(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 7 月 4 日 (04.07.2019)

WIPO | PCT

(10) 国际公布号
WO 2019/128248 A1

(51) 国际专利分类号:
G06F 17/16 (2006.01) *G06N 3/04* (2006.01)

(21) 国际申请号: PCT/CN2018/099733

(22) 国际申请日: 2018 年 8 月 9 日 (09.08.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201711481199.4 2017年12月29日 (29.12.2017) CN

(71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(72) 发明人: 许若圣 (XU, Ruosheng); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 陈静炜 (CHEN, Jingwei); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(74) 代理人: 北京中博世达专利商标代理有限公司 (BEIJING ZBSD PATENT & TRADEMARK AGENT LTD.); 中国北京市海淀区交大东路31号11号楼8层, Beijing 100044 (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,

(54) **Title:** SIGNAL PROCESSING METHOD AND APPARATUS

(54) 发明名称: 一种信号处理方法及装置

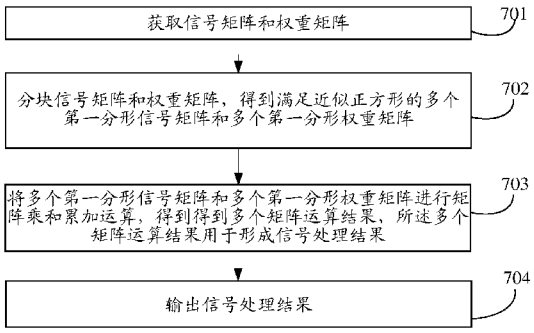

图 7

701 Acquire a signal matrix and a weight matrix
702 Divide the signal matrix and the weight matrix into blocks to obtain multiple first fractal signal matrices and multiple first fractal weight matrices satisfying a non-approximate square
703 Carry out matrix multiplication and accumulation operations on the multiple first fractal signal matrices and the multiple first fractal weight matrices to obtain multiple matrix operation results, wherein the multiple matrix operation results are used to form a signal processing result
704 Output the signal processing result

(57) **Abstract:** Provided are a signal processing method and apparatus, which relate to the technical field of computers and are used for improving the flexibility of a fractal matrix. The method is applied to a device including a processor, and comprises: acquiring a signal matrix and a weight matrix, wherein the signal matrix and the weight matrix are both two-dimensional matrices, and the number of columns of the signal matrix is equal to the number of rows of the weight matrix; dividing the signal matrix into blocks to obtain multiple first fractal signal matrices of X rows and H columns, and dividing the weight matrix into blocks to obtain multiple first fractal weight matrices of H rows and Y columns, wherein each of the first fractal signal matrices and each of the first fractal weight matrices satisfy a non-approximate square; and carrying out matrix multiplication and accumulation operations on the multiple first fractal signal matrices and the multiple first fractal weight matrices to obtain multiple matrix operation results, wherein the multiple matrix operation results are used to form a signal processing result.

(57) **摘要**: 本申请提供一种信号处理方法及装置，涉及计算机技术领域，用于提高分形矩阵的灵活性。该方法应用于包含处理器的设备中，包括：获取信号矩阵和权重矩阵，信号矩阵和权重矩阵均为二维矩阵，信号矩阵的列数与权重矩阵的行数相等；分块信号矩阵，得到X行H列的多个第一分形信号矩阵，以及分块权重矩阵，得到H行Y列的多个第一分形权重矩阵，每个第一分形信号矩阵和每个第一分形权重矩阵均满足非近似正方形；将多个第一分形信号矩阵和多个第一分形权重矩阵进行矩阵乘和累加运算，得到多个矩阵运算结果，所述多个矩阵运算结果用于形成信号处理结果。

WO 2019/128248 A1

[见续页]

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。

(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))。

一种信号处理方法及装置

本申请要求于2017年12月29日提交中国专利局、申请号为201711481199.4、申请名称为“一种信号处理方法及装置”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

本申请实施例涉及计算机技术领域，尤其涉及一种信号处理方法及装置。

背景技术

神经网络（Neural Network，NN），是一种模仿动物神经网络行为特征进行信息处理的网络结构。该结构由大量的节点（或称神经元）相互联接构成，基于特定运算模型通过对输入信息进行学习和训练达到处理信息的目的。一个神经网络包括输入层、隐藏层及输出层，输入层负责接收输入信号，输出层负责输出神经网络的计算结果，隐藏层负责学习、训练等计算过程，是网络的记忆单元，隐藏层的记忆功能由权重矩阵来表征，通常每个神经元对应一个权重系数。

其中，卷积神经网络（Convolutional Neural Network，CNN）是一种多层的神经网络，每层有多个二维平面组成，而每个平面由多个独立神经元组成，每个平面的多个神经元共享权重，通过权重共享可以降低神经网络中的参数数目。目前，在卷积神经网络中，处理器进行卷积操作通常是将输入信号特征与权重的卷积，转换为信号矩阵与权重矩阵之间的矩阵乘运算。在具体矩阵乘运算时，通常是根据条件|row-columns|≤1（也即是，|行数-列数|≤1，即矩阵的行数与列数的差值的绝对值小于或等于1），对信号矩阵和权重矩阵进行分形处理，得到多个近似于正方形的分形（Fractional）信号矩阵和分形权重矩阵，然后对多个分形信号矩阵和分形权重矩阵进行矩阵乘和累加运算。比如，如图1所示，假设C=AB，A为信号矩阵，B为权重矩阵，则进行矩阵乘运算时，由于处理器可能缺少对A和B这种大矩阵进行计算的能力或进行此类计算代价较大，可以将矩阵A根据条件划分为A00、A01、A10和A11，将矩阵B根据条件划分为B00、B01、B10和B11，相应的矩阵C可以由C00、C01、C10和C11四个矩阵块组成，矩阵C中每一矩阵块与分形信号矩阵和分形权重矩阵的关系可以如下公式所示。

$$C00 = A00B00 + A01B10$$

$$C01 = A00B01 + A01B11$$

$$C10 = A10B00 + A11B10$$

$$C11 = A10B01 + A11B11$$

上述方法中，在对矩阵C中的每个矩阵块进行计算时，可以通过数据复用的方式进行计算以减少功耗，例如，C00和C01的计算复用了数据A00，降低了读取数据A00的功耗开销。但是，根据条件|row-columns|≤1对信号矩阵和权重矩阵进行分形处理，得到的分形信号矩阵和分形权重矩阵的形状是固定的，其消耗的功耗也是固定的，设计灵活性不足。

发明内容

本发明的实施例提供一种信号处理方法及装置，用于提高分形矩阵的灵活性。

为达到上述目的，本发明的实施例采用如下技术方案：

第一方面，提供一种信号处理方法，应用于包含处理器的设备中，该方法包括：获取信号矩阵和权重矩阵，信号矩阵为二维矩阵且包括多个计算机可处理的待处理信号，权重矩阵为二维矩阵且包括多个权重系数，信号矩阵的列数与权重矩阵的行数相等；分块信号矩阵，得到 X 行 H 列的多个第一分形信号矩阵，以及分块权重矩阵，得到 H 行 Y 列的多个第一分形权重矩阵，每个第一分形信号矩阵和每个第一分形权重矩阵均满足非近似正方形；将多个第一分形信号矩阵和多个第一分形权重矩阵进行矩阵乘和累加运算，得到多个矩阵运算结果，所述多个矩阵运算结果用于形成信号处理结果，每个矩阵运算结果包括多个矩阵乘结果的累加，每个矩阵乘结果由一个第一分形信号矩阵与一个第一分形权重矩阵进行矩阵乘运算得到。可选地，所述方法还包括：输出信号处理结果，该信号处理结果包括所述多个矩阵运算结果。

上述技术方案中，处理器在获取到信号矩阵和权重矩阵时，分块信号矩阵和权重矩阵，得到 X 行 H 列的多个第一分形信号矩阵和 H 行 Y 列的多个第一分形权重矩阵，由于每个第一分形信号矩阵和每个第一分形权重矩阵均满足非近似正方形，因此，提高了分形矩阵的灵活性，以利于功耗优化设计。

在第一方面的一种可能的实现方式中，满足非近似正方形包括，矩阵的行数与列数的差值的绝对值大于或等于 2，即第一分形信号矩阵和第一分形权重矩阵的行数与列数的差值的绝对值均大于或等于 2。

在第一方面的一种可能的实现方式中，处理器包括第一缓存器和第二缓存器，分块信号矩阵，得到 X 行 H 列的多个第一分形信号矩阵，以及分块权重矩阵，得到 H 行 Y 列的多个第一分形权重矩阵，包括：通过第一缓存器分别多次从信号矩阵中读取 X 行 H 列的多个第一分形信号矩阵；通过第二缓存器分别多次从所述权重矩阵中读取 H 行 Y 列的多个第一分形权重矩阵。上述可能的技术方案中，处理器可以通过第一缓存器读取到非近似正方形的第一分形信号矩阵，通过第二缓存器读取到非近似正方形的第一分形权重矩阵，从而可以提供第一缓存器和第二缓存器读取的分形矩阵的灵活性。

在第一方面的一种可能的实现方式中，处理器还包括第三缓存器，该方法还包括：向第三缓存器中写入矩阵乘结果或至少两个矩阵乘结果的累加。

在第一方面的一种可能的实现方式中，对于一个第一分形信号矩阵与一个第一分形权重矩阵进行矩阵乘运算，得到一个矩阵乘结果，包括：分块第一分形信号矩阵，得到 x 行 h 列的多个第二分形信号矩阵，以及分块第一分形权重矩阵，得到 h 行 y 列的多个第二分形权重矩阵，每个第二分形信号矩阵和每个第二分形权重矩阵均满足非近似正方形；将多个第二分形信号矩阵和多个第二分形权重矩阵进行矩阵乘和累加运算，得到多个矩阵运算结果。上述可能的技术方案中，当处理器一次无法计算一个第一分形信号矩阵与一个第一分形权重矩阵的矩阵乘运算时，还可以进一步将其分为多个较小的非近似正方形的第二分形信号矩阵和第二权重矩阵，通过多个第二分形信号矩阵和多个第二权重矩阵进行矩阵乘和累加运算，从而可以进一步提高分形矩阵的灵

活性。

第二方面，提供一种信号处理装置，该装置包括：获取单元，用于获取信号矩阵和权重矩阵，信号矩阵为二维矩阵且包括多个计算机可处理的待处理信号，权重矩阵为二维矩阵且包括多个权重系数，信号矩阵的列数与权重矩阵的行数相等；处理单元，用于分块信号矩阵，得到 X 行 H 列的多个第一分形信号矩阵，以及分块权重矩阵，得到 H 行 Y 列的多个第一分形权重矩阵，每个第一分形信号矩阵和每个第一分形权重矩阵均满足非近似正方形，以及将多个第一分形信号矩阵和多个第一分形权重矩阵进行矩阵乘和累加运算，得到多个矩阵运算结果，所述多个矩阵运算结果用于形成信号处理结果，每个矩阵运算结果包括多个矩阵乘结果的累加，每个矩阵乘结果由一个第一分形信号矩阵与一个第一分形权重矩阵进行矩阵乘运算得到。可选地，该装置还包括输出单元，用于输出信号处理结果，该信号处理结果包括所述多个矩阵运算结果。

在第二方面的一种可能的实现方式中，满足非近似正方形包括，矩阵的行数与列数的差值的绝对值大于或等于 2。

在第二方面的一种可能的实现方式中，处理单元包括第一缓存器和第二缓存器，处理单元具体用于：通过第一缓存器分别多次从信号矩阵中读取 X 行 H 列的多个第一分形信号矩阵；通过第二缓存器分别多次从所述权重矩阵中读取 H 行 Y 列的多个第一分形权重矩阵。

在第二方面的一种可能的实现方式中，处理单元还包括第三缓存器，处理单元还用于：向第三缓存器中写入矩阵乘运算结果或至少两个矩阵乘运算结果的累加。

在第二方面的一种可能的实现方式中，对于一个第一分形信号矩阵与一个第一分形权重矩阵进行矩阵乘运算，处理单元还用于：分块第一分形信号矩阵，得到 x 行 h 列的多个第二分形信号矩阵，以及分块第一分形权重矩阵，得到 h 行 y 列的多个第二分形权重矩阵，每个第二分形信号矩阵和每个第二分形权重矩阵均满足非近似正方形；将多个第二分形信号矩阵和多个第二分形权重矩阵进行矩阵乘和累加运算，得到多个矩阵运算结果。

第三方面，提供一种信号处理装置，该装置包括：输入接口，用于获取信号矩阵和权重矩阵，信号矩阵为二维矩阵，权重矩阵为二维矩阵且包括多个权重系数，信号矩阵的列数与权重矩阵的行数相等；处理器，被配置为可处理如下操作：分块信号矩阵，得到 X 行 H 列的多个第一分形信号矩阵，以及分块权重矩阵，得到 H 行 Y 列的多个第一分形权重矩阵，每个第一分形信号矩阵和每个第一分形权重矩阵均满足非近似正方形；将多个第一分形信号矩阵和多个第一分形权重矩阵进行矩阵乘和累加运算，得到多个矩阵运算结果，所述多个矩阵运算结果用于形成信号处理结果，每个矩阵运算结果包括多个矩阵乘结果的累加，每个矩阵乘结果由一个第一分形信号矩阵与一个第一分形权重矩阵进行矩阵乘运算得到。可选地，该装置还包括输出接口，用于输出信号处理结果，信号处理结果包括所述多个矩阵运算结果。

在第三方面的一种可能的实现方式中，满足非近似正方形包括：矩阵的行数与列数的差值的绝对值大于或等于 2。

在第三方面的一种可能的实现方式中，处理器包括第一缓存器和第二缓存器，处理器还执行以下操作：通过第一缓存器分别多次从信号矩阵中读取 X 行 H 列的多个第

一分形信号矩阵；通过第二缓存器分别多次从权重矩阵中读取 H 行 Y 列的多个第一分形权重矩阵。

在第三方面的一种可能的实现方式中，处理器还包括第三缓存器，处理器还执行以下操作：向第三缓存器中写入矩阵乘结果或至少两个矩阵乘结果的累加。

在第三方面的一种可能的实现方式中，对于一个第一分形信号矩阵与一个第一分形权重矩阵进行矩阵乘运算，处理器还执行以下操作：分块第一分形信号矩阵，得到 x 行 h 列的多个第二分形信号矩阵，以及分块第一分形权重矩阵，得到 h 行 y 列的多个第二分形权重矩阵，每个第二分形信号矩阵和每个第二分形权重矩阵均满足非近似正方形；将多个第二分形信号矩阵和多个第二分形权重矩阵进行矩阵乘和累加运算，得到多个矩阵运算结果。

本申请的又一方面，提供了一种计算机可读存储介质，所述计算机可读存储介质中存储有指令，当其在计算机上运行时，使得该计算机执行上述第一方面或第一方面的任一种可能的实现方式所提供的信号处理方法。

本申请的又一方面，提供了一种包含指令的计算机程序产品，当其在计算机上运行时，使得该计算机执行上述第一方面或第一方面的任一种可能的实现方式所提供的信号处理方法。

本申请的又一方面，提供了一种处理器，该处理器用于：获取信号矩阵和权重矩阵，所述信号矩阵为二维矩阵且包括多个计算机可处理的待处理信号，所述权重矩阵为二维矩阵且包括多个权重系数，所述信号矩阵的列数与所述权重矩阵的行数相等；分块所述信号矩阵，得到 X 行 H 列的多个第一分形信号矩阵，以及分块所述权重矩阵，得到 H 行 Y 列的多个第一分形权重矩阵，每个第一分形信号矩阵和每个第一分形权重矩阵均满足非近似正方形；将所述多个第一分形信号矩阵和所述多个第一分形权重矩阵进行矩阵乘和累加运算，得到多个矩阵运算结果，所述多个矩阵运算结果用于形成信号处理结果，每个矩阵运算结果包括多个矩阵乘结果的累加，每个矩阵乘结果由一个第一分形信号矩阵与一个第一分形权重矩阵进行矩阵乘运算得到。

在一种可能的实现方式中，所述满足所述非近似正方形包括：矩阵的行数与列数的差值的绝对值大于或等于 2。

在一种可能的实现方式中，所述处理器包括第一缓存器和第二缓存器，所述处理器还执行以下操作：通过所述第一缓存器分别多次从所述信号矩阵中读取 X 行 H 列的多个第一分形信号矩阵；通过所述第二缓存器分别多次从所述权重矩阵中读取 H 行 Y 列的多个第一分形权重矩阵。

在一种可能的实现方式中，所述处理器还包括第三缓存器，所述处理器还执行以下操作：向所述第三缓存器中写入所述矩阵乘结果或至少两个所述矩阵乘结果的累加。

在一种可能的实现方式中，所述处理器还执行以下操作：分块所述第一分形信号矩阵，得到 x 行 h 列的多个第二分形信号矩阵，以及分块所述第一分形权重矩阵，得到 h 行 y 列的多个第二分形权重矩阵，每个第二分形信号矩阵和每个第二分形权重矩阵均满足所述非近似正方形；将所述多个第二分形信号矩阵和所述多个第二分形权重矩阵进行矩阵乘和累加运算，得到多个矩阵运算结果。

在一种可能的实现方式中，所述处理器包括用于执行之前所述计算处理的计算单

元。可选地，所述计算单元包括乘累加单元。所述乘累加单元是用于执行乘累加运算的硬件。

可以理解地，上述提供的任一种信息处理方法的装置、计算机存储介质或者计算机程序产品均用于执行上文所提供的对应的方法，因此，其所能达到的有益效果可参考上文所提供的对应的方法中的有益效果，此处不再赘述。

附图说明

图 1 为一种矩阵分块的示意图；

图 2 为本发明实施例提供的一种设备的结构示意图；

图 3 为本发明实施例提供的一种神经网络的结构示意图；

图 4 为本发明实施例提供的一种全连接神经网络的结构示意图；

图 5 为本发明实施例提供的一种卷积神经网络的结构示意图；

图 6 为本发明实施例提供的一种卷积操作的示意图；

图 7 为本发明实施例提供的一种信号处理方法的流程示意图；

图 8 为本发明实施例提供的一种矩阵分块的示意图；

图 9 为本发明实施例提供的一种第一分形信号矩阵的示意图；

图 10 为本发明实施例提供的一种处理器的结构示意图；

图 11 为本发明实施例提供的另一种处理器的结构示意图；

图 12 为本发明实施例提供的一种信号处理装置的结构示意图。

具体实施方式

图 2 为本申请实施例提供的一种设备的结构示意图，参见图 2，该设备可以包括存储器 201、处理器 202、通信接口 203 和总线 204。其中，存储器 201、处理器 202 以及通信接口 203 通过总线 204 相互连接。存储器 201 可用于存储数据、软件程序以及模块，主要包括存储程序区和存储数据区，存储程序区可存储操作系统、至少一个功能所需的应用程序等，存储数据区可存储该设备的使用时所创建的数据等。处理器 202 用于对该设备的动作进行控制管理，比如通过运行或执行存储在存储器 201 内的软件程序和/或模块，以及调用存储在存储器 201 内的数据，执行该设备的各种功能和处理数据。通信接口 203 用于支持该设备进行通信。

其中，处理器 202 可以包括中央处理器单元，通用处理器，数字信号处理器，专用集成电路，现场可编程门阵列或者其他可编程逻辑器件、晶体管逻辑器件、硬件部件或者其任意组合。其可以实现或执行结合本申请公开内容所描述的各种示例性的逻辑方框，模块和电路。所述处理器也可以是实现计算功能的组合，例如包含一个或多个微处理器组合，数字信号处理器和微处理器的组合等等。总线 204 可以是外设部件互连标准（Peripheral Component Interconnect，PCI）总线，或者扩展工业标准结构（Extended Industry Standard Architecture，EISA）总线等。所述总线可以分为地址总线、数据总线、控制总线等。为便于表示，图 2 中仅用一条粗线表示，但并不表示仅有一根总线或一种类型的总线。

如图 3 所示，是一种神经网络的结构示意图，该神经网络 300 具有 N 个处理层，$N \geq 3$ 且 N 取自然数，该神经网络的第一层为输入层 301，负责接收输入信号，该神经网络的最后一层为输出层 303，输出神经网络的处理结果，除去第一层和最后一层的

其他层为中间层304，这些中间层共同组成隐藏层302，隐藏层中的每一层中间层既可以接收输入信号，也可以输出信号，隐藏层负责输入信号的处理过程。每一层代表了信号处理的一个逻辑级别，通过多个层，数据信号可经过多级逻辑的处理。

为便于理解，下面对本申请实施例中神经网络的处理原理进行描述，神经网络的处理通常是非线性函数 $f(x_i)$，如 $f(x_i)=\max(0,x_i)$，在一些可行的实施例中，该处理函数可以是修正线性单元（Rectified Linear Units，ReLU）、双曲正切函数（tanh）或S型函数（sigmoid）等。假设 (x_1,x_2,x_3) 是一个一维信号矩阵，(h_1,h_2,h_3) 是输出信号矩阵，W_{ij} 表示输入 x_j 与输出 h_i 之间的权重系数，权重系数构成的矩阵为权重矩阵，则该一维信号矩阵与输出信号矩阵对应的权重矩阵 W 如式（1）所示：

$$W=\begin{pmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{22} & W_{23} \\ W_{31} & W_{32} & W_{33} \end{pmatrix} \quad (1)$$

输入信号与输出信号的关系如式（2）所示，其中 b_i 为神经网络处理函数的偏置值，该偏置值对神经网络的输入进行调整从而得到理想的输出结果。

$$\begin{aligned} h_1 &= f(W_{11}x_1+W_{12}x_2+W_{13}x_3+b_1) \\ h_2 &= f(W_{21}x_1+W_{22}x_2+W_{23}x_3+b_2) \\ h_3 &= f(W_{31}x_1+W_{32}x_2+W_{33}x_3+b_3) \end{aligned} \quad (2)$$

在一些可行的实施例中该神经网络的输入信号可以是语音信号、文本信号、图像信号、温度信号等各种形式的信号，该语音信号可以是录音设备录制的语音信号、移动手机或固定电话在通话过程中接收的语音信号、以及收音机接收的电台发送的语音信号等，文本信号可以是TXT文本信号、Word文本信号、以及PDF文本信号等，图像信号可以是相机拍摄的风景信号、显监控设备捕捉的社区环境的图像信号以及门禁系统获取的人脸的面部信号等，该神经网络的输入信号包括其他各种计算机可处理的工程信号，在此不再一一列举。该神经网络的隐藏层302进行的处理可以是去除语音信号中混杂的噪音信号从而增强语音信号、对文本信号中的特定内容进行理解、以及对人脸的面部图像信号进行识别等处理。

神经网络的每个层可以包括多个节点，也可以称为神经元。全连接神经网络是一种相邻层之间各神经元全连接的神经网络，即前一层中的全部神经元与后一层中的每个神经元都连接。示例性的，图4是一种包含三层的全连接神经网络的结构示意图，层1和层2均包括四个神经元，层3包括一个神经元。图4中“+1”表示偏置神经元，用于对神经网络中每一层的输入进行调整。由于全连接网络的相邻层中神经元是全连接的，当全连接神经网络的中间层较多时，则越靠后的处理层中的信号矩阵和权重矩阵的维度会很庞大，从而导致神经网络的网络尺寸过于庞大。

卷积神经网络可以采用较小的参数模板在输入信号空间域上滑动滤波，从而解决全连接神经网络中网络尺寸过于庞大的问题。卷积神经网络与普通神经网络的区别在于，卷积神经网络包含了一个由卷积层和子采样层构成的特征抽取器。在卷积神经网络的卷积层中，一个神经元只与部分邻层神经元连接。在卷积神经网络的一个卷积层中，通常包含若干个特征平面，每个特征平面由一些矩形排列的神经元组成，同一特征平面的神经元共享权值，这里共享的权值就是卷积核。卷

积核一般以随机小数矩阵的形式初始化，在网络的训练过程中卷积核将学习得到合理的权值。卷积核带来的直接好处是减少网络各层之间的连接。子采样也叫做池化，子采样可以看作一种特殊的卷积过程，卷积和子采样大大简化了模型复杂度，减少了模型的参数。如图 5 所示，卷积神经网络由三部分构成，第一部分是输入层，第二部分由多个卷积层和多个池化层的组合组成，第三部分是输出层，输出层可以由一个全连接的多层感知机分类器构成。

卷积神经网络中的卷积层可用于对输入信号阵列和权重阵列进行卷积操作。具体的，这里以一维输入信号为例，假设输入信号为 $f(u), u=0\sim N-1$，卷积核为 $h(v), v=0\sim n-1, n\le N$，则卷积运算可以通过以下公式（3）来描述。

$$y(i)=\sum_{u=0}^{N-1} f(u)h(i-u),\ i=0\sim N+n-1 \qquad (3)$$

卷积神经网络可以广泛应用于语音识别、人脸识别、通用物体识别、运动分析、图像处理等。示例性的，以输入信号为二维矩阵为例，如图 6 所示，假设该图像在某一卷积层中对应的输入特征包括 3 个三行三列的信号矩阵，卷积核包括 6 个二行二列的权重矩阵。图 6 中示出了卷积神经网络中进行卷积操作的两种具体操作方式，一种是传统卷积操作，另一种是矩阵变换后的卷积操作。其中，传统卷积操作是将每个信号矩阵与其对应的权重矩阵进行矩阵乘运算，并将对应的矩阵乘运算的结果进行累加，得到两个输出信号矩阵，即输出特征。另一种矩阵变换后的卷积操作，对不同的信号矩阵进行了转换，得到一个同时包括 3 个信号矩阵且矩阵维度较大的输入特征矩阵；同理，对 6 个权重矩阵也进行相应的转换操作，得到一个同时包括 6 个权重矩阵且矩阵维度较大的核矩阵；之后，通过变换得到的输入特征矩阵和核矩阵进行矩阵乘运算，得到输出特征矩阵。

通过对信号矩阵和权重矩阵进行矩阵变换，可以减小矩阵乘的操作次数，进而减小读取信号矩阵和权重矩阵的开销。但是，变换后的矩阵乘运算需要的计算开销较大，因此，需要通过矩阵分块将其转换为较小的分形矩阵，并通过分形矩阵相乘得到相应的结果，即将一个大的矩阵相乘拆分为多个分形矩阵的相乘和累加。

为了便于理解，下面对本申请实施例中一种具体的信号处理方法进行描述，该信号处理方法可以在神经网络的隐藏层中的任一个中间层中进行处理。可选的，该神经网络可以是全连接神经网络，该中间层也可以称为全连接层；或者，该神经网络也可以是卷积神经网络，该中间层中进行的处理具体可以是在卷积神经网络中的卷积层中处理。

图 7 为本申请实施例提供的一种信号处理方法的流程示意图，该方法的执行主体可以是设备，具体可以是设备中具有计算功能的单元，比如神经网络处理器等，该方法包括以下几个步骤。

步骤 701：获取信号矩阵和权重矩阵，信号矩阵的列数与权重矩阵的行数相等。

其中，信号矩阵可以来自神经网络的输入层或者信号处理所在中间层中的上一层中间层，该输入信号可以是语音信号、文本信号、图像信号以及温度信号等各种可以被采集并且被处理的信号，该矩阵可以是未进行矩阵转换的矩阵，也可以是经过矩阵转换后的矩阵，该信号矩阵可以是 M 行 K 列的二维矩阵，且矩阵中包括多个计算机可

处理的待处理信号，即每个元素对应一个信号。当该信号矩阵是经过转换后的矩阵时，该信号矩阵转换前的矩阵可以是一维列向量、一维行向量、二维矩阵（比如灰度图像）、以及三维矩阵（比如RGB彩色图像）等，本申请实施例对此不作具体限定。

另外，权重矩阵由一个个权重系数构成，该权重矩阵可以是由神经网络定义的，权重系数作用于输入信号，权重系数大对应的输入信号在神经网络学习训练的过程中会被加强，权重系数小对应的输入信号在学习训练的过程中会被减弱。该权重矩阵可以是未进行矩阵转换的权重矩阵，也可以是经过矩阵转换后的权重矩阵，且为K行N列的二维权重矩阵。

步骤702：分块信号矩阵，得到X行H列的多个第一分形信号矩阵，以及分块权重矩阵，得到H行Y列的多个第一分形权重矩阵，多个第一分形信号矩阵与多个第一分形权重矩阵存在对应关系，每个第一分形信号矩阵和每个第一分形权重矩阵均满足非近似正方形。

其中，由于原始的信号矩阵和权重矩阵通常都具有较大的维度，处理器无法直接对大维度的矩阵进行运算，因此需要对信号矩阵和权重矩阵分别进行分块，分块一个矩阵是指将矩阵分成多个子块，每个子块可以称为一个分形矩阵。分块得到的多个第一分形信号矩阵的个数与多个第一分形权重矩阵的个数相等，且多个第一分形信号矩阵与多个第一分形权重矩阵之间存在对应关系，该对应关系可以是一对多的关系，也可以是多对一的关系，或者是多对多的关系，即一个第一分形信号矩阵可以对应多个第一分形权重矩阵，或者多个第一分形信号矩阵对应一个第一分形权重矩阵、或者多个第一分形信号对应多个第一分形权重矩阵。

另外，第一分形信号矩阵的列数和第一分形权重矩阵的行数均为H，即一个第一分形信号矩阵与其对应的第一分形权重矩阵满足矩阵乘规则，该矩阵乘规则是指参加矩阵乘的第一个矩阵的列数等于参加矩阵乘的第二个矩阵的行数。X、H和Y与信号矩阵的行数和列数、以及权重矩阵的行数和列数有关，且第一分形信号矩阵和第一分形权重矩阵均满足非近似正方形。

如图8所示，为一种矩阵分块的示意图。假设信号矩阵为A、权重矩阵为B、矩阵C=AB。示例性的，信号矩阵A分块得到4个第一分形信号矩阵，分别表示为A00、A01、A10和A11，权重矩阵B分块得到4个第一分形权重矩阵，分别表示为B00、B01、B10和B11。以A00、A01、A10和A11为例，则A00和A10分别对应的第一分形权重矩阵包括B00和B01，A01和A11分别对应的第一分形权重矩阵包括B01和B10，对应的第一分形权重矩阵包括B00和B01。矩阵C可以由4个矩阵C00、C01、C10和C11组成，矩阵C中的每个组成矩阵与第一分形信号矩阵和第一分形权重矩阵的关系可以如下公式（4）所示。

$$\begin{aligned} C00 &= A00B00 + A01B10 \\ C01 &= A00B01 + A01B11 \\ C10 &= A10B00 + A11B10 \\ C11 &= A10B01 + A11B11 \end{aligned} \qquad (4)$$

其中，公式（4）中矩阵C的每个矩阵块的计算可以分两步执行，比如以C00为例，可以按照如下步骤（I）-（II）来执行，通过重复利用数据*C00 _ temp*可以减少数据的

读写量，降低处理器对带宽的需求，同时节省内存的读写功耗。

$$C00\_temp = A00B00 \quad (I)$$

$$C00\_temp = C00\_temp + A01B10 \quad (II)$$

或者，公式（4）中矩阵 C 的计算中按照如下步骤（i）-（ii）的示例执行，则处理器只需获取一次 A00，则可以通过复用数量 A00 的方式，节省处理器对内存的读写功耗。

$$C00\_temp = A00B00 \quad (i)$$

$$C01\_temp = A00B01 \quad (ii)$$

本申请实施例中，当处理器计算信号矩阵与权重矩阵的矩阵乘运算时，通过将信号矩阵分块为 X 行 H 列的多个第一分形信号矩阵，以及将权重矩阵分块为 H 行 Y 列的多个第一分形权重矩阵，每个第一分形信号矩阵和每个第一分形权重矩阵均满足近似正方形，从而可以提高了分形矩阵的灵活性，进而基于多个第一分形信号矩阵和多个第一分形权重矩阵进行矩阵乘和累加运算时，可以根据不同数据的读写功耗，实现处理器计算信号矩阵与权重矩阵的矩阵乘运算的最优设计。

可选的，满足非近似正方形包括矩阵的行数与列数的差值的绝对值大于或等于 2。即第一分形信号矩阵的行数 X 和列数 H 满足$|X-H|\geq 2$，第一分形权重矩阵的行数 H 和列数 Y 满足$|H-Y|\geq 2$，即第一分形信号矩阵和第一分形权重矩阵均可以为行数与列数之间的差值大于或等于 2 的长方形矩阵，即不满足近似正方形。示例性的，假设信号矩阵 A 为 M×K 矩阵、权重矩阵 B 为 K×N 矩阵，X、H 和 Y 与 M、K 和 N 有关，第一分形信号矩阵的行数 X 和列数 H 可以如图 9 所示。

步骤 703：将多个第一分形信号矩阵和多个第一分形权重矩阵进行矩阵乘和累加运算，得到多个矩阵运算结果，所述多个矩阵运算结果用于形成信号处理结果，每个矩阵运算结果包括多个矩阵乘结果的累加，每个矩阵乘结果由一个第一分形信号矩阵与一个第一分形权重矩阵进行矩阵乘运算得到。

其中，在上面的描述中已经提到过，多个第一分形信号矩阵的个数与多个第一分形权重矩阵的个数可以相等，也可以不相等。多个第一分形信号矩阵与多个第一分形权重矩阵之间可以存在对应关系，第一分形信号矩阵与第一分形权重矩阵之间满足矩阵乘规则，一个第一分形矩阵与对应于该第一分形矩阵的第一权重矩阵进行矩阵乘运算得到一个矩阵乘运算结果。因此，根据多个第一分形信号矩阵与多个第一分形权重矩阵的对应关系，将多个第一分形信号矩阵和多个第一分形权重矩阵进行矩阵乘和累加运算。以上计算过程可以得到包括多个矩阵乘运算结果的输出矩阵，一个矩阵运算结果包括多个矩阵乘结果的累加，每个矩阵乘运算结果可以包括多个计算机可处理的输出信号。

需要说明的是，如果将每个第一分形信号矩阵和每个第一分形权重矩阵各作为一个元素，则多个第一分形信号矩阵和多个第一分形权重矩阵的矩阵乘和累加运算，与包含多个元素的两个矩阵之间的相乘运算的计算方式类似。

为便于理解，这里以上述公式（4）中的为例进行说明，可以将矩阵 C 称为输出矩阵，C00、C01、C10 和 C11 称为矩阵运算结果，输出矩阵 C 包括四个矩阵运算结果。

以 C00 为例，A00 与 B00 的乘积为一个矩阵乘结果，A01 与 B10 的乘积也为一个矩阵乘结果，这两个矩阵乘结果在输出矩阵 C 中都对应 C00 的位置，则将这两个矩阵乘结果的累加称为一个矩阵乘运算结果。

步骤 704：输出信号处理结果，该信号处理结果包括多个所述矩阵运算结果。

当获取多个矩阵乘运算结果之后，处理器还可以输出信号处理结果，该信号处理结果包括多个矩阵运算结果。该多个矩阵运算结果组成的输出矩阵可以是二维矩阵（比如，灰度图像），该输出矩阵对应的输出信号可以是与输入信号对应的语音信号、文本信号、图像信号以及温度信号等各种可以被处理或者可以被播放、显示的信号。可选的，该信号处理结果可以去往信号处理所在中间层的下一层中间层或者神经网络的输出层。

进一步的，如图 10 所示，处理器可以包括乘累加（Multiply - Accumulator，MAC）单元、第一缓存器、第二缓存器和第三缓存器，处理器中的 MAC 单元可以与第一缓存器、第二缓存器和第三缓存器直接进行交互，处理器还可以包括第四缓存器，对于第一缓存器和第二缓存器相连，MAC 单元可以通过第一缓存器和第二缓存器与第三缓存器进行交互。其中，MAC 单元用于执行具体的乘加运算，第四缓存器可用于存储信号矩阵和权重矩阵，第一缓存器可用于存储信号矩阵的第一分形信号矩阵，第二缓存器可用于存储权重矩阵的第一分形权重矩阵，第三缓存器用于存储矩阵乘结果、或者至少两个矩阵乘结果的累加，至少两个矩阵乘结果的累加可以是一个矩阵运算结果。

例如，以上处理器中的各个单元可以电路硬件，包括电不限于晶体管、逻辑门、或基本运算单元等的一个或多个。再例如，信号矩阵和权重矩阵可以是来自处理器之前计算所生成的矩阵，也可以来自处理器之外的其他设备，如硬件加速器或其他处理器等。本实施例的处理器用于获取信号矩阵和权重矩阵并依照之前实施例的方法执行计算。图 10 中的处理器的具体运算过程可参照之前的方法实施例。

具体的，分块信号矩阵，得到 X 行 H 列的多个第一分形信号矩阵，包括：MAC 单元通过第一缓存器分别多次从信号矩阵中读取 X 行 H 列的第一分形信号矩阵，以得到 X 行 H 列的多个第一分形信号矩阵。其中，处理器可以从第四缓存器中读取 X 行 H 列的第一分形信号矩阵，并将 X 行 H 列的第一分形信号矩阵存储在第一缓存器中。第一缓存器的容量 V1 可以是固定的，且可以和 X 与 H 的乘积相等，即 V1=X×H，第一分形信号矩阵可以填充满第一缓存器。

具体的，分块权重矩阵，得到 H 行 Y 列的多个第一分形权重矩阵，包括：MAC 单元通过第二缓存器分别多次从权重矩阵中读取 H 行 Y 列的第一分形权重矩阵，以得到 H 行 Y 列的多个第一分形权重矩阵。其中，处理器可以从第四缓存器中读取 H 行 Y 列的第一分形权重矩阵，并将 H 行 Y 列的第一分形权重矩阵存储在第二缓存器中。第二缓存器的容量 V2 可以是固定的，且可以和 H 与 Y 的乘积等于，即 V2=H×Y，第一分形权重矩阵可以填充满第二缓存器。

其中，X 与第一缓存器的第一读写功耗正相关，Y 与第二缓存器的第一读写功耗正相关，H 分别与第一缓存器和第二缓存器的第一读写功耗反相关。X 与第一缓存器的第一读写功耗正相关是指，当 X 越大时，第一缓存器的第一读写功耗越大，当 X 越小时，第一缓存器的第一读写功耗越小，比如，X 与第一缓存器的第一读写功耗成正

比。Y 与第二缓存器的第一读写功耗正相关是指，当 Y 越大时，第二缓存器的第一读写功耗越大，当 Y 越小时，第二缓存器的第一读写功耗越小，比如，Y 与第二缓存器的第一读写功耗成正比。H 与第一缓存器和第二缓存器的第一读写功耗反相关是指，当 H 越大时，第一缓存器和第二缓存器的第一读写功耗越小，当 H 越小时，第一缓存器和第二缓存器的第一读写功耗越大，比如，H 与第一缓存器和第二缓存器的第一读写功耗成反比。

为便于理解，这里以一个第一分形信号矩阵和一个第一分形权重矩阵的矩阵乘为例，对 MAC 单元读取 X 行 H 列的第一分形信号矩阵和读取 H 行 Y 列的第一分形权重矩阵的过程中，X、Y 和 H 与第一缓存器和第二缓存器的读写功耗的关系进行详细说明。

当 MAC 单元读取第一分形信号矩阵时，MAC 单元需要先从第四缓存器存储的信号矩阵中通过第一缓存器读取一个 X 行 H 列的第一分形信号矩阵，即从第四缓存器读取的第一分形信号矩阵写入第一缓存器中，再从第一缓存器中读取第一分形信号矩阵。同理，当 MAC 单元读取第一分形权重矩阵时，MAC 单元需要先从第四缓存器存储的权重矩阵中通过第二缓存器读取一个 H 行 Y 列的第一分形权重矩阵，即从第四缓存器读取的第一分形权重矩阵写入第二缓存器中，再从第二缓存器中读取第一分形权重矩阵。当 MAC 单元在进行矩阵乘运算时，由于矩阵乘运算是用第一分形信号矩阵中的每一行分别乘以第一分形权重矩阵中的每一列，所以，MAC 单元需要通过 X 次读操作从第一缓存器中读取第一分形信号矩阵的 X 行，以及通过 Y 次读操作从第二缓存器中读取第一分形权重矩阵的 Y 列。由此可知，当 X 越大时，第一缓存器的读操作次数越大，进而第一缓存器的第一读写功耗越大，当 X 越小时，第一缓存器的读操作次数越小，进而第一缓存器的第一读写功耗越小，所以 X 与第一缓存器的第一读写功耗正相关。同理，当 Y 越大时，第二缓存器的读操作次数越大，进而第二缓存器的第一读写功耗越大，当 Y 越小时，第二缓存器的读操作次数越小，进而第二缓存器的第一读写功耗越小，所以 Y 与第二缓存器的第一读写功耗正相关。

由于第一缓存器和第二缓存器的容量通常是固定的，假设第一缓存器的容量 V1=X×H 固定，则当 X 越大时 H 越小，当 X 越小时 H 越大，所以 H 与第一缓存器的第一读写功耗反相关。假设第二缓存器的容量 V2=H×Y 固定，则当 Y 越大时 H 越小，当 Y 越小时 H 越大，所以 H 与第二缓存器的第一读写功耗反相关。可选的，第一缓存器的容量和第二缓存器的容量可以相等，即 X×H=H×Y，则 X 和 Y 相等。

具体的，当 MAC 单元进行矩阵乘运算，第三缓存器用于存储矩阵乘结果、或者至少两个矩阵乘结果的累加时，该方法还可以包括：向第三缓存器中写入矩阵乘结果或至少两个矩阵乘结果的累加；和/或，从第三缓存器中读取矩阵乘结果或至少两个矩阵乘结果的累加。

其中，X 和 Y 分别与第三缓存器的读写功耗反相关，H 与第三缓存器的读写功耗正相关。X 和 Y 分别与第三缓存器的读写功耗反相关，是指当 X 和 Y 越大时，第三缓存器的读写功耗越小，当 X 和 Y 越小时，第三缓存器读写功耗越大，比如，X 和 Y 分别与第三缓存器的读写功耗成反比。

为便于理解，这里以一个第一分形信号矩阵和一个第一分形权重矩阵的矩阵乘为

例，对 MAC 单元进行 X 行 H 列的第一分形信号矩阵与 H 行 Y 列的第一分形权重矩阵的矩阵乘运算过程中，X、Y 和 H 与第三缓存器的读写功耗的关系进行详细说明。

矩阵乘运算是用第一分形信号矩阵中的每一行分别乘以第一分形权重矩阵中的每一列，在进行行列相乘，以第一分形信号矩阵中的第 1 行（第 1 行中包括 H 个行元素）与第一分形权重矩阵中的第 1 列（第 1 行中包括 H 个列元素）相乘为例，则当 MAC 单元进行 H 个行元素与 H 个列元素的乘加运算时，MAC 单元先计算第一个行元素与第一个列元素的第一乘积之后，将第一乘积写入第三缓存器中，再计算第二个行元素与第二个列元素的第二乘积，之后从第三缓存器中读取第一乘积，将第一乘积和第二乘积进行累加后写入第三缓存器中，以此类推，直到计算得到 H 个行元素与 H 个列元素的乘加运算的结果。

由此可知，当 H 越大时，则 MAC 单元对第三缓存器的读写次数越大，进而第三缓存器的读写功耗越大，当 H 越小时，则 MAC 单元对第三缓存器的读写次数越小，进而第三缓存器的读写功耗越小，所以 H 与第三缓存器的读写功耗反相关。

由于第一缓存器和第二缓存器的容量通常是固定的，即 V1=X×H 和 V2=H×Y 是固定的，因此，当 H 越大时，则 X 和 Y 越小，当 H 越小时，则 X 和 Y 越大。因此，X 和 Y 分别与第三缓存器的读写功耗反相关，H 与第三缓存器的读写功耗正相关。

需要说明的是，上述是以一个第一分形信号矩阵和一个第一分形权重矩阵的矩阵乘为例进行说明，此时第三缓存器中可用于存储的是一个行元素与一个列元素的乘积，或者乘积的累加也是示例性的，并不对本申请实施例构成限定。当 MAC 单元将多个第一分形信号矩阵和多个第一分形权重矩阵进行矩阵乘和累加运算的过程中，第三缓存器用于存储一个矩阵乘结果、或者至少两个矩阵乘结果。

另外，这里的第一缓存器的第一读写功耗包括向第一缓存器写入第一分形信号矩阵的功耗，以及从第一缓存器中读取第一分形信号矩阵的功耗；第二缓存器的第一读写功耗包括向第二缓存器写入第一分形权重矩阵的功耗，以及通过从第二缓存器中读取第一分形权重矩阵的功耗；第三缓存器的读写功耗包括向第三缓存器写入和读取矩阵乘运算结果或者至少两个矩阵乘运算结果的累加的功耗。

比如，以图 8 所示的矩阵分块为例，当处理器在按照上述步骤（I）和步骤（II）确定 C00 时，处理器可以在步骤（I）执行完成后可以将矩阵乘结果 *C00 _ temp* 的结果存储在第三缓存器中，当完成 A01 与 B01 的矩阵乘运算得到第二个矩阵乘结果之后，处理器可以从第三缓存器中读取 *C00 _ temp*，并将其与第二个矩阵乘结果进行累加，以得到矩阵运算结果 C00，将 C00 存储在第三缓存器中。

进一步的，如果获取的第一分形信号矩阵和第一分形权重矩阵的维护仍然较大，处理器无法一次完成对一个第一分形信号矩阵和一个第一分形权重矩阵的运算，则还可以对其进一步分块处理，以得到处理器可以处理的粒度。

比如，以图 8 所示的矩阵分块为例，如果分块后的 A00、A01、A10 和 A11，以及 B00、B01、B10 和 B11 的粒度仍然较大，比如，处理器无法完成上述步骤（I）或步骤（II）的运算，则以步骤（I）的计算为例，处理器可以进一步将其分解为如下公式（5）所示。

$$\begin{aligned} C00(00) &= A00(00)B00(00) + A00(01)B00(10) \\ C00(01) &= A00(00)B00(01) + A00(01)B00(11) \\ C00(10) &= A00(10)B00(00) + A00(11)B00(10) \\ C00(11) &= A00(10)B00(01) + A00(11)B00(11) \end{aligned} \quad (5)$$

其中，矩阵 A00(00)、A00(01)、A00(10)和 A00(11)可以称为 A00 的分形矩阵，B00(00)、B00(01)、B00(10)和 B00(11)可以称为 B00 的分形矩阵；相应的，矩阵 C00 可以由 C00(00)、C00(01)、C00(10)和 C00(11)组成。

在本申请实施例中，对于一个第一分形信号矩阵与对应的一个第一分形权重矩阵进行矩阵乘运算，得到一个矩阵乘运算结果，包括：分块第一分形信号矩阵，得到 x 行 h 列的多个第二分形信号矩阵，以及分块第一分形权重矩阵，得到 h 行 y 列的多个第二分形权重矩阵，每个第二分形信号矩阵和每个第二分形权重矩阵均满足非近似正方形；将多个第二分形信号矩阵和多个第二分形权重矩阵进行矩阵乘和累加运算，得到多个矩阵运算结果。可选的，$|x-h|\geq 2$，$|h-y|\geq 2$。

结合图 10，参见图 11，当处理器还包括第一寄存器、第二寄存器和第三寄存器时，处理器可以通过第一寄存器与第一缓存器进行交互，通过第二寄存器与第二缓存器进行交互，以及通过第三寄存器与第三缓存器进行交互。其中，第一寄存器可以用于存储第二分形信号矩阵，即用于存储最小的分形信号矩阵，比如，用于图 8 中的存储 A00(00)、A00(01)、A00(10)或者 A00(11)。第二寄存器用于存储第二分形权重矩阵，即用于存储最小的分形权重矩阵，比如，用于存储图 8 中的 B00(00)、B00(01)、B00(10)或者 B00(11)。第三寄存器用于存储多个第二分形信号矩阵与多个第二分形权重矩阵的矩阵乘运算过程中的矩阵乘运算结果或者至少两个矩阵乘运算结果的累加，比如，用于存储图 8 中的 A00(00)B00(00)或者 A00(00)B00(01)等。

具体的，当 MAC 单元进行矩阵乘运算时，MAC 单元读取的第一分形权重矩阵存储在第一缓存器中，读取的第一分形权重矩阵存储在第二缓存器中，MAC 单元通过第一寄存器分别从第一缓存器中读取 x 行 h 列的第二分形信号矩阵，通过第二寄存器分别从第二缓存器中读取 x 行 h 列的第二分形权重矩阵。MAC 单元将多个第二分形信号矩阵与多个第二分形权重矩阵的矩阵乘运算过程中的矩阵乘结果或者至少两个矩阵乘结果的累加通过第三寄存器存储在第三缓存器中，和/或，从第三缓存器中读取该矩阵乘结果或者至少两个矩阵乘结果的累加。

相应的，当根据多个第二分形信号矩阵和多个第二分形权重矩阵进行矩阵乘和累加运算时，x 与第一缓存器的第二读写功耗正相关，y 与第二缓存器的第二读写功耗正相关，h 与第一缓存器和第二缓存器的第二读写功耗反相关。此外，x 和 y 分别与第三缓存器的读写功耗反相关，h 与第三缓存器的读写功耗正相关。其中，x、h 和 y 与不同缓存器之间的读写功耗的关系分析与上述 X、H 和 Y 与不同缓存器之间的读写功耗的关系的分析类似，具体参见上述描述，本申请实施例在此不再赘述。

需要说明的是，这里的第一缓存器的第二读写功耗包括向第一缓存器写入第一分形信号矩阵的功耗，以及通过第一寄存器从第一缓存器中读取第二分形信号矩阵的功耗；第二缓存器的第二读写功耗包括向第二缓存器写入第一分形权重矩阵的功耗，以

及通过第二寄存器从第二缓存器中读取第二分形权重矩阵的功耗；第三缓存器的读写功耗包括通过第三寄存器向第三缓存器写入和读取矩阵乘运算结果或者至少两个矩阵乘运算结果的累加的功耗。

综上所述，以第一分形信号矩阵和第一分形权重矩阵为例，若第一缓存器和第二缓存器的容量相等（即 X=Y，X×H=H×Y=常数）时，可以通过如下公式（6）表示 MAC 单元的总功耗与信号矩阵的行数和列数、权重矩阵的行数和列数、以及第一分形信号矩阵的行数 X 之间的关系。

$$E(X) = G_1(M, N, K)X + G_2(M, N, K)/X + G_3(M, N, K) \qquad (6)$$

其中，X 是自变量，M 和 K 分别为信号矩阵的行数和列数，K 和 N 分别为权重矩阵的行数和列数，G_1、G_2 和 G_3 是与 M、N 和 K 相关的子函数。

进而，在对信号矩阵和权重矩阵进行分块时，可以根据功耗最低原则确定对应的 X，相应的也即是确定 Y 和 H，进而获取多个第一分形信号矩阵和多个第一分形权重矩阵，进行矩阵乘和累加运算时，可以实现处理器的功耗的最优设计。由于不同设备的功耗参数不同，如何针对 X，Y 和 Z 进行最优功耗设计可以结合对缓存器的性能参数理解和实际测试进行，具体取决于实际应用场景和器件选型，本实施例对此不做过多展开。

在实际应用中，也可以根据不同缓存器的容量、处理器的功耗和不同缓存器的带宽等确定第一分形信号矩阵的行数和列数，以及第一分形权重矩阵的行数和列数，从而在根据多个第一分形信号矩阵和多个第一分形权重矩阵确定输出矩阵时，可以使不同缓存器的容量和带宽得到充分利用，同时尽可能的降低处理器的功耗。根据以上实施例的介绍，系统功耗与矩阵行数、列数、各缓存器的读写次数、或各缓存器的性能等各类参数存在一定关系，为了针对功耗进行优化，需要对各个缓存器的读写配置参数进行灵活调整，以利于降低功耗。为了适应这种灵活配置，本实施例设计了相关方法和装置，通过利用满足非近似正方形的分形信号矩阵和分形权重矩阵来执行计算，而不再将分形信号矩阵和分形权重矩阵严格限制在正方形，提高了设计灵活性，适应对缓存器的不同读写要求。

上述主要从设备的角度对本申请实施例提供的信号处理方法进行了介绍。可以理解的是，该设备为了实现上述功能，其包含了执行各个功能相应的硬件结构和/或软件模块。本领域技术人员应该很容易意识到，结合本文中所公开的实施例描述的各示例的网元及算法步骤，本申请能够以硬件或硬件和计算机软件的结合形式来实现。某个功能究竟以硬件还是计算机软件驱动硬件的方式来执行，取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本申请的范围。

本申请实施例可以根据上述方法示例对信号处理装置进行功能模块的划分，例如，可以对应各个功能划分各个功能模块，也可以将两个或两个以上的功能集成在一个处理模块中。上述集成的模块既可以采用硬件的形式实现，也可以采用软件功能模块的形式实现。需要说明的是，本申请实施例中对模块的划分是示意性的，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式。

在采用对应各个功能划分各个功能模块的情况下，图12示出了上述实施例中所涉

及的信号处理装置的一种可能的结构示意图，该信号处理装置包括：获取单元1201、处理单元1202和输出单元1203。其中，获取单元1201用于支持该信号处理装置执行图7中的步骤701；处理单元1202用于支持该信号处理装置执行图7中的步骤702和703，和/或用于本文所描述的技术的其他过程；输出单元1203用于该信号处理装置执行图7中的步骤704。

上面从模块化功能实体的角度对本申请实施例中的一种信号处理装置进行描述，下面从处理器硬件处理的角度对本申请实施例中的一种信号处理装置进行描述。

本申请实施例提供一种信号处理装置，该设备的结构可以如图 2 所示，该信号处理装置包括：存储器 201、处理器 202、通信接口 203 和总线 204。其中，通信接口 203 可以包括输入接口 2031 和输出接口 2032。

输入接口 2031：该输入接口用于获取信号矩阵和/或权重矩阵，该输入接口可以通过选择器实现获取信号矩阵和获取权重矩阵的切换；在一些可行的实施例中，该输入接口可用以分时复用的方式获取上述的信号矩阵或权重矩阵；在一些可行的实施例中，该输入接口可以有两个，分别实现信号矩阵和权重矩阵的获取，例如可实现同时获取信号矩阵和权重矩阵。

处理器 202：被配置为可处理上述信号处理方法的步骤 702-步骤 703 部分的功能。在一些可行的实施例中，该处理器可以是单处理器结构、多处理器结构、单线程处理器以及多线程处理器等，在一些可行的实施例中，该处理器可以集成在专用集成电路中，也可以是独立于集成电路之外的处理器芯片。

输出接口 2032：该输出接口用于输出上述信号处理方法中的信号处理结果，在一些可行的实施例中，该信号处理结果可以由处理器直接输出，也可以先被存储于存储器中，然后经存储器输出；在一些可行的实施例中，可以只有一个输出接口，也可以有多个输出接口。在一些可行的实施例中，该输出接口输出的信号处理结果可以送到存储器中存储，也可以送到下一个信号处理装置继续进行处理，或者送到显示设备进行显示、送到播放器终端进行播放等。

存储器 201：该存储器中可存储上述的信号矩阵、信号处理结果、权重矩阵、以及配置处理器的相关指令等。在一些可行的实施例中，可以有一个存储器，也可以有多个存储器；该存储器可以是软盘，硬盘如内置硬盘和移动硬盘，磁盘，光盘，磁光盘如 CD_ROM、DCD_ROM，非易失性存储设备如 RAM、ROM、PROM、EPROM、EEPROM、闪存、或者技术领域内所公知的任意其他形式的存储介质。

本申请实施例提供的上述信号处理装置的各组成部分分别用于实现相对应的前述信号处理方法的各步骤的功能，由于在前述的信号处理方法实施例中，已经对各步骤进行了详细说明，在此不再赘述。

本申请实施例还提供了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指令，当其在一个设备（比如，该设备可以是单片机，芯片、计算机等）上运行时，使得该设备执行上述信号处理方法的步骤701-步骤704中的一个或多个步骤。上述信号处理装置的各组成模块如果以软件功能单元的形式实现并作为独立的产品销售或使用时，可以存储在所述计算机可读取存储介质中。

基于这样的理解，本申请实施例还提供一种包含指令的计算机程序产品，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或者该技术方案的全部或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一个存储介质中，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）或其中的处理器执行本申请各个实施例所述方法的全部或部分步骤。

最后应说明的是：以上所述，仅为本申请的具体实施方式，但本申请的保护范围并不局限于此，任何在本申请揭露的技术范围内的变化或替换，都应涵盖在本申请的保护范围之内。因此，本申请的保护范围应以所述权利要求的保护范围为准。

权 利 要 求 书

1、一种信号处理方法，其特征在于，应用于包含处理器的设备中，所述方法包括：

获取信号矩阵和权重矩阵，所述信号矩阵为二维矩阵且包括多个计算机可处理的待处理信号，所述权重矩阵为二维矩阵且包括多个权重系数，所述信号矩阵的列数与所述权重矩阵的行数相等；

分块所述信号矩阵，得到X行H列的多个第一分形信号矩阵，以及分块所述权重矩阵，得到H行Y列的多个第一分形权重矩阵，每个第一分形信号矩阵和每个第一分形权重矩阵均满足非近似正方形；

将所述多个第一分形信号矩阵和所述多个第一分形权重矩阵进行矩阵乘和累加运算，得到多个矩阵运算结果，所述多个矩阵运算结果用于形成信号处理结果，每个矩阵运算结果包括多个矩阵乘结果的累加，每个矩阵乘结果由一个第一分形信号矩阵与一个第一分形权重矩阵进行矩阵乘运算得到。

2、根据权利要求1所述的信号处理方法，其特征在于，所述满足所述非近似正方形包括：矩阵的行数与列数的差值的绝对值大于或等于2。

3、根据权利要求1或2所述的信号处理方法，其特征在于，所述处理器包括第一缓存器和第二缓存器，所述分块所述信号矩阵，得到X行H列的多个第一分形信号矩阵，以及分块所述权重矩阵，得到H行Y列的多个第一分形权重矩阵，包括：

通过所述第一缓存器分别多次从所述信号矩阵中读取X行H列的多个第一分形信号矩阵；

通过所述第二缓存器分别多次从所述权重矩阵中读取H行Y列的多个第一分形权重矩阵。

4、根据权利要求1-3任一项所述的信号处理方法，其特征在于，所述处理器还包括第三缓存器，所述方法还包括：

向所述第三缓存器中写入所述矩阵乘结果或至少两个所述矩阵乘结果的累加。

5、根据权利要求1-4任一项所述的信号处理方法，其特征在于，对于一个第一分形信号矩阵与一个第一分形权重矩阵进行矩阵乘运算，得到一个矩阵乘结果，包括：

分块所述第一分形信号矩阵，得到x行h列的多个第二分形信号矩阵，以及分块所述第一分形权重矩阵，得到h行y列的多个第二分形权重矩阵，每个第二分形信号矩阵和每个第二分形权重矩阵均满足所述非近似正方形；

将所述多个第二分形信号矩阵和所述多个第二分形权重矩阵进行矩阵乘和累加运算，得到多个矩阵运算结果。

6、一种信号处理装置，其特征在于，所述装置包括：

获取单元，用于获取信号矩阵和权重矩阵，所述信号矩阵为二维矩阵且包括多个计算机可处理的待处理信号，所述权重矩阵为二维矩阵且包括多个权重系数，所述信号矩阵的列数与所述权重矩阵的行数相等；

处理单元，用于分块所述信号矩阵，得到X行H列的多个第一分形信号矩阵，以及分块所述权重矩阵，得到H行Y列的多个第一分形权重矩阵，每个第一分形信号矩阵和每个第一分形权重矩阵均满足非近似正方形，以及将所述多个第一分形信号矩阵

和所述多个第一分形权重矩阵进行矩阵乘和累加运算，得到多个矩阵运算结果，所述多个矩阵运算结果用于形成信号处理结果，每个矩阵运算结果包括多个矩阵乘结果的累加，每个矩阵乘结果由一个第一分形信号矩阵与一个第一分形权重矩阵进行矩阵乘运算得到。

7、根据权利要求 6 所述的信号处理装置，其特征在于，所述满足所述非近似正方形包括：矩阵的行数与列数的差值的绝对值大于或等于 2。

8、根据权利要求 6 或 7 所述的信号处理装置，其特征在于，所述处理单元包括第一缓存器和第二缓存器，所述处理单元，具体用于：

通过所述第一缓存器分别多次从所述信号矩阵中读取 X 行 H 列的多个第一分形信号矩阵；

通过所述第二缓存器分别多次从所述权重矩阵中读取 H 行 Y 列的多个第一分形权重矩阵。

9、根据权利要求 6-8 任一项所述的信号处理装置，其特征在于，所述处理单元还包括第三缓存器，所述处理单元，还用于：

向所述第三缓存器中写入所述矩阵乘结果或至少两个所述矩阵乘结果的累加。

10、根据权利要求 6-9 任一项所述的信号处理装置，其特征在于，对于一个第一分形信号矩阵与一个第一分形权重矩阵进行矩阵乘运算，所述处理单元，还用于：

分块所述第一分形信号矩阵，得到 x 行 h 列的多个第二分形信号矩阵，以及分块所述第一分形权重矩阵，得到 h 行 y 列的多个第二分形权重矩阵，每个第二分形信号矩阵和每个第二分形权重矩阵均满足所述非近似正方形；

将所述多个第二分形信号矩阵和所述多个第二分形权重矩阵进行矩阵乘和累加运算，得到多个矩阵运算结果。

11、一种信号处理装置，其特征在于，所述装置包括：

输入接口，用于获取信号矩阵和权重矩阵，所述信号矩阵为二维矩阵且包括多个计算机可处理的待处理信号，所述权重矩阵为二维矩阵且包括多个权重系数，所述信号矩阵的列数与所述权重矩阵的行数相等；

处理器，被配置为可处理如下操作：

分块所述信号矩阵，得到 X 行 H 列的多个第一分形信号矩阵，以及分块所述权重矩阵，得到 H 行 Y 列的多个第一分形权重矩阵，每个第一分形信号矩阵和每个第一分形权重矩阵均满足非近似正方形；

将所述多个第一分形信号矩阵和所述多个第一分形权重矩阵进行矩阵乘和累加运算得到多个矩阵运算结果，所述多个矩阵运算结果用于形成信号处理结果，每个矩阵运算结果包括多个矩阵乘结果的累加，每个矩阵乘结果由一个第一分形信号矩阵与一个第一分形权重矩阵进行矩阵乘运算得到。

12、根据权利要求 11 所述的信号处理装置，其特征在于，所述满足所述非近似正方形包括：矩阵的行数与列数的差值的绝对值大于或等于 2。

13、根据权利要求 11 或 12 所述的信号处理装置，其特征在于，所述处理器包括第一缓存器和第二缓存器，所述处理器还执行以下操作：

通过所述第一缓存器分别多次从所述信号矩阵中读取 X 行 H 列的多个第一分形信

号矩阵;

通过所述第二缓存器分别多次从所述权重矩阵中读取H行Y列的多个第一分形权重矩阵。

14、根据权利要求 11-13 任一项所述的信号处理装置，其特征在于，所述处理器还包括第三缓存器，所述处理器还执行以下操作:

向所述第三缓存器中写入所述矩阵乘结果或至少两个所述矩阵乘结果的累加。

15、根据权利要求 11-14 任一项所述的信号处理装置，其特征在于，对于一个第一分形信号矩阵与一个第一分形权重矩阵进行矩阵乘运算，所述处理器还执行以下操作:

分块所述第一分形信号矩阵，得到 x 行 h 列的多个第二分形信号矩阵，以及分块所述第一分形权重矩阵，得到 h 行 y 列的多个第二分形权重矩阵，每个第二分形信号矩阵和每个第二分形权重矩阵均满足所述非近似正方形;

将所述多个第二分形信号矩阵和所述多个第二分形权重矩阵进行矩阵乘和累加运算，得到多个矩阵运算结果。

16、一种可读存储介质，其特征在于，所述可读存储介质中存储有指令，当所述可读存储介质在设备上运行时，使得所述设备执行权利要求 1-5 任一项所述的信号处理方法。

17、一种计算机程序产品，其特征在于，当所述计算机程序产品在计算机上运行时，使得所述计算机执行权利要求 1-5 任一项所述的信号处理方法。

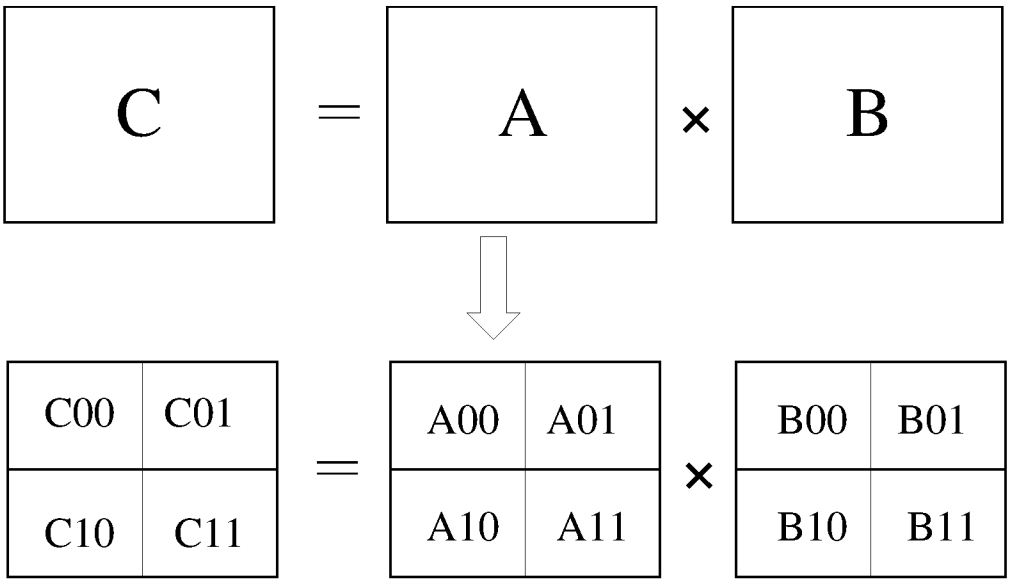
C
=
A
×
B
C00
C01
C10
C11
=
A00
A01
A10
A11
×
B00
B01
B10
B11

图 1

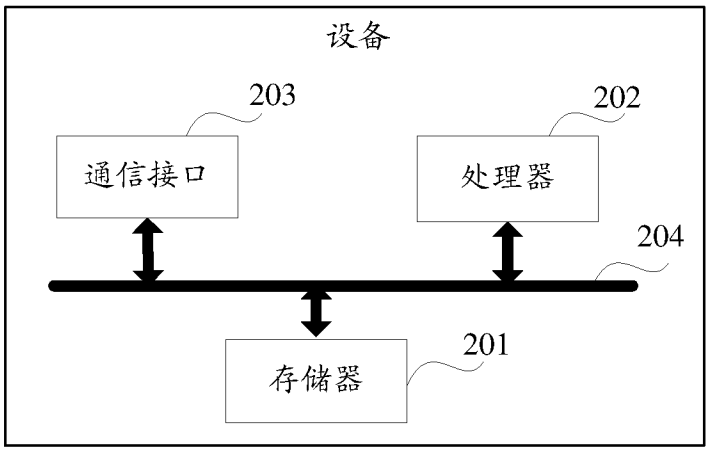
设备
203
通信接口
202
处理器
204
201
存储器

图 2

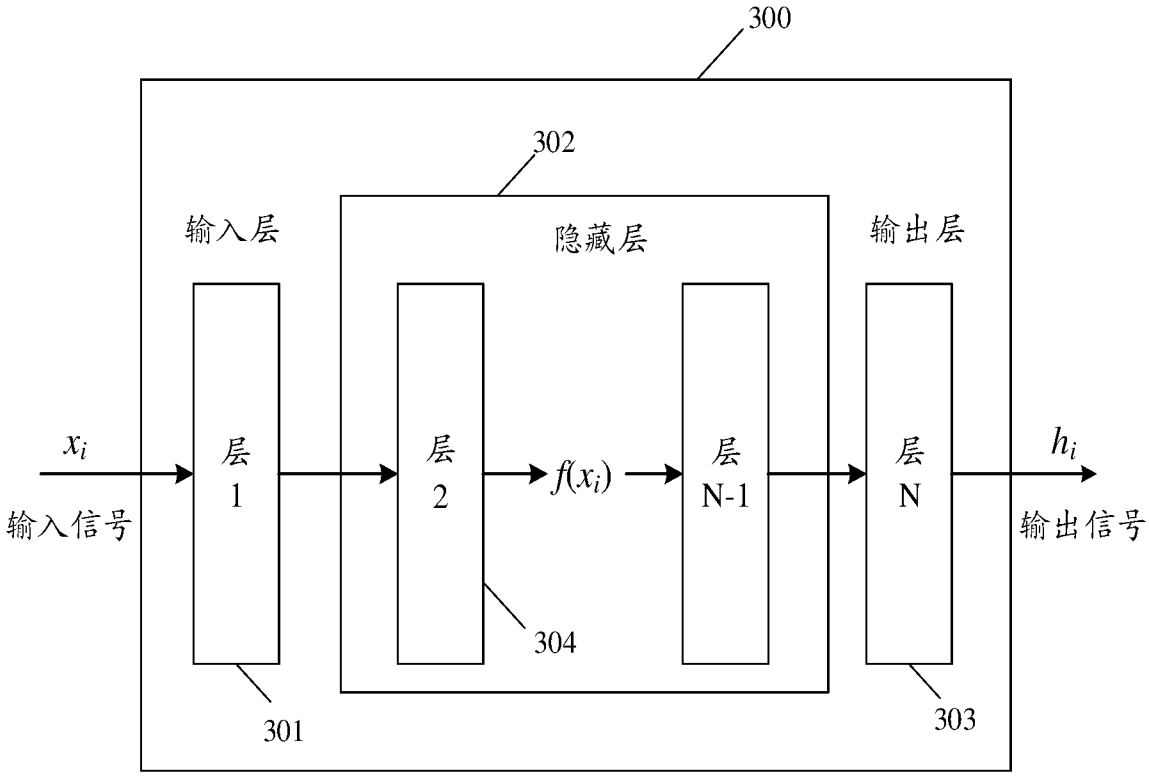
300
302
输入层
隐藏层
输出层
x_i
输入信号
层 1
层 2
$f(x_i)$
层 N-1
层 N
h_i
输出信号
304
301
303

图 3

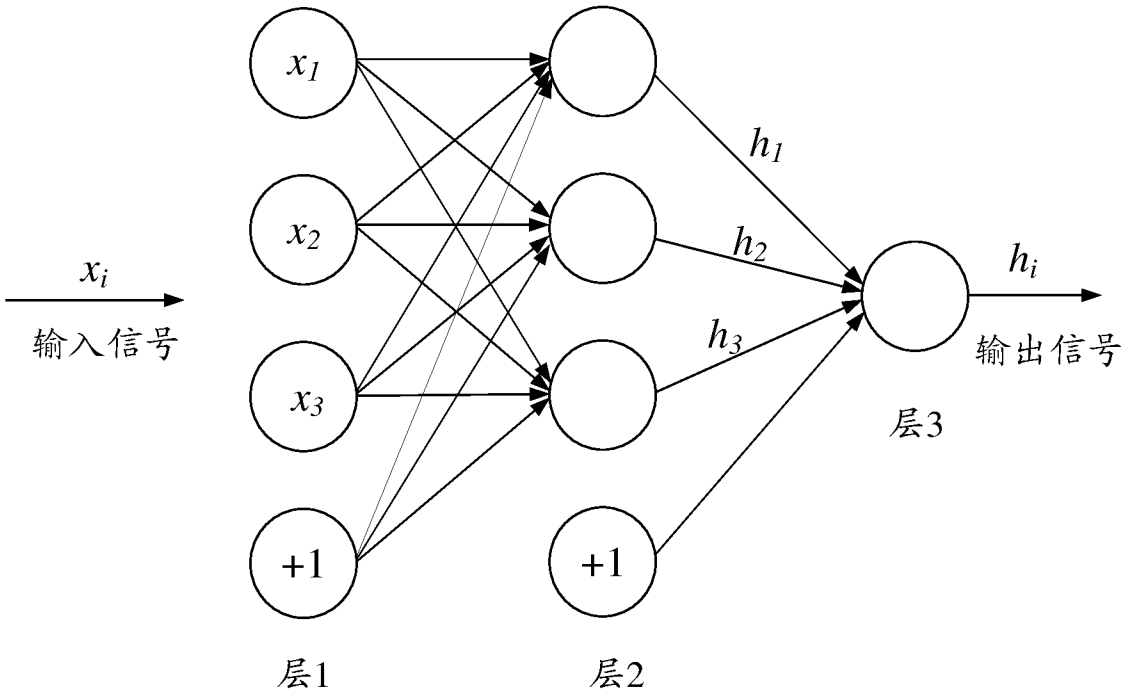
x_1
x_2
x_3
+1
+1
h_1
h_2
h_3
x_i
输入信号
h_i
输出信号
层1
层2
层3

图 4

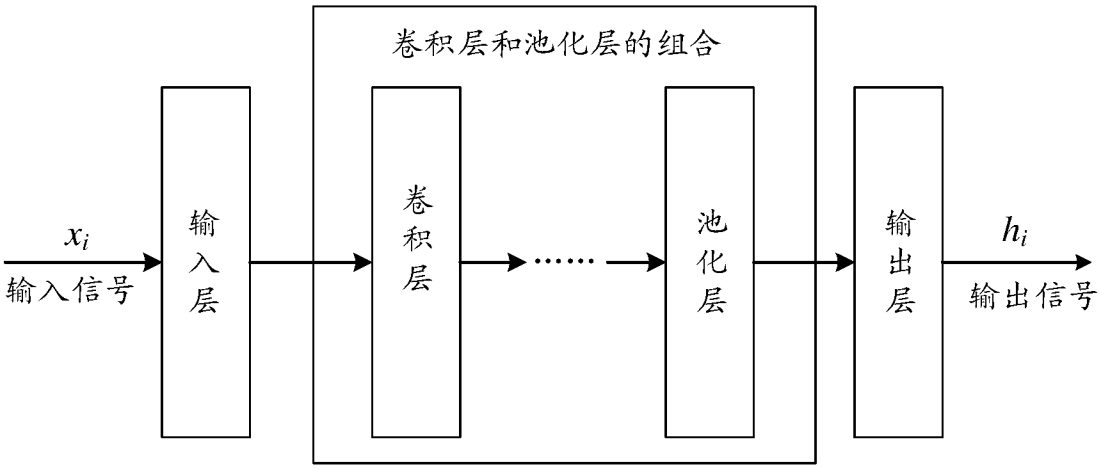
卷积层和池化层的组合
x_i
输入信号
输入层
卷积层
……
池化层
输出层
h_i
输出信号

图 5

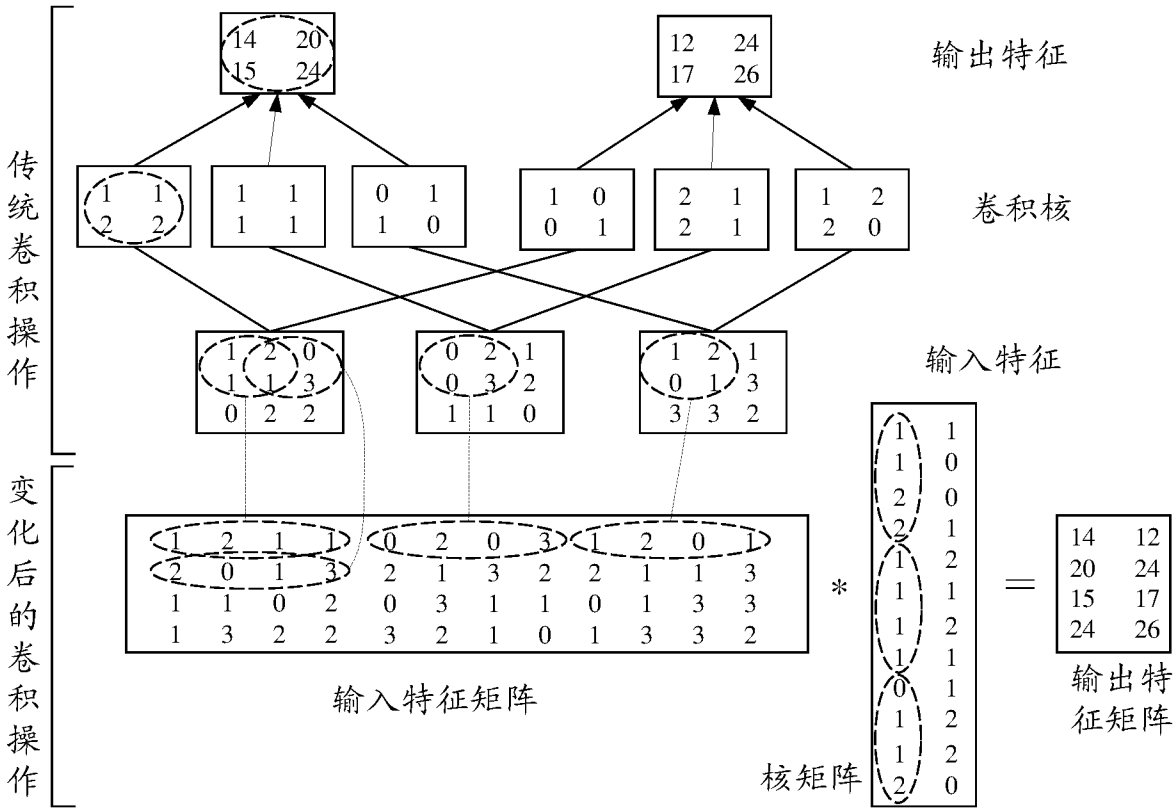
传统卷积操作
变化后的卷积操作
输出特征
卷积核
输入特征
输入特征矩阵
核矩阵
输出特征矩阵

图 6

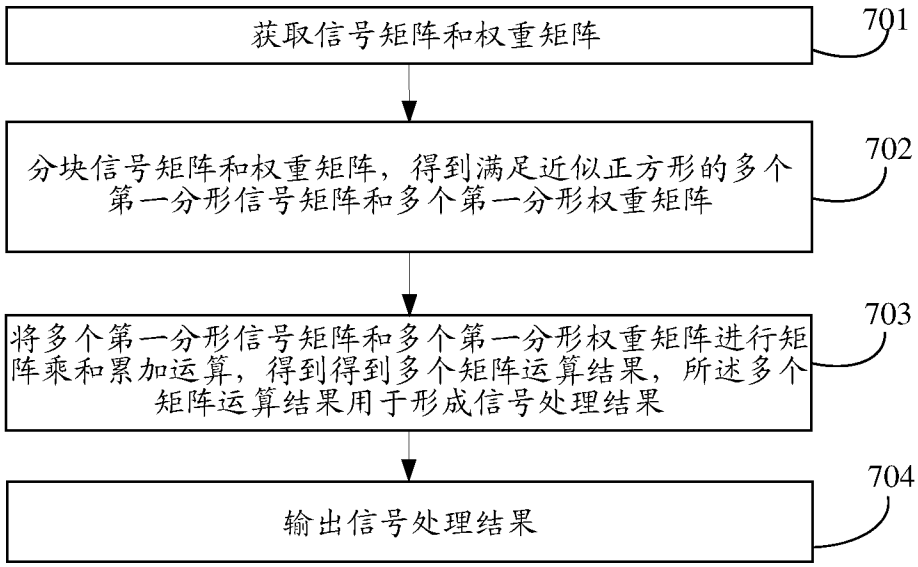
获取信号矩阵和权重矩阵
701
分块信号矩阵和权重矩阵，得到满足近似正方形的多个第一分形信号矩阵和多个第一分形权重矩阵
702
将多个第一分形信号矩阵和多个第一分形权重矩阵进行矩阵乘和累加运算，得到得到多个矩阵运算结果，所述多个矩阵运算结果用于形成信号处理结果
703
输出信号处理结果
704

图 7

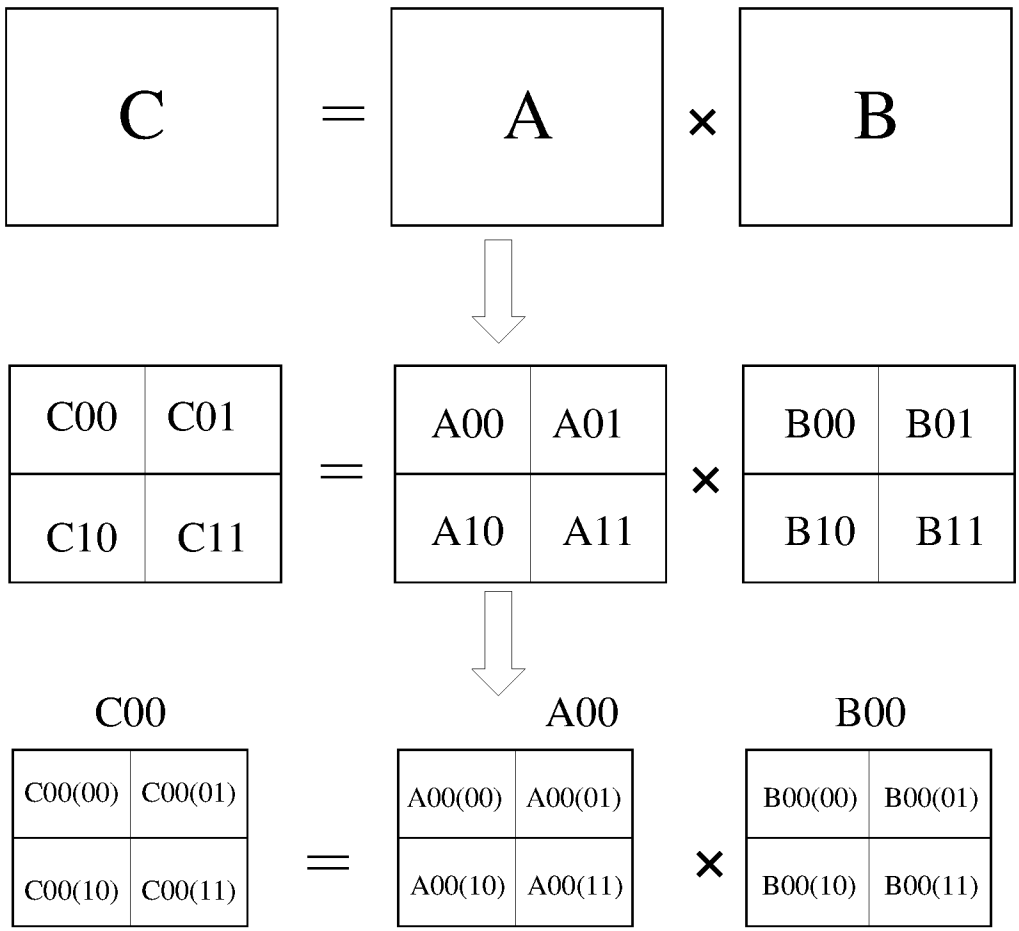
C = A × B
C00 C01 C10 C11 = A00 A01 A10 A11 × B00 B01 B10 B11
C00 = A00 × B00
C00(00) C00(01) C00(10) C00(11) = A00(00) A00(01) A00(10) A00(11) × B00(00) B00(01) B00(10) B00(11)

图 8

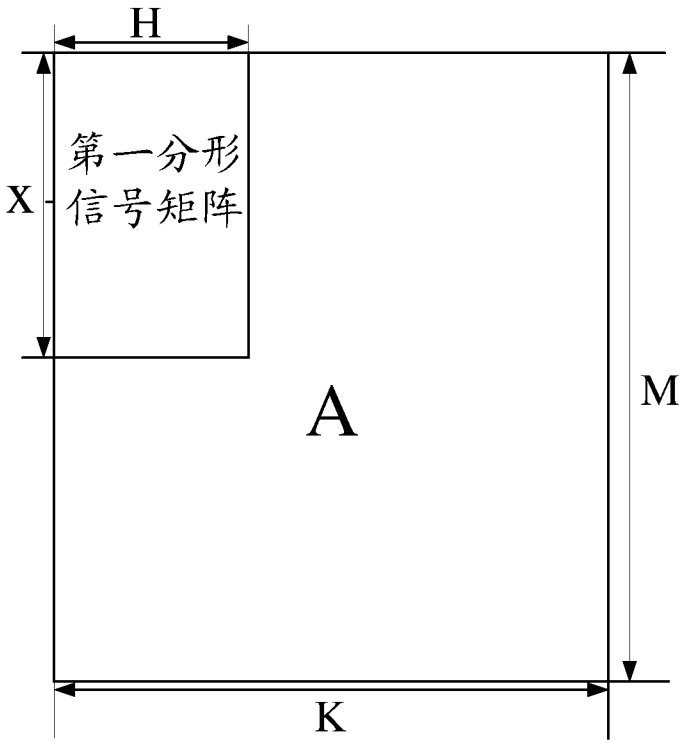
H
第一分形
信号矩阵
X
A
M
K

图 9

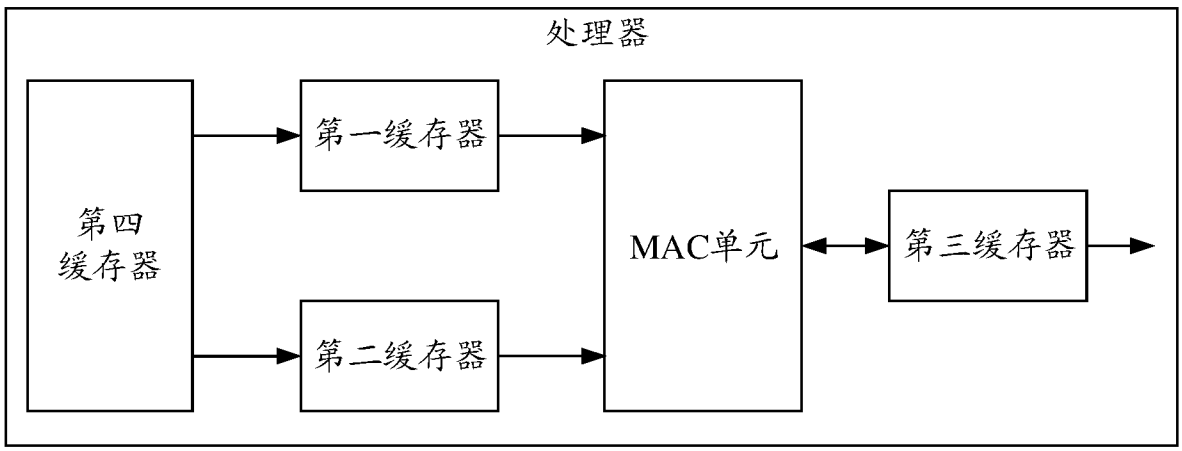
处理器
第四
缓存器
第一缓存器
第二缓存器
MAC单元
第三缓存器

图 10

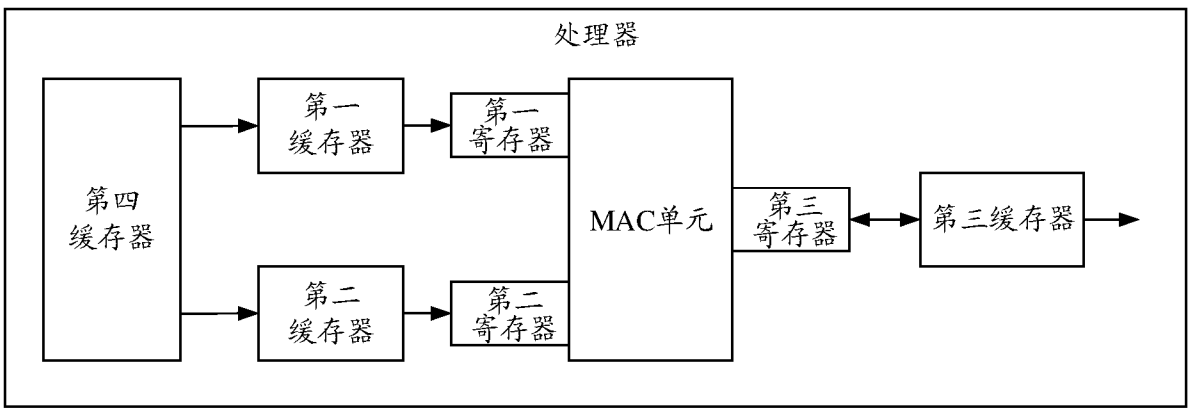
处理器
第四
缓存器
第一
缓存器
第二
缓存器
第一
寄存器
第二
寄存器
MAC单元
第三
寄存器
第三缓存器

图 11

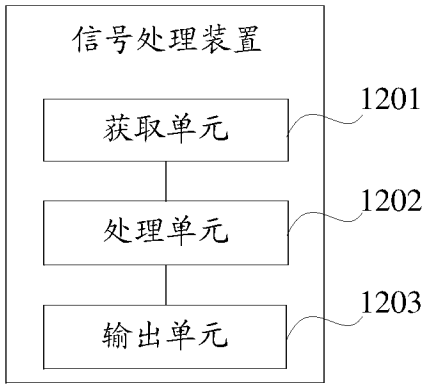

图 12

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2018/099733

A. CLASSIFICATION OF SUBJECT MATTER

G06F 17/16 (2006.01) i; G06N 3/04 (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)

G06F G06N H04L

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)

CNPAT, CNKI, WPI, EPODOC: CNN, convolutional, neural, network, 正方形, 正方型, 近似, 接近, square, fractional, 行, 列, row, column, 矩阵, matrix

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 107239824 A (BEIJING DEEPHI INTELLIGENT TECHNOLOGY CO., LTD.) 10 October 2017 (10.10.2017), description, paragraphs [0083]-[0124], and figures 6-9	1-17
A	CN 106127297 A (INSTITUTE OF AUTOMATION, CHINESE ACADEMY OF SCIENCES) 16 November 2016 (16.11.2016), entire document	1-17
A	CN 104170274 A (HUAWEI TECHNOLOGIES CO., LTD.) 26 November 2014 (26.11.2014), entire document	1-17
A	WO 2017051358 A1 (SISVEL TECHNOLOGY SRL et al.) 30 March 2017 (30.03.2017), entire document	1-17

☐ 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	Date of mailing of the international search report
13 October 2018	05 November 2018
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer YANG, Kaipeng Telephone No. (86-10) 53961727

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2018/099733

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 107239824 A	10 October 2017	US 20180157969 A1	07 June 2018
CN 106127297 A	16 November 2016	None	
CN 104170274 A	26 November 2014	WO 2015135117 A1	17 September 2015
WO 2017051358 A1	30 March 2017	EP 3354030 A1	01 August 2018
		CN 108028941 A	11 May 2018

国际检索报告

国际申请号 PCT/CN2018/099733

A. 主题的分类

G06F 17/16(2006.01)i; G06N 3/04(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)

G06F G06N H04L

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）)

CNPAT,CNKI,WPI,EPODOC:CNN, convolutional, neural, network, 正方形, 正方型, 近似, 接近, square, fractional, 行, 列, row, column, 矩阵, matrix

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
A	CN 107239824 A (北京深鉴智能科技有限公司) 2017年 10月 10日 (2017-10-10) 说明书第[0083]-[0124]段、图6-9	1-17
A	CN 106127297 A (中国科学院自动化研究所) 2016年 11月 16日 (2016-11-16) 全文	1-17
A	CN 104170274 A (华为技术有限公司) 2014年 11月 26日 (2014-11-26) 全文	1-17
A	WO 2017051358 A1 (SISVEL TECHNOLOGY SRL等) 2017年 3月 30日 (2017-03-30) 全文	1-17

☐ 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	国际检索报告邮寄日期
2018年 10月 13日	2018年 11月 5日
ISA/CN的名称和邮寄地址 中华人民共和国国家知识产权局(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	受权官员 杨凯鹏 电话号码 86-(10)-53961727

表 PCT/ISA/210 (第2页) (2015年1月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2018/099733

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	107239824	A	2017年 10月 10日	US	20180157969	A1	2018年 6月 7日
CN	106127297	A	2016年 11月 16日		无		
CN	104170274	A	2014年 11月 26日	WO	2015135117	A1	2015年 9月 17日
WO	2017051358	A1	2017年 3月 30日	EP	3354030	A1	2018年 8月 1日
				CN	108028941	A	2018年 5月 11日