

1、一种信息处理设备，它解码包括由每个具有传输包及其到达时间印记的多个源数据包组成的数据流的复用流，其中第二复用流的第一个图像作为第二图像，连接到作为第一图像的第一复用流的最后一个图像，从而无缝再现，该信息处理设备包括：

输出装置，用于根据复用流的到达时间印记输出源数据包；

视频缓冲器，用于缓存包括在源数据包中的视频数据；

音频缓冲器，用于缓存包括在源数据包中的音频数据；

视频解码装置，用于解码在视频缓冲器中缓存的视频数据；和

音频解码装置，用于解码在音频缓冲器中缓存的音频数据，其中

音频缓冲器具有能够缓存与输入第二图像到视频缓冲器所需时间对应的音频数据的容量，

其中满足 $EBn_max = (I_max/Rv) \times Ra$ ，其中 EBn_max 比特是音频缓冲器所需的容量； I_max 比特是第二图像的比特量， Rv 比特/秒是对视频缓冲器的输入比特率，并且 Ra 比特/秒是音频数据的比特率。

2、根据权利要求 1 的信息处理设备，其中第二图像是帧内编码图像。

3、根据权利要求 1 的信息处理设备，其中音频缓冲器具有能够缓存对应于 100 毫秒的音频数据的容量。

4、根据权利要求 1 的信息处理设备，其中复用流满足 $STC2^2_{start} > STC2^1_{end}$ ，其中定义 STC_delta 是在第一复用流的时间轴上的第一图像的呈现结束时间和在第二复用流的时间轴上的第二图像的呈现开始时间之间的时间差，则 $STC2^1_{end}$ 是通过使用时间差 STC_delta 转换 $STC1^1_{end}$ 到第二复用流的时间轴上的值而获得的值，其中 $STC2^1_{end} = STC1^1_{end} - STC_delta$ ， $STC1^1_{end}$ 是在第一复用流的最后的源数据包的最后字节从输出装置输出时的第一复用流的时间轴上的值，并且 $STC2^2_{start}$ 是在第二复用流的第一个源数据包的第一字节从输出装置输出时的第二复用流的时间轴上的值。

5、根据权利要求 1 的信息处理设备，其中复用流满足 $STC2^2_{start} > STC2^1_{end} + delta1$ ，其中定义 STC_delta 是第一复用流的时间轴上的第一图像的呈现结束时间和第二复用流的时间轴上的第二图像的呈现开始时间之间的时间差，则 $STC2^1_{end}$ 是通过使用时间差 STC_delta 转换 $STC1^1_{end}$ 到第二复用流的时间

轴上的值而获得的值，其中 $STC2^1_{end} = STC1^1_{end} - STC_delta$ ， $STC1^1_{end}$ 是在第一复用流的最后的源数据包的最后字节从输出装置输出时的第一复用流的时间轴上的值，并且 $STC2^2_{start}$ 是在第二复用流的第一个源数据包的第一字节从输出装置输出时的第二复用流的时间轴上的值，其中在第一复用流的最后的源数据包从输出装置输出后经过预定时间增量 $delta1$ 之后，第二复用流的第一源数据包从输出装置输出。

6、根据权利要求 1 的信息处理设备，其中

假定 STC_delta 是第一复用流的时间轴上的第一图像的呈现结束时间和第二复用流的时间轴上的第二图像的呈现开始时间之间的时间差，并且在第一复用流的最后一个源数据包输出开始后经过预定时间 ATC_delta 后，第二复用流的第一个源数据包从输出装置输出，

预定时间 ATC_delta 这样确定，使得满足时间差 STC_delta ，并且这样形成复用流使得满足时间差 STC_delta 。

7、根据权利要求 6 的信息处理设备，其中预定的时间 ATC_delta 作为第一复用流的附属信息管理。

8、根据权利要求 1 的信息处理设备，其中包括在第一和第二复用流中的音频数据是多信道音频数据。

9. 一种信息处理方法，解码包括由每个具有传输包及其达到时间印记的多个源数据包组成的数据流的复用流，并且其中第二复用流的第一个图像作为第二图像，与作为第一图像的第一复用流的最后一个图像连接，从而无缝再现，该信息处理方法包括：

根据复用流的到达时间印记输出源数据包的步骤；

缓存分别在视频和音频缓冲器中的源数据包中包含的视频和音频数据的步骤；和

对视频和音频缓冲器中缓存的视频和音频数据解码的步骤，其中

在缓存步骤中，在第二图像被缓存在视频缓冲器中之前，与输入第二图像到视频缓冲器所需时间对应的音频数据被缓存在音频缓冲器中，

其中满足 $EBn_max = (I_max/Rv) \times Ra$ ，其中 EBn_max 比特是音频缓冲器所需的容量； I_max 比特是第二图像的比特量， Rv 比特/秒是对视频缓冲器的输入比特率，并且 Ra 比特/秒是音频数据的比特率。

10、一种信息处理设备，产生包括由每个具有传输包及其到达时间印记

的多个源数据包组成的数据流的复用流，并且它由解码器基于到达时间印记读出并解码，该信息处理设备包括：

视频编码装置，用于产生第一视频编码流以结束第一图像的呈现和产生第二视频编码流，使得在第一图像之后立即开始将要呈现的第二图像的呈现；和

复用装置，用于复用第一视频编码流和与第一视频编码流同步的音频编码流，以产生第一复用流，复用第二视频编码流和与第二视频编码流同步的音频编码流，以产生第二复用流；并产生复用流，在该复用流中第二复用流的第一个图像作为第二图像，连接到作为第一图像的第一复用流的最后一个图像，以便无缝再现，以及

音频缓冲器，其容量满足 $EBn_max = (I_max/Rv) \times Ra$ ，其中 EBn_max 比特是音频缓冲器所需的容量； I_max 比特是第二图像的比特量， Rv 比特/秒是对视频缓冲器的输入比特率，并且 Ra 比特/秒是音频数据的比特率，使得与输入第二图像到解码器所需的时间对应的音频数据的输入能够在第二图像向解码器输入开始时完成。

11、根据权利要求 10 的信息处理设备，其中第二图像是帧内编码图像。

12、一种信息处理方法，产生包括由每个具有传输包及其到达时间印记的多个源数据包组成的数据流的复用流，并且它由解码器基于到达时间印记读出并解码，该信息处理方法包括：

产生步骤，产生第一视频编码流以结束第一图像的呈现和产生第二视频编码流使得在第一图像之后立即开始将要呈现的第二图像的呈现；和

复用步骤，复用第一视频编码流和与第一视频编码流同步的音频编码流，以产生第一复用流，复用第二视频编码流和与第二视频编码流同步的音频编码流以产生第二复用流，并产生一个复用流，该复用流中，第二复用流的第一个图像作为第二图像，连接到作为第一图像的第一复用流的最后一个图像，以便无缝再现，

缓冲音频的步骤，其中缓冲量满足 $EBn_max = (I_max/Rv) \times Ra$ ，其中 EBn_max 比特是音频缓冲器所需的容量； I_max 比特是第二图像的比特量， Rv 比特/秒是对缓冲视频的输入比特率，并且 Ra 比特/秒是音频数据的比特率，使得与输入第二图像到解码器所需的时间对应的音频数据的输入能够在第二图像向解码器输入开始时完成。

信息处理装置、信息处理方法

技术领域

本发明涉及一种用于编辑由具有视频帧精确度的视频和音频流组成的复用流以及无缝再现编辑点的信息处理装置、它的方法、一种程序和记录介质，以及一种以最优化的形式产生无缝再现的复用流的信息处理装置和存储复用流数据的记录介质。

本申请要求 2003 年 5 月 8 日申请的日本专利申请号 2003 - 130661 的优先权，其内容结合在本申请中作为参考。

背景技术

在日本专利申请公开号 2000 - 175152、日本专利申请公开号 2001 - 544118 和日本专利申请公开号 2002 - 158974 中描述了用于编辑具有视频帧精确度的视频和音频流的复用流以及无缝再现编辑点的方法。

图 1 是表示传统的 DVR - STD 模式 (DVR MPEG2 传输流播放器模式) (下文中称为“播放器”) 101 的框图。DVR - STD 是用于模拟产生并检查涉及无缝连接的两个播放项目 (PlayItems) 的 AV 流的解码处理的概念上的模式。

如图 1 所示，在播放器 101 中，从读出部分 (DVD 驱动器) 111 以比特率 R_{UD} 读出的 TS (传输流) 文件在读缓冲器 112 中缓存。源数据包从读缓冲器 112 以最大比特率 R_{MAX} 读出到源拆包器。

脉冲振荡器 (27MHz X-tal) 114 产生 27MHz 的脉冲。到达时间时钟计数器 115 是计数该 27MHz 频率脉冲的二进制计数器并且在时间 $t(i)$ 给源拆包器提供到达时间时钟计数器的计数值 $Arrival_time_clock(i)$ 。

一个源数据包具有一个传输包和它的 $arrival_time_stamp$ 。在当前源数据包的 $arrival_time_stamp$ 等于 $arrival_time_clock(i)$ 的 LSB (最小有效比特: 30 比特) 值时, 当前源数据包的传输包从源拆包器 113 输出。TS_recording_rate 是传输流的比特率 (下文中称为“TS”)。图 1 所示的 n 、 TB_n 、 MB_n 、 EB_n 、

TBsys、Bsys、Rxn、Rbxn、Rxsys、Dn、Dsys、On 和 Pn(k)的符号与在 ISO/IEC 13818 - 1 (MPEG2 系统规范) T-STD (由 ISO/IEC 13818 - 1 指定的传输流系统目标解码器) 中定义的一样。

接下来将描述上述传统播放器 101 中的解码处理。首先, 将描述单个 DVR MPEG2 TS 再现期间的解码处理。在单个 DVR MPEG2 TS 再现期间, 传输包从输出部分 110 输出从而被输入到 DVR - STD (即解码器 120) 的 TB1、TBn 或 TBsys 的定时由源数据包的 arrival_time_stamp 确定。涉及 TB1、MB1、EB1、TBn、Bn、TBsys 和 TBsys 的缓存操作的规范与由 ISO/IEC 13818 - 1 规定的 T-STD 中的情况相同。与解码和呈现操作相关的规范也与由 ISO/IEC 13818 - 1 规定的 T-STD 中的情况相同。

下面, 将描述无缝连接的播放项目再现期间的解码处理。这里, 将描述涉及无缝连接的播放项目的前一个流 TS1 和当前流 TS2 的再现。

在特定 AV 流(TS1)和接下来的与该 AV 流(TS1)无缝连接的 AV 流(TS2)之间切换期间, TS2 到达时基的时间轴与 TS1 到达时基的时间轴不同。而且 TS2 系统时基的时间轴与 TS1 系统时基的时间轴不同。视频图像的呈现需要无缝地继续下去。重叠可能在音频呈现单元的呈现时间存在。

接着, 将描述从源拆包器读出到 DVR - STD 的传输包的输入定时。

(1) 在 DVR - STD 的 TS1 到 TB1 的最后的视频包的输入完成的时刻 T1 之前

在时刻 T1 之前, 对 DVR - STD 的缓冲器 TB1、TB 或 TBsys 的输入定时由 TS1 的源数据包的 arrival_time_stamp 确定。

(2) 从时刻 T1 到 TS1 的剩余数据包的最后字节的输入完成的时刻 T2

TS1 的剩余数据包必须以 TS_recording_rate (TS1) 的比特率 (TS1 的最大比特率) 输入到 DVR - STD 的缓冲器 TBn 或 TBsys。TS_recording_rate (TS1) 是由对应于剪辑 1 的 ClipInfo () 定义的 TS_recording_rate 的值。TS1 的最后字节输入到缓冲器的时间是时刻 T2。因此, 从时刻 T1 到时刻 T2, 对源数据包的 arrival_time_stamp 忽略不计。

假定 N_1 是跟随着 TS1 的最后视频包的 TS1 的传输包的字节数, T1 和 T2 之间的时间 (时间 $T_{2-1} = T2 - T1$) 是完成以 TS_recording_rate (TS1) 的比特率的 N_1 字节输入所需的时间, 并且用以下的等式 (1) 表示。

$$T_{2-1} = T2 - T1 = N_1 / \text{TS_recording_rate} (\text{TS1}) \dots (1)$$

从时刻 T1 到时刻 T2, 图 1 所示的 Rxn 和 Rxsys 的值改变为 TS_recording_rate (TS1) 的值。除了以上的规则, 缓存操作与 T-STD 的相同。

因为图 1 所示的 Rxn 和 Rxsys 的值改变为时刻 T1 和 T2 之间的 TS_recording_rate (TS1) 的值, 除了由 T-STD 定义的缓存量, 还需要附加的缓存量 (对应于大约 1 秒的数据量), 以便音频解码器能够处理时间 T1 和 T2 之间的输入数据。

(3) 时刻 T2 之后

在时刻 T2, 到达时间时钟计数器 115 复位到 TS2 的第一源包的 arrival_time_stamp 的值。对 DVR-STD 的缓冲器 TB1、TBn 或 TBsys 的输入定时由 TS2 的源数据包的 arrival_time_stamp 确定。Rxn 和 Rxsys 改变到由 T-STD 确定的值。

接着, 将描述视频呈现定时。视频呈现单元必须通过其连接点被无缝地呈现。

这里假定

STC (系统时间时钟) 1: TS1 系统时基的时间轴

STC2: TS2 系统时基的时间轴 (正确地说, STC2 从 TS2 的第一 PCR (节目 (program) 时钟参考) 输入到 T-STD 时开始)。

按照以下方式确定 STC1 和 STC2 之间的偏移值。

假定

PTS1_{end}: 对应于 TS1 的最后的视频呈现单元的 STC1 上的 PTS

PTS2_{start}: 对应于 TS2 的第一个视频呈现单元的 STC2 上的 PTS

T_{pp}: 最后的视频呈现单元的呈现时间段, 两个系统时基之间的偏移值

STC_delta 由以下的等式 (2) 表示。

$$\text{STC_delta} = \text{PTS1}_{\text{end}} + T_{\text{pp}} - \text{PTS2}_{\text{start}} \dots (2)$$

接着, 将描述音频呈现定时。音频呈现单元的呈现定时的重叠可能出现在 TS1 和 TS2 的连接点上, 重叠从 0 到小于 2 的音频帧。播放器 101 必须选择一个音频样本, 并将音频呈现单元的呈现与连接点后的校正的时基再同步。

下面将说明当时间从 TS1 向与 TS1 无缝连接的 TS2 切换时, 由播放器 101 执行的 DVR-STD 的系统时间时钟控制过程。在呈现 TS1 的最后的音频呈现单元的时刻 T5, 系统时间时钟可能在时间 T2 和 T5 之间重叠。在时刻 T2 和

T5 之间，DVR - STD 从旧的时基值 (STC1) 向新的时基值 (STC2) 切换系统时间时钟。STC2 值能够由以下的等式 (3) 表示。

$$\text{STC2} = \text{STC1} - \text{STC_delta} \dots (3)$$

下面说明时间从 TS1 向与 TS1 无缝连接的 TS2 切换时，TS1 和 TS2 必须满足的编码条件。

假定

$\text{STC1}^1_{\text{video_end}}$: 当 TS1 的最后的视频包的最后字节到达 DVR - STD 的 TB1 时，系统时基 STC1 上的 STC 值

$\text{STC2}^2_{\text{video_start}}$: 当 TS2 的第一个的视频包的第一个字节到达 DVR - STD 的 TB1 时，系统时基 STC2 上的 STC 值

$\text{STC2}^1_{\text{video_end}}$: 通过把 $\text{STC1}^1_{\text{video_end}}$ 的值变换到系统时基 STC2 上的值所获得的值。

在这种情况下， $\text{STC2}^1_{\text{video_end}}$ 用以下等式 (4) 表示。

$$\text{STC2}^1_{\text{video_end}} = \text{STC1}^1_{\text{video_end}} - \text{STC_delta} \dots (4)$$

为了解码器 120 与 DVR - STD 兼容，必须满足以下两个条件。

(条件 1)

TS2 的第一个视频包到达 TB1 的定时必须符合以下的不等式 (5)。

$$\text{STC2}^2_{\text{video_start}} > \text{STC2}^1_{\text{video_end}} + T_{2-1} \dots (5)$$

剪辑 1 和/或剪辑 2 的部分流需要再次编码和/或再次复用，从而满足以上不等式 (5)。

(条件 2)

在通过变换 STC1 和 STC2 到互相相同的时间轴获得的系统时基的时间轴上，来自 TS1 的视频包输入和来自 TS2 的视频包的随后的输入不应当从视频缓冲器中上溢和下溢。

但是，如上所述，使用 DVR - STD 模式的传统播放器 101 能够处理在时刻 T1 和 T2 之间的输入数据。也就是说，因为 TS1 的剩余包以时刻 T1 和 T2 之间的 $\text{TS_recording_rate}(\text{TS1})$ 的比特率 (TS1 的最大比特率) 输入到 DVR - STD 的缓冲器 TBn 或 TBsys，所以，除了由 T - STD 确定的缓存量之外，还需要具有能够缓存对应于大约 1 秒的数据量的容量的附加的缓冲器。

该缓冲器容量基于以下因素。也就是说，在 MPEG2 TS 中，在与对应于特定字节位置的视频数据同步地再现的音频数据能够与预定区域内的复用相

位差分离存在，并且该复用相位差的最大值等于对应于1秒的数据量。因此，以上等式(1)的N1的最大值等于对应于多达1秒的音频数据。在时间T1和T2之间，忽略源数据包的arrival_time_stamp并且对应于N1数据量的源数据包以TS的最大比特率输入到音频缓冲器。因此，除了由T-STD确定的缓存量之外，还需要附加缓存量(对应于大约1秒的数据量)。

该附加缓冲器的量能够按照如下计算。也就是说，在按照杜比AC-3以例如640kbps编码的音频流的情况下，对应于1秒的音频数据是80k字节(=640k比特)。结果，需要80k字节的附加缓冲器。

在按照线性PCM方法编码的音频流的情况下(24比特样本，96kHz取样频率，8信道)，对应于1秒的音频数据大约是18M比特(=24比特样本×96,000样本/秒×8信道)。结果，需要大约3M字节的附加缓冲器。因而，在使用以上多信道音频数据的情况下，附加缓冲器的容量变得极其大。

发明内容

在考虑到传统的情况条件下提出本发明，并且其目的是提供一种信息处理设备、它的方法、程序和记录介质，所述信息处理设备包括音频缓冲器，音频缓冲器具有用于实现两个复用流的无缝解码的优化容量，其中在两个复用流的每一个中复用音频流和视频流，以及还提供产生对应于音频缓冲器容量的复用流的信息处理设备、它的方法和记录复用流的记录介质。

为了达到以上目的，根据本发明的一种信息处理设备，它解码包括由每个具有传输包及其到达时间印记的多个源数据包组成的数据流的复用流，其中第二复用流的第一个图像作为第二图像，连接到作为第一图像的第一复用流的最后一个图像，从而无缝再现，该信息处理设备包括：输出装置，用于根据复用流的到达时间印记输出源数据包；视频缓冲器，用于缓存包括在源数据包中的视频数据；音频缓冲器，用于缓存包括在源数据包中的音频数据；视频解码装置，用于解码在视频缓冲器中缓存的视频数据；和音频解码装置，用于解码在音频缓冲器中缓存的音频数据，其中音频缓冲器具有能够缓存与输入第二图像到视频缓冲器所需时间对应的音频数据的容量，其中满足

$EBn_max = (I_max/Rv) \times Ra$ ，其中EBn_max比特是音频缓冲器所需的容量；I_max比特是第二图像的比特量，Rv比特/秒是对视频缓冲器的输入比特率，并且Ra比特/秒是音频数据的比特率。本发明中，音频缓冲器具有能

够缓存与输入第二图像到视频缓冲器所需时间对应的音频数据的容量，并且即使在输入第一图像到视频缓冲器完成的时刻和输入第一复用流的最后的源数据包完成的时刻之间，源数据包也根据复用流中的源数据包的 arrival_time_stamp 输入到缓冲器。这去掉了传统上使用被忽略源数据包的 arrival_time_stamp 以 TS 的最大比特率输入传输包所需的 1 秒相对应的附加缓冲器。而且，有可能在已经输入第一复用流的最后的传输包后以其解码定时，输入首先在第二复用流中将被解码的图像到视频缓冲器。

而且能够满足 $EBn_max = (I_max/Rv) \times Ra$ ，其中 EBn_max 比特是音频缓冲器所需的容量； I_max 比特是第二图像的比特量， Rv 比特/秒是对视频缓冲器的输入比特率，并且 Ra 比特/秒是音频数据的比特率。假定 I_max 是例如 I 图像（即第二图像）的比特量，音频缓冲器的容量能够设置为至多 $EBn_max = (I_max/Rv) \times Ra$ 。

最好音频缓冲器具有能够缓存对应于 100 毫秒的音频数据的容量。因为根据 MPEG2 规范的 I 图像的数据量是通常在 1 秒中发送的数据量的 10% 或更少，有可能通过设置音频缓冲器的容量到对应于 100 毫秒的大小，以其解码定时输入 I 图像到视频缓冲器，从而允许音频缓冲器将音频数据向前移动那个数量。结果，降低了视频数据中的编码限制。也就是说，通过设置音频缓冲器的容量到上述大小，有可能形成复用流，以便音频数据的输入比其再现定时早 100 毫秒完成。

而且，假设 STC_delta 是在第一复用流的时间轴上的第一图像的呈现结束时间和在第二复用流的时间轴上的第二图像的呈现开始时间之间的时间差， $STC2^1_{end} (= STC1^1_{end} - STC_delta)$ 是通过使用时间差 STC_delta 转换 $STC1^1_{end}$ 到第二复用流的时间轴上的值而获得的值， $STC1^1_{end}$ 是在第一复用流的最后的源数据包的最后字节从输出装置输出时的第一复用流的时间轴上的值，并且 $STC2^2_{start}$ 是在第二复用流的第一个源数据包的第一字节从输出装置输出时的第二复用流的时间轴上的值。

而且，最好配置信息处理装置，使得在第一复用流的最后的源数据包从输出装置输出后经过预定时间增量 δ_1 之后，第二复用流的第一源数据包从输出装置输出。在这种情况下，满足 $STC2^2_{start} > STC2^1_{end}$ 。结果第二复用流的第一源数据包输入定时的确定变得更灵活，这使得很容易编码第二复用流。

而且假定 STC_delta 是第一复用流的时间轴上的第一图像的呈现结束时

间和第二复用流的时间轴上的第二图像的呈现开始时间之间的时间差，能够配置信息处理设备，使得在第一复用流的最后一个源数据包输出开始后经过预定时间 ATC_delta 后，第二复用流的第一个源数据包从输出装置输出，预定时间 ATC_delta 这样确定，使得满足时间差 STC_delta ，并且这样形成复用流使得满足时间差 STC_delta 。结果，第二复用流的第一源数据包输入定时的确定变得更灵活，这使得很容易编码第二复用流。

这种情况下，有可能作为第一复用流的附属信息管理预定的时间 ATC_delta 。

根据本发明的信息处理方法，解码包括由每个具有传输包及其达到时间印记的多个源数据包组成的数据流的复用流，并且其中第二复用流的第一个图像作为第二图像，与作为第一图像的第一复用流的最后一个图像连接，从而无缝再现，该信息处理方法包括：根据复用流的到达时间印记输出源数据包的步骤；缓存分别在视频和音频缓冲器中的源数据包中包含的视频和音频数据的步骤；和对视频和音频缓冲器中缓存的视频和音频数据解码的步骤，其中在缓存步骤中，在第二图像被缓存在视频缓冲器中之前，与输入第二图像到视频缓冲器所需时间对应的音频数据被缓存在音频缓冲器中，其中满足 $EBn_max = (I_max/Rv) \times Ra$ ，其中 EBn_max 比特是音频缓冲器所需的容量； I_max 比特是第二图像的比特量， Rv 比特/秒是对视频缓冲器的输入比特率，并且 Ra 比特/秒是音频数据的比特率。根据本发明的一种程序，允许计算机执行上述的信息处理。根据本发明的记录介质是记录该程序的计算机可读记录介质。

根据本发明的另一种记录介质，它记录包括由每个具有传输包及其达到时间印记的多个源数据包组成的数据流的复用流，其中复用流这样形成，使得作为第二复用流的第一个图像的第二图像与作为第一复用流的最后一个图像的第一图像连接，从而无缝再现，第一和第二复用流能够根据它们各自的到达时间印记输入到解码器，并且与输入第二图像到解码器所需时间对应的音频数据的输入能够在第二图像向解码器的输入开始时完成。

在本发明中，形成复用流使得与输入第二图像到解码器所需的时间对应的音频数据的输入在第二图像向解码器输入开始时完成。结果，能在第一复用流的最后一个传输包通过对复用流解码输入后，使用具有能够缓存与输入第二图像到视频缓冲器所需时间对应的音频数据的容量的音频缓冲器的解码

器，通过其解码定时输入首先在第二复用流中解码的图像到视频缓冲器。

根据本发明的一种信息处理设备，产生包括由每个具有传输包及其到达时间印记的多个源数据包组成的数据流的复用流，并且它由解码器基于到达时间印记读出并解码，该信息处理设备包括：视频编码装置，用于产生第一视频编码流以结束第一图像的呈现和产生第二视频编码流，使得在第一图像之后立即开始将要呈现的第二图像的呈现；和复用装置，用于复用第一视频编码流和与第一视频编码流同步的音频编码流，以产生第一复用流，复用第二视频编码流和与第二视频编码流同步的音频编码流，以产生第二复用流；并产生复用流，在该复用流中第二复用流的第一个图像作为第二图像，连接到作为第一图像的第一复用流的最后一个图像，以便无缝再现，以及音频缓冲器，其容量满足 $EBn_max = (I_max/Rv) \times Ra$ ，其中 EBn_max 比特是音频缓冲器所需的容量； I_max 比特是第二图像的比特量， Rv 比特/秒是对视频缓冲器的输入比特率，并且 Ra 比特/秒是音频数据的比特率，使得与输入第二图像到解码器所需的时间对应的音频数据的输入能够在第二图像向解码器输入开始时完成。在本发明中，复用这样执行，使得与例如 100 毫秒对应的音频数据的输入在第二图像向解码器输入开始时完成，100 毫秒对应于输入第二图像到解码器所需的时间。结果，在解码器中，音频数据在前移动到音频缓冲器，充分确保通过其解码定时发送第二图像，比如 I 图像的时间，这使得对复用流编码更容易。

根据本发明的另一种信息处理方法，产生包括由每个具有传输包及其到达时间印记的多个源数据包组成的数据流的复用流，并且它由解码器基于到达时间印记读出并解码，该信息处理方法包括：产生步骤，产生第一视频编码流以结束第一图像的呈现和产生第二视频编码流使得在第一图像之后立即开始将要呈现的第二图像的呈现；和复用步骤，复用第一视频编码流和与第一视频编码流同步的音频编码流，以产生第一复用流，复用第二视频编码流和与第二视频编码流同步的音频编码流以产生第二复用流，并产生一个复用流，该复用流中，第二复用流的第一个图像作为第二图像，连接到作为第一图像的第一复用流的最后一个图像，以便无缝再现，缓冲音频的步骤，其中缓冲量满足 $EBn_max = (I_max/Rv) \times Ra$ ，其中 EBn_max 比特是音频缓冲器所需的容量； I_max 比特是第二图像的比特量， Rv 比特/秒是对缓冲视频的输入比特率，并且 Ra 比特/秒是音频数据的比特率，使得与输入第二图像

到解码器所需的时间对应的音频数据的输入能够在第二图像向解码器输入开始时完成。本发明的以上和其他目的、优点和特征将通过与附图结合进行以下说明变得更明显。

附图说明

图 1 是表示传统的信息处理设备的框图；

图 2 是表示在使用桥式剪辑 (Bridge-Clip) 的情况下前一个播放项目和当前播放项目之间的关系模式的图；

图 3 是表示在不使用桥式剪辑的情况下前一个播放项目和当前播放项目之间的关系模式的图；

图 4 是表示以画面的呈现顺序的剪辑 1 和剪辑 2 的模式图，它们是将要互相无缝连接的视频流；

图 5 (a)、5 (b) 和 5 (c) 表示每个 AV 流中的数据流的例子，其中图 4 所示的视频流（剪辑 1 和剪辑 2）使用桥式序列（第一方法）无缝连接；

图 6 示出了每个 AV 流中的数据流的例子，其中图 4 所示的视频流（剪辑 1 和剪辑 2）通过不使用桥式序列的第二方法无缝连接；

图 7 是用于解释音频呈现的重叠的图，并示出了 TS1 和 TS2 中视频呈现单元和音频呈现单元的一种模式；

图 8 是表示根据本发明的实施例的信息处理设备的框图；

图 9 是在特定 AV 流（TS1）和与该 AV 流（TS1）无缝连接的下一个 AV 流（TS2）之间切换期间传输包的输入、解码和呈现的时序图；

图 10 是在特定 AV 流（TS1）和与该 AV 流（TS1）无缝连接的下一个 AV 流（TS2）之间切换期间传输包的输入、解码和呈现的时序图的另一个例子；

图 11 是在特定 AV 流（TS1）和与该 AV 流（TS1）无缝连接的下一个 AV 流（TS2）之间切换期间传输包的输入、解码和呈现的时序图的另一个例子；

图 12 示出了用于存储 ATC_delta 的附属信息 ClipInfo（）的数据格式；

图 13 是表示在多个 AV 流（TS2）将连接到特定 AV 流（TS1）的情况下附属信息 ClipInfo（）的模式图；

图 14 (a) 和 14 (b) 是表示在传统的 DVR - STD 中，在音频缓冲器的容量是 4k 字节的情况下，TS1 和无缝连接到 TS1 的 TS2 之间切换期间，DVR - STD 的视频缓冲器和音频缓冲器的比特占用量改变的例子的图；

图 15 (a) 和 15 (b) 是用于解释本发明实施例中的优点的图和表示在音频缓冲器的大小是 8k 字节的情况下，TS1 和无缝连接到 TS1 的 TS2 之间切换期间，DVR - STD 的视频缓冲器和音频缓冲器的比特占用量改变的例子的图。

具体实施方式

将在下文中参照附图详细说明本发明的一个实施例。该实施例通过把本发明应用到信息处理设备中获得，该信息处理设备以无缝方式连续再现由视频和音频流组成的两个复用的 AV 流。本发明将提出一种音频缓冲器，它具有能够在 DVR - STD（数字视频记录 - 系统目标解码器）中提供的互相无缝

连接的两个 AV 流再现时可用的优化的容量。

首先，定义在以下说明中使用的术语。“剪辑”表示视频和音频流的复用流。“播放列表”表示剪辑中的一组再现区域。某剪辑中的一个再现区域称为“播放项目”，播放项目”由时间轴上的一对 IN 点和 OUT 点表示。即，播放列表是一组播放项目。

“以无缝的方式再现播放项目”表示再现设备（播放器）呈现（再现）记录在盘上的音频/视频数据，同时防止在解码器的再现输出中的间隙或暂停的状态。

下面说明互相无缝连接的两个播放项目的结构。是否确保前一个播放项目和当前播放项目的无缝呈现能够通过当前播放项目中定义的 connection_condition 字段判断。作为用于实现播放项目之间无缝连接的方法，存在使用桥式剪辑（桥式序列）的一种方法（第一方法）和不使用桥式剪辑的另一种方法（第二方法）。

首先，将说明前一个播放项目和当前播放项目用桥式序列（第一方法）连接的情况下的 TS1 和 TS2。图 2 是表示使用桥式剪辑（第一方法）的情况下前一个播放项目（PI1）和当前播放项目（PI2）之间的关系模式的图。在图 2 中，在使用桥式剪辑的情况下将由播放器读出的流数据以阴影方式表示。DVR MPEG(活动图像专家组)2 传输流(TS)由整数个线列单元(aligned unit)组成。线列单元具有 6,144 字节(2048×3 字节)。一个线列单元包括 32 个源数据包并开始于源数据包的第一字节。

每个源数据包的长度为 192 字节。源数据包由 4 字节的 TP_extra_header 和 188 字节的传输包组成。TP_extra_header 具有 copy_permission_indicator 和 arrival_time_stamp。copy_permission_indicator 是表示传输包的有效载荷的复制限制的整数。arrival_time_stamp (ATS) 是表示 AV 流中对应的传输包到达解码器时的时间印记。基于构成 AV 流的每个源数据包的 arrival_time_stamp 的时间轴称为到达时基，并且它的时钟称为 ATC (到达时间时钟)。

图 2 所示的 TS1 (第一复用流) 由作为剪辑 1 (剪辑 AV 流) 的阴影部分的流数据 D1 和作为在阴影桥式剪辑中的 SPN_arrival_time_discontinuity 之前的部分的流数据 D2 组成。SPN_arrival_time_discontinuity 表示源数据包的地址，在该地址上桥式剪辑 AV 流文件中存在到达时基的不连续。

与剪辑 1 的阴影部分对应的 TS1 中的流数据 D1 是从解码与前一个播放项

目的 IN_time (图 2 中的 IN_time 1) 对应的呈现单元所需的流地址到 SPN_exit_from_previous_Clip 涉及的源数据包的流数据。

作为有阴影的桥式剪辑的 SPN_arrival_time_discontinuity 之前的部分的 TS1 中的流数据 D2 是从桥式剪辑中的第一源数据包到 SPN_arrival_time_discontinuity 涉及的源数据包紧接着前面的源数据包的流数据。

图 2 所示的 TS2 (第二复用流) 由作为剪辑 2 (剪辑 AV 流) 的阴影部分的数据流 D4 和作为有阴影的桥式剪辑中的 SPN_arrival_time_discontinuity 后面的一部分的流数据 D3 组成。

作为有阴影的桥式剪辑中的 SPN_arrival_time_discontinuity 后面的一部分的 TS2 中的流数据 D3 是从 SPN_arrival_time_discontinuity 涉及的源数据包到桥式剪辑的最后一个源数据包的流数据。

作为剪辑 2 的阴影部分的 TS2 中的流数据 D4 是从 SPN_enter_to_current_Clip 涉及的源数据包到解码与当前播放项目的 OUT_time (图 2 中的 OUT_time 2) 对应的呈现单元所需的流地址的流数据。

接着, 描述在通过不使用桥式序列的第二方法互相连接前一个播放项目和当前播放项目的情况下的 TS1 和 TS2。图 3 是表示通过不使用桥式序列的第二方法互相连接的前一个播放项目 (PI1) 和当前播放项目 (PI2) 之间的关系的模式的图。图 3 中, 将由播放器读出的流数据以阴影方式示出。

图 3 所示的 TS1 (第一复用流) 包括流数据 D5, 它是剪辑 1 (剪辑 AV 流) 的阴影部分。作为剪辑 1 的阴影部分的 TS1 中的流数据 D5 是从解码与前一个播放项目的 IN_time (图 3 中的 IN_time) 对应的呈现单元所需流地址到剪辑 1 的最后一个源数据包的数据。

图 3 所示的 TS2 (第二复用流) 由流数据 D6 组成, 它是剪辑 2 (剪辑 AV 流) 的阴影部分。作为剪辑 2 的阴影部分的 TS2 中的流数据 D6 是从剪辑 2 的第一源数据包到解码与当前播放项目的 OUT_time (图 3 中的 OUT_time 2) 对应的呈现单元所需的流地址的流数据。

图 2 和 3 中, TS1 和 TS2 构成数据流, 其中源数据包是连续的。接着, 说明 TS1 和 TS2 的流规则和它们之间的连接条件。

尽管 TS1 和 TS2 是通过复用视频和音频流获得的, 这里将说明无缝连接编码限制中的视频比特流的限制。

图 4 是表示以画面呈现顺序，作为互相无缝连接的视频流的剪辑 1 和剪辑 2 的模式图。当视频图像节目跳过一部分进行再现时，有必要在解码设备中执行视频流的再编码处理，从而无缝连接位于用作跳过再现的开始点的跳出点图像之前的出点一侧的节目和位于用作跳过再现的到达点的跳入时间图像之后的入点一侧的节目。

GOP (图像组)，根据 MPEG 规范的图像单元包括三种类型的编码后的图像：一个或多个 I (内部) 图像 (帧内编码图像)，它们每个都是参考图像，其中已经编码了图像而不需要根据其他图像进行预测编码；P (预测) 图像，它们每个都是通过用与呈现顺序相同的方向的预测编码对图像进行编码而获得的前向预测编码的图像；和 B (双向) 图像，它们每个都是通过使用前向和反向预测编码而获得的双向预测编码的图像。在图 4 中，剪辑 1 和剪辑 2 的每个标号表示呈现顺序，并且 I、P、B 或 i、p、b 表示图像类型。图 4 示出了连接剪辑 1 的 B7 和剪辑 2 的 b4 的情况。为了在连接点无缝地呈现视频流，位于 OUT_time 1 (剪辑 1 的 OUT_time) 之后和 IN_time 2 (剪辑 2 的 IN_time) 之前的不必要的图像必须通过编码连接点附近的剪辑的局部流的处理来消除。

图 5 (a)、5 (b) 和 5 (c) 示出了一个例子，其中图 4 所示的视频流 (剪辑 1 和剪辑 2) 使用桥式序列 (第一方法) 无缝连接。

SPN_arrival_time_discontinuity 之前的桥式剪辑的视频流由包括图 4 所示的剪辑 1 的 OUT_time 1 之前的图像的编码视频流组成。视频流连接到剪辑 1 的前一个视频流以后，两个视频流再次编码为符合 MPEG2 规范的一个连续的基本流。同样，SPN_arrival_time_discontinuity 后的桥式剪辑的视频流由包括图 4 所示的剪辑 2 的 IN_time 2 之后的图像的视频流组成。其解码能够正确开始的视频流连接到剪辑 2 的后续的视频流。在那之后，所连接的两个视频流再次编码为一个符合 MPEG 规范的连续的基本流。为了建立桥式剪辑，必须再次编码几个图像，并且其他图像能够通过复制原始剪辑来获得。

图 5 (a) 表示按呈现顺序的图 4 所示的剪辑 1。播放器从在前的剪辑 1 的 P5 的源数据包号 (SPN_exit_from_previous_Clip) 跳到图 5 (b) 所示的桥式剪辑。在图 2 所示的桥式剪辑的 D2 中，即在与按照桥式序列的 SPN_arrival_time_discontinuity 之前的视频数据相对应的剪辑 1 的 OUT_time 1 一侧的流数据中，包括直到 B4 的图像的数据 d1 由通过不做改变地复制剪辑

1 获得的数据组成, 并且由正常情况下原始剪辑 1 的图像 B6 和 B7 组成的数据 d2 包括通过解码剪辑 1 为非压缩的图像数据并再次对它们编码而获得的 P7 和 B6。而且, 在图 2 所示的桥式剪辑的 D3 中, 即在与按照桥式序列的 SPN_arrival_time_discontinuity 之后的视频数据相对应的剪辑 2 的 IN_time 2 一侧的流数据中, 原始剪辑 2 的图像 b4、p5、p8、b6、b7 成为通过一次解码剪辑 2 为非压缩的图像数据并再次编码它们而获得的新建立的数据 (i0, p1, p4, b2, b3)。在跳到 SPN_enter_to_current_Clip 之前的数据 d4 通过不改变地复制剪辑 2 而获得。

图 6 示出了一个例子, 其中图 4 所示的视频流 (剪辑 1 和剪辑 2) 通过不使用桥式序列的第二方法无缝连接。在图 6 所示的剪辑 1 和剪辑 2 中, 图像按照呈现顺序排列。即使在不使用桥式序列的情况下, 连接点附近的流被一次解码为非压缩的数据并再次解码优化图像类型, 这与图 5 所示的使用桥式序列的情况相同。也就是说, 剪辑 1 的视频流包括直到与图 4 所示的 OUT_time 1 相对应的图像的编码视频流, 其中原始剪辑 1 的 B6、B7 再次编码为数据 (P7, B6) d5, 从而成为一个符合 MPEG2 规范的连续的基本流。同样, 剪辑 2 的视频流包括与图 4 所示的剪辑 2 的 IN_time 2 相对应的图像之后的编码视频流, 其中原始剪辑 2 的 b4、p5、p8、b6、b7 再次编码为数据 (i0, p1, p4, b2, b3) d6, 从而成为一个符合 MPEG2 规范的连续的基本流。

接着, 将说明 TS1 和 TS2 的复用流的编码限制。图 7 是用于解释音频呈现的重叠的图, 并示出了 TS1 和 TS2 中的视频呈现单元 VPU1 和 VPU2 以及音频呈现单元 APU1 和 APU2 的模式。

如图 7 所示, TS1 的音频流的最后一个音频帧 A_end 包括具有等于 TS1 的最后一个呈现图像的呈现结束时间 (OUT_time1) 的呈现时间的音频样本。TS2 的音频流的第一个音频帧 A_start 包括具有等于 TS2 的第一呈现图像的呈现开始时间 (IN_time2) 的呈现时间的音频样本。因此, 在 TS1 和 TS2 之间的连接点的音频呈现单元序列中不存在间隙, 并且产生由不超过 2 个音频帧的音频呈现单元的长度确定的音频重叠。在连接点的 TS 是根据 DVR - STD (数字视频记录 - 系统目标解码器) 的 DVR MPEG2 TS, 这将在下文中说明。

DVR - STD 是用于模拟产生并检查已经互相无缝连接的两个播放项目涉及的 AV 流中的解码处理的概念模式。图 8 示出了 DVR - STD 模式 (DVR MPEG2 传输流播放器模式)。

如图 8 所示, 根据该实施例的信息处理设备 (DVR MPEG2 传输流播放器模式, 下文中称为“播放器”) 1 包括从为了无缝再现而连接的 TS 读出传输包并输出它们的输出部分 10 和解码来自输出部分 10 的传输包的解码器 (DVR - STD) 20。如后面描述的, 已经通过改变前述传统的 DVR - STD 的传输包输入定时和音频缓冲器的容量获得了解码器 20。在输出部分 10 中, 以读出速率 R_{UD} 从读出部分 (DVR 驱动器) 11 读出的 TS 文件在读缓冲器 12 中缓存。源数据包以比特率 R_{MAX} 从读缓冲器 12 读出到源拆包器 13。 R_{MAX} 是源数据包流的比特率。

脉冲振荡器 (27MHz X - tal) 14 产生 27MHz 脉冲。到达时间时钟计数器 15 是计数 27MHz 频率脉冲的二进制计数器, 并为源拆包器 13 提供 $Arrival_time_clock(i)$, 它是时刻 $t(i)$ 到达时间时钟计数器的计数值。

如上所述, 一个源数据包包括一个传输包和它的 $arrival_time_stamp$ 。在当前源数据包的 $arrival_time_stamp$ 等于 $arrival_time_clock(i)$ 的 LSB 30 比特的值时, 当前源数据包的传输包从源拆包器 13 输出。 $TS_recording_rate$ 是 TS 的比特率。

图 8 所示的符号 n 、 TB_n 、 MB_n 、 EB_n 、 TB_{sys} 、 B_{sys} 、 R_{xn} 、 R_{bxn} 、 R_{xsys} 、 D_n 、 D_{sys} 、 O_n 和 $P_n(k)$ 与 ISO/IEC13818 - 1 (MPEG2 系统规范) 的 T - STD (由 ISO/IEC 13818 - 1 规定的传输流系统目标解码器) 定义的那些相同。即如下所示。

n : 基本流的索引号

TB_n : 基本流 n 的传输缓冲器

MB_n (只在视频流中存在): 基本流 n 的复用缓冲器

EB_n : 只在视频流中存在的基本流 n 的基本流缓冲器

TB_{sys} : 用于被解码的节目的系统信息的输入缓冲器

B_{sys} : 用于被解码的节目的系统信息的系统目标解码器中的主缓冲器

R_{xn} : 从 TB_n 除去数据的传输率

R_{bxn} (只存在于视频流中): 从 MB_n 除去 PES 包有效载荷的传输率

R_{xsys} : 从 TB_{sys} 除去数据的传输率

D_n : 基本流 n 的解码器

D_{sys} : 涉及被解码的节目的系统信息的解码器

O_n : 视频流 n 的重新排序缓冲器

$P_n(k)$: 基本流 n 的第 k 个呈现单元

接着, 将描述解码器 20 的解码处理。首先, 将描述单个 DVR MPEG2 TS 再现期间的解码处理。

在单个 DVR MPEG2 TS 再现期间, 传输包输入到缓冲器 TB_1 、 TB_n 和 TB_{sys} 的定时由源数据包的 $arrival_time_stamp$ 确定。

TB_1 、 MB_1 、 EB_1 、 TB_n 、 B_n 、 TB_{sys} 和 B_{sys} 的缓存操作以与 ISO/IEC 13818-1 规定的 T-STD 中相同的方式规定。其解码和呈现操作也以在 ISO/IEC 13818-1 规定的 T-STD 中相同的方式规定。

接着, 将描述无缝连接的播放项目的再现期间的解码处理。图 9 是在特定 AV 流 (TS1) 和与该 AV 流 (TS1) 无缝连接的下一个 AV 流 (TS2) 之间切换期间传输包的输入、解码和呈现的时序图。

这里, 给出了无缝连接的播放项目涉及的两个 AV 流的描述。在后面的说明中, 将说明如图 2 和图 3 所示的已经无缝连接的 TS1 和 TS2 的再现。也就是说, TS1 是在前的流, TS2 是当前流。TS1 和 TS2 分割的各个包表示 TS1 和 TS2 的源数据包 SP1 和 SP2。

在特定 AV 流 (TS1) 和与该 AV 流 (TS1) 无缝连接的下一个 AV 流 (TS2) 之间切换期间, TS2 到达时基的时间轴 (图 9 中的 ATC2) 与 TS1 到达时基的时间轴 (图 9 中的 ATC1) 不相同。而且 TS2 系统时基的时间轴 (图 9 中的 STC2) 与 TS1 系统时基的时间轴 (图 9 中的 STC1) 不相同。视频图像的呈现需要是无缝的。重叠可能在音频呈现单元的呈现时间中存在。

在本实施例的播放器 1 中, 具有优化容量的音频缓冲器通过改变关于前述播放器 101 的随后两个点而获得, 前述播放器 101 在日本专利申请公开号 2000-175152, 日本专利申请公开号 2001-544118 和日本专利申请公开号 2002-158974 中有说明。该说明从第一变化点开始。第一变化点是特定 AV 流 (TS1) 和与该 AV 流 (TS1) 无缝连接的下一个 AV 流 (TS2) 之间切换期间, 直到最后一个包的 TS1 的包向解码器 20 的输入由它们的源数据包的 $arrival_time_stamp$ 确定。

也就是说, 如上所述, 在传统的播放器 101 中, 传输包以具有 $arrival_time_stamp$ 的 TS 的最大比特率输入到缓冲器中, 该 $arrival_time_stamp$ 在 TS1 的最后的视频包已经输入到 TB_1 的时刻 T1 和 TS1 的最后字节的输入已经完成的时刻 T2 之间被忽略, 然而在本实施例中, T1 和 T2 之间的源数据

包的输入由 TS1 的源数据包的 arrival_time_stamp 确定，与时刻 T1 之前的情况一样。这消除了与传统上使用被忽略的源数据包的 arrival_time_stamp 以 TS 的最大比特率 R_{max} 输入传输包所需的 1 秒时间相对应的附加缓冲器。

将参照图 9 说明在这种情况下对解码器 20 的输入定时。

(1) 在时刻 T1 之前

在时刻 T1 之前，即，直到 TS1 的最后的视频包输入到解码器 20 完成，解码器 20 的缓冲器 TB1、TBn 或 TBsys 的输入定时都由 TS1 的源数据包 SP1 的 arrival_time_stamp 确定。

(2) 从时刻 T1 到时刻 T2

TS1 的剩余的数据包向解码器 20 的输入定时也由 TS1 的源数据包 SP1 的 arrival_time_stamp 确定。当 TS1 的最后字节输入到缓冲器的时刻是时刻 T2。

(3) 时刻 T2 之后

在时刻 T2，到达时间时钟计数器 15 复位到 TS2 的第一源数据包的 arrival_time_stamp 的值。对解码器 20 的缓冲器 TB1、TBn 或 TBsys 的输入定时由 TS2 的源数据包 SP2 的 arrival_time_stamp 确定。

即，在 TS1 的最后字节向解码器 20 的输入完成的时刻 T2 之前，解码器 20 的缓冲器 TB1、TBn 或 TBsys 的输入定时由 TS1 的源数据包 SP1 的 arrival_time_stamp 确定，并且在时刻 T2 之后由 TS2 的源数据包 SP2 的 arrival_time_stamp 确定。

接着，将说明视频呈现定时。视频呈现单元必须通过图 2 和 3 所示的前述连接点无缝地呈现。即，TS1 的最后的视频数据（第一图像）和 TS2 的第一视频数据（第二图像）被无缝地再现。假定

STC1: TS1 系统时基的时间轴

STC2: TS2 系统时基的时间轴（正确地是，STC2 从 TS2 的第一 PCR（节目时钟参考）已经输入到 T-STD 的时刻开始）。

STC1 和 STC2 之间的偏移值按照如下方式确定。

假定

$PTS1_{end}$: 在对应于 TS1 的最后的视频呈现单元的 STC1 上的 PTS

$PTS2_{start}$: 在对应于 TS2 的第一个视频呈现单元的 STC2 上的 PTS

T_{pp} : TS1 的最后的视频呈现单元的呈现周期，

两个系统时基之间的偏移值 STC_delta 用以下的等式 (6) 表示。

$$STC_delta = PTS1_{end} + T_{pp} - PTS2_{start} \dots (6)$$

接着，将说明音频呈现定时。音频呈现单元的呈现定时的重叠可以存在于 TS1 和 TS2 的连接点上，重叠从 0 到小于 2 个音频帧（参照图 9 的音频重叠）。播放器 101 必须选择一个音频样本并对连接点之后的具有校正后的时基的音频呈现单元的呈现再次同步。

下面说明当时间从 TS1 向与 TS1 无缝连接的 TS2 切换时，用于控制由播放器执行的解码器 20 的系统时间时钟的过程。

在时刻 T5，呈现 TS1 的最后的音频呈现单元。系统时间时钟可以在时刻 T2 和 T5 之间重叠。在时刻 T2 和 T5 之间，解码器 20 从旧的时基值（STC1）向新的时基值（STC2）切换系统时间时钟。STC2 的值能够用以下等式（7）表示。

$$STC2 = STC1 - STC_delta \dots (7)$$

下面说明当时间从 TS1 向与 TS1 无缝连接的 TS2 切换时，TS1 和 TS2 必须符合的编码条件。

假定

STC1¹_{end}: 当 TS1 的最后的数据包的最后字节到达解码器 20 时，系统时基 STC1 上的 STC 的值

STC2²_{start}: 当 TS2 的第一数据包的第一个字节到达解码器 20 时，系统时基 STC2 上的 STC 的值

STC2¹_{end}: 通过把 STC1¹_{end} 值转换为系统时基 STC2 上的值而获得的值。

在这种情况下，STC2¹_{end} 用以下的等式（8）表示。

$$STC2^1_{end} = STC1^1_{end} - STC_delta \dots (8)$$

为了解码器 20 与 DVR - STD 兼容必须符合以下两个条件。

（条件 1）

TS2 的第一数据包到达解码器 20 的定时必须符合以下的不等式（9）。

$$STC2^2_{start} > STC2^1_{end} \dots (9)$$

剪辑 1 和/或剪辑 2 的部分数据流需要再次编码和/或再次复用，从而符合以上的不等式（9）。

（条件 2）

在通过把 STC1 和 STC2 转换为互相相同的时间轴而获得的系统时基的时间轴上，来自 TS1 的视频包和来自 TS2 的随后的视频包的输入不应当使视频

缓冲器上溢和下溢。而且，在通过把 STC1 和 STC2 转换到互相相同的时间轴而获得的系统时基的时间轴上，来自 TS1 的数据包和来自 TS2 的随后的数据包的输入不应当使解码器 20 的所有缓冲器上溢和下溢。

图 10 是在特定的 AV 流 (TS1) 和与 AV 流 (TS1) 无缝连接的下一个 AV 流 (TS2) 之间切换期间，传输包的输入、解码和呈现的时序图的另一个例子。在这种情况下，直到 TS1 的最后一个数据包的数据包向解码器 20 的输入定时也由它们的源数据包 `arrival_time_stamp` 确定。与图 9 所示的时序图不同的一点在于提供预定的时间间隔 (增量 `delta1`: 时刻 T2 和 T2' 之间的间隔)，如图 10 所示，从而消除了对 TS1 的最后一个数据包之后立即输入 TS2 的第一个数据包的需要。结果，TS2 的第一个数据包的输入定时的确定比图 9 的情况更灵活，这使得更容易编码 TS2。

将参照图 10 描述在这种情况下解码器 20 的输入定时。

(1) 时刻 T2 之前

在时刻 T2 之前，即，直到 TS1 的最后的数据包的最后字节向解码器 20 的输入完成，向解码器 20 的缓冲器 TB1、TBn 或 TBsys 的输入定时都由 TS1 的源数据包 SP1 的 `arrival_time_stamp` 确定。

(2) 时刻 T2' 之后

当时间通过时刻 T2 和增量 `delta1` 到达时间 T2' 时，到达时间时钟计数器 15 复位到 TS2 的第一源数据包的 `arrival_time_stamp` 值。解码器 20 的缓冲器 TB1、TBn 或 TBsys 的输入定时由 TS2 的源数据包 SP2 的 `arrival_time_stamp` 确定。

在增量 `delta1` 如图 10 所示提供的情况下，前述的 $STC2_{start}^2$ 和 $STC2_{end}^1$ 必须符合以下关系表达式 (10)。

$$STC2_{start}^2 > STC2_{end}^1 + \text{delta1} \dots (10)$$

图 11 是在特定的 AV 流 (TS1) 和与 AV 流 (TS1) 无缝连接的下一个 AV 流 (TS2) 之间切换期间，传输包的输入、解码和呈现的时序图的另一个例子。还在这种情况下，直到 TS1 的最后一个数据包的数据包向解码器 20 的输入定时也由它们的源数据包的 `arrival_time_stamp` 确定。与图 10 所示的时序图不同的一点在于提供预定的时间间隔 (`ATC_delta`: 时刻 T2 和 T2' 之间的间隔)，如图 11 所示。结果，TS2 的第一个数据包的输入定时的确定比图 9 的情况更灵活，这使得更容易编码 TS2。

将参照图 11 描述在这种情况下解码器 20 的输入定时。

(1) 时刻 T2 之前

在时刻 T2 之前，即，直到 TS1 的最后的的数据包的最后字节向解码器 20 的输入完成，向解码器 20 的缓冲器 TB1、TBn 或 TBsys 的输入定时由 TS1 的源数据包 SP1 的 arrival_time_stamp 确定。

(2) 从时刻 T2 到时刻 T2'

时刻 T2' 是 TS2 的第一数据包输入到解码器 20 的时刻。ATC_delta 是从 TS1 的最后的的数据包的 arrival_time_stamp (ATC1 的时刻) 到在 ATC1 上投影的时刻 T2' 的偏移时间。

(3) 时刻 T2' 之后

在时刻 T2'，到达时间时钟计数器 15 复位到 TS2 的第一源数据包的 arrival_time_stamp 的值。解码器 20 的缓冲器 TB1、TBn 或 TBsys 的输入定时由 TS2 的源数据包 SP2 的 arrival_time_stamp 确定。

这样确定 ATC_delta 的值以使得满足上述等式 (6) 的 STC_delta。

ATC_delta 的值作为流数据的附属信息进行管理。当 TS1 和 TS2 如图 11 所示互相无缝连接时，ATC_delta 的值作为 TS1 的附属信息进行管理。

图 12 示出了用于存储 ATC_delta 的附属信息 ClipInfo () 的数据格式。

在图 12 中，is_ATC_delta 是表示 ClipInfo () 是否具有 ATC_delta 的值的标记。多个值能够在 ClipInfo () 中寄存以允许如图 13 所示多个 TS2 连接到 TS1。当 is_ATC_delta 标记是 1 时，number_of_ATC_delta_entries 表示已经在 ClipInfo () 中寄存的 ATC_delta 的数量。

而且，在图 12 中，following_Clip_Information_file_name 是将连接到 TS1 的 TS2 的流的名字。当存在多个对应于 following_Clip_Information_file_name 的 TS2 时，对应于各个 TS2 的 ATC_delta 的值寄存在 ClipInfo () 中。

当 TS1 和 TS2 输入到图 8 中的 DVR - STD 模式时，输入其复用流和它们的附属信息 ClipInfo ()。ClipInfo () 包括上述的 ATC_delta 的信息，它用如上所述的在 TS1 和 TS2 之间切换的 DVR - STD 模式的控制器 (图 8 中未示出) 以预定的方法处理。

第二个变化点是解码器 20 的大小改变为满足以下条件的足够大的大小。条件是在 TS1 和与 TS1 无缝连接的 TS2 之间切换期间，TS1 的最后的传输包的输入已经完成之后，首先在 TS2 中将被解码的图像 (I 图像) 能够按照其

中的解码定时输入到视频缓冲器。

符合以上条件需要的音频缓冲器的容量的最大值如下所示。即，需要能够存储具有与“按照其解码定时足够输入 I 图像的最大比特量到视频缓冲器的时间”相对应的长度的音频数据量的大小。音频缓冲器需要的最大值 EBn_max 能够用以下等式 (11) 表示。

$$EBn_max = (I_max/Rv) * Ra[\text{比特}] \dots (11)$$

其中， I_max 是 I 图像的最大比特量，它对应于图 8 所示的视频编码缓冲器 EB1 的大小， Rv 是对视频编码缓冲器 EB1 的输入比特率，以及 Ra 是音频流的比特率。如以上等式 (11) 所示，将要计算的音频缓冲器的大小 EBn_max 是通过将对视频基本流缓冲器 (EB1) 以输入比特率把视频编码缓冲器 EB1 的缓冲器占用量从 0 增加到 I_max 所需的时间乘以 Ra 而获得的值。

作为具体值，推荐能够存储对应于至少 100 毫秒的音频数据的缓冲器大小。原因如下。即，当每 0.5 秒编码 I 图像时，I 图像的比特大小通常是编码比特率的 10% 或更少。假定编码比特率是例如 10Mbps，I 图像的大小通常是 1M 比特或更小。

因此，作为第一个原因，用至少 100 毫秒，I 图像能够按照其解码定时输入到解码器 20 的视频缓冲器。而且，作为第二个原因，如果解码器 20 的音频缓冲器能够存储对应于 100 毫秒的音频数据，有可能复用 TS1，以便音频数据向音频缓冲器的输入比音频数据的再现定时早 100 毫秒完成。因此，当音频缓冲器具有能够存储至少对应于 100 毫秒的音频数据的缓存大小时，基于以上的第一和第二原因，能在 TS1 和与 TS1 无缝连接的 TS2 之间切换期间，TS1 的最后的传输包输入完成之后，确保作为直到在 TS2 中首先将解码的图像 (I 图像) 的输入已经完成的时间的至少 100 毫秒。

能够存储对应于 100 毫秒的音频数据的音频缓冲器的容量具体如下计算。

在 640kbps 的杜比 AC3 音频流的情况下: $640\text{kbps} \times 0.1 \text{ 秒} = 64\text{k 比特} = 8\text{k 字节}$

在线性 PCM 音频流的情况下 (24 比特样本, 96KHz 取样频率, 8 信道):
(24 比特样本 \times 96,000 采样/秒 \times 8 信道) \times 0.1 秒 = 230,400 字节

下面将参照图 14 和 15 详细说明，以上描述的象本实施例的解码器 20 那样，通过改变 DVR - STD 的音频缓冲器的大小为“能够存储与按照其解码定时足够把首先在 TS2 中将要解码的图像 (I 图像) 输入到视频缓冲器的时间

相对应的音频数据量的大小”而获得的优点。

这里，将把具有 640kbps 的比特率和 48kHz 的取样频率的 AC3 音频流作为例子来解释。AC3 音频流的一个音频帧的样本数是 1, 536。因此，一个音频帧的时间长度是 32 毫秒。一个音频帧的字节大小是 2, 560 字节。

图 14 (a) 和 14 (b) 是表示在传统的 DVR - STD 中具有 4k 字节缓冲器大小的音频缓冲器的情况下，在 TS1 和与 TS1 无缝连接的 TS2 之间切换期间，DVR - STD 的视频缓冲器和音频缓冲器的比特占用量上的变化的例子的图。在图 14 (a) 和 14 (b) 中，虚线表示 TS1 的视频/音频数据的缓冲器变化，实线表示 TS2 的视频/音频数据的缓冲器变化。

4k 字节的音频缓冲器能够存储对应于 50 毫秒的音频数据。因此，在作为 TS1 的最后一个音频包的最后一个字节到达 DVR - STD 的时刻的 $STC1^{audio_end}$ ，能复用 TS1 以便音频数据的输入比其再现定时早 50 毫秒完成。但是，50 毫秒不足以按照其解码定时把首先在 TS2 中将要解码的图像 (I 图像) 输入到视频缓冲器。这时，编码受到限制，从而降低首先在 TS2 中将要解码的图像 (I 图像) 的大小，这使图像质量恶化。

因为 4k 字节的音频缓冲器能够将音频数据向前移动 50 毫秒，图 14 (a) 和 14 (b) 所示的起动延迟 t_1 变为至多 50 毫秒那么小， t_1 是输入 TS2 的第一 I 图像到视频缓冲器的时间。因此，能使用足够的时间输入 TS2 的第一 I 图像，降低 I 图像大小 S_1 ，结果 I 图像的图像质量由于编码的限制而恶化。如上所述，有必要除了 4k 字节之外提供对应于 1 秒的附加缓冲器并在 T1 和 T2 之间以 TS 的最大速率 R_{MAX} 输入 I 图像，从而增加起动延迟。这里，对具有 640kbps 比特率的 AC3 音频流给出了说明。但是，如上所述，对应于 1 秒的附加缓冲器对于多信道 LPCM 音频是很大的。

为了处理该问题，DVR - STD 的音频缓冲器大小改变为例如象本发明的解码器 20 的 8k 字节。图 15 (a) 和 15 (b) 每个表示优化音频缓冲器容量的一个例子。更特别的是，图 15 (a) 和 15 (b) 示出在音频缓冲器是 8k 字节的情况下，在 TS1 和无缝连接到 TS1 的 TS2 之间切换期间，本发明的 DVR - STD 的视频和音频缓冲器的比特占用量中的变化的图。在图 15 (a) 和 15 (b) 中，虚线表示 TS1 的视频/音频数据的缓冲器变化，实线表示 TS2 的视频/音频数据的缓冲器变化。

8k 字节的音频缓冲器能够存储对应于 100 毫秒的音频数据。因此，在

STC1¹_{audio_end}，它是 TS1 的最后一个音频包的最后一个字节到达 DVR - STD 的时刻，有可能复用 TS1，以便音频数据的输入比其再现定时早 100 毫秒完成。使用至少 100 毫秒，首先在 TS2 中将要解码的图像（I 图像）能够更容易在其解码定时输入到视频缓冲器。也就是说，能用足够的时间（起动延迟）t2 输入 TS2 中的第一个 I 图像，这能够增加首先在 TS2 中将要解码的图像（I 图像）的大小 S2。因此，I 图像的图像质量能够由于较低的编码限制而增加。

而且，在图 8 所示的播放器模式 1 中，TS 能够看作在复用器（信息处理设备）中产生并记录，TS 由每个具有传输包和到达时间印记的多个源数据包组成的数据流组成，并 TS 由解码器基于到达时间印记读出并解码。

按照例如参照图 4 到 6 所描述的，复用器包括：视频编码部分，产生再编码的剪辑 1（第一视频编码流）以在一预定图像结束呈现和在该图像之后立即呈现并再编码以开始呈现的剪辑 2（第二视频编码流）；复用部分，复用剪辑 1 和与剪辑 1 同步的音频编码流以产生 TS1 并复用剪辑 2 和与剪辑 2 同步的音频编码流以产生 TS2；和记录由 TS1 和 TS2 组成的复用流的记录部分。在复用部分中，TS1 和 TS2 这样被复用，以便音频数据向解码器 20 的输入能够在 I 图像向解码器 20 输入开始时完成，该音频数据对应于输入 I 图像，即向解码器 20 的视频缓冲器输入第二图像的时间。注意，如图 5 所示，能在编码部分产生桥式剪辑并在复用部分产生与 TS1 和 TS2 一起的复用桥式剪辑。

在用于记录上述复用器产生的复用流的记录介质上记录的是由在第一图像结束的 TS1 和从跟随着第一图像再现的第二图像开始的 TS1 组成的复用流，其中 TS1 和 TS2 能够根据它们的到达时间印记输入到解码器 20，并且 TS1 和 TS2 能够这样被复用，以便能够在第二图像，即输入到解码器 20 的 TS2 的第一图像输入开始时完成音频数据向解码器的输入，音频数据对应于输入第二图像到解码器 20 所需的时间。

在上述配置的本发明中，当再现互相无缝连接的 TS1 和 TS2 时，根据其从向解码器 20 的 TB1 输入 TS1 的最后一个视频包完成的时刻至到 TS1 剩余的数据包输入到解码器 20 的时刻的到达时间印记执行传输包的输入，并且音频缓冲器的大小从传统 DVR - STD 中的 4k 字节变化到能够存储具有与在其解码定时输入 I 图像的最大比特量到视频缓冲器所需时间相对应的长度的音频数据的数据量的大小。结果，能足以确保从 TS1 的最后一个数据包输入完成的时刻到在其解码定时完成作为 TS2 的第一图像的 I 图像输入的时刻所需

时间（启动延迟）。因此，I 图像的图像质量能够由于较低的编码限制而增加。

而且，在使用附加缓冲器的方法中，如传统上一样，当 TS 中的音频数据假定为例如多信道 LPCM 音频数据时，需要具有极大容量的附加缓冲器。然而，在本实施例中，能通过改变上述音频缓冲器的容量并通过根据到达时间印记输入传输包来消除对传统方法中不可缺少的附加缓冲器的需要。

本发明不限于以上参照附图描述的实施例，并且对本领域技术人员来说，很明显在不脱离这里随附的权利要求和本发明的精神和范围的条件能够作出各种修改、替换或等效物。

产业上的可利用性

根据上述本发明，能编辑由视频和具有视频帧精度的音频流组成的复用流并以无缝的方式再现它，并除去对应于 1 秒的附加缓冲器，比以前更加降低解码器所需的缓存量，1 秒是使用被忽略的源数据包的 arrival_time_stamp 以 TS 的最大比特率输入传输包所需的时间。而且，能改变音频缓冲器的大小到能够缓存具有与输入第二图像到视频缓冲器所需的时间相对应的长度的音频数据的大小，以便第二图像的图像质量能够由于较低的编码限制而增加。

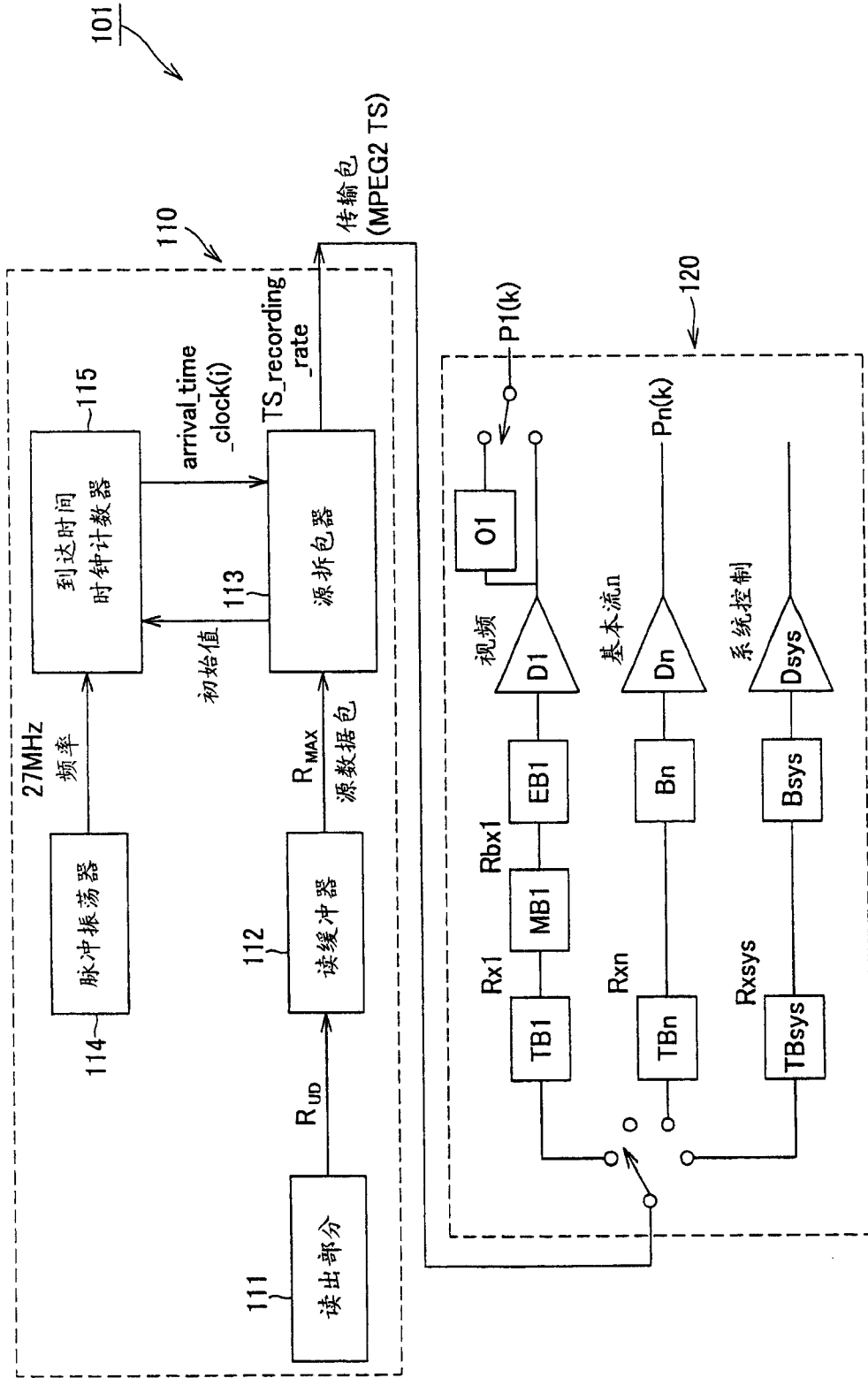


图 1

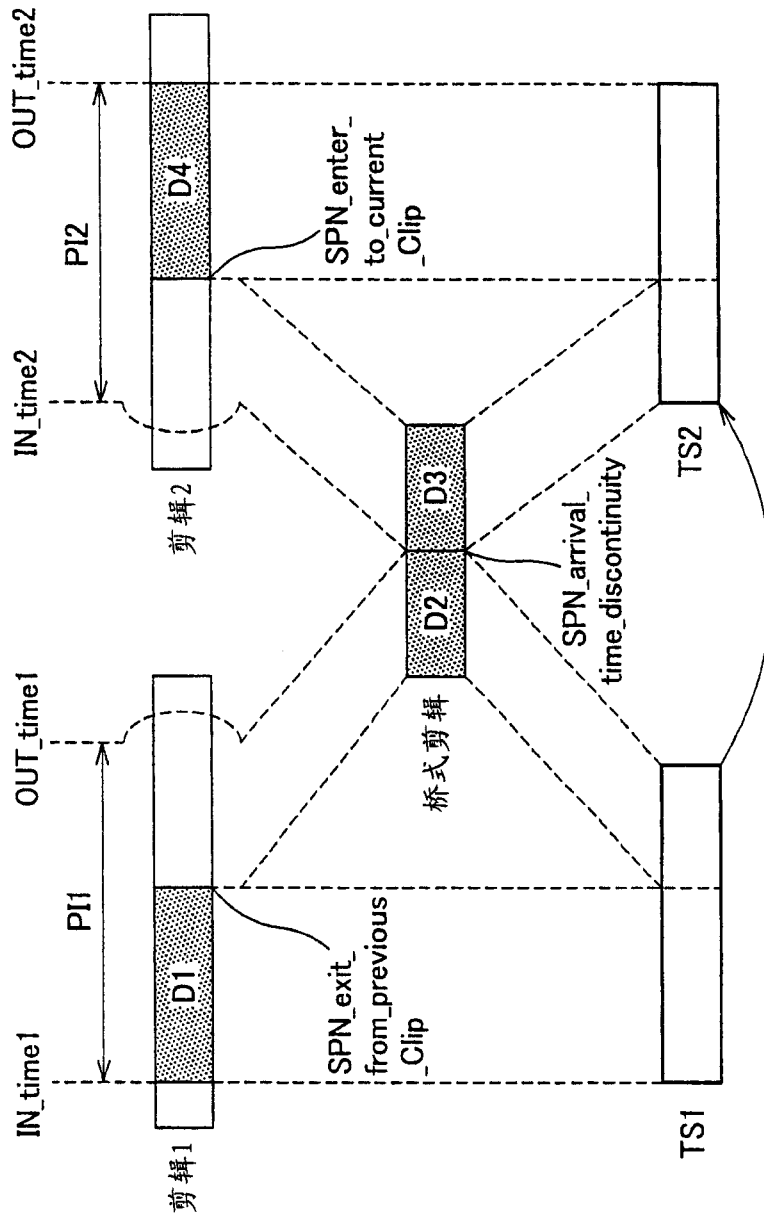


图 2

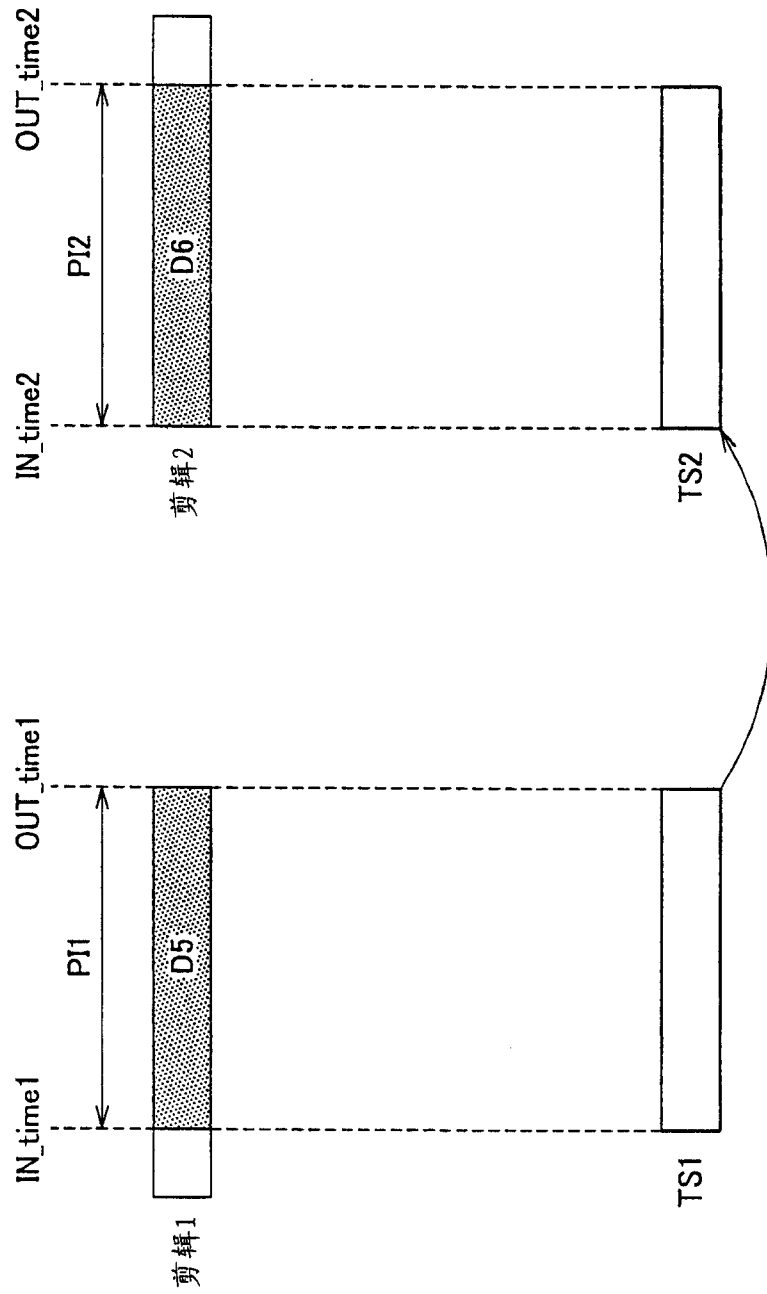


图 3

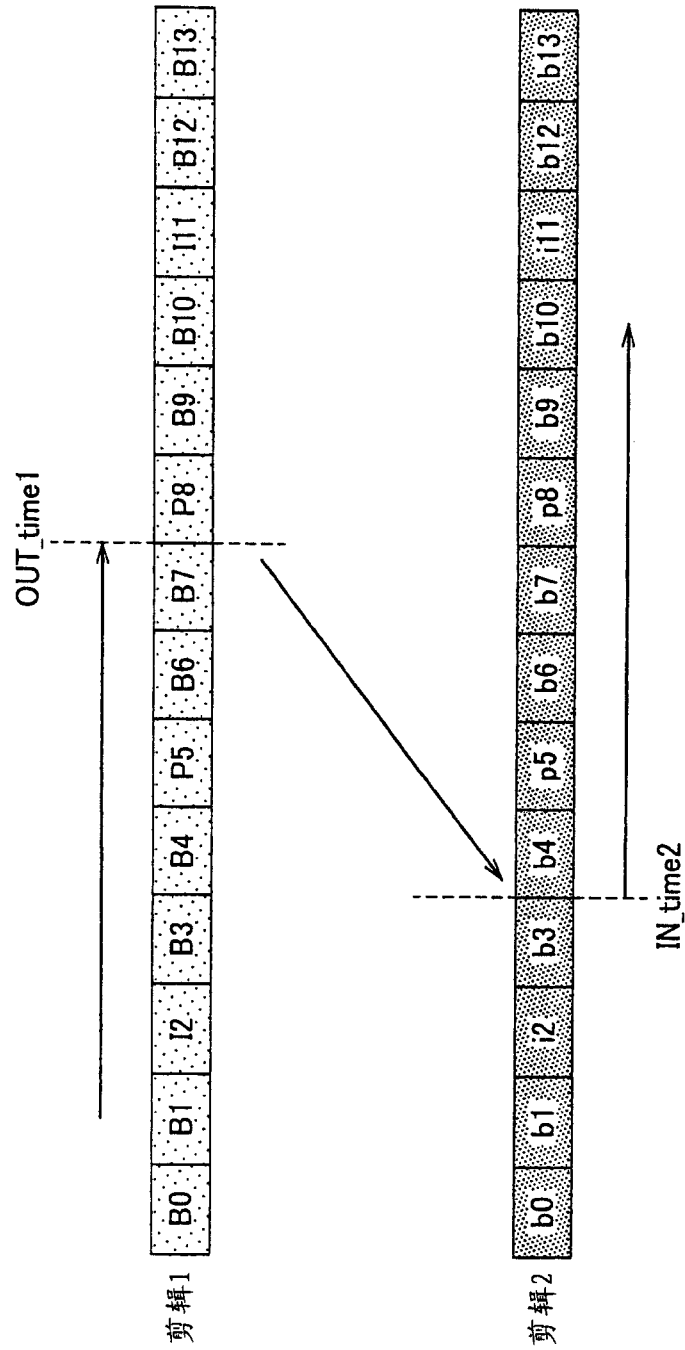


图 4

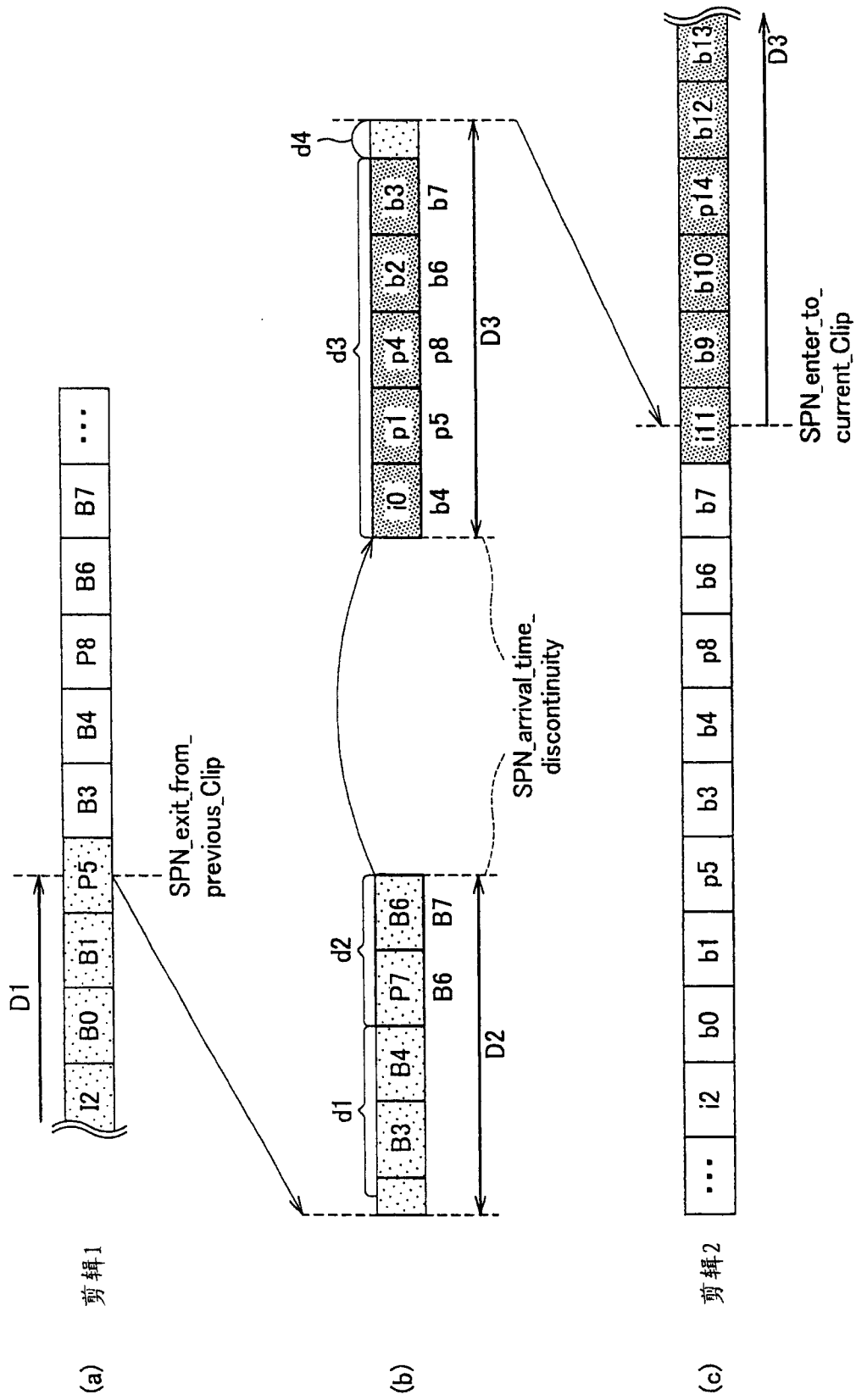


图 5

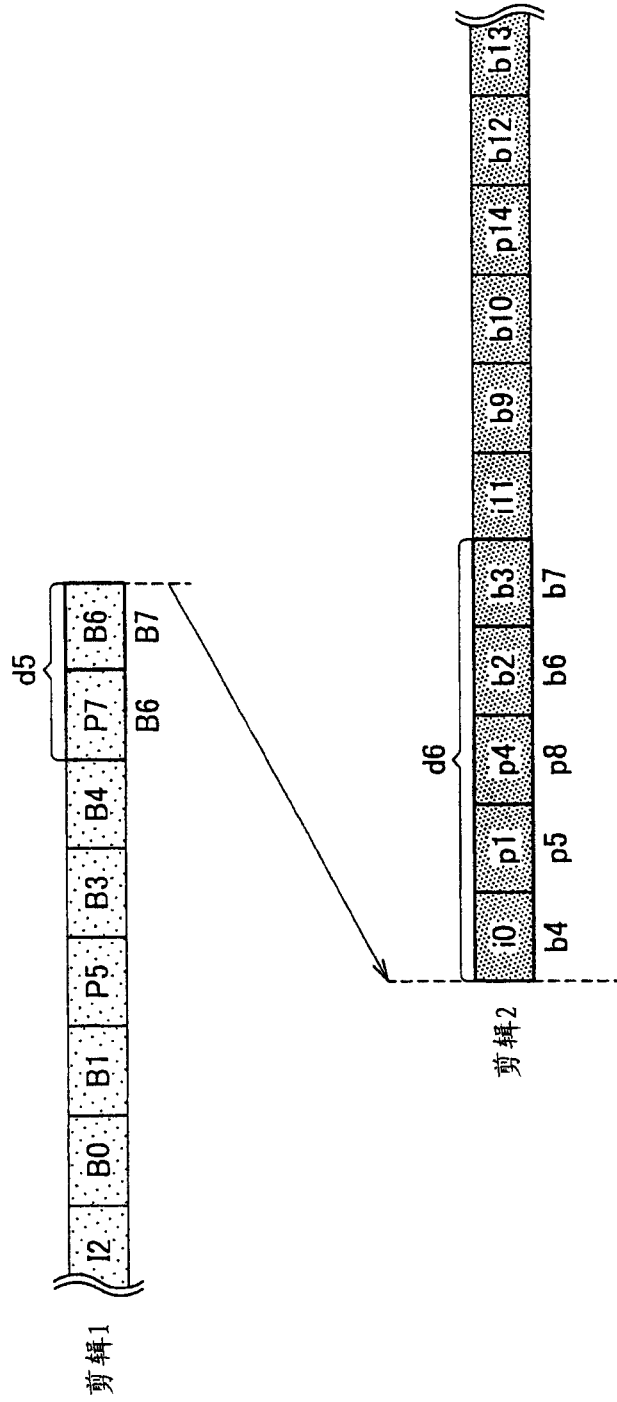


图 6

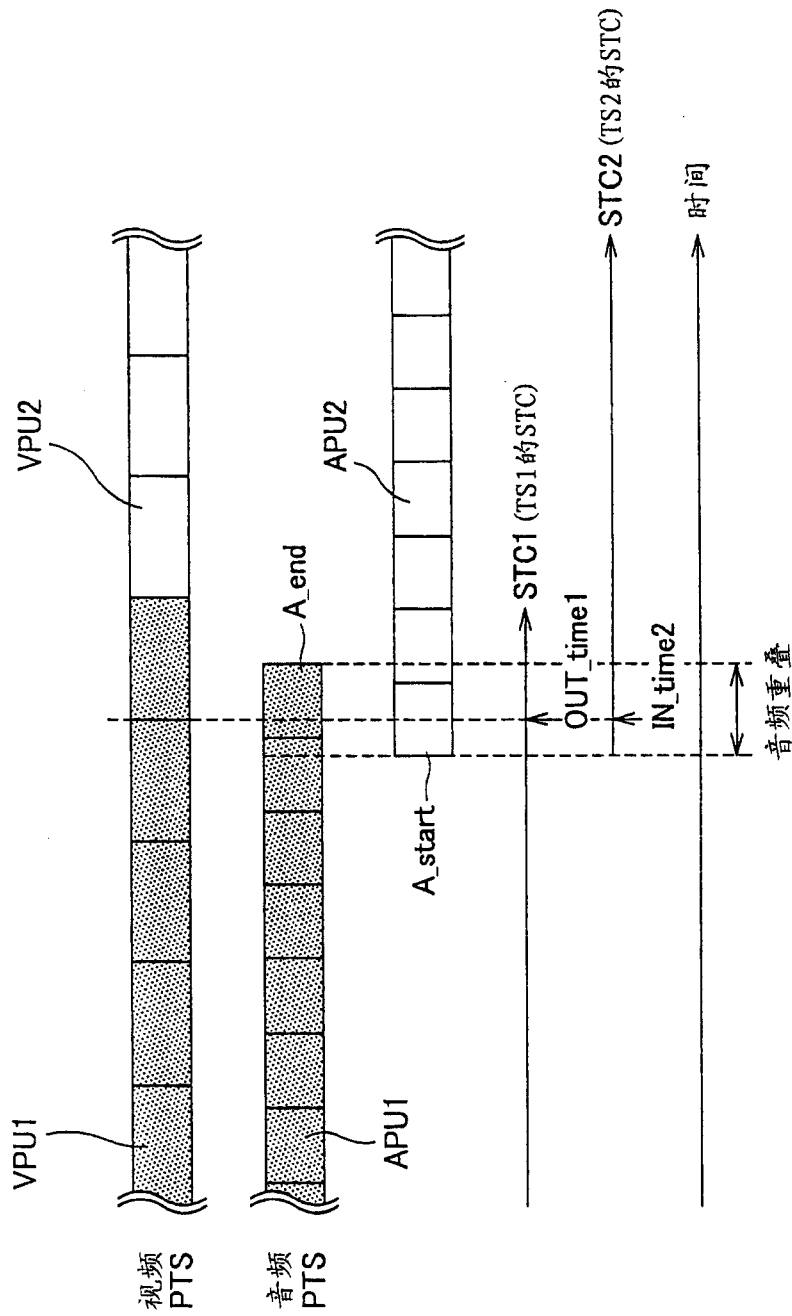


图 7

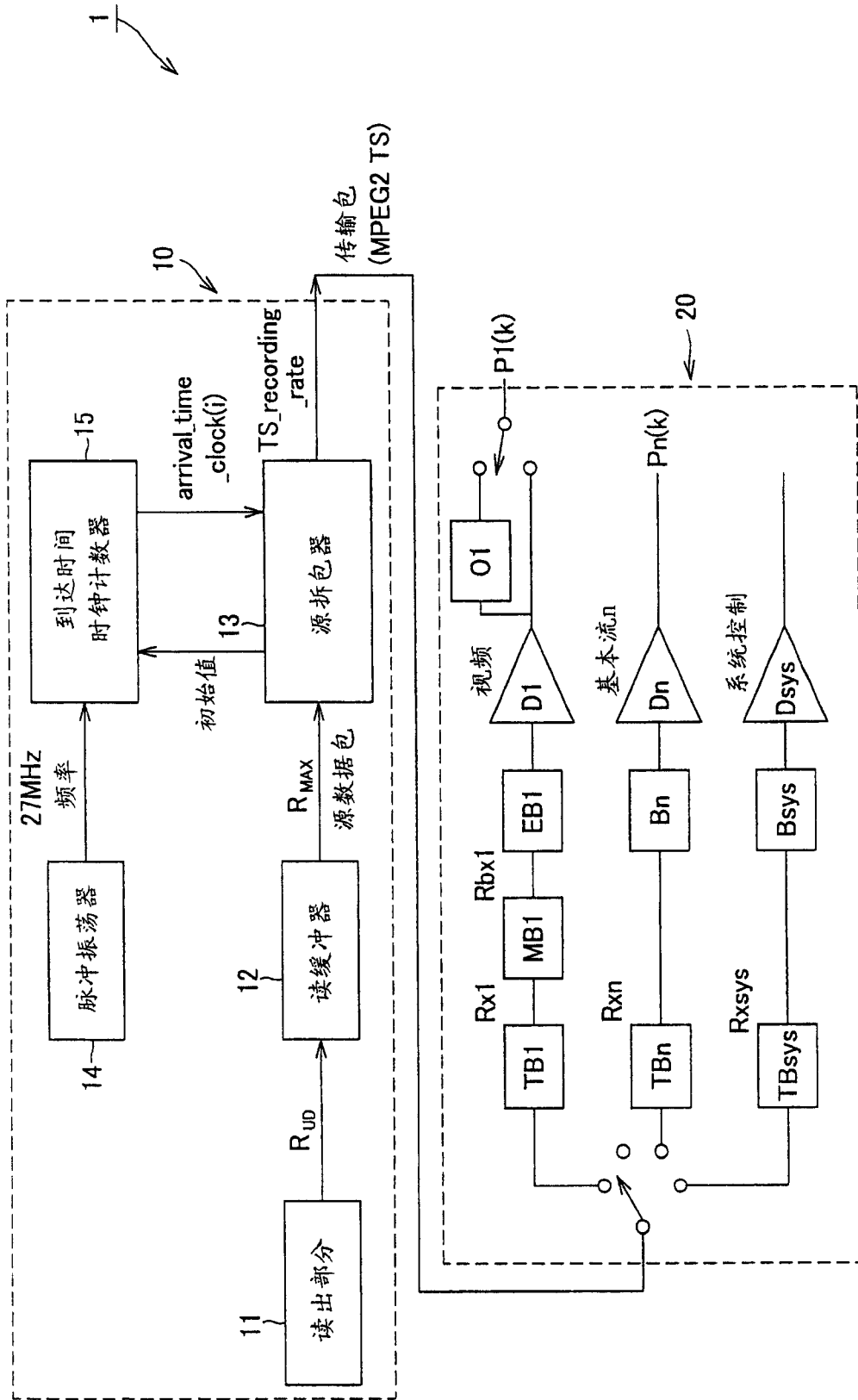


图 8

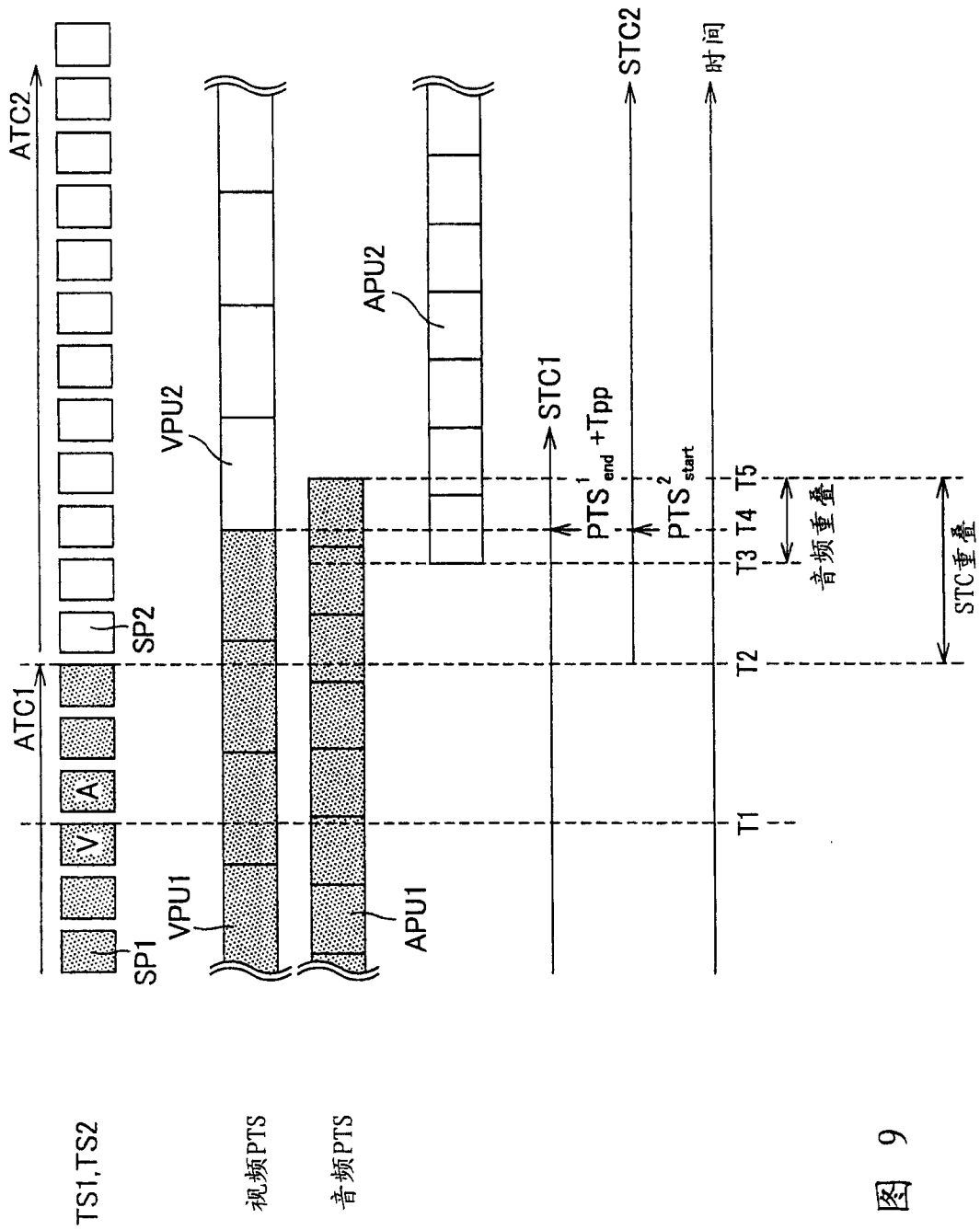


图 9

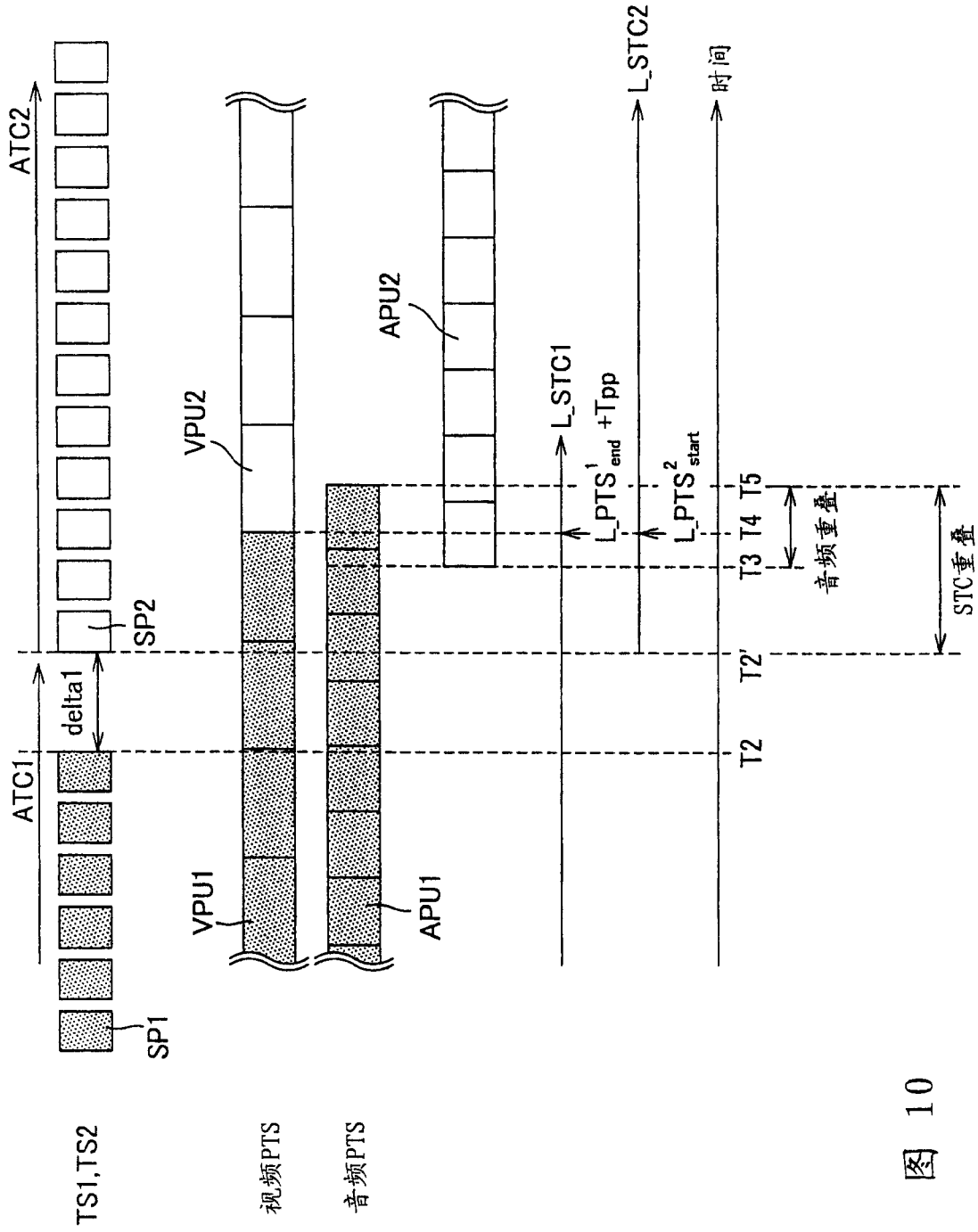


图 10

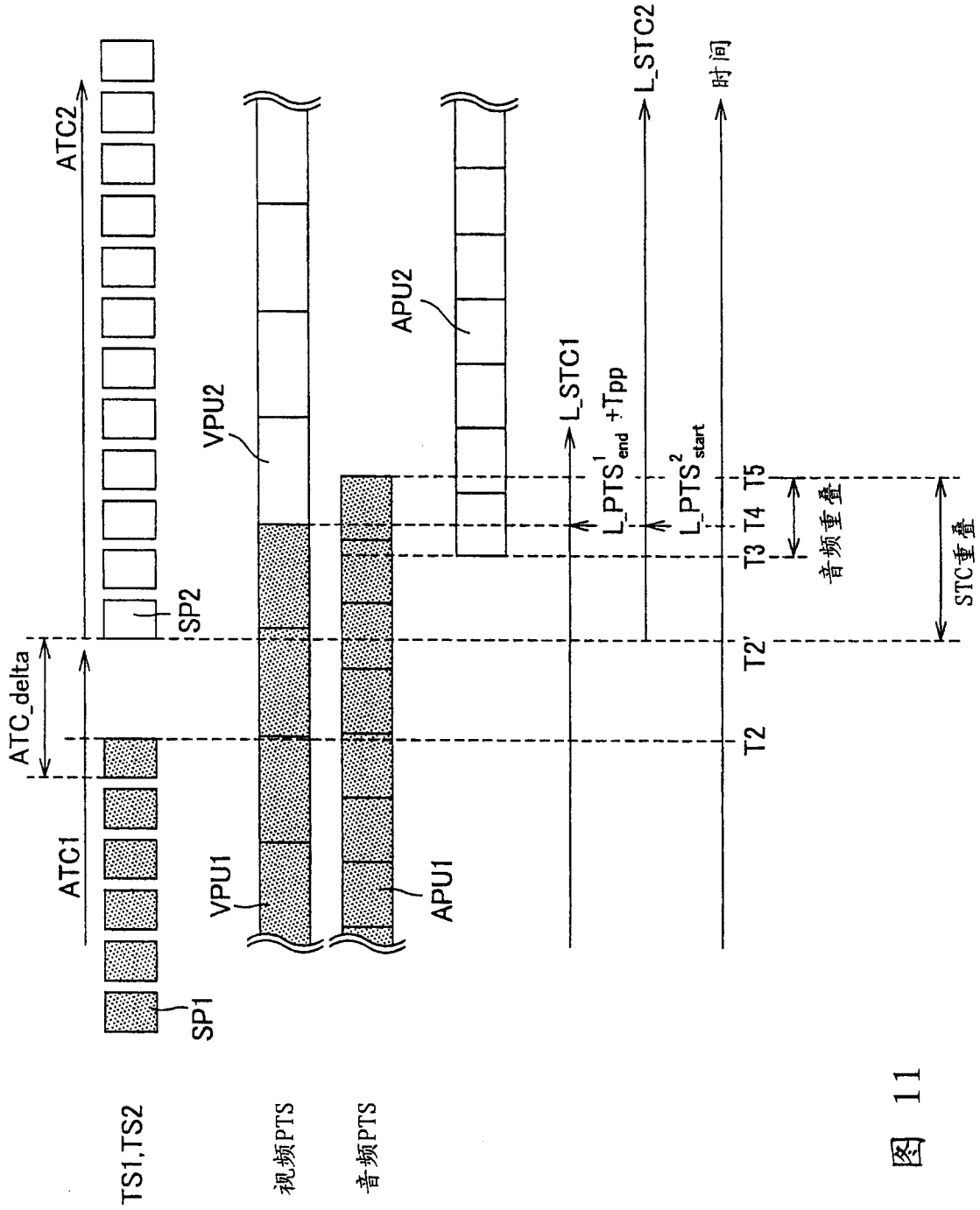


图 11

句法	比特数	记号
ClipInfo()		
length	32	uimsbf
reserved for future use	16	bslbf
Clip stream type	8	bslbf
application type	8	bslbf
reserved for future use	31	bslbf
is_ATC_delta	1	bslbf
TS recording rate	32	uimsbf
number of source packets	32	uimsbf
reserved for future use	1024	bslbf
TS type info block()		
if (is_ATC_delta=1_b){		
reserved for future use	8	bslbf
number of ATC delta entries	8	uimsbf
for (i=0; i<number of ATC delta entries; i++){		
ATC delta[i]	32	uimsbf
following Clip Information file name[i]	8*5	bslbf
Clip codec identifier	8*4	bslbf
reserved for future use	8	bslbf
}		
}		

图 12

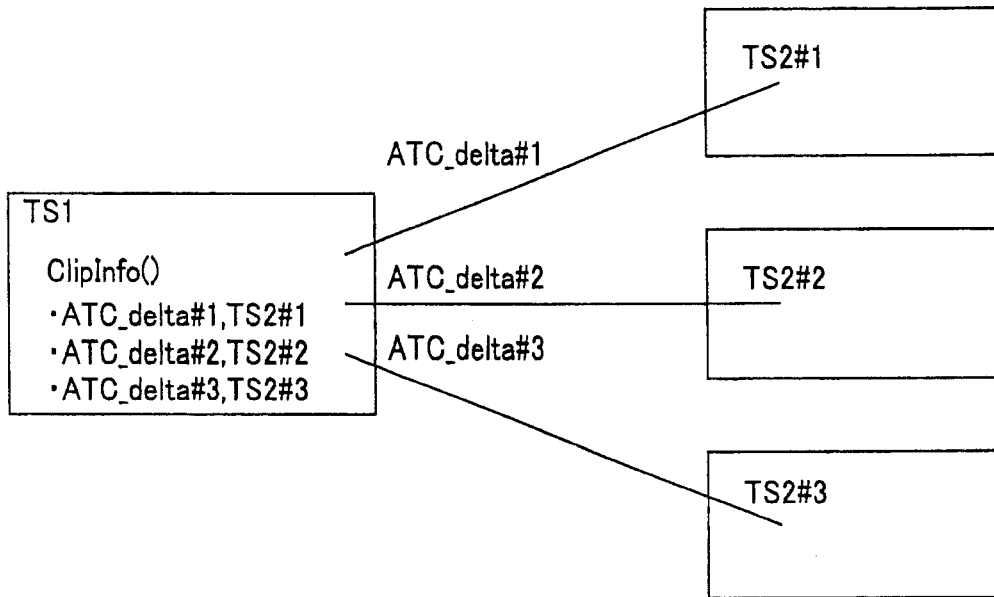


图 13

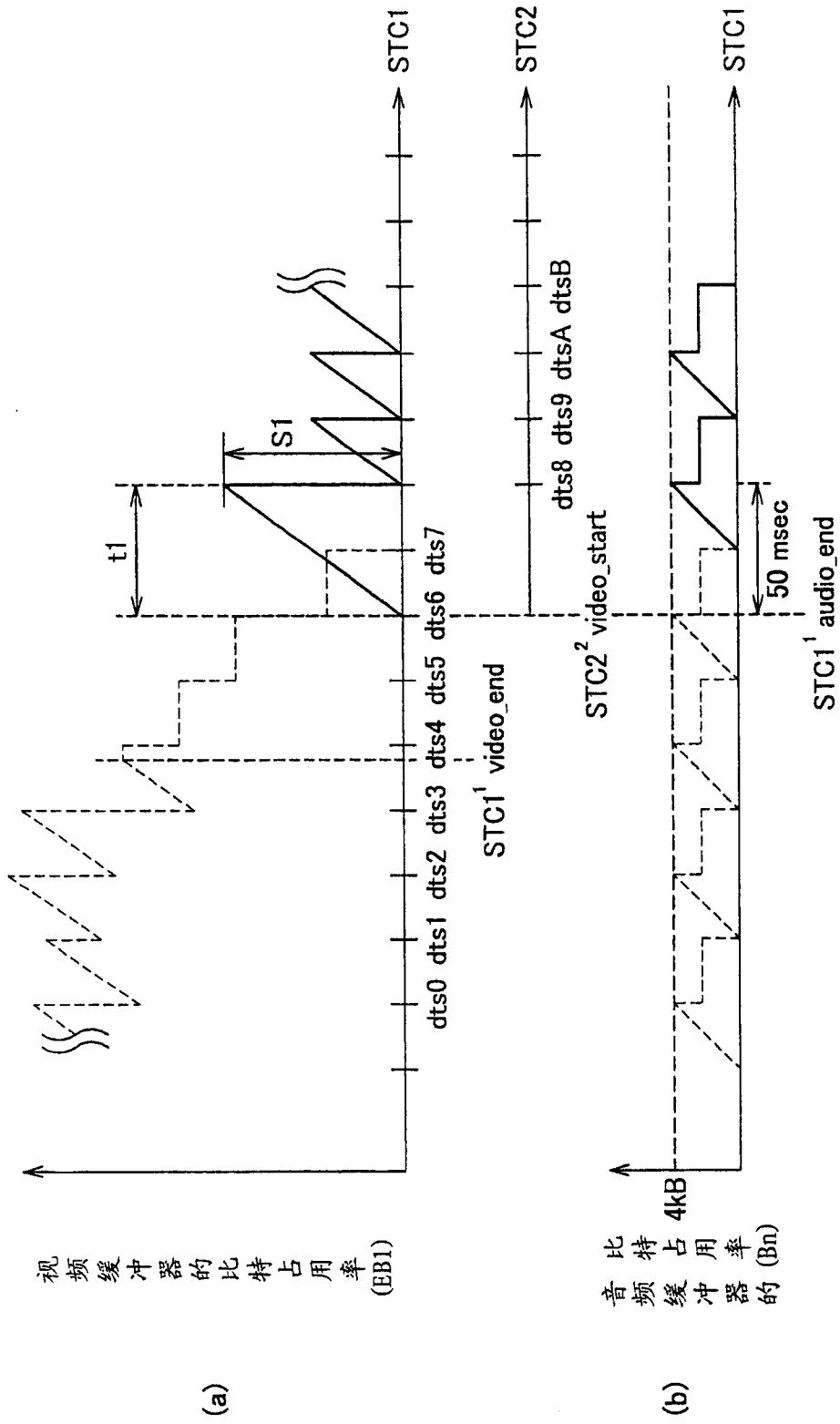


图 14

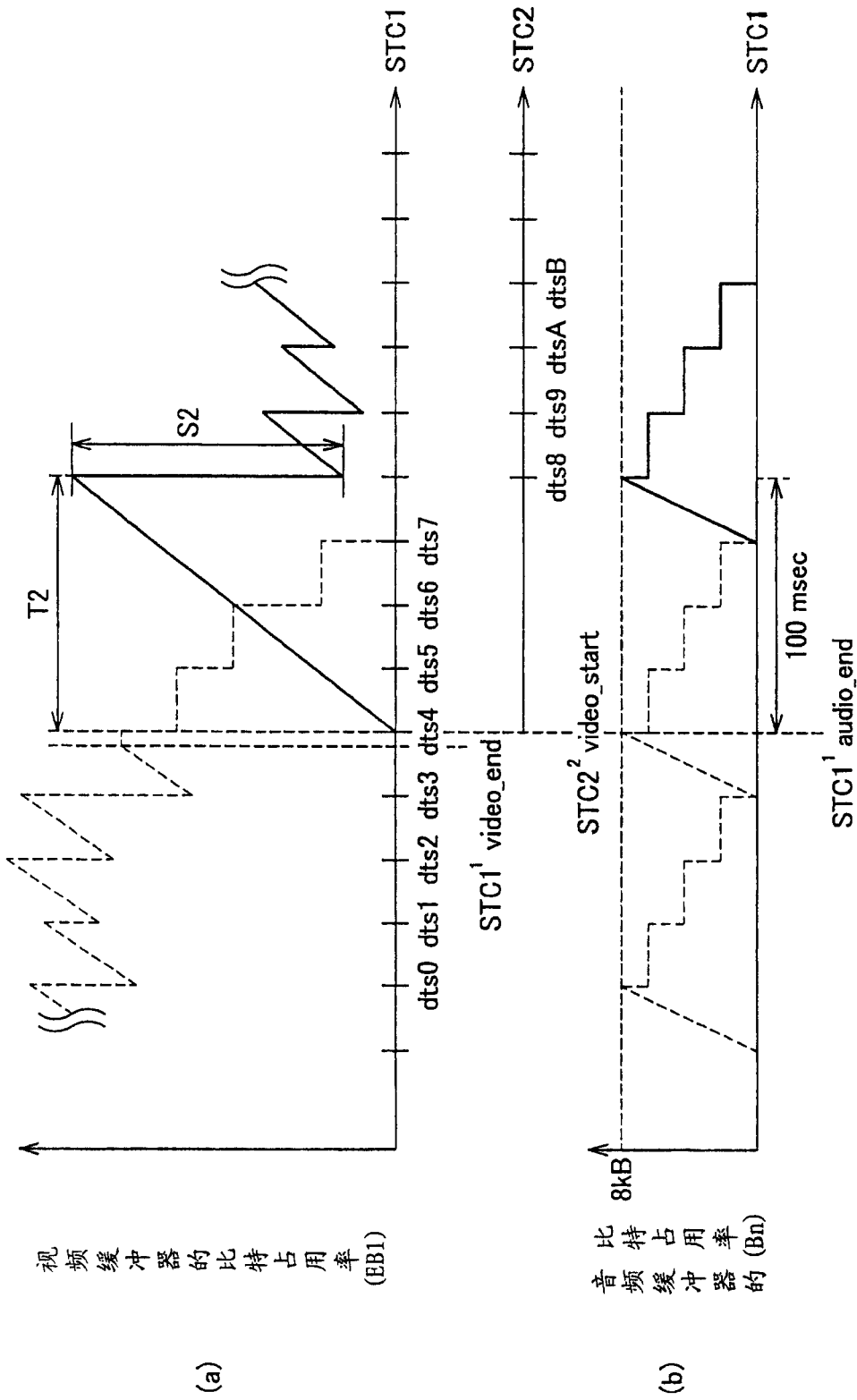


图 15