



(12) 发明专利申请

(10) 申请公布号 CN 104160419 A

(43) 申请公布日 2014. 11. 19

(21) 申请号 201380003686. 3

(22) 申请日 2013. 09. 13

(30) 优先权数据

2013110494 2013. 03. 11 RU

(85) PCT国际申请进入国家阶段日

2014. 04. 29

(86) PCT国际申请的申请数据

PCT/US2013/059704 2013. 09. 13

(87) PCT国际申请的公布数据

WO2014/143154 EN 2014. 09. 18

(71) 申请人 LSI 公司

地址 美国加利福尼亚

(72) 发明人 P·A·阿里西奇克 I·L·马祖仁克

A·A·莱图诺夫斯基

A·A·佩蒂尤什克

A·B·霍洛多恩克

(74) 专利代理机构 中国国际贸易促进委员会专利商标事务所 11038

代理人 秦晨

(51) Int. Cl.

G06T 1/00(2006. 01)

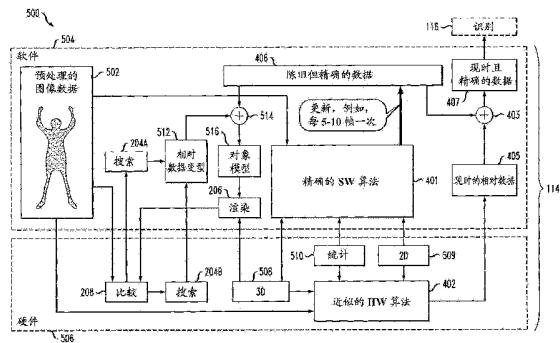
权利要求书2页 说明书10页 附图7页

(54) 发明名称

具有实现不同精度的软件及硬件算法的评价层的图像处理器

(57) 摘要

一种包括实现至少含有评价层和识别层的多个处理层的图像处理电路的图像处理器。评价层包括软件实现部分和硬件实现部分，评价层的软件实现部分被配置用于使用软件算法来生成第一精度等级的第一对象数据，而评价层的硬件实现部分被配置用于使用硬件算法来生成比第一精度等级低的第二精度等级的第二对象数据。评价层还包括被配置用于组合第一及第二对象数据以生成输出对象数据待传递给识别层的信号组合器。举例来说（仅为示例），评价层可以按照图像处理器的姿势识别系统的评价子系统的形式来实现。



1. 一种图像处理器,包括:
用于实现至少包括评价层和识别层的多个处理层的图像处理电路;
所述评价层包括软件实现部分和硬件实现部分;
所述评价层的所述软件实现部分被配置用于使用软件算法来生成第一精度等级的第一对象数据;
所述评价层的所述硬件实现部分被配置用于使用硬件算法来生成比所述第一精度等级低的第二精度等级的第二对象数据;
其中所述评价层还包括被配置用于组合所述第一及第二对象数据以生成用于传递给所述识别层的输出对象数据的信号组合器。
2. 根据权利要求 1 所述的图像处理器,其中所述评价层包括姿势识别系统的评价子系统。
3. 根据权利要求 1 所述的图像处理器,其中所述多个处理层还包括用于接收原始图像数据并将预处理的图像数据提供给所述评价层的预处理层。
4. 根据权利要求 1 所述的图像处理器,其中所述第一对象数据按第一速率更新,而所述第二对象数据按照比所述第一速率高的第二速率更新。
5. 根据权利要求 1 所述的图像处理器,其中所述第一速率小于所接收的图像数据的每个输入帧一次,而所述第二速率为每个输入帧一次。
6. 根据权利要求 1 所述的图像处理器,其中所述第一对象数据基于当前的输入帧与至少一个对象模型之间的比较而生成。
7. 根据权利要求 1 所述的图像处理器,其中所述第二对象数据基于当前的输入帧与至少一个先前输入帧之间的比较而生成。
8. 根据权利要求 1 所述的图像处理器,其中所述评价层在生成所述第一及第二对象数据时实现搜索-渲染-比较操作。
9. 根据权利要求 8 所述的图像处理器,其中与所述搜索-渲染-比较操作关联的单独搜索模块被实例化于所述软件实现部分和所述硬件实现部分中的每个中。
10. 根据权利要求 8 所述的图像处理器,其中与所述搜索-渲染-比较操作关联的渲染模块仅被实例化于所述软件实现部分内,而与与所述搜索-渲染-比较操作关联的比较模块仅被实例化于所述硬件实现部分内。
11. 根据权利要求 1 所述的图像处理器,其中所述评价层的所述软件实现部分实现多个高级处理操作,包括单点检测和跟踪、前表面检测和跟踪、边缘检测和跟踪以及快速试探例如质心检测中的一个或多个。
12. 根据权利要求 1 所述的图像处理器,其中所述评价层的所述硬件实现部分实现多个低级图像处理基元,包括诸如仿射或同度量变换的 3D 基元、诸如曲线跟踪的 2D 基元、诸如最小平方比较的比较基元、诸如运动预测或梯度搜索的搜索基元以及统计基元中的一个或多个。
13. 根据权利要求 1 所述的图像处理器,其中所述图像处理电路包括至少一个图形处理器集成电路。
14. 一种集成电路,包括根据权利要求 1 所述的图像处理器。
15. 一种方法,包括:

配置至少包括评价层和识别层的图像处理器的多个处理层；

在所述评价层的软件实现部分中使用软件算法来生成第一精度等级的第一对象数据；

在所述评价层的硬件实现部分中使用硬件算法来生成比所述第一精度等级低的第二精度等级的第二对象数据；以及

组合所述第一及第二对象数据以生成用于传递给所述识别层的输出对象数据。

16. 根据权利要求 15 所述的方法,其中生成第一及第二对象数据包括：

基于当前的输入帧与至少一个对象模型之间的比较来生成所述第一对象数据；以及

基于当前的输入帧与至少一个先前输入帧之间的比较来生成所述第二对象数据；

其中所述第一对象数据按第一速率更新,而所述第二对象数据按照比所述第一速率高的第二速率更新。

17. 一种具有实现于其内的计算机程序代码的计算机可读存储介质,其中所述计算机程序代码在被执行于处理设备内时促使所述处理设备执行根据权利要求 15 所述的方法。

18. 一种图像处理系统,包括：

提供图像数据的一个或多个图像源；

一个或多个图像目的地；以及

与所述一个或多个图像源以及所述一个或多个图像目的地耦接的图像处理器；

其中所述图像处理器包括：

用于实现至少包括评价层和识别层的多个处理层的图像处理电路；

所述评价层包括软件实现部分和硬件实现部分；

所述评价层的所述软件实现部分被配置用于使用软件算法来生成第一精度等级的第一对象数据；

所述评价层的所述硬件实现部分被配置用于使用硬件算法来生成比所述第一精度等级低的第二精度等级的第二对象数据；

其中所述评价层还包括被配置用于组合所述第一及第二对象数据以生成用于传递给所述识别层的输出对象数据的信号组合器。

19. 根据权利要求 18 所述的系统,其中所述一个或多个图像源中的至少一个包括深度传感器。

20. 根据权利要求 19 所述的系统,其中所述深度传感器是合并了所述图像处理器的深度成像器的一部分。

具有实现不同精度的软件及硬件算法的评价层的图像处理 器

技术领域

[0001] 本发明的领域一般地涉及图像处理,并且更特别地涉及图像数据在图像处理器中的处理。

背景技术

[0002] 图像处理在各种各样的不同应用中是重要的,并且这样的处理可以涉及二维(2D)图像、三维(3D)图像,或者多个不同类型的图像的组合。例如,空间场景的3D图像可以在图像处理器中使用三角剖分基于由使每个相机具有不同的场景视角而布置的各个相机捕获的多个2D图像来生成。作为选择,3D图像能够直接使用深度成像器(例如,结构光(SL)相机或飞行时间(ToF)相机)来生成。在此同样称为深度图像的这些及其他3D图像通常用于机器视觉应用(例如,姿势识别)中。

[0003] 在典型的常规布局中,来自图像传感器的原始图像数据通常要经过各种预处理操作。除了别的以外,这样的预处理操作还可以包括例如对比度增强、直方图均衡化、噪声降低、边缘突出显示以及坐标空间变换。所预处理的图像数据然后会受到为了实现姿势识别或一种机器视觉应用所需的附加处理。

发明内容

[0004] 在一种实施例中,图像处理器包括用于实现至少包括评价层和识别层的多个处理层的图像处理电路。评价层包括软件实现的部分和硬件实现的部分,评价层的软件实现部分被配置用于使用软件算法来生成第一精度等级的第一对象数据,而评价层的硬件实现部分被配置用于使用硬件算法来生成比第一精度等级低的第二精度等级的第二对象数据。评价层还包括被配置用于组合第一及第二对象数据以生成待传递给识别层的输出对象数据的信号组合器。

[0005] 评价层可以说明性地耦接于图像处理器的预处理层与识别层之间,该预处理层接收来自图像传感器的原始图像数据并且将预处理的图像数据提供给评价层,尽管众多别的布局都是可能的。

[0006] 举例来说(仅为示例),评价层可以按照图像处理器的姿势识别系统的评价子系统的形式来实现。

[0007] 本发明的其他实施例包括(但不限于)方法、装置、系统、处理设备、集成电路,以及具有实现于其内的计算机程序代码的计算机可读存储介质。

附图说明

[0008] 图1是在一种实施例中的图像处理系统的框图,包括配置有使用预处理层、评价层、识别层和应用层进行的姿势识别功能的图像处理器。

[0009] 图2至5示出了在图1的系统中的图像处理器的示例性评价层的各方面。

[0010] 图 6 至 8 是示出在图 1 的系统内的图像处理器的评价层的各个实施例中的各组处理操作的流程图,使用头部跟踪应用作为示例。

具体实施方式

[0011] 本发明的实施例在本文中结合示例性的图像处理系统来说明,该图像处理系统包括图像处理器或者使用多个处理层来实现姿势识别功能的其他类型的处理设备。但是,应当理解,本发明的实施例可更一般地应用于得益于更有效的图像处理的任何图像处理系统或者关联的设备或技术。

[0012] 图 1 示出了在本发明的一种实施例中的图像处理系统 100。图像处理系统 100 包括用于接收一个或多个图像源 105 的图像并将处理的图像提供给一个或多个图像目的地 107 的图像处理器 102。图像处理器 102 同样经由网络 104 与多个处理设备 106 通信。

[0013] 尽管图像源 105 和图像目的地 107 在图 1 中被显示为与处理设备 106 分离的,但是这样的源和目的地中的至少一个子集可以至少部分地使用一个或多个处理设备 106 来实现。因此,图像可以经由网络 104 提供给图像处理器 102,以便由一个或多个处理设备 106 处理。类似地,处理的图像可以由图像处理器 102 通过网络 104 传递给一个或多个处理设备 106。因此,可以将这样的处理设备视为图像源或图像目的地的实例。

[0014] 给定的图像源可以包括例如配置用于生成深度图像的 3D 成像器(例如,SL 相机或 ToF 相机),或者配置用于生成灰度图像、彩色图像、红外图像或其他类型的 2D 图像的 2D 成像器。同样可能的是,单个成像器或其他图像源能够提供深度图像和相应的 2D 图像(例如,灰度图像、彩色图像或红外图像)两者。例如,某些类型的现有 3D 相机能够产生给定场景的深度图,以及同一场景的 2D 图像。作为选择,用于提供给定场景的深度图的 3D 成像器能够布置于单独的高分辨率视频摄相机附近,或者是提供基本上相同场景的 2D 图像的其他 2D 成像器。

[0015] 还应当意识到,给定的图像源如同该术语在本文中所广泛使用那样可以代表用于并入图像处理器 102 的至少一部分的成像器的图像传感器部分。例如,一个或多个图像源 105 中的至少一个可以包括深度传感器,该深度传感器是 SL 相机、ToF 相机或者并入图像处理器 102 的其他深度成像器的一部分。许多可替换的布局都是可能的。例如,图像源的另一个实例是用于给图像处理器 102 提供待处理的图像的存储设备或服务器。

[0016] 给定的图像目的地可以包括例如计算机或移动电话的人机界面的一个或多个显示屏,或者用于接收来自图像处理器 102 的处理的图像的至少一个存储设备或服务器。

[0017] 因此,尽管图像源 105 和图像目的地 107 在图 1 中被示为与图像处理器 102 分离的,图像处理器 102 可以至少部分地与在共同处理设备上的该一个或多个图像源和该一个或多个图像目的地中的至少一个子集组合。因而,例如,给定的图像源和图像处理器 102 可以共同地实现于同一处理设备上。类似地,给定的图像目的地和图像处理器 102 可以共同地实现于同一处理设备上。

[0018] 在本实施例中,图像处理器 102 包括使用图像处理器 102 的多个处理层来实现的姿势识别 (GR) 系统 110。这些处理层包括在本文中更一般地称为图像处理器 102 的“图像处理电路”的至少一部分。在本实施例中,处理层在图中被示为各个椭圆形,并且包括预处理层 112 和多个较高的处理层,包括评价层 114、识别层 116 和应用层 118。

[0019] 处理层 112、114、116 和 118 在此还可以指的是 GR 系统 110 的各个子系统。但是，应当注意，本发明的实施例并不限于姿势识别，而是能够适合在各种各样的其他机器视觉应用中使用，并且在其他实施例中可以包括不同数量、类型及布局的层。

[0020] 此外，作为代替，图像处理器 102 的某些处理层在其他实施例中可以至少部分地实现于其他设备上。例如，预处理层 112 可以至少部分地实现于包括深度成像器或其他类型的成像器的图像源 105 内。同样可能的是，应用层 118 可以实现于与评价层 114 和识别层 116 不同的单独的处理设备上，例如，处理设备 106 之一。

[0021] 而且，应当意识到，图像处理器 102 自身可以包括多个不同的处理设备，使得处理层 112、114、116 和 118 使用两个或更多个处理设备来实现。本文所使用的术语“图像处理器”意指作广泛的理解，以便包含这些及其他布局。

[0022] 预处理层 112 对接收自一个或多个图像源 105 的图像数据执行预处理操作。在本实施例中假定，该接收到的图像数据包括接收自深度传感器的原始图像数据，但是在其他实施例中也可以处理其他类型的所接收的图像数据。预处理层 112 将预处理的图像数据提供给评价层 114。

[0023] 在预处理层 112 中接收自深度传感器的原始图像数据可以包括包含各自深度图像的帧流，每个这样的深度图像都包括多个深度图像像素。例如，给定的深度图像 D 可以按照实值矩阵的形式提供给预处理层 112。每个这样的实值都可以更具体地为深度图像的特定制素提供深度值 d_{ij} ，其中 i 和 j 指示像素索引，并且深度值表示到所成像的对象的距离。给定的这样的深度图像在本文中也称为深度图。

[0024] 具有索引 i, j 和深度值 d_{ij} 的给定像素能够被变换为 3D 空间中的 (x, y, z) 坐标。同样，如果给定像素的深度是未知的，则可以将预定值 u （例如，0）用作该像素的深度值。在其他实施例中可以使用各种各样的其他类型的图像数据。

[0025] 在某些实施例中，用于生成像素的深度值的传感器还可以为那些像素提供对应的可靠性值。例如，由该类型的传感器供应的每个像素 (i, j) 可以包括对 (d_{ij}, r_{ij}) ，其中 $0 \leq r_{ij} \leq 1$ 是深度图像的像素可靠性指标或者其他类型的可靠性值。作为选择，可靠性值可以在预处理层 112 中基于特定类型的传感器的已知特性而估计出或者被确定。可靠性值可以是用于一个或多个别的处理层中的单独的可靠性矩阵的一部分。

[0026] 在此说明，上述图像数据类型只是示例性的，并且所公开的技术能够适合用于许多不同类型的图像数据流，包括具有高帧频的流。

[0027] 如同以下将结合图 4 和 5 更详细地描述的，评价层 114 更具体地包括软件实现的部分和硬件实现的部分，评价层的软件实现部分被配置用于使用软件算法来生成第一精度等级的第一对象数据，而评价层的硬件实现部分被配置用于使用硬件算法来生成比第一精度等级低的第二精度等级的第二对象数据。评价层 114 还包括被配置用于组合第一及第二对象数据以生成待传递给识别层 116 的输出对象数据的信号组合器。对象数据可以具有多个部分，例如，常量部分和变量部分。

[0028] 举例来说，对象数据可以包括用于描述一个或多个成像对象的所谓的“骨架”数据，骨架数据的常量部分包括骨架尺寸，而骨架数据的变量部分包括骨架角度和质心位置。在其他实施例中可以使用许多其他类型的对象数据，并且这样的对象数据不需要包括单独的常量部分和变量部分。对象数据或其一些部分可以采取各种形式，例如，矩阵或向量形

式,如同本领域技术人员所应当意识到的。本文所使用的术语“对象数据”因此意指作广泛的理解,以便包含,例如,涉及在预处理的图像数据中检测到的或者与其关联的一个或多个对象的数据。

[0029] 识别层 116 在本实施例中使用由评价层 114 提供的对象数据来执行高级应用专用的图像处理,假定该图像处理至少包括姿势识别,但是除此之外或作为选择能够包括其他类型的高级应用专用的图像处理,例如,行为识别、表情识别和凝视跟踪。

[0030] 在本实施例中,假定应用层 118 包括使用识别层 116 的姿势识别输出的至少一个姿势识别应用,例如,成像场景的参数表示。其他类型的处理的图像数据输出可以提供给图像处理器 102 的一个或多个应用层或者相关的处理设备 106 或目的地 107。

[0031] 在上述骨架数据的情形中,这种数据由评价层 114 提供给识别层 116,该识别层 116 明确地将为特定的应用层 118 所需的当前对象状态进行分类。例如,如果由应用层实现的姿势识别应用包括响应于检测到各个不同的姿势而生成命令,则识别层 116 被配置用于检测这样的姿势并且给应用层提供适当的分类输出。

[0032] 处理层 112、114、116 和 118 可以包括图像处理器 102 的图像处理电路的不同部分,尽管给定的此类处理层可以被实现为硬件、固件和软件的组合。本文所使用的术语“层”因此意指作广泛的理解,并且可以包括,例如,专用硬件、处理内核、固件引擎及相关的固件,或者通用处理资源以及在这些资源上执行的相关软件,以及这些及其他类型图像处理电路的各种组合。假定给定的 GR 系统的子系统是处理层的示例,因为后一个术语在本文中被广泛地使用。

[0033] 另一种常规的图像处理集成电路或其他类型的图像处理电路可以被适当地修改以实现图像处理器 102 的处理层 112、114、116 和 118 中的一个或多个的至少一部分,如同本领域技术人员所应当意识到的。可以用于本发明的一种或多种实施例中的图像处理电路的一个可能实例是被适当地重新配置用于执行与处理层 112、114、116 和 118 中的一个或多个关联的功能的另一种常规的图形处理器。

[0034] 处理设备 106 可以任意组合地包括,例如,计算机、移动电话、服务器或存储设备。一个或多个这样的设备还可以包括,例如,显示屏或者用来呈现由图像处理器 102 生成的图像的其他用户界面。处理设备 106 可以因此包括被配置用于经由网络 104 来接收来自图像处理器 102 的处理的图像流或其他类型的输出信息的各种各样的不同目的地设备,包括例如用于接收来自图像处理器 102 的此类输出信息的至少一个服务器或存储设备。

[0035] “图像目的地”,如同该术语在本文中所广泛使用那样,意指包含用于接收从由图像处理器 102 处理的一个或多个图像得出的信息的任何设备或其部分,例如,用于接收由 GR 系统 110 生成的姿势识别输出的设备。因此,图像目的地可以接收至少一个处理的图像、至少一个处理的图像的一部分,或者从至少一个处理的图像的至少一部分得出的信息。

[0036] 尽管在本实施例中所示为与处理设备 106 分离的,图像处理器 102 可以至少部分与处理设备 106 中的一个或多个组合。因而,例如,图像处理器 102 可以至少部分地使用处理设备 106 中的给定一个来实现。举例来说,计算机或移动电话可以被配置用于并入图像处理器 102 以及可能为给定的图像源。因此,图像源 105 可以包括相机或者与计算机、移动电话或其他处理设备关联的其他成像器。如上所述,图像处理器 102 可以至少部分地与一个或多个图像源或图像目的地组合于共同处理设备上。

[0037] 在本实施例中,假定图像处理器 102 使用至少一个处理设备来实现并且包括与存储器 122 耦接的处理器 120。处理器 120 执行存储于存储器 122 内的软件代码,以便控制图像处理操作的性能。图像处理器 102 还包括用于支持经由网络 104 进行的通信的网络接口 124。

[0038] 处理层 112、114、116 和 118,以及处理器 120、存储器 122 和网络接口 124 被认为是在本文中更一般地称为图像处理器 102 的图像处理电路的电路的实例。

[0039] 处理器 120 可以任意组合地包括,例如,微处理器、专用集成电路 (ASIC)、现场可编程门阵列 (FPGA)、中央处理单元 (CPU)、算术逻辑单元 (ALU)、数字信号处理器 (DSP) 或其他类似的处理设备构件,以及其他类型和布局的图像处理电路。

[0040] 存储器 122 存储待由处理器 120 执行以实现图像处理器 102 的功能的若干部分的软件代码,例如,预处理层 112 的若干部分以及较高的处理层 114、116 和 118。用于存储待由相应的处理器执行的软件代码的此类给定的存储器是在本文中更一般地称为具有实现于其内的计算机程序代码的计算机可读介质或其他类型的计算机程序产品的存储器的实例,并且可以任意组合地包括,例如,电子存储器(例如,随机存取存储器 (RAM) 或只读存储器 (ROM))、磁存储器、光存储器或者其他类型的存储设备。如上所述,处理器可以包括微处理器、ASIC、FPGA、CPU、ALU、DSP 或其他图像处理电路的若干部分或组合。

[0041] 根据前面的描述应当清楚,本发明的实施例可以用集成电路的形式来实现。在给定的此类集成电路实现方式中,相同的管芯通常按照重复的图形形成于半导体晶圆的表面上。每个管芯都包括本文所描述的图像处理器或其他图像处理电路,并且可以包括其他结构或电路。个体管芯被从晶圆中切割或划片而得,然后被封装为集成电路。本领域技术人员应当知道如何对晶圆进行划片并封装管芯以产生集成电路。这样制造的集成电路被认为是本发明的实施例。

[0042] 图 1 所示的图像处理系统 100 的特定配置只是示例性的,并且系统 100 在其他实施例中除了具体示出的那些元件之外或者作为它们的代替可以包括其他元件,包括常见于该系统的常规实现方式中的类型的一个或多个元件。

[0043] 例如,在某些实施例中,图像处理系统 100 被实现为用于处理图像流以便识别用户姿势的视频游戏系统或者其他类型的基于姿势的系统。所公开的技术能够类似地适合用在需要基于姿势的人机界面的各种各样的其他系统中,并且还应用于除姿势识别外的应用,例如,在机器人中的机器视觉系统和其他工业应用。

[0044] 在说明性的实施例中的图像处理器 102 的操作现在将参照图 2 至 8 更详细地描述。应当意识到,结合这些附图所描述的示例性特征及功能不必存在于其他实施例中。

[0045] 首先参照图 2,图像处理器 102 的部分 200 包括与识别层 116 耦接的评价层 114。评价层 114 接收来自预处理层 112 的预处理的图像数据 202,作为它的输入。如前所述,假定这样的图像数据包括一个或多个预处理的深度图像,尽管能够使用其他类型的图像数据。

[0046] 在本实施例中,评价层 114 实现在生成对象数据中的搜索-渲染-比较操作。这些操作由评价层 114 的各个搜索、渲染和比较模块 204、206 和 208 来执行,并且包括所示的对象数据 210A、210B 和 210C 以及相关的对象数据变型的处理或生成。对象数据 210A 与最佳拟合的对象数据 210B 组合,以得到待传递给识别层 116 的输出的最佳拟合对象数据 210C。在本实施例中,假定对象数据 210A 包括具有常量部分和变量部分的实数的向量。例如,如

前所述,对象数据可以包括骨架数据,常量部分包括骨架尺寸,而变量部分包括骨架角度和其他变量数据(例如,质心位置)。

[0047] 评价层 114 在本实施例中还包括校准模块 212、运动预测模块 214 和试探(heuristics)模块 216。校准模块 212 使用当前输入的深度图像来计算骨架数据的常量部分,本文也称为当前输入帧。其他初始对象参数使用比较模块 208 和试探模块 216 来确定,该比较模块 208 和试探模块 216 两者均在当前输入的深度图像上操作。

[0048] 搜索-渲染-比较操作一般地包括按照由搜索模块 204 所实现的特定搜索界定的方式来改变对象参数,该特定搜索可以是梯度搜索或完全搜索,随后是由各个模块 206 和 208 执行的渲染和比较操作。完全搜索可以是同样称为穷尽搜索的那种类型的搜索。梯度搜索或完全搜索可以利用由运动预测模块 214 提供的运动预测信息。

[0049] 渲染(render)操作在本实施例中生成与当前一组对象参数对应的渲染的深度图像。搜索目标的特征可以是使用规定的比较基元(primitive)来最小化在所渲染的深度图像与当前输入的深度图像之间的距离,例如,基于距离度量(例如,欧氏距离或曼哈坦距离)的最小平方。换言之,搜索操作改变对象参数,直到使用比较基元确定对应的渲染的深度图像足够接近于输入深度图像。所渲染的深度图像和输入的深度图像典型地被表示为实数的各自度量,尽管能够使用其他图像数据布局。

[0050] 搜索-渲染-比较过程的复杂性将典型地取决于待确定的对象参数的数量,或者换言之,过程的自由度数量。例如,对于 N 个自由度,完全搜索将使用 k^N 次渲染和比较的迭代,其中 k 在上下文中指示每个自由度的变型数,而梯度搜索将使用 $(N+1) \cdot m$ 次渲染和比较的迭代,其中 m 是沿梯度方向的梯级数。梯度搜索能够被配置用于每 s 个梯级就重新计算梯度,这会使渲染和比较的迭代次数减少至 $(N+1) \cdot m/s+m$ 。

[0051] 在评价层 114 的给定实现方式中所使用的自由度及迭代的具体数量将依据具体的姿势识别应用来改变。例如,头部跟踪应用,假定头部是固体对象(例如,刚体),可以被配置用于使用 6 个自由度(例如, x , y 和 z 维度和 3 个欧拉角)以及缓慢的运动(即,少量的迭代)。在这种类型的应用中,完全搜索和梯度搜索两者都能够实时地操作,即使是以软件来实现。作为另一个实例,手部跟踪应用可以被配置用于每只手使用 25 个自由度以及快速的运动(即,许多迭代)。在这种类型的应用中,完全搜索典型地无法实时地操作,即使是以硬件来实现,尽管梯度搜索能够实时地利用并行性来操作。

[0052] 但是,应当注意,自由度的数量以及每个自由度的变型数能够通过使用试探模块 216 来减少。这可以包括例如计算某些对象参数。

[0053] 图 3 示出了在评价层 114 中对预处理的图像数据 302 执行的某些处理操作 300。在本实施例中,这些处理操作的特征一般在于包括高级的处理操作 304 和低级的图像处理基元 306。

[0054] 如同下文将结合图 5 更详细地描述的,高级的处理操作 304 一般在评价层 114 的软件实现部分中执行,而低级的图像处理基元一般在评价层 114 的硬件实现部分中执行。

[0055] 高级处理操作 304 在本实施例中包括单点检测和跟踪、前表面(frontal surface)检测和跟踪、边缘检测和跟踪,以及快速试探(例如,质心检测)。这些示例性的操作每个都使用输入图像的不同区域,如图所示,并且因此这些操作中的至少一个子集能够被有效地配置以彼此并行地工作。因而,图像处理器 102 的图像处理电路可以包括其中

高级处理操作并行进行的多处理器系统。

[0056] 单点检测和跟踪被用来识别对象,例如,成像个体的头部和手部。这样的对象对于识别层 116 通常被认为是非常重要的特征。

[0057] 前表面检测和跟踪能够为不同类型的深度成像器提供不同等级的可靠性。更特别地,取决于深度成像器类型,所产生的深度图像的不同部分可以含有不同等级的噪声,使得噪声等级能够在对象的前表面与其边缘之间急剧变动。例如,SL 相机一般地提供在前表面处具有相对低的噪声以及在边缘处具有相对高的噪声的深度图像。在此情况中,所检测的前表面比所检测的边缘更可靠。

[0058] 边缘检测和跟踪还能够为不同类型的深度成像器提供不同等级的可靠性。例如,ToF 相机与 SL 相机相比一般提供更少噪声的对象边缘。在此情况中,所检测的边缘比所检测的前表面更可靠。

[0059] 关于快速试探,这些试探一般对更大的图像区域进行操作,并且因此需要对图像数据进行更少的访问。

[0060] 低级的图像处理基元 306 包括:3D 基元(例如,仿射或同度量变换)、2D 基元(例如,曲线跟踪)、比较基元(例如,最小平方比较)、搜索基元(例如,运动预测或梯度搜索)及统计基元,以及可能别的基元。这些的低级处理基元 306 被一个或多个高级处理操作 304 利用。作为一个实例,边缘检测和跟踪操作可以利用曲线跟踪。

[0061] 如上所述,评价层 114 被配置以包括软件实现部分和硬件实现部分。此类布局的实例现在将更详细地参照图 4 和 5 来描述。

[0062] 首先参照图 4,评价层 114 包括精确的软件(SW)算法 401 以及近似的硬件(HW)算法 402,该软件算法 401 和硬件算法 402 被假定作为评价层的各自软件实现部分和硬件实现部分的一部分。在各自的算法 401 和 402 的背景下使用的术语“精确”和“近似”是在此更一般地称为各个精度等级的精度实例,而在其他实施例中能够将其他类型的精度等级用于各自软件及硬件算法中。精确的软件算法 401 和近似的硬件算法 402 在各自相对更高的以及相对更低的精度等级下生成对象数据。

[0063] 评价层 114 还包括被配置用于组合第一及第二对象数据来生成待传递给识别层 116 的输出对象数据的信号组合器 403。如前所述,对象数据可以具有多个部分,例如,常量部分和变量部分。元素 405、406 和 407 在图中指示不同对象数据集合,将在下文更详细地描述。

[0064] 使用精确的软件算法 401 生成的第一对象数据更具体地包括陈旧但精确的数据 406,而由近似的硬件算法 402 生成的第二对象数据包括现时的相对数据 405。第一及第二对象数据组合于信号组合器 403 内,以得到被提供给识别层 116 的现时的且精确的数据 407。现时的且精确的数据 407 在本实施例中被实时地生成,并且为在识别层中的后续处理提供可靠的精度。更特别地,在本实施例中,对象数据被计算作为由软件算法 401 提供的精确的非即时数据以及由硬件算法 402 提供的近似的相对数据的和或其他函数。软件及硬件算法一般相互并行地运行。

[0065] 由于其高精度等级,软件算法 401 无法实时运行,而是生成具有特定潜伏期(latency)的精确对象数据,在此情况中约为 5 至 10 帧。近似的硬件算法 402 在较低的精度等级下操作并且能够在很小的或没有潜伏期的情况下实时运行。它以现时的相对数据

405 的形式来计算相对的对象状态。该现时相对数据在信号组合器 403 内被加上形式为由软件算法 401 计算的陈旧但精确的数据 406 的最后已知的精确对象状态。结果是形式为现时的且精确的数据 407 的当前的对象状态。

[0066] 如上所述,包括由软件算法 401 生成的陈旧但精确的数据 406 的第一对象数据按照大约每 5 至 10 帧一次的速率来更新。术语“陈旧”在上下文中指示对应的对象数据涉及一个或多个先前帧,而不是涉及当前帧。

[0067] 应当注意,大约每 5 至 10 帧一次的更新率可以被看作是一种最坏情形的估计,并且可以使用不同的更新率,取决于图像处理器 102 的能力和配置。例如,如果使用更强大的图像处理电路或者降低处理操作的复杂性,则可以提高更新率,尽管仍然可预料到软件算法 401 将不会实时地操作。软件算法 401 的速度将同样取决于其他因素,例如,正被处理的现时图像数据的特性。例如,完全搜索在某些情况下可以在其第一步确定解决方案。

[0068] 在本实施例中,包括现时的相对数据 405 的第二对象数据假定以比第一对象数据显著要高的速率来更新,并且更具体地,对于每个输入帧都更新。因而,在本实施例中,使用软件算法 401 生成的第一对象数据按照低于每个输入帧一次的速率来更新,而使用硬件算法 402 生成的第二对象数据按照每个输入帧一次的速率来更新。在其他实施例中可以使用与不同精度的对象数据关联的其他速率。

[0069] 图 5 示出了图像处理器 102 的部分 500 的更详细视图,并且示出了在评价层 114 中用以使用精确的软件算法 401、近似的硬件算法 402 及其他关联的处理模块来处理预处理的图像数据 502 的方式。软件算法 401 和硬件算法 402 被布置于评价层 114 的各自的软件实现部分和硬件实现部分 504 和 506 内。评价层 114 的软件实现部分和硬件实现部分 504 和 506 在本文中分别称为软件部分和硬件部分。

[0070] 结合图 3 来描述的低级图像处理基元以硬件实现于评价层 114 的部分 506 内。这些低级基元包括,例如,3D 基元 508、2D 基元 509 和统计基元 510,以及搜索和比较功能。这允许低级基元由以软件实现于评价层 114 的部分 504 中的高级处理操作有效地共用。

[0071] 3D 基元 508 被用来执行仿射或同度量变换操作,例如,移位和旋转,以便支持渲染模块 206。3D 基元还能够支持其他高级处理操作,例如,前表面检测和跟踪。

[0072] 2D 基元 509 被用来支持高级处理操作,例如,边缘检测和跟踪。

[0073] 统计基元 510 由软件算法 401 和硬件算法 402 两者使用。例如,统计基元 510 能够被用来估计估计图像的不同部分的误差等级。

[0074] 在本实施例中,搜索-渲染-比较处理在评价层 114 的软件及硬件部分 504 和 506 之间是分离的。更特别地,图 2 的搜索模块 204 的功能被划分到实现于各自的软件和硬件部分 504 和 506 中的两个搜索模块 204A 和 204B 之内,这两个搜索模块 204A 和 204B 被用来生成相对数据变型 512。同样,渲染模块 206 被实现于软件部分 504 内,而比较模块 208 被实现于硬件部分 506 内。

[0075] 实现于软件部分 504 内的附加元件包括信号组合器 403,以及附加的信号组合器 514 和对象模型 516。对象模型 516 在本实施例中被用来给渲染模块 206 提供数据。

[0076] 尽管在图 5 中没有明确示出,但是可以在硬件部分 506 内实现仲裁器以便控制由软件及硬件搜索模块 204A 和 204B 对渲染和比较模块 206 和 208 的访问。这样的仲裁器能够使用用以通过最高优先级来处理来自软件搜索模块 204A 的访问请求的严格优先级方

法。

[0077] 图 6 至 8 说明性地示出了在 GR 系统 110 的头部跟踪应用的背景下可以实现于评价层 114 内的处理操作集合的三种不同实施例。在这些实施例中,假定评价层实现至少头部检测和跟踪功能,尽管附加的或可替换的高级处理操作类型也可以按照类似的方式来配置。还应当意识到,这些示例性过程的特定步骤、数据及其他特征只是示例性的,并且在其他实施例中能够使用附加的或可替换的步骤、数据及其他特征。

[0078] 在图 6 的实施例中,没有使用单独的软件和硬件算法 401 和 402。相反,当前的输入帧仅使用软件算法 401 来处理,该软件算法 401 在步骤 604 中通过与 3D 头部模型比较来找出头部位置。如图所示,该过程每帧都进行更新以提供精确的头部位置和取向 606。但是,假定使用的是完全搜索而不是梯度搜索,并且因此精确的头部位置和取向 606 在 k 帧的潜伏期下生成,其中通常 $k < 5$ 。此外,潜伏期一般将为以下因素的函数,例如,所使用的图像处理电路的类型、处理操作的复杂性,以及正被处理的现时图像数据的特性。注意, k 在上下文中指的是以帧数计的潜伏期,尽管同一变量在本文的另一上下文下具有不同的含义。

[0079] 在图 7 的实施例中,软件及硬件算法 401 和 402 两者均被用来处理当前的输入帧 702。步骤 704 一般地对应于图 6 的步骤 604,但是只是每隔四帧才执行,也称为“基本”帧,而不是每帧都执行。

[0080] 假定步骤 704 通过软件算法 401 来执行,并且执行与步骤 604 相同的完全搜索方法。对于基本帧,所产生的精确的头部位置和取向 706 每隔四帧才更新。当前的输入帧 702 以及精确的头部位置和取向 706 被存储于历史缓冲器 710 内。

[0081] 还假定,步骤 712 通过硬件算法 402 来执行。步骤 712 被应用于每个输入帧,包括当前的输入帧 702,并且包括找出在先前帧与当前输入帧之间的最佳拟合的同度量变换,有可能使用快速试探。先前帧是取自历史缓冲器 710 的先前帧数据 714 的一部分。所产生的同度量变换被应用于确定来自先前帧的相对映射 715。如同结合步骤 716 所示,相对映射 715 与先前帧数据 714 组合,以产生精确的头部位置和取向信息 718。

[0082] 如上所述,图 7 的实施例每 5 帧就更新精确的头部位置和取向 706,具有 k 帧潜伏期。误差累积得以限制,因为误差最多累积 $5+k$ 帧。所组合的软件和硬件算法处理的总潜伏期小于或等于 1 帧,并且一般地对应于用于确定最佳拟合的同度量变换以及将相对映射应用于当前输入帧所需的时间。同样,评价层 114 的配置允许步骤 704 和 712 异步操作,使得能够获得基本帧数据 706 就及早地更新输出数据 718。

[0083] 图 8 的实施例类似于图 7 的实施例,当前输入帧 802,步骤 804、812、816,历史缓冲器 810 以及数据 806、814、815 和 818 分别对应于图 7 的当前输入帧 702,步骤 704、712、716,历史缓冲器 710 以及数据 706、714、715 和 718。但是,作为如同在图 7 中那样于当前输入帧与先前帧之间找出最佳拟合的同度量变换的代替,图 8 的实施例在步骤 812 中找出在基本帧与当前帧之间的同度量变换,从而得到来自基本帧的相对映射 815,该相对映射 815 在步骤 816 中与基本帧组合。基本帧只是每隔四帧才获取,而对应的精确的头部和取向 806 具有 k 帧的潜伏期。因此,精确的头部取向和位置 818 在图 8 的处理的输出处具有有限的潜伏期及误差积累,类似于在图 7 的过程中生成的相应的输出对象数据的潜伏期及误差积累。

[0084] 图 6 至 8 所示的示例性过程的组合可以并行地执行,或者相互结合,输出可靠性指标被用来选择选择哪个过程输出将被使用。例如,如果用来为给定的基本帧生成精确的头

部位置和取向的完全搜索过程包括显著的误差,则使用该基本帧的图 8 的过程的全部输出将同样包括显著的误差。因此,输出验证模块可以被并入评价层 114 内,用于确定使用多个过程产生的输出的可靠性并且选择来自特定过程的输出以用于当前的输入帧。

[0085] 可靠性确定可以基于例如所算出的在使用来自给定过程的头部位置和取向来渲染的头部与来自预定头部模型的头部位置和取向之间的距离。这样的输出验证模块防止所不希望的误差传播,并且可以被配置以实时地操作。

[0086] 上述输出验证过程的一种可能的实现方式如下:

[0087] 1. 使用图 7 的过程来生成基于先前帧的输出。

[0088] 2. 估计在步骤 1 中生成的输出的可靠性。如果可靠性小于预定的阈值,则进行步骤 3,否则将步骤 1 的输出发送给识别层 116 并退出输出验证过程。

[0089] 3. 使用图 8 的过程来生成基于上一基本帧的输出。

[0090] 4. 估计在步骤 3 中生成的输出的可靠性。

[0091] 5. 基于除上一基本帧外的基本帧但否则使用图 8 的过程来生成输出。

[0092] 6. 估计在步骤 5 中生成的输出的可靠性。

[0093] 7. 比较来自步骤 2、4 和 6 的可靠性估计值,并且选择最高可靠性的输出以传递给识别层 116。

[0094] 该示例性的输出验证过程并不显著地增加评价层 114 的潜伏期,但是允许与对象数据输出关联的误差率被显著降低。

[0095] 如上所述,本发明的实施例能够涉及除姿势识别外的机器视觉应用。例如,本文所公开的技术能够应用于图像处理应用,在该图像处理应用中,对象数据参数能够作为先前状态与相对偏移的组合来计算,或者在该图像处理应用中,处理算法具有不同复杂性和潜伏期。其他实例包括这样的图像处理应用,其中相对简单的算法被用来计算具有相对多个帧的可靠精度的对象数据增量。

[0096] 应当再次强调,本文所描述的本发明的实施例意指仅为说明性的。例如,本发明的其他实施例能够使用除了本文描述的特定实施例所使用的那些之外的各种各样的不同类型及布局的图像处理电路、处理层、处理模块和处理操作来实现。另外,本文在描述某些实施例的背景下所作出的具体假定不必要应用于其他实施例中。本领域技术人员应当容易地清楚在随附的权利要求书的范围内的这些及许多其他可替换的实施例。

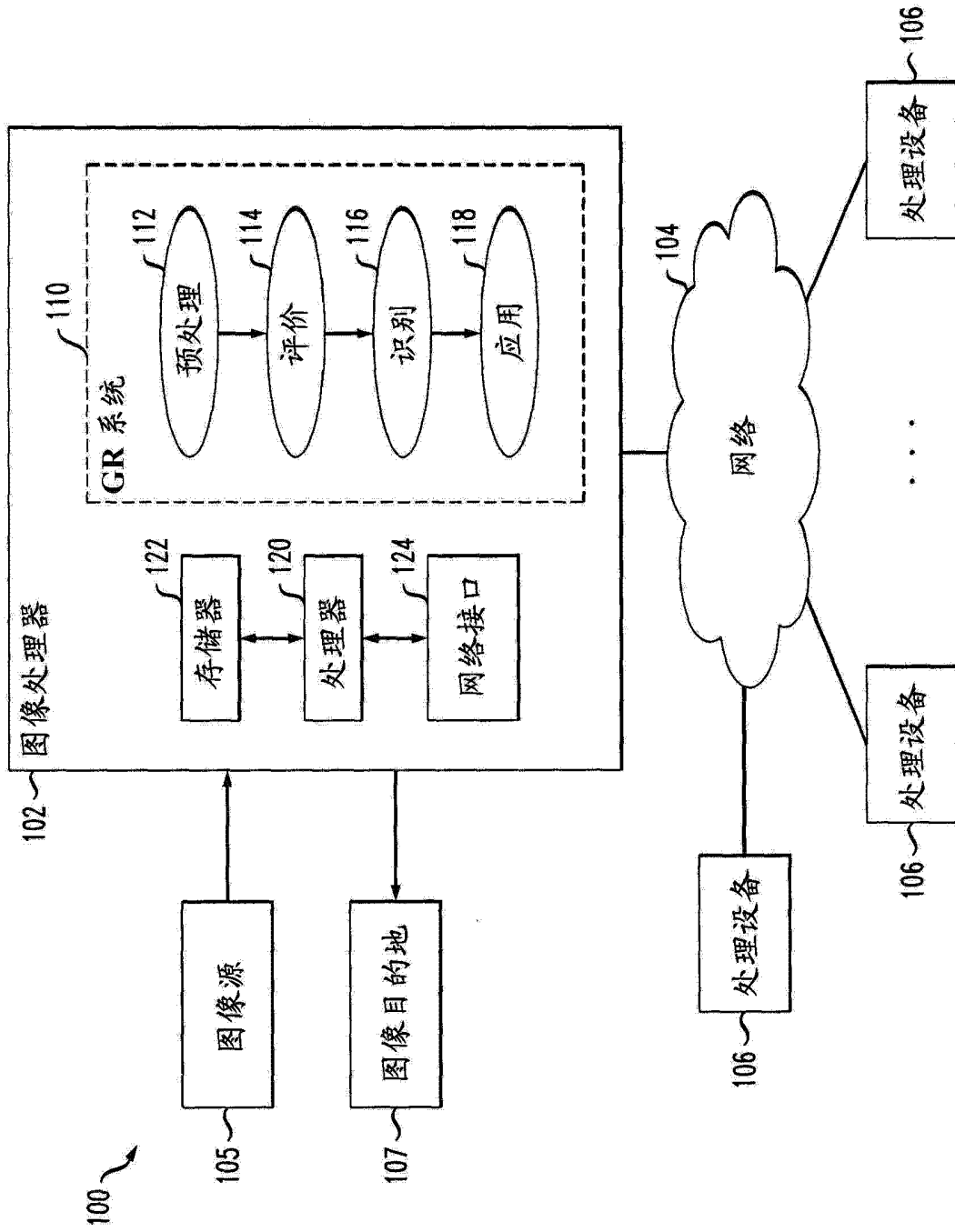


图 1

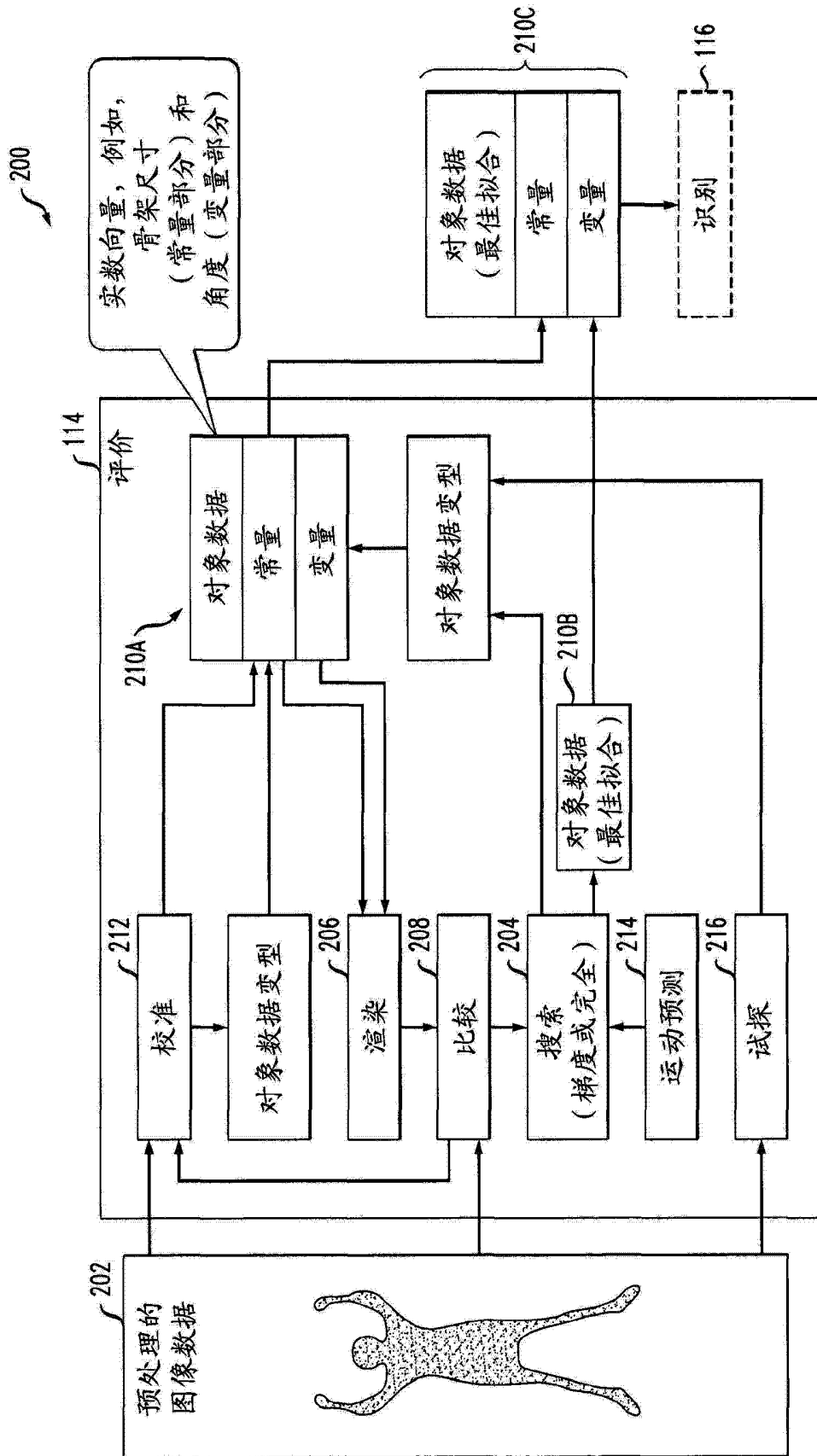


图 2

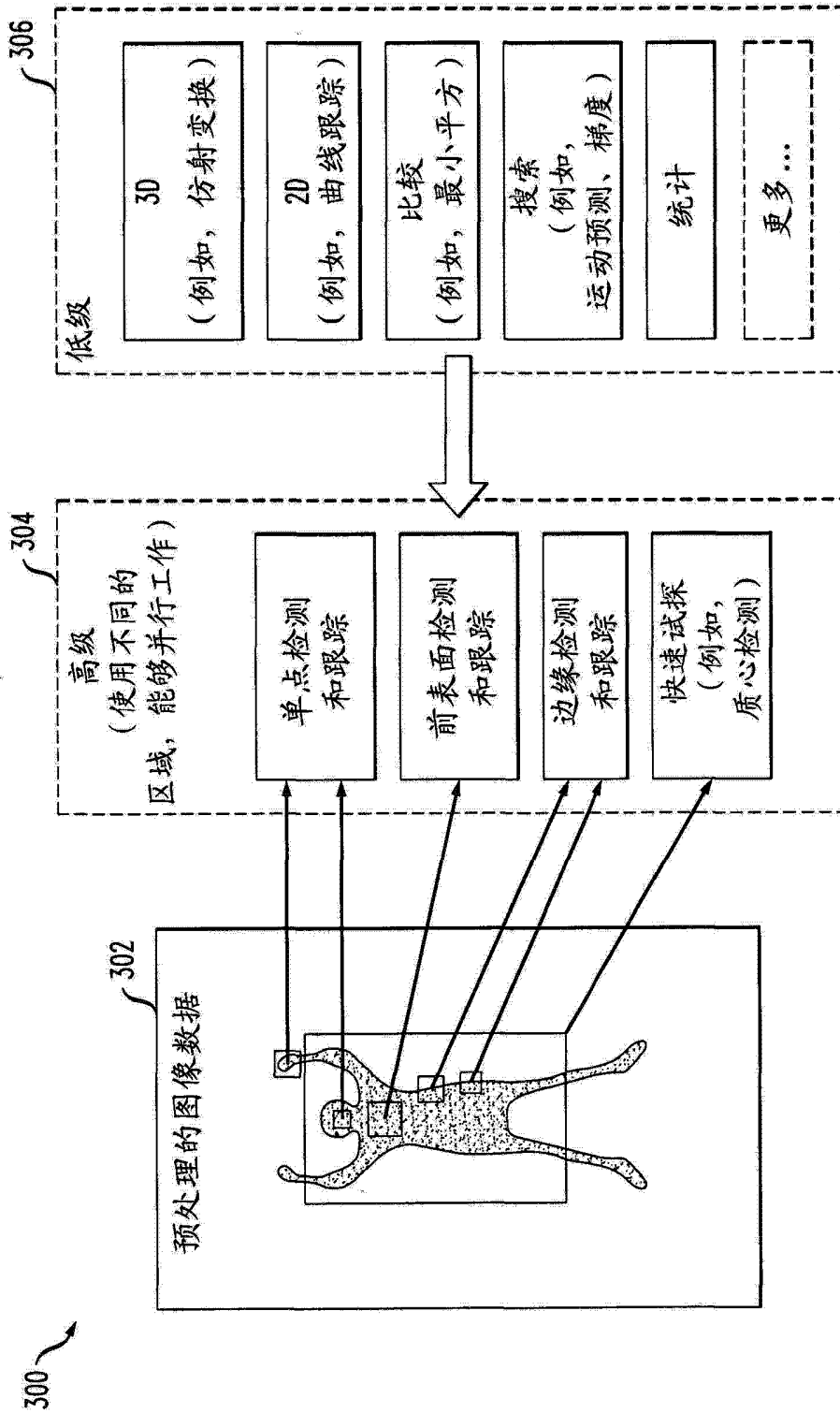


图 3

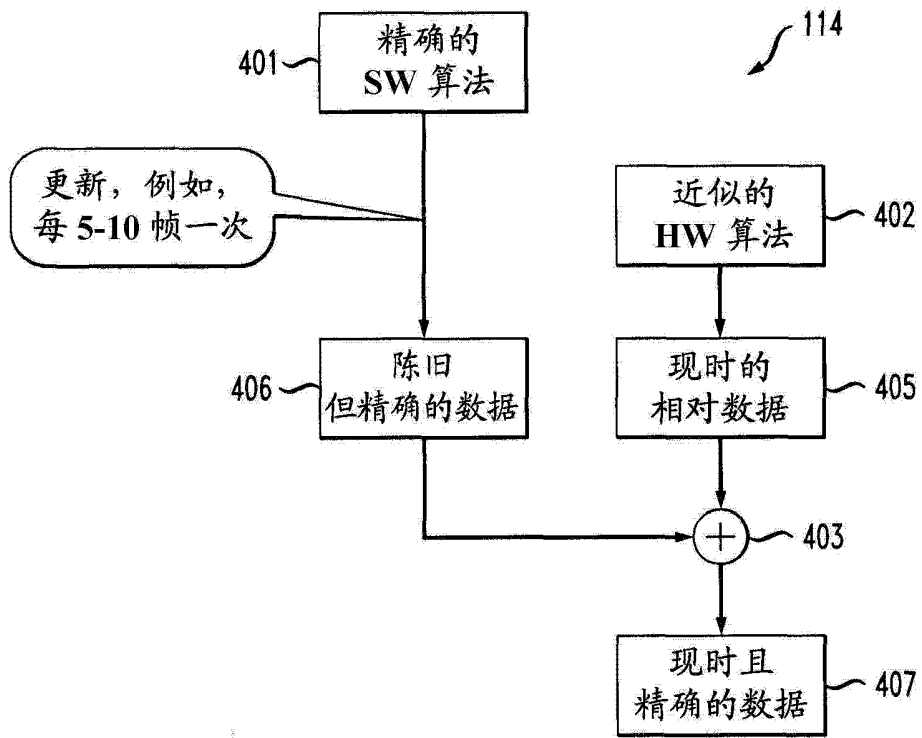


图 4

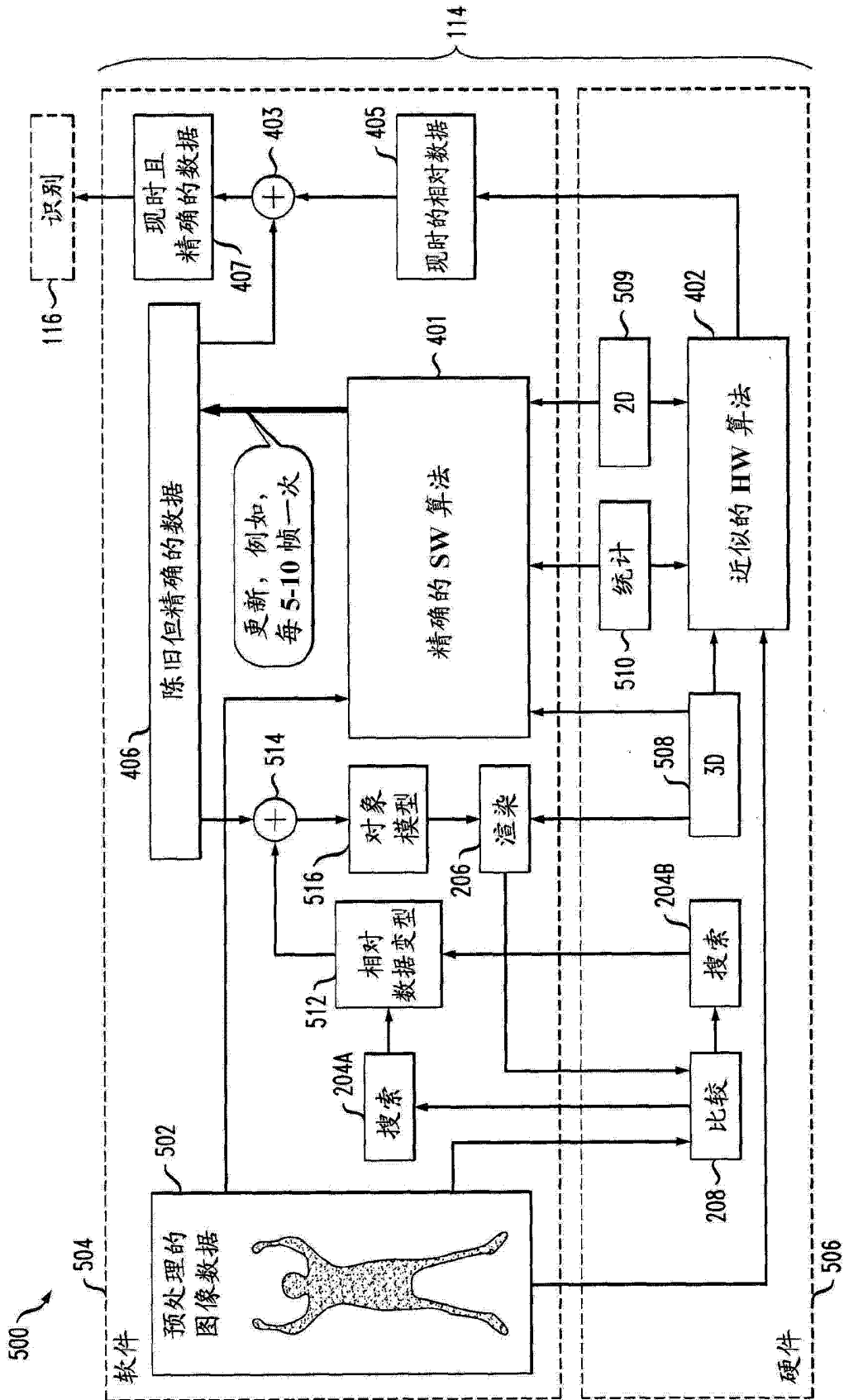


图 5

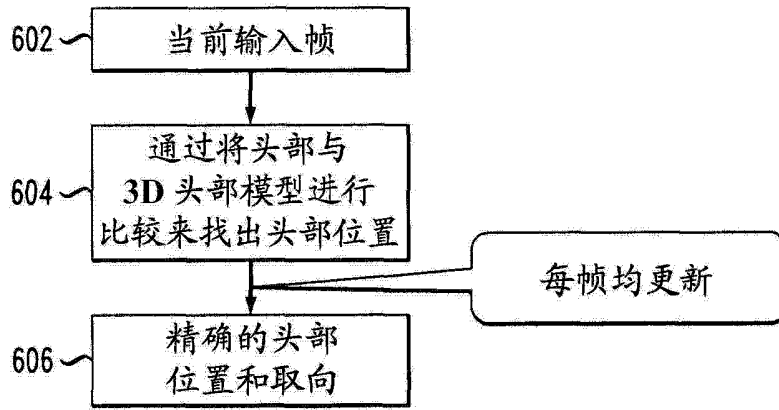


图 6

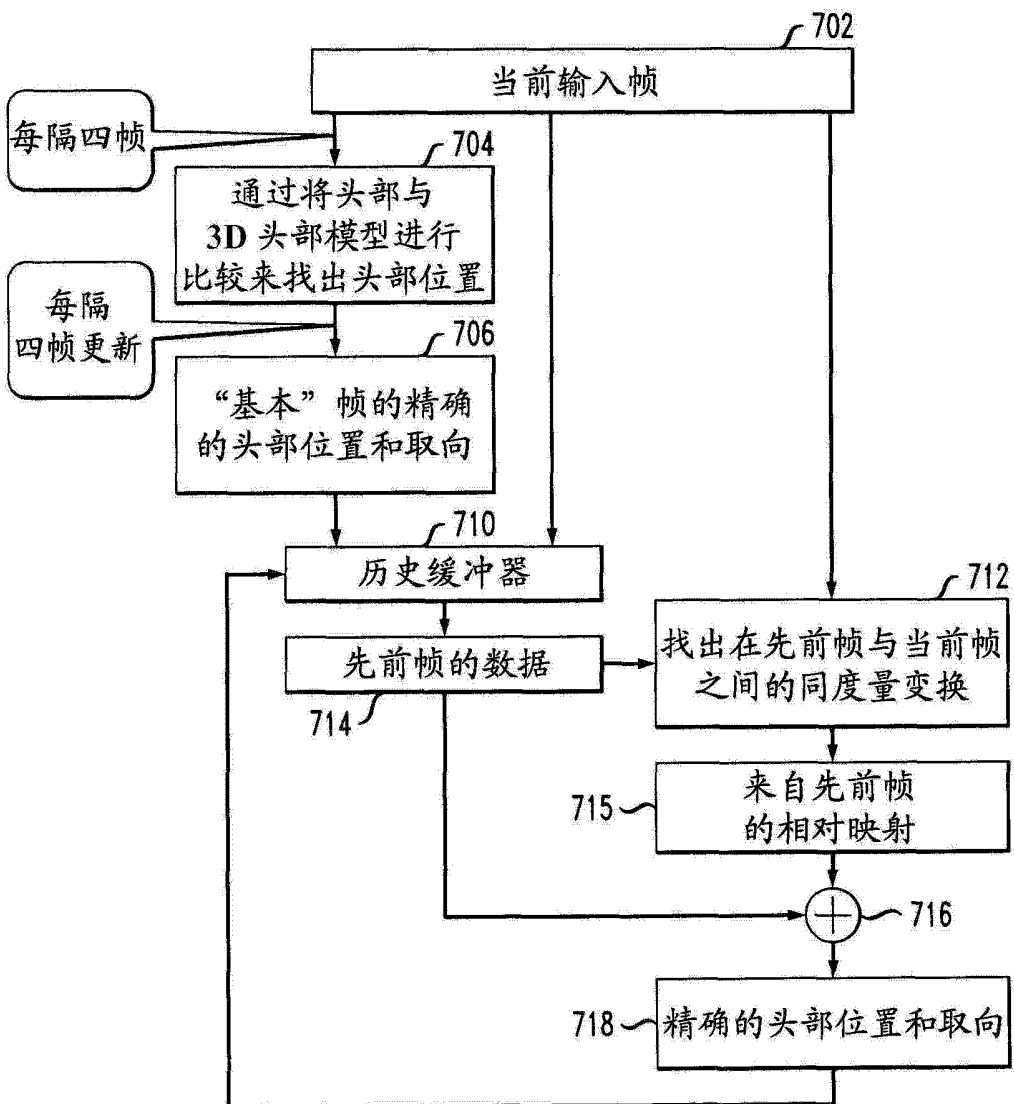


图 7

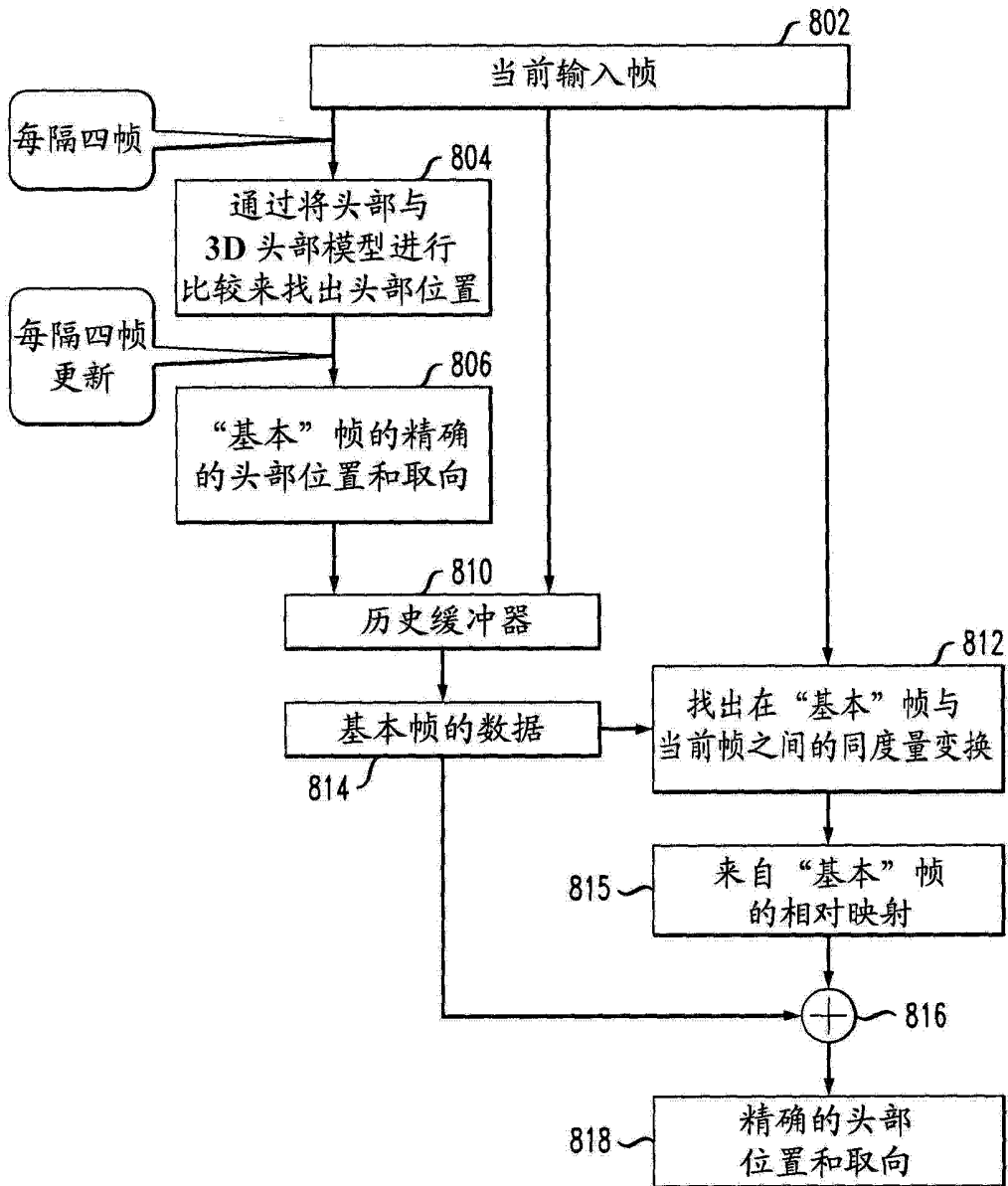


图 8