



(12) 发明专利申请

(10) 申请公布号 CN 102737251 A

(43) 申请公布日 2012. 10. 17

(21) 申请号 201210080952. X

(22) 申请日 2012. 03. 23

(30) 优先权数据

2011-077696 2011. 03. 31 JP

(71) 申请人 索尼公司

地址 日本东京都

(72) 发明人 伊藤真人 佐部浩太郎 横野顺

(74) 专利代理机构 北京集佳知识产权代理有限公司 11227

代理人 康建峰 郎晓虹

(51) Int. Cl.

G06K 9/62(2006. 01)

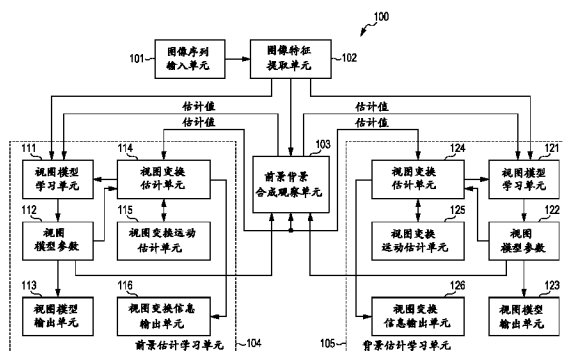
权利要求书 3 页 说明书 40 页 附图 26 页

(54) 发明名称

图像处理设备、图像处理方法、程序和记录介质

(57) 摘要

公开了图像处理设备、图像处理方法、程序和记录介质。该设备包括：图像特征输出单元，用于与帧时间对应地输出每一图像特征；前景估计单元，用于通过在前景视图模型上进行视图变换来估计时间 s 处的前景图像并输出估计的前景视图；背景估计单元，用于通过在背景视图模型上进行视图变换来估计时间 s 处的背景图像并输出估计的背景视图；合成视图生成单元，用于通过合成估计的前景视图和估计的背景视图来生成合成视图；前景学习单元，用于基于通过合成视图和时间 s 处的图像特征间进行比较而得到的估计值并通过更新前景视图模型的参数来学习前景视图模型；及背景学习单元，用于基于通过更新背景视图模型的参数而得到的估计值来学习背景视图模型。



1. 一种图像处理设备,包括:

图像特征输出单元,用于与帧时间对应地输出图像特征中的每一个,所述图像特征被形成作为输入活动图像数据中的每一帧的图像的多个特征点中的特征;

前景估计单元,用于通过在前景视图模型上进行作为几何变换的视图变换来估计时间 s 处的前景图像,并且输出所估计的前景视图,其中,所述前景视图模型具有所述图像中的前景的图像特征,作为与所述时间 s 处的图像特征有关的参数;

背景估计单元,用于通过在背景视图模型上进行作为几何变换的视图变换来估计所述时间 s 处的背景图像,并输出所估计的背景视图,所述背景视图模型具有所述图像中的背景的图像特征,作为与所述时间 s 处的图像特征有关的参数;

合成视图生成单元,用于通过合成所估计的前景视图和所估计的背景视图来生成合成视图;

前景学习单元,用于基于通过所述合成视图和所述时间 s 处的图像特征之间进行比较而得到的估计值并通过基于随机生成模型来更新所述前景视图模型的参数,来学习所述前景视图模型;以及

背景学习单元,用于基于所述估计值、通过基于随机生成模型来更新所述背景视图模型的参数,来学习所述背景视图模型。

2. 根据权利要求 1 所述的图像处理设备,其中,所述前景估计单元和所述背景估计单元中的每一个

基于所述估计值来计算多个视图变换中的每一个的后验概率,

基于所述图像特征和所述后验概率来估计该视图变换,以及

通过在所述视图模型上进行所述视图变换来分别输出所估计的前景视图和所估计的背景视图。

3. 根据权利要求 1 所述的图像处理设备,其中,所述前景学习单元和所述背景学习单元中的每一个

基于根据所述估计值来计算的每一视图变换的后验概率,对经过多次视图变换的视图模型的参数进行加权,

基于加权的参数来分别更新所述前景视图模型的参数和所述背景视图模型的参数。

4. 根据权利要求 1 所述的图像处理设备,其中,所述视图模型被配置为通过多个视图模型而形成的多视图模型,所述多个视图模型与通过以不同的角度来观看所述前景和所述背景中的每一个而得到的图像对应。

5. 根据权利要求 1 所述的图像处理设备,其中,所述前景视图模型和所述背景视图模型之一被配置为通过多个视图模型而形成的多视图模型,所述多个视图模型与通过以不同的角度来观看所述前景和所述背景中的每一个而得到的图像对应。

6. 根据权利要求 4 所述的图像处理设备,其中,所述前景估计单元和所述背景估计单元中的每一个

基于所述估计值来计算所述多个视图模型中的每一个的后验概率,

基于所述图像特征和所述后验概率来估计所述视图变换,以及

通过在所估计的与转变对应的视图模型上进行所述视图变换,来输出所估计的前景视图和所估计的背景视图,

其中,在计算所述视图模型的后验概率时,使用隐藏马尔科夫模型、前向神经网络、自递归神经网络、粒子滤波器和卡尔曼滤波器中的一个作为基于视图转变来估计先验概率的动态学习估计模型。

7. 根据权利要求 1 所述的图像处理设备,其中,所述前景估计单元和所述背景估计单元中的每一个

基于所述估计值来计算多个视图变换中的每一个的后验概率,以及

基于所述图像特征和所述后验概率来估计该视图变换。

8. 根据权利要求 1 所述的图像处理设备,其中,所述和前景学习单元和所述背景学习单元中的每一个

基于根据所述估计值而计算的每一转变的后验概率,对与多个转变对应的视图模型的参数进行加权,

基于根据所述估计值而计算的、每一视图变换的后验概率,对经过多个视图变换的视图模型的参数进行加权,以及

基于经加权的参数来分别更新所述前景视图模型的参数和所述背景视图模型的参数。

9. 根据权利要求 1 所述的图像处理设备,其中,在计算所述随机生成模型中的视图变换的后验概率时,使用隐藏马尔科夫模型、前向神经网络、自递归神经网络、粒子滤波器和卡尔曼滤波器中的一个作为基于视图变换运动来估计先验概率的动态学习估计模型。

10. 根据权利要求 1 所述的图像处理设备,其中,在所述前景学习单元中使用的随机生成模型不同于在所述背景学习单元中使用的随机生成模型。

11. 根据权利要求 1 所述的图像处理设备,其中,所述图像特征输出单元输出每一像素的像素位置和像素值彼此对应的信息,作为所述图像特征。

12. 根据权利要求 1 所述的图像处理设备,其中,所述图像特征输出单元输出利用哈里角点检测方法而检测到的特征点位置集合,作为所述图像特征。

13. 一种图像处理方法,包括:

由图像特征输出单元来与帧时间对应地输出图像特征中的每一个,所述图像特征被形成作为输入活动图像数据中的每一帧的图像的多个特征点中的特征;

由前景估计单元通过在前景视图模型上进行作为几何变换的视图变换来估计时间 s 处的前景图像,并且输出所估计的前景视图,所述前景视图模型具有所述图像中的前景的图像特征,作为与所述时间 s 处的图像特征有关的参数;

由背景估计单元通过在背景视图模型上进行作为几何变换的视图变换来估计所述时间 s 处的背景图像,并输出所估计的背景视图,所述背景视图模型具有所述图像中的背景的图像特征,作为与所述时间 s 处的图像特征有关的参数;

由合成视图生成单元通过合成所估计的前景视图和所估计的背景视图来生成合成视图;

由前景学习单元基于通过所述合成视图和所述时间 s 处的图像特征之间进行比较而得到的估计值并通过基于随机生成模型来更新所述前景视图模型的参数,来学习所述前景视图模型;以及

由背景学习单元基于通过基于随机生成模型来更新所述背景视图模型的参数而得到的估计值,来学习所述背景视图模型。

14. 一种程序,该程序使得计算机作为图像处理设备来工作,该图像处理设备包括:

图像特征输出单元,用于与帧时间对应地输出图像特征中的每一个,所述图像特征被形成作为输入活动图像数据中的每一帧的图像的多个特征点中的特征;

前景估计单元,用于通过在前景视图模型上进行作为几何变换的视图变换来估计时间 s 处的前景图像,并且输出所估计的前景视图,所述前景视图模型具有所述图像中的前景的图像特征,作为与所述时间 s 处的图像特征有关的参数;

背景估计单元,用于通过在背景视图模型上进行作为几何变换的视图变换来估计所述时间 s 处的背景图像,并输出所估计的背景视图,所述背景视图模型具有所述图像中的背景的图像特征,作为与所述时间 s 处的图像特征有关的参数;

合成视图生成单元,用于通过合成所估计的前景视图和所估计的背景视图来生成合成视图;

前景学习单元,用于基于通过在所述合成视图和所述时间 s 处的图像特征之间进行比较而得到的估计值并通过基于随机生成模型来更新所述前景视图模型的参数,来学习所述前景视图模型;以及

背景学习单元,用于基于通过基于随机生成模型来更新所述背景视图模型的参数而得到的估计值,来学习所述背景视图模型。

15. 一种记录介质,用于存储根据权利要求 14 所述的程序。

图像处理设备、图像处理方法、程序和记录介质

技术领域

[0001] 本技术涉及图像处理设备、图像处理方法、程序和记录介质,更具体地,涉及被配置为进行学习图像处理的图像处理设备、图像处理方法、程序和记录介质。

背景技术

[0002] 例如,用于面部识别或对象识别等的模式识别技术已经被视为与学习图像处理有关的当前可商业获得的技术。

[0003] 在根据相关技术的学习图像处理中,学习是在识别目标的标记被赋予大量图像数据作为学习数据之后进行的,此时识别目标的模型被学习,以形成识别器。

[0004] 例如,当对图像进行学习以便进行面部识别时,需要赋予用于确定人的姓名、他或她的面部的取向以及显示有该人面部图像的区域的信息等,作为标记。另外,当对图像进行学习以便进行目标识别时,需要赋予用于确定目标的名称、该目标的取向以及显示有该目标的区域的信息等,作为标记。

[0005] 由于标记的赋予是手动进行的,因此,难以准备大量的学习数据。

[0006] 因此,例如,已出现了根据包括活动图像的多个图像来自动学习目标模型而不进行如上所述的标记赋予的学习图像处理技术。

[0007] 作为学习图像处理的示例,已出现了一种自动学习前景和背景学习模型的方法,其中,根据包括活动图像的多个图像来自动学习目标模型(例如,参见 Michalis K. Titsias、Christopher K. I. Williams 所著的 "Unsupervised Learning of Multiple Aspects of Moving Objects from Video", 刊于 Panhellenic Conference on Informatics 2005,第 746-756 页)。

[0008] 另外,作为学习图像处理的另一示例,已出现了一种自动学习与前景对应的多视图目标模型的方法,其中,根据包括活动图像的多个图像来自动学习目标模型(例如,参见 H. Su、M. Sun、L. Fei-Fei 和 S. Savarese 所著的 "Learning a dense multi-view representation for detection, viewpoint classification and synthesis of object categories", 刊于 International Conference on Computer Vision (ICCV), 2009 年)。根据该方法,对多个视图之间的几何关系进行建模。

发明内容

[0009] 然而,在例如 "Unsupervised Learning of Multiple Aspects of Moving Objects from Video" 所公开的技术中,由于没有对多视图中的各个视图之间的关系进行建模,因而需要完成对每一视图的学习。因此,可以想像,当学习多视图时,如果视图的数量增多,则学习会很困难。

[0010] 在 "Learning a dense multi-view representation for detection, viewpoint classification and synthesis of object categories" 所披露的技术中,当在实际活动图像中包括与前景接近的背景时,由于学习是在没有分解前景和背景的情况下进行的这一

事实,因而,学习会很困难。另外,由于多个视图之间的几何关系被准确建模,因此,会出现操作于拍摄活动图像的相机的方法被强加了限制或者计算量大的问题。

[0011] 期望提供一种假定 (supposition) 较少且计算量较小的执行学习图像处理的技术。

[0012] 根据本技术的一个实施例,提供了一种图像处理设备,该图像处理设备包括:图像特征输出单元,用于与帧时间对应地输出图像特征中的每一个,所述图像特征被形成作为输入活动图像数据中的每一帧的图像的多个特征点中的特征;前景估计单元,用于通过在前景视图模型上进行作为几何变换的视图变换来估计时间 s 处的前景图像,并且输出所估计的前景视图,所述前景视图模型具有所述图像中的前景的图像特征,作为与所述时间 s 处的图像特征有关的参数;背景估计单元,用于通过在背景视图模型上进行作为几何变换的视图变换来估计所述时间 s 处的背景图像,并输出所估计的背景视图,所述背景视图模型具有所述图像中的背景的图像特征,作为与所述时间 s 处的图像特征有关的参数;合成视图生成单元,用于通过合成所估计的前景视图和所估计的背景视图来生成合成视图;前景学习单元,用于基于通过所述合成视图和所述时间 s 处的图像特征之间进行比较而得到的估计值并通过基于随机生成模型来更新所述前景视图模型的参数,来学习所述前景视图模型;以及背景学习单元,用于基于通过基于随机生成模型来更新所述背景视图模型的参数而得到的估计值,来学习所述背景视图模型。

[0013] 所述前景估计单元和所述背景估计单元中的每一个可基于所述估计值来计算多个视图变换中的每一个的后验概率,基于所述图像特征和所述后验概率来估计该视图变换,以及通过在所述视图模型上进行所述视图变换来分别输出所估计的前景视图和所估计的背景视图。

[0014] 所述前景学习单元和所述背景学习单元中的每一个可基于根据所述估计值来计算的每一视图变换的后验概率,对经过多次视图变换的视图模型的参数进行加权,并基于加权的参数来分别更新所述前景视图模型的参数和所述背景视图模型的参数。

[0015] 所述视图模型可被配置为通过多个视图模型而形成的多视图模型,所述多个视图模型与通过以不同的角度来观看所述前景和所述背景中的每一个而得到的图像对应。

[0016] 所述前景视图模型和所述背景视图模型之一可被配置为通过多个视图模型而形成的多视图模型,所述多个视图模型与通过以不同的角度来观看所述前景和所述背景中的每一个而得到的图像对应。

[0017] 所述前景估计单元和所述背景估计单元中的每一个可基于所述估计值来计算所述多个视图模型中的每一个的后验概率,基于所述图像特征和所述后验概率来估计所述视图变换,并通过在所估计的与转变 (transition) 对应的视图模型上进行所述视图变换,来输出所估计的前景视图和所估计的背景视图。在计算所述视图模型的后验概率时,使用 HMM、FNN、RNN、粒子滤波器和卡尔曼滤波器中的一个作为用于基于视图转变 (view transition) 来估计先验概率的动态学习估计模型。

[0018] 所述前景估计单元和所述背景估计单元中的每一个可基于所述估计值来计算多个视图变换中的每一个的后验概率,以及基于所述图像特征和所述后验概率来估计该视图变换。

[0019] 所述和前景学习单元和所述背景学习单元中的每一个可基于根据所述估计值而

计算的每一转变的后验概率,对与多个转变对应的视图模型的参数进行加权,基于根据所述估计值而计算的、每一视图变换的后验概率,对经过多个视图变换的视图模型的参数进行加权,以及基于经加权的参数来分别更新所述前景视图模型的参数和所述背景视图模型的参数。

[0020] 在计算所述随机生成模型中的视图变换的后验概率时,可使用 HMM、FNN、RNN、粒子滤波器和卡尔曼滤波器中的一个作为基于视图变换运动来估计先验概率的动态学习估计模型。

[0021] 在所述前景学习单元中使用的随机生成模型可不同于在所述背景学习单元中使用的随机生成模型。

[0022] 所述图像特征输出单元可输出每一像素的像素位置和像素值彼此对应的信息,作为所述图像特征。

[0023] 所述图像特征输出单元可输出利用哈里斯角点检测方法而检测到的特征点位置集合,作为所述图像特征。

[0024] 根据本技术的另一实施例,提供了一种图像处理方法,该方法包括:由图像特征输出单元来与帧时间对应地输出图像特征中的每一个,所述图像特征被形成作为输入活动图像数据中的每一帧的图像的多个特征点中的特征;由前景估计单元通过在前景视图模型上进行作为几何变换的视图变换来估计时间 s 处的前景图像,并且输出所估计的前景视图,所述前景视图模型具有所述图像中的前景的图像特征,作为与所述时间 s 处的图像特征有关的参数;由背景估计单元通过在背景视图模型上进行作为几何变换的视图变换来估计所述时间 s 处的背景图像,并输出所估计的背景视图,所述背景视图模型具有所述图像中的背景的图像特征,作为与所述时间 s 处的图像特征有关的参数;由合成视图生成单元通过合成所估计的前景视图和所估计的背景视图来生成合成视图;由前景学习单元基于通过所述合成视图和所述时间 s 处的图像特征之间进行比较而得到的估计值并通过基于随机生成模型来更新所述前景视图模型的参数,来学习所述前景视图模型;以及由背景学习单元基于通过基于随机生成模型来更新所述背景视图模型的参数而得到的估计值,来学习所述背景视图模型。

[0025] 根据本技术的又一实施例,提供了一种程序,该程序使得计算机作为图像处理设备来工作,该图像处理设备包括:图像特征输出单元,用于与帧时间对应地输出图像特征中的每一个,所述图像特征被形成作为输入活动图像数据中的每一帧的图像的多个特征点中的特征;前景估计单元,用于通过在前景视图模型上进行作为几何变换的视图变换来估计时间 s 处的前景图像,并且输出所估计的前景视图,所述前景视图模型具有所述图像中的前景的图像特征,作为与所述时间 s 处的图像特征有关的参数;背景估计单元,用于通过在背景视图模型上进行作为几何变换的视图变换来估计所述时间 s 处的背景图像,并输出所估计的背景视图,所述背景视图模型具有所述图像中的背景的图像特征,作为与所述时间 s 处的图像特征有关的参数;合成视图生成单元,用于通过合成所估计的前景视图和所估计的背景视图来生成合成视图;前景学习单元,用于基于通过所述合成视图和所述时间 s 处的图像特征之间进行比较而得到的估计值并通过基于随机生成模型来更新所述前景视图模型的参数,来学习所述前景视图模型;以及背景学习单元,用于基于通过基于随机生成模型来更新所述背景视图模型的参数而得到的估计值,来学习所述背景视图模型。

[0026] 根据本公开的实施例,与帧时间对应地输出图像特征中的每一个(所述图像特征被形成作为输入活动图像数据中的每一帧的图像的多个特征点中的特征);通过在前景视图模型(所述前景视图模型具有所述图像中的前景的图像特征,作为与所述时间 s 处的图像特征有关的参数)上进行作为几何变换的视图变换来估计时间 s 处的前景图像,并且输出所估计的前景视图;通过在背景视图模型(所述背景视图模型具有所述图像中的背景的图像特征,作为与所述时间 s 处的图像特征有关的参数)上进行作为几何变换的视图变换来估计所述时间 s 处的背景图像,并输出所估计的背景视图;通过合成所估计的前景视图和所估计的背景视图来生成合成视图;基于通过在所述合成视图和所述时间 s 处的图像特征之间进行比较而得到的估计值并通过随机生成模型来更新所述前景视图模型的参数,来学习所述前景视图模型;以及基于通过基于随机生成模型来更新所述背景视图模型的参数而得到的估计值,来学习所述背景视图模型。

[0027] 根据本技术的所述实施例,能够以较少的假定和较小的计算量来执行学习图像处理。

附图说明

[0028] 图 1 是示出输入图像的序列的一个示例的图;

[0029] 图 2 是示出根据相关技术的图像学习和识别以及根据本技术的图像学习和识别的图;

[0030] 图 3 为示出根据本技术的一个实施例的视图模型学习设备的配置的一个示例的框图;

[0031] 图 4 为示出根据本技术的另一实施例的视图模型学习设备的配置的一个示例的框图;

[0032] 图 5 为示出视图模型学习处理的一个示例的流程图;

[0033] 图 6 为示出前景估计处理的一个示例的流程;

[0034] 图 7 为示出背景估计处理的一个示例的流程图;

[0035] 图 8 为示出合成观察处理的一个示例的流程图;

[0036] 图 9 为示出前景学习处理的一个示例的流程图;

[0037] 图 10 为示出背景学习处理的一个示例的流程图;

[0038] 图 11 为示出视图模型学习处理的另一示例的流程图;

[0039] 图 12 为示出前景估计处理的另一示例的流程图;

[0040] 图 13 为示出背景估计处理的另一示例的流程图;

[0041] 图 14 为示出合成观察处理的另一示例的流程图;

[0042] 图 15 为示出前景学习处理的另一示例的流程图;

[0043] 图 16 为示出背景学习处理的另一示例的流程图;

[0044] 图 17 为示出前景图像的一个示例的图;

[0045] 图 18 为示出背景图像的一个示例的图;

[0046] 图 19 为示出输入图像序列中包括的图像的一个示例的图;

[0047] 图 20 为示出视图模型学习和图像识别中的视图变换估计的结构的图;

[0048] 图 21 为示出图像识别结果的图;

- [0049] 图 22 为示出另一输入图像序列的示例的图；
[0050] 图 23 为示出视图模型学习的一个示例的图；
[0051] 图 24 为示出图像识别结果的另一示例的图；
[0052] 图 25 为示出又一输入图像序列中包括的图像的示例的图；
[0053] 图 26 为示出视图模型学习的另一示例的图；
[0054] 图 27 为示出图像识别结果的另一示例的图；以及
[0055] 图 28 为示出图像识别结果的又一示例的图。

具体实施方式

[0056] 下文参考附图来描述在此公开的本技术的实施例。

[0057] 首先将描述根据相关技术的学习图像处理与根据本技术的实施例的学习图像处理之间的差别。

[0058] 在本技术的实施例中，例如，如图 1 所示，作为活动图像的、按时间顺序拍摄的图像 21-1 到 21-5 被提供，作为输入图像序列，并且前景视图模型和背景视图模型被配置为根据该输入图像序列来自动生成。

[0059] 在图 1 所示的示例中，在图像 21-1 到 21-5 中显示了笔记本类型的个人计算机。例如，图像 21-1 和 21-5 是通过在不同位置、围绕该笔记本类型的个人计算机来移动相机、从各个角度来拍摄该笔记本类型的个人计算机而得到的。例如，所述图像是通过将相机面对该笔记本类型的个人计算机并在从左侧向右侧移动该相机的情况下拍摄活动图像而得到的。

[0060] 在图 1 中，将笔记本类型的个人计算机作为前景来学习。放置笔记本类型的个人计算机的桌子以及放置于笔记本类型的个人计算机旁边的书籍作为背景来学习。前景学习的结果指的是前景视图模型，而背景学习的结果指的是背景视图模型。另外，后面会详细描述视图模型。

[0061] 在本技术的实施例中，假定在前景和背景之间存在运动的独立性。也就是说，当要学习的前景进行了运动（用 M01 来表示）时，假设要学习的背景进行与 M01 不同的运动（用 M02 来表示）。

[0062] 在本技术的实施例中，假定了光学前后关系（optical anterior-posterior relation）。也就是说，假定要学习的背景可能隐藏在前景后面，但前景不会隐藏在背景后面。

[0063] 在本技术的实施例中，如上所述，前景和背景之间的运动独立性以及光学前后关系被设为约束因素，并且通过在例如图像 21-1 到 21-5 的前景（所述笔记本类型的个人计算机）上执行视图变换估计来学习前景视图模型。另外，通过在图像 21-1 到 21-5 的背景（桌子和书籍等）上执行视图变换估计来学习背景视图模型。

[0064] 相应地，在根据本公开的技术中，如图 2 所示，可仅通过输入活动图像来进行图像识别，而不存储并学习大量的标记图像。

[0065] 例如，在根据相关技术的图像识别中，学习为从多个方向拍摄上述笔记本类型的个人计算机而得到的每一图像赋予标记。例如，所述标记包括用于指定所述笔记本类型的个人计算机的取向以及图像中显示所述笔记本类型的个人计算机的区域的信息。因此，如

图 2 所示,在本技术的实施例中,可通过仅输入没有标记的活动图像来识别所述笔记本类型的个人计算机。

[0066] 接下来,将描述本技术的实施例中使用的基本模型。

[0067] 当识别目标图像时,本技术的实施例可对应于单视图模型和多视图模型二者。

[0068] 在此,所述单视图模型指的是假定要学习的前景和背景可以用一个视图模型来表示的模型。

[0069] 另一方面,所述多视图模型指的是假定要学习的前景和背景可用多个视图模型来表示的模型。另外,视图模型被视为表示多视图模型中的一个图像。

[0070] 例如,输入图像序列用 $I_1, I_2, I_3, \dots, I_S$ 来表示,并且输入图像序列的特征用 $X_1, X_2, X_3, \dots, X_S$ 来表示。另外,由于输入图像序列可被设置为例如活动图像的图像数据,因此, $I_1, I_2, I_3, \dots, I_S$ 分别对应于活动图像的各帧的图像数据。此外,图像特征 of 输入图像序列的图像特征可被设置为输入图像序列(活动图像)的各图像(各帧的图像)中的像素值的集合。

[0071] 在本技术的实施例中,假定输入图像序列的图像特征(也称为观察特征)可通过每次变换前景视图模型或背景视图模型来获得。也就是说,在时间 s 处的观察特征 X_s 可通过利用等式 (1) 在时间 s 处对视图模型 M 进行视图变换 T_s 来获得。

$$[0072] \quad X_s = T_s M \quad (1)$$

[0073] 在等式 (1) 中, M 为单视图模型。视图变换 T_s 为用仿射变换(仿射变换)等来表示的坐标变换。

[0074] 另一方面,当使用多视图模型时,存在多个视图模型。例如,包括 L 个视图模型的多视图模型 M_v 用 $M_v = \{M_1, M_2, M_3, \dots, M_L\}$ 来表示。例如,这多个视图模型被视为对应于通过从前面、后面和侧面等来观看要学习的前景而得到的图像。

[0075] 在这种情况下,例如,假定图像特征(也称为观察特征)可通过每次对几个视图模型进行视图变换来获得。也就是说,在时间 s 处的观察特征 X_s 可通过利用等式 (2) 在时间 s 处对视图模型 M_v 进行视图变换 T_s 来获得。

$$[0076] \quad X_s = T_s M_v \quad (2)$$

[0077] 在本技术的实施例中,等式 (1) 或等式 (2) 中示出的视图模型通过下文描述的处理来学习。

[0078] 图 3 为示出根据本技术的一个实施例的视图模型学习设备 100 的配置的一个示例的框图。假设图 3 中的视图模型学习设备 100 在前景和背景学习中使用单视图模型。

[0079] 如图 3 所示,视图模型学习设备 100 包括图像序列输入单元 101,图像特征提取单元 102,前景背景合成观察单元 103,前景估计学习单元 104 和背景估计学习单元 105。

[0080] 前景估计学习单元 104 包括视图模型学习单元 111,视图模型参数 112,视图模型输出单元 113,视图变换估计单元 114,视图变换运动估计单元 115 以及视图变换信息输出单元 116。

[0081] 背景估计学习单元 105 包括视图模型学习单元 121,视图模型参数 122,视图模型输出单元 123,视图变换估计单元 124,视图变换运动估计单元 125 以及视图变换信息输出单元 126。

[0082] 图 3 中的图像序列输入单元 101 是控制输入图像序列的输入并提供被使得对应于

各个时间的、作为输入图像序列的图像到如上所述的图像特征提取单元 102 的功能块。

[0083] 图像特征提取单元 102 从各个时间的图像（即输入图像序列）中提取图像特征。图像特征的具体示例包括每一像素的亮度值和由多个像素形成的每一区域的亮度直方图、色彩直方图以及边缘直方图。或者，可使用通过哈里斯角点检测（Harris corner detection）检测到的每一特征点的坐标和 / 或具有方向可控滤波器等的特征的局部特征，作为所述图像特征。

[0084] 原则上，图像特征包括几何元素（几何模型）和特征量元素（特征量模型），如下文描述的视图模型的视图模型参数等。例如，图像特征可包括与一帧对应的图像中的每一像素的像素位置（几何模型）和像素值（特征量模型）。或者，图像特征可包括与一帧对应的图像中的预定特征点的像素的像素位置（几何模型）和像素值（特征量模型）。

[0085] 当使用局部特征时，如下所述的，图像特征可仅包括与一帧对应的图像中的预定特征点的像素的像素位置（几何模型）。

[0086] 前景背景合成观察单元 103 合成前景估计学习单元 104 输出（估计）的所估计的前景视图和背景估计学习单元 105 输出（估计）的所估计的背景视图。前景背景合成观察单元 103 将合成图像的图像特征与从图像特征提取单元 102 提供的图像特征相比较，并分别向前景估计学习单元 104 和背景估计学习单元 105 提供有关所述估计的估计值的反馈信息。

[0087] 也就是说，前景背景合成观察单元 103 从前景估计学习单元 104 获取所估计的前景视图。在此，当视图模型参数 112 存储前景视图模型 M_{FG} 并且前景视图模型在时间 s 的视图变换为 $T_{FG, s}$ 时，所估计的在时间 s 的前景视图可用 $T_{FG, s} M_{FG}$ 来表示。

[0088] 前景背景合成观察单元 103 从背景估计学习单元 105 获取所估计的背景视图。在此，当视图模型参数 122 存储背景视图模型 M_{BG} 并且背景视图模型在时间 s 的视图变换为 $T_{BG, s}$ 时，所估计的在时间 s 的背景视图可用 $T_{BG, s} M_{BG}$ 来表示。

[0089] 然后，前景背景合成观察单元 103 形成前景和背景的合成视图模型 $M_{FGB, s}$ 。

[0090] 前景背景合成观察单元 103 确定从图像特征提取单元 102 提供的图像特征的特征参数与合成视图模型的特征参数之间的对应关系。例如，在时间 s 处的图像特征 X_s 有 $x_1, x_2, x_3, \dots, x_N$ 作为参数，并且合成视图模型 $M_{FGB, s}$ 有 $m_1, m_2, m_3, \dots, m_N$ 作为参数。

[0091] 例如，当图像特征是每一像素的亮度值（每一像素值）时，可使得输入图像序列中时间 s 处的图像对应于与时间 s 处的合成视图模型对应的坐标的像素值。

[0092] 例如，通过使得时间 s 处的图像的在坐标 $(0, 0)$ 处的像素 x_1 对应于时间 s 处的合成视图模型的在坐标 $(0, 0)$ 处的像素 m_1 、使得时间 s 处的图像的在坐标 $(0, 1)$ 处的像素 x_2 对应于时间 s 处的合成视图模型的在坐标 $(0, 1)$ 处的像素 m_2 等来确定所述对应关系。时间 s 处的视图模型与图像之间的对应关系表示为 $C_{XFG, s}$ 。

[0093] 前景背景合成观察单元 103 在上述对应关系 $C_{XFG, s}$ 下针对图像特征 X_s 计算合成视图模型 $M_{FGB, s}$ 的估计值 E_s 。估计值 E_s 为对应的像素值之间的差的绝对值（该绝对值实际上是通过更为复杂的计算来计算的）的和。

[0094] 图 3 中的前景估计学习单元 104 如下操作。

[0095] 例如，视图变换估计单元 114 估计适用于时间 t 处的图像特征 X_s 的视图变换 T_s 。也就是说，视图变换估计单元 114 在前景视图模型上进行视图变换并估计通过其可获得图

像特征 X_s 中的前景图像的视图变换。

[0096] 此时,视图变换估计单元 114 输出(估计)多个合适的视图变换的候选。例如,视图变换估计单元 114 输出具有上百个不同参数的仿射变换,作为视图变换的候选。

[0097] 在此,例如,通过对前景视图模型进行几何变换来进行视图变换。具体地,用于平移(translate)视图模型的变换、用于扩展并缩小视图模型的变换、用于转动视图模型的变换、用于进行变换的组别的仿射变换或投影变换等被用作所述视图变换。由于视图变换的候选近乎无限存在,要输出的视图变换的候选,基于下面描述的视图变换运动估计单元 115 的估计结果来指定要输出的视图变换的候选。

[0098] 视图变换运动估计单元 115 被配置用于根据时间 s 处的视图变换 T_s 来估计时间 $s+1$ 处的视图变换 T'_{s+1} 。一般而言,当假定前景运动规则连续时,变换动态(transform dynamics)被假定作为满足等式 $T'_{s+1} = F_T(T_s)$ 的模型。例如,诸如 HMM(隐藏马尔科夫模型, Hidden Markov Model)、FNN(前向神经网络, Feed Forward Neural Network) 或 RNN(自递归神经网络, Recurrent Neural Network) 等动态学习估计模型可以用于该模型。诸如粒子滤波器(particle filter)或卡尔曼滤波器(Kalman filter)等动态估计模型可用于该模型。

[0099] 更具体地,视图变换运动估计单元 115 被配置用于基于前景背景合成观察单元 103 在给定时间输出的估计值 E_s 来计算每一视图变换的后验概率。视图变换估计单元 114 被配置用于通过基于每一视图变换选择并在给定时间输出合适的视图变换的后验概率来估计视图变换。

[0100] 视图模型参数 112 存储前景视图模型的参数。在此,这样说,由于前景包括多个特征点的几何模型的参数以及每一特征点的特征量模型的参数,前景模型的参数是通过对图像的特征进行建模来获得的。

[0101] 几何模型是用于多个特征点之间的相对位置关系的统计模型,而特征量模型是用于每一特征点的特征的统计模型。例如,当图像特征是每一像素的亮度值(每一像素值)时,几何模型表示每一像素的坐标值。另一方面,特征量模型表示通过将从多个前景中获得的相应坐标值的像素值乘以权重而得到的值的平均值。仅使用均值的方法和使用正态分布(均值和方差(dispersion))的方法可用作几何模型和特征量模型的统计模型。

[0102] 视图模型学习单元 111 针对输入图像序列中在每一时间处的图像的图像特征,相对于基于前景背景合成观察单元 103 确定的对应关系而选择的特征点,来学习上述几何模型和特征量模型。此时,视图模型学习单元 111 通过视图变换估计单元 114 估计的视图变换(例如,仿射变换)的逆变换来学习几何模型。然后,视图模型学习单元 111 基于前景背景合成观察单元 103 输出的估计值,在所有时间处以统计的方式来学习特征量模型。这样,视图模型的参数得以学习。

[0103] 下面描述用于获得视图模型的参数(包括几何模型的参数和特征量模型)的计算以及这些参数的学习。

[0104] 视图模型学习单元 111 学习的几何模型和特征量模型被存储于视图模型参数 112。相应地,前景视图模型 M_{FG} 中的实数(actual numbers)等被存储于视图模型参数 112。

[0105] 视图变换信息输出单元 116 输出由视图变换估计单元 114 估计的在每一时间处的视图变换以及每一时间处的视图模型与图像之间的对应关系。例如,所输出的视图变换以

及所输出的视图模型与图像之间的对应关系可以表示学习的前景在图像中所位于的位置、学习的前景的大小以及学习的前景的取向。

[0106] 视图模型输出单元 113 输出由视图模型学习单元 111 学习的前景的视图模型。在此,在识别另一输入活动图像时,可以使用所输出的视图模型(几何模型和特征量模型)。例如,当获得学习的前景在给定图像中所位于的位置、学习的前景的大小以及学习的前景的取向时,可使用从视图模型输出单元 113 输出的视图模型。

[0107] 图 3 中背景估计学习单元 105 的视图模型学习单元 121 到视图变换信息输出单元 126 是与前景估计学习单元 104 的视图模型学习单元 111 到视图变换信息输出单元 116 相同的功能块,这里不再重复描述。当然,背景估计学习单元 105 估计背景视图模型的视图变换并学习背景视图模型。

[0108] 图 4 为示出根据本技术的另一实施例的视图模型学习设备 200 的配置的一个示例的框图。当学习前景和背景时,图 4 的视图模型学习设备 200 使用多视图模型。

[0109] 图 4 的视图模型学习设备 200 包括图像序列输入单元 201、图像特征提取单元 202、前景背景合成观察单元 203、前景估计学习单元 204 以及背景估计学习单元 205。

[0110] 前景估计学习单元 204 包括多视图模型视图转变学习单元 211、多视图模型参数 212、视图转变模型参数 213 以及多视图模型视图转变输出单元 214。前景估计学习单元 204 还包括视图和变换估计单元 215、视图变换运动估计单元 216,视图转变估计单元 217 以及视图和变换信息输出单元 218。

[0111] 背景估计学习单元 205 包括多视图模型视图转变学习单元 221、多视图模型参数 222、视图转变模型参数 223 以及多视图模型视图转变输出单元 224。背景估计学习单元 205 还包括视图和变换估计单元 225、视图变换运动估计单元 226、视图转变估计单元 227 以及视图和变换信息输出单元 228。

[0112] 如上所述,图 4 的图像序列输入单元 201 是对输入图像序列输入进行控制并向提供图像特征提取单元 20 提供使得与相应的时间对应的图像作为输入图像序列的功能块。

[0113] 图像特征提取单元 202 从作为输入图像序列的在各时间处的图像中提取图像特征。图像特征的具体示例罢课每一像素的亮度值和多个像素形成的每一区域的亮度直方图、色彩直方图及边缘直方图。或者,通过哈里斯角点检测而检测到的每一特征点的坐标以及具有方向可控滤波器等的特征的局部特征可被用作图像特征。

[0114] 原则上,图像特征包括几何元素(几何模型)和特征量元素(特征量模型),如下文描述的视图模型的视图模型参数。例如,图像特征可包括与一帧对应的图像的每一像素的像素位置(几何模型)和像素值(特征量模型)。或者,图像特征可包括与一帧对应的图像中的位于预定特征点的像素的像素位置(几何模型)和像素值(特征量模型)。

[0115] 当使用局部特征时,如下文所述的,图像特征可仅包括与一帧对应的图像中预定特征点的像素的像素位置(几何模型)。

[0116] 前景背景合成观察单元 203 合成前景估计学习单元 204 输出(估计)的估计的前景视图以及背景估计学习单元 205 输出(估计)的估计的背景视图。前景背景合成观察单元 203 将合成图像的图像特征与从图像特征提取单元 202 提供的图像特征相比较,并分别向前景估计学习单元 204 和背景估计学习单元 205 提供有关所述估计的估计值的反馈信息。

[0117] 也就是说,前景背景合成观察单元 203 从前景估计学习单元 204 获得所估计的前景视图。另外,不同于视图模型学习设备 100,由于视图模型学习设备 200 使用多视图模型,视图模型学习设备 200 通过在从多个视图模型中选择的预定视图模型上进行视图变换来获取所估计的前景视图。

[0118] 前景背景合成观察单元 203 从背景估计学习单元 205 获得所估计的背景视图。另外,不同于视图模型学习设备 100,由于视图模型学习设备 200 使用多视图模型,视图模型学习设备 200 通过在从多个视图模型中选择的预定视图模型上进行视图变换来获取所估计的背景视图。

[0119] 如同视图模型学习设备 100,前景背景合成观察单元 203 形成前景和背景的合成视图模型。

[0120] 如同视图模型学习设备 100,前景背景合成观察单元 203 确定从图像特征提取单元 202 提供的图像特征与合成视图模型的参数之间的对应关系。

[0121] 前景背景合成观察单元 203 针对上述对应关系下的图像特征 X_s 来计算合成视图模型的估计值 E_s 。

[0122] 图 4 的前景估计学习单元 204 如下操作。

[0123] 如同图 3 的视图变换估计单元 114,例如,视图和变换估计单元 215 估计适合于时间 s 处的图像特征 X_s 的视图变换 T_s 。不同于视图变换估计单元 114,视图和变换估计单元 215 基于下述视图转变估计单元 217 的估计来估计适当的视图模型。

[0124] 例如,视图和变换估计单元 215 输出适当的视图模型与视图变换的组合的多个候选。例如,视图和变换估计单元 215 输出诸如视图模型 VM1 和视图变换 T_1 的组合以及视图模型 VM2 和视图变换 T_2 的组合等候选。

[0125] 如同图 3 的视图变换运动估计单元 115,视图变换运动估计单元 216 被配置用于根据时间 s 的视图变换 T_s 来估计的时间 $s+1$ 的视图变换 T'_{s+1} 。

[0126] 更具体地,视图变换运动估计单元 216 被配置用于基于前景背景合成观察单元 203 在给定时间输出的估计值 E_s 来计算每一视图变换的后验概率。视图和变换估计单元 215 被配置用于通过基于每一视图变换的后验概率来选择并输出在随后的时间的适当的视图变换来估计视图变换。

[0127] 多视图模型参数 212 例如具有 L 个视图模型,并通过多视图模型 $M_v = \{M_1, M_2, M_3, \dots, M_L\}$ 来配置。例如, $M_1, M_2, M_3, \dots, M_L$ 被视为对应于通过从正面、后面和侧面等来观看要学习的前景而得到的图像。

[0128] 在多视图模型的情况下, $M_1, M_2, M_3, \dots, M_L$ 的每一个具有几何模型和特征量模型。 $M_1, M_2, M_3, \dots, M_L$ 中的每一个被称为视图模型。

[0129] 视图转变估计单元 217 根据时间 s 处的视图模型 $M_{v,s}$ 来估计时间 $s+1$ 处的视图模型 $M'_{v,s+1}$ 。在此,由于假定要学习的前景具有三维结果,因此可假定例如满足等式 $M'_{v,s+1} = F(M_{v,s})$ 的模型。例如,可使用诸如 HMM 的动态学习估计模型作为该模型。

[0130] 更具体地,视图转变估计单元 217 被配置用于基于前景背景合成观察单元 203 在给定时间输出的估计值 E_s 来计算对应于每一视图转变的视图模型的后验概率。视图和变换估计单元 215 被配置用于通过基于每一视图模型的后验概率在随后的时间选择并输出适当的视图模型来估计视图模型(视图转变, view transition)。

[0131] 如同视图模型学习单元 111,多视图模型视图转变学习单元 211 学习前景多视图模型。也就是说,多视图模型视图转变学习单元 211 相对于基于前景背景合成观察单元 203 确定的对应关系来选择的特征点,针对输入图像序列中在每一时间的图像图像特征,来学习上述几何模型和特征量模型。在这种情况下,例如,与多个视图模型 $M_1, M_2, M_3, \dots, M_L$ 中的每一个对应地来学习几何模型和特征量模型。

[0132] 不同于视图模型学习单元 111,多视图模型视图转变学习单元 211 根据需要来学习视图转变模型。

[0133] 在此,视图转变模型被视为动态模型。具体地,根据需要来学习预定的动态模型(如 HMM)的参数。另外,可以预先给出动态模型的参数。

[0134] 下面会详细描述动态模型。另外,下文会详细描述用于视图模型的参数(包括几何模型的参数和特征量模型的参数)的计算以及这些参数的学习。

[0135] 视图转变模型参数 213 存储上述的动态模型的参数。例如,将 HMM 的状态转变的每一值作为参数来存储。相应地,前景视图转变模型中的实数(actula numbers)等也存储于视图转变模型参数 213。

[0136] 此时,存储于视图转变模型参数 213 的参数是与要学习的前景对应的方式存储的。例如,当将笔记本类型的个人计算机作为前景来学习时,赋予用于指定所述笔记本类型的个人计算机的索引等,并且 HMM 的状态转变概率的每一值等被作为参数来存储。另外,例如,当将玩具作为前景来学习时,赋予用于指定该玩具的索引等,并且 HMM 的状态转变概率的每一值等被作为参数来存储。

[0137] 这样,能够根据要学习的对象的运动来有效地学习例如视图转变的动态模型。

[0138] 视图和变换信息输出单元 218 输出由视图和变换估计单元 215 估计的每一时间处的视图模型和视图变换以及每一时间处的视图模型与图像之间的对应关系。例如,输出的视图变换以及输出的图像与视图模型之间的对应关系可表示学习的前景在图像中所处的位置、学习的前景的大小以及学习的前景的取向。

[0139] 多视图模型视图转变输出单元 214 输出由多视图模型视图转变学习单元 211 学习的前景多视图模型。在此,输出的多视图模型和视图转变模型可在识别另一输入活动图像时使用。例如,从视图模型输出单元 113 输出的视图模型可在获得了学习的前景在给定图像中的位置、学习的前景的大小以及学习的前景的取向时使用。

[0140] 由于图 4 中的背景估计学习单元 205 的多视图模型视图转变学习单元 221 到视图和变换信息输出单元 228 是与前景估计学习单元 204 的多视图模型视图转变学习单元 221 到视图和变换信息输出单元 228 相同的功能个,因此,这里不再重复描述。当然,背景估计学习单元 205 估计背景多视图模型的视图转变和视图变换,并学习多视图模型和背景视图转变模型。

[0141] 接下来将描述图 3 的视图模型学习单元 111 的视图模型参数的学习。

[0142] 在本技术的实施例中,使用 EM 算法来计算视图模型参数。也就是说,视图模型参数是通过用随机生成模型的模型等式来取代上述等式(1)并应用 EM 算法来计算和学习的。另外,例如,下列文献中详细披露了所述 EM 算法:Michalis K.Titsias、Doctor Thesis 所著的"Unsupervised Learning of Multiple Objects in Images"(爱丁堡大学,2005 年)。

[0143] 首先,视图模型参数 M 表示为 $m_1, m_2, m_3, \dots, m_N$ 。实质上,视图模型参数 M 包括几何模型参数 (例如,每一特征点的坐标值等) $mG_1, mG_2, mG_3, \dots, mG_N$ 以及特征量模型参数 (例如,每一特征点的像素值等) $mF_1, mF_2, mF_3, \dots, mF_N$ 。下文中,视图模型参数 M 被表示为 $M = \{m_1, m_2, m_3, \dots, m_N\}$,几何模型参数被表示为 $MG = \{mG_1, mG_2, mG_3, \dots, mG_N\}$ 并且特征量模型参数被表示为 $MF = \{mF_1, mF_2, mF_3, \dots, mF_N\}$ 。

[0144] 假设输入图像序列的在时间 s 处的图像特征 X_s 包括 $x_1, x_2, x_3, \dots, x_N$, 作为参数。实质上,图像特征 X_s 的参数包括几何模型参数 (例如,各特征点的坐标值) $xG_1, xG_2, xG_3, \dots, xG_N$ 以及特征量模型参数 (例如,各特征点的像素值) $xF_1, xF_2, xF_3, \dots, xF_N$ 。下文中,图像特征 X_s 的参数被表示为 $X_s = \{x_1, x_2, x_3, \dots, x_N\}$,几何模型参数被表示为 $X_{G,s} = \{xG_1, xG_2, xG_3, \dots, xG_N\}$, 并且特征量模型参数被表示为 $X_{F,s} = \{xF_1, xF_2, xF_3, \dots, xF_N\}$ 。

[0145] 另外, $T_1, T_2, \dots, T_{100}$ 作为在时间 s 输出的视图变换的候选。下文中,视图变换的候选被表示为 $T = \{T_1, T_2, \dots, T_{100}\}$ 。假设存在 NT 种类型的视图变换。

[0146] 基于上述假定,对应于等式 (1) 的随机生成模型的模型等式可被表示为等式 (3)。

$$[0147] \quad P(X_s | M, T_{s,k}) = \sum_{k=1}^{100} P_{T_k} \cdot C_{s,k} T_k M \quad (3)$$

[0148] 在该等式中, P_{T_k} 表示视图变换 T_k 的先验概率并且 T_k 表示在时间 s 处输出的视图变换的第 k 个候选。另外, $C_{s,k}$ 表示当在时间 s 处进行视图变换 T_k 时图像特征参数与视图模型参数的参数之间的对应关系。

[0149] 如上所述,视图模型参数包括几何模型参数和特征量模型参数。相应地,根据等式 (3),针对几何模型的随机生成模型的模型等式可表示为等式 (4),并且针对特征量模型的随机生成模型的模型等式可表示为等式 (5)。

$$[0150] \quad P(X_{G,s} | M_G, T_{s,k}) = \sum_{k=1}^{100} P_{T_k} \cdot C_{s,k} T_k M_G \quad (4)$$

$$[0151] \quad P(X_{F,s} | M_F, T_{s,k}) = \sum_{k=1}^{100} P_{T_k} \cdot C_{s,k} M_F \quad (5)$$

[0152] 被表示为上述随机生成模型的视图模型的学习可被归纳为相对于视图模型参数、针对所有时间处的图像特征 $X_1, X_2, \dots, X_s$ 最大化 $P(X|M)$ 的对数似然性 LH 的最大似然估计的问题。可利用 EM 算法来解该问题。另外, $P(X|M)$ 的对数似然性 LH 用等式 (6) 来表示。在此,假设 N_x 为所提取的图像特征的数量 (时间 s 的数量)。

$$[0153] \quad LH = \sum_{s=1}^{N_x} \log P(X_s | M) \quad (6)$$

[0154] 当给定视图模型 $M = \{m_1, m_2, m_3, \dots, m_N\}$ 时, EM 算法的 E 步骤对应于针对在每一时间处的图像特征 $X_1, X_2, \dots, X_s$ 来计算视图变换 $T = \{T_1, T_2, \dots, T_{100}\}$ 的后验概率

的步骤。也就是说,当提取(观察)图像特征 X_s 时,可用等式 (7) 将后验概率计算为视图变换为 T_k 时的概率 $P(T_k|X_s)$ 。

$$[0155] \quad P(T_k|X_s) = \frac{P_{T_k} \cdot P(X_s|T_k)}{\sum_{k=1}^{N_k} P_{T_k} \cdot P(X_s|T_k)} \quad (7)$$

[0156] 在等式 (7) 中, $P(X_s|T_k)$ 为似然性并且由图 3 的前景背景合成观察单元 103 来计算。似然性 $P(X_s|T_k)$ 被用作上述估计值 E_s 。在等式 (7), P_{T_k} 为视图变换 T_k 由图 3 的视图变换运动估计单元 115 来输出和计算的先验概率。然后,视图变换估计单元 114 最后计算等式 (7)。

[0157] 另一方面,在 EM 算法的 M 步骤,当给出在每一时间处的每一视图变换的后验概率 $P(T_k|X_s)$ 时,计算视图模型参数。 $M = \{m_1, m_2, m_3, \dots, m_N\}$ 可用等式 (8) 来计算。

$$[0158] \quad M = \frac{1}{S} \sum_{s=1}^S \sum_{k=1}^{N_T} P(T_k|X_s) \cdot T_k^{-1} C_{s,k} X_s \quad (8)$$

[0159] 由图 3 的视图模型学习单元 111 来计算等式 (8)。基于用等式 (8) 计算的参数 $\{m_1, m_2, m_3, \dots, m_N\}$ 来逐渐更新已经存储于视图模型参数 112 中的视图模型参数。这样,视图模型参数被学习。另外,在等式 (8),假定存在总共 N_T 个视图变换。

[0160] 也就是说, N_T 个视图变换中的每一个的后验概率每次由视图变换估计单元 114 计算,并被存储于视图变换运动估计单元 115。视图变换估计单元 114 基于从视图变换运动估计单元 115 输出的后验概率来估计视图变换,以生成所估计的前景视图,从而形成合成视图模型。

[0161] 由前景背景合成观察单元 103 将合成视图模型的视图模型参数与从实际输入的图像中提取的图像特征相比较,并计算估计值。基于计算的估计值,视图变换估计单元 114 计算每一视图变换的后验概率,并且视图模型学习单元 111 通过所述后验概率以加权的方式更新视图模型参数。

[0162] 至此已经描述了图 3 的视图模型学习单元 111 的视图模型参数的学习。然而,图 3 的视图模型学习单元 121 的视图模型参数以这样的方式进行学习。

[0163] 接下来描述图 4 的多视图模型视图转变单元 211 中的视图模型参数的学习。

[0164] 首先,假设多视图模型 M_v 包括 $\{M_1, M_2, M_3, \dots, M_L\}$ 。在此,视图模型 M_1 包括视图模型参数 $\{m_{11}, m_{12}, m_{13}, \dots, m_{1N}\}$, 并且视图模型 M_2 包括视图模型参数 $\{m_{21}, m_{22}, m_{23}, \dots, m_{2N}\}$ 。这样,每一视图模型包括视图模型参数。如上所述,实质上,视图模型参数包括几何模型的参数和特征量模型的参数。

[0165] 另外,假设输入图像序列在时间 s 处的图像特征 X_s 包括 $x_1, x_2, x_3, \dots, x_N$, 作为参数。实质上,图像特征 X_s 的参数包括几何模型的参数(例如,特征点的坐标值) $xG_1, xG_2, xG_3, \dots, xG_N$ 以及特征量模型的参数(例如,特征点的像素值) $xF_1, xF_2, xF_3, \dots, xF_N$ 。下文中,图像特征 X_s 的参数被表示为 $X_s = \{x_1, x_2, x_3, \dots, x_N\}$, 几何模型的参数被表示为 $X_G = \{xG_1, xG_2, xG_3, \dots, xG_N\}$, 以及特征量模型的参数被表示为 $X_{F,s} = \{xF_1, xF_2, xF_3, \dots,$

xFN}。

[0166] 另外, $T_1, T_2, \dots, T_{100}$ 被示出作为在时间 s 处输出的视图变换的候选。下文中, 视图变换的候选被表示为 $T = \{T_1, T_2, \dots, T_{100}\}$ 。假设存在 NT 种类型的视图变换。

[0167] 基于上述假定, 对应于等式 (2) 的随机生成模型的模型等式可表示为等式 (9)。等式 (9) 表示视图模型 M_v 以及相对于图像特征量 X_s 的、视图变换 T_k 的估计值。等式 (9) 由图 4 的前景背景合成观察单元 203 来计算, 并被提供到视图和变换估计单元 215 和多视图模型视图转变学习单元 211 作为估计值 E_s 。

$$[0168] \quad P(X_s | M_v, T_k) = \sum_{v=1}^L P_{M_v} \sum_{k=1}^{100} P_{T_k} \cdot C_{s,k} T_k M_v \quad (9)$$

[0169] 在该等式中, P_{M_v} 表示多视图模型中视图模型 M_v 的先验概率, P_{T_k} 表示视图变换 T_k 的先验概率, 并且 T_k 表示在时间 s 处输出的、视图变换的第 k 候选。另外, $C_{s,k}$ 表示当在时间 s 处执行视图变换 T_k 时图像特征参数与视图模型参数的参数之间的对应关系。

[0170] 可象上述单视图的情况那样来表示用于几何模型的随机生成模型的模型等式以及用于特征量模型的随机生成模型的模型等式, 这里不再重复。

[0171] 被表示为上述随机生成模型的视图模型的学习可被归纳为相对于视图模型参数、针对所有时间处的图像特征 $X_1, X_2, \dots, X_s$ 最大化 $P(X|M)$ 的对数似然性 LH 的最大似然估计的问题。可用 EM 算法来解该问题。另外, $P(X|M)$ 的对数似然性 LH 用等式 (10) 来表示。

$$[0172] \quad LH = \sum_{x=1}^{N_x} \sum_{l=1}^L \log P(X_s | M_v) \quad (10)$$

[0173] EM 算法的 E 步骤对应于在多视图模型 $M_v = \{M_1, M_2, M_3, \dots, M_L\}$ 的视图模型 $M_1 = \{m_{11}, m_{12}, m_{13}, \dots, m_{1N}\}$ 、视图模型 $M_2 = \{m_{21}, m_{22}, m_{23}, \dots, m_{2N}\}$ 等的情况下、针对每一时间处的图像特征 $X_1, X_2, \dots, X_s$ 计算视图变换 $T = \{T_1, T_2, \dots, T_{100}\}$ 和多视图模型 $M_v = \{M_1, M_2, M_3, \dots, M_L\}$ 的后验概率的步骤。也就是说, 当提取 (观察) 图像特征 X_s 时, 可用等式 (11) 将后验概率计算作为视图变换为 M_v 时的概率。另外, 当提取 (观察) 图像特征 X_s 时, 可用等式 (12) 将后验概率计算作为视图变换为 T_k 时的概率 $P(T_k | X_s)$ 。

$$[0174] \quad P(M_v | X_s) = \frac{P_{M_v} \cdot P(X_s | M_v)}{\sum_{v=1}^L P_{M_v} \cdot P(X_s | M_v)} \quad (11)$$

$$[0175] \quad P(T_k | X_s) = \frac{P_{T_k} \cdot P(X_s | T_k)}{\sum_{k=1}^{N_T} P_{T_k} \cdot P(X_s | T_k)} \quad (12)$$

[0176] 在等式 (11), $P(X_s | T_v)$ 是与视图变换关联的似然性, 并且由视图和变换估计单元

215 基于图 4 的前景背景合成观察单元 203 用等式 (9) 计算的估计值来计算。也就是说,在通过等式 (9) 计算的估计值中,与每一视图模型 M_v 关联的似然性可通过提取一个视图模型 M_v 并计算各个视图模型的估计值的总和来获得。在等式 (11), P_{mv} 为视图模型 M_v 由图 4 的视图转变估计单元 217 估计和计算时的概率。然后,视图和变换估计单元 215 最后计算等式 (11)。

[0177] 在等式 (12), $P(X_s|Tk)$ 是与视图变换关联的似然性,并且由视图和变换估计单元 215 基于图 4 的前景背景合成观察单元 203 用等式 (9) 计算的估计值来计算。也就是说,在用等式 (9) 计算的估计值中,与视图变换 Tk 中的每一视图模型关联的似然性可通过提取一个视图变换 Tk 并计算各视图模型的估计值的总和来获得。在等式 (12), P_{Tk} 是视图变换 Tk 由图 4 的视图变换运动估计单元 216 来估计和计算时的概率。然后,视图和变换估计单元 215 最后计算等式 (12)。

[0178] 另一方面,在 EM 算法的 M 步骤中,当给定每一时间处的每一视图变换的后验概率 $P(M_v|X_s)$ 以及在每一时间处的每一视图变换的后验概率 $P(Tk|X_s)$ 时,计算视图模型参数。 $M_v = \{mv1, mv2, mv3, \dots, mvN\}$ 可用等式 (13) 计算。

$$[0179] \quad M_v = \frac{1}{S} \sum_{s=1}^S P(M_v|X_s) \sum_{k=1}^{N_T} P(Tk|X_s) \cdot T_k^{-1} C_{s,k} X_s \quad (13)$$

[0180] 等式 (13) 由图 4 的多视图模型视图转变学习单元 211 计算。基于通过等式 (13) 计算的视图模型参数 $\{mv1, mv2, mv3, \dots, mvN\}$, 逐渐更新已经存储于多视图模型参数 212 中的各模型的视图模型参数。另外,在等式 (13), 假定存在总共 N_T 个视图变换。

[0181] 也就是说, N_T 个视图变换中的每一个的后验概率每次由视图和变换估计单元 215 计算, 并被存储于视图变换运动估计单元 216 中。视图和变换估计单元 215 每次计算 L 个视图模型 (视图转变的结果) 中的每一个的后验概率, 并将 L 个视图模型中的每一个的后验概率存储在视图转变估计单元 217 中。另外, 视图和变换估计单元 215 基于从视图变换运动估计单元 216 输出的后验概率来估计视图变换, 并基于从视图转变估计单元 217 输出的后验概率来估计视图转变, 以生成所估计的前景视图, 从而形成合成视图模型。

[0182] 由前景背景合成观察单元 203 将合成视图模型的视图模型参数与从实际输入的图像中提取的图像特征相比较, 并计算估计值。基于所计算的估计值, 视图和变换估计单元 215 计算每一视图变换的后验概率, 并计算每一视图模型的后验概率。另外, 多视图模型视图转变学习单元 211 以加权的方式用所述后验概率来更新视图模型参数。

[0183] 多视图模型的 L 个视图模型中的一个视图模型的视图模型参数通过计算等式 (13) 来计算。相应地, 等式 (13) 的计算是利用多视图模型的多个视图模型来计算的。

[0184] 这样, 视图模型参数得以学习。

[0185] 如上所述, 多视图模型视图转变学习单元 211 学习视图转变模型。例如, 基于等式 (11) 中的 $P(M_v|X_s)$ 来计算和学习视图转变模型的参数。

[0186] 至此已经描述了图 4 的多视图模型视图转变学习单元 211 的学习。然而, 图 4 的多视图模型视图转变学习单元 221 以这样的方式进行学习。

[0187] 接下来, 参考图 5 的流程图, 描述图 3 的视图模型学习设备 100 执行的多视图学习处理的一个示例。

[0188] 在步骤 S21, 图像序列输入单元 101 接收活动图像的输入。然后, 与各时间对应的帧图像作为输入图像序列被提供到图像特征提取单元 102。

[0189] 在步骤 S22, 图像特征提取单元 102 从在步骤 S21 的处理中输入的输入图像序列的、在每一时间处的图像中提取图像特征。

[0190] 在步骤 S23, 前景估计学习单元 104 进行下文参考图 6 的流程图描述的前景估计处理。然后, 前景估计学习单元 104 输出通过在前景视图模型上进行多个视图变换而得到的所估计的前景视图。

[0191] 在步骤 S24, 背景估计学习单元 105 进行下文参考图 7 的流程图描述的背景估计处理。然后, 背景估计学习单元 105 输出通过在背景视图模型上进行多个视图变换而得到的所估计的背景视图。

[0192] 在步骤 S25, 前景背景合成观察单元 103 进行下文参考图 8 的流程图描述的合成观察处理。然后, 前景背景合成观察单元 103 合成在步骤 S23 的处理中输出的所估计的前景视图以及在步骤 S24 的处理中输出的所估计的背景视图, 以生成合成视图模型, 并针对在步骤 S22 的处理中提取的图像特征来计算合成视图模型的估计值。

[0193] 实质上, 步骤 S23 到步骤 S25 的处理被反复执行活动图像的长度, 然后视图模型学习处理进行到步骤 S26。

[0194] 在步骤 S26, 前景估计学习单元 104 进行下文参考图 9 的流程图描述的前景学习处理。然后, 基于在步骤 S25 的处理中得到的估计值来更新前景的视图模型参数。

[0195] 在步骤 S27, 背景估计学习单元 105 进行下文参考图 10 的流程图描述的背景学习处理。然后, 基于在步骤 S25 的处理中得到的估计值来更新背景的视图模型参数。

[0196] 实质上, 例如, 步骤 S23 到步骤 S27 的处理被反复执行预定次, 或被反复执行直到等式 (6) 中表示的对数似然性的变化量等于或小于预定阈值。

[0197] 这样, 执行了视图模型学习处理。

[0198] 接下来, 参考图 6 来描述图 5 的步骤 S23 的前景估计处理的一个具体示例。

[0199] 在步骤 S41, 视图变换运动估计单元 115 根据在时间 $s-1$ 处的视图变换 T_{s-1} 来估计在时间 s 处的视图变换 T'_s 。一般而言, 由于假定前景的运动是规则或者连续的, 因此, 通过使用诸如粒子滤波器等动态估计模型, 针对变换的动态来估计前景的运动。

[0200] 在步骤 S42, 视图变换估计单元 114 基于视图变换运动估计单元 115 在步骤 S41 的估计结果来估计在时间 s 处的适当的视图变换 T_s 。此时, 视图变换估计单元 114 输出适当的视图变换 T_s 的多个候选。例如, 视图变换估计单元 114 输出具有上百个不同参数的仿射变换, 作为视图变换的候选。具体地, 视图变换估计单元 114 输出用于平移 (translate) 前景视图模型的变换、用于扩展并缩小前景视图模型的变换、用于转动前景视图模型的变换、用于进行变换的組合的仿射变换或投影变换。

[0201] 在步骤 S43, 视图变换估计单元 114 在前景视图模型上执行在步骤 S42 的处理中估计的多个视图变换。此时, 前景视图模型的参数从视图模型参数 112 中读出, 并经历视图变换。

[0202] 在步骤 S44, 视图变换估计单元 114 输出作为步骤 S43 的处理的结果而得到的估计的前景视图。在此, 输出与视图变换的候选对应的多个估计的前景视图。

[0203] 这样, 执行了前景估计处理。

[0204] 接下来,参考图 7 来描述图 5 的步骤 S24 的背景估计处理一个具体示例。

[0205] 在步骤 S61,视图变换运动估计单元 125 根据时间 $s-1$ 处的视图变换 T_{s-1} 来估计在时间 s 处的视图变换 T'_s 。一般而言,一般而言,由于假定背景的运动是规则或者连续的,因此,通过使用诸如粒子滤波器等动态估计模型,针对变换的动态来估计背景的运动。

[0206] 在步骤 S62,视图变换估计单元 124 基于视图变换运动估计单元 125 在步骤 S61 的估计结果来估计在时间 s 处的适当的视图变换 T_s 。此时,视图变换估计单元 124 输出适当的视图变换的多个候选。例如,视图变换估计单元 124 输出具有上百个不同参数的仿射变换,作为视图变换的候选。具体地,例如,视图变换估计单元 124 输出用于平移 (translate) 前景的图像的仿射变换、用于扩展并缩小前景的图像的仿射变换、用于转动前景的图像的仿射变换以及模式仿射变换 (如用于投影前景的图像的仿射变换)。

[0207] 在步骤 S63,视图变换估计单元 124 在背景视图模型上执行在步骤 S62 的处理中估计的多个视图变换。此时,背景视图模型的参数从视图模型参数 122 读出,并经历视图变换。

[0208] 在步骤 S64,视图变换估计单元 124 输出作为步骤 S63 的處理的结果而得到的估计的背景视图。在此,输出与视图变换的候选对应的多个估计的背景视图。

[0209] 或者,在背景估计处理的情况下,在步骤 S62 中仅输出视图变换的一个候选,而在步骤 S64 中可输出所估计的一个背景视图。这是因为,当作为前景的对象正在移动时,与前景相比,背景的运动足够小。

[0210] 这样,执行了背景估计处理。

[0211] 接下来,参考图 8 的流程图来描述在图 5 的步骤 S25 中执行的合成观察处理的一个具体示例。

[0212] 在步骤 S81,前景背景合成观察单元 103 形成前景和背景的合成视图模型。

[0213] 此时,前景背景合成观察单元 103 合成在图 6 的步骤 S44 的处理中输出的所估计的前景视图以及在图 7 的步骤 S64 的处理中输出的所估计的背景视图。另外,前景背景合成观察单元 103 确定在图 5 的步骤 S22 的处理中提取的图像特征的特征与合成视图模型的参数之间的对应关系。

[0214] 在步骤 S82,前景背景合成观察单元 103 相对于在步骤 S22 的处理中提取的图像特征,计算在步骤 S81 的处理中形成的合成视图模型的估计值。

[0215] 这样,执行了合成观察处理。

[0216] 接下来,参考图 9 的流程图来描述图 5 的步骤 S26 的前景学习处理的一个具体示例。

[0217] 在步骤 S101,视图变换估计单元 114 和视图模型学习单元 111 中的每一个获取估计值,该估计值是在图 8 的步骤 S82 的处理中计算的估计值并且是在每一时间处的合成视图模型的估计值。在这种情况下,获取似然性 $P(X_s|T_k)$,作为与每一时间对应的估计值。

[0218] 在步骤 S102,视图变换估计单元 114 计算每一视图变换的后验概率。此时,当提取 (观察) 图像特征 X_s 时,用上述等式 (7) 来将后验概率计算作为视图变换为 T_k 时的概率 $P(T_k|X_s)$ 。

[0219] 在等式 (7), P_{T_k} 为通过视图变换 T_k 输出的概率。概率 P_{T_k} 由视图变换运动估计单元 115 以与每一视图变换对应的方式计算,并被提供到视图变换估计单元 114。

[0220] 在步骤S103,视图模型学习单元111计算视图模型参数。也就是说,当给定了在步骤S102的处理中得到的在每一时间处的每一视图变换的后验概率 $P(T_k|X_s)$ 时,视图模型学习单元111利用上述等式(8)来计算视图模型参数。

[0221] 在步骤S104,视图模型学习单元111基于在步骤S103的处理中得到的视图模型参数来更新视图模型参数112。

[0222] 这样,进行了前景学习处理。

[0223] 接下来,参考图10的流程图描述图5的步骤S27的背景学习处理的一个具体示例。

[0224] 在步骤S121,视图变换估计单元124和视图模型学习单元121中的每一个获取估计值,该估计值是在图8的步骤S82的处理中计算的估计值并且是在每一时间处的合成视图模型的估计值。在这种情况下,获得似然性 $P(X_s|T_k)$,作为与每一时间对应的估计值。

[0225] 在步骤S122,视图变换估计单元124计算每一视图变换的后验概率。此时,当提取(观察)图像特征 X_s 时,通过上述等式(7)将后验概率计算作为视图变换为 T_k 时的概率 $P(T_k|X_s)$ 。

[0226] 在等式(7), P_{T_k} 是通过视图变换 T_k 输出的概率。概率 P_{T_k} 由视图变换运动估计单元125以与每一视图变换对应的方式来计算,并被提供到视图变换估计单元124。

[0227] 在步骤S123,视图模型学习单元121计算视图模型参数。也就是说,当给定在步骤S122的处理中获得的在每一时间处的每一视图变换的后验概率 $P(T_k|X_s)$ 时,视图模型学习单元121利用上述的等式(8)来计算视图模型参数。

[0228] 在步骤S124,视图模型学习单元121基于在步骤S123的处理中得到的视图模型参数来更新视图模型参数122。

[0229] 这样,执行了背景学习处理。

[0230] 接下来,参考图11的流程图来描述图4的视图模型学习设备200执行的视图模型学习处理。

[0231] 在步骤S151,图像序列输入单元201接收活动图像的输入。然后,被使得与每一时间对应的图像作为输入图像的序列而被提供到图像特征提取单元202。

[0232] 在步骤S152,图像特征提取单元202从在步骤S151中输入的输入图像序列的在每一时间处的图像中提取图像特征。

[0233] 在步骤S153,前景估计学习单元204执行下文参考图12的流程图描述的前景估计处理。然后,前景估计学习单元204输出前景多视图模型的视图模型中的适当的视图模型的组合的多个候选以及多个视图变换中的适当的视图变换。例如,前景估计学习单元204输出视图模型VM1和视图变换T1的组合的候选以及视图模型VM2和视图变换T2的组合f的候选等。输出通过这样的组合而得到的每一估计的前景视图。

[0234] 在步骤S154,背景估计学习单元205执行下文参考图13的流程图描述的背景估计处理。然后,背景估计学习单元205输出背景多视图模型的视图模型中的适当的视图模型的组合的多个候选以及多个视图变换终端的适当的视图变换。输出通过这样的组合而得到的每一估计的背景视图。

[0235] 在步骤S155,前景背景合成观察单元203进行下文参考图14的流程图描述的合成观察处理。然后,前景背景合成观察单元203合成在步骤S153的处理中输出的所估计的前

景视图以及在步骤 S154 的处理中输出的所估计的背景视图,以生成合成视图模型,并针对在步骤 S152 的处理中提取的图像特征来计算合成视图模型的估计值。

[0236] 实质上,步骤 S153 到步骤 S155 的处理被反复执行活动图像的长度,然后,多视图模型学习处理进行到步骤 S156。

[0237] 在步骤 S156,前景估计学习单元 204 执行下文参考图 15 的流程图描述的前景学习处理。然后,基于在步骤 S155 的处理中得到的估计值来更新前景的视图模型参数。

[0238] 在步骤 S157,背景估计学习单元 205 执行下文参考图 16 的流程图描述的背景学习处理。然后,基于在步骤 S155 的处理中得到的估计值来更新背景的视图模型参数。

[0239] 实质上,例如,步骤 S23 到步骤 S27 的处理被反复执行预定次数,或者被反复执行直到等式 (10) 表示的对数似然性的变化量小于或等于预定阈值。

[0240] 这样,执行了视图模型学习处理。

[0241] 接下来,参考图 12 来描述图 11 的步骤 S153 的前景估计处理的一个具体示例。

[0242] 在步骤 S171,视图转变估计单元 217 根据在时间 $s-1$ 处的视图模型 $M_{v,s-1}$ 来估计在时间 s 处视图模型 $M'_{v,s}$ 。在此,由于假定要学习的前景中存在三维结构,因此,例如,使用诸如 HMM 等随机生成模型来估计视图模型的转变。

[0243] 在步骤 S172,视图和变换估计单元 215 和视图转变估计单元 217 基于在步骤 S171 的估计结果来估计视图模型的候选。此时,视图和变换估计单元 215 输出适当的视图模型的多个候选。

[0244] 在步骤 S173,视图变换运动估计单元 216 根据在时间 $s-1$ 处的视图变换 T_{s-1} 来估计在时间 s 处的视图变换 T'_s 。一般而言,由于假定前景的运动是规则连续的,因此,利用诸如粒子滤波器等动态模型针对变换的动态来估计前景的运动。

[0245] 在步骤 S174,视图和变换估计单元 215 基于视图变换运动估计单元 216 在步骤 S172 的估计结果来估计在时间 s 处的适当的视图变换 T_s 。此时,视图和变换估计单元 215 输出适当的视图变换的多个候选。例如,视图和变换估计单元 215 输出具有多个模式的仿射变换,作为仿射变换的候选。具体地,例如,视图和变换估计单元 215 输出用于平移前景的图像的仿射变换、用于扩展和缩小前景的图像的仿射变换、用于转动前景的图像的仿射变换以及模式的仿射变换(如用于投影前景的图像的仿射变换)。

[0246] 在步骤 S175,视图和变换估计单元 215 在步骤 S172 的处理中输出的每一视图模型上执行在步骤 S174 的处理中输出的视图变换。此时,所估计的每一视图模型的参数从多视图模型参数 212 中读出,并经历视图变换。

[0247] 此时,视图模型和视图变换的预定数量的组合被选择,并经历视图变换。例如,视图模型和视图变换的上百个组合(即视图模型 VM1 和视图变换 T1 的组合和视图模型 VM2 和视图变换 T2 的组合等)被选择,并且每一组合经历视图变换。

[0248] 在步骤 S176,视图和变换估计单元 215 输出作为步骤 S175 的处理的结果而得到的所估计的前景视图。在此,视图和变换估计单元 215 输出对应于视图模型和视图变换的组合的多个估计的前景视图。

[0249] 这样,执行了前景估计处理。

[0250] 接下来,参考图 13 来描述图 11 的步骤 S154 的背景估计处理的一个具体示例。

[0251] 在步骤 S191,视图转变估计单元 227 根据在时间 $s-1$ 处的视图模型 $M_{v,s-1}$ 来估计在

时间 s 处的视图模型 $M'_{v,s}$ 。在此,由于假定要学习的背景中存在三维结构,因此,例如,利用诸如 HMM 的随机生成模型来估计视图模型的转变。

[0252] 在步骤 S192,视图和变换估计单元 225 基于视图转变估计单元 227 在步骤 S191 的估计结果来估计视图模型的候选。此时,视图和变换估计单元 225 输出适当的视图模型的多个候选。

[0253] 在步骤 S193,视图变换运动估计单元 226 根据在时间 $s-1$ 处的视图变换 T_{s-1} 来估计在时间 s 处的视图变换 T'_s 。一般而言,由于假定背景的运动是规则或连续的,因此,通过利用诸如粒子滤波器的动态估计模型,针对变换的动态来估计背景的运动。

[0254] 在步骤 S194,视图和变换估计单元 225 基于视图变换运动估计单元 226 在步骤 S192 的估计结果来估计在时间 s 处的适当的视图变换 T_s 。此时,视图和变换估计单元 225 输出适当的视图变换的多个候选。例如,视图和变换估计单元 225 输出具有上百个不同模式的仿射变换,作为仿射变换的候选。具体地,视图变换估计单元 225 输出用于平移背景视图模型的变换、用于扩展和缩小背景视图模型的变换、用于转动背景视图模型的变换、用于执行变换的组的仿射变换或投影变换。

[0255] 在步骤 S195,视图和变换估计单元 225 在步骤 S192 的处理中输出的每一视图模型上执行在步骤 S194 的处理中输出的视图变换。此时,所估计的每一视图模型的参数从多视图模型参数 222 中读出,并经历视图变换。

[0256] 此时,视图模型和视图变换的预定数量的组合被选择,并经历视图变换。例如,视图模型和视图变换的上百个组合被选择,且每一组合经历视图变换。

[0257] 在步骤 S196,视图和变换估计单元 225 输出作为步骤 S195 的处理的结果而得到的所估计的背景视图。在此,视图和变换估计单元 215 输出与视图模型和视图变换的组合对应的所估计的多个背景视图。

[0258] 或者,在背景估计处理的情况下,在步骤 S192 中仅输出视图模型的一个候选,在步骤 S194 中仅输出视图变换的一个候选,而在步骤 S196 中可输出一个估计的背景视图。这是因为,例如当作为前景的对象正在移动时,与前景相比,背景的运动足够小。

[0259] 这样,执行了背景估计处理。

[0260] 接下来,参考图 14 的流程图来描述在图 11 的步骤 S155 中执行的合成观察处理的一个具体示例。

[0261] 在步骤 S211,前景背景合成观察单元 203 形成前景和背景的合成视图模型。

[0262] 此时,前景背景合成观察单元 203 合成在图 12 的步骤 S176 的处理中输出的所估计的前景视图以及在图 13 的步骤 S196 的处理中输出的所估计的背景视图。另外,前景背景合成观察单元 203 确定在图 5 的步骤 S152 中提取的图像特征的参数与合成视图模型的参数之间的对应关系。

[0263] 在步骤 S212,前景背景合成观察单元 203 相对于在步骤 S152 的处理中提取的图像特征,计算在步骤 S211 的处理中形成的合成视图模型的估计值。此时,前景背景合成观察单元 203 利用上述等式 (9) 来计算估计值。

[0264] 这样,执行了合成观察处理。

[0265] 接下来,参考图 11 的流程图来描述图 15 的步骤 S156 的前景学习处理的一个具体示例。

[0266] 在步骤 S231, 视图和变换估计单元 215 和多视图模型视图转变学习单元 211 中的每一个获取估计值, 该估计值是在图 14 的步骤 S212 的处理中计算的估计值并且是在每一时间处的合成视图模型的估计值。

[0267] 在步骤 S232, 视图和变换估计单元 215 基于在步骤 S231 中得到的估计值来计算与视图模型关联的似然性。在这种情况下, 以与每一时间对应的方式来计算似然性 $P(X_s | M_v)$ 。

[0268] 在步骤 S233, 视图和变换估计单元 215 基于在步骤 S231 中得到的估计值来计算与视图变换关联的似然性。在这种情况下, 以与每一时间对应的方式来计算似然性 $P(X_s | T_k)$ 。

[0269] 在步骤 S234, 视图和变换估计单元 215 计算每一视图模型的后验概率。此时, 当提取 (观察) 图像特征 X_s 时, 利用上述的等式 (11) 将后验概率计算作为视图模型为 M_v 时的概率 $P(M_v | X_s)$ 。

[0270] 在步骤 S235, 视图和变换估计单元 215 计算每一视图变换的后验概率。此时, 当提取 (观察) 图像特征 X_s 时, 利用上述的等式 (12) 将后验概率计算作为视图变换为 T_k 时的概率 $P(T_k | X_s)$ 。

[0271] 在等式 (11), P_{M_v} 是通过视图模型 M_v 输出的概率。概率 P_{M_v} 由视图转变估计单元 217 以与每一视图变换对应的方式计算, 并被提供到视图和变换估计单元 215。在等式 (12), P_{T_k} 是通过视图变换 T_k 输出的概率。概率 P_{T_k} 由视图变换运动估计单元 216 以与每一视图变换对应的方式计算, 并被提供到视图和变换估计单元 215。

[0272] 在步骤 S236, 多视图模型视图转变学习单元 211 计算视图模型参数。也就是说, 当给定在步骤 S234 的处理中得到的在每一时间的每一视图模型的后验概率 $P(M_v | X_s)$ 以及在步骤 S235 的处理中得到的在每一时间处的每一视图变换的后验概率 $P(T_k | X_s)$ 时, 多视图模型视图转变学习单元 211 利用上述等式 (13) 的来计算视图模型参数。

[0273] 在步骤 S237, 多视图模型视图转变学习单元 211 基于在步骤 S236 的处理中得到的视图模型参数来更新多视图模型参数 212。

[0274] 这样, 执行了前景学习处理。

[0275] 接下来, 参考图 16 的流程图来描述图 11 的步骤 S157 的背景学习处理一个具体示例。

[0276] 在步骤 S251, 视图和变换估计单元 225 和多视图模型视图转变学习单元 221 的每一个获取估计值, 该估计值是在图 14 的步骤 S212 的处理中计算的估计值并且是在每一时间处的合成视图模型的估计值。

[0277] 在步骤 S252, 视图和变换估计单元 225 基于在步骤 S251 中得到的估计值来计算与视图模型关联的似然性。此时, 以与每一时间对应的方式来计算似然性 $P(X_s | M_v)$ 。

[0278] 在步骤 S253, 视图和变换估计单元 225 在步骤 S251 中得到的估计值来计算与视图变换关联的似然性。此时, 以与每一时间对应的方式来计算似然性 $P(X_s | T_k)$ 。

[0279] 在步骤 S254, 视图和变换估计单元 225 计算每一视图模型的后验概率。此时, 当提取 (观察) 图像特征 X_s 时, 通过上述的等式 (11) 将后验概率计算作为视图模型为 M_v 时的概率 $P(M_v | X_s)$ 。

[0280] 在步骤 S255, 视图和变换估计单元 225 计算每一视图模型的后验概率。此时, 当提取 (观察) 图像特征 X_s 时, 通过上述的等式 (12) 将后验概率计算作为视图变换为 T_k 时的概率 $P(T_k | X_s)$ 。

[0281] 在等式 (11), P_{M_v} 是通过视图模型 M_v 输出的先验概率。The 先验概率 P_{M_v} 由视图转变估计单元 227 以与每一视图模型对应的方式计算, 并被提供到视图和变换估计单元 225。在等式 (12), P_{T_k} 是通过视图变换 T_k 输出的概率。概率 P_{T_k} 由视图变换运动估计单元 226 以与每一视图变换对应的方式来计算, 并被提供到视图和变换估计单元 225。

[0282] 在步骤 S256, 多视图模型视图转变学习单元 221 计算视图模型参数。也就是说, 当给定了在步骤 S254 的处理中得到的在每一时间处的每一视图模型的后验概率 $P(M_v|X_s)$ 以及在步骤 S255 的处理中得到的每一视图变换的后验概率 $P(T_k|X_s)$ 时, 多视图模型视图转变学习单元 221 利用上述等式 (13) 来计算视图模型参数。

[0283] 在步骤 S257, 多视图模型视图转变学习单元 221 基于在步骤 S256 的处理中得到的视图模型参数来更新多视图模型参数 222。

[0284] 这样, 执行了背景学习处理。

[0285] 例如, 在根据相关技术的学习图像处理中, 当学习识别目标的模型以形成识别器时, 学习是通过向作为学习数据的大量图像数据赋予识别目标的标记来执行的。

[0286] 例如, 当为面部识别而学习图像时, 需要赋予用于指定人名、他或她的面部取向、显示这人的面部图像的区域的信息等, 作为标记。另外, 当为对象识别而学习图像时, 需要赋予用于指定对象名称、对象取向以及显示对象的区域的信息等, 作为标记。

[0287] 例如, 当在学习图像处理中学习图像时, 赋予标记会增加处理负担。因此, 难以在短时间内完成学习。

[0288] 然而, 在本技术的实施例, 无需通过向大量的图像数据赋予识别目标的标记来学习图像。相应地, 例如, 通过简单地输入活动图像等来容易地完成学习。

[0289] 另外, 在本技术的实施例, 实现了建模, 使得对多视图模型中的各个视图模型的视图转变估计能够通过利用 HMM 等以统计的方式执行。因此, 例如, 无需完成针对每一视图的学习。另外, 即使当视图的数量增加时, 也可以简单地完成学习。

[0290] 此外, 在本技术的实施例, 图像中的前景和背景被分解, 并分别被学习为前景视图模型和背景视图模型。因此, 例如, 即使当实际活动图像中包含与前景接近的背景时, 也能够简单地完成学习。

[0291] 相应地, 与相关技术相比, 在本技术的实施例, 能够以较少的计算量来进行学习图像处理。

[0292] 至此, 已经基于前景和背景的每一个用单视图模型来表示并且在前景和背景上进行相同的视图变换的假定描述了以上参考图 5 到图 10 示出的处理。已经基于前景和背景的每一个用多视图模型来表示并且在前景和背景上进行相同的视图变换的假定描述了以上参考图 11 到图 16 示出的处理。

[0293] 然而, 实质上, 可向背景和前景应用不同的视图模型和不同的视图变换。

[0294] 例如, 可通过用多视图模型来表示前景而分别用单视图模型和多视图模型来表示背景来学习视图模型。

[0295] 或者, 可将诸如 HMM、FNN 或 RNN 等动态学习估计模型用于前景视图变换的运动估计。可将诸如粒子滤波器或卡尔曼滤波器等动态估计模型用于背景视图变换的运动估计。

[0296] 也就是说, 作为一个变型, 图 4 中的视图模型学习设备 200 的背景估计学习单元 205 的配置可被替代为图 3 的视图模型学习设备 100 的背景估计学习单元 105 的配置。另

外,例如,作为一个变型,前景估计学习单元 204 的视图变换运动估计单元 216 和视图转变估计单元 217 可集成在一起,并且可根据与背景估计学习单元 205 的视图转变估计和视图变换运动估计不同的方法来进行估计。

[0297] 这样,可以各种形式来修改根据本技术的实施例的视图模型学习和利用学习的结果来执行的图像识别。下文中,将描述一些示例,其中,具体地设置了前景的视图模型和视图变换以及背景的视图模型和视图变换,并且学习了视图模型,以识别图像。

[0298] 也就是说,将描述作为具体的视图模型学习和应用了本技术的实施例的图像处理设备的图像识别的示例的第一到第三具体示例。

[0299] 第一具体示例

[0300] 在第一具体示例,通过输入并学习活动图像(其中图 17 所示的前景图像正相对于图 18 所示的背景图像而移动)来识别存在前景图像的区域。

[0301] 在这种情况下,图 3 的视图模型学习设备 100 将前景学习为单视图模型,并且也将背景学习为单视图模型,然后利用作为学习结果的视图模型来识别前景图像。

[0302] 当在视图模型的学习之后进行图像识别时,象视图模型学习一样,也估计视图变换,或者也估计视图转变。然而,当图像被识别时,通常仅示出所述视图变换或视图变换的一个候选。

[0303] 图 17 示出的前景图像是通过 5 乘 5 个像素的矩形区域而形成的图像,并且具有相应的像素为白或黑的预定图案。

[0304] 图 18 示出的背景图像是通过 10 乘 20 个像素的矩形区域而形成的图像,并且具有相应的像素为白或黑的预定图案。

[0305] 在此,假设后面的活动图像被输入到图像序列输入单元 101。也就是说,在每一时间处的输入图像是在背景的预定位置处反复显示前景并且通过在随后的时间沿着向上、向下、向右和向左的方向中的一个方向将前景的位置随机移动一个像素而使前景相对于背景正在移动的输入图像序列。

[0306] 在第一具体示例,相应地,假定背景(或相机)不移动,而前景只平移。在第一具体示例,通过将输入图像序列的二进制图案(0 或 1)转换为灰度图案(0 到 1 的范围中的连续值)而得到的结果被用作图像特征。另外,前景视图模型 M_{FG} 和背景视图模型 M_{BG} 为具有预定数量的像素且每一像素的亮度值(像素值)为在 0 到 1 的范围中的连续值的矩形区域。

[0307] 图 17 所示的前景被合成到图 18 所示的背景上,以形成图 19 所示的按时间顺序的图像,然后,所述图像被输入到视图模型学习设备 100。图 19 所示的图像为从作为活动图像输入的图像中提取的在三个时间处的图像。在图 19,在时间 $s = s_1$,前景图像显示在区域 301 中,该区域 301 位于图像的中间的下部。在时间 $s = s_2$,前景图像显示在区域 302 中,该区域 302 位于图像的中间的略靠下的部分。在时间 $s = s_3$,前景图像显示在区域 303 中,该区域 303 位于图像的中间。

[0308] 如上所述,图像特征提取单元 102 提取通过将输入图像序列的二进制图案转换成灰度图案而得到的结果,作为图像特征。例如,图像特征提取单元 102 提取在时间 s 处的图像特征 X_s 。

[0309] 前景背景合成观察单元 103 合成由前景估计学习单元 104 估计的所估计的前景视图和由背景估计学习单元 105 估计的所估计的背景视图。所估计的前景视图可通过在前景

视图模型 M_{FG} 上进行预定视图变换 T_k 来获得。在这种情况下,由于背景中有运动,因此,背景视图模型 M_{BG} 可控用作所估计的没有改变的背景视图。前景背景合成观察单元 103 形成背景和前景的合成视图模型 $M_{k, FGBG}$ 。

[0310] 合成视图模型 $M_{k, FGBG}$ 通过表达式 (14) 来得到。

$$[0311] \quad \mathbf{M}_{k, FGBG} = \mathbf{M}_{BG} \oplus \mathbf{T}_k \mathbf{M}_{FG} \quad (14)$$

[0312] 等式 (14) 中,运算符 $\oplus$ 表示后面的计算。即,得到经过视图变换 T_k 的前景视图模型 M_{FG} 的几何模型 $T_k M_{FG, G}$ 以及与几何模型 $T_k M_{FG, G}$ 几何对应的背景视图模型 M_{BG} 的特征点 $M'_{BG, G}$ 。在这种情况下,获得背景中与前景的每一像素重叠的坐标位置,作为特征点 $M'_{BG, G}$ 。另外,特征点 $M'_{BG, G}$ 的特征量 $M'_{BG, F}$ 用前景视图模型 M_{FG} 的特征量模型 $M_{FG, F}$ 来替代。

[0313] 也就是说,用前景的每一像素的亮度值来替代背景中与前景的每一像素重叠的坐标位置处的每一像素的亮度值。

[0314] 由于如上所述在此假定前景仅平移,因此,等式 (16) 中的视图变换 T_k 用等式 (15) 中的仿射变换 (affine transform) 来替代。

$$[0315] \quad T_k = \begin{pmatrix} 1 & 0 & u_k \\ 0 & 1 & v_k \\ 0 & 0 & 0 \end{pmatrix} \quad (15)$$

[0316] 前景背景合成观察单元 103 计算前景和背景的合成视图模型 $M_{k, FGBG}$ 的似然性 $P(X_s | M_{k, FGBG})$, 作为估计值。另外,计算似然性 $P(X_s | M_{k, FGBG})$, 作为估计值。在这种情况下,由于前景和背景的合成视图模型 $M_{k, FGBG}$ 的似然性与视图变换 T_k 的似然性 $P(X_s | T_k)$ 相同,因此上述估计值可用等式 (16) 来表示。

$$[0317] \quad P(X_s | T_k) = P(X_s | M_{k, FGBG}) = \frac{1}{D} \exp \left[- \frac{\sum_{i,j}^{C_{FGBG,s}} (x_{s,i} - (m_{k, FGBG,j}))^2}{2\sigma^2} \right] \quad (16)$$

[0318] 在等式 (16), σ 和 D 为预先确定的参数。在等式 (16), $C_{FGBG,s}$ 表示输入图像序列中在时间 s 处的图像与合成视图模型之间的对应关系。也就是说,等式 (16) 的最右边的分子表示输入图像序列中在时间 s 处的图像的各个像素的亮度值与背景和前景的合成视图模型对应的各个像素的亮度值之间的差的平方和。

[0319] 视图变换估计单元 114 的视图变换估计以及视图模型学习单元 111 的视图模型学习是基于用等式 (16) 计算的似然性来执行的。

[0320] 在第一具体示例,视图变换估计单元 114 和视图变换运动估计单元 115 的估计是根据 HMM 算法来执行的。

[0321] 由于如上所述地假定前景仅平移,因此,准备了 HMM 状态转变概率表,其中前景视图模型的位置对应于背景视图模型中。在这种情况下,描述了状态转变概率表,例如,从视图变换 T_a 到视图变换 T_b 的转变概率以及从视图变换 T_a 到视图变换 T_c 的转变概率。也就是说,该表被准备为描述了当多个视图变换中的每一个 (例如,到右边的一次平移以及到

左边的一次平移等)为 HMM 的内部状态节点时、从在时间 s 处的节点到在时间 $s+1$ 处的节点的转变的概率。

[0322] 此时,在状态转变概率表,添加了如下约束:HMM 的内部状态节点排列成二维格状,并且内部状态节点之间的并非转变概率的二维格状中的从一个节点到相邻节点的转变概率的转变概率为 0。

[0323] 下文描述的视图变换估计通过在视图变换运动估计单元 115 中预先存储这样的状态转变概率表并与视图变换估计单元 114 合作地操作该状态转变概率表来进行。

[0324] 利用 HMM 进行的视图变换估计是通过计算要在给定时间的输出的视图变换的后验概率来进行的。要在给定时间的输出的视图变换的后验概率可用等式 (17) 计算。也就是说,该后验概率可基于与每一视图变换对应的每一状态的观察似然性 $P(X_s|T_k)$ 以及状态之间的转变概率 $P(T_k|T_{k-1})$ 来计算。另外,该后验概率可基于 HMM 中的前向状态概率 $\alpha(T_k)$ 和后向状态概率 $\beta(T_k)$ 来计算。

$$\begin{aligned}
 [0325] \quad P(T_k|X_s) &= \frac{P_{T_k} \cdot P(X_s|T_k)}{\sum_{k=1}^{N_X} P_{T_k} \cdot P(X_s|T_k)} \\
 [0326] \quad &= \frac{\alpha(T_k)\beta(T_k)}{\sum_{k=1}^{N_X} \alpha(T_k)\beta(T_k)} \quad (17)
 \end{aligned}$$

[0327] 前向状态概率可以用等式 (18) 根据前向算法(通常用于计算 HMM 状态概率)来计算。另外,后向状态概率可以用等式 (19) 根据后向算法(通常用于计算 HMM 状态概率)来计算。

[0328] 下列文献例如详细披露了上述前向算法和后向算法:Christopher M. Bishop 所著的"Pattern Recognition and Machine Learning"(P.335,(Information Science and Statistics),纽约,2006 年)(下面也称为文献 A)。

$$[0329] \quad \alpha(T_k) = P(X_s|T_k) \sum_{T_{k-1}} \alpha(T_{k-1}) P(T_k|T_{k-1}) \quad (18)$$

$$[0330] \quad \beta(T_k) = \sum_{T_{k+1}} \beta(T_{k+1}) P(X_s|T_{k+1}) P(T_k|T_{k+1}) \quad (19)$$

[0331] 也就是说,当图像特征提取单元 102 提取图像特征 X_s 时,例如,视图变换估计单元 114 估计多个视图变换的候选。用等式 (18) 和等式 (19)、参考 HMM 状态转变概率表来计算所估计的每一候选。这样,能够获得针对视图变换的每一候选的前向状态概率 $\alpha(T_k)$ 和后向状态概率 $\beta(T_k)$ 。

[0332] 视图变换运动估计单元 115 通过进行等式 (17) 的计算来计算针对视图变换的每一候选的后验概率。例如,当视图变换估计单元 114 输出所述候选时,视图变换估计单元

114 基于变换的后验概率、通过要输出的候选的数量来选择并输出视图变换。

[0333] 这样,利用所估计的视图变换的候选在前景视图模型上进行了视图变换。

[0334] 另外,视图模型学习如下进行。

[0335] 视图模型学习单元 111 获得相对于在每一时间的图像特征 X_s 在前景视图模型上进行的每一视图变换 $T_{FG,k}$ 的后验概率 $P(T_{FG,k}|X_s)$, 并通过等式 (20) 表示的计算来计算前景的前景视图模型 M_{FG} 的视图模型参数。

$$[0336] \quad M_{FG} = \frac{1}{S} \sum_{s=1}^S \sum_{k=1}^{N_T} P(T_{FG,k}|X_s) \cdot T_{FG,k}^{-1} C_{X, FGBG, s} X_s \quad (20)$$

[0337] 在等式 (20), $C_{X, FGBG, s}$ 表示与背景对应的图像特征是从在时间 s 的图像特征 X_s 中提取的。

[0338] 视图模型学习单元 121 获得相对于在每一时间的图像特征 X_s 在背景视图模型 M_{BG} 上进行的每一视图变换 $T_{BG,k}$ 的后验概率 $P(T_{BG,k}|X_s)$, 并通过等式 (21) 表示的计算来计算背景的背景视图模型 M_{BG} 的视图模型参数。

$$[0339] \quad M_{BG} = \frac{1}{S} \sum_{s=1}^S C_{X, FGBG, s} X_s \quad (21)$$

[0340] 图 20 的图示出第一具体示例中图像识别和视图模型学习中的视图变换估计的机制。

[0341] 如图 20 所示,在前景视图模型 M_{FG} 上进行视图变换 T_k , 并且作为估计的前景视图而得到的结果被合成到背景视图模型 M_{BG} 。这样,形成了合成视图模型 $M_{k, FGBG}$ 。

[0342] 图 20 的左上侧的有阴影的矩形图表示前景视图模型 M_{FG} 。图 20 的左下侧的白色矩形图表示背景视图模型 M_{BG} 。图 20 的右上侧的矩形图表示合成视图模型 $M_{k, FGBG}$ 。

[0343] 由于如上所述地假定前景仅平移,因此,视图变换 T_k 用等式 (15) 的仿射变换来表示。相应地,所估计的前景视图是当前前景视图模型 M_{FG} 移动到几个位置而没有变化时示出的图像。此时,矩形前景的左上顶点的位置可用从矩形背景的左边开始测量的距离 (或像素数) u 以及从矩形背景的上边开始测量的距离 (或像素数) v 来表示。

[0344] 在这种情况下,状态转变概率表为图 20 的右下侧示出的表。也就是说,上述 v 值用状态转变概率表的水平轴来规定,上述 u 值用状态转变概率表的垂直轴来规定。状态转变概率表的每一节点表示将前景的左上顶点的位置移动 (平移) 到位置 (u, v) 的视图变换。然后,描述节点之间的转变概率。

[0345] 在状态转变概率表,如上所述,添加了如下约束:HMM 的内部状态节点排列成二维格状,并且内部状态节点之间的并非转变概率的二维格状中的从一个节点到相邻节点的转变概率的转变概率为 0。因此,能够改善状态估计的效率。

[0346] 在第一具体示例,例如,由于在视图变换估计中使用了 HMM 的算法,因此,可简单获得针对视图变换的每一候选的前向状态概率 $\alpha(T_k)$ 和后向状态概率 $\beta(T_k)$ 。结果,由于可简单地计算等式 (17) 中的针对视图变换的每一候选的后验概率,因此,能够更有效地估计视图变换。另外,如同普通的 HMM 学习一样,基于前向状态概率和后向状态概率来更新 HMM 状态转变概率。

[0347] 图 21 的图示出第一具体示例中的图像识别的结果的一个示例。图 21 示出活动图像的另一示例,其中图 17 所示的前景图像正在图 18 所示的背景图像中移动。图 21 所示的图像是从作为活动图像输入的图像中提取的在三个时间处的图像。假设图 21 所示的前景的运动不同于图 19 所示的前景的运动。

[0348] 如图 21 所示,在时间 $s = s_{11}$ 、时间 $s = s_{12}$ 及时间 $s = s_{13}$ 处的图像中的每一个的帧线内识别出前景图像。

[0349] 第二具体示例

[0350] 在第二具体示例,使用图 4 的视图模型学习设备 200 的背景估计学习单元配置,其被配置为与图 3 的背景估计学习单元 105 的配置相同。假设视图模型学习设备 200 将前景学习为多视图模型,而将背景学习为单视图模型,然后利用作为学习结果的视图模型来识别前景图像。

[0351] 当在视图模型的学习之后识别图像时,与视图模型学习的情况一样,估计视图变换或者估计视图转变。然而,当识别了图像时,通常仅输出所述视图变换的一个候选或所述视图变换。

[0352] 在第二具体示例,在前景的视图转变估计中使用 HMM,在前景的视图变换估计中使用粒子滤波器,并且在背景的视图变换估计使用粒子滤波器。

[0353] 图 22 的图示出第二具体示例中用作输入图像序列的图像的示例。如图 22 所示,输入图像序列是例如通过将相机移动到围绕笔记本类型的个人计算机的不同位置并以各种角度拍摄所述笔记本类型的个人计算机而得到的图像。例如,在时间 $s = s_1$ 处到时间 $s = s_6$ 处的图像是通过将相机从左边移动到面对所述笔记本类型的个人计算机并拍摄活动图像而得到的。

[0354] 在第二具体示例的情况下,所述笔记本类型的个人计算机作为前景来学习。放置所述笔记本类型的个人计算机的桌子以及形成所述笔记本类型的个人计算机的背景的书籍等作为背景来学习。

[0355] 图 22 所示的输入图像序列包括通过以各种角度来观看作为前景的所述笔记本类型的个人计算机而得到的图像(视图)。另外,假定输入图像序列的前景由于相机的移动而相对于背景相对平移为被扩展或缩小。

[0356] 另一方面,假定输入图像序列的背景由于相机的移动而平移。

[0357] 在第二具体示例,使用通过将输入图像序列的 RGB 图案转换成灰度图案(0 到 1 的范围中的连续值)而得到的结果。

[0358] 前景多视图模型 $M_{FG,V}$ 的视图模型 $\{M_{FG,1}, M_{FG,2}, \dots, M_{FG,L}\}$ 的视图模型参数中的每一个包括预定大小的矩形区域的每一像素的坐标值(位置)和亮度值(像素值)。

[0359] 背景视图模型 M_{BG} 为面积大于在图 22 中的每一时间处的图像中显示的背景图像的矩形图像。每一像素的坐标值(位置)和亮度值(像素值)为视图模型参数。

[0360] 在第二具体示例的情况下,图 22 示出的输入图像序列被输入到图像序列输入单元 201。图像特征提取单元 202 提取通过将输入图像序列中的每一时间处的图像的 RGB 图案转换成灰度图案而得到的结果,作为图像特征。

[0361] 前景背景合成观察单元 203 合成从前景估计学习单元 204 输出的所估计的前景视图和从背景估计学习单元 105 输出的所估计的背景视图,以形成合成视图模型 M_{FGBG} 。

[0362] 合成视图模型 M_{FGBG} 通过表达式 (22) 来获得。

$$[0363] \quad \mathbf{M}_{\text{FGBG}} = \mathbf{T}_{\text{BG},k} \mathbf{M}_{\text{BG}} \oplus \mathbf{T}_{\text{FG},k} \mathbf{M}_{\text{FG},v} \quad (22)$$

[0364] 在等式 (22) 中运算符 $\oplus$ 表示后面的计算。也就是说, 获得经过视图变换 $T_{\text{FG},k}$ 的前景多视图模型 $M_{\text{FG},v}$ 的几何模型 $T_{\text{FG},k} M_{\text{FG},v,G}$ 以及背景视图模型 M_{BG} 的特征点 $M'_{\text{BG},G}$ (其为与几何模型 $T_{\text{FG},k} M_{\text{FG},v,G}$ 几何对应的特征点并经过视图变换 $T_{\text{BG},k}$)。在这种情况下, 获得背景中与前景的每一像素重叠的坐标位置, 作为特征点 $M'_{\text{BG},G}$ 。另外, 特征点 $M'_{\text{BG},G}$ 的特征量 $M'_{\text{BG},F}$ 用前景视图模型 $M_{\text{FG},v}$ 的特征量模型 $M_{\text{FG},v,F}$ 来替代。

[0365] 也就是说, 背景中与前景的每一像素重叠的坐标位置处的每一像素亮度值被替代为前景的每一像素的亮度值。

[0366] 在此由于如上所述假定前景被平移为被扩展或缩小, 等式 (22) 中的视图变换 $T_{\text{FG},k}$ 通过等式 (23) 的仿射变换来表示。

$$[0367] \quad T_{\text{FG},k} = \begin{pmatrix} s1_{\text{FG},k} & 0 & u_{\text{FG},k} \\ 0 & s2_{\text{FG},k} & v_{\text{FG},k} \\ 0 & 0 & 0 \end{pmatrix} \quad (23)$$

[0368] 另外, 如上所述, 由于在此假设背景仅平移, 在等式 (22) 中的视图变换 $T_{\text{BG},k}$ 用在等式 (24) 中的仿射变换来表示。

$$[0369] \quad T_{\text{BG},k} = \begin{pmatrix} 1 & 0 & u_{\text{BG},k} \\ 0 & 1 & v_{\text{BG},k} \\ 0 & 0 & 0 \end{pmatrix} \quad (24)$$

[0370] 前景背景合成观察单元 203 通过等式 (25) 和等式 (26) 计算在时间 s 处的合成视图模型 M_{FGBG} 的似然性 $P(X_s | M_{\text{FGBG}})$ 。另外, 基于等式 (26) 的最右侧 (并非针对其需要通过边缘化 (marginalization) 来计算似然性的目标的变量), 来计算与下述等式 (27) 的计算所需的视图变换关联的似然性以及等式 (30) 的计算所需的视图模型关联的似然性。

$$[0371] \quad P(X_s | T_{\text{BG},k}, T_{\text{FG},k}, M_{\text{FG},v}) = P(X_s | M_{\text{FGBG}}) \quad (25)$$

$$[0372] \quad P(X_s | T_{\text{BG},k}, T_{\text{FG},k}, M_{\text{FG},v}) = P(X_s | M_{\text{FGBG}}) = \frac{1}{D} \exp \left(- \frac{\sum_{i,j}^{C_{\text{FGBG},s}} (x_{s,i} - (m_{\text{FGBG},j}))^2}{2\sigma^2} \right) \quad (26)$$

[0373] 在等式 (26), σ 和 D 是预先确定的参数。在等式 (26), $C_{\text{FGBG},s}$ 表示输入图像序列中在时间 s 处的图像与合成视图模型 M_{FGBG} 之间的对应关系。通过等式 (26) 得到的似然性被用作估计值。

[0374] 接下来, 描述前景的视图变换估计和视图转变估计。

[0375] 视图和变换估计单元 215 和视图变换运动估计单元 216 如下所述地进行视图变换估计。

[0376] 准备了粒子滤波器,其中,基于平移、扩展和缩小的假定的视图变换 $T_{FG,k}$ 的 4 个参数被用作状态量。在此,这 4 个参数对应于上述等式 (23) 中的右侧的矩阵中的位于第一行和第一列的分量、位于第一行和第三列的分量、位于第二行和第二列的分量以及位于第二行和第三列的分量。由于指定前景视图变换的 4 个参数的组合可对应于粒子滤波器的各个粒子,因此通过指定一个粒子来指定一个视图变换。

[0377] 另外,例如,粒子滤波器可较好地用于图像识别中的目标跟踪。文献 A 的第 364 页等详细披露了粒子滤波器。

[0378] 使用粒子滤波器执行的视图变换估计是通过计算在给定时间的视图变换的后验概率并计算在随后的时间的视图变换的后验概率来进行的。

[0379] 在给定时间 (时间 s) 的视图变换的后验概率可用此时采样的粒子的后验概率 $w_s^{(1)}$ 来近似。在时间 s 处采样的粒子的后验概率可通过等式 (27) 基于与在时间 s 处的视图变换关联的似然性来计算。另外,如上所述,视图和变换估计单元 215 基于前景背景合成观察单元 203 计算的估计值来计算与视图变换关联的似然性。

$$[0380] \quad P(T_{s,k}^{(l)} | X_s) \approx w_s^{(l)} = \frac{P(X_s | T_{s,k}^{(l)})}{\sum_{m=1}^M P(X_s | T_{s,k}^{(m)})} \quad (27)$$

[0381] 在等式 (27) 中,与在时间 s 处的视图变换关联的似然性被表示为 $P(X_s | T_{s,k}^{(1)})$ 或者 $P(X_s | T_{s,k}^{(m)})$ 。在等式 (27), (1) 和 (m) 表示粒子的索引,并且总共有 M 个粒子。如上所述,由于视图变换可对应于每一粒子,因此,可将粒子的索引赋予视图变换。

[0382] 用于采样在随后的时间 (时间 $s+1$) 的视图变换的概率基于在时间 s 处采样的粒子的后验概率 $w_s^{(1)}$ 来计算,并且用等式 (27) 以及等式 (28) 的状态更新规则 $F(T_{s+1,k}^{(1)} | T_{s,k}^{(1)})$ 来计算。

$$[0383] \quad P(T_{s+1,k} | X_s) \approx \sum_{l=1}^M w_s^{(l)} F(T_{s+1,k} | T_{s,k}^{(l)}) \quad (28)$$

[0384] 在此,例如,状态更新规则是针对等式 (29) 中在时间 s 处的视图变换 $T_{s,k}$ 的状态量 x_s 给出的。

$$[0385] \quad x_{s+1} = x_s + ND(0, \sigma) \quad (29)$$

[0386] 该等式表示状态量 x_s 邻近在时间 $s+1$ 处的状态量 x_s 的方差 σ 的正态分布。例如,当要学习和识别的前景的运动为轻柔时,状态更新规则是有效的。另一方面,当要学习和识别的前景的运动偏离假定时,需要考虑要学习和识别的前景的运动来设计状态更新规则,或者根据样本学习状态更新模型。

[0387] 视图变换运动估计单元 216 针对每种视图变换 (4 个参数的组合) 计算在时间 $s+1$ 处的视图变换的后验概率。例如,当视图和变换估计单元 215 输出视图变换的候选时,基于在时间 $s+1$ 的处视图变换的后验概率通过要输出的多个候选来选择和输出视图变换。

[0388] 这样,利用所估计的视图变换的候选在前景视图模型上进行了视图变换。

[0389] 在此,已经描述了粒子滤波器被分别应用于前景和背景的情况。然而,为了生成合成视图模型,粒子是从用于前景的粒子和用于背景的粒子中随机(例如,基于均匀分布概率)选择的。然后,基于所选粒子的视图变换参数的状态量来形成前景和背景的合成视图模型。

[0390] 视图和变换估计单元 215 和视图转变估计单元 217 如下所述地进行视图转变估计。

[0391] 准备了 HMM 状态转变概率表,其中前景多视图模型中的多个视图模型的每一个是 HMM 的内部状态节点。在这种情况下,状态转变概率表描述例如从视图模型 Ma 到视图模型 Mb 的转变概率、从视图模型 Ma 到视图模型 Mc 的转变概率。

[0392] 此时,在状态转变概率表,添加了如下约束:HMM 的内部状态节点排列成二维格状,并且内部状态节点之间的并非转变概率的二维格状中的从一个节点到相邻节点的转变概率的转变概率为 0。这是因为,可假定前景被假定为具有三维结构并且因此从三维结构中进行的多个视图之间的转变被假定为在覆盖三维结构的二维结构上进行。更准确地,可使用在球形表面上排列节点的方法。

[0393] 例如,如下文所述的,视图转变估计是通过将状态转变概率表存储于视图转变估计单元 217 中并与视图和变换估计单元 215 合作地来进行的操作来进行的。

[0394] 利用 HMM 的视图估计是通过计算在给定时间的视图模型的后验概率来进行的。在给定时间的视图模型的后验概率可基于在每一状态的观察似然性 $P(X_s | M_{FG,v})$ 以及状态间的转变概率 $P(M_{FG,v} | M_{FG,v'})$ 来计算的。另外,在给定时间的视图模型的后验概率可通过等式 (30) 利用 HMM 中的前向状态概率 $\alpha(M_{FG,v})$ 和后向状态概率 $\beta(M_{FG,v})$ 来计算。

$$\begin{aligned}
 [0395] \quad P(M_{FG,v} | X_s) &= \frac{P_{M_{FG,v}} \cdot P(X_s | M_{FG,v})}{\sum_{1}^L P_{M_{FG,v}} \cdot P(X_s | M_{FG,v})} \\
 [0396] \quad &= \frac{\alpha(M_{FG,v}) \beta(M_{FG,v})}{\sum_{1}^L \alpha(M_{FG,v}) \beta(M_{FG,v})} \quad (30)
 \end{aligned}$$

[0397] 前向状态概率可通过根据通常用于利用等式 (31) 计算 HMM 状态概率的前向算法来计算。后向状态概率可通过根据通常用于利用等式 (32) 计算 HMM 状态概率的后向算法来计算。

$$[0398] \quad \alpha(M_{FG,v}) = P(X_s | M_{FG,v}) \sum \alpha(M_{FG,v'}) P(M_{FG,v} | M_{FG,v'}) \quad (31)$$

$$[0399] \quad \beta(M_{FG,v}) = \sum \beta(M_{FG,v''}) P(X_s | M_{FG,v''}) P(M_{FG,v} | M_{FG,v''}) \quad (32)$$

[0400] 也就是说,当图像特征提取单元 202 提取在时间 s 处的图像特征 X_s 时,例如,视图和变换估计单元 215 估计多个视图转变的候选。利用等式 (31) 和等式 (32)、参考 HMM 状态转变概率表来计算所估计的每一候选。这样,可得到针对视图变换的每一候选的前向状态概率 $\alpha(M_{FG,v})$ 和后向状态概率 $\beta(M_{FG,v})$ 。

[0401] 这样,视图转变估计单元 217 针对每一视图模型来计算要在随后的时间转变的视

图变换的后验概率。例如,当视图和变换估计单元 215 输出视图转变的候选时,基于要在随后的时间转变的视图模型的后验概率通过要输出的多个候选来选择和输出视图模型。

[0402] 在针对其估计并输出视图转变的前景视图模型上进行通过以所述方式估计的视图变换的候选的视图变换。

[0403] 接下来,描述背景视图变换。

[0404] 视图变换估计单元 124 和视图变换运动估计单元 125 如下所述地进行视图变换估计。

[0405] 准备粒子滤波器,其中,假定了平移的视图变换 $T_{BG, k}$ 的两个参数被用作状态量。在此,两个参数对应于上述等式 (24) 的右侧的矩阵中的位于第一行和第三列的分量以及位于第二行和第三列的分量。

[0406] 与前景视图变换一样,计算要在随后的时间采样的粒子的后验概率,并基于该后验概率来输出每一视图变换作为候选。

[0407] 这样,利用所估计的视图变换的候选在背景视图模型进行了视图变换。

[0408] 在此,已经描述了粒子滤波器被分别应用于前景和背景的情况。然而,为了生成合成视图模型,粒子是从用于前景的粒子和用于背景的粒子中随机(例如,基于均匀分布概率)选择的。然后,基于所选粒子的视图变换参数的状态量来形成前景和背景的合成视图模型。

[0409] 接下来,将描述视图模型学习。

[0410] 多视图模型视图转变学习单元 211 如下所述地计算前景多视图模型的每一视图模型的视图模型参数。

[0411] 也就是说,多视图模型视图转变学习单元 211 基于以上述方式利用等式 (33) 得到的视图变换的后验概率和视图模型的后验概率来计算多视图模型的视图模型 $M_{FG, v}$ 的视图模型参数。

$$[0412] \quad M_{FG, v} = \frac{1}{S} \sum_{s=1}^S \sum_{k=1}^{N_T} P(M_{FG, v} | X_s) P(T_{FG, k} | X_s) \cdot T_{FG, k}^{-1} C_{X, FGBG, s} X_s \quad (33)$$

[0413] 在等式 (33),在前景视图模型上进行的视图变换表示为 $T_{FG, k}$,且存在总共 N_T 个视图变换。例如,通过等式 (33) 计算多视图模型的 L 个视图模型中的一个视图模型的视图模型参数。相应地,等式 (13) 的计算通过多视图模型的多个视图模型来进行。

[0414] 另外,视图模型学习单元 121 如下所述地计算背景视图模型的视图模型参数。

[0415] 也就是说,视图模型学习单元 121 利用等式 (34) 基于以上述方式计算的视图变换的后验概率来计算视图模型 M_{BG} 的视图模型参数。

$$[0416] \quad M_{BG} = \frac{1}{S} \sum_{s=1}^S \sum_{k=1}^{N_T} P(T_{BG, k} | X_s) \cdot T_{BG, k}^{-1} C_{X, FGBG, s} X_s \quad (34)$$

[0417] 在等式 (34),在背景视图模型上进行的视图变换被表示为 $T_{BG, k}$ 且存在总共 N_T 个视图变换。

[0418] 图 23 的图示出第二具体示例中的视图模型学习。

[0419] 在第二具体示例中,如图 23 所示,在前景多视图模型 $M_{FG, v}$ 的预定视图模型上进行

视图变换 $T_{FG,k}$ 。该视图模型被视为所估计的前景视图。另外,在背景视图模型 M_{BG} 上进行视图变换 $T_{BG,k}$ 。该视图模型被视为所估计的背景视图。

[0420] 在该示例中,在图的左上侧概念性地示出了前景多视图模型。也就是说,在这种情况下,前景多视图模型包括 9 个视图模型。在图的左上部,作为前景多视图模型 $M_{FG,V}$ 的视图模型的这 9 个视图模型被示出为矩形图。

[0421] 在该示例中,背景视图模型 M_{BG} 被配置为面积大于合成视图模型 $M_{FGBG,k}$ 中所显示的背景图像的面积的矩形图像。也就是说,在图的左下部,背景视图模型 M_{BG} 的整体尺寸(面积)被示出为较大的矩形,并且作为所估计的背景视图输出的部分的尺寸(面积)被示出为较小的矩形。

[0422] 然后,合成所估计的前景视图和所估计的背景视图,以形成合成视图模型 $M_{FGBG,k}$ 。也就是说,在图的右部,白色矩形和由阴影的矩形以重叠的方式示出,作为合成视图模型 $M_{FGBG,k}$ 。白色矩形为通过在从示出于图的左下部的背景视图模型 M_{BG} 中提取的图像上进行视图变换 $T_{BG,k}$ 而得到的所估计的背景视图。另外,有阴影的矩形为通过在示出于图的左上部的前景多视图模型 M_{FG} 中的所选视图模型上进行视图变换 $T_{FG,k}$ 而得到的所估计的前景视图。

[0423] 在第二具体示例中,粒子滤波器用于背景和前景的视图变换估计。因此,例如,即使有大量的视图变换的参数,也能有序地进行视图变换估计。另外,例如,可简单地计算在等式(28)中示出的视图变换的后验概率。

[0424] 在第二具体示例中,HMM 算法被用于前景视图转变的估计。因此,例如,可简单地获得针对视图变换的每一候选的前向状态概率 $\alpha(T_k)$ 和后向状态概率 $\beta(T_k)$ 。结果,由于可简单地计算等式(30)中所示的针对视图转变的每一候选的后验概率,因此,能更有效地估计视图变换。

[0425] 在状态转变概率表,如上所述,添加了如下约束:HMM 的内部状态节点排列成二维格状,并且内部状态节点之间的并非转变概率的二维格状中的从一个节点到相邻节点的转变概率的转变概率为 0。因此,可提高状态估计的效率。

[0426] 图 24 的图示出了第二具体示例中通过输入图 22 中示出的输入图像序列、学习前景多视图模型和背景视图模型并识别图像而得到的结果。示出于图 24 的右上部的图像时作为要学习的输入而输入的图像。

[0427] 如图 24 的左上部所示,对应于通过以不同的角度来观看笔记本类型的个人计算机而得到的图像的 9 个视图模型被学习作为前景多视图模型。另外,如图 24 的左下部所示,放置所述笔记本类型的个人计算机的桌子和形成所述笔记本类型的个人计算机的背景的书籍等被学习作为背景视图模型。

[0428] 与实际图像不同的是,在图 24 所示的视图模型中,模糊地显示了背景和前景的对象。如上所述,这是因为,基于视图转变或者视图变换的后验概率以加权的方式来更新作为视图模型的参数的每一像素的亮度值。

[0429] 如图 24 的右下部所示,形成了合成视图模型。也就是说,合成视图模型是通过以重叠的方式显示通过在背景的预定视图模型上进行预定视图变换并在背景视图模型进行预定视图变换而得到的所估计的前景视图来形成的。

[0430] 在图 24,可以理解,形成了与输入图像几乎相同的合成视图模型,因而可适当地学

习和识别图像。

[0431] 第三具体示例

[0432] 在第三具体示例中,图 3 的视图模型学习设备 100 将前景学习为单视图模型,并且也将背景学习为单视图模型,然后利用作为学习结果的视图模型来识别前景图像。

[0433] 当在视图模型的学习之后识别图像时,如同在视图模型学习中一样,也估计视图变换。然而,当识别了图像时,通常仅输出视图变换的一个候选或者所述视图变换。

[0434] 在第三具体示例中,局部特征(哈里斯角点)被用作视图模型,而粒子滤波器被用于视图变换估计。

[0435] 图 25 的图示出第三具体示例中作为输入图像序列的图像的示例。如图 25 所示,输入图像序列是通过手动固定桌上的玩偶(填充玩具)的姿势、移动玩偶的正面、后面、右面和左面并用相机拍摄该玩偶来得到的。例如,时间 $s = s_1$ 到时间 $s = s_3$ 的图像时通过逐渐移动玩偶并拍摄活动图像而得到的。

[0436] 在第三具体示例中,将玩偶作为前景来学习。另外,放置玩偶的桌子以及形成玩偶的背景的咖啡杯、桌面等作为背景来学习。

[0437] 在第三具体示例中,假定图 25 中所示的输入图像序列中作为前景的玩偶相对于背景被相对平移,以被扩展或缩小。

[0438] 另一方面,假定输入图像序列中的背景由于相机的移动而被平移。

[0439] 在第三具体示例中,输入图像序列的每一图像的 RGB 图案被转换成灰度图案(0 到 1 的范围中的连续值),并且通过哈里斯角点检测来检测特征点的位置。检测到的 f 个特征点的位置的集合被用作图像特征。在图 25 的时间 $s = s_1$ 到时间 $s = s_3$ 的图像中,以交叉形状等表示的点被绘在玩偶、桌子、咖啡杯以及桌面中。所绘的点为通过哈里斯角点检测而检测到的特征点的部分。

[0440] 在此,已经描述了通过哈里斯角点检测来检测特征点的示例。然而,还可使用其他方法来检测特征点。

[0441] 相应地,在第三具体示例中,无需在图像特种包括像素的亮度值等。图像特征仅包括每一特征点的坐标位置等。例如,在每一时间的图像特征 $X_1, X_2, \dots, X_s$ 中的图像特征 X_1 包括有关从时间 1 处的图像中检测到的特征点的坐标位置的集合 (x_1, y_9) 、 (x_{18}, y_{31}) 的信息。另外,图像特征 X_1 包括有关从时间 1 处的图像中检测到的特征点的坐标位置的集合 (x_3, y_6) 、 (x_{11}, y_{38}) 等的信息。因此,在第三具体示例中,无需包括特征量模型,这是因为视图模型参数仅包括几何模型就足够了。

[0442] 在哈里斯角点检测中,图像的边缘通常被检测作为特征点。例如,作为前景的玩偶的图像中的边缘的点以及作为背景的咖啡杯和桌面的图像中的边缘的点被检测作为特征点。

[0443] 在第三具体示例中,包括玩偶的图像中的第一特征点、第二边缘点等的特征点的集合以及二维空间中各个特征点之间的相对位置被学习作为前景视图模型。在第三具体示例中,包括咖啡杯和桌面的图像中的第一特征点、第二边缘点等的特征点的集合以及二维空间中各个特征点之间的相对位置被学习作为背景视图模型。

[0444] 与几何模型(坐标位置等)相似,诸如方向可控滤波器等的特征量模型可被添加作为有关特征点的信息。

[0445] 在第三具体示例中,图 25 所示的输入图像序列被输入到图像序列输入单元 101。图像特征提取单元 102 通过哈里斯角点检测从输入图像序列中在每一时间的图像中检测特征点。

[0446] 前景背景合成观察单元 103 合成由前景估计学习单元 104 估计的所估计的前景视图以及由背景估计学习单元 105 估计的所估计的背景视图,以形成合成视图模型 M_{FGBG} 。

[0447] 所估计的前景视图可通过在前景视图模型 M_{FG} 上进行视图变换 $T_{FG,k}$ 而得到,并被表示为 $T_{FG,k}M_{FG}$ 。所估计的背景视图可通过在背景视图模型 M_{BG} 上进行视图变换 $T_{BG,k}$ 来得到,并被表示为 $T_{BG,k}M_{BG}$ 。

[0448] 合成视图模型 M_{FGBG} 通过等式 (35) 而得到。

$$[0449] \quad \mathbf{M}_{FGBG} = \mathbf{T}_{BG,k} \mathbf{M}_{BG} \oplus \mathbf{T}_{FG,k} \mathbf{M}_{FG} \quad (35)$$

[0450] 在等式 (35) 中的操作符 $\oplus$ 表示两个特征点集合的相加计算。也就是说,图像特征和合成视图模型二者包括有关特征点的位置的信息。相应地,合成视图模型 M_{FGBG} 是通过将由所估计的前景视图 $T_{FG,k}M_{FG}$ 指示的特征点的集合与由所估计的背景视图 $T_{BG,k}M_{BG}$ 指示的特征点的集合相加而得到的。

[0451] 在此,如上所述,由于假定前景被平移或被扩展或缩小,因此,等式 (35) 的视图变换 $T_{FG,k}$ 用等式 (36) 的仿射变换来表示。

$$[0452] \quad T_{FG,k} = \begin{pmatrix} s1_{FG,k} & 0 & u_{FG,k} \\ 0 & s2_{FG,k} & v_{FG,k} \\ 0 & 0 & 0 \end{pmatrix} \quad (36)$$

[0453] 另外,如上所述,由于在此假定背景仅被平移,因此,等式 (35) 的视图变换 $T_{BG,k}$ 用等式 (37) 的仿射变换来表示。

$$[0454] \quad T_{BG,k} = \begin{pmatrix} 1 & 0 & u_{BG,k} \\ 0 & 1 & v_{BG,k} \\ 0 & 0 & 0 \end{pmatrix} \quad (37)$$

[0455] 前景背景合成观察单元 103 通过等式 (38) 和等式 (39) 来计算在时间 s 处的合成视图模型 M_{FGBG} 的似然性 $P(X_s | M_{FGBG})$ 。

$$[0456] \quad P(X_s | T_{BG,k}, T_{FG,k}) = P(X_s | M_{FGBG}) \quad (38)$$

$$[0457] \quad P(X_s | T_{BG,k}, T_{FG,k}) = P(X_s | M_{FGBG}) = \frac{1}{D} \exp \left(- \frac{\sum_{i,j}^{C_{FGBG,s}} (x_{s,i} - (m_{FGBG,j}))^2}{2\sigma^2} \right) \quad (39)$$

[0458] 在等式 (39), σ 和 D 是预先确定的参数。在等式 (39), $C_{FGBG,s}$ 表示输入图像序列中在时间 s 处的图像的特征点与合成视图模型 M_{FGBG} 的特征点之间的对应关系。通过等式 (39) 得到的似然性被用作估计值。

[0459] 此时,例如,以下面的方式来确定输入图像序列中在时间 s 处的图像的特征点与合成视图模型 M_{FGBG} 的特征点之间的对应关系。也就是说,通过组合优化算法的拍卖算法 (auction algorithm) 来选择特征点对,使得彼此对应的特征点对之间的距离作为一个整体被缩短。

[0460] 下列文献详细披露了拍卖算法:Kubo Mikio 和 Matsui Tosiki 所著的 "Combination Optimization [collection of short articles], Chapter 3 Assignment Problem" (ISBN 978-4-254-12617-4C3341, Asakura 出版社,1999 年 1 月 10 日)。

[0461] 如上所述,基于彼此对应的特征点对之间的距离的似然性用 b_y 等式 (38) 和等式 (39) 来计算。

[0462] 接下来,描述前景和背景视图变换。

[0463] 关于前景视图变换,视图变换估计单元 114 和视图变换运动估计单元 115 执行如下的视图变换估计。

[0464] 准备粒子滤波器,其中基于平移、扩展和缩小的假定的视图变换 $T_{\text{FG},k}$ 的 4 个参数被用作特征量。在此,这 4 个参数对应于上述等式 (36) 的右边的矩阵中的位于第一行和第一列的分量、位于第一行和第三列的分量、位于第二行和第二列的分量以及位于第二行和第三列的分量。由于指定前景视图变换的 4 个参数的组合可对应于粒子滤波器的各个粒子,因此,通过指定一个粒子来指定一个视图变换。

[0465] 如在第二具体示例中描述的,利用粒子滤波器进行的视图变换估计是通过计算在给定时间的视图变换的后验概率并计算在随后的时间的视图变换的后验概率来进行的。

[0466] 在给定时间 (时间 s) 的视图变换的后验概率可用此时采样的粒子的后验概率 $w_s^{(1)}$ 来表示。在时间 s 处采样的粒子的后验概率可基于与在时间 s 处的视图变换关联的似然性来计算。另外,如上所述,视图变换估计单元 114 基于由前景背景合成观察单元 103 计算的估计值来计算与视图变换关联的似然性。

[0467] 用于采样在随后的时间 (时间 $s+1$) 的视图变换的概率是基于在时间 s 处采样的粒子的后验概率 $w_s^{(1)}$ 以及状态更新规则 $F(T_{s+1}^{(1)}, k | T_{s,k})$ 来计算的。

[0468] 视图变换运动估计单元 115 针对每种视图变换 (4 个参数的组合) 来计算在时间 $s+1$ 处的视图变换的后验概率。例如,当视图变换估计单元 114 输出视图变换的候选时,基于在时间 $s+1$ 处的视图变换的后验概率通过要输出的多个候选来选择并输出视图变换。

[0469] 这样,利用所估计的视图变换的候选在前景视图模型上进行了视图变换。

[0470] 关于背景视图变换,视图变换估计单元 124 和视图转变估计单元 125 执行如下的视图转变估计。

[0471] 准备了粒子滤波器,其中假定平移的视图变换 $T_{\text{BG},k}$ 两个参数被用作状态量。在此,这两个参数对应于上述等式 (37) 右边的矩阵中的位于第一行和第三列的分量以及位于第二行和第三列的分量。

[0472] 如同前景视图变换中一样,计算在随后的时间采样的粒子的后验概率,并基于该后验概率输出每一视图变换作为候选。

[0473] 这样,利用所估计的视图变换的候选在前景视图模型上进行了视图变换。

[0474] 接下来,将描述视图模型学习。

[0475] 视图模型视图转变学习单元 111 如下所述的计算前景视图模型的视图模型参数。

[0476] 也就是说,视图模型学习单元 111 通过等式 (40) 基于以上述方式得到的视图变换的后验概率来计算视图模型 M_{FG} 的视图模型参数。

$$[0477] \quad M_{FG} = \frac{1}{S} \sum_{s=1}^S \sum_{k=1}^{N_T} P(T_{FG,k} | X_s) \cdot T_{FG,k}^{-1} C_{X, FGBG, s} X_s \quad (40)$$

[0478] 在等式 (40), 在前景视图模型上进行的视图变换被表示为 $T_{FG, k}$, 并且存在总共 N_T 个视图变换。

[0479] 另外, 视图模型学习单元 121 如下所述地计算背景视图模型的视图模型参数。

[0480] 也就是说, 视图模型学习单元 121 基于以上述方式计算的视图变换的后验概率用等式 (41) 来计算视图模型 M_{BG} 的视图模型参数。

$$[0481] \quad M_{BG} = \frac{1}{S} \sum_{s=1}^S \sum_{k=1}^{N_T} P(T_{BG,k} | X_s) \cdot T_{BG,k}^{-1} C_{X, FGBG, s} X_s \quad (41)$$

[0482] 在等式 (41), 在背景视图模型上进行的视图变换被表示为 $T_{BG, k}$, 并且存在总共 N_T 个视图变换。

[0483] 图 26 到 28 的图示出在第三具体示例中的视图模型学习。

[0484] 在第三具体示例中, 如图 26 所示, 在前景视图模型 M_{FG} 上进行视图变换 $T_{FG, k}$ 。该视图模型被视为所估计的前景视图。另外, 在背景视图模型 M_{BG} 上进行视图变换 $T_{BG, k}$ 。该视图模型被视为所估计的背景视图。

[0485] 在该示例中, 图中的左上部示出的矩形内示出了多个点, 作为前景视图模型 M_{FG} 。各个点为形成前景视图模型的特征点。

[0486] 在该示例中, 图中的左下部示出的矩形内示出了多个点, 作为背景视图模型 M_{BG} 。各个点为形成背景视图模型的特征点。

[0487] 然后, 所估计的前景视图和所估计的背景视图被合成, 以形成合成视图模型 $M_{FGBG, k}$ 。如上所述, 合成视图模型是通过将所估计的前景视图中的特征点的集合与所估计的背景视图中的特征点的集合相加而得到的。在该示例中, 在图中的中右部所示的矩形内示出多个点, 作为合成视图模型 $M_{FGBG, k}$ 。这些点是形成合成视图模型的特征点。

[0488] 不同于第一和第二具体示例, 在第三具体示例中, 局部特征量被用于图像特征或视图模型参数。相应地, 与第一和第二具体示例相比, 即使针对环境中诸如照明条件等的变化, 也能够进行鲁棒的识别。结果, 例如, 能够实现能以高速和较低的开销来学习和识别图像的设备。

[0489] 图 27 和 28 的图示出第三具体示例中通过输入图 25 所示的输入图像序列、学习前景视图模型和背景视图模型以及识别图像而得到的结果。

[0490] 在图 27 和 28, 被识别为前景和背景的特征点的点被显示为图中交叉形状或图中圆形的绘制点。在图 27 和 28, 显示在图中左侧 (前景估计) 的图像中的用交叉形状等表示的绘制点为作为前景的特征点来学习和识别 (估计) 的点。在图 27 和 28, 显示在图中右侧 (背景估计) 的图像中的用圆形等表示的绘制点为作为背景的特征点来学习和识别 (估计) 的点。

[0491] 图 27 示出在学习的早期阶段中通过识别图像而得到的结果, 并且图 28 示出在充

分学习后通过识别图像而得到的结果。

[0492] 如图 27 所示,在学习的早期阶段中,例如,在所估计的前景图像中背景的一部分被绘制为前景的特征点,因此,背景的该部分被错误地识别为前景。也就是说,应理解,在学习的早期阶段中不能适当地识别背景和前景。

[0493] 然而,如图 28 所示,应理解,在充分学习之后几乎适当地识别了背景和前景。也就是说,根据图 28 所示的识别结果可以理解,获得了合成视图模型,其中,可以几乎正确地绘制出输入图像中的前景的特征量和背景的特征点。

[0494] 至此已经描述了应用了根据本技术的第一、第二和第三具体示例的图像处理设备中的具体视图模型学习和图像识别的示例。

[0495] 当然,实质上可以实施除了上述示例之外的具体示例。例如,可实现用多视图模型来学习前景并用多视图模型来学习背景的具体示例。例如,可实现假定前景的运动为平移、扩展及缩小且假定背景的运动为平移、扩展及缩小的具体示例。

[0496] 或者,可假定转动或投影等作为前景和背景的运动。

[0497] 这样,可以用各种形式来修改根据本技术的实施例的视图模型学习以及利用该学习结果进行的图像识别。可参考上述具体示例来实现其他具体示例。相应地,所述具体示例可以用相同的方式来披露。

[0498] 另外,至此已经描述了一个目标(例如笔记本类型的个人计算机)作为前景来学习和识别并且一个目标(例如桌子或书籍)作为背景来学习和识别的示例。然而,例如,可将多个目标作为前景来学习和识别并可将多个目标作为背景来学习和识别。

[0499] 至此已经描述了将本技术的实施例应用到图像识别的示例。然而,例如,本技术的实施例可应用于传感器信息,而不是图像。例如,本技术的实施例可应用到诸如利用压力传感器的触觉传感器等的传感器信息。

[0500] 上述处理序列可以用硬件或软件来执行。当上述处理序列用软件来执行时,该软件的程序被安装到嵌入专用硬件的计算机或者能够通过从网络或记录介质中安装各种程序来执行各种功能的通用个人计算机 700(如图 29 所示)中。

[0501] 在图 29,CPU(中央处理单元)701 根据存储于 ROM(只读存储器)702 中的程序或者从存储单元 708 加载到 RAM(随机存取存储器)703 上的程序来执行各种处理。在 RAM 703 中不仅适当地存储有 CPU 701 执行的各种程序,还存储有所需的数据等。

[0502] CPU 701、ROM 702 以及 RAM 703 通过总线 704 彼此互连。总线 704 还连接到输入/输出接口 705。

[0503] 输入单元 706(配置为键盘或鼠标等)、输出单元 707(配置为扬声器或显示器等)、存储单元 708(配置为硬盘等)以及通信单元 709(配置为网络接口卡,如调制解调器或 LAN 卡)连接到输入/输出接口 705。通信单元 709 通过包括互联网的网络进行通信。

[0504] 驱动器 710 根据需要连接到输入/输出接口 705。以适当的方式来安装可拆卸介质 711(如磁盘、光盘、磁光盘或者半导体存储器等),使得从所安装的可拆卸介质 711 中读出的计算机程序可根据需要而被安装于存储单元 708 中。

[0505] 当上述处理序列通过软件来执行时,从诸如互联网等网络或者诸如可拆卸介质 711 等记录介质中安装该软件的程序。

[0506] 记录介质包括可拆卸介质 711,如磁盘(包括软盘(注册商标))、光盘(包括

CD-ROM(致密盘-只读存储器)和DVD(数字通用盘)、磁光盘(包括MD(微小盘)(注册商标))或者半导体存储器等,该介质存储程序并被分发以将程序给予远离图29所示的设备主体的用户。记录介质还包括在ROM702或存储单元708中包括的硬盘,其存储程序并以嵌入在设备主体的方式而被递送到用户。

[0507] 在本说明书中,上述处理序列不仅包括以上述的顺序按照时间顺序进行的处理,还包括平行或以独立的方式执行的程序(不必按照时间顺序来执行)。

[0508] 本技术的实施例不限于上述实施例,而是可以在本技术的范围内用各种形式来修改,而不脱离本技术的核心。

[0509] 本技术的实施例可以如下所述地实施。

[0510] (1) 一种图像处理设备包括:图像特征输出单元,用于与帧时间对应地输出图像特征中的每一个,所述图像特征被形成作为输入活动图像数据中的每一帧的图像的多个特征点中的特征;前景估计单元,用于通过在前景视图模型上进行作为几何变换的视图变换来估计时间 s 处的前景图像,并且输出所估计的前景视图,其中,所述前景视图模型具有所述图像中的前景的图像特征,作为与所述时间 s 处的图像特征有关的参数;背景估计单元,用于通过在背景视图模型上进行作为几何变换的视图变换来估计所述时间 s 处的背景图像,并输出所估计的背景视图,所述背景视图模型具有所述图像中的背景的图像特征,作为与所述时间 s 处的图像特征有关的参数;合成视图生成单元,用于通过合成所估计的前景视图和所估计的背景视图来生成合成视图;前景学习单元,用于基于通过所述合成视图和所述时间 s 处的图像特征之间进行比较而得到的估计值并通过基于随机生成模型来更新所述前景视图模型的参数,来学习所述前景视图模型;以及背景学习单元,用于基于所述估计值、通过基于随机生成模型来更新所述背景视图模型的参数,来学习所述背景视图模型。

[0511] (2) 在(1)中所述的图像处理设备中,所述前景估计单元和所述背景估计单元中的每一个基于所述估计值来计算多个视图变换中的每一个的后验概率,基于所述图像特征和所述后验概率来估计该视图变换,以及通过在所述视图模型上进行所述视图变换来分别输出所估计的前景视图和所估计的背景视图。

[0512] (3) 在(1)或(2)中所述的图像处理设备中,所述前景学习单元和所述背景学习单元中的每一个基于根据所述估计值来计算的每一视图变换的后验概率,对经过多次视图变换的视图模型的参数进行加权,基于加权的参数来分别更新所述前景视图模型的参数和所述背景视图模型的参数。

[0513] (4) 在(1)到(3)中的任一项所述的图像处理设备中,所述视图模型被配置为通过多个视图模型而形成的多视图模型,所述多个视图模型与通过以不同的角度来观看所述前景和所述背景中的每一个而得到的图像对应。

[0514] (5) 在(1)到(4)中的任一项所述的图像处理设备中,所述前景视图模型和所述背景视图模型之一被配置为通过多个视图模型而形成的多视图模型,所述多个视图模型与通过以不同的角度来观看所述前景和所述背景中的每一个而得到的图像对应。

[0515] (6) 在(4)或(5)中所述的图像处理设备中,所述前景估计单元和所述背景估计单元中的每一个基于所述估计值来计算所述多个视图模型中的每一个的后验概率,基于所述图像特征和所述后验概率来估计所述视图变换,以及通过在所估计的与转变对应的视图模

型上进行所述视图变换,来输出所估计的前景视图和所估计的背景视图。在计算所述视图模型的后验概率时,使用 HMM、FNN、RNN、粒子滤波器和卡尔曼滤波器中的一个作为基于视图转变来估计先验概率的动态学习估计模型。

[0516] (7) 在(6)中所述的图像处理设备中,所述前景估计单元和所述背景估计单元中的每一个基于所述估计值来计算多个视图变换中的每一个的后验概率,以及基于所述图像特征和所述后验概率来估计该视图变换。

[0517] (8) 在(4)或(7)中所述的图像处理设备中,所述和前景学习单元和所述背景学习单元中的每一个基于根据所述估计值而计算的每一转变的后验概率,对与多个转变对应的视图模型的参数进行加权,基于根据所述估计值而计算的、每一视图变换的后验概率,对经过多个视图变换的视图模型的参数进行加权,以及基于经加权的参数来分别更新所述前景视图模型的参数和所述背景视图模型的参数。

[0518] (9) 在(1)到(8)中的任一项所述的图像处理设备中,在计算所述随机生成模型中的视图变换的后验概率时,使用 HMM、FNN、RNN、粒子滤波器和卡尔曼滤波器中的一个作为基于视图变换运动来估计先验概率的动态学习估计模型。

[0519] (10) 在(1)到(9)中的任一项所述的图像处理设备中,在所述前景学习单元中使用的随机生成模型不同于在所述背景学习单元中使用的随机生成模型。

[0520] (11) 在(1)到(10)中的任一项所述的图像处理设备中,所述图像特征输出单元输出每一像素的像素位置和像素值彼此对应的信息,作为所述图像特征。

[0521] (12) 在(1)到(10)中的任一项所述的图像处理设备中,所述图像特征输出单元输出利用哈里斯角点检测方法而检测到的特征点位置集合,作为所述图像特征。

[0522] (13) 一种图像处理方法包括:由图像特征输出单元来与帧时间对应地输出图像特征中的每一个,所述图像特征被形成作为输入活动图像数据中的每一帧的图像的多个特征点中的特征;由前景估计单元通过在前景视图模型上进行作为几何变换的视图变换来估计时间 s 处的前景图像,并且输出所估计的前景视图,所述前景视图模型具有所述图像中的前景的图像特征,作为与所述时间 s 处的图像特征有关的参数;由背景估计单元通过在背景视图模型上进行作为几何变换的视图变换来估计所述时间 s 处的背景图像,并输出所估计的背景视图,所述背景视图模型具有所述图像中的背景的图像特征,作为与所述时间 s 处的图像特征有关的参数;由合成视图生成单元通过合成所估计的前景视图和所估计的背景视图来生成合成视图;由前景学习单元基于通过所述合成视图和所述时间 s 处的图像特征之间进行比较而得到的估计值并通过基于随机生成模型来更新所述前景视图模型的参数,来学习所述前景视图模型;以及由背景学习单元基于通过基于随机生成模型来更新所述背景视图模型的参数而得到的估计值,来学习所述背景视图模型。

[0523] (14) 一种程序使得计算机作为图像处理设备来工作,该图像处理设备包括:图像特征输出单元,用于与帧时间对应地输出图像特征中的每一个,所述图像特征被形成作为输入活动图像数据中的每一帧的图像的多个特征点中的特征;前景估计单元,用于通过在前景视图模型上进行作为几何变换的视图变换来估计时间 s 处的前景图像,并且输出所估计的前景视图,所述前景视图模型具有所述图像中的前景的图像特征,作为与所述时间 s 处的图像特征有关的参数;背景估计单元,用于通过在背景视图模型上进行作为几何变换的视图变换来估计所述时间 s 处的背景图像,并输出所估计的背景视图,所述背景视图模

型具有所述图像中的背景图像特征,作为与所述时间 s 处的图像特征有关的参数;合成视图生成单元,用于通过合成所估计的前景视图和所估计的背景视图来生成合成视图;前景学习单元,用于基于通过所述合成视图和所述时间 s 处的图像特征之间进行比较而得到的估计值并通过基于随机生成模型来更新所述前景视图模型的参数,来学习所述前景视图模型;以及背景学习单元,用于基于通过基于随机生成模型来更新所述背景视图模型的参数而得到的估计值,来学习所述背景视图模型。

[0524] (15) 一种记录介质,用于存储(14)中所述的程序。

[0525] 本公开包含与2011年3月31日提交于日本专利局的日本优先权专利申请JP 2011-077696中披露的主题相关的主题,该日本申请的所有内容通过引用合并于此。

[0526] 本领域的技术人员应理解,根据设计需求或者其他因素可以进行各种修改、组合、子组合和变更,只要这些修改、组合、子组合和变更在所附权利要求及其等同的范围内即可。

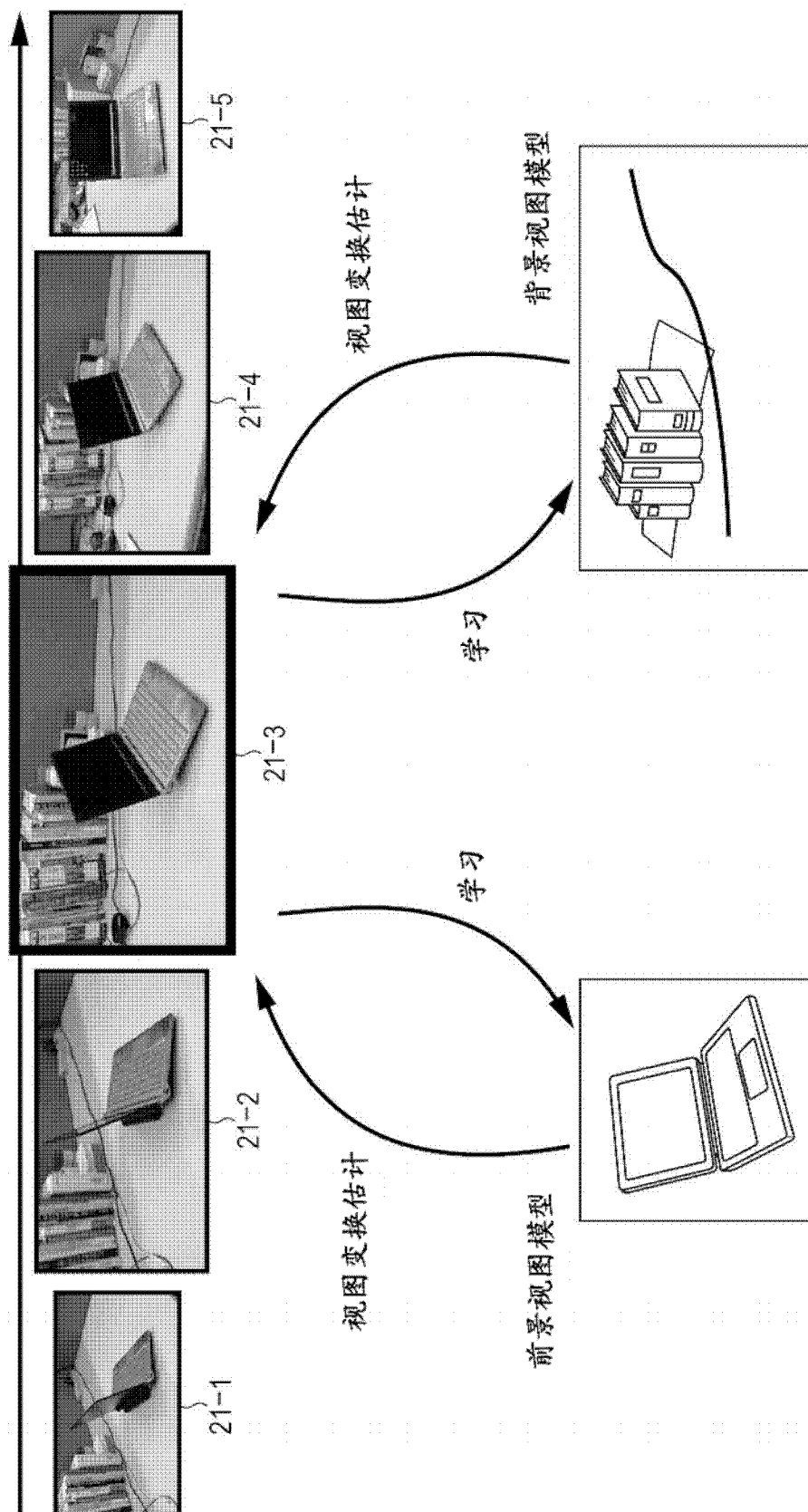


图 1

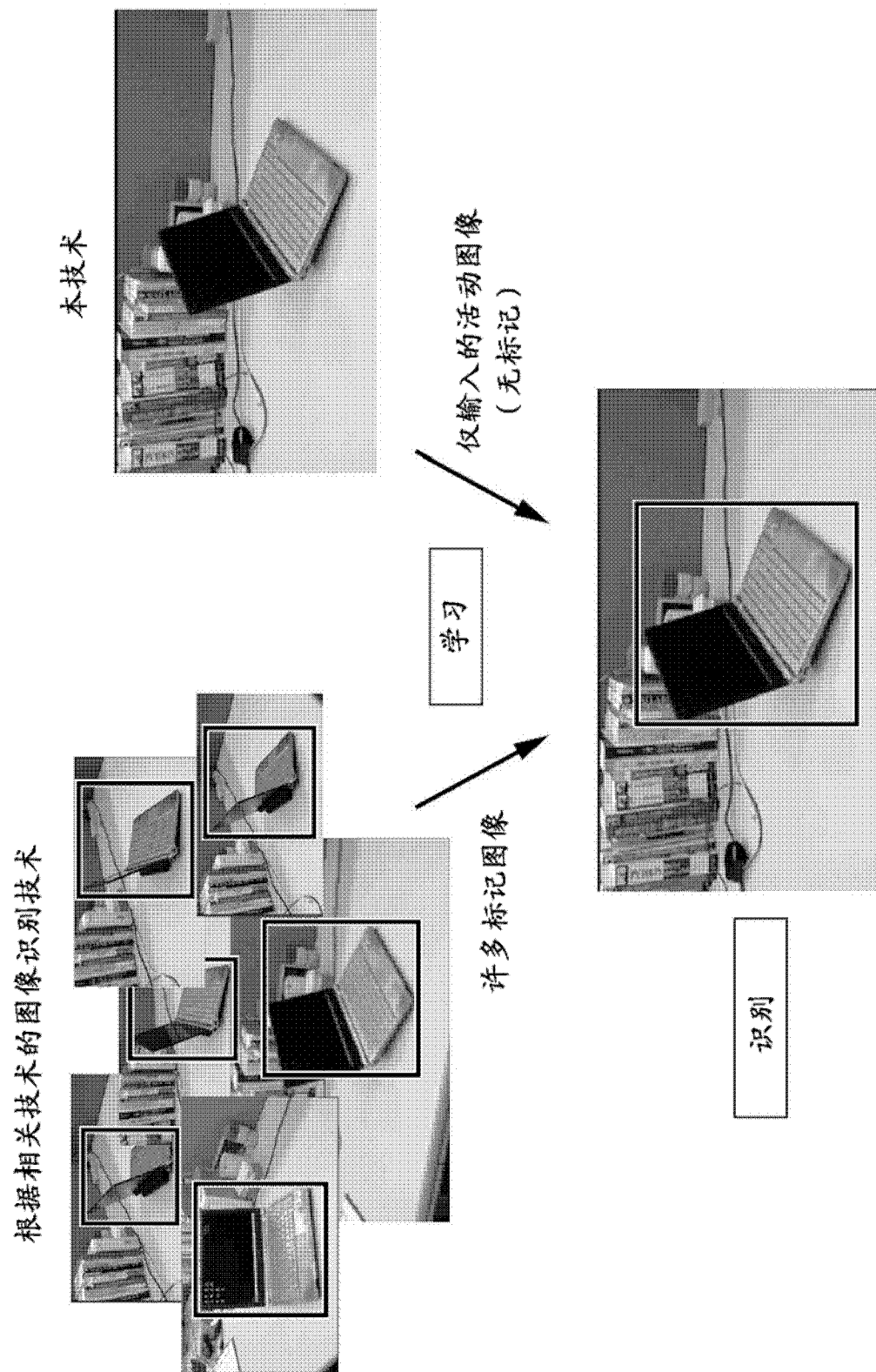


图 2

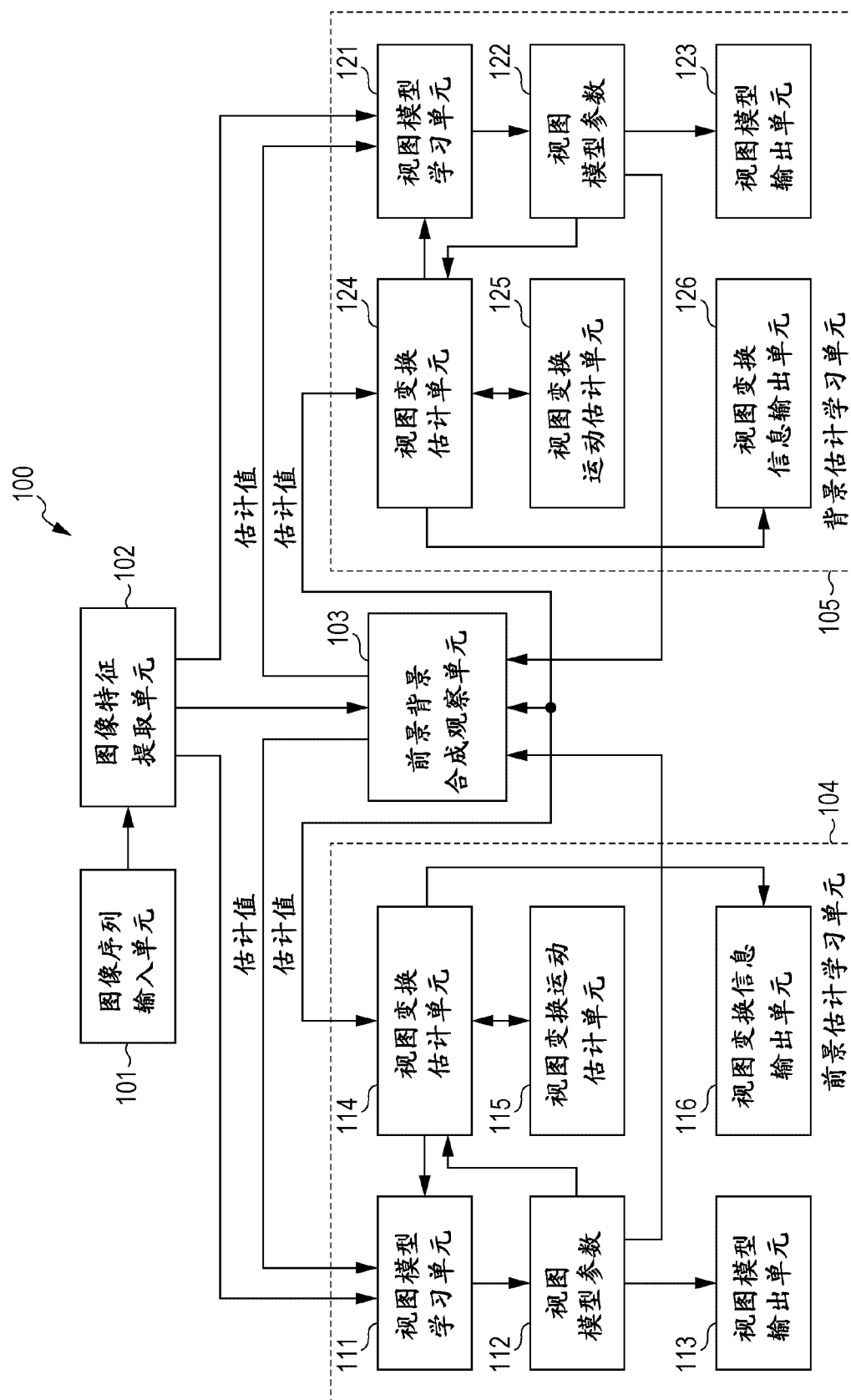


图 3

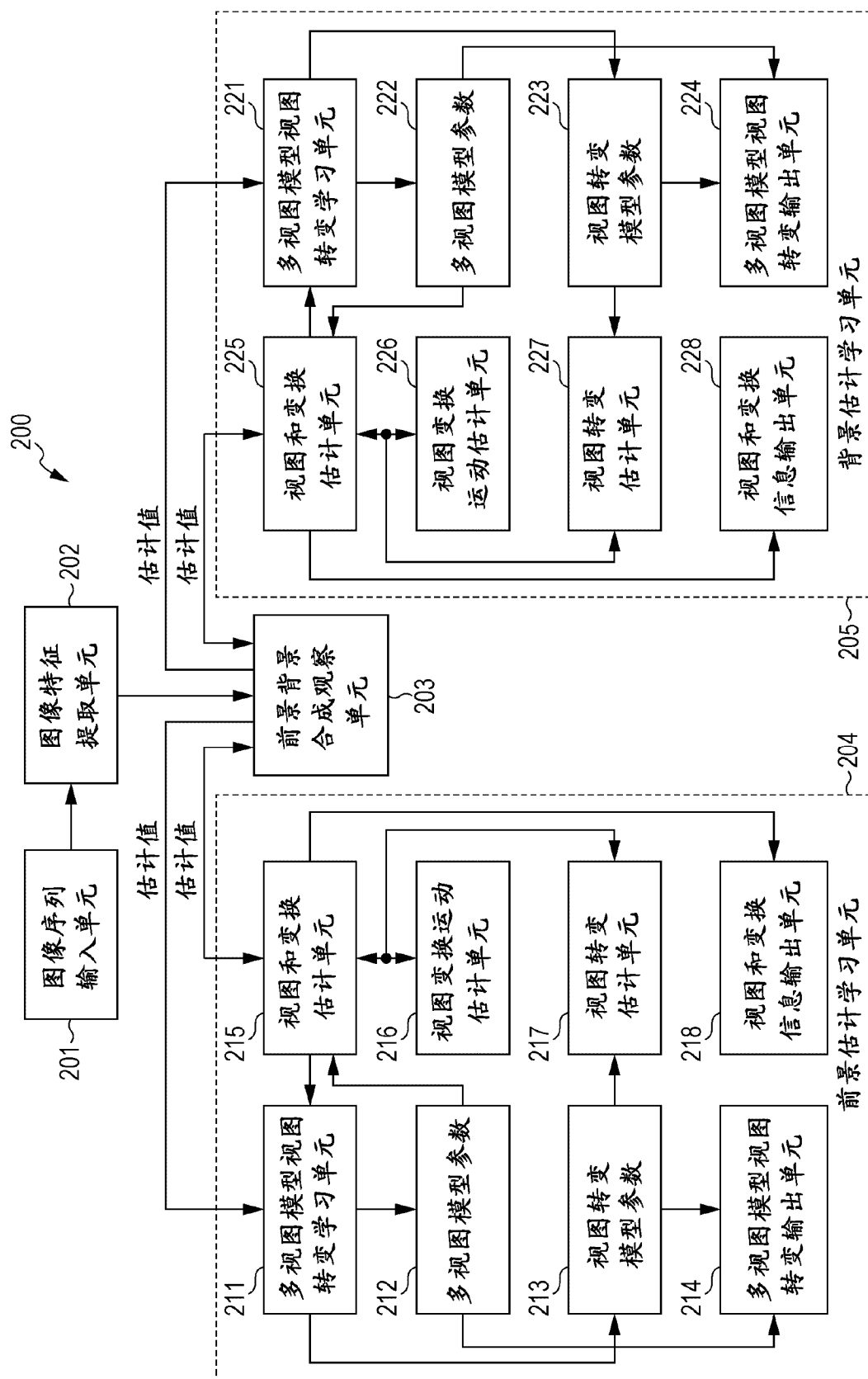


图 4

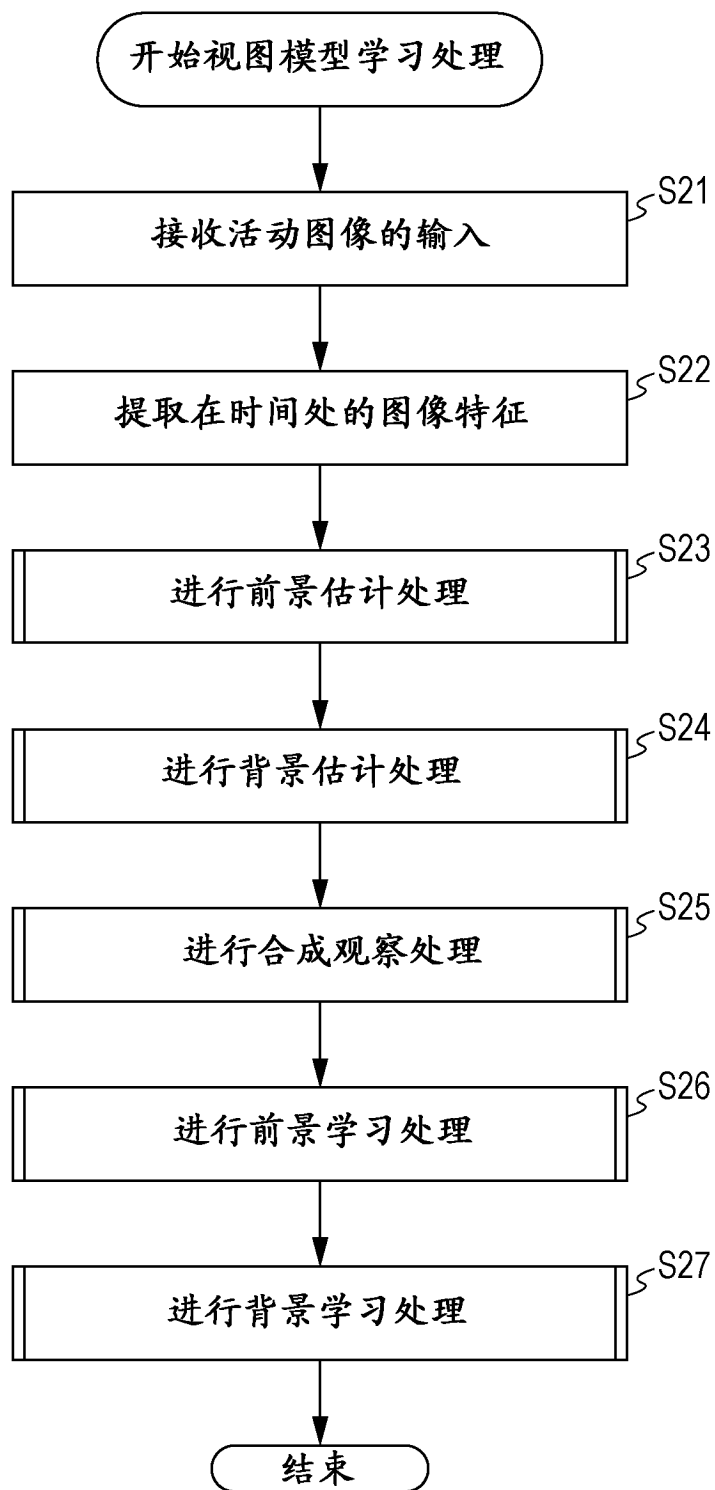


图 5

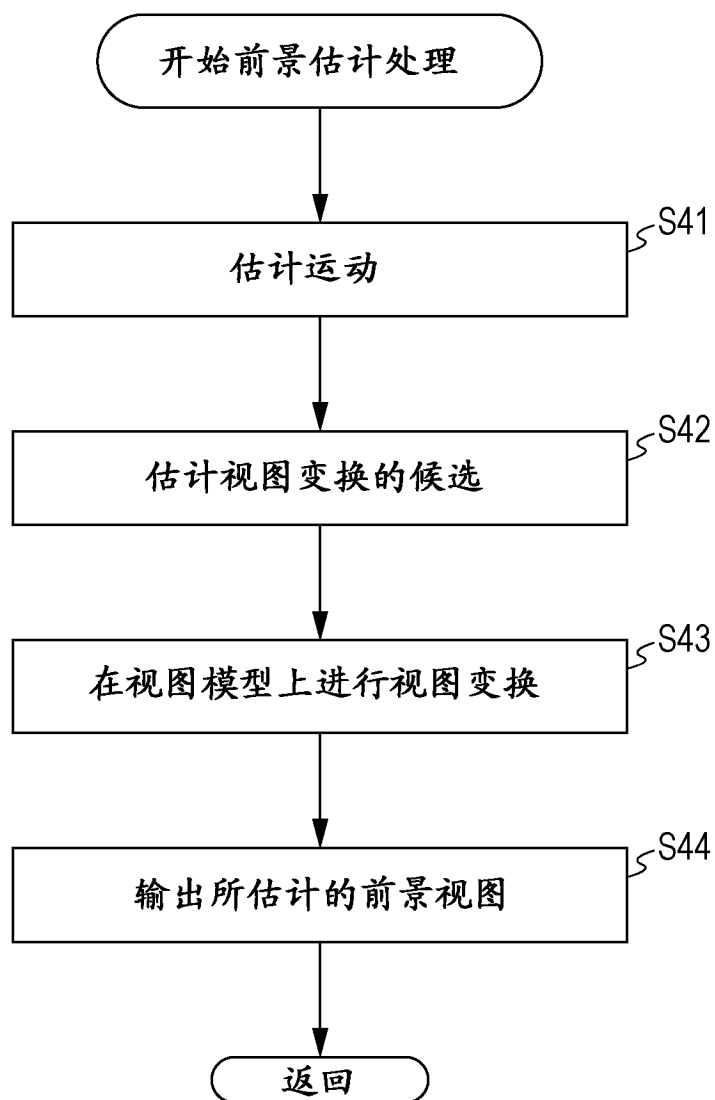


图 6

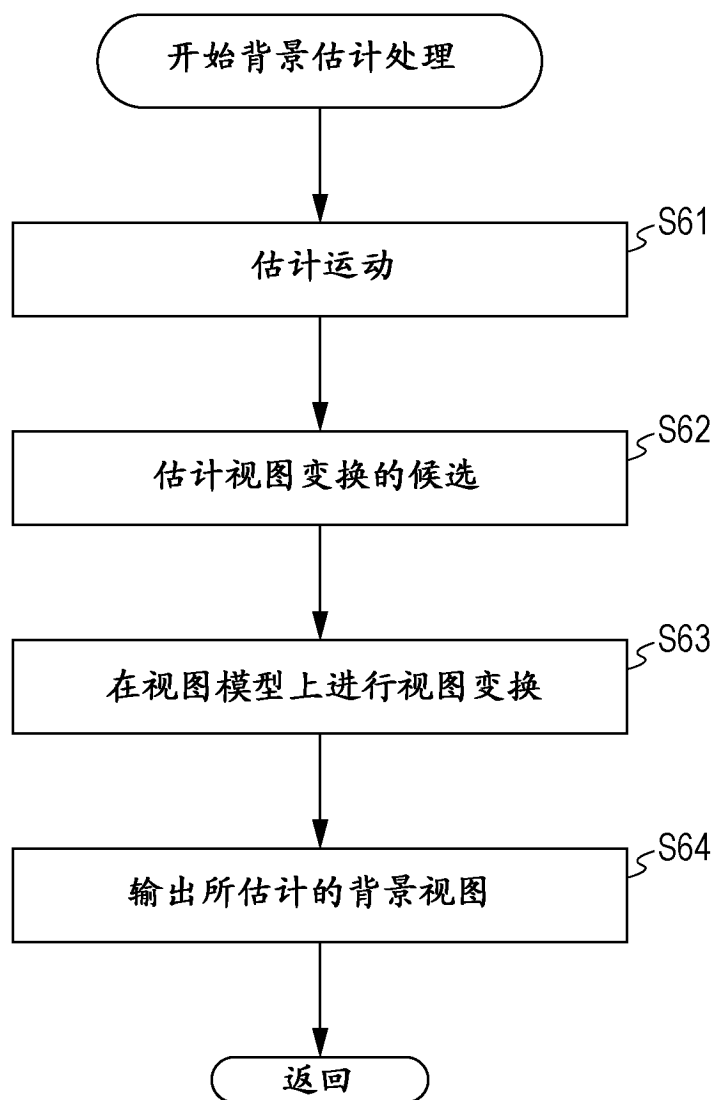


图 7

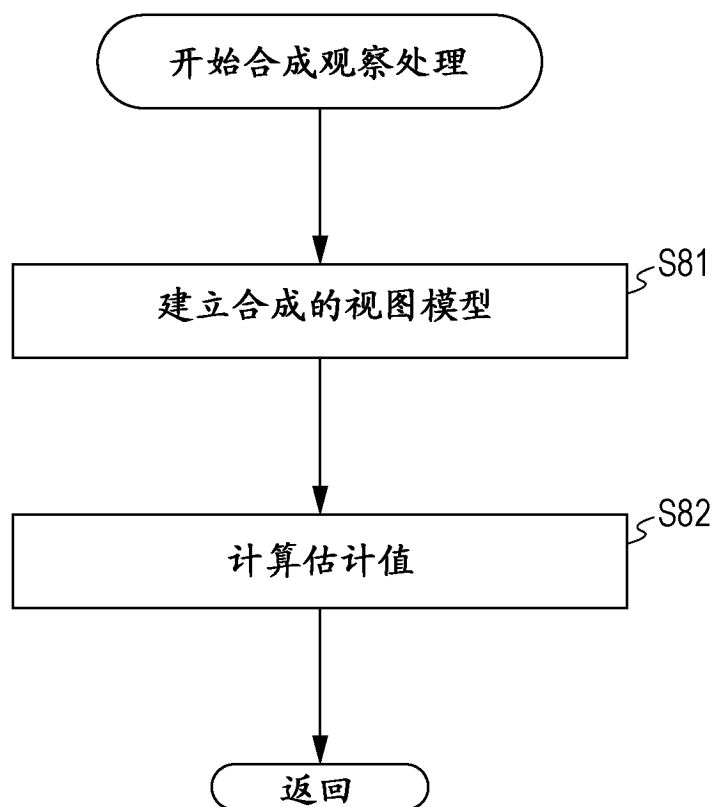


图 8

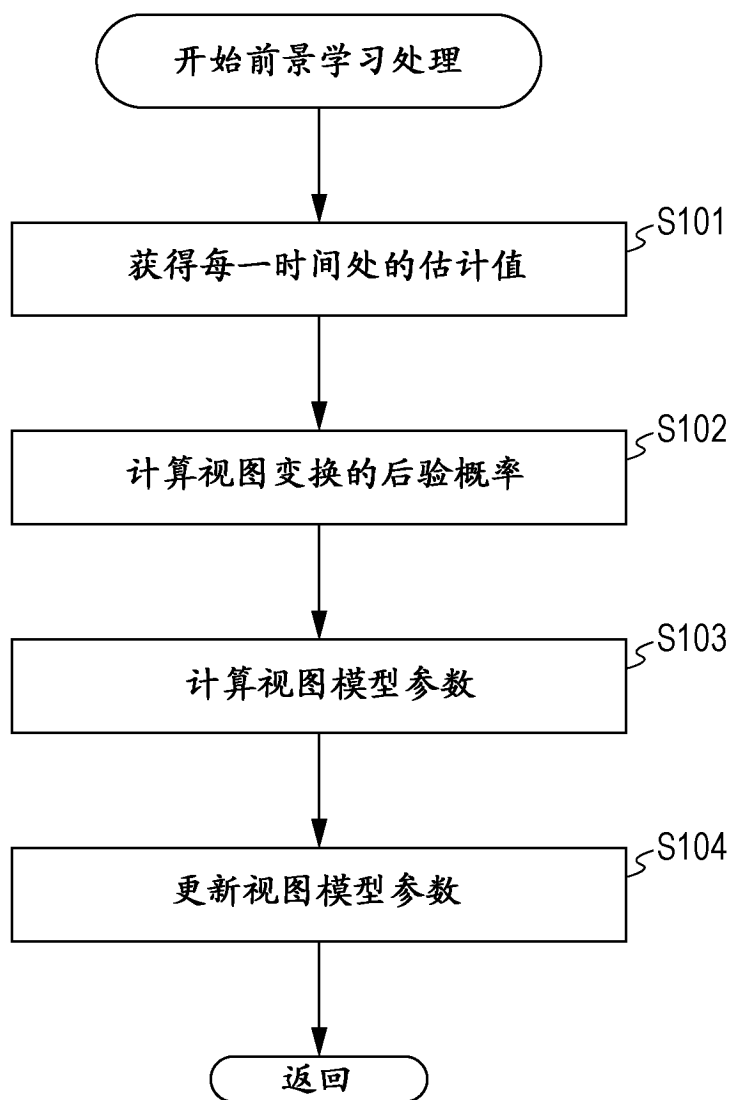


图 9

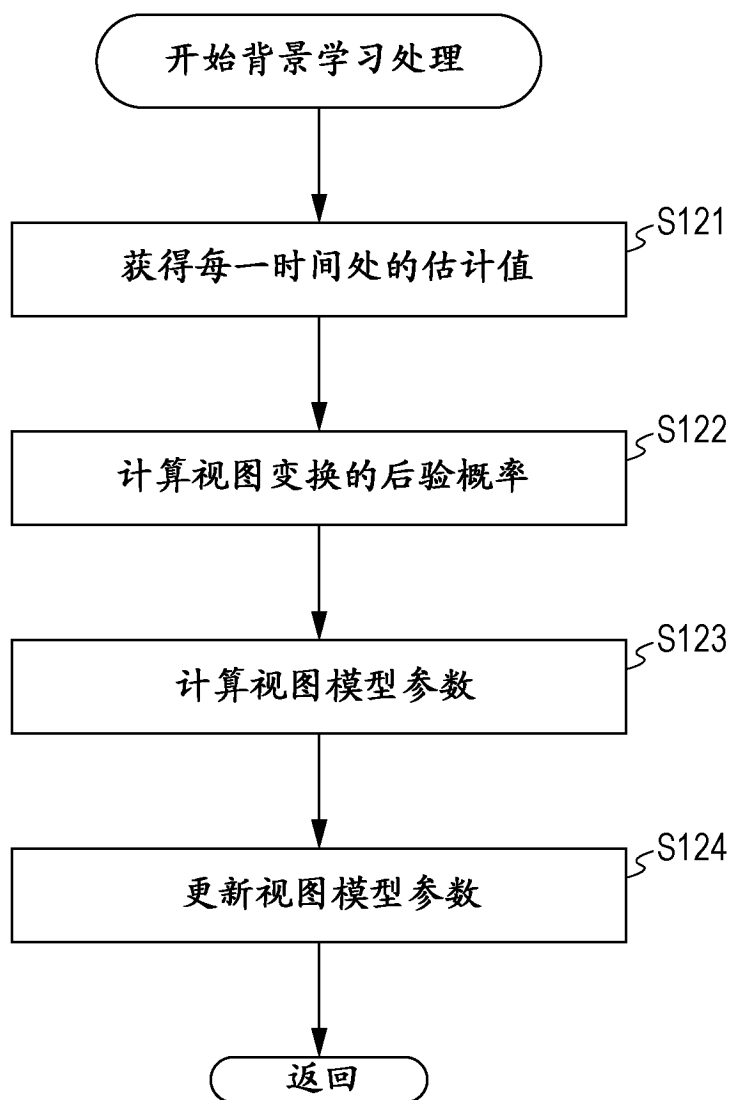


图 10

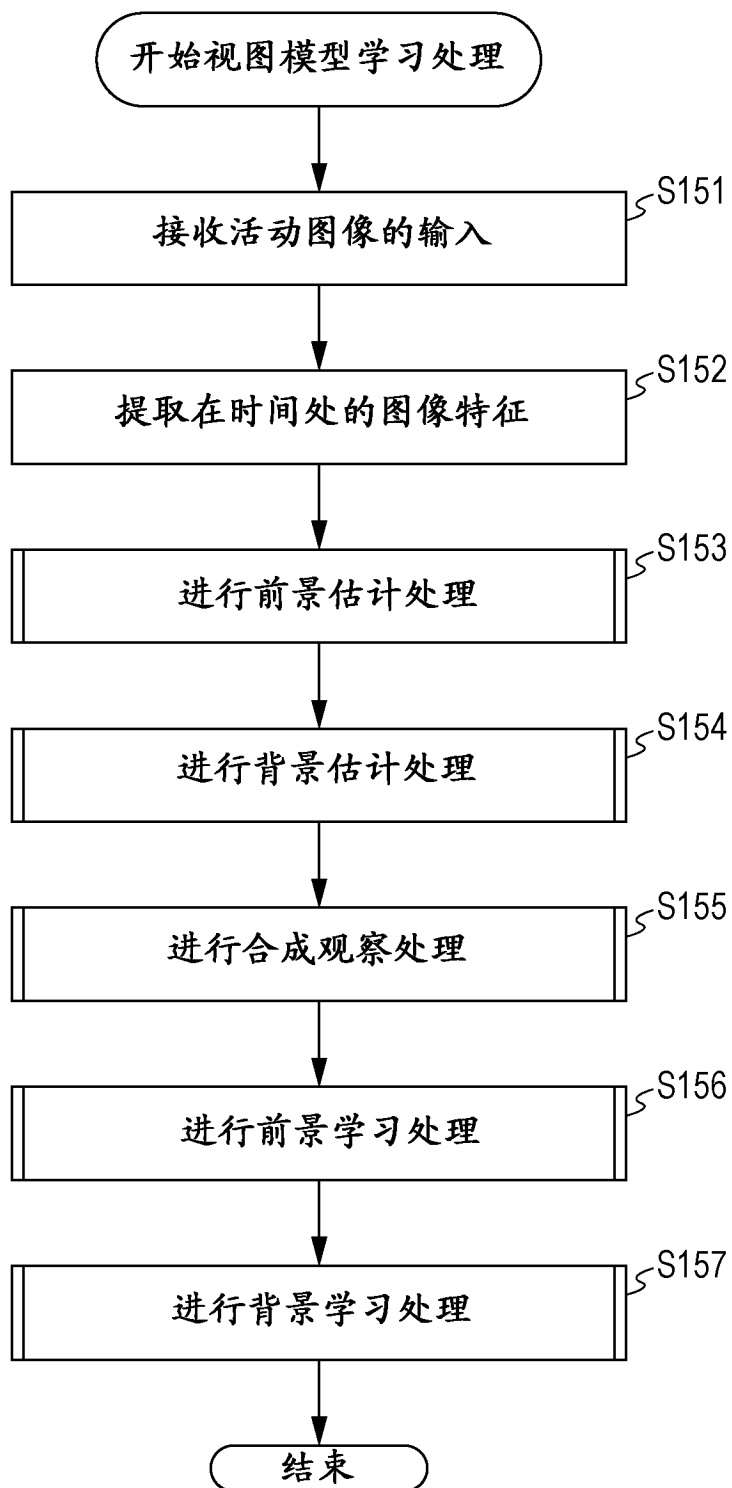


图 11

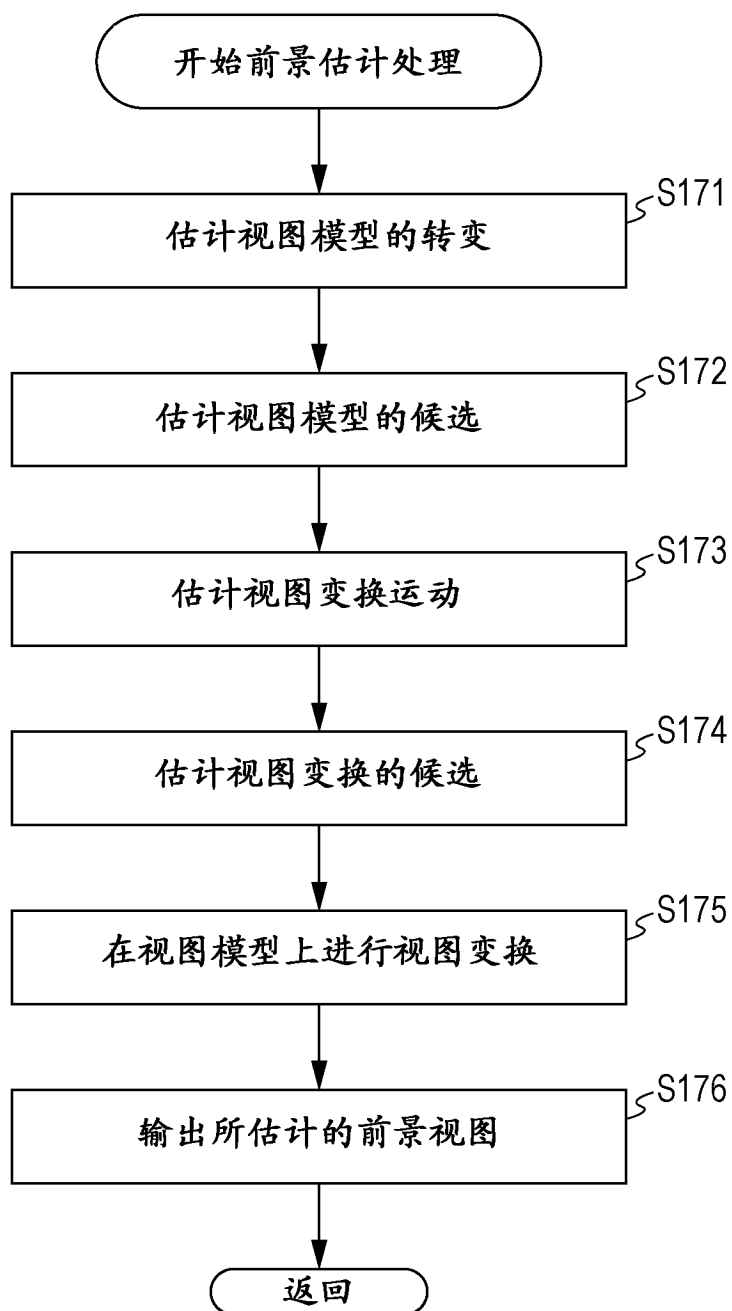


图 12

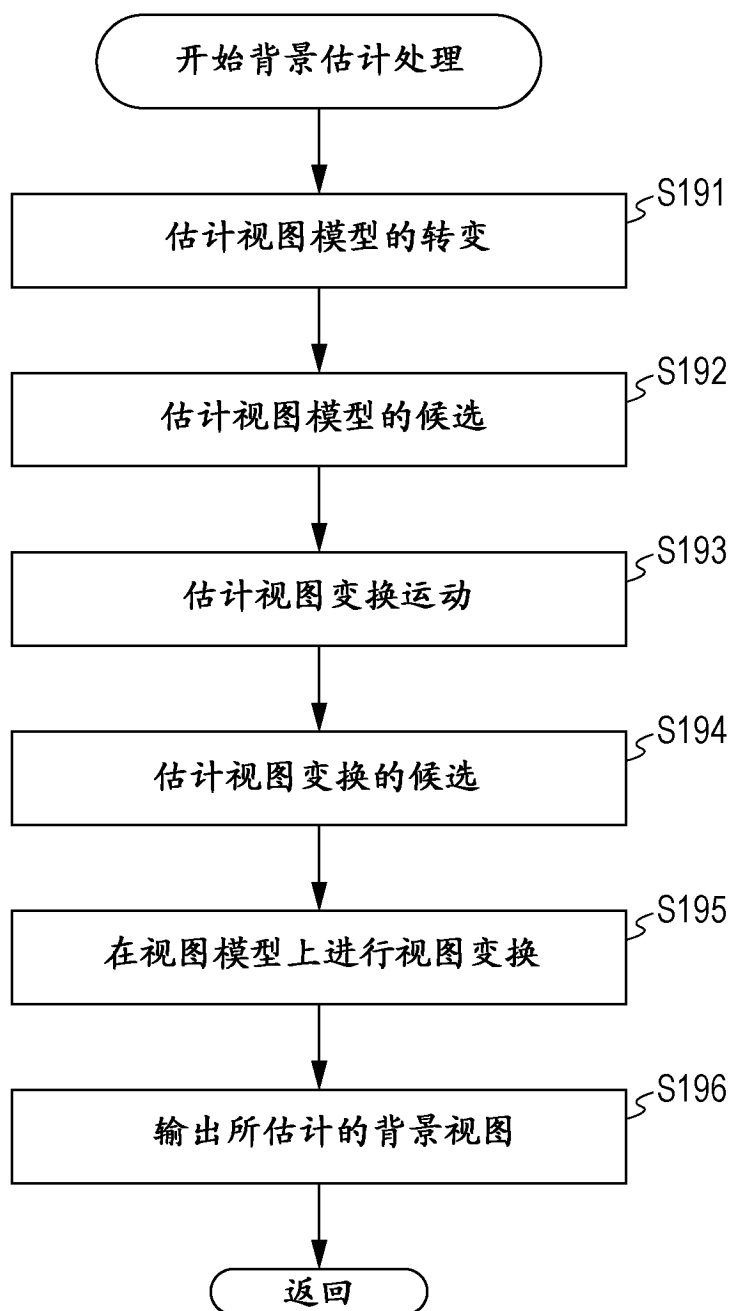


图 13

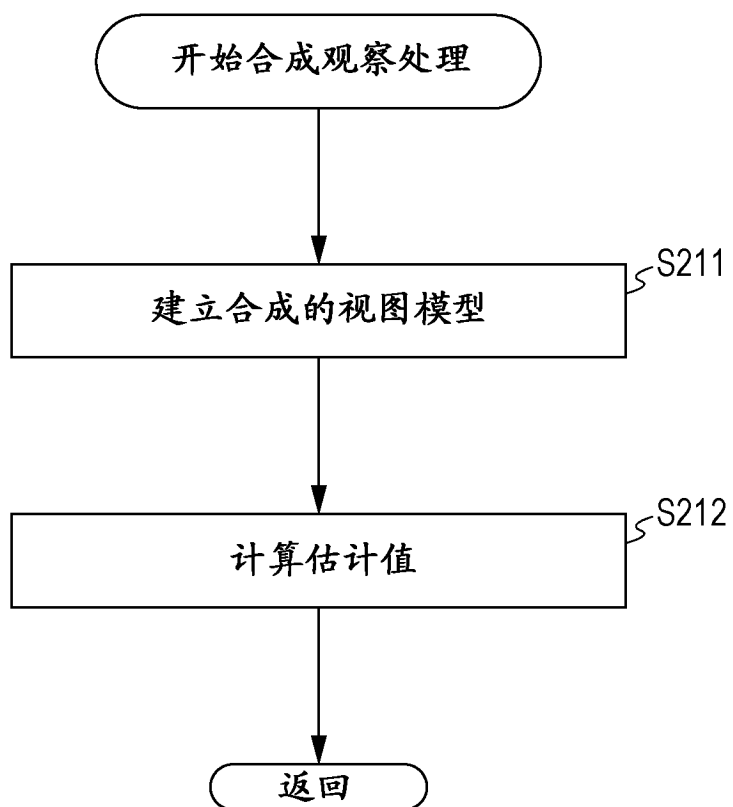


图 14

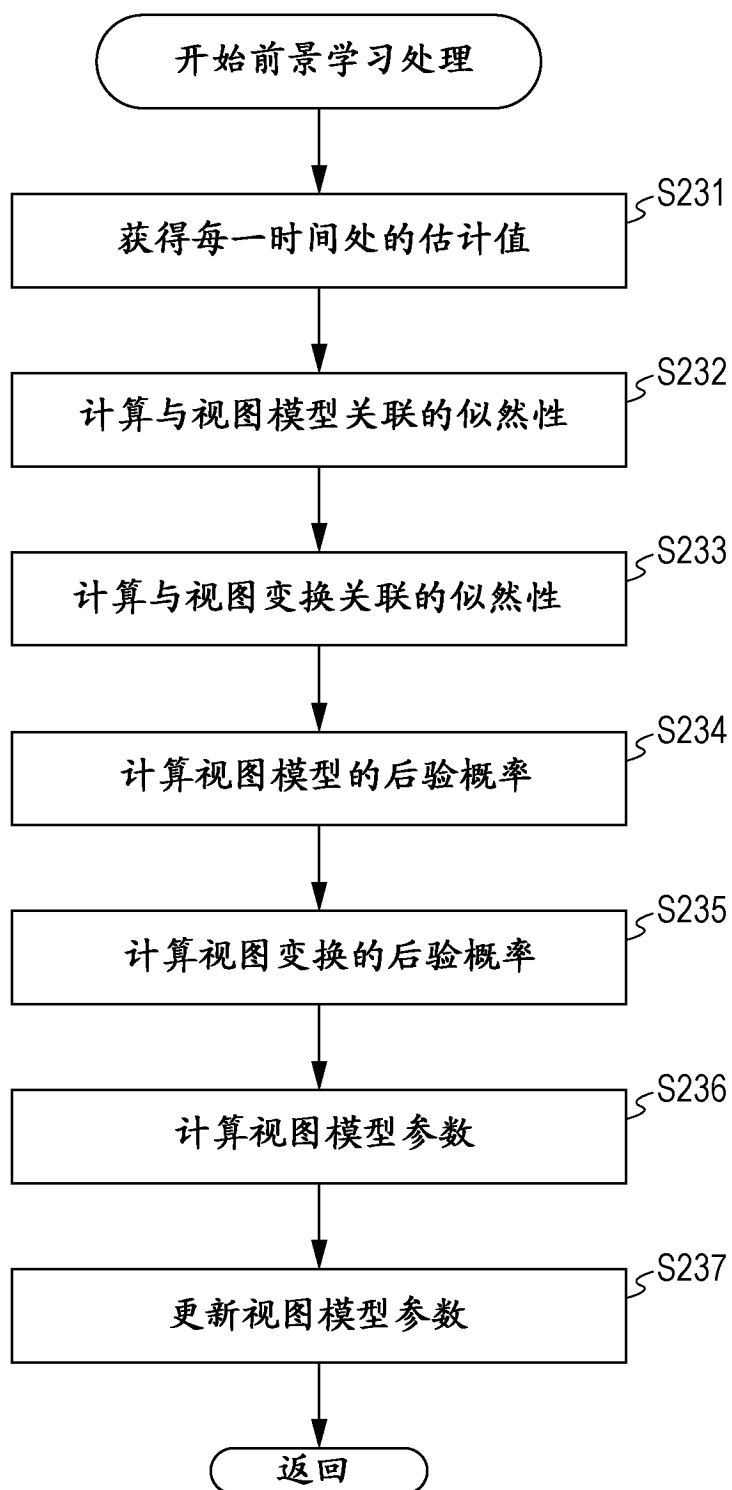


图 15

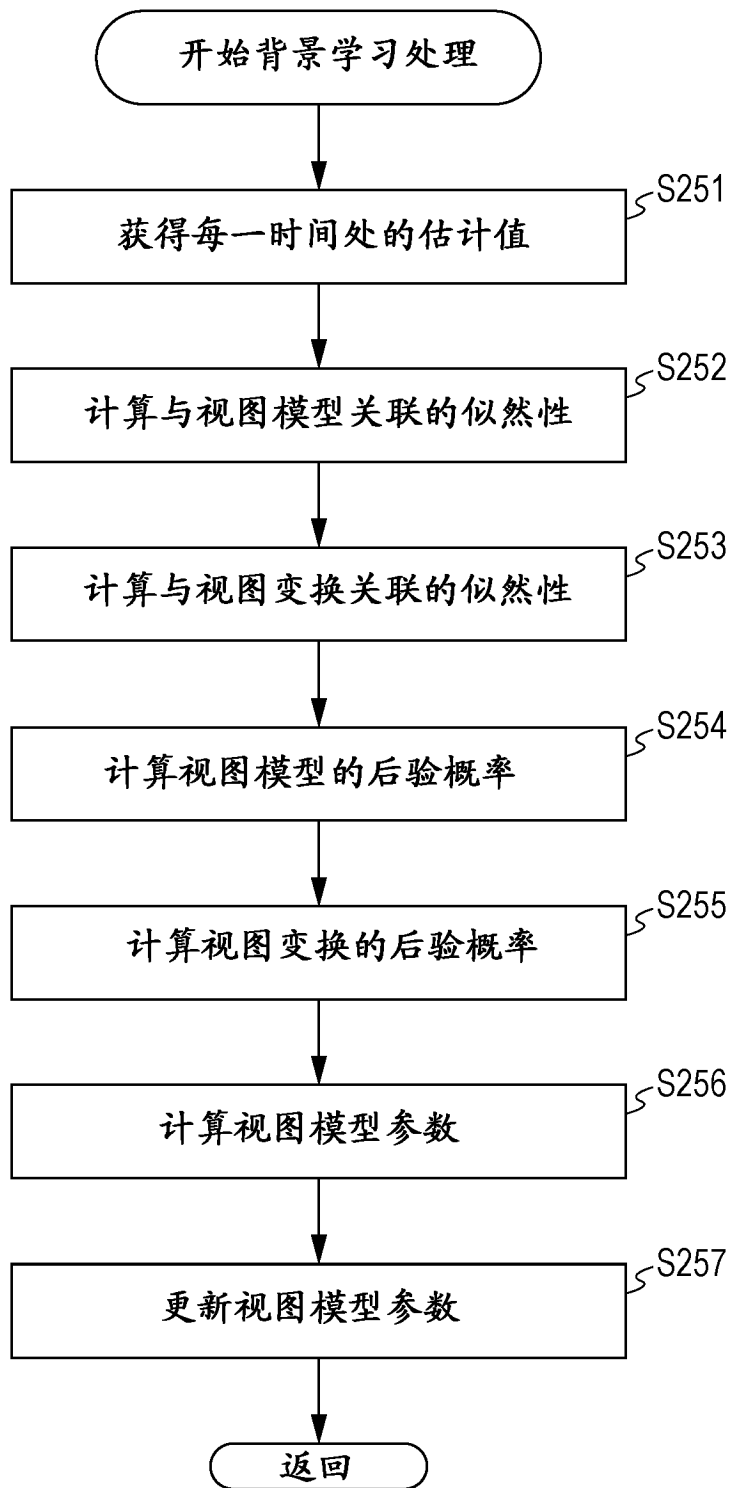


图 16

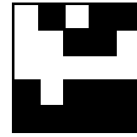


图 17

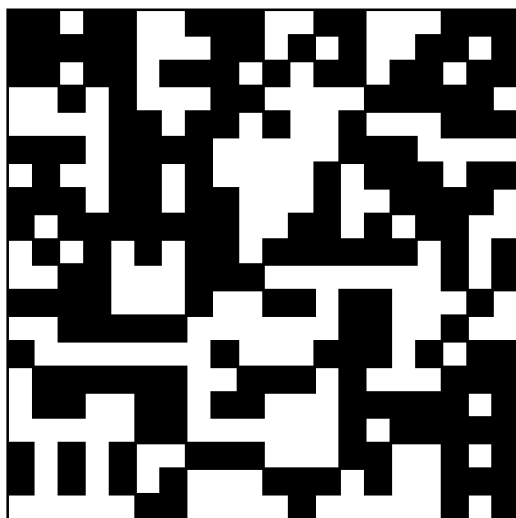


图 18

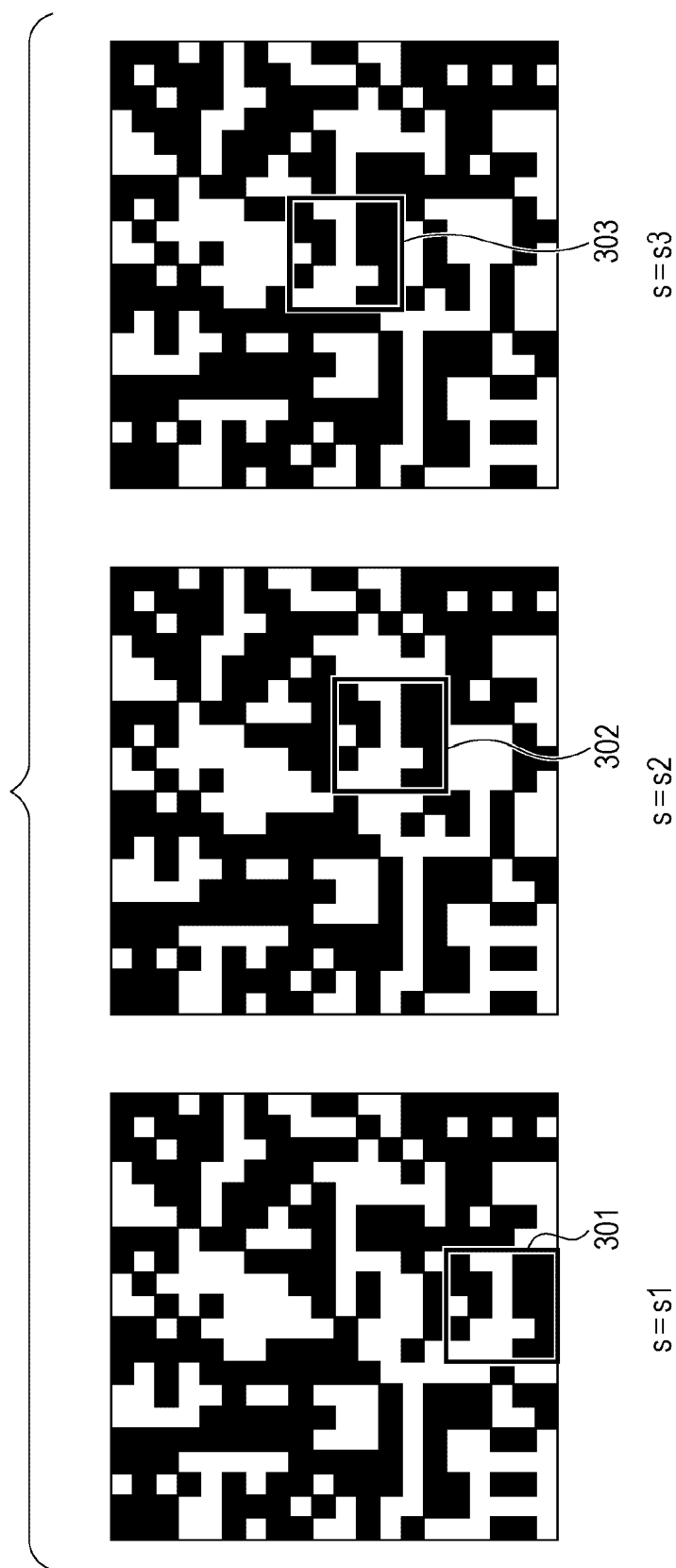


图 19

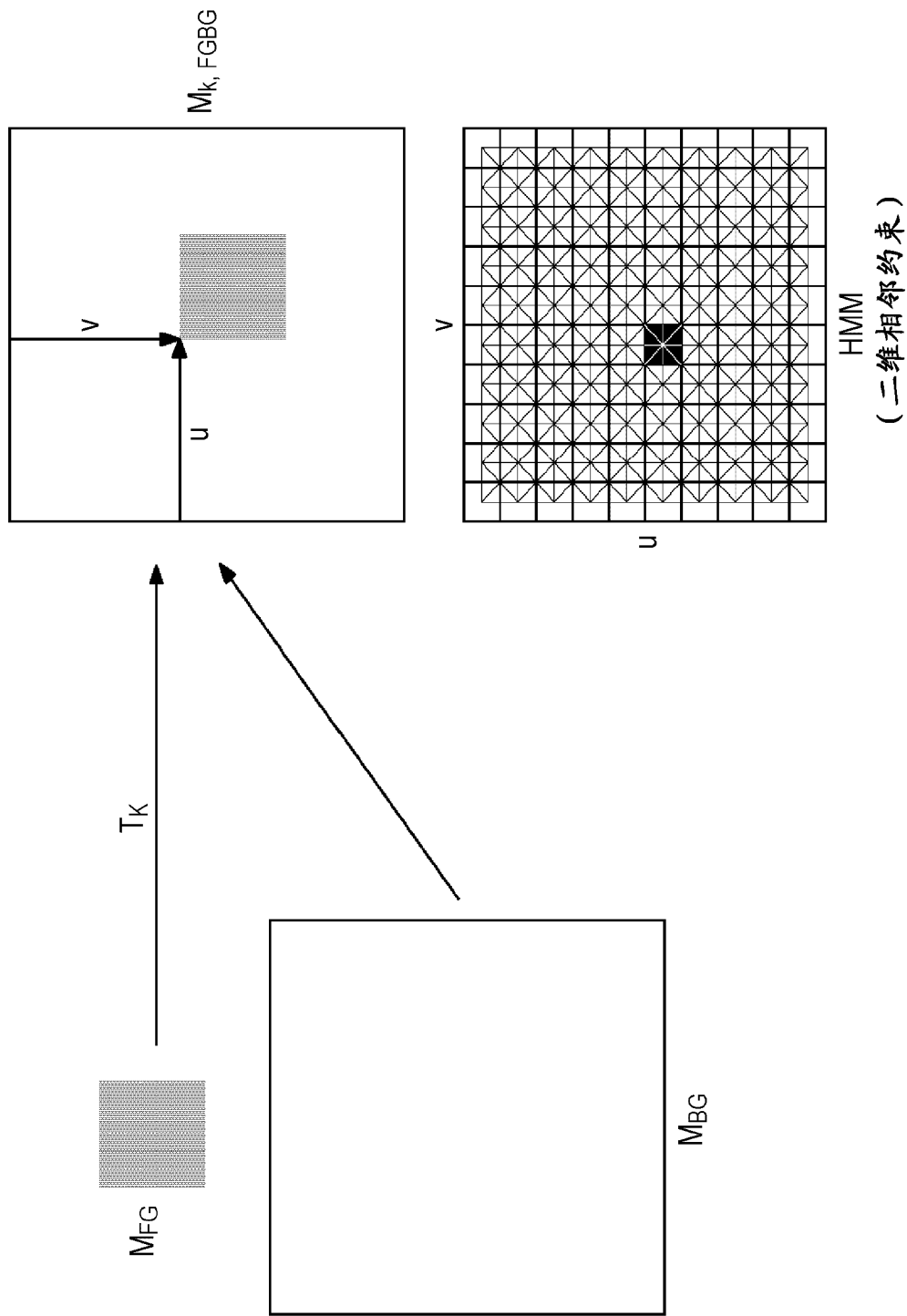


图 20

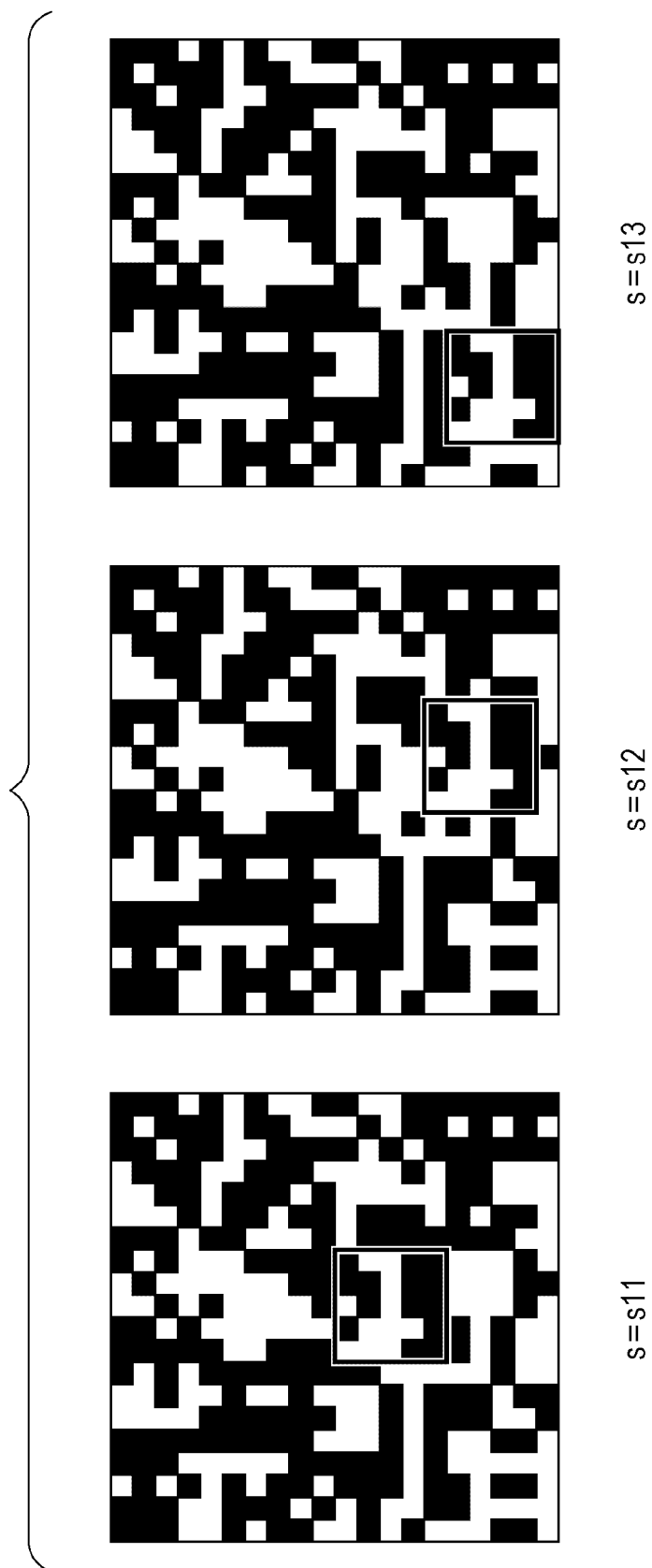


图 21

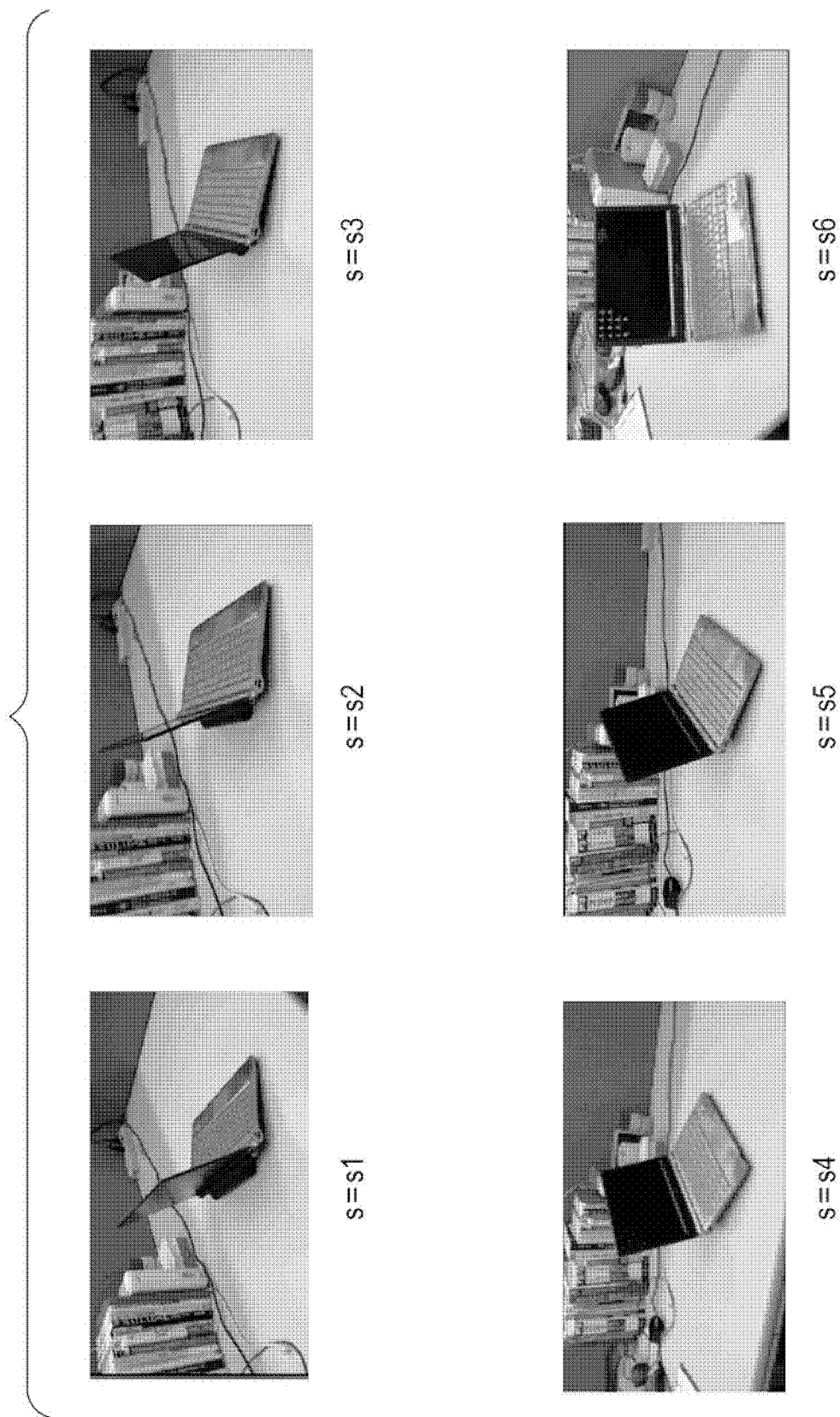


图 22

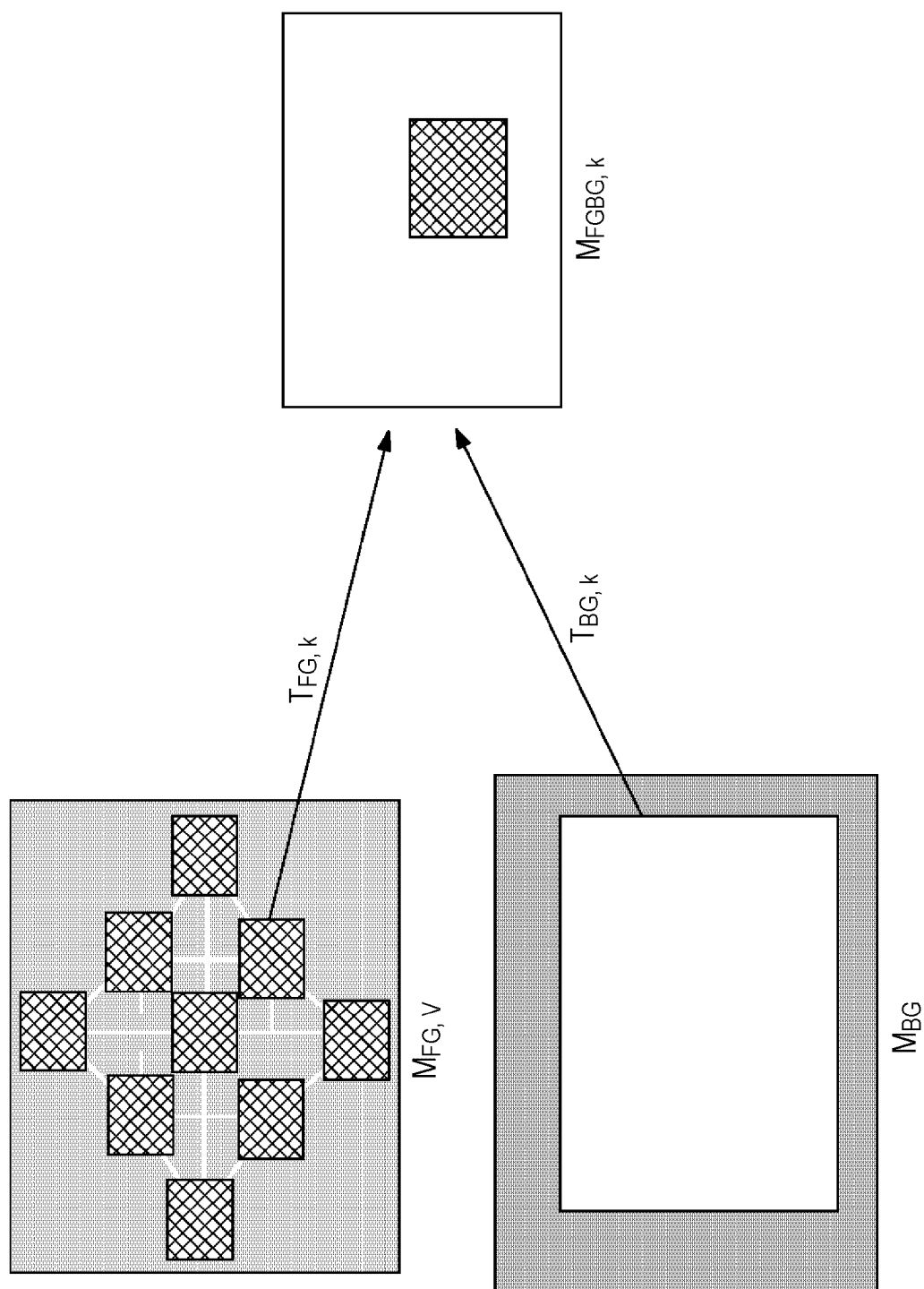


图 23

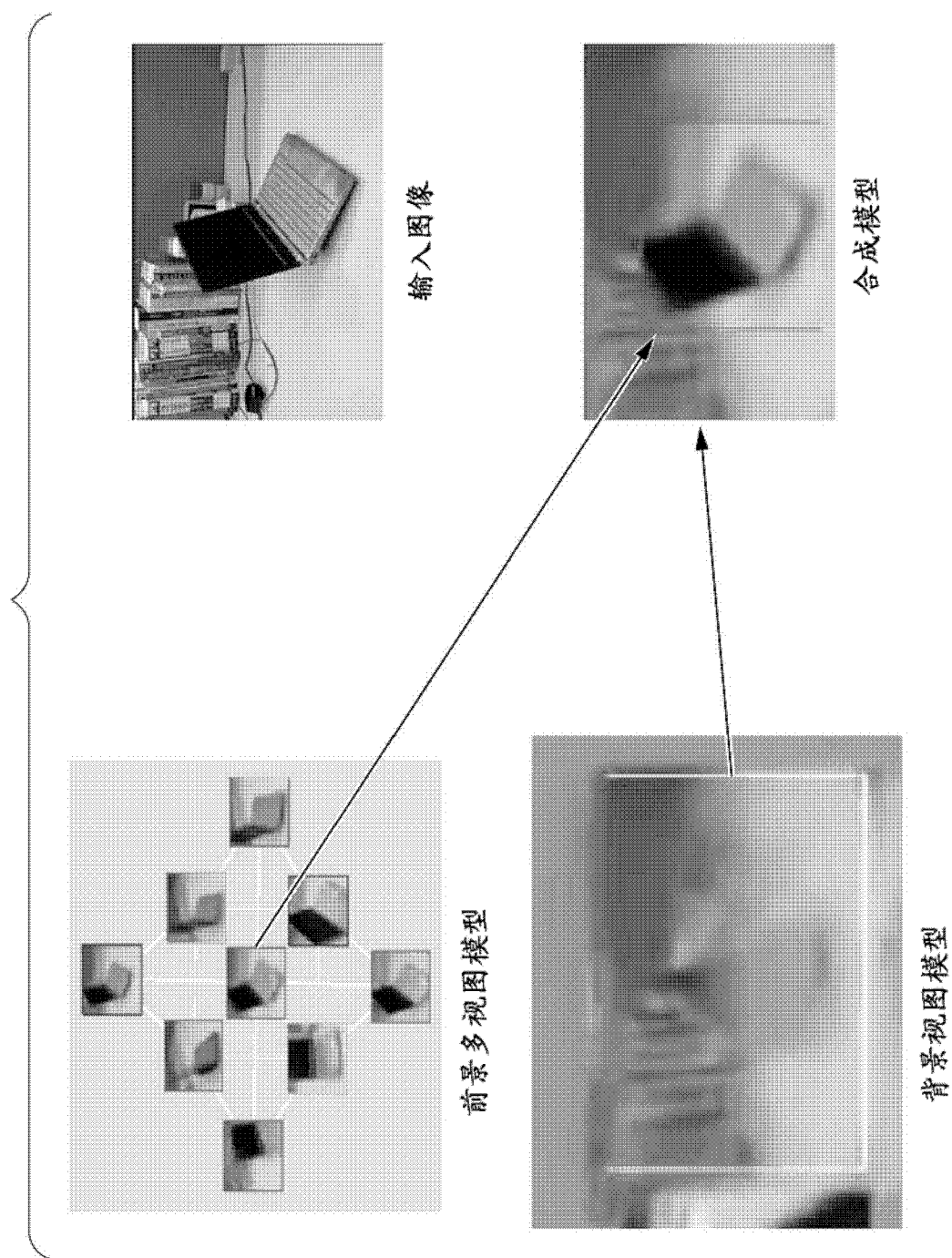


图 24

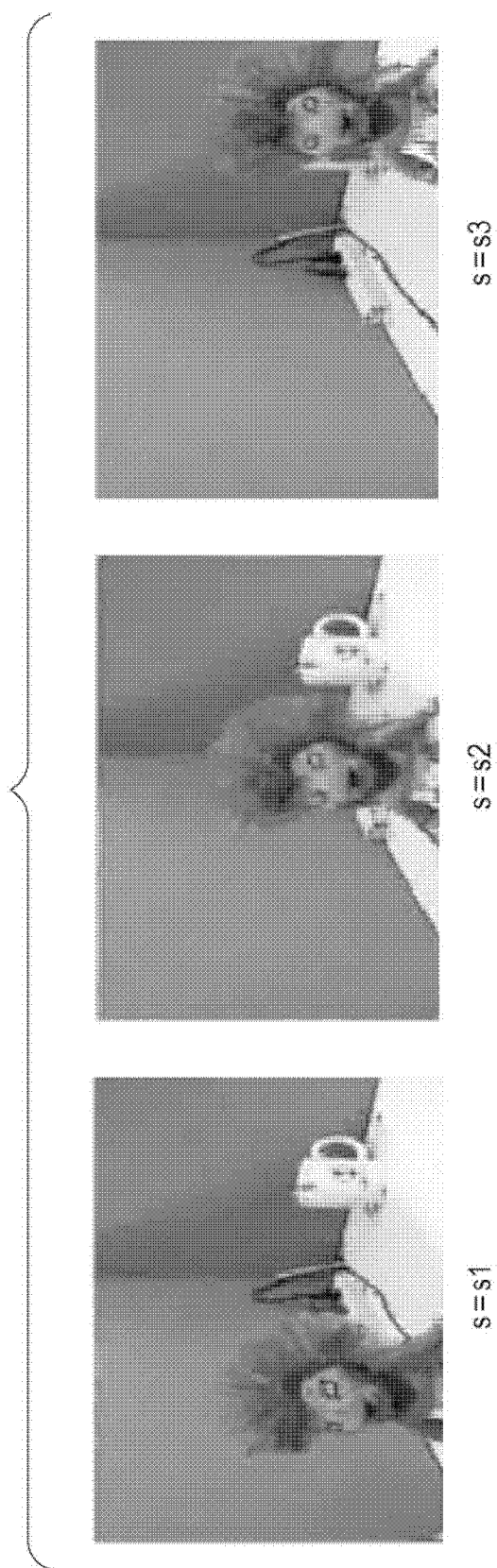


图 25

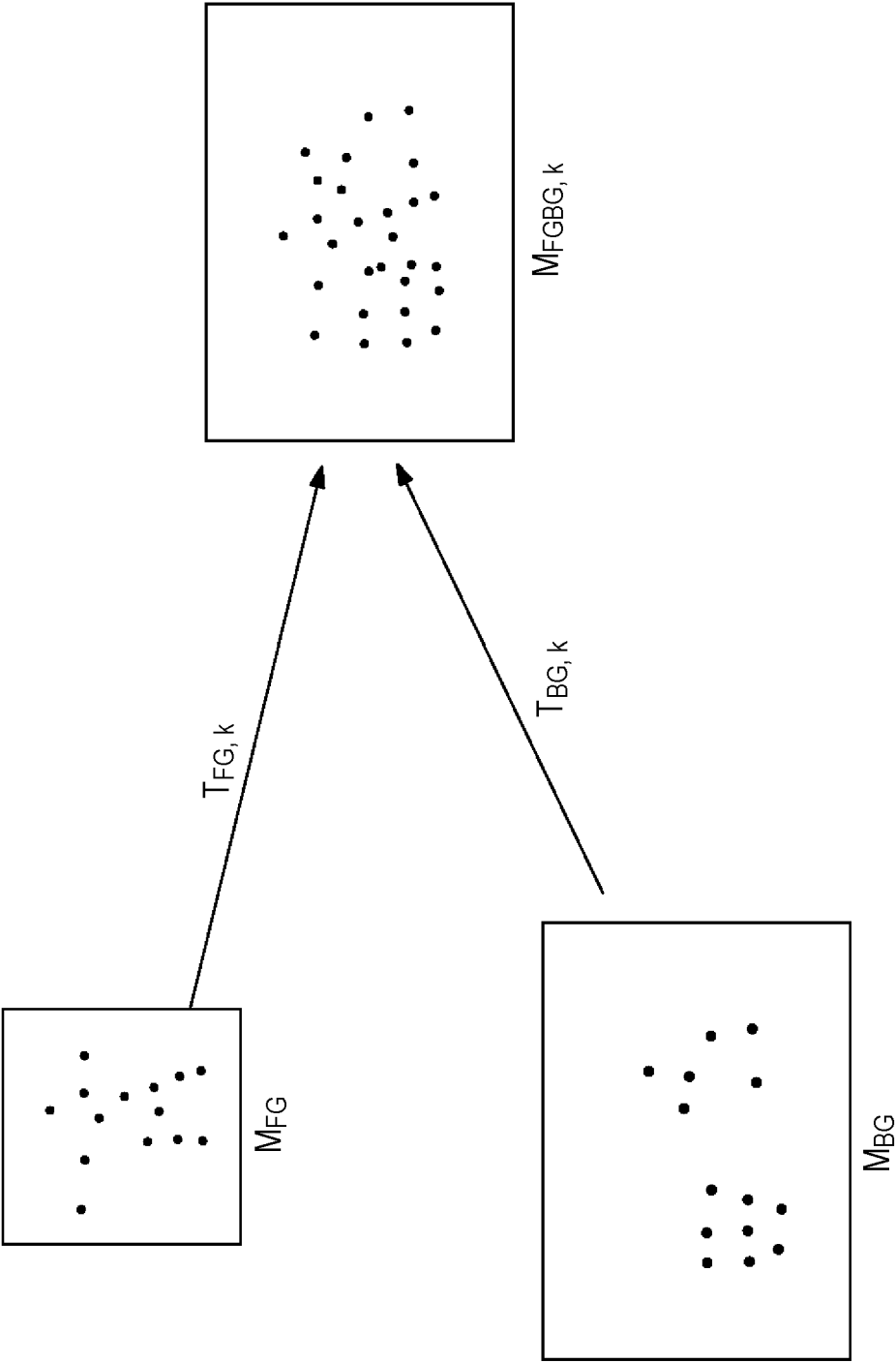


图 26

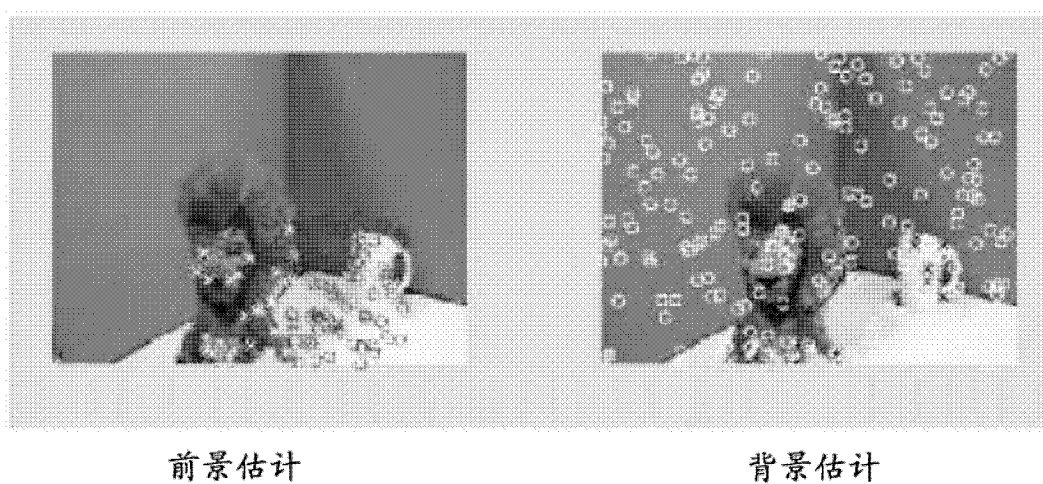


图 27

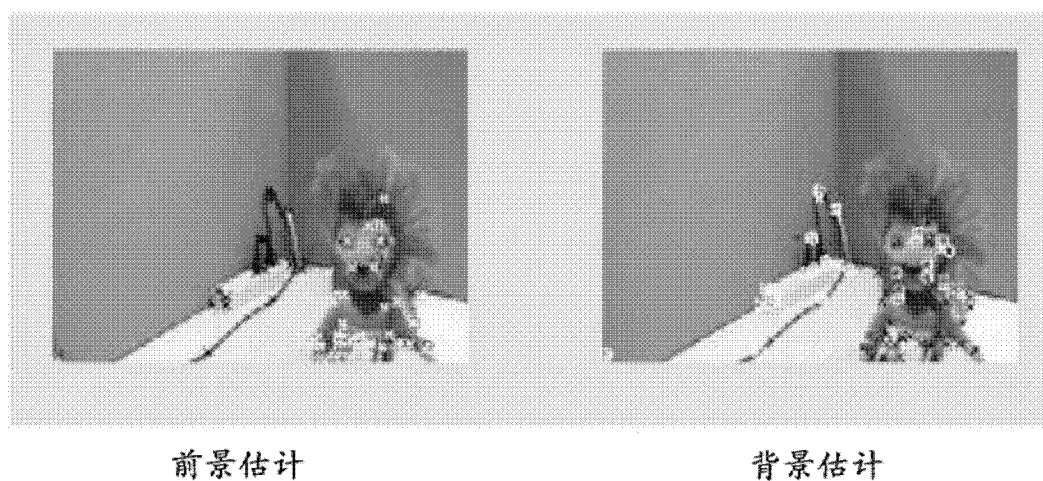


图 28