[19] 中华人民共和国国家知识产权局

[51] Int. Cl[7]
H04L 12/16

[12] 发明专利说明书

[21] ZL 专利号 01133068.6

[45] 授权公告日 2005 年 5 月 4 日

[11] 授权公告号 CN 1200537C

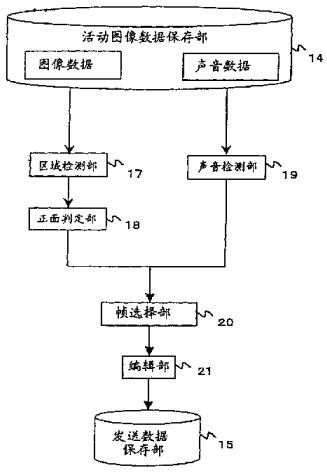
[22] 申请日 2001.9.12 [21] 申请号 01133068.6
[30] 优先权
[32] 2000. 9.12 [33] JP [31] 275991/2000
[71] 专利权人 松下电器产业株式会社
地址 日本国大阪府
[72] 发明人 今川和幸 高田雄二 松尾英明
岩佐克博 吉村哲也
审查员 徐 刚
[74] 专利代理机构 上海专利商标事务所有限公司
代理人 孙敬国

权利要求书 3 页 说明书 19 页 附图 20 页

[54] 发明名称 媒体编辑方法及其装置

[57] 摘要

本发明揭示一种媒体编辑方法及其装置，包括活动图像数据保存部(14)保存用户进行录像、录音所得的活动图像数据。区域提取部(17)从该数据提取包含用户的区域。正面判定部(18)检测所提取区域内用户是否朝向正面。声音检测部(19)检测活动图像数据中是否有规定电平以上的声音。帧选择部(20)根据正面判定部(18)和声音检测部(19)的输出结果，确定起始帧和结束帧。编辑部(21)根据确定的起始帧和结束帧，剪裁媒体，并进行图像变换的编辑。发送数据保存部(15)将编辑的媒体作为发送数据保存。所提供的媒体编辑装置能在便携终端的通信终端方便地编制包含消息的媒体。



ISSN1008-4274

1、一种媒体编辑方法，对包含将用户的一部分或全部作为被摄体拍摄的多幅图像组成的图像系列的媒体进行编辑，所述方法由媒体编辑装置执行，其特征在于，所述方法包含：

区域提取步骤，从所述图像提取包含用户的一部分或全部的区域；

脸朝向指定方向判定步骤，判定所述区域提取步骤中提取的区域包含的用户是否朝向预先规定的方向；

帧选择步骤，选择对应于往所述图像系列起始时刻到结束时刻的方向和结束时刻到起始时刻的方向，分别在所述脸朝向指定方向判定步骤首先判定为用户朝向所述指定方向的时刻之间的图像系列；

编辑步骤，对包含由所述帧选择步骤中选择的图像系列的媒体进行编辑。

2、如权利要求 1 所述的媒体编辑方法，其特征在于，

所述脸朝向指定方向判定步骤判定所述区域提取步骤中提取的区域包含的用户是否朝向正面。

3、如权利要求 1 或 2 所述的媒体编辑方法，其特征在于，

所述媒体包含声音，并且还包含检测该媒体所含声音的声音检测步骤；

所述帧选择步骤选择对应于往所述图像系列起始时刻到结束时刻的方向和结束时刻到起始时刻的方向，分别在所述脸朝向指定方判定步骤最先判定为用户朝向所述指定方向的时刻之间，而且对应于分别在所述声音检测步骤最先检测到声音的时刻之间的图像系列。

4、如权利要求 1 所述的媒体编辑方法，其特征在于，

所述编辑步骤利用元数据形式的描述规定在所述帧选择步骤选择的图像系列。

5、如权利要求 1 所述的媒体编辑方法，其特征在于，

所述编辑步骤从所述媒体剪裁图像系列，所述图像系列是从所述帧选择步骤中选择的。

6、如权利要求 1 所述的媒体编辑方法，其特征在于，

所述编辑步骤选择所述帧选择步骤中所选图像系列的第 1 图像作为初始显示用图像。

7、如权利要求 1 所述的媒体编辑方法，其特征在于，

所述编辑步骤根据所述区域提取步骤中所提取区域的位置和大小，计算所述图像系列中对应的部分区域，并且用所述部分区域进行编辑。

8、如权利要求 7 所述的媒体编辑方法，其特征在于，

所述编辑步骤利用元数据形式的描述规定所述部分区域。

9、如权利要求 1 所述的媒体编辑方法，其特征在于，

所述编辑步骤在设定为对所述区域提取步骤提取的区域完全不重叠或重叠区域较少的配置区域，配置所述媒体包含的文本。

10、如权利要求 9 所述的媒体编辑方法，其特征在于，

所述区域提取步骤提取包含用户脸部的脸部区域，

所述编辑步骤在设定为对所述提取步骤中提取的脸部区域完全不重叠或重叠区域较少的冒出区域配置所述媒体包含的文本。

11、如权利要求 1 所述的媒体编辑方法，其特征在于，

所述编辑步骤对 1 个以上的所述图像进行图像变换处理。

12、如权利要求 1 所述的媒体编辑方法，其特征在于，

所述区域提取步骤提取包含用户脸部的脸部区域；

所述编辑步骤参照根据所述区域提取步骤中提取的脸部区域算出的脸部特征，从存储多个人物图像及其对应脸部特征的人物数据库选择关联的 1 个以上所述人物图像。

13、如权利要求 12 所述的媒体编辑方法，其特征在于，

所述编辑步骤根据用户输入的文字数据，计算所述脸部特征。

14、如权利要求 12 所述的媒体编辑方法，其特征在于，

所述编辑步骤根据所述区域提取步骤中提取的脸部区域的脸长宽比和脸的部位特征，计算所述脸部特征。

15、一种媒体编辑装置，对包含将用户的一部分或全部作为被摄体拍摄的多幅图像组成的图像系列的媒体进行编辑，其特征在于，所述装置包含：

区域提取部，从所述图像提取包含用户的一部分或全部的区域；

脸朝向指定方向判定部，判定所述区域提取部中提取的区域包含的用户是否朝向预先规定的方向；

帧选择部，选择对应于往所述图像系列起始时刻到结束时刻的方向和结束时刻到起始时刻的方向，分别在所述脸朝向指定方向判定部首先判定为用户朝向所述指定方向的时刻之间的图像系列；以及

编辑部，对包含由所述帧选择部选择的图像系列的媒体进行编辑。

16、如权利要求 15 所述的媒体编辑装置，其特征在于，

所述媒体包含声音，并且还具有检测该媒体所含声音的声音检测部；

所述帧选择部选择对应于往所述图像系列起始时刻到结束时刻的方向和结束时刻到起始时刻的方向，分别在所述脸朝向指定方判定部最先判定为用户朝向所述指定方向的时刻之间，而且对应于分别在所述声音检测部最先检测到声音的时刻之间的图像系列。

17、如权利要求 15 或 16 所述的媒体编辑装置，其特征在于，

所述编辑部选择所述帧选择部选择的图像系列的第 1 图像作为初始显示用图像。

18、如权利要求 15 所述的媒体编辑装置，其特征在于，

所述编辑部根据所述区域提取部中所提取区域的位置和大小，计算所述图像系列中对应的部分区域，并且用所述部分区域进行编辑

19、如权利要求 15 所述的媒体编辑装置，其特征在于，

所述编辑部在设定为对所述区域提取部提取的区域完全不重叠或重叠区域较少的配置区域，配置所述媒体包含的文本。

20、如权利要求 15 所述的媒体编辑装置，其特征在于，

所述区域提取部提取包含用户脸部的脸部区域；

所述编辑部参照根据所述区域提取部中提取的脸部区域算出的脸部特征，从存储多个人物图像及其对应脸部特征的人物数据库选择关联的 1 个以上所述人物图像。

媒体编辑方法及其装置

技术领域

本发明涉及包含静止图像或活动图像的媒体的编辑方法及其装置，特别涉及可视电话、电视会议、可视函件、可视交谈、门电话(内部对讲电话)等，至少采用图像进行通信时，编辑包含有关图像的媒体用的媒体编辑方法及其装置。

背景技术

以往已提出的装置在数字光盘、数字镜物照像机、录像带、半导体存储器等中记录并重放会议、演讲、采访、会面、使用电话和可视电话等的会话等的声音和图像、来自电视机和监视摄像机等的图像等。与记录者通过笔记仅记录应记录信息的要点的方法相比，这种装置具有可用声音和图像的形式无遗漏地记录信息的优点，因而近年来广泛采用。

近来，随着广泛利用宽带通信，以个人为主，用可视电话、门电话或带摄像机的便携终端的信息设备处理声音和图像以进行通信的场所不断增多。例如，以文本相互通信的已有电子函件正在改变为采用声音和活动图像的可视函件等。随着可视电话的普及，用声音的录音电话(录音秘书电话)正在改变为采用图像和声音的记录电话等，同时用声音和活动图像进行通信的场所不断增多。

这里，例如采用可视函件和活动图像的记录电话那样，暂存消息等数据用作媒体时，通常经历以下的步骤：

(1) 首先按压进行录像的设备的录像启动键。

(2) 其次，对要传送的消息进行录像；

(3) 最后，按压结束键。

本发明各实施形态中，媒体是指静止图像和活动图像等构成的通信用消息(或其数据)。

通过通信线路将存储的消息送到别处时，通常经历以下的步骤：

(4) 确定递送所存储消息中的起止处，仅剪裁该部分递送。

如可视函件那样，对朋友等寄送消息时，有时也经历以下的步骤：

(5) 进行对消息施加花纹等装饰，增添人物等图画，部分剪裁图像等的媒体编辑。

上述步骤中，在步骤(4) 确定实际递送消息起止处(即剪裁的起点和终点时)，用户必须重放存储的消息，一面进行确认，一面确定剪裁的起点和终点。但是，使用带摄像机的便携终端的录音电话时，难以进行甚至不可能进行上述那样的媒体编辑作业。

因此，最好在执行上述步骤(3)的时刻，不再次重放消息，而确定要递送的范围，并加以剪裁。作为这种方法，日本特开平 6-343146 号公告揭示的方法是消息记录中进行用户输入，并从进行该输入的时刻开始，仅重放一定时间的信号。但是利用该方法，作为消息，只能重放记录者进行用户输入的时刻所决定的一定时间内的声音和图像，超过该一定时间的消息，完全不能重放。而且，仅由用户输入的时刻决定剪裁范围，因而要求慎重操作，同时操作本身也麻烦。

因此，最好不象上述已有例那样要求用户输入，而自动检测符合预先设定的规定条件的范围，并剪裁活动图像，作为这种方法，日本特开平 9-294239 号公告揭示的方法是在输入话音信息和图像信息中检测符合预先设定的规定条件的范围，并进行剪裁。而且，揭示了该规定条件为有无规定电平以上的声音信号、图像亮度变化和色分布变化和摄像机(图像)的动态。

发明内容

然而，在采用几乎单人面对摄像机进行消息录像之类的通常可视函件和活动图像的录音电话的情况下，用上述已有例的方法，存在以下问题。

首先，将有无声音信号作为条件对活动图像进行，不适合门电话和可视电话中的记录。其原因在于，尽管想知道谁来访问(或谁打来电话)也完全没有对声音消息录音时，不可能剪裁。将图像亮度变化和摄像机(图像)的动态作为条件也不合适。这是因为通常的可视函件和采用活动图像的记录电话的情况下，图像亮度变化和图像活动小的居多，很难剪裁适当的范围。

上述已有例中，检测符合预先设定的规定条件的全部范围并进行剪裁，但在可视函件和采用活动图像的记录电话的情况下，在消息中途进行多次剪裁，则分解为多个消息。例如将有无声音信号作为条件时，一件消息中途暂时无声

或重新启动消息时，消息分解为多个。然而，在电子函件和录音电话的情况下，希望这些构成一件消息。既使假设组合上述多个消息，作为一个消息，也包含不自然的间断。

在如上述步骤(5)那样，可视函件等中又进行装饰又编辑显示画面的情况下，有关操作通常繁杂，因此，当前可视函件装饰和编辑不普遍。例如，象娱乐中心等可看到的娱乐用自动摄像装置那样，都停留在对静止图像施加墙纸等装饰的形态和对文本电子函件加静止的人物图像的形态。尤其是使用市售便携电话终端时，希望仅用拇指操作，因而装饰和编辑操作容易变回更繁杂。然而，为了在便携终端的小显示画面内适当配置需要的信息，媒体编辑处理尤其重要。可是，适合便携终端的简单媒体编辑方法，不存在已有技术例。

因此，本发明要解决的技术问题是提供一种媒体编辑方法，能对包含几乎单一人物面对摄像机录制的活动图像的媒体进行适当且简单编辑。

本发明为解决前述的技术问题，具有以下所述的特征。

本发明是一种媒体编辑方法，对包含将用户的一部分或全部作为被摄体拍摄的多幅图像组成的图像系列的媒体进行编辑，该方法包含以下的步骤。区域提取步骤，从图像提取包含用户的一部分或全部的区域。脸朝向指定方向判定步骤，判定区域提取步骤中提取的区域包含的用户是否朝向预先指定的方向。帧选择步骤，选择对应于往图像系列起始时刻到结束时刻的方向和结束时刻到起始时刻的方向，分别在脸朝向指定方向判定步骤首先判断为用户朝向指定方向的时刻之间的图像系列。编辑步骤，对包含由帧选择步骤选择的图像系列的媒体进行编辑。

脸朝向指定方向判定步骤也可判定用户是否朝向正面。媒体还包含声音，并且还可具有检测该媒体所含声音的声音检测步骤。于是，帧选择步骤可选择对应于往图像系列起始时刻到结束时刻的方向和结束时刻到起始时刻的方向，分别在脸朝向指定方向判定步骤最先判定为用户朝向指定方向的时刻之间，而且对应于分别在声音检测步骤最先检测到声音的时刻之间的图像系列。

编辑步骤可利用无数据形式描述规定在帧选择步骤选择的图像系列，也可从媒体剪裁，还可将所选图像系列的第 1 图像作为初始显示用图像选择。而且，编辑步骤可根据区域提取步骤所提取区域位置和大小，算出图像系列中对应的部分区域，并用部分区域进行编辑，还可利用元数据形式的描述规定该部分区域。

编辑步骤在设定为对区域提取步骤提取的区域或脸部区域完全不重叠或重叠区域较少的配置区域或冒出区域，配置媒体包含的文本。

编辑步骤可对 1 个以上的图像进行图像变换处理，也可参照根据区域提取步骤提取的脸区域算出的脸特征，以存储多个人物图像和对应的脸特征的人物数据库选择相关的 1 个以上人物图像。

本发明还提供一种媒体编辑装置，对包含将用户的一部分或全部作为被摄体拍摄的多幅图像组成的图像系列的媒体进行编辑，其特征在于，所述装置包含：区域提取部，从所述图像提取包含用户的一部分或全部的区域；脸朝向指定方向判定部，判定所述区域提取部中提取的区域包含的用户是否朝向预先规定的方向；帧选择部，选择对应于往所述图像系列起始时刻到结束时刻的方向和结束时刻到起始时刻的方向，分别在所述脸朝向指定方向判定部首先判定为用户朝向所述指定方向的时刻之间的图像系列；以及编辑部，对包含由所述帧选择部选择的图像系列的媒体进行编辑。

通过与附图一起阅读以下的详细说明，会明白本发明的其他目的、特征和优点。

附图说明

图 1 为示出能进行实现本发明媒体编辑方法的图像通信的媒体编辑终端硬件组成框图。

图 2 为示出进行本发明媒体编辑时的信息流程和处理过程的框图。

图 3 为示出实施例形态 1 所涉及媒体编辑装置功能结构的框图。

图 4 为说明对某活动图像数据进行剪裁处理用的图。

图 5 为根据 MPEG-7 规范记述图 4 的索引信息的元数示例图。

图 6 为接收包含活动图像数据，发送者信息，件名等信息的可视函件的终端中画面显示例示意图。

图 7 为示出实施形态 2 记所涉及媒体编辑装置功能结构的框图。

图 8 为修饰处理例和显示画面例的示意图。

图 9 为部分区域元数据记述例示意图。

图 10 为活动图像不另外有显示文件名和文本的空间的画面显示例示意图。

图 11 为对与包含用户的图像区域不重叠的区域配置件名的显示例。

图 12 为对与包含用户的图像区域重叠程度小的区域配置件名的显示例。

图 13 为对活动图像写入文本的布局处理所涉及的元数据描述例示意图。

图 14 为加入人物的可视函件在接收终端的显示画面例示意图。

图 15 为示出实施形态 4 所涉及媒体编辑装置功能结构的框图。

图 16 为采用头发特征的脸特征量示意说明图。

图 17 为选择所用人物的编辑画面示例图。

图 18 为接收人物画件时的接收终端画面示例图。

图 19 为另一接收人物画件时的接收终端画面示例图。

具体实施方式

各实施形态基本结构

下面，参照附图概括说明本发明的各实施形态。

本发明的媒体编辑方法及其装置在使用可视电话、便携终端或门电话之类以个人为主或面向家用的图像通信终端，用户编制可视函件，或访问者留下记录时，能提供使用方便的接口。

图 1 为示出进行实现本发明媒体编辑方法的图像通信的媒体编辑终端硬件组成的框图。图 1 中，本媒体编辑终端具有对用户进行信息输入输出的输入部 1、图像输入部 2、图像显示部 3、声输入部 4、声输出部 5、处理输入输出信息的摄像控制部 6、声输入输出控制部 7、显示控制部 8、通信部 9、记录部 10、记录控制部 11、信号处理部 12 和控制部 13。这些部分通过系统总线和外部总线相互连接，而且为普通计算机结构或类似于计算机结构。

输入部 1 由键盘(包含 10 个数字键等)和鼠标等构成，输入使用者编辑媒体所需要的各种信息和文本形式的文章等。

图像输入部 2 由 CCD 摄像机等构成，输入由至少 1 幅以上图像组成的活动图像。图像输入部 2 基本上与后面说明的声输入部 4 联动，通常取得对录制消息的用户拍摄的活动图像。

声输入部 4 由话筒等构成，用于输入装置外部的声音。声输入部 4 基本上与图像输入部 2 联动，通常取得对消息进行录音的用户的声音。

图像显示部 3 由液晶显示器等构成，对用户显示其记录的活动图像和文学、别人发送来的活动图像和文学，以及操作上需要的各种信息等。

声输出部 5 由扬声器等构成，对用户输出其记录的声音、别人发送来的声音，以及操作上需要的告警音和振铃音等。

摄像控制部6进行图像输入部2的通/断控制和曝光控制等对图像输入部2的各种控制。声输入输出部7进行声输入部4和声输出部5的通/断等各种控制。显示控制部8对图像显示部3进行控制。

通信部9通过公用电话线路和无线等通信路由，与对端的信息处理装置进行后文所述实施形态编制的各种数据的收发。通信部9可用任何通信模式，例如可以是可视电话那样的同步通信，也可以先电子函件那样的非同步通信。

记录部10由存储器和硬盘等记录媒体构成，至少记录图像输入部2和声输入部4输入的数据。记录部10也可包含CD-ROM和DVD等记录媒体及其驱动装置。记录控制部11对记录部10进行输入输出控制。

信号处理部12由数字信号处理器等构成，对图像输入部2输入的图像信号、声输入部4输入的声音信号和记录部10记录的数据至少进行后所述各实施形态中编辑需要的各种处理。

控制部13由微计算机或CPU等构成，控制上述数据的流程，并进行各种处理。

本媒体编辑终端可以是将上述全部组成部分收装在1个壳体中的一体型装置，也可以是分散型装置，各组成部分通过网络和信号线收发数据。例如，带摄像机的便携电话终端为全部收装到1个壳体的一体型装置。门电话为了与访问者接口，至少将图像输入部2、声输入部4和声输出部5设置在门的外侧附近，其他组成部分则收装到设置在起居室等处的另一壳体中，因而是分散型装置。此外，还可考虑后文所述那样的外部具有人物数据库的分散型装置。

下面，概括说明用户采用本发明媒体编辑方法编制发送用数据时全部步骤的流程。图2为示出进行本发明媒体编制时的信息流程和处理步骤。首先，用户对上述图1的输入部1输入命令，以便开始取得视频和音频数据。然后，用户通过图像输入部2和声输入部4输入消息等，编制活动图像数据。

编制活动图像数据时，多数的情况下，通常起始区间和末尾区间包含与用户要输入的信息无关的区间。因此，进行省略活动图像数据始端和末端非所需区间的剪裁处理。下面详细说明。

其次，进行布局处理，使对发送数据的对端有用的信息(例如数据编制的日期和时间、编制者名称)与制成的活动图像一起，能用一幅画面显示。例如，进行修饰处理，从剪裁所得的活动图像数据剪切出以发消息的用户为中心的特定区域等，并对修剪所得区域进一步配置作为素材数据生成的文本和人物图像

等。素材数据是指图像、文本、计算机图形等对活动图像数据加入的全部数据。例如，素材数据可以是预先制作的图像图案，也可以是文字和符号的图案。该布局处理常用称为多目的互联网消息(函件)扩展(MIME:Multi-Purpose Internet Message(mail) Extensions)、超文本置标语言、可扩充置标语言(XML: Xtensible Markup Language)、MPEG-7 的元数据形式描述。经过以上的处理，最后制成发送数据，作为给数据发送对方的消息。

在图 1 的信号处理部 12、控制部 13、记录控制部 11 进行上述剪载处理和布局处理。这些处理通常用计算机可执行的程序实现。该程序从图中未示出的 CD-ROM 和半导体存储器卡等计算机可读取的记录媒体输入到记录部 10 等，或者通过通信线路下载。

可是，进行以上那样的处理时，如上文所述，利用已有的方法，用户必须输入信息的地方非常多。较详细而言，剪裁处理中，用户必须一面校验活动图像和声音，一面输入要剪裁的范围。而且，在布局处理中，用户必须进行编辑如何布局用的操作。尤其是布局处理的修饰中，用户必须对活动图像的全部帧指定要剪出的区域，加入素材数据时，用户需要一面确认活动图像数据中被摄体的位置和大小等，一面输入应添加处。

本发明各实施形态的媒体编辑装置和方法在信号处理部 12 和控制部 13 的一方或双方，进行后文所述的区域提取处理，正面判定处理和声音检测处理。利用该处理，对剪裁处理和布局处理的全部或部分，可提供灵巧的接口。

例如，在便携终端编制可视函件时，拍摄消息后希望立即发出函件的需求增多。这种情况下，通过提供灵巧的接口，即使立即发出函件时，用户也能用简单的操作(例如操作 1 个按键)编制可视函件，其编制方法是预先剪裁消息的始端和末端，而且修剪成用户驱于中心，并且根据情况配置墙纸和冒出框，从而形成接收者也容易观看的布局。上述处理，本发明各实施形态中未必全部包含，可根据用途仅组合采用需要的部分。下面，单独详述发明各实施形态。

实施形态 1

实施形态 1 的媒体编辑装置可自动进行上述图 2 的剪裁处理。图 3 为示出实施形态 1 所涉及媒体编辑装置功能结构的框图。图 3 中，本媒体编辑装置具有活动图像数据保存部 14，发送数据保存部 15，区域提取部 17，正面判定部 18、声音检测部 19，帧选择部 20 和编辑部 21。

活动图像数据保存部 14 对应于图 1 记录部 10，用户保存录像、录音所得的活动图像数据作为消息。区域提取部 17 从保存的活动图像数据提取包含被摄体(用户)的一部分或全部的特定区域。正面判定部 18 检测区域提取部 17 所提取区域内的用户是否朝正面。声音检测部 19 检测活动图像数据的各帧中是否有在规定电平以上的声音信号。帧选择部 20 根据正面判定部 18 和声音检测部 19 的输出结果，决定起始帧和结束帧。编辑部 21 根据帧选择部 20 确定的起始帧和结束帧，剪裁媒体，对该媒体进行图像变换处理等。发送数据保存部 15 对应于图 1 的记录部 10，保存编辑后的媒体，作为发送数据。适当发送保存的数据。

下面，说明这些组成部分的运作。图 4 用于说明对某活动图像数据进行的剪裁处理。图 4 中，对活动图像数据保存部 14 保存的活动图像数据进行剪裁处理。活动图像数据包含具有用户录时的消息的声音数据和与该数据同步录像的图像数据。可将图像数据和声音数据组合构成为 1 个数据，也可由图像数据、声音数据和记述其同步关系的数据分别构成。图 4 所示的数据例是用门电话等取得的消息的典型例，包含图中 A~H 的 8 个地方(时刻)。在这些时刻发生以下的情况：

A、开始取得数据

B、第 1 消息开始

C、第 1 消息结束

D、下一消息前正在考虑的区间

E、第 2 消息开始

F、声音停顿(进行换气等)

G、第 2 消息结束

H、数据获取结束

图中的曲线将横轴取为经过时间，纵轴取为输入声音电平，A~H 表示规定时刻。曲线上方添加的画是声音同时录像的几乎相同时刻的图像数据的 1 个场面，例如表示从注入门电话消息到该消息失期间对用户拍摄的图像。

参阅图 4 可知，编制的活动图像数据，多数情况下，其起始部分和末尾部分包含与用户想输入的信息时，存在开始交谈消息前的时间和谈话结束后的时间。着眼于包含这种消息的活动图像数据的特点，本媒体编辑装置利用以下方法对图像数据和声音数据自动确定应剪裁的区间。

首先，说明检测剪裁起始时刻的方法。检测起始时刻时，首先，区域检测部 17 按开始到结束的方向(即图 4 中从 A 观看 H 的方向)，对图像数据的各帧依次检测包含被摄体(用户)的一部分或全部的区域。

以往在提取被摄体区域方面，有各种方法。例如，特开平 5-91407 号公告揭示的方法是利用比较活动图像相邻帧间的视频信号，将活动量小的视像部分作为背景部分，其他的部分作为被摄体区域提取。特开平 5-161131 号公告揭示的方法是预先保持仅有背景的图像，各像素分别运算仅有背景的图像与活动图像各帧的差，将该差小的区域作为背景区域，并提取差大的区域作为被摄体区域。关注的区域为被摄体的头部和脸部时，可通过例如检测图像中的椭圆区域测出头部。关于这种方法，谷内等人者的“基于采用视点固定型全景倾斜度变焦距摄像机的自适应观看法模型的人物头部检测跟踪”(MIRU2000 图像识别理解讨论会演讲论文集 I(PP-9~14)中已揭示。其他还有利用彩色信息检测脸面的方法，侧重于眼、口等脸一部分的方法、基于模板匹配的方法等公知的种种方法。本区域提取部 17 用以上的已有方法能方便地提取被摄体区域。

接着，正面判定部 18 对区域提取部 17 提取的区域检测该区域存在的用户是否朝正面。以往在检测人物正面方面，有种种方法。例如，仅对正面图像预先准备模板，利用模板匹配。可检测是否正面。马场等人著的“头部分类器：人物脸图像实时分类”(第 7 次图像检测讨论会演讲论文集 PP-411~416)中揭示的方法用一种统计特征识别法，即支持矢量机(SVN：Support Vector Machine)，识别图像中脸的朝向。正面判定部 18 用以上的已有方法能判定人物是否朝下面。而且，用以上的方法，也可设置脸朝指定方向判定部代替正面判定部 18，检测图像区域中的用户是否朝预定方向(例如对正面往右偏斜 45 度的方向）。这样，可选择用朝斜方的用户脸部等预先决定的美观角度的图像。

声音检测部 19 将是否存在规定电平以上的声音信号作为条件，检测是否存在声音。例如，对环境噪声电平与输入和声音平均电平的比较决定该规定电平。还可用公知的声音识别等方法检测是否存在人的声音。

帧选择部 20 号参照正面判定部 18 和声音检测部 19 的判定结果，往起始时刻到结束时刻的方向校验各帧，将首先满足条件的时刻的方向校验各帧，将首先满足条件的时刻对应的帧作为起始帧。又往结束时刻到起始时刻的方向检验各帧，同样将最先满足条件的时刻对应的帧作为结束帧。根据正面判定部的判定结果，则图 4 所示图像中用户最先朝正面的时刻为 B 点，最后朝正面的时

刻为 G 点。于是，帧选择部 20 选择共同满足条件的帧，从而起始帧为 B 点的帧，结束帧为 G 点的帧。

如上文那样，本媒体编辑装置往起始时刻到结束时刻的方向和结束时刻到起始时刻的方向，分别进行扫描，选择最先满足条件的时刻，不切断 D 时刻附近存在的下一消息前的过渡期，就能对全部消息进行剪裁。因此，适合如可视函件等那样，将用户输入的数据汇总为 1 件发送的用途。

本媒体编辑装置进行图像中被摄体正面判定和声音检测两种处理，从而能可靠地剪裁作为消息录像的部分。例如，用户随意朝正面沉思时，也能根据该消息开始的时刻可靠地剪裁。然而，本媒体编辑装置中，也可省略声音检测。即，用户开始谈消息时，通常正面朝终端(的摄像机)用户开始谈消息时，通常正面朝终端(的摄像机)行走，判定为正面特别有效。用户注入消息前的考虑过程中无意间发出声音时，往往声音检测无效。因此，即使省略声音检测部 19，本媒体编辑装置也具有与上文所述大致相同的效果。

接着，编辑部 21 根据帧选择部 20 判定的起始帧和结束帧，对媒体(活动图像数据)进行剪裁。这里，编辑部 21 可编制完全删除剪裁范围以外的区间的活动图像数据，也可编制不删除活动图像数据而将剪裁范围作为系列包含在内的之数据。编制元数据时，在未剪裁的部分包含重要信息的情况下，由于存在数据，可利用该部分。下面，以采用 MPEG-7 格式时为例，说明元数据格式。

作为描述元数据的格式，存在种种标准规范或独立规范，但作为最新标准规范，有例如 MPEG-7 规范。MPEG-7 是称为多媒体内容描述接口(Multimedia Contemt Description Interface)(ISO/IEC15938)的活动图像专家组(MPEG:Moving Picture Experts Group: ISO/IEC JTCI/SC29/WG11)制订的 MPEG-1、MPEG-2、MPEG-4 后续的第 4 个标准规范。该标准规范规定描述多媒体信息内容用的框架，是以用于数字库检索和编辑应用程序为目的的规范。MPEG-7 还规定描述以图像、声音为中心的多媒体信息的内容用的描述符(Descripter)的标准集。对内容提供该描述，从而可根据多媒体信息内容进行检索。本规范的实际描述定义语言是对 XML Schema 语言(XML 图解语言)进行必要扩充的语言。但是，该扩充也保持与 XML Schema 语言的语法有互换性。

以上那样的 MPEG－7 中，为了描述内容的特征，可组合以下的基本要素。

(1)描述符(Descriptor：简称为 D)

描述符号描述多媒体内容中某单一特征用的基本工具。MPEG－7 中规定描

述符的语法(Symtax)和语义(Semantics)。

(2)描述图(Description Schema：简称为 DS)

描述图 4 规定多个描述工具间的结构和语义关系的图文框。MPEG－7 中，同样也规定描述图的语法和语义。作为构成描述图的描述工具，除描述符外，还包含别的描述图本身。

(3)描述定义语言(Description Definition Language)

描述定义语言是规定描述符和描述图的语法用的语言。MPEG－7 中，由将 W3C(World Wide Wed Consortium：万维网论坛)加以标准化的图解语言“XLM Schema”作为基础，追加描述多媒体内容特征时需要的各种数据型，从而规定了描述定义语言。作为扫描图(DS)的例子，有作为活动图像数据某部分的指示符数“Video Segnent DS”和描述图像内的区域的“StillRegion DS”等。作为描述符(D)的例子，有描述媒体格式的“MdiaFormat”等。MPEG－7 中，除标准规定的描述符和描述图外，还规定对新描述符和描述图进行定义或扩充用的语言(Description Definition Language)。因此，本媒体编辑装置中采用元数据时，根据上述语言描述元数据，就成为 MPEG－7 格式的描述。

图 5 为根据 MPEG－7 规范描述图 4 的索引信息的元数据示例图。图 5 中，该元数据关联的图像帧为对图 4 的 B 和 G 时刻采用“Vidro Segment DS”而形成的对象。这里，与实际图像帧的关联，在“Media Time”的“Media Time Point”(“媒体时间”的“媒体时间点”)进行，并描述相应 Video Segment(图像段)的时刻。例如，图 5 所示那样的“T13:20:01:1 F15”的描述，其含义为“13 时 20 分 01 秒的第 1 帧(但每秒 15 帧中存在 0 帧到 14 帧)”。利用这样的元数据描述，可进行仅选择活动图像中特定区间的重放控制。

以上的描述毕竟是一个例子，不限定描述方式。即，只要用与内容关联的方式描述就可以，因而也可以按上述以外的各种方式的元数据进行描述。元数据的描述为了在附图上表示并说明方便，特地用文本格式示出，但不受该格式限制，也可以是二进制格式。

如上文那样，不删除活动图像数据，而制成剪裁范围作为索引包含在内的元数据时，与编制完全删除剪裁范围以外的区间的活动图像数据时相比，想自动剪裁后进行再修改时，仅再编辑元数据即可，能自由编辑。

可是，上文中，为了进行自动剪裁，利用帧选择部 20 输出的起始帧和结束帧，但帧选择部 20 输出的起始帧在接收编辑媒体方的终端画面也可定义为

首先显示的图像。因此，本媒体编辑装置的剪裁方法进一步有用。即，设想将接收中的活动图像或保存的活动图像首先以静止图像(例如预映画面或简略画面等)示出的情况。又将这时间的图像定义为初始显示用图像。这时，在接收终端的画面显示活动图像的第 1 帧图像，例如图 4 中 A 点的图像，然而 A 点的图像，其发送源用户朝横向放映，不适合用作预映或简略介绍等的初始显示图像。因此，如图 5 那样，用元数据将起始帧定义为上述初始显示用图像。利用这种方法，本媒体编辑装置不对接收终端重新发送初始显示用静止图像，就能使放映适合作为初始显示用图像的朝正面的用户的 B 点所对应图像帧最先显示。在重新发送初始显示用的静止图像时，用区域提取部 17 和正面判定部 18 往起始时刻到结束时刻的方向进行扫描，检测出 B 点。将该 B 点作为初始显示用图像发送，则从接收着可看到朝正面的用户图像。

编辑部 21 可对该初始用显示图像或活动图像的帧进行图像变换处理，例如进行析像度变换处理。这样，就能配合接收终端侧显示的彩色进行显示。

由上文可知，根据本媒体编辑装置由于判定用户是否朝正面，能可靠地自动剪裁相对端的消息部分，并且接收者可仅重放需要的部分。作为初始显示用图像，也可方便地设定合适的图像。

实施形态 2

实施形态 2 的媒体编辑装置自动进行上述修饰处理，从而在接收终端画面小的情况下，也能有效编制容易观看的布局。

首先，说明本实施形态设想的地方。利用可视函件等给对端发送媒体时，该媒体通常不仅包含相应的活动图像数据，而且包含活动图像发送者的信息和件名等信息。图 6 为接收这种可视函件的终端的画面显示例，如图 6 所示，显示画面 100 上不权显示对端发来的活动图像画面叫,而且显示包含发送者姓名、接收方、件名等的首部画面 101、文本的文章画面 102、涉及适当装饰的装饰画面 103。

用便携终端那样的窄小显示画面重放图 6 的画面时，一般将总体缩小后显示。然而，在窄小的显示画面内进行进一步将总体缩小的显示时，难以观看文本，而且活动图像数据中的对端脸部也变小，不合适。

作为消息输入的活动图像数据，往往采用广角摄像机取得，用户可不必很介意照像机与被摄体的位置关系。因此，如图 6 的活动图像画面 104 那样，包

含很多用户图像以外的背景区。进行总体缩小的显示，则用户的脸部进一步减小，观看困难，不合适。

因此，本实施形态的媒体编辑装置为了进行显示画面仅显示包含用户的部分区域的布局处理，至少具有以下的组成部分。进行实施形态 1 中所述的剪裁处理时，可增加对应的图 3 的组成部分，因而省图其说明。

图 7 为显出实施形态 2 所涉及媒体编辑装置功能结构的框图。图 7 中，本媒体编辑装置具有活动图像数据保存部 14、发送数据保存部 15、区域提取部 17、布局编制部分 22 和素材数据保存部 23。这些组成部分在图 2 中执行所述布局处理的一部分或全部。

本媒体编辑装置具有与实施形态 1 的媒体编辑装置共同的结构和运作，因而共同的组成部分标准相同的符号，省略其说明。本实施形态中，声音数据不是特别需要，因而活动图像数据保存部 14 可保存与实施形态 1 相同的活动图像数据，也可保存不包含声音的图像数据。

图 7 中，素材数据保存部 23 对应于图 1 的记录部 10，保存图 6 所示的文本和装饰用图像数据等素材数据。布局编制部 22 根据用户的操作，从素材数据 23 适当读出素材数据，进行包含修饰处理的布局处理。后文详述。

图 8 为修饰处理例和显示画面例的示意图。图 8 中，其上部示出图 6 所示对方发来的活动图像画面 104。由于以上说明的原因，该图像为包含较多用户区域以外的背景区的图像。因此，如下文所述那样仅对用户区域进行修饰，编制布局。

首先，区域提取部 17 从活动图像数据保存部 14 所保存活动图像的各帧图像提取包含用户的一部分(例如脸部)或全部的区域。区域提取部 17 的这种动作用以上所述的方法能方便地实现。该提取区域可为任何形状。

其次，布局编制部 22 根据区域提取部 17 提取的区域，计算活动图像数据中要显示的部分区域。图 8 中，用活动图像画面 104 内的粗线方框示出部分区域。

布局编制部 22 进而组合该部分区域对应的图像和用户预先设定的素材数据(文本和图像等)，编制布局。图 8 中，编制成的显示画面 200 除部分区域所对应的活动图像 204 外，和图 6 相同，还显示首部画面 201、文章画面 202 和装饰画面 203。这样，在决定布局时，活动图像数据自动形成小的部分区域的图像，从而在小的显示画面也能容易到显示。

这里，布局编制部 22 通常制成决定活动图像数据和素材数据的布局的元数据的布局的元数据。因此，即使对布局编制部 22 设定的部分区域，同样是元数据方式的，容易处理，较方便。

图 9 为这种部分区域的元数据描述例示意图。图 9 的例子与实施形态 1 相同，也用 MPEG－7 格式进行描述。这里，描述为各帧用实施形态 1 所述的“Video Segment DS”，并且用“Still Region DS”设定分别位于各帧的部分区域。这种情况下，关于区域指定信息，则用“Contour Shape”描述作为部分区域的矩形形状(图中顶点数为 4)和矩形座标(图中未示出)。

这样采用元数据时，不像从活动图像数据切出部分区域编制新活动图像数据时那样，活动数据容量变小。然而，接收终端的用户可设定为根据爱好切换部分区域和全部区域进行显示，或者显示其他部分区，从而能配合接收终端的画面大小和接收者的爱好，自由改变布局。当然，这种情况下，对布局编制部 22 设定的部分区域，进行初始设定由于最先显示包含消息发送者的区域，较方便。

利用 MPEG－7，则不仅是图 9 那样每帧分别设定“StillRegion DS”的方法，而且也可采用活动区域信息“Moving Region DS”和声像综合区域信息“Audio VisualRegion DS”等。作为包含这些信息的基本定义，有表示多媒体内容的一部分的“Segment DS”的定义，但如果是基于该定义的 DS，则可用数量较小的描述进行与图 9 所示描述相同的描述。

如上文那样，本实施形态的媒体编辑装置解决定部分显示区域，因而即使在带摄像机的便携终端那样显示画面小的情况下，也能从容易观看仅有被摄体的区域的方式在画面内进行显示。在按照元数据描述决定布局时，即使在带摄像机的便携终端和 PC 终端那样显示画面大小不同的情况下，接收者也能以最佳方式在画面上进行显示。

实施形态 3

实施形态 3 的媒体编辑装置与实施形态 2 的修饰处理不同，可编制使活动图像显示区域尽量多而且也显示所需文本的布局。

首先，说明本实施形态设想的地方。例如考虑将上述图 6 的显示画面 100 修饰成活动图像画面 104 的部分尽量大，并且在小画面例如便携电话的显示画面)进行显示的情况。这里，假设该小画面上希望显示“件名”、“文本”和

活动图像，作为最低限度的信息。然而，实际上小画面只能显示活动图像的全部区域，除活动图像外，不存在显示件名和文件的空间。图 10 为这样只能显示活去图像部分的画面显示例的示意图。

因此，本媒体编辑装置结构与实施形态 2 的媒体编辑装置相同，同时区域提取部 17 和布局编制部 22 的运作加加以改变，以便显示上述文本信息。即，布局编制部 22 配置“件名”和“文本”等文本信息，使其对区域提取部 17 检测的包含用户的图像区域不重叠或至少重叠程度小。下面详细说明该运作。

首先，区域提取部 17 从活动图像数据检测包含用户的图像区域，并计算该区域的位置和大小。其次，布局编制部 22 接收区域提取部 17 算出的该区域的位置和大小、素材数据保存部 23 保存的“件名”和“文本”等素材数据。进而，布局编制部 22 对不与上述图像区域重叠(或重叠程度小)的区域设定配置素材数据的区域。图 11 为对与包含用户的图像区域不重叠的区域配置件名的显示例，如图 11 所示文件的件名配置在用户头部上方的空间，不重叠。因此，可编制不拘活动图像显示区域大，也能进行所需文件显示的布局。

布局编制部 22 也可将配置素材数据的区域设定为任意形态。该区域定义为冒出区域。常冒出区域带边框，并具有背景色(例如白色)。布局编制部 22 在该冒出区域内写入素材数据的内容，即文本。图 12 为对与包含用户的图像区域重叠程度小的区域配置文本的显示例。如图 12 所示，文本配置在用户左侧的空间，大致不重叠。因此，可编制拘活动图像显示区域大，也能显示所需文本的布局。

图 12 所示冒出区域的形状，如温画等中常用的那样，用户口部附近具有成锐角的凸出部分，并且参照文本的字数，对不与上述图像区域重叠(或重叠程度小)的区域设定冒出区域。

最好上述布局所得画面作为接收者终端的第 1 图像(上述初始显示用图像)在画面上显示。即，接收者打开收到的函件时，最先显示图 11 和图 12 那样的画面。然后，接收者仅确认件名和文本内容。这里，在 1 员不能容纳完文本内容等情况下，进行例如滚动处理等。这样，接收者仅在第 1 显示画面进行文本等的确认，活动图像重放过程中则不进行。然而，当然也可构成在活动图像重放中，以添加文本和件名等的方式进行重放，从而能一面听实际的活动图像消息，一面进行确认。

显示的文件内容不限于件名和文本。而且，可首先显示图 11 那样的画面，

接着根据接收者的操作依次显示图 12 那样的画面；也可以同时合成这些画面进行显示。这样，对与包含用户的图像区域不重叠(或重叠程度小)的区域进行配置，则可形成任何结构。

如上文那样，本媒体编辑装置即使在活动图像和文本混合的显示画面也能编制成接收者不混淆而且容易通过观看确认内容的布局。通过采用冒出框的方式，可形成宛如图像中的用户正在谈文件内容那样的画面，从而激活与通信对端之间的通信。

其次，最好布局编制部 22 与第 1 和第 2 实施形态一样，编制决定布局的元数据，以便进行上述布局处理，即对活动图像写入文本的处理。

图 13 为有关这种布局处理的元数据描述例示意图。图 13 的例子与实施形态 1 和 2 相同，用 MPEG－7 格式进行描述，按照“Media Duration”的值，即按照由媒体规定点表示的长度，以叠加的形式表示夹在“Text”标记中间的字面含义。这样，利用元数据进行描述，可显示文本，而不实际进行将文本嵌入活动图像的处理。

实施形态 4

实施形态 4 的媒体编辑装置用区域提取部 17 提取脸部区域选择该脸面区域对应的人物图像，从而扩大消息表现变化式样，用户能表现快乐，可促进通信顺利。

首先说明本实施形态设想的地方。如上文所述，利用图像的函件通过增加装饰，可制成提高欢乐性的函件。尤其如娱乐中心的自动摄像装置那样，那本人的图像添加漫画、三维图画等构成的人物，更具有提高欢乐性和便于亲近感的效果。图 14 为这样添加人物的可视函件接收终端显示画面例的示意图。如图 14 所示，显示画面 400 上添加包含发送者姓名、接收方、件名等的首部画面 401、文本的文章画面 402，对方发来的活动图像画面 404，而且和人物画面 403 一起示出。以上那样构成可视函件的接收画面，则可以说接收者增加对该函件的亲近感。

用户选择这种人物时，添加与活动图像的图像和内容无关的人物相比，要求希望选择与图像有关的人物的居多。本实施形态的媒体编辑装置在显示画面的图像是脸部图像时，布局处理中选择与该脸部图像对应的人物。下文中，将上术述那样增添人物的函件称为“人物函件”。

图 15 为示出实施形态 4 所涉及媒体编辑装置功能结构的框图。本媒体编辑装置具有活动图像数据保存部 14、发送数据保存部 15、区域提取部 17、正面判定部 18、编辑部 26、人物选择部 24、人物数据库 25。这些组成部分执行图 2 中所由局处理的一部分或全部。

本媒体编辑装置具有与实施形态 1 的媒体编辑装置共同的结构和运作，因而共同的组成部分标注可同的符号，省略其说明。但是，本实施形态中，声音数据非特别需要，因而活动图像数据保存部 14 可保存与实施形态 1 相同的活动图像数据，也可保存不包含声音的图像数据。进行实施形态 1 中所述的剪裁处理时可增加对应的图 3 的组成部分，因而省略其说明。与上文所述相同，正面判定部 18 也可以是脸朝指定向判定部。

正面，说明本媒体编辑装置的运作。区域提取部 17 和正面定部 18，其运作与实施形态 1 时相同，判定活动图像中的图像是否用户正面图像，并将判定结果输入编辑部 26。编辑部 26 将判定为正面图像的图像输入到人物选择部 24。人物选择部 24 根据该图像从各种人物数据库化后加以保存的人物数据库 25 选择 1 个或多个后补人物，并将所选人物对应的人物 ID 输入到编辑部 26。

这里，人物选择部 24 从人物数据库 25 保存数多个人物选择后补人物。本实施形态中，提取所述正面画像中脸部的特征，选择人物数据库 25 保存的 1 个或多个人物。

也就是说，人物数据库 25 预先登记 2 维人物图像数据和用计算机图形法制成的 3 维人物结构数据等人物图，该人物对应的脸部特征和相应的人物 ID，作为人物信息。人物选择部 24 参照从编辑部输入的正面图像中的脸部特征，从人物数据库 25 选择具有与该脸部特征一致或类似的脸部特征的 1 个以上人物图像。作为上述脸部特征，可考虑脸部大小、脸部长宽比、脸的部位特征等特征量。作为脸的部位特征，可考虑例如眼鼻口大小、眼鼻口位置关系、毛发量和颜色等特征量。是否戴眼镜等也可用作脸部特征。

进一步详细说明脸部特征量。图 16 为采用头发特征的脸部特征量的示例说明图。图 16 示出 6 个不同用户的摄影图像、相应的图像处理结果、相应的特征量和相应的特征表现。摄影图像与照片图像相同，但为了便于看图，画成肖像画。作为脸部特征，示出特征量和特征表现，但仅为其中一方也可，还可用其他值或表现。

图 16 中，处理结果以黑头发为前提，示出进行提取摄影图像黑区的处理

的结查。当然，即使头发为黑色以外的颜色，通过提取了发色同色的区或也同样有进行处理。作为特征量，示出归一化面积和归一化周长两个例子。归一化面积是指用脸部面积除头发面积进行归一化后的值。归一化周长是指用脸部周长除发部周长进行归一化的值，发量为与一般发量平均值比较进行分类所得的类型表现，具体而言，利用上述归一化周长，比平均值小时，发型表现为短发；比平均值大时，发型表现为长发。这样利用提取的脸部特征量或其特征表现，可从人物数据库选择类似的 1 个以上人物图像。

作为上述脸部特征量提取方法，以往已揭示种种技术。例如，山口等人著“脸部朝向和表情变化方面单靠的脸部识别系统：Smartface“(电子信息通信学会论文志 Vol. J84-D-11-No.6)揭示的方法用部分空间检测脸部区域，并利用隔离度滤波器检测眼、鼻、口等脸的部位。本媒体编辑装置通过利用至少一种公知的技术，能方便地提取脸部特征量。

为了参照提取的脸部特征量，选择人物数据库 25 登记的各后补人物，可用上述特征表现，也可计算与登记的脸部特征量的相关值，这里，超过计算设定为适合作后补的提了门限值的相关值时，提取该人物作为后补。人物选择部 24 将提取的后补人物后对应的人物 ID 通知编辑部 26。

编辑部 26 根据通知的人物 ID，对用户显示作为后补的人物图像。图 17 为选择所用人物的编辑画面示意图。图 17 中，示出 3 个后补人物，箭号为表示用户要选择的人物的光标。当然，可以不用光标，而用使要选择的人物图像翻转，或者加粗边框等方法。

在图 17 那样的编辑画面的基础上，用户选将使用的人物。编辑部 26 对生成描述所选人物 ID 的元数据的媒体进行编辑，从而产生发送数据。当然，也可将人物图像本身编入发送数据。这样产生的发送数据保存在发送数据保存部 15，并以适当的定时按地址发送给接收者。

图 18 为接收上述那样产生的发送数据时接收终端画面的示意图。如图 18 所示，在接收终端画面的下侧，显示作为发送者的用户选择的人物。在例如重放活动图像消息时，可不显示该人物，而在该显示位置显示活动图像。这样的布局，可由编辑部编制，也可在接收端设定。

成为后补的人物可以是 1 个，这时不进行后补选择，可较简单方便地编制函件。

编辑部 26 也可对人选择部 24 通知表示用户输入的特征量的字符串(上述

特征表现等)。例如，用户输入“毛发量多，短发”的字符串。这样，人物选择部 24 就将其以人物数据库 25 所保存图 16 所示那样特征表现核对，从而能选择后补人物。这样通过用字符串表示特征量，本媒体编辑装置能选择进一步接近用户意图的人物的后补人物。

如上所述，本媒体编辑装置不仅是全部组成部分收装在一个壳体的一体型装置，而且可以是各组成部分通过网络和通信线路进行数据收发的分散型装置。因此，可考虑例如人物选择部 24 和人物数据库 25 可通过网络接入且拉于与本媒体编辑装置不同的地方的情况。图 20 为示出这种分散型媒体编装置或系统的结构的框图。

图 20 中，这种分散型媒体编辑装置包含具有实施形态 1－3 所涉及媒体编辑装置的部分或全部功能的人物函件编辑终端 501、设置在与人物函件编辑终 501 不同的地方的人物选择部 724 和人物数据库 725。这些组成部分通过网络 600 相互连接。这种分散型媒体编辑装置也具有与上述一体型媒体编辑装置相同的结构以进行运行，因而可得相同的效果。图 20 的分散型媒体编辑装置，其人物选择部 724 和人物数据库 725 也可由人物函件编辑终端以外的终端(例如对函件进行接收的人物函件接收终端 502 等使用。即，在进行接收时，接收人物函件所含人物 ID 的人物函件接收终端 502 可仅接收来自人物数据库 725 的人物图像，因而终端本身不需要保持许多数据。即使在人物函件接收终端 502 回送信息时作为媒体编辑装置运作的情况下，也能共用人物选择部 724 和人物数据库 725。

这样，分散型媒体编辑装置，其人物选择部 724 和人物数据库 725 可由多个用户共用，因而各终端不需要内设这些组成部分，而且能利用保存许多人物的数据库。

总之，利用本媒体编辑装置，则用户可用从所拍摄活动图像提取的正面图像，从登记的许多人物缩小后补人物的范围，方便地选择喜欢的人物，编制人物函件，借助采用这种人物函件，能激活与通信对端之间的通信。

发明联系当前认为最现实且最佳的实施例进行了说明，但本发明不受所揭示实施例的限制。

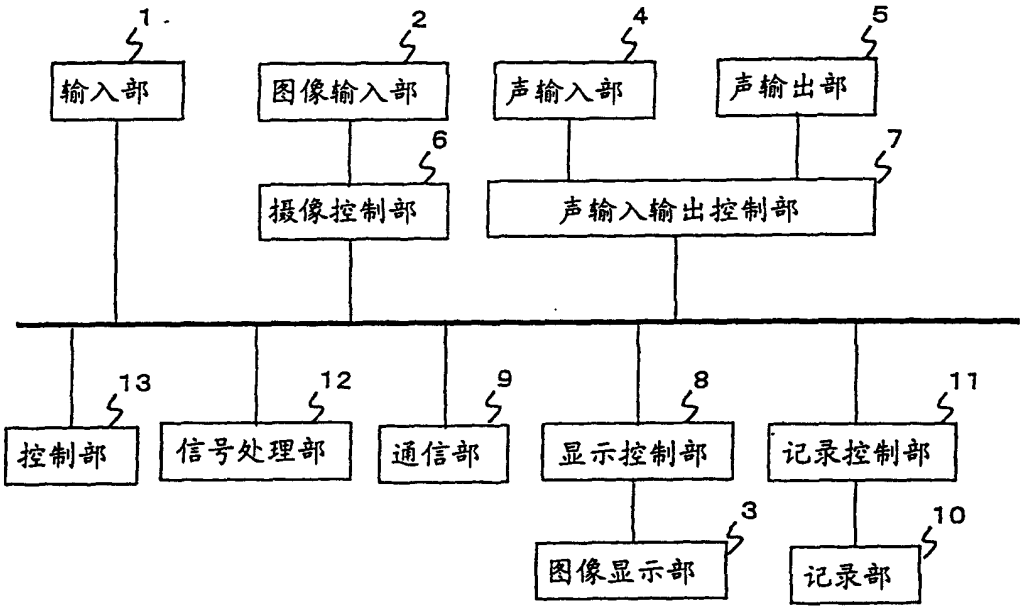

1
输入部
2
图像输入部
4
声输入部
5
声输出部
6
摄像控制部
7
声输入输出控制部
13
控制部
12
信号处理部
9
通信部
8
显示控制部
11
记录控制部
3
图像显示部
10
记录部

图 1

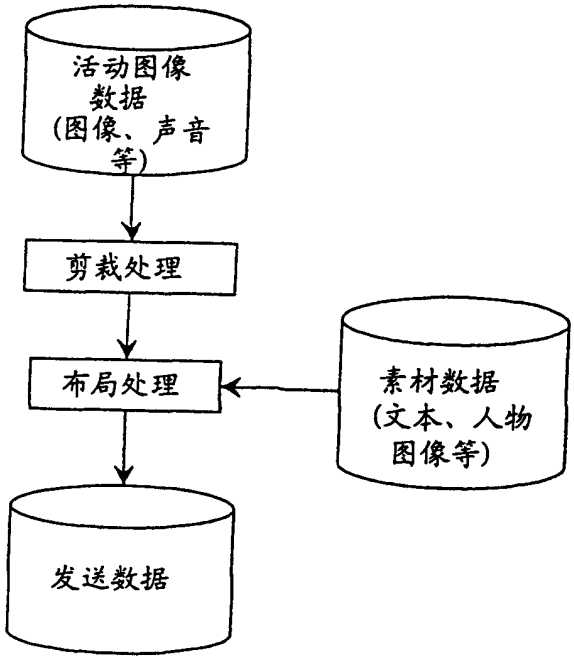


图 2

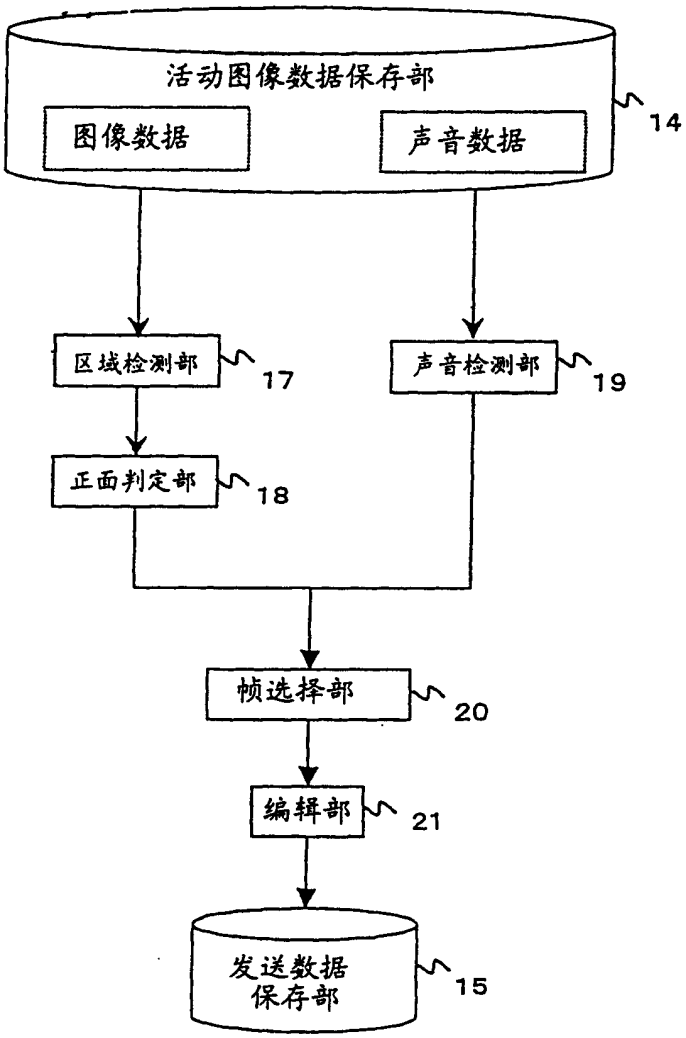

活动图像数据保存部
图像数据
声音数据
14
区域检测部
17
声音检测部
19
正面判定部
18
帧选择部
20
编辑部
21
发送数据
保存部
15

图 3

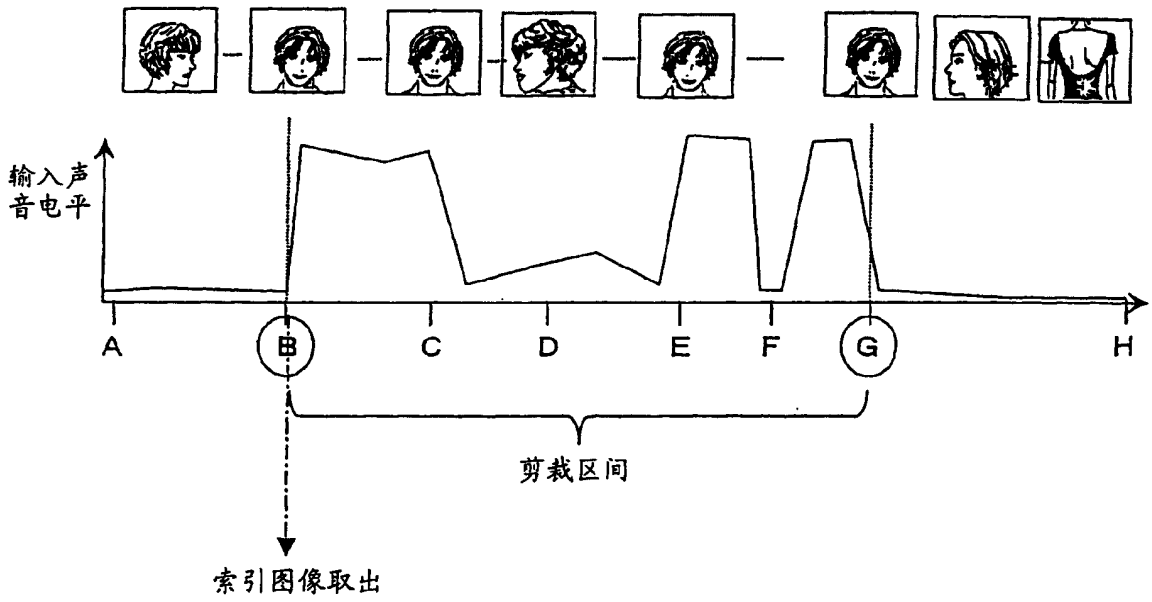


图 4

```
...
<VideoSegment id=" 起点 ">
        <MediaTime>
                <MediaTimePoint>T13:20:01:1F15</MediaTimePoint>
        </MediaTime>
</VideoSegment>
...
<VideoSegment id=" 终点 ">
        <MediaTime>
                <MediaTimePoint> T13:23:20:15F15 </MediaTimePoint>
        </MediaTime>
</VideoSegment>
```

图 5

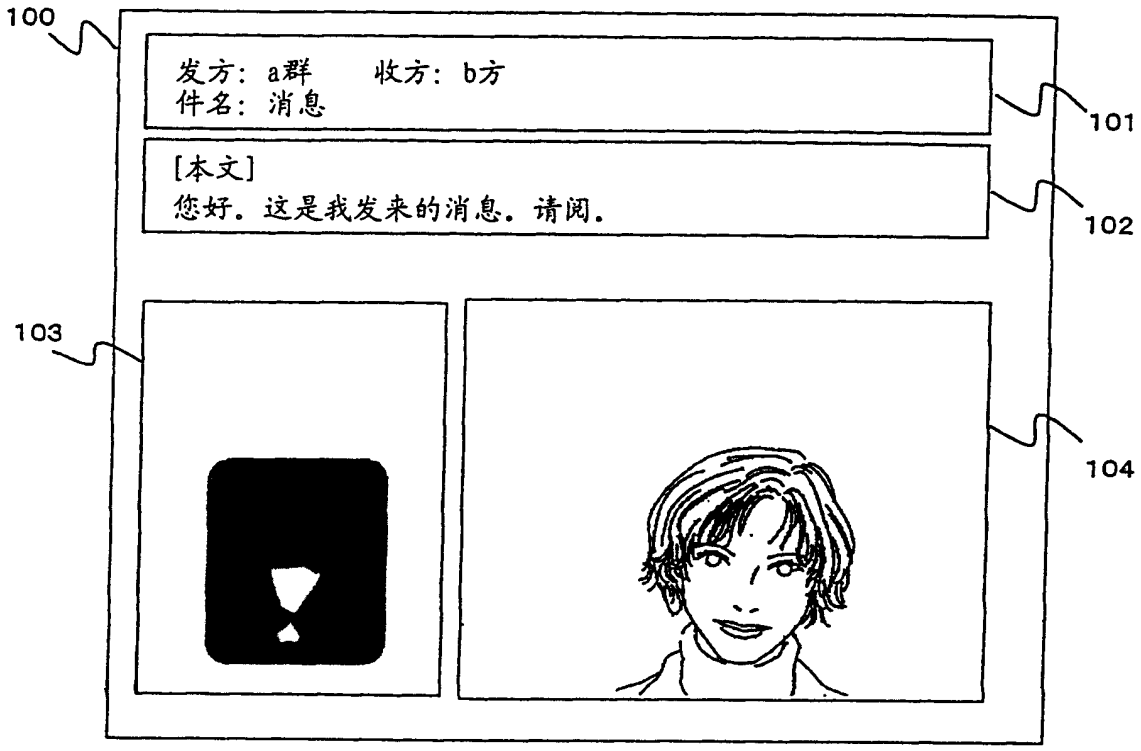


图 6

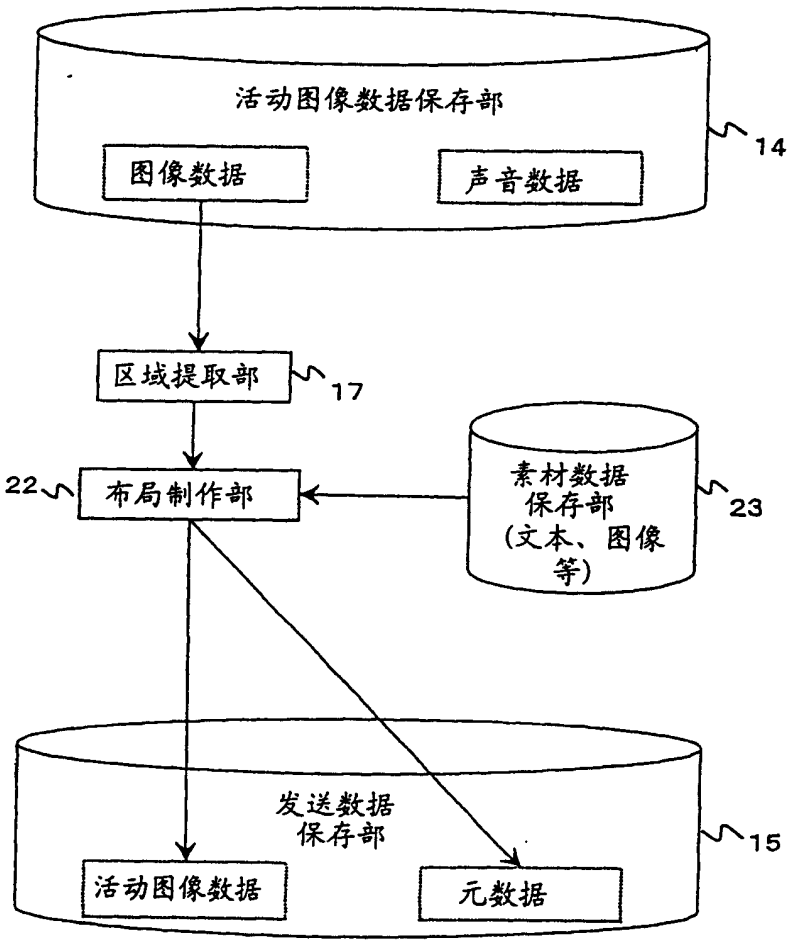

活动图像数据保存部
图像数据
声音数据
14
区域提取部
17
22
布局制作部
素材数据
保存部
(文本、图像
等)
23
发送数据
保存部
15
活动图像数据
元数据

图 7

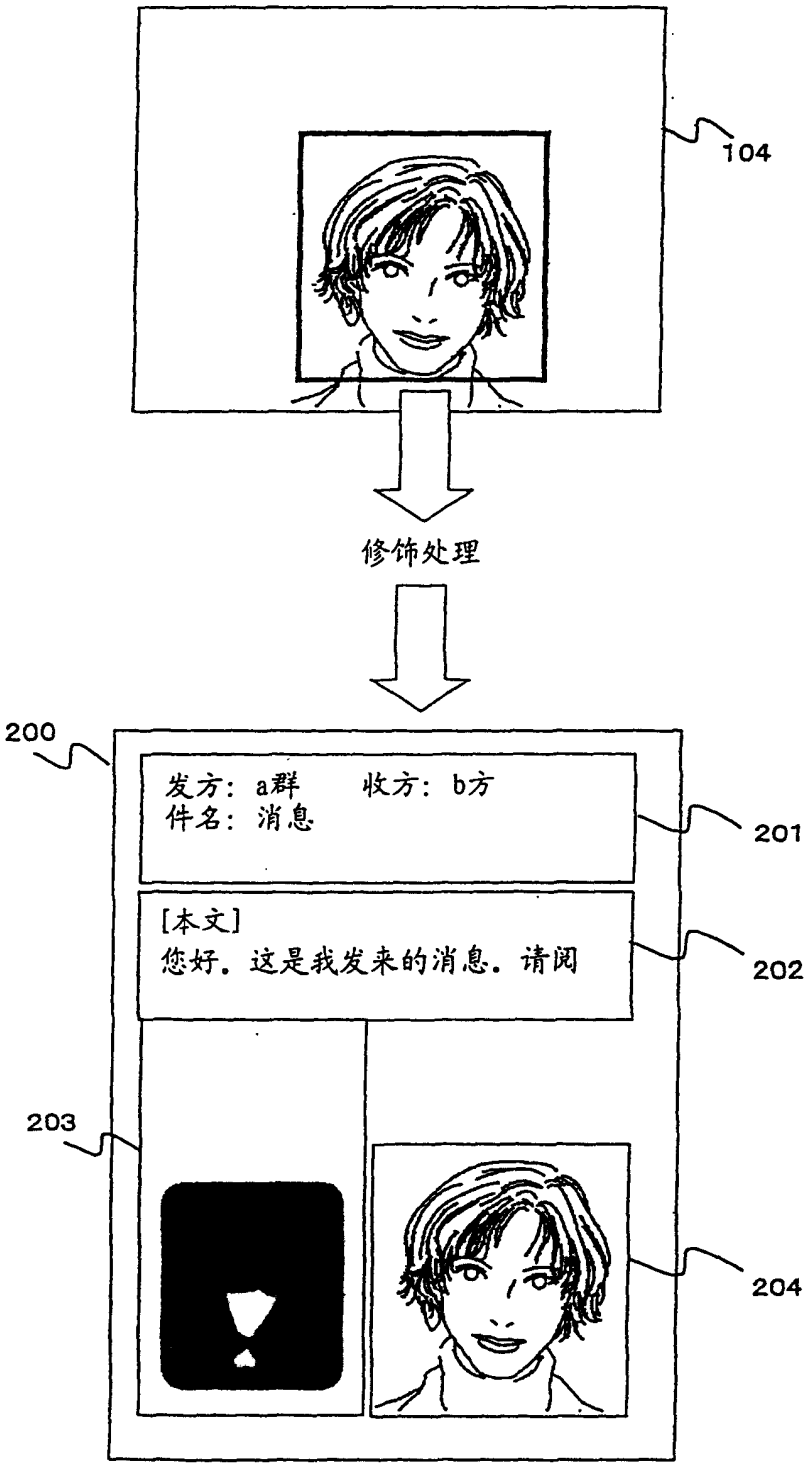


图 8

```
...
<VideoSegment id=" 第1帧 ">
    <MediaTime>
        <MediaTImePoint>T13:20:01:1F15</MediaTimePoint>
    </MediaTime>
    <StillRegion id=" 部分区域信息 ">
        <ContourShape numberOfPeaks="4">
            . . .
        </ContourShape>
    </StillRegion>
</VideoSegment>
...
<VideoSegment id=" 第2帧 ">
    <MediaTime>
        <MediaTImePoint> T13:23:20:15F15 </MediaTimePoint>
    </MediaTime>
    <StillRegion id=" 部分区域信息 ">
        <ContourShape numberOfPeaks="4">
            . . .
        </ContourShape>
    </StillRegion>
</VideoSegment>
...
<VideoSegment id=" 第3帧 ">
    <MediaTime>
        <MediaTImePoint>T13:29:02:3F15</MediaTimePoint>
    </MediaTime>
    <StillRegion id="部分区域信息 ">
        <ContourShape numberOfPeaks="4">
            . . .
        </ContourShape>
    </StillRegion>
</VideoSegment>
```

图 9

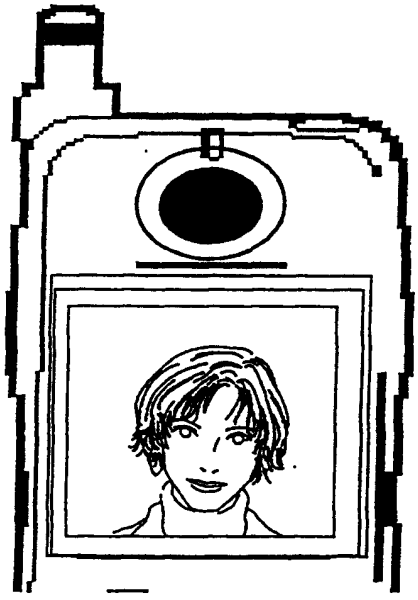


图 10

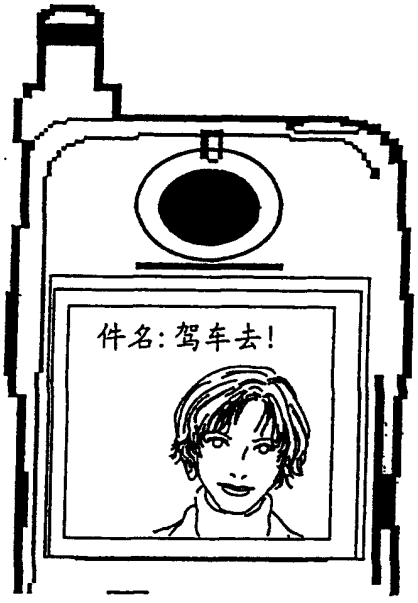


图 11

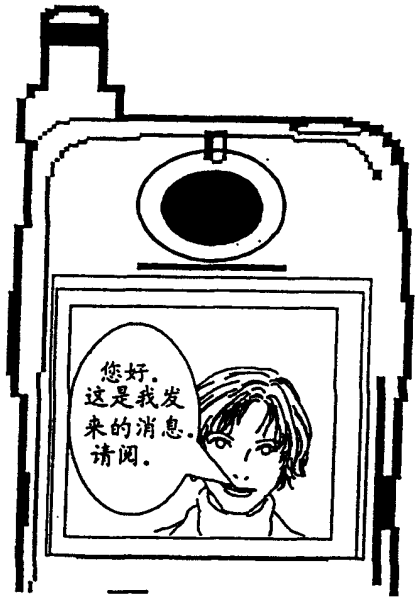
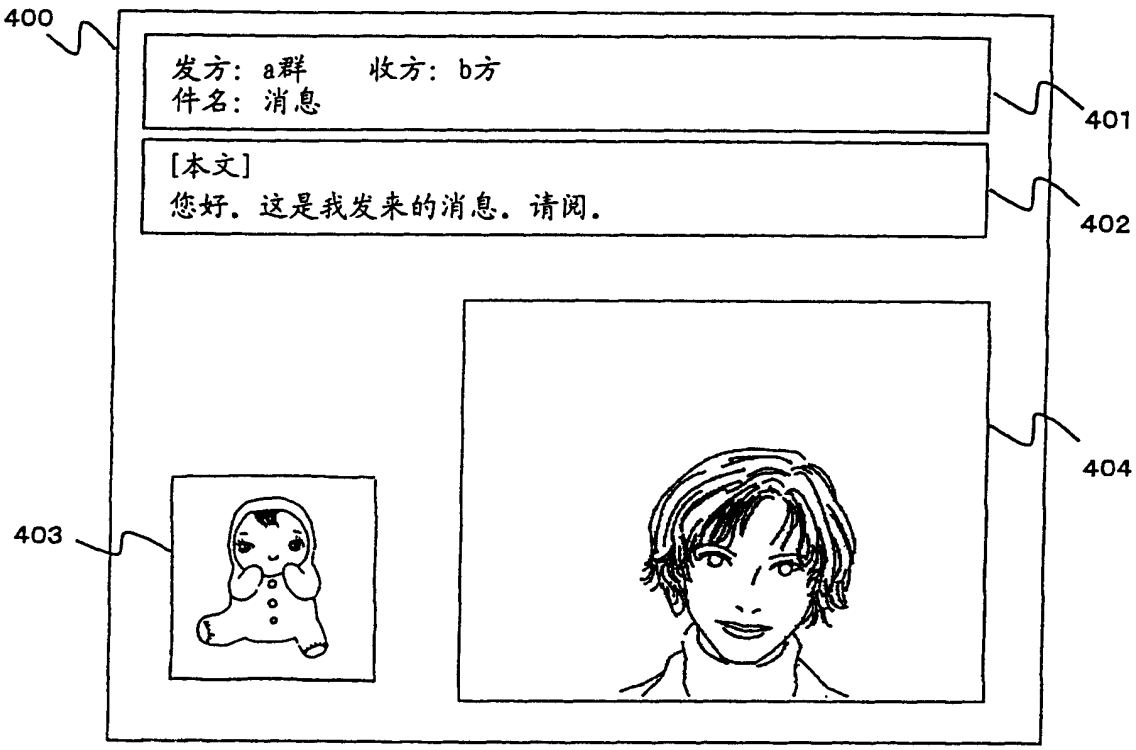


图 12

```
...
<VideoText id="VideoText1"
        textType = "Superimposed" fontSize = "40" fontType="ゴシック">
        <MediaTime>
                <MediaTimePoint>T13:20:01:1F15</MediaTimePoint>
                <MediaDuration>PT6S</MediaDuration>
        </MediaTime>
        ...
        <Text> 您好．这是我发来的消息．请阅．    </Text>
</VideoText>
```

图 13



显示画面

图 14

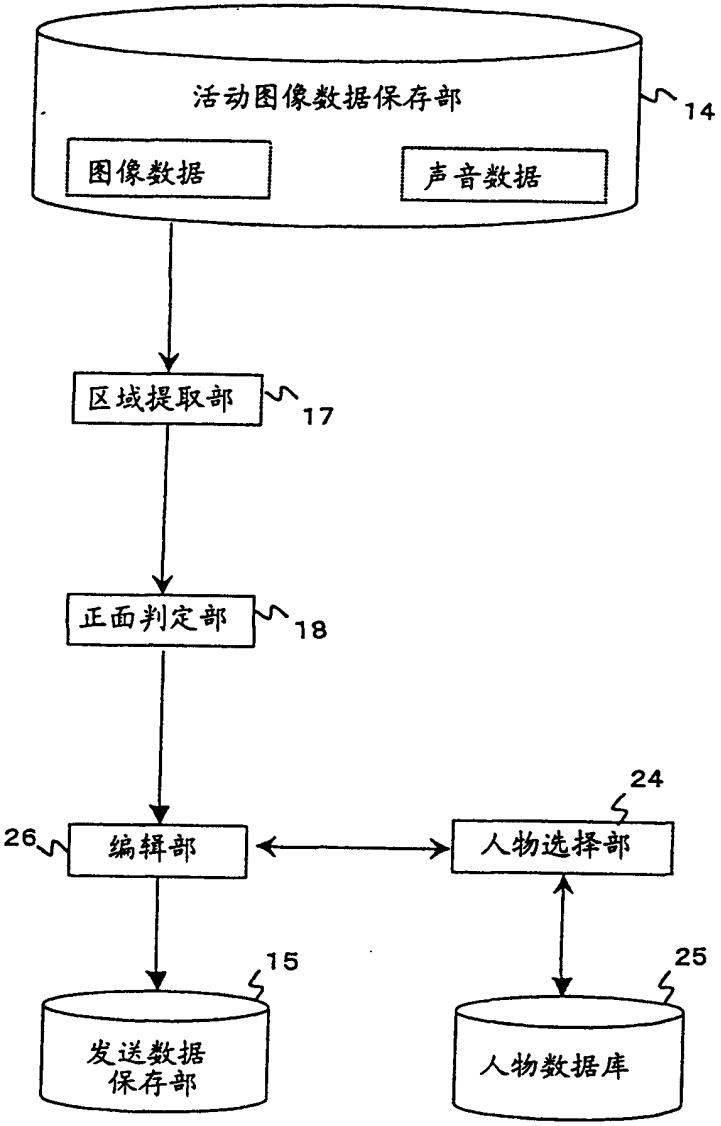
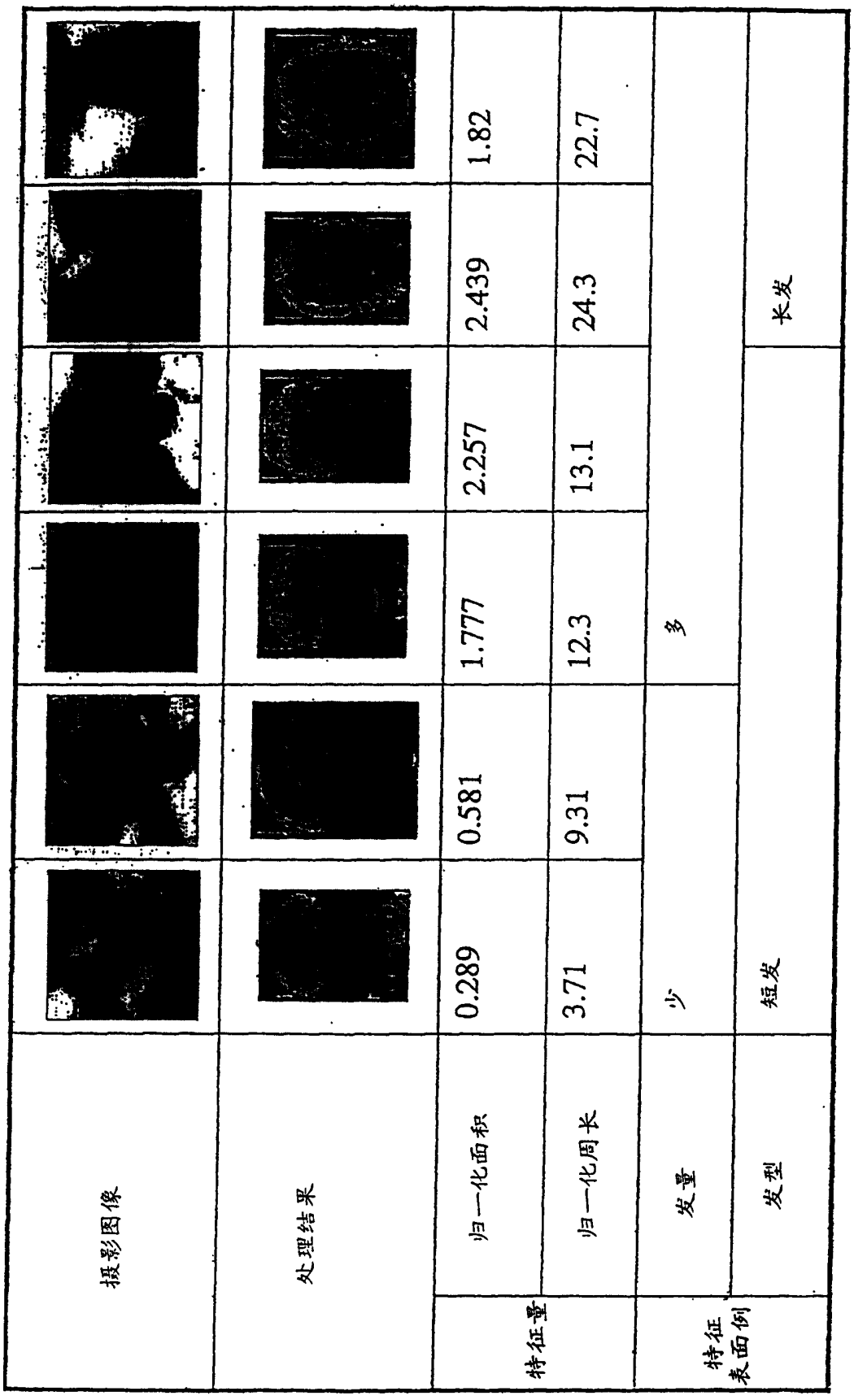

活动图像数据保存部
14
图像数据
声音数据
区域提取部
17
正面判定部
18
24
26
编辑部
人物选择部
15
25
发送数据
保存部
人物数据库

图 15



摄影图像						
处理结果						
特征量	归一化面积	0.289	0.581	1.777	2.257	2.439	1.82
	归一化周长	3.71	9.31	12.3	13.1	24.3	22.7
特征表面例	发量	少		多			
	发型	短发				长发	

图 16

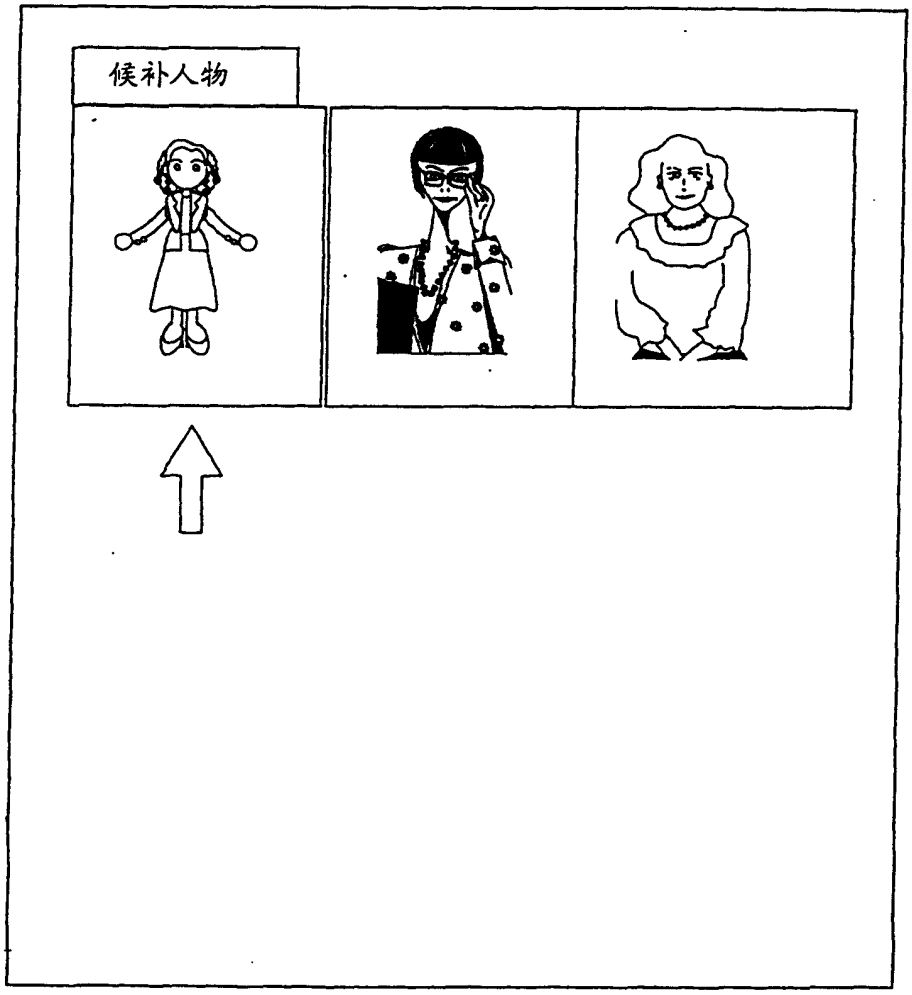


图 17



图 18

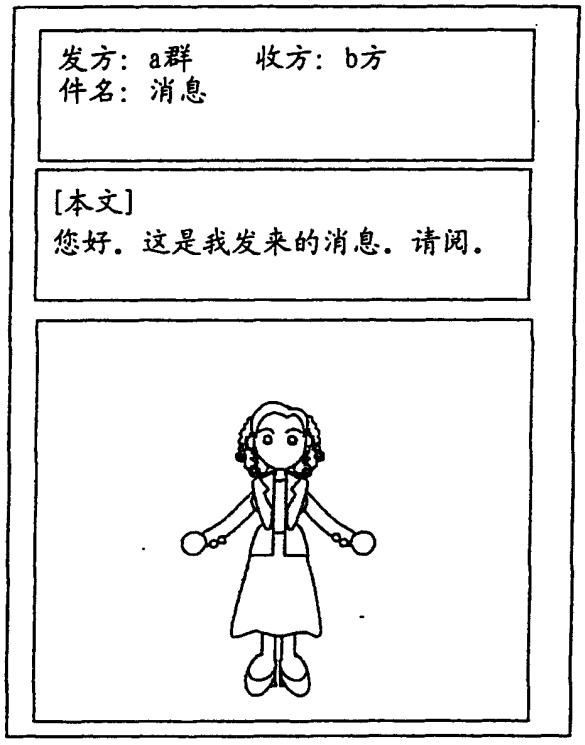


图 19

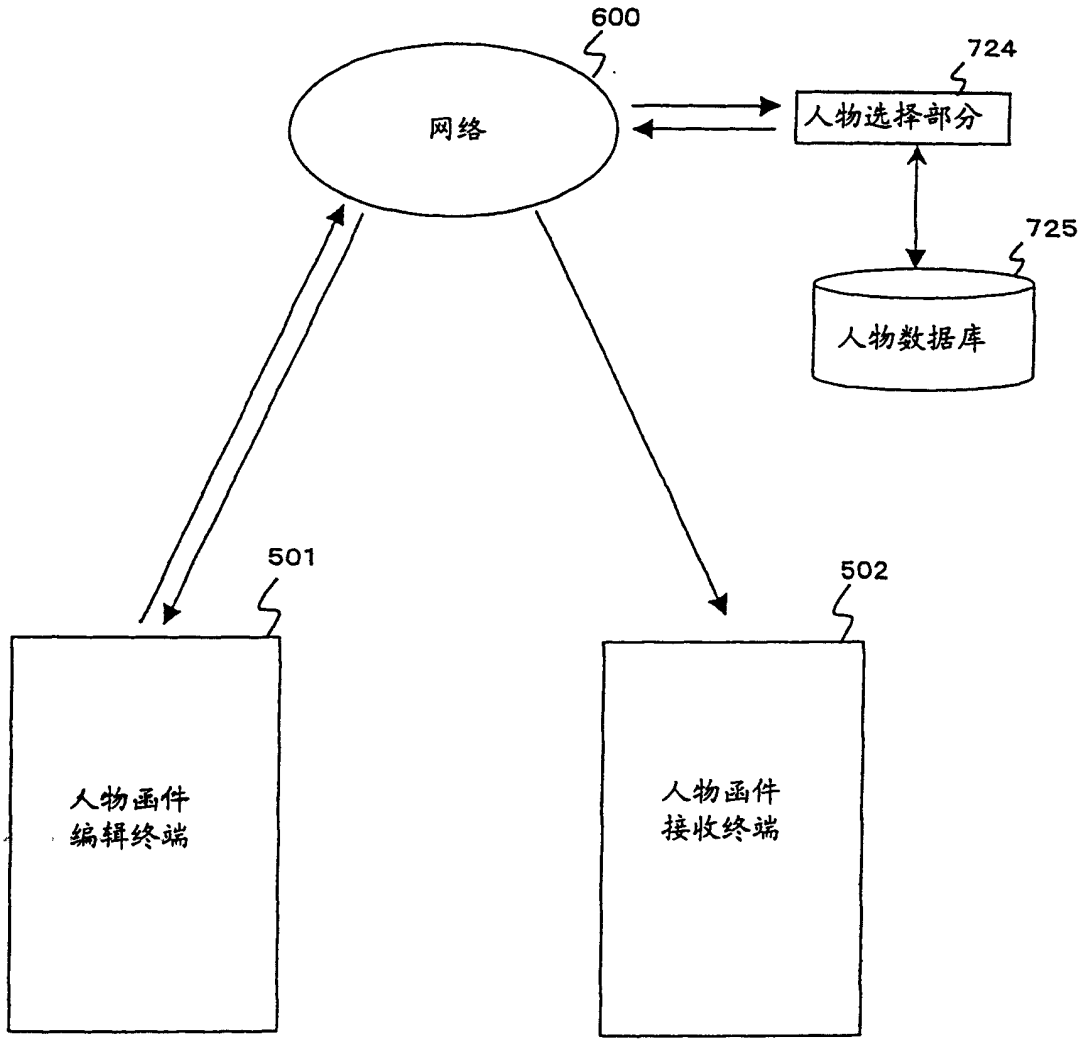


图 20