



[12]发明专利申请公开说明书

[21]申请号 96112495.4

[43]公开日 1997年8月20日

[11]公开号 CN 1157451A

[22]申请日 96.10.31

[74]专利代理机构 中国专利代理(香港)有限公司

[30]优先权

代理人 程天正 叶恺东

[32]95.10.31 [33]JP[31]283426 / 95

[71]申请人 兄弟工业株式会社

地址 日本名古屋市

共同申请人 爱克斯股份有限公司

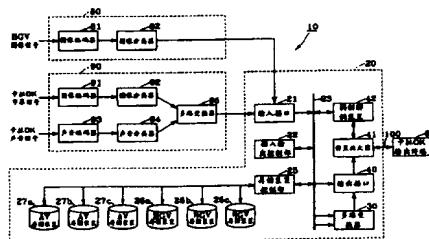
[72]发明人 谷口利典 铃木孝充

权利要求书 1 页 说明书 9 页 附图页数 7 页

[54]发明名称 图象声音数据提供装置

[57]摘要

当输入输出控制部 22 通过调制解调器 42 接收到使用者的点播要求时，就指示存储装置控制部 25 读出并传送 AV 流和 BGV 流。存储装置控制部 25 从 BGV 存储装置 26a～26c 读出 BGV 流、从 AV 存储装置 27a～27c 读出 AV 流并进行传送时，多路变换器 30 将 AV 流分离为字幕流和声音流后利用字幕流和 BGV 流构筑图像流，将声音流和图像流分时进行多路传送，作为图象声音流而输出。



权 利 要 求 书

1. 一种具有存储图像声音数据的数据存储装置、从该数据存储装置读出图像声音数据的读出装置和传送该读出的图像声音数据的传送装置的图像声音数据提供装置，其特征在于：上述数据存储装置除了上述图像声音数据外还存储图像数据；上述读出装置是也读出上述图像数据的结构，并且还设有将由上述读出装置读出的图像声音数据和图像数据综合后作为1个综合图像声音数据向上述传送装置输出的数据综合装置。

2. 按权利要求1所述的图像声音数据提供装置的特征在于：上述图像声音数据由字幕用的文字图像数据和演奏用的演奏声音数据构成。

3. 按权利要求1或2所述的图像声音数据提供装置的特征在于：设有向上述读出装置指示应读出的上述图像声音数据和图像数据的指示装置。

4. 按权利要求3所述的图像声音数据提供装置的特征在于：上述指示装置向上述读出装置指示读出输出终端要求的上述图像声音数据，该输出终端是接收该图像声音数据提供装置提供的上述综合图像声音数据的装置。

说 明 书

图像声音数据提供装置

本发明涉及有时称为图像服务系统或媒体服务系统的图像声音数据提供装置。
5

以往，在称为图像服务系统等的图像声音数据提供装置中，其基本功能是预先将图像和声音一起进行编码，将其综合的数据（系统流）存储到硬盘等存储装置中，根据用户的要求顺序读出，传送给用户（严格地说是终端装置）。其中的数据形式通常是按国际标准规格确定的M P E G 标准，
10 按照从规格上确定的约束条件，在1个数据中图像数据和声音数据保持图像和声音的同步性进行多路传送。

在这样的图像声音数据提供装置中，若考虑将卡拉O K 软件编码后作为系统流进行存储的情况，则存在如下问题。

如所周知，卡拉O K 是将卡拉O K 歌词迭加到背景图像（通常是动图像）上进行显示，未输入歌唱的音乐与卡拉O K 歌词同步地进行演奏。例如，设想有1万个乐曲的卡拉乐曲时，如果对每个乐曲准备特定的图像，假定1个乐曲为3分钟，则图像的总时间也有3 0 , 0 0 0 分钟，是一个庞大的数量。因此，在通信卡拉O K 等装置中，通过使乐曲的风格（演歌、通俗音乐、摇滚音乐等）与背景图像的风格相关联，将X 风格的背景图像与A 风格的乐曲组合循环使用图像数据，例如用数小时～数十小时的图像数据提供大量的卡拉O K 乐曲的背景图像，避免具有大量的图像数据的负担。
20
25

但是，在先有的图像声音数据提供装置中，为了存储卡拉O K 软件，预先将图像和声音一起进行编码，所以，假定有1万个乐曲的卡拉O K 乐曲，则图像也必须有1万个乐曲的图像。即使对多个卡拉O K 乐曲使用1种图像，编码后的图像数据的总量，也会达到1万个乐曲的演奏时间，如果是上述例子，就会达到3 0 , 0 0 0 分钟。即使利用M P E G 的技术，这也会达到数千兆字节以上的数据量，图像声音数据提供装置的存储容量及图像的编码处理所需要的时间等也很庞大，从而在费用和效果方面成为很大的问题。
30

这样的问题不限于卡拉O K 软件，将别的图像数据附加到图像声音数

据上作为系统流进行存储时也一样。

本发明就是为了解决上述问题而提出的，目的旨在提供例如既可以提供与大量的卡拉OK乐曲对应的图像数据又可以减少存储的图像数据的总量的性能价格比优异的图像声音数据提供装置。

作为用于解决上述问题的手段，本发明的第一实施例的图像声音数据提供装置的特征在于：在具有存储图像声音数据的数据存储装置、从该数据存储装置读出图像声音数据的读出装置和传送该读出的图像声音数据的传送装置的图像声音数据提供装置中，上述数据存储装置除了上述图像声音数据外还存储图像数据，上述读出装置是也读出上述图像数据的结构，并设有将由该读出装置读出的图像声音数据和图像数据综合后作为1个综合图像声音数据向上述传送装置输出的数据综合装置。

本发明的第二实施例的图像声音数据提供装置的特征在于：在上述本发明第一实施例的图像声音数据提供装置中，上述图像声音数据由字幕用的文字图像数据和演奏用的演奏声音数据构成。

本发明的第三实施例的图像声音数据提供装置的特征在于：在本发明的第一或第二实施例所述的图像声音数据提供装置中，设置指示上述读出装置应读出的上述图像声音数据和图像数据的指示装置。

本发明的第四实施例的图像声音数据提供装置的特征在于：在本发明的第三实施例所述的图像声音数据提供装置中，上述指示装置指示上述读出装置读出输出终端要求的上述图像声音数据，该输出终端是接收该图像声音数据提供装置提供的上述综合图像声音数据的装置。

在上述第一实施例所述的图像声音数据提供装置中，数据存储装置存储图像声音数据和图像数据，读出装置读出图像声音数据和图像数据。数据综合装置将由读出装置读出的图像声音数据和图像数据综合后作为1个综合图像声音数据输出给传送装置。传送装置传送该综合图像声音数据。传送的综合图像声音数据和先有的图像声音数据一样，可供用户使用。

例如，如果将卡拉OK的字幕和与演奏声音对应的数据作为图像声音数据，将与背景画面对应的数据作为图像数据，则可将1个图像数据用于多个卡拉OK乐曲，所以，不必准备卡拉OK的总演奏时间的图像数据。

即，由于不必准备与图像声音数据的总量相当的图像数据，所以，该图像声音数据提供装置的性能价格比非常优异。

在第二实施例所述的图像声音数据提供装置中，由于图像声音数据由

字幕用的文字图像数据和演奏用的演奏声音数据构成，所以，极适合于提供卡拉OK软件。

在第三实施例所述的图像声音数据提供装置中，指示装置向读出装置指示应读出的图像声音数据和图像数据。

这样，例如，便可提供所要求的卡拉OK软件等，可以提供所希望的综合图像声音数据。

在第四实施例所述的图像声音数据提供装置中，指示装置指示读出装置读出输出终端要求的上述图像声音数据，该输出终端是接收图像声音数据提供装置提供的上述综合图像声音数据的装置。因此，可以将把输出终端所要求的图像声音数据加工后的综合图像声音数据传送给输出终端。该结构非常适合于可以在与输出终端之间进行双向通信的系统。

图1是表示具体例的信息提供系统的概况和信息提供装置的结构的框图。

图2是表示具体例的信息提供装置内装的多路变换器的结构的框图。

图3是表示作为具体例的信息提供系统的终端的信息输出装置的结构的框图。

图4是表示具体例的信息提供装置存储的BGV流和AV流的结构的说明图。

图5是具体例的信息提供装置内装的多路变换器的数据加工处理的说明图。

图6是表示从具体例的信息提供装置内装的多路变换器输出的图像声音流的结构的说明图。

图7是在具体例的信息输出装置中显示的画面的说明图。

下面，利用将本发明的图像声音数据提供装置作为用于提供卡拉OK用的图像声音数据的信息提供装置而具体化了的具体例，详细说明本发明的实施形态。

如图1所示，本具体例的信息提供系统10是利用同轴电缆的通信线路100将作为图像声音数据提供装置的信息提供装置20和信息输出装置60连接而成，该信息输出装置60则被作为利用从该信息提供装置20供给的数据输出卡拉OK演奏声音和卡拉OK用图像的输出终端。另外，具有图像编码器81和图像分类器82的数据输入装置80和具有图像编码器91、图像分类器92、声音编码器93、声音分类器94和多路变

换器 9 5 的数据输入装置 9 0 附属于信息提供装置 2 0 , 但是, 这两个数据输入装置 8 0 , 9 0 可以与信息提供装置 2 0 自由装卸, 信息输出装置 6 0 图中只示出了 1 台, 但是, 通常是一个信息输出装置 6 0 与 1 台信息提供装置 2 0 连接。

5 数据输入装置 8 0 是用于将背景图像 (B G V) 用的 B G V 图像信号数字化后进行编码并将其分类的装置, 图像编码器 8 1 可以将模拟的 B G V 图像信号数字化后进行编码变换为 B G V 数据, 图像分类器 8 2 将 B G V 数据如图 4 所示的那样分类为背景图像包 (B G V 1 、 B G V 2 、 B G V 3 、 . . .), 可以作为 B G V 流输出。

10 在本具体例中, B G V 图像信号如图 4 所示的那样按 720×320 像素的图像尺寸进行编辑, B G V 流按照 M P E G 2 规格进行分类。该 720×320 像素的尺寸是与图 7 所示的标准的 M P E G 2 图像的 1 帧的上侧 $2/3$ 相当的图像尺寸。

15 数据输入装置 9 0 是用于分别将歌词字幕用的卡拉 O K 字幕信号和卡拉 O K 演奏用的卡拉 O K 声音信号数字化后进行编码、分类并进行分时多路传送的装置。

图像编码器 9 1 可以将模拟的卡拉 O K 字幕信号数字化后进行编码变换为字幕数据, 图像分类器 9 2 可以将字幕数据分类后形成字幕流。

20 卡拉 O K 字幕信号不是按文本数据的形式而是如图 4 所示的那样按 720×160 像素的图像尺寸例如使在蓝色的背景上写入了空心的字幕文字状态下的图像信号化后进行编辑, 该 720×160 像素的尺寸是与图 7 所示的标准的 M P E G 2 图像的 1 帧的下侧 $1/3$ 相当的图像尺寸。

25 另外, 声音编码器 9 3 可以将模拟的卡拉 O K 声音信号数字化后进行编码变换为声音数据, 声音分类器 9 4 可以将声音数据分类后形成声音流。

此外, 多路变换器 9 5 按照 M P E G 2 规格将从图像分类器 9 2 传送来的字幕流和从声音分类器 9 4 传送来的声音流进行多路传送, 可以如图 4 所示的那样形成由字幕包 (V 1 、 V 2 、 V 3 、 . . .) 和声音包 (A 1 、 A 2 、 A 3 、 . . .) 构成的字幕声音流 (A V 流) 进行输出。如图 30 4 所示, 该 A V 流与 720×160 像素的字幕图像和卡拉 O K 声音对应。

数据输入装置 8 0 与信息提供装置 2 0 连接时作为数据输入装置 8 0

的输出的 BGV 流输入输入接口 21。输入输入接口 21 的 BGV 流根据
既是控制装置又是指示装置的输入输出控制部 22 的指示通过总线 23 传
送给作为读出装置的存储装置控制部 25，由存储装置控制部 25 写入到
作为数据存储装置的背景图像存储装置（BGV 存储装置）26a，26
5 b，26c 的指定的区域进行存储。存储在 BGV 存储装置 26a，26b，
26c 内的 BGV 流可以由存储装置控制部 25 读出。

用于识别各 BGV 流的识别序号及表示该 BGV 流的图像的属性（风
景、人物等的图像种类和最好将该图像作为背景图像使用的卡拉OK 乐曲
的风格等）的数据附加到 BGV 流上。

10 同样，数据输入装置 90 与信息提供装置 20 连接时作为数据输入装
置 90 的输出的 AV 流也输入输入接口 21。输入输入接口 21 的 AV 流
根据输入输出控制部 22 的指示通过总线 23 传送给存储装置控制部 2
5，由存储装置控制部 25 写入作为数据存储装置的图像声音存储装置（AV
15 存储装置）27a，27b，27c 的指定的区域进行存储。存储在 AV
存储装置 27a，27b，27c 内的 AV 流可以由存储装置控制部 25 读出。

20 用于识别各 AV 流的数据附加在各 AV 流上。另外，将各 AV 流分类
为该卡拉OK 声音的种类（演歌、通俗音乐、摇滚音乐、民歌等种类），
按各分类将作为背景图像理想的 BGV 流的识别序号已列表的对称表存储
在 AV 存储装置 27a、27b、27c 中。该对称表每当有 AV 流加减和 BG
V 流加减时进行更新。

25 存储装置控制部 25 可以将存储在 BGV 存储装置 26a，26b，
26c 内的 BGV 流和存储在 AV 存储装置 27a，27b，27c 内的
AV 流读出并传送给作为数据综合装置的多路变换器 30。

下面，详细说明该多路变换器 30 的结构。如图 2 所示，多路变换器
30 具有缓冲器 31，32、分离装置 33、合成装置 34 和多路传送装
置 35。

30 缓冲器 31 是用于暂时存储从 AV 存储装置 27a～27c 传来的
AV 流并传送给分离装置 33 的缓冲存储器。缓冲器 32 是用于暂时存储
从 BGV 存储装置 26a～26c 传来的 BGV 流并传送给合成装置 3
4 的缓冲存储器。

分离装置 33 是用于将 AV 流分离为声音流和字幕流的装置。

如图 5 所示，该分离装置 3 3 将由字幕包 (V 1、V 2、V 3、...) 和声音包 (A 1、A 2、A 3、...) 构成的 A V 流分离为只有声音包的声音流和只有字幕包的字幕流，如图 2 和图 5 所示，声音流可以传送给多路传送装置 3 5，字幕流可以传送给合成装置 3 4。

如图 5 所示，合成装置 3 4 可以将从分离装置 3 3 传送来的字幕流的包解除形成字幕数据，将从缓冲器 3 2 传送来的 B G V 流的包解除形成 B G V 数据。并且，可以将字幕数据和 B G V 数据合成，形成 1 条图像数据。

该图像数据是将与图 7 所示的标准的 M P E G 2 图像的 1 帧的上侧 2 / 3 区域的图像对应的 B G V 数据和与同一帧的下侧 1 / 3 区域的图像对应的字幕数据交替地连接的结构，只要再生该图像数据，就可以如图 7 (b) 所示的那样，成为 1 帧的上侧 2 / 3 配上 B G V 图像而同一帧的下侧 1 / 3 配上字幕图像的图像。虽然是再生上下分割状态的图像，但是，图像数据本身是一连串的数据。

此外，合成装置 3 4 还可以将该图像数据按照 M P E G 2 规格进行分类后作为由图像包 (M V 1、M V 2、M V 3、...) 构成的图像流如图 2 和图 5 所示的那样传送给多路传送装置 3 5。

如图 5 所示，多路传送装置 3 5 可以将从分离装置 3 3 传送来的声音流和从合成装置 3 4 传送来的图像流按照 M P E G 2 规格分时进行多路传送，并作为图像声音流输出。如果将该图像声音流译码后进行再生，便可如图 6 所示的那样获得在 1 帧的上侧 2 / 3 配上 B G V 图像而在同一帧的下侧 1 / 3 配上字幕图像的图像和卡拉OK 声音。

如图 2 所示，多路变换器 3 0 可以将该图像声音流传送给输出接口 4 0。

输出接口 4 0 与作为输出装置的前置放大器 4 1 连接，从多路变换器 3 0 输出的图像声音流在前置放大器 4 1 中进行高频调制后利用指定的频道传送给通信线路 1 0 0。

另外，前置放大器 4 1 利用双向通信用的频道传送给作为通信装置的调制解调器 4 2 传送来的信号，可以将从通信线路 1 0 0 利用双向通信用的频道输入的信号传送给调制解调器 4 2。

调制解调器 4 2 通过总线 2 3 与输入输出控制部 2 2 连接，输入输出控制部 2 2 可以通过调制解调器 4 2 和前置放大器 4 1 将数据传送给利用通信线路 1 0 0 连接的信息输出装置 6 0，或接收信息输出装置 6 0 的数

据。另外，输入输出控制部 2 2 还可以控制上述信息提供装置 2 0 的各部分的动作。

下面，说明信息输出装置 6 0 的结构。

如图 3 所示，信息输出装置 6 0 具有作为用于控制信息输出装置 6 0 的各部分的动作的控制装置的中央控制装置 6 1。该中央控制装置 6 1 由众所周知的微机构成。作为具有键盘（图中未示出）的输入设备的输入装置 6 3 与中央控制装置 6 1 连接，只要操作输入装置 6 3，就可以向中央控制装置 6 1 输入卡拉OK 乐曲的点播和各种指示。

另外，中央控制装置 6 1 通过作为通信装置的调制解调器 6 2 与通信线路 1 0 0 连接，可以使用上述双向通信用的频道与信息提供装置 2 0 的输入输出控制部 2 2 进行通信。这样，便可将从输入装置 6 3 输入的点播传送给输入输出控制部 2 2，从而可以要求传送该卡拉OK 乐曲用的图像声音流。

此外，作为通信设备的调谐器 6 4 与通信线路 1 0 0 连接。调谐器 6 4 可以根据中央控制装置 6 1 的指示选择频道，接收利用所选择的频道传来的图像声音流，传送给系统译码器 6 6。

系统译码器 6 6 可以将由调谐器 6 4 接收并传来的图像声音流分离为声音流和图像流，将声音流向声音译码器 6 7 输出，将图像流向图像译码器 6 8 输出。

声音译码器 6 7 可以将声音流译码后并进行数 / 模变换（D / A 变换）后作为卡拉OK 声音信号向放大器 6 9 输出。

放大器 6 9 可以将卡拉OK 声音信号放大后传送给作为声音输出装置的扬声器 7 0，作为卡拉OK 演奏声音输出。另外，放大器 6 9 具有混合功能，可以将从麦克风 7 1 输入的歌唱声音信号和从声音译码器 6 7 输入的卡拉OK 声音信号混合后进行放大并向扬声器 7 0 输出。这样，便可从扬声器 7 0 输出卡拉OK 演奏声音与使用者的歌声混合的声音。

图像译码器 6 8 可以将图像流译码后并进行 D / A 变换后作为图像信号向监视器 7 5 输出。监视器 7 5 可以显示与图像信号对应的图像。

下面，说明该信息提供系统 1 0 的动作。

首先，当信息输出装置 6 0 的使用者操作输入装置 6 3 要求提供所希望的卡拉OK 乐曲时，中央控制装置 6 1 就将该卡拉OK 乐曲的图像声音流的传送要求通过调制解调器 6 2 利用通信线路 1 0 0 传送给信息提供装

置 2 0 。

在信息提供装置 2 0 中，输入输出控制部 2 2 通过调制解调器 4 2 接收该要求。其次，输入输出控制部 2 2 向存储装置控制部 2 5 指示读出点播的卡拉OK 乐曲的 AV 流和读出并传送输入输出控制部 2 2 根据包含该卡拉OK 乐曲的 AV 流的分类的对称表选定的 BGV 流。
5

此外，输入输出控制部 2 2 向前置放大器 4 1 指示传送点播的卡拉OK 乐曲的图像声音流的频道，并将该频道通知传送点播的信息输出装置 6 0 。

在信息输出装置 6 0 中，中央控制装置 6 1 指示调谐器 6 4 选择从信息提供装置 2 0 通知的频道。这样，信息输出装置 6 0 就等待图像声音流由该频道传送来。
10

其次，在信息提供装置 2 0 中，存储装置控制部 2 5 按照输入输出控制部 2 2 的指示，从 BGV 存储装置 2 6 a ~ 2 6 c 读出 BGV 流，并传送给多路变换器 3 0 ，从 AV 存储装置 2 7 a ~ 2 7 c 读出 AV 流，并传送给多路变换器 3 0 。
15

多路变换器 3 0 将 AV 流暂时分离为声音流和字幕流，利用字幕流和 BGV 流构筑图像流，并进而将声音流和图像流分时进行多路传送，作为图像声音流输出。

该图像声音流经过输出接口 4 0 传送给前置放大器 4 1，利用由输入输出控制部 2 2 指示的频道从前置放大器 4 1 传送给通信线路 1 0 0 。
20

在信息输出装置 6 0 中，调谐器 6 4 接收到图像声音流后将其传送给系统译码器 6 6 。

系统译码器 6 6 将图像声音流分离为声音流和图像流，并将声音流向声音译码器 6 7 输出，将图像流向图像译码器 6 8 输出。
25

图像译码器 6 8 将图像流译码并进行 D / A 变换后作为模拟的图像信号向监视器 7 5 输出，监视器 7 5 显示与图像信号对应的图像。监视器 7 5 的图像如图 7 (b) 所示的那样，以在画面的上侧 2 / 3 配上背景图像而在下侧 1 / 3 配上字幕图像的状态进行显示。

声音译码器 6 7 将声音流译码并进行 D / A 变换后作为模拟的卡拉OK 声音向放大器 6 9 输出。放大器 6 9 将从麦克风 7 1 输入的歌唱声音信号和从声音译码器 6 7 输入的卡拉OK 声音信号混合后放大并向扬声器 7 0 输出。这样，便可从扬声器 7 0 输出卡拉OK 演奏声音和使用者的歌声。
30

混合的声音。

这时，使用者可以看着监视器 7 5 的字幕进行歌唱。另外，还可以观看与字幕一起显示的背景图像。

如上所述，在信息提供装置 2 0 中，分别预先存储着卡拉OK的字幕和演奏声音用的AV流和背景图像用的BGV流，是由多路变换器 3 0 将AV流和根据该AV流选择的BGV流处理综合后作为图像声音流进行传送的，所以，可以将1个BGV流用于多个AV流（卡拉OK乐曲）。因此，不必准备卡拉O的总演奏时间的BGV流。

即，由于不必准备与AV流的总量相当的BGV流，所以，装置的性能价格比优异。

以上，按照具体例说明了本发明的实施形态，但是，本发明不限于这些具体例，在不超出本发明的范围内可以按各种形式实施。

在具体例中举出了分配卡拉OK软件的例子，但是，本发明也可以用于分配卡拉OK软件以外的情况。

另外，在具体例的信息提供装置中，预先存储了按MPEG2规格分类的AV流和BGV流，将它们综合的图像声音流也按照MPEG2规格进行分类，但是，也可以采用MPEG2以外的分类方式。或者，也可以不进行分类。

图像声音数据提供装置与输出终端的通信不限于具体例那样使用同轴电缆的通信，不论有线还是无线方式都可以采用。另外，不仅可以是电磁的通信，而且也可以是以光为媒体的通信。

如上所述，按照本发明第一实施例所述的图像声音数据提供装置，由于将分别存储的图像声音数据和图像数据综合后作为1个综合图像声音数据进行传送，所以，不需要准备与图像声音数据的总量相当的图像数据。因此，图像声音数据提供装置的性能价格比优异。

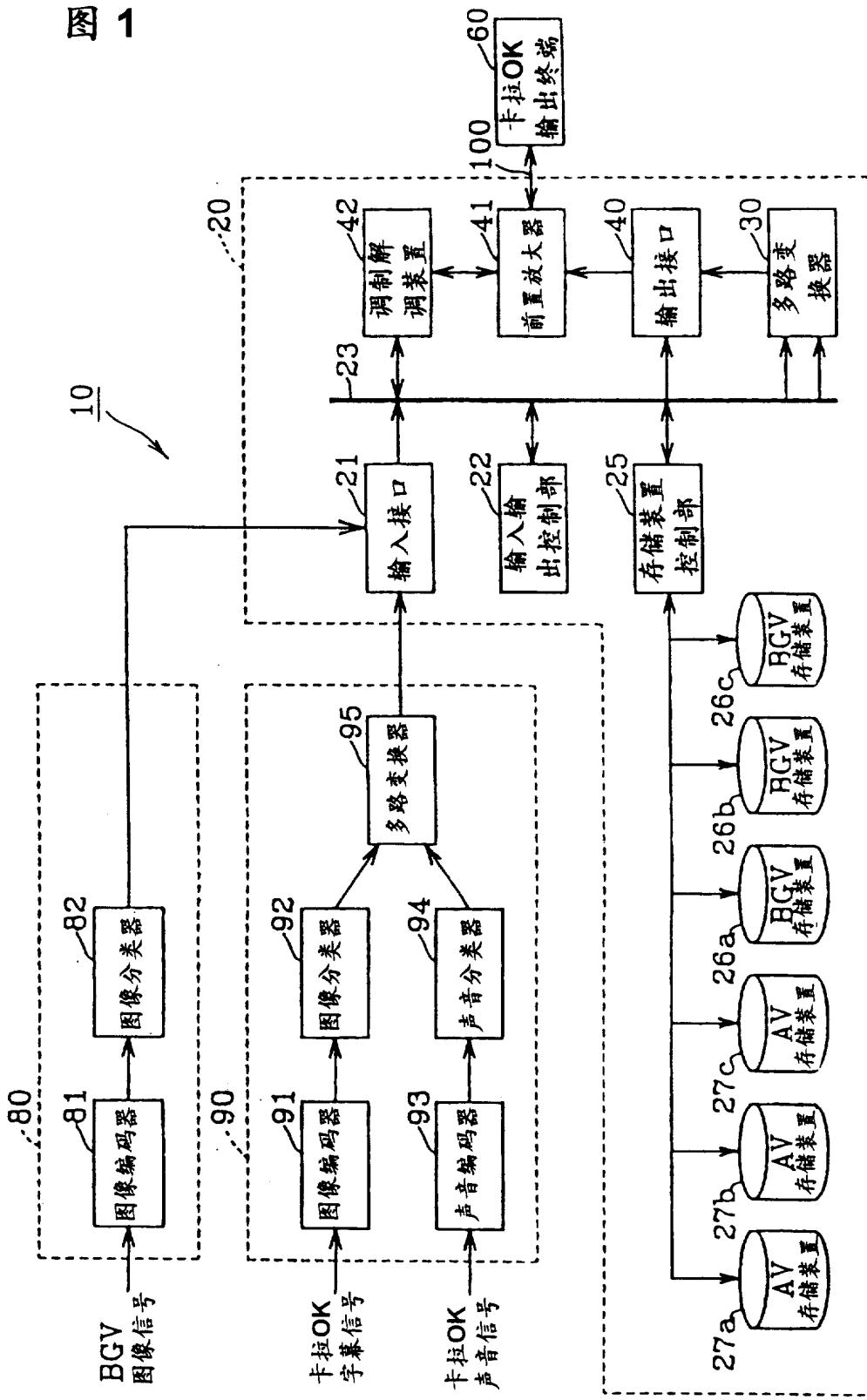
按照本发明第二实施例所述的图像声音数据提供装置极适合于提供卡拉OK软件。

按照本发明第三实施例所述的图像声音数据提供装置，可以提供例如点播的卡拉OK软件等从而可以提供所希望的综合图像声音数据。

按照本发明的第四实施例所述的图像声音数据提供装置由于可以将对由输出终端点播的图像声音数据加工后的综合图像声音数据传送给输出终端，所以，极适合于可以在与输出终端之间进行双向通信的系统。

说 明 图

图 1



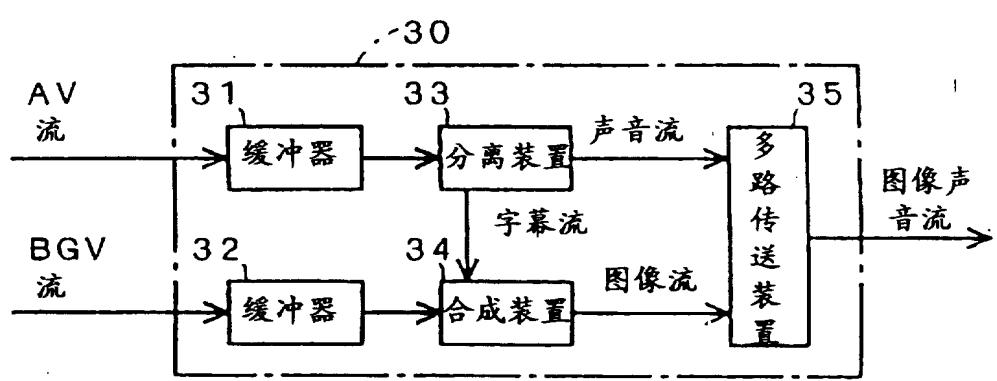


图 2

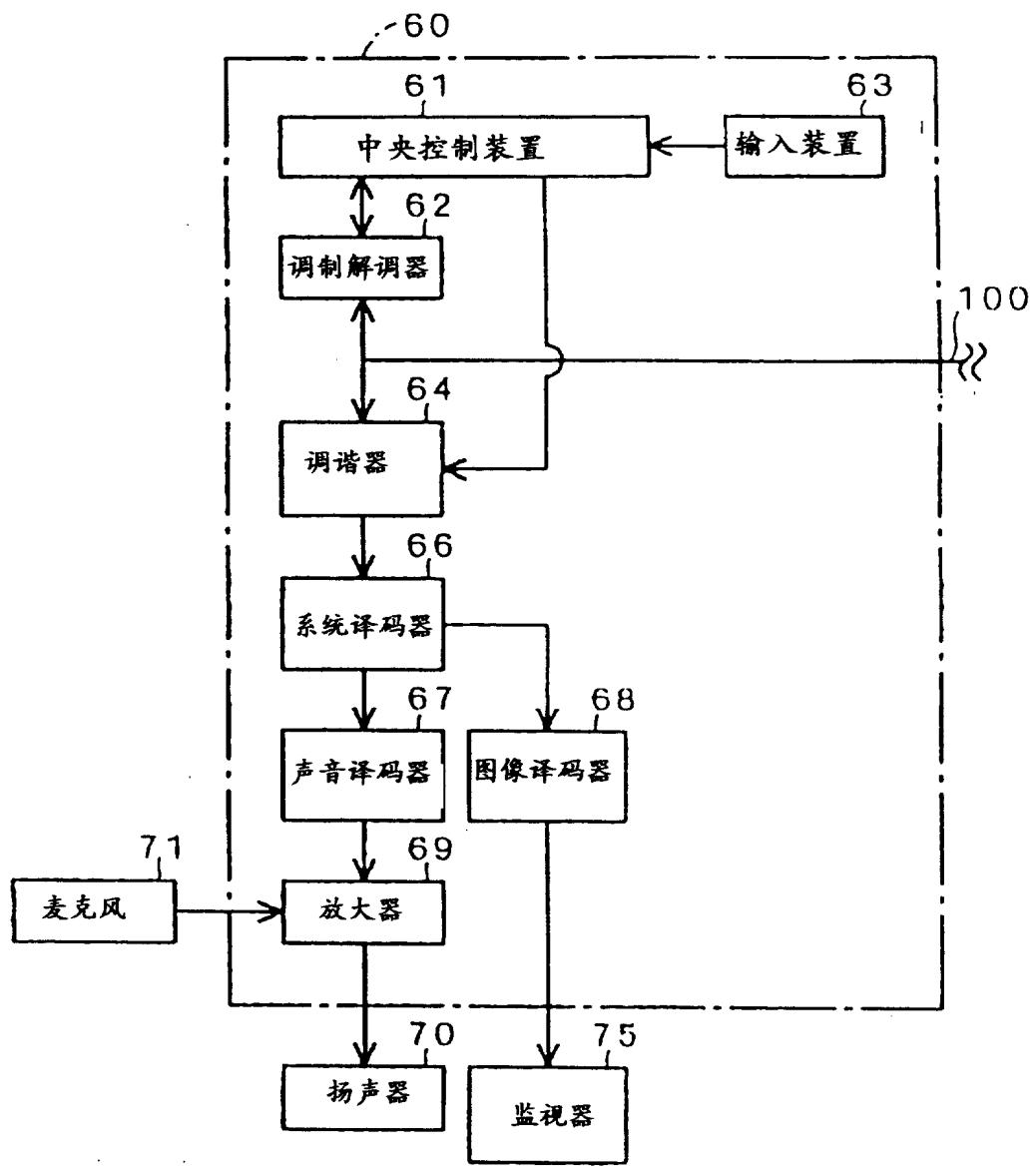


图 3

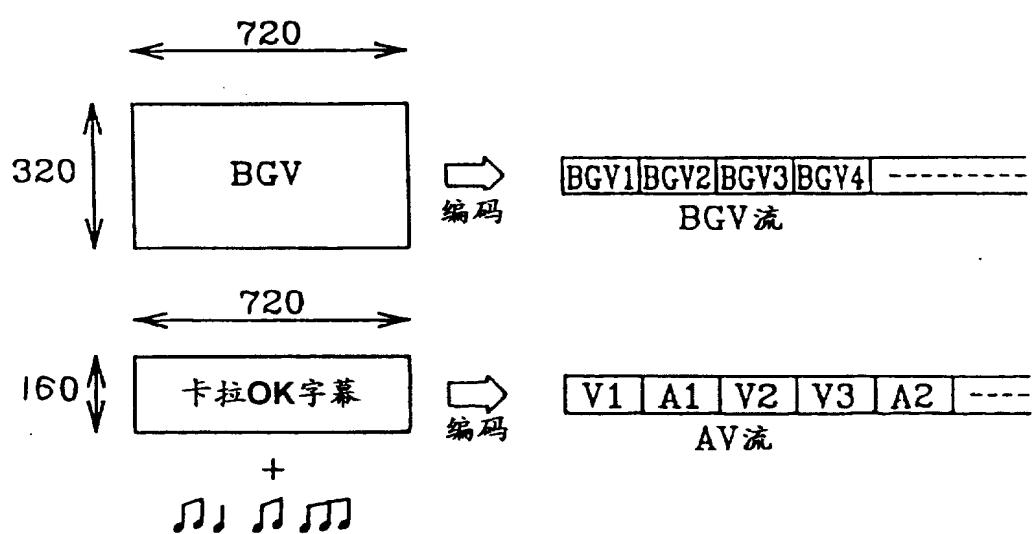


图 4

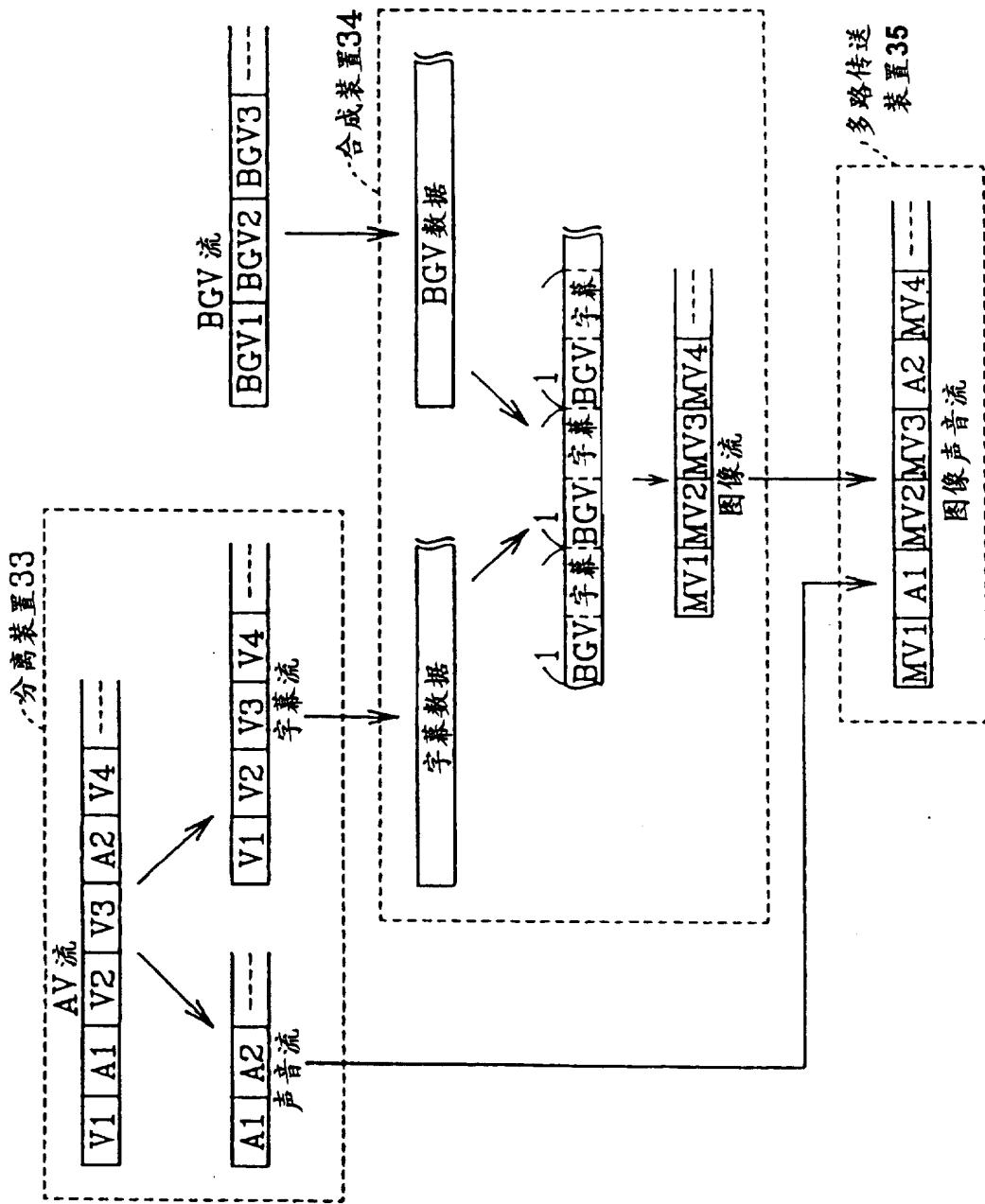


图 5

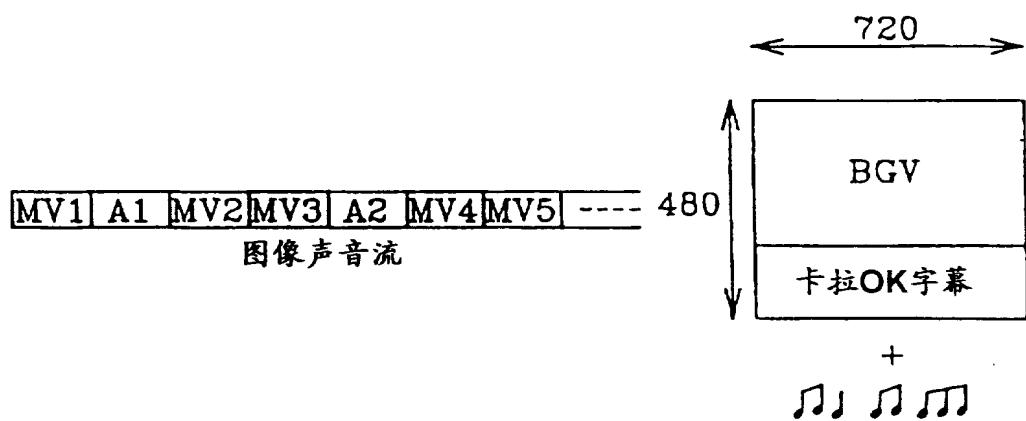


图 6

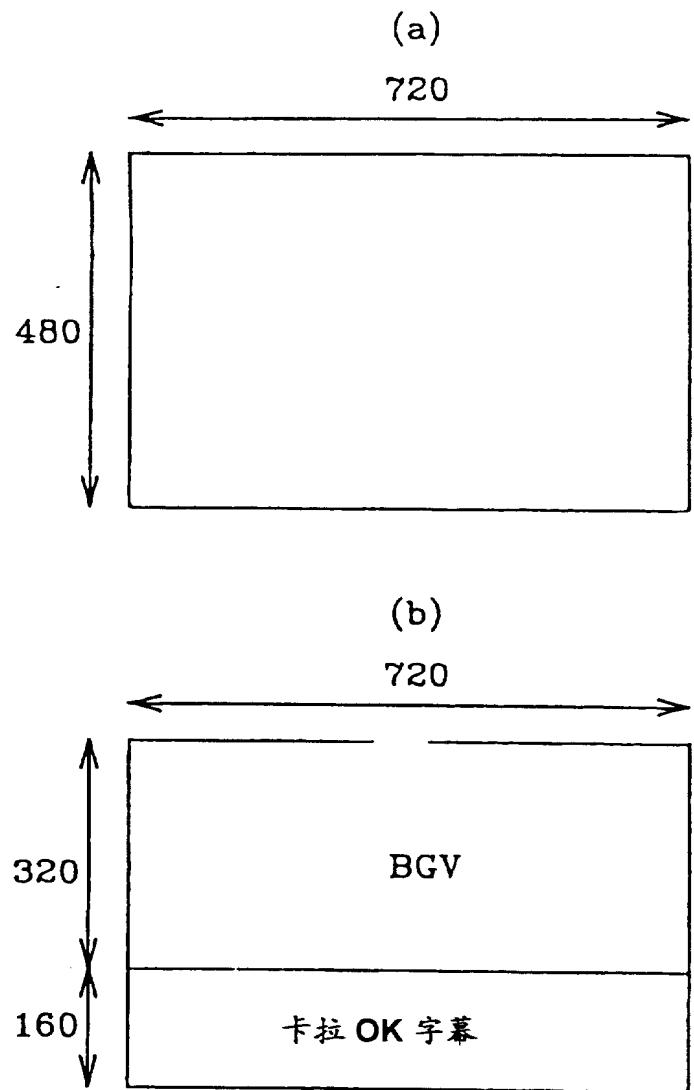


图 7