



(12) 发明专利

(10) 授权公告号 CN 102138333 B

(45) 授权公告日 2014. 09. 24

(21) 申请号 200980134021. X

(22) 申请日 2009. 08. 28

(30) 优先权数据

61/092, 967 2008. 08. 29 US

61/192, 612 2008. 09. 19 US

(85) PCT国际申请进入国家阶段日

2011. 02. 28

(86) PCT国际申请的申请数据

PCT/US2009/004924 2009. 08. 28

(87) PCT国际申请的公布数据

W02010/024938 EN 2010. 03. 04

(73) 专利权人 汤姆逊许可公司

地址 法国布洛涅 - 比扬古市

(72) 发明人 倪泽峰 田栋

西塔拉姆·布哈加瓦泰 琼·拉奇

(74) 专利代理机构 北京市柳沈律师事务所

11105

代理人 吕晓章

(51) Int. Cl.

H04N 13/00 (2006. 01)

(56) 对比文件

CN 1462561 A, 2003. 12. 17, 全文.

Voicu Popescu. "Forward Rasterization: A reconstruction algorithm for image based rendering". 《Department of computer science, Chapel Hill》. 2001, (第TR01-019期), 第4.2部分.

Eddie Cook, Peter Kauff, Thomas Sikora. "Multi-view synthesis: A novel view creation approach for free viewpoint video". 《SIGNAL PROCESSING: IMAGE COMMUNICATION》. 2006, 第21卷(第6期), 第480 - 485页, Fig. 2.

Eddie Cook, Peter Kauff, Thomas Sikora. "Multi-view synthesis: A novel view creation approach for free viewpoint video". 《SIGNAL PROCESSING: IMAGE COMMUNICATION》. 2006, 第21卷(第6期), 第480 - 485页, Fig. 2.

审查员 李雁

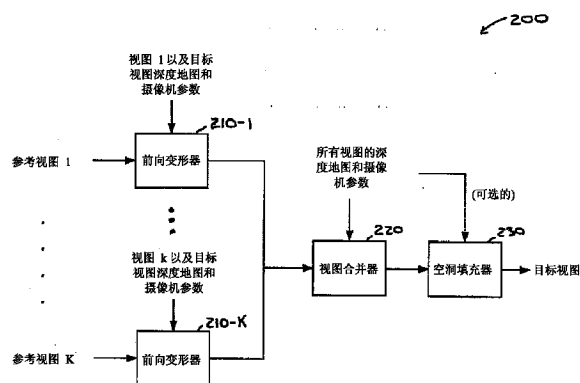
权利要求书2页 说明书14页 附图8页

(54) 发明名称

利用启发式视图合并的视图合成

(57) 摘要

几个实施方式涉及用于3D视频(3DV)应用的具有启发式视图合并的视图合成。根据一个方面,从至少一个参考视图位置到虚拟视图位置变形至少一个参考图像或其一部分,以产生至少一个变形的参考。识别所述至少一个变形的参考中的第一候选像素和第二候选像素。该第一候选像素和第二候选像素是来自所述虚拟视图位置的虚拟图像中目标像素位置的候选像素。基于所述第一和第二候选像素的值来确定所述目标像素位置处的像素值。



1. 一种视图合成方法,包括:

从第一参考视图位置到虚拟视图位置变形(705)至少一部分第一参考图像,以产生第一变形的参考;

从第二参考视图位置到虚拟视图位置变形至少一部分第二参考图像,以产生第二变形的参考,其中,所述第二参考视图位置不同于所述第一参考视图位置;

识别(710)所述第一变形的参考中的第一候选像素,以及识别所述第二变形的参考中的第二候选像素,所述第一候选像素和所述第二候选像素是来自所述虚拟视图位置的虚拟图像中目标像素位置的候选像素;以及

基于所述第一和第二候选像素的值来确定(725)所述目标像素位置处的像素值,其中,确定所述值包括对所述第一和第二候选像素的每一个使用加权因子对来自所述第一和第二候选像素值的目标像素值进行线性插值。

2. 根据权利要求1的方法,其中,所述加权因子是基于第一距离和第二距离确定的,该第一距离是所述第一参考视图位置与所述虚拟视图位置之间的距离,该第二距离是所述第二参考视图位置与所述虚拟视图位置之间的距离(725)。

3. 根据权利要求1的方法,其中,所述加权因子是根据所述第一候选像素与所述目标像素位置之间的距离进一步确定的。

4. 根据权利要求1的方法,其中,所述加权因子是根据与所述第一候选像素相关联的深度进一步确定的。

5. 根据权利要求1的方法,其中识别第一候选像素包括基于所述第一候选像素与所述目标像素位置之间的距离来识别所述第一候选像素(710)。

6. 根据权利要求1的方法,其中识别第一候选像素包括基于与所述第一候选像素相关联的深度来识别所述第一候选像素(710)。

7. 根据权利要求1的方法,其中识别第一候选像素包括从所述第一变形的参考的多个像素选择所述第一候选像素,以及该多个像素都在所述目标像素位置的临界值距离内,并且所述第一候选像素是基于最靠近摄像机的所述第一候选像素的深度来选择的(710)。

8. 根据权利要求7的方法,还包括基于另一像素的深度是否在所述第一候选像素的深度临界值内,从所述多个像素选择该另一像素作为另一候选像素,并且其中确定所述目标像素位置处的像素值还基于该另一候选像素的值(725)。

9. 根据权利要求1的方法,还包括:

在所述虚拟图像中的所有子像素位置插入相应的新目标像素,以获得多个相应的新目标像素(1055);

基于与最靠近摄像机的所述第一候选像素和所述第二候选像素的每个相关联的相应的深度,来估计所述多个相应的新目标像素的每个的相应值(1060);以及

利用下采样来产生对应于所述虚拟图像的最终虚拟视图(1065)。

10. 根据权利要求9的方法,其中所述插入包括进一步在所述虚拟图像中的所有剩余的子像素位置,插入另一相应的新目标像素。

11. 一种视图合成设备,包括:

从第一参考视图位置到虚拟视图位置变形至少一部分第一参考图像,以产生第一变形的参考的装置;

从第二参考视图位置到虚拟视图位置变形至少一部分第二参考图像,以产生第二变形的参考的装置,其中,所述第二参考视图位置不同于所述第一参考视图位置;

识别所述第一变形的参考中的第一候选像素,以及识别所述第二变形的参考中的第二候选像素的装置,所述第一候选像素和所述第二候选像素是来自所述虚拟视图位置的虚拟图像中目标像素位置的候选像素;以及

基于所述第一和第二候选像素的值来确定所述目标像素位置处的像素值的装置,其中,确定所述值包括对所述第一和第二候选像素的每一个使用加权因子对来自所述第一和第二候选像素值的目标像素值进行线性插值。

12. 根据权利要求 11 的设备,其中所述设备包括一个或多个编码器 (310) 或解码器。

利用启发式视图合并的视图合成

[0001] 相关申请的交叉参考

[0002] 本申请要求 (1) 2008 年 9 月 19 日提交的美国临时申请 61/192,612 号标题为“用于 3DV 应用的带有边界泼溅与启发式视图合并的视图合成”以及 (2) 2008 年 8 月 29 日提交的美国临时申请 61/092,967 号标题为“用于 3D 视频 (3DV) 应用的带有适应性泼溅的视图合成”的权益。这两篇美国临时申请的内容全部并入作为参考。

技术领域

[0003] 所描述的实施方式涉及编码系统。各种特别的实施方式涉及用于 3D 视频 (3DV) 应用的具有启发式视图合并的视图合成。

背景技术

[0004] 三维视频 (3DV) 是一种包括多视图视频的编码表示和深度信息并且目的在于例如在接收器处生成高质量的 3D 渲染的新的架构。这使得自动立体显示、自由视点应用和立体显示的 3D 视觉体验成为可能。需要更进一步的技术来生成额外的视图。

发明内容

[0005] 根据一个总体的方面,从至少一个参考视图位置到虚拟视图位置变形至少一个参考图像或其一部分,以产生至少一个变形的参考。识别所述至少一个变形的参考中的第一候选像素和第二候选像素。该第一候选像素和第二候选像素是来自所述虚拟视图位置的虚拟图像中目标像素位置的候选像素。基于所述第一和第二候选像素的值来确定所述目标像素位置处的像素值。。

[0006] 在下面的附图和说明书部分详细解释一个或多个实施方式的细节。即使是以一种特殊方式来描述实施方式,应当清楚的是,可以以各种方式配置或者体现实施方式。例如,一种实施方式可以用方法的方式来执行,或者体现为诸如例如是被配置来执行一组操作的设备或者是存储用于执行一组操作的指令的设备之类的设备,或者体现为信号。从接下来的详细描述结合附图和权利要求会容易理解其它方面和特征。

附图说明

[0007] 图 1A 是未矫正的视图的合成的实施方式的图示。

[0008] 图 1B 是矫正的视图的合成的实施方式的图示。

[0009] 图 2 是视图合成器实施方式的图示。

[0010] 图 3 是视频发送系统实施方式的图示。

[0011] 图 4 是视频接收系统实施方式的图示。

[0012] 图 5 是视频处理装置实施方式的图示。

[0013] 图 6 是用于发送和接收具有深度信息的多视图视频的系统的实施方式的图示。

[0014] 图 7 是视图合成过程的实施方式的图示。

[0015] 图 8 是用于矫正的视图的视图合并过程的实施方式的图示。

[0016] 图 9 是由 3D 点 $0r_i-P_i-O_s$ 确定的角度的图示。

[0017] 图 10A 是用于矫正的视图的上采样的实施方式的图示。

[0018] 图 10B 是基于上采样和 Z 缓冲合并过程的实施方式的图示。

具体实施例

[0019] 一些三维视频 (3DV) 应用对输入视图有严格的限制。输入视图通常必须得到很好的矫正,使得一维 (1D) 视差 (disparity) 可以描述如何将像素从一个视图置换到另一个视图。

[0020] 基于深度图像的渲染 (DIBR) 是一种使用从多个校准后的摄像机捕捉的多个图像以及相关连的逐像素深度信息的视图合成技术。从概念上讲,这种视图生成方法可以理解为一个两步骤的过程:(1)3D 图像变形;和(2)重建和再取样。关于 3D 图像变形,使用深度数据和相关连的摄像机参数将来自于参考图像的像素反投影到适当的 3D 位置,并重新将它们投影到新的图像空间。关于重建和再取样,同样涉及确定合成视图中的像素值。

[0021] 所述渲染方法可以是基于像素的泼溅或者是基于网格的(三角形的)。对于 3DV,逐像素深度典型的是用诸如立体之类的被动式计算机视觉技术来估计,而不是从激光测距扫描或者计算机图形模型生成的。因此,对于 3DV 中的实时处理,在只给定有噪声的深度信息时,基于像素的方法应当是优选的,以避免复杂性和生成耗费计算的网格,因为鲁棒的 3D 三角剖分 (triangulation) (表面重建) 是困难的几何问题。

[0022] 现有的泼溅算法达到了一些令人印象非常深刻的结果。但是,它们被设计来与高精度的深度一起使用,对低质量深度而言,它们可能不是充分的。另外,很多现有的算法将诸如逐像素垂直面或者三维点云之类 3DV 中不存在的技术方面视为理所当然。这样,需要新的合成方法来处理这些具体问题。

[0023] 给定深度信息和摄像机参数,就容易将参考像素变形到合成的视图上。最突出的问题是,如何从变形的参考视图像素估计目标视图中的像素值。图 1A 和图 1B 图示的是这个基本问题。图 1A 显示的是未矫正的视图合成 100。图 1B 显示的是矫正后的视图合成 150。在图 1A 和图 1B 中,字母“X”代表目标视图中要估计的像素,圆和方块表示从不同的参考视图变形的像素,其中,不同的形状表示不同的参考视图。

[0024] 一种简单的方法是,将变形的样本像素舍入到目的视图中其最接近的像素位置。当多个像素映射到合成视图中的同一个位置时,Z-缓冲是一种典型的解决方案,即,选择最靠近摄像机的那个像素。这种方案(舍入最近的像素位置)经常会在略为欠采样的表面中,尤其是沿着物体的边界处导致针孔。处理这个针孔问题最常见的方法是将参考视图内的一个像素映射为目标视图内的几个像素。这个过程称作泼溅。

[0025] 如果将一个参考像素映射到目标视图中的多个周边目标像素上,可以消除大部分的针孔。但是,会丢失某些图像细节。当使用透明的泼溅类型的重建核心时,同样要在消除针孔和丢失细节之间进行权衡。问题是:“我们如何来控制泼溅的程度?”例如,对于每个变形的像素,我们是该将它映射到它所有的周边目标像素上,还是仅将它映射到最靠近它的一个像素上?文献很大程度上并未处理这个问题。

[0026] 当应用多个参考视图时,一种常见的方法会单独处理来自于每个参考视图的合

成,然后将多个合成视图合并在一起。问题是如何来合并它们,例如,可以使用某类加权方案。例如,可以基于角距离、图像分辨率等等之类,给不同的参考视图应用不同的权重。注意,应当以对有噪声的深度信息鲁棒的方式来处理这些问题。

[0027] 使用 DIBR,可以从捕捉到的视图(在上下文中也称作参考视图)产生虚拟视图。产生虚拟视图是一项艰难的任务,尤其是当输入的深度信息有噪声并且其它诸如场景的 3D 表面属性之类的场景信息未知的时候。

[0028] 其中一个最困难的问题经常是,在对参考视图内的样本像素变形后,如何估计合成视图内每个像素的值。例如,对于每个目标合成像素,应当利用什么样的参考像素,以及如何组合它们。

[0029] 在至少一种实施方式中,我们提议用一种用于 3DV 应用的具有启发式视图合并的视图合成框架。发明人已经注意到,在涉及产生虚拟视图的 3DV 应用中(例如,使用 DIBR),这样产生虚拟视图是一项挑战性的任务,尤其是当输入的深度信息具有噪声并且诸如场景的 3D 表面属性之类的其它场景信息未知的时候。发明人进一步注意到,在产生这种虚拟视图时突出的问题是在对参考视图中的样本像素进行变形后如何估计合成视图中每个像素的值。例如,对于每个目标合成的像素,应该应用什么参考像素,以及如何组合这些参考像素?

[0030] 所以,在至少一种实施方式中,提供了基于例如其深度信息、其变形的 2D 图像位置和摄像机参数来合并多个变形的参考像素的启发式方法。当然,本发明的方法不仅限于以上方法,其他选项(信息、位置、参数等)也可以用于合并多个变形的参考像素,而保持本发明原理的精神。所建议的方案不限制多少参考视图被作为输入,并且无论摄像机视图是否被矫正都可被应用。

[0031] 在至少一种实施方式中,允许将单视图的合成和合并组合为单个合并方案。

[0032] 另外,发明人注意到,要从参考视图合成虚拟视图,通常需要三个步骤,即:(1) 前向变形;(2) 混合(单个视图的合成和多个视图的合并);和(3) 填充空洞。

[0033] 至于上文提到的与从参考视图合成虚拟视图有关的三个步骤的变形步骤,关于如何处理所述变形结果,基本上存在两种可以考虑的选项,即,合并和混合。

[0034] 关于合并,可以考虑完全变形每个视图,以给每个参考视图形成最终的变形视图。然后,可以将这些最终的变形视图“合并”,得到单个的真正最终的合成视图。“合并”会涉及,例如在 N 个候选像素(假设有 N 个最终的变形视图)之间进行挑选或者以某种方式将它们组合。当然,可以理解,用于确定目标像素值的候选像素数目不需要和变形视图的数目相同。也就是说,多个候选像素(或者一个也没有)可以来自于单个视图。

[0035] 关于混合,可以仍然变形每个视图,但是不为每个参考视图形成最终变形的视图。通过不进行到最终过程,当混合时,可以保留更多的可选项。这是有优点的,因为在某些情况下不同的视图可以为合成的目标视图的不同部分提供最佳信息。因而,混合可提供灵活性以在每个像素处从不同的视图选择信息的正确组合。因此,可以认为,合并是两步骤混合的特例,其中,首先分别处理来自于每个视图的候选,然后对处理结果进行组合。

[0036] 再次参考图 1A,可以用图 1A 来显示到典型的混合操作的输入,因为图 1A 包括从不同的参考视图变形的像素(分别是圆形和方块)。相反,对于典型的合并应用,人们会期望只看见圆形或者方块,因为会典型地分开变形每个参考视图然后进行处理,为各个参考视

图形成最终的变形视图。然后在典型的合并应用中组合多个参考视图的最终变形视图。

[0037] 返回到混合,作为一种涉及混合的可能的可选项 / 考虑,可能不执行泼溅,因为还不想填满所有的空洞。这些及其它可选项可以由本领域和相关领域的普通技术人员所确定,而保持本发明原理的精神。

[0038] 因此,可以理解,本发明原理的一个或多个实施例针对的是合并,而本发明原理的其它实施例可能针对的是混合。当然,进一步的实施例可以涉及合并和混合的组合。本申请中所讨论的特征和构思通常应用在合并和混合两种情况下,即使在上下文中只讨论合并或者混合中的一种。给定此处所提供的本发明原理的教导,本领域及相关技术领域的普通技术人员会容易理解与合并和 / 或混合相关的、保持本发明原理的精神的各种应用。

[0039] 应当理解的是,本发明通常涉及通信系统,更特别的是涉及无线系统,例如地面广播、手机、无线 - 保真 (Wi-Fi)、卫星等等。应当进一步理解的是,本发明原理可以实现在例如编码器、解码器、预处理器、后处理器和接收器 (它可以包括前述的一种或多种) 中。例如,在需要产生虚拟图像来进行编码时,本发明原理可以用在编码器中。关于编码器的进一步例子,这样一种编码器可以用于合成虚拟视图,用于从虚拟视图位置编码实际的图像,或者从靠近虚拟视图位置的视图位置编码图像。在涉及两个参考图像的实施方式中,可以对两个参考图像以及对应于虚拟视图的虚拟图像都进行编码。当然,给定此处所提供的本发明原理的教导,在保持本发明原理的精神的同时,本领域及相关领域的普通技术人员会考虑这些和各种其它应用,以及可应用本发明原理的对前面描述的应用的各种变形。

[0040] 另外,要理解的是,虽然此处是根据 H. 264/M PEG-4 AVC (AVC) 标准来描述一个或多个实施例的,但本发明原理不完全局限于同样的实施例,因此,给定此处所提供的本发明原理的教导,在保持本发明原理的精神的同时,可以容易地将本发明原理应用到多视图视频编码 (MVC)、当前和将来的 3DV 标准、以及其它视频编码标准、规范和 / 或建议,而保持本发明原理的精神。

[0041] 注意,“泼溅”指的是,将来自参考视图的一个变形的像素映射到目标视图中的几个像素的过程。

[0042] 注意,“深度信息”是指与深度有关的各类信息的通用术语。一种类型的深度信息是“深度地图”,它通常指的是逐像素深度图像。其它类型的深度信息包括,例如,为每个编码块而不是为每个编码的像素使用单个的深度值。

[0043] 图 2 显示的是根据本发明原理的实施例的可以应用本发明原理的示例的视图合成器 200。视图合成器 200 包括前向变形器 210-1 至 210-K、视图合并器 220 和空洞填充器 230。前向变形器 210-1 至 210-K 的各个输出与视图合并器 220 的第一输入以信号通信的方式进行连接。视图合并器 220 的输出与空洞填充器 230 的第一输入以信号通信的方式进行连接。前向变形器 210-1 至 210-k 的第一各个输入可以用于作为视图合成器 200 的输入,用于接收 1 至 K 的各个参考视图。前向变形器 210-1 至 210-K 的第二各个输入可用于作为视图合成器 200 的输入,用于分别接收视图 1 和与它对应的目标视图深度地图和摄像机参数直至视图 K 和与它对应的目标视图深度地图和摄像机参数。视图合并器 220 的第二输入可用于作为所述视图合成器的输入,用于接收所有视图的深度地图和摄像机参数。空洞填充器 230 的第二 (可选的) 输入可用于作为视图合成器 200 的输入,用于接收所有视图的深度地图和摄像机参数。空洞填充器 230 的输出可用于作为视图合成器 200 和输出,用于

输出目标视图。

[0044] 视图合并器 220 可以执行各种功能和操作中的一种或多种。例如,在一个实施例中,视图合并器 220 识别在至少一个变形的参考中的第一候选像素和第二候选像素,该第一候选像素和第二候选像素是来自虚拟视图位置的虚拟图像中目标像素位置的候选像素。此外,在实施例中,视图合并器 220 还基于第一和第二候选像素的值来确定目标像素位置处的像素的值。

[0045] 图 2 中的元件,例如前向变形器 210 和视图合并器 220 可以以各种方式实施。例如,可以在通用计算机上或者诸如视频编码器之类的专用机器上,或者在特殊用途的集成电路(诸如专用的集成电路(ASIC))上执行前向变形或视图合并功能的软件算法。这些实施也可以应用软件、硬件和固件的组合。前向变形和视图合并的通用功能对于本领域普通技术人员而言是周知的。这些通用功能可以按本申请中所描述的方式进行修改,以执行例如本申请所描述的前向变形和视图合并操作。

[0046] 图 3 显示可以根据本发明原理的实施方式的应用本发明原理的示例的视频发送系统 300。视频发送系统 300 可以是,例如用于使用诸如卫星、缆线、电话线或者地面广播之类的多种介质中的任何一种介质来发送信号的头端或者发送系统。可以通过因特网或者其它网络进行所述发送。

[0047] 视频发送系统 300 能够生成并输送使用视图间跳跃模式编码的视频内容。这是通过生成包括深度信息或者包括能够用于在接收端合成所述深度信息的信息的编码的信号来实现的,所述接收端例如可以具有解码器。

[0048] 视频发送系统 300 包括编码器 310 和能够发送编码的信号的发送器 320。编码器 310 接收视频信息,并使用视图间跳跃模式从所接收的视频信息生成具有深度的编码的信号。编码器 310 可以是,例如 AVC 编码器。编码器 310 可以包括子模块,具有例如用于接收各种信息片并将它们组合为用于存储或者发送的结构化形式的组合单元。各种信息片可以包括,例如编码的或者解码的视频、编码的或者解码的深度信息以及诸如运动矢量、编码模式指示符和语法元素之类的编码或者解码的元素。

[0049] 发送器 320 可以是,例如适用于发送具有代表编码图像的一个或多个比特流以及/或者与其相关的信息的程序信号。典型的发送器执行诸如提供纠错编码、在信号中插入数据、随机化信号中的能量以及将信号调制到一个或多个载波上中的一种或多种的功能。所述发送器可以包括天线(未示)或者与天线通过接口连接。相应地,发送器 320 的实施方式可以包括或者被限制为调制器。

[0050] 图 4 显示的是根据本发明原理的实施例的可以应用本发明原理的示例视频接收系统 400。可以将视频接收系统 400 配置为,接收例如通过卫星、缆线、电话线或者地面广播之类的各种介质传输的信号。所述信号可以从因特网或者其它网络接收。

[0051] 视频接收系统 400 可以是,例如手机、计算机、机顶盒、电视机或者其它接收编码的视频并提供例如用于显示给用户或者用于存储的解码的视频的装置。因此,视频接收系统 400 可以将它的输出提供给,例如电视屏幕、计算机显示器、计算机(用于存储、处理或者显示)或者其它存储器、处理或者显示装置。

[0052] 视频接收系统 400 能够接收和处理包括视频信息的视频内容。视频接收系统 400 包括能够接收诸如本申请的实施方式中所描述的信号之类的编码的信号的接收器 410,以

及能够解码所接收的信号的解码器 420。

[0053] 接收器 410 可以是,例如适用于接收具有代表编码的图像的多个比特流的程序信号。典型的接收器执行诸如,接收调制的和编码的数据信号、从一个或者多个载波解调出所述数据信号、去随机化 (de-randomize) 所述信号中的能量、去插值 (de-interleave) 信号中的数据以及对信号进行纠错解码中的一种或多种功能。接收器 410 可以包括天线 (未示) 或者与天线通过接口相接。接收器 410 的实施方式可以包括解调器或者被限制为解调器。

[0054] 解码器 420 输出包括视频信息和深度信息的视频信号。解码器 420 可以是例如 AVC 解码器。

[0055] 图 5 显示的是根据本发明原理的实施例的可以应用本发明原理的示例的视频处理装置 500。视频处理装置 500 可以是,例如机顶盒或者是其它接收编码的视频并提供例如用于显示给用户或者用于存储的解码的视频的装置。因此,视频处理装置 500 可以将它的输出提供给电视机、计算机显示器或者计算机或者其它处理装置。

[0056] 视频处理装置 500 包括前端 (FE) 装置 505 和解码器 510。前端装置 505 可以是例如一个接收器,适用于接收具有代表编码的图像的多个比特流的程序信号,以及从所述多个比特流中选择一个或多个比特流进行解码。典型的接收器执行例如接收调制的和编码的数据信号、解调所述数据信号、解码所述数据信号的一种或多种编码 (例如,信道编码和 / 或源编码) 以及 / 或者对所述数据信号进行纠错其中的一种或多种的功能。前端装置 505 可以从例如天线 (未示) 接收程序信号。前端装置 505 将所接收的数据信号提供给解码器 510。

[0057] 解码器 510 接收数据信号 520。数据信号 520 可以包括,例如一个或多个与先进视频编码 (AVC)、可伸缩的视频编码 (SVC) 或者多视图视频编码 (MVC) 兼容的流。

[0058] AVC 更具体的是指,现有的国际标准化组织 / 国际电工委员会 (ISO/IEC) 运动图像专家组 -4 (MPEG-4) 部分 10 先进视频编码 (AVC) 标准 / 国际电信联盟电信部分 (ITU-T) H. 264 建议书 (之后,“H. 264/MPEG-4 AVC 标准”或者其变形,诸如“AVC 标准”或者简单地是“AVC”)。

[0059] MVC 更具体的是指 AVC 标准的多视图视频编码 (“MVC”) 扩展 (附录 H),指的是 H. 264/MPEG-4 AVC、MVC 扩展 (“MVC 扩展”或者简单地是“MVC”)。

[0060] SVC 更具体的指的是指 AVC 标准的可伸缩视频编码 (“SVC”) 扩展 (附录 G),称作 H. 264/MPEG-4 AVC、SVC 扩展 (“SVC 扩展”或者简单地是“SVC”)。

[0061] 解码器 510 解码所接收到的信号 520 的全部或者一部分,并提供解码的视频信号 530 作为输出。解码的视频 530 被提供给选择器 550。装置 500 还包括接收用户输入 570 的用户界面 560。用户界面 560 基于用户输入 570 提供图像选择信号 580 给选择器 550。图像选择信号 580 和用户输入 570 表明用户希望显示多个图像、序列、可伸缩版本、视图或者可用的解码的数据的其它选项中的哪一个。选择器 550 提供所选择的图像作为输出 590。选择器 550 使用图像选择信息 580 选择要提供解码的视频 530 中的哪个图片以作为输出 590。

[0062] 在各种实施方式中,选择器 550 包括用户界面 560,在其它实施方式中不需要用户界面 560,因为选择器 550 直接接收用户输入 570,不单独执行界面功能。选择器 550 例如可以实施为软件或者集成电路。在一种实施方式中,选择器 550 与解码器 510 合并在一起,

在另一种实施方式中,解码器 510、选择器 550 和用户界面 560 全部集成在一起。

[0063] 在一个应用中,前端 505 接收广播的各种各样的电视节目并选择一个来处理。基于用户输入的期望观看的频道来选择一个节目。虽然图 5 没有显示用户给前端装置 505 的输入,但是前端装置 505 接收用户输入 570。前端 505 接收所述广播,并通过解调所述广播波谱的相关部分以及解码所解调的节目的任何外部编码来处理所期望的节目。前端 505 提供解码的节目给解码器 510。解码器 510 是包括装置 560 和 550 的集成单元。因此,解码器 510 接收用户输入,该用户输入是用产提供的表示所期望观看的节目中视图的指示。解码器 510 解码所选择的视图,以及来自于其它视图的任何所需要的参考图像,并提供用于在电视上(未示)显示的解码的视图 590。

[0064] 继续上述的应用,用户可能会期望切换所显示的视图,并可能会提供新的输入给解码器 510。在接收到来自于用户的“改变视图”之后,解码器 510 解码旧的视图和新的视图二者,以及在旧的视图和新的视图之间的所有视图。也就是说,解码器 510 解码物理位置上位于照取新的视图的摄像机和照取旧的视图的摄像机之间的摄像机所照取的所有视图。前端装置 505 还接收用于识别旧的视图、新的视图以及它们之间的视图的信息。这样的信息例如可以由具有与所述视图的位置有关的信息的控制器(图 5 中未示)或者解码器 510 来提供。其它实施方式可以使用一个前端装置,该前端装置具有与它集成在一起的控制器。

[0065] 解码器 510 提供所有这些解码的视图作为输出 590。后期处理器(图 5 中未示)在视图之间进行插值,以提供从所述旧的视图到所述新的视图的顺利过渡,并将这个过渡显示给用户。在过渡到新的视图之后,所述后期处理器(通过一个或多个未示的通信链路)告知解码器 510 和前端装置 505 只需要新的视图。之后,解码器 510 仅提供所述新的视图作为输出 590。

[0066] 系统 500 可以用于接收一系列图像的多个视图,以及用于呈现供显示的单个视图,并在各种视图之间以顺畅的方式进行切换。所述顺畅的方式可以包括在视图之间进行插值,以移到另一个视图。另外,系统 500 可以允许用户转动物体或者场景,或者相反,允许用户看物体或者场景的三维呈现。转动所述物体,例如可以对应于从视图移到视图,在视图之间进行插值,以获得在视图之间的顺畅的过渡,或者简单地获得三维呈现。也就是说,用户可以“选择”一个插入的视图作为要显示的“视图”。

[0067] 图 2 的元件可以被包含在图 3-5 中的各种位置。例如,图 2 的一个或多个元件可以位于编码器 310 和解码器 420 中。作为进一步的例子,视频处理装置 500 的实施方式可以将图 2 的一个或多个元件包括在图 5 的讨论中所提到的解码器 510 或者后期处理器中,从而在接收到的视图之间进行插值。

[0068] 返回到本发明原理和这些原理所应用的环境的描述,要理解的是,有利的是,本发明原理可以应用在 3D 视频(3DV)。3D 视频是新的架构,其包括多视图视频的编码表示和深度信息,目的是在接收器处生成高质量的 3D 渲染。这样使得自动多视觉显示的 3D 视觉体验成为可能。

[0069] 图 6 显示的是根据本发明原理的实施例的可以应用本发明原理的用于发送和接收具有深度信息的多视图视频的示例系统 600。在图 6 中,视频数据由实线表示,深度数据由短划线表示,元(meta)数据由虚线表示。系统 600 可以是,例如但是不局限于自由视点电视系统。在发送器侧 610,系统 600 包括三维(3D)内容制作器 620,它具有多个输入,用于

分别从多个源接收视频、深度和元数据中的一种或多种。这样的源可以包括,但是不局限于立体摄像机 611、深度摄像机 612、多摄像机设置 613 和 2- 维 /3- 维 (2D/3D) 过渡过程 614。可以使用一个或多个网络 630 发送与多视图视频编码 (MVC) 和数字视频广播 (DVB) 相关的视频、深度和的元数据中的一种或多种。

[0070] 在接收器侧 640,基于深度图像的渲染器 650 执行基于深度图像的渲染,以将信号投射到各种类型的显示器上。这种应用场景会造成特定的限制,诸如窄角捕像 (< 20 度)。基于深度图像的渲染器 650 能够接收显示器配置信息和用户喜好。基于深度图像的渲染器 650 的输出可以提供给 2D 显示器 661、M 视图 3D 显示器 662 和 / 或头部跟踪立体显示器 663 中的一个或者多个。

[0071] 图 7 显示的是根据本发明原理的实施例的视图合成的方法 700。在步骤 705 中,从第一参考视图位置到虚拟视图位置变形第一参考图像或其一部分,以产生第一变形的参考。

[0072] 在步骤 710 中,识别第一变形的参考中的第一候选像素。第一候选像素是来自虚拟视图位置的虚拟图像中目标像素位置的候选像素。可以理解步骤 710 可以包括例如基于第一候选像素和目标像素位置之间的距离来识别第一候选像素,其中该距离可包括门限 (例如距离低于临界值)。此外,可以理解步骤 710 可以包括例如基于与第一候选像素相关联的深度来识别第一候选像素。另外,可以理解步骤 710 可以包括例如基于选自第一变形的参考的多个像素的像素 (作为第一候选像素) 的距离来识别第一候选像素,该距离是距目标像素位置的的门限距离,该距离是与摄像机最近的距离。

[0073] 在步骤 715 中,从第二参考视图位置到虚拟视图位置变形第二参考图像或其一部分,以产生第二变形的参考。在步骤 720 中,识别第二变形的参考中的第二候选像素,该第二候选像素是来自虚拟视图位置的虚拟图像中目标像素位置的候选像素。

[0074] 在步骤 725 中,基于第一和第二候选像素值来确定目标像素位置处的像素值。可以理解步骤 725 可以包括插值,包括例如线性插值,第一和第二像素值。此外,可以理解步骤 725 可以包括例如对每个候选像素使用加权因子。这种加权因子例如可以基于摄像机参数来确定,摄像机参数可包括例如第一参考视图位置和虚拟视图位置之间的第一距离,以及第二参考视图位置和虚拟视图位置之间的第二距离。此外,这种加权因子例如可以基于由 3D 点 $0r_i-P_i-0_s$ 所确定的角度来确定 (如以下对于实施例 2 的详细描述)。此外,可以理解步骤 725 还可以基于,根据在第一候选像素的门限深度内的所选像素的深度,选自第一变形的参考的多个像素的另一候选像素的值。

[0075] 在步骤 730 中,编码一个或多个第一参考图像、第二参考图像和虚拟图像。

[0076] 可以理解虽然图 7 的实施例包括第一参考图像和第二参考图像,给定此处所提供的本发明原理的教导,本领域及相关技术领域的普通技术人员会容易理解本发明原理可应用于包括单个参考图像或多于两个参考图像的实施例,同时保持本发明原理的精神。作为可能变化的另一实施例,在单个参考图像的情况下,单个参考视图位置可用于产生第一和第二候选像素,在变形过程中利用一些变化而获得了第一和第二候选像素的不同的值,尽管应用了相同的单个参考视图位置。在包括单个参考图像的其他实施例中,可应用两个或多个 (不同的) 参考视图位置。给定此处所提供的本发明原理的教导,本领域及相关技术领域的普通技术人员会容易理解本发明原理的这些和其他变化,同时保持本发明原理的精

神。

[0077] 如上所述,在至少一个实施例中提供了基于例如深度信息、变形的 2D 图像位置和摄像机参数来混合多个变形的参考像素 / 视图的启发式方法。

[0078] 在 3DV 应用中,由于传输带宽或存储限制,降低了数量的视图和深度地图被发送或存储。由于期望在实际视图之间渲染虚拟视图,基于渲染的深度图像技术 (DIBR) 可用于产生中间的视图。

[0079] 要从参考视图合成虚拟视图,通常需要三个步骤,即:(1) 前向变形;(2) 混合(组合);和(3) 填充空洞。在实施一个实施例中,提供了启发式方案来解决由于有噪声的深度信息所产生的问题。仿真结果显示与现有的 3DV 中的方案相比获得了较高的质量。

[0080] 1. 背景信息 - 前向变形

[0081] 执行视图合成的第一步骤是前向变形,其涉及为所述参考视图中的每个像素找到它在目标视图内的对应位置。3D 图像变形是计算机图像领域内周知的。根据输入视图是否经过矫正,可以使用不同的方程式。

[0082] (a) 未矫正的视图

[0083] 如果我们根据 3D 点的齐次坐标 $P = [x, y, z, 1]^T$ 来定义它,并且它在参考图像平面中的透视投影(即,2D 图像位置)是 $p_r = [u_r, v_r, 1]^T$,然后我们得到:

$$[0084] \quad w_r \cdot p_r = \text{PPM}_r \cdot P, \quad (1)$$

[0085] 其中, w_r 是深度因子, PPM_r 是 3×4 透视投影矩阵,由摄像机参数可知。对应地,我们得到如下的合成的(目标)视图的方程式:

$$[0086] \quad w_s \cdot p_s = \text{PPM}_s \cdot P. \quad (2)$$

[0087] 我们将 PPM_r 的 12 个元素记为 q_{ij} , $i = 1, 2, 3$, 且 $j = 1, 2, 3, 4$ 。从图像点 p_r 及其深度 z , 可以用如下的线性方程式来估计 3D 点 P 的其余两个分量:

$$[0088] \quad \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad (3)$$

[0089] 其中,

$$[0090] \quad b_1 = (q_{14} - q_{34}) + (q_{13} - q_{33})z, \quad a_{11} = u_r q_{31} - q_{11}, \quad a_{12} = u_r q_{32} - q_{12}.$$

$$[0091] \quad b_2 = (q_{24} - q_{34}) + (q_{23} - q_{33})z, \quad a_{21} = v_r q_{31} - q_{21}, \quad a_{22} = v_r q_{32} - q_{22}.$$

[0092] 注意,将所述参考视图中的每个像素的输入深度水平量化为 3DV 中的 8 比特(即, 256 个水平,其中较大的值意味着更靠近摄像机)。用下面的公式将变形过程中所用的深度因子 z 直接链接到它的输入深度水平 Y :

$$[0093] \quad z = \frac{1}{\frac{Y}{255} \left(\frac{1}{Z_{\text{near}}} - \frac{1}{Z_{\text{far}}} \right) + \frac{1}{Z_{\text{far}}}}, \quad (4)$$

[0094] 其中, Z_{near} 和 Z_{far} 分别对应于场景中最远的像素和最远的像素的深度因子。当使用多于(或少于)8 比特来量化深度信息时,方程式(4)内的值 255 应当由 $2^B - 1$ 来代替,其中 B 是比特深度。

[0095] 当 P 的 3D 位置已知,通过方程式(2)将 P 重新投射到合成的图像平面上时,我们得到它在目标视图中的位置 p_s (即,变形的像素位置)。

[0096] (b) 矫正的视图

[0097] 对于矫正的视图, 1-D 视差 (典型的是沿着一条水平线) 描述如何将像素从一个视图替换到另一个视图。假设给定下面的摄像机参数:

[0098] (i) f , 摄像机透镜的焦距

[0099] (ii) l , 基线间距, 也称作摄像机距离; 和

[0100] (iii) du , 主点偏移的差异

[0101] 就输入视图得到了很好的矫正而言, 可以使用下面的公式从参考视图内的像素 $p_r = [u_r, v_r, 1]^T$ 计算目标视图内的变形的的位置 $p_s = [u_s, v_s, 1]^T$:

[0102]

$$u_s = u_r - \frac{f \cdot l}{z} + du; \quad v_s = v_r. \quad (5)$$

[0103] 2. 提议的方法: 视图混合

[0104] 在图 1A 和图 1B 中图示了视图变形的结果。在该步骤中处理如何从像素周围的变形的参考像素 (候选像素) 来估计目标视图中的像素 (目标像素) 值。在至少一个实施例中, 鉴于以上, 提供了一种基于深度信息、变形的像素位置和摄像机参数来合并几个变形的参考像素的启发式方法。

[0105] 实施例 1: 矫正的视图

[0106] 为了简单起见, 将矫正的视图合成作为例子, 即在相同的水平线上 (图 1B) 从候选像素来估计目标像素值。

[0107] 对于每个目标像素, 将距离该目标像素 $\pm a$ 像素内的变形的像素选择作为候选像素。找到具有最大深度级 $\max Y$ (最靠近虚拟摄像机) 的像素。此处参数 a 非常重要。如果太小, 将会出现针孔。如果太大, 将会丢失图像细节。如果关于场景或输入深度精度的一些现有知识是已知的, 例如使用深度噪声的差异, 则可以调整该参数。如果没有已知的知识, 则大多数时间应用值 1。

[0108] 在典型的 Z 缓冲算法中, 最大深度级别 (即, 最靠近摄像机) 的候选像素将确定目标位置处的像素值。这里, 也可以保留其他候选像素, 只要他们的深度级别与最大深度相当接近, 即 ($Y \geq \max Y - \text{thres} Y$)。其中 $\text{thres} Y$ 是临界参数。在实验中, $\text{thres} Y$ 设置为 10, 其可以根据 $\max Y$ 的大小或关于输入深度的精度的一些现有的知识来进行改变。我们用 m 表示所找到的候选像素的数量。

[0109] 为了进一步保持图像细节, 如果在距离目标像素 $\pm a/2$ 像素内有“足够”数量的候选像素, 则只有这些像素被用于估计目标像素颜色。我们定义这些候选像素的数量为 n 。为了确定 n 是否足够, 可以应用不同的标准, 如:

[0110] (i) 如果 $n \geq N$, 即如果 n 大于预定的临界值 N (当 $\text{thres} Y$ 设置为 10 并且有两个参考视图时, 我们推荐将其设置为 4)。在图 8 中显示了该推荐标准。

[0111] (ii) 如果 $m - n < M$, 即如果 m 不远大于 n , 其中 M 是预设的临界值。

[0112] 当然, 本发明原理不只限于以上的差值标准, 并且因此也可以使用其他差值标准, 同时保持本发明原理的精神。

[0113] 在选择了 n_p 个候选像素之后, 下一个任务是插值目标像素 C_s 。定义候选像素 i 的值为 C_i , 其是从参考视图 r_i 变形得到的, 并且到目标像素的相应的距离为 d_i 。则以下的线

性插值非常适用：

$$[0114] \quad C_s = (\sum_{i=1}^{n_p} w_i \cdot C_i) / \sum_{i=1}^{n_p} w_i, \text{ with } w_i = (a - d_i) \cdot W(r_i, i) \quad (6)$$

[0115] 其中 $W(r_i, i)$ 是分配给不同视图的加权因子。可以将其简单地设置为 1。对于矫正的视图, 建议根据基线间隔 l_r (视图 r_i 与目标视图之间的摄像机距离) 进行设置, 例如 $W(r_i, i) = 1/l_r$ 。

[0116] 图 8 显示的是根据本发明原理的实施例的用于矫正的视图的启发式视图合并过程 800。在步骤 805 中, 只选择距离目标像素 $\pm a$ 像素内的候选像素, 并且选择具有最大深度级 $\max Y$ 的候选像素 (即, 距摄像机最近)。在步骤 810, 去除深度级为 $Y < \max Y - \text{thres} Y$ 的候选像素 (即, 去除背景像素)。在步骤 815, 计算候选像素的总数 m , 以及与目标像素的距离在 $\pm a/2$ 内的候选像素的数量 n 。在步骤 820, 确定是否 $n \geq N$, 如果是, 则控制前进到步骤 825。否则, 控制前进到步骤 830。在步骤 825, 只保留与目标像素的距离在 $\pm a/2$ 内的候选像素。在步骤 830, 通过线性插值 eq (6) 估计目标像素 C_s 的颜色。

[0117] 实施例 2: 未矫正的视图

[0118] 图 8 中的合并方案容易延伸到未矫正的视图, 只是区别在于候选像素不在目标像素的相同的线上 (图 1A)。但是, 可以应用相同的原理来基于其深度和到目标像素的距离选择候选像素。

[0119] 可以使用相同的插值方案, 即图 6。为了更精确的加权, $W(r_i, i)$ 可以在像素级上进一步确定。例如, 应用由 3D 点 $Or_i - P_i - O_s$ 确定的角度, 其中 P_i 是对应于像素 i (利用方程 (3) 估计的) 的点的 3D 位置, Or_i 和 O_s 分别是参考视图 r_i 和合成视图的光学焦点中心 (从摄像机参数可知)。我们推荐设置 $W(r_i, i) = 1/\text{angle}(Or_i - P_i - O_s)$ 或对于 $q > 2$, $W(r_i, i) = \cos^q(\text{angle}(Or_i - P_i - O_s))$ 。图 9 显示根据本发明的原理的实施例由 3D 点 $Or_i - P_i - O_s$ 确定的角度 900。图 7 的方法 700 中的步骤 725 显示了根据一个实施例的基于角度 900 的加权因子的确定。

[0120] 实施例 3: 利用上采样的近似

[0121] 上述两个实施例中的方案看起来可能对于一些应用太复杂。有一些方法可以近似这些实施例用于快速实施。图 10A 显示的是根据本发明原理的实施例的用于矫正的视图的情况下简化的上采样实施 1000。在图 10A 中, “+” 表示插入半像素位置的新的目标像素。图 10B 显示的是根据本发明原理的实施例的基于 Z 缓冲的合并方案 1050。在步骤 1055 中, 在每个水平线 (例如上采样图 10A) 的半像素位置生成新样本。在步骤 1060 中, 从距目标像素 $\pm 1/2$ 内的候选像素中, 寻找具有最大深度级的像素并将其颜色作为目标像素 C_s (即 Z 缓存) 的颜色。在步骤 1065 中, 利用滤波器进行下采样 (例如 $\{1, 2, 1\}$)。

[0122] 在合成的视图中, 首先在所有半像素位置插入新的目标像素 (图 10A), 即沿水平方向上采样。然后对于每个目标像素, 应用简单的 Z 缓冲方案以估计像素值。该方案相当于在通常情况下 (图 8) 设置 $\text{thres} Y = 0$ 。为了产生最终的合成视图, 应用简单的下采样滤波器 (例如 $\{1, 2, 1\}$)。该滤波器近似在方程 (6) 中的加权 w_i 。

[0123] 相同的方法也可以应用于未矫正的视图。只是区别在于沿着水平和垂直方向上采样图像。

[0124] 可以理解虽然关于半像素和半像素位置描述了一个或多个实施例, 但是本发明原

理可应用于任何大小的子像素（并因此对应于子像素位置），同时保持本发明原理的精神。

[0125] 实施例 4：两步混合

[0126] 虽然在 3DV 中通常应用两个参考视图，但是以上描述的混合方案没有限制多少参考视图被提供作为输入。为了使所述方案易于实施，上述方案也可转换为两步，即单独地利用每个参考视图合成虚拟视图（例如应用上述任何方案），然后将所有合成的图像合并在一起。对于实施例 3 的一个实施例，该实施例利用上采样图像进行合并然后对合并的图像进行下采样。

[0127] 对于合并部分，可以应用简单的 Z 缓冲方案（即利用来自不同视图的候选像素，选择最靠近摄像机的一个）。可以选择，也可以应用上述关于 $W(r_i, i)$ 的加权方案。当然，在合并期间可以应用任何其他现有的视图加权方案。

[0128] 3. 后期处理：填充空洞

[0129] 目标视图内的一些像素在所述混合步骤期间从未被分配有值。这些位置称为空洞，通常是由未遮挡造成的（参考视图内先前几个看不见的场景点在合成视图内没有被覆盖到）。处理这些空洞最简单的方法是检查沿着这些空洞边缘的像素，并使用这些边缘像素的一些来填充这些空洞。由于该步骤不涉及所述混合方案，可以使用任何现有的空洞填充方案。

[0130] 因此，总之在一个或多个实施方式中，我们提议启发式混合方案：(1) 基于深度级别及其变形的图像位置来选择候选像素；和 (2) 利用由变形的图像位置和摄像机参数确定的加权因子来应用线性插值。

[0131] 由于我们的方法是启发式的，可能有一些潜在的变化。例如在实施例 1 和 2 中，如果有“足够”数量的候选像素则只选择在距离目标像素 $\pm a/2$ 像素内的候选像素。 $1/2$ 是为了便于实施。实际上，其可能是 $1/k$ ， k 可以为任何值。另一方面，可以增加一个或多个选择级别，例如只找到在距离目标像素 $\pm a/3$ ， $\pm a/4$ 或 $\pm a/6$ 像素内的候选像素等等。可以选择，跳过一步一步的选择过程，从距离目标像素最近的候选像素开始直到有足够的候选像素。另一个更常用的选择是基于到目标像素的距离来分组候选像素，并使用最靠近的组作为候选。

[0132] 作为实施例 3 中的另一实施例，在最后的下采样期间，目标视图被上采样到半像素位置以近似线性插值。以增加更多复杂性为代价，可以采用更多级别的上采样来获得更好的精度。此外，沿水平和垂直方向的上采样级别可以是不同的。

[0133] 我们描述了至少一个实施例，其中从至少一个参考视图位置到虚拟视图位置变形至少一个参考图像或其一部分，以产生至少一个变形的参考。这种实施识别至少一个变形的参考中的第一候选像素和第二候选像素，该第一候选像素和第二候选像素是来自所述虚拟视图位置的虚拟图像中目标像素位置的候选像素。该实施例还基于第一和第二候选像素的值来确定目标像素位置处的像素值。该实施例可以有很多变化。例如，在第一变化中，单个参考图像被变形以产生单个变形的参考，从中获得两个候选像素并用于确定在目标像素位置的像素值。作为另一个例子，在第二变化中，多个参考图像被变形以产生多个变形的参考，并且从每个变形的参考获得单个候选像素并用于确定在目标像素位置的像素值。

[0134] 我们已经描述了各种实施方式。鉴于以上，前述仅阐释的是本发明原理，因此应当理解本领域技术人员仍能够改动多个替换的解决方法，虽然在此处没有明确地描述，但是

这些替换的解决方法体现本发明的原理,并在本发明的精神和范畴内。因此,我们提供具有特定特征和技术方面的一个或多个实施方式。但是,所描述的実施方式的特征和技术方面也可以被改动以用于其它实施方式。相应地,虽然此处所描述的实施例可能是在特定的场景内描述的,但是这样的描述决不当被視為是限制所述特征和构想为这样的实施方式或者场景。

[0135] 说明书中提到的本发明原理的“一个实施例”或者“实施例”或者“一种实施方式”或者“实施方式”及它们的其它变化,意思是结合所述实施例所描述的特定特征、结构、特征等等之类包含在本发明原理的至少一个实施例中。因此,出现短语“在一个实施例中”、“在实施例中”或者“在一种实施方式中”,或者“在实施方式中”,以及出现在整个说明书中各个地方的其它变化,并不是所有都必须指的是同一个实施例。

[0136] 要理解的是,使用“/”“和 / 或”和“至少一个”,例如,在“A/B”、“A 和 / 或 B”和“A 和 B 中的至少一个”旨在涵盖只选择第一罗列的选项 (A),或者只选择第二罗列的选项 (B),或者两个选项都有 (A 和 B)。作为进一步的例子,在“A、B 和 / 或 C”和“A、B 和 C 中至少一个”的情况下,这样的表述旨在涵盖只选择第一罗列的选项 (A),或者仅选择第二罗列的选项 (B),或者仅选择第三罗列的选项 (C),或者仅选择第一和第二罗列的选项 (A 和 B),或者仅选择第一和第三罗列的选项 (A 和 C),或者仅选择第二和第三罗列的选项 (B 和 C),或者选择所有的选项 (A 和 B 和 C)。对于本领域普通技术人员而言,这显而易见是可以被扩展的。

[0137] 实施方式可以使用各种技术以信号的形式发出信息,包括但不限于,带内信息、带外信息、数据流数据、隐式信令和显式信令。对于各种实施方式和 / 或标准,带内信息和显式信令可以包括,片头、SEI 消息、其它高级语法和非高级的语法。相应地,虽然可能此处描述的实施方式是在特定的场景中描述的,但是,这样的描述绝不当被視為是限制所述特征和构思在这样的实施方式或者场景中。

[0138] 此处所描述的實施方式和特征可以用于 MPEG-4AVC 标准、或者有 MVC 扩展的 MPEG-4AVC 标准,或者有 SVC 扩展的 MPEG-4AVC 标准的环境中。但是,这些实施方式和特征可以用于另一种标准和 / 或推荐 (现有的或将来的) 的环境中,或者用于不涉及标准和 / 或推荐的环境中。

[0139] 此处所描述的实施例可以被实施在,例如,方法或者过程、设备、软件程序、数据流或者信号中。即使所讨论的特征的实施方式仅在单个形式的實施方式的环境下 (例如,仅作为方法来讨论),所讨论的特征的实施方式可以以其它形式来实施 (例如,设备或者程序)。例如,可以以适当的硬件、软件和固件来实施设备。可以以例如处理器之类的装置来实现所述方法,处理器指的是通用的处理装置,例如,包括计算机、微处理器、集成电路或者可编程逻辑装置。处理器还包括诸如,计算机、手机、便携式 / 个人数据助理 (“PDA”) 和其它便于在终端用户之间交流信息的装置之类的通信装置。

[0140] 此处描述的各种步骤和特征的實施方式可以体现在各种不同的装备或者应用中,特别是,例如,与数据编码和解码相关联的装备或者应用。这样的装备示例包括编码器、解码器、处理来自于解码器的后期处理器、给编码器提供输入的预处理器、视频编码器、视频解码器、视频编解码器、网络服务器、机顶盒、笔记本电脑、个人计算机、手机、PDA 和其它通信装置。应当清楚的是,所述装备可以是移动的,甚至是安装在移动交通工具内。

[0141] 另外,所述方法可以通过由处理器所执行的指令来实施,这样的指令(和/或由一种实施方式产生的数据值)可以存在处理器可读的介质中,诸如,例如集成电路、软件载体或者其它诸如硬盘、压缩光盘、随机存取存储器(“RAM”)或者只读存储器(“ROM”)之类的存储设备。所述指令可以形成有形地体现在处理器可读介质上的应用程序。指令可以是,例如,硬件、固件、软件或者组合。指令可以在例如,操作系统内、单独的应用中或者二者的组合内。因此,处理器的特征在于,例如,配置为执行处理的装置和包括具有执行处理的指令的处理器可读介质(诸如存储装置)。进一步地,除了或者代替指令,处理器可读介质可以存储一个实施方式所产生的数据值。

[0142] 实施方式可以产生被格式化来承载例如可以被存储或者被发送的信息的多种信号,这对本领域技术人员而言是显而易见的。所述信息可以包括,例如用于执行一个方法的指令,或者由所描述的一个实施方式生成的数据。例如,信号可以被格式化为承载混合的或者合并的变形的-参考-视图的数据,或者用于混合或者合并变形的参考视图的算法。这样的信号可以被格式化为,例如,作为电磁波(例如,使用波谱的无线射频部分)或者作为基带信号。所述格式化可以包括,例如,对数据流进行编码,并用编码的数据流调制载波。所述信号承载的信息可以是,例如,模拟或者数字信息。所述信号可以在各种不同的有线或者无线链路上传输,这是周知的。所述信号可以存储在处理器可读介质上。

[0143] 已经描述了一些实施方式。但是,要理解的是可以进行各种修改。例如,可以组合、补充、修改或者移去不同实施方式中的要素,来产生其它实施方式。另外,普通技术人员会明白,可以使用其它结构和处理过程来代替所揭示的这些,并且,如所揭示的实施方式,所获得的实施方式可以以至少基本上相同的方式,执行至少基本上相同的功能,并达到至少基本上相同的结果。相应地,本申请也包含这些及其它实施方式,它们也在权利要求的范围内。

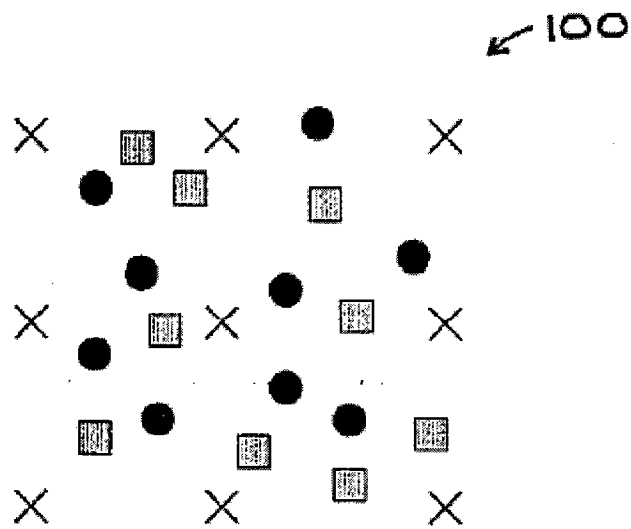


图 1A

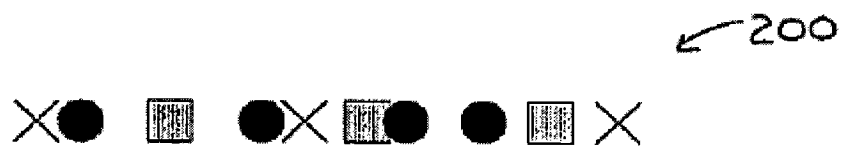


图 1B

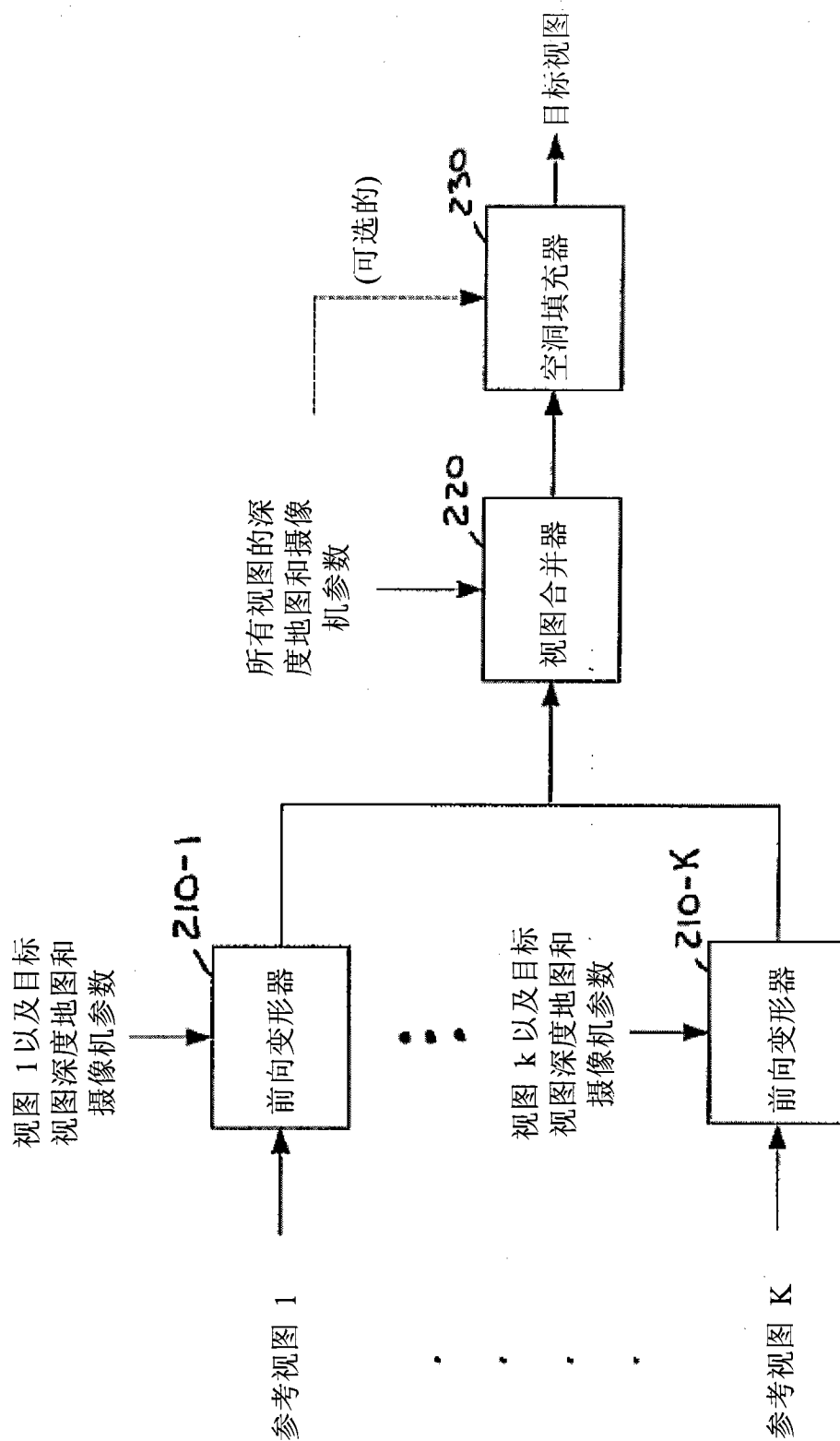


图 2

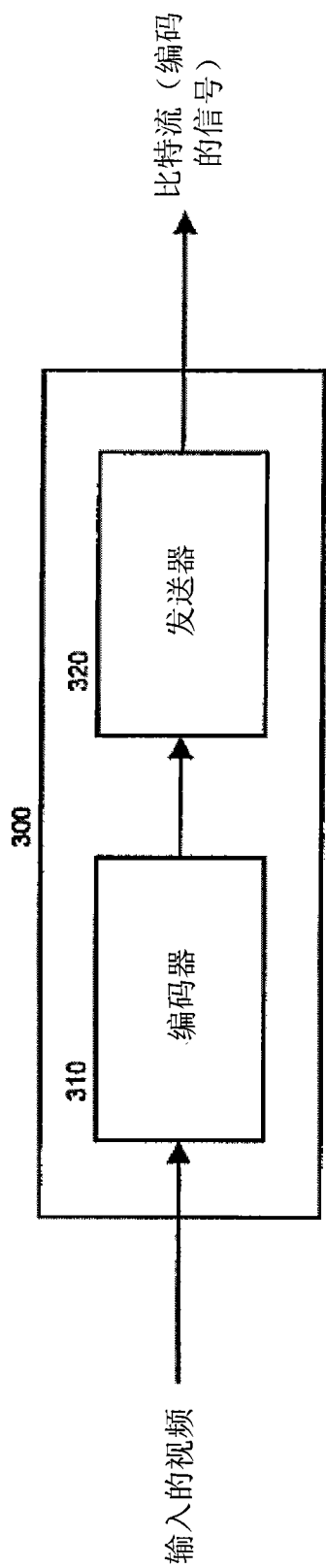


图 3

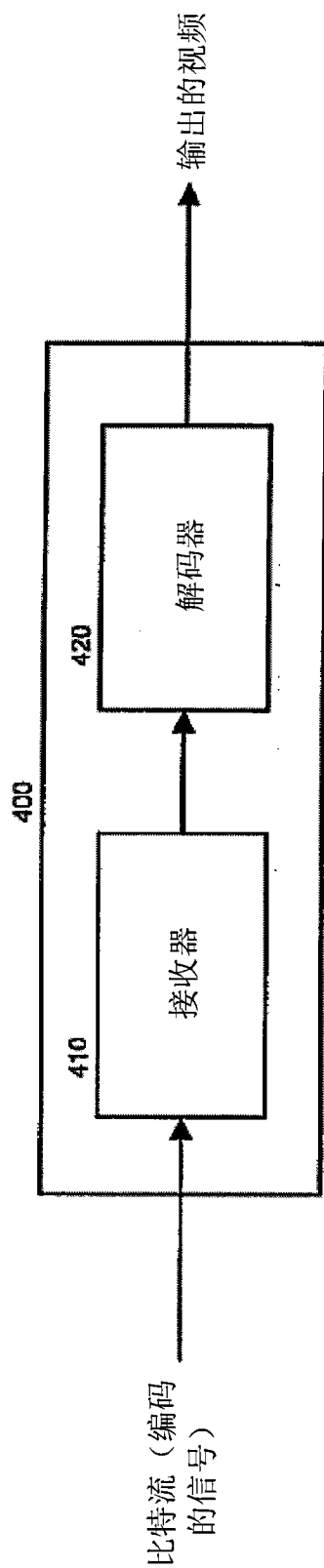


图 4

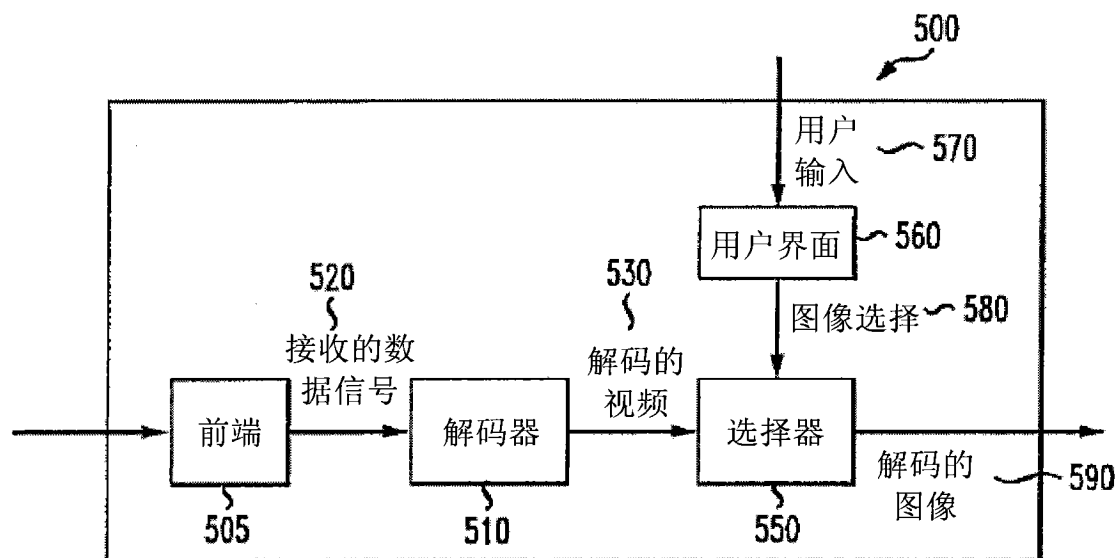


图 5

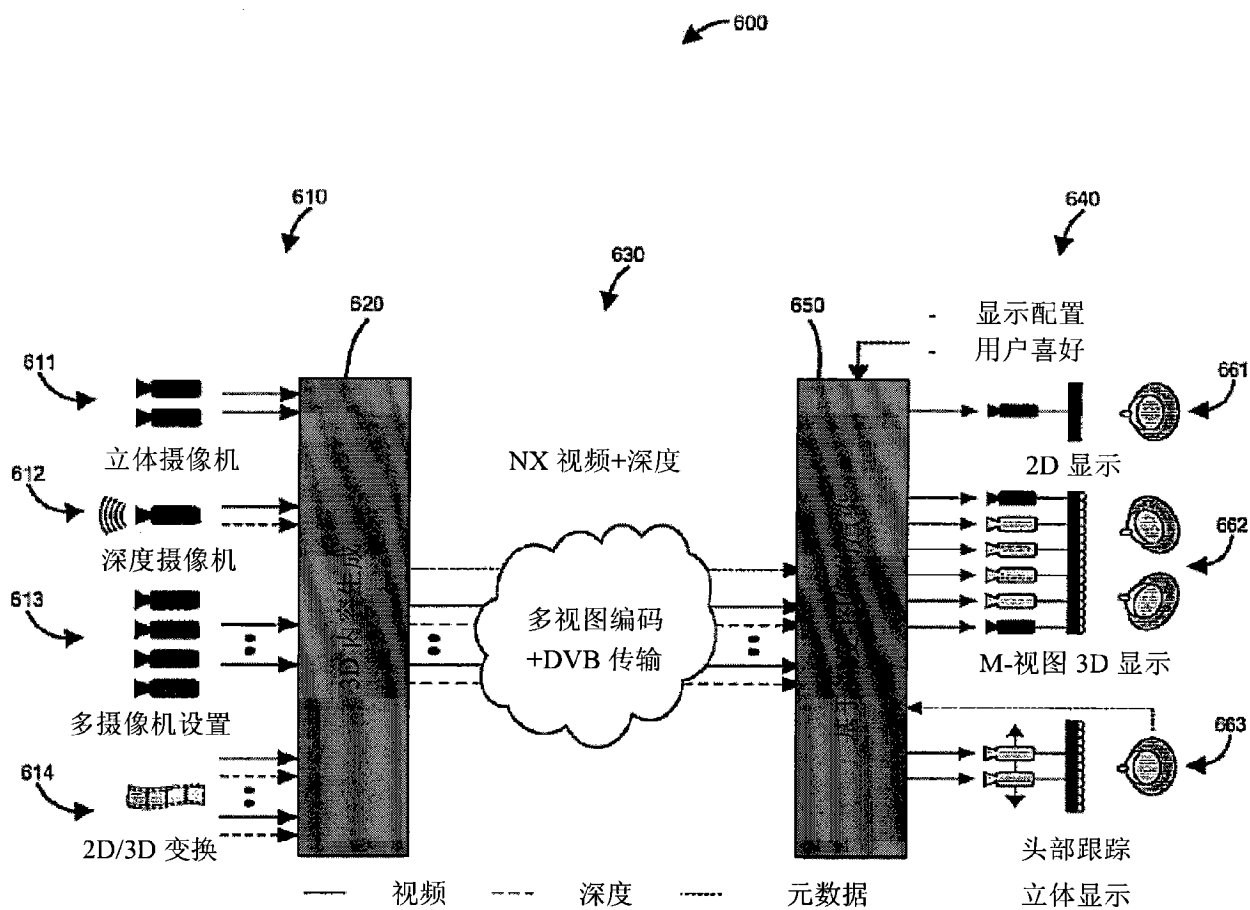


图 6

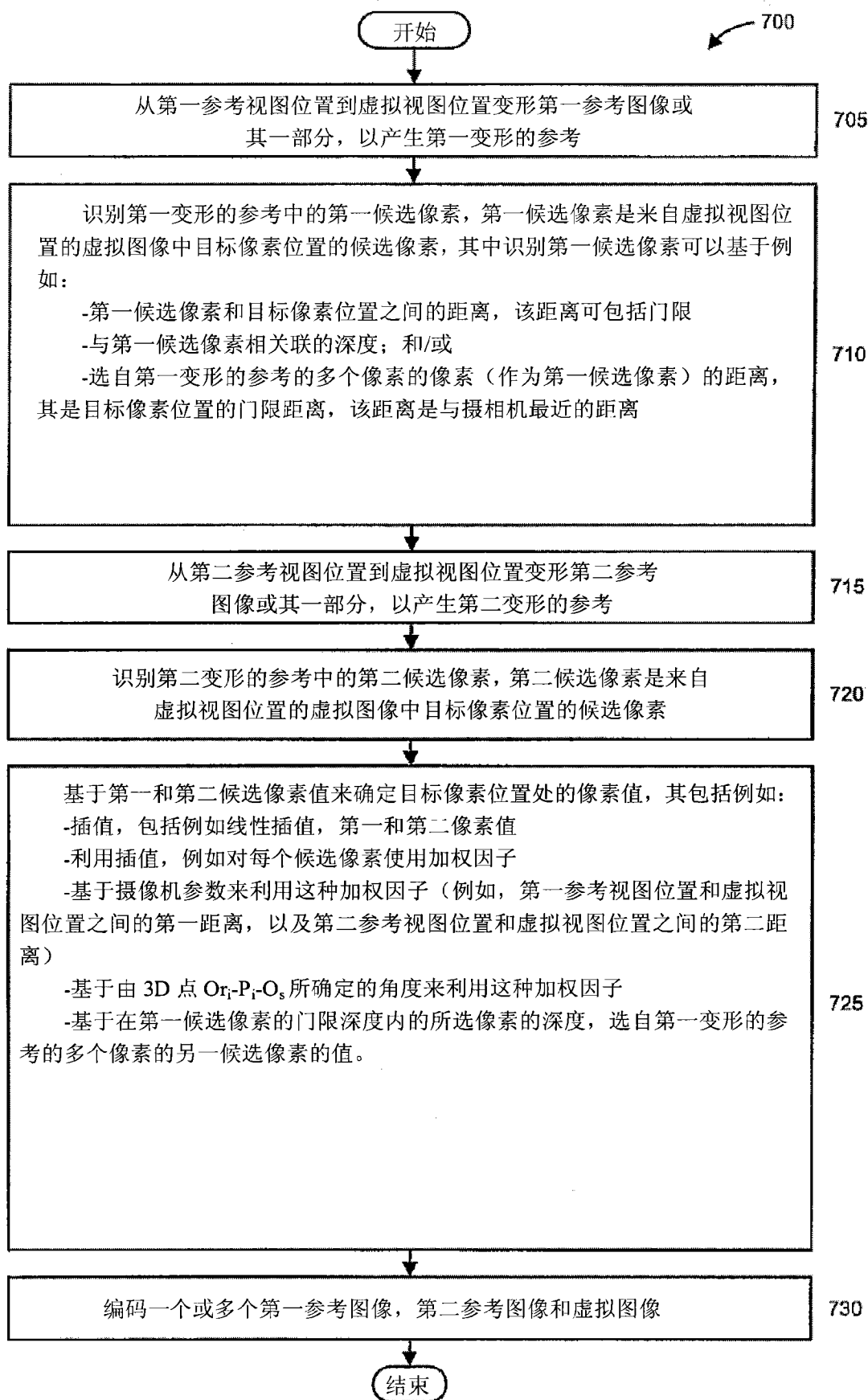


图 7

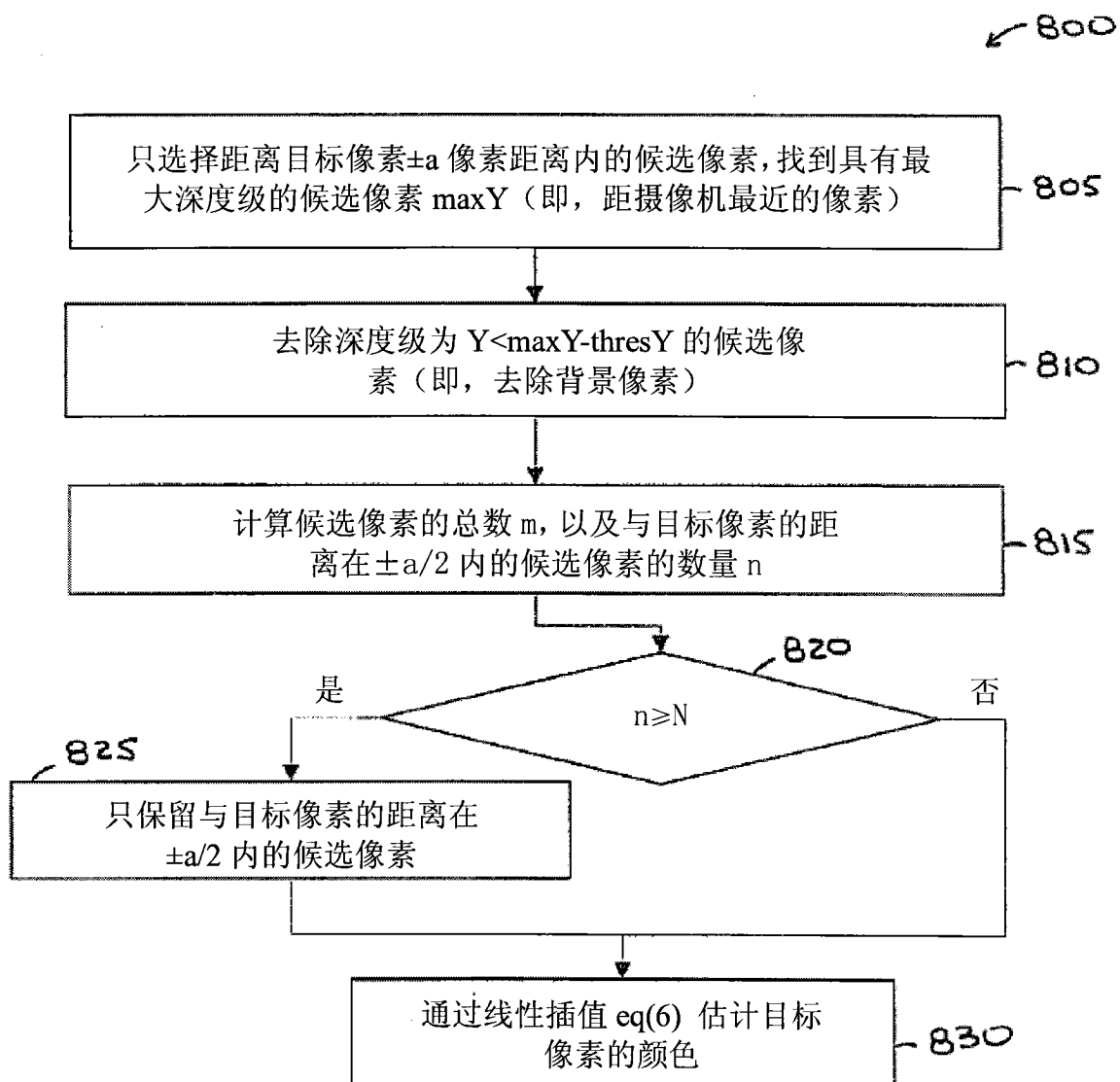


图 8

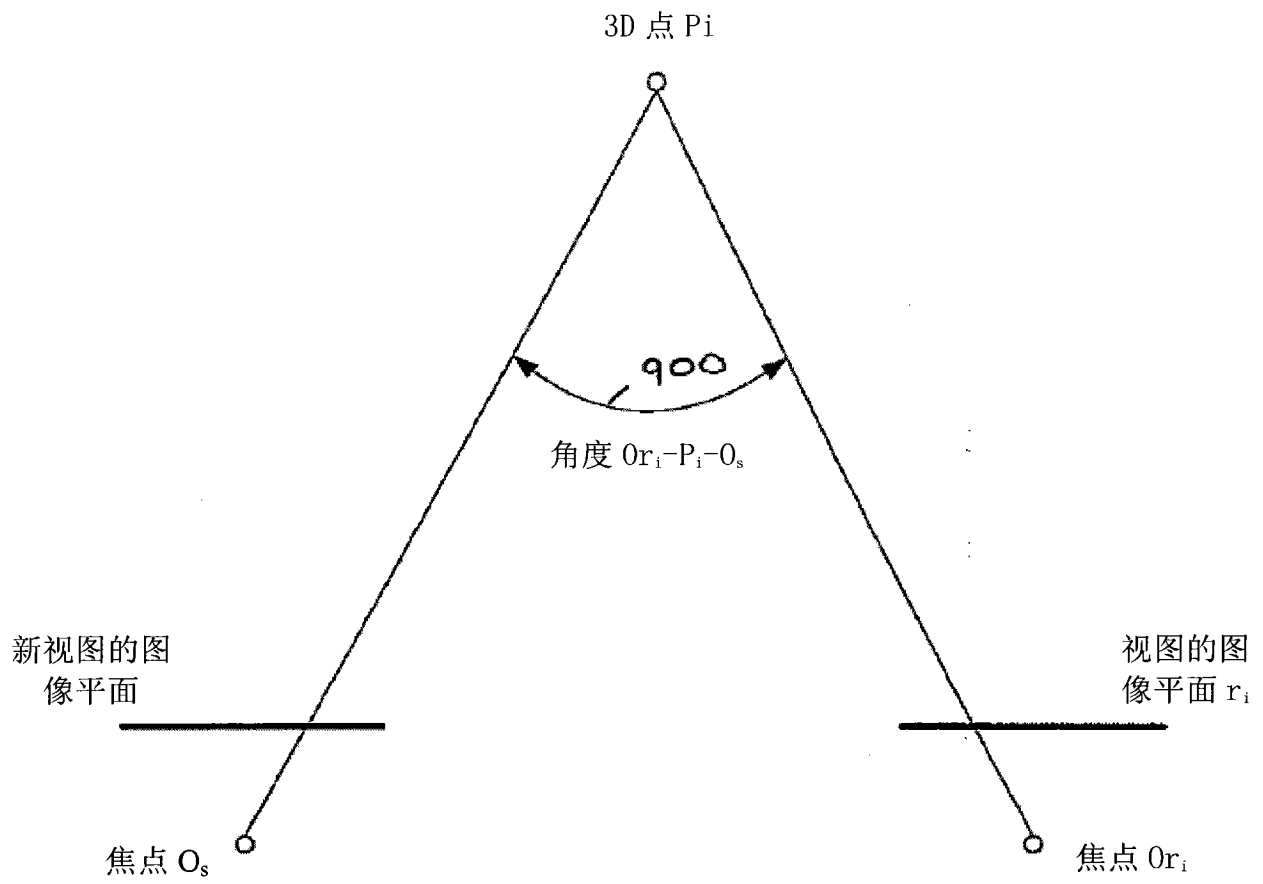


图 9

