

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

H04N 7/26

H04N 7/46 H04N 7/50

H04N 7/36



[12] 发明专利申请公开说明书

[21] 申请号 01803961.8

[43] 公开日 2003 年 3 月 5 日

[11] 公开号 CN 1401189A

[22] 申请日 2001.11.16 [21] 申请号 01803961.8

[30] 优先权

[32] 2000.11.23 [33] EP [31] 00403281.9

[32] 2001.9.11 [33] EP [31] 01402341.0

[86] 国际申请 PCT/EP01/13450 2001.11.16

[87] 国际公布 WO02/43399 英 2002.5.30

[85] 进入国家阶段日期 2002.7.22

[71] 申请人 皇家飞利浦电子有限公司

地址 荷兰艾恩德霍芬

[72] 发明人 Y·R·拉曼津

[74] 专利代理机构 中国专利代理(香港)有限公司

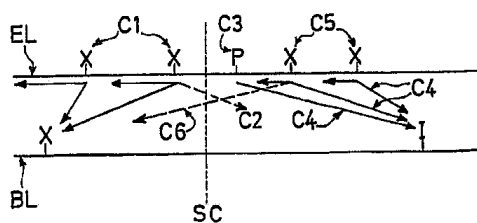
代理人 陈 霁

权利要求书 2 页 说明书 7 页 附图 1 页

[54] 发明名称 视频编码方法和相应的编码器

[57] 摘要

MPEG-4 视频标准包括一个预测编码方案。当在由所述的编码方案处理的序列中发生场景剪辑的时候，紧随着它的第一视频对象平面(VOP)被编码作为 I-VOP，而不是根据与前一个其完全不同的 VOP 对其进行预测。在时间可量测性的情况下，当在增强层的两个 VOP 之间发生场景剪辑的时候，定义了在该增强层中的预测操作期间选择时间参考的特定规则。



ISSN 1008-4274

1. 一种用在视频编码器中的用于视频序列的视频编号方法，所述的视频编码器包括基础层编码装置，用于接收所述视频序列并产生基础层信号，该信号与包含在所述序列的视频帧中的视频对象（VO）相应，并构成了适于以基础层比特率传送给视频解码器的第一比特流，以及增强层编码装置，用于接收所述的视频序列以及所述的基础层信号的解码版本，产生增强层信号，该信号与相应的基础层信号相联系并适于以增强层比特率传送给所述的视频解码器，所述的视频编码方法包括以下步骤：
- 5
- 10 (1) 将所述视频序列分割成所述的VO；
- (2) 编码每一个所述的VO的连续视频对象平面（VOP），所述的编码步骤本身包括编码所述的VOP的纹理和形状的子步骤，所述的纹理编码子步骤本身包括不对VOP进行预测的第一编码操作，被称为内编码或者I-VOP，不用任何对其它VOP的时间参考而编码，对VOP进行单向预测的第二编码操作，被称为预测或P-VOP，仅使用前一个或者后一个I-或P-VOP作为时间参考编码，以及对VOP进行双向预测的第三编码操作，被称为双向预测或B-VOP，使用前一和后一I-或P-VOP作为时间参考编码，当场景剪辑发生并且所述增强层VOP位于场景的最后一个基础层VOP以及下一场景的第一个基础层VOP之间的时候，所述的增强层VOP的时间参考的选择要遵循下面的特定处理规则：
- 15
- 20 (A) 位于场景剪辑以前的VOP：
- (a) 对编码类型没有限制；
- (b) 禁止使用基础层显示顺序中的下一个VOP作为时间参考；
- (B) 正好紧挨着场景剪辑以后的VOP：
- (a) P编码时间被执行；
- 25 (b) 将基础层显示顺序中的下一个VOP用作为时间参考；
- (C) 位于场景剪辑之后的其它VOP：
- (a) 对编码类型没有限制；
- (b) 禁止使用基础层显示顺序中的前一个VOP作为时间参考。
2. 存储在计算机可读介质上的并用于执行根据权利要求1的视频编码方法的计算机可执行处理步骤。
- 30

3. 一种在视频编码器中用于在所述的编码器中执行权利要求1的所述编码方法的计算机程序产品, 所述产品包括一组指令, 当其被装载到所述的编码器时, 使其执行所述方法的步骤。

4. 一种视频编码器, 包括基础层编码装置, 用于接收视频序列并产生基础层信号, 该信号与包含在所述序列的视频帧中的视频对象 (VO) 相应, 并构成了适于以基础层比特率传送给视频解码器的第一比特流, 以及增强层编码装置, 用于接收所述的视频序列以及所述的基础层信号的解码版本, 产生增强层信号, 该信号与相应的基础层信号相联系并适于以增强层比特率传送给所述的视频解码器, 所述的视频编码器包括:

10 (1) 将所述视频序列分割成所述的VO的装置;

(2) 用于编码连续视频对象平面 (VOP) 的纹理和形状的装置, 所述的纹理编码装置执行不对所述VOP进行预测的第一编码操作, 被称为内编码或者I-VOP, 不用任何对其它VOP的时间参考而编码, 对VOP进行单向预测的第二编码操作, 被称为预测或P-VOP, 仅使用前一个或者后一个I-或P-VOP作为时间参考编码, 以及对VOP进行双向预测的第三编码操作, 被称为双向预测或B-VOP, 使用前一和后一I-或P-VOP作为时间参考编码, 其特征在于当场景剪辑发生并所述且增强层VOP位于场景的最后一个基础层VOP以及下一场景的第一个基础层VOP之间的时候, 所述的增强层VOP的时间参考被选择, 这要遵循下面的特定的处理规则:

20 (A) 位于场景剪辑以前的VOP:

(a) 对编码类型没有限制;

(b) 禁止使用基础层显示顺序中的下一个VOP作为时间参考;

(B) 正好紧挨着场景剪辑以后的VOP:

(a) P编码时间被执行;

25 (b) 将基础层显示顺序中的下一个VOP用作为时间参考;

(C) 位于场景剪辑之后的其它VOP:

(a) 对编码类型没有限制;

(b) 禁止使用基础层显示顺序中的前一个VOP作为时间参考。

视频编码方法和相应的编码器

5 技术领域

本发明通常涉及视频压缩，尤其是涉及一种应用于视频序列以及在视频编码器中使用的视频编码方法，该编码器包括：基础层编码装置，用于接收所述的视频序列，并从中产生基础层信号，该信号与包含在所述序列的视频帧中的视频对象(VO)相应、并构成适于以基础层比特率传送给视频解码器的第一比特流；以及增强层编码装置，用于接收所述的视频序列以及所述的基础层信号的解码版本，并从中产生增强层信号，该信号与相应的基础层信号相联系的并适于以一个增强层比特率传送给所述的视频解码器。更精确的说，它涉及一种允许编码所述序列的VO的方法，并且该方法包括以下步骤：

- (1) 将视频序列分割成所述的VO；
- 15 (2) 编码每个所述的VO的连续视频对象平面 (VOP)，所述的编码步骤本身包括编码所述VOP的纹理和形状的子步骤，所述的纹理编码子步骤本身包括不对VOP进行预测的第一编码操作，被称为内编码或者I-VOP，不对另一个VOP进行任何时间参考而编码，对VOP进行单向预测的第二编码操作，被称为预测或P-VOP，仅使用前一个或者后一个I-或P-VOP作为时间参考编码，以及
- 20 对VOP进行双向预测的第三编码操作，被称为双向预测或B-VOP，使用前一和后一I-或P-VOP作为时间参考编码。

本发明也涉及一种存储在计算机可读介质之上并用于执行这样的编码方法的计算机可执行处理步骤，一种相应的计算机程序产品以及执行所述的方法的视频编码器。

25 背景技术

在一个应用MPEG-4标准(所述的标准例如在文章“MPEG-4版本1标准的回顾”中被描述,ISO/IEC JTC1/SC29/WG11 N1909,1997年10月, Fribourg)的编码器中，使用了三种类型的图像：内编码图像 (I)，独立于其它的图像进行编码，预测编码图像 (P)，通过运动补偿预测根据前一个参考图像(I或P)进行预

30 测，以及双向预测编码图像 (B)，根据前一和接下来的参考图像 (I或P) 进行

预测。I图像是最重要的，因为它们是参考图像并能够提供入点（在比特流中），在该点解码可以不参照任何先前的图像而进行（在这样的图像中只有空间冗余被消除）。为了既减少空间冗余，又减少时间冗余，P图像提供了一个比仅减少空间冗余的I图像更好的压缩。B图像提供了最好的压缩程度。

- 5 在MPEG-4中，使用了几种结构，例如视频对象（VO），它们是用户可以访问和操作的实体，以及视频对象平面（VOP），它们是在一个给定时间的视频对象的实例。在一编码比特流中，可以发现几种不同类型的VOP：内编码VOP，仅使用空间冗余（就比特而言是最昂贵的），预测编码VOP，根据前一参考VOP使用运动估计和补偿，以及双向预测编码VOP，根据前一以及接下来的
- 10 参考VOP使用运动估计和补偿。

- 对于P-VOP以及B-VOP，只有当前的VOP以及它的参考VOP（s）之间的差才被编码。只有P-和B-VOP涉及到运动补偿，根据所谓的“块匹配法则”执行：对于当前帧的每个宏块，在预定的搜索区域内寻找参考VOP中最匹配的宏块，并且计算一个运动矢量MV。下面通过绝对差的和（SAD）给出了类似的
- 15 标准。对于一个NxN宏块，SAD被表示为：

$$SAD = \sum_{i=0}^{NxN} |A(i) - B(i)|$$

这样，选中的宏块就与在一个搜索区域中计算出来的SAD最小的块相对应。对于所述的估计，存在不同的模式，这依赖于帧的类型：

- (a) 对于P-VOP宏块，只有“前向模式”适用(使用前面的参考I-VOP或者P-VOP)
- 20 (b) 对于B-VOP宏块，在宏块估计中可以使用四种模式：
- “前向模式”，（如对于P-VOP）；
 - “后向模式”：如前向模式，只是参考不再是前面的图像，而是接下来的P-或I-VOP；
 - “内插模式”或“双向模式”：它包括前向和后向模式并且使用前一以及
- 25 接下来的VOP；
- “直接模式”，每一运动矢量的计算都源于接下来的参考VOP的运动矢量以及不同的VOP之间的时间距离。

在MPEG-4中，提供一个非常重要的功能，可量测性。

可升级编码，也被称为“可分层编码”，允许以这样的方式产生一个编码表示，从而使得能够进行可升级解码操作。可伸缩性是允许对数据的合适子集进行解码的比特流特性，这导致了分辨率和/或质量与解码比特流的比例相适应的完整图像的产生。这样的功能在许多需要视频序列以各种分辨率和/或质量和/或复杂性上同时获得的应用中是有用的。实际上，如果比特流是可升级的，用户可以根据他的解码器或者显示器，或者根据他的可用带宽仅访问其一部分从而提供基础的视频，而其它的用户可以使用整个比特流从而产生一个较好的视频质量。

可伸缩性，就编码过程而言，成本要低于根据几个独立的比特流的分别编码的解决方案，其优势在于，它允许将一个比特流至少分成两个不同的比特流（并且，其中，一个的比特率要大于另一个）发送。因此可伸缩性的每一种类型要涉及一个以上的层。在时间可伸缩的情况下，至少要考虑包括一个低级层以及一个高级层的两层。低级层被称为基础层，以一个给定的帧速率进行编码，剩下的层被称为增强层，被编码从而提供在基础层中的信息丢失（为了形成一个具有较高的帧速率的视频信号），并且在显示端提供一个较高的瞬时清晰度。一个解码器可以仅解码基础层，其与解码视频流所需的最小数量的数据相应，或者也解码增强层（除了基础层），所述的增强层与提供一个增强视频信号所需的附加数据相应，并且如果需要一个更高的分辨率就每秒输出更多的帧。

如上所述，MPEG-4视频标准具有一个预测编码方案。当发生一个场景剪辑的时候，编码紧随所述场景剪辑的第一个VOP作为I-VOP是非常高效的，而不是试图根据前一个与其是完全不同的VOP进行预测。在时间可伸缩的情况下，问题更复杂，因为场景剪辑可能发生在增强层的两个VOP之间，而它不得不在基础层进行处理。如果在每一层第一个VOP被编码为I-VOP，这将导致比特的浪费以及编码效率的损失。

发明内容

因此本发明的一个目的是提供一种允许在场景剪辑的情况下降低所述编码效率的损失的方法。

为此，本发明涉及一种在本说明书的介绍部分定义的编码方法，并且其特征在于当场景剪辑发生并且增强层VOP位于场景的最后一个基础层VOP以及下

一场景的第一个基础层VOP之间的时候，所述的增强层VOP的时间参考被选择，这要遵循下面的特定的处理规则：

- (A) 位于场景剪辑以前的VOP：
 - (a) 对编码类型没有限制；
- 5 (B) 禁止使用基础层显示顺序中的下一个VOP作为时间参考；
- (B) 正好紧挨着场景剪辑发生以后的VOP：
 - (a) P编码时间被执行；
 - (b) 将基础层显示顺序中的下一个VOP用作为时间参考；
- (C) 位于场景剪辑之后的其它VOP：
 - 10 (a) 对编码类型没有限制；
 - (b) 禁止使用基础层显示顺序中的前一个VOP作为时间参考；

这种解决方案的主要优点在于它允许仅编码一个内VOP，而避免了无效的场景间预测。

- 本发明也涉及存储在计算机可读介质上的并用于执行所述视频编码方法的
- 15 计算机可执行处理步骤，也涉及一种包括一组指令的计算机程序产品，当它们被装载到一个所述的编码器中时，就使其执行该方法的步骤。本发明也涉及一种视频编码器，它包括：基础层编码装置，用于接收视频序列并产生基础层信号，该信号与包含在所述序列的视频帧中的视频对象（VO）相应，并构成了适于以基础层比特率传送给视频解码器的第一比特流；以及增强层编码装置，
- 20 用于接收所述的视频序列以及所述的基础层信号的解码版本，产生增强层信号，该信号与相应的基础层信号相联系并适于以增强层比特率传送给所述的视频解码器，所述的视频编码器包括：

- (1) 将视频序列分割成所述的VO的装置；
- (2) 用于编码连续视频对象平面（VOP）的纹理和形状的装置，所述
- 25 的纹理编码装置执行不对VOP进行预测的第一编码操作，被称为内编码或者I-VOP，不用任何对其它VOP的时间参考编码，对VOP进行单向预测的第二编码操作，被称为预测或P-VOP，仅使用前一个或者后一个I-或P-VOP作为时间参考编码，以及对VOP进行双向预测的第三编码操作，被称为双向预测或B-VOP，使用之前的和之后的I-或P-VOP作为时间参考编码，其特征在于当场
- 30 景剪辑发生并且增强层VOP位于场景的最后一个基础层VOP以及下一场景的第

一个基础层VOP之间的时候，所述的增强层VOP的时间参考被选择，这要遵循下面的特定处理规则：

- (A) 位于场景剪辑以前的VOP：
 - (a) 对编码类型没有限制；
 - 5 (b) 禁止使用基础层显示顺序中的下一个VOP作为时间参考；
- (B) 正好紧挨着场景剪辑后的VOP：
 - (a) P编码时间被执行；
 - (b) 将基础层显示顺序中的下一个VOP用作为时间参考；
- (C) 位于场景剪辑后的其它VOP：
 - 10 (a) 对编码类型没有限制；
 - (b) 禁止使用基础层显示顺序中的前一个VOP作为时间参考；

附图说明

本发明将参照随后的附图以更为详细的方式进行描述：

- 15 附图1描绘了在一个可升级MPEG-4视频流的基础层和增强层中时间参考的选择；

附图2描绘了根据本发明为时间参考的所述选择处理而定义的特定规则。

具体实施方式

- 如上所述，在一个编码比特流中可以发现不同类型的VOP：内编码VOP，预测编码VOP以及双向预测VOP。由于MPEG-4视频标准是一个预测编码方案，因此必须为每一个编码的非内VOP定义时间参考。在单层的情况下或者在可升级流的基础层（BL）中，时间参考以唯一方式被所述标准所定义，如图1所示(TR11,TR12,TR13)。相反，对于MPEG-4流的时间增强层（EL），有三种VOP可以作为运动预测中的可能的时间参考：增强层的最接近的解码VOP（TR21,TR22），或者基础层(以显示顺序)的前一VOP（TR23,TR24），或者基础层(以显示顺序)的下一个VOP（TR25,TR26）。对于时间增强层的P-VOP以及B-VOP的这三个可能的选择被显示在图1中（每一个箭头对应于一个可能的时间参考）：对增强层的每一个P-VOP必须选择一个参考，对每一个B-VOP必须选择两个。
- 20
- 25

- 现在本发明将与图2相联系进行描述。理论上，为了满足压缩效率的限制，每一场景变化仅需要一个I-VOP。由于在基础层中没有时间参考的可能选
- 30

择，因此决定编码在场景剪辑以后的基础层的第一个VOP为I-VOP（所述的场景剪辑参照图2中的SC），以及在增强层中禁止内编码以及场景间预测，特别是对位于场景的最后一个基础层VOP以及接下来的场景的第一个基础层VOP之间的所有的增强层VOP。这些编码规则是这样的：

- 5 (1) 位于场景剪辑之前的VOP：(a) 条件C1:对编码类型没有限制，以及 (b) 条件C2: 禁止使用在基础层的显示顺序中的下一个VOP作为时间参考：
 - (2) 紧挨着场景剪辑之后的VOP：(a)条件C3: 执行P编码类型，以及 (b) 条件C4: 所述的P编码使用基础层的显示顺序中的下一个VOP作为时间
- 10 参考：
 - (3) 位于场景剪辑之后的其它VOP：(a)条件C5: 对编码类型没有限制，以及(b)条件C6: 禁止使用基础层显示序列的前一VOP作为时间参考。

与这三种情况相联系的六种条件被显示在图2中，其中X相应于任何编码类型的VOP。比较图1和图2，可以清楚的看出下述条件被满足：条件C2，即对

15 于位于场景剪辑以前的VOP不使用下一个VOP；C4，即将基础层的下一个VOP作为时间参考；以及C6，即没有基础层的前一VOP被用作时间参考。

这样描述的视频编码方法然后通过一个视频编码器来执行，该编码器一方面包括基础层编码装置，用于接收视频序列并产生基础层信号，该信号与包含在所述序列的视频帧中的视频对象（VO）相应，并构成了适于以基础层比特

20 率传送给视频解码器的第一比特流，以及另一方面包括增强层编码装置，用于接收所述的视频序列以及所述的基础层信号的解码版本，产生增强层信号，该信号与相应的基础层信号相联系并适于以增强层比特率传送给所述的视频解码器。该编码器包括将视频序列分割成所述的VO的装置；以及用于编码每个所述VO的连续视频对象平面(VOP)的纹理和形状的装置，所述的纹理编码装置

25 执行三种类型的编码操作：第一种不对VOP进行预测，被称为内编码或者I-VOP，第二种对VOP进行单向预测，被称为预测或P-VOP，以及第三种对VOP进行双向预测，被称为双向预测或B-VOP。

根据本发明在所述的编码器中，当场景剪辑发生，并且所述增强层VOP位于场景的最后一个基础层VOP以及下一场景的第一个基础层VOP之间的时候，

30 所述的增强层VOP的时间参考将遵循上面定义的处理规则进行选择：

(A) 位于场景剪辑以前的VOP: 对编码类型没有限制并且禁止将基础层的下一个VOP用作为时间参考;

(B) 正好紧挨着场景剪辑以后的VOP: P编码时间被执行并且将基础层的下一个VOP用作为时间参考;

5 (C) 位于场景剪辑之后的其它的VOP:对编码类型没有限制并且禁止将基础层的前一个VOP用作为时间参考;

应该明白, 上面描述的视频解码器能够采用硬件或者软件来实现, 或者通过硬件与软件的结合来实现。然后可以通过任何类型的计算机系统或者其他适于执行上述的方法的装置来实现。一个典型的硬件与软件的结合是通用计算机系统
10 系统和计算机程序, 当被下载和执行的时候, 控制计算机系统使其执行上述的方法。此外, 可以利用一个包括用于执行本发明的一个或者多个功能的特定硬件的特定用途计算机。本发明也可以被嵌入到一个计算机程序介质或者产品中, 它包括能够执行上面描述的方法和功能的所有特征, 以及当被装载到一个计算机系统时, 它能够执行这些方法和功能。本发明也涉及存储在这样的计算机可
15 读介质或产品上的并且用于执行所述的视频编码方法的计算机可执行处理步骤。计算机程序, 软件程序, 程序, 程序产品或者软件, 在本文中的意思是一组指令的任何语言、编码或者符号的任何表达, 该组指令使得一个具有信息处理能力的系统直接或者在下面的两者之一或全部完成以后执行一个特定的功能, (a) 转换成其它语言, 编码或者符号, 和/或(b)以一个不同的物质形式再
20 现。

前面的对本发明的描述是为了举例和描述, 并不是穷举, 也不是将发明限制在公开的精确形式中, 在上面的指教下可以做出各种变形。这样的变形对于本领域的技术人员来说是明显的, 并也包含在本发明的范围内。

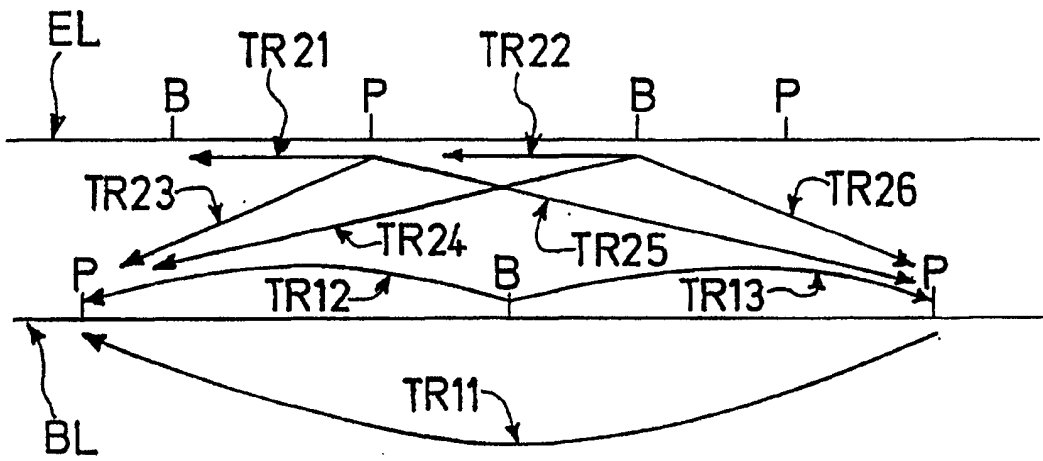


图 1

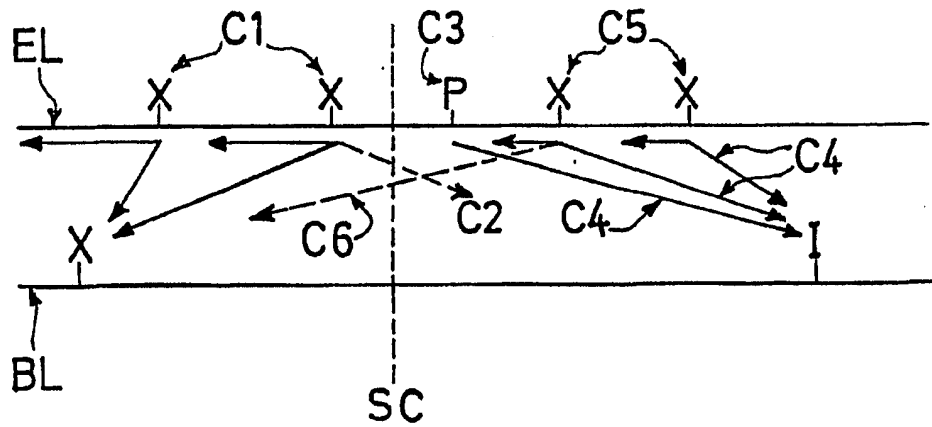


图 2