

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
H04N 7/12 (2006.01)



[12] 发明专利申请公开说明书

[21] 申请号 200480006390.8

[43] 公开日 2006年4月12日

[11] 公开号 CN 1759610A

[22] 申请日 2004.1.9

[21] 申请号 200480006390.8

[30] 优先权

[32] 2003.1.9 [33] US [31] 60/438,894

[32] 2003.11.5 [33] US [31] 60/517,467

[86] 国际申请 PCT/US2004/000446 2004.1.9

[87] 国际公布 WO2004/064373 英 2004.7.29

[85] 进入国家阶段日期 2005.9.9

[71] 申请人 加利福尼亚大学董事会

地址 美国加利福尼亚州

[72] 发明人 帕梅拉·科斯曼

阿萨纳西奥斯·利昂塔里斯

维杰伊·切拉帕

[74] 专利代理机构 北京市柳沈律师事务所

代理人 黄小临 王志森

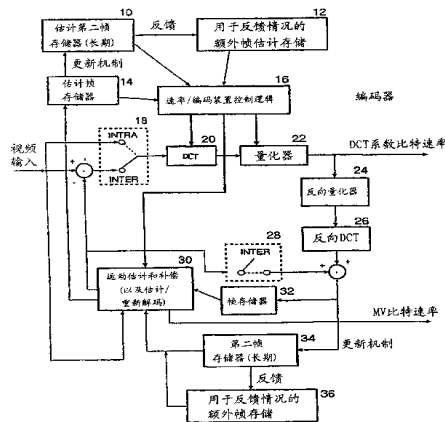
权利要求书4页 说明书19页 附图6页

[54] 发明名称

视频编码方法及设备

[57] 摘要

本发明使用了一种双帧以及可能多帧的方法(图1A)。本发明的实施例包括做出使用长期参考帧进行 INTER 编码还是进行 INTRA 帧编码的决定(18),该长期参考帧是不同于紧邻过去参考帧的帧。其它的实施例包括使用长期和短期参考块(34),以及在两种类型的 INTER 编码块和 INTRA 编码之间做出决定(18)。根据本发明的实施例,长期帧是高质量帧。高质量帧在特定条件下可以用作参考帧。



1. 一种视频编码器，包括：
编码装置，用于对向量编码以相对于至少一个参考块描述至少一个图像
5 块；
短期参考块缓冲器，其存储至少一个短期参考块；和
至少一个长期参考块缓冲器，其存储至少一个长期参考块；
其中所述编码装置基于在编码的时候检查的一个或多个因素来选择性地在相对于所述短期参考块缓冲器中的至少一个短期参考块的编码与相对于所述长期参考缓冲器中的至少一个长期参考块的编码之间进行选择，以最大化
10 压缩、视频质量、以及权衡压缩和视频质量的度量标准中的一个。
2. 如权利要求1所述的编码器，其中，当所述视频编码器所使用的连接改变到较低质量时，所述用于编码的编码装置选择性地选择所述至少一个长期参考块。
- 15 3. 如权利要求1所述的编码器，其中，当预见到所述视频编码器所使用的连接将改变到较低质量时，所述用于编码的编码装置选择性地选择所述至少一个长期参考块。
4. 如权利要求1所述的编码器，其中，所述用于编码的编码装置选择性地选择所述至少一个长期参考块来对背景数据编码并且选择性地选择更新近的参考块来对前景数据编码。
20
5. 如权利要求4所述的编码器，其中，所述更新近的参考块包括一个紧邻的过去的参考块。
6. 如权利要求5所述的编码器，其中，所述至少一个长期参考块包括至少一个就在所述紧邻的过去的参考块之前的块。
- 25 7. 如权利要求1所述的编码器，其中，所述在编码的时候检查的一个或多个因素包括下述因素的一个或多个：所述编码器对解码器处的失真的预期；编码器中帧缓冲器的数目；编码器中帧缓冲器的大小；来自解码器的任何反馈；改变数据信道质量的历史；改变图像区域质量从而对于每个至少一个正被编码的块在所述至少一个长期参考块和所述至少一个短期参考块之间选择
30 性地进行选择以最大化压缩、视频质量、权衡压缩和视频质量的度量标准中的一个的历史。

8. 如权利要求 7 所述的编码器, 其中, 所述在编码的时候检查的一个或多个因素还用于确定何时更新所述至少一个长期参考缓冲器。

9. 如权利要求 1 所述的编码器, 其中, 所述编码器包括多个长期参考块缓冲器。

5 10. 如权利要求 9 所述的编码器, 其中, 新近的参考块包括紧邻的过去的参考块。

11. 如权利要求 1 所述的编码器, 其中, 所述编码装置选择性地在使用所述至少一个长期参考块的编码(间编码)与使用内编码之间进行选择。

12. 如权利要求 11 所述的编码器, 其中, 所述编码装置对于所述至少一个长期参考块和在分像素网格上通过确定下述内容来进行分像素精确编码:

10 包括与实际像素位置一致的像素的初始像素位置;
包括位于两个初始像素位置之间的像素的水平或垂直插值的像素位置;
以及

对角线插值的像素位置。

15 13. 如权利要求 12 所述的编码器, 其中:
直接计算水平或垂直插值的像素位置和对角线插值的像素位置的一阶矩; 以及

估计水平或垂直插值的像素位置和对角线插值的像素位置的二阶矩。

14. 如权利要求 1 所述的编码器, 其中, 所述至少一个长期参考块缓冲器包括多帧缓冲器, 并且所述编码器的编码装置选择性地在使用参考块的编码(间编码)和使用内编码之间进行选择。

15. 如权利要求 14 所述的编码器, 其中, 所述编码装置在两种类型的间编码和内编码之间选择, 所述两种类型的间编码包括使用所述至少一个短期参考块(ST)的编码和使用所述至少一个长期参考块(LT)的编码, 并且其中:

25 所述编码装置通过使用将在先块的元素视为随机变量的递归最优每个像素估计来计算用于内编码和 ST 块的矩; 以及

所述编码装置通过使用将在先块的元素视为随机变量的递归最优每个像素估计来计算用于 LT 块的矩。

16. 如权利要求 15 所述的编码器, 其中, 更新所述 LT 块并且所述编码装置接收解码器反馈并且使用该反馈来确定何时更新所述 LT 块。

17. 如权利要求 16 所述的编码器, 其中, 所述编码装置使用所述反馈来

同步所述长期参考缓冲器。

18. 如权利要求 14 所述的编码器, 其中, 所述至少一个长期参考块包括在感兴趣区域中的块。

5 19. 如权利要求 1 所述的编码器, 其中, 所述至少一个长期参考块包括复合帧。

20. 如权利要求 1 所述的编码器, 其中, 所述至少一个长期参考块包括长期参考帧并且所述编码装置以逐块为基础来对帧进行编码。

21. 一种用于视频编码的方法, 所述方法包括下述步骤:

正常地编码并且存储正常质量参考帧;

10 还编码并存储高质量参考帧;

使用该高质量参考帧来对正被编码的帧的全部或部分进行编码。

22. 如权利要求 21 所述的方法, 其中, 所述高质量参考帧包括来自高质量信道状况的帧并且当高质量信道状况被低质量信道状况代替时使用该帧。

15 23. 如权利要求 22 所述的方法, 其中, 所述高质量参考帧在低质量信道状况时间段期间对背景数据进行编码。

24. 如权利要求 22 所述的方法, 包括周期性地或不定期地分配带宽以准备所述高质量参考帧的步骤。

25. 如权利要求 24 所述的方法, 其中, 所述周期性地或不定期地分配的步骤包括客户周期性地或不定期地剥夺在先和后继的带宽以创建高质量帧。

20 26. 如权利要求 24 所述的方法, 其中, 所述周期性地或不定期地分配的步骤包括服务器周期性地或不定期地提供突发带宽用于由客户进行高质量编码。

27. 如权利要求 21 所述的方法, 其中, 所述使用的步骤是响应于或者预见到将不利地影响在其上将传送正被编码的帧的信道的状况而进行的。

25 28. 一种用于对视频进行解码的方法, 所述方法包括下述步骤:

接收经编码的视频; 以及

不管在对正被解码的当前帧编码时经编码的视频是否是通过使用高质量长期帧进行编码的, 选择性地选择使用所存储的高质量长期帧来改进正被解码的帧, 以改进正被解码的当前帧。

30 29. 一种视频编码器, 包括:

多个帧缓冲器, 其中至少两个帧缓冲器存储来自非连续帧的视频帧信息;

编码装置，用于参考存储在所述多个帧缓冲器的一个或多个中的信息来对视频帧进行编码；以及

5 控制逻辑，用于通过选择跳跃更新、或连续更新、或一般更新、或者根据自由可选的更新参数的任意非周期性更新中的一个来更新所述多个帧缓冲器。

30. 如权利要求 29 所述的编码器，其中，在对一个视频序列进行编码时所述自由可选的更新参数保持固定。

视频编码方法及设备

5 技术领域

本发明的领域是图像和视频编码。本发明的其它领域包括图像和视频的数字通信和数字存储。

背景技术

10 由于精确描述图像的单个帧所需要的数字数据量呈向上曲线(sheer),因而图像数据的数字通信和存储是一项困难的任务。在视频中,数据量很快变得非常大。图像编码试图通过压缩图像数据即通过减少表示图像所必需的数据量来使得图像数据的通信和/或存储可管理。例如,通信资源具有有限的带宽。在无线通信介质中尤其如此。在图像数据编码中存在折衷。例如,减少
15 数据大小不应该使图像质量下降超过可接收的度。此外,必需管理计算成本和速度,尤其是在应节省计算资源和电力资源的设备中。视频编码/压缩方法的现代示例包括 MPEG-4 和 H.264。特别地,后者是专门设计用于分组网络上的视频传输。

许多设备访问不止一种通信介质。如膝上式电脑、个人数字助理(PDA)、
20 工作站或视频会议系统等设备可以访问多个网络。例如,一个设备可以访问几种不同类型的有线网络和无线网络。

现今许多视频压缩算法利用运动补偿来获得充分的压缩。运动补偿的基本思想如下。宏块(macroblock)表示图像数据块,例如,图像中 16×16 像素的方形区域。将待编码的当前帧中的宏块与参考帧中的一些宏块集合进行比较以找到最相似的一个。参考帧通常是图像中的在先帧。通常用像素差的绝对值之和或者像素之间的平方差来度量相似性。可以通过给定偏移向量来指定该最佳匹配块的位置,该偏移向量称作运动向量,其描述了待编码的当前
25 宏块与参考帧中的最佳匹配宏块之间的水平和垂直位置差。可能仅使用该运动向量来表示待编码的当前宏块。解码器一接收到该运动向量就可以从参考
30 帧取出参考块并且将该参考块粘贴到表示当前块的位置。如果参考块与当前块足够相似,则该直接替换可以提供适当的质量。如果它们不够接近,则解

码器可以可选地一起发送一些描述如何修改参考块以便使其与当前块更接近的额外信息。

在这两者中的任一种情况下，这都被称作 INTER (间) 编码。当解码器没有找到当前宏块的良好匹配时，解码器可以选择独自地对当前宏块进行编码，而不参考任何其它过去的块。这被称作 INTRA (内) 编码。在 INTER 和 INTRA 编码之间的选择是视频编码标准 MPEG、MPEG-2、MPEG-4 [T. Sikora, “The MPEG-4 Video Standard Verification Model”, IEEE Transactions on Circuits and Systems for Video Technology, vol. 7, no. 1, pp. 19-31, February 1997.]、H. 263 [G. Cote, B. Erol, M. Gallant and F. Kossentini, “H. 263+: Video Coding at Low Bit Rates,” IEEE Trans. Circ. And Systems for Video Techn, vol. 8, no. 7, pp. 849-865, Nov. 1998.] 和最新及当前技术 H. 264 中可得到的基本途径。

INTER 编码趋于比 INTRA 编码需要更少的比特，但是会传播误差。由于 INTRA 编码不参考在先帧，因而其不会传播在先帧中存在的误差。在 INTER 和 INTRA 编码之间进行选择涉及满足使用较少比特和对错误鲁棒的竞争目的。

在 INTER 和 INTRA 编码之间做出明智的选择是本领域关注的问题。涉及该主题的一篇论文是 R. Zhang, S. L. Regunathan, K. Rose 的 “Video Coding with Optimal Inter/Intra-Mode Switching for Packet Loss Resilience,”，IEEE Journal on Selected Areas in Communications, vol. 18, no. 6, pp. 966-76, June 2000。这篇论文提供了一种称作 ROPE (recursive optimal per pixel estimate, 递归最优每个像素估计) 的失真估计方法，其在估计失真中考虑两个因素：信道误差概率和被编码块的隐蔽能力 (concealability)。然后通过权衡减少失真 (如 ROPE 估计的) 和仅使用少数目的比特用于编码 (尤其是保持在目标速率约束之内) 的竞争目的来做出对于给定块 (诸如宏块) 的 INTER 和 INTRA 编码之间的选择。速率约束青睐 INTER 编码，误差约束青睐 INTRA 编码，而隐蔽误差的能力青睐 INTER 编码。

当用于传送视频数据的连接经受质量改变时，结果的视频解码会产生非常差的结果。当参考帧提供差质量的参考时，解码结果快速恶化。已提出的解决该问题的一种技术是保留多个帧。但是，这会使得编码负担和复杂度非常高。

存在许多多个参考帧方法的示例。[例如请参见, N.Vasconcelos and A.Lippman, "Library-based Image Coding," IEEE International Conference on Acoustics, Speech, and Signal Processing, vol.v, pp.V/489-V/492, 1994; T.Wiegand, X.zhang, and B. Girod, "Long term
5 Memory Motion-Compensated Prediction," IEEE Trans. Circ. And Systems for Video Techn., vol.9, no.1, pp.70-84, Feb.1999.]。在一个示例中, 当对视频的帧 N 中的块进行编码时, 编码器可能在帧 N-1、N-2、N-3、和 N-4 中寻找最佳可能匹配块。即, 可以搜索 4 个紧邻的过去帧以进行匹配。然后编码器会告诉解码器哪个参考帧提供了最佳匹配。例如, 可以分配 2 个
10 比特来描述 4 个帧中哪个提供了最佳匹配, 然后提供通常的运动向量来给出待编码的当前块与指定参考帧中最佳匹配块的位置之间的偏移。

在 T.Fukuhara, K.Asai, and T.Murakami, "Very Low Bit-Rate Video Coding with Block Partitioning and Adaptive Selection of Two Time-Differential Frame Memories," IEEE Trans. Circ. And Systems for
15 Video Techn. Vol.7, no.1, pp.212-220, Feb.1997 中, 仅使用了两个时间差分(time-differential)帧, 因此在计算复杂度上需要相对不大的增加。该双帧缓冲器是多帧缓冲器的特例, 其中只有两个参考帧。例如, 可以有一个短期参考帧(紧邻的过去帧)和一个长期参考帧(来自更远的过去的帧)。在 Fukuhara 等人的论文中, 一个帧是在先帧, 如许多混合编解码器中那样, 而
20 第二个帧包含根据预定规则周期性更新的来自更远的过去的帧。已经显示出多参考帧可以以计算负担和存储复杂度增加为代价在重构 PSNR(Peak Signal-to-Noise Ratio, 峰值信噪比)方面产生显著的增益。运动估计是混合视频编码系统中的主要性能瓶颈, 并且占据高于 80-90%的总编码时间。从而, 增加即使一个额外的帧缓冲器也会使编码时间加倍。对存储器需求来说
25 同样如此, 其中该增加也是线性的从而随着参考帧的数目增大是不允许的。

一种始终最佳连接(Always Best Connected, ABC)方法是在设备接入多个连接时使用的方法。诸如膝上式电脑或 PDA 的设备可能能够接入运行在不同速率的几种不同类型的无线或有线网络。例如, 该设备可能能够使用以太网连接(10 Mbps)、无线 LAN(11 Mbps)、HDR(400-500 Kbps)、1xRTT(64 Kbps)、
30 和 GPRS(16 Kbps)进行通信。在任何给定时刻, 假定该用户没有选择其用户汇总目录(profile)中任何其它网络的话, 该设备将使用在该特定时刻其可接

入的最佳连接来运行。最佳连接经常是具有最高数据速率的连接，但是也包
括其它因素(例如，错误率、延迟等等)。如果最佳连接变得不可用，或者其
质量恶化到不再是最佳连接的点，则预期设备无缝切换到另外的连接，即新
的最佳连接。该连接经常是处于较低速率的连接。还预期该设备周期性地探
5 查所有连接以查看哪个可用。如果高速率连接变得不可用并且然后再次变得
可用，则预期该设备发现该可用性，并且切换回去以使用该网络。

发明内容

本发明使用了双帧并且可能是多帧的方法。本发明的实施例包括做出使
10 用长期参考帧来进行 INTER 编码还是进行 INTRA 编码的决定，该长期参考帧
是不同于紧邻的过去参考帧的帧。其它实施例包括使用长期和短期参考块，
并且在 INTER 编码块和 INTRA 编码这两种类型之间做出决定。

根据本发明的实施例，长期帧是高质量帧。该高质量帧在特定条件下可
以用作参考帧。

15

附图说明

图 1A 和 1B 是图示根据本发明实施例的示例性系统的框图；

图 2 是图示在速率切换情景中的本发明的示例性长期参考帧编码方法的
示意图；

20 图 3A 和 3B 示意性地图示了两种调度器；图 3B 的调度器分配额外的带
宽给用户以产生高质量参考帧；

图 4 示意性地图示了根据本发明实施例的带有反馈的长期参考帧编码方
法；以及

25 图 5 示意性地图示了根据本发明实施例的带有反馈的长期参考帧编码方
法。

具体实施方式

本发明针对用于对诸如视频的图像数据进行编码的方法、设备和系统。
本发明还针对用于对数据进行解码的方法、设备和系统。本发明可用在利用
30 通信介质的许多类型的系统中。例如，本发明可以用在对等型通信中和服务
器客户端型通信中，并且还可以用于例如对图像数据进行编码以存储。本发

明的一个特定示例性实施例是视频会议系统和方法。本发明的实施例可能特别适用于其中视频连接改变速率显著量的情形。

在描述本发明实施例的过程中，讨论将着重于对帧的处理。这将是对本发明的普通实现。然而，如同标准中那样，基于块来进行处理和比较。对于本发明的目的而言，块大小通常是任意的。如上述标准中那样，普通实现将使用宏块。如本文中所采用的，可以将一帧划分成多个块，诸如宏块。然而，对帧大小没有特别限制，尽管本发明的实施例可能使用标准中的帧大小。

本发明可以作为许多上述标准的修正来应用。本发明可以利用在这些标准的任一个中计算的运动补偿向量，并且也可以利用其它方法进行向量计算。通过本发明，改进了公知标准，尽管本发明利用了通用标准框架和公知的硬件实现。

本发明使用了一种双帧以及可能多帧的方法。本发明的实施例包括做出使用长期参考帧来进行 INTER 编码还是进行 INTRA 编码的决定，该长期参考帧是不同于紧邻的过去参考帧的帧。其它实施例包括使用长期和短期参考块，并且在 INTER 编码块和 INTRA 编码这两种类型之间做出决定。递归最优每个像素估计(ROPE)用来计算用于 INTRA 编码和短期参考块的矩(moment)，而修正的递归最优每个像素估计用来计算用于长期参考块的矩，其中将在先块的元素视作随机变量。

根据本发明的实施例，长期帧是高质量帧。高质量帧在特定条件下可以用作参考帧。例如，该条件可以包括正在通过其接收视频的连接的质量改变。高质量帧例如可以是从较高质量连接的时刻起的过去帧。在另一示例性实施例中，在通信质量的预期改变之前保留了一个高质量帧或多个高质量帧。这可以是对低劣质的量预先警告或者预测的结果，或者可以周期性地，例如，以确保高质量帧可得到。例如，可以通过强迫使在前帧和后继帧不占用其典型占用的所分配的比特速率而以在前和后继帧为代价偶尔对高质量参考帧进行编码。在额外的实施例中，将帧划分成静态和动态部分。例如，可以识别背景部分和前景部分。背景部分的高质量参考帧由于其静态特性因而可以使用充分长的时间段，而前景部分由最近的或者新近接收的帧进行运动补偿。

本发明的实施例包括双帧或多帧保留实施例和单个参考帧保留实施例。例如，双帧实施例将(待编码的当前帧中的)每个块与过去的高质量参考帧和新

近的(短期)参考帧两者进行比较,并且通过确定产生较好结果的帧来进行运动补偿。将过去的高质量帧保持在长期参考帧缓冲器中而将新近的参考帧保持在新近的参考帧缓冲器中。当新近的参考帧比过去的参考帧质量高时,则将其作为过去的高质量参考帧保留在长期参考帧缓冲器中,并且丢弃先前的过去的高质量参考帧。在示例性单帧实施例中,保持单个参考帧缓冲器。在该单个参考帧缓冲器中保持高质量长期参考帧直到其废弃为止,例如,该高质量长期参考帧与待编码的当前帧的相似性降到特定阈值之下,从而做出了开始使用紧邻的过去帧作为参考帧的决定。

本发明的双帧或多帧实施例中的编码器复杂度没有现有技术中的复杂,这是因为保留了高质量帧作为水准基点(bench mark)。因此,例如只要长期高质量帧满足某尺度就不需要进行大量的比较。此外,可以不定期地或周期性地长期帧与新近帧比较的测试。本发明的单个帧实施例提供了低复杂度编码器,而不需要额外的编码器复杂度。

本发明的优选实施例方法通过使用来自过去的较高速率连接的数据进行长期运动补偿。考虑下述情形,其中,视频编码装置通过使用例如始终最佳连接(ABC)网络传输,并且当前的最佳连接质量恶化从而用户必须切换到较低速率的连接。考虑从1一直到N的帧是高质量的,因为它们是在高速率信道上传输的,并且从帧 N+1 开始,用户现在具有低速率连接。利用诸如 MPEG-2 或 H.263 的标准视频编码装置,在低速率连接持续时的期间视频质量将较低。帧 N+1(其必须以低速率传输)可以使用来自帧 N(其具有高质量)的 INTER 编码,并且因而帧 N+1 的质量将受益于通过使用高质量参考帧来完成块匹配运动补偿这一事实。但是由于削减了速率,所以平均而言帧 N+1 将比帧 N 质量低。

当对帧 N+2 编码时,现在通过使用帧 N+1 完成 INTER 编码,因而质量下降得更多。相应于较低速率,较低质量的传输,人们可以预计该质量下降相当快直到获得某个新的稳定状态。在这种情景下,本发明的实施例使用双帧运动补偿方法,其中长期帧缓冲器将包含例如帧 N,即在良好连接失效且较差连接替换该良好连接之前来自高质量高速率连接的最后一帧。从而,对于编码帧 M,其中 $M > N+1$,将有两个可用来进行运动补偿的帧:短期帧 M-1(低质量的紧邻的过去帧)、和长期帧 N(高质量的较远的过去的帧)。

考虑其中待编码的场景是很静态的,如很多视频会议应用那样的情形。

当设备从比方说 10Mbps 的无线 LAN 连接向下切换到 16Kbps 的 GPRS 连接时，场景中的背景可以保持相当固定。使用来自过去的超高质量的长期的运动补偿，可以提供对当前帧的非常精确的匹配。另一方面，由于 16Kbps 是非常低的速率并且将产生低质量的视频，所以仅使用过去的短期帧用于运动补偿可能给出很差的结果。过去的高质量帧可以提供优势还存在其它原因。以不变的连接也好还是以变化的连接也好，过去的高质量帧提供优势的另一种可能的情形都会发生。例如，让考虑其中人或物离开场景并且然后稍后返回到该场景的视频会议应用。仅使用紧邻的过去帧作为编码参考不会提供高质量信息。实际上，在他们返回到场景的第一帧中，根本没有用于他们的参考信息，无论是高质量的还是低质量的。如果保持了高质量过去帧，则即使当他们第一次返回到场景中时，也将具有高质量的参考信息。因而，这就是即使在场景是非静态的时候来自长期过去的高质量参考帧也可以是有用的情况。这种类型的优势还可能发生在视频会议情形中的人不时转过脸从而他们的脸对一些帧是不可见的然后他们又转回来的时候。如果仅使用紧邻的过去帧用作参考信息，那么当他们转回来的时候，最初将没有用于他们脸部的参考。

使用过去的较高速率连接的修正和额外的实施例包括其中存在例如对连接质量变化的预先警告或预测的情形。考虑其中设备具有某预警或者能够预测当前最佳连接的即将发生故障并且需要切换到低速率连接的情况。了解到解码器可以很好地利用特别高质量帧用于长期帧缓冲器，则发送器事实上可以在进行切换之前对具有特别高质量的一帧进行编码。即，通过强迫使在先帧和后继帧不占用它们通常占用的比特速率，编码器可以例如使用更精细的量化来产生特别高质量的帧。当切换到较低速率连接时，解码器可以使用该特定帧作为双帧运动补偿中的长期帧缓冲器。

在另一修正中，不定期地或周期性地对高质量帧编码。这样做的一个原因是预防连接质量的变化。在其它情况下，即使在不考虑连接质量变化的场合这也可能是有用的。考虑使得单个用户使用(例如)较少的比特用于 9 个帧以便将额外的比特赋予第 10 帧(由此产生较高质量的长期参考帧以用于接下来的 10 个帧)。这可以用来在整体上产生较高质量的视频而在所需带宽上没有增加。额外的实施例包括在低流量时间期间或者在额外比特速率可得时编码高质量帧。这也可以通过服务器例如分配用于高质量帧的比特速率来控制。

例如，如果许多用户正在共享带宽，由服务器完成的调度器可以在用户(客户端)之间轮转带宽的一部分。每个客户端周期性地被分配一段额外带宽。该额外的带宽用于对要用作参考帧的高质量帧进行编码。另一实施例涉及到通过强迫使某些在先帧或者后继帧不占用为它们分配的比特以创建高质量参考帧而使客户端或者对等用户使用它们自身的带宽分配来补偿。

5 本发明的实施例还可以识别数据的静态和动态部分并且基于这种识别来调整编码。假设存在这种情形，其中待编码和传输的视频由静止的或近似静止的背景部分和运动的前景部分组成。这可以发生在例如视频电话和视频会议应用中。如果待编码的当前帧可以被划分成背景区域和前景区域，则可以通过使用长期过去中的高质量帧对背景(假定静态)进行运动补偿，而通过使用最新近的帧对前景(假定非静止)进行运动补偿。此外，存在下述情况，其中数据的动态部分诸如前景受益于对高质量长期帧的参考。例如，其中对象或者人离开且稍后返回的视频中的情况给出了其中动态信息受益于一个或多个长期高质量帧的情况。

15 本发明的额外实施例与前述实施例或者下面列出的实施例的任一个相结合来使用单个帧长期预测。取替应用双帧概念，使用单个帧，即直到其变得废弃为止用于运动补偿的具有高质量的长期过去帧(即，其与当前帧如此不同以致最好不再使用该紧临过去的低质量帧用于预测)的时刻为止。与双帧缓冲器概念不同，这在编码器处不需要任何额外的复杂性。

20 本发明的额外实施例与解码器隐蔽(concealment)有关。当视频信号被压缩并且在不可靠的信道上传输时，必须采取一些处理误差的策略。一个策略是误差隐蔽(error concealment)，这是解码器可采用的一系列后处理方法。当接收到的帧的一部分被毁坏时，解码器的后处理方法试图对观众隐蔽此点。存在各种替换性办法：空间域插值、频域估计、涉及在参考帧中定位适当块的时间隐蔽(temporal concealment)。在本发明的实施例中，假定编码器根据标准过程编码，即在先帧用作参考帧。然而，解码器可以利用额外的参考帧(长期过去的高质量帧)以及在先帧。不是只在有丢失的时候才隐蔽，解码器可以在接收到质量非常低劣的宏块时利用隐蔽，可以通过使用过去的高质量参考帧用较高质量的块替代该质量非常低劣的宏块。这可以称作改进而不是隐蔽，

25 但是可以将其视为时间隐蔽的变体。本质上，当质量足够差时，可以选择认为丢失了块即使事实上该块并没有丢失，并且可以借用其中使用高质量过去

30

帧代替丢失块的丢失隐蔽(loss concealment)方法。

本发明的实施例还包括上述编码决定和框架的各种组合。现在将详细讨论某些特定优选实施例，对本领域技术人员来说各种额外的发明特征将是显而易见的。

5 图 1A 和 1B 示出了用于实施本发明的示例性系统。图 1A 和 1B 的系统可以使用本发明上面实施例的任一个。图 1A 示出了编码器。图 1B 示了解码器。在编码器中，长期帧估计存储器 10 存储对于用于模式选择的长期帧的估计。这些估计是通过修正的 ROPE 算法确定的。另一个构成额外存储 12 的存储器可以用于存储比存储在长期存储器 10 中的帧估计更早的帧估计。在使
10 用了反馈的本发明实施例中采用额外存储 12。短期帧估计存储在短期帧估计存储器 14 中。

速率/编码装置控制逻辑 16 使用帧估计来计算失真并且进行速率-失真优化。速率/编码装置控制逻辑还控制量化参数(quantization parameter, QP)选择、运动补偿和重新解码(re-decoding)控制。开关 18 由速率/编码装置控制逻辑控
15 制来发送图像像素(INTRA 编码)或者图像像素的差(INTER 编码)到离散余弦变换器(discrete cosine transformer, DCT)20。给定由速率/编码装置控制逻辑 16 选择的 QP，量化器 22 对 DCT 20 输出的图像 DCT 系数进行量化并且输出量化指数(系数)。给定量化指数，反向量化器 24 利用 QP 重构图像 DCT 系数。反向 DCT 26 接收 DCT 系数(例如 8×8 块的 DCT 系数)并且将它们转换成图像
20 像素或者像素差值。

在 INTER 编码的情况下，开关 28 添加用来自运动补偿器 30 的运动补偿获得的预测。在 INTRA 编码的情况下开关 28 保持断开状态。运动补偿器 30 通过使用来自短期帧存储器 32 的所存储的短期帧和来自长期帧存储器 34 的所存储的长期帧作为预测来对当前帧进行运动估计。等价地，运动补偿器可
25 以使用在先前阶段中获得的运动向量进行运动补偿。运动补偿器还对于对其反馈可用并且其参考帧已经在存储器 36 中进行缓冲的帧进行重写解码，其中存储器 36 存储额外过去的重新编码后的帧，该额外过去的重新编码后的帧用于对于对其反馈当前可用的帧进行重新解码(在本发明的反馈实施例中)。则额外存储 12 中所存储的估计等于重新编码的，即等于存储在 36 中的帧。将其
30 单独示出，这是因为这里其用作模式选择而在 36 中用于重新解码。

在解码器(图 1B)处，反向量化器 38 接收量化指数并且利用 QP 重构 DCT

系数。该反向量化器在每个方面都等价于编码器中的反向量化器 24。反向 DCT 40 接收 DCT 系数并且将它们转换成图像像素或者差值像素值。反向 DCT 等价于编码器中的反向 DCT 26。开关 42 添加预测或者不添加预测，并且该开关 42 等价于编码器的开关 28。短期帧存储器 44 存储在先的重构帧，
5 运动向量解码器 46 对接收到的运动向量进行解码，运动补偿器 48 通过使用由 46 提供的运动向量以及由短期帧存储器 44 和长期帧存储器 50 提供的参考进行运动补偿，该长期帧存储器 50 存储重构后的长期帧。

现在将讨论本发明的优选的额外的特定方法，并且该特定方法也可以用在图 1A 和 1B 或其它系统中。图 1A 和 1B 的系统 10 中用于对视频编码的方法使用高质量帧。将视频数据划分成帧。可以将帧分割成区域，诸如背景和前景。如果编码器拥有关于视频数据的智能(intelligence)则这将是有益的。还可以将视频数据划分成诸如块或宏块的单元，这使编码计算更简单。也可以将帧作为一个整体来对待。编码器确定在解码器处的预期失真。这可以以信道误差概率和过去编码模式的历史为基础。编码器使用高质量过去帧以改进
10 正被编码的帧。

编码器还可以识别数据信道质量变化并且存储正被编码的图像或图像区域的内容中的改变。可以使用反馈来存储解码器处的预期失真的更新版本。然后通过使用编码器对解码器处失真的预期、帧缓冲器的数目、帧缓冲器的大小、来自解码器的任何反馈、改变数据信道质量的历史、改变图像或区域
20 质量从而对于所传送的图像的正被编码的每个帧或区域(诸如宏块)来确定存储在帧缓冲器中的内容的历史、来确定哪个帧缓冲器用作参考缓冲器。进行确定以最大化压缩比或者视频质量或者权衡两者的度量标准。

INTER 和 INTRA 编码之间的切换是本发明实施例中实现的决定。利用了递归最优每个像素估计(ROPE)，其在运行在分组擦除(packet erasure)信
25 道上的混合式视频编码装置中提供模式决定。

假设视频比特流在分组擦除信道上传送。将每个帧分割成块组(Groups of Blocks, GOB)。每个 GOB 包含单个水平片断(slice)的宏块(macroblock, MB)并作为单个分组传送。由于再同步标记(resynchronization marker)所以可以独立接收每个分组并对其解码。这样，单个分组丢失就抹去了一个 MB 片断，
30 但是保持帧的剩余部分无损。

设 p 是分组擦除的概率，其也是每单个像素的擦除概率。当解码器检测

到擦除时，应用误差隐蔽方法。解码器用来自在先帧的一个宏块替换丢失的宏块，使用在丢失的 GOB 上面的 GOB 中的三个最近的宏块的运动向量(MV)的中值作为运动向量。如果上面的 GOB 也丢失了(或者最近的 MB 都是内编码的因而不存在运动向量)，则使用全零(0,0)MV，并且使用来自在先帧的处于共同位置(co-located)的宏块来替代丢失的宏块。

由 f_n 表示初始视频信号的帧 n ，其在编码器处被压缩并重构为 \hat{f}_n 。接收器处帧 n 的解码后的(可能是误差隐蔽后)重构用 \tilde{f}_n 表示。编码器并不知道 \tilde{f}_n 并且将其作为随机变量来对待。

设 f_n^i 表示帧 n 中像素 i 的初始值，并且设 \hat{f}_n^i 表示它的编码器重构。解码器处的重构后的值(可能在错误隐蔽之后的)用 \tilde{f}_n^i 表示。像素 i 的预期失真为：

$$d_n^i = E\{(f_n^i - \tilde{f}_n^i)^2\} = (f_n^i)^2 - 2f_n^i E\{\tilde{f}_n^i\} + E\{(\tilde{f}_n^i)^2\} \quad (1)$$

d_n^i 的计算需要被估计的图像序列 \tilde{f}_n^i 的随机变量的一阶矩和二阶矩。为计算这些值，在 ROPE 中提出了递归函数，其中有必要将内编码和间编码的 MB 的情况分开。

对于内编码的 MB，在相应于正确接收分组的 $1-p$ 的概率下 $\tilde{f}_n^i = \hat{f}_n^i$ 。如果分组丢失，但是在先的 GOB 正确，则基于中值运动向量的隐蔽致使解码器将当前帧中的像素 i 与在先帧中的像素 k 相关联。这样，在 $p(1-p)$ 的概率下 $\tilde{f}_n^i = \tilde{f}_{n-1}^k$ 。最后，如果当前的和在先的 GOB 分组都丢失了，则 $\tilde{f}_n^i = \tilde{f}_{n-1}^i$ (发生概率为 p^2)。则在内编码的 MB 中像素的两个矩为：

$$E\{\tilde{f}_n^i\} = (1-p)\hat{f}_n^i + p(1-p)E\{\tilde{f}_{n-1}^k\} + p^2 E\{\tilde{f}_{n-1}^i\} \quad (2)$$

$$E\{(\tilde{f}_n^i)^2\} = (1-p)(\hat{f}_n^i)^2 + p(1-p)E\{(\tilde{f}_{n-1}^k)^2\} + p^2 E\{(\tilde{f}_{n-1}^i)^2\} \quad (3)$$

对于间编码的 MB，假设它的真实运动向量是使得从在先帧中的像素 j 预测像素 i 。则，该像素的编码器预测是 \hat{f}_{n-1}^j 。预测误差 e_n^i 被压缩，量化余量(quantized residue)是 \hat{e}_n^i 。编码器重构为：

$$\hat{f}_n^i = \hat{f}_{n-1}^j + \hat{e}_n^i \quad (4)$$

编码器传送 \hat{e}_n^i 和 MB 的运动向量。如果正确接收到分组，则解码器知道了 \hat{e}_n^i 和 MV，但是仍必须使用其自身对在先帧中的像素 j 的重构 \tilde{f}_{n-1}^j ，而该重构可能不同于编码器值 \hat{f}_{n-1}^j 。这样，像素 i 的解码器重构如下给出：

$$\tilde{f}_n^i = \tilde{f}_{n-1}^j + \hat{e}_n^i \quad (5)$$

编码器再次将 \tilde{f}_{n-1}^j 作为随机变量建模。对后两种情况矩的偏差与内编码的 MB

类似，但是对于其中没有传送误差(概率 $1-p$)的第一种情况则不同。则对于间编码的 MB 中的像素 \tilde{f}_n^i 的一阶矩和二阶矩如下给出：

$$E\{\tilde{f}_n^i\} = (1-p)(\hat{e}_n^i + E\{\tilde{f}_{n-1}^j\}) + p(1-p)E\{\tilde{f}_{n-1}^k\} + p^2E\{\tilde{f}_{n-1}^i\} \quad (6)$$

$$E\{\tilde{f}_n^i\}^2 = (1-p)((\hat{e}_n^i)^2 + 2\hat{e}_n^i E\{\tilde{f}_{n-1}^j\} + E\{\tilde{f}_{n-1}^j\}^2) + p(1-p)E\{\tilde{f}_{n-1}^k\}^2 + p^2E\{\tilde{f}_{n-1}^i\}^2 \quad (7)$$

- 5 在编码器处执行这些递归以便计算解码器处预期的失真。编码器可以在它的编码决定中采用该结果，以最优地选取用于每个 MB 的编码模式。

ROPE 考虑由于压缩和传送误差两者所导致的预期失真以进行最优的模式切换。编码器对于给定的比特速率和分组丢失率以最优方式以宏块为基础在内编码或间编码之间进行切换。目标是在比特速率约束 R 下最小化总体失真 D 。通过使用拉格朗日乘子 λ (Lagrange multiplier λ)，ROPE 算法最小化总体成本 $J=D+\lambda R$ 。各个 MB 对该成本的贡献是累加的，从而可以以宏块为基础对其最小化。因而，通过下述最小化来选取每个 MB 的编码模式

$$\min_{\text{mode}} J_{MB} = \min_{\text{mode}} (D_{MB} + \lambda R_{MB}) \quad (8)$$

15 其中，MB 的失真 D_{MB} 是各个像素的失真贡献的总和。通过如 J.Choi 和 D.Park 所著“A Stable Feedback Control of the Butter State Using the Controlled Lagrange Multiplier Method”，IEEE Trans. Image Proc., vol.3, pp.546-58, September, 1994 中那样修改 λ 来实现速率控制。

本发明的实施例使用双帧缓冲器以及速率-失真框架下的最优模式切换。双帧缓冲器的基本效用如下。当对帧 n 编码时，编码器和解码器都在存储器中保持两个参考帧。短期参考帧是帧 $n-1$ 。长期参考帧是比方说帧 $n-k$ ，其中 k 可以是可变的但总是大于 1。现在描述如何选取长期参考帧。

25 在一种称之为跳跃更新(jump updating)的方法中，长期参考帧在从最新帧 $n-2$ 到最旧帧 $n-N-1$ 的范围内变化。当对帧 n 编码时，如果长期参考帧是 $n-N-1$ ，则当编码器继续前进到对帧 $n+1$ 编码时，短期参考帧将向前行进 1 个到帧 n ，并且长期参考帧将向前跳 N 个到帧 $n-1$ 。长期参考帧将然后对 N 个帧保持静止，并且然后又向前跳跃。将 N 称作跳跃更新参数。

另一方法是连续更新长期帧缓冲器从而长期帧缓冲器包含距离当前缓冲器固定时间距离的帧。因而，该缓冲器对每个帧 n 来说总是包含 $n-D$ 帧。将 D 称作连续更新参数。

注意到跳跃更新和连续更新都可以看作更一般的(N,D)更新策略的特例,在(N,D)更新策略中,长期参考帧向前跳跃N从而成为与待编码的当前帧向后距离为D的帧,并且然后对N个帧保持静止,然后再次向前跳跃。对于一般的(N,D)更新,帧k可以使LT帧作为新近帧k-D或者较早帧k-N-D+1。

- 5 在定义的跳跃更新中,对每个序列可以自由地选择N,而D=2,(意思是当更新发生时,LT帧向前跳跃N从而成为帧n-2)。在连续更新中,对于每个序列可以自由选择D,而N固定为1。最一般的更新策略将没有固定的N或D;而是无规律地在需要时将长期帧缓冲器更新为最有用的任何帧。在另一种情况中,在对一个序列编码时(N,D)保持固定。现在描述如何在编码模式中做出
- 10 选择。

- 每个宏块可以以三种编码模式之一进行编码:内编码、使用短期缓冲器的间编码(inter-ST编码)、使用长期缓冲器的间编码(inter-LT编码)。将使用修正的ROPE算法在这三种编码模式中做出选择。一旦选择了编码模式,则对比特流进行编码的语法就几乎等同于单个帧缓冲器的标准情况了。唯一的修正
- 15 正在于,如果选取了间编码,则将发送单个比特以指示使用了短期或长期帧。

现在描述在编码模式中如何做出选择。如前,分别使用 f_n 、 \hat{f}_n 、 \tilde{f}_n 来表示初始帧n、压缩帧的编码器重构、以及帧的解码器版本。假定长期帧缓冲器在m个帧之前更新过。这样,其在发送器处包含 \hat{f}_{n-m} 而在接收器处包含 \tilde{f}_{n-m} 。帧n中的像素i的预期失真由公式1给出。

- 20 为了计算公式1中的矩,用于内编码和inter-ST编码的MB中的像素的递归步骤与常规ROPE算法中的相应步骤等同。对于经inter-LT编码的MB中的像素,假定MB的真实运动向量是使得从帧n-m(其中m>1)中的像素j预测帧n中的像素i的运动向量。该像素的编码器预测是 \hat{f}_{n-m}^j 。压缩预测误差 e_n^i ,并且量化余量用 e_n^i 表示。像素的编码器重构是:

$$25 \quad \tilde{f}_n^i = \hat{e}_n^i + \tilde{f}_{n-m}^j \quad (9)$$

因为接收器无法访问 \hat{f}_{n-m}^j ,所以其使用 \tilde{f}_{n-m}^j :

$$\hat{f}_n^i = \hat{e}_n^i + \hat{f}_{n-m}^j \quad (10)$$

- 当MB丢失时,计算三个最近MB的中值运动向量,并且该中值运动向量用于将当前帧中的像素i与在先帧中的像素k相关联。使用与原始ROPE算法
- 30 中相同的参数,计算经inter-LT编码的MB中的像素的一阶矩和二阶矩 \tilde{f}_n^j ,

$$E\{\tilde{f}_n^i\} = (1-p)(\hat{e}_n^i + E\{\tilde{f}_{n-m}^j\}) + p(1-p)E\{\tilde{f}_{n-1}^k\} + p^2E\{\tilde{f}_{n-1}^i\} \quad (11)$$

$$E\{\tilde{f}_n^j\} = (1-p)((\hat{e}_n^j)^2 + 2\hat{e}_n^j E\{\tilde{f}_{n-m}^j\} + E\{(\tilde{f}_{n-m}^j)^2\}) + p(1-p)E\{\tilde{f}_{n-1}^k\} + p^2 E\{\tilde{f}_{n-1}^i\} \quad (12)$$

注意到仍然是使用在先帧 n-1 而不是长期帧来完成误差隐蔽。进行该误差隐蔽而不管上面的三个 MB 是 inter-ST 编码的或 inter-LT 编码的还是两者的某种组合。运动向量可以是高度不相关的。如果上面的 GOB 也丢失了，则使用来自在先帧的处于共同位置(co-located)的块来隐蔽 MB。

邻近的不相关运动向量的存在对运动向量编码效率产生消极影响。由于来自它们邻近的运动向量的对运动向量的不正确预测导致存在比特速率损失(bit rate loss)。此外，因为对于每个间编码的 MB 都使用一个比特来指定帧缓冲器所以降低了压缩效率。尽管如此，实验结果表明速率-失真优化对这些额外的比特建模并且仍然能够得到优越的压缩性能。

常规 ROPE 的另一修正包括延伸半像素(half-pel)运动向量(或者其它分像素(fractional pixel)运动向量)的收益(benefit)，但是避免完全精确的半像素建模或者其它分数建模这是因为其在 ROPE 中产生太高的惩罚。假设仅使用运动向量的整数部分来仍然实现误差隐蔽(EC)，并且因而不改变用于内编码的 MB 的公式 2 和 3。返回到用于间编码的 MB 的公式 6 和 7，观察到项 \hat{e}_n^j 、 $E\{\tilde{f}_{n-1}^k\}$ 、 $E\{\tilde{f}_{n-1}^i\}$ 、 $E\{(\tilde{f}_{n-1}^k)^2\}$ 、和 $E\{(\tilde{f}_{n-1}^i)^2\}$ 保持不变。然而， $E\{\tilde{f}_{n-1}^j\}$ 和 $E\{(\tilde{f}_{n-1}^j)^2\}$ 的计算变得很关键。像素坐标 j 现在指向覆盖了 4 倍于原始图像的面积的内插网格中的位置。

对于该计算，区分半像素网格上的三种类型的像素：与实际(初始)像素位置一致的像素(称作整数指数像素，它们不需要插值)、(水平地或垂直地)位于两个整数指数像素中间的像素、位于四个整数指数像素对角线中间的像素。使用双线性插值，从而插值后的值简单地是两个或四个近邻整数指数像素的平均。对于整数指数像素，递归方程与常规 ROPE 算法的等同，并且该估计是最优的。

对于水平或垂直插值后的像素，假定插值后的像素域上的 j 与原始像素域中使用像素 k_1 和 k_2 插值后的像素相对应。一阶矩在计算上是易处理的：

$$E\{\tilde{f}_{n-1}^j\} = \frac{1}{2} [1 + E\{\tilde{f}_{n-1}^{k_1}\} + E\{\tilde{f}_{n-1}^{k_2}\}] \quad (13)$$

但是二阶矩的表达式为：

$$E\{(\tilde{f}_{n-1}^j)^2\} = \frac{1}{4} \left[1 + E\{(\tilde{f}_{n-1}^{k_1})\} + E\{(\tilde{f}_{n-1}^{k_2})^2\} + 2E\{\tilde{f}_{n-1}^{k_1}\} + 2E\{\tilde{f}_{n-1}^{k_2}\} + 2E\{\tilde{f}_{n-1}^{k_1}\tilde{f}_{n-1}^{k_2}\} \right] \quad (14)$$

最后一项要求计算其水平/垂直维等于图像中像素数目的矩阵的相关。这可以在小尺寸的图像上或者利用充分计算资源实施。但是，使用典型的计算资源对于典型大小的图像，这计算上是不可行的。优选地使用余弦不等式来取其

5 近似:

$$E\{(\tilde{f}_{n-1}^j)^2\} \leq \frac{1}{4} \left[1 + E\{(\tilde{f}_{n-1}^{k_1})^2\} + E\{(\tilde{f}_{n-1}^{k_2})^2\} + 2E\{\tilde{f}_{n-1}^{k_1}\} + 2E\{\tilde{f}_{n-1}^{k_2}\} + 2\sqrt{E\{(\tilde{f}_{n-1}^{k_1})^2\}E\{(\tilde{f}_{n-1}^{k_2})^2\}} \right] \quad (15)$$

对于经对角插值的像素，假定经插值的像素网格上的 j 是对原始像素域中的像素 k_1 、 k_2 、 k_3 、 k_4 插值的结果。一阶矩可以精确地计算如下:

$$E\{\tilde{f}_{n-1}^j\} = \frac{1}{4} \left[2 + E\{\tilde{f}_{n-1}^{k_1}\} + E\{\tilde{f}_{n-1}^{k_2}\} + E\{\tilde{f}_{n-1}^{k_3}\} + E\{\tilde{f}_{n-1}^{k_4}\} \right] \quad (16)$$

10 二阶矩的精确表达式如下:

$$E\{(\tilde{f}_{n-1}^j)^2\} = \frac{1}{16} \left[4 + E\{(\tilde{f}_{n-1}^{k_1})^2\} + E\{(\tilde{f}_{n-1}^{k_2})^2\} + E\{(\tilde{f}_{n-1}^{k_3})^2\} + E\{(\tilde{f}_{n-1}^{k_4})^2\} + 4E\{\tilde{f}_{n-1}^{k_1}\} + 4E\{\tilde{f}_{n-1}^{k_2}\} + 4E\{\tilde{f}_{n-1}^{k_3}\} + 4E\{\tilde{f}_{n-1}^{k_4}\} + 2E\{\tilde{f}_{n-1}^{k_1}\tilde{f}_{n-1}^{k_2}\} + 2E\{\tilde{f}_{n-1}^{k_1}\tilde{f}_{n-1}^{k_3}\} + 2E\{\tilde{f}_{n-1}^{k_1}\tilde{f}_{n-1}^{k_4}\} + 2E\{\tilde{f}_{n-1}^{k_2}\tilde{f}_{n-1}^{k_3}\} + 2E\{\tilde{f}_{n-1}^{k_2}\tilde{f}_{n-1}^{k_4}\} + 2E\{\tilde{f}_{n-1}^{k_3}\tilde{f}_{n-1}^{k_4}\} \right] \quad (17)$$

应用如与水平/垂直的情况相同的近似，得到:

$$E\{(\tilde{f}_{n-1}^j)^2\} \leq \frac{1}{16} \left[4 + E\{(\tilde{f}_{n-1}^{k_1})^2\} + E\{(\tilde{f}_{n-1}^{k_2})^2\} + E\{(\tilde{f}_{n-1}^{k_3})^2\} + E\{(\tilde{f}_{n-1}^{k_4})^2\} + 4E\{\tilde{f}_{n-1}^{k_1}\} + 4E\{\tilde{f}_{n-1}^{k_2}\} + 4E\{\tilde{f}_{n-1}^{k_3}\} + 4E\{\tilde{f}_{n-1}^{k_4}\} + 2\sqrt{E\{(\tilde{f}_{n-1}^{k_1})^2\}E\{(\tilde{f}_{n-1}^{k_2})^2\}} + 2\sqrt{E\{(\tilde{f}_{n-1}^{k_1})^2\}E\{(\tilde{f}_{n-1}^{k_3})^2\}} + 2\sqrt{E\{(\tilde{f}_{n-1}^{k_1})^2\}E\{(\tilde{f}_{n-1}^{k_4})^2\}} + 2\sqrt{E\{(\tilde{f}_{n-1}^{k_2})^2\}E\{(\tilde{f}_{n-1}^{k_3})^2\}} + 2\sqrt{E\{(\tilde{f}_{n-1}^{k_2})^2\}E\{(\tilde{f}_{n-1}^{k_4})^2\}} + 2\sqrt{E\{(\tilde{f}_{n-1}^{k_3})^2\}E\{(\tilde{f}_{n-1}^{k_4})^2\}} \right] \quad (18)$$

并且使用该上限来近似二阶矩。

15 本发明的一个实施例是修正的 H.263+视频编解码器。这在实验中已经过验证。以两种方式修正了现有的 H.263+视频编解码器。在单帧(single-frame, SF)运动补偿的情况下，使用 ROPE 算法用于模式切换决定。结果的比特流完全符合 H.263+标准。其次，修正 H.263+视频编解码器以利用一个额外的(长期)帧缓冲器。如上面所引用的 Fukuhara 等的文章中那样根据更新参数 N 来周期

性地更新该帧缓冲器。对于单帧和双帧两种情况，测量整数和半像素运动向量的性能。半像素向量被采用并且使用上面给出的近似来建模。

实验表明，长期帧缓冲器的增加提高了编码器的压缩效率并且使得比特流对丢包更加鲁棒。然而，对于某些序列该效果很微小，并且其依赖于更新参数 N 。固定的 N 对于所有序列来说并不是最优的。最优化包括对于特定序列选取特定的更新参数 N 。

本发明的一个多帧编码实施例应用于速率切换网络。可以使用长期或高质量参考帧。例如，考虑无线网络经常经历网络容量上的显著跃变。产生这种跃变的示例事件是使用诸如始终最佳连接(ABC)方法的服务时的网络切换(handoff)。在这种情境下使用长期或高质量帧的双帧编码可以显著提高在紧接着网络切换之后传送的帧的质量并且对网络容量中的陡峭和剧烈的跃变起平滑作用。

在优选实施例中，指定长期过去帧为就在从高带宽向低带宽模式的网络切换之前编码的最后的帧。图 2 中对此进行了说明，在图 2 中发生从 10Mbps 到 10-20Kbps 连接的跃变。利用来自高速率连接的长期帧 $n-D$ 来对当前帧 n 编码。对于预测帧中的每个宏块(MB)，不仅在一个(或多个)紧邻过去帧而且在一个(或多个)长期过去帧上执行搜索，并且选取更佳匹配的块。

现在考虑修正的实际实施例，将图 2 的方法应用于遵照 MPEG-4 编码运行的 ABC 网络。在如 ABC 的服务下运行的视频编码器必须对多个数量级的带宽变化很鲁棒。假定 ABC 网络以最小的丢失来提供及时的分组传递。为了对抗带宽中的骤变，假定每个帧的量化参数可以在其全范围(1-31)上改变，这和将量化参数值的改变限定于在先值的 25%的符合标准的编码器形成对比。

通过修改标准的 MPEG-4 编码装置对多帧带宽切换实施例进行仿真。MPEG-4 编码装置使用 A.Vetro, H.Sun, Y.Wang, "MPEG-4 Rate Control for Multiple Video Objects", IEEE Transactions on Circuits and Systems for Video Technology, Vol.9, No.1, pp.186-99, February 1999 中所描述的速率控制方法。将每个帧视为 MPEG-4 编码器的单个对象。为长期帧分配额外的存储器。对每个间编码的 MB 传送额外的比特以通知解码器其参考哪个帧。内刷新周期(intra refresh period)设置为 100。降低内刷新周期增强了双帧编码器的性能。但是，频繁的内刷新导致更高的将超过 GPRS(General Packet Radio Service, 通用分组无线服务)系统可用的比特速率的比特速率。

测试了具有静态和动态背景以及情景变化的不同类型的视频序列。测试序列的格式是 QCIF(quarter common intermediate formate, 四分之一公用中间图像格式)。帧速率是 10 帧/秒。为调查切换到不同的低带宽网络的影响, 仿真了从 1Mbps 切换到范围从 10Kbps(GPRS)到 150Kbps(1xRTT CDMA)的低带宽网络。通过使用双帧缓冲器编码器以及为比较还通过使用常规的 MPEG-4 编码器对各种序列进行编码。发现, 如由解码后序列的 PSNR 所量化的, 保留高质量帧以用作双帧编码器的长期过去帧致使获得了在切换到较低带宽连接之后的高达几百帧的更好的视频质量。该技术在保留双参考帧的存储器方面需要较小的代价(不仅在编码器处而且在解码器处), 并且在搜索第二参考帧来寻找最佳匹配块的编码器复杂度上需要较小代价。

在一些实例中, 可以诸如通过预警或者预测而预见带宽的即将发生的切换。这提供了在较高带宽时对长期帧编码的时机。

在另一实施例中, 周期性地或者偶而地(即使在带宽或者信道/网络状况上没有变化时)取长期帧, 并且分配额外的带宽以产生作为长期帧的高质量帧。这是一种偶而地或周期性地(例如每 N 个帧)获取高质量长期参考帧的主动方法。

分配额外的比特以产生高质量参考帧从而其将具有比正被编码的常规质量更优的质量。然后该长期高质量参考帧可以在扩展的时间段期间用作高质量匹配块的源。在一个示例中, 如果发生状况改变则扩展的时间段可以超过 N 个帧。

额外的比特的分配可以来自调度器。例如, 假定存在使用由调度器分配的系统资源的多个用户。所考虑的系统是共享的无线介质如具有 400Kbps 的速率的 HDR(高数据速率)系统。调度器 S_1 在 k 个用户之间等分来自网络的可用带宽 B 。每个用户具有 B/k 的带宽。调度器 S_2 保留总带宽的某部分 b , 而将剩余的在所有用户之中划分。所保留的部分 b 将轮流分配给每个用户, 而在用户之间周期循环。每个用户在 $k-1$ 个时隙(time slot)期间具有带宽 $(B-b)/k$, 而在一个时隙期间具有带宽 $b+(B-b)/k$ 。图 3A 和 3B 代表两个调度器, 其中用高度代表带宽而水平轴是时间。其对于图 3A 的调度器 S_1 是均匀的, 但是在图 3B 中对于用户 i 则很高, 因为 S_2 中的调度器提供了额外的带宽。这用代表分配给用户 i 的额外带宽的厚黑条来显示。在不同的时间该额外的带宽被分配给系统中其它 $k-1$ 个用户。在任何时间点处所有用户上的平均带宽保持相

同。由 S_2 分配的额外的周期性带宽用于创建高质量帧，该高质量帧在多帧例如双帧运动补偿方案中用作长期帧。

5 另一实施例并不使用调度器，但是允许具有固定平均带宽的用户对于一个帧使用额外的比特(对于之前和之后的帧使用较少的比特)，如果视频中可容忍一定量额外的延迟的话。用于该帧的额外的比特简单地转化成用于传输该帧的更多的时间。存在压缩性能和延迟的折衷。在调度器实施例中，可以操作调度器使得由于周期性地提供额外的带宽所以对于高质量帧不导致额外的延迟。然而，会将总平均带宽制约为等于均匀分配的 S_1 系统的带宽。

10 特定实施例的修正 MPEG-4 编码器具有可选取的三种编码模式：内编码、使用短期参考帧的间编码、和使用长期参考帧的间编码。对于每个宏块(MB)，通过如下首先在内编码和间编码之间做出选择来在三种编码模式之中做出选择。计算短期帧或长期帧中的最佳匹配 MB 与待编码的当前 MB 之间的失真 d_{min} 。计算当前 MB 的标准偏差 σ 。如果 $\sigma < d_{min} - 512$ 则选取内编码，否则选取间编码。在根据短期和长期高质量帧的间编码之间的选择是基于运动补偿后的 MB 和待编码的当前 MB 之间的失真而做出的。选取得到较低失真的参考帧。

20 另一实施例包括使用来自解码器的反馈。作为示例，编码器可以从解码器接收指示对接收到的分组的确认的反馈。设 i 是当前帧的指数。使用具有固定延迟 d 的反馈，编码器可以对解码器的第 $(i-d)$ 个重构的帧具有良好的认识。将使用术语重新解码(re-decode)来描述编码器使用反馈信息对过去帧进行解码使其等同于该帧的解码器版本的处理。因为编码器知道哪个 GOB 被完整无缺的接收到以及哪个丢掉了，所以编码器可以精确地模拟解码器的包括误差隐蔽在内的操作。经重新解码的帧是在编码器处的与解码器版本等同的帧，然而使用术语估计来描述在编码器处的对于其反馈信息尚不可得的帧，从而强迫编码器估计解码器版本。利用反馈信息，使用上面用于内编码、inter-ST 编码、inter-LT 编码的 MB 的公式来仍然进行中间帧(intermediate frame)中像素值的估计。然而，现在关于这些公式所需的过去解码器帧的信息可以通过使用已 ACK/NACK 的经重新解码的帧来重新初始化(reinitialize)。然后编码器可以更加可靠地重新计算像素估计并且对于最后的 d 个帧跟踪潜在错误。将实际预测余量(residual)或内系数(intra coefficients)输入 ROPE 估计算法中，其中参考帧或者是同样递归计算得到的 ROPE 估计或者经重新解码的帧。

图 4 中图示了一个示例。这里，跳跃更新参数和反馈延迟分别是 $N=2$ 和 $d=5$ 。跳跃更新参数 $N=2$ 意味着帧 0 将是用于帧 2 和帧 3 的长期参考，帧 2 将是帧 4 和帧 5 的长期参考，帧 4 将用于帧 6 和帧 7。

由于 $d=5$ ，则开始对帧 7 编码时，帧 2 将被重新解码，并且该新近重新解码的帧将迅速地用于更新帧 3、4、5、6 的估计。关于对帧 7 编码，长期帧是帧 4，短期帧是帧 6，并且这两个帧的新估计将被编码器用来计算由于有关帧 7 的丢包所导致的预期失真。

一种替代的方法是使得长期帧缓冲器前进以包含最近的精确知道的帧即 $(i-d)$ 帧。反馈信息改进了 ST 帧的估计，并且将 LT 帧的估计误差减少为零。确保编码器和解码器长期帧缓冲器两者总是包含等同的重构。在延迟 d 的情况下，可以使用 $D=d$ 且 $N>1$ 的一般 (N,D) 更新策略，或者 $D=d$ 且 $N=1$ 的连续更新策略。图 5 中描述了 $N=2$ ， $d=5$ 的示例。在图 5 中，帧 12 正在被编码。它的 LT 帧是帧 7，帧 7 也是被重新解码了的。然而，对帧 7 重新解码需要帧 1 和帧 6 的经重新解码的版本，即其分别的 ST 帧和 LT 帧。现在可以获得帧 8、9、10、11 的估计。对于帧 8，将需要经重新解码的 7 和重新解码的 3。对于 9，将需要经估计的 8(ST)和经重新解码的 3(LT)。对于 10，将需要经估计的 9 和经重新解码的 5。类似地，11 需要经估计的 10 和经重新解码的 5。

通过使发射器和接收器处的长期帧缓冲器同步，可以完全消除由于丢包累积而导致的漂移误差(drift error)。如果经 Inter-LT 编码的宏块到达，则它们将在编码器和解码器处被以等同的方式重构。通常，这仅通过传输内编码的宏块来保证。但是，在这里，反馈信号使得能够使用长期帧缓冲器作为额外的误差鲁棒因子而不会很大地牺牲压缩效率。

取代使用反馈仅仅改进失真估计及进而改进模式选择，现在还可以使用该信息来在编码器处对 LT 帧重新解码并且从而改进运动估计，并且使用更真实的参考帧。实验显示在各种条件下其都表现良好。

虽然已经示出和描述了本发明的特定实施例，但是应该理解，其它修改、替换和变更对本领域普通技术人员很清楚。在不背离应该由所附权利要求确定的本发明的精神和范围的情况下，可以进行这样的修改、替换和变更。

在所附的权利要求中阐述了本发明的各种特征。

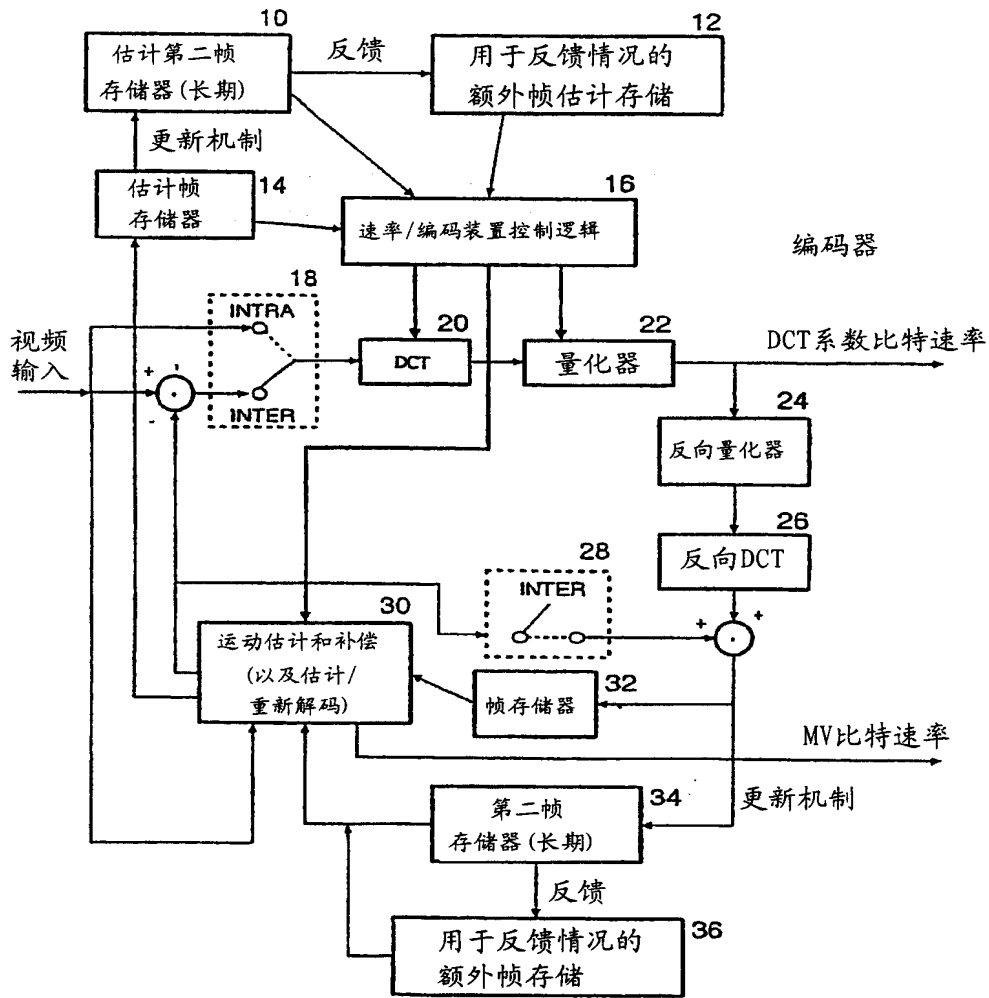


图 1A

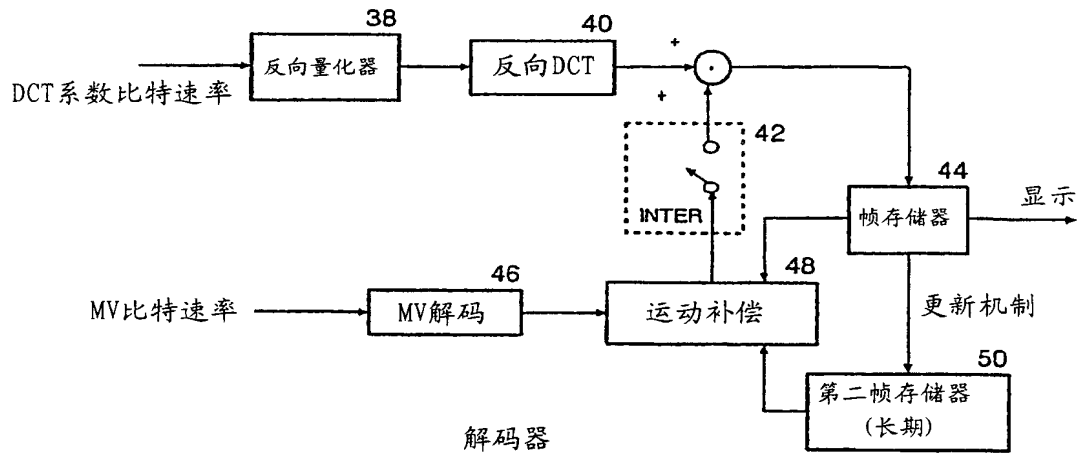


图 1B

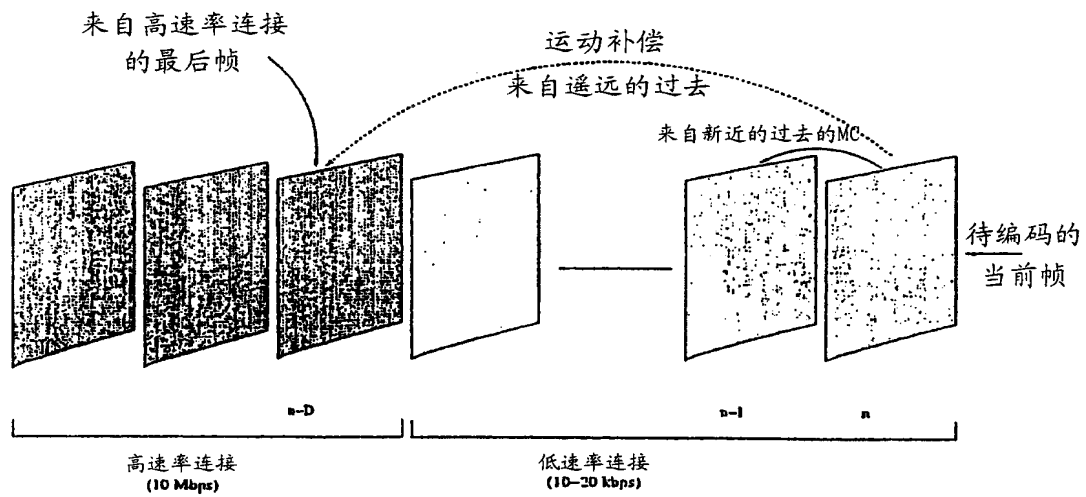


图 1: 双帧运动补偿示例

图 2

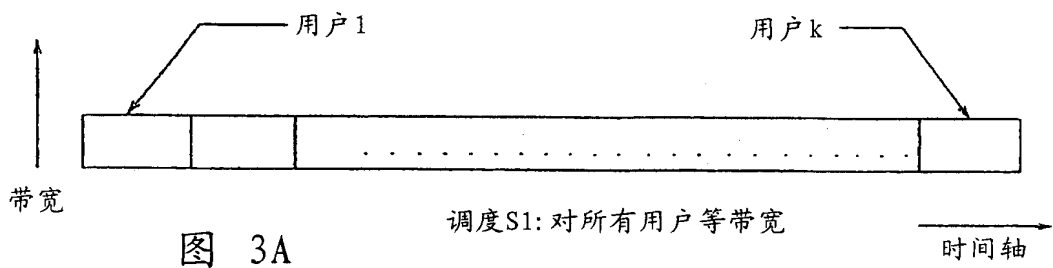
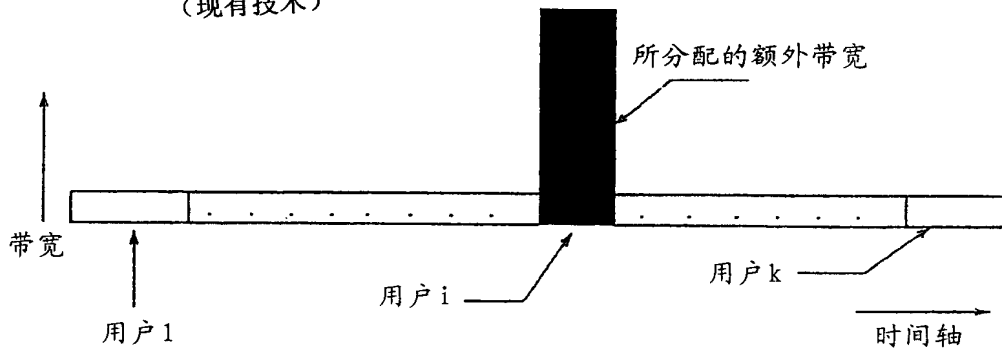


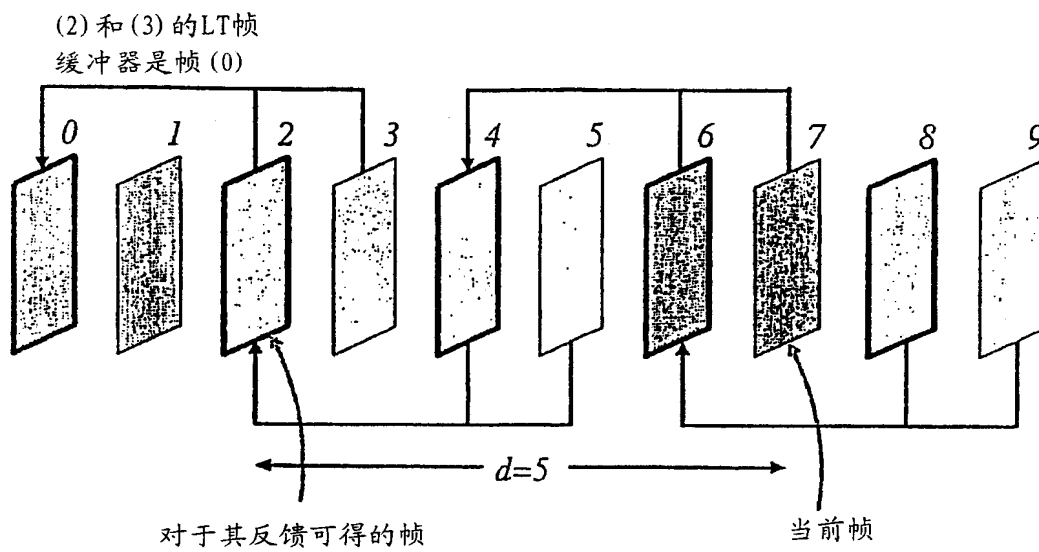
图 3A

(现有技术)



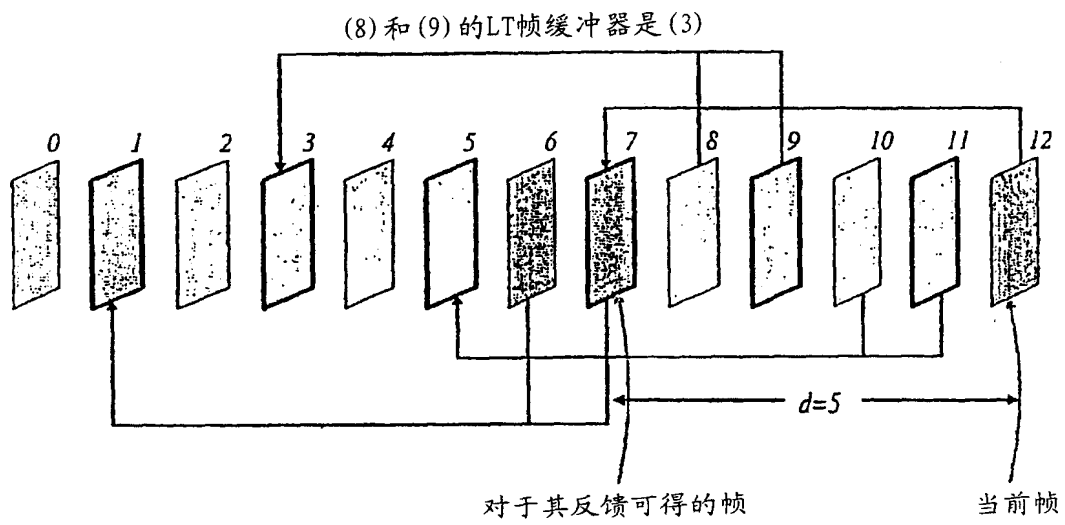
调度S2: 对N-1个用户等带宽, 对用户i额外的带宽

图 3B



参数 $N=2, D=2, d=5$ 的具有反馈的编码方法

图 4



参数 $N=2, D=5, d=5$ 的具有反馈的编码方法

图 5