

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.  
H04N 7/26 (2006.01)



# [12] 发明专利说明书

专利号 ZL 200580002762.4

[45] 授权公告日 2010年1月20日

[11] 授权公告号 CN 100584013C

[22] 申请日 2005.12.28

[21] 申请号 200580002762.4

[30] 优先权

[32] 2005.1.7 [33] JP [31] 002423/2005

[86] 国际申请 PCT/JP2005/024093 2005.12.28

[87] 国际公布 WO2006/073116 日 2006.7.13

[85] 进入国家阶段日期 2006.7.19

[73] 专利权人 日本电信电话株式会社

地址 日本东京都

[72] 发明人 北原正树 上仓一人 八岛由幸

[56] 参考文献

CN1450816A 2003.10.22

JP6225274A 1994.8.12

JP9275578A 1997.10.21

审查员 慈 雪

[74] 专利代理机构 中国专利代理(香港)有限公司  
代理人 浦柏明 刘宗杰

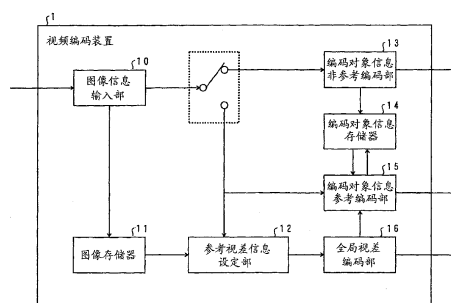
权利要求书4页 说明书22页 附图17页

[54] 发明名称

视频编码方法及装置、视频解码方法及装置

[57] 摘要

一种将多个图像编码为一个视频的视频编码方法，其具备：参考视差信息设定步骤，在对分配给第1图像内预定位置的编码对象信息进行编码时，设定表示该第1图像与编码完毕的第2图像之间的视点偏差的参考视差信息；对应位置设定步骤，利用上述参考视差信息，设定上述第2图像中与上述预定位置相关联的对应位置；代码表设定步骤，设定与上述对应位置的编码对象信息相对应的代码表；以及编码对象信息编码步骤，使用上述代码表对上述第1图像内的上述编码对象信息进行编码。



1. 一种视频编码方法，将多个图像编码为一个视频时，将从编码对象图像求得的编码对象信息进行熵编码，其特征在于，具备：

参考视差信息设定步骤，在对分配给第 1 图像内预定位置的编码对象信息进行编码时，设定表示该第 1 图像与编码完毕的第 2 图像之间的视点偏差的参考视差信息；

对应位置设定步骤，利用上述参考视差信息，设定上述第 2 图像中与上述预定位置相关联的对应位置；

代码表设定步骤，设定与上述第 2 图像中的上述对应位置的编码对象信息相对应的代码表；以及

编码对象信息编码步骤，使用上述代码表对上述第 1 图像内的上述编码对象信息进行编码。

2. 如权利要求 1 所述的视频编码方法，其中，在上述参考视差信息设定步骤中，将表示上述第 1 及第 2 图像间的平均视点偏差的全局视差信息设定为上述参考视差信息，并且该视频编码方法还具备对上述全局视差信息进行编码的全局视差信息编码步骤。

3. 如权利要求 1 所述的视频编码方法，其中，在上述参考视差信息设定步骤中，利用上述第 2 图像和另外的第 3 图像来推断两幅图像之间的视差，基于该视差和上述第 1~第 3 图像的摄像装置间的位置关系，设定上述参考视差信息。

4. 如权利要求 1 所述的视频编码方法，其中，在上述代码表设定步骤中，对代码表进行设定，该代码表是与上述第 2 图像内的上述对应位置的编码对象信息和对上述第 1 图像内上述预定位置的周边位置编码完毕的编码对象信息相对应的代码表。

5. 如权利要求 1 所述的视频编码方法，其中，在上述代码表设定步骤中，还参考与上述第 2 图像内上述对应位置的周边位置相对应的编码对象信息来设定上述代码表。

6. 如权利要求 1 所述的视频编码方法，其中，所述第 1 图像和第 2 图像的上述编码对象信息针对图像分割后所得的每个区块而设定，在上述对应位置设定步骤中，当将上述第 2 图像中与上述预定位置相对应的区块的位置基于上述参考视差信息进行偏移时，将与该区块的重叠面积为最大的区块的位置设定为对应位置。

7. 如权利要求 1 所述的视频编码方法, 其中, 所述第 1 图像和第 2 图像的上述编码对象信息针对图像分割后所得的每个区块而设定, 在上述对应位置设定步骤中, 将上述第 2 图像中与上述预定位置相对应的区块的位置基于上述参考视差信息进行偏移, 将偏移后的位置周边的多个区块的位置设定为对应位置。

8. 如权利要求 4 或 5 所述的视频编码方法, 其中, 在上述代码表设定步骤中, 设定与上述代码表设定中所使用的编码对象信息的众数相对应的代码表。

9. 如权利要求 4 或 5 所述的视频编码方法, 其中, 在上述代码表设定步骤中, 设定与上述代码表设定中所使用的编码对象信息的平均值相对应的代码表。

10. 一种视频解码方法, 对将多个图像编码为一个视频时通过将编码对象图像求得的编码对象信息进行熵编码所生成的编码数据进行解码, 其特征在于, 具备:

参考视差信息设定步骤, 在对分配给第 1 图像内预定位置的编码对象信息进行解码时, 设定表示该第 1 图像与编码完毕的第 2 图像之间的视点偏差的参考视差信息;

对应位置设定步骤, 利用上述参考视差信息, 设定上述第 2 图像中与上述预定位置相关联的对应位置;

代码表设定步骤, 设定与上述第 2 图像中的上述对应位置的编码对象信息相对应的代码表; 以及

编码对象信息解码步骤, 使用上述代码表对上述第 1 图像内的上述编码对象信息进行解码。

11. 如权利要求 10 所述的视频解码方法, 其中, 在上述参考视差信息设定步骤中将表示上述第 1 及第 2 图像间的平均视点偏差的全局视差信息解码, 并将其设定为上述参考视差信息。

12. 如权利要求 10 所述的视频解码方法, 其中, 在上述参考视差信息设定步骤中, 利用上述第 2 图像和另外的第 3 图像来推断两幅图像之间的视差, 基于该视差和上述第 1~第 3 图像的摄像装置间的位置关系, 设定上述参考视差信息。

13. 如权利要求 10 所述的视频解码方法, 其中, 在上述代码表设定步骤中, 对代码表进行设定, 该代码表是与上述第 2 图像内的上述

对应位置的编码对象信息和对上述第 1 图像内上述预定位置的周边位置编码完毕的编码对象信息相对应的代码表。

14. 如权利要求 10 所述的视频解码方法, 其中, 在上述代码表设定步骤中, 还参考与上述第 2 图像内上述对应位置的周边位置相对应的编码对象信息来设定上述代码表。

15. 如权利要求 10 所述的视频解码方法, 其中, 所述第 1 图像和第 2 图像的上述编码对象信息针对图像分割后所得的每个区块而设定, 在上述对应位置设定步骤中, 当将上述第 2 图像中与上述预定位置相对应的区块的位置基于上述参考视差信息进行偏移时, 将与该区块的重叠面积为最大的区块的位置设定为对应位置。

16. 如权利要求 10 所述的视频解码方法, 其中, 所述第 1 图像和第 2 图像的上述编码对象信息针对图像分割后所得的每个区块而设定, 在上述对应位置设定步骤中, 将上述第 2 图像中与上述预定位置相对应的区块的位置基于上述参考视差信息进行偏移, 将偏移后的位置周边的多个区块的位置设定为对应位置。

17. 如权利要求 13 或 14 所述的视频解码方法, 其中, 在上述代码表设定步骤中, 设定与上述代码表设定中所使用的编码对象信息的众数相对应的代码表。

18. 如权利要求 13 或 14 所述的视频解码方法, 其中, 在上述代码表设定步骤中, 设定与上述代码表设定中所使用的编码对象信息的平均值相对应的代码表。

19. 一种视频编码装置, 将多个图像编码为一个视频时, 将从编码对象图像求得的编码对象信息进行熵编码, 其特征在于, 具备:

参考视差信息设定部, 在对分配给第 1 图像内预定位置的编码对象信息进行编码时, 设定表示该第 1 图像与编码完毕的第 2 图像之间的视点偏差的参考视差信息;

对应位置设定部, 利用上述参考视差信息, 设定上述第 2 图像中与上述预定位置相关联的对应位置;

代码表设定部, 设定与上述第 2 图像中的上述对应位置的编码对象信息相对应的代码表; 以及

编码对象信息编码部, 使用上述代码表对上述第 1 图像内的上述编码对象信息进行编码。

20. 如权利要求 19 所述的视频编码装置, 其中, 上述参考视差信息设定部将表示上述第 1 及第 2 图像间的平均视点偏差的全局视差信息设定为上述参考视差信息, 并且该视频编码装置还具备对上述全局视差信息进行编码的全局视差信息编码部。

21. 如权利要求 19 所述的视频编码装置, 其中, 上述参考视差信息设定部利用上述第 2 图像和另外的第 3 图像来推断两幅图像之间的视差, 基于该视差和上述第 1~第 3 图像的摄像装置间的位置关系, 设定上述参考视差信息。

22. 一种视频解码装置, 对将多个图像编码为一个视频时通过将编码对象图像求得的编码对象信息进行熵编码所生成的编码数据进行解码, 其特征在于, 具备:

参考视差信息设定部, 在对分配给第 1 图像内预定位置的编码对象信息进行解码时, 设定表示该第 1 图像与编码完毕的第 2 图像之间的视点偏差的参考视差信息;

对应位置设定部, 利用上述参考视差信息, 设定上述第 2 图像中与上述预定位置相关联的对应位置;

代码表设定部, 设定与上述第 2 图像中的上述对应位置的编码对象信息相对应的代码表; 以及

编码对象信息解码部, 使用上述代码表对上述第 1 图像内的上述编码对象信息进行解码。

23. 如权利要求 22 所述的视频解码装置, 其中, 上述参考视差信息设定部将表示上述第 1 及第 2 图像间的平均视点偏差的全局视差信息解码, 并将其设定为上述参考视差信息。

24. 如权利要求 22 所述的视频解码装置, 其中, 上述参考视差信息设定部利用上述第 2 图像和另外的第 3 图像来推断两幅图像之间的视差, 基于该视差和上述第 1~第 3 图像的摄像装置间的位置关系, 设定上述参考视差信息。

## 视频编码方法及装置、视频解码方法及装置

### 技术领域

本发明涉及用于多视点运动图像编码的视频编码方法及其装置、用于实现该视频编码方法的视频编码程序及记录该程序的计算机可读的记录介质、对利用该视频编码技术所生成的编码数据进行解码的视频解码方法及其装置、以及用于实现该视频解码方法的视频解码程序及记录该程序的计算机可读的记录介质。

本申请要求享有 2005 年 1 月 7 日提出的日本专利申请特愿 2005-002423 号的优先权，在此援引其内容。

### 背景技术

多视点运动图像是使用位于各种位置的多个摄像机拍摄同一被拍摄体和背景所得的多个运动图像。以下将使用一个摄像机拍摄的运动图像称为“二维运动图像”，将针对同一被拍摄体和背景所拍摄的二维运动图像的集合称为“多视点运动图像”。

多视点运动图像中所包含的各个摄像机的二维运动图像在时间方向上具有很强的相关性。另一方面，当各摄像机同步时，各摄像机在同一时间的帧所拍摄的是完全相同状态的被拍摄体和背景，因此，各摄像机之间具有很强的相关性。

首先，叙述与二维运动图像的编码方式相关的现有技术。

在以国际编码标准 H.264、MPEG-4、MPEG-2 为代表的现有的大多数的二维运动图像编码方式中，利用动态补偿、正交变换、量子化、熵编码这样的技术来实现高效编码。

以 H.264 为例，在 I 帧中可以利用帧内相关进行编码，在 P 帧中可以利用过去的多个帧的帧间的相关进行编码，在 B 帧中则可以利用过去或未来的多个帧的每一个帧的帧间相关进行编码。

在 I 帧中，帧被分割为块（这种块称为宏块，块大小为 16×16），在各个宏块内进行帧内预测。在执行帧内预测时，将各宏块进一步分割为小块（以后称为子块），对各个子块可以执行不同的帧内预测方

法。

另一方面，在 P 帧中，可以对各个宏块进行帧内预测或帧间预测。P 帧中的帧内预测与 I 帧的情形相同。另一方面，在帧间预测时进行动态补偿。在动态补偿中，也将宏块分割为更小的区块，各子块可以具备不同的运动矢量、参考图像。

另外，在 B 帧中也可以进行帧内预测和帧间预测，但在 B 帧的帧间预测中，除了过去的帧之外，还可以将未来的帧也用于动态补偿的参考图像。例如，在以 I 帧→B 帧→B 帧→P 帧这样的帧构成进行编码的情况下，编码顺序为 I→P→B→B。此外，在 B 帧中可以参考 I 帧及 P 帧进行动态补偿。另外，与 P 帧的情形相同，将宏块分割后所得的各个子块可以具备不同的运动矢量。

执行帧内或帧间预测时会得到预测残差，在各个宏块，对预测残差块执行 DCT（离散余弦变换），进行量子化。具体地说，块大小为  $16 \times 16$  的宏块被分割为  $4 \times 4$  的子块，执行  $4 \times 4$  的 DCT。此外，DCT 系数的量子化值序列用以下信息表示。

- 编码区块图形 (Coded block pattern): 表示宏块内可定义的 4 个  $8 \times 8$  区块之中哪一个区块内存在非零 DCT 系数（以后称为非零系数）的信息

- 编码区块标志 (Coded block flag): 表示存在非零系数的  $8 \times 8$  区块内 4 个  $4 \times 4$  区块之中哪一个  $4 \times 4$  区块内存在非零系数的信息

- 重要映射 (Significance Map): 表示编码区块标志所示的  $4 \times 4$  区块内存在的 DCT 系数之中哪一个系数是非零系数的信息

- 级别信息: 表示重要映射所示的非零系数的值

除了与上述 DCT 系数相关的信息之外，针对各宏块分割为子块的方法或运动矢量等实施被称为熵编码 (Entropy Encoding) 的可逆编码，结束编码。

这里，将像素区域的量子化值或针对图像区块进行正交变换后的变换系数的量子化值（如果是 DCT 系数，则对应于上述级别信息）之外的熵编码的对象信息称为“辅助信息”。如果是 H.264，例如以下信息都可以作为上述 DCT 系数相关的辅助信息之外的辅助信息。这些辅助信息以宏块为单位提供。

- 宏块类型及子宏块类型: 宏块类型通过索引表示在宏块内执行

帧内预测还是帧间预测的指定、预测方法、块的分割方式等的组合。另一方面，子宏块类型通过索引表示在子块内的预测方法、块的分割方式等的组合。

· 参考图像索引：是各子块的动态补偿中所使用的参考图像的索引值。

· 运动矢量：是各子块中的运动矢量。在 H.264 中，表示为以周围的运动矢量预测的残差。

接着说明与熵编码相关的概要。

熵编码是可逆编码。这里，一般来说，所谓的可逆编码是将编码对象的符号 (symbol) (可以理解为从整数集合中提取的一个值) 变换为由 1 和 0 形成的比特串的处理。例如，当编码对象的符号为 {0, 1, 2, 3} 这一整数集合中的一个数值时，将符号 0 编码为 00、将符号 1 编码为 01、将符号 2 编码为 10、将符号 3 编码为 11，成为可逆编码。这种编码方法称为固定长度编码。这里，用来对符号进行编码的代码的集合 (在本例中是 {00, 01, 10, 11}) 称为“代码表”。

虽然固定长度编码是可逆编码，但从编码效率方面来看，其效率很差。在信息理论中，人们已经知道可以通过符号的出现概率 (关于整数集合的概率分布) 来实现高效率的可逆编码。大致是，为出现概率高的符号分配较短的编码长度，为出现概率低的符号分配较长的编码长度，由此，平均起来就可以获得比固定长度编码更高效的编码。依照此种方式，利用概率分布进行的可逆编码称为熵编码。

但是，实现如上述那样的高效率的熵编码的前提是编码对象的符号的概率分布是已知的。因此，一直以来所采用的方法是凭经验决定概率分布或边编码边学习符号的概率分布。此外，根据符号的概率分布求取最优代码表的方法一直以来也是存在的 (哈夫曼编码或算术编码)。因此，以下将概率分布和代码表看作是同义的。

在将熵编码用于运动图像编码的辅助信息、像素值、变换系数值的编码时，这些编码对象的信息在图像内概率会变化。因此，为了实现高效率的编码，必须随着在图像内的位置不同而适当地切换编码中使用的代码表。即，必须适当地切换概率分布。

在 H.264 中，利用所谓的内容自适应二进制算术编码 (Context-Adaptive Binary Arithmetic Coding(CABAC)) 方法来实现 (参考文献

1: 下述非专利文献 1)。下面以对宏块类型进行编码的情形为例表示 H.264 中的 CABAC 的概要。

在 CABAC 中, 在对某个宏块的宏块类型进行编码时, 参考该宏块的上方及左方的宏块的已完成编码的宏块类型来切换代码表。

图 17 表示这一参考关系的概念图。图中描述为 A、B 的宏块的宏块类型与编码对象的宏块类型具有强相关。

在 CABAC 中, 利用该相关预测最优代码表。具体地说, 对于宏块 A、B, 为所有可能的宏块类型值的组合分别分配代码表, 并利用与实际的 A、B 宏块类型值相对应的代码表(概率分布)对编码对象的宏块类型进行熵编码。其他编码对象的信息也根据同样的考虑方法进行熵编码。

接着, 说明现有的多视点运动图像的编码方法。

在现有的多视点运动图像的编码中, 利用上述时间方向及视差方向(摄像机之间)的相关来提高编码效率, 为此, 采用在时间方向进行了预测并在摄像机之间执行了补偿的编码。参考文献 2(下述非专利文献 2)中所示的手法就是这样的一个实例。

在该手法中, 将被称作“GOP”的帧集合分类为 Base GOP 和 Inter GOP 这两种 GOP 进行编码。在被称为 Base GOP 的 GOP 中, 同一摄像机的 GOP 图像中所包含的全部帧都使用帧内或帧间预测进行编码, 但在被称为 Inter GOP 的 GOP 所包含的帧中, 除了这样的帧内、帧间预测之外, 也可以使用视差方向预测, 其中利用了其他摄像机的 GOP 图像。这里所说的视差方向预测是指: 在对某个摄像机的图像的宏块进行编码时, 使用其他摄像机的图像作为参考图像, 进行与动态补偿相同的处理。

图 18 表示这种方法中的 GOP 结构的一个实例。在该图例中, 各个 GOP 具有 2 个图像, 各个箭头表示图像之间的参考关系。在该方法中, 编码时利用了时间方向及视差方向的相关, 故与只利用时间方向的相关的方法相比, 能够提高编码效率。

非专利文献 1: Detlev Marpe, et.al., "Context-Based Adaptive Binary Arithmetic Coding in the H.264/AVC Video Compression Standard", IEEE Transactions on Circuits and Systems for Video Technology, Vol.13. No.7, pp.620-636, July, 2003.

非专利文献 2: Hideaki Kimata and Masaki Kitahara, "Preliminary results on multiple view video coding(3DAV), "document M10976 MPEG Redmond Meeting, July, 2004.

但是, 在参考文献 2 的手法中, 在对编码对象的多视点运动图像进行编码时, 开始时首先根据编码效率决定是使用时间方向的预测进行编码还是使用视差方向的预测进行编码, 据此, 如果决定使用时间方向的预测进行编码, 其后的编码中就概不考虑视差方向的相关了。

这种情况下, 在编码对象的多视点运动图像中如果被拍摄体和背景随时间而产生的变化平缓、时间方向的相关超过了摄像机之间的相关, 则按照参考文献 2 的方法, 其存在的问题是, 与只利用时间方向的相关的方法相比, 无法提高编码效率。

这是因为, 如果将参考文献 2 的方法应用于如上述那样的多视点运动图像, 就总是会使用时间方向的预测, 这样就成了与仅使用时间方向的相关的方法几乎相同的编码。

但是, 在仅执行时间方向预测的情况下, 预测残差及其正交变换系数、运动矢量及宏块类型这样的辅助信息在摄像机之间也存在相关。因此, 在这些信息的编码中, 可以考虑使用摄像机间的相关。

另一方面, 在以将动态补偿的预测误差按照像素区域量子化、对该量子化值进行熵编码这样的编码方法为前提的情况下, 可以考虑应用参考文献 3 (盐寺太一郎、松田一朗、伊东晋, “动态补偿和基于三维预测的运动图像的可逆编码~关于内容建模的研究~) 的手法, 这是一种对一个二维运动图像进行编码的方法。

在该参考文献 3 中, 在对预测残差的量子化值进行熵编码时, 利用在各区块中求取的运动矢量, 通过参考上一帧已经编码结束的量子化值来切换代码表。

具体地说, 假设当前编码对象的像素位置为  $(x, y)$ , 该像素所包含的区块的运动矢量为  $(m_x, m_y)$ , 则参考与上一帧  $(x+m_x, y+m_y)$  位置周边的像素所对应的量子化值之和切换代码表。与  $(x+m_x, y+m_y)$  位置周边的像素所对应的量子化值与编码对象的像素的量子化值存在相关, 因此使用这种手法能够提高编码效率。

如果将该手法应用于多视点运动图像编码中的 (像素区域中) 预测残差的量子化值的熵编码, 则可以求取各区块中相邻摄像机在同一

帧中的视差矢量，在视差方向上执行参考文献 3 的手法中在时间方向上所执行的处理。

但是，在这种可简单类推的方法中，虽然能够提高预测残差的量子化值本身的编码效率，但必须对与各区块相应的视差矢量进行编码，因此，难以提高整体的编码效率。

另外，在这种可简单类推的方法中，无法对像素区域的预测残差之外的编码对象信息（例如正交变换系数、运动矢量或宏块类型等辅助信息）进行高效的编码。这是因为，像素区域的预测残差与其他编码对象信息之间不存在相关。

### 发明内容

本发明是鉴于这些情况而作成的，目的是提供一种新的编码技术，以便在对多视点运动图像进行编码时，能够在熵编码中利用视差方向的相关进行编码，由此，针对包括借助于现有的视差方向预测无法提高编码效率的多视点运动图像在内的各种多视点运动图像，都能够提高其编码效率。

鉴于上述情况，本发明所提供的视频编码方法是一种将多个图像（image）编码为一个视频（video image）的视频编码方法，其具备：参考视差（reference parallax）信息设定步骤，在对分配给第 1 图像内预定位置的编码对象信息进行编码时，设定表示该第 1 图像与编码完毕的第 2 图像之间的视点偏差的参考视差信息；对应位置设定步骤，利用上述参考视差信息，设定上述第 2 图像中与上述预定位置相关联的对应位置；代码表设定步骤，设定与上述对应位置的编码对象信息相对应的代码表；以及编码对象信息编码步骤，使用上述代码表对上述第 1 图像内的上述编码对象信息进行编码。

一个典型实例是，在上述参考视差信息设定步骤中，将表示上述第 1 及第 2 图像间的平均视点偏差的全局视差（global parallax）信息设定为上述参考视差信息，并且该视频编码方法还具备对上述全局视差信息进行编码的全局视差信息编码步骤。

另一典型实例是，在上述参考视差信息设定步骤中，利用上述第 2 图像和另外的第 3 图像来推断两幅图像之间的视差，基于该视差和上述第 1~第 3 图像的摄像装置间的位置关系，设定上述参考视差信息。

借助于具备这种结构的本发明，可以不对所参考的视差信息（用来表示视差矢量之类的视差大小的信息）进行编码，或者即使对其编码，数据量也很小；并且，能够利用视差方向的相关对成为编码对象的编码对象信息进行编码，由此，能够高效地对多个视频（例如多视点运动图像）进行编码。

在其实现中，不是针对每个区块求取视差信息进行编码，而是推断出表示不同摄像机之间的视点偏差的视差信息，或者针对一个图像求取一个平均视差信息（全局视差信息）进行编码。

各摄像机的摄像机参数是已知的，当有 2 个以上的摄像机的解码图像可以参考的情况下，能够不使用编码对象图像（第 1 图像）而推断出与编码对象图像的各像素相关的视差信息（视差矢量）。

另外，当摄像机排成一行而视轴平行的情况下，各摄像机在同一时刻的图像列等价于被拍摄体平行移动时所拍摄下来的运动图像，因此，仅使用参考视差信息就能够近似求取局部的视差信息。

此外，利用上述那样的视差信息，参考其他图像的编码对象信息，切换编码对象信息的熵编码所使用的代码表。

一个恰当的实例是，在上述代码表设定步骤中，对代码表进行设定，该代码表是与上述第 2 图像内的上述对应位置的编码对象信息和对上述第 1 图像内的上述预定位置周边的位置编码完毕的编码对象信息相对应的代码表。

另一个恰当的实例是，在上述代码表设定步骤中，还参考与上述第 2 图像内的上述对应位置周边的位置相对应的编码对象信息来设定上述代码表。由此，就可以获得稳固的代码表切换。

由此所获得的效果是，即使在存在参考视差信息的精度等的影响的情况下，在设定上述第 2 图像内的对应位置时，也能够应对上述第 1 图像内的上述预定位置在统计上的偏差。

即，通过参考周围的多个对应位置来设定代码表，就可以消除参考视差信息的精度等的影响。

一个典型实例是，上述编码对象信息针对图像分割后所得的每个区块而设定，在上述对应位置设定步骤中，当将上述第 2 图像中与上述预定位置相对应的区块的位置基于上述参考视差信息进行偏移时，将与该区块的重叠面积为最大的区块的位置设定为对应位置。

这种情况下，代码表设定单元通过参考该对应位置的区块的编码对象信息来切换代码表。通过采用这种结构，在基于区块的编码中，能够实现针对各区块求得的编码对象信息（正交变换系数、宏块类型等）的高效编码。

另一个典型实例是，上述编码对象信息针对图像分割后所得的每个区块而设定，在上述对应位置设定步骤中，将上述第 2 图像中与上述预定位置相对应的区块的位置基于上述参考视差信息进行偏移，将偏移后的位置周边的多个区块的位置设定为对应位置。

这时，代码表设定单元通过参考该多个区块的编码对象信息来切换代码表。通过采用这种结构，在基于区块的编码中，基于上述同样理由，可以获得稳固的代码表切换。

在上述代码表设定步骤中，也可以设定与上述代码表设定中所使用的编码对象信息的众数（most frequently appearing value）相对应的代码表。

通过采用这种结构，既防止了代码表数量的爆炸式增加，又能够对宏块类型及预测模式之类的、虽然表示为数值但只是用作索引数值的、无法定义大小关系的编码对象信息进行高效编码。

即，当所参考的对应位置数量为  $N$  时，如果为  $N$  个符号的全部组合准备代码表，则针对所参考的对应位置的数量为 1 时的代码表数量  $c$ ，就需要具备  $c^{N-1}$  个代码表。就是说，代码表数量呈指数式增加。与此相对，如果采用这种结构，则参考多个与对应位置相关的编码对象信息的众数来切换代码表，因此代码表的数量有  $c$  个即可。

另外，在上述代码表设定步骤中，也可以设定与上述代码表设定中所使用的编码对象信息的平均值相对应的代码表。

通过采用这种结构，既防止了代码表数量的爆炸式增加，又能够对正交变换系数的量子化值或非零系数的数量之类的可定义大小关系的编码对象信息进行高效编码。

如上所述，当所参考的对应位置数量为  $N$  时，如果为  $N$  个符号的全部组合准备代码表，则针对所参考的对应位置的数量为 1 时的代码表数量  $c$ ，就需要具备  $c^{N-1}$  个代码表。与此相对，如果采用这种结构，则参考多个与对应位置相关的编码对象信息的平均值来切换代码表，因此代码表的数量有  $c$  个即可。

本发明又提供一种视频解码方法，对将多个图像编码为一个视频所生成的编码数据进行解码，其特征在于，具备：参考视差信息设定步骤，在对分配给第1图像内预定位置的编码对象信息进行解码时，设定表示该第1图像与编码完毕的第2图像之间的视点偏差的参考视差信息；对应位置设定步骤，利用上述参考视差信息，设定上述第2图像中与上述预定位置相关联的对应位置；代码表设定步骤，设定与上述对应位置的编码对象信息相对应的代码表；以及编码对象信息解码步骤，使用上述代码表对上述第1图像内的上述编码对象信息进行解码。

一个典型实例是，在上述参考视差信息设定步骤中，将表示上述第1及第2图像间的平均视点偏差的全局视差信息解码，并将其设定为上述参考视差信息。

另一个典型实例是，在上述参考视差信息设定步骤中，利用上述第2图像和另外的第3图像来推断两幅图像之间的视差，基于该视差和上述第1~第3图像的摄像装置间的位置关系，设定上述参考视差信息。

一个恰当的实例是，在上述代码表设定步骤中，对代码表进行设定，该代码表是与上述第2图像内的上述对应位置的编码对象信息和对上述第1图像内的上述预定位置周边的位置编码完毕的编码对象信息相对应的代码表。

另一个恰当的实例是，在上述代码表设定步骤中，还参考与上述第2图像内的上述对应位置周边的位置相对应的编码对象信息来设定上述代码表。

一个典型实例是，上述编码对象信息针对图像分割后所得的每个区块而设定，在上述对应位置设定步骤中，当将上述第2图像中与上述预定位置相对应的区块的位置基于上述参考视差信息进行偏移时，将与该区块的重叠面积为最大的区块的位置设定为对应位置。

另一个典型实例是，上述编码对象信息针对图像分割后所得的每个区块而设定，在上述对应位置设定步骤中，将上述第2图像中与上述预定位置相对应的区块的位置基于上述参考视差信息进行偏移，将偏移后的位置周边的多个区块的位置设定为对应位置。

在上述代码表设定步骤中，既也可以设定与上述代码表设定中所

使用的编码对象信息的众数相对应的代码表，另外也可以设定与上述代码表设定中所使用的编码对象信息的平均值相对应的代码表。

本发明又提供一种视频编码装置，这是一种将多个图像编码为一个视频的视频编码装置，其具备：参考视差信息设定部，在对分配给第1图像内预定位置的编码对象信息进行编码时，设定表示该第1图像与编码完毕的第2图像之间的视点偏差的参考视差信息；对应位置设定部，利用上述参考视差信息，设定上述第2图像中与上述预定位置相关联的对应位置；代码表设定部，设定与上述对应位置的编码对象信息相对应的代码表；以及编码对象信息编码部，使用上述代码表对上述第1图像内的上述编码对象信息进行编码。

一个典型实例是，上述参考视差信息设定部将表示上述第1及第2图像间的平均视点偏差的全局视差信息设定为上述参考视差信息，并且该视频编码装置还具备对上述全局视差信息进行编码的全局视差信息编码部。

另一个典型实例是，上述参考视差信息设定部利用上述第2图像和另外的第3图像来推断两幅图像之间的视差，基于该视差和上述第1~第3图像的摄像装置间的位置关系，设定上述参考视差信息。

本发明又提供一种视频解码装置，这是一种对将多个图像作为一个视频编码所生成的编码数据进行解码的视频解码装置，其具备：参考视差信息设定部，在对分配给第1图像内预定位置的编码对象信息进行解码时，设定表示该第1图像与编码完毕的第2图像之间的视点偏差的参考视差信息；对应位置设定部，利用上述参考视差信息，设定上述第2图像中与上述预定位置相关联的对应位置；代码表设定部，设定与上述对应位置的编码对象信息相对应的代码表；以及编码对象信息解码部，使用上述代码表对上述第1图像内的上述编码对象信息进行解码。

一个典型实例是，上述参考视差信息设定部将表示上述第1及第2图像间的平均视点偏差的全局视差信息解码，并将其设定为上述参考视差信息。

另一个典型实例是，上述参考视差信息设定部利用上述第2图像和另外的第3图像来推断两幅图像之间的视差，基于该视差和上述第1~第3图像的摄像装置间的位置关系，设定上述参考视差信息。

本发明的各个方法可以利用计算机程序实现，该计算机程序可以记录在适当的计算机可读的记录介质而提供、或经由网络提供，在实施本发明时安装，通过在 CPU 等控制单元上执行动作而实现本发明。

如上述说明，借助于本发明，针对通过视差方向预测无法提高编码效率的多视点运动图像，可以利用熵编码中视差方向的相关，因此，对于各种多视点运动图像，都能够提高其编码效率。

### 附图说明

图 1A 是表示用于代码表切换的编码对象信息的参考关系的一个实例的图。

图 1B 是同样地表示用于代码表切换的编码对象信息的参考关系的一个实例的图。

图 2A 是表示用于代码表切换的编码对象信息的参考关系的一个实例的图。

图 2B 是同样地表示用于代码表切换的编码对象信息的参考关系的一个实例的图。

图 3 是表示本发明的一个实施方式中摄像机的结构的概念图。

图 4 是表示同一实施方式中视频编码装置的结构图。

图 5 是表示同一实施方式中编码对象信息参考编码部的结构的图。

图 6 是表示同一实施方式中视频编码装置所执行的动作流程的图。

图 7 是表示同一实施方式中摄像机 C2 的图像中的编码对象信息的编码处理的动作流程的图。

图 8A 是表示用于代码表切换的编码对象信息的参考关系的一个实例的图。

图 8B 是同样地表示用于代码表切换的编码对象信息的参考关系的一个实例的图。

图 9 是表示同一实施方式中视频解码装置的结构图。

图 10 是表示同一实施方式中编码对象信息参考解码部的结构的图。

图 11 是表示同一实施方式中视频解码装置所执行的动作流程的

图。

图 12 是表示同一实施方式中摄像机 C2 的图像中的编码对象信息的解码处理的动作流程的图。

图 13A 是表示同一实施方式中用于代码表切换的编码对象信息的参考关系的一个实例的图。

图 13B 是同样地表示用于代码表切换的编码对象信息的参考关系的一个实例的图。

图 14 是表示不使用编码对象图像而推断出与该图像的各像素相关的视差信息（视差矢量）的原理的图。

图 15 是表示在推断视差信息的情况下摄像机 C2 的图像中编码对象信息的编码处理的动作流程的图。

图 16 是表示在推断视差信息的情况下摄像机 C2 的图像中编码对象信息的解码处理的动作流程的图。

图 17 是表示 H.264 中用于代码表切换的宏块的参考关系的图。

图 18 是表示现有技术的 GOP 结构的一个实例的图。

符号说明

1 视频编码装置

2 视频解码装置

10 图像信息输入部

11 图像存储器

12 参考视差信息设定部

13 编码对象信息非参考编码部

14 编码对象信息存储器

15 编码对象信息参考编码部

16 全局视差编码部

20 编码对象信息非参考解码部

21 编码对象信息存储器

22 编码对象信息参考解码部

23 全局视差解码部

24 图像信息复原部

25 图像信息输出部

### 具体实施方式

下面，按照实施方式详细说明本发明。

首先，图 1A、1B 及图 2A、2B 表示了用于代码表切换的编码对象信息的参考关系的一个实例。这里，为便于说明，假定在这些图中使用全局视差矢量（全局视差信息）作为参考视差信息。

在这些图中，假定对 2 个摄像机（摄像机 C1、摄像机 C2）所拍摄的视频进行编码，对与摄像机 C1 相关的编码对象信息（所谓的编码对象信息是成为熵编码的对象的信息的总称，有预测残差、正交变换系数、宏块类型等信息）不参考摄像机 C2 的编码对象信息地进行编码，另一方面，对摄像机 C2 的编码对象信息的熵编码则参考摄像机 C1 的已编码完毕的编码对象信息进行熵编码。

在图 1A、1B 的实例中，将与摄像机 C2 中当前时刻成为编码对象的编码对象信息所属的区块 502（在该实例中假定为 MPEG-2 之类的基于区块的编码）的位置相对应的摄像机 C1 图像中的位置 501，在摄像机 C1 的图像中偏移全局视差矢量部分，按照与由此所确定的位置相对应的 1 个区块（图中记述为 C 的区块）的编码对象信息切换代码表，使用该代码表对编码对象信息进行编码。

此外，在上述实例中，将与区块 502 的位置相对应的摄像机 C1 的图像中的位置 501 在摄像机 C1 的图像中偏移全局视差矢量部分，与由此所确定的位置相对应的 1 个区块被确定下来的情况下，一个恰当的实例是，将与区块 501 的重叠面积最大的区块（例如 C）的位置设定为对应位置。

在图 2A、2B 的实例中，将与摄像机 C2 中当前时刻成为编码对象的编码对象信息所属的区块 502（在该实例中假定为 MPEG-2 之类的基于区块的编码）的位置相对应的摄像机 C1 图像中的位置 501，在摄像机 C1 的图像中偏移全局视差矢量部分，按照与由此所确定的位置周边相对应的多个区块（图中记述为 C、D、E、F 的区块）的编码对象信息切换代码表，使用该代码表对编码对象信息进行编码。

在下面的实施方式实例中，详细说明对使用 2 个摄像机拍摄所得的多视点运动图像进行编码的情况下的具体实例。

图 3 表示本实施方式实例中使用的摄像机结构的概念图。这里，图中所示的四角形图形表示各摄像机的帧（图像）。

如本图所示，在本实施方式实例中，假定使用摄像机 C1、C2，并且摄像机 C1 的帧的全部编码对象信息先于同一时刻的摄像机 C2 的帧进行编码，在摄像机 C2 的帧的编码对象信息的熵编码中参考摄像机 C1 的（已经编码完毕的）编码对象信息。

下面，为了简化说明，假定按照图 3 中帧图中所描述的顺序进行编码。另外，在对编码对象信息进行编码时区块的扫描方法假定为从左开始扫描各行，并从上面的行开始处理。

另外，在本发明中，所采用的结构是针对每个编码对象信息切换最适合的代码表，在本实施方式实例中，假定为对宏块类型这样的编码对象信息进行编码的情形，具体说明其代码表的切换。宏块类型是表示各区块的动态补偿方法及帧内预测方法的信息，表示为整数索引的 MBtype。在本实施方式实例中，假定 MBtype 有 12 种（即，MBtype = 0, 1, ....., 11）。

另外，在对编码对象信息进行熵编码时，假定执行的是利用了与编码对象信息相对应的区块周边的已经编码完毕的编码对象信息的内容依赖型熵编码、或者执行的是利用了这些已经编码完毕的编码对象信息及其他摄像机的帧的已经编码完毕的编码对象信息的内容依赖型熵编码。此外，假定适当地预先准备了当时使用的各个代码表。

图 4 表示本实施方式实例中使用的视频编码装置 1 的结构。

该图中所示的视频编码装置 1 具备：图像信息输入部 10，按照图 3 所示顺序输入摄像机 C1、C2 的帧；图像存储器 11，存储摄像机 C1 的帧；参考视差信息设定部 12，在摄像机 C1 和摄像机 C2 的帧之间求取参考视差信息（在本实施方式中是全局视差矢量）；编码对象信息非参考编码部 13，对摄像机 C1 的编码对象信息进行编码；编码对象信息存储器 14，存储摄像机 C1 的区块信息（编码对象信息）；编码对象信息参考编码部 15，对摄像机 C2 的编码对象信息进行编码；以及全局视差编码部 16，对全局视差矢量进行编码。

图 5 表示视频编码装置 1 中编码对象信息参考编码部 15 的详细结构。

如该图所示，编码对象信息参考编码部 15 具备：对应位置设定部 150，根据全局视差矢量求取对应位置；参考参数设定部 151，基于摄像机 C1 和摄像机 C2 的编码对象信息来设定用于设定代码表的参考参

数；代码表设定部 152，基于参考参数设定代码表；编码对象信息设定部 153，根据摄像机 C2 的图像设定编码对象信息；以及编码对象信息编码部 154，对摄像机 C2 的编码对象信息进行编码。

图 6 表示这种结构的视频编码装置 1 所执行的动作流程。

如该动作流程所示，首先，摄像机 C1 和摄像机 C2 的帧按照图 3 所示的顺序交替输入到图像信息输入部 10 (S10)。

所输入的帧如果是摄像机 C1 的帧，则在编码对象信息非参考编码部 13 将各区块的编码对象信息进行编码 (S11/S12)。另外，在图像存储器 11 中有 1 帧大小的存储器，总是存储着摄像机 C1 最新的输入完毕的帧。

另一方面，所输入的帧如果是摄像机 C2 的帧，则在编码对象信息参考编码部 15 将各区块的编码对象信息进行编码 (S11/S13)。此外，在摄像机 C2 的帧的编码对象信息的解码中使用事先由参考视差信息设定部 12 取得的全局视差矢量。

此外，最后输出依照此种方式生成的编码对象信息的编码数据 (S14)。

图 7 表示摄像机 C2 的编码对象信息的编码处理的详细动作流程。

参考视差信息设定部 12 中输入来自图像存储器 11 的摄像机 C1 的图像、来自图像信息输入部 10 的摄像机 C2 的图像。此外，利用参考视差信息设定部 12 求得全局视差矢量 (S130)。

全局视差矢量的计算方法可以有多种，这里假定为针对各个区块 (例如宏块) 分别求取视差矢量的平均矢量。

即，如果将摄像机 C2 的第  $b$  ( $b=0,1,\dots,B-1$ ) 个区块的像素值表示为：

$$I_b(i,j) \quad i=0,1,\dots,I-1, \quad j=0,1,\dots,J-1$$

摄像机 C1 的图像中左上方像素的位置为  $(i',j')$ 、大小为  $I \times J$  的区块的像素值表示为：

$$I_{i',j'}(i,j) \quad i=0,1,\dots,I-1, \quad j=0,1,\dots,J-1$$

那么，在参考视差信息设定部 12 中，针对摄像机 C2 的各区块，根据以下的估价函数进行区块匹配。

$$E(b,i',j') = \sum_i \sum_j (I_b(i,j) - I_{i',j'}(i,j))^2$$

其中， $\sum_i$  是  $i=0 \sim I-1$  的总和， $\sum_j$  是  $j=0 \sim J-1$  的总和。

这里，假定在各区块中将  $E(b,i',j')$  最小化时所求得的视差矢量为  $\langle d_b \rangle$ ，则全局视差矢量  $\langle d \rangle$  可以通过例如下式求取。

$$\langle d \rangle = \Sigma \langle d_b \rangle / B$$

其中， $\Sigma$  是  $b=0 \sim B-1$  的总和。

依照此种方式求得的全局视差矢量被输入到全局视差编码部 16，通过全局视差编码部 16 进行编码 (S131)。

在本实施方式实例中，假定全局视差矢量的编码被非可逆编码。这里，假定量子化步长取 1 进行量子化 (等价于被映射到最接近的整数)，进行熵编码。

接着，针对各区块，利用编码对象信息设定部 153 求取编码对象信息 (S133)，以便将编码对象信息进行熵编码。具体地说，求取动态补偿的运动矢量、MBtype 等信息及正交变换系数等编码对象信息。

接着，针对各区块，利用编码对象信息编码部 154 将编码对象信息进行编码。此时，对应位置设定部 150 利用被编码后又被解码出来的全局视差矢量判定摄像机 C1 的帧内是否存在对应位置 (S134)。

在本实施方式实例中，如图 8A、8B 所示，对应位置定义为最接近下述位置的 4 个区块的位置，即，该位置是使与摄像机 C2 图像中的区块  $b$  相对应的摄像机 C1 图像中的位置 501 偏移全局视差矢量部分后所得的位置。具体地说，如果将当前区块  $b$  的左上角的位置定义为  $I_b$ ，则将摄像机 C1 的图像中区块左上角位置最靠近  $I_b + \langle d \rangle$  的 4 个区块的位置定义为对应位置。不过，如果  $I_b + \langle d \rangle$  与区块左上角位置的距离大于预先设定的阈值  $\maxDist$ ，则不将其定义为对应位置。

当对应位置设定部 150 判断为对应位置不存在时，熵编码时代码表的切换仅参考相同图像内已经编码完毕的区块执行。

即，在对应位置不存在的情况下，参考参数设定部 151 求取图 8B 中记述为 A、B 的区块的 MBtype 的组合。例如，可能求得 (0, 5) 这样的 MBtype 组合。

代码表设定部 152 接收到之后，设定使用与该 (0, 5) MBtype 组合相对应的代码表进行熵编码 (S137)。此外，假定这里使用的代码表与摄像机 C1 所使用的相同。另外，代码表设定部 152 针对所有可能的 MBtype 值的组合分别保存代码表。

此外，接收到该代码表设定后，编码对象信息编码部 154 输出该

代码表中与当前区块的 MBtype 相对应的编码 (S138)。即, 当前区块的编码对象信息被编码。

另一方面, 当对应位置设定部 150 判断为存在对应位置时 (图 8A 中记述为 C、D、E、F 的区块), 参考参数设定部 151 求取与 4 个对应位置相关的 MBtype 的众数, 然后基于该 MBtype 和图 8B 中记述为 A、B 的区块的 MBtype 的组合求取最终的 MBtype 组合。即, 编码对象图像中当前欲编码的区块 b 周边的已经编码完毕的编码对象信息和已经编码完毕的参考图像中对应位置的编码对象信息被参考。

例如, 假定 4 个对应位置的 MBtype 为 (0, 5, 5, 4), 则众数为 5。另外, 如果图 8B 中记述为 A、B 的区块的 MBtype 的组合为 (0, 5), 则参考参数设定部 151 就会根据这些 MBtype 最终求取 (0, 5, 5) 这样的 MBtype 的组合。

代码表设定部 152 接收到之后, 设定使用与该 (0, 5, 5) MBtype 组合相对应的代码表进行熵编码 (S135)。

此外, 接收到该代码表设定后, 编码对象信息编码部 154 输出该代码表中与当前区块的 MBtype 相对应的编码 (S136)。

执行这些处理, 直到图像中全部区块的编码对象信息被编码 (S132/S139/S140)。

依照此种方式, 本发明的视频编码装置 1 利用视差方向的相关执行处理, 以便对成为熵编码的对象的编码对象信息进行编码。

其次, 说明本发明的视频解码装置。

图 9 表示本实施方式实例中使用的视频解码装置 2 的结构。

该图中所示的视频解码装置 2 具备: 编码对象信息非参考解码部 20, 对与摄像机 C1 相关的编码对象信息进行解码; 编码对象信息存储器 21, 存储摄像机 C1 被解码后的编码对象信息; 编码对象信息参考解码部 22, 对与摄像机 C2 相关的编码对象信息进行解码; 全局视差编码部 23, 对全局视差矢量进行解码; 图像信息复原部 24, 根据编码对象信息复原图像; 以及图像信息输出部 25, 将复原后的图像输出。

图 10 表示视频解码装置 2 中编码对象信息参考解码部 22 的详细结构。

如该图所示, 编码对象信息参考解码部 22 具备: 编码对象信息解码部 220, 对摄像机 C2 的编码对象信息进行解码; 对应位置设定部

221, 根据解码后的全局视差矢量求取对应位置; 参考参数设定部 222, 基于摄像机 C1 和摄像机 C2 的已经解码完毕的编码对象信息来设定用于设定代码表的参考参数; 以及代码表设定部 223, 基于参考参数设定代码表。

图 11 表示这种结构的视频解码装置 2 所执行的动作流程。

首先, 在视频解码装置 2 中, 摄像机 C1 和摄像机 C2 的编码数据按照图 3 所示的顺序输入 (S20)。

这时, 摄像机 C1 的编码数据被输入到编码对象信息非参考解码部 20, 摄像机 C2 的编码数据则输入到编码对象信息参考解码部 22。这里, 摄像机 C2 的编码数据被输入时, 全局视差矢量的编码数据也被输入到全局视差解码部 23。

所输入的编码数据如果是摄像机 C1 的编码数据, 各区块的编码对象信息则在编码对象信息非参考解码部 20 被解码 (S21/S22)。

另一方面, 所输入的编码数据如果是摄像机 C2 的编码数据, 各区块的编码对象信息则在编码对象信息参考解码部 22 被解码 (S21/S23)。

最后, 基于依照此种方式解码后的编码对象信息将图像复原后输出 (S24)。

图 12 表示摄像机 C2 的编码对象信息的解码处理的详细动作流程。

在摄像机 C2 的编码对象信息的解码处理中, 全局视差矢量的编码数据首先输入到全局视差解码部 23, 在全局视差解码部 23 将全局视差矢量解码 (S230)。

此外, 在对应位置设定部 221, 使用解码后的全局视差矢量判断摄像机 C1 的帧内是否存在与摄像机 C2 的帧内当前区块 b 相对应的对应位置 (S232)。对应位置的決定方法与编码时相同。

当对应位置设定部 221 判断为对应位置不存在时, 仅参考摄像机 C2 的已经解码完毕的编码对象信息来执行代码表的切换, 将当前区块的编码对象信息解码 (S235/S236)。参考关系与编码时相同。

这时, 参考参数设定部 222 和代码表设定部 223 所执行的动作与在视频编码装置 1 中当对应位置不存在的情况下参考参数设定部 151 和代码表设定部 152 所执行的动作相同, 将当前区块的 MBtype 解码。

另一方面，当对应位置设定部 221 判断为对应位置存在时（与编码的情况相同，图 8A 中记作 C、D、E、F 的区块），通过参考摄像机 C2 的已经解码完毕的编码对象信息及摄像机 C1 的对应位置的已经解码完毕的编码对象信息来切换代码表，当前区块的编码对象信息被解码（S233/S234）。

这时，参考参数设定部 222 和代码表设定部 223 所执行的动作与在视频编码装置 1 中当对应位置存在的情况下参考参数设定部 151 和代码表设定部 152 所执行的动作相同，将当前区块的 MBtype 解码。

当前区块的编码对象信息全部被解码后，在图像信息复原部 24，根据这些信息将当前区块的图像信息复原，从图像信息输出部 25 输出（S237）。

执行这些处理，直到图像中的全部区块完成解码（S231/S238/S239）。

依照此种方式，本发明的视频解码装置 2 利用视差方向的相关执行处理，以便将本发明的视频编码装置 1 所生成的编码数据进行解码。

上述实施方式实例中的编码处理与宏块类型信息的熵编码相关，但本发明也可以直接应用于对其他的成为熵编码对象的信息的熵编码。下面描述几个实例。

具体描述本发明在 H.264 中的编码区块图形、编码区块标志、重要映射、级别信息中的应用，这些信息成为熵编码的对象，是与 DCT 系数关联的信息。

如上所述，在 H.264 中，成为熵编码对象的所有信息都表示为比特串，利用内容依赖型算术编码对每个比特进行熵编码。即，在应用本发明的情况下，可以用来切换对各个比特进行熵编码时的代码表（由于是算术编码，因此具体地就是概率表）。

如上所述，编码区块图形信息是表示在宏块内的 4 个  $8 \times 8$  区块中哪一个区块中存在非零系数的信息。具体地说，表示为总计 4 比特的信息，用来表示在  $8 \times 8$  区块中非零系数的有无。

在对某个  $8 \times 8$  区块的编码区块图形的比特进行编码时，以  $8 \times 8$  区块为单位执行例如图 13A、13B 所示形式的参考即可。图 13A、13B 是分别对应于图 8A、8B 的用来表示区块的位置关系的图。即，分别准备好对应于  $8 \times 8$  区块 A、B、C 的比特的全部组合的代码表，利用其

中与编码对象  $8 \times 8$  区块  $b$  最适合的 A、B、C 的比特组合相对应的代码表进行熵编码即可。

另外，编码区块标志信息如上所述是表示存在非零系数的  $8 \times 8$  区块中的 4 个  $4 \times 4$  区块中哪一个区块中存在非零系数的信息。因此，编码区块标志是与编码区块图形相同形式的信息，也可以用同样的方式应用本发明。

接着，描述重要映射的熵编码中本发明的应用方法的实例。重要映射信息如上所述是表示利用编码区块标志所表示的  $4 \times 4$  区块中哪一个 DCT 系数非零的信息。具体地说，利用 1 比特表示某个 DCT 系数是否非零，按照某个顺序扫描该  $4 \times 4$  区块中的 16 个 DCT 系数，当 DCT 系数非零时取值为 1，否则取值为 0。这些比特称作有效比特 (Significance bit)。

此外，当 DCT 系数非零从而取值为 1 时，使用 1 比特表示该 DCT 系数是否为扫描过程中最后的非零系数 (在是非零系数的情况下取值为 1，否则取值为 0)，包含到重要映射中。这些比特称作末位比特 (Last bit)。

当扫描过程中出现了末位比特的值为 1 这样的状况的情况下 (即该位置的 DCT 系数是最后的非零系数的情况下)，就不再需要后面的与 DCT 系数相关的有效比特。另外，关于扫描过程中最后的 DCT 系数是否非零这一点，根据当时的有效比特和末位比特即可明确，因此通常不需要关于该 DCT 系数的信息。

在 H.264 中，将有效比特和末位比特单独编码，代码表根据扫描过程中的位置而切换 (如上所述，不需要关于最后的 DCT 系数的信息，因此需要 15 个代码表)。在应用本发明的情况下，可以例如如图 13A、13B 所示那样参考其他摄像机的  $4 \times 4$  区块 C 中非零系数的数量切换代码表。这是因为，根据多视点运动图像的性质，在区块 C 中非零的 DCT 系数的数量越多，编码对象的  $4 \times 4$  区块  $b$  的 16 个有效比特为 1 (即 DCT 系数非零) 的概率越高，末位比特为 0 的概率越高。

最后，描述级别信息的熵编码中本发明的应用方法的实例。在 H.264 中，将整数 DCT 系数的量子化值变换为 2 进制比特串，与重要映射的情况下同样地一边扫描一边将各比特进行熵编码。

这时所使用的代码表根据在区块中到目前为止的扫描过程中所出

现的绝对值大于 1 的量子化值的数量进行切换。在扫描过程中的某个时刻，如果出现了很多绝对值大于 1 的量子化值，则该时刻成为编码对象的量子化值具有较大值的概率很高，因此，预先准备反映了该性质的代码表。

在对多视点运动图像进行编码的情况下，其性质是，如图 13A、13B 所示的其他摄像机 (C1) 的区块 C 中 DCT 系数的量子化值越大，当前成为编码对象的区块 (b) 的量子化值大的值就越多。因此，在应用本发明的情况下，预先准备与例如区块 C 中量子化值的平均值相对应的代码表，利用与编码对象的区块 b 相关的其他摄像机的区块 C 的量子化值的平均值相对应的代码表进行熵编码即可。

以上根据实施方式实例对本发明进行了说明，但本发明并不限于此。例如，在实施方式实例中，将摄像机 C1、C2 这两台摄像机所拍摄的视频作为处理对象对本发明进行了说明，但本发明也可以直接应用于 3 台以上的摄像机所拍摄的视频。

另外，在上述实施方式实例中记述的是在视频编码装置中求取参考视差信息的情况。另一方面，也可以输入利用各种视差推断手法预先在外部求得的视差信息，使用该视差信息作为参考视差进行编码处理，由此实现本发明。

另外，在实施方式实例中，假定使用视差矢量作为视差信息，但本发明也可以使用视差矢量以外的信息作为视差信息。

具体地，例如也可以使用将参考文献 4 (Aljoscha Smolic, Thomas Sikora and Jens-Rainer Ohm, "Long-Term Global Motion Estimation and its Application for Sprite Coding, Content Description, and Segmentation", IEEE Transactions on Circuits and Systems for Video Technology, Vol.9, No.8, pp.1227-1242, December, 1999) 中所说明的全局动态模型应用于视差方向的全局视差模型。

这种情况下，在视频编码装置中求取全局视差模型的参数，对这些参数进行编码。另外，在求取对应位置时，将这些参数应用于全局视差模型来求取全局视差矢量，确定对应位置。另一方面，在视频解码装置中，将解码得到的全局视差模型的参数应用于全局视差模型，以确定对应位置。

另外，在上述实施方式实例中，摄像机有 2 台，所记述的是参考

其中一台摄像机的解码图像的情形；当各摄像机的摄像机参数（摄像机的位置或姿势、镜头拍照角度等）是已知的、并且可以参考 2 台以上不同的摄像机的解码图像的情况下，就能够不使用编码对象图像而使用图 14 所示的方法推断出与编码对象的图像的各像素相关的视差信息（视差矢量）。

在图 14 中，针对被拍摄体 600 拍摄编码对照图像的摄像机 C2 的视点位置假定为 620，对同一被拍摄体 600 进行拍摄（即输出参考图像）的另外 2 台摄像机 C1、C3 的视点位置分别假定为 610、630。另外，假定在编码对照图像中需要求取视差的像素位置为 640。如该图中所示，在相对于被拍摄体 600 的真实视差中，各参考图像中与像素位置 640 相对应的对应点的像素值应该是大致相等的值。因此，在很多视差推断方法中，将相对于各种纵深的各个参考图像的对应点的像素值进行比较，基于像素值最接近的纵深（图 14 中以符号 650 来表示）和已知的摄像机参数，就能够推断出来自摄像机 C1 或 C3 的视差。该处理能够以编码对象图像的像素为单位进行。

如果在上述实施方式实例的结构中实施该方法，则不再需要视频编码装置 1 的全局视差编码部 16、视频解码装置 2 的全局视差解码部 23。此外，使用例如相对于编码对象区块的各个像素推断出来的视差矢量的平均，通过与上述实施方式实例同样的处理来实施编码即可。

具体地说，在编码对象信息的编码处理中，执行如图 15 所示的处理。这里，设置步骤 S1300，以取代图 7 的步骤 S130、S131 的处理；在该步骤中，利用与摄像机 C2 所拍摄的编码对象图像的相同时刻所拍摄的摄像机 C1、C3 的图像的解码图像，借助于上述方法推断出来自摄像机 C1 或 C3 的视差。

另外，在解码处理中，执行如图 16 所示的处理。这里，设置步骤 S2300，以取代图 12 的步骤 S230 的处理；在该步骤中，利用与摄像机 C2 所拍摄的编码对象图像相同的时刻所拍摄的摄像机 C1、C3 的图像的解码图像，借助于同样的方法推断出来自摄像机 C1 或 C3 的视差。

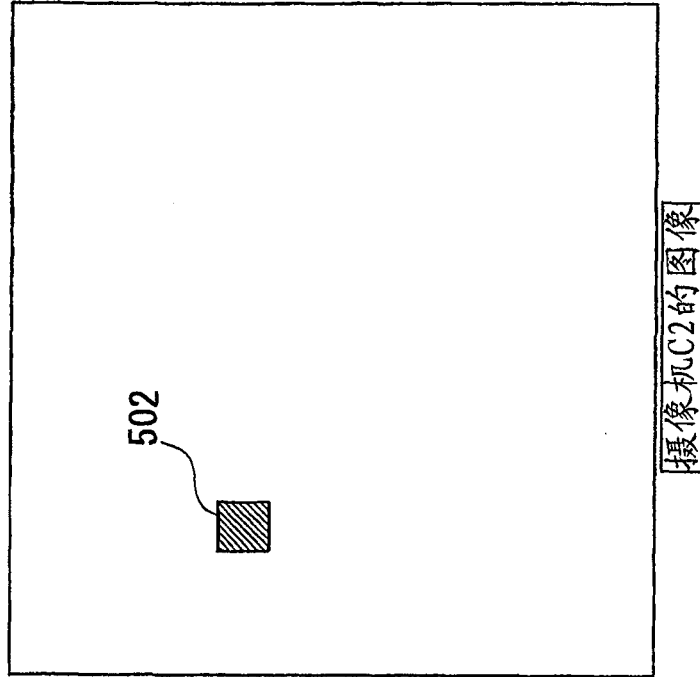


图 1B

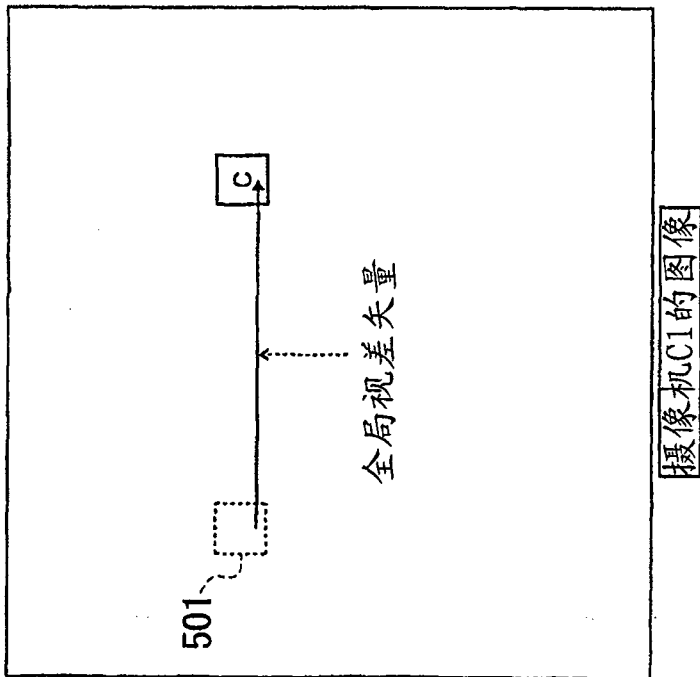


图 1A

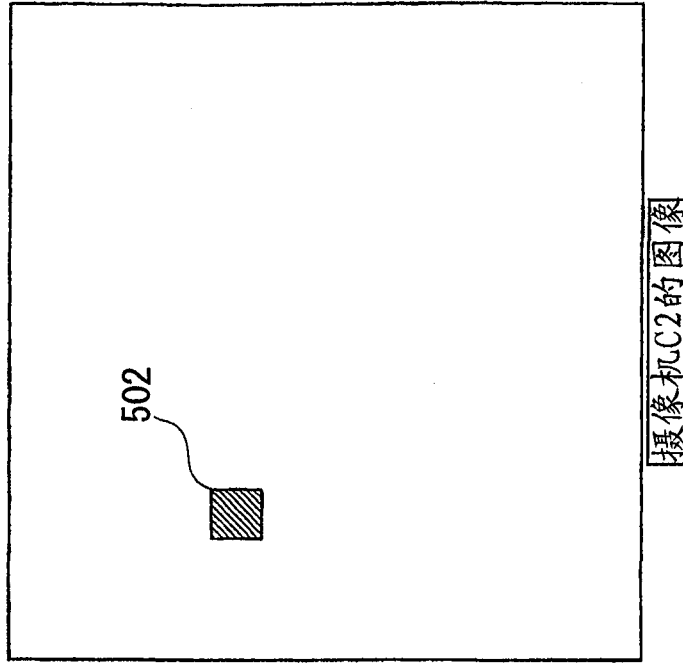


图 2B

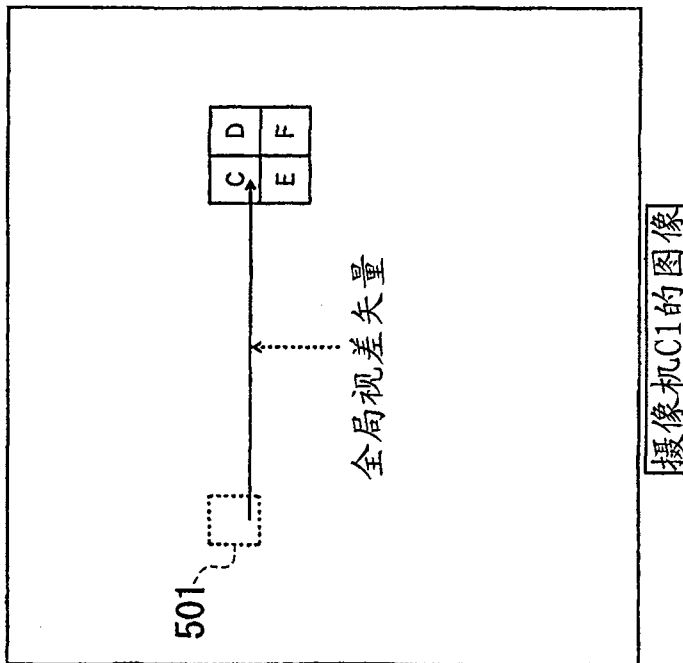


图 2A

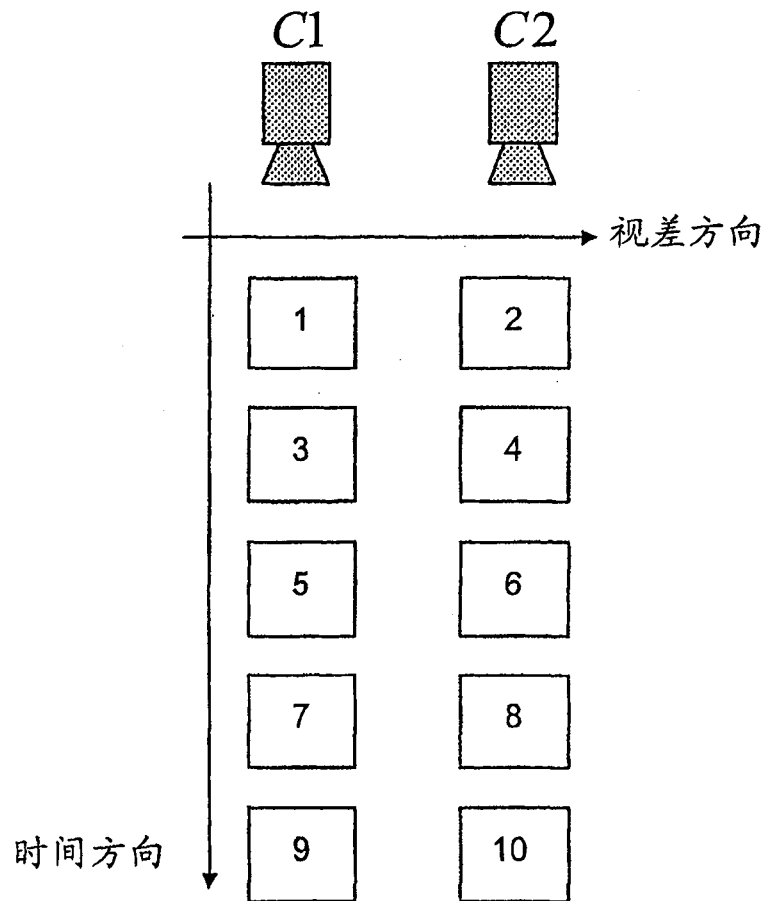


图 3

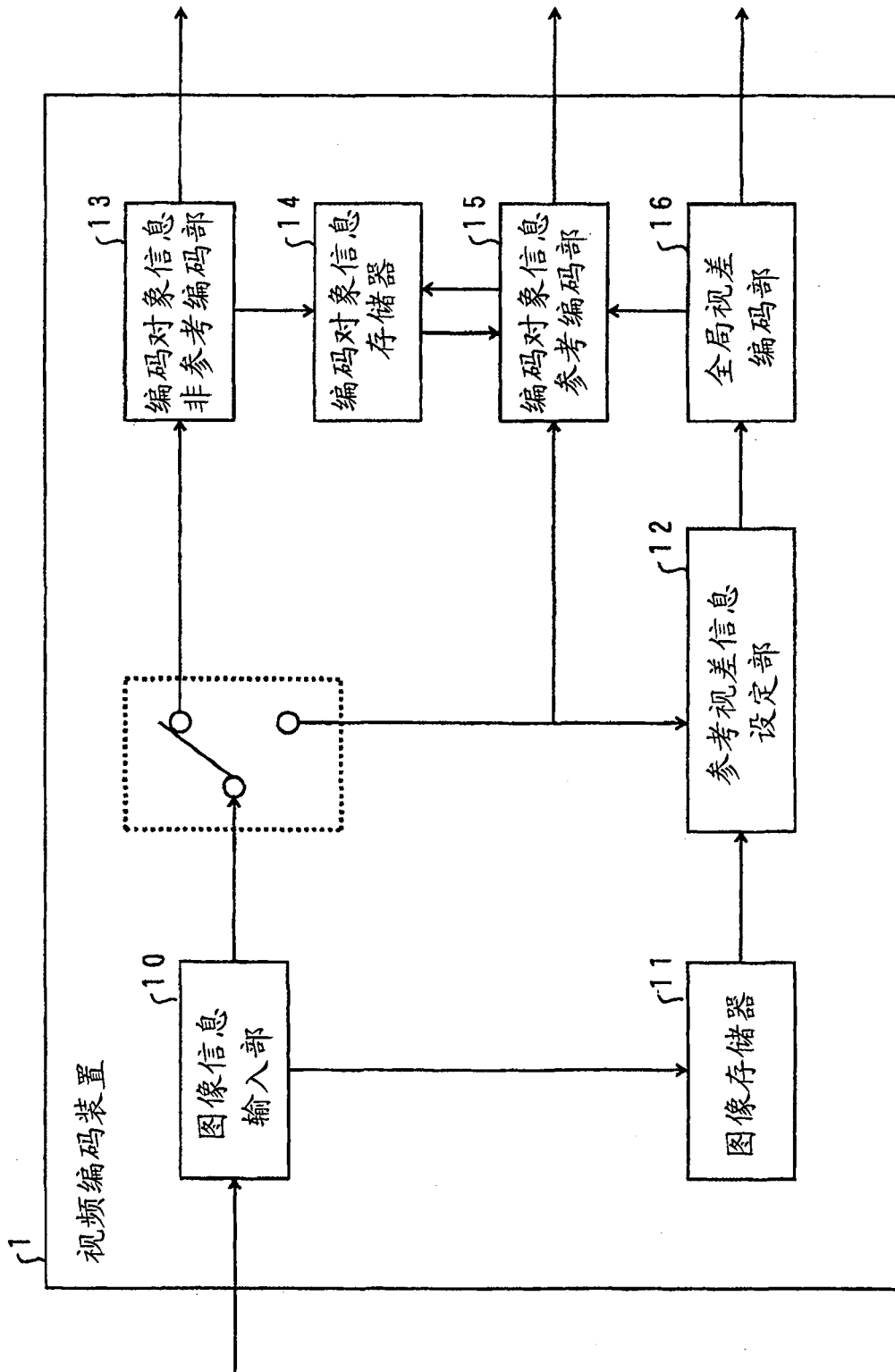


图 4

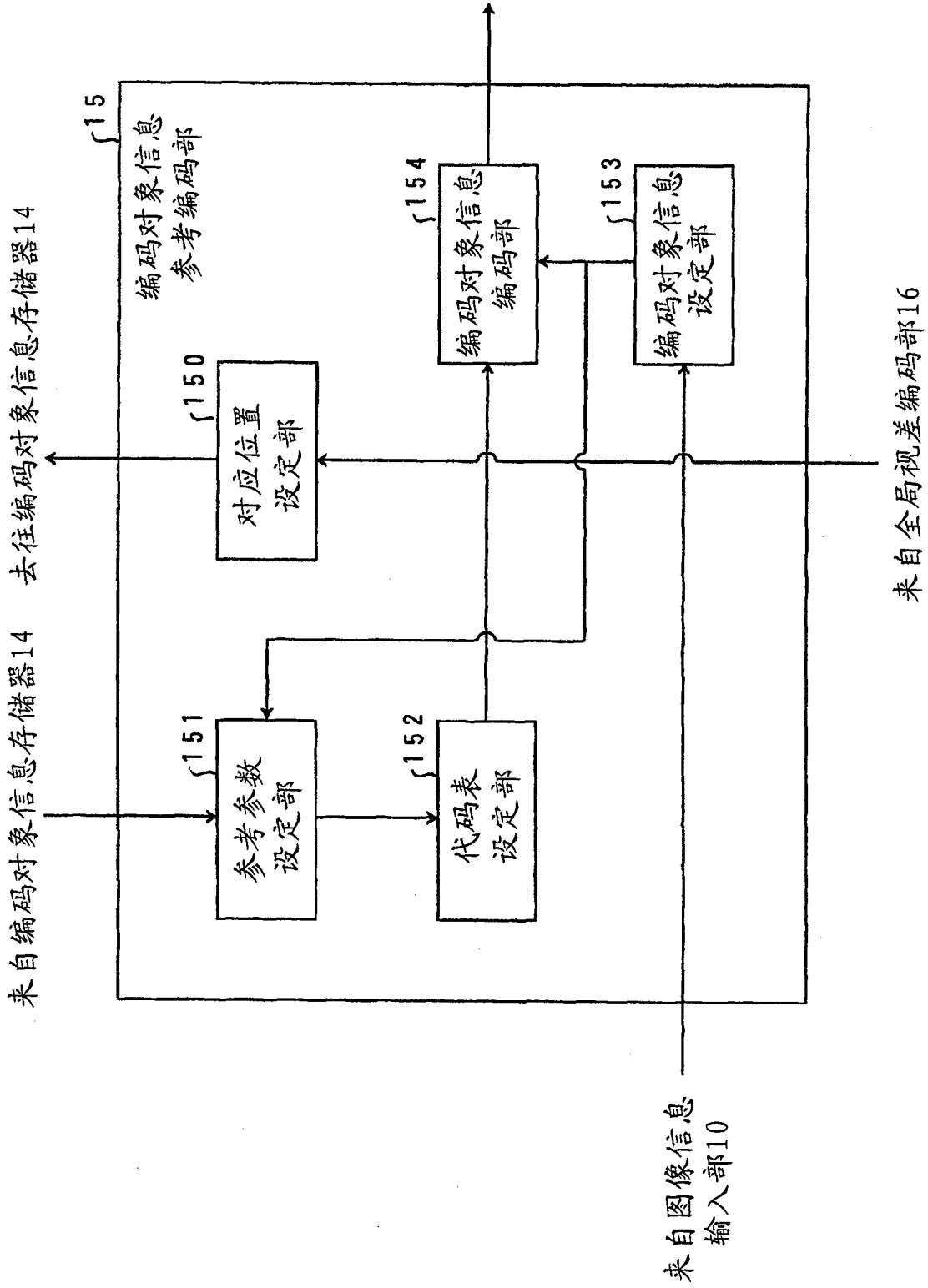


图 5

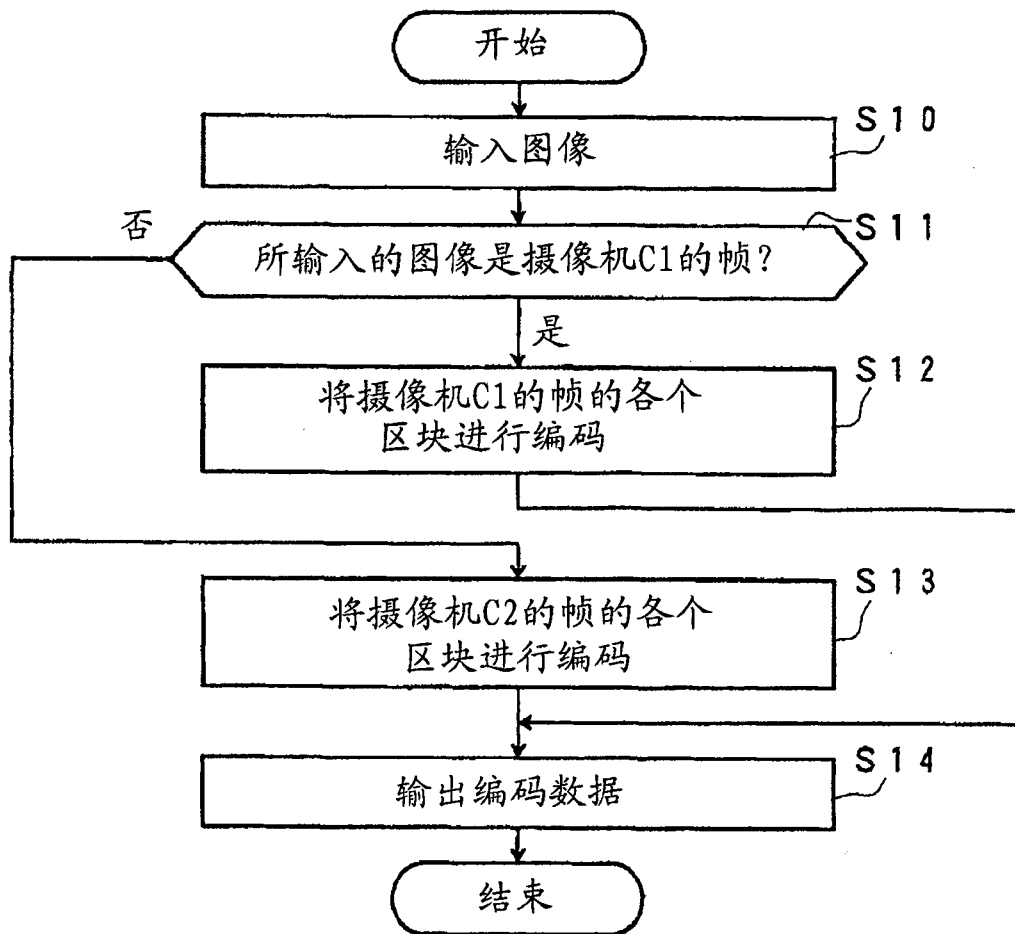


图 6

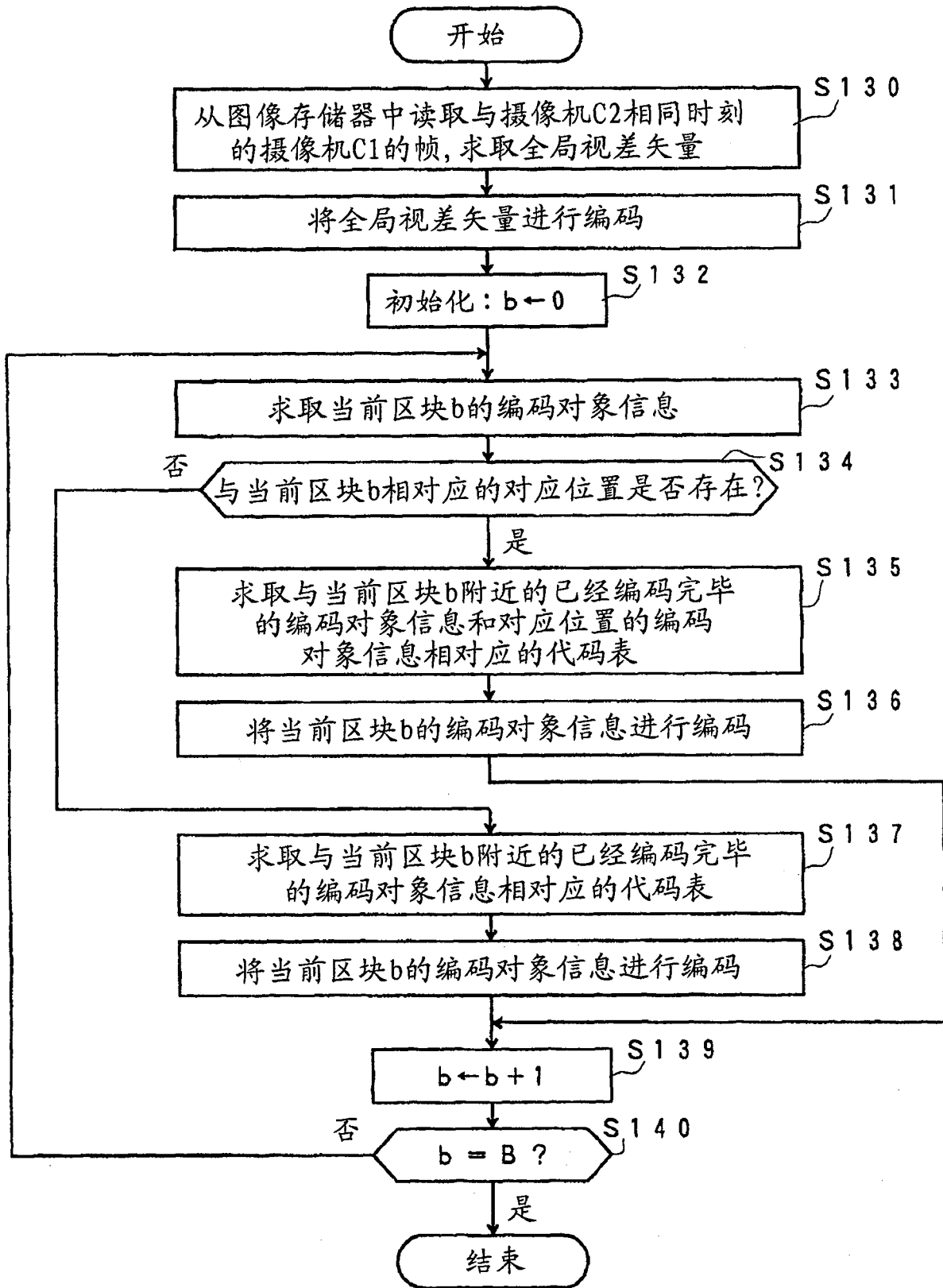
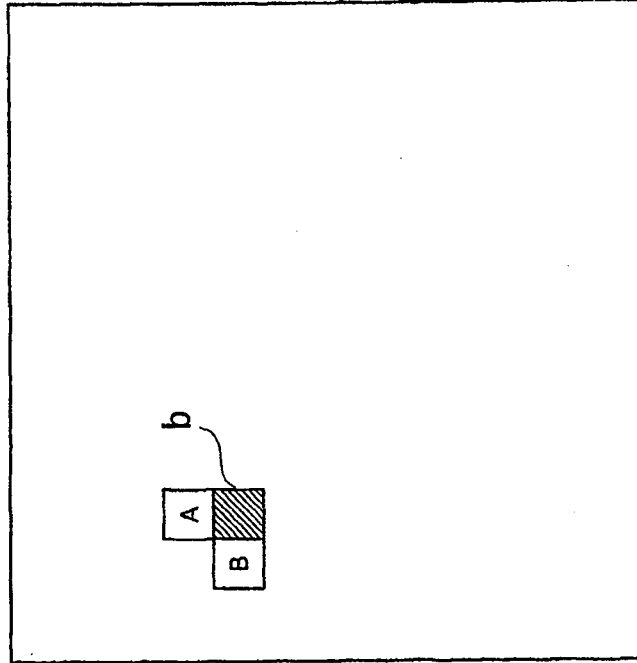
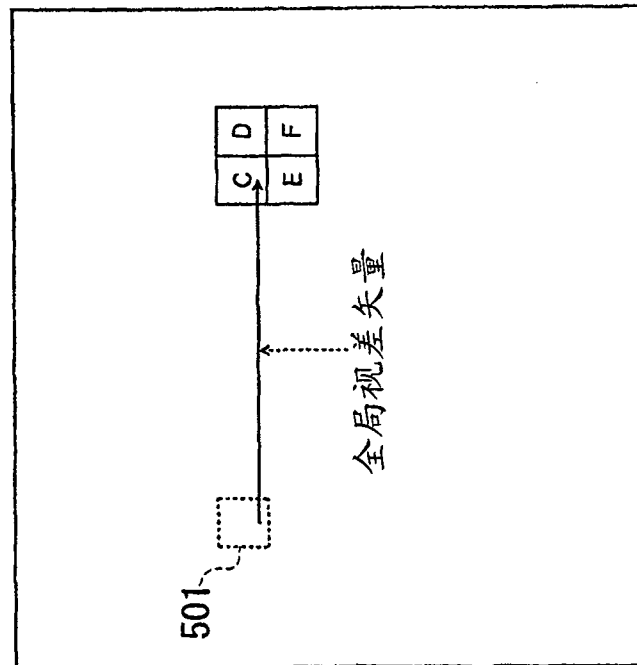


图 7



摄像机C2的图像



摄像机C1的图像

图 8B

图 8A

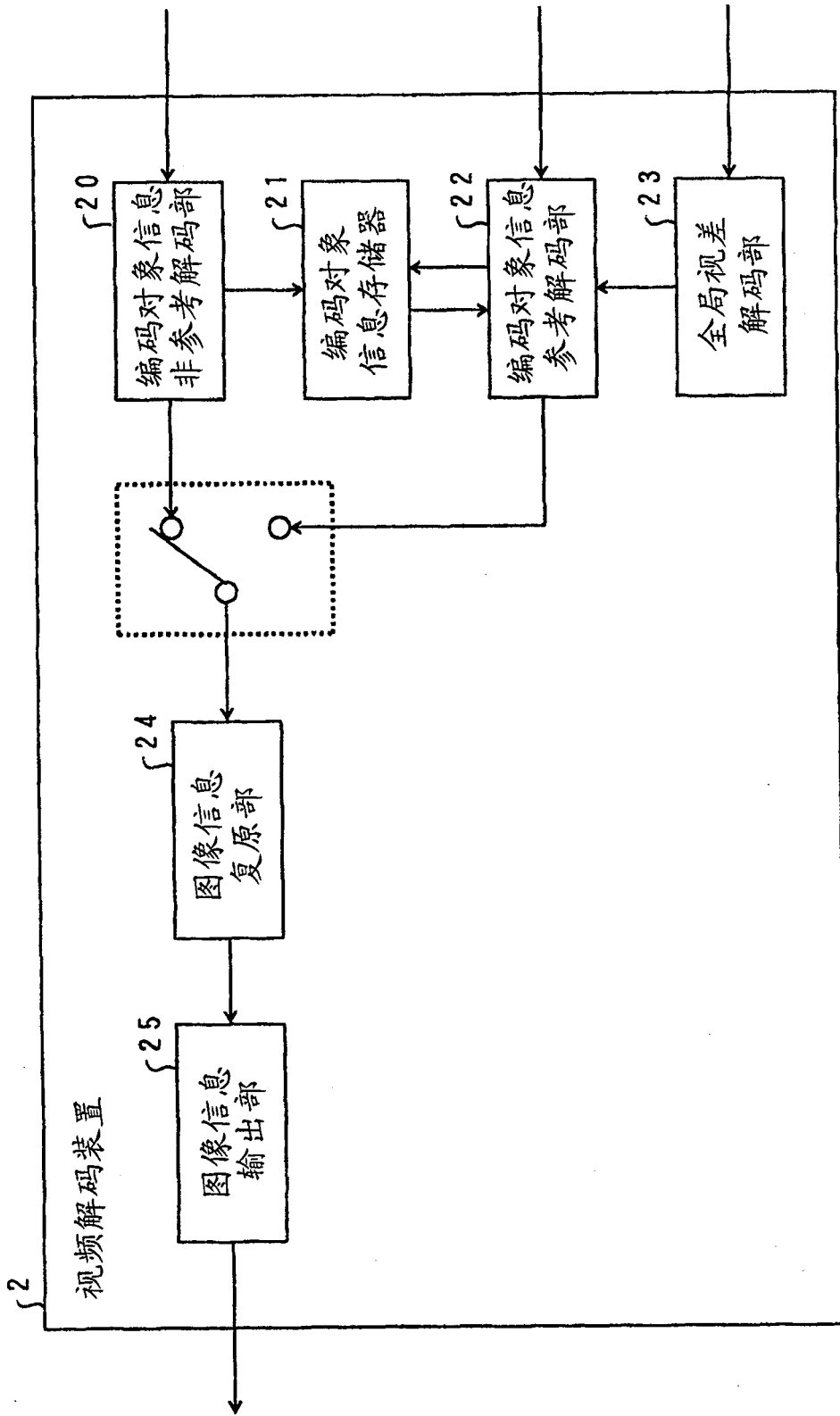


图 9

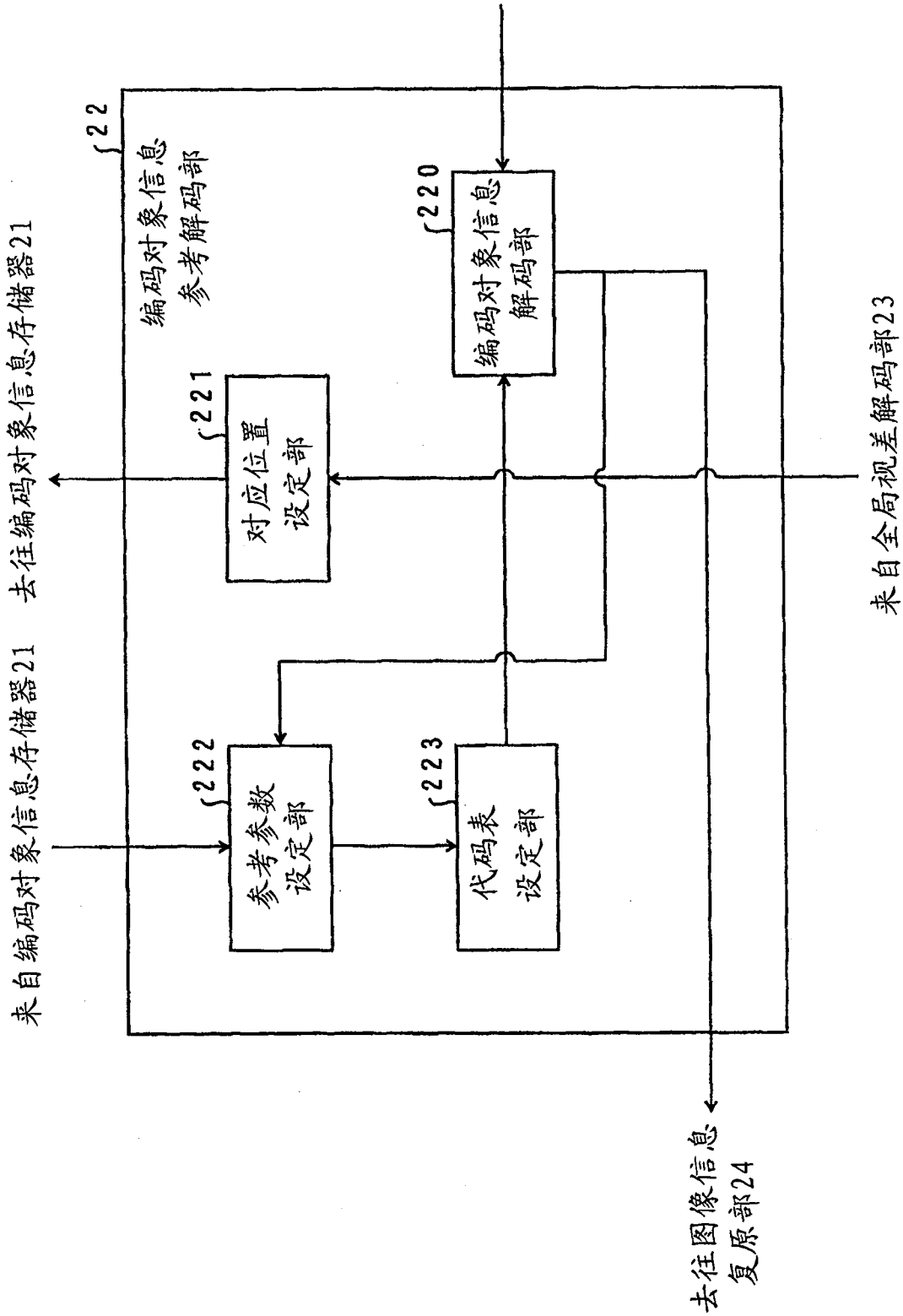


图 10

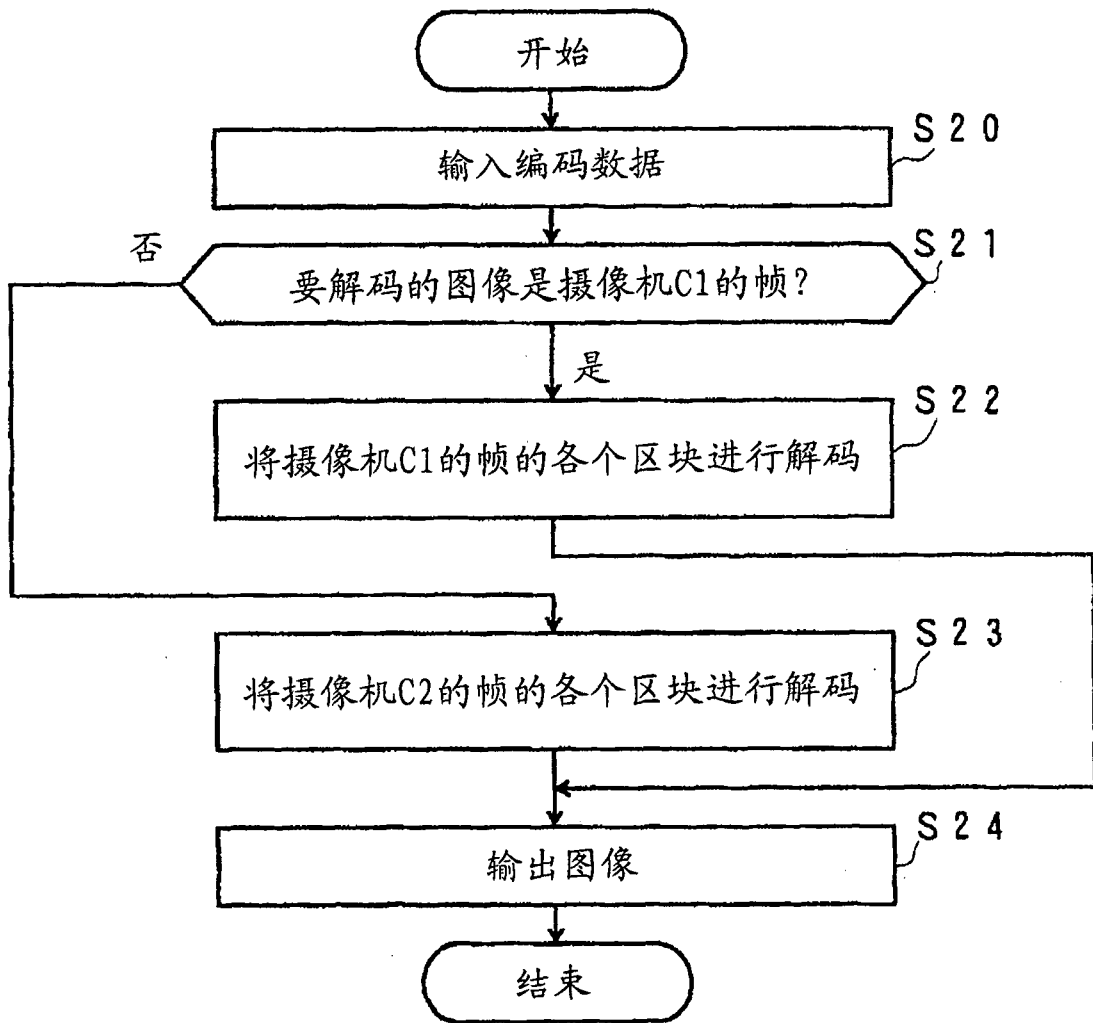


图 11

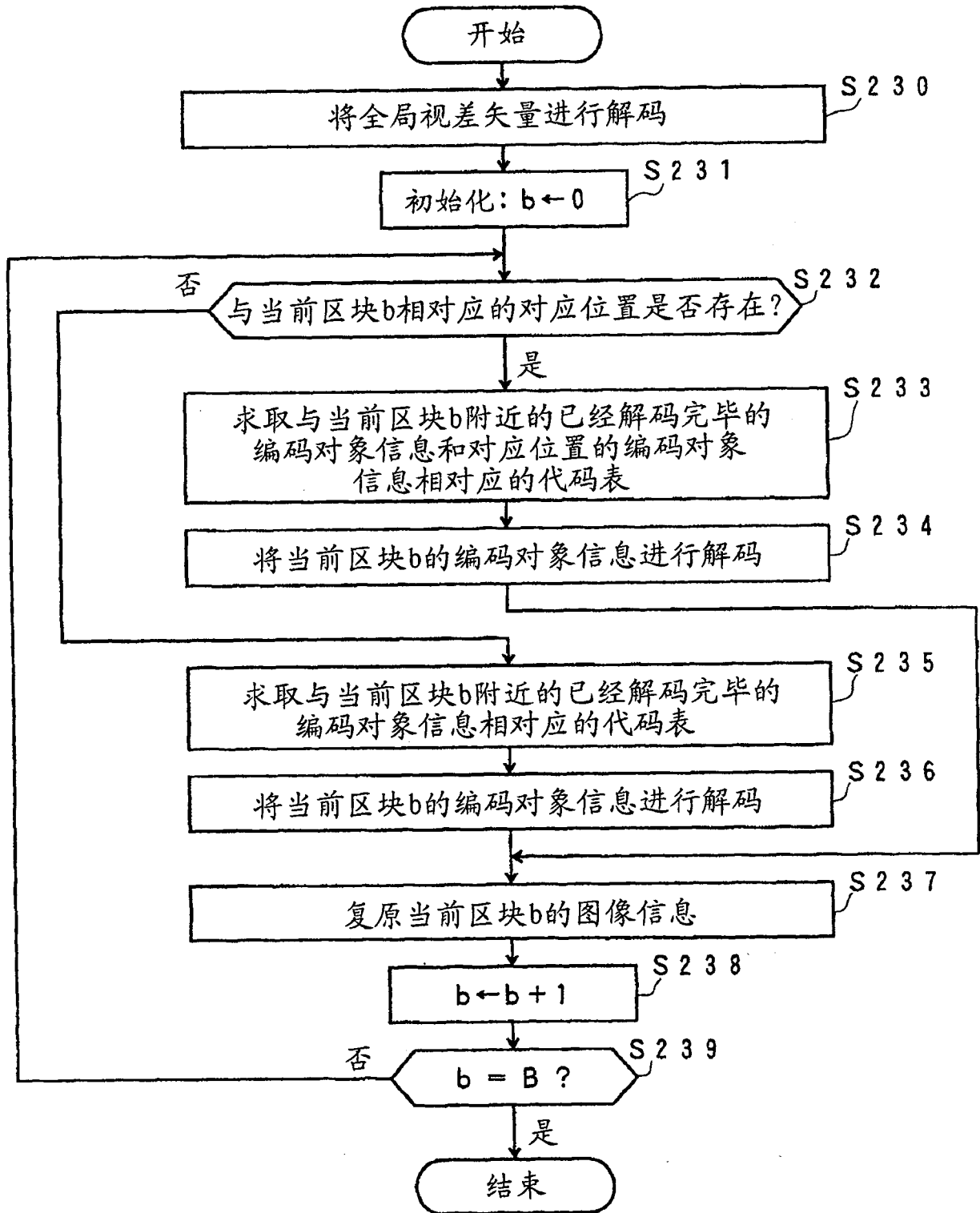


图 12

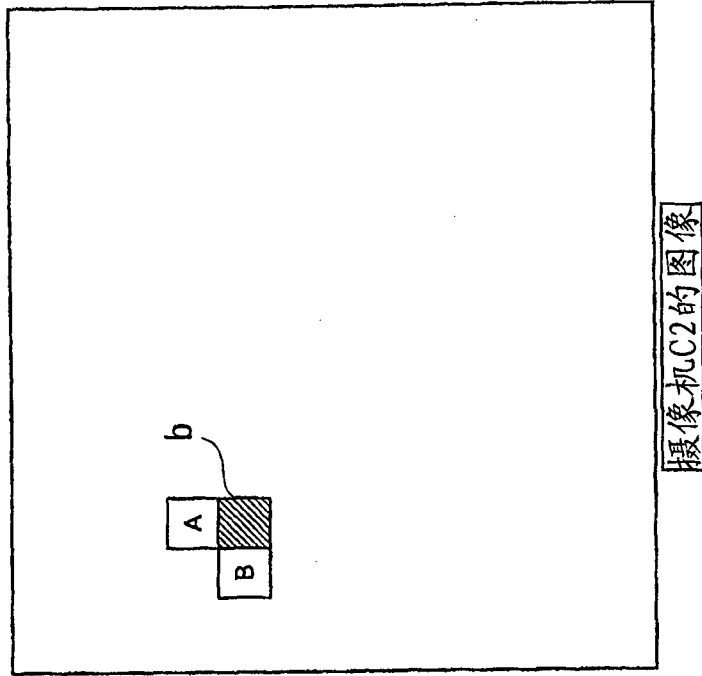


图 13B

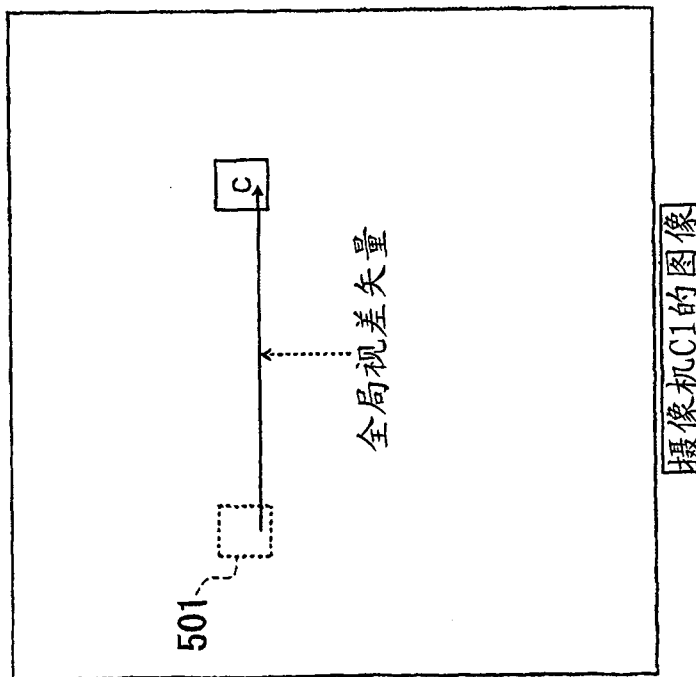


图 13A

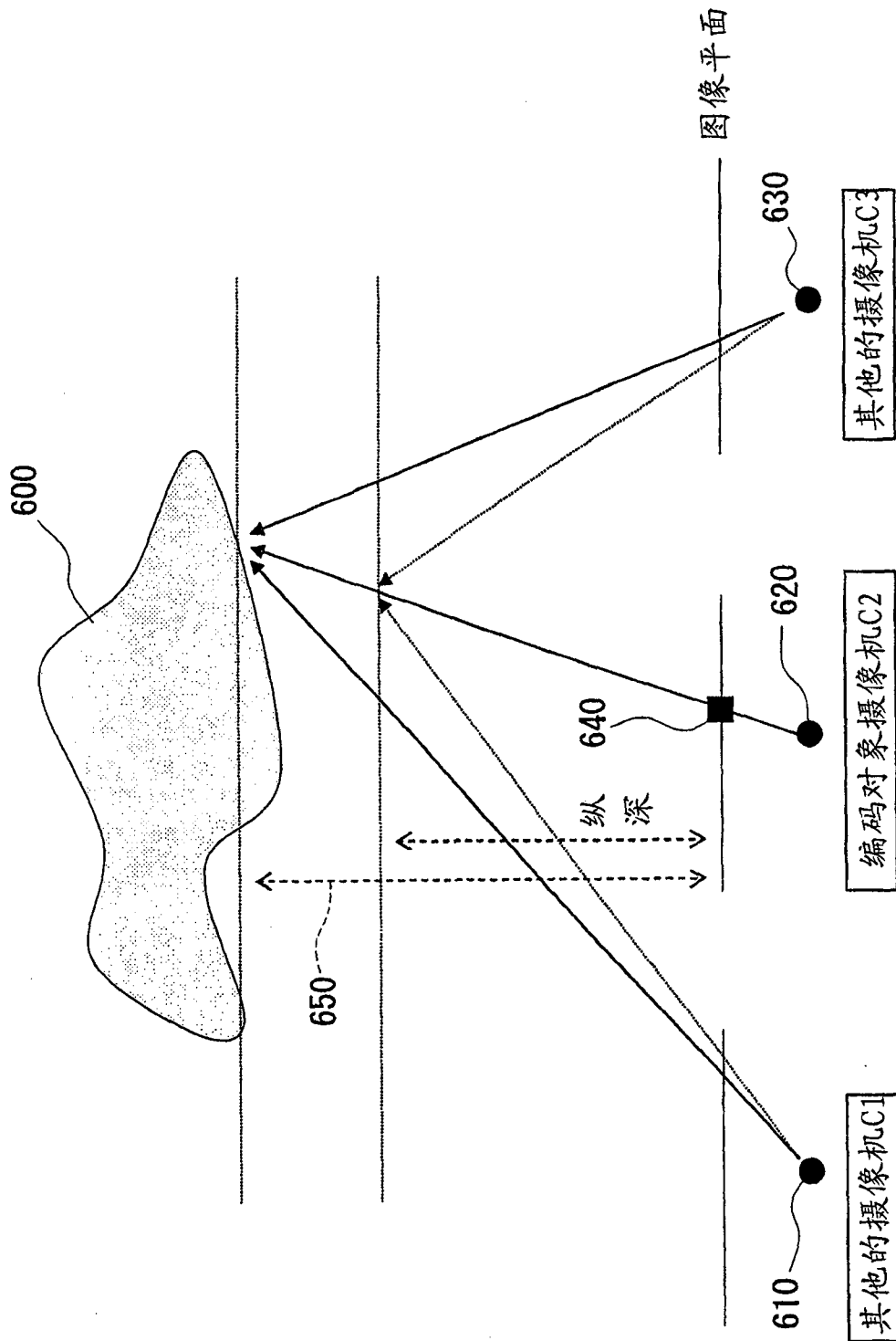


图 14

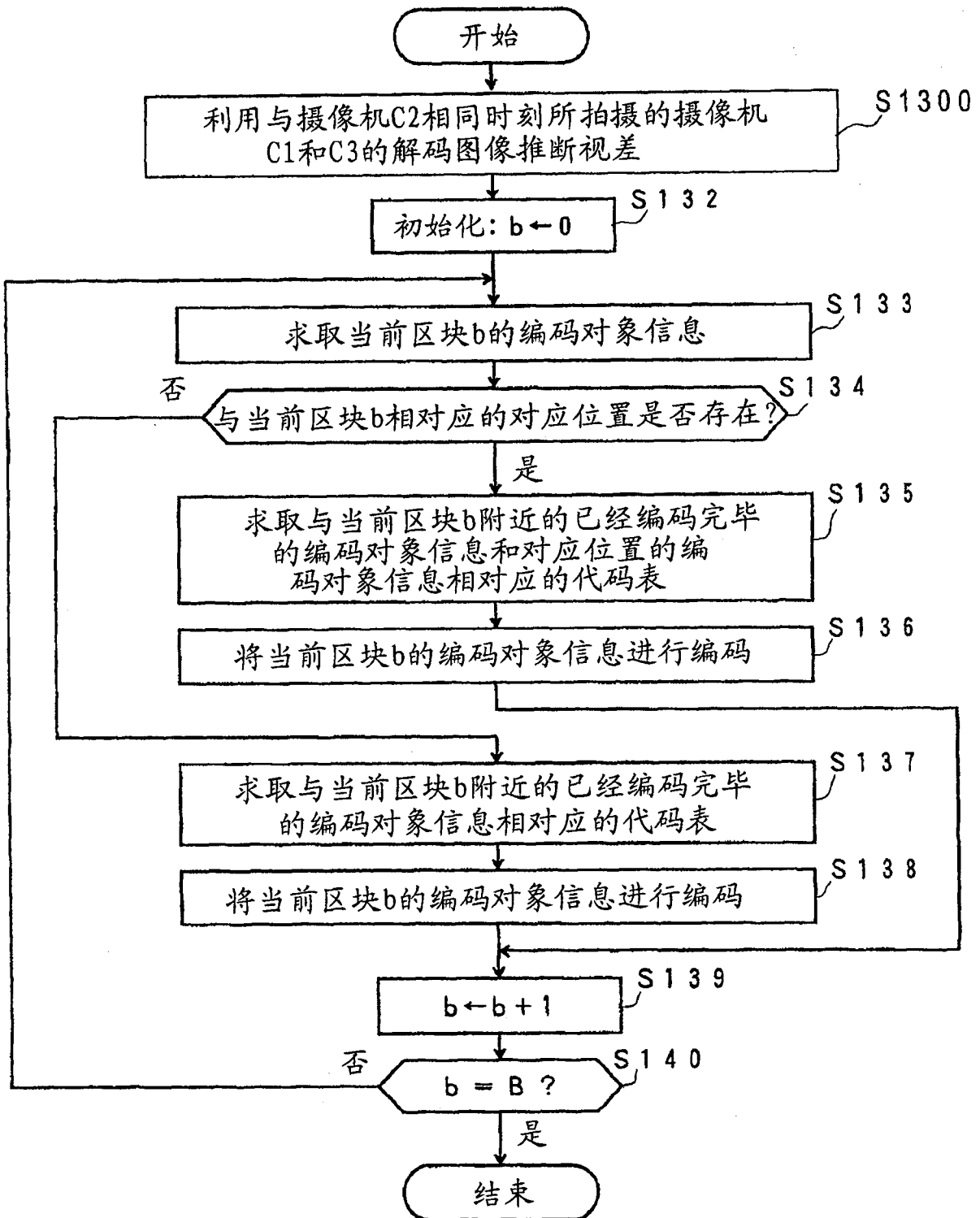


图 15

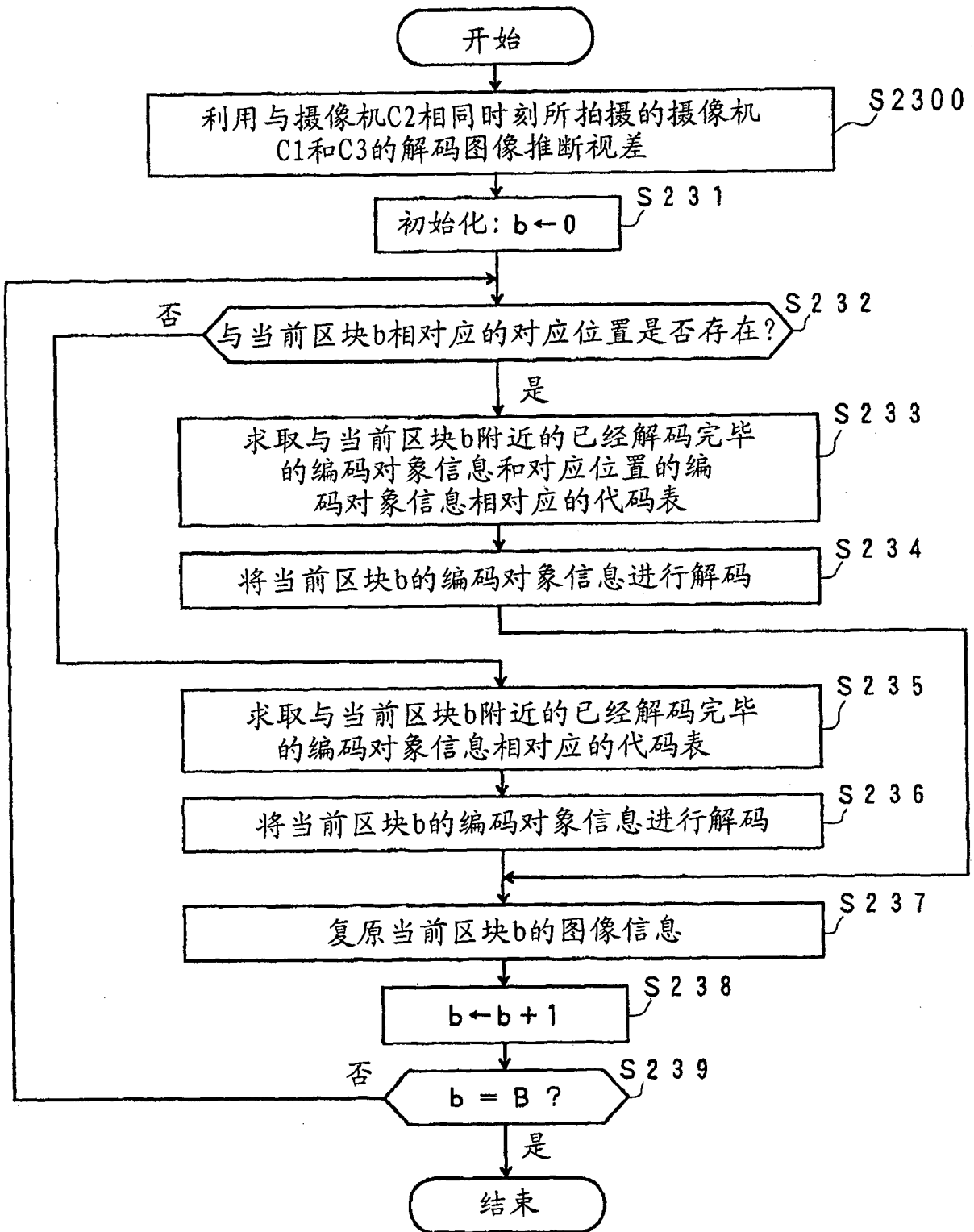


图 16

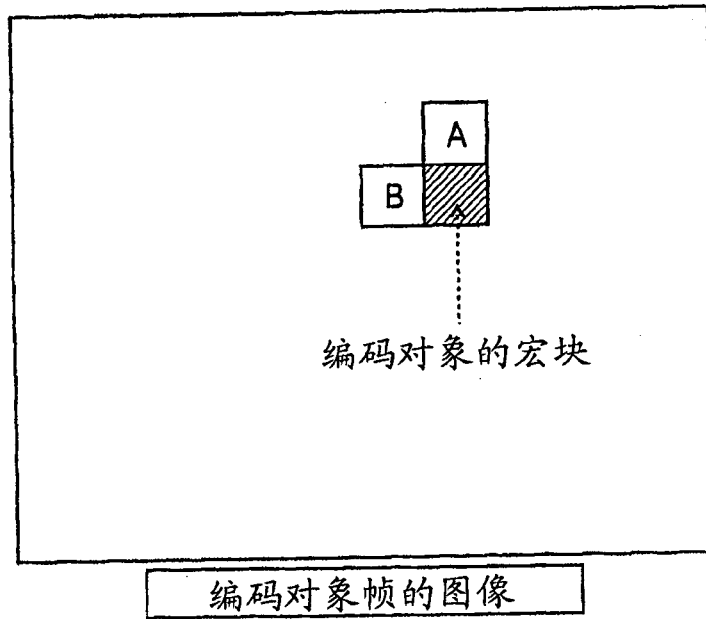


图 17

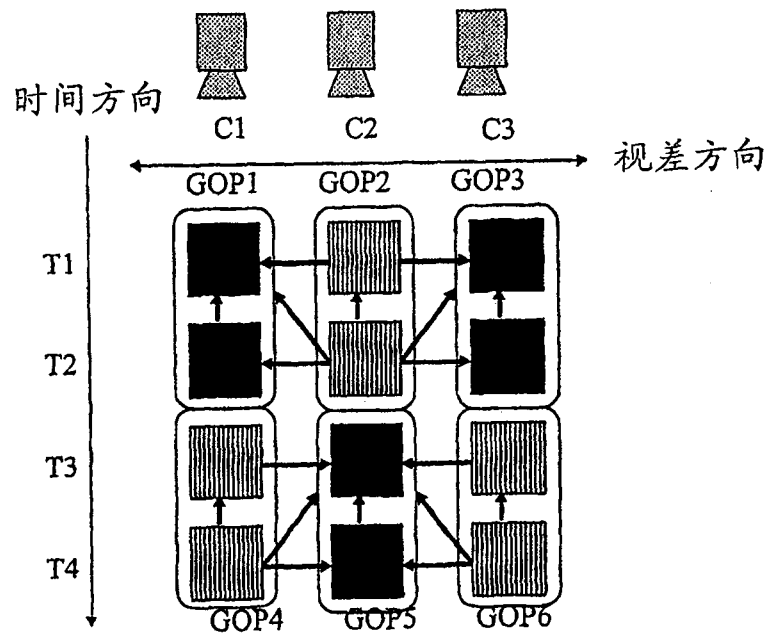


图 18