

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
H04N 7/01 (2006.01)



[12] 发明专利说明书

专利号 ZL 200480040820.8

[45] 授权公告日 2010 年 1 月 13 日

[11] 授权公告号 CN 100581235C

[22] 申请日 2004.11.4

[21] 申请号 200480040820.8

[30] 优先权

[32] 2004.1.21 [33] EP [31] 04090021.9

[86] 国际申请 PCT/EP2004/012483 2004.11.4

[87] 国际公布 WO2005/071956 英 2005.8.4

[85] 进入国家阶段日期 2006.7.21

[73] 专利权人 汤姆逊许可公司

地址 法国布洛涅

[72] 发明人 卡斯滕·赫佩尔

海因茨·沃纳·基森

安德烈·舍韦佐 马尔科·温特

[56] 参考文献

WO9952281A2 1999.10.14

US6181382B1 2001.1.30

US20020047937A1 2002.4.25

US20020118296A1 2002.8.29

审查员 戚颖

[74] 专利代理机构 北京康信知识产权代理有限责
任公司

代理人 章社杲

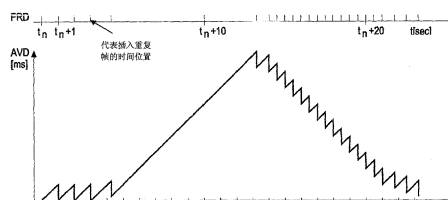
权利要求书 5 页 说明书 12 页 附图 5 页

[54] 发明名称

用于控制向第一格式图像序列插入附加场或
帧以从中建立第二格式图像序列的方法和装
置

[57] 摘要

本发明涉及用于向第一格式图像序列插入附加场或者帧以从中建立第二格式图像序列的控制方法和装置。世界上主要的电视系统采用隔行扫描并且采用 50Hz 或者 60Hz 的场频率。然而，电影是采用 24Hz 的帧频率以及逐行扫描制作而成的，该格式将被用于未来在采用 50Hz 的国家销售的数字视频光盘。在 50Hz 播放设备中，所述光盘内容的呈现是以原始的音频音调但却以重复的视频帧或场来实现，以平均地达到所述原始视频源的速度。然而，对于所述帧或者场的插入并不是以常规的排列进行的，而是以适应性的方式进行，从而减少可见的运动抖动。



1、一种用于控制 (CTRL, VDEC) 向第一格式图像序列插入附加的场或者帧以从中建立第二格式图像序列的方法, 所述第二格式图像序列的帧频率恒定、并且大于所述第一格式图像序列的帧频率, 所述方法包括下列步骤:

- 确定 (CTRL, VDEC, ADEC) 所述第一格式图像序列中的场或者帧位置, 在该位置上插入相应的附加场或者帧, 从而使得在第二格式图像序列中可见的运动抖动最小;

- 以非常规的场或者帧插入距离在所述第一格式图像序列中的一些所述位置入场或者帧, 从而使得总的来说任何临近帧之间平均距离对应于所述第二格式图像序列的临近帧之间的平均距离;

- 采用所述第二格式图像序列的格式一起呈现所述第一格式图像序列以及所述非常规插入的场和/或帧,

其特征在于,

为了达到可感知的唇同步, 确定所述视频轨迹的相对于所述音频轨迹的正方向的最大视频延迟 VD_{max} 和负方向的最大视频延迟 VD_{min} , 并且对所述第一格式图像序列插入所述场或者帧的位置进行控制, 从而在满足下列条件的情况下, 使在所述第二格式图像序列中由于插入的非常规性所引起的所述视频轨迹相对于所述音频轨迹的正方向的最大视频延迟 VD_{max} 和负方向的最大视频延迟 VD_{min} 位于唇同步可接受值 $VD_{minLipSync}$ 和 $VD_{maxLipSync}$ 之间:

- 检测缓慢移动或者静态的场景, 以及

- 在被分配到所述第一格式图像序列的音频信息中检测语音或者短音峰。

2、如权利要求 1 所述的方法, 其中, 对分配给所述第一格式图像序列的所述音频信息中的语音的检测, 是通过在多声道音频中进行评估, 以判

断所述中央声道相对于左和右声道是否表现出随时间爆发的能量分布，该分布明显区别于所述左和右声道的能量分布。

3、如权利要求 1 或 2 所述的方法，其中，所述第一格式图像序列储存或者记录在存储介质（D）上，或者作为数字电视信号被广播或者传播。

4、如权利要求 3 所述的方法，其中，所述存储介质（D）是光盘或者硬盘。

5、如权利要求 1 或 2 所述的方法，其中，在所述第一格式图像序列中插入场或者帧的位置是不包含大块的运动图像内容区域的场或者帧，所述运动可以通过对运动矢量的评估来进行确定。

6、如权利要求 1 或 2 所述的方法，其中，在所述第一格式图像序列中插入场或者帧的位置是发生场景变化或者逐渐变黑或者逐渐变白或者逐渐变为任一颜色的场或帧。

7、如权利要求 1 或 2 所述的方法，其中，在所述被插入的场或者帧被输出到所述第二格式图像序列中之前，对所述被插入的场或者帧进行运动补偿。

8、如权利要求 1 或 2 所述的方法，其中，所述第一格式图像序列为 MPEG-2 图像序列，并且其中，通过评估标记对于向所述第一格式图像序列插入（CTRL, VDEC）场或帧进行控制，所述评估标记指示场的时间顺序或者指示为了进行显示而对第一场进行重复，所述标记在所述第一格式图像序列每幅图像的用户数据场中被传送。

9、如权利要求 1 或 2 所述的方法，其中，所述第一帧频率是 24Hz 并且所述第二格式图像序列的帧频率是 25Hz。

10、如权利要求 1 或 2 所述的方法，其中，发送对于将要被重复或内插的场或者帧的位置的信号的标记在所述第一格式图像序列的用户数据场中被传送。

11、一种用于控制（CTRL, VDEC）向第一格式图像序列插入附加的场或者帧以从中建立第二格式图像序列的方法，所述第二格式图像序列的

帧频率恒定、并且大于所述第一格式图像序列的帧频率，其中，所述第一格式图像序列是具有基本为 24Hz 帧频率的 MPEG-2 图像序列，以及其中为了达到可以感知的唇同步，在所述 MPEG-2 图像序列中插入场或帧的位置由被传送的标记所控制，从而由插入的非常规性引起的最大图像内容延迟保持在唇同步可接受值 $VD_{\min LipSync}$ 和 $VD_{\max LipSync}$ 之间，以防出现缓慢移动或者静态的场景以及分配给所述 MPEG-2 图像序列的音频信息中的语音，对于所述 MPEG-2 图像序列中的每幅图像来说，所述方法包括在用户数据场中插入指示场时间顺序的标记或者指示为了进行显示而对第一场进行重复的标记的步骤。

12、一种用于控制 (CTRL, VDEC) 向具有第一帧频率的第一格式图像序列插入附加的场或者帧以从中建立第二格式图像序列的装置，所述第二格式图像序列的帧频率恒定、并且大于所述第一帧频率，所述装置包括部件 (CTRL, VDEC, ADEC) 用于确定在所述第一格式图像序列中的场或帧位置，在该位置上插入对应的附加场或帧，从而使所述第二格式图像序列中的可见运动抖动最小，

所述装置还用于以非常规场或帧插入距离在一些所述位置上向所述第一格式图像序列插入场或帧，从而使得总体来说任何临近帧之间的平均距离对应于所述第二格式图像序列中任何临近帧之间的平均距离，

所述装置还用于以所述第二格式图像序列的格式一起呈现所述第一格式图像序列以及所述非常规插入的场和/或帧，

其特征在于，

为了达到可感知的唇同步，由所述部件确定所述视频轨迹的相对于所述音频轨迹的正方向的最大视频延迟 $VD_{\max}$ 和负方向的最大视频延迟 $VD_{\min}$ ，并且控制在所述第一格式图像序列内插入场或帧的位置，从而在满足下列条件的情况下，使在所述第二格式图像序列中由插入的非常规性引起的所述视频轨迹相对于所述音频轨迹的正方向的最大视频延迟 $VD_{\max}$ 和负方向的最大视频延迟 $VD_{\min}$ 位于唇同步可接受值 $VD_{\min LipSync}$ 和

VD_{maxLipSync} 之间:

- 检测缓慢移动或者静态的场景, 以及
- 在被分配到所述第一格式图像序列的音频信息中检测语音或者短音峰。

13、如权利要求 12 所述的装置, 所述装置为光盘播放器或者光盘刻录机, 或者硬盘录像机, 或者 PC, 或者机顶盒, 或者电视接收器。

14、如权利要求 12 或 13 所述的装置, 其中所述装置输出所述原始的第一格式图像序列或者所述第二格式图像序列, 这个选择是由所接收到的重放模式信息来控制的, 所述重放模式信息自动地接收于与包括显示设备在内的设备相连的接口 (IF), 或者接收于用户接口 (UI)。

15、如权利要求 12 或 13 所述的装置, 其中对分配给所述第一格式图像序列的所述音频信息中的语音的检测, 是通过在多声道音频中进行评估, 以判断所述中央声道相对于左和右声道是否表现出随时间爆发的能量分布, 该分布明显区别于所述左和右声道的能量分布。

16、如权利要求 12 或 13 所述的装置, 其中所述第一格式图像序列储存或者记录在存储介质 (D) 上, 或者作为数字电视信号被广播或者传播。

17、根据权利要求 16 所述的装置, 其中, 所述存储介质 (D) 是光盘或者硬盘。

18、如权利要求 12 或 13 所述的装置, 其中在所述第一格式图像序列中插入场或者帧的位置是不包含大块的运动图像内容区域的场或者帧, 所述运动可以通过对运动矢量的评估来进行确定。

19、如权利要求 12 或 13 所述的装置, 其中在所述第一格式图像序列中插入场或者帧的位置是发生场景变化或者逐渐变黑或者逐渐变白或者逐渐变为任一颜色的场或帧。

20、如权利要求 12 或 13 所述的装置, 其中在所述被插入的场或者帧被输出到所述第二格式图像序列中之前, 对所述被插入的场或者帧进行运

动补偿。

21、如权利要求 12 或 13 所述的装置，其中所述第一格式图像序列为 MPEG-2 图像序列，并且通过评估标记对于向所述第一格式图像序列插入（CTRL，VDEC）场或帧进行控制，所述评估标记指示场的时间顺序或者指示为了进行显示而对第一场进行重复，该标记在所述第一格式图像序列每幅图像的用户数据场中被传送。

22、根据权利要求 12 或 13 所述的装置，其中，所述第一帧频率是 24Hz 并且所述第二格式图像序列的帧频率是 25Hz。

用于控制向第一格式图像序列插入附加场或帧 以从中建立第二格式图像序列的方法和装置

技术领域

本发明涉及用于控制向具有例如每秒 24 帧逐行扫描帧的第一格式图像序列插入附加场或者帧，以从中建立具有例如每秒 25 帧的第二格式图像的方法和装置。

背景技术

世界上主要的电视系统采用隔行扫描并且采用 50Hz 的场频率（例如在欧洲和中国采用 PAL 和 SECAM），或者 60Hz 或接近 60Hz 的场频率（例如在美国和日本采用 NTSC），分别以 50i 和 60i 来表示。然而，电影是采用由 24p 代表的 24Hz 的帧频率及逐行扫描制作而成的，这个值在隔行扫描格式下对应于 48i。

目前，从 24p 的电影到 60Hz 隔行扫描显示的转换是由图 2 所示的“3:2 下拉”技术完成的，其中由场循环在每五个场后加入一个 3:2 下拉场。隔行扫描场 ILF 源自于原始的电影帧 ORGFF。从第一原始电影帧 OFR1 产生三个输出场 OF1 到 OF3，并且从第三原始电影帧 OFR3 产生三个输出场 OF6 到 OF8。从第二原始电影帧 OFR2 产生两个输出场 OF4 和 OF5，并且从第四原始电影帧 OFR4 产生两个输出场 OF9 和 OF10，等等。

比较理想的是所发行媒体具备可以在全世界播放的单一视频音频格式，而不是像目前的情况，针对每一套装媒体例如 DVD 至少存在 50Hz 和 60Hz 两个版本。因为很多媒体源都是由 24fps（帧每秒）的电影组成，该 24p 的格式就优选地成为理想的单一视频格式，因此该单一的格式需要适应在所述采用 50Hz 和 60Hz 的国家的显示设备上正确显示的回放时间。

以下是用于从 24p 到 25p 或者 50i，或者更概括地说到 25fps 的已知转

换方案：

- 以快 4.2% 的速度重放：这会改变所述内容的长度并且要求昂贵的实时音频音调转换，因此并不适用于消费电子产品。目前，电影播放以及 DVD 确实将该种方案应用于视频，但是所必需的音频速度或者音调转换已经在内容提供商一端进行了处理，从而在消费者一端不需要进行音频音调转换。在采用 50Hz 的国家销售的 DVD 视频光盘包括音频数据流，这些音频数据流已经被编码从而使得所述 DVD 播放器的解码器自动的输出所述音频信号的正确速度或者音调。
- 采用常规场/帧复制机制：该方案会导致不可接受的常规的运动抖动，并且因此在实际中无法应用。
- 采用运动补偿帧速度转换：该方案是解决该转换问题的普遍方案，该方案十分昂贵，因此并不适用于消费电子产品。

发明内容

目前，从原始的 24p 格式电影视频及音频数据流到 50Hz 的隔行扫描显示的转换是通过以快 4% 的速度重放所述电影来实现的。然而这意味着，要对采用 50Hz 的国家的电影的艺术内容（它的持续时间，声音的音调）进行修改。类似于 3:2 下拉技术的场/帧重复机制并没有被采用，因为在常规情况下采用所述机制时，例如每 12 帧插入一个额外的场，所述机制表现出不可接受的运动抖动现象。

本发明要解决的问题在于提供场或帧插入机制，以改进的方式实现从 24p 格式到 25fps 格式的转换，从而使运动抖动现象最小化。

对当前电影场景的特性例如全局运动，亮度/强度级别以及场景变换位置进行评估，以在主观认为不令人厌烦的位置上应用复制或重复的帧/场。也就是说，本发明采用相对容易得到的待从 24p 转换到 25fps 的所述源材料的信息，用于在导致插入现象最小的非等距位置适应性的插入重复的场/帧。有利地，在源的帧速率和目标帧速率之间具有微小差距的情况下，本发明

可以用于所有的帧速率转换问题。如果该帧速率差别很大，例如在 24fps 到 30fps 的转换中，几乎就没有在时域进行移动或者对帧进行重复的自由度。

本发明促进了以低计算成本实现从 24fps 到 25fps 格式图像序列（示例值）的转换，并使运动抖动最小化。

总的来说，本发明所记载的方法适用于控制向第一格式图像序列插入附加的场或者帧，以从中建立第二格式图像序列，所述第二格式图像序列的帧频率是恒定的并且大于所述第一格式图像序列的帧频率，所述方法包括下列步骤：

- 确定所述第一格式图像序列中的场或者帧位置，在该位置上插入相应的附加场或者帧，从而使得在第二格式图像序列中可见的运动抖动最小；
- 以非常规的场或者帧插入距离在所述第一格式图像序列中的一些所述位置插入场或者帧，从而使得总的来说任何临近帧之间平均距离对应于所述第二格式图像序列的临近帧之间的平均距离；
- 采用所述第二格式图像序列的格式一起呈现所述第一格式图像序列以及所述非常规插入的场和/或帧，

总的来说，本发明所记载的装置适用于控制向第一格式图像序列插入附加的场或者帧，以从中建立第二格式图像序列，所述第二格式图像序列的帧频率是恒定的并且大于所述第一格式图像序列的帧频率，所述装置包括用于确定在所述第一格式图像序列中的场或帧位置的部件，在该位置上插入对应的附加场或帧，从而使所述第二格式图像序列中的可见运动抖动最小，

所述装置还用于以非常规场或帧插入距离在一些所述位置上向所述第一格式图像序列插入场或帧，从而使得总体来说任何临近帧之间的平均距离对应于所述第二格式图像序列中任何临近帧之间的平均距离，

所述装置还用于以所述第二格式图像序列的格式一起呈现所述第一格式图像序列以及所述非常规插入的场和/或帧。

如前所述的装置，其中所述装置为光盘播放器或者光盘刻录机，或者硬盘录像机，例如 HDD 录像机或者 PC，或者机顶盒，或者电视接收器。

如前所述的装置，所述装置为光盘播放器或者光盘刻录机或者硬盘录像机或者机顶盒，其中所述装置输出所述原始的第一格式图像序列或者所述第二格式图像序列，这个选择是由所接收到的重放模式信息来确定的，所述重放模式信息自动地接收于与包括显示设备在内的设备相连的接口，或者接收于用户接口。

如前所述的方法或者装置，其中对分配给所述第一格式图像序列的所述音频信息中的语音的检测，是通过在多声道音频中进行评估，以判断所述中央声道相对于左和右声道是否表现出随时间爆发的能量分布，该分布明显区别于所述左和右声道的能量分布。

如前所述的方法或者装置，其中所述第一格式图像序列储存或者记录在存储介质上，例如光盘或者硬盘，或者作为数字电视信号被广播或者传播。

如前所述的方法或者装置，其中在所述第一格式图像序列中插入场或者帧的位置是不包含大块的运动图像内容区域的场或者帧，所述运动可以通过对运动矢量的评估来进行确定。

如前所述的方法或者装置，其中在所述第一格式图像序列中插入场或者帧的位置是发生场景变化或者逐渐变黑或者逐渐变白或者逐渐变为任一颜色的场或帧。

如前所述的方法或者装置，其中在所述被插入的场或者帧被输出到所述第二格式图像序列中之前，对所述被插入的场或者帧进行运动补偿。

如前所述的方法或者装置，其中所述第一格式图像序列为 MPEG-2 图像序列，并且通过评估标记对于向所述第一格式图像序列插入场或帧进行控制，所述评估标记指示场的时间顺序或者指示为了进行显示而对第一场进行重复，该标记在所述第一格式图像序列每幅图像的用户数据场中被传送。

附图说明

参照以下附图对本发明具体实施例进行详细描述，其中：

图 1 所示为光盘播放器的简要框图；

图 2 所示为在 24p 源图像序列上采用 3:2 下拉技术以提供 60i 的图像序列的应用；

图 3 所示为重复帧的常规排列；

图 4 所示为重复场的常规排列；

图 5 所示为根据图 3 进行常规帧重复的时间轴；

图 6 所示为视频序列对于运动抖动的示例容许值；

图 7 所示为进行场或帧重复的示例非常规时间位置以及所导致的变化呈现延迟；

图 8 所示为作为视频延迟以及运动抖动容许值函数的帧或场重复距离；

图 9 所示为图 8 所示的帧或场重复距离函数，其中所述最大和最小视频延迟取决于所要求的唇同步程度；

图 10 所示为包括没有进行运动补偿的重复帧的 24fps 格式帧；

图 11 所示为与图 10 相关的 25fps 格式帧输出；

图 12 所示为包括进行了运动补偿的重复帧的 24fps 格式帧；

图 13 所示为与图 12 相关的 25fps 格式帧输出。

具体实施方式

在图 1 中一个光盘驱动器从光盘 D 读入以 24p 格式编码的视频和音频信号，所述光盘驱动器包括读取头和纠错 PEC。所述输出信号经过一个轨道缓冲器以及多路分离器 TBM 分别到达视频解码器 VDEC 以及音频解码器 ADEC。控制器 CTRL 可以控制 PEC、TBM、VDEC 和 ADEC。用户接口 UI，和/或在电视接收器或显示器（没有示出）与所述光盘播放器之间的接口 IF，被用于将所述播放器的输出转换到 24fps 模式或者 25fps 模式。所述

接口 IF 可以自动检查所述电视接收器或者显示器可以处理并呈现的模式。所述重放模式信息自动地源自于接口 IF 接收到的特征数据（即关于所述电视接收器或者显示器可用的显示模式的数据），所述接口 IF 通过有线、无线电波或者光学方式与所述电视接收器或者显示设备相连。所述特征数据可以常规地由所述接口 IF 接收，或者通过向所述电视接收器或者显示设备发出相应的请求而得到。作为选择，所述重放模式信息可以通过向用户显示相应请求而通过所述用户接口 UI 被输入。在 25fps 的输出来自于所述视频解码器 VDEC、所述控制器 CTRL，或者来自于所述视频解码器 VDEC 自身的情况下，根据所述被解码视频信号的特性来确定在哪个时间位置上由所述视频解码器对场或帧进行重复。在本发明的一些实施例中，这些时间位置也如下所述地被所述来自于音频解码器 ADEC 的音频信号所控制。除了光盘播放器，本发明也可以被用于其他类型的设备，例如数字机顶盒或者数字电视接收器，在这种情况下包括所述光盘驱动器以及轨道缓冲器的前端会被数字信号调谐器所替代。

图 3 所示为重复帧的常规排列，其中为了实现从已知 24p 到 25fps 的转换，每 24 帧重复一帧，即在 t_n , t_n+1 , t_n+2 , t_n+3 , 等秒进行重复。图 4 所示为重复场的常规排列，其中为了实现从已知 24p 到 25fps 的转换，每 24 场重复一场，即在 t_n , $t_n+0.5$, t_n+1 , $t_n+1.5$, t_n+2 , 等秒进行重复。如果所述显示设备具有隔行扫描输出则可以采用该种处理方法。在所述时间轴上发生抖动的位置数目加倍，但是相对于所述帧重复，每次“抖动事件”的强度减半。顶部的场源自于所述源序列指定帧的第一、三、五等线，且底部的场来自于所述源序列指定帧的第二、四、六等线。图 5 所示为根据图 3 进行常规帧重复的时间轴，在标记着 t_n , t_n+1 , t_n+2 , t_n+3 , 等秒的时间位置发生帧重复。

为了实现在非等距（或者不规律）的位置适应性地插入重复场或者帧，就需要相应的控制信息。有关源材料的内容信息以及图像信号特征，在根据例如 MPEG-2 视频，MPEG-4 视频或者 MPEG-4 视频第 10 部分的机制对

所述图像序列进行压缩时即可获得，该信息和特征将不仅被用于产生广播和套装媒体例如 DVD，还可被用于未来的媒体例如基于蓝光技术的光盘。

在本发明中有用的图像信号特征或信息是：

- 被产生和/或传输的运动矢量，
- 由编码器产生的场景变换信息，
- 平均亮度或者强度信息，该信息可以源自于对 DC 转换系数的分析，
- 平均纹理强度信息，该信息可以源自于对 AC 转换系数的分析。

该图像信号特征可以作为 MPEG 用户数据或者专用数据通过光盘或者广播从所述编码器传递到所述解码器。作为选择，所述视频解码器可以收集或者计算并提供该信息。

为了对运动矢量信息进行利用，对每一帧的运动矢量组 MV 进行收集和处理，以判断当前帧是否包含大块的可见移动区域，因为这样的区域在复制帧或者场时受运动抖动影响最大。为了判断该种区域的出现情况，可以对一帧的平均绝对矢量长度 $AvgMV_i$ 进行计算，以作为平移运动的指示：

$$AvgMV_i = \frac{1}{VX \cdot VY} \sum_{x=0}^{VX-1} \sum_{y=0}^{VY-1} |MV_{x,y}|, \quad (1)$$

其中“i”代表帧号码，“VX”以及“VY”代表在图像 x（水平）和 y（竖直）方向的运动矢量数目。因此，一般将各方向上的图像大小除以运动估计的块大小可以分别得到 VX 和 VY。

如果在一帧内的运动矢量指向与当前帧相距不同的时间距离的不同参考帧，还需要一个对该距离进行规格化的因数 $RDist_{x,y}$ ：

$$AvgMV_i = \frac{1}{VX \cdot VY} \sum_{x=0}^{VX-1} \sum_{y=0}^{VY-1} \frac{|MV_{x,y}|}{RDist_{x,y}}. \quad (2)$$

在采用了更加复杂的处理过程的本发明的另一实施例中，对每幅图像的运动分割进行了计算，即确定由邻近并且具有相似的运动矢量长度和方向的块组成的一个或多个簇，从而对具有不同运动方向的足够大的运动区

域进行检测。在该种情况下，对所述平均运动矢量的计算如下：

$$AvgMV_i = \frac{\sum_{c=1}^{nClusters} AvgMV_c \cdot ClusterSize_c}{\sum_{c=1}^{nClusters} ClusterSize_c}, \quad (2a)$$

其中 $AvgMV_c$ 是确定簇“c”的平均运动矢量长度。

有利地，该方案排除了在图像内部随机移动的小对象带来的运动矢量影响，所述小对象并不是任何确定的块簇运动的组成部分，也不会对运动抖动的可见度带来很大影响。

所述处理过程会将所述移动区域的纹理是否非常粗糙或者具有非常尖锐的边缘作为 $AvgMV_i$ 的权重因数，因为这也会提高运动抖动的可见度。有关纹理强度的信息可以最方便的源自于对被传输或被接收或被重放的估计误差的AC转换系数的统计分析。总的来说，应该通过对原始图像块进行分析来确定纹理强度，然而在很多情况下，所述具有强烈纹理的块在采用运动补偿估计进行编码后，在这些块的AC系数中也会比具有较弱纹理的块更多的估计误差能量。因此，在所述视频序列的特定时间位置上的所述运动抖动容许值MJT可以被表达为：

$$MJT=f(AvgMV, \text{纹理强度}, \text{边缘强度}) \quad (3)$$

并具备以下一般特性：

- 已知确定的纹理强度值和边缘强度值，MJT 与 $1/AvgMV$ 成比例；
- 已知确定的 $AvgMV$ 值和边缘强度值，MJT 与 $1/(\text{纹理强度})$ 成比例；
- 已知确定的 $AvgMV$ 值和纹理强度值，MJT 与 $1/(\text{边缘强度})$ 成比例。

图 6 所示为源序列的运动抖动的示例容许值 MJT(t)。

优选地，如图 7a 所示，当前所述运动抖动容许值的大小影响重复帧或者场的分布，即所述帧或者场的重复距离 FRD，这些重复帧或者场被插入以得到 25fps 序列。如图 7b 所示，提早或者延迟插入重复帧会引起所述音频轨道相对于视频轨道的负向或者正向延迟，即视频的变化呈现延迟。当从运动抖动容许值 MJT 到帧或者场重复距离 FRD 之间建立映射的时候，要

对在两个方向上能容许的最大的视频相对于音频的延迟加以考虑。

图 8 所示为解决该控制问题的一个可能的方案。所述帧或者场重复距离 FRD 可以表达为所述视频延迟 VD 以及所述运动抖动容许值 MJT 的函数：

$$FRD = f(VD, MJT), \quad (4)$$

并具备以下一般特性：

- 已知确定的 VD 值，FRD 与 $1/MJT$ 成比例；
- 已知确定的 MJT 值，FRD 与 $1/VD$ 成比例；

如图 8 所示，这个关系可以通过 $FRD=f(VD)$ 的特性来表达，该特性的变化取决于所述运动抖动容许值，在运动抖容许值低的情况下（例如高度运动的情况），倾向于比被插入帧之间的最优距离长的距离，并在运动抖动容许值高的情况下（例如低于平均运动水平的情况），倾向于比最优距离短的距离。所述最优场或者帧重复距离由 FRD_{opt} 所代表。所允许的最大视频延迟由 VD_{max} 所代表。在负方向所允许的最大视频延迟由 VD_{min} 所代表。

由于在场景变化位置的短暂停帧效果并不被认为是令人厌烦的，由视频编码器产生的（或由视频解码器产生的）场景变化信息可以用于在该位置上插入一个或多个重复的场或者帧，所重复的次数取决于当前视频延迟的程度。由于相同的原因，可以在逐渐变黑序列、逐渐变白序列或者逐渐趋近于任一颜色的序列之后插入重复的场或者帧。所有这样的特定位置都具有非常高的 MJT 值。

很显然，在所述位置可以采用重复的帧，尽管在其他图像内容中在个别的位置上仅对场进行重复以降低运动抖动强度。总的来说，重复的帧和重复的场可以在被转换的图像序列中共存。

对于可感知的唇同步来说，通常能够接受的延迟范围仅仅在至少有一个演讲者出现在所述场景以内的时候才需要被遵守。因此，当没有演讲者出现的时候，在音频和视频呈现之间的延迟可以大于上述范围。在快速运动场景的情况下通常是这种情况。因此，可以如图 9 所示进行附加控制，

从而可以使所述视频延迟范围 VD_{min} 和 VD_{max} 在以下数值之间进行平稳的转换或者调整:

- 在唇同步可接受值 $VD_{minLipSync}$ 和 $VD_{maxLipSync}$ 之间, 如果所检测到语音或者短的声音峰值(由特殊事件如拍门引起的)以及缓慢移动或者静态的场景;
- 否则就在更大的 VD 值 VD_{min} 和 VD_{max} 之间。

对语音的检测可以源于在例如最常用的多声道音频情况下对与左右声道相关的中央声道进行的评估, 因为在电影中语音通常被编码入所述中央声道。如果所述中央声道表现出随着时间爆发的能量分布, 并且该分布显著的区别于所述左右声道的能量分布, 那么出现语音的可能性就很高。

上述为了适应性地确定所述本地帧重复距离而进行的控制对于通过视频序列的一次运算来说是有效的。然而, 所述控制受益于在许多专业MPEG-2编码器中进行的二重运算编码处理。在该种情况下, 所述二重运算的第一次运算用于收集所述运动强度曲线, 场景切换位置以及需要严格唇同步的场景的序号、标号、位置以及长度, 黑帧, 等等。随后应用经过改进的控制机制, 该机制不仅考虑到目前处理的帧及其过去帧的可用信息, 还对过去以及未来帧的相邻帧信息加以考虑:

$$FRD(i)=f(VD, MJT(i-k) \dots MJT(i+k)), \quad (5)$$

其中“i”代表当前帧的号码, 且“k”代表以临近帧为参考的流水号。每个这样的函数的普遍特性是如果 $MJT(i)$ 小于周围的 MJT 值则 FRD 增大, 且如果 $MJT(i)$ 大于周围的 MJT 值的则 FRD 减小。相关的图像信号特性可以作为MPEG用户数据或者专用数据通过光盘或者广播信号从所述编码器向所述解码器进行传递。

在本发明的另一个实施例中, 在特定情况下采用运动补偿插值帧而不是重复帧并不会产生运算方面的花费。该运动补偿插值可以将被传输的运动矢量应用于当前帧。一般来说, 该运动矢量不适用于运动补偿帧插值, 因为这些运动矢量都被优化以实现最优的估计增益而不是指示真实的运动

场景。然而，如果经过解码器对所接收到的运动矢量进行分析后表明所述场景发生了同类的平移，则可以在当前帧和前一帧之间插入高度精确的一帧。平移的含义是一帧内所有的运动矢量在长度以及方向上是相同或者几乎相同的。因此，被插值帧的产生可以通过以当前帧的平均运动矢量所指示距离的一半对前一帧进行转换来实现。所述前一帧被认为是所述当前帧运动补偿估计的参考帧，并且所述被插值帧被置于与所述前一帧和所述当前帧等距的位置。如果所述估计帧不是所述前一帧，则要对平均运动矢量进行必要的缩放。要对根据接收到的运动矢量而确定存在缩放的情况加以相应的考虑。缩放的特征是缩放中心的运动矢量为零，并且在该缩放中心周围按照从中心向外（或者从外向中心）的方向运动矢量长度增加，所述运动矢量长度的增加与到所述缩放中心的距离相关。

有利地，如图 10 到 13 所示，相对于重复帧该种运动补偿插值在运动抖动表现方面作出了改进。图 10 和图 11 显示了带有竖直运动对象的运动轨迹的帧（由竖直的条代表）以及帧重复的一种情况，这会导致‘停顿’现象，其中图 10 为采用了 24fps 的格式的情况，图 11 所示为进行 25fps 格式转换后的情况。图 12 所示为以图 13 所示的提高了的 25fps 目标帧速率进行呈现时插入运动插值帧的情况，这将导致‘缓慢运动帧’而不是‘停顿’。

如上所述为了实现帧速率转换可以实施对帧和/或场重复以及插值的控制，该控制可以应用于 MPEG-2（或者类似）压缩系统的编码器和解码器端，因为在该两端可以获得大部分的端信息，可能除了可靠的场景变化指示以外。然而，为了利用所述解码器的高级图像序列特性信息，可以在（MPEG-2 或者其他）被压缩的 24fps 视频信号中传送所要进行重复或者插值的场或者帧位置信息。在 MPEG-2 语法中已经存在指示场时间顺序的标记（top_field_first）和为了进行显示而对第一场进行重复的标记（repeat_first_field）。如果要求以信号发出针对同一视频信号的从 24fps 到 30fps 和从 24fps 到 25fps 的转换模式，在每幅图像的合适的用户数据场中要对上述两系列标记中的一系列进行传送。

上面提到的 24fps 值和 25fps 值以及其他的数字都是示例性的数值，这些值可以相应的用于本发明的其他应用之中。

本发明可以应用于：

- 套装媒体（DVD，蓝光光盘，等），
- 下载媒体包括视频点播，近视频点播，等，
- 广播媒体。

本发明可以用于光盘播放器或者光盘刻录机，或者硬盘录像机例如 HDD 录像机或者 PC，还可以应用于机顶盒，或者电视接收器。

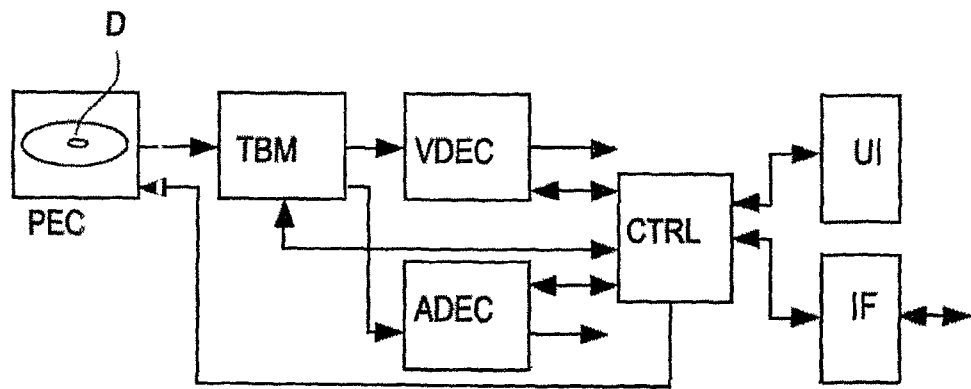


图 1

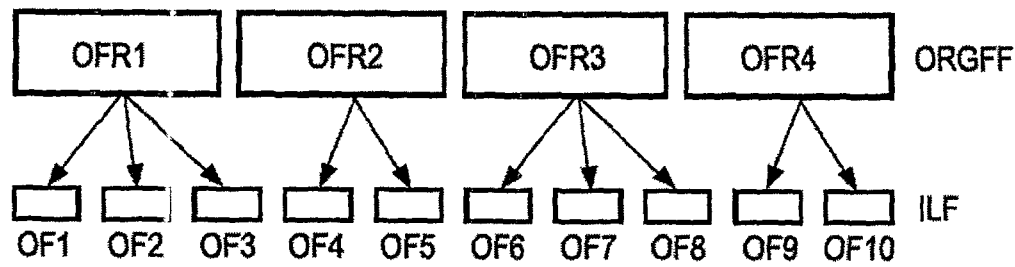


图 2

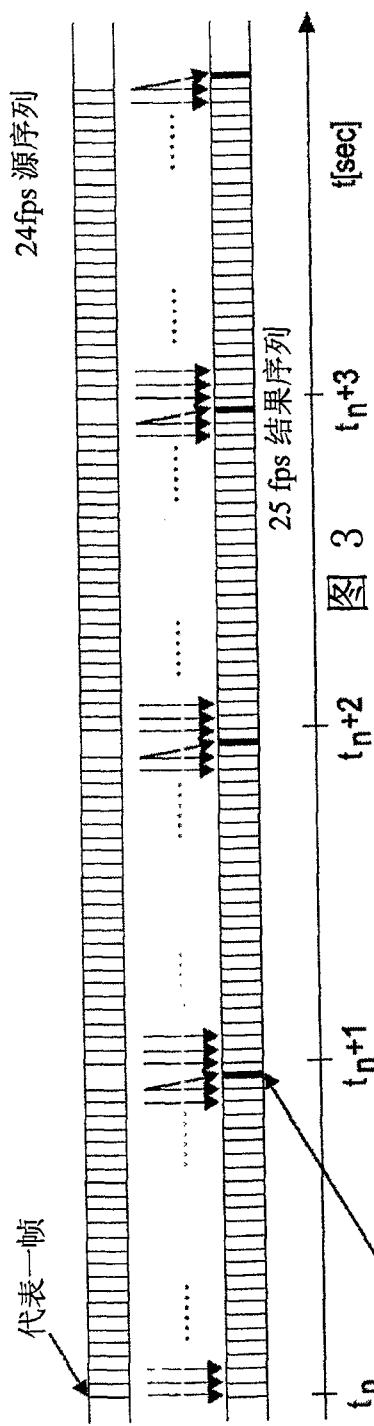


图 3

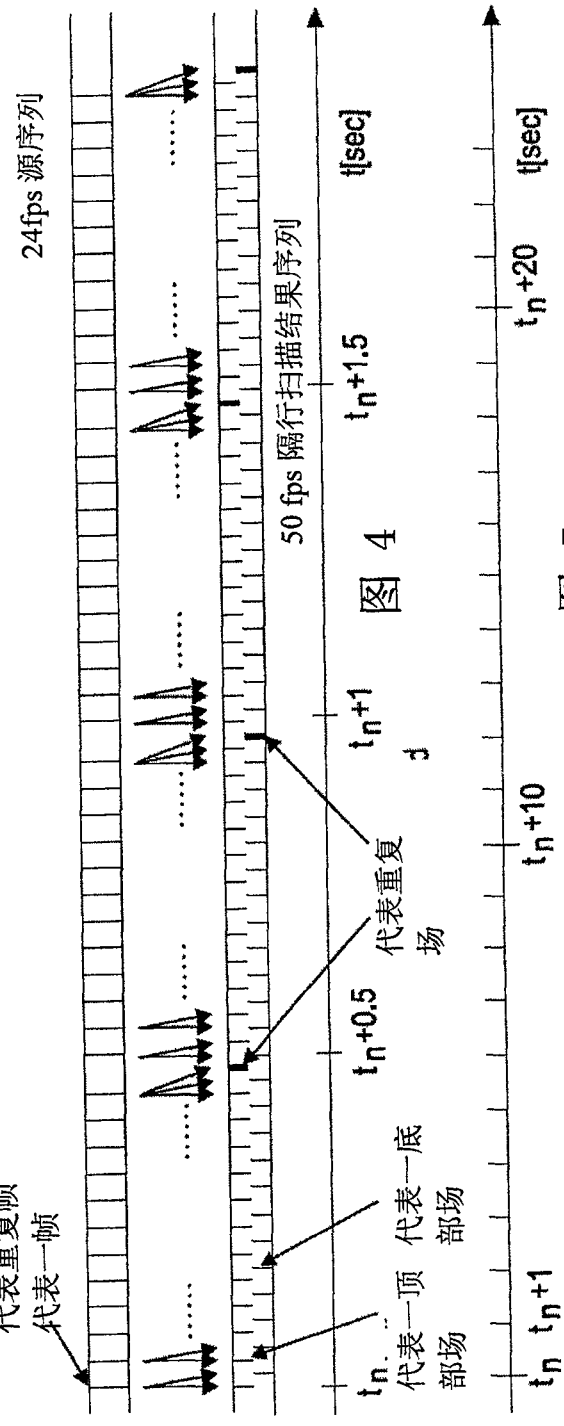
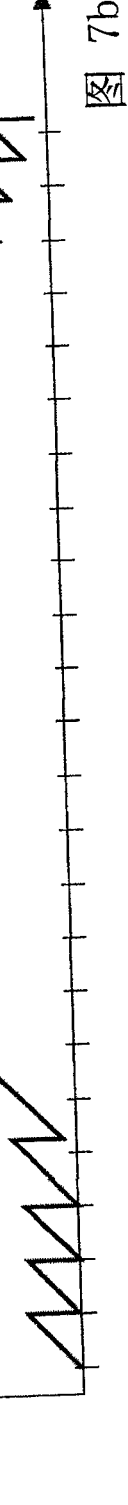
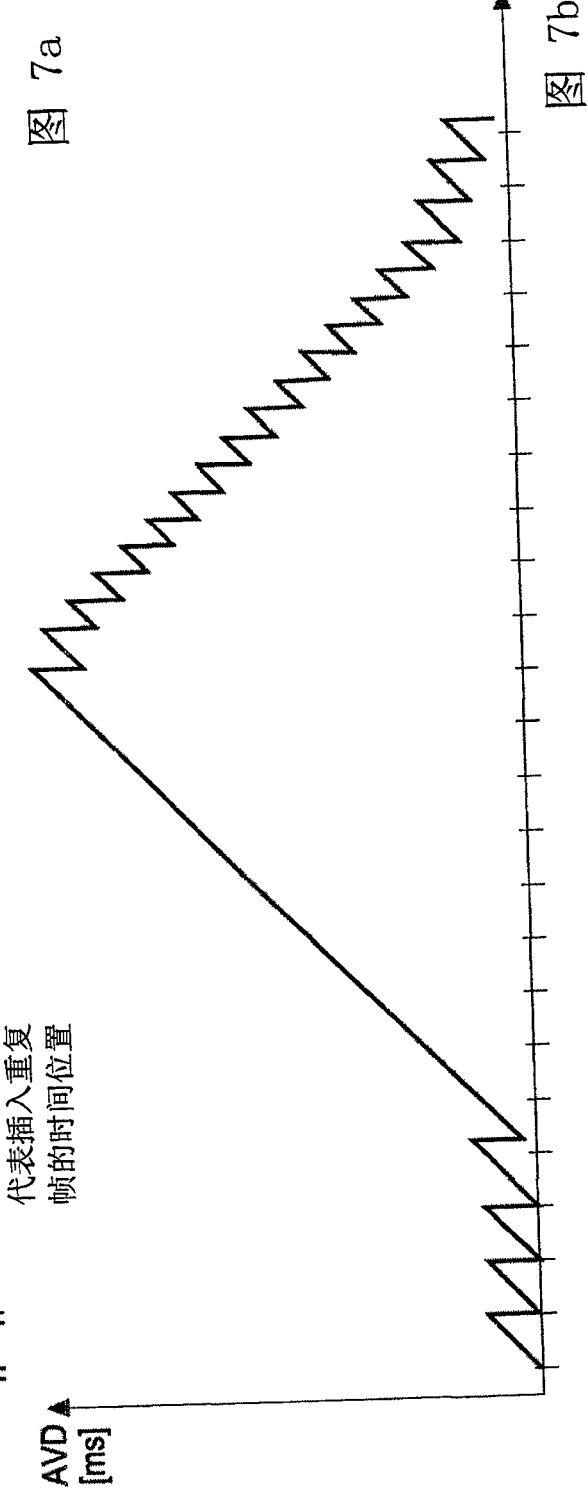
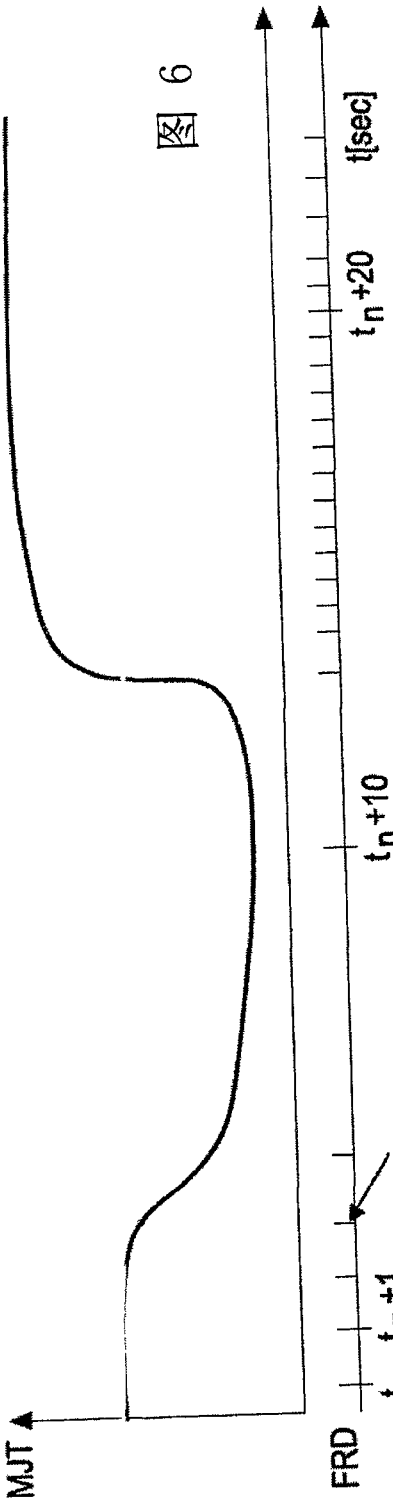


图 4

图 5



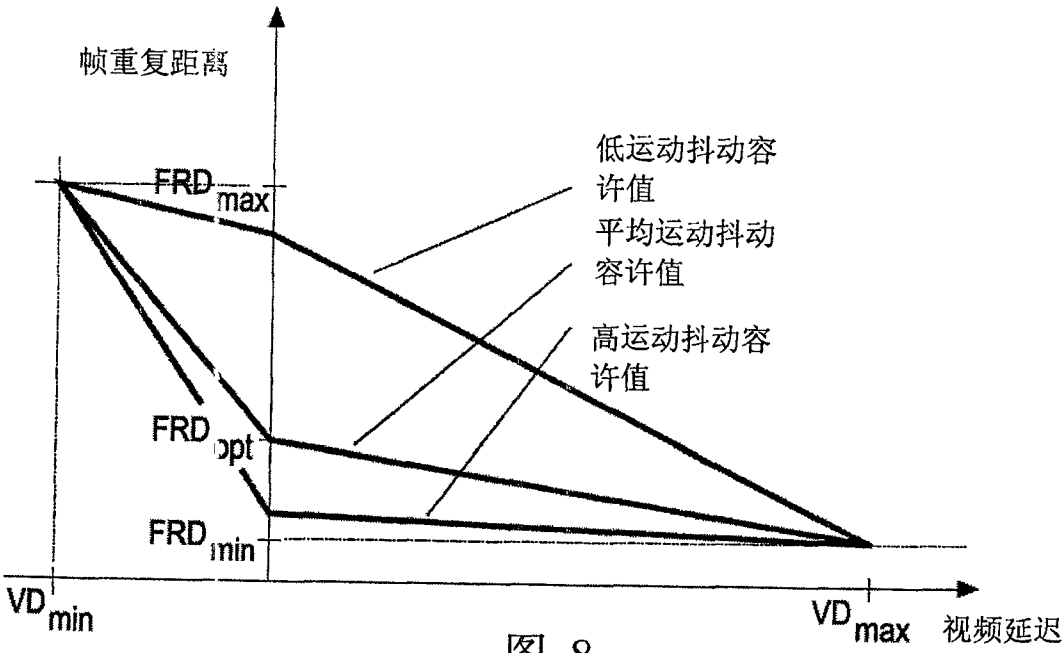


图 8

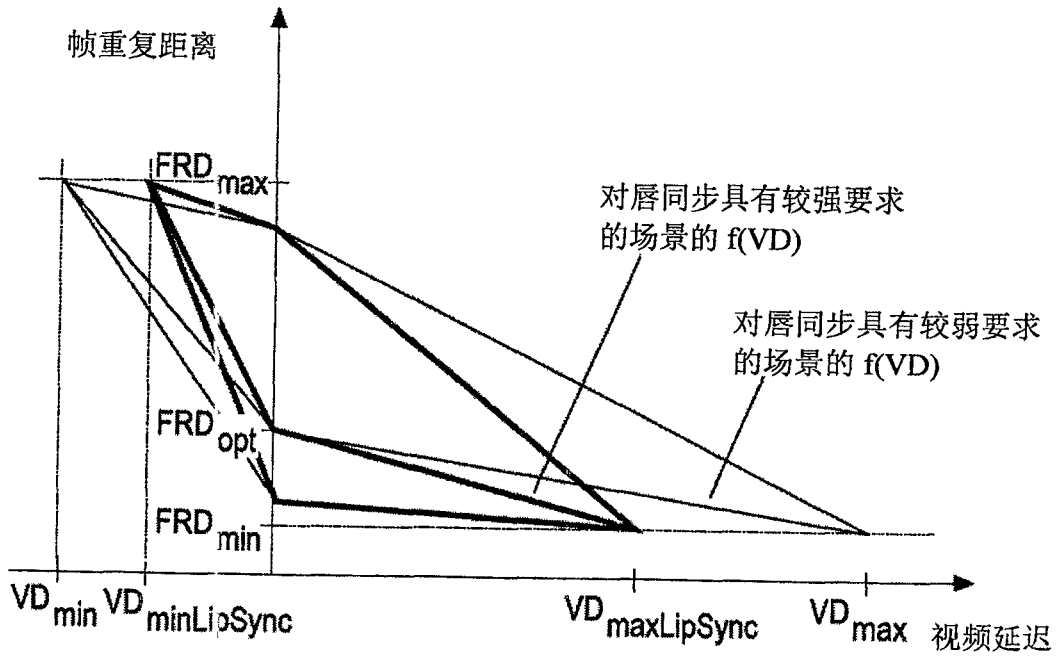


图 9

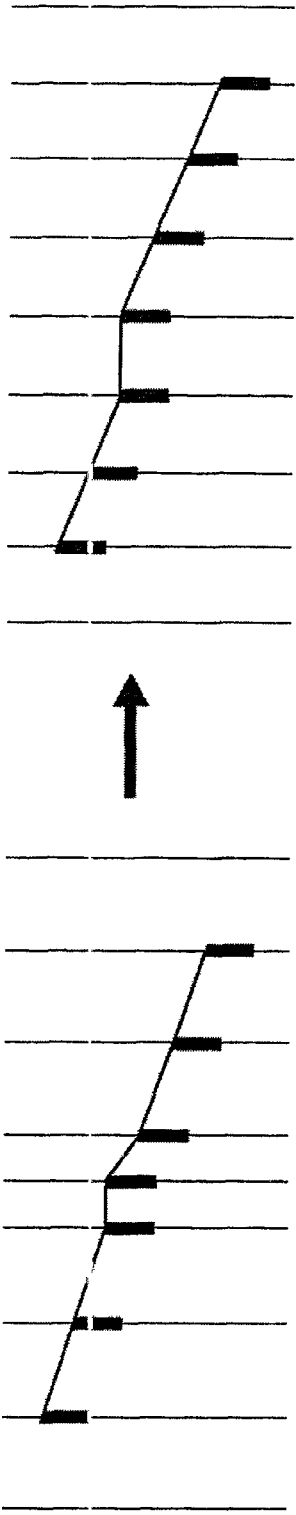


图 10

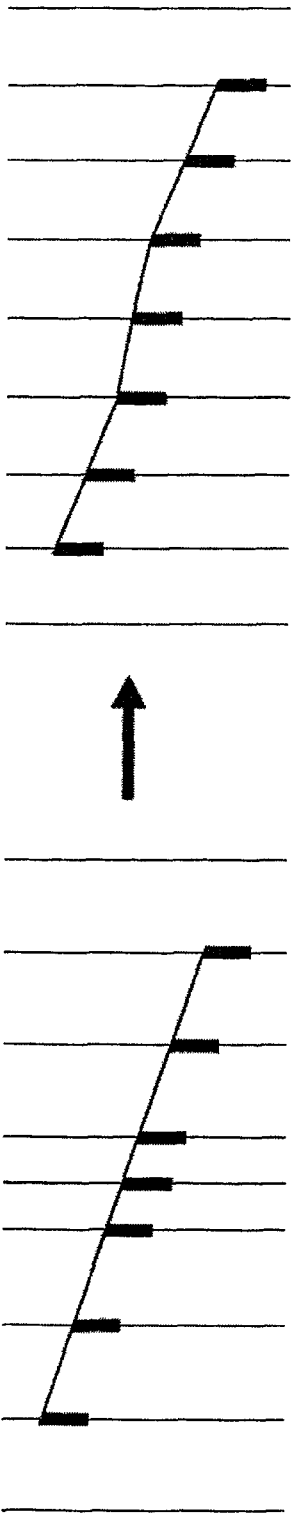


图 12



图 11

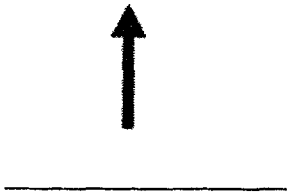


图 13