



(12) 发明专利

(10) 授权公告号 CN 101147399 B

(45) 授权公告日 2011. 11. 30

(21) 申请号 200680008133. 7

(22) 申请日 2006. 03. 20

(30) 优先权数据

05300253. 1 2005. 04. 06 EP

(85) PCT申请进入国家阶段日

2007. 09. 13

(86) PCT申请的申请数据

PCT/EP2006/060873 2006. 03. 20

(87) PCT申请的公布数据

W02006/106039 EN 2006. 10. 12

(73) 专利权人 汤姆森许可贸易公司

地址 法国布洛涅-比郎库尔

(72) 发明人 陈颖 翟杰夫 尹鹏

普尔温·比贝哈斯·潘迪特

(74) 专利代理机构 中科专利商标代理有限责任

公司 11021

代理人 戎志敏

(51) Int. Cl.

H04N 7/26 (2006. 01)

H04N 7/32 (2006. 01)

H04N 7/46 (2006. 01)

(56) 对比文件

US 6510177 B1, 2003. 01. 21, 全文.

CN 1568009 A, 2005. 01. 19, 全文.

US 6332002 B1, 2001. 12. 18, 全文.

REICHEL J ET AL. Joint Scalable Video Model JSVM 0. JOINT VIDEO TEAM(JVT) OF ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 AND ITU-T SG16 06). 2005, 1-73.

审查员 潘云

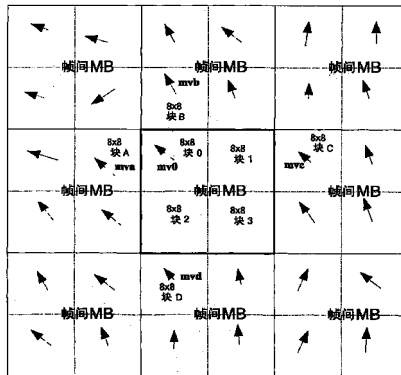
权利要求书 2 页 说明书 5 页 附图 2 页

(54) 发明名称

编码增强层视频数据的方法和设备

(57) 摘要

一种用于在 SVC 中提高 BLSkip 模式性能的方法, 包括以下步骤: 对基本层的运动场进行上采样, 内插帧内编码 MB 的运动矢量, 将 8x8 块运动场内插至 4x4 块运动场, 以及使用相邻备选项, 在 BLSkip 模式中生成用于 4x4 块的 MV 预测因子。



1. 一种编码空间增强层视频数据而预测运动矢量的方法,包括以下步骤:

对基本层的运动场进行上采样,

将上采样后的运动场的运动矢量分配给增强层的 16×16 宏块;

根据 16×16 宏块的运动矢量来为帧间编码的 8×8 块和帧内编码的 8×8 块生成运动矢量,其中,通过对四个最近的帧间编码的 8×8 块(块 A、...、块 D)的运动矢量执行内插来获得用于四个帧内编码的 8×8 块(块 0、...、块 3)的运动矢量;

通过针对四个帧内编码的 8×8 块(块 0、...、块 3)的运动矢量执行内插来获得增强层的 4×4 块的运动矢量;以及

执行编码模式判决,其中,对于基本层跳过模式,使用所获得的增强层的所述 4×4 块的运动矢量作为预测因子。

2. 如权利要求 1 所述的方法,其中,为了生成帧内编码的 8×8 块的运动矢量,使用双线性内插。

3. 如权利要求 2 所述的方法,其中,根据 $mv_0 = ((D_c \cdot mv_a + D_a \cdot mv_c) / (D_a + D_c) + (D_d \cdot mv_b + D_b \cdot mv_d) / (D_b + D_d)) / 2$,双线性内插考虑了从当前帧内编码的 8×8 块至四个最近帧间编码的 8×8 块(块 A、...、块 D)的距离,其中, mv_0 是用于当前帧内编码的 8×8 块内插的运动矢量, D_a 、 D_b 、 D_c 、 D_d 是当前帧内编码的 8×8 块至四个最近帧间编码的 8×8 块的距离, mv_a 、 mv_b 、 mv_c 、 mv_d 是四个最近帧间编码的 8×8 块的运动矢量。

4. 如权利要求 1 所述的方法,其中,为了通过内插获得增强层的每个 4×4 块($b14 \times 4A$ 、...、 $b14 \times 4D$)的运动矢量,仅使用当前 16×16 宏块的 8×8 块的运动矢量,其中, 8×8 块是帧内编码的或帧间编码的。

5. 如权利要求 4 所述的方法,其中,通过内插获得所述每个 4×4 块($b14 \times 4A$ 、...、 $b14 \times 4D$)的运动矢量包括:将权重分配给当前 16×16 宏块的 8×8 块的运动矢量,其中,分配给 8×8 块的运动矢量的权重取决于与该 4×4 块直接相邻或包括在该 4×4 块中的 2×2 子块的个数。

6. 一种编码空间增强层视频数据而预测运动矢量的设备,包括:

- 上采样装置,用于对基本层的运动场进行上采样;

- 装置,用于将上采样后的运动场的运动矢量分配给增强层的 16×16 宏块;

- 装置,用于根据 16×16 宏块的运动矢量来为帧内编码的 8×8 块和帧间编码的 8×8 块生成运动矢量,其中,通过对四个最近的帧间编码的 8×8 块(块 A、...、块 D)的运动矢量执行内插来获得四个帧内编码的 8×8 块(块 0、...、块 3)的运动矢量;以及

- 内插装置,用于通过针对四个帧内编码的 8×8 块(块 0、...、块 3)的运动矢量执行内插来获得增强层的 4×4 块的运动矢量;以及

- 装置,用于执行编码模式判决,其中,对于基本层跳过模式,使用所获得的增强层的所述 4×4 块的运动矢量作为预测因子。

7. 如权利要求 6 所述的设备,其中,为了所述生成帧内编码的 8×8 块的运动矢量,使用双线性内插。

8. 如权利要求 7 所述的设备,其中,根据 $mv_0 = ((D_c \cdot mv_a + D_a \cdot mv_c) / (D_a + D_c) + (D_d \cdot mv_b + D_b \cdot mv_d) / (D_b + D_d)) / 2$,双线性内插考虑了从当前帧内编码的 8×8 块至四个最近帧间编码的 8×8 块(块 A、...、块 D)的距离,其中, mv_0 是用于当前帧内编码的 8×8

块内插的运动矢量, D_a 、 D_b 、 D_c 、 D_d 是当前帧内编码 8×8 块至四个最近帧间编码的 8×8 块的距离, mv_a 、 mv_b 、 mv_c 、 mv_d 是四个最近帧间编码的 8×8 块的运动矢量。

9. 如权利要求 6 所述的设备, 其中, 为了通过内插获得增强层的每个 4×4 块 ($b14 \times 4A$ 、...、 $b14 \times 4D$) 的运动矢量, 仅使用当前 16×16 宏块的 8×8 块的运动矢量, 其中, 8×8 块是帧内编码的或帧间编码的。

10. 如权利要求 9 所述的设备, 其中, 通过内插获得所述每个 4×4 块 ($b14 \times 4A$ 、...、 $b14 \times 4D$) 的运动矢量包括: 将权重分配给当前 16×16 宏块的 8×8 块的运动矢量, 其中, 分配给 8×8 块的运动矢量的权重取决于与该 4×4 块直接相邻或包括在该 4×4 块中的 2×2 子块的个数。

编码增强层视频数据的方法和设备

技术领域

[0001] 本发明涉及一种用于对视频数据进行编码的方法。

背景技术

[0002] H. 264/AVC 标准提供了极好的编码效率,但是并没有考虑可缩放的视频编码(SVC)。SVC 可以提供不同的层,通常是基本层(BL)和增强层(EL)。为了给出视频编解码器的更多功能,运动图像专家组(MPEG)考虑为 SVC 提供标准。建议了各种技术,最终联合视频工作组(JVT)启动了称为 JSVC 的标准,相应的参考软件描述称为 JVSM。SVC 为应用程序提供了时间、SNR 和空间的可缩放性。JVSM 的基本层与 H. 264 兼容,如所规定的,H. 264 的大部分组件用于 JSVC,从而仅有个别组件需要根据子带结构进行调整。在所有的可缩放性中,由于很难使用两个空间可缩放层之间的冗余,所以空间可缩放性是最具挑战性和感兴趣的主题。

[0003] SVC 为空间可缩放性提供了多种技术,如 IntraBL 模式、残余预测或 BLSkip(基本层跳过)模式。可以在宏块(MB)级上选择这些模式。

[0004] IntraBL 模式使用上采样的重构 BL 画面来预测 EL 中的 MB,并仅对残余进行编码。残余预测尝试通过减去 BL 的上采样 MC 残余来减小 EL 的运动补偿(MC)残余的能量。

[0005] BLSkip 模式利用了 EL 中 MB 的上采样 MV,并且如果 MB 选择了该模式,则仅需要将残余写入比特流。因此,在空间可缩放情况下,BLSkip 模式利用了 BL 和 EL 的 MV 之间的冗余。

[0006] 在 SVC 的 JSVC 中,BLSkip 模式用于帧间编码预测(P)帧和帧间编码双预测(B)帧的 MB。将通常为每个 4×4 块存储的 BL MV 通过与 2 相乘来进行上采样。然后,上采样后的 MV 将与更高分辨率 EL 的 8×8 块相对应。即,如果 QCIF(176×144)BL 帧具有 (11×9) 个 MB,以及每个 MB 具有十六个 4×4 块,则 BL 中存在 $11 \times 9 \times 16$ 个 MV(如果没有帧内 MB)。当所选 MV 是 (h, v) ,以及相应的 4×4 块具有起始坐标 (x, y) ,则上采样后的 MV 是 $(h*2, v*2)$,以及在高分辨率帧(CIF: 352×288)中相应的 8×8 块是 $(2*x, 2*y)$ 。因此,将向具有起始坐标 $(2x, 2y)$ 、 $(2x+4, 2y)$ 、 $(2x, 2y+4)$ 和 $(2x+4, 2y+4)$ 的四个 4×4 块分配相同的 MV $(2h, 2v)$ 。

[0007] 然后,在模式判决过程中,当 BLSkip 模式是当前的备选项时,MB(如果它具有起始坐标 $(2x, 2y)$)将通过使用实际四个 MV 来设置 MV。对于四个 8×8 子块,为当前的 MB 设置四个 MV,它们具有相应的起始坐标 $(2x, 2y)$ 、 $(2x+8, 2y)$ 、 $(2x, 2y+8)$ 和 $(2x+8, 2y+8)$ 。

发明内容

[0008] 用于所描述的 MB 的 MV 结构模式实际上是 MODE_8x8,这表示将在分辨率深度 8×8 而非 4×4 中规定 MV 的差别。

[0009] 上述结构表示可以通过使用四个 MV 来处理每个 8×8 块的 MC 过程。然而,这 4 个 MV 实际上是相同的。使用这种模式,由于在解码过程中可以按照相同的方式来预测这些 MV,所以 MV 并不用于被编码。所以,如果预测因子与真实的运动稍有类似,则节省了许多比

特。SVC 所使用的上述解决方案是简单的,并具有低复杂度。然而,上采样后的 MV 由于两个原因而并不非常精确。一个原因在于, MV 仅具有 1/2 的像素精度,而主要为 MC 所使用的 MV 具有 1/4 的像素精度;另一原因在于,对于每个 8×8 块,它的四个 4×4 子块使用相同的 MV,这从分辨率的角度来看也可能并不足够准确。

[0010] 因此,由本发明所解决的问题是提高预测 MV 的精度,尤其对于 BLSkip 模式。

[0011] 本发明给出了对于空间可缩放性的更好的运动场上采样和内插方法,从而提高了 SVC 帧内画面的编码效率。

[0012] 本发明首先对由 8×8 MV 所构造的运动场进行平滑并内插到由 4×4 MV 所构造的更好粒度运动场,然后在预测过程期间,使用标量中值、矢量中值和平均方法来为 BLSkip 模式选择更好的预测因子。结果,所公开的 MV 预测技术实现了改进的编码效率,尤其对于 SVC JSVM 1.0 中的 BLSkip 模式。

[0013] 根据本发明,用于对更高层视频数据进行编码的方法包括以下步骤:

[0014] 对基本层的运动场进行上采样,

[0015] 内插帧内编码宏块的运动矢量,

[0016] 将 8×8 块运动场内插至 4×4 块运动场,以及

[0017] 在基本层跳过模式中生成用于 4×4 块的运动矢量预测因子,其中,使用来自相邻块的运动矢量。

[0018] 此外,根据本发明,用于对更高层视频数据进行编码的设备包括:

[0019] 装置,用于对基本层的运动场进行上采样,

[0020] 装置,用于内插帧内编码宏块的运动矢量,

[0021] 装置,用于将 8×8 块运动场内插至 4×4 块运动场,以及

[0022] 装置,用于在基本层跳过模式中生成用于 4×4 块的运动矢量预测因子,其中,使用来自相邻块的运动矢量。

[0023] 在从属权利要求、以下描述和附图中公开了本发明的有利实施例。

附图说明

[0024] 参照附图描述了本发明的示例性实施例,其中:

[0025] 图 1 是用于帧内 MB 的 8×8 MV 内插;

[0026] 图 2 是用于 4×4 块的 MV 的双线性内插;以及

[0027] 图 3 是用于 4×4 块的双线性内插的相关块。

具体实施方式

[0028] 目前,仅将根据基本层(低分辨率层)的协同定位的 MV 上采样的 MV 用作增强层(高分辨率层)的预测因子,上采样后的 MV 的精度很低(通常,如果在运动估计(ME)过程中使用 1/4 像素精度,则是 1/2 像素精度)。以及 MV 分割仅用于 8×8 块。即, 8×8 块中的每 4 个 4×4 块将共享相同的 MV。本发明关注于:尤其对于 BLSkip 模式,通过首先使用运动场平滑技术,然后给出更多的 MV 作为当前 MB 的预测因子,以确定是否将要选择 BLSkip 模式,从而提高 MV 的精度。给出 BLSkip 模式的更好的(多个)预测因子(它对于画面中的 MB 具有高百分比),可以提高编码效率。

[0029] 为了实现更高的编码效率,本发明通过对于选择 BLSkip 模式的 MB 给出更好的 MV 预测因子,改进了 BLSkip 模式。为了获得将更加接近于真实运动的精确的 MV,首先利用多种方法来平滑 BL 的运动场,然后利用策略从运动场以及来自当前层的相邻 MV 中选择用于 EL 的 MB 的 MV。

[0030] 在以下的 2D 区域 P 中定义运动场: $MV(P) \subset N^2$, 其中, D 是点阵像素集,以及 N 是整数字段。

[0031] 由于默认的 MV 精度是四分之一 (1/4) 像素,所以每个 MV $mv(p)$ 实际上是 4 和 ME 的替换的乘积。

[0032] 对于来自 BL 的上采样字段,可以将它看作在由 8×8 点阵像素组成的区域 P 中所限定的函数。所以 P 是 $\{(P_x, P_y) | P_x = 8x, P_y = 8y, \text{对于任何 } 0 \leq x < 22, \text{及 } 0 \leq y < 18\}$ 。这里, x 和 y 的上界实际上是由上采样之后的帧的分辨率的宽度和高度所确定的:宽度 /8 和高度 /8。

[0033] 实际上,如果 MB 是帧内,则不具有 MV,所以不存在对于 P 中相应的四个值的 MV(P) 的限定。所以通过内插首先给出对于 P 中那些值的定义,然后使用相邻的 MV。

[0034] 当具有对于在 P 的 8×8 点阵定义区域中任何值的限定时,将构造另一运动场: $MV(P_4) \subset N^2$, 其中,在 4×4 点阵像素的区域内定义该函数。

[0035] 1. 在帧内区域中定义运动场

[0036] 在 H.264/AVC 中,甚至帧间画面 (P 或 B) 也具有帧内 MB。在 SVC 中,保持相同的特征。所以在第一步骤中,给出了对于与帧内 MB 相对应的点阵的限定。

[0037] 可能的解决方案是使用双线性内插以获得用于帧内 MB 的 8×8 MV。第一步骤是对于来自每个方向的帧内 MB 中的 8×8 块,发现不属于帧内 MB 的最近的 8×8 块。

[0038] 如果 16×16 MB 的起始坐标是 (x, y), 则如果附近 (相邻) 的 MB 是帧间 MB, 则所谓最近的 MV 具有以下坐标:

[0039] 块 0 : $A(x-8, y), B(x, y-8), C(x+16, y), D(x, y+16)$

[0040] 块 1 : $A(x-8, y), B(x+8, y-8), C(x+16, y), D(x+8, y+16)$

[0041] 块 2 : $A(x-8, y+8), B(x, y-8), C(x+16, y+8), D(x, y+16)$

[0042] 块 3 : $A(x-8, y+8), B(x, y-8), C(x+16, y+8), D(x+8, y+16)$

[0043] 在图 1 中,不同方向的四个最近的 MV (mv_a, mv_b, mv_c, mv_d) 用于内插块 0 的 MV (mv_0)。

[0044] 实际上,在上采样后的运动场中,由于 16×16 帧内 MB (无论它是 INTRA_16x16 还是 INTRA_4x4) 将变成可分为四个 16×16 帧内 MB 的 32×32 块,所以四个帧内 MB 将同时出现。如果附近的 MB 仍然是帧内编码的,则本发明的方法继续沿相同方向进行搜索,直至遇到帧间 MB,然后使用该帧间 MB。在得到四个 8×8 块之后,可以执行 MV 内插过程。

[0045] 定义了两个块的距离,要由 $Dist(P_A - P_B) = |x_a - x_b| + |y_a - y_b|$ 计算获得,其中 P_A 和 P_B 是 8×8 块的起始坐标。

[0046] 例如,用于图 1 中块 0 的最近 8×8 块是块 A、B、C、D。以及 0 与 A、B、C、D 的距离分别是 $D_a = 8, D_b = 8, D_c = 16, D_d = 16$ 。块 0 中的内插 MV 由以下计算获得:

[0047] $mv_0 = ((D_c \cdot mv_a + D_a \cdot mv_c) / (D_a + D_c) + (D_d \cdot mv_b + D_b \cdot mv_d) / (D_b + D_d)) / 2$

[0048] 2. 对 8×8 运动场进行平滑并内插至 4×4 运动场

[0049] 现在,在 8×8 点阵的任何像素中定义运动场。下一步骤是创建在 4×4 点阵中定义的运动场。一个可能的解决方案在图 2 中示出。例如,用于 4×4 块 A ($b14 \times 4A$) 的权重分别是关于 8×8 块 0、1、2、3 ($b18 \times 80 \dots 3$) 的 $9/16$ 、 $3/16$ 、 $3/16$ 、 $1/16$ 。

[0050] 换言之, 8×8 块在 4×4 块上的权重或影响取决于它的 2×2 子块的个数,该子块包括在 4×4 块内或直接与 4×4 块相邻。如在与图 2 相对应的图 3 中所示,这些是 $b18 \times 8_0$ 中的 9、 $b18 \times 8_1$ 和 $b18 \times 8_2$ 的每个中的 3、以及 $b18 \times 8_3$ 中的 1。

[0051] 对于 4×4 块的每个中心点,存在 8×8 块的四个最近中心点,它们包含 MV (mv_a, mv_b, mv_c, mv_d)。 4×4 块的 MV 通过以下公式进行内插:

$$[0052] \quad mv = \sum_{i=0}^3 mv_i \cdot weight_v \cdot weight_h$$

[0053] 其中,下标 x 是图 2 中所示的 A、B、C、D。对于公式中的每个预测因子 MV,存在两个权重。

$$[0054] \quad Weight_v = (x_p - x_0) / (x_1 - x_0)$$

$$[0055] \quad Weight_h = (y_p - y_0) / (y_4 - y_0)$$

[0056] 其他解决方案也可用于运动场的内插和平滑。

[0057] 3. BLSkip 模式的运动矢量预测

[0058] 当在模式判决过程期间,当前模式备选项是 BLSkip 时,由于 BL 运动场实际上是基于 4×4 点阵的,所以将把运动补偿模式设为 MODE_4x4。

[0059] 简单的解决方案是将上采样运动场的 4×4 块的协同定位 MV 用作 MB 的每个 4×4 块的 BLSkip 模式。

[0060] 另一解决方案是使用相邻的 MV 来预测用于当前 4×4 块的 MV。备选项可以是四个块的 MV (左 4×4 块、上 4×4 块、下 4×4 块和右 4×4 块) 或者八个块的 MV (再加上左上 4×4 块、左下 4×4 块、右上 4×4 块和右下 4×4 块)。

[0061] 对于特定备选 MV,可以使用多种方法来获得预测因子:对备选项进行平均、取备选项的缩放中值、或者取备选项的矢量中值。

[0062] 平均意味着通过对备选项的各个水平或垂直值进行平均来生成 MV 的水平值和垂直值。

[0063] 缩放中值也通过分别获得水平值和垂直值的中值来生成。为了得到多个值的中值,简单的解决方案是首先对值进行分类,然后使用中值索引获得所分类的列表中的值。

[0064] 矢量中值方法得出了与其他具有最小距离之和的 MV。所以以下等式的解决方案是矢量中值 MV。

$$[0065] \quad mv_{VM} = \arg \min_{mv_i} (\sum_{i \neq j} Dist(mv_i, mv_j))$$

[0066] 为了低通画面,或者如果参考帧远离当前帧,则平均方法将得到更好的结果。如果两个帧的距离相靠近,则中值或矢量中值方法将会更好。一种自适应 MV 预测策略是对于低通帧利用平均方法,以及对于高通帧利用中值或矢量中值方法。

[0067] 原理上,用于提高 SVC 中 BLSkip 模式的性能的本发明的方法包括以下步骤:

[0068] - 对基本层的运动场进行上采样;

[0069] - 内插帧内 MB 的运动矢量;

- [0070] - 将 8×8 块运动场内插至 4×4 块运动场；
- [0071] - 使用相邻备选项, 在 BLSkip 模式中生成 4×4 块的 MV 预测因子。
- [0072] 用于提高 SVC 中 BLSkip 模式的性能的本发明的设备具有：
- [0073] - 装置, 用于对基本层的运动场进行上采样；
- [0074] - 装置, 用于内插帧内 MB 的运动矢量；
- [0075] - 装置, 用于将 8×8 块运动场内插至 4×4 块运动场；
- [0076] - 装置, 用于使用相邻备选项, 在 BLSkip 模式中生成 4×4 块的 MV 预测因子。
- [0077] 在本发明的一个实施例中, 来自不同方向的四个最近的 8×8 块的 MV 用于预测帧内 MB 中的 8×8 块。
- [0078] 在本发明的一个实施例中, 双线性内插用于将 8×8 运动场内插到 4×4 运动场。
- [0079] 在一个实施例中, 选择策略 (例如, 平均、中值和矢量中值) 用于从相邻 MV 备选项中得到 MV。
- [0080] 在一个实施例中, 使用自适应 MV 选择策略, 它考虑了当前帧与参考帧之间的距离: 如果距离短, 则优选中值或矢量中值, 否则优选平均。
- [0081] 简而言之, LSVM 将上采样后的 BLMV 用作用于 BLSkip 模式的更高分辨率层的 MV。然而, 上采样后的 MV 并不足够好, 还可以进行改进。所有上采样后的 MV 实际上是一半像素精度, 以及在 8×8 块中的每个四个 4×4 块共享相同的 MV。在本发明中, 提供了给出用于 BLSkip 模式的最佳运动矢量的方法和设备。
- [0082] 本发明的核心在于以下: 为了实现更高的编码效率, 本发明通过为选择 BLSkip 模式的 MB 给出更好的 MV 预测因子来改进 BLSkip 模式。为了得到将更加接近于真实运动的更加准确的 MV, 首先利用多种方法来平滑 BL 的运动场, 然后利用策略, 根据运动场以及根据来自当前层的相邻 MV, 为 EL 的 MB 选择 MV。
- [0083] 本发明的主要优点在于, 改进了 SVC 的帧间编码画面的编码效率, 这通过更好的运动场上采样和自适应内插方法来实现。

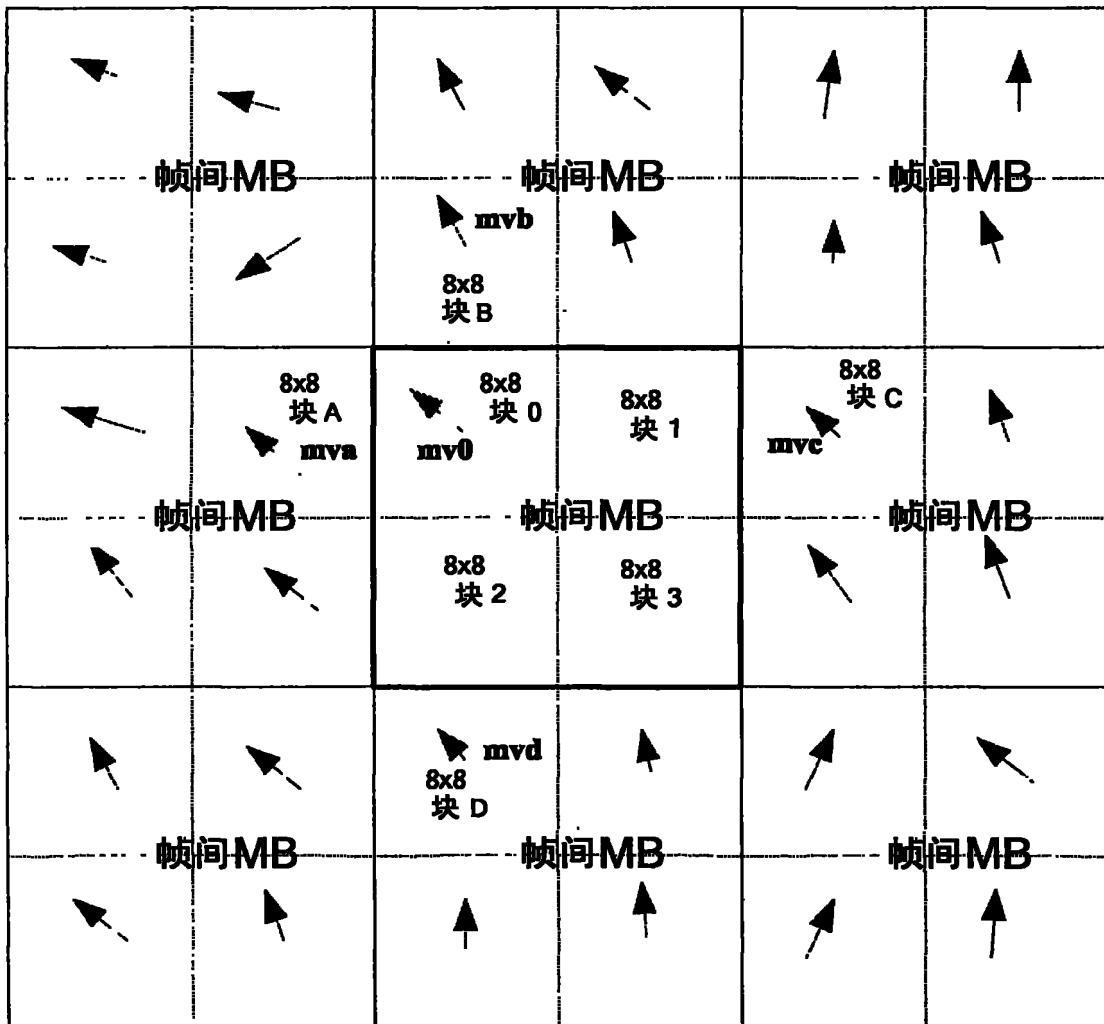


图 1

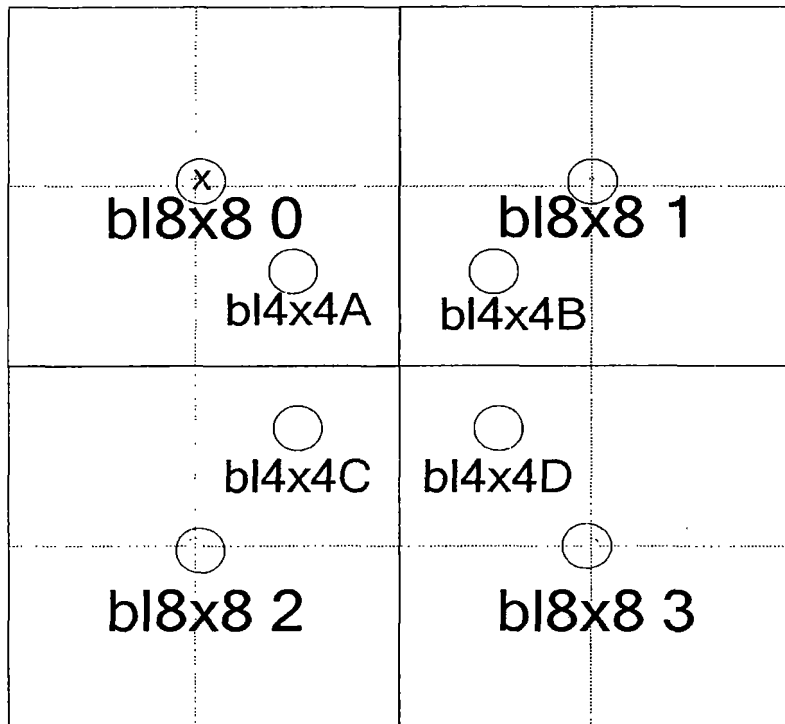


图 2

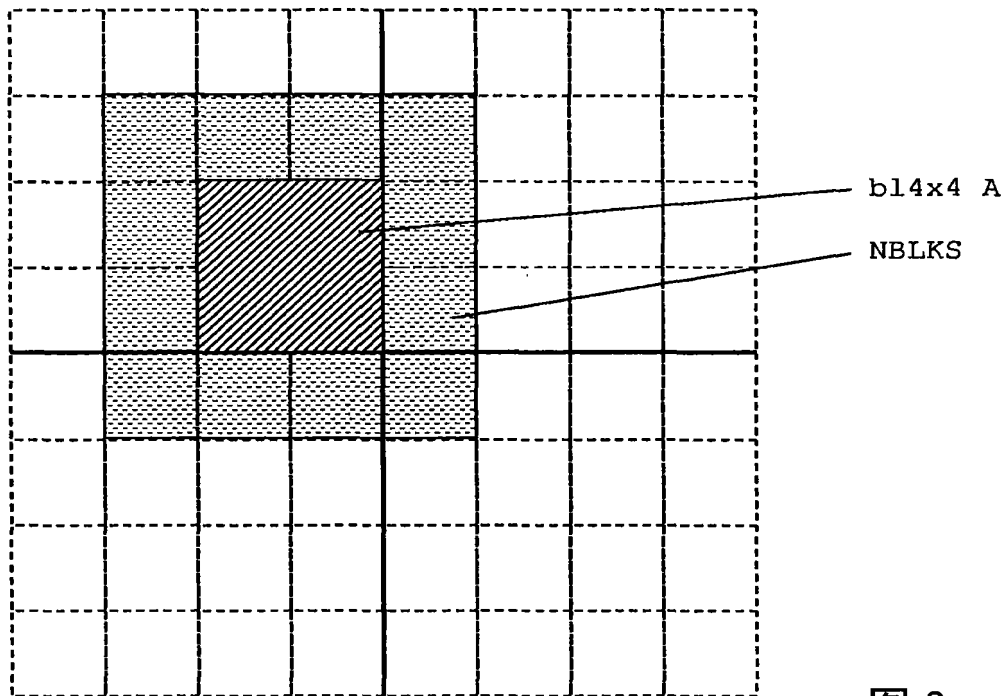


图 3