

一种数据编解码方法及相关装置

5 本申请要求于 2023 年 01 月 09 日提交国家知识产权局、申请号为 202310028884.0、发明名称为“一种数据编解码方法及相关装置”的中国专利申请的优先权，和要求于 2023 年 03 月 22 日提交国家知识产权局、申请号为 202310295323.7、发明名称为“一种数据编解码方法及相关装置”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

10 本申请涉及数据编解码领域，尤其涉及数据编解码方法及其相关装置。

背景技术

15 人工智能(artificial intelligence, AI)是利用数字计算机或者数字计算机控制的机器模拟、延伸和扩展人的智能，感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及应用系统。换句话说，人工智能是计算机科学的一个分支，它企图了解智能的实质，并生产出一种新的能以人类智能相似的方式作出反应的智能机器。人工智能也就是研究各种智能机器的设计原理与实现方法，使机器具有感知、推理与决策的功能。

20 媒体压缩通常包括预测、频率变换和量化的一个或多个阶段，随后是熵编码。对应的媒体解压缩通常包括熵解码，随后是逆量化、逆频率变换和预测的一个或多个阶段。通常，熵编码通过利用输入符号中的冗余(例如，利用具有共同值的许多输入符号和具有罕见值的多个输入符号的模式)将输入符号转换成具有较低比特率的编码数据。熵解码将编码数据转换成对应于输入符号的输出符号。存在熵编码/解码的许多变型，其在压缩效率和计算复杂度方面提供不同的折衷。例如，霍夫曼编码/解码在计算上是简单的，但是对于输入符号的值的某些分布具有较差的压缩效率。另一方面，算术编码/解码通常具有好得多的压缩效率，代价是高得多的计算复杂度。

25 非对称数字系统(ANS)编码/解码潜在地提供了高的压缩效率和低的计算复杂度。然而现有基于 ANS 的编码有除法、取余操作、判断，解码有二分查找操作，均耗时较长。且编码和解码过程中使用的编码表、解码表的项目过多，在 AI 压缩等分布较多的场景上，用于编码表、解码表占用的内存消耗很大。

发明内容

30 本申请提供了一种数据编码方法以及数据解码方法，可以实现内存与时延的平衡。

35 第一方面，本申请提供了一种数据编码方法，所述方法包括：获取目标符号；通过第一映射关系，得到所述目标符号对应的第一信息和第二信息；所述第一信息和所述第二信息为根据所述目标符号对应的概率信息得到；根据所述第一信息以及所述编码器的状态信息，通过第一运算，得到第一编码数据；根据所述第二信息，通过第二运算，更新所述状态信息，得到第二编码数据；其中，所述第一编码数据和所述第二编码数据用于得到所述目标符号的编码结果，且所述第一运算和所述第二运算不包括除法运算。

本申请中，编码表中的索引信息仅包括符号，且映射的对象仅包括第一信息和第二信息，相当于降低了编码表中内存占用，且避免了消耗较大的操作，实现内存与时延的平衡。

40 在一种可能的实现中，所述第一信息或所述第二信息表示一个整数数值。

在一种可能的实现中，所述第一运算和所述第二运算不包括判断运算。

在一种可能的实现中，所述第一运算包括：对所述第一信息与所述状态信息进行的加和操作，以及对所述加和操作的加和结果进行的比特移位操作；从所述状态信息中选择部分比特作为第一编码数据，所述部分比特在所述状态信息中的位置为根据所述比特移位操作的移位结果确定。

45 在一种可能的实现中，所述第二运算包括：对所述状态信息与数值 M 进行的加和操作，以及对所述加和操作的加和结果进行的比特移位操作；其中，所述概率信息包括量化概率，所述 M 为所述量化概率的分母；对所述比特移位操作的移位结果与所述第二信息进行的加和操作。

在一种可能的实现中,所述概率信息包括概率质量函数 PMF 的数值;所述 PMF 的数值为通过如下方式构建的:依次确定多个符号中每个符号的 PMF 的数值,所述多个符号包括所述目标符号;其中,

在确定所述目标符号的 PMF 的数值时,将除去已确定的符号的 PMF 之外的剩余 PMF 的总数值与所述目标符号的概率进行乘积,并计算所述乘积的乘积结果与除去已确定的符号的剩余概率之间的比值,根据所述比值与预设的第五数值之间的关系,确定所述目标符号的 PMF。

在一种可能的实现中,所述第五数值为 1,所述关系为取较大值。

在一种可能的实现中,所述概率信息包括累积分布函数 CDF 的数值;所述 CDF 的数值为通过如下方式构建的:基于预设顺序,依次确定每个符号的 CDF 的数值,其中所述预设顺序为根据多个符号对应的 PMF 的数值之间的关系确定,所述多个符号包括所述目标符号;其中,在所述目标符号的概率信息对应的概率分布为对称分布时,所述预设顺序为沿着所述对称分布中峰值点的两侧相交错的顺序;或者,在所述目标符号的概率信息对应的概率分布为离散分布时,所述预设顺序具体为从大到小的顺序;或者,在所述目标符号的概率信息对应的概率分布为离散分布时,所述预设顺序具体为从小到大的顺序。

在一种可能的实现中,所述目标符号为对图像的像素数据进行变换得到,所述对图像的像素数据进行变换得到的数据满足预设分布,所述变换还用于得到所述像素数据的分布特征,所述分布特征包括均值或方差,所述方法还包括:根据所述分布特征,从多个映射关系中确定与所述分布特征对应的所述第一映射关系。

在一种可能的实现中,所述多个分布特征包括第一分布特征和第二分布特征,所述第一分布特征的熵大于所述第二分布特征的熵;所述第二分布特征对应的映射关系中的量化概率的分母的数值大于所述第一分布特征对应的映射关系中的量化概率的分母。

在多种分布场景下,可以采用不同的量化概率分母,来减小编码长度,在进行数据编码时,可以包括预测模型和熵编码,预测模型可以输入原数据,输出分布索引以及待编码数据(例如本申请中的目标符号)。根据分布索引,可计算出熵。对于越小熵的分布,可以用越大的量化概率分母。

在一种可能的实现中,所述第一映射关系中的所述第一信息为通过如下方式构建的:将所述目标符号的 PMF 的数值中最高位连续为 0 的比特数量与预设的第六数值的加和运算;将所述加和运算的加和结果进行比特移位操作,并将所述比特移位操作的移位结果与对所述目标符号的 PMF 的数值进行比特移位操作的移位结果进行减法操作,得到所述第一信息;或者,将所述加和运算的加和结果进行比特移位操作,并将所述比特移位操作的移位结果与对所述目标符号的 PMF 的数值进行比特移位操作的移位结果进行减法操作,将所述减法操作的操作结果与数值 M 进行加和操作,得到所述第一信息,所述概率信息包括量化概率,所述 M 为所述量化概率的分母。

本申请实施例中 PMF 和 CDF 的计算方法可以保证在内存、时延不变的前提下,减小码字长度。

在一种可能的实现中,所述第一映射关系中的所述第二信息为通过如下方式构建的:将所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值进行减法运算,得到所述第二信息;或者,将所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值进行减法运算,将所述减法操作的操作结果与数值 M 进行加和操作,得到所述第二信息,所述概率信息包括量化概率,所述 M 为所述量化概率的分母。

在一种可能的实现中,所述第一映射关系中的所述第二信息为通过如下方式构建的:在所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值的减法运算的结果大于或等于 0 时,将所述减法运算的结果作为所述第二信息;或者,在所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值的减法运算的结果小于 0 时,将所述减法操作的操作结果与数值 M 进行加和操作,得到所述第二信息,所述概率信息包括量化概率,所述 M 为所述量化概率的分母。

在一种可能的实现中,为了进一步降低内存的占用,可以将第一信息和第二信息合并后存储,编码器可以在获取到合并的数据后,基于一定的运算规则还原得到第一信息和第二信息。

在一种可能的实现中,可以通过第一映射关系,得到所述目标符号对应的一个整数数值;基于第三运算,将所述整数数值还原为所述第一信息和所述第二信息,所述第一信息和所述第二信息分别为一个整数数值。通过上述方式,可减小内存消耗;减小内存读取时间,在部分硬件设备(如服务器)上可进一步提升吞吐率。

示例性的,可以将多整数参数(其中至多一个为有符号整数)合并成单个参数。方法为:将有符号整数放在最左边,之后通过移位、及比特或操作,按各自参数的比特数依次按比特存储;读取符号时,再通过移位、比特与操作,还原出各自参数。

在一种可能的实现中,所述概率信息包括每个符号对应的 PMF 信息,所述 PMF 信息包括第一数值和
5 第二数值;所述第一数值和所述第二数值的加和为所述 PMF 的数值;所述概率信息还包括每个符号对应的 CDF 信息,所述 CDF 信息包括第三数值和第四数值;所述第三数值和第四数值为通过如下方式构建的:

基于预设顺序,根据多个符号的第一数值,通过累加依次确定每个符号的对应的第三数值;

基于所述预设顺序,根据多个符号的第二数值和第三数值,通过累加依次确定每个符号对应的第四数值。

10 在一种可能的实现中,所述第二信息包括第一子信息(也可以称之为状态加数 1)、第二子信息(也可以称之为状态阈值)和第三子信息(也可以称之为状态加数 2);

所述第一映射关系中的所述第一子信息为通过如下方式构建的:将所述目标符号的第三数值和所述目标符号的 PMF 的数值进行减法运算,得到所述第一子信息;

15 所述第一映射关系中的所述第二子信息为通过如下方式构建的:将所述目标符号的第三数值和所述目标符号的第一数值进行加法运算,将所述加法运算的操作结果与 1 进行减法操作,得到所述第二子信息;

所述第一映射关系中的所述第三子信息为通过如下方式构建的:将所述目标符号的第四数值和所述目标符号的第三数值进行减法运算,将所述减法运算的操作结果与所述目标符号的第一数值进行减法操作,得到所述第三子信息。

20

在一种可能的实现中,所述第二运算包括:

对所述状态信息与数值 M 进行的加和操作,以及对所述加和操作的加和结果进行的比特移位操作;其中,所述概率信息包括量化概率,所述 M 为所述量化概率的分母;对所述比特移位操作的移位结果与
所述第一子信息进行的加和操作;

25 在所述加和的操作结果大于所述第二子信息时,将所述加和的操作结果与所述第三子信息的加和结果作为所述第二编码数据;

在所述加和的操作结果小于所述第二子信息时,将所述加和的操作结果作为所述第二编码数据。

30 上述方式可以但不限于适用于单峰对称分布(如高斯分布,逻辑斯谛分布,拉普拉斯分布等),用于减小编码长度,但编码计算量会提升。

在一种可能的实现中,在所述数值 M 为 256 时,所述第一信息通过 10 比特存储,所述第一子信息通过 8 比特存储,所述第二子信息通过 7 比特存储,所述第三子信息通过 7 比特存储,多个符号的第一数值的加和为 128,所述第三子信息小于 128。

35 本申请还提供了一种数据编码方法,应用于基于非对称数字系统 ANS 的编码器,所述方法包括:获取第一符号,所述第一符号对应的量化概率和第二符号对应的量化概率的和为 1,且所述第一符号对应的量化概率和 1 的差异小于阈值,所述第二符号对应的量化概率和 0 的差异小于阈值;根据所述编码器的状态信息和数值 M 之间的数值大小关系满足第一条条件,将所述状态信息和所述数值 M 之间进行第一运算,得到第一编码数据,并将所述状态信息更新为 0;根据所述编码器的状态信息和数值 M 之间的数值
40 大小关系不满足所述第一条条件,将所述状态信息加 1。

在一种可能的实现中,所述方法还包括:获取所述第二符号;将所述编码器的状态信息作为第一编码数据,并将所述状态信息更新为所述数值 M 与 1 的差值。

上述实施例可以针对于共两种符号,且单一符号概率接近 1 时的编、解码方案。此实施例无需构建编码、解码表,因此无需额外内存存储表格。

45

第二方面,本申请提供了一种数据解码方法,所述方法包括:获取所述解码器的状态信息以及存储

比特, 所述状态信息与第一比特相对应, 所述第一比特为当前待解码的比特; 根据第一映射关系, 确定所述状态信息对应的第一信息、第二信息以及第三信息; 所述第一信息为所述第一比特对应的目标符号, 所述目标符号为所述第一比特的解码结果, 所述第二信息表示从所述存储比特中选择的比特位数, 所述第三信息用于和根据所述第三信息从所述存储比特中选择的比特进行加和运算, 所述加和运算结果用于更新所述状态信息, 更新后的状态信息与第二比特相对应, 所述第二比特为所述第一比特之后的待解码比特。

在一种可能的实现中, 所述方法还包括: 从所述存储比特中剔除根据所述选择的比特, 得到更新后的存储比特。

在一种可能的实现中, 所述第一信息、所述第二信息或所述第三信息表示一个整数数值。

在一种可能的实现中, 所述第一映射关系中的所述第一信息为通过如下方式构建的: 将满足 $C(x) \leq s < (C(x) + P(x))$ 的 x 作为所述第一信息; 其中, 所述 C 表示 CDF 的数值, 所述 P 表示 PMF 的数值, 所述 s 表示所述状态信息。

在一种可能的实现中, 所述第一映射关系中的所述第一信息为通过如下方式构建的:

将满足 $C1(x) \leq s < C1(x) + P1(x)$ 或 $C2(x) \leq s < C2(x) + P2(x)$ 的 x 作为所述第一信息; 其中, 所述 $C1$ 表示 CDF 对应的第三数值, 所述 $P1$ 表示 PMF 对应的第一数值, 所述 $C2$ 表示 CDF 对应的第四数值, 所述 $P2$ 表示 PMF 对应的第二数值, 所述 s 表示所述状态信息; 所述第一数值和所述第二数值的加和为所述 PMF 的数值。

在一种可能的实现中, 所述第一映射关系中的所述第二信息为通过如下方式构建的:

将 $\text{clz}(s' - C(x) + P(x)) + X$ 的数值作为所述第二信息; 其中, 所述 clz 表示取数值中最高位连续为 0 的比特数量, 所述 X 为预设的第五数值, 所述 s' 为通过如下得到:

若 $C2(x) \leq s$, 则 $s' = P1(x) + s - C2(x) + P(x)$, 否则 $s' = s - C1(x) + P(x)$ 。

在一种可能的实现中, 所述第一映射关系中的所述第三信息为通过如下方式构建的: 将 $s' - C(x) + P(x)$ 的数值进行比特移位操作, 得到所述第三信息, 所述比特移位操作的移位数量为根据所述第二信息得到。

一种可能的实现中, 所述第一映射关系中的所述第二信息为通过如下方式构建的: 将 $\text{clz}(s - C(x) + P(x)) + X$ 的数值作为所述第二信息; 其中, 所述 clz 表示取数值中最高位连续为 0 的比特数量, 所述 X 为预设的第五数值。

在一种可能的实现中, 所述第一映射关系中的所述第三信息为通过如下方式构建的:

将 $s - C(x) + P(x)$ 的数值进行比特移位操作, 得到所述第三信息, 所述比特移位操作的移位数量为根据所述第二信息得到; 或者,

将 $s - C(x) + P(x)$ 的数值进行比特移位操作, 并将所述比特移位操作的操作结果减去数值 M , 得到所述第三信息, 所述 M 为所述量化概率的分母, 所述比特移位操作的移位数量为根据所述第二信息得到。

在一种可能的实现中, 所述第一信息、所述第二信息或所述第三信息为根据概率信息得到; 所述概率信息包括概率质量函数 PMF 的数值; 所述 PMF 的数值为通过如下方式构建的: 依次确定多个符号中每个符号的 PMF 的数值, 所述多个符号包括所述目标符号; 其中, 在确定所述目标符号的 PMF 的数值时, 将除去已确定的符号的 PMF 之外的剩余 PMF 的总数值与所述目标符号的概率进行乘积, 并计算所述乘积的乘积结果与除去已确定的符号的剩余概率之间的比值, 根据所述比值与预设数值之间的关系, 确定所述目标符号的 PMF。

在一种可能的实现中, 所述预设数值为 1, 所述关系为取较大值。

在一种可能的实现中, 所述第一信息、所述第二信息或所述第三信息为根据概率信息得到; 所述概率信息包括累积分布函数 CDF 的数值; 所述 CDF 的数值为通过如下方式构建的:

基于预设顺序, 依次确定每个符号的 CDF 的数值, 其中所述预设顺序为根据多个符号对应的 PMF 的数值之间的关系确定, 所述多个符号包括所述目标符号; 其中,

在所述目标符号的概率信息对应的概率分布为对称分布时, 所述预设顺序为沿着所述对称分布中峰值点的两侧相交错顺序; 或者,

在所述目标符号的概率信息对应的概率分布为离散分布时, 所述预设顺序具体为从大到小的顺序;

或者，

在所述目标符号的概率信息对应的概率分布为离散分布时，所述预设顺序具体为从小到大的顺序。

在一种可能的实现中，所述方法还包括：

- 5 获取编码结果；所述编码结果为对图像的像素数据进行编码得到的，所述状态信息以及所述存储比特为根据所述编码结果得到；
- 对所述编码结果进行处理，所述像素数据的分布特征，所述分布特征包括均值或方差；
- 根据所述分布特征，从多个映射关系中确定与所述分布特征对应的所述第一映射关系。
- 本申请还提供了一种数据解码方法，应用于基于非对称数字系统 ANS 的解码器，所述方法包括：
- 10 获取所述解码器的状态信息，所述状态信息与第一比特相对应，所述第一比特为当前待解码的比特；
- 若所述状态信息为数值 M 与 1 的差值，则确定所述第一比特的解码结果为第二符号，并将所述状态信息更新为从存储中读取所述数值 M 位的结果；
- 若所述状态信息不为数值 M 与 1 的差值，则确定所述第一比特的解码结果为第一符号，若所述状态信息为 0，则将所述状态信息更新为所述数值 M 与 2 的差值与从存储中读出 1 个比特的值的加和结果；
- 15 若所述状态信息不为 0，则将所述状态信息更新为所述状态信息与 1 的差值；更新后的状态信息与第二比特相对应，所述第二比特为所述第一比特之后的待解码比特。

第三方面，本申请提供了一种数据编码装置，所述装置包括：

- 获取模块，用于获取目标符号；
- 20 编码模块，用于通过第一映射关系，得到所述目标符号对应的第一信息和第二信息；所述第一信息和所述第二信息为根据所述目标符号对应的概率信息得到；
- 根据所述第一信息以及所述编码器的状态信息，通过第一运算，得到第一编码数据；
- 根据所述第二信息，通过第二运算，更新所述状态信息，得到第二编码数据；其中，
- 所述第一编码数据和所述第二编码数据用于得到所述目标符号的编码结果，且所述第一运算和所述
- 25 第二运算不包括除法运算。

在一种可能的实现中，所述第一信息或所述第二信息表示一个整数数值。

在一种可能的实现中，所述第一运算和所述第二运算不包括判断运算。

- 30 在一种可能的实现中，所述第一运算包括：
- 对所述第一信息与所述状态信息进行的加和操作，以及对所述加和操作的加和结果进行的比特移位操作；
- 从所述状态信息中选择部分比特作为第一编码数据，所述部分比特在所述状态信息中的位置为根据
- 35 所述比特移位操作的移位结果确定。

在一种可能的实现中，所述第二运算包括：

- 对所述状态信息与数值 M 进行的加和操作，以及对所述加和操作的加和结果进行的比特移位操作；
- 其中，所述概率信息包括量化概率，所述 M 为所述量化概率的分母；
- 40 对所述比特移位操作的移位结果与所述第二信息进行的加和操作。

在一种可能的实现中，所述概率信息包括概率质量函数 PMF 的数值；所述 PMF 的数值为通过如下方式构建的：

- 依次确定多个符号中每个符号的 PMF 的数值，所述多个符号包括所述目标符号；其中，
- 45 在确定所述目标符号的 PMF 的数值时，将除去已确定的符号的 PMF 之外的剩余 PMF 的总数值与所述目标符号的概率进行乘积，并计算所述乘积的乘积结果与除去已确定的符号的剩余概率之间的比值，根

据所述比值与预设的第五数值之间的关系，确定所述目标符号的 PMF。

在一种可能的实现中，所述第五数值为 1，所述关系为取较大值。

- 5 在一种可能的实现中，所述概率信息包括累积分布函数 CDF 的数值；所述 CDF 的数值为通过如下方式构建的：
- 基于预设顺序，依次确定每个符号的 CDF 的数值，其中所述预设顺序为根据多个符号对应的 PMF 的数值之间的关系确定，所述多个符号包括所述目标符号；其中，
- 在所述目标符号的概率信息对应的概率分布为对称分布时，所述预设顺序为沿着所述对称分布中峰值点的两侧相交错的顺序；或者，
- 10 在所述目标符号的概率信息对应的概率分布为离散分布时，所述预设顺序具体为从大到小的顺序；或者，
- 在所述目标符号的概率信息对应的概率分布为离散分布时，所述预设顺序具体为从小到大的顺序。
- 15 在一种可能的实现中，所述目标符号为对图像的像素数据进行变换得到，所述对图像的像素数据进行变换得到的数据满足预设分布，所述变换还用于得到所述像素数据的分布特征，所述分布特征包括均值或方差，所述获取模块，还用于：
- 根据所述分布特征，从多个映射关系中确定与所述分布特征对应的所述第一映射关系。
- 20 在一种可能的实现中，所述多个分布特征包括第一分布特征和第二分布特征，所述第一分布特征的熵大于所述第二分布特征的熵；
- 所述第二分布特征对应的映射关系中的量化概率的分母的数值大于所述第一分布特征对应的映射关系中的量化概率的分母。
- 25 在一种可能的实现中，所述第一映射关系中的所述第一信息为通过如下方式构建的：
- 将所述目标符号的 PMF 的数值中最高位连续为 0 的比特数量与预设的第六数值的加和运算；
- 将所述加和运算的加和结果进行比特移位操作，并将所述比特移位操作的移位结果与对所述目标符号的 PMF 的数值进行比特移位操作的移位结果进行减法操作，得到所述第一信息；或者，
- 将所述加和运算的加和结果进行比特移位操作，并将所述比特移位操作的移位结果与对所述目标符号的 PMF 的数值进行比特移位操作的移位结果进行减法操作，将所述减法操作的操作结果与数值 M 进行加和操作，得到所述第一信息，所述概率信息包括量化概率，所述 M 为所述量化概率的分母。
- 30 在一种可能的实现中，所述第一映射关系中的所述第二信息为通过如下方式构建的：
- 将所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值进行减法运算，得到所述第二信息；或者，
- 35 者，
- 将所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值进行减法运算，将所述减法操作的操作结果与数值 M 进行加和操作，得到所述第二信息，所述概率信息包括量化概率，所述 M 为所述量化概率的分母。
- 40 在一种可能的实现中，所述第一映射关系中的所述第二信息为通过如下方式构建的：
- 在所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值的减法运算的结果大于或等于 0 时，将所述减法运算的结果作为所述第二信息；或者，
- 在所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值的减法运算的结果小于 0 时，将所述减法操作的操作结果与数值 M 进行加和操作，得到所述第二信息，所述概率信息包括量化概率，所述 M 为
- 45 所述量化概率的分母。

在一种可能的实现中，所述编码模块，具体用于：
通过第一映射关系，得到所述目标符号对应的一个整数数值；
基于第三运算，将所述整数数值还原为所述第一信息和所述第二信息，所述第一信息和所述第二信息分别为一个整数数值。

5

在一种可能的实现中，所述概率信息包括每个符号对应的 PMF 信息，所述 PMF 信息包括第一数值和第二数值；所述第一数值和所述第二数值的加和为所述 PMF 的数值；所述概率信息还包括每个符号对应的 CDF 信息，所述 CDF 信息包括第三数值和第四数值；所述第三数值和第四数值为通过如下方式构建的：

基于预设顺序，根据多个符号的第一数值，通过累加依次确定每个符号的对应的第三数值；
10 基于所述预设顺序，根据多个符号的第二数值和第三数值，通过累加依次确定每个符号对应的第四数值。

在一种可能的实现中，所述第二信息包括第一子信息、第二子信息和第三子信息；
所述第一映射关系中的所述第一子信息为通过如下方式构建的：将所述目标符号的第三数值和所述
15 目标符号的 PMF 的数值进行减法运算，得到所述第一子信息；

所述第一映射关系中的所述第二子信息为通过如下方式构建的：将所述目标符号的第三数值和所述目标符号的第一数值进行加法运算，将所述加法运算的操作结果与 1 进行减法操作，得到所述第二子信息；

所述第一映射关系中的所述第三子信息为通过如下方式构建的：将所述目标符号的第四数值和所述
20 目标符号的第三数值进行减法运算，将所述减法运算的操作结果与所述目标符号的第一数值进行减法操作，得到所述第三子信息。

在一种可能的实现中，所述第二运算包括：
对所述状态信息与数值 M 进行的加和操作，以及对所述加和操作的加和结果进行的比特移位操作；
25 其中，所述概率信息包括量化概率，所述 M 为所述量化概率的分母；对所述比特移位操作的移位结果与所述第一子信息进行的加和操作；

在所述加和的操作结果大于所述第二子信息时，将所述加和的操作结果与所述第三子信息的加和结果作为所述第二编码数据；

在所述加和的操作结果小于所述第二子信息时，将所述加和的操作结果作为所述第二编码
30 数据。

在一种可能的实现中，在所述数值 M 为 256 时，所述第一信息通过 10 比特存储，所述第一子信息通过 8 比特存储，所述第二子信息通过 7 比特存储，所述第三子信息通过 7 比特存储，多个符号的第一数值的加和为 128，所述第三子信息小于 128。

35 本申请还提供了一种数据编码装置，应用于基于非对称数字系统 ANS 的编码器，所述装置包括：
获取模块，用于获取第一符号，所述第一符号对应的量化概率和第二符号对应的量化概率的和为 1，且所述第一符号对应的量化概率和 1 的差异小于阈值，所述第二符号对应的量化概率和 0 的差异小于阈值；

编码模块，用于根据所述编码器的状态信息和数值 M 之间的数值大小关系满足第一条件，将所述状态信息和所述数值 M 之间进行第一运算，得到第一编码数据，并将所述状态信息更新为 0；
40 根据所述编码器的状态信息和数值 M 之间的数值大小关系不满足所述第一条件，将所述状态信息加 1；

在一种可能的实现中，所述获取模块，还用于：
45 获取所述第二符号；
所述编码模块，还用于将所述编码器的状态信息作为第一编码数据，并将所述状态信息更新为所述

数值 M 与 1 的差值。

第四方面，本申请提供了一种数据解码装置，所述装置包括：

获取模块，用于获取所述解码器的状态信息以及存储比特，所述状态信息与第一比特相对应，所述
5 第一比特为当前待解码的比特；

解码模块，用于根据第一映射关系，确定所述状态信息对应的第一信息、第二信息以及第三信息；
所述第一信息为所述第一比特对应的目标符号，所述目标符号为所述第一比特的解码结果，所述第二信
息表示从所述存储比特中选择的比特位数，所述第三信息用于和根据所述第三信息从所述存储比特中选
10 择的比特进行加和运算，所述加和运算结果用于更新所述状态信息，更新后的状态信息与第二比特相对
应，所述第二比特为所述第一比特之后的待解码比特。

在一种可能的实现中，所述解码模块，还用于：

从所述存储比特中剔除根据所述选择的比特，得到更新后的存储比特。

15 在一种可能的实现中，所述第一信息、所述第二信息或所述第三信息表示一个整数数值。

在一种可能的实现中，所述第一映射关系中的所述第一信息为通过如下方式构建的：

将满足 $C(x) \leq s < (C(x) + P(x))$ 的 x 作为所述第一信息；其中，所述 C 表示 CDF 的数值，所述 P 表示
PMF 的数值，所述 s 表示所述状态信息；或者，

20 将满足 $C1(x) \leq s < C1(x) + P1(x)$ 或 $C2(x) \leq s < C2(x) + P2(x)$ 的 x 作为所述第一信息；其中，所
述 $C1$ 表示 CDF 对应的第三数值，所述 $P1$ 表示 PMF 对应的第一数值，所述 $C2$ 表示 CDF 对应的第四数值，
所述 $P2$ 表示 PMF 对应的第二数值，所述 s 表示所述状态信息；所述第一数值和所述第二数值的加和为所
述 PMF 的数值。

25 在一种可能的实现中，所述第一映射关系中的所述第二信息为通过如下方式构建的：

将 $\text{clz}(s - C(x) + P(x)) + X$ 的数值作为所述第二信息；其中，所述 clz 表示取数值中最高位连续为 0
的比特数量，所述 X 为预设的第五数值；或者，

将 $s - C(x) + P(x)$ 的数值进行比特移位操作，得到所述第三信息，所述比特移位操作的移位数量为
根据所述第二信息得到；或者，

30 将 $s - C(x) + P(x)$ 的数值进行比特移位操作，并将所述比特移位操作的操作结果减去数值 M ，得到
所述第三信息，所述 M 为所述量化概率的分母，所述比特移位操作的移位数量为根据所述第二信息得到；
或者，

将 $\text{clz}(s' - C(x) + P(x)) + X$ 的数值作为所述第二信息；其中，所述 clz 表示取数值中最高位连续为 0
的比特数量，所述 X 为预设的第五数值，所述 s' 为通过如下得到：

35 若 $C2(x) \leq s$ ，则 $s' = P1(x) + s - C2(x) + P(x)$ ，否则 $s' = s - C1(x) + P(x)$ 。

在一种可能的实现中，所述第一映射关系中的所述第三信息为通过如下方式构建的：

将 $s - C(x) + P(x)$ 的数值进行比特移位操作，得到所述第三信息，所述比特移位操作的移位数量为
根据所述第二信息得到；或者，

40 将 $s - C(x) + P(x)$ 的数值进行比特移位操作，并将所述比特移位操作的操作结果减去数值 M ，得到
所述第三信息，所述 M 为所述量化概率的分母，所述比特移位操作的移位数量为根据所述第二信息得到；
或者，

将 $s' - C(x) + P(x)$ 的数值进行比特移位操作，得到所述第三信息，所述比特移位操作的移位数量为
根据所述第二信息得到。

45

在一种可能的实现中，所述第一信息、所述第二信息或所述第三信息为根据概率信息得到；所述概

率信息包括概率质量函数 PMF 的数值；所述 PMF 的数值为通过如下方式构建的：

依次确定多个符号中每个符号的 PMF 的数值，所述多个符号包括所述目标符号；其中，

在确定所述目标符号的 PMF 的数值时，将除去已确定的符号的 PMF 之外的剩余 PMF 的总数值与所述目标符号的概率进行乘积，并计算所述乘积的乘积结果与除去已确定的符号的剩余概率之间的比值，根据所述比值与预设数值之间的关系，确定所述目标符号的 PMF。

在一种可能的实现中，所述预设数值为 1，所述关系为取较大值。

在一种可能的实现中，所述第一信息、所述第二信息或所述第三信息为根据概率信息得到；所述概率信息包括累积分布函数 CDF 的数值；所述 CDF 的数值为通过如下方式构建的：

基于预设顺序，依次确定每个符号的 CDF 的数值，其中所述预设顺序为根据多个符号对应的 PMF 的数值之间的关系确定，所述多个符号包括所述目标符号；其中，

在所述目标符号的概率信息对应的概率分布为对称分布时，所述预设顺序为沿着所述对称分布中峰值点的两侧相交错的顺序；或者，

在所述目标符号的概率信息对应的概率分布为离散分布时，所述预设顺序具体为从大到小的顺序；或者，

在所述目标符号的概率信息对应的概率分布为离散分布时，所述预设顺序具体为从小到大的顺序。

在一种可能的实现中，所述获取模块，还用于：

获取编码结果；所述编码结果为对图像的像素数据进行编码得到的，所述状态信息以及所述存储比特为根据所述编码结果得到；

对所述编码结果进行处理，所述像素数据的分布特征，所述分布特征包括均值或方差；

根据所述分布特征，从多个映射关系中确定与所述分布特征对应的所述第一映射关系。

本申请还提供了一种数据解码装置，应用于基于非对称数字系统 ANS 的解码器，所述装置包括：

获取模块，用于获取所述解码器的状态信息，所述状态信息与第一比特相对应，所述第一比特为当前待解码的比特；

解码模块，用于若所述状态信息为数值 M 与 1 的差值，则确定所述第一比特的解码结果为第二符号，并将所述状态信息更新为从存储中读取所述数值 M 位的结果；

若所述状态信息不为数值 M 与 1 的差值，则确定所述第一比特的解码结果为第一符号，若所述状态信息为 0，则将所述状态信息更新为所述数值 M 与 2 的差值与从存储中读出 1 个比特的值的加和结果；若所述状态信息不为 0，则将所述状态信息更新为所述状态信息与 1 的差值；更新后的状态信息与第二比特相对应，所述第二比特为所述第一比特之后的待解码比特。

第五方面，本申请实施例提供了一种数据编码装置，可以包括存储器、处理器以及总线系统，其中，存储器用于存储程序，处理器用于执行存储器中的程序，以执行如上述第一方面任一可选的方法。

第六方面，本申请实施例提供了一种数据解码装置，可以包括存储器、处理器以及总线系统，其中，存储器用于存储程序，处理器用于执行存储器中的程序，以执行如上述第二方面任一可选的方法。

第七方面，本申请实施例提供了一种计算机可读存储介质，计算机可读存储介质中存储有计算机程序，当其在计算机上运行时，使得计算机执行上述第一方面及任一可选的方法、以及上述第二方面任一可选的方法。

第八方面，本申请实施例提供了一种计算机程序产品，包括代码，当代码被执行时，用于实现上述第一方面及任一可选的方法、以及上述第二方面任一可选的方法。

第九方面，本申请提供了一种芯片系统，该芯片系统包括处理器，用于支持执行设备或训练设备实现上述方面中所涉及的功能，例如，发送或处理上述方法中所涉及的数据；或，信息。在一种可能的设计中，芯片系统还包括存储器，所述存储器，用于保存执行设备或训练设备必要的程序指令和数据。该芯片系统，可以由芯片构成，也可以包括芯片和其他分立器件。

附图说明

- 图 1A 为用于实现本发明实施例的视频译码系统实例的框图；
图 1B 为用于实现本发明实施例的另一种视频译码系统实例的框图；
5 图 2 为用于实现本发明实施例的编码器实例结构的框图；
图 3 为用于实现本发明实施例的解码器实例结构的框图；
图 4 为用于实现本发明实施例的视频译码设备实例的框图；
图 5 为本申请实施例提供的一种数据编码方法的流程示意图；
图 6 为本申请实施例提供的一种数据编码方法的流程示意图；
10 图 7 为本申请实施例提供的一种数据编码方法的流程示意图；
图 8 为本申请实施例提供的一种数据编码方法的流程示意图；
图 9 为本申请实施例提供的一种数据解码方法的流程示意图；
图 10A 为本申请实施例提供的一种数据解码方法的流程示意图；
图 10B 为本申请实施例提供的一种数据解码方法的流程示意图；
15 图 11 为本申请实施例提供的一种数据编码装置的结构示意图；
图 12 为本申请实施例提供的一种数据解码装置的结构示意图；
图 13 为本申请实施例提供的一种执行设备的示意图；
图 14 为本申请实施例提供的一种服务器的示意图；
图 15 为本申请实施例提供的一种芯片的示意图。

具体实施方式

下面结合本发明实施例中的附图对本发明实施例进行描述。本发明的实施方式部分使用的术语仅用于对本发明的具体实施例进行解释，而非旨在限定本发明。

- 25 下面结合附图，对本申请的实施例进行描述。本领域普通技术人员可知，随着技术的发展和场景的出现，本申请实施例提供的技术方案对于类似的技术问题，同样适用。

- 本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。应该理解这样使用的术语在适当情况下可以互换，这仅仅是描述本申请的实施例中具有相同属性的对象在描述时所采用的区分方式。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，以便包含一系列单元的过程、方法、系统、产品或设备不必限于那些单元，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单元。

- 30 本发明实施例中，“至少一个”是指一个或者多个，“多个”是指两个或两个以上。“和/或”，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：单独存在 A，同时存在 A 和 B，单独存在 B 的情况，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，是指的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-b，a-c，b-c，或 a-b-c，其中 a，b，c 可以是单个，也可以是多个。

- 本申请实施例描述的技术和工具中的许多是参考媒体编码器/解码器系统来说明的，媒体编码器/解码器系统诸如是视频编码器/解码器系统、音频编码器/解码器系统、或纹理编码器/解码器系统。备选地，本申请实施例描述的技术和工具通常可以在用于在编码/解码文本数据或其它类型的数据中使用的数据编码器/解码器系统中实现。

40 本发明的实施方式部分使用的术语仅用于对本发明的具体实施例进行解释，而非旨在限定本发明。下面先对本发明实施例可能涉及的一些概念进行简单介绍。

应理解，本申请并不局限于视频编解码，还可以是音频、纹理或者是图像编解码，作为示例，下面以视频编解码为例进行说明。

- 45 视频编码通常是指处理形成视频或视频序列的图片序列。在视频编码领域，术语“图片(picture)”、“帧(frame)”或“图像(image)”可以用作同义词。本文中使用的视频编码表示视频编码或视频解码。

视频编码在源侧执行,通常包括处理(例如,通过压缩)原始视频图片以减少表示该视频图片所需的数据量,从而更高效地存储和/或传输。视频解码在目的地侧执行,通常包括相对于编码器作逆处理,以重构视频图片。实施例涉及的视频图片“编码”应理解为涉及视频序列的“编码”或“解码”。编码部分和解码部分的组合也称为编解码(编码和解码)。

- 5 视频序列包括一系列图像(image),图像被进一步划分为切片(slice),切片再被划分为块(block)。视频编码以块为单位进行编码处理,在一些新的视频编码标准中,块的概念被进一步扩展。比如,在 H. 264 标准中有宏块(macroblock, MB),宏块可进一步划分成多个可用于预测编码的预测块(partition)。在高性能视频编码(high efficiency video coding, HEVC)标准中,采用编码单元(coding unit, CU),预测单元(prediction unit, PU)和变换单元(transform unit, TU)等基本概念,从功能上划分了多种块单元,并采用全新的基于树结构进行描述。比如 CU 可以按照二叉树进行划分为更小的 CU,而更小的 CU 还可以继续划分,从而形成一种二叉树结构,CU 是对编码图像进行划分和编码的基本单元。对于 PU 和 TU 也有类似的树结构,PU 可以对应预测块,是预测编码的基本单元。对 CU 按照划分模式进一步划分成多个 PU。TU 可以对应变换块,是对预测残差进行变换的基本单元。然而,无论 CU,PU 还是 TU,本质上都属于块(或称图像块)的概念。本申请实施例中涉及图像块的概念(例如第一图像块和第二图像块),具体可以参照此处描述。

- 例如在 HEVC 中,通过使用表示为编码树的二叉树结构将 CTU 拆分为多个 CU。在 CU 层级处作出是否使用图片间(时间)或图片内(空间)预测对图片区域进行编码的决策。每个 CU 可以根据 PU 拆分类型进一步拆分为一个、两个或四个 PU。一个 PU 内应用相同的预测过程,并在 PU 基础上将相关信息传输到解码器。在通过基于 PU 拆分类型应用预测过程获取残差块之后,可以根据类似于用于 CU 的编码树的其它二叉树结构将 CU 分割成变换单元(transform unit, TU)。在视频压缩技术最新的发展中,使用二叉树和二叉树(Quad-tree and binary tree, QTBT)分割帧来分割编码块。在 QTBT 块结构中,CU 可以为正方形或矩形形状。

- 本文中,为了便于描述和理解,可将当前编码图像中待处理的图像块(简称待处理图像块)称为当前块,例如在编码中,待处理图像块指当前正在编码的块;在解码中,待处理图像块指当前正在解码的块。将参考图像中用于对当前块进行预测的已解码的图像块称为参考块,即参考块是为当前块提供参考信号的块,其中,参考信号表示图像块内的像素值。可将参考图像中为当前块提供预测信号的块为预测块,其中,预测信号表示预测块内的像素值或者采样值或者采样信号。例如,在遍历多个参考块以后,找到了最佳参考块,此最佳参考块将为当前块提供预测,此块可称为预测块。

- 无损视频编码情况下,可以重构原始视频图片,即经重构视频图片具有与原始视频图片相同的质量(假设存储或传输期间没有传输损耗或其它数据丢失)。在有损视频编码情况下,通过例如量化执行进一步压缩,来减少表示视频图片所需的数据量,而解码器侧无法完全重构视频图片,即经重构视频图片的质量相比原始视频图片的质量较低或较差。

- H. 261 的几个视频编码标准属于“有损混合型视频编解码”(即,将样本域中的空间和时间预测与变换域中用于应用量化的 2D 变换编码结合)。视频序列的每个图片通常分割成不重叠的块集合,通常在块层级上进行编码。换句话说,编码器侧通常在块(视频块)层级处理亦即编码视频,例如,通过空间(图片内)预测和时间(图片间)预测来产生预测块,从当前块(当前处理或待处理的块)减去预测块以获取残差块,在变换域变换残差块并量化残差块,以减少待传输(压缩)的数据量,而解码器侧将相对于编码器的逆处理部分应用于经编码或经压缩块,以重构用于表示的当前块。另外,编码器复制解码器处理循环,使得编码器和解码器生成相同的预测(例如帧内预测和帧间预测)和/或重构,用于处理亦即编码后续块。

- 下面描述本发明实施例所应用的系统架构。参见图 1A,图 1A 示例性地给出了本发明实施例所应用的视频编码及解码系统 10 的示意性框图。如图 1A 所示,视频编码及解码系统 10 可包括源设备 12 和目的地设备 14,源设备 12 产生经编码视频数据,因此,源设备 12 可被称为视频编码装置。目的地设备 14 可对由源设备 12 所产生的经编码的视频数据进行解码,因此,目的地设备 14 可被称为视频解码装置。源设备 12、目的地设备 14 或两个的各种实施方案可包含一或多个处理器以及耦合到所述一或多个处理器的存储器。所述存储器可包含但不限于 RAM、ROM、EEPROM、快闪存储器或可用于以可由计算机存取的指令或数据结构的形式存储所要的程序代码的任何其它媒体,如本文所描述。源设备 12 和目的地设备

14 可以包括各种装置, 包含桌上型计算机、移动计算装置、笔记型(例如, 膝上型)计算机、平板计算机、机顶盒、例如所谓的“智能”电话等电话手持机、电视机、相机、显示装置、数字媒体播放器、视频游戏控制台、车载计算机、无线通信设备或其类似者。

虽然图 1A 将源设备 12 和目的地设备 14 绘示为单独的设备, 但设备实施例也可以同时包括源设备 12 和目的地设备 14 或同时包括两者的功能性, 即源设备 12 或对应的功能性以及目的地设备 14 或对应的功能性。在此类实施例中, 可以使用相同硬件和/或软件, 或使用单独的硬件和/或软件, 或其任何组合来实施源设备 12 或对应的功能性以及目的地设备 14 或对应的功能性。

源设备 12 和目的地设备 14 之间可通过链路 13 进行通信连接, 目的地设备 14 可经由链路 13 从源设备 12 接收经编码视频数据。链路 13 可包括能够将经编码视频数据从源设备 12 移动到目的地设备 14 的一或多个媒体或装置。在一个实例中, 链路 13 可包括使得源设备 12 能够实时将经编码视频数据直接发射到目的地设备 14 的一或多个通信媒体。在此实例中, 源设备 12 可根据通信标准(例如无线通信协议)来调制经编码视频数据, 且可将经调制的视频数据发射到目的地设备 14。所述一或多个通信媒体可包含无线和/或有线通信媒体, 例如射频(RF)频谱或一或多个物理传输线。所述一或多个通信媒体可形成基于分组的网络的一部分, 基于分组的网络例如为局域网、广域网或全球网络(例如, 因特网)。所述一或多个通信媒体可包含路由器、交换机、基站或促进从源设备 12 到目的地设备 14 的通信的其它设备。

源设备 12 包括编码器 20, 另外可选地, 源设备 12 还可以包括图片源 16、图片预处理器 18、以及通信接口 22。具体实现形态中, 所述编码器 20、图片源 16、图片预处理器 18、以及通信接口 22 可能是源设备 12 中的硬件部件, 也可能是源设备 12 中的软件程序。分别描述如下:

图片源 16, 可以包括或可以为任何类别的图片捕获设备, 用于例如捕获现实世界图片, 和/或任何类别的图片或评论(对于屏幕内容编码, 屏幕上的一些文字也认为是待编码的图片或图像的一部分)生成设备, 例如, 用于生成计算机动画图片的计算机图形处理器, 或用于获取和/或提供现实世界图片、计算机动画图片(例如, 屏幕内容、虚拟现实(virtual reality, VR)图片)的任何类别设备, 和/或其任何组合(例如, 实景(augmented reality, AR)图片)。图片源 16 可以为用于捕获图片的相机或者用于存储图片的存储器, 图片源 16 还可以包括存储先前捕获或产生的图片和/或获取或接收图片的任何类别的(内部或外部)接口。当图片源 16 为相机时, 图片源 16 可例如为本地的或集成在源设备中的集成相机; 当图片源 16 为存储器时, 图片源 16 可为本地的或例如集成在源设备中的集成存储器。当所述图片源 16 包括接口时, 接口可例如为从外部视频源接收图片的外部接口, 外部视频源例如为外部图片捕获设备, 比如相机、外部存储器或外部图片生成设备, 外部图片生成设备例如为外部计算机图形处理器、计算机或服务器。接口可以为根据任何专有或标准化接口协议的任何类别的接口, 例如有线或无线接口、光接口。

其中, 图片可以视为像素点(pixel element)的二维阵列或矩阵。阵列中的像素点也可以称为采样点。阵列或图片在水平和垂直方向(或轴线)上的采样点数目定义图片的尺寸和/或分辨率。为了表示颜色, 通常采用三个颜色分量, 即图片可以表示为或包含三个采样阵列。例如在 RGB 格式或颜色空间中, 图片包括对应的红色、绿色及蓝色采样阵列。但是, 在视频编码中, 每个像素通常以亮度/色度格式或颜色空间表示, 例如对于 YUV 格式的图片, 包括 Y 指示的亮度分量(有时也可以用 L 指示)以及 U 和 V 指示的两个色度分量。亮度(luma)分量 Y 表示亮度或灰度水平强度(例如, 在灰度等级图片中两者相同), 而两个色度(chroma)分量 U 和 V 表示色度或颜色信息分量。相应地, YUV 格式的图片包括亮度采样值(Y)的亮度采样阵列, 和色度值(U 和 V)的两个色度采样阵列。RGB 格式的图片可以转换或变换为 YUV 格式, 反之亦然, 该过程也称为色彩变换或转换。如果图片是黑白的, 该图片可以只包括亮度采样阵列。本发明实施例中, 由图片源 16 传输至图片处理器的图片也可称为原始图片数据 17。

图片预处理器 18, 用于接收原始图片数据 17 并对原始图片数据 17 执行预处理, 以获取经预处理的图片 19 或经预处理的图片数据 19。例如, 图片预处理器 18 执行的预处理可以包括整修、色彩格式转换(例如, 从 RGB 格式转换为 YUV 格式)、调色或去噪。

编码器 20(或称编码器 20), 用于接收经预处理的图片数据 19, 采用相关预测模式对经预处理的图片数据 19 进行处理, 从而提供经编码图片数据 21(下文将进一步基于图 2 或图 4 或图 5 描述编码器 20 的结构细节)。在一些实施例中, 编码器 20 可以用于执行后文所描述的各个实施例, 以实现本发明所描

述的视频编码方法。

通信接口 22, 可用于接收经编码图片数据 21, 并可通过链路 13 将经编码图片数据 21 传输至目的地设备 14 或任何其它设备(如存储器), 以用于存储或直接重构, 所述其它设备可为任何用于解码或存储的设备。通信接口 22 可例如用于将经编码图片数据 21 封装成合适的格式, 例如数据包, 以在链路 13 上传输。

目的地设备 14 包括解码器 30, 另外可选地, 目的地设备 14 还可以包括通信接口 28、图片后处理器 32 和显示设备 34。分别描述如下:

通信接口 28, 可用于从源设备 12 或任何其它源接收经编码图片数据 21, 所述任何其它源例如为存储设备, 存储设备例如为经编码图片数据存储设备。通信接口 28 可以用于藉由源设备 12 和目的地设备 14 之间的链路 13 或藉由任何类别的网络传输或接收经编码图片数据 21, 链路 13 例如为直接有线或无线连接, 任何类别的网络例如为有线或无线网络或其任何组合, 或任何类别的私网和公网, 或其任何组合。通信接口 28 可以例如用于解封封装通信接口 22 所传输的数据包以获取经编码图片数据 21。

通信接口 28 和通信接口 22 都可以配置为单向通信接口或者双向通信接口, 以及可以用于例如发送和接收消息来建立连接、确认和交换任何其它与通信链路和/或例如经编码图片数据传输的数据传输有关的信息。

解码器 30(或称为解码器 30), 用于接收经编码图片数据 21 并提供经解码图片数据 31 或经解码图片 31(下文将进一步基于图 3 或图 4 或图 5 描述解码器 30 的结构细节)。在一些实施例中, 解码器 30 可以用于执行后文所描述的各个实施例, 以实现本发明所描述的视频解码方法。

图片后处理器 32, 用于对经解码图片数据 31(也称为经重构图片数据)执行后处理, 以获得经后处理图片数据 33。图片后处理器 32 执行的后处理可以包括: 色彩格式转换(例如, 从 YUV 格式转换为 RGB 格式)、调色、整修或重采样, 或任何其它处理, 还可用于将经后处理图片数据 33 传输至显示设备 34。

显示设备 34, 用于接收经后处理图片数据 33 以向例如用户或观看者显示图片。显示设备 34 可以为或可以包括任何类别的用于呈现经重构图片的显示器, 例如, 集成的或外部的显示器或监视器。例如, 显示器可以包括液晶显示器(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 LED 显示器、硅基液晶(liquid crystal on silicon, LCoS)、数字光处理器(digital light processor, DLP)或任何类别的其它显示器。

虽然, 图 1A 将源设备 12 和目的地设备 14 绘示为单独的设备, 但设备实施例也可以同时包括源设备 12 和目的地设备 14 或同时包括两者的功能性, 即源设备 12 或对应的功能性以及目的地设备 14 或对应的功能性。在此类实施例中, 可以使用相同硬件和/或软件, 或使用单独的硬件和/或软件, 或其任何组合来实施源设备 12 或对应的功能性以及目的地设备 14 或对应的功能性。

本领域技术人员基于描述明显可知, 不同单元的功能性或图 1A 所示的源设备 12 和/或目的地设备 14 的功能性的存在和(准确)划分可能根据实际设备 and 应用有所不同。源设备 12 和目的地设备 14 可以包括各种设备中的任一个, 包含任何类别的手持或静止设备, 例如, 笔记本或膝上型计算机、移动电话、智能手机、平板或平板计算机、摄像机、台式计算机、机顶盒、电视机、相机、车载设备、显示设备、数字媒体播放器、视频游戏控制台、视频流式传输设备(例如内容服务服务器或内容分发服务器)、广播接收器设备、广播发射器设备等, 并可以不使用或使用任何类别的操作系统。

编码器 20 和解码器 30 都可以实施为各种合适电路中的任一个, 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、硬件或其任何组合。如果部分地以软件实施所述技术, 则设备可将软件的指令存储于合适的非暂时性计算机可读存储介质中, 且可使用一或多个处理器以硬件执行指令从而执行本公开的技术。前述内容(包含硬件、软件、硬件与软件的组合等)中的任一者可视为一或多个处理器。

在一些情况下, 图 1A 中所示视频编码及解码系统 10 仅为示例, 本发明的技术可以适用于不必包含编码和解码设备之间的任何数据通信的视频编码设置(例如, 视频编码或视频解码)。在其它实例中, 数据可从本地存储器检索、在网络上流式传输等。视频编码设备可以对数据进行编码并且将数据存储在存储器, 和/或视频解码设备可以从存储器检索数据并且对数据进行解码。在一些实例中, 由并不彼此通

信而是仅编码数据到存储器和/或从存储器检索数据且解码数据的设备执行编码和解码。

应理解，编码器 20 可以部署在终端设备或者云侧的服务器上，解码器 30 可以部署在终端设备或者云侧的服务器上，编码器 20 和解码器 30 也可以共同部署在终端设备或者云侧的服务器上。

在一种场景中，编码器 20 和解码器 30 可以部署在终端设备上，编码器 20 可以对终端设备上的视频进行编码压缩，并基于本申请实施例提供的视频编码方法进行二次压缩（或者称之为增量存储压缩），并存储压缩后的数据，在需要播放视频时，可以对存储的压缩后的数据进行解码。

在一种场景中，编码器 20 和解码器 30 可以部署在多个终端设备上，编码器 20 可以对终端设备上的视频进行编码压缩，并基于本申请实施例提供的视频编码方法进行二次压缩（或者称之为增量存储压缩），并将压缩后的数据传输至其他终端设备，其他设备在需要播放视频时，可以对存储的压缩后的数据进行解码。

在一种场景中，编码器 20 和解码器 30 可以部署在终端设备和云侧的服务器上，编码器 20 可以对终端设备上的视频进行编码压缩，并基于本申请实施例提供的视频编码方法进行二次压缩（或者称之为增量存储压缩），并将压缩后的数据传输至服务器。

参见图 1B，图 1B 是根据一示例性实施例的包含图 2 的编码器 20 和/或图 3 的解码器 30 的视频译码系统 40 的实例的说明图。视频译码系统 40 可以实现本发明实施例的各种技术的组合。在所说明的实施方式中，视频译码系统 40 可以包含成像设备 41、编码器 20、解码器 30（和/或藉由处理单元 46 的逻辑电路实施的视频编/解码器）、天线 42、一个或多个处理器 43、一个或多个存储器 44 和/或显示设备 45。

如图 1B 所示，成像设备 41、天线 42、处理单元 46、逻辑电路、编码器 20、解码器 30、处理器 43、存储器 44 和/或显示设备 45 能够互相通信。如所论述，虽然用编码器 20 和解码器 30 绘示视频译码系统 40，但在不同实例中，视频译码系统 40 可以只包含编码器 20 或只包含解码器 30。

在一些实例中，天线 42 可以用于传输或接收视频数据的经编码编码数据。另外，在一些实例中，显示设备 45 可以用于呈现视频数据。在一些实例中，逻辑电路可以通过处理单元 46 实施。处理单元 46 可以包含专用集成电路(application-specific integrated circuit, ASIC)逻辑、图形处理器、通用处理器等。视频译码系统 40 也可以包含可选的处理器 43，该可选处理器 43 类似地可以包含专用集成电路(application-specific integrated circuit, ASIC)逻辑、图形处理器、通用处理器等。在一些实例中，逻辑电路可以通过硬件实施，如视频编码专用硬件等，处理器 43 可以通过通用软件、操作系统等实施。另外，存储器 44 可以是任何类型的存储器，例如易失性存储器（例如，静态随机存取存储器(Static Random Access Memory, SRAM)、动态随机存储器(Dynamic Random Access Memory, DRAM)等）或非易失性存储器（例如，闪存等）等。在非限制性实例中，存储器 44 可以由超速缓存内存实施。在一些实例中，逻辑电路可以访问存储器 44（例如用于实施图像缓冲器）。在其它实例中，逻辑电路和/或处理单元 46 可以包含存储器（例如，缓存等）用于实施图像缓冲器等。

在一些实例中，通过逻辑电路实施的编码器 20 可以包含（例如，通过处理单元 46 或存储器 44 实施的）图像缓冲器和（例如，通过处理单元 46 实施的）图形处理单元。图形处理单元可以通信耦合至图像缓冲器。图形处理单元可以包含通过逻辑电路实施的编码器 20，以实施参照图 2 和/或本文中所描述的任何其它编码器系统或子系统所论述的各种模块。逻辑电路可以用于执行本文所论述的各种操作。

在一些实例中，解码器 30 可以以类似方式通过逻辑电路实施，以实施参照图 3 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 30 可以包含（通过处理单元 2820 或存储器 44 实施的）图像缓冲器和（例如，通过处理单元 46 实施的）图形处理单元。图形处理单元可以通信耦合至图像缓冲器。图形处理单元可以包含通过逻辑电路实施的解码器 30，以实施参照图 3 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，天线 42 可以用于接收视频数据的经编码编码数据。如所论述，经编码编码数据可以包含本文所论述的与编码视频帧相关的数据、指示符、索引值、模式选择数据等，例如与编码分割相关的数据（例如，变换系数或经量化变换系数，（如所论述的）可选指示符，和/或定义编码分割的数据）。视频译码系统 40 还可包含耦合至天线 42 并用于解码经编码编码数据的解码器 30。显示设备 45 用于呈现视频帧。

应理解，本发明实施例中对于参考编码器 20 所描述的实例，解码器 30 可以用于执行相反过程。关

于信令语法元素，解码器 30 可以用于接收并解析这种语法元素，相应地解码相关视频数据。在一些例子中，编码器 20 可以将语法元素熵编码成经编码视频编码数据。在此类实例中，解码器 30 可以解析这种语法元素，并相应地解码相关视频数据。

参见图 2，图 2 示出用于实现本发明实施例的编码器 20 的实例的示意性/概念性框图。在图 2 的实例中，编码器 20 包括残差计算单元 204、变换处理单元 206、量化单元 208、逆量化单元 210、逆变换处理单元 212、重构单元 214、缓冲器 216、环路滤波器单元 220、经解码图片缓冲器(decoded picture buffer, DPB) 230、预测处理单元 260 和熵编码单元 270。预测处理单元 260 可以包含帧间预测单元 244、帧内预测单元 254 和模式选择单元 262。帧间预测单元 244 可以包含运动估计单元和运动补偿单元(未图示)。图 2 所示的编码器 20 也可以称为混合型视频编码器或根据混合型视频编解码器的视频编码器。

例如，残差计算单元 204、变换处理单元 206、量化单元 208、预测处理单元 260 和熵编码单元 270 形成编码器 20 的前向信号路径，而例如逆量化单元 210、逆变换处理单元 212、重构单元 214、缓冲器 216、环路滤波器 220、经解码图片缓冲器(decoded picture buffer, DPB) 230、预测处理单元 260 形成编码器的后向信号路径，其中编码器的后向信号路径对应于解码器的信号路径(参见图 3 中的解码器 30)。

编码器 20 通过例如输入 202，接收图片 201 或图片 201 的图像块 203，例如，形成视频或视频序列的图片序列中的图片。图像块 203 也可以称为当前编码块或待处理图像块，图片 201 可以称为当前图片或待编码图片(尤其是在视频编码中将当前图片与其它图片区分开时，其它图片例如同一视频序列亦即也包括当前图片的视频序列中的先前经编码和/或经解码图片)。

编码器 20 的实施例可以包括分割单元(图 2 中未绘示)，用于将图片 201 分割成多个例如图像块 203 的块，通常分割成多个不重叠的块。分割单元可以用于对视频序列中所有图片使用相同的块大小以及定义块大小的对应栅格，或用于在图片或子集或图片群组之间更改块大小，并将每个图片分割成对应的块。

在一个实例中，编码器 20 的预测处理单元 260 可以用于执行上述分割技术的任何组合。

如图片 201，图像块 203 也是或可以视为具有采样值的采样点的二维阵列或矩阵，虽然其尺寸比图片 201 小。换句话说，图像块 203 可以包括，例如，一个采样阵列(例如黑白图片 201 情况下的亮度阵列)或三个采样阵列(例如，彩色图片情况下的一个亮度阵列和两个色度阵列)或依据所应用的色彩格式的任何其它数目和/或类别的阵列。图像块 203 的水平和垂直方向(或轴线)上采样点的数目定义图像块 203 的尺寸。

如图 2 所示的编码器 20 用于逐块编码图片 201，例如，对每个图像块 203 执行编码和预测。

残差计算单元 204 用于基于图片图像块 203 和预测块 265(下文提供预测块 265 的其它细节)计算残差块 205，例如，通过逐样本(逐像素)将图片图像块 203 的样本值减去预测块 265 的样本值，以在样本域中获取残差块 205。

变换处理单元 206 用于在残差块 205 的样本值上应用例如离散余弦变换(discrete cosine transform, DCT)或离散正弦变换(discrete sine transform, DST)的变换，以在变换域中获取变换系数 207。变换系数 207 也可以称为变换残差系数，并在变换域中表示残差块 205。

变换处理单元 206 可以用于应用 DCT/DST 的整数近似值，例如为 HEVC/H. 265 指定的变换。与正交 DCT 变换相比，这种整数近似值通常由某一因子按比例缩放。为了维持经正变换和逆变换处理的残差块的范数，应用额外比例缩放因子作为变换过程的一部分。比例缩放因子通常是基于某些约束条件选择的，例如，比例缩放因子是用于移位运算的 2 的幂、变换系数的位深度、准确性和实施成本之间的平衡等。例如，在解码器 30 侧通过例如逆变换处理单元 212 为逆变换(以及在编码器 20 侧通过例如逆变换处理单元 212 为对应逆变换)指定具体比例缩放因子，以及相应地，可以在编码器 20 侧通过变换处理单元 206 为正变换指定对应比例缩放因子。

量化单元 208 用于例如通过应用标量化或向量量化来量化变换系数 207，以获取经量化变换系数 209。经量化变换系数 209 也可以称为经量化残差系数 209。量化过程可以减少与部分或全部变换系数 207 有关的位深度。例如，可在量化期间将 n 位变换系数向下舍入到 m 位变换系数，其中 n 大于 m。可通过调整量化参数(quantization parameter, QP)修改量化程度。例如，对于标量化，可以应用不同的标度来实现较细或较粗的量化。较小量化步长对应较细量化，而较大量化步长对应较粗量化。可以通过量化参数(quantization parameter, QP)指示合适的量化步长。例如，量化参数可以为合适的量化步

长的预定义集合的索引。例如，较小的量化参数可以对应精细量化(较小量化步长)，较大量化参数可以对应粗糙量化(较大量化步长)，反之亦然。量化可以包含除以量化步长以及例如通过逆量化 210 执行的对应的量化或逆量化，或者可以包含乘以量化步长。根据例如 HEVC 的一些标准的实施例可以使用量化参数来确定量化步长。一般而言，可以基于量化参数使用包含除法的等式的定点近似来计算量化步长。

5 可以引入额外比例缩放因子来进行量化和反量化，以恢复可能由于在用于量化步长和量化参数的等式的定点近似中使用的标度而修改的残差块的范数。在一个实例实施方式中，可以合并逆变换和反量化的标度。或者，可以使用自定义量化表并在例如编码数据中将其从编码器通过信号发送到解码器。量化是有损操作，其中量化步长越大，损耗越大。

逆量化单元 210 用于在经量化系数上应用量化单元 208 的逆量化，以获取经反量化系数 211，例如，

10 基于或使用与量化单元 208 相同的量化步长，应用量化单元 208 应用的量化方案的逆量化方案。经反量化系数 211 也可以称为经反量化残差系数 211，对应于变换系数 207，虽然由于量化造成的损耗通常与变换系数不相同。

逆变换处理单元 212 用于应用变换处理单元 206 应用的变换的逆变换，例如，逆离散余弦变换 (discrete cosine transform, DCT) 或逆离散正弦变换 (discrete sine transform, DST)，以在样本域中获取逆变换块 213。逆变换块 213 也可以称为逆变换经反量化块 213 或逆变换残差块 213。

15

重构单元 214(例如，求和器 214)用于将逆变换块 213(即经重构残差块 213)添加至预测块 265，以在样本域中获取经重构块 215，例如，将经重构残差块 213 的样本值与预测块 265 的样本值相加。

可选地，例如线缓冲器 216 的缓冲器单元 216(或简称“缓冲器”216)用于缓冲或存储经重构块 215 和对应的样本值，用于例如帧内预测。在其它的实施例中，编码器可以用于使用存储在缓冲器单元 216 中的未经滤波的经重构块和/或对应的样本值来进行任何类别的估计和/或预测，例如帧内预测。

20

例如，编码器 20 的实施例可以经配置以使得缓冲器单元 216 不只用于存储用于帧内预测 254 的经重构块 215，也用于环路滤波器单元 220(在图 2 中未示出)，和/或，例如使得缓冲器单元 216 和经解码图片缓冲器单元 230 形成一个缓冲器。其它实施例可以用于将经滤波块 221 和/或来自经解码图片缓冲器 230 的块或样本(图 2 中均未示出)用作帧内预测 254 的输入或基础。

环路滤波器单元 220(或简称“环路滤波器”220)用于对经重构块 215 进行滤波以获取经滤波块 221，从而顺利进行像素转变或提高视频质量。环路滤波器单元 220 旨在表示一个或多个环路滤波器，例如去块滤波器、样本自适应偏移(sample-adaptive offset, SAO)滤波器或其它滤波器，例如双边滤波器、自适应环路滤波器(adaptive loop filter, ALF)，或锐化或平滑滤波器，或协同滤波器。尽管环路滤波器单元 220 在图 2 中示出为环内滤波器，但在其它配置中，环路滤波器单元 220 可实施为环后滤波器。

25

30 经滤波块 221 也可以称为经滤波的经重构块 221。经解码图片缓冲器 230 可以在环路滤波器单元 220 对经重构编码块执行滤波操作之后存储经重构编码块。

编码器 20(对应地，环路滤波器单元 220)的实施例可以用于输出环路滤波器参数(例如，样本自适应偏移信息)，例如，直接输出或由熵编码单元 270 或任何其它熵编码单元熵编码后输出，例如使得解码器 30 可以接收并应用相同的环路滤波器参数用于解码。

熵编码单元 270 被配置为根据来自量化单元 208 的输出(例如，量化的变换系数)以及边信息进行熵编码。熵编码单元 270 可以被配置为确定表示量化的变换系数、边信息等参数。熵编码单元 270 可以被配置为基于上下文信息来预测参数的值，然后对实际值和预测值之间的差进行编码。对于表示要被编码的值的符号，熵编码单元 270 可以被配置为以各种方式执行熵编码。典型的熵编码技术包括指数哥伦布(Exponential-Golomb)编码、哥伦布赖斯(Golomb-Rice)编码、上下文自适应二进制算术编码 (“CABAC”)、差分编码、霍夫曼编码、游程长度编码、Lempel-Ziv (“LZ”) 编码、字典编码、RANS 编码、ANS 编码、TANS 编码或者 ANS 编码的其它变型、以及上述的组合。本申请实施例中，熵编码单元 270 可以执行 ANS 编码的一种或多种变型。

35

40

经解码图片缓冲器(decoded picture buffer, DPB)230 可以为存储参考图片数据供编码器 20 编码视频数据之用的参考图片存储器。DPB230 可由多种存储器设备中的任一个形成，例如动态随机存储器(dynamic random access memory, DRAM)(包含同步 DRAM(synchronous DRAM, SDRAM)、磁阻式 RAM(magnetoresistive RAM, MRAM)、电阻式 RAM(resistive RAM, RRAM)或其它类型的存储器设备。可

45

以由同一存储器设备或单独的存储器设备提供 DPB230 和缓冲器 216。在某一实例中，经解码图片缓冲器 (decoded picture buffer, DPB) 230 用于存储经滤波块 221。经解码图片缓冲器 230 可以进一步用于存储同一当前图片或例如先前经重构图片的不同图片的其它先前的经滤波块，例如先前经重构和经滤波块 221，以及可以提供完整的先前经重构亦即经解码图片 (和对应参考块和样本) 和/或部分经重构当前图片 (和对应参考块和样本)，例如用于帧间预测。在某一实例中，如果经重构块 215 无需环内滤波而得以重构，则经解码图片缓冲器 (decoded picture buffer, DPB) 230 用于存储经重构块 215。

预测处理单元 260，也称为块预测处理单元 260，用于接收或获取图像块 203 (当前图片 201 的当前图像块 203) 和经重构图片数据，例如来自缓冲器 216 的同一 (当前) 图片的参考样本和/或来自经解码图片缓冲器 230 的一个或多个先前经解码图片的参考图片数据 231，以及用于处理这类数据进行预测，即提供可以为经帧间预测块 245 或经帧内预测块 255 的预测块 265。

模式选择单元 262 可以用于选择预测模式 (例如帧内或帧间预测模式) 和/或对应的用作预测块 265 的预测块 245 或 255，以计算残差块 205 和重构经重构块 215。

模式选择单元 262 的实施例可以用于选择预测模式 (例如，从预测处理单元 260 所支持的那些预测模式中选择)，所述预测模式提供最佳匹配或者说最小残差 (最小残差意味着传输或存储中更好的压缩)，或提供最小信令开销 (最小信令开销意味着传输或存储中更好的压缩)，或同时考虑或平衡以上两者。模式选择单元 262 可以用于基于码率失真优化 (rate distortion optimization, RDO) 确定预测模式，即选择提供最小码率失真优化的预测模式，或选择相关码率失真至少满足预测模式选择标准的预测模式。

下文将详细解释编码器 20 的实例 (例如，通过预测处理单元 260) 执行的预测处理和 (例如，通过模式选择单元 262) 执行的模式选择。

如上文所述，编码器 20 用于从 (预先确定的) 预测模式集合中确定或选择最好或最优的预测模式。预测模式集合可以包括例如帧内预测模式和/或帧间预测模式。

在可能的实现中，帧内预测模式集合可以包括多种不同的帧内预测模式，例如，如 DC (或均值) 模式和平面模式的非方向性模式，或如 H. 265 中定义的方向性模式，或者可以包括 67 种不同的帧内预测模式，例如，如 DC (或均值) 模式和平面模式的非方向性模式，或如正在发展中的 H. 266 中定义的方向性模式。

在可能的实现中，帧间预测模式集合取决于可用参考图片 (即，例如前述存储在 DPB230 中的至少部分经解码图片) 和其它帧间预测参数，例如取决于是否使用整个参考图片或只使用参考图片的一部分，例如围绕当前块的区域的搜索窗区域，来搜索最佳匹配参考块，和/或例如取决于是否应用如半像素和/或四分之一像素内插的像素内插，帧间预测模式集合例如可包括先进运动矢量 (Advanced Motion Vector Prediction, AMVP) 模式和融合 (merge) 模式。具体实施中，帧间预测模式集合可包括本发明实施例所描述的基于仿射运动模型的预测模式，例如基于仿射运动模型的先进运动矢量预测模式 (Affine AMVP mode) 或者基于仿射运动模型的融合模式 (Affine Merge mode)，具体的，基于控制点的 AMVP 模式 (继承的控制点运动矢量预测方法或者构造的控制点运动矢量预测方法)，基于控制点的 merge 模式 (继承的控制点运动矢量预测方法或者构造的控制点运动矢量预测方法)；以及，高级时域运动矢量预测 (advanced temporal motion vector prediction, ATMVP) 方法，PLANAR 方法等等；或者，通过上述仿射运动模型的融合模式、ATMVP 和/或 PLANAR 方法的综合形成的子块融合模式 (Sub-block based merging mode)，等等。本发明实施例中，对待处理推图像块的帧间预测可应用于单向预测 (前向或后向)、双向预测 (前向和后向) 或多帧预测中，当应用于双向预测时，可采用双向预测块级的广义双向预测 (Generalized Bi-prediction, GBi)，或称加权预测方法，在一个实例中，帧内预测单元 254 可以用于执行下文描述的帧间预测技术的任意组合。

除了以上预测模式，本发明实施例也可以应用跳过模式和/或直接模式。

预测处理单元 260 可以进一步用于将图像块 203 分割成较小的块分区或子块，例如，通过迭代使用四叉树 (quad-tree, QT) 分割、二进制树 (binary-tree, BT) 分割或三叉树 (triple-tree, TT) 分割，或其任何组合，以及用于例如为块分区或子块中的每一个执行预测，其中模式选择包括选择分割的图像块 203 的树结构和选择应用于块分区或子块中的每一个的预测模式。

帧间预测单元 244 可以包含运动估计 (motion estimation, ME) 单元 (图 2 中未示出) 和运动补偿

(motion compensation, MC)单元(图2中未示出)。运动估计单元用于接收或获取图片图像块203(当前图片201的当前图片图像块203)和经解码图片231,或至少一个或多个先前经重构块,例如,一个或多个其它/不同先前经解码图片231的经重构块,基于所确定的帧间预测模式来进行运动估计。例如,视频序列可以包括当前图片和先前经解码图片31,或换句话说,当前图片和先前经解码图片31可以是形成视频序列的图片序列的一部分,或者形成该图片序列。

例如,编码器20可以用于从多个其它图片(参考图像)中的同一或不同图片的多个参考块中选择参考块,并向运动估计单元(图2中未示出)提供参考图片和/或提供参考块的位置(X、Y坐标)与当前块的位置之间的偏移(空间偏移)作为帧间预测参数。该偏移也称为运动向量(motion vector, MV)。

运动补偿单元用于获取帧间预测参数,并基于或使用帧间预测参数执行帧间预测来获取帧间预测块245。由运动补偿单元(图2中未示出)执行的运动补偿可以包含基于通过运动估计(可能执行对子像素精确度的内插)确定的运动/块向量取出或生成预测块(预测值)。内插滤波可从已知像素样本产生额外像素样本,从而潜在地增加可用于编码图片块的候选预测块的数目。一旦接收到用于当前图片块的PU的运动向量,运动补偿单元246可以在一个参考图片列表中定位运动向量指向的预测块。运动补偿单元246还可以生成与块和视频条带相关联的语法元素,以供解码器30在解码视频条带的图片块时使用。

具体的,上述帧间预测单元244可向熵编码单元270传输语法元素,所述语法元素例如包括帧间预测参数(比如遍历多个帧间预测模式后选择用于当前块预测的帧间预测模式的指示信息)、候选运动矢量列表的索引号,可选地还包括GBI索引号、参考帧索引等等。可能应用场景中,如果帧间预测模式只有一种,那么也可以不在语法元素中携带帧间预测参数,此时解码端30可直接使用默认的预测模式进行解码。可以理解的,帧间预测单元244可以用于执行帧间预测技术的任意组合。

帧内预测单元254用于获取,例如接收同一图片的图片块203(当前图片块)和一个或多个先前经重构块,例如经重构相邻块,以进行帧内估计。例如,编码器20可以用于从多个(预定)帧内预测模式中选择帧内预测模式。

编码器20的实施例可以用于基于优化标准选择帧内预测模式,例如基于最小残差(例如,提供最类似于当前图片块203的预测块255的帧内预测模式)或最小码率失真。

帧内预测单元254进一步用于基于如所选择的帧内预测模式的帧内预测参数确定帧内预测块255。在任何情况下,在选择用于块的帧内预测模式之后,帧内预测单元254还用于向熵编码单元270提供帧内预测参数,即提供指示所选择的用于块的帧内预测模式的信息。在一个实例中,帧内预测单元254可以用于执行帧内预测技术的任意组合。

具体的,上述帧内预测单元254可向熵编码单元270传输语法元素,所述语法元素包括帧内预测参数(比如遍历多个帧内预测模式后选择用于当前块预测的帧内预测模式的指示信息)。可能应用场景中,如果帧内预测模式只有一种,那么也可以不在语法元素中携带帧内预测参数,此时解码端30可直接使用默认的预测模式进行解码。

熵编码单元270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(context adaptive VLC, CAVLC)方案、算术编码方案、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)应用于经量化残差系数209、帧间预测参数、帧内预测参数和/或环路滤波器参数中的单个或所有上(或不应用),以获取可以通过输出272以例如经编码编码数据21的形式输出的经编码图片数据21。可以将经编码编码数据传输到解码器30,或将其存档稍后由解码器30传输或检索。熵编码单元270还可用于熵编码正被编码的当前视频条带的其它语法元素。

编码器20的其它结构变型可用于编码视频流。例如,基于非变换的编码器20可以在没有针对某些块或帧的变换处理单元206的情况下直接量化残差信号。在另一实施方式中,编码器20可具有组合成单个单元的量化单元208和逆量化单元210。

应当理解的是,编码器20的其它的结构变化可用于编码视频流。例如,对于某些图像块或者图像帧,编码器20可以直接地量化残差信号而不需要经变换处理单元206处理,相应地也不需要经逆变换

处理单元 212 处理；或者，对于某些图像块或者图像帧，编码器 20 没有产生残差数据，相应地不需要经变换处理单元 206、量化单元 208、逆量化单元 210 和逆变换处理单元 212 处理；或者，编码器 20 可以将经重构图像块作为参考块直接地进行存储而不需要经滤波器 220 处理；或者，编码器 20 中量化单元 208 和逆量化单元 210 可以合并在一起。环路滤波器 220 是可选的，以及针对无损压缩编码的情况下，变换处理单元 206、量化单元 208、逆量化单元 210 和逆变换处理单元 212 是可选的。应当理解的是，根据不同的应用场景，帧间预测单元 244 和帧内预测单元 254 可以是被选择性的启用。

参见图 3，图 3 示出用于实现本发明实施例的解码器 30 的实例的示意性/概念性框图。解码器 30 用于接收例如由编码器 20 编码的经编码图片数据（例如，经编码编码数据）21，以获取经解码图片 231。在解码过程期间，解码器 30 从编码器 20 接收视频数据，例如表示经编码视频条带的图片块的经编码视频编码数据及相关联的语法元素。

在图 3 的实例中，解码器 30 包括熵解码单元 304、逆量化单元 310、逆变换处理单元 312、重构单元 314（例如求和器 314）、缓冲器 316、环路滤波器 320、经解码图片缓冲器 330 以及预测处理单元 360。预测处理单元 360 可以包含帧间预测单元 344、帧内预测单元 354 和模式选择单元 362。在一些实例中，解码器 30 可执行大体上与参照图 2 的编码器 20 描述的编码遍次互逆的解码遍次。

熵解码单元 304 用于对经编码图片数据 21 执行熵解码，以获取例如经量化系数 309 和/或经解码的编码参数（图 3 中未示出），例如，帧间预测、帧内预测参数、环路滤波器参数和/或其它语法元素中（经解码）的任意一个或全部。熵解码单元 304 进一步用于将帧间预测参数、帧内预测参数和/或其它语法元素转发至预测处理单元 360。解码器 30 可接收视频条带层级和/或视频块层级的语法元素。

熵解码单元 304 被配置为对编码数据进行熵解码，从而产生针对参数的输出符号。熵解码单元 304 可以被配置为重建表示量化的变换系数和边信息的参数。熵解码单元 304 可以被配置为以各种方式来执行熵解码。典型的熵解码技术包括指数哥伦布解码、哥伦布赖斯 解码、上下文自适应二进制算术解码、霍夫曼解码、游程长度解码、Lempel-Ziv（“LZ”）解码、字典解码、RANS 解码和 ANS 解码、TANS 解码或者 ANS 解码的其它变型、以及上述的组合。本申请实施例中，熵解码单元 304 可以执行 ANS 解码的一种或多种变型。

逆量化单元 310 功能上可与逆量化单元 110 相同，逆变换处理单元 312 功能上可与逆变换处理单元 212 相同，重构单元 314 功能上可与重构单元 214 相同，缓冲器 316 功能上可与缓冲器 216 相同，环路滤波器 320 功能上可与环路滤波器 220 相同，经解码图片缓冲器 330 功能上可与经解码图片缓冲器 230 相同。

预测处理单元 360 可以包括帧间预测单元 344 和帧内预测单元 354，其中帧间预测单元 344 功能上可以类似于帧间预测单元 244，帧内预测单元 354 功能上可以类似于帧内预测单元 254。预测处理单元 360 通常用于执行块预测和/或从经编码数据 21 获取预测块 365，以及从例如熵解码单元 304（显式地或隐式地）接收或获取预测相关参数和/或关于所选择的预测模式的信息。

当视频条带经编码为经帧内编码（I）条带时，预测处理单元 360 的帧内预测单元 354 用于基于信号表示的帧内预测模式及来自当前帧或图片的先前经解码块的数据来产生用于当前视频条带的图片块的预测块 365。当视频帧经编码为经帧间编码（即 B 或 P）条带时，预测处理单元 360 的帧间预测单元 344（例如，运动补偿单元）用于基于运动向量及从熵解码单元 304 接收的其它语法元素生成用于当前视频条带的视频块的预测块 365。对于帧间预测，可从一个参考图片列表内的一个参考图片中产生预测块。解码器 30 可基于存储于 DPB330 中的参考图片，使用默认建构技术来建构参考帧列表：列表 0 和列表 1。

预测处理单元 360 用于通过解析运动向量和其它语法元素，确定用于当前视频条带的视频块的预测信息，并使用预测信息产生用于正经解码的当前视频块的预测块。在本发明的一实例中，预测处理单元 360 使用接收到的一些语法元素确定用于编码视频条带的视频块的预测模式（例如，帧内或帧间预测）、帧间预测条带类型（例如，B 条带、P 条带或 GPB 条带）、用于条带的参考图片列表中的一个或多个的建构信息、用于条带的每个经帧间编码视频块的运动向量、条带的每个经帧间编码视频块的帧间预测状态以及其它信息，以解码当前视频条带的视频块。在本公开的另一实例中，解码器 30 从编码数据接收的语法元素包含接收自适应参数集（adaptive parameter set, APS）、序列参数集（sequence parameter set, SPS）、图片参数集（picture parameter set, PPS）或条带标头中的一个或多个中的语法元素。

逆量化单元 310 可用于逆量化(即,反量化)在编码数据中提供且由熵解码单元 304 解码的经量化变换系数。逆量化过程可包含使用由编码器 20 针对视频条带中的每一视频块所计算的量化参数来确定应该应用的量化程度并同样确定应该应用的逆量化程度。

5 逆变换处理单元 312 用于将逆变换(例如,逆 DCT、逆整数变换或概念上类似的逆变换过程)应用于变换系数,以便在像素域中产生残差块。

重构单元 314(例如,求和器 314)用于将逆变换块 313(即经重构残差块 313)添加到预测块 365,以在样本域中获取经重构块 315,例如通过将经重构残差块 313 的样本值与预测块 365 的样本值相加。

10 环路滤波器单元 320(在编码循环期间或在编码循环之后)用于对经重构块 315 进行滤波以获取经滤波块 321,从而顺利进行像素转变或提高视频质量。在一个实例中,环路滤波器单元 320 可以用于执行下文描述的滤波技术的任意组合。环路滤波器单元 320 旨在表示一个或多个环路滤波器,例如去块滤波器、样本自适应偏移(sample-adaptive offset, SAO)滤波器或其它滤波器,例如双边滤波器、自适应环路滤波器(adaptive loop filter, ALF),或锐化或平滑滤波器,或协同滤波器。尽管环路滤波器单元 320 在图 3 中示出为环内滤波器,但在其它配置中,环路滤波器单元 320 可实施为环后滤波器。

15 随后将给定帧或图片中的经解码视频块 321 存储在存储用于后续运动补偿的参考图片的经解码图片缓冲器 330 中。

解码器 30 用于例如,藉由输出 332 输出经解码图片 31,以向用户呈现或供用户查看。

20 解码器 30 的其它变型可用于对压缩的编码数据进行解码。例如,解码器 30 可以在没有环路滤波器单元 320 的情况下生成输出视频流。例如,基于非变换的解码器 30 可以在没有针对某些块或帧的逆变换处理单元 312 的情况下直接逆量化残差信号。在另一实施方式中,解码器 30 可以具有组合成单个单元的逆量化单元 310 和逆变换处理单元 312。

在具体实施例中,解码器 30 可用于实现后文图 9 对应实施例中描述的数据解码方法。

应当理解的是,本发明的编码器 20 和解码器 30 中,针对某个环节的处理结果可以经过进一步处理后,输出到下一个环节,例如,在插值滤波、运动矢量推导或环路滤波等环节之后,对相应环节的处理结果可以进一步进行 Clip 或移位 shift 等操作。

25 参见图 4,图 4 是本发明实施例提供的视频译码设备 400(例如视频编码设备 400 或视频解码设备 400)的结构示意图。视频译码设备 400 适于实施本文所描述的实施例。在一个实施例中,视频译码设备 400 可以是视频解码器(例如如图 1A 的解码器 30)或视频编码器(例如如图 1A 的编码器 20)。在另一个实施例中,视频译码设备 400 可以是上述图 1A 的解码器 30 或图 1A 的编码器 20 中的一个或多个组件。

30 视频译码设备 400 包括:用于接收数据的入口端口 410 和接收单元(Rx)420,用于处理数据的处理器、逻辑单元或中央处理器(CPU)430,用于传输数据的发射器单元(Tx)440 和出口端口 450,以及,用于存储数据的存储器 460。视频译码设备 400 还可以包括与入口端口 410、接收器单元 420、发射器单元 440 和出口端口 450 耦合的光电转换组件和电光(E/O)组件,用于光信号或电信号的出口或入口。

35 处理器 430 通过硬件和软件实现。处理器 430 可以实现为一个或多个 CPU 芯片、核(例如,多核处理器)、FPGA、ASIC 和 DSP。处理器 430 与入口端口 410、接收器单元 420、发射器单元 440、出口端口 450 和存储器 460 通信。处理器 430 包括译码模块 470(例如编码模块 470 或解码模块 470)。编码/解码模块 470 实现本文中所公开的实施例,以实现本发明实施例所提供的色度块预测方法。例如,编码/解码模块 470 实现、处理或提供各种编码操作。因此,通过编码/解码模块 470 为视频译码设备 400 的功能提供了实质性的改进,并影响了视频译码设备 400 到不同状态的转换。或者,以存储在存储器 460 中并由处理器 430 执行的指令来实现编码/解码模块 470。

40 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,用于在选择性地执行这些程序时存储程序,并存储在程序执行过程中读取的指令和数据。存储器 460 可以是易失性和/或非易失性的,可以是只读存储器(ROM)、随机存取存储器(RAM)、随机存取存储器(ternary content-addressable memory, TCAM)和/或静态随机存取存储器(SRAM)。

45 参见图 5,图 5 是根据一示例性实施例的可用作图 1A 中的源设备 12 和目的地设备 14 中的任一个或两个的装置的简化框图。装置可以实现本发明的技术。换言之,图 5 为本发明实施例的编码设备或解码设备(简称为译码设备)的一种实现方式的示意性框图。其中,译码设备可以包括处理器 510、存储器 530

和总线系统 550。其中，处理器和存储器通过总线系统相连，该存储器用于存储指令，该处理器用于执行该存储器存储的指令。译码设备的存储器存储程序代码，且处理器可以调用存储器中存储的程序代码执行本发明描述的各种视频编码或解码方法。为避免重复，这里不再详细描述。

在本发明实施例中，该处理器 510 可以是中央处理单元 (Central Processing Unit, 简称为“CPU”)，该处理器 510 还可以是其他通用处理器、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。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。

该存储器 530 可以包括只读存储器 (ROM) 设备或者随机存取存储器 (RAM) 设备。任何其他适宜类型的存储设备也可以用作存储器 530。存储器 530 可以包括由处理器 510 使用总线 550 访问的代码和数据 531。存储器 530 可以进一步包括操作系统 533 和应用程序 535，该应用程序 535 包括允许处理器 510 执行本发明描述的视频编码或解码方法的至少一个程序。例如，应用程序 535 可以包括应用 1 至 N，其进一步包括执行在本发明描述的视频编码或解码方法视频编码或解码应用 (简称视频译码应用)。

该总线系统 550 除包括数据总线之外，还可以包括电源总线、控制总线和状态信号总线等。但是为了清楚说明起见，在图中将各种总线都标为总线系统 550。

可选的，译码设备还可以包括一个或多个输出设备，诸如显示器 570。在一个示例中，显示器 570 可以是触感显示器，其将显示器与可操作地感测触摸输入的触感单元合并。显示器 570 可以经由总线 550 连接到处理器 510。

虽然图 5 中将装置的处理器 510 和存储器 530 绘示为集成在单个单元中，但是还可以使用其它配置。处理器 510 的运行可以分布在多个可直接耦合的机器中 (每个机器具有一个或多个处理器)，或分布在本地区域或其它网络中。存储器 530 可以分布在多个机器中，例如基于网络的存储器或多个运行装置的机器中的存储器。虽然此处只绘示单个总线，但装置的总线 550 可以由多个总线形成。进一步地，从存储器 530 可以直接耦合至装置的其它组件或可以通过网络访问，并且可包括单个集成单元，例如一个存储卡，或多个单元，例如多个存储卡。因此，可以以多种配置实施装置。

参照图 6，图 6 为本申请实施例提供的一种数据编码方法的实施例示意，如图 6 示出的那样，本申请实施例提供的一种数据编码方法，包括：

601、获取目标符号；

其中，图 6 对应的数据编码方法可以应用于基于非对称数字系统 ANS 的编码器，具体的，基于 ANS 的编码器可以作为数据编码过程中的熵编码单元，为了方便描述，以下将基于非对称数字系统 ANS 的编码器简称为 ANS 编码器。

在一种可能的实现中，ANS 编码器可以接收输入的符号流 (包括目标符号)，对输入符号流进行编码，并且输出编码数据 (或者可以称之为编码结果) 作为比特流的部分。通常，输出的比特的总数小于输入的比特的总数，从而提供压缩效果。

在一种可能的实现中，ANS 编码器可以包括输入符号缓冲器，用于存储待编码的输入符号。以目标符号为例，目标符号可以为来自媒体 (例如，视频、图像、音频、图形纹理) 的针对量化的变换系数的参数、来自媒体的针对其它残差数据的参数或其它数据。

在一种可能的实现中，对输入符号流编码得到的比特 (或者可以称之为码字) 可以包括状态信息和存储比特，其中，状态信息 s 为整数，支持加减乘除、移位、与或非等运算，存储 (memory) 比特为一个先进后出的比特流，支持写入指定的比特数 (write) 和读出指定的比特数 (read) 操作。需保证在将任何合法的数字写入一个特定的比特数后，读出相同的比特数能够恢复此数字。通过情况下，编码顺序与解码顺序相反，例如反向编码，正向解码。

在 ANS 中，需要计算概率信息，概率信息可以包括多个输入符号中每个输入符号对应的量化概率、概率质量函数 (probability mass function, PMF) 和累积分布函数 (cumulative distribution function, CDF)，具体的，可以基于量化概率计算 PMF 和 CDF。

在一种可能的实现中，可以将每种符号的概率量化成分数，例如，所有量化概率的分母 (记为数值 M) 均相等，且为 2 的整数次幂 (记为 $M = 2^m$)。根据量化后的概率计算每种符号 x 的概率质量函数 PMF

(记为 $P(x)$)及累积分布函数 CDF (记为 $C(x)$)。其中 $P(x)$ 为量化概率的分子,在一种可能的实现中, $C(x)$ 为小于 x 的符号的 $P(x)$ 的加和。

在一种可能的实现中,所述多个分布特征包括第一分布特征和第二分布特征,所述第一分布特征的熵大于所述第二分布特征的熵;所述第二分布特征对应的映射关系中的量化概率的分母的数值大于所述第一分布特征对应的映射关系中的量化概率的分母。

在多分布场景下,可以采用不同的量化概率分母,来减小编码长度,在进行数据编码时,可以包括预测模型和熵编码,预测模型可以输入原数据,输出分布索引以及待编码数据(例如本申请中的目标符号)。根据分布索引,可计算出熵。对于越小熵的分布,可以用越大的量化概率分母。

在编码或者解码时,不同的分布可以采用不同的量化概率分母。具体为,更集中的分布采用更大的量化概率分母。例如:

1. 在 高斯分布类型下,方差较小的分布利用更大的量化概率分母,方差较大的用的更小的量化概率分母;
2. 在 逻辑斯谛分布类型下,散布程度较小的分布利用较大的量化概率分母,散布程度较大的分布利用较小的量化概率分母;
3. 在 拉普拉斯分布类型下,尺度参数较小的分布利用较大的量化概率分母,尺度参数较大的分布利用较小的量化概率分母。

在一种可能的实现中,所述概率信息包括概率质量函数 PMF 的数值;所述 PMF 的数值为通过如下方式构建的:依次确定多个符号中每个符号的 PMF 的数值,所述多个符号包括所述目标符号;其中,在确定所述目标符号的 PMF 的数值时,将除去已确定的符号的 PMF 之外的剩余 PMF 的总数值与所述目标符号的概率进行乘积,并计算所述乘积的乘积结果与除去已确定的符号的剩余概率之间的比值,根据所述比值与预设的第五数值之间的关系,确定所述目标符号的 PMF。

在一种可能的实现中,所述第五数值为 1,所述关系为取较大值。

示例性的,可以初始化剩余总 PMF 值为 M ,剩余总概率为 1;将取值范围内的所有符号按概率从小到大排序;对于排序后的每个符号:其 PMF 值为:剩余总 PMF 值,乘以此符号的概率,除以剩余总概率,取整后与 1 取较大者;将剩余总 PMF 值减去此符号 PMF,剩余总概率减去此符号概率。

在一种可能的实现中,所述概率信息包括累积分布函数 CDF 的数值;所述 CDF 的数值为通过如下方式构建的:

基于预设顺序,依次确定每个符号的 CDF 的数值,其中所述预设顺序为根据多个符号对应的 PMF 的数值之间的关系确定,所述多个符号包括所述目标符号;其中,在所述目标符号的概率信息对应的概率分布为对称分布时,所述预设顺序为沿着所述对称分布中峰值点的两侧相交错的顺序;或者,在所述目标符号的概率信息对应的概率分布为离散分布时,所述预设顺序具体为从大到小的顺序;或者,在所述目标符号的概率信息对应的概率分布为离散分布时,所述预设顺序具体为从小到大的顺序。

例如,针对于单峰对称分布(例如,高斯分布、拉普拉斯分布、逻辑斯谛分布等),可以将峰值点 x 的 $C(x) = 0$,从峰值点向两侧,即 $x, x-1, x+1, x-2, x+2, \dots$ 的顺序,对 PMF 的值做累加,得到每个符号对应的 CDF 值。

例如,针对于离散分布,可以将取值范围内的所有符号按 PMF 值从大到小排序,记为 $x(0), x(1), x(2), \dots$,其中, $x(0)$ 的 $C(x(0)) = 0$,按 $x(1), x(2), \dots$ 的顺序,对 PMF 的值做累加,得到每个符号对应的 CDF 值。

可选的,在进行 CDF 的计算时,中间到两边的顺序也可以为 $x, x+1, x-1, x+2, x-2, \dots$;当 PMF 值相等时,可选多种可能的方式打破平局,如索引由小到大(由大到小),随机种子等。

参照表 1,表 1 为一种概率信息的示意。

表 1

符号	概率	量化概率	PMF ($P(x)$)	CDF ($C(x)$)
A	0.7	5/8	5	0
B	0.1	1/8	1	5
C	0.1	1/8	1	6

D	0.1	1/8	1	7
---	-----	-----	---	---

本申请实施例中 PMF 和 CDF 的计算方法可以保证在内存、时延不变的前提下，减小码字长度。

在一种可能的实现中，还需要获取到编码表（也就是本申请实施例中的映射关系，例如，第一映射关系），接下来以第一映射关系为例介绍如何构建编码表。

在一种可能的实现中，所述第一映射关系可以包括符号以及对应的比特加数（可以称之为第一信息）和状态加数（可以称之为第二信息）。

以符号为目标符号为例，在一种可能的实现中，所述第一映射关系中的所述第一信息为通过如下方式构建的：将所述目标符号的 PMF 的数值中最高位连续为 0 的比特数量与预设的第六数值的加和运算；将所述加和运算的加和结果进行比特移位操作，并将所述比特移位操作的移位结果与对所述目标符号的 PMF 的数值进行比特移位操作的移位结果进行减法操作，得到所述第一信息。

在一种可能的实现中，第六数值可以为 $m - 31$ ，其中， m 为量化概率的分母上 2 的整数次幂。

在一种可能的实现中，将所述加和运算的加和结果进行比特移位操作具体可以为将所述加和运算的加和结果进行向左 m 位的比特移位操作，对所述目标符号的 PMF 的数值进行比特移位操作具体可以为对所述目标符号的 PMF 的数值进行向左 $b(x)$ 位的比特移位操作， $b(x)$ 可以为上述将所述目标符号的 PMF 的数值中最高位连续为 0 的比特数量与预设的第六数值的加和运算的加和结果。

示例性的，在构建编码表时，对于每个符号 x ，构建方法如下（方式 1）：

计算出每次编码 x ，写入存储中的基础比特数 $b(x) = \text{clz}(P(x)) + m - 31$ 。其中 clz 函数计算一个 32 位整数的最高位的连续为 0 的比特的数量。

计算比特加数的结果： $(b(x) \ll m) - (P(x) \ll b(x)) + M$ ，范围 $[0, m \times M]$ ，用 $m + \lceil \log_2(m + 1) \rceil$ 比特的无符号整数表示。

以符号为目标符号为例，在一种可能的实现中，所述第一映射关系中的所述第一信息为通过如下方式构建的：将所述加和运算的加和结果进行比特移位操作，并将所述比特移位操作的移位结果与对所述目标符号的 PMF 的数值进行比特移位操作的移位结果进行减法操作，将所述减法操作的操作结果与数值 M 进行加和操作，得到所述第一信息，所述概率信息包括量化概率，所述 M 为所述量化概率的分母。

示例性的，在构建编码表时，对于每个符号 x ，第一信息的构建方法如下（方式 2）：

“比特加数”的结果相比方式 1 由 $(b(x) \ll m) - (P(x) \ll b(x)) + M$ 变为 $(b(x) \ll m) - (P(x) \ll b(x))$ ，范围改变为 $[-M, (m - 1) \times M]$ ，用 $m + 1 + \lceil \log_2 m \rceil$ 比特的有符号整数表示；由于减少了比特或运算，降低了运算的开销和复杂度。可在增加少量内存的前提下，进一步提升编码效率。适用于对编码效率要求更高的场景。

在一种可能的实现中，所述第一映射关系中的所述第二信息为通过如下方式构建的（方式 1）：将所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值进行减法运算，得到所述第二信息。

示例性的，在构建编码表时，对于每个符号 x ，第二信息的构建方法如下：

“状态加数”的结果由 $C(x) - P(x)$ 变为 $C(x) - P(x) + M$ ，范围改变为 $[1, M \times 2 - 2]$ ，用 $m + 1$ 比特的无符号整数表示。

在一种可能的实现中，所述第一映射关系中的所述第二信息为通过如下方式构建的（方式 2）：将所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值进行减法运算，将所述减法操作的操作结果与数值 M 进行加和操作，得到所述第二信息，所述概率信息包括量化概率，所述 M 为所述量化概率的分母。

示例性的，在构建编码表时，对于每个符号 x ，第一信息的构建方法如下（方式 2）：

“状态加数”的计算结果相比方式 1 由 $C(x) - P(x)$ 变为“若 $C(x) - P(x) \geq 0$ ，则“状态加数”为 $C(x) - P(x)$ ，否则为 $C(x) - P(x) + M$ ”。此时“状态加数”的取值范围为 $[0, M - 1]$ ，可用 m 比特的无符号整数表示，多一个比特或的运算，提高了计算开销，可在增加少量内存的前提下，进一步提升编码效率。适用于对编码效率要求更高的场景。

在一种可能的实现中，所述第一映射关系中的所述第二信息为通过如下方式构建的（方式 3）：在所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值的减法运算的结果大于或等于 0 时，将所述减法运算的结果作为所述第二信息。通过上述方式，可在编码效率有少量下降的前提下，进一步减小内存消耗。适用于极限省内存场景。

在一种可能的实现中，所述第一映射关系中的所述第二信息为通过如下方式构建的：在所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值的减法运算的结果小于 0 时，将所述减法操作的操作结果与数值 M 进行加和操作，得到所述第二信息，所述概率信息包括量化概率，所述 M 为所述量化概率的分母。

5 示例性的，在构建编码表时，对于每个符号 x，第一信息的构建方法如下（方式 2）：
“状态加数”的计算结果相比方式 1 由 $C(x) - P(x)$ 变为“若 $C(x) - P(x) \geq 0$ ，则“状态加数”为 $C(x) - P(x)$ ，否则为 $C(x) - P(x) + M$ ”。此时“状态加数”的取值范围为 $[0, M - 1]$ ，可用 m 比特的无符号整数表示。

参照表 2，表 2 为基于上述方式 1 构建的第一映射关系的示意：

10

表 2

x	比特加数	状态加数
A	-2	3
B	16	12
C	16	13
D	16	14

参照表 3，表 3 为基于上述方式 2 构建的第一映射关系的示意：

表 3

x	比特加数	状态加数
A	-2	3
B	16	12
C	16	13
D	16	14

参照表 4，表 4 为基于上述方式 3 构建的第一映射关系的示意：

表 4

x	比特加数	状态加数
A	6	3
B	24	4
C	24	5
D	24	6

15 在一种可能的实现中，所述第一信息或者所述第二信息可以表示一个整数数值，也就是单一的一个整数，例如，所述第一信息可以表示一个整数数值，第二信息可以表示一个整数数值。本申请实施例中，编码表中的索引信息仅包括符号，且映射的对象仅包括第一信息和第二信息，降低了编码表中内存占用。

应理解，还可以针对于不同的概率分布配置不同的映射关系（包括第一映射关系）。

在一种可能的实现中，所述目标符号为对图像的像素数据进行变换得到，所述对图像的像素数据进行变换得到的数据满足预设分布，所述变换还用于得到所述像素数据的分布特征，所述分布特征包括均值或方差，可以根据所述分布特征，从多个映射关系中确定与所述分布特征对应的所述第一映射关系。也就是说，可以将分布特征作为索引，从多个映射关系中确定与所述分布特征对应的所述第一映射关系。

参照图 8，图 8 为一种编码过程的示意。

25 模型可以包括两个部分，一是图像数据处理模块，模型对图像数据进行变换，变换数据的分布更为简单，有利于熵编码过程。模型可以为神经网络，对图像数据处理成变换数据，并得到每个维度变换数据的分布索引。模型也可以为图像色彩变换的预处理及量化过程，如对 RGB 图像进行 YUV，离散傅里叶变换等，得到变换数据，并针对这些变换数据建立特定的分布索引。

二是预先定义若干概率分布（如不同方差的高斯分布、逻辑斯谛分布等），并建立对应的分布索引和编码表，其中编码表根据计算的离散 PMF、CDF 进行建立。

在编码过程中，对于每个维度的变换数据，可以找到对应的分布索引，由此找到对应的编码表格，使用熵编码，形成编码后的码字并存储。码字还可以存储基本的头文件，如图像的长、宽，一些必要的参数等。

在一种可能的实现中，为了进一步降低内存的占用，可以将第一信息和第二信息合并后存储，编码器可以在获取到合并的数据后，基于一定的运算规则还原得到第一信息和第二信息。

在一种可能的实现中，可以通过第一映射关系，得到所述目标符号对应的一个整数数值；基于第三运算，将所述整数数值还原为所述第一信息和所述第二信息，所述第一信息和所述第二信息分别为一个整数数值。通过上述方式，可减小内存消耗；减小内存读取时间，在部分硬件设备（如服务器）上可进一步提升吞吐率。

示例性的，可以将多整数参数（其中至多一个为有符号整数）合并成单个参数。方法为：将有符号整数放在最左边，之后通过移位、及比特或操作，按各自参数的比特数依次按比特存储；读取符号时，再通过移位、比特与操作，还原出各自参数。

参照表 7，表 7 为进行信息合并前的编码表的一个示意，表 8 为进行信息合并后的编码表的一个示意，表 9 为编码表格参数长度汇总。

表 7		
x	比特加数	状态加数
A	6	-5
B	24	4
C	24	5
D	24	6

表 8	
x	编码参数
A	-154
B	152
C	184
D	216

表 9				
编码表构建方式	合并顺序	“比特加数”长度	“状态加数”长度	合并参数总长度
方式 1	[状态加数, 比特加数]	$\geq m + \lceil \log_2(m + 1) \rceil$	$\geq m + 1$	$m \leq 9, 24/32 \text{ bits}$ $m \leq 13, 32 \text{ bits}$
方式 2	[比特加数, 状态加数]	$\geq m + 1 + \lceil \log_2(m + 1) \rceil$	$\geq m + 1$	
方式 3	不限	$\geq m + \lceil \log_2(m + 1) \rceil$	$\geq m$	$m \leq 10, 24/32 \text{ bits}$ $m \leq 14, 32 \text{ bits}$

603、根据所述第一信息以及所述编码器的状态信息，通过第一运算，得到第一编码数据。

604、根据所述第二信息，通过第二运算，更新所述状态信息，得到第二编码数据；其中，所述第一编码数据和所述第二编码数据用于得到所述目标符号的编码结果，且所述第一运算和所述第二运算不包括除法运算。

其中，第一编码数据可以为存储比特，第二编码数据可以为状态 s。

在一种可能的实现中，所述第一运算和第二运算中的至少一个可以不包括除法运算，从而降低了计

算时延。

在一种可能的实现中，所述第一运算和第二运算中的至少一个可以不包括判断运算，从而降低了计算时延。

在一种可能的实现中，所述第一运算和第二运算中的至少一个可以不包括二分查找运算，从而降低了计算时延。

在一种可能的实现中，所述第一运算包括：对所述第一信息与所述状态信息进行的加和操作，以及对所述加和操作的加和结果进行的比特移位操作；从所述状态信息中选择部分比特作为第一编码数据，所述部分比特在所述状态信息中的位置为根据所述比特移位操作的移位结果确定。

在一种可能的实现中，所述第二运算包括：对所述状态信息与数值 M 进行的加和操作，以及对所述加和操作的加和结果进行的比特移位操作；其中，所述概率信息包括量化概率，所述 M 为所述量化概率的分母；对所述比特移位操作的移位结果与所述第二信息进行的加和操作。

示例性的，在编码期间， $0 \leq s < M$ 。s 初始化为 0。

对于编码符号 x 来说：由 x 查表得到“比特加数”和“状态加数”；计算写入存储中的比特数 b ：当前状态 s 与“比特加数”的和，右移 m 位；将当前状态 s 最低的 b 位写入存储；将状态 s 更新为：当前状态 s 加上 M ，再右移 b 位后，与“状态加数”的和。

可以将 s 用 m 比特无符号整数的二进制表示，与存储的比特合并在一起，得到编码结果，也就是码字。

参照表 5，表 5 为利用上述方式将 AAABA 编码为 00101110（其中，前三位 001 表示状态为 1，01110 为存储比特）。

表 5

s	Memory	x	写入存储比特数
0		A	0
3		B	3
5	011	A	1
1	0111	A	0
4	0111	A	1
1	01110		

参照图 7，图 7 为通过上述方式 1 进行编码表构建以及通过上述介绍的第一运算和第二运算进行的编码过程的流程示意。

在一种可能的实现中，在通过上述方式 2 进行编码表构建时，所述第二运算可以包括：在所述状态信息初始化为所述 M 的情况下，对所述状态信息进行的比特移位操作；对所述比特移位操作的移位结果与所述第二信息进行的加和操作；对所述加和操作的加和结果与所述 M 进行减法运算。

示例性的，在整个编码期间， $M \leq s < M \times 2$ ；编码时， s 初始化为 M ；

可以将“当前状态 s 加上 M ，再右移 b 位后，与状态加数的和”变为“当前状态 s 右移 b 位后，与状态加数的和”；将 $s - M$ 用 m 比特无符号整数的二进制表示，与存储的比特合并在一起。

参照表 6，表 6 为利用上述方式将 AAABA 编码为 00101110（其中，前三位 001 表示状态为 1，01110 为存储比特）。

表 6

s	Memory	x	写入存储比特数
8		A	0
11		B	3
13	011	A	1
9	0111	A	0

12	0111	A	1
9	01110		

本申请实施例提供了一种数据编码方法,应用于基于非对称数字系统ANS的编码器,所述方法包括:获取目标符号;通过第一映射关系,得到所述目标符号对应的第一信息和第二信息;所述第一信息和所述第二信息为根据所述目标符号对应的概率信息得到;根据所述第一信息以及所述编码器的状态信息,通过第一运算,得到第一编码数据;根据所述第二信息,通过第二运算,更新所述状态信息,得到第二编码数据;其中,所述第一编码数据和所述第二编码数据用于得到所述目标符号的编码结果,且所述第一运算和所述第二运算不包括除法运算。本申请中,编码表中的索引信息仅包括符号,且映射的对象仅包括第一信息和第二信息,相当于降低了编码表中内存占用,且避免了消耗较大的操作,实现内存与时延的平衡。

10 在一种可能的实现中,所述第二信息包括第一子信息(也可以称之为状态加数1)、第二子信息(也可以称之为状态阈值)和第三子信息(也可以称之为状态加数2);

所述第一映射关系中的所述第一子信息为通过如下方式构建的:将所述目标符号的第三数值和所述目标符号的PMF的数值进行减法运算,得到所述第一子信息;

15 所述第一映射关系中的所述第二子信息为通过如下方式构建的:将所述目标符号的第三数值和所述目标符号的第一数值进行加法运算,将所述加法运算的操作结果与1进行减法操作,得到所述第二子信息;

所述第一映射关系中的所述第三子信息为通过如下方式构建的:将所述目标符号的第四数值和所述目标符号的第三数值进行减法运算,将所述减法运算的操作结果与所述目标符号的第一数值进行减法操作,得到所述第三子信息。

20

在一种可能的实现中,所述第二运算包括:

对所述状态信息与数值M进行的加和操作,以及对所述加和操作的加和结果进行的比特移位操作;其中,所述概率信息包括量化概率,所述M为所述量化概率的分母;对所述比特移位操作的移位结果与所述第一子信息进行的加和操作;

25 在所述加和操作的加和结果大于所述第二子信息时,将所述加和操作的加和结果与所述第三子信息的加和结果作为所述第二编码数据;

在所述加和操作的加和结果小于所述第二子信息时,将所述加和操作的加和结果作为所述第二编码数据。

30 上述方式可以但不限于适用于单峰对称分布(如高斯分布,逻辑斯谛分布,拉普拉斯分布等),用于减小编码长度,但编码计算量会提升。

1. 将每个符号的PMF值分成两部分 $P(x) = P1(x) + P2(x)$

2. 计算 $C1(x)$ 为小于 x 的符号的 $P1(x)$ 的加和; $C2(x)$ 为小于 x 的符号的 $P2(x)$ 的加和,加上所有 $P1(x)$ 的值。

如下表格示出了一个具体的示意:

符号	概率	量化概率	PMF (P(x))	CDF (C(x))	P1(x)	P2(x)	C1(x)	C2(x)
A	0.7	5/8	5	0	3	2	0	4
B	0.1	1/8	1	5	0	1	3	6
C	0.1	1/8	1	6	1	0	3	7
D	0.1	1/8	1	7	0	1	4	7

35 在构建编码表时,对于每个符号 x ,构建方法的一个示意如下:

1) 计算出每次编码 x ,写入存储中的基础比特数 $b(x) = \text{clz}(P(x)) + m - 31$ 。其中 clz 函数计算一个32位整数的最高位的连续为0的比特的数量。

2) 计算比特加数的结果: $(b(x) \ll m) - (P(x) \ll b(x)) + M$, 范围 $[2, m \times M]$, 用 $m + \lceil \log_2(m + 1) \rceil$

比特的无符号整数表示； z

- 3) 计算状态加数 1 的结果： $C1(x) - P(x)$
- 4) 计算状态阈值的结果： $C1(x) + P1(x) - 1$
- 5) 计算状态加数 2 的结果： $C2(x) - C1(x) - P1(x)$

5 如下表格示出了一个编码表格的示意：

x	比特加数	状态加数 1	状态阈值	状态加数 2
A	6	-5	2	1
B	24	2	2	3
C	24	2	3	3
D	24	3	3	3

在编码时，码字可以由状态 s 及存储 memory 组成。

示例性的，在整个编码期间， $0 \leq s < M$ 。 s 初始化为 0。

编码符号 x ：

- 1) 由 x 查表得到“比特加数”，“状态加数 1”，“状态阈值”和“状态加数 2”；
- 10 2) 计算写入存储中的比特数 b ：当前状态 s 与“比特加数”的和，右移 m 位；
- 3) 将当前状态 s 最低的 b 位写入存储；
- 4) 将状态 s 更新为：当前状态 s 加上 M ，再右移 b 位后，与“状态加数 1”的和；
- 5) 若状态 s 大于“状态阈值”，则将 s 更新为 s 与“状态加数 2”的和。
- 码字：将 s 用 m 比特无符号整数的二进制表示，与存储的比特合并在一起。

15

s	Memory	x	写入存储比特数
0		A	0
4		B	3
6	100	A	1
2	1000	A	1
0	10000	A	0
4	10000		

在一种可能的实现中，在所述数值 M 为 256 时，所述第一信息通过 10 比特存储，所述第一子信息通过 8 比特存储，所述第二子信息通过 7 比特存储，所述第三子信息通过 7 比特存储，多个符号的第一数值的加和为 128，所述第三子信息小于 128。

20 当 $M=256$ ($m=8$) 时，满足如下条件：

每个符号的“比特加数”，“状态加数 1”，“状态阈值”，“状态加数 2”分别需要 10 比特，8 比特，7 比特，7 比特，共 32 比特。所有 $P1(x)$ 的和恰为 128，“状态加数 2”的值恒小于 128。

当 $M = 256$ ($m = 8$) 时，“比特加数”，“状态加数 1”，“状态阈值”，“状态加数 2”分别需要 12 比特，8 比特，8 比特，8 比特，8 比特。共 36 比特，无法用 4 字节表示。用如下方法可将 4 部分用共 32

25 比特表示：

1. 比特加数 B 原为 $[2, 2048]$ 间的偶数，需 12 比特，可用 10 比特表示 $B/2-1$ ；
2. 调整 $P1(x)$ 与 $P2(x)$ 的值，使得所有 $P1(x)$ 的和恰为 128，且“状态加数 2”的值恒小于 128。则“状态阈值”、“状态加数 2”的范围均为 $[0, 127]$ ，可用 7 比特表示；

应用此方案，每个符号的“比特加数”，“状态加数 1”，“状态阈值”，“状态加数 2”分别需要 10 比特，8 比特，7 比特，7 比特，共 32 比特。

30

本申请还提供了一种数据编码方法，应用于基于非对称数字系统 ANS 的编码器，所述方法包括：获取第一符号，所述第一符号对应的量化概率和第二符号对应的量化概率的和为 1，且所述第一符号对应的量化概率和 1 的差异小于阈值，所述第二符号对应的量化概率和 0 的差异小于阈值；根据所述编码器的状态信息和数值 M 之间的数值大小关系满足第一条件，将所述状态信息和所述数值 M 之间进行第一运算，得到第一编码数据，并将所述状态信息更新为 0；根据所述编码器的状态信息和数值 M 之间的数值大小关系不满足所述第一条件，将所述状态信息加 1。

在一种可能的实现中，所述方法还包括：获取所述第二符号；将所述编码器的状态信息作为第一编码数据，并将所述状态信息更新为所述数值 M 与 1 的差值。

上述实施例可以针对于共两种符号，且单一符号概率接近 1 时的编、解码方案。此实施例无需构建编码、解码表，因此无需额外内存存储表格。假设共两种符号 a 和 b，且符号 a 的概率接近 1，则编码算法如下：

压缩符号 a：若状态 $s \geq M - 2$ ，则将 $s - M + 2$ （0 或 1）写入 1 比特到存储，然后将 s 更新为 0；否则将 s 更新为 $s + 1$ ；

压缩符号 b：将 s 的全部 m 位写入存储，并更新 s 为 $M - 1$ 。

参照图 9，图 9 为本申请实施例提供的一种数据解码方法的实施例示意，如图 9 示出的那样，本申请实施例提供的一种数据解码方法，包括：

901、获取所述解码器的状态信息以及存储比特，所述状态信息与第一比特相对应，所述第一比特为当前待解码的比特；

示例性的，参照图 10A，图 10A 为本申请实施例提供的一种数据解码方法的具体示意。

其中，状态信息（解码状态）可以为从编码侧获取到的，也可以是解码过程中对于状态信息进行更新得到的一个状态信息。其中，存储比特可以为从编码侧获取到的，也可以是解码过程中对于存储比特进行更新得到的一个存储比特。

902、根据第一映射关系，确定所述状态信息对应的第一信息、第二信息以及第三信息；所述第一信息为所述第一比特对应的目标符号，所述目标符号为所述第一比特的解码结果，所述第二信息表示从所述存储比特中选择的比特位数，所述第三信息用于和根据所述第三信息从所述存储比特中选择的比特进行加和运算，所述加和运算结果用于更新所述状态信息，更新后的状态信息与第二比特相对应，所述第二比特为所述第一比特之后的待解码比特。

应理解，图 6 中描述的第一映射关系和图 9 描述的第一映射关系为不同的映射关系。图 9 中描述的第一映射关系为解码表。

其中，第一信息可以称之为解码符号、第二信息可以称之为读取比特数、第三信息可以称之为状态基数。

在一种可能的实现中，所述第一信息、所述第二信息或所述第三信息为根据概率信息得到；所述概率信息包括概率质量函数 PMF 的数值；所述 PMF 的数值为通过如下方式构建的：依次确定多个符号中每个符号的 PMF 的数值，所述多个符号包括所述目标符号；其中，在确定所述目标符号的 PMF 的数值时，将除去已确定的符号的 PMF 之外的剩余 PMF 的总数值与所述目标符号的概率进行乘积，并计算所述乘积的乘积结果与除去已确定的符号的剩余概率之间的比值，根据所述比值与预设数值之间的关系，确定所述目标符号的 PMF。

在一种可能的实现中，所述第一信息、所述第二信息或所述第三信息为根据概率信息得到；所述概率信息包括累积分布函数 CDF 的数值；所述 CDF 的数值为通过如下方式构建的：基于预设顺序，依次确定每个符号的 CDF 的数值，其中所述预设顺序为根据多个符号对应的 PMF 的数值之间的关系确定，所述多个符号包括所述目标符号；其中，在所述目标符号的概率信息对应的概率分布为对称分布时，所述预设顺序为沿着所述对称分布中峰值点的两侧相交叉的顺序；或者，在所述目标符号的概率信息对应的概率分布为离散分布时，所述预设顺序具体为从大到小的顺序；或者，在所述目标符号的概率信息对应的概率分布为离散分布时，所述预设顺序具体为从小到大的顺序。

在一种可能的实现中，所述第一映射关系中的所述第一信息为通过如下方式构建的：将满足

$C(x) \leq s < (C(x) + P(x))$ 的 x 作为所述第一信息；其中，所述 C 表示 CDF 的数值，所述 P 表示 PMF 的数值，所述 s 表示所述状态信息。

在一种可能的实现中，所述第一映射关系中的所述第二信息为通过如下方式构建的：将 $\text{clz}(s - C(x) + P(x)) + X$ 的数值作为所述第二信息；其中，所述 clz 表示取数值中最高位连续为 0 的比特数量，所述 X 为预设的第五数值。

在一种可能的实现中，所述第一映射关系中的所述第三信息为通过如下方式构建的：将 $s - C(x) + P(x)$ 的数值进行比特移位操作，得到所述第三信息，所述比特移位操作的移位数量为根据所述第二信息得到；或者，将 $s - C(x) + P(x)$ 的数值进行比特移位操作，并将所述比特移位操作的操作结果减去数值 M ，得到所述第三信息，所述 M 为所述量化概率的分母，所述比特移位操作的移位数量为根据所述第二信息得到。

示例性的，在构建解码表时，对于每一种状态 s ，构建方法如下：

“解码符号”为满足 $C(x) \leq s < C(x) + P(x)$ 的 x 值；“读取比特数”的值为 $\text{clz}(s - C(x) + P(x)) + m - 31$ ；“状态基数”的值为 $s - C(x) + P(x)$ 左移“读取比特数”位数的结果，与 M 的差。

参照表 10，表 10 为利用上述方式得到的一个解码表的示意：

表 10			
s	解码符号	读取比特数	状态基数
0	A	1	2
1	A	1	4
2	A	1	6
3	A	0	0
4	A	0	1
5	B	3	0
6	C	3	0
7	D	3	0

在一种可能的实现中，所述第一信息、所述第二信息或所述第三信息表示一个整数数值，也就是单一的一个整数，例如，所述第一信息可以表示一个整数数值，第二信息可以表示一个整数数值，第三信息可以表示一个整数数值。本申请实施例中，解码表中的索引信息仅包括符号，且映射的对象仅包括第一信息、第二信息和第三信息，降低了编码表中内存占用。

应理解，还可以针对于不同的概率分布配置不同的映射关系（包括第一映射关系）。

在一种可能的实现中，可以获取编码结果；所述编码结果为对图像的像素数据进行编码得到的，所述状态信息以及所述存储比特为根据所述编码结果得到；对所述编码结果进行处理，所述像素数据的分布特征，所述分布特征包括均值或方差；根据所述分布特征，从多个映射关系中确定与所述分布特征对应的所述第一映射关系。

参照图 10B，图 10B 为一种解码过程的示意。

其中，模型包括三个部分。

一是对码字中的头文件进行解码，如图像的长、宽，一些必要的参数等。这些参数可以对码字内的信息，建立解码数据（变换数据）对应的分布索引。

二是预先定义若干概率分布（如不同方差的高斯分布、逻辑斯谛分布等），并建立对应的分布索引和解码表，其中编码表根据计算的离散 PMF、CDF 进行建立。

三是图像解码器，对变换数据进行解码，恢复图像。其中，解码器可以为神经网络，也可以为图像色彩变换的后处理过程，如反傅里叶变换，YUV 转 RGB 等。

在解码过程中，对于待解码的码字，首先用模型读取头文件中基本配置，并根据码字建立必要的解码数据的分布索引。对每个维度的解码数据（变换数据），通过分布索引找到对应解码表，利用熵解码得到解码后的变换数据。最后使用模型，将变换数据恢复成原始图像。

在一种可能的实现中，可以通过第一映射关系，得到所述状态信息对应的一个整数数值；将所述整数数值还原为所述第一信息、所述第二信息和所述第三信息，所述第一信息、所述第二信息和所述第三信息分别为一个整数数值。

示例性的，可以将“解码符号”、“读取比特数”、“状态基数”按上述规则合并。单个参数的长度和由状态数 $M = 2^m$ 决定，为方便计算机读取，一般设为 24bits 或 32bits。

参照表 11，表 11 为码表格参数长度汇总的示意：

表 11

合并顺序	“解码符号”长度	“读取比特数”长度	“状态基数”长度	合并参数总长度
不限	$\geq m$	$\geq \lceil \log_2(m+1) \rceil$	$\geq m$	$m \leq 10, 24/32 \text{ bits}$ $m \leq 14, 32 \text{ bits}$

在一种可能的实现中，所述第一信息为所述第一比特对应的目标符号，所述目标符号为所述第一比特的解码结果，所述第二信息表示从所述存储比特中选择的比特位数，所述第三信息用于和根据所述第三信息从所述存储比特中选择的比特进行加和运算，所述加和运算结果用于更新所述状态信息，更新后的状态信息与第二比特相对应，所述第二比特为所述第一比特之后的待解码比特。此外，还可以从所述存储比特中剔除根据所述选择的比特，得到更新后的存储比特。

示例性的，可以根据当前状态值，查表得到“解码符号”、“读取比特数”、“状态基数”；将 s 更新为：从存储中读取“读取比特数”比特得到的数字，与“状态基数”的和。

参照表 12，表 12 为一个解码示意：解码 00101110（状态 001 表示 1，01110 为存储比特），解码结果为 AAABA。

表 12

s	Memory	x	存储中读取比特数
1	01110	A	1
4	0111	A	0
1	0111	A	1
5	011	B	3
3		A	

应理解，图 6 对应的实施例和图 9 实施例中的技术方案之间可以相互组合。示例性的，可以构建三个子表：

子表 1：对于每个状态 s ，求出“解码符号”，方法、结果与上述解码过程相同；

子表 2：对于每个符号 x ，求出“状态加数”，方法、结果与上述编码过程中描述的方式 1 的“状态加数”相同；

子表 3（所有分布共用一张表）：对于每个整数 $i \in [1, M \times 2 - 1]$ ，求出“状态比特数”为 $\text{clz}(i) + m - 31$ 。

解码单个符号过程：

将当前状态 s 代入子表 1，得到解码出的 x “解码符号”；

将解码出的 x 代入子表 2，得到“状态加数”；

计算需从存储读取的比特数：将 s 减“状态加数”的值代入子表 3，得到“状态比特数”；

更新状态 s 为：当前 s 的值减去“状态加数”后，左移“状态比特数”的位数，加上存储是读取“状态比特数”位数的值，对 M 取模。

用上述方法，每步解码的中间结果均与图 9 介绍的解码过程相同。

参照表 13，表 13 为利用上述方式得到的一个解码表的示意：

表 13

状态 s	0	1	2	3	4	5	6	7
解 码 符 号	A	A	A	A	A	B	C	D
符号 x	A		B		C		D	
状态加数	-5		4		5		6	
数字 i	1	2	3	4	5	6	7	8~15
状 态 比 特 数	3	2	2	1	1	1	1	0

本申请实施例能够取得更短的压缩时间。相对于 rANS，编解码时间较 rANS 提升 3-4 倍。
参照表 14，表 14 为一个有益效果的示意：

表 14

METHOD	Encode Time (us)	Decode Time (us)	Encode Bandwidth (M sym/s)	Decode Bandwidth (M sym/s)	Rel. Encode bandwidth	Rel. Decode bandwidth
Range Coder (baseline)	578736	1517031	36.4	13.9	108%	22%
rANS (baseline)	623609	341141	33.8	61.8	100%	100%
ANSAI MASS=10	129752	110278	162.6	191.3	481%	309%
ANSAI MASS=8	133935	101123	157.5	208.6	466%	337%

- 5 本申请实施例的表格内存更小，对于总符号数 X，状态数 M，如果表格每个元素用 32bits 存储，则编码表格大小 4X，解码表格大小 3M；如果压缩至 24bits 存储，则编码表格大小 3X，解码表格大小 3M。
对于特定分布构建表格，表格内存占用小于 100KB (MASS=8, M=256)。对应 tANS 的 10MB~20MB 的表格大小，表格内存占用大大缩减。
参照表 15，表 15 为一个有益效果的示意：

表 15

MASS / log2(M)	Encode	Decode	Encode opt.	Decode opt.
Table bits	32	32	24	24
8	93.9	65	70.4	48.75
10		260		195

在一种可能的实现中，所述第一映射关系中的所述第一信息为通过如下方式构建的：

将满足 $C1(x) \leq s < C1(x) + P1(x)$ 或 $C2(x) \leq s < C2(x) + P2(x)$ 的 x 作为所述第一信息；其中，所述 C1 表示 CDF 对应的第三数值，所述 P1 表示 PMF 对应的第一数值，所述 C2 表示 CDF 对应的第四数值，所述 P2 表示 PMF 对应的第二数值，所述 s 表示所述状态信息；所述第一数值和所述第二数值的加和为所述 PMF 的数值。

在一种可能的实现中，所述第一映射关系中的所述第二信息为通过如下方式构建的：

将 $clz(s' - C(x) + P(x)) + X$ 的数值作为所述第二信息；其中，所述 clz 表示取数值中最高位连续为 0 的比特数量，所述 X 为预设的第五数值，所述 s' 为通过如下得到：

若 $C2(x) \leq s$ ，则 $s' = P1(x) + s - C2(x) + P(x)$ ，否则 $s' = s - C1(x) + P(x)$ 。

在一种可能的实现中，所述第一映射关系中的所述第三信息为通过如下方式构建的：将 $s' - C(x) + P(x)$ 的数值进行比特移位操作，得到所述第三信息，所述比特移位操作的移位数量为根据所述第二信息得到。

对于每一种状态 s，构建方法如下：

1. “解码符号”为满足 $C1(x) \leq s < C1(x) + P1(x)$ 或 $C2(x) \leq s < C2(x) + P2(x)$ 的 x 值；
2. 求出 s 对应的 s' ：
- a) 若 $C2(x) \leq s$ ，则 $s' = P1(x) + s - C2(x) + P(x)$ ；
- b) 否则 $s' = s - C1(x) + P(x)$
- 5 3. 利用 s' ，用与实施例二相同的方法构建解码表：
- 1) “读取比特数”的值为 $clz(s' - C(x) + P(x)) + m - 31$ ；
- 2) “状态基数”的值为 $s' - C(x) + P(x)$ 左移“读取比特数”位数的结果，与 M 的差。

s	s'	解码符号	读取比特数	状态基数
0	0	A	1	2
1	1	A	1	4
2	2	A	1	6
3	6	C	3	0
4	3	A	0	0
5	4	A	0	1
6	5	B	3	0
7	7	D	3	0

此外，每个解码符号可以对应 1~2 段连续的状态 s 。如以上表格中，A 对应的是 0~2，以及 4~5。

解码单个符号时，可以根据当前状态值 s ，查表得到“解码符号”、“读取比特数”、“状态基数”；将

- 10 s 更新为：从存储中读取“读取比特数”比特得到的数字，与“状态基数”的和；

s	Memory	x	存储中读取比特数
4	10000	A	0
0	10000	A	1
2	1000	A	1
6	100	B	3
4		A	0

本申请还提供了一种数据解码方法，应用于基于非对称数字系统 ANS 的解码器，所述方法包括：

获取所述解码器的状态信息，所述状态信息与第一比特相对应，所述第一比特为当前待解码的比特；

- 15 若所述状态信息为数值 M 与 1 的差值，则确定所述第一比特的解码结果为第二符号，并将所述状态信息更新为从存储中读取所述数值 M 位的结果；

若所述状态信息不为数值 M 与 1 的差值，则确定所述第一比特的解码结果为第一符号，若所述状态信息为 0，则将所述状态信息更新为所述数值 M 与 2 的差值与从存储中读出 1 个比特的值的加和结果；

若所述状态信息不为 0，则将所述状态信息更新为所述状态信息与 1 的差值；更新后的状态信息与第二

- 20 比特相对应，所述第二比特为所述第一比特之后的待解码比特。

上述实施例为针对于共两种符号，且单一符号概率接近 1 时的编、解码方案。

两种符号 a, b ，此实施例无需构建编码、解码表，因此无需额外内存存储表格。假设共两种符号 a 和 b ，且符号 a 的概率接近 1，则编码算法如下：

压缩符号 a ：若状态 $s \geq M - 2$ ，则将 $s - M + 2$ （0 或 1）写入 1 比特到存储，然后将 s 更新为 0；否

- 25 则将 s 更新为 $s + 1$ ；

压缩符号 b ：将 s 的全部 m 位写入存储，并更新 s 为 $M - 1$ 。

解码算法为：若当前状态 s 为 $M - 1$ ，则解码出符号 b ，并将 s 更新为从存储中读取 m 位的结果；否

则则解码出 a ，然后更新 s ：1) 若 $s=0$ ，则将 s 更新为 $M-2$ 与从存储中读出一比特的值和；2) 否则，将 s 更新为 $s-1$ 。

接下来从装置的角度介绍本申请实施例提供的一种数据编码装置，参照图 11，图 11 为本申请实施例提供的一种数据编码装置的结构示意，如图 11 所示，本申请实施例提供的一种数据编码装置 1100 包括：

获取模块 1101，用于获取目标符号；

关于获取模块 1101 的具体描述可以参照上述实施例中步骤 601 的描述，这里不再赘述。

编码模块 1102，用于通过第一映射关系，得到所述目标符号对应的第一信息和第二信息；所述第一信息和所述第二信息为根据所述目标符号对应的概率信息得到；

根据所述第一信息以及所述编码器的状态信息，通过第一运算，得到第一编码数据；

根据所述第二信息，通过第二运算，更新所述状态信息，得到第二编码数据；其中，

所述第一编码数据和所述第二编码数据用于得到所述目标符号的编码结果，且所述第一运算和所述第二运算不包括除法运算。

关于编码模块 1102 的具体描述可以参照上述实施例中步骤 602 的描述，这里不再赘述。

在一种可能的实现中，所述第一信息或所述第二信息表示一个整数数值。

在一种可能的实现中，所述第一运算和所述第二运算不包括判断运算。

在一种可能的实现中，所述第一运算包括：

对所述第一信息与所述状态信息进行的加和操作，以及对所述加和操作的加和结果进行的比特移位操作；

从所述状态信息中选择部分比特作为第一编码数据，所述部分比特在所述状态信息中的位置为根据所述比特移位操作的移位结果确定。

在一种可能的实现中，所述第二运算包括：

对所述状态信息与数值 M 进行的加和操作，以及对所述加和操作的加和结果进行的比特移位操作；其中，所述概率信息包括量化概率，所述 M 为所述量化概率的分母；

对所述比特移位操作的移位结果与所述第二信息进行的加和操作。

在一种可能的实现中，所述概率信息包括概率质量函数 PMF 的数值；所述 PMF 的数值为通过如下方式构建的：

依次确定多个符号中每个符号的 PMF 的数值，所述多个符号包括所述目标符号；其中，

在确定所述目标符号的 PMF 的数值时，将除去已确定的符号的 PMF 之外的剩余 PMF 的总数值与所述目标符号的概率进行乘积，并计算所述乘积的乘积结果与除去已确定的符号的剩余概率之间的比值，根据所述比值与预设的第五数值之间的关系，确定所述目标符号的 PMF。

在一种可能的实现中，所述第五数值为 1，所述关系为取较大值。

在一种可能的实现中，所述概率信息包括累积分布函数 CDF 的数值；所述 CDF 的数值为通过如下方式构建的：

基于预设顺序，依次确定每个符号的 CDF 的数值，其中所述预设顺序为根据多个符号对应的 PMF 的数值之间的关系确定，所述多个符号包括所述目标符号；其中，

在所述目标符号的概率信息对应的概率分布为对称分布时，所述预设顺序为沿着所述对称分布中峰值点的两侧相交错的顺序；或者，

在所述目标符号的概率信息对应的概率分布为离散分布时，所述预设顺序具体为从大到小的顺序；或者，

在所述目标符号的概率信息对应的概率分布为离散分布时，所述预设顺序具体为从小到大的顺序。

在一种可能的实现中，所述目标符号为对图像的像素数据进行变换得到，所述对图像的像素数据进行变换得到的数据满足预设分布，所述变换还用于得到所述像素数据的分布特征，所述分布特征包括均值或方差，所述获取模块，还用于：

根据所述分布特征，从多个映射关系中确定与所述分布特征对应的所述第一映射关系。

在一种可能的实现中，所述多个分布特征包括第一分布特征和第二分布特征，所述第一分布特征的熵大于所述第二分布特征的熵；

5 所述第二分布特征对应的映射关系中的量化概率的分母的数值大于所述第一分布特征对应的映射关系中的量化概率的分母。

在一种可能的实现中，所述第一映射关系中的所述第一信息为通过如下方式构建的：

将所述目标符号的 PMF 的数值中最高位连续为 0 的比特数量与预设的第六数值的加和运算；

将所述加和运算的加和结果进行比特移位操作，并将所述比特移位操作的移位结果与对所述目标符号的 PMF 的数值进行比特移位操作的移位结果进行减法操作，得到所述第一信息；或者，

10 将所述加和运算的加和结果进行比特移位操作，并将所述比特移位操作的移位结果与对所述目标符号的 PMF 的数值进行比特移位操作的移位结果进行减法操作，将所述减法操作的操作结果与数值 M 进行加和操作，得到所述第一信息，所述概率信息包括量化概率，所述 M 为所述量化概率的分母。

在一种可能的实现中，所述第一映射关系中的所述第二信息为通过如下方式构建的：

15 将所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值进行减法运算，得到所述第二信息；或者，

将所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值进行减法运算，将所述减法操作的操作结果与数值 M 进行加和操作，得到所述第二信息，所述概率信息包括量化概率，所述 M 为所述量化概率的分母。

在一种可能的实现中，所述第一映射关系中的所述第二信息为通过如下方式构建的：

20 在所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值的减法运算的结果大于或等于 0 时，将所述减法运算的结果作为所述第二信息；或者，

在所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值的减法运算的结果小于 0 时，将所述减法操作的操作结果与数值 M 进行加和操作，得到所述第二信息，所述概率信息包括量化概率，所述 M 为所述量化概率的分母。

25 在一种可能的实现中，所述编码模块，具体用于：

通过第一映射关系，得到所述目标符号对应的一个整数数值；

基于第三运算，将所述整数数值还原为所述第一信息和所述第二信息，所述第一信息和所述第二信息分别为一个整数数值。

30 在一种可能的实现中，所述概率信息包括每个符号对应的 PMF 信息，所述 PMF 信息包括第一数值和第二数值；所述第一数值和所述第二数值的加和为所述 PMF 的数值；所述概率信息还包括每个符号对应的 CDF 信息，所述 CDF 信息包括第三数值和第四数值；所述第三数值和第四数值为通过如下方式构建的：

基于预设顺序，根据多个符号的第一数值，通过累加依次确定每个符号的对应的第三数值；

基于所述预设顺序，根据多个符号的第二数值和第三数值，通过累加依次确定每个符号对应的第四数值。

35 在一种可能的实现中，所述第二信息包括第一子信息、第二子信息和第三子信息；

所述第一映射关系中的所述第一子信息为通过如下方式构建的：将所述目标符号的第三数值和所述目标符号的 PMF 的数值进行减法运算，得到所述第一子信息；

40 所述第一映射关系中的所述第二子信息为通过如下方式构建的：将所述目标符号的第三数值和所述目标符号的第一数值进行加法运算，将所述加法运算的操作结果与 1 进行减法操作，得到所述第二子信息；

所述第一映射关系中的所述第三子信息为通过如下方式构建的：将所述目标符号的第四数值和所述目标符号的第三数值进行减法运算，将所述减法运算的操作结果与所述目标符号的第一数值进行减法操作，得到所述第三子信息。

在一种可能的实现中，所述第二运算包括：

45 对所述状态信息与数值 M 进行的加和操作，以及对所述加和操作的加和结果进行的比特移位操作；其中，所述概率信息包括量化概率，所述 M 为所述量化概率的分母；对所述比特移位操作的移位结果与

所述第一子信息进行的加和操作；

在所述加和的操作结果大于所述第二子信息时，将所述加和的操作结果与所述第三子信息的加和结果作为所述第二编码数据；

在所述加和的操作结果小于所述第二子信息时，将所述加和的操作结果作为所述第二编码数据。

在一种可能的实现中，在所述数值 M 为 256 时，所述第一信息通过 10 比特存储，所述第一子信息通过 8 比特存储，所述第二子信息通过 7 比特存储，所述第三子信息通过 7 比特存储，多个符号的第一数值的加和为 128，所述第三子信息小于 128。

本申请还提供了一种数据编码装置，应用于基于非对称数字系统 ANS 的编码器，所述装置包括：
获取模块，用于获取第一符号，所述第一符号对应的量化概率和第二符号对应的量化概率的和为 1，且所述第一符号对应的量化概率和 1 的差异小于阈值，所述第二符号对应的量化概率和 0 的差异小于阈值；

编码模块，用于根据所述编码器的状态信息和数值 M 之间的数值大小关系满足第一条件，将所述状态信息和所述数值 M 之间进行第一运算，得到第一编码数据，并将所述状态信息更新为 0；

根据所述编码器的状态信息和数值 M 之间的数值大小关系不满足所述第一条件，将所述状态信息加 1；

在一种可能的实现中，所述获取模块，还用于：

获取所述第二符号；

所述编码模块，还用于将所述编码器的状态信息作为第一编码数据，并将所述状态信息更新为所述数值 M 与 1 的差值。

接下来从装置的角度介绍本申请实施例提供的一种数据解码装置，参照图 12，图 12 为本申请实施例提供的一种数据解码装置的结构示意，如图 12 所示，本申请实施例提供的一种数据解码装置 1200 包括：

获取模块 1201，用于获取所述解码器的状态信息以及存储比特，所述状态信息与第一比特相对应，所述第一比特为当前待解码的比特；

关于获取模块 1201 的具体描述可以参照上述实施例中步骤 901 的描述，这里不再赘述。

解码模块 1202，用于根据第一映射关系，确定所述状态信息对应的第一信息、第二信息以及第三信息；所述第一信息为所述第一比特对应的目标符号，所述目标符号为所述第一比特的解码结果，所述第二信息表示从所述存储比特中选择的比特位数，所述第三信息用于和根据所述第三信息从所述存储比特中选择的比特进行加和运算，所述加和运算结果用于更新所述状态信息，更新后的状态信息与第二比特相对应，所述第二比特为所述第一比特之后的待解码比特。

关于解码模块 1202 的具体描述可以参照上述实施例中步骤 902 的描述，这里不再赘述。

在一种可能的实现中，所述解码模块，还用于：

从所述存储比特中剔除根据所述选择的比特，得到更新后的存储比特。

在一种可能的实现中，所述第一信息、所述第二信息或所述第三信息表示一个整数数值。

在一种可能的实现中，所述第一映射关系中的所述第一信息为通过如下方式构建的：

将满足 $C(x) \leq s < (C(x) + P(x))$ 的 x 作为所述第一信息；其中，所述 C 表示 CDF 的数值，所述 P 表示 PMF 的数值，所述 s 表示所述状态信息；或者，

将满足 $C1(x) \leq s < C1(x) + P1(x)$ 或 $C2(x) \leq s < C2(x) + P2(x)$ 的 x 作为所述第一信息；其中，所述 $C1$ 表示 CDF 对应的第三数值，所述 $P1$ 表示 PMF 对应的第一数值，所述 $C2$ 表示 CDF 对应的第四数值，所述 $P2$ 表示 PMF 对应的第二数值，所述 s 表示所述状态信息；所述第一数值和所述第二数值的加和为所述 PMF 的数值。

在一种可能的实现中，所述第一映射关系中的所述第二信息为通过如下方式构建的：

将 $\text{clz}(s - C(x) + P(x)) + X$ 的数值作为所述第二信息；其中，所述 clz 表示取数值中最高位连续为 0 的比特数量，所述 X 为预设的第五数值；或者，

将 $s - C(x) + P(x)$ 的数值进行比特移位操作, 得到所述第三信息, 所述比特移位操作的移位数量为根据所述第二信息得到; 或者,

将 $s - C(x) + P(x)$ 的数值进行比特移位操作, 并将所述比特移位操作的操作结果减去数值 M , 得到所述第三信息, 所述 M 为所述量化概率的分母, 所述比特移位操作的移位数量为根据所述第二信息得到;

5 或者,

将 $\text{clz}(s' - C(x) + P(x)) + X$ 的数值作为所述第二信息; 其中, 所述 clz 表示取数值中最高位连续为 0 的比特数量, 所述 X 为预设的第五数值, 所述 s' 为通过如下得到:

若 $C2(x) \leq s$, 则 $s' = P1(x) + s - C2(x) + P(x)$, 否则 $s' = s - C1(x) + P(x)$ 。

在一种可能的实现中, 所述第一映射关系中的所述第三信息为通过如下方式构建的:

10 将 $s - C(x) + P(x)$ 的数值进行比特移位操作, 得到所述第三信息, 所述比特移位操作的移位数量为根据所述第二信息得到; 或者,

将 $s - C(x) + P(x)$ 的数值进行比特移位操作, 并将所述比特移位操作的操作结果减去数值 M , 得到所述第三信息, 所述 M 为所述量化概率的分母, 所述比特移位操作的移位数量为根据所述第二信息得到; 或者,

15 将 $s' - C(x) + P(x)$ 的数值进行比特移位操作, 得到所述第三信息, 所述比特移位操作的移位数量为根据所述第二信息得到。

在一种可能的实现中, 所述第一信息、所述第二信息或所述第三信息为根据概率信息得到; 所述概率信息包括概率质量函数 PMF 的数值; 所述 PMF 的数值为通过如下方式构建的:

依次确定多个符号中每个符号的 PMF 的数值, 所述多个符号包括所述目标符号; 其中,

20 在确定所述目标符号的 PMF 的数值时, 将除去已确定的符号的 PMF 之外的剩余 PMF 的总数值与所述目标符号的概率进行乘积, 并计算所述乘积的乘积结果与除去已确定的符号的剩余概率之间的比值, 根据所述比值与预设数值之间的关系, 确定所述目标符号的 PMF。

在一种可能的实现中, 所述预设数值为 1, 所述关系为取较大值。

25 在一种可能的实现中, 所述第一信息、所述第二信息或所述第三信息为根据概率信息得到; 所述概率信息包括累积分布函数 CDF 的数值; 所述 CDF 的数值为通过如下方式构建的:

基于预设顺序, 依次确定每个符号的 CDF 的数值, 其中所述预设顺序为根据多个符号对应的 PMF 的数值之间的关系确定, 所述多个符号包括所述目标符号; 其中,

在所述目标符号的概率信息对应的概率分布为对称分布时, 所述预设顺序为沿着所述对称分布中峰值点的两侧相交错顺序; 或者,

30 在所述目标符号的概率信息对应的概率分布为离散分布时, 所述预设顺序具体为从大到小的顺序; 或者,

在所述目标符号的概率信息对应的概率分布为离散分布时, 所述预设顺序具体为从小到大的顺序。

在一种可能的实现中, 所述获取模块, 还用于:

35 获取编码结果; 所述编码结果为对图像的像素数据进行编码得到的, 所述状态信息以及所述存储比特为根据所述编码结果得到;

对所述编码结果进行处理, 所述像素数据的分布特征, 所述分布特征包括均值或方差;

根据所述分布特征, 从多个映射关系中确定与所述分布特征对应的所述第一映射关系。

本申请还提供了一种数据解码装置, 应用于基于非对称数字系统 ANS 的解码器, 所述装置包括:

40 获取模块, 用于获取所述解码器的状态信息, 所述状态信息与第一比特相对应, 所述第一比特为当前待解码的比特;

解码模块, 用于若所述状态信息为数值 M 与 1 的差值, 则确定所述第一比特的解码结果为第二符号, 并将所述状态信息更新为从存储中读取所述数值 M 位的结果;

若所述状态信息不为数值 M 与 1 的差值, 则确定所述第一比特的解码结果为第一符号, 若所述状态信息为 0, 则将所述状态信息更新为所述数值 M 与 2 的差值与从存储中读出 1 个比特的值的加和结果;

45 若所述状态信息不为 0, 则将所述状态信息更新为所述状态信息与 1 的差值; 更新后的状态信息与第二比特相对应, 所述第二比特为所述第一比特之后的待解码比特。

接下来介绍本申请实施例提供的一种执行设备，请参阅图 13，图 13 为本申请实施例提供的执行设备的一种结构示意图，执行设备 1300 具体可以表现为手机、平板、笔记本电脑、智能穿戴设备、服务器等，此处不做限定。其中，执行设备 1300 实现图 6 以及图 9 对应实施例中数据编解码方法的功能。具体的，执行设备 1300 包括：接收器 1301、发射器 1302、处理器 1303 和存储器 1304（其中执行设备 1300 中的处理器 1303 的数量可以一个或多个），其中，处理器 1303 可以包括应用处理器 13031 和通信处理器 13032。在本申请的一些实施例中，接收器 1301、发射器 1302、处理器 1303 和存储器 1304 可通过总线或其它方式连接。

存储器 1304 可以包括只读存储器和随机存取存储器，并向处理器 1303 提供指令和数据。存储器 1304 的一部分还可以包括非易失性随机存取存储器（non-volatile random access memory, NVRAM）。存储器 1304 存储有处理器和操作指令、可执行模块或者数据结构，或者它们的子集，或者它们的扩展集，其中，操作指令可包括各种操作指令，用于实现各种操作。

处理器 1303 控制执行设备的操作。具体的应用中，执行设备的各个组件通过总线系统耦合在一起，其中总线系统除包括数据总线之外，还可以包括电源总线、控制总线和状态信号总线等。但是为了清楚说明起见，在图中将各种总线都称为总线系统。

上述本申请实施例揭示的方法可以应用于处理器 1303 中，或者由处理器 1303 实现。处理器 1303 可以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 1303 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 1303 可以是通用处理器、数字信号处理器（digital signal processing, DSP）、微处理器或微控制器、以及视觉处理器（vision processing unit, VPU）、张量处理器（tensor processing unit, TPU）等适用于 AI 运算的处理器，还可进一步包括专用集成电路（application specific integrated circuit, ASIC）、现场可编程门阵列（field-programmable gate array, 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。该处理器 1303 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。软件模块可以位于随机存储器，闪存、只读存储器，可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器 1304，处理器 1303 读取存储器 1304 中的信息，结合其硬件完成上述实施例中图 6 以及图 9 对应实施例中数据编解码方法。

接收器 1301 可用于接收输入的数字或字符信息，以及产生与执行设备的相关设置以及功能控制有关的信号输入。发射器 1302 可用于通过第一接口输出数字或字符信息；发射器 1302 还可用于通过第一接口向磁盘组发送指令，以修改磁盘组中的数据；发射器 1302 还可以包括显示屏等显示设备。

本申请实施例还提供了一种训练设备，请参阅图 14，图 14 是本申请实施例提供的服务器一种结构示意图，具体的，服务器 1400 由一个或多个服务器实现，服务器 1400 可因配置或性能不同而产生比较大的差异，可以包括一个或一个以上中央处理器（central processing units, CPU）1414（例如，一个或一个以上处理器）和存储器 1432，一个或一个以上存储应用程序 1442 或数据 1444 的存储介质 1430（例如一个或一个以上海量存储设备）。其中，存储器 1432 和存储介质 1430 可以是短暂存储或持久存储。存储在存储介质 1430 的程序可以包括一个或一个以上模块（图示没标出），每个模块可以包括对服务器中的一系列指令操作。更进一步地，中央处理器 1414 可以设置为与存储介质 1430 通信，在服务器 1400 上执行存储介质 1430 中的一系列指令操作。

服务器 1400 还可以包括一个或一个以上电源 1426，一个或一个以上有线或无线网络接口 1450，一个或一个以上输入输出接口 1458；或，一个或一个以上操作系统 1441，例如 Windows Server™，Mac OS X™，Unix™，Linux™，FreeBSD™ 等等。

具体的，服务器可以进行上述实施例中模型训练相关的步骤。

本申请实施例中还提供一种包括计算机程序产品，当其在计算机上运行时，使得计算机执行图 6 以

及图 9 对应实施例中的数据编解码方法。

本申请实施例中还提供一种计算机可读存储介质，该计算机可读存储介质中存储有用于进行信号处理的程序，当其在计算机上运行时，使得计算机执行图 6 以及图 9 对应实施例中的数据编解码方法。

本申请实施例提供的执行设备、训练设备或终端设备具体可以为芯片，芯片包括：处理单元和通信单元，所述处理单元例如可以是处理器，所述通信单元例如可以是输入/输出接口、管脚或电路等。该处理单元可执行存储单元存储的计算机执行指令，以使执行设备内的芯片执行上述实施例描述的数据处理方法，或者，以使训练设备内的芯片执行上述实施例描述的数据处理方法。可选地，所述存储单元为所述芯片内的存储单元，如寄存器、缓存等，所述存储单元还可以是所述无线接入设备端内的位于所述芯片外部的存储单元，如只读存储器（read-only memory, ROM）或可存储静态信息和指令的其他类型的静态存储设备，随机存取存储器（random access memory, RAM）等。

具体的，请参阅图 15，图 15 为本申请实施例提供的芯片的一种结构示意图，所述芯片可以表现为神经网络处理器 NPU1500，NPU 1500 作为协处理器挂载到主 CPU（Host CPU）上，由 Host CPU 分配任务。NPU 的核心部分为运算电路 1503，通过控制器 1504 控制运算电路 1503 提取存储器中的矩阵数据并进行乘法运算。

NPU 1500 可以通过内部的各个器件之间的相互配合，来实现图 6 以及图 9 对应实施例中的数据编解码方法。

更具体的，在一些实现中，NPU 1500 中的运算电路 1503 内部包括多个处理单元（Process Engine, PE）。在一些实现中，运算电路 1503 是二维脉动阵列。运算电路 1503 还可以是一维脉动阵列或者能够执行例如乘法和加法这样的数学运算的其它电子线路。在一些实现中，运算电路 1503 是通用的矩阵处理器。

举例来说，假设有输入矩阵 A，权重矩阵 B，输出矩阵 C。运算电路从权重存储器 1502 中取矩阵 B 相应的数据，并缓存在运算电路中每一个 PE 上。运算电路从输入存储器 1501 中取矩阵 A 数据与矩阵 B 进行矩阵运算，得到的矩阵的部分结果或最终结果，保存在累加器（accumulator）1508 中。

统一存储器 1506 用于存放输入数据以及输出数据。权重数据直接通过存储单元访问控制器（Direct Memory Access Controller, DMAC）1505，DMAC 被搬运到权重存储器 1502 中。输入数据也通过 DMAC 被搬运到统一存储器 1506 中。

BIU 为 Bus Interface Unit 即，总线接口单元 1510，用于 AXI 总线与 DMAC 和取指存储器（Instruction Fetch Buffer, IFB）1509 的交互。

总线接口单元 1510（Bus Interface Unit，简称 BIU），用于取指存储器 1509 从外部存储器获取指令，还用于存储单元访问控制器 1505 从外部存储器获取输入矩阵 A 或者权重矩阵 B 的原数据。

DMAC 主要用于将外部存储器 DDR 中的输入数据搬运到统一存储器 1506 或将权重数据搬运到权重存储器 1502 中或将输入数据搬运到输入存储器 1501 中。

向量计算单元 1507 包括多个运算处理单元，在需要的情况下，对运算电路 1503 的输出做进一步处理，如向量乘，向量加，指数运算，对数运算，大小比较等等。主要用于神经网络中非卷积/全连接层网络计算，如 Batch Normalization（批归一化），像素级求和，对特征平面进行上采样等。

在一些实现中，向量计算单元 1507 能将经处理的输出的向量存储到统一存储器 1506。例如，向量计算单元 1507 可以将线性函数；或，非线性函数应用到运算电路 1503 的输出，例如对卷积层提取的特征平面进行线性插值，再例如累加值的向量，用以生成激活值。在一些实现中，向量计算单元 1507 生成归一化的值、像素级求和的值，或二者均有。在一些实现中，处理过的输出的向量能够用作到运算电路 1503 的激活输入，例如用于在神经网络中的后续层中的使用。

控制器 1504 连接的取指存储器（instruction fetch buffer）1509，用于存储控制器 1504 使用的指令；

统一存储器 1506，输入存储器 1501，权重存储器 1502 以及取指存储器 1509 均为 On-Chip 存储器。外部存储器私有于该 NPU 硬件架构。

其中，上述任一处提到的处理器，可以是一个通用中央处理器，微处理器，ASIC，或一个或多个用

于控制上述程序执行的集成电路。

另外需说明的是，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实施例方案的目的。另外，本申请提供的装置实施例附图中，模块之间的连接关系表示它们之间具有通信连接，具体可以实现为一条或多条通信总线或信号线。

通过以上的实施方式的描述，所属领域的技术人员可以清楚地了解到本申请可借助软件加必需的通用硬件的方式来实现，当然也可以通过专用硬件包括专用集成电路、专用 CPU、专用存储器、专用元器件等来实现。一般情况下，凡由计算机程序完成的功能都可以很容易地用相应的硬件来实现，而且，用来实现同一功能的具体硬件结构也可以是多多种多样的，例如模拟电路、数字电路或专用电路等。但是，对本申请而言更多情况下软件程序实现是更佳的实施方式。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品存储在可读取的存储介质中，如计算机的软盘、U 盘、移动硬盘、ROM、RAM、磁碟或者光盘等，包括若干指令用以使得一台计算机设备（可以是个人计算机，训练设备，或者网络设备等）执行本申请各个实施例所述的方法。

在上述实施例中，可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时，可以全部或部分地以计算机程序产品的形式实现。

所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机程序指令时，全部或部分地产生按照本申请实施例所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中，或者从一个计算机可读存储介质向另一计算机可读存储介质传输，例如，所述计算机指令可以从一个网站站点、计算机、训练设备或数据中心通过有线（例如同轴电缆、光纤、数字用户线（DSL））或无线（例如红外、无线、微波等）方式向另一个网站站点、计算机、训练设备或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存储的任何可用介质或者是包含一个或多个可用介质集成的训练设备、数据中心等数据存储设备。所述可用介质可以是磁性介质，（例如，软盘、硬盘、磁带）、光介质（例如，DVD）、或者半导体介质（例如固态硬盘（Solid State Disk，SSD））等。

权 利 要 求 书

1. 一种数据编码方法，其特征在于，应用于基于非对称数字系统 ANS 的编码器，所述方法包括：
获取目标符号；
通过第一映射关系，得到所述目标符号对应的第一信息和第二信息；所述第一信息和所述第二信息
5 为根据所述目标符号对应的概率信息得到；
根据所述第一信息以及所述编码器的状态信息，通过第一运算，得到第一编码数据；
根据所述第二信息，通过第二运算，更新所述状态信息，得到第二编码数据；其中，
所述第一编码数据和所述第二编码数据用于得到所述目标符号的编码结果，且所述第一运算和所述
第二运算不包括除法运算。
- 10 2. 根据权利要求 1 所述的方法，其特征在于，所述第一信息或所述第二信息表示一个整数数值。
3. 根据权利要求 1 或 2 所述的方法，其特征在于，所述第一运算和所述第二运算不包括判断运算。
- 15 4. 根据权利要求 1 至 3 任一所述的方法，其特征在于，所述第一运算包括：
对所述第一信息与所述状态信息进行的加和操作，以及对所述加和操作的加和结果进行的比特移位
操作；从所述状态信息中选择部分比特作为第一编码数据，所述部分比特在所述状态信息中的位置为根
据所述比特移位操作的移位结果确定。
- 20 5. 根据权利要求 1 至 4 任一所述的方法，其特征在于，所述第二运算包括：
对所述状态信息与数值 M 进行的加和操作，以及对所述加和操作的加和结果进行的比特移位操作；
其中，所述概率信息包括量化概率，所述 M 为所述量化概率的分母；对所述比特移位操作的移位结果与
所述第二信息进行的加和操作；或者，
在所述状态信息初始化为所述 M 的情况下，对所述状态信息进行的比特移位操作；对所述比特移位
25 操作的移位结果与所述第二信息进行的加和操作；对所述加和操作的加和结果与所述 M 进行减法运算。
6. 根据权利要求 1 至 5 任一所述的方法，其特征在于，所述概率信息包括概率质量函数 PMF 的数
值；所述 PMF 的数值为通过如下方式构建的：
依次确定多个符号中每个符号的 PMF 的数值，所述多个符号包括所述目标符号；其中，
30 在确定所述目标符号的 PMF 的数值时，将除去已确定的符号的 PMF 之外的剩余 PMF 的总数值与所述
目标符号的概率进行乘积，并计算所述乘积的乘积结果与除去已确定的符号的剩余概率之间的比值，根
据所述比值与预设的第五数值之间的关系，确定所述目标符号的 PMF。
7. 根据权利要求 6 所述的方法，其特征在于，所述第一数值为 1，所述关系为取较大值。
- 35 8. 根据权利要求 1 至 7 任一所述的方法，其特征在于，所述概率信息包括累积分布函数 CDF 的数
值；所述 CDF 的数值为通过如下方式构建的：
基于预设顺序，依次确定每个符号的 CDF 的数值，其中所述预设顺序为根据多个符号对应的 PMF 的
数值之间的关系确定，所述多个符号包括所述目标符号；其中，
40 在所述目标符号的概率信息对应的概率分布为对称分布时，所述预设顺序为沿着所述对称分布中峰
值点的两侧相交错的顺序；或者，
在所述目标符号的概率信息对应的概率分布为离散分布时，所述预设顺序具体为从大到小的顺序；
或者，
在所述目标符号的概率信息对应的概率分布为离散分布时，所述预设顺序具体为从小到大的顺序。
- 45

9. 根据权利要求 1 至 8 任一所述的方法, 其特征在于, 所述目标符号为对图像的像素数据进行变换得到, 所述对图像的像素数据进行变换得到的数据满足预设分布, 所述变换还用于得到所述像素数据的分布特征, 所述分布特征包括均值或方差, 所述方法还包括:

根据所述分布特征, 从多个映射关系中确定与所述分布特征对应的所述第一映射关系。

10. 根据权利要求 9 所述的方法, 其特征在于, 所述多个分布特征包括第一分布特征和第二分布特征, 所述第一分布特征的熵大于所述第二分布特征的熵;

所述第二分布特征对应的映射关系中的量化概率的分母的数值大于所述第一分布特征对应的映射关系中的量化概率的分母。

11. 根据权利要求 1 至 10 任一所述的方法, 其特征在于, 所述第一映射关系中的所述第一信息为通过如下方式构建的:

将所述目标符号的 PMF 的数值中最高位连续为 0 的比特数量与预设的第六数值的加和运算;

将所述加和运算的加和结果进行比特移位操作, 并将所述比特移位操作的移位结果与对所述目标符号的 PMF 的数值进行比特移位操作的移位结果进行减法操作, 得到所述第一信息; 或者,

将所述加和运算的加和结果进行比特移位操作, 并将所述比特移位操作的移位结果与对所述目标符号的 PMF 的数值进行比特移位操作的移位结果进行减法操作, 将所述减法操作的操作结果与数值 M 进行加和操作, 得到所述第一信息, 所述概率信息包括量化概率, 所述 M 为所述量化概率的分母。

12. 根据权利要求 1 至 11 任一所述的方法, 其特征在于, 所述第一映射关系中的所述第二信息为通过如下方式构建的:

将所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值进行减法运算, 得到所述第二信息; 或者,

将所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值进行减法运算, 将所述减法操作的操作结果与数值 M 进行加和操作, 得到所述第二信息, 所述概率信息包括量化概率, 所述 M 为所述量化概率的分母。

13. 根据权利要求 1 至 11 任一所述的方法, 其特征在于, 所述第一映射关系中的所述第二信息为通过如下方式构建的:

在所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值的减法运算的结果大于或等于 0 时, 将所述减法运算的结果作为所述第二信息; 或者,

在所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值的减法运算的结果小于 0 时, 将所述减法操作的操作结果与数值 M 进行加和操作, 得到所述第二信息, 所述概率信息包括量化概率, 所述 M 为所述量化概率的分母。

14. 根据权利要求 1 至 13 任一所述的方法, 其特征在于, 所述通过第一映射关系, 得到所述目标符号对应的第一信息和第二信息, 包括:

通过第一映射关系, 得到所述目标符号对应的一个整数数值;

基于第三运算, 将所述整数数值还原为所述第一信息和所述第二信息, 所述第一信息和所述第二信息分别为一个整数数值。

15. 根据权利要求 1 至 14 任一所述的方法, 其特征在于, 所述概率信息包括每个符号对应的 PMF 信息, 所述 PMF 信息包括第一数值和第二数值; 所述第一数值和所述第二数值的加和为所述 PMF 的数值; 所述概率信息还包括每个符号对应的 CDF 信息, 所述 CDF 信息包括第三数值和第四数值; 所述第三数值和第四数值为通过如下方式构建的:

基于预设顺序, 根据多个符号的第一数值, 通过累加依次确定每个符号的对应的第三数值;

基于所述预设顺序，根据多个符号的第二数值和第三数值，通过累加依次确定每个符号对应的第四数值。

16. 根据权利要求 15 所述的方法，其特征在于，所述第二信息包括第一子信息、第二子信息和第三子信息；

所述第一映射关系中的所述第一子信息为通过如下方式构建的：将所述目标符号的第三数值和所述目标符号的 PMF 的数值进行减法运算，得到所述第一子信息；

所述第一映射关系中的所述第二子信息为通过如下方式构建的：将所述目标符号的第三数值和所述目标符号的第一数值进行加法运算，将所述加法运算的操作结果与 1 进行减法操作，得到所述第二子信息；

所述第一映射关系中的所述第三子信息为通过如下方式构建的：将所述目标符号的第四数值和所述目标符号的第三数值进行减法运算，将所述减法运算的操作结果与所述目标符号的第一数值进行减法操作，得到所述第三子信息。

17. 根据权利要求 16 所述的方法，其特征在于，所述第二运算包括：

对所述状态信息与数值 M 进行的加和操作，以及对所述加和操作的加和结果进行的比特移位操作；其中，所述概率信息包括量化概率，所述 M 为所述量化概率的分母；对所述比特移位操作的移位结果与所述第一子信息进行的加和操作；

在所述加和操作的加和结果大于所述第二子信息时，将所述加和操作的加和结果与所述第三子信息的加和结果作为所述第二编码数据；

在所述加和操作的加和结果小于所述第二子信息时，将所述加和操作的加和结果作为所述第二编码数据。

18. 根据权利要求 16 或 17 所述的方法，其特征在于，在所述数值 M 为 256 时，所述第一信息通过 10 比特存储，所述第一子信息通过 8 比特存储，所述第二子信息通过 7 比特存储，所述第三子信息通过 7 比特存储，多个符号的第一数值的加和为 128，所述第三子信息小于 128。

19. 一种数据编码方法，其特征在于，应用于基于非对称数字系统 ANS 的编码器，所述方法包括：获取第一符号，所述第一符号对应的量化概率和第二符号对应的量化概率的和为 1，且所述第一符号对应的量化概率和 1 的差异小于阈值，所述第二符号对应的量化概率和 0 的差异小于阈值；

根据所述编码器的状态信息和数值 M 之间的数值大小关系满足第一条件，将所述状态信息和所述数值 M 之间进行第一运算，得到第一编码数据，并将所述状态信息更新为 0；

根据所述编码器的状态信息和数值 M 之间的数值大小关系不满足所述第一条件，将所述状态信息加 1。

20. 根据权利要求 19 所述的方法，其特征在于，所述方法还包括：

获取所述第二符号；
将所述编码器的状态信息作为第一编码数据，并将所述状态信息更新为所述数值 M 与 1 的差值。

21. 一种数据解码方法，其特征在于，应用于基于非对称数字系统 ANS 的解码器，所述方法包括：获取所述解码器的状态信息以及存储比特，所述状态信息与第一比特相对应，所述第一比特为当前待解码的比特；

根据第一映射关系，确定所述状态信息对应的第一信息、第二信息以及第三信息；所述第一信息为所述第一比特对应的目标符号，所述目标符号为所述第一比特的解码结果，所述第二信息表示从所述存储比特中选择的比特位数，所述第三信息用于和根据所述第三信息从所述存储比特中选择的比特进行加和运算，所述加和运算结果用于更新所述状态信息，更新后的状态信息与第二比特相对应，所述第二比

特为所述第一比特之后的待解码比特。

22. 根据权利要求 21 所述的方法, 其特征在于, 所述方法还包括:
从所述存储比特中剔除根据所述选择的比特, 得到更新后的存储比特。

5

23. 根据权利要求 21 或 22 所述的方法, 其特征在于, 所述第一信息、所述第二信息或所述第三信息表示一个整数数值。

24. 根据权利要求 21 至 23 任一所述的方法, 其特征在于, 所述第一映射关系中的所述第一信息为
10 通过如下方式构建的:

将满足 $C(x) \leq s < (C(x) + P(x))$ 的 x 作为所述第一信息; 其中, 所述 C 表示 CDF 的数值, 所述 P 表示 PMF 的数值, 所述 s 表示所述状态信息。

25. 根据权利要求 21 至 23 任一所述的方法, 其特征在于, 所述第一映射关系中的所述第一信息为
15 通过如下方式构建的:

将满足 $C1(x) \leq s < C1(x) + P1(x)$ 或 $C2(x) \leq s < C2(x) + P2(x)$ 的 x 作为所述第一信息; 其中, 所述 $C1$ 表示 CDF 对应的第三数值, 所述 $P1$ 表示 PMF 对应的第一数值, 所述 $C2$ 表示 CDF 对应的第四数值, 所述 $P2$ 表示 PMF 对应的第二数值, 所述 s 表示所述状态信息; 所述第一数值和所述第二数值的加和为所述 PMF 的数值。

20

26. 根据权利要求 21 至 25 任一所述的方法, 其特征在于, 所述第一映射关系中的所述第二信息为
通过如下方式构建的:

将 $\text{clz}(s - C(x) + P(x)) + X$ 的数值作为所述第二信息; 其中, 所述 clz 表示取数值中最高位连续为 0 的比特数量, 所述 X 为预设的第五数值。

25

27. 根据权利要求 21 至 25 任一所述的方法, 其特征在于, 所述第一映射关系中的所述第二信息为
通过如下方式构建的:

将 $\text{clz}(s' - C(x) + P(x)) + X$ 的数值作为所述第二信息; 其中, 所述 clz 表示取数值中最高位连续为 0 的比特数量, 所述 X 为预设的第五数值, 所述 s' 为通过如下得到:

30 若 $C2(x) \leq s$, 则 $s' = P1(x) + s - C2(x) + P(x)$, 否则 $s' = s - C1(x) + P(x)$ 。

28. 根据权利要求 21 至 27 任一所述的方法, 其特征在于, 所述第一映射关系中的所述第三信息为
通过如下方式构建的:

35 将 $s - C(x) + P(x)$ 的数值进行比特移位操作, 得到所述第三信息, 所述比特移位操作的移位数量为根据所述第二信息得到; 或者,

将 $s - C(x) + P(x)$ 的数值进行比特移位操作, 并将所述比特移位操作的操作结果减去数值 M , 得到所述第三信息, 所述 M 为所述量化概率的分母, 所述比特移位操作的移位数量为根据所述第二信息得到; 或者,

40 将 $s' - C(x) + P(x)$ 的数值进行比特移位操作, 得到所述第三信息, 所述比特移位操作的移位数量为根据所述第二信息得到。

29. 根据权利要求 21 至 28 任一所述的方法, 其特征在于, 所述第一信息、所述第二信息或所述第三信息为根据概率信息得到; 所述概率信息包括概率质量函数 PMF 的数值; 所述 PMF 的数值为通过如下方式构建的:

45 依次确定多个符号中每个符号的 PMF 的数值, 所述多个符号包括所述目标符号; 其中, 在确定所述目标符号的 PMF 的数值时, 将除去已确定的符号的 PMF 之外的剩余 PMF 的总数值与所述

目标符号的概率进行乘积，并计算所述乘积的乘积结果与除去已确定的符号的剩余概率之间的比值，根据所述比值与预设数值之间的关系，确定所述目标符号的 PMF。

30. 根据权利要求 21 至 29 任一所述的方法，其特征在于，所述第一信息、所述第二信息或所述第三信息为根据概率信息得到；所述概率信息包括累积分布函数 CDF 的数值；所述 CDF 的数值为通过如下方式构建的：

基于预设顺序，依次确定每个符号的 CDF 的数值，其中所述预设顺序为根据多个符号对应的 PMF 的数值之间的关系确定，所述多个符号包括所述目标符号；其中，

在所述目标符号的概率信息对应的概率分布为对称分布时，所述预设顺序为沿着所述对称分布中峰值点的两侧相交错的顺序；或者，

在所述目标符号的概率信息对应的概率分布为离散分布时，所述预设顺序具体为从大到小的顺序；或者，

在所述目标符号的概率信息对应的概率分布为离散分布时，所述预设顺序具体为从小到大的顺序。

31. 根据权利要求 21 至 30 任一所述的方法，其特征在于，所述方法还包括：

获取编码结果；所述编码结果为对图像的像素数据进行编码得到的，所述状态信息以及所述存储比特为根据所述编码结果得到；

对所述编码结果进行处理，所述像素数据的分布特征，所述分布特征包括均值或方差；

根据所述分布特征，从多个映射关系中确定与所述分布特征对应的所述第一映射关系。

32. 根据权利要求 21 至 31 任一所述的方法，其特征在于，所述根据第一映射关系，确定所述状态信息对应的第一信息、第二信息以及第三信息，包括：

通过第一映射关系，得到所述状态信息对应的一个整数数值；

将所述整数数值还原为所述第一信息、所述第二信息和所述第三信息，所述第一信息、所述第二信息和所述第三信息分别为一个整数数值。

33. 一种数据解码方法，其特征在于，应用于基于非对称数字系统 ANS 的解码器，所述方法包括：获取所述解码器的状态信息，所述状态信息与第一比特相对应，所述第一比特为当前待解码的比特；

若所述状态信息为数值 M 与 1 的差值，则确定所述第一比特的解码结果为第二符号，并将所述状态信息更新为从存储中读取所述数值 M 位的结果；

若所述状态信息不为数值 M 与 1 的差值，则确定所述第一比特的解码结果为第一符号，若所述状态信息为 0，则将所述状态信息更新为所述数值 M 与 2 的差值与从存储中读出 1 个比特的值的加和结果；若所述状态信息不为 0，则将所述状态信息更新为所述状态信息与 1 的差值；更新后的状态信息与第二比特相对应，所述第二比特为所述第一比特之后的待解码比特。

34. 一种数据编码装置，其特征在于，应用于基于非对称数字系统 ANS 的编码器，所述装置包括：获取模块，用于获取目标符号；

编码模块，用于通过第一映射关系，得到所述目标符号对应的第一信息和第二信息；所述第一信息和所述第二信息为根据所述目标符号对应的概率信息得到；

根据所述第一信息以及所述编码器的状态信息，通过第一运算，得到第一编码数据；

根据所述第二信息，通过第二运算，更新所述状态信息，得到第二编码数据；其中，

所述第一编码数据和所述第二编码数据用于得到所述目标符号的编码结果，且所述第一运算和所述第二运算不包括除法运算。

35. 根据权利要求 34 所述的装置，其特征在于，所述第一信息或所述第二信息表示一个整数数值。

36. 根据权利要求 34 或 35 所述的装置，其特征在于，所述第一运算和所述第二运算不包括判断运算。

37. 根据权利要求 34 至 36 任一所述的装置，其特征在于，所述第一运算包括：

5 对所述第一信息与所述状态信息进行的加和操作，以及对所述加和操作的加和结果进行的比特移位操作；

从所述状态信息中选择部分比特作为第一编码数据，所述部分比特在所述状态信息中的位置为根据所述比特移位操作的移位结果确定。

10 38. 根据权利要求 34 至 37 任一所述的装置，其特征在于，所述第二运算包括：

对所述状态信息与数值 M 进行的加和操作，以及对所述加和操作的加和结果进行的比特移位操作；其中，所述概率信息包括量化概率，所述 M 为所述量化概率的分母；

对所述比特移位操作的移位结果与所述第二信息进行的加和操作。

15 39. 根据权利要求 34 至 38 任一所述的装置，其特征在于，所述概率信息包括概率质量函数 PMF 的数值；所述 PMF 的数值为通过如下方式构建的：

依次确定多个符号中每个符号的 PMF 的数值，所述多个符号包括所述目标符号；其中，

在确定所述目标符号的 PMF 的数值时，将除去已确定的符号的 PMF 之外的剩余 PMF 的总数值与所述目标符号的概率进行乘积，并计算所述乘积的乘积结果与除去已确定的符号的剩余概率之间的比值，根据所述比值与预设的第五数值之间的关系，确定所述目标符号的 PMF。

20

40. 根据权利要求 39 所述的装置，其特征在于，所述第五数值为 1，所述关系为取较大值。

25 41. 根据权利要求 34 至 40 任一所述的装置，其特征在于，所述概率信息包括累积分布函数 CDF 的数值；所述 CDF 的数值为通过如下方式构建的：

基于预设顺序，依次确定每个符号的 CDF 的数值，其中所述预设顺序为根据多个符号对应的 PMF 的数值之间的关系确定，所述多个符号包括所述目标符号；其中，

在所述目标符号的概率信息对应的概率分布为对称分布时，所述预设顺序为沿着所述对称分布中峰值点的两侧相交错顺序；或者，

30 在所述目标符号的概率信息对应的概率分布为离散分布时，所述预设顺序具体为从大到小的顺序；或者，

在所述目标符号的概率信息对应的概率分布为离散分布时，所述预设顺序具体为从小到大的顺序。

35 42. 根据权利要求 34 至 41 任一所述的装置，其特征在于，所述目标符号为对图像的像素数据进行变换得到，所述对图像的像素数据进行变换得到的数据满足预设分布，所述变换还用于得到所述像素数据的分布特征，所述分布特征包括均值或方差，所述获取模块，还用于：

根据所述分布特征，从多个映射关系中确定与所述分布特征对应的所述第一映射关系。

40 43. 根据权利要求 42 所述的装置，其特征在于，所述多个分布特征包括第一分布特征和第二分布特征，所述第一分布特征的熵大于所述第二分布特征的熵；

所述第二分布特征对应的映射关系中的量化概率的分母的数值大于所述第一分布特征对应的映射关系中的量化概率的分母。

45 44. 根据权利要求 34 至 43 任一所述的装置，其特征在于，所述第一映射关系中的所述第一信息为通过如下方式构建的：

将所述目标符号的 PMF 的数值中最高位连续为 0 的比特数量与预设的第六数值的加和运算；

将所述加和运算的加和结果进行比特移位操作，并将所述比特移位操作的移位结果与对所述目标符号的 PMF 的数值进行比特移位操作的移位结果进行减法操作，得到所述第一信息；或者，

将所述加和运算的加和结果进行比特移位操作，并将所述比特移位操作的移位结果与对所述目标符号的 PMF 的数值进行比特移位操作的移位结果进行减法操作，将所述减法操作的操作结果与数值 M 进行加和操作，得到所述第一信息，所述概率信息包括量化概率，所述 M 为所述量化概率的分母。

45. 根据权利要求 34 至 44 任一所述的装置，其特征在于，所述第一映射关系中的所述第二信息为通过如下方式构建的：

将所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值进行减法运算，得到所述第二信息；或者，

将所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值进行减法运算，将所述减法操作的操作结果与数值 M 进行加和操作，得到所述第二信息，所述概率信息包括量化概率，所述 M 为所述量化概率的分母。

46. 根据权利要求 34 至 44 任一所述的装置，其特征在于，所述第一映射关系中的所述第二信息为通过如下方式构建的：

在所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值的减法运算的结果大于或等于 0 时，将所述减法运算的结果作为所述第二信息；或者，

在所述目标符号的 CDF 的数值和所述目标符号的 PMF 的数值的减法运算的结果小于 0 时，将所述减法操作的操作结果与数值 M 进行加和操作，得到所述第二信息，所述概率信息包括量化概率，所述 M 为所述量化概率的分母。

47. 根据权利要求 34 至 46 任一所述的装置，其特征在于，所述编码模块，具体用于：

通过第一映射关系，得到所述目标符号对应的一个整数数值；

基于第三运算，将所述整数数值还原为所述第一信息和所述第二信息，所述第一信息和所述第二信息分别为一个整数数值。

48. 根据权利要求 34 至 47 任一所述的装置，其特征在于，所述概率信息包括每个符号对应的 PMF 信息，所述 PMF 信息包括第一数值和第二数值；所述第一数值和所述第二数值的加和为所述 PMF 的数值；所述概率信息还包括每个符号对应的 CDF 信息，所述 CDF 信息包括第三数值和第四数值；所述第三数值和第四数值为通过如下方式构建的：

基于预设顺序，根据多个符号的第一数值，通过累加依次确定每个符号的对应的第三数值；

基于所述预设顺序，根据多个符号的第二数值和第三数值，通过累加依次确定每个符号对应的第四数值。

49. 根据权利要求 48 所述的装置，其特征在于，所述第二信息包括第一子信息、第二子信息和第三子信息；

所述第一映射关系中的所述第一子信息为通过如下方式构建的：将所述目标符号的第三数值和所述目标符号的 PMF 的数值进行减法运算，得到所述第一子信息；

所述第一映射关系中的所述第二子信息为通过如下方式构建的：将所述目标符号的第三数值和所述目标符号的第一数值进行加法运算，将所述加法运算的操作结果与 1 进行减法操作，得到所述第二子信息；

所述第一映射关系中的所述第三子信息为通过如下方式构建的：将所述目标符号的第四数值和所述目标符号的第三数值进行减法运算，将所述减法运算的操作结果与所述目标符号的第一数值进行减法操作，得到所述第三子信息。

50. 根据权利要求 49 所述的装置, 其特征在于, 所述第二运算包括:

对所述状态信息与数值 M 进行的加和操作, 以及对所述加和操作的加和结果进行的比特移位操作; 其中, 所述概率信息包括量化概率, 所述 M 为所述量化概率的分母; 对所述比特移位操作的移位结果与所述第一子信息进行的加和操作;

5 在所述加和操作的加和结果大于所述第二子信息时, 将所述加和操作的加和结果与所述第三子信息的加和结果作为所述第二编码数据;

在所述加和操作的加和结果小于所述第二子信息时, 将所述加和操作的加和结果作为所述第二编码数据。

10 51. 根据权利要求 49 或 50 所述的装置, 其特征在于, 在所述数值 M 为 256 时, 所述第一信息通过 10 比特存储, 所述第一子信息通过 8 比特存储, 所述第二子信息通过 7 比特存储, 所述第三子信息通过 7 比特存储, 多个符号的第一数值的加和为 128, 所述第三子信息小于 128。

52. 一种数据编码装置, 其特征在于, 应用于基于非对称数字系统 ANS 的编码器, 所述装置包括:

15 获取模块, 用于获取第一符号, 所述第一符号对应的量化概率和第二符号对应的量化概率的和为 1, 且所述第一符号对应的量化概率和 1 的差异小于阈值, 所述第二符号对应的量化概率和 0 的差异小于阈值;

编码模块, 用于根据所述编码器的状态信息和数值 M 之间的数值大小关系满足第一条件, 将所述状态信息和所述数值 M 之间进行第一运算, 得到第一编码数据, 并将所述状态信息更新为 0;

20 根据所述编码器的状态信息和数值 M 之间的数值大小关系不满足所述第一条件, 将所述状态信息加 1。

53. 根据权利要求 52 所述的装置, 其特征在于, 所述获取模块, 还用于:

获取所述第二符号;

25 所述编码模块, 还用于将所述编码器的状态信息作为第一编码数据, 并将所述状态信息更新为所述数值 M 与 1 的差值。

54. 一种数据解码装置, 其特征在于, 应用于基于非对称数字系统 ANS 的解码器, 所述装置包括:

30 获取模块, 用于获取所述解码器的状态信息以及存储比特, 所述状态信息与第一比特相对应, 所述第一比特为当前待解码的比特;

解码模块, 用于根据第一映射关系, 确定所述状态信息对应的第一信息、第二信息以及第三信息; 所述第一信息为所述第一比特对应的目标符号, 所述目标符号为所述第一比特的解码结果, 所述第二信息表示从所述存储比特中选择的比特位数, 所述第三信息用于和根据所述第三信息从所述存储比特中选择的比特进行加和运算, 所述加和运算结果用于更新所述状态信息, 更新后的状态信息与第二比特相对应, 所述第二比特为所述第一比特之后的待解码比特。

55. 根据权利要求 54 所述的装置, 其特征在于, 所述解码模块, 还用于:

从所述存储比特中剔除根据所述选择的比特, 得到更新后的存储比特。

40 56. 根据权利要求 54 或 55 所述的装置, 其特征在于, 所述第一信息、所述第二信息或所述第三信息表示一个整数数值。

57. 根据权利要求 54 至 56 任一所述的装置, 其特征在于, 所述第一映射关系中的所述第一信息为通过如下方式构建的:

45 将满足 $C(x) \leq s < (C(x) + P(x))$ 的 x 作为所述第一信息; 其中, 所述 C 表示 CDF 的数值, 所述 P 表示 PMF 的数值, 所述 s 表示所述状态信息; 或者,

将满足 $C1(x) \leq s < C1(x) + P1(x)$ 或 $C2(x) \leq s < C2(x) + P2(x)$ 的 x 作为所述第一信息；其中，所述 $C1$ 表示 CDF 对应的第三数值，所述 $P1$ 表示 PMF 对应的第一数值，所述 $C2$ 表示 CDF 对应的第四数值，所述 $P2$ 表示 PMF 对应的第二数值，所述 s 表示所述状态信息；所述第一数值和所述第二数值的加和为所述 PMF 的数值。

5

58. 根据权利要求 54 至 57 任一所述的装置，其特征在于，所述第一映射关系中的所述第二信息为通过如下方式构建的：

将 $\text{clz}(s - C(x) + P(x)) + X$ 的数值作为所述第二信息；其中，所述 clz 表示取数值中最高位连续为 0 的比特数量，所述 X 为预设的第五数值；或者，

10 将 $s - C(x) + P(x)$ 的数值进行比特移位操作，得到所述第三信息，所述比特移位操作的移位数量为根据所述第二信息得到；或者，

将 $s - C(x) + P(x)$ 的数值进行比特移位操作，并将所述比特移位操作的操作结果减去数值 M ，得到所述第三信息，所述 M 为所述量化概率的分母，所述比特移位操作的移位数量为根据所述第二信息得到；或者，

15 将 $\text{clz}(s' - C(x) + P(x)) + X$ 的数值作为所述第二信息；其中，所述 clz 表示取数值中最高位连续为 0 的比特数量，所述 X 为预设的第五数值，所述 s' 为通过如下得到：

若 $C2(x) \leq s$ ，则 $s' = P1(x) + s - C2(x) + P(x)$ ，否则 $s' = s - C1(x) + P(x)$ 。

20 59. 根据权利要求 54 至 58 任一所述的装置，其特征在于，所述第一映射关系中的所述第三信息为通过如下方式构建的：

将 $s - C(x) + P(x)$ 的数值进行比特移位操作，得到所述第三信息，所述比特移位操作的移位数量为根据所述第二信息得到；或者，

将 $s - C(x) + P(x)$ 的数值进行比特移位操作，并将所述比特移位操作的操作结果减去数值 M ，得到所述第三信息，所述 M 为所述量化概率的分母，所述比特移位操作的移位数量为根据所述第二信息得到；

25 或者，

将 $s' - C(x) + P(x)$ 的数值进行比特移位操作，得到所述第三信息，所述比特移位操作的移位数量为根据所述第二信息得到。

30 60. 根据权利要求 54 至 59 任一所述的装置，其特征在于，所述第一信息、所述第二信息或所述第三信息为根据概率信息得到；所述概率信息包括概率质量函数 PMF 的数值；所述 PMF 的数值为通过如下方式构建的：

依次确定多个符号中每个符号的 PMF 的数值，所述多个符号包括所述目标符号；其中，

在确定所述目标符号的 PMF 的数值时，将除去已确定的符号的 PMF 之外的剩余 PMF 的总数值与所述目标符号的概率进行乘积，并计算所述乘积的乘积结果与除去已确定的符号的剩余概率之间的比值，根据所述比值与预设数值之间的关系，确定所述目标符号的 PMF。

35

61. 根据权利要求 60 所述的装置，其特征在于，所述预设数值为 1，所述关系为取较大值。

40 62. 根据权利要求 54 至 61 任一所述的装置，其特征在于，所述第一信息、所述第二信息或所述第三信息为根据概率信息得到；所述概率信息包括累积分布函数 CDF 的数值；所述 CDF 的数值为通过如下方式构建的：

基于预设顺序，依次确定每个符号的 CDF 的数值，其中所述预设顺序为根据多个符号对应的 PMF 的数值之间的关系确定，所述多个符号包括所述目标符号；其中，

45 在所述目标符号的概率信息对应的概率分布为对称分布时，所述预设顺序为沿着所述对称分布中峰值点的两侧相交错顺序；或者，

在所述目标符号的概率信息对应的概率分布为离散分布时，所述预设顺序具体为从大到小的顺序；

或者，
在所述目标符号的概率信息对应的概率分布为离散分布时，所述预设顺序具体为从小到大的顺序。

63. 根据权利要求 54 至 62 任一所述的装置，其特征在于，所述获取模块，还用于：
5 获取编码结果；所述编码结果为对图像的像素数据进行编码得到的，所述状态信息以及所述存储比特为根据所述编码结果得到；
对所述编码结果进行处理，所述像素数据的分布特征，所述分布特征包括均值或方差；
根据所述分布特征，从多个映射关系中确定与所述分布特征对应的所述第一映射关系。

64. 根据权利要求 54 至 63 任一所述的装置，其特征在于，所述解码模块，具体用于：
10 通过第一映射关系，得到所述状态信息对应的一个整数数值；
将所述整数数值还原为所述第一信息、所述第二信息和所述第三信息，所述第一信息、所述第二信息和所述第三信息分别为一个整数数值。

65. 一种数据解码装置，其特征在于，应用于基于非对称数字系统 ANS 的解码器，所述装置包括：
15 获取模块，用于获取所述解码器的状态信息，所述状态信息与第一比特相对应，所述第一比特为当前待解码的比特；
解码模块，用于若所述状态信息为数值 M 与 1 的差值，则确定所述第一比特的解码结果为第二符号，并将所述状态信息更新为从存储中读取所述数值 M 位的结果；
20 若所述状态信息不为数值 M 与 1 的差值，则确定所述第一比特的解码结果为第一符号，若所述状态信息为 0，则将所述状态信息更新为所述数值 M 与 2 的差值与从存储中读出 1 个比特的值的加和结果；
若所述状态信息不为 0，则将所述状态信息更新为所述状态信息与 1 的差值；更新后的状态信息与第二比特相对应，所述第二比特为所述第一比特之后的待解码比特。

66. 一种计算设备，其特征在于，所述计算设备包括存储器和处理器；所述存储器存储有代码，所述处理器被配置为获取所述代码，并执行如权利要求 1 至 33 任一所述的方法。

67. 一种计算机存储介质，其特征在于，所述计算机存储介质存储有一个或多个指令，所述指令在由一个或多个计算机执行时使得所述一个或多个计算机实施权利要求 1 至 33 任一所述的方法。

68. 一种计算机程序产品，包括代码，其特征在于，在所述代码被执行时用于实现如权利要求 1 至 33 任一所述的方法。