

(12) 发明专利申请

(10) 申请公布号 CN 103069452 A

(43) 申请公布日 2013. 04. 24

(21) 申请号 201180040102. 0

(51) Int. Cl.

(22) 申请日 2011. 08. 19

G06T 1/00 (2006. 01)

(30) 优先权数据

61/375, 511 2010. 08. 20 US

61/509, 038 2011. 07. 18 US

(85) PCT申请进入国家阶段日

2013. 02. 19

(86) PCT申请的申请数据

PCT/JP2011/068734 2011. 08. 19

(87) PCT申请的公布数据

W02012/023596 JA 2012. 02. 23

(71) 申请人 国立大学法人东京大学

地址 日本东京都

(72) 发明人 渡边义浩 石川正俊 柴山裕树

(74) 专利代理机构 北京市柳沈律师事务所  
11105

代理人 车玲玲

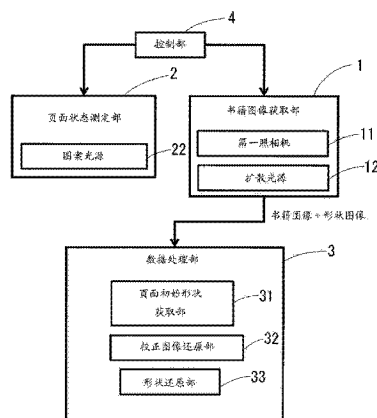
权利要求书1页 说明书12页 附图8页

(54) 发明名称

书籍读取系统及书籍读取方法

(57) 摘要

提供一种一边翻开装订的书籍一边高速读取记载于书籍的信息的技术。页面状态测定部(2)测定已翻开的书籍的页面状态。书籍图像获取部(1)的多个第一照相机(11)从彼此不同的方向多次分别获取有关书籍的页面的书籍图像。数据处理部(3)使用由页面状态测定部(2)测定的书籍的页面状态的信息修正由书籍图像获取部(1)获取的书籍图像。



1. 一种书籍读取系统,其特征在于,  
具备页面状态测定部、书籍图像获取部和数据处理部,  
所述页面状态测定部的结构为测定已翻开的书籍的页面状态,  
所述书籍图像获取部具备多个第一照相机,  
所述多个第一照相机的结构为从彼此不同的方向分别获取有关所述书籍的页面的书籍图像,

所述数据处理部的结构为使用由所述页面状态测定部测定的所述书籍的页面状态的信息校正由所述书籍图像获取部获取的所述书籍图像。

2. 如权利要求 1 所述的书籍读取系统,其中,  
所述页面状态测定部具备用于朝向所述书籍的页面投影既定的图案的图案光源,  
所述多个第一照相机的结构为进一步获取投影了所述图案的页面的形状图像。

3. 如权利要求 2 所述的书籍读取系统,其中,  
所述多个第一照相机的结构为在从所述图案光源投影所述图案的定时获取所述形状图像且在未从所述图案光源投影所述图案的定时获取所述书籍图像。

4. 如权利要求 1 所述的书籍读取系统,其中,  
所述页面状态测定部具备形状图像用的第二照相机,  
所述第二照相机具备使从所述图案光源投影的波长的光透过且遮蔽其以外的环境光的滤波器,

所述图案光源的结构为使用既定的波长的光朝向所述书籍的页面投影所述既定的图案,

由此,所述第二照相机能够获取投影于所述书籍的页面的所述既定的图案。

5. 如权利要求 1 ~ 4 中任一项所述的书籍读取系统,其中,  
所述第一照相机的结构为在所述书籍的每一页面获取多张的书籍图像,  
所述数据处理部的结构为通过统一所述多张的书籍图像的信息从而提高校正了页面形状的失真的书籍图像的精确度。

## 书籍读取系统及书籍读取方法

### 技术领域

[0001] 本发明涉及书籍读取系统及书籍读取方法。尤其,本发明涉及适用于一边翻开装订的书籍一边自动读取书籍中所记载的信息的技术。

### 背景技术

[0002] 利用多张图像还原形状的技术从模型的建模到建筑物的还原持续产生大量具有魅力的应用。但是,该大量的技术一直以对象物体为刚体为前提来解决各种课题。在对象为非刚体的情况下,由于形状变形的大前提的不同,要求采用新的方法解决问题。作为对象为非刚体时的形状还原,近年来提出了一些方法,但尚未确立如刚体的情况那般标准的方法。因此,在计算机视觉的领域成为挑战的课题。

[0003] 认为非刚体的变形估计具有技术挑战性,并且即使在其应用中,从产业检查到娱乐领域均可以广泛地做出创新的贡献。在目前为止的相同的方法涉及的报告中未提议确定的应用开展,但提出了高适用性的途径。例如,进行用于估计纸、罐、T恤衫、游艇的帆等的变形的评价实验。特别是以材质的非伸缩性为限制条件进行利用且将焦点置于只变形的还原。

[0004] 在此,认为现有的变形模型所导入的对象的非伸缩性的意思是,只要对象本来是从平面变形的,就可以展开成平坦的状态。该展开的操作在还原问题上可能具有重要的意义。具体而言,可以利用对所有的观测图像展开的纹理图像只存在一种的条件。这样,利用观测到的图像一并实施对象表面的纹理的还原时,认为在应用和技术上的两个侧面具有大的优点。

[0005] 认为这种技术对于书籍电子化的领域有贡献。近年来,面向书籍电子化的需求在全世界变高。书籍电子化需要高速性和高精度,但兼备其两者的系统尚未实现。至此,通常将书籍按一页固定并依次对相对平坦的状态的纸面进行电子化的样式,存在需要时间的问题。

[0006] 为解决该问题,提倡不在书籍的页面翻开动作中停止动作而进行电子化的书籍翻开扫描(Book Flipping Scanning)(下述非专利文献1)。

[0007] 认为通过该新的方法,从业务用途到个人用途根据各种必要性产生几个系统结构。例如,提出了通过使用三维传感器,实时获取形状,使用该形状的信息校正通过照相机获取的失真的书籍图像的类型(下述非专利文献1)。但是,该系统仅由1台照相机构成,并且每一页仅拍摄一次,因此,在书籍图像的分辨率低这一点及纸面重叠产生死角这一点上有改善的余地。

[0008] 现有技术文献

[0009] 非专利文献

[0010] 非专利文献1:[1]T. Nakashima, Y. Watanabe, T. Komuro, and M. Ishikawa. Book flipping scanning. 22nd Symposium on User Interface Software and Technology (UIST2009) (Victoria, 2009. 10. 5) /Adjunct Proceedings, pp. 79 — 80, 2009.

## 发明内容

[0011] 发明要解决的课题

[0012] 但是,在执行将翻开动作中的书籍连续电子化的系统时,有时因翻开中的某页面的背后隐藏下一页面,或者没有预料的页面失真,因而不能获取适宜的书籍图像。于是,可能不能适当地还原书籍图像。另外,高速的照相机由于分辨率会降低,所以存在电子化的书籍图像的精确度低的问题。

[0013] 即,为了书籍电子化,在这一点上有改善的余地。

[0014] 本发明是鉴于所述的情况而创立的。本发明的目的在于,提供一种为了一边翻开页面一边获取正确的书籍图像而可利用的技术。

[0015] 用于解决课题的方案

[0016] 解决所述的课题的发明如以下项目所记载。

[0017] (项目 1)

[0018] 一种书籍读取系统,其特征在于,

[0019] 具备页面状态测定部、书籍图像获取部和数据处理部,

[0020] 所述页面状态测定部的结构为测定已翻开的书籍的页面状态,

[0021] 所述书籍图像获取部具备多个第一照相机,

[0022] 所述多个第一照相机的结构为从彼此不同的方向分别获取有关所述书籍的页面的书籍图像,

[0023] 所述数据处理部的结构为使用由所述页面状态测定部测定的所述书籍的页面状态的信息校正由所述书籍图像获取部获取的所述书籍图像。

[0024] (项目 2)

[0025] 如项目 1 所述的书籍读取系统,其中,

[0026] 所述页面状态测定部具备用于朝向所述书籍的页面投影既定的图案的图案光源,

[0027] 所述多个第一照相机的结构为进一步获取投影了所述图案的页面的形状图像。

[0028] (项目 3)

[0029] 如项目 2 所述的书籍读取系统,其中,

[0030] 所述多个第一照相机的结构为在从所述图案光源投影所述图案的定时获取所述形状图像且在未从所述图案光源投影所述图案的定时获取所述书籍图像。

[0031] (项目 4)

[0032] 如项目 1 所述的书籍读取系统,其中,

[0033] 所述页面状态测定部具备形状图像用的第二照相机,

[0034] 所述第二照相机具备使从所述图案光源投影的波长的光透过且遮蔽其以外的环境光的滤波器,

[0035] 所述图案光源的结构为使用既定的波长的光朝向所述书籍的页面投影所述既定的图案,

[0036] 由此,所述第二照相机能够获取投影于所述书籍的页面的所述既定的图案。

[0037] (项目 5)

[0038] 如项目 1 ~ 4 中任一项所述的书籍读取系统,其中,

[0039] 所述第一照相机的结构为在所述书籍的每一页面获取多张的书籍图像，

[0040] 所述数据处理部的结构为通过统一所述多张的书籍图像的信息从而提高校正了页面形状的失真的书籍图像的精确度。

[0041] 发明效果

[0042] 根据本发明，获取利用多个第一照相机拍摄印刷于书籍的纸面的字符或绘画的图像(书籍图像)，因此，与单一的照相机的情况相比，能够更高精度地校正书籍图像而将书籍的信息电子化。进而，本发明的其它目的在于，通过统一对变形中的纸面拍摄了多次的书籍图像，从而与仅根据对一个页面拍摄了一次的书籍图像进行校正的情况相比，进行高质量的书籍电子化。

#### 附图说明

[0043] 图 1 是表示本发明的第一实施方式中的书籍读取系统的概略性结构的方框图。

[0044] 图 2 是用于说明本发明第一实施方式中的照相机及图案光源的配置的说明图。

[0045] 图 3 是用于说明本发明第一实施方式中的书籍读取系统的整体的动作的流程图。

[0046] 图 4 是用于说明获取页面初始形状的步骤的流程图。

[0047] 图 5 是用于说明可展面的说明图。

[0048] 图 6 是用于说明可展面的平面展开的说明图。

[0049] 图 7 是用于说明页面位置的估计的说明图。

[0050] 图 8 是用于说明可展面的估计的说明图。

[0051] 图 9 是用于说明还原展开的校正图像的步骤的说明图。

[0052] 图 10 是用于说明页面形状的还原的流程图。

[0053] 图 11 是用于说明还原页面的三维形状的步骤的说明图。

[0054] 图 12 是表示本发明第二实施方式的书籍读取系统的概略性结构的方框图。

#### 具体实施方式

[0055] 下面，参照附图说明本发明第一实施方式的书籍读取系统。

[0056] (第一实施方式的结构)

[0057] 本实施方式的书籍读取系统具备书籍图像获取部 1、页面状态测定部 2、数据处理部 3、控制部 4 (参照图 1)。

[0058] 书籍图像获取部 1 具备多个第一照相机 11 和扩散光源 12。

[0059] 多个第一照相机 11 为从彼此不同的方向分别获取有关书籍的页面的书籍图像的结构。第一照相机 11 在书籍的信息中被阵列状配置。另外，第一照相机 11 朝向书籍配置，可拍摄书籍的各页面。

[0060] 在第一照相机 11 为多台的情况下，优选以能够将书籍的页面分割拍摄的方式进行配置。另外，各照相机的拍摄区域优选为重复。第一照相机 11 为奇数台的情况下，在上述的偶数台照相机的配置的基础上，优选剩余的一台以能够拍摄页面整体的方式进行配置。

[0061] 另外，上述第一照相机优选在每次翻开一张页面时，可以以能够拍摄多张书籍图像的程度进行高速动作。

[0062] 扩散光源 12 通过控制部 4 与第一照相机 11 的动作时期同步进行动作。即，扩散

光源 12 在第一照相机 11 获取书籍图像时接通,容易获取鲜明的书籍图像。扩散光源 12 照射第一照相机 11 应拍摄的书籍。但是,在环境光充足的情况下,可以省略扩散光源 12。

[0063] 页面状态测定部 2 为测定已翻开的书籍的页面状态的结构。更详细而言,本实施方式的页面状态测定部 2 具备图案光源 22。图案光源 22 用于朝向书籍的页面投影既定的图案。在本实施方式中,由多个第一照相机 11 获取照射了来自图案光源 22 的图案的页面的图像(形状图像)。另外,作为图案光源 22,在本实施方式中使用投影仪。更详细而言,多个第一照相机 11 为在从图案光源投影图案的定时(第 f 帧)获取形状图像,在未从图案光源投影图案的定时(第 f+1 帧)获取书籍图像的结构。图案光源 22 可以设置多台。但是,在以下的说明中,假定图案光源 22 为一台。此外,关于形状图像,也如后述,优选以能够按每一页面拍摄多张的方式构成。

[0064] 图 2 表示第一照相机 11 和图案光源 22 的配置例。图 2 是从侧面观测系统和书籍的动作的图。假设书籍被一页一页地翻开。图中表示某瞬间的书籍的样态,但纸面总是变动。在此,将书籍的装订部中心设为原点 0,将水平轴设为 x 轴,将垂直轴设为 y 轴。将到纸面的中心的距离设为 r,根据翻开动作描绘同轴心的半径为 r 的半圆。此时,若将翻开的纸面的旋转角设为  $\theta$ ,则期望将照相机 11 和图案光源设置在成为如下的位置。

[0065]

$$(x, y) = (r \cos \theta \mp D \sin \theta, y = r \sin \theta \pm D \cos \theta)$$

[0066] 在此, D 是书籍的纸面的中心和至设置照相机 / 图案光源的距离,是由各器件(照相机至光源)的画角决定的参数。因此,在同模型描绘的曲线上任意并排照相机和光源。

[0067] 在此,假设所有的照相机和光源的位置关系和内部参数被预先获取。将该获取作业称作校准(calibration)。因此,在本实施方式中可以进行所有的照相机图像之间的坐标变换。

[0068] 数据处理部 3 为使用由页面状态测定部 2 测定的书籍的页面状态的信息校正由书籍图像获取部 1 获取的书籍图像的结构。更具体而言,该实施方式的数据处理部 3 具备页面初始形状获取部 31、校正图像还原部 32、形状还原部 33。有关数据处理部 3 的具体的动作后述。数据处理部 3 例如可以由安装于计算机硬件的计算机程序构成。

[0069] 控制部 4 用于在既定的定时使书籍图像获取部 1 和页面状态测定部 2 动作。关于控制部 4 的书籍图像获取部 1 及页面状态测定部 2 的动作定时后述。另外,控制部 4 例如也可以由安装于计算机硬件的计算机程序构成。

[0070] (第一实施方式的动作)

[0071] 下面,参照图 3 说明本实施方式的书籍读取系统的动作。

[0072] (图 3 的步骤 SA-1 及 2)

[0073] 首先,通过控制部 4 驱动页面状态测定部 2 和书籍图像获取部 1,分别获取书籍的形状图像和书籍图像。

[0074] 更具体而言,例如,在 t 时刻扩散光源发光,同时由多个第一照相机 11 获取书籍图像。接着,在 t+1 时刻图案光源发光,同时由多个第一照相机获取形状图像。通过重复该动作,能够获取多张书籍图像及形状图像。书籍图像和形状图像的拍摄周期可以根据页面翻开速度适宜决定,优选以在从各页面的页面翻开开始到页面翻开终点为止的期间能够获取多个书籍图像和形状图像的方式设定拍摄周期。

[0075] 获取的书籍图像及形状图像依次被送入数据处理部 3。

[0076] (图 3 的步骤 SA-3)

[0077] 下面,通过数据处理部 3 的页面初始形状获取部 31 获取有关形状图像的各拍摄时刻的页面初始形状。进一步参照图 4 说明该处理。

[0078] (图 4 的步骤 SB-1)

[0079] (关于可展面)

[0080] 在此,说明在本实施方式的说明中提及的可展面。可展面是指在作为直线向量  $\xi$  移动时的包络面而描绘的曲面中曲面上的任意点的高斯曲率为 0 的曲面。在此,将描绘可展面的各直线称作母线。另外,将连接母线(rulings)的曲线向量  $\gamma$  称作准线(directrix)(参照图 5)。可展面的一般式以  $X(s, r) = \gamma(s) + r \xi(s)$  记述。 $s$  是曲线  $\gamma$  的曲线长, $r$  表示单位向量  $\xi$  的方向和长度。作为可展面的记述例,有如下的式子。

$$[0081] \quad \mathbf{X}(s, r) = \boldsymbol{\gamma}(s) + r \cdot \boldsymbol{\xi}(s)$$

$$[0082] \quad \boldsymbol{\xi}(s) = \frac{\boldsymbol{\gamma}''(s) \times \boldsymbol{\gamma}'''(s)}{\|\boldsymbol{\gamma}''(s) \times \boldsymbol{\gamma}'''(s)\|}$$

[0083] 该情况下,通过确定准线,能够设定成为可展面的曲面,因此,决定可展面的模型参数被定义为定义准线的参数。也考虑其它定义,但本说明中主要以同模型为前提使用。

[0084] 另外,可展面能够在平面上展开而不会伸缩断裂。在本实施方式中,在估计了表示纸面变形的可展面后,通过进行该平面展开,进行观测图像上的失真的校正。之后,将校正了失真的图像称作校正图像。在此,参照图 6 说明平面展开。在展开前和展开后的坐标系中保存距离。因此,展开的面的坐标表示为  $Q(s, r) = L(s) + r \eta(s)$ 。另外,准线为测地线,因此,在展开后的平面上成为直线  $L$ 。进而,准线  $\gamma$  和母线  $\xi$  间的角度与展开后的直线  $L$  和对应于母线的  $\eta$  相同。使用这些关系,可求取可展面上的各点在展开后所处的坐标。

[0085] (使用形状图像计算三维形状的方法:SB-1)

[0086] 下面说明根据形状图像计算可展面的三维形状的方法。本步骤中,确定形状图像的各点与实际投影的光源图案的哪一位置对应,利用该对应,通过三角测量导出各点的三维位置。

[0087] 例如,形状图像中的点和光源图案中的点的对应如下解决,即,在书籍的纸面变形的范围预先已知的情况下,以在形状图像上能够唯一地确定图案的方式设计图案,同时设置照相机和光源图案。作为简单的例子,对图案为由平行的多个线段构成的多线的图案的情况进行说明。从光源投影的各线由对象物体的表面反射并被照相机捕获。由照相机捕获的线的位置根据反射的场所和照相机的位置变化。在照相机被固定,对象的测量范围预先已知的情况下,各线在照相机上变化的范围可以设为已知。可以以该各线的变化范围不与其它线重复的方式设置照相机和光源。由此,可以唯一地确定在照相机上观测到的线相当于投影的图案的哪一位置。

[0088] 或者,通过利用成为样品的书籍图像的方法,解决形状图像和光源图案的对应。该方法中,在系统设置后预先拍摄因书籍的变形而产生的形状图像的多个事例图像。同事例图像和光源图案的对应手动进行。在观测时,进行输入的形状图像和同事例图像的对照,发现相关最高的图像。之后,以输入的形状图像的各点,在检测出的事例图像的各点附近搜索

并发现以同一点为中心的局部区域的相关最高的位置。将相当于在事例图像中发现的点的图案的位置对应至形状图像的点。

[0089] 在第一实施方式中,通过定时控制获取形状图像和书籍图像,因此,获取这些图像的时刻不同。通过相对于纸面变形的速度充分高速地进行拍摄,考虑该时刻偏差带来的精度的劣化通过下述的方法被补偿。即,在照相机的帧率非常高的情况下,将由形状图像算出的三维点群作为对于连续的時刻的书籍图像的数据进行分配。另一方面,在该假定难以成立的情况下,使用下述方法的任一方法估计书籍图像在被获取的時刻的形状。本方法对各观点的数据适用。

[0090] (方法 1)

[0091] 图 7 表示该方法的概要。假设输入图像的透镜失真已被校正。首先,求通过透视投影的关系式

$$[0092] \quad \omega u_t = P X_t$$

[0093] 所导出的连接照相机的光学中心和图像点的三维直线与算出的時刻  $t-1$  和  $t+1$  的曲面形状的交点  $X_{t-1}$ ,  $X_{t+1}$ 。曲面形状通过根据利用上述的方法算出的三维点群求近似的连续曲面而获取。将两点的中央值设为如下。

$$[0094] \quad X_t(u_t)$$

[0095] 对书籍图像上的多个位置重复相同步骤,计算時刻  $t$  的三维点的集合。

[0096] (方法 2)

[0097] 通过以连续的三个時刻拍摄的图像来获取对应点。为获取对应点,准备三张连续的图像。在此,假设准备了時刻  $t-1$ 、 $t$ 、 $t+1$  的图像。此外,以第  $t-1$ 、 $t+1$  時刻的图像为形状图像,以第  $t$  時刻的图像为书籍图像。

[0098] 首先,在時刻  $t$  的图像内任意设定代表点。其次,切出以同点为中心的局部区域的图像  $I_s$ 。在時刻  $t-1$  的图像内设定以与代表点相同的坐标为中心的附近范围。在同范围内扫描图像  $I_s$ ,在范围内的各点计算切出的图像  $I_s$  和時刻  $t-1$  的图像  $I_{t-1}$  的图像相关。决定相关最高的時刻  $t-1$  的图像上的位置。将同位置设为与時刻  $t$  的代表点对应的位置。对于第  $t+1$  時刻的图像也进行同样的步骤。另外,设定多个代表点,重复相同步骤,获取对应点的集合。此外,假设投影了图案的位置不作为代表点进行设定。

[0099] 接着,求获取的图像点  $u_{t \pm 1}$  的三维位置  $X_{t \pm 1}$ 。计算式设为下述二式。

$$[0100] \quad \omega u_{t \pm 1} = P X_{t \pm 1}, g(X_{t \pm 1}) = 0 \quad \text{在此, } g(X) = 0 \text{ 表示三维点 } X \text{ 为曲面形状上的点。}$$

[0101] 求出時刻  $t-1$ 、 $t+1$  的三维点后,求对应的時刻  $t$  的三维点。因此,求解

[0102] 评价函数  $\min_{X_t} |X_t - X_{t-1}|^2 + |X_t - X_{t+1}|^2$ , 约束条件  $\omega u_t = P X_t$  的附带制约的最小化问题。

[0103] (图 3 的步骤 SB-2)

[0104] 在本实施方式中,第一照相机有多个。因此,使用通过校准得到的坐标变换将由各照相机的坐标系记述的形状与基准坐标系对齐。基准坐标系选择一个主照相机,设定同照相机的坐标系。这样,每時刻使所有的照相机的形状在同一坐标系移动。

[0105] (图 4 的步骤 SB-3)



[0106] 在步骤 SB-2, 叙述以对齐到同一坐标的三维点群为输入的可展面的形状估计。作为可展面的模型, 利用设为如下式的展直可展面。

$$[0107] \quad \mathbf{X}(s, r) = \gamma(s) + r \cdot \xi(s)$$

$$[0108] \quad \xi(s) = \frac{\gamma''(s) \times \gamma'''(s)}{\|\gamma''(s) \times \gamma'''(s)\|}$$

[0109] 在展直可展面中, 以准线为测地线。以下说明具体例。

[0110] 首先, 将输入数据 P (测量的三维点的集合) 应用于暂时的参量连续曲面。作为参量连续曲面, 例如有多项式近似  $z=f(x, y)$  及线性网格等。应用后, 获取母线的集合。图 8 表示方法的概要。

[0111] 首先, 说明求点  $\sim c_i$  上的母线  $\sim \xi_i$  的方法。穿过点  $\sim c_i$  的母线  $\sim \xi_i$  与点  $\sim c_i$  中的法线向量  $\sim n_i$  正交。另外, 母线群与曲面相接。根据这些条件搜索点  $\sim c_i$  中的母线  $\sim \xi_i$ 。

[0112] 其次, 说明在决定了第  $i$  个的母线  $\sim \xi_i$  后, 搜索新的附近点  $\sim c_{i+1}$  的方法。本次, 准线  $\gamma$  为测地线。因此, 需要两点 ( $\sim c_i, \sim c_{i+1}$ ) 间的距离在曲面上为最短。根据该条件, 假定关注的点  $\sim c_i$  的周围区域为平面, 且以连接两点的向量和新的母线  $\sim \xi_{i+1}$  正交的方式选择新的点  $\sim c_{i+1}$ 。

[0113] 通过上述方法, 改变起点  $\sim c_0$ , 获取多个成为解的候选的母线的集合。例如, 在  $xy$  平面 (将三维空间的基底轴称作  $x, y, z$  轴) 上将起点均等地配置。

[0114] 使用在此得到的初始值, 解如下的最小化问题。

$$[0115] \quad \min_{\theta, s} \sum_i |\bar{\xi}_i - \xi(s_i)|^2 + |\bar{c}_i - M(s_i, 0)|^2$$

[0116] 在此,  $\theta$  为可展面的模型参数。另外,  $s$  为决定可展面的准线上的位置的参变量。使用求出的可展面求与测量出的点群分别对应的参变量  $s, r$ 。此外, 这里的参变量  $s$  表示曲线长, 但也可以根据使用的曲线模型使用同模型的参变量。在决定了参变量后, 应用下一最小化问题。

$$[0117] \quad \min_{\theta} \sum_{i=1}^N |P_i - M(s_i, r_i)|^2$$

[0118] 在估计精度不足的情况下, 使用由上述的最小化问题所估计的可展面重新计算与测量点对应的参变量, 之后再次估计可展面。

[0119] 将以上的步骤 SB-3 应用于对齐到各时刻的同一坐标系的三维点群, 从而能够估计第一时刻  $\sim$  第 T 时刻的可展面。

[0120] (图 4 的步骤 SB-4)

[0121] 根据  $T \times N_2$  张书籍图像生成多个在图像中观测曲面上的相同范围的两张图像的对。根据取出的两张图像检测同一点的位置, 作为对应点保存。在此, T 为拍摄时间数,  $N_2$  为第一照相机的台数。

[0122] 叙述对应点的检测方法。首先, 从两张图像提取对应点。对应点的检测使用对照局部区域的图像特征量的方式 (Scale - Invariant Feature Transform (参照 D. G. Lowe. Object recognition from local scale - invariant features. In Proceedings of the

International Conference on Computer Vision, pages 1150 – 1157, 1999.) 等)。另外, 在两张图像的变动小的情况下, 除去对应点间的距离大的对。但是, 在此得到的对应点的对也包含错误的对。

[0123] 为除去该脱离值, 应用如下步骤。使用 RANSAC 法(参照 Martin A. Fischler and Robert C. Bolles(June 1981). 'Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography'. Comm. of the ACM 24: 381 – 395. ), 估计两张图像的对应点间的对照率最高的映射变换。该映射变换可利用单应变换或二维仿射变换等。使用该映射变换使两张图像粗略对齐。

[0124] 接着, 将粗略对齐的两张图像补片分割。补片分割是指将图像分割为小的区划的操作。将各区划称作补片。在两张图像中, 补片的大小相同。

[0125] 在分割后, 对各补片应用如下步骤。在两张图像间获取相同位置的补片内的对应点对。针对该对应点对, 与上述同样地使用 RANSAC 法, 从两个补片间的对应点对除去脱离值。对所有的补片重复该步骤。

[0126] 与在图像整体假定了映射变换的情况相比, 仅在补片内假定的情况由于该假定近似性以更高的精度满足, 所以认为脱离值的检验有效发挥作用。结束以上的步骤后, 将由各补片获取的对应点的集合作为特征点使用。

[0127] (图 4 的步骤 SB-5)

[0128] 在此, 求在步骤 SB-4 检测到的图像上的各对应点在可展面上的位置。这可作为求连接各照相机的光轴中心和图像点的三维直线与可展面的交点的问题进行计算。接着, 根据计算出的可展面上的三维点求展开到平面上时的同点的坐标。这使用在上述的可展面的概要中叙述的方法。接着, 决定展开的坐标系上的基准坐标系, 求向同坐标系的变换。通过以上的步骤, 求与图像上的点  $q$  对应的可展面上的三维点  $X$  和展开面上的点  $u$ 、及用于在展开各时刻的可展面后与基准的展开面的坐标系对齐的变换  $\psi$ 。

[0129] (图 4 的步骤 SB-6)

[0130] 在步骤 SB-4 检测到的两张书籍图像间的对应点在通过可展面的展开而生成的校正图像上成为相同位置。因此, 通过对对应点解如下的最佳化问题, 从而再次估计  $T$  时刻量的可展面。该处理是统一以多个照相机视点拍摄了多次的观测信息从而估计各时刻的可展面的处理, 因此, 相当于本发明的“统一多张的书籍图像的信息的处理”的一例。

$$[0131] \quad \min_{\{\Theta_k\}_{k=1}^T, \{\eta_i\}_{i=1}^N} Q + U_1 + U_2 \quad (1)$$

$$[0132] \quad Q = \sum_i \|q_i - \tilde{q}_i(\Theta_{k_i}, \eta_i)\|^2 \quad (2)$$

$$[0133] \quad U_1 = \sum_j \|\tilde{u}_{i_j,1}(\Theta_{k_j,1}, \eta_{i_j,1}) - \tilde{u}_{i_j,2}(\Theta_{k_j,2}, \eta_{i_j,2})\|^2 \quad (3)$$

$$[0134] \quad U_2 = \sum_c \|\tilde{u}_{i_c}(\Theta_{k_c}, \eta_{i_c}) - u_{i_c}\|^2 \quad (4)$$

$$[0135] \quad \Theta_k = \begin{bmatrix} \theta_k \\ \psi_k \end{bmatrix} \quad (5)$$

$$[0136] \quad \eta_i = [s, r] \quad (6)$$

[0137] 这些式中,  $\Theta_k$  是表示时刻  $k$  的可展面的参数。同参数如上述式(5), 由表示可展面的变形的参数  $\theta$  和对齐到展开后的基准坐标系的变换参数  $\psi$  这两个构成。该参数准备  $T$  时刻量。另外,  $\eta_i$  是表示可展面上的位置的参数。该参数准备相应于在步骤 SB-4 中获取的所有点的数目。在上述式(1)中, 对在步骤 SB-4 中获取的所有点求使评价函数最小化的上述参数。

[0138] 上述式(1)的第一项如上述式(2)所示。该项为了使所估计的可展面与观测到的书籍图像匹配, 计算观测图像上的点和对应于该点的估计图像点的位置之间的误差。 $q_i$  表示图像点,  $\sim q_i$  表示由参数  $\Theta$  和  $\eta$  算出的“估计的图像点”。该计算通过在由参数  $\Theta$  和  $\eta$  计算出可展面上的三维位置后在对应的视点的照相机面上投影同一位置而进行。此外,  $\Theta_{ki}$  表示与第  $i$  个的图像点的时刻相当的可展面参数。

[0139] 上述式(1)的第二项如上述式(3)所示。该项对在步骤 SB-4 中作为书籍纸面上的同一点检测的对应点的对进行计算。这些点在展开后成为相同位置。在上述式(3)中, 计算这些点在展开坐标上的位置误差。第  $j$  个对应点的展开坐标的对由  $\sim u_{ij,1}$  和  $\sim u_{ij,2}$  表示。将相当于对的第一个的点的编号由  $i_{j,1}$  表示, 将相当于第二个的点的编号由  $i_{j,2}$  表示。在此的展开坐标使用估计参数  $\Theta$  和  $\eta$  计算。有关与对的各点对应的估计参数的分配的下标也以与上述相同的规则记载。

[0140] 上述式(1)的第三项如上述式(4)所示。该项是针对校正后的坐标位置已知这一点进行计算的项。例如, 就纸面的四角的角等而言, 如果纸面的尺寸已知, 则可以确定校正后的坐标。因此, 同项中计算使用估计参数  $\Theta$  和  $\eta$  计算的坐标  $\sim u_{ic}$  和已知的目标坐标  $u_{ic}$  的位置的误差。此外,  $i_c$  表示与第  $c$  个已知点相当的点的编号。

[0141] (图 3 的步骤 SA-4)

[0142] 接着, 重复图 3 的步骤 SA-4 和 SA-5 直至收敛。这些处理是统一多张书籍图像而生成一张校正了的书籍图像的处理, 因此, 与本发明中的“统一多张书籍图像的信息的处理”的一例相当。

[0143] 在步骤 SA-4 中, 通过数据处理部 3 的校正图像还原部 32 还原展开的校正图像。图 9 表示该处理。即, 使用  $T$  时刻量的可展面和  $T$  时刻量的书籍图像来解图 9 的式, 由此可以得到展开的校正图像(校正了书籍图像后的图像)。

[0144] 图 9 的式中,  $I(q)$  表示坐标  $q$  的亮度值。 $I_k^j$  表示以时刻  $k$  的第  $j$  个视点观测的图像。 $\sim I(q)$  表示根据参数  $\Theta$  和校正图像  $J$  计算的估计观测图像。该计算由下式表示。

$$[0145] \quad \bar{I}_k(q) = \sum_{u_i, \|q-g(u_i)\| < \epsilon} B_i J(u_i) \quad (7)$$

[0146] 同式的  $g(u)$  表示校正图像和观测图像的坐标变换。这根据可展面的形状而成为不同的变换。该变换函数表示如下。

$$[0147] \quad \mathbf{q} = \mathbf{g}(\mathbf{u}; \Theta) \quad (8)$$

[0148] 该式根据可展面的展开和将可展面上的点投影于观测照相机面上的两个操作计算。在进行了坐标系的变换后,基于透镜模糊模型计算位于观测照相机的各点的附近的校正图像的亮度值的加权总和,计算估计观测图像 $\sim I$ 。

[0149] 图 9 的  $f(J)$  表示有关校正图像  $J$  的限制条件。预先了解有关  $J$  的信息的情况下,追加  $f(J)$ 。

[0150] (图 3 的步骤 SA-5)

[0151] 接着,通过数据处理部 3 的形状还原部 33 还原书籍图像的形状。该步骤可通过将在步骤 SA-4 中估计了校正图像  $J$  的问题切换为估计可展面参数  $\Theta$  的问题来实现。这由下式表示。

$$[0152] \quad \min_{\Theta_k} \sum_{j=1}^{N_2} \sum_{\mathbf{q}} \|\mathbf{I}_k^j(\mathbf{q}) - \tilde{\mathbf{I}}_k^j(\mathbf{q}; \Theta_k, J)\|^2 \quad (9)$$

[0153] 但是,该最小化问题在解的搜索中需要时间。在短时间内实施数据处理的情况下,应用以后的还原步骤。参照图 10 说明该还原的步骤。

[0154] (图 10 的步骤 SC-1)

[0155] 首先,使用估计出的变形参数  $\Theta_k$  生成估计拍摄图像 $\sim I_k$ 。该图像的生成方法与在步骤 SA-4 中所叙述的相同。

[0156] (图 10 的步骤 SC-2)

[0157] 其次,在输入图像  $I_k$  和生成的估计拍摄图像 $\sim I_k$  之间获取特征点的对应。特征点的对应使用与在步骤 SB-4 中所叙述的相同的方法。

[0158] (图 10 的步骤 SC-3)

[0159] 本实施方式的方法中,不改变校正图像,而搜索使照相机图像之间的对应点之间的位置的误差最小化的可展面参数,还原正确的可展面。

[0160] 图 11 表示正确的可展面的还原的概念图。下式表示将误差最小化且还原可展面的问题。该问题在各时刻求解。

$$[0161] \quad \min_{\Theta_k, \{\eta_k^{j,i}\}_{i=1}^{N_c}} \sum_{j=1}^{N_2} \sum_{i=1}^{N_c} \|\mathbf{V}_k^{j,i} - \tilde{\mathbf{V}}_k^{j,i}(\Theta_k, \eta_k^{j,i})\|^2$$

[0162] 下式表示该式的  $\mathbf{V}_k^{j,i}$ 。

$$[0163] \quad \mathbf{V}_k^{j,i} = \begin{bmatrix} \mathbf{q}_k^{j,i} \\ \mathbf{u}_k^{j,i} \end{bmatrix} \quad (10)$$

[0164] 这是针对时刻  $k$  的第  $j$  个视点的第  $i$  个关注点(在 SC-2 中作为对应点获取的点),并拍了其在该照相机上的坐标  $\mathbf{q}$  和在校正图像上的坐标  $\mathbf{u}$ 。照相机上的坐标在步骤 SC-2 获取。校正图像上的坐标在步骤 SC-2 中发现了对应点之后使用在该时点的对应的校正图像上的位置。该坐标在步骤 SC-3 的估计中为常数而不变化。 $\sim \mathbf{V}_k^{j,i}$  表示使用可展面参数  $\Theta$  和表示可展面上的位置的参变量  $\eta_k^{j,i}$  计算的观测图像上的位置和校正图像的位置。此外,

如在步骤 SB-6 中也叙述的那样,针对关注点中预先知晓展开后的坐标的关注点,作为  $V_k^{j,i}$  的校正图像的坐标设定已知的值。

[0165] 通过将把这些数据的误差最小化,可在满足校正图像为唯一的条件的同时,还原可展现。

[0166] (第二实施方式)

[0167] 下面,参照图 12 说明本发明的第二实施方式的系统。此外,在第二实施方式的说明中,对于与上述的第一实施方式的说明基本上共通的要素使用同一符号,由此避免说明的繁杂。

[0168] 在上述的第一实施方式中,由第一照相机 11 一同拍摄了书籍图像和形状图像。但是,第二实施方式的第一照相机 11 仅拍摄书籍图像。

[0169] 另外,第二实施方式的页面状态测定部 2 具备形状图像用的第二照相机 21。第二照相机 21 具备使从图案光源投影的波长的光透过且遮蔽其以外的环境光的滤波器(未图示)。

[0170] 第二实施方式的图案光源 22 为使用既定波长的光朝向书籍的页面投影既定的图案的结构。由此,第二照相机 21 可将投影于书籍的页面的既定的图案作为形状图像获取。

[0171] 在第二实施方式中,第一照相机 11 在扩散光源 12 发光的定时动作,可以获取书籍图像。另外,第二照相机 21 在图案光源 22 发光的定时动作,可经由滤波器获取形状图像。在此,在该第二实施方式中,可通过波长区分来自扩散光源 12 的光和来自图案光源 22 的光,因此,通过使扩散光源 12 和图案光源 22 同时发光且使第一照相机 11 和第二照相机 21 同时动作,可以同时获取书籍图像和形状图像。由此,在该实施方式中,具有可以增加能够在单位时间内获取的书籍图像和形状图像的张数的优点。

[0172] 另外,第二实施方式的书籍图像获取部 1 的第一照相机 11 优选具备遮蔽图案光源 22 的波长的光且使来自扩散光源 12 及环境光的波长的光透过的滤波器(未图示)。通过该构成,能够防止图案照到书籍图像。但是,考虑也可以不在书籍图像获取部 1 的第一照相机 11 设置滤波器而在获取书籍图像后通过图像处理除去图案。该情况下,也可以省略滤波器的设置。此外,第二实施方式中,由于可通过波长来区分来自扩散光源 12 的光和来自图案光源 22 的光,所以还能够使各光源总是接通。

[0173] 第二实施方式中的其它结构及动作与上述的第一实施方式基本上相同,因此,省略其以上的详细说明。但是,在步骤 SB-1 中叙述的形状图像和书籍图像的时刻偏移的削除步骤在本实施方式中不进行。

[0174] 此外,本发明不限于上述的实施方式,在不脱离本发明的宗旨的范围内可以增加各种变更。

[0175] 例如,上述的各构成要素只要作为功能块存在即可,也可以作为独立的硬件存在。另外,作为安装方法,可以使用硬件也可以使用计算机软件。另外,本发明的一个功能要素也可以通过多个功能要素的集合实现,本发明的多个功能要素也可以通过一个功能要素实现。

[0176] 另外,构成本发明的各功能要素也可以分开存在。在分开存在的情况下,例如可以经由网络交接所需的数据。各部的内部的各功能同样也可以分开存在。例如,也可以使用网络计算或云计算实现本实施方式中的各功能要素或其一部分。

- [0177] 标号说明
- [0178] 1 书籍图像获取部
- [0179] 11 第一照相机
- [0180] 12 扩散光源
- [0181] 2 页面状态测定部
- [0182] 21 第二照相机
- [0183] 22 图案光源
- [0184] 3 数据处理部
- [0185] 31 页面初始形状获取部
- [0186] 32 校正图像还原部
- [0187] 33 形状还原部
- [0188] 4 控制部

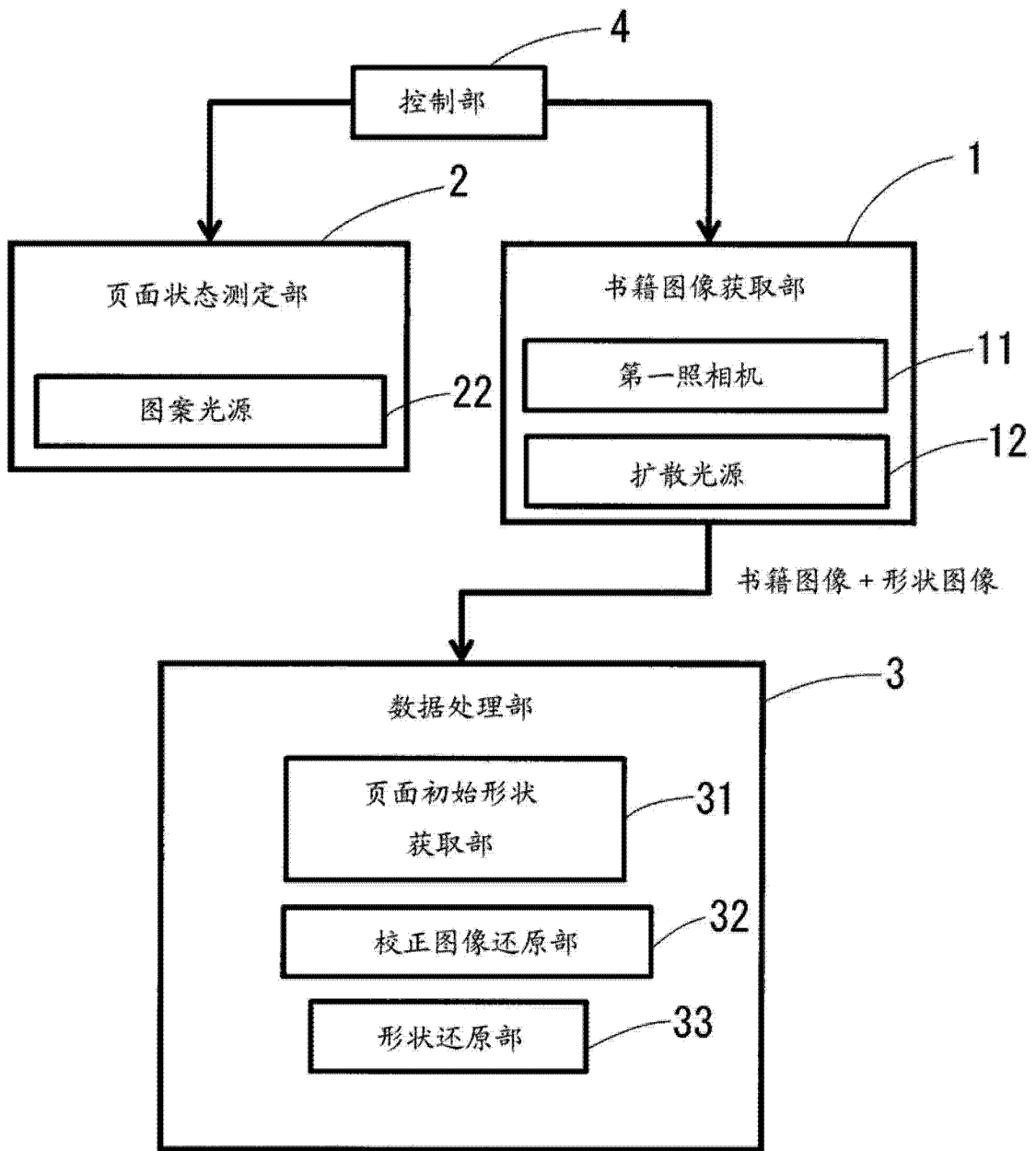


图 1

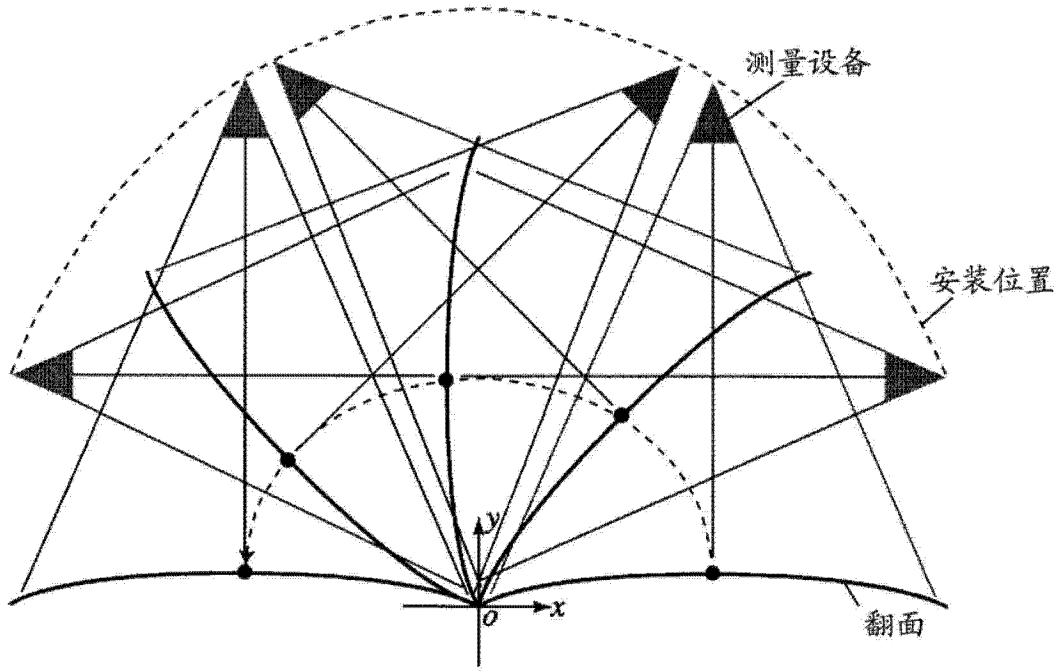


图 2

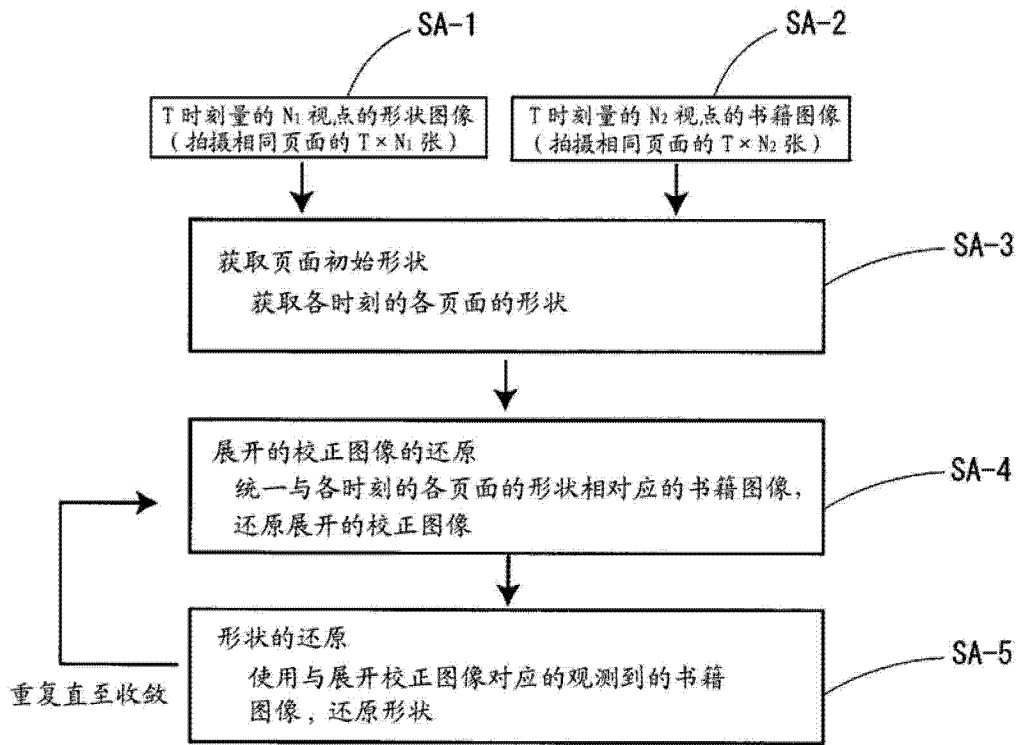


图 3



获取页面初始形状

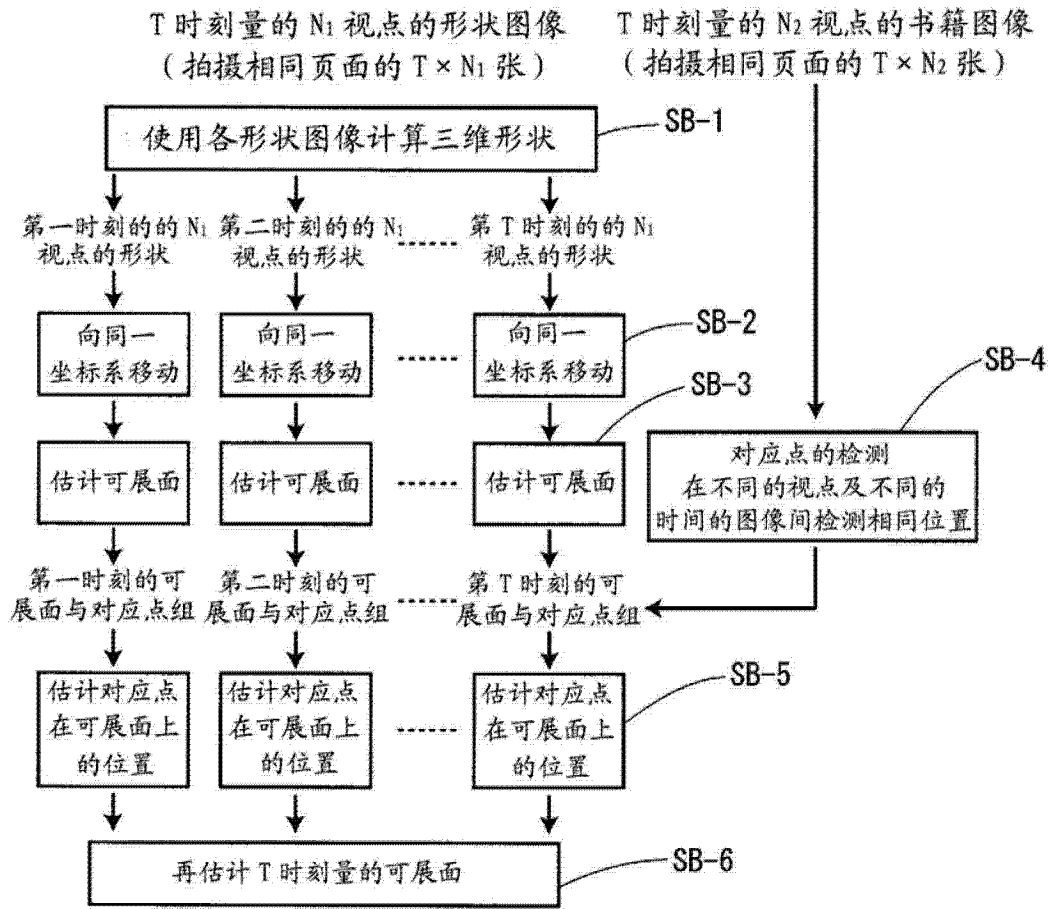


图 4

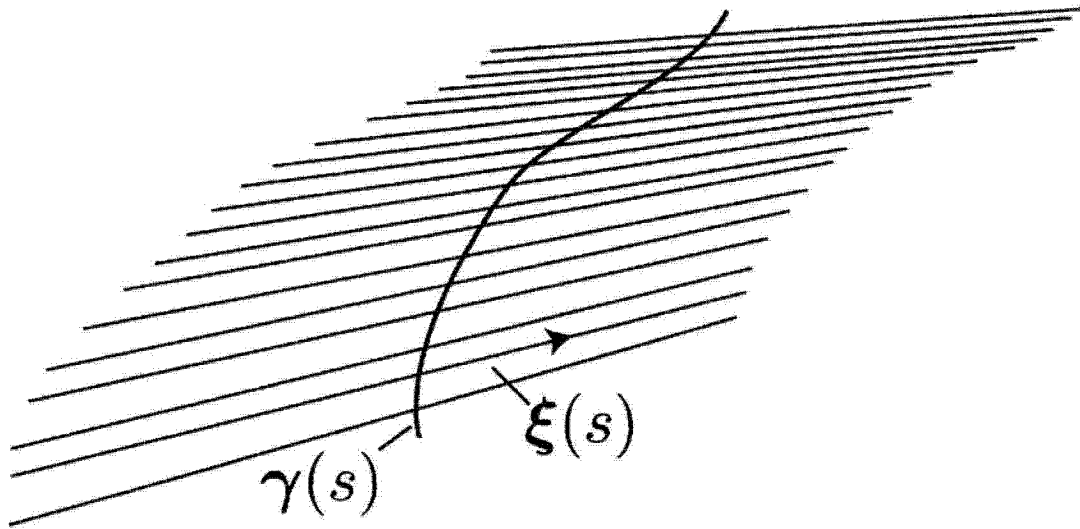


图 5

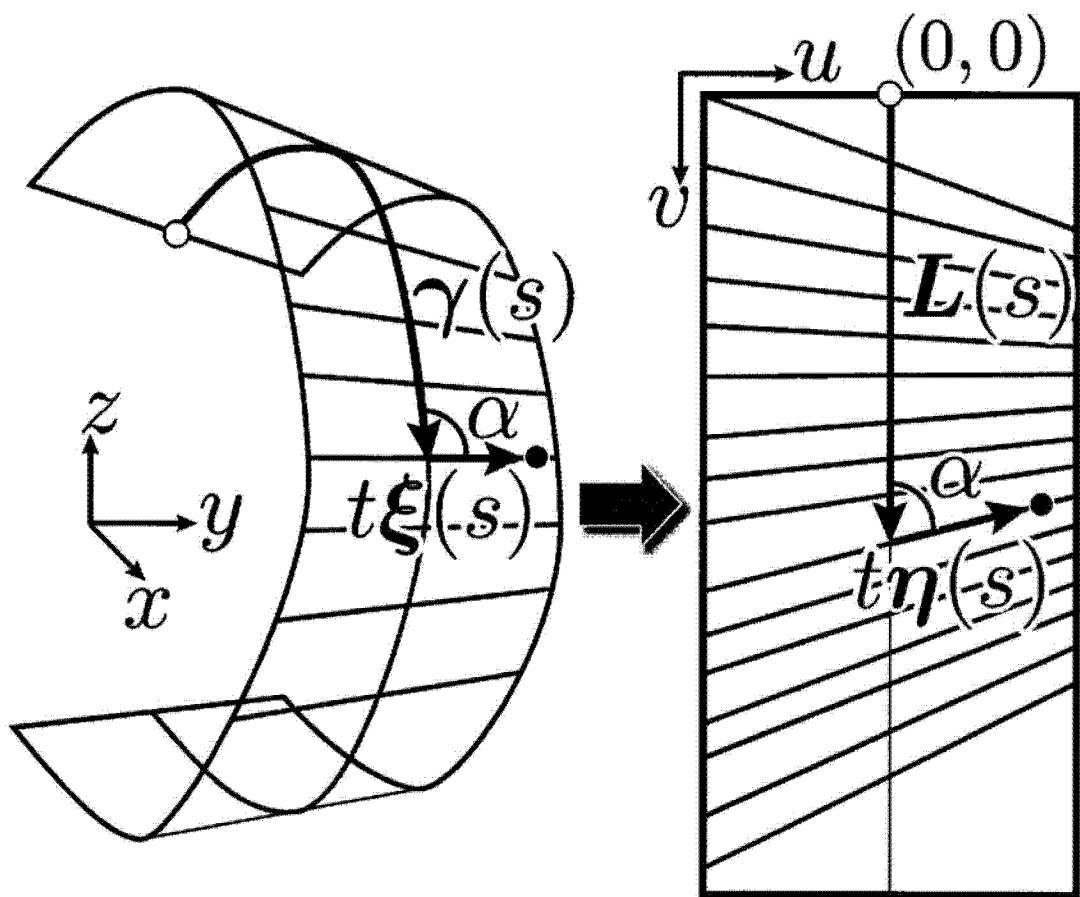


图 6

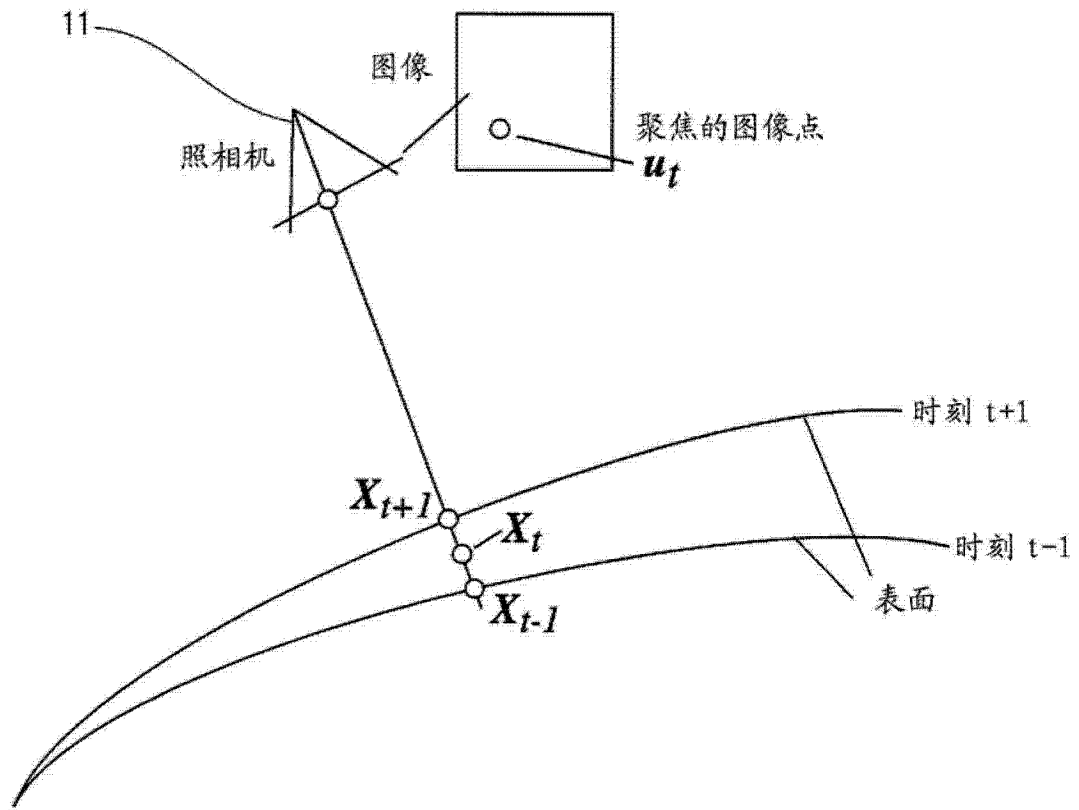


图 7

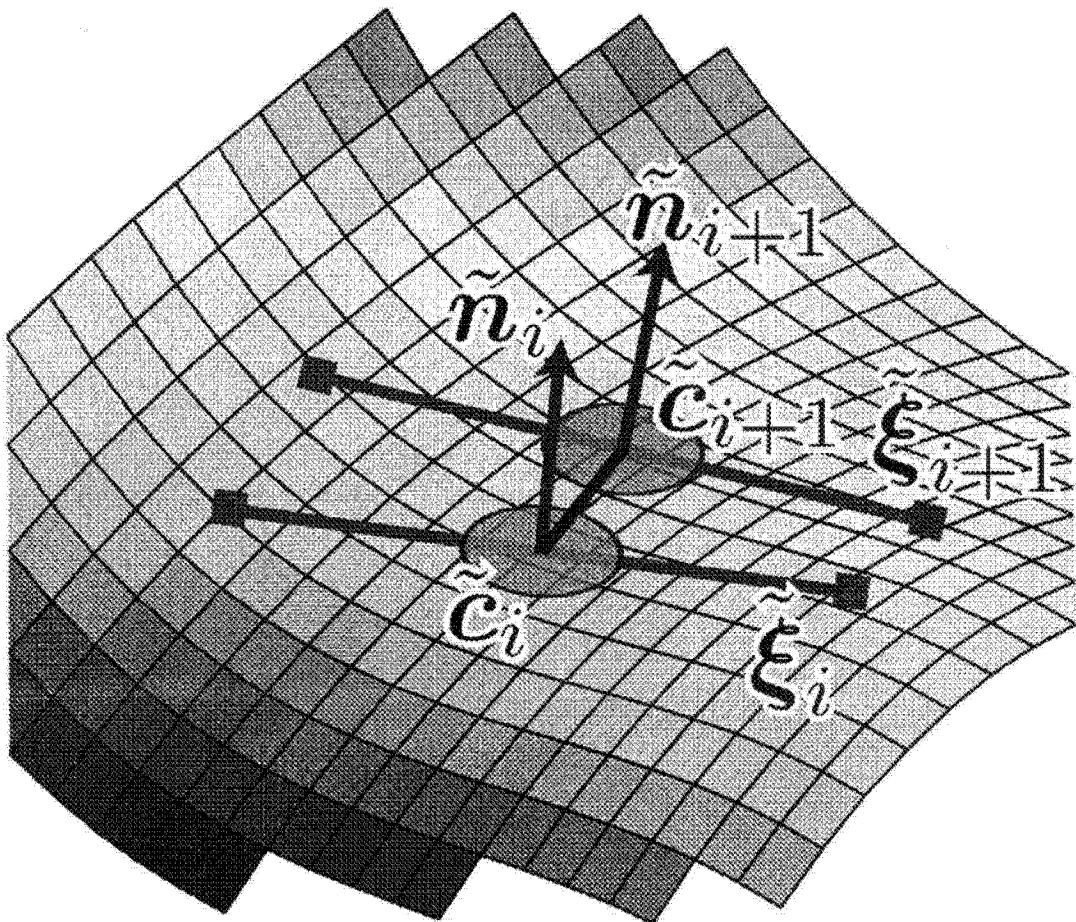


图 8

展开的校正图像的还原

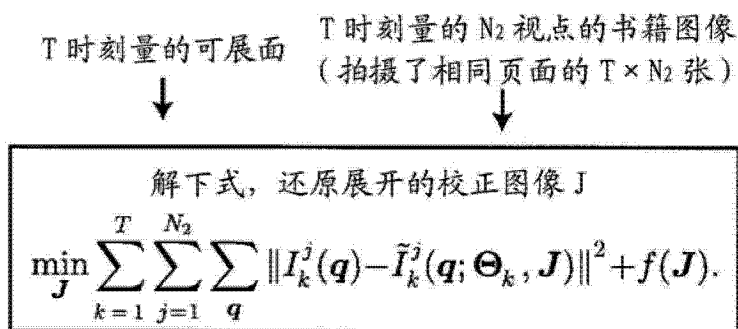


图 9

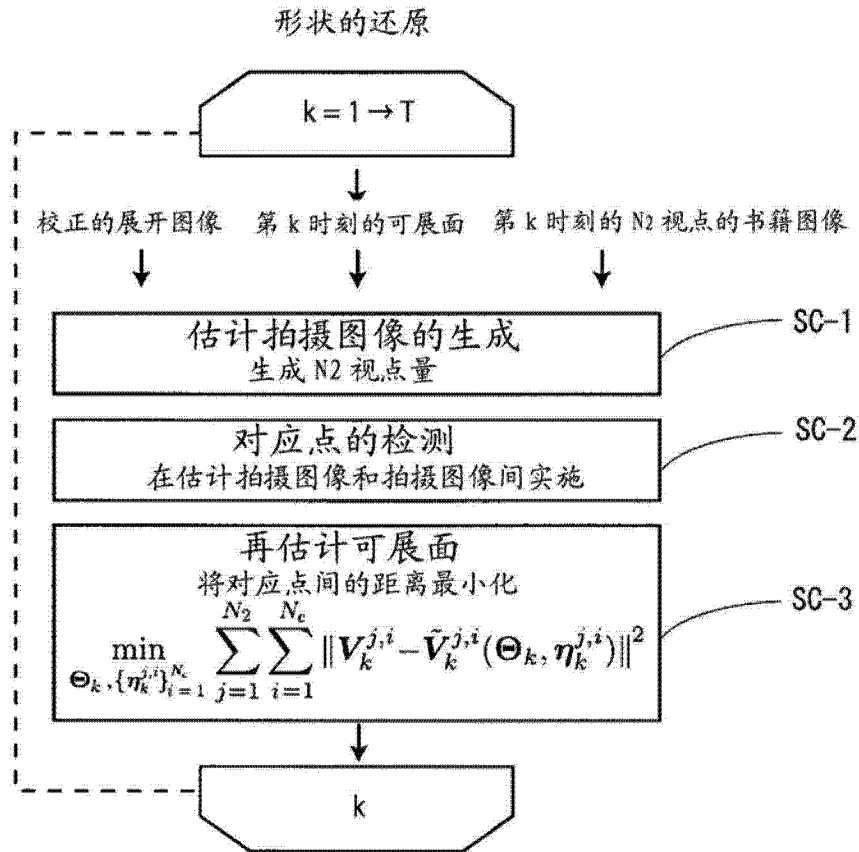


图 10

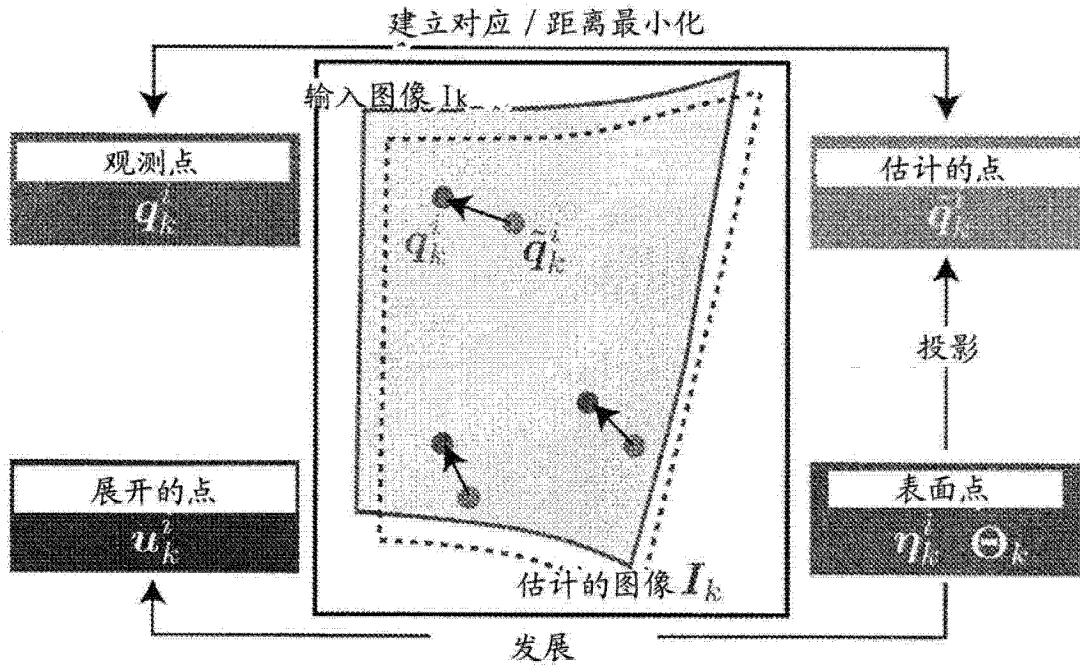


图 11

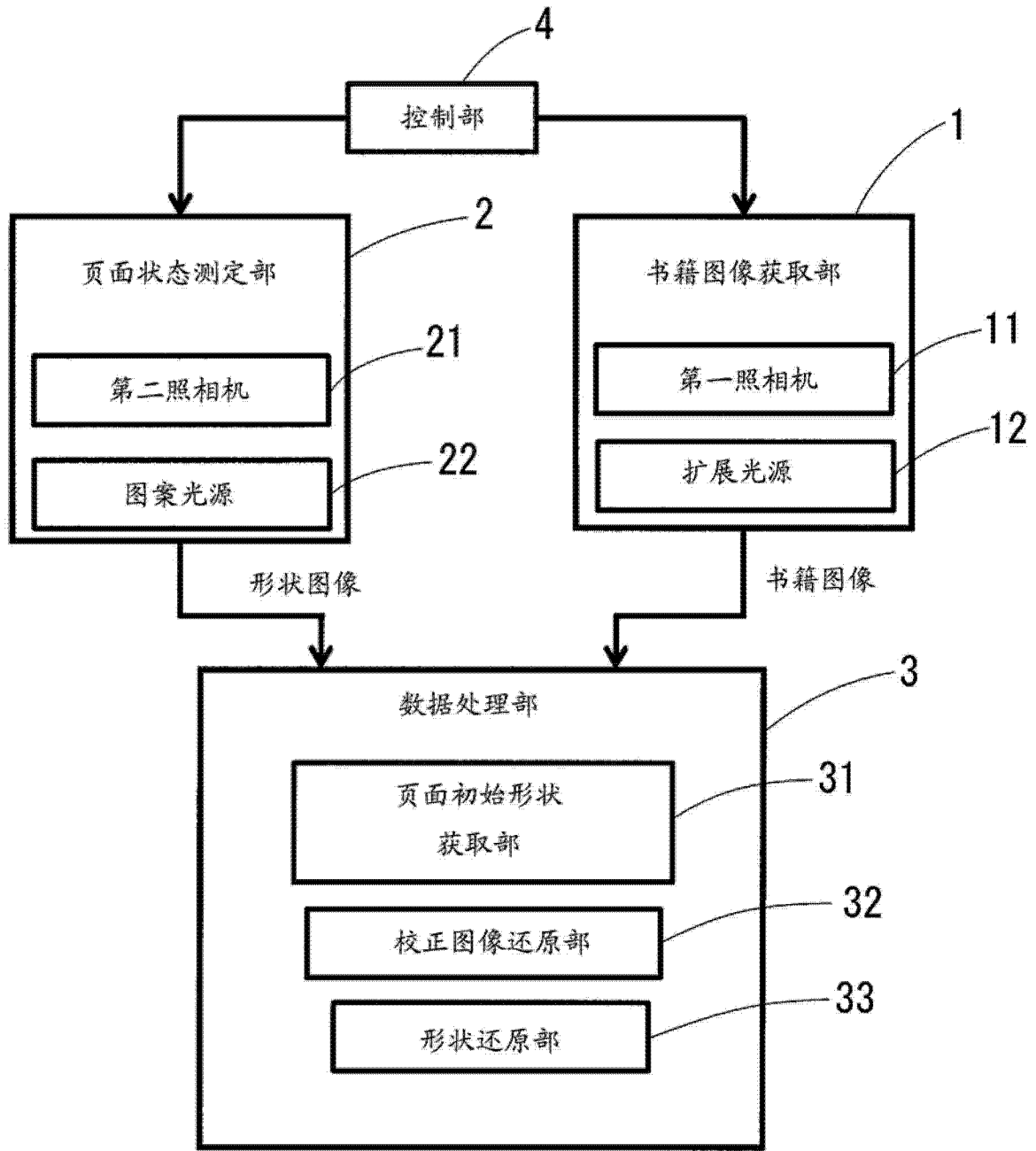


图 12