



(12) 发明专利申请

(10) 申请公布号 CN 102075680 A

(43) 申请公布日 2011. 05. 25

(21) 申请号 201010549989. 3

(22) 申请日 2010. 11. 15

(30) 优先权数据

264616/09 2009. 11. 20 JP

(71) 申请人 索尼公司

地址 日本东京都

(72) 发明人 河井哲郎 中尾大辅

(74) 专利代理机构 北京市柳沈律师事务所

11105

代理人 郭定辉

(51) Int. Cl.

H04N 5/232(2006. 01)

H04N 5/91(2006. 01)

H04N 5/92(2006. 01)

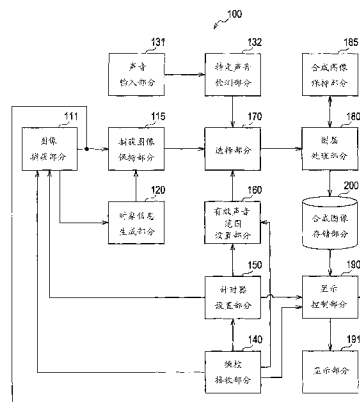
权利要求书 2 页 说明书 35 页 附图 30 页

(54) 发明名称

图像处理设备、图像处理方法和程序

(57) 摘要

提供图像处理设备及图像处理方法。该图像处理设备包括：特定声音检测部分，其检测在由包括在图像组中的目标对象进行的特定动作中在动作期间生成的特定声音，该图像组包括以时间顺序方式连续的多个图像；有效声音范围设置部分，其将在时间轴上用于基于用户操控确定由特定声音检测部分所检测到的特定声音是否有效的范围设置为有效声音范围；以及选择部分，其在所设置的有效声音范围内检测到特定声音的情况下，参考在时间轴上特定声音的检测位置，从时间轴上图像组的选择范围中选择表示特定动作的过渡的预定数量的动作过渡图像。



1. 一种图像处理设备,包括:

特定声音检测部分,其检测在由包括在图像组中的目标对象进行的特定动作中在动作期间生成的特定声音,该图像组包括以时间顺序方式连续的多个图像;

有效声音范围设置部分,其将在时间轴上用于基于用户操控确定由特定声音检测部分所检测到的特定声音是否有效的范围设置为有效声音范围;以及

选择部分,其在所设置的有效声音范围内检测到特定声音的情况下,参考在时间轴上特定声音的检测位置,从时间轴上图像组的选择范围中选择表示特定动作的过渡的预定数量的动作过渡图像。

2. 根据权利要求 1 的设备,

其中,用户操控是在特定动作终止时接收到的特定操控或用于使特定动作的终止时间变成时间计数器的终止时间的设置的计时器设置操控,并且

其中,有效声音范围设置部分在用户操控为特定操控的情况下,参考时间轴上接收到特定操控的位置来设置有效声音范围,而在用户操控为计时器设置操控的情况下,借助于计时器设置操控、参考时间计数器的终止时间来设置有效声音范围。

3. 根据权利要求 2 的设备,

其中,与在用户操控为计时器设置操控的情况下所设置的有效声音范围相比,有效声音范围设置部分将在用户操控为特定操控的情况下所设置的有效声音范围变窄。

4. 根据权利要求 1 的设备,还包括确定部分,其确定在所设置的有效声音范围中检测到的特定声音在时间轴上的检测位置之前和之后的预定范围作为选择范围,

其中,选择部分在所确定的选择范围内选择动作过渡图像。

5. 根据权利要求 1 的设备,还包括:

对象距离计算部分,其计算到目标对象的距离;以及

延迟时间计算部分,其基于所计算出的距离计算所检测到的特定声音的延迟时间,

其中,选择部分基于所计算出的延迟时间校正在时间轴上所检测到的特定声音的检测位置,并且在经校正的检测位置存在于所设置的有效声音范围内的情况下,参考经校正的检测位置,从选择范围中选择动作过渡图像。

6. 根据权利要求 1 的设备,

其中,关于选择范围中动作过渡图像的选择间隔,与其他选择间隔相比,选择部分使得接近在比具有在时间轴上特定声音的检测位置作为中心位置的选择范围更窄的范围内的选择间隔。

7. 根据权利要求 1 的设备,还包括合成图像生成部分,其通过以时间顺序方式安排与合成多个所选择的动作过渡图像,生成表示目标对象的动作过渡的合成图像。

8. 根据权利要求 1 的设备,还包括:

存储部分,其存储涉及多种类型的特定动作的特定声音和有效声音范围;以及

操控接收部分,其接收在其中存储特定声音和有效声音范围的、用于从多种类型的特定动作之中指定想要的特定动作的指定操控,

其中,特定声音检测部分检测涉及所指定的特定动作的特定声音,并且

其中,有效声音范围设置部分基于用户操控设置涉及所指定的特定动作的有效声音范围。

9. 根据权利要求 1 的设备，

其中，在生成图像组时，有效声音范围设置部分基于涉及图像捕获操作的用户操控设置有效声音范围。

10. 一种图像处理设备，其包括：

特定声音检测部分，其检测在包含在由图像组中的目标对象进行的特定动作中在动作期间生成的特定声音，该图像组包括以时间顺序方式连续的多个图像；

特定变化检测部分，其在形成该图像组的各图像之间检测在时间轴上的特定变化；

有效声音范围设置部分，其将在时间轴上用于基于在其中在时间轴上检测到特定变化的位置确定由特定声音检测部分检测到的特定声音是否有效的范围设置为有效声音范围；以及

选择部分，其在所设置的有效声音范围内检测到特定声音的情况下，参考在时间轴上特定声音的检测位置，从时间轴上图像组的选择范围中选择表示特定动作的过渡的预定数量的动作过渡图像。

11. 根据权利要求 9 的设备，

其中，特定变化检测部分使用从形成图像组的每个图像提取的特征量和基于涉及形成图像组的每个图像的声音而提取的特征量中至少之一来检测特定变化。

12. 一种图像处理方法，其包括如下步骤：

基于用户操控，将在时间轴上用于确定在由包括在图像组中的目标对象进行的特定动作中在动作期间生成的特定声音的检测是否有效的范围设置为有效声音范围，该图像组包括以时间顺序方式连续的多个图像；以及

在所设置的有效声音范围内检测到特定声音的情况下，参考在时间轴上特定声音的检测位置，从时间轴上图像组的选择范围中选择表示特定动作的过渡的预定数量的动作过渡图像。

13. 一种使得在计算机上执行过程的程序，该过程包括如下步骤：

基于用户操控，将在时间轴上用于确定在由包括在图像组中的目标对象进行的特定动作中在动作期间生成的特定声音的检测是否有效的范围设置为有效声音范围，该图像组包括以时间顺序方式连续的多个图像；以及

在所设置的有效声音范围内检测到特定声音的情况下，参考在时间轴上特定声音的检测位置，从时间轴上图像组的选择范围中选择表示特定动作的过渡的预定数量的动作过渡图像。

图像处理设备、图像处理方法和程序

技术领域

[0001] 本发明涉及图像处理设备,更具体地说,涉及用于处理以时间顺序方式连续的多个图像的图像处理设备和图像处理方法,以及用于允许在计算机上执行该方法的程序。

背景技术

[0002] 最近几年来,广泛使用了诸如通过成像诸如人之类的对象生成图像数据并将所生成的图像数据记录为诸如图像文件、运动图像文件之类的内容的数字照相机、数字视频摄像机(例如,集成式摄录机(integrated camera recorder))之类的图像捕获设备。此外,作为内容的体育运动(例如,高尔夫球、棒球等)的图像的记录广泛使用这样的图像捕获设备。以这种方式记录的内容(例如,运动图像内容)可以用于例如检查体育运动中的动作(例如,高尔夫球挥杆或击球(棒球))。

[0003] 例如,在确认体育运动动作(例如,高尔夫球挥杆或击球(棒球))的情况下,已经提出了合成图像生成方法,其生成表示动作过渡(motion transition)的图像作为合成图像,以便可以容易地抓住动作过渡。例如,提出了从以预定间隔连续拍摄的图像中选择多个图像并使用该多个图像生成合成图像的图像捕获设备(例如,参看考日本未审专利申请公开第2009-44573号(图9))。

发明内容

[0004] 根据上述相关技术,例如,可以生成表示进行高尔夫球挥杆的人的动作过渡的合成图像。

[0005] 在这个方面,例如,由于在比较短时间内进行高尔夫球挥杆或击球(棒球)的动作,因此当生成表示动作过渡的合成图像时,适当选择形成合成目标图像的图像很重要。然而,在上述相关技术中,参考用户的开始操控的时间确定关于用于生成合成图像的多个图像的选择的范围。因此,例如,在用户不习惯于拍摄或者是初学者的情况下,可能难以适当地进行关于在比较短时间内进行的高尔夫球挥杆或击球(棒球)的开始操控。

[0006] 从而,例如,当选择表示高尔夫球挥杆或击球(棒球)的动作过渡的图像时,选择包括在适当范围内的图像很重要。

[0007] 从而,有必要提供适当地选择表示目标对象进行的特定动作的过渡的图像的技术。

[0008] 根据本发明的一个实施例,提供一种图像处理设备、图像处理方法以及用于允许在计算机上执行该方法的程序,该图像处理设备包括:特定声音检测部分,其检测在由包含在图像组中的目标对象进行的特定动作中在动作期间生成的特定声音,该图像组包括以时间顺序方式连续的多个图像;有效声音范围设置部分,其将在时间轴上用于基于用户操控确定由特定声音检测部分所检测到的特定声音是否有效的范围设置为有效声音范围;以及选择部分,其在所设置的有效声音范围内检测到特定声音的情况下,参考在时间轴上特定声音的检测位置,从时间轴上图像组的选择范围中选择表示特定动作的过渡的预定数量的

动作过渡图像。因此,可以基于用户操控设置有效声音范围,并且在所设置的有效声音范围内检测到特定声音的情况下,参考特定声音的检测位置,从该选择范围中选择预定数量的动作过渡图像。

[0009] 在该实施例中,用户操控可以是在特定动作终止时接收到的特定操控或用于使特定动作的终止时间变成时间计数器的终止时间的设置的计时器设置操控,并且有效声音范围设置部分在用户操控为特定操控的情况下可以参考时间轴上接收到特定操控的位置来设置有效声音范围,而在用户操控为计时器设置操控的情况下可以借助于计时器设置操控、参考时间计数器的终止时间来设置有效声音范围。因此,可以在用户操控为特定操控的情况下参考接收到特定操控的位置来设置有效声音范围,而在用户操控为计时器设置操控的情况下可以借助于计时器设置操控、参考时间计数器的终止时间来设置有效声音范围。

[0010] 在该实施例中,与在用户操控为计时器设置操控的情况下所设置的有效声音范围相比,有效声音范围设置部分将在用户操控为特定操控的情况下所设置的有效声音范围变窄。因此,与在用户操控为计时器设置操控的情况下所设置的有效声音范围相比,在用户操控为特定操控的情况下所设置的有效声音范围变窄。

[0011] 在该实施例中,该设备还可以包括确定部分,其确定在所设置的有效声音范围中检测到的特定声音在时间轴上的检测位置之前和之后的预定范围作为选择范围,并且选择部分可以在所确定的选择范围内选择动作过渡图像。因此,可以确定在所设置的有效声音范围中检测到的特定声音的检测位置之前和之后的预定范围作为选择范围,并且在所确定的选择范围内选择动作过渡图像。

[0012] 在该实施例中,该设备还可以包括:对象距离计算部分,其计算到目标对象的距离;以及延迟时间计算部分,其基于所计算出的距离计算所检测到的特定声音的延迟时间,并且选择部分可以基于所计算出的延迟时间校正在时间轴上所检测到的特定声音的检测位置,并且在经校正的检测位置存在于所设置的有效声音范围内的情况下,参考经校正的检测位置,从选择范围中选择动作过渡图像。因此,可以计算到目标对象的距离,基于所计算出的距离计算所检测到的特定声音的延迟时间,基于所计算出的延迟时间校正所检测到的特定声音的检测位置,并且在经校正的检测位置存在于所设置的有效声音范围内的情况下,参考经校正的检测位置,从选择范围中选择动作过渡图像。

[0013] 在该实施例中,关于选择范围中动作过渡图像的选择间隔,与其他选择间隔相比,选择部分可以压缩选择间隔在比具有在时间轴上特定声音的检测位置作为中心位置的选择范围更窄的范围内。因此,关于选择范围中动作过渡图像的选择间隔,与其他选择间隔相比,可以压缩选择间隔在比具有特定声音的检测位置作为中心位置的选择范围更窄的范围内。

[0014] 在该实施例中,该设备还可以包括合成图像生成部分,其通过以时间顺序方式安排与合成多个所选择的动作过渡图像,生成表示目标对象的动作过渡的合成图像。因此,可以以时间顺序方式安排与合成多个所选择的动作过渡图像,从而生成表示目标对象的动作过渡的合成图像。

[0015] 在该实施例中,该设备还包括:存储部分,其存储涉及多种类型的特定动作的特定声音和有效声音范围;操控接收部分,其接收在其中存储特定声音和有效声音范围的、用于从多种类型的特定动作之中指定想要的特定动作的指定操控。这里,特定声音检测部分可

以检测涉及所指定的特定动作的特定声音,并且有效声音范围设置部分可以基于用户操控设置涉及所指定的特定动作的有效声音范围。因此,当接收到用于从多种类型的特定动作之中指定想要的特定动作的指定操控时,可以检测涉及所指定的特定动作的特定声音,并且基于用户操控设置涉及所指定的特定动作的有效声音范围。

[0016] 在该实施例中,在生成图像组时,有效声音范围设置部分可以基于涉及图像捕获操作的用户操控设置有效声音范围。

[0017] 根据本发明的另一实施例,提供一种图像处理设备、图像处理方法以及用于允许在计算机上执行该方法的程序,该图像处理设备包括:特定声音检测部分,其检测在包含在由图像组中的目标对象进行的特定动作中在动作期间生成的特定声音,该图像组包括以时间顺序方式连续的多个图像;特定变化检测部分,其在形成该图像组的各图像之间检测在时间轴上的特定变化;有效声音范围设置部分,其将在时间轴上用于基于在其中在时间轴上检测到特定变化的位置确定由特定声音检测部分检测到的特定声音是否有效的范围设置为有效声音范围;以及选择部分,其在所设置的有效声音范围内检测到特定声音的情况下,参考在时间轴上特定声音的检测位置,从时间轴上图像组的选择范围中选择表示特定动作的过渡的预定数量的动作过渡图像。因此,可以基于检测到特定变化的位置设置有效声音范围,并且在所设置的有效声音范围内检测到特定声音的情况下,参考特定声音的检测位置,从该选择范围中选择预定数量的动作过渡图像。

[0018] 在该实施例中,特定变化检测部分可以使用从形成图像组的每个图像提取的特征量和基于涉及形成图像组的每个图像的声音而提取的特征量中至少之一来检测特定变化。因此,可以使用从形成图像组的每个图像提取的特征量和基于涉及形成图像组的每个图像的声音而提取的特征量中至少之一来检测特定变化。

[0019] 根据本发明的各实施例,可以适当地选择表示目标对象进行的特定动作的过渡的图像。

附图说明

[0020] 图 1 是说明根据本发明的第一实施例的、图像捕获设备的功能性配置的示例的方块图。

[0021] 图 2 是说明根据本发明的第二实施例的、选择部分和图层处理部分的功能性配置的示例的方块图。

[0022] 图 3A、3B 和 3C 是示意地说明根据本发明的第一实施例的、图像捕获设备与作为图像捕获目标的人之间的位置关系以及从该位置关系生成的动作图像之间的关系图。

[0023] 图 4A 和 4B 是说明根据本发明的第一实施例的、由合成目标图像选择部分选择的合成目标图像和由图层处理部分生成的合成图像的示例图。

[0024] 图 5A、5B 和 5C 是示意地说明根据本发明的第一实施例的、有效声音范围设置部分的有效声音范围的设置方法以及合成目标图像选择范围确定部分的合成目标图像选择范围的确定方法的图。

[0025] 图 6A 和 6B 是说明根据本发明的第一实施例的、显示在显示部分上的显示屏幕的过渡的示例图。

[0026] 图 7A、7B 和 7C 是示意地说明根据本发明的第一实施例的、有效声音范围设置部分

的有效声音范围的设置方法和合成目标图像选择范围确定部分的合成目标图像选择范围的确定方法的图。

[0027] 图 8A、8B、8C 和 8D 是示意地说明根据本发明的第一实施例的、选择部分的合成目标图像的选择处理和图层处理部分的合成图像的生成处理的流程的示例的图。

[0028] 图 9A 和 9B 是示意地说明根据本发明的第一实施例的、图层分离部分的合成目标图像的分离方法的图。

[0029] 图 10A、10B 和 10C 是示意地说明根据本发明的第一实施例的、作为图层修改部分的修改处理的目标的图像和由图层修改设备修改了的图像的图。

[0030] 图 11A 和 11B 是示意地说明根据本发明的第一实施例的、图层合成部分的合成方法的图。

[0031] 图 12A 和 12B 是示意地说明根据本发明的第一实施例的、图层合成部分的合成目标图像的合成方法的图。

[0032] 图 13 是说明根据本发明的第一实施例的、由图层处理部分生成的合成图像的示例的图。

[0033] 图 14 是说明根据本发明的第一实施例的、图像捕获设备的合成图像生成处理的处理过程的示例的流程图。

[0034] 图 15 是说明根据本发明的第一实施例的、图像捕获设备的合成图像生成处理的处理过程的示例的流程图。

[0035] 图 16A 和 16B 是说明根据本发明的第二实施例的、图像捕获设备与目标对象之间的距离和碰撞声音的生成位置及其达到位置之间的关系关系的图。

[0036] 图 17 是说明根据本发明的第二实施例的、图像捕获设备的功能性配置的示例的方块图。

[0037] 图 18A 和 18B 是示意地说明根据本发明的第二实施例的、作为图像捕获部分的图像捕获目标的图像捕获范围以及显示在显示部分上的捕获图像的图。

[0038] 图 19 是示意地说明根据本发明的第二实施例的、作为图像捕获部分的图像捕获目标的图像捕获范围与显示在显示部分上的捕获图像之间的位置关系的顶视图。

[0039] 图 20A 和 20B 是示意地说明根据本发明的第二实施例的、有效声音范围设置部分的有效声音范围的设置方法和合成目标图像选择范围确定部分的合成目标图像选择范围的确定方法的图。

[0040] 图 21 是说明根据本发明的第二实施例的、显示部分中的显示示例的图。

[0041] 图 22 是说明根据本发明的第二实施例的、图像捕获设备的合成图像生成处理的处理过程的示例的流程图。

[0042] 图 23 是说明根据本发明的第三实施例的、图像捕获设备的功能性配置的示例的方块图。

[0043] 图 24 是说明根据本发明的第三实施例的、目标对象信息存储部分的存储内容的示例的图。

[0044] 图 25 是说明根据本发明的第三实施例的、显示部分的显示示例（显示屏幕）的图。

[0045] 图 26A 和 26B 是说明根据本发明的第三实施例的、由图层处理部分生成的合成图

像的示例的图。

[0046] 图 27 是说明根据本发明的第四实施例的、图像捕获设备的功能性配置的示例的方块图。

[0047] 图 28 是说明根据本发明的第四实施例的、目标对象信息存储部分的存储内容的示例的图。

[0048] 图 29A 和 29B 是示意地说明根据本发明的第四实施例的、有效声音范围设置部分的有效声音范围的设置方法和合成目标图像选择范围确定部分的合成目标图像选择范围的确定方法的图。

[0049] 图 30A 和 30B 是示意地说明根据本发明的第四实施例的、有效声音范围设置部分的有效声音范围的设置方法和合成目标图像选择范围确定部分的合成目标图像选择范围的确定方法的图。

具体实施方式

[0050] 在下文中,将按照以下顺序描述实现本发明的实施例:

[0051] 1. 第一实施例(合成图像生成控制:基于用户操控设置有效声音范围的示例)

[0052] 2. 第二实施例(合成图像生成控制:计算对象距离并校正碰撞声音检测位置的示例)

[0053] 3. 第三实施例(合成图像生成控制:从多个对象中指定想要的对象并生成涉及该对象的合成图像的示例)

[0054] 4. 第四实施例(合成图像生成控制:以所记录的图像内容生成合成图像的示例)

[0055] 1. 第一实施例

[0056] [图像捕获设备的配置示例]

[0057] 图 1 是说明根据本发明的第一实施例的、图像捕获设备 100 的功能性配置的示例的方块图。图像捕获设备 100 包括图像捕获部分 111、对象信息生成部分 120、捕获图像保持部分 115、声音输入部分 131、特定声音检测部分 132、操控接收部分 140 和计时器设置部分 150。此外,图像捕获设备 100 包括有效声音范围设置部分 160、选择部分 170、图层处理部分 180、合成图像保持部分 185、显示控制部分 190、显示部分 191 和合成图像存储部分 200。例如,图像捕获设备 100 可以用数字视频摄像机来实现,该数字视频摄像机可以成像对象来生成捕获图像(图像数据)、通过图像分析提取关于该图像数据的每个特征量并使用每一个所提取的特征量进行各种图像处理。此外,图像捕获设备 100 是公开在权利要求书中的图像处理设备的示例。

[0058] 图像捕获部分 111 包括:图像捕获元件(例如,如图 19 所示的图像捕获元件 112),其将通过镜头(未示出)入射的对象的光转换成电信号;以及信号处理部分(未示出),其处理图像捕获元件的输出信号来生成捕获图像(图像数据)。也就是说,在图像捕获部分 111 中,通过镜头入射的对象的光学图像被形成在图像捕获元件的成像表面上,图像捕获元件在这种状态下进行图像捕获操作,而且该信号处理部分对于成像信号进行信号处理,因此,生成捕获图像。基于从操控接收部分 140 或计时器设置部分 150 输出的图像捕获操作的开始指令信息来生成捕获图像。此外,将所生成的捕获图像供应到并保持在捕获图像保持部分 115 中。另外,将所生成的捕获图像供应到显示控制部分 190,然后显示在显示部分

191 上。此外,将所生成的捕获图像以及在捕获图像生成时每个镜头(例如,聚焦镜头和变焦镜头)的位置和聚焦位置供应到对象信息生成部分 120。

[0059] 捕获图像保持部分 115 是用于保持由图像捕获部分 111 生成的捕获图像预定时间的环形缓存器,并将所保持的捕获图像供应到选择部分 170。此外,捕获图像保持部分 115 与图像捕获部分 111 生成的捕获图像相关联地保持对象信息生成部分 120 生成的对象信息。这里,例如,捕获图像被保持在捕获图像保持部分 115 中的预定时间可以是 3 至 10 秒。也就是说,将图像捕获部分 111 所生成的最新捕获图像和对象信息生成部分 120 所生成的对象信息在捕获图像保持部分 115 中保持 3 至 10 秒。

[0060] 对象信息生成部分 120 基于图像捕获部分 111 所生成的捕获图像和在捕获图像生成时的每个信息(例如,镜头的位置和聚焦位置),生成涉及捕获图像的对象信息。此外,对象信息生成部分 120 将所生成的对象信息供应到捕获图像保持部分 115,并允许将其保持在其中。例如,对象信息是用于从包括在图像捕获部分 111 所生成的捕获图像中的对象中辨别目标对象区域和其背景区域的信息。例如,生成景深图(depth map)作为对象信息。这里,景深图是指示从图像捕获位置(例如,图像捕获设备 100 的位置)到包括在捕获图像中的对象的距离(对象距离)的数据。例如,可以使用诸如 TOF(飞行时间(time of flight))、与散焦的景深(depth from defocus)之类的技术作为生成景深图的技术。例如,TOF 技术是用于基于从对象反射从光源发射的光并到达传感器的延迟时间以及光速度计算到对象的距离的技术。

[0061] 例如,声音输入部分 131 获得图像捕获设备 100 周围的声音,并将所获得的声音(声音数据)输出到特定声音检测部分 132。用多个麦克风等实现声音输入部分 131。

[0062] 特定声音检测部分 132 基于从声音输入部分 131 输出的声音来检测特定声音,并且在检测到特定声音的情况下,将其输出到选择部分 170。在本发明第一实施例中示范了在高尔夫球挥杆期间高尔夫球棒接触高尔夫球的瞬间生成的声音(碰撞声音)为特定声音的情况。例如,特定声音检测部分 132 预先存储涉及特定声音的特征量作为特定声音标识信息(参考数据)。然后,特定声音检测部分 132 从声音输入部分 131 输出的声音中提取特征量,并将所提取的特征量与特定声音标识信息比较以计算相似性。在所计算出的相似性超过阈值的情况下,特定声音检测部分 132 确定从声音输入部分 131 输出的声音是特定声音。具体地说,从声音输入部分 131 输出的声音数据通过 AD(模-数)转换处理采样,并转换为数字数据。此外,以适当时间间隔对数字数据进行诸如频率分析之类的处理,以便将该数字数据转换为指示声音的谱或其他声学特征的参数。因此,提取了声音的时间顺序特征量。此外,使用所保持的参考数据进行与所提取的时间顺序特征量的匹配处理,然后将声音识别结果输出作为匹配处理的结果。可以使用各种其他现存方法进行声音分析和识别。

[0063] 操控接收部分 140 接收由用户操控的操控内容,并将基于接收到的操控内容的操控信号输出到图像捕获部分 111、计时器设置部分 150、有效声音范围设置部分 160 和显示控制部分 190。例如,当接收到指令用于生成合成图像的合成图像记录方式的设置的合成图像记录方式设置操控时,操控接收部分 140 将用于指令图像捕获操作的开始的控制信号输出到图像捕获部分 111。然后,操控接收部分 140 将用于指令捕获图像的显示开始的控制信号输出到显示控制部分 190。例如,通过合成图像记录方式的设置按钮的按压操控来进行合成图像记录方式的设置操控。此外,例如,当接收到用于指令合成图像记录方式的释放的

合成图像记录方式释放操控时,操控接收部分 140 将用于停止图像捕获操作的控制信号输出到图像捕获部分 111。例如,通过合成图像记录方式的释放按钮进行合成图像记录方式释放操控。此外,例如,当在设置了合成图像记录方式的情况下接收到合成图像的记录指令操控时,操控接收部分 140 将指示接收到合成图像的记录指令操控的控制信号输出到有效声音范围设置部分 160。然后,操控接收部分 140 将用于停止图像捕获操作的控制信号输出到图像捕获部分 111。例如,通过视频记录按钮(REC 按钮)进行合成图像的记录指令操控。此外,例如,当接收到用于通过计时器设置在合成图像记录方式下进行图像捕获操作的计时器设置操控时,操控接收部分 140 将用于指令计时器设置的控制信号输出到计时器设置部分 150。例如,通过计时器设置按钮进行计时器设置操控。此外,当按压计时器设置按钮时,可以进行计数器值(例如,8 至 17 秒的设置)的设置操控。这些操控可以使用诸如配备在图像捕获设备 100 的按钮之类的操控构件输入,并且可以通过经由遥控器操控的无线通信等进行。

[0064] 计时器设置部分 150 设置用于在合成图像记录方式下进行图像捕获操作的计时器(例如,自拍器(self timer)功能)。具体地说,如果操控接收部分 140 接收到计时器设置操控,则计时器设置部分 150 将用于指令图像捕获操作的开始的控制信号输出到图像捕获部分 111。此外,计时器设置部分 150 将用于指令捕获图像的显示开始的控制信号以及由计时器设置引起的计数值输出到显示控制部分 190。另外,在计数值的显示开始之后,计时器设置部分 150 顺序输出根据设置内容的计数值到显示控制部分 190。因此,如图 6A 和 6B 所示,图像捕获部分 111 所生成的捕获图像被显示在显示部分 191 上,并且计数值被重叠和显示在捕获图像上。例如,将每一秒递增一地加起来的计数值显示在显示部分 191 上。此外,在所设置的计数器的计数值为“零”的情况下计时器设置部分 150 将指示该事实的控制信号输出到有效声音范围设置部分 160。此外,计时器设置部分 150 将用于停止图像捕获操作的控制信号输出到图像捕获部分 111。

[0065] 在设置了合成图像记录方式的情况下,有效声音范围设置部分 160 设置用于确定特定声音检测部分 132 检测到的特定声音是否有效的有效声音范围。此外,有效声音范围设置部分 160 将关于所设置的有效声音范围的信息(有效声音范围和在设计该范围时用作参考的时间轴上的位置)输出到选择部分 170。具体地说,在操控接收部分 140 接收到合成图像的记录指令操控的情况下,或者在计时器设置部分 150 所设置的计时器的计数值变成“零”的情况下,有效声音范围设置部分 160 设置有效声音范围。这里,在操控接收部分 140 接收到合成图像的记录指令操控的情况下,接收到合成图像的记录指令操控的时间轴上的位置变成参考位置。此外,计时器设置部分 150 所设置的计时器的计数值变成“零”,计数值变成“零”的时间轴上的位置变成参考位置。将参考图 5A 至 7C 等详细描述有效声音范围的这些设置方法。

[0066] 选择部分 170 从保持在捕获图像保持部分 115 中的捕获图像中,选择在借助于图层处理部分 180 的合成图像的生成处理中使用的合成目标图像。此外,选择部分 170 将关于所选择的合成目标图像的信息(例如,合成目标图像、关于这些图像的对象信息以及在合成图像的生成处理中使用的坐标)输出到图层处理部分 180。将参考图 2 详细描述选择部分 170 的内部配置。此外,合成目标图像是公开在权利要求书中的动作过渡图像的示例。

[0067] 图层处理部分 180 使用选择部分 170 所选择的合成目标图像生成合成图像,然后

将所生成的合成图像存储在合成图像存储部分 200 中。将参考图 2 详细描述图层处理部分 180 的内部配置。此外, 图层处理部分 180 是公开在权利要求书中的合成图像生成部分的示例。

[0068] 合成图像保持部分 185 在借助于图层处理部分 180 的合成图像的生成处理时在合成处理期间保持合成图像(历史图像), 并将所保持的合成图像供应到图层处理部分 180。将参考图 2 详细描述合成图像保持部分 185。

[0069] 显示控制部分 190 根据操控接收部分 140 接收到的操控输入, 将存储在合成图像存储部分 200 中的合成图像或从图像捕获部分 111 输出的捕获图像显示在显示部分 191 上。此外, 显示控制部分 190 将计时器设置部分 150 所设置的计时器的计数值与捕获图像重叠, 以显示在显示部分 191 上。

[0070] 显示部分 191 在显示控制部分 190 的控制下显示每一个图像。

[0071] 合成图像存储部分 200 存储图层处理部分 180 所生成的合成图像, 并将所存储的合成图像供应到显示控制部分 190。

[0072] 图 2 是说明根据本发明的第二实施例的、选择部分 170 和图层处理部分 180 的功能性配置的示例的方块图。选择部分 170 包括合成目标图像选择范围确定部分 171、合成目标图像选择部分 172、坐标计算部分 173 以及计数器部分 174。此外, 图层处理部分 180 包括图层分离部分 181、图层修改部分 182 以及图层合成部分 183。

[0073] 合成目标图像选择范围确定部分 171 确定用于在保持在捕获图像保持部分 115 中的捕获图像之中, 选择在借助于图层处理部分 180 的合成图像的生成处理中使用的合成目标图像的范围(合成目标图像选择范围)。具体地说, 合成目标图像选择范围确定部分 171 基于特定声音检测部分 132 所检测到的特定声音的检测位置(时间轴上的位置)以及有效声音范围设置部分 160 所设置的有效声音范围, 确定合成目标图像选择范围。例如, 在特定声音的检测位置包括在有效声音范围中的情况下, 合成目标图像选择范围确定部分 171 基于特定声音的检测位置, 确定合成目标图像选择范围。另一方面, 在特定声音的检测位置不包括在有效声音范围中的情况下, 合成目标图像选择范围确定部分 171 基于在设置有效声音范围时变成参考的时间轴上的位置, 确定合成目标图像选择范围。这里, 操控接收部分 140 接收到合成图像的记录指令操控的情况下, 接收到合成图像的记录指令操控的时间轴上的位置变成参考位置。此外, 在计时器设置部分 150 所设置的计时器的计数值变成“零”的情况下, 计数值变成“零”的时间轴上的位置变成参考位置。此外, 合成目标图像选择范围确定部分 171 将关于所确定的合成目标图像选择范围的信息(例如, 合成目标图像选择范围和关于包括在该范围中的图像的对象信息)输出到合成目标图像选择部分 172。合成目标图像选择范围确定部分 171 是公开在权利要求书中的确定部分的示例。

[0074] 合成目标图像选择部分 172 从包括在合成目标图像选择范围确定部分 171 所确定的合成目标图像选择范围中的相应图像中, 选择在借助于图层处理部分 180 的合成图像的生成处理中使用的合成目标图像。例如, 合成目标图像选择部分 172 从包括在合成目标图像选择范围确定部分 171 所确定的合成目标图像选择范围中的相应图像(保持在捕获图像保持部分 115 中的捕获图像)中, 以预定时间间隔选择图像作为合成目标图像。此外, 例如, 合成目标图像可以被选择成具有适合于肉眼观察目标对象的动作过渡的间隔。此外, 合成目标图像选择部分 172 将关于所选择的合成目标图像的信息(例如, 合成目标图像, 关于这

些图像的对象信息)输出到坐标计算部分 173 和图层分离部分 181。

[0075] 关于从合成目标图像选择部分 172 输出的每一个合成目标图像,坐标计算部分 173 计算在图层处理部分 180 的合成图像的生成处理中使用的坐标,并将所计算出的坐标输出到图层修改部分 182。在这种坐标计算中,计算要与涉及作为计算目标的合成目标图像的合成图像合成的像素位置。也就是说,在这种坐标计算中,与最后生成的合成图像(例如,如图 8D 所示的合成图像 402)之中、作为计算目标的合成目标图像对应地计算像素范围(包括如图 8D 所示的 F0 和 F10 的范围)。

[0076] 计数器部分 174 将在合成目标图像选择范围确定部分 171 的合成目标图像选择范围确定时使用的计数器(碰撞声音检测计数器)的值供应到合成目标图像选择范围确定部分 171。这里,计数器部分 174 连续地增加碰撞声音检测计数器。此外,在从特定声音检测部分 132 检测到特定声音的事实被输出到合成目标图像选择范围确定部分 171 的情况下,由合成目标图像选择范围确定部分 171 将碰撞声音检测计数器的值重置为“零”。类似地,在重置后,计数器部分 174 还连续增加碰撞声音检测计数器。此外,合成目标图像选择范围确定部分 171 在合成目标图像选择范围的确定时,使用碰撞声音检测计数器的值估计碰撞声音检测位置。也就是说,合成目标图像选择范围的确定时,将比确定时间在前碰撞声音检测计数器的值的位置(在时间轴上的在前位置)估计为碰撞声音检测位置。

[0077] 对于合成目标图像选择部分 172 所选择的合成目标图像,图层分离部分 181 使用对应的对象信息进行图层分离。这里,术语“层”指的是通过对每一个图像(帧)分离目标对象部分和背景部分而获得的每一个图像。此外,图层分离部分 181 将所分离的每一图层图像输出到图层修改部分 182。

[0078] 图层修改部分 182 使用坐标计算部分 173 所计算出的坐标,进行用于将从图层分离部分 181 输出的相应图层图像生成合成图像的各种修改处理。作为这种修改处理,进行诸如剪切、缩放、旋转和坐标移动之类的几何算术处理。例如,缩放让其处理内容根据合成目标图像的数量、合成图像的尺寸等来确定。此外,作为修改处理,可以进行诸如强调目标对象的动作部分之类的图像处理。另外,图层修改部分 182 将经处理的图层图像输出到图层合成部分 183。

[0079] 图层合成部分 183 对从图层修改部分 182 输出的图层图像进行图像合成处理,并存储所生成的合成图像到合成图像存储部分 200。具体地说,图层合成部分 183 合成相应图层图像,以便将从图层修改部分 182 输出的图层图像安排在对应的像素范围中。此外,图层合成部分 183 合成紧接在合成目标之前的图层图像,然后在合成图像保持部分 185 顺序保持合成图像(历史图像)。此外,当进行变成下一合成目标的图层图像合成处理时,图层合成部分 183 从合成图像保持部分 185 获得保持在合成图像保持部分 185 中的合成图像(历史图像),并且在合成处理中使用它们。

[0080] 合成图像保持部分 185 顺序保持图层合成部分 183 所生成的合成图像(历史图像),并将所保持的合成图像供应到图层合成部分 183。也就是说,将图层合成部分 183 所生成的合成图像顺序更新并保持在合成图像保持部分 185 中。

[0081] [合成图像生成处理的过渡的示例]

[0082] 接下来,简短地描述在生成合成图像的情况下生成处理的过渡的示例。

[0083] 图 3A 至 3C 是示意地说明根据本发明的第一实施例的、图像捕获设备 100 与作为

图像捕获目标的人 300 的位置关系以及从该位置关系生成的动作图像的关系的图。

[0084] 图 3A 是示意地说明图像捕获设备 100 与作为关于图像捕获设备 100 的图像捕获目标的人 300 之间的位置关系的图。例如,人 300 在高尔夫练习场中进行高尔夫球棒 301 的练习挥杆。当人 300 从如图 3A 所示的姿势挥动高尔夫球棒 301 时,人 300 用高尔夫球棒 301 打击高尔夫球 302 来以想要的方向驱动高尔夫球 302。

[0085] 图 3B 示意地说明由图像捕获部分 111 以矩形形状生成的运动图像 310。此外,在图 3B 中,在与运动图像 310 对应的矩形形状中,沿着时间轴安排形成运动图像 310 的多帧之中的某些帧(图像 311 至 313 等)。此外,运动图像 310 是通过在如图 3A 所示的状态下通过由图像捕获设备 100 成像人 300 的高尔夫练习挥杆的样子而获得的运动图像。这里,图像 311 是通过成像人 300 准备好高尔夫球挥杆的状态而获得的图像,而图像 312 是通过成像人 300 观察高尔夫球 302 在高尔夫球挥杆完成之后飞行的方向的状态而获得的图像。此外,图像 313 是在人 300 进行高尔夫球挥杆时、通过成像高尔夫球棒 301 接触高尔夫球 302 的瞬间的状态而获得的图像。这里,当人 300 进行高尔夫球挥杆时,在高尔夫球棒 301 接触高尔夫球 302 的瞬间生成特定声音(碰撞声音)。将生成碰撞声音的运动图像 310 中的位置表示为碰撞声音生成位置 315。

[0086] 在使用这样生成的运动图像 310 生成指示人 300 的高尔夫球挥杆的过渡的合成图像的情况下,例如,选择用于选择合成目标图像的预定范围(合成目标图像选择范围)。合成目标图像选择范围是包括例如从高尔夫球挥杆开始到其结束的一系列动作过渡的范围。此外,在该预定范围内,选择并合成作为合成目标的图像。图 4B 示出了这种合成示例。

[0087] 例如,如图 3B 所示,在运动图像 310 中确定合成目标图像选择范围 320,并将包括在合成目标图像选择范围 320 中的相应帧之中满足预定条件的帧确定为合成目标图像。

[0088] 在图 3C 中,沿着时间轴安排包括在如图 3B 所示的运动图像 310 中确定的合成目标图像选择范围 320 中的相应帧之中的某些帧(图像 321 至 326)。这里,运动图像 321 是人 300 开始高尔夫球挥杆时捕获的图像,而图像 326 是人 300 完成高尔夫球挥杆时捕获的图像。此外,图像 323 是与如图 3B 所示的图像 313 对应的图像。另外,图像 322、324 和 325 是在图像 321 和 326 之间、通过以时间顺序方式顺序成像进行高尔夫球挥杆的人 300 的动作过渡而获得的图像。

[0089] 图 4A 和 4B 是说明根据本发明的第一实施例的、由合成目标图像选择部分 172 选择的合成目标图像和由图层处理部分 180 生成的合成图像的示例的图。在图 4A 中,以时间顺序方式安排被选为包括在如图 3B 所示的运动图像 310 中的合成目标图像选择范围 320 中的相应帧之中的合成目标图像的 24 个图像。在如图 4A 所示的 24 个图像之中包括与如图 3C 所示的图像 321 至 326 相同的对象的图像被赋予相同的参考号。这里,作为合成目标图像的选择方法,例如,可以使用用于以预定间隔选择帧作为合成目标图像的选择方法。此外,例如,可以认为在某些情况下精确地确认高尔夫球棒 301 接触高尔夫球 302 的时间周围的动作过渡。在这种情况下,生成碰撞声音时的图像(例如,图像 323)的邻近图像的间隔可以比其他间隔更接近,以选择合成目标图像。

[0090] 图 4B 示意地说明图层处理部分 180 所生成的合成图像 330。按照这种方式,确定包括从人 300 的高尔夫球挥杆开始到其结束的动作过渡的合成目标图像选择范围,并且在合成目标图像选择范围内选择合成目标图像以便合成。因此,可以生成指示人 300 的高尔

夫球挥杆的动作过渡的合成图像。

[0091] 如上所述,在从运动图像 310 生成合成图像的情况下,适当地选择用于合成目标图像的选择的合成目标图像选择范围很重要。然而,由于在比较短时间内进行高尔夫球挥杆,因此可能难以从运动图像 310 适当地选择合成目标图像选择范围。也就是说,可能难以从包括范围从包括人 300 的高尔夫球挥杆开始之前的状态的图像到包括高尔夫球挥杆结束之后的状态的图像的图像的运动图像 310 中选择合成目标图像选择范围。

[0092] 这里,如上所述,在高尔夫球挥杆中间生成碰撞声音。按照这种方式,由于在高尔夫球挥杆期间频繁地在预定位置上生成碰撞声音,因此可以基于碰撞声音选择合成目标图像选择范围。然而,在人 300 在有很多其他人的高尔夫练习场中练习的情况下,当人 300 准备好高尔夫球挥杆等时,非常可能由人 300 之外的人的高尔夫球挥杆生成碰撞声音。也就是说,在适当地记录涉及引起碰撞声音的体育运动的即时 (instant) 动作的情况下,如果所检测到的碰撞声音总是有效,那么非常可能错误地检测到来自检测目标之外的人的碰撞声音。由此,在生成涉及人 300 的合成图像的情况下,在基于碰撞声音选择合成目标图像选择范围时适当地检测人 300 的高尔夫球挥杆所生成的碰撞声音很重要。因此,在本发明的第一实施例中,将提供有效声音范围的、在其中只有基于用户操控的拍摄触发周围的邻近定时有效的示例设置成减少碰撞声音的错误检测。仅仅使用在有效声音范围中检测到的碰撞声音设置合成目标图像选择范围,从而选择合适的合成目标图像。

[0093] [基于视频记录按钮操控的合成目标图像选择范围的确定示例]

[0094] 图 5A、5B 和 5C 是示意地说明根据本发明的第一实施例的、有效声音范围设置部分 160 的有效声音范围的设置方法以及合成目标图像选择范围确定部分 171 的合成目标图像选择范围的确定方法的图。

[0095] 图 5A 以矩形形状示意地说明图像捕获部分 111 所生成的运动图像 350。此外,按照与如图 3B 所示的示例类似的方式,运动图像 350 是通过由图像捕获设备 100 成像人 (例如, Koda Goro) 的高尔夫练习挥杆的样子而获得的运动图像。这里,假设 Koda Goro 要求他的朋友 (例如, Otogawa Rokuro) 拍摄他的高尔夫球挥杆。在这种情况下, Otogawa Rokuro 用手握住图像捕获设备 100, 其光轴方向朝向 Koda Goro, 并且按压合成图像记录方式的设置按钮。如果操控接收部分 140 接收到按压操控, 那么图像捕获部分 111 生成捕获图像, 并将所生成的捕获图像顺序地保持在成像图像保持部分 115。此外, 显示控制部分 190 在显示部分 191 上顺序显示所生成的捕获图像 (所谓直接图像 (through image))。按照这种方式, 在 Otogawa Rokuro 使用设置了合成图像记录方式的图像捕获设备 100 进行拍摄的状态下, Koda Goro 进行高尔夫球挥杆。在 Koda Goro 刚刚完成高尔夫球挥杆时, Otogawa Rokuro 迅速按压视频记录按钮。如果操控接收部分 140 接收到按压操控, 那么有效声音范围设置部分 160 设置有效声音范围。

[0096] 例如, 假设 Otogawa Rokuro 按压视频记录按钮的位置 (时间轴上的位置) 是如图 5A 所示的运动图像 350 的时间轴上的视频记录按钮的按压位置 (视频记录按钮按压位置 351)。在这种情况下, 有效声音范围设置部分 160 基于视频记录按钮按压位置 351 设置有效声音范围 352。具体地说, 有效声音范围设置部分 160 设置在时间轴上领先视频记录按钮按压位置 351 时间 L1 的范围, 作为有效声音范围 352。在这个示例中, 由于可以在一系列高尔夫球挥杆动作结束之后生成视频记录按钮的按压 (拍摄触发), 因此可以在拍摄触发的

定时之前存在碰撞声音的生成位置。因此,这个示例是在视频记录按钮的按压之前设置有效声音范围的示例。

[0097] 这里,例如,时间 L1 可以是 1.0 秒。基于在以这种方式设置的有效声音范围检测到的碰撞声音确定合成目标图像选择范围。将参考图 5B 描述合成目标图像选择范围的确定方法。此外,在未在以这种方式设置的有效声音范围检测到的碰撞声音的情况下,基于视频记录按钮按压位置 351 确定合成目标图像选择范围。参考图 5C 描述合成目标图像选择范围的确定方法。

[0098] 图 5B 示意地说明在由有效声音范围设置部分 160 设置的有效声音范围 352 中检测到碰撞声音的情况下合成目标图像选择范围的确定方法。在图 5B 中,将特定声音检测部分 132 检测到碰撞声音的位置(时间轴上的位置)表示为碰撞声音检测位置 353。如图 5B 所示,在由有效声音范围设置部分 160 设置的有效声音范围 352 中检测到碰撞声音的情况下,合成目标图像选择范围确定部分 171 基于碰撞声音检测位置 353 确定合成目标图像选择范围。也就是说,合成目标图像选择范围确定部分 171 参考碰撞声音检测位置 353 将在时间轴上碰撞声音检测位置 353 之前和之后的预定范围确定为合成目标图像选择范围 354。具体地说,合成目标图像选择范围确定部分 171 确定时间轴上的一个范围 L4 作为合成目标图像选择范围 354,该范围 L4 包括领先碰撞声音检测位置 353 时间 L2 的范围和超过碰撞声音检测位置 353 时间 L3 的范围。

[0099] 这里,考虑到高尔夫球棒的旋转速度,在高尔夫球棒接触高尔夫球之后的高尔夫球挥杆期间的时间短于高尔夫球棒接触高尔夫球之前的时间。也就是说,在生成碰撞声音之后的高尔夫球挥杆期间的时间短于生成碰撞声音之前的时间。因此,确定合成目标图像选择范围 354,以便碰撞声音检测位置 353 之前的时间 L2 长于这以后的时间 L3。例如,时间 L2 可以是 1.5 秒,而时间 L3 可以是 0.5 秒。按照这种方式,当基于在有效声音范围中检测到的碰撞声音确定合成目标图像选择范围时,可以确定合适的合成目标图像选择范围。

[0100] 这里,假设在由有效声音范围设置部分 160 设置的有效声音范围 352 之外的位置(例如,由如图 5B 所示的箭头 358 和 359 所指示的位置)中检测碰撞声音。在有效声音范围 352 之外的位置中检测到碰撞声音的情况下,碰撞声音不被用于合成目标图像选择范围的确定。如图 5C 所示,在不使用碰撞声音的情况下,可以基于视频记录按钮的按压位置确定合成目标图像选择范围。此外,在有效声音范围 352 之外的位置中检测到碰撞声音的情况下,碰撞声音不被用于合成目标图像选择范围的确定,但可以用于下一合成图像的生成处理。

[0101] 图 5C 示意地说明不在由有效声音范围设置部分 160 设置的有效声音范围 352 中检测到碰撞声音的情况下合成目标图像选择范围的确定方法。在图 5C 中,按照与图 5A 的情况类似的方式,将按压视频记录按钮的位置表示为视频记录按钮按压位置 351。如图 5C 所示,在未在有效声音范围 352 中检测到碰撞声音的情况下,合成目标图像选择范围确定部分 171 基于视频记录按钮按压位置 351 确定合成目标图像选择范围。也就是说,合成目标图像选择范围确定部分 171 参考视频记录按钮按压位置 351,将在时间轴上视频记录按钮按压位置 351 之前的预定范围确定为合成目标图像选择范围 355。具体地说,合成目标图像选择范围确定部分 171 确定在时间轴上领先碰撞声音检测位置 353 时间 L7 (L5+L6) 的范围,作为合成目标图像选择范围 355。按照这种方式,在未在有效声音范围 352 中检测到

碰撞声音的情况下,将在时间轴上领先视频记录按钮的按压位置时间 L6 的位置认为碰撞声音检测位置,以确定合成目标图像选择范围 355。也就是说,时间 L5 和 L6 对应于如图 5B 所示的时间 L2 和 L3。这里,例如,按照与图 5B 的情况类似的方式,时间 L7 可以是 2.0 秒。也就是说,时间 L5 为 1.5 秒,而时间 L6 为 0.5 秒。然而,在基于视频记录按钮按压位置确定合成目标图像选择范围的情况下,由于基于用户操控确定合成目标图像选择范围,因此与如图 5B 所示的合成目标图像选择范围 354 相比,精确度可能降低。因此,在基于视频记录按钮按压位置确定合成目标图像选择范围的情况下,例如,时间 L7 可以长于时间 L4,以确定合成目标图像选择范围。

[0102] 此外,这些时间 L1 至 L7 可以通过用户操控改变。

[0103] [基于计时器的合成目标图像选择范围的确定示例]

[0104] 图 6A 和 6B 是说明根据本发明的第一实施例的、显示在显示部分 191 上的显示屏幕的过渡的示例的图。图 6A 说明其中由计时器设置部分 150 设置的计时器的计数值“17”(372)与人 300 的高尔夫球挥杆开始之前由图像捕获部分 111 生成的图像重叠的显示示例(显示屏幕 370)。图 6B 说明其中由计时器设置部分 150 设置的计时器的计数值“0”(373)与人 300 的高尔夫球挥杆结束时由图像捕获部分 111 生成的图像重叠的显示示例(显示屏幕 371)。

[0105] 在这个示例中,假设人 300(例如,Koda Goro)自己拍摄他的高尔夫练习挥杆的样子。在这种情况下,例如,正在练习的 Koda Goro 将图像捕获设备 100 安放成使得其光轴方向定向到他自己(如图 3A 所示),然后按压计时器设置按钮。通过这种按压操控,图像捕获部分 111 生成捕获图像,而显示控制部分 190 在显示部分 191 上显示所生成的捕获图像(所谓直接图像)。此外,计时器设置部分 150 通过计时器设置按钮的按压操控设置计时器,显示控制部分 190 将与所设置的内容对应的计数值(例如,17 秒)和直接图像重叠,以显示在显示部分 191 上。例如,如图 6A 所示,在计时器设置按钮的按压之后,显示显示屏幕 370,在其中重叠了由计时器设置部分 150 设置的计时器的计数值“17”(372)和包括准备好高尔夫球挥杆的人 300 的捕获图像。之后,类似地,在显示部分 191 上显示一个显示屏幕,在其中计时器的计数值(以一秒为间隔从 17 秒减少的值)与包括人 300 的捕获图像重叠。人 300 在以这种方式观察显示在显示部分 191 上的显示屏幕的同时,进行高尔夫球挥杆,以便在计时器的计数值变成“0”的时间周围终止高尔夫球挥杆。此外,在计时器的计数值为“0”的情况下,有效声音范围设置部分 160 设置有效声音范围。然后,在预定时间之后停止成像动作。将参考图 7A、7B 和 7C 详细描述有效声音范围的设置方法。

[0106] 当如上所述那样进行计时器设置时,例如,即使在朋友不进行拍摄的情况下,人 300 也可以容易地拍摄他自己的高尔夫球挥杆的样子。

[0107] 图 7A、7B 和 7C 是示意地说明根据本发明的第一实施例的、有效声音范围设置部分 160 的有效声音范围的设置方法和合成目标图像选择范围确定部分 171 的合成目标图像选择范围的确定方法的图。这里,第一实施例提供在其中基于由计时器设置部分 150 设置的计时器来设置有效声音范围的示例。

[0108] 图 7A 以矩形形状示意地说明图像捕获部分 111 所生成的运动图像 380。按照与如图 3B 所示的示例类似的方式,运动图像 380 是通过由图像捕获设备 100 成像人(例如,Koda Goro)的高尔夫练习挥杆的样子而获得的运动图像。例如,如图 6A 和 6B 所示,在人 300 自

已拍摄他的高尔夫练习挥杆的样子的情况下,假设计时器设置部分 150 所设置的计时器的计数值为“0”。在这种情况下,如上所述,有效声音范围设置部分 160 设置有效声音范围。

[0109] 例如,假设在其中计时器设置部分 150 所设置的计时器的计数值为“0”的位置是如图 7A 所示的运动图像 380 的时间轴上的时间计数器“0”的位置(时间计数器“0”位置 381)。在这种情况下,有效声音范围设置部分 160 基于时间计数器“0”位置 381 设置有效声音范围 382。具体地说,有效声音范围设置部分 160 参考时间计数器“0”位置 381,将时间轴上时间计数器“0”位置 381 之前和之后的预定范围设置为有效声音范围 382。例如,有效声音范围设置部分 160 设置在时间轴上的范围 L13 作为有效声音范围 382,该范围 L13 包括领先时间计数器“0”位置 381 时间 L11 的范围和超过时间计数器“0”位置 381 时间 L12 的范围。

[0110] 这里,人 300 可能难以进行高尔夫球挥杆来在时间计数器“0”的定时处终止该系列高尔夫球挥杆动作。因此,在这个示例中,生成碰撞声音瞬间可以对应于时间计数器“0”的位置,并且有效声音范围被设置成覆盖时间计数器“0”的位置之前和之后的范围。按照这种方式,基于视频记录按钮的按压位置设置有效声音范围的情况以及基于时间计数器“0”的位置设置有效声音范围的情况可以具有不同的设置内容。当根据多个拍摄触发特征设置有效声音范围时,可以减少错误检测到要检测的碰撞声音之外的声音的危险。

[0111] 在这个方面,时间 L11 可以是 0.5 秒,而时间 L12 可以是 0.5 秒,例如。也就是说,例如,有效声音范围 382 的时间 L13 可以是 1.0 秒。人 300 在观察显示在显示部分 191 上的显示屏幕的同时,进行高尔夫球挥杆,以便在计时器的计数值变成“0”的时间周围终止高尔夫球挥杆。由此,与基于按压视频记录按钮的位置设置有效声音范围的情况相比,可能降低精确度。因此,基于计时器设置操控的有效声音范围可以相对宽于基于视频记录按钮按压操控的有效声音范围。也就是说,基于视频记录按钮按压操控的有效声音范围可以比基于计时器设置操控的有效声音范围窄。基于在按照这种方式设置的有效声音范围中检测到的碰撞声音确定合成目标图像选择范围。将参考图 7B 描述合成目标图像选择范围的确定方法。此外,在未在按照这种方式设置的有效声音范围中检测到碰撞声音的情况下,基于时间计数器“0”位置 381 确定合成目标图像选择范围。参考图 7C 描述合成目标图像选择范围的确定方法。

[0112] 图 7B 示意地说明在由有效声音范围设置部分 160 设置的有效声音范围 382 中检测到碰撞声音的情况下合成目标图像选择范围的确定方法。在图 7B 中,将特定声音检测部分 132 检测到碰撞声音的位置(时间轴上的位置)表示为碰撞声音检测位置 383。如图 7B 所示,在由有效声音范围设置部分 160 设置的有效声音范围 382 中检测到碰撞声音的情况下,合成目标图像选择范围确定部分 171 基于碰撞声音检测位置 383 确定合成目标图像选择范围。也就是说,合成目标图像选择范围确定部分 171 参考碰撞声音检测位置 383 将在时间轴上碰撞声音检测位置 383 之前和之后的预定范围确定为合成目标图像选择范围 354。具体地说,将时间轴上、包括领先碰撞声音检测位置 383 时间 L14 的范围和超过碰撞声音检测位置 383 时间 L15 的范围的范围 L16 确定为合成目标图像选择范围 384。

[0113] 这里,如上所述,在高尔夫球棒接触高尔夫球之后的高尔夫球挥杆期间的时间短于高尔夫球棒接触高尔夫球之前的时间。因此,确定合成目标图像选择范围 384,以便碰撞声音检测位置 383 之前的时间 L14 长于这以后的时间 L15。例如,时间 L14 可以是 1.5 秒,

而时间 L15 可以是 0.5 秒。按照这种方式,当基于在有效声音范围中检测到的碰撞声音确定合成目标图像选择范围时,可以确定合适的合成目标图像选择范围。

[0114] 假设在由有效声音范围设置部分 160 设置的有效声音范围 382 之外的位置(例如,由如图 7B 所示的箭头 388 和 389 所指示的位置)中检测碰撞声音。在按照这种方式在有效声音范围 382 之外的位置中检测到碰撞声音的情况下,以与如图 5B 所示的示例类似方式,不将碰撞声音用于合成目标图像选择范围的确定。如图 7C 所示,在不使用碰撞声音的情况下,可以基于时间计数器“0”位置确定合成目标图像选择范围。此外,在有效声音范围 382 之外的位置中检测到碰撞声音的情况下,碰撞声音不被用于合成目标图像选择范围的确定,但可以用于下一合成图像的生成处理。

[0115] 图 7C 示意地说明不在由有效声音范围设置部分 160 设置的有效声音范围 382 中检测到碰撞声音的情况下合成目标图像选择范围的确定方法。在图 7C 中,按照与图 7A 的示例类似的方式,将在其中由计时器设置部分 150 设置的计时器的计数值为“0”的位置表示为时间计数器“0”位置 381。如图 7C 所示,在未有效声音范围 382 中检测到碰撞声音的情况下,合成目标图像选择范围确定部分 171 基于时间计数器“0”位置 381 确定合成目标图像选择范围。也就是说,合成目标图像选择范围确定部分 171 参考时间计数器“0”位置 381,将在时间轴上在时间计数器“0”位置 381 之前和之后的预定范围确定为合成目标图像选择范围 385。具体地说,有效声音范围设置部分 160 将在时间轴上的一个范围 L19 设置为合成目标图像选择范围 385,该范围 L19 包括领先时间计数器“0”位置 381 时间 L17 的范围和超过时间计数器“0”位置 381 时间 L18 的范围。

[0116] 这里,人 300 在观察显示部分 191 上的显示屏幕的同时,进行高尔夫球挥杆,以便在计时器的计数值变成“0”的时间周围终止高尔夫球挥杆。由此,如图 5C 所示,与基于按压视频记录按钮的位置确定合成目标图像选择范围的情况相比,可能降低精确度。因此,在未有效声音范围 382 中检测到碰撞声音的情况下,将在时间轴上计时器的计数值为“0”的位置(时间计数器“0”位置 381)认为碰撞声音检测位置,以确定合成目标图像选择范围 385。

[0117] 此外,时间 L17 和 L18 对应于如图 7B 所示的时间 L14 和 L15。此外,例如,按照与图 7B 的示例类似的方式,时间 L17 可以是 1.5 秒,而时间 L18 可以是 0.5 秒。如上所述,在基于计时器的计数值为“0”的位置确定合成目标图像选择范围的情况下,基于由用户的感受引起的行为确定合成目标图像选择范围。由此,与如图 7B 所示的合成目标图像选择范围 384 相比,精确度可能降低。因此,在基于计时器的计数值为“0”的位置确定合成目标图像选择范围的情况下,例如,时间 L19 可以长于时间 L16,以确定合成目标图像选择范围。

[0118] 此外,相应时间 L11 至 L19 可以通过用户操控改变。

[0119] [合成图像的生成示例]

[0120] 接下来,将参考附图使用包括在所确定的合成目标图像选择范围中的每一帧,详细描述生成合成图像的示例。在本发明的第一实施例中,将作为示例,描述生成静止图像的合成图像的情况。

[0121] 图 8A、8B、8C 和 8D 是示意地说明根据本发明的第一实施例的、选择部分 170 的合成目标图像的选择处理和图层处理部分 180 的合成图像的生成处理的流程的示例的图。

[0122] 图 8A 示意地说明由图像捕获部分 111 生成的运动图像 400 以及运动图像 400 中

由合成目标图像选择范围确定部分 171 确定的合成目标图像选择范围 401。按照与如图 3B 所示的示例类似的方式等,运动图像 400 是通过由图像捕获设备 100 成像人的高尔夫练习挥杆的样子而获得的运动图像。此外,按照与如图 5B 和 5C 以及图 7B 和 7C 所示的示例类似的方式,假设合成目标图像选择范围 401 是由合成目标图像选择范围确定部分 171 确定的合成目标图像选择范围。

[0123] 图 8B 示意地说明包括在运动图像 400 中的合成目标图像选择范围 401 中的相应帧。在图 8B 中,包括在合成目标图像选择范围 401 中的相应帧以矩形形状示出。这里,将指示相应帧的 F1 至 F90 给定为矩形形状。

[0124] 合成目标图像选择部分 172 基于预定条件,从包括在合成目标图像选择范围 401 中的相应帧中选择合成目标图像。例如,合成目标图像选择部分 172 在帧 F1 至 F90 之中将预定间隔(例如,10-帧间隔)上的帧选择为合成目标图像。在这种情况下,例如,包括在检测到碰撞声音的位置周围的预定范围(例如,0.05 秒)中的帧的间隔可以比其他间隔更接近,以便被选择。因此,可以使用在要观察的动作过渡周围选择的图像生成合成图像。以这种方式选择的合成目标图像示出在图 8C 中。在这个示例中,为了描述的简单,示范了比较少量的帧。

[0125] 图 8C 示意地说明了从包括在合成目标图像选择范围 401 中的相应帧中选择为合成目标图像的多个合成目标图像(帧 F0、F10 等)。在图 8C 中,以矩形形状示出合成目标图像。这里,指示每一个合成目标图像的 F0、F10 或其他帧以矩形形状给出。在这个示例中,为了描述的简单,示范了比较少量的帧(例如,10 帧)。

[0126] 图 8D 示意地说明了由合成目标图像(多个帧 F0、F10 等)生成的合成图像 402。在图 8D 中,按照与图 8C 的示例类似的方式,以矩形形状示出合成目标图像。这里,指示每一个合成目标图像的 F1、F10 或其他帧以矩形形状给出。

[0127] 图层处理部分 180 通过修改处理来合成由合成目标图像选择部分 172 选择的合成目标图像(帧 F0、F10 等),从而生成合成图像。例如,图层处理部分 180 对关于由合成目标图像选择范围确定部分 171 选择的十个合成目标图像的合成,进行修改处理(例如,相对端的剪切(cutout)),从而以时间顺序方式在修改后合成合成目标图像。例如,图层处理部分 180 合成要被从左上端部分起按时间顺序方式安排的五个合成目标图像(F0、F10、F20、F30 和 F40),并且合成要被从左下端部分起按时间顺序方式安排的五个合成目标图像(F50、F60、F70、F80 和 F90)。基于坐标计算部分 173 的计算结果确定合成目标图像的合成位置。因此,如图 4B 所示,生成指示练习高尔夫球挥杆的人的动作过渡的合成图像。

[0128] 在图 8A、8B、8C 和 8D 中,已经描述了在其中容易地进行图像合成的示例(在其中在合成目标图像的相对端的剪切之后进行图像合成的示例)。在下文中,将描述在其中将合成目标图像分离成目标对象图层和背景图层以生成合成图像的示例。

[0129] [合成目标图像的图层分离示例]

[0130] 图 9A 和 9B 是示意地说明根据本发明的第一实施例的、图层分离部分 181 的合成目标图像的分离方法的图。在这个示例中,基于对象信息生成部分 120 所生成的对象信息,将由合成目标图像选择部分 172 选择的合成目标图像被分离为两个图层。

[0131] 图 9A 说明了合成目标图像 410 和为合成目标图像 410 生成的对象信息 412。合成目标图像 410 是包括人 411 的图像。此外,作为对象信息生成部分 120 所生成的对象信息,

例如,生成指示将“1”赋予与目标对象区域对应的像素而将“0”赋予与目标对象之外的区域(背景区域)对应的像素的信息。在图9A中,与人411对应的区域(目标对象区域)是白区域413,而人411之外的区域(背景区域)是黑区域,从而示意地说明了对象信息412。

[0132] 如上所述,在捕获图像保持部分115中,顺序保持图像捕获部分111所生成的捕获图像,并且与捕获图像相关联地保持对象信息生成部分120所生成的对象信息。此外,合成目标图像选择部分172将所选择的合成目标图像和为合成目标图像生成的对象信息输出到图层分离部分181。然后,图层分离部分181使用对应的对象信息,将从合成目标图像选择部分172输出的合成目标图像分离成两个图层。

[0133] 图9B说明了由图层分离部分181分离出的目标对象图层420和背景图层422。目标对象图层420是通过提取如图9A所示的合成目标图像410中的目标对象区域(也就是说,被赋予“1”作为对象信息412的像素区域)而获得的图像。也就是说,提取了与包括在合成目标图像410中的人411对应的区域421。此外,背景图层422是通过提取如图9A所示的合成目标图像410中的背景区域(也就是说,被赋予“0”作为对象信息412的像素区域)而获得的图像。也就是说,提取了与包括在合成目标图像410中人411之外的区域(背景区域)对应的区域(区域423之外的区域)。

[0134] 对按照这种方式分离出来的图层图像,用图层修改部分182进行各种修改处理。将参考图10A至10C详细描述这种修改示例。

[0135] [合成目标图像的修改示例]

[0136] 图10A、10B和10C是示意地说明根据本发明的第一实施例的、作为图层修改部分182的修改处理的目标的图像和由图层修改设备182修改了的图像的图。这里,作为一个示例,描述了为诸如高尔夫球这样的、以站立姿势进行的体育运动生成合成图像的情况。此外,在图10A、10B和10C中,为了描述的简单,作为一个示例,示出了使用图层分离之前的状态的合成目标图像进行修改处理的情况。

[0137] 图10A说明了借助于图层修改部分182的修改之前的合成目标图像430。此外,图10B和10C说明了借助于图层修改部分182的修改之后的合成目标图像432和433。进行高尔夫球挥杆的人431被包括在合成目标图像430、432和433中。

[0138] 例如,如图10A所示的合成目标图像430具有宽度W1和高度H1的图像尺寸。例如,合成目标图像430的图像尺寸可以是VGA尺寸(640像素×480像素)。也就是说,宽度W1可以是640,而高度H1可以是480。

[0139] 如图10B所示的合成目标图像432是与如图10A所示的合成目标图像430中的过渡动作有效区域对应的图像,并且例如具有宽度W2和高度H2的图像尺寸。这里,过渡动作有效区域是用于从包括在修改之前的合成目标图像中的对象中剪切目标对象的区域。例如,合成目标图像432的图像尺寸可以是320像素×480像素。也就是说,宽度W2可以是320,而高度H2可以是480。

[0140] 如图10C所示的合成目标图像433是与如图10A所示的合成目标图像430中的剩余背景区域对应的图像,并且例如宽度W3和高度H3的图像尺寸。剩余背景区域是在合成处理中用作背景图像的区域。例如,合成目标图像433的图像尺寸可以是160像素×480像素。也就是说,宽度W3可以是160,而高度H3可以是480。

[0141] 假设过渡动作有效区域与剩余背景区域之间的中心位置与合成目标图像的中心

位置一致。此外,根据目标对象的过渡方向,过渡动作有效区域的尺寸($W2 \times H2$)与剩余背景区域尺寸($W3 \times H3$)可以在其尺寸上相同(至少宽度和高度之一)。例如,在图像被表示为动作过渡图像时,目标对象的过渡方向是在其上图像以时间方式领先的方向。

[0142] 例如,如图 4B 所示,在为诸如高尔夫球这样的、以站立姿势进行的体育运动的动作过渡生成合成图像的情况下,很可能每一个目标对象图像在横向方向上过渡。由此,如图 10B 和 10C 所示,可能最好过渡动作有效区域的高度 $H2$ 和剩余背景区域的高度 $H3$ 与合成目标图像的高度 $H1$ 相同。另一方面,在为水平方向状态下进行的体育运动的动作过渡生成合成图像的情况下,每一个目标对象图像都可能在纵向方向上过渡。在这种情况下,可能最好过渡动作有效区域的宽度 $W2$ 和剩余背景区域的宽度 $W3$ 与合成目标图像的宽度 $W1$ 相同。例如,在柔道的寝技等变成目标对象的情况下,很有可能在纵向方向上进行动作过渡。

[0143] 图 10A、10B 和 10C 示出了一个示例,在其中最好根据对象的尺寸或动作进行过渡动作有效区域和剩余背景区域的宽度和高度的适当设置,以便将图像剪切为用于合成处理的图像。

[0144] [图像合成示例]

[0145] 图 11A 和 11B 是示意地说明根据本发明的第一实施例的、图层合成部分 183 的合成方法的图。这里,该实施例提供一个示例,在其中对以时间顺序方式连续的两个合成目标图像进行合成处理。

[0146] 图 11A 示意地说明由图层分离部分 181 分离并由图层修改部分 182 修改的合成目标图像(以时间顺序方式连续的两个合成目标图像。如上所述,合成目标图像被图层分离部分 181 分离成目标对象图层和背景图层。此外,图层修改部分 182 对图层分离部分 181 分离出来的目标对象图层和背景图层进行修改处理。按照这种方式,当对以时间顺序方式连续的两个合成目标图像进行分离和修改时,如图 11A 所示,生成四个图层(第一优先级图像 441 至第四优先级图像 444)。

[0147] 假设第一优先级图像 441 是合成目标图像中在时间轴上来得较晚的目标对象图层,而第三优先级图像 443 是合成目标图像的背景图层。此外,假设第二优先级图像 442 是合成目标图像中在时间轴上来得较早的目标对象图层,而第四优先级图像 444 是合成目标图像的背景图层。而且,在合成这些图像的情况下,进行合成以便盖写具有较高优先级的图像。

[0148] 图 11B 说明了基于相应图层图像的优先级生成的合成图像 450。也就是说,图层合成部分 183 基于优先级合成四个图层(第一优先级图像 441 至第四优先级图像 444),从而生成合成图像 450。合成图像 450 是通过合成包括在第一优先级图像 441 中的人像 445 和包括在第二优先级图像 442 中的人像 446 以被盖写在由第三优先级图像 443 和第四优先级图像 444 合成的背景区域上而获得的图像。在这种情况下,合成第三优先级图像 443 以盖写在第四优先级图像 444 上。此外,合成包括在第一优先级图像中的人像 445 以被盖写在包括在第二优先级图像 442 中的人像 446 上。

[0149] 如图 12A 和 12B 所示,在顺序合成三个或多个合成目标图像来生成合成图像的情况下,以时间顺序方式顺序合成合成目标图像,从而生成合成图像。此外,将在以时间顺序方式连续的两个合成目标图像之中在时间轴上来得较早的合成目标图像保持在合成图像保持部分 185 中,到来得较晚的合成目标图像的合成处理时间为止。

[0150] 图 12A 和 12B 是示意地说明根据本发明的第一实施例的、图层合成部分 183 的合成目标图像的合成方法的图。这里,示出了到生成如图 12B 所示的合成图像 460 为止的合成目标图像的合成示例。

[0151] 图 12A 说明了在到生成如图 12B 所示的合成图像 460 为止的某个时间点处的合成状态。例如,假设如图 12A 所示的状态是对合成目标图像 462(两个图层图像)进行合成处理的时间点处的状态。也就是说,图 12A 示出了与合成目标图像 462 对应的过渡动作有效区域 E10,以及在其中对刚来的合成目标图像 463 进行合成处理的合成图像 461。在这个示例中,将刚变成合成目标的相应图层图像保持在合成图像保持部分 185 中。

[0152] 这里,在合成目标图像 462 的区域 E11 中,在邻近合成目标图像 463 和合成目标图像 462 之间不存在重叠区域。由此,在将合成目标图像 462 与合成图像 461 合成的情况下,关于合成目标图像 462 中的区域 E11,合成合成目标图像 462 的过渡动作有效区域的像素以盖写在合成图像 461 上,作为现在示出那样的新合成图像。

[0153] 然而,合成目标图像 462 中的区域 E12 与邻近合成目标图像 463 的一部分重叠。因此,如图 11A 和 11B 所示,根据优先级在合成目标图像 462 的区域 E12 中合成相应图层图像。

[0154] 也就是说,图层合成部分 183 使用作为当前合成目标的合成目标图像 462 和保持在合成图像保持部分 185 中的邻近合成目标图像 463 的相应图层图像,根据针对区域 E12 的优先级合成相应图层图像。通过这样的合成,从合成目标图像 462 和 463 生成区域 E11+E12 的合成图像。此外,图层合成部分 183 将要盖写的、以这种方式生成的区域 E11+E12 的合成图像与保持在合成图像保持部分 185 中的合成图像(与邻近合成目标图像 463 合成的合成图像)合成。也就是说,图层合成部分 183 进行粘贴在与保持在合成图像保持部分 185 中的合成图像对应的区域 E1 中由这样的合成所生成的区域 E11+E12 的合成图像的处理。

[0155] 此外,变成当前合成目标和由当前合成处理所生成的合成图像的合成目标图像 462(两个图层图像)被保持在合成图像保持部分 185 中,并且用于下一合成处理。按照这种方式生成的合成图像被示出在图 12B 中。

[0156] 按照这种方式,当根据优先级合成相应图层图像时,关于在合成处理中在连续图像之间重叠的区域(例如,区域 E12),可以生成合成图像,而不会损坏一部分目标对象。因此,在使用多个合成目标图像生成合成图像的情况下,也可以生成能够适当地表达目标对象的动作的动作过渡图像。

[0157] 按照这种方式,由图层合成部分 183 生成的合成图像 460 被存储在合成图像存储部分 200 中。此外,例如,根据用户操控在显示部分 191 上显示合成图像 460。此外,例如,只要图层合成部分 183 生成合成图像,它可以被自动显示在显示部分 191 上。图 13 说明了以这种方式生成的合成图像的示例。

[0158] [合成图像示例]

[0159] 图 13 是说明根据本发明的第一实施例的、由图层处理部分 180 生成的合成图像的示例的图。图 13 所示的合成图像 470 指示练习高尔夫球挥杆的人的动作过渡。在合成图像 470 中,箭头 471 所指示的区域是包括与生成碰撞声音的位置对应的图像的区域。

[0160] [图像捕获设备的操作示例]

[0161] 图 14 是说明根据本发明的第一实施例的、图像捕获设备 100 的合成图像生成处理的处理过程的示例的流程图。

[0162] 首先,确定是否进行合成图像记录方式的设置操控(步骤 S901)。例如,确定是否进行合成图像记录方式的设置按钮或计时器设置按钮的按压操控。在不进行合成图像记录方式的设置操控的情况下(步骤 S901),继续监视到进行合成图像记录方式的设置操控为止。在进行合成图像记录方式的设置操控的情况下(步骤 S901),图像捕获部分 111 进行捕获图像的生成处理(步骤 S902),并将所生成的捕获图像保持在捕获图像保持部分 115 中(步骤 S903)。也就是说,进行缓存处理。

[0163] 随后,确定是否进行视频记录按钮的按压操控(步骤 S904)。在进行视频记录按钮的按压操控的情况下,有效声音范围设置部分 160 基于按压位置设置有效声音范围(步骤 S905)。例如,如图 5A 所示,基于视频记录按钮按压位置 351 设置有效声音范围 352。另一方面,在不进行视频记录按钮的按压操控的情况下(步骤 S904),确定由计时器设置部分 150 设置的计时器的计数值是否为“0”(步骤 S906)。在计时器的计数值为“0”的情况下(步骤 S906),有效声音范围设置部分 160 基于计时器的计数值为“0”的位置设置有效声音范围(步骤 S907)。例如,如图 7A 所示,基于时间计数器“0”位置 381 设置有效声音范围 382。此外,在计时器的计数值不为“0”,或者不由计时器设置部分 150 进行计时器设置的情况下(步骤 S906),过程返回步骤 S902。步骤 S904 至 S907 表示公开在权利要求书中的有效声音范围设置过程的示例。

[0164] 接下来,进行缓存终止处理(步骤 S908),并确定是否在所设置的有效声音范围中检测到碰撞声音(步骤 S909)。在所设置的有效声音范围中检测到碰撞声音的情况下(步骤 S909),合成目标图像选择范围确定部分 171 基于检测到碰撞声音的位置确定合成目标图像选择范围(步骤 S910)。例如,如图 5C 所示,基于碰撞声音检测位置 353 确定合成目标图像选择范围 355。此外,例如,如图 7C 所示,基于碰撞声音检测位置 383 确定合成目标图像选择范围 385。步骤 S909 和 S910 表示公开在权利要求书中的选择过程的示例。

[0165] 此外,在所设置的有效声音范围中未检测到碰撞声音的情况下(步骤 S909),合成目标图像选择范围确定部分 171 基于设置有效声音范围时作为参考的位置确定合成目标图像选择范围(步骤 S911)。也就是说,在基于视频记录按钮的按压位置设置有效声音范围的情况下,基于按压位置确定合成目标图像选择范围。例如,如图 5C 所示,基于视频记录按钮按压位置 351 确定合成目标图像选择范围 355。另一方面,在基于计时器的计数值为“0”的位置设置有效声音范围的情况下,基于该位置确定合成目标图像选择范围。例如,如图 7C 所示,基于时间计数器“0”位置 381 确定合成目标图像选择范围 385。

[0166] 随后,合成目标图像选择部分 172 从包括在所确定的合成目标图像选择范围中的相应图像选择合成目标图像(步骤 S912)。然后,使用所选择的合成目标图像进行合成图像的生成处理(步骤 S913)。

[0167] 接下来,确定是否进行合成图像记录方式的释放操控(步骤 S194)。在不进行合成图像记录方式的释放操控的情况下,过程返回步骤 S902,随后进行涉及下一合成图像的图像合成处理。另一方面,在进行合成图像记录方式的释放操控的情况下(步骤 S914),终止合成图像生成处理的操作。

[0168] 按照这种方式,在本发明的第一实施例中,将其中变成检测目标的碰撞声音可以

存在的时间跨度设置为有效声音范围,因此,可以减少除了变成检测目标的碰撞声音之外的声音的错误检测。因此,生成使用诸如高尔夫球挥杆之类的、引起碰撞声音的极快速动作作为目标的合成图像的情况下,可以高精度地检测到碰撞声音的生成位置。此外,可以生成强调在碰撞声音的精确生成位置周围的合成图像,并且生成具有高可视性(visibility)的合成图像。在这种情况下,由于碰撞声音的生成位置的检测不需要用户的手工工作,所以可以减少用户的麻烦。

[0169] 此外,如上所述,可以确定是否根据是否在有效声音范围中检测到碰撞声音来生成合成图像。这种示例示出在图 15 中。

[0170] 图 15 是说明根据本发明的第一实施例的、图像捕获设备 100 的合成图像生成处理的处理过程的示例的流程图。这个示例是图 14 中所示的合成图像生成处理的处理过程的修改示例,在其中省略了步骤 S911。因此,将相同的参考数字赋予与图 14 相同的处理步骤,并省略其描述。

[0171] 确定是否在步骤 S905 或 S907 所设置的有效声音范围中检测到碰撞声音(步骤 S909)。在所设置的有效声音范围中未检测到碰撞声音的情况下(步骤 S909),过程转到步骤 S914。也就是说,在所设置的有效声音范围中未检测到碰撞声音的情况下,使用下一视频记录按钮的按压位置或者计时器的计数值为“0”的位置,进行合成图像的生成处理,而不生成合成图像。也就是说,在有效声音范围中检测到碰撞声音的情况下,选择部分 170 进行控制,以便在图层处理部分 180 中生成使用特定声音的合成图像,并进行控制以便在未有效声音范围中检测到碰撞声音的情况下,不在图层处理部分 180 生成合成图像。

[0172] 按照这种方式,不能在有效声音范围中检测到碰撞声音的情况下,通过返回到下一可拍摄状态而不生成合成图像,可以减少用户拍摄机会的损失。此外,由于可以使用精确的碰撞声音生成合成图像,因此可以生成合适的合成图像。

[0173] 当按照这种方式生成合成图像时,即使不习惯图像捕获设备的操作的初学者也可以容易地生成合适的合成图像。

[0174] 在图像捕获部分 111 所生成的捕获图像被记录为图像内容(例如,运动图像文件或连续静止图像文件)的情况下,可以将碰撞声音检测位置记录为帧的元数据。因此,例如,在显示图像内容列表的情况下,可以将与碰撞声音检测位置对应的帧用作代表性缩略图。

[0175] 2. 第二实施例

[0176] 在本发明的第一实施例中,已经通过示例的方式描述了图像捕获设备与目标对象(练习高尔夫球挥杆的人)之间的距离比较短的情况。然而,图像捕获设备可以被定位于与目标对象的距离比较长的位置上并且可以借助于图像捕获设备的变焦功能拍摄目标对象。在图像捕获设备与目标对象之间的距离比较长的这种情况下,到碰撞声音到达图像捕获设备为止的时间根据该距离的长度变长。在这种情况下,在生成碰撞声音的时间与碰撞声音到达图像捕获设备的时间之间发生延迟。因此,本发明的第二实施例提供一个示例,在其中在图像捕获设备与目标对象之间的距离比较长的情况下,根据该距离的长度,校正时间轴上的碰撞声音的检测位置。

[0177] 图 16A 和 16B 是说明根据本发明的第二实施例的、图像捕获设备 500 与目标对象之间的距离和碰撞声音的生成位置及其达到位置之间的关系的图。图 16A 示意地说明了图

像捕获设备 500 与变成图像捕获设备 500 的图像捕获目标的人 540 之间的位置关系。这里,图像捕获设备 500 与人 540 之间的距离被称为距离 d 。除了距离 d 不同外,图像捕获设备 500 与人 540 之间的关系与图 3A 中所示的示例近似相同,因此省略其描述。

[0178] 图 16B 示意地说明了形成由图 17 中的图像捕获部分 111 所生成的运动图像的相应图像与人 540 的动作过渡之间的关系。图像组 550 表示在图 16A 中所示的状态下人 540 的动作过渡,在其中以时间轴安排形成图像组 550 的相应图像。图像组 560 表示在图 16A 中所示的该状态下形成由图像捕获部分 111 所生成的运动图像的相应图像,在其中以时间轴安排形成图像组 560 的相应图像。这里,将通过人 540 的高尔夫球挥杆生成碰撞声音的位置(时间轴上的位置)表示为碰撞声音生成位置 555。

[0179] 在这个方面,声音在空气中的速度为大约 340 米/秒。因此,在图像捕获设备 500 与人 540 之间的距离 d 比较长的情况下。到碰撞声音到达图像捕获设备 500 为止的时间变长,因此在生成碰撞声音的时间与碰撞声音被输入到图像捕获设备 500 的时间之间发生延迟。

[0180] 例如,假设形成图 16B 中所示的图像组 550 的图像之中的图像 551 的位置是碰撞声音生成位置 555。在这种情况下,例如,如箭头 552 所指,在距离 d 比较短(例如,几米)的情况下,与生成碰撞声音的位置对应的图像 551 和在碰撞声音被输入到图像捕获设备 500 时所生成的图像 561 相互近似相同。另一方面,如箭头 553 所指,在距离 d 比较长(例如,340 米或更长)的情况下,与生成碰撞声音的位置对应的图像 551 和在碰撞声音被输入到图像捕获设备 500 时所生成的图像 562 相互不同,从而引起延迟。在这个方面,即使在距离 d 比较长的情况下,也很可能以与距离 d 比较短的情况下近似相同的观察角度来观察在显示在显示部分 191 上的显示屏幕上的图像。由此,可能难以通过用户的手工操控来校正碰撞声音的延迟。因此,在下文中,示出校正根据图像捕获设备与目标对象之间的距离生成的碰撞声音的延迟的示例。

[0181] [图像捕获设备的配置示例]

[0182] 图 17 是说明根据本发明的第二实施例的、图像捕获设备 500 的功能性配置的示例的方块图。图像捕获设备 500 不同于图 1 中所示的图像捕获设备 100 在于:添加了对象距离计算部分 510,并且代替选择部分 170 安装了选择部分 520。除了这些组件之外的配置几乎与图像捕获设备 100 中的相同。因此,将相同的参考数字赋予共用组件,并且将省略其一部分的描述。

[0183] 对象距离计算部分 510 计算与包括在图像捕获部分 111 所生成的捕获图像中的对象的距离(对象距离),并将所计算出的对象距离输出到选择部分 520。例如,对象距离计算部分 510 检测包括在图像捕获部分 111 所生成的捕获图像中的人的面孔,并使用所检测到的面孔和捕获图像生成时的成像信息(例如,每个镜头的位置和聚焦位置)计算目标对象的对象距离。例如,可以将通过记录面孔的亮度分布信息的模板和内容图像的匹配的面孔检测方法(例如,参看日本未审专利申请公开第 2004-133637 号)用作包括在捕获图像中的面孔的检测方法。此外,可以使用基于包括在捕获图像中的肤色部分或人类面孔的特征量的面孔检测方法。通过这样的面孔检测方法,可以计算人的面孔在捕获图像中的位置和尺寸。

[0184] 选择部分 520 基于从对象距离计算部分 510 输出的对象距离估计碰撞声音的延迟

时间,并且在考虑到延迟时间的情况下选择合成目标图像。具体地说,选择部分 520 的合成目标图像选择范围确定部分 171(图 2 中所示)基于从对象距离计算部分 510 输出的对象距离,校正特定声音检测部分 132 所检测到的特定声音的检测位置(时间轴上的位置)。此外,合成目标图像选择范围确定部分 171 基于校正之后的位置(校正位置)和由有效声音范围设置部分 160 设置的有效声音范围,确定合成目标图像选择范围。合成目标图像选择范围确定部分 171 是公开在权利要求书中的延迟时间计算部分的示例。

[0185] [对象距离的计算示例]

[0186] 接下来,将参考附图详细描述用于计算图像捕获设备 500 与目标对象之间的距离的距离计算方法。这里,将描述用于使用平常人类面孔的尺寸(参考值)估计图像捕获设备 500 与面孔之间的距离的距离计算方法。

[0187] 图 18A 和 18B 是示意地说明根据本发明的第二实施例的、作为图像捕获部分 111 的图像捕获目标的图像捕获范围以及显示在显示部分 191 上的捕获图像的图。图 18A 是示意地说明在成像以高尔夫球场为背景的人 571 的情况下的图像捕获范围 570 的图,而图 18B 是说明与图 18A 所示的图像捕获范围 570 对应的捕获图像 575 的、显示部分 191 的显示示例的图。

[0188] 例如,如图 18A 所示,在人 571 以高尔夫球场为背景被成像为对象的情况下,根据图像捕获设备 500 的位置和变焦镜头的位置或诸如此类的信息来确定成像人 571 的范围(图像捕获范围 570)。此外,如图 18B 所示,将从包括在图像捕获范围 570 中的对象入射的光由图像捕获部分 111 转换成捕获图像,并且在显示部分 191 上显示捕获图像(所谓直接图像)。

[0189] 这里,在图 18A 中所示的图像捕获范围 570 中,将图像捕获设备 500 在左右方向上的宽度表示为图像捕获范围宽度 W_A ,而将包括在图像捕获范围 570 中的人 571 的面孔 572 在左右方向上的宽度表示为面孔宽度 W_{ref} 。面孔宽度 W_{ref} 对应于平常人类面孔的尺寸,不是人 571 的实际面孔宽度。此外,将图 18B 中所示的捕获图像 575 的左右方向上的宽度表示为图像宽度 W_w ,而将包括在捕获范围 575 中的人 576 的面孔 577 在左右方向上的宽度表示为面孔图像宽度 W_f 。在这种情况下,图像捕获范围宽度 W_A 与面孔宽度 W_{ref} 的比率通常和图像宽度 W_w 与面孔图像宽度 W_f 的比率相同。

[0190] 图 19 是示意地说明根据本发明的第二实施例的、作为图像捕获部分 111 的图像捕获目标的图像捕获范围与显示在显示部分 191 上的捕获图像之间的位置关系的顶视图。图 19 所示的示例中,示意地示出了图 18A 中所示的人 571 与提供在图像捕获设备 500 中的图像捕获元件 112 和显示部分 191 之间的位置关系。如图 19 所示,将从图像捕获设备 500 到面孔 572 的距离表示为对象距离 D_f ;将图像捕获设备 500 中的焦距表示为焦距 f ;并且将图像捕获元件 112 的宽度表示为图像捕获元件宽度 W_i 。图像捕获范围宽度 W_A 和面孔宽度 W_{ref} 与图 18A 中所示的那些相同。

[0191] 如图 19 所示,如果从包括在图像捕获范围 570 中的对象入射的光进入图像捕获元件 112,那么生成与包括在图像捕获范围 570 中的对象对应的捕获图像,然后在显示部分 191 上显示所生成的捕获图像 575。此外,在从人 571 的面孔 572 入射的光进入图像捕获元件 112 的情况下图像捕获元件 112 的面孔的宽度是该图像捕获元件上的面孔宽度 W_b 。

[0192] 这里,从具有平行线的比例关系建立以下两个等式。

[0193] $Df/W_{ref} = f/W_b$... 等式 1

[0194] $W_f/W_w = W_b/W_i$... 等式 2

[0195] 这里,将等式 1 改变成 $Df = f \times W_{ref}/W_b$,而将等式 2 改变成 $W_b = W_f \times W_i/W_w$ 。然后,将通过改变等式 2 所获得的 $W_b = W_f \times W_i/W_w$ 代入等式 1,从而计算出以下等式 3。基于镜头的基本物理规律计算这个等式 3。

[0196] $Df = W_{ref} \times (f/W_i) \times (W_w/W_f)$... 等式 3

[0197] 这里, W_i (图像捕获元件宽度)和 W_w (图像宽度)是常数,而平常人类面孔的尺寸被用作 W_{ref} 。在这种情况下,当检测到 W_f (面孔图像宽度)时,可以使用等式 3 计算 Df (对该面孔的估计距离)。

[0198] 例如,如图 18A 所示,在一个人的面孔 572 被包括在图像捕获范围 570 中的情况下,通过与对象距离计算部分 510 从与图像捕获范围 570 对应的捕获图像 575 中检测面孔 577。按照这种方式,在检测到面孔的情况下,对象距离计算部分 510 基于等式 3,使用所检测到的面孔的宽度(面孔图像宽度)计算对象距离 Df 。

[0199] 本发明的第二实施例提供这样一个示例,在其中使用一个参考值作为平常人类面孔尺寸来计算对象距离。这里,根据人的属性(例如,年龄和性别)变成图像捕获目标的人具有不同的面孔尺寸。例如,在比较孩子面孔与成人面孔的情况下,面孔尺寸可能相互不同。此外,在比较女性面孔与男性面孔的情况下,面孔尺寸可能相互不同。因此,根据人的属性的多个参考值可以被保持在对象距离计算部分 510 中。然后,对象距离计算部分 510 可以检测关于从捕获图像中检测到的面孔的属性,并且对象距离计算部分 510 可以使用根据该属性的参考值计算距离 d 。为了检测该属性,例如,可以使用这样一种检测方法,在其中使用面孔图像中的两个点之间的亮度的差值,由弱学习器(weak learner)检测相应属性(例如,参看日本未审专利申请公开第 2009-118009 号)。

[0200] [延迟时间的校正示例]

[0201] 图 20A 和 20B 是示意地说明根据本发明的第二实施例的、有效声音范围设置部分 160 的有效声音范围的设置方法和合成目标图像选择范围确定部分 171 的合成目标图像选择范围的确定方法的图。由于这个示例是图 5A 和 5B 的修改示例,因此将省略共用组件的描述。

[0202] 图 20A 示意地以矩形形状说明了图像捕获部分 111 所生成的运动图像 580。此外,假设按压视频记录按钮的位置是图 20A 中所示的运动图像 580 在时间轴上的视频记录按钮的按压位置(视频记录按钮按压位置 581)。在这种情况下,有效声音范围设置部分 160 基于视频记录按钮按压位置 581 设置有效声音范围 582。具体地说,有效声音范围设置部分 160 设置这样一个范围作为有效声音范围 582,该范围在时间轴上领先视频记录按钮按压位置 581 时间 L_{21} 。这里,例如,按照与图 5A 所示的时间 L_1 类似方式,时间 L_{21} 可以是 1.0 秒。此外,时间 L_{21} 可以长于时间 L_1 ,并且可以根据对象距离计算部分 510 所计算出的对象距离的尺寸来改变。

[0203] 图 20B 示意地说明了在由有效声音范围设置部分 160 设置的有效声音范围 582 中检测到碰撞声音的情况下合成目标图像选择范围的确定方法。在图 20B 中,将由特定声音检测部分 132 检测到碰撞声音的位置(时间轴上的位置)表示为碰撞声音检测位置 583。

[0204] 在本发明的第二实施例中,在是否在有效声音范围 582 中检测到碰撞声音的确定

之前,合成目标图像选择范围确定部分 171 基于对象距离计算部分 510 所估计的对象距离 $d(m)$ 来估计声音的延迟时间。具体地说,合成目标图像选择范围确定部分 171 基于对象距离计算部分 510 所估计的对象距离 $d(m)$ 来估计声音的延迟时间。如上所述,由于声音的速度是 340(米/秒),因此可以使用以下等式 4 计算到碰撞声音到达图像捕获设备 500 为止的时间 x (秒)。

[0205] $x = d/340 \quad \dots \quad$ 等式 4

[0206] 合成目标图像选择范围确定部分 171 在时间轴上将碰撞声音检测位置 583 移动使用等式 4 计算出的时间 x 。将移动之后的位置表示为校正位置 584。

[0207] 随后,合成目标图像选择范围确定部分 171 确定该校正位置 584 是否被包括在由有效声音范围设置部分 160 设置的有效声音范围 582 中。如图 20B 所示,在校正位置 584 被包括在由有效声音范围设置部分 160 设置的有效声音范围 582 中的情况下,合成目标图像选择范围确定部分 171 基于该校正位置 584 确定合成目标图像选择范围。也就是说,合成目标图像选择范围确定部分 171 参考校正位置 584,将时间轴上在校正位置 584 之前和之后的预定范围确定为合成目标图像选择范围 585。具体地说,合成目标图像选择范围确定部分 171 确定时间轴上的一个范围 L24 作为合成目标图像选择范围 585,该范围 L24 包括领先校正位置 584 时间 L22 的范围和超过校正位置 584 时间 L23 的范围。

[0208] 这里,例如,时间 L22 至 L24 可以与图 5B 中所示的时间 L2 至 L4 相同。此外,时间 L24 可以长于时间 L4,并且可以根据对象距离计算部分 510 计算出的对象距离的尺寸来改变。

[0209] 此外,例如,在图像捕获设备 500 与人 540 之间的距离比较远的情况下,可以通过遥控器操控来进行计时器设置。因此,在基于计时器设置部分 150 所设置的计时器的计数值来设置有效声音范围并且使用该有效声音范围来确定合成目标图像选择范围的情况下,类似地,可以估计声音的延迟时间,并且可以基于延迟时间进行校正。

[0210] 按照这种方式,即使在图像捕获设备 500 与人 540 之间的距离比较远的情况下,碰撞声音的检测位置也被校正来基于校正之后的位置确定合成目标图像选择范围。因此,可以确定合适的合成目标图像选择范围。

[0211] 这里,校正位置可能不存在于由有效声音范围设置部分 160 设置的有效声音范围 582 中。在这种情况下,按照与图 5C 中所示的示例类似的方式,不使用碰撞声音。此外,在不使用碰撞声音的情况下,如图 5C 所示,可以基于视频记录按钮的按压位置确定合成目标图像选择范围。

[0212] 在图 18A、18B 和图 19 中,示出了使用人的面孔的尺寸计算对象距离的示例,但如图 21 所示,对象距离可以用除面孔之外的部分的尺寸来计算。

[0213] [对象距离的计算示例]

[0214] 图 21 是说明根据本发明的第二实施例的、显示部分 191 中的显示示例的图。在图 21 中所示的显示屏幕 590 中,将用于在适当位置安排进行高尔夫球挥杆的人 591 的操控支持图像 592 和 593 显示成重叠在捕获图像上。这个示例是图 18A 和 18B 以及图 19 中所示的对象距离的计算方法的修改示例,并且是基于在显示屏幕 590 上对象占据的区域的尺寸(垂直方向上的长度)计算对象距离的示例。

[0215] 操控支持图像 592 和 593 表示进行高尔夫球挥杆的人 591 被安排在显示屏幕 590

中的推荐区域。操控支持图像 592 在显示屏幕 590 用黑线指示,并且表示进行高尔夫球挥杆的人 591 在左右方向上的推荐区域。此外,操控支持图像 593 在显示屏幕 590 用点状黑框指示,并且表示进行高尔夫球挥杆的人 591 在上下方向上的推荐区域。这里,用操控支持图像 593 指示的上下方向上的推荐区域可以是显示屏幕 590 在垂直方向上的长度的预定比率(例如,70%)的值 H11。

[0216] 例如,在使用图像捕获设备 500 通过设置合成图像记录方式进行拍摄的情况下,摄影师确认包括在显示屏幕 590 中的操控支持图像 592 和 593,并进行调节以便将目标对象的人包括在操控支持图像 592 和 593 中。此外,在进行变焦操控等的情况下,类似地,摄影师进行调节以便将目标对象的人包括在操控支持图像 592 和 593 中。

[0217] 按照这种方式,通过应用于像以上那样基于涉及面孔和包括在捕获图像中的面孔的尺寸的参考值计算对象距离的对象距离计算方法,可以计算对象距离。也就是说,由于包括在捕获图像中的人的尺寸可以在图 21 中所示的示例中的预定值 H11,因此可以基于涉及打高尔夫球的人的平常高度和预定值 H11 的参考值(例如,170cm)来计算对象距离。

[0218] 在图 18A、18B、图 19 和图 21 中所示的示例中,使用了人的每个部分的尺寸来计算对象距离,但可以使用其他对象距离计算方法。例如,可以生成涉及捕获图像的景深图并使用该景深图计算对象距离。此外,可以将距离测量传感器安装在图像捕获设备 500 中,并且可以使用距离测量传感器所测量的对象距离。而且,可以使用聚焦位置信息计算对象距离。

[0219] 此外,可以根据拍摄情形选择使用聚焦位置信息的对象距离计算方法和使用变焦位置信息的对象距离计算方法中的任何一种。例如,在焦点形成在宽边缘(wide edge)的变焦位置中的情况下,很可能聚焦位置信息具有比变焦位置信息更高的精度。由此,在焦点形成在宽边缘的变焦位置中的情况下,可以选择使用聚焦位置信息的对象距离计算方法。此外,可以用特定方法估计对象距离,并且可以确定是使用估计结果还是参考估计结果用其他方法进行重新估计。

[0220] [图像捕获设备的操作示例]

[0221] 图 22 是说明根据本发明的第二实施例的、图像捕获设备 500 的合成图像生成处理的处理过程的示例的流程图。这个示例是图 14 中所示的合成图像生成处理的处理过程的修改示例,在其中添加了步骤 S921 并进行步骤 S922 和 S923 的处理过程代替步骤 S910。因此,将相同的参考数字赋予与图 14 相同的处理过程,并省略其描述。

[0222] 在进行了缓存终止处理之后(步骤 S908),对象距离计算部分 510 计算对象距离(步骤 S921)。

[0223] 此外,在所设置的有效声音范围中检测到碰撞声音的情况下(步骤 S909),合成目标图像选择范围确定部分 171 基于所计算出的对象距离校正碰撞声音检测位置(步骤 S922)。具体地说,合成目标图像选择范围确定部分 171 基于对象距离计算延迟时间,并且把从计数器部分 174 供应的计数值减去所计算出的延迟时间来计算校正值。然后,合成目标图像选择范围确定部分 171 基于该校正值指定在碰撞声音检测位置的校正之后的位置。例如,如图 20B 所示,将碰撞声音检测位置 583 移动校正值 x 来计算校正位置 584。

[0224] 随后,合成目标图像选择范围确定部分 171 基于该校正后的位置(校正位置)确定合成目标图像选择范围(步骤 S923)。例如,如图 20B 所示,基于校正位置 584 确定合成目标图像选择范围 585。

[0225] 按照这种方式,根据本发明的第二实施例,可以高精度地检测碰撞声音的生成位置,而不用依赖于由变焦操控引起的图像捕获设备 500 与目标对象之间的距离的改变。因此,可以生成具有高可视性的合成图像,而不用依赖于由变焦操控引起的图像捕获设备 500 与目标对象之间的距离的改变。此外,由于不需要用户的手工工作,因此可以减少用户的麻烦。

[0226] 此外,例如,在将图像捕获部分 111 所生成的捕获图像记录为图像内容的情况下,当与内容相关联地记录元数据时,可以记录校正之后的碰撞声音检测位置。因此,由于校正之后的碰撞声音检测位置可以被用在再现图像文件的情况中,因此可以检测碰撞声音的校正生成位置。此外,例如,在显示图像内容列表的情况下,可以将与校正后的碰撞声音检测位置对应的帧用作代表性缩略图。

[0227] 3. 第三实施例

[0228] 本发明的第一和第二实施例主要提供了将进行高尔夫球挥杆的人的动作过渡生成合成图像的示例。这里,例如,在除高尔夫球之外的体育运动(诸如球击比赛(ball hitting game)或对象碰撞比赛(object collision game))情况下,在选手击球或对象相互碰撞的瞬间生成特征声音。此外,例如,在焰火的情况下,在壮观的爆炸瞬间生成爆炸声音。当使用这些特定声音确定合成目标图像选择范围时,可以生成合适的合成图像。在这个方面,本发明的第三实施例提供了根据用户喜好容易地生成这样的合成图像的示例。

[0229] [图像捕获设备的配置示例]

[0230] 图 23 是说明根据本发明的第三实施例的、图像捕获设备 600 的功能性配置的示例的方块图。图像捕获设备 600 不同于图 17 中所示的图像捕获设备 500 在于:添加了目标对象信息存储部分 610,并且修改了一部分其他组件。也就是说,代替图像捕获设备 500 中的特定检测部分 132、有效声音范围设置部分 160 和选择部分 520,图像捕获设备 600 配备了特定声音检测部分 620、有效声音范围设置部分 630 和选择部分 640。除这些组件之外的配置几乎与图像捕获设备 500 相同。因此,将相同的参考数字赋予共用组件,并省略其一一部分的描述。

[0231] 在下文中,将描述适合于根据本发明的各实施例的合成目标图像的选择处理的特定动作(体育运动等)的条件(例如,第一条件和第二条件)的示例。该第一条件是目标对象的主体部分在整个全部动作中基本存在于相同位置。此外,该第二条件是牵涉碰撞声音的快速动作存在于一系列动作当中。作为除了高尔夫球之外满足这些条件的体育运动的示例,可以示范棒球击球、网球发球、屋瓦打碎(breaking of rooftiles,一种硬气功)等。此外,可以示范除体育运动之外的、诸如焰火之类的对象。因此,本发明的第三实施例提供这样的对象或体育运动变成目标对象的示例。

[0232] 目标对象信息存储部分 610 关于多种类型的目标对象的特定动作存储关于特定声音和有效声音范围的信息,并将相应被存储信息供应到特定声音检测部分 620、有效声音范围设置部分 630 和选择部分 640。将参考图 24 详细描述目标对象信息存储部分 610 的所存储的内容。目标对象信息存储部分 610 是公开在权利要求书中的存储部分的示例。

[0233] 特定声音检测部分 620 基于从声音输入部分 131 输出的声音检测特定声音,并在检测到特定声音的情况下将该事实输出到选择部分 640。此外,特定声音检测部分 620 检测涉及网球(发球)和屋瓦打碎的特定值或更高的声音。在这个方面,本发明的第三实施例

提供除了涉及高尔夫球挥杆的碰撞声音之外,在图 24 中的碰撞声音 611 中示出的每一个碰撞声音都变成特定声音的示例。例如,在棒球(击球)和网球(发球)中,在动作中球棒或球拍接触球的瞬间生成的声音变成碰撞声音。此外,例如,在屋瓦打碎中,在动作中手接触屋瓦的瞬间生成的声音变成碰撞声音。另外,在焰火中,在焰火在天空中爆炸的瞬间生成的声音变成碰撞声音。为了检测这些声音,可以使用根据本发明的第一实施例的特定声音检测方法。

[0234] 此外,特定声音检测部分 620、有效声音范围设置部分 630 和选择部分 640 从目标对象信息存储部分 610,获取关于多种类型的目标对象的特定动作之中由用户操控指定的特定动作的目标对象信息。然后,特定声音检测部分 620、有效声音范围设置部分 630 和选择部分 640 使用所获取的目标对象信息进行相应处理。在这些相应处理中,这个示例除了碰撞声音的值和有效声音范围等不同外,与本发明的第一和第二实施例中所示的示例几乎相同,因此将省略其描述。

[0235] [目标对象信息存储部分的存储示例]

[0236] 图 24 是说明根据本发明的第三实施例的、目标对象信息存储部分 610 的存储内容的示例的图。在目标对象信息存储部分 610 中存储碰撞声音 611、高速动作范围 612、有效声音范围 613 和合成目标图像选择范围 614。

[0237] 将用于通过特定声音检测部分 620 检测特定声音(碰撞声音)的特定声音标识信息(参考数据)存储在碰撞声音 611 中。图 24 中所示的碰撞声音 611 仅仅公开了指示相应特定声音的文字。

[0238] 在高速动作范围 612 中存储了在合成目标图像选择范围中、具有比用于合成目标图像的选择的其他间隔更接近的范围。例如,在由目标对象引起的特定动作是棒球(击球)的情况下,可以认为精确地确认了棒球球棒接触球的瞬间周围的动作过渡。因此,在合成目标图像选择范围中,设置在其中在该瞬间周围图像的间隔更接近用于合成目标图像的选择的其他间隔的高速动作范围。例如,在这种高速动作范围中,设置碰撞声音周围的预定范围。

[0239] 有效声音范围 613 存储通过有效声音范围设置部分 630 变成设置目标的有效声音范围。本发明的第三实施例提供将视频记录按钮的按压操控用作拍摄触发的示例。按照与本发明的第一和第二实施例中所示的示例类似的方式,在高尔夫球和棒球(击球)中,可以在作为目标的动作终止之后生成拍摄触发。在焰火的情况下,可以在焰火在天空中爆炸的瞬间生成拍摄触发。此外,按照与本发明的第一和第二实施例中所示的示例类似的方式,在高尔夫球、棒球(击球)和焰火中,可以将拍摄触发之前的特定范围设置为有效声音范围。

[0240] 另外,网球(发球)的情况下,可以在选手上仰(turn up)预定时间的瞬间生成拍摄触发。在屋瓦打碎中,可以在打碎屋瓦的人的面孔战斗精神饱满(full of fighting spirit)的瞬间生成拍摄触发。此外,在网球(发球)和屋瓦打碎中,可以将从拍摄触发初次检测到特定值或更高的声音之前的特定范围设置为有效声音范围。按照这种方式,在网球(发球)和屋瓦打碎的情况下,相应于检测到特定值或更高的声音的时间设置有效声音范围的后边缘。然而,在超过预定时间未检测到该声音的情况下,声音检测可以变成无效,然后可以设置新有效声音范围。对于这种新的有效声音范围,重新生成拍摄触发。因此,可以减少碰撞声音的错误检测。在这个方面,用特定声音检测部分 620 检测在网球(发球)

和屋瓦打碎中的特定值或更高的声音。

[0241] 在合成目标图像选择范围 614 中存储用合成目标图像选择范围确定部分 171 确定的合成目标图像选择范围。本发明的第三实施例提供将关于在有效声音范围中检测到的碰撞声音的特定范围设置为合成目标图像选择范围的示例。

[0242] [目标对象的指定屏幕的显示示例]

[0243] 图 25 是说明根据本发明的第三实施例的、显示部分 191 的显示示例（显示屏幕 660）的图。显示屏幕 660 配备有高尔夫球（挥杆）按钮 661、棒球（击球）按钮 662、网球（发球）按钮 663、屋瓦打碎按钮 664 和焰火按钮 665。此外，显示屏幕 660 配备有返回按钮 666 和确定按钮 667。

[0244] 按压高尔夫球（挥杆）按钮 661、棒球（击球）按钮 662、网球（发球）按钮 663、屋瓦打碎按钮 664 和焰火按钮 665 来指定作为在为动作过渡生成合成图像时的目标的体育运动类型。例如，在显示部分 191 由触摸板构成的情况下，可以通过想要的按钮的按压操控进行该指定。

[0245] 按压确定按钮 667 来确定用于指定作为合成图像的目标的体育运动类型的按压操控之后的指定。通过这种按压，设置合成目标图像记录方式。

[0246] 例如，按压返回按钮 666 来返回紧接在之前显示的显示屏幕。

[0247] 例如，在显示屏幕 660 中，在按压了想要的按钮（例如，网球（发球）按钮 663）并按压了确定按钮 667 的情况下，将指示该事实的操控信号从操控接收部分 140 输出到相应部分。然后，特定声音检测部分 620、有效声音范围设置部分 630 和选择部分 640 从目标对象信息存储部分 610 获取所指定的体育运动类型的目标对象信息（例如，关于网球（发球）的目标对象信息）。然后，特定声音检测部分 620、有效声音范围设置部分 630 和选择部分 640 使用所获取的目标对象信息进行相应处理。

[0248] [合成图像示例]

[0249] 图 26A 和 26B 是说明根据本发明的第三实施例的、由图层处理部分 180 生成的合成图像的示例的图。

[0250] 图 26A 说明了在进行棒球击球的人被成像的情况下的合成图像 671。也就是说，合成图像 671 是使用按压了在图 25 中所示的棒球（击球）按钮 662 并且按压了确定按钮 667 之后开始的图像捕获操作获得的运动图像生成的合成图像。

[0251] 图 26B 说明了在进行网球发球的人被成像的情况下的合成图像 672。也就是说，合成图像 672 是使用按压了在图 25 中所示的网球（发球）按钮 663 并且按压了确定按钮 667 之后开始的图像捕获操作获得的运动图像生成的合成图像。

[0252] 按照这种方式，在除了高尔夫球挥杆之外的特定动作期间，将变成检测目标的碰撞声音可以存在的时间跨度设置为有效声音范围，因此可以减少除了变成检测目标的碰撞声音之外的声音的错误检测。因此，在生成使用引起作为目标的碰撞声音的极快速动作的合成图像的情况下，可以高精度地检测碰撞声音的生成位置。此外，可以生成强调在碰撞声音的精确生成位置周围的合成图像，并生成具有高可见性的合成图像。在这种情况下，由于碰撞声音的生成位置的检测不需要用户的手工工作，因此可以减少用户的麻烦。当按照这种方式生成合成图像时，即使不习惯于图像捕获设备的操作的初学者也可以容易地根据用户喜好生成合适的合成图像。

[0253] 4. 第四实施例

[0254] 本发明的第一至第三实施例提供在其中用图像捕获设备进行的图像捕获操作并且使用在图像捕获操作的时间生成的碰撞声音生成合成图像的示例。这里,在为图像捕获设备记录的内容(例如,运动图像内容)生成合成图像的情况下,可以使用包括在内容中的碰撞声音生成合成图像。此外,例如,在高尔夫球比赛中,可以生成除了碰撞声音之外的特征声音,诸如高尔夫球挥杆之后观众的高呼。由此,可以代替视频记录按钮的按压操控或计时器设置,使用该特征声音作为触发设置有效声音范围。因此,本发明的第四实施例提供了这样一个示例,在其中代替视频记录按钮的按压操控或计时器设置,使用该特征声音设置有效声音范围,并且为图像捕获设备所记录的运动图像内容生成合成图像。

[0255] [图像处理设备的配置示例]

[0256] 图 27 是说明根据本发明的第四实施例的、图像捕获设备 700 的功能性配置的示例的方块图。图像处理设备 700 不同于图 23 中所示的图像捕获设备 600 在于:省略了声音输入部分 131 和计时器设置部分 150,并且修改了某些其他组件。也就是说,代替图像捕获设备 600 中的目标对象信息存储部分 610、图像捕获部分 111 和对象信息生成部分 120,图像捕获设备 700 配备有目标对象信息存储部分 710、输入部分 720 和对象信息生成部分 730。此外,代替图像捕获设备 600 中的特定声音检测部分 620、有效声音范围设置部分 630 和选择部分 640,图像捕获设备 700 配备有特定声音检测部分 740、有效声音范围设置部分 750 和选择部分 760。除了这些组件之外的配置几乎与图像捕获设备 600 相同。因此,相同的参考数字被赋予共用组件,并且将省略其一部分的描述。

[0257] 目标对象信息存储部分 710 存储关于特定声音或针对多个类型的目标对象的特定动作的有效声音范围的信息,并将相应所存储的信息供应到特定声音检测部分 740、有效声音范围设置部分 750 和选择部分 760。将参考图 28 详细描述目标对象信息存储部分 710 的存储内容。

[0258] 输入部分 720 是这样一个输入部分,通过它输入由诸如数字摄像机之类的图像捕获设备所记录的图像内容(例如,运动图像文件)。例如,假设该图像内容是这样的内容,在其中将涉及动作图像或连续静止图像的图像捕获操作时的相应帧的元数据(例如,声音信息,变焦信息或聚焦位置信息)与相应帧相关联地进行记录。输入部分 720 以预定间隔,将形成输入的图像内容的相应图像(帧)供应到捕获图像保持部分 115、显示控制部分 190 和特定声音检测部分 740。此外,输入部分 720 将形成输入的图像内容的相应图像和关于图像的属性信息(例如,镜头位置和聚焦位置)输出到对象距离计算部分 510 和对象信息生成部分 730。

[0259] 对象信息生成部分 730 基于从输入部分 720 供应的相应图像和关于图像的属性信息,生成关于相应图像的对象信息。此外,对象信息生成部分 730 将所生成的对象信息供应到捕获图像保持部分 115 以在那里保持对象信息,还将对象信息供应到有效声音范围设置部分 750 和选择部分 760。作为对象信息,例如,除了用于辨别目标对象的区域与其背景区域的信息之外,还生成用于检测相应图像之间在时间轴上的特定变化的信息。例如,检测包括在图像中的面孔,并且生成关于面孔的表情、方向等的属性信息。例如,可以检测诸如高兴、愤怒、悲伤或愉快的表情、紧张、面孔方向嘴巴表情(口/闭)、眼睛表情(例如,眨眼)等之类的属性,并且可以基于检测结果生成属性信息。作为属性检测方法,例如,可

以使用基于从涉及变成检测目标的属性的面孔图像中提取的特征量的标识方法。也就是说,从涉及变成检测目标的属性的面孔图像中提取的特征量被事先存储为标识词典。此外,从包括所检测到的面孔的面孔图像中提取特征量,并将所提取的特征量与包括在标识词典中的特征量比较,从而计算这些特征量之间的相似性。另外,在所计算出的相似性超过阈值的情况下,确定所检测到的面孔是与变成超过阈值的相似性的计算目标的标识词典对应的属性。而且,例如,可以使用上述属性检测方法(例如,参考日本未审专利申请公开第2009-118009号)。

[0260] 有效声音范围设置部分 750 基于按照这种方式生成的属性信息检测相应图像之间在时间轴上的特定变化。作为这种特定变化,例如,可以检测面孔变化。例如,可以将高兴、愤怒、悲伤、愉快或紧张时的表情变化、面孔方向的变化(例如在面孔方向从下面改变到正面的情况下)、开口/闭口时的变化或眨眼时的眼睛变化检测为面孔变化。此外,可以将人的面孔出现在屏幕或人的面孔离开屏幕的情况检测为相应图像之间在时间轴上的特定变化。此外,作为用于检测相应图像之间在时间轴上的特定变化的信息,可以生成指示整个图像的明亮度(brightness)的明亮度信息和指示图像中的亮度值(luminance value)的分布状态的亮度值分布信息。然后,可以将信息中的特定变化检测为相应图像之间在时间轴上的特定变化。

[0261] 特定声音检测部分 740 基于包括在从输入部分 720 提供的图像内容中的声音检测特定声音,然后在检测到特定声音的情况下,将该事实输出到有效声音范围设置部分 750 和选择部分 760。通过特定声音检测部分 740 变成检测目标的碰撞声音与本发明的第三实施例中的相同。此外,在本发明的第四实施例中,当检测到相应图像之间在时间轴上的特定变化时,使用通过特定声音检测部分 740 检测到的特定声音。例如,可以将举行高尔夫球或棒球赛的情况下观众的高呼(狂热欢呼)或在使战斗精神注入屋瓦打碎的瞬间的喊叫检测为特定声音。此外,例如,可以将吹奏人的吹奏声(whistling sound)或泼水声(例如,在目标对象涉及水的情况下)检测为特定声音。此外,有效声音范围设置部分 750 检测特定声音的变化作为相应图像在时间轴上的特定变化。例如,可以将从声音刚刚存在的状态到检测到高呼的时间,或者从检测到高呼的状态到声音刚刚存在的时间,检测为相应图像之间在时间轴上的特定变化。作为特定声音检测方法,可以应用本发明的第一实施例所示的特定声音检测方法。也就是说,有效声音范围设置部分 750 可以使用从相应图像提取的特征量和基于与相应图像相关联的声音提取的特征量中至少之一来检测特定变化。有效声音范围设置部分 750 是公开在权利要求书中的特定变化检测部分的示例。

[0262] 特定声音检测部分 740、有效声音范围设置部分 750 和选择部分 760 从目标对象信息存储部分 710,获取关于在多种类型的目标对象的特定动作之中由用户操控指定的特定动作的目标对象信息。此外,特定声音检测部分 740、有效声音范围设置部分 750 和选择部分 760 使用所获取的目标对象信息进行相应处理。除了当设置了碰撞声音时的触发、碰撞声音、有效声音范围的值等不同之外,相应处理几乎与本发明的第一至第三实施例中所示的示例相同,因此将省略其描述。

[0263] [目标对象信息存储部分的存储示例]

[0264] 图 28 是说明根据本发明的第四实施例的、目标对象信息存储部分 710 的存储内容的示例的图。在目标对象信息存储部分 710 中存储碰撞声音 711、高速动作范围 712、用于

确定有效声音范围的触发 713、有效声音范围 714 和合成目标图像选择范围 715。碰撞声音 711、高速动作范围 712 和合成目标图像选择范围 715 与图 24 中所示的碰撞声音 611、高速动作范围 612 和合成目标图像选择范围 614 相同,因此省略其描述。

[0265] 用于确定有效声音范围的触发 713 存储由有效声音范围设置部分 750 设置了有效声音范围时的触发。例如,在高尔夫球和棒球(击球)的情况下,可以在从输入部分 720 所输入的内容中检测到观众的高呼的瞬间生成触发。基于特定声音检测部分 740 所检测到的特定声音,由有效声音范围设置部分 750 进行观众欢呼的检测。此外,在焰火的情况下,可以在从输入部分 720 输入的内容中在屏幕中的黑暗状态被过渡到明亮状态的瞬间生成触发。基于由对象信息生成部分 730 所生成的属性信息,由有效声音范围设置部分 750 进行屏幕中的黑暗状态被过渡到明亮状态的瞬间的检测。此外,在网球(发球)的情况下,可以在从输入部分 720 输入的内容中在选手上仰预定时间的瞬间生成触发。基于由对象信息生成部分 730 所生成的属性信息,由有效声音范围设置部分 750 进行选手上仰预定时间的瞬间的检测。此外,在屋瓦打碎的情况下,可以在从输入部分 720 输入的内容中在进行屋瓦打碎的人的面孔战斗精神饱满的瞬间生成触发。基于由对象信息生成部分 730 所生成的属性信息或由特定声音检测部分 740 检测到的特定声音,由有效声音范围设置部分 750 进行进行屋瓦打碎的人的面孔战斗精神饱满的瞬间的检测。

[0266] 有效声音范围 714 存储通过有效声音范围设置部分 750 变成设置目标的有效声音范围。本发明的第四实施例提供这样一个示例,在其中存储在用于确定有效声音范围的触发 713 中的特定变化的检测时间是有效声音范围的触发。除了存储在用于确定有效声音范围的触发 713 中的特定变化的检测时间是有效声音范围的触发之外,这个有效声音范围几乎与图 24 中所示的有效声音范围 613 相同,因此省略其描述。

[0267] [合成目标图像选择范围的确定示例]

[0268] 图 29A 和 29B 是示意地说明根据本发明的第四实施例的、有效声音范围设置部分 750 的有效声音范围的设置方法和合成目标图像选择范围确定部分 171 的合成目标图像选择范围的确定方法的图。这是在其中在指定图 28 中所示的棒球(击球)的情况下确定合成目标图像选择范围的示例。此外,这个示例是图 5A 和 5B 的修改示例,因此将省略一部分共用组件的描述。

[0269] 图 29A 以矩形形状示意地说明经由输入部分 720 输入的运动图像 800。例如,运动图像 800 是记录棒球赛的样子运动图像。此外,在形成运动图像 800 的相应图像中,将特定声音检测部分 740 检测到观众的高呼的位置表示为在图 29A 中所示的运动图像 800 在时间轴上的高呼检测位置 801。在这种情况下,有效声音范围设置部分 750 基于高呼检测位置 801 设置有效声音范围 802。具体地说,有效声音范围设置部分 750 将一个范围设置为有效声音范围 802,该范围在时间轴上领先高呼检测位置 801 时间 L31。这里,按照与图 5A 中所示的时间 L1 类似的方式,例如,时间 L31 可以是 1.0 秒。此外,时间 L31 可以长于时间 L1。

[0270] 例如,考虑到在高尔夫球赛中选手具有非常大的射程(great shot)时的碰撞声音的生成位置出现在观众的高呼之前。因此,将有效声音范围设置(时间轴上领先的位置)在观众的高呼之前,从而使得可以适当地检测碰撞声音。

[0271] 图 29B 示意地说明在由有效声音范围设置部分 750 所设置的有效声音范围 802 中检测到碰撞声音的情况下合成目标图像选择范围的确定方法。在图 29B 中,将由特定声音

检测部分 740 检测到碰撞声音的位置（时间轴上的位置）表示为碰撞声音检测位置 803。

[0272] 例如，合成目标图像选择范围确定部分 171 确定碰撞声音检测位置 803 是否包括在由有效声音范围设置部分 750 所设置的有效声音范围 802 中。如图 29B 所示，在碰撞声音检测位置 803 包括在有效声音范围 802 中的情况下，合成目标图像选择范围确定部分 171 基于碰撞声音检测位置 803 确定合成目标图像选择范围。也就是说，合成目标图像选择范围确定部分 171 参考碰撞声音检测位置 803，将时间轴上在碰撞声音检测位置 803 之前和之后的特定范围确定为合成目标图像选择范围 804。具体地说，将这样一个范围 L34 确定为合成目标图像选择范围 804，该范围 L34 包括参考碰撞声音检测位置 803 在时间轴上领先时间 L32 的范围以及超过碰撞声音检测位置 803 时间 L33 的范围。

[0273] 这里，例如，时间 L32 至 L34 可以是与图 5B 中所示的时间 L2 至 L4 相同的值。此外，时间 L34 可长于时间 L4。

[0274] 图 30A 和 30B 是示意地说明根据本发明的第四实施例的、有效声音范围设置部分 750 的有效声音范围的设置方法和合成目标图像选择范围确定部分 171 的合成目标图像选择范围的确定方法的图。这个示例是在其中在指定图 28 中的网球（发球）的情况下确定合成目标图像选择范围的示例。此外，这个示例是图 5A 和 5B 中所示的修改示例，因此将省略一部分共用组件的描述。

[0275] 图 30A 示意地以矩形形状说明经由输入部分 720 输入的运动图像 810。例如，运动图像 810 是在其中记录网球赛的样子运动图像。此外，在用于形成运动图像 810 的相应图像中，将由对象信息生成部分 730 检测到的、网球选手的面孔上仰预定时间或更长的位置表示为时间轴上的运动图像 800 的面孔上仰状态检测位置 811。此外，将特定声音检测部分 740 检测到特定值或更高的声音的位置（时间轴上的位置）表示为特定值或更高的声音检测位置 812。在这种情况下，有效声音范围设置部分 750 基于面孔上仰状态检测位置 811 和特定值或更高的声音检测位置 812 设置有效声音范围 813。具体地说，有效声音范围设置部分 750 参考面孔上仰状态检测位置 811 和特定值或更高的声音检测位置 812，在时间轴上设置有效声音范围 813（时间 L41）。这里，例如，时间 L41 是面孔上仰状态检测位置 811 是开始点而特定值或更高的声音检测位置 812 是结束点的范围

[0276] 在这个方面，在涉及网球（发球）的有效声音范围的端点上，特定值或更高的声音的检测是设置条件。这里，在不从有效声音范围的开始点检测到特定值或更高的声音预定时间或更长的情况下，有效声音范围设置部分 750 使得涉及开始点的有效声音范围失效，并且可以等待用于确定有效声音范围的新触发。因此，可以减少碰撞声音的错误检测。此外，可以类似地应用涉及屋瓦打碎的有效声音范围的端点。

[0277] 图 30B 示意地说明在由有效声音范围设置部分 750 所设置的有效声音范围 813 中检测到碰撞声音的情况下合成目标图像选择范围的确定方法。在图 30B 中，将特定声音检测部分 740 检测到碰撞声音的位置（时间轴上的位置）表示为碰撞声音检测位置 814。

[0278] 例如，合成目标图像选择范围确定部分 171 确定碰撞声音检测位置 814 是否包括在由有效声音范围设置部分 750 所设置的有效声音范围 813 中。如图 30B 所示，在碰撞声音检测位置 814 包括在有效声音范围 813 的情况下，合成目标图像选择范围确定部分 171 基于碰撞声音检测位置 814 确定合成目标图像选择范围。也就是说，合成目标图像选择范围确定部分 171 参考碰撞声音检测位置 814，将在时间轴上在碰撞声音检测位置 814 之前和之

后的特定范围确定为合成目标图像选择范围 815。具体地说,将这样一个范围 L44 确定为合成目标图像选择范围 815,该范围 L44 包括在时间轴上领先碰撞声音检测位置 814 时间 L42 的范围和超过碰撞声音检测位置 814 时间 L43 的范围。

[0279] 这里,如图 28 所示,时间 L42 可以例如是 2.5 秒,而时间 L43 可以例如是 0.5 秒。也就是说,时间 L44 可以是 3.0 秒。时间 L42 至 L44 的设置内容可以根据用户的喜好改变。

[0280] 此外,在图 29A、29B 以及图 30A 和 30B 中所示的示例中,在像如本发明的第二实施例所示那样在有效声音范围中是否检测到碰撞声音的确定之前,可以估计声音的延迟时间,并且可以基于估计结果校正碰撞声音检测位置。按照这种方式,校正碰撞声音的检测位置,从而使得可以基于校正后的位置确定合成目标图像选择范围。此外,在作为有效声音范围的触发的特定声音(例如,观众的高呼)中,可以估计声音延迟时间,并且可以基于估计结果校正检测位置。例如,在特定声音为观众的高呼的情况下,使用涉及相应帧的元数据(例如,变焦信息或聚焦位置信息)计算到目标对象的背景的对象距离,并且可以基于该对象距离来估计观众的高呼的延迟时间。因此,即使在其中比较远的目标对象通过变焦功能变成图像捕获目标的运动图像内容中,也可以确定合适的合成目标图像选择范围。

[0281] 此外,在图 29A 和 29B 以及图 30A 和 30B 中所示的示例中,可能不能在有效声音范围设置部分 750 所设置的有效声音范围中检测到碰撞声音。在这种情况下,不使用碰撞声音生成合成图像,并且可以进行下一合成图像处理。

[0282] 按照这种方式,关于由诸如数字视频摄像机之类的图像捕获设备记录的图像内容,可以使用在相应图像之间在时间轴上的特定变化设置有效声音范围。因此,在生成在其中引起碰撞声音的极快速动作是目标的合成图像的情况下,可以以高水平的精度检测碰撞声音的生成位置,而不用按压视频记录按钮等。此外,可以生成强调在碰撞声音的精确生成位置周围的合成图像,并生成具有高可视性的合成图像。此外,由于碰撞声音的生成位置的检测也不需要用户的手工工作,因此可以减少用户的麻烦。当按照这种方式生成合成图像时,即使不习惯于图像捕获设备的操作的初学者也可以容易地根据用户的喜好生成合适的合成图像。

[0283] 用于使用在相应图像之间在时间轴上的特定变化设置有效声音范围的设置方法可以被使用到在本发明的第一至第三实施例中所示的图像捕获设备中。

[0284] 此外,在本发明的实施例中,由合成目标图像选择部分 172 选择的合成目标图像作为静止图像用于合成图像的生成处理。这里,例如,所选择的合成目标图像可以用于在其中将生成合成图像的处理显示为动作图像的合成图像生成处理。此外,所选择的合成目标图像可以用作合成目标,也可以用作指示特定动作的过渡的多个图像(例如,用在幻灯片中的图像)。

[0285] 本实施例可以应用到诸如具有成像功能的移动电话、个人计算机、视频系统、编辑设备之类的图像处理设备。此外,本发明所呈现的实施例中的处理过程也可以以处理程序提供。

[0286] 本发明的实施例示范了用于实现本发明的示例。这里,像在本发明的实施例中显然看出来那样,本发明的实施例中的公开分别对应于权利要求书中的公开。类似地,权利要求书中的公开分别对应于具有相同参考数字的本发明的实施例中的公开。这里,本发明不限于这些实施例,可以在不脱离本发明的精神的范围内进行各种修改。

[0287] 此外,本发明的实施例中描述的处理过程可以提供为具有过程序列的方法。此外,这些处理过程可以提供为程序,该程序使得过程序列在计算机和记录该程序的记录介质中执行。可以将 CD(致密盘), MD(迷你盘), DVD(数字多功能盘), 存储器卡, 蓝光盘(注册商标)等用作这种记录介质。

[0288] 本申请包含涉及在于 2009 年 11 月 20 日在日本专利局提交的日本优先权专利申请 JP 2009-264616 中公开的内容的主题,该优先权专利申请的全部内容通过参考合并在这里。

[0289] 本领域技术人员应该理解,依赖于设计要求和因素,可以发生各种修改、组合、部分组合和变更,只要它们在所附的权利要求书或其等价物的范围内。

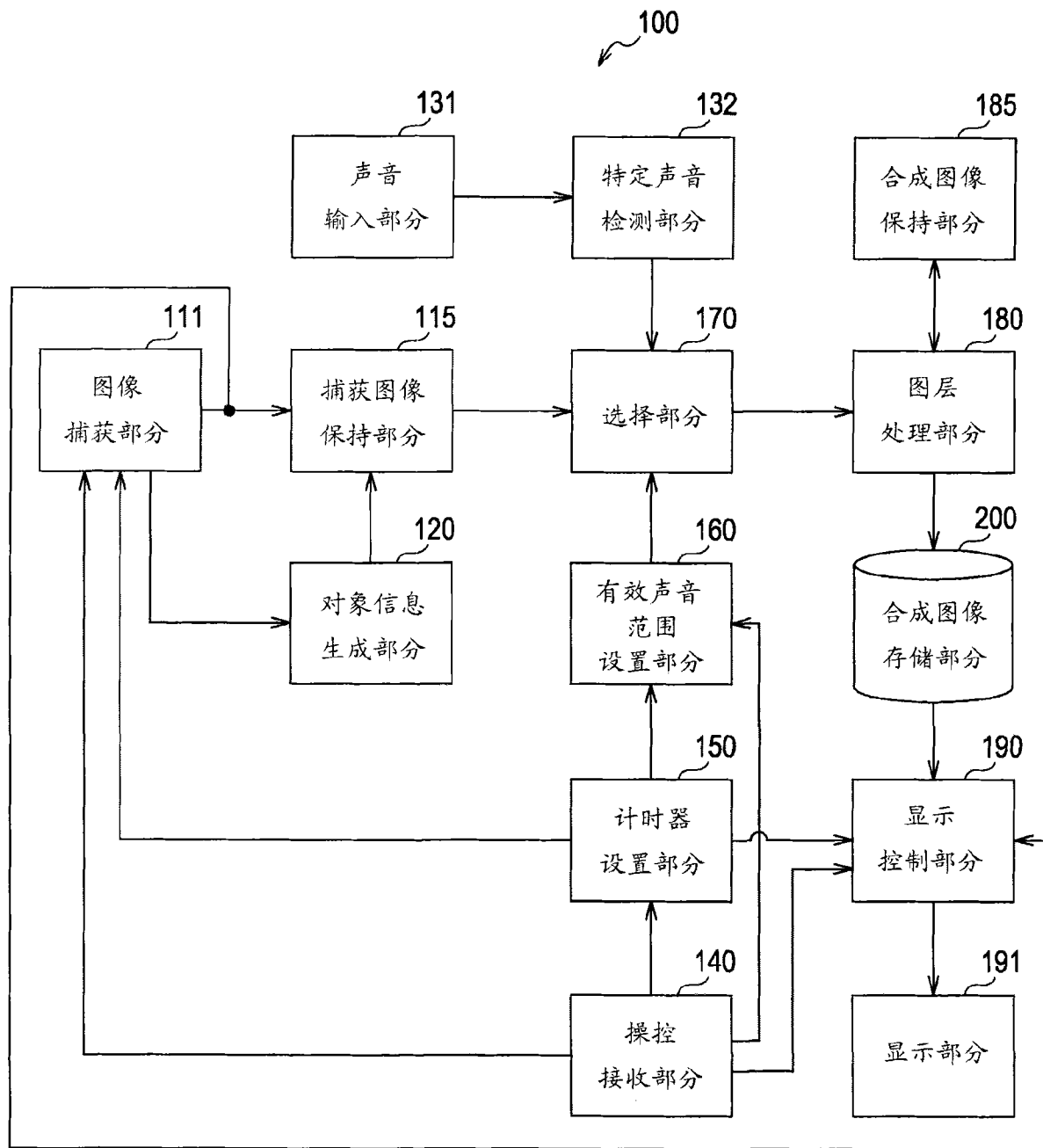


图 1

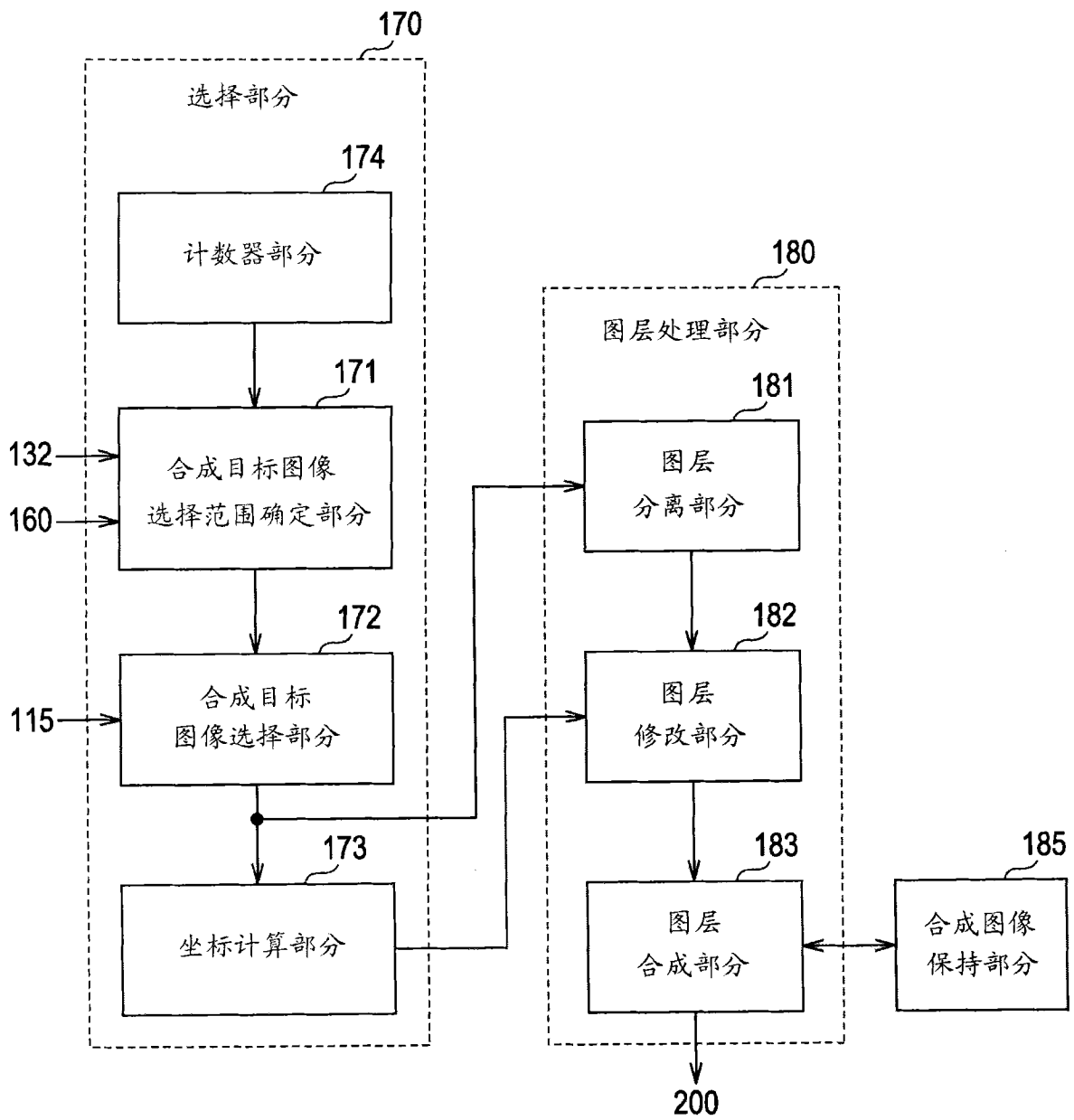


图 2

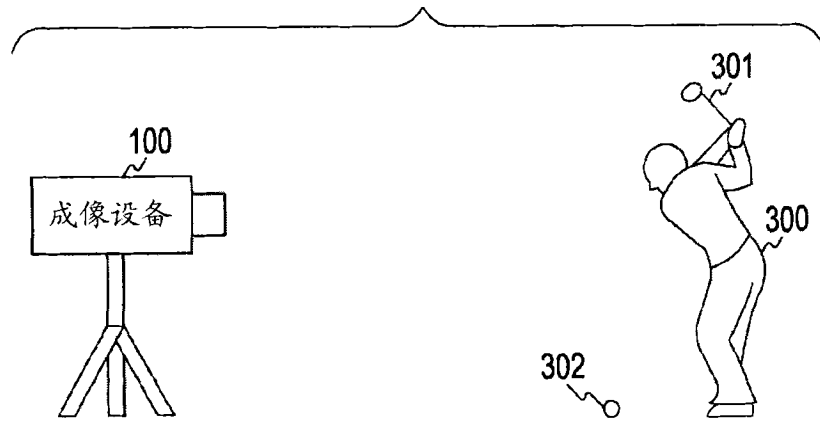


图 3A

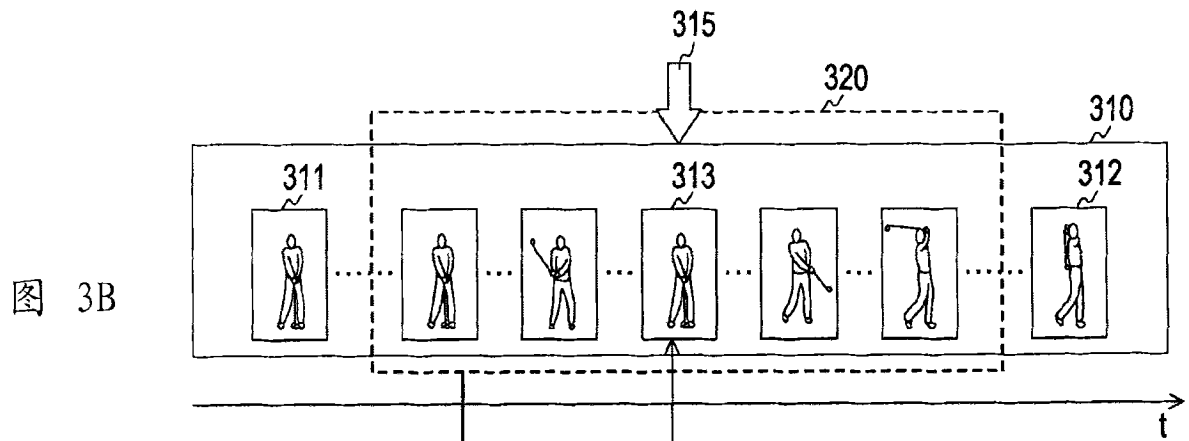


图 3B

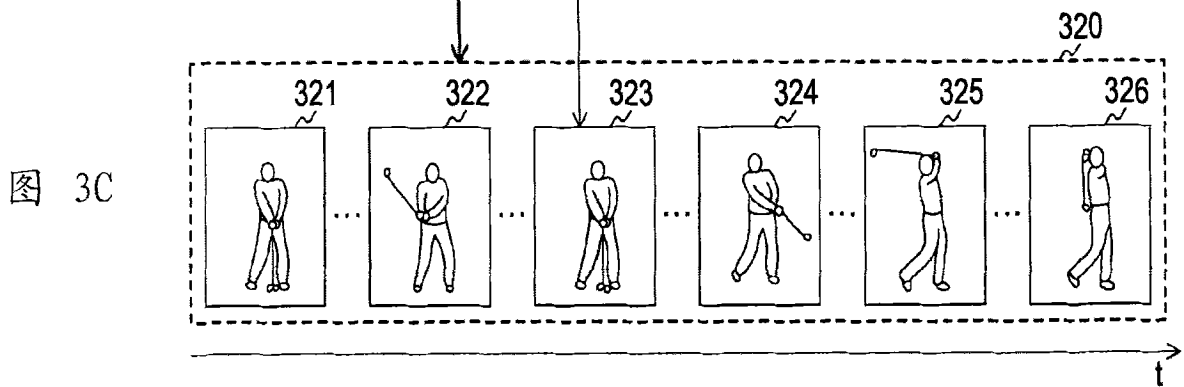


图 3C

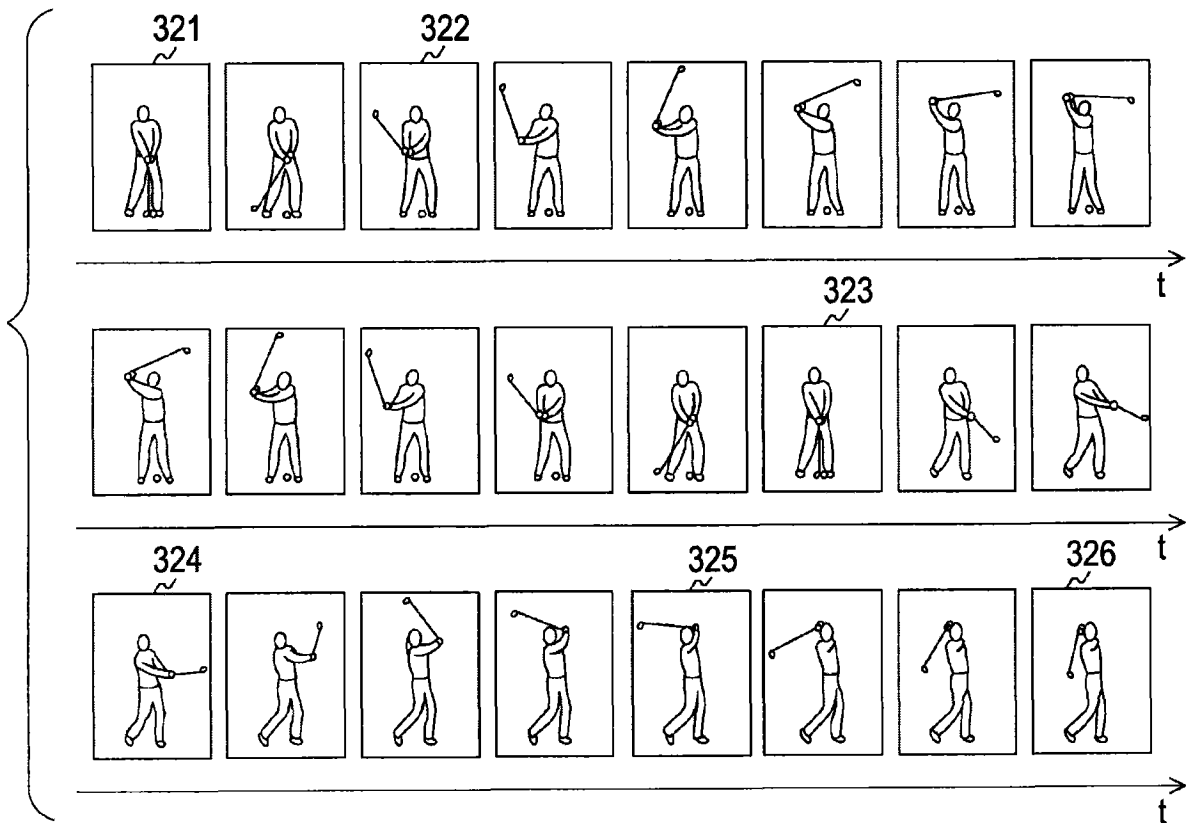


图 4A

图像合成

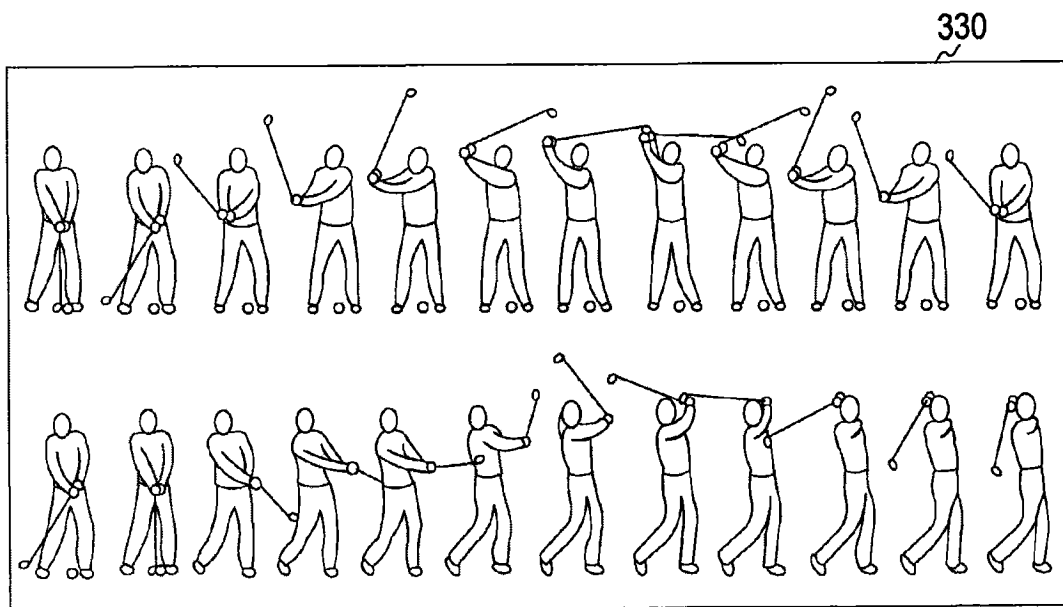


图 4B

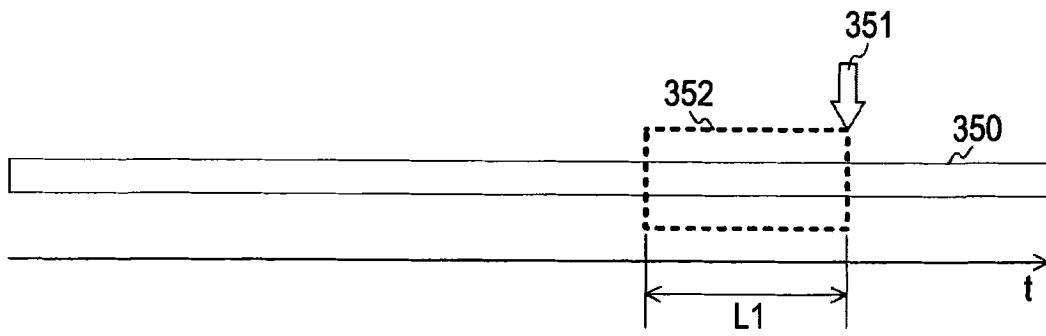


图 5A

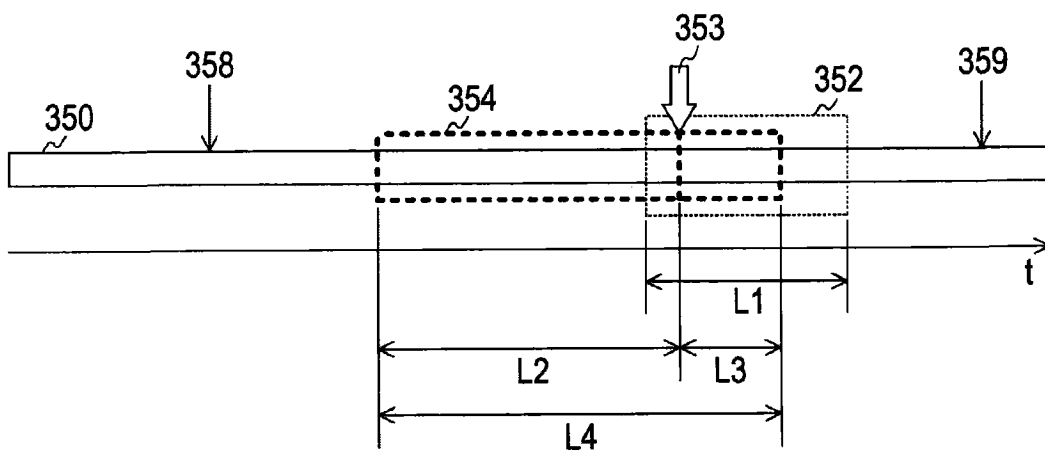


图 5B

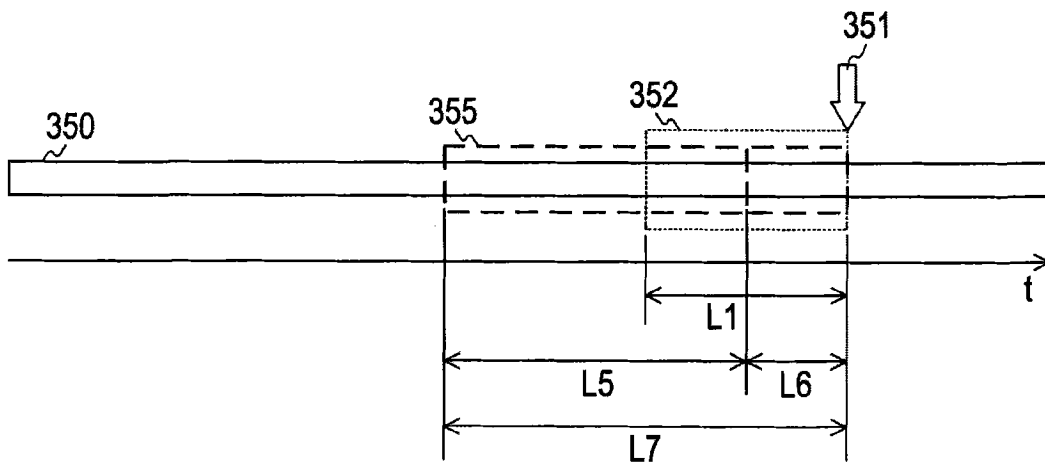


图 5C

图 6A

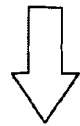
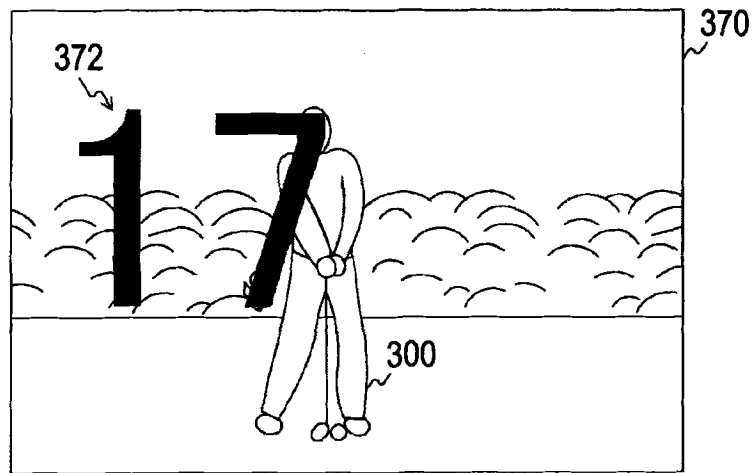
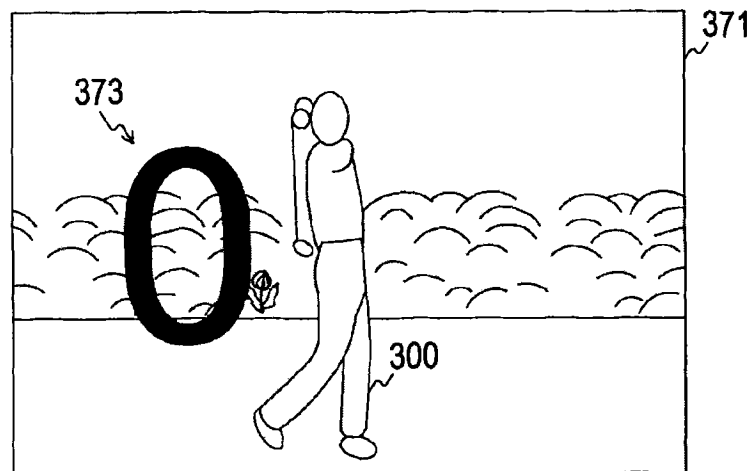


图 6B



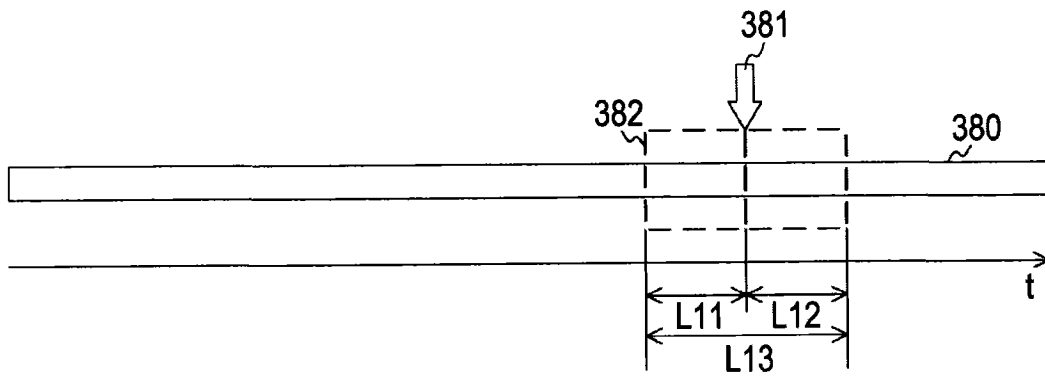


图 7A

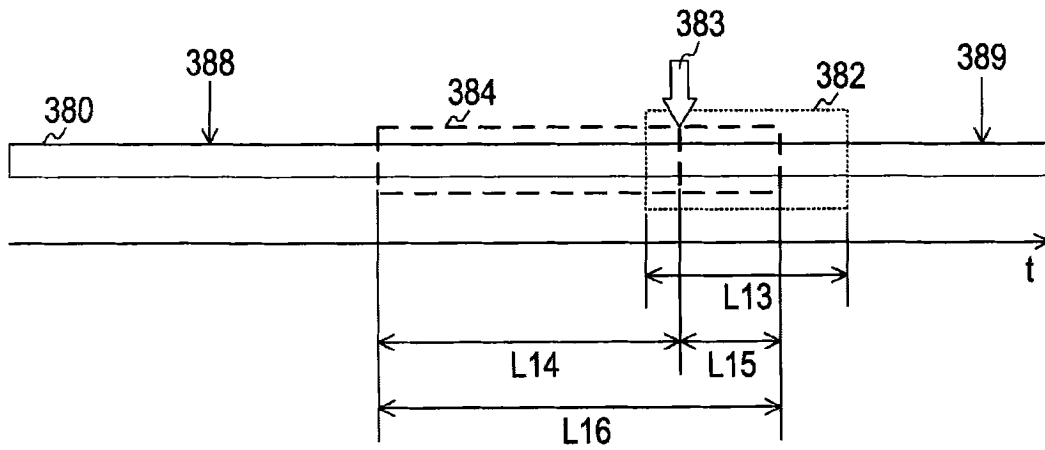


图 7B

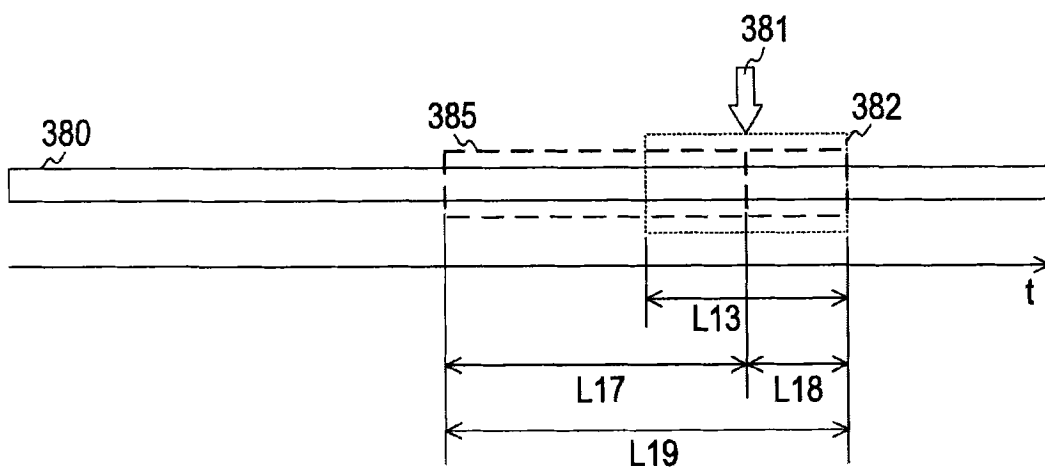


图 7C

图 8A

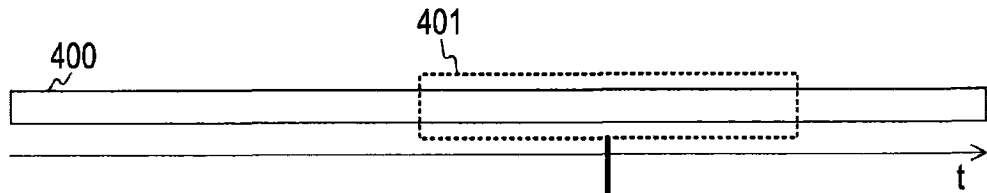


图 8B

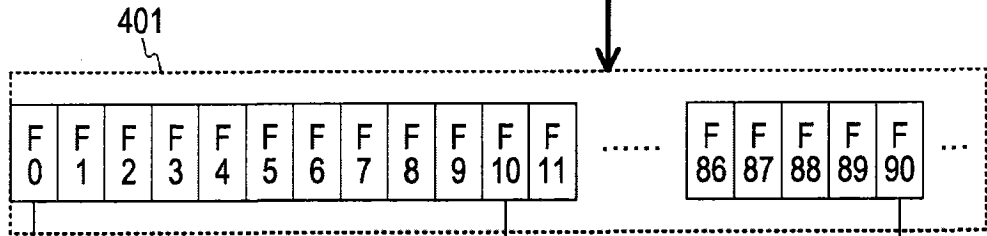


图 8C

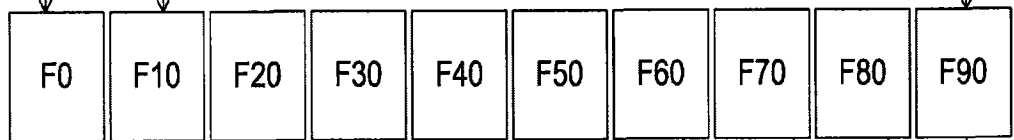
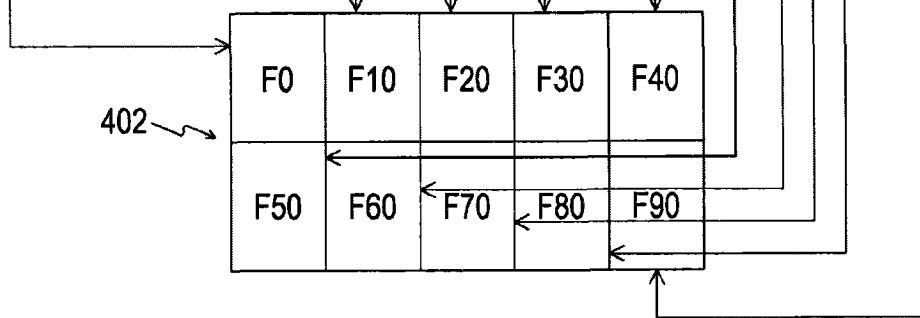
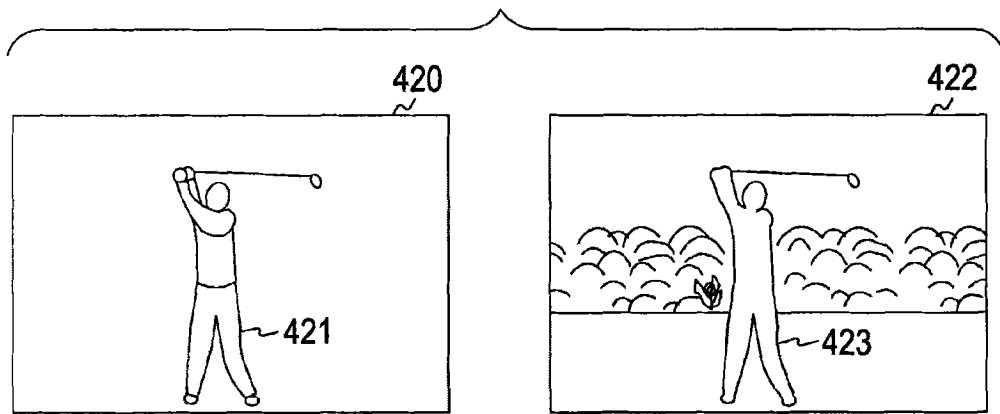
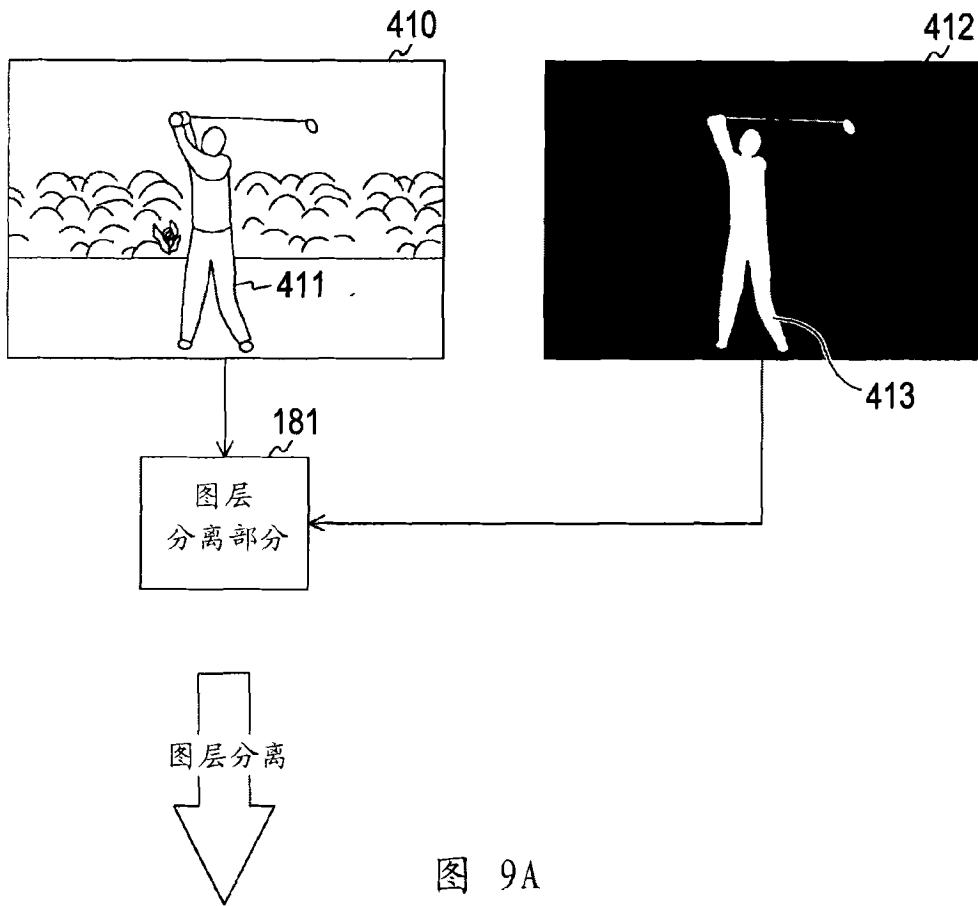


图 8D





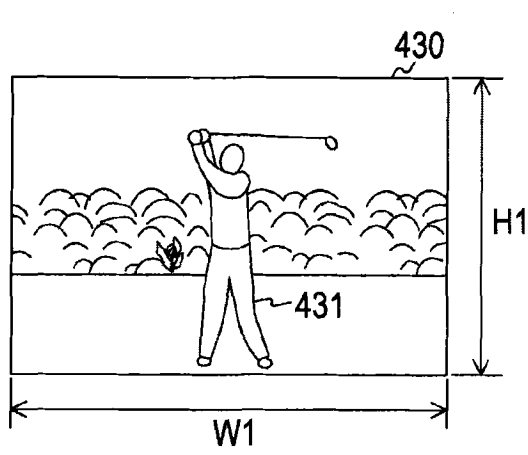


图 10A

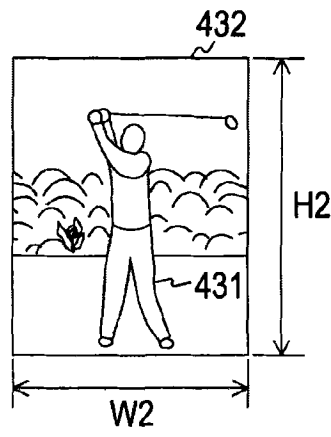


图 10B

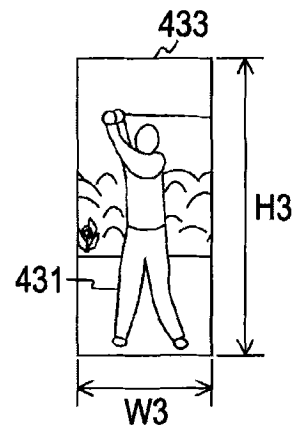


图 10C

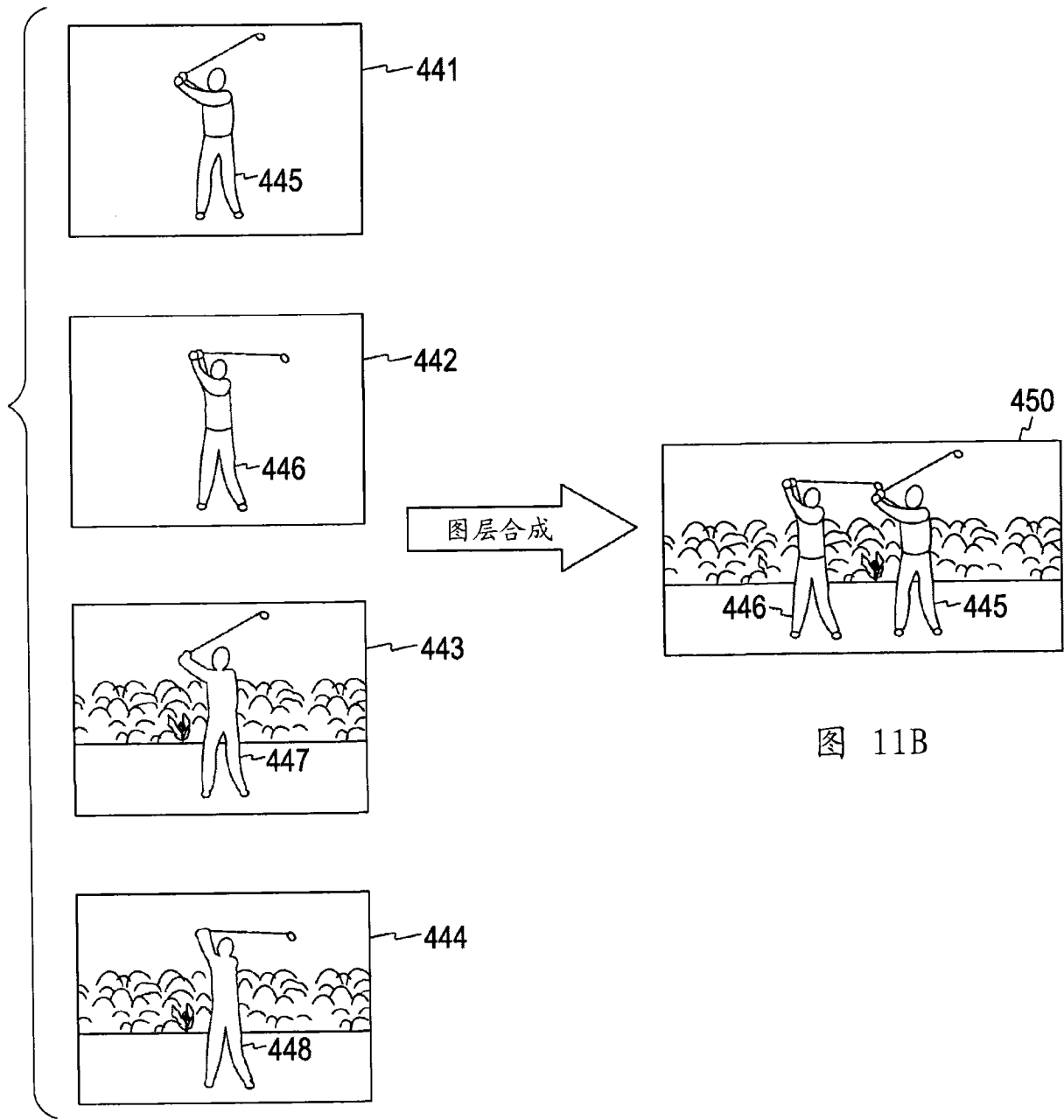


图 11A

图 11B

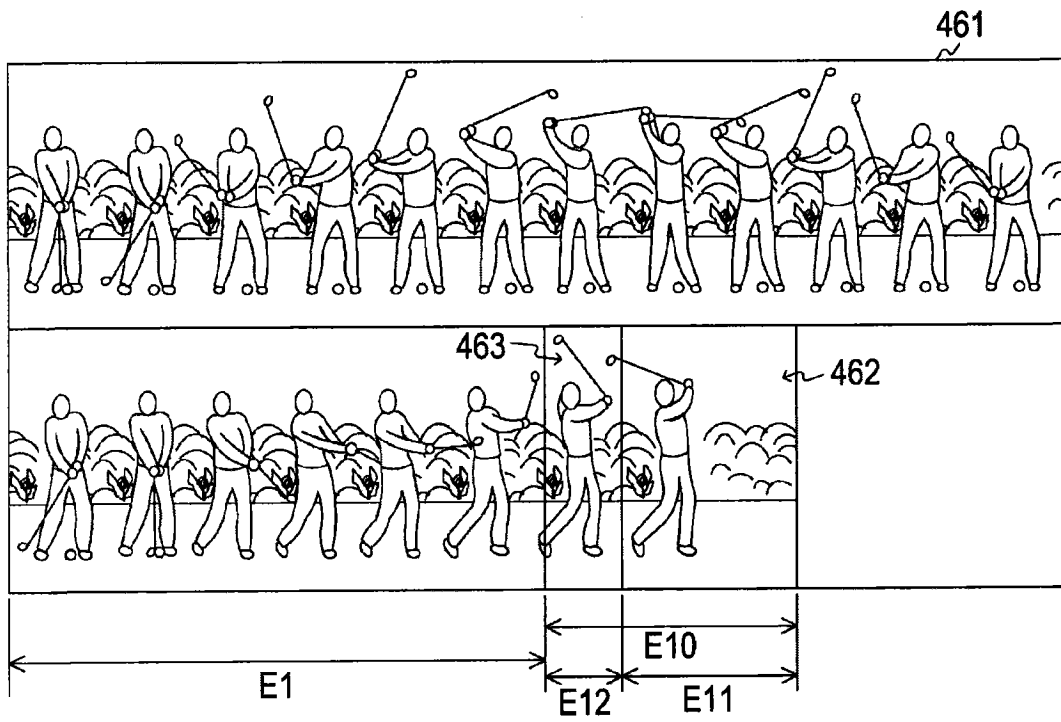


图 12A

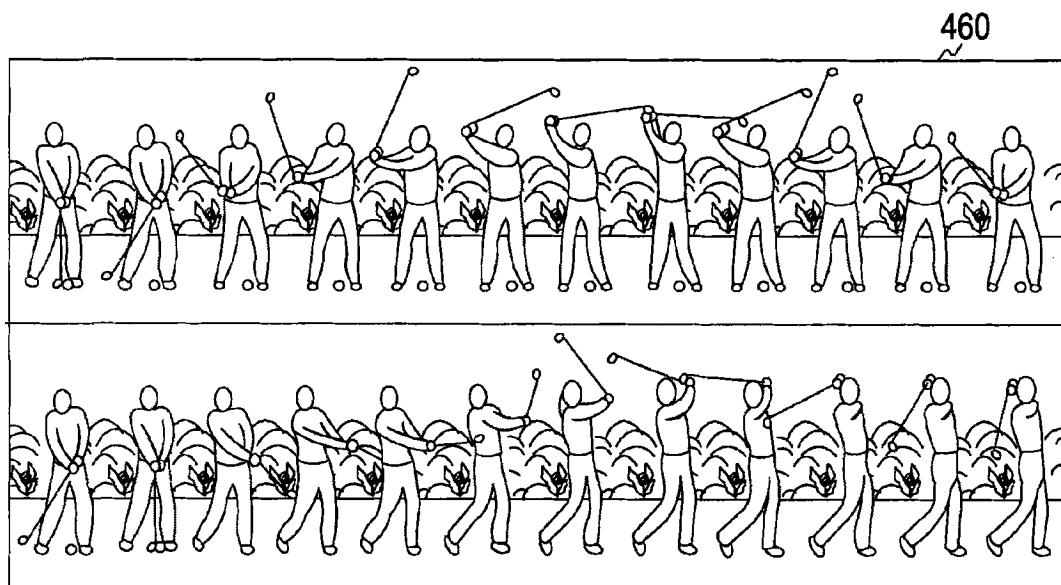


图 12B

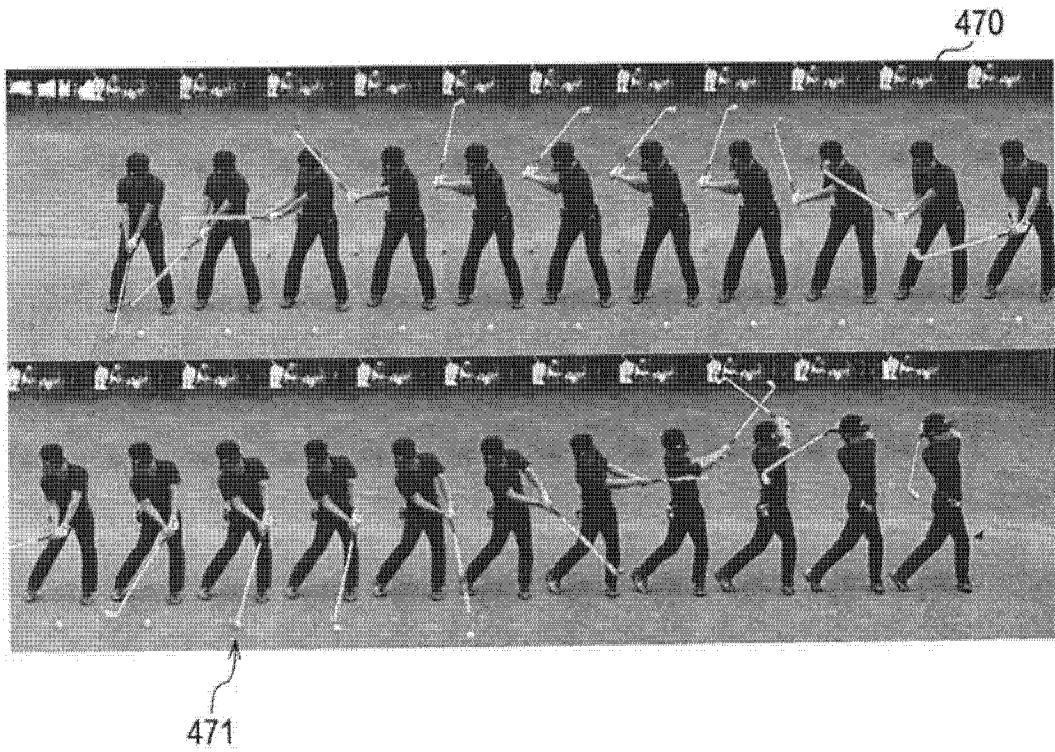


图 13

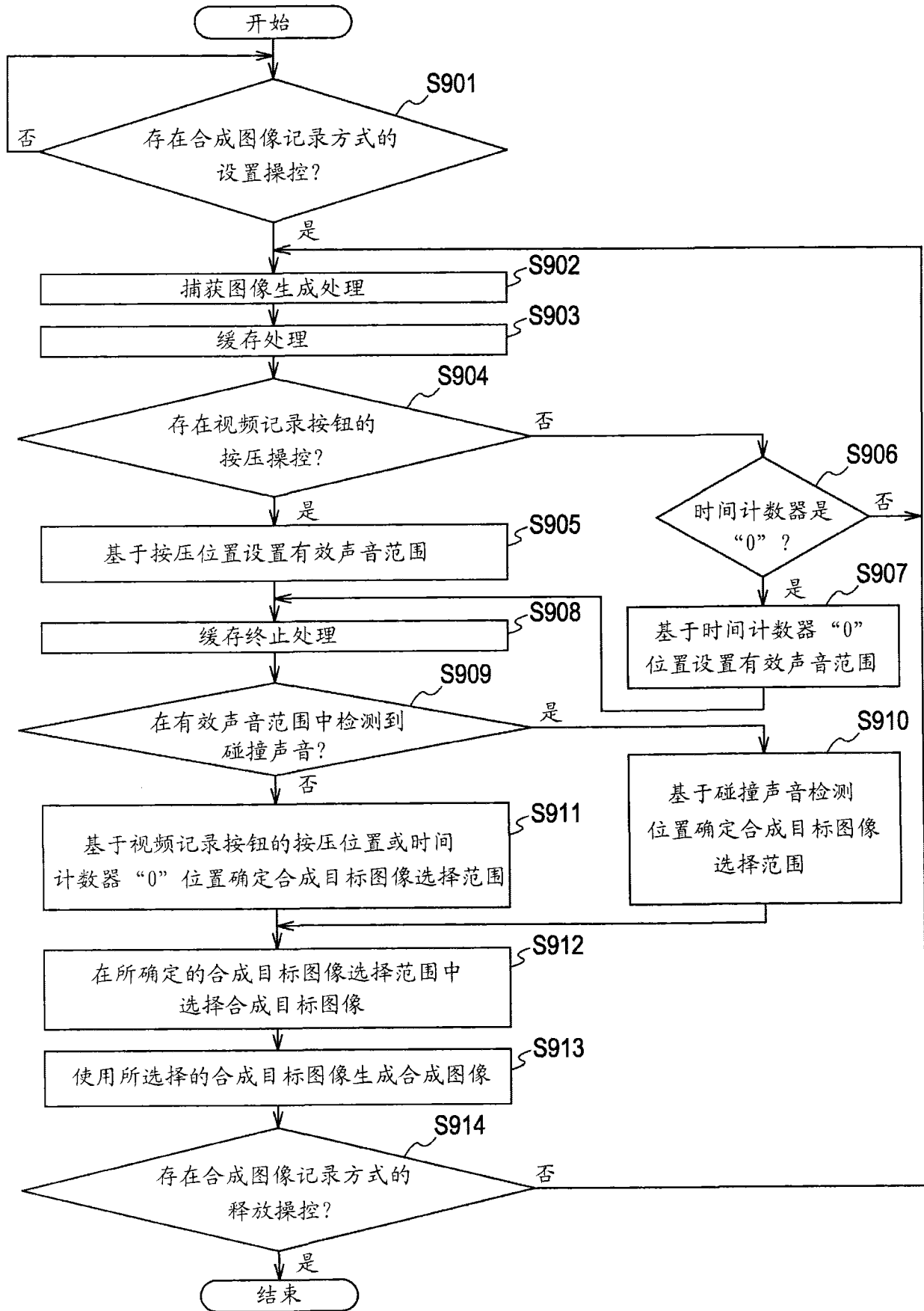


图 14

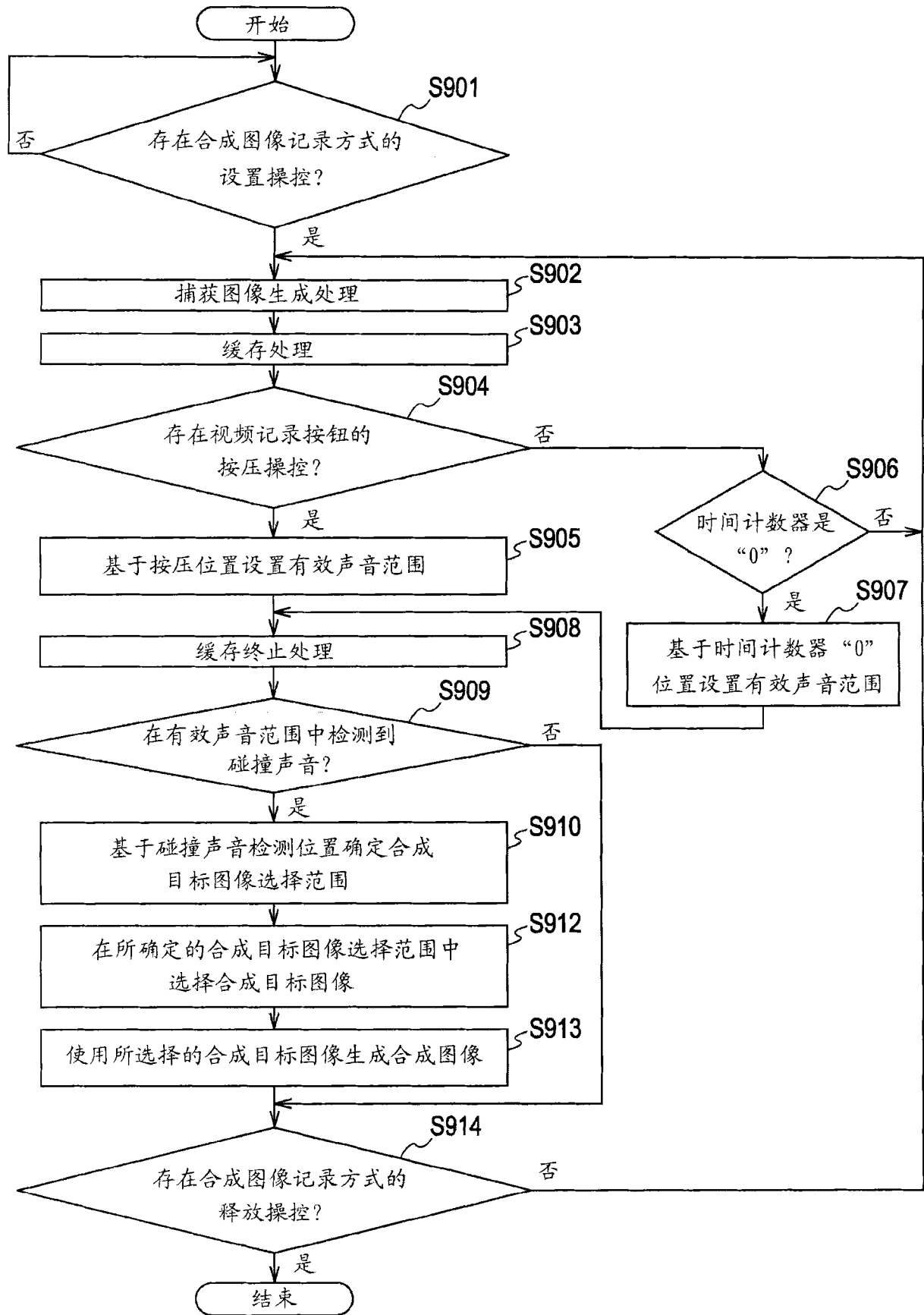


图 15

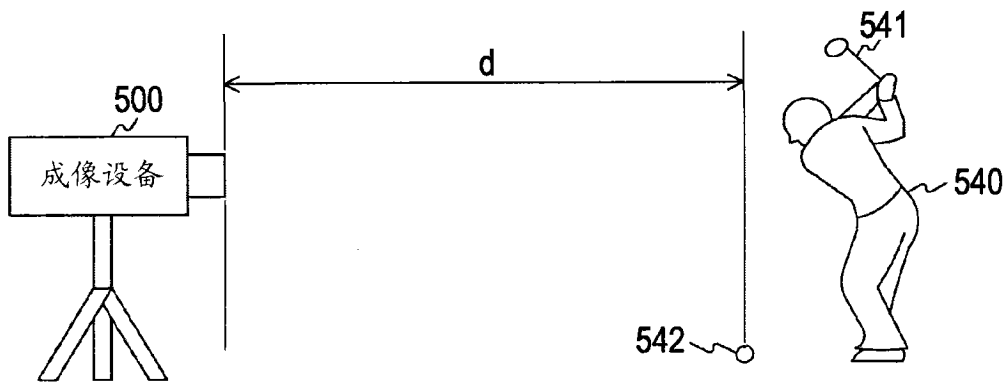


图 16A

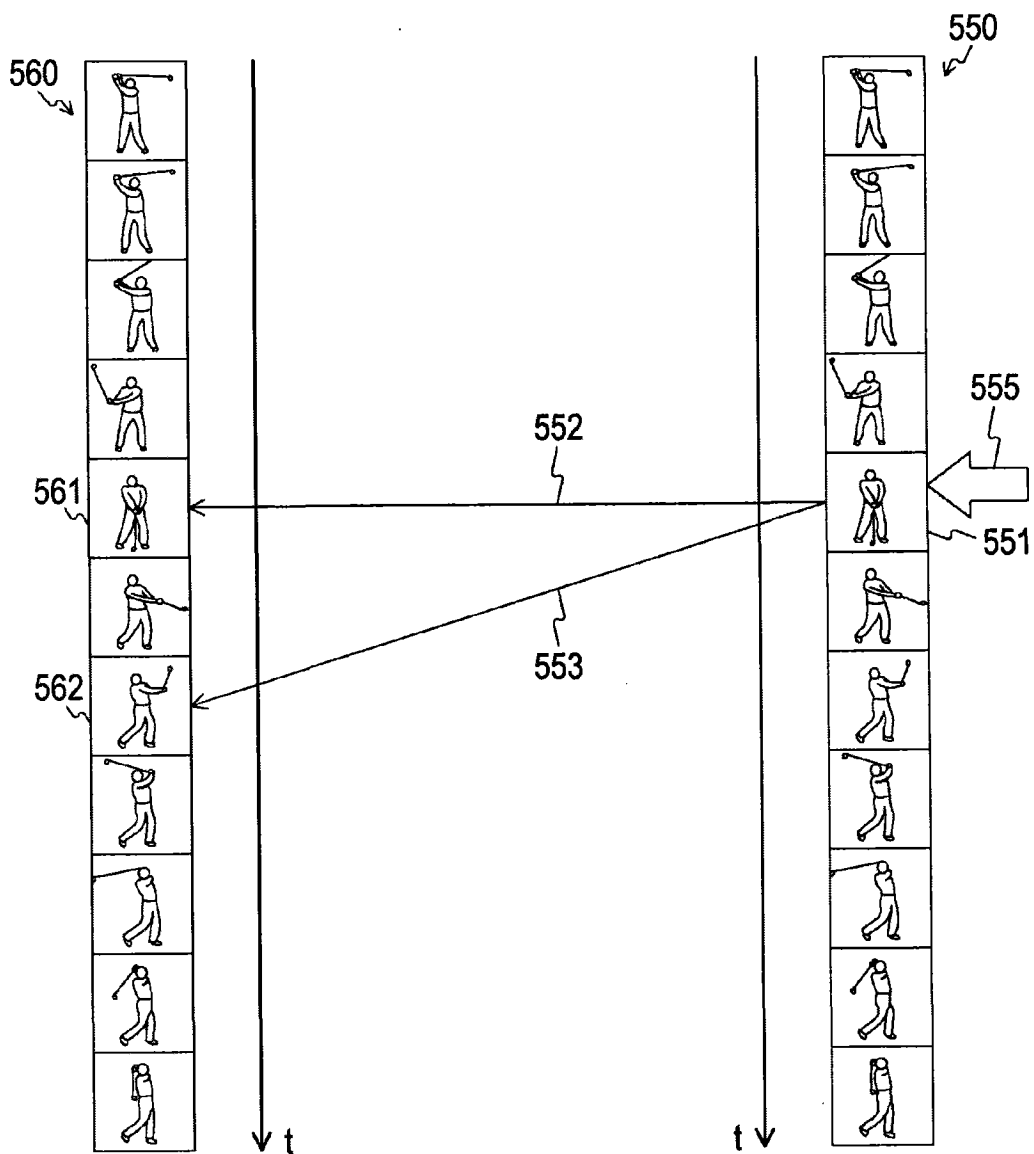


图 16B

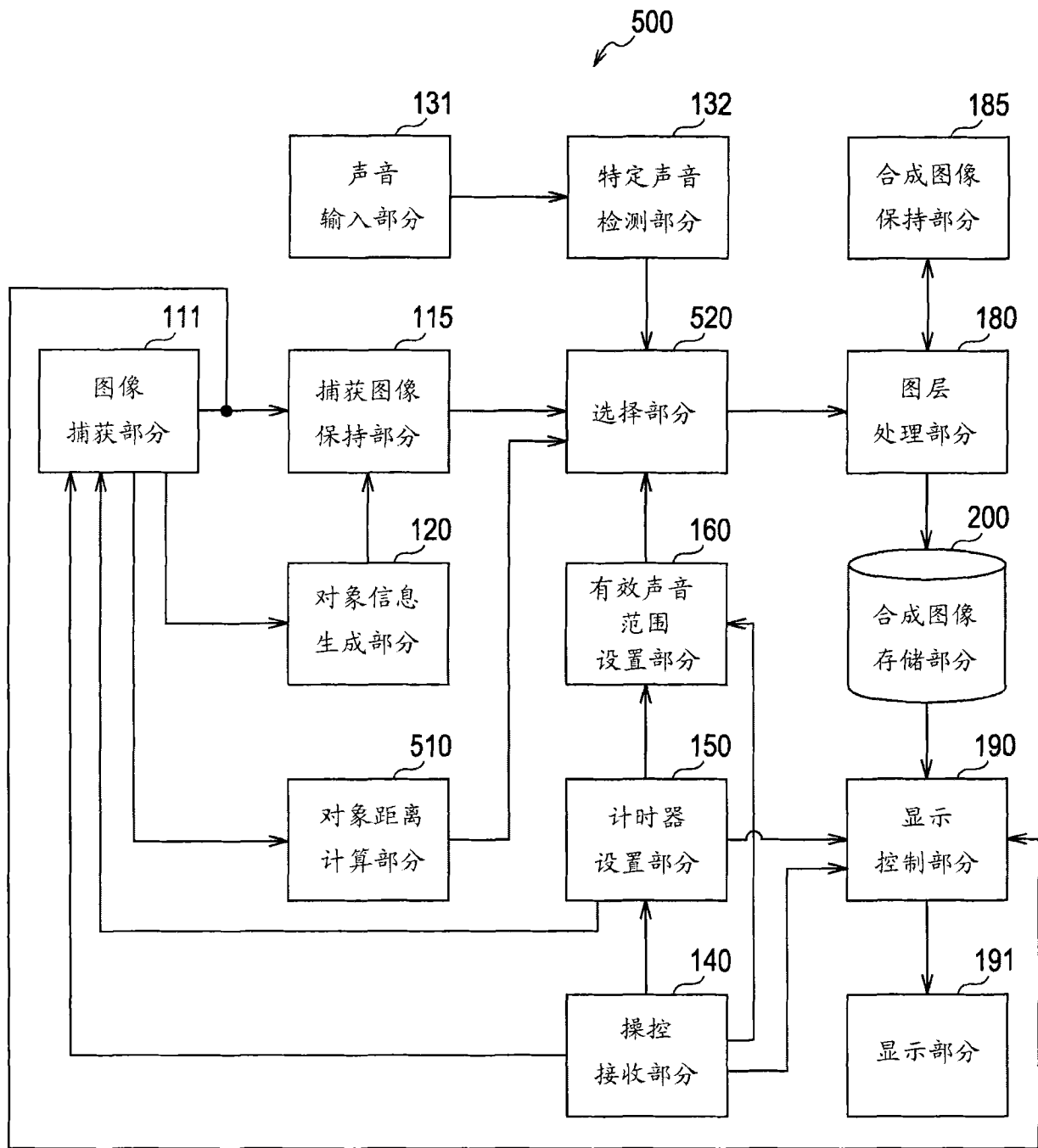


图 17

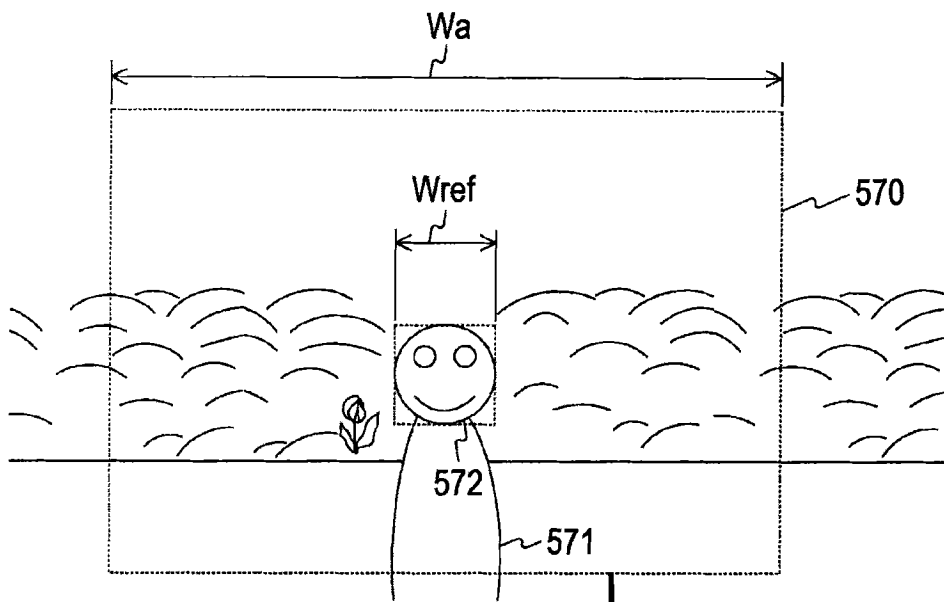


图 18A

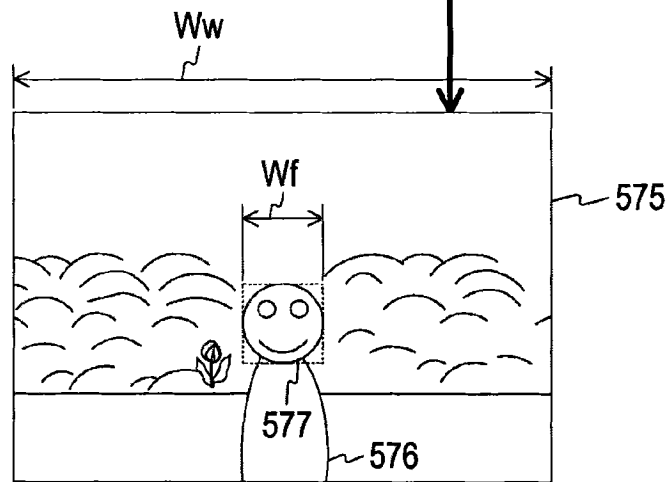


图 18B

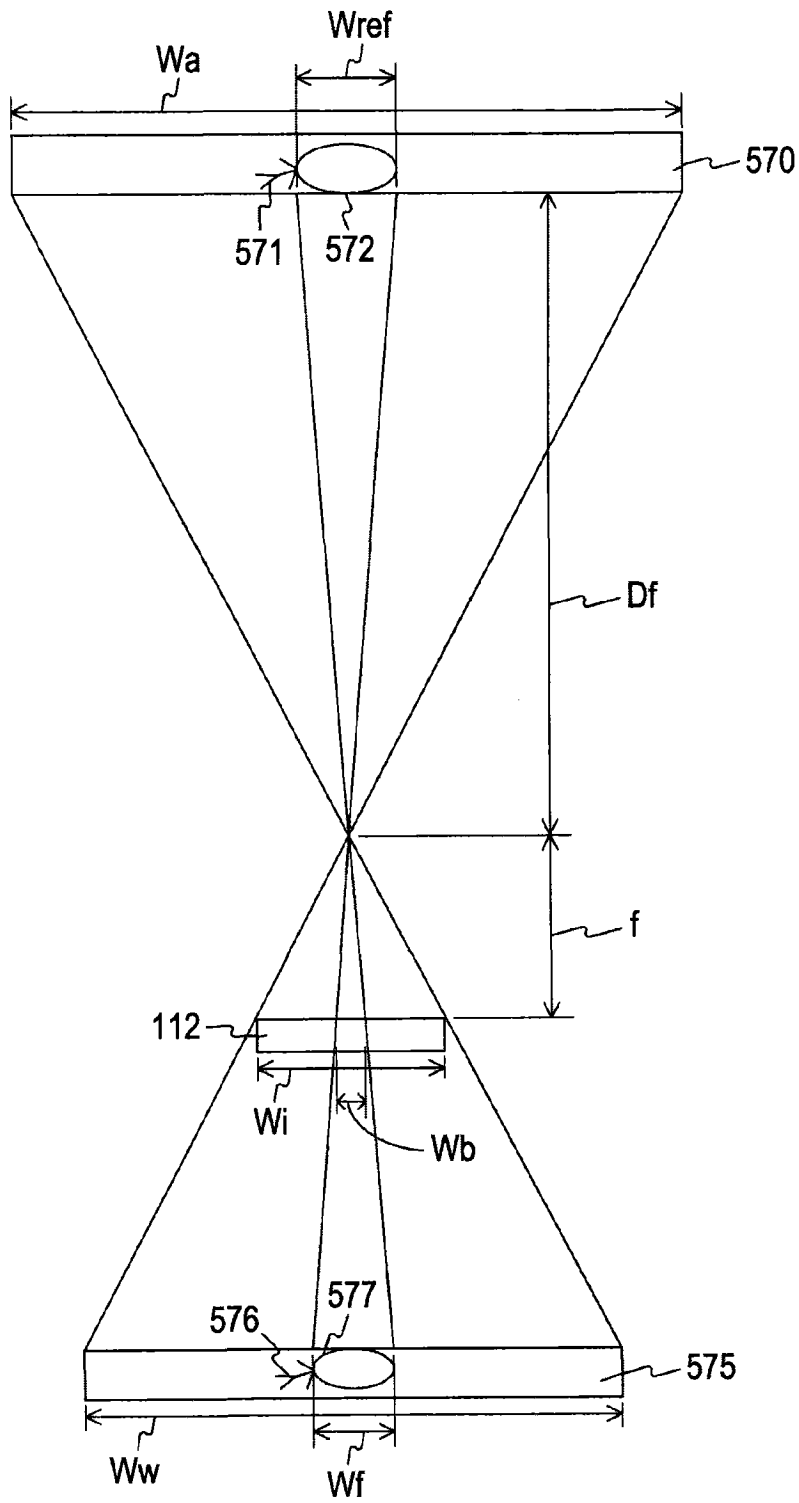


图 19

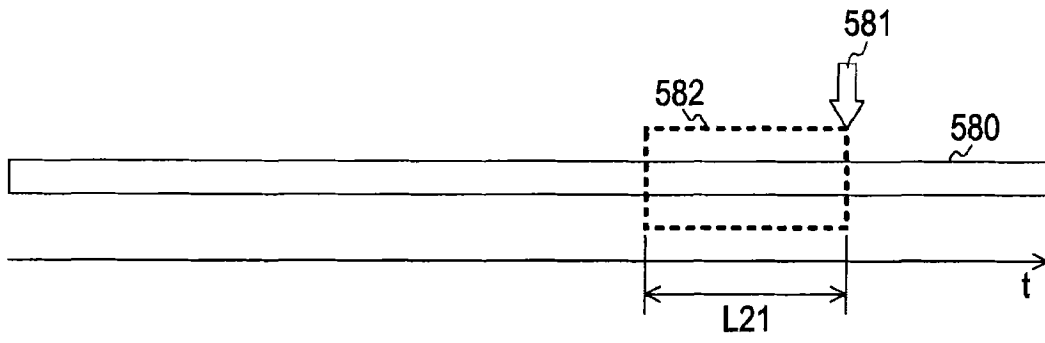


图 20A

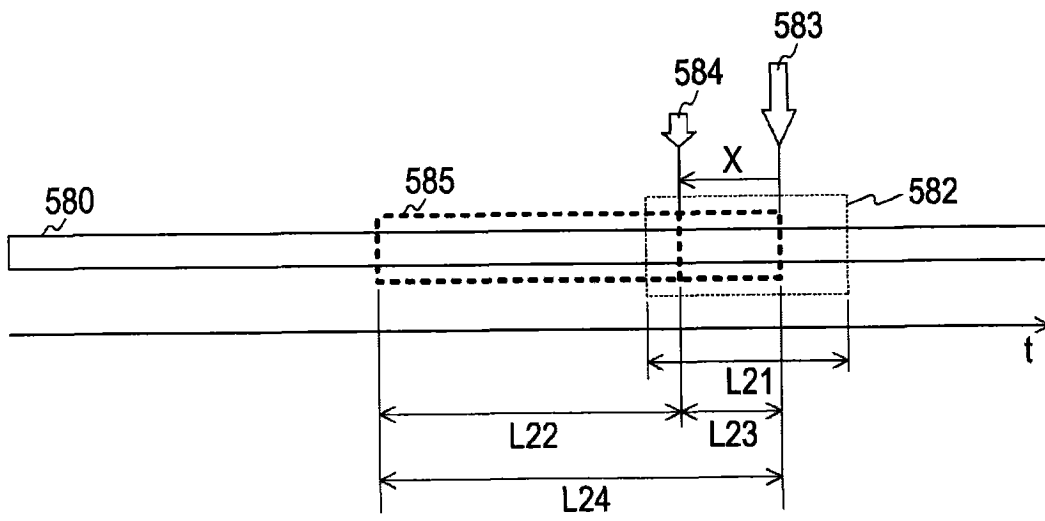


图 20B

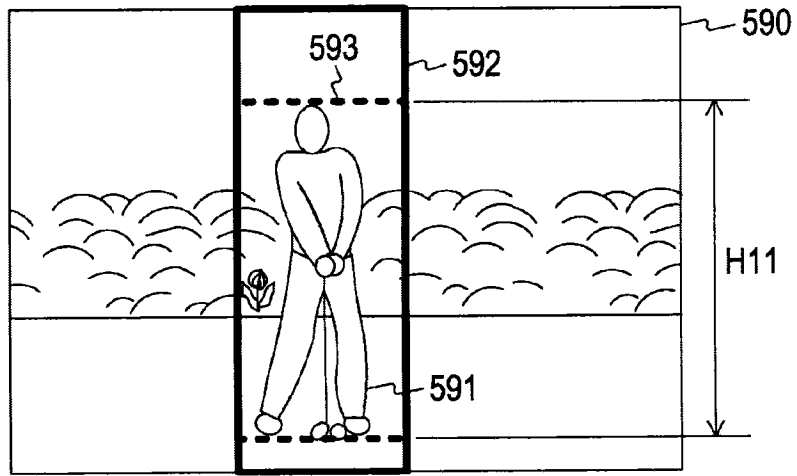


图 21

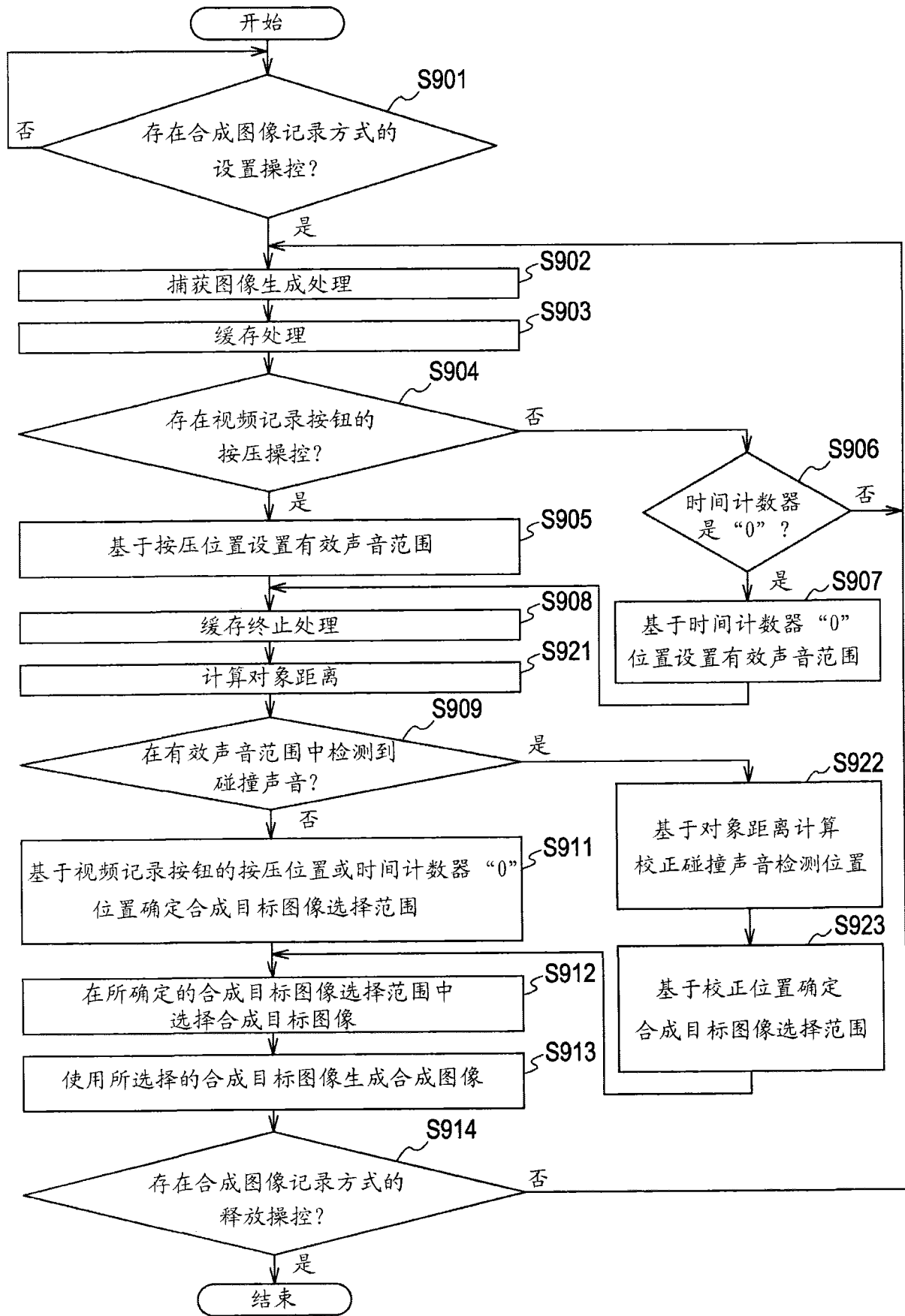


图 22

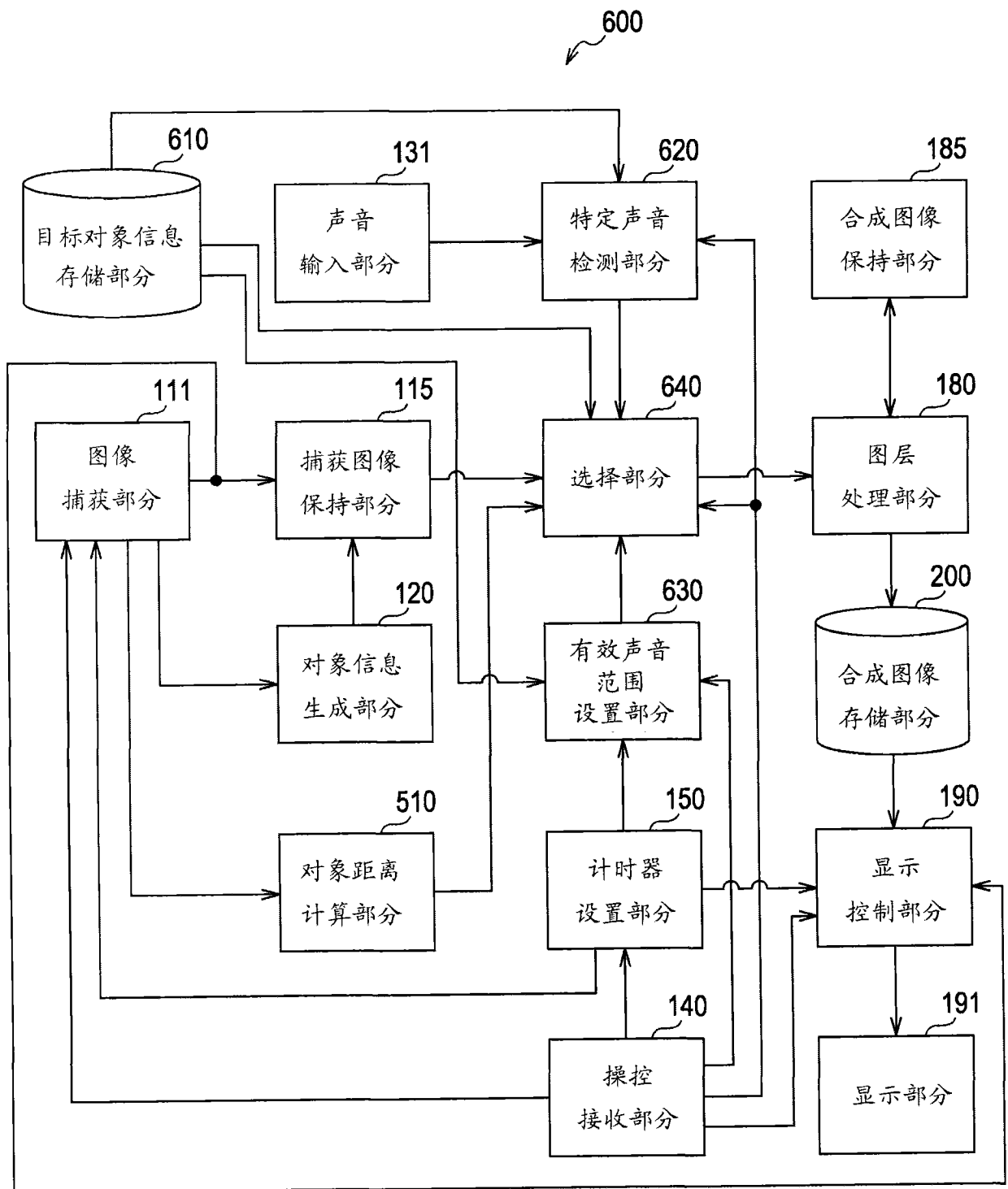


图 23

	611 碰撞声音	612 高速动作范围	613 有效声音范围	614 合成目标图像 选择范围
高尔夫球	通过挥杆期间击球生成的声音	碰撞声音周围 0.05 秒	拍摄触发之前的特定范围	碰撞声音之前 1.5 秒和以后 0.5 秒 (总共 2.0 秒)
棒球 (击球)	通过挥杆期间击球生成的声音	碰撞声音周围 0.05 秒	拍摄触发之前的特定范围	碰撞声音之前 1.5 秒和以后 0.5 秒 (总共 2.0 秒)
网球 (发球)	通过击球生成的声音	碰撞声音周围 0.05 秒	从拍摄触发到首次检测到特定值或更高的声音为止的范围	碰撞声音之前 2.5 秒和以后 0.5 秒 (总共 3.0 秒)
屋瓦打碎	通过用手打碎屋瓦所生成的声音	碰撞声音周围 0.1 秒	从拍摄触发到首次检测到特定值或更高的声音为止的范围	碰撞声音之前 1.5 秒和以后 0.5 秒 (总共 2.0 秒)
焰火	在天空中爆炸的瞬间的声音	碰撞声音周围 0.1 秒	拍摄触发之前的特定范围	碰撞声音之前 0.1 秒和以后 5.1 秒 (总共 5.2 秒)

图 24

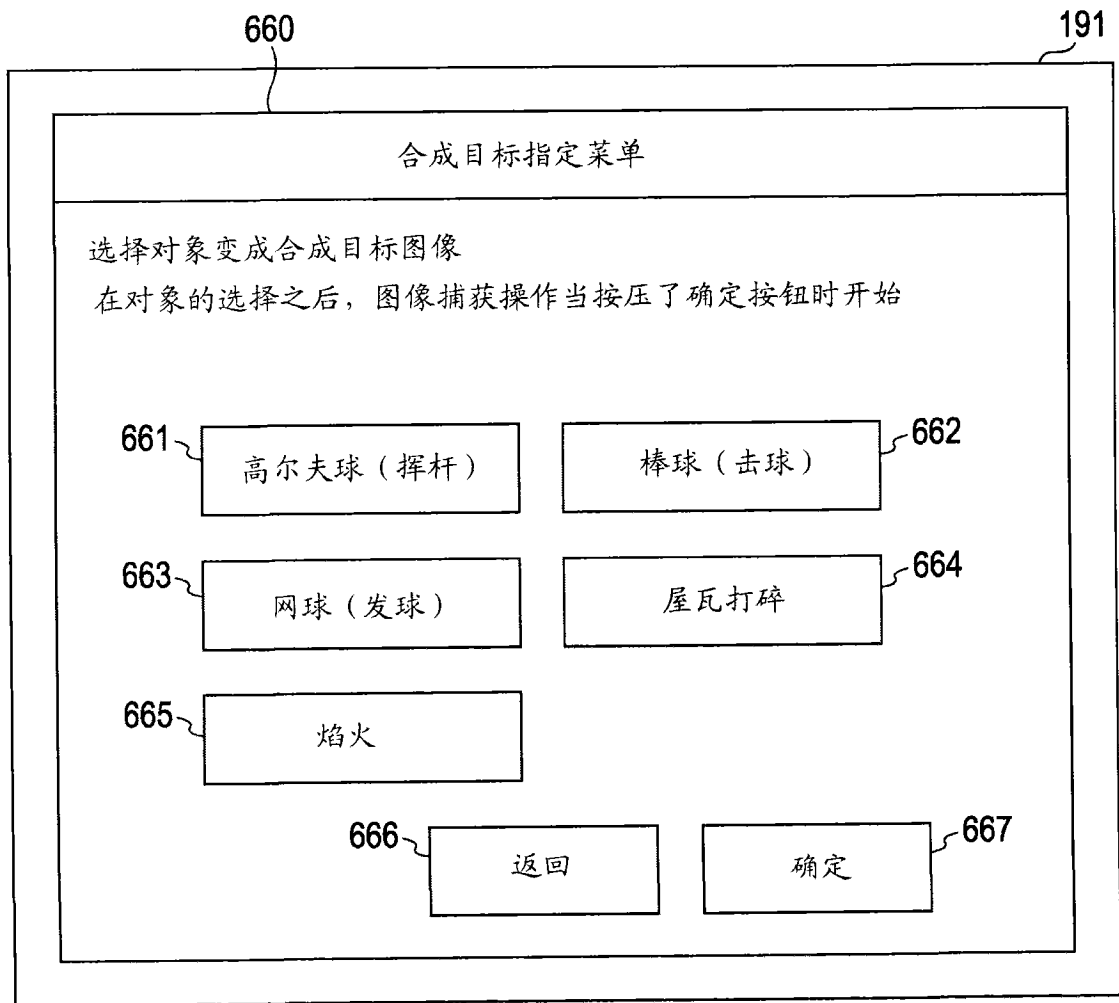


图 25

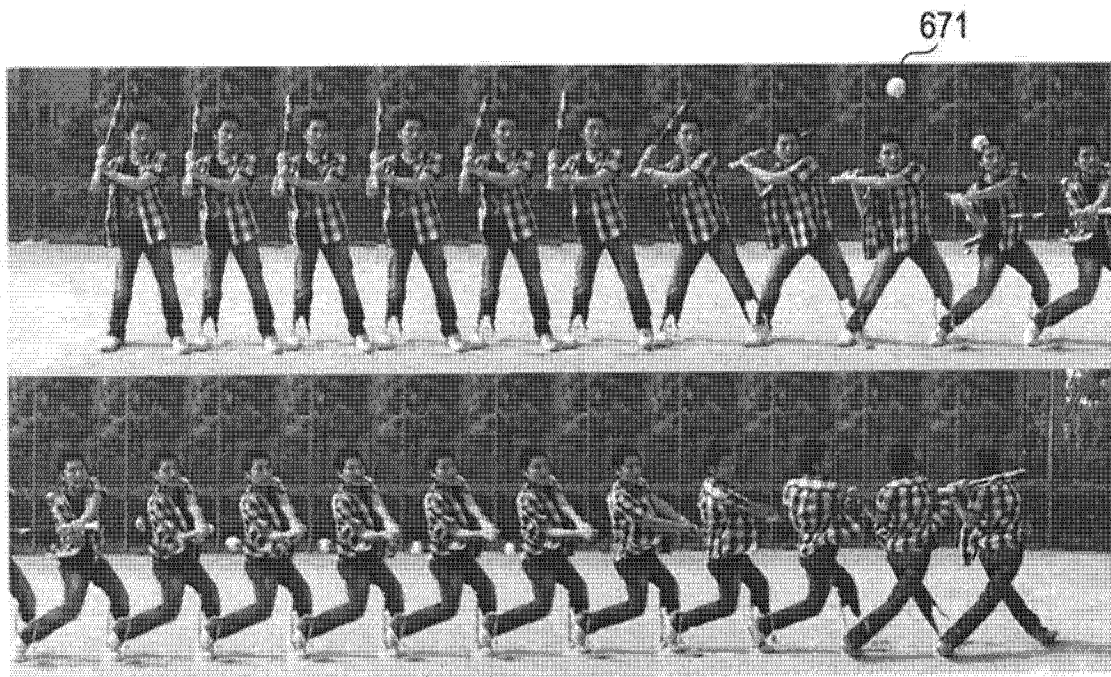


图 26A

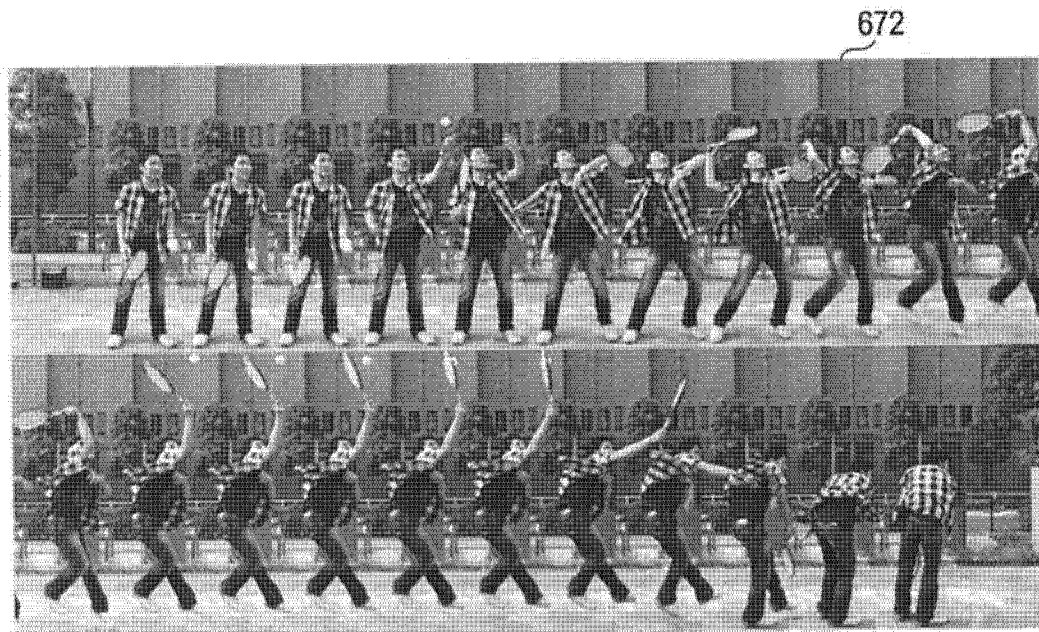


图 26B

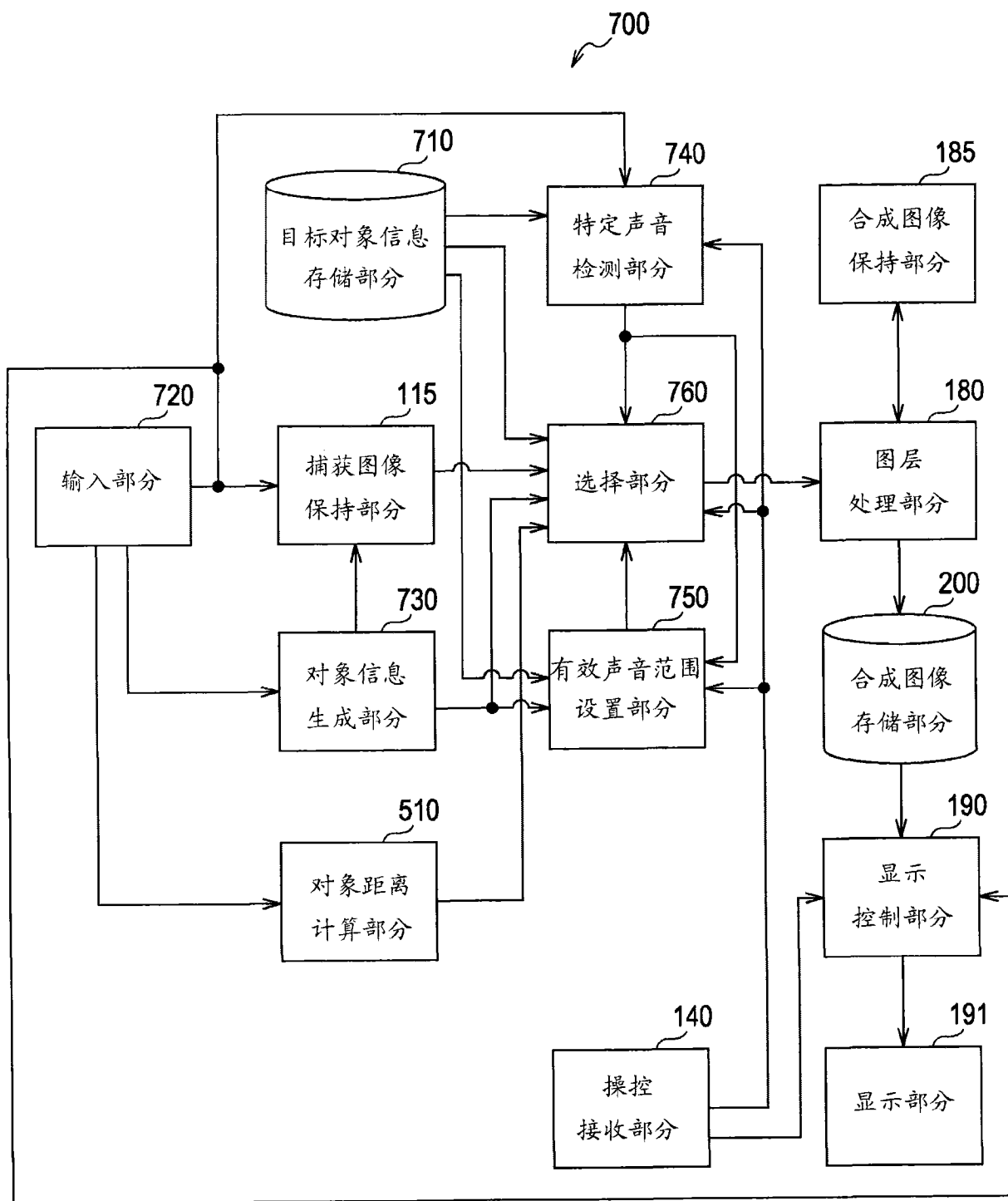


图 27

	710 碰撞声音	711 高速动作范围	712 用于确定有效声音的触发	713 有效声音范围	714 合成目标图像选择范围	715
高尔夫球	通过挥杆期间击球生成的声音	碰撞声音周围 0.05 秒	当检测到观众的高呼时	触发之前的特定范围	碰撞声音之前 1.5 秒和以后 0.5 秒 (总共 2.0 秒)	
棒球 (击球)	通过挥杆期间击球生成的声音	碰撞声音周围 0.05 秒	当检测到观众的高呼时	触发之前的特定范围	碰撞声音之前 1.5 秒和以后 0.5 秒 (总共 2.0 秒)	
网球 (发球)	通过击球生成的声音	碰撞声音周围 0.05 秒	当面孔上仰了预定时间时	从触发到首次检测到特定值或更高的声音为止的范围	碰撞声音之前 2.5 秒和以后 0.5 秒 (总共 3.0 秒)	
屋瓦打碎	通过用手打碎屋瓦所生成的声音	碰撞声音周围 0.1 秒	当面孔战斗精神饱满时	从触发到首次检测到特定值或更高的声音为止的范围	碰撞声音之前 1.5 秒和以后 0.5 秒 (总共 2.0 秒)	
焰火	在天空中爆炸的瞬间的声音	碰撞声音周围 0.1 秒	当屏幕中的黑暗状态被过渡到明亮状态时	触发之前的特定范围	碰撞声音之前 0.1 秒和以后 5.1 秒 (总共 5.2 秒)	

图 28

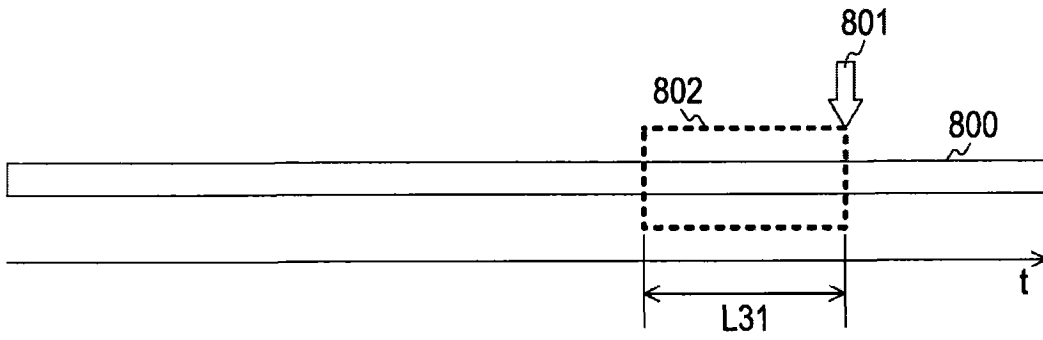


图 29A

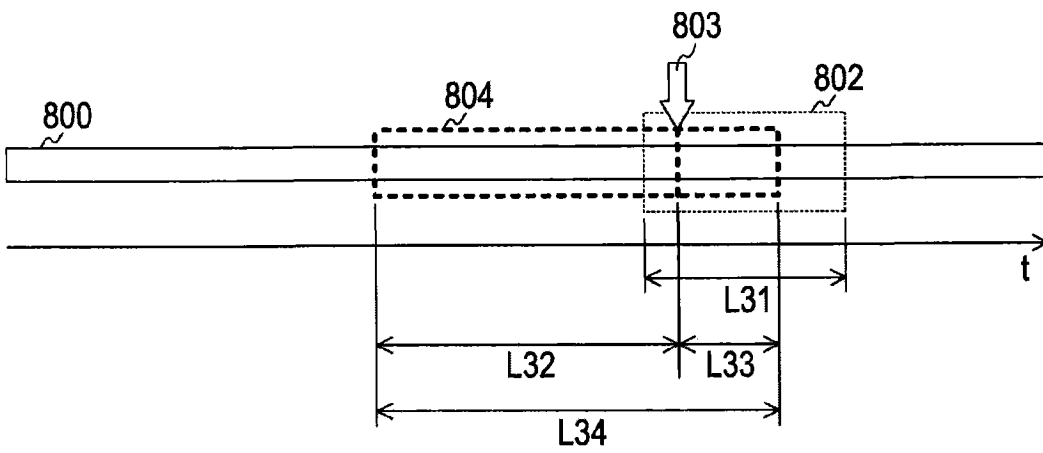


图 29B

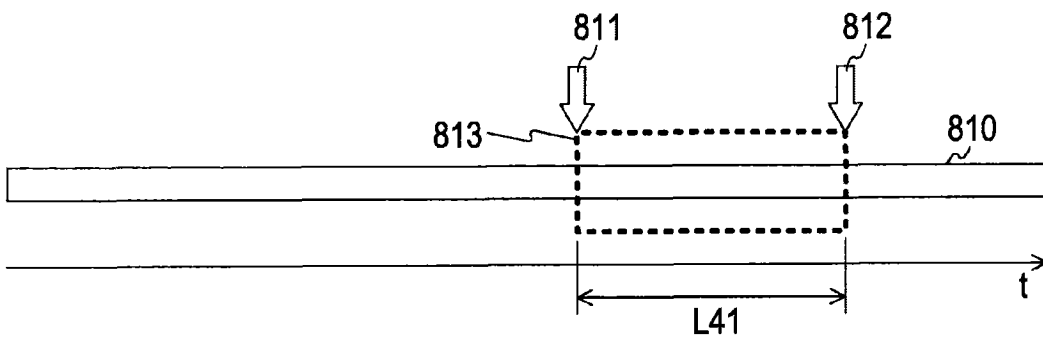


图 30A

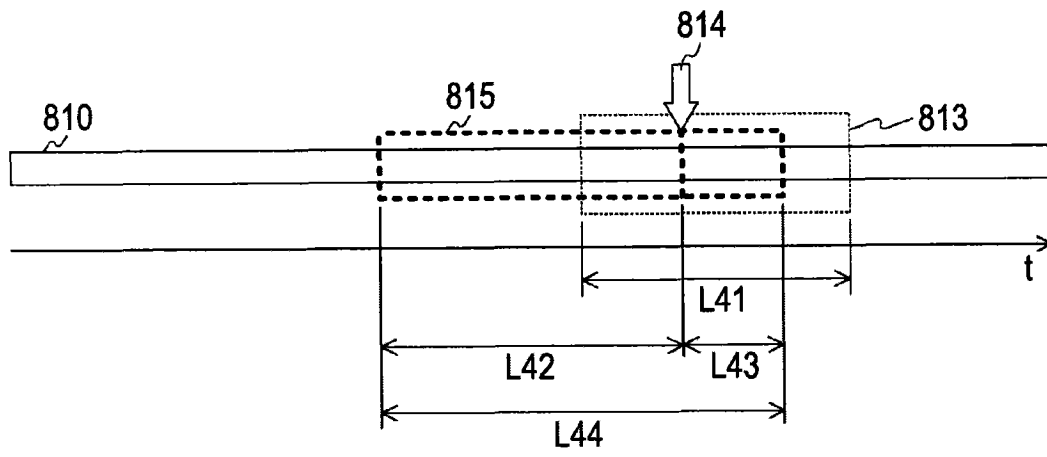


图 30B