

(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
2019年5月16日 (16.05.2019)



(10) 国际公布号
WO 2019/091318 A1

(51) 国际专利分类号:
G06K 9/62 (2006.01)

中国北京市北京经济技术开发区地泽路9号, Beijing 100176 (CN)。

(21) 国际申请号: PCT/CN2018/113115

(74) 代理人: 北京市柳沈律师事务所 (LIU, SHEN & ASSOCIATES); 中国北京市海淀区彩和坊路10号1号楼10层, Beijing 100080 (CN)。

(22) 国际申请日: 2018年10月31日 (31.10.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201711100238.1 2017年11月9日 (09.11.2017) 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, IR, IS, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(71) 申请人: 京东方科技集团股份有限公司
(BOE TECHNOLOGY GROUP CO., LTD.) [CN/CN];
中国北京市朝阳区酒仙桥路10号,
Beijing 100015 (CN)。

(72) 发明人: 那彦波 (NAVARRETE MICHELINI, Pablo);
中国北京市北京经济技术开发区地泽路9号, Beijing 100176 (CN)。 刘瀚文 (LIU, Hanwen);

(54) **Title:** IMAGE CLASSIFICATION AND CONVERSION METHOD AND DEVICE, IMAGE PROCESSOR AND TRAINING METHOD THEREFOR, AND MEDIUM

(54) 发明名称: 图像分类和转换方法、装置、图像处理器及其训练方法和介质

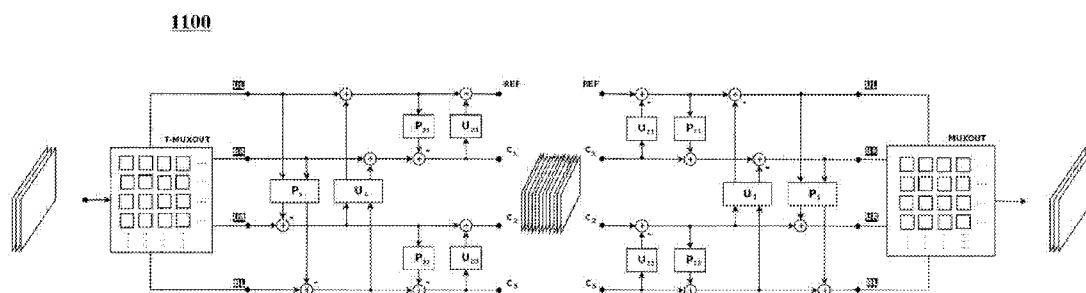


图11

(57) **Abstract:** Disclosed are an image classification and conversion method and device, an image processor and a training method therefor, and a medium. The image classification method comprises: receiving an input image; performing image encoding on the input image by using n stages of cascaded encoding units to generate an output image, with n being an integer greater than 1, wherein with regard to $1 \leq i < n$, an input of an i th-stage encoding unit is an i th-stage encoded input image and comprises m^{i-1} image components, an output of the i th-stage encoding unit is an i th-stage encoded output image and comprises m^i image components, and the output of the i th-stage encoding unit is an input of an $(i+1)$ th-stage encoding unit, with m being an integer greater than 1; and outputting the output image, wherein the output image comprises m^n output sub-images, the m^n output sub-images respectively correspond to m^n output image components of an n th-stage encoding unit, and each of the m^n output sub-images corresponds to an image category.

(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))。

(57) 摘要: 公开了一种图像分类和转换方法、装置、图像处理器及其训练方法和介质。所述图像分类方法包括: 接收输入图像; 利用级联的 n 级编码单元对所述输入图像进行图像编码以产生输出图像, n 为大于1的整数, 对于 $1 \leq i < n$, 第 i 级编码单元的输入为第 i 级编码输入图像并包括 m^{i-1} 个图像成分, 第 i 级编码单元的输出为第 i 级编码输出图像并包括 m^i 个图像成分, 以及第 i 级编码单元的输出是第 $i+1$ 级编码单元的输入, 其中 m 是大于1的整数; 输出所述输出图像, 所述输出图像包括 m^n 个输出子图像, 所述 m^n 个输出子图像分别对应于第 n 级编码单元的 m^n 个输出图像成分, 并且所述 m^n 个输出子图像的每一个对应于一种图像类别。

图像分类和转换方法、装置、图像处理器及其训练方法和介质

相关申请的交叉引用

- 5 本申请要求于2017年11月09日递交的中国专利申请第201711100238.1号的优先权，在此以引用的方式并入上述中国专利申请公开的内容以作为本公开的一部分。

技术领域

- 10 本公开涉及图像处理领域，具体涉及一种图像分类方法、图像分类装置、图像转换方法、包括图像分类装置和图像还原装置的图像处理器及其训练方法和介质。

背景技术

- 15 基于现有的图像分类方法，可以对输入图像进行分析并输出用于该输入图像的标签，该标签代表了输入图像的图像类别。然而，根据现有的图像分类方法可以得到图像类别的信息，但不能获得输入图像中对应于其类别的图像像素信息。

20 发明内容

针对以上问题，本公开提供一种新的用于图像分类和转换的方法、装置、图像处理器及其训练方法和介质。

- 根据本公开的一方面，提出了一种图像分类方法，包括：接收输入图像；利用级联的 n 级编码单元对所述输入图像进行图像编码以产生输出图像， n 为大于1的整数，对于 $1 \leq i < n$ ，第 i 级编码单元的输入为第 i 级编码输入图像并包括 m^{i-1} 个图像成分，第 i 级编码单元的输出为第 i 级编码输出图像并包括 m^i 个图像成分，以及第 i 级编码单元的输出是第 $i+1$ 级编码单元的输入，其中 m 是大于1的整数；输出所述输出图像，所述输出图像包括 m^n 个输出子图像，所述 m^n 个输出子图像的每一个对应于一种图像类别；获取所述输出图像中 m^n 个输出子图像的每一个的像素值，并根据像素值确定所述 m^n 个输出
- 25
- 30

子图像中的至少一个是所述输入图像的分类子图像，以及确定所述输入图像的分类为对应于所述分类子图像的图像类别。

根据本公开的另一方面，提出了一种图像分类装置，包括：输入端，配置成接收输入图像；级联的 n 级编码单元，配置成对所述输入图像进行图像编码以产生输出图像， n 为大于 1 的整数，对于 $1 \leq i < n$ ，第 i 级编码单元的输入为第 i 级编码输入图像并包括 m^{i-1} 个图像成分，第 i 级编码单元的输出为第 i 级编码输出图像并包括 m^i 个图像成分，以及第 i 级编码单元的输出是第 $i+1$ 级编码单元的输入，其中 m 是大于 1 的整数；输出端，配置成输出所述输出图像，所述输出图像包括 m^n 个输出子图像，所述 m^n 个输出子图像的每一个对应于一种图像类别；分类单元，配置成获取所述输出图像中 m^n 个输出子图像的每一个的像素值，并根据像素值确定所述 m^n 个输出子图像中的至少一个是所述输入图像的分类子图像，并确定所述输入图像的分类为对应于所述分类子图像的图像类别。

根据本公开的另一方面，提出了一种图像处理器，包括：图像编码装置，所述图像编码装置包括：编码输入端，配置成接收输入图像；级联的 n 级编码单元，配置成对所述输入图像进行图像编码以产生输出图像， n 为大于 1 的整数，对于 $1 \leq i < n$ ，第 i 级编码单元的输入为第 i 级编码输入图像并包括 m^{i-1} 个图像成分，第 i 级编码单元的输出为第 i 级编码输出图像并包括 m^i 个图像成分，以及第 i 级编码单元的输出是第 $i+1$ 级编码单元的输入，其中 m 是大于 1 的整数；编码输出端，配置成输出所述输出图像，所述输出图像包括 m^n 个输出子图像，所述 m^n 个输出子图像的每一个对应于一种图像类别；图像解码装置，所述图像解码装置包括：解码输入端，配置成所述输入图像的图像成分，所述输入图像的图像成分包括 m^n 个图像成分，其中 m 是大于 1 的整数， n 是大于 1 的整数；级联的 n 级解码单元，配置成对所述解码输入图像进行图像解码以产生还原图像， n 为大于 1 的整数，对于 $1 \leq i < n$ ，第 i 级解码单元的输入为第 i 级解码输入图像并包括 m^i 个图像成分，第 i 级解码单元的输入为第 i 级解码输出图像并包括 m^{i-1} 个图像成分，以及第 i 级解码单元的输入是第 $i+1$ 级解码单元的输入；解码输出端，配置成输出对应于所述解码输入图像的还原图像。

的实施例的图像分类装置可以利用最新的深度学习的发展和表现的优点对输入图像进行分类，并提取输入图像中对应该类别的像素信息。更进一步地，可以使用对应该类别的像素信息对其他图像进行图像类别转换。

5 附图说明

为了更清楚地说明本公开实施例的技术方案，下面将对实施例描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本公开的一些实施例，对于本领域普通技术人员而言，在没有做出创造性劳动的前提下，还可以根据这些附图获得其他的附图。以下附图并未刻意按实际尺寸等比例缩放绘制，重点在于示出本公开的主旨。

图 1 是图示用于图像处理的卷积神经网络的示意图；

图 2 是图示用于图像处理的卷积神经网络的示意图；

图 3 是图示用于多分辨率图像变换的小波变换的示意图；

图 4 是利用卷积神经网络实现小波变换的图像处理器的结构示意图；

15 图 5 示出了根据本公开的实施例的一种图像分类装置的示意图；

图 6 图示了根据本公开的实施例的一种拆分单元的示意图；

图 7 图示了根据本公开的实施例的一种变换单元的示意图；

图 8 示出了根据本公开实施例的一种图像还原装置的示意图；

图 9 示出了根据本公开的实施例的一种逆变换单元的示意图；

20 图 10 示出了根据本公开的实施例的一种复合单元的示意图；

图 11 示意性地示出了对图像进行变换编码和变换解码的过程；

图 12 示出了根据本公开的实施例的一种图像分类方法的流程图；

图 13 示出了根据本公开的实施例的图像编码过程的流程图；

25 图 14 示出了根据本公开的实施例的第 i 级变换编码单元中的图像变换过程的流程图；

图 15A 示出了根据本公开的实施例的第 i 级变换编码单元中的图像变换过程的流程图；

图 15B 示出了根据本公开的实施例的第 i 级变换编码单元中的图像变换过程的流程图；

图 16 示出了根据本公开的实施例的基于更新图像的小波变换的流程图；

5 图 17 示出了根据本公开的实施例的基于差别图像的小波变换的流程图；

图 18 示出了根据本公开的实施例的一种图像还原方法的流程图；

图 19 示出了根据本公开的实施例的第 i 级变换解码单元的图像解码方法的流程图；

10 图 20 示出了根据本公开的实施例的图像逆变换方法的流程图；

图 21A 示出了根据本公开的实施例的第 i 级变换解码单元的图像解码方法的流程图；

图 21B 示出了根据本公开的实施例的第 i 级变换解码单元的图像解码方法的流程图；

15 图 22 示出了基于第一解码输入成分和第二解码输入成分的逆小波变换方法的流程图；

图 23 中示出了基于第三解码输入成分和第四解码输入成分的逆小波变换方法的流程图；以及

图 24 示出了根据本公开的实施例的一种图像处理器的示意图。

20

具体实施方式

为使本公开实施例的目的、技术方案和优点更加清楚，下面将结合本公开实施例的附图，对本公开实施例的技术方案进行清楚、完整地描述。显然，所描述的实施例是本公开的一部分实施例，而不是全部的实施例。基于所描述的本公开的实施例，本领域普通技术人员在无需创造性劳动的前提下所获得的所有其他实施例，都属于本公开保护的范围。

25

除非另外定义，本公开使用的技术术语或者科学术语应当为本公开所属领域内具有一般技能的人士所理解的通常意义。本公开中使用的“第一”、“第二”以及类似的词语并不表示任何顺序、数量或者重要性，而只是用来

区分不同的组成部分。同样，“包括”或者“包含”等类似的词语意指出现该词前面的元件或者物件涵盖出现在该词后面列举的元件或者物件及其等同，而不排除其他元件或者物件。“连接”或者“相连”等类似的词语并非限定于物理的或者机械的连接，而是可以包括电性连接或信号连接，不管是直接的还是间接的。

深度学习中一种常用的网络结构是卷积神经网络。卷积神经网络是经常使用图像作为输入并且使用卷积核来替换权重的神经网络结构。作为示例，一个简单的神经网络结构在图 1 中示出。图 1 示出了一种卷积神经网络的简单示意图。该卷积神经网络例如用于图像处理，使用图像作为输入和输出，例如卷积核替代权重。图 1 中示出了简单结构的卷积神经网络。如图 1 所示，该结构在左侧的四个输入端子处获取 4 个输入图像，在中心的隐藏层 102 具有 3 个单元（输出图像），并且在输出层 103 具有 2 个单元，产生 2 个输出

图像。具有权重 w_{ij}^k 的每个框对应于卷积核（例如， 3×3 或 5×5 内核），其中 k 是指示输入层编号的标签， i 和 j 分别是指示输入和输出单元的标签。偏置（bias） b_i^k 是添加到卷积的输出的标量。加入几个卷积和偏置的结果然后通过激活层的激活函数（Activation Function），激活函数通常对应于整流线性单元（rectifying linear unit, ReLU）或 S 形函数或双曲正切（Sigmoid）函数等。卷积核的权重和偏置在系统的操作期间是固定的，通过使用一组输入/输出示例图像的训练过程获得，并且被调整以适合取决于应用的一些优化准则。典型的配置涉及每层中的数十或数百个卷积核。

图 2 中示出了由于图 1 所示的卷积神经网络中的激活函数的激活结果而等效的示例图。其中假设某个特定的输入，其仅在第一层中的第二 ReLU（对应于图 2 中偏置 b_2^1 所指向的节点）和第二层中的第一 ReLU（对应于图 2 中偏置 b_1^2 所指向的节点）的输出大于 0。对于该特定输入，到其他 ReLU 的输入是 0，因此可以在图 2 中省略。

本公开介绍了一种将深度学习网络用于图像分类的方法，这里的图像分类是可逆的。例如，可以将包括手写数字的图像分类为 10 类：0、1.....9。本公开提供的系统可以输出多个（例如 10 个）被称为“潜在空间”（latent

space) 的小分辨率的图像, 在其中只有一个图像将显示数字, 该数字是对应于输入图像的正确数字。潜在空间的图像可以是相对于输入图像更低分辨率的图像。存在许多用于解决分类问题的系统, 但是在本公开中, 可以利用这些低分辨率的输出并对其进行解码以恢复原始的输入图像。在这种场景中人们可以操作分类的输出用于不同的目的。一种应用是组合潜在空间信息以将对应于一个类别的图像转换为另一个类别。例如, 人们可以将一个数字的图像转换为另一个数字, 或将一个男人的图像转换为一个女人的图像, 或将狗转换为猫, 同时保留与类别无关的其他所有特征。

图 3 是图示用于多分辨率图像变换的小波变换的示意图。小波变换是一种用于图像编解码处理的多分辨率图像变换, 其应用包括 JPEG 2000 标准中的变换编码。在图像编码处理中, 小波变换用于以更小的低分辨率图像 (例如, 原始图像的一部分图像) 代表原始的高分辨率图像。在图像解码处理中, 逆小波变换用于利用低分辨率图像以及恢复原始图像所需的差异特征, 恢复得到原始图像。

图 3 示意性地示出了 3 级小波变换和逆变换。如图 3 所示, 更小的低分辨率图像之一是原始图像的缩小版本 A, 而其他的低分辨率图像代表恢复原始图像所需的细节 (Dh、Dv 和 Dd)。

图 4 是利用卷积神经网络实现小波变换的图像处理器的结构示意图。提升算法 (Lifting Scheme) 是小波变换的一种有效实施方式, 并且是构造小波时的一种灵活的工具。图 4 示意性地示出了用于一维 (1D) 数据的标准结构。图 4 的左侧为编码器 41。编码器 41 中的拆分单元 402 将输入的原始图像 401 变换为低分辨率图像 A 和细节 D。更具体地, 编码器 41 使用预测滤波器 p 和更新滤波器 u。对于压缩应用, 希望细节 D 约为 0, 使得大部分的信息包含在图像 A 中。图 4 的右侧为解码器 42。解码器 42 的参数可以与来自编码器 41 的滤波器 p 和 u 完全相同, 而仅仅是滤波器 p 和 u 相反地布置。由于编码器 41 和解码器 42 的严格对应, 该配置确保了经由解码器 42 的拼接单元 403 拼接得到的解码图像 404 与原始图像 401 完全相同。此外, 图 4 所示的结构也不是限制性的, 可以替代地在编码器和解码器按照交换更新滤波器 u 和预测滤波器 p 的顺序进行配置。在本公开中, 更新滤波器 u 和预测滤波器 p 可以使用如图 1 所示的卷积神经网络实现。

以下，将参照附图进一步详细描述根据本公开实施例的图像分类装置、图像还原装置、包括图像分类装置、图像还原装置的图像处理器、相应的图像分类方法、图像还原方法以及用于配置上述图像处理器的训练方法。

图 5 示出了根据本公开实施例的一种图像分类装置的示意图。图像分类装置 500 包括变换编码单元 510。

变换编码单元 510 可以包括用于接收编码输入图像的输入端，输入图像可以是仅包括单通道（如 R、G、B 或灰度等）的图像，也可以是包括任意多个通道（如 R、G、B 和灰度等）的图像。

变换编码单元 510 还可以包括级联的 n 级编码单元 510-1、510-2、..... 510- n ， n 为大于等于 1 的整数，对于 $1 \leq i < n$ ，第 i 级编码单元的输入为第 i 级编码输入图像，其包括 m^{i-1} 个图像成分，第 i 级编码单元的输入为第 i 级编码输入图像并包括 m^i 个图像成分，以及第 i 级编码单元的输入是第 $i+1$ 级编码单元的输入，其中 m 是大于 1 的整数。

在一些实施例中，级联的 n 级编码单元中的每一个可以包括拆分单元 512 和变换单元 514。也就是说，第 i 级编码单元 510- i 中包括拆分单元 512- i 、变换单元 514- i 。拆分单元 512- i 用于对第 i 级编码单元接收的 m^{i-1} 个图像成分中的每一个执行拆分操作，将第 i 级编码输入图像中的每个图像成分拆分为 m 个图像成分，即将第 i 级编码输入图像中的 m^{i-1} 个图像成分拆分为 m^i 个图像成分。变换单元 514- i 用于对由第 i 级编码单元拆分得到的 m^i 个图像成分进行变换。

变换编码单元 510 还可以包括用于输出编码输出图像的输出端，其中编码输出图像包括对应于编码输入图像的 m^n 个输出图像成分，即 m^n 个输出子图像，并且 m^n 个输出子图像的每一个对应于一种图像类别。

图像分类装置 500 还可以包括分类单元 520，配置成获取编码输出图像中 m^n 个输出子图像的每一个的像素值，并根据像素值确定 m^n 个输出子图像中的至少一个是输入图像的类别子图像，以及确定输入图像的类别为对应于所述类别子图像的图像类别。这里所说的像素值可以是图像中所有像素的像素值的和，在下文中，图像中各像素值的和也可以被简称为该图像的像素值。在一些实施例中，分类单元 520 可以配置成比较 m^n 个输出子图像中每一个的像素值与第一阈值，将像素值大于第一阈值的输出子图像确定为类别子图

像，并输出对应于类别子图像的类别标签，其中类别子图像包括输入图像中对应于类别标签的像素信息。

上述变换编码单元能够对输入图像拆分以及图像变换，并提取图像中的多个图像成分作为输出。输出的每一个图像成分对应于一种图像分类的类别。

5 根据图像分类装置输出的多个图像成分的像素值可以确定输入图像的类别。

在一些实施例中，经过上述图像分类装置的处理后，输出图像中的多个输出子图像中有至少一个输出子图像的像素值不为零，而其他输出子图像的像素值都接近于零，那么可以认为上述像素值不为零的输出子图像代表了输入图像的类别。例如，如果输出图像包括4个输出子图像 C_0 (REF)、 C_1 、 C_2 、

10 C_3 ，如果其中 C_1 的像素值不为零，而 C_0 、 C_2 、 C_3 的像素值都接近于零，则将输入图像的类别确定为 C_1 类。具体 C_1 类代表的图像特征由图像分类装置的配置所决定。经过下文中将介绍的训练方法的配置后，经过编码变换后输出的输出子图像代表了输入图像的类别信息以及对应于类别信息的像素信息。

15 例如， C_0 、 C_1 、 C_2 、 C_3 可以分别代表数字0、1、2、3。当图像分类装置的输入图像中包括数字0的图像时，其输出的4个输出子图像中 C_0 的像素值不为零，而 C_1 、 C_2 、 C_3 的像素值接近于零。并且， C_0 的像素信息可以表示数字0在输入图像中的像素信息，如数字0的图像的形状、大小、位置等。例如，如果在输入图像的左上角存在数字0的图像，那么，在输出的输出子

20 图像 C_0 的左上角存在对应于输入图像中包括的数字0的像素信息。当图像分类装置的输入图像中同时包括数字0和数字2的图像时，输出的4个输出子图像中 C_0 、 C_2 的像素值不为零，而 C_1 、 C_3 的像素值接近于零。并且， C_0 的像素信息可以表示数字0在输入图像中的像素信息，如数字0的图像的形状、大小、位置等。 C_2 的像素信息可以表示数字2在输入图像中的像素信息，如

25 数字2的图像的形状、大小、位置等。例如，如果在图像分类装置的输入图像的左上角存在数字0的图像，右下角存在数字2的图像，那么在输出的输出子图像 C_0 的左上角存在对应于输入图像中包括的数字0的像素信息，在输出的输出子图像 C_2 的右下角存在对应于输入图像中包括的数字2的像素信息。

30 又例如， C_0 、 C_1 、 C_2 、 C_3 可以分别代表男人、女人、猫、狗。那么，类

似地，当输入图像中包括男人和狗时，输出的4个输出子图像中 C_0 、 C_3 的像素值不为零，而 C_1 、 C_2 的像素值接近于零。并且， C_0 的像素信息可以表示输入图像中的男人的像素信息， C_2 的像素信息可以表示输入图像中的狗的像素信息。

5 也就是说，经过训练的图像分类装置能够对输入图像进行分类并提取相应的类别信息。当输入图像中包含符合多个类别的图像信息时，图像分类装置将输出多个对应类别的输出子图像，输出子图像中包括对应于输入图像中该类别的像素信息。

10 经过上述分类过程提取的对应于图像类别的像素信息能够进一步用于图像转换。在下文中将参考图8-图10对图像转换的过程进行进一步的说明。

图6中示例性的示出了一种能够将一张图像拆分为4个分辨率更低的小图像的拆分单元512。如图6所示，拆分单元T-MUXOUT可以将原始图像以 2×2 的基本像素矩阵为单位进行划分，其中每个基本像素矩阵包括4个原始像素。拆分单元512进一步提取所有划分好的 2×2 的基本像素矩阵中的特定位置的像素，并根据每个基本像素矩阵中特定位置的像素确定拆分后的图像。例如，如图6中示出的，拆分单元512的输入图像包括16个原始像素，拆分单元512将输入图像划分为基本像素矩阵 A_{11} 、 A_{12} 、 A_{21} 、 A_{22} ，其中基本像素矩阵 A_{11} 中包括像素 a_{11} 、 b_{11} 、 c_{11} 、 d_{11} ，基本像素矩阵 A_{12} 中包括像素 a_{12} 、 b_{12} 、 c_{12} 、 d_{12} ，基本像素矩阵 A_{21} 中包括像素 a_{21} 、 b_{21} 、 c_{21} 、 d_{21} ，基本像素矩阵 A_{22} 中包括像素 a_{22} 、 b_{22} 、 c_{22} 、 d_{22} 。拆分单元512可以通过提取所有基本像素矩阵中左上角（即[1,1]位置处）的原始像素，如图5中示出的像素 a_{11} 、 a_{12} 、 a_{21} 、 a_{22} ，并将提取的像素按照拆分前各像素在图像中排列的顺序进行排列，以生成第一张拆分后的低分辨率图像。进一步地，拆分单元512可以通过提取所有基本像素矩阵中[1, 2]位置处的原始像素，如图5中示出的像素 b_{11} 、 b_{12} 、 b_{21} 、 b_{22} ，并将提取的像素按照拆分前各像素在图像中排列的顺序进行排列，以生成第二张拆分后的低分辨率图像。类似地，拆分单元可以生成其余的拆分后的低分辨率小图像。

可以理解，如图6所示的拆分单元可以将任意大小的图像拆分为4个分辨率更低的小图像。在一些实施例中，拆分后的多个低分辨率图像尺寸相等。30 例如，如图6所示的拆分单元512可以将原始尺寸为 128×128 的图像拆分

成 4 个尺寸均为 64×64 的低分辨率图像。

也可以理解，如图 6 所示的拆分单元只是根据本公开的原理的拆分单元的一个示例。事实上，可以通过调整划分的基本像素矩阵的大小和形状将图像拆分为多个分辨率更低的小图像。例如，当基本像素矩阵的大小为 3×3 时，拆分单元可以将输入图像拆分为 $3 \times 3 = 9$ 个分辨率更低的小图像。又例如，当基本像素矩阵的大小是 3×4 时，拆分单元可以将输入图像拆分为 $3 \times 4 = 12$ 个分辨率更低的小图像。也就是说，当基本像素矩阵的大小是 $a \times b$ 时，拆分单元可以将输入图像拆分为 $a \times b = c$ 个分辨率更低的小图像。本领域技术人员可以了解，根据本公开的原理，拆分单元 512 可以将一张图像拆分为任意多个分辨率更低的缩小的图像。

也可以理解，图 6 示出的是利用拆分单元对二维的图像数据进行拆分的示意图。根据本公开的原理，拆分单元 512 也可以对其他任意维度（如一维、三维等）的图像数据进行拆分。

为了描述方便，在下文中将以图 6 示出的拆分单元为例进行描述，并将拆分后的四个低分辨率的图像分别称为左上（UL）、右上（UR）、左下（BL）和右下（BR）。也就是说，对于第 i 级编码单元来说，输入图像包括 4^{i-1} 个图像成分，经过第 i 级编码单元中的拆分单元 512- i ，第 i 级编码输入图像被拆分为 4^i 个图像成分。

图 7 图示了根据本公开的实施例的一种变换单元 514 的示意图。如前所述，拆分单元可以将原始图像拆分为 4 个低分辨率图像 UL、UR、BL 和 BR。变换单元 514 可以对上述四个低分辨率图像 UL、UR、BL 和 BR 进行图像变换，从而提取出输入图像中的图像成分。

变换单元 514 可以包括第一预测单元 710，配置成基于 UL 图像和 BR 图像生成用于 UR 图像和 BL 图像的预测图像并获取 UR 图像和 BL 图像和该预测图像之间的差别图像；第一更新单元 720，配置成基于 UR 图像和 BL 图像和预测图像之间的差别图像生成用于 UL 图像和 BR 图像的更新图像；第一小波变换单元 730，配置成执行基于用于 UL 图像和 BR 图像的更新图像的小波变换，并基于小波变换的结果生成第一编码输出成分和第二编码输出成分；第二小波变换单元 740，配置成执行基于 UR 图像和 BL 图像和上述预测图像之间的小波变换，并基于小波变换的结果生成第三编码输出成分

和第四编码输出成分。

在一些实施例中，如图 7 所示，第一预测单元 710 可以进一步包括第一预测卷积网络 P_1 和第一去叠加单元 712。第一预测单元 P_1 配置成接收 UL 图像和 BR 图像作为输入，并生成用于 UR 图像和 BL 图像的第一预测特征和
5 第二预测特征。第一预测特征和第二预测特征可以是相同的，也可以是不同的。第一去叠加单元 712 配置成对 UR 图像和第一预测特征执行去叠加操作以获得第一差别特征，以及对 BL 图像和第二预测特征执行去叠加操作以获得第二差别特征。

在一些实施例中，第一更新单元 720 可以进一步包括第一更新卷积网络
10 U_1 和第一叠加单元 722。第一更新卷积网络 U_1 配置成接收第一差别特征和第二差别特征作为输入，并生成用于 UL 图像和 BR 图像的第一更新特征和第二更新特征。第一更新特征和第二更新特征可以是相同的，也可以是不同的。第一叠加单元 722 配置成对 UL 图像与第一更新特征执行叠加操作以获得第一叠加特征，以及对 BR 图像与第二更新特征执行叠加操作以获得第二叠加
15 特征。

在一些实施例中，第一小波变换单元 730 可以进一步包括第二预测卷积网络 P_{21} ，配置成接收第一叠加特征作为输入，并生成用于第二叠加特征的第三预测特征；第二去叠加单元 732，配置成对第二叠加特征与第三预测特征执行去叠加操作以获得第二编码输出成分；第二更新卷积网络 U_{21} ，配置成
20 接收第二编码输出成分作为输入，并生成用于第一编码输出成分的第三更新特征；第二叠加单元 734，配置成对第一叠加特征与第三更新特征执行叠加操作以获得第一编码输出成分。

在一些实施例中，第二小波变换单元 740 可以进一步包括第三预测卷积网络 P_{22} ，配置成接收第一差别特征作为输入，并生成用于第二差别特征的第四预测特征；第三去叠加单元 742，配置成对第二差别特征与第四预测特征执行去叠加操作以获得第四编码输出成分；第三更新卷积网络 U_{22} ，配置成
25 接收第四编码输出成分作为输入，并生成用于第一差别特征的第四更新特征；第三叠加单元 644，配置成对第一差别特征与第四更新特征执行叠加操作以获得第三编码输出成分。

30 图 7 中示出的结构不是限制性的。例如，可以在变换单元 514 中对换第

一预测单元 710 和第一更新单元 720 的结构。

利用图 7 示出的图像处理装置可以对拆分后的低分辨率图像进行图像变换并提取输入图像中的图像成分。这里的图像变换中不损失图像信息，经过相应的逆变换可以无损的还原图像信息。

5 图 8 示出了根据本公开实施例的一种图像还原装置的示意图。图像还原装置 800 可以包括变换解码单元 810。

图 8 示出的变换解码单元 810 对应于图 5 中示出的变换编码单元，能够将经过变换编码单元 510 变换的图像数据无损地还原为原始数据。

变换解码单元 810 可以包括用于接收解码输入图像的输入端，解码输入
10 图像包括 m^n 个图像成分，其中 m 是大于 1 的整数， n 是大于等于 1 的整数。其中 m^n 个图像成分中的每个图像成分可以包括多个通道（如 RGB 三个通道）。

变换解码单元 810 还可以包括级联的 n 级解码单元 810-1、810-2.....
810- n ，对于 $1 \leq i < n$ ，第 i 级解码单元的输入为第 i 级解码输入图像并包括
15 m^i 个图像成分，第 i 级解码单元的输入为第 i 级解码输入图像并包括 m^{i-1} 个图像成分，以及第 i 级解码单元的输入是第 $i+1$ 级解码单元的输入。

在一些实施例中，级联的 n 级解码单元中的每一个可以包括逆变换单元
812 和复合单元 814。也就是说，第 i 级解码单元 810- i 中包括逆变换单元 812- i
和复合单元 814- i 。逆变换单元用于对第 i 级解码单元的输入的 m^i 个图像成分
20 执行逆变换，从而无损地还原解码输入图像包括的 m^i 个图像成分对应的还原图像。复合单元 814 用于对 m^i 个经过逆变换的解码输出成分执行复合操作，从而将 m^i 个图像成分复合为 m^{i-1} 个图像成分。

变换解码单元 810 还可以包括输出端，配置成输出对应于解码输入图像
中 m^n 个图像成分的还原图像。

25 图 9 示出了根据本公开的实施例的一种逆变换单元 812 的示意图。当 $m=4$ 时，第 i 级解码输入图像包括第一解码输入成分、第二解码输入成分、第三解码输入成分和第四解码输入成分，其中每个解码输入成分包括 4^{i-1} 个图像成分。

逆变换单元 812 可以包括第一逆小波变换单元 930，配置成执行基于第
30 一解码输入成分和第二解码输入成分的逆小波变换，并基于逆小波变换的结

果获得第一差别特征和第一叠加特征；第二逆小波变换单元 940，配置成执行基于第三解码输入成分和第四解码输入成分的逆小波变换，并基于逆小波变换的结果获得第二差别特征和第二叠加特征；第二更新单元 920，配置成基于第二差别特征和第二叠加特征生成更新图像，并基于该更新图像、第一差别特征和第一叠加特征生成第一解码输出成分和第二解码输出成分；第二预测单元 910，配置成基于第一解码输出成分和第二解码输出成分生成预测图像，并基于预测图像、第二差别特征和第二叠加特征生成第三解码输出成分和第四解码输出成分。

在一些实施例中，第二更新单元 920 进一步包括第一更新卷积网络 U'_1 和第一去叠加单元 922。第一更新卷积网络 U'_1 配置成接收第二差别特征和第二叠加特征作为输入，并生成关于第二差别特征和第二叠加特征的第一更新特征和第二更新特征。第一更新特征和第二更新特征可以是相同的，也可以是不同的。第一去叠加单元 922 配置成对第一差别特征和第一更新特征执行去叠加操作以获得第一解码输出成分，以及对第一叠加特征和第二更新特征执行去叠加操作以获得第二解码输出成分。

在一些实施例中，第二预测单元 910 进一步包括第一预测卷积网络 P'_1 和第一叠加单元 912。第一预测卷积网络 P'_1 配置成接收第一解码输出成分和第二解码输出成分作为输入，并生成关于第一解码输出成分和第二解码输出成分的第一预测特征和第二预测特征。第一预测特征和第二预测特征可以是相同的，也可以是不同的。第一叠加单元 912 配置成对第二差别特征和第一预测特征执行叠加操作以获得第三解码输出成分，以及对第二叠加特征和第二预测特征执行叠加操作以获得第四解码输出成分。

在一些实施例中，第一逆小波变换单元 930 可以进一步包括第二更新卷积网络 U'_{21} ，配置成接收第二解码输入成分作为输入，并生成关于第二解码输入成分的第三更新特征；第二去叠加单元 934，配置成对第一解码输入成分和第三更新特征执行去叠加操作以获得第一差别特征；第二预测卷积网络 P'_{21} ，配置成接收第一差别特征作为输入，并生成关于第一差别特征的第三预测特征；第二叠加单元 932，配置成对第二解码输入成分和第三预测特征执行叠加操作以获得第一叠加特征。

在一些实施例中，第二逆小波变换 940 可以进一步包括第三更新卷积网

络 U'_{22} ，配置成接收第四解码输入成分作为输入，并生成关于第四图像解码输入的第四更新特征；第三去叠加单元 942，配置成对第三解码输入成分和第四更新特征执行去叠加操作以获得第二差别特征；第三预测卷积网络 P'_{22} ，配置成接收第二差别特征作为输入，并生成关于第二差别特征的第四预测特征；第三叠加单元 944，配置成对第四解码输入成分和第四预测特征执行叠加操作以获得第二叠加特征。

由于逆变换单元 812 可以用于恢复经过变换单元 514 的处理的图像，因此，在一些实施例中，逆变换单元 812 中的卷积网络与变换单元 514 中的卷积网络完全对应。也就是说，逆变换单元 812 中的第一预测卷积网络 P'_1 、第一更新卷积网络 U'_1 、第二更新卷积网络 U'_{21} 、第二预测卷积网络 P'_{21} 、第三更新卷积网络 U'_{22} 、第三预测卷积网络 P'_{22} 与变换单元 514 中的第一预测卷积网络 P_1 、第一更新卷积网络 U_1 、第二更新卷积网络 U_{21} 、第二预测卷积网络 P_{21} 、第三更新卷积网络 U_{22} 、第三预测卷积网络 P_{22} 具有相同的结构和配置参数。

图 9 中示出的结构是非限制性的。例如，可以在逆变换单元 812 中对换第二预测单元 910 和第二更新单元 920 的结构。

图 10 示出了根据本公开的实施例的一种复合单元的示意图。复合单元可以将多个低分辨率的小图像复合为分辨率更高的复合图像。复合单元配置成执行如前所述的拆分单元的逆变换，从而将拆分后的低分辨率的小图像还原为高分辨率的原始图像。

在一些实施例中，当通过图像分类装置获得输入图像的多个输出子图像后，如果将得到的输出子图像按照输出的顺序输入图像还原装置，那么图像还原装置将输出还原的输入图像。例如，输入图像经过图像分类装置的处理后得到输出子图像 C_0 (REF)、 C_1 、 C_2 、 C_3 ，如果将输出子图像 C_0 (REF)、 C_1 、 C_2 、 C_3 对应地输入逆变换单元的输入端，复合单元 MUXOUT 将依序聪哥输出子图像中提取对应的像素点，并将提取的像素点依次排列以生成还原图像。经过上述逆变换过程和复合操作，图像还原装置将输出与输入图像相同的还原图像。

在另一些实施例中，如果改变图像分类装置输出的输出子图像，并将改变后的多个输出子图像输入图像还原装置，那么图像还原装置将输出与输入

图像不同的还原装置。

例如，如果输入图像的左上角存在数字 0，右下角存在数字 2，那么经过图像分类装置的处理后得到的输出子图像中的 C_0 、 C_2 将包括左上角存在数字 0，右下角存在数字 2 的像素信息。在还原过程中，如果将 C_0 和 C_2 的像素信息对换后输入图像还原装置，那么图像还原装置输出的是将输入图像中数字 0 和数字 2 对换后的图像。

又例如，如果将输入图像经过图像分类装置的处理后得到的输出子图像中的至少一个的像素信息（例如，对应于 0 的像素信息）替换为对应与其他类别的像素信息（例如，对应于 5 的像素信息），那么经过图像还原装置输出的还原图像是将原始的输入图像中的数字 0 替换为数字 5 的图像。

图 11 示意性地示出了对图像进行变换编码和变换解码的过程。在变换编码单元的输入端接收输入图像。如图 11 所示，输入图像可以包括任意多个通道，例如，RGB 三通道。经过第 1 级变换解码单元的处理，输入图像经由拆分单元被拆分为四个分辨率更低的子图像。如上文所述，输入图像可以被拆分为任意多个子图像。通过变换单元对拆分后的子图像进行图像变换，获得图像成分。可以看出，对于包括多个通道的输入图像，如图 11 所示的第 1 级变换编码单元的每个箭头可以处理多个通道的数据。例如，对于包括 RGB 3 个通道输入图像，第 1 级变换编码单元中的每个箭头表示在该次输入或输出中的数据存在 3 个通道。经过第 1 级变换编码单元后，输入图像的每个通道均被变换为四个图像成分。

根据图像处理的实际需要，可以使用多级变换编码单元对图像进行处理。例如，如图 11 所示的输入图像经过 n 级变换编码单元后，可以得到 4^n 个图像成分，其中一个或多个图像成分包含了输入图像类别信息，其余是包含了其他细节信息的图像成分。与类别信息相比，其余图像成分的像素信息接近于零。也就是说，经过如本公开实施例提供的图像变换方法，通过变换后得到的图像成分的像素信息可以确定输入图像的类别，以及对应于该类别的输入图像中的像素信息。

此外，由于每一级变换编码单元都将输入图像拆分为更多的低分辨率的子图像，因此，每一级变换编码单元都可以比上一级变换编码单元具有更多的通道。例如，对于如图 11 中示出的输入图像，第 1 级变换编码单元中的每

个箭头表示输入/输出包括 3 个通道的数据,第 2 级变换编码单元中的每个箭头表示输入/输出包括 12 个通道的数据,依次类推,第 n 级变换编码单元中的每个箭头表示输入/输出包括 $3 \times 4^{n-1}$ 个通道的数据。

如上所述的图像变换编码过程是可逆的，对应于 n 级变换编码单元，使用相同配置的 n 级变换解码单元可以在不丢失图像信息的情况下还原输入图像。每一级变换解码单元用于对输入的多个图像成分进行逆变换，并对变换后的图像成分执行复合操作，将图像成分还原为分辨率更高的图像成分。对应于编码过程，经过相同级数的解码过程的处理，可以将多个图像成分还原为原始的输入图像。在此不再赘述。

图 12 示出了根据本公开的实施例的一种图像分类方法的流程图。可以利用如图 5 所示的图像分类装置执行图像分类方法 1200。在步骤 S1202 中，接收输入图像。然后，在步骤 S1204 中，利用级联的 n 级编码单元对输入图像进行图像编码以产生输出图像， n 为大于等于 1 的整数，对于 $1 \leq i < n$ ，第 i 级编码单元的输入为第 i 级编码输入图像并包括 m^{i-1} 个图像成分，第 i 级编码单元的输入为第 i 级编码输入图像并包括 m^i 个图像成分，以及第 i 级编码单元的输入是第 $i+1$ 级编码单元的输入，其中 m 是大于 1 的整数。在步骤 S1206 中，输出所述输出图像，所述输出图像包括 m^n 个输出子图像，所述 m^n 个输出子图像分别对应于第 n 级编码单元的 m^n 个输出图像成分，并且所述 m^n 个输出子图像的每一个对应于一种图像类别。在步骤 S1208 中，获取输出图像中 m^n 个输出子图像的每一个的像素值，根据像素值确定 m^n 个输出子图像中的一个为输入图像的类别子图像。在一些实施例中，步骤 S1208 包括比较所述 m^n 个输出子图像中每一个的像素值与第一阈值，将像素值大于第一阈值的输出子图像确定为类别子图像，并输出对应于所述类别子图像的类别标签，其中所述类别子图像包括输入图像中对应于所述类别标签的像素信息。在步骤 S1210 中，确定输入图像的类别为对应于所述类别子图像的图像类别。

上述图像分类方法能够对输入图像进行分类，确定输入图像的类别，并同时输出对应于输入图像的图像类别的像素信息。

图 13 示出了根据本公开的实施例的第 i 级变换编码单元的图像编码过程的流程图。可以利用如图 5 中示出的变换编码单元 510-i 执行图像编码过程 1300。在步骤 S1302 中，接收第 i 级编码输入图像。在步骤 S1304 中，对于

第 i 级编码输入图像中的每个图像成分，将该图像成分拆分为 m 个编码输入成分。在步骤 S1306 中，对于第 i 级编码输入图像中的每个图像成分，对从该图像成分拆分得到的 m 个编码输入成分进行图像变换，生成对应于该图像成分的 m 个编码输出成分。在步骤 S1308 中，输出对应于第 i 级编码输入的 m^{i-1} 个图像成分的 m^i 个编码输出成分作为第 i 级编码输出图像。

图 14 示出了当 $m=4$ 时，根据本公开的实施例的第 i 级变换编码单元中的图像变换过程的流程图。可以使用如图 5 或图 7 中示出的变换单元 514 执行图像变换过程 1400。

当 $m=4$ 时，第 i 级编码输入图像中的每个图像成分被拆分为第一编码输入成分、第二编码输入成分、第三编码输入成分和第四编码输入成分。因此在步骤 S1402 中，变换单元 514 接收第一编码输入成分、第二编码输入成分、第三编码输入成分和第四编码输入成分。在步骤 S1404 中，基于第一编码输入成分和第二编码输入成分生成预测图像并获取第三编码输入成分和第四编码输入成分和预测图像的差别图像。

其中，如图 15A 中所示出的，步骤 S1404 可以进一步包括：在步骤 S1502 中，利用将第一编码输入成分和第二编码输入成分作为输入的第一预测卷积网络 P_1 生成第一预测特征和第二预测特征。第一预测特征和第二预测特征可以是相同的，也可以是不同的。在步骤 S1504 中，对第三编码输入成分和第一预测特征执行去叠加操作以获得第一差别特征。在步骤 S1506 中，对第四编码输入成分和第二预测特征执行去叠加操作以获得第二差别特征。

在步骤 S1406 中，基于差别图像、第一编码输入成分和第二编码输入成分生成第一编码输入成分和第二编码输入成分的更新图像。

其中，如图 15B 中示出的，步骤 S1404 可以进一步包括：在步骤 S1508 中，利用将第一差别特征和第二差别特征作为输入的第一更新卷积网络 U_1 生成第一更新特征和第二更新特征。第一更新特征和第二更新特征可以是相同的，也可以是不同的。在步骤 S1510 中，对第一编码输入成分与第一更新特征执行叠加操作以获得第一叠加特征。在步骤 S1512 中，对第二编码输入成分与第二更新特征执行叠加操作以获得第二叠加特征。

在步骤 S1408 中，执行基于更新图像的小波变换，并基于小波变换的结果生成第一编码输出成分和第二编码输出成分。

在步骤 S1410 中, 执行基于差别图像的小波变换, 并基于小波变换的结果生成第三编码输出成分和第四编码输出成分。

图 16 示出了根据本公开的实施例的基于更新图像的小波变换的流程图。可以利用图 7 中示出的第一小波变换单元 730 实施基于更新图像的小波变换 1600。在步骤 S1602 中, 利用将第一叠加特征作为输入的第二预测卷积网络 P_{21} 生成关于第一叠加特征的第三预测特征。在步骤 S1604 中, 对第二叠加特征与第三预测特征执行去叠加操作以获得第二编码输出成分。在步骤 S1606 中, 利用将第二编码输出成分作为输入的第二更新卷积网络 U_{21} 生成关于第二编码输出成分的第三更新特征。在步骤 S1608 中, 对第一叠加特征与第三更新特征执行叠加操作以获得第一编码输出成分。

图 17 示出了根据本公开的实施例的基于差别图像的小波变换的流程图。可以利用图 7 中示出的第二小波变换单元 740 实施基于差别图像的小波变换 1700。在步骤 S1702 中, 利用将第一差别特征作为输入的第二预测卷积网络 P_{22} 生成第四预测特征。在步骤 S1704 中, 对第二差别特征与第四预测特征执行去叠加操作以获得第四编码输出成分。在步骤 S1706 中, 利用将第四编码输出成分作为输入的第二更新卷积网络 U_{22} 生成第四更新特征。在步骤 S1708 中, 对第一差别特征与第四更新特征执行叠加操作以获得第三编码输出成分。

根据本公开的实施例提供的图像变换方法, 可以提取输入图像中的多个图像成分, 并且基于提取后的多个图像成分的像素值确定输入图像的分类。

图 18 示出了根据本公开的实施例的一种图像还原方法的流程图。可以利用如图 8 所示的图像还原装置执行图像还原方法 1800。在步骤 S1802 中, 接收输入图像, 输入图像包括 m^n 个图像成分。在步骤 S1804 中, 利用级联的 n 级解码单元对输入图像进行图像解码以产生还原图像, 对于 $1 \leq i < n$, 第 i 级解码单元的输入为第 i 级解码输入图像并包括 m^i 个图像成分, 第 i 级解码单元的输入为第 i 级解码输出图像并包括 m^{i-1} 个图像成分, 以及第 i 级解码单元的输出是第 $i+1$ 级解码单元的输入。在步骤 S1806 中, 输出对应于输入图像的还原图像。

为了无损地还原图像, 图像还原方法 1800 是对应于图像分类方法 1200 的。也就是说, 例如, 当图像分类方法 1200 中包括 n 级编码单元时, 图像还原方法 1800 中也相应地包括 n 级解码单元。

图 19 示出了根据本公开的实施例的第 i 级变换解码单元的图像解码方法的流程图。可以利用如图 8 中示出的变换解码单元 810 执行图像解码方法 1900。在步骤 S1902 中，接收第 i 级解码输入图像，其中第 i 级输入图像包括 m^i 个输入子图像。在步骤 S1904 中，对 m^i 个图像成分进行图像逆变换，
5 生成对应于第 i 级解码输入图像的 m^i 个解码输出成分。在步骤 S1906 中，将 m^i 个解码输出成分复合为 m^{i-1} 个解码输出子图像。在步骤 S1908 中，将对应于第 i 级解码输入图像的 m^i 个图像成分的 m^{i-1} 个解码输出子图像输出作为第 i 级解码输出图像。

图 20 示出了当 $m=4$ 时，根据本公开的实施例的图像逆变换方法的流程图。可以利用如图 8 或图 9 中示出的逆变换单元 812 执行图像逆变换方法 2000。在步骤 S2002 中，逆变换单元 812 接收第一解码输入成分、第二解码输入成分、第三解码输入成分以及第四解码输入成分。在步骤 S2004 中，执行基于第一解码输入成分和第二解码输入成分的逆小波变换，并基于逆小波变换的结果获得第一差别特征和第一叠加特征。在步骤 S2006 中，执行基于
10 第三解码输入成分和第四解码输入成分的逆小波变换，并基于逆小波变换的结果获得第二差别特征和第二叠加特征。

在步骤 S2008 中，基于第二差别特征和第二叠加特征生成更新图像，并基于更新图像、第一差别特征和第一叠加特征生成第一解码输出成分和第二解码输出成分。

如图 21A 中所示出的，步骤 S2008 可以进一步包括：在步骤 S2102 中，利用将第二差别特征和第二叠加作为输入的第一更新卷积网络 U'_1 生成第一更新特征和第二更新特征。第一更新特征和第二更新特征可以是相同的，也可以是不同的。在步骤 S2104 中，对第一差别特征和第一更新特征执行去叠加操作以获得第一解码输出成分。在步骤 S2106 中，对第一叠加特征和第二
20 更新特征执行去叠加操作以获得第二解码输出成分。

在步骤 S2010 中，基于第一解码输出成分和第二解码输出成分生成预测图像，并基于预测图像、第二差别特征和第二叠加特征生成第三解码输出成分和第四解码输出成分。

如图 21B 中所示出的，步骤 S2010 可以进一步包括：在步骤 S2108 中，
30 利用将第一解码输出成分和第二解码输出成分作为输入的第一预测卷积网络

P'_1 生成第一预测特征和第二预测特征。第一预测特征和第二预测特征可以是相同的，也可以是不同的。在步骤 S2110 中，对第二差别特征和第一预测特征执行叠加操作以获得第三解码输出成分。在步骤 S2106 中，对第二叠加特征和第二预测特征执行叠加操作以获得第四解码输出成分。

5 图 22 示出了基于第一解码输入成分和第二解码输入成分的逆小波变换方法的流程图。可以利用如图 9 中示出的逆小波变换单元 930 执行逆小波变换方法 2200。在步骤 S2202 中，利用将第二解码输入成分作为输入的第二更新卷积网络 U'_{21} 生成第三更新特征。在步骤 S2204 中，对第一解码输入成分和第三更新特征执行去叠加操作以获得第一差别特征。在步骤 S2206 中，利用将第一差别特征作为输入的第二预测卷积网络 P'_{21} 生成第三预测特征。在
10 步骤 S2208 中，对第二解码输入成分和第三预测特征执行叠加操作以获得第一叠加特征。

图 23 中示出了基于第三解码输入成分和第四解码输入成分的逆小波变换方法的流程图。可以利用如图 9 中示出的逆小波变换单元 940 执行逆小波变换方法 2300。在步骤 S2302 中，利用将第四解码输入成分作为输入的第三更新卷积网络 U'_{22} 生成第四更新特征。在步骤 S2304 中，对第三解码输入成分和第四更新特征执行去叠加操作以获得第二差别特征。在步骤 S2306 中，利用将第二差别特征作为输入的第二预测卷积网络 P'_{22} 生成第四预测特征。步骤 S2308 中，对第四解码输入成分和第四预测特征执行叠加操作以获得第
15 二叠加特征。
20

利用本公开的实施例提供的图像还原方法，可以在不丢失信息的情况下将图像成分还原为原始图像。

图 24 示出了根据本公开的一种实施例的一种图像处理器的示意图。如图 24 所示，图像处理器 2400 的前半部分可以是如图 5 所示的图像分类装置中的变换编码单元，用于从输入图像中提取多个图像成分。图像处理器 2400 的后半部分可以是如图 8 所示的图像还原装置，用于还原图像成分。利用图像处理器 2400 可以实现对图像的分类及还原过程，图像分类装置和还原装置的具体
25 的结构已在上文中详细阐述，在此不再赘述。

利用深度学习方法可以实现对图像处理器 2400 中各卷积网络的参数的
30 配置。

根据本公开的实施例的训练方法包括以下步骤:

将训练图像输入图像处理器, 调整 n 级编码单元和 n 级解码单元中各卷积层中各卷积网络的权值, 运行有限次迭代以使目标函数最优化。

对于如图 24 所示出的图像处理器, 针对每一级编码单元和解码单元输入训练图像。例如, 在图像处理器的输入端输入原始的高分辨率图像 HR 图像。

在一些实施例中, 目标函数可以包括图像处理器中的编码损失、解码损失、风格损失、以及权重正则化系数中的一项或任意几项的和。下文中将介绍上述损失函数的计算方法。

在 HR 图像经过第 1 级编码单元的处理后, 计算第 1 级编码单元输出的参考图像 REF_1 和第 1 级编码单元的训练图像 LR_1 之间的编码损失。上述编码损失可以通过式 (1) 中示出的编码损失函数 L_{ENCk} 进行计算:

$$L_{ENCk} = \|REF_k - LR_k\|_2 + \sum_{\substack{i=1 \\ i \neq C_k}}^{C_0 \cdot m^k} \|C_{ki}\|_1 \quad (1)$$

在这里 $\|x\|_2 = \sqrt{\sum_i x_i^2}$, $\|x\|_1 = \sum_i |x_i|$ 。

其中 REF_k 是第 k 级编码单元输出的第一图像成分, LR_k 是第 k 级编码单元的训练图像, 其中 LR_k 是所述图像处理器的训练图像的下采样图像, 并具有与 REF_k 相同的尺寸; C_0 是所述训练图像的数量; C_{ki} 是第 k 级编码单元输出的图像成分, 其中 $1 \leq i \leq m^k - 1$ 。

相应地, 在解码过程中可以计算第 k 级解码单元输出的参考图像 REF_k 和第 k 级解码单元的训练图像之间的解码损失。解码单元的输入的训练图像包括 $M = m \cdot n$ 个图像成分, 其中一个图像成分是 HR 图像的下采样图像, 其余图像成分均为 0。

上述解码损失可以通过式 (2) 中示出的解码损失函数 L_{DECk} 进行计算:

$$L_{DECk} = IQ(REF_k, LR_k) \quad (2)$$

其中 IQ 函数评价 REF_k 与 LR_k 之间的差别。在一些实施例中, IQ 函数可以是 MSE 函数:

$MSE(X, Y) = \|X - Y\|_2$, 其中 X、Y 分别代表 REF_k 与 LR_k 的图像数据。

在一些实施例中，IQ 函数可以是 SSIM 函数：

$$SSIM(X, Y) = \frac{(2\mu_X\mu_Y + c_1)(2\sigma_{XY} + c_2)}{(\mu_X^2 + \mu_Y^2 + c_1)(\sigma_X^2 + \sigma_Y^2 + c_2)},$$

其中 X、Y 分别代表 REF_k 与 LR_k

的图像数据。 μ_X 和 μ_Y 分别代表 X 和 Y 的平均值， σ_X 和 σ_Y 分别代表 X 和 Y 的标准偏差， $c_1 = (0.01 \times D)^2$ ， $c_2 = (0.03 \times D)^2$ ，D 表示图像的动态范围，例如，

5 对于浮点数来说，D 的值通常是 1.0。

此外，根据第 i 级编码单元的输出以及对应一级的解码单元的输入可以计算这一级的风格损失函数。例如，可以根据第 1 级编码单元的输出以及第 n 级解码单元的输入计算第 1 级的风格损失函数。根据第 2 级编码单元的输出以及第 n-1 级解码单元的输入可以计算第 2 级的风格损失函数。风格损失函数可以通过式 (3) 定义：

$$L_{STYLE_k}(X, Y) = \|G_X - G_Y\|_2, \quad (3)$$

其中对于具有 m 个通道的图像成分 F，

$$G_F(k, l) = \frac{1}{hw} \sum_{i=1}^m F^k(i) F^l(i);$$

其中 G_X 、 G_Y 分别是 X 图像、Y 图像的格拉姆矩阵的特征量，X 是第 k 级编码单元的输出图像，Y 是第 i+1-k 级编码单元的输出图像，其中 $1 \leq k \leq n$ 。

此外，系统的权重正则化系数可以通过式 (4) 定义：

$$L_{REG} = \frac{\|W\|_1}{\|b\|_1} \quad (4)$$

其中 W 是图像处理器中所有卷积网络的权重参数，b 是图像处理器中所有卷积网络的偏置。

20 基于以上损失函数中的一项或多项可以计算图像处理器的总损失函数。可以将图像处理器的总损失函数应用于任何深度学习的优化策略，如随机梯度下降 SGD 或其变型（如动量 SGD、Adam、RMSProp 等）。

通过本公开的实施例提供的图像处理器的训练方法，可以利用深度学习的策略对图像处理器中的卷积神经网络进行参数配置。通过计算训练图像与图像处理器中生成的图像之间的损失函数作为目标函数，调整图像处理器中卷积神经网络的参数使得目标函数最优化，从而实现更好的图像分类效果。

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

最后，还需要说明的是，上述一系列处理不仅包括以这里所述的顺序按时间序列执行的处理，而且包括并行或分别地、而不是按时间顺序执行的处理。

10 通过以上的实施方式的描述，本领域的技术人员可以清楚地了解到本公开可借助软件加必需的硬件平台的方式来实现，当然也可以全部通过硬件来实施。基于这样的理解，本公开的技术方案对背景技术做出贡献的全部或者部分可以以软件产品的形式体现出来，该计算机软件产品可以存储在存储介质中，如 ROM/RAM、磁碟、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，服务器，或者网络设备等等）执行本公开各个实施例或者实施例的某些部分所述的方法。

15 以上对本公开进行了详细介绍，本文中应用了具体实施例对本公开的原理及实施方式进行了阐述，以上实施例的说明只是用于帮助理解本公开的方法及其核心思想；同时，对于本领域的一般技术人员，依据本公开的思想，20 在具体实施方式及应用范围上均会有改变之处，综上所述，本说明书内容不应理解为对本公开的限制。

权利要求书

1. 一种图像分类方法, 包括:

接收输入图像;

5 利用级联的 n 级编码单元对所述输入图像进行图像编码以产生输出图像,
 n 为大于 1 的整数,

 对于 $1 \leq i < n$, 第 i 级编码单元的输入为第 i 级编码输入图像并包括 m^{i-1}
个图像成分, 第 i 级编码单元的输出为第 i 级编码输出图像并包括 m^i 个图像
成分, 以及第 i 级编码单元的输出是第 $i+1$ 级编码单元的输入, 其中 m 是大
10 于 1 的整数;

 输出所述输出图像, 所述输出图像包括 m^n 个输出子图像, 所述 m^n 个输
出子图像分别对应于第 n 级编码单元的 m^n 个输出图像成分, 并且所述 m^n 个
输出子图像的每一个对应于一种图像类别;

 获取所述输出图像中 m^n 个输出子图像的每一个的像素值, 并根据像素值
15 确定所述 m^n 个输出子图像中的至少一个是所述输入图像的类别子图像, 以及
 确定所述输入图像的类别为对应于所述类别子图像的图像类别。

2. 如权利要求 1 所述的图像分类方法, 其中根据所述 m^n 个输出子图像中
每一个的像素值, 确定输入图像的类别包括:

 比较所述 m^n 个输出子图像中每一个子图像中的每个像素点的像素值之和
20 与第一阈值, 将其像素值之和大于第一阈值的输出子图像确定为类别子图像,
 并输出对应于所述类别子图像的类别标签, 其中所述类别子图像包括输入图
像中对应于所述类别标签的像素信息。

3. 如权利要求 1-2 所述的图像分类方法, 其中所述输入图像、所述 m^n 个
输出子图像的每一个包括 R、G、B 三个通道。

25 4. 如权利要求 1-3 所述的图像分类方法, 其中, 利用第 i 级编码单元进行
图像编码包括:

 接收第 i 级编码输入图像;

 对于所述第 i 级编码输入图像中的每个图像成分, 将该图像成分拆分为 m
个编码输入成分, 其中所述 m 个编码输入成分中的每一个的尺寸是第 i 级编

码输入图像中的每个图像成分的尺寸的 $1/m$ 倍, 对所述 m 个编码输入成分进行图像变换, 生成对应于该图像成分的 m 个编码输出成分, 其中所述 m 个编码输出成分中的每一个的尺寸与所述 m 个编码输入成分中的每一个的尺寸相同;

- 5 输出对应于第 i 级编码输入的 m^{i-1} 个图像成分的 m^i 个编码输出成分作为第 i 级编码输出图像。

5. 如权利要求 4 所述的图像分类方法, 其中 $m=4$, 所述第 i 级编码输入图像中的每个图像成分被拆分为第一编码输入成分、第二编码输入成分、第三编码输入成分和第四编码输入成分, 利用第 i 级编码单元对所述第一编码输入成分、所述第二编码输入成分、所述第三编码输入成分和所述第四编码输入成分进行图像变换包括:

接收所述第一编码输入成分、所述第二编码输入成分、所述第三编码输入成分和所述第四编码输入成分;

15 基于所述第一编码输入成分和所述第二编码输入成分生成预测图像并获取所述第三编码输入成分和所述第四编码输入成分和所述预测图像的差别图像;

基于所述差别图像、所述第一编码输入成分和所述第二编码输入成分生成关于所述第一编码输入成分和所述第二编码输入成分的更新图像;

20 执行基于所述更新图像的小波变换, 并基于小波变换的结果生成第一编码输出成分和第二编码输出成分;

执行基于所述差别图像的小波变换, 并基于小波变换的结果生成第三编码输出和第四编码输出。

6. 如权利要求 5 所述的图像分类方法,

25 其中, 基于所述第一编码输入成分和所述第二编码输入成分生成预测图像并获取所述第三编码输入成分和所述第四编码输入成分和所述预测图像的差别图像包括:

利用将所述第一编码输入成分和所述第二编码输入成分作为输入的第一预测卷积网络生成第一预测特征和第二预测特征;

对所述第三编码输入成分和所述第一预测特征执行去叠加操作以获得第

一差别特征;

对所述第四编码输入成分和所述第二预测特征执行去叠加操作以获得第二差别特征;

5 其中, 基于所述差别图像、所述第一编码输入成分和所述第二编码输入成分生成更新图像包括:

利用将所述第一差别特征和所述第二差别特征作为输入的第一更新卷积网络生成第一更新特征和第二更新特征;

对所述第一编码输入成分与所述第一更新特征执行叠加操作以获得第一叠加特征;

10 对所述第二编码输入成分与所述第二更新特征执行叠加操作以获得第二叠加特征。

7. 如权利要求 6 所述的图像分类方法, 其中执行所述基于所述更新图像的小波变换, 并基于小波变换的结果生成第一编码输出成分和第二编码输出成分包括:

15 利用将所述第一叠加特征作为输入的第二预测卷积网络生成第三预测特征;

对所述第二叠加特征与所述第三预测特征执行去叠加操作以获得所述第二编码输出成分;

20 利用将所述第二编码输出成分作为输入的第二更新卷积网络生成第三更新特征;

对所述第一叠加特征与所述第三更新特征执行叠加操作以获得所述第一编码输出成分。

25 8. 如权利要求 6-7 任一项所述的图像分类方法, 其中执行所述基于所述差别图像的小波变换, 并基于小波变换的结果生成第三编码输出成分和第四编码输出成分包括:

利用将所述第一差别特征作为输入的第三预测卷积网络生成第四预测特征;

对所述第二差别特征与所述第四预测特征执行去叠加操作以获得所述第四编码输出成分;

利用将所述第四编码输出成分作为输入的第三更新卷积网络生成第四更新特征;

对所述第一差别特征与所述第四更新特征执行叠加操作以获得所述第三编码输出成分。

5 9. 一种图像分类装置, 包括:

输入端, 配置成接收输入图像;

级联的 n 级编码单元, 配置成对所述输入图像进行图像编码以产生输出图像, n 为大于 1 的整数, 对于 $1 \leq i < n$, 第 i 级编码单元的输入为第 i 级编码输入图像并包括 m^{i-1} 个图像成分, 第 i 级编码单元的输入为第 i 级编码输出图像并包括 m^i 个图像成分, 以及第 i 级编码单元的输入是第 $i+1$ 级编码单元的输入, 其中 m 是大于 1 的整数;

输出端, 配置成输出所述输出图像, 所述输出图像包括 m^n 个输出子图像, 所述 m^n 个输出子图像分别对应于第 n 级编码单元的 m^n 个输出图像成分, 并且所述 m^n 个输出子图像的每一个对应于一种图像类别;

15 分类单元, 配置成获取所述输出图像中 m^n 个输出子图像的每一个的像素值, 并根据像素值确定所述 m^n 个输出子图像中的至少一个是所述输入图像的类别子图像, 并确定所述输入图像的类别为对应于所述类别子图像的图像类别。

10. 如权利要求 9 所述的图像分类装置, 其中分类单元进一步配置成:

20 比较所述 m^n 个输出子图像中每一个子图像中的各像素的像素值之和与第一阈值, 将像素值之和大于第一阈值的输出子图像确定为类别子图像, 并输出对应于所述类别子图像的类别标签, 其中所述类别子图像包括输入图像中对应于所述类别标签的像素信息。

11. 如权利要求 9-10 任一项所述的图像分类装置, 其中所述输入图像、
25 所述 m^n 个输出子图像中的每一个包括 R、G、B 三个通道。

12. 如权利要求 9-11 任一项所述的图像分类装置, 其中第 i 级编码单元包括:

输入端, 配置成接收第 i 级编码输入图像;

拆分单元, 配置成对于所述第 i 级编码输入图像中的每个图像成分, 将该

图像成分拆分为 m 个编码输入成分，其中所述 m 个编码输入成分中的每一个的尺寸是第 i 级编码输入图像中的每个图像成分的尺寸的 $1/m$ 倍；

变换单元，配置成对于所述第 i 级编码输入图像中的每个图像成分，对从该图像成分拆分得到的所述 m 个编码输入成分进行图像变换，生成对应于该
5 图像成分的 m 个编码输出成分，其中所述 m 个编码输出成分中的每一个的尺寸与所述 m 个编码输入成分中的每一个的尺寸相同；

编码输出端，配置成输出对应于第 i 级编码输入的 m^{i-1} 个图像成分的 m^i 个编码输出成分作为第 i 级编码输出图像。

13. 如权利要求 12 所述的图像分类装置，其中， $m=4$ ，所述第 i 级编码
10 输入图像中的每个图像成分被拆分为第一编码输入成分、第二编码输入成分、第三编码输入成分和第四编码输入成分，

第 i 级编码单元的变换单元进一步包括：

第一预测单元，配置成基于所述第一编码输入成分和所述第二编码输入成分生成预测图像并获取所述第三编码输入成分和所述第四编码输入成分和
15 所述预测图像的差别图像；

第一更新单元，配置成基于所述差别图像、所述第一编码输入成分和所述第二编码输入成分生成关于所述第一编码输入成分和所述第二编码输入成分的更新图像；

第一小波变换单元，配置成执行基于所述更新图像的小波变换，并基于
20 小波变换的结果生成第一编码输出成分和第二编码输出成分；

第二小波变换单元，配置成执行基于所述差别图像的小波变换，并基于小波变换的结果生成第三编码输出成分和第四编码输出成分。

14. 如权利要求 13 所述的图像分类装置，其中

第一预测单元进一步包括：

25 第一预测卷积网络，配置成接收所述第一编码输入成分和所述第二编码输入成分作为输入，并生成第一预测特征和第二预测特征；

第一去叠加单元，配置成对所述第三编码输入成分和所述第一预测特征执行去叠加操作以获得第一差别特征，以及对所述第四编码输入成分和所述第二预测特征执行去叠加操作以获得第二差别特征；

第一更新单元进一步包括:

第一更新卷积网络,配置成接收所述第一差别特征和所述第二差别特征作为输入,并生成第一更新特征和第二更新特征;

5 第一叠加单元,配置成对所述第一编码输入成分与所述第一更新特征执行叠加操作以获得第一叠加特征,以及对所述第二编码输入成分与所述第二更新特征执行叠加操作以获得第二叠加特征。

15. 如权利要求 14 所述的图像分类装置,其中所述第一小波变换单元进一步包括:

10 第二预测卷积网络,配置成接收所述第一叠加特征作为输入,并生成第三预测特征;

第二去叠加单元,配置成对所述第二叠加特征与所述第三预测特征执行去叠加操作以获得所述第二编码输出成分;

第二更新卷积网络,配置成接收所述第二编码输出成分作为输入,并生成第三更新特征;

15 第二叠加单元,配置成对所述第一叠加特征与所述第三更新特征执行叠加操作以获得所述第一编码输出成分。

16. 如权利要求 14-15 所述的图像分类装置,其中所述第二小波变换单元进一步包括:

20 第三预测卷积网络,配置成接收所述第一差别特征作为输入,并生成第四预测特征;

第三去叠加单元,配置成对所述第二差别特征与所述第四预测特征执行去叠加操作以获得所述第四编码输出成分;

第三更新卷积网络,配置成接收所述第四编码输出成分作为输入,并生成第四更新特征;

25 第三叠加单元,配置成对所述第一差别特征与所述第四更新特征执行叠加操作以获得所述第三编码输出成分。

17. 一种图像处理器,包括:

图像编码装置,所述图像编码装置包括:

编码输入端,配置成接收输入图像;

级联的 n 级编码单元, 配置成对所述输入图像进行图像编码以产生输出图像, n 为大于 1 的整数, 对于 $1 \leq i < n$, 第 i 级编码单元的输入为第 i 级编码输入图像并包括 m^{i-1} 个图像成分, 第 i 级编码单元的输出为第 i 级编码输出图像并包括 m^i 个图像成分, 以及第 i 级编码单元的输出是第 $i+1$ 级编码单元的输入, 其中 m 是大于 1 的整数;

编码输出端, 配置成输出所述输出图像, 所述输出图像包括 m^n 个输出子图像, 所述 m^n 个输出子图像分别对应于第 n 级编码单元的 m^n 个输出图像成分, 并且所述 m^n 个输出子图像的每一个对应于一种图像类别;

图像解码装置, 所述图像解码装置包括:

解码输入端, 配置成接收解码输入图像, 所述解码输入图像包括 m^n 个图像成分, 其中 m 是大于 1 的整数, n 是大于 1 的整数;

级联的 n 级解码单元, 配置成对所述解码输入图像进行图像解码以产生还原图像, n 为大于 1 的整数, 对于 $1 \leq i < n$, 第 i 级解码单元的输入为第 i 级解码输入图像并包括 m^i 个图像成分, 第 i 级解码单元的输出为第 i 级解码输出图像并包括 m^{i-1} 个图像成分, 以及第 i 级解码单元的输出是第 $i+1$ 级解码单元的输入;

解码输出端, 配置成输出对应于所述解码输入图像的还原图像。

18. 一种用于如权利要求 17 所述的图像处理器的训练方法, 包括:

将训练图像输入所述图像处理器, 调整所述 n 级编码单元和所述 n 级解码单元中各卷积层中各卷积网络的权值, 运行有限次迭代以使目标函数最优化。

19. 如权利要求 18 所述的训练方法, 其中所述目标函数是以下各项中的一项或多项的和:

编码损失函数

$$L_{ENC_k} = \|REF_k - LR_k\|_2 + \sum_{\substack{i=1 \\ i \neq C_k}}^{C_0 \cdot m^k} \|C_{ki}\|_1$$

其中 REF_k 是第 k 级编码单元输出的第一图像成分, LR_k 是第 k 级编码单元的训练图像, 其中 LR_k 是所述图像处理器的训练图像的下采样图像, 并具有与 REF_k 相同的尺寸; C_0 是所述训练图像的数量; C_{ki} 是第 k 级编码单元输

出的图像成分, 其中 $1 \leq i \leq m^k - 1$;

解码损失函数

$$L_{DEC_k} = IQ(REF_k, LR_k)$$

其中 IQ 函数评价 REF_k 与 LR_k 之间的差别;

5 风格损失函数

$$L_{STYLE_k}(X, Y) = \|G_X - G_Y\|_2$$

$$\text{其中 } G_F(k, l) = \frac{1}{hw} \sum_{i=1}^m F^k(i) F^l(i);$$

其中 G_X 、 G_Y 分别是 X 图像、Y 的格拉姆矩阵的特征量, X 是第 k 级编码单元的输出图像, Y 是第 $i+1-k$ 级编码单元的输出图像, 其中 $1 \leq k \leq n$;

10 权重正则化系数

$$L_{REG} = \frac{\|W\|_1}{\|b\|_1}$$

其中 W 是所述图像处理器中所有卷积网络的权重参数, b 是所述图像处理器中所有卷积网络的偏置。

20. 一种图像转换方法, 包括:

15 接收第一输入图像和第二输入图像;

利用级联的 n 级编码单元对所述第一输入图像进行图像编码以产生第一输出图像, n 为大于 1 的整数,

20 对于 $1 \leq i < n$, 第 i 级编码单元的输入为第 i 级编码输入图像并包括 m^{i-1} 个图像成分, 第 i 级编码单元的输出为第 i 级编码输出图像并包括 m^i 个图像成分, 以及第 i 级编码单元的输出是第 $i+1$ 级编码单元的输入, 其中 m 是大于 1 的整数;

输出第一输出图像, 所述第一输出图像包括 m^n 个输出子图像, 所述 m^n 个输出子图像分别对应于第 n 级编码单元的 m^n 个输出图像成分, 并且所述 m^n 个输出子图像的每一个对应于一种图像类别;

25 获取所述第一输出图像中 m^n 个输出子图像的每一个的像素值, 并根据像素值确定所述 m^n 个输出子图像中的至少一个是所述第一输入图像的分类子图像;

确定所述第一输入图像的分类为对应于所述分类子图像的图像类别;

获取所述类别子图像的像素信息;

基于所述类别子图像的像素信息对所述第二输入图像进行图像变换, 将所述第二输入图像变换为对应于所述第一输入图像的图像类别的第三图像。

21. 一种如权利要求 20 所述的图像转换方法, 其中利用第 i 级编码单元
5 进行图像编码包括:

接收第 i 级编码输入图像;

对于所述第 i 级编码输入图像中的每个图像成分, 将该图像成分拆分为 m
个编码输入成分, 其中所述 m 个编码输入成分中的每一个的尺寸是第 i 级编
码输入图像中的每个图像成分的尺寸的 $1/m$ 倍, 对所述 m 个编码输入成分进
10 行图像变换, 生成对应于该图像成分的 m 个编码输出成分, 其中所述 m 个编
码输出成分中的每一个的尺寸与所述 m 个编码输入成分中的每一个的尺寸相
同;

输出对应于第 i 级编码输入的 m^{i-1} 个图像成分的 m^i 个编码输出成分作为
第 i 级编码输出图像。

15 22. 一种计算机可读介质, 其上存储有指令, 当处理器执行所述指令时使
得计算机执行如权利要求 1-8 中任一所述的图像分类方法或权利要求 20-21 中
任一所述的图像转换方法。

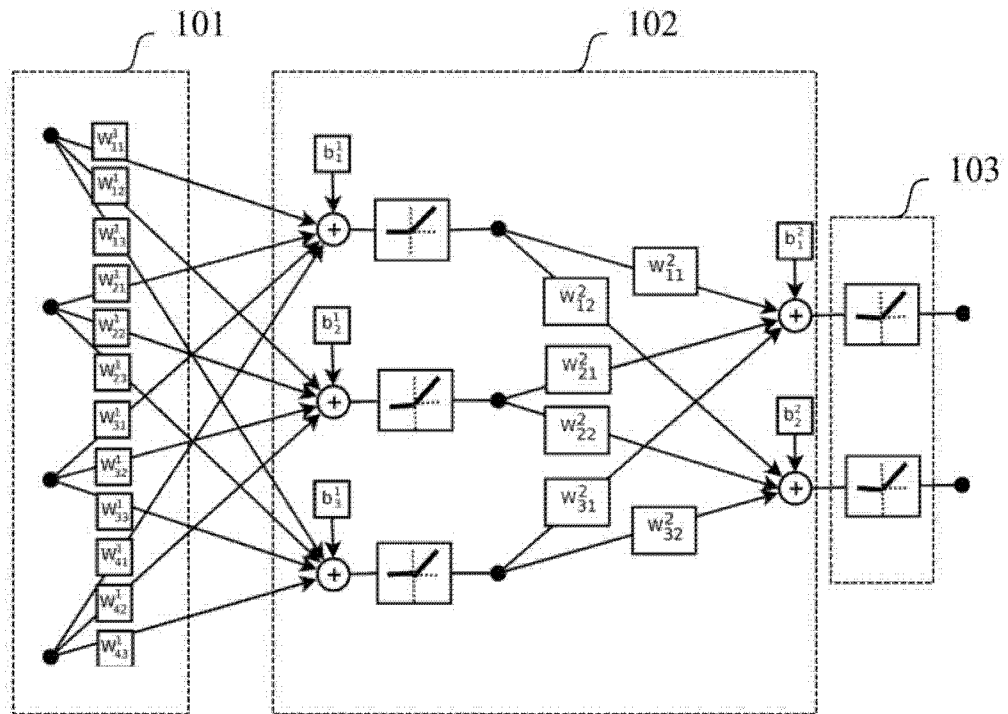


图1

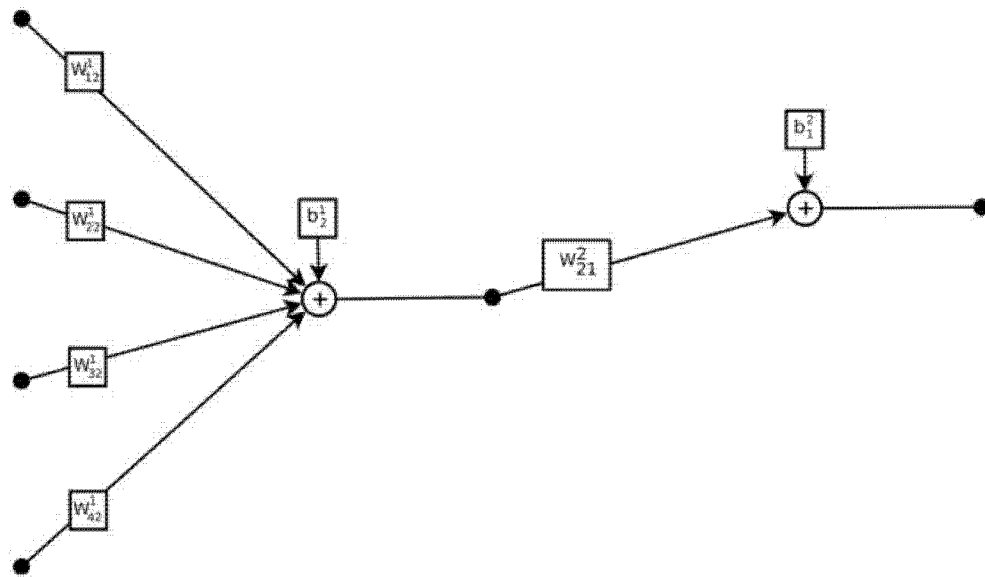


图2

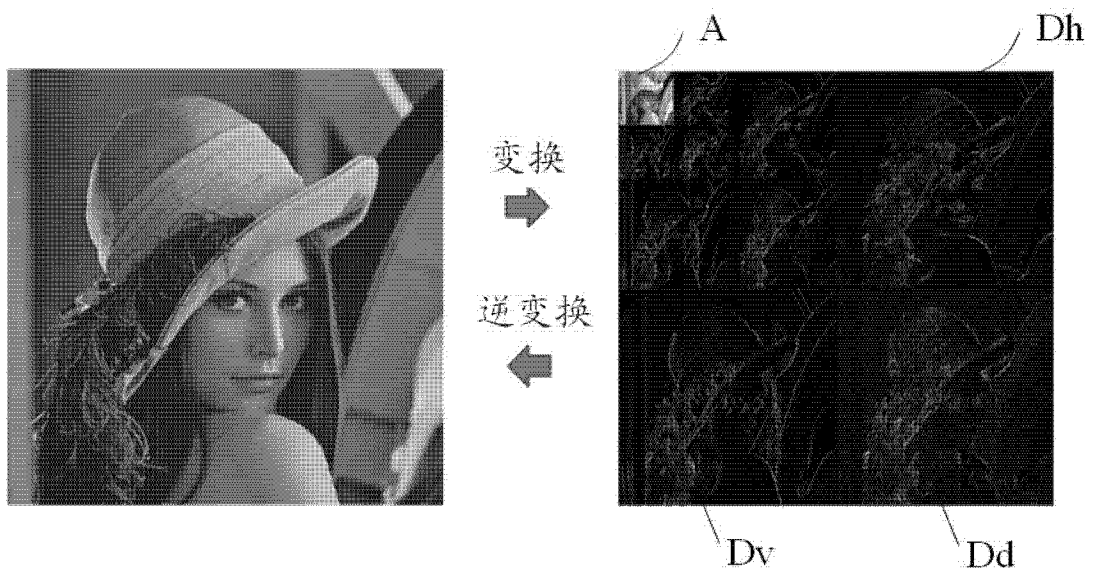


图3

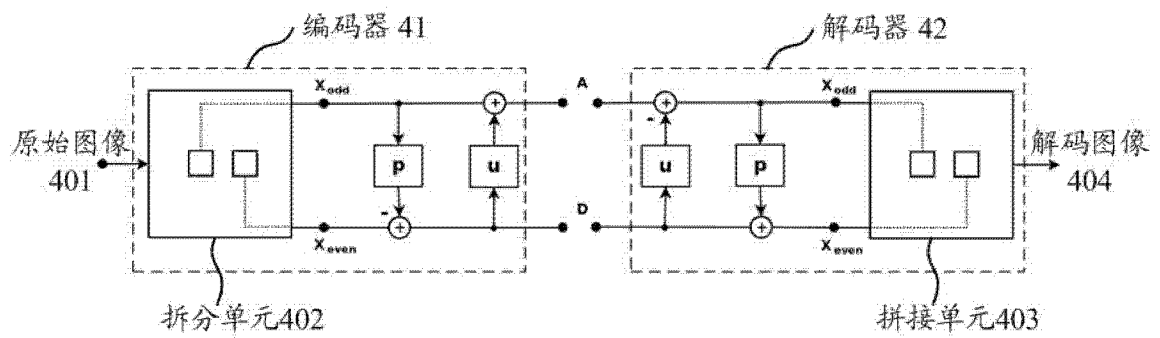


图4

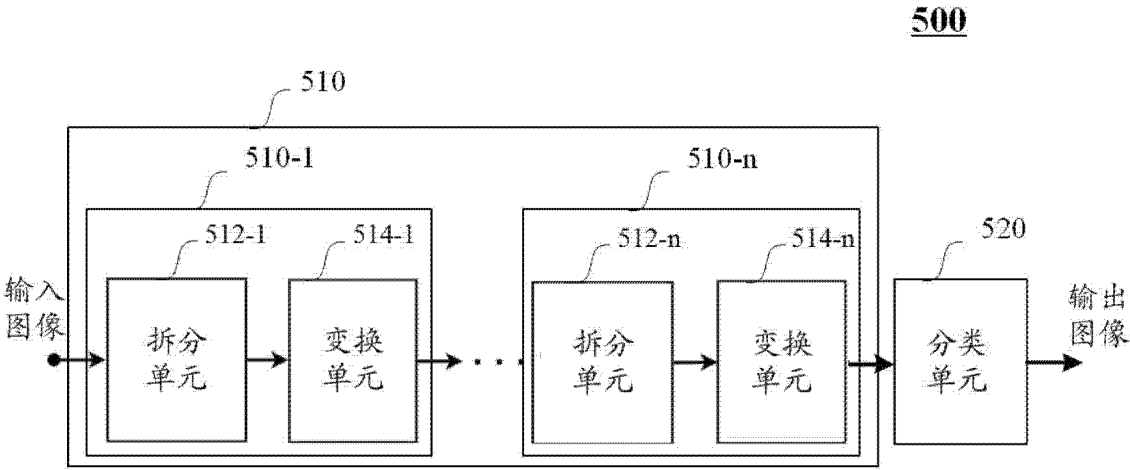


图5

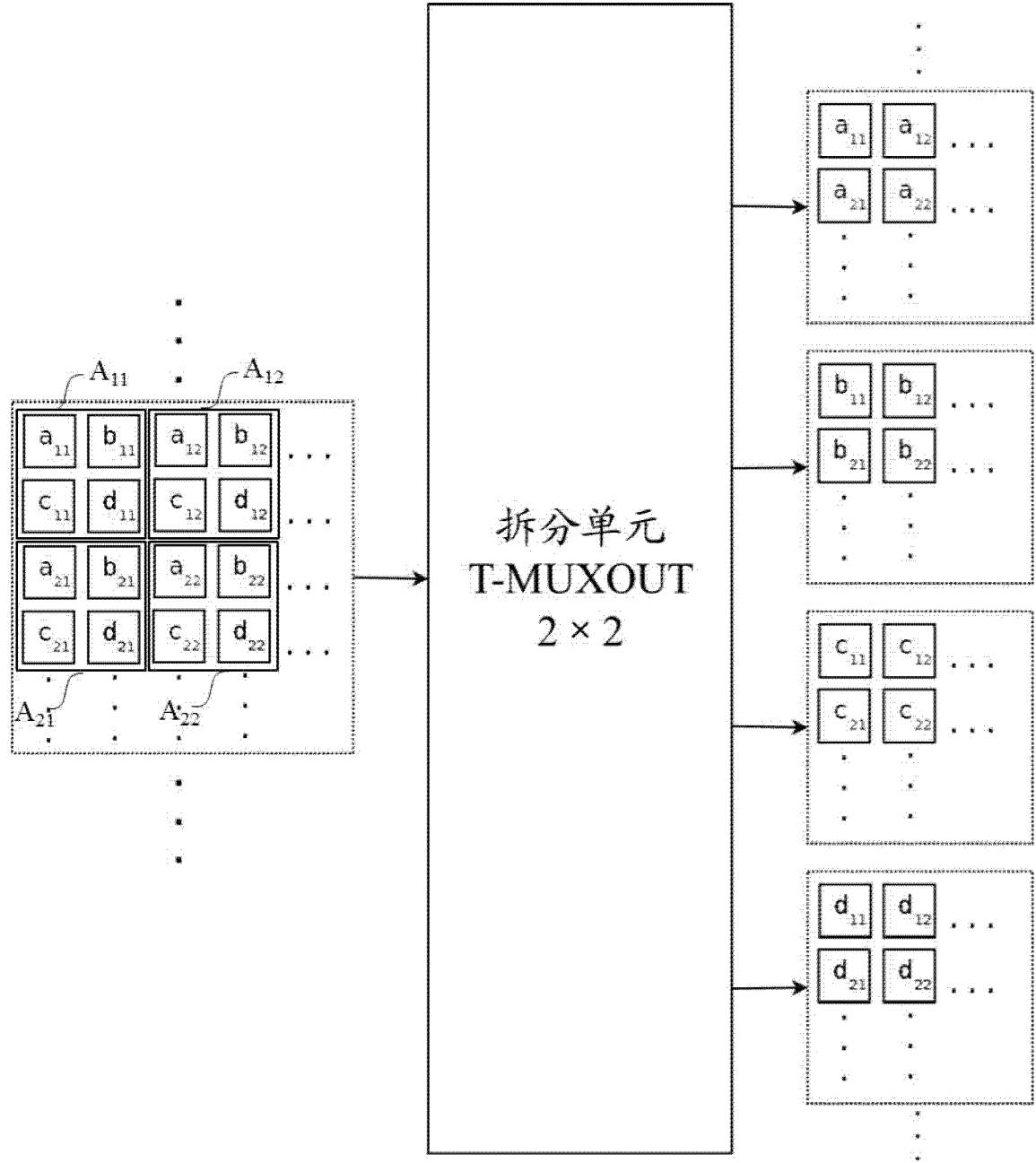


图6

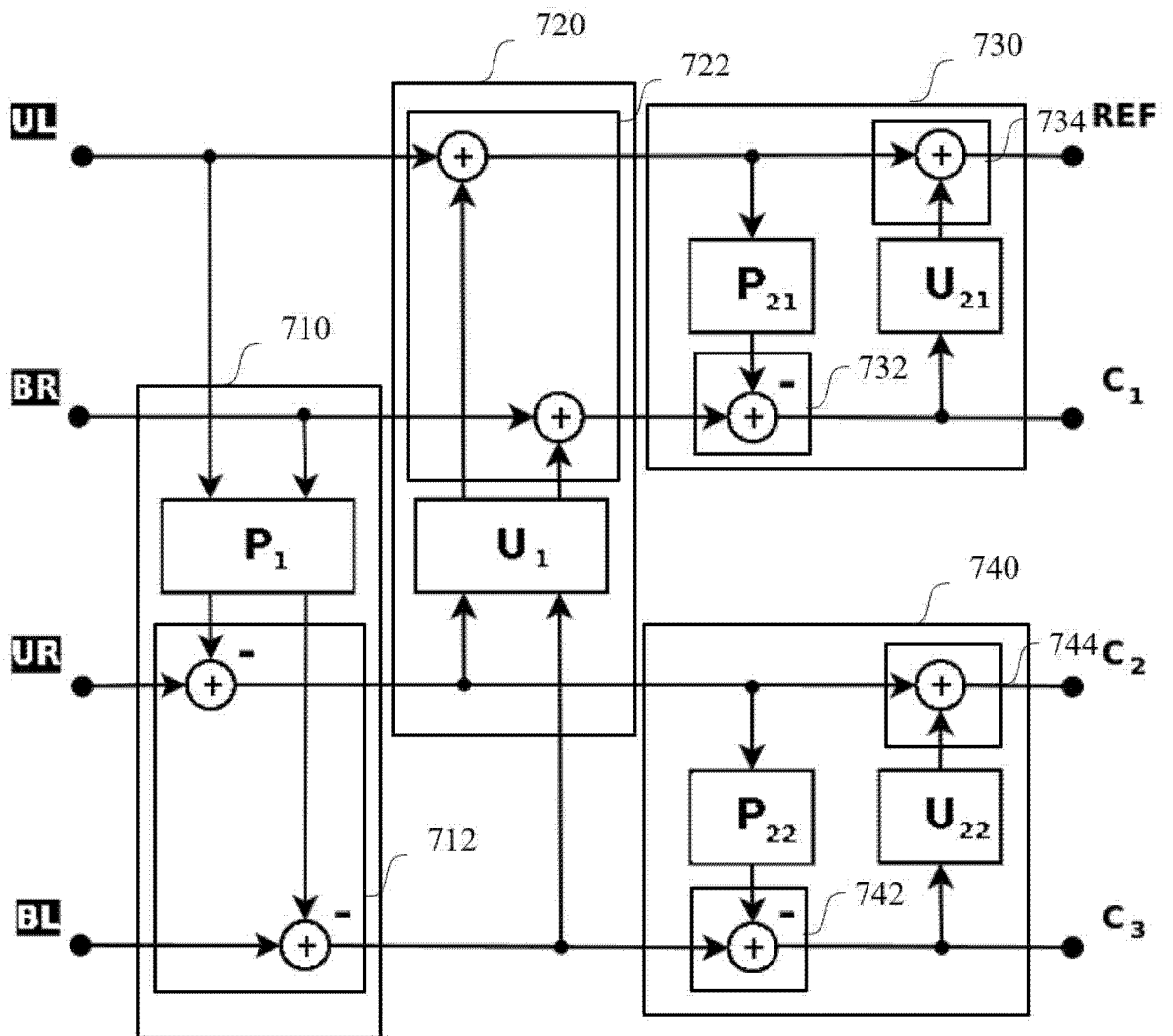


图7

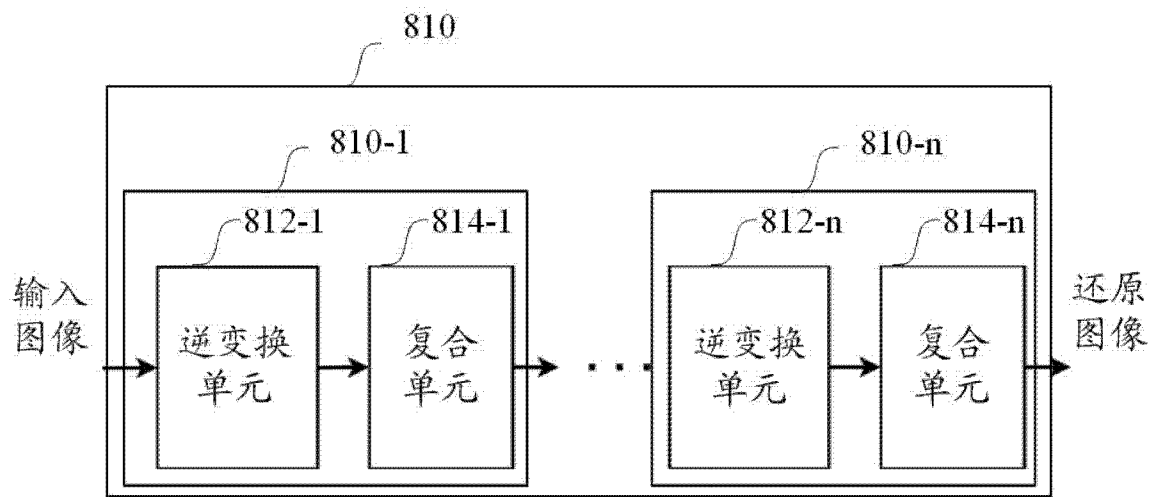
800

图8

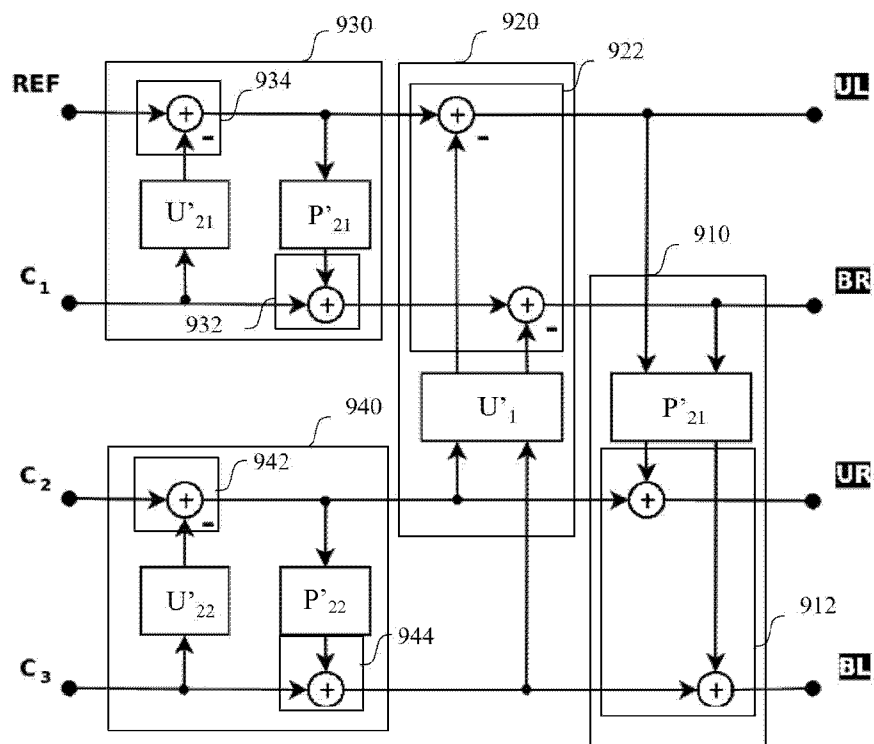


图9

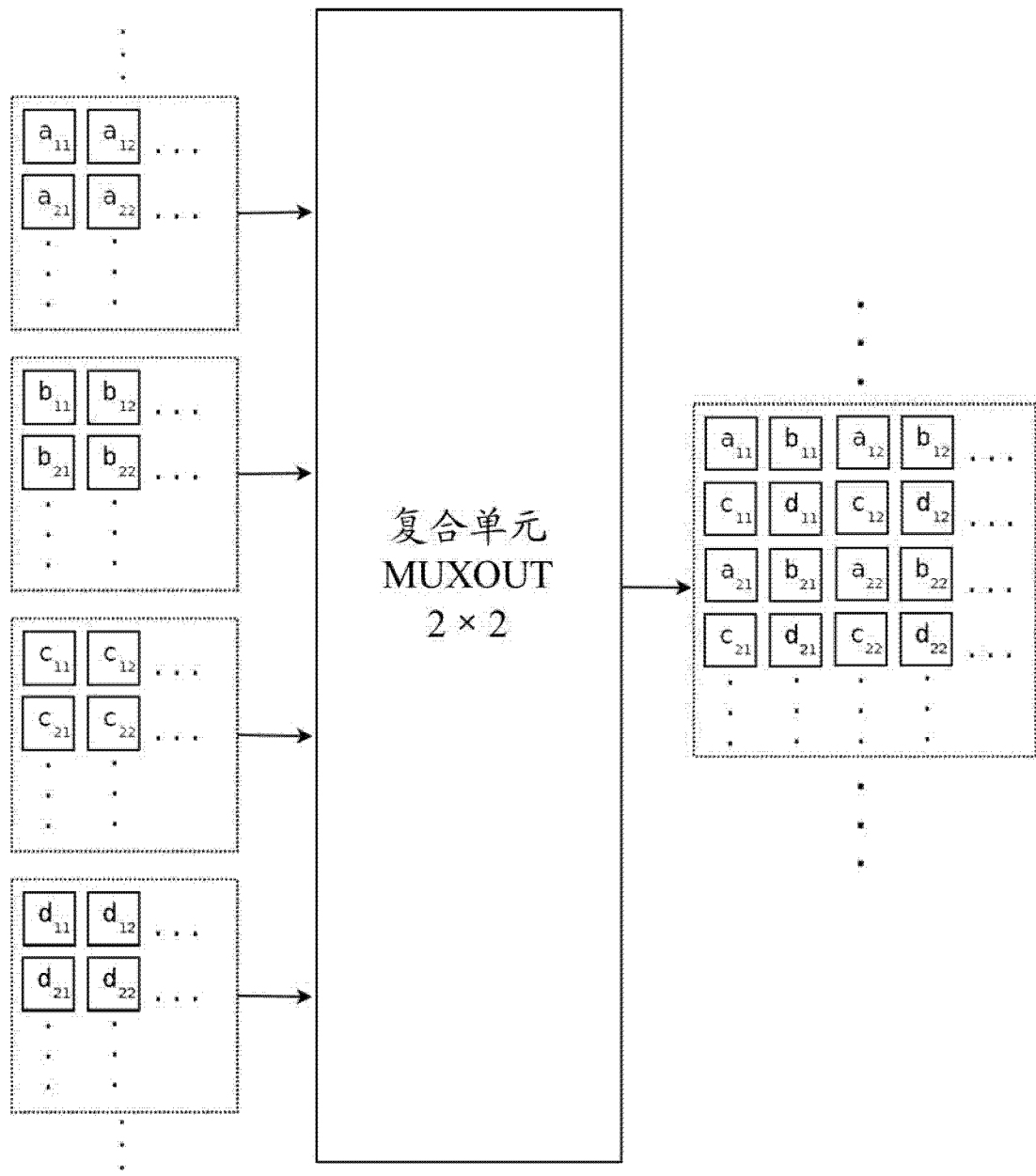


图10

1100

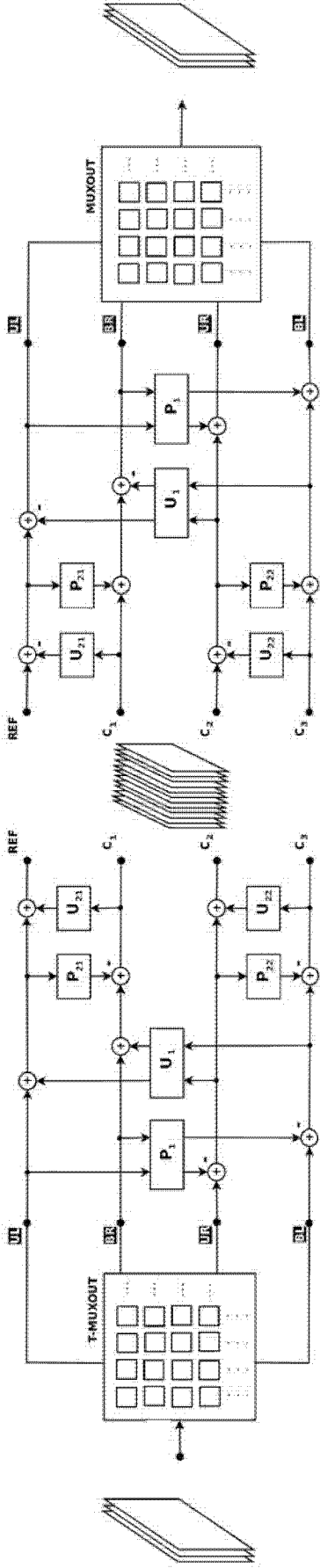


图11

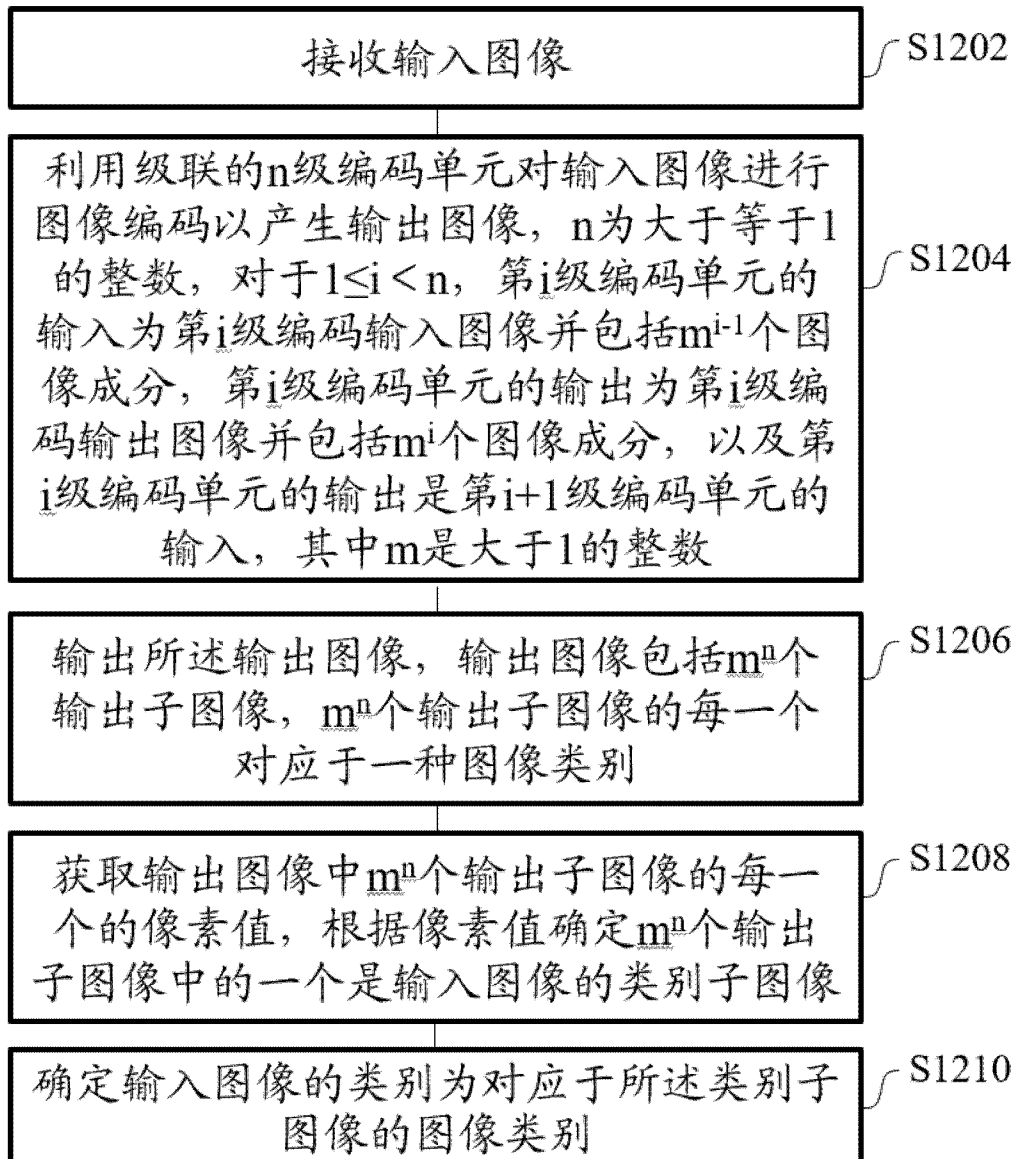
1200

图12

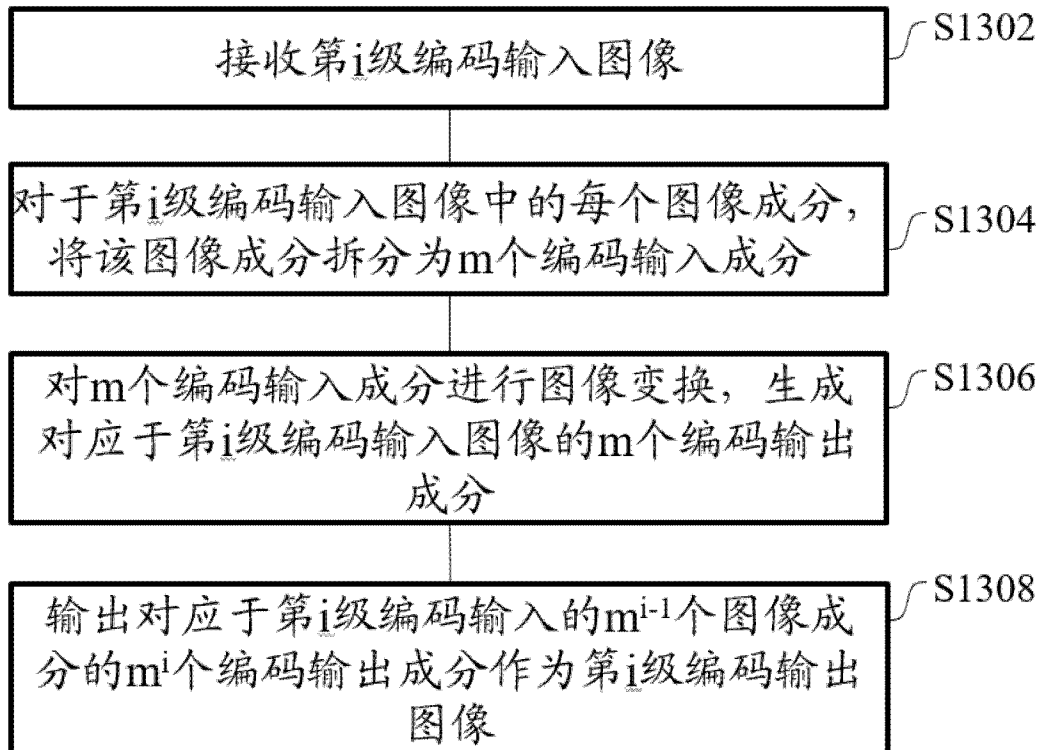
1300

图13

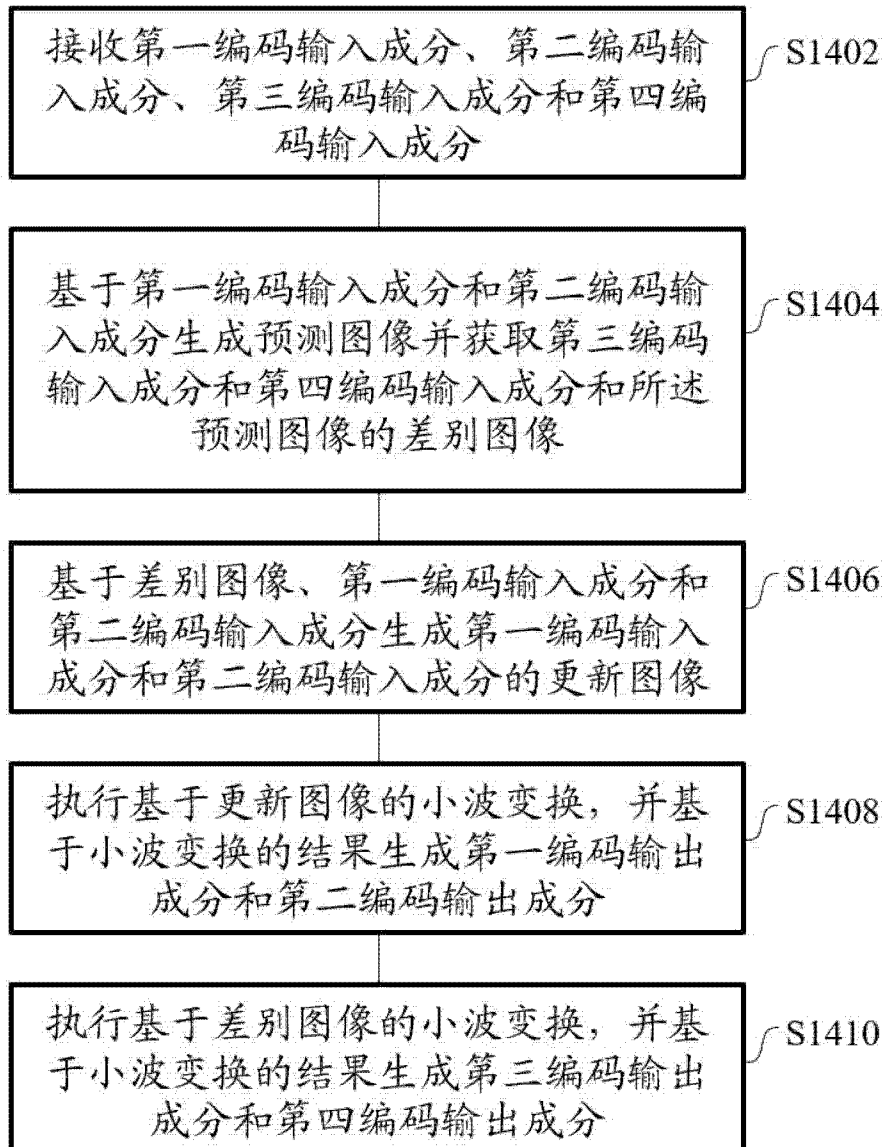
1400

图14

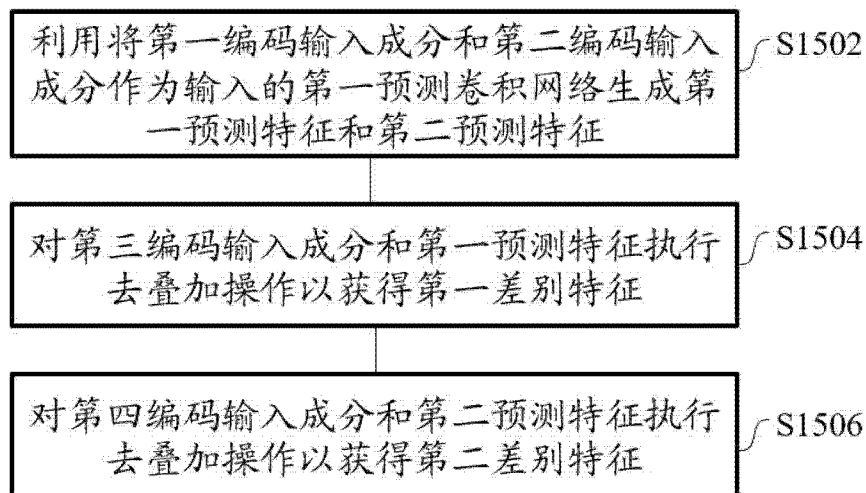


图15A

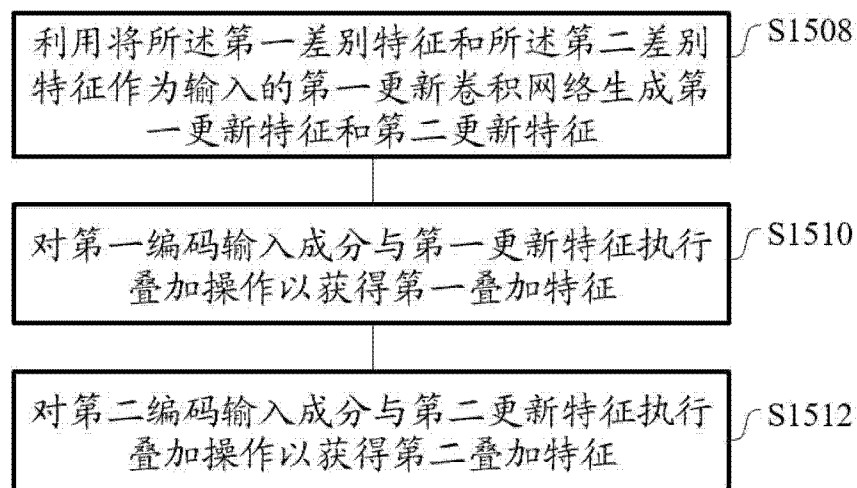


图15B

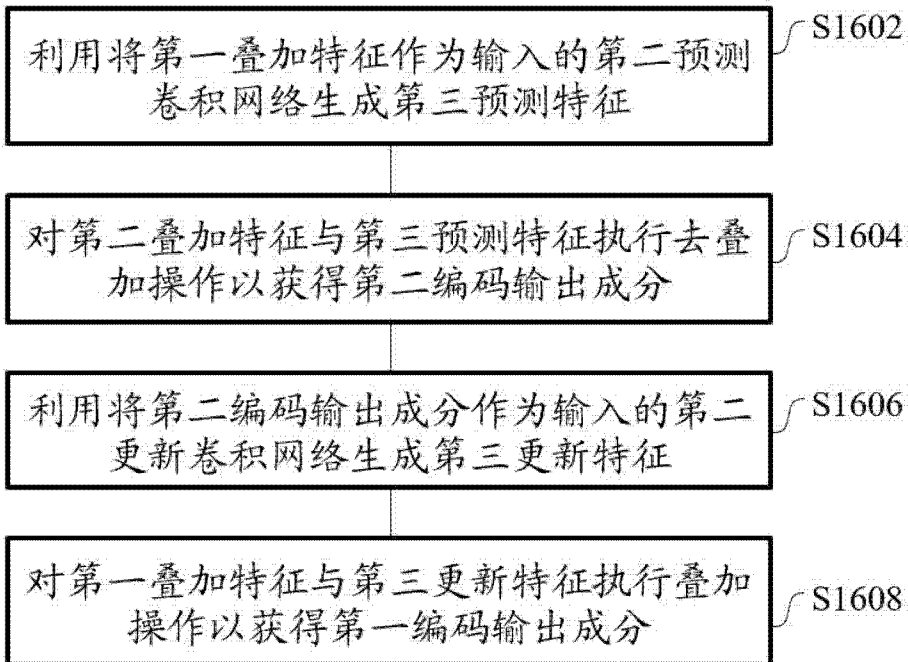
1600

图16

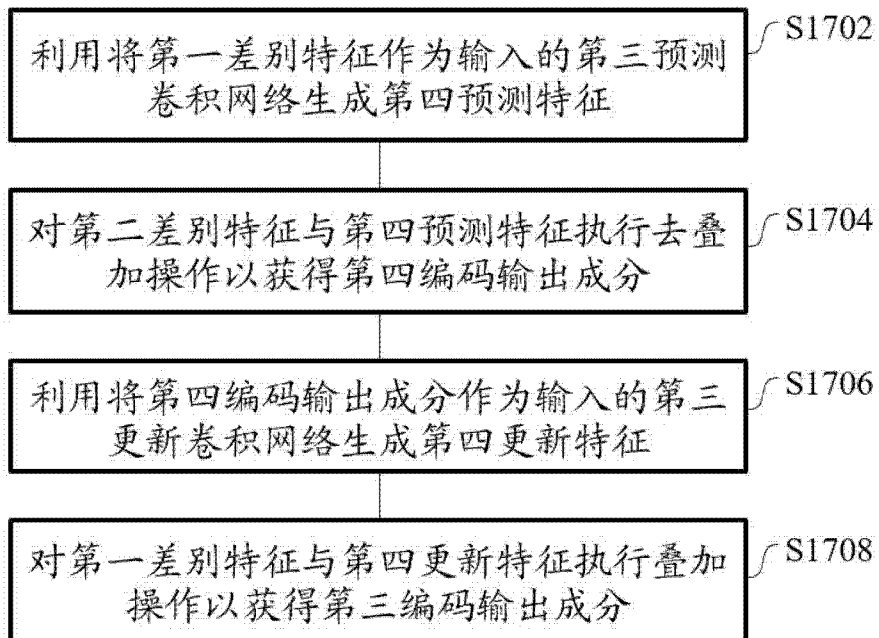
1700

图17

1800

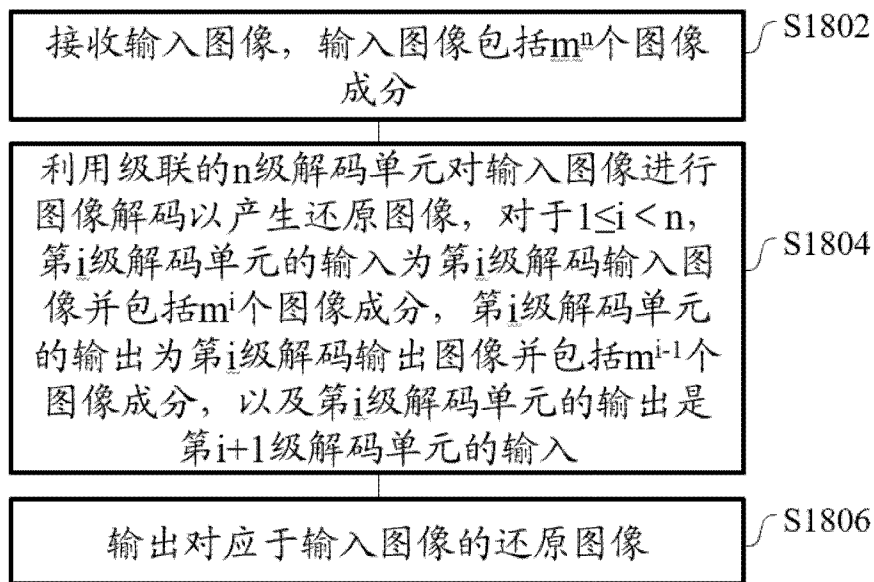


图 18

1900

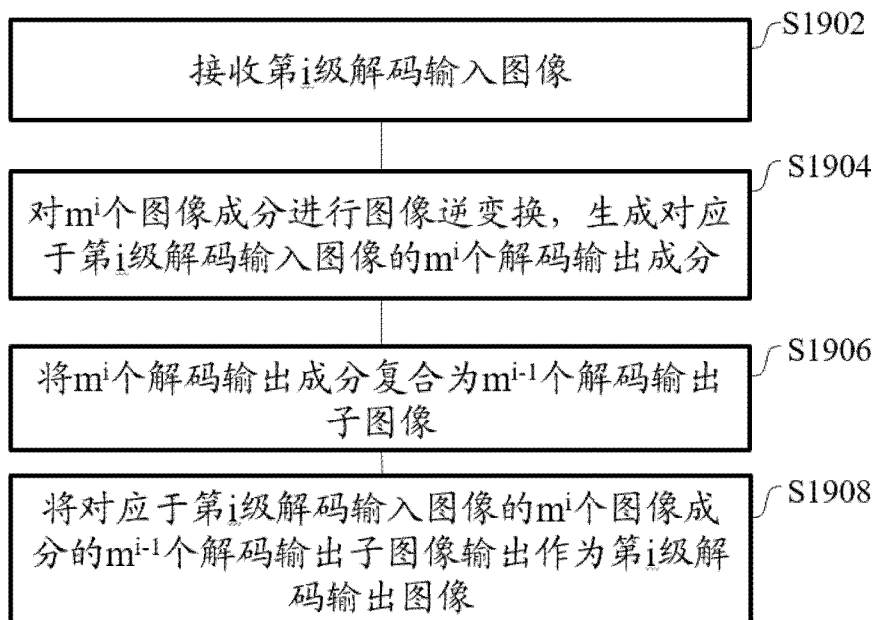


图 19

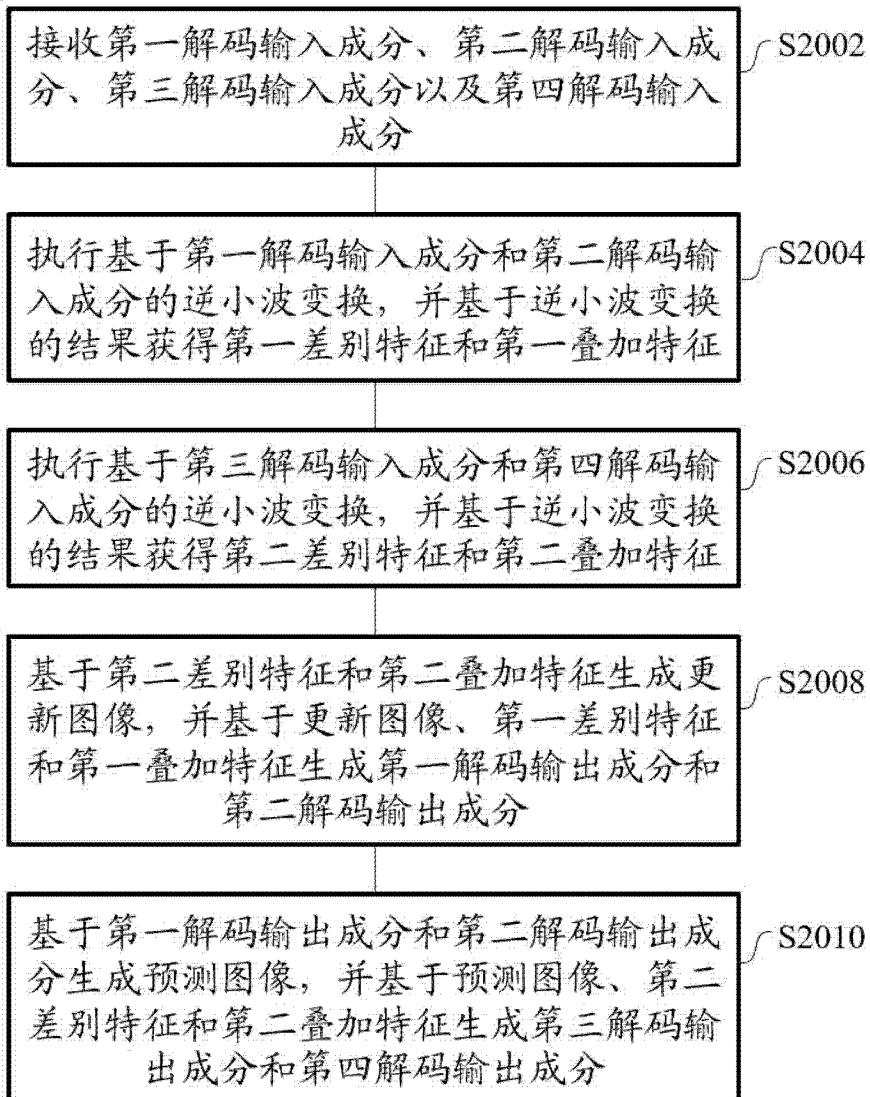
2000

图20

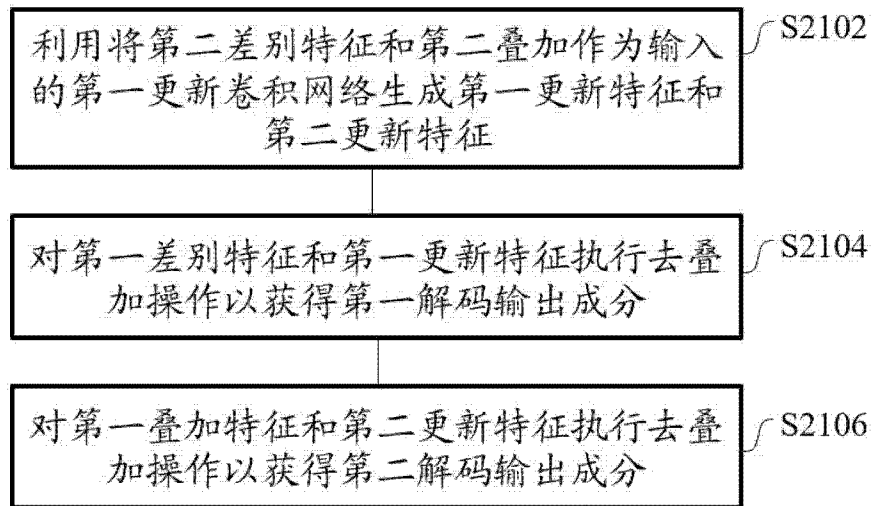


图21A

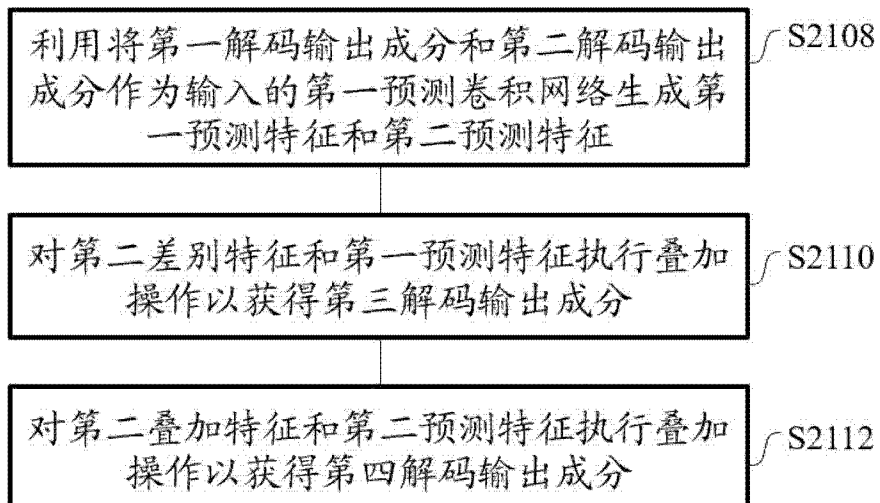


图21B

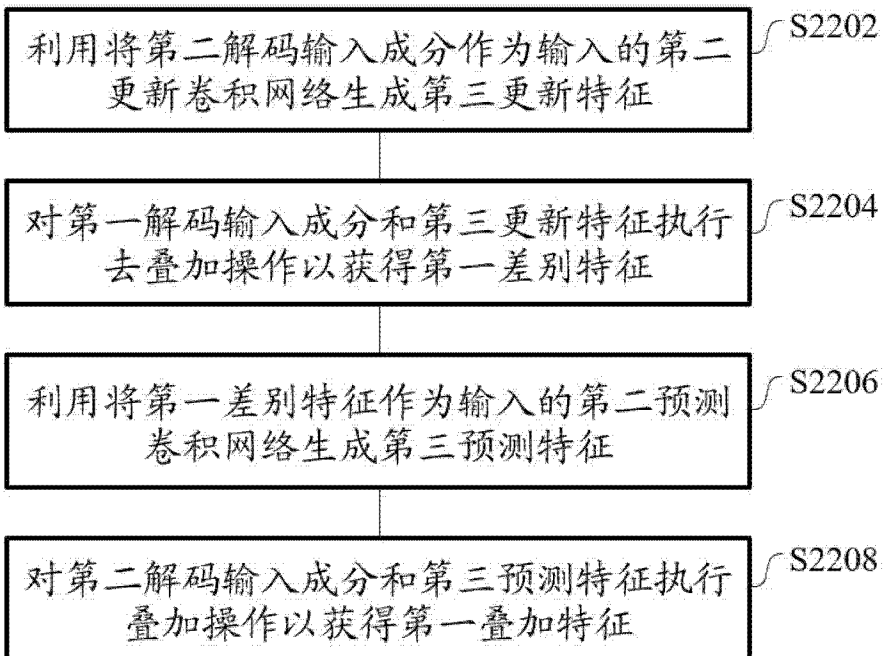
2200

图22

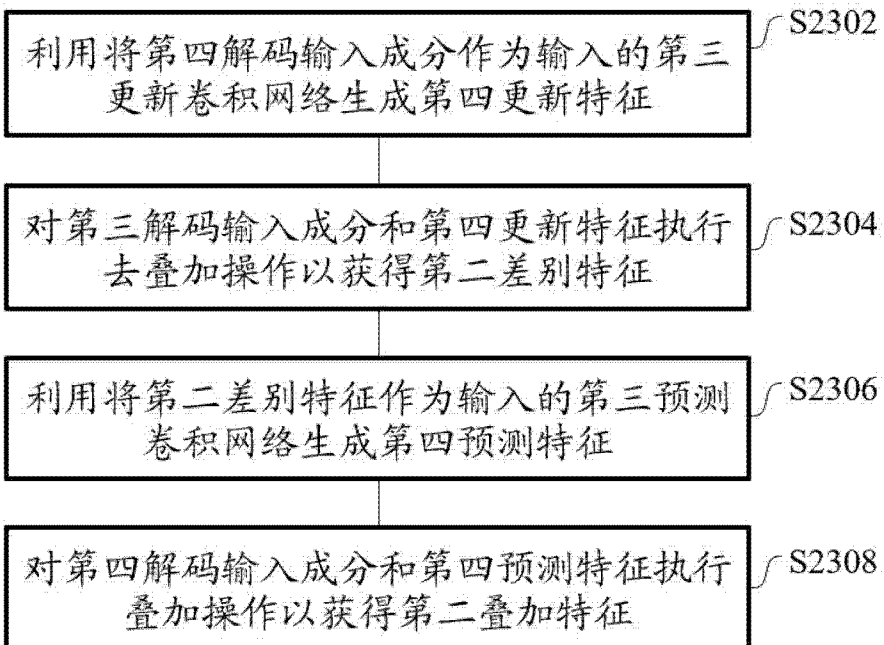
2300

图23

2400

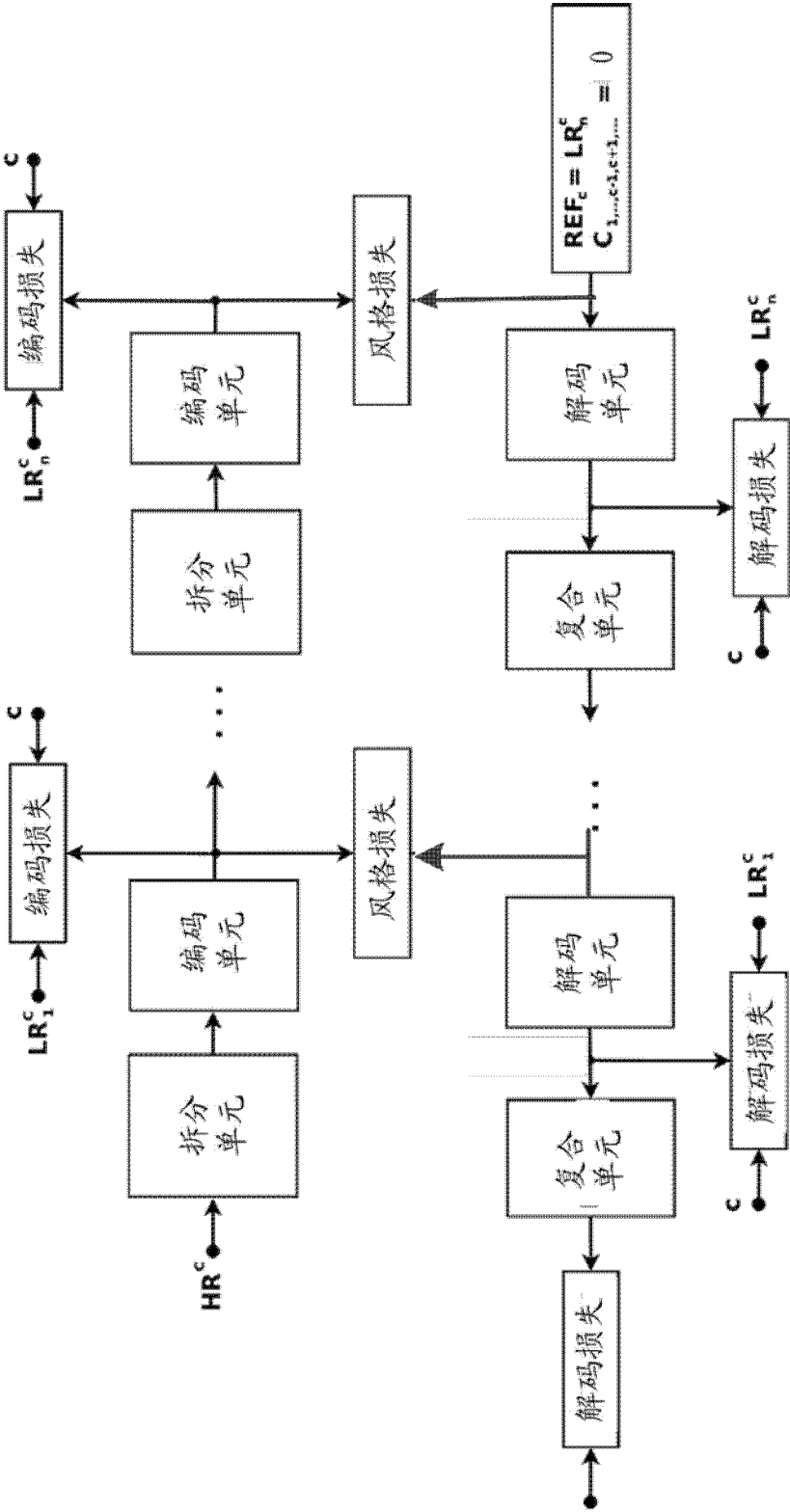


图 24

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/113115

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/62(2006.01)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)

G06K9/-

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, CNKI, VEN, USTXT, WOTXT, EPTXT: 图像, 分类, 编码, 子图像, 像素, 小波, 卷积, image, picture, class, coding, sub image, pixel, wavelet, convolution

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 107895174 A (BOE TECHNOLOGY GROUP CO., LTD.) 10 April 2018 (2018-04-10) entire document	1-22
A	EP 2977933 A1 (BADEN WUERTEMBERG STIFTUNG GGMBH) 27 January 2016 (2016-01-27) entire document	1-22
A	CN 103116766 A (NANJING UNIVERSITY) 22 May 2013 (2013-05-22) entire document	1-22
A	CN 106845418 A (BEIHANG UNIVERSITY) 13 June 2017 (2017-06-13) entire document	1-22
A	CN 105809182 A (SHENZHEN INSTITUTES OF ADVANCED TECHNOLOGY, CHINESE ACADEMY OF SCIENCES) 27 July 2016 (2016-07-27) entire document	1-22

☐ Further documents are listed in the continuation of Box C.☒ 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

09 January 2019

Date of mailing of the international search report

29 January 2019

Name and mailing address of the ISA/CN

State Intellectual Property Office of the P. R. China (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/CN2018/113115

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	107895174	A	10 April 2018	None	
EP	2977933	A1	27 January 2016	None	
CN	103116766	A	22 May 2013	CN 103116766	B 18 November 2015
CN	106845418	A	13 June 2017	None	
CN	105809182	A	27 July 2016	None	

国际检索报告

国际申请号

PCT/CN2018/113115

A. 主题的分类 G06K 9/62 (2006.01) i 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																				
B. 检索领域 检索的最低限度文献 (标明分类系统和分类号) G06K9/- 包含在检索领域中的除最低限度文献以外的检索文献 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)) CNABS, CNTXT, CNKI, VEN, USTXT, WOTXT, EPTXT: 图像, 分类, 编码, 子图像, 像素, 小波, 卷积, image, picture, class, coding, sub image, pixel, wavelet, convolution																				
C. 相关文件 <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>PX</td> <td>CN 107895174 A (京东方科技集团股份有限公司) 2018年 4月 10日 (2018 - 04 - 10) 全文</td> <td>1-22</td> </tr> <tr> <td>A</td> <td>EP 2977933 A1 (BADEN WUERTTEMBERG STIFTUNG GMBH) 2016年 1月 27日 (2016 - 01 - 27) 全文</td> <td>1-22</td> </tr> <tr> <td>A</td> <td>CN 103116766 A (南京大学) 2013年 5月 22日 (2013 - 05 - 22) 全文</td> <td>1-22</td> </tr> <tr> <td>A</td> <td>CN 106845418 A (北京航空航天大学) 2017年 6月 13日 (2017 - 06 - 13) 全文</td> <td>1-22</td> </tr> <tr> <td>A</td> <td>CN 105809182 A (中国科学院深圳先进技术研究院) 2016年 7月 27日 (2016 - 07 - 27) 全文</td> <td>1-22</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 107895174 A (京东方科技集团股份有限公司) 2018年 4月 10日 (2018 - 04 - 10) 全文	1-22	A	EP 2977933 A1 (BADEN WUERTTEMBERG STIFTUNG GMBH) 2016年 1月 27日 (2016 - 01 - 27) 全文	1-22	A	CN 103116766 A (南京大学) 2013年 5月 22日 (2013 - 05 - 22) 全文	1-22	A	CN 106845418 A (北京航空航天大学) 2017年 6月 13日 (2017 - 06 - 13) 全文	1-22	A	CN 105809182 A (中国科学院深圳先进技术研究院) 2016年 7月 27日 (2016 - 07 - 27) 全文	1-22
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																		
PX	CN 107895174 A (京东方科技集团股份有限公司) 2018年 4月 10日 (2018 - 04 - 10) 全文	1-22																		
A	EP 2977933 A1 (BADEN WUERTTEMBERG STIFTUNG GMBH) 2016年 1月 27日 (2016 - 01 - 27) 全文	1-22																		
A	CN 103116766 A (南京大学) 2013年 5月 22日 (2013 - 05 - 22) 全文	1-22																		
A	CN 106845418 A (北京航空航天大学) 2017年 6月 13日 (2017 - 06 - 13) 全文	1-22																		
A	CN 105809182 A (中国科学院深圳先进技术研究院) 2016年 7月 27日 (2016 - 07 - 27) 全文	1-22																		
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																				
* 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																				
国际检索实际完成的日期 2019年 1月 9日		国际检索报告邮寄日期 2019年 1月 29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 王国海 电话号码 86-(20)-28958137																		

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/113115

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	107895174	A	2018年 4月 10日	无	
EP	2977933	A1	2016年 1月 27日	无	
CN	103116766	A	2013年 5月 22日	CN 103116766 B	2015年 11月 18日
CN	106845418	A	2017年 6月 13日	无	
CN	105809182	A	2016年 7月 27日	无	