



(12) 发明专利

(10) 授权公告号 CN 113534948 B

(45) 授权公告日 2025. 02. 11

(21) 申请号 202011088589.7

(22) 申请日 2020.10.13

(65) 同一申请的已公布的文献号
申请公布号 CN 113534948 A

(43) 申请公布日 2021.10.22

(30) 优先权数据
10-2020-0046268 2020.04.16 KR

(73) 专利权人 三星电子株式会社
地址 韩国京畿道
专利权人 高丽大学校产学协力团

(72) 发明人 金允泰 薛湘勋 成基荣 李泓锡
全明帝

(74) 专利代理机构 北京市柳沈律师事务所
11105

专利代理师 王新宇

(51) Int.Cl.

G06F 3/01 (2006.01)

G06F 3/0346 (2013.01)

G06N 3/045 (2023.01)

G06N 3/0442 (2023.01)

G06N 3/084 (2023.01)

(56) 对比文件

E. Jared Shamwell等.“Unsupervised Deep Visual-Inertial Odometry with Online Error Correction for RGB-D Imagery”.《IEEE Transactions on Pattern Analysis and Machine Intelligence》.2019,第42卷(第10期),第2478-2493页,摘要、第2-7节.

审查员 夏玫

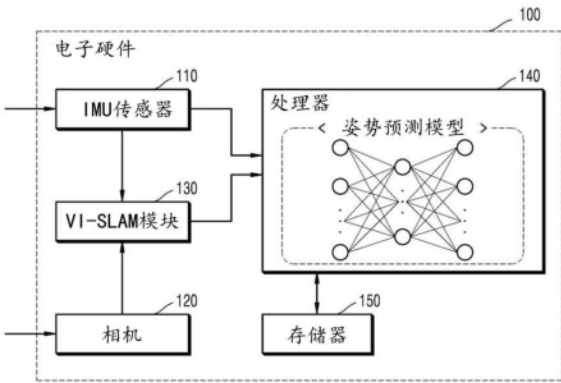
权利要求书3页 说明书10页 附图15页

(54) 发明名称

增强现实AR装置和预测其中的姿势的方法

(57) 摘要

提供增强现实AR装置和预测AR装置的姿势的方法。在增强现实(AR)装置中,以惯性测量单元IMU速率获得对应于所述AR装置的移动的IMU值;经由视觉惯性同步定位与建图(VI-SLAM)模块基于所述IMU值和所述AR装置周围的图像估计所述AR装置的中间6D6自由度(6D))姿势,并且通过使用深度神经网络执行学习来生成用于预测所述AR装置的相对6D姿势的姿势预测模型。



1. 一种预测增强现实AR装置的姿势的方法,所述方法包括:

使用在AR装置中设置的惯性测量单元IMU传感器以第一频率的IMU速率获得对应于AR装置移动的IMU值;

基于所获得的IMU值和所述AR装置周围的图像估计所述AR装置的中间6自由度6D姿势,所述图像通过所述AR装置中设置的相机以第二频率的帧速率获得;

通过处理器生成用于预测所述AR装置的相对6D姿势的姿势预测模型;

将作为第一输入的IMU值输入所述姿势预测模型,以及将在预测相对6D姿势之前估计的中间6D姿势作为第二输入单独输入到所述姿势预测模型;以及

通过使用由所述姿势预测模型实施的深度神经网络基于所获得的IMU值和中间6D姿势执行学习,输出所述AR装置的相对6D姿势。

2. 如权利要求1所述的方法,其中所述中间6D姿势包括:

帧速率的第一类型姿势数据,其通过使用将所获得的IMU值和以第二频率的帧速率获得的图像融合的同步定位与建图SLAM技术来估计,或者

第一频率的IMU速率的第二类型姿势数据,其包括对应于积分值的姿势数据,所述积分值通过预积分方法根据所获得的IMU值和通过使用SLAM技术估计的所述帧速率的基于图像的姿势数据来计算。

3. 如权利要求1所述的方法,其中所述姿势预测模型根据对所述深度神经网络的学习输出每个对应于以滑动方式移动的窗口的相对6D姿势,所述学习基于IMU值中的k个IMU值,以及中间6D姿势中的m个中间6D姿势,以及

其中在所述窗口中,k是自然数,以及m是自然数。

4. 如权利要求3所述的方法,其中所述窗口基于所述IMU速率滑动并且与滑动窗口对应的k个IMU值以及m个中间6D姿势被输入到所述深度神经网络,并且

所述姿势预测模型输出对应于所述IMU速率的每个窗口的相对6D姿势。

5. 如权利要求1所述的方法,还包括:基于完成对中间6D姿势的估计,在拍摄图像的实时时间点将所述估计的中间6D姿势更新并设置为地面实况GT姿势。

6. 如权利要求5所述的方法,其中所述姿势预测模型基于所述GT姿势的更新经由自我监督方法使用所述深度神经网络执行学习。

7. 如权利要求1所述的方法,其中所述姿势预测模型利用包括第一频率的IMU速率的IMU值和第二频率的帧速率的所述中间6D姿势的多速率输入使用所述深度神经网络来执行学习,并且

所述第一频率大于所述第二频率。

8. 如权利要求1所述的方法,其中所述深度神经网络包括使用长短期记忆或门控递归单元的递归神经网络。

9. 如权利要求8所述的方法,其中所述深度神经网络通过使用更新的地面实况GT姿势减少损失函数来执行随时间的反向传播。

10. 如权利要求1所述的方法,还包括:基于根据所述AR装置的移动新获得的IMU值和新估计的中间6D姿势被输入到所述姿势预测模型的所述学习到的深度神经网络,通过所述姿势预测模型的所述学习到的深度神经网络的处理实时地输出预测相对6D姿势的推断结果。

11. 一种非暂时性计算机可读记录介质,其上记录有程序以用于在计算机中执行预测

增强现实AR装置的姿势的方法,所述方法包括:

使用在AR装置中设置的惯性测量单元IMU传感器以第一频率的IMU速率获得对应于AR装置移动的IMU值;

基于所获得的IMU值和所述AR装置周围的图像估计所述AR装置的中间6自由度6D姿势,所述图像通过所述AR装置中设置的相机以第二频率的帧速率获得;

通过处理器生成用于预测所述AR装置的相对6D姿势的姿势预测模型;

将作为第一输入的IMU值输入所述姿势预测模型,以及将在预测相对6D姿势之前估计的中间6D姿势作为第二输入单独输入到所述姿势预测模型;以及

通过使用由所述姿势预测模型实施的深度神经网络基于所获得的IMU值和中间6D姿势执行学习,输出所述AR装置的相对6D姿势。

12. 一种增强现实AR装置,包括:

惯性测量单元IMU传感器,被配置为以第一频率的IMU速率获得对应于所述AR装置的移动的IMU值;

相机,被配置为以第二频率的帧速率获得所述AR装置周围的图像;

视觉惯性同步定位与建图VI-SLAM模块,被配置为基于所获得的IMU值和所述AR装置周围的图像估计所述AR装置的中间6自由度6D姿势;以及

处理器,被配置为生成用于预测所述AR装置的相对6D姿势的姿势预测模型,

其中,将作为第一输入的IMU值输入所述姿势预测模型,以及将在预测相对6D姿势之前估计的中间6D姿势作为第二输入单独输入到所述姿势预测模型,以及

其中,通过使用由所述姿势预测模型实施的深度神经网络基于所获得的IMU值和中间6D姿势执行学习输出所述AR装置的相对6D姿势。

13. 如权利要求12所述的AR装置,其中所述姿势预测模型根据所述深度神经网络的学习输出每个对应于以滑动方式移动的窗口的相对6D姿势,所述学习基于IMU值中的k个IMU值以及中间6D姿势中的m个中间6D姿势,以及

其中在所述窗口中,k是自然数,以及m是自然数。

14. 如权利要求13所述的AR装置,其中所述窗口基于所述IMU速率滑动并且与滑动窗口对应的k个IMU值以及m个中间6D姿势被输入到所述深度神经网络,并且

所述姿势预测模型输出对应于所述IMU速率的每个窗口的相对6D姿势。

15. 如权利要求12所述的AR装置,其中所述处理器进一步被配置为:基于完成所述VI-SLAM模块对中间6D姿势的估计,在拍摄图像的实时时间点将所述估计的中间6D姿势更新并设置为地面实况GT姿势。

16. 如权利要求15所述的AR装置,其中所述姿势预测模型基于所述GT姿势的更新经由自我监督方法使用所述深度神经网络执行学习。

17. 如权利要求12所述的AR装置,其中所述姿势预测模型利用包括第一频率的IMU速率的所述IMU值和第二频率的帧速率的所述中间6D姿势的多速率输入来使用所述深度神经网络执行学习,并且

所述第一频率大于所述第二频率。

18. 如权利要求12所述的AR装置,其中所述深度神经网络包括使用长短期记忆或门控递归单元的递归神经网络。

19. 如权利要求18所述的AR装置,其中所述深度神经网络通过使用更新的GT姿势减少损失函数来执行随时间的反向传播。

20. 如权利要求12所述的AR装置,其中所述处理器进一步被配置为:在根据所述AR装置的移动新获得的IMU值和新估计的中间6D姿势被输入到所述姿势预测模型的所述学习到的深度神经网络时,通过所述姿势预测模型的学习到的深度神经网络的处理实时地输出预测相对6D姿势的推断结果。

21. 一种处理器中预测姿势的方法,所述方法包括:

使用在增强现实AR装置中设置的惯性测量单元IMU传感器以第一频率的IMU速率获得对应于AR装置的移动的IMU值;

获得基于所获得的IMU值和所述AR装置周围的图像估计的所述AR装置的中间6自由度6D姿势,所述图像使用所述AR装置中设置的相机以第二频率的帧速率获得;

生成用于预测所述AR装置的相对6D姿势的姿势预测模型;

将作为第一输入的IMU值输入所述姿势预测模型,以及将在预测相对6D姿势之前获得的中间6D姿势作为第二输入单独输入到所述姿势预测模型;以及

通过使用由所述姿势预测模型实施的深度神经网络基于所获得的IMU值和中间6D姿势执行学习输出所述AR装置的相对6D姿势。

增强现实AR装置和预测其中的姿势的方法

[0001] 相关申请的交叉引用

[0002] 本申请要求在韩国知识产权局中于2020年4月16日提交的韩国专利申请号10-2020-0046268的优先权,所述申请的公开内容通过引用以其全部内容并入本文。

背景技术

[0003] 1.技术领域

[0004] 本公开的示例实施例涉及增强现实 (AR) 装置以及预测AR装置中的姿势的方法。

[0005] 2.相关技术的说明

[0006] 随着最近的技术发展,已经开发了可以佩戴在人体上的各种类型的可佩戴装置。眼镜型可佩戴装置被戴在用户的头上,并且可以通过在装置的显示器上显示关于虚拟对象的视觉信息来向用户提供增强现实服务。

[0007] 根据诸如增强现实 (AR) 技术的显示技术,装置上输出的图像可以被用户感知为真实的。因此,用户通过AR技术显示装置等可以如真实存在地体验现实生活中无法遇到的情况。为了促使这样的体验,重要的是通过AR装置将输出图像实时地提供给用户。输出图像必须与佩戴在用户身上的可佩戴装置的姿势紧密相关,并因此需要精确检测用户的姿势。当不实时显示输出图像时,由于检测到的姿势与输出图像之间的差异引起的错误,用户可能感到不舒服。因此,需要使用诸如计算机视觉或同步定位与建图 (SLAM) 的各种数字处理技术的精确姿势预测方法。

发明内容

[0008] 提供增强现实 (AR) 装置以及预测AR装置的姿势的方法。本公开将要实现的技术目标不限于上述目标,并且可以从以下实施例中推断出其他技术目标。

[0009] 附加的方面将部分地在下面的描述中阐述,并且部分地将从描述中显而易见,或者可以通过本公开所呈现的实施例的实践来了解。

[0010] 根据本公开的一方面,提供一种预测增强现实AR装置的姿势的方法,所述方法包括:使用在AR装置中设置的惯性测量单元IMU以第一频率的IMU速率获得对应于AR装置的移动的IMU值;基于所获得的IMU值和所述AR装置周围的图像估计所述AR装置的中间6自由度 (6D) 姿势,所述图像通过所述AR装置中设置的相机以第二频率的帧速率获得;以及通过处理器使用深度神经网络基于所获得的IMU值和中间6D姿势执行学习,生成用于预测所述AR装置的相对6D姿势的姿势预测模型。

[0011] 根据本公开的另一方面,提供一种非暂时性计算机可读记录介质,其具有记录于其上的程序以用于在计算机中执行预测增强现实AR装置的姿势的方法,所述方法包括:使用在AR装置中设置的惯性测量单元IMU传感器以第一频率的IMU速率获得对应于AR装置的移动的IMU值;基于所获得的IMU值和所述AR装置周围的图像估计所述AR装置的中间6自由度 (6D) 姿势,所述图像通过所述AR装置中设置的相机以第二频率的帧速率获得;以及通过处理器使用深度神经网络基于所获得的IMU值和中间6D姿势执行学习,生成用于预测所述

AR装置的相对6D姿势的姿势预测模型。

[0012] 根据本公开的另一方面,提供一种增强现实AR装置,包括:惯性测量单元IMU传感器,被配置为以第一频率的IMU速率获得对应于所述AR装置的移动的IMU值;相机,配置为以第二频率的帧速率获得所述AR装置周围的图像;视觉惯性同步定位与建图VI-SLAM模块,被配置为基于所获得的IMU值和所述AR装置周围的图像估计所述AR装置的中间6自由度(6D)姿势;以及处理器,配置为通过使用深度神经网络基于所获得的IMU值和中间6D姿势执行学习,生成用于预测所述AR装置的相对6D姿势的姿势预测模型。

[0013] 根据本公开的另一方面,提供一种处理器中预测姿势的方法,所述方法包括:使用在增强现实AR装置中设置的惯性测量单元IMU传感器以第一频率的IMU速率获得对应于AR装置的移动的IMU值;获得基于所获得的IMU值和所述AR装置周围的图像估计的所述AR装置的中间6自由度(6D)姿势,所述图像使用所述AR装置中设置的相机以第二频率的帧速率获得;以及使用深度神经网络基于所获得的IMU值和中间6D姿势执行学习,生成用于预测所述AR装置的相对6D姿势的姿势预测模型。

[0014] 根据本公开的另一方面,提供一种训练神经网络用于姿势预测的计算机实现的方法,所述方法包括:收集对应于增强现实AR装置的移动的惯性测量单元IMU值,以第一频率收集IMU值;收集所述AR装置周围的图像,以第二频率收集所述图像;对收集的IMU值和收集的图像执行一个或多个操作以创建所述AR装置的中间6自由度(6D)姿势;以及基于收集的IMU值和中间6D姿势训练所述神经网络。

[0015] 根据本公开的另一方面,提供一种训练神经网络用于姿势预测的设备,包括:存储器,存储一个或多个指令;和处理器,配置为执行一个或多个指令来:收集对应于增强现实AR装置的移动的惯性测量单元IMU值,以第一频率收集IMU值;收集所述AR装置周围的图像,以第二频率收集所述图像;对收集的IMU值和收集的图像执行一个或多个操作以创建所述AR装置的中间6自由度(6D)姿势;以及基于收集的IMU值和中间6D姿势训练所述神经网络。

附图说明

[0016] 通过以下结合附图的描述,本公开的某些示例实施例的上述和其它方面、特征和优点将变得更加显而易见,其中:

[0017] 图1示出了根据示例实施例的增强现实(AR)装置;

[0018] 图2是根据示例实施例的设置在AR装置中的电子硬件的框图;

[0019] 图3是根据示例实施例的用于经由惯性测量单位(IMU)传感器获得IMU值的时间点的时序图;

[0020] 图4是根据示例实施例的用于经由相机获得图像的时间点的时序图;

[0021] 图5A是根据示例实施例的用于经由虚拟惯性同步定位与建图(VI-SLAM)模块估计中间六度(6D)姿势的时间点的时序图;

[0022] 图5B是根据另一示例实施例的用于经由虚拟惯性同步定位与建图(VI-SLAM)模块估计中间6D姿势的时间点的时序图;

[0023] 图6是根据示例实施例的生成姿势预测模型的方法的时序图;

[0024] 图7A是根据示例实施例的通过滑动窗口方法在姿势预测模型中执行姿势预测的方法的时序图;

[0025] 图7B是根据另一示例实施例的通过滑动窗口方法在姿势预测模型中执行姿势预测的方法的时序图；

[0026] 图8是根据示例实施例的基于深度神经网络 (DNN) 的由递归神经网络 (RNN) 使用长短记忆 (LSTM) 实现的姿势预测模型的图；

[0027] 图9和图10是根据一些示例实施例的在处理器中执行姿势预测的框图；

[0028] 图11A和图11B是根据一些示例实施例的基于已经完成学习的DNN在姿势预测模型中执行姿势预测的推断的时间图；

[0029] 图12是根据示例实施例的预测AR装置中的姿势的方法的流程图。

具体实施方式

[0030] 现在将详细参考实施例,其示例在附图中示出,其中相同的附图标记始终表示相同的元件。在这点上,一些示例实施例可以具有不同的形式,并且不应该被解释为限于这里阐述的描述。因此,下面通过参考附图来描述示例实施例,以解释各方面。如本文中所使用,术语“和/或”包括相关联的所列项目中的一个或多个的任何和所有组合。诸如“……中的至少一个”等表达,当在元件列表之后时,修饰整个元件列表而不修饰列表中的各个元件。

[0031] 考虑到本公开中的功能,已经从当前广泛使用的通用术语中选择了示例实施例中使用的术语。然而,根据本领域的普通技术人员的意图、案件先例以及新技术的出现,这些术语可以变化。此外,对于特殊情况,在描述部分中详细描述了申请人选择的术语的含义。因此,本公开中使用的术语是基于它们相对于整个说明书中所讨论的内容的含义而不是其简单含义来定义的。

[0032] 在示例实施例中,尽管可以使用诸如“第一”、“第二”等的术语来描述各种部件,此类部件必须不限于以上术语。上述术语仅用来将一个部件与另一部件区分开。

[0033] 一些实施例中使用的术语仅用于描述示例实施例,并不旨在限制本公开。除非在上下文中具有明显不同的含义,否则单数形式的表达包括复数形式的表达。此外,在示例实施例中,当组成元件“连接”或“被连接”到另一组成元件时,所述组成元件不仅直接地接触或连接到另一组成元件,而且还通过插在两者之间的其他组成元件中的至少一个电接触或连接到所述另一组成元件。当零件可以“包括”某个组成组件时,除非另有说明,否则它可能不解释为排除另一个组成元件,而是可以解释为进一步包括其他组成元件。

[0034] 关于示例实施例,特别是在权利要求中使用的“以上”和类似指令可以指单数和复数。此外,可以按任何合适的顺序执行根据本文所描述的示例实施例的所有方法的操作或步骤,除非在此另外地指示或另外明显地与上下文矛盾。示例实施例不限于所描述的操作或步骤的描述顺序。

[0035] 可以根据功能块部件和各种处理步骤来描述本公开。此类功能块可以由被配置为执行指定功能的任意数量的硬件和/或软件部件来实现。例如,与示例实施例相关的功能块可以由一个或多个微处理器或用于特定功能的电路元件实现。此外,可以用各种编程或脚本语言来实现功能块。此外,功能块可以在一个或多个处理器上执行的算法中实现。示例实施例可以采用用于电子配置、信号处理和/或控制、数据处理等的任何数量的相关技术。

[0036] 此外,各种附图中所示的连接线或连接器旨在表示各种元件之间的功能关系和/或物理或逻辑耦合。应当注意,在实际装置中可以存在许多替代或附加的功能关系、物理连

接或逻辑连接。

[0037] 以下参考附图详细描述本公开的示例实施例。然而,示例实施例不限于此,并且将理解,可以在其中做出形式和细节的各种改变而不脱离所附权利要求书的精神和范围。

[0038] 图1示出了根据示例实施例的增强现实 (AR) 10 装置。

[0039] 参考图1,AR装置10可以是用户佩戴的可佩戴电子装置。例如,AR装置10可以是眼镜型可佩戴装置,但是实施例不限于此。例如,根据另一示例实施例,AR装置10可以是要佩戴在用户的其他人体部位上的不同类型的可佩戴装置,例如AR头盔。

[0040] AR装置10可以通过AR装置10上的显示器提供AR服务,所述AR服务融合用户周围的现实世界信息和数字或虚拟对象信息。例如,当通过AR装置10在用户的眼睛上显示AR对象(或AR场景)时,可以将AR对象与在以AR装置10的当前姿势看到的现实世界的场景一起显示。根据示例实施例,AR对象可以在AR设备10的当前姿势下由用户感知的现实世界环境的场景上显示,覆盖或叠加。尽管用户可以通过AR装置10看到AR对象,但是AR对象在现实世界中不存在。

[0041] 真实世界是观察者或用户可以通过AR装置10看到的真实场景,并且可以包括真实世界对象。另一方面,AR对象(或场景)是通过图形处理生成的图像,并且可以对应于静态图像或动态图像。例如,AR对象(或场景)可以是覆盖在真实场景上以提供关于真实场景中的真实对象的信息或者关于AR装置10的操作的信息或控制菜单的图像。

[0042] 如图1所示,AR装置10可以对应于如图1所示的头戴式可佩戴装置,并且当用户的头部移动时相应地可以一起移动。因此,由于可以根据用户的头部的移动来连续改变AR装置10的姿势,有必要考虑到用户头部的姿势变化来连续更新AR装置10的姿势。例如,当佩戴AR装置10的用户在道路上行走时,通过AR装置10看到的真实世界场景不断变化。因此,可以考虑当前姿势来重新渲染由AR装置10提供的AR对象(或AR场景),使得向用户提供了在现实世界场景中存在AR对象(或AR场景)的无缝视角。

[0043] 根据示例实施例,AR装置10的部分内部空间可以配备有电子硬件(HW) 100,其包括各种类型的感测模块,诸如惯性测量单元(IMU) 传感器、同步定位与建图(SLAM) 模块、加速度计或指南针,诸如相机的图像捕获模块、麦克风、GPS模块、通信接口、处理单元或电池。根据示例实施例,可以在AR装置10的部分内部空间配备所述电子HW。AR装置10可以进一步配备有显示AR对象的光学引擎或光学元件。图1示出电子HW 100设置在AR装置10的一侧的框架中,电子HW 100在AR装置10中的位置不限于此,并且电子HW 100可以设置在AR装置10上的各个位置处,并且硬件组成元件中的每一个可以设置在同一空间中或不同空间中。

[0044] 图2是根据实施例的设置在AR装置中的电子硬件的框图。

[0045] 参考图2,电子HW 100可以包括IMU传感器110、相机120、视觉惯性同步定位与建图(VI-SLAM) 模块130、处理器140以及存储器150。在图2的电子HW 100中,为了便于解释,主要示出与执行姿势预测的实施例相关的硬件组成元件。然而,根据一些示例实施例,电子HW 100可以进一步包括其他硬件组成元件,诸如麦克风、全球定位系统(GPS) 模块、通信接口或光学引擎或其他通用组成元件。

[0046] IMU传感器110是用于测量AR装置10的速度、方向、重力或加速度的传感器。IMU传感器110可以通过使用陀螺仪、加速计或地磁场来测量AR装置10在三维(3D) 空间中的自由移动,使得可以测量前进方向、横向方向和高度方向上的加速度、以及横滚、俯仰和偏航角

速度。可以通过对从IMU传感器110获得的加速度和角速度求积分,计算AR装置10的速度、相对位置或姿势角度。

[0047] IMU传感器110可以以第一频率的帧速率(即,IMU速率)获得对应于图1的AR装置10移动的IMU值。换句话说,可以从IMU传感器110获得IMU值的时间序列。

[0048] 图3是根据示例实施例的用于经由IMU传感器获得IMU值的时间点的时序图。参考图3,IMU传感器110可以以第一频率的IMU速率测量在前进方向(X轴)、横向方向(Y轴)和高度方向(Z轴)上的加速度,以及横滚、俯仰和偏航角速度,并且输出所测量的IMU值。根据示例实施例,IMU速率可以是500HZ,但是实施例不限于此。在图3的时序图300中,每个竖线表示获得IMU值的时间点。例如,在时间点I获得的IMU值可以表示为 IMU_i ,...,在时间点L获得的IMU值可以表示为 IMU_L ,在时间点L+1获得的IMU值可以表示为 IMU_{L+1} ,并且在时间点L+2获得的IMU值可以表示为 IMU_{L+2} ,其中i和L是自然数。

[0049] 再次参考图2,相机120是用于拍摄AR装置10周围环境的真实场景的图像捕获装置。例如,相机120可以是配置为捕获AR设备10前面的周围环境的前置相机。相机120以第二频率的帧速率获得AR装置10周围的图像。换句话说,可以从相机120获得AR装置10周围的真实场景的图像的时间序列。

[0050] 图4是根据示例实施例的用于经由相机获得图像的时间点的时序图。参考图4,相机120以第二频率的帧速率获得AR装置10的周围图像。根据示例实施例,帧速率可以是30fps,但是实施例不限于此。从相机120获得图像的第二频率低于从IMU传感器110获得IMU值的第一频率。换句话说,第一频率的IMU速率是相对较高的频率,并且第二频率的帧速率是相对较低的频率。在图4的时序图400中,每个竖线表示获得图像的时间点。例如,在时间点0获得的图像可以表示为 I_0 ,...,在时间点n获得的图像可以表示为 I_n ,在时间点N-1获得的图像可以表示为 I_{N-1} ,并且在时间点N获得的图像可以表示为 I_N ,其中n和N是自然数。

[0051] 再参考图2,根据示例实施例,VI-SLAM模块130是通过融合视觉数据和惯性数据(即,视觉-惯性融合)执行用于估计相机对于地图的相对位置的计算的硬件组件(如,硬件处理器)。

[0052] 当仅使用从IMU传感器110获得的IMU速率的IMU值来估计姿势时,由于获得的IMU值的噪声和缓慢变化的偏差,姿势可能不准确。此外,当通过仅用从相机120获得的图像执行SLAM处理来估计姿势时,可能难以执行高速姿势计算。因此,VI-SLAM模块130可以通过执行误差值补偿的VI-SLAM方法中的基于视觉惯性融合的姿势估计来提高SLAM处理的准确度,所述VI-SLAM方法通过使用从IMU传感器110获得的IMU值和从相机120获得的图像来补偿误差值。

[0053] 例如,由于各种因素,诸如由于平坦区域(例如墙壁、光线不足、运动模糊/聚焦问题、高速相机运动或手抖)导致的遮挡,可能无法计算良好的相机姿势。然而,在示例实施例中,由于执行通过基于视觉惯性融合的深度神经网络(DNN)来学习相机姿势,即使在无法获得适当相机图像的情况下,也可以提高AR装置10的姿势预测的准确度。

[0054] VI-SLAM模块130可以基于由相机120以第二频率的帧速率获得的图像(图像序列)和从IMU传感器110获得的IMU值(IMU序列)的输入来估计AR装置10的中间6自由度姿势(中间6D姿势)。“中间”6D姿势是指在预测最终6D姿势(即,相对6D姿势)之前根据中间计算结果的姿势。

[0055] 在一个实施例中,每当以第二频率的帧速率获得图像时,VI-SLAM模块130就可以执行中间6D姿势的估计。在这种情况下,可以以第二频率的帧速率执行中间6D姿势的估计。换句话说,这种情况下的中间6D姿势可以是第二频率的帧速率的第一类型姿势数据,其通过执行用于将IMU值和以第二频率的帧速率获得的图像融合的SLAM技术而估计。

[0056] 此外,在另一示例实施例中,在以第二频率图像的帧速率获得图像的时间点之间,VI-SLAM模块130可以通过以下方式以第一频率的IMU速率执行中间6D姿势的估计:通过根据以帧速率获得图像的时间点之间的IMU值的预积分方法计算角速度的积分值和每个轴的加速度的积分值。换句话说,这种情况下的中间6D姿势可以是IMU速率的第二类型姿势数据,其包括对应于积分值的姿势数据,所述积分值通过根据IMU值和通过执行上述SLAM技术估计的帧速率的基于图像的姿势数据的预积分方法计算。

[0057] 图5A是根据示例实施例的用于估计VI-SLAM模块中的中间6D姿势的时间点的时序图。参考图5A,VI-SLAM模块130可以接收以IMU速率获得的IMU值520的输入和以帧速率获得的图像,并且以帧速率估计中间6D姿势510。

[0058] 详细地,VI-SLAM模块130可以通过在计算时间 Δ_{TRK} 期间用在时间点0获得的图像 I_0 以及在时间点 A'_0 与时间点 A_0 之间获得的IMU值执行计算来获得中间6D姿势 P_0 。尽管中间6D姿势 P_0 的计算在时间点 A_0 完成,中间6D姿势 P_0 可以被认为是对应于时间点 A'_0 的姿势。以相同的方式,VI-SLAM模块130可以通过在计算时间 Δ_{TRK} 期间的计算来获得中间6D姿势 P_n 、中间6D姿势 P_{N-1} 和中间6D姿势 P_N 中的每一个,并且获得的中间6D姿势 P_n 、中间6D姿势 P_{N-1} 和中间6D姿势 P_N 可以被认为分别对应于时间点 A'_n 的姿势、时间点 A'_{N-1} 的姿势以及时间点 A'_N 的姿势。如下所述,每当VI-SLAM模块130完成中间6D姿势的估计时,所估计的中间6D姿势在拍摄图像的实际时间点,例如时间点 A'_0 、...时间点 A'_n 、时间点 A'_{N-1} 、时间点 A'_N 、被更新到地面实况(GT)并且可以用于学习。

[0059] 图5B是根据另一示例实施例的用于经由VI-SLAM模块估计中间6D姿势的时间点的时序图。参考图5B,与图5A不同,VI-SLAM模块130以IMU速率估计中间6D姿势515。

[0060] 详细地,VI-SLAM模块130通过对以IMU速率获得的IMU值520执行两次积分来获得每个轴的加速度的积分值和角速度的积分值。积分值可以对应于从IMU值520单个估计的6D姿势。换句话说,VI-SLAM模块130通过对IMU值520执行预积分基于以帧速率获得的中间6D姿势(P_0 、 P_n 、 P_{N-1} 、 P_N ...)之间的时间点处的IMU积分来预测以IMU速率获得的中间6D姿势($\dots$ 、 P'_{n-1} 、 P'_{n-2} 、 P'_{n-3} 、...、 P'_{N-1} 、 P'_{N-2} 、 P'_{N-3} 、...)。这样,由于除了通过SLAM过程估计②的基于相机图像的中间6D姿势(P_0 、 P_n 、 P_{N-1} 、 P_N ...)之外还通过预积分过程①估计了基于IMU积分的中间6D姿势($\dots$ 、 P'_{n-1} 、 P'_{n-2} 、 P'_{n-3} 、...、 P'_{N-1} 、 P'_{N-2} 、 P'_{N-3} 、...),即使在由于诸如遮挡或模糊的问题可能无法平稳使用相机图像时,VI-SLAM模块130可以连续输出中间6D姿势515的估计结果。

[0061] 再参考图2,处理器140是可以运行操作系统或应用程序来控制AR装置10的整体操作,并且可以执行包括姿势预测和AR图像生成的各种数据处理和操作的硬件组件。处理器140可以包括例如中央处理单元(CPU)、微处理器、图形处理单元(GPU)、专用集成电路(ASIC)、数字信号处理器(DSP)、数字信号处理装置(DSPD)、可编程逻辑装置(PLD)、现场可编程门阵列(FPGA)、应用处理器(AP)、神经处理单元(NPU)或张量处理单元(TPU),但是实施例不限于此。

[0062] 处理器140生成姿势预测模型,所述姿势预测模型通过基于中间6D姿势和通过使用DNN获得的IMU值的输入执行学习来预测AR装置10的相对6D姿势。“相对”6D姿势可以意指中间先前预测的6D姿势(原始姿势)与当前预测的6D姿势(变换姿势)之间的姿势差异。另外相对6D姿势可以意指通过由中间6D姿势计算最终获得的最终姿势。

[0063] 根据示例实施例,姿势预测模型可以利用包括IMU速率的IMU输入和帧速率的中间6D姿势输入的多速率输入来执行DNN学习。换句话说,由于IMU速率是高于帧速率的频率,姿势预测模型可以利用不同频率的输入来执行用于姿势预测的学习和用于姿势预测的推断。

[0064] 此外,根据另一示例实施例,姿势预测模型可以用中间6D姿势输入的多速率输入来执行DNN学习,所述多速率输入包括IMU速率的IMU输入、IMU速率的基于IMU积分的中间6D姿势输入,以及帧速率的基于图像的中间6D姿势输入。

[0065] 由于输入到姿势预测模型的IMU输入和中间6D姿势输入对应于时间数据序列,DNN可以使用长短记忆(LSTM)或门控递归单元(GRU)实现为递归神经网络(RNN)。姿势预测模型可以是基于图5A中上述的GT姿势的更新以自我监督方法执行DNN学习的模型。

[0066] 当根据AR装置10的移动新获得的IMU值和根据AR装置10的移动新估计的中间6D姿势被输入到姿势预测模型的学习到的DNN时,处理器140可以通过学习到的DNN对姿势预测模型的处理实时地输出预测相对6D姿势的推断结果。

[0067] 处理器140可以通过使用在使用AR装置10期间获得的IMU值和相机图像、或者VI-SLAM模块130在特定时段内不使用AR装置10的状态(例如,在充电状态)下计算的6D姿势来执行DNN学习。当用户再次佩戴AR装置10来使用时,处理器140可以完成DNN学习并执行DNN的推断以预测AR装置10的姿势。

[0068] 存储器150可以包括诸如易失性存储器或非易失性存储器的内部存储器。存储器150可以在处理器140的控制下存储用于驱动和控制AR装置10的各种数据、程序或应用,以及输入/输出信号或AR数据。此外,存储器150可以存储与DNN相关的数据。存储器150可以对应于存储器装置,诸如随机存取存储器(RAM)、只读存储器(ROM)、硬盘驱动器(HDD)、固态驱动器(SSD)、紧凑型闪存(CF)、安全数字(SD)、微型安全数字(微-SD)、迷你安全数字(迷你-SD)、极限数字(xD)、存储棒等,并且存储器150的类型不限于此,并且可以是各样的。

[0069] 图6是根据实施例的生成姿势预测模型的方法的时序图。

[0070] 参考图6,处理器140接收从图2的IMU传感器110获得的IMU值520和由图2的VI-SLAM模块130估计的中间6D姿势510作为输入。

[0071] 具体地,处理器140可以利用以滑动方式移动的窗口的输入对DNN执行用于姿势预测的学习。以滑动方式移动的窗口可以包括 k 个IMU值,其中 k 是自然数,以及 m 个中间6D姿势,其中 m 是自然数。例如,窗口1610(WINDOW1610)可以被设置为包括从 IMU_i 到 IMU_L 的IMU值和从 P_n 到 P_N 的中间6D姿势。换句话说,DNN利用包括IMU速率的IMU输入和帧速率的中间6D姿势输入的多速率输入来执行学习。

[0072] 处理器140可以通过输入到DNN的WINDOW1610的处理来输出AR装置10的相对6D姿势 O_1 。

[0073] 图7A是根据示例实施例的通过滑动窗口方法在姿势预测模型中执行姿势预测的方法的时序图。

[0074] 参考图7A,当通过对窗口1610的处理来预测相对6D姿势 O_1 时,窗口可以基于IMU速

率滑动。换句话说,窗口可以以时间点为单位在时间轴上滑动以获得一个IMU值。

[0075] 处理器140可以通过使用DNN对WINDOWS-1620滑动的处理来输出AR装置10的相对6D姿势 O_{s-1} ,并且通过使用DNN对WINDOWS630滑动的处理来输出AR装置10的相对6D姿势 O_s 。换句话说,基于DNN的姿势预测模型可以接收多速率(IMU速率和帧速率)的输入,但是以IMU速率输出对应于每个窗口的相对6D姿势。

[0076] 基于DNN的姿势预测模型在执行预测相对6D姿势的学习时,可以基于GT姿势710的更新以自监督学习方法(或非监督学习方法)执行学习。例如,在DNN学习期间,可以将与输出窗口的时刻匹配的的时刻的中间6D姿势用作用于输出窗口的GT姿势710。

[0077] 每当由图2的VI-SLAM模块130完成对中间6D姿势的估计时,可以通过在拍摄图像的实时时间点将估计的中间6D姿势更新为GT姿势(P_0 (GT) ... P_n (GT) ... P_{M-1} (GT) ... P_N (GT) ...)来设置GT姿势710。

[0078] DNN可以通过使用更新的GT姿势减少损失函数来执行随时间的反向传播(BPTT),并因此将DNN的权重参数更新为适当的值,使得可以提高基于DNN的姿势预测模型的姿势预测准确度。

[0079] 图7B是根据另一示例实施例的通过滑动窗口方法在姿势预测模型中执行姿势预测的方法的时序图。

[0080] 参考图7B,与图7A不同,处理器140接收中间6D姿势输入515作为DNN的多速率输入,所述中间6D姿势输入包括IMU值520、IMU速率的基于IMU积分的中间6D姿势输入(...、 P'_{n-1} 、 P'_{n-2} 、 P'_{n-3} ...、 P'_{N-1} 、 P'_{N-2} 、 P'_{N-3} ...)以及帧速率的基于图像的中间6D姿势输入(P_0 、 P_n 、 P_{N-1} 、 P_N)。因此,可以在比图7A的GT姿势710的更新短的时间间隔处更新GT姿势720,并因此即使在可能由于诸如遮挡或模糊的问题无法平稳使用相机图像的时刻,DDN也可以相对准确地输出相对6D姿势的预测结果。

[0081] 图8是根据示例实施例的基于DNN的由RNN使用LSTM实现的姿势预测模型的图。

[0082] 根据示例实施例,由于用于图1的AR装置10中的姿势预测的输入对应于时间数据序列,姿势预测模型可以由RNN实现。然而,即使这样,姿势预测模型也不一定限于基于RNN的神经网络,并且可以采用不同类型的神经网络。

[0083] RNN是一种对使用过去的信息来预测当前信息有用的神经网络。然而,由于RNN有诸如长期依赖的问题,引入了LSTM类型的RNN来解决所述问题。LSTM可以分别包括单元状态和隐藏层。在LSTM中,单元状态管理信息存储器并擦除不重要的信息(忘记门),并且隐藏层仅将重要信息更新到单元(输入门),使得可以解决RNN的长期依赖性问题。

[0084] 参考图8,示出了展开的多速率LSTM。例如500Hz的第一频率的IMU速率的IMU值和第二频率的帧速率(例如30fps)的中间6D姿势被输入到LSTM。这样,在LSTM利用多速率输入执行姿势预测的学习时,如虚线所示,存储低频(第二频率)的速率的输入(中间6D姿势)的LSTM可以在学习中间改变到未激活状态。LSTM中存储的过去IMU值和过去中间6D姿势的输入可以传输到要与在不同时刻预测相对6D姿势相关的其他LSTM的单元。

[0085] 在由LSTM型RNN执行姿势预测学习姿势预测模型中,由于通过使用更新的GT姿势(P_0 ... P_N ...)执行BPTT来减少损失函数,可以提高基于LSTM的姿势预测模型的姿势预测准确度。

[0086] 根据另一示例实施例,基于DNN的姿势预测模型可以由RNN使用门控递归单元

(GRU)代替LSTM来实现,并且本领域技术人员可以理解,GRU方法可以与上述关于LSTM的方法类似地操作。

[0087] 图9和图10是根据示例实施例的在处理器中执行姿势预测的框图。

[0088] 首先参考图9,IMU传感器110将包括以IMU速率获得的IMU值的IMU序列①的时间数据的时间数据传输到VI-SLAM模块130,并且相机120将包括以帧速率获得的图像的图像序列①的时间数据传输到VI-SLAM模块130。IMU传感器110将IMU序列①传输到处理器140以进行姿势预测。

[0089] VI-SLAM模块130基于IMU序列①和图像序列①估计中间6D姿势,并且将估计的中间6D姿势①传输到处理器140。在这种情况下,VI-SLAM模块130在拍摄图像的实时时间点(提前量 Δ_{TRK})将估计的中间6D姿势①更新为GT姿势①,并且将更新的GT姿势①传输到处理器140。

[0090] 处理器140通过使用基于DNN的姿势预测模型来计算关于IMU序列①和中间6D姿势①的输出,并且将相对姿势①作为其结果输出。

[0091] 接着,参考图10,VI-SLAM模块130基于从IMU传感器110获得的IMU序列②和从相机120获得的图像序列估计中间6D姿势②,并且将估计的中间6D姿势②传输到处理器140。处理器140通过使用基于DNN的姿势预测模型来计算关于IMU序列②和中间6D姿势②的输出,并且将相对姿势②作为其结果输出。

[0092] 当预测了相对姿势②时,处理器140可以基于更新的GT姿势①和更新的GT姿势②,通过向后传播(例如,DNN的BPTT)来更新权重参数。根据示例实施例,处理器140可以重复执行上述过程以预测随后的相对姿势。

[0093] 在图9和图10中,符号①和②表示包括不同时间点的数据。

[0094] 图11A和图11B是根据一些实施例的基于已经完成学习的DNN在姿势预测模型中执行姿势预测的推断的时间图。

[0095] 参考图11A和图11B,基于输入到DNN的IMU值和中间6D姿势来执行姿势预测的学习,并且当DNN的参数由更新的GT姿势更新时,DNN学习完成,并因此可以生成姿势预测模型。当根据图1的AR装置10的移动新获得的IMU值和新估计的中间6D姿势输入到基于学习到的DNN生成的姿势预测模型时,图2的处理器140可以通过学习到的DNN对姿势预测模型的处理实时地输出预测相对6D姿势的推断结果。

[0096] 例如,如上所述,可以以滑动窗口方法将IMU值和中间6D姿势输入到学习到的DNN。当输入WINDOW11110时,处理器140通过在计算时间 Δ_R 期间处理WINDOW11110中的输入来执行相对6D姿势的预测结果的推断,并且输出对应于WINDOW11110的推断结果,即预测的相对6D姿势。同样地,处理器140关于输入具有滑动的WINDOW21120的情况和输入具有滑动的WINDOW31130的情况执行推断,并且输出推断结果,即,预测的相对6D姿势。

[0097] 由于DNN已基于输入IMU值和中间6D姿势完成了对于姿势预测的学习,处理器140不仅可以基于DNN学习的结果来推断AR装置10的当前和未来6D姿势,而且还输出AR装置10的未来6D姿势的预测结果。

[0098] 图12是根据示例实施例的预测AR装置中的姿势的方法的流程图。参考图12,由于图1的AR装置10中的姿势预测方法与上述图中的示例实施例相关,以上附图中描述的内容,即使在以下描述中省略了,也可以应用到图12的方法。

[0099] 在操作1201中,处理器140以第一频率的IMU速率从IMU传感器110获得对应于AR装置10的移动的IMU值。

[0100] 在操作1202中,处理器140可以从VI-SLAM模块130获得AR装置10的中间6D姿势,其中基于IMU值的输入以及由相机120以第二频率的帧速率获得的AR装置10周围的图像来估计所述中间6D姿势。

[0101] 在操作1203中,处理器140基于IMU值和中间6D姿势的输入使用DNN执行学习,从而生成预测AR装置10的相对6D姿势的姿势预测模型。

[0102] 上述示例实施例可以被编写为在计算机上可执行的程序,并且可以在使用计算机可读记录介质来操作程序的通用数字计算机中实现。此外,可以通过各种手段将在上述实施例中使用的数据的结构记录在计算机可读记录介质上。计算机可读记录介质可以包括诸如磁存储介质的存储介质,例如,ROM、软盘或硬盘,或者诸如CD-ROM或DVD的光学读取介质。

[0103] 应当理解,本文描述的示例实施例应当仅被认为是描述性意义而不是为了限制的目的。通常应当认为每个示例实施例中的特征或方面的描述可用于其他示例实施例中的其他类似特征或方面。虽然已参考附图描述一个或多个示例实施例,但是本领域普通技术人员将理解,在不脱离如由以下权利要求限定的精神和范围的情况下,可以在其中对形式和细节进行各种改变。

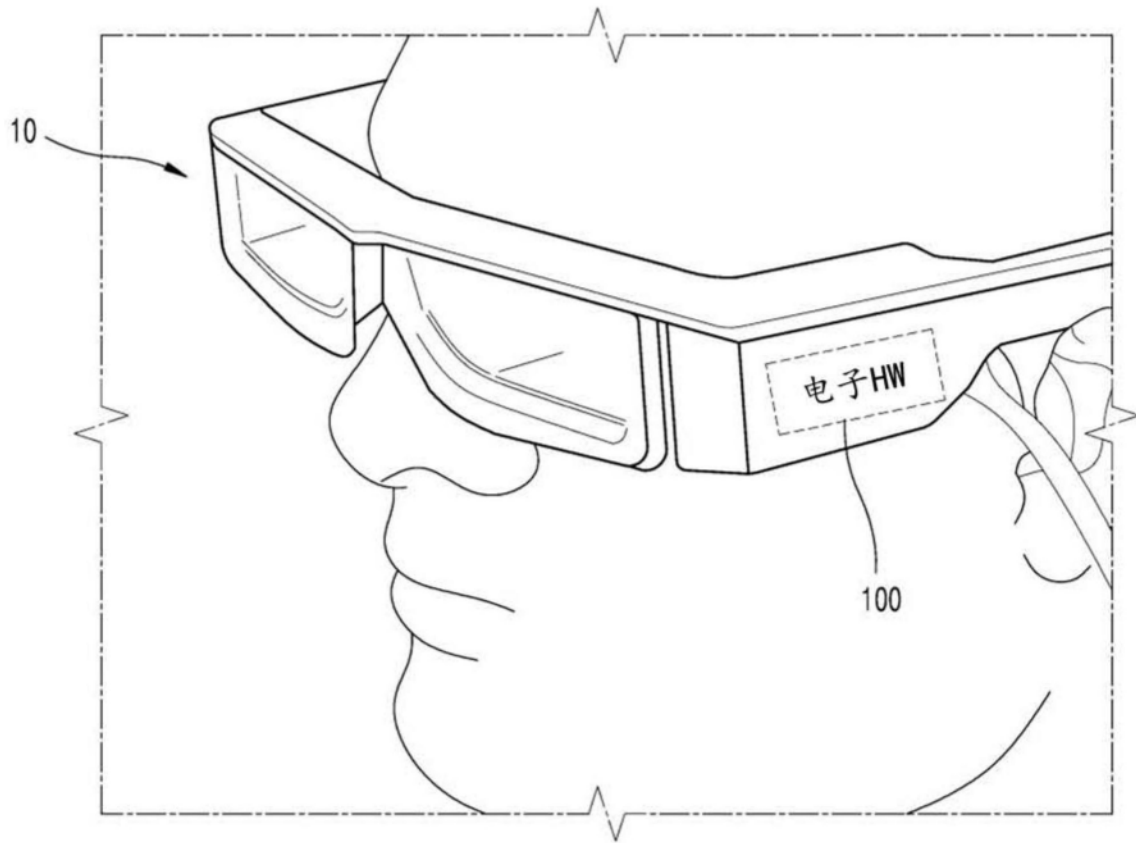


图1

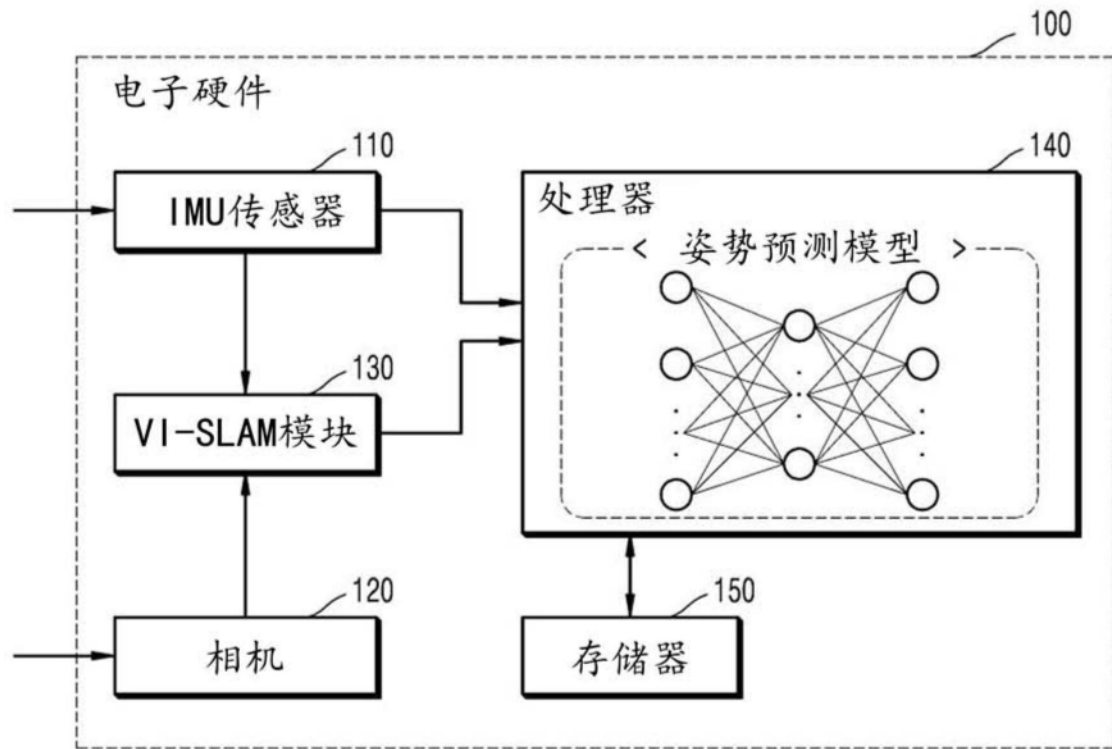


图2

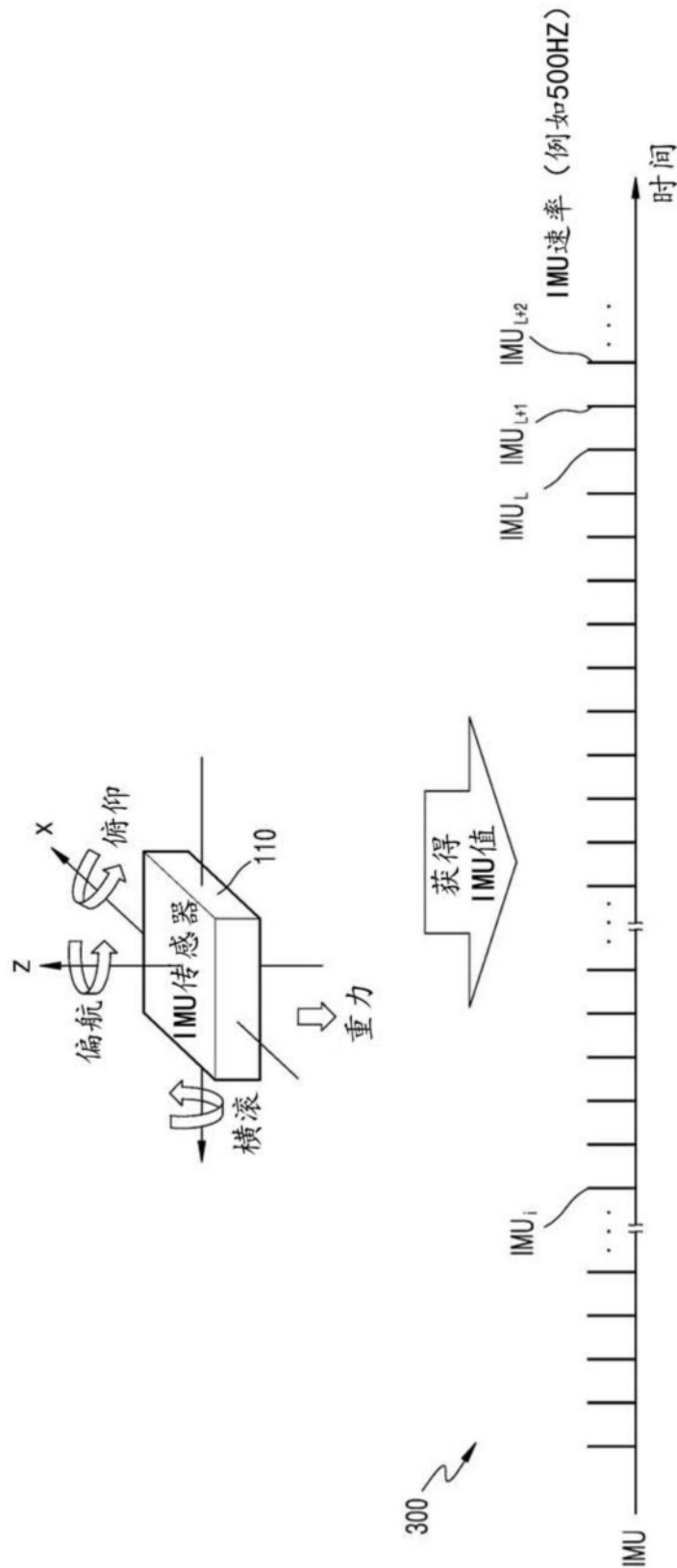


图3

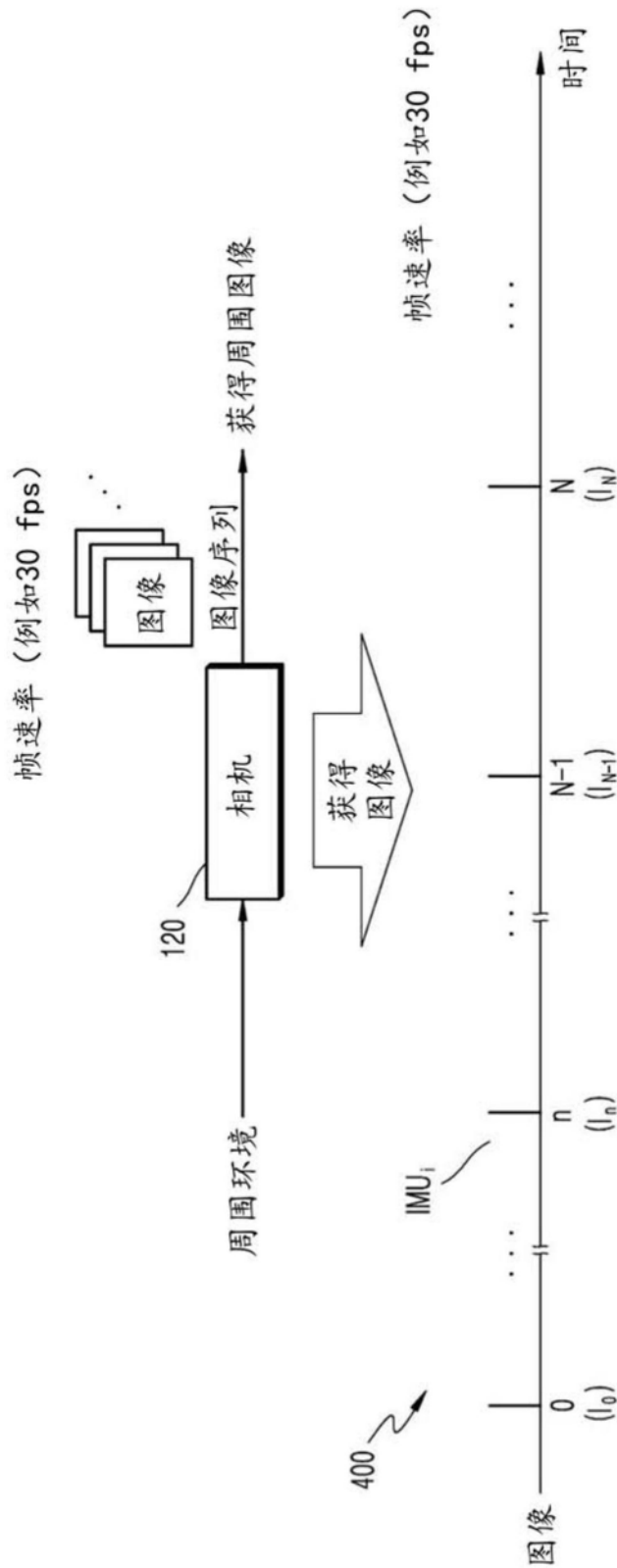


图4

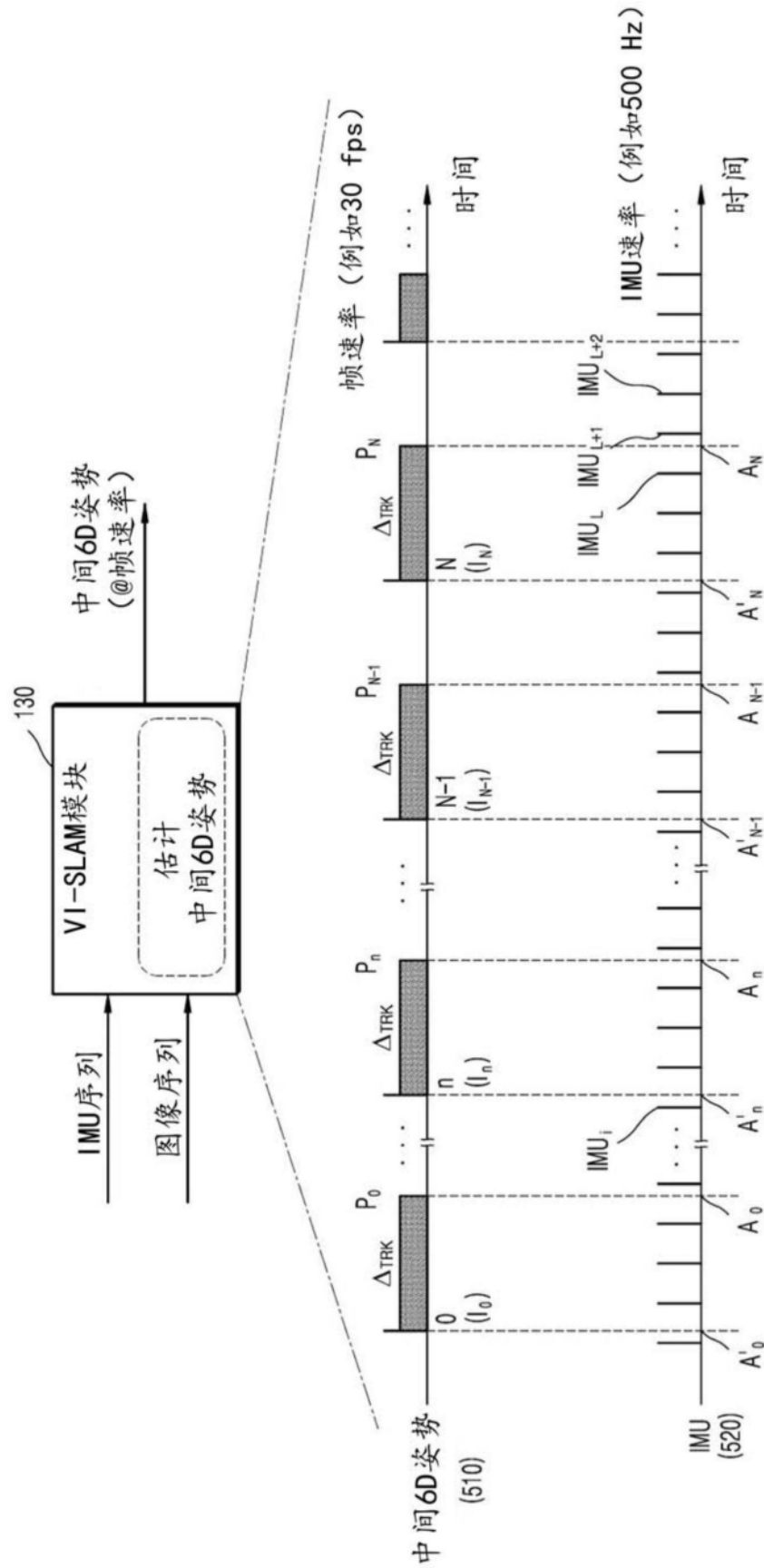


图5A

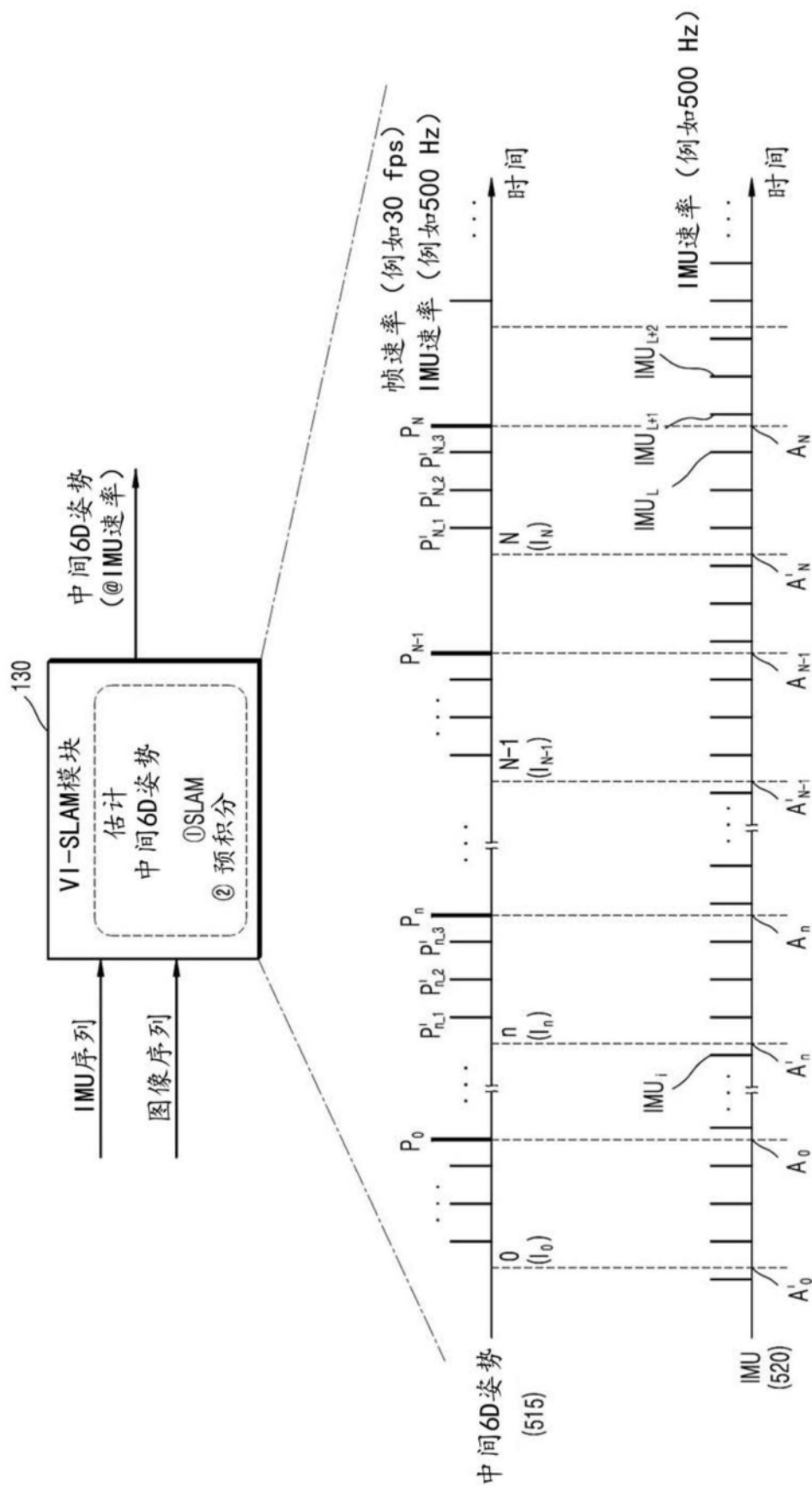


图5B

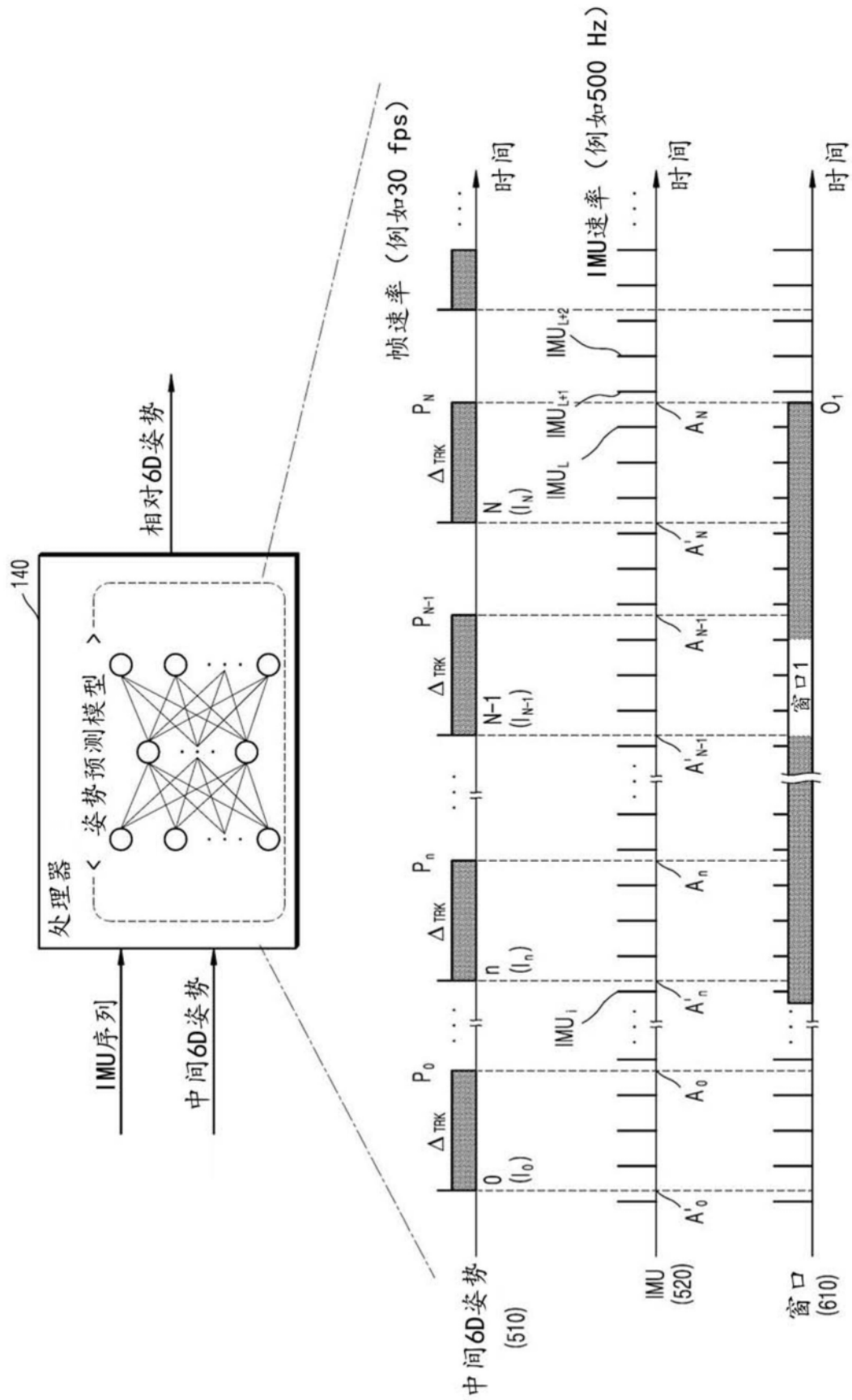


图6

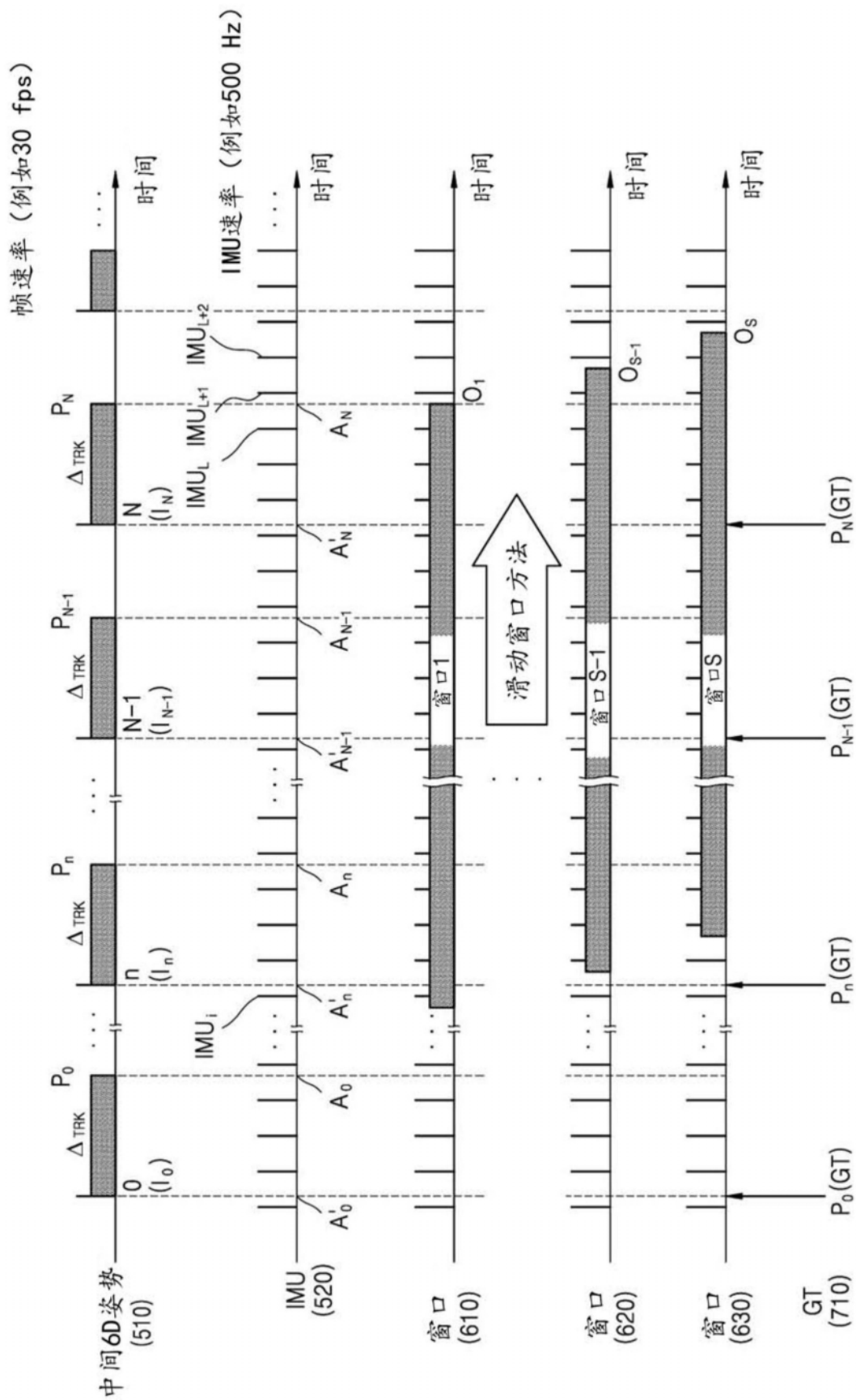


图7A

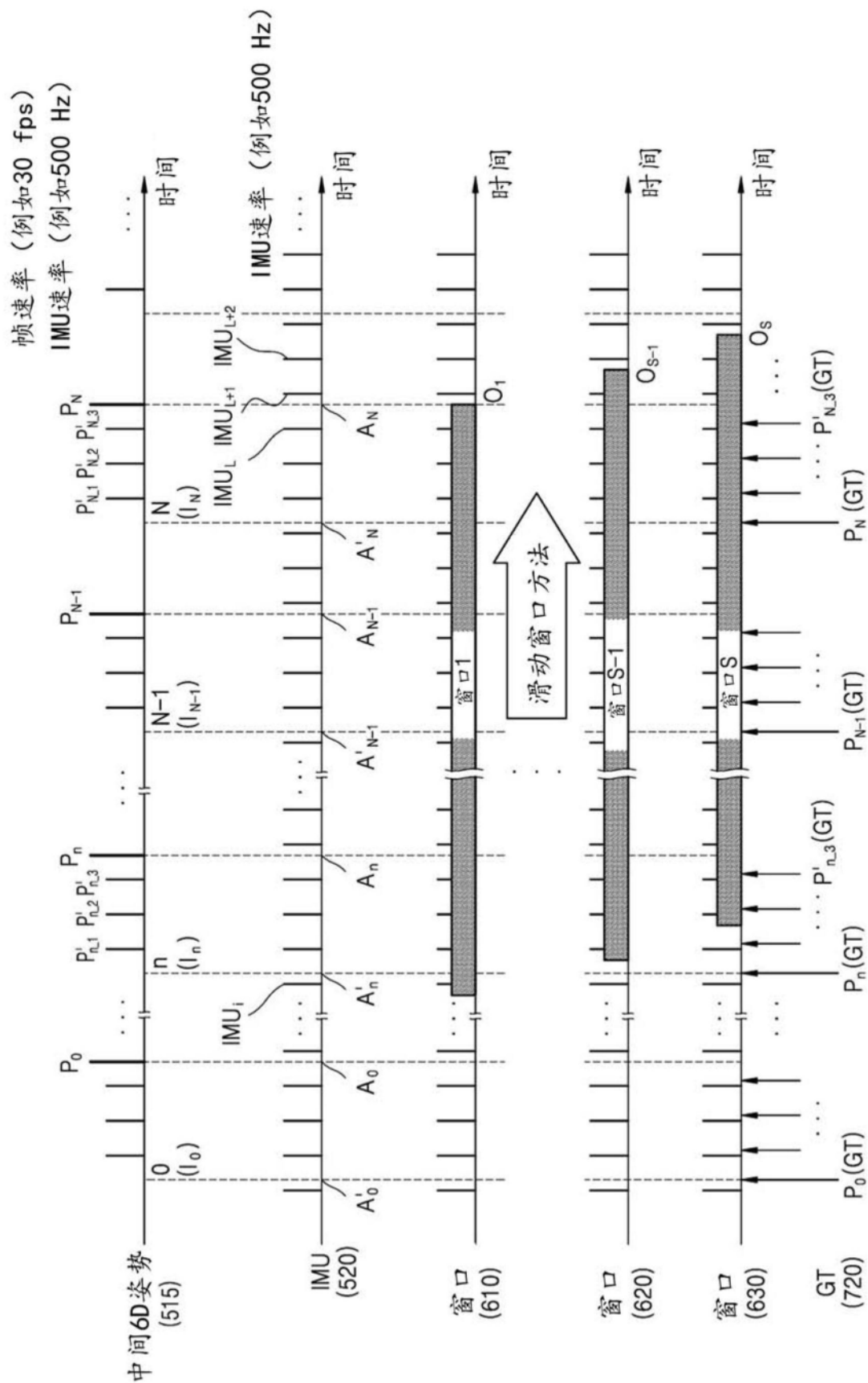


图7B

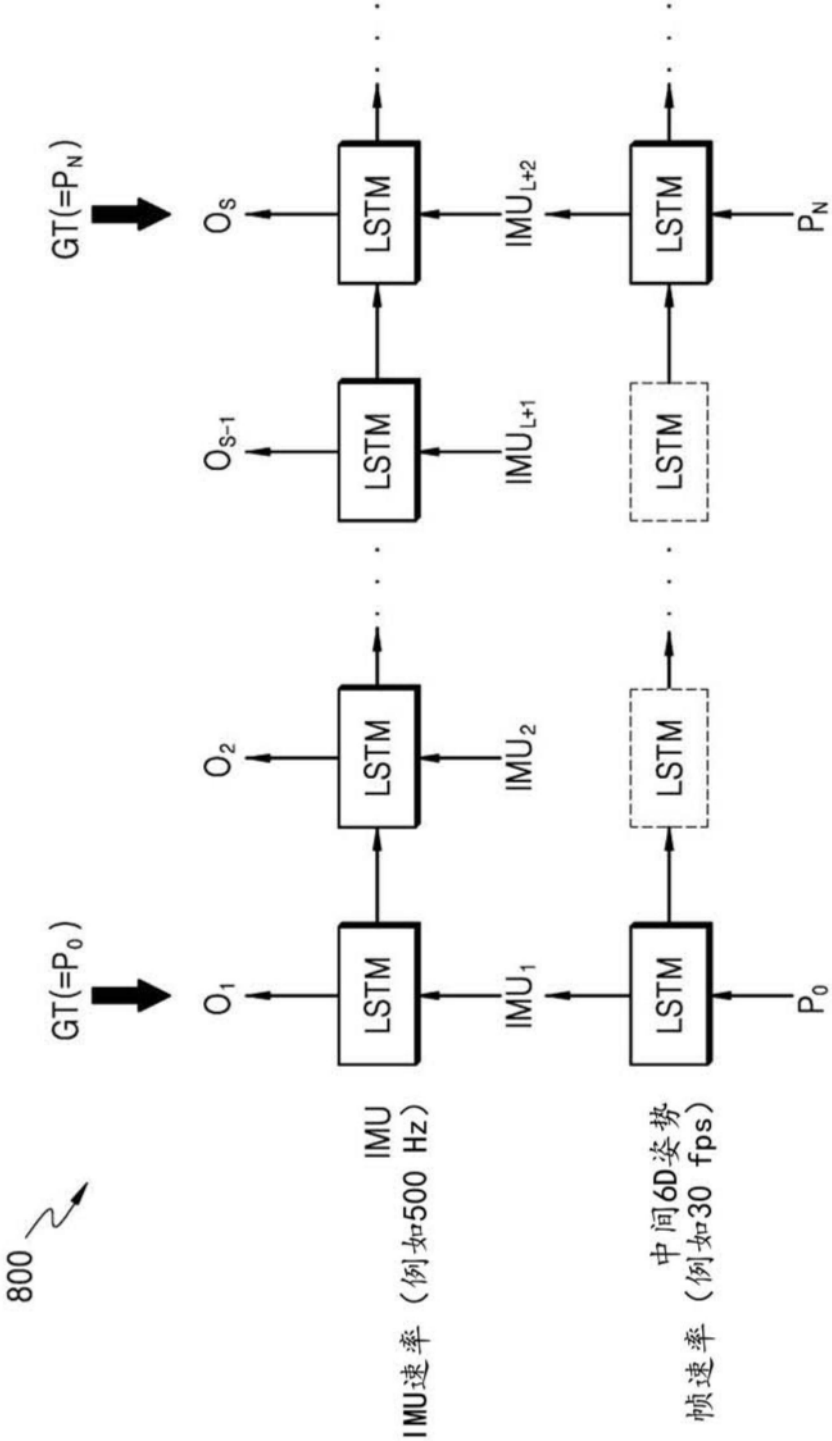


图8

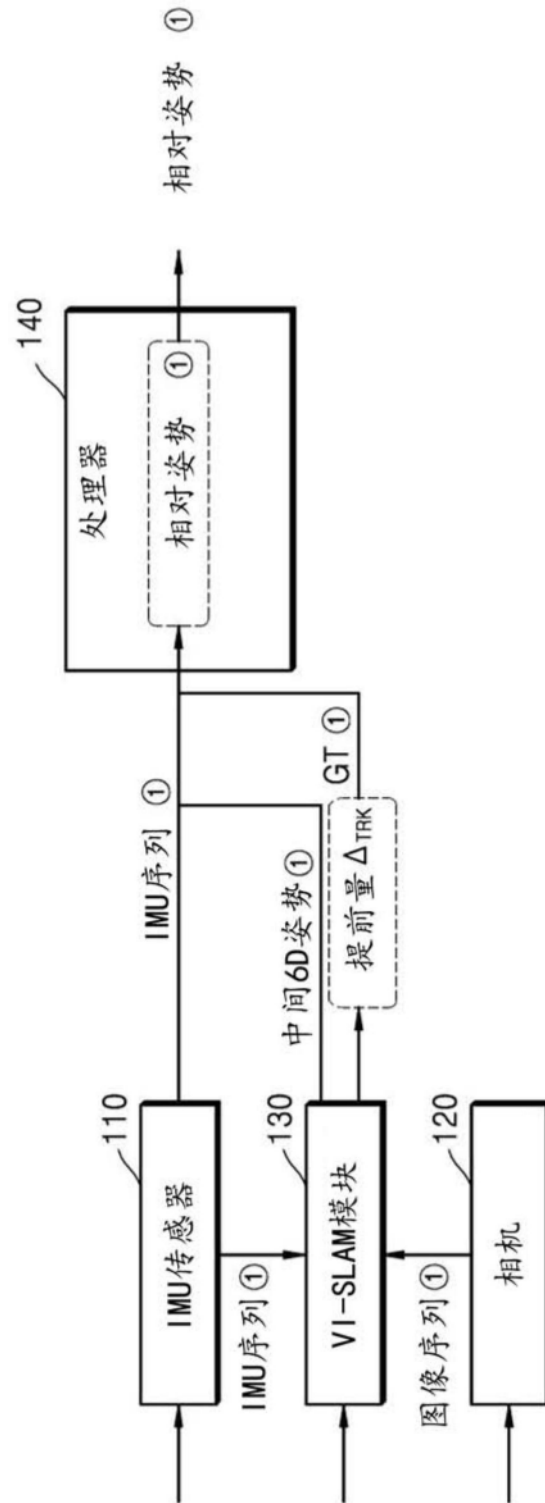


图9

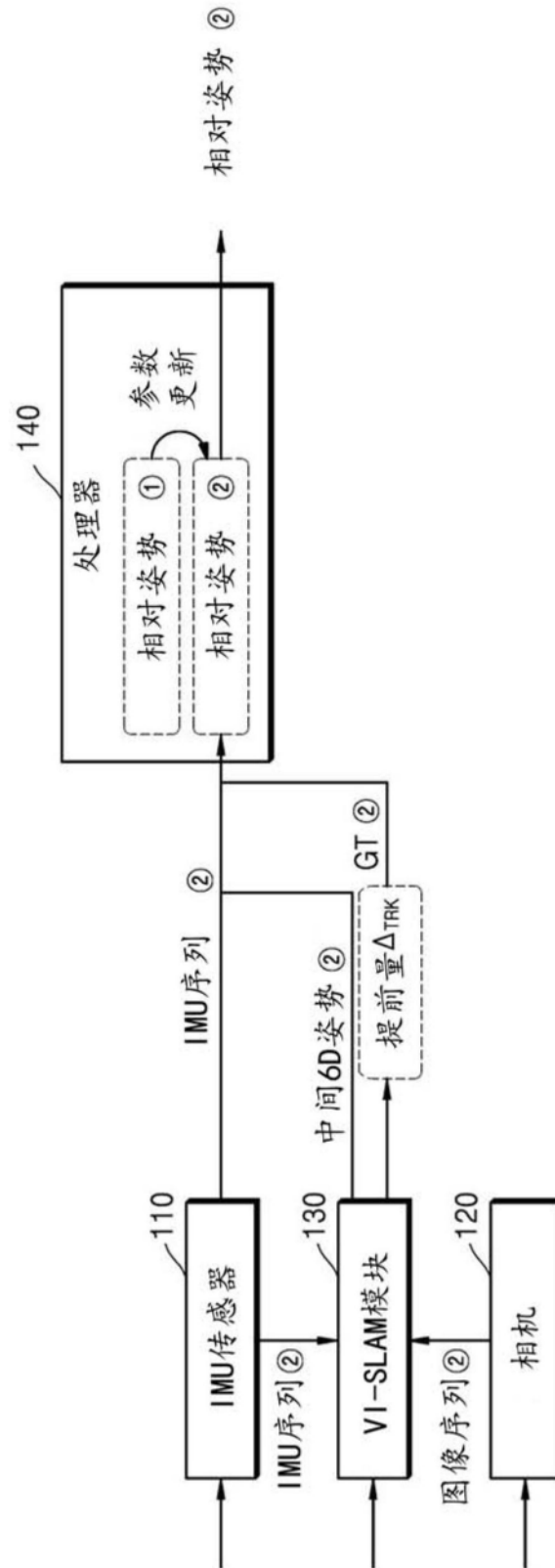


图10

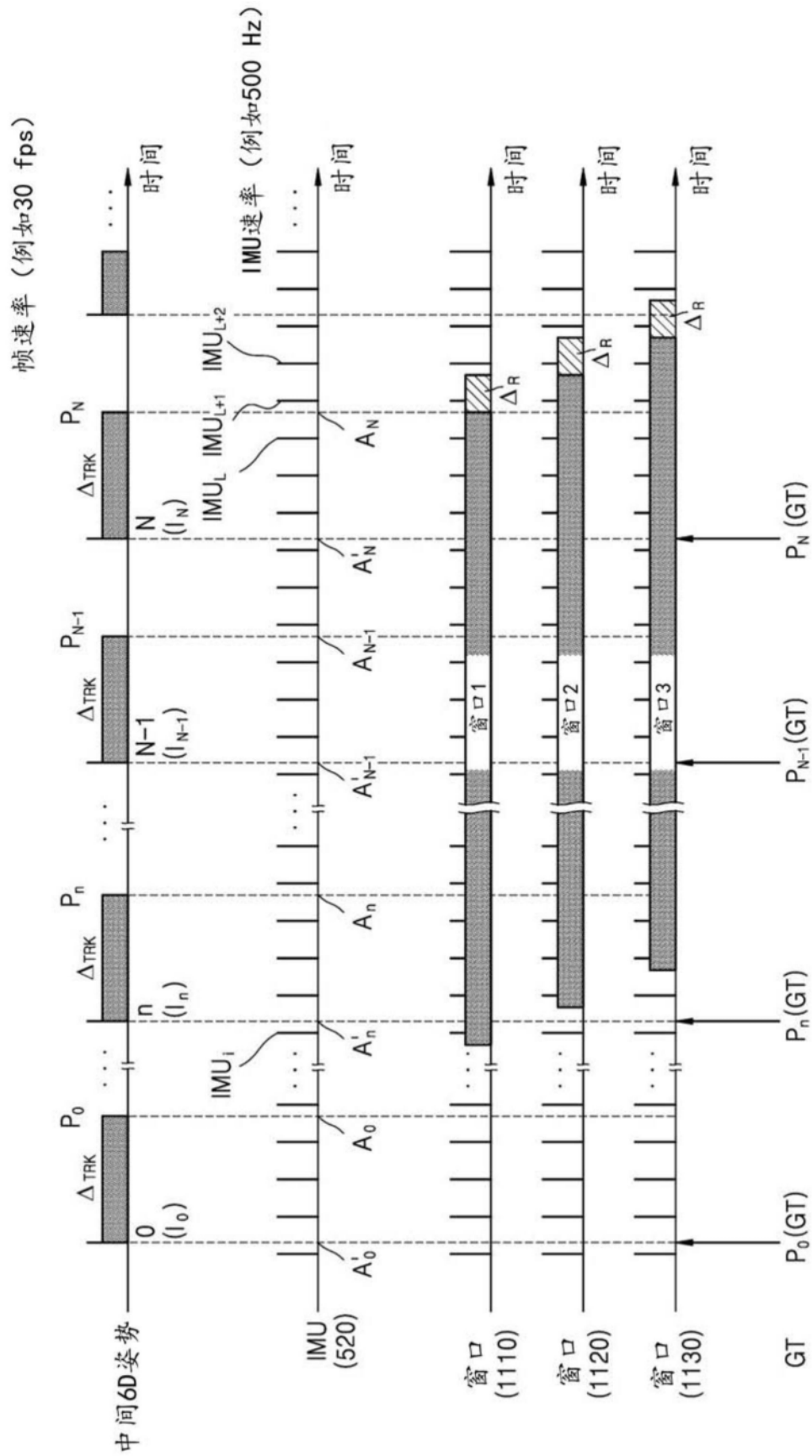


图11A

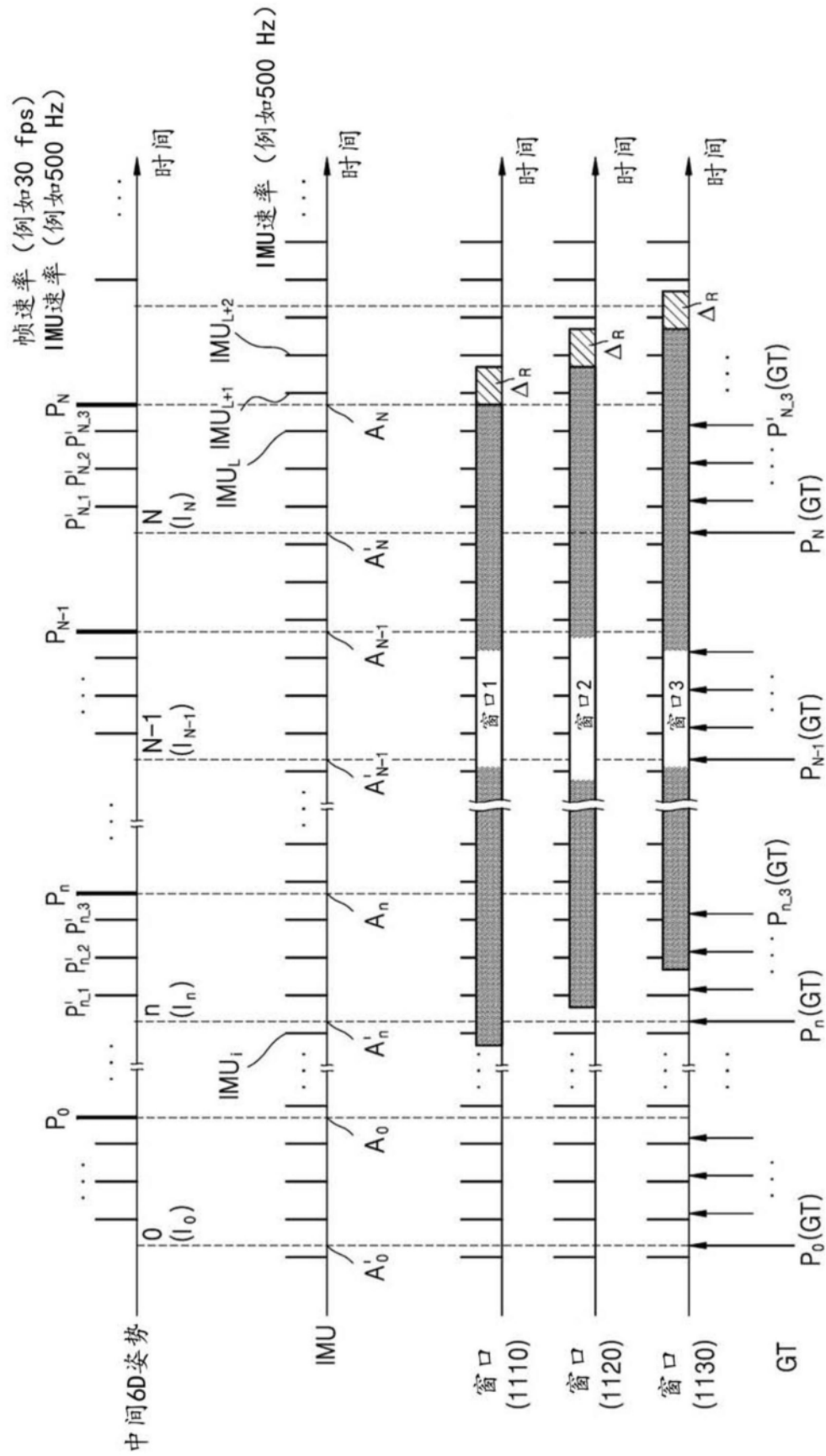


图11B



图12