



(12) 发明专利

(10) 授权公告号 CN 101366281 B

(45) 授权公告日 2011.06.22

(21) 申请号 200780001989.6

(22) 申请日 2007.01.08

(30) 优先权数据

60/756,846 2006.01.06 US

60/786,573 2006.03.27 US

60/829,515 2006.10.13 US

(85) PCT申请进入国家阶段日

2008.07.07

(86) PCT申请的申请数据

PCT/US2007/000195 2007.01.08

(87) PCT申请的公布数据

W02007/081752 EN 2007.07.19

(73) 专利权人 微软公司

地址 美国华盛顿州

(72) 发明人 G·J·苏利万

(74) 专利代理机构 上海专利商标事务所有限公

司 31100

代理人 顾嘉运

(51) Int. Cl.

H04N 7/24 (2006.01)

H04N 7/26 (2006.01)

H04N 7/46 (2006.01)

H04N 7/50 (2006.01)

(56) 对比文件

US 5555021 A, 1996.09.10, 全文.

CN 1203500 A, 1998.12.30, 全文.

JP 8009281 A, 1996.01.12, 全文.

CN 1208306 A, 1999.02.17, 全文.

CN 1274503 A, 2000.11.22, 全文.

审查员 张伯约

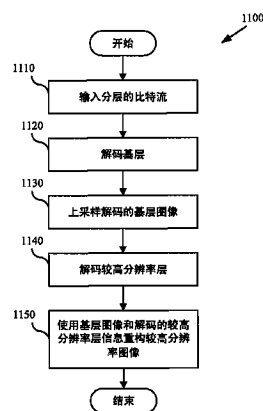
权利要求书 3 页 说明书 28 页 附图 12 页

(54) 发明名称

用于多分辨率视频编码和解码的重新采样和
调整图像大小操作的方法

(57) 摘要

描述了在诸如空间可伸缩视频编码和解码的应用中为调整图像大小计算高精度位置的技术和工具。在一个方面,根据重新采样比例因子执行视频图像的重新采样。重新采样包括计算重新采样的阵列中位置 i, j 处的样本值。计算包括以涉及通过用上采样比例因子的倒数(近似的或精确的)乘以 2^n 值来对值的一部分进行近似的方式来计算导出的水平或垂直子样本位置 x 或 y 。近似可以是舍入或者某些其他类型的近似,诸如近似到邻近整数的取上限整数或取下限整数函数。使用滤波器对样本值进行内插。



1. 一种对多分辨率视频编码或解码进行重新采样的方法,包括:

根据水平重新采样比例因子执行图像数据的重新采样,其中所述重新采样包括在重新采样的阵列中水平位置 i 处样本值的计算,并且其中所述计算包括:

以数学上等于公式 $x = (i * C + D) \gg S$ 的结果的方式计算导出的水平子样本位置 x ,其中 C 通过对水平重新采样比例因子的倒数乘以等于 2^{S+F} 的值进行近似而导出,并且其中 F 、 C 、 D 和 S 是整数值,其中, F 是基于水平子样本位置 x 的分数分量的位数, D 是一偏移量,而 S 设置动态范围和精度。

2. 一种对多分辨率视频编码或解码进行重新采样的方法,包括:

根据垂直重新采样比例因子执行图像数据的重新采样,其中所述重新采样包括在重新采样的阵列中垂直位置 j 处样本值的计算,并且其中所述计算包括:

以数学上等于公式 $y = (j * C + D) \gg S$ 的结果的方式计算导出的垂直子样本位置 y ,其中 C 通过对垂直重新采样比例因子的倒数乘以等于 2^{S+F} 的值进行近似而导出,并且其中 F 、 C 、 D 和 S 是整数值,其中, F 是基于垂直子样本位置 y 的分数分量的位数, D 是一偏移量,而 S 设置动态范围和精度。

3. 一种对多分辨率视频编码或解码进行重新采样的方法,包括:

根据水平上采样比例因子和垂直上采样比例因子执行视频图像的上采样,其中所述上采样包括在上采样的阵列中水平位置 i 和垂直位置 j 处内插样本值的计算,并且其中所述计算包括:

以数学上等于公式 $x = (i * C + D) \gg S$ 的结果的方式计算导出的水平子样本位置 x ,其中 C 通过对水平上采样比例因子的倒数乘以等于 2^{S+F} 的值进行近似而导出,并且其中 F 、 C 、 D 和 S 是整数值,其中, F 是基于水平子样本位置 x 的分数分量的位数, D 是一偏移量,而 S 设置动态范围和精度;

以数学上等于公式 $y = (j * C + D) \gg S$ 的结果的方式计算导出的垂直子样本位置 y ,其中 C 通过对垂直上采样比例因子的倒数乘以等于 2^{S+F} 的值进行近似而导出,其中, F 是基于垂直子样本位置 y 的分数分量的位数, D 是一偏移量,而 S 设置动态范围和精度;以及

在由导出的水平子样本位置 x 和导出的垂直子样本位置 y 所定义的位置 x, y 处内插一样本值。

4. 如权利要求 3 所述的方法,其特征在于,还包括:

基于所述导出的水平子样本位置 x 的 F 个最低有效位选择水平重新采样滤波器;以及
基于所述导出的水平子样本位置 x 的剩余更高有效位选择要被滤波的较低分辨率样本;以及

其中在所述位置 x, y 处内插一样本值包括:

基于所选的较低分辨率样本并使用所选的水平重新采样滤波器来内插所述样本值。

5. 如权利要求 4 所述的方法,其特征在于,应用于所述导出的水平子样本位置 x 的 F 个最低有效位的至少一个值的水平重新采样滤波器是带有两个以上非零滤波器抽头值的有限脉冲响应滤波器。

6. 如权利要求 4 所述的方法,其特征在于,应用于所述导出的水平子样本位置 x 的 F 个最低有效位的除零之外所有值的水平重新采样滤波器是带有四个非零滤波器抽头值的有限脉冲响应滤波器。

7. 如权利要求 3 所述的方法,其特征在于,还包括:

基于所述导出的垂直子样本位置 y 的 F 个最低有效位选择垂直重新采样滤波器;以及
基于所述导出的垂直子样本位置 y 的剩余更高有效位选择要被滤波的较低分辨率样本;以及

其中在所述位置 x, y 处内插一样本值包括:

基于所选的较低分辨率样本并使用所选的垂直重新采样滤波器来内插所述样本值。

8. 如权利要求 7 所述的方法,其特征在于,应用于所述导出的垂直子样本位置 y 的 F 个最低有效位的至少一个值的垂直重新采样滤波器是带有两个以上非零滤波器抽头值的有限脉冲响应滤波器。

9. 如权利要求 7 所述的方法,其特征在于,应用于所述导出的垂直子样本位置 x 的 F 个最低有效位的除零之外所有值的垂直重新采样滤波器是带有四个非零滤波器抽头值的有限脉冲响应滤波器。

10. 如权利要求 3 所述的方法,其特征在于,所述上采样是使用一个或多个 Mitchell-Netravalli 重新采样滤波器执行的。

11. 如权利要求 3 所述的方法,其特征在于,所述上采样是使用一个或多个 Catmull-Rom 重新采样滤波器执行的。

12. 如权利要求 3 所述的方法,其特征在于, F 、 C 、 D 和 S 的所述垂直值或水平值中的至少一个值至少部分基于所述样本值是色度样本值还是亮度样本值而不同。

13. 如权利要求 3 所述的方法,其特征在于,数学上等于所述公式 $x = (i * C + D) \gg S$ 的结果的方式包括所述公式 $x = ((i * C + D) \gg S) + E$ 的实现,其中 E 是第二偏移量。

14. 如权利要求 3 所述的方法,其特征在于,所述上采样是使用具有受控于带宽控制参数的滤波器抽头值的一个或多个重新采样滤波器执行的。

15. 如权利要求 3 所述的方法,其特征在于,所述上采样是在分层的空间可伸缩视频解码过程中执行的。

16. 如权利要求 3 所述的方法,其特征在于,所述上采样是在分层的空间可伸缩视频编码过程中执行的。

17. 如权利要求 3 所述的方法,其特征在于,所述上采样是针对参考图像重新采样执行的。

18. 如权利要求 3 所述的方法,其特征在于,所述 F 的值等于 4 而所述 S 的值等于 12。

19. 如权利要求 3 所述的方法,其特征在于,所述近似包括舍入。

20. 如权利要求 3 所述的方法,其特征在于,所述 F 、 C 、 D 和 S 的整数值中的至少一个值对于所述水平计算和所述垂直计算不同。

21. 一种对多分辨率视频编码或解码进行重新采样的方法,包括:

根据水平上采样比例因子和垂直上采样比例因子执行视频图像的上采样,其中所述上采样包括在上采样的阵列中水平位置 i 和垂直位置 j 处内插样本值的计算,并且其中所述计算包括:

以数学上等于所述公式 $x = ((2^J * i + Q) * C + D) \gg S$ 的结果的方式计算导出的水平子样本位置 x ,其中 C 通过对水平上采样比例因子的倒数乘以等于 2^{S+F} 的值进行近似而导出,并且其中 F 、 C 、 D 、 S 、 J 和 Q 是整数值,其中, F 是基于水平子样本位置 x 的分数分量的位数,

D 和 Q 是偏移量, J 是基于所述重采样是用于亮度还是色度, 而 S 设置动态范围和精度;

以数学上等于所述公式 $y = ((2^J * j + Q) * C + D) \gg S$ 的结果的方式计算导出的垂直子样本位置 y, 其中 C 是通过 C 通过对垂直上采样比例因子的倒数乘以等于 2^{S+F} 的值进行近似而导出, 其中 F、C、D、S、J 和 Q 是整数值, 其中, F 是基于垂直子样本位置 y 的分数分量的位数, D 和 Q 是偏移量, J 是基于所述重采样是用于亮度还是色度, 而 S 设置动态范围和精度; 以及

在由导出的水平子样本位置 x 和导出的垂直子样本位置 y 所定义的位置 x, y 处内插一样本值。

用于多分辨率视频编码和解码的重新采样和调整图像大小 操作的方法

技术领域

[0001] 描述了用于编码 / 解码数字视频的技术和工具。

[0002] 背景

[0003] 随着 DVD 的日益流行, 经因特网、数码相机、数字媒体传递音乐也变得愈发平常。工程师使用各种技术在维持质量的同时有效处理数字音频、视频和图像。为了理解这些技术, 理解音频、视频和图像信息是如何在计算机内表示和处理是有帮助的。

[0004] I. 计算机中媒体信息的表示

[0005] 计算机将媒体信息作为表示该信息的一系列数字来处理。例如, 单个数字可以表示图像的每个元素小区域的亮度强度和诸如红色、绿色或蓝色的色彩分量的强度, 使得图像的数字表示由这些数字的一个或多个阵列组成。每个这样的数字可以被称为一个样本。对于彩色图像, 习惯上用一个以上的样本来表示每个元素区域的颜色, 通常使用三个样本。针对该元素区域的这些样本的集合可被称之为像素, 其中单词“像素”指的是“图像元素”这一概念的缩写。例如, 一个像素可以由显示元素区域所必须的表示红光、绿光和蓝光强度的三个样本组成。这一像素类型被称为 RGB 像素。若干因素影响媒体信息的质量, 包括样本深度、分辨率和帧率 (用于视频)。

[0006] 样本深度是指示能够用来表示一样本的数量范围的通常以比特为单位测得的性质。当用于该样本的可能值越多, 质量就越高, 因为该数量能够捕捉更精细的强度变化和 / 或更大的值范围。分辨率通常指的是某一持续时间 (对于音频) 或空间 (对于图像或单独的视频图像) 内的样本数。空间分辨率更高的图像往往比其他图像看上去更明快, 并且包含更多可辨别的有用细节。帧率是视频时间分辨率的常用术语。较高帧率的视频往往比其他视频更能模拟自然物体的平滑运动, 并且也可以类似地认为其在时间维度上包含更多细节。鉴于所有这些因素, 如下表 1 所示按照表示样本深度、分辨率和帧率所需的比特率, 高质量对存储和传输信息的成本的折衷。

[0007]

每像素的比特数 (采样深度乘以每像素 的样本数)	分辨率 (以像素为单 位, 宽 × 高)	帧率 (帧 / 秒)	比特率 (百万比特 / 秒)
8 (值 0-255, 单色)	160x120	7.5	1.2
24 (值 0-255, RGB)	320x240	15	27.6
24 (值 0-255, RGB)	640x480	30	221.2
24 (值 0-255, RGB)	1280x720	60	1327.1

[0008] 表 1: 原始视频不同质量级的比特率

[0009] 不考虑存储和发送高质量视频（诸如 HDTV）所需的高比特率，公司和消费者愈发依赖计算机来创建、分布和回放高质量内容。为此，工程师使用压缩（也称为源译码或源编码）来降低数字媒体的比特率。压缩通过将信息转换成更低比特率的形式而降低了信息存储和传送的成本。压缩可以是无损的，其中视频质量不受损害，但是比特率的降低受到视频复杂度的限制。或者，压缩可以是有损的，其中视频质量受到损害，但是比特率的降低更显著。解压（也被称为解码）从压缩形式中重构原始信息版本。“编解码器”是编码器 / 解码器系统。

[0010] 一般而言，视频压缩技术包括“帧内”压缩和“帧间”或预测压缩。对于视频图像，帧内压缩技术压缩各个图像。帧间压缩技术参考在前和 / 或在后图像压缩图像。

[0011] II. 多分辨率视频和空间可伸缩性

[0012] 标准的视频编码器在目标比特率低于某一阈值时会经历大幅的性能劣化。量化和其他有损处理阶段引入失真。在低比特率的情况下，高频信息可能会严重失真或者完全丢失。结果是出现明显的伪像并使得重构视频的质量显著降低。虽然可用比特率随着传输和处理技术的改进而提升，但是以受限的比特率维持高视觉质量仍然是视频编解码设计的主要目标。现有的编解码器使用若干种方法来改善受限比特率下的视觉质量。

[0013] 多分辨率编码允许以不同的空间分辨率来编码视频。分辨率降低的视频能够以信息损失为代价按足够低的比特率编码。例如，在前视频编码器可以下采样（使用下采样滤波器）全分辨率视频并在垂直和 / 或水平方向上以降低的分辨率对其进行编码。每个方向上分辨率降低一半使得经编码图像的大小维度缩小一半。编码器发信号通知解码器这一分辨率降低的编码。解码器接收指示分辨率降低编码的信息并从接收到的信息确定该分辨率降低的视频应被如何上采样（使用上采样滤波器）来在显示前增加其图像大小。然而，在编码器下采样和编码时就已损失的信息在经上采样的图像中仍然丢失。

[0014] 空间可伸缩的视频使用多层方法，允许编码器在基层降低空间分辨率（并由此降低比特率），同时在一个或多个增强层中维持来自源视频的更高分辨率的信息。例如，基层帧内图像可以用降低的分辨率编码，而伴随的增强层帧内图像则可以用更高分辨率编码。类似地，基层预测的图像可由增强层预测的图像伴随。解码器可以选择（鉴于比特率约束和 / 或其他准则）仅解码较低分辨率的基层以图像获得较低分辨率的重构图像，或者解码基层和增强层的图像以获得更高分辨率的重构图像。当以低于显示图像的分辨率对基层进行编码（也被称为下采样）时，经编码的图像大小实际上小于显示的图像。解码器执行计算以调整重构图像的大小并使用上采样滤波器在重构图像内的合适位置处生成内插样本值。然而，使用空间可伸缩视频的早先的编解码器已经由于非柔性的上采样滤波器以及不精确或昂贵（在计算时间或比特率方面）的调整图像大小技术而蒙受损失。

[0015] 给定视频压缩和解压对于数字视频的关键重要性，视频压缩和解压是充分开发的领域并不令人惊奇。然而，不论早先的视频压缩和解压技术的好处如何，它们都没有以下技术和工具的优点。

发明内容

[0016] 提供本概述以便以简化形式介绍将在以下的详细描述中进一步描述的一些概念。本概述并不旨在标识要求保护的主题的关键特征或本质特征，也不旨在用于帮助限制要求

保护的主题的范围。

[0017] 在本概述中,详细描述涉及用于多分辨率和分层空间可伸缩视频编码和解码的各种技术和工具。

[0018] 例如,详细描述涉及在诸如空间可伸缩视频编码和解码的应用中用于调整图像大小的高精度位置计算的各种技术和工具。描述了在诸如空间可伸缩视频编码和解码的应用中为调整图像大小高精度位置计算的技术和工具。在一个方面,根据重新采样比例因子执行视频图像的重新采样。重新采样包括计算重新采样阵列中位置 i, j 处的样本值。计算包括以涉及通过用上采样比例因子的倒数(近似的或精确的)乘以 2^n 值来对值的一部分进行近似(或者用 2^n 值除以上采样比例因子或者对上采样比例因子的近似)的方式来计算导出的水平或垂直子样本位置 x 或 y 。指数 n 可以是包括表示分数分量位数的整数 F 在内的两个整数之和。近似可以是舍入或者某些其他类型的近似,诸如近似到邻近整数的取上限整数或取下限整数函数。使用滤波器对样本值进行内插。

[0019] 所描述的技术的一些替换提供了变化的样本位置计算,在一个实现中提供计算中近似一个额外位的精度而不会明显改变该样本位置计算过程或其复杂性。所描述的技术的一些进一步替换涉及如何对 4:2:2 和 4:4:4 采样结构进行样本位置计算。用于这些采样结构的这些替换技术只要在色度和亮度采样网格的分辨率在一特定维度上相同就把亮度和色度样本位置计算锁定在一起。

[0020] 根据本发明的第一方面,提供了一种对多分辨率视频编码或解码进行重新采样的方法,包括:根据水平重新采样比例因子执行图像数据的重新采样,其中所述重新采样包括在重新采样的阵列中水平位置 i 处样本值的计算,并且其中所述计算包括:以数学上等于公式 $x = (i * C + D) \gg S$ 的结果的方式计算导出的水平子样本位置 x ,其中 C 通过对水平重新采样比例因子的倒数乘以等于 2^{S+F} 的值进行近似而导出,并且其中 F, C, D 和 S 是整数值,其中, F 是基于水平子样本位置 x 的分数分量的位数, D 是一偏移量,而 S 设置动态范围和精度。

[0021] 根据本发明的第二方面,提供了一种对多分辨率视频编码或解码进行重新采样的方法,包括:根据垂直重新采样比例因子执行图像数据的重新采样,其中所述重新采样包括在重新采样的阵列中垂直位置 j 处样本值的计算,并且其中所述计算包括:以数学上等于公式 $y = (j * C + D) \gg S$ 的结果的方式计算导出的垂直子样本位置 y ,其中 C 通过对垂直重新采样比例因子的倒数乘以等于 2^{S+F} 的值进行近似而导出,并且其中 F, C, D 和 S 是整数值,其中, F 是基于垂直子样本位置 y 的分数分量的位数, D 是一偏移量,而 S 设置动态范围和精度。

[0022] 根据本发明的第三方面,提供了一种对多分辨率视频编码或解码进行重新采样的方法,包括:根据水平上采样比例因子和垂直上采样比例因子执行视频图像的上采样,其中所述上采样包括在上采样的阵列中水平位置 i 和垂直位置 j 处内插样本值的计算,并且其中所述计算包括:以数学上等于公式 $x = (i * C + D) \gg S$ 的结果的方式计算导出的水平子样本位置 x ,其中 C 通过对水平上采样比例因子的倒数乘以等于 2^{S+F} 的值进行近似而导出,并且其中 F, C, D 和 S 是整数值,其中, F 是基于水平子样本位置 x 的分数分量的位数, D 是一偏移量,而 S 设置动态范围和精度;以数学上等于公式 $y = (j * C + D) \gg S$ 的结果的方式计算导出的垂直子样本位置 y ,其中 C 通过对垂直上采样比例因子的倒数乘以等于 2^{S+F} 的值

进行近似而导出,其中,F是基于垂直子样本位置y的分数分量的位数,D是一偏移量,而S设置动态范围和精度;以及在由导出的水平子样本位置x和导出的垂直子样本位置y所定义的位置x,y处内插一样本值。

[0023] 根据本发明的第四方面,提供了一种对多分辨率视频编码或解码进行重新采样的方法,包括:根据水平上采样比例因子和垂直上采样比例因子执行视频图像的上采样,其中所述上采样包括在上采样的阵列中水平位置i和垂直位置j处内插样本值的计算,并且其中所述计算包括:以数学上等于所述公式 $x = ((2^J * i + Q) * C + D) >> S$ 的结果的方式计算导出的水平子样本位置x,其中C通过对水平上采样比例因子的倒数乘以等于 2^{S+F} 的值进行近似而导出,并且其中F、C、D、S、J和Q是整数值,其中,F是基于水平子样本位置x的分数分量的位数,D和Q是偏移量,J是基于所述重采样是用于亮度还是色度,而S设置动态范围和精度;以数学上等于所述公式 $y = ((2^J * j + Q) * C + D) >> S$ 的结果的方式计算导出的垂直子样本位置y,其中C是通过C通过对垂直上采样比例因子的倒数乘以等于 2^{S+F} 的值进行近似而导出,其中F、C、D、S、J和Q是整数值,其中,F是基于垂直子样本位置y的分数分量的位数,D和Q是偏移量,J是基于所述重采样是用于亮度还是色度,而S设置动态范围和精度;以及在由导出的水平子样本位置x和导出的垂直子样本位置y所定义的位置x,y处内插一样本值。

[0024] 当参考附图阅读以下各个实施例的详细描述时,可以清楚其它特征和优点。

附图说明

[0025] 图1是可结合其实现所描述的若干实施例的合适的计算环境的框图。

[0026] 图2是可结合其实现所描述的若干实施例的通用视频编码器系统的框图。

[0027] 图3是可结合其实现所描述的若干实施例的通用视频解码器系统的框图。

[0028] 图4是在所描述的若干实施例中使用的宏块格式的图示。

[0029] 图5A是隔行扫描视频帧的一部分的图示,示出了上半帧和下半帧的交错行。图5B是为编码/解码组织为帧的隔行扫描视频帧的图示,且图5C是为编码/解码组织为半帧的隔行扫描视频帧的图示。

[0030] 图5D示出了相对于视频帧每个半帧的亮度样本位置的4:2:0色度样本位置的六个示例性空间排列。

[0031] 图6是示出了用于多分辨率视频编码的概括性技术的流程图。

[0032] 图7是示出了用于多分辨率视频解码的概括性技术的流程图。

[0033] 图8是示出了多分辨率帧内图像和帧间预测图像的多分辨率编码技术的流程图。

[0034] 图9是示出了多分辨率帧内图像和帧间预测图像的多分辨率解码技术的流程图。

[0035] 图10是示出了用于编码空间可伸缩比特流层以允许按不同分辨率解码视频的技术的流程图。

[0036] 图11是示出了用于解码空间可伸缩比特流层以允许按不同分辨率解码视频的技术的流程图。

[0037] 图12和13是示出了用于示例性多阶段位置计算技术的伪码的代码图。

[0038] 图14是示出了用于示例性增量位置计算技术的伪码的代码图。

具体实施方式

[0039] 描述的实施例涉及用于多分辨率的和分层的空间可伸缩视频编码和解码的技术和工具。

[0040] 本文描述的各种技术和工具可以独立使用。某些技术和工具也可以结合使用（例如，在组合的编码和 / 或解码过程的不同相位）。

[0041] 如下将参考处理动作的流程图描述各种技术。在流程图中示出的各种处理动作可以合并为更少的动作或者分割成更多的动作。为了简明，在特定流程图中示出的各动作与在其它地方描述的各动作之间的关系通常不示出。在许多情况下，可以重排流程图中的动作。

[0042] 大部分详细描述旨在表示、编码和解码视频信息。本文中描述的用于表示、编码和解码视频信息的技术和工具可以应用于音频信息、静止图像信息或其他媒体信息。

[0043] I. 计算环境

[0044] 图 1 示出了适合在其中实现所描述的若干实施例的合适计算环境 100 的一个通用示例。计算环境 100 并非旨在对使用范围或功能提出任何限制，因为这些技术和工具可以在完全不同的通用或专用计算环境中实现。

[0045] 参考图 1，计算环境 100 包括至少一个处理单元 110 和存储器 120。在图 1 中，这一最基本配置 130 包括在虚线内。处理单元 110 执行计算机可执行指令，且可以是真实或虚拟处理器。在多处理系统中，多个处理单元执行计算机可执行指令以提高处理能力。存储器 120 可以是易失性存储器（例如，寄存器、高速缓存、RAM）、非易失性存储器（例如，ROM、EEPROM、闪存等）或两者的某一组合。存储器 120 储存用此处所描述的一个或多个技术或工具实现视频编码器或解码器的软件 180。

[0046] 计算环境可具有额外的特征。例如，计算环境 100 包括存储 140、一个或多个输入设备 150、一个或多个输出设备 160 以及一个或多个通信连接 170。诸如总线、控制器或网络等互连机制（未示出）将计算环境 100 的组件互连。通常，操作系统软件（未示出）为在计算环境 100 中执行的其它软件提供了操作环境，并协调计算环境 100 的组件的活动。

[0047] 存储 140 可以是可移动或不可移动的，且包括磁盘、磁带或磁带盒、CD-ROM、DVD、闪存或可用于储存信息并可在计算环境 100 内访问的任何其它介质。存储 140 储存用于软件 180 实现视频编码器或解码器的指令。

[0048] 输入设备 150 可以是诸如键盘、鼠标、笔、触摸屏或跟踪球等触摸输入设备、语音输入设备、扫描设备或可向计算环境 100 提供输入的另一设备。对于音频或视频编码，输入设备 150 可以是声卡、显卡、TV 调谐卡、或接受模拟或数字格式的音频或视频输入的类似的设备、或将音频或视频样值读入计算环境 100 的 CD-ROM、CD-RW 或 DVD。输出设备 160 可以是显示器、打印机、扬声器、CD 或 DVD 刻录机、或从计算环境 100 提供输出的另一设备。

[0049] 通信连接 170 允许通过通信介质到另一计算实体的通信。通信介质传达诸如计算机可执行指令、音频或视频输入或输出、或已调制数据信号形式的其它数据等信息。已调制数据信号是其一个或多个特征以在信号中编码信息的方式设置或改变的信号。作为示例而非局限，通信介质包括以电、光、RF、红外、声学或其它载波实现的有线或无线技术。

[0050] 各种技术和工具可以在计算机可读介质的一般上下文中描述。计算机可读介质可以是可在计算环境内访问的任何可用介质。作为示例而非局限，对于计算环境 100，计算机

可读介质包括存储器 120、存储 140、通信介质以及上述任一个的组合。

[0051] 各种技术和工具可以在诸如程序模块中所包括的在计算环境中的一个或多个目标真实或虚拟处理器上执行的计算机可执行指令的一般上下文中描述。一般而言,程序模块包括例程、程序、库、对象、类、组件、数据结构等,它们执行特定任务或实现特定抽象数据类型。程序模块的功能可以如各实施例中所需的组合或在程序模块之间分离。用于程序模块的计算机可执行指令可以在本地或分布式计算环境中执行。

[0052] 为了呈现,详细描述使用了如“编码”、“解码”和“选择”等术语来描述计算环境中的计算机操作。这些术语是由计算机执行的操作的高级抽象,且不应与人类所执行的动作混淆。对应于这些术语的实际的计算机操作取决于实现而不同。

[0053] II. 示例性视频编码器和解码器

[0054] 图 2 是可结合其实现所描述的某些实施例的示例性视频编码器 200 的框图。图 3 是可结合其实现所描述的某些实施例的概括性视频解码器 300 的框图。

[0055] 编码器 200 和解码器 300 内的模块之间所示的关系指示了编码器和解码器中的一般信息流;为简明起见,未示出其它关系。具体地,图 2 和 3 一般不示出指示用于视频序列、图像、片、宏块、块等的编码器设置、模式、表等辅助信息。这一辅助信息通常在该辅助信息的熵编码之后在输出比特流中发送。输出比特流的格式可以取决于实现变化。

[0056] 编码器 200 和解码器 300 处理视频图像,视频图像可以是视频帧、视频半帧或帧和半帧的组合。图像和宏块级的比特流句法和语法可取决于使用了帧还是半帧。也可以对宏块组织和总体时序有改变。编码器 200 和解码器 300 是基于块的并且对帧使用 4:2:0 的宏块格式,其中每一宏块包括四个 8×8 的亮度块(有时候作为一个 16×16 的宏块来对待)以及两个 8×8 的色度块。对于半帧,可使用相同或不同的宏块组织和格式。8×8 的块还可在不同阶段细分,例如在频率变换和熵编码阶段。示例性视频帧组织在以下更详细描述。或者,编码器 200 和解码器 300 是基于对象的,使用不同的宏块或块格式,或对与 8×8 的块和 16×16 的宏块不同大小或配置的样本集执行操作。

[0057] 取决于所需的实现和压缩类型,编码器或解码器的模块可被添加、省略、分成多个模块、与其它模块组合、和/或用相似的模块来替代。在替换实施例中,具有不同模块和/或其它模块配置的编码器或解码器执行一个或多个所描述的技术。

[0058] A. 视频帧组织

[0059] 在某些实现中,编码器 200 和解码器 300 处理如下组织的视频帧。帧包含视频信号的空间信息行。对于逐行扫描视频,这些行包含表示从一同时刻立即采样一场景内容快照并从该帧上到下覆盖整幅场景的样值。逐行扫描视频帧被划分成诸如图 4 所示的宏块 400 等宏块。宏块 400 包括四个 8×8 的亮度块(Y1 到 Y4)以及两个 8×8 的色度块,这些色度块与四个亮度块共同定位,但是水平和垂直分辨率都是一半,遵循常规的 4:2:0 的宏块格式。8×8 的块还可在不同阶段细分,例如在频率变换(例如,8x4、4x8 或 4x4 DCT)和熵编码阶段。逐行扫描 I 帧是帧内编码的逐行扫描视频帧,其中术语“帧内”指的是不涉及从其他在前解码的图像内容进行预测的编码方法。逐行扫描 P 帧是使用从与当前图像时间上不同的时刻处的一个或多个图像的预测(在某些上下文中有时也称为前向预测)来编码的逐行扫描视频帧,而逐行扫描 B 帧是使用涉及在某些区域内(可能加权的)多个预测的平均的帧间预测(有时称为双预测或双向预测)来编码的逐行扫描视频帧。逐行扫描 P 帧和 B 帧可

包括帧内编码宏块以及各种类型的帧间预测宏块。

[0060] 隔行扫描视频帧由一场景的两种扫描类型的交替序列构成——一种包括帧的偶数行（行号为 0、2、4 等），称之为上半帧，而另一种包括帧的奇数行（行号 1、3、5 等），称之为下半帧。这两个半帧可表示两个不同的快照时间瞬间。图 5A 示出了隔行扫描视频帧 500 的一部分，包括位于隔行扫描视频帧 500 的左上部分的上半帧和下半帧的交替行。

[0061] 图 5B 示出了为编码 / 解码组织为帧 530 的图 5A 的隔行扫描视频帧 500。隔行扫描视频帧 500 已被划分成诸如宏块 531 和 532 等的宏块或其他这类区域，它们使用如图 4 所示的 4:2:0 的格式。在亮度平面中，每一宏块 531、532 包括来自上半帧的 8 行，这 8 行与来自下半帧的 8 行交替，总共有 16 行，且每一行是 16 个样本长。（将该图像实际组织到宏块或其他这类区域以及宏块 531、532 内亮度块和色度块的实际组织和布置未示出，且实际上对不同的编码决策以及对不同的视频编码设计可以有所变化。）在给定宏块内，上半帧信息和下半帧信息可以联合编码或在各相位处的任一个单独编码。

[0062] 隔行扫描 I 帧是包括两个半帧的帧内编码的隔行扫描视频帧，其中每个宏块包括关于一个或两个半帧的信息。隔行扫描 P 帧是包括两个半帧的使用帧间预测编码的隔行扫描视频帧，其中每个宏块包括关于一个或两个半帧的信息，同隔行扫描 B 帧一样。隔行扫描 P 帧和 B 帧可包括帧内编码的宏块以及各种类型的帧间预测宏块。

[0063] 图 5C 示出了为编码 / 解码而被组织成半帧 560 的图 5A 的隔行扫描视频帧 500。隔行扫描视频帧 500 的两个半帧中的每一个被划分成宏块。上半帧被划分成诸如宏块 561 等的宏块，下半帧被划分成诸如宏块 562 等的宏块。（同样地，这些宏块也使用如图 4 所示的 4:2:0 格式，并且将图像组织到宏块或其他这类区域以及各宏块内亮度块和色度块的布置未示出且可以变化。）在亮度平面中，宏块 561 包括来自上半帧的 16 行，而宏块 562 包括来自下半帧的 16 行，且每一行是 16 个样本长。

[0064] 隔行扫描 I 半帧是隔行扫描视频帧的单个单独表示的半帧。隔行扫描 P 半帧是使用帧间图像预测编码的隔行扫描视频帧的单个单独表示的半帧，同隔行扫描 B 半帧一样。隔行扫描 P 半帧和 B 半帧可包括帧内编码的宏块以及不同类型的帧间预测宏块。

[0065] 为编码 / 解码组织为半帧的隔行扫描视频帧可包括不同半帧类型的各种组合。例如，这样的帧可在上半帧和下半帧两者中具有相同的半帧类型（I- 半帧、P- 半帧或 B- 半帧），或者在每一半帧中具有不同的半帧类型。

[0066] 术语图像一般指的是源、已编码或已重构图像数据的帧或半帧。对于逐行扫描视频，图像通常是逐行扫描视频帧。对于隔行扫描视频，图像取决于上下文可以指隔行扫描视频帧、帧的上半帧、或帧的下半帧。

[0067] 图 5D 示出了相对于视频帧每个半帧的亮度样本位置的 4:2:0 色度样本位置的六个示例性空间排列。

[0068] 或者，编码器 200 和解码器 300 是基于对象的，使用不同的宏块或格式（例如，4:2:2 或 4:4:4）或块格式，或对与 8×8 的块和 16×16 的宏块不同大小或配置的样本集执行操作。

[0069] B. 视频编码器

[0070] 图 2 是示例性视频编码器系统 200 的框图。编码器系统 200 接收包括当前图像 205（例如，逐行扫描视频帧、隔行扫描视频帧或隔行扫描视频帧的半帧）的视频图像序列，

并产生压缩的视频信息 295 作为输出。视频编码器的具体实施例通常使用示例编码器 200 的变化或补充版本。

[0071] 编码器系统 200 为帧内编码（帧内）图像（I- 图像）和帧间图像预测（帧间）图像（P 或 B 图像）使用编码过程。为演示起见，图 2 示出了 I 图像的通过编码器系统 200 的路径以及用于帧间图像预测图像的路径。编码器系统 200 的许多组件用于同时压缩 I 图像和帧间预测图像两者。由这些组件执行的确切操作可以取决于所压缩的信息类型而变化。

[0072] 帧间图像预测图像按照来自一个或多个其它图像（通常被称为参考图像）的预测（或差）来表示。预测残差是所预测的和原始图像之差。相反，I 图像不参考其它图像来压缩。I 图像可以使用空间预测或频域预测（即，帧间图像预测）从来自 I 图像本身其他部分的数据来预测 I 图像的某些部分。但为了简明，这些 I 图像在本文中将被不称为“预测”图像，因而短语“预测”图像可被理解为帧间图像预测图像（例如，P 图像或 B 图像）。

[0073] 如果当前图像 205 是预测图像，则运动估计器 210 估计当前图像 205 的宏块或其它样本集相对于一个或多个参考图像（例如，缓冲在图像存储 220 中的重构的前一图像 225）的运动。运动估计器 210 可以相对于一个或多个时间上在前的参考图像和一个或多个时间上在将来的参考图像（例如，在双向预测图像的情况下）估计运动。因此，编码器系统 200 可使用单独的存储 220 和 222 用于多个参考图像。

[0074] 运动估计器 210 可按照全样本、1/2 样本、1/4 样本或其它增量来估计运动，并可在逐图像的基础或其它基础上切换运动估计的分辨率。运动估计器 210（和补偿器 230）也可在每一帧或其它基础上在参考图像样本内插的类型之间切换（例如，在三次卷积内插和双线性内插之间）。运动估计的分辨率可以在水平和垂直上相同或不同。运动估计器 210 输出运动信息 215，诸如差分运动矢量信息作为辅助信息。编码器 200 通过例如为运动矢量计算一个或多个预测值，计算运动矢量和预测值之间的差，以及对差进行熵编码，来对运动信息 215 进行编码。为重构运动矢量，运动补偿器 230 将预测值与运动矢量差信息组合。

[0075] 运动补偿器 230 将重构的运动矢量应用于重构的图像 225，以形成经运动补偿的预测 235。然而，预测很少是完美的，且经运动补偿的预测 235 和原始的当前图像 205 之间的差是预测残差 245。在稍后的图像重构期间，将预测残差 245 的近似添加到经运动补偿的预测 235，以获得比经运动补偿预测 235 更接近于原始的当前图像 205 的重构图像。然而，在有损压缩中，某些信息仍从原始当前图像 205 中丢失。或者，运动估计器和运动补偿器应用另一类型的运动估计 / 补偿。

[0076] 频率变换器 260 将空间域视频信息转换成频域（即，频谱）数据。对于基于块的视频编码，频率变换器 260 通常向样本数据或预测残差数据的块应用离散余弦变换（DCT）、DCT 的变体或某些其它块变换，从而产生频域变换系数块。或者，频率变换器 260 应用诸如傅立叶变换等另一常规频率变换类型或使用小波或子带分析。频率变换器 260 可应用 8×8 、 8×4 、 4×8 、 4×4 或其它大小的频率变换。

[0077] 量化器 270 随后量化频域变换系数块。量化器根据在逐图像的基础、宏块基础或某些其它基础上变化的量化步长向变换系数应用标量量化，其中量化步长是管理在解码器反量化器过程中在可离散表示重构点之间均匀隔开的间隔的控制参数，而这也可以在编码器反量化器过程 276 中重复。或者，量化器向频域变换系数应用另一类型的量化，例如带有非均匀重构点的标量量化器、矢量量化器或非自适应量化，或直接在不使用频率变换的编

码器系统中量化空间域数据。除自适应量化之外,编码器 200 可为速率控制使用帧丢弃、自适应滤波或其它技术。

[0078] 当需要重构的当前图像用于后续的运动估计 / 补偿时,反量化器 276 在量化的频域变换系数上执行反量化。反频率变换器 266 然后执行频率变换器 260 的逆运算,从而产生重构的预测残差的近似(对于预测图像)或重构的 I 图片的近似。如果当前图像 205 是 I 图像,则重构的 I 图像的近似用作重构的当前图像的近似(未示出)。如果当前图像 205 是预测图像,则重构的预测残差的近似被添加到经运动补偿的预测 235,以形成重构的当前图像的近似。一个或多个图像存储 220、222 缓冲重构的当前图像的近似,以在后续图像的经运动补偿预测中用作参考图像。编码器可以在将图像近似存储到一个或多个图像存储 220、222 之前,向重构帧应用解块滤波器或其他图像细化过程以便自适应地平滑来自图像的不连续性并移除其它伪像。

[0079] 熵编码器 280 压缩量化器 270 的输出以及某些辅助信息(例如,运动信息 215、量化步长)。典型的熵编码技术包括算术编码、差分编码、哈夫曼编码、行程长度编码、Lempel-Ziv 编码、字典式编码以及上述的组合。熵编码器 280 通常对不同种类的信息(例如,低频系数、高频系数、零频系数、不同种类的辅助信息)使用不同的编码技术,并可从特定编码技术内的多个代码表中进行选择。

[0080] 熵编码器 280 向多路复用器 [“MUX”] 290 提供压缩的视频信息 295。MUX 290 可包括缓冲器,并且缓冲器完整性级别指示符可被反馈给比特率自适应模块用于速率控制。在 MUX 290 之前或之后,压缩的视频信息 295 可被信道编码用于通过网络发送。信道编码可向压缩的视频信息 295 应用检错和纠错数据。

[0081] C. 视频解码器

[0082] 图 3 是示例性视频解码器系统 300 的框图。解码器系统 300 接收关于压缩的视频图像序列的信息 395,并产生包括重构的图像 305(例如,逐行扫描视频帧、隔行扫描视频帧或隔行扫描视频帧的半帧)的输出。视频解码器的具体实施例通常使用概括性解码器 300 的变体或补充版本。

[0083] 解码器系统 300 解压预测图像和 I 图像。为演示起见,图 3 示出了通过解码器系统 300 用于 I 图片的路径以及用于预测图像的路径。解码器系统 300 的许多组件用于解压 I 图像和预测图像两者。由这些组件执行的确切操作可以取决于所解压的信息类型而变化。

[0084] DEMUX 390 接收关于压缩的视频序列的信息 395,并使得所接收的信息对熵解码器 380 可用。DEMUX 390 可包括抖动缓冲器以及其它缓冲器。在 DEMUX 390 之前或之中,压缩的视频信息可以被信道解码,并被处理用于检错和纠错。

[0085] 熵解码器 380 对熵编码的量化数据以及熵编码的辅助信息(例如,运动信息 315、量化步长)进行熵解码,通常应用编码器中执行的熵编码的逆运算。熵解码技术包括算术解码、差分解码、哈夫曼解码、行程长度解码、Lempel-Ziv 解码、字典式解码以及上述的组合。熵解码器 380 通常对不同种类的信息(例如,低频系数、高频系数、零频系数、不同种类的辅助信息)使用不同的解码技术,并可从特定解码技术内的多个代码表中进行选择。

[0086] 解码器 300 通过例如为运动矢量计算一个或多个预测值、对运动矢量差进行熵解码(在熵解码器 380 处)、以及将解码的运动矢量差与预测值组合以重构运动矢量,来对运动信息 315 进行解码。

[0087] 运动补偿器 330 向一个或多个参考图像 315 应用运动信息 325, 以形成所重构的图像 305 的预测 335。例如, 运动补偿器 330 使用一个或多个宏块运动矢量以找出样本块或者在参考图像 325 中的样本之间内插分数位置。一个或多个图像存储 (例如, 图像存储 320、322) 存储先前重构的图像以用作参考图像。通常, B 图像具有一个以上参考图像 (例如, 至少一个时间上先前的参考图像以及至少一个时间上未来的参考图像)。因此, 解码器系统 300 可对多个参考图像使用单独的图像存储 320 和 322。运动补偿器 330 可以按全样本、1/2 样本、1/4 样本或其它增量来补偿运动, 并可在逐图像的基础或其它基础上切换运动补偿的分辨率。运动补偿器 330 也可在每一帧或其它基础上在参考图像样本内插的类型之间 (例如, 在三次卷积内插和双线性内插之间) 切换。运动补偿的分辨率可以在水平和垂直上相同或不同。或者, 运动补偿器应用另一类型的运动补偿。运动补偿器的预测很少是完美的, 因此解码器 300 也重构预测残差。

[0088] 反量化器 370 对熵解码的数据进行反量化。一般而言, 反量化器向熵解码的数据应用均匀的标量反量化, 其中重构步长在逐图像的基础上、宏块基础上或某些其它基础上变化。或者, 反量化器向数据应用另一类型的反量化, 例如用于在非均匀矢量或非自适应反量化, 或直接在不使用反频率变换的解码器系统对空间域数据进行反量化。

[0089] 反频率变换器 360 将反量化频域变换系数转换成空间域视频信息。对于基于块的视频图像, 反频率变换器 360 向频率变换系数块应用反 DCT [“IDCT”]、IDCT 的变体或其它反块变换, 从而分别对 I 图像或预测图像产生样本数据或帧间图像预测残差数据。或者, 反频率变换器 360 应用另一类型的反频率变换, 诸如傅立叶反变换或使用小波或子带合成。反频率变换器 360 可应用 8×8 、 8×4 、 4×8 、 4×4 或其它大小的反频率变换。

[0090] 对于预测图像, 解码器 300 将重构的预测残差 345 与经运动补偿的预测 335 组合, 以形成重构的图像 305。当解码器需要重构的图像 305 用于后续的运动补偿时, 一个或多个图像存储 (例如, 图像存储 320) 缓冲重构的图像 305 以供预测下一图像时使用。在某些实施例中, 解码器 300 可以在将重构图像 305 存储到一个或多个图像存储 (例如, 图像存储 320) 之前或者在解码视频放出期间显示解码图像之前, 向重构图像应用解块滤波器或其他图像细化过程以便自适应地平滑来自图像的不连续性并移除其它伪像。

[0091] III. 多分辨率编码和解码的概览

[0092] 视频可以按不同的分辨率编码 (和解码)。出于本描述的目的, 多分辨率编码和解码可以被描述为基于帧的编码和解码 (例如, 参考图像重新采样) 或者分层的 (有时也称为可空间缩放的) 编码和解码。多分辨率编码和解码还可以涉及隔行扫描视频、基于半帧的编码和解码、以及在指定分辨率基础上或某些其他基础上在基于帧和基于半帧的编码和解码之间的切换。然而, 在此概览中出于简化概念描述的目的讨论了进行逐行扫描视频的帧编码。

[0093] A. 基于帧的多分辨率编码和解码

[0094] 在基于帧的多分辨率编码中, 编码器以不同的分辨率编码输入图像。编码器在逐图像的基础上或在某些其他基础上为各图像选择空间分辨率。例如, 在参考图像重新采样时, 如果该图像以不同于当前编码的分辨率编码则可以对参考图像进行重新采样。术语重新采样用于描述用于表示图像区域或者采样信号的某些其他部分的样本数增加 (上采样) 或降低 (下采样)。每单位面积或者每信号部分的样本数被称为采样的分辨率。

[0095] 空间分辨率可以基于例如可用比特率的减小 / 增大、量化步长的减小 / 增大、输入视频内容运动量的减小 / 增大、视频内容的其他属性（例如，在较低分辨率下可能会明显失真的强边缘、文本或其他内容的呈现）或在某些其他基础上来选择。空间分辨率可以在垂直、水平或者垂直和水平维度两者上变化。水平分辨率可以与垂直分辨率相同或不同。解码器使用互补技术对编码帧进行解码。

[0096] 一旦编码器已经为当前图像或者当前图像内的区域选择了空间分辨率，编码器将在对原始图像编码前将其重新采样至期望的分辨率。编码器随后可以发信号通知解码器这一空间分辨率的选择。

[0097] 图 6 示出了基于帧的多分辨率图像编码的技术 (600)。诸如图 2 所示编码器的编码器为一图像设置分辨率 (610)。例如，编码器考虑以上列出准则或其他准则。编码器随后按此分辨率编码图像 (620)。如果完成了要编码的所有图像的编码 (630)，则编码器退出。如果没有，则编码器就为下一幅图像设置分辨率 (610) 并继续编码。作为替换，编码器可以在图像级以外的某些级别上设置分辨率，诸如为图像的不同部分不同地设置分辨率或者为一组或一系列图像作出分辨率选择。

[0098] 编码器可以编码预测图像以及帧内图像。图 8 示出了基于帧的多分辨率帧内图像和图像间预测图像编码的技术 (800)。首先，编码器在 810 检查要编码的当前图像是帧内图像还是预测图像。如果当前图像是帧内图像，编码器就在 820 设置针对当前图像的分辨率。如果图像是预测图像，编码器就在 830 针对当前图像设置分辨率之前设置针对参考图像的分辨率。在为当前图像设置分辨率之后，编码器就按该分辨率编码当前图像 (840)。为一图像设置分辨率（无论是当前源图像或者是存储的参考图像）可涉及对该图像的重新采样以便与所选分辨率相匹配并且可以涉及对信号进行编码以将该所选分辨率指示给解码器。如果完成了要编码的所有图像的编码 (850)，则编码器退出。如果没有，编码器继续编码另外的图像。作为替换，编码器以不同的方式处理预测图像。

[0099] 解码器对经编码图像进行解码，并且如有必要，在显示前重新采样该图像。与经编码的图像的分辨率相类似，解码图像的分辨率可以用许多不同方式调整。例如，解码图像的分辨率可被调整为适应输出显示设备或者输出显示设备一区域（例如，用于“画中画”或 PC 桌面窗口显示）的分辨率。

[0100] 图 7 示出了基于帧的多分辨率图像解码的技术 (700)。诸如图 3 所示解码器的解码器为图像设置分辨率（在 710）。例如，解码器从编码器获取分辨率信息。解码器随后按此分辨率解码图像 (720)。如果完成了要解码的所有图像的编码 (730)，则解码器退出。如果没有，则解码器就为下一幅图像设置分辨率 (710) 并继续解码。作为替换，解码器在图像级以外的某些级别上设置分辨率。

[0101] 解码器可以解码预测图像以及帧内图像。图 9 示出了基于帧的多分辨率帧内图像和预测图像解码的技术 (900)。

[0102] 首先，解码器检查要解码的当前图像是帧内图像还是预测图像 (910)。如果当前图像是帧内图像，解码器设置针对当前图像的分辨率 (920)。如果图像是预测图像，编码器就在针对当前图像设置分辨率 (920) 之前设置针对参考图像的分辨率 (930)。设置参考图像的分辨率可以涉及对存储的参考图像进行重新采样以匹配所选分辨率。在为当前图像设置分辨率 (920) 之后，解码器就按该分辨率解码当前图像 (940)。如果完成了要解码的所有图

像的解码 (950), 则解码器退出。如果没有, 解码器继续解码。

[0103] 解码器通常按编码器使用的同一分辨率来解码图像。作为替换, 解码器可以按不同的分辨率解码图像, 诸如在解码器可用的分辨率无法与编码器中使用的分辨率完全相同的情况下。

[0104] B. 分层的多分辨率编码和解码

[0105] 在分层的多分辨率编码中, 编码器在各层中编码视频, 其中各层具有按不同分辨率解码视频的信息。以此方式, 编码器以一个以上的分辨率来对视频内的至少某些独立图像进行编码。解码器随后可以通过处理各层的不同组合来按一个或多个分辨率解码该视频。例如, 第一层 (有时被称为基层) 含有用于按较低分辨率解码视频的信息, 而一个或多个其他层 (有时被称为增强层) 含有用于按较高分辨率解码视频的信息。

[0106] 可以将基层本身设计为可独立解码的比特流。于是, 在这一设计中, 仅解码基层的解码器将按基层的较低分辨率生成有效解码的比特流。使用增强层恰当解码较高分辨率图像可能还需要解码部分或全部的经编码基层数据以及可能的一个或多个增强层。解码基层和一个或多个其他更高分辨率层的解码器将能够比仅解码基层的解码器生成更高分辨率的内容。两、三或更多层可用来允许两、三或更多不同的分辨率。作为替换, 更高分辨率层本身也可以是能独立解码的比特流。(这一设计通常被称为同时联播多分辨率编码方法。)

[0107] 图 10 示出了用于编码比特流层以便允许按不同分辨率解码的技术 (1000)。诸如图 2 所示编码器 200 的编码器获取全分辨率视频信息作为输入 (1010)。编码器下采样全分辨率视频信息 (1020) 并使用下采样信息编码基层 (1030)。编码器使用基层和更高分辨率视频信息编码一个或多个更高分辨率层 (1040)。更高分辨率层可以是允许按全分辨率解码的层或者是按某一中间分辨率解码的层。编码器随后输出包括两个以上编码层的分层比特流。作为替换, 更高分辨率层的编码 (1040) 可能不使用基层信息, 于是就能够为同时联播多分辨率编码方法来独立解码更高分辨率层的数据。

[0108] 编码器能够按照图 10 所示的基本概要用各种方式来实现多分辨率层的编码。更多信息, 请参见例如美国专利 No. 6, 510, 177, 或者 MPEG-2 标准或其他视频标准。

[0109] 图 11 示出了用于解码比特流层以便允许按不同分辨率解码视频的技术 (1100)。诸如图 3 所示解码器 300 的解码器获取分层比特流作为输入 (1110)。各层包括较低分辨率层 (基层) 和一个或多个包含更高分辨率信息的层。更高分辨率层无需包含可独立编码的图像; 一般而言, 更高分辨率层包括描述各图像较高和较低分辨率版本之间差异的残差信息。解码器解码基层 (1120), 并且如果期望更高分辨率的解码, 则解码器将解码的基层图像下采样至期望的分辨率 (1130)。解码器解码一个或多个较高分辨率层 (1140) 并将解码的较高分辨率信息与下采样的解码基层图片相结合以形成较高分辨率的图像 (1150)。取决于期望的分辨率级, 更高分辨率图像可以是全分辨率图像或者中间分辨率图像。更多信息, 请参见例如美国专利 No. 6, 510, 177, 或者 MPEG-2 标准或其他视频标准。

[0110] 解码器通常按编码器使用分辨率之一来解码图像。作为替换, 解码器可用的分辨率可以不与编码器中使用的分辨率完全相同。

[0111] IV. 可伸缩视频编码和解码的重新采样滤波器

[0112] 这一部分描述用于可伸缩视频编码和解码的技术和工具。虽然所述的一些技术和工具是在分层 (或空间可伸缩) 的上下文中描述的, 但是所述的一些技术和工具也可以在

基于帧（或者参考图像采样）的上下文中或者在涉及重新采样滤波器的某些其他上下文中使用。此外，虽然所述的一些技术和工具可以在重新采样图像的上下文中描述，但是所述的一些技术和工具也可以用于对来自更高分辨率信号预测的残差或差分信号进行重新采样。

[0113] 可伸缩视频编码 (SVC) 是允许对一较大比特流的子集进行解码以生成其帧质量在某些应用中可接受的解码图像（虽然这些图像质量要低于解码整个较高比特率比特流而生成的质量）的一种数字视频编码类型。一种周知类型的 SVC 被称为空间可伸缩性，或者分辨率可伸缩性。在空间 SVC 设计中，编码过程（或者在编码过程前执行的预处理功能，这取决于编码过程范围的精确定义）通常包括将该视频下采样至一较低分辨率并对该较低分辨率视频进行编码以启用较低分辨率解码过程，同时上采样该较低分辨率的解码图像来用作更高分辨率视频图像中各样本值的预测。针对更高分辨率视频的解码过程随后包括解码较低分辨率视频（或其部分），并使用上采样的视频作为更高分辨率视频图像中各样本值的预测。这些设计要求使用重新采样滤波器。更具体地，编解码器的设计包括在解码器和编码器两者中使用上采样滤波器并在编码器或编码预处理器中使用下采样滤波器。对这类设计中使用的上采样滤波器予以特别关注。一般而言，上采样过程被设计为在编码器和解码器中相同以防止漂移现象，所谓漂移是由于在编码和解码期间使用同一信号的不同预测而导致的误差累积。

[0114] 某些空间 SVC 设计的一大缺点是在解码过程中使用低质量滤波器（例如，双抽头双线性滤波器）。使用更高质量的滤波器会有益于视频质量。

[0115] 空间 SVC 可以包括能在滤波器的重新采样比率内实现更高程度灵活性的重新采样滤波器。然而，这可能需要在编码器和解码器实现中针对要开发的这一滤波器的每个不同“相位”以及要存储的这些滤波器的“抽头”值的大量特定的滤波器设计。

[0116] 此外，这对视频质量有益以允许编码器控制用于空间 SVC 的重新采样滤波器的模糊的量。于是，对设计用于上采样或下采样的重新采样的每个“相位”，取决于要在该过程中引入的模糊的期望程度来从若干不同的滤波器进行的选择会是有利的。要在上采样期间执行的模糊程度的选择可以作为传达以供解码过程使用的信息而从编码器发送至解码器。这一额外的灵活性使设计进一步复杂，因为它大幅增加来需要在编码器或解码器中存储的必要抽头值的数量。

[0117] 可以使用统一的设计来指定具有各种相位和各种模糊程度的各种重新采样的滤波器。一种可能的解决方案是使用 Mitchell-Netravali 滤波器设计方法。Mitchell-Netravali 滤波器设计方法对这些问题的直接应用可能会以针对要在编码器或解码器中计算的量的过大可能值动态范围的形式来要求过多的计算资源。例如，一种这样的设计可能要求使用 45 位算术处理，而不是在通用 CPU 和 DSP 中常用的 16 位或 32 位处理元件。为解决这一问题，提供了一些设计改进。

[0118] 典型的 SVC 设计要求针对空间可伸缩性的标准化上采样滤波器。为了支持任意重新采样比率（已知为扩展的空间可伸缩性特征），描述了一种结合了关于重新采样比率的大量灵活性的上采样滤波器设计。另一个关键方面是亮度和色度的相对对准。当在单层方法中发现各种对准结构（参见例如针对 4:2:0 色度的 H. 261/MPEG-1 相对于 MPEG-2 对准以及 H. 264/MPEG-4 AVC）时，所描述的技术和工具以一种编码器能够轻易向解码器指示如何恰当应用滤波的方式支持各类灵活对准。

[0119] 所述的技术和工具包括能够进行高质量上采样和良好抗图形失真的上采样滤波器。更具体地,所描述的技术和工具具有的质量要优于由在前针对空间可伸缩性的双线性滤波器设计提供的质量。所描述的技术和工具具有视觉上合意并能提供良好信号处理频率行为的高质量上采样滤波器。所描述的技术和工具包括指定简单且不要求大型存储器存储表来保持抽头值的滤波器设计,并且滤波操作本身在计算上也操作简单。例如,所描述的技术和工具具有不过于冗长且不要求过度数学精度或极度复杂数学函数的滤波器。

[0120] 这一部分描述具有一个或多个下列特征的设计:

- [0121] - 亮度 / 色度相位对准的灵活性;
- [0122] - 重新采样比率的灵活性;
- [0123] - 频率特性的灵活性;
- [0124] - 高视觉质量;
- [0125] - 不太少也不太多的滤波器抽头(例如,在 4 到 6 之间);
- [0126] - 指定简单;
- [0127] - 操作简单(例如,使用实用字长的算术)。

[0128] A. Mitchell-Netravali 上采样滤波器

[0129] 所描述的技术和工具采用可分的滤波方法——因此随后的讨论将主要集中在二维信号的处理,因为二维情况是一维情况的简单可分应用。首先基于由下式给出的概念连续脉冲响应 $h(x)$ 提出一组两参数滤波器集合:

$$[0130] \quad h(x) = \frac{1}{6} * \begin{cases} (12-9b-6c)|x|^3 - (18-12b-6c)|x|^2 + (6-2b) & |x| < 1 \\ -(b+6c)|x|^3 + (6b+30c)|x|^2 - (12b+48c)|x| + (8b+24c) & 1 \leq |x| < 2 \\ 0 & \text{otherwise} \end{cases} \quad (1),$$

[0131] 其中 b 和 c 是那两个参数 (otherwise 即其他)。对于相对相位偏移量位置 $0 \leq x < 1$, 该核生成由以下矩阵等式给出的抽头值的 4 抽头有限脉冲响应 (FIR) 滤波器:

$$[0132] \quad \frac{1}{6} * \begin{bmatrix} 1 & x & x^2 & x^3 \end{bmatrix} * \begin{bmatrix} b & 6-2b & b & 0 \\ -(3b+6c) & 0 & 3b+6c & 0 \\ 3b+12c & -18+12b+6c & 18-15b-12c & -6c \\ -(b+6c) & 12-9b-6c & -(12-9b-6c) & b+6c \end{bmatrix} \quad (2)$$

[0133] 实际上,只有考虑 x 的范围从 0 到 $1/2$ 就够了,因为用于 x 的 FIR 滤波器核完全就是逆序的用于 $1-x$ 的 FIR 滤波器核。

[0134] 这一设计具有多个感兴趣且有用的属性。如下列出其中的一部分:

[0135] - 计算该滤波器抽头值不需要三角函数、超越函数或者无理数处理。实际上,这一滤波器的抽头值可以用很少的简单操作直接计算。没有必要为要使用的各种可能参数值和相位存储这些抽头值,因为在需要是可以简单计算这些值。(因此,为了标准化这些滤波器的使用,只需要少量公式——而不需要对类似余弦或贝塞尔函数的近似函数的多次或标准化尝试的大型表。

[0136] - 所得的滤波器具有 4 个抽头。这是非常实用的数目。

[0137] - 滤波器在其主波瓣的每一侧只有单个旁瓣。于是就不会生成过度的边缘振荡效应。

[0138] - 滤波器具有平滑的脉冲响应。其值及其一阶导数都是连续的。

[0139] - 它具有单位增益 DC 响应,意味着被上采样的信息种不存在整体的亮度放大或衰减。

[0140] - 这些滤波器系列的成员包括对周知良好滤波器的相对良好的近似,诸如“Lanczos-2”设计和“Catmull-Rom”设计。

[0141] 此外,所描述的技术和工具包括在用于选择视觉合意滤波器的两参数之间的特定关系。这一关系可如下表达为:

$$[0142] \quad c = \frac{1}{2}(1-b) \quad (3)$$

[0143] 这将自由度降低至单个带宽控制参数 b 。这一参数控制由滤波器引入的额外模糊程度。注意与值 $b = 0$ 相关联的这一系列的成员是优秀且周知的 Catmull-Rom 上采样滤波器(也被称为密钥“三次卷积”内插滤波器)。

[0144] 除了 Mitchell-Netravali 滤波器系列的所有成员都能找到的基本优点之外, Catmull-Rom 上采样滤波器本身还具有诸多良好的属性:

[0145] - 它是“内插”滤波器——即,对于相位值 $x = 0$ 和 $x = 1$ 而言,滤波器具有等于 1 的单个非零抽头。换句话说,上采样信号将会正好在每个上采样曲线段的边缘处通过输入样本的值。

[0146] - 如果输入样本的设置形成抛物线(或直线或静态值),输出点将正好落在抛物线曲线(或直线或静态值)上。

[0147] 实际上,用某些方式, Catmull-Rom 上采样器出于这些原因可被认为是这一长度的最佳上采样滤波器——虽然引入某些额外模糊(增加 b) 有时可能更为视觉合意。同样地,引入某些额外模糊可以有助于抹掉低比特率压缩伪像,于是就能够更相似地充当真实上采样图像的 Wiener 滤波器(用于噪声滤波的周知滤波器)估计器。

[0148] 简单地将等式 (3) 代入等式 (2) 可得如下抽头值:

$$[0149] \quad \frac{1}{6} * \begin{bmatrix} 1 & x & x^2 & x^3 \end{bmatrix} * \begin{bmatrix} b & 6-2b & b & 0 \\ -3 & 0 & 3 & 0 \\ 6-3b & -15+9b & 12-9b & -3+3b \\ -(3-2b) & 9-6b & -(9-6b) & 3-2b \end{bmatrix} \quad (4).$$

[0150] 据报告,基于 9 位专家观察者和超过 500 个样本的主观性测试,可得:

[0151] - 可用范围被报告为 $0 \leq b \leq 5/3$;

[0152] - $0 \leq b \leq 1/2$ 被归类为视觉上“令人满意”,当 $b = 1/3$ 时报告为视觉合意;

[0153] - $b > 1/2$ 被归类为“模糊”,当 $b = 3/2$ 时报告为非常模糊

[0154] B. 带宽控制参数的整数化

[0155] 在等式 (4) 中被 6 除可能是不合宜的。相反,希望整数化带宽控制参数和滤波器抽头值,因为作为解码器设计一部分的无限精度是不可行的。考虑使用如下定义的新整数值变量 a 代入:

$$[0156] \quad a = (b/6) * 2^S \quad (5),$$

[0157] 其中 S 是整数移位因子而 a 是用作整数化带宽控制参数的无符号整数。参数 a 可以由编码器在比特流中的视频序列级处编码为语法元素。例如,参数 a 可以用可变长度或固定长度代码来显式编码,共同地用其他信息编码,或者显式地作为信号通知。可选地,参

数 a 可以在比特流中的某些其他级处作为信号通知。

[0158] 整数化导致整数化的抽头值：

$$[0159] \quad \begin{bmatrix} 1 & x & x^2 & x^3 \end{bmatrix} * \begin{bmatrix} a & 2^S - 2a & a & 0 \\ -2^{S-1} & 0 & 2^{S-1} & 0 \\ 2^S - 3a & -5*2^{S-1} + 9a & 2^{S+1} - 9a & -2^{S-1} + 3a \\ -(2^{S-1} - 2a) & 3*2^{S-1} - 6a & -(3*2^{S-1} - 6a) & 2^{S-1} - 2a \end{bmatrix} \quad (6)$$

[0160] 该结果随后需要在二进制算术处理中按比例减小 S 个位置。

[0161] 如果 a 的范围从 0 到 M , b 的范围则从 0 到 $6*M/2^S$ 。对 M 的某些可能有用的选择包括以下各项：

[0162] $-M = 2^{(S-2)} - 1$, 得到 b 的范围从 0 到 $3/2 - 6/2^S$ 。

[0163] $-M = \text{Ceil}(2^S/6)$, 返回大于或等于 $2^S/6$ 的最小整数, 得到 b 的范围从 0 到略大于 1。

[0164] $-M = 2^{(S-3)} - 1$, 导致 b 的近似范围从 0 到 $3/4 - 6/2^S$ 。

[0165] 这些对 M 的选择大到足以覆盖大多数有用情况, 其中第一个选择 ($M = 2^{(S-2)} - 1$) 是三个选择中较大的。 S 的有用范围在 6 到 8 之间。例如, 考虑 $S = 7$ 和 $M = 2^{(S-2)} - 1$, 即 $M = 31$ 。作为替换, 也可以使用 M 和 S 的其他值。

[0166] C. 分数样本定位的整数化

[0167] 接下来, 考虑 x 值的粒度。为了实用性, 也应该近似 x 。例如, 可以如下定义整数 i ：

$$[0168] \quad x = i \div 2^F \quad (7)$$

[0169] 其中 F 表示所支持的分数样本位置精度。对于一个足够精确的重新采样操作示例而言, 考虑 $F \geq 4$ (1/16 或更高的样本定位精度)。这就得到以下的整数化滤波器抽头值：

$$[0170] \quad \begin{bmatrix} 1 & i & i^2 & i^3 \end{bmatrix} * \begin{bmatrix} a*2^{3F} & (2^S - 2a)*2^{3F} & a*2^{3F} & 0 \\ -2^{2F+S-1} & 0 & 2^{2F+S-1} & 0 \\ (2^S - 3a)*2^F & (-5*2^{S-1} + 9a)*2^F & (2^{S+1} - 9a)*2^F & (-2^{S-1} + 3a)*2^F \\ -(2^{S-1} - 2a) & 3*2^{S-1} - 6a & -(3*2^{S-1} - 6a) & 2^{S-1} - 2a \end{bmatrix} \quad (8)$$

[0171] 例如, 考虑 $F = 4$ 。该结果随后需要按比例减小 $3F+S$ 个位置。

[0172] 注意到上述矩阵中的每个元都含有因子 2 (假设 S 大于 1)。于是能够如下代替地将抽头值公式化为：

$$[0173] \quad \begin{bmatrix} 1 & i & i^2 & i^3 \end{bmatrix} * \begin{bmatrix} a*2^{3F-1} & (2^{S-1} - a)*2^{3F} & a*2^{3F-1} & 0 \\ -2^{2F+S-2} & 0 & 2^{2F+S-2} & 0 \\ (2^S - 3a)*2^{F-1} & (-5*2^{S-1} + 9a)*2^{F-1} & (2^{S+1} - 9a)*2^{F-1} & (-2^{S-1} + 3a)*2^{F-1} \\ -(2^{S-2} - a) & 3*(2^{S-2} - a) & -3*(2^{S-2} - a) & 2^{S-2} - a \end{bmatrix} \quad (9),$$

[0174] 其中每个抽头值都已被除以 2。该结果随后仅需要按比例减小 $3F+S-1$ 个位置。

[0175] 对于按比例减小, 定义函数 $\text{RoundingRightShift}(p, R)$ 作为为输入值 p 计算的右移 R 位的输出 (带舍入), 其计算如下：

$$[0176] \quad \text{RoundingRightShift}(p, R) = \begin{cases} (p + 2^{R-1}) \gg R & \text{for } R = 2, 3, 4, \dots \\ (p) \gg R & \text{for } R = 0 \text{ or } 1 \end{cases} \quad (10)$$

[0177] 其中符号“>>”指的是使用 2 的补码二进制算术的二进制算术右移算符。作为替换,不同地执行舍入右移。

[0178] 如下提供舍入右移的一些示例性应用。

[0179] D. 动态范围考虑

[0180] 如果用 N 位样本位长度滤波图像并在执行任何舍入之前两维地这般执行,那么在将结果按比例减小 $2*(3F+S-1)$ 个位置并将输出限制为 N 位范围之前,将会在累加器内需要 $2*(3F+S-1)+N+1$ 位的动态范围。例如,如果 $F = 4$, $S = 7$ 且 $N = 8$,则需要使用 45 位的累加器来计算滤波的结果。

[0181] 会在以下各分段中讨论一些缓解这一问题的方法。这些方法可以彼此分开或彼此结合使用。应该理解基于本文对所述动态范围缓解方法进行变化是可能的。

[0182] 1. 第一示例性动态范围缓解方法

[0183] 考虑首先执行水平滤波随后进行垂直滤波的示例。考虑对二维处理流水线中任何点的 W 位最大字长。在第一动态范围缓解方法中,为了实现滤波,在该过程的第一(水平)阶段输出处使用 R_H 位舍入右移,而在该过程的第二(垂直)阶段输出处使用 R_V 位舍入右移。

[0184] 于是可以如下计算:

$$[0185] \quad 2*(3F+S-1)+N+1-R_H = W \quad (11),$$

[0186] 因此

$$[0187] \quad R_H = 2*(3F+S-1)+N+1-W \quad (12).$$

[0188] 随后从下式计算第二(垂直)阶段的右移:

$$[0189] \quad R_H+R_V = 2*(3F+S-1) \quad (13),$$

[0190] 因此

$$[0191] \quad R_V = 2*(3F+S-1)-R_H \quad (14).$$

[0192] 例如,对于 $F = 4$ 且 $S = 7$ 且 $N = 8$ 且 $W = 32$,得到 $R_H = 13$ 且 $R_V = 23$ 。于是,代替 45 位的动态范围,用舍入右移将动态范围缩减到 32 位。可以为不同的 W 值使用不同的右移位数。

[0193] 2. 第二示例性动态范围缓解方法

[0194] 第二动态范围缓解方法涉及降低抽头值的精度而非该相位定位的精度(即,缩减 F),降低滤波器带宽调整参数的粒度(即,缩减 S)或降低第一阶段输出的精度(即,增加 R_H)。

[0195] 将等式 (9) 生成的四个整数抽头值表示为 $[t_{-1}, t_0, t_1, t_2]$ 。注意到四个滤波器抽头值之和将会等于 2^{3F+S-1} ,即: $t_{-1}+t_0+t_1+t_2 = 2^{3F+S-1}$

[0196] (15).

[0197] 这是该示例性动态范围缓解方法的一个重要性质,因为只要当四个输入样本具有相同的值,输出就会具有相同的值。

[0198] 使用在等式 (10) 中找出的舍入右移的示例性定义并给出用于抽头值的右移量 R_t ,则如下定义:

$$[0199] \quad u_{-1} = \text{RoundingRightShift}(t_{-1}, R_t);$$

$$[0200] \quad u_1 = \text{RoundingRightShift}(t_1, R_t);$$

[0201] $u_2 = \text{RoundingRightShift}(t_2, R_t)$;

[0202] $u_0 = 2^{3F+S-1} - u_{-1} - u_1 - u_2$.

[0203] 随后用抽头值 $[u_{-1}, u_0, u_1, u_2]$ 而非 $[t_{-1}, t_0, t_1, t_2]$ 来执行滤波。 R_t 的值每增加 1 就表示算术累加器中所需的动态范围减少 1 位, 同时在后续处理阶段中要执行的右移也减少 1 位。

[0204] 3. 第三示例性动态范围缓解方法

[0205] 前一设计使用类似于第一示例性动态范围缓解方法概念的手段, 不同之处在于它在该过程的第一阶段之后让右移量作为相位定位变量 i 值的函数。

[0206] 可以认识到当 i 的值是 2^k 的整数倍时, 等式 (9) 所示的滤波器抽头值将含有 K 个零值 LSB。于是, 如果滤波过程的第二阶段使用 2^k 整数倍的相位定位变量 i , 那么就可以将第二阶段的抽头值右移 K 位并将第一阶段的右移量减小 K 位。

[0207] 这在操作一般的重新采样因子时将变得相当难以跟踪。然而, 当执行 2 : 1 的简单重新采样因子或其他简单因子时, 很容易确认在滤波过程第二阶段中使用的所有相位都含有相同的乘数 2^k , 从而允许将该方法应用于这些特定的情况。

[0208] V. 位置计算的技术和工具

[0209] 描述了为空间 SVC 计算定位信息的技术和工具。

[0210] 一些技术和工具涉及如何聚焦于字长 B 并在该字长约束内优化计算精度。并非只是选择精度并要求一些必要的字长, 应用这一新方法会在真实实现中得到更高的精度并将拓宽该技术的有效应用范围, 因为它使用所有可用字长来最大化该约束内的准确性。

[0211] 一些技术和工具涉及 a) 偏移坐标系的原点以及 b) 使用无符号整数而非带符号整数以实现精度和字长 / 动态范围之间更好的折衷。需要增加少量的计算用以将原点偏移项加入每个计算的位置。

[0212] 一些技术和工具涉及将要生成样本串的各不同部分的计算打入处理的不同阶段, 其中坐标系的原点在每阶段开始时改变。同样地, 它用另一个少量计算要求的增加 (因为在每阶段开始时执行某些额外计算) 来提供精度和字长 / 动态范围之间更好的折衷。如果到达该技术的逻辑极值, 就能够排除对乘法操作的需要, 从而进一步改进精度和字长 / 动态范围之间的折衷。然而, 将会需要为每个样本执行某些额外的操作 (因为“每一阶段”所需的额外计算在每阶段仅含一个样本的情况下就变成成为每个样本所需)。

[0213] 作为一概括主题, 描述的设计用于该处理位置计算部分以实现计算结果的精度、处理元素的字长 / 动态范围、以及处理中涉及数学操作的次数和类型 (例如, 移位、加法和乘法操作) 之间理想的折衷。

[0214] 例如, 所描述的技术和工具允许使用 B 位 (例如, 32 位) 算术的灵活精度计算。这允许空间 SVC 编码器 / 解码器灵活适应不同的图像大小而无需为了计算转换成不同的算术 (例如, 16 位或 64 位算术)。使用灵活精度 B 位 (例如, 32 位) 算术, 编码器 / 解码器能够将灵活位数用于分数分量。这允许随着表示整数分量所需的位数降低 (例如, 对于较小的帧大小) 而增加计算精度。随着表示整数分量的所需位数的增加 (例如, 对于较大的帧大小), 编码器 / 解码器能够为整数分量使用更多位而为分数分量使用更少位, 从而降低精度但维持 B 位算术。由此就能大大简化不同精度和不同帧大小之间的变化。

[0215] 这一部分包括用于一示例性实现的具体细节。然而, 应该注意到本文描述的具体

细节可以依据在此描述的原则而在其他实现中有所变化。

[0216] A. 介绍和位置计算原则

[0217] 描述了用于计算位置和相位信息以实现低得多的计算要求而无需任何显著准确性损失的技术。例如,所描述的技术能够显著降低计算要求,例如通过动态降低额定动态范围要求(例如,降低几十位)。考虑可在基层和增强层中使用的各种可能的色度位置,期望找出一种提供重新采样色度样本相对于亮度样本的合适定位的解决方案。因此,所描述的技术允许用亮度和色度位置之间的不同关系来作出调整以计算视频格式的位置。

[0218] 为扩展空间可伸缩性设计的在前的上采样方法使用一种相当麻烦的方法来计算上采样低分辨率层时的位置和相位信息;它按比例缩放分母的上移近似逆,从而在分子增加时(即,在上采样过程从左向右或从上往下移时)导致了反演近似中舍入误差的放大通过比较,本文描述的技术具有卓越的准确性和简化的计算。更具体地,所述技术使得动态范围和位置计算中的右移量减小了几十位。

[0219] 例如,描述的一种技术用于计算位置信息来获取整数位置和相位定位变量 i , 其中 $i = 0..2^F-1$, 以便在 SVC 空间上采样中使用。

[0220] 所描述的技术将重新采样过程应用于空间可伸缩视频编码应用,而非前向参考图像重新采样。在这一空间可伸缩编码的应用中,可以应用某些简化。并非是普通的变形过程,只需要调整图像大小操作。这可以是针对每一维度分开的设计。

[0221] B. 位置计算设计

[0222] 考虑问题陈述,在每一维度(x或y)中,因为样本串的生成概念地处于新的(上采样的)阵列中实值从 L 至 $R > L$ 的范围。这一实值范围对应于参考低分辨率阵列中从 L' 至 $R' > L'$ 的范围。

[0223] 对于在新阵列中的位置 T 其中 $L \leq T \leq R$, 于是需要计算对应于该新阵列中位置的参考阵列中的位置。这会是位置 $T' = L' + (T-L)*(R' - L') \div (R-L)$ 。

[0224] 现在代替考虑调整从 L 至 R 的范围大小,定义整数 $M > 0$ 并且考虑通过相同的大小调整比率 $(R' - L') \div (R-L)$ 来调整从 L 至 $L+2^M$ 的范围大小。参考样本坐标中对应的范围则是从 L' 至 R'' , 其中 $R'' = L' + 2^M*(R' - L') \div (R-L)$ 。如果 M 足够大,即 $\text{if } M \geq \text{Ceil}(\text{Log}_2(R-L))$, 则 $R'' \geq R'$ 。(目前假设维持这一约束以解释以下概念,虽然这一约束对于等式的适当功能并不真是必需的。)

[0225] 现在能够为定位计算使用位置 L' 和 R'' 之间的线性内插。位置 L 被映射至位置 L' , 并且位置 $T \geq L$ 被映射至位置 $((2^M - (T-L))*L' + (T-L)*R'') \div 2^M$ 。这将操作的分母转换为 2 的幂,从而通过允许用二进制右移代替除法操作而降低了除法操作的计算复杂性。

[0226] 可以作出适当修改来整数化这一计算。将 L' 和 R'' 的值舍入为 $1 \div 2^G$ 的整数倍,其中 G 是整数,使得 L' 由 $k \div 2^G$ 近似,而 R'' 由 $r \div 2^G$ 近似,其中 k 和 r 是整数。使用这一调整,就能将位置 T 映射至位置 $((2^M - (T-L))*k + (T-L)*r) \div 2^{(M+G)}$ 。

[0227] 现在假设 T 和 L 的相关值是 $1 \div 2^J$ 的整数倍,其中 J 是整数,使得 $T-L = j \div 2^J$ 。使用这一调整,就能将位置 T 映射至位置 $((2^{(M+J)} - j)*k + j*r) \div 2^{(M+G+J)}$ 。

[0228] 回想以上部分 IV, 重新采样滤波器的分数相位可以是单位为 $1 \div 2^F$ 的整数。所以在这些单位中,算出的位置是 $\text{Round}(((2^{(M+J)} - j)*k + j*r) \div 2^{(M+G+J-F)})$, 或者

[0229] $t' = ((2^{(M+J)} - j)*k + j*r + 2^{(M+G+J-F-1)}) \gg (M+G+J-F)$ (16),

[0230] 或者,更简单地,

$$[0231] \quad t' = (j * C + D) >> S \quad (17),$$

[0232] 其中

$$[0233] \quad S = M + G + J - F \quad (18),$$

$$[0234] \quad C = r - k \quad (19),$$

$$[0235] \quad D = (k << (M + J)) + (1 << (S - 1)) \quad (20)。$$

[0236] 此处描述的方法在将算出的位置舍入至 $1 \div 2^F$ 最近倍数之前 (在两个设计内都存在的误差) 产生的仅有误差 (假设在 L 和 R 以及 L' 和 R' 的表示中没有误差) 是从位置 R'' 舍入到最近倍数 $1 \div 2^G$ 的舍入误差。该量在 $G+M$ 相对较大的情况下会很小。实际上,这一误差源被紧紧绑定在约 $(T-L) \div 2^{(G+M+1)}$ 的大小,结果计算的字长要求适中,并且模算术允许将结果的整数部分分出以使得字长最小化,或者允许用其他类似方法分解该计算。

[0237] F 例如可以是 4 或更大。(对于某些应用, $F = 3$ 或 $F = 2$ 就足够了。) J 值的示例包括用于亮度位置计算的 $J = 1$ 以及用于色度样本位置的 $J = 2$ 。如下可以找出这些 J 值示例的基本原理。

[0238] 1. 使用带符号的 B 位算术的第一示例性简化位置计算技术

[0239] 如果 $R' > 0$ 且 $L' > -R'$, 那么在要上采样图像中计算的所有位置 t' , 作为以 $1 \div 2^F$ 为单位的整数, 则位于 -2^Z 和 $2^Z - 1$ 之间, 其中 $Z = \text{Ceil}(\text{Log}_2(R')) + F$ 。如果 $(j * C + D)$ 计算的字长是 B 位, 并且假设使用带符号的 2 的补码算术, 则可以要求 $B - 1 \geq Z + S$ 。如果这一约束严格, 即如果 $B - 1 = Z + M + G + J - F$, 则实现高精度。

[0240] 对于相当小的图像大小 (例如, 在当前 H. 264/MPEG-4 AVC 标准中达 4.2 级), 可以使用 $B = 32$ 作为字长。也可以使用其他的 B 值。对于很大的图像, 可以使用更大的 B 。该计算也可以被轻易分解成更小的字长子计算以便在 16 位或其他处理器上使用。

[0241] 剩下的两个自由度是 M 和 G 。它们的关系是灵活的, 只要 G 大到足以避免在将 L' 表示为 $k \div 2^G$ 时舍入误差的任何需要。于是, 基于下一部分对 SVC 讨论的问题, 可以选择 $G = 2$, 得到:

$$[0242] \quad M = B + F - (G + J + Z + 1)$$

[0243] 即,

$$[0244] \quad M = 32 + 4 - (2 + 1 + Z + 1)$$

[0245] 即,

$$[0246] \quad M = 32 - Z。$$

[0247] 例如, 如果希望用 $B = 32$ 和 $L' = 0$ 上采样具有 1000 个亮度样本宽度的图像亮度阵列, 那么可以使用这第一示例性位置计算技术来使用 $F = 4$ 、 $G = 2$ 、 $J = 1$ 、 $M = 18$ 、 $S = 17$ 和 $Z = 14$ 。

[0248] 当 T 很接近 (或等于) R 且 R' 很接近 (或等于) 2 的整数幂, 特别是当 $(T-L) * (R' - L') \div 2^F$ 较大 (例如, 大于 $1/2$) 时, 那么可以在假设上存在由 1 违背上限的可能。在此不进一步考虑这些情况, 虽然处理这些情况的调整是直接的。

[0249] 2. 使用无符号的 B 位算术的第二示例性位置计算技术

[0250] 如果在低分辨率图像中计算的所有位置都大于或等于 0, 这种情况有时通过将合适的偏移量加入坐标系原点时存在, 那么使用无符号的整数算术而非带符号的 2 的补码算

术来计算 $t' = (j * C + D)$ 会是更好的选择。这允许多一位的动态范围而不会在计算中溢出（即，能够使用 B 位的动态范围大小而非 B-1 位），从而将 M（或 G）和 S 分别加 1 并进一步增加计算结果的精度。于是，在包括偏移量 E 以调整坐标系原点之后，计算的形式可以是 $t' = ((j * C + D') \gg S') + E$ 而非只是 $t' = (j * C + D) \gg S$ 。

[0251] 通过标识何时将不需要原点偏移量 E，提供有关该涉及无符号算术的更精确方法的进一步细节如下。

[0252] - 如上所述选择用于 B、F、G、J 和 Z 的值。

[0253] - 设置 $M = B + F - (G + J + Z)$ 。

[0254] - 分别计算在上式 (18)、(19) 和 (20) 指定的 S、C 和 D，其中 D 按带符号数计算。

[0255] - 如果 D 大于或等于 0，则不需要原点偏移量（即，不使用 E）并且可以使用无符号算术简单执行计算为 $t' = (j * C + D) \gg S$ ，而所得结果的精度会高于以上在部分 V. B. 1 中描述的第一示例性位置计算技术。

[0256] 除了通过启用使用无符号整数的计算来增加精度之外，偏移原点有时也可用于通过启用 Z 值的下降来提供改善的精度。没有原点偏移量，Z 是 R' 的函数。但用原点偏移量，就可以让 Z 是 $R' - L'$ 的函数，则在其得到较小 Z 值的情况下将使计算更为精确。

[0257] 通过示出一种对原点的偏移的方式，导出 D' 和 E，提供有关该涉及无符号算术的更精确方法的进一步细节如下。

[0258] - 如上所述选择用于 B、F、G 和 J 的值。

[0259] - 设置 $Z = \text{Ceil}(\text{Log}_2(R' - L')) + F$ 。

[0260] - 设置 $M = B + F - (G + J + Z)$ 。

[0261] - 分别计算在上式 (18)、(19) 和 (20) 指定的 S、C 和 D，其中 D 按带符号数计算。

[0262] - 设置 $E = D \gg S$ 。

[0263] - Set $D' = D - (E \ll S)$ 。

[0264] - 该位置计算于是可执行为 $t' = ((j * C + D') \gg S) + E$ 。

[0265] 如果以此方式计算 D' 和 E（以及 M、S 和 Z），则等式 $t' = ((j * C + D') \gg S) + E$ 的算术结果实际将总是与等式 $t' = (j * C + D) \gg S$ 的结果理论上相同，除了 $(j * C + D)$ 的值有时会落入从 0 到 $2^B - 1$ 的范围之外，而 $(j * C + D')$ 的值则不会。

[0266] 例如，如果希望用 $B = 32$ 和 $L' = 0$ 上采样具有 1000 个亮度样本宽度的图像亮度阵列，那么可以使用这第二示例性位置计算技术来使用 $F = 4$ 、 $G = 2$ 、 $J = 1$ 、 $M = 19$ 、 $S = 18$ 和 $Z = 14$ 。也等效地起作用，并非偏移原点以使得 $j * C + D'$ 所有值都非负，并由此允许使用 B 位计算使用无符号算术来使用范围从 0 到 $2^B - 1$ 的 B 位计算的另一种可能性是通过另一个 $2^{(B-1)}$ 以允许使用带符号算术来使用范围从 $-2^{(B-1)}$ 到 $2^{(B-1)} - 1$ 的 B 位计算来进一步将原点向右偏移。

[0267] 而在前一部分的第一示例性位置计算技术中，存在当 T 很接近（或等于）R 且 $R' - L'$ 很接近（或等于）2 的整数幂时所需的“角情况”调整。

[0268] 3. 用于位置计算的示例性多阶技术

[0269] 已经讨论了在其中使设计能够为覆盖要生成样本范围的所有 j 值（即，为在 L 和 R 之间的所有 T 值）而用相同的变量值 C、 D' 、S 和 E 用相同的等式（例如， $t' = ((j * C + D') \gg S) + E$ ）执行计算的方法。现在讨论如何放松这一假设，从而允许更高的精度和 / 或降

低的计算动态范围要求。

[0270] 一般而言,重新采样过程从左向右(或从上向下)进行以在相等间隔的位置上生成连续的样本串。在以上部分 V. B. 2 中描述的第二示例性位置技术,示出了改变使用偏移量参数 E 的原点如何能用于利用寄存器的 B 位动态范围来计算位置计算的 $(j * C + D')$ 部分。

[0271] 回想前一部分,D 中仅有的 S 个最低有效位在 D' 中保留,其他则移至 E 中。于是,计算 $(j * C + D')$ 的主要剩余的问题是 $j * C$ 的大小。

[0272] 回想 T 和 L 是 $1 \div 2^J$ 的整数倍。一般在较高分辨率图像中执行上采样过程以生成整数值增量处的样本串,例如相继生成样本之间的 2^J 间隔。于是期望为 p 和 N 的一些值计算对于 $i = 0$ 至 $N-1$ 对应于位置 $T_i = (p + i * 2^J) \div 2^J$ 的位置 t'_i 。

[0273] 这一过程可以如图 12 的伪码 1200 所示为 p 和 N 的一些值而用伪码总结。随着 i 向 N 递增,q 值增加,并且 q 的最大值应被维持在 B 位可用动态范围内。为 q 计算的最大值为 $(p + (N-1) * 2^J) * C + D'$ 。

[0274] 现在,代替以此方式在一个循环中生成所有样本,考虑将该过程分成多个阶段,例如两个阶段。例如,在两阶段过程中,第一阶段生成第一 $N_0 < N$ 个样本,而第二阶段生成剩下的 $N - N_0$ 个样本。同样地,因为 p 是关于循环的约束,所以可以在第一阶段前将其影响移入 D' 和 E。这导致了在图 13 的伪码 1300 中示出的两阶段过程。

[0275] 在伪码 1300 中每阶段的开始处,原点已被重设使得针对此阶段的 q 的第一值中除了 S 个最低有效位之外的所有位都已移入 E (即,针对第一阶段的 E_0 和针对第二阶段的 E_1)。于是,在两阶段中每阶段的操作期间,q 要求更小的动态范围。在以此方式将该过程分成各阶段之后,q 的最大值将会是 $N_0 * C' + D_0$ 或 $((N - N_0 - 1) * C' + D_1)$ 中的较大值。但是因为 D_0 和 D_1 各自具有不多于 S 位的无符号动态范围,所以这通常将会是小于前述单阶段设计的最大值。在此阶段中生成的样本数(即,第一阶段 N_0 个而第二阶段 $N - N_0$ 个)可以影响关联计算的动态范围。例如,在每一阶段使用较少样本数将导致相关计算的较小动态范围。

[0276] 每个阶段可以进一步分成更多的阶段,于是总共 N 个样本的生成可以被进一步分解成任何数量的这些更小的阶段。例如,该过程可以被分成大小相等的各阶段,从而在每一阶段生成例如 8 或 16 个相继样本的块。这一技术或者可用于降低计算 q 时所需的动态范围 B 的位数,或者可用于增加计算的精度(增加 S 和 G+M)同时保持动态范围相同,或者是这两大优点的混合。

[0277] 将位置计算过程分解成各阶段的这一技术也能够用于沿着极长的输入样本串(概念上,该串可以是无限长)执行连续的重新采样过程,诸如随着样本从音频信号的模数转换器的到来执行采样率转换。无疑的是,如果不把该过程分成有限大小的阶段并从每一阶段到下一阶段递增地重新设置原点,则在先前各部分中描述的技术将无法处理无限长度的样本串,因为这需要处理字长中无限的动态范围。然而,将各技术应用于有效无限串长的困难并非这些技术的实质限制,因为应用于有效无限长度仅在以用代表 $1 \div 2^G$ 倍数为整数单位的假设基准位置 L' 和 R'' 的表示没有引入舍入误差时有用。

[0278] 在其中能够应用多阶段位置计算技术的场景下,提供一种沿着无限长度样本串执行计算的方式,该方式不带有舍入误差的“漂移”累积,无论在整个速率转换过程的位置计算操作中发生了什么。

[0279] 4. 位置计算的示例性增量操作

[0280] 针对上述多阶段分解概念的有趣的指定情况是当每个阶段要生成的样本数已经降低至每阶段一个样本。图 14 中的伪码 1400 表示针对 $i = 0$ 至 $N-1$ 生成 N 个位置 t'_i 的过程。

[0281] 因为将该过程描述为上采样过程（虽然相同的原理也可应用于下采样过程），所以知道对于每次 i 的增量都在较高分辨率图像中存在有 1 的间隔，因此在较低分辨率图像中存在有小于或等于 1 的增量。在较低分辨率图像中空间位置内 1 的增量对应于 C' 的 $2^{(S+F)}$ 的值。同样知道 $D' < 2^S$ 。因此， $q = C' + D'$ 具有从 0 到小于 $2^{(S+F)} + 2^S$ 的范围，因此可以使用无符号整数算术用不多于 $B = S+F+1$ 位的动态范围要求来计算 q 。在一个实现中，这一动态范围要求对图像大小是不变的（即，不依赖于 R' 或 $R' - L'$ 的值）。

[0282] 对于可伸缩视频编码和许多其他这类应用，并不真的需要支持上采样比率非常接近 1。在这类应用中，可以假设 C' 实际上要求不多于 $S+F$ 位。

[0283] 例如，如果希望用 $B = 32$ 和 $L' = 0$ 上采样具有 1000 个亮度样本宽度的图像亮度阵列，那么可以使用这一方法来使用 $F = 4, G = 2, J = 1, M = 29, S = 28$ 和 $Z = 14$ 。结果将会是如此的格外精确使得较小的 B 值看上去是更为合理的选择。

[0284] 作为替换，如果希望用 $B = 16$ 和 $L' = 0$ 上采样具有 1000 个亮度样本宽度的图像亮度阵列，那么可以使用这一方法来使用 $F = 4, G = 2, J = 1, M = 13, S = 12$ 和 $Z = 14$ 。

[0285] 关于进一步了解执行上采样操作场景可提供进一步的优化机会。例如，如果上采样比率明显大于 2，则动态范围要求将再减小一位，并且对于大于 4、16 等的上采样比率而言继续减小。

[0286] 参考这一部分中的示例性增量位置计算技术描述的改变（相对于上述的示例性多阶位置计算机技术）都不会影响在给定 C, D 和 S 的值时所实际算出的位置 t'_i 的值。只有改变支持该计算所需的动态范围。

[0287] 用于这一分解形式的伪码 1400 中的内部循环不要求任何乘法操作。这一事实有利于在某些计算处理器上提供缩短的计算时间。

[0288] 5. 额外的注释

[0289] 对于诸如 $2 : 1, 3 : 2$ 之类的常见重新采样比率——在其中无需为了将位置 L' 和 R'' 近似为以为 $1 \div 2^G$ 单位的整数而作出舍入的任何情况——在使用这些方法时根本不存在任何舍入误差（除了当将最终结果舍入以 $1 \div 2^F$ 为单位的整数时引入的任何舍入误差，该误差是无论位置计算方法都将存在的误差）。

[0290] C. 亮度和色度位置和关系

[0291] 假设全新（上采样）图像和参考图像阵列相对于亮度采样网格索引坐标精确对准，则当前图像坐标内的位置 L 和 R 是 $L = -\frac{1}{2}$ 和 $R = -\frac{1}{2}$ ，其中 W 是取决于相关重新采样维度图像在垂直或水平方向上的样本数。等效地，可以设置图像空间坐标系的原点向网格索引 0 位置左移（或上移）半个样本并在从图像空间坐标转换成网格索引值时加 $1/2$ ，由此免于在空间坐标系中执行计算时处理负数的需要。

[0292] 参考（低分辨率）图像中的位置 L' 和 R' 以相同方式参考采样网格坐标，其中在此情况下 W 是参考图像而非新图像中的样本数。

[0293] 对于色度采样网格（无论是在新图像还是在参考图像中），情况多少没那么简单。为了构建色度样本相对于亮度的指定对准，考虑由色度样本表示的图像矩形与由亮度样本表示的矩形相同。这会产生如下情况：水平

[0294] 地，对于 4:2:0 色度采样类型 0、2 和 4（参见图 5D），当前图像坐标由 $L = -\frac{1}{4}$ 和 $R = W - \frac{1}{4}$ 定义。

[0295] - 水平地，对于 4:2:0 色度采样类型 3、1 和 5（参见图 5D），当前图像坐标由 $L = -\frac{1}{2}$ 和 $R = W - \frac{1}{2}$ 定义。

[0296] - 垂直地，对于 4:2:0 色度采样类型 2 和 3（参见图 5D），当前图像坐标由 $L = -\frac{1}{4}$ 和 $R = W - \frac{1}{4}$ 定义。

[0297] - 垂直地，对于 4:2:0 色度采样类型 0 和 1（参见图 5D），当前图像坐标由 $L = -\frac{1}{2}$ 和 $R = W - \frac{1}{2}$ 定义。

[0298] - 垂直地，对于 4:2:0 色度采样类型 4 和 5（参见图 5D），当前图像坐标由 $L = -\frac{3}{4}$ 和 $R = W - \frac{3}{4}$ 定义。

[0299] - 水平地，对于 4:2:2 色度采样，通常用于工业实践的 4:2:2 采样的当前图像坐标由 $L = -\frac{1}{4}$ 和 $R = W - \frac{1}{4}$ 定义。

[0300] - 垂直地，对于 4:2:2 色度采样，通常用于工业实践的 4:2:2 采样的当前图像坐标由 $L = -\frac{1}{2}$ 和 $R = W - \frac{1}{2}$ 定义。

[0301] - 对于水平和垂直两者而言，对于 4:4:4 色度采样，当前图像坐标由 $L = -\frac{1}{2}$ 和 $R = W - \frac{1}{2}$ 定义。

[0302] 同样地，使用一偏移量充分的将坐标系的原点移动到位置 L 的左侧以免于处理负数。

[0303] 整数坐标和分数相位偏移量余数通过调整要在上采样阵列中生成的各样本的整数坐标位置以补偿分数偏移量 L，并在随后应用部分 V. B 结尾处所示的转换来计算。概念地，将结果右移 F 位导致指向参考图像的整数系数指针，并且减去左移的整数坐标（移位 F 位）提供相位偏移量余数。

[0304] D. 上采样位置计算的额外精度

[0305] 这一部分描述如何将以上部分 V. C. 4 的位置计算方法映射至特定的上采样过程，诸如用于 H. 264 SVC 扩展的上采样过程。位置计算以一种非常灵活的方式应用以最大化各

种色度格式以及逐行和隔行扫描帧格式下亮度和色度通道两者的精度。在这一部分中描述的技术能够取决于实现和不同的上采样过程而变化。

[0306] 在上述位置计算中（在以上部分 V. A-C 中），比例改变参数（是变量 C，其后在以下标记为 ΔX （或者 ΔY ）由等于 2^J 的比例因子按比例增加（其中对于亮度 $J = 1$ ，而对于色度则为 2）来形成增量，加入该增量以从左向右或从上到下生成每个样本位置。选择缩放比例以使得按比例放大的增量将符合 16 位。

[0307] 1. 比例缩放位置计算的最大精度

[0308] 应用位置计算方法的一种直接方式是将比例改变参数按比例放大等于 2^J 的比例因子，其中对于亮度 $J = 1$ 而对于色度则为 2，从而形成增量，加入该增量以从左向右或从上到下生成每个样本位置。随后选择比例缩放参数以确保按比例放大的增量将符合指定的字长，诸如 16 位。在如下部分将描述更灵活的设计来最大化位置精度。

[0309] a. 亮度通道

[0310] “直接”亮度位置计算方法可以在 $F = 4$ 和 $S = 12$ 时由以下示例性等式（沿着水平方向）概括：

[0311] $\Delta X = \text{Floor}(((\text{BasePicWidth} < 15) + (\text{ScaledBaseWidth} > 1)) \div \text{ScaledBaseWidth})$

[0312] $xf = ((2 * (xP - \text{ScaledBaseLeftOffset}) + 1) * \Delta X - 30720) >> 12$

[0313] 在此， BasePicWidth 是基层或低分辨率图像的水平分辨率； ScaledBaseWidth 是高分辨率图像区域或窗口的水平分辨率； ΔX 是中间比例改变参数，在此情况下是 32768 乘以上采样比率的倒数的舍入近似； xP 表示在高分辨率图像中的样本位置； $\text{ScaledBaseLeftOffset}$ 表示高分辨率图像中图像窗口的相对位置，以及 $\text{Floor}()$ 指示小于或等于其自变量的最大整数。常数值 30720 是由在右移之前添加 2^{S-1} 作为舍入偏移量并针对亮度采样网格参考位置的半个样本偏移量来减去 $2^S * 2^F / 2$ 而产生，如在上述部分 V. C 开始处所讨论的。

[0314] 值得注意的是 xP 的每次增量都导致等式内部 $2 * \Delta X$ 的增量。同样地，量 $2 * \Delta X$ 的 LSB 总是为零，使得计算精度的一位实质上被浪费。通过如下改变这些等式，可以在复杂度没有任何显著增加的情况下近似地获取额外一位的精度：

[0315] $\Delta X = \text{Floor}(((\text{BasePicWidth} < 16) + (\text{ScaledBaseWidth} > 1)) \div \text{ScaledBaseWidth})$

[0316] $xf = ((xP - \text{ScaledBaseLeftOffset}) * \Delta X + (\Delta X >> 1) - 30720) >> 12$

[0317] 或者（略微）更加精确的形式如下：

[0318] $\Delta Xa = \text{Floor}(((\text{BasePicWidth} < 16) + (\text{ScaledBaseWidth} > 1)) \div \text{ScaledBaseWidth})$

[0319] $\Delta Xb = \text{Floor}(((\text{BasePicWidth} < 15) + (\text{ScaledBaseWidth} > 1)) \div \text{ScaledBaseWidth})$

[0320] $xf = ((xP - \text{ScaledBaseLeftOffset}) * \Delta Xa + \Delta Xb - 30720) >> 12$

[0321] 推荐后两种形式，因为其更高的准确性和可以忽略的复杂度影响（虽然精度的差异看上去也很小）。

[0322] 注意到对于处理其上难以执行除法计算的体系结构，具有这些等式之一的结果能

够简化其他的计算。deltaXa 的值将总是在 $2 * \text{deltaXa}$ 加 1 或减 1 的范围内。因此可以导出如下的简化规则以避免对 deltaXa 的计算执行除法操作的需要：

[0323] $\text{deltaXa} = (\text{deltaXb} \ll 1)$

[0324] $\text{remainderDiff} = (\text{BasePicWidth} \ll 16) + (\text{ScaledBaseWidth} \gg 1) - \text{deltaXa}$

[0325] $\text{if}(\text{remainderDiff} < 0)$

[0326] $\text{deltaXa}--$

[0327] $\text{elseif}(\text{remainderDiff} \geq \text{ScaledBaseWidth})$

[0328] $\text{deltaXa}++$

[0329] b. 色度通道

[0330] 在这一部分的设计中可以为色度通道使用四因子乘法器来代替二因子乘法器，从而能够表示 4:2:0 采样的色度位置（为色度使用 $J = 2$ 而非如所述地为亮度使用的 $J = 1$ ）。因此，“直接”等式为：

[0331] $\text{deltaXC} = \text{Floor}(((\text{BasePicWidthC} \ll 14) + (\text{ScaledBaseWidthC} \gg 1)) \div$

[0332] $\text{ScaledBaseWidthC})$

[0333] $\text{xfC} = (((4 * (\text{xC} - \text{ScaledBaseLeftOffsetC}) +$

[0334] $(2 + \text{scaledBaseChromaPhaseX})) * \text{deltaXC}$

[0335] $+ 2048) \gg 12) - 4 * (2 + \text{baseChromaPhaseX})$

[0336] 在此，baseGhromaPhaseX 和 scaledBaseChromaPhaseX 分别表示用于低分辨率和高分辨率的色度采样网格位置偏移量。这些参数的值可以随着信息从编码器发送至解码器而被显式传达，或者可以具有由应用确定的特定值。所有其他的变量与为亮度通道定义的那些相似，而其带有的“C”后缀表示对色度通道的应用。

[0337] xC 的每次递增都导致等式内部 $4 * \text{deltaXC}$ 的增量。因此，通过如下改变这些等式，可以在复杂度没有任何实质增加的情况下近似地获取额外两位的精度：

[0338] $\text{deltaXC} = \text{Floor}(((\text{BasePicWidthC} \ll 16) + (\text{ScaledBaseWidthC} \gg 1)) \div$

[0339] $\text{ScaledBaseWidthC})$

[0340] $\text{xfC} = (((\text{xC} - \text{ScaledBaseLeftOffsetC}) * \text{deltaXC}$

[0341] $+ (2 + \text{scaledBaseChromaPhaseX}) * ((\text{deltaXC} + K) \gg 2)$

[0342] $+ 2048) \gg 12) - 4 * (2 + \text{baseChromaPhaseX})$

[0343] 其中 $K = 0, 1$ 或 2 。使用 $K = 0$ 可以避免一次额外操作。使用 $K = 1$ 或 $K = 2$ 可以具有稍高的准确性。

[0344] 相对地稍微更精确的形式可以如下：

[0345] $\text{deltaXCa} = \text{Floor}(((\text{BasePicWidthC} \ll 16) + (\text{ScaledBaseWidthC} \gg 1)) \div$

[0346] $\text{ScaledBaseWidthC})$

[0347] $\text{deltaXCb} = \text{Floor}(((\text{BasePicWidthC} \ll 14) + (\text{ScaledBaseWidthC} \gg 1)) \div$

[0348] $\text{ScaledBaseWidthC})$

[0349] $\text{xfC} = (((\text{xC} - \text{ScaledBaseLeftOffsetC}) * \text{deltaXCa} +$

[0350] $(2 + \text{scaledBaseChromaPhaseX}) * \text{deltaXCb}$

[0351] $+ 2048) \gg 12) - 4 * (2 + \text{baseChromaPhaseX})$

[0352] 如同亮度情况一样，推荐后一种变体，因为复杂度的差异看上去可以忽略不计

(虽然精度差异看上去也很小)。

[0353] c. 隔行扫描场坐标

[0354] 对图像坐标系的参考通常基于亮度帧坐标中半个样本位置,于是导致如上所述针对亮度坐标参考位置的为 2 的比例因子。亮度帧坐标中半个样本的位移对应于 4:2:0 色度帧坐标中四分之一一个样本的位移,这也是当前在用于上述色度坐标的比例缩放中使用为 4 的因子而非为 2 的因子的原因。

[0355] 水平地,对表示帧的编码图像和表示隔行扫描视频单个半帧的那些图像进行的操作不存在本质上的差别。然而,当编码的图像表示单个半帧时,在亮度帧垂直坐标中半个样本的位移对应于在亮度场垂直坐标中四分之一一个样本的位移。于是,应该在垂直亮度坐标位置的计算中应用为 4 而非为 2 的比例因子。

[0356] 类似地,当编码的图像表示单个半帧时,在亮度帧垂直坐标中半个样本的位移对应于在色度场垂直坐标中八分之一一个样本的位移。于是,应该在垂直色度坐标位置的计算中应用为 8 而非为 4 的比例因子。

[0357] 用于在编码场图像中计算垂直坐标位置的这些比例缩放因子可以与以上关于编码帧图像中增量计算所述相同的方式并入 ΔY 垂直增量计算。在此情况下,由于应用了增加的比例缩放因子,使得精度的改善对亮度位置近似为 2 位的精度增加而对色度(垂直地)近似为 3 位的精度增加。

[0358] 2. 4:2:2 和 4:4:4 色度限制和细化

[0359] 部分 V. D. 1. b 的位置计算方法要求为色度使用亮度不同的乘法因子。这对 4:2:0 视频有意义并且对 4:2:2 视频在水平上也合理,但是它对 4:2:2 视频在垂直上或者对 4:4:4 视频在水平或垂直上都不必要,因为在这些情况下亮度和色度分辨率相同并且亮度和色度样本因此推测起来在同一地点。

[0360] 结果,部分 V. D. 1. b 的方法可能要求分开计算来确定亮度和色度位置,即使在某些维度中亮度和色度分辨率相同且无意相移的情况下也是如此,这只是因为在这两种情况下将要执行的舍入略有不同。这是不期望的,所以在这一部分中建议为 4:2:2 和 4:4:4 采样结构使用不同的色度处理。

[0361] a. 4:2:2 垂直的和 4:4:4 水平的和垂直的位置

[0362] 对于 4:2:2 视频的垂直维度以及 4:4:4 视频的垂直和水平维度两者,不存在对色度相位的自定义控制的明显需要。因此,只要在某一维度中色度分辨率与亮度分辨率相同,就应该对用于计算色度位置的等式进行修改以便无论何时色度采样格式在某一具体维度上对亮度和色度具有相同的分辨率,为亮度和色度样本两者计算准确相同的位置。一个选项是仅设置色度位置变量等于亮度位置变量,另一选项是设置色度位置等式以使它们具有相同的结果。

[0363] b. 4:2:2 水平位置

[0364] 虽然水平地允许为 4:2:2 视频的色度相位调整不存在功能问题,但如果 4:2:2 只使用了一种类型的水平子采样结构,诸如对应于部分 V. D. 1. b 各等式中用于 $\text{scaledBaseChromaPhaseX}$ 或 BaseChromaPhaseX 的值 -1 的那个,那么可能期望考虑只要颜色采样格式是 4:2:2 就强迫使用这些值。

[0365] VI. 扩展和变化

[0366] 本文中的技术和工具也可应用于使用参考图像重新采样的多分辨率视频编码, 诸如在 ITU-T 国际标准推荐 H. 263 的附件 P 中找出的情况。

[0367] 本文中的技术和工具不仅还可应用于图像样本阵列的上采样, 还可用于残差数据信号或其他信号的上采样。例如, 本文中的技术和工具也可应用于降低分辨率更新编码的残差数据信号上采样, 例如可在 ITU-T 国际标准推荐 H. 263 的附件 Q 中找出。作为另一个示例, 本文中描述的技术和工具也可应用于在空间可伸缩视频编码设计中从较低分辨率残差信号中预测高分辨率残差信号的残差数据信号的上采样。作为又一个示例, 本文中描述的技术和工具也可应用于在空间可伸缩视频编码设计中运动矢量场的上采样作为再一个示例, 本文中描述的技术和工具也可应用于图形图像、静止照片图像、音频样本信号等的上采样。

[0368] 参考各所述实施例描述和示出了本发明的原理之后, 可以认识到, 可以在排列和细节上修改所述各实施例, 而不脱离这些原理。应当理解, 此处所描述的程序、过程或方法不相关于或不限于任何特定类型的计算环境, 除非另外指明。可依照此处所描述的教导来使用各种类型的通用或专用计算环境或执行操作。描述的实施例中以软件示出的元素可以用硬件来实现, 反之亦然。

[0369] 鉴于可应用本发明的原理的许多可能的实施例, 要求保护落入所附权利要求书及其等效技术方案的范围和精神之内的所有这样的实施例作为本发明。

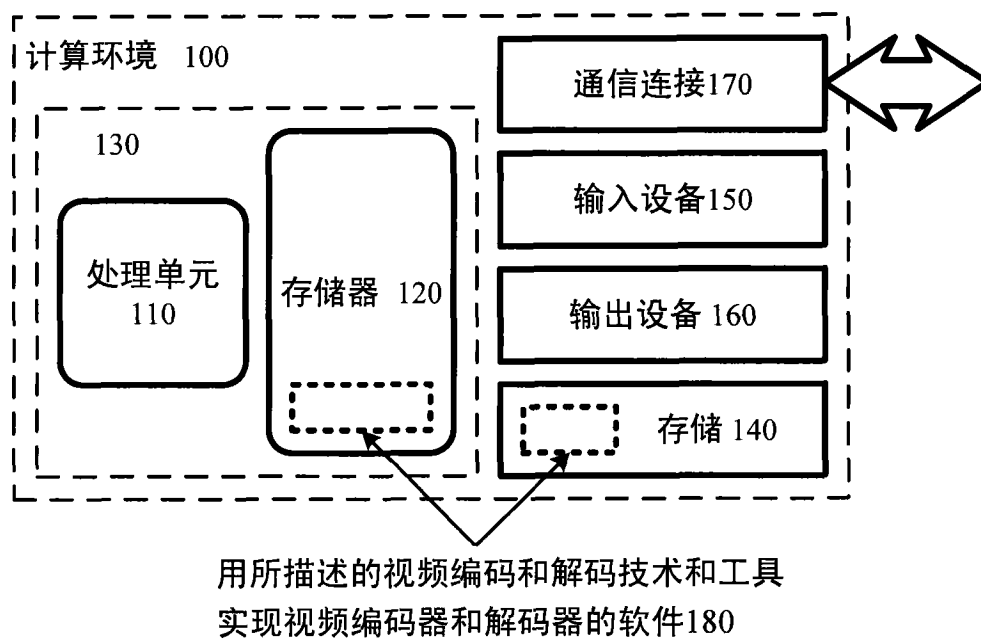


图 1

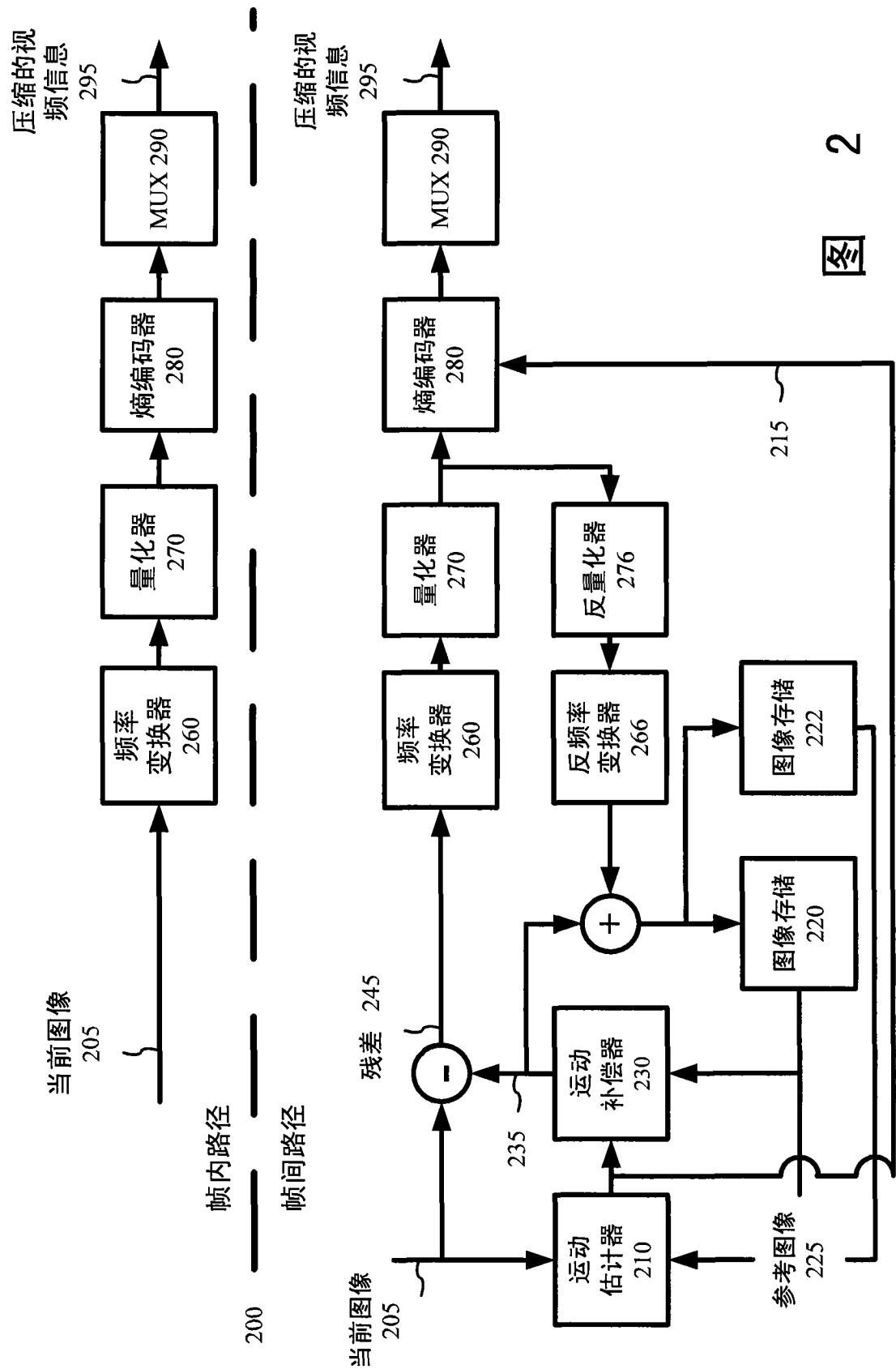


图 2

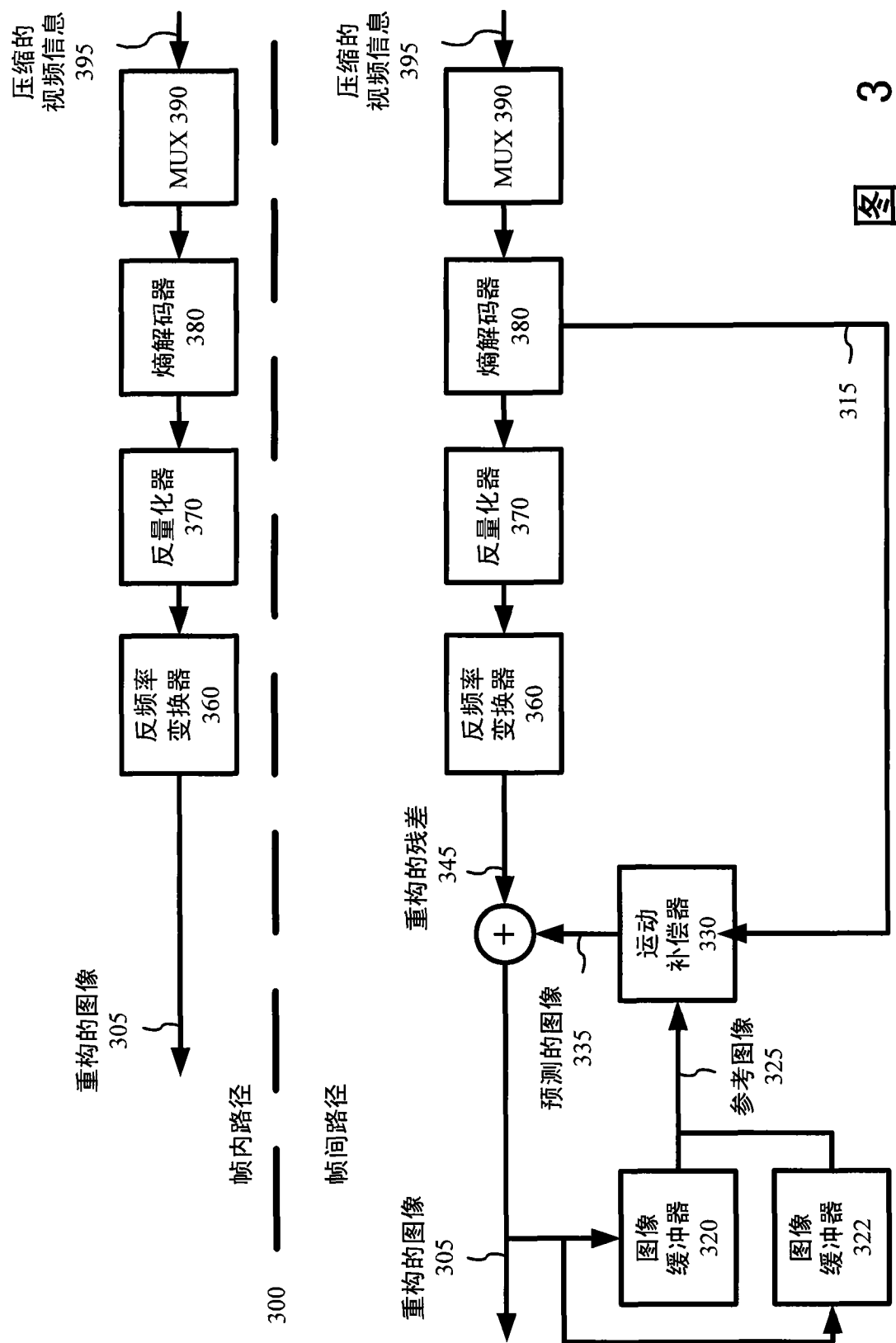
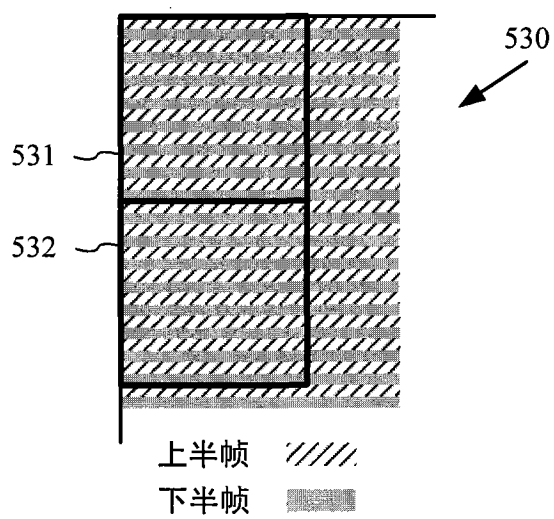
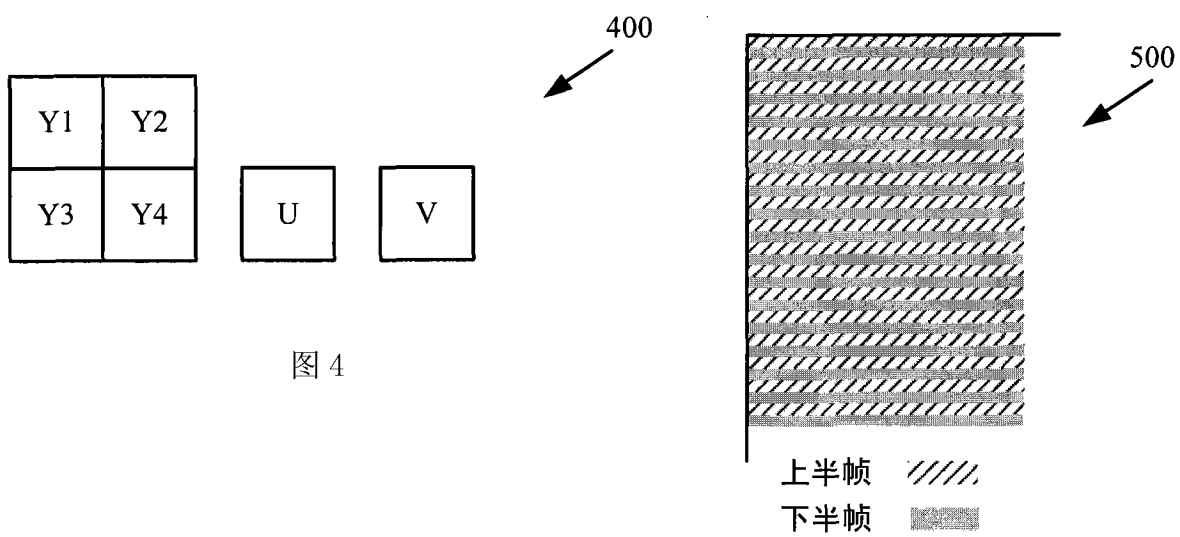


图 3



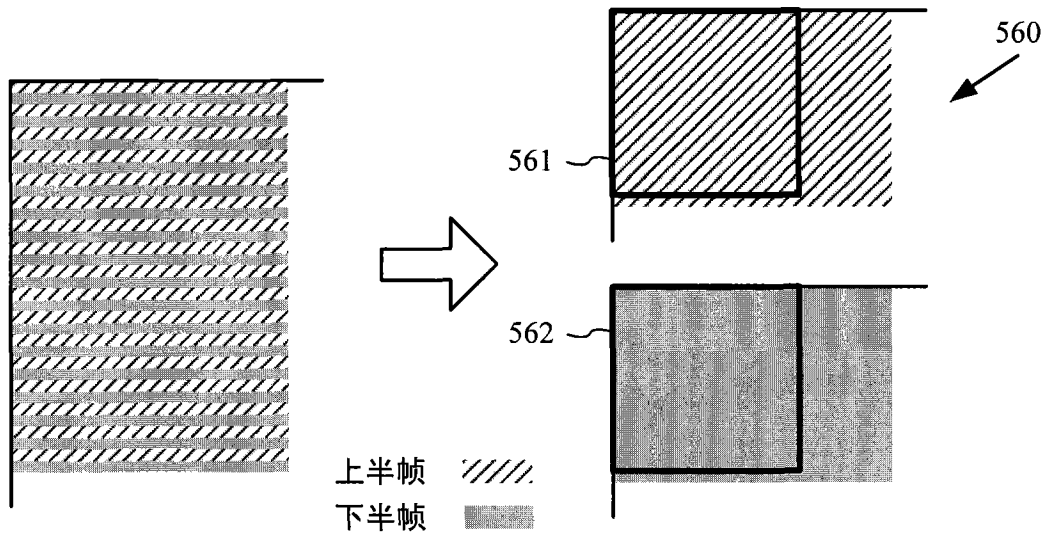


图 5C

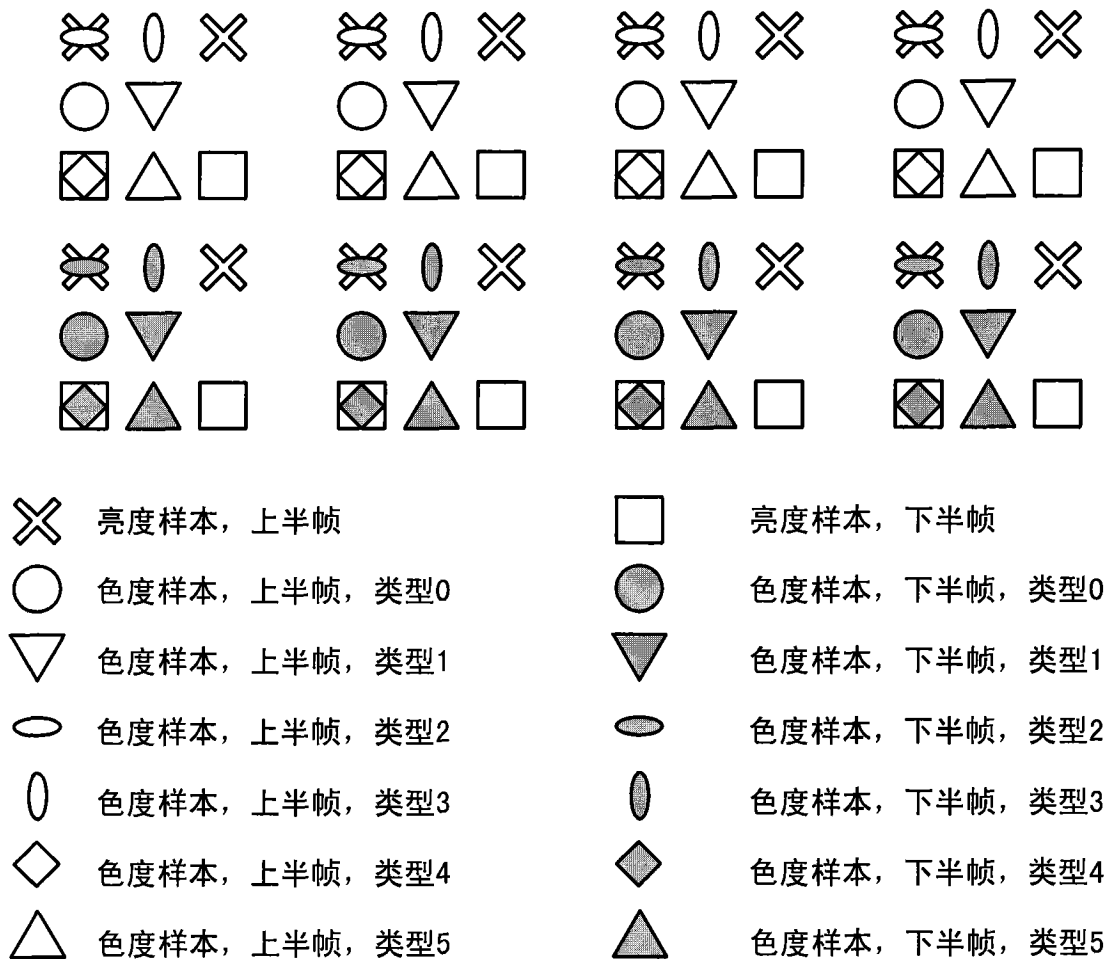
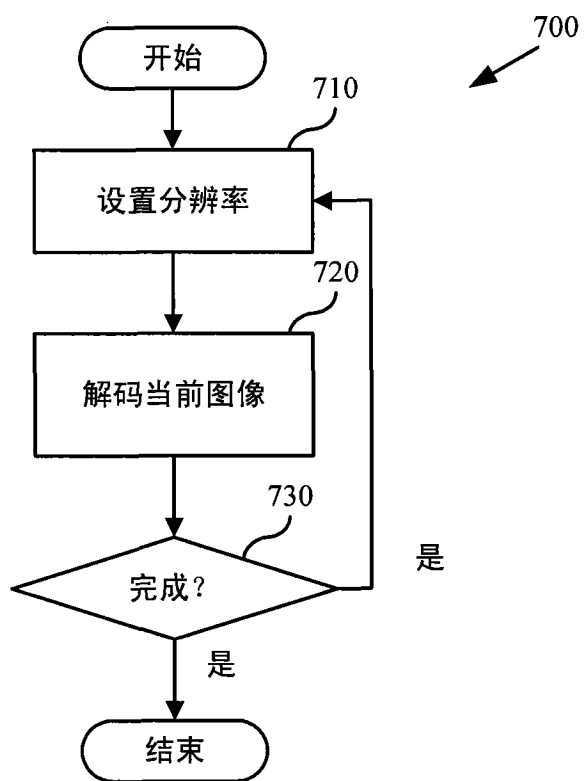
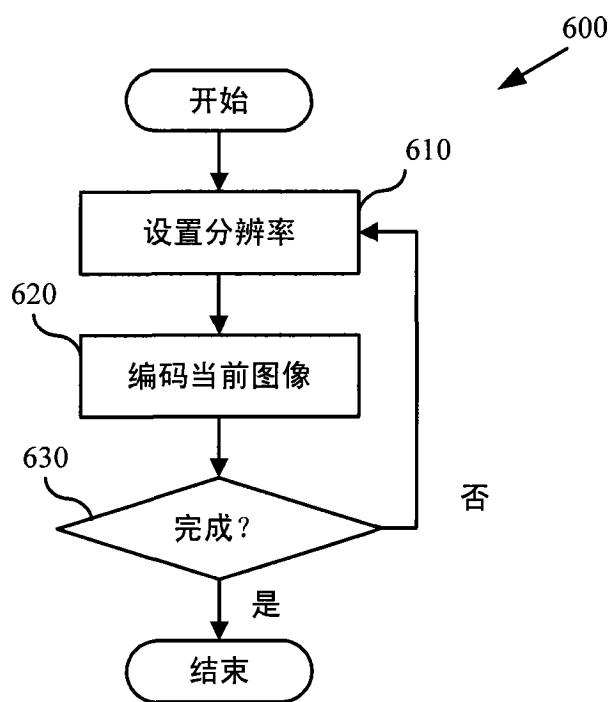


图 5D



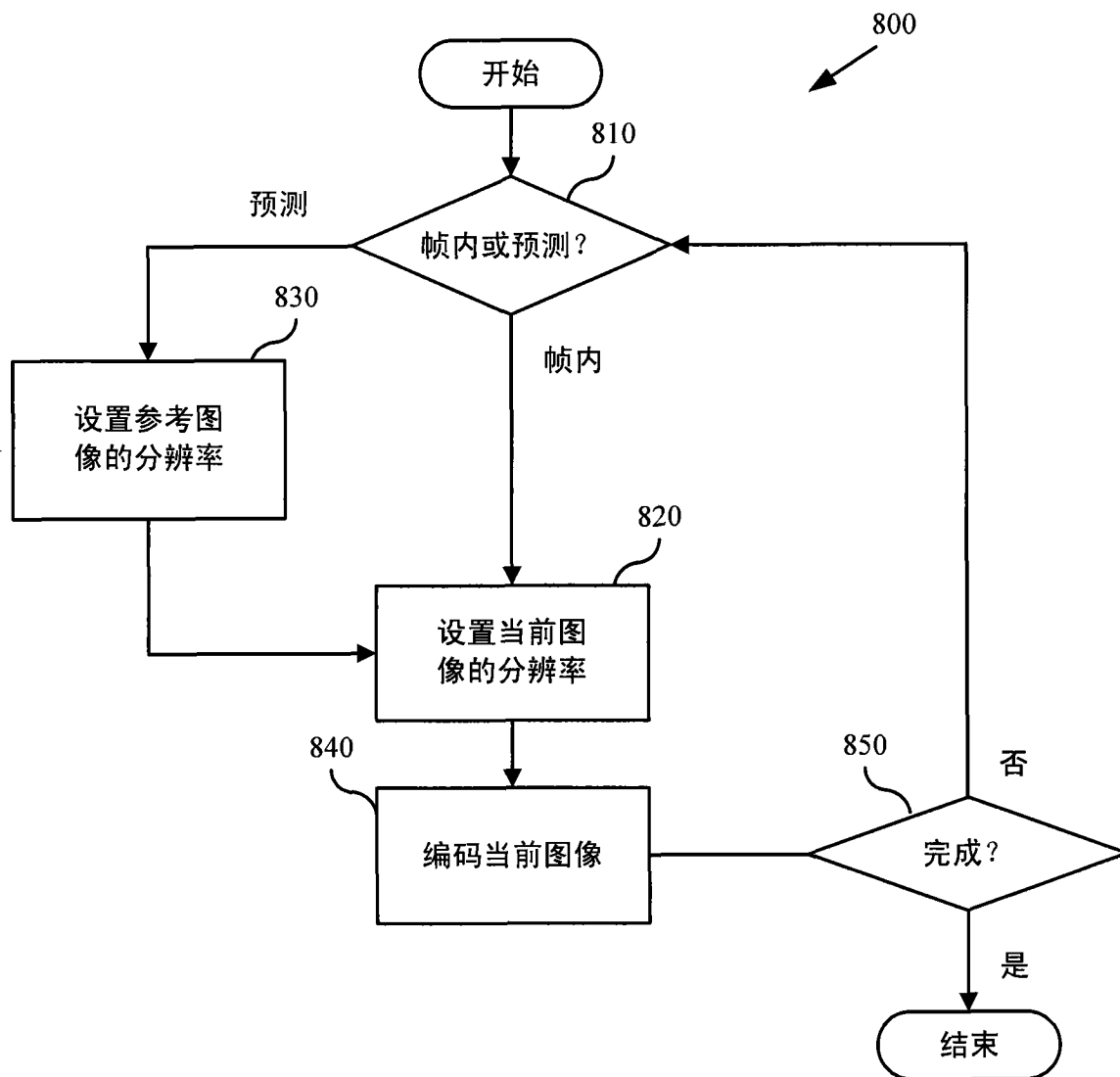


图 8

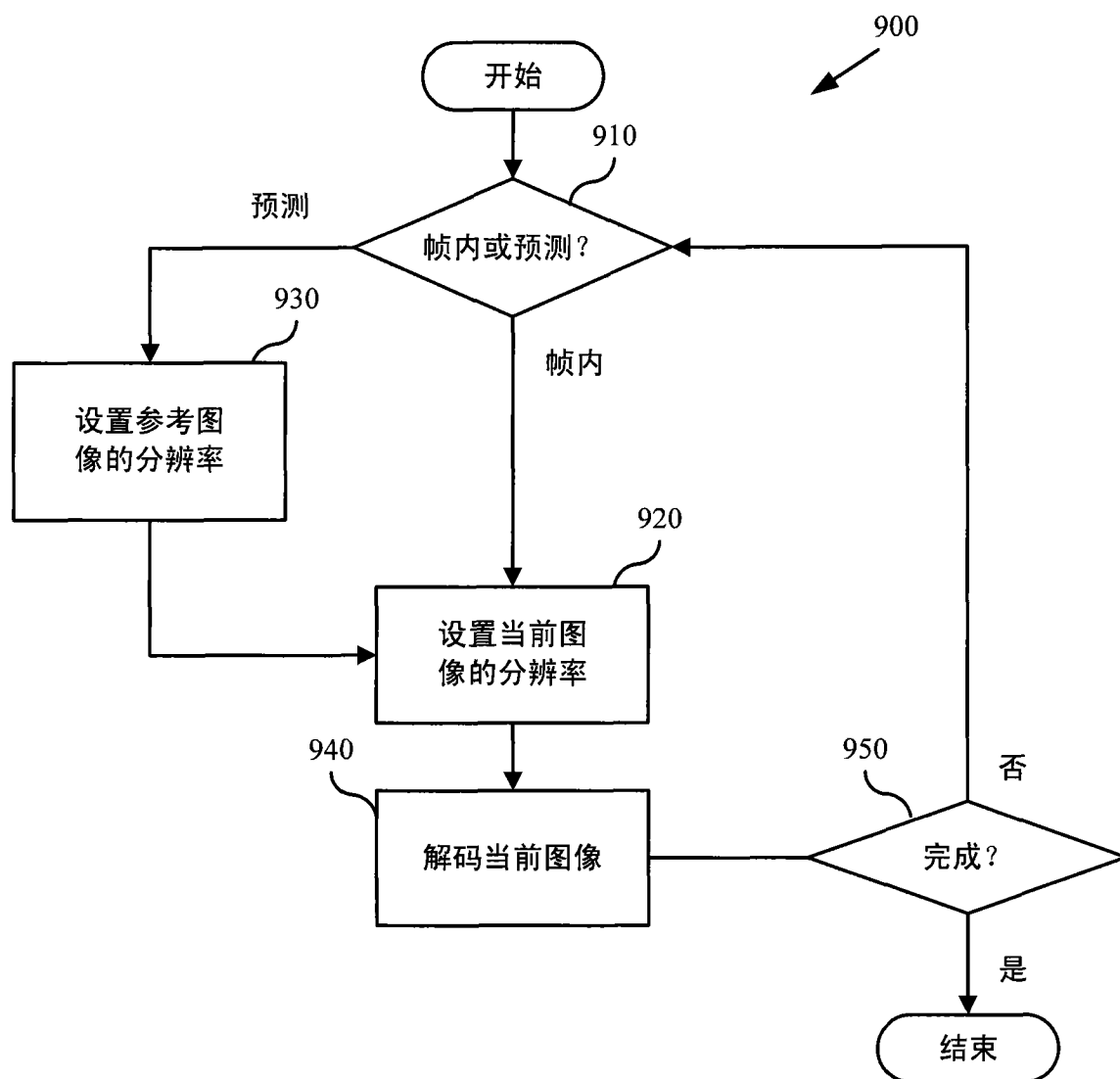


图 9

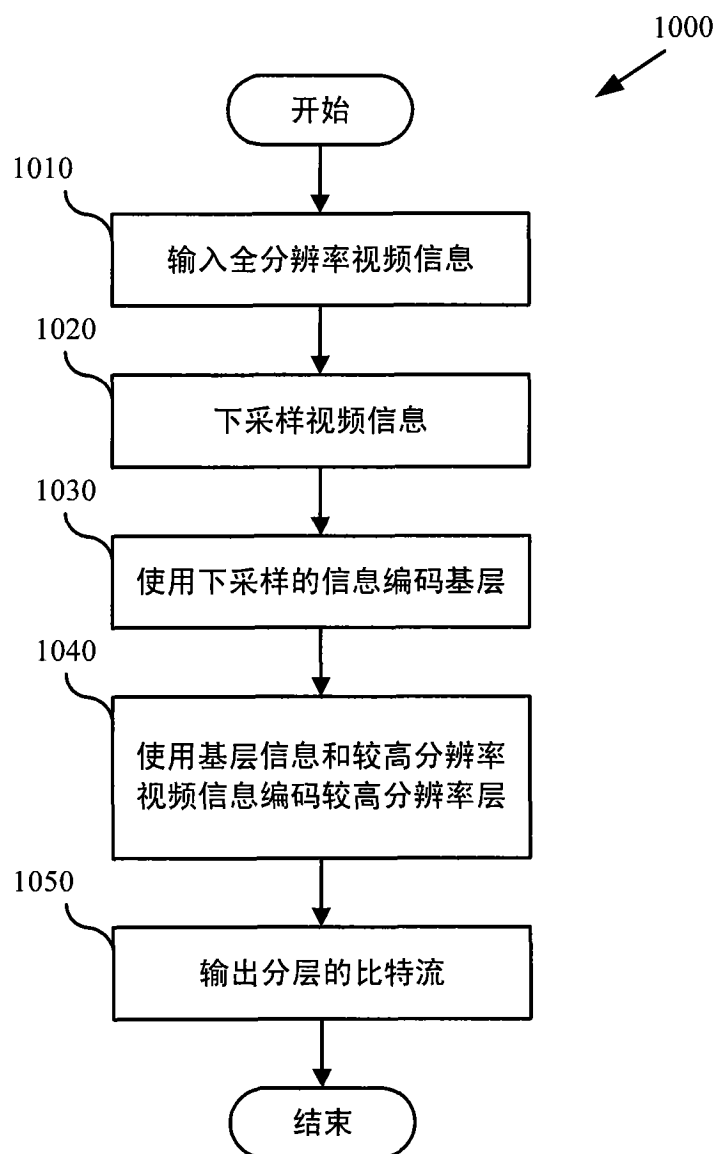


图 10

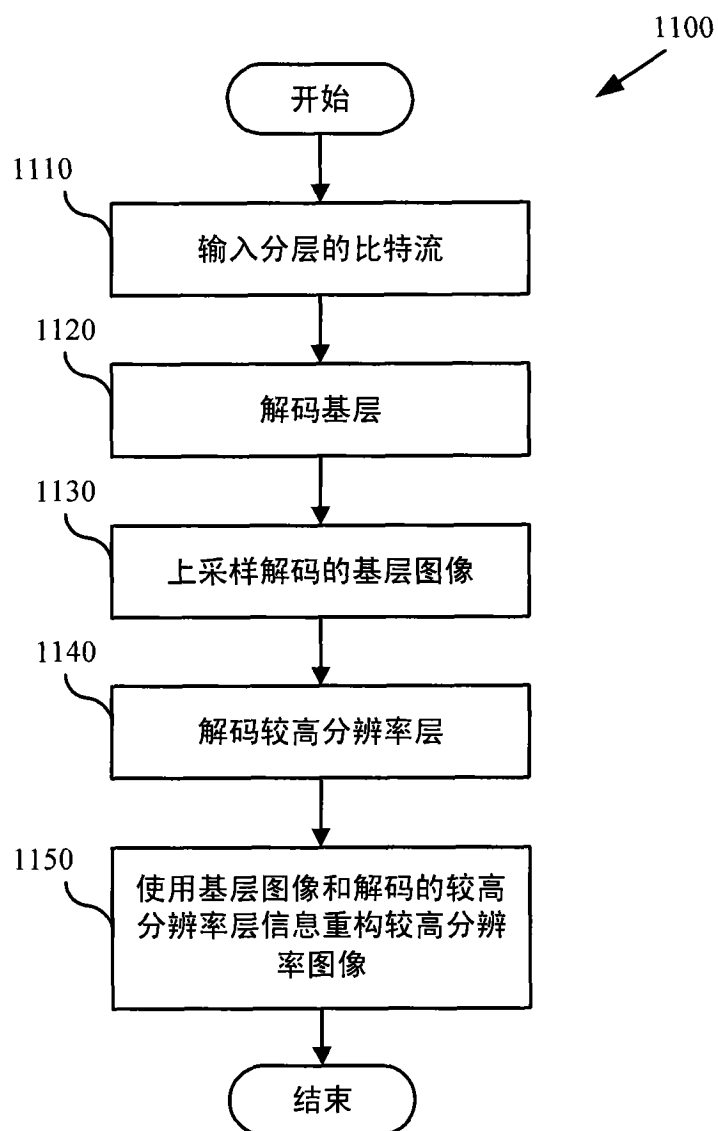


图 11

1200
↙

```

E = D >> S;
D' = D - ( E << S );
for( i = 0; i < N; i++ ) {
    q = ( p + i * 2J ) * C + D';
    t'i = ( q >> S ) + E;
}

```

图 12

1300
↙

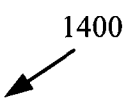
```

C' = C << J;
q = p * C + D;
/*第一阶段*/
w = q >> S;
E0 = w;
D0 = q - ( w << S );
t0 = E0
for( i = 1; i < N0; i++ ) {
    q = i * C' + D0;
    t'i = ( q >> S ) + E0;
}
/*第二阶段*/
q = N0 * C' + D0;
w = q >> S;
E1 = E0 + w;
D1 = q - ( w << S );
tN0 = E1;
for( i = 1; i < N - N0; i++ ) {
    q = i * C' + D1;
    tN0+i = ( q >> S ) + E1;
}

```

图 13

1400



```
C' = C << J;  
E = 0;  
  
q = p * C + D;  
w = q >> S;  
D' = q - ( w << S );  
E += w;  
t0 = E;  
  
for( i = 1; i < N; i++ ) {  
    q = C' + D'  
    w = q >> S;  
    D' = q - ( w << S );  
    E += w;  
    t'i = E;  
}
```

图 14