



[12] 发明专利说明书

专利号 ZL 200410036925.8

[45] 授权公告日 2007年1月10日

[11] 授权公告号 CN 1294764C

[22] 申请日 1998.6.5

[21] 申请号 200410036925.8

分案原申请号 98109558.5

[30] 优先权

[32] 1997.6.5 [33] US [31] 08/869,493

[73] 专利权人 通用仪器公司

地址 美国宾夕法尼亚

[72] 发明人 陈学敏 阿贾伊·卢特拉

加内什·拉詹 曼德阿姆·纳拉辛汉

[56] 参考文献

JP7-170514A 1995.7.4 H04N7/24

US5049991A 1991.9.17 H04N7/137

US5467136A 1995.11.14 H04N7/36

CN1134583A 1996.10.30 G11B7/00

审查员 陈荣华

[74] 专利代理机构 永新专利商标代理有限公司

代理人 刘炳胜

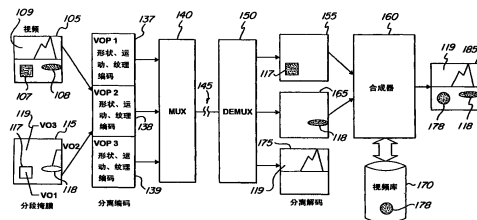
权利要求书 5 页 说明书 21 页 附图 8 页

[54] 发明名称

用于编码双向预测视频对象面的方法及其解码装置

[57] 摘要

一种用于对双向预测的视频对象面 (B-VOP) 编码的方法, 包括: 缩放包括多个视频对象面 (VOP) 的输入视频序列, 以在相应基层和增强层中传送; 在所述基层中提供对应于所述输入视频序列 VOP 的第一和第二基层 VOP; 根据运动向量 MV_p , 由所述第一基层 VOP 预测所述第二基层 VOP; 在增强层中处于所述第一和第二基层 VOP 中间的时间位置提供所述 B-VOP; 和利用通过缩放所述运动向量 MV_p 而得到的下列之一来编码所述 B-VOP: (a) 向前运动向量 MV_f , 和 (b) 向后运动向量 MV_b 。



1. 一种用于对双向预测的视频对象面编码的方法，包括下列各步：

缩放包括多个视频对象面的输入视频序列，以在相应基层和增强层中传送；

在所述基层中提供对应于所述输入视频序列视频对象面的第一和第二基层视频对象面；

根据运动向量 MV_p ，由所述第一基层视频对象面预测所述第二基层视频对象面；

在增强层中处于所述第一和第二基层视频对象面中间的时间位置提供所述双向预测的视频对象面；和

利用通过缩放所述运动向量 MV_p 而得到的下列之一来编码所述双向预测的视频对象面：

(a) 向前运动向量 MV_f ，和

(b) 向后运动向量 MV_b 。

2. 权利要求 1 的方法，其中：

所述第一和第二基层视频对象面之间相隔时间间隔 TR_p ；

所述第一基层视频对象面和所述双向预测的视频对象面之间相隔时间间隔 TR_b ；

至少下列之一：

(a) 所述向前运动向量 MV_f 根据关系式 $MV_f = (m/n) \bullet TR_b \bullet MV_p / TR_p$ 确定；且

(b) 所述向后运动向量 MV_b 根据关系式 $MV_b = (m/n) \bullet (TR_b - TR_p) \bullet MV_p / TR_p$ 确定；

其中， m/n 是第一和第二基层视频对象面的空间分辨率与双向预测的视

频对象面空间分辨率的比值。

3. 权利要求 1 或 2 的方法，还包括步骤：

利用至少下述之一编码所述双向预测的视频对象面：

(a) 所述第一基层视频对象面的搜索区，其中心根据所述向前运动向量 MV_f 确定；和

(b) 所述第二基层视频对象面的搜索区，其中心根据所述向后运动向量 MV_b 确定。

4. 一种用于恢复数据流中包括有多个视频对象面的输入视频序列的方法，所述输入视频序列在相应基层和增强层中经过缩放并传送，且其中视频对象面具有相关的空间和时间分辨率，其中：

在所述基层中提供对应于所述输入视频序列视频对象面的第一和第二基层视频对象面；

根据运动向量 MV_p ，由所述第一基层视频对象面预测所述第二基层视频对象面；

在增强层中处于所述第一和第二基层视频对象面中间的时间位置处提供双向预测的视频对象面；且

采用通过对所述运动向量 MV_p 进行缩放而得到的向前运动向量 MV_f 和向后运动向量 MV_b 编码所述双向预测的视频对象面；

所述方法包括下列步骤：

从所述数据流中恢复所述向前运动向量 MV_f 和向后运动向量 MV_b ；且

利用所述向前运动向量 MV_f 和所述向后运动向量 MV_b 解码所述双向预测的视频对象面。

5. 权利要求 4 的方法, 其中:

所述第一和第二基层视频对象面之间相隔时间间隔 TR_p ;

所述第一基层视频对象面和所述双向预测的视频对象面之间相隔时间间隔 TR_B ;

至少为下列之一:

(a) 所述向前运动向量 MV_f 根据关系式 $MV_f = (m/n) \bullet TR_B \bullet MV_p / TR_p$ 确定; 且

(b) 所述向后运动向量 MV_b 根据关系式 $MV_b = (m/n) \bullet (TR_B - TR_p) \bullet MV_p / TR_p$ 确定;

其中, m/n 是第一和第二基层视频对象面的空间分辨率与双向预测的视频对象面空间分辨率的比值。

6. 权利要求 4 或 5 的方法, 其中:

利用至少下述之一编码所述双向预测的视频对象面:

(a) 所述第一基层视频对象面的搜索区, 其中心根据所述向前运动向量 MV_f 确定; 和

(b) 所述第二基层视频对象面的搜索区, 其中心根据所述向后运动向量 MV_b 确定。

7. 一种用于恢复数据流中包括多个视频对象面的输入视频序列的解码装置, 所述输入视频序列在相应基层和增强层中经过缩放并传送, 且其中的视频对象面具有相关的空间和时间分辨率, 其中:

在所述基层中提供对应于所述输入视频序列视频对象面的第一和第二基层视频对象面;

根据运动向量 MV_p , 所述第二基层视频对象面由所述第一基层视频对象

面预测;

在增强层中处于所述第一和第二基层视频对象面中间的时间位置处提供双向预测视频对象面; 且

采用通过对所述运动向量 MV_p 进行缩放而得到的向前运动向量 MV_f 和向后运动向量 MV_b 编码所述双向预测的视频对象面;

所述装置包括:

恢复构件, 用于由所述数据流中恢复向前运动向量 MV_f 和向后运动向量 MV_b ; 和

解码构件, 利用所述向前运动向量 MV_f 和所述向后运动向量 MV_b 解码所述双向预测的视频对象面的装置。

8. 权利要求7的装置, 其中:

所述第一和第二基层视频对象面之间相隔时间间隔 TR_p ;

所述第一基层视频对象面和所述双向预测的视频对象面之间相隔时间间隔 TR_b ;

至少下列之一:

(a) 所述向前运动向量 MV_f 根据关系式 $MV_f = (m/n) \cdot TR_b \cdot MV_p / TR_p$ 确定; 且

(b) 所述向后运动向量 MV_b 根据关系式 $MV_b = (m/n) \cdot (TR_b - TR_p) \cdot MV_p / TR_p$ 确定;

其中, m/n 是第一和第二基层视频对象面的空间分辨率与双向预测的视频对象面的空间分辨率的比值。

9. 权利要求7或8的装置, 其中:

利用至少下述之一编码所述双向预测的视频对象面:

(a) 所述第一基层视频对象面的搜索区, 其中心根据所述向前运动向量 MV_f 确定; 和

(b) 所述第二基层视频对象面的搜索区, 其中心根据所述向后运动向量 MV_b 确定。

用于编码双向预测视频对象面的方法及其解码装置

技术领域

本发明涉及为数字视频序列中包括视频对象面的视频图像提供时间和空间缩放的方法和装置。特别是提出一种适用于可缩放帧模式或场模式视频的运动补偿方案。也提出一种利用空间变换如离散余弦变换(DCT)对场模式视频进行自适应压缩的方案。

背景技术

本发明尤其适用于各种多媒体应用,且与此处作为参考引用的文献ISO/IEC/JTC1/SC29/WG11 N1642中标题为" MPEG-4 视频校验模式 7.0版",1997年4月,所描述的MPEG-4校验模式(VM)7.0标准兼容。本发明还可提供立体视频编码、画中画、预览频道和异步传输模式(ATM)传送。

MPEG-4是一种提供灵活框架及用于对数字音频视频数据进行传送、访问和操作的开放式编码工具集合的新编码标准。这些工具支持广泛的特征。MPEG-4灵活的框架支持编码工具的各种组合及其由计算机、远程传送和娱乐(如,电视和电影)业所要求应用的相应功能,如数据库浏览、信息检索和交互式传送。

MPEG-4提供标准核心技术,允许在多媒体环境中高效存贮、发送和操纵视频数据。MPEG-4取得高效的压缩、目标缩放、空间和时间缩放及误差回弹。

MPEG-4视频VM编码器/解码器(编码解码器)是带有运动补偿的基于块和目标的混合编码器。通过采用重叠块运动补偿用 9×9 DCT对纹理进行编码。目标形状被表示为alpha图并利用基于内容的算术编码(CAE)算法或改进的DCT编码器进行编码,二者均利用时间预测。该编

码器可处理由计算机图形得到的子图形。其它编码方法，如小波和子图形编码，也可用于特殊应用。

运动补偿纹理编码是众所周知的视频编码方法。该方法可构造成三级过程模型。第一级是包括运动估计和补偿 (ME / MC) 和二维空间变换的信号处理。ME / MC 和空间变换的目标是利用视频序列中时间和空间相关性优化复杂限制情况下量化和熵编码的比率失真性能。最常见的 ME / MC 技术是块匹配，最常见的空间变换是 DCT。然而，对于任意形状 VOP 边界块的 ME / MC 和 DCT 编码产生特殊的关注。

MPEG-2 主要轮廓是 MPEG-4 标准的前身，且在文献 ISO / IEC JTC1 / SC29 / WG11 N0702 中，标题为 " 信息技术 - 运动图像和相关音频通用编码，建议 H. 262 "，1994 年 3 月 25 日，中描述，此处作为参考文献引用。对 MPEG-2 主要轮廓进行可缩放延伸已被定义，其为两个或更多独立位流或层提供。每一层可被结合以形成单个高质量信号。例如，基层可提供低质量视频信号，而增强层提供可增强基层图像的附加信息。

尤其是，时间和空间可缩放可提供不同视频标准或解码器功能间的兼容性。通过空间可缩放，基层视频可具有比输入视频序列低的空间分辨率，在该情况下，增强层载有将基层分辨率恢复到输入序列的分辨率水平的信息。例如，对应国际电信联盟 - 无线电部 (ITU - R) 601 标准 (分辨率 720×576 像素) 的输入视频序列可载于对应公共交换格式 (CIF) 标准 (分辨率 360×299 像素) 的基层。在该情况下，增强层载有被解码器用来将基层视频恢复到 ITU - R601 标准的信息。或者，增强层可能具有减小的空间分辨率。

利用时间可缩放，基层可具有比输入视频序列低的时间分辨率 (即，帧频)，而增强层载有丢失帧。当在解码器结合时，原有的帧频可被恢复。

因此，期望提供用于视频信号 (包括视频对象面 (VOP)，如用于 MPEG-4 标准中的) 编码的时间和空间可缩放功能。期望具备立体视频编码、画中画、预览频道和异步传输模式 (ATM) 传送功能。

还期望具有较低复杂性和低成本的编码解码器设计，其中，搜索范围

被减小，用于对双向预测 VOP (B - VOP) 进行增强层预测编码的运动估计。还期望在确定残余并对数据进行空间转换之前通过自适应重排序增强层 VOP 的像素线对交错视频输入信号进行高效编码。本发明提供一具备上述和其它优点的系统。

发明内容

根据本发明，提出一种为数字视频序列中的视频图像如视频对象面提供时间和空间的缩放的方法和装置。VOP 可包括整个帧和 / 或帧的子集，且可为任意形状。此外，多个 VOP 可在一帧中提供或瞬时重合。

按照本发明的一个方面，提出一种用于对包含视频对象面 (VOP) 的输入视频序列进行缩放的方法，用以在相应的基层和增强层中进行传送，其中，向下取样数据被载于基层。输入视频序列中的 VOP 具有相关的空间分辨率和时间分辨率 (例如帧频)；

输入视频序列的第一特定 VOP 的像素数据被向下取样，以提供具有减小的空间分辨率的第一基层 VOP；第一基层 VOP 的至少一部分的像素数据被向上取样，以在增强层提供第一向上取样 VOP；第一向上取样 VOP 利用输入视频序列的第一特定 VOP 被进行差分编码，且被提供于增强层中对应于第一基层 VOP 的时间位置；

差分编码包括根据第一向上取样 VOP 的像素数据与输入视频序列的第一特定 VOP 的像素数据间的差确定残余的步骤，残余利用例如 DCT 被空间转换以提供转换因子；其中

输入视频序列的 VOP 是场模式 VOP 时，若像素数据行满足重排序准则，差分编码包括在确定残余之前，将第一向上取样 VOP 的像素数据行重排序为场模式。该准则是，是否对立场行 (例如，奇对偶，偶对奇) 的亮度值之差的和值大于相同场行 (奇对奇，偶对偶) 的亮度数据之差与偏差项的和值。

第一基层 VOP 的向上取样像素数据可为整个第一基层 VOP 的子集，以致于未向上取样的第一基层 VOP 的余下部分比已向上取样的像素数据具有

较低的空间分辨率。

第二基层 VOP 和增强层中的向上取样 VOP 可以类似方式提供。第一基层 VOP 和第二基层 VOP 之一或两者可被用于预测对应于第一和第二向上取样 VOP 的中间 VOP。中间 VOP 被编码, 用于在增强层中进行第一和第二向上取样 VOP 间的传送。

而且, 当在第一和第二基层 VOP 间无中间基层 VOP 时, 增强层可能比基层具有较高的时间分辨率。

在特定应用中, 基层和增强层提供画中画 (PIP) 功能 (其中, PIP 图像被载于基层), 或预览频道功能 (其中, 预览图像被载于基层)。在此应用中, 可以接受 PIP 图像或自由预览图像具有减少的空间和/或时间分辨率。在 ATM 应用中, 较高优先级、低比特率数据可在基层中提供, 而较低优先级、高比特率数据可在增强层中提供。在此情况下, 基层被分配一保用带宽, 但增强层可能偶然丢失。

一种用于对包括视频对象面的输入视频序列进行可缩放的方法 (其中, 向下取样数据被载于增强层, 而不是基层) 被提出。通过此方法, 输入视频序列的第一特定 VOP 被作为第一基层 VOP 提供于基层, 例如不改变空间分辨率。至少一部分第一基层 VOP 的像素数据被向下取样, 以在增强层中对应于第一基层 VOP 的时间位置提供相应的第一向下取样 VOP。第一特定 VOP 的相应像素数据被向下取样, 以提供一比较 VOP, 且第一向下取样 VOP 利用比较 VOP 被差分编码。

基层和增强层可提供立体视频功能, 其中, 增强层中的图像数据比基层中的图像数据具有较低的空间分辨率。

本发明提供一种用于对双向预测的视频对象面编码的方法, 包括下列步骤: 缩放包括多个视频对象面的输入视频序列, 以在相应基层和增强层中传送; 在所述基层中提供对应于所述输入视频序列视频对象面的第一和第二基层视频对象面; 根据运动向量 MV_p , 由所述第一基层视频对象面预测所述第二基层视频对象面; 在增强层中处于所述第一和第二基层视频对象面中间的时间位置提供所述双向预测的视频对象面; 和利用通过缩放所述

运动向量 MV_p 而得到的向前运动向量 MV_r 或向后运动向量 MV_b 来编码所述双向预测的视频对象面。

对应输入视频序列 VOP 的第一和第二基层 VOP 被提供于基层中。第二基层 VOP 是 P - VOP (其可根据运动向量 MV_p 从第一基层 VOP 预测)。B - VOP 被提供于增强层中第一和第二基层 VOP 之间。B - VOP 利用通过对运动向量 MV_p 按缩放所得的向前运动向量 MV_r 和向后运动向量 MV_b 中的至少一个编码。此高效编码技术避免需在参考 VOP 中进行独立的耗时的搜索。时间间隔 TR_p 分离第一和第二基层 VOP, 而时间间隔 TR_b 分离第一基层 VOP 和 B - VOP。

比值 m/n 被定义为第一和第二基层 VOP 的空间分辨率与 B - VOP 的空间分辨率的比值。即, 基层 VOP 或增强层中的 B - VOP 可相对输入视频序列的 VOP 以比值 m/n 被向下取样。假定基层或增强层 VOP 与输入视频序列具有相同的空间分辨率。向前运动向量 MV_r 根据关系式 $MV_r = (m/n) \cdot TR_b \cdot MV_p / TR_p$ 确定, 向后运动向量 MV_b 根据关系式 $MV_b = (m/n) \cdot (TR_b - TR_p) \cdot MV_p / TR_p$ 确定。 m/n 是任一正数, 包括分数值。

B - VOP 利用第一基层 VOP 的搜索区 (其中心根据向前运动向量 MV_r 确定) 和第二基层 VOP 的搜索区 (其中心根据向后运动向量 MV_b 确定) 编码。

相应的解码器方法和装置也被提出。

本发明还提出一种用于恢复数据流中包括有多个视频对象面的输入视频序列的方法, 所述输入视频序列在相应基层和增强层中经过缩放并传送, 且其中视频对象面具有相关的空间和时间分辨率, 其中: 在所述基层中提供对应于所述输入视频序列视频对象面的第一和第二基层视频对象面; 根据运动向量 MV_p , 由所述第一基层视频对象面预测所述第二基层视频对象面; 在增强层中处于所述第一和第二基层视频对象面中间的时间位置处提供双向预测的视频对象面; 且采用通过对所述运动向量 MV_p 进行缩放而得到的向前运动向量 MV_r 和向后运动向量 MV_b 编码所述双向预测的视频对象面;

所述方法包括下列步骤:

从所述数据流中恢复所述向前运动向量 MV_f 和向后运动向量 MV_b ; 且利用所述向前运动向量 MV_f 和所述向后运动向量 MV_b 解码所述双向预测的视频对象面。

本发明还提出一种用于恢复数据流中包括多个视频对象面的输入视频序列的解码装置, 所述输入视频序列在相应基层和增强层中经过缩放并传送, 且其中的视频对象面具有相关的空间和时间分辨率, 其中: 在所述基层中提供对应于所述输入视频序列视频对象面的第一和第二基层视频对象面; 根据运动向量 MV_b , 所述第二基层视频对象面由所述第一基层视频对象面预测; 在增强层中处于所述第一和第二基层视频对象面中间的时间位置处提供双向预测视频对象面; 且采用通过对所述运动向量 MV_b 进行缩放而得到的向前运动向量 MV_f 和向后运动向量 MV_b 编码所述双向预测的视频对象面; 所述装置包括:

恢复构件, 用于由所述数据流中恢复向前运动向量 MV_f 和向后运动向量 MV_b ; 和解码构件, 利用所述向前运动向量 MV_f 和所述向后运动向量 MV_b 解码所述双向预测的视频对象面的装置。

附图说明

图 1 是根据本发明的视频对象面 (VOP) 编码和解码过程图。

图 2 是根据本发明的 VOP 编码器和解码器框图。

图 3 是根据本发明的像素向上取样图。

图 4 是基层与增强层 VOP 间预测过程的示例图。

图 5 是根据本发明的 VOP 空间和时间缩放图。

图 6 描述根据本发明由帧到场模式像素行重排序。

图 7 是根据本发明的带有空间和时间缩放的画中画 (PIP) 或预览频道应用图。

图 8 是根据本发明的立体视频应用图。

具体实施方式

提出一种为数字视频序列中包括视频对象面的视频图像提供时间和空间缩放的方法和装置。

图1描述了根据本发明的对视频目标进行编码和解码过程的图。帧105包括三个图形单元，包括方形前景单元107，椭圆形前景单元108，和背景单元109。在帧115中，该单元是采用分段掩膜设计，从而VOP117代表方形前景单元107，VOP118代表椭圆形前景单元108，VOP119代表背景单元109。VOP可具有任意形状，一系列VOP被称作一视频目标。一整个矩形视频帧也可被称作一VOP。由此，术语"VOP"在此处使用以表明任意或非任意图像区域形状。可利用已知技术获得分段掩膜，且具有类似于ITU-R601亮度数据的格式。每一像素被确认为属于视频帧中某一区域。

帧105和来自帧115的VOP数据被提供给独立的编码功能。尤其是，VOP117、118和119分别在编码器137、138和139处进行形状、运动和纹理编码。通过形状编码，二进制和灰度级形状信息被编码。通过运动编码，利用一帧中的运动估计来编码形状信息。通过纹理编码，进行空间变换如DCT以获取可被变长编码而用于压缩的变换系数。

编码的VOP数据此后在多路器(MUX)140结合以通过信道145传输。或者，该数据可被存贮于记录媒质。所接收到的编码VOP数据通过解多路复用器150分离，以便被分离的VOP117-119被解码并恢复。帧155、165和175分别示出VOP117、118和119已被解码和恢复，且由此可利用合成器160(其例如与视频库170接口)各自独立处理。

合成器可为一例如可安放于用户家中的个人计算机以允许编辑所接收的数据从而提供特制的图像。例如，用户的个人视频库170可包括先前存贮的与所接收VOP不同的VOP178(例如，一个圆)。用户可制作其中圆形VOP178替换方形VOP117的帧185。帧185由此包括所接收的VOP118和119，和当地存贮的VOP179。

在另一例中，背景VOP109可由用户所选择的背景取代。例如，当观看电视新闻时，播音员可被作为与背景，如新闻演播室分离的VOP编码。用

户可由库 170 或其它电视节目, 如带有股票价格或天气信息的频道, 选择背景。用户由此可作为视频编辑者。

视频库 170 也可存贮通过信道 145 所接收的 VOP, 且可通过网络如 Internet 访问 VOP 或其它图像单元。

可以理解, 帧 105 可能包括不是 VOP 的区域, 由此不能被独立操作。而且, 帧 105 无需具有 VOP。通常, 一视频段可包含单个 VOP, 或一个 VOP 序列。

图 1 中的视频目标编码和解码过程可带来许多娱乐、商业和教育应用, 包括个人计算机游戏、虚拟环境、图形用户界面、可视会议、Internet 应用及其它。尤其是, 根据本发明的 VOP 时间和空间缩放的能力可提供更强大的功能。

图 2 是根据本发明的视频目标编码器和解码器的框图。编码器 201, 其对应图 1 所示单元 137-139, 包括接收输入视频数据序列 " i_n " 的可缩放预处理器 205。为对比增强层具有较低空间分辨率的基层取得空间缩放, " i_n " 被空间向下取样, 以获得信号 " i_{n-0} ", 它又通过路径 217 被提供给基层编码器 220。" i_{n-0} " 在基层编码器 220 处编码, 且被编码的数据被提供给多路器 (MUX) 230。MPEG-4 系统和描述语言 (MSDL) MUX 可被使用。

应注意, 当输入视频序列 " i_n " 是场 (交错) 模式时, 由于向下取样不保留奇和偶场的像素数据, 向下取样信号 " i_{n-0} " 将处于帧 (顺序) 模式。当然, " i_n " 为帧模式时, " i_{n-0} " 也将处于帧模式。

经路径 218 将重建的图像数据从基层编码器 220 提供给完成像素向上取样的中处理器 215, 联系图 3 下文将详细描述。处于帧模式的向上取样图像数据此后经路径 212 提供给增强层编码器 210, 在其中利用从预处理器 205 经路径 207 提供给编码器 210 的输入图像数据进行差分编码。具体说, 该向上取样像素数据 (例如, 亮度数据) 从输入图像数据中减去以获得残余, 此后利用 DCT 或其它空间变换可对该残余编码。

根据本发明, 当输入视频序列为场模式时, 通过将对应输入视频序列

的初始偶（顶部）和奇（底部）场的向上取样增强层图像的像素行分组，编码效率可被提高。这样在某些情形下可减小残余幅度，因为一场内的像素数据与相同场内其它像素数据常具有比对立场数据更大的相关性。因此，通过减小残余的幅度，更少的比特被要求用于对图像数据编码。参考图 6 和相关讨论，下面进行更详尽的描述。

增强层中向上取样图像的编码残余被提供给 MUX 230，用以通过信道 245 与基层数据一起传输。或者数据可本地存贮。注意，MUX 230、信道 245 和 DEMUX 250 分别对应图 1 中的单元 140、145 和 150。

注意，从基层编码器 220 提供给中处理器 215 的图像数据可为整个视频图像，如全帧 VOP，或为整个图像一子集的 VOP。而且，多个 VOP 可提供给中处理器 215。MPEG-4 目前支持到 256VOP。

在解码器 299，编码数据在解多路复用器 (DEMUX) 250，如一 MPEG-4 MSDL DEMUX 被接收。在本例中比基层数据具有更高的空间分辨率的增强层数据被提供给增强层解码器 260。基层数据被提供给基层解码器 270，其中信号 "out_0" 被恢复并经路径 267 提供给中处理器 265，经路径 277 提供给缩放后处理器 280。该中处理器通过对基层数据向上取样来恢复全分辨率图像，以与编码器侧的中处理器 215 类似的方式工作。该图像经路径 262 被提供给增强层解码器 260 用于恢复增强层数据信号 "out_1"，其此后经路径 272 被提供到缩放后处理器 280。缩放后处理器 280 完成例如对解码基层数据进行空间向上取样以作为信号 "output_0" 显示，增强层数据被输出以作为信号 "output_1" 显示的操作。

当编码器 201 被用于时间缩放时，预处理器 205 对于例如基层完成时间分路（例如，拉开过程或帧抛弃）以减小帧频。例如，为将帧频由 30 帧/秒减小到 15 帧/秒，每隔一帧被抛弃。

下面的表 1 示出中处理器 215 和 265、缩放预处理器 205 和缩放后处理器 280 的 24 个可能的配置。

表 1

配置	层	时间分辨率	空间分辨率	可缩放预处理器	中处理器	可缩放预处理器
1	基层	低(高)	低	向下取样滤波	向上取样滤波	N/C
	增强层	低(高)	高	N/C	N/A	N/C
2	基层	低	低	向下取样滤波和拉开处理	向上取样滤波	N/C
	增强层	高	高	N/C	N/A	N/C
3	基层	高	低	向下取样滤波	向上取样滤波	N/C
	增强层	低	高	拉开处理	N/A	N/C
4	基层	低(高)	高	N/C	向下取样滤波	N/C
	增强层	低(高)	低	向下取样滤波	N/A	向上取样滤波
5	基层	低	高	拉开处理	向下取样滤波	N/C
	增强层	高	低	向下取样滤波	N/A	向上取样滤波
6	基层	高	高	N/C	向下取样滤波	N/C
	增强层	低	低	向下取样滤波和拉开处理	N/A	向上取样滤波
7	基层	低(高)	高	N/C	N/C	N/C
	增强层	低(高)	高	N/C	N/A	N/C
8	基层	低	高	拉开处理	N/C	N/C
	增强层	高	高	N/C	N/A	N/C
9	基层	高	高	N/C	N/C	N/C
	增强层	低	高	拉开处理	N/A	N/C
10	基层	低(高)	低	向下取样滤波	N/C	向上取样滤波
	增强层	低(高)	低	向下取样滤波	N/A	向上取样滤波
11	基层	低	低	向下取样滤波和拉开处理	N/C	向上取样滤波
	增强层	高	低	向下取样滤波	N/A	向上取样滤波
12	基层	高	低	向下取样滤波	N/C	向上取样滤波
	增强层	低	低	向下取样滤波和拉开处理	N/A	向上取样滤波

表 1 中，第一列指明配置数，第二列指明层，第三列指明层的时间分辨率（例如，高或低）。当“低（高）”被列出，基层和增强层的时间分辨率或均为高或均为低。第四列指明空间分辨率。第五、第六和第七列指明缩放预处理器 205、中处理器 215 和 265 及后处理器 280 的相应动作。“N / C”表示在时间或空间分辨率上无变化，即进行通常处理过程。“N / A”意味着“不可用”。中处理器 215、265 动作不影响增强层。

空间可缩放编码采用配置 1 作为实例描述。如上所述，当可缩放编码器 201 被用于编码 VOP 时，预处理器 205 生成两个具有不同空间分辨率的 VOP 子流。如表 1 所示，基层具有较低的空间分辨率，增强层具有对应输入序列分辨率的较高的空间分辨率。因此，基层序列“ i_{n-0} ”在可缩放预处理器 205 处由输入视频序列“ i_n ”的向下取样过程生成。增强层序列通过在可缩放中处理器 215、265 处对向下取样基层序列进行向上取样滤波生成，以取得与“ i_n ”相同的高空间分辨率。后处理器 280 完成常规过程，即，它不改变“ out_{-1} ”或“ out_{-0} ”的时间或空间分辨率。

例如，基层 CIF 分辨率序列（ 360×298 像素）可由对 ITU-R 601 分辨率输入序列（ 720×576 像素）进行 2:1 向下取样滤波生成。可使用任意整数或非整数比率向下取样。

时间和空间可缩放编码采用配置 2 作为实例描述。此处，输入视频序列“ i_n ”，其具有较高的空间和时间分辨率，被转换为具有较低的空间和时间分辨率的基层序列，和具有较高的空间和时间分辨率的增强层序列。由表 1 所示，通过在预处理器 205 进行向下取样滤波和拉开处理以提供信号“ i_{n-0} ”，在中处理器 215 和 265 中进行向上取样滤波，在后处理器 280 中进行常规处理，来获得上述转换。

通过配置 3，具有低或高时间分辨率和高空间分辨率的输入视频序列“ i_n ”被转换为具有相应的低或高的时间分辨率和高空间分辨率的基层序列，和具有相应低或高时间分辨率和低空间分辨率的增强层序列。通过在预处理器 205 进行对于增强层序列“ i_{n-1} ”的向下取样滤波，在中处理器 215 和 265 进行向下取样滤波，在后处理器 280 进行增强层序列“ ou

t - 1" 的向上取样滤波, 来获得上述转换。

其它配置可由上述实例理解。

图 3 是根据本发明的像素向上取样图。向上取样滤波可通过表 1 的配置 1 由中处理器 215 和 265 完成。例如, 具有 CIF 分辨率 (360 × 298 像素) 的 VOP 可通过 2: 1 向上取样转换为 ITU - R 601 分辨率 (720 × 576 像素)。CIF 图像的像素 310、320、330 和 340 可被取样以生成 ITU - R 601 图像的像素 355、360、365、370、375、380、395 和 390。具体说, ITU - R 601 的像素 360 通过分别如箭头 312 和 322 所示对 CIF 像素 310 和 320 取样获得。类似地, ITU - R 601 的像素 365 也通过分别如箭头 314 和 324 所示对 CIF 像素 310 和 320 取样获得。

图 4 是基层和增强层中 VOP 间预测过程的示例图。在图 2 所示增强层编码器 210 中, 增强层 VOP 可被编码为 P - VOP 或 B - VOP。在本例中, 增强层中的 VOP 具有比基层 VOP 更大的空间分辨率, 由此被画成较大区域。时间分辨率 (例如帧频) 对于两层是相同的。VOP 以出现顺序从左到右示出。

基层包括 I - VOP405, B - VOP415 和 420, 及 P - VOP430。增强层包括 P - VOP450 和 490, B - VOP 460 和 480。B - VOP415 如箭头 410 和 440 所示由其它基层 VOP 预测, B - VOP420 如箭头 425 和 435 所示也由其它基层 VOP 预测。P - VOP430 如箭头 445 所示由 I - VOP405 预测。P - VOP 450 如箭头 455 所示通过对基层 VOP 向上取样来得到, P - VOP480 如箭头 485 所示通过对基层 VOP 向上取样得出。B - VOP460 如箭头 465 和 475 所示由基层 VOP 预测, B - VOP490 如箭头 470 和 495 所示由基层 VOP 预测。

通常, 与基层 I - VOP 瞬间重合 (如在显示或出现顺序中) 的增强层 VOP 被编码为 P - VOP。例如, VOP450 与 I - VOP405 瞬时重合, 由此被编码为 P - VOP。与基层 P - VOP 瞬间重合的增强层 VOP 被编码为 P - 或 B - VOP。例如, VOP490 与 P - VOP430 瞬间重合而被编码为 P - VOP。与基层 B - VOP 瞬间重合的增强层 VOP 被编码为 B - VOP。例如, 见 B - VOP460 与 490。

I - VOP405 和 P - VOP430 已知为固定 (anchor) VOP, 因为它们被用作增强层 VOP 的预测参考。I - VOP405 和 P - VOP430 由此在增强层对应预

测 VOP 编码之前被编码。增强层 P - VOP 的预测参考由 MPEG-4 兼容语法中的向前 (预测) 时间参考指示器 `forward-temporal-ref` 指定。该指示器是指向瞬时重合的基层 I - VOP 的非负整数。增强层 B - VOP 的预测参考由 `ref-select-code`、`forward-temporal-ref` 和 `backward-temporal-ref` 指定。如下面表 2 所示。注意该表对于 MPEG-2 和 MPEG-4 VM 3.0 可缩放方案是不同的。

表 2

Ref-select-code	向前时间参考 VOP	向后前时间参考 VOP
00	基层	基层
01	基层	增强层
10	增强层	基层
11	增强层	增强层

表 2 示出了对于增强层中 B - VOP 的预测参考选择。例如, 假定基层中的 I - VOP405 和 P - VOP430 的时间参考码 `temporal-ref` 分别为 0 和 3。同时, 令增强层中 P - VOP450 的 `temporal-ref` 为 0。则图 4 中, 对于 P - VOP 450 有 `forward-temporal-ref` = 0。分别如箭头 465 和 475, 470 和 485 所指示的 B - VOP460 和 490 的预测参考由 `ref-select-code` = 00, `forward-temporal-ref` = 0 和 `backward-temporal-ref` = 3 确定。P - VOP490 的预测参考由 `ref-select-code` = 10, `forward-temporal-ref` = 0 和 `backward-temporal-ref` = 3 指定。

在对基层和增强层编码中, 预测模式可由下面表 3 所示 2 位字 `VOP-prediction-type` 指明。

表 3

VOP-prediction-type	Code
I	00
P	01
B	10

" I " 预测类型指明一内部编码 VOP, " P " 预测类型指明一预测的 VOP, " B " 预测类型指明双向预测 VOP。基层序列 " i n _ 0 " 的编码与不可缩放编码过程相同, 例如, 根据 MPEG-2 主框架或 H. 263 标准。

图 6 描述根据发明将像素行由帧到场重排序或置换。如上所述, 当输入 VOP 为场模式且被向下取样时, 所产生的 VOP 将为帧模式。由此, 当向下取样图像被在空间向上取样时, 所产生的 VOP 也将为帧模式。然而, 当向上取样 VOP 通过从向上取样 VOP 中减去输入 VOP 而被差分编码时, 若空间变换如 D C T 被随后在残余上操作时, 所产生的残余可能不会产生最优的编码效率。即, 在许多情形下, 残余值可通过对向上取样图像行置换(即重排序)以对偶数和奇数行分组而减小, 因为在相同场的像素之间比对立场像素间可能有更大的相关性。

一代表增强层中向上取样像素(例如亮度)数据的图像示于 600。例如, 假定图像 600 是一通过对 9×9 的块进行 2: 1 向上取样而得到的 16×16 的宏块。该宏块包括偶数行 602、604、606、608、610、612、614 和 616, 以及奇数行 603、605、607、609、611、613、615 和 617。偶数和奇数行分别组成顶部场和底部场。宏块 600 包括四个 9×9 的亮度块, 包括由区域 620 与行 602 至 609 相交所定义的第一块, 由区域 625 与行 602 至 609 相交所定义的第二块, 由区域 620 与行 610 至 617 相交所定义的第三块, 由区域 625 与行 610 至 617 相交所定义的第四块。

根据本发明, 在确定残余并进行 D C T 之前, 当图像 600 中的像素行被置换以形成相同场亮度块时, 650 处所示的宏块被形成。示于 645 的箭头

指明行 602 至 617 的重排序。例如，偶数行 602，其为宏块 600 的第一行，也是宏块 650 的第一行。偶数行 604 为宏块 650 的第二行。类似地，偶数行 606、608、610、612、614 和 616 被分别作为宏块 650 的第三行至第八行。因此，带有偶数行的 16×9 的亮度区 680 被作成。第一 9×9 块由区域 690 与 670 相交定义，而第二 9×9 的块由区域 690 与 675 相交定义。

类似地，奇数行被移至 16×9 的区域 685。区域 685 包括由区域 685 与 670 相交所定义的第一 9×9 的块，而第二 9×9 的块由区域 685 与 675 相交所定义。区域 685 由此包括奇数行 603、605、607、609、611、613、615 和 617。

根据宏块 600 是否如宏块 650 所示被重排序，作用于残余的 D C T 此处被称为 "场 D C T" 或 "帧 D C T" 或其它。但是，应当理解，本发明可适用于其它空间变换。当采用场 D C T 时，宏块空间域的亮度行（或亮度误差）被由帧 D C T 取向置换到顶部（偶）和底部（奇）场 D C T 配置。所产生的宏块被按常规转换、量化并变长编码。当一场 D C T 宏块被解码时，在所有亮度块由反向 D C T (I D C T) 获得之后进行反向置换。4:2:0 的色度数据不受该模式影响。

根据本发明选择场模式或帧模式 D C T 的准则如下所述。在下面情况下应选择场 D C T:

$$\sum_{i=0}^6 \sum_{j=0}^{15} |P_{2i,j} - P_{2i+1,j}| + |P_{2i+1,j} - P_{2i+2,j}| > \sum_{i=0}^6 \sum_{j=0}^{15} |P_{2i,j} - P_{2i+2,j}| + |P_{2i+1,j} - P_{2i+3,j}| + bias$$

其中， $p_{i,j}$ 是在 D C T 作用于每一 9×9 亮度块之前的空间亮度差分（如残余）数据。其优势在于，该式仅使用一阶差分，由此可使运算简便而廉价。项 " b i a s " 是考虑未被计算在内的非线性影响的因子。例如， $b i a s = 64$ 可被采用。若上述关系不成立，则采用帧 D C T。

注意，在上式的左手侧，误差项指对立场像素差（如，偶对奇，奇对偶）。因此，左手侧是对立场行亮度值之差的和。在右手侧，误差项指相同场像素差（如，偶对偶，奇对奇）。因此，右手侧是相同场行亮度数据差值与偏差项的和。

或者，通过修正上式以取每一误差项的平方值而不是绝对值，二阶式可被用于确定是帧还是场 D C T 可被使用。在此情况下，" b i a s " 项不被要求。

图 5 是描述根据本发明的 VOP 空间和时间缩放的图。通过基于目标的缩放，所选 VOP 的帧频和空间分辨率可被增强，以使其比其余区域具有更高的质量，例如，所选目标的帧频和 / 或空间分辨率可比其余区域高。例如，新闻播音员的 VOP 可被提供比演播室背景更高的分辨率。

轴 505 和 506 指明帧数。在基层，包括 VOP520 的帧 510 被提供于帧 0 的位置，而带有 VOP532 (对应于 VOP520) 的帧 530 被提供于帧 3 的位置。而且，帧 530 由帧 510 预测，如箭头 512 所示。增强层包括 VOP522、524、526 和 542。这些 VOP 相对于 VOP520 和 532 具有增加的空间分辨率，由此被画成较大区域。

P - VOP522 如箭头 570 所示由向上取样 VOP520 得出。B - VOP524 和 526 由基层 VOP520 和 532 预测，分别如箭头 572 和 576，及 574 和 578 所示。

用于生成基层和增强层序列的输入视频序列具有全分辨率 (例如，对于对应国家电视标准委员会 (N T S C) 的 I T U - R 601 为 720×490 或对于对应相位变化线 (P A L) 的 I T U - R 为 720×576) 和全帧频 (对于对应 N T S C 的 I T U - R 为 30 帧 / 60 场或对于对应 P A L 的 I T U - R 601 为 25 帧 / 50 场)。可缩放编码被实行，以通过利用增强层编码使物体的分辨率和帧频被保留。基层的视频目标，包括 VOP520 和 532，具有较低的分辨率 (如全分辨率 VOP 的四分之一) 和较低的帧频 (如初始帧频的三分之一)。而且，在增强层，只有 VOP520 被增强。帧 510 的其余部分未被增强。虽然仅一个 VOP 被示出，实际上，可提供任意数量的 VOP。而且，当两个或更多数量的 VOP 被提供时，所有或只有所选择的 VOP 可被增强。

基层序列可由对原始序列进行向下取样和帧抛弃而生成。然后基层 VOP 通过顺序编码工具被编码为 I - VOP 或 P - VOP。当输入视频序列是交错的，交错的编码工具如场 / 帧运动估计和补偿，及场 / 帧 D C T 不被使用，因

为输入交错视频序列的向下取样生成顺序视频序列。增强层 VOP 利用时间和空间可缩放工具编码。例如，在增强层，VOP522 和 VOP542 被利用空间可缩放编码为 P - VOP。VOP524 和 VOP526 利用时间和空间可缩放工具分别由基层参考 VOP 的向上取样 VOP，即 VOP520 和 VOP532，被编码为 B - VOP。

本发明的另一方面揭示了一种通过减少运动向量搜索范围以减少 B - VOP 的运动估计编码复杂性的技术。该技术可应用于帧模式和场模式输入视频序列。尤其是，参考 VOP 的搜索中心通过对相应基层 VOP 的运动向量调整，而不是通过在参考 VOP 中进行独立的耗时的搜索来确定。该耗时的搜索一般覆盖，例如水平 + / - 64 像素和垂直 + / - 49 像素的范围，因此没有所揭示技术高效。

增强层 B - VOP 524 和 526 的运动向量搜索中心由下式决定：

$$MV_r = (m / n \cdot TR_B \cdot MV_p) / TR_p$$

$$MV_b = (m / n \cdot (TR_B - TR_p) \cdot MV_p) / TR_p$$

其中， MV_r 是向前运动向量， MV_b 是向后运动向量， MV_p 是基层中 P - VOP（如 VOP 532）的运动向量， TR_B 是已往参考 VOP（如 VOP520）与增强层中当前 B - VOP 之间的时间间隔， TR_p 是已往参考 VOP 与基层未来参考 P - VOP（如 VOP532）之间的时间间隔。 m / n 是基层 VOP 空间分辨率对增强层 VOP 空间分辨率之比值。即，基层 VOP 或增强层中的 B - VOP 可相对于输入视频序列以比值 m / n 被向下取样。在图 5 的实例中， m / n 是随后被向上取样以提供增强层 VOP 的基层 VOP 的向下取样率。 m / n 可能小于、等于或大于 1。例如，对于 B - VOP524， $TR_B = 1$ ， $TR_p = 3$ ，且 2: 1 向下取样（即， $m / n = 2$ ），我们可得 $MV_r = 2 / 3 MV_p$ ，且 $MV_b = -4 / 3 MV_p$ 。注意所有运动向量是二维的。运动向量搜索范围是 16×16 方形区域，例如，其中心由 MV_r 和 MV_b 确定。运动向量与传输数据流中的增强层和基层视频数据传送，且由用于对视频数据解码的解码器恢复。

通常，为根据本发明对交错视频编码，交错编码工具被用于获取更好的性能。这些工具包括用于宏块内和差分间宏块的场 / 帧 DCT，以及场预测，即，顶部场到底部场、顶部场到顶部场、底部场到顶部场和底部场

到底部场。

对于上述表 1 中的配置，这些交错编码工具被结合如下。

(1) 对于两层均具有低空间分辨率的配置，只有顺序（帧模式）编码工具被使用。在此情况下，两层将编码不同的视频序列，例如，在立体视频信号中。为编码立体视频，对右视图（增强层）序列的运动估计搜索范围是 9×9 像素。该 9×9 （全像素）搜索范围集中于对应 VOP 的基层中相应宏块的同类型运动向量周围。

(2) 对于基层具有低空间分辨率、增强层具有高空间分辨率的配置，交错编码工具只用于增强层序列。用于对增强层序列编码的运动估计搜索范围是 9×9 （全像素）。该 9×9 搜索范围集中于对应 VOP 的基层中相应宏块的重可缩放（即因子为 2）同类型运动向量周围。基于场的估计和预测将仅被用于增强层的搜索和补偿。

(3) 对于基层具有高空间分辨率、增强层具有低空间分辨率的配置，交错编码工具只用于基层序列，如 MPEG-2 主框架的主层。用于对增强层序列编码的运动估计搜索范围是 4×4 （全像素）。该 4×4 搜索范围集中于对应 VOP 的基层中相应宏块的重可缩放（即， $1/2$ 倍）同类型运动向量周围。对于表 1 中的配置 2，如上所述，例如，两层序列的编码具有不同的时间单位频率。

图 7 是根据本发明带有空间和时间缩放的画中画（PIP）或预览频道应用图。通过 PIP，次级节目被作为电视上所见到的主要节目的子集而提供。由于次级节目占有较小区域，观众不关心减小了分辨率的图像，于是 PIP 图像的时间和/或空间分辨率可被减小以保留带宽。

类似地，预览频道节目可给用户（该节目可用一定费用购买）提供免费低分辨率样本。此应用提供授权频道（如，Pay-Per-View 为每幅画面付费）的几分钟免费浏览以进行预览。在预览频道编码的视频可具有较低的分辨率和较低的帧频。解码器为此类预览频道控制浏览时间。

表 1 中时间和空间可缩放编码的配置 2，如上所述，可被用于提供具有

比基层和增强层解码输出更低空间分辨率的基层解码输出。基层的视频序列可用较低的帧频编码，而增强层用较高帧频编码。

例如，基层中的视频序列可具有 C I F 分辨率和 15 帧 / 秒的帧频，增强层中的对应视频序列具有 I T U - R 601 分辨率和 30 帧 / 秒的帧频。在此情况下，增强层可符合 N T S C 标准，而 PIP 或预览功能由基层提供，其可符合 C I F 标准。由此，PIP 功能可由具有与 MPEG-2 主框架 M a i n L e v e l 标准相类似的编码复杂度和效率的可缩放编码提供。

基层包括低空间分辨率 VOP705 和 730。而且，基层的时间分辨率是增强层的 1 / 3。增强层包括高空间分辨率 VOP750、760、780 和 790。P - VOP 750 由向上取样 I - VOP705 得到，如箭头 755 所示。B - VOP760 如箭头 765 和 775 所示由基层 VOP 预测。B - VOP780 如箭头 770 和 785 所示由基层 VOP 预测。P - VOP 790 由向上取样 P - VOP 730 而获得，如箭头 795 所示。

如图 8 所示是根据本发明立体视频应用图。立体视频功能被提供于 MPEG-2 多视图框架 (M V P) 系统，在文献 I S O / I E C / J T C 1 / S C 29 / W G 11 N 1196 中描述。基层被分配到左视图，增强层被分配到右视图。

为提高编码效率，增强层图像可用比基层低的分辨率编码。例如，表 1 的配置 4 可被用于基层具有 I T U - R 601 空间分辨率，增强层具有 C I F 空间分辨率的情况。用于预测增强层图像的基层参考图像被向下取样。由此，增强层的解码器包括向上取样过程。此外，自适应帧 / 场 D C T 编码可被用于基层，但不可用于增强层。

基层包括 VOP805、815、820 和 830，增强层包括 VOP850、860、880 和 890。B - VOP 815 和 820 分别如箭头 810、840，和 835、825 所示由其它基层 VOP 预测。P - VOP830 如箭头 845 所示由 I - VOP805 预测。P - VOP 850 如箭头 855 所示由 I - VOP805 向下取样得出。B - VOP 860 如箭头 865 和 875 所示由基层 VOP 预测。B - VOP 880 如箭头 870 和 885 所示由基层 VOP 预测。P - VOP 890 如箭头 895 所示由 P - VOP830 向下取样得出。

或者，为使基层和增强层具有相同的空间分辨率和帧频，表 1 中的配

置 7 可被采用。在此情况下, 基层的编码过程可与非可缩放编码过程相同, 例如, 如 MPEG-4 V M 非可缩放编码或 MPEG-2 主框架 M a i n L e v e l 标准所描述, 自适应帧/场 D C T 编码被用于增强层中。

在本发明的进一步应用中, 异步传输模式 (A T M) 传送技术被提供。一般地说, 通过 A T M 网络传输视频信号的趋势迅速增长。这是由于这些网络较固定比特率 (C B R) 传输可提供许多优势的变比特率 (V B R) 性质。例如, 在 V B R 信道, 可获得大致稳定的图像质量。而且, 在 A T M 网络中的视频源可被按统计特性多路传输, 要求比通过 C B R 信道传输低的传输比特率, 因为 C B R 系统中的弹性缓冲使得视频信号的长期平均数据率小于短期平均值。

但是, 尽管 A T M 网络具有许多优势, 其主要不足在于拥塞。在拥塞的网络中, 视频包被排队以寻找输出路径。长期延迟包可能到达太晚从而在接收机处无用, 结果被解码器抛弃。视频编码解码器必须设计成可经受包丢失。

为使视频编码器几乎不受包丢失的影响, 可使用本发明的时间和空间可缩放编码技术。尤其是, 来自基层的视频数据可以高优先级被发送, 并被提供 A T M 网络保用的比特率。若发生拥塞, 由于信道不是保证的, 来自增强层的视频数据包可能丢失。若增强层包被接收到, 则图像质量提高。一采用表 1 中配置 1 的编码方案可取得该结果。该方案可如图 4 所示被实现, 如上结合预测模式的描述, 其中, 基层是高优先级层。因此, 在基层进行高优先级低比特率传送, 在增强层进行低优先级高比特率传送。

类似地, 该可缩放编码可用于视频编码和通过 Internet、i n t r a n e t 和其它传送网的传输。

由此, 可以看出, 本发明提供为一数字视频序列中包括视频对象面 (V O P) 的视频图像提供时间和空间缩放的方法和装置。本发明一方面通过自适应压缩可缩放的场模式输入视频序列使编码效率提高。基于线性准则, 增强层中的向上取样 V O P 被重排序以提供与原始视频序列更大的相关性。通过使用例如 D C T 的空间变换对所产生的残余编码。本发明另一方面, 提

出通过可缩放运动向量（其已对基层 VOP 确定）对增强层 VOP 编码的运动补偿方案。减小的搜索区域被定义（其中心由可缩放运动向量定义）。此技术适用于可缩放帧模式或场模式输入视频信号。

此外，各种编码解码器处理器配置被提出以取得特定的可缩放编码结果。可缩放编码的应用，包括立体视频、画中画、预览频道和 A T M 传送，也被描述。

尽管本发明已结合特定的实施例被描述，本领域的技术人员可理解，在不偏离权利要求中所提出的发明的精神与范围的情况下，可对其进行各种改进。例如，虽然两可缩放层被描述，也可提供多于两层。而且，虽然为简便起见，在一些图中采用方形或长方形 VOP，本发明同样适用于任意形状的 VOP。

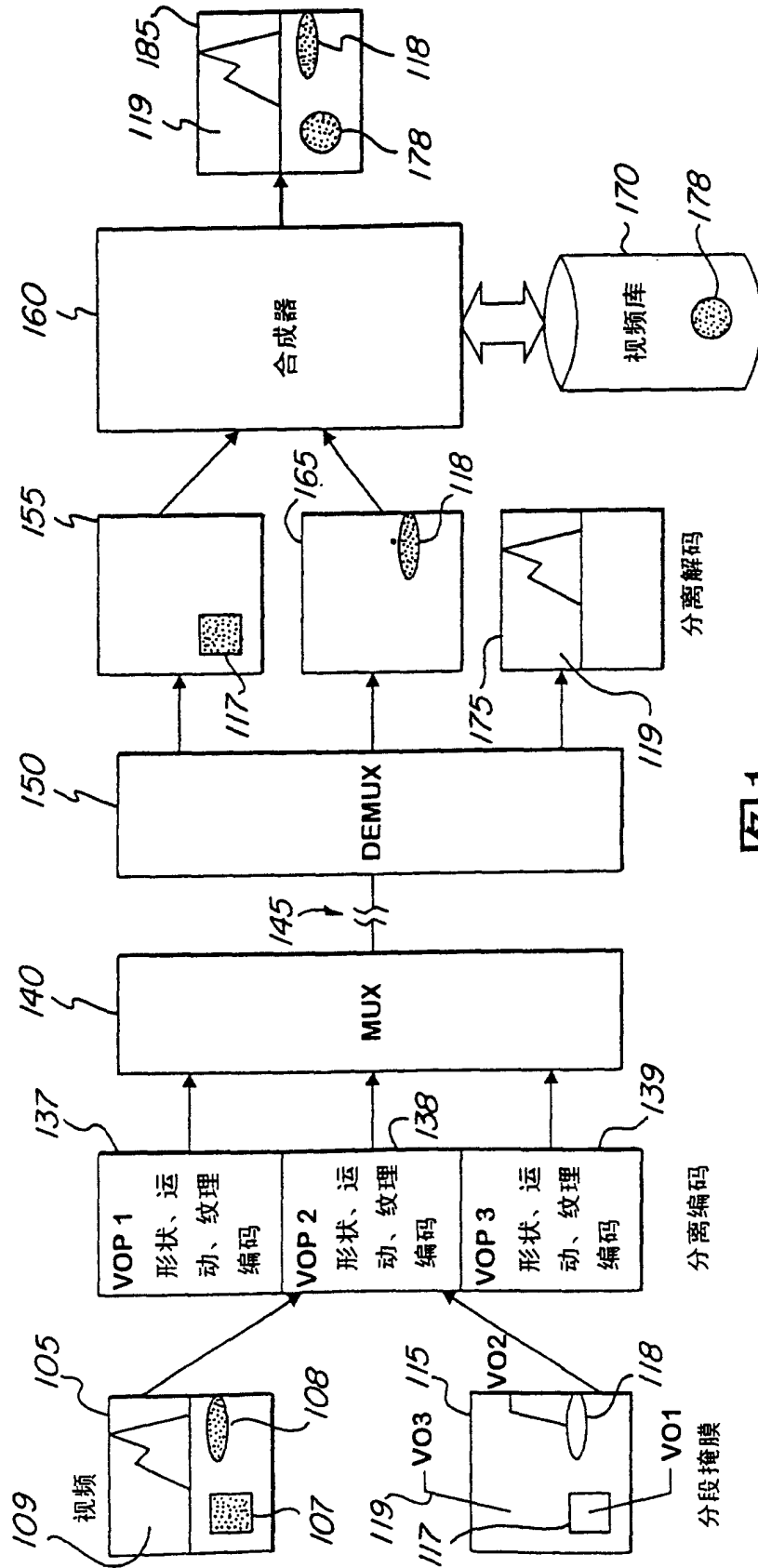


图1

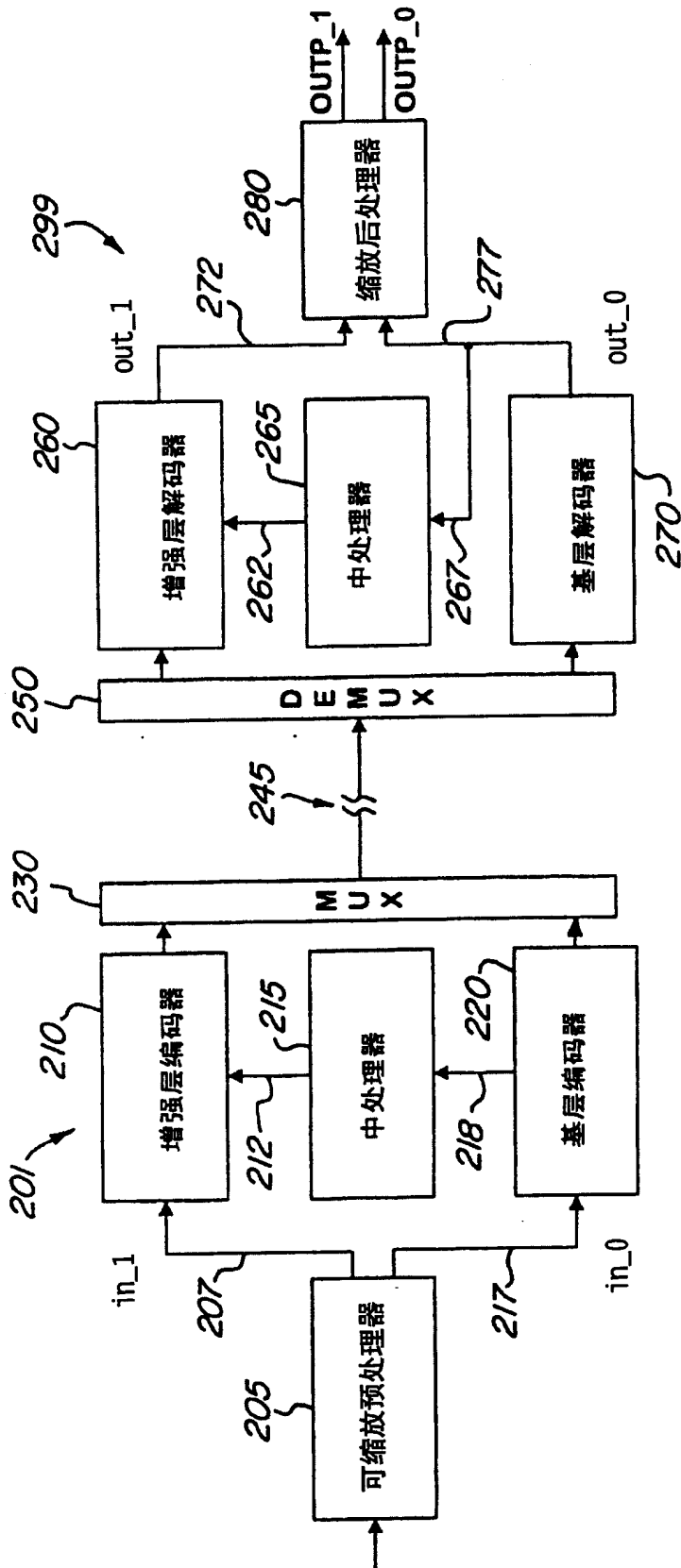


图2

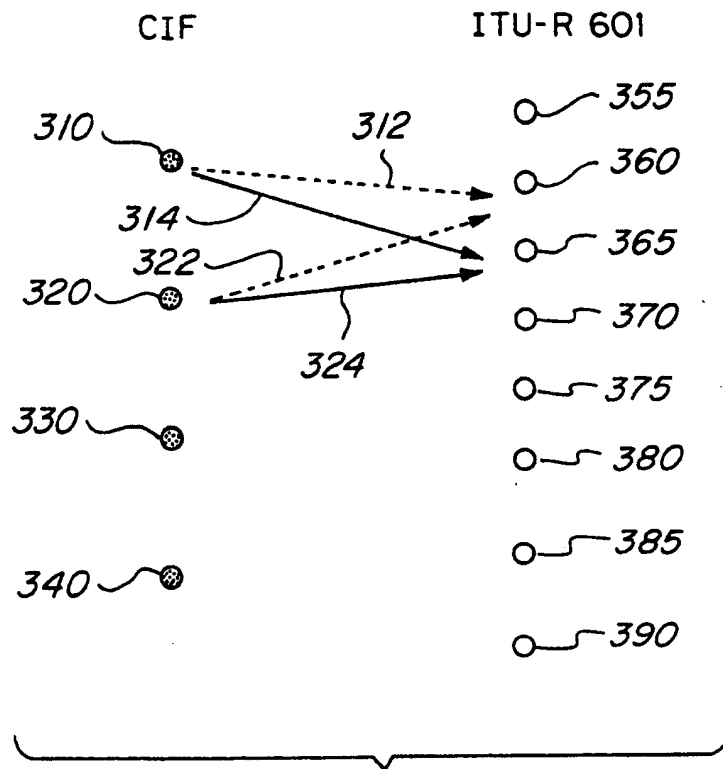


图3

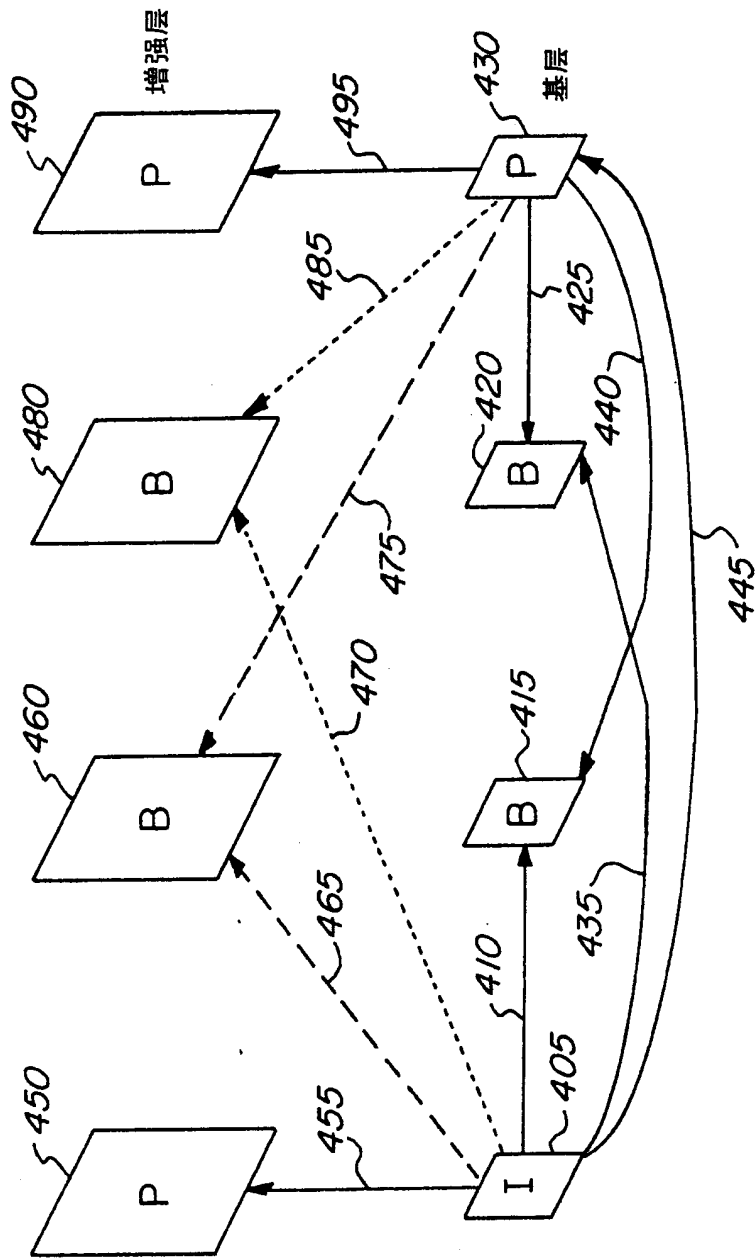


图4

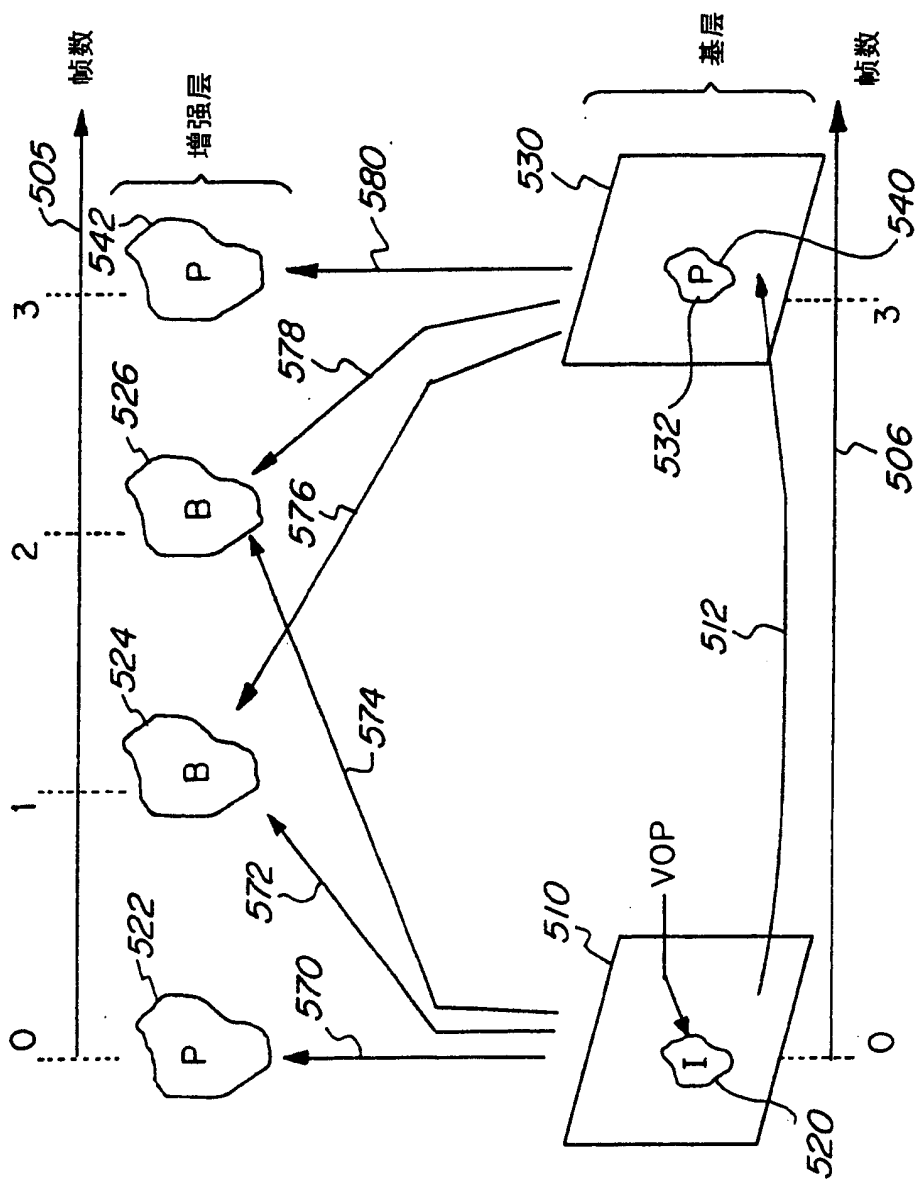


图5

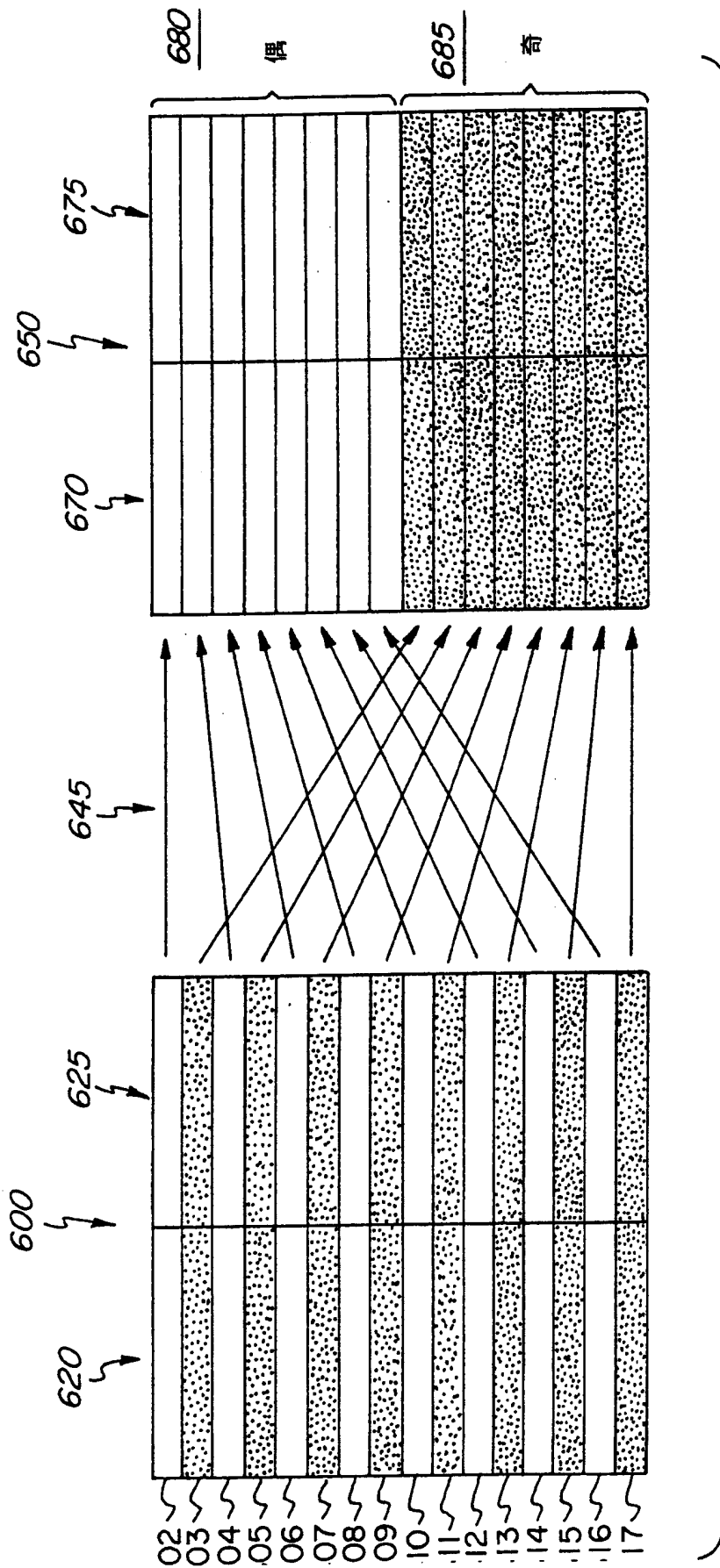


图6

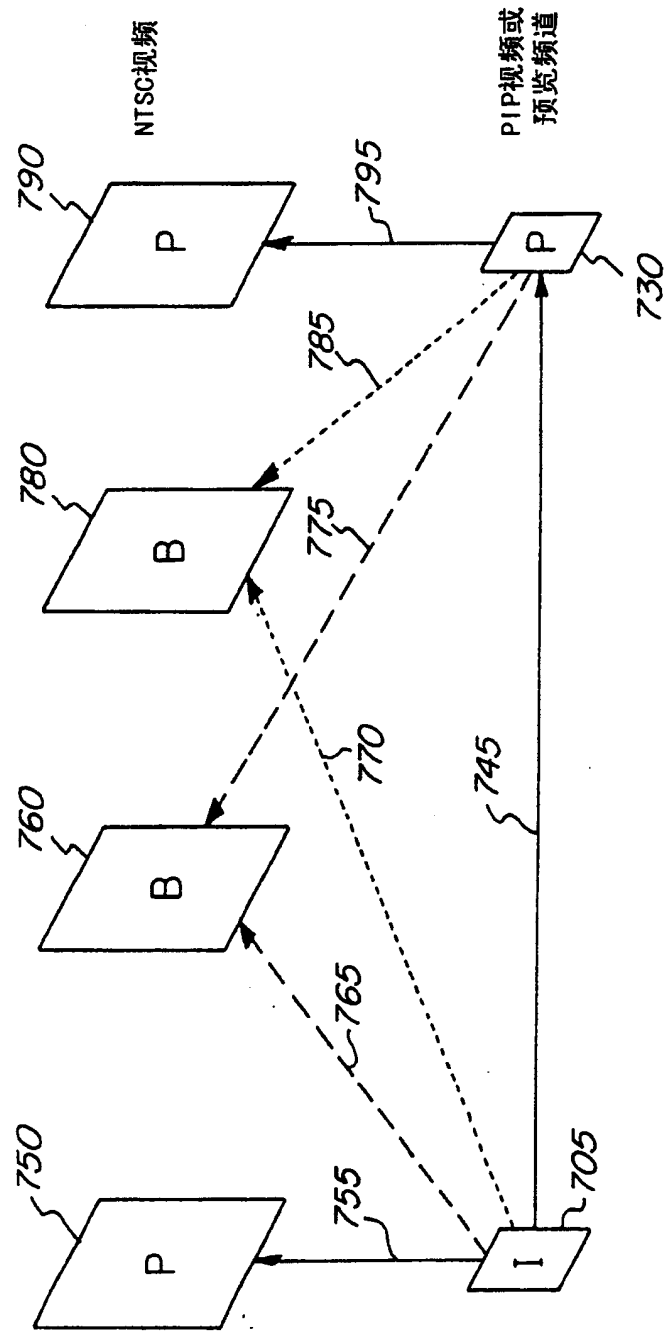


图7

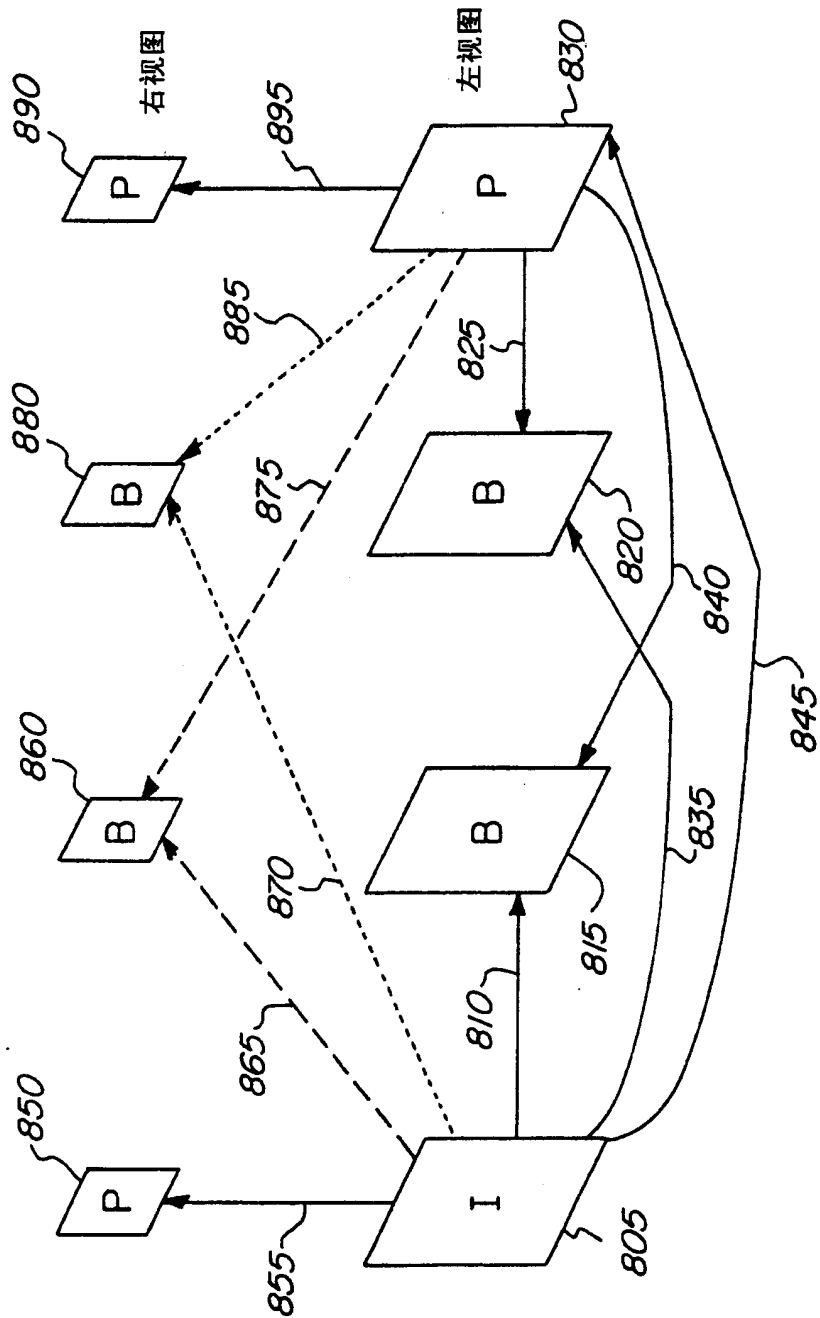


图 8