



(12)发明专利

(10)授权公告号 CN 105474666 B

(45)授权公告日 2018.11.16

(21)申请号 201480044955.5

(22)申请日 2014.04.25

(65)同一申请的已公布的文献号

申请公布号 CN 105474666 A

(43)申请公布日 2016.04.06

(85)PCT国际申请进入国家阶段日

2016.02.06

(86)PCT国际申请的申请数据

PCT/JP2014/002334 2014.04.25

(87)PCT国际申请的公布数据

W02015/162645 JA 2015.10.29

(73)专利权人 松下知识产权经营株式会社

地址 日本大阪

(72)发明人 岸本伦典 渡边泰章 高桑诚

中村学 首藤秀树 田村健二

山崎龙次 齐藤宽夫 秋山明宽

(74)专利代理机构 中原信达知识产权代理有限

责任公司 11219

代理人 穆德骏 谢丽娜

(51)Int.Cl.

H04R 3/00(2006.01)

G10L 21/0272(2006.01)

G10L 25/51(2006.01)

H04N 7/18(2006.01)

H04R 1/40(2006.01)

(56)对比文件

JP 2013240000 A, 2013.11.28,

JP 2004109361 A, 2004.04.08,

JP 2003348569 A, 2003.12.05,

JP 2006339741 A, 2006.12.14,

JP 2007228459 A, 2007.09.06,

JP 2008219227 A, 2008.09.18,

JP 2004153754 A, 2004.05.27,

JP 2008271157 A, 2008.11.06,

JP 2008271157 A, 2008.11.06,

JP 2007300220 A, 2007.11.15,

JP 2006238110 A, 2006.09.07,

审查员 陈诗华

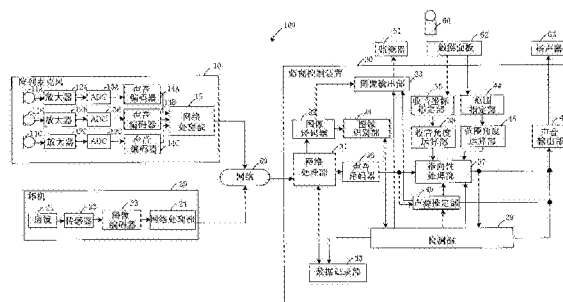
权利要求书3页 说明书23页 附图10页

(54)发明名称

声音处理系统及声音处理方法

(57)摘要

提供一种能够促进声音数据及图像数据的有效利用并提高便利性的声音处理装置。具备：数据取得部，取得由包括多个麦克风的收音部从预定区域内的声源收音到的声音数据及由拍摄上述预定区域内的至少一部分的图像的拍摄部拍摄到的图像数据；第一指定部，指定与基于上述图像数据而显示的图像上的指定部位对应的以上述收音部为基准的方向；第二指定部，指定与基于上述图像数据而显示的图像上的指定部位对应的上述预定区域内的预定范围；及指向性处理部，在所指定的上述预定范围内，增强所指定的上述方向的上述声音数据中的声音成分。



1. 一种声音处理方法,是声音处理装置的声音处理方法,所述声音处理方法具备:

数据取得步骤,取得由包括多个麦克风的收音部从预定区域内的声源收音到的声音数据及由拍摄所述预定区域内的至少一部分的图像的拍摄部拍摄到的图像数据;

第一指定步骤,指定与基于所述图像数据而显示的图像上的指定部位对应的以所述收音部为基准的所述收音部的指向性的方向;

第二指定步骤,指定与基于所述图像数据而显示的图像上的指定部位对应的所述预定区域内的所述收音部的指向性的范围的大小;

指向性处理步骤,在所指定的所述收音部的指向性的范围内,以增强所指定的所述方向的所述声音数据中的声音成分的方式,在所指定的所述方向上形成所指定的大小的范围的所述收音部的指向性并生成增强声音数据;

声音检测步骤,根据所指定的所述方向及所述指向性的范围内的增强声音数据检测监视对象音;及

处理步骤,在通过声音检测步骤检测到所述监视对象音的情况下,将检索用标志与所述声音数据及所述图像数据建立对应地记录于所述记录部,所述检索用标志用于从记录所述声音数据及所述图像数据的记录部检索包括所述监视对象音的声音数据或包括所述监视对象音的声源的图像数据,

所述检索用标志包括:以所述收音部为基准的所述监视对象音的声源的方向和包括以所述收音部为基准的所述监视对象音的声源的所述范围的大小中的至少一个信息。

2. 根据权利要求1所述的声音处理方法,其中,

所述处理步骤取得与记录于所述记录部的检索用标志所包含的预定检索用标志对应的记录于所述记录部的声音数据或图像数据。

3. 根据权利要求1所述的声音处理方法,其中,

所述检索用标志还包括所述监视对象音的类别及通过所述声音检测步骤检测到所述监视对象音的时刻中的至少一个信息。

4. 根据权利要求1~3中任一项所述的声音处理方法,其中,

在检测到所述监视对象音的情况下,所述处理步骤使提示部提示包括检测到所述监视对象音的内容的警告信息。

5. 根据权利要求1~3中任一项所述的声音处理方法,其中,

在检测到所述监视对象音的情况下,所述处理步骤将包括所述监视对象音的声音数据记录于记录部。

6. 根据权利要求1~3中任一项所述的声音处理方法,其中,

在检测到所述监视对象音的情况下,所述处理步骤对通过所述指向性处理步骤增强声音成分的方向及所述范围的大小中的至少一方进行变更。

7. 根据权利要求1~3中任一项所述的声音处理方法,其中,

所述声音处理方法还具备推定步骤,推定发出所述监视对象音的声源的位置,并使提示部提示推定出的位置的信息。

8. 根据权利要求1~3中任一项所述的声音处理方法,其中,

所述声音处理方法还具备推定发出所述监视对象音的声源的位置的推定步骤,

所述指向性处理步骤增强从由所述推定步骤推定出的所述声源的位置的方向到来的

声音成分。

9. 根据权利要求1~3中任一项所述的声音处理方法,其中,

在由所述指向性处理步骤增强后的声音成分的信号等级为第一预定信号等级以上或第二预定信号等级以下的情况下,所述声音检测步骤检测该声音成分作为所述监视对象音。

10. 根据权利要求1~3中任一项所述的声音处理方法,其中,

所述声音检测步骤根据由所述指向性处理步骤增强后的声音成分中的至少一个声音成分检测预定关键字作为所述监视对象音。

11. 根据权利要求10所述的声音处理方法,其中,

所述处理步骤对包括检测到的所述预定关键字的声音数据中的所述预定关键字的部分进行加工。

12. 根据权利要求10所述的声音处理方法,其中,

所述处理步骤将包括检测到的所述预定关键字的声音数据记录于记录部。

13. 根据权利要求1~3中任一项所述的声音处理方法,其中,

所述声音检测步骤检测由所述指向性处理步骤增强后的声音成分中的至少一个声音成分所包含的预定异常音作为所述监视对象音。

14. 根据权利要求1所述的声音处理方法,其中,

所述声音处理方法具备对所述图像数据进行图像识别的图像识别步骤,

所述处理步骤根据所述图像识别步骤的图像识别结果来进行记录所述检索用标志的处理。

15. 根据权利要求14所述的声音处理方法,其中,

所述图像识别步骤在所述图像数据中识别所述监视对象音的声源的类别。

16. 根据权利要求14或15所述的声音处理方法,其中,

所述图像识别步骤在所述图像数据中识别所述监视对象音的声源的移动的有无。

17. 根据权利要求14所述的声音处理方法,其中,

在通过所述图像识别步骤对所述图像数据进行了图像识别的情况下,所述处理步骤将检索用标志记录于记录部。

18. 根据权利要求17所述的声音处理方法,其中,

所述处理步骤使用记录于所述记录部的检索用标志所包含的预定检索用标志,与所述预定检索用标志对应地取得记录于所述记录部的声音数据或图像数据。

19. 根据权利要求17或18所述的声音处理方法,其中,

所述检索用标志包括所述声源的类别、所述声源的移动的有无及包括所述声源的缩略图像中的至少一个。

20. 根据权利要求14所述的声音处理方法,其中,

在检测到所述监视对象音的情况下,所述处理步骤根据所述图像识别步骤的图像识别结果,使提示部提示包括检测到所述监视对象音的内容的警告信息。

21. 根据权利要求14所述的声音处理方法,其中,

在检测到所述监视对象音的情况下,所述处理步骤根据所述图像识别步骤的图像识别结果,将包括所述监视对象音的声音数据记录于记录部。

22. 根据权利要求14所述的声​​音处理方法, 其中,

在检测到所述监视对象音的情况下, 所述处理步骤根据所述图像识别步骤的图像识别结果, 对通过所述指向性处理步骤增强声音成分的方向及范围中的至少一方进行变更。

23. 一种声音处理系统, 具备:

收音装置, 包括使用多个麦克风从预定区域内的声源对声音数据进行收音的收音部;

拍摄装置, 包括拍摄所述预定区域内的至少一部分的图像的拍摄部; 及

声音处理装置, 对由所述收音部收音到的声音数据进行处理,

所述声音处理装置具备:

数据取得部, 取得由所述收音部收音到的声音数据及由所述拍摄部拍摄到的图像数据;

第一指定部, 指定与基于所述图像数据而显示的图像上的指定部位对应的以所述收音部为基准的所述收音部的指向性的方向;

第二指定部, 指定与基于所述图像数据而显示的图像上的指定部位对应的所述预定区域内的所述收音部的指向性的范围的大小; 及

指向性处理部, 在所指定的所述收音部的指向性的范围内, 以增强所指定的所述方向的声音数据中的声音成分的方式, 在所指定的所述方向上形成所指定的的大小的范围的所述收音部的指向性并生成增强声音数据;

声音检测部, 根据所指定的所述方向及所述指向性的范围内的增强声音数据来检测监视对象音; 及

处理部, 在通过声音检测部检测到所述监视对象音的情况下, 将检索用标志与所述声音数据及所述图像数据建立对应地记录于所述记录部, 所述检索用标志用于从记录所述声音数据及所述图像数据的记录部检索包括所述监视对象音的声音数据或包括所述监视对象音的声源的图像数据,

所述检索用标志包括: 以所述收音部为基准的所述监视对象音的声源的方向和包括以所述收音部为基准的所述监视对象音的声源的所述范围的大小中的至少一个信息。

声音处理系统及声音处理方法

技术领域

[0001] 本发明涉及声音处理装置、声音处理系统及声音处理方法。

背景技术

[0002] 以往,例如,在特定的室内或远距离地点对工场、店铺、公共的场所的状况进行监视的情况下,利用监视系统。监视系统具备例如拍摄图像的相机、对声音进行收音的麦克风、存储预定数据(例如,拍摄到的图像、收音到的声音)的记录器装置。通过使用监视系统,例如在发生事件或事故的情况下,对记录器装置记录的过去的数据进行重放,记录的图像或声音能够有助于过去的时刻的状况掌握。

[0003] 作为以往的监视系统,已知有全方位相机及麦克风阵列用的系统。该系统利用由多个麦克风形成的阵列麦克风,通过滤波仅提取来自特定的方向的声音,而形成波束(例如,参照专利文献1)。

[0004] 专利文献1:日本国特开2004-32782号公报

发明内容

[0005] 使用阵列麦克风收音到的声音数据可能包括各种有益的信息。在专利文献1的监视系统中,声音数据及图像数据的有效利用并不充分,期待提高利用监视系统的利用者的便利性。

[0006] 本发明鉴于上述情况而作出,提供能够促进声音数据及图像数据的有效利用并提高便利性的声音处理装置、声音处理系统及声音处理方法。

[0007] 本发明的一方式的声音处理装置具备:数据取得部,取得由包括多个麦克风的收音部从预定区域内的声源收音到的声音数据及由拍摄所述预定区域内的至少一部分的图像的拍摄部拍摄到的图像数据;第一指定部,指定与基于上述图像数据而显示的图像上的指定部位对应的以上述收音部为基准的方向;第二指定部,指定与基于上述图像数据而显示的图像上的指定部位对应的上述预定区域内的预定范围;及指向性处理部,在所指定的上述预定范围内,增强所指定的上述方向的上述声音数据中的声音成分。

[0008] 而且,本发明的另一方式的声音处理系统具备:收音装置,包括使用多个麦克风从预定区域内的声源对声音数据进行收音的收音部;拍摄装置,包括拍摄上述预定区域内的至少一部分的图像的拍摄部;及声音处理装置,对由上述收音部收音到的声音数据进行处理,上述声音处理装置具备:数据取得部,取得由上述收音部收音到的声音数据及由上述拍摄部拍摄到的图像数据;第一指定部,指定与基于上述图像数据而显示的图像上的指定部位对应的以上述收音部为基准的方向;第二指定部,指定与基于上述图像数据而显示的图像上的指定部位对应的上述预定区域内的预定范围;及指向性处理部,在所指定的上述预定范围内,增强所指定的上述方向的声音数据中的声音成分。

[0009] 而且,本发明的又一方式的声音处理方法是声音处理装置的声音处理方法,该声音处理方法包括以下步骤:取得由包括多个麦克风的收音部从预定区域内的声源收音到的

声音数据及由拍摄上述预定区域内的至少一部分的图像的拍摄部拍摄到的图像数据;指定与基于上述图像数据而显示的图像上的指定部位对应的以上述收音部为基准的方向;指定与基于上述图像数据而显示的图像上的指定部位对应的上述预定区域内的预定范围;及在所指定的上述预定范围内,增强所指定的上述方向的上述声音数据中的声音成分。

[0010] 发明效果

[0011] 根据本发明,能够促进声音数据及图像数据的有效利用,提高便利性。

附图说明

[0012] 图1是第一实施方式的监视系统的概要图。

[0013] 图2是表示第一实施方式的监视系统的构成例的框图。

[0014] 图3是表示第一实施方式的阵列麦克风、相机及各声源的配置状态的一例的俯视图。

[0015] 图4是表示第一实施方式的监视控制装置的动作例的流程图。

[0016] 图5是第一实施方式的指向性处理的概要图。

[0017] 图6是第二实施方式的监视系统的概要图。

[0018] 图7是表示第二实施方式的监视系统的构成例的框图。

[0019] 图8是表示第二实施方式的监视控制装置的动作例的流程图。

[0020] 图9(A)、图9(B)是表示使用了第一实施方式的PC的收音范围的大小的变更例的示意图,图9(C)是表示使用了第一实施方式的平板终端的收音范围的大小的一例的示意图。

[0021] 图10(A)、图10(B)是表示第一实施方式的收音范围及收音角度的一例的示意图。

具体实施方式

[0022] 关于本发明的实施方式,以下使用附图进行说明。

[0023] (第一实施方式)

[0024] 图1是第一实施方式的监视系统100的概要图。在监视系统100中,阵列麦克风10(Array microphones)、相机20及监视控制装置30经由有线或无线的网络50而连接。

[0025] 阵列麦克风10是收音部、收音装置的一例。相机20是拍摄部、拍摄装置的一例。监视控制装置30是声音处理装置的一例。监视系统100是声音处理系统的一例。

[0026] 阵列麦克风10包括多个麦克风11(11A、11B、11C、•••),对阵列麦克风10的周围的声音进行收音,即从预定区域内的声源进行收音,得到声音数据。相机20拍摄能够由相机20拍摄的上述预定区域内的至少一部分,得到图像数据。图像数据包括例如活动图像或静止图像。监视控制装置30根据阵列麦克风10的收音结果、相机20的拍摄结果,进行与监视相关的各种处理。

[0027] 在监视系统100中,在单元框体91内一体地装入1个相机20和阵列麦克风10所包含的16个麦克风11(11A、11B、11C、•••),从而形成收音单元90。另外,阵列麦克风10中的麦克风的个数可以为15个以下,也可以为17个以上。而且,阵列麦克风10与相机20也可以不形成收音单元90而分体形成。

[0028] 相机20例如将拍摄方向的中心(光轴方向)设为铅垂下方向地配置在单元框体91的大致中央部。阵列麦克风10中的多个麦克风11以沿着单元框体91的设置面将相机20的周

围包围的方式以一定的间隔配置在1个圆周上。多个麦克风11也可以不配置在圆周上而配置在例如矩形上。另外,这样的相机20及多个麦克风11的配置关系、配置形状是一例,也可以是其他的配置关系、配置形状。

[0029] 相机20构成为例如能够同时拍摄大范围(例如全方位)的被摄体。各麦克风11构成为例如能够检测从大范围(例如全方位)到来的声波。

[0030] 图2是表示监视系统100的构成例的框图。

[0031] 监视系统100具备阵列麦克风10、相机20及监视控制装置30。阵列麦克风10、相机20及监视控制装置30经由网络50连接成相互能够进行数据通信的状态。而且,在监视控制装置30上连接例如监视器61、触摸面板62及扬声器63。

[0032] 图2的结构可设想例如在实时地监视图像及声音的情况下监视控制装置30记录图像数据及声音数据的情况。另外,也可以是,相机20记录图像数据,阵列麦克风10记录声音数据,在记录后能够参照图像数据及声音数据。

[0033] 另外,以下,主要以阵列麦克风10所包含的多个麦克风11中的3个麦克风11A、11B、11C为代表进行说明。关于3个麦克风11A~11C以外的麦克风,也具有与麦克风11A~11C同样的结构及功能。

[0034] 阵列麦克风10将多个麦克风11A、11B、11C以相互接近的状态规则地(例如在圆周上)排列、形成。麦克风11A~11C是将声音转换成电信号(声音数据)的转换器。在阵列麦克风10中,多个麦克风11A、11B、11C也可以不是规则地配置。在这种情况下,例如,可以将各麦克风11A~11C的位置的信息保持于监视系统100,来进行指向性处理。

[0035] 在麦克风11A~11C的输出上连接有放大器12A~12C、A/D转换器(ADC: Analog to Digital Converter)13A~13C及声音编码器14A~14C。而且,在声音编码器14A~14C的输出上连接有网络处理部15。

[0036] 麦克风11A~11C生成与从各种方向输入的声响的振动对应的声音数据。该声音数据是模拟声音数据。放大器12A~12C对麦克风11A~11C输出的声音数据进行放大。A/D转换器(ADC)13A~13C对放大器12A~12C输出的声音数据周期性地采样,将声音数据转换成数字数据。声音编码器14A~14C对A/D转换器13A~13C输出的声音数据(声音数据的波形的时序变化)进行编码,生成适合于传送的预定形式的声音数据。

[0037] 另外,本实施方式中的“声音”除了包括由于人类的发声而得到的声音之外,还可以包括例如由于机械性的振动而产生的一般的声响或噪音的成分。而且,“声音”可以包括例如监视对象的声音以外的声音。即,由麦克风11A~11C收音到的声响的信号有时不区分声响的种类地记载为“声音”。

[0038] 网络处理部15取得声音编码器14A~14C生成的声音数据,并将声音数据向网络50送出。例如,相对于麦克风11A~11C收集到的声音,声音编码器14A~14C生成独立的聲音数据。由此,网络处理部15将与各麦克风11A~11C对应的多个声道的声音数据向网络50送出。

[0039] 相机20具备透镜21、传感器22、图像编码器23及网络处理部24。

[0040] 透镜21为例如全方位透镜、鱼眼透镜。传感器22是拍摄设备,包括例如CCD(Charge Coupled Device)影像传感器、CMOS(Complementary Metal Oxide Semiconductor)影像传感器。传感器22根据经由透镜21向传感器22的拍摄面入射的被摄体的光像来生成图像数据。

[0041] 图像编码器23对传感器22输出的图像数据依次进行处理,生成适合于预定规格的图像数据。网络处理部24将图像编码器23生成的图像数据向网络50送出。

[0042] 监视控制装置30为例如PC(Personal Computer)30 α (参照图9(A)、图9(B))、平板终端30 β (参照图9(C))。监视控制装置30包括例如CPU(Central Processing Unit)或DSP(Digital Signal Processor)。监视控制装置30包括例如ROM(Read Only Memory)或RAM(Random Access Memory)。

[0043] 监视控制装置30例如通过CPU或DSP执行在ROM或RAM中记录的控制用的程序(例如,应用程序、ActiveX形式的程序),由此实现各种功能。而且,ROM或RAM形成未图示的存储器。

[0044] 监视控制装置30具备网络处理部31、图像译码器32、图像输出部33、图像识别部34、收音坐标指定部35、范围指定部44、声音译码器36及指向性处理部37。而且,监视控制装置30具备收音角度运算部38、范围角度运算部45、检测部39、声源推定部40、声音输出部42及数据记录部43。

[0045] 网络处理部31经由网络50,与阵列麦克风10及相机20之间进行数据通信。通过数据通信,网络处理部31从阵列麦克风10取得多个声道的声音数据,从相机20取得图像数据。网络处理部31是数据取得部的一例。

[0046] 网络处理部31可以从阵列麦克风10及相机20直接取得阵列麦克风10送出的声音数据和相机20送出的图像数据。网络处理部31也可以在任意的时刻从数据记录部43读出并取得数据记录部43所记录的声音数据或图像数据(至少声音数据)。网络处理部31也可以将从阵列麦克风10及相机20直接取得的声音数据或图像数据在任意的时刻记录于数据记录部43。

[0047] 图像译码器32对来自网络处理部31的图像数据进行译码,生成能够重放的图像数据。

[0048] 图像输出部33将来自图像译码器32的图像数据转换成监视器61能够显示的形式的数据,向监视器61送出。而且,图像输出部33也可以控制监视器61的显示。而且,图像输出部33可以将与来自检测部39的检测信息对应的图像数据向监视器61送出。

[0049] 监视器61显示各种图像数据。监视器61例如根据来自图像输出部33的图像数据来显示图像。例如,相机20拍摄到的图像显示于监视器61。监视器61是提示部的一例。

[0050] 图像识别部34可以对于来自图像输出部33的图像数据执行预定图像处理,并识别与例如未图示的存储器中事先登记的各种图案的图像是否一致。例如,执行图案匹配的处理,从图像所包含的各种物体之中,提取与预定人物或预定人物的脸部类似的图案。也可以提取人物以外的物体的图案。

[0051] 而且,图像识别部34例如也可以识别图像数据所包含的物体的种类(例如,人物、男性、女性)。而且,图像识别部34可以具有VMD(Video Motion Detector)功能,检测图像数据内的移动的有无。

[0052] 收音坐标指定部35例如从触摸面板62或图像识别部34接受输入,并导出与输入位置或输入范围的基准位置(例如中心位置)对应的坐标。例如,在监视器61的画面上显示的图像中,收音坐标指定部35接受操作者60应关注的位置(例如图1的附图标记P1)的坐标作为收音坐标(x,y)。收音坐标指定部35是第一指定部的一例,指定与基于图像数据而显示的

图像上的指定部位(例如收音坐标)对应的以收音部(例如阵列麦克风10)为基准的方向。

[0053] 操作者60例如边观察监视器61边操作触摸面板62。由于伴随着触摸面板62上的移动操作(例如拖曳操作)而画面上显示的指示器(未图示)的位置进行移动,因此操作者60能够在画面上的显示范围内变更收音坐标。例如,通过操作者60对触摸面板62的触摸操作,指示器的坐标作为收音坐标而向收音坐标指定部35提供。操作者60是使用监视系统100进行监视的监视者的一例。

[0054] 范围指定部44例如从触摸面板62接受输入,导出与输入范围对应的坐标,导出输入范围的大小。例如,范围指定部44在监视器61的画面上显示的图像中,接受操作者60应关注的范围的坐标作为收音范围A1,导出收音范围A1的大小。范围指定部44是第二指定部的一例,指定与基于图像数据而显示的图像上的指定部位(例如收音范围A1)对应的预定区域内的预定范围。

[0055] 范围指定部44例如可以在触摸面板62的画面上,通过使用多个手指同时地指定收音范围A1的操作,来指定收音范围A1,从而指定(导出)收音范围A1的大小。例如,可以通过在触摸面板62的画面上实施缩小操作(例如参照图9(A))或放大操作(例如参照图9(B)),来指定收音范围A1,从而导出收音范围A1的大小。画面所包含的范围是预定区域的一例,收音范围A1是预定范围的一例。可以将指定收音范围A1的大小的情况简称为指定收音范围A1。缩小操作是对收音范围A1的大小进行缩小的操作的一例,放大操作是对收音范围A1的大小进行放大的操作的一例。

[0056] 范围指定部44例如可以指定以2根手指的触摸位置为直径的圆或椭圆的形状的范围作为收音范围A1,从而导出收音范围A1的大小。范围指定部44例如指定以2根手指的触摸位置为对角的点的多边形的形状的范围作为收音范围A1,并导出收音范围A1的大小。范围指定部44例如可以指定3根以上的手指触摸位置作为收音范围A1的轮廓点,也可以指定复杂的形状的范围并导出收音范围A1的大小。设为何种形状收音范围A1例如可以预先确定并存储于未图示的存储器。而且,范围指定部44例如也可以在触摸面板62的画面上,使用1根手指以时间差来指定多个位置,导出通过这多个位置的收音范围A1,并导出收音范围A1的大小。

[0057] 在收音坐标与收音范围A1的关系预先确定的情况下,可以根据预先确定的关系来指定收音范围A1或收音坐标。例如,在指定了收音坐标的情况下,可以指定以收音坐标为中心的预定形状收音范围A1,并导出收音范围A1的大小。例如,在收音范围A1指定了形成轮廓的多个点的情况下,可以导出收音范围A1的大小,并指定收音范围A1的中心位置的坐标作为收音坐标。

[0058] 也可以使用触摸面板62以外的输入单元来指定收音坐标及收音范围A1的大小。例如,也可以在监视控制装置上连接鼠标,而操作者60使用鼠标来触摸所希望的图像范围。例如,还可以通过手指以外的输入单元(例如指示笔),对触摸面板62进行触摸,来指定收音坐标及收音范围A1的大小。

[0059] 而且,图像识别部34在识别到事先登记的图案包括于图像数据的情况下,可以将识别到的图案所存在的监视器61上的位置(例如,图1的附图标记P1)的坐标作为收音坐标向收音坐标指定部35赋予。识别到的图案例如是人物的整体、人物的脸部。

[0060] 而且,图像识别部34在识别到事先登记的图案包括于图像数据的情况下,可以将

识别到的图案所存在的监视器61上的范围作为收音范围A1,并将收音范围A1的大小的信息向范围指定部44赋予。识别到的图案例如是人物的整体、人物的脸部。

[0061] 声音译码器36将来自网络处理部15的多个声道的声音数据输入并译码。而且,在声音译码器36中,对多个声道的声音数据进行处理的声音译码器可以分别独立地设置。在这种情况下,能够同时处理阵列麦克风10的麦克风11A~11C分别收集到的多个声道的声音数据。

[0062] 收音角度运算部38基于收音坐标指定部35决定的收音坐标,导出(例如算出)表示阵列麦克风10的指向性的方向的收音角度 θ_1 。收音角度运算部38导出的收音角度 θ_1 作为指向性处理部37的参数而输入。例如,可以是收音坐标与收音角度 θ_1 一一对应,包括该对应信息的转换表存储于未图示的存储器中。收音角度运算部38可以参照该转换表来导出收音角度 θ_1 。

[0063] 范围角度运算部45基于范围指定部44决定的收音范围A1的大小,导出(例如算出)表示阵列麦克风10的指向性的范围的大小(指向性的扩展、指向性的强弱)的范围角度 θ_2 。范围角度运算部45导出的范围角度 θ_2 作为指向性处理部37的参数而输入。例如,可以是收音范围A1的大小与范围角度 θ_2 一一对应,并将包括该对应信息的转换表存储于未图示的存储器。范围角度运算部45可以参照该转换表,并导出范围角度 θ_2 。

[0064] 在此,说明收音范围A1的大小与范围角度 θ_2 的关系。

[0065] 图10(A)、图10(B)是表示收音范围A1及范围角度 θ_2 的一例的示意图。

[0066] 图10(A)例示范围角度 θ_{2a} 作为收音范围A1比较小(窄)的情况下的范围角度 θ_2 。在图10(A)中,指定包括位于监视器61的中央部的人物P11的方向且不包括人物P12、P13的方向的范围作为收音范围A1。因此,在图10(A)的状态下3人的人物P11~P13进行交谈的情况下,能够听取中央部的人物P11的声音。因此,能够保护人物P12、P13的隐私。

[0067] 图10(B)例示范围角度 θ_{2b} 作为收音范围A1比较大(广)的情况下的范围角度 θ_2 。在图10(B)中,指定监视器61所显示的包括3人的人物P11~P13的方向的范围作为收音范围A1。因此,在图10(B)的状态下3人的人物P11~P13进行交谈的情况下,能够听取人物P11~P13的声音。因此,能够提高监视器61上所显示的区域的监视精度,能够确保安全性。

[0068] 指向性处理部37从收音角度运算部38取得收音角度 θ_1 的信息,从范围角度运算部45取得范围角度 θ_2 的信息,从声音译码器36取得声音数据。指向性处理部37根据收音角度 θ_1 及范围角度 θ_2 ,将从声音译码器36输出的多个声道的声音数据按照预定算法进行合成,形成指向性(指向性处理)。

[0069] 例如,指向性处理部37提升监视对象的人物所存在的场所(关注点)的方向(指向性的方向)及范围(指向性的范围)的声音成分的信号等级,并降低除此以外的方向的声音成分的信号等级。指向性处理部37将指向性处理后的声音数据向检测部39及声音输出部42输出。

[0070] 指向性处理部37例如根据范围角度 θ_2 ,通过公知的方法,进行指向性处理。例如,指向性处理部37可以根据范围角度 θ_2 ,来决定指向性处理使用的声音数据的声道数、即使用收音到的声音的麦克风11的个数。例如,指向性处理部37可以在范围角度 θ_2 越小时,越增强指向性,因此容易增加声音数据的声道数。例如,指向性处理部37也可以在范围角度 θ_2 越大时,越减弱指向性,因此减少声音数据的声道数。

[0071] 例如,在图10(A)中,指向性处理部37使用由阵列麦克风10所包含的16个麦克风11中的8个麦克风11收音到的声音数据,进行指向性处理。例如,在图10(B)中,指向性处理部37使用由阵列麦克风10所包含的16个麦克风11中的4个麦克风11收音到的声音数据,进行指向性处理。

[0072] 指向性处理部37可以根据由声源推定部40推定出的声源(例如监视对象的人物、异常音)的位置,进行指向性处理。指向性处理部37例如可以从声源推定部40多次取得声源的推定位置的信息,且每次取得时变更(例如切换)指向性的方向。由此,即使在声源移动的情况下,也能够追踪并监视声源的位置。即,在声源的位置的追踪中,将指向性转向推定出的声源的位置。

[0073] 检测部39取得通过指向性处理部37进行了指向性处理的声音数据。该声音数据包括例如对第一指向性的方向及范围的声音成分进行了增强的第一声音数据和对第二指向性的方向及范围的声音成分进行了增强的第二声音数据。检测部39根据取得的声音数据检测监视对象音(预定声音的一例)。即,检测部39具有作为声音检测部的功能。另外,在本实施方式中,声音成分的增强是指例如利用由多个麦克风形成的阵列麦克风,通过滤波仅提取来自特定的方向及范围的声音。

[0074] 另外,检测部39在检测到监视对象音的情况下,进行各种处理。关于检测部39的详情,在后文叙述。检测部39是在检测到监视对象音的情况下进行预定处理的处理部的一例。

[0075] 声源推定部40取得来自声音译码器36的声音数据,推定发出由检测部39检测到的监视对象音的声源的位置。声源广泛地包括例如交谈中的人物、发出响声的人物、特定的人物(男性、女性)、物体(例如紧急车辆)、异常音(例如,警铃、警笛)的产生源、特定的环境音的产生源、其他的声源。声源推定部40是推定部的一例。

[0076] 声源推定部40例如通过公知的声源推定技术来推定声源的位置。声源推定部40产生的声源的位置的推定结果用于例如基于指向性处理部37的异常音的追踪、指向性的切换。

[0077] 声源推定部40可以将声源的位置的推定结果向例如图像输出部33或声音输出部42输出。图像输出部33或声音输出部42对声源的位置的推定结果进行提示,由此操作者60能够容易地掌握声源的位置。

[0078] 声音输出部42例如将来自指向性处理部37的声音数据从数字声音数据转换成模拟声音数据,将声音数据放大,向扬声器63提供。

[0079] 扬声器63输出与来自声音输出部42的声音数据相当的声音。因此,操作者60能够从扬声器63听见对由阵列麦克风10收音到的声音数据进行了处理的声音。扬声器63是提示部的一例。

[0080] 数据记录部43例如可以包括HDD(Hard Disk Drive)、SSD(Solid State Drive),并依次记录网络处理部31取得的多个声道的声音数据或图像数据。数据记录部43在记录声音数据及图像数据的情况下,将声音数据的生成时刻与图像数据的生成时刻建立对应地记录。而且,可以将上述生成时刻的信息与声音数据或图像数据一起记录。数据记录部43可以设置在监视控制装置30的内部,或者可以设置在监视控制装置30的外部作为外部记录介质。

[0081] 而且,数据记录部43记录例如用于对记录的声音数据或图像数据进行检索的检索

用标志的信息。记录于数据记录部43的检索用标志由监视控制装置30中的其他的结构部适当参照。

[0082] 接下来,说明检测部39的详情。

[0083] 检测部39例如在指向性处理后的声音数据的信号等级为第一预定阈值以上或第二预定阈值以下的情况下,检测该声音数据作为监视对象音。与声音数据的信号等级比较的阈值的信息例如保持在未图示的存储器中。在声音数据的信号等级成为第二预定阈值以下的情况下,例如,虽然产生机械的动作音,但是也包括该机械停止而动作音消失的情况。

[0084] 检测部39例如检测进行了指向性处理的声音数据所包含的异常音作为监视对象音。例如,异常音的图案存储于未图示的存储器,检测部39在声音数据包括异常音的图案的情况下,检测异常音。

[0085] 检测部39例如检测进行了指向性处理的声音数据所包含的预定关键字作为监视对象音。例如,将关键字的信息存储于未图示的存储器,检测部39在声音数据中包括记录于存储器的关键字的情况下,检测关键字。另外,在检测关键字的情况下,例如,可以使用公知的声音识别技术。在这种情况下,检测部39具有公知的声音识别功能。

[0086] 另外,监视对象音可以预先设定。例如,检测部39可以将信号等级为第一预定阈值以上或第二预定阈值以下的声音、异常音、关键字中的至少一个设定为监视对象音。该设定信息例如存储在未图示的存储器中。

[0087] 检测部39在检测到上述的监视对象音的情况下,将检测到监视对象音的内容的信息(检测信息)向图像输出部33及声音输出部42中的至少一方发送。检测信息包括例如异常音、具有第一预定阈值以上或第二预定阈值以下的信号等级的声音、检测到预定关键字的内容的警告信息(警报)。

[0088] 而且,检测部39在检测到监视对象音的情况下,向数据记录部43发送预定信息。检测部39在检测到监视对象音的情况下,例如,可以将检索用标志的信息向数据记录部43发送、保持。检索用标志是用于从数据记录部43中检索包括监视对象音的声音数据或与该声音数据对应的图像数据的标志。

[0089] 检索用标志例如可以在与实时取得的声音数据或图像数据相同的时机,记录于数据记录部43。而且,检索用标志例如也可以与已经记录于数据记录部43的声音数据或图像数据建立对应地记录于数据记录部43。

[0090] 例如,操作者60经由触摸面板62输入与检索用标志一致或对应的信息,由此图像译码器32或声音译码器36在记录于数据记录部43的声音数据或图像数据中,检索、取得与检索用标志一致或对应的数据。因此,例如,即使在对声音数据或图像数据进行长时间录音或录像的情况下,也能够缩短检索时间。

[0091] 而且,操作者60例如可以经由触摸面板62,从时序地排列有多个检索用标志的列表中选择特定的检索用标志。在这种情况下,操作者60可以从生成时刻最旧的或最新的检索用标志起依次选择特定的检索用标志。而且,操作者60例如可以经由触摸面板62,选择在与由计时部(未图示)计时的时刻对应的时刻生成的检索用标志作为特定的检索用标志。图像译码器32或声音译码器36在记录于数据记录部43的声音数据或图像数据中,检索、取得与上述特定的检索用标志一致或对应的数据。列表例如记录于数据记录部43。

[0092] 检索用标志包括例如由检测部39检测到监视对象音的时刻的信息。检索用标志包

括例如产生监视对象音的声源的方向(指向性的方向)的信息。检索用标志包括例如包括产生了监视对象音的声源的范围的大小(指向性的范围的大小)的信息。检索用标志包括例如监视对象音的类别(包括异常音、关键字的声音、预定阈值以上或预定阈值以下的信号等级的声音)的信息。声音的类别例如通过检测部39使用公知的声音识别技术来判定。

[0093] 检索用标志包括例如通过VMD功能检测到的监视对象音的声源的移动的有无或移动的方向的信息。检测移动的有无等的声源例如包括于在上述的监视对象音的产生时刻或产生时间带由相机20拍摄到的图像数据中。通过VMD功能检测到的信息例如每当检测到移动时,从图像识别部34向检测部39发送。

[0094] 检索用标志包括例如由图像识别部34进行了图像识别的监视对象音的声源的类别的信息。识别声源的类别的图像数据例如是在监视对象音的产生时刻或产生时间带由相机20拍摄到的图像数据。声源的类别的信息从图像识别部34向检测部39发送。

[0095] 检索用标志包括例如缩略图像(静止图像)。缩略图像是例如在监视对象音的产生时刻或产生时间带由相机20拍摄到的图像数据的至少一部分。缩略图像从图像识别部34向检测部39发送。

[0096] 检测部39可以在检测到监视对象音的情况下,开始由网络处理部31接收到的声音数据或图像数据的录音或录像。例如,网络处理部31暂时蓄积预定期间(例如30秒钟)的声音数据或图像数据,在由检测部39未检测到监视对象音的情况下,将暂时蓄积的声音数据或图像数据废弃。检测部39在检测到监视对象音的情况下,对网络处理部31作出指示,进行控制而将暂时蓄积中的声音数据或图像数据包括在内地向数据记录部43记录声音数据或图像数据(预录音、预录像)。而且,数据记录部43记录来自网络处理部31的声音数据或图像数据。预录音、预录像可以在经过预定时间之后结束。

[0097] 检测部39在检测到预定关键字作为监视对象音的情况下,可以不将包括该关键字的声音数据向数据记录部43记录而消去。或者,检测部39在检测到预定关键字作为监视对象音的情况下,可以从声音数据消去该关键字的部分,也可以通过该关键字以外的声音进行置换。检测部39可以将消去或置换了关键字的部分的声音数据向数据记录部43记录。由此,在关键字应隐匿的信息的情况下,能够保护隐匿信息或隐私。这样的与关键字的消去或置换相关的处理也称为“关键字加工”。另外,关键字加工也可以对于向数据记录部43记录完的声音数据进行。

[0098] 检测部39在检测到监视对象音的情况下,可以对指向性处理部37作出指示,来切换指向性的方向及指向性的范围的大小中的至少一方。在这种情况下,指向性处理部37可以将指向性的方向切换为预定方向,也可以将指向性的范围的大小切换为预定大小。

[0099] 例如,可以将相机20的可拍摄范围所包含的多个地点(地点A、地点B)的信息预先登记在未图示的存储器中。在地点A的方向上检测到监视对象音的情况下,指向性处理部37可以将指向性的方向从地点A的方向切换为地点A以外的地点(例如地点B)的方向。

[0100] 例如,将相机20的可拍摄范围所包含的多个范围(范围A、范围B)的大小的信息预先登记于未图示的存储器。在范围A中检测到监视对象音的情况下,指向性处理部37可以将指向性的范围的大小从范围A的大小切换为范围A以外的大小(例如范围B的大小)。

[0101] 检测部39在检测到预定关键字作为监视对象音的情况下,可以将包括该关键字的声音数据向数据记录部43记录。该记录可以包括预录音、预录像。由此,操作者60通过将应

监视的关键字预先登记,能够以关键字为触发而开始记录,能够提高监视精度。

[0102] 接下来,说明阵列麦克风10、相机20及各声源的配置状态。

[0103] 图3是表示阵列麦克风10、相机20及各声源的配置状态的一例的示意图。

[0104] 在图3中,例如,收音单元90固定于屋内的顶棚面101。在图3中,阵列麦克风10所包含的多个麦克风11A~11C沿着顶棚面101(收音单元90的设置面)排列。附图标记PA表示声源。

[0105] 而且,收音单元90以使阵列麦克风10的基准方向与相机20的基准方向(例如光轴方向)一致的方式安装于顶棚面101。相对于阵列麦克风10的基准方向的水平方向及垂直方向与相对于相机20的基准方向的水平方向及垂直方向一致。该水平方向是x轴方向、y轴方向,垂直方向是z轴方向。

[0106] 阵列麦克风10的基准方向例如是阵列麦克风10中的各麦克风11排列的排列方向。收音角度 θ_1 是通过阵列麦克风10的基准方向和指向性的方向而形成的角度。通过阵列麦克风10的基准方向和指向性的方向而形成的收音角度 θ_1 的水平方向成分是水平角 θ_{1h} 。通过阵列麦克风10的基准方向和指向性的垂直方向而形成的收音角度 θ_1 的垂直方向成分是垂直角 θ_{1v} 。

[0107] 在收音单元90中,阵列麦克风10中的各麦克风11在圆周上以一定的间隔排列,因此相对于沿着排列面(x-y面)的水平方向,无论对于哪个方向,声音数据的频率特性都相同。因此,在图3的例子中,收音角度 θ_1 实质上依赖于垂直角 θ_{1v} 。因此,在以下的说明中,主要作为收音角度 θ_1 ,不考虑水平角 θ_{1h} 进行说明。

[0108] 如图3所示,收音单元90中的阵列麦克风10的收音角度 θ_1 (垂直角 θ_{1v})是平行于麦克风11A~11C的排列面的方向(x轴、y轴)与指向性的灵敏度最大的方向所成的角度。

[0109] 麦克风11A~11C对朝向麦克风11A~11C到来的声音进行收音。而且,相机20以正下方(z轴方向)的方向为基准方向(光轴方向),对相机20的周围的例如全方位进行拍摄。

[0110] 另外,阵列麦克风10的收音对象或相机20的拍摄对象可以不是全方位而限制为一部分的方向。而且,阵列麦克风10或监视控制装置30可以对收音对象限制为一部分的方向的状态下收音到的声音数据进行合成,生成与收音对象为全方位的情况同样的声音数据。而且,相机20或监视控制装置30也可以对拍摄对象限制为一部分的方向的状态下拍摄到的图像信号进行合成,生成与拍摄对象为全方位的情况同样的图像信号。

[0111] 另外,例如,在阵列麦克风10的基准方向与相机20的基准方向不一致的情况下,可以考虑水平角 θ_{1h} 。在这种情况下,可以考虑水平角 θ_{1h} 和垂直角 θ_{1v} ,例如根据三维(x,y,z)的位置或方向来形成指向性。

[0112] 接下来,说明监视控制装置30的动作例。

[0113] 图4是表示监视控制装置30的动作例的流程图。

[0114] 图4示出实时动作例。实时动作是例如操作者60使用监视控制装置30实时地对阵列麦克风10收音到的声音数据及相机20拍摄到的图像进行监视的情况的动作。

[0115] 在图4中,首先,网络处理部31经由网络50来接收相机20送出的图像数据。而且,网络处理部31经由网络50来接收阵列麦克风10送出的多个声道的声音数据(S11)。

[0116] 网络处理部31接收到的图像数据由图像译码器32译码,向图像输出部33发送。图像输出部33将译码后的图像数据向监视器61输出,并进行控制以使监视器61显示图像

(S12)。而且,网络处理部31可以将图像数据及声音数据向数据记录部43记录。

[0117] 接下来,收音坐标指定部35例如接受来自触摸面板62的坐标输入(S13)。范围指定部44例如接受基于来自触摸面板62的坐标输入的收音范围,并导出收音范围的大小(S13)。例如,操作者60识别辨认监视器61上显示的图像的显示位置,操作触摸面板62而指定应关注的图像位置及图像范围。图像位置的指定及图像范围的指定哪个先进行均可。

[0118] 收音坐标指定部35导出与指定的图像范围(图像位置)对应的收音坐标。范围指定部44导出与指定的图像范围对应的收音范围的大小。操作者60例如触摸监视器61上显示的图像所包含的特定的人物的位置(例如图1的附图标记P1),通过缩小操作或放大操作来指定范围的大小(例如图1的附图标记A1的大小)。由此,收音坐标指定部35取得收音坐标及范围坐标。上述图像范围是例如监视者应监视的监视区域的一例。

[0119] 收音坐标指定部35可以取代操作者60对图像范围的位置的指定,使图像识别部34从图像中识别预定图案,并取得预定图案所存在的坐标作为收音坐标。

[0120] 范围指定部44可以取代操作者60对图像范围的大小的指定,使图像识别部34从图像中识别预定图案,并取得预定图案所存在的范围的大小的信息作为收音范围的大小的信息。

[0121] 收音角度运算部38基于由收音坐标指定部35取得的收音坐标,例如参照转换表,或进行公知的运算处理,由此导出收音角度 θ_1 (S14)。

[0122] 范围角度运算部45基于由范围指定部44取得的收音范围的大小的信息,例如参照转换表,或进行公知的运算处理,由此导出范围角度 θ_2 (S14)。

[0123] 导出的收音角度 θ_1 及范围角度 θ_2 向指向性处理部37输入。指向性处理部37根据收音角度 θ_1 及范围角度 θ_2 ,导出阵列麦克风10的指向性处理用的参数。并且,指向性处理部37对于来自声音译码器36的声音数据,使用导出的参数进行指向性处理(S15)。由此,在指向性处理部37输出的声音数据中,例如,相对于收音角度 θ_1 的方向且与范围角度 θ_2 对应的的大小的范围而阵列麦克风10的收音灵敏度变得最大。

[0124] 接下来,检测部39从进行了指向性处理的声音数据中,检测监视对象音(例如,异常音、预定关键字、第一预定阈值以上或第二预定阈值以下的信号等级的声音)(S16)。在检测到监视对象音之前,在S16中等待。

[0125] 接下来,图像识别部34例如可以对包括检测到的监视对象音的声源的图像数据进行图像识别,来识别监视对象音的声源的类别(例如,人、男性、女性、物体、其他的声源)(S17)。由此,操作者60根据声源的类别能够容易地判断是否应进行监视,因此能够减轻操作者60的负担,能够提高监视精度。

[0126] 图像识别部34例如可以使用VMD功能来检测监视对象音的声源的移动(S17)。由此,操作者60能够容易地关注声源的移动,因此能够减轻操作者60的负担,能够提高监视精度。

[0127] 图像识别部34可以将图像识别后的结果(例如,监视对象音的声源的类别的信息、监视对象音的声源的移动的信息)向检测部39发送。

[0128] 另外,S17的处理可以省略。例如,可以是使用者经由触摸面板62来设定是否省略S17的处理的信息,也可以是未图示的控制部根据监视等级来设定是否省略S17的处理的信息。是否省略S17的处理的信息例如保持在未图示的存储器中。

[0129] 接下来,监视控制装置30根据检测部39的检测结果及图像识别部34的图像识别结果中的至少一方,进行预定处理(动作)(S18)。

[0130] 例如,在检测到监视对象音的情况、识别到声源的类别的情况或者检测到声源的移动的情况、即产生监视触发的情况下,检测部39可以对图像输出部33进行指示以通过图像来警告信息。而且,在产生监视触发的情况下,检测部39也可以对声音输出部42进行指示,以通过声音来警告信息(S18)。而且,检测部39可以根据监视触发的类别,进行不同的警告音的鸣叫、警告信息的显示。由此,监视控制装置30的操作者60能够容易地识别监视对象音的产生等,能够减轻操作者60的负担,能够提高监视精度。

[0131] 例如,检测部39在产生了监视触发的情况下,可以将检索用标志的信息向数据记录部43记录(S18)。由此,即使在操作者60将来再看声音数据或图像数据的情况下,也能够容易地检索所希望的声音数据或图像数据的特定的部位,能够缩短例如验证时间。

[0132] 例如,在产生了监视触发的情况下,检测部39可以对网络处理部31进行指示,以进行预录音及预录像中的至少一方(S18)。由此,在产生监视触发之前,不进行向数据记录部43的录音或录像,由此能够提高数据记录部43的有效利用效率。而且,在产生了监视触发的情况下,能够可靠地记录监视触发产生时刻的声音数据或图像数据,例如,将来能够作为验证材料进行确认。

[0133] 例如,在检测到预定关键字作为监视对象音的情况下,检测部39可以进行关键字加工(S18)。由此,即使在关键字为隐匿信息的情况下,也能够保护隐匿信息。而且,在消去或置换关键字而记录包括关键字的声音数据的情况下,能够保护隐匿信息并保存声音数据。

[0134] 例如,在产生监视触发的情况下,检测部39可以对指向性处理部37进行指示,以切换指向性的方向(S18)。由此,例如,通过以朝向预设的方向的方式变更指向性的方向,能够提高在预想到声源的移动的情况下能够追踪监视对象音的可能性。

[0135] 例如,在产生监视触发的情况下,检测部39可以对指向性处理部37进行指示,以切换指向性的范围的大小(S18)。由此,例如,通过以将预设的范围的大小作为指向性的范围的大小的方式进行变更,能够提高在预想到声源的移动的情况下能够追踪监视对象音的可能性。例如,通过指向性的范围的大小的变更,即使在相机20与声源的距离发生变化的情况下,也能够提高监视对象音的追踪精度。

[0136] 接下来,声源推定部40推定监视对象音的声源的位置(S19)。由此,能够提高操作者60的监视精度。

[0137] 接下来,指向性处理部37在预定时机(例如每预定时间),取得从声源推定部40推定出的监视对象音的声源的位置的信息,以使指向性朝向该声源的位置的方式切换指向性的方向(S20)。由此,能够追踪监视对象音的声源,操作者60能够容易地监视声源的动向,能够提高监视精度。

[0138] 另外,S19、S20可以省略。

[0139] 根据图4的动作例,操作者60经由监视器61及扬声器63,能够同时监视当前的监视区域中的图像和声音。尤其是能够监视包括监视对象音的声音及监视对象音的声源的图像。而且,能够边确认图像边指定任意的监视区域,并作为监视对象。监视区域的指定包括例如监视区域的位置、范围的大小的指定。而且,根据监视对象音的检测而进行各种处理,

由此声音数据及图像数据的有效利用性增加,能够提高便利性。

[0140] 接下来,说明监视系统100的指向性处理的详情。

[0141] 图5是表示指向性处理的基本构成例的示意图。在图5中,指向性处理部37包括多个延迟器37bA、37bB、37bC和加法运算器37c,可以通过这些部件的处理来形成指向性。

[0142] A/D转换器13A、13B、13C将麦克风11A~11C输出的模拟声音数据转换成数字声音数据,指向性处理部37对于转换后的数字声音数据进行指向性处理。另外,阵列麦克风10所包含的麦克风的个数(n)、A/D转换器的个数(n)及指向性处理部37所包含的延迟器的个数(n)根据需要而增减。

[0143] 在图5中,多个麦克风11A~11C配置在相互分离预定距离的位置,因此从1个声源80发出的声波到达各个麦克风11A~11C为止的时间产生相对的时间差(到达时间差)。声源80是例如监视对象音的声源。

[0144] 由于存在上述到达时间差的影响,若将多个麦克风11A~11C分别检测到的声音数据直接相加,则由于存在相位差的多个声音数据的相加而信号等级有时会衰减。因此,通过延迟器37bA~37bC向多个声音数据分别赋予时间延迟来调整相位,并通过加法运算器37c将调整了相位的声音数据彼此相加。由此,将相位一致的多个声音数据相加,信号等级增大。

[0145] 在图5中,上述到达时间差根据从声源80向阵列麦克风10的框体入射面121入射的声波的到达方向(相当于收音角度 θ_1)而变化。例如,在多个麦克风11A~11C检测到从特定的方向(θ_1)到来的声波的情况下,向加法运算器37c输入的多个声音数据的相位一致,加法运算器37c输出的声音数据的信号等级增大。另一方面,在从特定的方向(θ_1)以外到来的声波中,向加法运算器37c输入的多个声音数据产生相位差,加法运算器37c输出的声音数据的信号等级衰减。由此,能够以提高对于从特定的方向(θ_1)到来的声波的灵敏度的方式形成阵列麦克风10的指向性。

[0146] 在监视对象音的声波从收音角度 θ_1 的方向到达框体入射面121的情况下,由(式1)表示的各延迟时间D1、D2、D3被分配为各延迟器37bA、37bB、37bC的延迟时间。

[0147] $D1 = L1/Vs = d(n-1) \cos\theta_1/Vs$

[0148] $D2 = L2/Vs = d(n-2) \cos\theta_1/Vs \cdots \cdots$ (式1)

[0149] $D3 = L3/Vs = d(n-3) \cos\theta_1/Vs$

[0150] 其中,

[0151] L1:第1个麦克风与第n个麦克风之间的声波到达距离差(已知的常数)

[0152] L2:第2个麦克风与第n个麦克风之间的声波到达距离差(已知的常数)

[0153] L3:第3个麦克风与第n个麦克风之间的声波到达距离差(已知的常数)

[0154] Vs:音速(已知的常数)

[0155] d:麦克风的配置间隔(已知的常数)

[0156] 另外,作为一例,在图2所示的系统结构的情况下, $n=3$,在图1所示的收音单元90的情况下, $n=16$ 。

[0157] 如(式1)所示,在指向性与从特定的方向 θ_1 到达阵列麦克风10的声波一致的情况下,对应于框体入射面121处的相对于各麦克风11A~11C的声波的到达时间差,向各延迟器37bA、37bB、37bC分配延迟时间D1~D3。

[0158] 例如,指向性处理部37基于来自收音角度运算部38的收音角度 θ_1 和(式1),取得各延迟时间D1~D3,并将延迟时间D1~D3向各延迟器37bA~37bC分配。由此,能够增强从收音角度 θ_1 的方向到达框体入射面121的声波的声音数据,形成阵列麦克风10的指向性。

[0159] 另外,例如,分配的延迟时间D1~D3、(式1)中的已知的常数存储在监视控制装置30内的未图示的存储器中。

[0160] 根据监视系统100,例如,在实时地接收到的图像数据中,接受来自监视控制装置30的操作者60的监视区域的指定,在指向性朝向与监视区域对应的方向及范围的状态下,能够监视异常的有无。在产生了监视触发的情况下,监视控制装置30进行各种处理,由此能促进由阵列麦克风10收音到的声音数据及由相机20拍摄到的图像数据的有效利用,能够提高操作者60的便利性。

[0161] (第二实施方式)

[0162] 在第二实施方式中,假定将记录声音数据或图像数据的记录器与监视控制装置分体地设置于监视系统的情况。

[0163] 图6是实施方式的监视系统100B的概要图。图6与图1相比,不同点在于监视系统100B具备记录器70。记录器70与网络50连接。记录器70是存储装置的一例。记录器70存储例如由阵列麦克风10收音到的声音数据、由相机20拍摄到的图像数据。

[0164] 图7是表示监视系统100B的构成例的框图。在图7的监视系统100B中,关于与图2所示的监视系统100同样的结构,标注同一附图标记,省略或简化说明。

[0165] 监视系统100B具备阵列麦克风10、相机20、监视控制装置30B及记录器70。

[0166] 与图2所示的监视控制装置30相比,监视控制装置30B不具备数据记录部43。监视控制装置30B取代向数据记录部43记录数据或者从数据记录部43读出数据,而访问记录器70具备的数据记录部72来记录数据或者从数据记录部72读出数据。在监视控制装置30B与记录器70之间进行数据通信的情况下,数据经由监视控制装置30B的网络处理部31、网络50及记录器70的网络处理部71进行通信。

[0167] 记录器70具备网络处理部71及数据记录部72。记录器70包括例如CPU、DSP、ROM或RAM,通过CPU或DSP,执行记录在ROM或RAM中的控制用的程序,由此实现各种功能。

[0168] 网络处理部71例如经由网络50取得从阵列麦克风10送出的多个声道的声音数据或者从相机20送出的图像数据。网络处理部71例如将数据记录部72记录的声音数据或图像数据向网络50送出。

[0169] 数据记录部72具有与图2所示的监视控制装置30的数据记录部43同样的结构及功能。而且,数据记录部72记录与数据记录部43记录的数据同样的数据(例如,声音数据、图像数据、检索用标志的信息)。

[0170] 例如,在网络处理部71从监视控制装置30B接收到声音数据、图像数据、检索用标志的信息的情况下,数据记录部72可以将接收到的数据建立对应地记录。而且,网络处理部71从监视控制装置30B接收检索用标志的信息,在数据记录部72已经记录有声音数据或图像数据的情况下,数据记录部72可以将检索用标志的信息与声音数据或图像数据建立对应地记录。

[0171] 而且,例如通过CPU执行预定命令,由此将记录于数据记录部72的声音数据、图像数据及检索用标志的信息从数据记录部72读出,经由网络处理部71及网络50向监视控制装

置30B发送。

[0172] 例如,在经由网络50从监视控制装置30B接收到预定信息的情况下,数据记录部72判定作为检索用标志而记录的信息与接收到的预定信息是否一致或对应。数据记录部72在判定为两者一致的情况下,检索与检索用标志建立对应的声音数据或图像数据,并将检索到的声音数据或图像数据向网络50送出。

[0173] 这样,通过使用记录于记录器70的检索用标志,能够容易地检索过去记录的声音数据或图像数据,能够缩短检索时间,能够提高操作者60的便利性。

[0174] 接下来,说明监视控制装置30B的动作例。

[0175] 图8是表示监视控制装置30B的动作例的流程图。

[0176] 图8示出记录器输出重放动作例。记录器输出重放动作例例如是操作者60使用监视控制装置30B对记录于记录器70的过去的声音数据及图像信号进行分析的情况的动作。另外,在图8中,关于进行与图4的处理同样的处理的步骤,标注同一步骤编号,省略或简化说明。

[0177] 在监视系统100B中,在过去相机20拍摄到的图像数据及阵列麦克风10收音到的多个声道的声音数据记录于记录器70的情况下,能够从记录器70读出记录的图像数据及声音数据。

[0178] 监视控制装置30B例如按照来自操作者60的输入操作,对记录器70作出指示,以读出记录于记录器70的特定的图像数据及声音数据。在这种情况下,上述特定的图像数据及声音数据从记录器70被读出,经由网络50而由网络处理部31接收(S21)。

[0179] 接下来,进行图8的S12~S20的处理。另外,S17、S19、S20的处理可以省略。

[0180] 根据图8的动作例,操作者60经由监视器61及扬声器63,能够同时监视过去的监视区域中的图像和声音。尤其是能够监视包括监视对象音的声音及监视对象音的声源的图像。而且,能够边确认图像,边指定任意的监视区域,并作为监视对象。监视区域的指定包括例如监视区域的位置、范围的大小的指定。而且,根据监视对象音的检测来进行各种处理,由此,声音数据及图像数据的有效利用性增加,且能够提高便利性。

[0181] 另外,检索用标志与记录的图像数据或声音数据建立对应地记录,由此例如在之后检索与监视对象音相关的数据的情况下,能够快速检索。这样,通过根据监视对象音的检测来进行各种处理,声音数据及图像数据的有效利用性增加,且能够提高便利性。

[0182] 另外,图8的记录器输出重放动作例也可以应用于第一实施方式中对记录于数据记录部43的数据进行处理的情况的动作。

[0183] 根据监视系统100B,例如,在过去记录的图像数据中,接受来自监视控制装置30B的操作者60的监视区域的指定,在将指向性朝向与监视区域对应的方向及范围的状态下,能够监视异常的有无。在产生了监视触发的情况下,监视控制装置30B通过进行各种处理,能够促进由阵列麦克风10收音到的声音数据及由相机20拍摄到的图像数据的有效利用,提高操作者60的便利性。

[0184] 另外,本发明并不局限于上述实施方式的结构,只要是能够实现权利要求书所示的功能或本实施方式的结构具有的的功能的结构,则无论是何种结构都能够应用。

[0185] 例如,在上述实施方式中,阵列麦克风10或相机20可以具备监视控制装置30、30B所具备的与声音处理相关的一部分的结构部。阵列麦克风10例如可以具有图像识别部34、

收音坐标指定部35、范围指定部44、收音角度运算部38、范围角度运算部45、指向性处理部37、检测部39、声源推定部40的一部分或全部。由此,能够降低监视控制装置30、30B的处理负荷。另外,在阵列麦克风10具有与声音处理相关的一部分的结构部的情况下,在监视控制装置30、30B与阵列麦克风10之间,经由网络50进行适当必要的通信。

[0186] 例如,在上述实施方式中,例示了多个麦克风11在1个圆周上以一定的间隔配置的阵列麦克风10,但是各麦克风11的排列可以是不同的排列。例如,可以沿着单方向(例如x轴方向)以一定的间隔排列成一行。而且,可以沿着两个方向(例如x轴方向、y轴方向)以一定的间隔排列成十字型。而且,可以在直径不同的2个圆周上以一定的间隔排列。

[0187] 例如,在上述实施方式中,监视控制装置30、30B可以不使用相机20,而将实际的空间的监视范围的方向与阵列麦克风10的收音角度 θ_1 建立对应并预设。即,监视控制装置30、30B的未图示的存储器可以保持上述监视范围的方向与收音角度 θ_1 的对应信息。在这种情况下,例如,当使用者经由触摸面板62等指定监视范围的方向时,收音角度运算部38可以参照保持于存储器的对应信息,导出收音角度 θ_1 。而且,例如,使用者可以经由触摸面板62等而直接指定收音角度 θ_1 ,并将指定的数据作为由收音角度运算部38导出的数据进行处理。由此,能够不使用相机20而决定指向性的方向。

[0188] 例如,在上述实施方式中,监视控制装置30、30B可以不使用相机20,而将实际的空间的监视范围的大小与阵列麦克风10的范围角度 θ_2 建立对应并预设。即,监视控制装置30、30B的未图示的存储器可以保持上述监视范围的大小与范围角度 θ_2 的对应信息。在这种情况下,例如,当使用者经由触摸面板62等而指定监视范围的大小时,收音角度运算部38可以参照保持于存储器的对应信息,来导出范围角度 θ_2 。而且,例如,使用者可以经由触摸面板62等而直接指定范围角度 θ_2 ,并将指定的数据作为由范围角度运算部45导出的数据进行处理。由此,能够不使用相机20而决定指向性的范围的大小。

[0189] 例如,在上述实施方式中,监视系统100、100B可以是不使用图像而使用声音进行监视的系统。在这种情况下,在监视系统100、100B中,例如,可以省略相机20或者用于实现与显示关联的功能的结构部。

[0190] 例如,在上述实施方式中,例示了收音单元90固定于屋内的顶棚面101的情况,但是收音单元90也可以固定于其他的位置(例如,屋内的壁面)。而且,在监视系统100、100B中,可以设置多个相机20。而且,监视器61、触摸面板62及扬声器63可以包括于监视控制装置30、30B。

[0191] 例如,在上述实施方式中,可以在监视器61上显示用于进行音量调整的软键盘(屏幕键盘)。通过触摸面板62对软键盘进行操作,由此能够调整例如进行了指向性处理的声音数据的音量。

[0192] 例如,在上述实施方式中,监视控制装置30、30B的未图示的控制部可以校正根据设置收音单元90的环境而产生的声音数据的变形。而且,未图示的控制部可以校正由相机20(例如具有鱼眼透镜的相机)拍摄到的图像数据所产生的变形。

[0193] 例如,在上述实施方式中,例如,收音坐标指定部35可以通过触摸面板62来触摸监视区域,在指向性朝向了该监视区域之后通过触摸面板62再次触摸监视区域的情况下,将该监视区域从监视对象排除。即,收音坐标指定部35在由监视器61显示的图像数据中的同一位置或区域被进行了多次触摸的情况下,可以结束收音坐标的导出,使基于指向性处理

部37的指向性处理结束。

[0194] 在上述实施方式中,例如,收音坐标指定部35在以由触摸面板62触摸了监视区域的状态接受到拖曳操作的情况下,可以使监视区域移动。

[0195] 在上述实施方式中,监视系统100、100B可以具备多个收音单元90。在这种情况下,各收音单元90可以协作地形成图像数据,并形成声音数据。而且,由各收音单元90的相机20拍摄到的图像可以同时显示在监视器61的分割画面上。监视控制装置30、30B在各分割画面上,即使在通过触摸面板62跨多个分割画面地接受到拖曳操作的情况下,也可以使用由各收音单元90收音到的声音数据进行指向性处理。

[0196] 在上述实施方式中,收音坐标指定部35可以同时指定多个收音坐标。范围指定部44可以同时指定多个收音范围。在这种情况下,可以是,收音角度运算部38运算多个收音角度 θ_1 ,范围角度运算部45运算多个范围角度 θ_2 ,指向性处理部37生成多个方向及范围的声音成分被增强后的多个声音数据。

[0197] (本发明的一方式的概要)

[0198] 本发明的一方式的声音处理装置具备:数据取得部,取得由包括多个麦克风的收音部从预定区域内的声源收音到的声音数据及由拍摄上述预定区域内的至少一部分的图像的拍摄部拍摄到的图像数据;第一指定部,指定与基于上述图像数据而显示的图像上的指定部位对应的以上述收音部为基准的方向;第二指定部,指定与基于上述图像数据而显示的图像上的指定部位对应的上述预定区域内的预定范围;及指向性处理部,在所指定的上述预定范围内,增强所指定的上述方向的上述声音数据中的声音成分。

[0199] 根据该结构,操作声音处理装置的监视者在指定监视区域的方向及范围的大小且在与监视区域对应的方向及大小的范围内形成了指向性的状态下,能够监视异常的有无。由此,能够促进收音到的声音数据及图像数据的有效利用,提高监视者的便利性。

[0200] 另外,本发明的一方式的声音处理装置具备:声音检测部,根据由上述指向性处理部增强后的上述方向及上述范围的声音成分检测预定声音;及处理部,在通过上述声音检测部检测到上述预定声音的情况下,进行预定处理。

[0201] 根据该结构,声音处理装置伴随着预定声音的检测而实施各种动作,因此能够促进收音到的声音数据及拍摄到的图像数据的有效利用,能够提高监视者的便利性。

[0202] 另外,本发明的一方式的声音处理装置可以构成为,在检测到上述预定声音的情况下,上述处理部将检索用标志记录于记录部,该检索用标志用于从记录上述声音数据及上述图像数据的上述记录部检索包括上述预定声音的声音数据或包括上述预定声音的声源的图像数据。

[0203] 根据该结构,即使在监视者将来再看声音数据或图像数据的情况下,也能够容易地检索所希望的声音数据或图像数据、或者声音数据或图像数据的特定的部位,能够缩短例如验证时间。

[0204] 而且,本发明的一方式的声音处理装置可以构成为,上述处理部取得与记录于上述记录部的检索用标志所包含的预定检索用标志对应的记录于上述记录部的声音数据或图像数据。

[0205] 根据该结构,例如,能够使用经由触摸面板而使用者指定的检索用标志,检索发生了同样的现象的过去的声音数据或图像数据。因此,能够促进声音数据或图像数据的有效

利用,提高监视者的便利性,也能够提高监视精度。

[0206] 另外,本发明的一方式的声音处理装置可以构成为,上述检索用标志包括上述预定声音的类别、以上述收音部为基准的上述预定声音的声源的方向、包括以上述收音部为基准的上述预定声音的声源的上述范围的大小及由上述声音检测部检测到上述预定声音的时刻中的至少1个信息。

[0207] 根据该结构,能够记录各种信息作为与声音数据对应的检索用标志,在必要时能够迅速地检索所希望的声音数据。

[0208] 而且,本发明的一方式的声音处理装置可以构成为,在检测到上述预定声音的情况下,上述处理部使提示部提示包括检测到上述预定声音的内容的警告信息。

[0209] 根据该结构,例如,能够将产生了与异常相伴的声音的情况向监视者通知,催促强化监视等的注意。

[0210] 而且,本发明的一方式的声音处理装置可以构成为,在检测到上述预定声音的情况下,上述处理部将包括上述预定声音的声音数据记录于记录部。

[0211] 根据该结构,在未检测到预定声音的时刻不记录声音数据,由此能够提高记录部的有效利用效率。而且,在检测到预定声音的情况下,能够可靠地记录预定声音的产生时刻的声音数据,例如,将来能够作为验证材料进行确认。而且,在同样也记录图像数据的情况下,能够进一步提高监视精度。

[0212] 而且,本发明的一方式的声音处理装置可以构成为,在检测到上述预定声音的情况下,上述处理部对通过上述指向性处理部增强声音成分的方向及上述范围的大小中的至少一方进行变更。

[0213] 根据该结构,例如,在预定声音的声源移动且预想到其移动的情况下,以朝向预设的方向及大小的范围的方式变更指向性的方向及范围的大小,由此能够提高能够追踪预定声音的可能性。

[0214] 而且,本发明的一方式的声音处理装置可以构成为,具备推定部,该推定部推定发出上述预定声音的声源的位置,并使提示部提示推定出的位置的信息。

[0215] 根据该结构,监视者能够容易地确认推定出的声源的位置,因此能够减轻监视者的负担,能够提高监视精度。

[0216] 而且,本发明的一方式的声音处理装置可以构成为,具备推定发出上述预定声音的声源的位置的推定部,上述指向性处理部增强从由上述推定部推定出的上述声源的位置的方向到来的声音成分。

[0217] 根据该结构,监视者依靠声音数据而能够追踪预定声音的声源,因此能够容易地监视声源的动向,能够提高监视精度。

[0218] 而且,本发明的一方式的声音处理装置可以构成为,在由上述指向性处理部增强后的声音成分的信号等级为第一预定信号等级以上或第二预定信号等级以下的情况下,上述声音检测部检测该声音成分作为上述预定声音。

[0219] 根据该结构,例如通过能够检测比通常预想的生活音大的声音地设定阈值,能够提高检测异常音的可能性。

[0220] 而且,本发明的一方式的声音处理装置可以构成为,上述声音检测部根据由上述指向性处理部增强后的声音成分中的至少1个声音成分检测预定关键字作为预定声音。

[0221] 根据该结构,例如,能够检测指向性的方向的声音成分中的隐匿性高的关键字,能够对关键字实施各种处理,因此能够提高使用者的便利性。

[0222] 而且,本发明的一方式的声音处理装置可以构成为,上述处理部对包括检测到的上述预定关键字的声音数据中的上述预定关键字的部分进行加工。

[0223] 根据该结构,即使在关键字为隐匿信息的情况下,也能够保护隐匿信息。而且,例如,在将关键字消去或置换而记录包括关键字的声音数据的情况下,能够保护隐匿信息,并保存声音数据。

[0224] 而且,本发明的一方式的声音处理装置可以构成为,上述处理部将包括检测到的上述预定关键字的声音数据记录于记录部。

[0225] 根据该结构,能够以在监视区域中发出的关键字为触发来开始声音数据的记录。由此,在未检测到关键字的时刻不记录声音数据,由此能够提高记录部的有效利用效率。而且,在检测到关键字的情况下,也可以记录包括关键字的声音数据,例如,将来可以作为验证材料进行确认。

[0226] 而且,本发明的一方式的声音处理装置可以构成为,上述声音检测部检测由上述指向性处理部增强后的声音成分中的至少一个声音成分所包含的预定异常音作为上述预定声音。

[0227] 根据该结构,例如,能够检测指向性的方向的声音成分中表现出异常的异常音,能够使用异常音来实施各种处理,因此能够提高使用者的便利性。

[0228] 而且,本发明的一方式的声音处理装置可以构成为,具备对上述图像数据进行图像识别的图像识别部,上述处理部根据上述图像识别部的图像识别结果来进行上述预定处理。

[0229] 根据该结构,通过与预定声音的检测一并也考虑图像识别结果,例如,能够提高监视区域的异常检测精度。因此,监视者能够适当地应对,能够提高监视精度。

[0230] 而且,本发明的一方式的声音处理装置可以构成为,上述图像识别部在上述图像数据中识别上述预定声音的声源的类别。

[0231] 根据该结构,监视者能够根据声源的类别容易地判断是否应监视,因此能够减轻监视者的负担,能够提高监视精度。而且,在附加表示声源的类别的检索用标志的情况下,即使在将来再看声音数据的情况下,也能够根据声源的类别容易地检索,例如能够缩短验证时间。

[0232] 而且,本发明的一方式的声音处理装置可以构成为,上述图像识别部在上述图像数据中识别上述预定声音的声源的移动的有无。

[0233] 根据该结构,监视者能够容易地关注声源的移动,因此能够减轻监视者的负担,能够提高监视精度。而且,在附加表示声源的移动的有无的检索用标志的情况下,即使在将来再看声音数据的情况下,也能够根据声源的移动的有无容易地检索,例如能够缩短验证时间。

[0234] 而且,本发明的一方式的声音处理装置可以构成为,在通过上述图像识别部对上述图像数据进行了图像识别的情况下,上述处理部将检索用标志记录于记录部,上述检索用标志用于从记录上述声音数据及上述图像数据的上述记录部检索包括上述预定声音的声音数据或包括上述预定声音的声源的图像数据。

[0235] 根据该结构,监视者即使在将来再看图像数据的情况下,也能够容易地检索所希望的图像数据或图像数据的特定的部位,例如能够缩短验证时间。

[0236] 而且,本发明的一方式的声音处理装置可以构成为,上述处理部使用记录于上述记录部的检索用标志所包含的预定检索用标志,与上述预定检索用标志对应地取得记录于上述记录部的上述声音数据或图像数据。

[0237] 根据该结构,例如,能够使用经由触摸面板而由使用者指定的检索用标志,检索发生了同样的现象的过去的图像数据。因此,能够促进与声音数据对应的图像数据的有效利用,能够提高监视者的便利性,也能够提高监视精度。

[0238] 而且,本发明的一方式的声音处理装置可以构成为,上述检索用标志包括上述声源的类别、上述声源的移动的有无及包括上述声源的缩略图像中的至少1个。

[0239] 根据该结构,能够记录各种信息作为与图像数据对应的检索用标志,在必要时迅速地检索所希望的图像数据。

[0240] 而且,本发明的一方式的声音处理装置可以构成为,在检测到上述预定声音的情况下,上述处理部根据上述图像识别部的图像识别结果,使提示部提示包括检测到上述预定声音的内容的警告信息。

[0241] 根据该结构,能够利用图像识别结果,例如,将产生了与异常相伴的声音的情况向监视者通知,催促强化监视等的注意。

[0242] 而且,本发明的一方式的声音处理装置可以构成为,在检测到上述预定声音的情况下,上述处理部根据上述图像识别部的图像识别结果,将包括上述预定声音的声音数据记录于记录部。

[0243] 根据该结构,在未检测到预定声音的时刻不记录声音数据,由此能够提高记录部的有效利用效率。而且,在检测到预定声音的情况下,利用图像识别结果,能够可靠地记录预定声音的产生时刻的声音数据,例如,将来能够作为验证材料进行确认。而且,在同样也记录图像数据的情况下,能够进一步提高监视精度。

[0244] 而且,本发明的一方式的声音处理装置也可以构成为,在检测到上述预定声音的情况下,上述处理部根据上述图像识别部的图像识别结果,对通过上述指向性处理部增强声音成分的方向及上述范围的大小中的至少一方进行变更。

[0245] 根据该结构,例如,在预定声音的声源移动并预想其移动的情况下,利用图像识别结果,以朝向预设的方向及大小的范围的方式变更指向性的方向及范围的大小,由此能够提高追踪预定声音的可能性。

[0246] 而且,本发明的一方式的声音处理系统具备:收音装置,包括使用多个麦克风从预定区域内的声源对声音数据进行收音的收音部;拍摄装置,包括对上述预定区域内的至少一部分的图像进行拍摄的拍摄部;及声音处理装置,对由上述收音部收音到的声音数据进行处理,上述声音处理装置具备:数据取得部,取得由上述收音部收音到的声音数据及由上述拍摄部拍摄到的图像数据;第一指定部,指定与基于上述图像数据而显示的图像上的指定部位对应的以上述收音部为基准的方向;第二指定部,指定与基于上述图像数据而显示的图像上的指定部位对应的上述预定区域内的预定范围;及指向性处理部,在所指定的上述预定范围内,增强所指定的上述方向的声音数据的声音成分。

[0247] 根据该结构,监视者指定监视区域的方向及范围的大小,在与监视区域对应的方

向及大小的范围内形成了指向性的状态下,能够监视异常的有无。由此,能够促进收音到的声音数据及图像数据的有效利用,提高监视者的便利性。

[0248] 而且,本发明的一方式的声音处理系统可以构成为,上述声音处理装置还具备:声音检测部,根据由上述指向性处理部增强后的上述方向及上述范围的声音成分检测预定声音;及处理部,在通过上述声音检测部检测到上述预定声音的情况下,进行预定处理。

[0249] 根据该结构,声音处理装置伴随着预定声音的检测而实施各种动作,因此能够促进收音到的声音数据及拍摄到的图像数据的有效利用,能够提高监视者的便利性。

[0250] 而且,本发明的一方式的声音处理系统可以构成为,上述数据取得部从上述收音装置取得上述声音数据,从上述拍摄装置取得上述图像数据,上述声音处理装置具备记录部,该记录部记录上述声音数据、上述图像数据及用于检索包括上述预定声音的声音数据的检索用标志。

[0251] 根据该结构,例如,根据实时的声音数据或图像数据,能够清晰地听取到监视区域的声音。因此,能够提高实时的监视精度。而且,例如,将实时的声音数据或图像数据与检索用标志一起记录,由此,将来能够容易地检索所希望的声音数据或图像数据,能够缩短验证时间。

[0252] 而且,本发明的一方式的声音处理系统可以构成为,具备记录数据的记录装置,上述记录装置具备记录部,该记录部将由上述收音部收音到的声音数据与由上述拍摄部拍摄到的图像数据建立对应地记录,并记录用于检索包括上述预定声音的声音数据的检索用标志,上述数据取得部从上述记录部取得上述声音数据、上述图像数据及上述检索用标志。

[0253] 根据该结构,能够清晰地听取过去的图像的监视区域的声音。因此,例如,能够提高使用了过去的图像的监视区域的声音解析的精度。而且,通过对记录的声音数据或图像数据附加检索用标志,将来能够使所希望的声音数据或图像数据的检索容易。而且,在已经记录检索用标志的情况下,能够使用检索用标志,容易地检索所希望的声音数据或图像数据。

[0254] 而且,本发明的一方式的声音处理方法是声音处理装置的声音处理方法,该声音处理方法包括以下步骤:取得由包括多个麦克风的收音部从预定区域内的声源收音到的声音数据及由拍摄上述预定区域内的至少一部分的图像的拍摄部拍摄到的图像数据;指定与基于上述图像数据而显示的图像上的指定部位对应的以上述收音部为基准的方向;指定与基于上述图像数据而显示的图像上的指定部位对应的上述预定区域内的预定范围;及在所指定的上述预定范围内,增强所指定的上述方向的上述声音数据的声音成分。

[0255] 根据该方法,监视者在指定监视区域的方向及大小的范围且在与监视区域对应的方向及大小的范围内形成了指向性的状态下,能够监视异常的有无。由此,能够促进收音到的声音数据及图像数据的有效利用,提高监视者的便利性。

[0256] 工业实用性

[0257] 本发明对于能够促进声音数据及图像数据的有效利用并提高便利性的声音处理装置、声音处理系统及声音处理方法等是有用的。

[0258] 附图标记说明

[0259] 10 阵列麦克风

[0260] 11A、11B、11C 麦克风

- [0261] 12A、12B、12C 放大器
- [0262] 13A、13B、13C A/D转换器
- [0263] 14A、14B、14C 声音编码器
- [0264] 15 网络处理部
- [0265] 20 相机
- [0266] 21 透镜
- [0267] 22 传感器
- [0268] 23 图像编码器
- [0269] 24 网络处理部
- [0270] 30、30B 监视控制装置
- [0271] 31 网络处理部
- [0272] 32 图像译码器
- [0273] 33 图像输出部
- [0274] 34 图像识别部
- [0275] 35 收音坐标指定部
- [0276] 36 声音译码器
- [0277] 37 指向性处理部
- [0278] 38 收音角度运算部
- [0279] 39 检测部
- [0280] 40 声源推定部
- [0281] 42 声音输出部
- [0282] 43 数据记录部
- [0283] 44 范围指定部
- [0284] 45 范围角度运算部
- [0285] 50 网络
- [0286] 60 操作者
- [0287] 61 监视器
- [0288] 62 触摸面板
- [0289] 63 扬声器
- [0290] 70 记录器
- [0291] 71 网络处理部
- [0292] 72 数据记录部
- [0293] 80 声源
- [0294] 90 收音单元
- [0295] 91 单元框体
- [0296] 100、100B 监视系统
- [0297] 101 顶棚面
- [0298] 102 地板面
- [0299] 121 框体入射面

- [0300] PA 声源
- [0301] P11、P12、P13 人物
- [0302] A1 收音范围
- [0303] $\theta 1$ 收音角度
- [0304] $\theta 2$ 、 $\theta 2a$ 、 $\theta 2b$ 范围角度

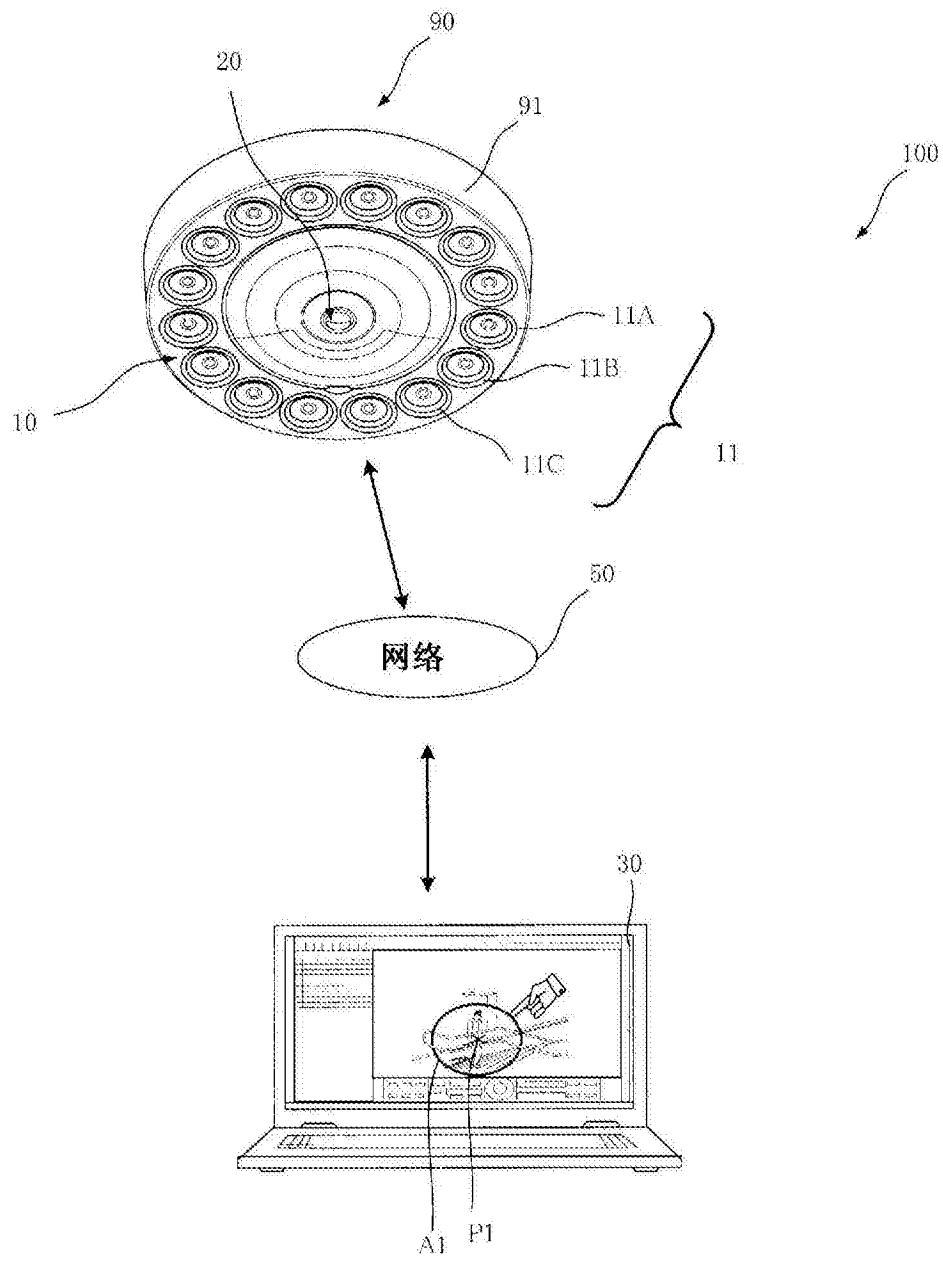


图1

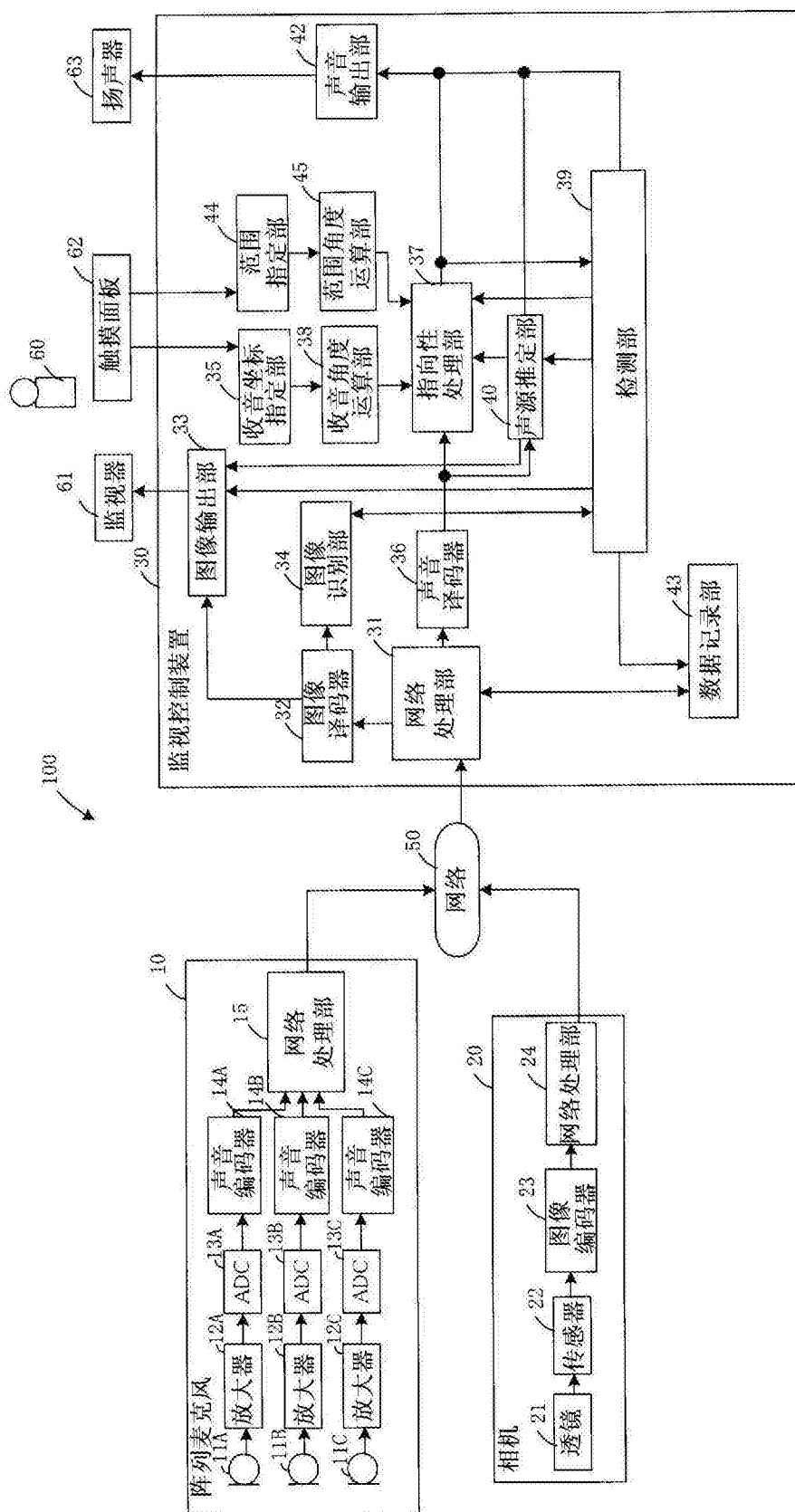


图2

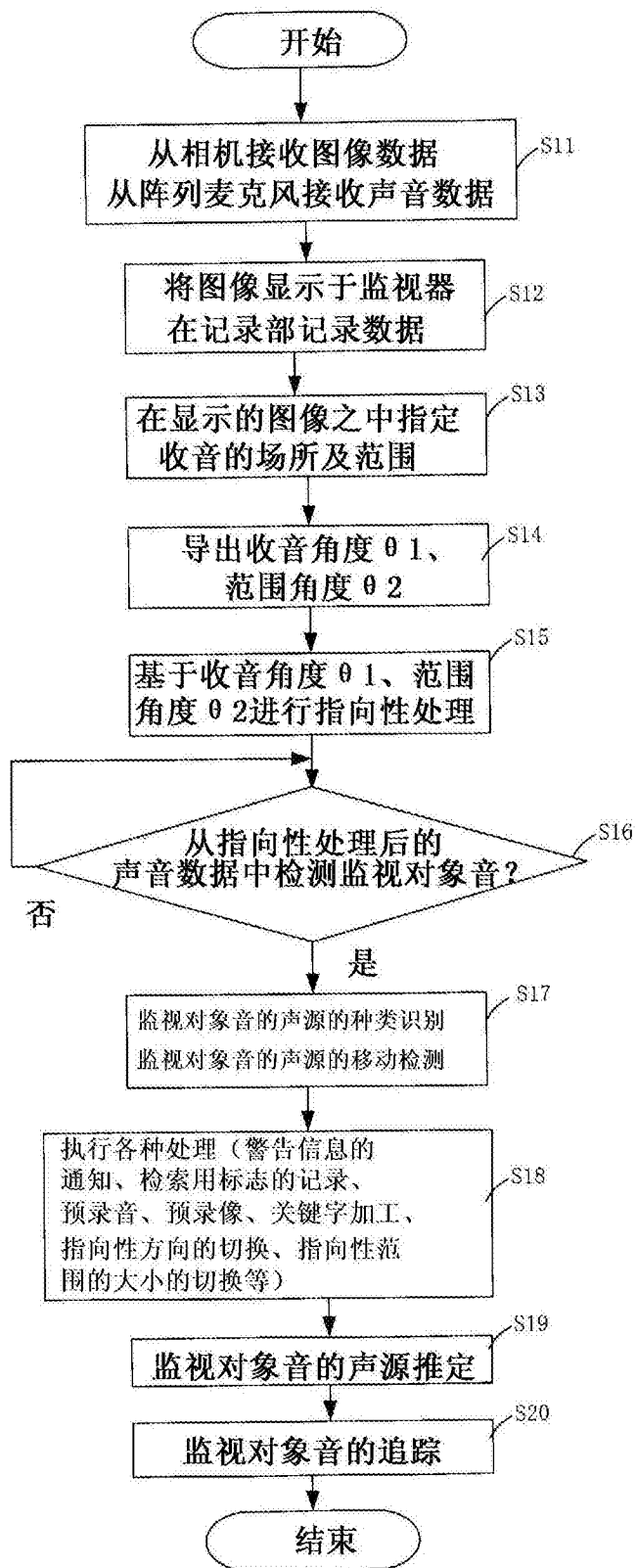


图4

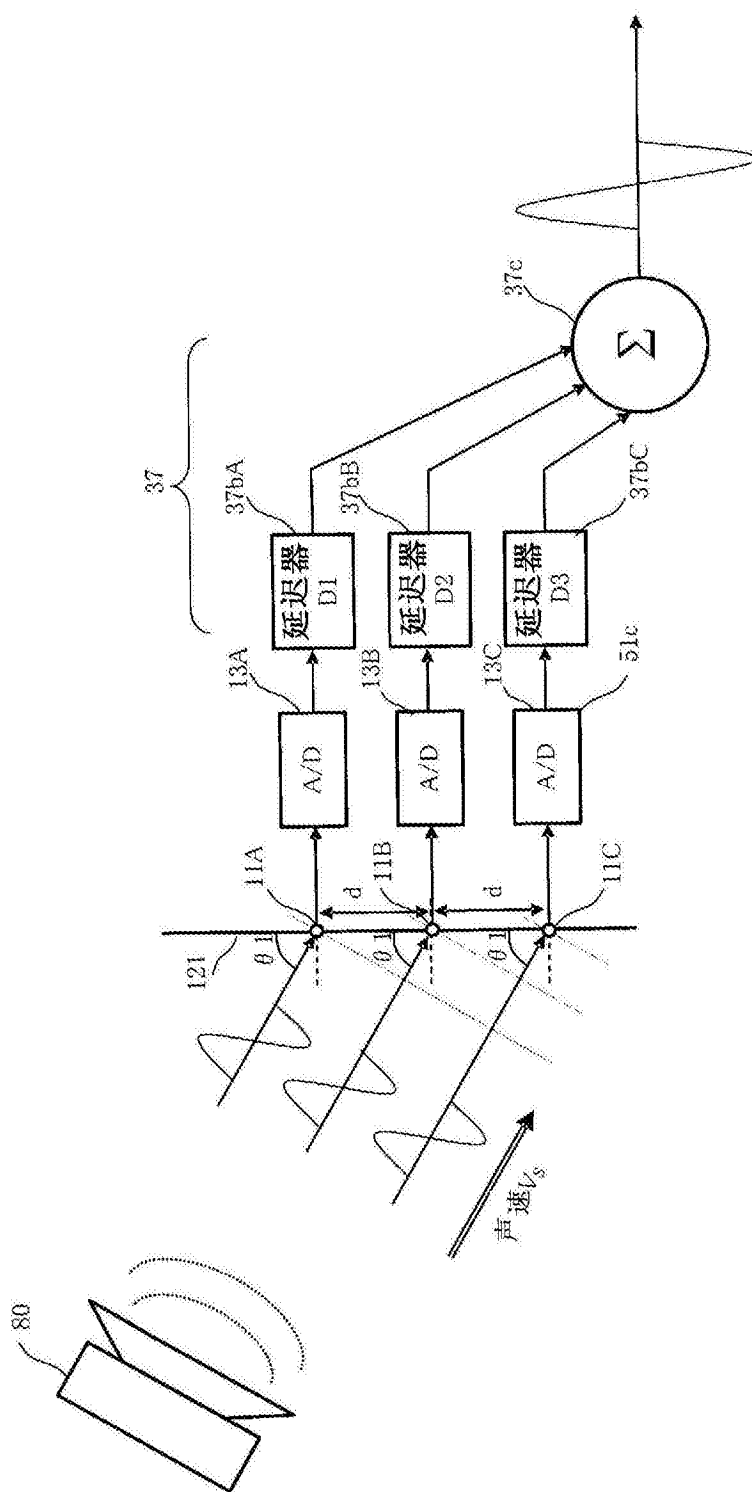


图5

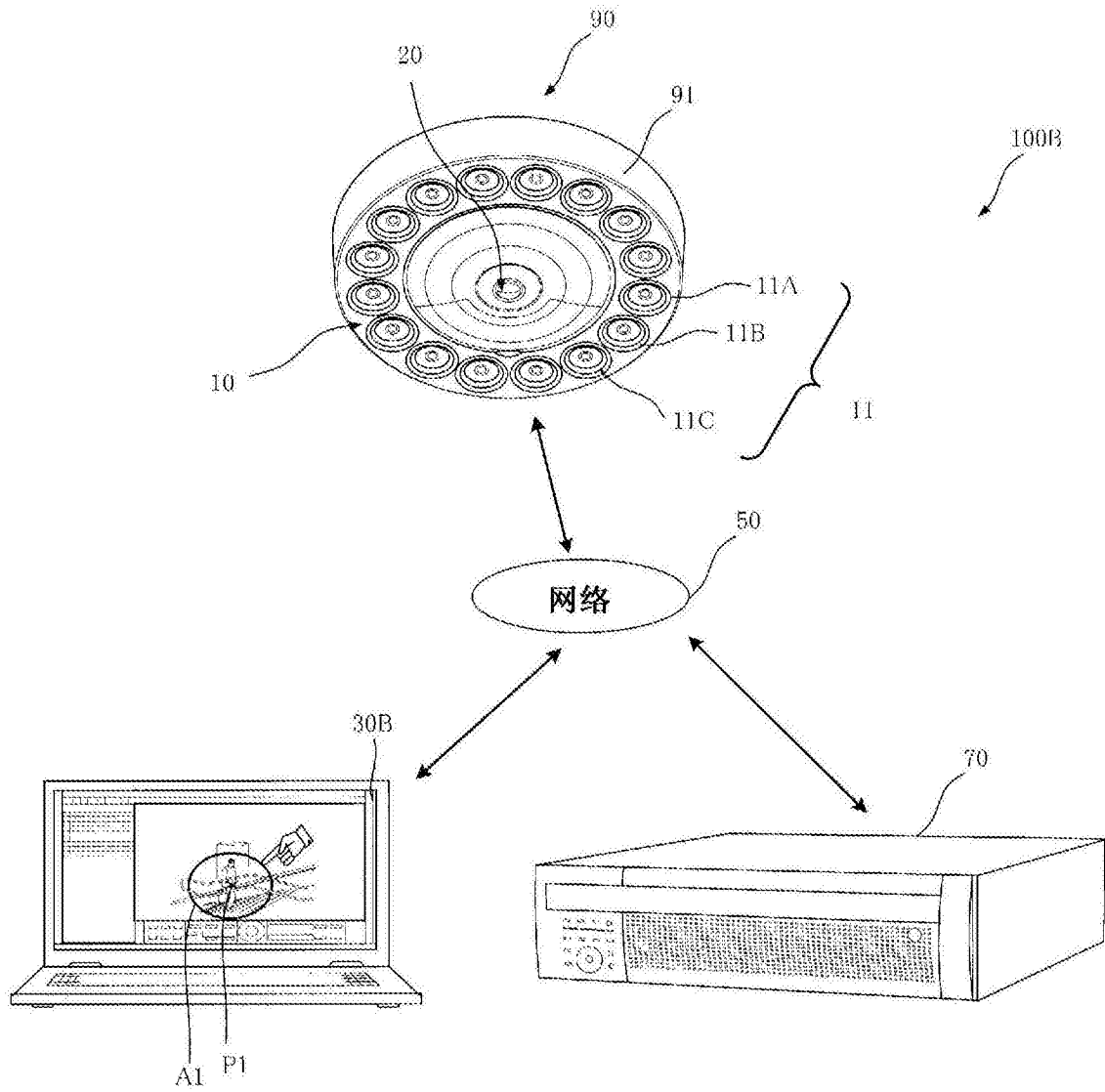


图6

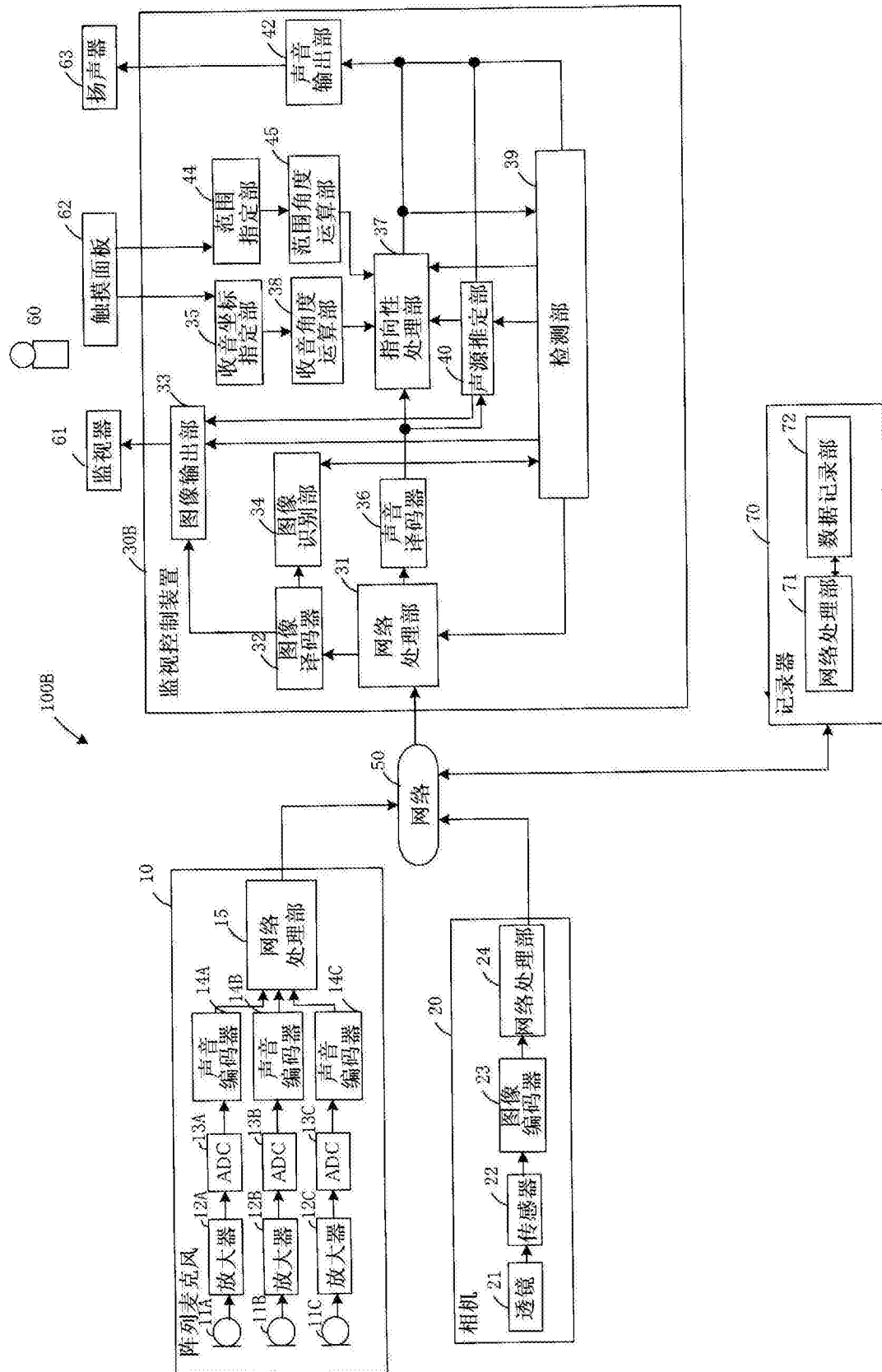


图7

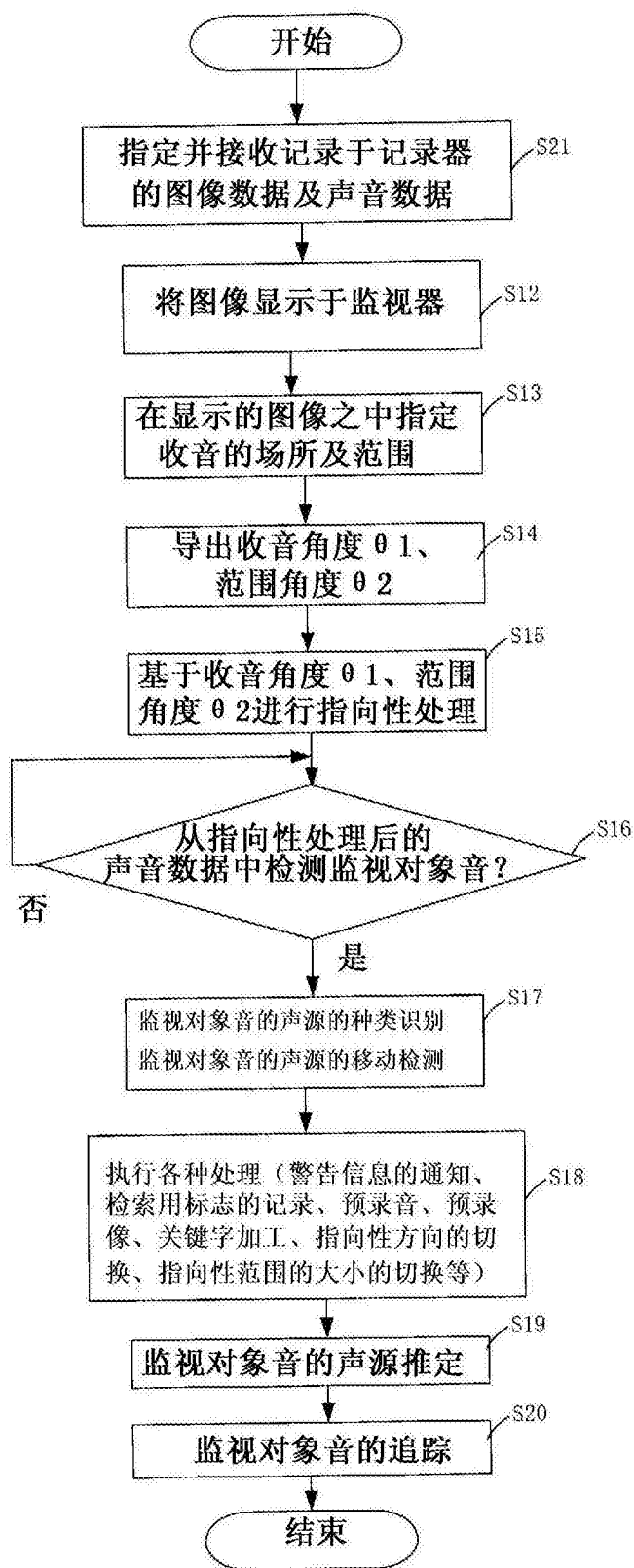


图8

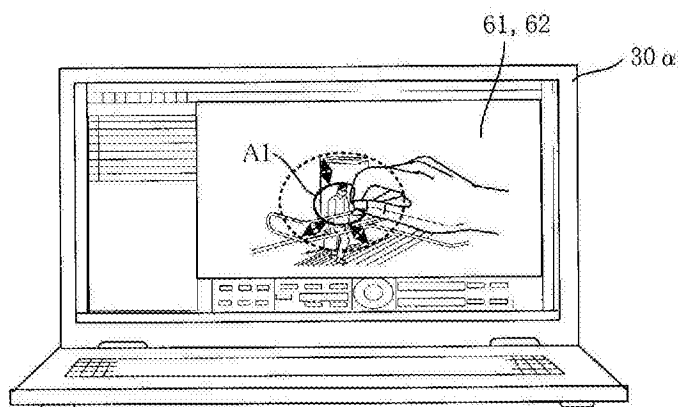


图9(A)

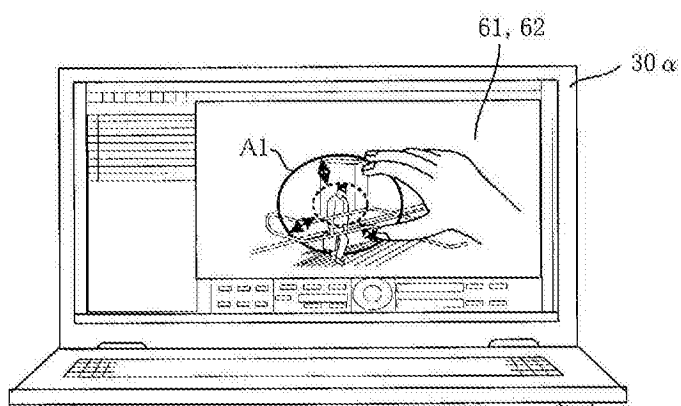


图9(B)

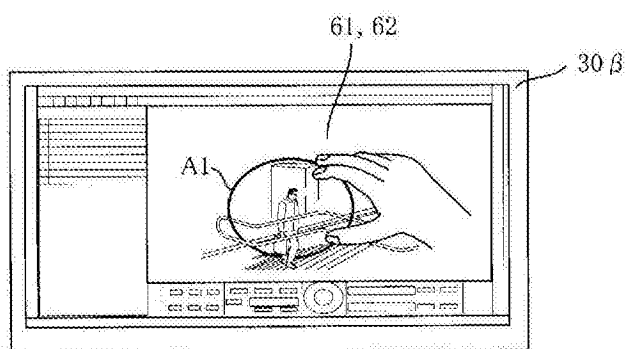


图9(C)

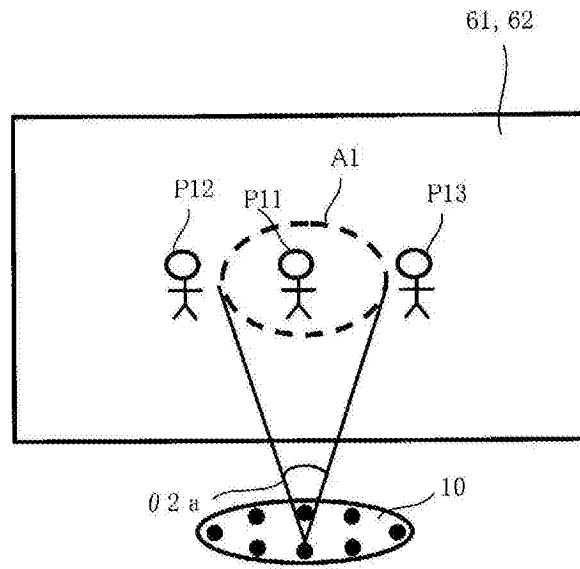


图10 (A)

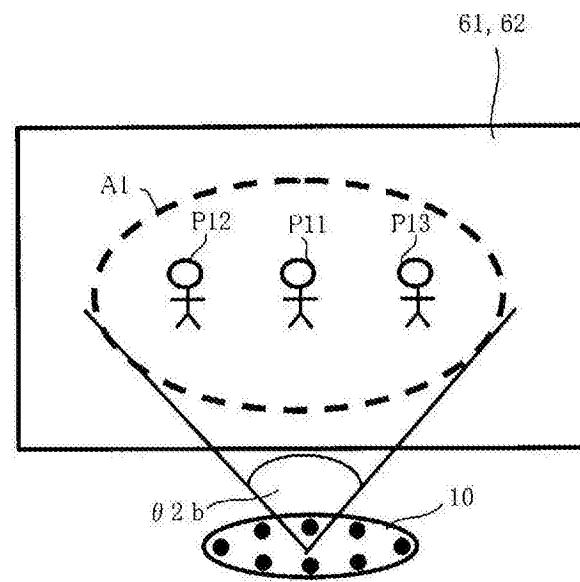


图10 (B)