



(12)发明专利

(10)授权公告号 CN 104159060 B

(45)授权公告日 2017.10.24

(21)申请号 201410438251.8

(22)申请日 2007.03.13

(65)同一申请的已公布的文献号
申请公布号 CN 104159060 A

(43)申请公布日 2014.11.19

(30)优先权数据
60/789,048 2006.04.03 US
60/789,266 2006.04.04 US
60/789,377 2006.04.04 US

(62)分案原申请数据
200780010753.9 2007.03.13

(73)专利权人 高通股份有限公司
地址 美国加利福尼亚州

(72)发明人 田涛 刘方 石方
维贾雅拉克希米·R·拉韦恩德拉
恩

(74)专利代理机构 北京律盟知识产权代理有限
责任公司 11287

代理人 宋献涛

(51)Int.Cl.
H04N 7/01(2006.01)
H04N 5/14(2006.01)
H04N 19/523(2014.01)
H04N 19/87(2014.01)
H04N 19/19(2014.01)
H04N 19/107(2014.01)
H04N 19/176(2014.01)
H04N 19/86(2014.01)
H04N 19/147(2014.01)

(56)对比文件
US 5864369 A,1999.01.26,
CN 1372769 A,2002.10.02,
EP 1005227 A2,2000.05.31,
EP 1164792 A3,2003.08.13,

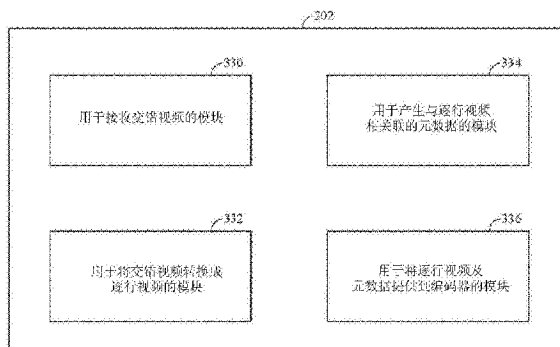
审查员 李春彦

权利要求书2页 说明书35页 附图40页

(54)发明名称
预处理器方法及设备

(57)摘要

本发明涉及预处理器方法及设备,且更明确地说,涉及在数据压缩处理之前或与数据压缩处理一起执行的处理操作。一种处理多媒体数据的方法包括接收交错视频帧,获得用于所述交错视频帧的元数据,使用所述元数据的至少一部分将所述交错视频帧转换成逐行视频;及将所述逐行视频及所述元数据的至少一部分提供到编码器以用于编码所述逐行视频。所述方法还可包括产生用于所述交错视频帧的空间信息及双向运动信息,及使用所述空间信息及所述双向运动信息基于所述交错视频帧而产生逐行视频。



1. 一种处理多媒体数据的方法,其包含:
接收数字交错视频帧;
通过解交错所述数字交错视频帧将所述数字交错视频帧转换成数字逐行视频帧,
其中,所述解交错包括:
产生用于所述数字交错视频帧的空间信息和用于所述数字交错视频帧中的至少一个帧的运动信息;以及
使用所述空间信息和所述运动信息来产生所述数字逐行视频帧;
其中使用所述空间信息和运动信息两者产生的第一帧所对应的像素信息,相比紧接在所述第一帧之后的未使用所述运动信息而产生的第二帧所对应的像素信息,被施加更多权重。
2. 根据权利要求1所述的方法,其中解交错进一步包含:
产生用于所述数字交错视频帧的双向运动信息;及
使用所述双向运动信息基于所述数字交错视频帧而产生所述数字逐行视频。
3. 根据权利要求1所述的方法,其中转换所述数字交错视频帧包含反电视电影处理3:2下拉视频帧。
4. 根据权利要求1所述的方法,其进一步包含对所述数字逐行视频调整大小。
5. 根据权利要求1所述的方法,其进一步包含用去噪声滤波器滤波所述数字逐行视频。
6. 根据权利要求1所述的方法,进一步包括:基于经转换数字逐行视频帧产生元数据;
基于所述元数据确定编码参数;以及,根据所述编码参数来编码所述数字逐行视频帧。
7. 一种用于处理多媒体数据的设备,其包含:
接收器,其经配置以接收数字交错视频帧;
解交错器,其经配置以通过解交错所述数字交错视频帧将所述数字交错视频帧转换成数字逐行视频帧,其中,所述解交错包括:
产生用于所述数字交错视频帧的空间信息和用于所述数字交错视频帧中的至少一个帧的运动信息;以及
使用所述空间信息和所述运动信息来产生所述数字逐行视频帧;
其中使用所述空间信息和运动信息两者产生的第一帧所对应的像素信息,相比紧接在所述第一帧之后的未使用所述运动信息而产生的第二帧所对应的像素信息,被施加更多权重。
8. 根据权利要求7所述的设备,其进一步包含编码器,所述编码器经配置以接收所述数字逐行视频帧且根据分割器产生的压缩信息来编码所述数字逐行视频,该分割器经配置以产生与所述数字逐行视频帧相关联的元数据。
9. 根据权利要求7所述的设备,其进一步包含用于对所述数字逐行视频帧进行去噪声的去噪声滤波器。
10. 根据权利要求7所述的设备,其中所述解交错器包含反电视电影处理器。
11. 根据权利要求7所述的设备,其进一步包含经配置以对所述数字逐行视频帧的逐行视频帧调整大小的重取样器。
12. 根据权利要求7所述的设备,其中解交错器经配置以产生用于所述数字交错视频帧的双向运动信息;且

使用所述双向运动信息基于所述数字交错视频帧而产生所述数字逐行视频帧。

13. 根据权利要求7所述的设备,进一步包括分割器,该分割器经配置以:产生与所述数字逐行视频帧相关联的元数据,将所述数字逐行视频帧和所述元数据提供给编码器以用于所述数字逐行视频帧的编码,其中所述元数据包括压缩信息。

14. 一种用于处理多媒体数据的设备,其包含:

接收装置,其用于接收数字交错视频帧;

转换装置,其用于通过解交错所述数字交错视频帧将所述数字交错视频帧转换成数字逐行视频帧,其中,所述解交错包括:

产生用于所述数字交错视频帧的空间信息和用于所述数字交错视频帧中的至少一个帧的运动信息;以及

使用所述空间信息和所述运动信息来产生所述数字逐行视频帧;

其中使用所述空间信息和运动信息两者产生的第一帧所对应的像素信息,相比紧接在所述第一帧之后的未使用所述运动信息而产生的第二帧所对应的像素信息,被施加更多权重。

15. 根据权利要求14所述的设备,其中所述转换装置包含反电视电影处理器。

16. 根据权利要求14所述的设备,其进一步包含重取样装置,所述重取样装置用于重取样以对逐行帧调整大小。

17. 根据权利要求14所述的设备,其进一步包含编码装置,所述编码装置用于使用所提供的与所述数字逐行视频帧相关联的元数据来编码所述数字逐行视频帧。

18. 根据权利要求14所述的设备,其进一步包含去噪声装置,所述去噪声装置用于对所述数字逐行视频帧去噪声。

19. 根据权利要求14所述的设备,其中所述转换装置经配置以:产生用于所述数字交错视频帧的双向运动信息;及使用所述双向运动信息基于所述交错视频帧而产生所述数字逐行视频帧。

20. 根据权利要求14所述的设备,进一步包括:用于产生与所述数字逐行视频帧相关联的元数据的装置;及用于将所述数字逐行视频帧和至少部分所述元数据提供给编码器以用于所述数字逐行视频帧的编码的装置,其中编码参数是基于所述至少部分元数据确定的。

21. 一种处理器,经配置用于:

接收数字交错视频帧;

通过解交错所述数字交错视频帧将所述数字交错视频帧转换成数字逐行视频帧,

其中,所述解交错包括:

产生用于所述数字交错视频帧的空间信息和用于所述数字交错视频帧中的至少一个帧的运动信息;以及

使用所述空间信息和所述运动信息来产生所述数字逐行视频帧;

其中使用所述空间信息和运动信息两者产生的第一帧所对应的像素信息,相比紧接在所述第一帧之后的未使用所述运动信息而产生的第二帧所对应的像素信息,被施加更多权重。

预处理器方法及设备

[0001] 分案申请的相关信息

[0002] 本申请为发明名称为“预处理器方法及设备”的原中国发明专利申请的分案申请。原申请的申请号为200780010753.9;原申请的申请日为2007年3月13日。

[0003] 根据35U.S.C.§119主张优先权

[0004] 本专利申请案主张2006年4月3日申请的第60/789,048号临时申请案、2006年4月4日申请的第60/789,266号临时申请案及2006年4月4日申请的第60/789,377号临时申请案的优先权,所有所述申请案均转让给本受让人且在此以引用的方式并入本文中。

技术领域

[0005] 本发明大体来说涉及多媒体数据处理,且更明确地说,涉及在数据压缩处理之前或与数据压缩处理一起执行的处理操作。

背景技术

[0006] 无

发明内容

[0007] 本文所描述的发明设备及方法中的每一者均具有若干方面,所述方面中的任何单个方面均不单独对其所要属性负责。在不限制本发明的范围情况下,现将简要论述其更突出特征。在考虑此论述之后,且特定来说,在阅读标题为“实施方式”的部分之后,应了解本发明的特征如何提供对多媒体数据处理设备及方法的改进。

[0008] 在一方面中,一种处理多媒体数据的方法包含接收交错视频帧,将所述交错视频帧转换成逐行视频(progressive video),产生与所述逐行视频相关联的元数据,及将所述逐行视频及所述元数据的至少一部分提供到编码器以用于编码所述逐行视频。所述方法可进一步包括使用元数据来编码逐行视频。在一些方面中,交错视频帧(interlaced video)包含NTSC视频。转换视频帧可包括解交错所述交错视频帧。

[0009] 在一些方面中,元数据可包括带宽信息、双向运动信息、带宽比(bandwidth ratio)、复杂性值(例如,时间复杂性值或空间复杂性值或两者)、亮度信息(luminance information),且空间信息可包括亮度及/或色度信息。所述方法还可包括产生用于所述交错视频帧的空间信息及双向运动信息,及使用所述空间信息及所述双向运动信息基于所述交错视频帧而产生逐行视频。在一些方面中,转换所述交错视频帧包含反电视电影处理3/2下拉视频帧,及/或对逐行视频调整大小。所述方法可进一步包含分割逐行视频以确定图像群组(group of picture)信息,其中所述分割可包括对逐行视频的分镜头检测(shot detection)。在一些方面中,所述方法还包括用去噪声滤波器来滤波逐行视频。

[0010] 在另一方面中,一种用于处理多媒体数据的设备可包括经配置以接收交错视频帧的接收器、经配置以将所述交错视频帧转换成逐行视频的解交错器,及经配置以产生与所述逐行视频相关联的元数据且将所述逐行视频及所述元数据提供到编码器以用于编码所

述逐行视频的分割器。在一些方面中,所述设备可进一步包括经配置以从通信模块接收逐行视频且使用所提供的元数据来编码逐行视频的编码器。所述解交错器可经配置以执行空间时间解交错及/或反电视电影处理(inverse telecining)。所述分割器可经配置以执行分镜头检测且基于所述分镜头检测产生压缩信息。在一些方面中,所述分割器可经配置以产生带宽信息。所述设备还可包括经配置以对逐行帧调整大小的重取样器。所述元数据可包括带宽信息、双向运动信息、带宽比、亮度信息、与内容相关的空间复杂性值,及/或与内容相关的时间复杂性值。在一些方面中,所述解交错器经配置以产生用于所述交错视频帧的空间信息及双向运动信息且使用所述空间信息及所述双向运动信息基于所述交错视频帧而产生逐行视频。

[0011] 另一方面包含一种用于处理多媒体数据的设备,所述设备包括用于接收交错视频帧的装置、用于将所述交错视频帧转换成逐行视频的装置、用于产生与所述逐行视频相关联的元数据的装置,及用于将所述逐行视频及所述元数据的至少一部分提供到编码器以用于编码所述逐行视频的装置。在一些方面中,所述转换装置包含反电视电影处理器及/或空间时间解交错器。在一些方面中,所述产生装置经配置以执行分镜头检测且基于所述分镜头检测产生压缩信息。在一些方面中,所述产生装置经配置以产生带宽信息。在一些方面中,所述产生装置包括用于重取样以对逐行帧调整大小的装置。

[0012] 另一方面包含一种机器可读媒体,所述机器可读媒体包含用于处理多媒体数据的指令,所述指令在执行时致使机器:接收交错视频帧,将所述交错视频帧转换成逐行视频,产生与所述逐行视频相关联的元数据,且将所述逐行视频及所述元数据的至少一部分提供到编码器以用于编码所述逐行视频。

[0013] 另一方面包括一种处理器,所述处理器包含一配置,所述配置用以接收交错视频、将交错视频转换成逐行视频、产生与所述逐行视频相关联的元数据及将所述逐行视频及所述元数据的至少一部分提供到编码器以用于编码所述逐行视频。对交错视频的转换可包括执行空间时间解交错。在一些方面中,对交错视频的转换包含执行反电视电影处理。在一些方面中,元数据的产生包括基于检测镜头变化而产生压缩信息。在一些方面中,元数据的产生包括确定逐行视频的压缩信息。在一些方面中,所述配置包括用以对视频重取样以产生经调整大小的逐行帧的配置。在一些方面中,所述元数据可包括带宽信息、双向运动信息、复杂性信息(例如,基于内容的时间或空间复杂性信息)及/或压缩信息。

附图说明

- [0014] 图1为用于递送串流多媒体数据的通信系统的框图;
- [0015] 图2为包括预处理器的数字传输设施的框图;
- [0016] 图3A为预处理器的说明性方面的框图;
- [0017] 图3B为说明用于处理多媒体数据的过程的流程图;
- [0018] 图3C为说明用于处理多媒体数据的装置的框图;
- [0019] 图4为说明例示性预处理器的操作的框图;
- [0020] 图5为反电视电影处理过程中的相位决策的图;
- [0021] 图6为说明反电视电影处理视频的过程的流程图;
- [0022] 图7为对展示相位转变的格子的说明;

- [0023] 图8为对识别用以创建多个度量的相应帧的指导；
- [0024] 图9为说明如何创建图8的度量的流程图；
- [0025] 图10为展示对度量的到达所估计相位的处理的流程图；
- [0026] 图11为说明用于产生决策变量的系统的数据流图；
- [0027] 图12为描绘用于评估分支信息的变量的框图；
- [0028] 图13A、13B及13C为展示如何计算下包络的流程图；
- [0029] 图14为展示一致性检测器的操作的流程图；
- [0030] 图15为展示计算决策变量的偏移的过程的流程图，所述偏移用以补偿相位决策中的不一致性；
- [0031] 图16呈现已估计下拉相位之后的反电视电影处理的操作。
- [0032] 图17为解交错器装置的框图；
- [0033] 图18为另一解交错器装置的框图；
- [0034] 图19为交错图像的子取样样式的图式；
- [0035] 图20为使用Wmed滤波运动估计来产生解交错帧的解交错器装置的框图；
- [0036] 图21说明用于确定多媒体数据的静态区的光圈的一方面；
- [0037] 图22为说明用于确定多媒体数据的缓慢运动区的光圈的一方面的图；
- [0038] 图23为说明运动估计的一方面的图；
- [0039] 图24说明在确定运动补偿时所使用的两个运动向量图；
- [0040] 图25为说明解交错多媒体数据的方法的流程图；
- [0041] 图26为说明使用空间时间信息产生解交错帧的方法的流程图；
- [0042] 图27为说明对于解交错执行运动补偿的方法的流程图；
- [0043] 图28为根据一些方面的预处理器的框图，所述预处理器包含经配置用于分镜头检测及其它预处理操作的处理器；
- [0044] 图29说明编码复杂性C与所分配位B之间的关系；
- [0045] 图30为说明对图像群进行操作且在一些方面中可用于基于视频帧中的分镜头检测而编码视频的过程的流程图；
- [0046] 图31为说明用于分镜头检测的过程的流程图；
- [0047] 图32为说明用于确定视频中的镜头的不同分类的过程的流程图；
- [0048] 图33为说明用于基于分镜头检测结果将帧压缩方案指派给视频帧的过程的流程图；
- [0049] 图34为说明用于确定突然的场景变化的过程的流程图；
- [0050] 图35为说明用于确定缓慢变化的场景的过程的流程图；
- [0051] 图36为说明用于确定含有相机闪光的场景的过程的流程图；
- [0052] 图37说明当前帧与前一帧 MV_P 之间及当前帧与下一帧 MV_N 之间的运动补偿向量；
- [0053] 图38为说明用于确定帧差异度量时所使用的变量的关系的图表；
- [0054] 图39为说明编码数据及计算残余的框图；
- [0055] 图40为说明确定帧差异度量的框图；
- [0056] 图41为说明其中将压缩类型指派给帧的程序的流程图；
- [0057] 图42说明1-D多相位重取样的实例；

[0058] 图43为说明数据帧的安全动作区及安全标题区的图表;及

[0059] 图44为说明数据帧的安全动作区的图表。

具体实施方式

[0060] 以下描述包括用于提供对实例的透彻理解的细节。然而,所属领域的技术人员将了解,即使本文中未描述或说明实例或方面中的过程或装置的每一细节,仍可实践所述实例。举例来说,可在框图中展示电组件,所述框图不说明所述组件的每一电连接或每一电元件以便不会以不必要的细节混淆所述实例。在其它情况下,可详细展示所述组件、其它结构及技术以进一步解释所述实例。

[0061] 本文中描述了某些发明方面及预处理器及预处理器操作方法的方面,其可改进现存预处理及编码系统的性能。所述预处理器可处理元数据及视频以为编码作准备,其包括执行解交错、反电视电影处理、滤波、识别镜头类型、处理及产生元数据及产生带宽信息。本文中对“一个方面”、“一方面”、“一些方面”或“某些方面”的参考意味着结合所述方面所描述的特定特征、结构或特性中的一者或一者以上可包括于预处理器系统的至少一个方面中。所述短语在说明书中的多个位置中的出现不一定指同一方面,也不一定指与其它方面相互排斥的单独或替代方面。此外,描述了一些方面可能展示出且其它方面可能未展示出的各种特征。类似地,描述了可为一些方面的步骤而非其它方面的步骤的各种步骤。

[0062] 如本文中所使用的“多媒体数据”或“多媒体”为广义术语,其包括视频数据(其可包括音频数据)、音频数据,或视频数据与音频数据两者。如本文中所使用的“视频数据”或“视频”为广义术语,其指含有文本、图像及/或音频数据的图像或者一个或一个以上系列或序列的图像,且除非另外规定,否则其可用于指多媒体数据或所述术语可互换使用。

[0063] 图1为用于递送串流多媒体的通信系统100的框图。所述系统可应用于将数字压缩视频传输到多个终端(如图1中所展示)。数字视频源可为(例如)数字电缆或卫星馈入或经数字化的模拟源。在传输设施120中处理视频源,在所述传输设施中视频源经编码并调制到载波上以用于经由网络140传输到一个或一个以上终端160。终端160解码所接收的视频且通常显示所述视频的至少一部分。网络140指适合于传输经编码数据的任何类型的通信网络(有线或无线)。举例来说,网络140可为蜂窝电话网络、有线或无线局域网络(LAN)或广域网络(WAN)或因特网。终端160可为能够接收并显示数据的任何类型的通信装置,其包括(但不限于)蜂窝电话、PDA、家用或商用视频显示设备、计算机(便携型、膝上型、手持型、PC及较大的基于服务器的计算机系统)及能够使用多媒体数据的个人娱乐装置。

[0064] 图2及图3说明预处理器202的样本方面。在图2中,预处理器202处于数字传输设施120中。解码器201解码来自数字视频源的经编码数据且将元数据204及视频205提供到预处理器202。预处理器202经配置以对视频205及元数据204执行特定类型的处理且将经处理的元数据206(例如,基本层参考帧、增强层参考帧、带宽信息、内容信息)及视频207提供到编码器203。对多媒体数据的所述预处理可改进视觉清晰度、反混叠(anti-aliasing)及数据压缩效率。通常,预处理器202接收由解码器201提供的视频序列且将所述视频序列转换成逐行视频序列以供由编码器进一步处理(例如,编码)。在一些方面中,预处理器202可经配置以用于许多操作,所述操作包括反电视电影处理、解交错、滤波(例如,假影移除、去环(de-ringing)、去块(de-blocking)及去噪声(de-noising))、调整大小(例如,从标准定义

到四分之一视频图形阵列(QVGA)的空间分辨率的下取样),及GOP结构产生(例如,计算复杂性映射产生、场景变化检测及衰落/闪光检测)。

[0065] 图3A说明预处理器202,其经配置有模块或组件(此处共同称作“模块”)以执行其所接收的元数据204及视频205的预处理操作且接着提供经处理的元数据206及逐行视频207以供进一步处理(例如,提供到编码器)。可以硬件、软件、固件或其组合实施所述模块。预处理器202可包括各种模块,所述模块包括所说明的模块中的一者或一者以上,所说明的模块包括反电视电影处理301、解交错器302、去噪声器303、混叠抑制器304、重取样器305、去块器/去环器306,及GOP分割器307,以下进一步描述所有模块。预处理器202还可包括可用于处理视频及元数据的其它适当模块,其包括存储器308及通信模块309。软件模块可驻留于RAM存储器、快闪存储器、ROM存储器、EPROM存储器、EEPROM存储器、寄存器、硬盘、可装卸磁盘、CD-ROM或此项技术中已知的任何其它形式的存储媒体中。例示性存储媒体耦合到处理器,以使得所述处理器可从存储媒体读取信息且将信息写入到存储媒体。在替代实施例中,存储媒体可与处理器成一体。处理器及存储媒体可驻留于ASIC中。所述ASIC可驻留于用户终端中。在替代实施例中,处理器及存储媒体可作为离散组件驻留于用户终端中。

[0066] 图3B为说明用于处理多媒体数据的过程300的流程图。过程300开始且进行到框320,在框320中,接收交错视频。图2及图3中所说明的预处理器202可执行此步骤。在一些方面中,解码器(例如,图2的解码器201)可接收交错数据且将交错数据提供到预处理器202。在一些方面中,图3C中所展示的数据接收模块330(其为预处理器202的一部分)可执行此步骤。过程300接着进行到框322,在框322中,将交错视频转换成逐行视频。图2及图3A中的预处理器202及图3C的模块332可执行此步骤。如果交错视频已经过电视电影处理,则框322处理可包括执行反电视电影处理以产生逐行视频。过程300接着进行到框324以产生与逐行视频相关联的元数据。图3A中的GOP分割器307及图3C中的模块334可执行所述处理。过程300接着进行到框326,在框326中,将逐行视频及元数据的至少一部分提供到编码器以用于编码(例如,压缩)。图2及图3A中所展示的预处理器202及图3C中的模块336可执行此步骤。在将逐行视频及相关联元数据提供到另一组件以用于编码之后,过程300可结束。

[0067] 图3C为说明用于处理多媒体数据的装置的框图。此处展示所述装置并入于预处理器202中。预处理器202包括用于接收视频的装置(例如,模块330)。预处理器202还包括用于将交错数据转换成逐行视频的装置(例如,模块332)。所述装置可包括(例如)空间时间解交错器及/或反电视电影处理器。预处理器202还包括用于产生与逐行视频相关联的元数据的装置(例如,模块334)。所述装置可包括如本文中所描述的可产生多种类型的元数据的GOP分割器307(图3A)。预处理器202还可包括用于将逐行视频及元数据提供到编码器以用于编码的装置,如通过模块336所说明。在一些方面中,所述装置可包括图3A中所说明的通信模块309。如所属领域的技术人员将了解,可以许多标准方式实施所述装置。

[0068] 预处理器202可将所获得的元数据(例如,从解码器201或从另一源所获得)用于所述预处理操作中的一者或一者以上。元数据可包括与描述或分类多媒体数据的内容有关的信息(“内容信息”)。具体来说,元数据可包括内容分类。在一些方面中,元数据不包括编码操作所需要的内容信息。在所述状况下,预处理器202可经配置以确定内容信息且将所述内容信息用于预处理操作及/或将内容信息提供到其它组件(例如,编码器203)。在一些方面中,预处理器202可使用所述内容信息来影响GOP分割,确定适当的滤波类型,及/或确定被

传送到编码器的编码参数。

[0069] 图4展示可包括于预处理器中的过程框的说明性实例,且说明可由预处理器202执行的处理。在此实例中,预处理器202接收元数据及视频204、205且将包含(经处理的)元数据及视频的输出数据206、207提供到编码器228。由预处理器接收的视频通常有三种类型。第一,所接收的视频可为逐行视频且不必执行解交错。第二,视频数据可为经电视电影处理的视频,从24fps电影序列转换的交错视频,在此状况下为视频。第三,视频可为未经电视电影处理的交错视频。预处理器226可如以下所描述处理所述类型的视频。

[0070] 在框401处,预处理器202确定所接收的视频204、205是否为逐行视频。在一些状况下,此可从元数据加以确定(如果元数据含有所述信息),或通过处理视频本身加以确定。举例来说,以下所描述的反电视电影处理过程可确定所接收的视频205是否为逐行视频。如果所接收的视频205为逐行视频,则过程进行到框407,在框407中,对视频执行滤波操作以减少噪声(例如,白高斯(Gaussian)噪声)。如果视频不为逐行视频,则在框401处,过程进行到框404到相位检测器。

[0071] 相位检测器404辨别起源于电视电影处理的视频与以标准广播格式开始的视频。如果作出视频是经电视电影处理的决策(从相位检测器404退出的是(YES)决策路径),则在反电视电影处理406中使经电视电影处理的视频返回到其最初格式。识别并消除冗余场且将从同一视频帧导出的场重新编成完整图像。因为以1/24秒的规则间隔摄影记录经重构的胶片图像的序列,所以在GOP分割器412或解码器中执行的运动估计过程更准确,所述过程是使用经反电视电影处理的图像而不是使用具有不规则时基的经电视电影处理的数据。

[0072] 在一方面中,相位检测器404在接收到视频帧之后作出某些决策。所述决策包括:(i)来自电视电影处理输出的当前视频及3:2下拉相位是否为图5中所展示的五個相位 P_0 、 P_1 、 P_2 、 P_3 及 P_4 中的一者;及(ii)视频经产生为常规的NTSC。所述决策经表示为相位 P_5 。所述决策表现为图4中所展示的相位检测器404的输出。来自相位检测器404的经标注为“是”的路径启动反电视电影处理406,指示反电视电影处理406已具备正确的下拉相位以使得其可选出由同一摄影图像所形成的场且组合所述场。来自相位检测器404的经标注为“否”的路径类似地启动解交错器405以将明显的NTSC帧分成若干场以用于最佳处理。反电视电影处理经进一步描述于标题为“基于状态机的反电视电影处理算法(Inverse Telecine Algorithm Based on State Machine)”的同在申请中的美国专利申请案[代理人案号为QFDM.021A(050943)]中,所述申请案归本发明的受让人所拥有且以全文引用的方式并入本文中。

[0073] 相位检测器404可连续分析视频帧,因为可在任何时间接收到不同类型的视频。作为例示,可将符合NTSC标准的视频插入视频中作为商业节目。在反电视电影处理之后,将所得的逐行视频发送到可用于减少白高斯噪声的去噪声器(滤波器)407。

[0074] 当辨识出常规NTSC视频时(来自相位检测器401的否(NO)路径),将所述视频传输到解交错器405以用于压缩。解交错器405将交错场变换成逐行视频,且可接着对逐行视频执行去噪声操作。

[0075] 在适当的反电视电影处理或解交错处理之后,在框408处,处理逐行视频以进行混叠抑制及重取样(例如,调整大小)。

[0076] 在重取样之后,逐行视频接着进行到框410,在框410中,执行去块及去环操作。两

种类型的假影“成块(blocking)”及“成环(ringing)”常见在视频压缩应用中出现。成块假影的出现是因为压缩算法将每一帧划分成若干块(例如,8×8个块)。重构每一块具有一些小误差,且块的边缘处的误差常常与邻近块的边缘处的误差形成对比,使得块边界可见。相比之下,成环假影表现为围绕图像特征的边缘的失真。成环假影的出现是因为编码器在量化高频DCT系数时丢弃了太多信息。在一些说明性实例中,去块与去环两者可使用低通FIR(有限脉冲响应)滤波器来隐藏所述可见假影。

[0077] 在去块及去环之后,由GOP分割器412处理逐行视频。GOP分割可包括检测镜头变化,产生复杂性映射(例如,时间、空间带宽映射)及自适应GOP分割。分镜头检测涉及确定图像群组(GOP)中的帧何时展示出指示场景变化已出现的数据。场景变化检测可用于视频编码器以确定适当GOP长度且基于所述GOP长度插入I帧,而不是以固定间隔插入I帧。预处理器202还可经配置以产生可用于编码多媒体数据的带宽映射。在一些方面中,改为由位于预处理器外部的内容分类模块产生带宽映射。自适应GOP分割可自适应地改变被编码在一起的图像群组的组成。以下描述图4中所展示的所述操作的说明性实例。

[0078] 反电视电影处理

[0079] 以下描述反电视电影处理过程且参看图4到图16提供反电视电影处理的说明性实例。当已知源属性且使用所述源属性来选择理想地匹配的处理形式时,视频压缩给出最佳结果。不在播送中(Off-the-air)的视频(例如)可以若干方式起源。视频相机、广播播音室等等中按照惯例所产生的广播视频在美国符合NTSC标准。根据所述标准,每一帧由两个场组成。一个场由奇数线组成,另一场由偶数线组成。此可被称作“交错”格式。虽然以大致30帧/秒的速率产生帧,但所述场为电视相机的图像的记录,所述记录相隔1/60秒。另一方面,以24帧/秒的速率拍摄胶片,每一帧由完整图像组成。此可被称作“逐行”格式。对于NTSC设备中的传输,“逐行”视频经由电视电影处理过程而转换成“交错”视频格式。以下进一步论述,在一方面中,系统有利地确定视频何时已经电视电影处理且执行适当变换以重新产生最初的逐行帧。

[0080] 图4展示电视电影处理已转换成交错视频的逐行帧的效应。 F_1 、 F_2 、 F_3 及 F_4 为逐行图像,其为电视电影处理器的输入。相应帧下的数字“1”及“2”为对奇数或偶数场的指示。注意,鉴于帧速率之间的不均等性,重复一些场。图4还展示下拉相位 P_0 、 P_1 、 P_2 、 P_3 及 P_4 。通过具有相同的第一场的两个NTSC兼容帧中的第一者标记相位 P_0 。随后的四个帧对应于相位 P_1 、 P_2 、 P_3 及 P_4 。注意,由 P_2 及 P_3 标记的帧具有相同的第二场。因为胶片帧 F_1 经扫描三遍,所以形成两个相同的相继输出的NTSC兼容第一场。从胶片帧 F_1 导出的所有NTSC场均从同一胶片图像取得且因此是在同一时刻被取得。从胶片导出的其它NTSC帧可具有相隔1/24秒的相邻场。

[0081] 图4中所说明的相位检测器404在接收到视频帧之后作出某些决策。所述决策包括:(i)来自电视电影处理输出的当前视频及3:2下拉相位是否为图5的定义512中所展示五个相位 P_0 、 P_1 、 P_2 、 P_3 及 P_4 中的一者;及(ii)视频经产生为常规的NTSC——所述决策被表示为相位 P_5 。

[0082] 所述决策表现为图4中所展示的相位检测器401的输出。来自相位检测器401的经标注为“是”的路径启动反电视电影处理406,其指示反电视电影处理406已具备正确的下拉相位以使得其可选出由同一摄影图像形成的场且组合所述场。来自相位检测器401的经标注为“否”的路径类似地启动解交错器框405以将明显的NTSC帧分成若干场以用于最佳处

理。

[0083] 图6为说明反电视电影处理视频流的过程600的流程图。在一方面中,由图3的反电视电影处理301执行过程600。在步骤651处开始,反电视电影处理器301基于所接收的视频确定多个度量。在此方面中,形成四个度量,所述四个度量为从同一帧或相邻帧取出的场之间的差异的和。所述四个度量经进一步汇编成对从所接收的数据导出的四个度量与对于六个假设相位的每一者来说所述度量的最可能的值之间的距离的欧几里德(Euclidean)测量。欧几里德和被称为分支信息;对于每一所接收的帧,存在六个所述量。每一假设相位具有后续相位,所述后续相位在可能的下拉相位的状况下随每一所接收的帧而改变。

[0084] 可能的转变路径展示于图7中且表示为767。存在六个所述路径。决策过程维持六个测量,所述测量等效于假设相位的每一路径的欧几里德距离的和。为使程序响应于已改变的条件,当所述和中的每一欧几里德距离变旧时,减小所述欧几里德距离。欧几里德距离的和最小的相位轨迹被认为是可操作的相位轨迹。此轨迹的当前相位被称为“适用相位”。基于经选择的相位的反电视电影处理现在可发生,只要其不是P₅。如果选择了P₅,则使用框405(图4)处的解交错器解交错当前帧。总之,利用适用相位作为当前下拉相位,或者作为用以命令对已经估计为具有有效NTSC格式的帧进行解交错的指示符。

[0085] 对于从视频输入接收到的每一帧,计算四个度量中的每一者的新值。所述度量经定义为:

$$[0086] \quad SAD_{FS} = \sum | \text{当前场一值}(i, j) - \text{先前场一值}(i, j) | \quad (1)$$

$$[0087] \quad SAD_{SS} = \sum | \text{当前场二值}(i, j) - \text{先前场二值}(i, j) | \quad (2)$$

$$[0088] \quad SAD_{P0} = \sum | \text{当前场一值}(i, j) - \text{先前场二值}(i, j) | \quad (3)$$

$$[0089] \quad SAD_{C0} = \sum | \text{当前场一值}(i, j) - \text{当前场二值}(i, j) | \quad (4)$$

[0090] 术语SAD是术语“合计绝对差”的缩写。图8中图解展示经区别以形成度量的场。下标指场号;字母表示先前(=P)或者当前(=C)。图8中的括号指场的成对差。SAD_{FS}指当前帧的经标注为C₁的场一与先前帧的经标注为P₁的场一之间的差,在图8中所提供的定义中经标注为FS的括号横跨所述场;SAD_{SS}指当前帧的经标注为C₂的场二与先前帧的经标注为P₂的场二之间的差,经标注为SS的括号横跨所述两个场;SAD_{C0}指当前帧的经标注为C₂的场2与当前帧的经标注为C₁的场一之间的差,经标注为C0的括号横跨所述场;且SAD_{P0}指当前帧的场一与先前帧的场2之间的差,经标注为P0的括号横跨所述两个场。

[0091] 以下描述用以评估每一SAD的计算负载。常规NTSC中存在大约480条有效水平线。为了使水平方向中的分辨率相同,具有4:3的纵横比,应存在 $480 \times 4/3 = 640$ 条均等的垂直线或自由度。 640×480 像素的视频格式是高级电视制式委员会所接受的格式的一者。因此,每隔1/30秒(一帧的持续时间),产生 $640 \times 480 = 307,200$ 个新像素。以 9.2×10^6 像素/秒的速率产生新数据,其暗示运行此系统的硬件或软件以大约为10MB或以上的速率处理数据。此为系统的高速部分的一者。其可通过硬件、软件、固件、中间件、微码或其任何组合予以实施。SAD计算器可为作为硬件、固件、中间件并入于另一装置的组件中的单独组件,或以在处理器上执行的微码或软件予以实施,或其组合。当以软件、固件、中间件或微码予以实施时,可将执行计算的程序码或码段存储于例如存储媒体的机器可读媒体中。码段可表示程序、函数、子程序、程序、例程、子例程、模块、软件封装、类,或指令、数据结构或程序语句的任何组合。可通过传递及/或接收信息、数据、自变量(argument)、参数或存储器内容而将码段耦

合到另一码段或硬件电路。

[0092] 图9中的流程图900使图8中的关系明确且图9是等式1到4的图解表示。图9展示分别保持SAD_{FS}、SAD_{CO}、SAD_{SS}及SAD_{PO}的最近值的存储位置941、942、943及944。所述值各自由绝对差计算器940的四个和所产生,所述绝对差计算器940处理先前第一场数据931的亮度值、当前第一场数据932的亮度值、当前第二场数据933的亮度值及先前第二场数据934的亮度值。在定义度量的求和中,术语“值(i, j)”意味着为位置i, j处的亮度值,求和对所有有效像素的求和,但是不排除对有效像素的有意义的子集的求和。

[0093] 图10中的流程图100为说明用于检测经电视电影处理的视频且反转经电视电影处理的视频以恢复到经最初扫描的胶片图像的过程的详细流程图。在步骤1030中,评估图9中所定义的度量。继续到步骤1083,寻找四个度量的下包络值。SAD度量的下包络为经动态确定的量,其是最高数字底限,在其下SAD不会穿过。继续到步骤1085,确定以下在等式5到10中所定义的分支信息量,其可使用先前所确定的度量、下包络值及用实验方法确定的常数A。因为连续的相位值可能不一致,所以在步骤1087中确定量 Δ 以减少此明显的不稳定性。当相位决策的序列与图7中所展示的问题模型一致时,认为所述相位是一致的。在所述步骤之后,过程进行到步骤1089以使用 Δ 的当前值计算决策变量。决策变量计算器1089使用通向其在框1080中产生的所有信息来评估决策变量。步骤1030、1083、1085、1087及1089为图6中的度量确定651的扩展。由相位选择器1090从所述变量寻找适用相位。如图所展示,决策步骤1091使用适用相位来反转经电视电影处理的视频或者解交错经电视电影处理的视频。其是对图4中的相位检测器404的操作的更明确叙述。在一方面中,由图4的相位检测器404执行图10的处理。在步骤1030处开始,检测器404通过上述参看图8所描述的过程确定多个度量,且继续通过步骤1083、1085、1087、1089、1090及1091。

[0094] 流程图1000说明用于估计当前相位的过程。所述流程图在步骤1083处描述使用经确定的度量及下包络值来计算分支信息。分支信息可经辨识为先前所论述的欧几里德距离。可用于产生分支信息的例示性等式为以下等式5到10。在图12的步骤1209中计算分支信息量。

[0095] 经处理的视频数据可存储于存储媒体中,所述存储媒体可包括(例如)芯片配置式存储媒体(例如,ROM、RAM)或连接到处理器的盘型存储媒体(例如,磁性的或光学的)。在一些方面中,反电视电影处理406及解交错器405可各自含有部分或所有存储媒体。通过以下等式定义分支信息量。

$$[0096] \quad \text{分支信息(0)} = (\text{SAD}_{\text{FS}} - \text{H}_s)^2 + (\text{SAD}_{\text{SS}} - \text{H}_s)^2 + (\text{SAD}_{\text{PO}} - \text{H}_p)^2 + (\text{SAD}_{\text{CO}} - \text{L}_c)^2 \quad (5)$$

$$[0097] \quad \text{分支信息(1)} = (\text{SAD}_{\text{FS}} - \text{L}_s)^2 + (\text{SAD}_{\text{SS}} - \text{H}_s)^2 + (\text{SAD}_{\text{PO}} - \text{L}_p)^2 + (\text{SAD}_{\text{CO}} - \text{H}_c)^2 \quad (6)$$

$$[0098] \quad \text{分支信息(2)} = (\text{SAD}_{\text{FS}} - \text{H}_s)^2 + (\text{SAD}_{\text{SS}} - \text{H}_s)^2 + (\text{SAD}_{\text{PO}} - \text{L}_p)^2 + (\text{SAD}_{\text{CO}} - \text{H}_c)^2 \quad (7)$$

$$[0099] \quad \text{分支信息(3)} = (\text{SAD}_{\text{FS}} - \text{H}_s)^2 + (\text{SAD}_{\text{SS}} - \text{L}_s)^2 + (\text{SAD}_{\text{PO}} - \text{L}_p)^2 + (\text{SAD}_{\text{CO}} - \text{L}_c)^2 \quad (8)$$

$$[0100] \quad \text{分支信息(4)} = (\text{SAD}_{\text{FS}} - \text{H}_s)^2 + (\text{SAD}_{\text{SS}} - \text{H}_s)^2 + (\text{SAD}_{\text{PO}} - \text{H}_p)^2 + (\text{SAD}_{\text{CO}} - \text{L}_c)^2 \quad (9)$$

$$[0101] \quad \text{分支信息(5)} = (\text{SAD}_{\text{FS}} - \text{L}_s)^2 + (\text{SAD}_{\text{SS}} - \text{L}_s)^2 + (\text{SAD}_{\text{PO}} - \text{L}_p)^2 + (\text{SAD}_{\text{CO}} - \text{L}_c)^2 \quad (10)$$

[0102] 在图12中的分支信息计算器1209中展示分支计算的细微细节。如在计算器1209中所展示,得出分支信息使用量 L_s (SAD_{FS}及SAD_{SS}的下包络值)、 L_p (SAD_{PO}的下包络值),及 L_c (SAD_{CO}的下包络值)。下包络在分支信息计算中用作距离偏移,以便单独或者与预定常数A一起产生 H_s 、 H_p 及 H_c 。在以下所描述的下包络跟踪器中保持下包络的值为最新的。 H 偏移经定义

为:

$$[0103] \quad H_s = L_s + A \quad (11)$$

$$[0104] \quad H_{p0} = L_p + A \quad (12)$$

$$[0105] \quad H_c = L_c + A \quad (13)$$

[0106] 在图13A、13B及13C中呈现跟踪 L_s 、 L_p 及 L_c 的值的过 程。考虑(例如)图13A顶部处所 展示的用于 L_p 的跟踪算法1300。在比较器1305中将度量 SAD_{p0} 与 L_p 的当前值加上阈值 T_p 进行 比较。如果 SAD_{p0} 超过 L_p 的当前值加上阈值 T_p ,则如框1315中所展示,不改变 L_p 的当前值。如果 SAD_{p0} 未超过 L_p 的当前值加上阈值 T_p ,则如框1313中所见, L_p 的新值变成 SAD_{p0} 与 L_p 的线性组 合。在另一方面中,对于步骤1315, L_p 的新值为 $L_p + T_p$ 。

[0107] 类似地计算图13B及图13C中的量 L_s 及 L_c 。图13A、13B及13C中具有相同功能的处理 框经相同地编号,但给出撇号('或")以展示其对不同变量组进行操作。举例来说,当形成 SAD_{p0} 与 L_c 的线性组合时,在框1313'中展示所述操作。对于 L_p 的状况,另一方面对于1315'将 用 $L_c + T_c$ 替换 L_c 。

[0108] 然而,在 L_s 的状况下,图13B中的算法交替地处理 SAD_{Fs} 及 SAD_{Ss} ,依次标注每一X,因 为此下包络适用于两个变量。当将框1308中的 SAD_{Fs} 的当前值读入框1303中X的位置中,随后 将1307中 SAD_{Ss} 的当前值读入框1302中X的位置中时,发生 SAD_{Fs} 值与 SAD_{Ss} 值的交替。对于 L_p 的状况,另一方面对于1315"将用 $L_s + T_s$ 替换 L_s 。通过实验预定用于测试当前下包络值的量A 及阈值。

[0109] 图11为说明用于执行图10的步骤1089的例示性过程的流程图。图11大致展示用于 更新决策变量的过程。在图11中用从度量导出的新信息来更新六个决策变量(对应于六个 可能的决策)。如下得到所述决策变量:

$$[0110] \quad D_0 = \alpha D_4 + \text{分支信息}(0) \quad (14)$$

$$[0111] \quad D_1 = \alpha D_0 + \text{分支信息}(1) \quad (15)$$

$$[0112] \quad D_2 = \alpha D_1 + \text{分支信息}(2) \quad (16)$$

$$[0113] \quad D_3 = \alpha D_2 + \text{分支信息}(3) \quad (17)$$

$$[0114] \quad D_4 = \alpha D_3 + \text{分支信息}(4) \quad (18)$$

$$[0115] \quad D_5 = \alpha D_5 + \text{分支信息}(5) \quad (19)$$

[0116] 量 α 小于一且限制决策变量对其过去值的依赖性; α 的使用等效于当欧几里德距离 的数据变旧时减小每一欧几里德距离的效应。在流程图1162中,在左侧将待更新的决策变 量在线1101、1102、1103、1104、1105及1106上列出为可用的。接着在框1100的一者中将相位 转变路径的一者上的决策变量的每一者乘以 α , α 为小于一的数;接着将旧的决策变量的衰 减值添加到通过相位转变路径上的下一相位编索引的分支信息变量的当前值,衰减的决策 变量在所述相位转变路径上。此在框1110中发生。在框1193中使变量 D_5 偏移一量 Δ ; Δ 是在 框1112中计算出。如以下所述,选择所述量以减少由此系统所确定的相位序列中的不一致 性。在框1120中得到最小的决策变量。

[0117] 总之,将每一决策所特有的新信息添加到已被乘以 α 的适当的决策变量的先前值 以得到当前决策变量的值。当已得到新度量时,可作出新的决策;因此,此技术能够在接收 到每一帧的场1及场2时作出新的决策。所述决策变量为先前所提到的欧几里德距离的和。

[0118] 适用相位经选择为具有最小决策变量的下标的相位。在图10的框1090中明确作出

基于决策变量的决策。决策空间中允许某些决策。如框1091中所描述,所述决策为:(i)适用相位不是 P_5 ——则反电视电影处理视频及(ii)适用相位为 P_5 ——则解交错视频。

[0119] 因为度量是从本质上可变的视频中取出,所以决策的相干串中可能存在偶然误差。此技术检测到与图7不一致的相位序列。在图14中概述其操作。算法1400在框1405中存储当前相位决策的下标(=x)且在框1406中存储先前相位决策的下标(=y)。在框1410中,测试是否 $x=y=5$;在框1411中,测试以下值:

[0120] 是否

[0121] $x=1, y=0$;或

[0122] $x=2, y=1$;或

[0123] $x=3, y=2$;或

[0124] $x=4, y=3$;或

[0125] $x=0, y=4$ 。

[0126] 如果两个测试的任一者为肯定的,则在框1420中宣告所述决策是一致的。如果任一测试均不是肯定的,则在图15中计算展示于图11的框1193中的偏移且将所述偏移添加到与 P_5 相关联的决策变量 D_5 。

[0127] 对 D_5 的修改作为过程1500的一部分还出现于图15中,所述修改对相位序列中的不一致性提供校正作用。假定流程图1500中的框1510中的一致性测试已失败。沿着从框1510引出的“否”分支进行,框1514中的下一测试为:对于所有 $i < 5$,是否 $D_5 > D_i$;或者为:对于 $i < 5$,所述变量的至少一者 D_i 是否大于 D_5 。如果第一状况有效,则在框1516中将初始值为 δ_0 的参数 δ 改变成 $3\delta_0$ 。如果第二状况有效,则在框1517中将 δ 改变成 $4\delta_0$ 。在框152B中,将 Δ 的值更新为 Δ_B ,其中

[0128] $\Delta_B = \text{最大}(\Delta - \delta, -40\delta_0)$ (20)

[0129] 再返回框15210,假定所述决策串经判定为一致。在框15215中,将参数 δ 改变成由下式定义的 δ_+

[0130] $\delta_+ = \text{最大}(2\delta, 16\delta_0)$ (21)

[0131] 在框152A中将 δ 的新值插入用于 Δ 的更新关系 Δ_A 中。此是

[0132] $\Delta_A = \text{最大}(\Delta + \delta, 40\delta_0)$ (22)

[0133] 接着在框1593中将 Δ 的更新值添加到决策变量 D_5 。

[0134] 图16展示一旦确定了下拉相位,反电视电影处理过程如何进行。使用此信息,将场1605及1605'识别为表示视频的同一场。对所述两个场一起求平均且将其与场1606组合以重构帧1620。经重构的帧为1620'。类似过程将重构帧1622。不复制从帧1621及1623导出的场。通过将所述帧的第一及第二场重新编织在一起而重构所述帧。

[0135] 在以上所描述的方面中,每当接收到新帧时,得到度量的四个新值且使用新近计算出的决策变量来测试六重假设组。其它处理结构可适合于计算所述决策变量。维特比(Viterbi)解码器将组成路径的分支的度量添加在一起以形成路径度量。此处所定义的决策变量通过类似规则所形成:每一者为新信息变量的“漏泄”和。(在漏泄求和中,在将新信息数据添加到决策变量的先前值之前,将所述决策变量的先前值乘以小于一的数。)维特比解码器结构可经修改以支持此程序的操作。

[0136] 虽然就处理常规视频(其中,每隔1/30秒出现一新帧)来说描述本方面,但应注意,

此过程可适用于在时间上向后记录并处理的帧。决策空间保持相同,但存在较小变化,所述变化反映出输入帧的序列的时间反转。举例来说,来自时间反转模式的一串相干电视电影处理决策(此处所展示的)

[0137] P₄P₃P₂P₁P₀

[0138] 将还在时间上被反转。

[0139] 使用对第一方面的此变化将允许决策在作出成功决策时进行两种尝试:一种尝试是在时间上向前,另一种尝试是在时间上向后。虽然两种尝试不是独立的,但其不相同,因为每一种尝试将以不同的次序处理度量。

[0140] 此想法可与缓冲器一起应用,所述缓冲器经维持以存储可能另外需要的将来的视频帧。如果发现视频片段在前向处理方向中给出不可接受地不一致结果,则程序将从所述缓冲器中取出将来的帧且试图通过在反方向中处理所述帧来克服视频的拉伸困难。

[0141] 此专利中所描述的对视频的处理还可适用于PAL格式的视频。

[0142] 解交错器

[0143] 如本文中所使用的“解交错器”是广义术语,其可用于描述完全或大部分地处理交错多媒体数据以形成逐行多媒体数据的解交错系统、装置或过程(包括例如经配置以执行过程的软件、固件或硬件)。

[0144] 视频相机、广播播音室等中按照惯例所产生的广播视频在美国符合NTSC标准。一种压缩视频的常见方式是将其交错。在交错数据中,每一帧由两个场中的一者组成。一个场由帧的奇数线组成,另一场由偶数线组成。虽然以大致30帧/秒产生帧,但所述场为电视相机的图像的的记录,所述记录相隔1/60秒。交错视频信号的每一帧展示图像的每隔一水平线。当所述帧投影于屏幕上时,视频信号在展示偶数线与奇数线之间交替。当足够快地执行此交替时(例如,大约每秒60个帧),视频图像在人眼看来是平滑的。

[0145] 在基于NTSC(美国)及PAL(欧洲)格式的模拟电视广播中已使用交错达数十年。因为用每一帧发送仅一半的图像,所以交错视频使用的带宽大约为其发送整个图像将会使用的带宽的一半。在终端16内部的视频的最后的显示格式不必要与NTSC兼容且不能容易地显示交错数据。相反,现代基于像素的显示器(例如,LCD、DLP、LCOS、等离子体等)是逐行扫描且显示经逐行扫描的视频源(而许多较旧的视频装置使用较旧的交错扫描技术)。一些常使用的解交错算法的实例描述于P·哈维斯托(P.Haavisto),J.祝霍拉(J.Juhola)及Y.约奥(Y.Neuvo)的“使用自适应加权中值滤波的扫描速率向上转换(Scan rate up-conversion using adaptive weighted median filtering)”(HDTV II的信号处理(Signal Processing of HDTV II),第703-710页,1990年)及R·西蒙内提(R.Simonetti),S.卡拉托(S.Carrato),G.蓝波尼(G.Ramponi)及A.保罗费里森(A.Polo Filisan)的“用于多媒体应用的HDTV图像的解交错(Deinterlacing of HDTV Images for Multimedia Applications)”(HDTV II的信号处理(Signal Processing of HDTV IV),第765-772页,1993年)中。

[0146] 以下描述用于可单独使用或组合使用以改进解交错的性能且可用于解交错器405(图4)中的系统及方法的解交错方面的实例。所述方面可包括使用空间时间滤波来解交错选定帧以确定第一临时解交错帧、使用双向运动估计及运动补偿从所述选定帧确定第二临时解交错帧,及接着组合所述第一与第二临时帧以形成最终的逐行帧。所述空间时间滤波

可使用加权中值滤波 (“Wmed”) 滤波器, 所述滤波器可包括防止使水平或近水平边缘模糊的水平边缘检测器。对“当前”场的先前及随后的邻近场的空间时间滤波产生强度运动等级映射, 所述强度运动等级映射将选定帧的若干部分分类成不同的运动等级, 例如, 静态、缓慢运动及快速运动。

[0147] 在一些方面中, 通过Wmed滤波使用包括来自五个邻近场(前两个场、当前场及后两个场)的像素的滤波光圈产生所述强度映射。Wmed滤波可确定可有效地处置场景变化及物体出现及消失的前向、后向及双向静态区检测。在各方面中, 可在场间滤波模式中跨越具有相同奇偶性的一个或一个以上场利用Wmed滤波器, 且可通过调节阈值标准而将其切换到场内滤波模式。在一些方面中, 运动估计及补偿使用亮度(像素的强度或明亮度)及色度数据(像素的色彩信息)来改进选定帧的解交错区域, 其中明亮度等级几乎一致但色彩不同。去噪声滤波器可用于增加运动估计的准确度。可将去噪声滤波器应用于经Wmed解交错的临时帧以移除由Wmed滤波所产生的混叠假影。以下所描述的解交错方法及系统产生优良的解交错结果且具有相对低的计算复杂性, 其允许快速运行解交错实施方案, 使所述实施方案适合于各种各样的解交错应用, 所述应用包括用于利用显示器将数据提供到蜂窝电话、计算机及其它类型的电子或通信装置的系统。

[0148] 本文中参考用于解交错多媒体数据的各种组件、模块及/或步骤来描述解交错器及解交错方法的方面。

[0149] 图17为说明可用作图4中的解交错器405的解交错器1700的一方面的框图。解交错器1700包括在空间及时间上(“空间时间”)滤波交错数据的至少一部分且产生空间时间信息的空间滤波器1730。举例来说, 可在空间滤波器1730中使用Wmed。在一些方面中, 解交错器1700还包括去噪声滤波器(未图示), 例如, 维纳(Weiner)滤波器或小波收缩(wavelet shrinkage)滤波器。解交错器1700还包括提供对交错数据的选定帧的运动估计及补偿且产生运动信息的运动估计器1732。组合器1734接收并组合空间时间信息与运动信息以形成逐行帧。

[0150] 图18为解交错器1700的另一框图。解交错器1700中的处理器1836包括空间滤波器模块1838、运动估计模块1840及组合器模块1842。来自外部源48的交错多媒体数据可被提供到解交错器1700中的通信模块44。可通过硬件、软件、固件、中间件、微码或其任何组合来实施解交错器及其组件或步骤。举例来说, 解交错器可为作为硬件、固件、中间件并入于另一装置的组件中的单独组件, 或以在处理器上执行的微码或软件予以实施, 或其组合。当以软件、固件、中间件或微码予以实施时, 可将执行解交错器任务的程序码或码段存储于例如存储媒体的机器可读媒体中。码段可表示过程、函数、子程序、程序、例程、子例程、模块、软件封装、类, 或指令、数据结构或程序语句的任何组合。可通过传递及/或接收信息、数据、自变量、参数或存储器内容而将码段耦合到另一码段或硬件电路。

[0151] 所接收的交错数据可存储于解交错器1700中的存储媒体1846中, 存储媒体1846可包括(例如)芯片配置式存储媒体(例如, ROM、RAM)或连接到处理器1836的盘型存储媒体(例如, 磁性的或光学的)。在一些方面中, 处理器1836可含有部分或所有存储媒体。处理器1836经配置以处理交错多媒体数据以形成随后被提供到另一装置或过程的逐行帧。

[0152] 类似电视的传统模拟视频装置以交错方式再现视频, 即, 所述装置传输编号为偶数的扫描线(偶数场)及编号为奇数的扫描线(奇数场)。从信号取样观点看, 这等效于以如

下所描述的样式进行的空间时间子取样 (subsampling) :

$$[0153] \quad F(x, y, n) = \begin{cases} \Theta(x, y, n), & \text{如果 } y \bmod 2 = 0 \text{ 对于偶数场,} \\ \Theta(x, y, n), & \text{如果 } y \bmod 2 = 1 \text{ 对于奇数场,} \\ \text{擦除,} & \text{否则,} \end{cases} \quad (23)$$

[0154] 其中 Θ 代表最初的帧图像, F 代表交错场, 且 (x, y, n) 分别表示像素的水平、垂直及时间位置。

[0155] 在不损失一般性的情况下, 可假定在本发明中 $n=0$ 始终为偶数场, 因此以上等式 23 经简化为

$$[0156] \quad F(x, y, n) = \begin{cases} \Theta(x, y, n), & \text{如果 } y \bmod 2 = n \bmod 2, \\ \text{擦除,} & \text{否则,} \end{cases} \quad (24)$$

[0157] 因为不是在水平尺寸中进行抽取, 所以可以接下来的 $n \sim y$ 坐标描绘子取样式。在图 19 中, 圆圈与星号均表示位置, 在所述位置中最初的全帧图像具有样本像素。解交错过程抽取星号像素, 而保留圆圈像素完整无缺。请注意, 我们从零开始对垂直位置编索引, 因此, 偶数场为顶部场, 且奇数场为底部场。

[0158] 解交错器的目标为将交错视频 (场序列) 变换成非交错逐行帧 (帧序列)。换句话说, 内插偶数及奇数场以“恢复”或产生全帧图像。此可通过等式 25 来表示:

$$[0159] \quad F_o(x, y, n) = \begin{cases} F(x, y, n), & y \bmod 2 = n \bmod 2, \\ F_i(x, y, n), & \text{否则,} \end{cases} \quad (25)$$

[0160] 其中 F_i 表示缺少的像素的解交错结果。

[0161] 图 20 为说明解交错器的一方面的某些方面的框图, 所述解交错器使用 W_{med} 滤波及运动估计从交错多媒体数据产生逐行帧。图 20 的上部分展示可使用来自当前场、前两个场 (PP 场及 P 场) 及后两个场 (下一场及再下一场) 的信息而产生的运动强度映射 2052。运动强度映射 2052 将当前帧分类或分割成两个或两个以上的不同运动等级, 且可通过下文中进一步详细描述的空间时间滤波而产生。在一些方面中, 产生运动强度映射 2052 以识别如以下参考等式 4 到 8 所描述的静态区、缓慢运动区及快速运动区。空间时间滤波器 (例如, W_{med} 滤波器 2054) 使用基于运动强度映射的标准来滤波交错多媒体数据, 且产生空间时间临时解交错帧。在一些方面中, W_{med} 滤波过程涉及 $[-1, 1]$ 的水平邻域、 $[-3, 3]$ 的垂直邻域, 及通过图 20 中所说明的五个场 (PP 场、P 场、当前场、下一场、再下一场) 表示的五个相邻场的时间邻域, 其中 Z^{-1} 表示一个场的延迟。相对于当前场, 下一场及 P 场为非奇偶场且 PP 场及再下一场为奇偶场。用于空间时间滤波的“邻域”是指在滤波操作期间实际所使用的场及像素的空间及时间位置, 且可说明为如 (例如) 图 21 及图 22 中所展示的“光圈”。

[0162] 解交错器还可包括去噪声器 (去噪声滤波器) 2056。去噪声器 2056 经配置以滤波由 W_{med} 滤波器 2054 产生的空间时间临时解交错帧。对空间时间临时解交错帧去噪声使随后的运动搜索过程更准确, 尤其是当源交错多媒体数据序列受白噪声污染时。去噪声器 2056 还可至少部分地移除 W_{med} 图像中偶数行与奇数行之间的混叠。可将去噪声器 2056 实施为多种

滤波器,包括同样在下文中进一步加以描述的基于小波收缩及小波维纳(Wiener)滤波器的去噪声器。

[0163] 图20的下部分说明用于确定交错多媒体数据的运动信息(例如,运动向量候选者、运动估计、运动补偿)的方面。具体来说,图20说明运动估计及运动补偿方案,所述运动估计及运动补偿方案用于产生选定帧的经运动补偿的临时逐行帧,且接着将其与Wmed临时帧组合以形成所得的“最终”逐行帧,其展示为经解交错的当前帧2064。在一些方面中,交错多媒体数据的运动向量(“MV”)候选者(或估计)从外部运动估计器提供到解交错器且用于为双向运动估计器及补偿器(“ME/MC”)2068提供起始点。在一些方面中,MV候选选择器2072对于正被处理的块的MV候选者使用先前所确定的用于邻近块的MV,例如先前经处理的块(例如,经解交错的先前帧2070中的块)的MV。可基于先前经解交错的帧70及下一(例如,将来的)Wmed帧2058而双向进行运动补偿。由组合器2062合并或组合当前Wmed帧2060与经运动补偿(“MC”)的当前帧2066。所得的经解交错的当前帧2064(现为逐行帧)经提供回ME/MC 2068以用作经解交错的先前帧2070且还经传送到解交错器外部以供进一步处理(例如,压缩及传输到显示终端)。以下更详细描述图20中所展示的各种方面。

[0164] 图25说明用于处理多媒体数据以从交错帧序列产生逐行帧序列的过程2500。在一方面中,由图4中所说明的解交错器405产生逐行帧。在框2502处,过程2500(过程“A”)产生用于选定帧的空间时间信息。空间时间信息可包括用于分类多媒体数据的运动等级且产生运动强度映射的信息,且包括Wmed临时解交错帧及用于产生所述帧的信息(例如,用于等式26到33中的信息)。可由如图20的上部分中所说明的Wmed滤波器2054及其相关处理(以下进一步详细描述)执行此过程。在图26中所说明的过程A中,在框2602处将区域分类成具有不同运动等级的场,如以下进一步加以描述。

[0165] 接下来,在框2504处(过程“B”),过程2500产生用于选定帧的运动补偿信息。在一方面中,在图20的下部分中所说明的双向运动估计器/运动补偿器2068可执行此过程。过程2500接着进行到框2506,其中过程基于空间时间信息及运动补偿信息而解交错选定帧的场以形成与选定帧相关联的逐行帧。此可由在图20的下部分中所说明的组合器2062来执行。

[0166] 运动强度映射

[0167] 对于每一帧,可通过处理当前场中的像素以确定不同“运动”的区来确定运动强度2052映射。以下参看图21到图24描述确定三类运动强度映射的说明性方面。运动强度映射基于比较相同奇偶性场及不同奇偶性场中的像素而指定每一帧的区为静态区、缓慢运动区及快速运动区。

[0168] 静态区

[0169] 确定运动映射的静态区可包含处理相邻场的邻域中的像素以确定某(些)像素的亮度差是否满足某种标准。在一些方面中,确定运动映射的静态区包含处理五个相邻场(当前场(C)、时间上在所述当前场之前的两个场及时间上在所述当前场之后的两个帧)的邻域中的像素以确定某(些)像素的亮度差是否满足某些阈值。所述五个场说明于图20中, Z^{-1} 表示一个场的延迟。换句话说,通常将以具有 Z^{-1} 的时间延迟的所述序列显示所述五个相邻场。

[0170] 图21说明根据一些方面识别所述五个场的每一者的某些像素的光圈,所述光圈可用于空间时间滤波。所述光圈包括(从左到右)再前一场(PP)、前一场(P)、当前场(C)、下一场(N)及再下一场(NN)的 3×3 像素群组。在一些方面中,如果当前场的区满足等式26到28中

所描述的标准,则在运动映射中认为当前场的所述区为静态区,图21中说明像素位置及对应场:

$$[0171] \quad |L_P - L_N| < T_1 \quad (26)$$

[0172] 且

$$[0173] \quad \left(\left| \frac{L_{BPP} - L_B}{2} \right| + \left| \frac{L_{EPP} - L_E}{2} \right| \right) < T_1 \text{(前向静态)} \quad (27)$$

[0174] 或

$$[0175] \quad \left(\left| \frac{L_{BNN} - L_B}{2} \right| + \left| \frac{L_{ENN} - L_E}{2} \right| \right) < T_1 \text{(后向静态)} \quad (28)$$

[0176] 其中 T_1 为阈值,

[0177] L_P 为位于P场中的像素P的亮度,

[0178] L_N 为位于N场中的像素N的亮度,

[0179] L_B 为位于当前场中的像素B的亮度,

[0180] L_E 为位于当前场中的像素E的亮度,

[0181] L_{BPP} 为位于PP场中的像素B_{PP}的亮度,

[0182] L_{EPP} 为位于PP场中的像素E_{PP}的亮度,

[0183] L_{BNN} 为位于NN场中的像素B_{NN}的亮度,且

[0184] L_{ENN} 为位于NN场中的像素E_{NN}的亮度。

[0185] 阈值 T_1 可经预定并设定为特定值,可通过除解交错以外的过程加以确定及提供(例如,作为用于正被解交错的视频的元数据),或阈值 T_1 可在解交错期间加以动态确定。

[0186] 由于至少两个原因,以上在等式26、27及28中所描述的静态区标准使用比常规解交错技术多的场。首先,与不同奇偶性场之间的比较相比,相同奇偶性场之间的比较具有较低的混叠及相位失配。然而,正被处理的场与其最相邻的相同奇偶性场近邻之间的最小时间差(因此,相关性)为两个场,比正被处理的场与其最相邻的不同奇偶性场近邻之间的时间差大。较可靠的不同奇偶性场与混叠较低的不同奇偶性场的组合可改进静态区检测的准确度。

[0187] 此外,所述五个场相对于当前帧C中的像素X可对称地分布于过去及将来,如图21中所展示。静态区可经细分成三类:前向静态(相对于先前帧为静态)、后向静态(相对于下一帧为静态),或双向(如果满足前向与后向标准两者)。对静态区的此较细分类可改进性能,尤其是在场景变化时及物体出现/消失时。

[0188] 缓慢运动区

[0189] 如果某些像素的亮度值不满足将被指定为静态区的标准但满足将被指定为缓慢运动区的标准,则在运动映射中可认为运动映射的一区为缓慢运动区。以下等式29定义可用于确定缓慢运动区的标准。参看图22,在以像素X为中心的光圈中展示等式29中所识别的像素I_a、I_c、J_a、J_c、K_a、K_c、L_a、L_c、P及N的位置。所述光圈包括当前场(C)的3×7像素邻域,以及下一场(N)、前一场(P)的3×5邻域。如果像素X不满足上述所列的用于静态区的标准且如果光圈中的像素满足等式29中所展示的以下标准,则认为像素X为缓慢运动区的一部分:

$$[0190] \quad (|L_{Ia} - L_{Ic}| + |L_{Ja} - L_{Jc}| + |L_{Ka} - L_{Kc}| + |L_{La} - L_{Lc}| + |L_P - L_N|) / 5 < T_2 \quad (29)$$

[0191] 其中 T_2 为阈值,且

[0192] L_{Ia} 、 L_{Ic} 、 L_{Ja} 、 L_{Jc} 、 L_{Ka} 、 L_{Kc} 、 L_{La} 、 L_{Lc} 、 L_P 、 L_N 分别为像素Ia、Ic、Ja、Jc、Ka、Kc、La、Lc、P及N的亮度值。

[0193] 阈值 T_2 还可经预定并设定为特定值,还可通过除解交错以外的过程加以确定及提供(例如,作为用于正被解交错的视频的元数据),或阈值 T_2 可在解交错期间加以动态确定。

[0194] 请注意,由于滤波器的边缘检测能力的角度,滤波器可使水平边缘模糊(例如,距垂直对准 45° 以上)。举例来说,图22中所说明的光圈(滤波器)的边缘检测能力受像素“A”与“F”或“C”与“D”所形成的角度影响。最佳将不并入比此角度更为水平的任何边缘且因此阶梯假影可出现在所述边缘处。在一些方面中,缓慢运动类可经划分成两个子类“水平边缘”及“其它”以计及此边缘检测效应。如果满足以下所展示的等式30中的标准,则可将缓慢运动像素分类为水平边缘,且如果不满足等式30中的标准,则可将缓慢运动像素分类为所谓的“其它”类。

$$[0195] \quad |(LA+LB+LC) - (LD+LE+LF)| < T_3 \quad (30)$$

[0196] 其中 T_3 为阈值,且LA、LB、LC、LD、LE及LF为像素A、B、C、D、E及F的亮度值。

[0197] 可对水平边缘及其它边缘中的每一者使用不同的内插方法。

[0198] 快速运动区

[0199] 如果不满足用于静态区的标准及用于缓慢运动区的标准,则可认为像素是在快速运动区中。

[0200] 在对选定帧中的像素分类之后,过程A(图26)接着进行到框2604且基于运动强度映射产生临时解交错帧。在此方面中, W_{med} 滤波器2054(图20)滤波选定场及必要的相邻场以提供候选全帧图像 F_0 ,所述候选全帧图像可经定义如下:

[0201]

$$F_0(\bar{x}, n) = \begin{cases} F(\bar{x}, n), & (y \bmod 2 = n \bmod 2) \\ \frac{1}{2}(F(\bar{x}, n-1) + F(\bar{x}, n+1)), & \text{(静态后向及前向)} \\ F(\bar{x}, n-1) & \text{(静态前向但非前向)} \\ F(\bar{x}, n+1) & \text{(静态后向但非后向)} \\ med(A, B, C, D, E, F), & \text{(缓慢运动w/o水平边缘)} \\ med\left(\alpha_0 \frac{A+F}{2}, \alpha_1 \frac{B+E}{2}, \alpha_2 \frac{C+D}{2}, \alpha_3 \frac{G+H}{2}\right), & \text{(缓慢运动w/水平边缘)} \\ \frac{B+E}{2}, & \text{(快速运动)} \end{cases} \quad (31)$$

[0202] 其中, α_i ($i=0, 1, 2, 3$)为整数权重,其经计算如下:

$$[0203] \quad \alpha_i = \begin{cases} 2 & \text{如果 } \beta_1 = \min\{\beta_0, \beta_1, \beta_2, \beta_3\} \\ 1, & \text{否则} \end{cases} \quad (32)$$

$$[0204] \quad \beta_0 = \frac{A+F}{|A-F|}, \quad \beta_1 = \frac{B+E}{|B-E|}, \quad \beta_2 = \frac{C+D}{|C-D|}, \quad \beta_3 = \frac{G+H}{|G-H|} \quad (33)$$

[0205] 提供经 W_{med} 滤波的临时解交错帧以用于与运动估计及运动补偿处理一起进行的进一步处理,如图20的下部分中所说明。

[0206] 如以上所描述且如等式31中所展示,静态内插包含场间内插且缓慢运动及快速运动内插包含场内内插。在其中不需要相同奇偶性场的时间(例如,场间)内插的某些方面中,可通过将阈值 T_1 (等式4到6)设定为零($T_1=0$)而“停用”时间内插。在时间内插被停用的情况下对当前场的处理会导致不将运动等级映射的任何区分类为静态区,且Wmed滤波器2054(图20)使用图22中的光圈中所说明的三个场,其对一当前场及两个相邻的非奇偶场进行操作。

[0207] 去噪声

[0208] 在某些方面中,去噪声器可用于在使用运动补偿信息进一步处理候选Wmed帧之前从候选Wmed帧移除噪声。去噪声器可移除存在于Wmed帧中的噪声且保留信号存在,而不管信号的频率内容如何。可使用各种类型的去噪声滤波器,包括小波滤波器。小波用于在空间与比例换算域中定位给定信号的一类函数。基于小波的基本想法为以不同比例(scale)或分辨率分析信号以使得小波表示中的小变化在最初的信号中产生对应的小变化。

[0209] 在一些方面中,去噪声滤波器基于(4,2)双正交立方B样条(spline)小波滤波器的方面。可通过以下正变换及逆变换来定义一此种滤波器:

$$[0210] \quad h(z) = \frac{3}{4} + \frac{1}{2}(z + z^{-1}) + \frac{1}{8}(z + z^{-2}) \quad (\text{正变换}) \quad (34)$$

[0211] 及

$$[0212] \quad g(z) = \frac{5}{4}z^{-1} - \frac{5}{32}(1 + z^{-2}) - \frac{3}{8}(z + z^{-3}) - \frac{3}{32}(z^2 + z^{-4}) \quad (\text{逆变换}) \quad (35)$$

[0213] 去噪声滤波器的应用可增加有噪声环境中的运动补偿的精确度。假定视频序列中的噪声为相加性白高斯。通过 σ 表示经估计的噪声变化。其可经估计为最高频率子带系数的绝对中位差除以0.6745。所述滤波器的实施方案进一步描述于D.L.杜诺夫(D.L.Donoho)及I.M.约翰斯顿(I.M.Johnstone)的“通过小波收缩进行的理想空间适应(Ideal spatial adaptation by wavelet shrinkage)”(生物统计学(Biometrika),第8卷,第425-455页,1994年)中,其以全文引用的方式并入本文中。

[0214] 小波收缩或小波维纳滤波器还可用作去噪声器。小波收缩去噪声可涉及小波变换域中的收缩,且通常包含三个步骤:线性正小波变换、非线性收缩去噪声及线性逆小波变换。维纳滤波器为MSE最佳线性滤波器,其可用于改进由于相加性噪声及模糊而降级的图像。所述滤波器通常是此项技术中所已知的且描述于(例如)以上所参考的“通过小波收缩进行的理想空间适应(Ideal spatial adaptation by wavelet shrinkage)”及S.P.戈亥尔(S.P.Ghael),A.M.萨伊德(A.M.Sayeed)及R.G.巴拉尼克(R.G.Baraniuk)所著的“经由试验维纳滤波进行的改进小波去噪声(Improvement Wavelet denoising via empirical Wiener filtering)”(SPIE学报(Proceedings of SPIE),第3169卷,第389-399页,圣地亚哥(San Diego),1997年7月)中。

[0215] 运动补偿

[0216] 参看图27,在框2702处,过程B执行双向运动估计,且接着在框104处使用运动估计来执行运动补偿,运动补偿进一步说明于图20中且描述于下文中的说明性方面中。在Wmed滤波器与基于运动补偿的解交错器之间存在一场“滞后(lag)”。从如图23中所展示的前一帧“P”与后一帧“N”中的信息预测用于当前场“C”的“缺少的”数据(像素数据的非最初行)的

运动补偿信息。在当前场(图23)中,实线表示最初的像素数据所存在的行且虚线表示经Wmed内插的像素数据所存在的行。在某些方面中,在4行乘8列的像素邻域中执行运动补偿。然而,此像素邻域为一用于解释目的的实例,且所属领域的技术人员应明白,可在基于包含不同数目个行与不同数目个列的像素邻域的其他方面中执行运动补偿,像素邻域的选择可基于许多因素,所述因素包括(例如)计算速度、可用处理功率或正被解交错的多媒体数据的特征。因为当前场仅具有所述行的一半,所以待匹配的四个行实际上对应于8像素×8像素的区。

[0217] 参看图20,双向ME/MC 2068可使用均方误差和(SSE),其可用于测量相对于Wmed下一帧2058及经解交错的当前帧2070的Wmed当前帧2060的一预测块与一被预测块之间的类似性。经运动补偿的当前帧2066的产生接着使用来自最类似的匹配块的像素信息来填充最初像素线之间的缺失数据。在一些方面中,双向ME/MC 2068偏置来自经解交错的先前帧2070信息的像素信息或给予所述像素信息更多权重,因为所述像素信息由运动补偿信息及Wmed信息产生,而仅通过空间时间滤波解交错Wmed下一帧2058。

[0218] 在一些方面中,为改进具有类似亮度区域但不同色度区域的场的区域中的匹配性能,可使用一度量,所述度量包括像素的一个或一个以上亮度群组(例如,一4行×8列的亮度块)及像素的一个或一个以上色度群组(例如,两个2行×4列的色度块U及V)的像素值的组成。所述方法有效地减少色彩敏感性区域处的失配。

[0219] 运动向量(MV)在垂直尺寸中具有1/2像素的粒度且在水平尺寸中具有1/2或1/4像素的粒度。可使用内插滤波器来获得分数像素(fractional-pixel)样本。举例来说,可用于获得半像素样本的一些滤波器包括双线性滤波器(1,1)、H.263/AVC所推荐的内插滤波器:(1,-5,20,20,-5,1),及六分支海明窗式(Hamming windowed)正弦函数滤波器(3,-21,147,147,-21,3)。可通过应用双线性滤波器从全像素及半像素样本产生1/4像素样本。

[0220] 在一些方面中,运动补偿可使用多种类型的搜索过程以匹配在当前帧的某一位置处的数据(例如,描绘一物体)与在另一帧(例如,下一帧或前一帧)中的不同位置处的对应数据,相应帧内的位置的差异指示所述物体的运动。举例来说,搜索过程使用可覆盖较大搜索区的全运动搜索或可使用较少像素的快速运动搜索,且/或用于搜索样式中的选定像素可具有特定形状(例如,菱形)。对于快速运动搜索,搜索区可以运动估计或运动候选者为f中心,运动估计或运动候选者可用作搜索相邻帧的起始点。在一些方面中,可从外部运动估计器产生MV候选者且将MV候选者提供到解交错器。来自先前经运动补偿的相邻帧中的对应邻域的宏块的运动向量还可用作运动估计。在一些方面中,可从搜索对应的先前帧及接下来的帧的宏块(例如,3宏块×3宏块)的一邻域产生MV候选者。

[0221] 图24说明如图23中所展示的可在运动估计/补偿期间通过搜索前一帧及下一帧的邻域而产生的两个MV映射 MV_P 及 MV_N 的一实例。在 MV_P 与 MV_N 中,待处理以确定运动信息的块是通过“X”表示的中心块。在 MV_P 与 MV_N 中,存在可在正被处理的当前块X的运动估计期间使用的九个MV候选者。在此实例中,所述MV候选者中的四个存在于来自先前所执行的运动搜索的同一场中且通过 MV_P 与 MV_N 中颜色较淡的块(图24)加以描绘。通过颜色较深的块描绘的五个其它MV候选者是从先前经处理的帧的运动信息(或映射)所复制。

[0222] 在完成运动估计/补偿之后,针对缺少的行(通过图23中的虚线所表示)可产生两个内插结果:由Wmed滤波器(图20中的Wmed当前帧2060)产生的一内插结果及由运动补偿器

(MC当前帧2066)的运动估计处理产生的一内插结果。组合器2062通常通过使用Wmed当前帧2060及MC当前帧2066的至少一部分来合并Wmed当前帧2060与MC当前帧2066以产生当前经解交错的帧2064。然而,在某些条件下,组合器2062可仅使用当前帧2060或MC当前帧2066中的一者来产生当前经解交错的帧。在一实例中,组合器2062合并Wmed当前帧2060与MC当前帧2066以产生如等式36中所展示经解交错的输出信号:

$$[0223] \quad F_o(\bar{x}, n) = \begin{cases} F(\bar{x}, n), & (y \bmod 2 = n \bmod 2) \\ k_2 F_{Wmed}(\bar{x}, n) + (1 - k_2) F_{MC}(\bar{x} - \vec{D}, n - 1), & (\text{否则}) \end{cases} \quad (36)$$

[0224] 其中 $F(\bar{x}, n)$ 用于位置 $x = (x, y)^t$ 处的场 n_1 中的亮度值,其中 t 用于转置。使用定义如下的截割函数

$$[0225] \quad \text{截割}(0, 1, a) = 0, \text{如果}(a < 0); 1, \text{如果}(a > 1); a, \text{否则} \quad (37)$$

[0226] k_1 可经计算为:

$$[0227] \quad k_1 = \text{截割}\left(0, C_1 \sqrt{\text{Diff}}\right) \quad (38)$$

[0228] 其中 C_1 为稳健性参数,且Diff为预测帧像素与经预测的帧中的可用像素(从现有场取得)之间的亮度差。通过适当选择 C_1 ,有可能调谐均方误差的相对重要性。可如等式39中所展示计算 k_2 :

$$[0229] \quad k_2 = 1 - \text{截割}\left(0, 1, (1 - k_1) \frac{|F_{Wmed}(\bar{x} - \vec{y}_u, n) - F_{MC}(\bar{x} - \vec{y}_u - \vec{D}, n - 1)| + \delta}{|F_{Wmed}(\bar{x}, n) - F_{MC}(\bar{x} - \vec{D}, n - 1)| + \delta}\right) \quad (39)$$

[0230] 其中 $\bar{x} = (x, y)$ 、 $\vec{y}_u = (0, 1)$ 、 \vec{L} 为运动向量, δ 为用来防止被除以零的小常数。使用截割函数(clipping function)来滤波的解交错进一步描述于G.D.汉(G.D.Haan)及E.B.贝勒斯(E.B.Bellers)的“视频数据的解交错(De-interlacing of video data)”(电气电子工程师协会关于消费电子产品的交易(IEEE Transactions on Consumer Electronics),第43卷,第3期,第819-825页,1997年)中,其以全文引用的方式并入本文中。

[0231] 在一些方面中,组合器2062可经配置以尝试并维护以下等式以实现高PSNR及稳健结果:

$$[0232] \quad |F_o(\bar{x}, n) - F_{Wmed}(\bar{x}, n)| = |F_o(\bar{x} - \vec{y}_u, n) - F_{Wmed}(\bar{x} - \vec{y}_u, n)| \quad (40)$$

[0233] 有可能用Wmed+MC解交错方案来解耦合包含场间内插及场内内插的解交错预测方案。换句话说,空间时间Wmed滤波可主要用于场内内插目的,而可在运动补偿期间执行场间内插。此减少了Wmed结果的峰值信噪比,但在应用运动补偿之后视觉质量更令人满意,因为来自不准确场间预测模式决策的不良像素将从Wmed滤波过程被移除。

[0234] 色度处理可与经布置的亮度处理一致。根据运动映射产生,通过观察色度像素的四个经布置的亮度像素的运动等级而获得所述色度像素的运动等级。所述操作可基于表决(色度运动等级借用主要的亮度运动等级)。然而,我们建议使用如下保守方法。如果四个亮度像素的任一者具有快速运动等级,则色度运动等级应为快速运动;否则,如果四个亮度像素的任一者具有缓慢运动等级,则色度运动等级将为缓慢运动;否则色度运动等级为静态的。所述保守方法可能无法实现最高PSNR,但无论色度运动等级中是否存在歧义,所述保守

方法均避免使用INTER预测的风险。

[0235] 使用所描述的经单独描述的Wmed算法及本文所描述的组合的Wmed与运动补偿算法来解交错多媒体数据序列。还使用像素掺混(或平均)算法及“非解交错”状况(其中,仅组合场而不进行任何内插或掺混)来解交错相同的多媒体数据序列。分析所得帧以确定PSNR且在以下表中展示PSNR:

[0236]

PSNR (dB) 序列	无解交错	掺混	Wmed	Wmed+MC
足球	8.955194	11.38215	19.26221	19.50528
城市	11.64183	12.93981	15.03303	15.09859
组员	13.32435	15.66387	22.36501	22.58777

[0237] 即使通过使用MC加上Wmed的解交错仅能改进边缘PSNR,通过组合Wmed与MC内插结果所产生的解交错图像的视觉质量仍由于以上所提及的原因而在视觉上更令人满意,组合Wmed结果与MC结果会抑制偶数场与奇数场之间的混叠及噪声。

[0238] 在一些重取样方面中,多相位重取样器经实施用于图像大小重新设定。在下取样的一实例中,最初的图像与经调整大小的图像之间的比可为 p/q ,其中 p 及 q 是互为质数的整数。相位的总数目为 p 。对于约为0.5的调整大小因数,多相位滤波器的截止频率在一些方面中为0.6。截止频率不完全匹配调整大小比,以便提高经调整大小的序列的高频响应。此不可避免地允许一些混叠。然而,众所周知,与模糊且无混叠的图像相比,人眼更喜欢清晰但有一点混叠的图像。

[0239] 图42说明多相位重取样的一实例,其展示调整大小比为 $3/4$ 时的相位。图42中所说明的截止频率也为 $3/4$ 。在以上图42中用垂直轴说明最初的像素。还将正弦函数(sinc function)绘制为以所述轴为中心以表示滤波器波形。因为我们选择截止频率与重取样比完全相同,所以正弦函数的零在重新调整像素大小之后与像素的位置重叠,在图42中用十字形线说明所述重叠。为在调整大小之后得到像素值,可如以下等式中所展示从最初的像素合计组成:

$$[0240] \quad v(x) = \sum_{i=-\infty}^{\infty} u(i) \times \text{sinc}(\pi f_c (i-x)) \quad (41)$$

[0241] 其中 f_c 为截止频率。以上1-D多相位滤波器可适用于水平尺寸与垂直尺寸。

[0242] 重取样(调整大小)的另一方面计及过扫描。在NTSC电视信号中,一图像具有486条扫描线,且在数字视频中,每一条扫描线上可具有720个像素。然而,由于所述大小与屏幕格式之间的失配,并非所有完整图像均可在电视上看见。图像的不可见部分被称为过扫描。

[0243] 为帮助广播台将有用信息置于对尽可能多的电视可见的区中,电影及电视工程师协会(SMPTE)定义了被称为安全动作区及安全标题区的动作帧的特定大小。见SMPTE推荐的关于用于电视系统的安全动作区及安全标题区测试样式的说明书的实践RP 27.3-1989 (practice RP 27.3~1989on Specifications for Safe Action and Safe Title Areas Test Pattern for Television Systems)。安全动作区被SMPTE定义为“所有显著动作必须发生”的区。安全标题区被定义为“可限制所有有用信息以确保在大多数家用电视接收器上

的可见性”的区。举例来说,如图43中所说明,安全动作区4310占据屏幕的中心90%,周围给出5%的边界。安全标题区4305占据屏幕的中心80%,给出10%的边界。

[0244] 现参看图44,因为安全标题区如此小,所以为了在图像中添加更多内容,一些广播将在安全动作区中包括文本,所述文本在白色矩形窗4415内部。通常在过扫描中可看见黑色边界。举例来说,在图44中,黑色边界出现在图像的上侧4420及下侧4425。可在过扫描中移除所述黑色边界,因为H.264视频在运动估计中使用边界延伸。经延伸的黑色边界可增加残余。我们可适当地将边界消减2%,且接着进行调整大小。可因此产生用于调整大小的滤波器。在多相位下取样之前执行截断以移除过扫描。

[0245] 去块/去环

[0246] 在去块处理的一实例中,可对一帧的所有 4×4 个块边缘(所述帧的边界处的边缘及被停用去块滤波器过程的任何边缘除外)应用去块滤波器。将在完成帧构造过程之后以宏块为基础执行此滤波过程,一帧中的所有宏块是以递增的宏块地址的次序予以处理。对于每一宏块,首先从左到右滤波垂直边缘,且接着从顶到底滤波水平边缘。对于水平方向且对于垂直方向,在四个16样本边缘上执行亮度去块滤波过程且在两个8样本边缘上执行用于每一色度分量的去块滤波过程,如图39中所展示。可能已经通过对先前宏块的去块过程操作予以修改的当前宏块的上方及到左方的样本值将用作对当前宏块的去块滤波过程的输入且可在对当前宏块的滤波期间进一步予以修改。在对垂直边缘的滤波期间加以修改的样本值可作用于同一宏块的水平边缘的滤波的输入。对于亮度及色度分量可单独调用去块过程。

[0247] 在去环处理的一实例中,可自适应地应用2-D滤波器以使边缘附近的区平滑。边缘像素经历很少滤波或不经历滤波以便避免模糊。

[0248] GOP分割器

[0249] 以下描述包括可包括于GOP分割器中的带宽映射产生、分镜头检测及自适应GOP分割的处理的说明性实例。

[0250] 带宽映射产生

[0251] 人类视觉质量 V 可为编码复杂性 C 与经分配的位 B (还称作带宽)的函数。图29是说明此关系的图表。请注意,从人类视觉观点看,编码复杂性度量 C 考虑空间及时间频率。因为人眼对失真更敏感,所以复杂性值相应较高。通常可假定: V 在 C 中单调递减,且在 B 中单调递增。

[0252] 为实现恒定的视觉质量,将带宽(B_i)指派给待编码的第 i 个物体(帧或MB),所述带宽(B_i)满足正下方的两个等式中所表示的标准:

$$[0253] \quad B_i = B(C_i, V) \quad (42)$$

$$[0254] \quad B = \sum_i B_i \quad (43)$$

[0255] 在正上方的两个等式中, C_i 为第 i 个物体的编码复杂性, B 为总的可用带宽,且 V 为对于一物体所实现的视觉质量。

[0256] 人类视觉质量难以用等式表示。因此,未精确定义以上等式组。然而,如果假定3-D模型在所有变量中是连续的,则可认为带宽比(B_i/B)在(C, V)对的邻域内不变。在以下所展示的等式中定义带宽比 β_i :

[0257] $\beta_i = B_i/B$ (44)

[0258] 可接着如以下等式中所表示定义位分配:

[0259] $\beta_i = \beta(C_i)$

[0260] $1 = \sum_i \beta_i$ (45)

[0261] 其中

[0262] $(C_i, V) \in \delta(C_0, V_0)$

[0263] 其中 δ 指示“邻域”。

[0264] 编码复杂性在空间与时间上受人类视觉敏感性的影响。吉洛德 (Girod) 的人类视觉模型是可用于定义空间复杂性的模型的一实例。此模型考虑局部空间频率及环境照明。所得度量被称为 D_{csat} 。在所述过程中的预处理点处,不知道图像将经帧内编码还是经帧间编码且产生用于两者的带宽比。根据不同视频物体的 β_{INTRA} 之间的比而分配位。对于经帧内编码的图像,带宽比表示于以下等式中:

[0265] $\beta_{INTRA} = \beta_{0INTRA} \log_{10}(1 + \alpha_{INTRA} Y^2 D_{csat})$ (46)

[0266] 在以上等式中, Y 为宏块的平均亮度分量, α_{INTRA} 为用于亮度平方及其后的 D_{csat} 项的加权因数, β_{0INTRA} 为用于保证 $1 = \sum_i \beta_i$ 的正规化因数。举例来说, $\alpha_{INTRA} = 4$ 的值实现优良视觉质量。内容信息(例如,内容分类)可用于将 α_{INTRA} 设定为一值,所述值对应于视频的特定内容的所要的优良视觉质量等级。在一实例中,如果视频内容包含“发言者头部(talking head)”新闻广播,则因为可能认为所述视频的信息图像或可显示的部分不如音频部分重要,所以可将视觉质量等级设定为较低,且可分配较少位用来编码数据。在另一实例中,如果视频内容包含体育运动事件,则因为对于观察者来说所显示的图像更为重要,所以内容信息可用于将 α_{INTRA} 设定为对应于较高视觉质量等级的值,且因此可分配更多位用来编码数据。

[0267] 为理解此关系,请注意,带宽是以对数方式分配有编码复杂性。亮度平方项 Y^2 反映具有较大量值的系数使用较多位来编码的事实。为防止对数得到负值,添加一到括号中的项。还可使用具有其它基数的对数。

[0268] 时间复杂性是通过对帧差异度量的测量加以确定,所述测量在考虑运动量(例如,运动向量)以及例如绝对差的和(SAD)的帧差异度量的情况下测量两个连续帧之间的差异。

[0269] 用于经帧间编码的图像的位分配可考虑空间复杂性以及时间复杂性。此表示于以下:

[0270] $\beta_{INTER} = \beta_{0INTER} \log_{10}(1 + \alpha_{INTER} \cdot SSD \cdot D_{csat} \exp(-\gamma ||MV_P + MV_N||^2))$ (47)

[0271] 在以上等式中, MV_P 及 MV_N 为用于当前MB之前向运动向量及后向运动向量。可注意到,经帧内编码的带宽公式中的 Y^2 由平方差的和(SSD)所替换。为理解以上等式中 $||MV_P + MV_N||^2$ 的作用,注意人类视觉系统的以下特征:经历平滑、可预测的运动(小的 $||MV_P + MV_N||^2$)的区吸引注意力且可由眼睛跟踪且通常不可忍受比静态区域还多的失真。然而,经历快速或不可预测的运动(大的 $||MV_P + MV_N||^2$)的区不可被跟踪且可忍受显著量化。实验表明: $\alpha_{INTER} = 1$ 、 $\gamma = 0.001$ 实现优良视觉质量。

[0272] 分镜头检测

[0273] 以下描述分镜头检测的一说明性实例。所述组件及过程可包括于GOP分割器412(图4)中。

[0274] 运动补偿器23可经配置以确定关于视频中的帧的双向运动信息。运动补偿器23还可经配置以确定一个或一个以上差异度量,例如绝对差的和(SAD)或平方差的和(SSD),且计算包括用于一个或一个以上帧的亮度信息(例如,宏块(MB)亮度平均值或差异)、一亮度直方图差异及一帧差异度量的其它信息,参考等式1到3描述所述其它信息的实例。镜头分类器可经配置以使用由运动补偿器所确定的信息将视频中的帧分类成两类或两类以上的“镜头”。编码器经配置以基于所述镜头分类而自适应地编码多个帧。以下参考等式1到10描述运动补偿器、镜头分类器及编码器。

[0275] 图28为根据一些方面的预处理器202的框图,所述预处理器202包含经配置以用于分镜头检测及其它预处理操作的处理器2831。可由在如图4中所展示的预处理器202外部的源提供数字视频源且将数字视频源传送到预处理器202中的通信模块2836。预处理器202含有与处理器2831通信的存储媒体2825,处理器2831与存储媒体2825均与通信模块2836通信。处理器2831包括可操作以如本文中所描述般产生运动信息、分类视频数据的帧中的镜头并执行其它预处理测试的运动补偿器2032、镜头分类器2833及用于预处理的其它模块2034。运动补偿器、镜头分类器及其它模块可含有类似于图4中的对应模块的过程,且可处理视频以确定以下所描述的信息。具体来说,处理器2831可具有一配置以:获得指示多个视频帧的相邻帧之间的差异的度量(所述度量包含双向运动信息及亮度信息),基于所述度量而确定所述多个视频帧中的镜头变化,并基于所述镜头变化而自适应编码所述多个帧。在一些方面中,可通过在处理器2831外部的装置或过程来计算所述度量,所述装置或过程还可在预处理器202外部且直接或间接经由另一装置或存储器与处理器2831通信。还可由处理器2831计算所述度量,例如,由运动补偿器2832计算所述度量。

[0276] 预处理器202将用于进一步处理、编码及传输的视频及元数据提供到其它装置,例如,终端6(图1)。经编码的视频在一些方面中可为可包含基本层及增强层的可缩放的经多层编码的视频。可缩放的层编码进一步描述于标题为“具有两层编码和单层编码的可缩放视频编码(Scalable Video Coding With Two Layer Encoding And Single Layer Decoding)”的同在申请中的美国专利申请案[代理人案号为050078]中,所述申请案归本发明的受让人所拥有且其以全文引用的方式并入本文中。

[0277] 结合图28及本文所揭示的其它实例及图式所描述的各种说明性逻辑块、组件、模块及电路在一些方面中可用以下装置加以实施或执行:通用处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)或其它可编程逻辑装置、离散门或晶体管逻辑、离散硬件组件或其经设计以执行本文所描述的功能的任何组合。例如图28中所展示的处理器的通用处理器可为微处理器,但在替代实施例中,处理器可为任何常规处理器、控制器、微控制器或状态机。处理器还可实施为计算装置的组合,例如,DSP与微处理器的组合、多个微处理器的组合、一个或一个以上微处理器与DSP核心的联合,或任何其它此种配置。

[0278] 视频编码通常对经构造的图像群组(GOP)进行操作。GOP通常由经帧内编码的帧(I帧)开始,其后是一系列P(预测)或B(双向)帧。通常,I帧可存储用于显示所述帧的所有数据,B帧依赖于前一帧及后一帧中的数据(例如,仅含有从前一帧变化的数据或不同于下一

帧中的数据),且P帧含有已从前一帧变化的数据。

[0279] 在常见使用中,在经编码的视频中,I帧散布有P帧及B帧。就大小(例如,用于编码所述帧的位的数目)来说,I帧通常比P帧大得多,P帧又比B帧大。对于有效率的编码、传输及解码处理,GOP的长度应足够长而能够减少来自大的I帧的有效损失,且应足够短而能够阻止编码器与解码器之间的失配,或信道损害。此外,由于相同的原因,P帧中的宏块(MB)可经帧内编码。

[0280] 场景变化检测可用于视频编码器以确定适当的GOP长度且基于GOP长度插入I帧,而不是以固定间隔插入I帧。在实际串流视频系统中,通信信道通常由于位误差或封包损失而受损害。将I帧或I MB置于哪里可显著影响解码视频质量及观察体验。一种编码方案是对于具有从经布置的先前图像或图像部分的显著改变的图像或图像部分使用经帧内编码的帧。通常不可用运动估计有效且有效率地预测所述区域,且如果所述区域不受帧间编码技术(例如,使用B帧及P帧的编码)作用,则可更有效率地进行编码。在信道损害的情况下,所述区域可能遭受误差传播,可通过帧内编码减少或消除(或几乎如此)所述误差传播。

[0281] 可将GOP视频的部分分类成两类或两类以上,其中每一区域可具有不同的帧内编码标准,其可取决于特定实施方案。作为一实例,可将视频分类成三类:突然的场景变化,交叉衰落及其它缓慢的场景变化,及相机闪光灯。突然的场景变化包括通常由相机操作造成的显著不同于先前帧的帧。因为所述帧的内容不同于先前帧的内容,所以突然的场景变化帧应编码为I帧。交叉衰落及其它缓慢的场景变化包括通常由相机镜头的计算机处理造成的场景的缓慢切换。两种不同场景的逐渐掺混在人眼看来可能更令人满意,但其提出对视频编码的挑战。运动补偿不可有效地减少所述帧的位速率,且对于所述帧可更新更多的内部MB。

[0282] 当帧的内容包括相机闪光时,相机闪光灯或相机闪光事件出现。所述闪光在持续时间上相对短(例如,一帧)且极明亮,以致一描绘所述闪光的帧中的像素相对于相邻帧上的对应区展示出通常高的亮度。相机闪光灯突然且迅速地改变图像的亮度。相机闪光灯的持续时间通常比人类视觉系统(HVS)的暂时遮蔽持续时间短,人类视觉系统(HVS)的暂时遮蔽持续时间通常被定义为44ms。人眼对所述短的明亮度丛发的质量不敏感,且因此可对其进行粗略编码。因为不可用运动补偿有效地处理闪光灯帧且闪光灯帧是将来帧的不良的预测候选者,所以所述帧的粗略编码不会减少将来帧的编码效率。由于“人工的”高亮度,经分类为闪光灯的场不应用于预测其它帧,且由于同一原因,其它帧不可有效地用于预测所述帧。一旦识别所述帧,便可取出所述帧,因为所述帧需要相对高的处理量。一选项为移除相机闪光灯帧且在相机闪光灯帧的位置中编码DC系数;此种解决方法是简单的,在计算上是快速的且节省许多位。

[0283] 当检测到以上帧的任一者时,宣告镜头事件。分镜头检测不仅有助于改进编码质量,而且还可帮助识别视频内容搜索及编索引。在下文中描述场景检测过程的一方面。

[0284] 图30说明对GOP进行操作且在一些方面中可用于基于视频帧中的分镜头检测而编码视频的过程3000,其中过程3000的部分(或子过程)参看图30到图40予以描述及说明。处理器2831可经配置以并有过程3000。在过程3000开始之后,过程3000进行到框3042,在框3042中,获得用于视频帧的度量(信息),所述度量包括指示相邻帧之间的差异的信息。所述度量包括随后用于确定出现在相邻帧之间的变化的双向运动信息及基于亮度的信息,所述

信息可用于镜头分类。可从另一装置或过程获得所述度量,或由(例如)处理器2831计算所述度量。参考图31中的过程A描述度量产生的说明性实例。

[0285] 过程3000接着进行到框3044,在框3044中,基于所述度量确定视频中的镜头变化。可将视频帧分类成两类或两类以上的包含于帧中的镜头类型,例如,突然的场景变化、缓慢变化的场景或含有高亮度值的场景(相机闪光)。某些实施方案编码可能需要其它类。参考图32中的过程B且更详细参考图34到图36中的过程D、E及F描述镜头分类的一说明性实例。

[0286] 一旦对帧进行了分类,过程3000便进行到框3046,在框3046中,可使用镜头分类结果来编码帧或为编码指定帧。所述结果可影响是用经帧内编码的帧来编码帧还是用预测帧(例如,P帧或B帧)来编码帧。图33中的过程C展示使用镜头结果的编码方案的一实例。

[0287] 图31说明用于获得视频度量的过程的一实例。图31说明出现于图30的框3042中的某些步骤。仍参看图31,在框3152中,过程A获得或确定视频的双向运动估计及补偿信息。图28的运动补偿器2832可经配置以对帧执行双向运动估计且确定可用于随后的镜头分类的运动补偿信息。过程A接着进行到框3154,在框3154中,过程A产生包括用于当前或选定帧与一个或一个以上相邻帧的亮度差异直方图的亮度信息。最后,过程A接着继续到框3156,在框3156中,计算指示帧中所含有的镜头的度量。一此种度量为在两个实例中于等式4及10中展示的帧差异度量。以下描述确定运动信息、亮度信息及帧差异度量的说明性实例。

[0288] 运动补偿

[0289] 为执行双向运动估计/补偿,可用双向运动补偿器预处理一视频序列,所述双向运动补偿器将当前帧的每一 8×8 块与最相邻的邻近帧中的两个帧(一个在过去,且一个在将来)中的块匹配。运动补偿器产生每一块的运动向量及差异度量。图37说明此概念,图37展示将当前帧C的像素与过去帧P及将来的(或下一)帧N的像素匹配的一实例,且图37描绘到经匹配的像素的运动向量(过去运动向量 MV_P 及将来运动向量 MV_N)。以下是对双向运动向量产生及有关编码的一说明性方面的简短描述。

[0290] 图40说明运动向量确定过程及(例如)MPEG-4中的预测帧编码的一实例。图40中所描述的过程是对可在图31的框3152中发生的实例过程的更详细说明。在图40中,当前图像4034由 5×5 个宏块组成,其中此实例中的宏块的数目是任意的。一宏块由 16×16 个像素组成。可通过8位亮度值(Y)及两个8位色度值(Cr及Cb)来定义像素。

[0291] 在MPEG中,可以4:2:0格式存储Y、Cr及Cb分量,其中在X及Y方向中以因数2对Cr及Cb分量进行下取样。因此,每一宏块将由256个Y分量、64个Cr分量及64个Cb分量组成。在不同于当前图像4034的时间点处,从参考图像4032预测当前图像4034的宏块4036。在参考图像4032中进行搜索以定位在Y、Cr及Cb值上最接近正被编码的当前宏块4036的最佳匹配宏块4038。在运动向量4040中编码最佳匹配宏块4038在参考图像4032中的位置。参考图像4032可以是解码器在构造当前图像4034之前便已重构的I帧或P帧。从当前宏块4036中减去最佳匹配宏块4038(计算Y、Cr及Cb分量的每一者的差异),从而产生残余误差4042。用二维(2D)离散余弦变换(DCT)4044编码残余误差4042且接着将其量化4046。可执行量化4046以通过(例如)向高频系数分配较少的位而向低频系数分配较多的位而提供空间压缩。残余误差4042的经量化的系数以及运动向量4040及参考图像4034识别信息是表示当前宏块4036的编码信息。编码信息可存储于存储器中以供将来使用或出于(例如)误差校正或图像增强的目的加以操作,或在网络140上加以传输。

[0292] 残余误差4042的经编码经量化的系数以及经编码的运动向量4040可用于在编码器中重构当前宏块4036以用作用于随后运动估计及补偿的参考帧的一部分。编码器可仿效用于此P帧重构的解码器的程序。仿效解码器将导致编码器与解码器以相同参考图像工作。此处呈现重构过程,无论所述重构过程在编码器中进行以用于进一步帧间编码或者是在解码器中进行。可在重构参考帧(或正被参考的图像或帧的一部分)之后开始重构P帧。解量化4050经编码经量化的系数且接着执行二维离散余弦逆变换DCT或IDCT 4052,从而产生经解码或重构的残余误差4054。经编码的运动向量4040经解码且用于在已经重构的参考图像4032中定位已经重构的最佳匹配宏块4056。接着将经重构的残余误差4054添加到经重构的最佳匹配宏块4056以形成经重构的宏块4058。经重构的宏块4058可存储于存储器中,独立地或与其它经重构的宏块一起显示于一图像中,或经进一步处理以用于图像增强。

[0293] 使用B帧的编码(或用双向预测予以编码的任何区段)可利用当前图像中的一区域与前一图像中的一最佳匹配预测区域及后一图像中的一最佳匹配预测区域之间的时间冗余。组合后一最佳匹配预测区域与前一最佳匹配预测区域以形成组合的双向预测区域。当前图像区域与最佳匹配的组的双向预测区域之间的差异为残余误差(或预测误差)。可在两个运动向量中编码最佳匹配预测区域在后一参考图像中的位置及最佳匹配预测区域在前一参考图像中的位置。

[0294] 亮度直方图差异

[0295] 运动补偿器可产生每一块的差异度量。所述差异度量可为平方差的和(SSD)或绝对差的和(SAD)。在不损失一般性的情况下,此处SAD用作一实例。

[0296] 对于每一帧,将SAD比计算如下:

$$[0297] \quad \gamma = \frac{\varepsilon + SAD_P}{\varepsilon + SAD_N} \quad (48)$$

[0298] 其中SAD_P及SAD_N分别为前向及后向差异度量的绝对差的和。请注意,分母含有一小的正数 ε 以防止“被除以零”的误差。分子还含有 ε 以平衡分母中一的效应。举例来说,如果前一帧、当前帧及下一帧是相同的,则运动搜索应产生SAD_P=SAD_N=0。在此状况下,以上计算产生 $\gamma = 1$ 而不是0或无限。

[0299] 可为每一帧计算一亮度直方图。多媒体图像通常具有8位的亮度深度(例如,“区间(bin)”的数目)。根据一些方面用于计算亮度直方图的亮度深度可经设定为16以获得直方图。在其它方面中,亮度深度可经设定为适当数目,所述适当数目可取决于正被处理的数据的类型、可用的计算功率或其它预定标准。在一些方面中,可基于计算出或接收到的度量(例如,数据的内容)而动态设定亮度深度。

[0300] 等式49说明计算亮度直方图差异(希腊字母 λ)的一实例:

$$[0301] \quad \lambda = \frac{\sum_{i=1}^{16} |N_{Pi} - N_{Ci}|}{N} \quad (49)$$

[0302] 其中N_{Pi}为用于前一帧的第i个区间中的块的数目,且N_{Ci}为用于当前帧的第i个区间中的块的数目,且N为一帧中的块的总数目。如果前一帧与当前帧的亮度直方图差异完全不同(或不相交),则 $\lambda = 2$ 。

[0303] 可如等式50中所展示计算参考图5的块56所论述的帧差异度量D:

$$[0304] \quad D = \frac{\gamma_C}{\gamma_P} + A\lambda(2\lambda + 1) \quad (50)$$

[0305] 其中A为根据应用所选择的常数,且 $\gamma_C = \frac{\varepsilon + SAD_P}{\varepsilon + SAD_N}$,且 $\gamma_P = \frac{\varepsilon + SAD_{PP}}{\varepsilon + SAD_C}$ 。

[0306] 图32说明使用对于视频所获得或确定的度量来确定三类镜头(或场景)变化的过程B的一实例。图32说明在图30的框3044的一方面中出现的某些步骤。再参看图32,在框3262中,过程B首先确定帧是否满足将被指定为突然的场景变化的标准。图34中的过程D说明此确定的一实例。过程B接着进行到框3264,在框3264中,确定所述帧是否为缓慢变化的场景的一部分。图35中的过程C说明确定缓慢变化的场景的一实例。最后,在框3266处,过程B确定帧是否含有相机闪光(换句话说,不同于前一帧的大的亮度值)。图36中的过程F说明确定含有相机闪光的帧的一实例。以下描述所述过程的一说明性实例。

[0307] 突然的场景变化

[0308] 图34为说明确定突然的场景变化的过程的流程图。图34进一步详细阐述可在图32的框3262的一些方面中出现的某些步骤。在框3482处检查帧差异度量D是否满足等式51中所展示的标准:

$$[0309] \quad D = \frac{\gamma_C}{\gamma_P} + A\lambda(2\lambda + 1) \geq T_1 \quad (51)$$

[0310] 其中A为根据应用所选择的常数,且 T_1 为阈值。如果满足所述标准,则在框3484处,过程D指定所述帧为突然的场景变化,且在此实例中,不需要任何其它的镜头分类。

[0311] 在一实例中,模拟展示设定 $A=1$ 且 $T_1=5$ 实现优良检测性能。如果当前帧为突然的场景变化帧,则 γ_C 应大且 γ_P 应小。可使用比 $\frac{\gamma_C}{\gamma_P}$ 而不单独使用 γ_C 以使得度量经正规化为上下文的活动等级。

[0312] 请注意,以上标准以非线性方法使用亮度直方图差异(λ)。图39说明 $\lambda*(2\lambda+1)$ 为凸函数。当 λ 小(例如,接近于零)时,其仅仅为预强调。 λ 变大时,通过所述函数进行更多强调。在此预强调的情况下,对于大于1.4的任何 λ ,如果阈值 T_1 经设定为5,则检测到突然的场景变化。

[0313] 交叉衰落及缓慢的场景变化

[0314] 图35进一步说明可在图32的步骤3264中出现的一些方面的其它细节。参看图35,在框3592处,过程E确定帧是否为描绘缓慢的场景变化的一系列帧的一部分。如果帧差异度量D小于第一阈值 T_1 且大于或等于第二阈值 T_2 (如等式52中所说明),则过程E确定当前帧为交叉衰落或其它缓慢的场景变化:

$$[0315] \quad T_2 \leq D < T_1 \quad (52)$$

[0316] 对于某数目个连续帧,其中 T_1 为以上所用的同一阈值且 T_2 为另一阈值。由于实施方案中可能的差异, T_1 及 T_2 的确切值通常通过正常实验加以确定。如果满足标准,则在框94处,过程E将帧分类为用于选定帧末端的缓慢变化的场景镜头分类的一部分。

[0317] 相机闪光灯事件

[0318] 图36中所展示的过程F为可确定当前帧是否包含相机闪光灯的过程的一实例。在

此说明性方面相机中,亮度直方图统计用于确定当前帧是否包含相机闪光灯。如框3602处所展示,过程F通过首先确定当前帧的亮度是否大于前一帧的亮度及下一帧的亮度来确定相机闪光事件是在选定帧中。如果答案为否,则帧不是相机闪光事件;但如果答案为是,则帧可能为相机闪光事件。在框3604处,过程F确定后向差异度量是否大于阈值 T_3 ,且前向差异度量是否大于阈值 T_4 ;如果两个所述条件均得到满足,则在框3606处,过程F将当前帧分类为具有相机闪光灯。在一实例中,在框3602处,过程F确定当前帧的平均亮度减去前一帧的平均亮度是否等于或超过阈值 T_3 ,且过程F确定当前帧的平均亮度减去下一帧的平均亮度是否大于或等于阈值 T_3 ,如等式53及54中所展示:

$$[0319] \quad \bar{Y}_C - \bar{Y}_P \geq T_3 \quad (53)$$

$$[0320] \quad \bar{Y}_C - \bar{Y}_N \geq T_3 \quad (54)$$

[0321] 如果不满足标准,则不将当前帧分类为包含相机闪光灯且过程F返回。如果满足标准,则过程F进行到框3604,在框3604中,确定后向差异度量 SAD_P 及前向差异度量 SAD_N 是否大于特定阈值 T_4 ,如以下等式55及56中所说明:

$$[0322] \quad SAD_P \geq T_4 \quad (55)$$

$$[0323] \quad SAD_N \geq T_4 \quad (56)$$

[0324] 其中 \bar{Y}_C 为当前帧的平均亮度, \bar{Y}_P 为前一帧的平均亮度, \bar{Y}_N 为下一帧的平均亮度,且 SAD_P 及 SAD_N 为与当前帧相关联的前向及后向差异度量。如果不满足标准,则过程F返回。

[0325] 因为实施所描述的过程可导致包括阈值的操作参数的差异,所以 T_3 值通常通过正常实验加以确定。因为相机闪光通常仅进行一个帧,所以 SAD 值包括于确定中,且由于亮度差异,不可能使用运动补偿从前向与后向方向良好地预测此帧。

[0326] 在一些方面中,阈值 T_1 、 T_2 、 T_3 及 T_4 中的一者或一者以上经预定且所述值经并入于编码装置中的镜头分类器中。通常经由分镜头检测的一特定实施方案的测试来选择所述阈值。在一些方面中,可基于供应到镜头分类器的使用信息(例如,元数据)或基于由镜头分类器本身计算出的信息在处理期间(例如,动态地)设定阈值 T_1 、 T_2 、 T_3 及 T_4 中的一者或一者以上。

[0327] 现参看图33,图33展示用于基于选定帧的镜头分类而确定用于视频或用于编码所述视频的编码参数的过程C。在框3370处,过程C确定选定帧是否经分类为突然的场景变化。如果答案为是,则在框3371处,将当前帧分类为突然的场景变化,且可将帧编码为I帧并可确定GOP边界。如果答案为否,则过程C进行到框3372;如果当前帧经分类为缓慢变化的场景的一部分,则在框3373处可将缓慢变化的场景中的当前帧及其它帧编码为预测帧(例如,P帧或B帧)。过程C接着进行到框3374,在框3374处,检查当前帧是否经分类为包含相机闪光的闪光灯场景。如果答案为是,则在框3375处可识别帧用于特殊处理,例如,移除、复制前一帧,或编码用于所述帧的特定系数。如果答案为否,则不进行当前帧的任何分类且可根据其它标准来编码选定帧,将选定帧编码为I帧或丢弃。可在编码器中实施过程C。

[0328] 在上述方面中,通过帧差异度量 D 指示待压缩的帧与所述帧的相邻的两个帧之间的差异量。如果检测到显著的单向亮度变化量,则其表示帧中的交叉衰落效应。交叉衰落越显著,可通过使用B帧实现越大的增益。在一些方面中,使用如以下等式57中所展示的经修

改的帧差异度量:

$$[0329] \quad D_1 = \begin{cases} \left(1 - \alpha + 2\alpha \frac{|d_P - d_N|}{d_P + d_N}\right) \times D, & \text{如果 } Y_{P-\Delta} \geq Y_C \geq Y_{N+\Delta} \text{ 或 } Y_{P+\Delta} \leq Y_C \leq Y_{N-\Delta}, \\ D, & \text{否则,} \end{cases} \quad (57)$$

[0330] 其中 $d_P = |Y_C - Y_P|$ 及 $d_N = |Y_C - Y_N|$ 分别为当前帧与前一帧之间的亮度差异及当前帧与下一帧之间的亮度差异, Δ 表示可在正常实验中加以确定的常数(因为其可取决于实施方案), 且 α 为具有在 0 与 1 之间的值的加权变量。

[0331] 如果观察到亮度变化的一致趋势且偏移强度足够大, 则经修改的帧差异度量 D_1 仅不同于最初的帧差异度量 D 。 D_1 等于或小于 D 。如果亮度的变化是稳定的 ($d_P = d_N$), 则经修改的帧差异度量 D_1 小于最初的帧差异度量 D , 最低比为 $(1 - \alpha)$ 。

[0332] 以下表 1 展示通过添加突然的场景变化检测得到的性能改进。非场景变化 (NSC) 与场景变化 (SC) 状况中, I 帧的总数目近似相同。在 NSC 状况中, I 帧均匀分布于整个序列中, 而在 SC 状况中, I 帧仅指派给突然的场景变化帧。

[0333] 可见在 PSNR 方面通常可实现 0.2~0.3dB 的改进。模拟结果展示: 分镜头检测器在确定以上所提及的镜头事件中非常准确。对具有正常交叉衰落效应的五个码片的模拟展示: 在 $\Delta = 5.5$ 及 $\alpha = 0.4$ 的情况下, 在同一位速率下实现 0.226031dB 的 PSNR 增益。

[0334]

序列\度量	位速率 (kbps)	平均QP	PSNR (dB)
动画NSC	226.2403	31.1696	35.6426
动画SC	232.8023	29.8171	36.4513
音乐NSC	246.6394	32.8524	35.9337
音乐SC	250.0994	32.3209	36.1202
标题NSC	216.9493	29.8304	38.9804
标题新闻SC	220.2512	28.9011	39.3151
篮球NSC	256.8726	33.1429	33.5262
篮球SC	254.9242	32.4341	33.8635

[0335] 表 1: 突然的场景变化检测的模拟结果

[0336] 自适应 GOP 结构

[0337] 以下描述自适应 GOP 结构操作的一说明性实例。所述操作可包括于图 412 的 GOP 分割器 412 中。虽然可强加一规则结构, 但 MPEG2 (较旧的视频压缩标准) 不要求 GOP 具有一规则结构。MPEG2 序列总是以 I 帧开始, 即, 已在不参考先前图像的情况下予以编码的帧。通常通过固定在 I 帧之后的 P 图像或预测图像的 GOP 中的间距在编码器处预先布置 MPEG2 GOP 格式。P 帧是已从先前 I 图像或 P 图像予以部分预测的图像。起始的 I 帧与后续的 P 帧之间的帧经编码为 B 帧。“B” 帧 (B 代表双向) 可个别或同时使用先前的及接下来的 I 图像或 P 图像作为参考。用于编码 I 帧的位的数目平均会超过用于编码 P 帧的位的数目; 同样, 用于编码 P 帧的位的数目平均会超过用于编码 B 帧的位的数目。如果使用被跳过的帧, 则所述帧可不使用任何位用于其表示。

[0338] 使用 P 帧及 B 帧及 (在更近的压缩算法中的) 帧的跳跃的一益处在于, 有可能减小视频传输大小。当时间冗余较高时 (例如, 当图像间存在小变化时), P 图像、B 图像或被跳过的

图像的使用有效地表示视频流,因为早先经解码的I图像或P图像稍后用作解码其它P图像或B图像的参考。

[0339] 图像群组分割器自适应地编码帧以最小化时间冗余。量化帧之间的差异且在对经量化的差异执行适宜的测试之后自动作出通过I帧、P帧、B帧或被跳过的帧来表示图像的决策。通过预处理器202的其它操作来辅助GOP分割器中的处理,所述处理提供滤波以用于噪声移除。

[0340] 自适应编码过程具有在“固定的”编码过程中不可获得的优点。固定的过程忽略内容中已发生小变化的可能性;然而,自适应程序允许将更多的B帧插入每一I帧与P帧之间或两个P帧之间,藉此减少用于充分表示帧序列的位的数目。相反,(例如)在固定的编码过程中,当视频内容中的变化较显著时,因为预测帧与参考帧之间的差异太大,所以P帧的效率被大大减少。在所述条件下,匹配的物体可能从运动搜索区域中掉出,或由于由相机角度的变化造成的失真而使匹配的物体之间的类似性减少。自适应编码过程可有利地用于视需要确定应在何时编码P帧。

[0341] 在本文所揭示的系统中,自动感测以上所描述的条件类型。本文所描述的自适应编码过程是灵活的且使其适合于内容的所述变化。自适应编码过程评估一帧差异度量,所述帧差异度量可被认为是对具有相同的相加性距离属性的帧之间的距离的测量。在概念上,倘若帧 F_1 、 F_2 及 F_3 具有帧间距离 d_{12} 及 d_{23} ,则认为 F_1 与 F_3 之间的距离至少为 $d_{12}+d_{23}$ 。在此类似距离的度量及其它测量的基础上进行帧指派。

[0342] GOP分割器412通过当接收到帧时将图像类型指派给帧而操作。图像类型指示可用于编码每一块的预测方法:

[0343] 在不参考其它图像下编码I图像。因为I图像是独立的,所以其在数据流中提供存取点,于所述存取点处可开始解码。如果到帧的前导帧的“距离”超过场景变化阈值,则将I编码类型指派给所述帧。

[0344] P图像可使用先前的I图像或P图像来进行运动补偿预测。P图像使用先前场或帧中的可从正被预测的块移动的块作为编码的基础。在从正被考虑的块减去参考块之后,通常使用用于消除空间冗余的离散余弦变换来编码残余块。如果帧与经指派为P帧的最后帧之间的“距离”超过通常小于第一阈值的第二阈值,则将P编码类型指派给所述帧。

[0345] B帧图像可如上所述使用先前的及接下来的P图像或I图像来进行运动补偿。可前向、后向或双向地预测B图像中的块;或可在不参考其它帧的情况下对所述块进行帧内编码。在H.264中,参考块可为来自多达32个帧的多达32个块的线性组合。如果不可将帧指派为I类型或P类型,如果从所述帧到所述帧的紧靠的前导帧的“距离”大于通常小于所述第二阈值的第三阈值,则将所述帧指派为B类型。如果不可将帧指派为变成经编码的B帧,则将所述帧指派为“跳帧”状态。可跳过此帧,因为其实际上为前一帧的复本。

[0346] 评估以显示次序量化相邻帧之间的差异的度量是此处理的第一部分,其发生于GOP分割器412中。此度量是以上所提及的距离;用此度量对每一帧的适当类型进行评估。因此,I帧与相邻的P帧之间或两个相继的P帧之间的间距可变化。计算所述度量由用基于块的运动补偿器处理视频帧开始,虽然例如 8×8 、 4×4 及 8×16 的其它块大小是可能的,但为视频压缩的基本单位的块通常包含 16×16 像素。对于呈现于输出端处的由两个解交错场组成的帧,以场为基础进行运动补偿,对参考块的搜索在场中发生而不是在帧中发生。对于当前

帧的第一场中的一块,在当前帧之后的帧的场中寻找一前向参考块;同样,在紧靠于当前场之前的帧的场中寻找一后向参考块。将所述当前块组合成经补偿的场。所述过程以帧的第二场继续。组合两个经补偿的场以形成前向及后向补偿帧。

[0347] 对于在反电视电影处理406中所创建的帧,对参考块的搜索可仅以帧为基础,因为仅产生经重构的胶片帧。得到两个参考块及两个差异(前向与后向),从而还产生前向及后向补偿帧。总之,运动补偿器产生用于每一块的运动向量及差异度量。注意,视正在评估前向差异还是后向差异而定,在正被考虑的场或帧中的块与最佳地匹配所述块的块(所述块在先前的场或帧中或在紧靠于其后的场或帧中)之间评估度量的差异。仅亮度值参与此计算。

[0348] 运动补偿步骤因此产生两组差异。所述差异在具有当前亮度值的块之间以及具有从在时间上紧靠于当前帧之前及紧靠于当前帧之后的帧获得的参考块中的亮度值之间。为块中的每一像素确定每一前向差异及每一后向差异的绝对值且分别在整個帧上合计每一者。当处理包含帧的解交错NTSC场时,两种求和包括两个场。以此方法,得到前向差异及后向差异的合计绝对值 SAD_P 及 SAD_N 。

[0349] 对于每一帧,使用以下关系计算SAD比,

$$[0350] \quad \gamma = \frac{\varepsilon + SAD_P}{\varepsilon + SAD_N} \quad (58)$$

[0351] 其中 SAD_P 及 SAD_N 分别为前向差异及后向差异的合计绝对值。将小的正数 ε 添加到分子以防止“被除以零”的误差。将类似的 ε 项添加到分母,进一步减少当 SAD_P 或 SAD_N 接近于零时 γ 的敏感性。

[0352] 在一替代方面中,差异可为SSD(平方差的和)及SAD(绝对差的和)或SATD(其中,通过在获得块元素中的差异之前对块应用二维离散余弦变换来变换像素值的块)。虽然在其它方面中可使用较小的区,但所述和是在有效视频的区上加以评估。

[0353] 还计算所接收的每一帧(未经运动补偿)的亮度直方图。所述直方图对系数的 16×16 阵列中的DC系数(即,(0,0)系数)(如果其可用)进行操作,所述阵列是对亮度值的块应用二维离散余弦变换的结果。等效地, 16×16 块中的亮度的256个值的平均值可用于直方图中。对于亮度深度为八位的图像,区间的数目经设定为16。接下来的度量评估直方图差异

$$[0354] \quad \lambda = \frac{1}{N} \sum_{i=1}^{16} |N_{Pi} - N_{Ci}| \quad (59)$$

[0355] 以上, N_{Pi} 为来自第*i*区间中的前一帧的块的数目,且 N_{Ci} 为来自属于第*i*区间中的当前帧的块的数目, N 为帧中的块的总数目。

[0356] 如下组合所述中间结果以形成当前帧差异度量

$$[0357] \quad M = \frac{\gamma_C}{\gamma_P} + \lambda(2\lambda + 1) \quad (60)$$

[0358] 其中 γ_C 为基于当前帧的SAD比,且 γ_P 为基于前一帧的SAD比。如果场景具有平滑运动且其亮度直方图几乎不改变,则 $M \approx 1$ 。如果当前帧显示突然的场景变化,则 γ_C 将大且 γ_P 应小。使用比 $\frac{\gamma_C}{\gamma_P}$ 而不单独使用 γ_C 以使得度量经正规化为上下文的活动等级。

[0359] 图40中的数据流4100说明可用于计算帧差异度量的特定组件。预处理器4125将交错场(在具有NTSC源的视频状况下)及胶片图像的帧(当视频源是反电视电影处理的结果时)递送到双向运动补偿器4133。双向运动补偿器4133通过将场分裂成 16×16 像素的块且将每一块与前一帧的场的已定义区中的所有 16×16 块相比较而对所述场(或在电影视频源的状况下的帧)进行操作。选择提供最佳匹配的块且从当前块减去所述块。获得所述差异的绝对值且在包含当前块的256个像素上合计结果。当对所述场的所有当前块进行此操作且接着对两个场进行此操作时,已由后向差异模块4137计算出量 SAD_N (后向差异度量)。可由前向差异模块4136执行类似程序。前向差异模块4136使用在时间上紧靠于当前帧之前的帧作为参考块的源以得出 SAD_P (前向差异度量)。虽然使用经恢复的胶片帧进行估计过程,但同一估计过程还发生于当输入帧在反电视电影处理中形成时。可在直方图差异模块4141中形成可用于完成帧差异度量的计算的直方图。基于块的亮度的平均值将每一 16×16 块指派给一区间。此信息是通过以下所形成:在一块中将所有256个像素亮度值相加,通过256对其进行正规化(必要时)且使已置有平均值的区间的计数递增。对每一经预先运动补偿的帧进行一次计算,当新的当前帧到达时,用于当前帧的直方图变成用于前一帧的直方图。通过直方图差异模块4141中的块的数目区别并正规化所述两个直方图以形成通过等式59定义的 λ 。在帧差异组合器4143中组合所述结果以评估等式60中定义的当前帧差异,所述帧差异组合器4143使用在直方图差异模块4141、前向及后向差异模块4136及4137中得到的中间结果。

[0360] 可通过硬件、软件、固件、中间件、微码或其任何组合实施流程图4100的系统及其组件或步骤。流程图4100的每一功能组件(包括预处理器4135、双向运动补偿器4133、前向及后向差异度量模块4136及4137、直方图差异模块4141及帧差异度量组合器4143)可实现为单独组件,作为硬件、固件、中间件并入于另一装置的组件中,或以在处理器上执行的微码或软件予以实施,或其组合。当以软件、固件、中间件或微码予以实施时,可将执行所要任务的程序码或码段存储于例如存储媒体的机器可读媒体中。码段可表示过程、函数、子程序、程序、例程、子例程、模块、软件封装、类,或指令、数据结构或程序语句的任何组合。可通过传送及/或接收信息、数据、自变量、参数或存储器内容而将码段耦合到另一码段或硬件电路。

[0361] 经接收及处理的数据可存储于存储媒体中,所述存储媒体可包括(例如)芯片配置式存储媒体(例如,ROM、RAM)或连接到处理器的盘型存储媒体(例如,磁性的或光学的)。在一些方面中,组合器4143可含有部分或所有所述存储媒体。图41中的流程图4200说明将压缩类型指派给帧的过程。在一方面M中,等式3中所定义的当前帧差异是用于相对于帧指派所作的所有决策的基础。当决策框4253指示:如果考虑中的帧系一序列中的第一帧,标记为是的决策路径进行到框4255,藉此宣告所述帧为I帧。在框4257中将累积的帧差异设定为零,且过程返回(在框4258中)到起始框4253。如果被考虑的帧不是一序列中的第一帧,则标记为否的路径从作决策的框4253开始,且在测试框4259中针对场景变化阈值测试当前帧差异。如果当前帧差异大于所述阈值,则标记为是的决策路径进行到框4255,再次导致I帧的指派。如果当前帧差异小于场景变化阈值,则否路径进行到框4261,在框4261中,将当前帧差异添加到累积的帧差异。

[0362] 继续所述流程图,在决策框4263处,将累积的帧差异与通常小于场景变化阈值的

阈值 t 比较。如果累积的帧差异大于 t ,则控制转移到框4265,且将帧指派为 P 帧;接着在框4267中将累积的帧差异重设为零。如果累积的帧差异小于 t ,则控制从框4263转移到框4269。在框4269中将当前帧差异与小于 t 的 τ 比较。如果当前帧差异小于 τ ,则在框4273中指派跳过所述帧;如果当前帧差异大于 τ ,则将帧指派为 β 帧。

[0363] 在一替代方面中,将另一帧编码复杂性指示符 M^* 定义为

$$[0364] \quad M^* = M \times \text{最小}(1, \alpha \text{最大}(0, \text{SAD}_P - s) \times \text{最大}(0, \text{MV}_P - m)), \quad (61)$$

[0365] 其中 α 为换算因数(scaler), SAD_P 为具有前向运动补偿的SAD, MV_P 为在来自前向运动补偿的运动向量的像素中测量到的长度的和,且 s 及 m 为当 SAD_P 小于 s 或 MV_P 小于 m 时将帧编码复杂性指示符再现为零的两个阈值数。将使用 M^* 代替图41的流程图4200中的当前帧差异。如图可见,仅当前向运动补偿展示出缓慢的运动等级时, M^* 才不同于 M 。在此状况下, M^* 小于 M 。

[0366] 应注意,本文所描述的分镜头检测及编码方面可作为过程予以描述,所述过程描绘为流程图(flowchart, flow diagram)、结构图或框图。虽然图中所展示的流程图可将操作描述为一循序过程,但可并行或同时执行许多操作。此外,可重新布置操作的次序。当完成过程的操作时,通常终止所述过程。过程可对应于方法、函数、程序、子例程、子程序等。当过程对应于函数时,其终止对应于所述函数到调用函数或主函数的返回。

[0367] 所属领域的技术人员还应明白,可在不影响装置的操作的情况下重新布置本文所揭示的装置的一个或一个以上元件。类似地,可在不影响装置的操作的情况下组合本文所揭示的装置的一个或一个以上元件。所属领域的技术人员将了解,可使用多种不同技术和技法中的任一者来表示信息及多媒体数据。所属领域的技术人员将进一步了解,结合本文所揭示的实例加以描述的各种说明性逻辑块、模块及算法步骤可实施为电子硬件、固件、计算机软件、中间件、微码或其组合。为清楚说明硬件与软件的此互换性,以上已大致在功能性方面描述了各种说明性组件、块、模块、电路及步骤。所述功能性是实施为硬件还是软件取决于特定应用及强加于整个系统的设计约束。对于每一特定应用,所属领域的技术人员可以变化的方法实施所描述的功能性,但所述实施决策不应被解释为导致违背所揭示的方法的范围。

[0368] 举例来说,结合本文所揭示的分镜头检测及编码实例及图式所描述的方法或算法的步骤可直接实施于硬件中、由处理器执行的软件模块中,或两者的组合中。所述方法及算法特别适用于通信技术,其包括视频到移动电话、计算机、膝上型计算机、PDA及所有类型的个人及商业通信装置的无线传输。软件模块可驻留于RAM存储器、快闪存储器、ROM存储器、EPROM存储器、EEPROM存储器、寄存器、硬盘、抽取式磁盘、CD-ROM或此项技术中已知的任何其它形式的存储媒体中。例示性存储媒体耦合到处理器,以使得处理器可从存储媒体读取信息且将信息到写入存储媒体。在替代实施例中,存储媒体可与处理器成一体。处理器及存储媒体可驻留于专用集成电路(ASIC)中。所述ASIC可驻留于调制解调器中。在替代实施例中,处理器及存储媒体可作为离散组件驻留于调制解调器中。

[0369] 此外,结合本文所揭示的实例所描述的多种说明性逻辑块、组件、模块及电路可用以下装置予以实施或执行:通用处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)或其它可编程逻辑装置、离散门或晶体管逻辑、离散硬件组件或其经设计以执行本文所描述的功能的任何组合。通用处理器可为微处理器,但在替代实施例中,

处理器可为任何常规的处理器、控制器、微控制器或状态机。处理器还可实施为计算装置的组合,例如,DSP与微处理器的组合、多个微处理器的组合、一个或一个以上微处理器与DSP核心的联合,或任何其它此种配置。

[0370] 提供对所揭示的实例的先前描述以使任何所属领域的技术人员能够制作或使用所揭示的方法及设备。所属领域的技术人员将容易明白对所述实例的各种修改,且本文所定义的原理可适用于其它实例且可在不脱离所揭示的方法及设备的精神或范围的情况下添加额外元件。对所述方面的描述旨在为说明性的,且无意限制权利要求书的范围。

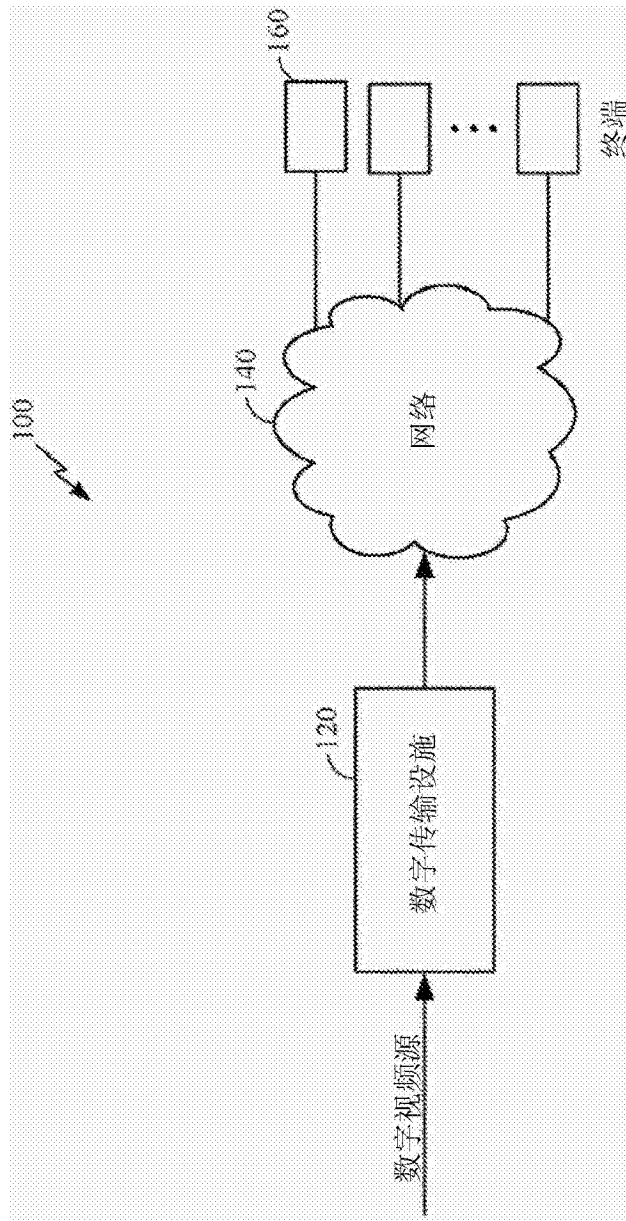


图1

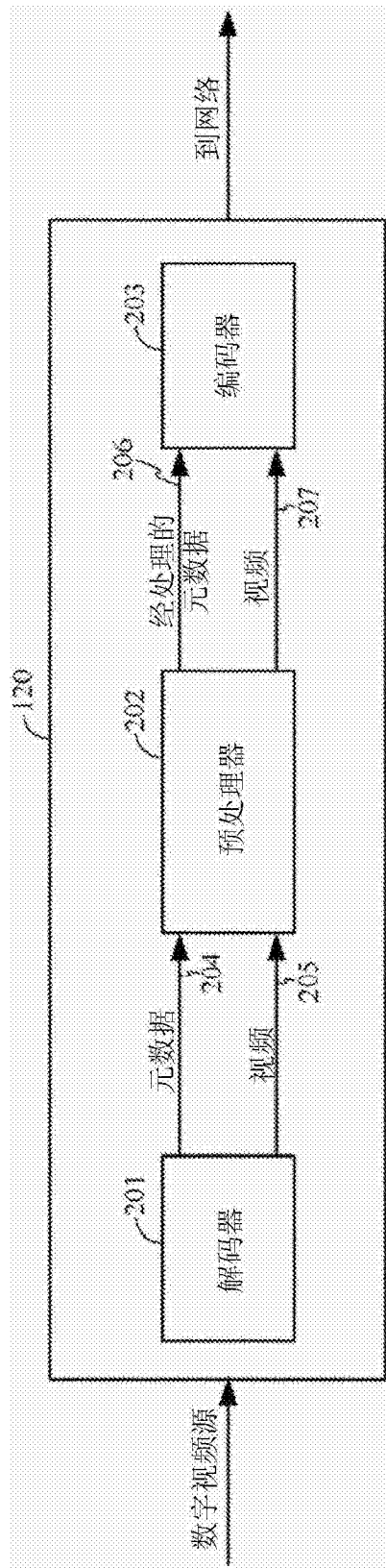


图2

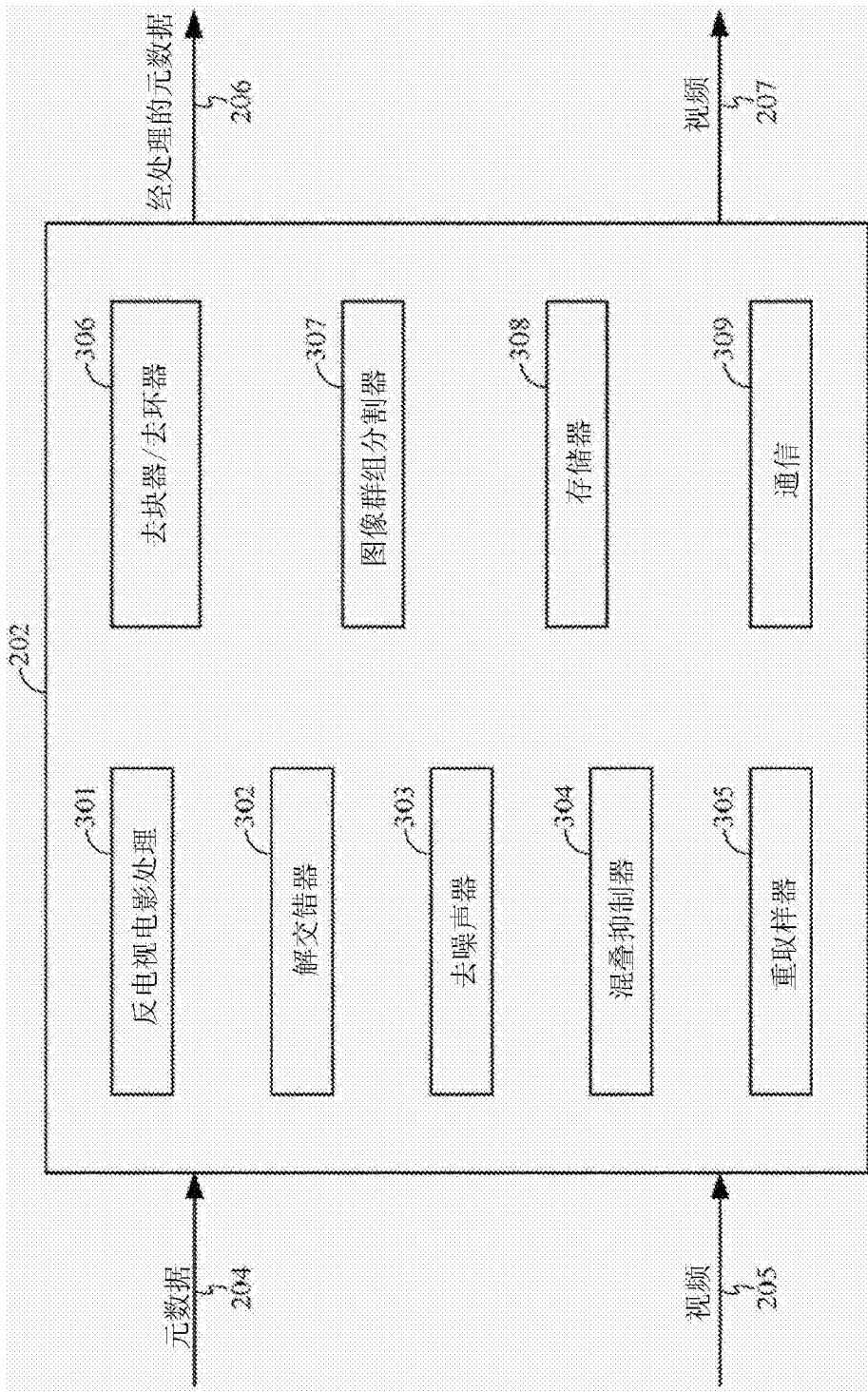


图3A

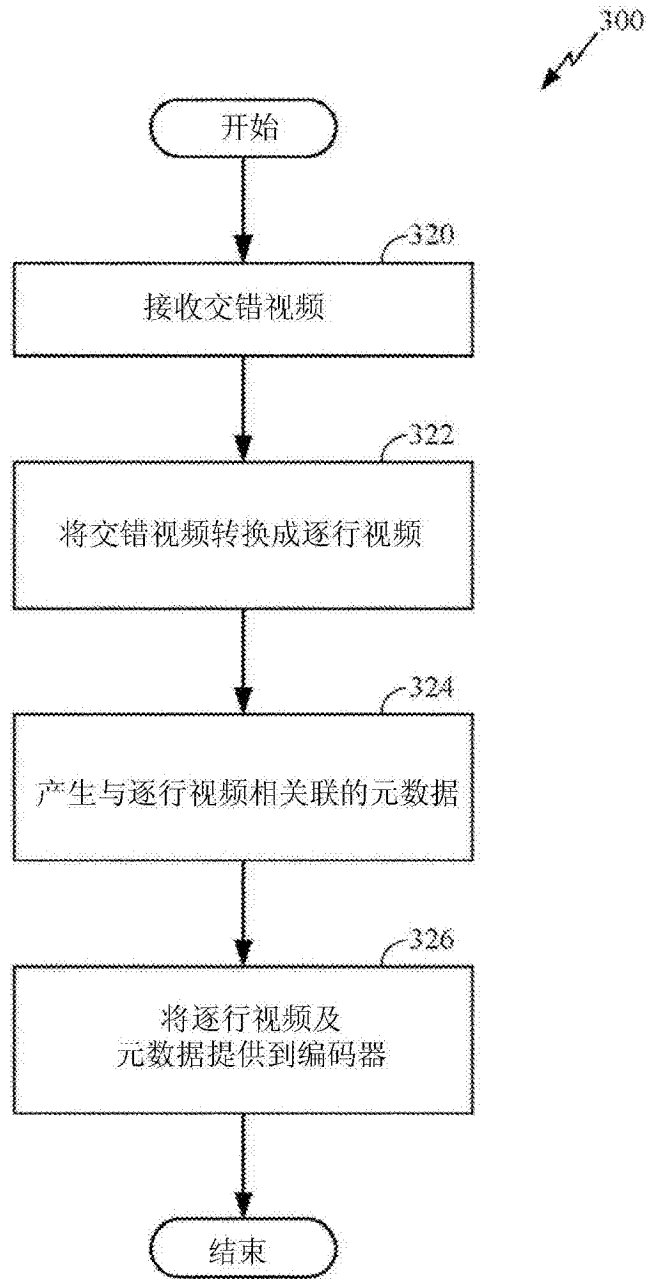


图3B

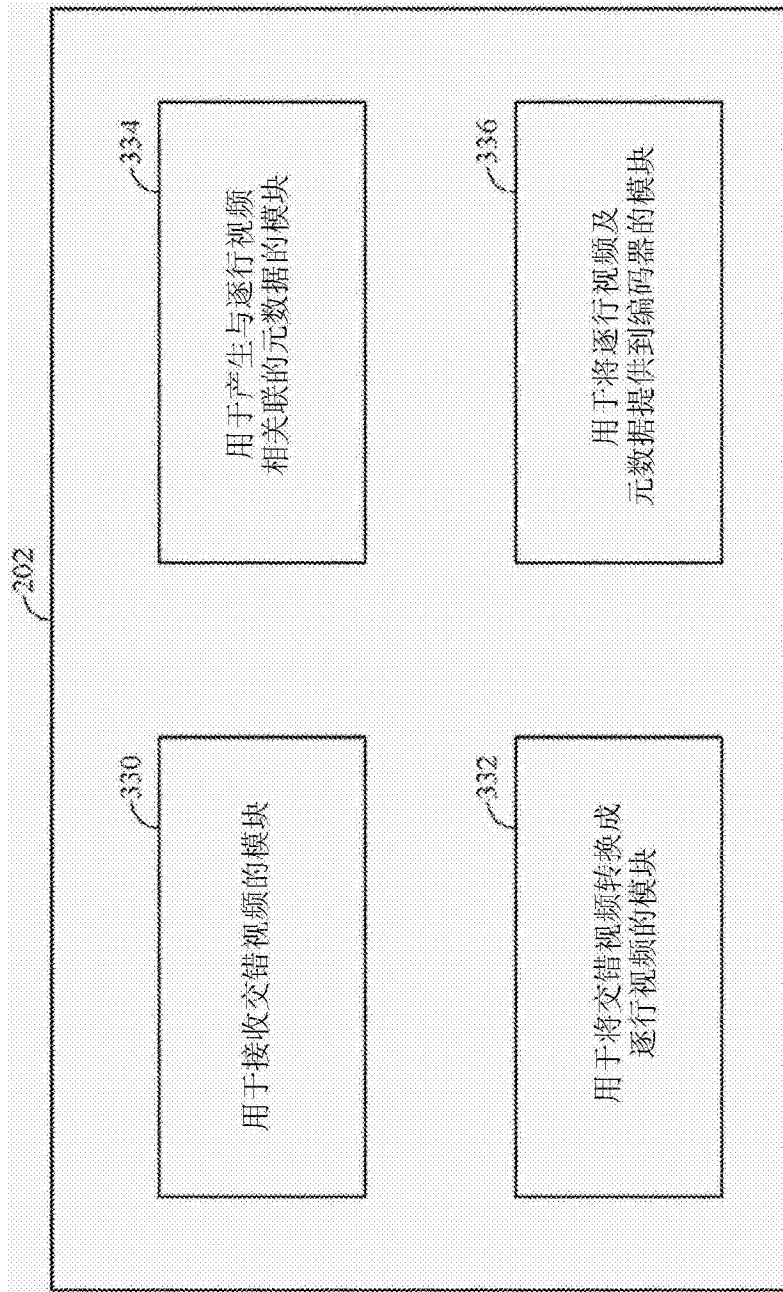


图3C

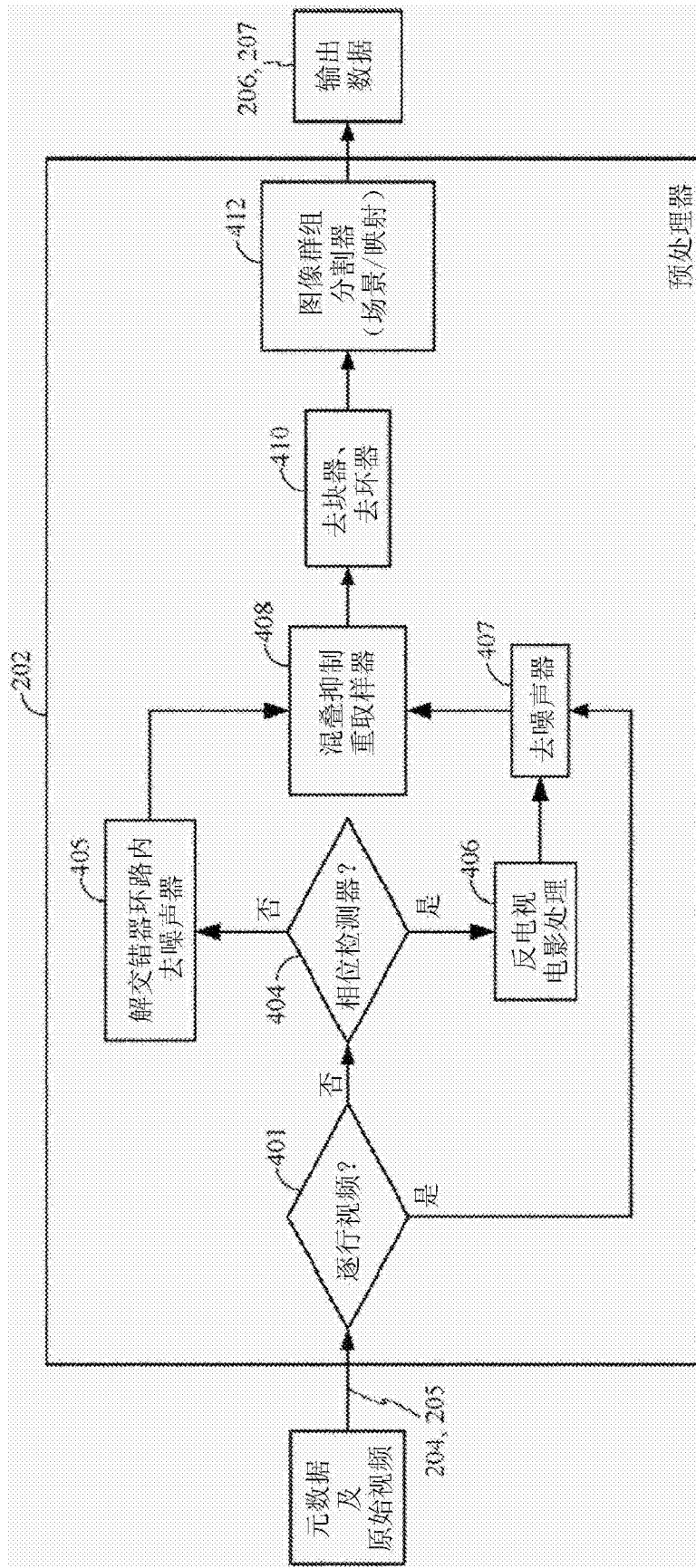


图4

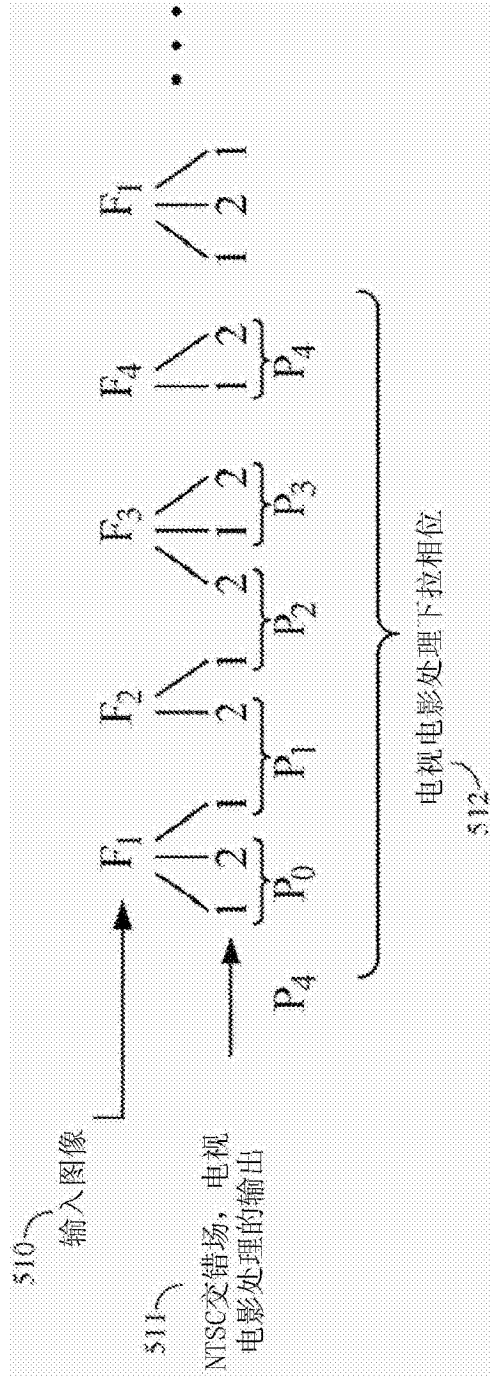
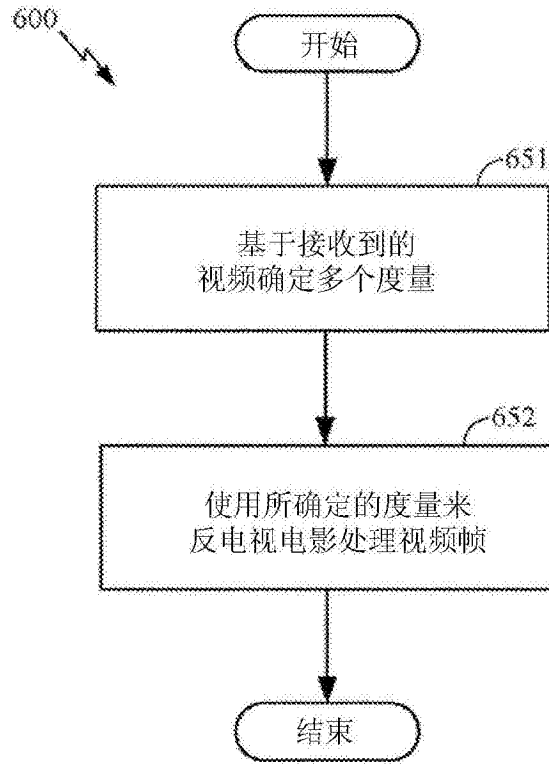


图5



系统概述

图6

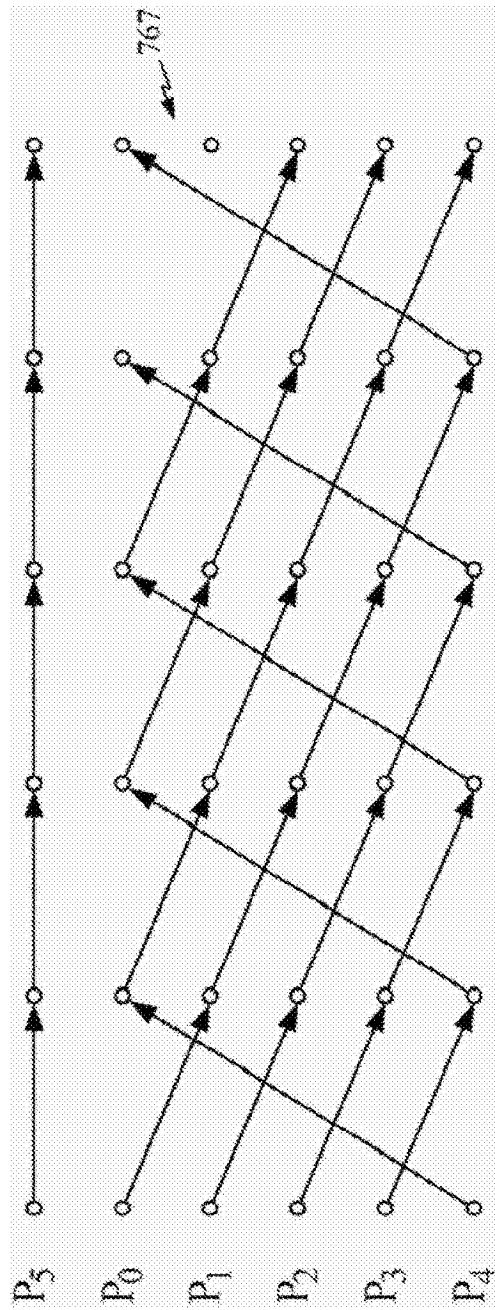


图7

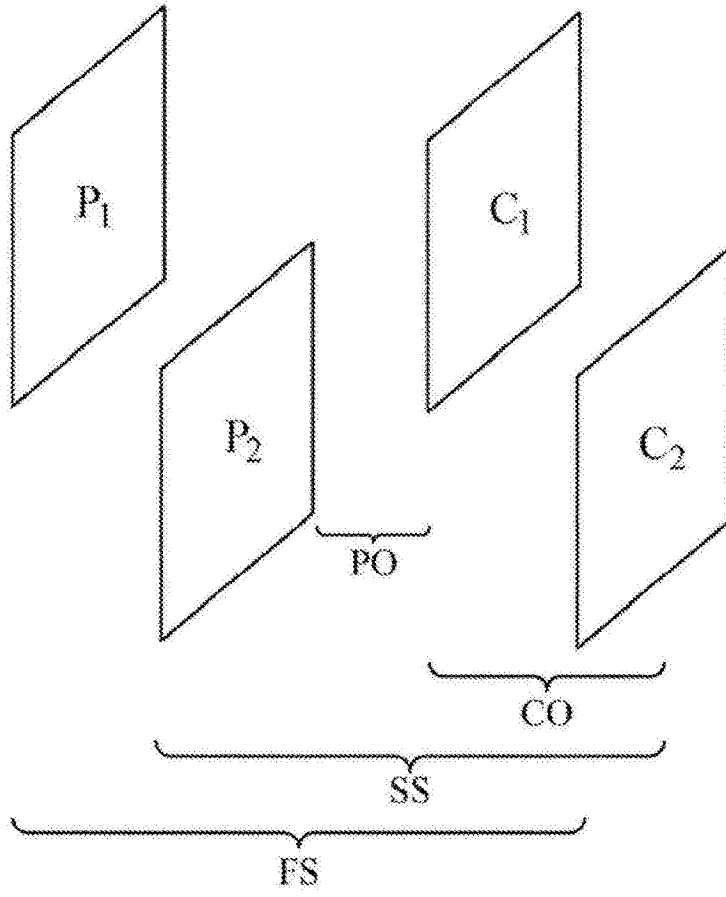


图8

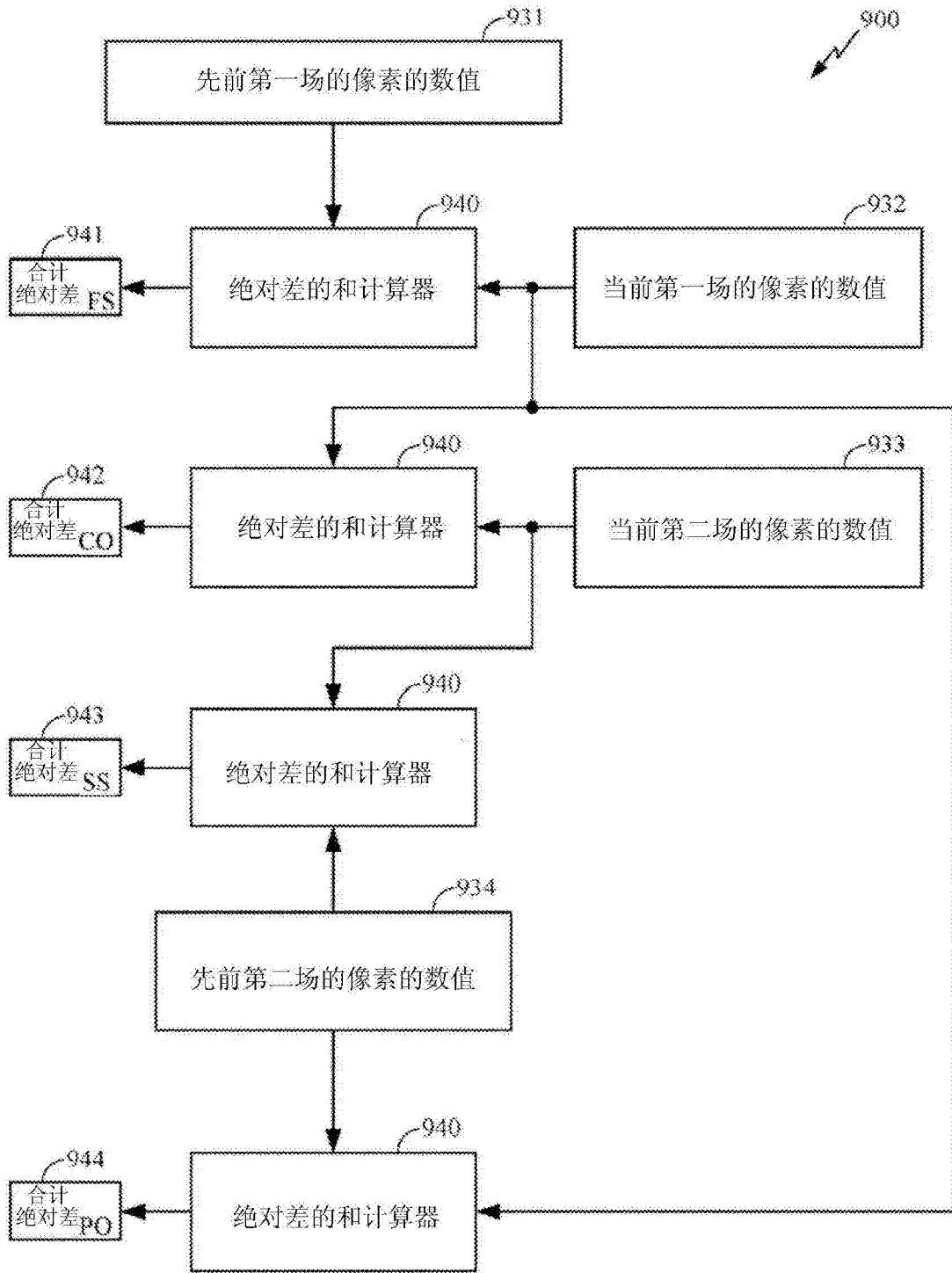


图9

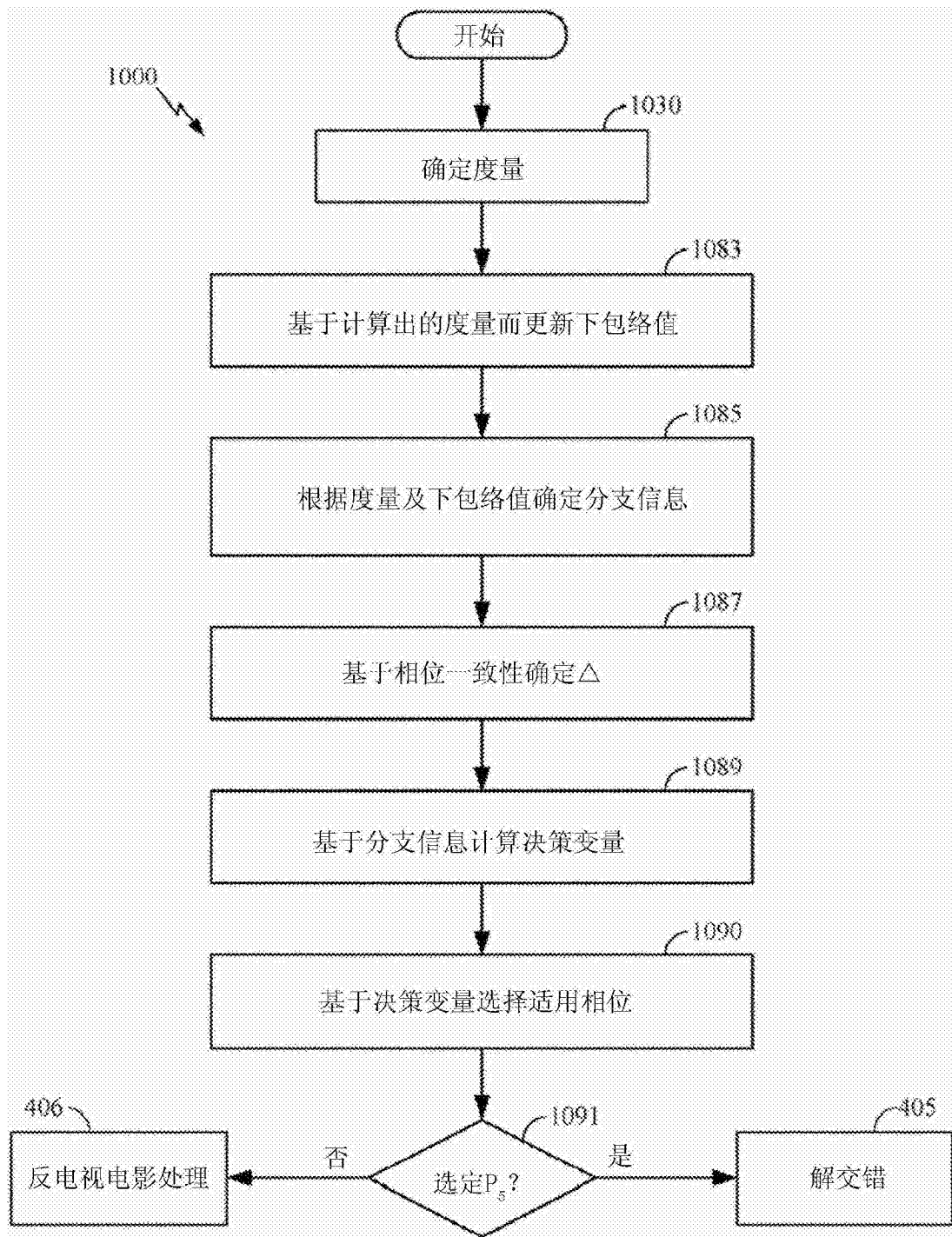


图10

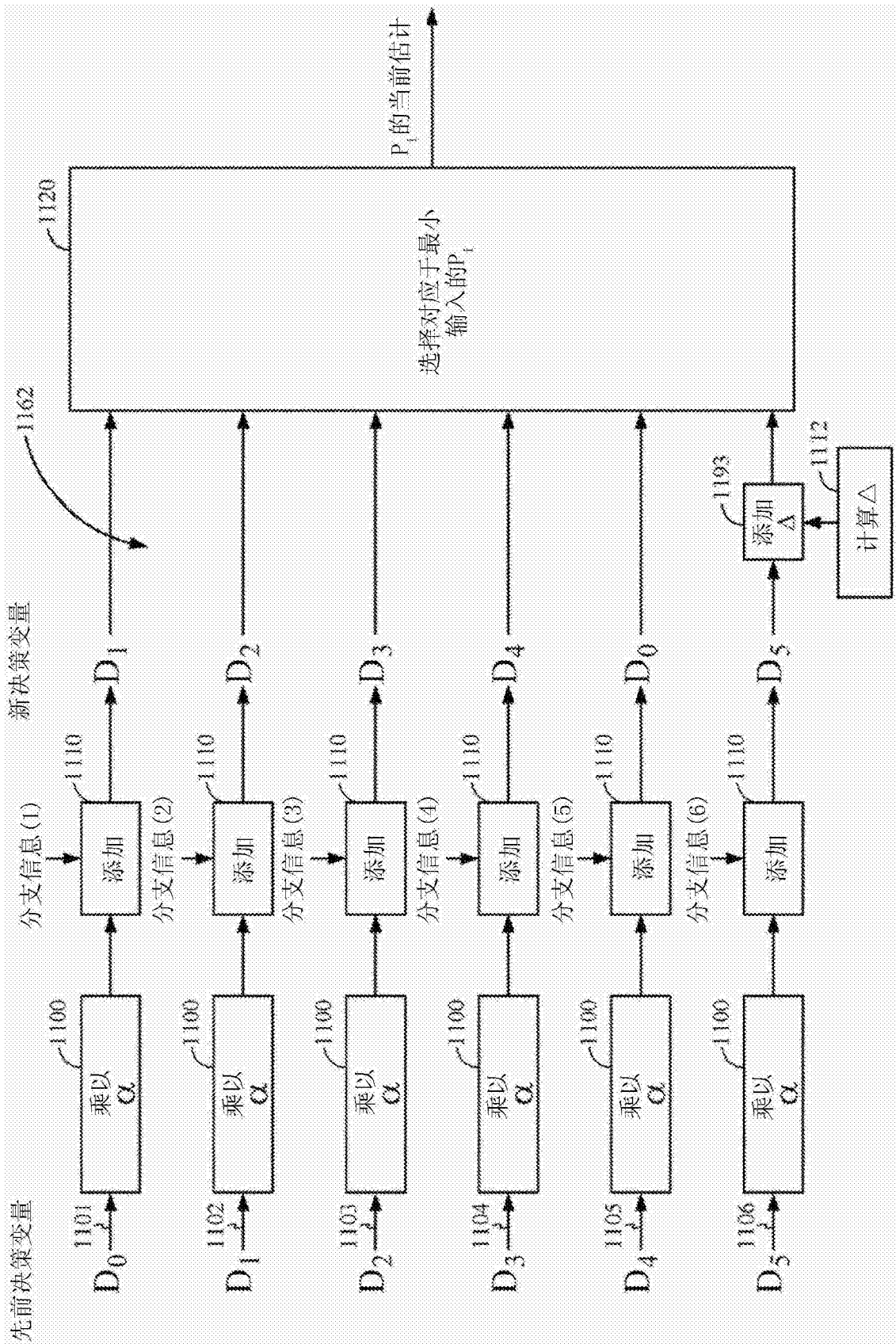


图11

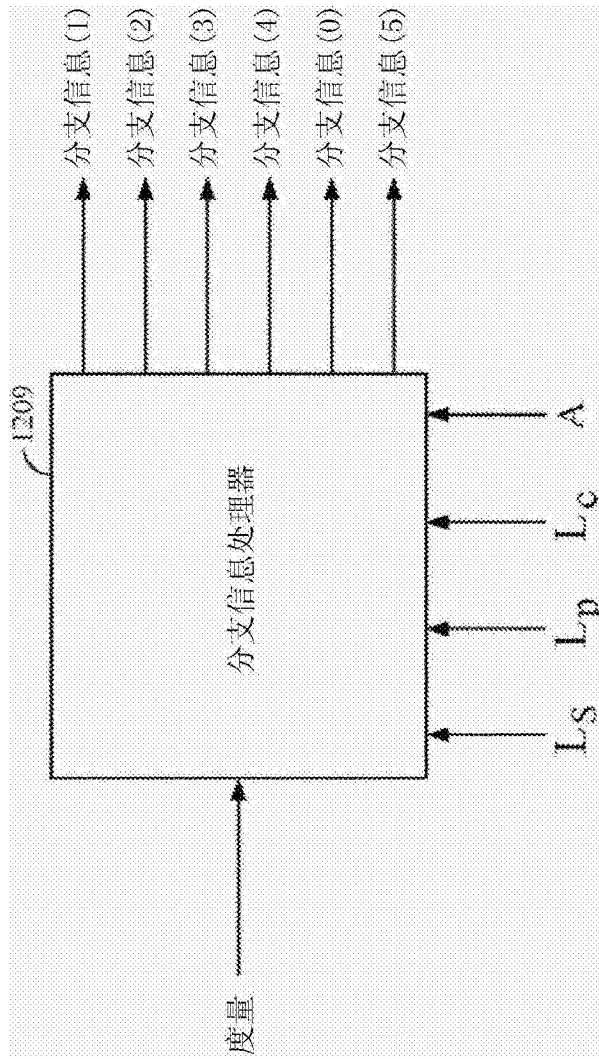


图12

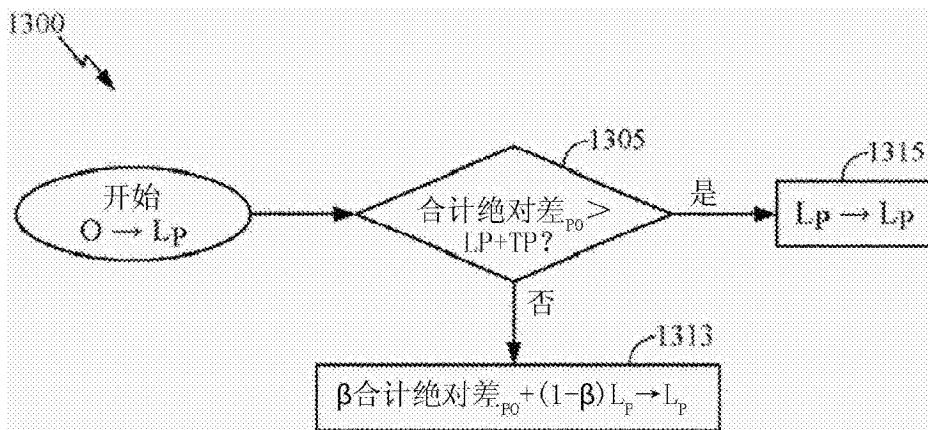


图13A

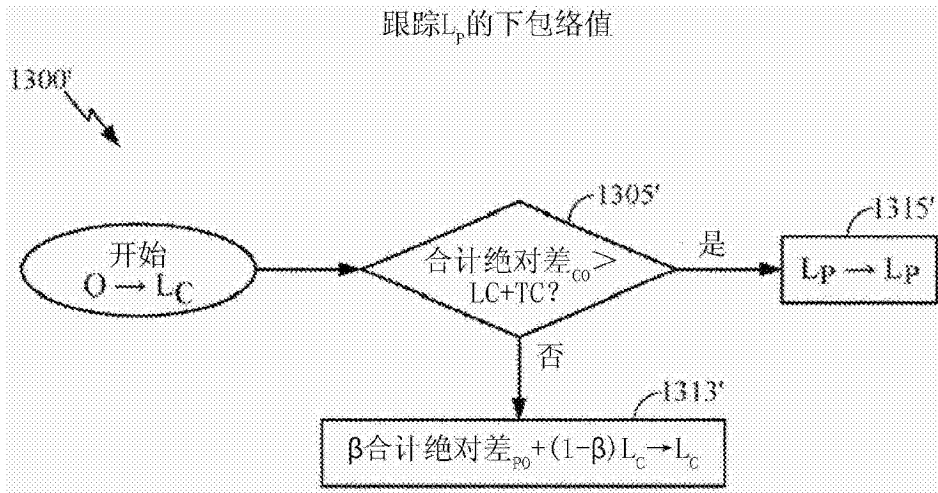


图13B

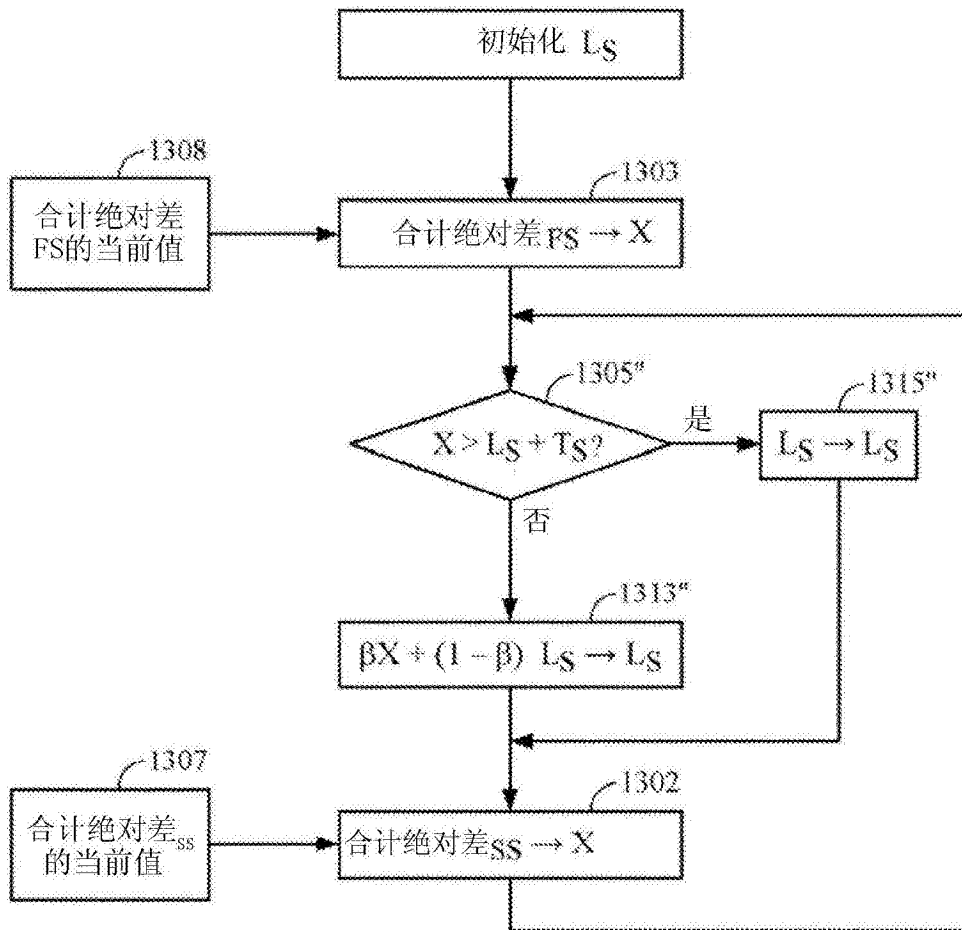


图13C

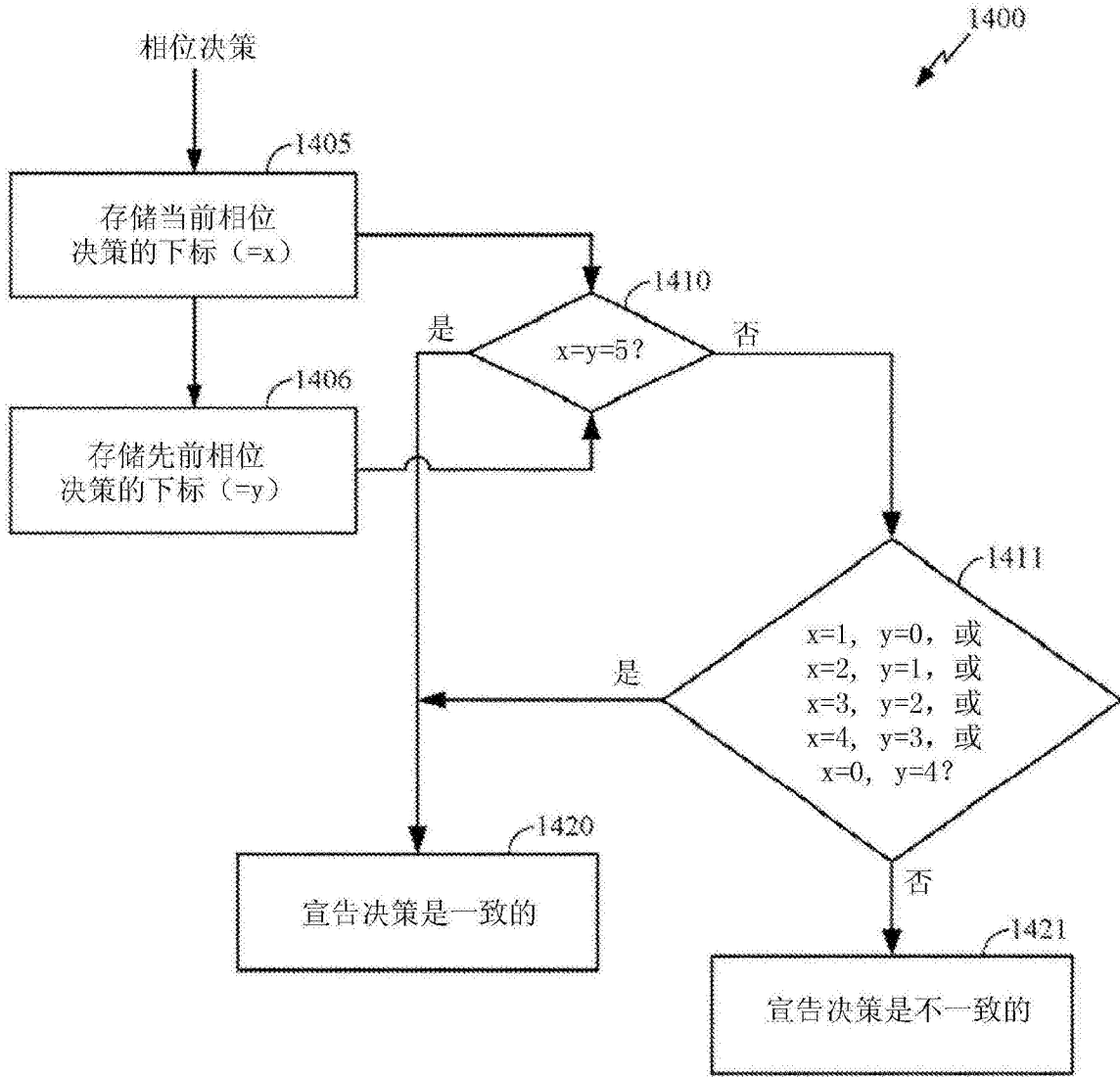


图14

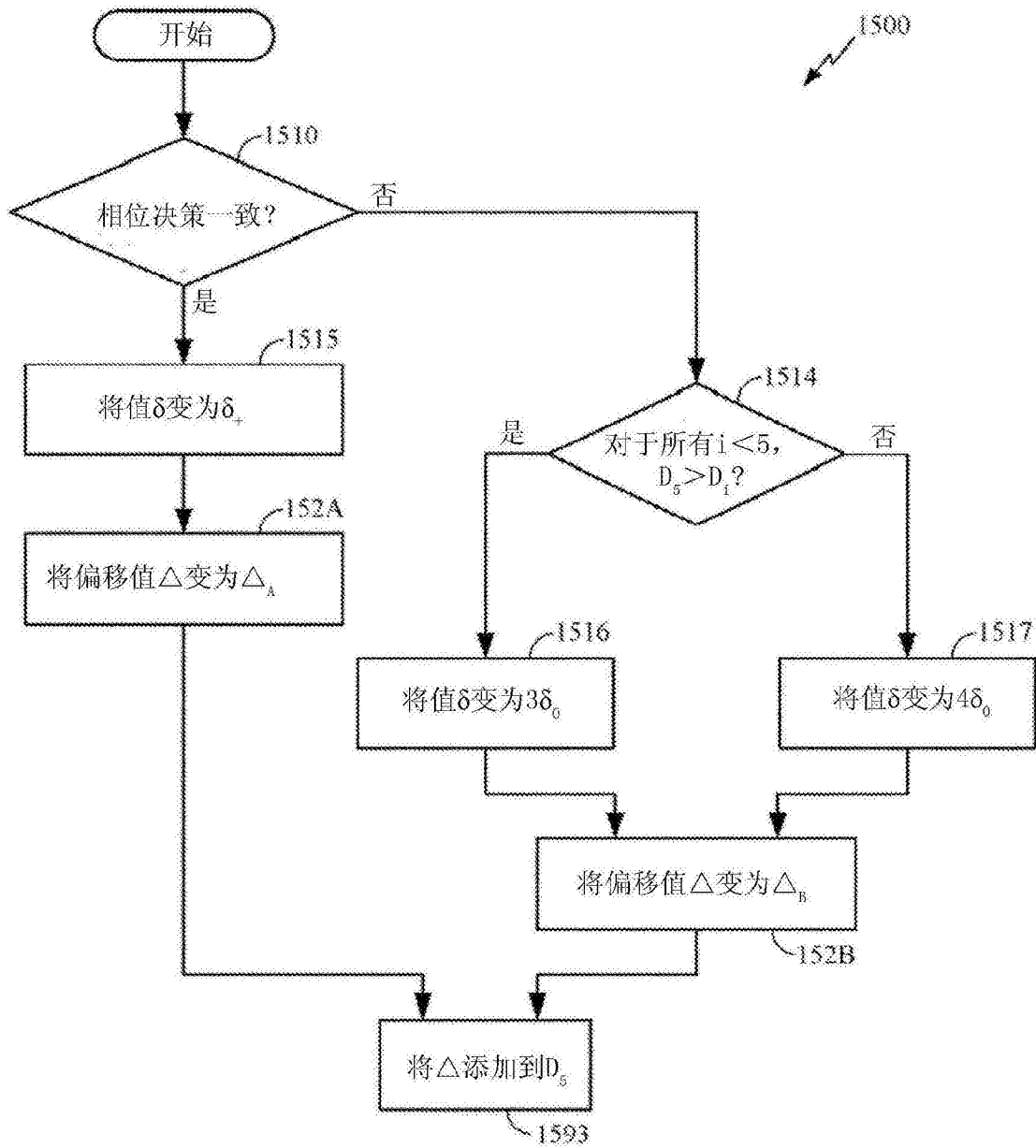


图15

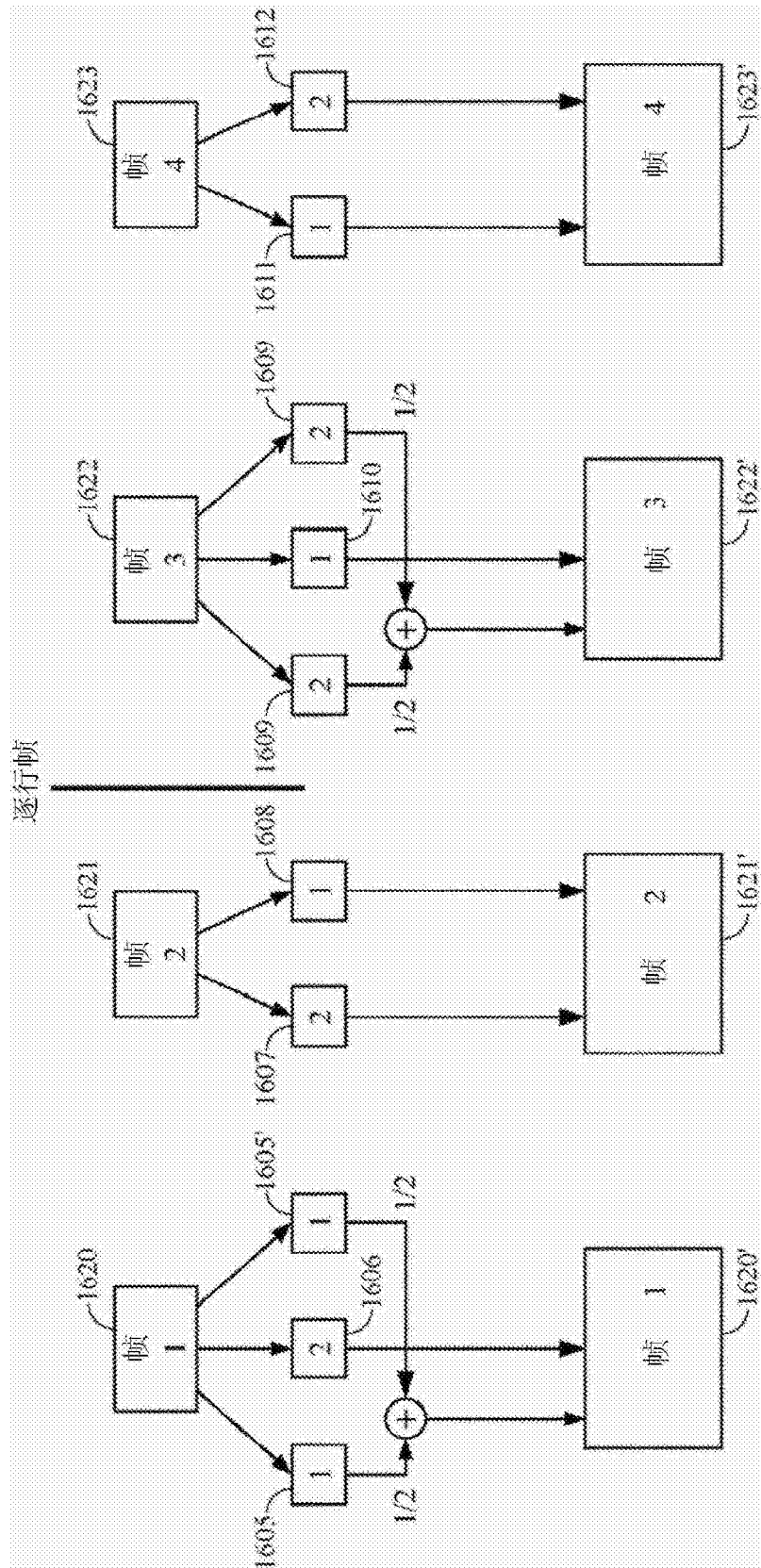


图16

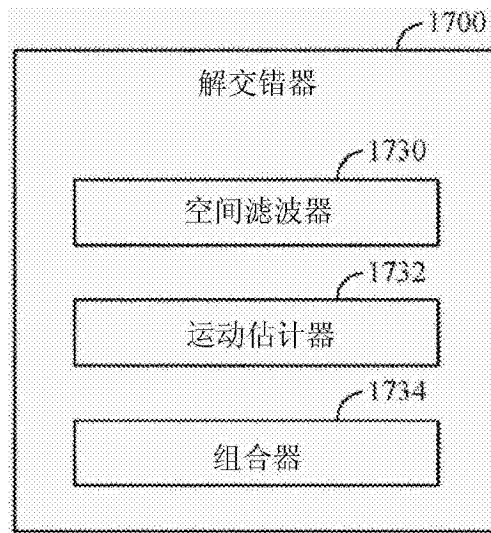


图17

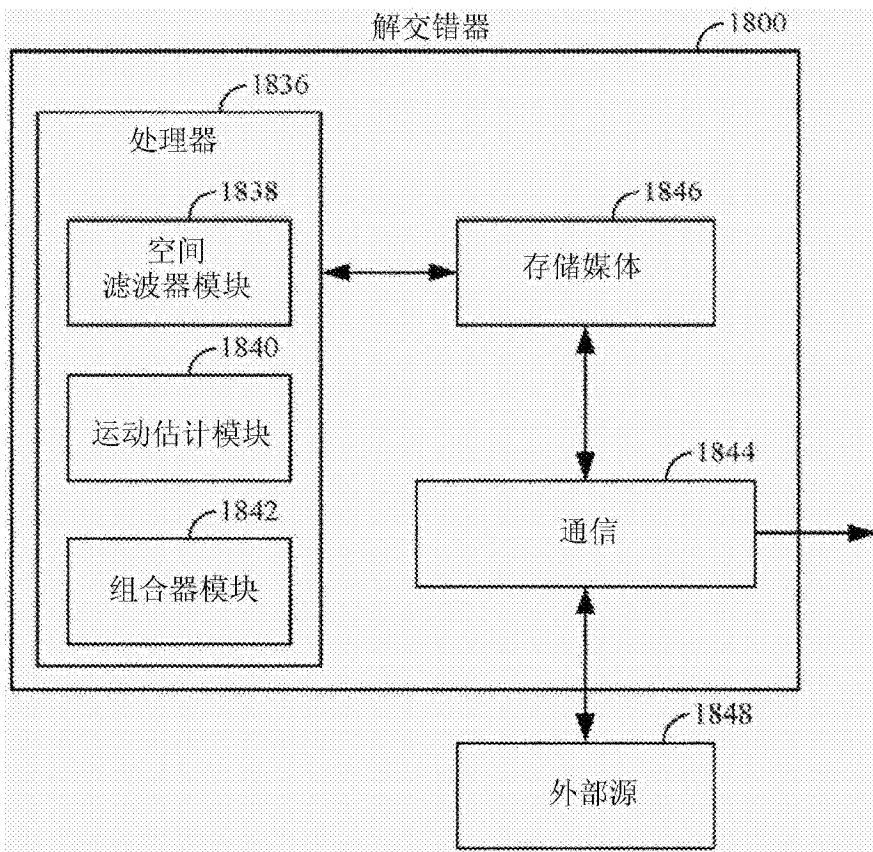


图18

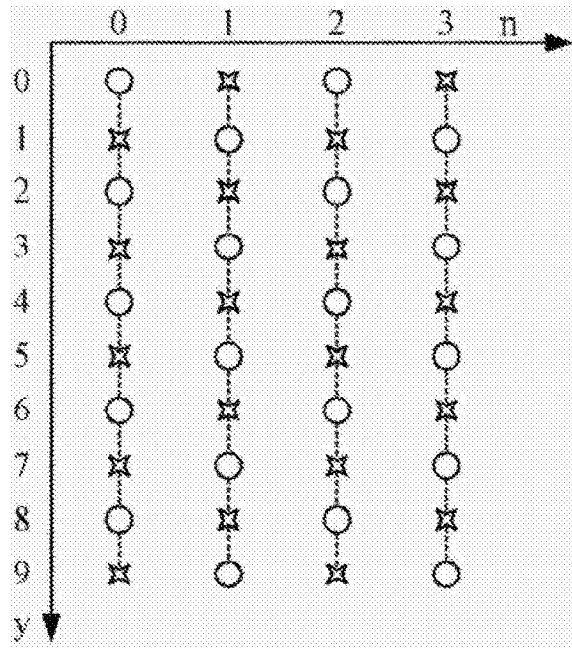


图19

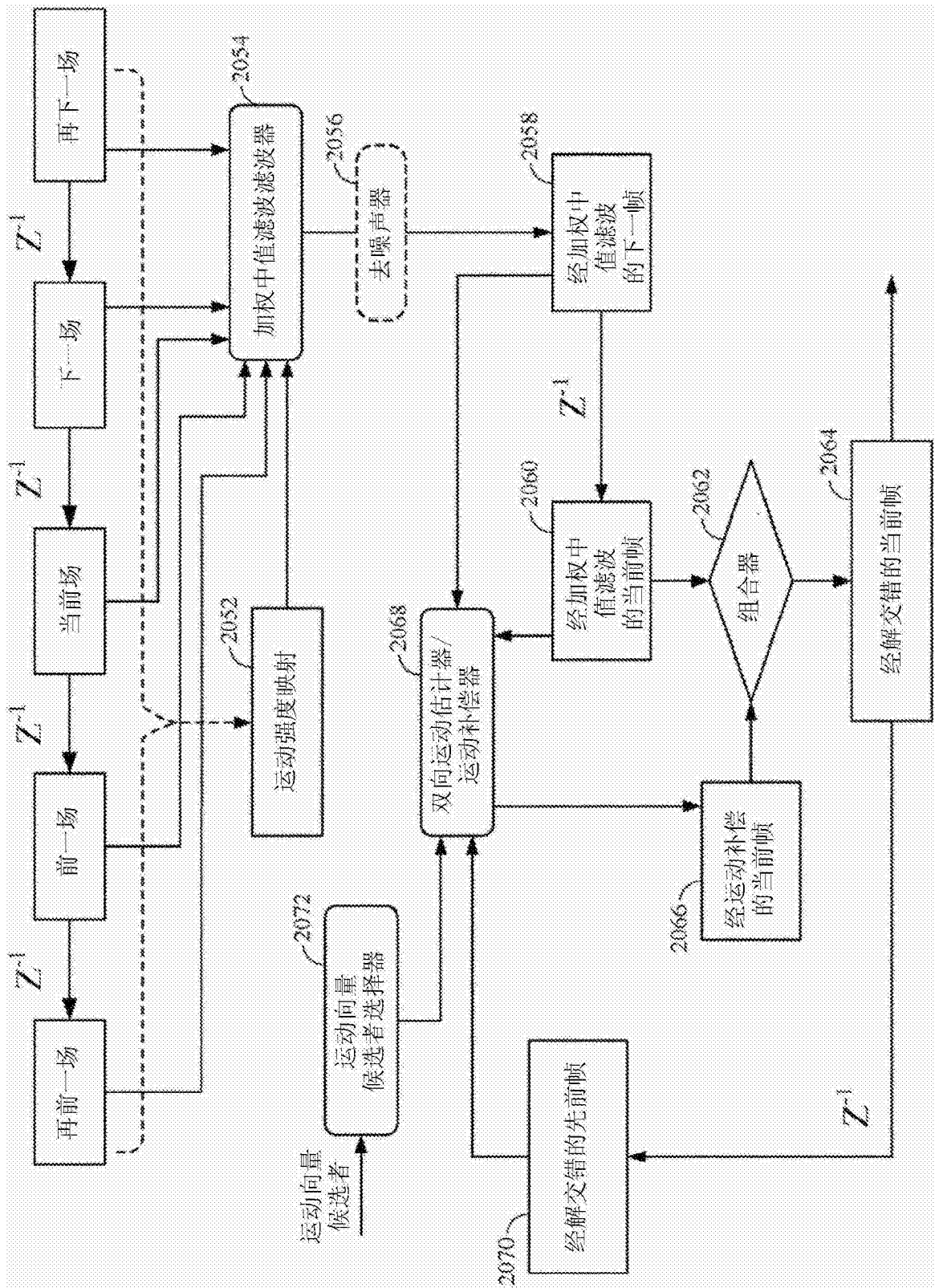


图20

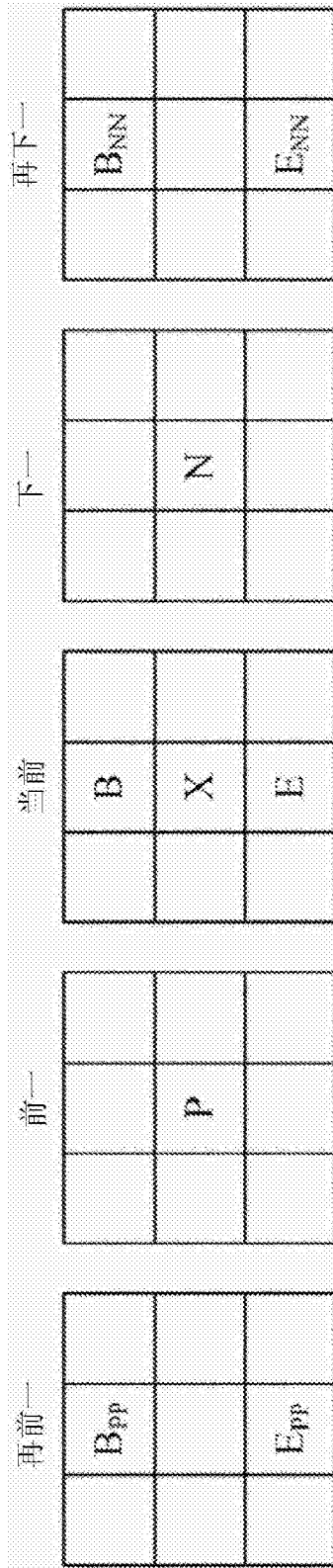


图21

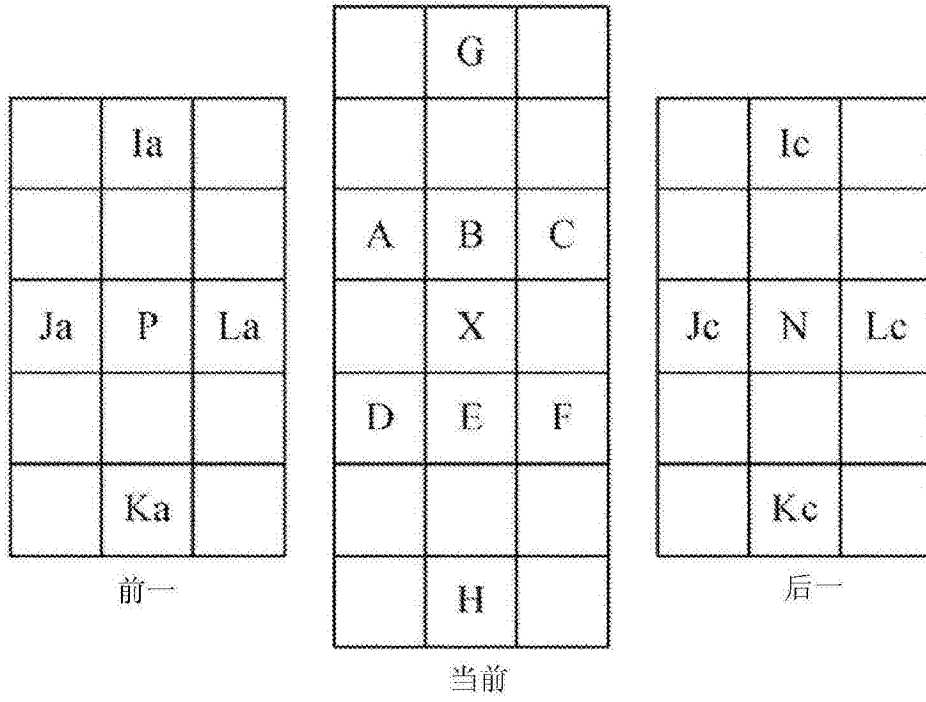


图22

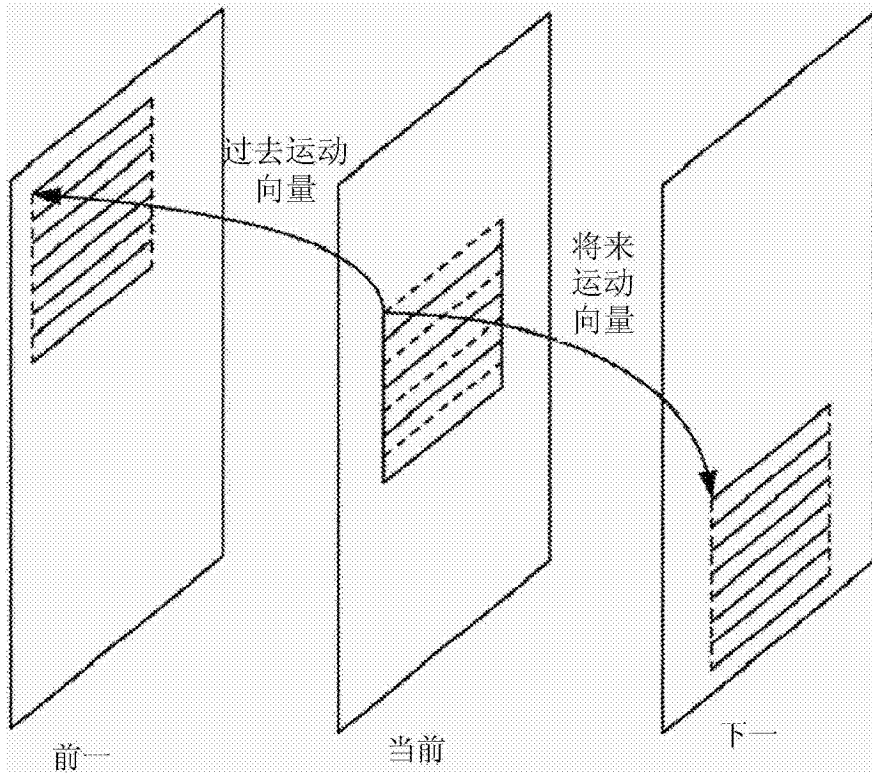


图23

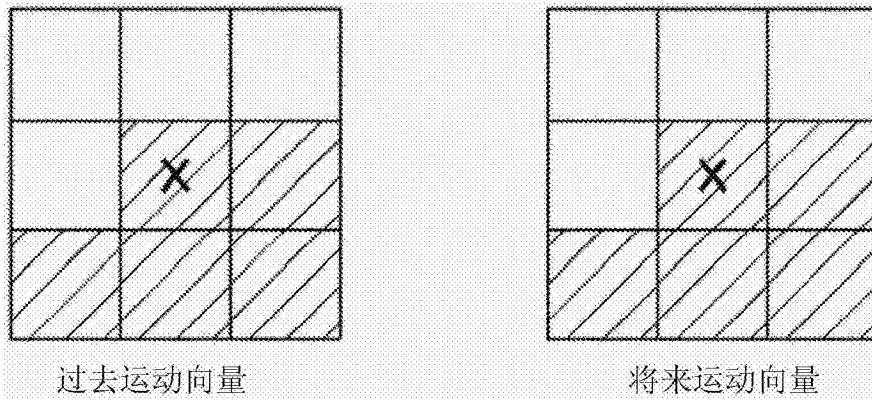


图24

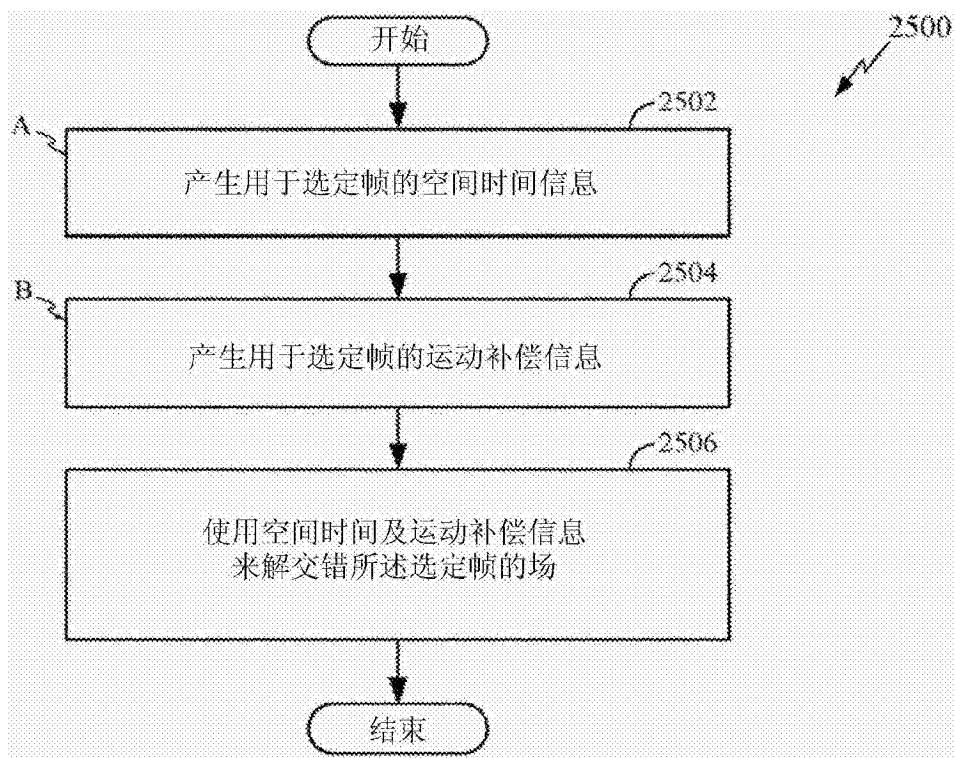


图25

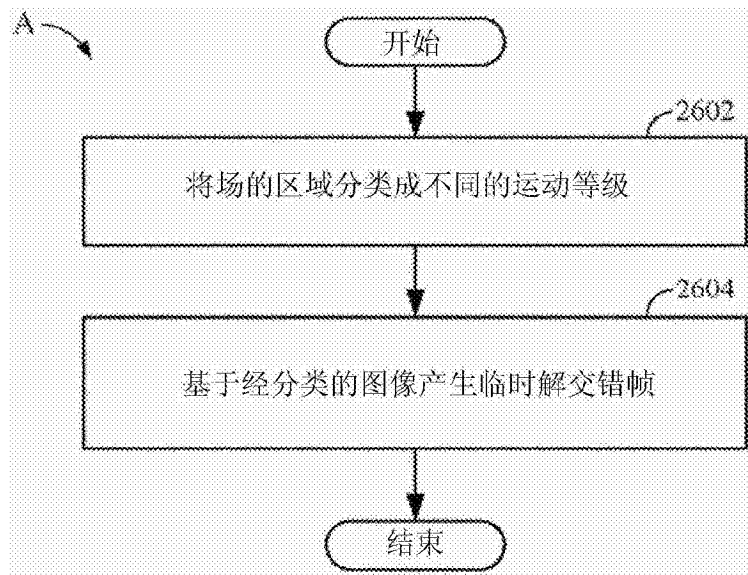


图26

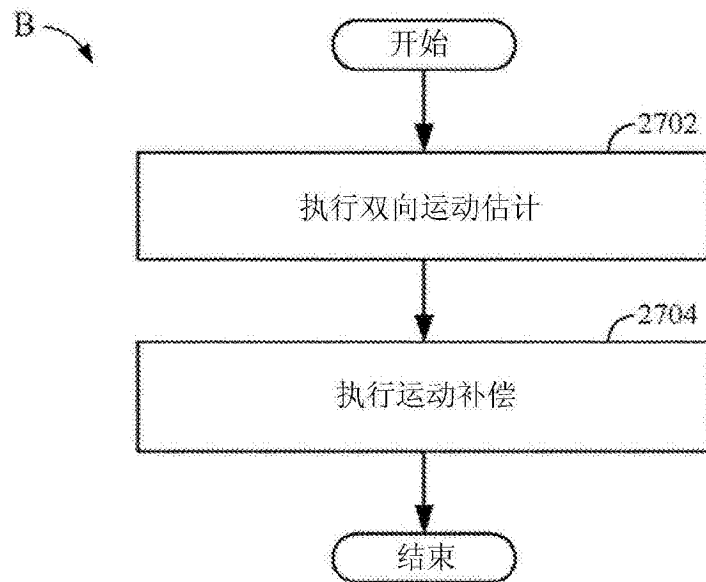


图27

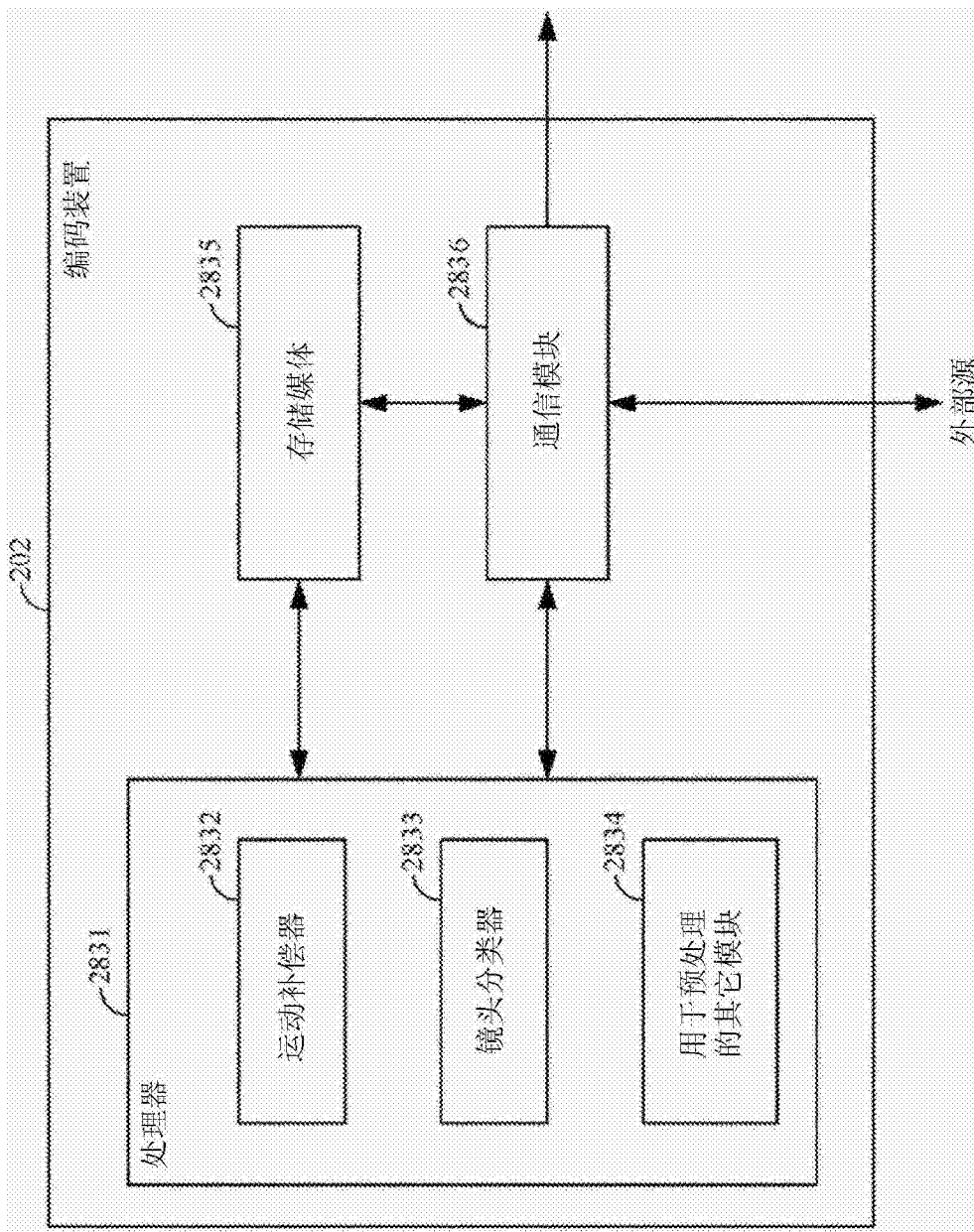


图28

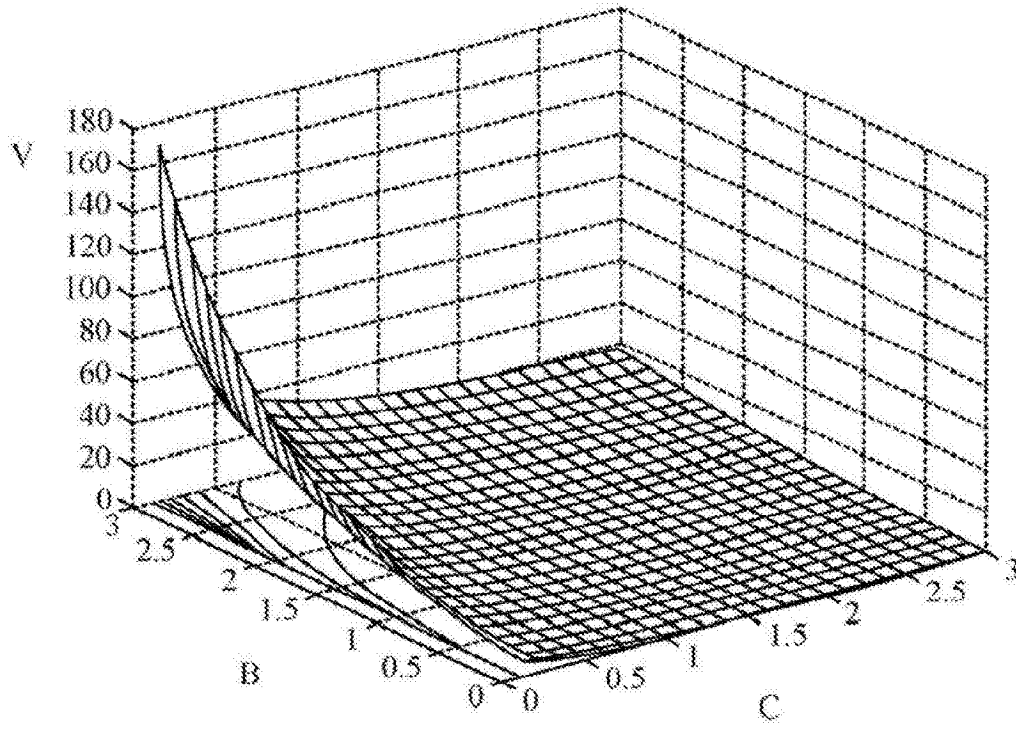


图29

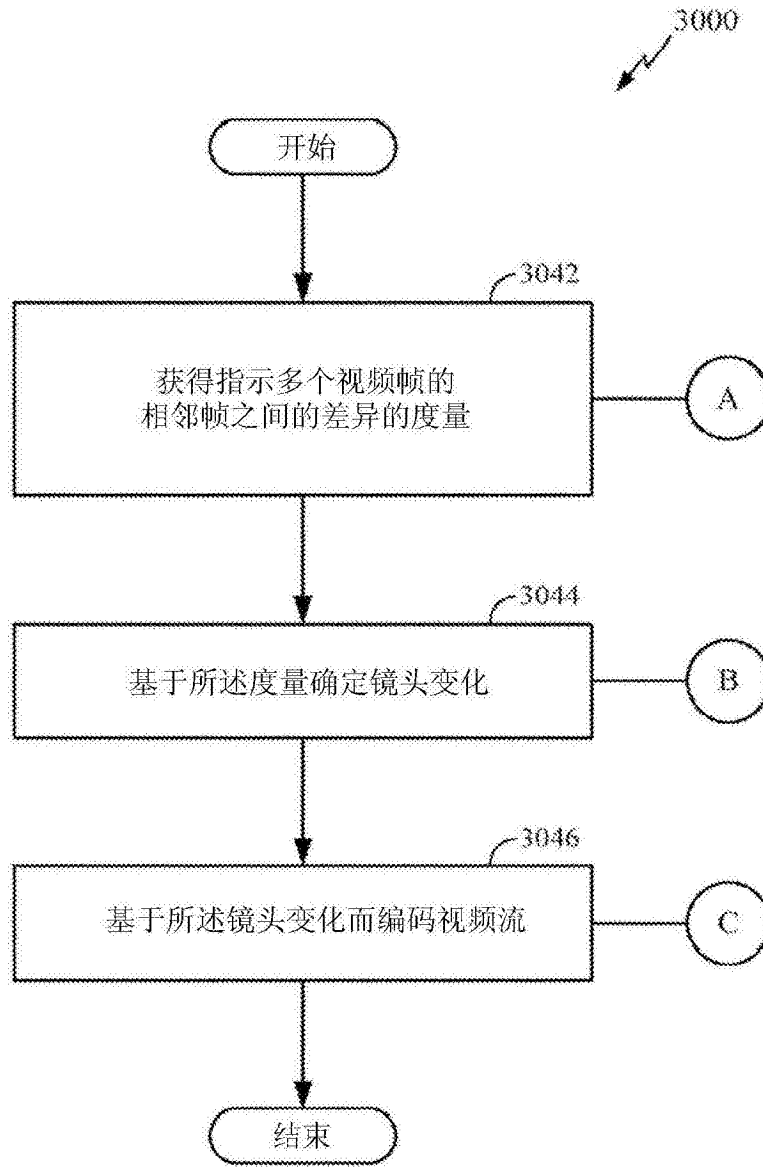


图30

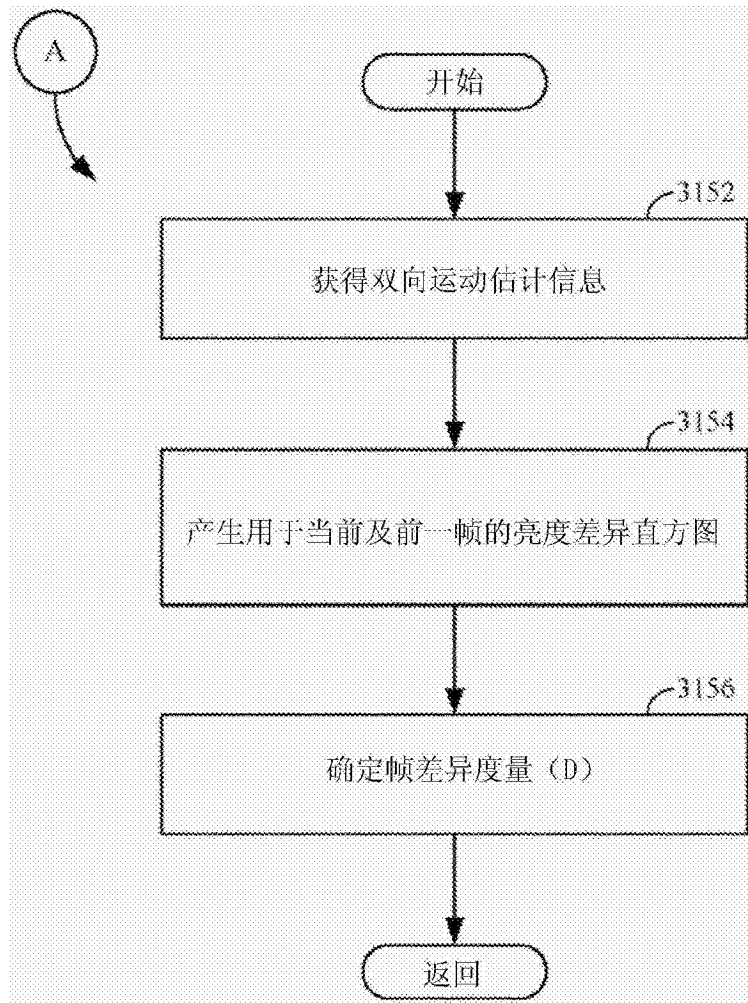


图31

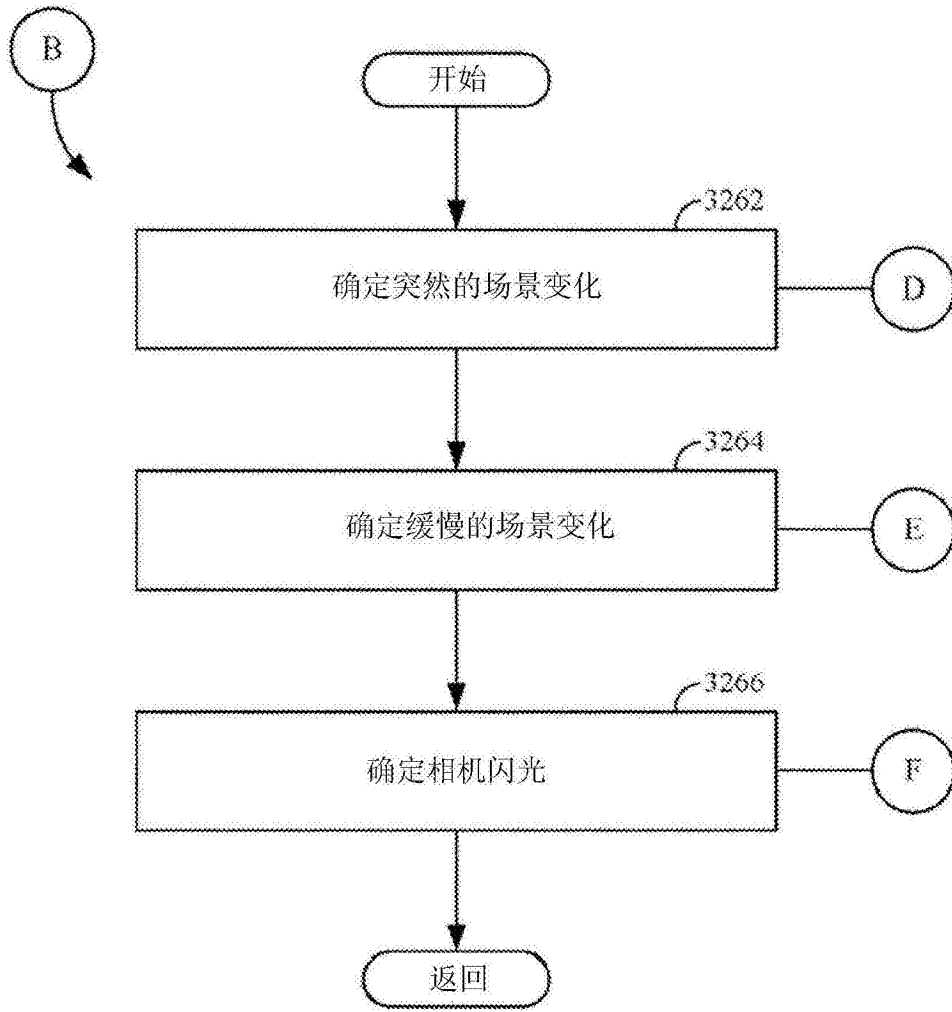


图32

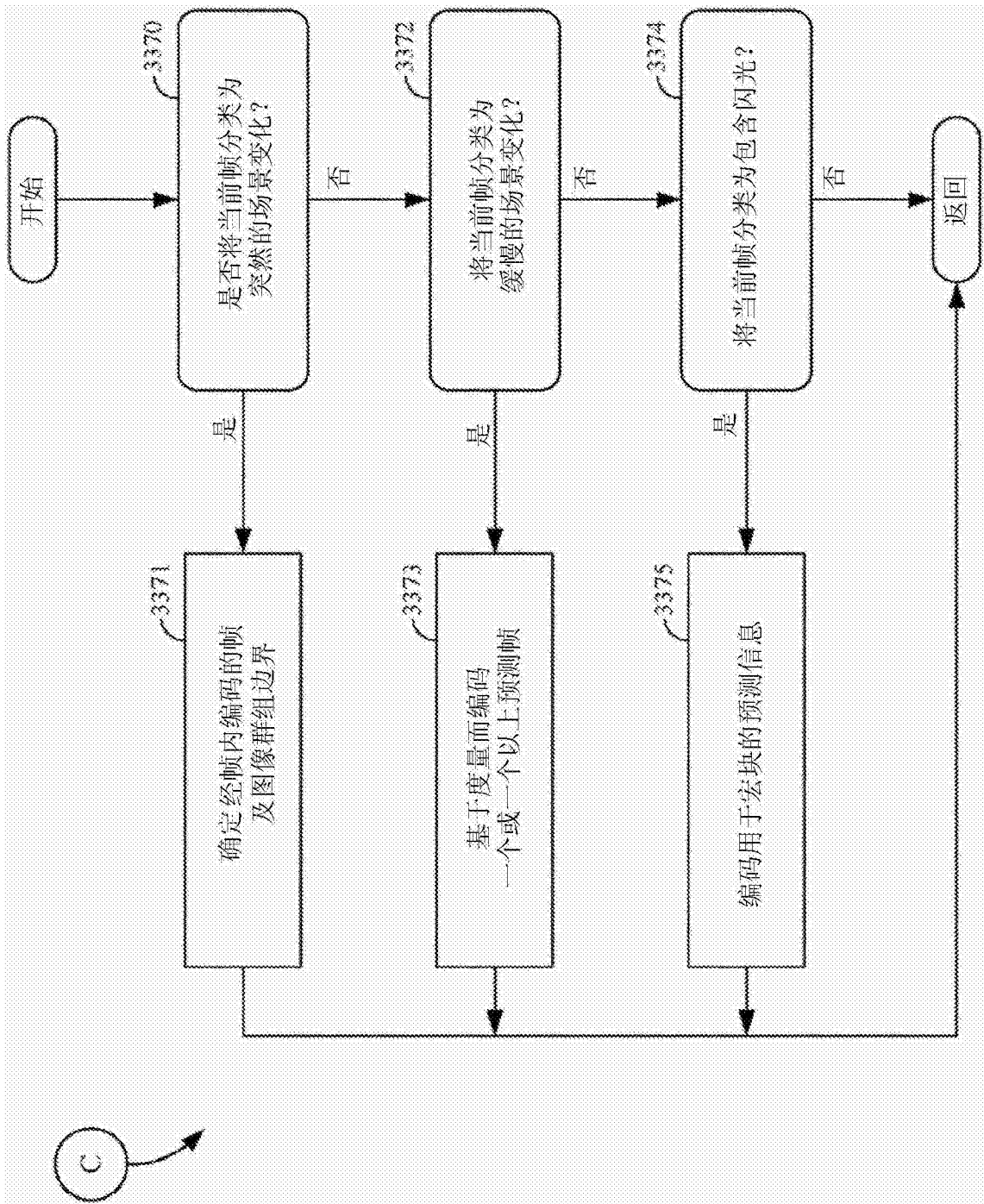


图33

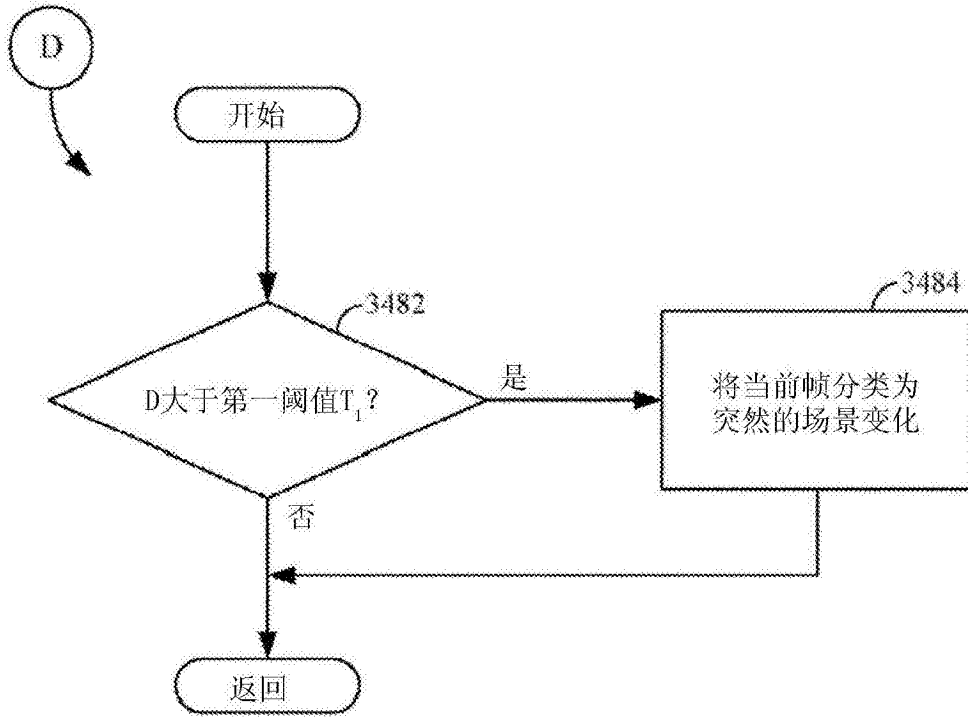


图34

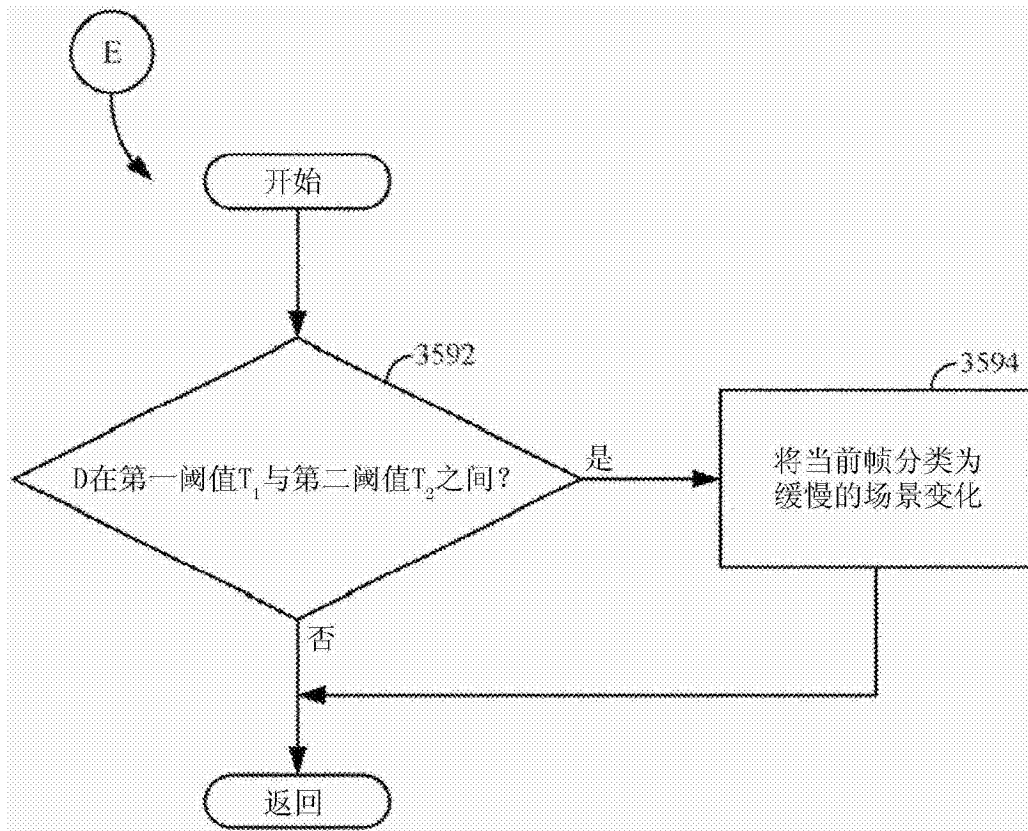


图35

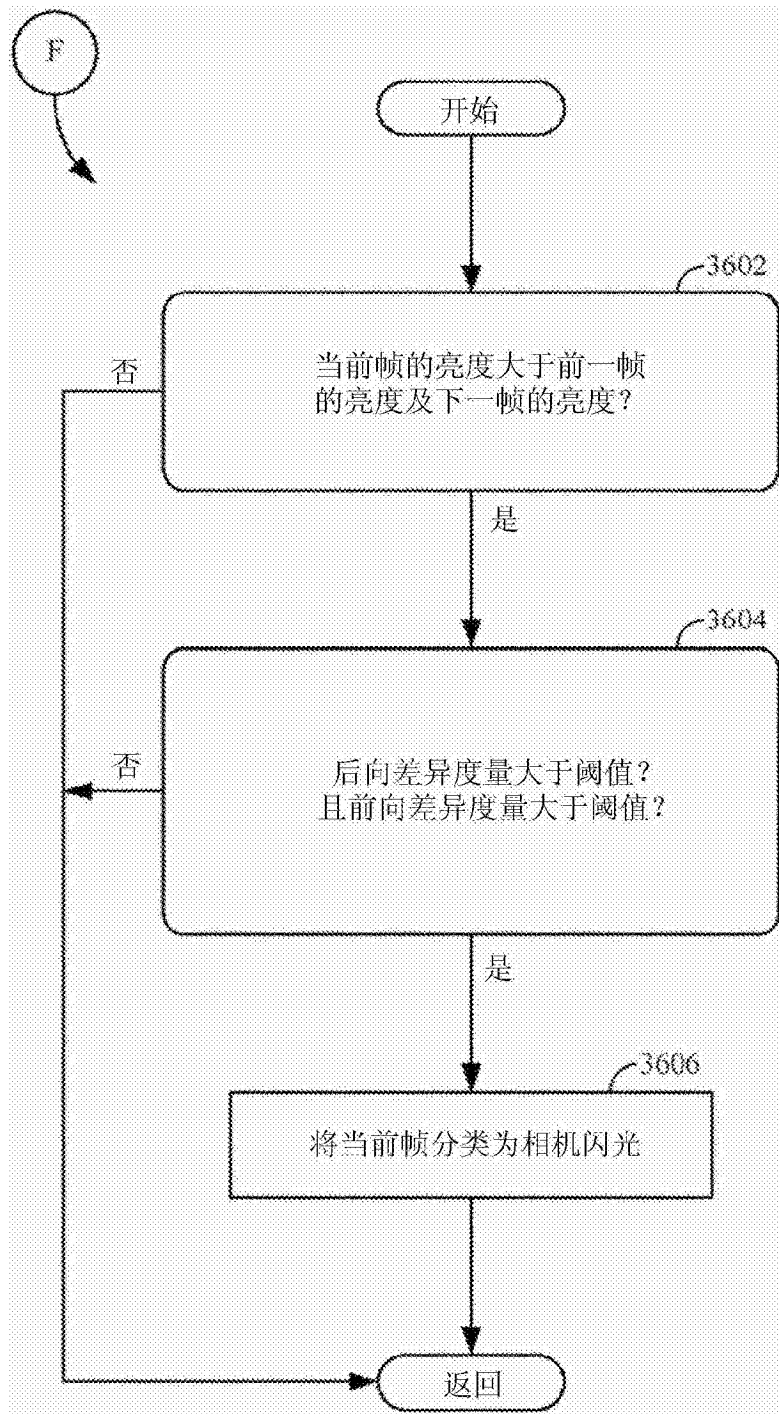


图36

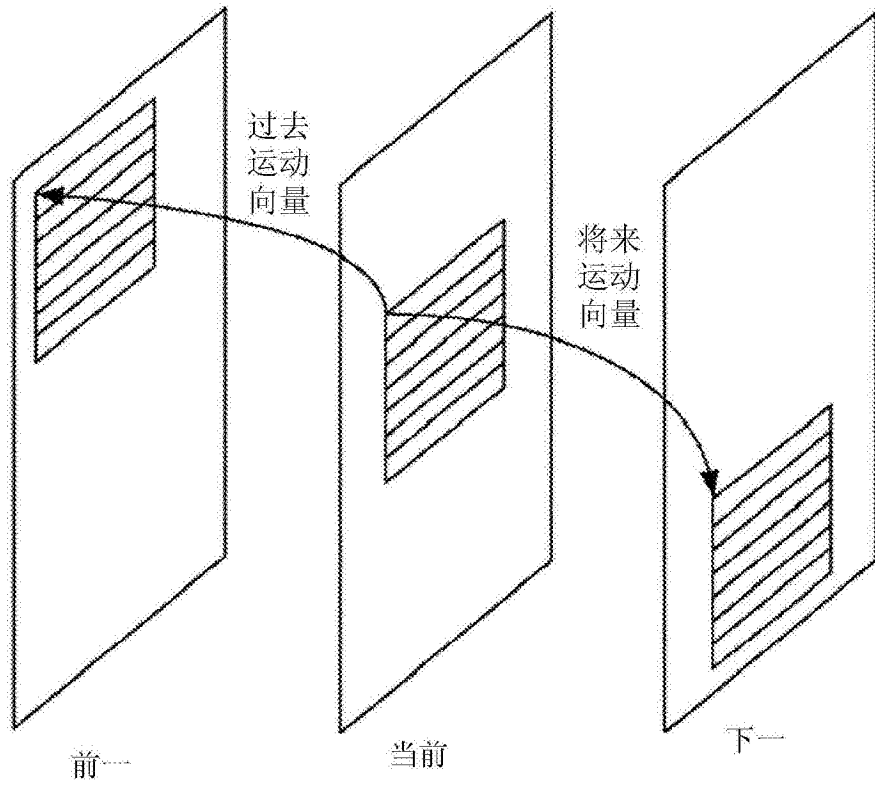


图37

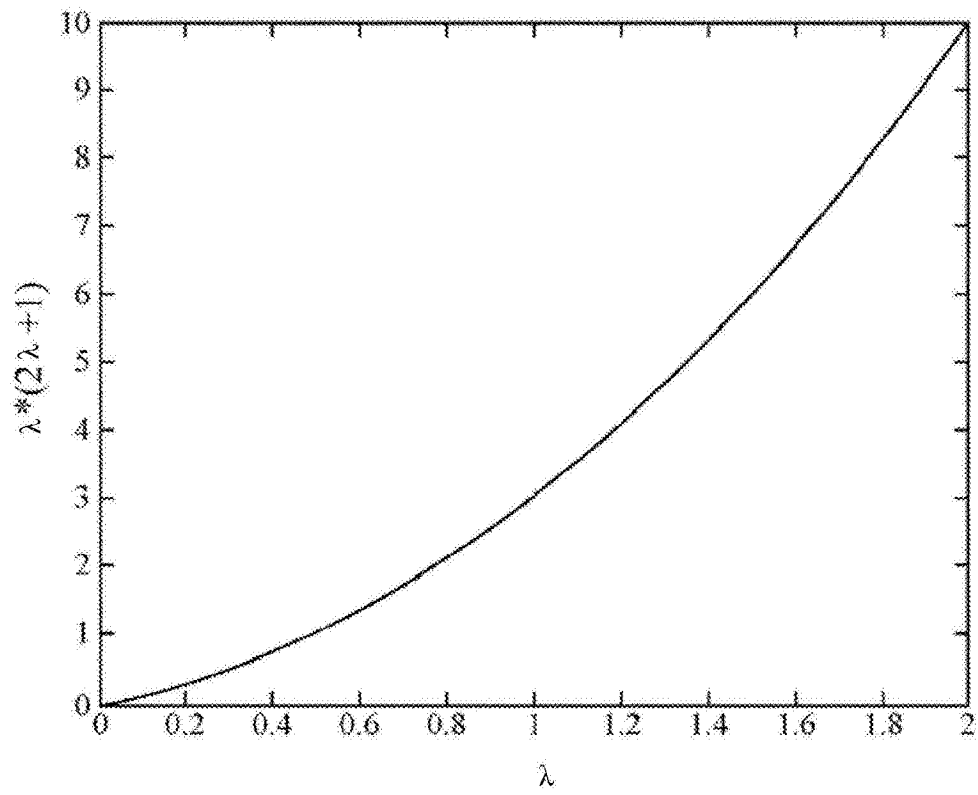


图38

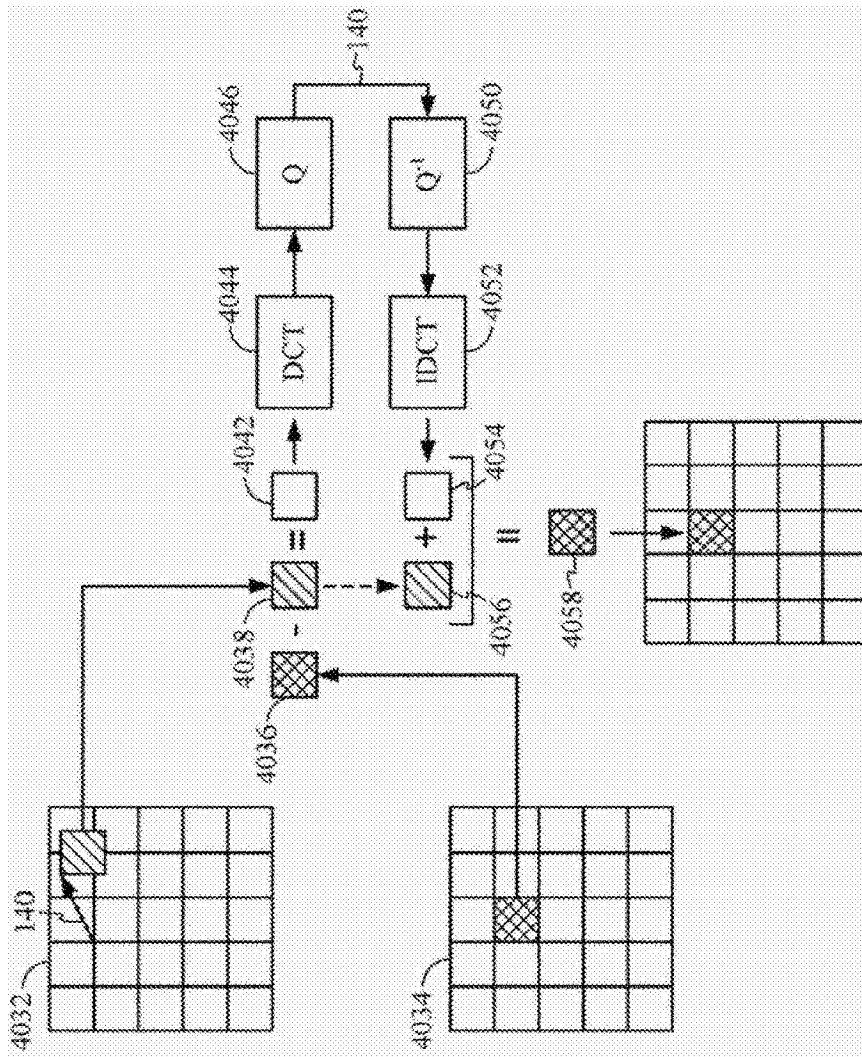


图39

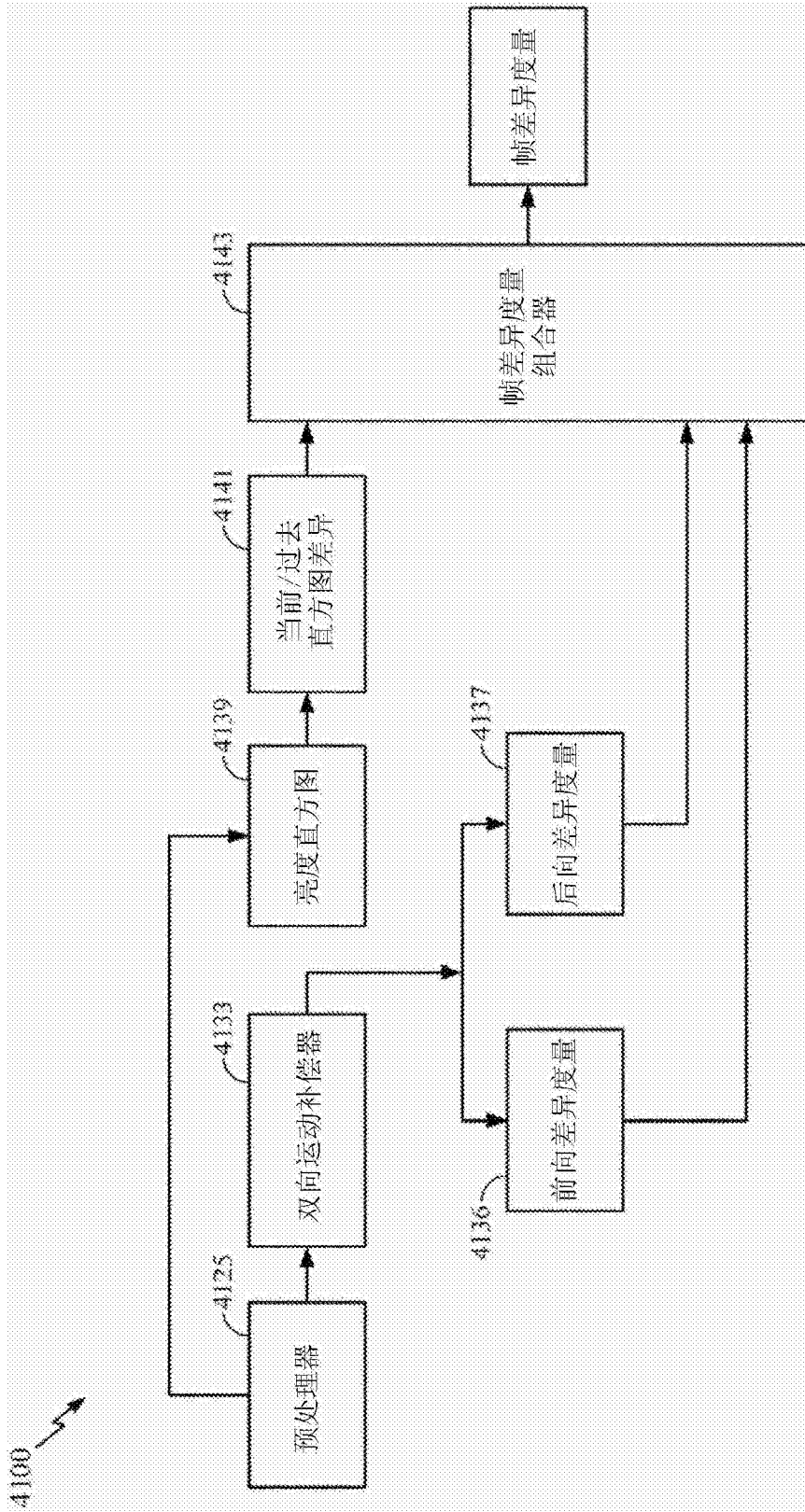


图40

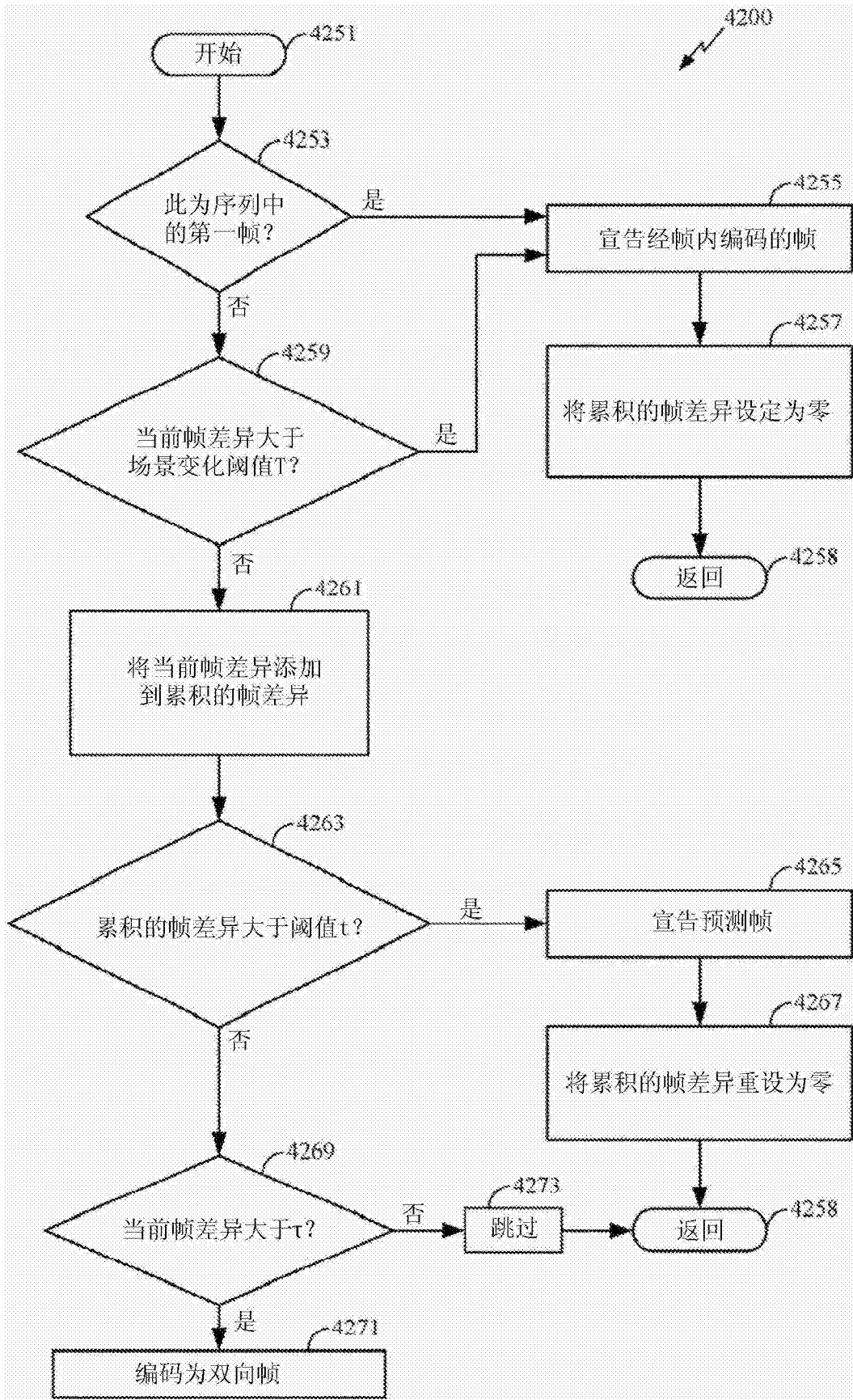


图41

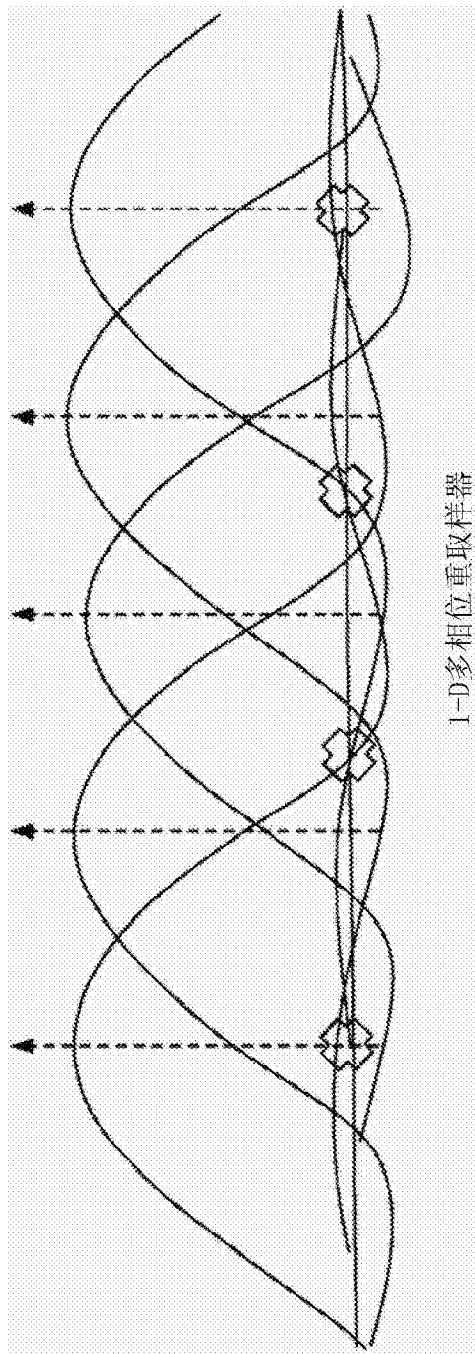


图42

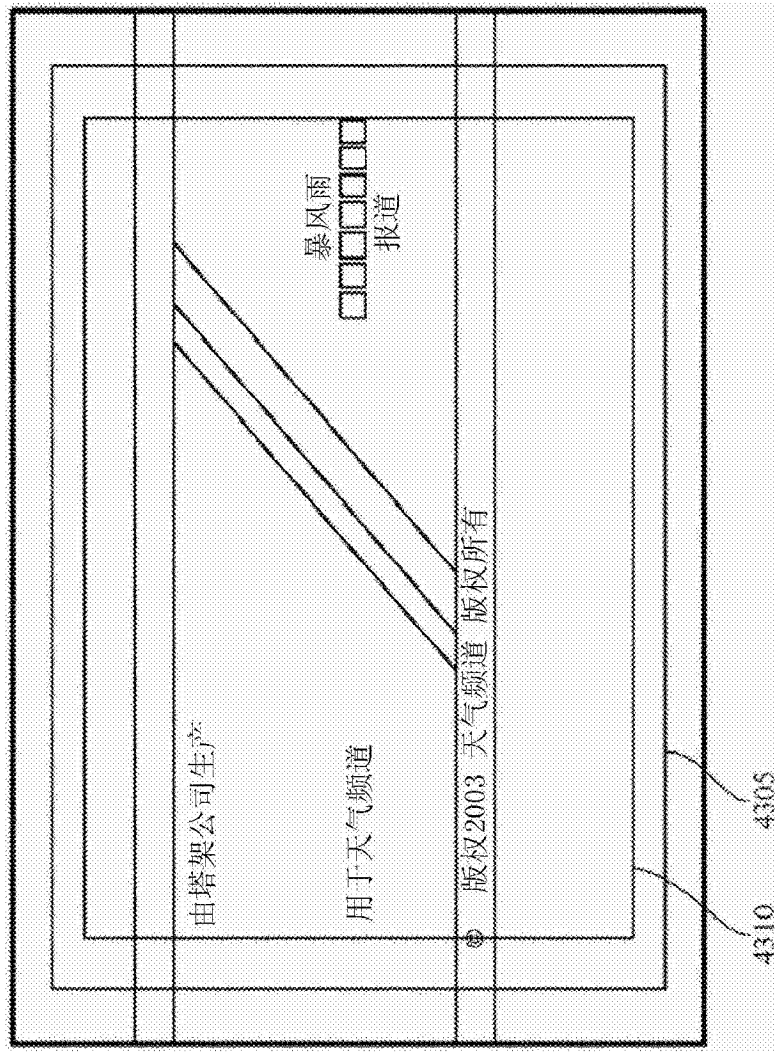


图43

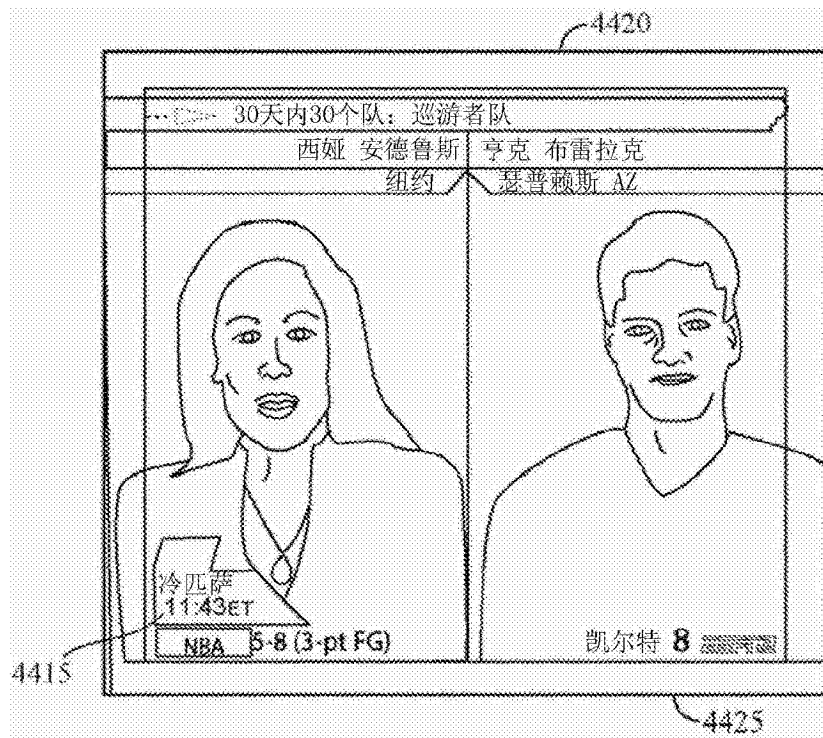


图44