

(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 2024/239503 A1

(51) 国际专利分类号:

G16B 20/30 (2019.01)

号自贸试验区大兴机场片区自贸创新服务中心一层0002号, Beijing 102615 (CN)。

(21) 国际申请号:

PCT/CN2023/120768

(22) 国际申请日:

2023 年 9 月 22 日 (22.09.2023)

(25) 申请语言:

中文

(26) 公布语言:

中文

(30) 优先权:

PCT/CN2023/096109

2023年5月24日 (24.05.2023) CN

(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)。北京华大生命科学研究院 (BGI RESEARCH BEIJING) [CN/CN]; 中国北京市大兴区礼贤镇元平北路1

(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)。

(54) **Title:** REGISTRATION METHOD AND APPARATUS FOR SPATIO-TEMPORAL TRANSCRIPTOME SLICES, ELECTRONIC DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 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]

获取目标切片每个目标spot, 分别在相邻切片上对应的概率传输矩阵中, 概率转移值最大的spot的索引向量

101



基于概率传输矩阵, 对所述目标切片进行刚性配准, 得到配准后的目标切片

102



根据所述索引向量对应的相邻切片上的spot点坐标, 及所述目标spot的坐标, 计算所述配准后的目标切片在所述相邻切片上的形变场

103



在相邻两张切片的所述形变场中分别对应查找到所述坐标的变化值, 并利用所述变化值对所述目标切片各spot的坐标进行配准

104

101 Acquire each target spot of a target slice, and respectively acquire, in a probabilistic transfer matrix corresponding to an adjacent slice, an index vector of a spot having the maximum probability transfer value

102 Perform rigid registration on the target slice on the basis of the probabilistic transfer matrix to obtain a registered target slice

103 Calculate a deformation field of the registered target slice on the adjacent slice according to coordinates of the spot on the adjacent slice corresponding to the index vector and coordinates of the target spot

104 Respectively find change values of coordinates in deformation fields of two adjacent slices, and register the coordinates of each spot of the target slice by using the change values

(57) **Abstract:** Disclosed are a registration method and apparatus for spatio-temporal transcriptome slices, an electronic device, and a storage medium. The method comprises: acquiring each target spot of a target slice, and respectively acquiring, in a probabilistic transfer matrix corresponding to an adjacent slice, an index vector of a spot having the maximum probability transfer value; performing rigid registration on the target slice on the basis of the probabilistic transfer matrix to obtain a registered target slice; calculating a deformation field of the registered target slice on the adjacent slice according to coordinates of the spot on the adjacent slice corresponding to the index vector and coordinates of the target spot; and respectively finding change values of coordinates in deformation fields of two adjacent slices, and registering the coordinates of each spot of the target slice by using the change values. After rigid registration is performed on the target slice, the deformation of the slice is registered to restore the real shapes of a plurality of slices in an organism, thereby further improving the accuracy of a registration result.

[见续页]

(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。

(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: 本公开公开了时空转录组切片的配准方法及装置、电子设备和存储介质, 主获取目标切片每个目标spot, 分别在相邻切片上对应的概率传输矩阵中, 概率转移值最大的spot的索引向量; 基于概率传输矩阵, 对目标切片进行刚性配准, 得到配准后的目标切片; 根据索引向量对应的相邻切片上的spot点坐标, 及目标spot的坐标, 计算配准后的目标切片在相邻切片上的形变场; 在相邻两张切片的形变场中分别对应查找到坐标的变化值, 并利用变化值对目标切片各spot的坐标进行配准。在对目标切片进行刚性配准之后, 再通过对切片的形变进行配准, 还原了多张切片在生物体内的真实形状, 进一步提高了配准结果的准确性。

说明书

发明名称: 时空转录组切片的配准 方法及装置、电子设备和存储介质

[0001] 相关技术的交叉引用

[0002] 本申请基于申请号为PCT/CN2023/096109, 申请日为2023年05月24日的PCT专利申请提出, 并要求该PCT专利申请的优先权, 该PCT专利申请的全部内容通过引用结合在本申请中。

技术领域

[0003] 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、电子设备和存储介质。

背景技术

[0004] 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。

[0005] 该PASTE算法基于概率传输矩阵, 计算和实现刚性配准, 但是, 该PASTE算法仅做了刚性配准, 在形状方面并未进行配准。

[0006] 发明内容

[0007] 本公开提供了一种时空转录组切片的配准方法、装置、电子设备和存储介质。其主要目的在于实现通过对时空转录组切片的形变进行配准, 还原了多张切片在生物体内的真实形状, 进一步提高了配准结果的准确性。

[0008] 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, 其中, 包括:

[0009] 获取目标切片每个目标spot, 分别在相邻切片上对应的概率传输矩阵中, 概率转移值最大的spot的索引向量; 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；

- [0010] 基于所述概率传输矩阵，对所述目标切片进行刚性配准，得到配准后的目标切片；
- [0011] 根据所述索引向量对应的相邻切片上的spot点坐标，及所述目标spot的坐标，计算所述配准后的目标切片在所述相邻切片上的形变场；
- [0012] 在相邻两张切片的所述形变场中分别对应查找到所述坐标的变化值，并利用所述变化值对所述目标切片各spot的坐标进行配准。
- [0013] 在一些实施例中，所述相邻切片包括左相邻的第一切片；
- [0014] 所述获取目标切片每个目标spot，分别在相邻切片上对应的概率传输矩阵中，概率转移值最大的spot的索引向量包括：
- [0015] 获取第一概率传输矩阵，所述第一概率传输矩阵为第一切片与所述目标切片计算得到的概率传输矩阵，其中，所述目标切片每个spot对应所述第二概率传输矩阵上的一列，所述第一切片每个spot对应所述第一概率传输矩阵上的一行；
- [0016] 在所述第一概率传输矩阵中查找概率转移值最大的元素所在的行，并将所述概率转移值最大的元素所在的行的行标信息，确定为概率转移值最大的spot的索引向量。
- [0017] 在一些实施例中，所述相邻切片包括右相邻的第二切片；
- [0018] 所述获取目标切片每个目标spot，分别在相邻切片上对应的概率传输矩阵中，概率转移值最大的spot的索引向量包括：
- [0019] 获取第二概率传输矩阵，所述第二概率传输矩阵为所述目标切片与第二切片计算得到的概率传输矩阵，其中，所述第二切片每个spot对应所述第三概率传输矩阵上的一列，所述目标切片每个spot对应所述第二概率传输矩阵上的一行；
- [0020] 在所述第二概率传输矩阵中查找概率转移值最大的元素所在的列，并将所述概率转移值最大的元素所在的列的列标信息，确定为概率转移值最大的spot的索引向量。
- [0021] 在一些实施例中，所述相邻切片包括左相邻的第三切片以及右相邻的第四切片；

- [0022] 所述获取目标切片每个目标spot，分别在相邻切片上对应的概率传输矩阵中，概率转移值最大的spot的索引向量包括：
- [0023] 获取第三概率传输矩阵，所述第三概率传输矩阵为第三切片与所述目标切片计算得到的概率传输矩阵，其中，所述目标切片每个spot对应所述第二概率传输矩阵上的一列，所述第三切片每个spot对应所述第三概率传输矩阵上的一行；
- [0024] 在所述第三概率传输矩阵中查找概率转移值最大的元素所在的行，并将所述概率转移值最大的元素所在的行的行标信息，确定为概率转移值最大的spot的索引向量；
- [0025] 获取第四概率传输矩阵，所述第四概率传输矩阵为所述目标切片与第四切片计算得到的概率传输矩阵，其中，所述第四切片每个spot对应所述第四概率传输矩阵上的一列，所述目标切片每个spot对应所述第四概率传输矩阵上的一行；
- [0026] 在所述第四概率传输矩阵中查找概率转移值最大的元素所在的列，并将所述概率转移值最大的元素所在的列的列标信息，确定为概率转移值最大的spot的索引向量。在一些实施例中，所述根据所述索引向量对应的相邻切片上的spot点坐标，及所述目标spot的坐标，计算所述配准后的目标切片在所述相邻切片上的形变场包括：
- [0027] 根据所述索引向量从预设坐标矩阵中查找在所述目标切片在所述相邻切片上对应的spot点坐标；其中，所述预设坐标矩阵中第一列对应所述索引向量的横坐标，第二列对应所述索引向量的纵坐标，每行表示一个spot；
- [0028] 基于所述spot点坐标计算在所述相邻切片上的形变场。
- [0029] 在一些实施例中，在在相邻两张切片的所述形变场中分别对应查找到所述坐标的变化值之前，所述方法还包括：
- [0030] 调用预设滤波函数对所述目标切片的所述形变场进行平滑处理，得到滤波后的形变场；
- [0031] 所述在相邻两张切片的所述形变场中分别对应查找到所述坐标的变化值包括：
- [0032] 在相邻两张切片的所述滤波后的形变场中分别对应查找到所述坐标的变化值。
- [0033] 在一些实施例中，在获取目标切片中同一类型的spot对应的第一概率传输矩阵之后，所述方法还包括：

- [0034] 查找所述第一概率传输矩阵中元素值小于预设阈值的目标元素值；
- [0035] 将所有所述目标元素值配置为0。
- [0036] 在一些实施例中，在获取目标切片每个目标spot，分别在相邻切片上对应的概率传输矩阵中，概率转移值最大的spot的索引向量之前，所述方法还包括：
- [0037] 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；
- [0038] 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；
- [0039] 根据所述每一类型spot对应的权重及预设正则化系数，分别计算第一概率传输矩阵，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。
- [0040] 在一些实施例中，在获取两张相邻切片之前，所述方法还包括：
- [0041] 去掉两张相邻切片中的不共有的spot。
- [0042] 根据本公开的第二方面，提供了一种时空转录组切片的配准装置，包括：
- [0043] 第一获取单元，用于获取目标切片每个目标spot，分别在相邻切片上对应的概率传输矩阵中，概率转移值最大的spot的索引向量；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；
- [0044] 第二获取单元，用于获取所述目标切片每个目标spot，分别在相邻切片上对应的第二概率传输矩阵中，概率转移值最大的spot的索引向量；
- [0045] 第一配准单元，用于基于所述第一概率传输矩阵及所述第二概率传输矩阵，对所述目标切片进行刚性配准，得到配准后的目标切片；
- [0046] 第一计算单元，用于根据所述索引向量对应的相邻切片上的spot点坐标，及所述目标spot的坐标，计算所述配准后的目标切片在所述相邻切片上的形变场；
- [0047] 第一查找单元，用于在相邻两张切片的所述形变场中分别对应查找到所述坐标的变化值；
- [0048] 第二配准单元，用于利用所述变化值对所述目标切片各spot的坐标进行配准。
- [0049] 在一些实施例中，所述相邻切片包括左相邻的第一切片；

[0050] 所述第二获取单元还用于：

[0051] 获取第一概率传输矩阵，所述第一概率传输矩阵为第一切片与所述目标切片计算得到的概率传输矩阵，其中，所述目标切片每个spot对应所述第二概率传输矩阵上的一列，所述第一切片每个spot对应所述第一概率传输矩阵上的一行；

[0052] 在所述第一概率传输矩阵中查找概率转移值最大的元素所在的行，并将所述概率转移值最大的元素所在的行的行标信息，确定为概率转移值最大的spot的索引向量。

[0053] 在一些实施例中，所述相邻切片包括右相邻的第二切片；

[0054] 所述第二获取单元还用于：

[0055] 获取第二概率传输矩阵，所述第二概率传输矩阵为所述目标切片与第二切片计算得到的概率传输矩阵，其中，所述第二切片每个spot对应所述第三概率传输矩阵上的一列，所述目标切片每个spot对应所述第二概率传输矩阵上的一行；

[0056] 在所述第二概率传输矩阵中查找概率转移值最大的元素所在的列，并将所述概率转移值最大的元素所在的列的列标信息，确定为概率转移值最大的spot的索引向量。

[0057] 在一些实施例中，所述相邻切片包括左相邻的第三切片以及右相邻的第四切片；

[0058] 所述第二获取单元还用于：

[0059] 获取第三概率传输矩阵，所述第三概率传输矩阵为第三切片与所述目标切片计算得到的概率传输矩阵，其中，所述目标切片每个spot对应所述第二概率传输矩阵上的一列，所述第三切片每个spot对应所述第三概率传输矩阵上的一行；

[0060] 在所述第三概率传输矩阵中查找概率转移值最大的元素所在的行，并将所述概率转移值最大的元素所在的行的行标信息，确定为概率转移值最大的spot的索引向量；

[0061] 获取第四概率传输矩阵，所述第四概率传输矩阵为所述目标切片与第四切片计算得到的概率传输矩阵，其中，所述第四切片每个spot对应所述第四概率传输矩阵上的一列，所述目标切片每个spot对应所述第四概率传输矩阵上的一行；

- [0062] 在所述第四概率传输矩阵中查找概率转移值最大的元素所在的列，并将所述概率转移值最大的元素所在的列的列标信息，确定为概率转移值最大的spot的索引向量。在一些实施例中，所述计算单元还用于：
- [0063] 根据所述索引向量从预设坐标矩阵中查找在所述配准后的目标切片在所述相邻切片上对应的spot点坐标；其中，所述预设坐标矩阵中第一列对应所述索引向量的横坐标，第二列对应所述索引向量的纵坐标，每行表示一个spot；
- [0064] 基于所述spot点坐标计算在所述相邻切片上的形变场。
- [0065] 在一些实施例中，所述装置还包括：
- [0066] 处理单元，用于在所述第一查找单元在相邻两张切片的所述形变场中分别对应查找到所述坐标的变化值之前，调用预设滤波函数对所述目标切片的所述形变场进行平滑处理，得到滤波后的形变场；
- [0067] 所述第一查找单元，还用于在相邻两张切片的所述滤波后的形变场中分别对应查找到所述坐标的变化值。
- [0068] 在一些实施例中，所述装置还包括：
- [0069] 第二查找单元，用于在所述第一获取单元获取目标切片中同一类型的spot对应的第一概率传输矩阵之后，查找所述第一概率传输矩阵中元素值小于预设阈值的目標元素值；
- [0070] 配置单元，用于将所有所述目标元素值配置为0。
- [0071] 在一些实施例中，所述装置还包括：
- [0072] 第三获取单元，用于在获取目标切片每个目标spot，分别在相邻切片上对应的概率传输矩阵中，概率转移值最大的spot的索引向量之前，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；
- [0073] 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；
- [0074] 计算单元，用于根据所述每一类型spot对应的权重及预设正则化系数，分别计算第一概率传输矩阵，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。

[0075] 在一些实施例中，所述装置还包括：

[0076] 去除单元，用于在所述第三获取单元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。

[0077] 根据本公开的第三方面，提供了一种电子设备，包括：

[0078] 至少一个处理器；以及

[0079] 与所述至少一个处理器通信连接的存储器；其中，

[0080] 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。

[0081] 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。

[0082] 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。

[0083] 本公开提供的时空转录组切片的配准方法、装置、电子设备和存储介质，获取目标切片中同一类型的spot对应的第一概率传输矩阵，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，获取所述目标切片每个目标spot，分别在相邻切片上对应的第二概率传输矩阵中，概率转移值最大的spot的索引向量，根据所述索引向量对应的相邻切片上的spot点坐标，及所述目标spot的坐标，基于概率传输矩阵，对所述目标切片进行刚性配准，得到配准后的目标切片，计算所述配准后的目标切片在所述相邻切片上的形变场，在相邻两张切片的所述形变场中分别对应查找到所述坐标的变化值，并利用所述变化值对所述目标切片各spot的坐标进行配准。在对目标切片进行刚性配准之后，再通过对切片的形变进行配准，还原了多张切片在生物体内的真实形状，进一步提高了配准结果的准确性。

[0084] 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

[0085] 附图用于更好地理解本方案，不构成对本公开的限定。其中：

[0086] 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；

[0087] 图2为本公开实施例所提供的一种受力导致的形变的示意图；

[0088] 图3为本申请实施例提供的一种经滤波处理后的形变场示意图；

[0089] 图4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；

[0090] 图5为本申请实施例提供的一种果蝇幼虫在进行配准前后的示意图；

[0091] 图6为本申请实施例提供的一种未执行配准前所有细胞占据面积的变化情况示意图；

[0092] 图7为本申请实施例提供的一种执行配准前所有细胞占据面积的变化情况示意图；

[0093] 图8示出了本申请实施例与相关技术的比对示意图；

[0094] 图9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；

[0095] 图10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；

[0096] 图11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

[0097] 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。

[0098] 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、装置、电子设备和存储介质。

[0099] 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。

[0100] 如图1所示,该方法包含以下步骤:

[0101] 步骤101,获取目标切片每个目标spot,分别在相邻切片上对应的概率传输矩阵中,概率转移值最大的spot的索引向量。

[0102] 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。

[0103] 按照切片的顺序依次计算目标切片与相邻切片之间的概率传输矩阵,例如第一片和第二片之间进行计算,然后第二片和第三片之间进行计算……等等,当前进行计算的切片为目标切片。概率传输矩阵中存储的是目标切片与前一个相邻切片之间每个spot的表达量相似性,或者,第一概率传输矩阵中存储的是目标切片与后一个相邻切片之间每个spot的表达量相似性。

[0104] 针对目标切片,找到其中每一个spot,在该目标切片的相邻切片上的对应概率转移矩阵上,概率转移值最大的spot点。

[0105] 本申请实施例中,相邻切片包含但不限于目标切片的前一张切片和后一张切片,可以是与目标切片最相近的两张相邻切片:例如应用在相邻或相近的时空转录组切片的配准任务中时,如切片的切片顺序为S1、S2、S3,那么S2作为目标切片的前一张切片和后一张切片分别为S1和S3。

[0106] 作为本申请实施例的另一实施例,如果目标切片位于末位,没有后一张切片,那么相邻切片为前一张切片,因此只计算前一张切片(即S1)对应spot点;如果位于首位,目标切片没有前一张切片,那么只计算后一张切片(即S3)对应的spot点。

[0107] 实际应用中,以相邻切片包含前一张切片和后一张切片为例,N为目标切片上spot的个数, N_{pre} 为前一张切片spot的个数, N_{next} 为后一张切片spot的个数; $M^{(pre)} \in R^{N_{pre} \times N}$,为前一张切片和目标切片计算得到的概率转移矩阵; $M^{(next)} \in R^{N \times N_{next}}$,为目标切片和后一张切片计算得到的概率转移矩阵。

[0108] 计算 $I_{pre} \in Z^N$,为描述目标切片在前一张切片上各对应spot的索引的向量;

$$I_{pre}^n \in \arg\max_{i \in \{0, \dots, N_{pre}-1\}} M_{i,n}^{(pre)}$$

[0109] 计算 $I_{next} \in Z^N$,为描述目标切片在后一张切片上各对应spot的索引的向量;

$$I_{\text{next}}^n \in \arg\max_{j \in \{0, \dots, N_{\text{next}}-1\}} M_{n,j}^{(\text{next})}$$

[0110] 实际应用中，可以是取最大的概率值的某个百分比(不限于某个固定的百分比数值)作为阈值等等。除了使用阈值的方法外，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。具体的，本申请实施例对此不做限定。

[0111] 步骤102，基于概率传输矩阵，对所述目标切片进行刚性配准，得到配准后的目标切片。

[0112] 基于上述步骤获取的概率传输矩阵，基于方法Procrustes Analysis进行刚性配准。具体内容为：对按顺序排列的若干张切片，首先将目标切片和前一张切片的中心进行对齐，然后根据如下公式将目标切片的角度进行旋转，使目标切片对齐于前一张切片。

$$\hat{\theta} = \arctan\left(\frac{\text{Tr}\left(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} W \Pi^T Z^T\right)}{\text{Tr}(W \Pi^T Z^T)}\right)$$

[0113] 其中 Π 是第二输出的概率传输矩阵， $W \in \mathbb{R}^{n_2 \times 2}$ 是目标切片的所有坐标， $Z \in \mathbb{R}^{n_1 \times 2}$ 是前一张切片的所有坐标。

[0114] 以上完成了目标切片向前一张切片的刚性配准。然后，将后一张切片配准到配准后的目标切片上，以此类推，直到遍历所有切片。

[0115] 需要说明的是，上述实例是以目标切片与前一张切片为例进行的说明，应当明确，该种说明方式仅为了示例，而并非用以限定。

[0116] 步骤103，根据所述索引向量对应的相邻切片上的spot点坐标，及所述目标spot的坐标，计算所述配准后的目标切片在所述相邻切片上的形变场。

[0117] 目标切片在相邻切片上得到的形变场，由目标切片在相邻切片上对应的spot点的坐标，以及目标切片上spot点坐标，计算得到。

[0118] 为用数学形式描述该过程，定义： X 为目标切片上所有点的坐标，其中第一列存储x坐标，第二列存储y坐标，每行表示一个spot，顺序和 $M^{(\text{pre})}$ 中列方向，以及 $M^{(\text{next})}$ 对应的行方向上的spot顺序相一致；定义 $X^{(\text{pre})}$ 为前一片上所有点的坐标，其中第一列存储x坐标，第二列存储y坐标，每行表示一个spot，顺序和 $M^{(\text{pre})}$ 中行方

向对应的spot顺序相一致；定义 $X^{(next)}$ 为后一片上所有点的坐标，其中第一列存储x坐标，第二列存储y坐标，每行表示一个spot，顺序和 $M^{(next)}$ 中列方向对应的spot顺序相一致。

- [0119] psize为对应形变场上同一像素的两点的最远横、纵距离；该值和切片上两点的平均距离接近即可。
- [0120] 步骤104，在相邻两张切片的所述形变场中分别对应查找到所述坐标的变化值，并利用所述变化值对所述目标切片各spot的坐标进行配准。
- [0121] 基于每片切片各spot的坐标，在其前后形变场中，分别对应查找到坐标的变化值，并将前后变化值的平均值应用在目标切片各spot的坐标上。
- [0122] 本公开提供的时空转录组切片的配准方法，获取目标切片中同一类型的spot对应的第一概率传输矩阵，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，获取所述目标切片每个目标spot，分别在相邻切片上对应的第二概率传输矩阵中，概率转移值最大的spot的索引向量，根据所述索引向量对应的相邻切片上的spot点坐标，及所述目标spot的坐标，计算所述配准后的目标切片在所述相邻切片上的形变场，在相邻两张切片的所述形变场中分别对应查找到所述坐标的变化值，并利用所述变化值对所述目标切片各spot的坐标进行配准。通过对切片的形变进行配准，还原了多张切片在生物体内的真实形状，进一步提高了配准结果的准确性。
- [0123] 示例一，当所述相邻切片包括左相邻的第一切片，所述获取目标切片每个目标spot，分别在相邻切片上对应的概率传输矩阵中，概率转移值最大的spot的索引向量包括：获取第一概率传输矩阵，所述第一概率传输矩阵为第一切片与所述目标切片计算得到的概率传输矩阵，其中，所述目标切片每个spot对应所述第二概率传输矩阵上的一列，所述第一切片每个spot对应所述第一概率传输矩阵上的一行；在所述第一概率传输矩阵中查找概率转移值最大的元素所在的行，并将所述概率转移值最大的元素所在的行的行标信息，确定为概率转移值最大的spot的索引向量。。

[0124] 呈由上述示例，在找目标切片S2片上每一个spot，在第一切片(S1)上对应的spot点时，首先取得S1和S2输出的第一概率转移矩阵M12；S2上的每一个spot，对应M12上的一列。每一列对应若干个元素，而元素值最大的元素所在的行，对应的是该spot在S1上，对应第一概率转移矩阵上，概率转移值最大的spot点，将其确定为概率转移值最大的spot的索引向量。

[0125] 实际应用中，以相邻切片包含第一切片为例，N为目标切片上spot的个数， N_{pre} 为第一切片spot的个数， $M^{(pre)} \in \mathbb{R}^{N_{pre} \times N}$ ，为第一切片和目标切片计算得到的第一概率转移矩阵；计算 $I_{pre} \in \mathbb{Z}^N$ ，为描述目标切片在第一切片上各对应spot的索引的向量；

$$I_{pre}^n \in \arg\max_{i \in \{0, \dots, N_{pre}-1\}} M_{i,n}^{(pre)}$$

[0126] 示例二，所述相邻切片包括右相邻的第二切片；所述获取所述目标切片每个目标spot，分别在相邻切片上对应的概率传输矩阵中，概率转移值最大的spot的索引向量包括：获取第二概率传输矩阵，所述第二概率传输矩阵为所述目标切片与第二切片计算得到的概率传输矩阵，其中，所述第二切片每个spot对应所述第三概率传输矩阵上的一列，所述目标切片每个spot对应所述第二概率传输矩阵上的一行，在所述第二概率传输矩阵中查找概率转移值最大的元素所在的列，并将所述概率转移值最大的元素所在的列的列标信息，确定为概率转移值最大的spot的索引向量。

[0127] 呈由上述示例，在找目标切片S2片上每一个spot，在第二切片(S3)上对应的spot点时，首先取得S2和S3在输出的第二概率转移矩阵M23；S2上的每一个spot，对应M23上的一行。每一行对应若干个元素，而元素值最大的元素所在的列，对应的是该spot在S3上，对应第二概率转移矩阵上，概率转移值最大的spot点，确定为概率转移值最大的spot的索引向量。

[0128] 实际应用中，以相邻切片包含第二切片为例，N为目标切片上spot的个数， N_{next} 为第二切片spot的个数； $M^{(next)} \in \mathbb{R}^{N \times N_{next}}$ ，为目标切片和第二切片计算得到的第二概率转移矩阵。

[0129] 计算 $I_{next} \in \mathbb{Z}^N$ ，为描述目标切片在第二切片上各对应spot的索引的向量：

$$I_{\text{next}}^n \in \arg\max_{j \in \{0, \dots, N_{\text{next}}-1\}} M_{n,j}^{(\text{next})}$$

- [0130] 示例三，所述相邻切片包括左相邻的第三切片以及右相邻的第四切片；所述获取所述目标切片每个目标spot，分别在相邻切片上对应的概率传输矩阵中，概率转移值最大的spot的索引向量包括：获取第三概率传输矩阵，所述第三概率传输矩阵为第三切片与所述目标切片计算得到的概率传输矩阵，其中，所述目标切片每个spot对应所述第二概率传输矩阵上的一列，所述第三切片每个spot对应所述第三概率传输矩阵上的一行；在所述第三概率传输矩阵中查找概率转移值最大的元素所在的行，并将所述概率转移值最大的元素所在的行的行标信息，确定为概率转移值最大的spot的索引向量；获取第四概率传输矩阵，所述第四概率传输矩阵为所述目标切片与第四切片计算得到的概率传输矩阵，其中，所述第四切片每个spot对应所述第四概率传输矩阵上的一列，所述目标切片每个spot对应所述第四概率传输矩阵上的一行；在所述第四概率传输矩阵中查找概率转移值最大的元素所在的列，并将所述概率转移值最大的元素所在的列的列标信息，确定为概率转移值最大的spot的索引向量。
- [0131] 该实现过程为示例一和示例二的结合，具体的，可参阅示例一和示例二的详细说明，本申请实施例在此不再进行赘述。
- [0132] 作为本申请实施例的一种可实现方式，在步骤103执行根据所述索引向量对应的相邻切片上的spot点坐标，及所述目标spot的坐标，计算所述配准后的目标切片在所述相邻切片上的形变场时，可以采用但不局限于以下方式：根据所述索引向量从预设坐标矩阵中查找在所述目标切片在所述相邻切片上对应的spot点坐标；其中，所述预设坐标矩阵中第一列对应所述索引向量的横坐标，第二列对应所述索引向量的纵坐标，每行表示一个spot；基于所述spot点坐标计算在所述相邻切片上的形变场。
- [0133] 呈由上述示例，依旧以相邻切片包括左相邻的第三切片(前一张切片)以及右相邻的第四切片(后一张切片)。定义： X 为目标切片上所有点的坐标，其中第一列存储 x 坐标，第二列存储 y 坐标，每行表示一个spot，顺序和 $M^{(\text{pre})}$ 中列方向，以及 $M^{(\text{next})}$ 对应的行方向上的spot顺序相一致；定义 $X^{(\text{pre})}$ 为前一片上所有点的坐标，其中第一列存储 x 坐标，第二列存储 y 坐标，每行表示一个spot，顺序和 $M^{(\text{pre})}$ 中行方

向对应的spot顺序相一致；定义 $X^{(next)}$ 为后一片上所有点的坐标，其中第一列存储x坐标，第二列存储y坐标，每行表示一个spot，顺序和 $M^{(next)}$ 中列方向对应的spot顺序相一致，psize为对应形变场上同一像素的两点的最远横、纵距离；该值只要和切片上两点的平均距离接近即可。

[0134] 在计算在所述前一张切片上的形变场 $\Delta X^{(pre)} \in \mathbb{R}^{N \times 2}$ ，和在所述后一张切片上的形变场 $\Delta X^{(next)} \in \mathbb{R}^{N \times 2}$ 时，采用下述计算方式：

[0135] 其中，在预设坐标矩阵中第一列存储x坐标，第二列存储y坐标，每行表示一个spot，顺序和 $M^{(pre)}$ 中列方向对应的spot顺序相一致；

$$X_{n,:}^{(corres_pre)} = X_{I_{pre}^n}^{(pre)}$$

$$X_{n,:}^{(corres_next)} = X_{I_{next}^n}^{(next)}$$

$$\Delta X_{n,:}^{(pre)} = X_{n,:}^{(corres_pre)} - X_{n,:}$$

$$\Delta X_{n,:}^{(next)} = X_{n,:}^{(corres_next)} - X_{n,:}$$

[0136] 计算 $X^{(pixel)} \in \mathbb{R}^{N \times 2}$ ，每一行表示一个spot，第一列为该spot对应形变场的i索引(行位置)，第二列对应该spot对应形变场的j索引(列位置)：

$$X_{i,:}^{(pixel)} = X_{i,:} - [\min_{i \in \{0, \dots, N\}} X_{i,0}, \min_{i \in \{0, \dots, N\}} X_{i,1}]^T$$

$$X_{i,j \in \{0,1\}}^{(pixel)} = \lceil X_{i,j}^{(pixel)} / \text{psize} \rceil$$

[0137] 计算弹性场的高H，和宽W：其中：

$$H = \lceil (\max_{i \in \{0, \dots, N\}} X_{i,0} - \min_{i \in \{0, \dots, N\}} X_{i,0}) / \text{psize} \rceil$$

$$W = \lceil (\max_{i \in \{0, \dots, N\}} X_{i,1} - \min_{i \in \{0, \dots, N\}} X_{i,1}) / \text{psize} \rceil$$

[0138] 计算弹性场的网络节点 $\text{mesh}_{\text{trans}} \in \mathbb{N}^{n_{\text{grids}} \times 2}, n_{\text{grids}} = H \times W$ 。例如采用网格节点计算方法, f_{meshgrid} 计算得到网格节点: $\text{mesh} = f_{\text{meshgrid}}([0:H, 0:W])$, 注意 $\text{mesh} \in \mathbb{Z}^{2 \times H \times W}$; 并采用格式转换方法 f_{flatten} 和 $f_{\text{concatenate}}$ (对该张量的格式进行转换:

$$\begin{aligned} \text{mesh}_{\text{trans}} &= f_{\text{concatenate}}((f_{\text{expanddims}}(f_{\text{flatten}}(\text{mesh}_{0,:}), \text{axis}=1), \\ &f_{\text{expand_dims}}(f_{\text{flatten}}(\text{mesh}_{1,:}), \text{axis}=1)), \text{axis}=1) \end{aligned}$$

[0139] 计算整个形变场时, 可以采用但不限于使用插值方法, $f_{\text{interp_grid}}$ 得到整个形变场:

$$F(x_{\text{pre}}) = f_{\text{interp_grid}}(X^{(\text{pixel})}), \Delta X_{:,0}^{(\text{pre})}, \text{mesh}_{\text{trans}})$$

$$F(y_{\text{pre}}) = f_{\text{interp_grid}}(X^{(\text{pixel})}), \Delta X_{:,1}^{(\text{pre})}, \text{mesh}_{\text{trans}})$$

$$F(y_{\text{next}}) = f_{\text{interp_grid}}(X^{(\text{pixel})}), \Delta X_{:,0}^{(\text{next})}, \text{mesh}_{\text{trans}})$$

$$F(y_{\text{next}}) = f_{\text{interp_grid}}(X^{(\text{pixel})}), \Delta X_{:,1}^{(\text{next})}, \text{mesh}_{\text{trans}})$$

[0140] 插值方法除了上述实现方式外还可以采用线性插值, 三次插值, B-Spline插值、TPS插值等等方法, 具体的, 本申请实施例不做限定。

[0141] 在实际应用中, 上述计算得到的形变场, 如图2所示, 在一些位置存在形变值突变, 或梯度较为明显的变化。切片由受力导致的形变, 应当平滑变化的规律; 实际上体现的是输出的形变场的误差, 因此它会导致下游配准结果的不准确。这种变化主要来自于计算得到的概率转移矩阵存在的误差, 该误差又是由本方法无法影响的上游数据质量, 以及方法本身的误差导致。

[0142] 因此减少这种变化对下游配准结果带来的消极影响, 在在相邻两张切片的所述形变场中分别对应查找到所述坐标的变化值之前, 调用预设滤波函数对所述目标切片的所述形变场进行平滑处理, 得到滤波后的形变场; 在相邻两张切片的所述滤波后的形变场中分别对应查找到所述坐标的变化值。

[0143] 实际应用中预设滤波函数可以采用但不局限于高斯滤波，对形变场进行平滑，达到降噪的效果。如图3为例，图3为本申请实施例提供的一种经滤波处理后的形变场。

[0144] 具体的对形变场进行滤波平滑，从而达到降噪的效果。采用高斯滤波方法 $f_{\text{gaussian_filter}}$ ，对每张切片，计算得到4个滤波后的形变场：

$$F^{(x_pre)} = f_{\text{gaussian_filter}}(F^{(x_pre)})$$

$$F^{(y_pre)} = f_{\text{gaussian_filter}}(F^{(y_pre)})$$

$$F^{(x_next)} = f_{\text{gaussian_filter}}(F^{(x_next)})$$

$$F^{(y_next)} = f_{\text{gaussian_filter}}(F^{(y_next)})$$

[0145] 实际应用中，除了高斯滤波的平滑方法。该平滑方法也可以采用均值滤波，中位数滤波等方法，具体的本申请实施例不做限定。

[0146] 执行完滤波处理后，本发明滤波后的形变场，应用到每片的spot的坐标上。具体做法：基于每片切片各spot的坐标，在其前后滤波后的形变场中，分别对应查找到坐标的变化值；并将前后变化值的平均值应用在坐标上。

[0147] 基于计算得到的 $X^{(\text{pixel})}$ ，以及最终输出的 $F^{(x_pre)}$, $F^{(y_pre)}$, $F^{(x_next)}$, $F^{(y_next)}$ ，计算得 到

$$\Delta X^{(\text{pre_final})} \in \mathbb{R}^{N \times 2}, \quad \Delta X^{(\text{next_final})} \in \mathbb{R}^{N \times 2};$$

$$\Delta X_{i,0}^{(\text{pre_final})} = F_{X_{i,0}^{(\text{pixel})}}^{(x_pre)}$$

$$\Delta X_{i,1}^{(\text{pre_final})} = F_{X_{i,1}^{(\text{pixel})}}^{(y_pre)}$$

$$\Delta X_{i,0}^{(\text{next_final})} = F_{X_{i,0}^{(\text{pixel})}}^{(x_next)}$$

$$\Delta X_{i,1}^{(\text{next_final})} = F_{X_{i,1}^{(\text{pixel})}}^{(y_next)}$$

[0148] 需要说明的是，上述示例是以目标切片存在前一张切片和后一张切片为例进行的说明，若不存在前一张切片，则只计算 $\Delta X^{(next_final)}$ ；若不存在后一张切片，则只计算 $\Delta X^{(pre_final)}$ ；计算方法不变。

[0149] 将上述结果应用于坐标上，就得到了重构后的坐标，具体公式如下(以目标切片存在前一张切片和后一张切片为例)：

$$X = X + \frac{1}{2} \Delta X^{(pre_final)} + \frac{1}{2} \Delta X^{(next_final)}$$

[0150] 若目标切片不存在前一张切片，计算公式如下：

$$X = X + \Delta X^{(next_final)}$$

[0151] 若目标切片不存在后一张切片，计算公式如下：

$$X = X + \Delta X^{(pre_final)}$$

[0152] 实际应用中，第一概率转移矩阵中有些元素不为0的位置，所表示的spot点配准关系，并不对应跨切片上距离最近的两个spot，而仅仅是最优输运算法得到的，具较小概率的输运关系。把第一概率转移矩阵上所有的值用于后面的刚性配准，使得配准结果具有准确率低的问题。

[0153] 为了解决上述问题，本申请实施例将原始概率传输矩阵中元素值小于预设阈值的元素值统一置0，得到第一概率转移矩阵，并将第一概率转移矩阵中的每个值重新除以矩阵所有元素的和。

[0154] 其中，预设阈值的计算方法如下：

$$thresh = \min(pi[pi > 0]) + (\max(pi) - \min(pi[pi > 0])) \times perc.$$

[0155] 举例而言，对应的perc的值是0.2，实际上，perc.也可以取其他0-1的值(注意不能等于1，否则将滤掉全部的锚)。

[0156] 本申请实施例还提供另一种时空转录组切片的配准方法，如图4所示，所述方法包括：

[0157] 步骤201，去掉两张相邻切片中的不共具的spot。

[0158] 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。

- [0159] 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。
- [0160] 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。
- [0161] 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。
- [0162] 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。
- [0163] 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。
- [0164] 在判定类型是否相同时, 可以有步骤202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。
- [0165] 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。
- [0166] 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。

[0167] 步骤205，获取目标切片每个目标spot，分别在相邻切片上对应的概率传输矩阵中，概率转移值最大的spot的索引向量；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；

[0168] 步骤206，基于所述概率传输矩阵，对所述目标切片进行刚性配准，得到配准后的目标切片；

[0169] 步骤207，根据所述索引向量对应的相邻切片上的spot点坐标，及所述目标spot的坐标，计算所述配准后的目标切片在所述相邻切片上的形变场；

[0170] 步骤208，在相邻两张切片的所述形变场中分别对应查找到所述坐标的变化值，并利用所述变化值对所述目标切片各spot的坐标进行配准。

[0171] 有关步骤205至步骤208，可参阅图1的相关详细说明，本申请实施例在此不再赘述。

[0172] 上述实施例已详细说明了时空转录组切片的配准流程，以下以实验数据进行说明。

[0173] 实验一：

[0174] 如图5所示，图5示出了一种果蝇幼虫(L3)在进行配准前后的示意图，效果的变化。同样地，弯箭头由加入弹性部分前，指向加入弹性部分后。由图5可看出切片的形状得到改变，其前后的结构相似性得到提高，第三列切片顶部的明显形变得修复，此两个案例均体现了弹性部分的工作还原了多张切片在生物体内的真实形状，因此进一步提高了配准结果的准确性。

[0175] 在加入本申请实施例所述的配准方法前后，对果蝇胚胎不同细胞类型占据切面的面积进行统计，分析其在沿着z轴的不同切面上的变化情况。

[0176] 图6为本申请实施例提供的一种未执行配准前所有细胞占据面积的变化情况示意图，包括上皮(epidermis)细胞、中肠(midgut)细胞、前肠(foregut)细胞，以及所有细胞占据面积的变化情况，图7为本申请实施例提供的一种执行配准前所有细胞占据面积的变化情况示意图，描述的是加入弹性部分(配准)之后，对应的4种

类型占据面积的变化情况。可以看到，加入弹性部分之后，各细胞类型占据面积随着z轴的变化过程更加平滑。对果蝇胚胎来说，各器官在生物体内的表面是平滑的，整个样本也是平滑形状的：因此，沿着切刀方向，上皮细胞、中肠细胞、前肠细胞占据面积，以及完整样本占据面积，应当呈现平滑的变化趋势。加入弹性部分之后的结果更符合这种趋势——这体现了加入弹性工作后，配准结果的准确性被进一步提高。

[0177] 实验二：

[0178] 如图8所示，图8示出了本申请实施例与相关技术的比对示意图，对比了相关技术PASTE，PASTE2以及本申请实施例在鼠胚胎数据上的配准效果。

[0179] a图描述了未经过配准的连续6张切片的原始数据的形状；b图描述的是PASTE方法配准后输出的数据；c图描述的是PASTE2方法配准后输出的数据，d图描述的是本申请实施例配准输出的数据。可以看到，PASTE方法在第5片和第6片的配准中，出现了超过 100° 的误差，PASTE2方法在第3片的配准中，出现了超过 10° 的，肉眼可见的误差，相比之下，本申请实施例配准的角度误差更小。同时，PASTE和PASTE2均没有将第二片顶部的形变进行修复，但本申请实施例所述的方法对该形变进行了修复(如箭头所指)。

[0180] 本申请实施例所述的配准方法，可以不限于对同一个实验样本的3d重构，还可以是跨对象(例如切片不来自同一只动物)、跨时间(例如切片采样自不同的发育时期)、跨物种(例如鼠脑切片和兔脑切片)的配准。

[0181] 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。

[0182] 图9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9所示，包括：

[0183] 第一获取单元31，用于获取目标切片每个目标spot，分别在相邻切片上对应的概率传输矩阵中，概率转移值最大的spot的索引向量；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；

- [0184] 第二获取单元32，用于获取所述目标切片每个目标spot，分别在相邻切片上对应的第二概率传输矩阵中，概率转移值最大的spot的索引向量；
- [0185] 第一配准单元33，用于基于所述第一概率传输矩阵及所述第二概率传输矩阵，对所述目标切片进行刚性配准，得到配准后的目标切片；
- [0186] 第一计算单元34，用于根据所述索引向量对应的相邻切片上的spot点坐标，及所述目标spot的坐标，计算所述配准后的目标切片在所述相邻切片上的形变场；
- [0187] 第一查找单元35，用于在相邻两张切片的所述形变场中分别对应查找到所述坐标的变化值；
- [0188] 第二配准单元36，用于利用所述变化值对所述目标切片各spot的坐标进行配准。
- [0189] 进一步地，在本实施例一种可能的实现方式中，如图10所示，所述相邻切片包括 左相邻的第一切片；
- [0190] 所述第二获取单元32还用于：
- [0191] 获取第一概率传输矩阵，所述第一概率传输矩阵为第一切片与所述目标切片计算得到的概率传输矩阵，其中，所述目标切片每个spot对应所述第二概率传输矩阵上的一列，所述第一切片每个spot对应所述第一概率传输矩阵上的一行；
- [0192] 在所述第一概率传输矩阵中查找概率转移值最大的元素所在的行，并将所述概率转移值最大的元素所在的行的行标信息，确定为概率转移值最大的spot的索引向量。进一步地，在本实施例一种可能的实现方式中，如图10所示，所述相邻切片包括右相邻的第二切片；
- [0193] 所述第二获取单元32还用于：
- [0194] 获取第二概率传输矩阵，所述第二概率传输矩阵为所述目标切片与第二切片计算得到的概率传输矩阵，其中，所述第二切片每个spot对应所述第三概率传输矩阵上的一列，所述目标切片每个spot对应所述第二概率传输矩阵上的一行；
- [0195] 在所述第二概率传输矩阵中查找概率转移值最大的元素所在的列，并将所述概率转移值最大的元素所在的列的列标信息，确定为概率转移值最大的spot的索引向量。

- [0196] 进一步地，在本实施例一种可能的实现方式中，如图10所示，所述相邻切片包括左相邻的第三切片以及右相邻的第四切片；
- [0197] 所述第二获取单元32还用于：
- [0198] 获取第三概率传输矩阵，所述第三概率传输矩阵为第三切片与所述目标切片计算得到的概率传输矩阵，其中，所述目标切片每个spot对应所述第二概率传输矩阵上的一列，所述第三切片每个spot对应所述第三概率传输矩阵上的一行；
- [0199] 在所述第三概率传输矩阵中查找概率转移值最大的元素所在的行，并将所述概率转移值最大的元素所在的行的行标信息，确定为概率转移值最大的spot的索引向量；
- [0200] 获取第四概率传输矩阵，所述第四概率传输矩阵为所述目标切片与第四切片计算得到的概率传输矩阵，其中，所述第四切片每个spot对应所述第四概率传输矩阵上的一列，所述目标切片每个spot对应所述第四概率传输矩阵上的一行；
- [0201] 在所述第四概率传输矩阵中查找概率转移值最大的元素所在的列，并将所述概率转移值最大的元素所在的列的列标信息，确定为概率转移值最大的spot的索引向量。进一步地，在本实施例一种可能的实现方式中，如图10所示，所述计算单元33还用于：
- [0202] 根据所述索引向量从预设坐标矩阵中查找在所述目标切片在所述相邻切片上对应的spot点坐标；其中，所述预设坐标矩阵中第一列对应所述索引向量的横坐标，第二列对应所述索引向量的纵坐标，每行表示一个spot；
- [0203] 基于所述spot点坐标计算在所述相邻切片上的形变场。
- [0204] 进一步地，在本实施例一种可能的实现方式中，如图10所示，所述装置还包括：
- [0205] 处理单元37，用于在所述第一查找单元35在相邻两张切片的所述形变场中分别对应查找到所述坐标的变化值之前，调用预设滤波函数对所述目标切片的所述形变场进行平滑处理，得到滤波后的形变场；
- [0206] 所述第一查找单元35，还用于在相邻两张切片的所述滤波后的形变场中分别对应查找到所述坐标的变化值。

[0207] 进一步地，在本实施例一种可能的实现方式中，如图10所示，所述装置还包括：

[0208] 第二查找单元38，用于在所述第一获取单元获取目标切片中同一类型的spot对应的第一概率传输矩阵之后，查找原始概率传输矩阵中元素值小于预设阈值的目标元素值；

[0209] 配置单元39，用于将所有所述目标元素值配置为0，得到所述第一概率传输矩阵。

[0210] 进一步地，在本实施例一种可能的实现方式中，如图10所示，所述装置还包括：

[0211] 第三获取单元310，用于在获取目标切片每个目标spot，分别在相邻切片上对应的概率传输矩阵中，概率转移值最大的spot的索引向量之前，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；

[0212] 确定单元311，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；

[0213] 第二计算单元312，用于根据所述每一类型spot对应的权重及预设正则化系数，分别计算第一概率传输矩阵，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。

[0214] 进一步地，在本实施例一种可能的实现方式中，如图10所示，所述装置还包括：

[0215] 去除单元313，用于在所述第三获取单元310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。

[0216] 本公开提供的时空转录组切片的配准装置，获取目标切片中同一类型的spot对应的第一概率传输矩阵，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，获取所述目标切片每个目标spot，分别在相邻切片上对应的第二概率传输矩阵中，概率转移值最大的spot的索引向量，根据所述索引向量对应的相邻切片上的spot点坐标，及所述目标spot的坐标，计算所述配准后的目标切片在所

述相邻切片上的形变场，在相邻两张切片的所述形变场中分别对应查找到所述坐标的变化值，并利用所述变化值对所述目标切片各spot的坐标进行配准。在对目标切片进行刚性配准之后，再通过对切片的形变进行配准，还原了多张切片在生物体内的真实形状，进一步提高了配准结果的准确性。

[0217] 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。

[0218] 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。

[0219] 图11示出了可以用来实施本公开的实施例的示例电子设备400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。

[0220] 如图11所示，设备400包括计算单元401，其可以根据存储在ROM(Read-Only Memory，只读存储器)402中的计算机程序或者从存储单元408加载到RAM(Random Access Memory，随机访问/存取存储器)403中的计算机程序，来执行各种适当的动作和处理。在RAM 403中，还可存储设备400操作所需的各种程序和数据。计算单元401、ROM 402以及RAM 403通过总线404彼此相连。I/O(Input/Output，输入/输出)接口405也连接至总线404。

[0221] 设备400中的多个部件连接至I/O接口405，包括：输入单元406，例如键盘、鼠标等；输出单元407，例如各种类型的显示器、扬声器等；存储单元408，例如磁盘、光盘等；以及通信单元409，例如网卡、调制解调器、无线通信收发机等。通信单元409允许设备400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。

[0222] 计算单元401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401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401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408。在一些实施例中, 计算机程序的部分或者全部可以经由ROM 402和/或通信单元409而被载入和/或安装到设备400上。当计算机程序加载到RAM 403并由计算单元401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401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。

[0223] 本文中以上描述的系统和技术和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)、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。

[0224] 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, 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。

- [0225] 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、或上述内容的任何合适组合。
- [0226] 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(Liquid Crystal 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。
- [0227] 可以将此处描述的系统和技术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。
- [0228] 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器的关系。服务器可以是云服务器，又称为云计算服务器或云主机，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。

- [0229] 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。
- [0230] 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。
- [0231] 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

[权利要求 1]

一种时空转录组切片的配准方法，其特征在于，包括：
获取目标切片每个目标spot，分别在相邻切片上对应的概率传输矩阵中，概率转移值最大的spot的索引向量；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；
基于所述概率传输矩阵，对所述目标切片进行刚性配准，得到配准后的目标切片；
根据所述索引向量对应的相邻切片上的spot点坐标，及所述目标spot的坐标，计算所述配准后的目标切片在所述相邻切片上的形变场；
在相邻两张切片的所述形变场中分别对应查找到所述坐标的变化值，并利用所述变化值对所述目标切片各spot的坐标进行配准。

[权利要求 2]

根据权利要求1所述的方法，其特征在于，所述相邻切片包括左相邻的第一切片；
所述获取目标切片每个目标spot，分别在相邻切片上对应的概率传输矩阵中，概率转移值最大的spot的索引向量包括：
获取第一概率传输矩阵，所述第一概率传输矩阵为第一切片与所述目标切片计算得到的概率传输矩阵，其中，所述目标切片每个spot对应所述第二概率传输矩阵上的一列，所述第一切片每个spot对应所述第一概率传输矩阵上的一行；
在所述第一概率传输矩阵中查找概率转移值最大的元素所在的行，并将所述概率转移值最大的元素所在的行的行标信息，确定为概率转移值最大的spot的索引向量。

[权利要求 3]

根据权利要求1所述的方法，其特征在于，所述相邻切片包括右相邻的第二切片；
所述获取目标切片每个目标spot，分别在相邻切片上对应的概率传输矩阵中，概率转移值最大的spot的索引向量包括：

获取第二概率传输矩阵，所述第二概率传输矩阵为所述目标切片与第二切片计算得到的概率传输矩阵，其中，所述第二切片每个spot对应所述第三概率传输矩阵上的一列，所述目标切片每个spot对应所述第二概率传输矩阵上的一行；

在所述第二概率传输矩阵中查找概率转移值最大的元素所在的列，并将所述概率转移值最大的元素所在的列的列标信息，确定为概率转移值最大的spot的索引向量。

[权利要求 4]

根据权利要求1所述的方法，其特征在于，所述相邻切片包括左相邻的第三切片以及右相邻的第四切片；

所述获取目标切片每个目标spot，分别在相邻切片上对应的概率传输矩阵中，概率转移值最大的spot的索引向量包括：

获取第三概率传输矩阵，所述第三概率传输矩阵为第三切片与所述目标切片计算得到的概率传输矩阵，其中，所述目标切片每个spot对应所述第二概率传输矩阵上的一列，所述第三切片每个spot对应所述第三概率传输矩阵上的一行；

在所述第三概率传输矩阵中查找概率转移值最大的元素所在的行，并将所述概率转移值最大的元素所在的行的行标信息，确定为概率转移值最大的spot的索引向量；

获取第四概率传输矩阵，所述第四概率传输矩阵为所述目标切片与第四切片计算得到的概率传输矩阵，其中，所述第四切片每个spot对应所述第四概率传输矩阵上的一列，所述目标切片每个spot对应所述第四概率传输矩阵上的一行；

在所述第四概率传输矩阵中查找概率转移值最大的元素所在的列，并将所述概率转移值最大的元素所在的列的列标信息，确定为概率转移值最大的spot的索引向量。

[权利要求 5]

根据权利要求1所述的方法，其特征在于，所述根据所述索引向量对应的相邻切片上的spot点坐标，及所述目标spot的坐标，计算所述配准后的目标切片在所述相邻切片上的形变场包括：

根据所述索引向量从预设坐标矩阵中查找在所述配准后的目标切片在所述相邻切片上对应的spot点坐标；其中，所述预设坐标矩阵中第一列对应所述索引向量的横坐标，第二列对应所述索引向量的纵坐标，每行表示一个spot；

基于所述spot点坐标计算在所述相邻切片上的形变场。

[权利要求 6]

根据权利要求1所述的方法，其特征在于，在在相邻两张切片的所述形变场中分别对应查找到所述坐标的变化值之前，所述方法还包括：

调用预设滤波函数对所述目标切片的所述形变场进行平滑处理，得到滤波后的形变场；

所述在相邻两张切片的所述形变场中分别对应查找到所述坐标的变化值包括：

在相邻两张切片的所述滤波后的形变场中分别对应查找到所述坐标的变化值。

[权利要求 7]

根据权利要求1所述的方法，其特征在于，在获取目标切片中同一类型的spot对应的第一概率传输矩阵之后，所述方法还包括：

查找所述第一概率传输矩阵中元素值小于预设阈值的目标元素值；

将所有所述目标元素值配置为0。

[权利要求 8]

根据权利要求1-7中任一项所述的方法，其特征在于，在获取目标切片每个目标spot，分别在相邻切片上对应的概率传输矩阵中，概率转移值最大的spot的索引向量之前，所述方法还包括：

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，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。

[权利要求 9] 根据权利要求8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。

[权利要求 10] 一种时空转录组切片的配准装置，其特征在于，包括：

第一获取单元，用于获取目标切片每个目标spot，分别在相邻切片上对应的概率传输矩阵中，概率转移值最大的spot的索引向量；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；

第二获取单元，用于获取所述目标切片每个目标spot，分别在相邻切片上对应的第二概率传输矩阵中，概率转移值最大的spot的索引向量；

第一配准单元，用于基于所述第一概率传输矩阵及所述第二概率传输矩阵，对所述目标切片进行刚性配准，得到配准后的目标切片；

第一计算单元，用于根据所述索引向量对应的相邻切片上的spot点坐标，及所述目标spot的坐标，计算所述配准后的目标切片在所述相邻切片上的形变场；

第一查找单元，用于在相邻两张切片的所述形变场中分别对应查找所述坐标的变化值；

第二配准单元，用于利用所述变化值对所述目标切片各spot的坐标进行配准。

[权利要求 11] 一种电子设备，其特征在于，包括：

至少一个处理器；以及

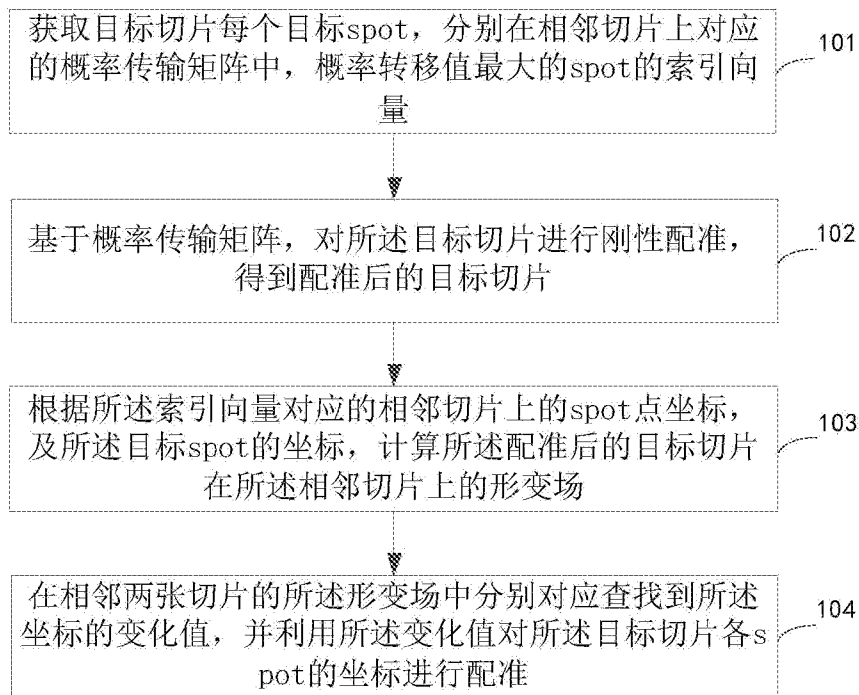
与所述至少一个处理器通信连接的存储器；其中，

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1-9中 任一项所述的方法。

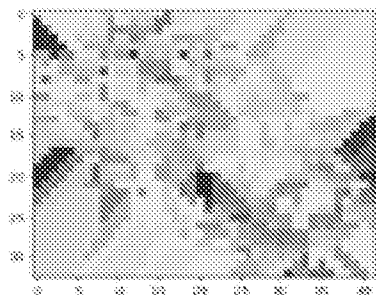
[权利要求 12] 一种存储有计算机指令的非瞬时计算机可读存储介质，其特征在于，所述计算机指令用于使所述计算机执行根据权利要求1-9中任一项所述的方法。

[权利要求 13] 一种计算机程序产品，其特征在于，包括计算机程序，所述计算机程序在被处理器执行时实现根据权利要求1-9中任一项所述的方法。

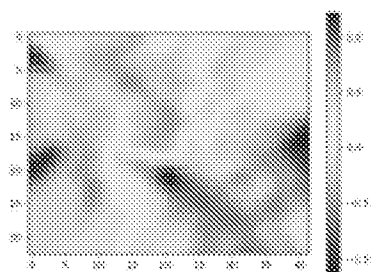
[图 1]

细则 91,
28.12.2023

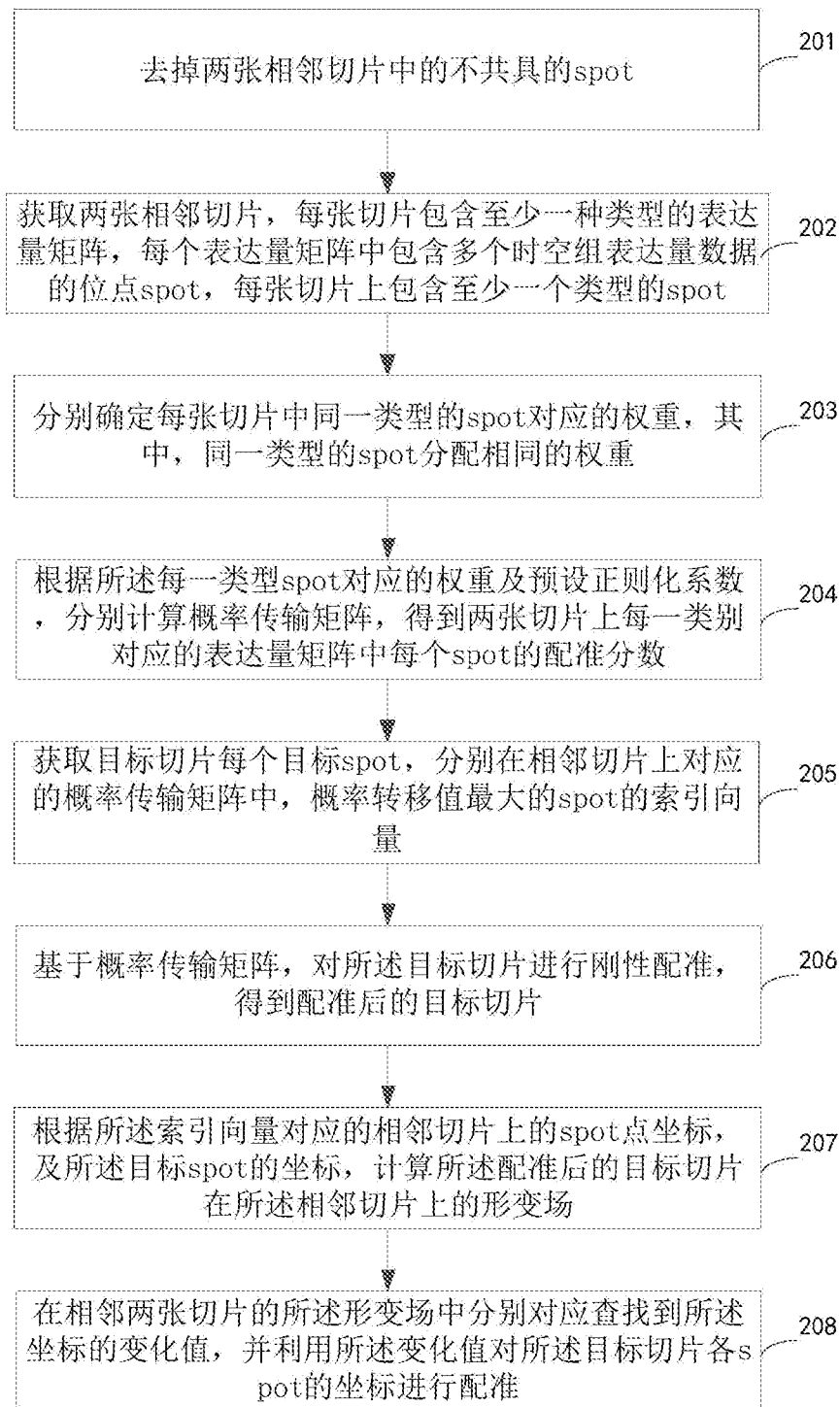
[图 2]

细则 91,
28.12.2023

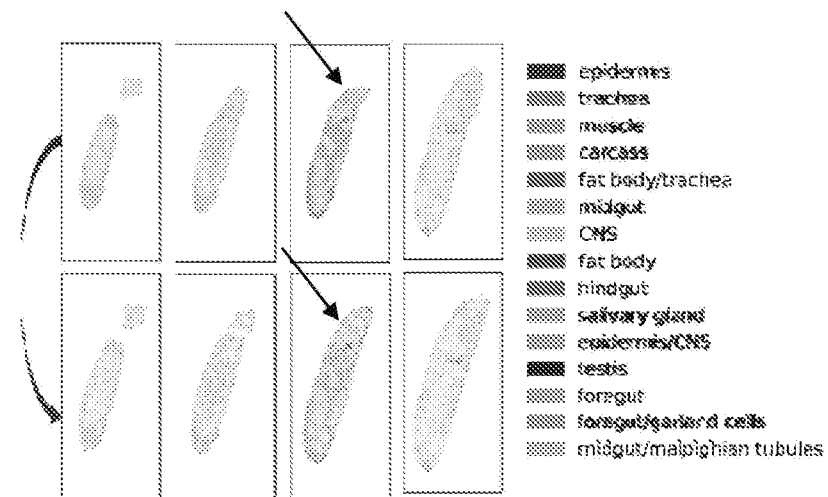
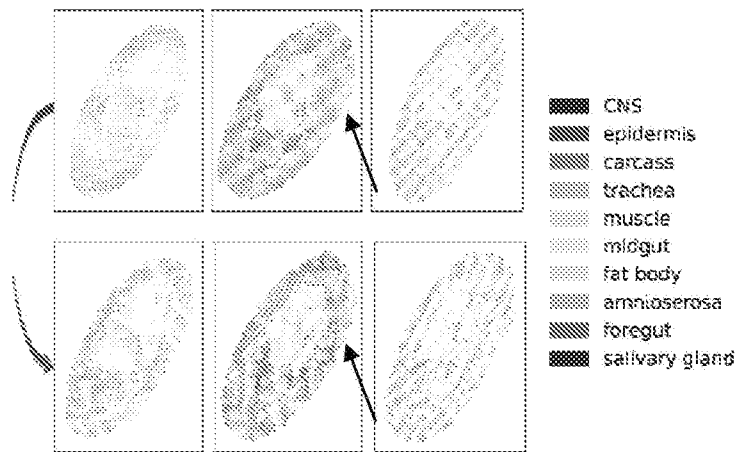
[图 3]

细则 91,
28.12.2023

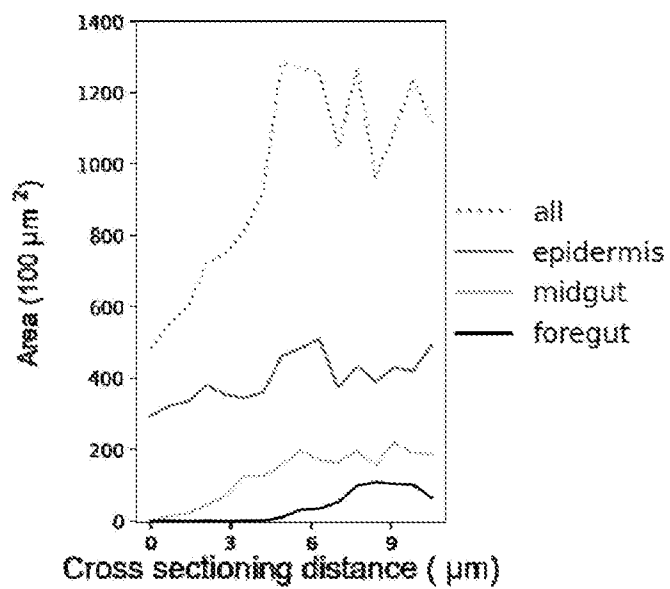
[图 4]



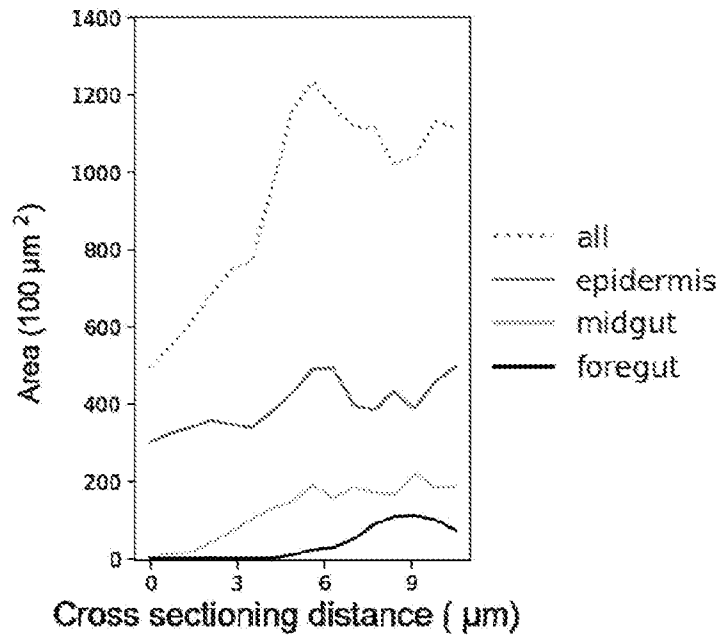
[图 5]



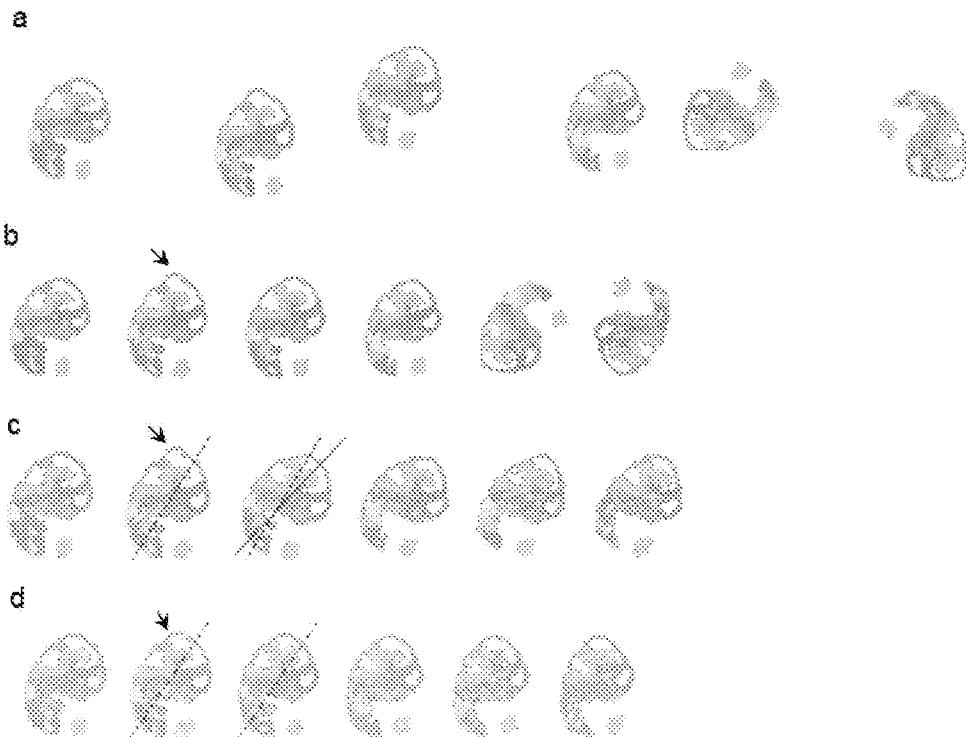
[图 6]



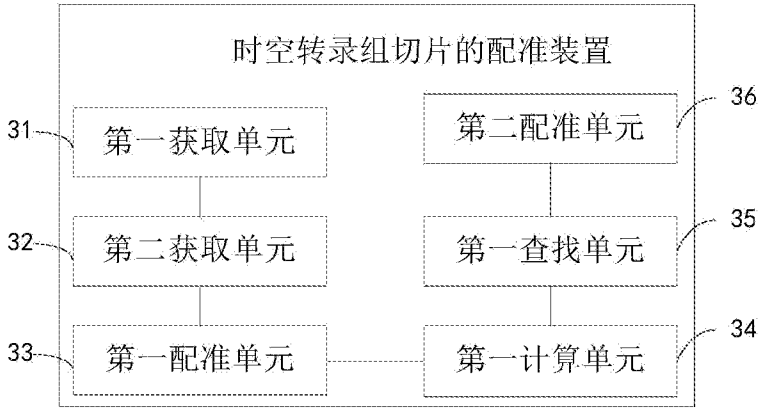
[图 7]



[图 8]

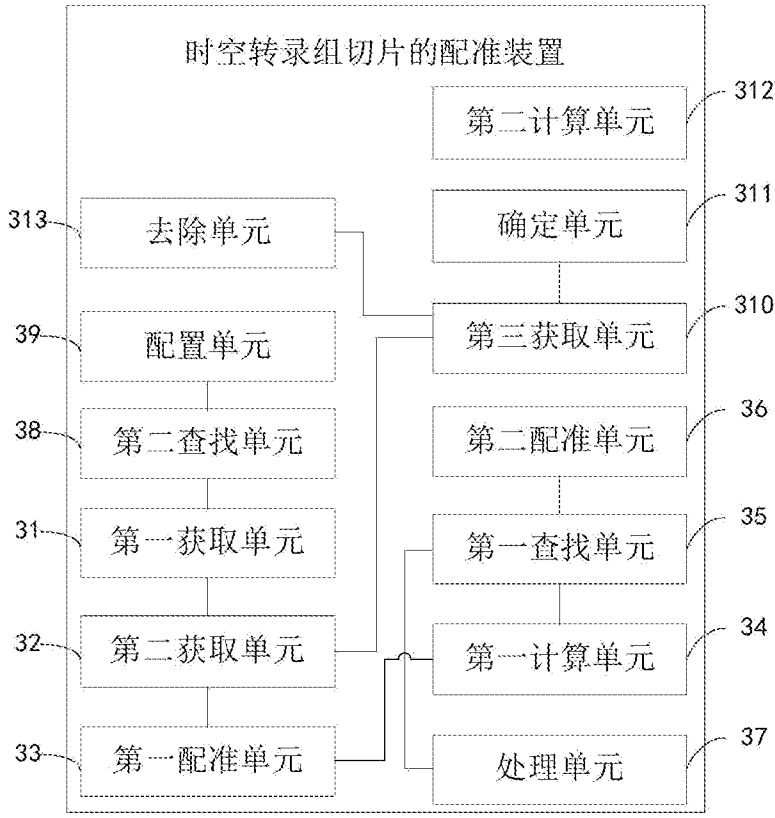


[图 9]



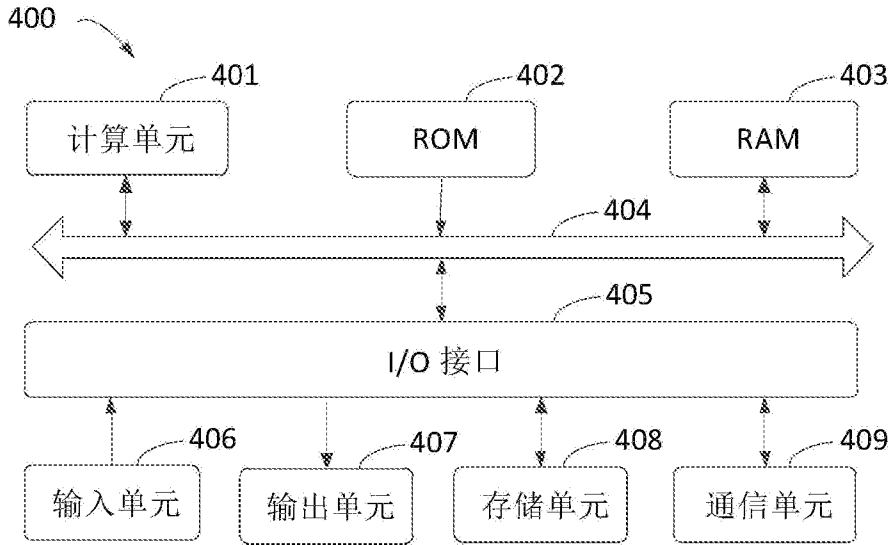
细则 91,
28.12.2023

[图 10]



细则 91,
28.12.2023

[图 11]



细则 91,
28.12.2023

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/120768

A. CLASSIFICATION OF SUBJECT MATTER

G16B 20/30(2019.01)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

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)

VEN, CNABS, CNTXT, WOTXT, USTXT, EPTXT, CNKI, IEEE: 时空, 空间, 转录, 切片, 配准, 概率, 矩阵, 索引, 最大, 邻, spatial, transcriptomics, PASTE, section, slice, spot, alignment, probabilistic, matrix, index, key, adjacent, neighbouring

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 107545567 A (INSTITUTE OF AUTOMATION, CHINESE ACADEMY OF SCIENCES) 05 January 2018 (2018-01-05) description, paragraphs [0046]-[0065]	1-13
A	CN 115798587 A (SHANGHAI LUMING BIOLOGICAL TECHNOLOGY CO., LTD.) 14 March 2023 (2023-03-14) entire document	1-13
A	CN 108038874 A (INSTITUTE OF AUTOMATION, CHINESE ACADEMY OF SCIENCES) 15 May 2018 (2018-05-15) entire document	1-13
A	CN 114817363 A (BEIJING BAIDU NETCOM SCIENCE AND TECHNOLOGY CO., LTD.) 29 July 2022 (2022-07-29) entire document	1-13
A	US 2018182103 A1 (INTERNATIONAL BUSINESS MACHINES CORP.) 28 June 2018 (2018-06-28) entire document	1-13



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

26 January 2024

Date of mailing of the international search report

31 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

International application No.

PCT/CN2023/120768

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 113269813 A (GENOME WISDOM INC.) 17 August 2021 (2021-08-17) entire document	1-13
A	CN 110874849 A (ANHUI UNIVERSITY) 10 March 2020 (2020-03-10) entire document	1-13

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2023/120768

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	107545567	A	05 January 2018	None	
CN	115798587	A	14 March 2023	None	
CN	108038874	A	15 May 2018	None	
CN	114817363	A	29 July 2022	None	
US	2018182103	A1	28 June 2018	None	
CN	113269813	A	17 August 2021	None	
CN	110874849	A	10 March 2020	None	

A. 主题的分类

G16B 20/30(2019.01)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

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))

VEN, CNABS, CNTXT, WOTXT, USTXT, EPTXT, CNKI, IEEE: 时空, 空间, 转录, 切片, 配准, 概率, 矩阵, 索引, 最大, 邻, spatial, transcriptomics, PASTE, section, slice, spot, alignment, probabilistic, matrix, index, key, adjacent, neighbouring

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 107545567 A (中国科学院自动化研究所) 2018年1月5日 (2018 - 01 - 05) 说明书第[0046] - [0065]段	1-13
A	CN 115798587 A (上海鹿明生物科技有限公司) 2023年3月14日 (2023 - 03 - 14) 全文	1-13
A	CN 108038874 A (中国科学院自动化研究所) 2018年5月15日 (2018 - 05 - 15) 全文	1-13
A	CN 114817363 A (北京百度网讯科技有限公司) 2022年7月29日 (2022 - 07 - 29) 全文	1-13
A	US 2018182103 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 2018年6月28日 (2018 - 06 - 28) 全文	1-13
A	CN 113269813 A (志诺维思(北京)基因科技有限公司) 2021年8月17日 (2021 - 08 - 17) 全文	1-13
A	CN 110874849 A (安徽大学) 2020年3月10日 (2020 - 03 - 10) 全文	1-13

☐ 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

2024年1月26日

国际检索报告邮寄日期

2024年1月31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

赵婷

电话号码 (+86) 010-53961350

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/CN2023/120768

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	107545567	A	2018年1月5日	无	
CN	115798587	A	2023年3月14日	无	
CN	108038874	A	2018年5月15日	无	
CN	114817363	A	2022年7月29日	无	
US	2018182103	A1	2018年6月28日	无	
CN	113269813	A	2021年8月17日	无	
CN	110874849	A	2020年3月10日	无	