

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

H04N 7/46 (2006.01)

H04N 7/36 (2006.01)



[12] 发明专利说明书

专利号 ZL 02805555.1

[45] 授权公告日 2006年12月6日

[11] 授权公告号 CN 1288914C

[22] 申请日 2002.1.25 [21] 申请号 02805555.1

[30] 优先权

[32] 2001.1.26 [33] FR [31] 01/01174

[86] 国际申请 PCT/FR2002/000322 2002.1.25

[87] 国际公布 WO2002/060184 法 2002.8.1

[85] 进入国家阶段日期 2003.8.26

[73] 专利权人 法国电信公司

地址 法国巴黎

[72] 发明人 亨利·桑塞

纳撒雷·劳伦特 - 查特奈特

亚历山大·布森

审查员 戴惠英

[74] 专利代理机构 中国国际贸易促进委员会专利

商标事务所

代理人 郭思宇

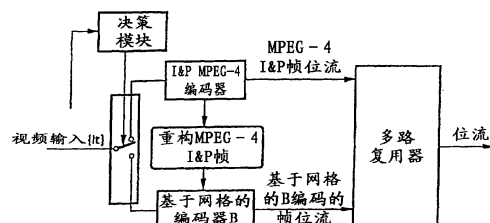
权利要求书 5 页 说明书 16 页 附图 5 页

[54] 发明名称

图象编码与解码方法以及对应的装置和应用

[57] 摘要

本发明涉及一种图象编码方法，包括：选择由包含至少一个图象的序列隔开的两个图象 IT_1 和 IT_2 ；对所述 IT_1 和 IT_2 图象进行编码，通过使用没有运动补偿的参考图象的第一图象编码限定所述序列图象，优化一图象的光度表示，所述第一编码实现对图象块的变换及各块的临时预测；通过使用带有图象内运动补偿的参考图象的第二图象编码对图象序列的每一图象进行编码，优化至少两个图象之间运动的表示，所述第二编码考虑从使用所述第一编码被编码的前面最近的一个图象计算的运动向量场、和/或从使用所述第一编码被编码的后面最近的一个图象计算的运动向量场，所述运动向量场应用于由节点形成的一网格，并被用来确定与使用所述第二编码被编码的一图象相关的推演的运动向量场，使得每一图象只根据一种编码被编码。



1. 一种图象编码方法，其特征在于以下步骤：
 - 选择由包含至少一个图象的序列隔开的两个图象 IT_1 和 IT_2 ；
 - 对所述 IT_1 和 IT_2 图象进行编码，通过使用没有运动补偿的参考图象的第一图象编码限定所述序列图象，优化一图象的光度表示，所述第一编码实现对图象块的变换及各块的临时预测；
 - 通过使用带有图象内运动补偿的参考图象的第二图象编码对图象序列的每一图象进行编码，优化至少两个图象之间运动的表示，所述第二编码考虑从使用所述第一编码被编码的前面最近的一个图象计算的运动向量场、和/或从使用所述第一编码被编码的后面最近的一个图象计算的运动向量场，所述运动向量场应用于由节点形成的一网格，并被用来确定与使用所述第二编码被编码的一图象相关的推演的运动向量场，
使得每一图象只根据一种编码被编码。
2. 如权利要求 1 中所述的编码方法，其中根据属于包括以下的组的技术至少之一，解码器知道表示所述编码之一的选择的一条信息：
 - 在编码和解码等级知道的预定选择；
 - 表示所述选择并包含在包括至少某些编码图象数据的数据流中的一条信息；
 - 表示所述选择并包含在与编码图象数据无关的数据流中的一条信息；
 - 通过解码器以固有的方式确定所述选择。
3. 如权利要求 1 和 2 任何之一所述的编码方法，其中包括选择用于所述图象的编码的步骤，其中至少：
 - 第一编码优化图象的光度表示，
 - 第二编码优化至少两个图象之间的运动表示。
4. 如权利要求 1 所述的编码方法，其中所述第二编码在 I_{11} 图

象和 I_{t_2} 后图象之间实现运动前估计；及运动补偿步骤，其中在所述运动估计期间获得的且不导致变换的运动向量按 $k=m/(t_1+t_2)$, $0 < k < 1$ 标量加权，以便内插至少一个 I_{m1} 图象， m 属于 $[t_1, t_2]$ 。

5. 如权利要求 1 所述的编码方法，其中所述第二编码在 I_{t_2} 图象和 I_{t_1} 前图象之间实现运动后估计，及运动补偿阶段，其中在所述运动估计期间获得的且不导致变换的运动向量按 $k'=1-m/(t_1+t_2)$, $0 \leq k' \leq 1$ 标量加权，以便内插至少一个 I_{m2} 图象， m 属于 $[t_1, t_2]$ 。

6. 如权利要求 4 或 5 所述的编码方法，其中该方法实现一种后估计和前估计，以便获得估计的 I_m 图象， $I_m = aE_1 + (1-a)E_2$ ，其中 $0 \leq a \leq 1$ 。

7. 如权利要求 3 所述的编码方法，其中所述选择阶段基于 N 个固定因子子采样的执行， N 中的一个图象使用所述第一编码被编码。

8. 如权利要求 7 所述的编码方法，其中 N 大于 2。

9. 如权利要求 7 或 8 所述的编码方法，其中 N 是可变的。

10. 如权利要求 1 所述的编码方法，其中所述第一编码是 MPEG-4 或 H26L 编码。

11. 如权利要求 10 所述的编码方法，其中通过所述 MPEG-4 或 H26L 编码获得的图象包括内部类型和/或可预测类型图象。

12. 如权利要求 1 所述的编码方法，其中所述第二编码基于 M 级分层网格化， M 大于或等于 1。

13. 如权利要求 12 所述的编码方法，其中所述网格是三角形。

14. 如权利要求 12 或 13 所述的编码方法，其中该方法包括隐藏区域管理阶段。

15. 如权利要求 1 所述的编码方法，其中该方法产生能够通过分开的传输通路传输的至少两个数据流。

16. 如权利要求 15 所述的编码方法，其中所述数据流属于如下的一组，包括：

- 全局标题；
- 根据所述第一编码被编码的图象数据；

- 根据所述第二编码被编码的图象数据。

17. 如权利要求 1 所述的编码方法，其中所述图象属于包括以下各项的组中的至少一个信号：

- 数字 TV 信号；
- 通过 IP 网络实时传播的视频信号；
- 通过网络到可移动站的实时传播的视频信号；
- 将被储存在一个数据存储器上的图象的数据信号。

18. 一种用于对图象信号进行解码的方法，所述图象信号包括一个标题和一个位流，

所述位流包括一个根据第一图象编码被编码的帧序列和根据第二图象编码被编码的帧组，

第一图象编码使用没有运动补偿的参考图象，优化一图象的光度表示，所述第一编码实现对图象块的变换及各块的临时预测；

第二图象编码使用带有图象内运动补偿的参考图象，优化至少两个图象之间运动的表示，所述第二编码考虑从使用所述第一编码被编码的前面最近的一个图象计算的运动向量场、和/或从使用所述第一编码被编码的后面最近的一个图象计算的运动向量场，所述运动向量场应用于由节点形成的一网格，并被用来确定与使用所述第二编码被编码的一图象相关的推演的运动向量场，

其中所述用于解码的方法包括：

读取标题，允许混合解码被激活的步骤；

对于对应于一个独立实体的位流的每一部分，识别被使用的编码类型的步骤；

根据识别的编码类型，用于对独立实体进行解码的步骤。

19. 一种用于对图象信号进行编码的设备，包括以下装置：

- 用于选择由包含至少一个图象的序列隔开的两个图象 IT_1 和 IT_2 的装置；

- 用于对所述 IT_1 和 IT_2 图象进行编码、通过使用没有运动补偿的参考图象的第一图象编码限定所述序列图象、优化一图象的光度表

示的第一编码装置，所述第一编码实现对图象块的变换及各块的临时预测；

- 用于通过使用带有图象内运动补偿的参考图象的第二图象编码对图象序列的每一图象进行编码、优化至少两个图象之间运动的表示的第二编码装置，所述第二编码考虑从使用所述第一编码被编码的前面最近的一个图象计算的运动向量场、和/或从使用所述第一编码被编码的后面最近的一个图象计算的运动向量场，所述运动向量场应用于由节点形成的一网格，并被用来确定与使用所述第二编码被编码的一图象相关的推演的运动向量场，

使得每一图象只根据一种编码被编码。

20. 一种用于对图象信号进行解码的设备，所述图象信号包括一个标题和一个位流，

所述位流包括一个根据第一图象编码被编码的帧序列和根据第二图象编码被编码的帧组，

第一图象编码使用没有运动补偿的参考图象，优化一图象的光度表示，所述第一编码实现对图象块的变换及各块的临时预测；

第二图象编码使用带有图象内运动补偿的参考图象，优化至少两个图象之间运动的表示，所述第二编码考虑从使用所述第一编码被编码的前面最近的一个图象计算的运动向量场、和/或从使用所述第一编码被编码的后面最近的一个图象计算的运动向量场，所述运动向量场应用于由节点形成的一网格，并被用来确定与使用所述第二编码被编码的一图象相关的推演的运动向量场，

其中所述用于解码的设备包括：

读取标题，允许混合解码被激活的装置；

对于对应于一个独立实体的位流的每一部分，识别被使用的编码类型的装置；

根据识别的编码类型，用于对独立实体进行解码的装置。

21. 如权利要求 20 所述解码装置，其中包括，类似于编码期间所实现的，用于确定至少一部分向量场和/或至少一部分隐藏区域的

装置。

22. 一种用于对图象信号进行编码、传输和/或解码的系统，包括如权利要求 19 所述的用于对图象信号进行编码的设备和/或如权利要求 20 或 21 所述的用于对图象信号进行解码的设备。

23. 如权利要求 22 中所述的系统，其中根据属于包括以下的组的技术至少之一，解码器知道关于所述编码之一的选择的一条信息：

- 在编码和解码级知道的预定选择；
- 表示该选择并包含在包括至少某些编码图象数据的一个数据流中的一条信息；
- 表示该选择并包含在与编码图象数据无关的一个数据流中的一条信息；
- 通过解码器以固有的方式确定所述选择。

图象编码与解码方法以及对应的装置和应用

技术领域

本发明的技术领域是通过降低动画图象序列特别是视频的位速率进行编码。

大量的视频编码应用。这些包括(列表并未穷尽):

- 数字 TV 传输
- 几种类型的网络上的实时视频传输: IP, 移动, (“流动 IP”)
- 视频的计算机存储

本发明可特别用于实现 MPEG 型编码的系统框架中。MPEG-型编码意思是基于临时预测和离散余弦变换的编码, 基于硬块结构, 常常有固定尺寸, 但可以有可变尺寸。对这一编码族的两种表示标准是版本 1 到 4 的 MPEG-4 标准和直到版本 2 的 ITU-T/H.263。本发明还可用于 CCITT H26L 建议框架内(例如, 参见对应的 VCEG-N83d1 文献)。

当前提供的视频编码和解码方案划分为两类:

- 通过 ISO/MPEG, 或通过 ITU-T 标准化的编码, 都基于相同类型的技术(基于块结构的临时预测及离散余弦变换)
- 在研发中由研究实验室提供的编码, 这涉及大量技术: 小波编码, 区域性编码, 分形编码, 网格编码等等。

当前, MPEG-4 编码被认为是商业标准, 不仅是标准化的编码中, 而且是公开的编码。

MPEG-4 或 ITU-T/H.263++型编码认为是达到其极限, 特别是由于用作为对所有编码计算和操作的支持的固定尺寸的硬块结构。特别地, 序列内图象的临时预测不能被充分利用。

此外, 公开的另一种编码还没有达到充分的优化程度。

因而，为了允许被编码的视频序列以低位速率获得，编码器通常要降低图象的尺寸，并生成原始视频序列的临时子采样。然而，第二种技术的不便在于，取决于子采样水平要恢复对于用户或多或少成问题的不确定的运动。

为了避免这些不确定的运动，必须在解码器中使用临时内插重新产生失去的图象(未被编码的)。

然而，当前用于图象临时内插的技术不能获得满意的结果，特别是当它们以单解码器实现时。实际上，这些技术是与基于块的运动补偿技术相关的可视人工产物的来源，这些技术只对块的所有象素定义单个的运动向量。

发明内容

本发明的目的具体是补救以往技术的限制。

更准确地说，本发明的一个目的是要提供一种对图象数据编码和解码的技术，允许相对于已知的技术获得降低的位速率和/或改进的重新产生的图象质量。

根据本发明，这一目的是借助于一种图象编码过程实现，该过程有选择地实现至少两种图象编码方法，每一方法根据各种优化准则优化至少一个视频序列图象的压缩。

根据几个有用的实施例，关于选择所述解码器编码方式之一的信息，解码器可根据属于包括以下诸技术的组的至少一种技术知道：

- 关于编码和解码已知的预定选择；
- 包含在由至少某些编码图象数据组成的数据流中的选择的表示信息；
- 包含在与编码图象数据无关的数据流中的选择的表示信息；
- 以预定的方式通过解码器选择的确定。

该过程最好涉及用于选择施加到所述图象的编码方式的一个阶段，这至少包括：

- 优化图象光度表示的一种初始编码。

- 优化至少两个图象之间运动表示的第二编码。

这样，本发明通过编码混成涉及一种新的视频编码过程，特别是 MPEG-型编码和使用基于由网格表示的临时内插的编码，以及解码过程及相关的二进制表示结构。

所述第二编码，最好考虑利用所述初始编码被编码的至少一个先前的图象和/或至少一个后继的图象。

所述第二编码宜考虑从使用所述初始编码被编码的紧挨的前一图象计算的运动向量场，和/或从使用所述初始编码被编码的后继紧跟的图象计算的运动向量场。

这些向量场宜应用于网格。

这种情形下，所述运动向量场可用来确定与使用所述第二编码被编码的图象相关的推演的运动向量场。

根据本发明一优选实施例，所述选择阶段基于一个 N 固定因子子样本的实现， N 上的图象使用所述初始编码被编码。

这一 N 值与至少一个预定准则相关，并最好为一变量。

根据一特定的实施例，所述初始编码实现对图象块的一种变换，及各块的一种预测。例如这一变换为 DCT-型变换，Hadamard 变换，子波变换等。

将观察到，图象块不必是方形的，而是可取适于需要的任何形式和可用的方法。

这一初始编码具体可以是 MPEG-4 或 H26L 编码。

在后者的情形下，最好使用类型 I(内)图象和/或类型 P(可预测)图象(最好不是类型 B 图象)。

根据本发明的另一具体方式，所述第二编码宜基于 M -级分层网格来实现，其中 M 大于或等于 1，并例如为三角形网格。

这种情形下，该过程最好包含隐藏区域或消隐区域的管理阶段。

所产生的数据可被组成单个的流。最好预期有至少两个能够在独立传输通路上传输的数据流。

所述数据流最好属于包括以下内容的组:

- 全局标题;
- 根据所述初始编码被编码的图象数据;
- 根据所述第二编码被编码的图象数据。

因而流传输可独立进行。特别地, 这允许根据使得和需要图象的渐进的和/或局部的解码。

因而, 根据本发明特定的实施例, 开发了以下方面:

- 通过标准化 MPEG 或 ITU-T/H.263-型编码的构成模块推进的优化

- 基于网格技术的临时预测和相关的误差编码。

确实, 基于网格的方法由于属于连续的运动场而有助于避免通常的阻塞。而且, 网格技术允许消隐被检测的“对象”, 以及正确适用于这些区域的误差编码。此外, 通过围绕这些区域组合 MPEG-型误差编码, 以大大低于通过 MPEG 提供的双向图象(类型 B 图象)的成本可明显改进内插的效率。

因而, 使用 MPEG-型编码能够以低的临时分辨率对基本的信息有效编码, 质量良好, 并然后使用通过网格内插方式的编码能够恢复序列的全部流。

当然, 本发明还涉及:

- 对使用上述编码过程被编码的图象信号解码的过程;
- 对使用上述编码过程被编码的图象信号编码的过程;
- 用于对使用上述编码过程被编码的图象信号解码的装置(最好包括用于确定至少部分的向量场和/或至少部分的消隐区的装置, 类似于编码期间实现的那些装置);

- 用于存储使用上述编码过程被编码的至少一个图象信号的装置

- 使用上述编码过程被编码的图象信号的编码, 传输和/或解码系统(根据属于包括以下技术的组的技术至少之一, 所述编码方式之一的选择最好对于解码器是已知的:

- 在编码和解码期间已知的预定选择;
- 包含在至少包括某些编码图象数据数据流中的选择的表示信息;
- 包含在数据流中的选择的表示信息, 与被编码的图象数据无关;
- 通过解码器以固有的方式选择的确定;
- 计算机程序产品, 用于对使用编码过程被编码的图象信号进行编码和/或解码;
- 对这种程序的数据支持。

本发明还涉及包含根据上述过程编码的数据的图象数据信号。

这一信号最好包括至少一个指示符, 规定该过程是否已被激活。

在视频序列和/或在每一信号帧的开头, 该信号最好包括规定帧结构的数据。

使用所述第二编码被编码的序列, 最好以指示根据这一第二编码被编码的帧的数目的标题开始。

根据一特定的实施例, 信号包括指示两个可在分开的传输通路传输的数据流。

这种情形下, 所述数据流最好属于包括以下内容的组:

- 全局标题
- 根据所述初始编码被编码的图象数据;
- 根据所述第二编码被编码的图象数据。

本发明发现在各种领域的应用, 特别是在属于包括以下内容的组的领域:

- 数字化 TV
- IP 网络上的实时视频;
- 通过网络到移动站的实时视频;
- 图象数据存储。

附图说明

在阅读作为简单示例性、非限制性例子，及附加的附图而提供的本发明的优选实施例的说明，本发明其它的特性和优点将是明显的，其中：

- 图 1 是本发明编码的示意图；
- 图 2 示出对于运动的分层网格结构的一例；
- 图 3 示出关于三角形网格的仿射内插 (interpolation affine) 原理；

- 图 4 是通过覆盖三角形检测的消隐一例；
- 图 5 示出对称方阵中任何图象三角形的变换过程；
- 图 6 示出任何三角形到直角等腰三角形的变换；
- 图 7 通过相关四元树示出分层网格和表示；
- 图 8 是对分层网格编码决策的一例；
- 图 9 示出根据本发明的位流全局结构；
- 图 10 示出根据本发明的解码器的块程序。

具体实施方式

以下所述本发明的实施例基本上由 MPEG-型编码的混合组成，例如 MPEG-4，使用以内插方式操作的网格编码，也称为 B 方式或 MPEG-标准的 B 图象。

应当注意，这里所述的 MPEG-4 编码可由任何基于等价技术的编码器代替，就是说使用基于块结构、临时预测和离散余弦变换及对于产生的信息量化和熵编码。特别地，ITU-T/H.263++编码能够代替 MPEG-4 编码。

对于进入编码器的序列的每一图象，编码器根据特定决策过程 (例如固定因子临时子采样) 决定使用具有 MPEG-4 编码模块的编码器，还是使用基于网格的编码模块编码器。

以网格方式编码的图象，作为它们的临时预测参照，使用按 MPEG-4 方式编码的图象，这些图象位于紧靠按它们所属的网格网格

方式编码的图象之前或之后。

本发明的压缩效率关键点在于，基于网格的运动补偿导致非常有力的临时预测，其相关编码成本非常低。

确实，这种技术：

- 考虑图象内各种类型的运动
- 特别处理由于对象的运动所至区域的覆盖和去覆盖

图 1 提供了编码器的原理总体图。

首先，根据给定的决策方法，例如根据预定的节奏： N 中 1 个图象按 MPEG 编码，其余的按网格内插方式编码，输入的图象或被导向 MPEG 编码模块或基于网格的编码模块。应注意， N_k 表示按 MPEG 方式编码的图象数。

所有其余 I_l , $N_k < l < N_{k+1}$ 图象通过以称为 B 方式的内插方式工作的网格编码器编码，例如三角形。这一编码器一般原理如下：

1. 在 N_k 和 N_{k+1} 图象之间的之前和之后运动场的计算。对这些场以三角形网格形式建模。

2a. 可预测区域的估计：在内插期间，估计 I_t 和 $I_{t'}$ 之间的运动。如果运动向量的应用不导致一种变换(表示存在于可预测区域中)，则通过使用 $k(0 \leq k \leq 1)$ 标量对向量加权进行运动补偿，以便内插 I_{t+k} , $t+k$ 属于 $[t, t']$ 。对运动补偿有如下所述的三种可能的方法。

2b. 不可预测区域的估计：在 l 个待编码的图象中，从知道这些运动场检测不可预测消隐区域。

3. 根据以下三种可能的方式之一这些消隐区域的特定编码：

- 使用没有运动补偿的参照图象(N_k , N_{k+1} ，或在以它们的运动场运动期间被补偿的这些图象)之一预测，然后使用基于三角形网格的技术对预测误差编码；

- 使用带有图象内运动补偿的参照图象(N_k , N_{k+1} ，或在以它们的运动场运动期间被补偿的这些图象)之一预测，然后使用基于三角形网格的技术对预测误差编码；

- 基于三角形网格技术的图象内编码。

4. 可选地, 残余预测误差或编码的方式 P MPEG-型编码, 限于围绕消隐区域的一区域。

如上所述, 能够使用三种方法进行运动补偿: 之前估计, 之后估计, 或之前与之后估计。

1) 之前估计:

在内插期间, 估计 I_{t1} 和 I_{t2} 之间的运动。如果该运动向量的施加不导致变换(这对应于可预测的区域), 则通过以标量 $k=m/(t1+t2)(0<k<1)$ 对向量加权进行运动补偿, 以便内插 I_m , m 属于 $[t1, t2]$ 。获得 E1 估计的图象。

2) 之后估计:

在内插期间, 估计 I_{t2} 和 I_{t1} 之间的运动。如果该运动向量的施加不导致变换(这对应于可预测的区域), 则通过以标量 $k'=1-m/(t1+t2)(0\leq k'\leq 1)$ 对向量加权进行运动补偿, 以便内插 I_m , m 属于 $[t1, t2]$ 。获得 E2 估计的图象。

3) 两者一同:

在内插期间, 估计 I_{t1} 和 I_{t2} 以及 I_{t2} 和 I_{t1} 之间的运动。如果该运动向量的施加不导致变换(这对应于可预测的区域), 则通过以标量 $k(0\leq k\leq 1)$ 对“之前”的向量加权进行运动补偿, 且获得 E1。对“之后”的向量使用标量 $k'(0\leq k'\leq 1)$ 同样进行补偿, 并获得 E2。然后估计的图象为 $E=aE1+(1-a)E2$, 其中 $0\leq a\leq 1$ 。

每 2 位, 在计算估计的解和相关源图象之间的 PSNR 之后取最佳解。

1. 计算 N_k 与 N_{k+1} 图象之间的之前和之后的运动场

以分层网格形式, 例如三角形, 如图 2 中所示的 T_k^b 与 T_{k+1}^b , 计算 N_k 与 N_{k+1} 图象之间的之前和之后的运动场。

根据运动估计过程期间特定的准则, 通过划分一定的网格, 例如三角形网格被划分为 4 个子三角形, 获得这些网格。在每一分层级, 对每一网格取或不取划分决策。一旦这些划分已经决定, 已划分的网格相邻的网格被划分, 以便维护一致的网格结构。划分之前的初

始网格(分层的顶部)可取任何形式。

在图 2 的例子中, 运动估计器决定划分三角形 3 和 8。这导致三角形 2, 4, 7 和 9 的划分。重复该过程直到预定的分层级。

在三角形网格的情形下, 由三角形网格 T 定义的运动场表示由以下对每一三角形 e 给出:

$$\bar{d}(\bar{p}, D) = \sum_{n \in \text{ver}(e)} \Psi_n^e(\bar{p}) \bar{d}_n, \quad \forall \bar{p}(x, y) \in e$$

其中

- e 指示具有 x 和 y 坐标的当前点 \bar{p} 的三角形元素 T
- $\{\text{ver}(e)\}$ 指示位置 $\bar{p}_i, \bar{p}_j, \bar{p}_k$ 的标号为 i, j, k 的所有三个结点或峰。
- $\Psi_l(l = i, j, k)$ 表示 $e_{i,j,k}$ 三角形元素中点 $\bar{p}(x, y)$ 的重心坐标, 其中

$$\begin{cases} \Psi_l^e(x, y) = \alpha_l^e + \beta_l^e x + \gamma_l^e y, \alpha_l^e, \beta_l^e, \gamma_l^e \in \mathfrak{R} & \text{若 } \bar{p}(x, y) \in e_{i,j,k} \\ \sum_{l=i,j,k} \Psi_l^e(x, y) = 1 \\ \Psi_l^e(x, y) = 0 & \text{否则} \end{cases}$$

这种类型模块定义了一种全连续的场。进而, 这允许表示精确性的充分控制, 这是对于压缩的重要特征。

在网格分层的每一等级, 计算结点的运动向量以便使预测误差最小化。可使用各种基于网格的运动估计器, 例如专利 FR No. 98 11227, 或 FR No. 99 15568 中所述的那些估计器。

重要的是, 最终的网格是来自初始网格通过划分的分层过程的结果。这一分层性的使用确实有利于结点与其父结点之间(其被插入的弧的边缘)结点运动向量的差分编码。在解码器中根据初始网格, 以及网格划分指示器重新计算网格结构。

因而, 在过程的结尾, 对于由 N_k 与 N_{k+1} 图象组成的图象组的每一组获得 2 个运动网格, 用来重新产生组的所有图象。

2. 消隐区域的检测

从这 2 个网格，检测消隐区域，这是从覆盖或去覆盖对象的 N_{k+1} 图象获得的 N_k 图象中不可预测区域或反之。

在它们由它们的结点向量移动之后，这些区域简单地三角形的覆盖定义。

附图示出基于运动后三角形去覆盖的消隐的检测。

通过对消隐区域中的三角形去激活，编码器能够继续进行运动估计，以便获得偏移较小的位移向量。

然而，严格来说这是编码器策略的一内部特征，并在最后，是两个完全的运动网格 T_k^b 和 T_{k+1}^f 被编码并插入到位流。然后解码器根据这 2 个网格能够找出消隐区域。

这些消隐区域在 N_k 与 N_{k+1} 图象中定义，并一旦它们被检测到，它们对应的三角形在编码器和解码器中都被标为结果。

然而，编码器必须知道在 N_{k+1} 与 N_{k+1-1} 图象中哪里找到这些区域。这些图象通过在由施加结点运动向量被编码的图象上投影 T_k^b 和 T_{k+1}^f 网格而简单获得，这些向量被重新标称化，以考虑当前图象与参照图象 N_k 或 N_{k+1} 之间的临时距离。

3. 消隐区域编码

对于每一消隐区域，用于可能预测的参照图象在 I_{N_k} 、 $I_{N(k+1)}$ 以及在 1 处通过以 T_k^b 和 T_{k+1}^f 网格在还没有网格覆盖之处的运动补偿而获得的图象 I_1^c 之中选择。

更精确来说， I_{N_k} 和 $I_{N(k+1)}$ 之间的选择简单地取决于曾产生当前消隐区域的 T_k^b 和 T_{k+1}^f 网格。然后，这一网格用来预测图象并提供 I_1^c 。 I_1^c 与 I_{N_k} 或 $I_{N(k+1)}$ 之间的选择基于预测误差准则：要保持引起最弱误差的图象。因而，这只不过是在二进制流中每区域插入 1 位，以便对所保持的预测的选择编码。

不取选择的基准图象 I_r 。

这些区域的编码的其余部分包括两个阶段：

- 预测阶段，
- 预测误差或在不正确的预测情形下原始纹理的编码阶段。

3.1 消隐区域纹理残余预测

可互相排他地使用三种方法。其决策基于最小误差准则。

方式 1:

区域中的象素的值 Y , U 和 V 就是位于同一参照图象 I_r 中的象素的值。就是说, \hat{I}_1 结果图象。然后对 \hat{I}_1 与 I_1 之间的预测误差编码。

方式 2:

因而运动估计在消隐区域上 I_1 (要被编码的图象)与 \hat{I}_1 (来自方式 1 预测的结果)之间进行。因而, 在网格覆盖之前从上一网格级 T_r , $r=k$ 或 $k+1$, 所得的网格, 作为其结点运动被编码。最后, 根据稍后定义的过程编码预测残余误差。

方式 3:

不进行预测, 并对区域的象素原始值编码。

4. 对消隐区域的纹理或预测误差编码。

原始纹理和预测误差线得到相同的编码, 其原理如下:

假设初始三角形网格能够从对于被编码的区域预测保持的 T_r 网格运动定义。获得这一原始网格的途径将在稍后说明。

因而纹理是根据一种选择在每一网格逼近的:

- 带有高频的网格是基于称为 DCT 的离散余弦变换被编码的。
- 使用精细的最终单元模型对较平滑的网格编码。

又分层方法将用来降低通过网格对表示编码的成本。

所使用的方法允许与网格规则分层结构相关的编码低成本维护, 同时允许网格的非规则分解图象内容的局部适配。

从区域的粗初始网格, 网格被再划分为 4 个三角形子网格, 直到给定的等级。

在最后一个等级, 如果由 2 个相邻的三角形产生的四边形对角线的可选的改变可以降低近似误差, 则实施该改变。

4.1 消隐区域上的纹理网格的初始化

这一网格在所考虑的区域中变换出现之前简单地由最后的 T_r 等级给出(根据保持的方向从 T_k 或 T_{k+1} 所得的网格)。因而, 因为是从

运动网格抽取的，故获得自动插入到运动网格的纹理网格。

4.2 用于在三角形上纹理的表示

两个表示被组合：仿射内插与三角形 4 DCT

仿射内插

三角形网格结点包含光度(颜色, 误差)信息, 并通过拉格朗日修正元素进行对三角形内的点的内插。也称为仿射内插。

由三个结点 $\vec{p}_l, l=i, j, k$ 定义的三角形 $e_{i,j,k}$ 的点 $\vec{p}(x, y)$ 的 $v(\vec{p})$ 值由以下方程式提供:

$$v(\vec{p}) = \sum_{l=i,j,k} \Psi_l(x,y)(\vec{p}_l)$$

$\Psi_l(l=i, j, k)$ 表示该点的重心坐标。

$V(\vec{p})$ 可以是点的 Y, U 或 V 光度分量之一, 或这些分量的预测误差。

可使用几种方法计算结点值, 特别是最小二乘法。

三角形上的离散余弦变换

该方法的原理在于将任何三角形变换为直角等腰三角形基准。该直角等腰三角形对于斜边是对称的, 以便提供对称的方阵(图 4)。

因而经典的 DCT(正方形)施加到这一矩阵。可能证明, 被变换的矩阵也是对称的。因而只有其下三角形的系数被量化然后被统计编码(熵编码)。

图 4 描绘了该过程各阶段: T 三角形的选择, T 三角形仿射变换为 T' 直角等腰三角形。通过仿射变换, 三角形象素不再位于规则正交图, 且建议对参考三角形内的光度值重新采样。为此, 使用类似于图象运动补偿过程的过程(在仿射变换期间), 例如使用双线性内插器。

仿射变换 F 及其逆 F^{-1} 由以下方程式定义:

$$F: \begin{cases} X = \frac{(x_2 - x_1)(y_1 - y) + (y_1 - y_2)(x_1 - x)}{(x_3 - x_1) + (y_2 - y_1) + (y_1 - y_3)(x_2 - x_1)} N \\ Y = \frac{(x_1 - x_3) + (y_1 - y) + (y_3 - y_1)(x_1 - x)}{(x_3 - x_1)(y_2 - y_1) + (y_1 - y_3)(x_2 - x_1)} N \end{cases}$$

$$F^{-1}: \begin{cases} x = x_1 + (x_3 - x_1) \frac{X}{N} + (x_2 - x_1) \frac{Y}{N} \\ y = y_1 + (y_3 - y_1) \frac{X}{N} + (y_2 - y_1) \frac{Y}{N} \end{cases}$$

通过 F^{-1} 逆变换获得 T' 三角形的光度(因而 T' 的呈对称的 M 块的)值 $M(i, j)$, 以及然后要被编码的图象纹理的 Γ 内插:

$$M(i_x, j_x) = I_{\Gamma}(P_{(x,y)}) = I_{\Gamma}(F^{-1}(Q_{(x,y)}))$$

其中 $M(i_x, j_y)$ 系数是在 $Q(X, Y)$ 点的值, 其中变换的 $P(x, y)$ 为 $F^{-1}(Q)$

I_{Γ} 指示用来计算在该点的图象值的内插器, 潜在的非整数坐标。

Γ 纹理的 $\hat{\Gamma}$ 再生由以下给出:

$$\hat{\Gamma}(P_{(x,y)}) = I_{\Gamma}(Q_{(x,y)}) = I_{\Gamma}(F(P_{(x,y)}))$$

其中 I_{Γ} 指示来自 M' 块值的内插纹理, M 的量化版。

这一技术只能用于非零区域三角形。但按定义, 这类三角形不需要纹理编码。

与 SADCT 相反(DCT 适用形成), 即使没有量化, 这一变换也不能保证在逆变换后最好的再生。

为了减少再生的误差, 引入一标度因子 α 以对于三角形 i 计算 M_i 块(尺寸 $N_i \times N_i$):

$$N_i = E(\alpha \sqrt{2A_i})$$

其中

- E 是过量完全部分,
- A_i 是三角形 i 区域。

实际上, $\alpha = 1$ 提供有益的折衷, 其对于接近等腰三角形的三角形更有效。在 $\alpha < 1$ 的情形与量化步骤结合使用, 以便压缩信息量。

一旦 M_i 块对于每一三角形定义，一个经典的 DCT 变换施加到它们，并根据几个可能的过程例如一致标量量化，或甚至带有变换系数频率的增量量化，来量化这些变换系数。也可使用著名的 MPEG 或 JPEG 量化矩阵。

DCT 表达式由以下给出：

$$F(u,v) = \frac{1}{\sqrt{NM}} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \lambda(u)\lambda(v) \cos\left(\frac{(2i+1)\pi}{2N}\right) \cos\left(\frac{(2j+1)\pi}{2M}\right) f(i,j)$$

$$\lambda(\xi) = \begin{cases} \frac{1}{\sqrt{2}} & \text{若 } \xi = 0 \\ 1 & \text{否则} \end{cases}$$

$F(u, v) = F(v, u)$ 关系出现，因为：

$$f(i, j) = f(j, i) \quad (\forall u, v, i, j=0, \dots, N-1) \text{ 被定义。}$$

结果，足以只计算变换矩阵的下部分的系数。

4.3 纹理的整体编码

如上指出，使用通过划分给定等级分层的每一三角形为 4 个子三角形获得的一致分层网格，并在弧的中点插入结点。重复该过程直到达到最大等级。三角形的这一分层结构还以四元树的形式由编码器表示和管理(图 5)。注意，只考虑包含在编码区域中的三角形。基本初始网格的构成过程保证了网格分层结构的所有三角形属于编码区域。

使消隐区域网格化的编码过程可概括如下：

1. 通过生成规则初始网格化，交错分层网格化在要编码的区域上被定义，并然后重复划分三角形为 4 个子三角形，并在弧的中点插入新的结点。计算结点值以便使由于网格化的逼近误差最小化。

2. 对包含使用结点值的象素的三角形使用仿射内插，移动象素值彼此接近。

对于分层结构中的每一三角形，计算逼近误差 E ，并然后基于 2 个阈值 σ_1 和 σ_2 决定各种表示和编码方法：

1. 如果 $E < \sigma_1$ ，则对三角形的仿射内插即足够。
2. 如果 $\sigma_1 < E < \sigma_2$ ，则必须使用三角形最精细的分解以获得良好的逼近，总是使用仿射内插。

3. 如果 $E > \sigma_2$, 则三角形被纹理化, 且使用 DCT 对仿射内插误差编码。

最后, 对最细的网格化, 测试通过 2 个相邻的三角形形成的四边形对角线的置换产生的误差的降低。如果获得正的结果, 这一置换得到验证。

根据对不同三角形选择的编码方法, 按以下方式对各种数据编码。

首先使用父结点(其中当前结点被插入的弧的边缘)值预测 YUV 结点值。因而结点值与预测值之间的差被量化。

最后使用一种算法编码器, 对四元树结构(包括三角形的划分和非划分指示器), 对角线置换指示器, YUV 的差分结点值及量化 DCT 系数编码, 并插入到位流。

5. 通过网格化编码的帧的二进制流中编码信息的概述

N_{k+1} 与 $N_{k+1}-1$ 之间按网格化方式编码的每一组帧(其中 N_{k+1} 与 $N_{k+1}-1$ 分别是以 MPEG 方式编码的先前和后继的帧)在二进制流中整体表示。

在编码的形式中被驱动的数据包括:

帧组集合的标题, 其中包括被编码的帧的实际数目。

运动的网格(结点的结构和运动向量) T_k^b 和 T_{k+1}^f 。

对于每一组图象的预测或原始误差纹理。

6. 位流的整体结构

如图 8 所指出, 整体位流包括以 MPEG 方式编码的帧序列, 及按网格化编码的帧组。

表示编码序列的位流的整体标题其中包含混合编码的指示。

对应于以网格化方式编码的帧组的位流的部分, 以指示其中实际被编码的帧数的标题开始, 并可以是零。

分别对应于编码序列的整体标题的各种数据流(位流), 以 MPEG 编码的图象及以内插网格化方式编码的 i 个图象的组, 如果需要可通过不同的独立的通路发送。特别地, 该编码过程允许该序列分

层结构的(或可标量化的)解码,即只使用总位速率一部分的解码。

7. 解码过程

图9提供了解码原理的示意图。

首先,标题解码允许激活混合解码。

然后,如果是按 MPEG-4 编码的帧或按网格化编码的帧组,则解码器识别对应于独立实体的位流的每一部分。

MPEG-4 帧提供给 MPEG-解码模块。而以网格化方式编码的帧组提供给通过网格化解码的模块。

7.1 基于网格化的解码

首先对于 I_l , $N_{k+1} < l < N_{k+1}$ 图象组的 T_k^b 和 $T_{(k+1)}^f$ 运动网格化解码。

然后根据编码器中相同的过程找出这些图象的消隐区域。

然后,使用 I_{Nk} 和 I_{Nk+1} 图象及 T_k^b 和 T_{k+1}^f 运动场简单地内插消隐区域之外的象素。

根据与编码器过程相同的过程对每一消隐区域,找出最粗的纹理的网格化(分层结构的顶层)。

因而对与对应的分层网格(三角形划分因子,仿射内插决策或 DCT 编码,差分结点 YUV 值和量化的 DCT 系数)相关的信息解码,并再生这些区域的象素 YUV 值。

图1

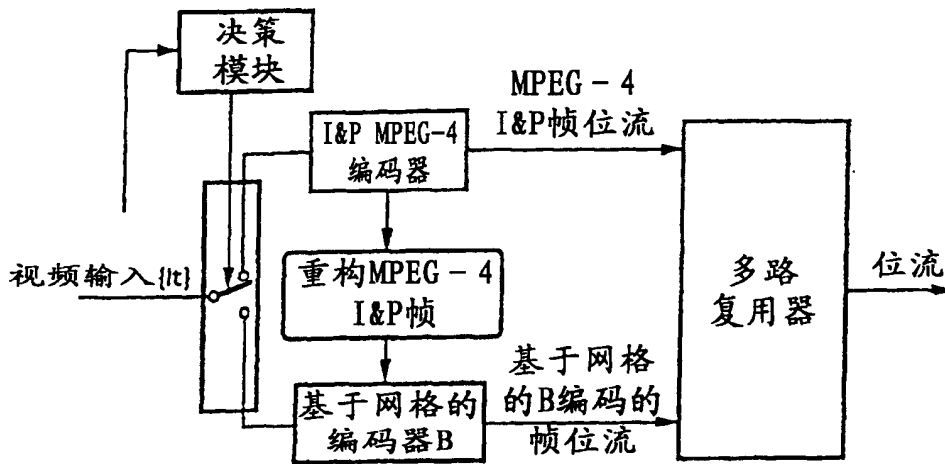


图2

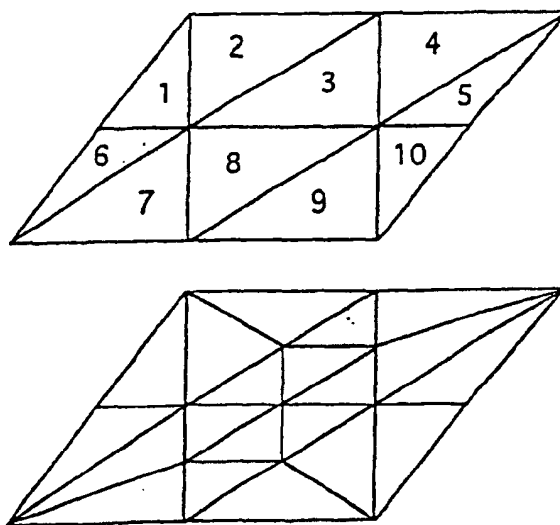


图 3

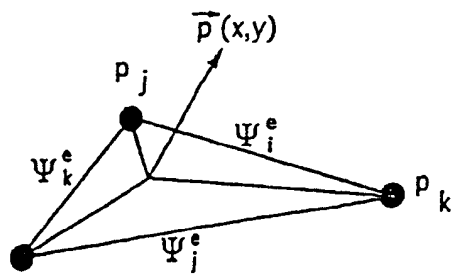


图 4

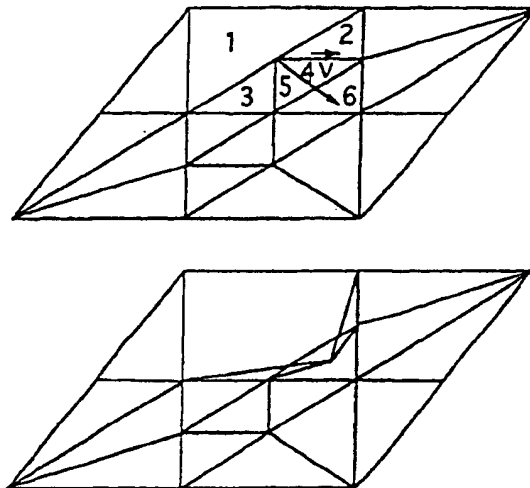


图5

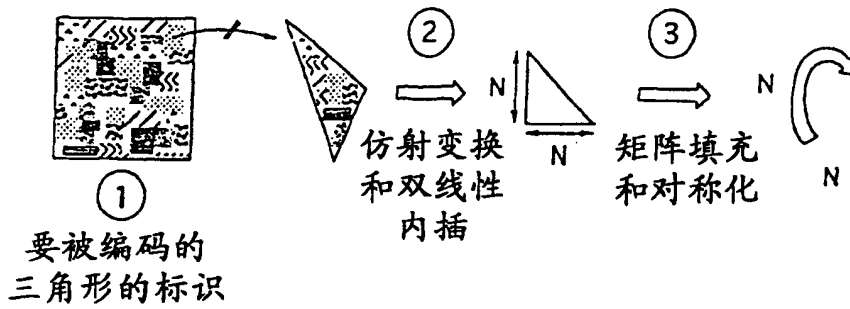


图6

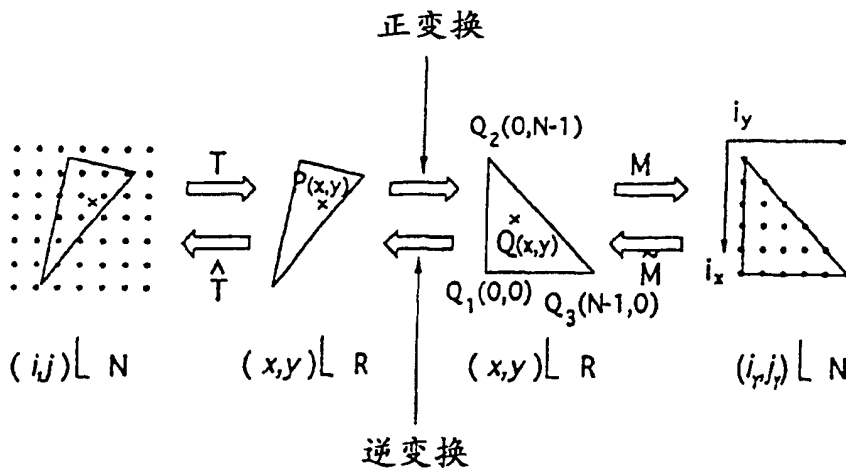


图7

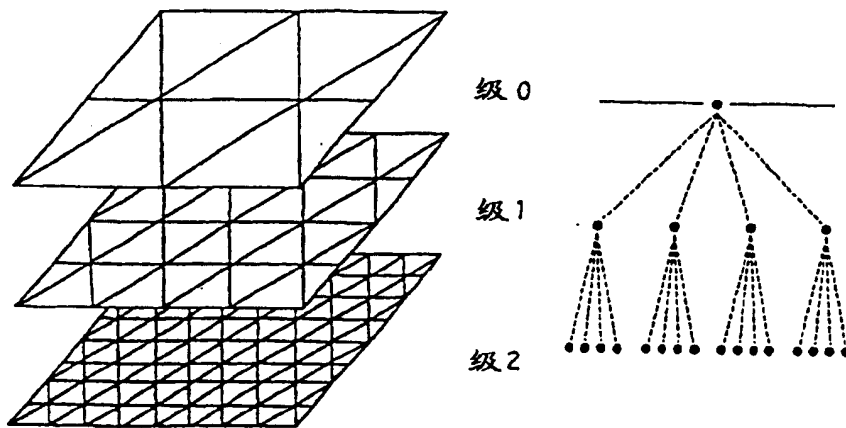


图8

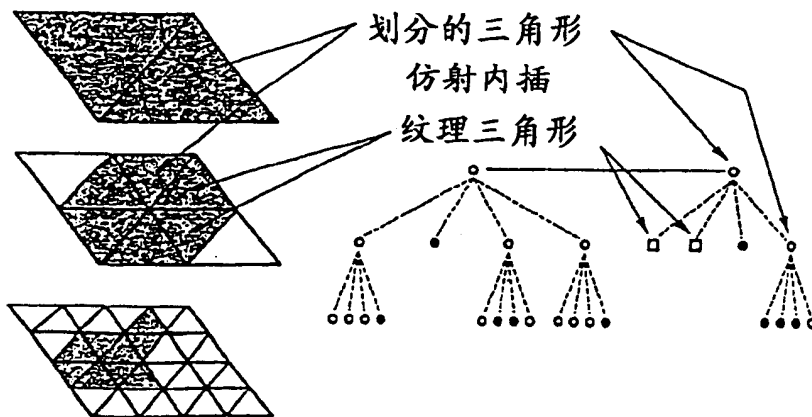


图9

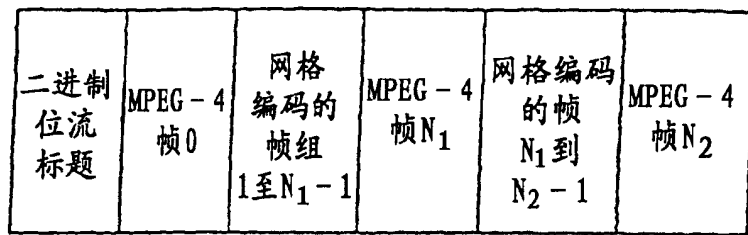


图10

