



(12) 发明专利申请

(10) 申请公布号 CN 115443473 A

(43) 申请公布日 2022. 12. 06

(21) 申请号 202080099442.X

(51) Int.Cl.

(22) 申请日 2020.04.08

G06N 20/00 (2006.01)

(85) PCT国际申请进入国家阶段日

2022.09.30

(86) PCT国际申请的申请数据

PCT/JP2020/015877 2020.04.08

(87) PCT国际申请的公布数据

W02021/205581 JA 2021.10.14

(71) 申请人 富士通株式会社

地址 日本神奈川县

(72) 发明人 宫原昭一

(74) 专利代理机构 北京集佳知识产权代理有限公司

公司 11227

专利代理师 舒艳君

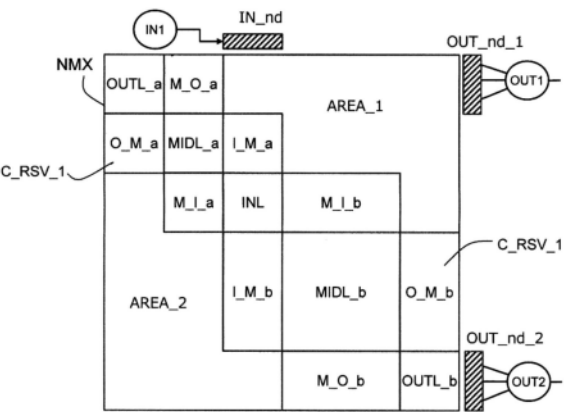
权利要求书3页 说明书11页 附图14页

(54) 发明名称

信息处理系统、信息处理装置、信息处理方法以及信息处理程序

(57) 摘要

本发明抑制储层计算模型的过度学习。储层具有共用的输入层、基于给予至输入层的输入来输出第一读出值和第二读出值的第一输出层和第二输出层，包括具有输入层和第一输出层的第一部分储层、输入层与第二输出层之间的尺寸大于输入层与第一输出层之间的尺寸的第二部分储层，学习处理具有：(1) 计算减小第一输出与教师数据的差分的第三输出权重的第一处理，第一输出是由第一输出层基于第一输入而输出的第三读出值与第一输出权重的乘积和值；(2) 计算减小第二输出与差分教师数据的差分的第四输出权重的第二处理，第二输出是由第二输出装置基于第一输入而输出的第四读出值与第二输出权重的乘积和值，差分教师数据是第三读出值与第三输出权重的乘积和值与教师数据的差分。



1. 一种信息处理系统,具有:
信息处理装置,具有存储器和访问上述存储器的处理器;以及
储层,能够与上述信息处理装置进行通信,
上述储层具有:
输入层,被给予输入;
第一输出层,基于上述输入来输出第一读出值;
第二输出层,基于上述输入来输出第二读出值;
第一部分储层,具有上述输入层和上述第一输出层;以及
第二部分储层,具有上述输入层和上述第二输出层,并且上述输入层与上述第二输出层之间的尺寸大于上述输入层与上述第一输出层之间的尺寸,
上述处理器执行上述储层的学习处理,
上述学习处理包括:
第一处理,计算减小第一输出与对应于上述第一输入的教师数据的差分的第三输出权重,其中,上述第一输出是由上述第一输出层基于第一输入而输出的第三读出值与第一输出权重的乘积和值;以及
第二处理,计算减小第二输出与差分教师数据的差分的第四输出权重,其中,上述第二输出是由上述第二输出层基于上述第一输入而输出的第四读出值与第二输出权重的乘积和值,上述差分教师数据是上述第三读出值与上述第三输出权重的乘积和值与上述教师数据的差分。
2. 根据权利要求1所述的信息处理系统,其中,
上述处理器执行被执行了上述学习处理的上述储层的推理处理,
上述推理处理具有计算第三输出与第四输出的和的处理,其中,上述第三输出是由上述第一输出层基于第二输入而输出的第五读出值与上述第三输出权重的乘积和值,上述第四输出是由上述第二输出层基于上述第二输入而输出的第六读出值与上述第四输出权重的乘积和值。
3. 根据权利要求1或2所述的信息处理系统,其中,
上述储层具有多个节点,
上述多个节点具有在任意的节点间具有权重的耦合,
上述多个节点中的任意的节点具有自耦合,上述自耦合具有权重,
上述第一部分储层包含上述多个节点的一部分,
上述第二部分储层包含上述多个节点的另一部分,上述多个节点的上述另一部分的节点数多于上述多个节点的上述一部分的节点数。
4. 根据权利要求3所述的信息处理系统,其中,
上述储层是交叉开关,上述交叉开关具有对应于上述多个节点的输入的多个列线、对应于上述多个节点的输出的多个行线、以及在上述多个列线与上述多个行线的任意的交叉位置具有上述自耦合或者上述耦合。
5. 根据权利要求1或2所述的信息处理系统,其中,
上述储层是基于输入来输出非线性特性的材料储层,
上述输入层设置于上述材料储层的任意的输入位置,

上述第一输出层设置于上述材料储层的第一输出位置，

上述第二输出层设置于上述材料储层的第二输出位置，其中，上述输入位置与上述第二输出位置之间的距离比上述输入位置与上述第一输出位置之间的距离长。

6. 一种信息处理装置，具有存储器和访问上述存储器的处理器，并且能够与储层进行通信，其中，

上述储层具有：

输入层，被给予输入；

第一输出层，基于上述输入来输出第一读出值；

第二输出层，基于上述输入来输出第二读出值；

第一部分储层，具有上述输入层和上述第一输出层；以及

第二部分储层，具有上述输入层和上述第二输出层，并且上述输入层与上述第二输出层之间的尺寸大于上述输入层与上述第一输出层之间的尺寸，

上述处理器执行上述储层的学习处理，

上述学习处理包括：

第一处理，计算减小第一输出与对应于上述第一输入的教师数据的差分的第三输出权重，其中，上述第一输出是由上述第一输出层基于第一输入而输出的第三读出值与第一输出权重的乘积和值；以及

第二处理，计算减小第二输出与差分教师数据的差分的第四输出权重，其中，上述第二输出是由上述第二输出层基于上述第一输入而输出的第四读出值与第二输出权重的乘积和值，上述差分教师数据是上述第三读出值与上述第三输出权重的乘积和值与上述教师数据的差分。

7. 一种信息处理方法，是信息处理装置的信息处理方法，上述信息处理装置具有存储器和访问上述存储器的处理器并且能够与储层进行通信，其中，

上述储层具有：

输入层，被给予输入；

第一输出层，基于上述输入来输出第一读出值；

第二输出层，基于上述输入来输出第二读出值；

第一部分储层，具有上述输入层和上述第一输出层；以及

第二部分储层，具有上述输入层和上述第二输出层，并且上述输入层与上述第二输出层之间的尺寸大于上述输入层与上述第一输出层之间的尺寸，

上述处理器执行上述储层的学习处理，

上述学习处理包括：

第一处理，计算减小第一输出与对应于上述第一输入的教师数据的差分的第三输出权重，其中，上述第一输出是由上述第一输出层基于第一输入而输出的第三读出值与第一输出权重的乘积和值；以及

第二处理，计算减小第二输出与差分教师数据的差分的第四输出权重，其中，上述第二输出是由上述第二输出层基于上述第一输入而输出的第四读出值与第二输出权重的乘积和值，上述差分教师数据是上述第三读出值与上述第三输出权重的乘积和值与上述教师数据的差分。

8. 一种信息处理程序,使具有存储器和访问上述存储器的处理器并且能够与储层进行通信的信息处理装置的上述处理器执行上述储层的学习处理,其中,

上述储层具有:

输入层,被给予输入;

第一输出层,基于上述输入来输出第一读出值;

第二输出层,基于上述输入来输出第二读出值;

第一部分储层,具有上述输入层和上述第一输出层;以及

第二部分储层,具有上述输入层和上述第二输出层,并且上述输入层与上述第二输出层之间的尺寸大于上述输入层与上述第一输出层之间的尺寸,

上述学习处理包括:

第一处理,计算减小第一输出与对应于上述第一输入的教师数据的差分的第三输出权重,其中,上述第一输出是由上述第一输出层基于第一输入而输出的第三读出值与第一输出权重的乘积和值;以及

第二处理,计算减小第二输出与差分教师数据的差分的第四输出权重,其中,上述第二输出是由上述第二输出层基于上述第一输入而输出的第四读出值与第二输出权重的乘积和值,上述差分教师数据是上述第三读出值与上述第三输出权重的乘积和值与上述教师数据的差分。

信息处理系统、信息处理装置、信息处理方法以及信息处理程序

技术领域

[0001] 本发明涉及信息处理系统、信息处理装置、信息处理方法以及信息处理程序。

背景技术

[0002] 作为机器学习模型之一,神经网络非常普及。特别是,基于深度学习的深度神经网络(以下,称为DNN(Deep Neural Network)。)在图像处理等中报告高性能。这些DNN是分层神经网络。另外,作为DNN之一的递归神经网络(以下,称为RNN(Recurrent Neural Network)。)具有将过去时刻的神经元的状态反馈输入到自己或者其他的神元元的结构。反馈输入是取决于过去的输入历史的过去时刻的状态,适用于基于时间序列输入的前后关系来进行推理的模型。

[0003] 另一方面,众所周知,在RNN中,由于运算处理复杂化,而在DNN的学习工序中应用为了优化参数而进行的反向误差传播法是不现实的。

[0004] 在储层计算(Reservoir Computing)中,利用能够通过RNN所具有的递归反馈将过去的状态反映到当前的状态中的性质,在学习工序中不更新构成RNN的储层的权重,而更新由储层的输出节点构成的全耦合层的权重(输出权重)。由此,能够大幅削减学习工序的运算的负荷。

[0005] 在储层计算中,由于储层内的神经元具有非线性特性,所以提出了代替储层而利用具有非线性的特性的物理结构物。在本发明中,能够应用于通过处理器执行计算机程序来实现储层的情况和通过物理储层来实现储层的情况中的任意一个情况。

[0006] 专利文献1:日本特表2004-511866号公报

[0007] 专利文献2:国际公开第2018/212201号

[0008] 专利文献3:日本特开2018-180701号公报

[0009] 非专利文献1:http://eeip.t.u-tokyo.ac.jp/gtanaka/siryo/tanaka201902_ieice.pdf电子信息通信学会学报Vol.102, No.2, 2019《储层计算的概念和最近的动向》田中刚平

[0010] 非专利文献2:<https://www.nature.com/articles/s41467-017-02337-y.pdf>

[0011] 越增大储层的尺寸,储层具有越高维函数的特性。但是,若以具有高维函数特性的储层进行学习,则会成为过度学习,模型成为过度拟合状态。

发明内容

[0012] 因此,本实施方式的第一方面的目的在于提供一种与能够抑制由过度学习引起的过度拟合状态的储层计算相关的信息处理系统、信息处理装置、信息处理方法以及信息处理程序。

[0013] 本实施方式的第一方面是一种信息处理系统,具有:

[0014] 信息处理装置,具有存储器和访问上述存储器的处理器;以及

- [0015] 储层,能够与上述信息处理装置进行通信,
- [0016] 上述储层具有:
- [0017] 输入层,被给予输入;
- [0018] 第一输出层,基于上述输入来输出第一读出值;
- [0019] 第二输出层,基于上述输入来输出第二读出值;
- [0020] 第一部分储层,具有上述输入层和上述第一输出层;
- [0021] 第二部分储层,具有上述输入层和上述第二输出层,上述输入层与上述第二输出层之间的尺寸大于上述输入层与上述第一输出层之间的尺寸,
- [0022] 上述处理器执行上述储层的学习处理,
- [0023] 上述学习处理包括:
- [0024] 第一处理,计算减小第一输出与对应于上述第一输入的教师数据的差分的第三输出权重,其中,上述第一输出是由上述第一输出层基于第一输入而输出的第三读出值与第一输出权重的乘积和值;以及
- [0025] 第二处理,计算减小第二输出与差分教师数据的差分的第四输出权重,其中,上述第二输出是由上述第二输出层基于上述第一输入而输出的第四读出值与第二输出权重的乘积和值,上述差分教师数据是上述第三读出值与上述第三输出权重的乘积和值与上述教师数据的差分。
- [0026] 根据第一侧面,能够抑制由过度学习引起的过度拟合状态。

附图说明

- [0027] 图1是实现本实施方式中的储层计算模型的信息处理系统的示意图。
- [0028] 图2是表示储层计算模型的网络的一个例子的图。
- [0029] 图3是表示作为储层计算模型而使用了作为物理储层之一的交叉开关储层的信息处理系统的图。
- [0030] 图4是表示交叉开关储层(Crossbar Reservoir)的一个例子和作为读出部的全耦合层的一个例子的图。
- [0031] 图5是表示对应于图2的例子的交叉开关储层的相邻矩阵的一个例子的图。
- [0032] 图6是分类表示交叉开关储层的相邻矩阵所表示的交叉点的耦合的图。
- [0033] 图7是表示图6的相邻矩阵中示出的九个区域的关系的图。
- [0034] 图8是表示第一实施方式中的储层计算模型的图。
- [0035] 图9是表示第二实施方式中的储层计算模型的图。
- [0036] 图10是表示第一实施方式中的储层计算模型的相邻矩阵的区域的图。
- [0037] 图11是表示图10的相邻矩阵中示出的九个区域的关系的图。
- [0038] 图12是表示相邻矩阵的具体例的图。
- [0039] 图13是表示第一实施方式和第二实施方式的储层计算模型的学习工序的流程图。
- [0040] 图14是表示第一实施方式和第二实施方式的储层计算模型的推理工序的处理的流程图。
- [0041] 图15是进行学习工序的输出权重的更新的FORCE学习的流程图。
- [0042] 图16是表示第三实施方式中的储层计算模型的图。

具体实施方式

[0043] 图1是实现本实施方式中的储层计算模型的信息处理系统的示意图。信息处理装置10例如是PC、HPC等计算机。信息处理装置10具有处理器12、处理器所访问的主存储器14、以及与外部的接口装置16。另外，信息处理装置10在大容量的辅助存储装置20~26中储存储层模型的学习程序20、推理程序22、用于学习的训练数据24以及具有储层模型的构成数据和输出权重等参数的储层模型数据26。处理器12通过执行从辅助存储装置读出并在主存储器14中展开的学习程序20和推理程序22，来执行学习工序的运算以及推理工序的运算。另外，处理器12执行储层模型数据26所包含的运算程序，对时间序列的输入运算储层的状态的迁移。

[0044] 另外，在图1的信息处理系统中，设置物理储层18。由此，处理器12也能够经由接口16参照物理储层18的物理上的变化，来代替参照储层模型数据26运算储层的状态的迁移。在该情况下，处理器12在学习工序中不进行储层的状态迁移的运算，代替其而获取物理储层18的输出，通过运算来进行减少教师数据与输出的误差的输出权重的更新。

[0045] 从多个终端装置30、32经由网络NW访问信息处理装置10。响应于来自终端装置的学习处理的请求、推理处理的请求，处理器12执行学习模型、推理模型，执行所请求的学习处理以及推理处理的运算。

[0046] 图2是表示储层计算模型的网络的一个例子的图。以下，将储层的神经元称为节点。储层RSV具有节点N1~N5和节点间耦合或者节点的自耦合。耦合被称为边缘。另外，耦合具有用箭头表示的方向和与各方向对应的权重。在图中，储层RSV内的箭头是节点间耦合或者节点的自耦合。例如，节点N1、N3具有自耦合。在节点N1与节点N2之间具有双向的耦合。在节点N2与N3之间，具有从节点N2向节点N3的朝向的一个方向的耦合。

[0047] 在学习之前，随机或者根据规定的规则来决定储层RSV的节点间耦合或者自耦合的有无和各耦合的权重。所谓的规定的规则，例如，以节点间的耦合概率、耦合的权重的方差为参数。

[0048] 若在时刻 t 从输入端子IN1向节点N1给予输入，则基于该输入和时刻 t 之前的时刻的节点的状态推移节点N1~N5的状态。然后，输出节点N4和N5的状态，将基于节点N4和N5的输出(读出数)和输出权重 w_4 、 w_5 运算出的活性输入至输出端子OUT。节点N4、N5、输出权重 w_4 、 w_5 的耦合以及输出端子OUT构成全耦合层FCL。像这样，储层计算模型具有储层RSV和针对储层的输出构成全耦合层FCL的读出部。

[0049] 在图2的储层RSV中，节点N1对应于被给予输入的输入层，节点N4、N5对应于输出层，其状态作为输出被读出。而且，输入层的节点N1与输出层的节点N4、N5之间的节点N2、N3对应于中间层。

[0050] 在储层计算模型的学习工序中，在固定了储层内的节点间耦合、节点自耦合的状态下，用梯度法更新作为读出部的全耦合层FCL的输出权重。即，通过梯度法更新输出权重 w_4 、 w_5 ，以使教师数据与输出节点OUT的误差最小化。储层内的耦合的权重不会通过学习工序中的梯度法来更新。因此，在储层计算模型的学习工序中，通过更新而调整的输出权重少，学习处理被简化，能够以少的训练数据完成学习工序。

[0051] 储层计算模型也能够以物理上构成储层RSV的物理储层来实现，也能够通过基于计算机程序的运算来计算储层的状态的迁移。以下，作为物理储层的一个例子，以利用非线性

性元件的交叉开关储层的例子进行说明。

[0052] 图3是表示作为储层计算模型而使用了作为物理储层之一的交叉开关储层的信息处理系统的图。图3的储层计算模型的信息处理系统具有包含处理器的信息处理装置10和作为物理储层的交叉开关储层C_RSV。交叉开关储层C_RSV具有对应于节点的输入的列线 and 对应于节点的输出的行线,虽然未图示,但在列线与行线之间设置具有非线性特性的忆阻器(Memristor)、CR电路等模拟电路等物理电路。

[0053] 另一方面,若对应于图2的储层计算模型,则从处理器向交叉开关储层C_RSV内的输入层的节点N1的输入供给输入IN1,并将输出层的节点N4、N5的输出(读出值)输出至处理器。处理器通过对作为读出部的全耦合层执行计算机程序,来执行学习工序以及推理工序的运算。

[0054] 因此,在储层计算模型中,执行学习程序的处理器将训练数据的输入IN1给予至交叉开关储层C_RSV的输入层的节点N1,响应于此,交叉开关储层C_RSV的状态从过去的状态迁移至当前的状态。而且,处理器输入表示交叉开关储层C_RSV的输出层的节点N4、N5的状态的输出,并执行学习程序,对作为读出部的全耦合层FCL执行基于梯度法的输出权重的更新。通过反复该学习,输出权重被优化,生成储层计算模型,该储层计算模型生成与教师数据相同或者接近的输出。

[0055] 图4是表示交叉开关储层的一个例子和作为读出部的全耦合层的一个例子的图。交叉开关储层C_RSV针对图2所示的五个节点N1~N5,具有对应于节点的输入的五根列线(纵线)N1~N5、对应于节点的输出的五根行线(横线)N1~N5、以及配置于列线与行线的交叉点的25个耦合电路f11~f55。耦合电路f11~f55例如是忆阻器。

[0056] 在图4的交叉开关储层C_RSV中,在所有的五根列线和五根行线之间设置有耦合电路f11~f55,但该耦合电路也可以仅设置于任意的列线和行线之间。例如,在图2的储层RSV中,节点间耦合和节点自耦合仅设置在一部分的节点间以及节点。相同的节点的列线和行线之间的耦合电路对应于节点自耦合。另外,不同的节点的列线和行线之间的耦合电路对应于从列线的节点到行船的节点的方向的单向节点间耦合。而且,耦合电路对应于图2的储层内的耦合的权重。

[0057] 在图4的例子中,向对应于节点的输入的五根列线,输入通过DAC(数字-模拟转换器)对从输入部IN_1的输入端子IN1供给的数字信号转换后的模拟信号X1~X5。输入部IN_2、IN_3也同样。其中,在向储层内的一部分节点给予输入的情况下,对输入端子IN1的数字信号进行DAC转换后的模拟信号被给予至一部分节点的输入。

[0058] 另一方面,对应于节点的输出的五根列线的模拟信号y1~y5被ADC(模拟-数字转换器)转换,并输出至全耦合层FCL_1,全耦合层FCL_1的乘积和值的数字信号作为活性输出至输出节点OUT1。乘积和值如下。

[0059] $OUT1 = y1 * w1 + y2 * w2 + y3 * w3 + y4 * w4 + y5 * w5$

[0060] 其他全耦合层FCL_2、FCL_3也同样。其中,在将输出从储层内的一部分节点输出的情况下,五根行线的一部分节点的输出输出至全耦合层。

[0061] 在图2的储层的情况下,25个耦合电路中的f11、f12、f21、f24、f32、f33、f35、f45、f51、f54具有有限的值,剩余的无限大,换句话说非耦合。在这样的交叉开关储层C_RSV中,若在时间序列的输入中的时刻t1将输入x1给予至输入层的节点N1,则全节点N1~N5的状态

迁移至以下的式子所示的输出 $y_1 \sim y_5$ 。

$$[0062] \quad y_{1_t1} = \{f_{11}(x_1) + f_{12}(x_2)\} + \alpha * y_{1_t0}$$

$$[0063] \quad y_{2_t1} = \{f_{21}(x_1) + f_{24}(x_4)\} + \alpha * y_{2_t0}$$

$$[0064] \quad y_{3_t1} = \{f_{32}(x_2) + f_{33}(x_3) + f_{35}(x_5)\} + \alpha * y_{3_t0}$$

$$[0065] \quad y_{4_t1} = \{f_{45}(x_5)\} + \alpha * y_{4_t0}$$

$$[0066] \quad y_{5_t1} = \{f_{51}(x_1) + f_{54}(x_4)\} + \alpha * y_{5_t0}$$

[0067] 在这里, $f_{11}(x) \sim f_{55}(x)$ 是对节点的输入 x 求出输出至节点的输出 y 的值的函数。另外, 时刻 t_0 是输入时的时刻 t_1 的前一时刻, y_{1_t1} 是时刻 t_1 的状态, y_{1_t0} 是时刻 t_0 的状态。而且, α 对应于节点的状态的衰减率。在耦合电路为忆阻器的情况下, 函数 $f_{11}(x) \sim f_{55}(x)$ 的特性基于某一时刻的输入而非线性地变化。

[0068] 如上所述, 在作为物理储层的交叉开关储层中, 各节点保持基于过去的输入的状态, 并且针对时间序列的输入逐个迁移至新的状态。像这样, 输入被交叉开关储层 C_RSV 非线性转换, 转换后的状态被映射到高维的特征空间。另一方面, 交叉开关储层 C_RSV 的输出层的节点的状态在全耦合层 FCL 中线性耦合, 输出输出 OUT 。

[0069] 在这里, 对用矩阵表示交叉开关储层 C_RSV 的节点间耦合和节点自耦合的权重的相邻矩阵进行说明。通过利用相邻矩阵, 容易理解具有交叉开关储层的输入层、中间层、输出层的节点、以及节点间耦合和节点自耦合的交叉开关储层的结构。

[0070] 图5是表示对应于图2的例子的交叉开关储层的相邻矩阵的一个例子的图。相邻矩阵在对应于节点 $N1 \sim N5$ 的输入的列和对应于节点 $N1 \sim N5$ 的输出的行的交叉位置示出节点间耦合的权重和节点自耦合的权重。例如, 在节点 $N1$ 的列与节点 $N1$ 的行的交叉位置, 示出节点 $N1$ 的自耦合的权重。另一方面, 在节点 $N1$ 的列和节点 $N2$ 的行的交叉位置, 示出从节点 $N1$ 向节点 $N2$ 的节点间耦合的权重。相反, 在节点 $N2$ 的列和节点 $N1$ 的行的交叉位置, 示出从节点 $N2$ 向节点 $N1$ 的节点间耦合的权重。因此, 在图5中, 在通过上述的式子表示的函数 f_{11} 、 f_{12} 、 f_{21} 、 f_{24} 、 f_{32} 、 f_{33} 、 f_{35} 、 f_{45} 、 f_{51} 、 f_{54} 的位置示出各个权重。

[0071] 图6是分类示出交叉开关储层的相邻矩阵所表示的交叉点的耦合的图。在图2的储层的例子中, 节点 $N1$ 为输入层 INL , 节点 $N4$ 、 $N5$ 是输出层 $OUTL$, 输入层与输出层之间的节点 $N2$ 、 $N3$ 为中间层 $MIDL$ 。因此, 在相邻矩阵中, 节点 $N1$ 的列与行的交叉位置表示为输入层 INL , 节点 $N4$ 、 $N5$ 的列与行的交叉位置表示为输出层 $OUTL$, 节点 $N2$ 、 $N3$ 的列与行的交叉位置表示为中间层 $MIDL$ 。输入层 INL 、中间层 $MIDL$ 、输出层 $OUTL$ 为构成各层的节点的数量为1的正方矩阵, 对应于节点的自耦合区域。

[0072] 进一步,

[0073] 将节点 $N1$ 的列与 $N2$ 、 $N3$ 的行的交叉位置表示为从输入层向中间层的节点间耦合区域 I_M

[0074] 将节点 $N2$ 、 $N3$ 的列与 $N4$ 、 $N5$ 的行的交叉位置表示为从中间层向输出层的节点间耦合区域 M_O

[0075] 将节点 $N2$ 、 $N3$ 的列与 $N1$ 的行的交叉位置表示为从中间层向输入层的节点间耦合区域 M_I

[0076] 将节点 $N4$ 、 $N5$ 的列与 $N2$ 、 $N3$ 的行的交叉位置表示为从输出层向中间层的节点间耦合区域 O_M

[0077] 进一步,

[0078] 将节点N1的列与N4、N5的行的交叉位置表示为从输入层向输出层的节点间耦合区域I_0

[0079] 将节点N4、N5的列与节点N1的行的交叉位置表示为从输出层向输入层的节点间耦合区域O_I。

[0080] 图7是表示图6的相邻矩阵中示出的九个区域的关系的图。在图7中,示出输入层INL、中间层MIDL、输出层OUTL和这些层之间的节点间耦合的区域I_M、M_0、I_0、O_M、M_I、I_0的关系。

[0081] 因此,参照图4、图5、图6、图7的交叉开关储层的结构以及相邻矩阵,针对图2的储层RSV,对针对时间序列的输入的节点的状态的迁移进行说明。其中,在以下的说明中,从图2的储层RSV的节点N1向节点N5的耦合为非耦合。因此,图7的节点间耦合的区域O_I、I_0为非耦合。

[0082] 若在时刻t1向节点N1给予输入,则

[0083] (1) 输入层INL的节点N1推移为如下变化之和,即:

[0084] 时刻t1的输入x1的变化;

[0085] 基于时刻t0的输入层INL的节点状态x1和自耦合的N1的变化;以及

[0086] 基于时刻t0的中间层MIDL的节点状态x2、x3和N1的区域M_I的节点间耦合M_I (N2、N3到N1) 的N1的变化。

[0087] (2) 中间层MIDL的节点N2、N3推移为如下变化之和,即:

[0088] 基于时刻t1的输入x1和区域I_M的节点间耦合 (N1到N2、N3) 的N2、N3的变化;

[0089] 基于时刻t0的中间层MIDL的节点状态x2、x3和中间层的自耦合的N2、N3的变化;以及

[0090] 基于时刻t0的输出层OUTL的节点状态x4、x5和区域O_M的节点间耦合 (N4、N5到N2、N3) 的N2、N3的变化。

[0091] (3) 输出层OUTL的节点N4、N5推移为如下变化之和,即:基于时刻t0的中间层MIDL的节点状态x2、x3和区域M_0的节点间耦合 (N2、N3到N4、N5) 的N4、N5的变化;以及

[0092] 基于时刻t0的输出层OUTL的节点状态x4、x5和输出层的自耦合的N4、N5的变化。

[0093] 如上所述,交叉开关储层的输入层INL、中间层MIDL、输出层OUTL成为相邻矩阵内的正方形的区域,分别具有自耦合。另外,这些层间的节点间耦合位于相邻矩阵内的区域I_M、M_0、O_M、M_I、I_0、O_I。

[0094] [本实施方式中的储层计算模型]

[0095] 越增大储层的尺寸,储层具有越高维的函数特性。但是,若通过训练数据对具有高维的函数特性的储层进行学习,则成为过度学习,模型成为对训练数据过度拟合状态。

[0096] 图8是表示第一实施方式中的储层计算模型的图。图8的模型是具有作为物理储层的一个例子的交叉开关储层C_RSV和信息处理装置10的信息处理系统。

[0097] 交叉开关储层C_RSV具有第一部分储层C_RSV_1和第二部分储层C_RSV_2,第二部分储层C_RSV_2与第一部分储层C_RSV_1相比,输入层与输出层之间的尺寸较大,且具有较高维函数特性。第一部分储层C_RSV_1和第二部分储层C_RSV_2具有共用的输入节点IN_nd,具有各自的输出节点OUT_nd_1、OUT_nd_2,且分别具有共用的输入节点与各个输出节点间

的中间层的节点。共用的输入节点IN_nd构成共用的输入层,各个输出节点OUT_nd_1、OUT_nd_2构成各个输出层。对于交叉开关储层C_RSV的结构,在后面示出相邻矩阵进行说明。储层的尺寸例如是中间层的节点数。

[0098] 来自各个输出层的输出节点OUT_nd_1、OUT_nd_2的读出值被输出至图1所记载的信息处理装置10,处理器12运算读出值与输出权重的线性耦合的乘积和值,生成两个输出OUT1、OUT2,计算两个输出的和作为交叉开关储层C_RSV的输出。另外,处理器12在后述的步骤中进行基于学习工序中的训练数据的输出权重的优化。

[0099] 图9是表示第二实施方式中的储层计算模型的图。图9的模型是具有利用作为物理储层的一个例子的材料的非线性特性的材料储层M_RSV和信息处理装置10的信息处理系统。材料储层的例子例如是具有电非线性的材料的块、具有光学的非线性的材料的块等。

[0100] 材料储层M_RSV是材料的块。若在材料的任意位置设定输入输入IN1的输入层的输入节点IN_nd和取出读出值的输出层的输出节点OUT_nd,则输入层和输出层的节点间的材料成为中间层的虚拟的节点。与第一实施方式同样,对于第二实施方式,材料储层M_RSV也具有第一部分储层M_RSV_1和第二部分储层M_RSV_2,第二部分储层M_RSV_2与第一部分储层M_RSV_1相比,输入与输出之间的尺寸较大且具有较高维函数特性。第一部分储层M_RSV_1和第二部分储层M_RSV_2具有部分重合的区域,在该区域设置共用的输入层的输入节点IN_nd。另外,在第一部分储层M_RSV_1和第二部分储层M_RSV_2设置各个输出层的节点OUT_nd_1、OUT_nd_2。材料储层M_RSV的尺寸例如是输入节点IN_nd与输出节点OUT_nd间的距离。

[0101] 与第一实施方式同样,来自各个输出层的输出节点OUT_nd_1、OUT_nd_2的读出值被输出至信息处理装置10,处理器运算读出值和输出权重的线性耦合的乘积和值,生成两个输出OUT1、OUT2,计算两个输出的和作为材料储层M_RSV的输出。另外,处理器在后述的步骤中进行基于学习工序中的训练数据的输出权重的优化。

[0102] 图10是表示第一实施方式中的储层计算模型的相邻矩阵的区域的图。如图8所示,作为物理储层的一个例子的交叉开关储层C_RSV具有第一部分储层C_RSV_1和第二部分储层C_RSV_2,第二部分储层C_RSV_2的尺寸大于第一部分储层C_RSV_1的尺寸。另外,图11是表示图10的相邻矩阵所示的九个区域的关系的图。

[0103] 第一部分储层C_RSV_1具备:具有输入节点IN_nd的输入层INL、具有第一输出节点OUT_nd_1的第一输出层OUTL_a以及具有输入层INL与第一输出层OUTL_a之间的第一中间节点组的第一中间层MIDL_a。而且,与图6、图7同样,第一部分储层C_RSV_1具有:包含从输入层INL的节点朝向第一中间层MIDL_a的节点的节点间耦合的区域I_M_a、以及包含从第一中间层MIDL_a的节点朝向第一输出层OUTL_a的节点的节点间耦合的区域M_O_a。同样地,具有:包含从输出层OUTL的节点朝向第一中间层MIDL_a的节点的节点间耦合的区域O_M_a和包含从第一中间层MIDL_a的节点朝向输入层INL的节点的节点间耦合的区域M_I_a。此外,图10的区域AREA_1与AREA_2内的区域I_O_a和O_I_a的节点间耦合均是非耦合。这是因为,在物理储层的情况下,一般地输入节点组与输出节点组之间的距离长,所以这些节点间耦合为非耦合。

[0104] 第一输出层的第一输出节点OUT_nd_1的读出值和输出权重的乘积和值为第一输出OUT1。该全耦合层的输出权重在学习工序中被更新为最佳值。

[0105] 第二部分储层C_RSV_2具备:具有输入节点IN_nd的输入层INL、具有第二输出节点

OUT_nd_2的第二输出层OUTL_b以及具有输入层INL和第二输出层OUTL_b之间的第二中间节点组的第二中间层MIDL_b。而且,与图6、图7同样,第二部分储层C_RSV_2具有:包含从输入层INL的节点朝向第二中间层MIDL_b的节点的节点间耦合的区域I_M_b和包含从第二中间层MIDL_b的节点朝向第二输出层OUTL_b的节点的节点间耦合的区域M_O_b。同样地,具有:包含从输出层OUTL的节点朝向第二中间层MIDL_b的节点的节点间耦合的区域O_M_b和包含从第二中间层MIDL_b的节点朝向输入层INL的节点的节点间耦合的区域M_I_b。此外,图10的区域AREA_1与AREA_2内的区域I_O_b和O_I_b的节点间耦合均为非耦合。理由如上所述。

[0106] 第二输出层的第二输出节点OUT_nd_2的读出值和输出权重的乘积和值为第二输出OUT2。该全耦合层的输出权重在学习工序中被更新为最佳值。

[0107] 图12是表示相邻矩阵的具体例的图。该相邻矩阵NMX是 20×20 行的例子,因此是对应于20个节点的交叉开关储层的相邻矩阵。根据该相邻矩阵NMX,具有20个节点的交叉开关储层被分割为节点N7、N8的输入节点组、第一输出节点组N1、N2、第二输出节点组N19、N20、第一中间节点组N3~N6以及第二中间节点组N9~N18。

[0108] 而且,第一部分储层具有输入节点组N7、N8 (IN_nd)、第一输出节点组N1、N2以及第一中间节点组N3~N6。第一输出节点组N1、N2 (OUT_nd_1) 的读出值与输出权重的乘积和值为第一输出OUT1。同样地,第二部分储层具有输入节点组N7、N8 (IN_nd)、第二输出节点组N19、N20 (OUT_nd_2) 以及第二中间节点组N9~N18。第二输出节点组N19、N21的读出值与输出权重的乘积和值为第二输出OUT2。

[0109] 如图12所示,在除了区域AREA_1和AREA_2以外的区域,在任意交叉位置分配有权重的值。除此以外的交叉位置的权重=0是指非耦合。根据在储层计算模型中实现的目的,适当地初始设计将交叉储层的哪个交叉位置设为怎样的权重。

[0110] 图10、图11、图12的相邻矩阵也可以应用于材料储层。材料储层由于针对被给予至输入节点的输入在输出节点生成非线性特性的读出值,所以通过设定与该非线性特性对应的权重,能够涉及具有作为目的的非线性特性的储层的相邻矩阵。

[0111] 图13是表示第一实施方式和第二实施方式的储层计算模型的学习工序的流程图。处理器执行学习程序并执行以下的处理。首先,处理器对储层的输出权重进行初始化(S1)。接下来,处理器将训练数据的输入值输入至储层的输入节点(S2)。若将输入值输入至输入节点,则在物理储层中,节点数少或者小尺寸的第一部分储层基于输入值和过去的节点的状态以及节点间耦合和节点自耦合,生成各节点的状态(S3)。同样地,在物理储层中,输入层与输出层之间的节点数多于第一部分储层或者输入层与输出层之间的尺寸更大的第二部分储层基于输入值和过去的节点的状态以及节点间耦合和节点自耦合,生成各节点的状态(S4)。换句话说,响应于相同的输入值的输入,第一部分储层和第二部分储层的各自的状态迁移,第一部分储层的第一输出节点和第二部分储层的第二输出节点的状态迁移。

[0112] 对于上述的第一部分储层和第二部分储层的状态迁移,也可以处理器进行第一部分储层和第二部分储层的状态迁移的运算,并进行第一部分储层的第一输出节点和第二部分储层的第二输出节点的状态。

[0113] 接下来,处理器首先执行将第一部分储层的输出权重调整为最佳值的第一参数调整(S5)。在该第一参数调整中,对于将通过基于第一部分储层的第一输出节点的读出值的第一输出权重的线性耦合求出的乘积和值作为第一输出的第一全耦合层,处理器更新第一

输出权重,以使第一输出与教师数据的差分减少。输出权重的更新能够通过梯度法来进行,或者使用线性回归等简单的算法来学习输出权重。

[0114] 并且,处理器将第一部分储层的第一输出权重设定为更新后的值,运算针对工序S2的输入值的第一输出,并运算该第一输出与教师数据的差分作为差分教师数据(S6)。

[0115] 接下来,处理器执行将第二部分储层的输出权重调整为最佳值的第二参数调整(S7)。在该第二十一参数调整中,对于将基于在工序S4中获取或者运算出的第二部分储层的第二输出节点的读出值和第二输出权重的乘积和值作为第二输出的第二全耦合层,处理器更新第二输出权重,以使第二输出和差分教师数据的差分变小。该输出权重的更新除了使用差分教师数据以外与第一部分储层相同。

[0116] 处理器对所有教师数据反复执行从输入上述的训练数据的输入值的工序S2到第二参数调整的工序S7的处理(S8)。或者,对多个教师数据反复执行,直到模型的输出精度达到目标精度。模型的输出精度通过第一部分储层的读出部的第一输出和第二部分储层的读出部的第二输出的和与教师数据的差分除以上述和等来求出。

[0117] 在由第一部分储层进行的学习工序中,由于第一部分储层的尺寸小且节点数少,所以生成相对低维的函数模型。其结果是,通过学习而被优化的第一部分储层和读出部生成低频成分的第一输出。由于是低维的函数,所以在学习工序中能够避免过度学习。另一方面,由于第二部分储层的尺寸大于第一部分储层的尺寸且节点数多,所以生成比第一部分储层高维的函数模型。其结果是,通过学习而被优化的第二部分储层和读出部生成高频成分的第二输出。而且,第一输出和第二输出的和成为储层计算模型的输出。因此,储层计算模型能够生成对低频成分的第一输出加上高频成分的第二输出所得的输出。

[0118] 图14是表示第一实施方式和第二实施方式的储层计算模型的推理工序的处理的流程图。处理器将输入值输入至储层(S10)。在物理储层中,节点数少的小尺寸的第一部分储层和节点数多的大尺寸的第二部分储层根据基于输入值和过去的输入历史的各个储层的状态和节点间耦合以及节点的自耦合来迁移其状态(S11、S12)。进一步,处理器基于第一部分储层的输出节点的值和输出权重来运算读出部的第一输出(S13),并基于第二部分储层的输出节点的值和输出权重来运算读出部的第二输出(S14)。此时的输出权重是在学习工序中调整后的最佳的输出权重。而且,处理器运算第一输出与第二输出的和,计算推理值(S15)。处理器反复上述的推理工序S10~S15,直到推理结束(S16)。

[0119] 与学习工序同样地,在推理工序中,处理器通过第一部分储层和其读出部生成推理输出的低频成分,并通过第二部分储层和其读出部生成推理输出的高频成分。根据本实施方式,由于在学习工序中未成为过度学习状态,所以能够生成泛化性高的储层计算模型。

[0120] 图15是进行学习工序的输出权重的更新的FORCE学习的流程图。在FORCE学习(First Order Reduced and Controlled Error Learning:一阶减少和受控错误学习)中,首先,处理器输入储层输出向量(输出节点所输出的读出值的向量) $r(t)$ 和教师数据 $f(t)$ (S20)。在 n 个输出节点的情况下,储层输出向量 $r(t)$ 是时刻 t 下的1行 n 列的矩阵, n 个输出节点的读出值是矩阵的 n 个要素。

[0121] 接下来,处理器在初次的学习中(S21的“是”),对矩阵 $P(t)$ 和输出权重 $w(t)$ 分别进行初始化(S22、S23)。矩阵 P 是用于计算输出权重向量 $w(t)$ 的矩阵,以单位矩阵 I 进行初始化。然后,进行差分 e 的更新(S26)。

[0122] 接下来,处理器输入储层输出向量 $r(t)$ 和教师数据 $f(t)$ (S29),由于是初次以外的情况(S21的“否”),所以进行矩阵 P 的更新和输出权重 $w(t)$ 的更新(S24、S25)。各个更新的计算式被示于图15。在矩阵 P 的更新中,处理器根据前一时刻 $t-\Delta t$ 的矩阵 $P(t-\Delta t)$,通过S24所示的计算式,来计算当前时刻 t 的矩阵 $P(t)$ (S24)。另外,在输出权重 w 的更新中,处理器根据前一时刻 $t-\Delta t$ 的输出权重 $w(t-\Delta t)$,通过S25所示的计算式,来计算当前时刻 t 的输出权重 $w(t)$ (S25)。然后,处理器进行差分 e 的更新(S26)。在差分 e 的更新中,处理器对在前一时刻 $t-\Delta t$ 更新的输出权重 w 的转置矩阵乘以当前时刻 t 的储层输出向量 $r(t)$ 来进行乘积和运算,并将乘积和运算值与教师数据 $f(t)$ 的差分作为当前时刻 t 的差分 $e(t)$ 来输出。

[0123] 之后,每当输入新的教师数据和储层输出向量,处理器就反复执行工序S24~S26,并更新输出权重 $w(t)$ 的处理。上述的输出权重的更新处理通过图13的第一部分储层和读出部的输出权重的更新处理S5以及第二部分储层和读出部的输出权重的更新处理S7来进行。在第二部分储层和读出部的输出权重的更新处理S7中,教师数据为差分教师数据。

[0124] 图16是表示第三实施方式中的储层计算模型的图。图16的模型与图9同样是具有材料储层 M_RSV 和信息处理装置10的信息处理系统。

[0125] 材料储层 M_RSV 具有第一部分储层 M_RSV_1 、尺寸大于第一部分储层 M_RSV_1 的尺寸且具有更高维的函数特性的第二部分储层 M_RSV_2 、以及尺寸大于第二部分储层 M_RSV_2 的尺寸且具有更高维的函数特性的第三部分储层 M_RSV_3 。第一部分储层 M_RSV_1 、第二部分储层 M_RSV_2 以及第三部分储层 M_RSV_3 具有部分重合的区域,且在该区域设置共用的输入层的输入节点 IN_nd 。另外,在第一部分储层 M_RSV_1 、第二部分储层 M_RSV_2 、第三部分储层 M_RSV_3 ,设置各个输出层的节点 OUT_nd_1 、 OUT_nd_2 、 OUT_nd_3 。

[0126] 与第一实施方式同样,来自各个输出层的输出节点 OUT_nd_1 、 OUT_nd_2 、 OUT_nd_3 的读出值被输出至信息处理装置10,处理器运算读出值和输出权重的线性耦合的乘积和值,生成三个输出 $OUT1$ 、 $OUT2$ 、 $OUT3$,并计算三个输出 $OUT1$ 、 $OUT2$ 、 $OUT3$ 的和作为交叉开关储层的输出。

[0127] 另外,处理器进行基于学习工序中的训练数据的输出权重的优化。在该学习工序中,如图13所示,将训练数据的输入值输入至输入节点 IN_nd ,并使三个部分储层的状态迁移,并且,从尺寸小的部分储层开始按尺寸增大的顺序,分别更新三个部分储层的读出部的权重。在该情况下,在第二部分储层和读出部中,将基于第一部分储层的读出部的更新后的输出权重的第一输出与教师数据的差分作为差分教师数据。另外,在第三部分储层和读出部中,将基于第一部分储层的读出部的更新后的输出权重的第一输出与第二部分储层中的同样的第二输出的和与教师数据的差分作为差分教师数据。而且,处理器将与三个部分储层对应的第一输出、第二输出、第三输出的和作为材料储层的输出来输出。

[0128] [实施例]

[0129] 在实施例中,通过处理器执行交叉开关储层的状态迁移运算程序来实现图8的交叉开关储层。交叉开关储层的节点间耦合和自耦合的权重以及输入节点和两个输出节点的例子例如与图12相同。其中,与图12的节点数为20个不同,在本实施例中,节点数为100个,是图12的5倍多。处理器执行在交叉开关结构的交点设置有具有非线性特性的忆阻器等电路元件的情况下的状态迁移运算程序。

[0130] 另外,在本实施例中,将输入节点设为10个,将第一输出节点设为10个,将第二输

出节点设为10个,将第一部分储层的第一中间层的节点数设为20个,将第二部分储层的第二中间层的节点数设为50个。交叉开关储层的节点间耦合和自耦合随机地选择,如输入节点和输出节点那样,未接近的节点间的耦合成为未耦合。

[0131] 作为比较例,使用未分割为多个部分储层的储层,采用所有节点数为100个、输入节点数为10个、输出节点数为10个、中间层的节点数为80个的储层模型。

[0132] 在实施例的实验中,将被称为NARMA10的非线性函数的预测问题作为模型的任务。NARMA10作为RNN的评价方法一般来说是已知的。NARMA10的非线性函数若输入某个长度的时间序列的输入数据,则输出针对输入的时间序列的数值数据。该输入和输出的时间序列的数值数据是训练数据的输入数据和教师数据。

[0133] 在实施例的实验中,向本实施例的储层计算模型按时间序列输入上述训练数据的输入数据,并反复执行了使用对应的教师数据来优化模型的输出权重的学习。然后,计算出在将NARMA10的输入数据给予至学习后的模型时由模型生成的输出数据的与NARMA10的函数输出的正解数据的误差的归一化双均方根误差(NRMSE)。另外,比较例的储层模型也计算了同样的学习和学习后的误差。其结果是,在比较例的储层模型中,误差为0.2033,相对于此,在本实施方式的储层模型中,误差小到0.1718。

[0134] 如上所述,本实施方式的储层计算模型向将储层按各个尺寸依次增大的多个部分储层输入共用的输入,并进行按尺寸从小到达的顺序更新各部分储层和读出部的输出权重以使输出和教师数据的差分变小的学习。在该情况下,在最小的第一部分储层和读出部中,更新其输出权重,以使读出部的第一输出与教师数据的差分变小,在第二小的第二部分储层和读出部中,更新其输出权重,以使读出部的第二输出与差分教师数据(从教师数据中减去以更新完毕输出权重求出的第一输出所得的值)的差分,并且,在第K部分储层和读出部中,更新其输出权重,以使读出部的第K输出与差分教师数据(从教师数据中减去以更新完毕输出权重求出的第一输出~第K-1输出的和所得的值)的差分变小。

[0135] 根据本实施方式,储层计算模型在尺寸最小的部分储层和其读出部生成输出最低的周波成分,随着尺寸增大,生成输出更高的频率成分。由此,能够抑制学习工序中的过度学习,生成泛化性高的储层计算模型。

[0136] 附图标记说明:10…信息处理装置;12…处理器;18…物理储层;RSV…储层;C_RSV…交叉开关储层;M_RSV…材料储层;C_RSV_1…第一部分储层;C_RSV_2…第二部分储层;INL…输入层;OUTL…输出层;MIDL…中间层;FCL…全耦合层、读出部;IN_nd…输入节点;OUT_nd_1、OUT_nd_2…输出节点;OUT1、OUT2…输出节点;w1、w2…输出权重;N1~N5…节点。

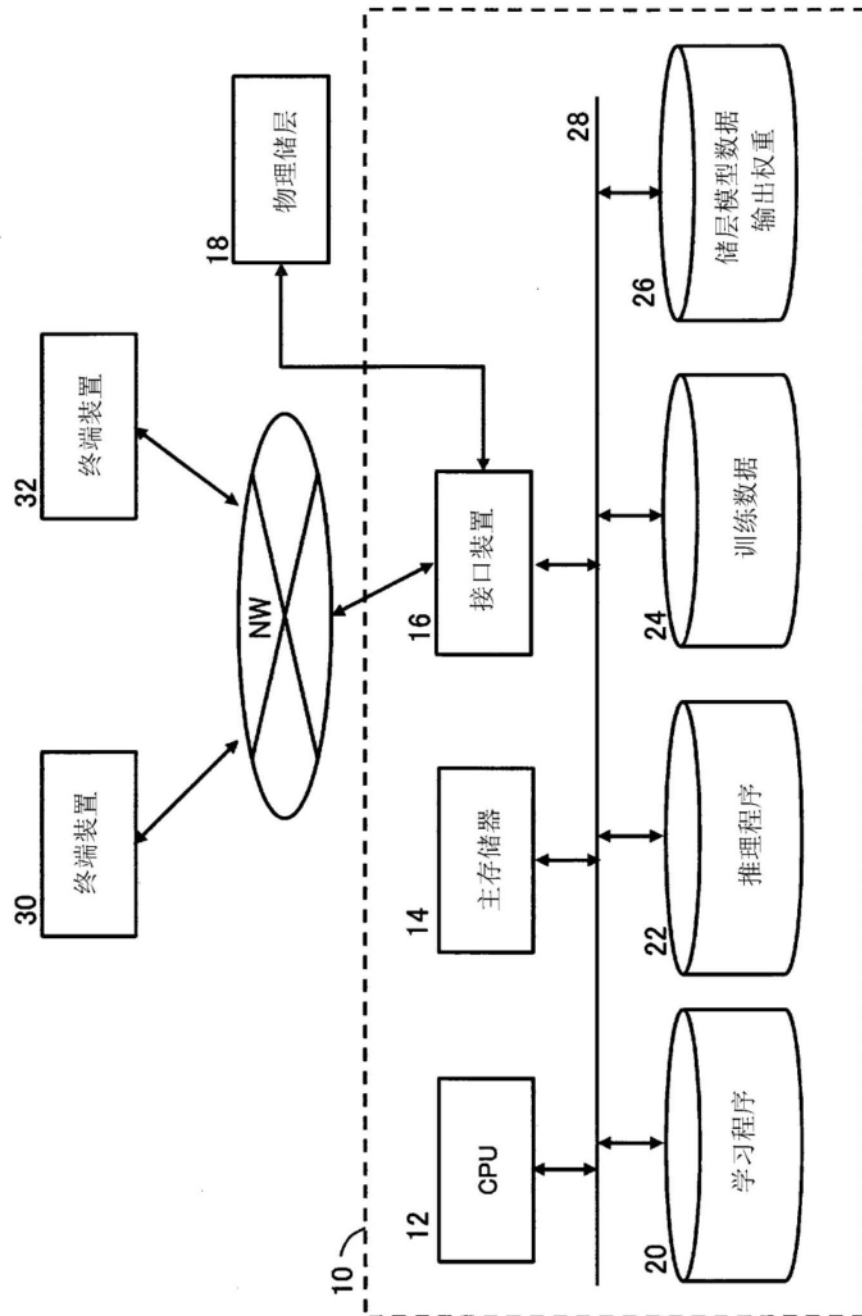


图1

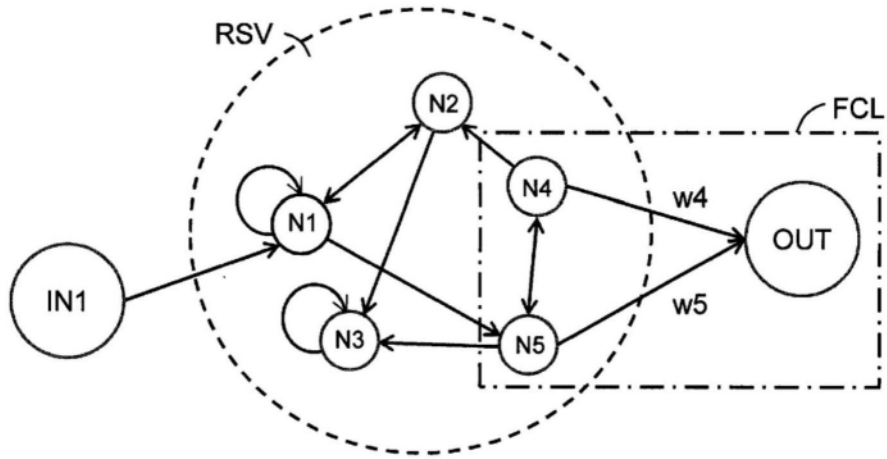


图2

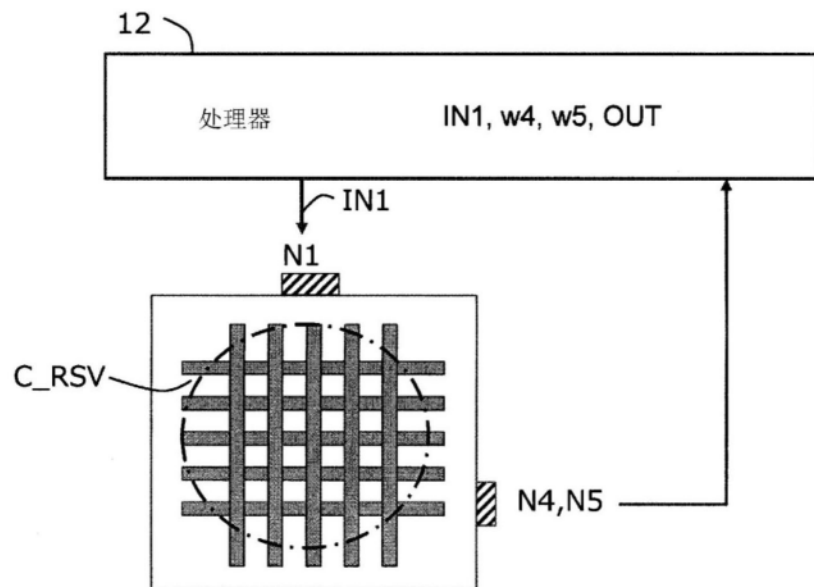


图3

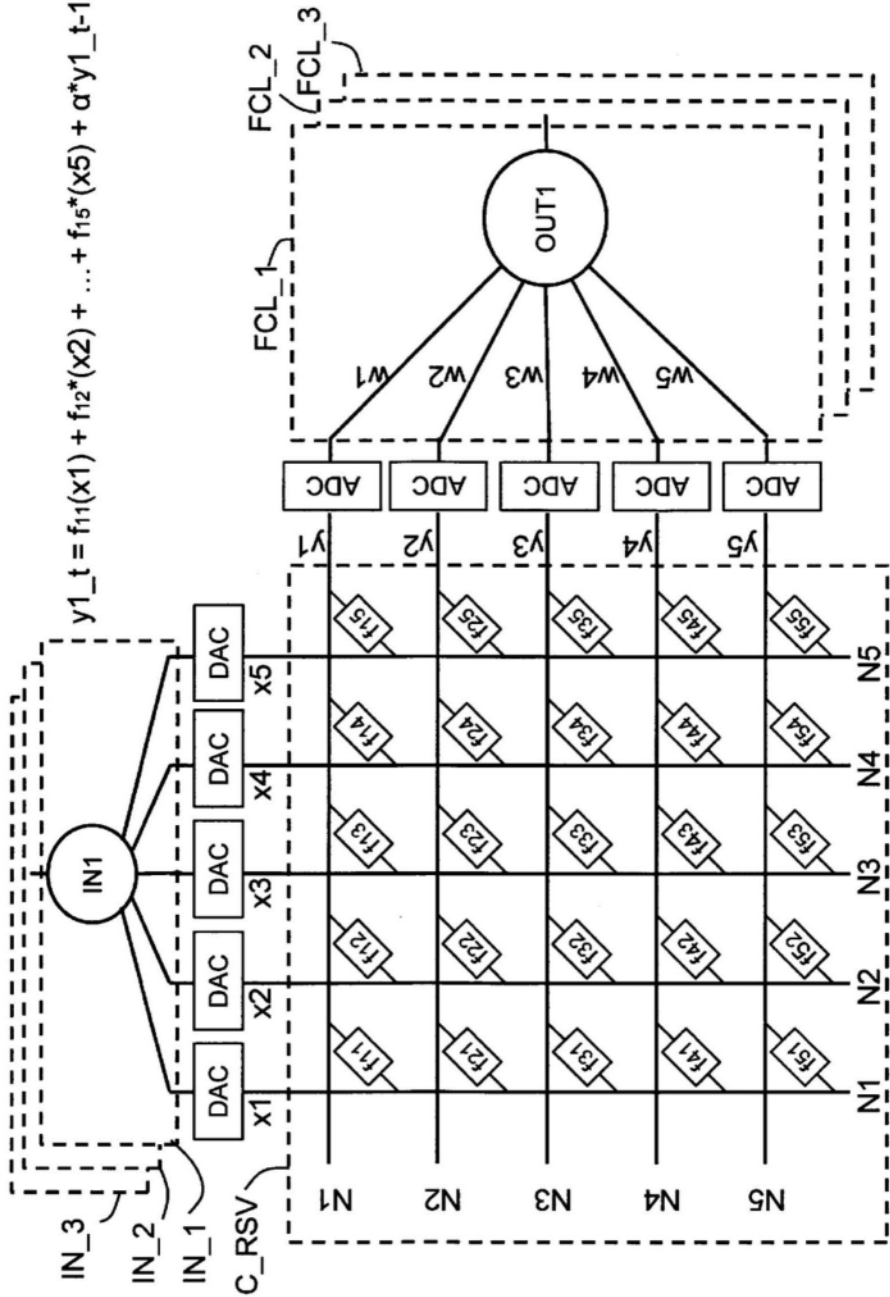


图4

节点（输入）

NMX

	N1	N2	N3	N4	N5
N1	0.87	-0.64			
N2	0.43			0.90	
N3		-0.41	-0.27		0.07
N4					0.51
N5	0.65			-0.43	

节点（输出）

图5

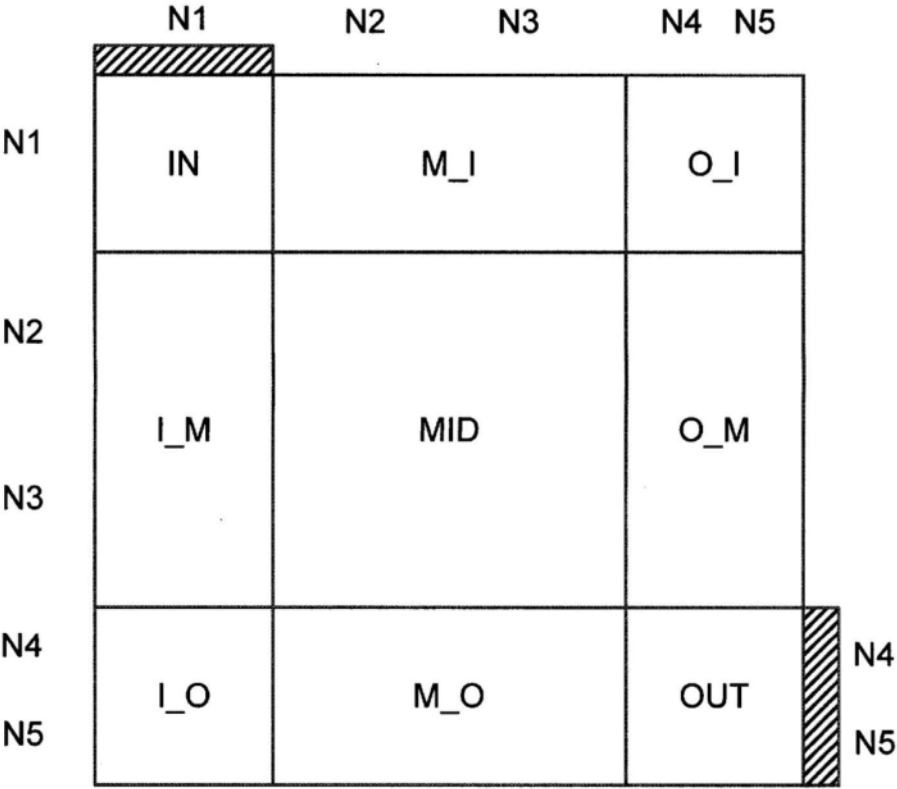


图6

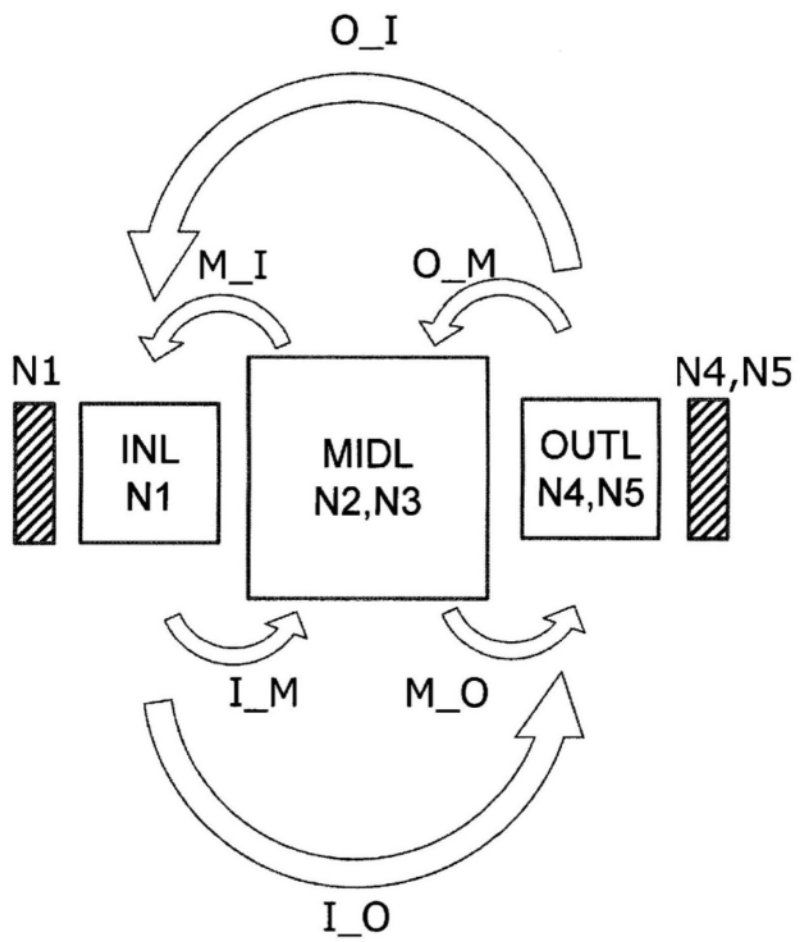


图7

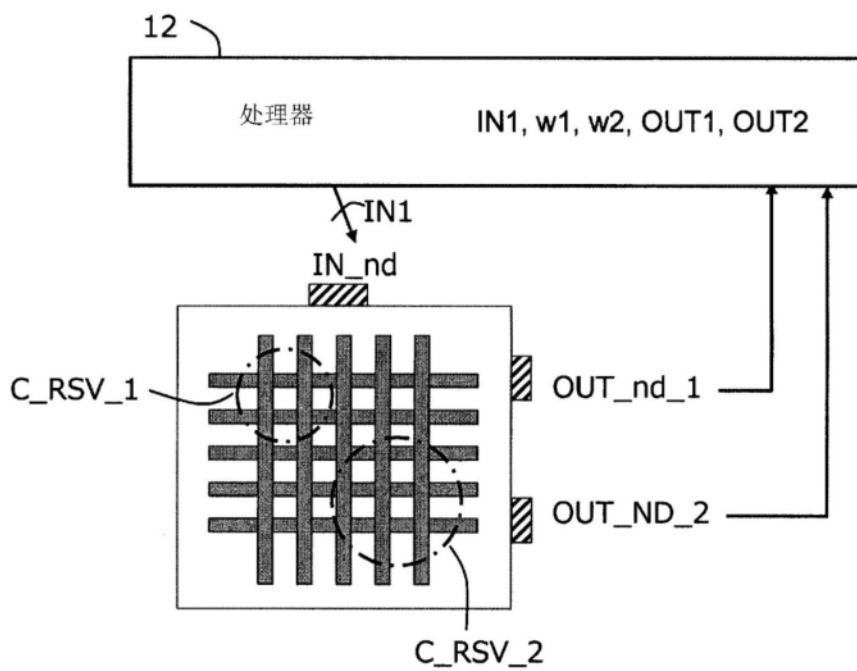


图8

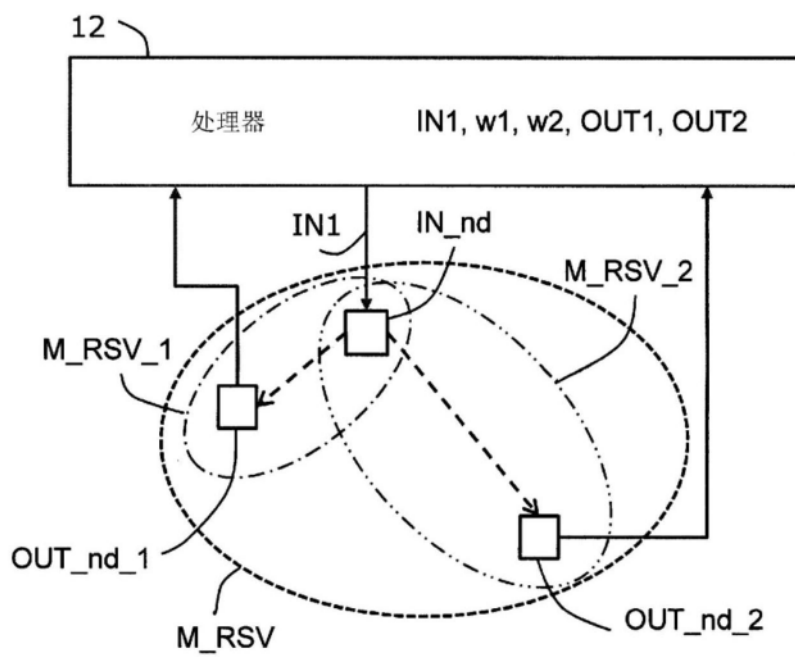


图9

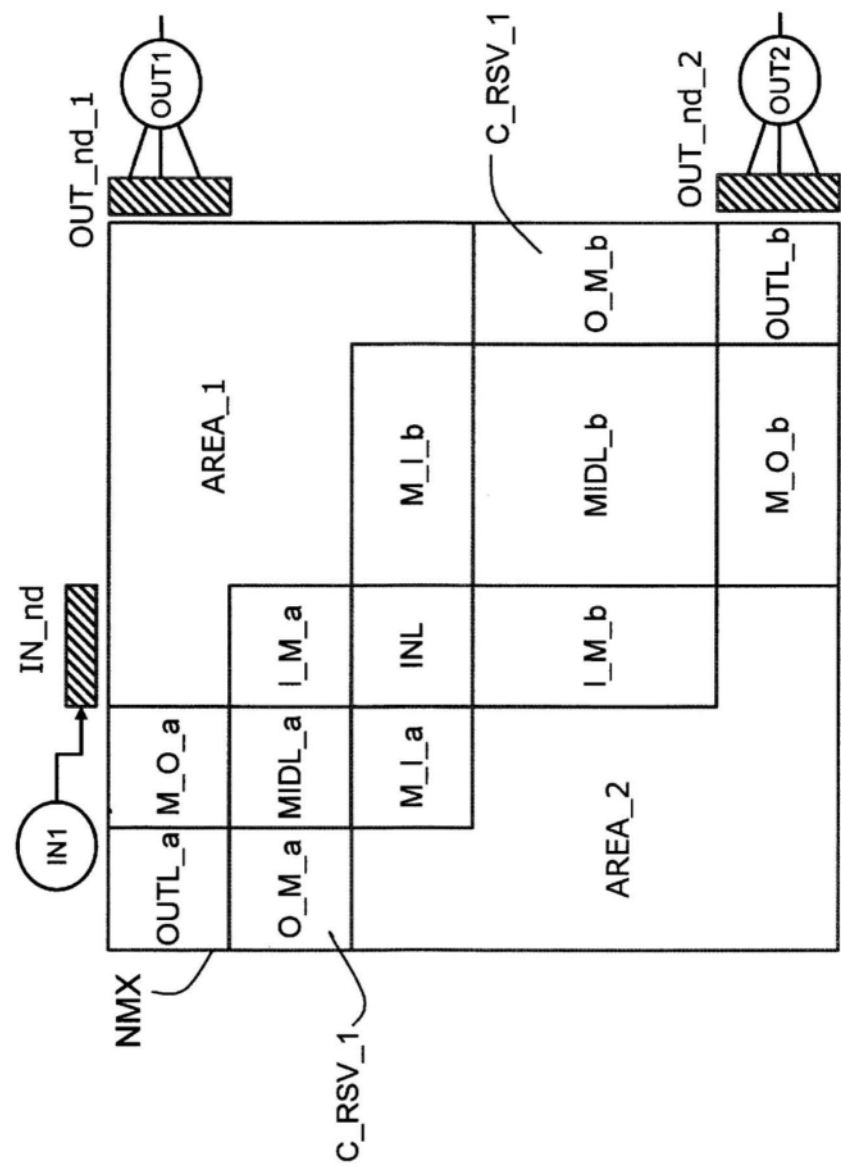


图10

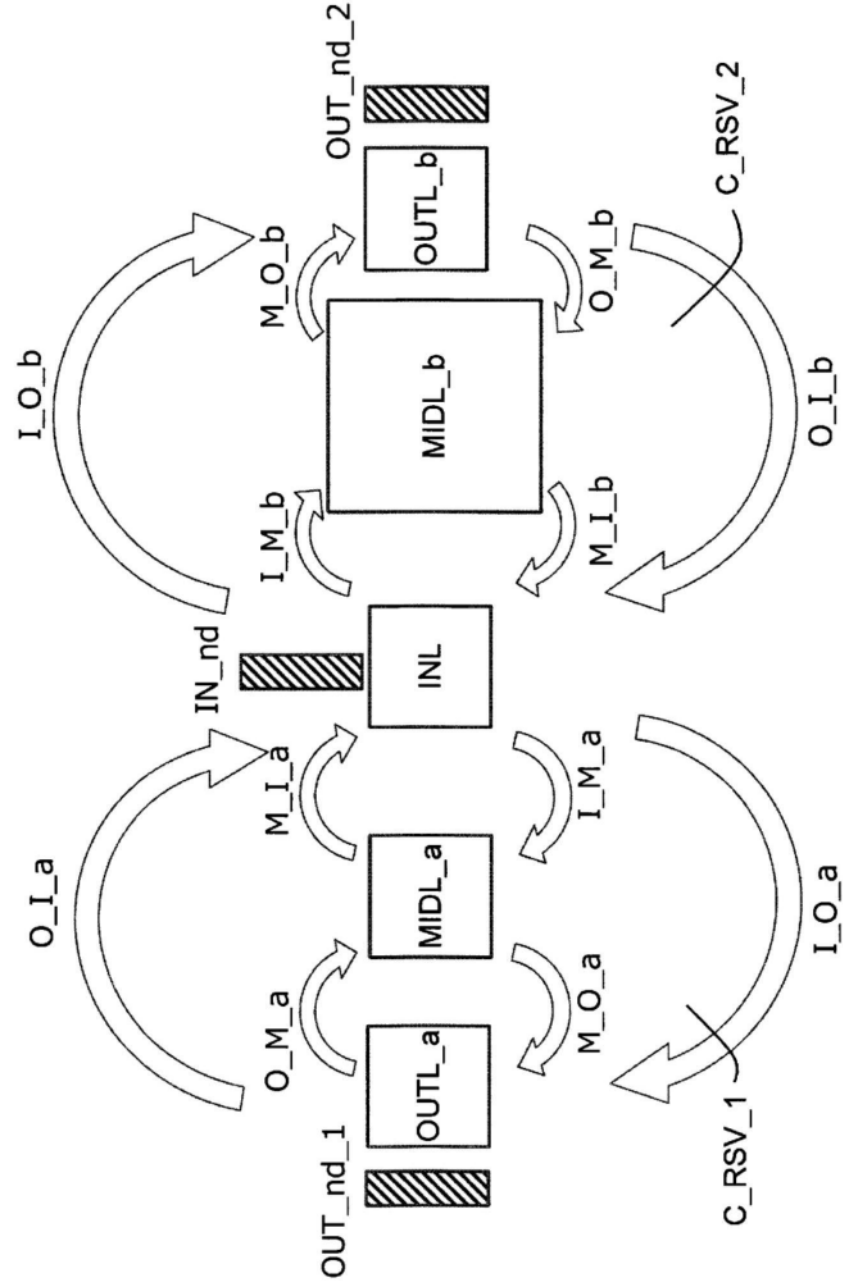


图11

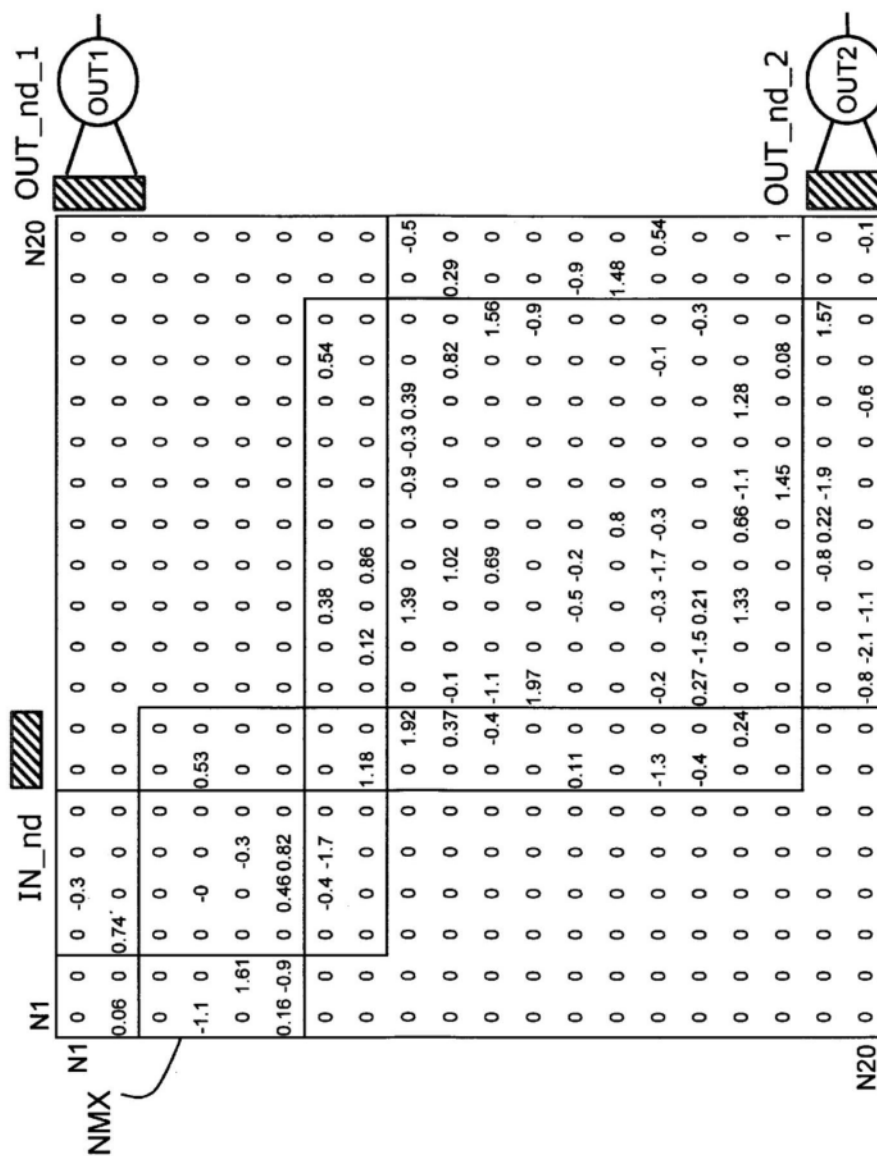


图12

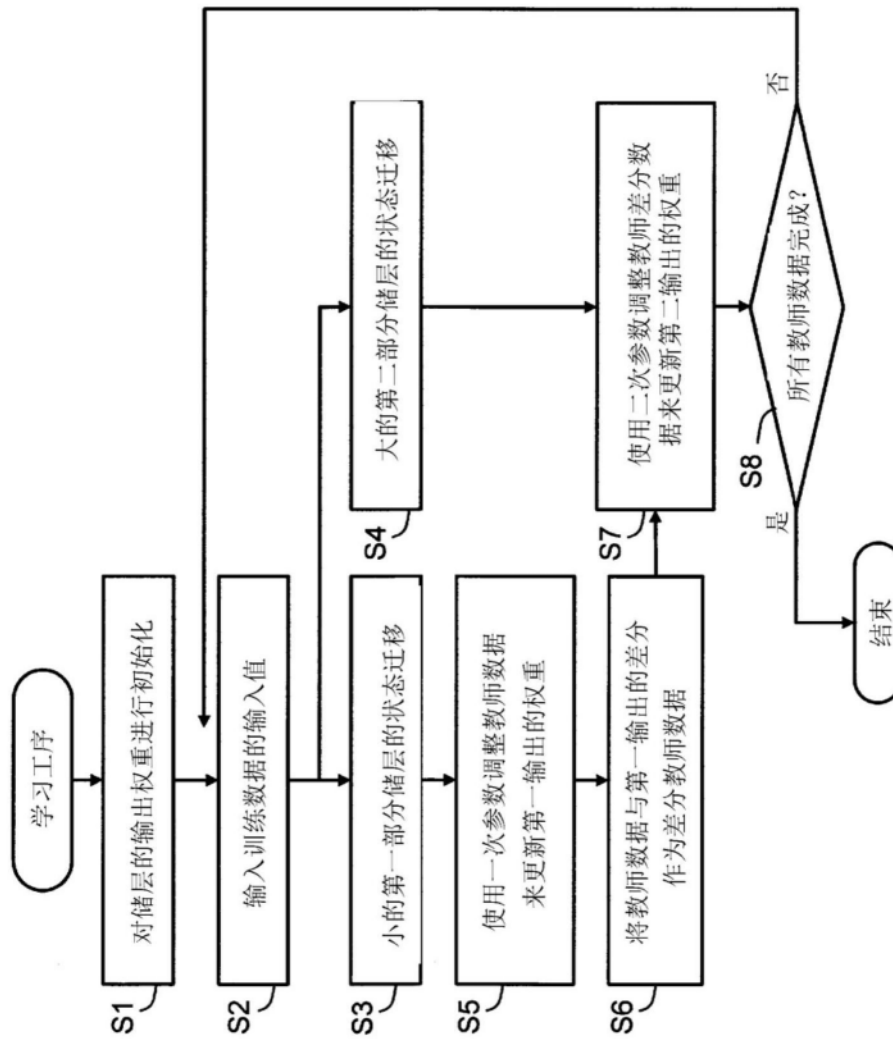


图13

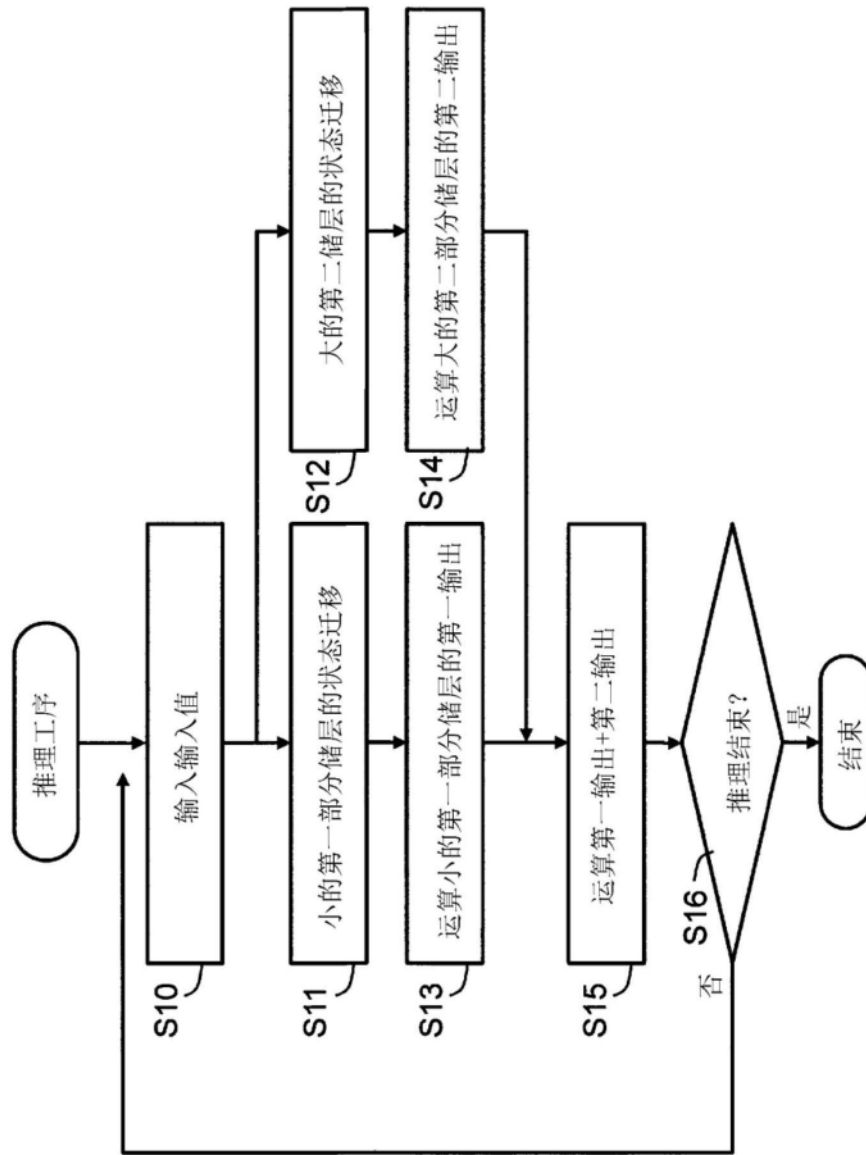


图14

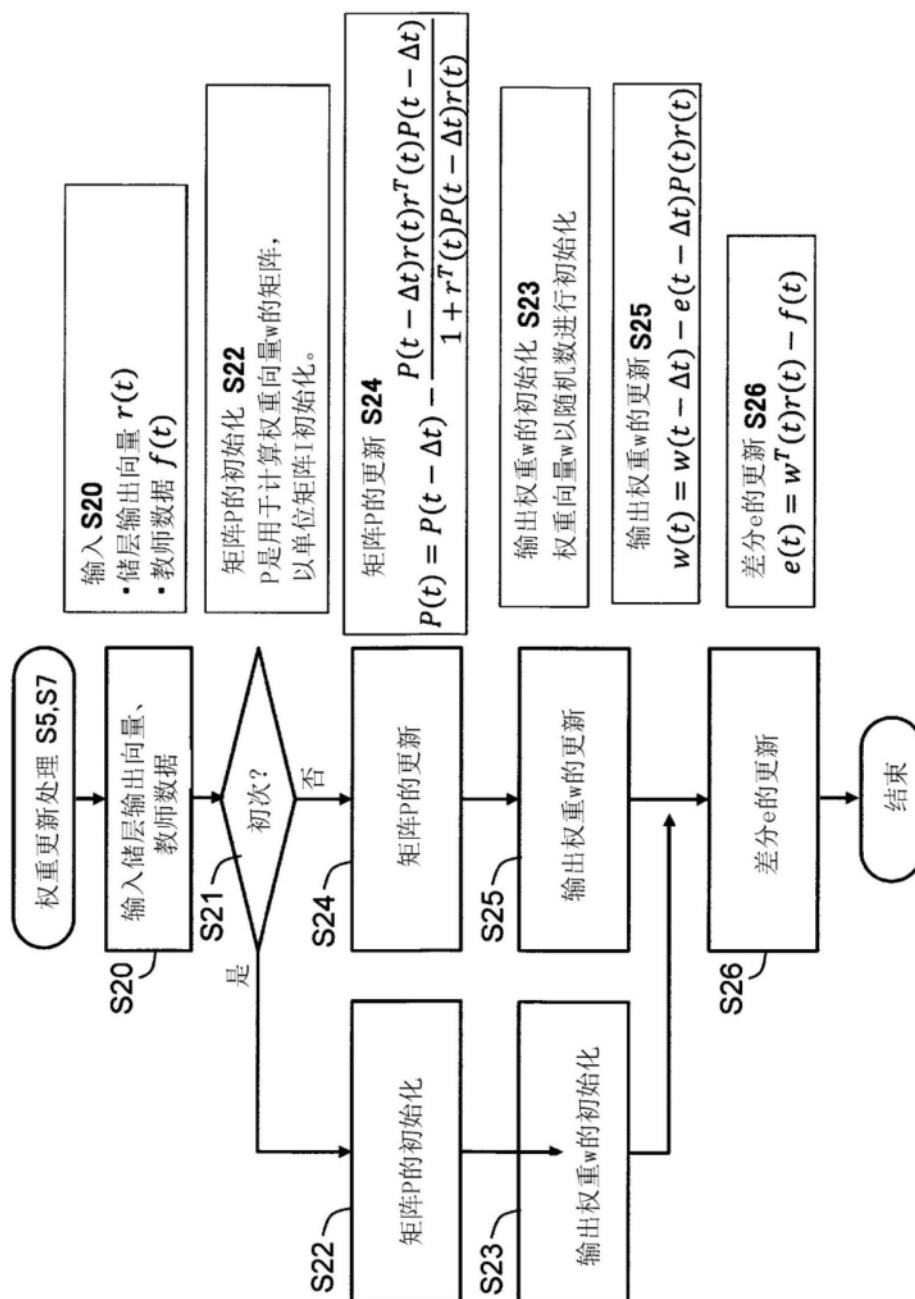


图15

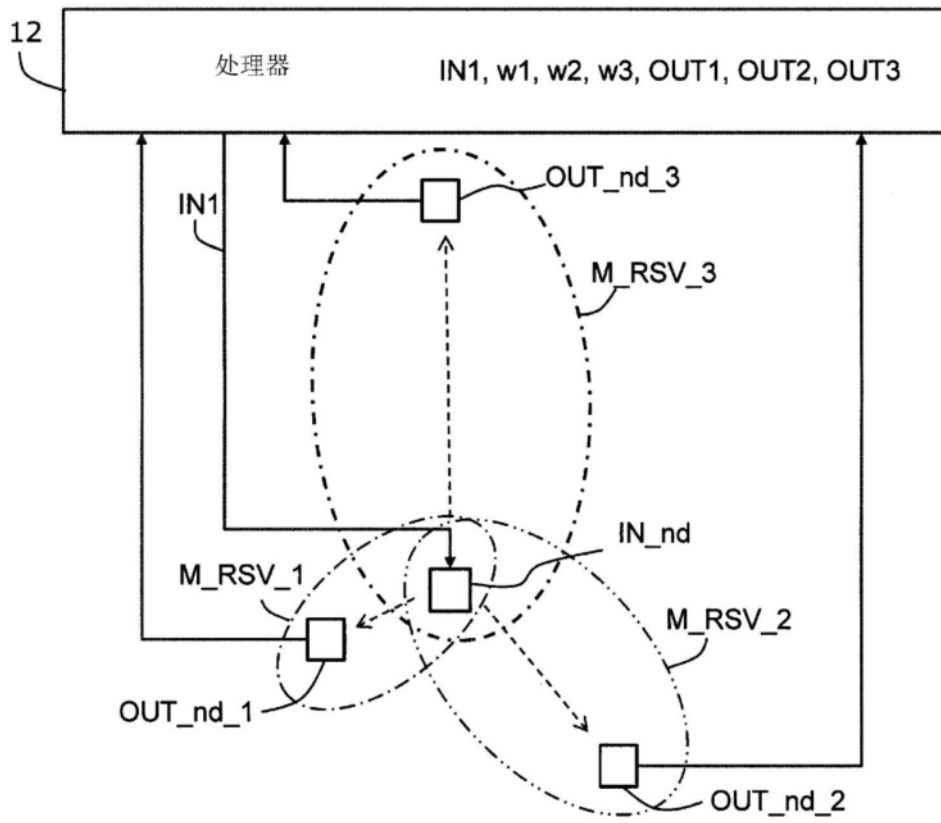


图16