



(12) 发明专利

(10) 授权公告号 CN 109104551 B

(45) 授权公告日 2021.06.25

(21) 申请号 201810619113.8

(51) Int.CI.

(22) 申请日 2018.06.15

H04N 5/225 (2006.01)

(65) 同一申请的已公布的文献号

H04N 5/232 (2006.01)

申请公布号 CN 109104551 A

H04N 5/265 (2006.01)

(43) 申请公布日 2018.12.28

(56) 对比文件

(30) 优先权数据

CN 102982515 A, 2013.03.20

17176772.6 2017.06.20 EP

CN 106385536 A, 2017.02.08

(73) 专利权人 安讯士有限公司

CN 103997609 A, 2014.08.20

地址 瑞典浪德

US 2013147910 A1, 2013.06.13

(72) 发明人 T·亨宁松 袁嵩

CN 104754228 A, 2015.07.01

(74) 专利代理机构 北京德琦知识产权代理有限公司

US 2008024614 A1, 2008.01.31

11018

审查员 王田

代理人 康泉 宋志强

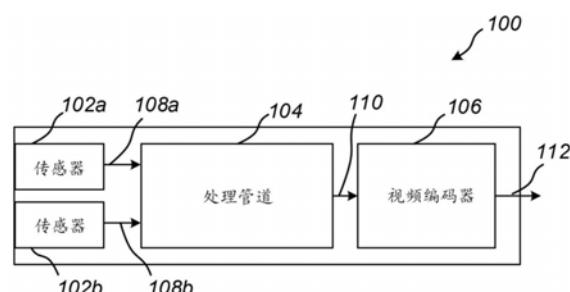
权利要求书4页 说明书11页 附图4页

(54) 发明名称

多传感器摄像机、用于多传感器摄像机的方法和处理管道

(57) 摘要

本公开提供了多传感器摄像机、用于多传感器摄像机的方法和处理管道。提供了一种在包括具有部分重叠视野的第一传感器和第二传感器的多传感器摄像机中执行的方法。第一视频帧和第二视频帧各自具有非重叠部分和重叠部分。通过将来自第一和第二视频帧的非重叠部分的图像数据与仅来自第一视频帧的重叠部分的图像数据结合来生成第一视频流的帧；生成第二视频流的帧以包括来自至少第二视频帧的重叠部分的图像数据。并行处理第一视频流的帧和第二视频流的帧。将重叠添加到第一视频流中的处理后的帧。在准备重叠的步骤和添加重叠的步骤中的至少一个步骤中，将来自第一视频帧的重叠部分的图像数据与来自第二视频帧的重叠部分的图像数据融合。



1. 一种在包括具有部分重叠视野的第一传感器(102a)和第二传感器(102b)的多传感器摄像机(100)中执行的方法,包括:

接收(S02)分别通过所述第一传感器(102a)和所述第二传感器(102b)同时捕获到的第一视频帧(108a)和第二视频帧(108b),

其中所述第一视频帧(108a)和所述第二视频帧(108b)中的每个具有不与所述第一视频帧(108a)和所述第二视频帧(108b)中的另一个重叠的非重叠部分(118a、118b)、以及与所述第一视频帧(108a)和所述第二视频帧(108b)中的另一个重叠的重叠部分(128a、128b),

通过将来自所述第一视频帧(108a)和所述第二视频帧(108b)的所述非重叠部分(118a、118b)的图像数据(A、B)与仅来自所述第一视频帧(108a)的所述重叠部分(128a)的图像数据(A')结合在一起,生成(S04)第一视频流(214)的帧(302),

生成(S06)第二视频流(216)的帧(304、404),其中所述第二视频流的所述帧(304、404)包括来自至少所述第二视频帧(108b)的所述重叠部分(128b)的图像数据(B'),

并行处理(S08)所述第一视频流的所述帧(302)和所述第二视频流的所述帧(304、404),其中所述第二视频流的所述帧(304、404)的处理包括基于来自至少所述第二视频帧(108b)的所述重叠部分(128b)的所述图像数据(B')来准备重叠(306、406),以及

在所述第一视频流中的处理后的帧(308)的与所述第一视频帧(108a)的所述重叠部分(128a)相对应的部分(328a)处,将所述重叠(306、406)添加(S10)到所述第一视频流中的所述处理后的帧(308),

其中所述第二视频流的所述帧(304)包括来自所述第二视频帧(108b)和所述第一视频帧(108a)的所述重叠部分(128a、128b)的图像数据(B'、A'),准备所述重叠的步骤包括将来自所述第二视频帧(108b)的所述重叠部分(128b)的图像数据(B')与来自所述第一视频帧(108a)的所述重叠部分(128a)的图像数据(A')融合,并且在添加所述重叠的步骤中,所述重叠(306)替换所述第一视频流中的所述处理后的帧(308)的与所述第一视频帧(108a)的所述重叠部分(128a)相对应的部分(328a)中的图像数据(A'),

或者

其中所述重叠(406)包括仅来自所述第二视频帧(108b)的所述重叠部分(128b)的图像数据(B'),添加所述重叠的步骤包括:在所述第一视频流(108a)中的所述处理后的帧(308)的与所述第一视频帧(108a)的所述重叠部分(128a)相对应的部分(328a)处,根据掩码(408)将所述重叠(406)与所述第一视频流(108a)中的所述处理后的帧(308)融合,其中所述掩码(408)为所述重叠(406)中的每个像素定义融合度,并且所述掩码(408)在准备所述重叠的步骤中被计算。

2. 根据权利要求1所述的方法,其中在准备所述重叠的步骤包括将来自所述第二视频帧(108b)的所述重叠部分(128b)的图像数据(B')与来自所述第一视频帧(108a)的所述重叠部分(128a)的图像数据(A')融合的情况下,所述融合包括:针对所述重叠部分(128a、128b)中的每个像素,计算来自所述第一视频帧(108a)的图像数据(A')和来自所述第二视频帧(108b)的图像数据(B')的加权平均值,

其中与来自所述第二视频帧(108b)的所述重叠部分(128b)的所述图像数据(B')相关的权重取决于从所述像素到所述第二视频帧(108b)的所述重叠部分(128b)和所述非重

叠部分(118b)之间的边界的最短距离,并且

其中与来自所述第一视频帧(108a)的所述重叠部分(128a)的所述图像数据(A')相关联的权重取决于从所述像素到所述第一视频帧(108a)的所述重叠部分(128a)和所述非重叠部分(118a)之间的边界的最短距离。

3.根据权利要求1或2所述的方法,其中在准备所述重叠的步骤包括将来自所述第二视频帧(108b)的所述重叠部分(128b)的图像数据(B')与来自所述第一视频帧(108a)的所述重叠部分(128a)的图像数据(A')融合的情况下,所述融合包括:

计算来自所述第一视频帧(108a)的所述重叠部分(128a)的图像数据(A')和来自所述第二视频帧(108b)的所述重叠部分(128b)图像数据(B')的第一加权平均值,其中所述第一加权平均值是关于低于第一阈值的空间频率来计算的,

计算来自所述第一视频帧(108a)的所述重叠部分(128a)的图像数据(A')和来自所述第二视频帧(108b)的所述重叠部分(128b)的图像数据(B')的第二加权平均值,其中所述第二加权平均值是关于等于或高于第二阈值的空间频率来计算的,所述第二阈值大于或等于所述第一阈值,

其中,为所述第一加权平均值和所述第二加权平均值不同地选择权重,并且

通过将所述第一加权平均值与所述第二加权平均值相加来计算所述重叠。

4.根据权利要求1或2所述的方法,其中在准备所述重叠的步骤包括将来自所述第二视频帧(108b)的所述重叠部分(128b)的图像数据(B')与来自所述第一视频帧(108a)的所述重叠部分(128a)的图像数据(A')融合的情况下,所述融合包括:

识别所述第一视频帧(108a)的所述重叠部分(128a)中的对象和所述第二视频帧(108b)的所述重叠部分(128b)中的对象,

计算所述重叠部分(128a、128b)中来自所述第一视频帧(108a)的图像数据(A')和来自所述第二视频帧(108b)的图像数据(B')的加权平均值,

其中,如果对象被识别为位于所述第一视频帧(108a)的所述重叠部分(128a)和所述第二视频帧(108b)的所述重叠部分(128b)二者中,则对于被识别为属于所述对象的像素,将与所述第一视频帧(108a)的所述图像数据(A')和所述第二视频帧(108b)的所述图像数据(B')中的一个相关联的权重设置为0。

5.根据权利要求1所述的方法,其中在添加所述重叠的步骤包括在所述第一视频流(108a)中的所述处理后的帧(308)的与所述第一视频帧(108a)的所述重叠部分(128a)相对应的部分(328a)处,根据掩码(408)将所述重叠(406)与所述第一视频流(108a)中的所述处理后的帧(308)融合的情况下,所述重叠(406)中的每个像素与所述第二视频帧(108b)的所述重叠部分(128b)的像素相对应,并且其中所述融合度取决于从所述像素到所述第二视频帧(108b)的所述重叠部分(128b)和所述非重叠部分(118b)之间的边界的最短距离。

6.根据权利要求1所述的方法,其中在添加所述重叠的步骤包括在所述第一视频流(108a)中的所述处理后的帧(308)的与所述第一视频帧(108a)的所述重叠部分(128a)相对应的部分(328a)处,根据掩码(408)将所述重叠(406)与所述第一视频流(108a)中的所述处理后的帧(308)融合的情况下,所述第二视频流的所述帧(404)包括来自所述第二视频帧(108b)和所述第一视频帧(108a)的所述重叠部分(128a、128b)的图像数据(A'、B'),并且其中准备所述重叠的步骤进一步包括:基于来自所述第二视频帧(108b)和所述第一视频帧

(108a) 的所述重叠部分 (128a、128b) 的所述图像数据 (A'、B') 来计算所述掩码 (408)。

7. 根据权利要求6所述的方法,其中以低于从所述第一传感器 (102a) 和第二传感器 (102b) 接收视频帧的速率的速率来更新所述掩码 (408)。

8. 根据权利要求1所述的方法,其中所述第一视频流的所述帧 (302) 的处理和所述第二视频流的所述帧 (304、404) 的处理包括:将所述第一视频帧 (108a) 的图像数据与所述第二视频帧 (108b) 的图像数据对准。

9. 一种用于包括具有部分重叠视野的第一传感器 (102a) 和第二传感器 (102b) 的多传感器摄像机 (100) 的处理管道 (104),包括:

接收器 (202),配置为接收分别通过所述第一传感器 (102a) 和所述第二传感器 (102b) 同时捕获到的第一视频帧 (108a) 和第二视频帧 (108b),

其中所述第一视频帧 (108a) 和所述第二视频帧 (108b) 中的每个具有不与所述第一视频帧 (108a) 和所述第二视频帧 (108b) 中的另一个重叠的非重叠部分 (118a、118b)、以及与所述第一视频帧 (108a) 和所述第二视频帧 (108b) 中的另一个重叠的重叠部分 (128a、128b);

帧生成组件 (206),配置为通过将来自所述第一视频帧 (108a) 和所述第二视频帧 (108b) 的所述非重叠部分 (118a、118b) 的图像数据 (A、B) 与仅来自所述第一视频帧 (108a) 的所述重叠部分 (128a) 的图像数据结合在一起生成第一视频流 (214) 的帧 (302),以及生成第二视频流 (216) 的帧 (304、404),其中所述第二视频流 (216) 的所述帧 (304、404) 包括来自至少所述第二视频帧 (108b) 的所述重叠部分 (128b) 的图像数据 (B'),

第一处理组件 (208),配置为处理所述第一视频流 (214) 的所述帧 (302),

第二处理组件 (210),配置为与处理所述第一视频流 (214) 的所述帧 (302) 的所述第一处理组件 (208) 并行地处理所述第二视频流 (216) 的所述帧 (304、404),其中所述第二视频流 (216) 的所述帧 (304、404) 的处理包括基于来自至少所述第二视频帧 (108b) 的所述重叠部分 (128b) 的图像数据 (B') 来准备重叠 (306、406);以及

重叠组件 (212),配置为在所述第一视频流 (214) 中的处理后的帧 (308) 的与所述第一视频帧 (108a) 的所述重叠部分 (128a) 相对应的部分 (328a) 处,将所述重叠 (306、406) 添加到所述第一视频流 (214) 中的所述处理后的帧 (308),

其中所述第二视频流的所述帧 (304) 包括来自所述第二视频帧 (108b) 和所述第一视频帧 (108a) 的所述重叠部分 (128a、128b) 的图像数据 (B'、A'),所述第二处理组件 (210) 被配置为在准备所述重叠时,将来自所述第二视频帧 (108b) 的所述重叠部分 (128b) 的图像数据 (B') 与来自所述第一视频帧 (108a) 的所述重叠部分 (128a) 的图像数据 (A') 融合,并且所述重叠组件 (212) 被配置为在添加所述重叠时,替换所述第一视频流中的所述处理后的帧 (308) 的与所述第一视频帧 (108a) 的所述重叠部分 (128a) 相对应的部分 (328a) 中的图像数据 (A'),

或者

其中所述重叠 (406) 包括仅来自所述第二视频帧 (108b) 的所述重叠部分 (128b) 的图像数据 (B'),所述重叠组件 (212) 被配置为在添加所述重叠时,在所述第一视频流 (108a) 中的所述处理后的帧 (308) 的与所述第一视频帧 (108a) 的所述重叠部分 (128a) 相对应的部分 (328a) 处,根据掩码 (408) 将所述重叠 (406) 与所述第一视频流 (108a) 中的所述处理后的帧

(308) 融合,其中所述掩码为所述重叠(406)中的每个像素定义融合度,并且所述第二处理组件(210)被配置为在准备所述重叠时计算所述掩码(408)。

10. 一种多传感器摄像机(100),包括:

第一传感器(102a),布置成捕获描绘第一视野的视频帧;

第二传感器(102b),布置成与所述第一传感器(102a)同时捕获视频帧,所述第二传感器(102b)描绘与所述第一视野部分地重叠的第二视野;以及

根据权利要求9所述的处理管道(104),所述处理管道(104)能操作地连接到所述第一传感器(102a)和所述第二传感器(102b),以便接收通过所述第一传感器(102a)和所述第二传感器(102b)捕获到的视频帧。

11. 一种计算机可读介质,所述计算机可读介质具有存储在其上的计算机代码指令,所述计算机代码指令在由具有处理能力的设备执行时用于实施根据权利要求1-8中任一项所述的方法。

## 多传感器摄像机、用于多传感器摄像机的方法和处理管道

### 技术领域

[0001] 本发明涉及多传感器摄像机领域，并且具体地涉及通过多传感器摄像机捕获到的图像的拼接。

### 背景技术

[0002] 包括具有部分重叠视野的多个传感器的摄像机日益受到监控市场的青睐。通过这样的摄像机的多个传感器同时捕获到的图像可以被拼接在一起形成全景图像，该全景图像描绘了该多个传感器的组合视野。优选进行拼接以便提供图像之间的无缝衔接。

[0003] 将来自多个传感器的图像拼接在一起的处理通常是处理密集的 (processing intense)，因此倾向于在摄像机的处理管道 (processing pipeline) 中引入延迟。最终，这可能会限制摄像机的帧速率，原因在于摄像机的处理管道可能无法足够快地处理捕获到的帧。

[0004] 在市场上可利用的摄像机中发现的这个问题的一个解决方案是在摄像机的处理管道的上游包括现场可编程门阵列 (FPGA)，从而在将同时捕获到的图像输入到处理管道之前执行图像的拼接。然而，这种方案是昂贵的，尤其是在FPGA应当足够快从而不会引入由图像拼接引起的不期望的延迟的情况下。而且，这种方案排除了以下情况：摄像机的处理管道中实施的一些图像处理是在执行拼接之前对单个图像进行的。在某些情况下，这可能导致在摄像机的处理管道的终端处输出的最终的拼接图像的品质较差。因此，存在改进的空间。

### 发明内容

[0005] 鉴于以上所述，因此本发明的目的是提供一种用于将通过摄像机的多个传感器同时捕获到的图像拼接在一起的改进过程。具体地，目的是减少在摄像机的处理管道中由拼接图像的过程造成的延迟。另一目的是提供一种以成本效益的方式实施的拼接过程。

[0006] 根据本发明的第一方面，上述目的是通过一种在包括具有部分重叠视野的第一传感器和第二传感器的多传感器摄像机中执行的方法实现，该方法包括：

[0007] 接收分别通过第一传感器和第二传感器同时捕获到的第一视频帧和第二视频帧，

[0008] 其中第一视频帧和第二视频帧中的每个具有不与该第一视频帧和该第二视频帧中的另一个重叠的非重叠部分、以及与该第一视频帧和该第二视频帧中的另一个重叠的重叠部分，

[0009] 通过将来自第一视频帧和第二视频帧的非重叠部分的图像数据与仅来自第一视频帧的重叠部分的图像数据结合在一起，生成第一视频流的帧，

[0010] 生成第二视频流的帧，其中第二视频流的帧包括来自至少第二视频帧的重叠部分的图像数据，

[0011] 并行处理第一视频流的帧和第二视频流的帧，其中第二视频流的帧的处理包括基于来自至少第二视频帧的重叠部分的图像数据来准备重叠 (overlay)；以及

[0012] 在第一视频流中的处理后的帧的与第一视频帧的重叠部分相对应的部分处，将重

叠添加到该第一视频流中的处理后的帧，

[0013] 其中在准备重叠的步骤和添加重叠的步骤中的至少一个步骤中，将来自第一视频帧的重叠部分的图像数据与来自第二视频帧的重叠部分的图像数据融合。

[0014] 任何拼接过程的处理密集的部分在于计算如何融合图像数据以及然后将捕获到的图像的图像数据融合在该图像重叠的部分中，以便提供图像之间的无缝衔接。然而，发明人已经意识到，仅需要捕获到的视频帧的极小部分图像数据以便计算如何融合图像数据、准备用于融合的图像数据、或者甚至实施融合。更具体地，根据特定实施例，仅需要来自视频帧中的一个或二者的重叠部分的图像数据来实施这些步骤。因此，提议将极小部分图像数据分离成第二视频流的帧，该极小部分图像数据可以与包括在第一视频流的帧中的大部分图像数据并行处理。以这种方式，仅在极小部分图像数据上操作的拼接过程的处理密集的部分可以与在大部分图像数据上实施的该拼接过程的其他较少处理密集的部分并行实施。因此减少了延迟。

[0015] 市场上可利用的摄像机通常具有适于在视频流上添加重叠的硬件，该重叠例如是商标标志等类似物的形式。发明人已经意识到，这种可利用的功能可以与图像的拼接结合使用。例如，第二视频流的帧（即，来自第一视频帧和第二视频帧中的一个或二者的重叠部分的图像数据）可以用于准备重叠，然后通过替换第一视频流中该帧的图像数据或者通过与第一视频流中该帧的图像数据融合来将该重叠添加到第一视频流的帧上，从而使该实施方式具有成本效益。

[0016] 第一传感器和第二传感器具有部分重叠视野。换言之，第一传感器和第二传感器描绘相同的场景，但具有不同的而且重叠的视野。这意味着在第一视频帧和第二视频帧中将存在这样的像素：该像素保存来自（即描绘）场景的相同部分的图像信息。第一视频帧和第二视频帧中包括这样的像素的部分在此被称为重叠部分。反过来，第一视频帧的以下部分被称为非重叠部分：该部分中的像素描绘未被第二视频帧描绘的场景，反之亦然。

[0017] 重叠通常是指旨在被重叠在另一图像上的图像。将重叠添加到帧或图像通常是指：重叠被重叠在该帧或该图像上或者被插入在该帧或该图像上。可通过替换帧中的图像数据或者通过将重叠与帧中的图像数据融合（例如，通过设定重叠的透明度）来添加重叠。

[0018] 融合通常是指：计算加权平均值。

[0019] 在第一组实施例中，第二视频流的帧包括来自第二视频帧和第一视频帧的重叠部分的图像数据，其中准备重叠的步骤包括将来自第二视频帧的重叠部分的图像数据与来自第一视频帧的重叠部分的图像数据融合。在这样的实施例中，因此在准备重叠的过程中将来自第一视频帧和第二视频帧的重叠部分的图像数据融合。由于这个过程是与第一视频流的处理并行进行的，所以融合过程不会造成任何额外的延时。

[0020] 在第一组实施例中，与第一视频帧和第二视频帧的重叠部分相对应的图像数据的融合是在重叠的准备期间实施的。因此，在将重叠添加到第一视频流中的处理后的帧时，不需要再进行进一步的融合。在这种情况下，可通过替换来添加重叠。更详细地，添加重叠的步骤，重叠替换第一视频流中的处理后的帧的在与第一视频帧的重叠部分相对应的部分中的图像数据。第一组实施例的总体优势在于，可更加自由地选择融合算法。甚至可以选择执行重叠的添加的可用硬件不支持的算法。将在下文中描述算法的示例。

[0021] 可通过各种方式在重叠的准备期间将来自第一视频帧和第二视频帧的重叠部分

的图像数据融合。

[0022] 通常,融合可以包括:针对重叠部分中的每个像素,计算来自第一视频帧的图像数据和来自第二视频帧的图像数据的加权平均值。

[0023] 根据在此被称为“阿尔法融合 (alpha-blending)”的一个示例,加权平均值的权重取决于从像素到第一视频帧和第二视频帧的重叠部分和非重叠部分之间的边界的距离。更详细地,与来自第二视频帧的重叠部分的图像数据相关联的权重取决于从像素到第二视频帧的重叠部分和非重叠部分之间的边界的最短距离,并且与来自第一视频帧的重叠部分的图像数据相关联的权重取决于从像素到第一视频帧的重叠部分和非重叠部分之间的边界的最短距离。

[0024] 根据在此被称为“多频带融合”的进一步的示例,可以为空间高频内容和空间低频内容不同地选择权重。例如,阿尔法融合可以应用于较低的频率;然而,用于较高的频率的权重被设置为0或1,这取决于像素是更靠近第一视频帧的非重叠部分的边界还是更靠近第二视频帧的非重叠部分的边界。更详细地,融合可以包括:

[0025] 计算来自第一视频帧的重叠部分的图像数据和来自第二视频帧的重叠部分的图像数据的第一加权平均值,其中该第一加权平均值是关于低于第一阈值的空间频率来计算的,

[0026] 计算来自第一视频帧的重叠部分的图像数据和来自第二视频帧的重叠部分的图像数据的第二加权平均值,其中该第二加权平均值是关于等于或高于第二阈值的空间频率来计算的,该第二阈值大于或等于第一阈值,

[0027] 其中,为第一加权平均值和第二加权平均值不同地选择权重,并且

[0028] 通过将第一加权平均值与第二加权平均值相加来计算重叠。

[0029] 根据在此被称为“基于特征的融合”的又一示例,该融合包括:

[0030] 识别第一视频帧的重叠部分中的对象和第二视频帧的重叠部分中的对象,

[0031] 计算重叠部分中来自第一视频帧的图像数据和来自第二视频帧的图像数据的加权平均值,

[0032] 其中,如果对象被识别为位于第一视频帧的重叠部分和第二视频帧的重叠部分二者中,则对于被识别为属于该对象的像素,将与第一视频帧的图像数据和第二视频帧的图像数据中的一个相关联的权重设置为0。利用这种类型的融合,仅从第一视频帧和第二视频帧中的一个包括表示诸如人员等对象的图像数据,从而避免该对象在融合图像中在稍微不同的位置处出现两次。由于多个传感器的视差效应,可能是这样的情况。

[0033] 根据第二组实施例,融合发生在将重叠添加到第一视频流中的处理后的帧上时,而不是在准备重叠时。在这样的实施例中,重叠仅包括来自第二视频帧的重叠部分的图像数据,并且添加重叠的步骤包括:在第一视频流中的处理后的帧的与第一视频帧的重叠部分相对应的部分处,将重叠与该第一视频流中的处理后的帧融合。重叠的添加通常是通过专用硬件实施的,该专用硬件通常是专用集成电路 (ASIC) 的形式。通过在添加重叠时执行融合,不需要在例如中央处理单元或图形处理单元中运行自定义算法。相反,诸如ASIC等专用硬件可以用于在添加重叠的同时实施融合,从而加快融合过程。

[0034] 例如,可通过设定在将重叠添加到第一视频流中的处理后的帧时该重叠的透明度来实施融合。对于不同的像素,重叠的透明度可被赋予不同的数值。透明度为1导致来自重

叠的图像数据不融合,并且添加重叠后像素中的结果值是第一视频帧的值。透明度为0导致重叠完全融合。在该情况下,添加重叠后像素中的结果值是第二视频帧的值。介于0和1之间的透明度将导致重叠融合到各种程度,即,添加重叠后像素中的图像数据结果值是第一视频帧的图像数据和第二视频帧的图像数据的加权平均值。因此,对于重叠中的每个像素,可以在重叠的准备期间指定例如与0和1之间的数值相对应的融合度。然后,该融合度可以用于设定在将重叠添加到第一视频流中的处理后的帧上时该重叠的透明度。

[0035] 例如,可以根据掩码将重叠与第一视频流中的处理后的帧融合,掩码为重叠中的每个像素定义融合度。然后,该掩码可以用于设定在将重叠添加到第一视频流中的处理后的帧上时该重叠的透明度。

[0036] 通常,在准备重叠时确定掩码。根据一些实施例,独立于任何图像数据来确定掩码。例如,融合度可以对应于上述阿尔法融合,以使得该融合度仅取决于从像素到第二视频帧的非重叠部分的边界的距离。更详细地,重叠中的每个像素可以与第二视频帧的重叠部分的像素相对应,其中融合度取决于从该像素到第二视频帧的重叠部分和非重叠部分之间的边界的最短距离。

[0037] 根据一些实施例,附加地或替代地,可以基于图像数据来确定掩码。图像数据可以与第二视频帧的重叠部分相对应,但优选地,该图像数据与第一视频帧和第二视频帧二者的重叠部分相对应。在后一种情况下,第二视频流的帧包括来自第二视频帧和第一视频帧的重叠部分的图像数据,并且准备重叠的步骤还包括:基于来自第二视频帧和第一视频帧的重叠部分的图像数据来计算掩码。以这样的方式,融合可以适于重叠部分的图像内容。

[0038] 如上面关于第一组实施例进一步讨论的,融合是指:形成第一视频帧的重叠部分中的图像数据和第二视频帧的重叠部分中的图像数据的加权平均值。应注意的是,融合度(即,掩码表示的值)与要应用于来自第二视频帧的图像数据的权重相对应。可以通过从1中减去融合度来获得要应用于来自第一视频帧的图像数据的权重。

[0039] 因此,基于来自第二视频帧和第一视频帧的重叠部分的图像数据来计算掩码的步骤对应于计算与来自第二视频帧和第一视频帧的图像数据相关联的权重。因此,上面结合第一组实施例的关于计算这些权重的内容同样适于计算融合度,即,掩码的元素。例如,可以计算掩码以便包括与多频带融合和/或基于特征的融合相关联的权重。

[0040] 取决于具体情况,基于图像数据计算掩码的过程可能需要时间。因此,为了进一步减少延迟,不针对被处理的每个帧计算新的掩码可能是有利的。相反,一旦已经计算出掩码,就可将该掩码重复用于多个帧。更详细地,可以以低于从第一传感器和第二传感器接收视频帧的速率来更新掩码。

[0041] 在将来自第一视频帧和第二视频帧的重叠部分的图像数据融合之前,优选地,将该第一视频帧和该第二视频帧对准。因此,第一视频流的帧的处理和第二视频流的帧的处理可以包括将该第一视频帧的图像数据和该第二视频帧的图像数据对准。除其他之外,对准可以包括:执行圆桶校正(barrel-correction)和旋转帧(通常在安装传感器以拍摄场景的站立式图片时)、以及执行第一视频帧和第二视频帧在例如圆柱上的投影。后者用于补偿第一传感器和第二传感器未安装在相同位置的情况,这意味着从不同视点描绘该场景。这可以通过例如执行投影来校正,从而给观察者留下各帧是从相同视点捕获到的印象。对每个帧应用的具体对准通常是相同的,并且可以在传感器的校正期间确定该具体对准。

[0042] 根据本发明的第二方面,上述目的是通过一种用于包括具有部分重叠视野的第一传感器和第二传感器的多传感器摄像机的处理管道来实现的,该处理管道包括:

[0043] 接收器,配置为接收分别通过第一传感器和第二传感器同时捕获到的第一视频帧和第二视频帧,

[0044] 其中第一视频帧和第二视频帧中的每个具有不与该第一视频帧和该第二视频帧中的另一个重叠的非重叠部分、以及与该第一视频帧和该第二视频帧中的另一个重叠的重叠部分;

[0045] 帧生成组件,配置为通过将来自第一视频帧和第二视频帧的非重叠部分的图像数据与仅来自第一视频帧的重叠部分的图像数据结合在一起生成第一视频流的帧,以及生成第二视频流的帧,其中第二视频流的帧包括来自至少第二视频帧的重叠部分的图像数据,

[0046] 第一处理组件,配置为处理第一视频流的帧,

[0047] 第二处理组件,配置为与处理第一视频流的帧的第一处理组件并行地处理第二视频流的帧,其中第二视频流的帧的处理包括基于来自至少第二视频帧的重叠部分的图像数据来准备重叠;以及

[0048] 重叠组件,配置为在第一视频流中的处理后的帧的与第一视频帧的重叠部分相对应的部分处,将重叠添加到该第一视频流中的处理后的帧,

[0049] 其中第二处理组件和重叠组件中的至少一个被配置为:在准备重叠或者添加重叠时,将来自第一视频帧的重叠部分的图像数据与来自第二视频帧的重叠部分的图像数据融合。

[0050] 根据本发明的第三方面,上述目的是通过一种多传感器摄像机来实现的,该多传感器摄像机包括:

[0051] 第一传感器,布置成捕获描绘第一视野的视频帧;

[0052] 第二传感器,布置成与第一传感器同时捕获视频帧,该第二传感器描绘了与第一视野部分地重叠的第二视野;以及

[0053] 根据第二方面的处理管道,该处理管道可操作地连接到第一传感器和第二传感器,以便接收通过第一传感器和第二传感器捕获到的视频帧。

[0054] 根据本方面的第四方面,上述目的是通过一种包括计算机可读介质的计算机程序产品来实现的,该计算机可读介质具有存储在其上的计算机代码指令,该计算机代码指令在由具有处理能力的设备执行时用于实施第一方面的方法。

[0055] 第二、第三和第四方面通常可具有与第一方面相同的特征和优点。还应注意的是,本发明涉及特征的所有可能的组合,除非另外明确指明。

[0056] 通常,权利要求书中使用的所有术语应根据其在本技术领域中的通用含义来解释,除非在此另外明确定义。在此公开的任何方法中的步骤不必以所公开的精确顺序执行,除非明确指明。

## 附图说明

[0057] 参照附图,通过以下对本发明的优选实施例的说明性和非限制性的详细描述,将更好地理解本发明的上述以及其他目的,特征和优点,其中相同的附图标记将用于类似的

元素,附图中:

- [0058] 图1示意性地示出了根据实施例的多传感器摄像机。
- [0059] 图2示意性地示出了图1的多传感器摄像机的处理管道。
- [0060] 图3示意性地示出了根据第一组实施例的处理管道中的处理流程。
- [0061] 图4示意性地示出了根据第二组实施例的处理管道中的处理流程。
- [0062] 图5示意性地示出了如何选择用于将第一视频帧和第二视频帧的重叠部分的图像数据融合的权重的示例。
- [0063] 图6示意性地示出了如何选择用于将第一视频帧和第二视频帧的重叠部分的图像数据融合的权重的示例。
- [0064] 图7是根据实施例的在多传感器摄像机中执行的方法的流程图。

## 具体实施方式

[0065] 现在将在下文中参照其中示出了本发明的实施例的附图来更全面地描述本发明。在此公开的系统和装置将在操作过程中进行描述。

[0066] 图1示出了多传感器摄像机100。摄像机100包括多个图像传感器102a、102b以及处理管道104。摄像机100还可以包括视频编码器106。这里,为了说明和简化的目的,示出了两个图像传感器102a、102b。然而,应理解的是,摄像机100可以包括多于两个图像传感器。

[0067] 图像传感器102a、102b彼此相关地布置,使得他们具有部分重叠的场景视野。这暗示了通过图像传感器102a、102b捕获到的视频帧将部分地重叠,意味着将在通过该图像传感器102a、102b中的多于一个图像传感器(但如果存在多于两个的图像传感器,则不必为所有传感器)捕获到的视频帧中描绘场景的一部分。因此,在通过至少两个图像传感器捕获到的视频帧中,存在描绘场景的相同部分的像素。

[0068] 摄像机100被布置成以特定帧率捕获视频帧。具体地,摄像机100以特定速率、同时地使用多个传感器102a、102b来捕获视频帧。具体地,摄像机100被布置成同时地使用第一图像传感器102a来捕获第一视频帧108a以及使用第二图像传感器102b来捕获第二视频帧108b。然后,将捕获到的视频帧108a、108b输入到被布置在摄像机100中的图像传感器102a、102b下游的处理管道104。

[0069] 处理管道104被布置成处理通过图像传感器102a、102b捕获到的视频帧108a、108b。“处理管道”通常是指按顺序连接的一组数据处理元件,其中一个元件的输出是下一个元件的输入。具体地,处理管道104被配置为将视频帧108a、108b拼接在一起以创建场景的全景图像110。因此,全景图像110可以对应于图像传感器102a、102的组合视野,即,联合视野。一旦已经创建了全景图像110,就可将该全景图像110转发给视频编码器106,该视频编码器106在该全景图像110例如以比特流112的形式从摄像机100输出之前对该全景图像110编码,该比特流112通过网络被发送到视频解码器。

[0070] 图2更详细地示出了处理管道104。处理管道包括接收器202、帧生成组件206、第一处理组件208、第二处理组件210和重叠组件212。处理管道还可以包括图像处理组件204。

[0071] 接收器被布置成接收通过图1中所示的图像传感器102a、102b捕获到的视频帧108a、108b。帧生成组件206被布置在接收器202的下游,并且被布置成可选地经由图像处理组件204从接收器202接收视频帧108a、108b,该图像处理组件204被布置在接收器202的下

游、但在帧生成组件206的上游。例如,图像处理组件204可以执行通常在摄像机的图像处理管道中执行的各种类型的图像处理,例如缺陷像素校正、像列固定模式噪音补偿和串音补偿的除伪(artefact removal)、白平衡、渐晕校正、噪声过滤、去马赛克、锐化、颜色矩阵处理、暗电流校正和伽马校正(gamma correction)。然而,图像处理中的一些或全部(例如白平衡处理)替代地可以通过第一处理组件208和/或第二处理组件210来执行。帧生成组件206相应地将生成第一视频流214和第二视频流216的帧。第一视频流214的帧被转发给第一处理组件208,并且第二视频流216的帧被转发给第二处理组件210,该第二处理组件210与该第一处理组件208并排布置。第一处理组件208和第二处理组件210的输出然后被转发给重叠组件212,该重叠组件212相应地输出全景图像110。将在下文中更详细地描述这个过程。

[0072] 因此,处理管道104包括被配置为实现该处理管道104的功能的各种组件202、204、206、208、210、212。具体地,每个所示组件对应于处理管道104的功能。然而,如本领域技术人员所理解的,各种组件也可以与包括在处理管道104中的结构相关联以实现组件的功能。如本领域技术人员所意识到的,具体结构可以取决于组件的具体实施方式,例如,该具体结构是否以硬件、软件或者它们的组合来实现。

[0073] 通常,处理管道104可以包括被配置为实现组件202、204、206、208、210、212,并且更具体地实现他们的功能的电路系统。

[0074] 在硬件实施方式中,组件202、204、206、208、210、212中的每个可以对应于电路系统,该电路系统是专用的并且特别设计为提供组件的功能。该电路系统可以是一个或多个集成电路(例如,一个或多个专用集成电路)的形式。作为示例,帧生成组件206因此可以包括在使用时生成第一视频流214和第二视频流216的电路系统。

[0075] 在软件实施方式中,电路系统替代地可以是一个或多个处理器(例如,一个或多个微处理器)的形式,该电路系统与存储在诸如非易失性存储器等(非暂时性)计算机可读介质上的计算机代码指令相关联,使得处理管道104实施在此公开的任何方法。在此情况下,组件202、204、206、208、210、212因此可以分别与存储在该计算机可读介质上的计算机代码指令的一部分相对应,计算机代码指令在由处理器执行时使处理管道104实施组件的功能。

[0076] 应理解的是,还可以具有硬件和软件实施方式的组合,意味着组件202、204、206、208、210、212中的一些组件的功能是以硬件来实现的,并且其它组件的功能是以软件来实现的。

[0077] 现在将参照图1-6和图7的流程图在下面描述处理管道104的操作。

[0078] 在步骤S02中,接收器202从图像传感器102a、102b接收第一视频帧108a和第二视频帧108b。第一视频帧108a和第二视频帧108b通过摄像机100同时捕获到。

[0079] 图3和图4更详细地示出了第一视频帧108a和第二视频帧108b。第一视频帧108a具有不与第二视频帧108b重叠的部分118a。换言之,非重叠部分118a的像素描绘了场景中未被第二视频帧108b描绘的部分。第一视频帧108a的非重叠部分118a包括图像数据A。类似地,第二视频帧108b具有不与第二视频帧108b重叠的部分118b。第二视频帧108b的非重叠部分118b包括图像数据B。第一视频帧108a还具有与第二视频帧108b的部分128b重叠的部分128a。换言之,部分128a、128b中的像素描绘了场景的相同部分。第一视频帧108a的重叠部分128a包括图像数据A',并且第二视频帧108b的重叠部分128b包括图像数据B'。

[0080] 在步骤S04中,帧生成组件206继续生成第一视频流的帧302。通过将来自第一视频帧108a的非重叠部分118a的图像数据A、来自第一视频帧108a的重叠部分128a的图像数据A'以及来自第二视频帧108b的非重叠部分118b的图像数据B结合在一起生成帧302。以这样的方式,帧302包括来自整个第一视频帧108a的图像数据A、A'以及来自第二视频帧108b的非重叠部分的图像数据B。尤其应注意的是,帧302中不包括来自第二视频帧108b的重叠部分128b的图像数据B'。因此,对于视频帧108a、108b重叠的部分,帧302中包括来自视频帧中的仅一个视频帧的图像数据。

[0081] 第一视频流的帧302是覆盖了图像传感器102a、102b的组合视野的全景图像。然而,在生成这个全景图像时,没有对来自输入视频帧108a、108b的图像数据实施混合。相反,以上面描述的方式为全景图像中的每个像素选择来自一个视频帧的图像数据。更详细地,通过将来自第一视频帧108a和第二视频帧108b的部分118、128a、118b的图像数据A、A'结合在一起生成帧302。这意味着图像数据A、A'、B是以适当的顺序并排布置的,从而构成新的图像(全景图像),该图像覆盖了图像传感器102a、102b的组合视野。

[0082] 然后,将第一视频流的帧302转发给第一处理组件208。

[0083] 帧生成组件206还生成第二视频流的帧304、404。帧304、404包括至少第二视频帧128b的重叠部分的图像数据。

[0084] 在图3所示的第一组实施例中,第二视频流的帧304包括来自第一视频帧108a和第二视频帧108b二者的重叠部分128a、128b的图像数据A'、B'。在图4所示的第二组实施例中,第二视频流的帧404可以包括仅来自第二视频帧108b的图像数据B',或者可以包括来自第一视频帧108a和第二视频帧108b二者的图像数据。这两个替代方案在图4中以符号“/”隔开。

[0085] 然后,将帧304、404输入第二处理组件210。

[0086] 在步骤S08中,第一处理组件208处理帧302,并且第二处理组件210处理帧304、404。第一视频流的帧302的处理和第二视频流的帧304、404的处理是并行进行的。

[0087] 如上面进一步描述的,第一视频流的帧302的处理可以包括各种图像处理步骤,例如缺陷像素校正、像列固定模式噪音补偿和串音补偿的除伪、白平衡、渐晕校正、噪声过滤、去马赛克、锐化、颜色矩阵处理、暗电流校正和伽马校正。图像处理中的一些或全部可以在生成第一视频流和第二视频流之前通过图像处理组件204执行,并且图像处理中的一些或全部可以在生成第一视频流和第二视频流之后通过第一处理组件208执行。第一视频流的帧302的处理通常还包括本领域已知的对准。这可以包括帧302或其部分的圆桶畸变校正(即,来自第一视频帧108a和第二视频帧108b的部分可能需要不同的校正),帧302或其部分(在摄像机被布置成以站立式格式拍摄图片时)的旋转,以及帧302或者更准确地说是来自视频帧108a、108b中的不同视频帧的部分在表面(例如圆柱)上的投影。进行后者以补偿视频帧108a、108b是从不同视点捕获到的这一事实。通常,在相机的初始校准之后,对每个帧进行同样的对准。因此,可以在相机的图像传感器的校准的参数的基础上进行对准。

[0088] 在图3和图4中,帧302的处理后的版本由308表示。

[0089] 第二视频流的帧304、404的处理还可以包括根据上面已经描述过的对准,以及重叠306、406的准备。通常,帧302比帧304、404大得多,使得可以在通过第一处理组件208对准并且亦处理帧302的同时实施帧304、404的对准和重叠的准备。

[0090] 在图3所示的第一组实施例中,重叠306的准备包括融合,即,混合来自第一视频帧108a和第二视频帧108b的重叠部分128a、128b的图像数据A' 和B'。在这个实施例中,从第二处理组件210输出的重叠306包括融合后的图像数据。第二处理组件210例如可以通过为重叠部分128a、128b中的每个对应像素计算图像数据A' 和图像数据B' 的加权平均值来融合该图像数据A' 、B'。不同像素的加权平均值的权重可能是不同的。用w(x) 表示应用于像素x中的图像数据B' 的权重w,0≤w≤1,融合可以描述为:

$$[0091] (1-w(x)) A'(x) + w(x) B'(x)$$

[0092] 这里,B'(x) 表示像素x中重叠部分128b中的图像数据,并且A'(x) 表示与像素x相对应的像素中重叠部分128a中的图像数据。

[0093] 可以独立于图像数据A' 、B' 来选择权重。随着时间的推移,权重还可以保持不变。根据归属于图3的第一组实施例的一个实施例,与图像数据B' 相关联的权重w取决于从第二视频帧108b的重叠部分128b中的像素到第二视频帧108b的非重叠部分118b的边界的最短距离。这意味着与图像数据A' 相关联的权重(1-w) 将取决于重叠部分128a中对应像素与第一视频帧108a的重叠部分128a和非重叠部分118a之间的边界之间的最短距离。这在图5中进一步被示出,图5示出了应用于图像数据B' 的权重502b是到重叠部分128b和非重叠部分118b之间的边界的距离的函数。进一步,如图5中所示,应用于图像数据A' 的权重502a是到重叠部分128a和非重叠部分118a之间的边界的距离的函数。这里,权重与到边界的接近度成比例,即,权重随着到边界的接近而线性地增加。然而,应理解的是,其它函数关系同样也是可能的。然而,权重通常随着到边界的接近而增加。

[0094] 对于低频内容(例如低于第一阈值的空间频率)和高频内容(例如等于或高于第二阈值的频率),可以进一步不同地选择权重。第二阈值可以等于或大于第一阈值。在此情况下,可以对图像数据A' 、B' 进行空间高通滤波和空间低通滤波,以提取图像数据A' 、B' 的低频内容和高频内容。通过以第一方式选择权重,可以计算低频内容的第一加权平均值。例如,可以如关于图5所述来选择低频内容的权重。通过以第二方式选择权重,可以计算高频内容的第二加权平均值。例如,可以如图6中所示来选择高频内容的权重,图6示出了应用于图像数据B' 的权重602b和应用于图像数据A' 的权重602a。在这种情况下,权重要么是0,要么是1。对于部分128b中的如下像素,权重602b等于1:该像素到部分118b、128b之间的边界的距离小于部分128a中的对应像素到部分118a、128a之间的边界的距离。对于部分128b中的其它像素,权重为0。

[0095] 也可以根据图像数据A' 、B' 来选择权重。在一个示例中,权重可以由重叠部分118a、118b中的对象(例如人员)的存在来控制。更详细地,由于图像传感器102a、102b从稍微不同的位置和角度查看场景,所以第一视频帧和第二视频帧108a、108b将受到视差效应。因此,如果场景中存在对象,则即使已实施过对准,该对象也可能出现在重叠部分128a、128b中的稍微不同的位置处。因此,如果通过形成加权平均值来将重叠部分128a、128b中的图像数据A' 、B' 融合,则对象可能在融合图像中出现两次。为了避免这种情形,可以选择仅包括来自重叠部分128a、128b中的一个重叠部分的对象。更详细地,可以使用标准对象检测技术来检测重叠部分128a、128b中的对象。对于被识别为属于重叠部分128a、128b中的对象的像素,可以将来自第一视频帧108a的图像数据A' 和来自第二视频帧108b的图像数据B' 中的一个的权重设置为0,这意味着融合图像仅包括来自第一视频帧108a和第二视频帧108b

中的一个视频帧的图像数据。

[0096] 在图4所示的第二组实施例中,重叠406的准备不包括图像数据A'、B'的融合。在这些实施例中,重叠406仅包括来自第二视频帧108b的重叠部分128b的图像数据B'。事实上,重叠406通常与第二视频帧108b的重叠部分128b直接对应,因为重叠406中的每个像素与第二视频帧108b的重叠部分128b的像素相对应。根据第二组实施例的第二处理组件210不是实施任何融合,而是计算掩码408。掩码408为重叠408中的每个值定义了数值,该数值可被称为融合度。融合度是在0和1之间的数值,其描述了重叠408与另一块图像数据的融合中应包括多大比重(即多高的百分比)的重叠408。由于重叠408与第二视频帧108b的重叠部分128b的图像数据B'相对应,并且另一块图像数据与第一视频帧108a的图像数据A'相对应(如后面将描述的),融合度与如上面关于第一组实施例描述的、要应用于来自第二视频帧108b的图像数据B'的权重w相当。换言之,在第二组实施例中,第二处理组件210计算权重w但不执行实际的融合。

[0097] 上面结合第一组实施例已经说过的关于计算要应用于图像数据B'的权重w因此同样适用于第二组实施例。在独立于图像数据来计算权重的情况下,输入到第二处理组件210的帧404仅需要包括来自第二视频帧108b的重叠部分128b的图像数据B'。在基于图像数据计算权重的情况下,输入到第二处理组件210的帧404包括来自重叠部分128a、128b二者的图像数据A'、B'。

[0098] 为了进一步加快处理,可以重复使用处理先前帧时使用的掩码。例如,可以以低于从第一传感器102a和第二传感器102b接收视频帧的速率的速率来更新掩码。以这样的方式,减少了处理时间,从而进一步减少了系统的延迟。

[0099] 在步骤S10中,重叠组件212继续将重叠306、406添加到第一视频流中的处理后的帧308。具体地,重叠组件212在处理后的帧308的与第一视频帧108a的重叠部分128a相对应的部分328a处添加重叠306、406。重叠的添加在第一组实施例的情况下可以通过替换来完成,或者在第二组实施例的情况下可以通过融合来完成。

[0100] 更具体地,在图3所示的第一组实施例中,在准备重叠306时,图像数据A'、B'已经融合。因此,在将重叠306添加到第一视频流中的处理后的帧308时,不需要进一步的融合。如图3所示,重叠组件212因此可以替换处理后的帧308的部分328a中的图像数据。例如,这可以通过将重叠的透明参数设置为等于0来实现,从而指示重叠组件212以不透明的方式添加重叠306,这在实践中与通过替换图像数据来添加重叠306是一样的。

[0101] 在图4所示的第二组实施例中,在将重叠406添加到处理后的帧308时,重叠组件212继续融合图像数据A'、B'。重叠组件212根据通过第二处理组件210计算得到的掩码408来实施融合。更详细地,重叠组件212可以计算重叠406和处理后的帧308的部分328a的加权平均值。通过掩码408来给出重叠406中的像素的权重,并且通过从1中减去对应像素位置处的掩码的数值来得到要应用于部分328a的权重。在实践中,这可以在像素间的对比(pixel-wise)的基础上通过使用掩码设置重叠的透明水平来实现。

[0102] 作为将重叠306、406添加到处理后的帧308的结果,生成了全景图像110。全景图像110是第一视频帧108a和第二视频帧108b的拼接结果。对于在第一视频帧108a中描绘但在第二视频帧108b中未描绘的场景的部分,全景图像110包括来自第一视频帧108a的图像数据A。类似地,全景图像110包括来自在第二视频帧108b中描绘但在第一视频帧108a中未描绘的场景的部分,全景图像110包括来自第二视频帧108b的图像数据B。

绘的场景的部分的图像数据B。对于通过第一视频帧108a和第二视频帧108b二者描绘了的场景的部分,全景图像110包括如上所述的、来自第一视频帧108a的图像数据A' 和来自第二视频帧108b的图像数据B' 的融合。

[0103] 应意识到的是,本领域技术人员可以以多种方式来修改上述实施例并且依然使用如上面实施例中所示的本发明的优点。例如,为了简化的目的,以上示例针对具有两个传感器的摄像机来给出。然而,本发明在存在多于两个传感器的情况下同等地适用。因此,本发明不应限于所示实施例,而应仅由所附权利要求限定。另外,如本领域技术人员所理解的,所示出的实施例可以被组合。

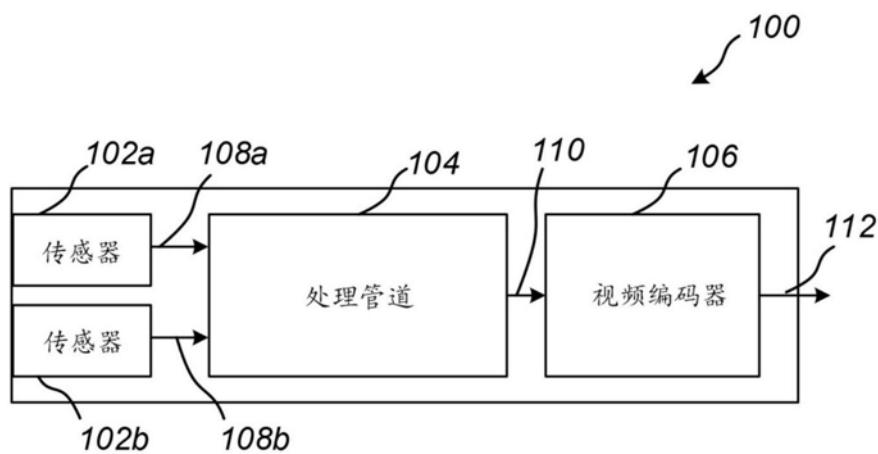


图1

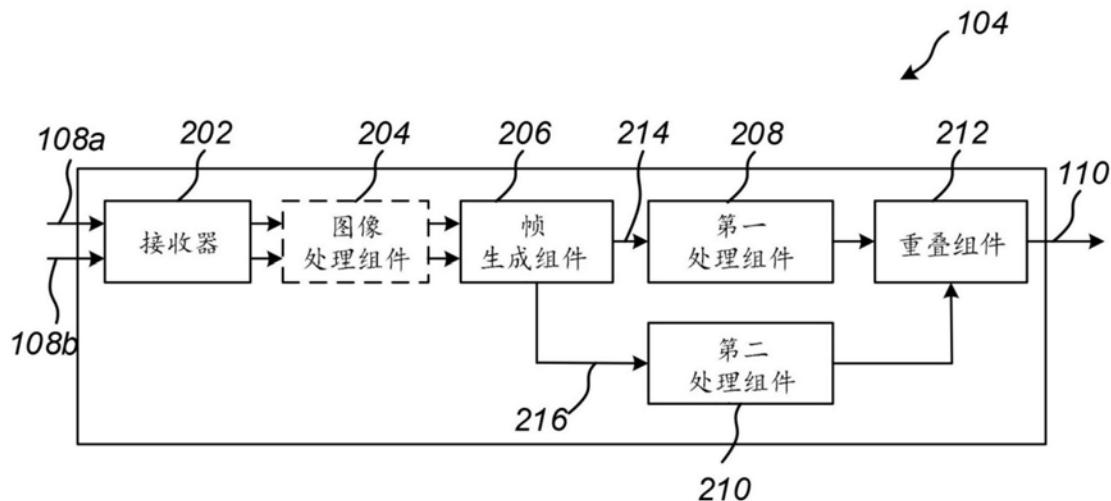


图2

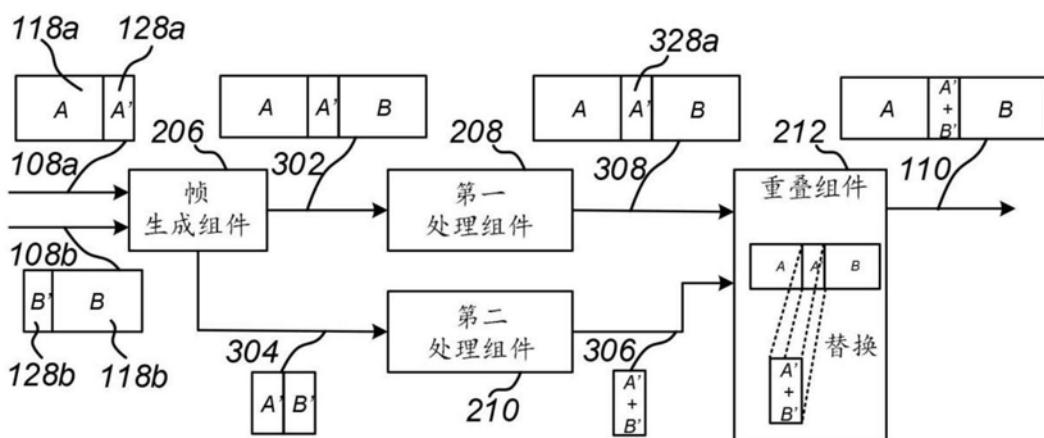


图3

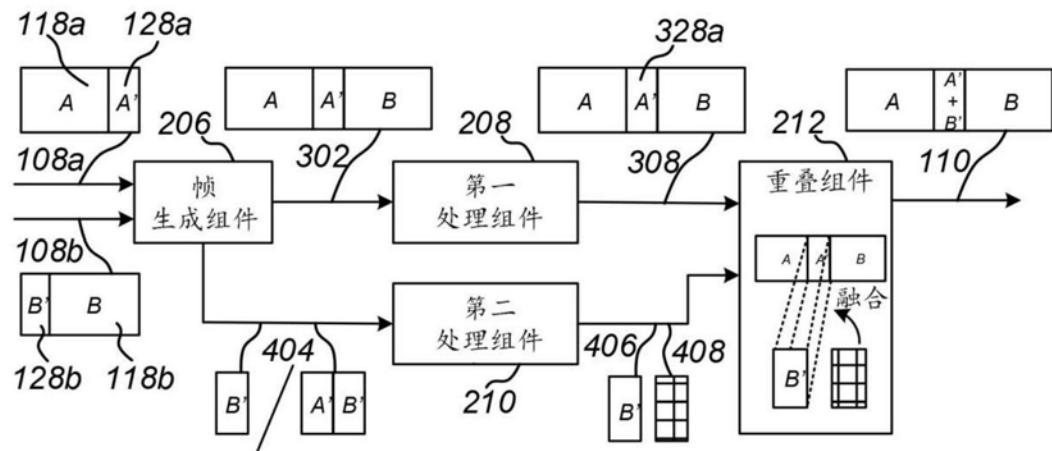


图4

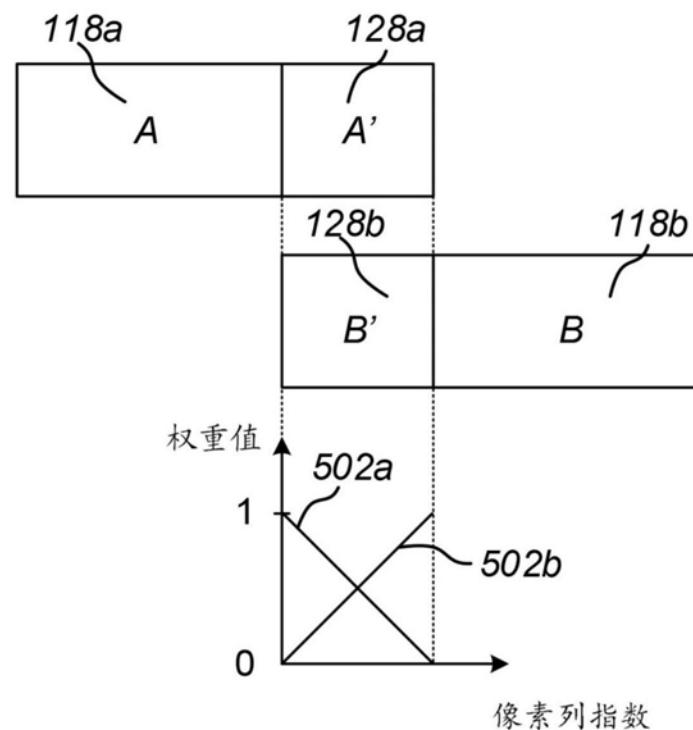


图5

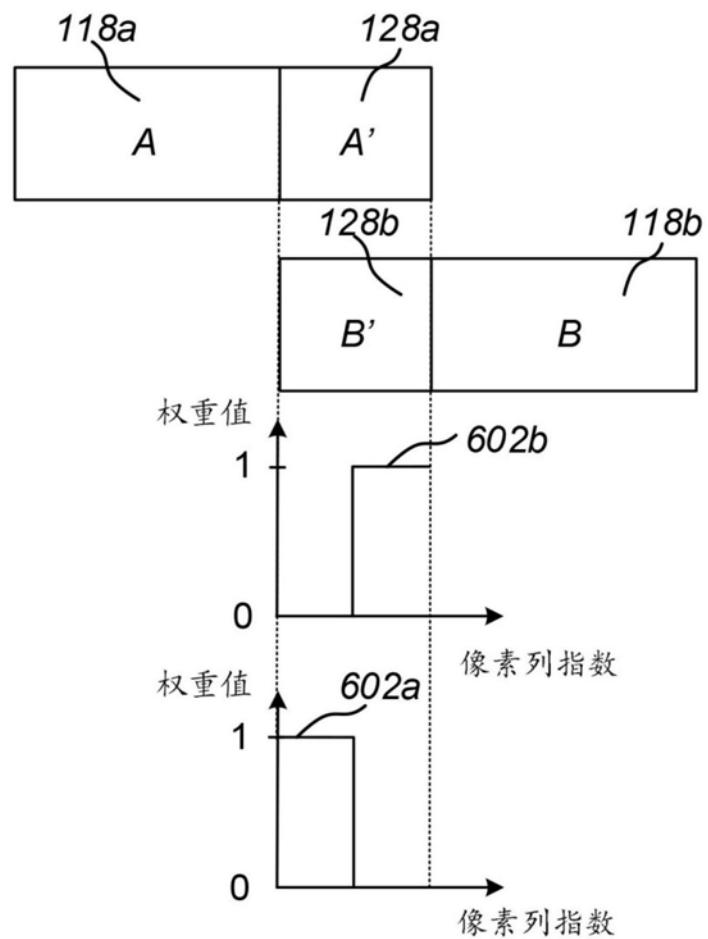


图6

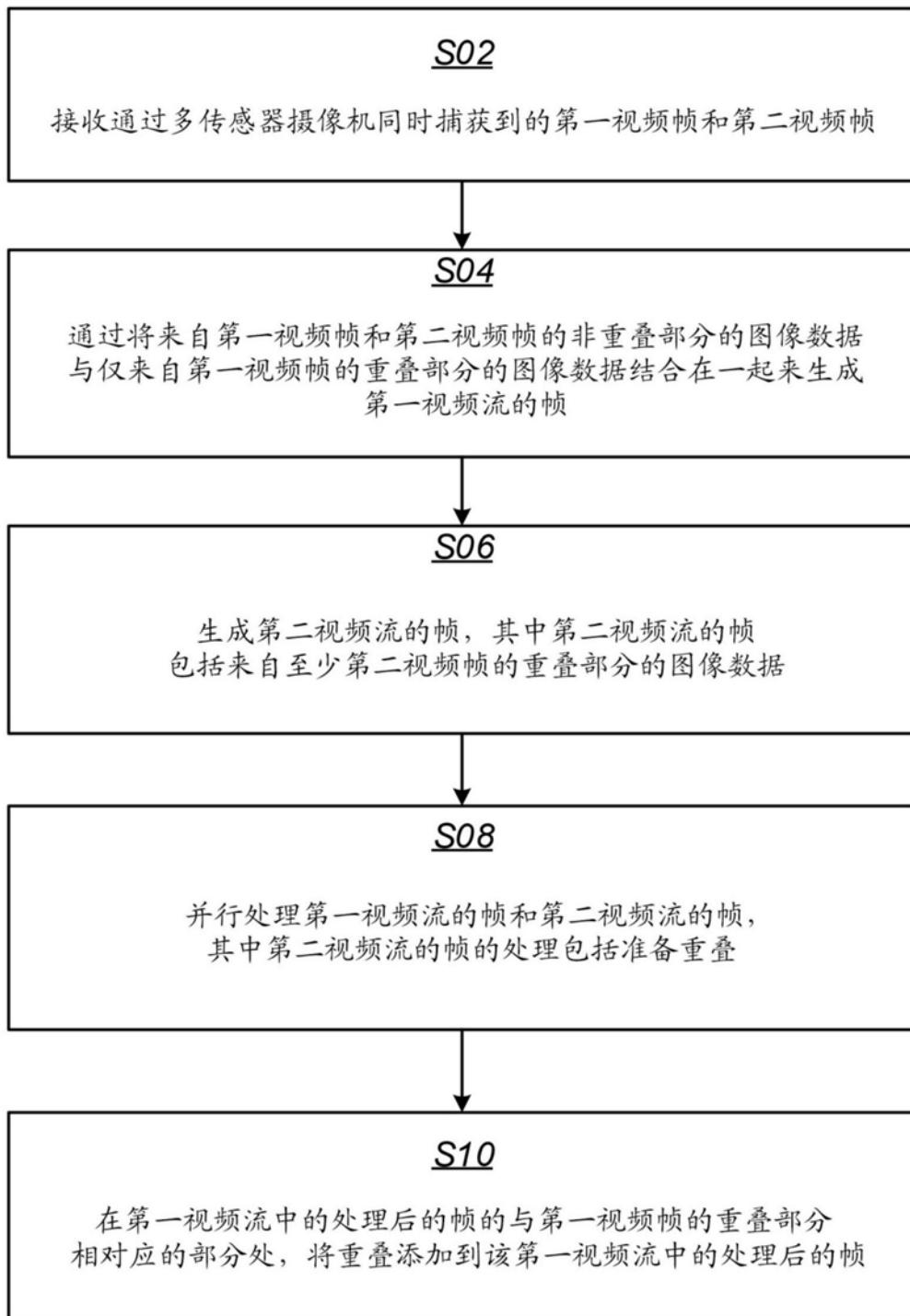


图7