



(12) 发明专利

(10) 授权公告号 CN 103004214 B

(45) 授权公告日 2015. 11. 25

(21) 申请号 201180034708. 3

(51) Int. Cl.

(22) 申请日 2011. 07. 15

H04N 13/00(2006. 01)

(30) 优先权数据

61/364, 940 2010. 07. 16 US

13/094, 607 2011. 04. 26 US

(56) 对比文件

CN 1981295 A, 2007. 06. 13,

WO 2010/025458 A1, 2010. 03. 04,

Chun-Chi Chen et al. A

(85) PCT国际申请进入国家阶段日

2013. 01. 14

Synthesis-Quality-Oriented Depth Refinement Scheme for MPEG Free Viewpoint Television (FTV). 《IEEE Xplore Digital Library》. 2009, 第 171-178 页.

(86) PCT国际申请的申请数据

PCT/US2011/044277 2011. 07. 15

David J. Scherba et al. Depth Map

Calibration by Stereo and Wireless Sensor Network Fusion. 《IEEE Xplore Digital Library》. 2005, 第 1540-1547 页.

(87) PCT国际申请的公布数据

W02012/009695 EN 2012. 01. 19

(73) 专利权人 高通股份有限公司

地址 美国加利福尼亚州

审查员 张永海

(72) 发明人 胡迪特·马蒂内·包萨

马什胡尔·M·苏勒赫

(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

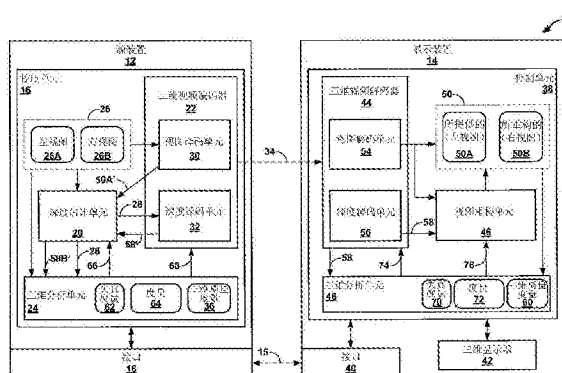
权利要求书7页 说明书17页 附图7页

(54) 发明名称

用于三维视频的基于视觉的质量度量

(57) 摘要

一般来说,本发明描述用于针对三维(3D)视频确定基于视觉的质量度量的技术。包括3D分析单元(24、48)的装置(12、14)可实施这些技术来计算3D客观度量(36、60)。所述3D分析单元(24、48)估计将使用基于DIBR的3D视频数据(34)产生失真受限的图像视图的理想深度图(62、70),且随后基于所述理想深度图(62、70)与在产生所述基于DIBR的3D视频数据(34)中使用的深度图(28、58)的定量比较来导出一个或一个以上失真度量(64、72)。基于所述导出的一个或一个以上失真度量(64、72),所述3D分析单元(24、48)计算所述客观度量(36、60)以量化基于DIBR的视频数据(34)的视觉质量。所述3D客观度量(36、60)可用于更改或以其他方式调整深度估计和编码/解码参数以校正预期的视觉不适。



CN 103004214 B

1. 一种用于获得客观度量 (36、60) 以量化基于基于深度图像再现 DIBR 的三维 3D 视频数据的视觉质量的方法, 所述方法包括:

至少部分基于以下参数来估计将使用基于 DIBR 的 3D 视频数据 (138、156) 产生失真受限图像视图的理想深度图 (62、70): 从所述基于 DIBR 的 3D 视频数据产生的视图、虚拟 3D 图像俘获装置的焦距和所述虚拟 3D 图像俘获装置的基线;

基于所述理想深度图 (62、70) 与在产生所述基于 DIBR 的 3D 视频数据 (34、140、158) 中使用的深度图 (28、58) 的定量比较来导出一个或一个以上失真度量 (64、72); 以及

基于所述导出的一个或一个以上失真度量 (64、142、160) 来计算所述客观度量 (36、60) 以量化所述基于 DIBR 的 3D 视频数据的视觉质量,

其中估计所述理想深度图包含根据以下等式中的一者来估计所述理想深度图:

$$Z_{IDEAL} = \frac{\pm FB}{k(I_o - I_g) \pm \frac{FB}{Z}},$$

其中  $I_g$  表示从所述基于 DIBR 的 3D 视频数据产生的视图,  $I_o$  表示原始图像,  $Z$  表示所述基于 DIBR 的 3D 视频数据的所述深度图,  $F$  表示用于从所述基于 DIBR 的 3D 视频数据产生不同视图的虚拟 3D 图像俘获装置的焦距, 而  $B$  表示用于从所述基于 DIBR 的 3D 视频数据产生不同视图的虚拟 3D 图像俘获装置的基线;

$$Z_{IDEAL} = \frac{\pm FB}{k(I_R - I_g) \pm FB \left( \frac{1}{Z_L} + \frac{1}{Z_R} \right) + h},$$

其中  $I_R$  表示参考视图,  $I_g$  表示从所述基于 DIBR 的 3D 视频数据产生的视图,  $Z_L$  表示来自原始左眼视图的所述基于 DIBR 的 3D 视频数据的深度图,  $Z_R$  表示来自原始右眼视图的所述基于 DIBR 的 3D 视频数据的深度图,  $k$  表示第一常数,  $h$  表示第二常数,  $F$  表示用于从所述基于 DIBR 的 3D 视频数据产生额外视图的所述虚拟 3D 图像俘获装置的焦距, 而  $B$  表示用于从所述基于 DIBR 的 3D 视频数据产生额外视图的所述虚拟 3D 图像俘获装置的基线; 或者

$$Z_{IDEAL} \approx \frac{\pm FB}{f(I_R, I_g) \pm \frac{FB}{Z}},$$

其中  $I_R$  表示参考视图,  $I_g$  表示从所述基于 DIBR 的 3D 视频数据产生的视图,  $F$  表示用于从所述基于 DIBR 的 3D 视频数据产生额外视图的所述虚拟 3D 图像俘获装置的焦距, 而  $B$  表示用于从所述基于 DIBR 的 3D 视频数据产生所述额外视图的所述虚拟 3D 图像俘获装置的基线,

其中函数  $f(I_R, I_g)$  表示在将水平移位应用于所述参考视图  $I_R$  的每一像素块之后所述参考视图中的所述每一像素块与所述产生的视图  $I_g$  中的每一对应像素块之间的亮度上的差,

其中所述水平移位的量是  $I_R$  和  $I_g$  的对应块之间的平均视差, 以及

其中所述平均视差随所述参考视图的深度图而变。

2. 根据权利要求 1 所述的方法, 其中所述失真度量中的每一个指示在观看所述基于

DIBR 的 3D 视频数据时预期的 3D 观看视觉不适的测量。

3. 根据权利要求 1 所述的方法, 其中导出所述一个或一个以上失真度量包括导出空间错误异常点 S0 失真度量和时间错误异常点 T0 失真度量中的一个或一个以上。

4. 根据权利要求 1 所述的方法,

其中导出所述一个或一个以上失真度量包括:

将空间错误异常点 S0 失真度量导出为理想深度图与所述基于 DIBR 的 3D 视频数据的所述深度图之间的差的标准偏差; 以及

将时间错误异常点 T0 失真度量导出为针对所述 3D 视频的两个连续帧的所述理想深度图与所述基于 DIBR 的 3D 视频数据的所述深度图之间的差的标准偏差,

其中所述方法进一步包括将时间不一致性 TI 失真度量导出为用于产生基于 DIBR 的 3D 视频数据和相继的基于 DIBR 的 3D 视频数据的深度图之间的所述差的标准偏差,

其中计算所述客观度量包括:

将所述 S0、T0 和 TI 失真度量中的每一个规格化到 0 到 1 的范围; 以及

根据以下等式计算所述客观度量:

$$3VQM = K(1-S0(S0 \cap T0))^a(1-TI)^b(1-T0)^c,$$

其中 3VQM 表示所述客观度量, S0、T0 和 SI 表示所述经规格化的 S0、T0 和 TI 失真度量, a、b 和 c 表示通过训练序列确定的常数, 而 k 表示缩放因子。

5. 根据权利要求 1 所述的方法,

其中所述基于 DIBR 的 3D 视频数据进一步包含提供的视图, 以及

其中所述方法进一步包括:

根据一个或一个以上参数从 3D 视频数据估计所述深度图, 其中所述 3D 视频数据包含左视图和右视图, 其中所述提供的视图选自所述左视图和右视图之一; 以及

修正用于估计所述 3D 视频数据的所述深度图的所述一个或一个以上参数以便校正由所述客观度量识别的缺陷。

6. 根据权利要求 1 所述的方法,

其中所述基于 DIBR 的 3D 视频数据进一步包含提供的视图, 以及

其中所述方法进一步包括:

根据一个或一个以上参数对所述基于 DIBR 的 3D 视频数据的所述深度图和所述提供的视图进行编码以产生经编码的基于 DIBR 的 3D 视频数据; 以及

修正用于对所述深度图和所述提供的视图中的一个或一个以上进行编码的所述一个或一个以上参数以校正由所述客观度量识别的缺陷。

7. 根据权利要求 1 所述的方法,

其中所述 3D 视频数据进一步包含提供的视图, 以及

其中所述方法进一步包括:

根据一个或一个以上参数对所述深度图的经编码版本和所述提供的视图的经编码版本进行解码以产生所述深度图和所述提供的视图; 以及

修正用于对所述深度图的所述经编码版本和所述提供的视图的所述经编码版本进行解码的所述一个或一个以上参数以便校正由所述客观度量识别的缺陷。

8. 根据权利要求 1 所述的方法,

其中所述基于 DIBR 的 3D 视频数据进一步包含提供的视图,以及

其中所述方法进一步包括:

根据一个或一个以上参数从所述深度图和所述提供的视图重构 3D 视频数据的左视图或右视图以产生所述 3D 视频数据,其中将所述提供的视图选择为所述 3D 视频数据的所述左视图或所述右视图;以及

修正用于重构所述左视图或所述右视图的所述一个或一个以上参数以校正由所述客观度量识别的缺陷。

9. 一种获得客观度量 (36、60) 以量化基于基于深度图像再现 DIBR 的三维 3D 视频数据的视觉质量的装置,所述装置 (12、14) 包括:

3D 分析单元 (24、48),其计算所述客观度量 (36、60),其中所述 3D 分析单元 (24、48) 包含:

理想深度估计单元 (82),其至少部分基于以下参数来估计将使用基于 DIBR 的 3D 视频数据 (34) 产生失真受限图像视图的理想深度图 (62、70):从所述基于 DIBR 的 3D 视频数据产生的视图、虚拟 3D 图像俘获装置的焦距和所述虚拟 3D 图像俘获装置的基线;

失真度量计算单元 (84),其基于所述理想深度 (62、70) 与在产生所述基于 DIBR 的 3D 视频数据 (34) 中使用的深度图 (28、58) 的定量比较来导出一个或一个以上失真度量 (64、72);以及

客观度量计算单元 (86),其基于所述导出的一个或一个以上失真度量 (64、72) 来计算所述客观度量 (36、60) 以量化基于 DIBR 的视频数据 (34) 的视觉质量,

其中所述理想深度图估计单元根据以下等式中的一者来估计所述理想深度图:

$$Z_{IDEAL} = \frac{\pm FB}{k(I_o - I_g) \pm \frac{FB}{Z}},$$

其中  $I_g$  表示从所述基于 DIBR 的 3D 视频数据产生的视图, $I_o$  表示原始图像, $Z$  表示所述基于 DIBR 的 3D 视频数据的所述深度图, $F$  表示用于从所述基于 DIBR 的 3D 视频数据产生不同视图的虚拟 3D 图像俘获装置的焦距,而  $B$  表示用于从所述基于 DIBR 的 3D 视频数据产生不同视图的虚拟 3D 图像俘获装置的基线;

$$Z_{IDEAL} = \frac{\pm FB}{k(I_R - I_g) \pm FB \left( \frac{1}{Z_L} + \frac{1}{Z_R} \right) + h},$$

其中  $I_R$  表示参考视图, $I_g$  表示从所述基于 DIBR 的 3D 视频数据产生的视图, $Z_L$  表示来自原始左眼视图的所述基于 DIBR 的 3D 视频数据的深度图, $Z_R$  表示来自原始右眼视图的所述基于 DIBR 的 3D 视频数据的深度图, $k$  表示第一常数, $h$  表示第二常数, $F$  表示用于从所述基于 DIBR 的 3D 视频数据产生额外视图的所述虚拟 3D 图像俘获装置的焦距,而  $B$  表示用于从所述基于 DIBR 的 3D 视频数据产生额外视图的所述虚拟 3D 图像俘获装置的基线;或者

$$Z_{IDEAL} \approx \frac{\pm FB}{f(I_R, I_g) \pm \frac{FB}{Z}},$$

其中  $I_R$  表示参考视图, $I_g$  表示从所述基于 DIBR 的 3D 视频数据产生的视图, $F$  表示用

于从所述基于 DIBR 的 3D 视频数据产生额外视图的所述虚拟 3D 图像俘获装置的焦距, 而 B 表示用于从所述基于 DIBR 的 3D 视频数据产生所述额外视图的所述虚拟 3D 图像俘获装置的基线,

其中函数  $f(I_r, I_g)$  表示在将水平移位应用于所述参考视图  $I_r$  的每一像素块之后所述参考视图中的所述每一像素块与所述产生的视图  $I_g$  中的每一对应像素块之间的亮度上的差,

其中所述水平移位的量是  $I_r$  和  $I_g$  的对应块之间的平均视差, 以及

其中所述平均视差随所述参考视图的深度图而变。

10. 根据权利要求 9 所述的装置, 其中所述失真度量中的每一个指示在观看所述基于 DIBR 的 3D 视频数据时预期的 3D 观看视觉不适的测量。

11. 根据权利要求 9 所述的装置,

其中所述失真度量计算单元将空间错误异常点 S0 失真度量导出为理想深度图与所述基于 DIBR 的 3D 视频数据的所述深度图之间的差的标准偏差, 将时间错误异常点 T0 失真度量导出为针对所述 3D 视频的两个连续帧的所述理想深度图与所述基于 DIBR 的 3D 视频数据的所述深度图之间的差的标准偏差, 且将时间不一致性 TI 失真度量导出为用于产生基于 DIBR 的 3D 视频数据和相继的基于 DIBR 的 3D 视频数据的深度图之间的所述差的标准偏差,

其中客观度量计算单元将所述 S0、T0 和 TI 失真度量中的每一个规格化到 0 到 1 的范围且根据以下等式来计算所述客观度量:

$$3VQM = K(1-S0(S0 \cap T0))^a(1-TI)^b(1-T0)^c,$$

其中 3VQM 表示所述客观度量, S0、T0 和 SI 表示所述经规格化的 S0、T0 和 TI 失真度量, a、b 和 c 表示通过训练序列确定的常数, 而 k 表示缩放因子。

12. 根据权利要求 9 所述的装置,

其中所述基于 DIBR 的 3D 视频数据进一步包含提供的视图, 以及

其中所述装置进一步包含深度估计单元, 所述深度估计单元根据一个或一个以上参数从 3D 视频数据估计所述深度图, 其中所述 3D 视频数据包含左视图和右视图, 其中所述提供的视图选自所述左视图和右视图之一, 且修正用于估计所述 3D 视频数据的所述深度图的所述一个或一个以上参数以便校正由所述客观度量识别的缺陷。

13. 根据权利要求 9 所述的装置,

其中所述基于 DIBR 的 3D 视频数据进一步包含提供的视图, 以及

其中所述装置进一步包含 3D 视频编码器, 所述 3D 视频编码器根据一个或一个以上参数对所述基于 DIBR 的 3D 视频数据的所述深度图和所述提供的视图进行编码以产生经编码的基于 DIBR 的 3D 视频数据, 且修正用于对所述深度图和所述提供的视图中的一个或一个以上进行编码的所述一个或一个以上参数以校正由所述客观度量识别的缺陷。

14. 根据权利要求 9 所述的装置,

其中所述 3D 视频数据进一步包含提供的视图, 以及

其中所述装置进一步包括 3D 视频解码器, 所述 3D 视频解码器根据一个或一个以上参数对所述深度图的经编码版本和所述提供的视图的经编码版本进行解码以产生所述深度图和所述提供的视图, 且修正用于对所述深度图的所述经编码版本和所述提供的视图的所

述经编码版本进行解码的所述一个或一个以上参数以便校正由所述客观度量识别的缺陷。

15. 根据权利要求 9 所述的装置，

其中所述基于 DIBR 的 3D 视频数据进一步包含提供的视图，以及

其中所述装置进一步包括视图重构单元，所述视图重构单元根据一个或一个以上参数从所述深度图和所述提供的视图重构 3D 视频数据的左视图或右视图以产生所述 3D 视频数据，其中将所述提供的视图选择为所述 3D 视频数据的所述左视图或所述右视图，且修正用于重构所述左视图或所述右视图的所述一个或一个以上参数以校正由所述客观度量识别的缺陷。

16. 一种获得客观度量 (36、60) 以量化基于基于深度图像再现 DIBR 的三维 3D 视频数据的视觉质量的设备，所述设备 (12、14) 包括：

用于至少部分基于以下参数来估计将使用基于 DIBR 的 3D 视频数据 (34) 产生失真受限图像视图的理想深度图 (62、70) 的装置：从所述基于 DIBR 的 3D 视频数据产生的视图、虚拟 3D 图像俘获装置的焦距和所述虚拟 3D 图像俘获装置的基线；

用于基于所述理想深度图与在产生所述基于 DIBR 的 3D 视频数据 (34) 中使用的深度图 (62、70) 的定量比较来导出一个或一个以上失真度量 (64、72) 的装置；以及

用于基于所述导出的一个或一个以上失真度量 (64、72) 来计算所述客观度量 (36、60) 以量化所述基于 DIBR 的 3D 视频数据 (34) 的视觉质量的装置，

其中用于估计理想深度图的装置进一步包括用于根据以下等式中的一者来估计所述理想深度图的装置：

$$Z_{IDEAL} = \frac{\pm FB}{k(I_o - I_g) \pm \frac{FB}{Z}},$$

其中  $I_g$  表示从所述基于 DIBR 的 3D 视频数据产生的视图， $I_o$  表示原始图像， $Z$  表示所述基于 DIBR 的 3D 视频数据的所述深度图， $F$  表示用于从所述基于 DIBR 的 3D 视频数据产生不同视图的虚拟 3D 图像俘获装置的焦距，而  $B$  表示用于从所述基于 DIBR 的 3D 视频数据产生不同视图的虚拟 3D 图像俘获装置的基线；

$$Z_{IDEAL} = \frac{\pm FB}{k(I_R - I_g) \pm FB \left( \frac{1}{Z_L} + \frac{1}{Z_R} \right) + h},$$

其中  $I_R$  表示参考视图， $I_g$  表示从所述基于 DIBR 的 3D 视频数据产生的视图， $Z_L$  表示来自原始左眼视图的所述基于 DIBR 的 3D 视频数据的深度图， $Z_R$  表示来自原始右眼视图的所述基于 DIBR 的 3D 视频数据的深度图， $k$  表示第一常数， $h$  表示第二常数， $F$  表示用于从所述基于 DIBR 的 3D 视频数据产生额外视图的所述虚拟 3D 图像俘获装置的焦距，而  $B$  表示用于从所述基于 DIBR 的 3D 视频数据产生额外视图的所述虚拟 3D 图像俘获装置的基线；或者

$$Z_{IDEAL} \approx \frac{\pm FB}{f(I_R, I_g) \pm \frac{FB}{Z}},$$

其中  $I_R$  表示参考视图， $I_g$  表示从所述基于 DIBR 的 3D 视频数据产生的视图， $F$  表示用于从所述基于 DIBR 的 3D 视频数据产生额外视图的所述虚拟 3D 图像俘获装置的焦距，而  $B$

表示用于从所述基于 DIBR 的 3D 视频数据产生所述额外视图的所述虚拟 3D 图像俘获装置的基线，

其中函数  $f(I_r, I_g)$  表示在将水平移位应用于所述参考视图  $I_r$  的每一像素块之后所述参考视图中的所述每一像素块与所述产生的视图  $I_g$  中的每一对应像素块之间的亮度上的差，

其中所述水平移位的量是  $I_r$  和  $I_g$  的对应块之间的平均视差，以及  
其中所述平均视差随所述参考视图的深度图而变。

17. 根据权利要求 16 所述的设备，其中所述失真度量中的每一个指示在观看所述基于 DIBR 的 3D 视频数据时预期的 3D 观看视觉不适的测量。

18. 根据权利要求 16 所述的设备，其进一步包括用于导出空间错误异常点 S0 失真度量 和时间错误异常点 T0 失真度量中的一个或一个以上的装置。

19. 根据权利要求 16 所述的设备，其进一步包括：

用于将空间错误异常点 S0 失真度量导出为理想深度图与所述基于 DIBR 的 3D 视频数据的所述深度图之间的差的标准偏差的装置；以及

用于将时间错误异常点 T0 失真度量导出为针对所述 3D 视频的两个连续帧的所述理想深度图与所述基于 DIBR 的 3D 视频数据的所述深度图之间的差的标准偏差的装置；

用于将时间不一致性 TI 失真度量导出为用于产生基于 DIBR 的 3D 视频数据和相继的基于 DIBR 的 3D 视频数据的深度图之间的所述差的标准偏差的装置；

用于将所述 S0、T0 和 TI 失真度量中的每一个规格化到 0 到 1 的范围的装置；以及

用于根据以下等式计算所述客观度量的装置：

$$3VQM = K(1-S0(S0 \cap T0))^a(1-TI)^b(1-T0)^c,$$

其中 3VQM 表示所述客观度量，S0、T0 和 SI 表示所述经规格化的 S0、T0 和 TI 失真度量，a、b 和 c 表示通过训练序列确定的常数，而 k 表示缩放因子。

20. 根据权利要求 16 所述的设备，

其中所述基于 DIBR 的 3D 视频数据进一步包含提供的视图，以及

其中所述设备进一步包括：

用于根据一个或一个以上参数从 3D 视频数据估计所述深度图的装置，其中所述 3D 视频数据包含左视图和右视图，其中所述提供的视图选自所述左视图和右视图之一；以及

用于修正用于估计所述 3D 视频数据的所述深度图的所述一个或一个以上参数以便校正由所述客观度量识别的缺陷的装置。

21. 根据权利要求 16 所述的设备，

其中所述基于 DIBR 的 3D 视频数据进一步包含提供的视图，以及

其中所述设备进一步包括：

用于根据一个或一个以上参数对所述基于 DIBR 的 3D 视频数据的所述深度图和所述提供的视图进行编码以产生经编码的基于 DIBR 的 3D 视频数据的装置；以及

用于修正用于对所述深度图和所述提供的视图中的一个或一个以上进行编码的所述一个或一个以上参数以校正由所述客观度量识别的缺陷的装置。

22. 根据权利要求 16 所述的设备，

其中所述 3D 视频数据进一步包含提供的视图，以及

其中所述设备进一步包括：

用于根据一个或一个以上参数对所述深度图的经编码版本和所述提供的视图的经编码版本进行解码以产生所述深度图和所述提供的视图的装置；以及

用于修正用于对所述深度图的所述经编码版本和所述提供的视图的所述经编码版本进行解码的所述一个或一个以上参数以便校正由所述客观度量识别的缺陷的装置。

23. 根据权利要求 16 所述的设备，

其中所述基于 DIBR 的 3D 视频数据进一步包含提供的视图，以及

其中所述设备进一步包括：

用于根据一个或一个以上参数从所述深度图和所述提供的视图重构 3D 视频数据的左视图或右视图以产生所述 3D 视频数据的装置，其中将所述提供的视图选择为所述 3D 视频数据的所述左视图或所述右视图；以及

用于修正用于重构所述左视图或所述右视图的所述一个或一个以上参数以校正由所述客观度量识别的缺陷的装置。



## 用于三维视频的基于视觉的质量度量

[0001] 本申请案主张 2010 年 7 月 16 申请的第 61/364,940 号美国临时申请案的权益,所述申请案的全文以引用的方式并入本文中。

### 技术领域

[0002] 本发明涉及视频再现,且更特定来说,涉及三维视频再现。

### 背景技术

[0003] 三维 (3D) 视频俘获装置一般在大体上模仿人眼布置的构造中包含两个相机。所述两个相机各自俘获场景的二维 (2D) 视频数据,但是从模仿从相应的人左眼和右眼对场景的感知的略微移位的角度来俘获。此被模仿的左眼和右眼 2D 视频数据常被分别称作左眼和右眼 2D 视图。在给定相机的焦距和相机中心之间的基线距离的情况下,可从此被模仿的左眼和右眼 2D 视图提取深度信息。此深度信息可用于扩充所述左眼和 / 或右眼 2D 视图中的一个或一个以上以形成 3D 视频数据。

[0004] 通常,仅结合所述视图中的一者来提供深度信息,因为可从提供的视图和深度信息产生另一视图。从提供的视图和深度信息再现另一视图的此技术被称作基于深度图像的再现 (DIBR)。考虑到仅需要一个视图且可将深度信息编码为灰度级图像 (其比全色 2D 视频数据消耗少得多的空间), DIBR 减小了 3D 视频数据的大小。可进一步压缩 DIBR 中的所得的 3D 视频数据以进一步减小视频数据的大小。对此 3D 视频数据的压缩可促进将此 3D 视频数据无线地递送到 (例如) 无线显示器。

[0005] 3D 视频编码器可实施深度图估计模块以产生包含来自两个所俘获的视图的单一视图和深度信息的 3D 视频数据。3D 视频解码器可实施 DIBR 以从提供的视图和深度信息再现额外视图以供 3D 显示装置呈现。3D 视频编码器和 3D 视频解码器中的每一者可另外执行对 3D 视频数据的一些分析以评估视图的质量。通常,3D 视频编码器和解码器利用现有的 2D 质量度量 (2DQM) 来评估这些视图中的每一者的质量,且以分别推测性地反映所俘获的 3D 视频和所再现的 3D 视频数据的质量的方式组合这些 2D 质量度量。这些 2D 质量度量中的一些已被扩充以考虑深度图度量,从而进一步精细化用于 3D 视频数据的所得的质量度量。响应于此制定的伪 3D 质量度量,3D 视频编码器可从两个所俘获的视图修正深度图的产生,且 3D 视频解码器可从提供的视图和深度信息修正视图的产生。

### 发明内容

[0006] 一般来说,本发明描述用于提供客观三维 (3D) 质量度量 (3DQM) 的技术,其能够实现对错误和其源的适当发现,且能够对深度图估计和深度图和视图编码器 / 解码进行精细调谐。也就是说,不是利用 2D 质量度量 (2DQM) 来个别地评估若干视图中的每一者的质量且组合这些 2D 质量度量以形成可被称作伪 3D 质量度量的度量,所述技术可避免对 2DQM 的推测性组合来支持从理想深度图的估计计算出的客观 3DQM,所述理想深度图使用基于 DIBR 的 3D 视频数据来提供失真受限的图像视图。另外,所述技术可隔离 3D 视频编码器、3D 视频

解码器和 / 或无线信道中的各种操作, 以便与通常用于评估基于 DIBR 的 3D 视频数据的推测性 3DQM 相比潜在地更好地识别错误的源。以此方式, 所述技术可提供客观三维 3DQM, 其能够实现对错误和其源的适当发现, 且能够对深度图估计和深度图和视图编码器 / 解码进行精细调谐。

[0007] 在一个方面中, 一种用于获得客观度量以量化基于基于深度图像的重现 (DIBR) 的三维 (3D) 视频数据的视觉质量的方法包括: 估计将使用基于 DIBR 的 3D 视频数据产生失真受限的图像视图的理想深度图; 基于所述理想深度图与在产生所述基于 DIBR 的 3D 视频数据中使用的深度图的定量比较来导出一个或一个以上失真度量; 以及基于所述导出的一个或一个以上失真度量来计算所述客观度量以量化基于 DIBR 的 3D 视频数据的视觉质量。

[0008] 在另一方面中, 一种装置获得客观度量以量化基于基于深度图像的重现 (DIBR) 的三维 (3D) 视频数据的视觉质量。所述装置包括计算 3D 客观度量的 3D 分析单元。所述 3D 分析单元包含: 理想深度估计单元, 其将使用基于 DIBR 的 3D 视频数据产生失真受限的图像视图的理想深度图; 失真度量计算单元, 其基于所述理想深度图与在产生所述基于 DIBR 的 3D 视频数据中使用的深度图的定量比较来导出一个或一个以上失真度量; 以及客观度量计算单元, 其基于所述导出的一个或一个以上失真度量来计算所述客观度量以量化基于 DIBR 的视频数据的视觉质量。

[0009] 在另一方面中, 一种设备获得客观度量以量化基于基于深度图像的重现 (DIBR) 的三维 (3D) 视频数据的视觉质量。所述设备包括用于估计将使用基于 DIBR 的 3D 视频数据产生失真受限的图像视图的理想深度图的装置; 用于基于所述理想深度图与在产生所述基于 DIBR 的 3D 视频数据中使用的深度图的定量比较来导出一个或一个以上失真度量的装置; 以及用于基于所述导出的一个或一个以上失真度量来计算所述客观度量以量化基于 DIBR 的 3D 视频数据的视觉质量的装置。

[0010] 在另一方面中, 一种非暂时性计算机可读媒体包括在执行时致使一个或一个以上处理器进行以下操作的指令: 估计将使用基于 DIBR 的 3D 视频数据产生失真受限的图像视图的理想深度图; 基于所述理想深度图与在产生所述基于 DIBR 的 3D 视频数据中使用的深度图的定量比较来导出一个或一个以上失真度量; 以及基于所述导出的一个或一个以上失真度量来计算所述客观度量以量化基于 DIBR 的 3D 视频数据的视觉质量。

[0011] 所述技术的一个或一个以上方面的细节在附图及以下描述中陈述。从描述和图式, 且从权利要求书将明白所述技术的其它特征、目的和优势。

## 附图说明

[0012] 图 1 是说明实施本发明中所描述的客观三维 (3D) 质量度量 (3DQM) 导出技术的示范性系统的方框图。

[0013] 图 2 是说明实施本发明中所描述的技术的各种方面的 3D 分析单元的实例的方框图。

[0014] 图 3 是说明在实施本发明中所描述的三维 (3D) 质量度量导出技术的各种方面中的源装置的示范性操作的流程图。

[0015] 图 4 是说明在实施本发明中所描述的技术的各种方面的显示装置的示范性操作的流程图。

[0016] 图 5 是在实施本发明中所描述的技术的各种方面以计算 3D 质量度量的 3D 分析单元的示范性操作的流程图。

[0017] 图 6 是说明提供根据本发明中所描述的技术而产生的 3DQM 的主观分析的图表的图。

[0018] 图 7 是展示移位传感器相机模型的图。

## 具体实施方式

[0019] 图 1 是说明实施本发明中所描述的客观三维 (3D) 质量度量 (3DQM) 导出技术的示范性系统 10 的方框图。在图 1 的实例中,系统 10 包含源装置 12 和显示装置 14,源装置 12 和显示装置 14 经由无线通信信道 15 而彼此通信。源装置 12 可包含通用多媒体装置,例如个人计算机、工作站、个人数字助理 (PDA)、移动电话 (包含所谓的“智能电话”),或包括能够执行软件且尤其是多媒体软件的通用处理器的任何其它类型的装置。源装置 14 可替代地包括专用多媒体装置,例如视频摄像机、数字视频光盘 (DVD) 播放器、电视、机顶盒 (STB)、压缩光盘 (CD) 播放器、数字媒体播放器 (例如,所谓的“MP3”播放器或 MP3/MP4 播放器的组合,以及播放其它格式 (包含高级音频译码 (AAC)、视窗媒体视频 (WMV) 和波形音频视频 (WAV) 格式) 的其它媒体播放器)、数字视频记录器 (DVR)、全球定位系统 (GPS) 装置,或专用于一组一个或一个以上多媒体应用且通常经由加载和执行多媒体软件而不实现用户控制的任何其它装置。

[0020] 显示装置 14 一般表示能够经由显示器进行视频回放的任何装置。显示装置 16 可包括电视 (TV) 显示器 (其可被称作 3D 视频显示装置或混合式 2D/3D 视频显示装置),或能够显示 3D 视频数据的任何其它类型的显示装置。显示装置 14 可替代地包括具有显示器的任何其它装置,例如膝上型计算机、个人媒体播放器 (PMP)、桌上型计算机、工作站、PDA 和便携式数字媒体播放器 (例如,便携式 DVD 播放器)。

[0021] 出于说明的目的,假设显示装置 14 表示与源装置 12 无线地进行通信的无线全 3D 电视。然而,本发明的技术不应限于无线全 3D 电视,而是可相对于各种装置的不同配置以许多不同方式来实施。举例来说,除了显示装置之外,可相对于与 3D 显示器或电视分离但与 3D 显示器或电视介接的机顶盒或其它离散视觉或音频 / 视觉装置来实施所述技术。如本发明中所使用,全 3D 显示器涉及集成全部必要的硬件逻辑、模块、单元、软件或其它组件以容易地显示并实现 3D 视频数据的观看的 3D 显示器。如本发明中所使用的 3D 显示器涉及未包含全部逻辑、模块、单元、软件或其它组件以实现 3D 视频数据的接收、解码、呈现和观看的 3D 显示器。3D 显示器一般需要单独的装置、硬件卡或其它组件来实现对 3D 视频数据的接收、解码、呈现和观看。

[0022] 源装置 12 包含控制单元 16 和接口 18。控制单元 16 可表示一个或一个以上处理器 (图 1 中未展示),所述一个或一个以上处理器执行软件指令,例如用于界定软件或计算机程序的软件指令,存储到计算机可读存储媒体 (再次,图 1 中未展示) 的软件指令,所述计算机可读存储媒体例如为存储装置 (例如,磁盘驱动器,或光学驱动器),或存储器 (例如,快闪存储器、随机存取存储器或 RAM) 或存储用以致使可编程处理器执行本发明中所描述的技术的指令 (例如,呈计算机程序或其它可执行体的形式) 的任何其它类型的易失性或非易失性存储器。或者,控制单元 16 可包括用于执行本发明中所描述的技术的专用硬

件,例如一个或一个以上集成电路、一个或一个以上专用集成电路(ASIC)、一个或一个以上专用特殊处理器(ASSP)、一个或一个以上现场可编程门阵列(FPGA)或专用硬件的前述实例的任何组合。

[0023] 接口 18 表示借以与另一装置(例如,显示装置 14)无线地通信的接口。虽然在图 1 的实例中未展示,但源装置 12 可包含用以经由有线或无线连接与其它装置进行通信的额外接口。这些额外接口可包括(例如)Bluetooth™接口、无线蜂窝式接口、高清晰度多媒体接口(HDMI)、微型 HDMI、通用系统总线(USB)接口、外部串行高级技术附件(eSATA)接口、串行接口、单独视频(s-视频)接口、分量视频接口、RCA 接口,或用于与另一装置进行通信的任何其它类型的有线或无线接口。

[0024] 在图 1 的实例中,源装置 12 的控制单元 16 包含许多单元 20 到 24,许多单元 20 到 24 可共同地实施或单独地实施为一个或一个以上硬件单元或软件与一个或一个以上硬件单元的组合。深度估计单元 20 表示基于包含左视图 26A 和右视图 26B 的 3D 视频数据 26 来估计深度信息 28 的单元。可使用 3D 视频俘获装置来俘获 3D 视频数据 26,所述 3D 视频俘获装置包含第一 2D 视频俘获装置(例如,第一相机)和第二 2D 视频俘获装置(例如,第二相机),其中所述第一和第二 2D 视频俘获装置经定位以模仿人左眼和右眼在人头部的放置。第一 2D 视频俘获装置可俘获模仿场景的来自人左眼的视图的 2D 视频数据,而第二 2D 视频俘获装置可俘获模仿同一场景的来自人右眼的视图的 2D 视频数据。左视图 26A 表示由第一 2D 视频俘获装置俘获的 2D 视频数据,且右视图 26B 表示由第二 2D 视频俘获装置俘获的 2D 视频数据。

[0025] 虽然出于容易说明的目的而未在图 1 中展示,但源装置 12 可包含用于俘获 3D 视频数据 26 的额外单元、模块、逻辑、硬件或其它组件,例如上述的 3D 视频俘获装置。或者,源装置 12 可充当用于存储 3D 视频数据 26 的存档或库。在一些例子中,源装置 12 可经由包含于源装置 12 内的接口 18 无线地接收 3D 视频数据 26。在一些例子中,这些外部装置可经由接口 18 与源装置 12 介接以将 3D 视频数据存储于源装置 12 内。

[0026] 3D 视频编码器 22 表示以压缩方式编码 3D 视频数据 26 和深度信息 28 的单元。更具体来说,3D 视频编码器 22 表示以促进基于深度图像的再现(DIBR)的方式编码 3D 视频数据 26 的视图 26A、26B 中的一者以及深度信息 28 的单元。DIBR 涉及从提供的视图和深度信息再现虚拟视图。DIBR 的益处是仅左视图和右视图中的单一者需要传输,而不是传输两个视图。另外,DIBR 提供通常作为灰度级图像或所谓的深度图(其可在大小上显著小于其它视图)而提供的深度信息。实施 DIBR 的解码器可随后再现或另外产生未从深度图和提供的视图发送的视图。3D 视频编码器 22 可进一步压缩视图 26A、26B 中的一者以及深度信息 28 以进一步压缩 3D 视频数据 26。3D 视频编码器包含用以压缩视图 26A、26B 中的一者的视图译码单元 30 以及用以压缩深度信息 28 的深度译码单元 32。3D 视频编码器可格式化或另外封装或包封经编码的视图和深度信息以作为经编码的基于 DIBR 的 3D 视频数据 34。

[0027] 3D 分析单元 24 表示执行或另外实施本发明中所描述的技术以产生客观地评估经编码的基于 DIBR 的 3D 视频数据 34 的质量的 3D 质量度量 36(“3DQM 36”)的单元。3DQM 还可在本发明中被称作 3D 视频质量度量(3VQM)。虽然展示为与深度估计单元 20 和 3D 视频编码器 22 分离的单元,但 3D 分析单元 24 可替代地被集成到深度估计单元 20 和 3D 视频编码器 22 中的一者或两者中,以执行或另外实施在本发明中在下文更详细地描述的技术。

[0028] 如图 1 的实例中进一步展示,显示装置 14 包含控制单元 38(其可大体上类似于源装置 12 的控制单元 16),和接口 40(其可大体上类似于源装置 12 的接口 18)。显示装置 14 还包含 3D 显示器 42,3D 显示器 42 表示能够呈现 3D 视频数据以供显示装置 14 的观看者消费的任何显示器,包含可需要快门、偏光或彩色眼镜的立体显示器和不需要眼镜或任何其它技术来观看 3D 中的内容的自动立体显示器。3D 显示器 42 可包括任何类型的显示器,包含有机发光二极管(OLED)显示器、发光二极管(LED)显示器、等离子体显示器和阴极射线管(CRT)显示器。源装置 14 的控制单元 16 包含许多单元 44 到 48,许多单元 44 到 48 可共同地实施或单独地实施为一个或一个以上硬件单元或软件与一个或一个以上硬件单元的组合。

[0029] 3D 视频解码器 44 表示对经编码的基于 DIBR 的 3D 视频数据 34 进行解码以产生提供的视图 50A(出于说明的目的,假设其为左视图 50A)和深度信息 52 的单元。3D 视频解码器 44 包含视图解码单元 54 和深度解码单元 56。3D 视频解码器 44 可将视图解码单元 54 和深度解码单元 56 个别地实施为单独的硬件单元或单一的硬件单元,所述硬件单元可通过或不通过执行软件或指令来实施本发明中所描述的一些功能性。视图解码单元 54 表示对基于 DIBR 的视频数据 34 的经编码的提供的视图进行解码以产生提供的视图 50A 的单元。深度解码单元 56 表示对经编码的深度图或深度信息进行解码以产生深度信息 58 的单元。

[0030] 视图重构单元 46 实施 DIBR 以基于提供的视图 50A 和深度信息 58 来重构右视图 50B。出于此原因,右视图 50B 可被称作重构视图 50B(“recon(右)视图 50B”)。视图重构单元 46 可实施与由深度估计单元 20 实施的操作相反的操作以重构右视图 50B。视图重构单元 46 可执行基于深度图像的再现(DIBR),其为视图重构单元 46 实施的操作或过程,涉及使用深度图将来自一个视图(即,在此实例中是提供的视图)的内容映射到其它视图中的给定位置。视图重构单元 46 可随后实施各种过程以将空位置填充在所产生的视图中(其通常被称作“孔”)。3D 分析单元 48 执行本发明中所描述的类似于 3D 分析单元 24 所执行的技术的技术,但出于下文更详细地描述的原因,其经过扩充以考虑到可用于 3D 分析单元 48 的信息中的差异。

[0031] 起初,源装置 12 可接收 3D 视频数据 26。在一些例子中,源装置 12 从包含于源装置 12 内或经由有线或无线连接耦合到源装置 12 的 3D 俘获装置接收 3D 视频数据 26。或者,如上文所述,源装置 12 经由接口 18 或上文所列举的其它额外接口中的一者接收 3D 视频数据 26。源装置 12 随后对 3D 视频数据 26 进行编码以压缩此 3D 视频数据 26,以用于经由接口 18 和无线通信信道 15 递送到显示装置 14。虽然在图 1 的实例中未展示,但控制单元 16 可包含呈现一个或一个以上用户接口的用户接口单元或模块,用户可用所述一个或一个以上用户接口进行交互以起始对 3D 视频数据 26 的编码和 3D 视频数据 26 到显示装置 14 的传递。

[0032] 一旦起始对 3D 视频数据 26 的编码和传递,源装置 12 的控制单元 16 便调用深度估计单元 20 以确定深度信息 28。控制单元 16 还调用 3D 视频编码器 22 来对视图 26A、26B 中的一者和深度信息 28 进行编码。如上文所述,假设 3D 视频编码器 22 对左视图 26A 进行编码。给定此假设,3D 视频编码器 22 调用用以对左视图 26A 进行编码的视图译码单元 20 和用以对深度信息 28 进行编码的深度译码单元 32。3D 视频编码器 22 输出经编码的基于

DIBR 的 3D 视频数据 34, 经编码的基于 DIBR 的 3D 视频数据 34 经由接口 18 和无线通信信道 15 被传输到显示装置 14。

[0033] 接口 40 接收此经编码的基于 DIBR 的 3D 视频数据 34 且将此视频数据 34 转发到包含于显示装置 14 内的控制单元 38 的 3D 视频解码器 44。3D 视频解码器 44 调用视图解码单元 54 以对经编码的左视图 26A 进行解码。视图解码单元 54 对经编码的左视图 26 进行解码以产生提供的视图 50A。3D 视频解码器 44 还调用深度解码单元 56 以对经编码的基于 DIBR 的 3D 视频数据 34 的经编码的深度信息 28 进行解码。深度解码单元 56 对经编码的深度信息 28 进行解码以产生深度信息 58。控制单元 38 调用视图重构单元 46 以重构右视图 50B, 于是, 视图重构单元 46 使用 DIBR 基于提供的视图 50A 和深度信息 58 而形成所重构的右视图 50B。左视图 50A 和右视图 50B 被作为 3D 视频数据 50 转发到 3D 显示器 42, 3D 显示器 42 将此 3D 视频数据 50 呈现给 3D 显示器 42 的观看者。

[0034] 在此基于 DIBR 的编码和基于 DIBR 的解码的整个过程中, 3D 分析单元 24、48 中的每一者根据本发明中所描述的技术分别确定 3DQM 36 和 60。首先参看源装置 12 的 3D 分析单元 24, 3D 分析单元 24 根据下文更详细地描述的本发明的技术估计将使用基于 DIBR 的 3D 视频数据产生失真受限的图像视图的理想深度图。简要说, 3D 分析单元 24 至少基于所产生或所重构的视图来估计理想深度图。也就是说, 虽然未在图 1 的实例中展示, 但 3D 视频编码器 22 还包含与显示装置 14 的 3D 视频解码器 44 类似的 3D 视频解码器。同样, 深度估计单元 20 还可包含与视图重构单元 46 类似的视图重构单元 46, 以便实施 DIBR 且从包含于 3D 视频编码器 22 内的 3D 视频解码器的输出重构所产生的视图。控制单元 16 包含这些解码和视图重构单元以促进对选定视图 (即, 在此实例中是左视图 26A) 和深度信息 28 的深度估计和后续编码。深度估计单元 20 可因此基于经解码的深度信息 58' 和经解码的左视图 50A' 而输出所重构的右视图 50B'。3D 分析单元 24 随后至少基于所产生或所重构的视图 50B' 而产生理想深度图 62 (“DM 62”)。

[0035] 在被给予全部 3D 视频数据 26, 而不是包含单一参考或提供的视图和深度图的仅基于 DIBR 的视频数据的情况下, 3D 分析单元 24 可依照原始视图 (即, 右视图 26B) 和所产生的或所重构的视图 50B' 两者而计算出理想深度图 62。此所谓的“全参考”背景 (其中 3D 分析单元 24 具有对所俘获的 3D 视频数据 26 的存取权) 可提供更准确的深度图 62。虽然关于此“全参考”背景进行描述, 但 3D 分析单元 24 可关于其它背景实施本发明中所描述的技术, 包含所谓的“简化参考”背景和“无参考”背景, 以上两者在下文更详细地描述。

[0036] 通常, 源装置 12 的 3D 分析单元 24 在全参考或简化参考背景中计算理想深度图 62, 而 3D 分析单元 48 在简化参考或部分参考背景中计算其深度图 70。这两个 3D 分析单元 24 如何计算理想深度图 62、70 的差随 3D 视频数据的可用性而变, 如下文更详细地描述。全参考背景涉及其中 3D 视频数据 26 完全可用的例子, 即, 视图 26A、26B 两者都可用而不是仅视图 26A、26B 中的一者可用。简化参考背景涉及其中 3D 视频数据 26 仅部分可用的例子, 其原因是为左视图 26A 和右视图 26B 二者均提供了深度图, 但仅提供了视图 26A、26B 中的单一者。部分参考背景指其中 3D 视频数据 26 仅包含基于 DIBR 的 3D 视频数据的例子, 即, 视频数据包含左视图和右视图中的单一者以及深度图或其它深度信息。

[0037] 3D 分析单元 24 随后通过理想深度图 62 与给定深度信息或深度图 28 的定量比较而导出一个或一个以上失真度量 64 (“度量 64”)。3D 分析单元 24 可将度量 64 中的一个

或一个以上导出为（例如）理想深度图 62 与深度图 28 之间的差的标准偏差。作为替代或附加，3D 分析单元 24 可将度量 64 中的一个或一个以上导出为针对 3D 视频数据 26 的第一帧的理想深度图 62 与深度图 28 之间的差与针对 3D 视频数据 26 的第二帧的理想深度图 62 与深度图 28 之间的差的改变的标准偏差。作为替代或附加，3D 分析单元 24 可将度量 64 中的一个或一个以上导出为针对 3D 视频数据 24 的第一帧而计算出的深度图 28 与针对 3D 视频数据 24 的第二帧而计算出的深度图 28 之间的差的标准偏差。

[0038] 3D 分析单元 24 随后基于导出的失真度量 64 来计算 3DQM 36。在一些例子中，3D 分析单元 24 将 3DQM 36 产生为所有失真度量 64 的数学组合。在其它例子中，3D 分析单元 24 将 3DQM 36 计算为失真度量 64 的仅一子集的组合。无论如何，3DQM 36 表示客观度量，原因在于其未被计算为主观导出的度量的组合。实际上，3D 分析单元 24 之所以将 3DQM 36 计算为客观导出的度量 64 的组合，原因是这些度量 64 是通过给定深度图 28 与客观理想深度图 62 的估计的比较而导出。此理想深度图 62 表示参考深度图 62，其将使用基于 DIBR 的视频数据来产生图像视图中的受限失真，且因此提供客观标准，相对于所述客观标准来评估所得的经编码的基于 DIBR 的 3D 视频数据 34。

[0039] 因此，除了执行深度图 28 与深度图 58' 的常规比较以便导出伪 3DQM（其中深度图 28 可固有地包含归因于由深度估计单元 20 执行的深度估计而引起的错误）之外，客观理想深度图 62 避免了这些固有错误，因为 3D 分析单元 24 以校正这些固有错误的方式计算出此深度图 62。另外，3D 分析单元 24 依照度量 64 而产生 3DQM 36，度量 64 经导出以考虑到可能被视为在观看所得的 3D 视频数据 50 中所感知的视觉不适的东西。换句话说，不是盲目地将一些深度度量与传统的 2D 度量进行组合以计算伪 3DQM，3D 分析单元 24 实施本发明中所描述的技术来计算 3DQM 36 以使得其考虑到所感知的视觉不适。

[0040] 一旦 3D 分析单元 24 计算出 3DQM 36，3D 分析单元 24 可将 3DQM 36 传输到深度估计单元 20 和 3D 视频编码器 22 中的每一者，深度估计单元 20 和 3D 视频编码器 22 中的每一者随后基于此 3DQM 36 来更新一个或一个以上参数以校正由 3DQM 36 识别的至少一些视觉不适。或者，3D 分析单元 24 可针对深度估计单元 20 和 3D 视频编码器 22 中的每一者产生一组新的一个或一个以上参数 66、68，且将这些新的参数 66 和 68 转发到深度估计单元 20 和 3D 视频编码器 22 中的每一者。这些新的参数 66 和 68 有效地重新配置深度估计单元 20 和 3D 视频编码器 22 以调整深度估计和编码，从而试图校正在观看由 3DQM 36 识别的经编码的基于 DIBR 的 3D 视频数据 34 时所预期的所感知的视觉不适。3D 分析单元 24 通常与由深度估计单元 20 所执行的深度估计和由 3D 视频编码器 22 所执行的编码并发地执行此整个过程，以便实现对这些单元 20、22 的动态（即，实时的或准实时的）重新配置，从而校正在观看经编码的基于 DIBR 的 3D 视频数据 34 时所预期的所感知的视觉不适。

[0041] 如上文所述，3D 分析单元 48 同样计算 3DQM 60。3D 分析单元 48 可大体上类似于 3D 分析单元 24，除了 3D 分析单元 24 可基于不同信息来估计理想深度图 (DM) 70 之外。3D 分析单元 24 仍至少部分基于所重构的右视图 50B 来估计深度图 70。然而，3D 分析单元 48 可不且通常不具有对原始右视图 26B 的存取权，因为这将阻挠使用基于 DIBR 的方案的目的，因为将发送两个视图 26A、26B，从而消除了对仅发送这些视图 26A、26B 中的单一者以及深度信息或深度图 28 的节省。因此，3D 分析单元 48 一般驻留于简化参考或部分参考背景中，且在不具有对右视图 26B 的存取权的情况下确定理想深度图 70。

[0042] 以大体上类似于 3D 分析单元 24 的方式,3D 分析单元 48 基于理想深度图 70 和所重构的深度图 58 而导出或另外计算出失真度量 72 中的一个或一个以上。再次,类似于 3D 分析单元 24,3D 分析单元 48 基于度量 72 而计算出 3DQM 60,且将此 3DQM 60 提供给 3D 视频解码器 44 和视图重构单元 46,使得这些单元 44、46 可更新其相应的参数来考虑到由 3DQM 60 识别的任何所感知的观看者不适。或者,3D 分析单元 48 可基于 3DQM 60 来确定参数 74 和 76,且用所确定的参数 74 来更新 3D 视频解码器 44,且用参数 76 来更新视图重构单元 46。

[0043] 这些新参数 74 和 76 有效地重新配置 3D 视频解码器 44 和视图重构单元 46 以进行解码和视图重构,从而试图校正在观看由 3DQM 60 识别的经编码的基于 DIBR 的 3D 视频数据 34 时所预期的所感知的视觉不适。3D 分析单元 48 通常与由 3D 视频解码器 44 所执行的解码和由视图重构单元 46 所执行的视图重构并发地执行此整个过程,以便实现对这些单元 44、46 的动态(即,实时的或准实时的)重新配置,从而校正在观看经编码的基于 DIBR 的 3D 视频数据 34 时所预期的所感知的视觉不适。3D 显示器 42 可随后呈现经解码的 3D 视频数据 50 以用于呈现给 3D 显示器 42 的一个或一个以上观看者。

[0044] 以此方式,不依赖于针对左视图和右视图所计算出的常规的 2D 质量度量且随后组合这些度量以形成忽略深度信息的某一类伪 3DQM,本发明中所描述的技术制定出真实的 3DQM,其相对于客观理想深度图来评估失真。在 DIBR 中,深度图中的错误(例如,归因于错误的估计、数值四舍五入和压缩假影)导致相对的像素位置中以及所重构的视图 50B 的像素值的量值中的错误。所重构的视图 50B 中的这些错误的视觉效应以显著亮度变化的形式在纹理区域周围在空间上引人注目,且以闪烁的形式在平坦区周围在时间上引人注目。基于 DIBR 的 3D 视频中的视觉不适可另外源自若干因素,包含过度视差、快速改变的视差、几何失真以及各种深度线索之间的不一致性(例如未匹配的对象色彩)。本发明中所描述的技术以 3DQM 36 和 60 的形式客观地量化基于 DIBR 的 3D 视频的视觉质量,以便潜在地动态地校正由 3DQM 36 和 60 识别的视觉不适。

[0045] 另外,可以类似于速率-失真压缩算法的方式的方式实施所述技术。也就是说,在速率-失真压缩算法中,可增加对视频的压缩(例如通过量化残余值),以便在潜在地在经解压的视频中引入更多假影的折衷下减少无线信道上的带宽消耗。在这些压缩系统中,可减少压缩以减少将假影引入到经解压的视频中,但此可导致消耗更多无线带宽的更大文件。3DQM(3DQM 36 或 60)可识别通过压缩和深度估计引入的假影,且驱动参数调整,其导致与上文关于速率-受约束的压缩算法所描述的折衷类似的折衷。在此意义上,可使用 3DQM 在引入更多假影的代价下增加压缩,或在增加带宽消耗的代价下减少压缩。

[0046] 在此方面中,3DQM 可实现对基于 DIBR 的 3D 视频数据产生和压缩的裁定以适合特定应用。也就是说,3DQM 可提供在估计准确的深度图中所涉及的计算的复杂性与针对这些图像引入的假影之间的折衷,其类似于压缩中的速率-失真折衷。通过实现此些折衷,3DQM 可促进在其中处理或电池资源是有限的移动或嵌入式实时应用中实施 3D 视频数据技术,例如在移动装置中实施 3D 视频数据技术,所述移动装置例如为蜂窝式电话、膝上型计算机、所谓的上网本、个人数字助理(PDA)等。

[0047] 虽然上文关于两个视图系统进行描述,但可在其中将两个以上视图提供给 3D 显示器从而允许显示多个不同视图的多视图 3D 系统中实施所述技术。举例来说,取决于观看



者相对于 3D 显示器在何处,一些 3D 显示器能够显示多个视图。在显示器的左边的观看者可接收到与关于 3D 显示器居中的观看者或在 3D 显示器左边的观看者不同的视图。这些 3D 显示器可并发地显示这些视图中的每一者,其中可使用本发明中所描述的 DIBR 过程进行对这些视图中的每一者的编码和解码。在此方面中,所述技术可用于针对这些视图中的每一者或针对作为一群组的视图提供 3DQM,其中这些一个或一个以上 3DQM 可促进对预期在观看这些视图时发生的至少一些视觉不适的识别和潜在后续的校正。因此,虽然在本发明中出于容易说明的目的关于二视图系统进行描述,但所述技术可用于多视图系统中。

[0048] 图 2 是说明实施本发明中所描述的技术的各种方面的 3D 分析单元 80 的实例的方框图。3D 分析单元 80 可表示在图 1 的实例中所展示的 3D 分析单元 24 或 48 中的任一者,且对 3D 分析单元 80 的以下论述的各个部分可应用于 3D 分析单元 24 和 48 中的每一者。在一些例子中,3D 分析单元 24 和 48 中的每一者可大体上彼此类似,使得 3D 分析单元 80 可一般表示 3D 分析单元 24 和 48 两者。在其它例子中,3D 分析单元 24 可适于特定背景,例如上述的全参考背景,使得 3D 分析单元 80 仅部分表示此经调适的 3D 分析单元 24。同样,3D 分析单元 48 可适于特定背景,例如上述的部分参考背景,使得 3D 分析单元 80 仅部分表示此经调适的 3D 分析单元 48。

[0049] 如图 2 的实例中所展示,3D 分析单元 80 包含理想深度估计单元 82、失真度量计算单元 84、3DQM 计算单元 86 和参数产生单元 88。这些单元 82 到 88 中的每一者可单独地实施为硬件或硬件与软件的组合。或者,单元 82 到 88 中的一个或一个以上可共同地实施为硬件或硬件与软件的组合。

[0050] 理想深度估计单元 82 一般表示估计理想深度且输出理想深度图 90 的单元。理想深度估计单元 82 包含许多单元 92 到 96,许多单元 92 到 96 基于 3D 视频数据的不同层级的可用性来估计理想深度图 90。当完全可使用 3D 视频数据而不是仅基于 DIBR 的 3D 视频数据时,全参考单元 92 估计理想深度图 90。也就是说,全参考单元 92 基于原始视图 98 以及原始视图的重构或所重构的视图 100 (“recon 视图 100”)来估计理想深度图 90。全参考单元 92 还基于所确定的深度信息 102 (“深度 info 102”)和 3D 俘获装置参数(例如,焦距 103 和基线 105(其为作为 3D 俘获装置的 2D 相机的中心之间的直线距离的量度))来估计此理想深度图 90。下文关于图 7 更详细地描述焦距 103 和基线 105。在一些例子中,这些 3D 俘获装置参数可反映虚拟 3D 俘获装置的参数,原因在于这些参数不是实际 3D 俘获装置的参数,而仅仅是表示通常用于将 3D 俘获装置建模的共同参数。在数学上,可通过以下等式 (1) 来表示由全参考单元 92 执行的全参考估计:

$$[0051] \quad Z_{IDEAL} = \frac{\pm FB}{k(I_o - I_g) \pm \frac{FB}{Z}}, \quad (1)$$

[0052] 其中变量  $Z_{IDEAL}$  是指理想深度图 90 且变量 F 是指焦距 103。以上等式 (1) 中的变量 B 是指基线 105,变量 k 是指常数值,变量  $I_o$  是指原始视图或图像 98,变量  $I_g$  是指所产生的或所重构的视图或图像 100,且变量 Z 是指深度信息 102 的深度图。在本发明中对视图或图像的参考是指形成每一视图或图像的实际像素值。将所计算的理想深度图计算为灰度级值的二维阵列。

[0053] 简化参考单元 94 也估计理想深度图 90,但当存在 3D 视频数据的简化或部分可用

性时,是在不同的简化参考背景中进行。也就是说,简化参考单元 94 基于提供的视图 104 和所重构的视图 100 而估计理想深度图 90。简化参考单元 94 还基于所确定的深度信息 102 和 3D 俘获装置参数、焦距 103 和基线 105 而估计此理想深度图 90。关于深度信息 102,简化参考单元 94 一般需要用于左视图的深度图和用于右视图的深度图两者。深度估计单元(例如,深度估计单元 20)一般能够提供对应于左视图和右视图中的每一者的这两个深度图,且在图 1 的实例中所展示的深度信息 28 可包含这两个深度图。深度译码单元 32 可对这两个深度图进行编码且将这些深度图作为经编码的基于 DIBR 的 3D 视频数据 34 进行传输。在数学上,可通过以下等式 (2) 来表示由简化参考单元 94 执行的简化参考估计:

$$[0054] \quad Z_{IDEAL} = \frac{\pm FB}{k(I_R - I_g) \pm FB \left( \frac{1}{Z_L} + \frac{1}{Z_R} \right) + h}, \quad (2)$$

[0055] 其中变量  $Z_{IDEAL}$  是指理想深度图 90 且变量 F 是指焦距 103。以上等式 (2) 中的变量 B 是指基线 105, 变量 k 是指常数值, 变量 h 是指根据下文关于图 7 更详细地描述的移位传感器相机模型而相对于传感器中心的投影中心的水平方向上的移位。变量  $I_R$  是指参考或提供的视图或图像 104 (“prov 视图 104”), 变量  $I_g$  是指所产生的或所重构的视图或图像 100, 且变量  $Z_L$  是指深度信息 102 的左深度图, 且  $Z_R$  是指深度信息 102 的右深度图。

[0056] 无参考单元 96 表示估计理想深度图 90 的又一单元, 但当仅存在基于 DIBR 的 3D 视频数据时, 是在被称作“无参考”背景的又一背景中进行。也就是说, 无参考单元 96 基于提供的视图 104 和所重构的视图 100 而估计理想深度图 90。无参考单元 94 还基于深度信息 102 和 3D 俘获装置参数、焦距 103 和基线 105 而估计此理想深度图 90。在数学上, 可通过以下等式 (3) 来表示由无参考单元 94 执行的无参考估计:

$$[0057] \quad Z_{IDEAL} \approx \frac{\pm FB}{f(I_R, I_g) \pm \frac{FB}{Z}}, \quad (3)$$

[0058] 其中变量  $Z_{IDEAL}$  是指理想深度图 90 且变量 F 是指焦距 103。以上等式 (3) 中的变量 B 是指基线 105, 变量  $I_R$  是指参考或提供的视图或图像 104 (“prov 视图 104”), 变量  $I_g$  是指所产生的或所重构的视图或图像 100, 且变量 Z 是指深度信息 102 的深度图。等式 (3) 中的函数 f 是指计算  $I_R$  和  $I_g$  的对应块之间的平均视差 (d) 且通过值 d 将一移位应用于  $I_R$  中的对应块的函数。函数 f 随后输出经移位的  $I_R$  与所产生的视图  $I_g$  之间的亮度差。

[0059] 失真度量计算单元 84 表示基于理想深度图估计 90 来计算失真度量的单元。更具体来说, 失真度量计算单元 84 可将失真度量计算为深度信息 102 的深度图与理想深度图估计 90 之间的差的定量度量。虽然深度信息 102 与理想深度图估计 90 之间的差并非总是识别视觉不适, 但深度中的不一致错误确实导致视觉不适。出于此原因, 许多失真度量测量不一致性而非错误, 因为一致的错误可能不导致视觉不适。

[0060] 举例来说, 在空间域中, 特定深度平面上的一致的 (或均一的) 错误导致整个平面在一个方向上移位, 且此类错误的感知效应将是在所感知的深度上的略微增加或减小, 其一般不提供较大的视觉不适。然而, 深度中的不一致错误导致离位的色彩像素 / 块, 其归因于未匹配的对象色彩而以深度线索中的不一致性的形式产生视觉不适。

[0061] 因此, 失真度量计算单元 84 计算评估不一致性而不是仅识别错误的失真度量。在

将理想深度估计 90 与一个或一个以上失真度量（例如，下文所描述的失真度量 112 到 116）进行比较的过程中，失真度量计算单元 84 俘获由深度图估计和压缩导致的错误以及由于处理所合成或所产生的彩色视频自身所导致的错误（例如，归因于填充算法和视频压缩过程）。由于处理所产生的彩色视频所导致的错误得以识别，因为理想深度图估计 90 是依照给定深度和彩色视频自身而产生。

[0062] 失真度量计算单元 84 包含空间错误异常点 (S0) 失真度量单元 106 (“S0 失真度量单元 106”)、时间错误异常点 (T0) 失真度量单元 108 (“T0 失真度量单元 108”)，和时间不一致性 (TI) 失真度量单元 110 (“TI 失真度量单元 110”)。S0 失真度量单元 106 表示基于理想深度图估计 90 来计算 S0 失真度量 112 的单元。具体来说，S0 失真度量单元 106 将 S0 失真度量 112 计算为深度信息 102 的给定深度图与理想深度图估计 90 之间的差的标准偏差，其可通过以下等式 (4) 在数学上表达：

$$[0063] \quad S0 = \text{std}(\Delta z), \quad (4)$$

[0064] 其中变量 S0 是指 S0 失真度量 112，且  $\text{std}(\Delta Z)$  是指深度信息 102 的深度图与理想深度图估计 90 之间的差的标准偏差。S0 失真度量单元 106 有效地量化针对 3D 视频数据的给定图像或帧的空间不一致性。

[0065] S0 失真度量 112 一般俘获由预处理导致的噪声以及还有在用于产生 3D 视频数据的卷绕过程中的不准确性。卷绕一般是指围绕 3D 体积（例如，多边形）卷绕 2D 图像。卷绕可涉及固有的近似以及相机建模的近似。不具有任何处理深度的所合成或所重构的图像的视差不与从相机理想上获取的图像的视差确切相同。因此，不同深度平面中的移位将不完美匹配。为了将深度图与理想深度图之间的差中的这些错误与归因于处理而引起的错误进行分离，使用此差的标准偏差来计算异常点。这些异常点是由深度图处理导致的噪声加上归因于不适当的平面移位而引起的边缘。以此方式，S0 识别由深度图处理和不适当的平面移位导致的异常点。

[0066] T0 失真度量单元 108 表示基于理想深度图估计 90 来计算 T0 失真度量 114 的单元。具体来说，T0 失真度量单元 108 将 T0 失真度量 114 计算为针对 3D 视频数据的给定帧的深度信息 102 的给定深度图与针对同一帧的理想深度图估计 90 之间的差减去针对 3D 视频数据的后续帧的深度信息 102 的给定深度图与针对同一后续帧的理想深度图估计 9 之间的差的标准偏差，所述标准偏差可通过以下等式 (5) 在数学上表达：

$$[0067] \quad T0 = \text{std}(\Delta Z_{t+1} - \Delta Z_t). \quad (5)$$

[0068] 以上等式 (5) 中的变量 T0 是指 T0 失真度量 114，且  $\text{std}(\Delta Z_{t+1} - \Delta Z_t)$  是指针对 3D 视频数据的给定帧 (t) 的深度信息 102 的给定深度图与针对同一帧的理想深度图估计 90 之间的差减去针对 3D 视频数据的后续帧 (t+1) 的深度信息 102 的给定深度图与针对同一后续帧的理想深度图估计 9 之间的差的标准偏差。

[0069] T0 失真度量单元 106 有效地量化时间不一致性，其可以显著亮度变化的形式在纹理区域周围，且以闪烁的形式在平坦区周围在空间上引人注目。T0 后面的推断是，由噪声引入的错误在时间上不一致，而边缘将在时间上一致，因为同一卷绕过程产生两个帧。在此方面中，T0 失真度量 114 滤出 S0 中的边缘且仅保持噪声贡献。

[0070] 快速改变的视差或不一致性是视觉不适的另一源，且主要是由深度估计和填充算法中的错误或压缩导致。这些不一致性可由 TI 失真度量 116 识别，其中 TI 失真度量单元

110 表示计算此 TI 失真度量 116 的单元。TI 失真度量单元 110 可根据以下等式 (6) 来计算 TI 失真度量 116：

$$[0071] \quad TI = \text{std}(Z_{t+1}-Z_t), \quad (6)$$

[0072] 其中变量 TI 是指 TI 失真度量 116, 且  $\text{std}(Z_{t+1}-Z_t)$  是指针对 3D 视频数据的给定帧 (t) 的深度信息 102 与针对 3D 视频数据的后续帧 (t+1) 的深度信息 102 之间的差的标准偏差。虽然已在上文描述了许多示范性失真度量 112 到 116, 但本发明中所描述的技术应不限于这些示范性失真度量 112 到 116, 而是可包含用于评估 3D 视频数据的所预期的视觉不适的任何其它类型的失真度量。

[0073] 3DQM 计算单元 86 表示基于失真度量 (例如, 失真度量 112 到 116) 中的一个或一个以上来计算 3DQM 118 的单元。3DQM 计算单元 86 可规格化或另外调整失真度量 112 到 116, 之后组合失真度量 112 到 116 以形成 3DQM 118。3DQM 计算单元 86 可根据以下示范性等式 (7) 计算 3DQM 118：

$$[0074] \quad 3DQM = K(1-S0(S0 \cap T0))^a(1-TI)^b(1-T0)^c. \quad (7)$$

[0075] 在以上等式 (7) 中, 变量 3DQM 是指 3DQM 118, 变量 S0 是指 S0 失真度量 112, 变量 T0 是指 T0 失真度量 114, 变量 TI 是指 TI 失真度量 116, 变量 K 是指用以在范围 [1-5] 内缩放最终的 3DQM 度量以映射到 MOS (平均意见得分) 范围上的常数, 且变量 a、b 和 c 表示由训练序列确定的常数值。实际上, 3DQM 计算单元 86 产生 3DQM 118 以识别预期观看者在观看 3D 视频数据 (针对其确定此 3DQM 118) 时将经历的至少一些视觉不适。

[0076] 参数产生单元 88 表示基于 3DQM 118 而产生参数 120 的单元。参数产生单元 88 产生参数 120 以校正由 3DQM 118 识别的至少一些视觉不适。参数产生单元 88 可确定参数 120, 参数 120 更新深度估计单元 (例如, 深度估计单元 20)、3D 视频编码器 (例如, 3D 视频编码器 22)、3D 视频解码器 (例如, 3D 视频解码器 44) 以及视图重构单元 (例如, 视图重构单元 46) 中的任一者, 此取决于 3D 分析单元 80 在其中操作的背景。

[0077] 举例来说, 3D 分析单元 80 可在发起 3D 视频数据的源装置 (例如, 源装置 12) 内操作。假设 3D 分析单元 80 表示 3D 分析单元 24 (出于说明的目的), 则 3D 分析单元 80 接收包含左视图 26A 和右视图 26B 两者的全 3D 视频数据 26。3D 分析单元 80 还用标准预先配置, 或接收焦距 103 和基线 105。另外, 3D 分析单元 80 从深度估计单元 20 接收呈所重构的右视图 50B' 的形式的所重构的右视图 100。3D 分析单元 80 进一步接收呈深度信息 28 的形式的深度信息 102, 深度信息 102 可包括对应于左视图 26A 的左深度图和对应于右视图 26B 的右深度图。

[0078] 基于此接收的信息, 考虑到 3D 分析单元 80 接收左视图 26A 和右视图 26B 两者以及包括左深度图和右深度图的深度信息 28, 3D 分析单元 80 的理想深度估计单元 82 调用全参考单元 92。一旦经调用, 全参考单元 92 随后根据上文列举的等式 (1) 计算理想深度图估计 90, 理想深度图估计 90 在图 1 的源装置 12 的背景中表示理想深度图 62。全参考单元 92 将理想深度图 62 输出到失真度量计算单元 84。

[0079] 在接收到理想深度图 62 后, 失真度量计算单元 84 即刻调用 S0 失真度量单元 106、T0 失真度量单元 108 和 TI 失真度量单元 110 中的每一者, 以根据对应的等式 (3) 到 (6) 来计算失真度量 112 到 116 中的相应者。在图 1 的源装置 12 的背景中, 失真度量 112 到 116 表示度量 64。单元 106 到 110 将失真度量 64 输出到 3DQM 计算单元 86。3DQM 计算单元 86

计算 3DQM 118, 3DQM 118 在图 1 的源装置 12 的背景中表示 3DQM 36。3DQM 计算单元 86 随后将 3DQM 36 输出到参数产生单元 88, 参数产生单元 88 输出在图 1 的源装置 12 的背景中包含参数 66、68 的参数 120。

[0080] 虽然关于全参考背景进行描述, 但 3D 分析单元 80 还可实施于不俘获或另外发起 3D 视频数据或内容而是仅存储 3D 视频数据的源装置内。此存储的 3D 视频数据可仅包含对应于左视图和右视图的深度图以及视图 26A、26B 中的单一者。在此所谓的“半参考”背景中, 不是调用全参考单元 92 来计算理想深度图 90, 理想深度估计单元 84 调用简化或半参考单元 94 来计算理想深度图 90, 理想深度图 90 再次在图 1 的源装置 12 的背景中表示理想深度图 62。然而, 在所有方面中, 3D 分析单元 80 大体上类似于上文关于全参考例子所描述的内容而操作。

[0081] 作为另一实例, 3D 分析单元 80 可在发起 3D 视频数据的显示装置 (例如, 显示装置 14) 内操作。假设 3D 分析单元 80 表示 3D 分析单元 48 (出于说明的目的), 则 3D 分析单元 80 接收包含提供的左视图 50A 和所重构的右视图 50B 的基于 DIBR 的 3D 视频数据 50。3D 分析单元 90 还接收焦距 103 和基线 105。3D 分析单元 80 进一步接收呈深度信息 58 的形式的深度信息 102, 深度信息 102 可包括对应于提供的左视图 50A 的左深度图。

[0082] 基于此接收的信息, 3D 分析单元 80 的理想深度估计单元 82 调用无参考单元 96, 同时考虑 3D 分析单元 80 仅接收单一提供的视图 50 以及呈深度信息 58 形式的对应左深度图。一旦经调用, 无参考单元 96 随后根据上文列举的等式 (3) 计算理想深度图估计 90, 其在图 1 的显示装置 14 的背景中表示理想深度图 70。全参考单元 92 将理想深度图 62 输出到失真度量计算单元 84。

[0083] 在接收到理想深度图 62 后, 失真度量计算单元 84 即刻调用 S0 失真度量单元 106、T0 失真度量单元 108 和 TI 失真度量单元 110 中的每一个, 以根据对应的等式 (4) 到 (6) 来计算失真度量 112 到 116 中的相应失真度量。在图 1 的显示装置 14 的背景中, 失真度量 112 到 116 表示度量 72。单元 106 到 110 将失真度量 72 输出到 3DQM 计算单元 86。3DQM 计算单元 86 计算 3DQM 118, 3DQM 118 在图 1 的显示装置 14 的背景中表示 3DQM60。3DQM 计算单元 86 随后将 3DQM 60 输出到参数产生单元 88, 参数产生单元 88 输出在图 1 的显示装置 14 的背景中包含参数 74、76 的参数 120。

[0084] 虽然关于全参考背景进行描述, 但 3D 分析单元 80 还可实施于接收基于 DIBR 的 3D 视频数据的显示装置内, 所述基于 DIBR 的 3D 视频数据包含呈深度信息 102 的形式的右深度图和左深度图两者。在此所谓的“半参考”背景中, 不是调用全参考单元 96 来计算理想深度图 90, 理想深度估计单元 84 调用简化或半参考单元 94 来计算理想深度图 90, 理想深度图 90 再次在图 1 的显示装置 12 的背景中表示理想深度图 70。然而, 在所有其它方面中, 3D 分析单元 80 可大体上类似于上文关于全参考和无参考例子所描述的内容而操作。

[0085] 图 3 是说明在实施本发明中所描述的三维 (3D) 质量度量导出技术的各种方面中的源装置 (例如, 图 1 的实例中所展示的源装置 12) 的示范性操作的流程图。虽然关于图 1 的源 12 进行描述, 但所述技术可由能够对 3D 视频数据进行编码以产生经编码的基于 DIBR 的 3D 视频数据的任何装置来实施。

[0086] 起初, 源装置 12 的控制单元 16 调用深度估计单元 20 以通过上文所描述的方式来计算深度信息 28 (130)。3D 视频编码器 22 调用视图译码单元 30 以对视图 26A、26B 中的一

者进行编码 (132), 其中出于说明的目的, 假设视图译码单元 30 对左视图 26A 进行编码。3D 视频编码器 22 还调用深度译码单元 32 以对深度信息 28 进行编码 (134)。3D 视频编码器 22 输出经编码的基于 DIBR 的 3D 视频数据 34。

[0087] 3D 视频编码器 22 还调用视图译码单元 30 以对经编码的左视图进行解码, 从而输出经解码的左视图 50A'。此解码是视频解码的例程方面, 因为视频编码对经编码的视图进行解码以用于确定残余数据, 所述残余数据是经解码的左视图 50A' 与后续左视图之间的差。类似地, 深度译码单元 32 对经编码的深度信息进行解码以产生经解码的深度信息 58'。深度译码单元 32 将经解码的深度信息 58' 输出到深度估计单元 20。

[0088] 基于经解码的左视图 50A' 和经解码的深度信息 58', 深度估计单元 20 重构视图 26A、26B 中的另一者, 所述另一者在图 1 的实例中被展示为所重构的右视图 50B' (136)。3D 分析单元 24 基于至少一个所重构的右视图 50B' 来计算理想深度图 62, 如上文所描述 (138)。3D 分析单元 24 随后以上文所描述的方式基于理想深度图 62 来导出失真度量 64 (140)。还如上文所描述, 3D 分析单元 24 基于失真度量 64 来计算 3DQM 36 (142)。另外, 3D 分析单元 24 基于 3DQM 36 来产生参数 66、68 (144)。3D 分析单元 24 产生参数 66、68 以便校正在观看 3D 视频数据 26 时所预期的所识别的视觉不适中的至少一些视觉不适。通过使用这些参数 66、68, 3D 分析单元 24 配置深度估计单元 20 和深度译码单元 32 (146)。

[0089] 图 4 是说明在实施本发明中所描述的技术的各种方面中的显示装置 (例如, 图 1 的实例中所展示的显示装置 14) 的示范性操作的流程图。虽然关于图 1 的显示装置 14 进行描述, 但所述技术可由能够对经编码的基于 DIBR 的 3D 视频数据进行解码以产生基于 DIBR 的 3D 视频数据的任何装置来实施。

[0090] 起初, 显示装置 14 的控制单元 38 调用 3D 视频解码器 44 以对经编码的基于 DIBR 的 3D 视频数据 34 进行解码。经编码的基于 DIBR 的 3D 视频数据 34 包含经编码的提供的视图和经编码的深度信息。3D 视频解码器 44 调用视图解码单元 54 以对提供的视图进行解码, 从而产生提供的视图 50A (150)。3D 视频解码器 44 还调用深度解码单元 56 以对经编码的深度信息进行解码, 从而产生经解码的深度信息 58 (152)。控制单元 38 随后调用视图重构单元 46, 视图重构单元 46 从经解码的深度信息 58 和经解码的提供的视图 50A 重构所重构的视图 50B (154)。

[0091] 控制单元 38 在重构视图 50B 之后进一步调用 3D 分析单元 48。3D 分析单元 48 以上文所描述的方式基于至少一个所重构的视图 50B 来计算理想深度图 70 (156)。3D 分析单元 48 随后如上文所描述基于理想深度图 70 来导出失真度量 72 (158)。还如上文所描述, 3D 分析单元 48 基于失真度量 72 来计算 3DQM 60 (160)。3D 分析单元 48 进一步基于 3DQM 60 来产生参数 74、76 (162)。3D 分析单元 24 产生参数 74、76 以便校正在观看 3D 视频数据 50 时所预期的所识别的视觉不适中的至少一些视觉不适。通过使用这些所产生的参数 74、76, 3D 分析模块 48 配置 3D 视频解码器 44 和视图重构单元 46 (164)。

[0092] 图 5 是在实施本发明中所描述的技术的各种方面以计算 3D 质量度量的 3D 分析单元 (例如, 在图 2 的实例中所展示的 3D 分析单元 80) 的示范性操作的流程图。起初, 3D 分析单元 80 接收呈所重构的视图 100 和深度信息 102 的形式的信息, 以及原始视图 98 和提供的视图 104 中的一个或一个以上 (170)。3D 分析单元 80 调用理想深度估计单元 82 以响应于接收到此信息。理想深度估计单元 82 随后基于所接收的信息来确定背景 (172)。

[0093] 举例来说,假设理想深度估计单元 82 接收原始视图 98、所重构的视图 100 和包含用于两个所述视图的深度图的深度信息 102,则理想深度估计单元 82 确定 3D 分析单元 80 在其中操作的背景是全参考背景。所接收的信息可因此被表征为全参考信息。在确定此全参考背景后 (“是 174”),理想深度估计单元 82 即刻调用全参考单元 92。全参考单元 92 实施上文所列举的等式 (1) 以基于所确定的全参考信息来计算理想深度图 90(176)。

[0094] 作为另一实例,假设理想深度估计单元 82 仅接收所重构的视图 100、提供的视图 104 和用于左视图和右视图两者的深度图,则理想深度估计单元 82 确定 3D 分析单元 80 在其中操作的背景是简化参考背景。所接收的信息可因此被表征为简化参考信息。理想深度估计单元 82 随后确定 3D 分析单元 80 不在全参考背景中操作 (“否”174),而是在简化参考背景中操作 (“是”178)。响应于此确定,理想深度估计单元 82 调用简化参考单元 94。简化参考单元 94 实施上文所列举的等式 (2) 以基于简化参考信息来计算理想深度图估计 90(180)。

[0095] 作为又一实例,假设理想深度估计单元 82 仅接收所重构的视图 100、提供的视图 104 和用于左视图和右视图中的一者的深度图,则理想深度估计单元 82 确定 3D 分析单元 80 在其中操作的背景是部分参考背景。所接收的信息可因此被表征为无参考信息。理想深度估计单元 82 随后确定 3D 分析单元 80 不在全参考背景中操作 (“否”174) 或在简化参考背景中操作 (“否”178)。结果,理想深度估计单元 82 调用无参考单元 96。无参考单元 96 实施上文所列举的等式 (3) 以基于无参考信息来计算理想深度图估计 90(182)。

[0096] 无论如何,理想深度估计单元 82 将理想深度图估计 90 输出到失真度量计算单元 84。响应于接收到此理想深度估计单元 82,失真度量计算单元 84 调用度量单元 106 到 110。度量单元 106 到 110 随后根据上文所列举的对应等式 (4) 到 (6) 基于理想深度图估计 90 来计算度量 112 到 116(184)。深度度量计算单元 84 将这些度量 112 到 116 输出到 3DQM 计算单元 86。响应于接收到这些度量 112 到 116,3DQM 计算单元 86 根据上文所列举的等式 (7) 基于失真度量 112 到 116 来计算 3DQM 118(186)。3D 计算单元 86 将 3DQM 118 输出到参数产生单元 88。响应于接收到 3DQM 118,参数产生单元 88 基于 3DQM 118 产生参数 120(188)。

[0097] 图 6 是说明提供根据本发明中所描述的技术而产生的 3DQM 的主观分析的图表 190 的图。图表 190 包含 y 轴 192,y 轴 192 界定 3DQM(例如,图 1 的 3DQM 36、60 和图 2 的 3DQM 118) 的可能值。图表 190 还包含 x 轴 194,x 轴 194 界定平均意见得分,其中得分 1 表示低劣的评估,得分 2 表示较差的评估,得分 3 表示中等的评估,得分 4 表示良好的评估,且得分 5 表示优秀的评估。

[0098] 图表 190 还包含展示为图表 190 中的实线的第一条线 196。这条线 196 识别曾由观看者(针对其产生平均意见得分)评估的相同 3D 视频数据的 3DQM 的理想值。值得注意的是,线 196 指示 3DQM 得分 0 对应于平均意见得分 0,且以此方式线性进行,使得 3DQM 得分 1 对应于平均意见得分 1,且依此类推。图表 190 进一步提供一组虚线 198A、198B,其识别远离理想线 196 的一个  $\delta$  标准偏差(正和负)。此外,图表 190 包含第二组线 200A、200B,其界定与理想线 196 的异常点边界。

[0099] 图表 190 还具有以下特征:在全参考背景和部分参考背景两者中计算出的 3DQM 的许多数据点。全参考 3DQM 和无参考 3DQM 一般驻留于理想线 196 的一个标准偏差内(如由线 198A、198B 识别)。具体来说,结果的均方根(RMS)误差等于约 0.6158,且平均意见得分

的标准偏差为约 0.7885。考虑到均方根误差小于平均意见得分的标准偏差,可一般得出结论:3DQM 非常准确。全参考和无参考 3DQM 的相关系数等于约 0.8942,从而指示结果是相干的,而异常点比率为 0,从而指示所有 3DQM 值是一致的。在此方面,3DQM 可近似主观观看者结果,且因此可促进对 3D 视频数据的自动校正以改进或至少促进 3D 视频数据的观看体验。

[0100] 3DTV 的客观质量度量(例如,上文所描述的 3DQM)对于 DIBR 算法、深度图压缩、深度图估计、填孔技术和显示增强的质量的提升可非常重要。理想深度图估计(特定为 3DQM)也可比较重要,因为其可作用于以上文所描述的方式在深度图估计中进行低劣像素校正的精细化步骤。具体来说,对深度估计不准确性如何影响特定图像的视觉质量进行量化可允许在估计准确的深度图中所涉及的计算的复杂性与针对这些图像引入的假影之间的折衷,其类似于压缩中的速率-失真折衷。此折衷方面在其中处理或电池资源是有限的嵌入式实时应用中可尤其适用。3DQM 技术的实施方案还可允许对归因于深度图假影而引起的感知性失真进行实际计算,其与基于错误的度量相比允许用于将若干位分配在深度图压缩中的更有效率的速率-失真准则。

[0101] 再次,虽然上文关于两个视图系统进行描述,但可在其中将两个以上视图提供给 3D 显示器从而允许显示多个不同视图的多视图 3D 系统中实施所述技术。在此方面中,所述技术可用于针对这些视图中的每一者或针对作为一群组的视图提供 3DQM,其中这些一个或一个以上 3DQM 可促进对预期在观看这些视图时发生的至少一些视觉不适的识别和潜在后续的校正。因此,虽然在本发明中出于容易说明的目的关于二视图系统进行描述,但所述技术可用于多视图系统中。

[0102] 图 7 是展示移位传感器相机模型的图 220。移位传感器相机模型可对 3D 俘获装置有效地建模,且出于此原因,可表示虚拟 3D 俘获装置。在图 220 中,点 222A、222B 分别表示左眼 2D 俘获装置和右眼 2D 俘获装置的中心点。如上文所述,这些点 222A、222B 之间的水平距离被称作基线,其由字母 'B' 表示。所述中心中的每一者到焦点的距离(其再次被称作焦距)在图 220 中被展示为字母 'F'。图 220 中的字母 'h' 是指相对于传感器中心的投影中心的水平方向上的移位。这些各种字母 'B'、'F' 和 'h' 是指移位传感器相机模型的变量,所述变量可响应于不同的 3DQM 值而更改以校正至少一些所预期的视觉不适。这些变量还对应于上文在所述等式(例如,等式(2)和(3))中的各个等式中所使用的变量。

[0103] 可以硬件、固件或其任何组合来实施本文中所描述的技术。在一些例子中,所述硬件还可执行软件。被描述为模块、单元或组件的任何特征可共同地实施于集成逻辑装置中或分开实施为离散但可互操作的逻辑装置。在一些状况下,各种特征可实施为集成电路装置,例如,集成电路芯片或芯片组。如果是软件实施,则可至少部分地通过包含在经执行时致使处理器执行上述方法中的一个或一个以上的指令的计算机可读媒体来实现所述技术。

[0104] 计算机可读媒体可形成计算机程序产品的部分,计算机程序产品可包含封装材料。计算机可读媒体可包括计算机数据存储媒体,例如,随机存取存储器(RAM)、同步动态随机存取存储器(SDRAM)、只读存储器(ROM)、非易失性随机存取存储器(NVRAM)、电可擦除可编程只读存储器(EEPROM)、快闪存储器、磁性或光学数据存储媒体等。另外或替代地,所述技术可至少部分由载运或传送呈指令或数据结构的形式的代码且可由计算机存取、读取和



/或执行的计算机可读通信媒体来实现。

[0105] 代码或指令可由一个或一个以上处理器执行,例如,一个或一个以上 DSP、通用微处理器、ASIC、现场可编程逻辑阵列 (FPGA) 或其它等效集成或离散逻辑电路。因此,如本文中所使用的术语“处理器”可指上述结构或适合于实施本文中所描述的技术的任一其它结构中的任一者。另外,在一些方面中,本文中所描述的功能性可提供于专用软件模块或硬件模块内。本发明还预期包含用以实施本发明中所描述的技术中的一个或一个以上的电路的多种集成电路装置中的任一者。此电路可提供于单个集成电路芯片中或提供于所谓的芯片组中的多个可互操作的集成电路芯片中。这些集成电路装置可用于多种应用中,所述应用中的一些可包含在无线通信装置(例如,移动电话手持机)中的使用。

[0106] 已描述了本发明的各种实例。这些和其它实例处于所附权利要求书的范围内。

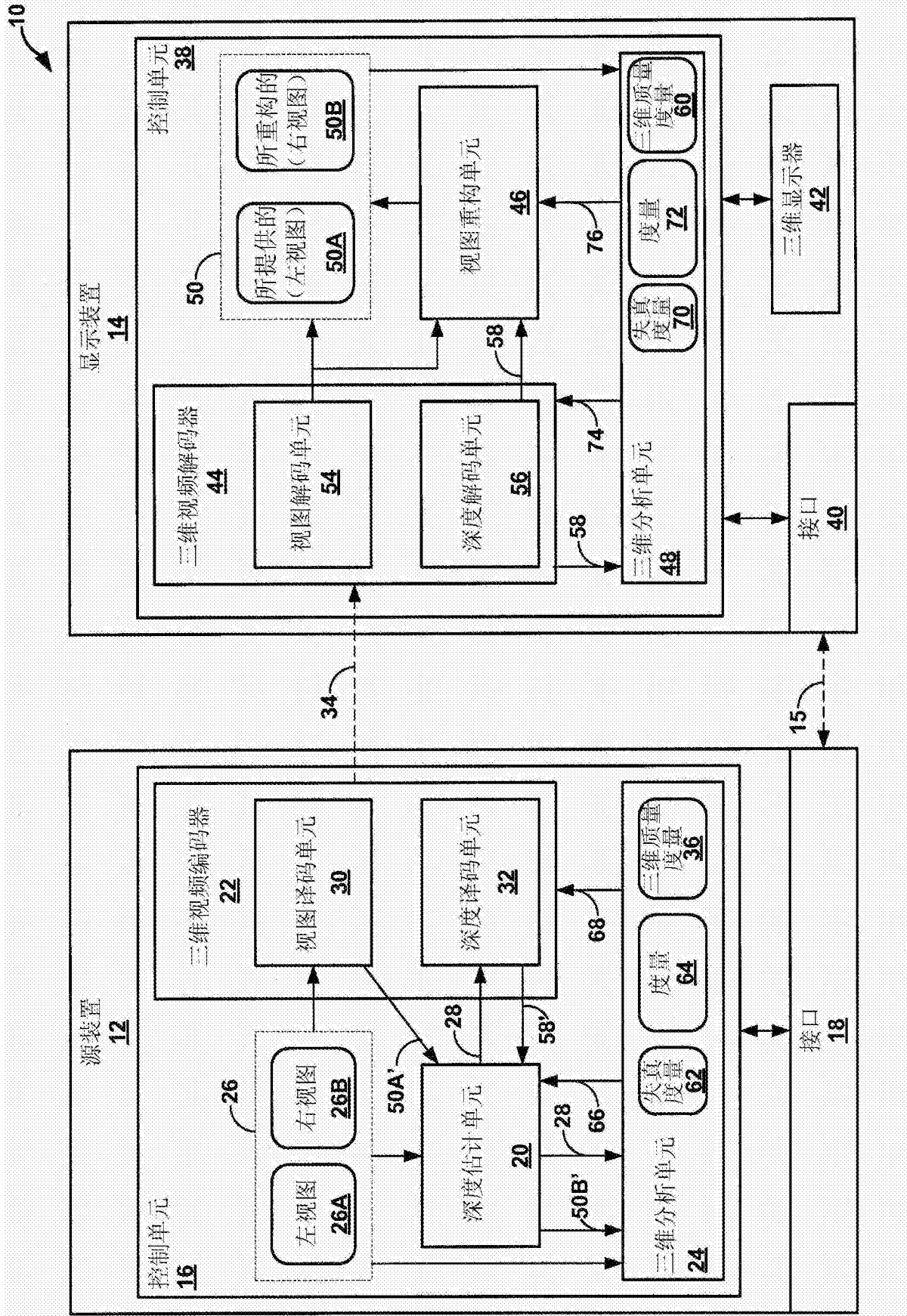


图 1

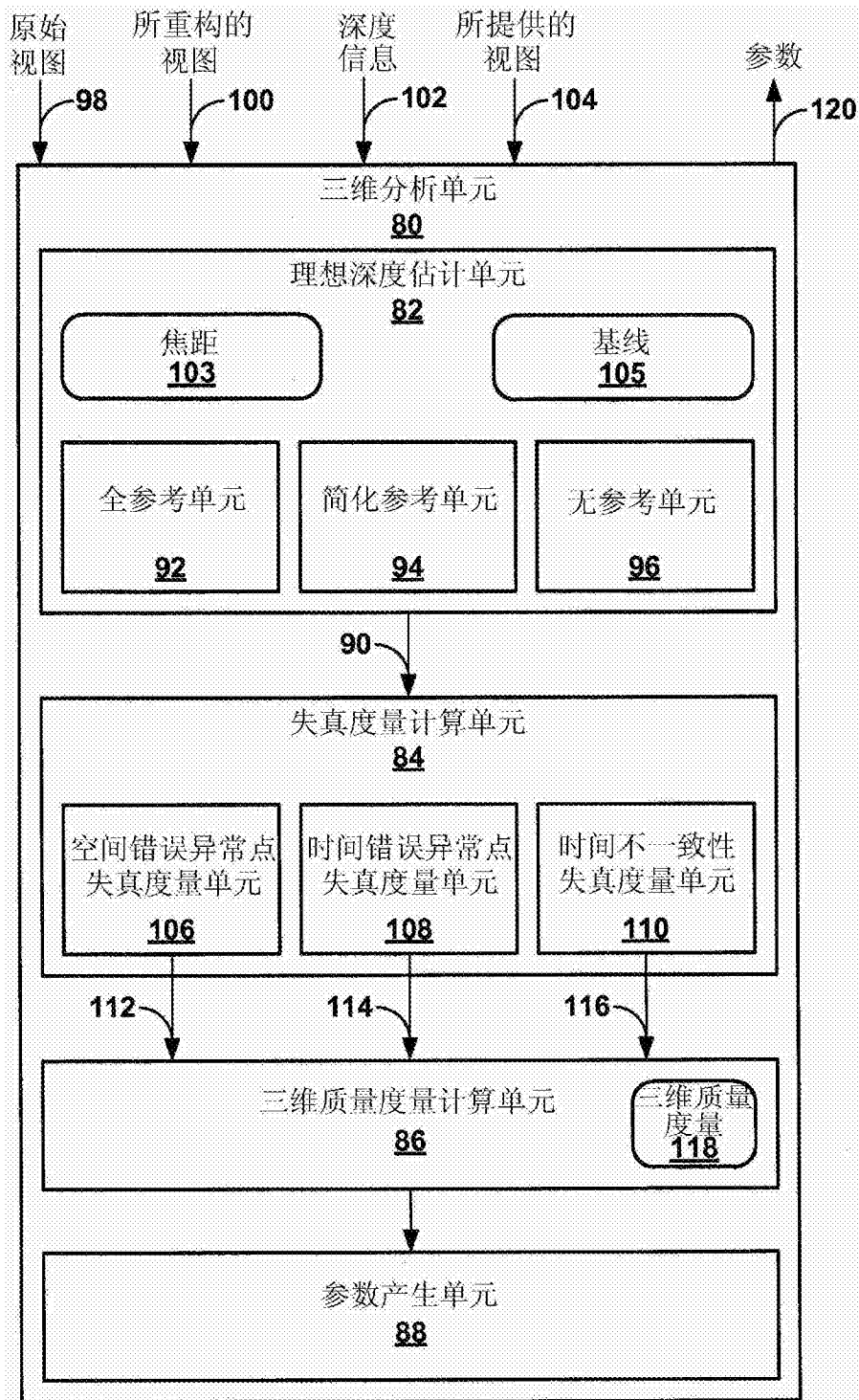


图 2

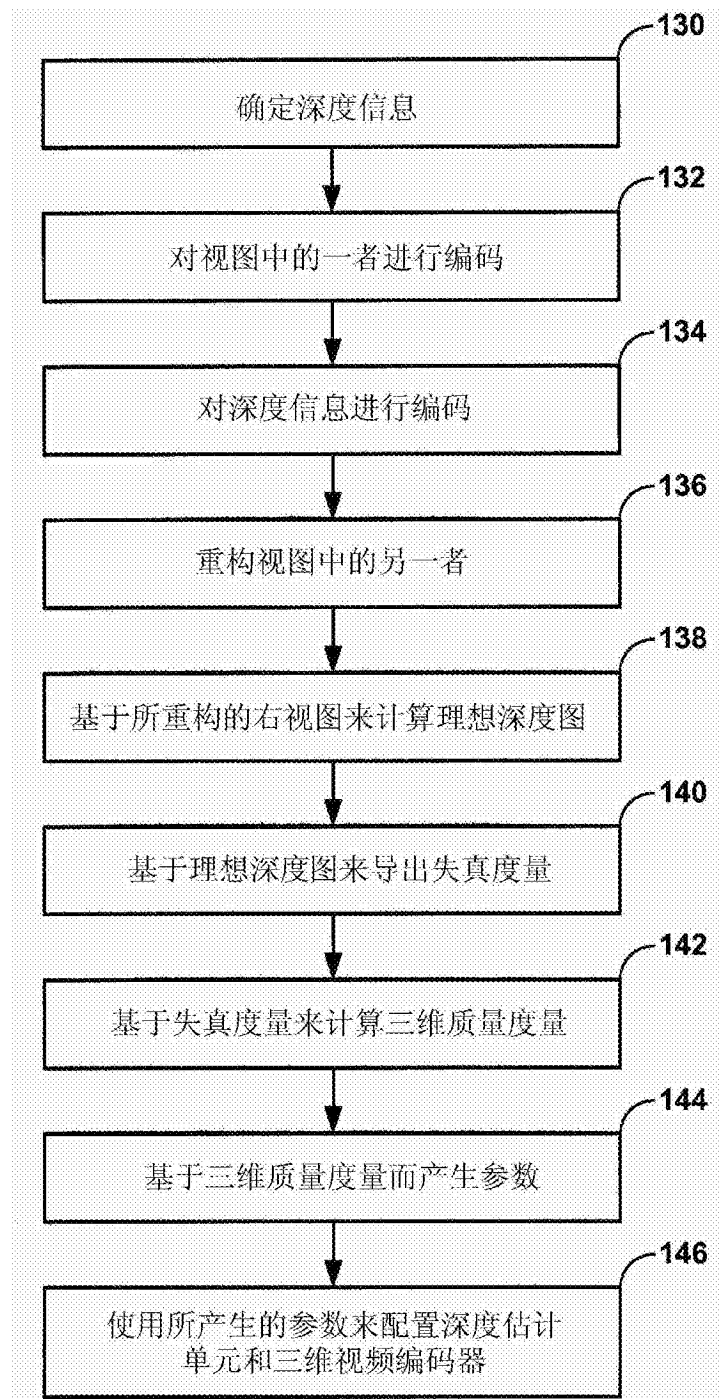


图 3

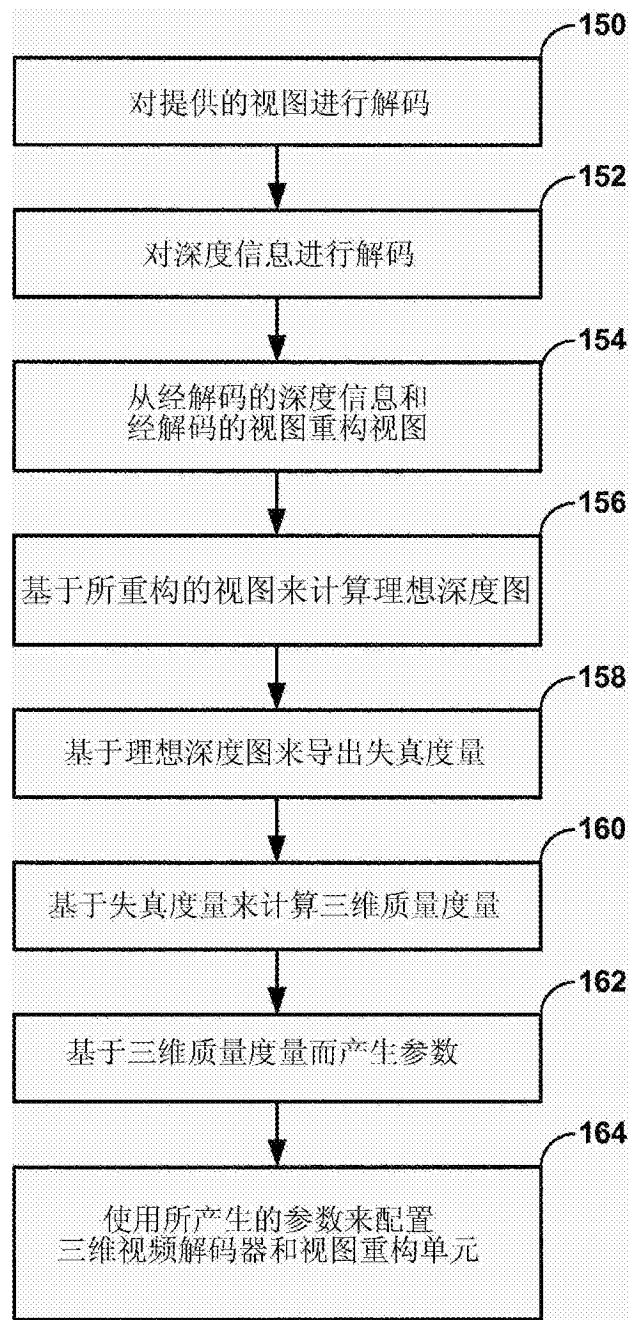


图 4

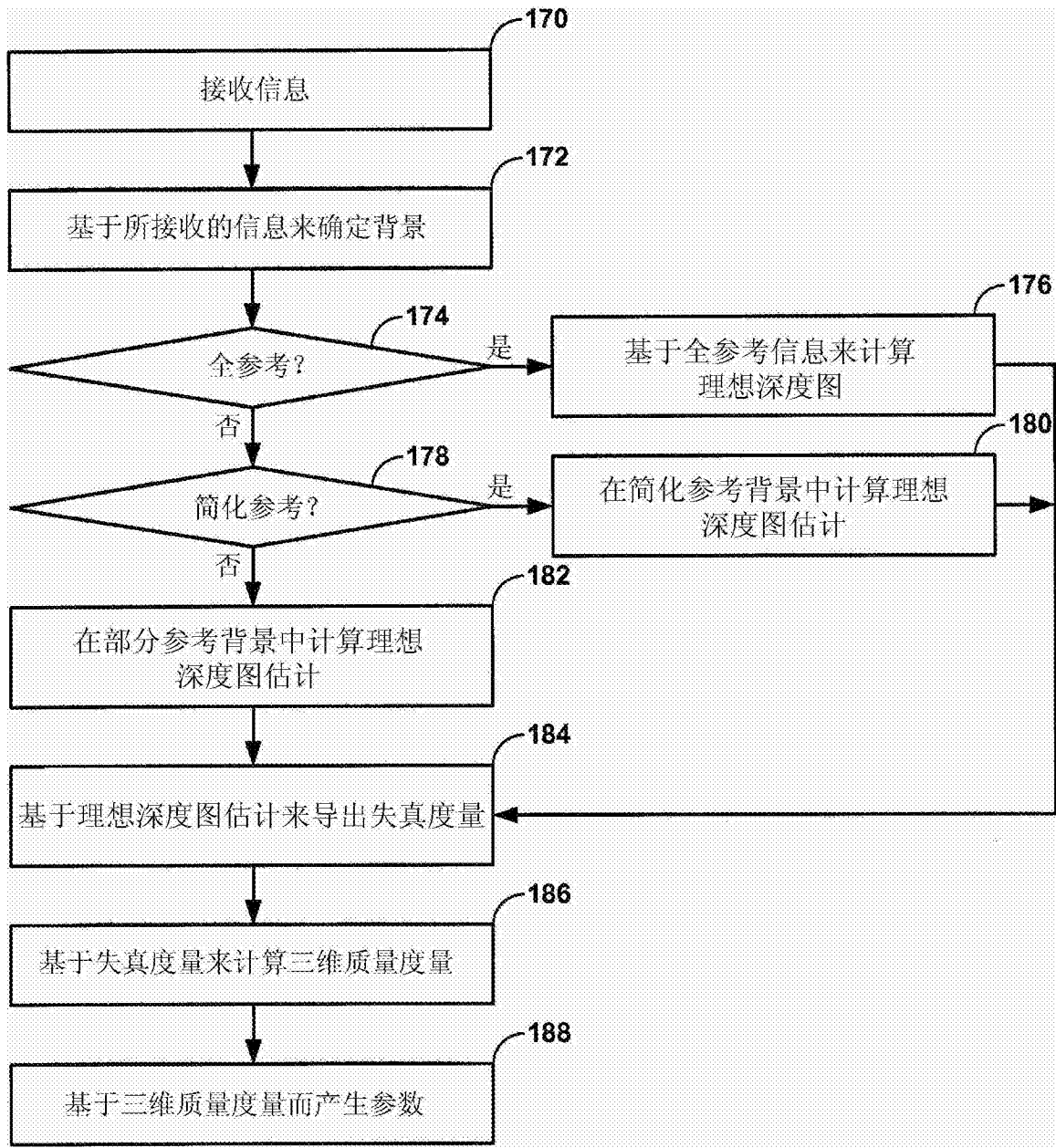


图 5

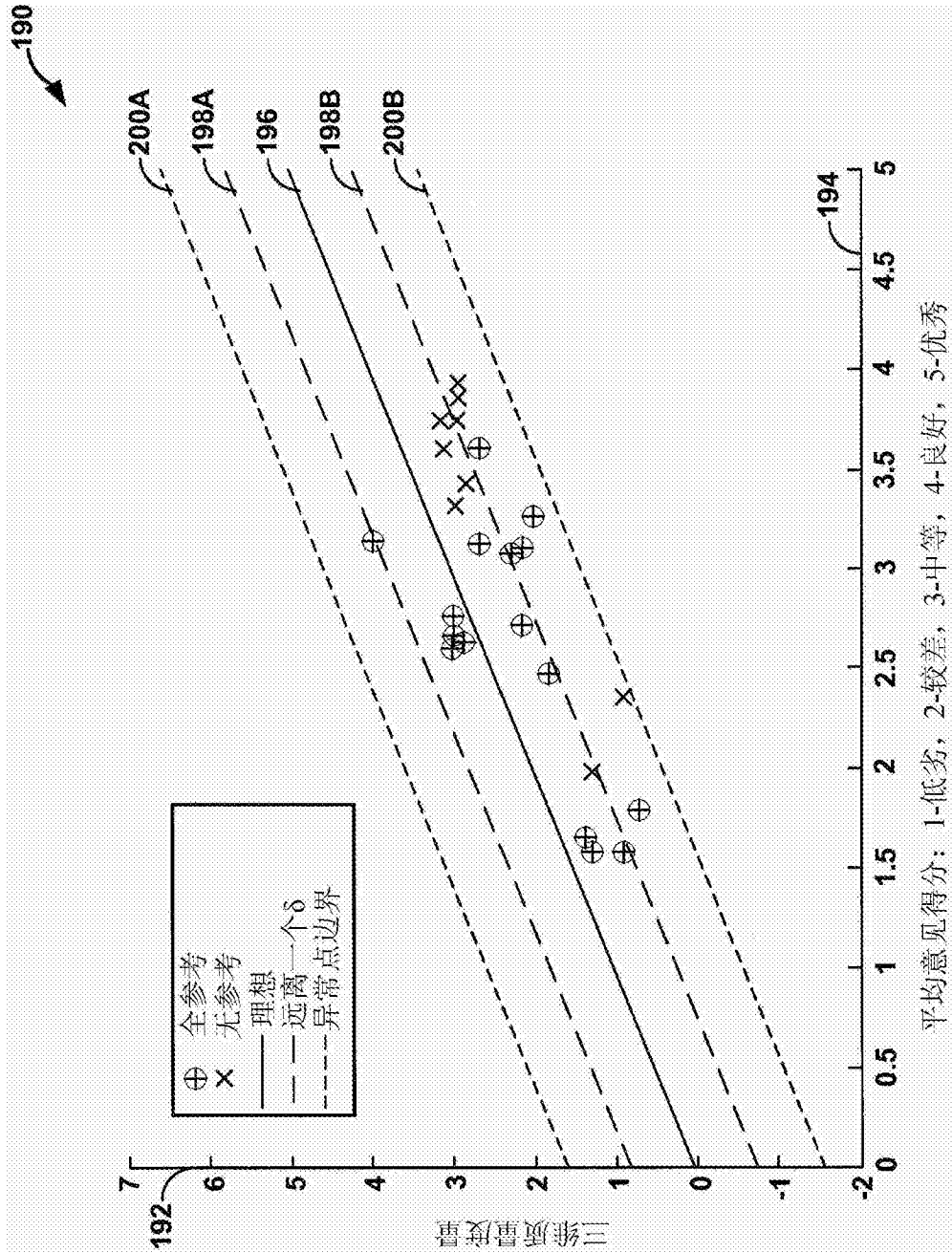


图 6

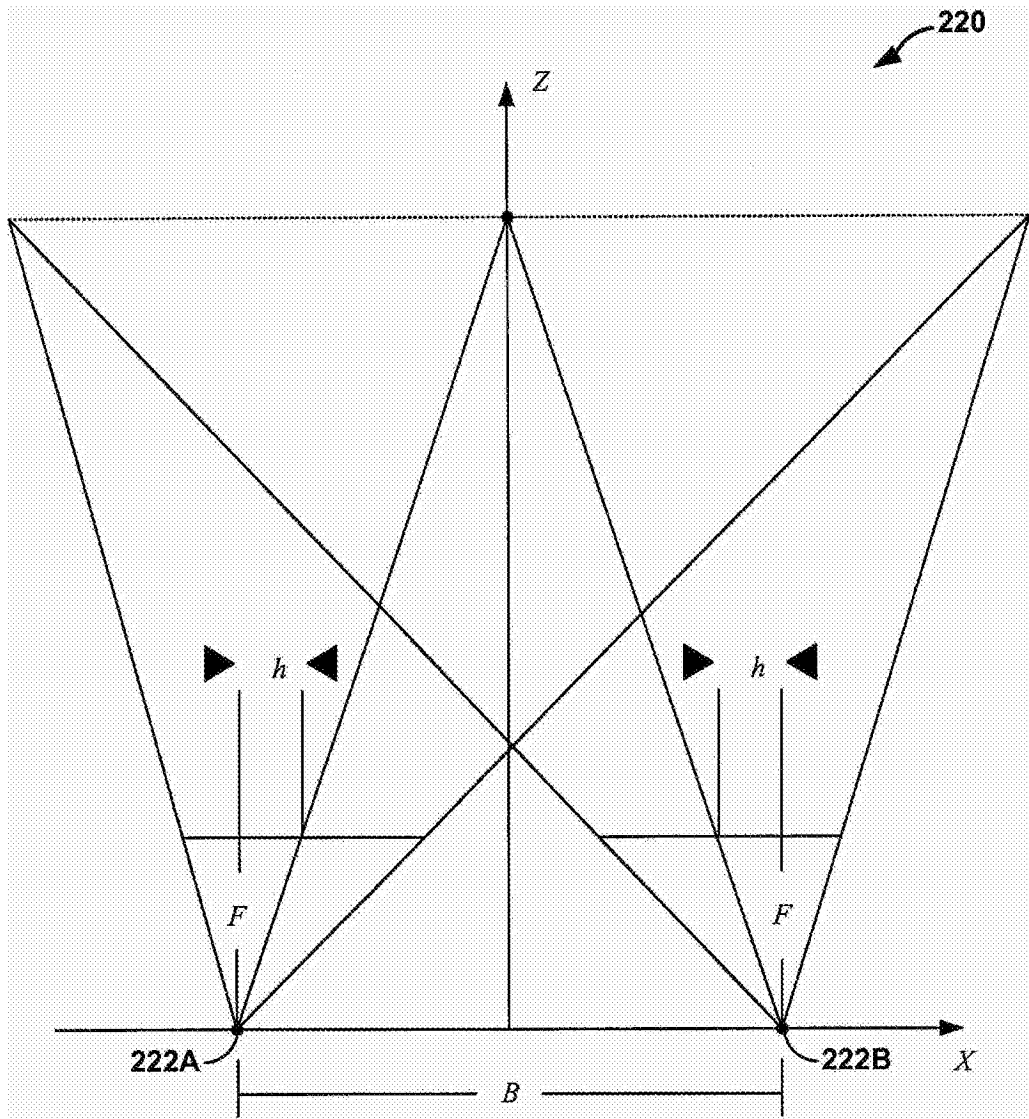


图 7