

(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
2024 年 11 月 28 日 (28.11.2024)



(10) 国际公布号
WO 2024/239284 A1

- (51) 国际专利分类号:
G16B 20/30 (2019.01) **G06T 7/33** (2017.01)
- (21) 国际申请号: PCT/CN2023/096109
- (22) 国际申请日: 2023 年 5 月 24 日 (24.05.2023)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (71) 申请人: 深圳华大生命科学研究院 (BGI SHENZHEN) [CN/CN]; 中国广东省深圳市盐田区北山工业区综合楼, Guangdong 518083 (CN)。
- (72) 发明人: 夏天祎 (XIA, Tianyi); 中国广东省深圳市盐田区北山工业区综合楼, Guangdong 518083 (CN)。刘超 (LIU, Chao); 中国广东省深圳市盐田区北山工业区综合楼, Guangdong 518083 (CN)。李美 (LI, Mei); 中国广东省深圳市盐田区北山工业区综合楼, Guangdong 518083 (CN)。张勇 (ZHANG, Yong); 中国广东省深圳市盐田区北山工业区综合楼, Guangdong 518083 (CN)。黎宇翔

(LI, Yuxiang); 中国广东省深圳市盐田区北山工业区综合楼, Guangdong 518083 (CN)。

- (74) 代理人: 北京法胜知识产权代理有限公司 (FASHENG INTELLECTUAL PROPERTY COMPANY, LTD.); 中国北京市海淀区悦秀路99号1单元4层401, Beijing 100085 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(54) Title: SPATIOTEMPORAL TRANSCRIPTOMICS SLICE ALIGNMENT METHOD AND APPARATUS

(54) 发明名称: 时空转录组切片的配准方法及装置

获取两张相邻切片, 每张切片包含至少一种类型的表达量矩阵, 每个表达量矩阵中包含多个时空组表达量数据的位点spot, 每张切片上包含至少一个类型的spot 101

↓

分别确定每张切片中同一类型的spot对应的权重, 其中, 同一类型的spot分配相同的权重, 且不同类型的spot分配的权重不同 102

↓

根据所述每一类型spot对应的权重及预设正则化系数, 分别计算概率传输矩阵, 得到两张切片上每一类别对应的表达量矩阵中每个spot的配准分数 103

图 1

101 Obtain two adjacent slices, wherein each slice comprises at least one type of expression level matrixes, each expression level matrix comprises spots of a plurality of pieces of spatiotemporal transcriptomics expression level data, and each slice comprises at least one type of spots

102 Respectively determine weights corresponding to the same type of spots in each slice, wherein the same weight is assigned to the same type of spots, and weights assigned to different types of spots are different

103 According to the weights corresponding to each type of spots and a preset regularization coefficient, respectively calculate probability transition matrixes to obtain an alignment score of each spot in the expression level matrix corresponding to each category on the two slices

(57) Abstract: The present disclosure provides a spatiotemporal transcriptomics slice alignment method and apparatus. The main technical solution comprises: obtaining two adjacent slices, wherein each slice comprises at least one type of expression level matrixes, and each expression level matrix comprises spots of a plurality of pieces of spatiotemporal transcriptomics expression level data; respectively determining weights corresponding to the same type of spots in each slice, wherein the same weight is assigned to the same type of spots, and weights assigned to different types of spots are different (102); and according to the weight corresponding to each type of spots and a preset regularization coefficient, respectively calculating probability transition matrixes to obtain an alignment score of each spot in the expression level matrix corresponding to each category on the two slices (103). Compared with the related art, the same weight is assigned to the same type of spots in each slice, and the weights assigned to different types of spots are different, so that the accuracy of the probability transition matrixes obtained on the basis of weight calculation is improved, and the accuracy of alignment realized on the basis of the probability transition matrixes is thus improved.

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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: 本公开提供一种时空转录组切片的配准方法及装置, 主要技术方案包括: 获取两张相邻切片, 每张切片包含至少一种类型的表达量矩阵, 每个表达量矩阵中包含多个时空组表达量数据的位点spot; 分别确定每张切片中同一类型的spot对应的权重, 其中, 同一类型的spot分配相同的权重, 且不同类型的spot分配的权重不同(102); 根据每一类型spot对应的权重及预设正则化系数, 分别计算概率传输矩阵, 得到两张切片上每一类别对应的表达量矩阵中每个spot的配准分数(103)。与相关技术相比, 为每张切片中同一类型的spot分配相同的权重, 且不同类型的spot分配的权重不同, 提高了基于权重计算得到的概率转移矩阵的准确性, 进而提高了基于概率转移矩阵所实现的配准的准确性。

时空转录组切片的配准方法及装置

5 技术领域

本公开涉及数据处理技术领域，尤其涉及一种时空转录组切片的配准方法及装置。

背景技术

空间转录组实验的概率配准 (Probabilistic Alignment of Spatial Transcriptomics Experiments, PASTE), 是一个对来及相邻或相近切片的时空组表达量数据的位点 (也称 spot), spot 既可以表示细胞, 也可以表示和细胞大小接近的、人工于时空芯片位点上划分的 n 个位点宽乘以 n 个位点高的表达量检测区域之间进行配准的算法。

该 PASTE 算法中每一个 spot 点具有相同的权重, 且每两片时空表达量数据中, 结构信息和特征信息的在目标函数中的占比是预先设定的、无法自动调节的。

发明内容

本公开提供了一种时空转录组切片的配准方法及装置、电子设备和存储介质。

根据本公开的第一方面, 提供了一种时空转录组切片的配准方法, 包括:

20 获取两张相邻切片, 每张切片包含至少一种类型的表达量矩阵, 每个表达量矩阵中包含多个时空组表达量数据的位点 spot, 每张切片上包含至少一个类型的 spot;

分别确定每张切片中同一类型的 spot 对应的权重, 其中, 同一类型的 spot 分配相同的权重;

25 根据所述每一类型 spot 对应的权重及预设正则化系数, 分别计算概率传输矩阵, 得到两张切片上每一类别对应的表达量矩阵中每个 spot 的配准分数。

可选地, 在获取两张相邻切片之前, 所述方法还包括:

去掉两张相邻切片中的不共有的 spot。

可选地, 所述分别确定每张切片中同一类型的 spot 对应的权重包括:

30 对于每张切片中同一类型的 spot, 分别将对应的表达量矩阵沿着 spot 点进行加和, 分别得到不同类型对应的表达量矩阵;

计算每一类型 spot 分别在所述两张切片上的表达量矩阵的距离, 并根据所述距离确定每一类型 spot 对应的权重。

可选地, 在对于每张切片中同一类型的 spot, 分别将对应的表达量矩阵沿着 spot 点进行加和, 分别得到不同类型对应的表达量矩阵之前, 所述方法还包括:

35 对每张切片中的非高变基因进行过滤, 保留每张切片中的高变基因。

- 可选地，在根据所述距离确定每一类型 spot 对应的权重之前，所述方法还包括：
查找大于预设距离阈值的距离值；
调用预设处理函数，将所述大于所述预设距离阈值的距离值进行压缩处理；
将压缩处理后的距离值与小于所述预设距离阈值的距离值，作为所述 spot 分别在
- 5 所述两张切片上的表达量矩阵的最终距离；
所述根据所述距离确定每一类型 spot 对应的权重包括：
计算所述最终距离的相反数，根据相反数计算结果确定每一类型 spot 对应的权重。
可选地，在所述根据所述每一类型 spot 对应的权重及预设正则化系数，分别计算
概率传输矩阵，得到两张切片上每一类别对应的表达量矩阵中每个 spot 的配准分数
- 10 之后，所述方法还包括：
查找所述概率传输矩阵中小于预设阈值的概率，并将所述小于预设阈值的概率设置
为 0；
将所述概率传输矩阵中的每个概率，除以所述概率传输矩阵之和，得到两张切片
上每一类别对应的表达量矩阵中每个 spot 的配准分数。
- 15 可选地，所述根据所述每一类型 spot 对应的权重及预设正则化系数，分别计算概
率传输矩阵包括：
基于不同的预设正则化系数，分别计算对应的概率传输矩阵；
从不同的概率传输矩阵中选择一个准确性分数高于预设分数阈值的概率矩阵作为
最终的概率传输矩阵。
- 20 第二方面，还提供一种时空转录组切片的配准装置，包括：
获取单元，用于获取两张相邻切片，每张切片包含至少一种类型的表达量矩阵，
每个表达量矩阵中包含多个时空组表达量数据的位点 spot，每张切片上包含至少一个
类型的 spot；
确定单元，用于分别确定每张切片中同一类型的 spot 对应的权重，其中，同一类
型的 spot 分配相同的权重；
- 25 配准单元，用于根据所述每一类型 spot 对应的权重及预设正则化系数，分别计算
概率传输矩阵，得到两张切片上每一类别对应的表达量矩阵中每个 spot 的配准分数。
可选地，所述装置还包括：
处理单元，用于在获取两张相邻切片之前，去掉两张相邻切片中的不共有的 spot。
- 30 可选地，所述确定单元，包括：
加和模块，用于对于每张切片中同一类型的 spot，分别将对应的表达量矩阵沿着
spot 点进行加和，分别得到不同类型对应的表达量矩阵；
计算模块，用于计算每一类型 spot 分别在所述两张切片上的表达量矩阵的距离；
第一确定模块，用于根据所述距离确定每一类型 spot 对应的权重。
- 35 可选地，所述确定单元还包括：
过滤模块，用于在对于每张切片中同一类型的 spot，分别将对应的表达量矩阵沿

着 spot 点进行加和，分别得到不同类型对应的表达量矩阵之前，对每张切片中的非高变基因进行过滤，保留每张切片中的高变基因。

可选地，所述确定单元还包括：

查找模块，用于在根据所述距离确定每一类型 spot 对应的权重之前，查找大于预设距离阈值的距离值；

处理模块，用于调用预设处理函数，将所述大于所述预设距离阈值的距离值进行压缩处理；

第二确定模块，用于将压缩处理后的距离值与小于所述预设距离阈值的距离值，作为所述 spot 分别在所述两张切片上的表达量矩阵的最终距离；

所述第一确定模块，还用于计算所述最终距离的相反数，根据相反数计算结果确定每一类型 spot 对应的权重。

可选地，所述装置还包括：

查找单元，用于在所述配准单元根据所述每一类型 spot 对应的权重及预设正则化系数，分别计算概率传输矩阵，得到两张切片上每一类别对应的表达量矩阵中每个 spot 的配准分数之后，查找所述概率传输矩阵中小于预设阈值的概率；

设置单元，用于将所述小于预设阈值的概率设置为 0；

将所述概率传输矩阵中的每个概率，除以所述概率传输矩阵之和，得到两张切片上每一类别对应的表达量矩阵中每个 spot 的配准分数。

可选地，所述配准单元包括：

基于不同的预设正则化系数，分别计算对应的概率传输矩阵；

从不同的概率传输矩阵中选择一个准确性分数高于预设分数阈值的概率矩阵作为最终的概率传输矩阵。

根据本公开的第三方面，提供了一种电子设备，包括：

至少一个处理器；以及

与所述至少一个处理器通信连接的存储器；其中，

所述存储器存储有可被所述至少一个处理器执行的指令，所述指令被所述至少一个处理器执行，以使所述至少一个处理器能够执行前述第一方面所述的方法。

根据本公开的第四方面，提供了一种存储有计算机指令的非瞬时计算机可读存储介质，其中，所述计算机指令用于使所述计算机执行前述第一方面所述的方法。

根据本公开的第五方面，提供了一种计算机程序产品，包括计算机程序，所述计算机程序在被处理器执行时实现如前述第一方面所述的方法。

本公开提供一种时空转录组切片的配准方法，获取两张相邻切片，每张切片包含至少一种类型的表达量矩阵，每个表达量矩阵中包含多个时空组表达量数据的位点 spot，每张切片上包含至少一个类型的 spot；分别确定每张切片中同一类型的 spot 对应的权重，其中，同一类型的 spot 分配相同的权重；根据所述每一类型 spot 对应的权重及预设正则化系数，分别计算概率传输矩阵，得到两张切片上每一类别对应的

表达量矩阵中每个 spot 的配准分数。与相关技术相比，本公开实施例为每张切片中同一类型的 spot 分配相同的权重，且不同类型的 spot 分配的权重不同，提高了基于权重计算得到的概率转移矩阵的准确性，进而提高了基于概率转移矩阵所实现的配准的准确性。

- 5 应当理解，本部分所描述的内容并非旨在标识本申请的实施例的关键或重要特征，也不用于限制本申请的范围。本申请的其它特征将通过以下的说明书而变得容易理解。

附图说明

附图用于更好地理解本方案，不构成对本公开的限定。其中：

- 10 图 1 为本公开实施例所提供的一种时空转录组切片的配准方法的流程示意图；
图 2 为本公开实施例所提供的另一种时空转录组切片的配准方法的流程示意图；
图 3 为本公开实施例所提供的一种确认每张切片中同一类型的 spot 对应的权重方法的流程示意图；
15 图 4 为本公开实施例提供的一种时空转录组切片的配准装置的结构示意图；
图 5 为本公开实施例提供的另一种时空转录组切片的配准装置的结构示意图；
图 6 为本公开实施例提供的示例电子设备的示意性框图。

具体实施方式

- 20 以下结合附图对本公开的示范性实施例做出说明，其中包括本公开实施例的各种细节以助于理解，应当将它们认为仅仅是示范性的。因此，本领域普通技术人员应当认识到，可以对这里描述的实施例做出各种改变和修改，而不会背离本公开的范围和精神。同样，为了清楚和简明，以下的描述中省略了对公知功能和结构的描述。

- 25 下面参考附图描述本公开实施例的时空转录组切片的配准方法及装置、电子设备和存储介质。

图 1 为本公开实施例所提供的一种时空转录组切片的配准方法的流程示意图。

如图 1 所示，该方法包含以下步骤：

- 30 步骤 101，获取两张相邻切片，每张切片包含至少一种类型的表达量矩阵，每个表达量矩阵中包含多个时空组表达量数据的位点 spot，每张切片上包含至少一个类型的 spot。

- 需要说明的是，本申请实施例所述的方法，并非是指仅适用于两张相邻切片之间的配准，而是在一次配准计算过程中，需要对两张切片进行两两配准，例如，需要进行配准计算的切片有 5 张，当前计算的为切片 1 和相邻的切片 2 之间的配准，在执行完切片 1 和切片 2 的配准后，继续执行切片 2 和相邻的切片 3 之间的配准，再执行切片 3 和相邻的切片 4 之间的配准，最后执行切片 4 和相邻的切片 5 之间配准。以上示例仅为便于给出的数量说明，在实际应用中对切片配准的数量不进行限定。
- 35

需要说明的是，本申请实施例所述的表达量数据具备聚类或注释结果，以辅助算法的实施；事实上，出于后期分析和展示的需要，绝大部分做 spot 配准的数据均具备聚类或注释结果。所述 spot 既可以表示细胞，也可以表示和细胞大小接近的、人工于时空芯片位点上划分的 n 个位点宽乘以 n 个位点高的表达量检测区域。

5 步骤 102，分别确定每张切片中同一类型的 spot 对应的权重，其中，同一类型的 spot 分配相同的权重。

本申请实施例中，对于相同类型的 spot 具有相同的权重值，该值取决于该类型在一张切片上的表达量和在另一张切片上的表达量的相似度，即相似度约大，权重值越大，相似度越小，权重值越小。

10 在判定类型是否相同时，可以有步骤 101 中表达量数据的聚类结果或注释结果确定。有关聚类算法的说明，本申请实施例在此不再进行赘述。步骤 103，根据所述每一类型 spot 对应的权重及预设正则化系数，分别计算概率传输矩阵，得到两张切片上每一类别对应的表达量矩阵中每个 spot 的配准分数。

15 为提升配准算法的准确性，本公开实施例引入了自动化调参机制：在计算中，算法依次采用指数变化的不同的正则化系数，分别计算概率传输矩阵，并得到其对应的准确性分数。在轮流使用过所有规定参与计算的正则化系数后，选择准确性分数最高的概率传输矩阵，作为两张切片上每一类别对应的表达量矩阵中每个 spot 的配准分数。

20 本公开提供一种时空转录组切片的配准方法，获取两张相邻切片，每张切片包含至少一种类型的表达量矩阵，每个表达量矩阵中包含多个时空组表达量数据的位点 spot，每张切片上包含至少一个类型的 spot；分别确定每张切片中同一类型的 spot 对应的权重，其中，同一类型的 spot 分配相同的权重，且不同类型的 spot 分配的权重不同；根据所述每一类型 spot 对应的权重及预设正则化系数，分别计算概率传输矩阵，得到两张切片上每一类别对应的表达量矩阵中每个 spot 的配准分数。与相关技术相比，本公开实施例为每张切片中同一类型的 spot 分配相同的权重，且不同类型的 spot 分配的权重不同，提高了基于权重计算得到的概率转移矩阵的准确性，进而提高了基于概率转移矩阵所实现的配准的准确性。

本申请实施例还提供另一种时空转录组切片的配准方法的流程示意图。

如图 2 所示，所述方法还包括：

30 步骤 201，去掉两张相邻切片中的不共具的 spot。

本申请实施例中增加了去掉不共具的 spot 类型的操作。将两片表达量数据间，只在一片上出现的 spot 进行去除，防止其成为噪音，干扰概率传输矩阵的计算。

35 步骤 202，获取两张相邻切片，每张切片包含至少一种类型的表达量矩阵，每个表达量矩阵中包含多个时空组表达量数据的位点 spot，每张切片上包含至少一个类型的 spot；

有关步骤 202 可参见步骤 101 的详细说明。

步骤 203, 分别确定每张切片中同一类型的 spot 对应的权重, 其中, 同一类型的 spot 分配相同的权重, 且不同类型的 spot 分配的权重不同;

在确认每张切片中同一类型的 spot 对应的权重时, 可以采用但不局限于以下方式, 如图 3 所示, 包括:

5 步骤 2031, 对每张切片中的非高变基因进行过滤, 保留每张切片中的高变基因。

对于 spot 基因表达量矩阵, 选取和保留高变基因, 去掉非高变基因的其他基因。其目的在于削减非高变基因对于计算 spot 类在不同切片上相似性的影响, 使得计算结果更准确。

10 为了便于对高变基因进行理解, 示例性的, 某张切片上, 一共捕获了 10000 种不同的基因, 基因表达量矩阵上记录了每个 spot 上, 每种基因各自被捕获了多少条, 因此, 该基因表达量矩阵的行数等于 spot 个数, 列数等于 10000。为筛选高变基因, 按顺序做: (1) 标准化, (2) 计算每个基因的方差, (3) 排序并筛选这三个步骤。首先, 以 z-score 标准化为例介绍标准化: 基因表达量矩阵上, 对每一行, 计算该行上的所有值的平均值和标准差, 让每个值都分别减去该平均值, 再除以该标准差; 每一行均做完此计算, 就完成了步骤 (1), 注意该矩阵的值由此被更新; 之后, 对于 15 这 10000 个基因, 计算每个基因的列上的所有数据的方差, 作为该基因的方差, 就完成了步骤 (2); 最后, 对于这 10000 个基因, 根据 (2) 计算出的方差将它们进行从大到小的排序, 取排名靠前的若干个, 就得到了高变基因。这 10000 个基因中, 不是高变基因的, 即非高变基因。具体应用该过程中若干个包含但不限于 10、20、50、100、20 200、500 等, 本申请实施例对具体数量不做限定。

步骤 2032, 对于每张切片中同一类型的 spot, 分别将对应的表达量矩阵沿着 spot 点进行加和, 分别得到不同类型对应的表达量矩阵;

对每一张切片上的每一类 spot, 将其表达量矩阵沿着 spot 点进行加和, 得到整个 spot 类上, 不同基因的表达量矩阵。

25 对每一张切片上的每一类 spot, 对得到的表达量矩阵进行对数归一化, 以控制数据的规模效应, 即减少芯片不同捕获位点测序深度不均一, 以及不同 spot 类群 spot 组成数目不均一, 所带来的各个 spot 类表达量矩阵不在同一尺度上的问题。

示例性的, 假设有 4 个 spot 点, spot 点共两类: 1 类和 2 类 (两个点一类), 3 个基因 a,b,c: 统计 1 类上的两个 spot, 总共表达了多少 a 基因, 统计 1 类上的 30 两个 spot, 总共表达了多少 b 基因, 统计 1 类上的两个 spot, 总共表达了多少 c 基因, 统计 2 类上的两个 spot, 总共表达了多少 a 基因, 统计 2 类上的两个 spot, 总共表达了多少 b 基因, 统计 2 类上的两个 spot, 总共表达了多少 c 基因。

在执行归一化处理时, 使用对数归一化对 spot 类的表达量进行规模效应控制, 也可以使用其他规模效应消除/控制的方法, 例如最小-最大归一化等等, 具体的本申请 35 实施例对此不进行限定。

步骤 2033, 计算每一类型 spot 分别在所述两张切片上的表达量矩阵的距离, 并查

找大于预设距离阈值的距离值。

计算每一类 spot 在两张切片上的表达量矩阵的距离, 将此距离值赋给两张切片上对应 spot 的距离值。该距离包含但不限于 KL 散度距离、欧式距离、余弦相似度, 詹森-香农距离 (Jensen-Shannon, JS) 距离, 或其他的距离衡量方式。

- 5 示例性的, 不同切片聚类 and 注释得到的类别信息, 是基于相同的参考系: 比如第一张切片, 有 a, b 两类 spot, 第二张切片, 有 a, b, c 三类 spot, 那么要统计第一张切片的 a 和第二张切片的 a 的距离, 要统计第一张切片的 b 和第二张切片的 b 的距离, 而无需在统计 c 的距离。

- 10 步骤 2034, 调用预设处理函数, 将所述大于所述预设距离阈值的距离值进行压缩处理; 将压缩处理后的距离值与小于所述预设距离阈值的距离值, 作为所述 spot 分别在所述两张切片上的表达量矩阵的最终距离。

- 15 将各 spot 类的距离值经过预设处理函数 (例如 logistic 函数) 的处理, 该函数在一定程度上压缩较大的距离值, 保持较小的距离值, 从而使下一步的权重计算更能突出具有小距离值 spot 类, 即权重比较大的 spot 类之间的距离值差异, 使对配准更重要的 spot 类得到更精细的权重分配; logistic 函数如下所示:

$$\text{dis}_{\text{logistic}} = \frac{1}{(1 + e^{-1 * \text{dis}})} - 0.5$$

需要说明的是, 预设处理函数除了经过 logistic 函数外, 也可以是其他导数在 $[R, +\infty)$ 范围内递减的函数 (其中 R 表示的是任意一个小于正无穷的实数, $+\infty$ 表示正无穷), 例如对数函数等; 具体的, 本公开实施例对此不做限定。

- 20 步骤 2035, 计算所述最终距离的相反数, 根据相反数计算结果确定每一类型 spot 对应的权重。

将经过步骤 2034 处理的距离值计算相反数, 得到权重, 使权重偏移至最小值为 0 后, 使每个权重除以所有权重的和。过程如下:

- 25 $\text{weight} = -1 * \text{dis}_{\text{logistic}}$
 $\text{weight} = \text{weight} - \min(\text{weight})$
 $\text{weight} = \text{weight} / \Sigma \text{weight}$

步骤 204, 根据所述每一类型 spot 对应的权重及预设正则化系数, 分别计算概率传输矩阵, 得到两张切片上每一类别对应的表达量矩阵中每个 spot 的配准分数。

- 30 基于不同的预设正则化系数, 分别计算对应的概率传输矩阵, 从不同的概率传输矩阵中选择一个准确性分数高于预设分数阈值的概率矩阵作为最终的概率传输矩阵。

对于每两片切片, 首先, 计算 M_{AB} , C_1 , C_2 矩阵: M_{AB} 矩阵中存储的是两片之间每两个 spot 的表达量相似性, C_1 矩阵中存储的是用于计算的两片中第一片的所有点之间的距离, C_2 矩阵中存储的是用于计算的两片中第二片的所有点之间的距离。

- 35 对于每两片切片, 首先, 计算 M_{AB} 矩阵, C_1 矩阵, C_2 矩阵, a 向量, b 向量: M_{AB} 矩阵中存储的是两片之间每两个 spot 的表达量相似性, C_1 矩阵中存储的是用于计算的两片中第一片的所有点之间的距离, C_2 矩阵中存储的是用于计算的两片中第二片

的所有点之间的距离, a, b 向量分别存储的是, 这两张切片上不同细胞所对应的权重。

然后, PASTE 算法使用的 FusedGromov-Wasserstein 版的 OT 算法, 会依据上述众矩阵和向量, 以及一个固定的正则化系数 alpha (其本质上是结构信息项和表达量信息项在代价函数中的权重比例), 迭代地计算概率传输矩阵。上述矩阵, 向量和系数统称为输入参数。在 PASTE 算法中, 用户需要指定一个 alpha 值, 然后 OT 算法会依据该固定值, 进行计算。多次迭代之后, 该矩阵中存储的数据趋于稳定。这两片切片的概率传输矩阵计算结束。然后开始计算下一对切片。

在计算概率传输矩阵时, 可以采用但不局限于以下方法, 例如: 将该概率传输矩阵的计算问题形成一个优化问题的数学形式, 包括优化的目标函数, 和约束条件。目标函数为:

$$FGW_{\eta, \alpha}(\mu, \nu) = \min_{\pi \in \Pi(h, g)} E_{\eta}(\mathcal{M}_{AB}, C_1, C_2; \pi) \quad (1)$$

where

$$\begin{aligned} E_{\eta}(\mathcal{M}_{AB}, C_1, C_2, \pi) &= \langle (1-\alpha)\mathcal{M}_{AB}^q + \alpha L(C_1, C_2)^q \otimes \pi, \pi \rangle \\ &= \sum_{i,j,k,l} (1-\alpha)d(a_i, b_j)^q + \alpha[C_1(i, k) - C_2(j, l)]^q \pi_{i,j,k,l} \end{aligned}$$

约束条件为:

$$\Pi(h, g) = \{\pi \in \mathbb{R}_+^{n \times m} \text{ s.t. } \sum_{i=1}^n \pi_{i,j} = h_j, \sum_{j=1}^m \pi_{i,j} = g_i\}$$

以上输入参数中, $\mathcal{M}_{AB}$ 矩阵对应公式中的 M_{AB} , C_1 矩阵对应公式中的 C_1 , C_2 矩阵对应公式中的 C_2 , a 向量对应公式中的 h, b 向量对应公式中的 g。

得到以上数学形式后, 使用 Fused Gromov Wasserstein 的条件梯度算法进行求解。求解的流程如下:

Algorithm 1 Conditional Gradient (CG) for FGW

```

1:  $\pi^{(0)} \leftarrow \mu \times \nu^T$ 
2: for  $i = 1, \dots, \tau$  do
3:    $G \leftarrow$  Gradient from Eq. (7) w.r.t.  $\pi^{(i-1)}$ 
4:    $\tilde{\pi}^{(i)} \leftarrow$  Solve OT with ground loss  $G$ 
5:    $\tau^{(i)} \leftarrow$  Line-search for loss (1) with  $\tau \in (0, 1)$  using Alg. 2
6:    $\pi^{(i)} \leftarrow (1 - \tau^{(i)})\pi^{(i-1)} + \tau^{(i)}\tilde{\pi}^{(i)}$ 
7: end for
```

Algorithm 2 Line-search for CG ($q = 2$)

```

1:  $c_{C_1}, c_{C_2}$  from Eq. (6) in (Frey et al., 2016)
2:  $a = -2\alpha(C_1\tilde{\pi}^{(i)}C_2, \tilde{\pi}^{(i)})$ 
3:  $b = ((1-\alpha)\mathcal{M}_{AB} + \alpha c_{C_1}C_2, \tilde{\pi}^{(i)}) - 2\alpha((C_1\tilde{\pi}^{(i)}C_2, \pi^{(i-1)}) + (C_1\pi^{(i-1)}C_2, \tilde{\pi}^{(i)}))$ 
4:  $c = E_2(\mathcal{M}_{AB}, C_1, C_2, \pi^{(i-1)})$ 
5: if  $a > 0$  then
6:    $\tau^{(i)} \leftarrow \min(1, \max(0, \frac{-a}{2c}))$ 
7: else
8:    $\tau^{(i)} \leftarrow 1$  if  $a + b < 0$  else  $\tau^{(i)} \leftarrow 0$ 
9: end if
```

其中, Eq. (7) 为:

$$G = (1-\alpha)\mathcal{M}_{AB}^q + 2\alpha L(C_1, C_2)^q \otimes \pi$$

Eq. (6)为:

$$\pi^* = \arg \min_{\pi \in \Pi(h, g)} \text{vec}(\pi)^T Q(\alpha) \text{vec}(\pi) + \text{vec}(D(\alpha))^T \text{vec}(\pi)$$

其中,

$$Q = -2\alpha C_2 \otimes_K C_1 \text{ and } D(\alpha) = (1 - \alpha)M_{AB}$$

- 5 上述输入参数中, M_{AB} 矩阵对应公式中的 M_{AB} , C_1 矩阵对应公式中的 C_1 , C_2 矩阵对应公式中的 C_2 , a 向量对应公式中的 μ_x , b 向量对应公式中的 μ_y 。

10 综上, 为提升算法准确性, 本申请实施例引入了自动化调参机制: 在计算中, 算法依次采用指数变化的, 不同的正则化系数, 分别计算概率传输矩阵, 并得到其对应的准确性分数。在轮流使用过所有规定参与计算的正则化系数后, 选择准确性分数最高的概率传输矩阵, 作为计算结果。

步骤 205, 查找所述概率传输矩阵中小于预设阈值的概率, 并将所述小于预设阈值的概率设置为 0。

本步骤也称为概率传输矩阵的后处理过程, 将元素值低于预设阈值的元素值(概率)统一设置为 0,

- 15 预设阈值的计算方法如下:

$$\text{thresh} = \min(\text{pi}[\text{pi} > 0]) + (\max(\text{pi}) - \min(\text{pi}[\text{pi} > 0])) / 10$$

20 实际应用中, 对概率传输矩阵进行后处理时, 预设阈值的计算不限于上述计算; 也可以是取最大的概率值的某个百分比作为阈值等等。除了对比阈值的方法外, 也可采用其他方法, 如对每个 spot 在概率传输矩阵结果中表示的与之配对的 spot 依据坐标进行聚类, 保留 spot 数量较多的类, 去掉较少的类等等。

步骤 206, 将所述概率传输矩阵中的每个概率, 除以所述概率传输矩阵之和, 得到两张切片上每一类别对应的表达量矩阵中每个 spot 的配准分数。

25 与上述的时空转录组切片的配准方法相对应, 本发明还提出一种时空转录组切片的配准装置。由于本发明的装置实施例与上述的方法实施例相对应, 对于装置实施例中未披露的细节可参照上述的方法实施例, 本发明中不再进行赘述。

图 4 为本公开实施例提供的一种时空转录组切片的配准装置的结构示意图, 如图 4 所示, 包括:

30 获取单元 41, 用于获取两张相邻切片, 每张切片包含至少一种类型的表达量矩阵, 每个表达量矩阵中包含多个时空组表达量数据的位点 spot, 每张切片上包含至少一个类型的 spot;

确定单元 42, 用于分别确定每张切片中同一类型的 spot 对应的权重, 其中, 同一类型的 spot 分配相同的权重, 且不同类型的 spot 分配的权重不同;

35 配准单元 43, 用于根据所述每一类型 spot 对应的权重及预设正则化系数, 分别计算概率传输矩阵, 得到两张切片上每一类别对应的表达量矩阵中每个 spot 的配准分数。

本公开提供一种时空转录组切片的配准装置，获取两张相邻切片，每张切片包含至少一种类型的表达量矩阵，每个表达量矩阵中包含多个时空组表达量数据的位点 spot，每张切片上包含至少一个类型的 spot；分别确定每张切片中同一类型的 spot 对应的权重，其中，同一类型的 spot 分配相同的权重，且不同类型的 spot 分配的权重不同；根据所述每一类型 spot 对应的权重及预设正则化系数，分别计算概率传输矩阵，得到两张切片上每一类别对应的表达量矩阵中每个 spot 的配准分数。与相关技术相比，本公开实施例为每张切片中同一类型的 spot 分配相同的权重，且不同类型的 spot 分配的权重不同，提高了基于权重计算得到的概率转移矩阵的准确性，进而提高了基于概率转移矩阵所实现的配准的准确性。

10 在本实施例的一些方式中，如图 5 所示，所述装置还包括：

处理单元 44，用于在获取两张相邻切片之前，去掉两张相邻切片中的不共具的 spot。

在本实施例的一些方式中，如图 5 所示，所述确定单元 42，包括：

15 加和模块 421，用于对于每张切片中同一类型的 spot，分别将对应的表达量矩阵沿着 spot 点进行加和，分别得到不同类型对应的表达量矩阵；

计算模块 422，用于计算每一类型 spot 分别在所述两张切片上的表达量矩阵的距离；

第一确定模块 423，用于根据所述距离确定每一类型 spot 对应的权重。

在本实施例的一些方式中，如图 5 所示，所述确定单元 42 还包括：

20 过滤模块 424，用于在对于每张切片中同一类型的 spot，分别将对应的表达量矩阵沿着 spot 点进行加和，分别得到不同类型对应的表达量矩阵之前，对每张切片中的非高变基因进行过滤，保留每张切片中的高变基因。

在本实施例的一些方式中，如图 5 所示，所述确定单元 42 还包括：

25 查找模块 425，用于在根据所述距离确定每一类型 spot 对应的权重之前，查找大于预设距离阈值的距离值；

处理模块 426，用于调用预设处理函数，将所述大于所述预设距离阈值的距离值进行压缩处理；

第二确定模块 427，用于将压缩处理后的距离值与小于所述预设距离阈值的距离值，作为所述 spot 分别在所述两张切片上的表达量矩阵的最终距离；

30 所述第一确定模块 423，还用于计算所述最终距离的相反数，根据相反数计算结果确定每一类型 spot 对应的权重。

在本实施例的一些方式中，如图 5 所示，所述装置还包括：

35 查找单元 45，用于在所述配准单元 44 根据所述每一类型 spot 对应的权重及预设正则化系数，分别计算概率传输矩阵，得到两张切片上每一类别对应的表达量矩阵中每个 spot 的配准分数之后，查找所述概率传输矩阵中小于预设阈值的概率；

设置单元 46，用于将所述小于预设阈值的概率设置为 0；

将所述概率传输矩阵中的每个概率，除以所述概率传输矩阵之和，得到两张切片上每一类别对应的表达量矩阵中每个 spot 的配准分数。

在本实施例的一些方式中，如图 5 所示，所述配准单元 44 包括：

基于不同的预设正则化系数，分别计算对应的概率传输矩阵；

- 5 从不同的概率传输矩阵中选择一个准确性分数高于预设分数阈值的概率矩阵作为最终的概率传输矩阵。

需要说明的是，前述对方法实施例的解释说明，也适用于本实施例的装置，原理相同，本实施例中不再限定。

- 10 根据本公开的实施例，本公开还提供了一种电子设备、一种可读存储介质和一种计算机程序产品。

- 图 6 示出了可以用来实施本公开的实施例的示例电子设备 500 的示意性框图。电子设备旨在表示各种形式的数字计算机，诸如，膝上型计算机、台式计算机、工作台、个人数字助理、服务器、刀片式服务器、大型计算机、和其它适合的计算机。电子设备还可以表示各种形式的移动装置，诸如，个人数字处理、蜂窝电话、智能电话、可穿戴设备和其它类似的计算装置。本文所示的部件、它们的连接和关系、以及它们的功能仅作为示例，并且不意在限制本文中描述的和/或者要求的本公开的实现。

- 20 如图 6 所示，设备 500 包括计算单元 501，其可以根据存储在 ROM (Read-Only Memory, 只读存储器) 502 中的计算机程序或者从存储单元 508 加载到 RAM (Random Access Memory, 随机访问/存取存储器) 503 中的计算机程序，来执行各种适当的操作和处理。在 RAM 503 中，还可存储设备 500 操作所需的各种程序和数据。计算单元 501、ROM 502 以及 RAM 503 通过总线 504 彼此相连。I/O (Input/Output, 输入/输出) 接口 505 也连接至总线 504。

- 25 设备 500 中的多个部件连接至 I/O 接口 505，包括：输入单元 506，例如键盘、鼠标等；输出单元 507，例如各种类型的显示器、扬声器等；存储单元 508，例如磁盘、光盘等；以及通信单元 509，例如网卡、调制解调器、无线通信收发机等。通信单元 509 允许设备 500 通过诸如因特网的计算机网络和/或各种电信网络与其他设备交换信息/数据。

- 30 计算单元 501 可以是各种具有处理和计算能力的通用和/或专用处理组件。计算单元 501 的一些示例包括但不限于 CPU (Central Processing Unit, 中央处理单元)、GPU (Graphic Processing Units, 图形处理单元)、各种专用的 AI (Artificial Intelligence, 人工智能) 计算芯片、各种运行机器学习模型算法的计算单元、DSP (Digital Signal Processor, 数字信号处理器)、以及任何适当的处理器、控制器、微控制器等。计算单元 501 执行上文所描述的各个方法和处理，例如时空转录组切片的配准方法。例如，在一些实施例中，时空转录组切片的配准方法可被实现为计算机软件程序，其被有形地包含于机器可读介质，例如存储单元 508。在一些实施例中，计算机程序的部分或者全部可以经由 ROM 502 和/或通信单元 509 而被载入和/或安装

到设备 500 上。当计算机程序加载到 RAM 503 并由计算单元 501 执行时，可以执行上文描述的方法的一个或多个步骤。备选地，在其他实施例中，计算单元 501 可以通过其他任何适当的方式（例如，借助于固件）而被配置为执行前述时空转录组切片的配准方法。

5 本文中以上描述的系统和技术各种实施方式可以在数字电子电路系统、集成电路系统、FPGA（Field Programmable Gate Array，现场可编程门阵列）、ASIC（Application-Specific Integrated Circuit，专用集成电路）、ASSP（Application Specific Standard Product，专用标准产品）、SOC（System On Chip，芯片上系统的系统）、CPLD（Complex Programmable Logic Device，复杂可编程逻辑设备）、
10 计算机硬件、固件、软件、和/或它们的组合中实现。这些各种实施方式可以包括：实施在一个或者多个计算机程序中，该一个或者多个计算机程序可在包括至少一个可
编程处理器的可编程系统上执行和/或解释，该可编程处理器可以是专用或者通用可
编程处理器，可以从存储系统、至少一个输入装置、和至少一个输出装置接收数据和
15 指令，并且将数据和指令传输至该存储系统、该至少一个输入装置、和该至少一个输出装置。

用于实施本公开的方法的程序代码可以采用一个或多个编程语言的任何组合来编写。这些程序代码可以提供给通用计算机、专用计算机或其他可编程数据处理装置的处理器或控制器，使得程序代码当由处理器或控制器执行时使流程图和/或框图中所规定的功能/操作被实施。程序代码可以完全在机器上执行、部分地在机器上执行，
20 作为独立软件包部分地在机器上执行且部分地在远程机器上执行或完全在远程机器或服务器上执行。

在本公开的上下文中，机器可读介质可以是有形的介质，其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。机器可读介质可以是机器可读信号介质或机器可读储存介质。机器可读介质可以包括
25 但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置或设备，或者上述内容的任何合适组合。机器可读存储介质的更具体示例会包括基于一个或多个线的电气连接、便携式计算机盘、硬盘、RAM、ROM、EPROM（Electrically Programmable Read-Only-Memory，可擦除可编程只读存储器）或快闪存储器、光纤、CD-ROM（Compact Disc Read-Only Memory，便捷式紧凑盘只读存储器）、光学储存设备、磁储存设备、
30 或上述内容的任何合适组合。

为了提供与用户的交互，可以在计算机上实施此处描述的系统和技术，该计算机具有：用于向用户显示信息的显示装置（例如，CRT（Cathode-Ray Tube，阴极射线管）或者 LCD（LiquidCrystal Display，液晶显示器）监视器）；以及键盘和指向装置（例如，鼠标或者轨迹球），用户可以通过该键盘和该指向装置来将输入提供给计算机。其它种类的装置还可以用于提供与用户的交互；例如，提供给用户的反馈可以是任何形式的传感反馈（例如，视觉反馈、听觉反馈、或者触觉反馈）；并且可以用
35

任何形式（包括声输入、语音输入或者、触觉输入）来接收来自用户的输入。

可以将此处描述的系统和技术的实施在包括后台部件的计算系统（例如，作为数据服务器）、或者包括中间件部件的计算系统（例如，应用服务器）、或者包括前端部件的计算系统（例如，具有图形用户界面或者网络浏览器的用户计算机，用户可以通过该图形用户界面或者该网络浏览器来与此处描述的系统和技术的实施方式交互）、或者包括这种后台部件、中间件部件、或者前端部件的任何组合的计算系统中。可以通过任何形式或者介质的数字数据通信（例如，通信网络）来将系统的部件相互连接。通信网络的示例包括：LAN（Local Area Network，局域网）、WAN（Wide Area Network，广域网）、互联网和区块链网络。

计算机系统可以包括客户端和服务端。客户端和服务端一般远离彼此并且通常通过通信网络进行交互。通过在相应的计算机上运行并且彼此具有客户端-服务器关系的计算机程序来产生客户端和服务端的关系。服务器可以是云服务器，又称为云服务器或云主机，是云计算服务体系中的一项主机产品，以解决了传统物理主机与 VPS 服务（“Virtual Private Server”，或简称“VPS”）中，存在的管理难度大，业务扩展性弱的缺陷。服务器也可以为分布式系统的服务器，或者是结合了区块链的服务器。

其中，需要说明的是，人工智能是研究使计算机来模拟人的某些思维过程和智能行为（如学习、推理、思考、规划等）的学科，既有硬件层面的技术也有软件层面的技术。人工智能硬件技术一般包括如传感器、专用人工智能芯片、云计算、分布式存储、大数据处理等技术；人工智能软件技术主要包括计算机视觉技术、语音识别技术、自然语言处理技术以及机器学习/深度学习、大数据处理技术、知识图谱技术等几大方向。

应该理解，可以使用上面所示的各种形式的流程，重新排序、增加或删除步骤。例如，本公开中记载的各步骤可以并行地执行也可以顺序地执行也可以不同的次序执行，只要能够实现本公开公开的技术方案所期望的结果，本文在此不进行限制。

上述具体实施方式，并不构成对本公开保护范围的限制。本领域技术人员应该明白的是，根据设计要求和因素，可以进行各种修改、组合、子组合和替代。任何在本公开的精神和原则之内所作的修改、等同替换和改进等，均应包含在本公开保护范围之内。

权 利 要 求 书

1. 一种时空转录组切片的配准方法，其特征在于，包括：

5 获取两张相邻切片，每张切片包含至少一种类型的表达量矩阵，每个表达量矩阵中包含多个时空组表达量数据的位点 spot，每张切片上包含至少一个类型的 spot；
分别确定每张切片中同一类型的 spot 对应的权重，其中，同一类型的 spot 分配相同的权重；

根据所述每一类型 spot 对应的权重及预设正则化系数，分别计算概率传输矩阵，得到两张切片上每一类别对应的表达量矩阵中每个 spot 的配准分数。

10 2. 根据权利要求 1 所述的方法，其特征在于，在获取两张相邻切片之前，所述方法还包括：

去掉两张相邻切片中的不共有的 spot。

3. 根据权利要求 1 所述的方法，其特征在于，所述分别确定每张切片中同一类型的 spot 对应的权重包括：

15 对于每张切片中同一类型的 spot，分别将对应的表达量矩阵沿着 spot 点进行加和，分别得到不同类型对应的表达量矩阵；

计算每一类型 spot 分别在所述两张切片上的表达量矩阵的距离，并根据所述距离确定每一类型 spot 对应的权重。

20 4. 根据权利要求 3 所述的方法，其特征在于，在对于每张切片中同一类型的 spot，分别将对应的表达量矩阵沿着 spot 点进行加和，分别得到不同类型对应的表达量矩阵之前，所述方法还包括：

对每张切片中的非高变基因进行过滤，保留每张切片中的高变基因。

5. 根据权利要求 3 所述的方法，其特征在于，在根据所述距离确定每一类型 spot 对应的权重之前，所述方法还包括：

25 查找大于预设距离阈值的距离值；

调用预设处理函数，将所述大于所述预设距离阈值的距离值进行压缩处理；

将压缩处理后的距离值与小于所述预设距离阈值的距离值，作为所述 spot 分别在所述两张切片上的表达量矩阵的最终距离；

所述根据所述距离确定每一类型 spot 对应的权重包括：

30 计算所述最终距离的相反数，根据相反数计算结果确定每一类型 spot 对应的权重。

6. 根据权利要求 1-5 中任一项所述的方法，其特征在于，在所述根据所述每一类型 spot 对应的权重及预设正则化系数，分别计算概率传输矩阵，得到两张切片上每一类别对应的表达量矩阵中每个 spot 的配准分数之后，所述方法还包括：

35 查找所述概率传输矩阵中小于预设阈值的概率，并将所述小于预设阈值的概率设置为 0；

将所述概率传输矩阵中的每个概率，除以所述概率传输矩阵之和，得到两张切片上每一类别对应的表达量矩阵中每个 spot 的配准分数。

7. 根据权利要求 6 所述的方法，其特征在于，所述根据所述每一类型 spot 对应的权重及预设正则化系数，分别计算概率传输矩阵包括：

5 基于不同的预设正则化系数，分别计算对应的概率传输矩阵；

从不同的概率传输矩阵中选择一个准确性分数高于预设分数阈值的概率矩阵作为最终的概率传输矩阵。

8. 一种时空转录组切片的配准装置，其特征在于，包括：

10 获取单元，用于获取两张相邻切片，每张切片包含至少一种类型的表达量矩阵，每个表达量矩阵中包含多个时空组表达量数据的位点 spot，每张切片上包含至少一个类型的 spot；

确定单元，用于分别确定每张切片中同一类型的 spot 对应的权重，其中，同一类型的 spot 分配相同的权重；

15 配准单元，用于根据所述每一类型 spot 对应的权重及预设正则化系数，分别计算概率传输矩阵，得到两张切片上每一类别对应的表达量矩阵中每个 spot 的配准分数。

9. 一种电子设备，其特征在于，包括：

至少一个处理器；以及

与所述至少一个处理器通信连接的存储器；其中，

20 所述存储器存储有可被所述至少一个处理器执行的指令，所述指令被所述至少一个处理器执行，以使所述至少一个处理器能够执行权利要求 1-7 中任一项所述的方法。

10. 一种存储有计算机指令的非瞬时计算机可读存储介质，其特征在于，所述计算机指令用于使所述计算机执行根据权利要求 1-7 中任一项所述的方法。

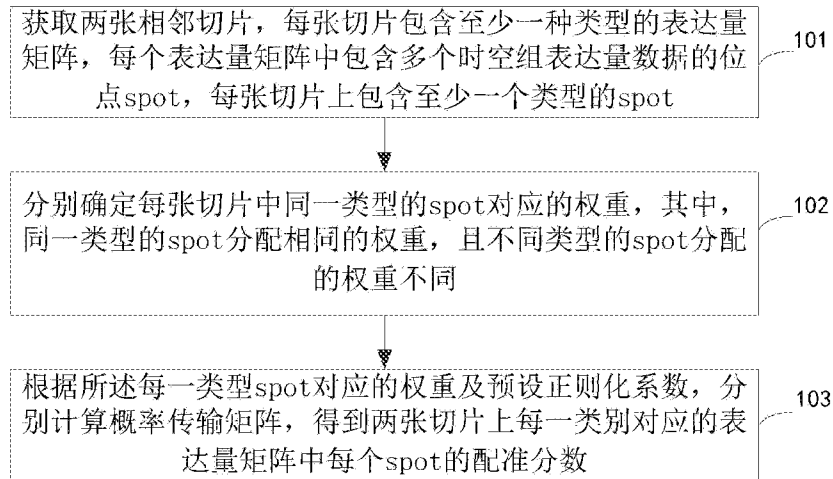


图 1

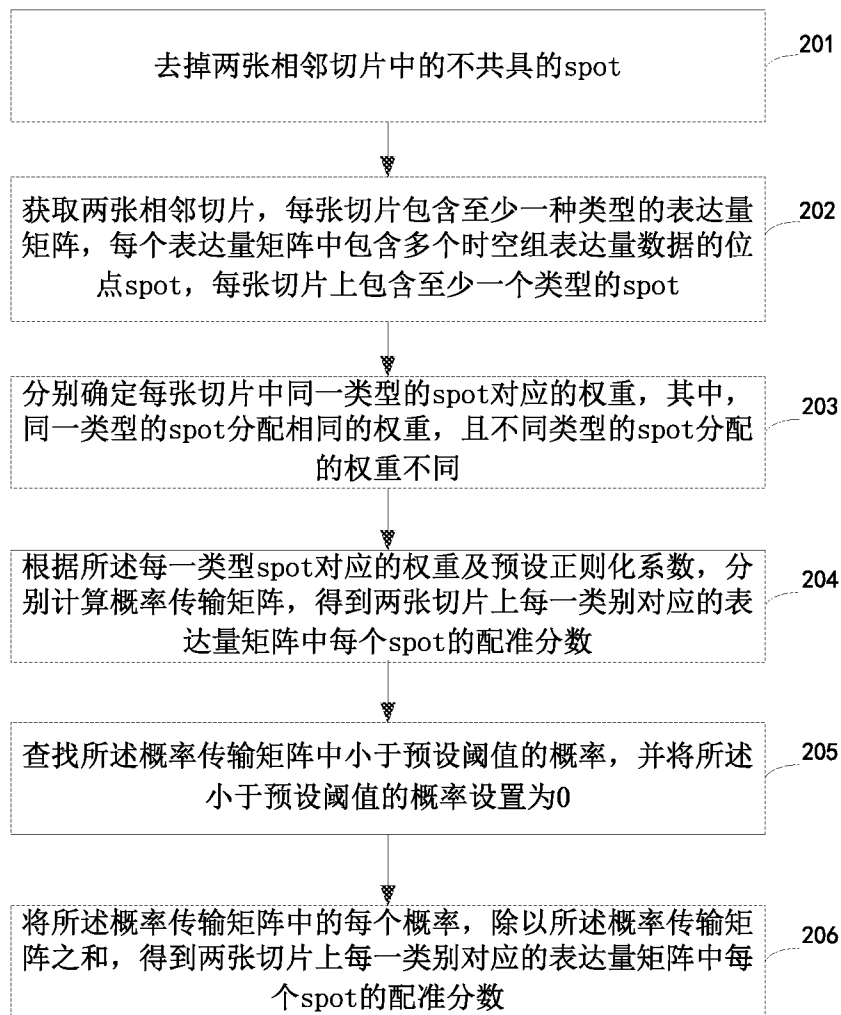


图 2

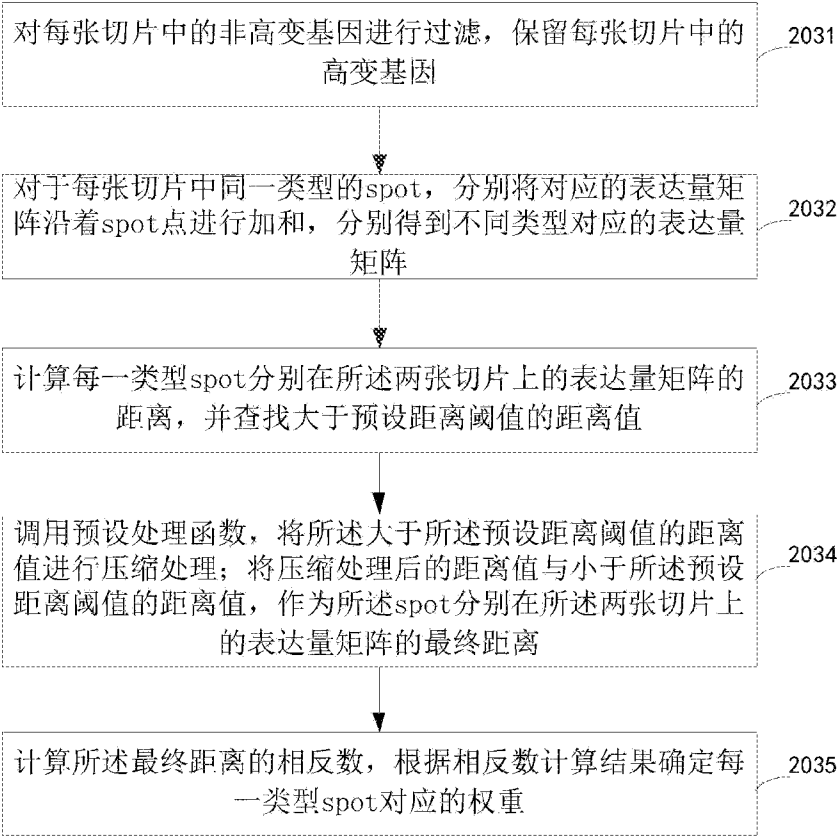


图 3

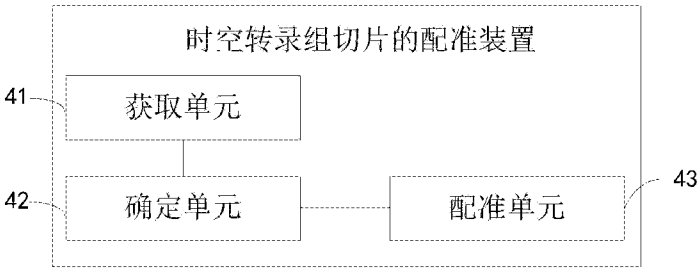


图 4



图 5

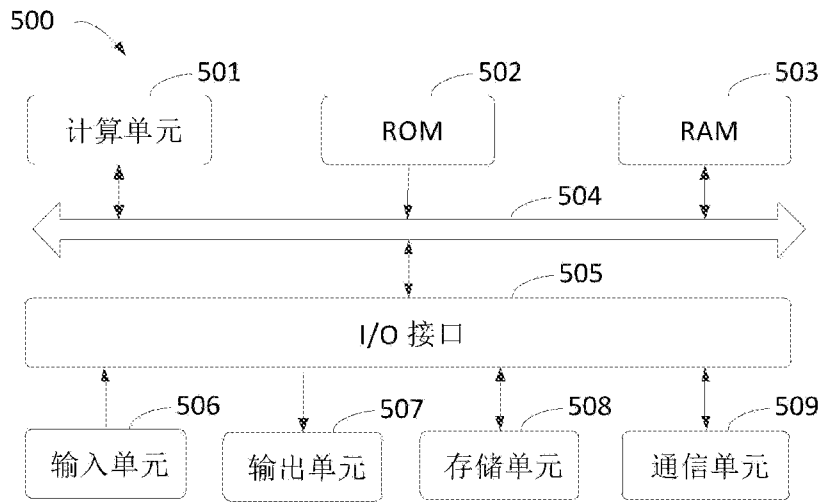


图 6

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/096109

A. CLASSIFICATION OF SUBJECT MATTER

G16B20/30(2019.01)i; G06T7/33(2017.01)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)

IPC: G16B, G06T

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; DWPI; WPABS; ENTXT; VEN; CNKI; IEEE: 切片, 配准, 对准, 校准, 矩阵, 向量, 权重, 权值, 系数, match +, calibrat+, matrix??. array?, weight?, coefficient?.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 114387319 A (BEIJING BAIDU NETCOM SCIENCE AND TECHNOLOGY CO., LTD.) 22 April 2022 (2022-04-22) description, paragraphs 0022-0118	1-10
A	CN 107194959 A (BEIJING HAZHI WANGJU INFORMATION TECHNOLOGY CO., LTD.) 22 September 2017 (2017-09-22) entire document	1-10
A	CN 107545567 A (INSTITUTE OF AUTOMATION, CHINESE ACADEMY OF SCIENCES) 05 January 2018 (2018-01-05) entire document	1-10
A	CN 113436238 A (HUBEI ECARX TECHNOLOGY CO., LTD.) 24 September 2021 (2021-09-24) entire document	1-10
A	US 2021035297 A1 (THE JOAN AND IRWIN JACOBS TECHNION-CORNELL INSTITUTE) 04 February 2021 (2021-02-04) entire document	1-10



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

“D” document cited by the applicant in the international application

“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

19 December 2023

Date of mailing of the international search report

20 January 2024

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
China No. 6, Xitucheng Road, Jimenqiao, Haidian District,
Beijing 100088

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2023/096109

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	114387319	A	22 April 2022	CN 114387319 B	14 November 2023
CN	107194959	A	22 September 2017	None	
CN	107545567	A	05 January 2018	CN 107545567 B	19 May 2020
CN	113436238	A	24 September 2021	CN 113436238 B	23 November 2021
US	2021035297	A1	04 February 2021	US 11830192 B2	28 November 2023
				WO 2021021430 A1	04 February 2021

A. 主题的分类

G16B20/30(2019.01)i; G06T7/33(2017.01)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)

IPC: G16B, G06T

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;DWPI;WPABS;ENTXT;VEN;CNKI;IEEE:切片, 配准, 对准, 校准, 矩阵, 向量, 权重, 权值, 系数, match+, calibrat+, matrix?-, array?, weight?, coefficient?。

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
X	CN 114387319 A (北京百度网讯科技有限公司) 2022年4月22日 (2022 - 04 - 22) 说明书第0022-0118段	1-10
A	CN 107194959 A (北京海致网聚信息技术有限公司) 2017年9月22日 (2017 - 09 - 22) 全文	1-10
A	CN 107545567 A (中国科学院自动化研究所) 2018年1月5日 (2018 - 01 - 05) 全文	1-10
A	CN 113436238 A (湖北亿咖通科技有限公司) 2021年9月24日 (2021 - 09 - 24) 全文	1-10
A	US 2021035297 A1 (THE JOAN AND IRWIN JACOBS TECHNION CORNELL INST) 2021年2月4日 (2021 - 02 - 04) 全文	1-10

☐ 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

“D” 申请人在国际申请中引证的文件

“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

2023年12月19日

国际检索报告邮寄日期

2024年1月20日

ISA/CN的名称和邮寄地址

中国国家知识产权局
中国北京市海淀区蓟门桥西土城路6号 100088

授权官员

陈敏

电话号码 (+86) 027-59371775

PCT/ISA/210 表(第2页) (2022年7月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2023/096109

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	114387319	A	2022年4月22日	CN 114387319 B	2023年11月14日
CN	107194959	A	2017年9月22日	无	
CN	107545567	A	2018年1月5日	CN 107545567 B	2020年5月19日
CN	113436238	A	2021年9月24日	CN 113436238 B	2021年11月23日
US	2021035297	A1	2021年2月4日	US 11830192 B2	2023年11月28日
				WO 2021021430 A1	2021年2月4日