

[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 说 明 书

专利号 ZL 03802209.5

[45] 授权公告日 2009 年 10 月 21 日

[11] 授权公告号 CN 100553321C

[22] 申请日 2003.1.14 [21] 申请号 03802209.5
[30] 优先权

CN 1255017 A 2000.5.31

审查员 王艳妮

[32] 2002. 1. 14 [33] US [31] 60/348,811

[74] 专利代理机构 北京市金杜律师事务所

[86] 国际申请 PCT/IB2003/000061 2003.1.14

代理人 王茂华 李 辉

[87] 国际公布 WO2003/058945 英 2003.7.17

〔85〕 进入国家阶段日期 2004.7.14

[73] 专利权人 诺基亚有限公司

地址 芬兰埃斯波

[72] 发明人 J·濑内马

[56] 参考文献

US 6122017 A 2000.9.19

CN 1159270 A 1997.9.10

US 5579054 A 1996.11.26

JP8 - 336146 A 1996. 12. 17

US 4901145 A 1990.2.13

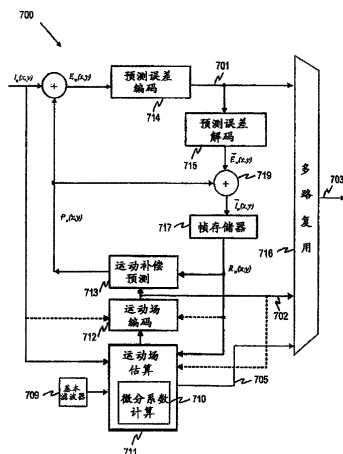
权利要求书 4 页 说明书 20 页 附图 13 页

[54] 发明名称

编码动态滤波器

[57] 摘要

一种视频编码器系统，用于编码包括一系列视频帧的视频序列，每一视频帧包括像素值阵列，其中定义内插滤波器，以便在解码处理中重构像素值。该系统相对于预定义基本滤波器差分地编码内插滤波器的系数，以提供一组差值。因为基本滤波器系数对于编码器和解码器是公知的，并且可以在统计上合理地接近在视频序列内使用的实际滤波器，因此解码器可以根据该组差值重构像素值。



1. 编码数字视频序列内的图像以提供编码视频数据的一种方法，其中数字视频序列包括一系列视频帧，每一帧具有多个像素值，以及其中将具有利用多个系数值表示的多个系数的内插滤波器用于根据编码视频数据重构所述数字视频序列的帧内的像素值，所述方法的特征在于：

相对于预定义的基本滤波器，差分地编码内插滤波器的系数值，以形成一组差值；和

将所述一组差值包括在编码视频数据内。

2. 根据权利要求1的方法，其中将编码视频数据从视频编码器发送到视频解码器，所述方法的特征在于：

编码视频数据包括表示所述一组差值的编码值，并且所述一组差值在从视频编码器发送到视频解码器之前被熵编码。

3. 根据权利要求1的方法，其特征在于，预定义基本滤波器具有多个系数，这些多个系数具有的值统计地类似于内插滤波器的系数值。

4. 根据权利要求1的方法，其特征在于，选择内插滤波器的系数用于选定图像分段内像素值的内插。

5. 根据权利要求1的方法，其特征在于，预定义基本滤波器具有固定系数值。

6. 根据权利要求1的方法，其特征在于，预定义基本滤波器具有适合于视频序列统计的多个系数。

7. 根据权利要求1的方法，其特征在于，内插滤波器是对称的，以致于仅编码一半的滤波器系数。

8. 根据权利要求1的方法，其特征在于，以从第一系数值到最后一个系数值的某一顺序，编码内插滤波器的系数值。

9. 根据权利要求8的方法，其特征在于，编码系数值的某一顺序不同于所述系数的空间顺序。

10. 根据权利要求8的方法，其特征在于，内插滤波器的系数值之和

是固定的。

11. 根据权利要求1的方法，其中预定义基本滤波器具有多个其它系数值，所述方法的特征在于，将恒定值添加到预定义基本滤波器的其它系数值，以降低内插滤波器的系数值和预定义基本滤波器的其它系数值之间差值的幅度。

12. 一种视频编码器，包括：

用于编码具有一系列视频帧的数字视频序列内的图像以提供表示视频序列的编码视频数据的装置，视频序列的每一帧包括多个像素值；和

用于定义用于在解码处理中重构所述数字视频序列的帧内的像素值的内插滤波器的装置，其中内插滤波器具有利用多个系数值表示的多个系数，所述视频编码器的特征在于：

响应于内插滤波器而用于计算所述内插滤波器和预定义基本滤波器的系数值之间的差值以提供一组差值的装置；和

用于将所述一组差值包括在编码视频数据内的装置。

13. 根据权利要求12的视频编码器，其特征还在于，

用于在编码视频数据内提供所述一组差值之前熵编码所述一组差值的装置。

14. 根据权利要求13的视频编码器，其特征在于，

内插滤波器是对称的，并且熵编码装置被安排为仅编码该一组差值的一半。

15. 解码表示包括一系列视频帧的数字视频序列的视频数据的一种方法，视频序列的每一帧包括多个像素值，其中将具有利用多个系数值表示的多个系数的内插滤波器用于重构所述数字视频序列的帧内的像素值，所述方法的特征在于：

从视频数据中检索一组差值，所述一组差值表示内插滤波器和预定义基本滤波器的系数值之间的差值；

根据所述一组差值和预定义基本滤波器，构建另一滤波器；和

根据另一滤波器，重构视频序列的帧内的像素值。

16. 根据权利要求 15 的方法, 其中预定义基本滤波器具有利用多个其它系数值代表的多个系数, 所述方法的特征还在于:

求和所述一组差值与预定义基本滤波器的其它系数值, 以构建另一滤波器。

17. 根据权利要求 16 的方法, 其特征在于, 通过熵解码从视频数据中检索所述一组差值。

18. 一种视频解码器, 包括:

用于接收比特流内的视频数据的装置, 所接收的视频数据表示包括一系列视频帧的数字视频序列, 视频序列的每一帧包括多个像素值, 所述视频解码器的特征在于:

用于从比特流中检索一组差值的装置;

用于根据预定义基本滤波器和所述一组差值构建内插滤波器的装置;

和

用于根据内插滤波器和所接收的视频数据重构视频序列的帧内的像素值的装置。

19. 根据权利要求 18 的视频解码器, 其中预定义基本滤波器具有利用其它系数值表示的多个系数, 所述视频解码器的特征还在于:

用于求和所述一组差值与预定义基本滤波器的其它系数值以构建内插滤波器的装置。

20. 根据权利要求 18 的视频解码器, 其特征还在于:

用于从比特流中熵解码所述一组差值的装置。

21. 一种视频编码系统, 包括:

编码器, 用于编码具有一系列视频帧的数字视频序列内的图像以提供表示视频序列的比特流内的编码视频数据, 视频序列的每一帧包括多个像素值, 该编码器具有用于定义用于在解码处理中重构所述数字视频序列的帧内像素值的内插滤波器的装置, 其中内插滤波器具有利用多个系数值表示的多个滤波器系数; 和

解码器, 用于接收比特流内的编码视频数据, 以便在解码处理中重构

视频序列的帧内的像素值，所述视频编码系统的特征在于：

该编码器还包括：

用于计算内插滤波器和预定义基本滤波器之间的差以提供一组差值的装置；和

用于将所述一组差值包括在比特流内的装置；和

该解码器包括：

用于从比特流中检索所述一组差值的装置；和

用于根据预定义基本滤波器和检索的一组差值构建另一滤波器的装置，以致于在解码处理中视频序列的帧内的像素值的重构基于另一滤波器。

编码动态滤波器

本申请基于于 2002 年 1 月 14 日提交的美国临时专利申请 60/348,811，并要求其优先权。

技术领域

本发明涉及视频编码中的运动补偿。更具体地说，本发明涉及一种用于编码在运动补偿视频编码器和解码器中用于重构图像像素的内插滤波器的系数的方法。本发明还涉及实现本发明方法的相应视频编码器、视频解码器和视频传输系统。

背景技术

当今存在许多种视频编码标准。这些标准包括国际电信联盟电信标准化部门 (ITU-T) 建议 H.263 和国际标准化组织 (ISO) 的运动图像专家组 (MPEG) 标准 MPEG-1、MPEG-2 和 MPEG-4。这些视频编码标准基于运动补偿预测和预测误差编码的使用。通过分析并编码在视频序列内连续帧之间的运动，并使用运动信息重构图像块，执行运动补偿预测。使用运动内插滤波器建立图像块的重构，所述滤波器能够生成所需要的用于像素和子像素位置的图像（像素）值。在下面的段落中更详细地描述使用内插滤波器的运动补偿预测和图像重构的基本原理。

类似于记录在胶片上的普通运动图像的数字视频序列包括一系列的静止图像，通常称作“帧”。通过在相对快的速率上通常是每秒 15 至 30 帧，逐帧地显示来建立运动的错觉。因为相对快的帧速率，连续帧的图像内容趋于非常类似，并因而连续帧包含大量的冗余信息。

数字视频序列的每一帧包含图像像素阵列。在通常使用的称作 Quarter Common Interchange Format（四分之一普通交换格式）(QCIF) 的数字视频格式中，一帧包括 176×144 个像素的阵列，并因而每帧包含 25344 个像素。该帧的每个像素利用一定数量的比特来表示，这些比特承载关于对应于该像素的图像区域的亮度和/或色彩内容（色度）的信息。通常，使用所谓的 YUV 彩色模型来表示图像的亮度和色度内容。亮度或 Y 分量代表图像的强度（明亮度），

而图像彩色内容利用标记为 U 和 V 的两个色度分量来表示。

基于图像内容的量度/色度表示的彩色模型与基于包含主要颜色（即红、绿和蓝，RGB）的表示的彩色模型相比提供了某些优点。因为人类视觉系统对强度变化比对色彩变化更为敏感，YUV 彩色模型通过将比亮度分量（Y）更低的空间分辨率用于色度分量（U,V）来利用了这个属性。这样，可以以图像质量的微小降低为代价来降低编码图像内色彩信息需要的信息量。

通常通过空间二次抽样获取色度分量的较低空间分辨率。通常，使用代表亮度信息的一块 16×16 值编码一块 16×16 图像像素，而两个色度分量分别用一块 8×8 值表示，所述值表示等效于 16×16 阵列的亮度值的图像区域。因而，在水平和垂直方向上以因子 2 在空间上二次抽样色度分量。所获得的一块 16×16 亮度块和两个 8×8 色度块的组合（assembly）通常称作 YUV 宏块或简称为宏块。

QCIF 图像包括 11×9 个宏块。如果用 8 比特分辨率（即通过 0 至 255 范围内的值）表示亮度块和色度块，则每宏块需要的比特总数是 $(16 \times 16 \times 8) + 2 \times (8 \times 8 \times 8) = 3072$ 比特。因而，使用每分量 8 比特数分辨率以 QCIF 格式表示视频帧需要的比特数是 $99 \times 3072 = 304128$ 比特。因此，发送、记录或显示以每秒 30 帧的速率包括一系列这样的 QCIF 格式帧的视频序列需要的数据量超过 9Mbps（每秒百万比特）。在视频记录、传输和显示中使用这个数据速率是不实际的，因为需要非常大的存储容量、传输信道容量和硬件性能。因此，已经开发了诸如上述的视频编码标准，以便在保持可接受的图像质量的同时降低表示和发送视频数据需要的信息量。

先前所述的视频编码标准中的每个标准分别为具有不同特性的视频记录或传输系统内的应用而定制。例如，ISO MPEG-1 标准被特别设计用于其中可用数据带宽高达大约 1.5M 比特/秒的情况。MPEG-2 视频编码标准主要可应用于具有可用数据带宽高达大约 10M 比特/秒的数字存储介质和视频广播与通信。ITU-T 推荐标准 H.263 预定用于可用带宽通常非常低的系统中。它尤其适合于用于将在诸如 ISDN（综合业务数字网）或常规 PSTN（公用业务电话网）等的固定线路网络上实时发送视频数据的情况中，其中可用数据传输带宽通常在 64k 比特/秒的量级上。在移动视频电话中，其中传输至少部分地发生在无线电通信链路上，可用带宽可以低到 20k 比特/秒。

尽管当前存在的各种视频编码标准为不同情况中的使用而定制，但是它们用于降低将要发送的信息量的机制具有许多共同特性。具体而言，它们全部如此工作以降低在将要发送的视频序列内的冗余和感觉上无关的信息量。在视频序列内存在基本上三种冗余度：空间、时间和频谱冗余度。空间冗余度是用于描述在一系列的各个帧内相邻像素之间的相关性的项。时间冗余度表示在一系列的帧内出现的目标可能在随后帧内出现的事实。频谱冗余度是指在同一图像的不同色彩分量之间的相关性。

通过简单地降低在给定图像序列内各种形式的冗余度，通常不能实现足够有效的压缩。因而，最新的编码器也降低了那些主观上最不重要的视频序列部分的质量。此外，利用高效的无损编码自身也降低了压缩视频比特流的冗余度。通常，这使用熵编码来实现。

运动补偿预测是一种形式的时间冗余降低，其中通过跟踪帧之间图像的目标或区域的运动，根据序列内的其它帧来“预测”视频序列内一些（通常很多）帧的内容。使用运动补偿预测压缩的帧通常称作 INTER-编码帧或 P-帧，而将不使用运动补偿预测压缩的帧称作 INTRA-编码帧或 I-帧。预测（运动补偿，INTER-编码）图像很少精确地足以以足够的质量表示图像内容，并因此空间补偿预测误差（PE）帧也与每个 INTER 帧相关。许多视频压缩方案也使用双向预测帧，这通常被称为 B-图像或 B-帧。将 B-图像插入在参考或所谓的“锚（anchor）”图像对（I 或 P 帧）之间，并根据锚图像之一或两者进行预测。

在附图的图 3 中图示了在典型的压缩视频序列内出现的不同类型的帧。如从图中可以看出的，该序列开始于 INTRA 或 I 帧 30。在图 3 中，箭头 33 表示“前向”预测处理，通过该处理形成 P-帧 34。用于形成 B-帧 36 的双向预测处理分别用箭头 31a 和 31b 来表示。

在图 1 和图 2 中图示使用运动补偿预测的通用视频编码系统的示意图。图 1 图示使用运动补偿预测的编码器 10，图 2 图示相应的解码器 20。图 1 所示的编码器 10 包括运动场估算块 11、运动场编码块 12、运动压缩预测块 13、预测误差编码块 14、预测误差解码块 15、多路复用块 16、帧存储器 17 和加法器 19。解码器 20 包括运动补偿预测块 21、预测误差解码块 22、多路分用块 23 和帧存储器 24。

使用运动补偿预测的视频编码器的操作原理是最小化预测误差帧 $E_n(x,y)$ 内

的信息量,所述预测误差帧 $E_n(x,y)$ 是正在被编码的当前帧 $I_n(x,y)$ 和预测帧 $P_n(x,y)$ 之间的差值。因而,将预测误差帧定义为如下:

$$E_n(x,y) = I_n(x,y) - P_n(x,y) \quad (1)$$

使用参考帧 $R_n(x,y)$ 的像素值建立预测帧 $P_n(x,y)$, 所述参考帧 $R_n(x,y)$ 通常是先前编码并发送的一个帧,例如恰在当前帧之前的帧,并可以从编码器 10 的帧存储器 17 中获得。更具体地说,通过查找参考帧 $R_n(x,y)$ 内的“预测像素”来构建预测帧 $P_n(x,y)$, 所述“预测像素”基本上对应于当前帧内的像素。导出描述当前帧内的像素和其在参考帧内相应的预测像素之间的关系(例如相对位置、旋转、定标等等)的运动信息,并通过根据运动信息移动预测像素来构建预测帧。这样,使用参考帧内的像素值,将预测帧构建为当前帧的近似表示。因此,上面提到的预测误差帧表示由预测帧提供的当前帧的近似表示和当前帧本身之间的差值。使用运动补偿预测的视频编码器所提供的基本优点是因为这样的事实,即通过形成其预测需要的运动信息以及在预测误差帧内的相关预测误差信息,可以获得当前帧的相对简洁的描述。

由于帧内的大量像素,将每个像素的单独运动信息发送给解码器通常是低效的。相反,在大多数视频编码方案中,将当前帧划分成较大的图像分段 S_k , 并将与这些分段相关的运动信息发送给解码器。例如,通常为一帧的每个宏块提供运动信息,并且随后将相同的运动信息用于该宏块内的所有像素。在一些视频编码标准中,例如在当前正在开发的 ITU-T 推荐标准 H.26L 中,能够将宏块划分成更小的块,向每个更小的块提供它自己的运动信息。

运动信息通常采用运动矢量 $[\Delta x(x,y), \Delta y(x,y)]$ 的形式。数值 $\Delta x(x,y)$ 和 $\Delta y(x,y)$ 对代表在当前帧 $I_n(x,y)$ 内的像素 (x,y) 相对于参考帧 $R_n(x,y)$ 内像素的水平和垂直位移。在运动场估算块 11 内计算运动矢量 $[\Delta x(x,y), \Delta y(x,y)]$, 并将当前帧的运动矢量组 $[\Delta x(\cdot), \Delta y(\cdot)]$ 称作运动矢量场。

典型地,通过其左上角的 (x,y) 坐标来指定当前视频帧内宏块的位置。因而,在运动信息与一帧的每个宏块相关的视频编码方案中,每个运动矢量描述表示当前帧 $I_n(x,y)$ 内宏块的左上角的像素相对于参考帧 $R_n(x,y)$ 内预测像素的基本上对应块的左上角内一个像素的水平和垂直位移 $\Delta x(x,y)$ 和 $\Delta y(x,y)$ (如图 4b 所示)。

运动估算是需要大量计算的工作。给定参考帧 $R_n(x,y)$ 和例如包括在当前

帧内 $N \times N$ 像素的正方形宏块（如图 4b 所示），运动估算的目的是查找根据某一标准匹配当前图片内宏块特性的参考帧内的 $N \times N$ 像素块。这个标准例如可以是当前帧内宏块像素和与之比较的参考帧内像素块之间的绝对差值之和（SAD）。通常将这个处理称作“块匹配”。应当指出，通常，将要匹配的块的几何结构和参考帧内的几何结构不必相同，因为实际情况下的目标可能经历比例变化以及旋转和翘曲。然而，在当前的国际视频编码标准中，例如上面提到的，仅使用平移运动模型（参见下文），并因此固定的矩形几何结构就足够了。

理想上，为了实现找到匹配的最佳机会，应当搜索整个参考帧。然而，这是不实际的，因为这在视频编码器上施加了过高的计算负担。相反，搜索区域通常限于围绕当前帧内宏块的原始位置的一个区域 $[-p,p]$ ，如图 4c 所示。

为了进一步降低将要从编码器 10 发送给解码器 20 的运动信息量，通过使用运动模型表示，在编码器 10 的运动场编码块 12 内编码运动矢量场。在这一处理中，使用某些预先确定的函数来重新表示图像分段的运动矢量，或者换句话说，使用模型表示运动矢量场。几乎所有当前使用的运动矢量场模型是加法运动模型，符合下述通用公式：

$$\Delta x(x, y) = \sum_{i=0}^{N-1} a_i f_i(x, y) \quad (2)$$

$$\Delta y(x, y) = \sum_{i=0}^{M-1} b_i g_i(x, y) \quad (3)$$

其中 a_i 和 b_i 是运动系数。将这些运动系数发送给解码器 20（图 1 和图 2 中的信息流 2）。函数 f_i 和 g_i 是运动场基本函数。它们对于编码器和解码器是公知的。使用这些系数和基本函数可以构建近似的运动矢量场 $(\tilde{\Delta x}(x, y), \tilde{\Delta y}(x, y))$ 。因为基本函数是编码器 10 和解码器 20 所知的（即存储在其中），仅需要将运动系数发送给编码器，从而降低表示帧的运动信息所需要的信息量。

最简单的运动模型是平移运动模型，它仅需要两个系数来描述每个分段的运动矢量。通过下式给出运动矢量的值：

$$\begin{aligned} \Delta x(x, y) &= a_0 \\ \Delta y(x, y) &= b_0 \end{aligned} \quad (4)$$

这是在 ITU-T 推荐标准 H.263 和 ISO 标准 MPEG-1、MPEG-2、MPEG-4 中用于描述 16×16 和 8×8 像素块的运动的模型。使用平移运动模型的系统通

常在全像素分辨率或者全像素分辨率的某一整数分数上例如在二分之一或四分之一像素分辨率上执行运动估算。

在编码器 10 的运动补偿预测块 13 内构建预测帧 $P_n(x,y)$ ，并且通过下式给出：

$$P_n(x,y) = R_n[x + \tilde{\Delta}x(x,y), y + \tilde{\Delta}y(x,y)] \quad (5)$$

在预测误差编码块 14 内，通常通过将其表示为某些二维函数的有限级数（变换）来压缩预测误差帧 $E_n(x,y)$ 。例如，可以使用二维离散余弦变换（DCT）。在将变换系数发送给解码器（图 1 和图 2 中的信息流 1）之前量化和熵（例如哈夫曼）编码这些变换系数。因为量化引入误差，该操作通常在预测误差帧 $E_n(x,y)$ 内生成某种劣化（信息损失）。为了补偿此劣化，编码器 10 还包括预测误差解码块 15，其中使用变换系数构建解码的预测误差帧 $\tilde{E}_n(x,y)$ 。由加法器 19 将这个本地解码的预测误差帧添加给预测帧 $P_n(x,y)$ ，并将所获得的解码的当前帧 $\tilde{I}_n(x,y)$ 存储在帧存储器 17 内，以便进一步用作下一个参考帧 $R_{n+1}(x,y)$ 。

在多路复用器 16 内组合承载运动矢量相关信息的信息流 2 与有关预测误差的信息，并将通常至少包含这两种类型信息的信息流 3 发送给解码器 20。

现在将描述相应视频解码器 20 的操作。

解码器 20 的帧存储器 24 存储先前重构的参考帧 $R_n(x,y)$ 。使用所接收的运动系数信息和先前重构的参考帧 $R_n(x,y)$ 的像素值，根据等式 5 在解码器 20 的运动补偿预测块 21 内构建预测帧 $P_n(x,y)$ 。在预测误差解码块 22 内使用发送的预测误差帧 $E_n(x,y)$ 的变换系数来构建解码的预测误差帧 $\tilde{E}_n(x,y)$ 。随后，通过相加预测帧 $P_n(x,y)$ 与解码的预测误差帧 $\tilde{E}_n(x,y)$ 来重构解码的当前帧 $\tilde{I}_n(x,y)$ 的像素：

$$\tilde{I}_n(x,y) = P_n(x,y) + \tilde{E}_n(x,y) = R_n[x + \tilde{\Delta}x(x,y), y + \tilde{\Delta}y(x,y)] + \tilde{E}_n(x,y) \quad (6)$$

将这个解码的当前帧作为下一个参考帧 $R_{n+1}(x,y)$ 存储在帧存储器 24 内。

在上面的数字视频的运动补偿编码和解码的描述中，描述当前帧内的宏块相对于参考帧 $R_n(x,y)$ 的运动运动矢量 $[\Delta x(x,y), \Delta y(x,y)]$ 可以指向参考帧内的任一像素。这意味着可以仅在由帧内的图像像素确定的分辨率（所谓全像素分辨率）上表示数字视频序列的帧之间的运动。然而，实际的运动具有任意的精确度，并因而上面描述的系统仅提供了数字视频序列的连续帧之间运动的近似模

型化。典型地，具有全像素分辨率的视频帧之间运动的模型化对于允许有效地最小化与每个宏块或帧相关的预测误差（PE）信息是不够精确的。因此，为了能够更精确地模型化实际运动和有助于降低必需从编码器发送给解码器的 PE 信息量，许多视频编码标准允许运动矢量指向“之间（in between）”图像像素。换句话说，运动矢量可以具有“子像素”分辨率。允许运动矢量将子像素分辨率增加到必须执行的编码和解码操作的复杂性，所以限制运动矢量可以具有的空间分辨率依然是有利的。因而，诸如先前所述的视频编码标准通常仅允许运动矢量具有全、半或四分之一像素分辨率。

具有子像素分辨率的运动估算可以实施为两阶段处理，如在图 5 中以示例性方式图示的，用于其中运动矢量可以具有全或半像素分辨率的通用视频编码方案。在第一步骤中，使用适当的运动估算方案例如上面描述的块匹配处理来确定具有全像素分辨率的运动矢量。所获得的具有全像素分辨率的运动矢量在图 5 中图示。

在第二阶段中，进一步精化在第一阶段内确定的运动矢量以获得所希望的半像素分辨率。在图 5 所示的例子中，这通过形成八个新的 16×16 像素的搜索块来进行，在图 5 中利用 X 标记每块的左上角的位置。将这些位置表示为 $[\Delta x + m/2, \Delta y + n/2]$ ，其中 m 和 n 可以采用值 -1 、 0 和 $+1$ ，但是不能同时为零。因为仅有原始图像像素的像素值是已知的，所以使用某种形式的内插方案为八个新搜索块的每个搜索块估算位于半像素位置上的子像素的值（例如，亮度和/或色度值）。

如果已经以半像素分辨率内插子像素的值，则比较八个搜索块中的每个块与寻找其运动矢量的宏块。当执行块匹配处理以确定具有全像素分辨率的运动矢量时，根据诸如 SAD 的某一标准比较宏块与八个搜索块中的每一块。作为比较结果，通常将获得最小 SAD 值。根据视频序列内运动的属性，这个最小值可以对应于由原始运动矢量（具有全像素分辨率）指定的位置，或者它可以对应于具有半像素分辨率的位置。因而，有可能确定运动矢量是否应当指向全像素或半像素位置，或者如果子像素分辨率合适的话，确定正确的子像素分辨率运动矢量。

实际上，通过从周围的像素值内插子像素的值来执行参考帧内子像素值的估算。通常，可以将位于非整数位置 $(x, y) = (n + \Delta x, m + \Delta y)$ 的子像素值 $F(x, y)$ 的

内插格式化二维操作，在数学上表示为：

$$F(x, y) = \sum_{k=-K}^{K-1} \sum_{l=-L}^{L-1} f(k+K, l+L) F(n+k, m+l) \quad (7)$$

其中 $f(k,l)$ 是滤波系数，而 n 和 m 分别通过将 x 和 y 截取到整数值来获得。典型地，滤波系数取决于 x 和 y 值，并且内插滤波器通常是所谓的“可分离的滤波器”，在这种情况下可以将子像素值 $F(x,y)$ 计算如下：

$$F(x, y) = \sum_{k=-K}^{K-1} f(k+K) \sum_{l=-L}^{L-1} f(l+L) F(n+k, m+l) \quad (8)$$

在编码器中计算运动矢量。一旦将相应的运动系数发送给解码器，直接的方法是使用与在编码器中使用的方法相同的内插方法来内插所需要的子像素。这样，可以根据参考帧和所发送的运动矢量重构帧存储器 24 内在参考帧之后的帧。

常规上，在视频编码器和解码器内使用的内插滤波器使用固定的滤波器系数值，并将相同的滤波器（即具有相同滤波器系数值的相同类型滤波器）用于所编码的视频序列的所有帧。可以将同一滤波器进一步用于所有的视频序列，而与它们的特性和如何获取（捕获）它们无关。Wedi（“用于运动补偿混合视频编码的自适应内插滤波器”，图像编码研讨会（PCS 2001），韩国汉城，2001 年 4 月）推荐使用具有自适应滤波器系数值的内插滤波器来补偿视频编码处理中的某些缺陷。具体而言，Wedi 描述图像获取处理中的混叠、所允许的运动矢量的有限分辨率和平移运动模型的有限有效性如何引入附加预测误差。视频图像中的混叠是由于图像获取处理中使用非理想的低通滤波器（和结果不满足 Nyquist 抽样理论）引起的。混叠干扰视频序列内的运动补偿预测并引起附加的预测误差分量。允许的运动矢量（例如，全像素，二分之一像素或四分之一像素）的有限精确度和平移运动模型只表示连续视频帧之间的水平与垂直平移运动的能力也引起附加的预测误差贡献。Wedi 还推荐了通过修改（adapt）内插滤波器的滤波器系数值以补偿由混叠、有限运动矢量精度和平移运动模型的有限有效性引入的附加预测误差而能够实现编码效率的改进。

更普通地，应当理解因为视频序列内运动变化的性质和特性，最佳内插滤波器作为时间和图像位置的函数而变化。Wedi 介绍了一个例子，其中将具有动态自适应滤波器系数值的内插滤波器集成到 H.26L 视频编译码器中，更具体

地说, 由 H.26L 的测试模型 (TML) 4.TML-4 定义的编译码器使用四分之一像素运动矢量分辨率和带有六个对称滤波器系数的维纳型 (Wiener-type) 内插滤波器 (6-抽头滤波器) 的版本。Wedi 介绍的例子推荐逐帧地修改内插滤波器的滤波器系数, 差分地编码滤波器系数, 并将它们作为主视频数据的辅助 (side) 信息发送给解码器。基于这种方法的建议包括在 H.26L 视频编译码器的测试模型 8 内使用带有动态自适应滤波器系数值的内插滤波器。这在 ITU-电信标准化部门的标题为“用于 H.26L 的自适应内插滤波器”, 研究组 16, 问题 6, 视频编码专家组 (VCEG), 文件 VCEG-N28, 2001 年 9 月和“用于 H.26L 的自适应内插滤波器的更多结果”, 研究组 16, 问题 6, 视频编码专家组 (VCEG), 文献 VCEG-O16rl, 2001 年 11 月中进行了介绍。

动态自适应内插滤波器的使用提出了与编码视频数据流的编码效率相关的重要问题, 并且还对编码视频数据的错误复原能力具有影响。能够以一种直接的方式理解编码效率的问题。在使用具有固定滤波器系数值的内插滤波器的视频编码系统中, 不需要包括任何与编码视频数据比特流内的滤波器系数值相关的任何信息。滤波器系数值可以简单地记录在视频编码器和视频解码器内。换句话说, 在根据使用固定内插滤波器的特定视频编码标准实现的视频编码系统内, 根据标准规范将系数值预先编程到编码器和解码器。然而, 如果允许动态的自适应滤波器系数, 则变得必需发送与这些系数值相关的信息。因为定期地更新滤波器系数 (例如在逐帧的基础上), 这必然添加到从视频编码器发送给解码器的信息量, 并对编码效率产生有害的影响。在低比特率视频编码应用中, 需要发送的信息量的任何增加通常都是不希望的。

因而, 为了最佳模型化和补偿运动, 需要动态内插滤波器的有效表示方法。

关于错误复原能力, 应当理解将与动态可变内插滤波器的系数相关的信息从编码器发送给解码器的方式可能影响视频数据对传输误差的敏感性。更具体地说, 在使用动态自适应内插滤波器的视频编码系统内, 在解码器上视频序列帧的正确重构依靠于滤波器系数值的正确接收和解码。如果与系数值相关的信息在它从编码器到解码器的发送期间经受错误, 很可能出现重构的视频数据的恶化。在现有技术中存在三种编码滤波器系数的方式。第一种是分别地熵编码滤波器系数值。第二种是相对于已解码滤波器的滤波器系数差分地编码滤波器系数值 (如在 Wedi 内推荐的), 第三种是定义一组滤波器并编码选定滤波器的

索引。

如上面所介绍的，可用于编码内插滤波器系数的现有技术解决方案在不同的使用情况下全都具有与之相关的问题。第一种方法，其中单独地编码内插滤波器系数，提供了较差的编码性能，因为它并不使用任何先验信息（即与先前编码的内插滤波器系数值相关的信息）。因此，这种方法需要添加给编码视频比特流的过多的信息以描述内插滤波器系数值。如在 Wedi 所推荐的系数的差分编码是有效的，但是不能在存在可能的传输错误的环境下使用，因为滤波器系数取决于先前滤波器系数的正确解码。如先前描述的，如果编码的视频比特流在从编码器到解码器的传输期间经受错误，则可能出现在解码器上重构视频数据的破坏。第三种使用预定义滤波器组的现有技术解决方案仅提供有限的选择，并因而降低了编码性能。换句话说，这种选择不能实现使用带有动态自适应滤波器系数的内插滤波器的全部优点，如在 Wedi 中阐述的。

因而，应当认识到需要一种有效的并且并不导致编码视频比特流的错误复原能力恶化的编码自适应内插滤波器的系数值的方法。

发明内容

本发明组合差分编码的好的编码效率与允许在所有环境下使用的错误复原能力特性。因此，特别适合于在易于出现错误的环境内使用的视频编码系统内实现，例如编码的视频比特流将在容易遭受干扰的无线电通信链路上传输。

因而，根据本发明的一个方面，提供一种编码数字视频序列内的图像以提供编码视频数据的方法，其中数字视频序列包括一系列的视频帧，每帧包含多个像素值，以及其中将具有利用多个系数值代表的多个系数的内插滤波器用于根据编码视频数据重构所述数字视频序列的帧内的像素值，该方法的特征在于：

相对于预定义的基本滤波器（base filter）差分地编码内插滤波器的系数值，以形成一组差值；和

修改编码视频数据内的所述一组差值，以便像素值的重构基于所述一组差值。

有利地，编码的视频数据包括表示所述一组差值的编码值，并且在从视频编码器发送给视频解码器之前熵编码所述一组差值。

有利地，预定义基本滤波器具有多个系数，这些系数具有的值统计地类似

于内插滤波器的系数值。

有利地，选择内插滤波器的系数用于选定图像分段内像素值的内插。

有利地，预定义基本滤波器具有固定的系数值。

有利地，预定义基本滤波器具有适合于视频序列统计的多个系数。

最佳地，内插滤波器是对称的，以便仅编码一半的滤波器系数。

有利地，以从第一系数值到最后一个系数值的某一顺序编码内插滤波器的系数值，并且该某一顺序不同于所述系数的空间顺序。

有利地，内插滤波器的系数值之和是固定的。

有利地，预定义基本滤波器具有多个系数值，并且将恒定值添加到预定义基本滤波器的系数值，以降低内插滤波器的系数值和预定义基本滤波器的系数值之间差值的幅度。

根据本发明的第二方面，提供一种视频编码器，包括：

用于编码包括一系列视频帧的数字视频序列内的图像以提供表示视频序列的编码视频数据的装置，视频序列的每一帧包括多个像素值；和

用于定义用于在解码处理中重构所述数字视频序列的帧内像素值的内插滤波器的装置，其中内插滤波器具有利用多个系数值表示的多个系数。该视频编码器的特征在于：

响应于内插滤波器而用于计算所述内插滤波器和预定义基本滤波器的系数值之间的差值以提供一组差值的装置；和

用于修改编码视频数据内的所述一组差值的装置，以便解码处理内像素值的重构基于所述一组差值。

有利地，该编码器包括用于在修改编码视频数据内的所述一组差值之前熵编码所述一组差值的装置。

根据本发明的第三方面，提供一种解码表示包括一系列视频帧的数字视频序列的视频数据的方法，视频序列的每一帧包括多个像素值，其中将具有利用多个系数值表示的多个系数的内插滤波器用于重构所述数字视频序列的帧内的像素值。该方法的特征在于：

从视频数据中检索一组差值，所述一组差值表示内插滤波器和预定义基本滤波器的系数值之间的差值；

根据所述一组差值和预定义基本滤波器构建另一滤波器；和

根据另一滤波器重构像素值。

有利地，预定义基本滤波器具有利用多个系数值代表的多个系数，和通过求和所述一组差值与预定义基本滤波器的系数值来执行另一滤波器的构建。

有利地，通过熵解码从视频数据中检索所述一组差值。

根据本发明的第四方面，提供一种视频解码器，包括用于接收比特流内的视频数据的装置，所接收的视频数据表示包括一系列视频帧的数字视频序列，视频序列的每一帧包括多个像素值。该视频解码器的特征在于：

用于从比特流中检索一组差值的装置；

用于根据预定义基本滤波器和所述一组差值构建内插滤波器的装置；和

用于根据内插滤波器和所接收的视频数据重构视频序列的帧内的像素值的装置。

有利地，视频解码器还具有用于求和所述一组差值与预定义基本滤波器的其它系数值以构建内插滤波器的装置，和用于从比特流中熵解码所述一组差值的装置。

根据本发明的第五方面，提供一种视频编码系统，包括：

编码器，用于编码具有一系列视频帧的数字视频序列内的图像，以提供表示视频序列的比特流内的编码视频数据，视频序列的每帧包括多个像素值，该编码器包括用于定义用于在解码处理中重构所述数字视频序列的帧内像素值的内插滤波器的装置，其中内插滤波器具有利用多个系数值表示的多个滤波器系数；和

解码器，用于接收比特流内的编码视频数据，以便在解码处理中重构视频序列的帧内的像素值。该视频编码系统的特征在于：

该编码器还包括：

用于计算所述内插滤波器和预定义基本滤波器之间的差值以提供一组差值的装置；和

用于修改比特流内的所述一组差值的装置；和

该解码器包括：

用于从比特流中检索所述一组差值的装置；和

用于根据预定义基本滤波器和所检索的一组差值构建另一滤波器的装置，以便解码处理中像素值的重构基于另一滤波器。

参考下述描述并结合附图，本发明的这些和其它特征将变得显而易见。然而，应当理解仅仅为了说明的目的而非作为本发明限制的发明设计了这些附图。

附图说明

图 1 是根据现有技术的通用视频编码器的方框图；
图 2 是根据现有技术的通用视频解码器的方框图；
图 3 是图示在视频编码内使用的帧类型的示意图；
图 4a 是图示当前帧内宏块的示意图；
图 4b 是图示用于块匹配的参考帧的示意图；
图 4c 是图示围绕当前帧内宏块的原始位置的搜索区域的示意图；
图 5 是图示根据现有技术的运动估算到子像素分辨率的处理的示意图；
图 6a 是图示最佳内插滤波器的示意图；
图 6b 是图示分解成基本滤波器和差值系数的最佳内插的示意图；
图 6c 是图示将被编码和被发送给解码器的差值系数的示意图；
图 7 是图示能够执行本发明的包括视频编码和解码设备的终端设备的方框图；

图 8a 是图示根据本发明优选实施例的视频编码器的方框图；
图 8b 是图示根据本发明另一实施例的视频编码器的方框图；
图 8c 是图示根据本发明又一实施例的视频编码器的方框图；
图 9a 是图示根据本发明优选实施例的视频解码器的方框图；
图 9b 是图示根据本发明另一实施例的视频解码器的方框图；
图 9c 是图示根据本发明又一实施例的视频解码器的方框图。

具体实施方式

根据本发明的编码器相对于预定义基本滤波器系数差分地编码滤波器系数。图 6a 至图 6c 图示了根据本发明的方法。在图 6a 中的条形 (bar) 图表示内插滤波器系数值，每个条形对应于滤波器系数之一。条形的高度表示相应的系数值，延伸到水平轴之上的条形表示正系数值，延伸到水平轴之下的条形表示负系数值。参见图 6a 和图 6b，条形图 110 表示编码器检测为最适合于选定图像分段的运动内插的滤波器，而条形图 140 表示基本滤波器。在图 6a 图示的例子中，滤波器是具有 6 个滤波器系数的 6 抽头对称滤波器。并不发送这样

的滤波器系数本身，仅编码和发送选定滤波器 110 和基本滤波器 140 之间的差值 130。发送的系数 120 在图 6c 中图示。

使用本发明，因为可以使用熵编码器有效地编码小幅度差值，所以获得编码增益。当这样的差值包括在由视频编码器生成的编码视频比特流内，并将视频比特流从编码器发送给相应的解码器时，通过从编码比特流检索差值，并将它们添加到在解码器内存储的预定义基本滤波器的系数值，可以在解码器上重构内插滤波器的系数。

应当指出，基本滤波器还适合于视频序列和所接收的滤波器系数的统计，以进一步改善编码效率。还有可能为整个编译器预定义基本滤波器。换句话说，同一预定义基本滤波器用于将要编码的所有视频序列，而与它们的特性或者获取它们的方式无关。可选择地，基本滤波器适合于视频数据，即不同的基本滤波器用于不同的视频序列，或者在根据某些预定义规则编码特定视频序列期间可以修改基本滤波器。

如果滤波器是对称的，如图 6a 至 6c 所示，仅需要编码一半的滤波器系数。通过拷贝可以获得其它的。在图 6c 图示的例子中，通过实现第四、第五和第六滤波器系数分别等同于第三、第二和第一滤波器系数，可以进一步地降低在编码视频比特流内表示自适应内插滤波器的系数值需要的信息量。因而，在这种情况下，通过三个值实际上可以编码内插滤波器的六个系数，第一个值表示第一内插滤波器系数和预定义基本滤波器的第一系数之间的差值，第二个值表示第二内插滤波器系数和预定义基本滤波器的第二系数之间的差值，而第三个值表示第三内插滤波器系数与预定义基本滤波器的第三系数之间的差值。因此，仅需要在从编码器发送给解码器的编码视频比特流内包括这三个差值，因为通过适当地拷贝前三个重构滤波器系数值，解码器可以获得其余的三个内插滤波器系数。如果基本滤波器和内插滤波器具有奇数个而非偶数个系数，但是也是对称的，则也可以采用类似的方法。在这种情况下，应当理解，将要编码的差值个数是 $(n/2)+1$ ，其中 n 是基本滤波器/内插滤波器内的系数的数量。

根据本发明的方法还可以与其它系数编码方法组合。例如，可以定义并用它们的索引编码一组最频繁使用的滤波器。可以使用允许对可用滤波器最大变化的所描述的方法来编码使用最不频繁使用的滤波器，并从而克服上面描述的用于编码系数值的第三种现有技术方法的缺点。

滤波器系数的编码顺序并不一定遵循空间顺序。例如，表示内插滤波器系数的差值不一定以与滤波器内出现的系数相同的顺序包括在编码视频比特流内。在这种情况下，必需定义描述差值在比特流内出现的顺序的预定义规则，并且对于编码器和解码器来说这是公知的。

基本滤波器可以适应于同一滤波器的所接收滤波器系数。例如，如果第一发送滤波器系数大于基本滤波器系数，则可以降低第二基本滤波器系数。如果知道这些滤波器系数之和，则尤其是真实的。

典型地，滤波器系数之和是固定。在这种情况下，不需要编码最后一个滤波器系数，而可以通过从总和中减去前面的系数之和来计算出最后一个滤波器系数。如果滤波器系数之和不是固定的，则可以将单独发送的一个或多个常数添加给基本滤波器系数或滤波器输出，以降低系数差值的幅度。

图7图示包括可以被修改以便根据本发明操作的视频编码和解码装置的终端设备。更精确地说，图7图示根据ITU-T推荐标准H.324实现的多媒体终端60。可以将该终端视为多媒体收发信机设备。它包括捕获、编码和多路复用多媒体数据流以便通过通信网络传输的单元以及接收、多路分用、解码和显示所接收的多媒体内容的单元。ITU-T推荐标准H.324定义了终端的全部操作，并参考了控制其各个组成部件操作的其它推荐标准。这种类型的多媒体终端可以在诸如可视电话的实时应用和诸如例如从互联网内的多媒体内容服务器检索或流传输视频剪辑的非实时应用中使用。

在本发明的上下文下，应当认识到，图7所示的H.324终端仅是适合于本发明方法应用的多种可选多媒体终端实现方式之一。还应当指出，在终端装置的位置和实施方式上还存在多种选择。如图7所示，多媒体终端可以定位在连接到诸如模拟PSTN（公用电话交换网）的固定线路电话网的通信设备内。在这种情况下，多媒体终端安装有符合ITU-T推荐标准V.8、V.34和可选的V.8bis的调制解调器71。可选择地，多媒体终端可以连接到外部调制解调器。调制解调器能够将多媒体终端生成的多路复用数字数据和控制数据变换为适合于在PSTN上传输的模拟形式。它还允许多媒体终端以模拟形式从PSTN接收数据和控制信号，并将它们转换成可以由终端以合适的方式多路分用和处理的数字数据流。

还可以以这样的方式实现H.324多媒体终端，以使之可以直接连接到数字

固定线路网络，例如 ISDN（综合业务数字网）。在这种情况下，使用 ISDN 用户网络接口替换调制解调器 71。在图 7 中，用可选方框 72 表示这个 ISDN 用户网络接口。

H.324 多媒体终端还可以适合于在移动通信应用中使用。如果与无线通信链路一起使用，则可以用如在图 7 中用可选方框 73 表示的任意适当的无线接口替换调制解调器 71。例如，H.324/M 多媒体终端可以包括能够连接到当前的第二代 GSM 移动电话网络或者推荐的第三代 UMTS（通用移动电话系统）的无线电收发信机。

应当指出，在设计用于双向通信即用于视频数据的发送和接收的多媒体终端内，提供根据本发明实现的视频编码器和视频解码器是有利的。这样的编码器和解码器对通常实施为单个组合功能单元，称作“编译码器”。

现在将参考图 7 更详细地描述典型的 H.324 多媒体终端。多媒体终端 60 包括多个称作“终端设备”的单元。这个终端设备包括分别用参考数字 61、62 和 63 一般表示的视频、音频和远程信息处理设备。视频设备 61 可以包括例如用于捕获视频图像的视频摄像机、用于显示所接收的视频内容的监视器和可选择的视频处理设备。音频设备 62 通常包括例如用于捕获话音消息的麦克风和用于再生所接收的音频内容的扬声器。音频设备还可以包括附加的音频处理单元。远程信息处理设备 63 可以包括数据终端、键盘、电子白板或者静止图像收发信机，例如传真机单元。

视频设备 61 耦合到视频编译码器 65。视频编译码器 65 包括都是根据本发明实现的视频编码器和相应的视频解码器。在下文中将描述这样的编码器和解码器。视频编译码器 65 负责以合适的形式编码捕获的视频数据以便在通信链路上传输，并解码从通信网络接收到的压缩视频内容。在图 7 所示的例子中，假设以这样的方式实现视频编译码器以包括动态自适应内插滤波器的使用。还假设视频编译码器的编码器部分适合于根据如先前所述的本发明方法的实施例编码和发送内插滤波器系数值给相应的解码器。类似地，视频编译码器的解码器部分适合于接收并解码根据本发明方法的同一实施例编码的滤波器系数值。

终端的音频设备耦合到音频编译码器，在图 7 中用参考数字 66 表示。类似于视频编译码器，音频编译码器包括编码器/解码器对。它将终端的音频设备捕获的音频数据转换成适合于在通信链路上传输的形式并将从网络接收到的编

码音频数据转换回到适合于例如在终端的扬声器上再生的形式。将音频编译码器的输出传送给延迟块 67。这补偿由视频编码处理引入的延迟，并从而保证音频和视频内容的同步。

多媒体终端的系统控制块 64 使用合适的控制协议（信令块 68）控制端到网络信令，以建立在发送和接收终端之间操作的模式。信令块 68 交换与发送和接收终端的编码和解码能力相关的信息，并可以用于实现视频编码器的各种编码模式。系统控制块 64 还控制数据加密的使用。将与在数据传输中将要使用的加密类型相关的信息从加密块 69 传送到多路复用器/多路分用器（MUX/DMUX 单元）70。

在从多媒体终端的数据传输期间，MUX/DMUX 单元 70 组合编码和同步的视频和音频流与从远程信息处理设备 63 输入的数据和可能的控制数据，以形成单个比特流。使用由加密块 69 提供的与将要应用于比特流的数据加密类型（如果有的话）相关的信息来选择加密模式。因此，当接收多路复用的和可能加密的多媒体比特流时，MUX/DMUX 单元 70 负责解密该比特流，将它划分成组成的多媒体分量，并将这些分量发送给合适的用于解码和再生的编译码器和/或终端设备。

图 8a 是根据本发明的优选实施例实现的视频编码器 700 的示意方框图。图 8a 所示的视频编码器的结构在许多方面类似于图 1 所示的现有技术的视频编码器的结构，但对执行与子像素值内插和编码视频比特流的形成相关的操作的那些编码器部分进行了适当修改。视频编码器 700 的大多数单元以类似于先前描述的现有技术的视频编码器 10（参见图 1）的对应单元的方式起作用和操作。为了简明省略这些单元的描述。具体而言，视频编码器 700 包括运动场估算块 711、运动场编码块 712、运动补偿预测块 713、预测误差编码块 714、预测误差解码块 715、多路复用块 716、帧存储器 717 和加法器 719。如图 8a 所示，运动场估算块 711 还包括微分系数计算块 710，它用于计算选定滤波器和基本滤波器 709 之间的差值。

现在将详细描述视频编码器 700 的操作。与现有技术的视频编码器相同，根据本发明这一实施例的视频编码器 700 相对于参考帧 $R_n(x,y)$ 使用运动补偿预测，以生成表示以 INTER 格式编码的视频帧的比特流。它执行到子像素分辨率的运动补偿预测，并且还使用具有动态可变滤波器系数值的内插滤波器，以

形成在运动补偿处理期间需要的子像素值。

作为用于每块的两阶段处理，视频编码器 700 逐块地执行运动补偿预测，并执行到子像素分辨率的运动补偿。在第一阶段中，通过块匹配，即搜索在参考帧 $R_n(x,y)$ 内最佳匹配将要编码的当前图像块的像素值的一块，确定具有全像素分辨率的运动矢量。由运动场估算块 711 结合帧存储器 717 执行块匹配操作，从所述帧存储器检索参考帧 $R_n(x,y)$ 的像素值。在运动补偿预测的第二阶段，将在第一阶段内确定的运动矢量重新精化到所希望的子像素分辨率。为此，运动场估算块 711 通过内插在先前识别为最佳匹配当前编码的图像块的区域（参见图 5）内的参考帧 $R_n(x,y)$ 的像素值来形成具有子像素分辨率的新的搜索块。作为这一处理的一部分，运动场估算块 711 确定用于内插子像素值的最佳内插滤波器。有利地，结合每个图像块的编码修改内插滤波器的系数值。在可选实施例中，可以不频繁地修改内插滤波器的系数，例如每帧一次，或者在将要编码的新视频序列的起始部分上。

在内插必需的子像素值和形成新搜索块之后，运动场估算块 711 执行进一步的搜索，以确定任一新搜索块是否代表与原先在全像素分辨率上标识为最佳匹配块相比更佳的与当前图像块的匹配。这样，运动场估算块 711 确定代表当前正在编码的图像块的运动矢量应当指向全像素位置还是子像素位置。

运动场估算块 711 将所识别出的运动矢量输出给运动场编码块 712，它使用运动模型来近似运动矢量，如先前描述的。随后，运动补偿预测块 713 使用近似的运动矢量和预测误差信息形成当前图像块的预测。随后在预测误差编码块 714 中编码该预测。随后，将用于当前图像块的编码预测误差信息从预测误差编码块 714 转发给多路复用器块 716。多路复用器块 716 还从运动场编码块 712 接收有关近似运动矢量的信息（以运动系数的形式），并从运动场估算块 711 接收在当前图像块的运动补偿预测期间使用的最佳内插滤波器的相关信息。根据本发明的这一实施例，运动场估算块 711 根据由微分系数计算块 710 计算出的计算结果发送一组差值 705，其表示用于当前块的最佳内插滤波器的滤波器系数与在编码器 700 内存储的预定义基本滤波器 709 的系数之间的差。多路复用器块 716 随后通过组合运动信息（运动系数）、预测误差数据、滤波器系数差值和可能的控制信息来形成代表图像当前块的编码比特流 703。在包括在比特流内之前可以使用熵编码器来编码不同种类信息的每种信息，并随后发送给

相应的解码器。

在本发明的可选实施例中，运动场估算块 711 发送表示最佳内插滤波器的滤波器系数的一组值 704 给位于运动场估算块 711 和多路复用器块 716 之间的微分系数计算块 710，如图 8b 所示。根据基本滤波器 709，微分系数计算块 710 计算差值 705，并将它们发送给多路复用器块 716。

在另一种可选实施例中，微分系数计算块 710 驻留在多路复用器块 716 内。在这种情况下，可以由运动场估算块 711 将最佳内插滤波器的滤波器系数 704 直接发送给多路复用器块 716，如图 8c 所示。

图 9a 是根据本发明优选实施例实施的并对应于图 8a 所示的编码器 700 的视频解码器 800 的方框图。解码器 800 包括运动补偿预测块 721、预测误差解码块 722、多路分用块 723 和帧存储器 824。解码器 800 内的大多数单元以类似于现有技术 20 中的相应单元（参见图 2）起作用和操作。然而，如图 9a 所示，本发明的解码器 800 包括滤波器重构块 810，它根据差值 130（图 6b 和图 6c）和预定义基本滤波器 809 重构最佳内插滤波器 110（参见图 6a）。预定义基本滤波器 809 最好等同于基本滤波器 709（图 8a—8c）。

现在将详细地描述视频解码器 800 的操作。多路分用器 823 接收编码比特流 803，将该比特流分割成其组成部分（运动系数、预测误差数据、滤波器系数差值和可能的控制信息），并执行各种数据类型的任何必要的熵解码。多路分用器 823 将从所接收的比特流 803 中检索出的预测误差信息转发给预测误差解码块 822。它还将所接收的运动信息转发给运动补偿预测块 821。在本发明的这一实施例中，多路分用器 823 通过信号 802 将所接收的（和熵解码的）差值转发给运动补偿预测块 821，以允许滤波器重构块 810 通过将所接收的差值添加到在解码器内存储的预定义基本滤波器 809 的系数来重构最佳内插滤波器 110（参见图 6a）。运动补偿预测块 821 随后使用由重构的系数值定义的最佳内插滤波器来构建用于当前被解码的图像块的预测。更具体地说，运动补偿预测块 821 通过检索在帧存储器 824 内存储的参考帧 $R_n(x,y)$ 的像素值并在需要时根据所接收的运动信息内插它们以形成任何需要的子像素值，形成当前图像块的预测。随后组合当前图像块的预测与相应的预测误差数据，以形成所关注的图像块的重构。

可选择地，滤波器重构块 810 驻留在运动补偿预测块 821 之外，如图 9b

所示。根据从多路分用器 823 接收到的信号 802 内包含的差值，滤波器重构块 810 重构最佳内插滤波器，并将重构滤波器系数 805 发送给运动补偿预测块 821。在另一种可选实施例中，滤波器重构块 810 位于多路分用器块 823 内。多路分用器块 823 将重构的最佳内插滤波器的系数转发给运动补偿预测块 821。

根据本发明的编码器相对于预定义基本滤波器系数差分地编码滤波器系数，以允许解码器根据这些差值重构最佳内插滤波器。基本滤波器系数应当是编码器和解码器所知的，并应当在统计上合理地靠近在视频序列内使用的实际滤波器，以产生好的编码性能。换句话说，根据本发明的方法，定义具有一组特定系数值的基本滤波器，并随后编码基本滤波器系数和实际使用的内插滤波器系数之间的差值，并包括在视频比特流内。这样，相对于分开编码每个自适应滤波器系数的方法，降低了表示编码视频比特流内自适应内插滤波器系数需要的信息量。如果基本滤波器的系数非常类似于实际使用的内插滤波器的系数，则将要编码的差值很小。因而，如果预定义基本滤波器在统计上类似于实际使用的内插滤波器，这是有利的，因为在这种情况下，降低了差值，并实现了编码效率的进一步改善。

与在 Wedi 内推荐的差分编码方法不同，根据本发明的方法保持相对好的误差复原能力。在从编码器向解码器传输编码视频比特流期间出现错误的情况下，仅仅是基本滤波器和实际使用的内插滤波器之间的差值受此错误的影响。

应当指出，可以将根据本发明的多媒体终端、视频编码器、解码器和视频编译码器的功能单元实施为软件或专用硬件或两者之组合。根据本发明的视频编码和解码方法尤其适合于以计算机程序的形式实现，所述计算机程序包括用于执行本发明的功能步骤的机器可读指令。因此，根据本发明的编码器、解码器和视频编译码器可以实施为在存储介质上存储的并在诸如个人台式计算机的计算机内执行的软件代码，以便给该计算机提供视频编码和/或解码功能。

尽管以具体实施例为上下文进行了描述，但是对于本领域的技术人员来说，显然可以对这些教导进行多种修改和各种改变。因而，虽然已经针对其一个或多个优选实施例图示和描述了本发明，但是本领域的技术人员将理解在不脱离上面阐述的本发明的范围和精神的情况下可以在其中进行某些修改或改变。

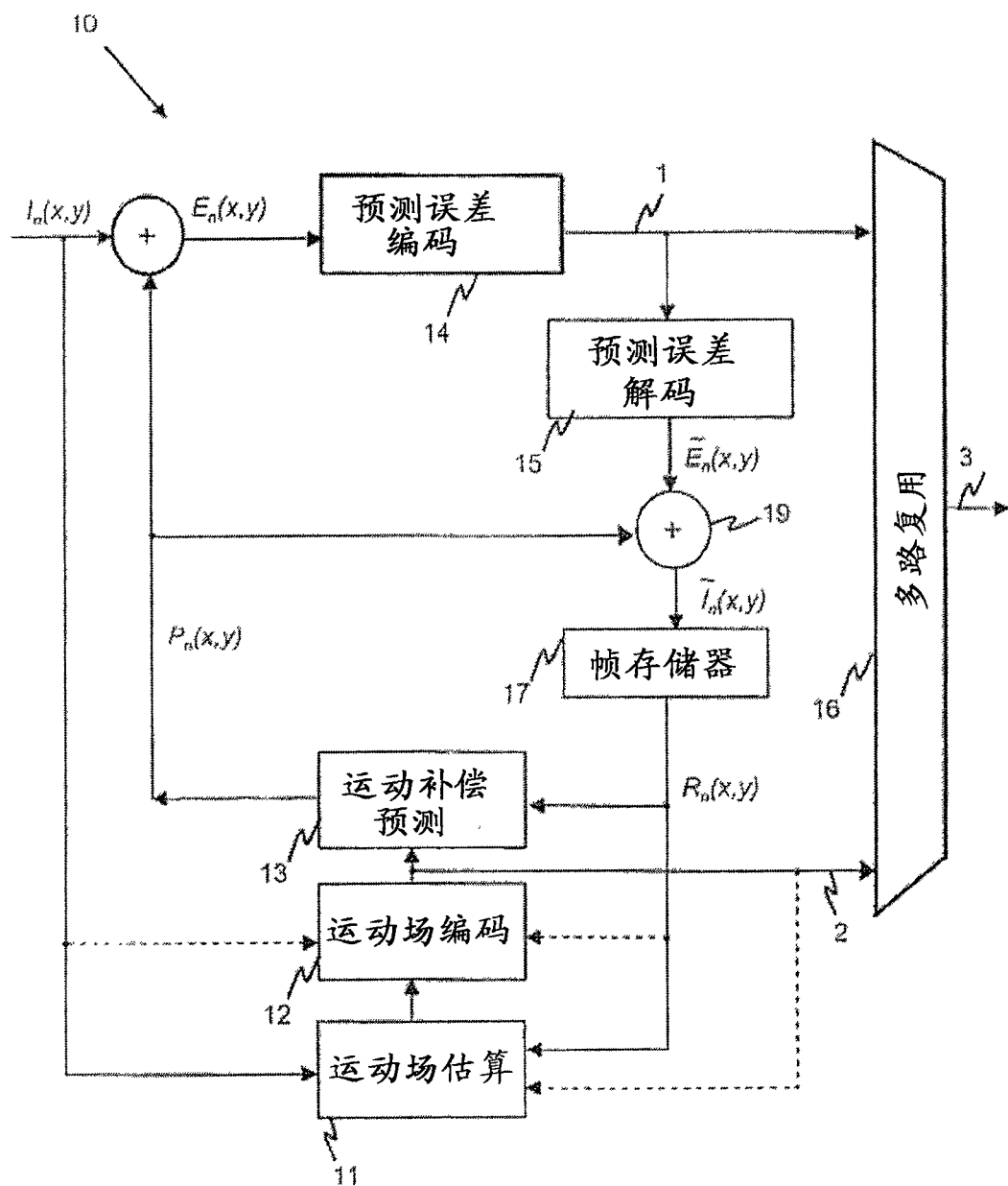


图 1
现有技术

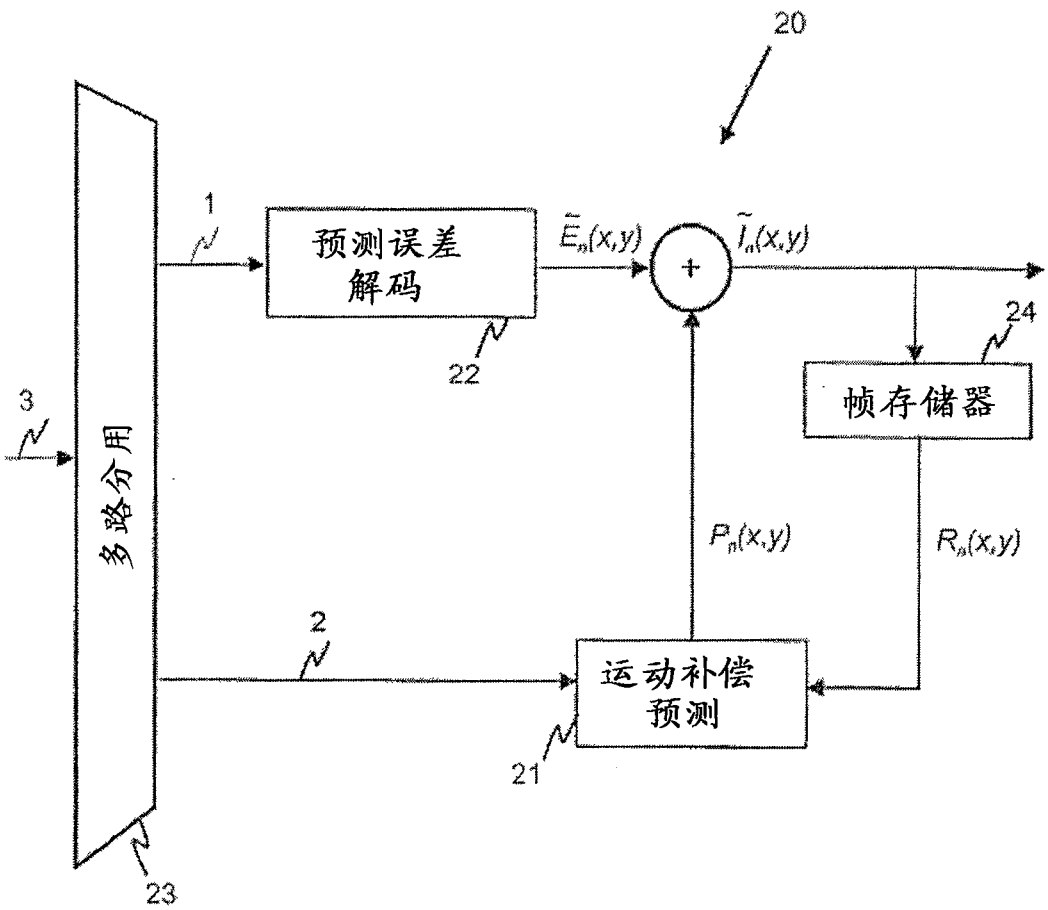


图 2
现有技术

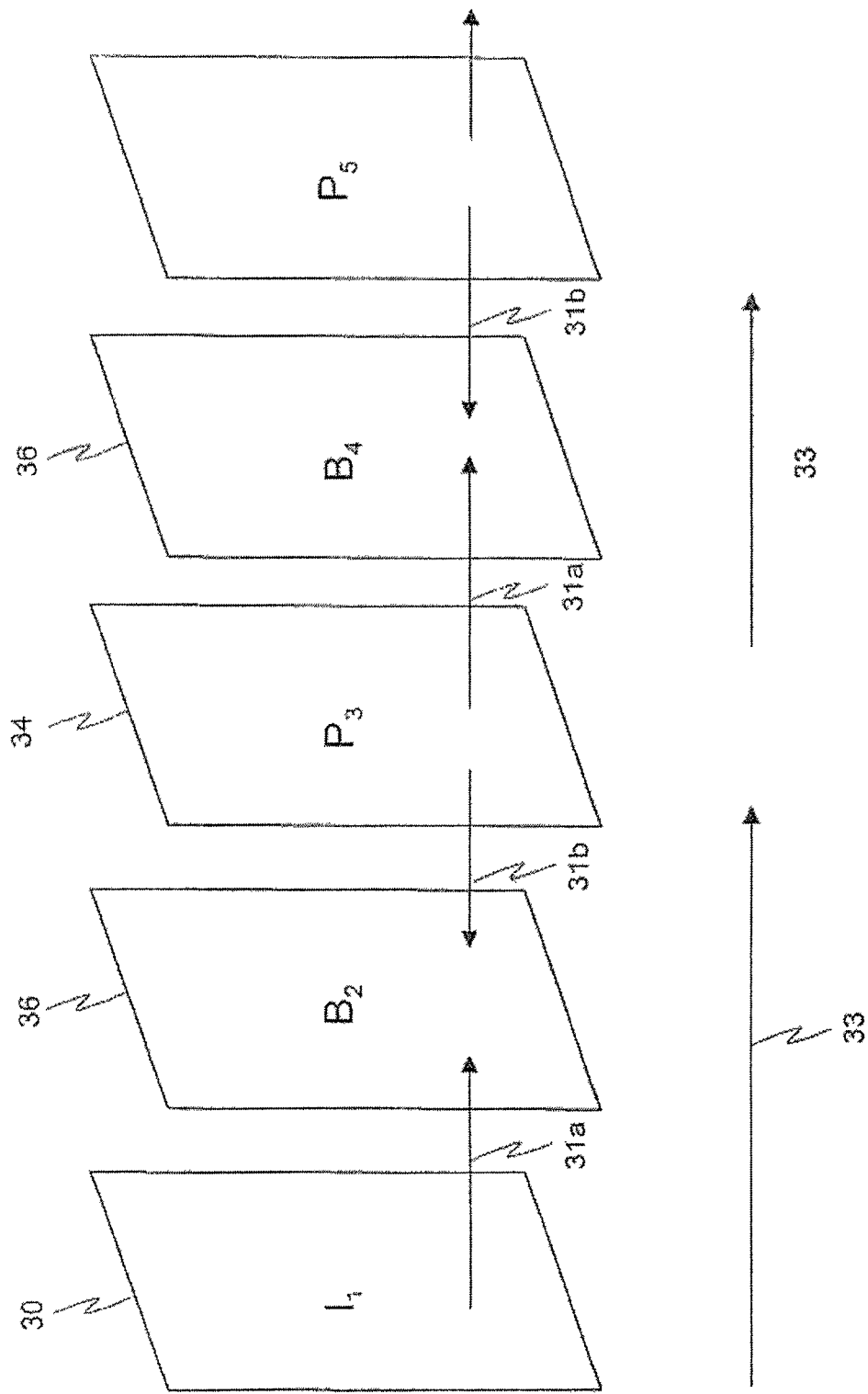


图 3

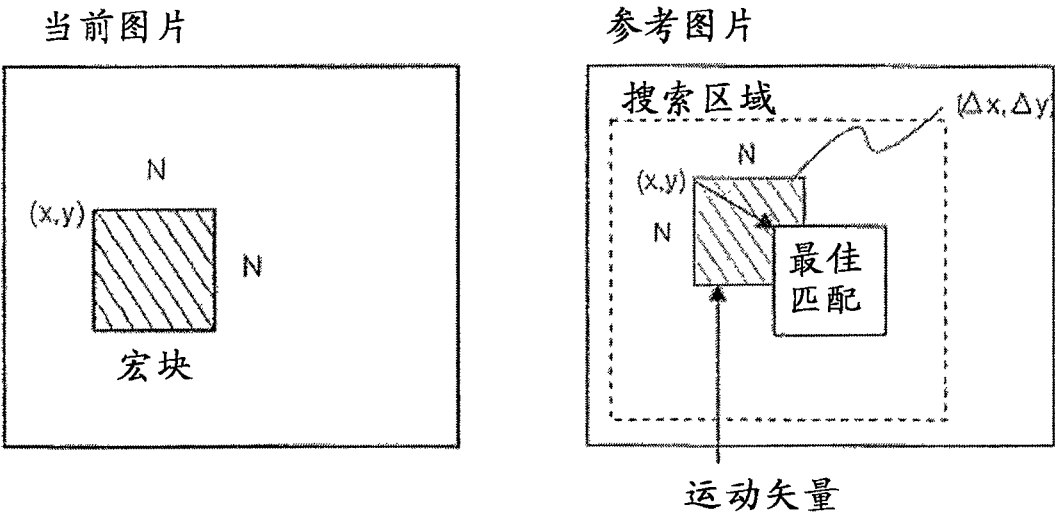


图 4a

图 4b

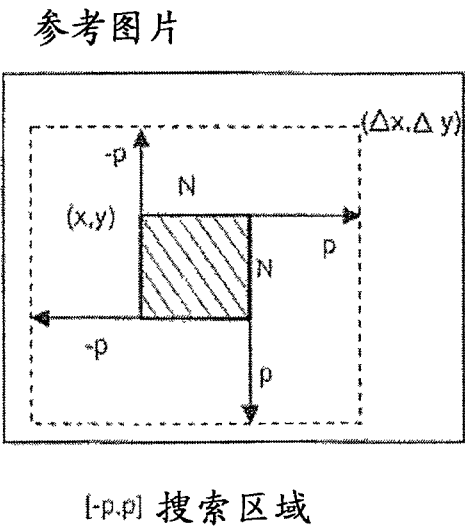


图 4c

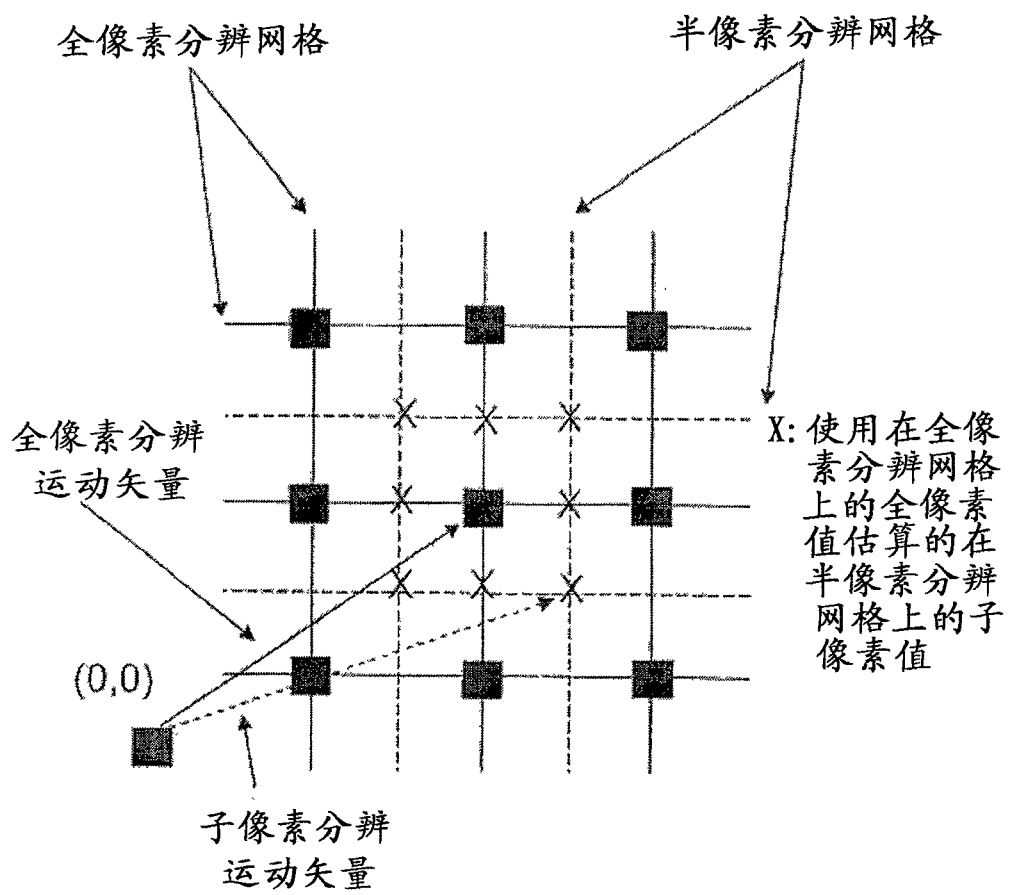


图 5

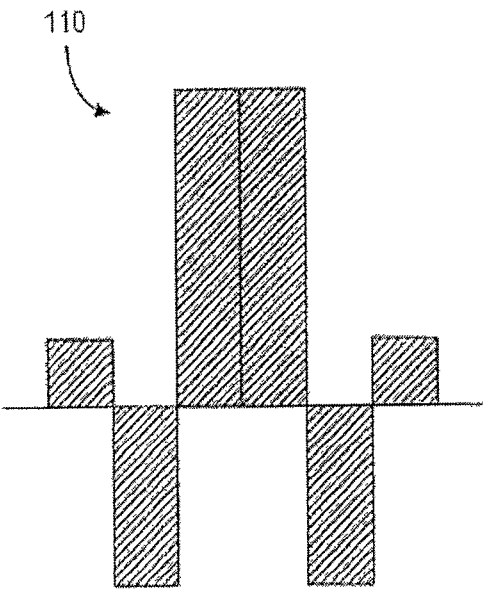


图 6a

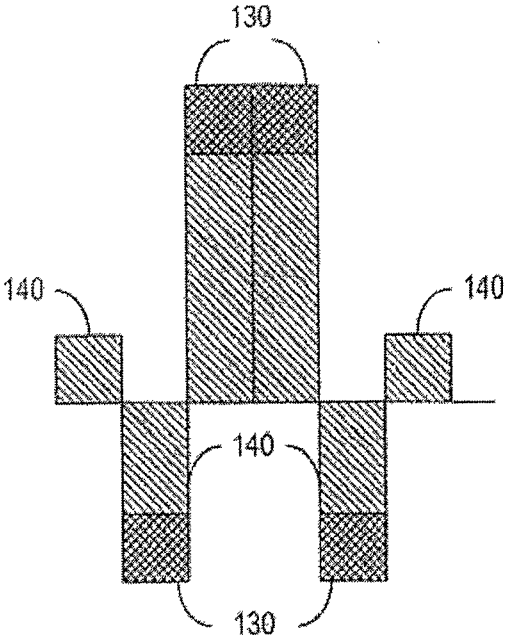


图 6b

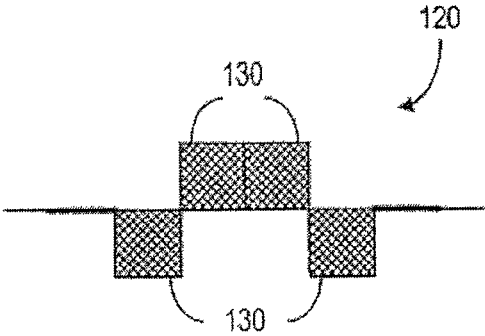


图 6c

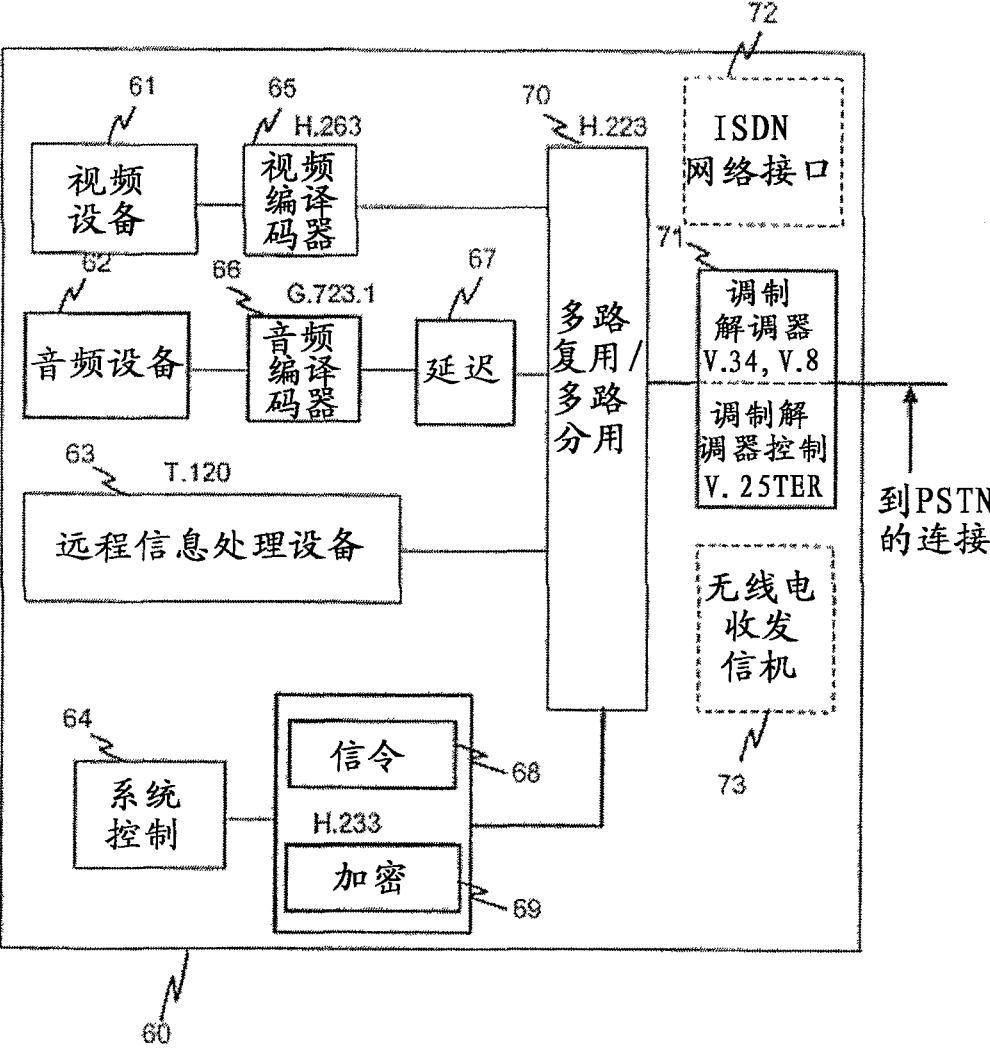


图 7

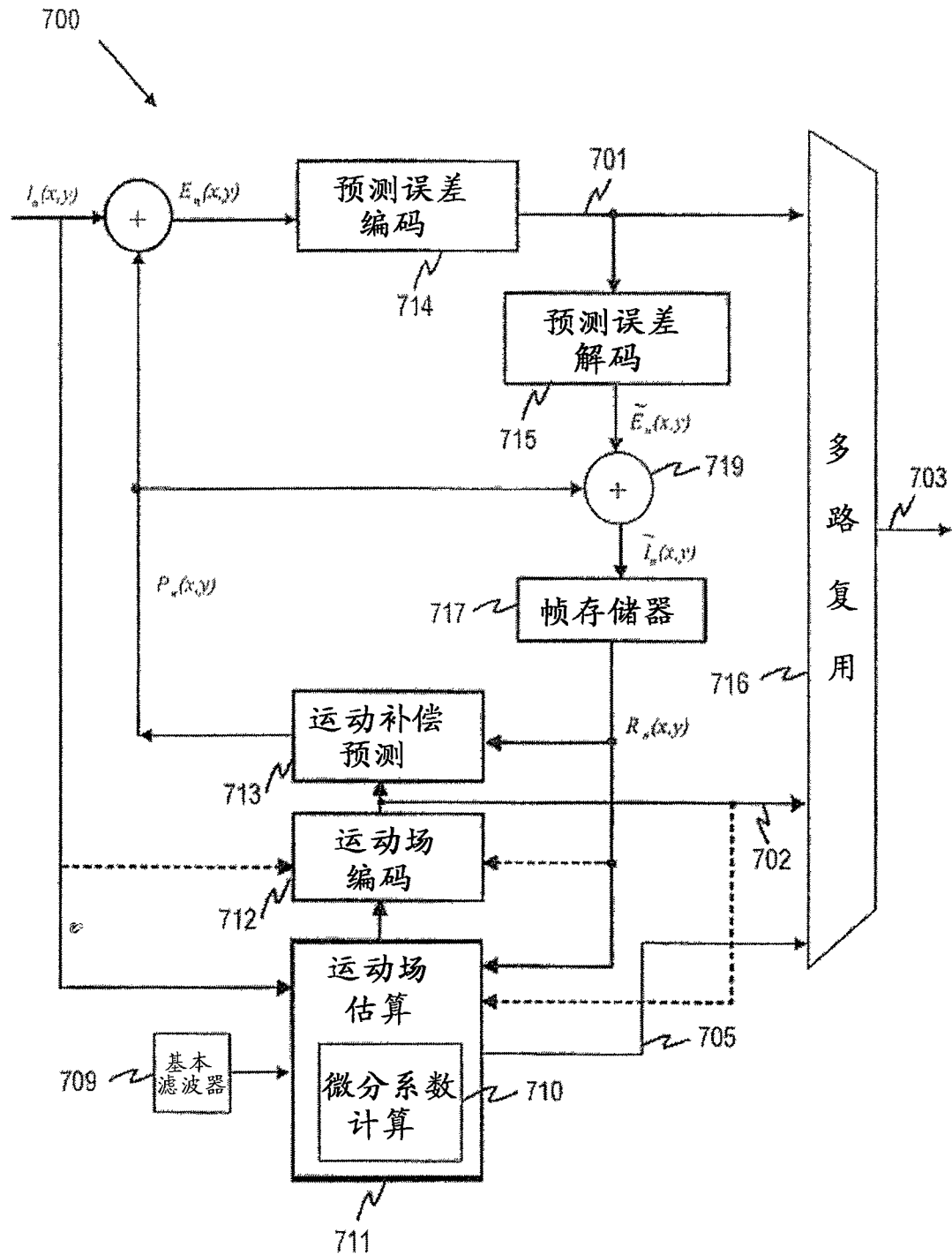


图 8a

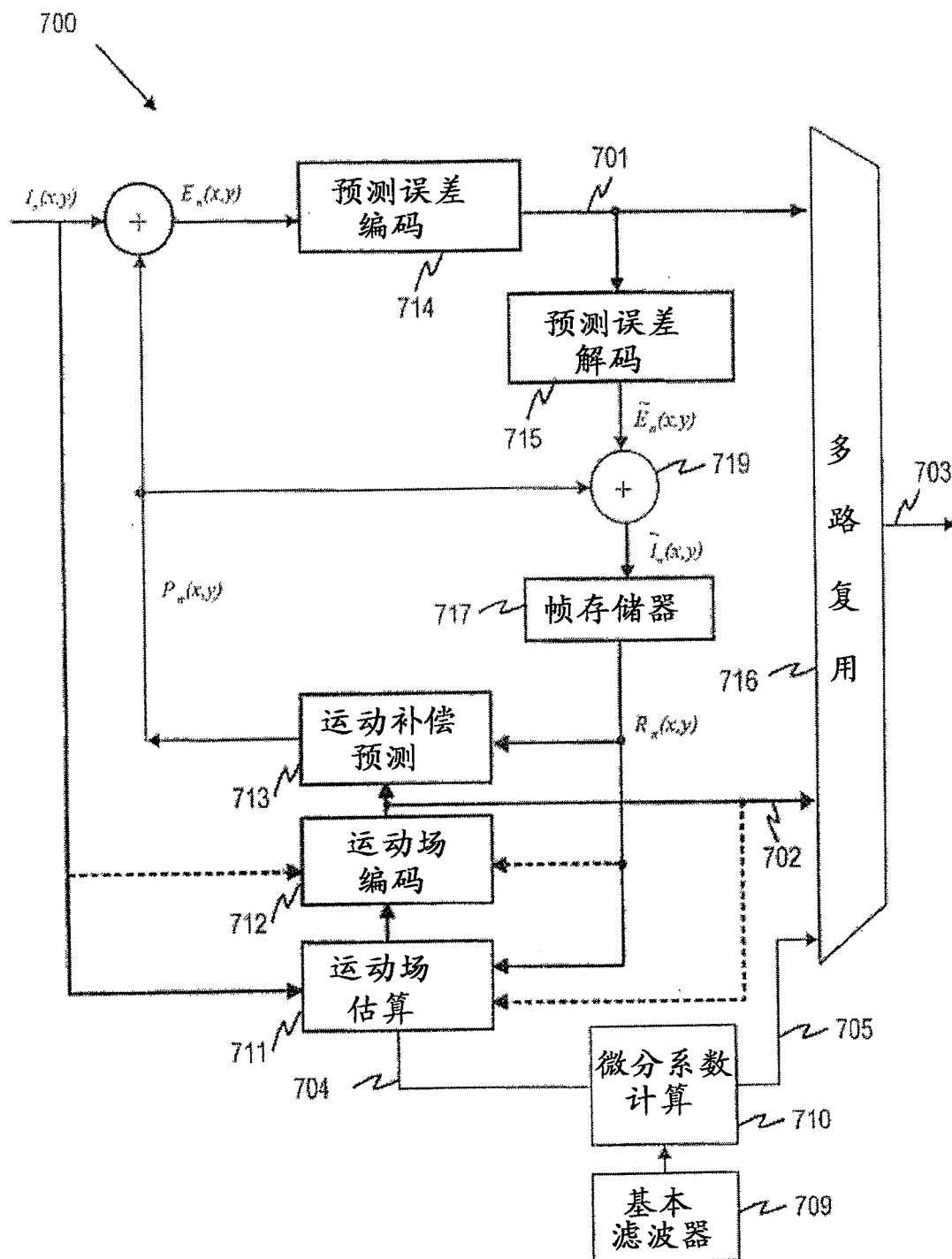


图 8b

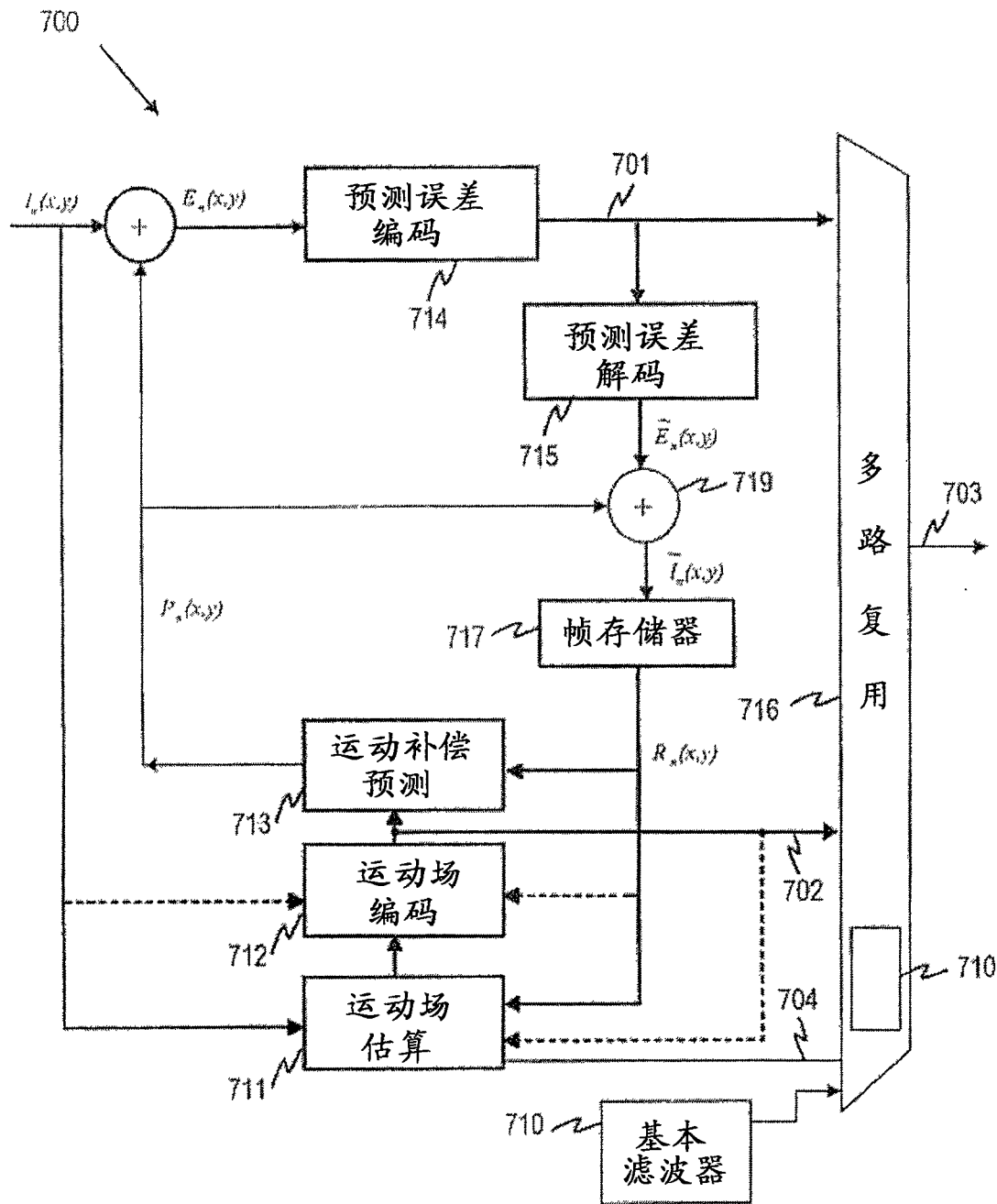


图 8c

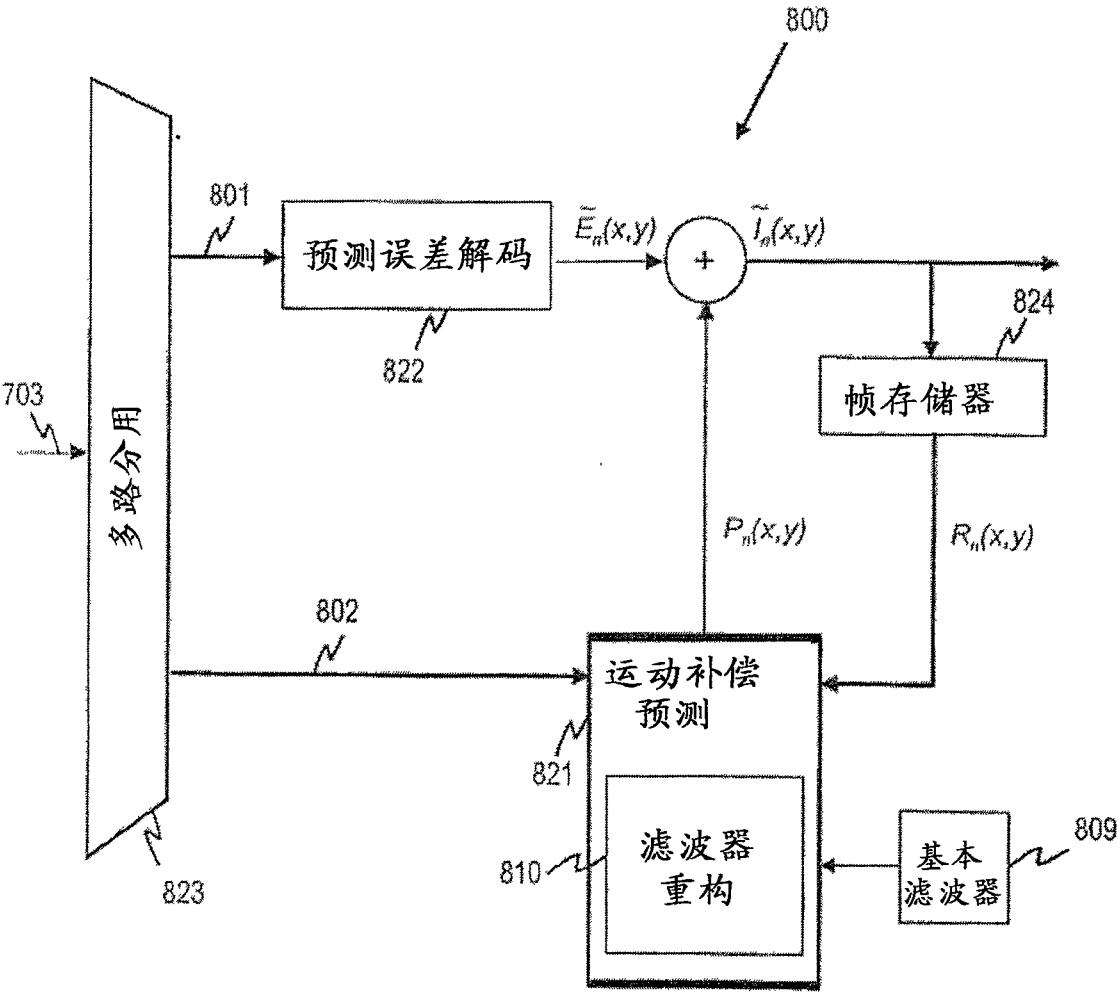


图 9a

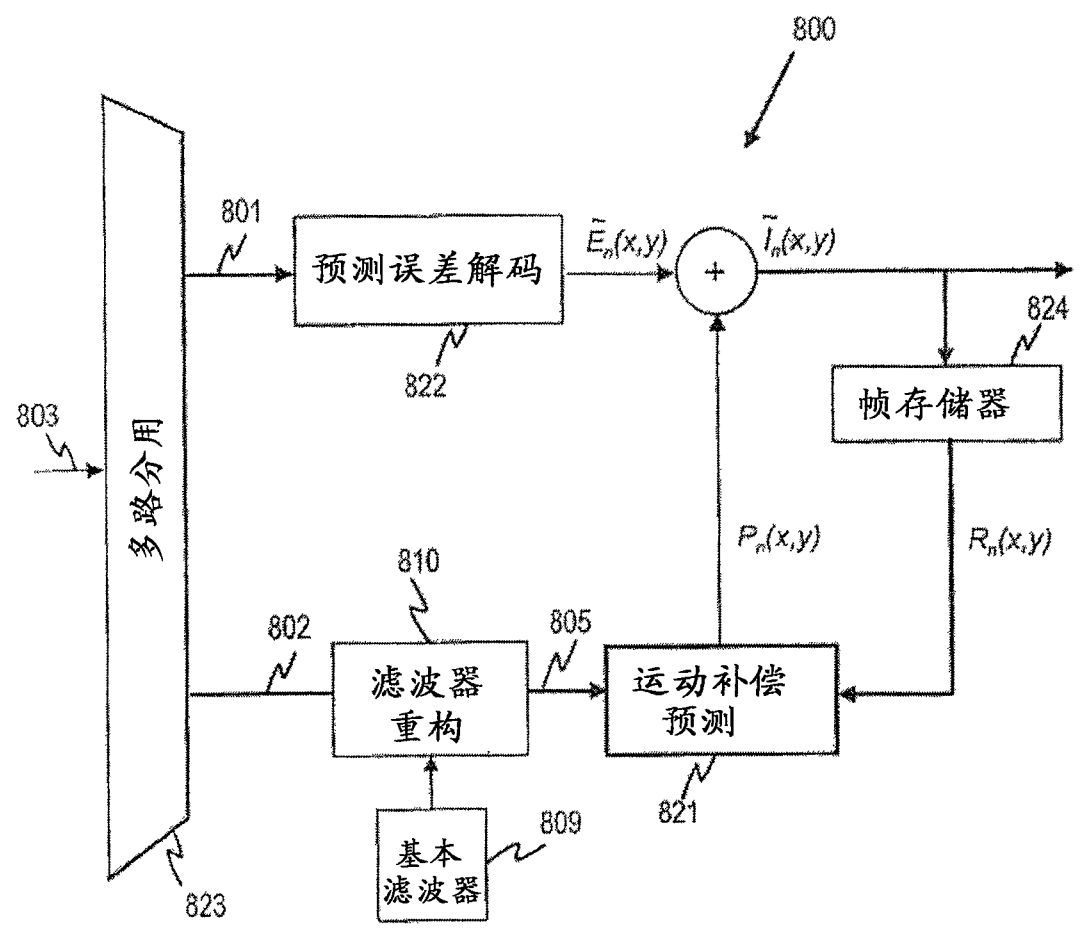


图 9b

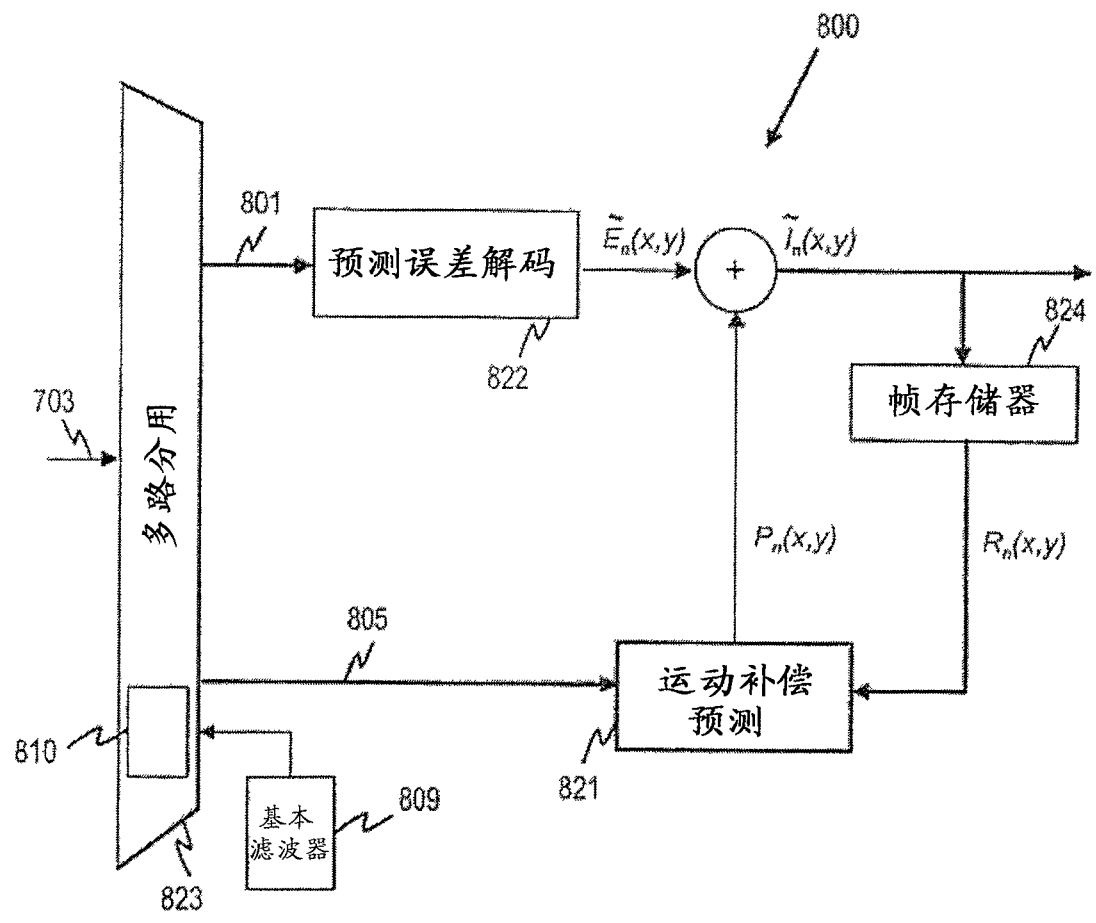


图 9c