



(12) 发明专利申请

(10) 申请公布号 CN 116745850 A

(43) 申请公布日 2023. 09. 12

(21) 申请号 202180089147.0

(22) 申请日 2021.11.22

(30) 优先权数据

2020-197046 2020.11.27 JP

(85) PCT国际申请进入国家阶段日

2023.07.04

(86) PCT国际申请的申请数据

PCT/JP2021/042833 2021.11.22

(87) PCT国际申请的公布数据

W02022/113945 JA 2022.06.02

(71) 申请人 株式会社力森诺科

地址 日本东京都

(72) 发明人 花冈恭平

(74) 专利代理机构 北京银龙知识产权代理有限公司 11243

专利代理师 孔博 胡玉美

(51) Int.Cl.

G16C 20/30 (2006.01)

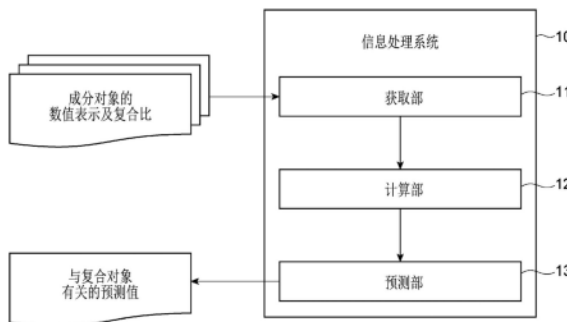
权利要求书2页 说明书9页 附图5页

(54) 发明名称

信息处理系统、信息处理方法及信息处理程序

(57) 摘要

一实施方式所涉及的信息处理系统具备至少一个处理器。至少一个处理器获取针对多个成分对象的每一个的数值表示及复合比,根据多个数值表示执行机器学习,计算与多个成分对象对应的多个回归参数,将多个复合比适用于由多个回归参数定义的回归模型,计算表示通过使多个成分对象复合而获得的复合对象的特性的预测值。



1. 一种信息处理系统,其具备至少一个处理器,
所述至少一个处理器进行如下处理:
获取针对多个成分对象的每一个的数值表示及复合比,
根据多个所述数值表示执行机器学习,计算与多个成分对象对应的多个回归参数,
将多个所述复合比适用于由所述多个回归参数定义的回归模型,计算表示通过使所述多个成分对象复合而获得的复合对象的特性的预测值。
2. 根据权利要求1所述的信息处理系统,其中,
所述至少一个处理器进行如下处理:
将所述多个数值表示输入到第1机器学习模型,计算与多个成分对象对应的多个特征矢量,
将所述多个特征矢量输入到第2机器学习模型,计算所述多个回归参数。
3. 根据权利要求2所述的信息处理系统,其中,
所述第1机器学习模型包括嵌入函数用机器学习模型和相互作用函数用机器学习模型,
所述至少一个处理器进行如下处理:
将所述多个数值表示输入到所述嵌入函数用机器学习模型,计算与多个成分对象对应的多个第1特征矢量,
将所述多个第1特征矢量输入到所述相互作用函数用机器学习模型,计算与多个成分对象对应的多个第2特征矢量,
将所述多个第2特征矢量输入到所述第2机器学习模型,计算所述多个回归参数。
4. 根据权利要求3所述的信息处理系统,其中,
所述嵌入函数用机器学习模型是根据作为非定型数据的所述数值表示生成作为固定长度矢量的所述第1特征矢量的机器学习模型。
5. 根据权利要求1至4中任一项所述的信息处理系统,其中,
所述回归模型是Scheffe多项式,
所述至少一个处理器计算所述Scheffe多项式的1次项的多个回归系数作为所述多个回归参数。
6. 根据权利要求5所述的信息处理系统,其中,
所述至少一个处理器还计算所述Scheffe多项式的2次项的多个回归系数作为所述多个回归参数。
7. 根据权利要求1至6中任一项所述的信息处理系统,其中,
所述成分对象是材料,所述复合对象是多成分物质。
8. 根据权利要求7所述的信息处理系统,其中,
所述材料是聚合物或单体,所述多成分物质是聚合物合金。
9. 一种信息处理方法,其由具备至少一个处理器的信息处理系统执行,所述信息处理方法包括如下步骤:
获取针对多个成分对象的每一个的数值表示及复合比;
根据多个所述数值表示执行机器学习,计算与多个成分对象对应的多个回归参

数;及

将多个所述复合比适用于由所述多个回归参数定义的回归模型,计算表示通过使所述多个成分对象复合而获得的复合对象的特性的预测值。

10.一种信息处理程序,其使计算机执行如下步骤:

获取针对多个成分对象的每一个的数值表示及复合比;

根据多个所述数值表示执行机器学习,计算与所述多个成分对象对应的多个回归参数;及

将多个所述复合比适用于由所述多个回归参数定义的回归模型,计算表示通过使所述多个成分对象复合而获得的复合对象的特性的预测值。

信息处理系统、信息处理方法及信息处理程序

技术领域

[0001] 本发明的一方式涉及一种信息处理系统、信息处理方法及信息处理程序。

背景技术

[0002] 使用利用机器学习对通过使多个成分对象复合而获得的复合对象进行分析的方法。例如,专利文献1中记载了预测生物高分子的立体结构与化合物的立体结构的键合性的方法。该方法包括:根据生物高分子的立体结构和化合物的立体结构,生成生物高分子与化合物的复合体的预测立体结构的步骤;将该预测立体结构变换为表示与相互作用模式的核对结果的预测立体结构矢量的步骤;及通过使用机器学习算法判断该预测立体结构矢量,预测生物高分子的立体结构与化合物的立体结构的键合性的步骤。

[0003] 以往技术文献

[0004] 专利文献

[0005] 专利文献1:日本特开2019-28879号公报

发明内容

[0006] 发明要解决的技术课题

[0007] 在成分对象多样或存在多个的情况下,针对这些成分对象无法准备足够量的数据,其结果,复合对象的分析精度有可能无法达到所期待的水准。因此,期望一种用于即使在针对成分对象无法准备足够量的数据的情况下也提高复合对象的分析精度的结构。

[0008] 用于解决技术课题的手段

[0009] 本发明的一方式所涉及的信息处理系统具备至少一个处理器。至少一个处理器获取针对多个成分对象的每一个的数值表示及复合比,根据多个数值表示执行机器学习,计算与多个成分对象对应的多个回归参数,将多个复合比适用于由多个回归参数定义的回归模型,计算表示通过使多个成分对象复合而获得的复合对象的特性的预测值。

[0010] 本发明的一方式所涉及的信息处理方法由具备至少一个处理器的信息处理系统执行。该信息处理方法包括如下步骤:获取针对多个成分对象的每一个的数值表示及复合比;根据多个数值表示执行机器学习,计算与多个成分对象对应的多个回归参数;及将多个复合比适用于由多个回归参数定义的回归模型,计算表示通过使多个成分对象复合而获得的复合对象的特性的预测值。

[0011] 本发明的一方式所涉及的信息处理程序使计算机执行如下步骤:获取针对多个成分对象的每一个的数值表示及复合比;根据多个数值表示执行机器学习,计算与多个成分对象对应的多个回归参数;及将多个复合比适用于由多个回归参数定义的回归模型,计算表示通过使多个成分对象复合而获得的复合对象的特性的预测值。

[0012] 在这样的方式中,根据各成分对象的数据执行机器学习,计算与多个成分对象对应的多个回归参数。并且,将复合比适用于由该回归参数定义的回归模型,预测复合对象的特性。通过使用机器学习模型及回归模型,即使在针对成分对象无法准备足够量的数据的

情况下,也能够提高复合对象的分析精度。

[0013] 发明效果

[0014] 根据本发明的一方式,即使在针对成分对象无法准备足够量的数据的情况下,也能够提高复合对象的分析精度。

附图说明

[0015] 图1是表示构成实施方式所涉及的信息处理系统的计算机的硬件结构的一例的图。

[0016] 图2是表示实施方式所涉及的信息处理系统的功能结构的一例的图。

[0017] 图3是表示实施方式所涉及的信息处理系统的动作的一例的流程图。

[0018] 图4是表示计算回归参数的步骤的一例的图。

[0019] 图5是表示计算回归参数的步骤的另一例的图。

具体实施方式

[0020] 以下,参考附图,对本发明中的实施方式详细地进行说明。在附图说明中,对于相同或等同的要件标注相同的符号,省略重复的说明。

[0021] [系统的概要]

[0022] 实施方式所涉及的信息处理系统10是执行与通过以给定的复合比使多个成分对象复合而获得的复合对象有关的分析的计算机系统。成分对象是指用于生成复合对象的有体物或无体物。复合对象可以是有体物或无体物。作为有体物的例子,可举出任意物质或物体。作为无体物的例子,可举出数据及信息。“使多个成分对象复合”是指将多个成分对象作为一个对象、即复合对象的处理。复合的方法没有限定,例如可以是配合、调配、合成、结合、混合、合并、组合、化合、或合体,也可以是其他方法。与复合对象有关的分析是指用于获得表示复合对象的某些特性的数据的处理。

[0023] 多个成分对象可以是任意多种材料,在该情况下,复合对象是由这些材料生成的多成分物质。材料是指用于生成多成分物质的任意构成要素。例如,多个材料可以是任意多种分子或原子,在该情况下,复合对象是通过以任意方法复合这些分子或原子而获得的多成分物质。例如,材料可以是聚合物或单体,相对于此,多成分物质可以是聚合物合金。材料可以是单体,相对于此,多成分物质可以是聚合物。材料可以是药物、即具有药理作用的化学物质,相对于此,多成分物质可以是药剂。

[0024] 信息处理系统10为了与复合对象有关的分析而执行机器学习。机器学习是指根据给定的信息进行学习而自主地找出定律或规则的方法。机器学习的具体的方法没有限定。例如,信息处理系统10可以使用作为包含神经网络而构成的计算模型的机器学习模型来执行机器学习。神经网络是指模拟人的脑神经系统的结构的信息处理模型。作为更具体的例子,信息处理系统10也可以使用图神经网络(GNN)、卷积神经网络(CNN)、循环神经网络(RNN)、注意力RNN(Attention RNN)及多头注意力(Multi-Head Attention)中的至少一个来执行机器学习。

[0025] [系统的结构]

[0026] 信息处理系统10由1台以上的计算机构成。在使用多台计算机的情况下,这些计算

机经由互联网、内部网等通信网络连接,从而逻辑上构建一个信息处理系统10。

[0027] 图1是表示构成信息处理系统10的计算机100的一般硬件结构的一例的图。例如,计算机100具备执行操作系统、应用程序等的CPU等处理器101、由ROM及RAM构成的主存储部102、由硬盘、闪存等构成的辅助存储部103、由网卡或无线通信模块构成的通信控制部104、键盘、鼠标等输入装置105、及监视器等输出装置106。

[0028] 信息处理系统10的各功能要件通过在处理器101或主存储部102上读入预先设定的程序并使处理器101执行该程序来实现。处理器101按照该程序,使通信控制部104、输入装置105、或输出装置106动作,进行主存储部102或辅助存储部103中的数据读出及写入。处理所需的数据或数据库存储在主存储部102或辅助存储部103内。

[0029] 图2是表示信息处理系统10的功能结构的一例的图。信息处理系统10具备获取部11、计算部12及预测部13作为功能要件。

[0030] 获取部11是获取与多个成分对象有关的数据的功能要件。具体而言,获取部11针对多个成分对象的每一个获取数值表示及复合比。成分对象的数值表示是指使用多个数值表示成分对象的任意属性的数据。成分对象的属性是指成分对象所具备的性质或特征。数值表示可以通过各种方法进行可视化,例如可以通过数字、英文字符、文本、分子图、矢量、图像、时序数据等方法进行可视化,也可以通过这些方法中的任意2个以上的组合进行可视化。构成数值表示的各个数值可以用十进制表示,也可以用二进制、十六进制等其他标记法表示。成分对象的复合比是指多个成分对象之间的比例。复合比的具体种类、单位及表示方法没有限定,可以根据成分对象或复合对象任意设定。例如,复合比可以由百分率等比率表示,也可以由直方图表示,也可以由各个成分对象的绝对量表示。

[0031] 计算部12是计算用于预测复合对象的特性的回归模型的回归参数的功能要件。具体而言,计算部12根据与多个成分对象对应的多个数值表示执行机器学习,计算回归参数。回归模型是指,在给定1以上的解释变量x的值的情况下,用于求出1以上的目标变量y的值的公式。回归模型可以是线性回归模型,也可以是非线性回归模型。作为回归模型的例子,可举出Scheffe多项式(Scheffe polynomial)。然而,回归模型也可以是其他参数模型。回归参数是回归模型中包含的数值。

[0032] 预测部13是预测复合对象的特性并输出其预测值的功能要件。复合对象的特性是指复合对象所具有的特有的性质。具体而言,预测部13将复合比适用于由计算出的回归参数定义的回归模型,计算预测值。换言之,预测部13将多个复合比代入该回归模型来计算预测值。

[0033] 在一例中,计算部12及预测部13的组合通过一个机器学习模型来实现。或者,计算部12可以通过机器学习模型来实现,预测部13可以通过不使用机器学习模型的算法来实现。

[0034] 在一例中,在本实施方式中使用的至少一个机器学习模型的每一个是期待推算精度最高的学习完毕模型,因此能够称为“最佳机器学习模型”。然而,应该注意,该学习完毕模型并不限于“现实中最佳”。通过由给定的计算机处理包含输入矢量和标签的多个组合的教师数据来生成学习完毕模型。给定的计算机将输入矢量输入到机器学习模型来计算输出值,求出该输出值与由教师数据表示的标签的误差。输出值例如是预测值。输出值与标签的误差可以说是推算结果与正确答案之差。计算机根据该误差更新机器学习模型内的给定的

参数。计算机通过重复这样的学习来生成学习完毕模型。生成学习完毕模型的计算机没有限定,例如可以是信息处理系统10,也可以是其他计算机系统。生成学习完毕模型的处理能够称为学习阶段,利用该学习完毕模型的处理能够称为运用阶段。

[0035] 在一例中,在本实施方式中使用的机器学习模型整体可以由不依赖于输入顺序的函数来描述。通过该结构,能够排除机器学习中多个矢量的排列顺序的影响。

[0036] [系统的动作]

[0037] 参考图3,对信息处理系统10的动作进行说明,并且对本实施方式所涉及的信息处理方法进行说明。图3是将信息处理系统10的动作的一例作为处理流程S1表示的流程图。处理流程S1相当于运用阶段。

[0038] 在步骤S11中,获取部11针对多个成分对象的每一个获取数值表示及复合比。作为一例,若输入与两个成分对象Ea、Eb有关的信息,则获取部11例如获取成分对象Ea的数值表示{1,1,2,3,4,3,3,5,6,7,5,4}、成分对象Eb的数值表示{1,1,5,6,4,3,3,5,1,7,0,0}、成分对象Ea、Eb的复合比{0.7、0.3}。在该例子中,各个数值表示用矢量表示。复合比{0.7、0.3}是指以7:3的比例使用成分对象Ea、Eb来获得复合对象。

[0039] 获取部11可以通过任意方法获取多个成分对象各自的数据。例如,获取部11可以访问给定的数据库而读出数据,也可以从其他计算机或计算机系统接收数据,还可以接受由信息处理系统10的用户输入的数据。或者,获取部11也可以通过这些方法中的任意2种以上来获取数据。

[0040] 在步骤S12中,计算部12针对多个成分对象的每一个,根据数值表示计算特征矢量。特征矢量是指表示成分对象的特征的矢量。成分对象的特征是指使该成分对象与其他对象不同的任意要素。矢量是指具有n个数值的n维的量,能够表示为1维的排列。

[0041] 在步骤S13中,计算部12根据计算出的多个特征矢量,计算与多个成分对象对应的多个回归参数。

[0042] 在步骤S14中,预测部13使用由计算出的多个回归参数定义的回归模型,计算表示复合对象的特性的预测值。简而言之,由回归参数定义的回归模型是指将特定具体的数值确定为回归参数的回归模型。预测部13将多个复合比适用于回归模型,计算预测值。

[0043] 在步骤S15中,预测部13输出该预测值。预测值的输出方法没有限定。例如,预测部13可以将预测值存储在给定的数据库中,也可以发送到其他计算机或计算机系统,还可以显示在显示装置上。或者,预测部13也可以为了信息处理系统10中的后续处理,将预测值输出到其他功能要件。

[0044] 参考图4及图5,对与回归模型有关的处理更详细地进行说明。图4及图5都是表示计算回归参数的步骤的例子的图。在任意例子中,成分对象都表示聚苯乙烯、聚丙烯酸及聚甲基丙烯酸丁酯这3种材料(聚合物)。针对这些材料的每一个,可以准备任意形式的数值表示。

[0045] 在与材料的配合有关的问题中,经常使用上述Scheffe多项式。因此,在图4及图5的例子中,假设回归模型是Scheffe多项式。

[0046] 对图4的例子进行说明。在作为步骤S12的一部分的步骤S121中,计算部12通过用于计算矢量的特征的嵌入函数(embedding function)用机器学习模型,针对多个成分对象的每一个根据数值表示计算特征矢量Z。该机器学习模型是学习完毕模型。在嵌入函数中,

输入矢量与输出矢量具有1对1的关系。在该例子中,输入矢量是数值表示,输出矢量是特征矢量Z。计算部12将与多个成分对象对应的多个数值表示输入到嵌入函数用模型,计算该多个成分对象各自的特征矢量Z。在一例中,计算部12针对多个成分对象的每一个,将与该成分对象对应的数值表示输入到嵌入函数用模型,计算该成分对象的特征矢量Z。在一例中,嵌入函数用模型可以根据作为非定型数据的数值表示生成作为固定长度矢量的特征矢量Z。非定型数据是指不由固定长度矢量表示的数据。在图4的例子中,计算部12计算与聚苯乙烯对应的特征矢量 Z_1 、与聚丙烯酸对应的特征矢量 Z_2 及与聚甲基丙烯酸丁酯对应的特征矢量 Z_3 。

[0047] 嵌入函数用机器学习模型没有限定,可以考虑成分对象及复合对象的种类等因素而以任意方针决定。例如,计算部12可以使用图神经网络(GNN)、卷积神经网络(CNN)或循环神经网络(RNN)来执行嵌入函数。

[0048] 在作为步骤S12的一部分的步骤S122中,计算部12通过用于使多个矢量相互作用的相互作用函数(interaction function)用机器学习模型,针对多个成分对象根据特征矢量Z计算其他特征矢量M。该机器学习模型是学习完毕模型。在相互作用函数中,输入矢量与输出矢量具有1对1的关系。在该例子中,输入矢量是特征矢量Z,输出矢量是特征矢量M。在一例中,计算部12将与多个成分对象对应的多个特征矢量Z的集合输入到相互作用函数用模型,针对该多个成分对象的每一个计算特征矢量M。在图4的例子中,计算部12计算与聚苯乙烯对应的特征矢量 M_1 、与聚丙烯酸对应的特征矢量 M_2 及与聚甲基丙烯酸丁酯对应的特征矢量 M_3 。

[0049] 相互作用函数用机器学习模型没有限定,可以考虑成分对象及复合对象的种类等因素而以任意方针决定。例如,计算部12可以使用注意力RNN(Attention RNN)或多头注意力(Multi-Head Attention)来执行相互作用函数用机器学习。在另一例子中,计算部12也可以使用不包含学习参数的相互作用函数来计算特征矢量M。

[0050] 在图4所示的步骤S13中,计算部12针对多个成分对象的每一个,根据特征矢量M计算线性回归模型的1次项的回归参数a。在一例中,计算部12通过机器学习模型计算回归参数。该机器学习模型是学习完毕模型。在计算1次项的回归参数的函数中,输入矢量与输出值具有1对1的关系。在该例子中,输入矢量是特征矢量M,输出值是回归参数a。在一例中,计算部12将与多个成分对象对应的多个特征矢量M的集合输入到机器学习模型,针对该多个成分对象的每一个计算回归参数a。在图4的例子中,计算部12计算与聚苯乙烯对应的回归参数 a_1 、与聚丙烯酸对应的回归参数 a_2 及与聚甲基丙烯酸丁酯对应的回归参数 a_3 。

[0051] 用于计算回归参数的机器学习模型没有限定,可以考虑成分对象及复合对象的种类等因素而以任意方针决定。例如,计算部12可以使用全连接神经网络(FCNN)来计算回归参数。

[0052] 在图4所示的步骤S14中,预测部13通过由3个回归参数 a_1 、 a_2 、 a_3 定义的下述Scheffe多项式(1)来计算预测值E。回归参数a也可以说是式(1)的1次项的回归系数。预测值E表示由聚苯乙烯、聚丙烯酸及聚甲基丙烯酸丁酯获得的多成分物质(聚合物合金)的特性。式(1)中的变量r是指复合比。聚苯乙烯、聚丙烯酸及聚甲基丙烯酸丁酯的复合比分别表示为 r_1 、 r_2 、 r_3 。

[0053] [数式1]

$$[0054] \quad E = \sum_{1 \leq i \leq 3} a_i r_i \quad \dots(1)$$

[0055] 对图5的例子进行说明。在图5的例子中,包括步骤S121及步骤S122的步骤S12与图4的例子相同,步骤S13、S14与图4的例子不同。

[0056] 在图5所示的步骤S13中,计算部12针对多个成分对象的每一个,根据特征矢量M计算线性回归模型的回归参数。具体而言,计算部12计算1次项的回归参数a及2次项的回归参数b。在一例中,计算部12通过FCNN等机器学习来计算回归参数。针对线性回归模型的1次项及2次项分别准备机器学习模型。

[0057] 与图4的例子同样地,在计算1次项的回归参数的函数中,输入矢量与输出值具有1对1的关系。在该例子中,输入矢量是特征矢量M,输出值是回归参数a。在一例中,计算部12将与多个成分对象对应的多个特征矢量M的集合输入到机器学习模型,针对该多个成分对象的每一个计算回归参数a。在图5的例子中也同样地,计算部12计算与聚苯乙烯对应的回归参数 a_1 、与聚丙烯酸对应的回归参数 a_2 及与聚甲基丙烯酸丁酯对应的回归参数 a_3 。

[0058] 在计算2次项的回归参数的函数中,各个输入矢量通过合成两个特征矢量而获得。该函数是根据两个矢量计算一个回归参数的函数。在该例子中,合成两个特征矢量M。在图5的例子中,计算部12合成两个特征矢量 M_1 、 M_2 而生成第1输入矢量,合成两个特征矢量 M_1 、 M_3 而生成第2输入矢量,合成两个特征矢量 M_2 、 M_3 而生成第3输入矢量。因此,第1输入矢量与聚苯乙烯及聚丙烯酸对应,第2输入矢量与聚苯乙烯及聚甲基丙烯酸丁酯对应,第3输入矢量与聚丙烯酸及聚甲基丙烯酸丁酯对应。在计算2次项的回归参数的函数中,输入矢量与输出值也具有1对1的关系。在该例子中,输入矢量是两个特征矢量M的合成,输出值是回归参数b。在一例中,计算部12将输入矢量的所有组合输入到机器学习模型,针对各个组合计算回归参数b。在图5的例子中,计算部12计算与聚苯乙烯及聚丙烯酸的组合对应的回归参数 b_{12} 、与聚苯乙烯及聚甲基丙烯酸丁酯的组合对应的回归参数 b_{13} 及与聚丙烯酸及聚甲基丙烯酸丁酯的组合对应的回归参数 b_{23} 。

[0059] 在图5所示的步骤S14中,预测部13通过由6个回归参数 a_1 、 a_2 、 a_3 、 b_{12} 、 b_{13} 、 b_{23} 定义的下述Scheffe多项式(2)来计算预测值E。在式(2)中,回归参数a也可以说是1次项的回归系数,回归参数b也可以说是2次项的回归系数。式(2)中的变量r的含义与式(1)相同,是复合比。

[0060] [数式2]

$$[0061] \quad E = \sum_{1 \leq i \leq 3} a_i r_i + \sum_{1 \leq i < j \leq 3} b_{ij} r_i r_j \quad \dots(2)$$

[0062] 在图4及图5中示出了3个成分对象,但成分对象的数量当然不受限定,信息处理系统10可以处理任意数量的成分对象。

[0063] 针对包含3次以上的项或其他参数的回归模型也同样地,信息处理系统10根据相关的所有成分对象的特征矢量输出各个回归参数即可。在如线性回归的截距那样计算不依赖于特定解释变量的回归参数的情况下,信息处理系统10可以根据所有成分对象的特征矢量来输出一个回归参数。

[0064] 在图4及图5的例子中,计算部12执行嵌入函数及相互作用函数这两者,但可以省略这两个函数中的一者。例如,计算部12也可以根据通过嵌入函数用机器学习模型获得的

特征矢量Z来计算回归参数。在任何情况下,计算部12执行机器学习来计算回归参数。

[0065] 在一例中,嵌入函数用机器学习模型、相互作用函数用机器学习模型、回归参数用机器学习模型及回归模型可以由一个神经网络构建,也可以由多个神经网络的集合构建。或者,嵌入函数用机器学习模型、相互作用函数用机器学习模型及回归参数用机器学习模型可以由一个神经网络构建,也可以由多个神经网络的集合构建。

[0066] [程序]

[0067] 用于使计算机或计算机系统作为信息处理系统10发挥功能的信息处理程序包含用于使该计算机系统作为获取部11、计算部12及预测部13发挥作用的程序代码。该信息处理程序可以在非临时性地记录在CD-ROM、DVD-ROM、半导体存储器等有形记录介质的基础上提供。或者,信息处理程序也可以作为叠加在载波上的数据信号而经由通信网络提供。所提供的信息处理程序例如存储在辅助存储部103中。处理器101从辅助存储部103读出并执行该信息处理程序,由此实现上述各功能要件。

[0068] [效果]

[0069] 如以上说明,本发明的一方式所涉及的信息处理系统具备至少一个处理器。至少一个处理器获取针对多个成分对象的每一个的数值表示及复合比,根据多个数值表示执行机器学习,计算与多个成分对象对应的多个回归参数,将多个复合比适用于由多个回归参数定义的回归模型,计算表示通过使多个成分对象复合而获得的复合对象的特性的预测值。

[0070] 本发明的一方式所涉及的信息处理方法由具备至少一个处理器的信息处理系统执行。该信息处理方法包括如下步骤:获取针对多个成分对象的每一个的数值表示及复合比;根据多个数值表示执行机器学习,计算与多个成分对象对应的多个回归参数;及将多个复合比适用于由多个回归参数定义的回归模型,计算表示通过使多个成分对象复合而获得的复合对象的特性的预测值。

[0071] 本发明的一方式所涉及的信息处理程序使计算机执行如下步骤:获取针对多个成分对象的每一个的数值表示及复合比;根据多个数值表示执行机器学习,计算与多个成分对象对应的多个回归参数;及将多个复合比适用于由多个回归参数定义的回归模型,计算表示通过使多个成分对象复合而获得的复合对象的特性的预测值。

[0072] 在这样的方式中,根据各成分对象的数据执行机器学习,计算与多个成分对象对应的多个回归参数。并且,将复合比适用于由该回归参数定义的回归模型,预测复合对象的特性。通过使用机器学习模型及回归模型,即使在针对成分对象无法准备足够量的数据的情况下,也能够提高复合对象的分析精度。

[0073] 一旦获得回归参数,就能够变更复合比,通过回归模型瞬间重新计算复合对象的特性。即,能够重新利用计算出的回归参数。采用通过机器学习推算回归参数的方法,由此能够高速执行在变更复合比的同时探索复合对象的特性的处理。

[0074] 在另一方式所涉及的信息处理系统中,至少一个处理器可以将多个数值表示输入到第1机器学习模型,计算与多个成分对象对应的多个特征矢量,将多个特征矢量输入到第2机器学习模型,计算多个回归参数。通过该一连串的步骤,即使在针对成分对象无法准备足够量的数据的情况下,也能够进一步提高复合对象的分析精度。

[0075] 在另一方式所涉及的信息处理系统中,第1机器学习模型可以包括嵌入函数用机

器学习模型和相互作用函数用机器学习模型。至少一个处理器可以将多个数值表示输入到嵌入函数用机器学习模型,计算与多个成分对象对应的多个第1特征矢量,将多个第1特征矢量输入到相互作用函数用机器学习模型,计算与多个成分对象对应的多个第2特征矢量,将多个第2特征矢量输入到第2机器学习模型,计算多个回归参数。通过这样构成第1机器学习模型,即使在针对成分对象无法准备足够量的数据的情况下,也能够进一步提高复合对象的分析精度。

[0076] 在另一方式所涉及的信息处理系统中,嵌入函数用机器学习模型可以是根据作为非定型数据的数值表示生成作为固定长度矢量的第1特征矢量的机器学习模型。通过使用该机器学习模型,能够从不能由固定长度矢量表示的数值表示获得特征矢量。

[0077] 在另一方式所涉及的信息处理系统中,回归模型可以是Scheffe多项式。至少一个处理器可以计算Scheffe多项式的1次项的多个回归系数作为多个回归参数。通过使用在与配合有关的问题中经常使用的Scheffe多项式,能够高精度地分析通过配合多个成分对象而获得的复合对象。此外,能够根据1次项的回归系数,计算考虑了成分对象单独的影响度的预测值。

[0078] 在另一方式所涉及的信息处理系统中,至少一个处理器还可以计算Scheffe多项式的2次项的多个回归系数作为多个回归参数。在该情况下,能够根据2次项的回归系数,计算进一步考虑了两个成分对象的合成的影响度的预测值。

[0079] 在另一方式所涉及的信息处理系统中,成分对象可以是材料,复合对象可以是多成分物质。在该情况下,即使在对材料无法准备足够量的数据的情况下,也能够提高多成分物质的分析精度。

[0080] 在另一方式所涉及的信息处理系统中,材料可以是聚合物或单体,多成分物质可以是聚合物合金。在该情况下,即使在对聚合物或单体无法准备足够量的数据的情况下,也能够提高聚合物合金的分析精度。聚合物或单体非常多样,相对于此,聚合物合金的种类也庞大。对于这样的聚合物、单体及聚合物合金,一般只能对可取的组合的一部分进行实验,因此大多数情况下无法获得足够量的数据。根据本方式,即使在这样数据不充分的情况下,也能够高精度地分析聚合物合金。

[0081] [变形例]

[0082] 以上,根据其实施方式对本发明详细地进行说明。然而,本发明并不限定于上述实施方式。本发明能够在不脱离其主旨的范围内进行各种变形。

[0083] 由至少一个处理器执行的信息处理方法的处理步骤并不限定于上述实施方式中的例子。例如,可以省略上述步骤或处理的一部分,也可以以其他顺序执行各步骤。并且,可以组合上述步骤中的任意2个以上的步骤,也可以修正或删除步骤的一部分。或者,除了上述各步骤以外,还可以执行其他步骤。

[0084] 在信息处理系统内比较两个数值的大小关系时,可以使用“以上”及“大于”这两个基准中的任一个,也可以使用“以下”及“小于”这两个基准中的任一个。这样的基准的选择不变更关于比较两个数值的大小关系的处理的技术意义。

[0085] 在本发明中,“至少一个处理器执行第1处理,执行第2处理,……执行第n处理。”的表述、或与其对应的表述表示包含执行从第1处理到第n处理的n个处理的处理器在中途改变的情况的概念。即,该表述表示包含n个处理全部由相同的处理器执行的情况和在n个处理

中处理器以任意方针改变的情况这两种情况的概念。

[0086] 符号说明

[0087] 10-信息处理系统,11-获取部,12-计算部,13-预测部。

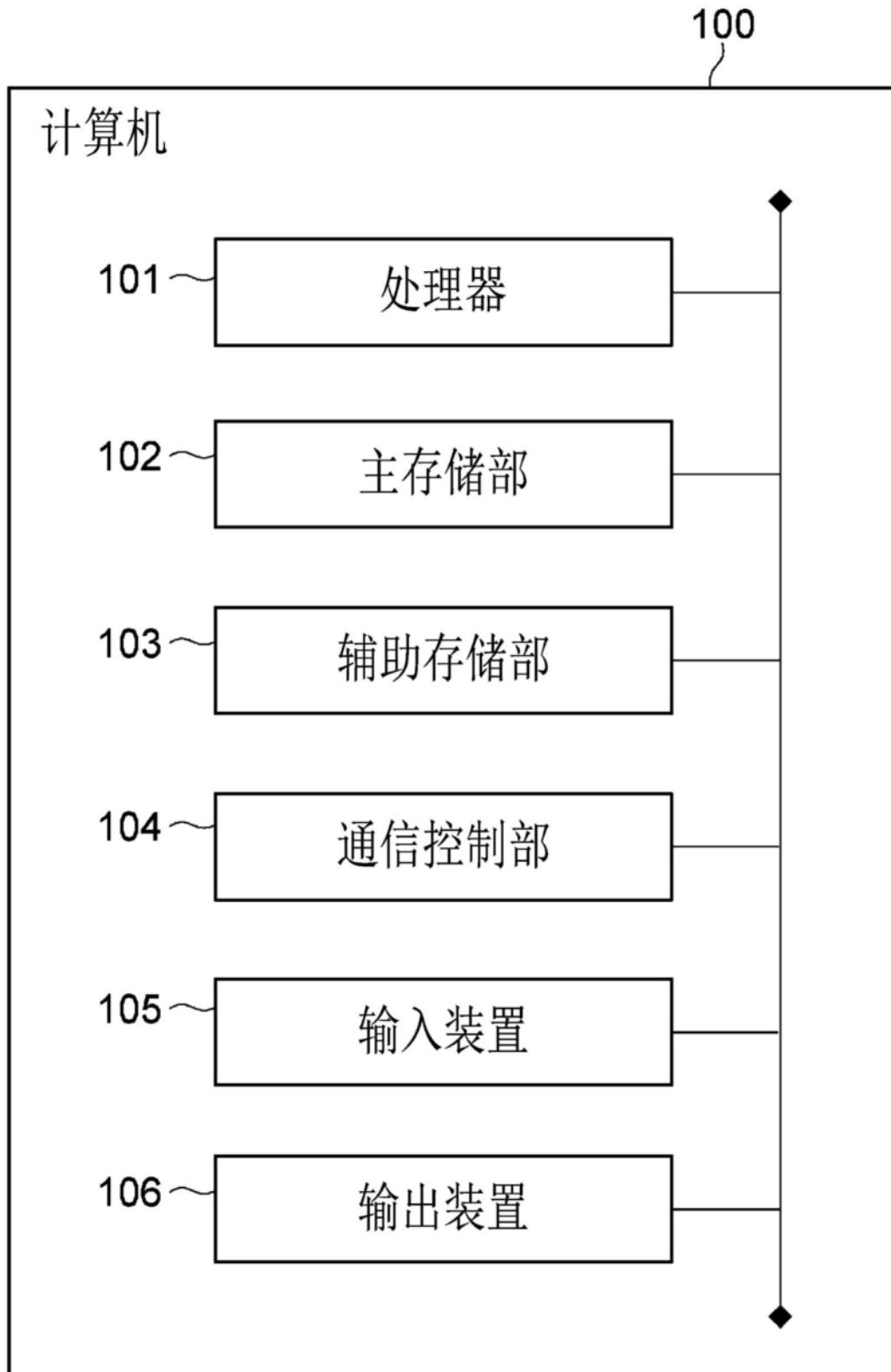


图1

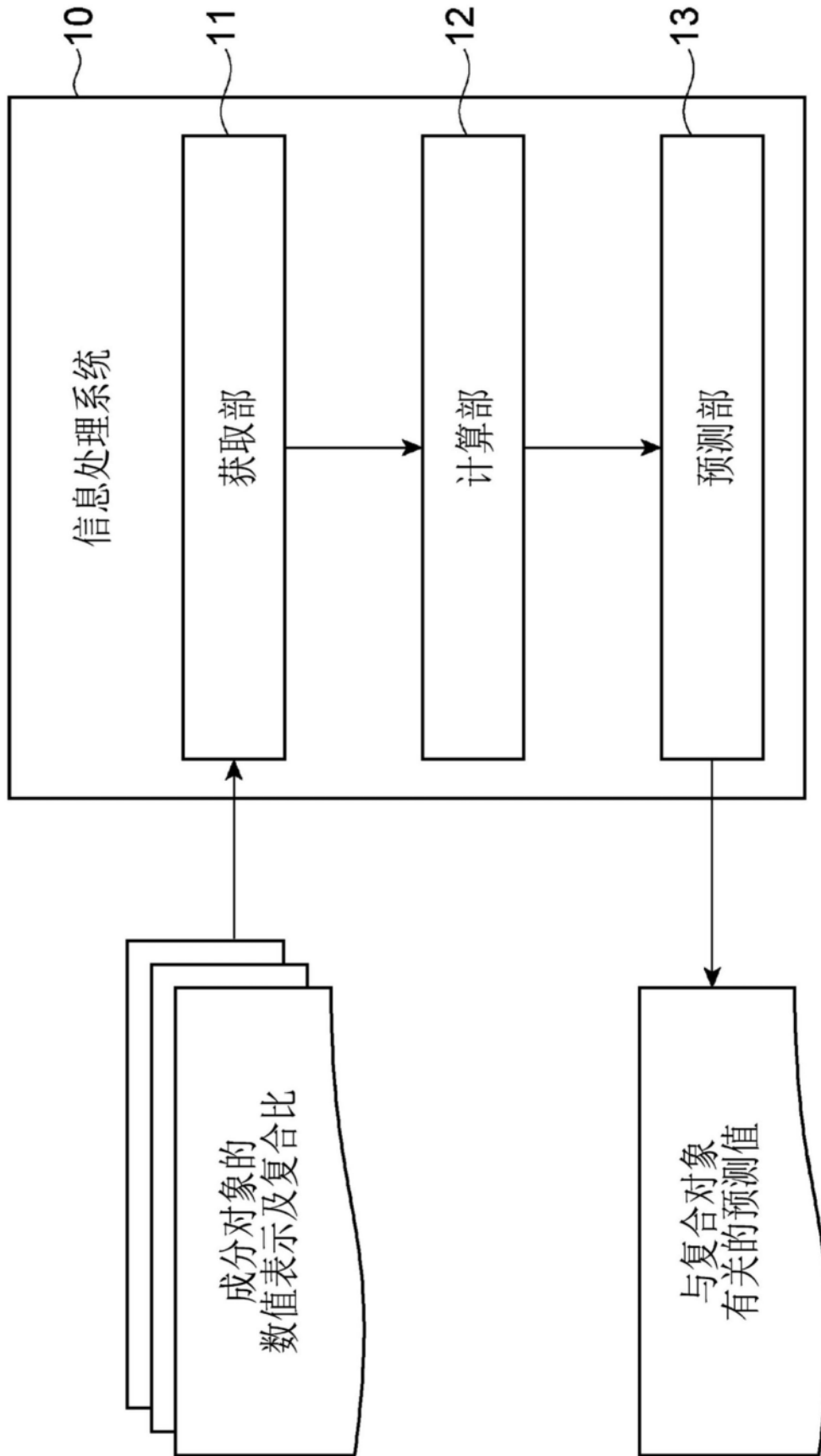


图2

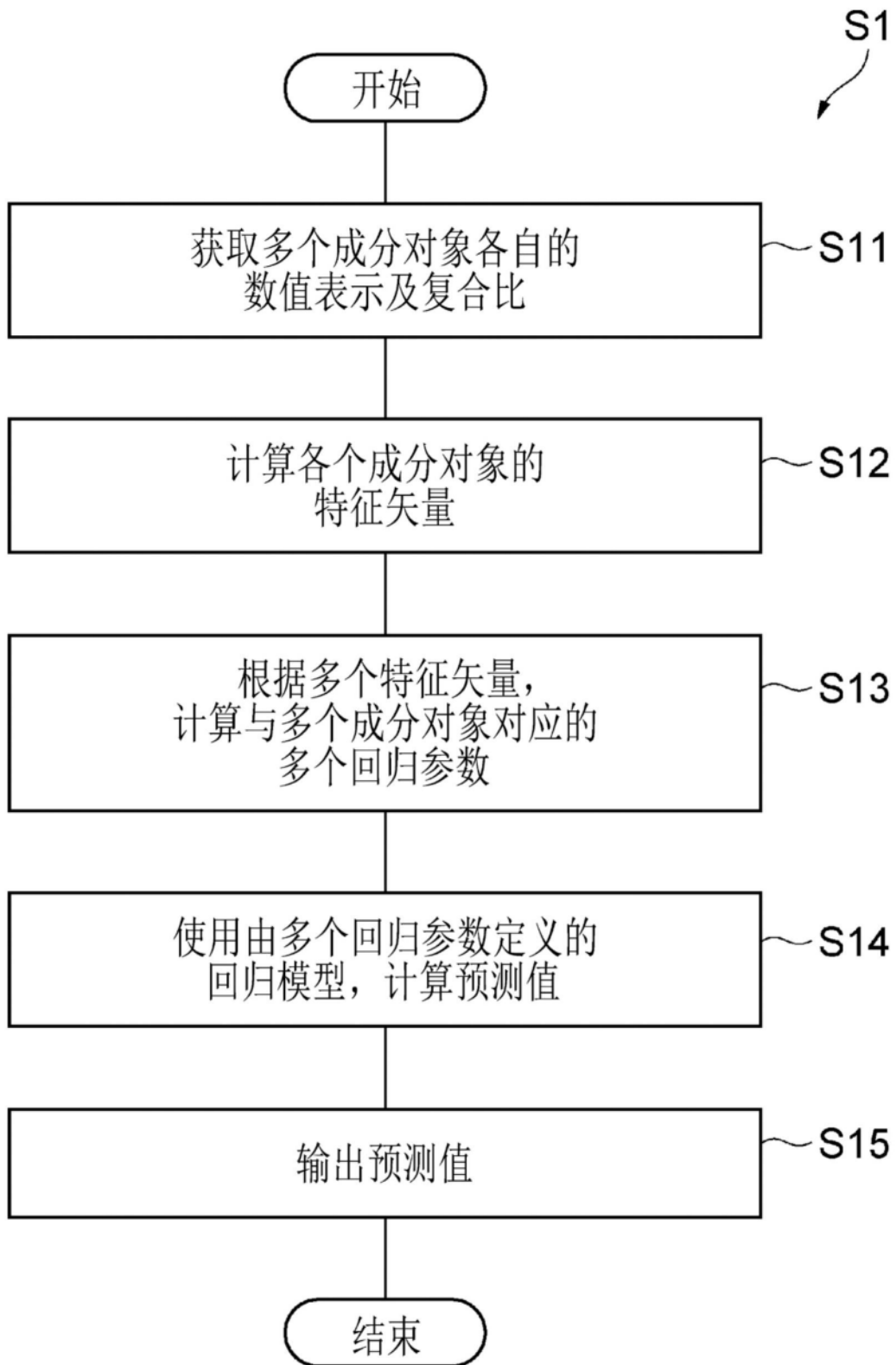


图3

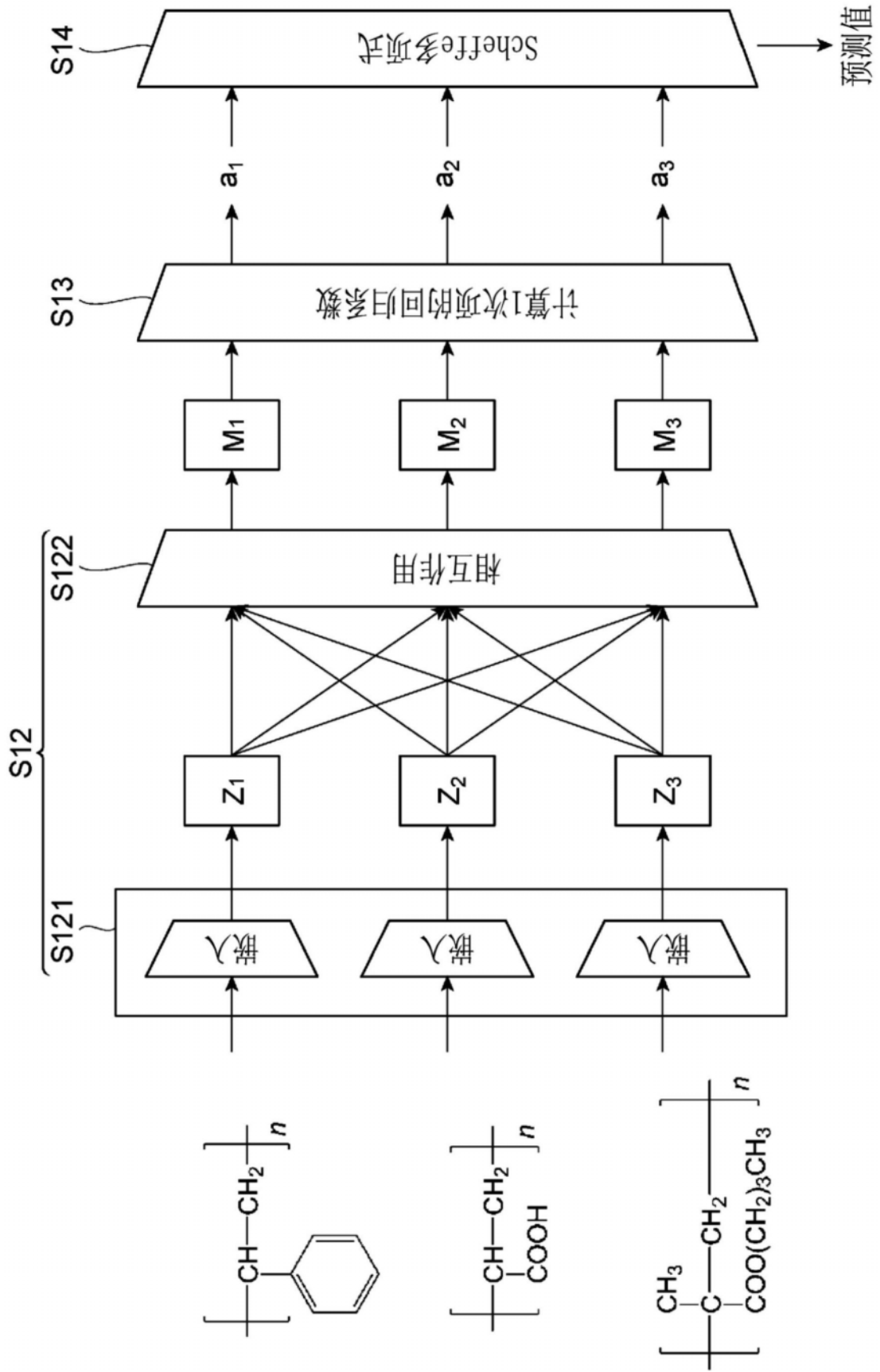


图4

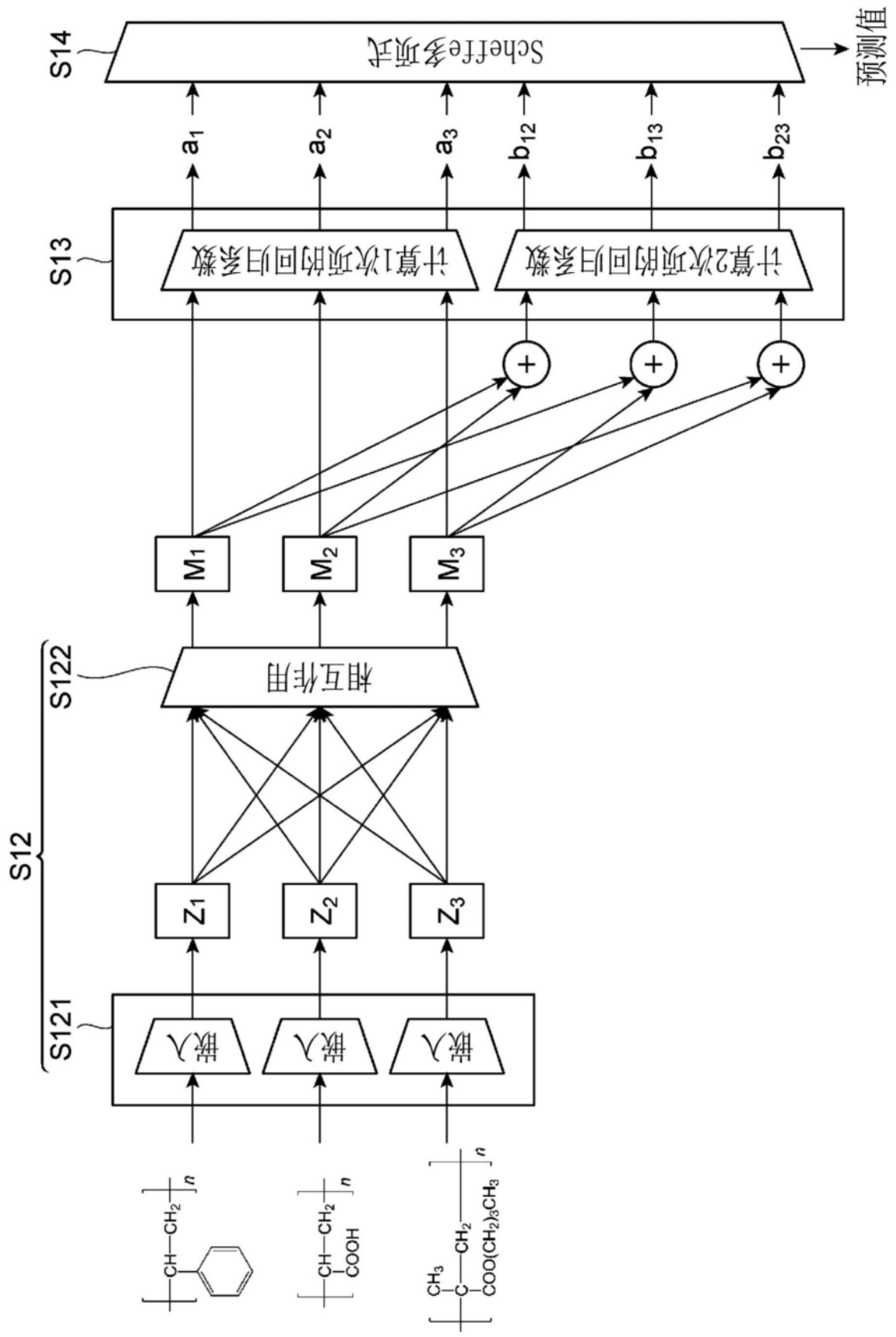


图5