 一种可能性中，所述特征数据各特征元素的决策信息可以以决策图等方式来进行表示。

其中决策图为一维、二维或多维图像数据且与所述特征数据的尺寸一致。

一种可能性中，联合网络还输出所述第一特征元素的概率估计结果，所述第一特征元素的概率估计结果包括所述第一特征元素的概率值，和/或所述概率分布的第一参数和所述概率分布的第二参数。

- 5 在一种可能性中，当所述决策图中对应第一特征元素位置的值为预设值时，需要对所述第一特征元素执行熵解码；当所述决策图中对应第一特征元素位置的值为不为预设值时，不需要对所述第一特征元素执行熵解码，并将所述第一特征元素的特征值设置为 k ，其中 k 为整数。

- 10 上述技术方案中，通过对每个待编码的特征元素进行判定是否需要执行熵解码，从而跳过某些特征元素的熵解码过程，可以显著减少需执行熵解码的元素个数。这样，可以降低熵解码复杂度。

- 15 在现有端到端特征数据编解码主流方案中，在熵编解码或算术编解码过程中存在着复杂度过高问题。本申请利用了待编码特征数据中特征点概率分布相关信息，对每个待编解码特征数据中的特征元素进行判定是否需要需要进行熵编解码，从而跳过某些特征元素的熵编解码过程，可以显著减少需进行编解码的元素个数，降低了编解码复杂度。另一方面，可以根据码流的码率实际大小要求，灵活的对阈值大小进行设置以控制生成码流的码率大小。

附图及以下说明中将详细描述一个或多个实施例。其它特征、目的和优点在说明、附图以及权利要求中是显而易见的。

20 附图说明

下面对本申请实施例用到的附图进行介绍。

图 1A 为图像译码系统示例性框图；

图 1B 为图像译码系统处理电路实现；

图 1C 为图像译码设备示意性框图；

- 25 图 1D 为本申请实施例装置实现图；

图 2A 为本申请一种可能场景的系统架构图；

图 2B 为本申请一种可能场景的系统架构图；

图 3A-3D 为编码器示意性框图；

图 4A 为编码网络单元示意图；

- 30 图 4B 为编码网络的网络结构示意图；

图 5 为编码决策实现单元结构示意图；

图 6 为联合网络输出示例图；

图 7 为生成网络输出示例图；

图 8 为解码决策实现示意图；

图 9 为解码网络的网络结构示例图；

5 图 10A 为本申请实施例译码方法的一个示例图；

图 10B 为本申请实施例图像特征图解码器示意性框图；

图 11A 为本申请实施例译码方法的一个示例图；

图 12 为边信息提取模块网络结构示例图；

图 13A 为本申请实施例译码方法的一个示例图；

10 图 13B 为本申请实施例图像特征图解码器示意性框图；

图 14 为本申请实施例译码方法的一个示例图；

图 15 为联合网络的网络结构示例图；

图 16 为本申请实施例图像特征图解码器示意性框图；

图 17 为本申请实施例译码方法的一个示例图；

15 图 18 为本申请编码装置的一个示例性的结构示意图；

图 19 为本申请解码装置的一个示例性的结构示意图。

具体实施方式

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

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

本申请实施例提供一种基于 AI 的特征数据编解码技术，尤其是提供一种基于神经网络

络的图像特征图和/或音频特征变量的编解码技术，具体提供一种基于端到端的图像特征图和/或音频特征变量的编解码系统。

在图像编码领域，术语“图像 (picture)”、或“图片 (image)”，可以用作同义词。图像编码（或通常称为编码）包括图像编码和图像解码两部分，其中视频由多个图像所组成，是连续图像的代表方式。图像编码在源侧判断，通常包括处理（例如，压缩）原始视频图像以减少表示该视频图像所需的数据量（从而更高效存储和/或传输）。图像解码在目的地侧判断，通常包括相对于编码器作逆处理，以重建图像。实施例涉及的图像或音频的“译码”应理解为图像或音频的“编码”或“解码”。编码部分和解码部分也合称为编解码（编码和解码，CODEC）。

在无损图像编码情况下，可以重建原始图像，即重建的图像与原始图像具有相同的质量（假设存储或传输期间没有传输损耗或其它数据丢失）。在传统有损图像编码情况下，通过量化等判断进一步压缩，来减少表示视频图像所需的数据量，而解码器侧无法完全重建视频图像，即重建的视频图像的质量比原始视频图像的质量较低或较差。

由于本申请实施例涉及大量神经网络的应用，为了便于理解，下面先对本申请实施例可能涉及的神经网络的相关术语和概念进行介绍。

(1) 神经网络

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

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

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

(2) 深度神经网络

深度神经网络 (deep neural network, DNN)，也称多层神经网络，可以理解为具有多层隐含层的神经网络。按照不同层的位置对 DNN 进行划分，DNN 内部的神经网络可以分为三类：输入层，隐含层，输出层。一般来说第一层是输入层，最后一层是输出层，中间

的层数都是隐含层。层与层之间是全连接的，也就是说，第 i 层的任意一个神经元一定与第 $i+1$ 层的任意一个神经元相连。

虽然 DNN 看起来很复杂，但是就每一层的工作来说，其实并不复杂，简单来说就是如

下线性关系表达式： $\vec{y} = \alpha(W\vec{x} + \vec{b})$ ，其中， \vec{x} 是输入向量， \vec{y} 是输出向量， \vec{b} 是偏移向量，

5 W 是权重矩阵（也称系数）， $\alpha(\cdot)$ 是激活函数。每一层仅仅是对输入向量 \vec{x} 经过如此简单的操作得到输出向量 \vec{y} 。由于 DNN 层数多，系数 W 和偏移向量 \vec{b} 的数量也比较多。这些参数在 DNN 中的定义如下所述：以系数 W 为例：假设在一个三层的 DNN 中，第二层的第 4 个神经元到第三层的第 2 个神经元的线性系数定义为 W_{24}^3 。上标 3 代表系数 W 所在的层数，而下标对应的是输出的第三层索引 2 和输入的第二层索引 4。

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

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

15 量 W 形成的权重矩阵）。

(3) 卷积神经网络

卷积神经网络（convolutional neuron network, CNN）是一种带有卷积结构的深度神经网络。卷积神经网络包含了一个由卷积层和子采样层构成的特征抽取器，该特征抽取器可以看作是滤波器。卷积层是指卷积神经网络中对输入信号进行卷积处理的神经元层。

20 在卷积神经网络的卷积层中，一个神经元可以只与部分邻层神经元连接。一个卷积层中，通常包含若干个特征平面，每个特征平面可以由一些矩形排列的神经单元组成。同一特征平面的神经单元共享权重，这里共享的权重就是卷积核。共享权重可以理解为提取图像信息的方式与位置无关。卷积核可以以随机大小的矩阵的形式初始化，在卷积神经网络的训练过程中卷积核可以通过学习得到合理的权重。另外，共享权重带来的直接好处是减少卷

25 积神经网络各层之间的连接，同时又降低了过拟合的风险。

(4) 熵编码

熵编码用于将熵编码算法或方案（例如，可变长度编码（variable length coding, VLC）方案、上下文自适应 VLC 方案（context adaptive VLC, CALVC）、算术编码方案、

二值化算法、上下文自适应二进制算术编码(context adaptive binary arithmetic coding, CABAC)、基于语法的上下文自适应二进制算术编码(syntax-based context-adaptive binary arithmetic coding, SBAC)、概率区间分割熵(probability interval partitioning entropy, PIPE) 编码或其它熵编码方法或技术) 应用于量化系数、其它语法元素, 得到
5 可以通过输出端以编码比特流等形式输出的编码数据, 使得解码器等可以接收并使用用于解码的参数。可将编码比特流传输到解码器, 或将其保存在存储器中稍后由解码器传输或检索。

在以下译码系统 10 的实施例中, 编码器 20A 和解码器 30A 根据图 1A 至图 15 进行描述。

10 图 1A 为示例性译码系统 10 的示意性框图, 例如可以利用本申请技术的图像(或音频) 译码系统 10 (或简称为译码系统 10)。图像译码系统 10 中的编码器 20A 和解码器 30A 代表可用于根据本申请中描述的各种示例判断各技术的设备等。

如图 1A 所示, 译码系统 10 包括源设备 12, 源设备 12 用于将编码图像(或音频) 等编码码流 21 提供给用于对编码码流 21 进行解码的目的设备 14。

15 源设备 12 包括编码器 20A, 另外即可选地, 图像源 16、预处理器(或预处理单元) 18、通信接口(或通信单元) 26 和概率估计(或概率估计单元) 40。

图像(或音频) 源 16 可包括或可以为任意类型的用于捕获现实世界图像(或音频) 等的图像捕获设备, 和/或任意类型的图像生成设备, 例如用于生成计算机动画图像的计算机图形处理器或任意类型的用于获取和/或提供现实世界图像、计算机生成图像(例如,
20 屏幕内容、虚拟现实(virtual reality, VR) 图像和/或其任意组合(例如增强现实(augmented reality, AR) 图像) 的设备。所述音频或图像源可以为存储上述任意音频或图像的任意类型的内存或存储器。

为了区分预处理器(或预处理单元) 18 判断的处理, 图像或音频(图像或音频数据) 17 也可称为原始图像或音频(原始图像数据或音频数据) 17。

25 预处理器 18 用于接收(原始) 图像(或音频) 数据 17, 并对图像(或音频) 数据 17 进行预处理, 得到预处理图像或音频(或预处理图像或音频数据) 19。例如, 预处理器 18 判断的预处理可包括修剪、颜色格式转换(例如从 RGB 转换为 YCbCr)、调色或去噪。可以理解的是, 预处理单元 18 可以为可选组件。

编码器 20A 包括编码网络 20、熵编码 24, 另外即可选地, 预处理器 22。

30 图像(或音频) 编码网络(或编码网络) 20 用于接收预处理图像(或音频) 数据 19

并提供编码图像（或音频）数据 21。

预处理器 22 用于接收待编码特征数据 21，并对待编码特征数据 21 进行预处理，得到预处理待编码特征数据 23。例如，预处理器 22 判断的预处理可包括修剪、颜色格式转换（例如从 RGB 转换为 YCbCr）、调色或去噪。可以理解的是，预处理单元 22 可以为可
5 选组件。

熵编码 24 用于接收待编码特征数据（或预处理待编码特征数据）23 并根据概率估计 40 提供的概率估计结果 41 生成编码码流 25。

源设备 12 中的通信接口 26 可用于：接收编码码流 25 并通过通信信道 27 向目的设备 14 等另一设备或任何其它设备发送编码码流 25（或其它任意处理后的版本），以便存储
10 或直接重建。

目的设备 14 包括解码器 30A，另外即可选地，可包括通信接口（或通信单元）28、后处理器（或后处理单元）36 和显示设备 38。

目的设备 14 中的通信接口 28 用于直接从源设备 12 或从存储设备等任意其它源设备接收编码码流 25（或其它任意处理后的版本），例如，存储设备为编码码流存储设备，
15 并将编码码流 25 提供给解码器 30A。

通信接口 26 和通信接口 28 可用于通过源设备 12 与目的设备 14 之间的直连通信链路，例如直接有线或无线连接等，或者通过任意类型的网络，例如有线网络、无线网络或其任意组合、任意类型的私网和公网或其任意类型的组合，发送或接收编码码流（或编码码流数据）25。

例如，通信接口 26 可用于将编码码流 25 封装为报文等合适的格式，和/或使用任意类型的传输编码或处理来处理所述编码码流，以便在通信链路或通信网络上进行传输。
20

通信接口 28 与通信接口 26 对应，例如，可用于接收传输数据，并使用任意类型的对应传输解码或处理和/或解封装对传输数据进行处理，得到编码码流 25。

通信接口 26 和通信接口 28 均可配置为如图 1A 中从源设备 12 指向目的设备 14 的对应通信信道 27 的箭头所指示的单向通信接口，或双向通信接口，并且可用于发送和接收消息等，以建立连接，确认并交换与通信链路和/或例如编码后的图像数据传输等数据传输相关的任何其它信息，等等。
25

解码器 30A 包括解码网络 34、熵解码 30，另外即可选地，后处理器 32。

熵解码 30 用于接收编码码流 25 并根据概率估计 40 提供的概率估计结果 42 提供解码
30 特征数据 31。

后处理器 32 用于对解码特征数据 31 进行后处理,得到后处理后的解码特征数据 33。后处理单元 32 判断的后处理可以包括例如颜色格式转换(例如从 YCbCr 转换为 RGB)、调色、修剪或重采样,可以理解的是,后处理单元 32 可以为可选组件。

5 解码网络 34 用于接收解码特征数据 31 或后处理后的解码特征数据 33 并提供重建图像数据 35。

后处理器 36 用于对重建图像数据 35 进行后处理,得到后处理后的重建图像数据 37。后处理单元 36 判断的后处理可以包括例如颜色格式转换(例如从 YCbCr 转换为 RGB)、调色、修剪或重采样,可以理解的是,后处理单元 36 可以为可选组件。

10 显示设备 38 用于接收重建图像数据 35 或后处理后的重建图像数据 37,以向用户或观看者等显示图像。显示设备 38 可以为或包括任意类型的用于表示重建后音频或图像的播放器或显示器,例如,集成或外部显示屏或显示器。例如,显示屏可包括液晶显示器(liquid crystal display, LCD)、有机发光二极管(organic light emitting diode, OLED)显示器、等离子显示器、投影仪、微型 LED 显示器、硅基液晶显示器(liquid crystal on silicon, LCoS)、数字光处理器(digital light processor, DLP)或任意类型的其它显示屏。

20 尽管图 1A 示出了源设备 12 和目的设备 14 作为独立的设备,但设备实施例也可以同时包括源设备 12 和目的设备 14 或同时包括源设备 12 和目的设备 14 的功能,即同时包括源设备 12 或对应功能和目的设备 14 或对应功能。在这些实施例中,源设备 12 或对应功能和目的设备 14 或对应功能可以使用相同硬件和/或软件或通过单独的硬件和/或软件或其任意组合来实现。

根据描述,图 1A 所示的源设备 12 和/或目的设备 14 中的不同单元或功能的存在和(准确)划分可能根据实际设备和应用而有所不同,这对技术人员来说是显而易见的。

25 特征数据编码器 20A(例如图像特征图编码器或音频特征变量编码器)或特征数据解码器 30A(例如图像特征图解码器或音频特征变量解码器)或两者都可通过如图 1B 所示的处理电路实现,例如一个或多个微处理器、数字信号处理器(digital signal processor, DSP)、专用集成电路(application-specific integrated circuit, ASIC)、现场可编程门阵列(field-programmable gate array, FPGA)、离散逻辑、硬件、图像编码专用处理器或其任意组合。特征数据编码器 20A 可以通过处理电路 56 实现,特征数据解码器 30A 可以通过处理电路 56 实现。所述处理电路 56 可用于判断下文论述的各种操作。如果
30 部分技术在软件中实施,则设备可以将软件的指令存储在合适的非瞬时性计算机可读存储

介质中，并且使用一个或多个处理器在硬件中判断指令，从而判断本发明技术。特征数据编码器 20A 和特征数据解码器 30A 中的其中一个可作为组合编解码器(encoder/decoder, CODEC)的一部分集成在单个设备中，如图 1B 所示。

源设备 12 和目的设备 14 可包括各种设备中的任一种，包括任意类型的手持设备或固定设备，例如，笔记本电脑或膝上型电脑、手机、智能手机、平板或平板电脑、相机、台式计算机、机顶盒、电视机、显示设备、数字媒体播放器、视频游戏控制台、视频流设备（例如，内容业务服务器或内容分发服务器）、广播接收设备、广播发射设备，等等，并可以不使用或使用任意类型的操作系统。在一些情况下，源设备 12 和目的设备 14 可配备用于无线通信的组件。因此，源设备 12 和目的设备 14 可以是无线通信设备。

在一些情况下，图 1A 所示的译码系统 10 仅仅是示例性的，本申请提供的技术可适用于图像特征图或音频特征变量编码设置（例如，图像特征图编码或图像特征图解码），这些设置不一定包括编码设备与解码设备之间的任何数据通信。在其它示例中，数据从本地存储器中检索，通过网络发送，等等。图像特征图或音频特征变量编码设备可以对数据进行编码并将数据存储到存储器中，和/或图像特征图或音频特征变量解码设备可以从存储器中检索数据并对数据进行解码。在一些示例中，编码和解码由相互不通信而只是编码数据到存储器和/或从存储器中检索并解码数据的设备来判断。

图 1B 是根据一示例性实施例包含图 1A 的特征数据编码器 20A 和/或图 1B 的特征数据解码器 30A 的译码系统 50 的实例的说明图。译码系统 50 可以包含成像（或产生音频）设备 51、编码器 20A、解码器 30A（和/或藉由处理电路 56 实施的特征数据编/解码器）、天线 52、一个或多个处理器 53、一个或多个内存存储器 54 和/或显示（或音频播放）设备 55。

如图 1B 所示，成像（或产生音频）设备 51、天线 52、处理电路 56、编码器 20A、解码器 30A、处理器 53、内存存储器 54 和/或显示（或音频播放）设备 55 能够互相通信。在不同实例中，译码系统 50 可以只包含编码器 20A 或只包含解码器 30A。

在一些实例中，天线 52 可以用于传输或接收特征数据的经编码比特流。另外，在一些实例中，显示（或音频播放）设备 55 可以用于呈现图像（或音频）数据。处理电路 56 可以包含专用集成电路（application-specific integrated circuit, ASIC）逻辑、图形处理器、通用处理器等。译码系统 50 也可以包含可选的处理器 53，该可选处理器 53 类似地可以包含专用集成电路（application-specific integrated circuit, ASIC）逻辑、图形处理器、音频处理器、通用处理器等。另外，内存存储器 54 可以是任何类型的

存储器，例如易失性存储器（例如，静态随机存取存储器（static random access memory, SRAM）、动态随机存储器（dynamic random access memory, DRAM）等）或非易失性存储器（例如，闪存等）等。在非限制性实例中，内存存储器 54 可以由超速缓存内存实施。在其它实例中，处理电路 56 可以包含存储器（例如，缓存等）用于实施图像缓冲器等。

5 在一些实例中，通过逻辑电路实施的编码器 20A 可以包含（例如，通过处理电路 56 或内存存储器 54 实施的）图像缓冲器和（例如，通过处理电路 56 实施的）图形处理单元。图形处理单元可以通信耦合至图像缓冲器。图形处理单元可以包含通过处理电路 56 实施的编码器 20A。逻辑电路可以用于判断本文所论述的各种操作。

10 在一些实例中，解码器 30A 可以以类似方式通过处理电路 56 实施，以实施参照图 1B 的解码器 30 和/或本文中所描述的任何其它解码器系统或子系统所论述的各种模块。在一些实例中，逻辑电路实施的解码器 30A 可以包含（通过处理电路 56 或内存存储器 54 实施的）图像缓冲器和（例如，通过处理电路 56 实施的）图形处理单元。图形处理单元可以通信耦合至图像缓冲器。图形处理单元可以包含通过处理电路 56 实施的图像解码器 30A。

15 在一些实例中，天线 52 可以用于接收图像数据的经编码比特流。如所论述，经编码比特流可以包含本文所论述的与编码音频或视频帧相关的数据、指示符、索引值、模式选择数据等，例如与编码分割相关的数据。译码系统 50 还可包含耦合至天线 52 并用于解码经编码比特流的解码器 30A。显示（或音频播放）设备 55 用于呈现图像（或音频）。

20 应理解，本申请实施例中对于参考编码器 20A 所描述的实例，解码器 30A 可以用于判断相反过程。关于信令语法元素，解码器 30A 可以用于接收并解析这种语法元素，相应地解码相关图像数据。在一些例子中，编码器 20A 可以将语法元素熵编码成经编码比特流。在此类实例中，解码器 30A 可以解析这种语法元素，并相应地解码相关图像数据。

图 1C 为本发明实施例提供的译码设备 400 的示意图。译码设备 400 适用于实现本文描述的公开实施例。在一个实施例中，译码设备 400 可以是解码器，例如图 1A 中的图像特征图解码器 30A，也可以是编码器，例如图 1A 中的图像特征图编码器 20A。

25 图像译码设备 400 包括：用于接收数据的入端口 410（或输入端口 410）和接收单元（receiver unit, Rx）420；用于处理数据的处理器、逻辑单元或中央处理器（central processing unit, CPU）430；例如，这里的处理器 430 可以是神经网络处理器 430；用于传输数据的发送单元（transmitter unit, Tx）440 和出端口 450（或输出端口 450）；用于存储数据的存储器 460。图像（或音频）译码设备 400 还可包括耦合到入端口 410、
30 接收单元 420、发送单元 440 和出端口 450 的光电（optical-to-electrical, OE）组件

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

处理器 430 通过硬件和软件实现。处理器 430 可实现为一个或多个处理器芯片、核(例如, 多核处理器)、FPGA、ASIC 和 DSP。处理器 430 与入端口 410、接收单元 420、发送单元 440、出端口 450 和存储器 460 通信。处理器 430 包括译码模块 470 (例如, 基于神经网络 NN 的译码模块 470)。译码模块 470 实施上文所公开的实施例。例如, 译码模块 470 判断、处理、准备或提供各种编码操作。因此, 通过译码模块 470 为译码设备 400 的功能提供了实质性的改进, 并且影响了译码设备 400 到不同状态的切换。或者, 以存储在存储器 460 中并由处理器 430 判断的指令来实现译码模块 470。

存储器 460 包括一个或多个磁盘、磁带机和固态硬盘, 可以用作溢出数据存储设备, 用于在选择判断程序时存储此类程序, 并且存储在程序判断过程中读取的指令和数据。存储器 460 可以是易失性和/或非易失性的, 可以是只读存储器 (read-only memory, ROM)、随机存取存储器 (random access memory, RAM)、三态内容寻址存储器 (ternary content-addressable memory, TCAM) 和/或静态随机存取存储器 (static random-access memory, SRAM)。

图 1D 为示例性实施例提供的装置 500 的简化框图, 装置 500 可用作图 1A 中的源设备 12 和目的设备 14 中的任一个或两个。

装置 500 中的处理器 502 可以是中央处理器。或者, 处理器 502 可以是现有的或今后将研发出的能够操控或处理信息的任何其它类型设备或多个设备。虽然可以使用如图所示的处理器 502 等单个处理器来实施已公开的实现方式, 但使用一个以上的处理器速度更快和效率更高。

在一种实现方式中, 装置 500 中的存储器 504 可以是只读存储器 (ROM) 设备或随机存取存储器 (RAM) 设备。任何其它合适类型的存储设备都可以用作存储器 504。存储器 504 可以包括处理器 502 通过总线 512 访问的代码和数据 506。存储器 504 还可包括操作系统 508 和应用程序 510, 应用程序 510 包括允许处理器 502 判断本文所述方法的至少一个程序。例如, 应用程序 510 可以包括应用 1 至 N, 还包括判断本文所述方法的图像译码应用。

装置 500 还可以包括一个或多个输出设备, 例如显示器 518。在一个示例中, 显示器 518 可以是将显示器与可用于感测触摸输入的触敏元件组合的触敏显示器。显示器 518 可以通过总线 512 耦合到处理器 502。

虽然装置 500 中的总线 512 在本文中描述为单个总线, 但是总线 512 可以包括多个总

线。此外，辅助储存器可以直接耦合到装置 500 的其它组件或通过网络访问，并且可以包括存储卡等单个集成单元或多个存储卡等多个单元。因此，装置 500 可以具有各种各样的配置。

图 2A 示出了一种可能的图像特征图或音频特征变量编解码场景下的系统架构 1800，
5 包括：

采集设备 1801：视频采集设备完成原始视频（或音频）采集；

采集前处理 1802：原始视频（或音频）采集经过一系列的前处理得到视频（或音频）
数据；

10 编码 1803：视频（或音频）编码用于降低编码冗余，降低图像特征图或音频特征变
量压缩过程中的数据传输量；

发送 1804：将编码后得到的压缩编码码流数据通过发送模块进行数据发送；

接收 1805：压缩编码码流数据经过网络传输被接收模块所接收；

码流解码 1806：对码流数据进行码流解码；

渲染显示（或播放）1807：对解码后的数据进行渲染显示（或播放）；

15 图 2B 示出了一种可能的图像特征图（或音频特征变量）面向机器任务场景下的系统
架构 1900，包括：

特征提取 1901：对图像（或音频）源进行特征提取；

边信息提取 1902：对特征提取数据进行边信息提取；

20 概率估计 1903：边信息作为概率估计的输入，对特征图（或特征变量）进行概率估
计以得到概率估计结果；

编码 1904：结合概率估计结果对特征提取数据执行熵编码以得到码流；

可选的，在进行编码以前对特征提取数据执行量化或取整操作，然后对量化或取整后
的特征提取数据进行编码。

可选的，对边信息执行熵编码，使得码流中包括了边信息数据。

25 解码 1905：结合概率估计结果对码流执行熵解码得到图像特征图（或音频特征变量）；

其中可选的，如码流中包括了边信息编码数据，对边信息编码数据执行熵解码，并结
合解码的边信息数据作为概率估计的输入以得到概率估计结果。

需要说明的是，当只有边信息作为概率估计的输入时，可以并行输出各特征元素的概
率估计结果；当概率估计的输入包括有上下文信息时，需要串行输出各特征元素概率估计
30 结果。其中所述边信息为图像特征图或音频特征变量输入神经网络进一步提取得到的特征

信息，所述边信息包含的特征元素的个数比图像特征图或音频特征变量的特征元素少。可选地，可以将图像特征图或音频特征变量的边信息编入码流。

机器视觉任务 1906：对解码特征图（或特征变量）执行机器视觉（或听觉）任务。

5 具体的，将解码特征数据输入值机器视觉（或听觉）任务网络，网路输出为视觉（或听觉）任务相关如分类、目标识别、语义分割等任务的一维、二维或多维数据。

在一种可能实现上，在系统架构 1900 实现过程中，特征提取、编码过程在终端上实现，对解码以及执行机器视觉任务在云端上实现。

编码器 20A 可用于通过输入端 202 等接收图像（或图像数据）或音频（或音频数据）
10 17。接收图像、图像数据、音频、音频数据也可以是预处理后的图像（或预处理后的图像数据）或音频（或预处理后的音频数据）19。为简单起见，以下描述使用图像（或音频）17。图像（或音频）17 也可称为当前图像或待编码的图像（尤其是在视频编码中将当前图像与其它图像区分开时，其它图像例如同一视频序列，即也包括当前图像的视频序列中的之前编码后图像和/或解码后图像）或当前音频或待编码的音频。

15 （数字）图像为或可以视为具有强度值的像素点组成的二维阵列或矩阵。阵列中的像素点也可以称为像素（pixel 或 pel）（图像元素的简称）。阵列或图像在水平方向和垂直方向（或轴线）上的像素点数量决定了图像的大小和/或分辨率。为了表示颜色，通常采用三个颜色分量，即图像可以表示为或包括三个像素点阵列。在 RGB 格式或颜色空间中，图像包括对应的红色、绿色和蓝色像素点阵列。同样，每个像素可以以亮度/色度格式或
20 颜色空间表示，例如 YCbCr，包括 Y 指示的亮度分量（有时也用 L 表示）以及 Cb、Cr 表示的两个色度分量。亮度（luma）分量 Y 表示亮度或灰度水平强度（例如，在灰度等级图像中两者相同），而两个色度（chrominance，简称为 chroma）分量 Cb 和 Cr 表示色度或颜色信息分量。相应地，YCbCr 格式的图像包括亮度像素点值（Y）的亮度像素点阵列和色度值（Cb 和 Cr）的两个色度像素点阵列。RGB 格式的图像可以转换或变换为 YCbCr 格
25 式，反之亦然，该过程也称为颜色变换或转换。如果图像是黑白的，则该图像可以只包括亮度像素点阵列。相应地，图像可以为例如单色格式的亮度像素点阵列或 4:2:0、4:2:2 和 4:4:4 彩色格式的亮度像素点阵列和两个相应的色度像素点阵列。图像编码器 20A 对图像的色彩空间不做限制。

在一个可能性中，编码器 20A 的实施例可包括图像（或音频）分割单元（图 1A 或图
30 1B 中未示出），用于将图像（或音频）17 分割成多个（通常不重叠）图像块 203 或音频

段。这些图像块在 H. 265/HEVC 和 VVC 标准中也可以称为根块、宏块 (H. 264/AVC) 或编码树块 (Coding Tree Block, CTB), 或编码树单元 (Coding Tree Unit, CTU)。分割单元可用于对视频序列中的所有图像使用相同的块大小和使用限定块大小的对应网格, 或在图像或图像子集或图像组之间改变块大小, 并将每个图像分割成对应块。

5 在另一个可能性中, 编码器可用于直接接收图像 17 的块 203, 例如, 组成所述图像 17 的一个、几个或所有块。图像块 203 也可以称为当前图像块或待编码图像块。

与图像 17 一样, 图像块 203 同样是或可认为是具有强度值 (像素点值) 的像素点组成的二维阵列或矩阵, 但是图像块 203 的比图像 17 的小。换句话说, 块 203 可包括一个像素点阵列 (例如, 单色图像 17 情况下的亮度阵列或彩色图像情况下的亮度阵列或色度
10 阵列) 或三个像素点阵列 (例如, 彩色图像 17 情况下的一个亮度阵列和两个色度阵列) 或根据所采用的颜色格式的任何其它数量和/或类型的阵列。块 203 的水平方向和垂直方向 (或轴线) 上的像素点数量限定了块 203 的大小。相应地, 块可以为 $M \times N$ (M 列 \times N 行) 个像素点阵列, 或 $M \times N$ 个变换系数阵列等。

在另一个可能性中, 图 1A-1B 或图 3A-3D 所示的编码器 20A 用于逐块对图像 17 进行
15 编码。

在另一个可能性中, 图 1A-1B 或图 3A-3D 所示的编码器 20A 用于对图像 17 进行编码。

在另一个可能性中, 图 1A-1B 或图 3A-3D 所示的编码器 20A 还可以用于使用片 (也称为视频片) 分割编码图像, 其中图像可以使用一个或多个片 (通常为不重叠的) 进行分割或编码。每个片可包括一个或多个块 (例如, 编码树单元 CTU) 或一个或多个块组 (例如
20 H. 265/HEVC/VVC 标准中的编码区块 (tile) 和 VVC 标准中的子图像 (subpicture))。

在另一个可能性中, 图 1A-1B 或图 3A-3D 所示的编码器 20A 还可以用于使用片/编码区块组 (也称为视频编码区块组) 和/或编码区块 (也称为视频编码区块) 对图像进行分割和/或编码, 其中图像可以使用一个或多个片/编码区块组 (通常为不重叠的) 进行分割或编码, 每个片/编码区块组可包括一个或多个块 (例如 CTU) 或一个或多个编码区块等,
25 其中每个编码区块可以为矩形等形状, 可包括一个或多个完整或部分块 (例如 CTU)。

编码网络 20

编码网络 20 用于通过编码网络根据输入数据来得到图像特征图或音频特征变量。

在一个可能性中, 编码网络 20 如图 4A 所示, 编码网络 20 包含多个网络层, 任意一网络层可以为卷积层、归一化层、非线性激活层等。

30 在一个可能性中, 编码网络 20 输入为至少一张待编码图像或至少一个待编码图像块。

待编码图像为可以为原始图像，有损图像或者为残差图像。

在一个可能性中，编码网络 20 中的编码网络的网络结构示例如图 4B 所示，可见示例中编码网络包含了 5 个网络层，具体包括了三个卷积层以及两个非线性激活层。

取整 24

- 5 取整用于通过例如标量量化或矢量量化对图像特征图或音频特征变量进行取整，得到取整后的图像特征图或音频特征变量。

在一个可能性中，编码器 20A 可用于输出取整参数 (quantization parameter, QP)，例如，直接输出或由编码决策实现单元进行编码或压缩后输出，例如使得解码器 30A 可接收并使用量化参数进行解码。

- 10 在一个可能性中，输出特征图或特征音频特征变量在进行取整前进行预处理，预处理可包括修剪、颜色格式转换 (例如从 RGB 转换为 YCbCr)、调色或去噪等。

概率估计 40

概率估计根据的输入特征图或特征变量信息以得到图像特征图或音频特征变量的概率估计结果。

- 15 概率估计用于对取整后的图像特征图或音频特征变量进行概率估计。

概率估计可以为概率估计网络，概率估计网络为卷积网络，卷积网络中包括了卷积层和非线性激活层。以图 4B 为例，概率估计网络包含了 5 个网络层，具体包括了三个卷积层以及两个非线性激活层。概率估计可以采用非网络的传统概率估计方法实现。概率估计方法包括且不限于等最大似然估计、最大后验估计、极大似然估计等统计方法。

- 20 编码决策实现 26

如图 5 所示，编码决策实现包括了编码元素判断以及熵编码。所述图像特征图或音频特征变量为编码网络所输出的一维、二维或多维数据，其中每个数据均为特征元素。编码元素判断 261

- 25 编码元素判断是根据概率估计的概率估计结果信息来对图像特征图或音频特征变量中的每一个特征元素进行判断并根据判断结果决定具体对哪些特征元素执行熵编码。

图像特征图或音频特征变量的第 P 个特征元素的元素判断过程完成后，开始图像特征图的第 P+1 个特征元素的元素判断过程，其中 P 为正整数且 P 小于 M。

熵编码 262

- 30 熵编码可以采用各种公开的熵编码算法进行编码，譬如采用方案例如，可变长度编码 (variable length coding, VLC) 方案、上下文自适应 VLC 方案 (context adaptive VLC,

CAVLC)、熵编码方案、二值化算法、上下文自适应二进制熵编码(context adaptive binary arithmetic coding, CABAC)、基于语法的上下文自适应二进制熵编码(syntax-based context-adaptive binary arithmetic coding, SBAC)、概率区间分割熵(probability interval partitioning entropy, PIPE)编码或其它熵编码方法或技术。得到可以通过
5 输出端 212 以编码比特流 25 等形式输出的编码图像数据 25, 使得解码器 30A 等可以接收并使用用于解码的参数。可将编码比特流 25 传输到解码器 30A, 或将其保存在存储器中稍后由解码器 30A 传输或检索。

另一种可能性中, 熵编码可以采用熵编码网络进行编码, 譬如采用卷积网络实现。

10 在一个可能性中, 熵编码由于不知道取整特征图的真实字符概率, 可以统计这些或相关的信息添加至熵编码中, 把这些信息传到解码端。

联合网络 44

联合网络是根据输入边信息得到图像特征图或音频特征变量的概率估计结果和决策信息。联合网络为多层网络, 联合网络可以为卷积网络, 卷积网络中包括了卷积层和非线性激活层。联合网络任意一网络层可以为卷积层、归一化层、非线性激活层等。

15 其中所述决策信息可以为一维、二维或多维数据, 所述决策信息尺寸可以与图像特征图尺寸一致。

所述决策信息可以在联合网络中任意一网络层后输出。

所述概率估计结果可以在联合网络中任意一网络层后输出。

20 图 6 为联合网络的网络结构输出示例, 网络结构包括 4 个网络层, 其中决策信息在第四网络层后进行输出, 概率估计结果在第二网络层后输出。

生成网络 46

生成网络是根据输入概率估计结果得到图像特征图中各特征元素的决策信息。生成网络为多层网络, 生成网络可以为卷积网络, 卷积网络中包括了卷积层和非线性激活层。生成网络任意一网络层可以为卷积层、归一化层、非线性激活层等。

25 所述决策信息可以在生成网络中任意一网络层后输出。所述决策信息可以为一维、二维或多维数据。

图 7 为生成网络的网络结构输出决策信息示例, 网络结构包括 4 个网络层。

解码决策实现 30

30 如图 8 所示, 解码决策实现包括了元素判断以及熵解码。所述图像特征图或音频特征变量为解码决策实现所输出的一维、二维或多维数据, 其中每个数据均为特征元素。

解码元素判断 301

解码元素判断根据概率估计的概率估计结果来对图像特征图或音频特征变量中的每一个特征元素进行判断并根据判断结果决定具体对哪些特征元素执行熵解码。解码元素判断对图像特征图或音频特征变量中的每一个特征元素进行判断并根据判断结果决定具体对哪些特征元素执行熵解码，可以看作是编码元素判断对图像特征图中的每一个特征元素进行判断并根据判断结果决定具体对哪些特征元素执行熵编码的逆过程。

熵解码 302

熵解码可以采用各种公开的熵解码算法进行编码，譬如采用方案例如，可变长度编码 (variable length coding, VLC) 方案、上下文自适应 VLC 方案 (context adaptive VLC, CAVLC)、熵解码方案、二值化算法、上下文自适应二进制熵解码 (context adaptive binary arithmetic coding, CABAC)、基于语法的上下文自适应二进制熵解码 (syntax-based context-adaptive binary arithmetic coding, SBAC)、概率区间分割熵 (probability interval partitioning entropy, PIPE) 编码或其它熵编码方法或技术。得到可以通过输出端 212 以编码比特流 25 等形式输出的编码图像 (或音频) 数据 25，使得解码器 30A 等可以接收并使用用于解码的参数。可将编码比特流 25 传输到解码器 30A，或将其保存在存储器中稍后由解码器 30A 传输或检索。

另一种可能性中，熵解码可以采用熵解码网络进行解码，譬如采用卷积网络实现。

解码网络 34

解码网络用于将解码图像特征图或音频特征变量 31 或后处理解码图像特征图或音频特征变量 33 通过解码网络 34 以在像素域中得到重建图像 (或音频) 数据 35 或面向机器任务数据。

解码网络包含多个网络层，任意一网络层可以为卷积层、归一化层、非线性激活层等。解码网络单元 306 中可以存在包括叠加 (concat)、相加、相减等操作。

在一个可能性中，解码网络中各网络层结构可以互为相同或者不同。

解码网络的结构示例如图 9 所示，可见示例中解码网络包含了 5 个网络层，具体包括了一个归一化层、两个卷积层以及两个非线性激活层。

解码网络输出重建图像 (或音频)，或者输出得到面向机器任务数据。具体的，所述解码网络可以包括目标识别网络，分类网络或者语义分割网络。

应理解，在编码器 20A 和解码器 30A 中，可以对当前步骤的处理结果进一步处理，然后输出到下一步骤。例如，在编码器单元或解码器单元之后，可以对编码器单元或解码器

单元的处理结果进行进一步的运算或处理，例如裁剪 (clip) 或移位 (shift) 运算或滤波处理。

5 基于上文的描述，下面给出本申请实施例提供的一些图像特征图或音频特征变量的编解码方法。对于下文描述的各方法实施例，为了方便起见，将其都表述为一系列的动作步骤的组合，但是本领域技术人员应该知悉，本申请技术方案的具体实现并不受所描述的一系列的动作步骤的顺序的限制。

下面结合附图，对本申请的流程进行详细的描述。需要说明的是，流程图中的编码端过程具体可以由上述的编码器 20A 来执行，流程图中的解码端过程具体可以由上述的解码器 10 30A 来执行。

实施例一至实施例五中，第一特征元素或第二特征元素即为当前待编码特征元素或为当前待解码特征元素或，比如 $\hat{y}[x][y][i]$ 。决策图也可以称为决策图 map。决策图优选的为二元图，二元图也可以称为二元图 map。

在本申请实施例一中，图 10A 示出了具体实现流程 1400，运行步骤如下：

15 编码端：

步骤 1401：获取图像的特征图

本步骤具体由图 3A 中的编码网络 204 来实现，具体可以参照上述对编码网络 20 的描述。将图像分别输入特征提取模块输出图像的特征图 y ，特征图 y 可以是维度为 $w \times h \times c$ 的三维数据。具体的，特征提取模块可以使用现有的神经网络来实现，在此不做限定。本步骤为现有技术。

特征量化模块对特征图 y 中的每个特征值进行量化，将浮点数的特征值进行四舍五入得到整数特征值，得到量化后的特征图 \hat{y} 。可以参照上述实施例对取整 24 的描述。

步骤 1402：对特征图 \hat{y} 进行概率估计得到各特征元素的概率估计结果，即特征图 \hat{y} 中的每个特征元素 $\hat{y}[x][y][i]$ 的概率分布：

25 其中，参数 x, y, i 为正整数，坐标 (x, y, i) 表示当前待编码特征元素的位置，具体的，坐标 (x, y, i) 表示当前待编码特征元素在当前三维特征图中相对于左上顶点的特征元素的位置。本步骤具体由图 3A 中的概率估计 210 来实现，具体可以参照上述对概率估计 40 的描述。具体可以使用概率分布模型来获得概率分布，例如使用单高斯模型 (Gaussian single model, GSM) 或者混合高斯模型 (Gaussian mixture model, GMM)

30 建模，首先将边信息 \hat{z} 和上下文信息输入概率估计网络，对特征图 \hat{y} 中的每个特征元素

$\hat{y}[x][y][i]$ 进行概率估计得到各特征元素 $\hat{y}[x][y][i]$ 的概率分布。概率估计网络可以使用基于深度学习网络，例如循环神经网络（Recurrent Neural Network, RNN）和卷积神经网络（Convolutional Neural Network, PixelCNN）等，在此不做限定。将模型参数代入概率分布模型中，得到概率分布。

5 步骤 1403：对特征图 \hat{y} 执行熵编码得到压缩码流，并生成压缩码流。

本步骤具体由图 3A 中的编码决策实现 208 来实现，具体可以参照上述对编码决策实现 26 的描述。根据所述概率分布，得到当前待编码特征元素 $\hat{y}[x][y][i]$ 取值为 k 的概率 P ，当当前待编码特征元素 $\hat{y}[x][y][i]$ 的概率估计结果 P 不满足预设条件：当 P 大于（或者等于）第一阈值 T_0 时，跳过当前待编码特征元素执行熵编码过程；否则，当当前待编码特征元素的概率估计结果 P 满足预设条件：当 P 小于第一阈值 T_0 时，对当前待编码特征元素执行熵编码写入码流。其中， k 可为任意整数，例如 0, 1, -1, 2, 3 等。所述第一阈值 T_0 为满足 $0 < T_0 < 1$ 中的任一数，例如取值为 0.99, 0.98, 0.97, 0.95 等。（可以认为每个特征元素的阈值都相同）。

步骤 1404、编码器发送或存储压缩码流。

15 解码端：

步骤 1411：获取解码图像特征图的码流

步骤 1412：根据码流进行概率估计得到各特征元素的概率估计结果

本步骤具体由图 10B 中的的概率估计 302 来实现，具体可以参照上述对概率估计 40 的描述。对待解码特征图 \hat{y} 中的每个特征元素 $\hat{y}[x][y][i]$ 进行概率估计，得到待解码特征元素 $\hat{y}[x][y][i]$ 的概率分布。待解码特征图 \hat{y} 中包括多个特征元素，所述多个特征元素包括当前待解码特征元素。

解码端使用的概率估计网络结构图与本实施例的编码端概率估计网络结构相同。

步骤 1413：对待解码特征图 \hat{y} 执行熵解码

本步骤具体由图 10B 中的解码决策实现 304 来实现，具体可以参照上述对解码决策实现 30 的描述。根据当前待解码特征元素的概率分布，得到当前待解码特征元素取值为 k 的概率 P ，即当前待解码特征元素的概率估计结果 P ，当概率估计结果 P 不满足预设条件： P 大于所述第一阈值 T_0 时，不需要对当前待解码特征元素执行熵解码，将当前待解码特征元素数值设置为 k ，否则，当当前待解码特征元素满足预设条件： P 小于或者等于所述第一阈值 T_0 时，对码流执行熵解码，得到当前待解码特征元素的值。

30 其中所述第一阈值 T_0 可以通过解析码流，从码流中获取索引号，解码端使用编码端

相同的方式来构建阈值候选列表，然后根据阈值候选列表中预置和索引号的对应关系得到对应的阈值。其中，从码流中获取索引号即从序列头、图像头、Slice/条带或 SEI 中获取索引号。

5 或者可以直接解析码流，从码流中获取阈值，具体地，从序列头、图像头、Slice/条带或 SEI 中获取阈值。

或者根据与解码的阈值约定策略，直接设置固定阈值。

步骤 1414：对解码后的特征图 \hat{y} 进行重建，或者输入面向机器视觉任务模块执行相应的机器任务。本步骤具体可以由图 10B 中的解码网络 306 来实现，具体可以参照上述对解码网络 34 的描述。

10 情况一：将熵解码后的特征图 \hat{y} 输入图像重建模块，神经网络输出重建图。所述神经网络可以采用任一结构，例如全连接网络、卷积神经网络、循环神经网络等。所述神经网络可以采用多层的结构深度神经网络结构来达到更好的估计效果。

情况二：将熵解码后的特征图 \hat{y} 输入面向机器视觉任务模块执行相应的机器任务。例如完成物体分类、识别、分割等机器视觉任务。

15 以上解码端的 k 值与编码端的 k 值相对应设置。

图 11A 示出了本申请实施例二的具体实现流程 1500，运行步骤如下：

需要说明的是，本实施例的方法一至方法六中：概率估计结果包括第一参数和第二参数；当所述概率分布为高斯分布时，第一参数为均值 μ ，第二参数为方差 σ ；当所述概率分布为拉普拉斯分布时，第一参数为位置参数 μ ，第二参数为尺度参数 b 。

编码端：

步骤 1501：获取图像的特征图

本步骤具体由图 3B 中的编码网络 204 来实现，具体可以参照上述对编码网络 20 的描述。将图像分别输入特征提取模块输出图像的特征图 y ，特征图 y 可以是维度为 $w \times h \times c$ 的三维数据。。具体的，特征提取模块可以使用现有的神经网络来实现，在此不做限定。本步骤为现有技术。

特征量化模块对特征图 y 中的每个特征值进行量化，将浮点数的特征值进行四舍五入得到整数特征值，得到量化后的特征图 \hat{y} 。

步骤 1502：图像的特征图 \hat{y} 输入边信息提取模块，输出边信息 \hat{z} 。

30 本步骤具体由图 3B 中的边信息提取单元 214 来实现。其中，边信息提取模块可以使

用图 12 所示的网络来实现, 边信息 z 可以理解为对特征图 \hat{y} 进行进一步提取得到的特征图 \hat{z} , \hat{z} 所含包含的特征元素的个数比特征图 \hat{y} 少。

需要说明的是, 可以在本步骤中, 对边信息 z 执行熵编码并写入码流, 也可以在后续的步骤 1504 中对边信息 z 执行熵编码并写入码流, 在此不做限定。

5 步骤 1503: 对特征图 \hat{y} 进行概率估计得到各特征元素的概率估计结果。

本步骤具体由图 3B 中的概率估计 210 来实现, 具体可以参照上述对概率估计 40 的描述。可以使用概率分布模型来获得概率估计结果及概率分布。其中, 概率分布模型可以为: 单高斯模型 (Gaussian single model, GSM) 或者非对称高斯模型或者混合高斯模型 (Gaussian mixture model, GMM) 或者拉普拉斯分布模型 (Laplace distribution)。

10 当概率分布模型为高斯模型时 (单高斯模型或者非对称高斯模型或者混合高斯模型), 首先将边信息 z 或者上下文信息输入概率估计网络, 对特征图 \hat{y} 中的每个特征元素 $\hat{y}[x][y][i]$ 进行概率估计得到均值参数 μ 和方差 σ 的值。进一步地, 将所述均值参数 μ 和方差 σ 输入所使用的概率分布模型中, 得到概率分布。此时概率估计结果为均值参数 μ 和方差 σ 。

15 当概率分布模型拉普拉斯分布模型时, 首先将边信息 z 或者上下文信息输入概率估计网络, 对特征图 \hat{y} 中的每个特征元素 $\hat{y}[x][y][i]$ 进行概率估计得到位置参数 μ 和尺度参数 b 的值。进一步地, 将所述位置参数 μ 和尺度参数 b 输入所使用的概率分布模型中, 得到概率分布。此时概率估计结果为位置参数 μ 和尺度参数 b 。

还可以将边信息 z 和/或上下文信息输入概率估计网络, 对待编码特征图 \hat{y} 中的每个特征元素 $\hat{y}[x][y][i]$ 进行概率估计得到当前待编码特征元素 $\hat{y}[x][y][i]$ 的概率分布。根据所述
20 概率分布, 得到当前待编码特征元素 $\hat{y}[x][y][i]$ 取值为 m 的概率 P 。此时概率估计结果为当前待编码特征元素 $\hat{y}[x][y][i]$ 取值为 m 的概率 P 。

其中, 概率估计网络可以使用基于深度学习网络, 例如循环神经网络 (Recurrent Neural Network, RNN) 和卷积神经网络 (Convolutional Neural Network, PixelCNN) 等, 在此不做限定。

25 步骤 1504: 根据概率估计结果判断当前待编码特征元素 $\hat{y}[x][y][i]$ 是否需要执行熵编码, 并根据判断结果执行熵编码写入压缩码流 (编码码流) 或者不执行熵编码。仅当判断出需要对所述当前待编码第一特征元素执行熵编码时, 对所述当前待编码特征元素执行熵编码。

本步骤具体由图 3B 中的编码决策实现 208 来实现, 具体可以参照上述编码决策实现
30 26 的描述。根据概率估计结果判断当前待编码特征元素 $\hat{y}[x][y][i]$ 是否需要执行熵编码可

以使用以下方法中的一项或者多项。其中，参数 x, y, i 为正整数，坐标 (x, y, i) 表示当前待编码特征元素的位置，具体的，坐标 (x, y, i) 表示当前待编码特征元素在当前三维特征图中相对于左上顶点的特征元素的位置。

方法一：当所述概率分布模型为高斯分布时，根据所述第一特元素的概率估计结果判断是否对所述当前待编码特征元素执行熵编码，当当前待编码特征元素的高斯分布的均值参数 μ 和方差 σ 的值不满足预设条件：当均值 μ 与 k 的差的绝对值小于第二阈值 $T1$ 且方差 σ 小于第三阈值 $T2$ 时，不需要对当前待编码特征元素 $\hat{y}[x][y][i]$ 执行熵编码过程，否则，当满足预设条件：当均值 μ 与 k 的差的绝对值大于或者等于第二阈值 $T1$ 或方差 σ 小于第三阈值 $T2$ 时，对当前待编码特征元素 $\hat{y}[x][y][i]$ 执行熵编码写入码流。其中， k 为任意的整数，例如 $0, 1, -1, 2, 3$ 等。 $T2$ 取值为满足 $0 < T2 < 1$ 中的任一数，例如取值为 $0.2, 0.3, 0.4$ 等。 $T1$ 是大于或等于 0 小于 1 的数，例如 $0.01, 0.02, 0.001, 0.002$ 。

特别地， k 取值为 0 时为最优值，可以直接判断当高斯分布的均值参数 μ 绝对值小于 $T1$ 且高斯分布的方差 σ 小于 $T2$ 时，则跳过对当前待编码特征元素 $\hat{y}[x][y][i]$ 执行熵编码过程，否则，对当前待编码特征元素 $\hat{y}[x][y][i]$ 执行熵编码写入码流。其中， $T2$ 的取值为满足 $0 < T2 < 1$ 中的任一数，例如取值为 $0.2, 0.3, 0.4$ 等。 $T1$ 是大于或等于 0 小于 1 的数，例如 $0.01, 0.02, 0.001, 0.002$ 。

方法二：当所述概率分布为高斯分布时，根据所述概率估计结果，得到当前待编码特征元素 $\hat{y}[x][y][i]$ 的高斯分布的均值参数 μ 和方差 σ 的值，当均值 μ 、方差 σ 与 k 的关系满足 $abs(\mu - k) + \sigma < T3$ 时（不满足预设条件），跳过对当前待编码特征元素 $\hat{y}[x][y][i]$ 执行熵编码过程，其中， $abs(\mu - k)$ 表示计算均值 μ 与 k 的差的绝对值；否则，当当前待编码特征元素的概率估计结果满足 $abs(\mu - k) + \sigma \geq T3$ 时（预设条件），对当前待编码特征元素 $\hat{y}[x][y][i]$ 执行熵编码写入码流。其中， k 为任意的整数，例如 $0, 1, -1, -2, 3$ 等。其中，第四阈值 $T3$ 是大于或等于 0 小于 1 的数，例如取值为 $0.2, 0.3, 0.4$ 等。

方法三：当所述概率分布为拉普拉斯分布时，根据所述概率估计结果，得到当前待编码特征元素 $\hat{y}[x][y][i]$ 的拉普拉斯分布的位置参数 μ 、尺度参数 b 的值。当位置参数 μ 、尺度参数 b 与 k 的关系满足 $abs(\mu - k) + \sigma < T4$ （不满足预设条件）时，跳过对当前待编码特征元素 $\hat{y}[x][y][i]$ 执行熵编码过程，其中， $abs(\mu - k)$ 表示计算位置参数 μ 与 k 的差的绝对值；否则，当当前待编码特征元素的概率估计结果满足 $abs(\mu - k) + \sigma \geq T4$ （预设条件），对当前待编码特征元素 $\hat{y}[x][y][i]$ 执行熵编码写入码流。其中， k 为任意的整数，例如 $0, 1, -1, -2, 3$ 等。第四阈值 $T4$ 是大于或等于 0 小于 0.5 的数，例如取值为 $0.05, 0.09, 0.17$

方法一：以阈值 T1 为例，取 T1 取值范围内的任意一个取值作为阈值 T1，将阈值 T1 写入码流。具体地，将所述阈值写入码流，可将其保存在序列头、图像头、Slice/条带或 SEI 中传送到解码端，还可以使用其他方法，在此不做限定。其余阈值 T0, T2, T3, T4, T5 和 T6 也可以使用类似方法。

5 方法二：编码端采用与解码端约定的固定阈值，无需写入码流，无需传输到解码端。例如，以阈值 T1 为例，直接取 T1 取值范围内任一值作为 T1 的取值。其余阈值 T0, T2, T3, T4, T5 和 T6 也可以使用类似方法。

方法三：构建阈值候选列表，将在 T1 取值范围内最有可能的取值放入阈值候选列表中，每个阈值对应一个阈值索引号，确定一个最优的阈值，将最优阈值作为 T1 的值，并将最优阈值的索引号作为 T1 的阈值索引号，将 T1 的阈值索引号写入码流。具体地，将所述阈值写入码流，可将其保存在序列头、图像头、Slice/条带或 SEI 中传送到解码端，还可以使用其他方法，在此不做限定。其余阈值 T0, T2, T3, T4, T5 和 T6 也可以使用类似方法。

步骤 1505：编码器发送或存储压缩码流。

15

解码端：

步骤 1511：获取待解码图像特征图的码流

步骤 1512：获取各特征元素的概率估计结果

本步骤具体由图 11A 中的概率估计单元 302 来实现，具体可以参照上述对概率估计 20 40 的描述。对边信息 \hat{z} 执行熵解码得到边信息 \hat{z} ，结合边信息 \hat{z} 对待解码特征图 \hat{y} 中的每个特征元素 $\hat{y}[x][y][i]$ 进行概率估计，得到当前待解码特征元素 $\hat{y}[x][y][i]$ 的概率估计结果。

需要说明的是，解码端使用的概率估计方法与本实施例编码端的概率估计方法对应相同，概率估计网络结构图与本实施例的编码端概率估计网络结构相同，在此不做赘述。

步骤 1513：本步骤具体由图 11A 中的解码决策实现 304 来实现，具体可以参照上述 25 对解码决策实现 30 的描述。根据概率估计结果判断当前待解码特征元素 $\hat{y}[x][y][i]$ 是否需要执行熵解码，并根据判断结果执行或者不执行熵解码，得到解码后的特征图 \hat{y} 。

根据概率估计结果判断当前待解码特征元素 $\hat{y}[x][y][i]$ 是否需要执行熵解码可以使用以下方法中的一项或者多项。

方法一：当所述概率分布模型为高斯分布时，根据所述概率估计结果，得到当前待解 30 码特征元素 $\hat{y}[x][y][i]$ 的均值参数 μ 和方差 σ 的值，当均值 μ 与 k 的差的绝对值小于第二阈

值 $T1$ 且方差 σ 小于第三阈值 $T2$ 时 (不满足预设条件), 将当前待解码特征元素 $\hat{y}[x][y][i]$ 的数值设置为 k , 跳过对当前待解码特征元素 $\hat{y}[x][y][i]$ 执行熵解码过程; 否则, 当均值 μ 与 k 的差的绝对值小于第二阈值 $T1$ 或方差 σ 大于或等于第三阈值 $T2$ 时 (预设条件), 对当前待解码特征元素 $\hat{y}[x][y][i]$ 执行熵解码, 得到当前待解码特征元素 $\hat{y}[x][y][i]$ 的值。

5 特别地, k 取值为 0 时为最优值, 可以直接判断当高斯分布的均值参数 μ 绝对值小于 $T1$ 且高斯分布的方差 σ 小于 $T2$ 时, 将当前待解码特征元素 $\hat{y}[x][y][i]$ 的数值设置为 k , 则跳过对当前待解码特征元素 $\hat{y}[x][y][i]$ 执行熵解码过程, 否则, 对当前待解码特征元素 $\hat{y}[x][y][i]$ 执行熵解码, 得到当前待解码特征元素 $\hat{y}[x][y][i]$ 的值。

10 方法二: 当所述概率分布为高斯分布时, 根据所述概率估计结果, 得到当前待解码特征元素 $\hat{y}[x][y][i]$ 的均值参数 μ 和方差 σ 的值, 当均值 μ 、方差 σ 与 k 的关系满足 $\text{abs}(\mu - k) + \sigma < T3$ 时 (不满足预设条件), $T3$ 为第四阈值, 将当前待解码特征元素 $\hat{y}[x][y][i]$ 的数值设置为 k , 跳过对当前待解码特征元素 $\hat{y}[x][y][i]$ 执行熵解码过程, 否则, 当当前待解码特征元素的概率估计结果满足 $\text{abs}(\mu - k) + \sigma \geq T3$ 时 (预设条件), 对当前待解码特征元素 $\hat{y}[x][y][i]$ 执行熵解码, 得到当前待解码特征元素 $\hat{y}[x][y][i]$ 的值。

15 方法三: 当所述概率分布为拉普拉斯分布时, 根据所述概率估计结果, 得到位置参数 μ 、尺度参数 b 的值。当位置参数 μ 、尺度参数 b 与 k 的关系满足 $\text{abs}(\mu - k) + \sigma < T4$ 时 (不满足预设条件), $T4$ 为第四阈值, 将当前待解码特征元素 $\hat{y}[x][y][i]$ 的数值设置为 k , 跳过对当前待解码特征元素 $\hat{y}[x][y][i]$ 执行熵解码过程, 否则, 当当前待解码特征元素的概率估计结果满足 $\text{abs}(\mu - k) + \sigma \geq T4$ (预设条件), 对当前待解码特征元素 $\hat{y}[x][y][i]$ 执行熵解
20 码, 得到当前待解码特征元素 $\hat{y}[x][y][i]$ 的值。

方法四: 当所述概率分布为拉普拉斯分布时, 根据所述概率估计结果, 得到位置参数 μ 、尺度参数 b 的值。当位置参数 μ 与 k 的差的绝对值小于第二阈值 $T5$ 且尺度参数 b 小于第三阈值 $T6$ 时 (不满足预设条件), 将当前待解码特征元素 $\hat{y}[x][y][i]$ 的数值设置为 k , 跳过对当前待解码特征元素 $\hat{y}[x][y][i]$ 执行熵解码过程, 否则, 当位置参数 μ 与 k 的差的绝对
25 值小于第二阈值 $T5$ 或尺度参数 b 大于或者等于第三阈值 $T6$ (预设条件) 时, 对当前待解码特征元素 $\hat{y}[x][y][i]$ 执行熵解码, 得到当前待解码特征元素 $\hat{y}[x][y][i]$ 的值。

特别地, k 取值为 0 时为最优值, 可以直接判断当位置参数 μ 绝对值小于 $T5$ 且尺度参数 b 小于 $T6$ 时, 将当前待解码特征元素 $\hat{y}[x][y][i]$ 的数值设置为 k , 跳过对当前待解码特征元素 $\hat{y}[x][y][i]$ 执行熵解码过程, 否则, 对当前待解码特征元素 $\hat{y}[x][y][i]$ 执行熵解码,
30 得到当前待解码特征元素 $\hat{y}[x][y][i]$ 的值。

方法五：当所述概率分布为混合高斯分布时，根据所述概率估计结果，得到当前待解码特征元素 $\hat{y}[x][y][i]$ 的混合高斯分布的所有均值参数 μ_i 和方差 σ_i 的值。当混合高斯分布的所有均值与当前待解码特征元素取值 k 的差的绝对值之和与所述混合高斯分布的任一方差的和小于第五阈值 $T7$ 时（不满足预设条件），将当前待解码特征元素 $\hat{y}[x][y][i]$ 的数值设置为 k ，跳过对当前待解码特征元素 $\hat{y}[x][y][i]$ 执行熵解码过程，否则，当混合高斯分布的所有均值与所述当前待解码特征元素取值 k 的差的绝对值之和与所述混合高斯分布的任一方差的和大于或者等于第五阈值 $T7$ （预设条件），对当前待解码特征元素 $\hat{y}[x][y][i]$ 执行熵解码，得到当前待解码特征元素 $\hat{y}[x][y][i]$ 的值。

方法六：根据当前待解码特征元素的概率分布，得到当前待解码特征元素取值为 k 的概率 P ，即当前待解码特征元素的概率估计结果 P ，当概率估计结果 P 不满足预设条件： P 大于所述第一阈值 $T0$ 时，不需要对当前待解码特征元素执行熵解码，将当前待解码特征元素数值设置为 k ，否则，当当前待解码特征元素满足预设条件： P 小于或者等于所述第一阈值 $T0$ 时，对码流执行熵解码，得到当前待解码特征元素的值。

以上解码端的 k 值与编码端的 k 值相对应设置。

其中，获取阈值 $T0, T1, T2, T3, T4, T5, T6$ 和 $T7$ 的方法与编码端对应，可使用以下方法之一：

方法一：从码流中获取阈值，具体地，从序列头、图像头、Slice/条带或SEI中获取阈值。

方法二：解码端采用与编码端约定的固定阈值。

方法三：从码流中获取阈值索引号，具体地，从序列头、图像头、Slice/条带或SEI中获取阈值索引号。然后解码端使用与编码端相同的方式构建阈值候选列表，根据阈值索引号在阈值候选列表中得到相应的阈值。

需要说明的是，在实际应用中，为保证平台的一致性，可以对所述阈值 $T1, T2, T3, T4, T5$ 和 $T6$ 进行整点化，即进行移位放大为整数。

步骤1514：与步骤1414相同。

图13A示出了本申请实施例三提供的具体实现流程1600，运行步骤如下：

编码端：

步骤1601：与步骤1501相同，本步骤具体由图3C中的编码网络204来实现，具体可以参照上述对编码网络20的描述；

步骤 1602: 与步骤 1502 相同, 本步骤具体由图 3C 中的边信息提取 214 来实现;

步骤 1603: 对特征图 \hat{y} 进行概率估计得到各特征元素的概率估计结果;

本步骤具体可以由图 3C 中的概率估计 210 来实现, 具体可以参照上述对概率估计 40 的描述。可以使用概率分布模型来获得概率估计结果。其中, 概率分布模型可以为: 单高
5 斯模型或者非对称高斯模型或者混合高斯模型或者拉普拉斯分布模型。

当概率分布模型为高斯模型时(单高斯模型或者非对称高斯模型或者混合高斯模型), 首先将边信息 z 或者上下文信息输入概率估计网络, 对特征图 \hat{y} 中的每个特征元素 $\hat{y}[x][y][i]$ 进行概率估计得到模型参数均值参数 μ 和方差 σ 的值, 即概率估计结果。

当概率分布模型拉普拉斯分布模型时, 首先将边信息 z 或者上下文信息输入概率估计
10 网络, 对特征图 \hat{y} 中的每个特征元素 $\hat{y}[x][y][i]$ 进行概率估计得到模型参数位置参数 μ 和尺度参数 b 的值, 即概率估计结果。

进一步地, 将所述概率估计结果输入所使用的概率分布模型中, 得到概率分布。或者,

将边信息 z 和/或上下文信息输入概率估计网络, 对待编码特征图 \hat{y} 中的每个特征元素
15 $\hat{y}[x][y][i]$ 进行概率估计得到当前待编码特征元素 $\hat{y}[x][y][i]$ 的概率分布。根据所述概率分布, 得到当前待编码特征元素 $\hat{y}[x][y][i]$ 取值为 m 的概率 P 。其中, m 为任意的整数, 例如 0, 1, -1, -2, 3 等。

其中, 概率估计网络可以使用基于深度学习网络, 例如循环神经网络和卷积神经网络等, 在此不做限定。

步骤 1604: 根据概率估计结果判断是否对所述当前待编码特征元素执行熵编码。根
20 据所述判断结果对所述当前待编码特征元素执行熵编码并写入编码码流或不执行熵编码。仅当判断出需要对所述当前待编码特征元素执行熵编码时, 对所述当前待编码特征元素执行熵编码。

本步骤具体由图 3C 中的生成网络 216 以及编码决策实现 208 来实现, 具体可以参照
25 上述对生成网络 46 和编码决策实现 26 的描述。将所述概率估计结果 211 输入判断模块, 输出与特征图 \hat{y} 维度相同的决策信息 217。本实施例中决策信息 217 可以为三维的决策图 map 。其中, 判断模块可以使用网络的方法来实现, 即将所述概率估计结果或者概率分布输入图 7 所示的生成网络, 网络输出决策图 map 。决策图 $map[x][y][i]$ 为预设值表示对应位置的当前待编码特征元素 $\hat{y}[x][y][i]$ 需要熵编码, 根据概率分布对当前待编码特征元素执行熵编码。决策图 $map[x][y][i]$ 为预设值表示对应位置的当前待编码特征元素 $\hat{y}[x][y][i]$ 高
30 概率取值为 k , 决策图 $map[x][y][i]$ 不为预设值表示对应位置的当前待编码特征元素

$\hat{y}[x][y][i]$ 不需要熵编码,即跳过熵编码的过程。其中,决策信息是与特征图 \hat{y} 维度相同的决策图 map 。决策图 $\text{map}[x][y][i]$ 表示决策图 map 中坐标位置为 (x,y,i) 处的值。当生成网络输出的决策图 map 中所述当前待编码特征元素 \hat{y} 只有两种可选值时,预设值为特定某一数值,例如当前待编码特征元素可选数值为0和1时,预设值为0或1;当生成网络输出决策图 map 中当前待编码特征元素 \hat{y} 有多种可选值时,预设值为一些特定数值,例如当前待编码特征元素 \hat{y} 元素可选数值为0~255时,预设值为0~255的真子集。

在一种可能实现的方式中,将所述当前待编码特征元素的概率估计结果或者概率分布输入判断模块,判断模块直接输出当前待编码特征元素是否需要执行熵编码的决策信息。例如,判断模块输出的决策信息为预设值时,表示当前待编码特征元素需要执行熵编码,判断模块输出的决策信息不为预设值时,表示当前待编码特征元素不需要执行熵编码。判断模块可以通过网络的方法来实现,即将所述概率估计结果或者概率分布输入图7所示的生成网络,网络输出决策信息,即预设值。

方法一:决策信息是与特征图 \hat{y} 维度相同的决策图 map ,决策图 $\text{map}[x][y][i]$ 为预设值表示对应位置的当前待编码特征元素 $\hat{y}[x][y][i]$ 需要熵编码,根据概率分布对当前待编码特征元素执行熵编码。决策图 $\text{map}[x][y][i]$ 不为预设值表示对应位置的当前待编码特征元素 $\hat{y}[x][y][i]$ 高概率取值为 k ,决策图 $\text{map}[x][y][i]$ 为0表示对应位置的当前待编码特征元素 $\hat{y}[x][y][i]$ 不需要熵编码,即跳过熵编码的过程。当决策图 map 中所述特征元素 \hat{y} 只有两种可选值时,预设值为特定某一数值,例如特征元素可选数值为0和1时,预设值为0或1;当决策图 map 中特征元素 \hat{y} 有多种可选值时,预设值为一些特定数值,例如特征元素 \hat{y} 元素可选数值为0~255时,预设值为0~255的真子集。

方法二:决策信息是与特征图 \hat{y} 维度相同的决策图 map ,决策图 $\text{map}[x][y][i]$ 大于或等于阈值 T_0 表示对应位置的当前待编码特征元素 $\hat{y}[x][y][i]$ 需要熵编码,根据概率分布对当前待编码特征元素执行熵编码。决策图 $\text{map}[x][y][i]$ 小于阈值 T_0 表示对应位置的当前待编码特征元素 $\hat{y}[x][y][i]$ 高概率取值为 k ,表示对应位置的当前待编码特征元素 $\hat{y}[x][y][i]$ 不需要熵编码,即跳过熵编码的过程。根据决策图的数值范围, T_0 可为数值范围内的均值。

方法三:决策信息还可以是所述联合网络直接输出的标识或者标识的值,当决策信息为预设值时,表示当前待编码特征元素需要执行熵编码,判断模块输出的决策信息不为预设值时,表示当前待编码特征元素不需要执行熵编码。例如标识或者标识的值的可选数值为0和1时,则相应地预设值为0或1。当标识或者标识的值也可以有多种可选值时,预

设值为一些特定数值，例如标识或者标识的值的可选数值为 0~255 时，预设值为 0~255 的真子集。

其中高概率是指：当前待编码特征元素 $\hat{y}[x][y][i]$ 取值为 k 时的概率很高，大于阈值 P ，其中 P 可以是大于 0.9 的数，例如 0.9, 0.95 或 0.98 等。

5 步骤 1605：编码器发送或存储压缩码流。

对特征图 \hat{y} 中的至少一个特征元素执行上述步骤 1601 到 1604，以得到压缩码流，并将压缩码流传输到解码端。

解码端：

步骤 1611：获取待解码的压缩码流

10 步骤 1612：对待解码特征图 \hat{y} 进行概率估计得到各特征元素的概率估计结果

本步骤具体可以由图 13B 中的概率估计 302 来实现，具体可以参照上述对概率估计 40 的描述。从码流中获取边信息 \hat{z} ，使用步骤 1603 中方法获取当前待解码特征元素的概率估计结果

步骤 1613：获取决策信息，并根据决策信息判断是否执行熵解码。

15 本步骤具体可以由图 13B 中的生成网络 310 以及解码决策实现 304 来实现，具体可以参照上述对生成网络 46 和解码决策实现 30 的描述。使用与本实施例编码端相同的方法获取决策信息 311。决策图 $map[x][y][i]$ 为预设值表示对应位置的当前待解码特征元素 $\hat{y}[x][y][i]$ 需要熵解码，根据概率分布对当前待解码特征元素执行熵解码。决策图 $map[x][y][i]$ 不为预设值表示对应位置的当前待解码特征元素 $\hat{y}[x][y][i]$ 不需要熵解
20 码，即表示对应位置 $\hat{y}[x][y][i]$ 为特定数值 k 。

在一种可能实现的方式中，将所述当前待解码特征元素的概率估计结果或者概率分布输入判断模块，判断模块直接输出当前待解码特征元素是否需要执行熵解码的决策信息。例如，判断模块输出的决策信息为预设值时，表示当前待解码特征元素需要执行熵解码，判断模块输出的决策信息不为预设值时，表示当前待解码特征元素不需要执行熵解码，将
25 当前待解码特征元素的值设置为 k 。判断模块可以通过网络的方法来实现，即将所述概率估计结果或者概率分布输入图 8 所示的生成网络，网络输出决策信息，即预设值。此决策信息用于指示是否对所述当前待解码特征元素执行熵解码，所述决策信息可以包括决策图 map 。

步骤 1614：步骤 1414 相同。

30 以上解码端的 k 值与编码端的 k 值相对应设置。

图 14 示出了本申请实施例四的具体实现流程 1700，运行步骤如下：

编码端：

步骤 1701：与步骤 1501 相同，本步骤具体可以由图 3D 中的编码网络 204 来实现，
5 具体可以参照上述对编码网络 20 的描述；

步骤 1702：与步骤 1502 相同，本步骤具体由图 3D 中的边信息提取 214 来实现；

步骤 1703：获取特征图 \hat{y} 中每个特征元素的概率估计结果和决策信息；

本步骤具体可以由图 3D 中的联合网络 218 来实现，具体可以参照上述对联合网络 34
10 的描述。具体的，将边信息 \hat{z} 和/或上下文信息输入联合网络，联合网络输出待编码特征图
 \hat{y} 中的每个特征元素 $\hat{y}[x][y][i]$ 的概率分布和/或概率估计结果，及与特征图 \hat{y} 维度相同的决
策信息。例如，当同时将边信息 \hat{z} 和上下文信息输入联合网络，可采用网络结构如图 15 所
示。

需要说明的是，本实施例对联合网络的具体结构不做约束。

需要说明的是，决策信息、概率分布和/或概率估计结果均可以从联合网络的不同层
15 输出。例如：情况 1) 网络中间层输出决策信息，最后层输出概率分布和/或概率估计结
果；情况 2) 网络中间层输出概率分布和/或概率估计结果，最后层输出决策信息；情况 3)
网络最后层一起输出决策信息及概率分布和/或概率估计结果。

当概率分布模型为高斯模型时(单高斯模型或者非对称高斯模型或者混合高斯模型)，
首先将边信息 \hat{z} 或者上下文信息输入联合网络，得到模型参数均值参数 μ 和方差 σ 的值，
20 即概率估计结果。进一步地，将概率估计结果输入高斯模型中，得到概率分布。

当概率分布模型拉普拉斯分布模型时，首先将边信息 \hat{z} 或者上下文信息输入联合网络，
得到模型参数位置参数 μ 和尺度参数 b 的值，即概率估计结果。进一步地，将概率估计结
果输入拉普拉斯分布模型中，得到概率分布。

或者，将边信息 \hat{z} 和/或上下文信息输入联合网络，得到当前待编码特征元素 $\hat{y}[x][y][i]$
25 的概率分布。根据所述概率分布，得到当前待编码特征元素 $\hat{y}[x][y][i]$ 取值为 m 的概率 P ，
为概率估计结果。其中， m 为任意的整数，例如 0, 1, -1, -2, 3 等。

步骤 1704，并根据决策信息判断是否执行熵编码；根据判断结果执行熵编码并写入
压缩码流(编码码流)或者不执行熵编码。仅当判断出需要对所述当前待编码特征元素执行
熵编码时，对所述当前待编码特征元素执行熵编码。本步骤具体可以由图 3D 中的编码决
30 策实现 208 来实现，具体可以参照上述编码决策实现 26 的描述。

方法一：决策信息是与特征图 \hat{y} 维度相同的决策图 map ，决策图 $\text{map}[x][y][i]$ 为预设值表示对应位置的当前待编码特征元素 $\hat{y}[x][y][i]$ 需要熵编码，根据概率分布对当前待编码特征元素执行熵编码。决策图 $\text{map}[x][y][i]$ 不为预设值表示对应位置的当前待编码特征元素 $\hat{y}[x][y][i]$ 高概率取值为 k ，决策图 $\text{map}[x][y][i]$ 为 0 表示对应位置的当前待编码特征元素 $\hat{y}[x][y][i]$ 不需要熵编码，即跳过熵编码的过程。当决策图 map 中所述当前待编码特征元素 \hat{y} 只有两种可选值时，预设值为特定某一数值，例如当前待编码特征元素可选数值为 0 和 1 时，预设值为 0 或 1；当决策图 map 中当前待编码特征元素 \hat{y} 有多种可选值时，预设值为一些特定数值，例如当前待编码特征元素 \hat{y} 元素可选数值为 0~255 时，预设值为 0~255 的真子集。

方法二：决策信息是与特征图 \hat{y} 维度相同的决策图 map ，决策图 $\text{map}[x][y][i]$ 大于或等于阈值 T_0 表示对应位置的当前待编码特征元素 $\hat{y}[x][y][i]$ 需要熵编码，根据概率分布对当前待编码特征元素执行熵编码。决策图 $\text{map}[x][y][i]$ 小于阈值 T_0 表示对应位置的当前待编码特征元素 $\hat{y}[x][y][i]$ 高概率取值为 k ，表示对应位置的当前待编码特征元素 $\hat{y}[x][y][i]$ 不需要熵编码，即跳过熵编码的过程。根据决策图 map 的数值范围， T_0 可为数值范围内的均值。

方法三：决策信息还可以是所述联合网络直接输出的标识或者标识的值，当决策信息为预设值时，表示当前待编码特征元素需要执行熵编码，判断模块输出的决策信息不为预设值时，表示当前待编码特征元素不需要执行熵编码。当联合网络输出决策图 map 中所述当前待编码特征元素只有两种可选值时，预设值为特定某一数值，例如所述当前待编码特征元素可选数值为 0 和 1 时，预设值为 0 或 1；当联合网络输出决策图 map 中所述当前待编码特征元素有多种可选值时，预设值为一些特定数值，例如所述当前待编码特征元素可选数值为 0~255 时，预设值为 0~255 的真子集。

其中高概率是指：当前待编码特征元素 $\hat{y}[x][y][i]$ 取值为 m 时的概率很高，例如取值为 k 时的概率大于阈值 P ，其中 P 可以是大于 0.9 的数，例如 0.9, 0.95 或 0.98 等。

步骤 1705：编码器发送或存储压缩码流。

解码端：

步骤 1711：获取待解码图像特征图的码流，从码流中获取边信息 \hat{z} 。

步骤 1712：获取特征图 \hat{y} 中每个特征元素的概率估计结果和决策信息

本步骤具体可以由图 16 中的联合网络 312 来实现，具体可以参照上述对联合网络 34 的描述。获取特征图 \hat{y} 中每个特征元素的概率估计结果和决策信息的方法同步骤 1703。

步骤 1713: 根据决策信息判断是否执行熵解码; 根据判断结果执行或者不执行熵解码, 本步骤具体可以由图 16 中的解码决策实现 304 来实现, 具体可以参照上述对解码决策实现 30 的描述。

方法一: 决策信息是决策图 map, 决策图 $map[x][y][i]$ 为预设值表示对应位置的当前待解码特征元素 $\hat{y}[x][y][i]$ 需要熵解码, 根据概率分布对当前待解码特征元素执行熵解码。
5 决策图 $map[x][y][i]$ 不为预设值表示对应位置的当前待解码特征元素 $\hat{y}[x][y][i]$ 不需要熵解码, 即表示对应位置 $\hat{y}[x][y][i]$ 设定为特定数值 k。

方法二: 决策信息是与特征图 \hat{y} 维度相同的决策图 map, 决策图 $map[x][y][i]$ 大于或等于阈值 T_0 表示对应位置的当前待解码特征元素 $\hat{y}[x][y][i]$ 需要熵解码。
10 决策图 $map[x][y][i]$ 小于阈值 T_0 表示对应位置的当前待解码特征元素 $\hat{y}[x][y][i]$ 高概率取值为 k, 表示对应位置的当前待解码特征元素 $\hat{y}[x][y][i]$ 不需要熵解码, 即表示对应位置 $\hat{y}[x][y][i]$ 设定为特定数值 k。 T_0 取值与编码端相同。

方法三: 决策信息还可以是所述联合网络直接输出的标识或者标识的值, 当所述决策信息为预设值时, 表示当前待解码特征元素需要执行熵解码, 判断模块输出所述决策信息
15 不为预设值时, 表示当前待解码特征元素不需要执行熵解码, 将当前待解码特征元素的值设置为 k。当联合网络输出决策图 map 中所述当前待解码特征元素只有两种可选值时, 预设值为特定某一数值, 例如所述当前待解码特征元素可选数值为 0 和 1 时, 预设值为 0 或 1; 当联合网络输出决策图 map 中所述当前待解码特征元素有多种可选值时, 预设值为一些特定数值, 例如所述当前待解码特征元素可选数值为 0~255 时, 预设值为 0~255 的真
20 子集。

步骤 1714: 与步骤 1414 相同, 本步骤具体可以由上述实施例的解码器 9C 中的解码网络单元 306 来实现, 具体可以参照上述实施例对解码网络单元 306 的描述。

以上解码端的 k 值与编码端的 k 值相对应设置。

25 图 17 示出了本申请实施例五的具体实现流程 1800, 运行步骤如下:

步骤 1801: 获取待编码音频数据的特征变量

待编码音频信号可以是时域音频信号; 待编码音频信号可以是时域信号经过时频变换后得到的频域信号, 例如频域信号可以是时域音频信号经过 MDCT 变换后的频域信号, 时域音频信号经过 FFT 变换后的频域信号; 待编码信号也可以是 QMF 滤波后的信号; 待编码
30 信号还可以是残差信号, 例如其他编码的残差信号或者 LPC 滤波后的残差信号。

获取待编码音频数据的特征变量：可以是根据待编码音频信号提取特征矢量，例如根据待编码音频信号提取梅尔倒谱系数；对提取的特征矢量进行量化，将量化后的特征矢量作为待编码音频数据的特征变量。

5 获取待编码音频数据的特征变量：还可以利用现有的神经网络来实现，例如将待编码音频信号经过编码神经网络处理获得潜在变量，对神经网络输出的潜在变量进行量化，将量化后的潜在变量作为待编码音频数据的特征变量。编码神经网络处理是预先训练好的，本发明对编码神经网络的具体网络结构和训练方法不做限定。例如编码神经网络可以选择全连接网络或者 CNN 网络。本发明对编码神经网络包含的层数和每一层的节点数也不做限定。

10 不同结构的编码神经网络输出的潜在变量的形式可能不同。例如，编码神经网络是全连接网络，输出的潜在变量为一个矢量，矢量的维数M是潜在变量的大小（latent size），例如 $y=[y(0), y(1), \dots, y(M-1)]$ 。编码神经网络是CNN网络，输出的潜在变量为一个 $N \times M$ 维矩阵，其中N为CNN网络的通道数（channel），M为CNN网络的每个通道潜在变量的大小（latent size），如

$$15 \quad y = \begin{bmatrix} y(0,0) & \dots & y(0, M-1) \\ \vdots & \ddots & \vdots \\ y(N-1,0) & \dots & y(N-1, M-1) \end{bmatrix}$$

对神经网络输出的潜在变量进行量化的具体方法可以是对潜在变量的每个元素进行标量量化，标量量化的量化步长可以根据不同的编码速率来确定。标量量化还可以存在偏置量，例如待量化的潜在变量经过偏置处理后再按照确定好的量化步长进行标量量化。对潜在变量进行量化的量化方法还可以使用其他的现有量化技术实现，本发明不做限定。

20 量化后的特征矢量或者量化后的潜在变量均可记作 \hat{y} ，即待编码音频数据的特征变量。
步骤 1802：待编码音频数据的特征变量 \hat{y} 输入边信息提取模块，输出边信息 \hat{z} 。

其中，边信息提取模块可以使用图 12 所示的网络来实现，边信息 \hat{z} 可以理解为对特征变量 \hat{y} 进行进一步提取得到的特征变量 \hat{z} ， \hat{z} 所含包含的特征元素的个数比特征变量 \hat{y} 少。

需要说明的是，可以在本步骤中，对边信息 \hat{z} 进行熵编码并写入码流，也可以在后续
25 的步骤 1804 中对边信息 \hat{z} 进行熵编码并写入码流，在此不做限定。

步骤 1803：对特征变量 \hat{y} 进行概率估计得到各特征元素的概率估计结果。

可以使用概率分布模型来获得概率估计结果及概率分布。其中，概率分布模型可以为：单高斯模型（Gaussian single model, GSM）或者非对称高斯模型或者混合高斯模型（Gaussian mixture model, GMM）或者拉普拉斯分布模型（Laplace distribution）。

下面以特征变量 \hat{y} 为 $N \times M$ 维矩阵为例进行说明。当前待编码特征变量 \hat{y} 中的特征元素记作 $\hat{y}[j][i]$, $j \in [0, N - 1]$, $i \in [0, M - 1]$ 。

当概率分布模型为高斯模型时(单高斯模型或者非对称高斯模型或者混合高斯模型), 5 首先将边信息 z 或者上下文信息输入概率估计网络, 对特征变量 \hat{y} 中的每个特征元素 $\hat{y}[j][i]$ 进行概率估计得到均值参数 μ 和方差 σ 的值。进一步地, 将所述均值参数 μ 和方差 σ 输入所使用的概率分布模型中, 得到概率分布。此时概率估计结果为均值参数 μ 和方差 σ 。

也可以值估计方差, 例如当概率分布模型为高斯模型时(单高斯模型或者非对称高斯模型或者混合高斯模型), 首先将边信息 z 或者上下文信息输入概率估计网络, 对特征变 10 量 \hat{y} 中的每个特征元素 $\hat{y}[j][i]$ 进行概率估计得到方差 σ 的值。进一步地, 将所述方差 σ 输入所使用的概率分布模型中, 得到概率分布。此时概率估计结果为方差 σ 。

当概率分布模型拉普拉斯分布模型时, 首先将边信息 z 或者上下文信息输入概率估计网络, 对特征图变量 \hat{y} 中的每个特征元素 $\hat{y}[j][i]$ 进行概率估计得到位置参数 μ 和尺度参数 b 15 的值。进一步地, 将所述位置参数 μ 和尺度参数 b 输入所使用的概率分布模型中, 得到概率分布。此时概率估计结果为位置参数 μ 和尺度参数 b 。

还可以将边信息 z 和/或上下文信息输入概率估计网络, 对待编码特征图 \hat{y} 中的每个特征元素 $\hat{y}[j][i]$ 进行概率估计得到当前待编码特征元素 $\hat{y}[j][i]$ 的概率分布。根据所述概率分布, 得到当前待编码特征元素 $\hat{y}[j][i]$ 取值为 m 的概率 P 。此时概率估计结果为当前待编码特征元素 $\hat{y}[j][i]$ 取值为 m 的概率 P 。

其中, 概率估计网络可以使用基于深度学习网络, 例如循环神经网络 (Recurrent 20 Neural Network, RNN) 和卷积神经网络 (Convolutional Neural Network, PixelCNN) 等, 在此不做限定。

步骤 1804: 根据概率估计结果判断当前待编码特征元素是否需要执行熵编码, 并根据判断结果执行熵编码写入压缩码流(编码码流)或者不执行熵编码。

根据概率估计结果判断当前待编码特征元素 $\hat{y}[j][i]$ 是否需要执行熵编码可以使用以下 25 方法中的一项或者多项。其中, 参数 j, i 为正整数, 坐标 (j, i) 表示当前待编码特征元素的位置。或者, 根据概率估计结果判断当前待编码特征元素 $\hat{y}[i]$ 是否需要执行熵编码可以使用以下方法中的一项或者多项。其中, 参数 i 为正整数, 坐标 i 表示当前待编码特征元素的位置。

下面以根据概率估计结果判断当前待编码特征元素 $\hat{y}[j][i]$ 是否需要执行熵编码为例进 30 行说明, 判断当前待编码特征元素 $\hat{y}[i]$ 是否需要执行熵编码的方法类似, 这里不再赘述。

方法一：当所述概率分布模型为高斯分布时，根据所述第一特元素的概率估计结果判断是否对所述当前待编码特征元素执行熵编码，当当前待编码特征元素的高斯分布的均值参数 μ 和方差 σ 的值满足第二条件：当均值 μ 与 k 的差的绝对值小于第二阈值 $T1$ 且方差 σ 小于第三阈值 $T2$ 时，不需要对当前待编码特征元素 $\hat{y}[j][i]$ 执行熵编码过程，否则，当满足第一条件：当均值 μ 与 k 的差的绝对值大于或者等于第二阈值 $T1$ 或方差 σ 小于第三阈值 $T2$ 时，对当前待编码特征元素 $\hat{y}[j][i]$ 进行熵编码写入码流。其中， k 为任意的整数，例如 $0, 1, -1, 2, 3$ 等。 $T2$ 取值为满足 $0 < T2 < 1$ 中的任一数，例如取值为 $0.2, 0.3, 0.4$ 等。 $T1$ 是大于或等于 0 小于 1 的数，例如 $0.01, 0.02, 0.001, 0.002$ 。

特别地， k 取值为 0 时为最优值，可以直接判断当高斯分布的均值参数 μ 绝对值小于 $T1$ 且高斯分布的方差 σ 小于 $T2$ 时，则跳过对当前待编码特征元素 $\hat{y}[j][i]$ 进行熵编码过程，否则，对当前待编码特征元素 $\hat{y}[j][i]$ 进行熵编码写入码流。其中， $T2$ 的取值为满足 $0 < T2 < 1$ 中的任一数，例如取值为 $0.2, 0.3, 0.4$ 等。 $T1$ 是大于或等于 0 小于 1 的数，例如 $0.01, 0.02, 0.001, 0.002$ 。

方法二：当所述概率分布为高斯分布时，根据所述概率估计结果，得到当前待编码特征元素 $\hat{y}[j][i]$ 的高斯分布的均值参数 μ 和方差 σ 的值，当均值 μ 、方差 σ 与 k 的关系满足 $abs(\mu - k) + \sigma < T3$ 时（第二条件），跳过对当前待编码特征元素 $\hat{y}[j][i]$ 进行熵编码过程，其中， $abs(\mu - k)$ 表示计算均值 μ 与 k 的差的绝对值；否则，当当前待编码特征元素的概率估计结果满足 $abs(\mu - k) + \sigma \geq T3$ 时（第一条件），对当前待编码特征元素 $\hat{y}[j][i]$ 进行熵编码写入码流。其中， k 为任意的整数，例如 $0, 1, -1, -2, 3$ 等。其中，第四阈值 $T3$ 是大于或等于 0 小于 1 的数，例如取值为 $0.2, 0.3, 0.4$ 等。

当所述概率分布为高斯分布时，如果对特征变量 \hat{y} 中的每个特征元素 $\hat{y}[j][i]$ 进行概率估计仅得到当前待编码特征元素 $\hat{y}[j][i]$ 的高斯分布的方差 σ 的值，当方差 σ 满足 $\sigma < T3$ 时（第二条件），跳过对当前待编码特征元素 $\hat{y}[j][i]$ 进行熵编码过程；否则，当当前待编码特征元素的概率估计结果满足 $\sigma \geq T3$ 时（第一条件），对当前待编码特征元素 $\hat{y}[j][i]$ 进行熵编码写入码流。其中，第四阈值 $T3$ 是大于或等于 0 小于 1 的数，例如取值为 $0.2, 0.3, 0.4$ 等。

方法三：当所述概率分布为拉普拉斯分布时，根据所述概率估计结果，得到当前待编码特征元素 $\hat{y}[j][i]$ 的拉普拉斯分布的位置参数 μ 、尺度参数 b 的值。当位置参数 μ 、尺度参数 b 与 k 的关系满足 $abs(\mu - k) + \sigma < T4$ （第二条件）时，跳过对当前待编码特征元素 $\hat{y}[j][i]$ 进行熵编码过程，其中， $abs(\mu - k)$ 表示计算位置参数 μ 与 k 的差的绝对值；否则，

当当前待编码特征元素的概率估计结果满足 $abs(\mu - k) + \sigma \geq T4$ （第一条件），对当前待编码特征元素 $\hat{y}[j][i]$ 进行熵编码写入码流。其中， k 为任意的整数，例如0, 1, -1, -2, 3等。第四阈值 $T4$ 是大于或等于0小于0.5的数，例如取值为0.05, 0.09, 0.17等。

方法四：当所述概率分布为拉普拉斯分布时，根据所述概率估计结果，得到当前待编码特征元素 $\hat{y}[j][i]$ 的拉普拉斯分布的位置参数 μ 、尺度参数 b 的值。当位置参数 μ 与 k 的差的绝对值小于第二阈值 $T5$ 且尺度参数 b 小于第三阈值 $T6$ （第二条件）时，跳过对当前待编码特征元素 $\hat{y}[j][i]$ 进行熵编码过程，否则，当位置参数 μ 与 k 的差的绝对值小于第二阈值 $T5$ 或尺度参数 b 大于或者等于第三阈值 $T6$ （第一条件）时，对当前待编码特征元素 $\hat{y}[j][i]$ 进行熵编码写入码流。其中， k 为任意的整数，例如0, 1, -1, -2, 3等。 $T5$ 取值为
10 $1e-2$ ， $T6$ 取值为满足 $T6 < 0.5$ 中的任一数，例如取值为0.05, 0.09, 0.17等。

特别地， k 取值为0时为最优值，可以直接判断当位置参数 μ 绝对值小于 $T5$ 且尺度参数 b 小于 $T6$ 时，则跳过对当前待编码特征元素 $\hat{y}[j][i]$ 进行熵编码过程，否则，对当前待编码特征元素 $\hat{y}[j][i]$ 进行熵编码写入码流。其中，阈值 $T5$ 取值为 $1e-2$ ， $T2$ 的取值为满足 $T6 < 0.5$ 中的任一数，例如取值为0.05, 0.09, 0.17等。

方法五：当所述概率分布为混合高斯分布时，根据所述概率估计结果，得到当前待编码特征元素 $\hat{y}[j][i]$ 的混合高斯分布的所有均值参数 μ_i 和方差 σ_i 的值。当混合高斯分布的所有均值与 k 的差的绝对值之和与所述混合高斯分布的任一方差的和小于第五阈值 $T7$ （第二条件）时，跳过对当前待编码特征元素 $\hat{y}[j][i]$ 进行熵编码过程；否则，当混合高斯分布的所有均值与 k 的差的绝对值之和与所述混合高斯分布的任一方差的和大于或者等于第
20 五阈值 $T7$ （第一条件），对当前待编码特征元素 $\hat{y}[j][i]$ 进行熵编码写入码流。其中， k 为任意的整数，例如0, 1, -1, -2, 3等。 $T7$ 是大于或等于0小于1的数，例如取值为0.2, 0.3, 0.4等。（可以认为每个特征元素的阈值都相同）

方法六：根据所述概率分布，得到当前待编码特征元素 $\hat{y}[j][i]$ 取值为 k 的概率 P ，当当前待编码特征元素的概率估计结果 P 满足第二条件：当 P 大于（或者等于）第一阈值
25 $T0$ 时，跳过当前待编码特征元素进行熵编码过程；否则，当当前待编码特征元素的概率估计结果 P 满足第一条件：当 P 小于第一阈值 $T0$ 时，对当前待编码特征元素进行熵编码写入码流。其中， k 可为任意整数，例如0, 1, -1, 2, 3等。所述第一阈值 $T0$ 为满足 $0 < T0 < 1$ 中的任一数，例如取值为0.99, 0.98, 0.97, 0.95等。（可以认为每个特征元素的阈值都相同）

需要说明的是，在实际应用中，为保证平台的一致性，可以对所述阈值 T1, T2, T3, T4, T5 和 T6 进行整点化，即进行移位放大为整数。

需要说明的是，阈值的获取方法还可以使用以下方法之一，在此不做限定：

方法一：以阈值 T1 为例，取 T1 取值范围内的任意一个取值作为阈值 T1，将阈值 T1 写入码流。具体地，将所述阈值写入码流，可将其保存在序列头、图像头、Slice/条带或 SEI 中传送到解码端，还可以使用其他方法，在此不做限定。其余阈值 T0, T2, T3, T4, T5 和 T6 也可以使用类似方法。

方法二：编码端采用与解码端约定的固定阈值，无需写入码流，无需传输到解码端。例如，以阈值 T1 为例，直接取 T1 取值范围内任一值作为 T1 的取值。其余阈值 T0, T2, T3, T4, T5 和 T6 也可以使用类似方法。

方法三：构建阈值候选列表，将在 T1 取值范围内最有可能的取值放入阈值候选列表中，每个阈值对应一个阈值索引号，确定一个最优的阈值，将最优阈值作为 T1 的值，并将最优阈值的索引号作为 T1 的阈值索引号，将 T1 的阈值索引号写入码流。具体地，将所述阈值写入码流，可将其保存在序列头、图像头、Slice/条带或 SEI 中传送到解码端，还可以使用其他方法，在此不做限定。其余阈值 T0, T2, T3, T4, T5 和 T6 也可以使用类似方法。

步骤 1805：编码器发送或存储压缩码流。

解码端：

步骤 1811：获取待解码音频特征变量的码流

步骤 1812：获取各特征元素的概率估计结果

对边信息 \hat{z} 进行熵解码得到边信息 \hat{z} ，结合边信息 \hat{z} 对待解码音频特征变量 \hat{y} 中的每个特征元素 $\hat{y}[j][i]$ 进行概率估计，得到当前待解码特征元素 $\hat{y}[j][i]$ 的概率估计结果。其中，参数 j, i 为正整数，坐标 (j, i) 表示当前待解码特征元素的位置。或者，对边信息 \hat{z} 进行熵解码得到边信息 \hat{z} ，结合边信息 \hat{z} 对待解码音频特征变量 \hat{y} 中的每个特征元素 $[i]$ 进行概率估计，得到当前待解码特征元素 $\hat{y}[i]$ 的概率估计结果。其中，参数 i 为正整数，坐标 i 表示当前待解码特征元素的位置。

需要说明的是，解码端使用的概率估计方法与本实施例编码端的概率估计方法对应相同，概率估计网络结构图与本实施例的编码端概率估计网络结构相同，在此不做赘述。

步骤 1813: 根据概率估计结果判断当前待解码特征元素是否需要执行熵解码, 并根据判断结果执行或者不执行熵解码, 得到解码后的特征变量 \hat{y} 。

根据概率估计结果判断当前待解码特征元素 $\hat{y}[j][i]$ 是否需要执行熵解码可以使用以下方法中的一项或者多项。或者, 根据概率估计结果判断当前待解码特征元素 $\hat{y}[i]$ 是否需要
5 执行熵解码可以使用以下方法中的一项或者多项。

下面以根据概率估计结果判断当前待解码特征元素 $\hat{y}[j][i]$ 是否需要执行熵解码为例进行说明, 判断当前待解码特征元素 $\hat{y}[i]$ 是否需要执行熵解码的方法类似, 这里不再赘述。

方法一: 当所述概率分布模型为高斯分布时, 根据所述概率估计结果, 得到当前待解码特征元素 $\hat{y}[j][i]$ 的均值参数 μ 和方差 σ 的值, 当均值 μ 与 k 的差的绝对值小于第二阈值
10 T1 且方差 σ 小于第三阈值 T2 时 (第二条件), 将当前待解码特征元素 $\hat{y}[j][i]$ 的数值设置为 k , 跳过对当前待解码特征元素 $\hat{y}[j][i]$ 进行熵解码过程; 否则, 当均值 μ 与 k 的差的绝对值小于第二阈值 T1 或方差 σ 大于或等于第三阈值 T2 时 (第一条件), 对当前待解码特征元素 $\hat{y}[j][i]$ 进行熵解码, 得到当前待解码特征元素 $\hat{y}[j][i]$ 的值。

特别地, k 取值为 0 时为最优值, 可以直接判断当高斯分布的均值参数 μ 绝对值小
15 于 T1 且高斯分布的方差 σ 小于 T2 时, 将当前待解码特征元素 $\hat{y}[j][i]$ 的数值设置为 k , 则跳过对当前待解码特征元素 $\hat{y}[j][i]$ 进行熵解码过程, 否则, 对当前待解码特征元素 $\hat{y}[j][i]$ 进行熵解码, 得到当前待解码特征元素 $\hat{y}[j][i]$ 的值。

方法二: 当所述概率分布为高斯分布时, 根据所述概率估计结果, 得到当前待解码特征元素 $\hat{y}[j][i]$ 的均值参数 μ 和方差 σ 的值, 当均值 μ 、方差 σ 与 k 的关系满足
20 $\text{abs}(\mu - k) + \sigma < T3$ 时 (第二条件), T3 为第四阈值, 将当前待解码特征元素 $\hat{y}[j][i]$ 的数值设置为 k , 跳过对当前待解码特征元素 $\hat{y}[j][i]$ 进行熵解码过程, 否则, 当当前待解码特征元素的概率估计结果满足 $\text{abs}(\mu - k) + \sigma \geq T3$ 时 (第一条件), 对当前待解码特征元素 $\hat{y}[j][i]$ 进行熵解码, 得到当前待解码特征元素 $\hat{y}[j][i]$ 的值。当所述概率分布为高斯分布时, 如果根据所述概率估计结果, 仅得到当前待解码特征元素 $\hat{y}[j][i]$ 的方差 σ 的值, 当方差 σ
25 关系满足 $\sigma < T3$ 时 (第二条件), T3 为第四阈值, 将当前待解码特征元素 $\hat{y}[j][i]$ 的数值设置为 0, 跳过对当前待解码特征元素 $\hat{y}[j][i]$ 进行熵解码过程, 否则, 当当前待解码特征元素的概率估计结果满足 $\sigma \geq T3$ 时 (第一条件), 对当前待解码特征元素 $\hat{y}[j][i]$ 进行熵解码, 得到当前待解码特征元素 $\hat{y}[j][i]$ 的值。

方法三: 当所述概率分布为拉普拉斯分布时, 根据所述概率估计结果, 得到位置参数
30 μ 、尺度参数 b 的值。当位置参数 μ 、尺度参数 b 与 k 的关系满足 $\text{abs}(\mu - k) + \sigma < T4$ 时 (第

二条件)，T4 为第四阈值，将当前待解码特征元素 $\hat{y}[j][i]$ 的数值设置为 k，跳过对特征元素 $\hat{y}[j][i]$ 进行熵解码过程，否则，当当前待解码特征元素的概率估计结果满足 $\text{abs}(\mu - k) + \sigma \geq T4$ （第一条条件），对特征元素 $\hat{y}[j][i]$ 进行熵解码，得到特征元素 $\hat{y}[j][i]$ 的值。

方法四：当所述概率分布为拉普拉斯分布时，根据所述概率估计结果，得到位置参数
5 μ 、尺度参数 b 的值。当位置参数 μ 与 k 的差的绝对值小于第二阈值 T5 且尺度参数 b 小于第三阈值 T6 时（第二条条件），将当前待解码特征元素 $\hat{y}[j][i]$ 的数值设置为 k，跳过对当前待解码特征元素 $\hat{y}[j][i]$ 进行熵解码过程，否则，当位置参数 μ 与 k 的差的绝对值小于第二阈值 T5 或尺度参数 b 大于或者等于第三阈值 T6（第一条条件）时，对当前待解码特征元素 $\hat{y}[j][i]$ 进行熵解码，得到当前待解码特征元素 $\hat{y}[j][i]$ 的值。

10 特别地，k 取值为 0 时为最优值，可以直接判断当位置参数 μ 绝对值小于 T5 且尺度参数 b 小于 T6 时，将当前待解码特征元素 $\hat{y}[j][i]$ 的数值设置为 k，跳过对当前待解码特征元素 $\hat{y}[j][i]$ 进行熵解码过程，否则，对当前待解码特征元素 $\hat{y}[j][i]$ 进行熵解码，得到当前待解码特征元素 $\hat{y}[j][i]$ 的值。

方法五：当所述概率分布为混合高斯分布时，根据所述概率估计结果，得到当前待解
15 码特征元素 $\hat{y}[j][i]$ 的混合高斯分布的所有均值参数 μ_i 和方差 σ_i 的值。当混合高斯分布的所有均值与 k 的差的绝对值之和与所述混合高斯分布的任一方差的和小于第五阈值 T7 时（第二条条件），将当前待解码特征元素 $\hat{y}[j][i]$ 的数值设置为 k，跳过对当前待解码特征元素 $\hat{y}[j][i]$ 进行熵解码过程，否则，当混合高斯分布的所有均值与 k 的差的绝对值之和与所述混合高斯分布的任一方差的和大于或者等于第五阈值 T7（第一条条件），对当前待解
20 码特征元素 $\hat{y}[j][i]$ 进行熵解码，得到当前待解码特征元素 $\hat{y}[j][i]$ 的值。

方法六：根据当前待解码特征元素的概率分布，得到当前待解码特征元素取值为 k 的概率 P，即当前待解码特征元素的概率估计结果 P，当概率估计结果 P 满足第二条条件：P 大于所述第一阈值 T0 时，不需要对第一特征元素执行熵解码，将当前待解码特征元素数值设置为 k，否则，当当前待解码特征元素满足第一条条件：P 小于或者等于所述第一阈
25 值 T0 时，对码流进行熵解码，得到第一特征元素的值。

以上解码端的 k 值与编码端的 k 值相对应设置。

其中，获取阈值 T0，T1，T2，T3，T4，T5，T6 和 T7 的方法与编码端对应，可使用以下方法之一：

方法一：从码流中获取阈值，具体地，从序列头、图像头、Slice/条带或 SEI 中获取
30 阈值。

方法二：解码端采用与编码端约定的固定阈值。

方法三：从码流中获取阈值索引号，具体地，从序列头、图像头、Slice/条带或SEI中获取阈值索引号。然后解码端使用与编码端相同的方式构建阈值候选列表，根据阈值索引号在阈值候选列表中得到相应的阈值。

5 需要说明的是，在实际应用中，为保证平台的一致性，可以对所述阈值T1，T2，T3，T4，T5和T6进行整点化，即进行移位放大为整数。

步骤1814：对解码后的特征变量 \hat{y} 进行重建，或者输入面向机器听觉任务模块执行相应的机器任务。本步骤具体可以由图10B中的解码网络306来实现，具体可以参照上述对解码网络34的描述。

10 情况一：将熵解码后的特征变量 \hat{y} 输入图像重建模块，神经网络输出重建音频。所述神经网络可以采用任一结构，例如全连接网络、卷积神经网络、循环神经网络等。所述神经网络可以采用多层的结构深度神经网络结构来达到更好的估计效果。

情况二：将熵解码后的特征变量 \hat{y} 输入面向机器听觉任务模块执行相应的机器任务。例如完成音频分类、识别等机器听觉任务。

15 以上解码端的k值与编码端的k值相对应设置。

图18为本申请编码装置的一个示例性的结构示意图，如图18所示，本示例的装置可以对应于编码器20A。该装置可以包括：获得模块2001和编码模块2002。获得模块2001可以包括前述实施例中的编码网络204、取整206(可选)、概率估计210、边信息提取214、生成网络216(可选)及联合网络218(可选)。编码模块2002包括前述实施例中的编码决策实现208。其中，

20 获得模块2001，用于获取待编码特征数据，所述待编码特征数据包括多个特征元素，所述多个特征元素包括第一特征元素，以及用于获取所述中第一特征元素的概率估计结果；编码模块2002，用于根据所述第一特征元素的概率估计结果，判断是否对所述第一特征元素执行熵编码；仅当判断出需要对所述第一特征元素执行熵编码时，对所述第一特征元素执行熵编码。

25 在一种可能的实现方式中，所述判断是否对所述特征数据的第一特征元素执行熵编码包括：当所述特征数据的第一特征元素的概率估计结果满足预设条件，需要对所述特征数据的第一特征元素熵编码；当所述特征数据的第一特征元素的概率估计结果不满足预设条件，不需要对所述特征数据的第一特征元素熵编码。

30 件，不需要对所述特征数据的第一特征元素熵编码。

在一种可能的实现方式中,所述编码模块,还用于根据所述特征数据的概率估计结果判断:所述特征数据的概率估计结果输入生成网络,网络输出决策信息。当所述第一特征元素的决策信息的取值为1时,需要对所述特征数据的第一特征元素编码;当所述第一特征元素的决策信息的取值不为1时,不需要对所述特征数据的第一特征元素编码。

5 在一种可能的实现方式中,所述预设条件为第一特征元素取值为 k 的概率值小于或等于第一阈值,其中 k 为整数。

在一种可能的实现方式中,所述预设条件为所述第一特征元素的概率分布的均值与第一特征元素取值 k 的差的绝对值大于或等于第二阈值或所述第一特征元素的方差大于或等于第三阈值,其中 k 为整数。

10 在另一种可能的实现方式中,所述预设条件为所述第一特征元素的概率分布的均值与第一特征元素取值 k 的差的绝对值与所述第一特征元素的概率分布的方差的和大于或等于第四阈值,其中 k 为整数。

在一种可能的实现方式中,所述第一特征元素取值为 k 的概率值为所述第一特征元素的所有可能的取值的概率值中的最大概率值。

15 在一种可能的实现方式中,对所述特征数据进行概率估计以得到所述特征数据中各特征元素的概率估计结果,其中所述第一特征元素的概率估计结果包括所述第一特征元素的概率值,和/或所述概率分布的第一参数和所述概率分布的第二参数。

在一种可能的实现方式中,将所述特征数据的概率估计结果输入生成网络以得到所述第一特征元素的决策信息。根据所述第一特征元素的决策信息,判断是否对所述第一特征元素执行熵编码。

20

在一种可能的实现方式中,当所述特征数据的决策信息为决策图时,则所述决策图中对应所述第一特征元素所在位置的值为预设值时,判断需要对所述第一特征元素执行熵编码;当所述决策图中对应所述第一特征元素所在位置的值为不为预设值时,判断不需要对所述第一特征元素执行熵编码。

25 在一种可能的实现方式中,当所述特征数据的决策信息为预设值时,判断需要对所述第一特征元素执行熵编码;当所述决策信息不为预设值时,判断不需要对所述第一特征元素执行熵编码。在一种可能的实现方式中,所述编码模块,还用于构建第一阈值的阈值候选列表,将所述第一阈值放入所述第一阈值的阈值候选列表中且对应有所述第一阈值的索引号,将所述第一阈值的索引号写入编码码流,其中所述第一阈值的阈值候选列表的长度

30 可以设置为 T ; T 为大于或等于1的整数。

本实施例的装置，可以用于图 3A-3D 所示方法实施例中由编码器实施的技术方案，其实现原理和技术效果类似，此处不再赘述。

图 19 为本申请解码装置的一个示例性的结构示意图，如图 19 所示，本示例的装置可以对应于解码器 30。该装置可以包括：获得模块 2101 和解码模块 2102。获得模块 2101
5 可以包括前述实施例中的概率估计 302、生成网络 310(可选)及联合网络 312。解码模块 2102 包括前述实施例中的解码决策实现 304 和解码网络 306。其中，

获得模块 2101，用于获取待解码特征数据的码流，所述待解码特征数据包括多个特征元素，所述多个特征元素包括第一特征元素；获取所述第一特征元素的概率估计结果；解码模块 2102，用于根据所述第一特征元素的概率估计结果，判断是否对所述第一特征
10 元素执行熵解码；仅当判断出需要对所述第一特征元素执行熵解码时，对所述第一特征元素执行熵解码。

在一种可能的实现方式中，所述判断是否对所述特征数据的第一特征元素熵解码包括：当所述特征数据的第一特征元素的概率估计结果满足预设条件，需要对所述特征数据的第一特征元素解码；或当所述特征数据的第一特征元素的概率估计结果不满足预设条件，不
15 需要对所述特征数据的第一特征元素解码，将第一特征元素的特征值设置为 k ；其中 k 为整数。

在一种可能的实现方式中，所述解码模块，还用于根据所述特征数据的概率估计结果判断：所述特征数据的概率估计结果输入判断网络模块，网络输出决策信息。当所述决策信息中对应所述特征数据的第一特征元素位置的取值为 1 时，对所述特征数据的第一特征
20 元素解码；当所述决策信息中对应所述特征数据的第一特征元素位置的取值不为 1 时，不对所述特征数据的第一特征元素解码，将第一特征元素的特征值设置为 k ，其中 k 为整数。

在一种可能的实现方式中，所述预设条件为第一特征元素取值为 k 的概率值小于等于第一阈值，其中 k 为整数。

在另一种可能的实现方式中，所述预设条件为所述第一特征元素的概率分布的均值与
25 所述第一特征元素取值 k 的差的绝对值大于或等于第二阈值或所述第一特征元素的概率分布的方差大于或等于第三阈值。

在另一种可能的实现方式中，所述预设条件为所述第一特征元素的概率分布的均值与
所述第一特征元素取值 k 的差的绝对值与所述第一特征元素的概率分布的方差的和大于
或等于第四阈值。

30 在一种可能的实现方式中，对所述特征数据进行概率估计以得到所述特征数据中各特

征元素的概率估计结果，其中所述第一特征元素的概率估计结果包括所述第一特征元素的概率值，和/或所述概率分布的第一参数和所述概率分布的第二参数。

在一种可能的实现方式中，所述第一特征元素取值为 k 的概率值为所述第一特征元素的所有可能的取值的概率值中的最大概率值。

- 5 在一种可能的实现方式中，所述第 N 个特征元素的概率估计结果包括至少以下一项：所述第 N 个特征元素的概率值，概率分布的第一参数和概率分布的第二参数以及决策信息。当所述决策信息中对应所述特征数据的第一特征元素位置的取值为 1 时，对所述特征数据的第一特征元素解码；当所述决策信息中对应所述特征数据的第一特征元素位置的取值不为 1 时，不对所述特征数据的第一特征元素解码，将第一特征元素的特征值设置为 k ，其中 k 为整数。
- 10

- 在一种可能的实现方式中，将所述特征数据的概率估计结果输入生成网络以得到所述第一特征元素的决策信息；当所述第一特征元素的决策信息的值为预设值时，判断需要对所述第一特征元素执行熵解码；当所述第一特征元素的决策信息的值不为预设值时，判断不需要对所述第一特征元素执行熵解码，将第一特征元素的特征值设置为 k ，其中 k 为整数且 k 为所述第一特征元素的多个候选取值中的一个。
- 15

在一种可能的实现方式中，所述获得模块，还用于构建第一阈值的阈值候选列表，通过对所述码流进行解码以得到所述第一阈值的阈值候选列表的索引号，将所述第一阈值的索引号所对应所述第一阈值的阈值候选列表位置的值作为所述第一阈值的值，其中所述第一阈值的阈值候选列表的长度可以设置为 T ； T 为大于或等于 1 的整数。

- 20 本实施例的装置，可以用于图 10B, 13B, 16 所示方法实施例中由解码器实施的技术方案，其实现原理和技术效果类似，此处不再赘述。

- 本领域技术人员能够领会，结合本文公开描述的各种说明性逻辑框、模块和算法步骤所描述的功能可以硬件、软件、固件或其任何组合来实施。如果以软件来实施，那么各种说明性逻辑框、模块、和步骤描述的功能可作为一或多个指令或代码在计算机可读媒体上存储或传输，且由基于硬件的处理单元判断。计算机可读媒体可包含计算机可读存储媒体，其对应于有形媒体，例如数据存储媒体，或包括任何促进将计算机程序从一处传送到另一处的媒体（例如，根据通信协议）的通信媒体。以此方式，计算机可读媒体大体上可对应于
- 25
- (1) 非暂时性的有形计算机可读存储媒体，或 (2) 通信媒体，例如信号或载波。数据存储媒体可为可由一或多个计算机或一或多个处理器存取以检索用于实施本申请中描述的技术的指令、代码和/或数据结构的任何可用媒体。计算机程序产品可包含计算机可读媒
- 30

体。

作为实例而非限制，此类计算机可读存储媒体可包括 RAM、ROM、EEPROM、CD-ROM 或其它光盘存储装置、磁盘存储装置或其它磁性存储装置、快闪存储器或可用来存储指令或数据结构的形式的所要程序代码并且可由计算机存取的任何其它媒体。并且，任何连接被
5 恰当地称作计算机可读媒体。举例来说，如果使用同轴缆线、光纤缆线、双绞线、数字订户线 (DSL) 或例如红外线、无线电和微波等无线技术从网站、服务器或其它远程源传输指令，那么同轴缆线、光纤缆线、双绞线、DSL 或例如红外线、无线电和微波等无线技术包含在媒体的定义中。但是，应理解，所述计算机可读存储媒体和数据存储媒体并不包括连接、载波、信号或其它暂时媒体，而是实际上针对于非暂时性有形存储媒体。如本文中
10 使用，磁盘和光盘包含压缩光盘 (CD)、激光光盘、光学光盘、数字多功能光盘 (DVD) 和蓝光光盘，其中磁盘通常以磁性方式再现数据，而光盘利用激光以光学方式再现数据。以上各项的组合也应包含在计算机可读媒体的范围内。

可通过例如一或多个数字信号处理器 (DSP)、通用微处理器、专用集成电路 (ASIC)、现场可编程逻辑阵列 (FPGA) 或其它等效集成或离散逻辑电路等一或多个处理器来判断指令。因此，如本文中所使用的术语“处理器”可指前述结构或适合于实施本文中所描述的技术的任一其它结构中的任一者。另外，在一些方面中，本文中所描述的各种说明性逻辑框、模块、和步骤所描述的功能可以提供于经配置以用于编码和解码的专用硬件和/或软件模块内，或者并入在组合编解码器中。而且，所述技术可完全实施于一或多个电路或逻辑元件中。
15

本申请的技术可在各种各样的装置或设备中实施，包含无线手持机、集成电路 (IC) 或一组 IC (例如，芯片组)。本申请中描述各种组件、模块或单元是为了强调用于判断所揭示的技术的装置的功能方面，但未必需要由不同硬件单元实现。实际上，如上文所描述，各种单元可结合合适的软件和/或固件组合在编码解码器硬件单元中，或者通过互操作硬件单元（包含如上文所描述的一或多个处理器）来提供。
20

以上所述，仅为本申请示例性的具体实施方式，但本申请的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本申请揭露的技术范围内，可轻易想到的变化或替换，都应涵盖在本申请的保护范围之内。因此，本申请的保护范围应该以权利要求的保护范围为准。
25

30

权 利 要 求 书

1、一种特征数据的编码方法，其特征在于，包括：

获取待编码特征数据，所述待编码特征数据包括多个特征元素，所述多个特征元素包括第一特征元素；

5 获取所述第一特征元素的概率估计结果；

根据所述第一特征元素的概率估计结果，判断是否对所述第一特征元素执行熵编码；
仅当判断出需要对所述第一特征元素执行熵编码时，对所述第一特征元素执行熵编码。

2、根据权利要求 1 所述的方法，其特征在于，所述根据所述第一特征元素的概率估计结果，判断是否对所述第一特征元素执行熵编码包括：

10 当所述第一特征元素的概率估计结果满足预设条件时，判断需要对所述第一特征元素执行熵编码；或

当所述第一特征元素的概率估计结果不满足预设条件时，判断不需要对所述第一特征元素执行熵编码。

3、根据权利要求 2 所述的方法，其特征在于，所述第一特征元素的概率估计结果为
15 所述第一特征元素取值为 k 的概率值，则所述预设条件为所述第一特征元素取值为 k 的概率值小于或等于第一阈值，其中 k 为整数且 k 为所述第一特征元素的多个候选取值中的一个。

4、根据权利要求 2 所述的方法，其特征在于，所述第一特征元素的概率估计结果包括所述第一特征元素概率分布的第一参数和第二参数，则所述预设条件为：

20 所述第一特征元素的概率分布的所述第一参数与所述第一特征元素取值为 k 的差的绝对值大于或等于第二阈值；或

所述第一特征元素的概率分布的所述第二参数大于或等于第三阈值；或

25 所述第一特征元素的概率分布的所述第一参数与所述第一特征元素取值为 k 的差的绝对值与所述第一特征元素的概率分布的所述第二参数的和大于或等于第四阈值，其中 k 为整数且 k 为所述第一特征元素的多个候选取值中的一个。

5、根据权利要求 4 项所述的方法，其特征在于：

当所述概率分布为高斯分布，所述第一特征元素概率分布的第一参数为所述第一特征元素高斯分布的均值，所述第一特征元素概率分布的第二参数为所述第一特征元素高斯分布的方差；或

30 当所述概率分布为拉普拉斯分布，所述第一特征元素概率分布的第一参数为所述第一

特征元素拉普拉斯分布的位置参数，所述第一特征元素概率分布的第二参数为所述第一特征元素拉普拉斯分布的尺度参数。

6、根据权利要求 3 所述的方法，其特征在于，所述方法还包括：

5 构建阈值候选列表，将所述第一阈值放入所述阈值候选列表中，且将对应有所述第一阈值的索引号写入编码码流，其中所述阈值候选列表的长度为 T ， T 为大于或等于 1 的整数。

7、根据权利要求 2 所述的方法，其特征在于，所述第一特征元素的概率估计结果通过混合高斯分布获得时，则所述预设条件为：

10 所述第一特征元素的混合高斯分布的所有均值与所述第一特征元素的取值为 k 的差的绝对值之和与所述第一特征元素的混合高斯分布的任一方差的和大于或等于第五阈值；或

所述第一特征元素的混合高斯分布的任一均值与所述第一特征元素的取值为 k 的差大于或等于第六阈值；或

所述第一特征元素的混合高斯分布的任一方差大于或等于第七阈值；

15 其中 k 为整数且 k 为所述第一特征元素的多个候选取值中的一个。

8、根据权利要求 2 所述的方法，其特征在于，所述第一特征元素的概率估计结果通过非对称高斯分布获得时，则所述预设条件为：

所述第一特征元素的非对称高斯分布的均值与所述第一特征元素的取值为 k 的差的绝对值大于或等于第八阈值；或

20 所述第一特征元素的非对称高斯分布的第一方差大于或等于第九阈值；或

所述第一特征元素的非对称高斯分布的第二方差大于或等于第十阈值；

其中 k 为整数且 k 为所述第一特征元素的多个候选取值中的一个。

9、根据权利要求 3-8 任一所述的方法，其特征在于：

25 所述第一特征元素取值为 k 的概率值为所述第一特征元素的所有候选取值的概率值中的最大概率值。

10、根据权利要求 1 所述的方法，其特征在于，所述根据所述第一特征元素的概率估计结果，判断是否对所述第一特征元素执行熵编码包括：

将所述特征数据的概率估计结果输入生成网络以得到所述第一特征元素的决策信息，根据所述第一特征元素的决策信息，判断是否对所述第一特征元素执行熵编码。

30 11、根据权利要求 10 所述的方法，其特征在于，当所述特征数据的决策信息为决策

图时，则所述决策图中对应所述第一特征元素所在位置的值为预设值时，判断需要对所述第一特征元素执行熵编码；

当所述决策图中对应所述第一特征元素所在位置的值为不为预设值时，判断不需要对所述第一特征元素执行熵编码。

5 12、根据权利要求 10 所述的方法，其特征在于，当所述特征数据的决策信息为预设值时，判断需要对所述第一特征元素执行熵编码；

当所述决策信息不为预设值时，判断不需要对所述第一特征元素执行熵编码。

13、根据权利要求 1-12 任一所述的方法，其特征在于，所述多个特征元素还包括第二特征元素，当判断出不需要对所述第二特征元素执行熵编码时，跳过对所述第二特征元
10 素执行熵编码。

14、根据权利要求 1-13 任一所述的方法，其特征在于，所述方法还包括：

将包括所述第一特征元素的多个特征元素的熵编码结果写入编码码流。

15、一种特征数据的解码方法，其特征在于，包括：

获取待解码特征数据的码流；

15 所述待解码特征数据包括多个特征元素，所述多个特征元素包括第一特征元素；

获取所述第一特征元素的概率估计结果；

根据所述第一特征元素的概率估计结果，判断是否对所述第一特征元素执行熵解码；

仅当判断出需要对所述第一特征元素执行熵解码时，对所述第一特征元素执行熵解码。

16、根据权利要求 15 所述的方法，其特征在于，所述根据所述第一特征元素的概率
20 估计结果，判断是否对所述第一特征元素执行熵解码包括：

当所述第一特征元素的概率估计结果满足预设条件时，判断需要对所述特征数据的第一特征元素执行熵解码；或

25 当所述第一特征元素的概率估计结果不满足预设条件时，判断不需要对所述特征数据的第一特征元素执行熵解码，将所述第一特征元素的特征值设置为 k ，其中 k 为整数且 k 为所述第一特征元素的多个候选取值中的一个。

17、根据权利要求 16 所述的方法，其特征在于，当所述第一特征元素的概率估计结果为所述第一特征元素取值为 k 的概率值，则所述预设条件为所述第一特征元素取值为 k 的概率值小于或等于第一阈值，其中 k 为整数且 k 为所述第一特征元素的多个候选取值中的一个。

30 18、根据权利要求 16 所述的方法，其特征在于，当所述第一特征元素的概率估计结

果包括所述第一特征元素概率分布的第一参数和第二参数，则所述预设条件为：

所述第一特征元素的概率分布的所述第一参数与所述第一特征元素取值为 k 的差的绝对值大于或等于第二阈值；或

所述第一特征元素的概率分布的所述第二参数大于或等于第三阈值；或

5 所述第一特征元素的概率分布的所述第一参数与所述第一特征元素取值为 k 的差的绝对值与所述第一特征元素的概率分布的所述第二参数的和大于或等于第四阈值，

其中 k 为整数且 k 为所述第一特征元素的多个候选取值中的一个。

19、根据权利要求 18 所述的方法，其特征在于：

10 当所述概率分布为高斯分布，所述第一特征元素概率分布的第一参数为所述第一特征元素高斯分布的均值，所述第二参数为所述第一特征元素高斯分布的方差；或

当所述概率分布为拉普拉斯分布，所述第一特征元素概率分布的第一参数为所述第一特征元素拉普拉斯分布的位置参数，所述第二参数为所述第一特征元素拉普拉斯分布的尺度参数。

15 20、根据权利要求 16 所述的方法，其特征在于，所述第一特征元素的概率估计结果通过混合高斯分布获得时，则所述预设条件为：

所述第一特征元素的混合高斯分布的所有均值与所述第一特征元素的取值为 k 的差的绝对值之和与所述第一特征元素的混合高斯分布的任一方差的和大于或等于第五阈值；或

20 所述第一特征元素的混合高斯分布的任一均值与所述第一特征元素的取值为 k 的差大于或等于第六阈值；或

所述第一特征元素的混合高斯分布的任一方差大于或等于第七阈值；

其中 k 为整数且 k 为所述第一特征元素的多个候选取值中的一个。

25 21、根据权利要求 16 所述的方法，其特征在于，所述第一特征元素的概率估计结果通过非对称高斯分布获得时，则所述预设条件为：

所述第一特征元素的非对称高斯分布的均值与所述第一特征元素的取值为 k 的差的绝对值大于或等于第八阈值；或

所述第一特征元素的非对称高斯分布的第一方差大于或等于第九阈值；或

所述第一特征元素的非对称高斯分布的第二方差大于或等于第十阈值；

30 其中 k 为整数且 k 为所述第一特征元素的多个候选取值中的一个。

22、根据权利要求 16-21 任一所述的装置，其特征在于，所述第一特征元素取值为 k 的概率值为所述第一特征元素的所有候选取值的概率值中的最大概率值。

23、根据权利要求 15 所述的方法，其特征在于，所述根据所述第一特征元素的概率估计结果，判断是否对所述第一特征元素执行熵解码包括：

5 将所述特征数据的概率估计结果输入生成网络以得到所述第一特征元素的决策信息，根据所述第一特征元素的决策信息，判断是否对所述第一特征元素执行熵解码。

24、根据权利要求 23 所述的方法，其特征在于，当所述特征数据的决策信息为决策图时，则所述决策图中对应所述第一特征元素所在位置的值为预设值时，判断需要对所述第一特征元素执行熵解码；

10 当所述决策图中对应所述第一特征元素所在位置的值为不为预设值时，判断不需要对所述第一特征元素执行熵解码。

25、根据权利要求 23 所述的方法，其特征在于，当所述特征数据的决策信息为预设值时，判断需要对所述第一特征元素执行熵解码；

当所述决策信息不为预设值时，判断不需要对所述第一特征元素执行熵解码。

15 26、根据权利要求 15-25 任一所述的方法，其特征在于，所述方法还包括：所述特征数据经过解码网络以得到所述重建数据或面向机器任务数据。

27、一种特征数据编码装置，其特征在于，包括：

获得模块，用于获取待编码特征数据，所述待编码特征数据包括多个特征元素，所述多个特征元素包括第一特征元素，以及用于获取所述中第一特征元素的概率估计结果；

20 编码模块，用于根据所述第一特征元素的概率估计结果，判断是否对所述第一特征元素执行熵编码；仅当判断出需要对所述第一特征元素执行熵编码时，对所述第一特征元素执行熵编码。

28、根据权利要求 27 所述的装置，其特征在于，所述根据所述第一特征元素的概率估计结果，判断是否对所述第一特征元素执行熵编码包括：

25 当所述第一特征元素的概率估计结果满足预设条件时，判断需要对所述特征数据的第一特征元素执行熵编码；或

当所述第一特征元素的概率估计结果不满足预设条件时，判断不需要对所述特征数据的第一特征元素执行熵编码。

29、根据权利要求 28 所述的装置，其特征在于，所述第一特征元素的概率估计结果为所述第一特征元素取值为 k 的概率值，则所述预设条件为所述第一特征元素取值为 k 的

30

概率值小于或等于第一阈值，其中 k 为整数且 k 为所述第一特征元素的多个候选取值中的一个。

30、根据权利要求 28 所述的装置，其特征在于，所述第一特征元素的概率估计结果包括所述第一特征元素概率分布的第一参数和第二参数，则所述预设条件为：

5 所述所述第一特征元素的概率分布的所述第一参数与所述第一特征元素取值为 k 的差的绝对值大于或等于第二阈值；或

所述所述第一特征元素的概率分布的所述第二参数大于或等于第三阈值；或

所述所述第一特征元素的概率分布的所述第一参数与所述第一特征元素取值为 k 的差的绝对值与所述第一特征元素的概率分布的所述第二参数的和大于或等于第四阈值，

10 其中 k 为整数且 k 为所述第一特征元素的多个候选取值中的一个。

31、根据权利要求 30 所述的装置，其特征在于：

当所述概率分布为高斯分布，所述第一特征元素概率分布的第一参数为所述第一特征元素高斯分布的均值，所述第一特征元素概率分布的第二参数为所述第一特征元素高斯分布的方差；或

15 当所述概率分布为拉普拉斯分布，所述第一特征元素概率分布的第一参数为所述第一特征元素拉普拉斯分布的位置参数，所述第一特征元素概率分布的第二参数为所述第一特征元素拉普拉斯分布的尺度参数。

32、根据权利要求 29 所述的装置，其特征在于：

20 所述编码模块，还用于构建阈值候选列表，将所述第一阈值放入所述阈值候选列表中，且将对有所述第一阈值的索引号写入编码码流，其中所述阈值候选列表的长度为 T ， T 为大于或等于 1 的整数。

33、根据权利要求 28 所述的装置，其特征在于，所述第一特征元素的概率估计结果通过混合高斯分布获得时，则所述预设条件为：

25 所述第一特征元素的混合高斯分布的所有均值与所述第一特征元素的取值为 k 的差的绝对值之和与所述第一特征元素的混合高斯分布的任一方差的和大于或等于第五阈值；或

所述第一特征元素的混合高斯分布的任一均值与所述第一特征元素的取值为 k 的差大于或等于第六阈值；或

所述第一特征元素的混合高斯分布的任一方差大于或等于第七阈值；

30 其中 k 为整数且 k 为所述第一特征元素的多个候选取值中的一个。

34、根据权利要求 28 所述的装置，其特征在于，所述第一特征元素的概率估计结果通过非对称高斯分布获得时，则所述预设条件为：

所述第一特征元素的非对称高斯分布的均值与所述第一特征元素的取值为 k 的差的绝对值大于或等于第八阈值；或

5 所述第一特征元素的非对称高斯分布的第一方差大于或等于第九阈值；或

所述第一特征元素的非对称高斯分布的第二方差大于或等于第十阈值；

其中 k 为整数且 k 为所述第一特征元素的多个候选取值中的一个。

35、根据权利要求 29-34 任一所述的方法，其特征在于：

10 所述第一特征元素取值为 k 的概率值为所述第一特征元素的所有候选取值的概率值中的最大概率值。

36、根据权利要求 27 所述的装置，其特征在于，所述根据所述第一特征元素的概率估计结果，判断是否对所述第一特征元素执行熵编码包括：

将所述特征数据的概率估计结果输入生成网络以得到所述第一特征元素的决策信息，根据所述第一特征元素的决策信息，判断是否对所述第一特征元素执行熵编码。

15 37、根据权利要求 36 所述的装置，其特征在于，当所述特征数据的决策信息为决策图时，则所述决策图中对应所述第一特征元素所在位置的值预设值时，判断需要对所述第一特征元素执行熵编码；

当所述决策图中对应所述第一特征元素所在位置的值不为预设值时，判断不需要对所述第一特征元素执行熵编码。

20 38、根据权利要求 36 所述的装置，其特征在于，当所述特征数据的决策信息为预设值时，判断需要对所述第一特征元素执行熵编码；

当所述决策信息不为预设值时，判断不需要对所述第一特征元素执行熵编码。

25 39、根据权利要求 27-38 任一所述的装置，其特征在于，所述多个特征元素还包括第二特征元素，当判断出不需要对所述第二特征元素执行熵编码时，跳过对所述第二特征元素执行熵编码。

40、根据权利要求 27-39 任一所述的装置，其特征在于，所述编码模块还包括：

将包括所述的第一特征元素的多个特征元素的熵编码结果写入编码码流

41、一种特征数据解码装置，其特征在于，包括：

30 获得模块，用于获取待解码特征数据的码流，所述待解码特征数据包括多个特征元素，所述多个特征元素包括第一特征元素；获取所述第一特征元素的概率估计结果；

解码模块，用于根据所述第一特征元素的概率估计结果，判断是否对所述第一特征元素执行熵解码；仅当判断出需要对所述第一特征元素执行熵解码时，对所述第一特征元素执行熵解码。

42、根据权利要求 41 所述的装置，其特征在于，所述根据所述第一特征元素的概率估计结果，判断是否对所述第一特征元素执行熵解码包括：

当所述第一特征元素的概率估计结果满足预设条件时，判断需要对所述特征数据的第一特征元素执行熵解码；或

当所述第一特征元素的概率估计结果不满足预设条件时，判断不需要对所述特征数据的第一特征元素执行熵解码，将所述第一特征元素的特征值设置为 k ，其中 k 为整数且 k 为多个候选取值中的一个。

43、根据权利要求 42 所述的装置，其特征在于，当所述第一特征元素的概率估计结果为所述第一特征元素取值为 k 的概率值，则所述预设条件为所述第一特征元素取值为 k 的概率值小于或等于第一阈值，其中 k 为整数且 k 为所述第一特征元素的多个候选取值中的一个。

44、根据权利要求 42 所述的装置，其特征在于，当所述第一特征元素的概率估计结果包括所述第一特征元素概率分布的第一参数和第二参数，则所述预设条件为：

所述第一特征元素的概率分布的所述第一参数与所述第一特征元素取值为 k 的差的绝对值大于或等于第二阈值；或

所述第一特征元素的概率分布的所述第二参数大于或等于第三阈值；或

所述第一特征元素的概率分布的所述第一参数与所述第一特征元素取值为 k 的差的绝对值与所述第一特征元素的概率分布的所述第二参数的和大于或等于第四阈值，

其中 k 为整数且 k 为所述第一特征元素的多个候选取值中的一个。

45、根据权利要求 44 所述的装置，其特征在于：

当所述概率分布为高斯分布，所述第一特征元素概率分布的第一参数为所述第一特征元素高斯分布的均值，所述第一特征元素概率分布的第二参数为所述第一特征元素高斯分布的方差；或

当所述概率分布为拉普拉斯分布，所述第一特征元素概率分布的第一参数为所述第一特征元素拉普拉斯分布的位置参数，所述第一特征元素概率分布的第二参数为所述第一特征元素拉普拉斯分布的尺度参数。

46、根据权利要求 42 所述的装置，其特征在于，所述第一特征元素的概率估计结果

通过混合高斯分布获得时，则所述预设条件为：

所述第一特征元素的混合高斯分布的所有均值与所述第一特征元素的取值为 k 的差的绝对值之和与所述第一特征元素的混合高斯分布的任一方差的和大于或等于第五阈值；
或

5 所述第一特征元素的混合高斯分布的任一均值与所述第一特征元素的取值为 k 的差大于或等于第六阈值；或

所述第一特征元素的混合高斯分布的任一方差大于或等于第七阈值；

其中 k 为整数且 k 为所述第一特征元素的多个候选取值中的一个。

47、根据权利要求 42 所述的装置，其特征在于，所述第一特征元素的概率估计结果
10 通过非对称高斯分布获得时，则所述预设条件为：

所述第一特征元素的非对称高斯分布的均值与所述第一特征元素的取值为 k 的差的绝对值大于或等于第八阈值；或

所述第一特征元素的非对称高斯分布的第一方差大于或等于第九阈值；或

所述第一特征元素的非对称高斯分布的第二方差大于或等于第十阈值；

15 其中 k 为整数且 k 为所述第一特征元素的多个候选取值中的一个。

48、根据权利要求 42-47 任一所述的装置，其特征在于，所述第一特征元素取值为 k 的概率值为所述第一特征元素的所有候选取值的概率值中的最大概率值。

49、根据权利要求 41 所述的装置，其特征在于，所述根据所述第一特征元素的概率估计结果，判断是否对所述第一特征元素执行熵解码包括：

20 将所述特征数据的概率估计结果输入生成网络以得到所述第一特征元素的决策信息，根据所述第一特征元素的决策信息，判断是否对所述第一特征元素执行熵解码。

50、根据权利要求 49 所述的装置，其特征在于，当所述特征数据的决策信息为决策图时，则所述决策图中对应所述第一特征元素所在位置的值预设值时，需要对所述第一特征元素执行熵解码；

25 当所述决策图中对应所述第一特征元素所在位置的值不为预设值时，不需要对所述第一特征元素执行熵解码。

51、根据权利要求 49 所述的装置，其特征在于，当所述特征数据的决策信息为预设值时，判断需要对所述第一特征元素执行熵解码；

当所述决策信息不为预设值时，判断不需要对所述第一特征元素执行熵解码。

30 52、根据权利要求 41-51 任一所述的装置，其特征在于：

所述解码模块，还用于所述特征数据经过解码网络以得到所述重建数据或面向机器任务数据。

53、一种特征数据的编码方法，其特征在于，包括：

5 获取待编码特征数据，所述特征数据包括多个特征元素，所述多个特征元素包括第一特征元素；

获取所述特征数据的边信息，对所述特征数据的边信息输入联合网络以得到所述第一特征元素的决策信息；

根据所述第一特征元素的决策信息，判断是否对所述第一特征元素执行熵编码；

仅当判断出需要对所述第一特征元素执行熵编码时，对所述第一特征元素执行熵编码。

10 54、根据权利要求 53 所述的方法，其特征在于，当所述特征数据的决策信息为决策图时，则所述决策图中对应所述第一特征元素所在位置的值为预设值时，判断需要对所述第一特征元素执行熵编码；

当所述决策图中对应所述第一特征元素所在位置的值为不为预设值时，判断不需要对所述第一特征元素执行熵编码。

15 55、根据权利要求 53 所述的方法，其特征在于，当所述特征数据的决策信息为预设值时，判断需要对所述第一特征元素执行熵编码；

当所述决策信息不为预设值时，判断不需要对所述第一特征元素执行熵编码。

20 56、根据权利要求 53-55 任一所述的方法，其特征在于，所述多个特征元素还包括第二特征元素，当判断出不需要对所述第二特征元素执行熵编码时，跳过对所述第二特征元素执行熵编码。

57、根据权利要求 53-56 任一所述的方法，其特征在于，所述方法还包括：

将包括所述的第一特征元素的多个特征元素的熵编码结果写入编码码流。

58、一种特征数据的解码方法，其特征在于，包括：

获取待解码特征数据的码流和所述待解码特征数据的边信息；

25 所述待解码特征数据包括多个特征元素，所述多个特征元素包括第一特征元素；

对所述待解码特征数据的边信息输入联合网络以得到所述第一特征元素的决策信息；

根据所述第一特征元素的决策信息，判断是否对所述第一特征元素执行熵解码；

仅当判断出需要对所述第一特征元素执行熵解码时，对所述第一特征元素执行熵解码。

30 59、根据权利要求 58 所述的方法，其特征在于，当所述特征数据的决策信息为决策图时，则所述决策图中对应所述第一特征元素所在位置的值为预设值时，判断需要对所述

第一特征元素执行熵解码；

当所述决策图中对应所述第一特征元素所在位置的值不为预设值时，判断不需要对所述第一特征元素执行熵解码，将所述第一特征元素的特征值设置为 k ，其中 k 为整数。

60、根据权利要求 58 所述的方法，其特征在于，当所述特征数据的决策信息为预设值时，判断需要对所述第一特征元素执行熵编码；

当所述决策信息不为预设值时，判断不需要对所述第一特征元素执行熵编码，将所述第一特征元素的特征值设置为 k ，其中 k 为整数。

61、根据权利要求 58-60 任一所述的方法，其特征在于，所述方法还包括：

所述特征数据经过解码网络以得到所述重建数据或面向机器任务数据。

62、一种编码器，其特征在于，包括处理电路，用于执行权利要求 1 至 14, 53 至 57 任一项所述的方法。

63、一种解码器，其特征在于，包括处理电路，用于执行权利要求 15 至 26, 58 至 61 任一项所述的方法。

64、一种计算机程序产品，其特征在于，包括程序代码，当其在计算机或处理器上判断时，用于判断权利要求 1 至 26, 53 至 61 任一项所述的方法。

65、一种非瞬时性计算机可读存储介质，其特征在于，包括根据权利要求 14 或 57 所述的编码方法获得的码流。

66、一种编码器，其特征在于，包括：

一个或多个处理器；

非瞬时性计算机可读存储介质，耦合到所述处理器并存储由所述处理器判断的程序，其中所述程序在由所述处理器判断时，使得所述编码器执行根据权利要求 1 至 14, 53 至 57 任一项所述的方法。

67、一种解码器，其特征在于，包括：

一个或多个处理器；

非瞬时性计算机可读存储介质，耦合到所述处理器并存储由所述处理器判断的程序，其中所述程序在由所述处理器判断时，使得所述解码器执行根据权利要求 15 至 26, 58 至 61 任一项所述的方法。

68、一种编码器，其特征在于，包括：

一个或多个处理器；

非瞬时性计算机可读存储介质，耦合到所述处理器并存储由所述处理器判断的程序，

其中所述程序在由所述处理器判断时，使得所述编码器执行根据权利要求 1 至 14, 53 至 57 任一项所述的方法。

69、一种图像或音频处理器，其特征在于，包括处理电路，用于执行根据权利要求 1 至 26, 53 至 61 任一项所述的方法。

5 70、一种非瞬时性计算机可读存储介质，其特征在于，包括程序代码，当其由计算机设备判断时，用于执行根据权利要求 1 至 26, 53 至 61 任一项所述的方法。

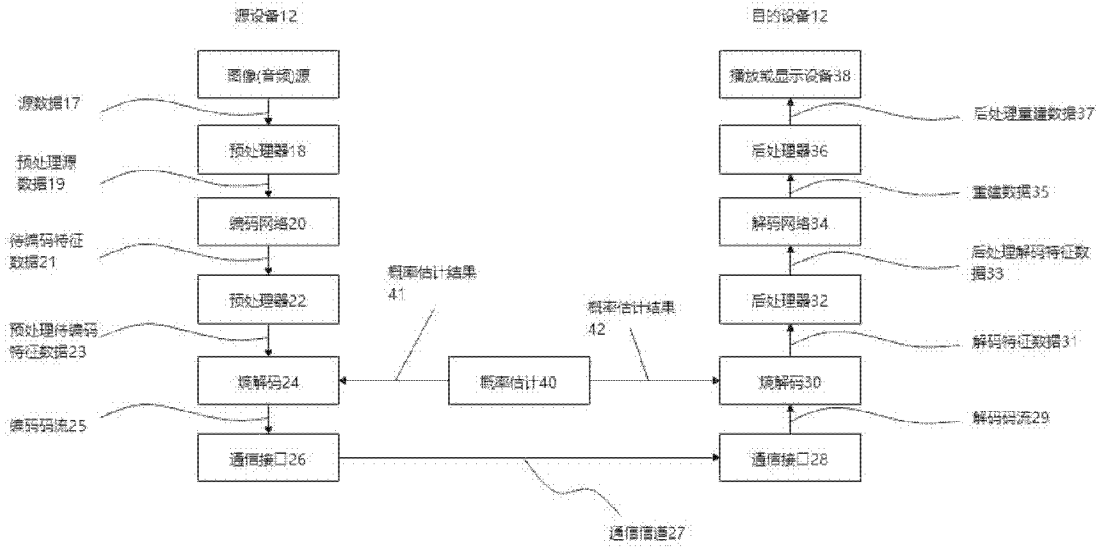


图 1A

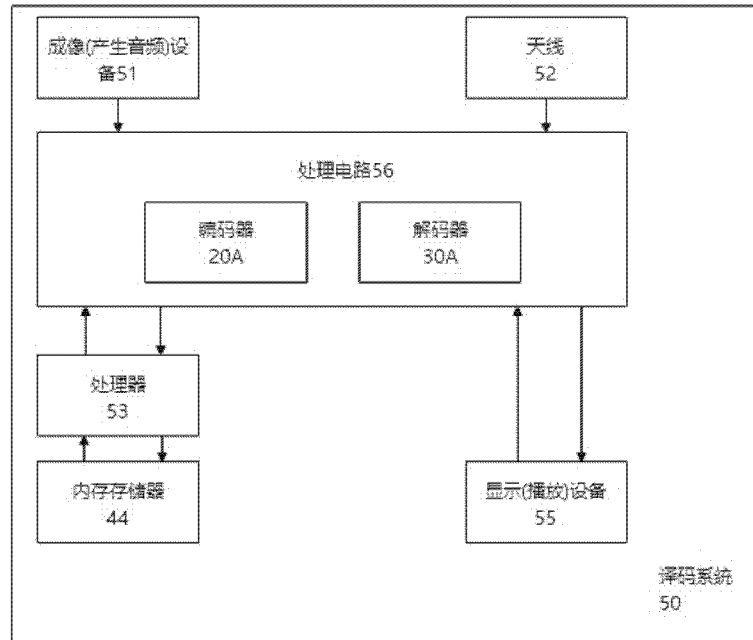


图 1B

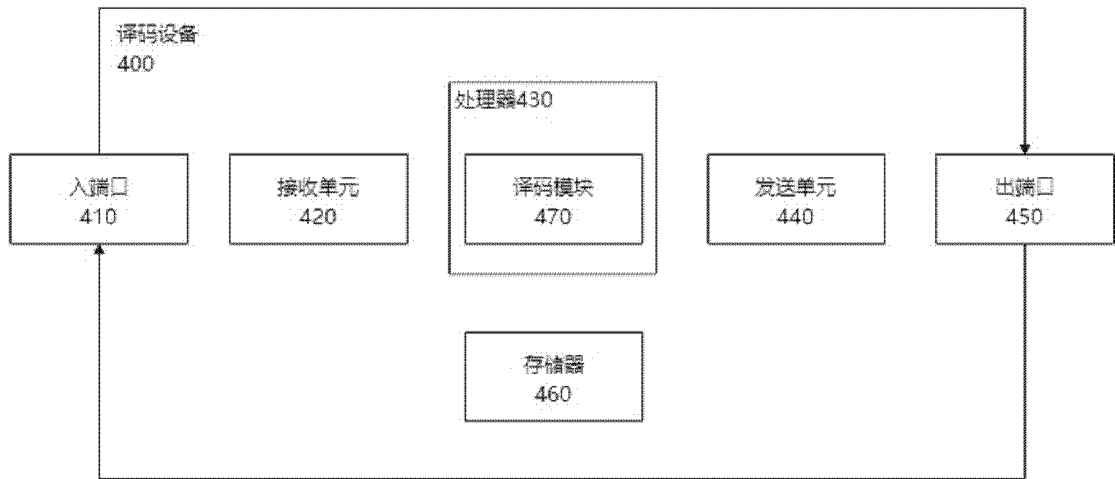


图 1C

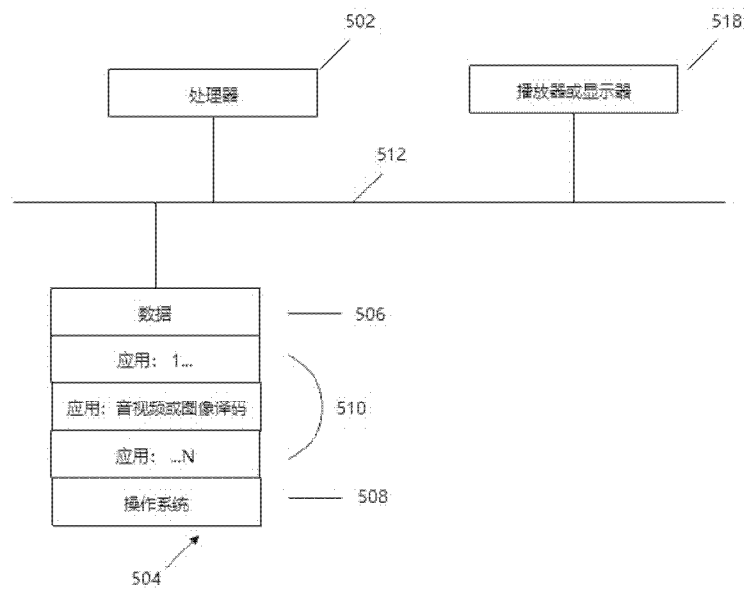


图 1D

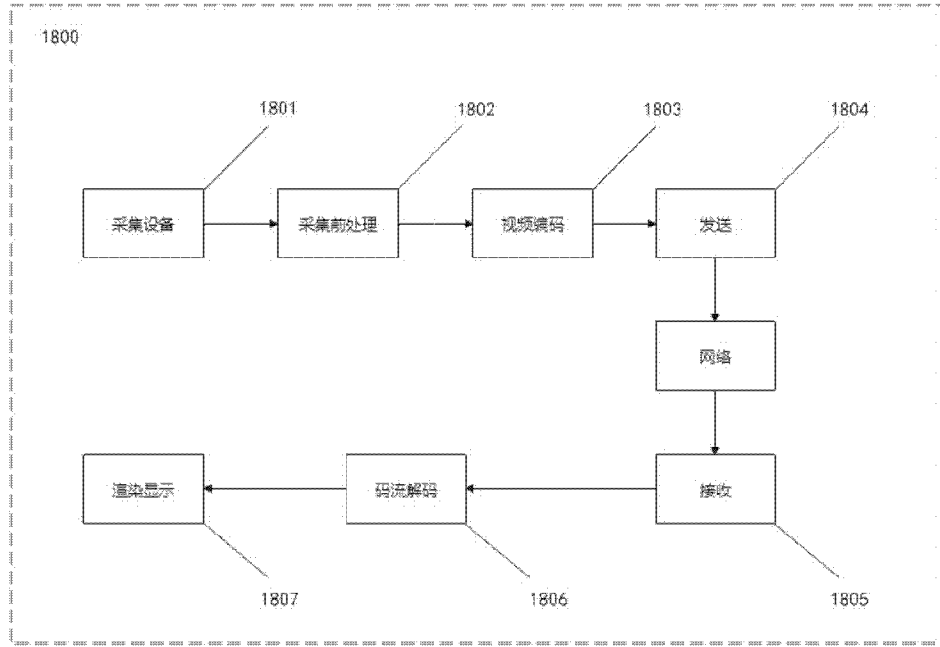


图 2A

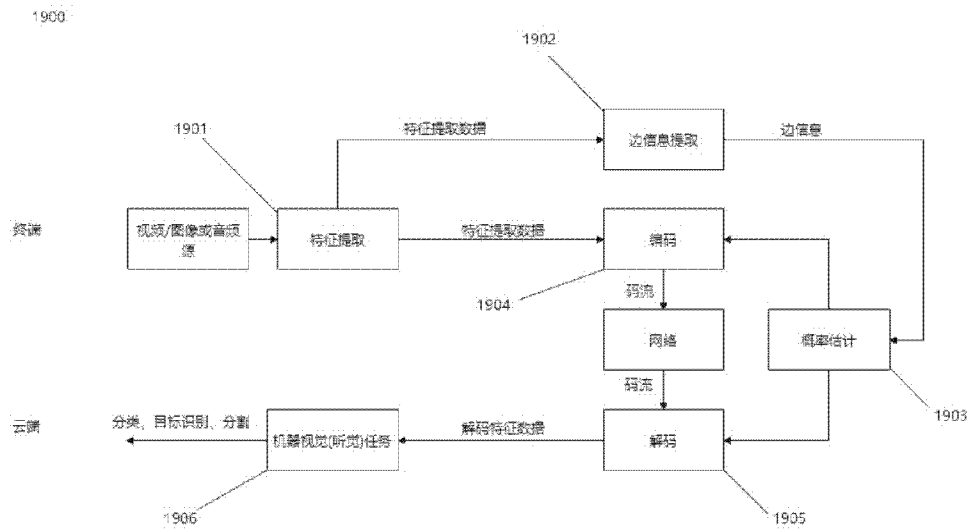


图 2B

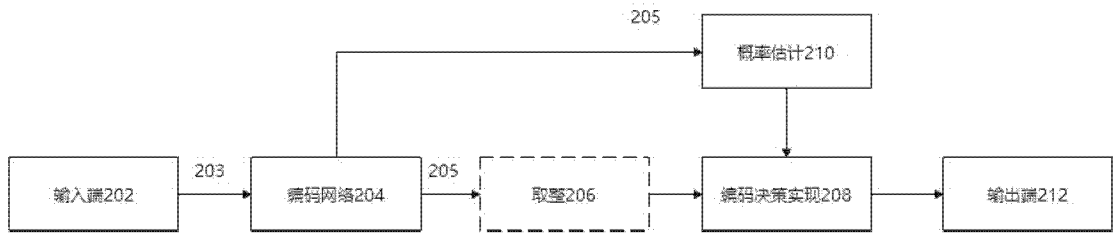


图 3A

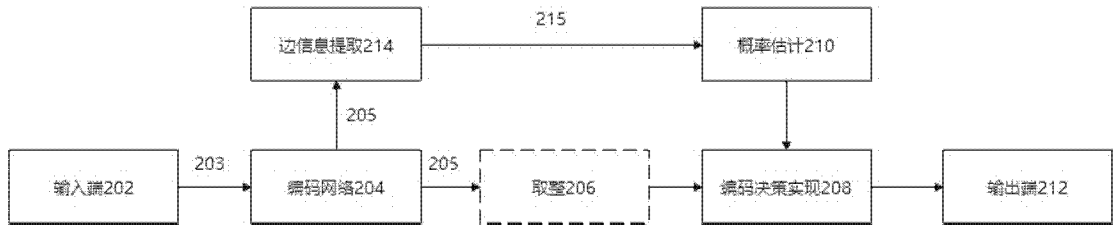


图 3B

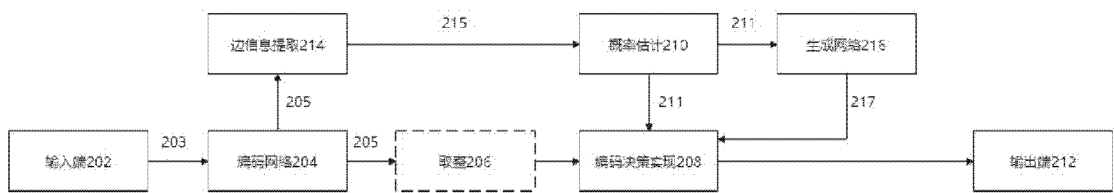


图 3C

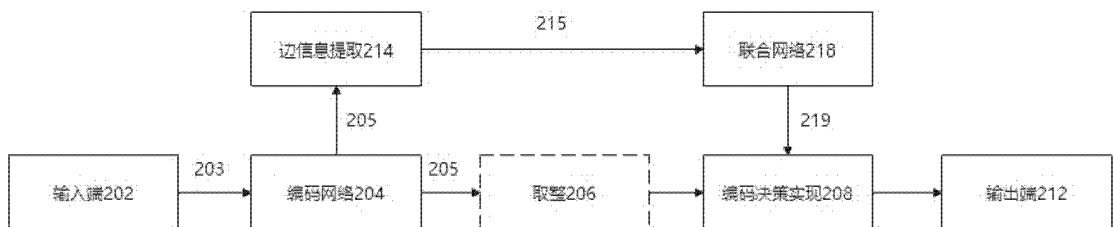


图 3D

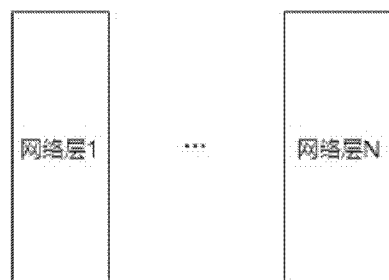


图 4A

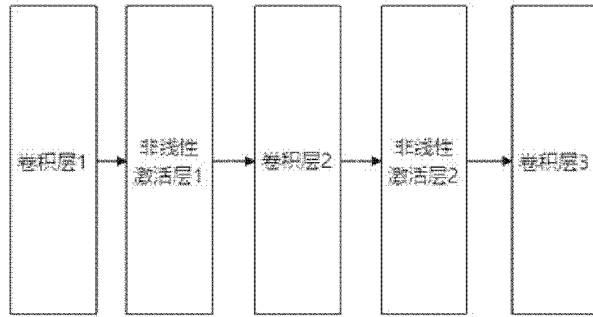


图 4B

编码决策实现26

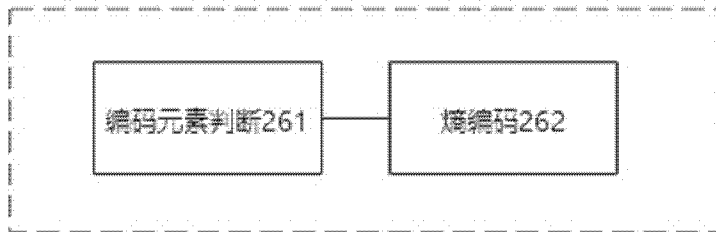


图 5

联合网络

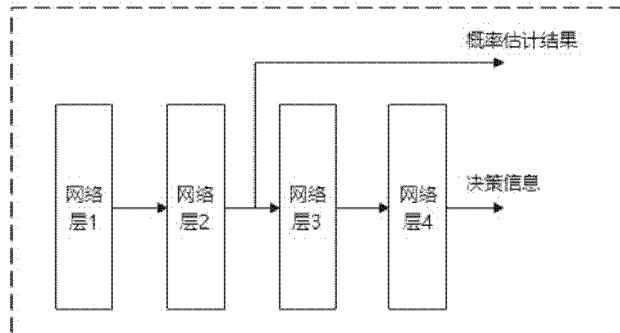


图 6

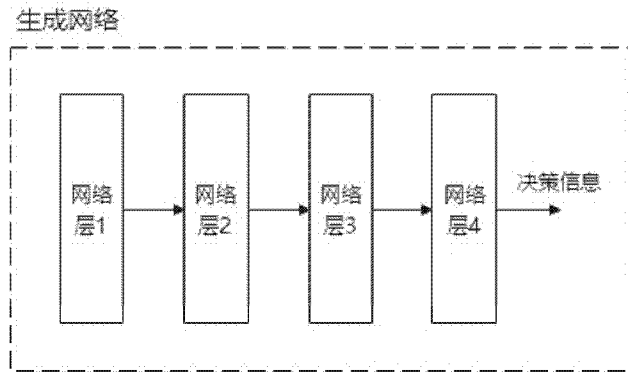


图 7

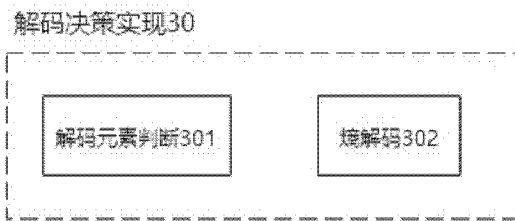


图 8

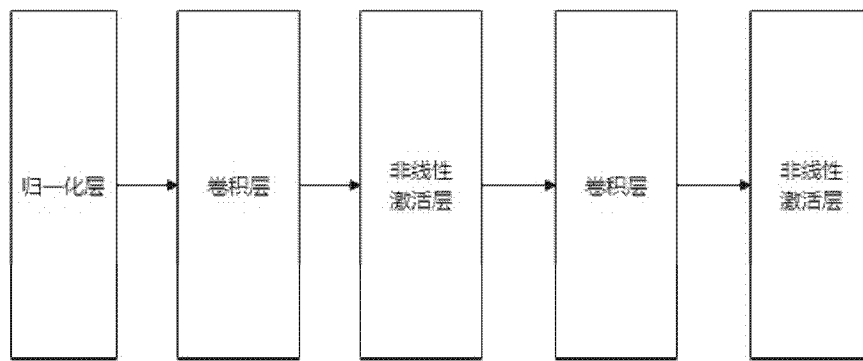


图 9

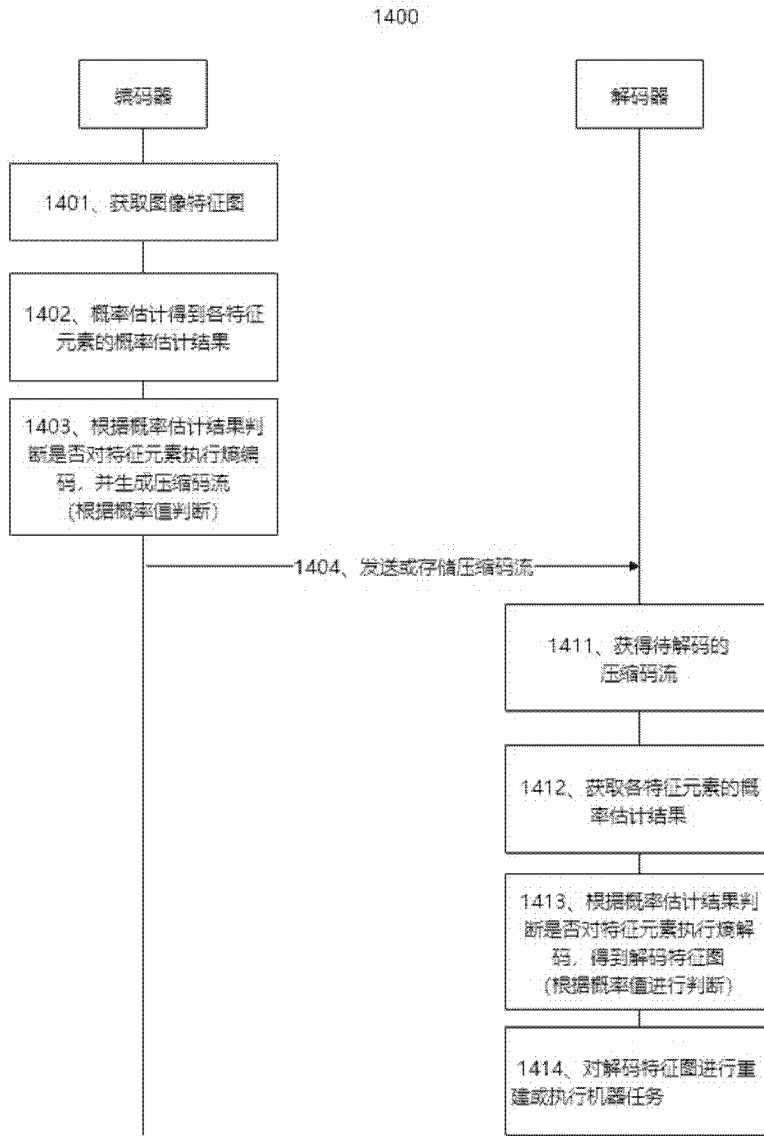


图 10A

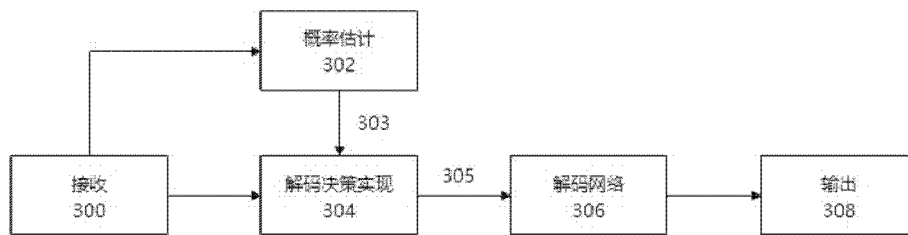


图 10B

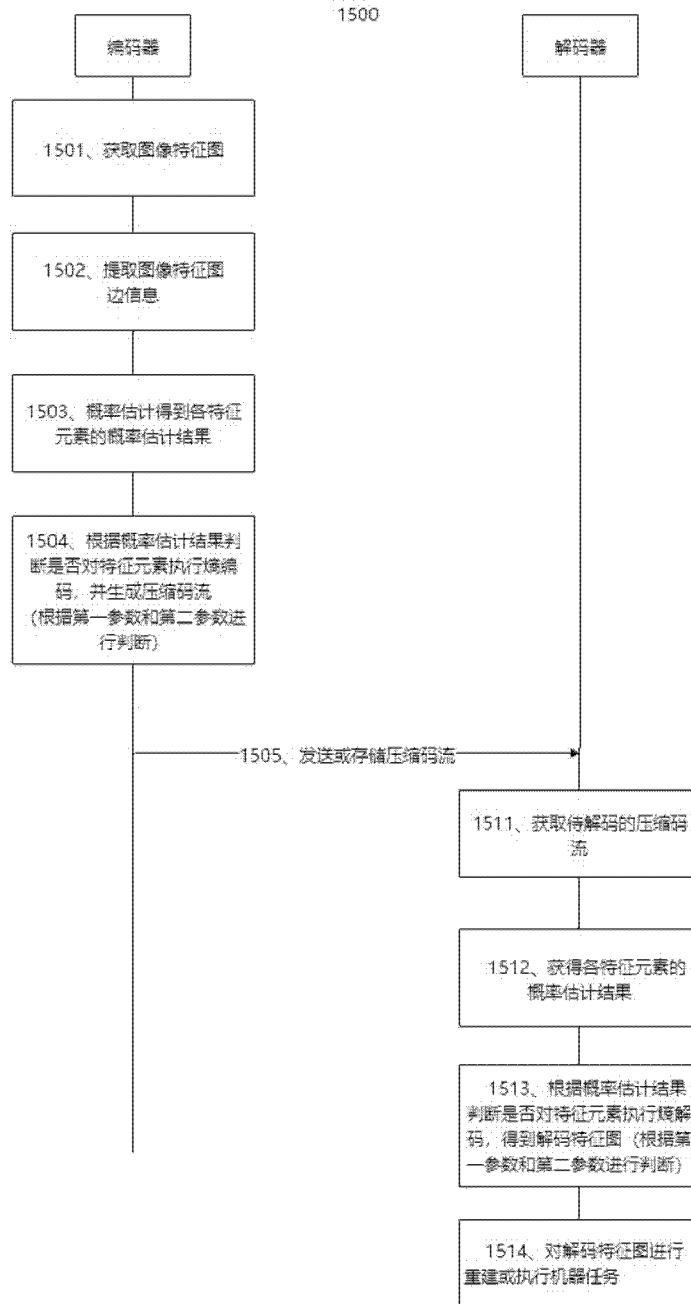


图 11A

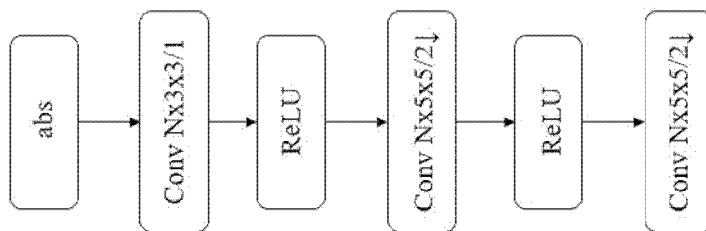


图 12

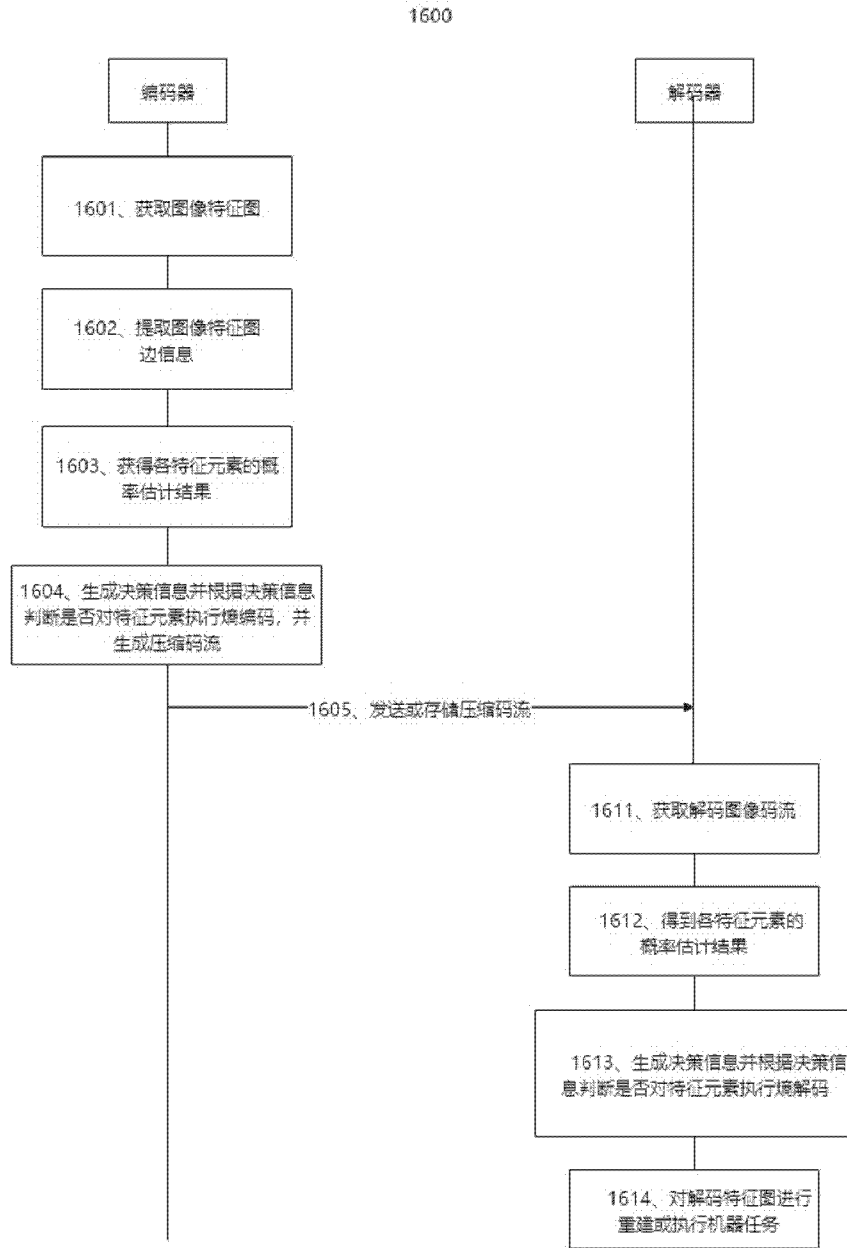


图 13A

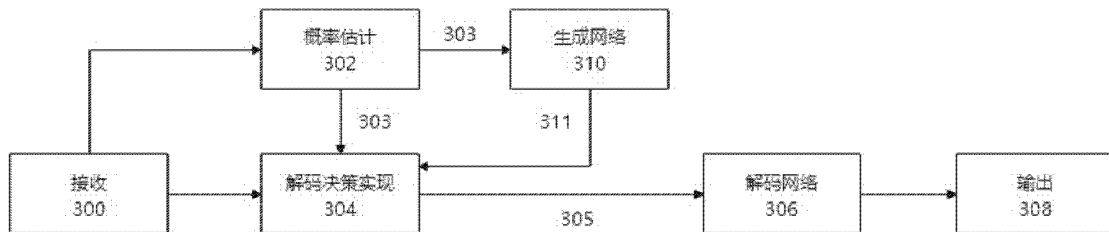


图 13B

1700

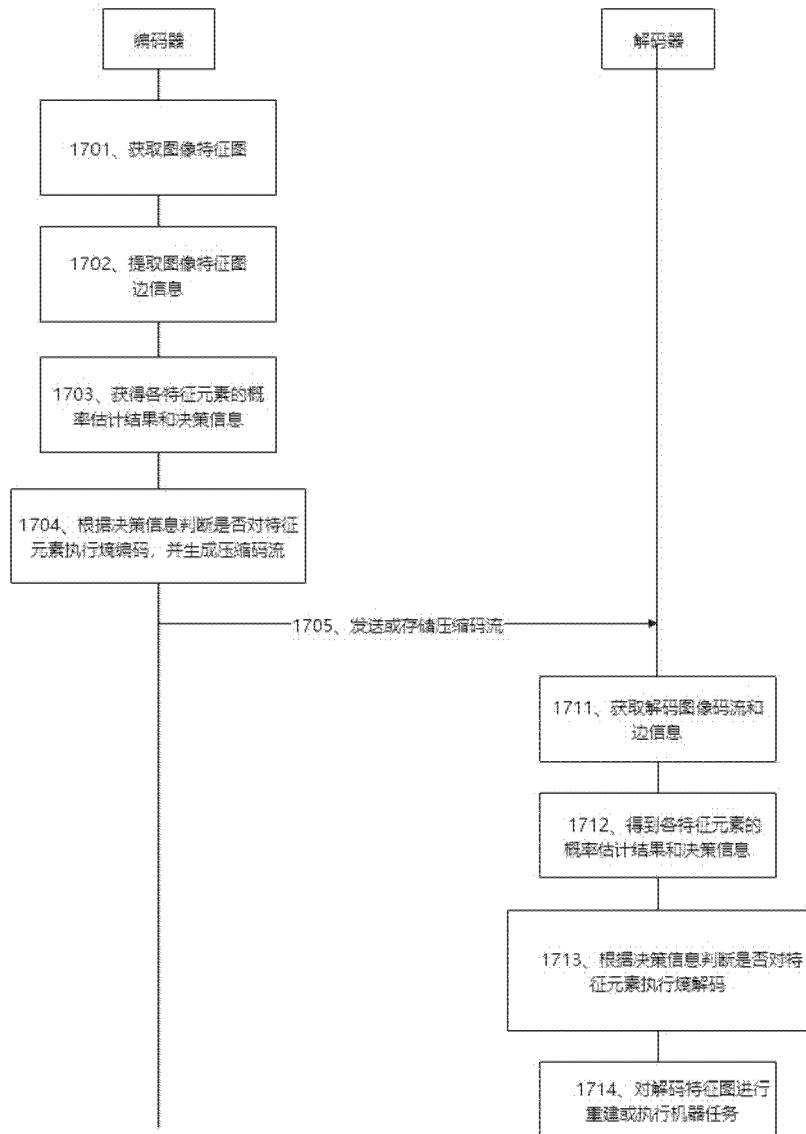


图 14

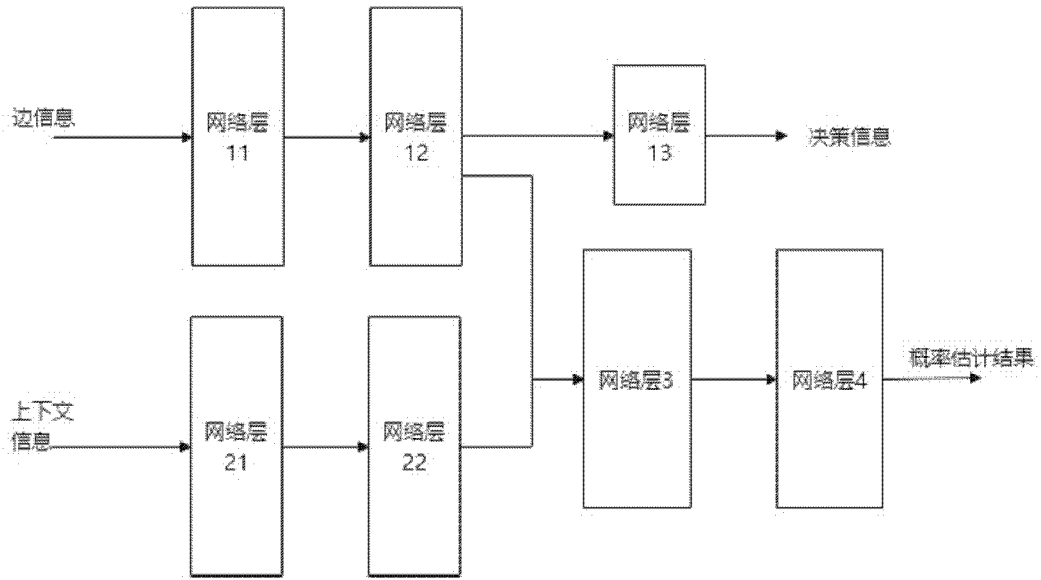


图 15

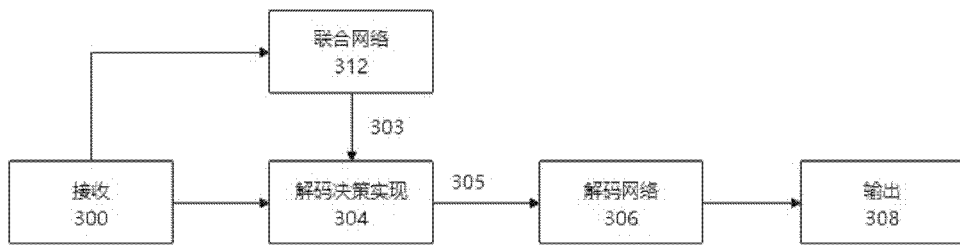


图 16

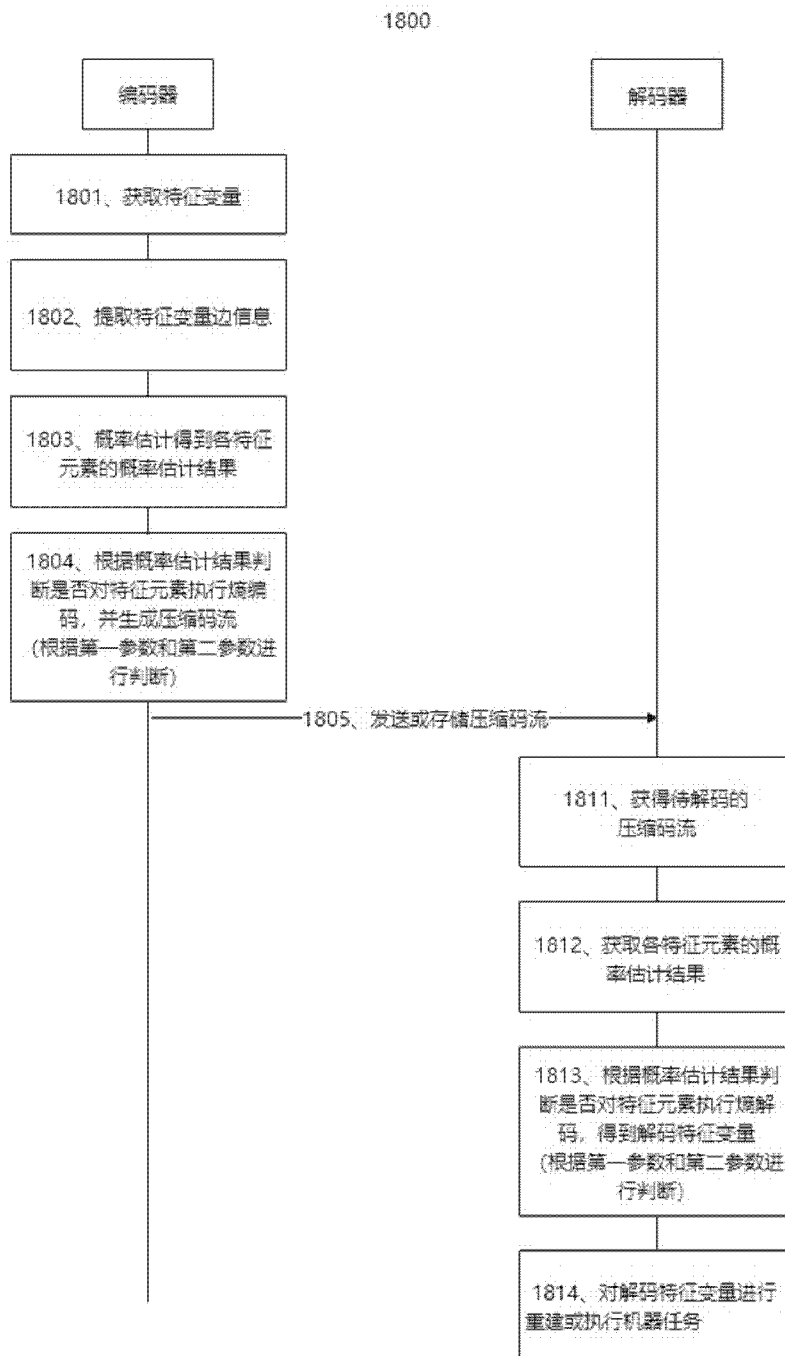


图 17

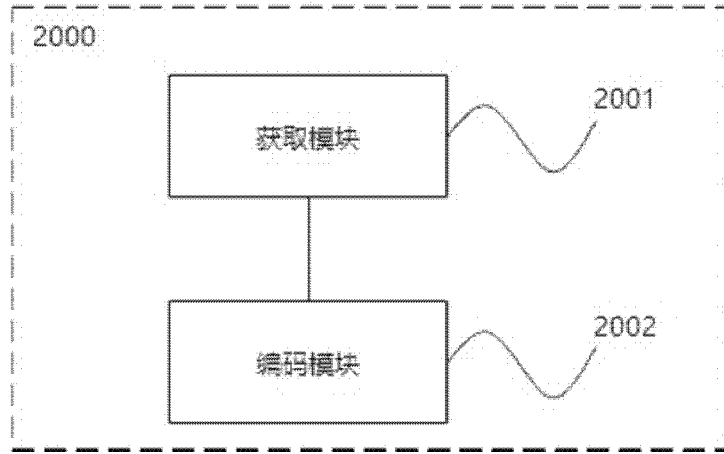


图 18

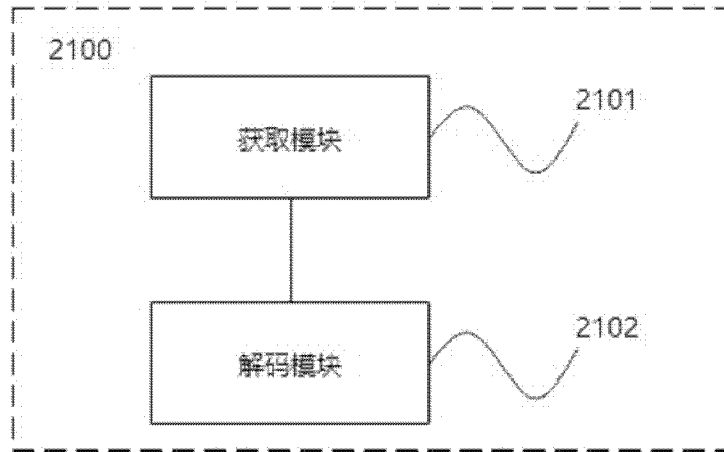


图 19

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/096510

A. CLASSIFICATION OF SUBJECT MATTER		
H04N 19/13(2014.01)i; G10L 25/30(2013.01)i		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) H04N,G10L		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) CNABS; CNTXT; WPABSC; ENTXTC; CNKI; DWPI; VEN; WOTXT; USTXT; IEEE; JNET: 特征, 概率, 熵编码, 判断, 决策, 高斯, 边信息, 神经网络, 跳过, 大于, 小于, feature, probability, entropy coding, decision, gauss, side information, neural network, skip, greater, smaller		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 111988629 A (FUJITSU LTD.) 24 November 2020 (2020-11-24) description, paragraphs 78-257, and figures 1-15	1-2, 10, 14-16, 23, 26-28, 36, 40-42, 49, 52-53, 57-58, 61-70
Y	US 10652581 B1 (GOOGLE L.L.C.) 12 May 2020 (2020-05-12) description, column 17	1-2, 10, 14-16, 23, 26-28, 36, 40-42, 49, 52-53, 57-58, 61-70
A	US 10127913 B1 (SIF CODEC LLC) 13 November 2018 (2018-11-13) entire document	1-70
A	CN 111107377 A (YAOKE INTELLIGENT TECHNOLOGY SHANGHAI CO., LTD.) 05 May 2020 (2020-05-05) entire document	1-70
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 26 July 2022		Date of mailing of the international search report 12 August 2022
Name and mailing address of the ISA/CN China National Intellectual Property Administration (ISA/ CN) No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088, China Facsimile No. (86-10)62019451		Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No. PCT/CN2022/096510

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	111988629	A	24 November 2020	None	
US	10652581	B1	12 May 2020	WO 2020176144 A1	03 September 2020
				CN 113287306 A	20 August 2021
				US 2020275130 A1	27 August 2020
				EP 3932055 A1	05 January 2022
US	10127913	B1	13 November 2018	None	
CN	111107377	A	05 May 2020	None	

国际检索报告

国际申请号

PCT/CN2022/096510

<p>A. 主题的分类</p> <p>H04N 19/13 (2014.01) i; G10L 25/30 (2013.01) i</p> <p>按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类</p>																	
<p>B. 检索领域</p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>H04N, G10L</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))</p> <p>CNABS;CNTXT;WPABSC;ENTXTC;CNKI;DWPI;VEN;WOTXT;USTXT;IEEE;JVET:特征, 概率, 熵编码, 判断, 决策, 高斯, 边信息, 神经网络, 跳过, 大于, 小于, feature, probability, entropy coding, decision, gauss, side information, neural network, skip, greater, smaller</p>																	
<p>C. 相关文件</p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>Y</td> <td>CN 111988629 A (富士通株式会社) 2020年11月24日 (2020 - 11 - 24) 说明书第78-257段以及附图1-15</td> <td>1-2、10、14-16、23、26-28、36、40-42、49、52-53、57-58、61-70</td> </tr> <tr> <td>Y</td> <td>US 10652581 B1 (GOOGLE LLC) 2020年5月12日 (2020 - 05 - 12) 说明书第17栏</td> <td>1-2、10、14-16、23、26-28、36、40-42、49、52-53、57-58、61-70</td> </tr> <tr> <td>A</td> <td>US 10127913 B1 (SIF CODEC LLC) 2018年11月13日 (2018 - 11 - 13) 全文</td> <td>1-70</td> </tr> <tr> <td>A</td> <td>CN 111107377 A (曜科智能科技上海有限公司) 2020年5月5日 (2020 - 05 - 05) 全文</td> <td>1-70</td> </tr> </tbody> </table> <p><input type="checkbox"/> 其余文件在C栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。</p> <p>* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件</p>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	Y	CN 111988629 A (富士通株式会社) 2020年11月24日 (2020 - 11 - 24) 说明书第78-257段以及附图1-15	1-2、10、14-16、23、26-28、36、40-42、49、52-53、57-58、61-70	Y	US 10652581 B1 (GOOGLE LLC) 2020年5月12日 (2020 - 05 - 12) 说明书第17栏	1-2、10、14-16、23、26-28、36、40-42、49、52-53、57-58、61-70	A	US 10127913 B1 (SIF CODEC LLC) 2018年11月13日 (2018 - 11 - 13) 全文	1-70	A	CN 111107377 A (曜科智能科技上海有限公司) 2020年5月5日 (2020 - 05 - 05) 全文	1-70
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求															
Y	CN 111988629 A (富士通株式会社) 2020年11月24日 (2020 - 11 - 24) 说明书第78-257段以及附图1-15	1-2、10、14-16、23、26-28、36、40-42、49、52-53、57-58、61-70															
Y	US 10652581 B1 (GOOGLE LLC) 2020年5月12日 (2020 - 05 - 12) 说明书第17栏	1-2、10、14-16、23、26-28、36、40-42、49、52-53、57-58、61-70															
A	US 10127913 B1 (SIF CODEC LLC) 2018年11月13日 (2018 - 11 - 13) 全文	1-70															
A	CN 111107377 A (曜科智能科技上海有限公司) 2020年5月5日 (2020 - 05 - 05) 全文	1-70															
国际检索实际完成的日期	2022年7月26日	国际检索报告邮寄日期	2022年8月12日														
ISA/CN的名称和邮寄地址	中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451	授权官员	史亦澍 电话号码 010-62411573														

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2022/096510

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	111988629	A	2020年11月24日	无			
US	10652581	B1	2020年5月12日	WO	2020176144	A1	2020年9月3日
				CN	113287306	A	2021年8月20日
				US	2020275130	A1	2020年8月27日
				EP	3932055	A1	2022年1月5日
US	10127913	B1	2018年11月13日	无			
CN	111107377	A	2020年5月5日	无			